Development of Auto-Immolative Spacers for Probes of Enzyme Activity

Oliver Thörn Seshold

Numéro national de thèse : 2013ENSL0817

THÈSE

VERSION DESTINEE A LA COMMISSION D'EXAMEN

En vue de l'obtention du grade de

Docteur de l'Université de Lyon, délivré par l'Ecole Normale Supérieure de Lyon

Discipline : Chimie

Laboratoire de Chimie, UMR 5182

Ecole Doctorale de Chimie de Lyon, ED 206

Presentée et soutenue publiquement le 27/06/2013

par Oliver Torvald THORN-SESHOLD

Development of Auto-Immolative Spacers for Probes of Enzyme Activity

---

Directeur de Thèse : Professeur Jens HASSERODT

Après l'avis de : Professeur Pierre-Yves RENARD, Rapporteur
Professeur Jean-Louis REYMOND, Rapporteur

Devant la Commission d'examen formée de :

Professeur Jens HASSERODT, Ecole Normale Supérieure de Lyon, Membre
Docteur Sébastien PAPOT, Université de Poitiers, Membre
Professeur Olivier PIVA, Université Lyon 1, Président
Professeur Pierre-Yves RENARD, Université de Rouen, Rapporteur
Professeur Jean-Louis REYMOND, Universität Bern, Rapporteur

Personal Acknowledgements

Firstly, and once again, my profound thanks to Jens Hasserodt for the opportunity to work with him on these and other projects, for his support and advice and all the other far more valuable things. These happy years have made the "I Love Science" tag a reality; I look forward to many more years of fruitful exchanges. And I'm truly sorry I never could fix that vacuum pump.

Thanks:
to all my coworkers in group CBO for their support and engagement: in particular, thanks to Jacek Kolanowski for żurek, support, and inspiration; to my office mates Guillaume Gros and Jinping Wang for the International Bureau; to Maxime Prost for his collaboration; to Fayçal Touti, whose advice comes at the right time; to Andreas Brunner, Claire Weisslinger and Nils Aronssohn for all they taught me; and to polymath CBO guest stars Anna Servat Gimenez, Eva Deront and Paula Adler. Your energy and reinforcement have helped me see this through, and I look forward to following your bright futures;

to those whose collaborative help has broadened my horizons: in particular, thanks to Malgorzata Borowiak; also Pierre Jalinot, Florian Albrieux, Véronique Josserand and Jean-Luc Coll;

to my coworkers at the Laboratoire de Chimie outside CBO: to Jean-Christophe for his involvement with phosphoniums, to Martine for IR and Sandrine for udeft NMR, to Philippe for his struggles to keep the LCMS alive, and to Jean-Pierre, Menaf, Yann, Delphine, Pascal and Julien for each their own *pierre apportée*;

to my referees and jury members for their time and their energy to pass on the sparks of science;

to my constant companions Bouvier, Janáček, Shostakovich, Galland, and the ONL;

to my friends for injecting life and energy into these years of organic metastability; to my wonderful flatmates Eva, Johanna and Cédric; to the fluffy ferret redactrix, Tif, sushi fish, cromo, green frog, red lory, kraseviye cat and butterfly-elflet;

to those *sine qua non*: George Bacskay; Bob Gilbert; Tom Grujic; Andrew Haines; Leif Holmlid; Elizabeth New; Anthony Phillips; Geoff Salem; Martijn de Sterke; Gerald Teng; Gah-Yi Vahn; Trent Wallis;

and to my parents, who grounded me *In Vovo* many years before *in vivo*.

Lastly, I dedicate this thesis to my inspiration and teacher, Michael Bishop.

Yeah?

Keywords

auto-immolative spacers · molecular probes · exopeptidases · glycosidases · imaging agents

Mots-Clés

espaceurs auto-effondrables · sondes moléculaires · exopeptidases · glycosidases · agents d'imagerie

Détails Pratiques

Ces travaux de thèse (10/2009-08/2012) ont été fait sous la direction de Prof. Jens Hasserodt, dans l'équipe de Chimie Bio-Organique, au Laboratoire de Chimie de l'Ecole Normale Supérieure de Lyon, 46 Allée d'Italie, Lyon 69007, France. Les recherches pratiques de Master 2 (01/2009-07/2009) auxquelles il est fait référence en partie II.2.1 ont été fait au même lieu. Tout travail pratique contenu dans ce manuscrit est du travail personnel et de thèse sauf mention explicite au contraire.

Remerciements Pratiques

Ce travail de thèse a été financé par une bourse doctorale du Ministère de l'Education Supérieure et de la Recherche.

Table of Contents

Textual / molecular abbreviations

38C2	aldolase catalytic antibody 38C2 developed by R. Lerner *et al*.
3H-DCDHF	(*E*)-2-(3-cyano-4-(3-hydroxy-4-methoxystyryl)-5,5-dimethylfuran-2(5*H*)-ylidene)malononitrile
7AMC	7-amino-4-methylcoumarin
AIBN	2,2'-azobis(*iso*-butyronitrile)
Anis	*para*-anisaldehyde TLC revealing dip
β-ALA	β-alanyl aminopeptidase
β-Ala	β-alanine (3-aminopropanoic acid)
β-GAL	β-galactosidase enzyme
βGal / Bn_4βGal / Ac_4βGal	β-D-galactopyranosyl / tetra-*O*-benzyl-βGal / tetra-*O*-acetyl-βGal. An external bond (*e.g.* βGalX for an external bond to X) is to the anomeric carbon
β-GLUC	β-glucuronidase enzyme
β-LACT	β-lactamase enzyme
BocLeuOH	*N*-*tert*-butoxycarbonyl-L-leucine
BTQC	ethyl 7-(benzo[d]thiazol-2-yl)-6-hydroxy-4-oxo-1,4-dihydroquinoline-3-carboxylate
CBO	Chimie Bio-Organique (research group headed by Prof. Hasserodt)
CD	circular dichroism
CDI	carbonyl diimidazole
Cl_2HPQ, diClHPQ	6-chloro-2-(2-hydroxyphenyl)quinazolin-4(3*H*)-one. An external bond is to the phenol (*e.g.* Cl_2HPQOCOCl is the chloroformate)
CPP	Cell-penetrating peptide
CPT	camptothecin
CT	X-ray Computed Tomography
Cy	cyclohexane
DCC	N,N'-dicyclohexylcarbodiimide
DCM	dichloromethane
DCU	N,N'-dicyclohexylurea
DDQ	2,3-dichloro-5,6-dicyano-1,4-benzoquinone
DMAc	dimethylacetamide
DMEM	Dulbecco's modified Eagle's medium
DMF	dimethylformamide
DMSO	dimethylsulfoxide
DOX	doxirubicin
dptacn	1,4-dipicolyl-1,4,9-triazacyclononane; an external bond (*e.g.* dptacnX for an external bond to X) is to the secondary amine, N-9
dttacn	1,4-bis(5-tetrazolyl)methyl-1,4,9-triazacyclononane
EA	ethyl acetate
EDC	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide
FDA	fluorescein diacetate
G	a general substituent, eg. for electron density modulation on an aromatic ring
Haness	Hanessian's TLC revealing dip
HIV	Human Immunodeficiency Virus
HOBt	N-hydroxybenzotriazole
HPLC	High performance liquid chromatography
HPQ	2-(2-hydroxyphenyl)quinazolin-4(3*H*)-one. An external bond is to the phenol (*e.g.* HPQOCOCl is the chloroformate)
IR	Infra-red spectroscopy
LAP	Porcine Liver Leucine Aminopeptidase, EC 3.4.11.1 (also known by 80 other names, including Aminopeptidase M, and Membrane Alanyl Peptidase)

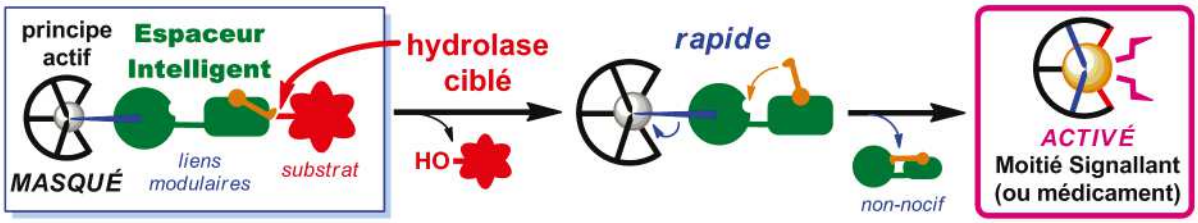
LG	leaving group
MRI	Magnetic Resonance Imaging
Ms	methanesulfonyl
MS [3Å/4Å]	molecular sieves; 3Å for MeOH, 4Å if otherwise unspecified
MS; DIMS(+/-); LCMS(+/-)	Mass Spectrometry, (positive/negative mode)Direct-Injection Mass Spectrometry, coupled High Performance Liquid Chromatography/(positive/negative mode)Mass Spectrometry
NBS / NCS	*N*-bromosuccinimide / *N*-chlorosuccinimide
Nin	ninhydrin TLC revealing dip
NMI	N-methylimidazole
NMR	Nuclear Magnetic Resonance
NTR	nitroreductase (bacterial enzyme)
PABA	*para*-aminobenzyl alcohol
PBS	phosphate buffer solution (for biological tests)
PET	Positron Emission Tomography
PG	protecting group
PHBA	*para*-hydroxybenzyl alcohol
picolyl, 2-picX	(2-pyridyl)methyl: eg, 2-picLi is 2-pyridylmethyllithium
picoline	2-methylpyridine
PMA	phosphomolybdic acid TLC revealing dip
PSBA	*para*-mercaptobenzyl alcohol
py	pyridine
pyridyl, 2-pyX	2-pyridyl with a bond to X; eg, 2-pyCHO is 2-pyridine carboxaldehyde
RFU	Random Fluorescence Units (arbitrary units for fluorescence readout, but using identical acquisition protocols, should be intercomparable)
RT	room temperature (~22-26°C)
sat.	a saturated aqueous solution of
S_N2	bimolecular nucleophilic substitution [reaction]
S_NAr	nucleophilic aromatic substitution [reaction]
SPECT	Single Photon Emission Computed Tomography
spont.	spontaneous
tacn	1,4,9-triazacyclononane
TBAF, TBA^+	*tert*-butylammonium fluoride, *tert*-butylammonium ion
TFA	trifluoroacetic acid / trifluoroacetate
THF	tetrahydrofuran
THP; DHP	2-tetrahydropyranyl / 3,4-dihydro-2*H*-pyran
TLC	Thin-layer chromatography
TML	trimethyl lock spacer
TMS; TBDMS/ TBS; TBDPS	trimethylsilyl; *tert*-butyl-dimethylsilyl; *tert*-butyl-diphenylsilyl
TPA	Two Photon Absorption [fluorescence imaging]
tptacn	1,4,9-tripicolyl-1,4,9-triazacyclononane
Tris	2-hydroxymethyl-2-amino-1,3-propanediol
Ts	*para*-toluenesulfonyl; thus TsOH: *para*-toluenesulfonic acid
Van	acidified vanillin TLC revealing dip

Abbreviations of analytical variables and parameters

A ($\mu M \cdot min^{-1}$)	Enzyme activity (in the section of kinetic analysis and modelling, sometimes approximated by v_{max} where appropriate)
$A_{X,I}(t)$, A_{PLAT} (RFU)	Raw fluorescence signals from the fluorimeter, $A_{X,I}(t)$=signal for replicate number I of data set X, A_{PLAT}=raw signal at plateau
B_{ACQ} (RFU)	experimental background signal not including any signal from the proagent
B_{PR} ($RFU \cdot M^{-1}$)	unwanted 'background' signal from the proagent form
C: C_{LIN}, $C_{PPT}(t)$, $C_{PROC}(t)$ (μM)	Concentration parameters: $C_{PROC}(t)$=probe concentration processed by time t; C_{LIN}=$C_{PROC}(t_{LIN})$, $C_{PPT}(t)$=probe concentration first hydrolysed then precipitated by time t.
F: $F_X(t)$, F_{PLAT}, $F_{NoEnz}(t)=F_N(t)$ (RFU)	Fluorescence timecourse signals after treatment by the data processing algorithm: $F_X(t)$=signals for data set X, F_{PLAT}=plateau fluorescence signal; $F_{NoEnz}(t)=F_N(t)$=signal from the no-enzyme control reaction
J (Hz)	The internuclear coupling constant in Hertz
k: k_{CYC}, k, $k_{CYC,LB}$ (min^{-1})	k_{CYC}= rate constant of spacer cyclisation; k=a general dummy constant used as needed; $k_{CYC,LB}$= lower bound for cyclisation rate as determined in initial kinetic analysis
K: K_S, K_M (μM)	K_S=solubility constant; K_M=Michaelis-Menten constant
m/z (Th)	The absolute magnitude of the mass-to-charge ratio of an ion in Thomson
m_{max} ($\mu M \cdot min^{-1}$)	maximal signal increase gradient determined from the linear signal region
P: P_0, P_R, P_X, P(t) (μM)	Probe (proagent) concentrations: P_0=initial concentration, P_R=P(t)=remaining concentration at time t, P_X=concentration hydrolysed (enzymatically+spontaneously) by time t.
Q: Q(t), Q_{LB} (μM)	Q(t)=concentration of **33** (processed probe before cyclisation) at time t; Q_{LB}=lower bound for $Q(t_{LIN})$
R	inherant signal-to-background ratio of an experiment
R(t) (μM)	Concentration of fluorophore dissolved in solution at time t
S ($RFU.M^{-1}$)	signal strength of the emitter/agent
S(t) (μM)	Effective concentration of fluorophore precipitated from solution by time t
t: t_{corr}, t_{lag}, $t_{C=0}$, t_{LIN}, $t_{1/2}$, $\tau_{1/2}$, t_{UB}, t_{LB}, $t_{CYC,UB}$ (min)	Time parameters: t_{corr}=correction time for initial cyclisation kinetic analysis; t_{lag}=signal onset time; $t_{C=0}$=t-intercept of linear signal region; t_{LIN}=linear signal onset time; $t_{1/2}$= $\tau_{1/2}$=a general halflife value (spontaneous hydrolysis, cyclisation, signal generation...), t_{UB}=upper bound for t_{corr}; t_{LB}=lower bound for t_{corr}; $t_{CYC,UB}$=upper-bound cyclisation halftime
t_{ret} (min)	The retention time of a product during LCMS, in minutes
U, mU	enzyme activity expressed in standard Units: eg. 1 $U \cdot mL^{-1}$ → 1 $\mu mol \cdot mL^{-1}$ of a specific substrate processed per minute under standard and saturating conditions (=1 $mM \cdot min^{-1}$)
v_{max}, v* ($\mu M \cdot min^{-1}$)	v_{max}=limiting enzymatic activity when a specific amount of enzyme is saturated by substrate; v*=estimated v_{max} of 1 nM enzyme
δ (ppm)	The chemical shift of a nucleus measured in parts per million from the nuclear reference frequency of tetramethylsilane at that field strength
ΔStokes (nm)	Difference between probe and autofluorescence stokes shifts: $\Delta Stokes=\Delta\lambda_P-\Delta\lambda_A$
$\Delta\lambda_A$, $\Delta\lambda_P$ (nm)	Stokes shifts of tissue autofluorescence ($\Delta\lambda_A$) and of a probe ($\Delta\lambda_P$)
ν (cm^{-1})	The wavenumber of an excitation in IR spectroscopy in reciprocal centimetres

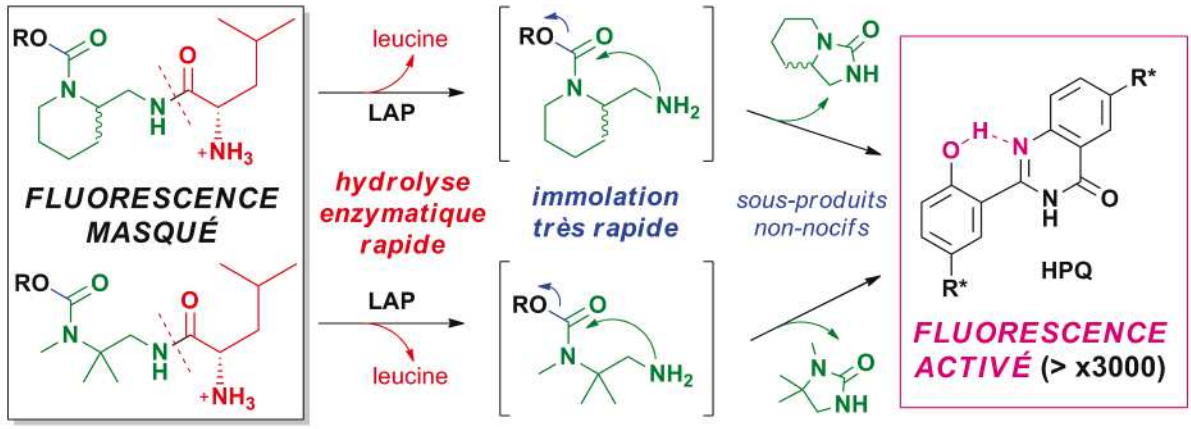
Résumé en français

L'imagerie moléculaire est une méthode pour l'étude des processus enzymatiques dans le vivant, qui peut promettre une meilleure compréhension de la biologie ainsi qu'un diagnostic plus fiable et précoce des maladies. Une approche particulièrement sensible et pratique s'appuie sur l'utilisation de sondes modulaires de type *éteint-ALLUMÉ*, ce qui se prête bien à une architecture à trois composants : moitié signalant—espaceur intelligent—substrat. En principe, selon le choix du substrat, on pourra étudier des enzymes différents arbitrairement ; et en fonction du moitié signalant, on pourra obtenir des sondes à caractéristiques ou modalités diverses adaptées à des fins spécifiques.



La connexion fonctionnelle entre ces substrats et moitiés signalants devrait être assuré par un espaceur intelligent approprié. Or, actuellement, l'espace chimique d'espaceurs connus manque de bonnes solutions pour de nombreuses connectivités nécessaires dans la conception de sondes pour des processus enzymatiques de grand intérêt. Cette thèse traite de la mise en œuvre d'espaceurs auto-effondrables novateurs pouvant résoudre des problèmes dans la conception de sondes pour l'activité de deux classes importantes d'hydrolases par deux modalités phares d'imagerie *in vivo*.

La première partie se focalise sur des espaceurs servant de lien entre un substrat de peptidase et un phénol, pour des sondes fluorogènes. Les espaceurs auto-effondrables connus susceptibles de servir dans cette connexion ne réunissaient pas toutes les qualités nécessaires pour de bonnes applications modulaires *in vivo* – eg. stabilité, cinétique enzymatique, cinétique d'effondrement, et non-toxicité de sous-produits. Deux espaceurs cyclisant novateurs de type 1,2-diamine ont été conçus pour faire face à ces défis. Des sondes fluorogènes basées sur ces espaceurs ont été conçues pour démontrer le potentiel de cette approche, et les sondes synthétisés et examinés par des tests *in vitro*.



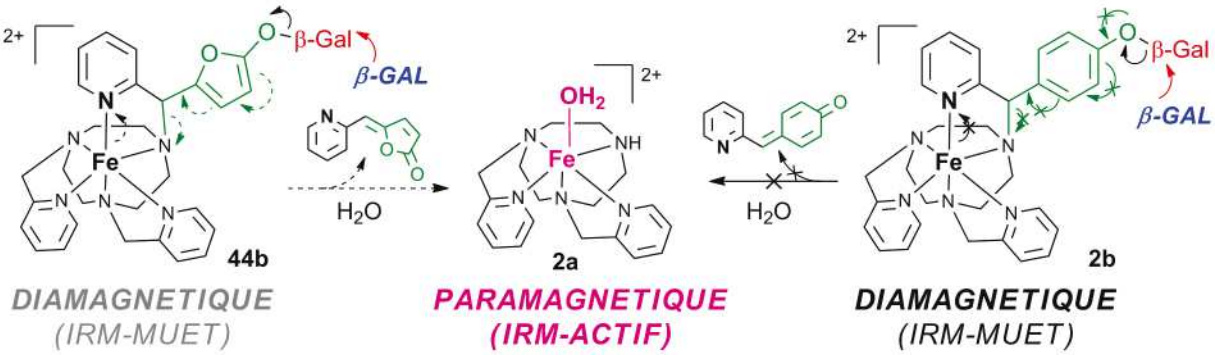
La conception générale bâtie sur ces espaceurs novateurs (sondes à 3 composants), a montré une performance très satisfaisante : les sondes étaient très robustes (demi-vie > 560 h), mais ont pu être enzymatiquement hydrolysées avec une bonne cinétique, et puis relâchaient rapidement (demi-temps ~ 3 min) leur fluorophore, qui était un **HPQ**. L'emploi de ces fluorophores dans des sondes d'activité enzymatique a jusqu'alors été empêché par l'instabilité de leur adduits, même s'ils sont connus pour présenter de nombreux avantages pratiques par rapport aux fluorophores phénoliques habituels tels les fluorescéines et resorufins. Clé pour ces études *in vitro*, leur fluorescence est de type ESIPT (Transfert de Proton Intramoléculaire dans l'Etat Excité), ces sondes opèrent alors selon un vrai mécanisme microscopique *éteint–ALLUMÉ*, pour maximiser leur sensibilité potentielle.

Grâce à la robustesse du présent système, ces sondes ont pu mettre à profit ces fluorophores exceptionnels, pour faire une première démonstration d'un système de sonde *macroscopiquement* binaire *éteint–ALLUMÉ* pour la libération de phénols sous activité d'aminopeptidases. Ces sondes avaient une excellente sensibilité par rapport aux méthodes actuelles, avec un rapport signal:contrôle de 3000:1 (ou 43000:1 avec soustraction du signal à t_0). La modélisation de plusieurs aspects du système a permis de conclure quant au potentiel de ces sondes précipitant ESIPT par rapport aux sondes classiques, et d'analyser de manière explicite les quatre cinétiques microscopiques intervenant dans la génération de signal ainsi que les paramètres du régime permanent.

Ce système pourrait permettre de faire de l'imagerie moléculaire ultra-sensible d'une gamme d'exopeptidases avec des fluorophores ESIPT, en changeant sensiblement les substrats utilisés ; ces espaceurs pourraient également servir dans des sondes comportant d'autres fluorophores phénoliques. Le potentiel de cette approche pour des thérapies ciblées est aussi actuellement en évaluation, avec un promédicament de phénol anticancéreux pour activation par un peptidase rapporteur. Ces espaceurs ont également du potentiel pour une utilisation comme adaptateurs chimiques en générale, par exemple dans des oligomères bio-répondeurs.

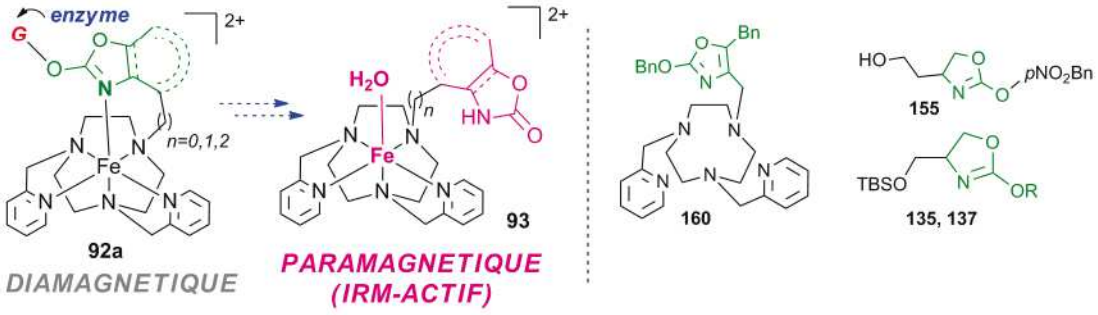
La deuxième partie se focalise sur des espaceurs auto-effondrables pour utilisation dans des sondes pour l'imagerie par résonance magnétique dans le mode *éteint–ALLUMÉ*. Le principe de cette magnétogenèse serait la transformation enzymatiquement déclenchée dans un complexe de fer(II), de l'état bas-spin (diamagnétique) vers l'état haut-spin (paramagnétique). Selon la conception du groupe, cette transformation serait le résultat d'un changement dans la première sphère de coordination d'un complexe ferreux chélatant, d'un motif N_6 à un motif N_5 (eg. **2b**→**2a**). Mes travaux ont ciblé la détection de l'activité de glycosidases, pour laquelle deux familles d'espaceurs ont été explorées.

La première famille d'espaceurs était les *glycosides 2-furanoliques*. Le choix de ceux-ci était motivé par une exploration précédente infructueuse des quinones méthydes dans le groupe : ces espaceurs n'ont pas immolé une fois le complexe ferreux formé (**2b**). Les espaceurs auto-effondrables furanoliques étaient conçus pour montrer une plus grande tendance à éliminer, mais la stratégie éliminante générale des travaux précédents a été retenue (donc, nouveaux cibles tels que **44b**).



Les présents travaux de thèse ont exploré trois voies synthétiques principales vers ces cibles de type glycoside 2-furanolique: cycloisomérisations soit avant soit après installation du groupement pyridyl coordonnant, et substitution nucléophile aromatique. Ainsi, les premiers glycosides 2-furanoliques attestés ont été synthétisés par deux voies différentes et leur fonctionnalisation a été explorée. Toutefois, après de nombreux essais, l'instabilité hydrolytique/éliminant de ces espèces et la réactivité élevée du motif alkoxyfuran, a conduit à l'abandon de cette stratégie considérée comme peu prometteuse pour d'applications *in vivo*.

Une deuxième famille d'espaceurs était alors abordée, avec un mode de fonctionnement de coordination→décoordination (**92a**→**93**) plutôt que d'élimination comme essayé précédemment. Ceci était choisi pour réduire le défi énergétique de la transformation conduisant au magnétogenèse, pour ainsi rendre les cibles moins instables que les furanols. Des carbimidates cycliques (oxazolidinones, oxazolones et benzoxazolones *O*-alkylés) ont été sélectionnés comme espaceurs tautomérisants.



L'attachement de substrats / modèles de substrats par *O*-alkylation a posé des défis synthétiques initials. Une voie de synthèse a été mise au point pour permettre l'attachement des substrats par l'intermédiaire d'un espaceur auxiliaire de type PHBA/PABA. Au moment de la rédaction, divers espaceurs alkylés avec des substrats modèles, tels que **135**/**137**/**155**, ont été isolés ; notamment, un ligand modèle (**160**) est prêt à la purification pour une étude évaluant l'obtention ou non de l'état diamagnétique par ce type d'espaceur. Des travaux synthétiques actuels espèrent pouvoir conclure rapidement sur la faisabilité de cette stratégie de coordination-décoordination, et orienter de futurs travaux dans l'équipe.

En somme, cette thèse présentera des explorations synthétiques et analytiques autour des sondes d'activité enzymatique basées sur des espaceurs auto-effondrables originaux.

Introduction

in vivo probes

for hydrolase imaging

Table of Contents: Introduction

I. INTRODUCTION

1. Molecular Imaging of Enzyme Activity

1.1. Motivation

Enzymes are the crucial catalysts for almost all biological processes. Despite the overwhelming complexity of biological pathways, involving thousands of actors to achieve a typical overall functional or anatomical change, key functions – eg. activation of signalling cascades or process-limiting reaction steps – are often performed by specific enzymes.[1,2] For example, HIV-1 protease is a crucial enzyme in the progress of HIV infection[3]; caspase-3 is a critical peptidase for initiating apoptosis in mammals[4]; and prostate-specific antigen is a peptidase which is well correlated with manifestations of prostate cancer[5].

Molecular Imaging is the quantification and imaging of the activity of specific target enzymes in the cuvette (*in vitro*), in cells (*in cellulo*), in tissues or in whole organisms (*in vivo*). By unpicking the networks of enzyme activity which underlie biology, molecular imaging can promise a functional understanding of the molecular bases of biological processes and disease. It can also promise early and precise diagnoses of functional changes or malfunctions, well before anatomical changes are evident, and therefore allows for earlier and better adapted treatments.[1,2]

Classical investigative techniques (eg. antibody conjugates, some reactive probes) only measure an enzyme's total concentration. However, a single enzyme may show significant differences of activity or even function, depending on a variety of factors including its biochemical form (eg. monomer/dimer, cofactor-associated, acetylated or phosphorylated or with more significant post-translational modifications), folding, the cell type in which it is expressed, or its cellular localisation.[6,7] Rather than examining enzyme concentration, molecular imaging instead aims to quantify enzymatic activity towards a given substrate, which allows for these biochemical variations in enzyme role and activity. Therefore molecular imaging is a significant and necessary functional advance over classical methods, as it alone can conclude on an enzyme's true functional effects.

1.2. High-sensitivity off-ON probes

Molecular imaging can exploit an enzyme's activity to profit from catalytically amplified signal readout: the enzyme processes multiple probe molecules, so sensitivity is potentially high even for small enzyme activities (rather than 1:1 detection by classical methods).[1,2] One promising amplified approach to high-sensitivity molecular imaging is the proagent strategy.[8] The probe is applied in a proagent form, which should be converted into an agent form upon the catalytic activity of the target enzyme, in a manner that reflects the enzyme's activity on its natural substrate. If the proagent and

agent forms show substantially different signal characteristics, then monitoring the signal output should give molecular information about the activity of the target enzyme.

A proagent may generate no signal whatsoever ("off"), but when its substrate is processed by the enzyme, the agent may then generate a signal ("ON"): this is an *off-ON mode* probe. The benefit of the off-ON mode is that all the observed signal can be attributed to the activation of the probe, and the more carefully the signal is collected, the better the detection sensitivity will be. If the acquisition parameters allow for a low inherent background, then the off-ON probe will give unambiguous, highly sensitive detection of enzymatic activity. Alternatively, if proagent and agent forms both generate signal, just in different amounts, this is a *relative change mode* probe; probes operating in an *ON-off mode* are also seen. Nearly all existing probes are relative change mode probes: eg. enzyme-sensitive Gd(III)-based MRI probes, and most FRET-based or electronic modulation fluorogenic probes. However, for molecular imaging *in vivo*, off-ON probes are an absolute requirement (Figure 1).

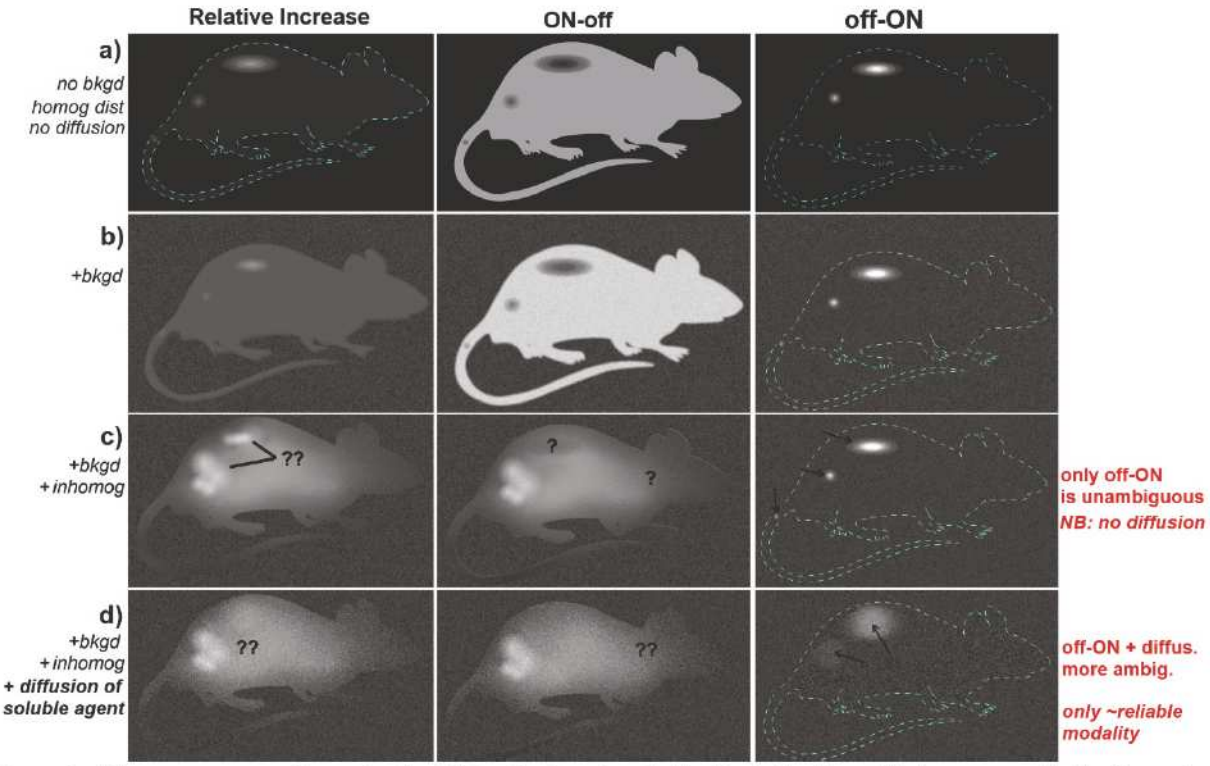


Figure 1 – Three common probe modes for Molecular Imaging, in the ideal situation of high enzyme activity, shown in practice in four common situations: a) *in vitro* style: static and homogenous proagent distribution, and background subtraction; b) colony assay style: no background subtraction; c) ideal *in vivo* style: proagent dynamically and inhomogenously distributed, but agent cellularly retained; d) typical *in vivo* style: active agent also diffuses.

There many arguments for the absolute necessity of the off-ON approach for *in vivo* imaging[1,2] (and both longer discussions, and analytical models, are given in sections I.1.3 and I.3):

Sensitivity: For *in vivo* imaging, small enzymatic activities must be quantified as sensitively as possible. So the ratio of the true signal (small amount of activated agent) to the background component (much proagent + background fluctuations) in the collected signal, must be maximised.[9,10] The best mathematical way to do this is to use a proagent which gives a minimum of signal.

Practically too, the absolute background signal limits the detection sensitivity by saturating the detector and drowning regions with small differences of signal. Both relative change and ON-off modes give a high background component from the proagent form, so are ill-suited to the sensitive detection of realistic amounts of enzymatic activity in biologically relevant situations (Figure 1b-1d).

Reliability: For true off-ON probes, the measured signal gives the concentration of the activated agent, which then correlates to the target enzyme's activity. Yet, for the other modes the measured signal is a function of the proagent concentration too. With static, homogenous distribution and background subtraction (cf. *in vitro*), all three modes can give good quantification, because the proagent signal component can be factored out (Figure 1a). Yet *in vivo*, proagents are inhomogenously and dynamically distributed in the organism from injection till acquisition.[11,12] Unless the instantaneous concentrations of the proagent can be orthogonally determined at all points in the subject, its contribution cannot be factored out of the measured signal: so the true enzymatic signal component of both relative change and ON-off probes cannot be reliably determined. As background subtraction is also usually impossible, experiments using such probes potentially give no true molecular information whatsoever (Figure 1c). Considering too the effects of diffusion of the activated agent, the unambiguity and high sensitivity of the off-ON mode are vital for reliable *in vivo* imaging (Figure 1d).

The simple, practical solution for reliable high sensitivity is to design off-ON mode probes.[13]

The two molecular imaging modalities which best lend themselves to the off-ON approach are optical imaging and MRI. SPECT, PET, CT and ultrasound currently rely more on the biodistribution of their contrast/signal-generating agents to localise potential enzymatic activity, rather than on that enzyme activity to molecularly trigger signal generation.[14] Therefore this thesis will treat off-ON molecular imaging in the stricter context of MRI and optical (specifically, fluorescence) probes.

1.3. Modularity: the role of spacers in probes (and prodrugs)

Consider first the highly specific, off-ON small-molecule proagent capecitabine[15]: three successive enzymatic reactions, in the liver then in its tumour target, convert the substrate-mimic proagent to the active agent 5-fluorouracil (Figure 2). This gives capecitabine high specificity for its enzyme pathway: but it is hard to apply its design to target other critical enzymes, or to release other active agents.

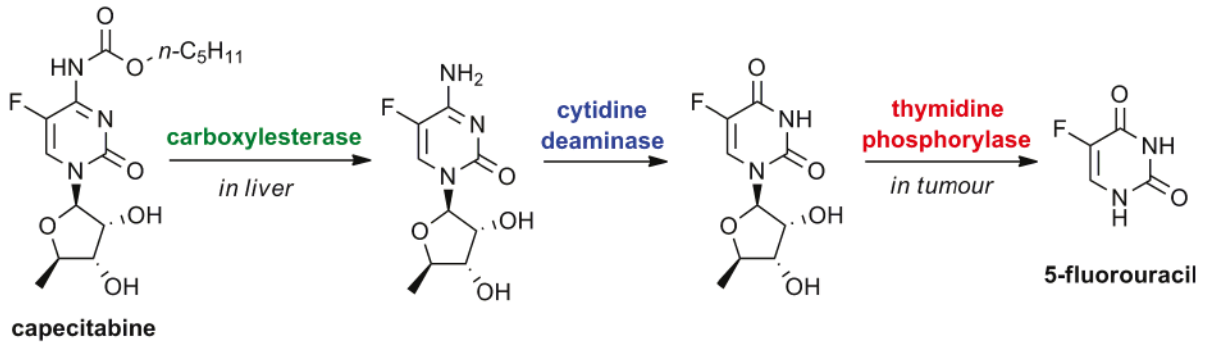


Figure 2 – The highly specific, non-modular design of capecitabine, a proagent for 5-fluorouracil.

The incredible specificity attainable with biomacromolecules (eg. antibody cojugates) also promises highly specific imaging or treatment, in theory: but smaller, cheaper, druglike[16] molecular probes and therapeutics may be the only practical way to achieve these outcomes *in vivo*:

> *"[Biomacromolecules are] very limited due to pharmacokinetic and drug disposition limitations at both the tissue and cellular level… [those] with intracellular targets will only be effective… if they avoid degradation, hepatic/reticuloendothelial system uptake and renal excretion, traverse the microvascular endothelium, cross target cell membranes and escape degradation in the endosome-lysosome system. Subverting these barriers has contributed to the requirement for large doses, and low therapeutic margins observed."*[17]

By contrast to these highly specific strategies, consider a *modular* off-ON molecular design, where a domain of agents can be 'mixed and matched' with a range of masking groups sensitive to different enzymes. By varying the choice of agent and masking group, probes/proagents can be developed for a wide variety of applications (Figure 3). Such modular designs have been explored for decades,[8] and provide many pharmacological as well as conceptual advantages in research and medicine.[18-20] For example, a molecular imaging proagent which selectively images diseased cells in the presence of healthy tissues, may provide a design for prodrugs offering as selective a treatment of that disease without systemic side effects (change of payload). Or, a proagent which releases a chromophore after the action of a reporter enzyme may later find a higher impact application releasing an anticancer drug, triggered by an enzyme of true *in vivo* interest[21,22] (change of payload and substrate). This modularity is especially important when an agent of interest cannot be directly linked to a desired substrate because the agent does not possess a suitable 'handle' group to give a stable, yet enzyme-activatable proagent (chemistry problem); or when that direct linkage results in bad enzyme processing of the substrate moiety, often due to steric congestion/low recognition (biological problem)[23,24].

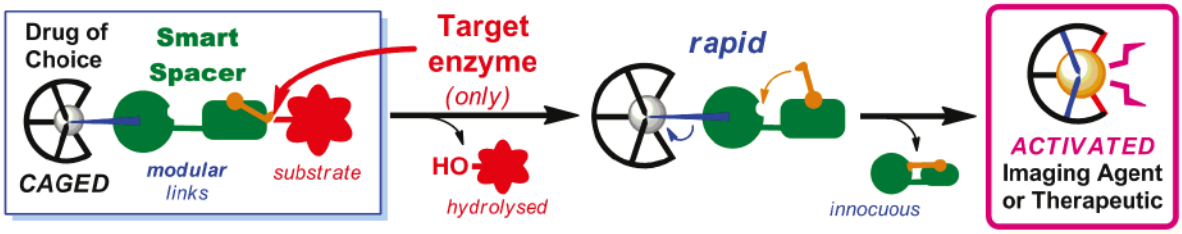


Figure 3 – Modular, three-component designs for enzyme-activated proagents, based around auto-immolative spacers (green), may permit a range of enzyme-specific applications in probes and prodrugs (illustrated for hydrolases).

The chief way to achieve modularity is to split a probe's design requirements into functional units which can each be implemented by a certain chemical moiety (eg. an enzyme-cleaved substrate, a signal-emitting moiety, a solubilising moiety, a chemical- or light-responsive switch). Ideally, a fully modular probe design would allow relatively orthogonal reactions to assemble finished proagents at a late synthetic stage, by synthetically feasible, cheap, and generally-applicable chemical reactions joining the components. Crucially, these joints may often involve the use of chemical adapters,[25] also known as smart spacers or intelligent linkers, connecting the parts in a functional way, much as do

logic junctions[21] in circuit design. **Auto-immolative spacers**[8,26] are the most widespread smart spacers, and may used as "traceless" inserts between two (or more) moieties when their direct attachment to each other is unfeasible or undesired. In a configuration A–*auto-immolative-spacer*–B, the spacer must ensure that once the bond A–*spacer* is broken, the bond *spacer*–B rapidly breaks too.

The functional chemical space of groups tolerated by a spacer at the A and B positions therefore defines the classes of proagent it can be used in. For example, a glycosidase-sensitive amine-releasing probe requires a spacer tolerating both a glycoside at A, and an amine at B (note: spacers can also be 'stacked' one after another to give overall connectivities A–*spacer1*–*spacer2*–B which are otherwise chemically impossible).[21,27] Two examples of spacer connectivities discussed more extensively later are now introduced: the cyclisation spacers featured in **1a**/**34b** adapt an N-terminal substrate (leucine) onto an O-terminal agent (fluorophore **HPQ**); and the eliminative spacer featured in **2b** adapts an O-terminal substrate (galactose) onto an N-terminal agent (macrocyclic amine complex **2a**; Figure 4).

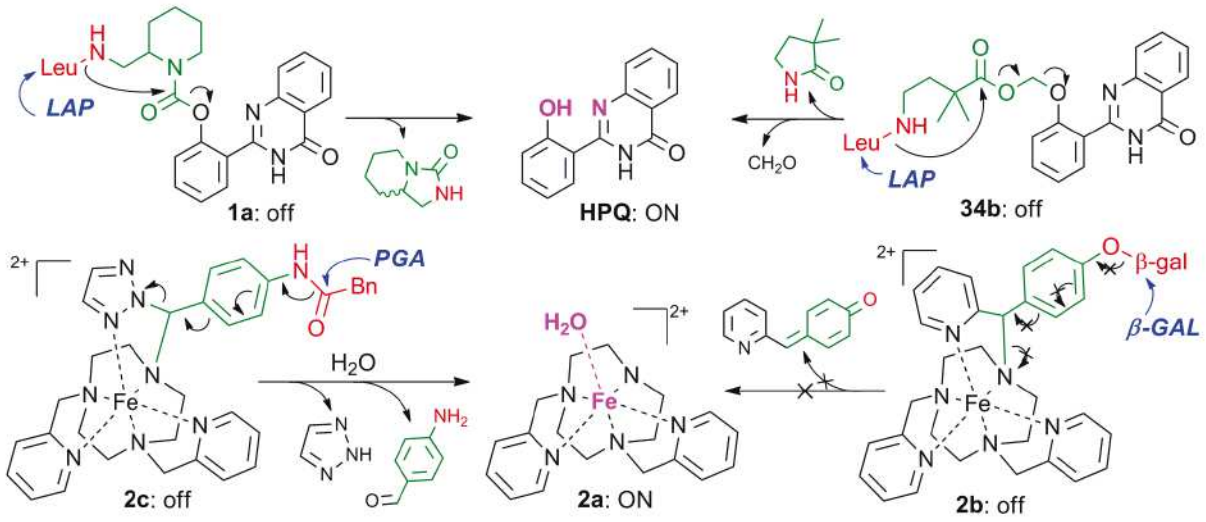


Figure 4 – Recent proagent designs in the Hasserodt group. LAP: leucine aminopeptidase; Leu: leucyl; PGA: penicillen amidase; β-GAL: β-galactosidase; β-gal: β-galactosyl.

Both proagents **1a**[28] and **34b**[29] cyclised rapidly after enzymatic cleavage of their leucine substrates to release their **HPQ** fluorophore, yet **34b** was not in practice a success (Part I). Also, the PHBA eliminative spacer has been used for more than 30 years, yet in design **2b** it failed to show the desired auto-immolative behaviour after enzymatic activity:[30] instead, redesigned probes such as **2c** were required to give enzyme-triggered release of **2a** (Part II).[31] **The careful design of functional auto-immolative spacers with a suitable adaptive chemical space is the limiting factor in the conception and application of enzyme-activated proagents based on modular designs**; and the development of spacers which can be modularly applied to the synthesis of proagents of type A–*spacer*–B where a direct bond A–B is undesirable or impossible, is a vital goal for both Molecular Imaging and Molecular Medicine. This PhD will describe research into just such spacers, at all times considering other vital criteria: including cost and synthetic feasibility, but chiefly, the *in vivo* suitability of the probes based around them.

2. Design Principles for Enzyme-Triggered off-ON Probes

2.1. Crucial general characteristics for in vivo *probe use*

> *"Emerging and more challenging biomedical applications such as* in vivo*... imaging also require the use of [off-ON] probes...* ***numerous [probes]*** *suitable for the visualization of biologically relevant molecules... and enzyme activities... in various* in vitro *and* in cellulo *models,* ***have been developed. However, few of them fulfill all the requirements for*** **in vivo** ***use.***"[10]

Selected *in vivo* requirements are briefly introduced here; comprehensive lists are available.[1,32,33] Primarily, *in vivo* detection and imaging of enzymatic activity is limited by the ability to **establish a sufficient concentration of the activated agent on-site**[1], especially considering that *in vivo* studies (should) aim at very low enzyme activities, present at native (not transfection-enhanced) levels. So:

(1) The proagent must be able to reach its target enzyme, which implies good water solubility[34,35], membrane permeability[36], and pharmacokinetics otherwise no *in vivo* experiment can be performed. If probes require *in vitro* characterisation in organic solvent mixtures,[10,37] this is an unpromising sign!

(2a) Upon encountering the enzyme, the proagent is quickly processed[38], and it quickly dissociates[39] from the enzyme to give optimal turnover. If the probe is processed to a metastable intermediate state, that intermediate must very quickly[26] transform to yield the active signalling agent. Probe designs cannot afford to ignore toxicology: designs which inhibit[39-41] or inactivate[42] the target enzyme (or kill the organism under study) are always unacceptable; those which eventually poison the cell are unacceptable except in the case of (extremely!) selectively-activated cytotoxics[43].

(2b) The observed signal must be correlated with a causative level of enzyme activity. Importantly, this requires that signal only be generated by activity of the specific target enzyme, and no signal be generated by spontaneous probe activation: "the *off* state STAYS *off* during the experimental lifetime."[44] Spontaneous signal generation really implies a relative change mode: the observed signal varies from [spontaneous level] to [spontaneous level plus enzyme contribution]. Therefore the detection sensitivity of unstable probes in real settings is just as low as that of the relative change probes introduced in section 1.2, but with the added disadvantage that the maximum relative change constantly decreases over the experimental lifetime, as more of the probe hydrolyses.

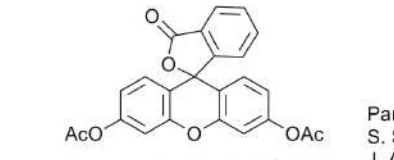
(2c) The activated probe should stay where it was activated and not diffuse, or else the signal cannot be reliably localised[44,45] or even retained at all[11] (Figure 1d). An ideal activated agent would accumulate[11] inside the cells which liberated it[46], over a long experimental time during which it must not biodegrade, or otherwise lose its signal-emitting capacity.[32] Once the activated probe is in place,

(3) A 'sufficient' concentration depends on the inherent signal-to-background ratio R of the experiment. For *in vivo* imaging, where the target enzymatic activity (and so, the potential signal) is small, this ratio usually depends far more on the experimental background signal B_{ACQ} rather than the

signal brightness S of the emitter. B_{ACQ} differs greatly between imaging modalities, as do other sensitivity-affecting factors like signal absorption and background increase by scattering: compare the required probe concentrations for PET (<nM) to MRI (~mM)[13]. However, an agent's specific signal-emission characteristics may allow selective signal collection only from the probe, and not from the background: this greatly increases R almost regardless of S, and promises high-sensitivity probes. The signal from the proagent must naturally also be minimised ("the *off* is REALLY *off*") particularly if we seek to visualise a small signal which may otherwise be swamped in a large proagent background.

An ideal off-ON probe could be applied in high concentration (bioavailability) without generating signal from unprocessed probe (off is really off); as rapidly as possible the activated form would accumulate (good processing speed and no loss by signal degradation/diffusion), over an arbitrarily long experimental time (no nonspecific or spontaneous signal generation), until even a small enzyme activity gave rise to a cumulative signal above the background level (inherent R).

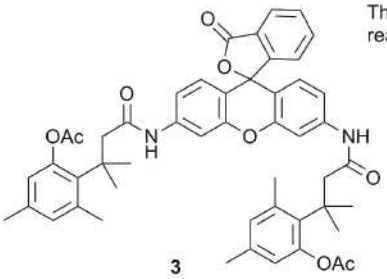
For an example relevant to this thesis, consider the commercialised, phenolic, esterase probe fluorescein diacetate (FDA), which Raines compared to his own, more stable esterase probe design **3**.[44] Both proagents exhibit low absolute signal relative to the respective activated agents (*potential* off-ON design), but FDA generates large nonenzymatic signal by spontaneous hydrolysis. Therefore by (2b), FDA really functions as a relative change probe whose potential sensitivity rapidly diminishes to zero. When incubated with cells for only two hours, it indeed generated so much nonenzymatic signal that cells could not be resolved. Therefore despite the potential of fluorescein diacetate as an off-ON probe, in a practical setting the probe was useless due to its low stability. By contrast, Raines' more stable off-ON probe **3**, exploiting smart spacers to create more stable phenolic acetates, gave selective imaging of cell esterase activity. As a general rule, probe robustness can easily outweigh traditional *in vitro* considerations such as signalling agent brightness, as the crucial determinant of probe success *in vivo* (Figure 5; see too section 2.3).



fluorescein diacetate (FDA)

Panels A, B and C in this figure had been reproduced from
S. S. Chandran, K. A. Dickson, R. T. Raines;
J. Am. Chem. Soc. 2005, 127, 1652-1653 (DOI: 10.1021/ja043736v);
they were originally Figures 2A, 2B and 2C respectively (page 1653).

These panels have been removed from this distribution for copyright reasons.



3

Figure 5 (from Chandran *et al.*[44]) – (A) Nonenzymatic signal generation for two-component phenolic probe fluorescein diacetate (FDA; upper curve, red) vs Raines' three-component probe, 3 (lower curve, blue), in DMEM medium (main) or PBS (inset). Cells were imaged after incubation with FDA (B), or probe 3 (C).

2.2. Probe architectural issues: introduction

A quick introduction to the major architectural classes of molecular probes is given (Figure 6).

Reactive Probes are here defined as single-component probes which may be processed by the target enzyme by effecting a functional group transformation, yielding a product which displays a signal output. Generally these probes are limited to nitroreductases (eg. **4**); reactive prodrugs also usually target oxidoreductases, eg. tirapazamine.[19,47,48] It is impossible to modulate either the targeted enzyme or the active agent generated as for multicomponent probes.[49]

Two-component probes are generally based on an enzyme substrate which largely mutes a signalling molecule to which it is attached (eg. commercial **5**); multiply-substituted probes such as FDA conceptually also belong to this class. Enzymatic action either changes the substrate into a non-muting form[50] or, more often, cleaves the muting group entirely from the signaller[51,52]. The 2-component approach cannot "mix and match" substrates and signalling molecules, either: an esterase probe cannot be made from an amine signalling moiety, nor an aminopeptidase targeted with a phenolic drug. Depending on the difference of chemical nature and of steric demand of the payload[53] relative to those of the enzyme's natural substrate, enzymatic processing may be slow *in vitro*:[54] if so, then when the probe is in competition with the natural substrate (*in vivo*), the turnover is likely to be minimal[55] so the probe will be useless. The majority of current substrates for reporting on enzyme activity are 2-component substrates.[56]

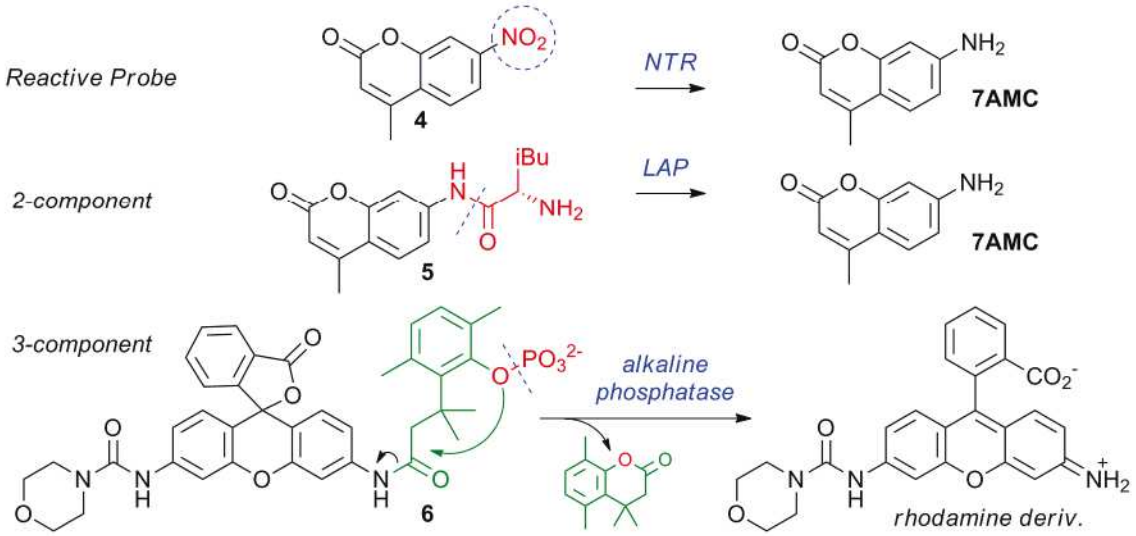


Figure 6 – Examples of fluorogenic probes illustrating chief probe architectures. 4: reactive probe for nitroreductase (NTR).[49] 5: 2-component probe for leucine aminopeptidase (LAP). 6: 3-component probe for alkaline phosphatase[41].

Three-component probes are broadly composed of a substrate bonded to a smart spacer, which is itself bonded to the agent of interest: a signalling moiety or drug. The smart spacer is crucial for achieving modular, mix-and-match design of a range of proagents or prodrugs targeting different enzymes (depending on its tolerated chemical space; cf. section 1.3).[25] For example, a peptide-to-phenol spacer achieves a connectivity that is impossible for 2-component probes; and it may as easily

allow molecular imaging of the activity of a reporter peptidase with a phenolic fluorophore (**1a**), as gating the release of a chemotherapeutic to the enzymatic activity of a cancer biomarker peptidase for targeted therapy. The phosphate-to-aniline spacer in **6** similarly addresses a connectivity problem.

Well-designed spacers should favour fast enzymatic processing of the probe[28], at comparable rates to that of the native substrate, so that similar results are obtained *in vivo* as *in vitro*. Modular spacers should furthermore ensure that this can be extended to any enzyme (in a similar class) with similar results. Robust enzymes such as common reporter enzymes may cleave almost any probe bearing their pre-scissile substrate – so they are often used for proofs-of-concept. However, if the post-scissile portion of the spacer/probe is not specifically planned to ensure it, a design will rarely be modular: it will not allow targeting other enzymes when the substrate is altered (eg. moving from targeting an *in vitro* reporter enzyme to a particular enzyme of *in vivo* interest). For example, a functional β-galactosidase prodrug design (**7**) may not translate to β-glucuronidase (**8**)[57]; or the N,N'-dimethylethylene diamine spacer in **9** that gives rapid auto-immolative behaviour after (indirect) liberation by eg. catalytic antibody 38C2[58] or nitroreductase[59] may not allow targeting plasmin[60] (**10a**) or penicillen amidase[23]. Careful design of extensible spacers is vital (Figure 7).

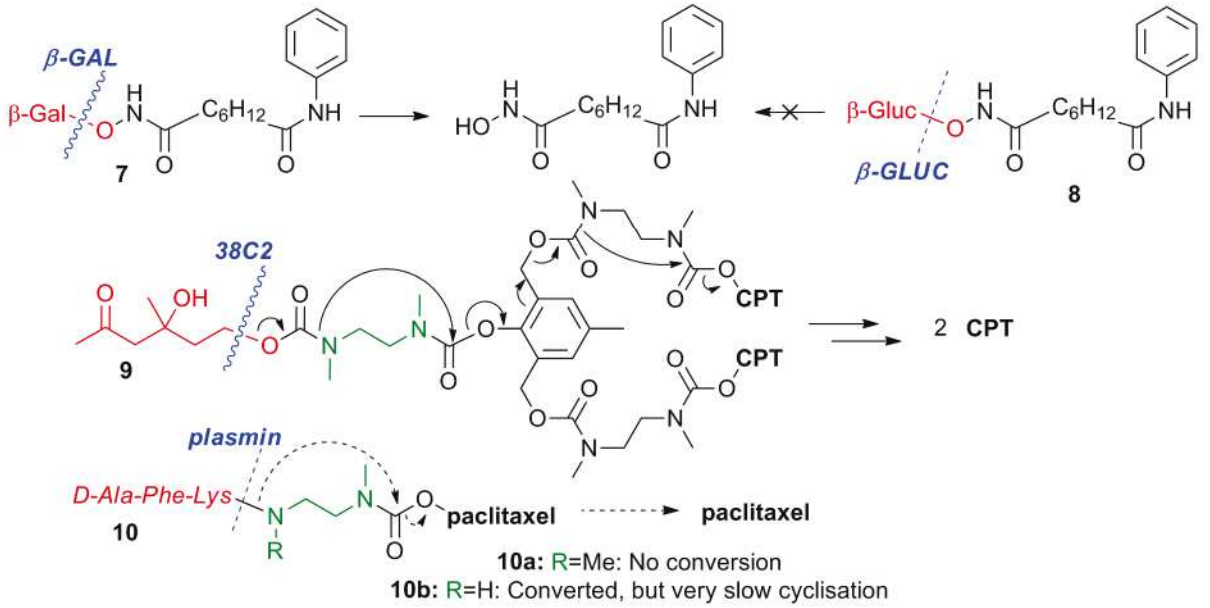


Figure 7 – For modularity, enzyme specificity must be considered, or else probes may not be processed by the target enzyme. β-GAL: β-galactosidase; β-GLUC: β-glucuronidase; 38C2: catalytic antibody 38C2; CPT: camptothecin.

Pseudo-Three-component probes: Multifunctional spacers which may be triggered by more than one enzyme (eg. diethylenetriamine[22]), or may release more than one molecule (eg. 2,4-dihydroxymethyl-aniline[43] or 2,6-dihydroxymethyl-phenol as in **9**) have been used in self-immolative dendrons[25,61]. As to their function, these are 'just' higher-dimension extensions of three-component probes and can be understood with the same considerations, so they are classed here. Similarly, probes with two or more linearly-stacked spacers (as featured in **9**) also function as 3-component probes.

2.3. Hydrolase probes: specific architectural and in vivo issues

This PhD was concerned with the design and synthesis of off-ON probes for hydrolase activity. Three specific problems plague hydrolase probes more than probes for other classes of enzymes, and each problem must be addressed by designing an appropriate probe architecture.

(1) **Spontaneous Hydrolysis**: The main problem which especially affects *in vivo* hydrolase probes is their intrinsic hydrolytic instability. These probes target enzymes which hydrolyse a crucial bond, whereupon the active agent must be released. If this bond is also susceptible to spontaneous hydrolysis by water (55 M, 37 °C, 5<pH<8), the probe will spontaneously generate signal which is uncorrelated to true enzymatic activity. This greatly reduces its detection sensitivity, and cannot warrant a description as truly off-ON (see Figure 5 and section III.3): so **instability must always be evaluated**.

Notably for this PhD, hydrolase probes usually display the following crucial bonds: phenolic phosphate ester (phosphatases)[62], phenolic ester (esterases)[63], anilinic amide (exopeptidases)[64] or true peptide (endopeptidases)[65,66], and phenolic glycoside (glycosidases)[67]. Note that although more stable crucial bonds are sometimes proposed (eg. aliphatic glycosides)[68], these may be such poor enzyme substrates, especially in competition with native substrates, that they are useless *in vivo*.[69]

There are some clear, general stability rules. Anilides and peptides (eg **5**) are usually stable to spontaneous hydrolysis. However, phenolic compounds are problematic: phenolic esters are known to be rather unstable in water[23,70]; phenolic glycosides less so; and phenolic phosphates the least of these three: but even phenolic phosphatase probes may hydrolyse at rates which demand precautions with their use[41,71]. Phenolic bond stability heavily depends on the leaving group quality of the phenol (and less so on pH). The problem especially for optical probes is that as a general rule, chromophore / fluorophore phenols are all well-delocalised and many are electron-withdrawn, therefore they are exceptionally good leaving groups (see Part I).[10] Therefore **the practical sensitivity of two-component phenolic hydrolase probes may be too limited** for true *in vivo* work. Two-component fluorogenic esterase probes are thoroughly undesirable and degrade too rapidly for *in vivo* experiments (recall FDA in Figure 5)[72,73]. To achieve robustness for *in vivo* use, fluorogenic glycosidase probes should also avoid two-component designs; even phosphatase probes[74] have been shown to give crucial improvements in robustness by using three-component designs.[41] Even phenolic bonds traditionally considered relatively stable[75], such as benzyl phenolic ethers (**11**), will degrade spontaneously too fast for use *in vivo* if good leaving-group phenols are used[10].

(2) **Proagent Signal Quenching**: Hydrolase-triggered activation often relies on removing a masking group from a latent active moiety[76]. However, total masking of the signal from a molecule is very difficult to achieve;[10] in this respect totally-masked prodrugs may be simpler to design, eg. by sterically blocking a critical pharmacophore with almost any type of masking group. Commonly, hydrolase probe publications claim off-ON functioning based on *liberating* a known active agent, but

actually fail to prove near-total signal quenching in the proagent form. Either a *mechanism* critical for signal generation must be entirely blocked by the masking group, or the masking group must shift the proagent signal so far from that of the free agent so that the two are *baseline-separated*: otherwise there is no justification for claiming off-ON mode sensitivity, but at best, only a large relative change.

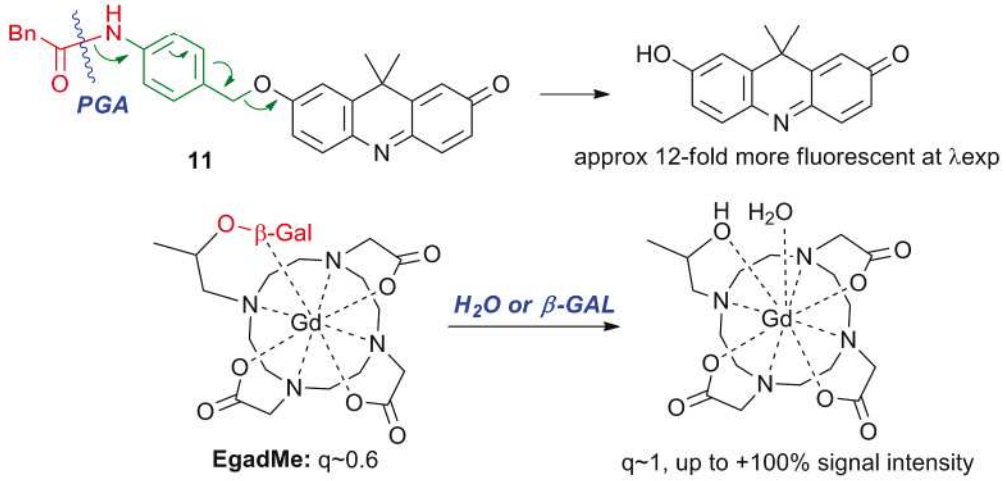


Figure 8 – Covalent modifications which do not give off-ON behaviour, as proagent signal quenching is incomplete.

Commonly, covalent modifications more or less strongly alter the properties of a signalling molecule – eg. by changing the electron density – but examples of total, mechanistic signal blocking are rare.

Example 1: The fluorescence of most soluble phenols depends on conjugation/electron donation from the phenol/phenolate. Trapping the phenol eg. as an ester, carbamate or ether, reduces this and so the less-conjugated and less electron-rich trapped phenol usually adopts a fluorescence spectrum which is usually both less intense and spectrally-shifted[72]. However: this spectral shift is almost never enough to give the baseline separation that is required for true off-ON behaviour (cf. **11**[10], Figure 8).[10,72] Aniline fluorophores too may give only small fluorescence changes upon covalent modifications, eg. only +500% enhancement for an amide→aniline unmasking.[56,77] Also, proagent quenching by covalent modification becomes more difficult as the proportional change on the signal molecule decreases: a particular problem for longer-wavelength fluorophores needed *in vivo*, as these are necessarily more delocalised/functionalised (discussion developed further in section I.1.3.1).

Compare these complications to true mechanistic signal blocking eg. by masking the Excited State Intramolecular Proton Transfer process (ESIPT; section I.1.3.4). Without the crucial proton, ESIPT signal generation is not reduced, or shifted: it is entirely eliminated.[78] ESIPT-unmasking probes are therefore truly off-ON: eg. **1a→HPQ** (Figure 4) gives >+200,000% signal enhancement on activation.

Example 2: The relaxivity change of many enzyme-sensitive MRI contrast agents relies on changes in water coordination. For example, in **EgadMe**[79], changing the inner-sphere water coordination number *q* upon (putative) enzymatic activity increases the MRI signal intensity. Yet, commonly, half of such agents' relaxivity (signal enhancement) is due to second-sphere effects which remain constant.[69] As no critical mechanism is un/blocked, their maximal signal change is only around +100%[79] (Figure 8).

Instead, magnetogenic probes rely on diamagnetic proagents incapable of signal enhancement, so only after activation of paramagnetism is signal seen (**2c→2a**, Figure 4): mechanistically off-ON.[31]

Non-total signal quenching in the proagent state gives (a) lowered sensitivity *in vitro*; though often, this is still considered acceptable, because *in vitro* background and proagent signals can be subtracted (Figure 1a) and diffusion can be ignored in this essentially homogenous medium (a cuvette is a single voxel). However, non-quenched designs actually give (b) minimal or zero sensitivity *in vivo*, because it is impossible to apply such 'corrections,' and diffusion can no longer be ignored (Figure 1d; see too sections I.1.3 and I.3.1). Therefore, **truly off-ON designs are an absolute requirement *in vivo*** (not a luxury); and mechanistically off-ON probes may represent the best approach to implementing them.

(3) **Balancing Specificity vs Kinetics vs Stability**: Hydrolase probes usually rely on attaching a natural substrate to a non-natural leaving group; for best enzymatic kinetics, this should be one with as low steric presence and as favourable leaving-group nature as possible. Therefore, avoiding **nonspecific enzymatic activity**, as well as spontaneous hydrolysis, while ***still* ensuring good enzymatic processing** kinetics, may be problematic. For example, although electron-rich phenolic esters and even carbonates may achieve long halflives in cell culture media or buffer (>24 h), they are so swiftly and nonspecifically hydrolysed in cells or serum (eg. 12 min halflife!) by a multitude of esterases,[80] that almost[76] no probe designs even aim at selective esterase targeting. The promiscuity and *in vivo* ubiquity of phosphatases is also problematic for their selective targeting[55,81,82]. In contrast, glycosidase,[83] and especially peptidase[66,84] substrate specificities can[85] be high, enabling unambiguous targeting (though for glycosidases, usually, the post-scissile leaving-group must be so good that probe stability becomes a problem).

Importantly for this PhD, many of the known biomarkers for diseased states, as well as healthy biological processes, are relatively specific peptidases and glycosidases: so probes which map their activity may not only increase understanding in biology, but also allow better diagnosis and monitoring of disease.[1,2,33,86-88] This PhD involved designing auto-immolative spacers for just such probes, specifically, for off-ON peptidase-sensing fluorogenic probes and off-ON glycosidase-sensing MRI probes: two promising "soft" imaging modalities for both *in cellulo* and *in vivo* studies/medicine.

In general though, whatever enzyme is chosen, the balance of enzyme processing kinetics with specificity and stability is a difficult problem for *in vivo* designs: thousands of probes have been reported *in vitro*, yet far fewer will ever actually be exploited *in cellulo* let alone *in vivo*. If probes fail to balance these criteria, problems can be swept under the carpet during *in vitro* studies, by using a high concentration of purified enzyme over a short time with no competing substrates. However, if studies using such conditions fail to present explicit evaluations of stability and enzyme kinetics, they almost certainly do so because these would reveal the flaws in their design. Probes whose efficiency is claimed based on such spurious validations will have no *in vivo* value whatsoever.

2.4. Auto-immolative spacers for 3-component hydrolase probes

The chemistry/functional chemical space tolerance of a spacer determines the scope of applications of probes incorporating it. Auto-immolative spacers are primarily divided into cyclisation and elimination spacers, which are briefly outlined here. Further details are given in Part I (cyclisation) and Part II (elimination) as needed; more extensive reviews are available.[15]

2.4.1. Cyclisation spacers for in vivo *probes*

In vivo-suitable cyclisation spacers are commonly based on five-membered cyclisation of aliphatic amines (or thiols), or six-membered cyclisation of phenols, onto carbamates or even amides.

1,2-diamines: Many probes, prodrugs, and auto-immolative polymers or dendrimers have relied on diamines as auto-immolative spacers for an **N→O** connection.[25,60,89] These may be either *enzyme-recognised*, eg. N'-methylethylenediamine **12** (which allows hydrolase processing of an N-terminal substrate eg. a peptide on its primary amine terminus[90]), or else *adaptive*,[89,91] eg. N,N'-dimethyl-ethylenediamine **13** (which does not[23,60]). The spacers connect via an N-methyl end to the carbamate of the alcohol to be released (the agent, or another spacer). **12** (a primary amine nucleophile) gives far slower cyclisation than **13** (secondary amine; eg. $\tau_{1/2}$(**13**)~35 min; $\tau_{1/2}$(**12**)~304 min; Figure 9).

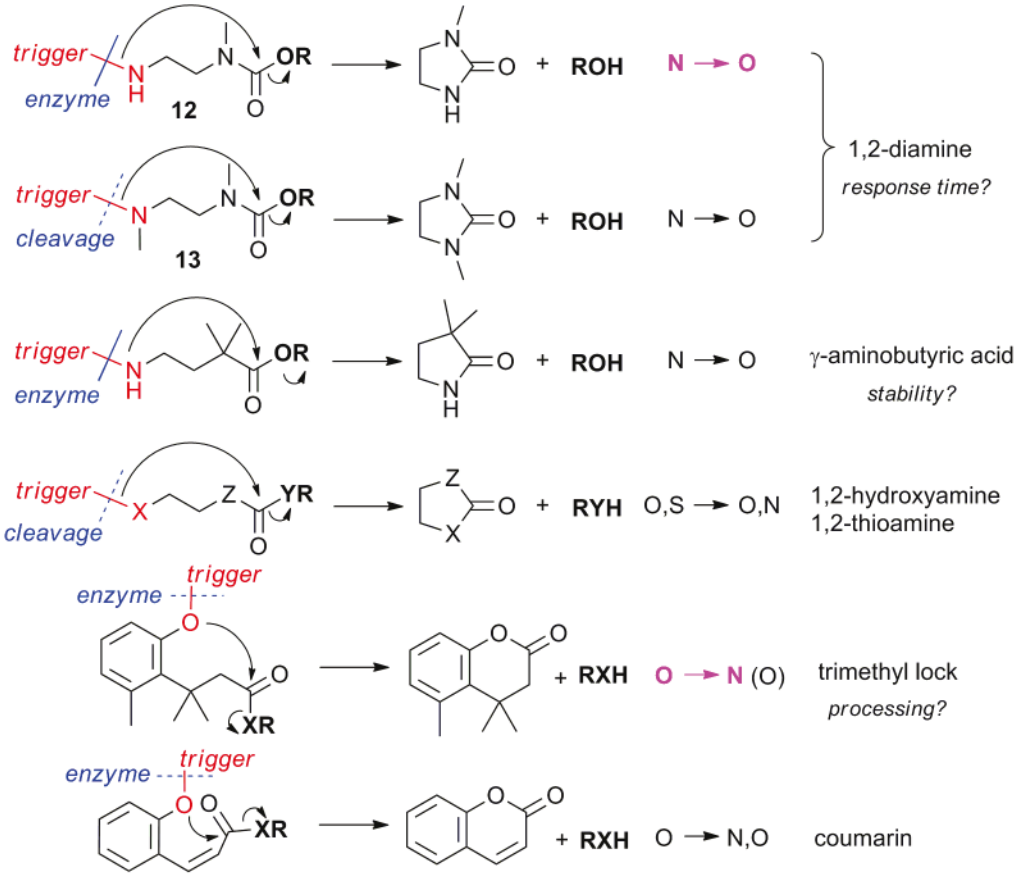


Figure 9 – selected cyclisation spacer types used in prodrug activation

Other cyclisation systems: N-cyclisations of γ-aminobutyric acid derivatives onto esters[92,93] have been used as alternatives to diamine cyclisations for N→O connections, though these may suffer from

nonspecific ester hydrolysis by endogenous esterases[27]. N-cyclisation of dipeptides onto acyl groups such as anilamides[77] or N-acyl nucleotides[94] to form diketopiperazides is also known, though slow. In aliphatic systems, other nucleophiles (especially thiols[95] but also alcohols) can cyclise similarly onto carbamates[96], esters[27], or carbonates[23], permitting the release of either alcohols[97] or amines[96] after hydrolase/cleavage activity on S- or O-terminal substrates. A more powerful general system for oxygen nucleophiles especially is Cohen's trimethyl lock (TML)[98,99], where the phenol cyclises onto an ester or amide with very favourable kinetics.[100,101] The TML phenol is less electron deficient than most chromophores/fluorophores, and may give correspondingly more stable probes (cf. the stability of TML ester **3** vs FDA in Figure 5; see too **14**). However, the steric hindrance of the *ortho*-quaternary carbon may slow enzymatic processing by all but the most permissive enzymes, and may even render synthesis difficult[41]; the solubility of TML species is usually also low[102]. The similar coumarin cyclisation spacers[103,104] are somewhat more sterically permissive but also involve some synthetic problems, and may be less hydrolytically stable (Figure 9).

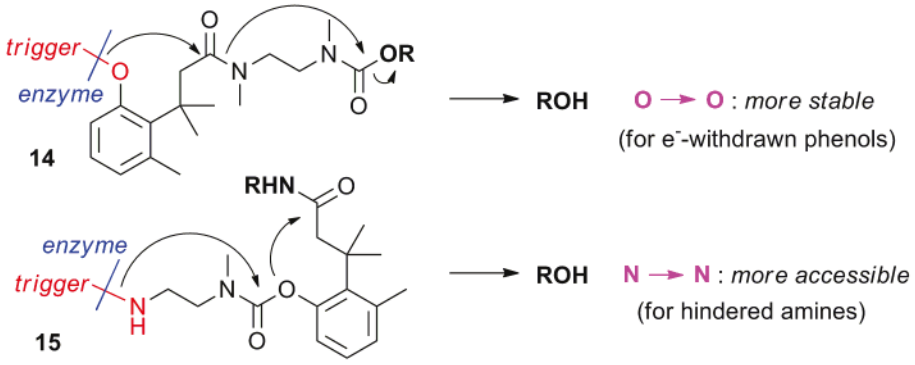


Figure 10 – Stacked cyclisation spacers for solving stability or steric problems with direct attachment

A (peptide-1,2-diamine-carbamyl-TML-amide) spacer stack may be a sufficient *in vivo* solution for **N→N** connection if direct attachment is unfeasible (**15**). For *permissive* O-terminal hydrolases, the TML spacer may be a good balance between hydrolytic stability and good leaving-group character for rapid processing, thereby providing a solution for an **O→N** connection. Likewise, a stacked (X-TML-1,2-diamine-carbamyl) spacer pair has potential as an adapter for **O→O** connection that is more robust than most 2-component designs (**14**), eg for fluorophore phenol release (Figure 10).

Specifically as regards Part I, cyclisations onto a 2°-amine carbamate provide a perfectly stable *in vivo* **N→O** connection, even with good leaving group phenols[28,80,105], and cyclisation spacer byproducts are nontoxic. This contrasts to quinone methide elimination spacers, whose benzyl aryl ethers can be unstable[10] and whose elimination byproducts are notoriously toxic. Therefore 1,2-diamines would be the preferred spacers for peptidase-triggered phenol release, except that the kinetics of prior art enzyme-recognising diamines are too slow for good applications *in vivo*, whereas quinone methides may even immolate before leaving the active site.[26,106] Part I of this thesis will develop new 1,2-diamines that address this **need for a stable, nontoxic amine→phenol spacer with fast response kinetics, finally allowing robust fluorogenic phenolic probes for *in vivo* imaging**.

2.4.2. Elimination spacers

Eliminative spacers are usually based on 1,2- or conjugated eliminations (1,4-; 1,6- etc) of an alkoxide, amide or thiolate, from direct or conjugated hemiaminal ethers, aminals or hemithioaminals.

1,2-elimination spacers allow adapting a terminal heteroatom (eg. N-terminal substrate→thiol via a hemithioaminal ether[34]), or adjusting a probe's stability while retaining the same heteroatom (eg. acyloxymethyl aryl ethers are more stable than aryl esters[72]). Such spacers may simultaneously tune the probe's properties (eg. solubilisation by a carboxylic acid on the spacer[34]). The C=X-bearing spacer byproduct (eg. imine, aldehyde) may be somewhat alkyating *in vivo*.[6] Enzyme-recognised 1,2-elimination spacers do not greatly increase the distance between the agent and the substrate, so may not provide a significant advantage in enzyme recognition (unlike conjugated spacers);[24] also, their different electronic character from natural hydrolase substrates (eg. hemiaminal ether rather than amine) may slow enzyme processing kinetics. Therefore 1,2-elimination spacers are more often used as downstream *adapters* in spacer stacks. Notably, the carbamic acid of an amine is a powerful yet robust and nontoxic **O→N** connection which expels CO_2 to give favourable amide elimination (cf. β-lactamase probe, Figure 11); likewise, the (less robust) carbonic acids eliminate alkoxides (O→O).

Conjugated elimination spacers are more widely used, and usually include an aromatic ring: thus, the quinone methide series, including the 1,6-eliminative *para*-aminobenzyl alcohol (PABA)[8] spacer and its phenol (PHBA)[107] and thiophenol (PSBA)[108] analogues. *Ortho*-elimination (eg. **9**)[107], 1,8-elimination[109] and alicyclic 1,4-elimination[110,111] are also seen;[8,112] and the pyridinone methide explored during the Master 2 thesis leading to this PhD work[113] and later reported by Shabat[114] also falls into this category. The structure of the byproduct quinone methides modulates their well-known toxic[115,116], alkylating behaviour[42,117], but in general, an azaquinone methide can be considered even more cytotoxic than the standard chemotherapeutic doxorubicin[118]. Either inductively[106]- or resonance[119]-activated quinone methides are especially powerful alkylating agents and are well-known as irreversible covalent inhibitors. More conjugated[61] or electron-rich derivatives (such coumarin-based fluorogenic elimination spacers which elegantly combine *spacer and signaller* functions[120]), or hindered derivatives (eg. 6-adamantyl[121] quinone methides), or **tautomerisable** byproducts[111], may be first generation strategies towards less toxic elimination byproducts.[122] However, except for the (low-solubility) adamantyl system, these latter may have slower response kinetics (less energetic incentive to fragment) and all have seen less development so far. A notably elegant (but non-generalisable) alicyclic elimination mechanism is shown by cephalosporin prodrugs, where β-lactamase action on the substrate also triggers its β-elimination of a leaving group[27], so the *substrate and spacer* functions are combined (Figure 11).

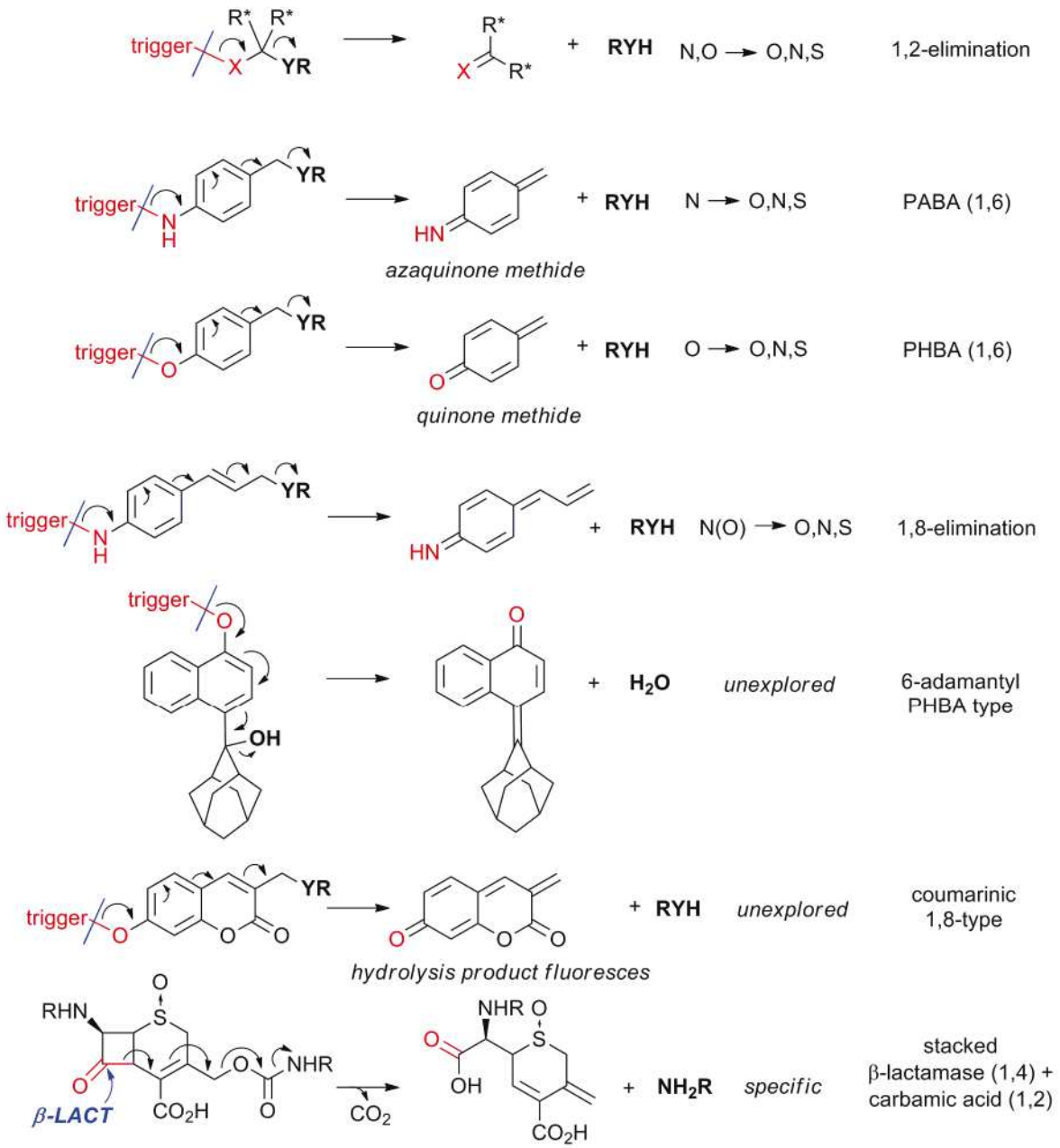


Figure 11 - selected eliminative spacers used in prodrug design

This PhD considered that the alkylating properties of current general-purpose elimination spacers made them highly undesirable for *in vivo* applications in diagnostics and therapeutics (except selective delivery of cytotoxics)[123]. Part II concerns efforts to address the non-fragmentation (insufficient energetic incentive) of the group's prior MRI probe design **2b**→**2a**, by developing new eliminative spacers, if possible with a lower toxicity profile than current eliminative spacers. One angle of attack was simply towards spacers which would eliminate more powerfully than the known repertoire of quinone methides. A second and more exploratory approach was to develop entirely novel tautomerisable eliminative spacers, themselves locked in a pro-spacer form unable to eliminate before enzymatic activity, to circumvent the main energetic challenge inherent in the **2b**→**2a** transformation.

3. In vivo off-ON Hydrolase Probes: Challenges and Visions

Despite more than thirty years of research, the situation in 2010 was accurately summed up by Meyer *et al*.[23]: "the design and optimisation of new enzymatically triggered self-immolative linkers take an ever more critical place in the development of efficient prodrug systems and smart… bioprobes."

As relevant to this PhD, phenol and aniline fluorophores which can work approximately in an off-ON mode and which are tolerated *in vivo* are well known, even if good NIR fluorophores are still lacking (see Part I).[1,10,100] Examples of MRI probes with mechanistic off-ON possibilities are known, even if their development is decades behind that of fluorescent probes (see Part II).[31,124,125] Many substrates that specifically target hydrolases of interest are known, especially for peptidases and glycosidases; more are hotly pursued by researchers seeking the next important suites of biomarkers.[1,7,33]

Yet the auto-immolative spacers are the missing links in the chain. There is no suitable and complete toolbox of spacers that can enable the *à la carte* creation of proagents, with the required features for *in vivo* use, for gating the release of the common classes of functional imaging agents to the activity of the major exohydrolases (terminal phosphatase, exopeptidase, glycosidase, esterase/lipase). The modular targeting of endohydrolases (endopeptidases, inner phosphatases such as ATPases) or minor hydrolases (eg C-S, C-N, ether, halide, or C-C hydrolases) is still less developed. Specific challenges promising wide-ranging applications include: gating alcohol release to exopeptidase activity on N-terminal spacers[23]; developing more robust O-terminal spacers which are rapidly processed by even unpermissive glycosidases and esterases[100]; and developing auto-immolative N-terminal spacers for probes which can be competitively processed alongside native substrates of key endopeptidases[28].

In general, the ultra-sensitive off-ON molecular imaging of specific hydrolase activity is a hope for the future not a reality, even in fundamental research rather than clinical use. Furthermore, without sensitive-enough tools for *in vivo* enzyme mapping, the current knowledge and understanding of biomarkers is lamentably deficient: so even the practical diagnostic *potential* of molecular imaging also remains hidden. Yet some of the major promises of modular, specific, robust, off-ON *in vivo* molecular systems are clear: (a) expanding the fundamental understanding of biological processes by revealing and spatio-temporally characterising the complex and interdependent enzymatic mechanisms underlying them; then (b) providing cheap, reliable and early detection of enzyme misregulation to enable earlier and better disease diagnosis; then (c) swapping molecular imaging's signalling molecules for molecular medicine's therapeutics, to achieve highly sensitive targeted therapies.

Many practical scenarios can be imagined for the application of enzyme-targeted molecular systems. *In vitro,* immobilised probes (eg. on paper support) could be used much as pH paper is today, to "dip & read" the specific activity of a sample towards a standard substrate – or supported probe arrays could fingerprint diseases by studying the activity of a multitude of enzymes at once. *In vivo*, probes could evaluate the progress of gene therapy directly or by quantifying the activity of a reporter gene;

or multimodal and multiplexed imaging cocktails could simultaneously characterise and localise complex multi-enzyme processes in their native state, to deliver new horizons for molecular biology. In the clinic, when diseases are detected, or even suspected, then off-ON molecularly-targeted drug cocktails, or logic-gated single prodrugs eg. *X*-AND-*Y* constructs, may be able to provide undreamt-of, molecular specificity for the diseased cells in the patient, and enable efficient therapy of hitherto untreatable disorders[126].

The greatest current challenge is to design the spacers that will allow these applications.

4. Objectives of this PhD

A concise summary of the background to and objectives of this PhD work is given to orient the reader, before proceeding to the relevant Parts of this PhD.

4.1.1. Part I: Fluorogenic Probes

The Hasserodt group's global objective has been to develop a modular 3-component system for off-ON proagents of phenols, activated by exopeptidases, suitable for use *in vivo*. The system should be applicable in probes and prodrugs, and should satisfy all the design criteria for *in vivo* use: eg. proagent solubility/bioavailability, proagent robustness, rapid and competitive enzymatic processing, rapid spacer immolation, and spacer byproduct non-toxicity. Relevant prior work in the group explored rapid-immolation γ-aminobutyryloxymethyl ether spacers (as in **34b**, Figure 4), but these hydrolysed spontaneously too quickly for good *in vivo* applications. Prof. Hasserodt's current design was for novel, similarly-preorganised 1,2-diamine cyclisation spacers, which were to overcome the sluggish immolation kinetics of extant *in vivo*-suitable N→O connection smart spacers, as well as the spontaneous hydrolytic cleavage of the the spacer-proagent link in the prior CBO group designs.

The task I was assigned was to synthesise and prove the latest 3-component design as applied to sensitive fluorogenic probes, then to examine their performance *in vitro*. The fluorogenic probes were designed to ensure the greatest probe sensitivity, by operating in a microscopically mechanistically off-ON mode using ESIPT fluorophore un/masking. The fluorophores' insolubility gave them the potential for subcellular localisation of enzyme activity with minimal photobleaching and environment effects, and their large Stokes shift could promise minimal tissue background signal, for still greater sensitivity. The quality of the 3-component design would be convincingly tested by creating probes for electron-withdrawn, insoluble ESIPT phenols known for their heightened spontaneous degradation (eg. **1a**, Figure 4). The synthesis also had to be as modular, cheap and tolerant as possible. Once I demonstrated the synthesis, I was to hand over this project to Maxime Prost for all further exploration.

To this assigned task I added my own personal objectives. To design proper enzymatic evaluations and analyse them correctly, I had to understand, anticipate and quantify the behaviour of the unusual precipitating fluorophores by performing original modelling and analysis. I also wished to perform a

rigorous, explicit and quantitative evaluation of the probes' performance as regards all the key design criteria, both microscopic (solubility, stability, enzymatic kinetics, spacer response time) and macroscopic (signal-to-control sensitivity ratios, and signal response behaviour as regards signal onset time, form, and signal stability), which required unprecedented analysis and modelling on several levels. Lastly, comparing the probes to the prior art in the literature again required novel modelling, of the fundamental possibilities of this precipitating off-ON system as compared to both relative change probes and (still mostly hypothetical) off-ON soluble probes.

My combined objectives were therefore to synthesise, understand, evaluate, and compare to the existing literature a fluorogenic exopeptidase probe system based around novel cyclisation spacers, suitable for modular use *in vivo*, operating by an unprecedented molecular mechanism, and attaining a new regime of *ultra*sensitive enzyme detection.

4.1.2. Part II: Magnetogenic Probes

The Hasserodt group's global objective has been to develop a modular 3-component design for enzyme-activated magnetogenesis in complexes of Fe(II), towards a system of MRI probes suitable for *in vivo* molecular imaging. The proof-of-concept system must overcome the challenging spacer design requirements imposed by working with a metal complex, which places specific demands on both the spacer's functional group tolerance and its energetic incentive to fragment. Relevant prior work in the group explored classic PHBA spacers (as in phenolic glycoside **2b**, Figure 4), but these failed to eliminate once the complex was formed. My own prior work in my M2 had examined more flexible pyridinone methide spacers which failed to eliminate even before complexation (section III.4.1.2). Prof. Hasserodt's current plan involved two unprecedented substrate-spacer designs: (1) furanol glycosides as spacers with more powerful eliminative incentive (as in **44b**, in Figure II.7 on page 116); and (2) cyclic carbimidate esters bearing glycosidase substrates on the exocyclic oxygen as novel tautomeric spacers potentially facing less energetic challenge to spin switching than elimination spacers (as in **92a**, in Figure II.33 on page 144).

The initial task I was assigned was to develop suitable syntheses for a range of these substrate-spacer designs, which was considered likely to be a challenge given their almost unattested O-alkylated structures. Prof. Hasserodt provided initial synthetic lead strategies (furanol cycloisomerisation/ α-functionalisation as per section II.3.1; carbimidate S_NAr as per section I.7.2.1 as well as reviews notably including that of Fearnley[127]), and it was my responsibility to troubleshoot these strategies and targets, and/or pursue new synthetic routes as needed. The ultimate goal was to carry on the syntheses until model ligands could be obtained for complexation studies, then to examine the functional suitability of these spacer systems for use in magnetogenic probe complexes; yet this initially depends on considerations of synthetic access, then modularity and aqueous stability, before the question of spin switching can be broached.

4.1.3. List of crucial results from this PhD

The essential results from this PhD (rather than the analyses behind them) are detailed below:

Part I: Fluorogenic aminopeptidase probes

Practical Results:

- Final synthetic routes to the fluorogenic probes (Figure I.25, page 75)
- Macroscopic observables in active and control experiments *in vitro* (Figures I.29-I.30, page 81)
- Macroscopic signal-to-control sensitivity ratios (Table I.1, page 86)
- Microscopic kinetics: enzyme processing, cyclisation, and precipitation parameters from multistep fit (Figure I.36 and Table I.4, pages 93-94)
- Comparison of probe performance to the prior art (section I.5.6.2, pages 96-100)

Modelling:

- Sensitivity as a function of probe mode (Figures I.9-I.12, pages 53-56)
- Sensitivity as a function of Stokes shift (Figure I.14, page 61)
- Considerations for the use of precipitating fluorophores (section I.3.3, pages 63-67)

Part II: Towards magnetogenic glycosidase probes

Design:

- Design of the 2-furanol-based spacer (section II.2, pages 120-122)
- Design of the cyclic carbamate-based spacers (section II.6, pages 148-150)

Syntheses:

- 5-methyl-2-furanol glycosides (section II.3.1.3, page 128)
- Cycloisomerisation towards 5-(2-picolyl)-2-furanol glycosides (section II.3.2.4, pages 135-136)
- S_NAr route towards (2-pyridyl)-(5-(2-furanol glycoside))methanol (Figure II.33, page 144)
- Initial O-alkylations of cyclic carbamates (Figure II.50, page 160)
- Final, tandem O-alkylation/desilylations of cyclic carbamates (Table II.1, page 163)
- Towards ligands with O-alkylated cyclic carbamates (Figures II.51-II.53, pages 161-164)

4.1.4. Note to the reader

Selected publications which in this author's opinion provide an especially good overview / further discussion of the topics relevant to each chapter, are given at the start of each chapter's bibliography, and may make for enriching complementary reading.

Key publications

- **An excellent, short resource on *in vivo* probe design and use**: Boonacker E, Van Noorden CJF. Enzyme Cytochemical Techniques for Metabolic Mapping in Living Cells, with Special Reference to Proteolysis. *J. Histo. Cyto.* **2001**, *49*, 1473-1486 (DOI weblink: 10.1177/002215540104901201).
- **A longer, comprehensive overview of the field of Molecular Imaging**: National Research Council. *Visualizing Chemistry: The Progress and Promise of Advanced Chemical Imaging,* The National Academies Press, Washington, D.C., **2006**.

A. Bibliography of the Introduction

DOI are indicated where possible (and failing that, PubMed ID [PMID]), for fastest lookup.

[1] E. Boonacker, C. J. F. Van Noorden; *J. Histo. Cyto.* **2001**, *49*, 1473-1486 *(10.1177/002215540104901201)*.

[2] R. Weissleder, B. D. Ross, A. Rehemtulla, S. S. Gambhir; *Molecular Imaging,* People's Medical Publishing House, Shelton, Connecticut, **2010**.

[3] C. Debouck; *AIDS research and human retroviruses* **1992**, *8*, 153-164 *(PMID: 1540403)*.

[4] D. W. Nicholson, A. Ali, N. A. Thornberry, J. P. Vaillancourt, C. K. Ding, M. Gallant, Y. Gareau, P. R. Griffin, M. Labelle, Y. A. Lazebnik, N. A. Munday, S. M. Raju, M. E. Smulson, T.-T. Yamin, V. L. Yu, D. K. Miller; *Nature* **1995**, *376*, 37-43 *(10.1038/376037a0)*.

[5] W. Friedmann, J. Steffens, H. Lobeck; *Onkologie* **1984**, *7*, 337-341 *(10.1159/000215476)*.

[6] L. Stryer, J. M. Berg, J. L. Tymoczko; *Biochemistry,* Freeman & Company Limited, W H **2002**.

[7] E. Boonacker, J. Stap, A. Koehler, C. J. Van Noorden; *Acta Histochem.* **2004**, *106*, 89-96 *(10.1016/j.acthis.2004.01.002)*.

[8] P. L. Carl, P. K. Chakravarty, J. A. Katzenellenbogen; *J. Med. Chem.* **1981**, *24*, 479-480 *(10.1021/jm00137a001)*.

[9] National Research Council; *Visualizing Chemistry: The Progress and Promise of Advanced Chemical Imaging,* The National Academies Press, Washington, D.C., **2006**.

[10] J. A. Richard, Y. Meyer, V. Jolivel, M. Massonneau, R. Dumeunier, D. Vaudry, H. Vaudry, P. Y. Renard, A. Romieu; *Bioconjugate Chem.* **2008**, *19*, 1707-1718 *(10.1021/bc8001997)*.

[11] S. Izumi, Y. Urano, K. Hanaoka, T. Terai, T. Nagano; *J. Am. Chem. Soc.* **2009**, *131*, 10189-10200 *(10.1021/ja902511p)*.

[12] P. Hermann, J. Kotek, V. Kubicek, I. Lukes; *J. Chem. Soc., Dalton Trans.* **2008**, 3027-3047 *(10.1039/B719704G)*.

[13] M. Baker; *Nature* **2010**, *463*, 977-980 *(10.1038/463977a)*.

[14] *Molecular Imaging and Contrast Agent Database (MICAD)*, PubMed-NLM, Bethesda MD, **2004** (PMID: 20641179).

[15] F. Kratz, I. A. Muller, C. Ryppa, A. Warnecke; *ChemMedChem* **2008**, *3*, 20-53 *(10.1002/cmdc.200700159)*.

[16] C. A. Lipinski; *Drug Discovery Today: Technologies* **2004**, *1*, 337-341 *(10.1016/j.ddtec.2004.11.007)*.

[17] *Innovative Medicines Initiative - 4th Call for proposals, (Vol. IMI-GB-Dec-2011-09.1),* Innovative Medicines Initiative Joint Undertaking, **2011**.

[18] K. M. Huttunen, H. Raunio, J. Rautio; *Pharmacological Reviews* **2011**, *63*, 750-771 *(10.1124/pr.110.003459)*.

[19] M. Rooseboom, J. N. M. Commandeur, N. P. E. Vermeulen; *Pharmacological Reviews* **2004**, *56*, 53-102 *(10.1124/pr.56.1.3)*.

[20] V. Stella, R. Borchardt, M. Hageman, R. Oliyai, H. Maag, J. Tilley; in *Prodrugs: Challenges and Rewards (Part 1)*, Springer, New York, **2007**.

[21] R. J. Amir, M. Popkov, R. A. Lerner, C. F. Barbas, D. Shabat; *Angew. Chem., Int. Ed.* **2005**, *44*, 4378-4381 *(10.1002/anie.200500842)*.

[22] R. J. Amir, D. Shabat; *Chem. Commun.* **2004**, 1614-1615 *(10.1039/B404946B)*.

[23] Y. Meyer, J.-A. Richard, B. Delest, P. Noack, P.-Y. Renard, A. Romieu; *Org. Biomol. Chem.* **2010**, *8*, 1777-1780 *(10.1039/B926316K)*.

[24] F. M. H. de Groot, W. J. Loos, R. Koekkoek, L. W. A. van Berkom, G. F. Busscher, A. E. Seelen, C. Albrecht, P. de Bruijn, H. W. Scheeren; *J. Org. Chem.* **2001**, *66*, 8815-8830 *(10.1021/jo0158884)*.

[25] D. Shabat, R. J. Amir, A. Gopin, N. Pessah, M. Shamis; *Chem.–Eur. J* **2004**, *10*, 2626-2634 *(10.1002/chem.200305715)*.

[26] W. S. Saari, J. E. Schwering, P. A. Lyle, S. J. Smith, E. L. Engelhardt; *J. Med. Chem.* **1990**, *33*, 97-101 *(10.1021/jm00163a016)*.

[27] V. M. Vrudhula, J. F. MacMaster, Z. Li, D. E. Kerr, P. D. Senter; *Bioorg. Med. Chem. Lett.* **2002**, *12*, 3591-3594 *(10.1016/S0960-894X(02)00784-9)*.

[28] O. Thorn-Seshold, M. Vargas-Sanchez, S. McKeon, J. Hasserodt; *Chem. Commun.* **2012**, *48*, 6253-6255 *(10.1039/C2CC32227G)*.

[29] X.-b. Zhang, M. Waibel, J. Hasserodt; *Chem.–Eur. J* **2010**, *16*, 792-795 *(10.1002/chem.200902412)*.

[30] V. Stavila, Y. Stortz, C. Franc, D. Pitrat, P. Maurin, J. Hasserodt; *Eur. J. Inorg. Chem.* **2008**, 3943-3947 *(10.1002/ejic.200800419)*.

[31] F. Touti, P. Maurin, J. Hasserodt; *Angew. Chem., Int. Ed.* **2013**, *52*, 4654–4658 *(10.1002/anie.201208848)*.

[32] J. V. Frangioni; *Curr. Opin. Chem. Biol.* **2003**, *7*, 626-634 *(10.1016/j.cbpa.2003.08.007)*.

[33] C. J. F. Van Noorden; *J. Histo. Cyto.* **2010**, *58*, 481-497 *(10.1369/jhc.2010.955518)*.

[34] Y. Meyer, J.-A. Richard, M. Massonneau, P.-Y. Renard, A. Romieu; *Org. Lett.* **2008**, *10*, 1517-1520 *(10.1021/ol800198f)*.

[35] Y. Yang, P. Babiak, J.-L. Reymond; *Org. Biomol. Chem.* **2006**, *4*, 1746-1754 *(10.1039/B601151A)*.

[36] V. J. Stella; in *Prodrugs*, (*Vol. 5*), (Eds.: V. J. Stella, R. T. Borchardt, M. J. Hageman, R. Oliyai, H. Maag and J. W. Tilley), Springer, New York, **2007**, pp. 3-33.

[37] A. L. James, J. D. Perry, S. P. Stanforth; *J. Heterocycl. Chem.* **2006**, *43*, 515-517 *(10.1002/jhet.5570430241)*.

[38] T. Jenkins, A. Kolesnikov; *Controlled Release of Phenolic Opioids (WO2007140272)*, **2007**.

[39] J. D. Tibodeau, L. M. Benson, C. R. Isham, W. G. Owen, K. C. Bible; *Antioxidants & redox signaling* **2009**, *11*, 1097-1106 *(10.1089/ars.2008.2318)*.

[40] D. H. Kwan, H.-M. Chen, K. Ratananikom, S. M. Hancock, Y. Watanabe, P. T. Kongsaeree, A. L. Samuels, S. G. Withers; *Angew. Chem., Int. Ed.* **2011**, *50*, 300-303 *(10.1002/anie.201005705)*.

[41] M. N. Levine, R. T. Raines; *Anal. Biochem.* **2011**, *418*, 247-252 *(10.1016/j.ab.2011.07.021)*.

[42] J. D. Sellars, M. Landrum, A. Congreve, D. P. Dixon, J. A. Mosely, A. Beeby, R. Edwards, P. G. Steel; *Org. Biomol. Chem.* **2010**, *8*, 1610-1618 *(10.1039/B920443A)*.

[43] M. Grinda, J. Clarhaut, B. Renoux, I. Tranoy-Opalinski, S. Papot; *MedChemComm* **2012**, *3*, 68-70 *(10.1039/C1MD00193K)*.

[44] S. S. Chandran, K. A. Dickson, R. T. Raines; *J. Am. Chem. Soc.* **2005**, *127*, 1652-1653 *(10.1021/ja043736v)*.

[45] A. V. Zaytsev, R. J. Anderson, A. Bedernjak, P. W. Groundwater, Y. Huang, J. D. Perry, S. Orenga, C. Roger-Dalbert, A. James; *Org. Biomol. Chem.* **2008**, *6*, 682-692 *(10.1039/B716978G)*.

[46] K. Okuda, Y. Okabe, T. Kadonosono, T. Ueno, B. G. M. Youssif, S. Kizaka-Kondoh, H. Nagasawa; *Bioconjugate Chem.* **2012**, *23*, 324-329 *(10.1021/bc2004704)*.

[47] E. M. Zeman, J. M. Brown, M. J. Lemmon, V. K. Hirst, W. W. Lee; *Int. J. Radiat. Oncol. Biol. Phys.* **1986**, *12*, 1239-1242 *(PMID: 3744945)*.

[48] E. O. Aboagye, A. D. Lewis, A. Johnson, P. Workman, M. Tracy, I. M. Huxham; *Br. J. Cancer* **1995**, *72*, 312-318 *(PMID: 7640211)*.

[49] A. L. James, J. D. Perry, C. Jay, D. Monget, J. W. Rasburn, F. K. Gould; *Lett. Appl. Microbiol.* **2001**, *33*, 403-408 *(PMID: 11737621)*.

[50] S. Takahashi, W. Piao, Y. Matsumura, T. Komatsu, T. Ueno, T. Terai, T. Kamachi, M. Kohno, T. Nagano, K. Hanaoka; *J. Am. Chem. Soc.* **2012**, *134*, 19588-19591 *(10.1021/ja310049d)*.

[51] K. Kiyose, K. Hanaoka, D. Oushiki, T. Nakamura, M. Kajimura, M. Suematsu, H. Nishimatsu, T. Yamane, T. Terai, Y. Hirata, T. Nagano; *J. Am. Chem. Soc.* **2010**, *132*, 15846-15848 *(10.1021/ja105937q)*.

[52] M. Kamiya, D. Asanuma, E. Kuranaga, A. Takeishi, M. Sakabe, M. Miura, T. Nagano, Y. Urano; *J. Am. Chem. Soc.* **2011**, *133*, 12960-12963 *(10.1021/ja204781t)*.

[53] G. M. Dubowchik, R. A. Firestone, L. Padilla, D. Willner, S. J. Hofstead, K. Mosure, J. O. Knipe, S. J. Lasch, P. A. Trail; *Bioconjugate Chem.* **2002**, *13*, 855-869 *(10.1021/bc025536j)*.

[54] A. Razgulin, N. Ma, J. Rao; *Chem. Soc. Rev.* **2011**, *40*, 4186-4216 *(10.1039/c1cs15035a)*.

[55] S. M. Hacker, N. Hardt, A. Buntru, D. Pagliarini, M. Mockel, T. U. Mayer, M. Scheffner, C. R. Hauck, A. Marx; *Chem. Sci.* **2013**, *4*, 1588-1596 *(10.1039/C3SC21916J)*.

[56] S. K. Grant, J. G. Sklar, R. T. Cummings; *J. Biomol. Screen.* **2002**, *7*, 531-540 *(10.1177/1087057102238627)*.

[57] M. Thomas, F. Rivault, I. Tranoy-Opalinski, J. Roche, J.-P. Gesson, S. Papot; *Bioorg. Med. Chem. Lett.* **2007**, *17*, 983-986 *(10.1016/j.bmcl.2006.11.042)*.

[58] M. Shamis, H. N. Lode, D. Shabat; *J. Am. Chem. Soc.* **2004**, *126*, 1726-1731 *(10.1021/ja039052p)*.

[59] Z. Zhang, K. Tanabe, H. Hatta, S.-i. Nishimoto; *Org. Biomol. Chem.* **2005**, *3*, 1905-1910 *(10.1039/B502813B)*.

[60] F. M. H. de Groot, L. W. A. van Berkom, H. W. Scheeren; *J. Med. Chem.* **2000**, *43*, 3093-3102 *(10.1021/jm0009078)*.

[61] F. M. H. de Groot, C. Albrecht, R. Koekkoek, P. H. Beusker, H. W. Scheeren; *Angew. Chem., Int. Ed.* **2003**, *42*, 4490-4494 *(10.1002/anie.200351942)*.

[62] Z. Huang, E. Terpetschnig, W. You, R. P. Haugland; *Anal. Biochem.* **1992**, *207*, 32-39 *(10.1016/0003-2697(92)90495-S)*.

[63] M. Kwolek-Mirek, S. Bednarska, R. Zadrag-Tecza, G. Bartosz; *Cell. Biol. Int.* **2011**, *35*, 1111-1119 *(10.1042/cbi20100763)*.

[64] B. W. Lee, G. L. Johnson, S. A. Hed, Z. Darzynkiewicz, J. W. Talhouk, S. Mehrotra; *BioTechniques* **2003**, *35*, 1080-1085 *(PMID:14628683)*.

[65] K. Kominami, T. Nagai, T. Sawasaki, Y. Tsujimura, K. Yashima, Y. Sunaga, M. Tsuchimochi, J. Nishimura, K. Chiba, J. Nakabayashi, K. Koyamada, Y. Endo, H. Yokota, A. Miyawaki, N. Manabe, K. Sakamaki; *PLoS One* **2012**, *7*, e50218 *(10.1371/journal.pone.0050218)*.

[66] S. Jin, E. Ellis, J. V. Veetil, H. Yao, K. Ye; *Biotechnol. Prog.* **2011**, *27*, 1107-1114 *(10.1002/btpr.628)*.

[67] E. Kohen, C. Kohen, J. G. Hirschberg, R. Santus, G. Grabowski, W. Mangel, S. Gatt, J. Prince; *Cell Biochem. Funct.* **1993**, *11*, 167-177 *(10.1002/cbf.290110304)*.

[68] D. Bini, M. Gregori, U. Cosentino, G. Moro, A. Canales, A. Capitoli, J. Jiménez-Barbero, L. Cipolla; *Carbohydr. Res.* **2012**, *354*, 21-31 *(10.1016/j.carres.2012.03.002)*.

[69] R. A. Moats, S. E. Fraser, T. J. Meade; *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 726-728 *(10.1002/anie.199707261)*.

[70] L. D. Lavis; *ACS Chem. Biol.* **2008**, *3*, 203-206 *(10.1021/cb800065s)*.

[71] Invitrogen; *Product Information Sheet, ELF-97 Endogenous Phosphatase Detection Kit (E6601)* **2004**.

[72] L. D. Lavis, T. Y. Chao, R. T. Raines; *Chem. Sci.* **2011**, *2*, 521-530 *(10.1039/c0sc00466a)*.

[73] D. Shallom, V. Belakhov, D. Solomon, G. Shoham, T. Baasov, Y. Shoham; *J. Biol. Chem.* **2002**, *277*, 43667-43673 *(10.1074/jbc.M208285200)*.

[74] Z. Huang, Q. Wang, H. D. Ly, A. Gorvindarajan, J. Scheigetz, R. Zamboni, S. Desmarais, C. Ramachandran; *J. Biomol. Screen.* **1999**, *4*, 327-334 *(10.1177/108705719900400608)*.

[75] G. Leary, D. Sawtell; *Holzforschung* **1984**, *38*, 53-54 *(10.1515/hfsg.1984.38.1.53)*.

[76] L. Tian, Y. Yang, L. M. Wysocki, A. C. Arnold, A. Hu, B. Ravichandran, S. M. Sternson, L. L. Looger, L. D. Lavis; *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 4756-4761 *(10.1073/pnas.1111943109)*.

[77] N.-H. Ho, R. Weissleder, C.-H. Tung; *Bioorg. Med. Chem. Lett.* **2006**, *16*, 2599-2602 *(10.1016/j.bmcl.2006.02.045)*.

[78] V. B. Paragas, J. A. Kramer, C. Fox, R. P. Haugland, V. L. Singer; *Journal of Microscopy* **2002**, *206*, 106-119 *(10.1046/j.1365-2818.2002.01017.x)*.

[79] A. Y. Louie, M. M. Huber, E. T. Ahrens, U. Rothbacher, R. Moats, R. E. Jacobs, S. E. Fraser, T. J. Meade; *Nat. Biotechnol.* **2000**, *18*, 321-325 *(10.1038/73780)*.

[80] R. B. Greenwald, Y. H. Choe, C. D. Conover, K. Shum, D. Wu, M. Royzen; *J. Med. Chem.* **2000**, *43*, 475-487 *(10.1021/jm990498j)*.

[81] E. S. Furfine, C. T. Baker, M. R. Hale, D. J. Reynolds, J. A. Salisbury, A. D. Searle, S. D. Studenberg, D. Todd, R. D. Tung, A. Spaltenstein; *Antimicrobial agents and chemotherapy* **2004**, *48*, 791-798 *(10.1128/AAC.48.3.791-798.2004)*.

[82] H. Yuan, N. Li, Y. Lai; *Drug. Metab. Dispos.* **2009**, *37*, 1443-1447 *(10.1124/dmd.108.026245)*.

[83] D. J. Weiss, D. Liggitt, J. G. Clark; *Histochem. J.* **1999**, *31*, 231-236 *(PMID: 10447064)*.

[84] M. Poreba, S. McGowan, T. S. Skinner-Adams, K. R. Trenholme, D. L. Gardiner, J. C. Whisstock, J. To, G. S. Salvesen, J. P. Dalton, M. Drag; *PLoS One* **2012**, *7*, e31938 *(10.1371/journal.pone.0031938)*.

[85] A. Fernandes, A. Viterisi, F. Coutrot, S. Potok, D. A. Leigh, V. Aucagne, S. Papot; *Angew. Chem., Int. Ed.* **2009**, *48*, 6443-6447 *(10.1002/anie.200903215)*.

[86] G. Wiederschain, S. Raghavan, E. Kolodny; *Clin. Chim. Acta* **1992**, *205*, 87-96 *(PMID: 1521344)*.

[87] G. A. Grabowski, S. Gatt, M. Horowitz; *Critical reviews in biochemistry and molecular biology* **1990**, *25*, 385-414 *(10.3109/10409239009090616)*.

[88] K. Y. Choi, M. Swierczewska, S. Lee, X. Y. Chen; *Theranostics* **2012**, *2*, 156-178 *(10.7150/thno.4068)*.

[89] E. K. Y. Chen, R. A. McBride, E. R. Gillies; *Macromolecules* **2012**, *45*, 7364-7374 *(10.1021/ma301667c)*.

[90] S. K. Kumar, S. A. Williams, J. T. Isaacs, S. R. Denmeade, S. R. Khan; *Bioorg. Med. Chem.* **2007**, *15*, 4973-4984 *(10.1016/j.bmc.2007.04.029)*.

[91] N. Pessah, M. Reznik, M. Shamis, F. Yantiri, H. Xin, K. Bowdish, N. Shomron, G. Ast, D. Shabat; *Bioorg. Med. Chem.* **2004**, *12*, 1859-1866 *(10.1016/j.bmc.2004.01.039)*.

[92] M. Waibel, X.-b. Zhang, J. Hasserodt; *Synthesis* **2009**, 318-324 *(10.1055/s-0028-1083296)*.

[93] M. A. DeWit, E. R. Gillies; *Org. Biomol. Chem.* **2011**, *9*, 1846-1854 *(10.1039/C0OB00890G)*.

[94] P. Wipf, W. Li, C. M. Adeyeye, J. M. Rusnak, J. S. Lazo; *Bioorg. Med. Chem.* **1996**, *4*, 1585-1596 *(PMID: 8931928)*.

[95] M. A. Dewit, A. Beaton, E. R. Gillies; *J. Polym. Sci., Part A: Polym. Chem.* **2010**, *48*, 3977-3985 *(10.1002/pola.24180)*.

[96] H. P. Ng, D. P. C. McGee, G. Wu, Z. Li, O. L. Saunders, V. Martichonok, G. T. Yarranton; *Cytotoxins, Prodrugs, Linkers and Stabilizers useful therefor (US7129261)*, **2002**.

[97] M. H. Lee, J. H. Han, J.-H. Lee, H. G. Choi, C. Kang, J. S. Kim; *J. Am. Chem. Soc.* **2012**, *134*, 17314-17319 *(10.1021/ja308446y)*.

[98] S. Milstien, L. A. Cohen; *J. Am. Chem. Soc.* **1972**, *94*, 9158-9165 *(10.1021/ja00781a029)*.

[99] R. T. Borchardt, L. A. Cohen; *J. Am. Chem. Soc.* **1972**, *94*, 9166-9174 *(10.1021/ja00781a030)*.

[100] M. N. Levine, R. T. Raines; *Chem. Sci.* **2012**, *3*, 2412-2420 *(10.1039/C2SC20536J)*.

[101] K. L. Amsberry, A. E. Gerstenberger, R. T. Borchardt; *Pharm Res* **1991**, *8*, 455-461 *(PMID: 1871039)*.

[102] K. Achilles; *Arch. Pharm. (Weinheim, Ger.)* **2001**, *334*, 209-215 *(10.1002/1521-4184(200106)334:6<209::aid-ardp209>3.0.co;2-y)*.

[103] B. Wang, H. Zhang, W. Wang; *Bioorg. Med. Chem. Lett.* **1996**, *6*, 945-950 *(10.1016/0960-894X(96)00147-3)*.

[104] B. Wang, W. Wang, H. Zhang, D. Shan, T. D. Smith; *Bioorg. Med. Chem. Lett.* **1996**, *6*, 2823-2826 *(10.1016/S0960-894X(96)00526-4)*.

[105] I. Christenson; *Acta Chem. Scand.* **1964**, *18*, 904-922 *(10.3891/acta.chem.scand.18-0904)*.

[106] S. Halazy, V. Berges, A. Ehrhard, C. Danzin; *Bioorg. Chem.* **1990**, *18*, 330-344 *(10.1016/0045-2068(90)90007-R)*.

[107] M. Wakselman; *N. J. Chimie* **1983**, *7*, 439-447.

[108] P. D. Senter, W. E. Pearce, R. S. Greenfield; *J. Org. Chem.* **1990**, *55*, 2975-2978 *(10.1021/jo00296a082)*.

[109] E. W. P. Damen, T. J. Nevalainen, T. J. M. van den Bergh, F. M. H. de Groot, H. W. Scheeren; *Bioorg. Med. Chem.* **2002**, *10*, 71-77 *(10.1016/S0968-0896(01)00235-8)*.

[110] S. Papot, D. Combaud, J.-P. Gesson; *Bioorg. Med. Chem. Lett.* **1998**, *8*, 2545-2548 *(10.1016/S0960-894X(98)00454-5)*.

[111] R. Sicart, M.-P. Collin, J.-L. Reymond; *Biotechnol. J.* **2007**, *2*, 221-231 *(10.1002/biot.200600181)*.

[112] M. Azoulay, F. Chalard, J.-P. Gesson, J.-C. Florent, C. Monneret; *Carbohydr. Res.* **2001**, *332*, 151-156 *(10.1016/S0008-6215(01)00083-0)*.

[113] O. Thorn-Seshold; *Towards new molecular candidates for an auto-immolative coordinating arm on macrocyclic ligands* (M2 Thesis), Ecole Normale Supérieure de Lyon, **2009**.

[114] R. Perry-Feigenbaum, P. S. Baran, D. Shabat; *Org. Biomol. Chem.* **2009**, *7*, 4825-4828 *(10.1039/B915265B)*.

[115] S. E. Rokita, J. Yang, P. Pande, W. A. Greenberg; *J. Org. Chem.* **1997**, *62*, 3010-3012 *(10.1021/jo9700336)*.

[116] M. Freccero; *Mini-Rev. Org. Chem.* **2004**, *1*, 403-415 *(10.2174/1570193043403091)*.

[117] E. E. Weinert, R. Dondi, S. Colloredo-Melz, K. N. Frankenfield, C. H. Mitchell, M. Freccero, S. E. Rokita; *J. Am. Chem. Soc.* **2006**, *128*, 11940-11947 *(10.1021/ja062948k)*.

[118] M. Grinda, J. Clarhaut, I. Tranoy-Opalinski, B. Renoux, A. Monvoisin, L. Cronier, S. Papot; *ChemMedChem* **2012**, *6*, 2137-2141 *(10.1002/cmdc.201100355)*.

[119] J. C. Briggs, A. H. Haines, R. J. K. Taylor; *J. Chem. Soc., Chem. Commun.* **1992**, 1039-1041 *(10.1039/C39920001039)*.

[120] R. Weinstain, E. Segal, R. Satchi-Fainaro, D. Shabat; *Chem. Commun.* **2010**, *46*, 553-555 *(10.1039/B919329D)*.

[121] J. Veljković, L. Uzelac, K. Molčanov, K. Mlinarić-Majerski, M. Kralj, P. Wan, N. Basarić; *J. Org. Chem.* **2012**, *77*, 4596-4610 *(10.1021/jo3002479)*.

[122] *This author hopes to pursue a different strategy altogether during his postdoc work.*

[123] T. J. Monks, D. C. Jones; *Curr. Drug Metab.* **2002**, *3*, 425-438 *(PMID: 12093358)*.

[124] V. Stavila, M. Allali, L. Canaple, Y. Stortz, C. Franc, P. Maurin, O. Beuf, O. Dufay, J. Samarut, M. Janier, J. Hasserodt; *New J. Chem.* **2008**, *32*, 428-435 *(10.1039/b715254j)*.

[125] J. Hasserodt; *Contrast Agents for Magnetic Resonance Imaging (WO2005094903)*, **2005**.

[126] *This author's fond hope is to develop an intelligently-targeted off-ON molecular system for selective imaging and destruction of tumour cells, regardless of tumour type: a project which will hopefully begin during the upcoming year (2013-2014).*

[127] S. P. Fearnley; *Curr. Org. Chem.* **2004**, *8*, 1289-1337 *(10.2174/1385272043369971)*.

Part I

Fluorogenic probes for aminopeptidases

Table of Contents: Part I

I. FLUORESCENCE DETECTION OF EXOPEPTIDASE ACTIVITY

1. State of the Art: Imaging with Fluorogenic Probes

1.1. Fluorescence imaging

1.1.1. General parameters and requirements

Optical – and particularly fluorescence – molecular imaging[1] is an attractive modality for chemical designers, since it can arguably offer the highest sensitivity of the biomolecularly-switchable imaging modalities. The sensitivity of an imaging modality is essentially limited by the signal background with which it must compete; that sensitivity of fluorescence imaging (~nM-μM imaging agent required) lies roughly halfway between that of the two other truly *molecular* imaging methods in most common use, PET (~pM) and MRI (~mM)[2]. Note that fluorescence imaging is however much cheaper, and probably intrinsically more versatile, than both of these methods, and it is already a mainstay technique in high-throughput screening. Fluorescence imaging, like MRI, also distinguishes itself from PET and other imaging methods by avoiding the use of intrinsically harmful radiation. Fluorescence molecular imaging therefore represents a thoroughly promising avenue of research for low-cost, high-throughput, routine biomedical procedures from *in vitro* assays to *in vivo* diagnostics.[3]

By implementing truly off-ON molecular functionality with an *appropriate* fluorophore, the signal background expected from biological tissue may be minimal, therefore the sensitivity may be maximised. The chief source of unwanted *in vivo* fluorescence background is tissue autofluorescence; this is strongly dependent on the excitation wavelength used and on the tissue type. Imaging with a large excitation/emission wavelength split is critical for minimising the autofluorescence background contribution.[1] For *in cellulo* and especially *in vitro* tests, this background may be much reduced regardless of the excitation and emission wavelengths, or even near-zero: but the biological *in vivo* situation *must* be remembered when designing probes for *in vivo* use: apparently promising probes *in vitro* can easily prove useless already *in cellulo* if they fail to take these requirements into account.[4,5]

The absolute wavelengths of the excitation/emission light are also critical for *in vivo* sensitivity. The wavelength-dependent light absorption profile of biological tissues has led to the common notion of the biological transparency window between ~650 – 900 nm, since light in this region is only weakly absorbed *in vivo*. Such long-wavelength excitation and emission are therefore further requirements for good *in vivo* sensitivity for several reasons. As regards sensitivity, longer wavelengths minimise the autofluorescence background, and also the depth penetration of both excitation and emission light, leading to higher intrinsic signal-to-background. As regards the desired *in vivo* applications of probes,

there are many arguments for imaging in the biological transparency window, and avoiding short-wavelength excitation. There is a limit on the excitation power per unit absorbing volume which can be tolerated before essentially cooking the tissue under study;[1] therefore if excitation/emission is performed with strongly-absorbed wavelengths, eg. much below 600 nm, strong tissue absorption limits the excitation power that can be used, so experiments are restricted to a depth of around a millimetre, regardless of the fluorophore brightness/concentration that can be systemically established. Photochemical damage is another consideration which is wavelength-dependent and favours longer wavelengths. For example, short-wavelength excitation (<<380 nm) cannot be freely used *in cellulo*, or even *in vitro* with very short wavelengths (<330 nm), as it results in extremely rapid photochemical damage to and degradation of cells and enzymes. Fluorophores themselves commonly accelerate this damage by acting as photosensitisers for eg. the generation of reactive oxygen species. Also, the time-integral total excitation energy applied to a cell is limited by photochemical damage that can be generated (eg. 10 $J \cdot cm^{-2}$ over the entire cell lifetime, at ~400 nm)[1]. By contrast, imaging in the near-infrared (NIR; ~700-900 nm) provides optimal depth penetration with minimal absorption risks, and even whole-animal fluorescence imaging is relatively easy under these conditions.[2]

Multiphoton fluorescence is an extremely attractive method for establishing the theoretical minimal signal background in a practical setting, and thus obtaining the greatest possible sensitivity, while avoiding short-wavelength excitation. However, current multiphoton organic fluorophores require enormous irradiance intensities to generate favourable absorption probabilities, thus are also effectively limited when applied to *in vivo* studies where the subject cannot be cooked. Upconverting quantum dot strategies require far lower illumination intensities for multiphoton processes, but they do not lend themselves as easily to molecular imaging strategies because their signal cannot yet be reliably be modulated by enzymatic activity; they also have problems of biodistribution and toxicity. This thesis will focus on the design of classical, single-photon, molecular fluorogenic probes.[5-10]

1.1.2. Practical requirements for fluorogenic probes in vivo

Fluorogenic probes, ie. probes which generate fluorescence after biomolecular activation, offer the greatest sensitivity attainable in fluorescence molecular imaging (see too section I.3.1). Powerful fluorogenicity is a primordial requirement for *in vivo* fluorescence imaging: due to the problems of small enzymatic activity and dynamically inhomogenous probe concentrations, a large and unambiguous change in fluorescence properties upon enzymatic activation is a necessity not a virtue; without it, the sensitivity obtainable is - without exception - minimal. Therefore, the following discussion assumes the design of truly off-ON fluorogenic probes as a *sine qua non*; note however that true off-ON designs are actually very rare (see section I.1.3).

Excellent reviews of the practical requirements for fluorogenic probes are available[1,5,10,11]; so a brief selection of the criteria which will be most relevant to this PhD work is given below without further explanation; see too the general requirements for *in vivo* probes outlined in the Introduction.

Probe requirements

- "High chemical stability of fluorogenic probes is vital for certain experiments, as spontaneous hydrolysis compromises the ability to assess enzymatic activity, raises background fluorescence, and diminishes membrane permeance."[4,12] Due to the small enzyme activity (thus small signal) that can be expected *in vivo*, exceptional probe stability over the experimental lifetime is required, or else the observed signal will be ambiguous and therefore of limited or zero value (see section I.3.1);

- The proagent must be selectively processed by its target enzyme, and processing must be at a rate comparable to and competitive with that of the natural substrate, "so that increase in absorbance or fluorescence in time is a proper reflection of physiological substrate fluxes;"[1] only then can strong catalytic signal amplification by the target enzyme be assured even at its low native activity;

- The proagent must be freely membrane-permeable and water-soluble;

- The probe and its byproducts (and metabolites) must be nontoxic for cells and tissues, so as not to poison the enzyme, cell or subject;

- A good probe will follow a cheap, modular design which is assembled at a late stage.[13]

Fluorophore characteristics

- A minimum requirement for reliable *in vivo* work is excitation/emission in the red/NIR; the closer that the fluorophore fulfils this requirement, the better will be the translation of its *in vitro* results to *in cellulo* and then *in vivo* settings;

- Then, to obtain high sensitivity *in vivo* absolutely requires fluorophores with Stokes shift $\Delta\lambda_P$ significantly larger than that of the tissue autofluorescence (usually $\Delta\lambda_A$~40 nm) caused by the excitation wavelength.[14] Fluorescence studies exciting and collecting with a small wavelength split will generate much unavoidable autofluorescence background and so suffer low sensitivity; if excitation/emission can instead be performed with a large wavelength split, very much larger signal to background levels may be obtained. For example, in section I.3.2 I derive a sensitivity model which shows that *in vivo* detection sensitivity depends exponentially on $\Delta\lambda_P$ (eg. a fluorophore with $\Delta\lambda_P$~30-50 nm gives *in vivo* sensitivity which is a factor of 300 lower than a fluorophore with $\Delta\lambda_P$=150 nm);

- Then, fluorophores must not diffuse away from the site of enzymatic activity, or else there is no correlation between spatially-observable signal and enzymatic activity (signal bleed[15]): ie, they must maintain a locally high concentration.[5,11] Fluorophores which precipitate upon release can mark the real site of enzymatic activity with with subcellular resolution, and do not suffer either diffusion, rapid excretion or bioconcentration in a different cell compartment/tissue[1,16], so they are a prime way to achieve good signal localisation;[17,18]

- Then, the fluorophore must not photobleach, as this vastly reduces the potential signal over time (and probably results in photosensitised toxicity). Photobleaching also prevents reliable time-resolved assays of enzymatic activity, which otherwise can deliver the most useful information (section I.5.5);

- Only then, does the fluorophore's signal brightness (quantum yield Φ and absorption coefficient ε) enter into consideration. The higher the brightness, the higher will be the collected signal, but this is a linear dependency, whereas the other factors listed above all create exponential or power-law effects on sensitivity, especially *in cellulo* and *in vivo*. I therefore consider the previous factors far more important in a practical, *in vivo* sense, even if results in the cuvette with bright fluorophores that otherwise fail the previous criteria (eg. many fluorescein probes) may seem initially promising.[19]

1.2. Fluorogenic probes for peptidase activity

This PhD was concerned with the synthesis of fluorogenic peptidase probes. Thousands of molecules have been reported as 'fluorogenic peptidase probes,' but there is still ample room for innovation, since comparatively few of these designs are actually appropriate for monitoring the activity of their claimed biological target in a relevant *in cellulo* or *in vivo* setting. As has been stated baldly for ATPase probes, "[most] probes are conditioned by the synthetic availability of their modifications rather than adjustment to the biological system of interest."[20] Likewise, peptidase probes usually rely closely on one of three synthetically well-established architectural approaches, each of which brings its own, substantial, unresolved challenges. These approaches and their intrinsic architectural challenges are briefly introduced below; chemical challenges are addressed in section I.1.3.

Exopeptidase targeting is usually designed by attaching the substrate specificity sequence of a targeted exopeptidase to the amine terminus of an amine probe (2-component design), or to a probe via an N-terminal enzyme-recognised spacer/adapter (3-component design). However, the amine terminus cannot deviate too strongly from the natural substrate: a primary peptide bond is an absolute requirement for enzymatic processing, and large differences in enzyme processing kinetics are seen even between relatively unhindered amines, eg. an order of magnitude slower processing of a PABA anilide than a primary aliphatic amide.[21] These differences in rate are not usually critical *in vitro* as more enzyme can be added to overcome bad kinetics, or the reaction can be run longer if the probe is stable. However, once taken to *in vivo* work, badly-processed probes may generate no signal whatsoever since with small enzymatic activity, and especially competition from native substrates, they are simply processed too slowly to generate the sufficient concentration required for imaging.[1]

Two-component exopeptidase probes for permissive aminopeptidases can be easily and cheaply synthesised, by directly attaching the peptide substrate to the fluorophore.[22] Examples include *p*-nitroanilide **16**, a standard substrate for the present work's leucine aminopeptidase model enzyme, and AMC probe **5** previously detailed. The 2-component design only allows for the use of amine fluorophores/drugs, so is not fully modular; and because most fluorescence-modulatable amine fluorophores are anilides, they may also suffer from critical enzyme recognition problems *in vivo*.

Three-component exopeptidase probes based on a substrate-*spacer*-fluorophore design can therefore present two important advantages: (1) enabling the use of non-amine-fluorophores/chromophores, and

eventually, drugs; (2) use of better-recognised amine terminal linkers for faster processing; note that even simple elongation of the substrate from a drug, even without changing the post-scissile steric load, can also improve recognition[23]. While two-component probes can claim to represent a general solution for releasing aniline fluorophores, at least *in vitro*, the 3-component approach is the only solution for applying phenols in peptidase-triggered probes. In an ideal case, successful designs for phenols which fulfil the above requirements (section I.1.1.2) could also be applied to general alcohols, as well as carbamates[24] of non-primary- or sterically-hindered amines[25]. This would complete the toolbox for peptidase-triggered release of all types of amines and alcohols: the two most common functionality-critical drug/probe functional groups.

The amine-terminal spacer most often used for alcohol and carbamic acid release is PABA (**11**), and 1,2-diamines are also widely used for phenol and alcohol release (cf. **10a**, **10b**). Both spacer systems, as well as the other common spacers for peptidase-triggered phenol release, have non-ideal performance features which are adressed in section I.1.4. *Multi-component* probes, eg. stacked-spacer (**18**), multiply-gated or multiple-delivery (**9**) dendritic systems can be made by stacking multiple spacers from the three-component toolbox, so are not treated separately.

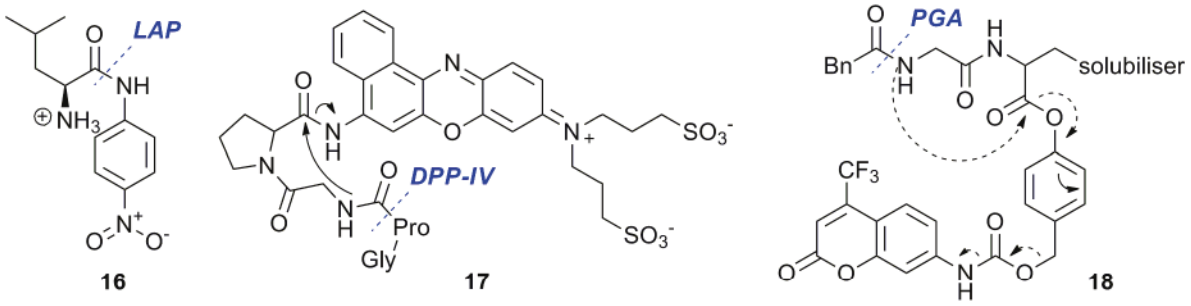


Figure I.1 – Two-component peptidase probes 16 and 17, and stacked-linker design 18

Endopeptidase targeting is usually also achieved with the peptidase's natural specificity sequence. However, most endopeptidases cannot tolerate variations in the immediate post-scissile portion of the specificity sequence, therefore the pre-scissile portion cannot be attached to a probe/linker's amine terminus as is possible for exopeptidase probes. Since the post-scissile peptide sequence is not generally functional, no auto-immolation or functional response can usually be expected. Therefore two-component designs (sequence-fluorophore) are generally not possible. Some endopeptidases do however leave a post-scissile peptide which can degrade to release an active agent, and can therefore be targeted with two-component designs, eg. that of Weissleder (**17**; Figure I.1)[26]; yet not only does this limit the scope of target endopeptidases, but post-triggering response kinetics may be dismal.

Resonant energy transfer probes such as Förster Resonance Energy Transfer (FRET) probes are therefore the principle design for endopeptidase targeting, as they are suitable for use with the widest possible range of endopeptidases, including important members such as the caspases and HIV-1 protease.[27-29] The substrate specificity sequence is attached between a first fluorophore and either a quencher moiety or a second fluorophore, so that spatially-favoured resonant energy transfer either

dampens or bathochromically shifts the first fluorophore's emission. Peptidolysis results in the separation of the two components, and therefore in an increase of fluorescence from the first fluorophore since the FRET effect is strongly distance-dependent (eg. **19**)[27]. Steric-spacer peptide sequences may commonly be inserted between the fluorophores and the scissile sequence to reduce steric hindrance.[30]

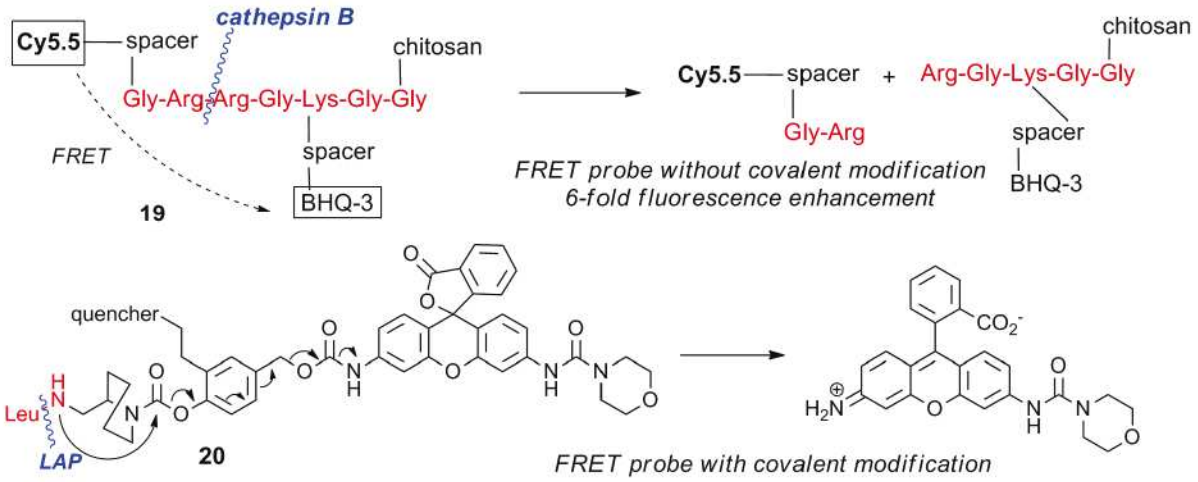


Figure I.2 - FRET probes usually function as relative change probes (19), though if the fluorophores are also covalently modified, substantial off-ON character may be seen (design 20).

FRET-based fluorescence probes have the potential to act in the mechanistically-quenched, off-ON mode *if* signal quenching in the proagent is quantitative[28]: but in reality, signal quenching in the proagent form is almost[31] never total. Considering the absolute signal from FRET probes, *in vitro* signal enhancements reported are typically only from +100%[32] to +500%[33,34] though higher enhancements are possible; yet even *in cellulo* these figures may already drop substantially (eg. +10%[35] to +35%[36]). Fluorophores for longer-wavelength *in vivo* work may also be harder to quench reliably by FRET. Therefore *in vivo* probes based only on the FRET effect[37] are, in general, not considered here as off-ON probes but rather as relative change probes[33] (see too the discussion of concentration effects in the introduction and section I.1.3.3).

FRET-based probes may of course combine a FRET-like quenching of the proagent with a covalent modification of the fluorophore during enzymatic activation, eg. by use of a suitable smart spacer (eg. hypothetical structure **20**, Figure I.2). If the change of characteristics is large enough, the probe may be essentially off-ON, however such designs will generally only be suitable for exopeptidases; as the cost, bulk and synthetic challenge of suitable FRET probes is well beyond that of most two- and three-component designs, I consider FRET an uncompetitive solution for exopeptidase targeting. Even for endopeptidase targeting, FRET is not really an ideal solution, but more of a stopgap until a more practical, higher-sensitivity approach for endopeptidase targeting is developed (see Future Directions).

1.3. Phenolic fluorophores for fluorogenic probes

1.3.1. General characteristics of standard prior art fluorophores

Most current fluorescent exopeptidase probes are based on releasing soluble fluorophores, which are usually heteroatom-functionalised aromatic or heterocyclic systems, usually anilines and phenols.[28]

Anilines, including 7-aminomethylcoumarin (AMC), disperse orange 11, and rhodamines, are rather easily, *and stably*, used as anilides in prior art 2-component fluorogenic probes[38], and where necessary, 3-component designs[25]. *Phenols* are the other major class of interest, both of fluorophores for fluorogenic probes (eg. fluoresceins, 7-hydroxycoumarins and resorufins), and also of drugs for prodrug strategies (alcohols/phenols). This PhD work was motivated by the consideration that essentially, **no designs for peptidase-triggered phenol-releasing probes or prodrugs had yet been published which would prove satisfactory general solutions for *in vivo* molecular imaging and molecular medicine**. This PhD work attempts to address this important problem with the prior art. Some more in-depth introduction to the general properties of phenols relevant to molecular imaging is therefore needed (recall too Introduction section 2, especially section 2.3).

Diffusion away from the activating cell or cell compartment is a perennial problem for most standard soluble fluorophores, especially common phenols like fluorescein, which may remain inside cells for only a few minutes (eg. Introduction Figure 5 and section 2.1).[1,16] However, a yet more severe problem is how to **quench their signal** completely and robustly in the proagent, for sensitive *in vivo* uses. Like anilines, most fluorescent phenols are moderately quenched by covalent derivatisation (see Introduction section 2.3). Commonly this gives phenolic proagents with a hypochromic shift of emission (eg. ~10-fold reduction), and usually a hypsochromic shift of both excitation and emission spectra (eg. at best, up to 50 nm for coumarins; the figures for anilines are usually less favourable). Note however that not all phenolic fluorophores show such a hypsochromic shift (eg Nile Red dyes do not)[30,39] and quenching may be especially problematic for the longer-wavelength fluorophores which are required for *in vivo* work. Still, by relying on the lower brightness of a covalently-modified phenol proagent, and also selecting a wavelength which more selectively excites the activated fluorophore, a relative ratio of fluorophore:proagent fluorescence of perhaps up to 100:1 can be obtained *in vitro*[40]. Yet this is far from being competitive with a true off-ON probe (see section I.3.1.4).

In passing, note that popular bisphenolic or bisanilinic fluorophores such as fluoresceins and rhodamines, which are of especial interest as they are the most common longer-wavelength fluorophores, represent an exception to this rule. When bis-substituted, their electronic properties often are modified enough to give closely off-ON behaviour when activated to the mono-/non-substituted forms.[22] However, the usual approach of bis-substrate attachment (eg. **3** or FDA, Introduction Figure 5) not only heavily **retards** the attainment of signal plateau, but also complicates the **signal↔activity relationship**.[1] Mono-hydrolysis gives a certain fluorescence increase already, but then bis-hydrolysis increases the signal further, so correlating observed signal with a microscopic information will require

cavalier assumptions about what proportion of the agent is partially or completely processed at any time; as hydrolysis of the second substrate may have different intrinsic kinetics (if the leaving-group nature has changed, cf. fluorescein diacetate), and also depends on diffusion from the active site (chances of performing both hydrolysis steps with the one enzyme, or returning to the bulk between them), deconvoluting a microscopically reasonable signal↔activity relationship becomes very difficult. For better designs such as **6**, a single processing event results in an increase of fluorescence, and no second processing is possible, which solves both issues although sacrificing some signal intensity. However, this design strategy is a comparatively recent effort: the majority of longer-wavelength fluorogenic probes are based on these or similar bis-functional fluorescent cores being bis-substrate-substituted; while the most popular mono-substitutable aniline/phenol cores (the coumarins) remain firmly anchored in the near-UV/blue wavelength domain, which involve other problems (see section I.1.1).

However, although the *potential* proagent:fluorophore or off:ON signal ratios are fairly well-established *in vitro*, these signal-to-control ratios can rarely be reproduced *in cellulo*, because phenols face greater instability than do the anilines (cf. Introduction section 2.3); given the suboptimal Stokes shifts and absolute wavelengths of **cell-compatible** phenols (see below), standard covalent modifications to standard phenolic fluorophores may, on the whole, provide **mixed or little success as a design for *in vivo* work**[4,41-44]:

> *"Phenolic fluorophores... and their derivatives, are in widespread use.*[11] *These dyes suffer, however, from two significant limitations. First, their fluorescence relies on the phenolate form and hence decreases at low pH... Secondly, although their fluorescence can be masked by acylation of the phenolic oxygen, such esters are unstable to spontaneous hydrolysis. The extensive conjugation makes the phenolate an exceptional leaving group. These limitations are necessarily intertwined, as substituents that decrease the* pK_a *of the phenolic hydroxyl group relieve the pH-dependence near physiological pH but enhance the instability..."*[45]

Probe stability must be ensured for *in vivo* applications, so the phenol pK_a cannot be lowered. Therefore **pH/environmental dependency** of the observed fluorescence signal will usually be a problem[40] for relating an observed amount of fluorescence to a molecular quantity of probe activation across different tissue environments. Yet even without pK_a modification, the **general instability** of almost all bonds to fluorescent phenols gives such elevated levels of probe degradation that with small *in vivo* enzymatic activities, the rate of spontaneous hydrolysis may strongly compete with or totally outweigh enzymatic processing, resulting in catastrophic loss of sensitivity beyond the *in vitro* situation. This instability is an especial challenge for the design of the auto-immolative spacers that are needed to adapt onto phenols for peptidase probes/prodrugs (detailed discussion in section I.4).[12]

Also, despite their brightness, soluble fluorescent phenols, especially fluoresceins, often suffer from more or less rapid **photobleaching**[17]. Their **Stokes shifts** are commonly 50 nm or substantially less, a significant sensitivity disadvantage. Lastly, a recent study on **far red/NIR**-emitting soluble phenols (mainly acridinone and oxazinone derivatives) has strongly concluded that they are **incompatible with *in cellulo / in vivo* work**: since they incorporate a thiol-reactive quinonimine ring, their fluorescence is totally quenched in the presence of free thiols (eg. glutathione, ~10 mM[46] in cells).[30]

Many problems face the design of phenolic probes for detection of enzyme activity, based on the use of standard phenols. I now single out the problem of proagent signal quenching for special attention.

1.3.2. Mechanistic quenching for off-ON probes

This selective argument for the necessity of total proagent quenching is expanded from the summary given in the Introduction (section 1.2).

(1) *In vivo* imaging must detect very small, native levels of enzymatic activity as sensitively as possible; unlike the *in vitro* case, one cannot simply "add more enzyme" to favour large probe conversion and therefore large signal. To maximise the difference between the regions with **true enzymatic signal** (small amount of activated agent, + much proagent + background) vs **control regions** (~just as much proagent + background), mathematically requires a proagent which gives a minimum of signal.[30,47]

(2) Practically too, the **absolute control signal** is a limiting factor as it saturates the detector if the sensitivity is set too high, and drowns regions of low enzymatic signal in background noise. Therefore both relative change and ON-off modes, with their high control signal levels, are ill-suited to the sensitive detection of realistic amounts of enzymatic activity in biologically relevant situations.

(3) In a low-background experiment, the measured signal from a true off-ON probe directly gives the **concentration of the activated agent**, which can be correlated to the enzyme activity. For the other modes, the measured signal is instead a strong function of the proagent concentration too. If the relative signal change is from 1 (proagent) to n (agent), then nothing distinguishes between sites with [agent]=X, [proagent]=nX, or eg. [agent]=X/2 + [proagent]=nX/2. As n increases, the probe approaches off-ON behaviour; or conversely, as n approaches 1, the probe becomes useless.

Then, the inhomogeneity of the proagent distribution must also be considered. *In vitro*, experiments with low-n relative change probes in homogenous media can assume that both agent and proagent are homogenously distributed throughout the cuvette, and so extract the signal due to the agent only, by $[agent]_t = [progent]_0 - [proagent]_t$ (Figure I.3a). But *in vivo*, this is impossible: probes are inhomogenously and dynamically distributed in the organism from the moment of injection till the time of acquisition. Not only are they continuously filtered out eg. through the kidneys, but their distribution may also be strongly inhomogenous both on a cellular level and on the scale of the organism due to preferential bioconcentration in certain cell compartments[16] or tissue types[48]. So

unless the instantaneous concentrations of both proagent and agent can be orthogonally determined at all points in the volume of interest, or confidently estimated, the true enzymatic activation cannot be reliably determined. Furthermore, the error in correlating an observed signal with a causative amount of enzyme activity increases as either the experiment time, or the ratio of applied probe concentration to real enzymatic activity, increase: since each increase the likelihood that variations in observed signal are due to biodistribution rather than enzymatic processing. Once inhomogenous, higher *in vivo* background levels are added to these considerations, the sensitivity of all probe modes is still further reduced, but that of the off-ON mode is by far the most robust (Figure I.3d).

Therefore, *in vitro* results from relative change and ON-off probes vastly over-represent the sensitivity obtainable when concentration corrections are no longer reliably possible - *in vivo* and sometimes even in cell culture. For example, "we have designed and synthesized various[49,50] highly sensitive [relative change] fluorescence probes… nevertheless, they were sometimes insufficiently sensitive to detect biomolecules in living cells, despite high chemical sensitivity in cuvette"[16] (this example also suffers from the non-specificity of its probes, which as per Introduction section 2.3, should also have been expected). This thesis argues that the drastic **loss of sensitivity in moving relative change probes from *in vitro* to *in cellulo* and *in vivo* applications** should always be anticipated.

Note however, that relative change probes whose concentrations can be determined orthogonally, eg. by ratiometric methods, do exist,[13,51] and these can hope to perform concentration correction *in vivo*. However, this usually implies either complex/expensive probes or complex orthogonal measurements: even so, the final result may still not be competitive with off-ON designs for practical reasons, eg. (2).

1.3.3. A qualitative illustration of the need for off-ON mode probes

I prepared Figure 1 in the Introduction (reproduced here as Figure I.3) to illustrate some hypothetical results which one might expect for the three different probe modalities, in the favourable case of large enzyme activity, so that most of the probe in a highly enzyme-active tissue is processed. This is a great overestimate of what can be expected with small *in vivo* enzyme activities: in which case the balance swings much further still to favour the off-ON mode (see section I.3.1).

There were twenty possible signal levels in my images, roughly correlated with equal step-sizes of brightness perception in greyscale for human vision (cf. Wyszecki's greyscale perception model[52] applied to HSL-coded colours); the background is initially at level zero; the relative increase probe changes from level 8 (mouse) to 16 (centre of largest tumour); the ON-off probe from 16 to 4; the off-ON probe from level 0 to 18, though these levels are arbitrary and are intended as illustration only.

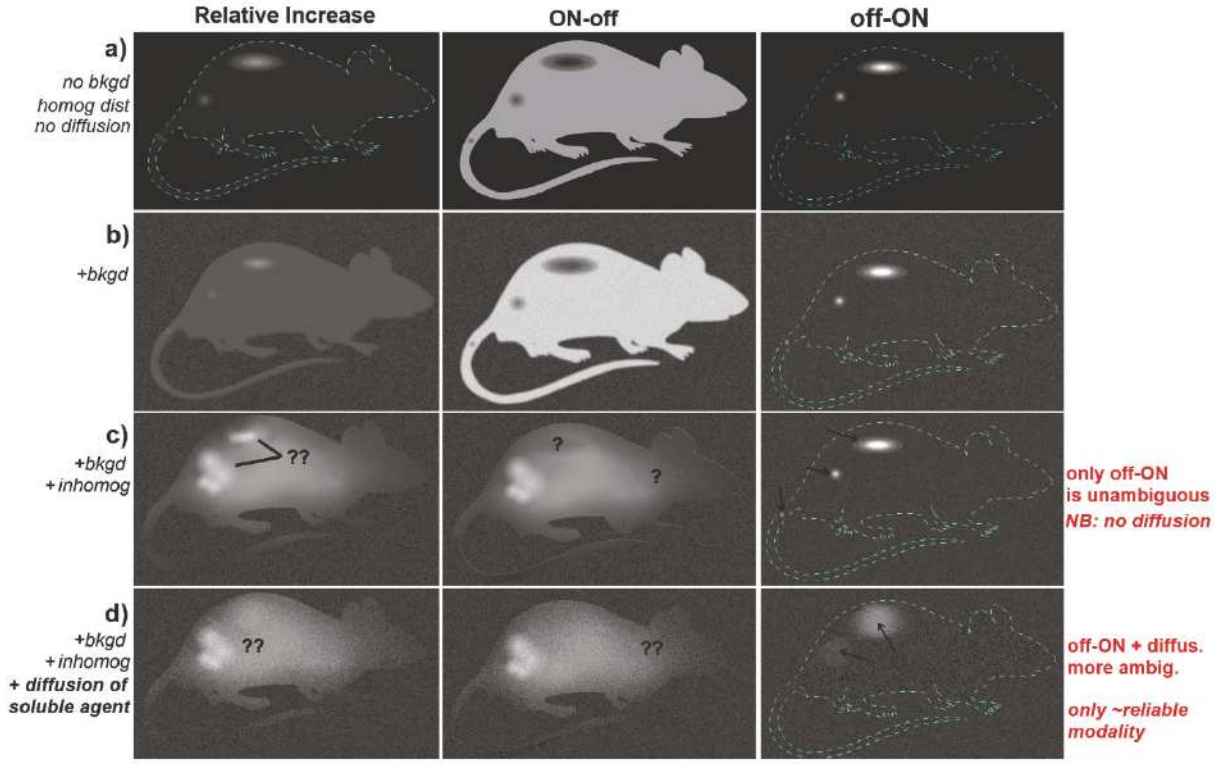


Figure I.3 – Practical ramifications of probe modalities in an *in vivo* setting: truly off-ON probes are required.

Row a: A mouse bears three tumours of increasing size, with increasing specific enzyme activity at their core. A probe is applied instantaneously and homogenously distributed; no subsequent diffusion of either activated or proagent forms is observed; and there is no signal background. We therefore obtain results as for *in vitro* experiments: clearly, both concentration correction and no-enzyme control subtraction are possible on these images, so the sensitivities of all three modes are similar when the enzyme activity is large. If the activity is small, clearly the off-ON mode is most sensitive, eg. examining the signal-to-background ratio around the small tumour in the tail: see point (1) above.

Row b: A uniform background signal plus some small scatter has been added to the images. By points (1) and (2) above, this visibly detracts most from the sensitivity of the ON-off and relative increase modes (again, see tail tumour), but much less from the off-ON probe.

Row c: Inhomogenous concentration of the probe before processing has been added (especially, towards the area of the kidneys); it is still assumed that after processing the activated probe accumulates without diffusion (rarely true). As can be seen, the results from the ON-off mode probe are useless: the darkest areas are the face and feet and tail (low proagent concentration), not the tumour (high agent concentration), so no interpretation is reliably possible. The relative increase probe also gives a large signal by bioconcentration, which makes characterising the large tumour doubtful (or else, mischaracterising the kidneys as tumours likely). The **off-ON probe is the only probe mode which retains unambiguous detection which is sensitive enough** to reveal all three tumours.

Row d: Diffusion of the activated agent after enzymatic processing, as well as of the proagent, is now shown (agent diffusion rate on the same order as the experimental timescale). This smears out the

signal from the activated agent and identification is difficult in all cases, which illustrates the critical problem of soluble fluorophore diffusion (cf. sections I.1.1.2 and I.1.3.1). Only the off-ON mode can hope to attain any selectivity under these circumstances, and even its sensitivity is much reduced.

Conclusions: (a) After data processing, *in vitro* tests with large enzyme activities can be used to show equal sensitivities for all three probe modes, but only off-ON probes reliably deliver *in vitro* tests whose sensitivity and unambiguity can reliably estimate what is possible *in vivo*. (b) Off-ON probes are mechanistically the most sensitive if the enzyme concentration is low. (c) Off-ON probes are needed *in vivo* when the concentration of proagent is dynamically inhomogenous. (d) For the realistic imaging of *in vivo* enzyme activity, there is a powerful need for probes which will accumulate without diffusion at the site of enzymatic activity, such as precipitating probes, so that the sensitivities can be as shown in panel c) rather than panel d). Of these, the **off-ON, non-diffusing probes clearly show the greatest *in vitro→in vivo* sensitivity**.

1.3.4. ESIPT mechanistic modulation fluorophores and the HPQ system

The Hasserodt group's research focus avoids soluble fluorophores, instead concentrating on rather unusual, Excited State Intramolecular Proton Transfer (ESIPT) solid-state phenolic fluorophores,[53,54] whose exceptional properties address ten key problems facing current phenolic probes. In particular, they provide an exceptionally favourable solution to the crucial problems of proagent signal quenching and agent diffusion.[1,17,54] However, until now no enzyme-responsive solid-state ESIPT probes have been exploited (ie., used in practice in more than a single report), except the excellent, widely-appreciated commercial probe "ELF-97," a nonspecific probe for any type of phosphatases.[1,55] This PhD chapter focusses on implementing auto-immolative spacers to finally permit aminopeptidase probes based on a solid-state ESIPT fluorophore core, hydroxyphenyl-quinazolinone (**HPQ**).

Fluorophores based on **HPQ**, notably including **diClHPQ** (the fluorophore of the ELF-97 probe, see Figure I.6), are highly insoluble, eg. K_S(**diClHPQ**)< 2 μM. (1) This stops them diffusing away from the site/cell compartment where enzymatic activity releases them, thus solving the problem of signal diffusion that prevents many soluble fluorophores from reliably localising enzyme activity *in cellulo* or *in vivo*.[15] While other approaches for preventing signal bleed out of cells do exist (eg. blocking export out of the cell, as will be seen in Part II, or reactive tagging), precipitation is clearly the simplest, most reliable, and least biochemically offensive. It also brings a host of other advantages: (2) **HPQ** fluorescence is totally environment-independent, because the crystal lattice is a constant environment no matter the external solution (compare to the variable results with soluble fluorophores dependent on pH, cell thiols, etc).[17] (3) **HPQ**s are also *immensely* photostable, with a photobleaching halflife that is orders of magnitude beyond that of standard soluble fluorophores. For example, under identical circumstances, fluorescein was photobleached at high power in only 5 seconds, while **diClHPQ** showed only a 4% loss in fluorescence over 100 s (ie. halflife ~2000 s).[17] This photostability can be attributed to the insulation of the excited fluorophore from the surrounding

medium when it is in the crystal lattice, preventing excited-state reactions with dissolved species. (4) Their insolubility also minimises the biochemical and photosensitising stress of the fluorophore on the system under study. For example, 100 µM of liberated **diClHPQ** will only expose the cell to <2 µM of the dissolved phenol, whereas for soluble fluorophores such as acridinones, the full liberated concentration is free to react with cell species so the potential toxic loading is far greater.

However, the **HPQ** system provides a second unusual molecular feature whose consequences on macroscopic probe properties are even more important: the principal mode of fluorescence generation is by the Excited State Intramolecular Proton Transfer (ESIPT) process (Figure I.4).[55,56]

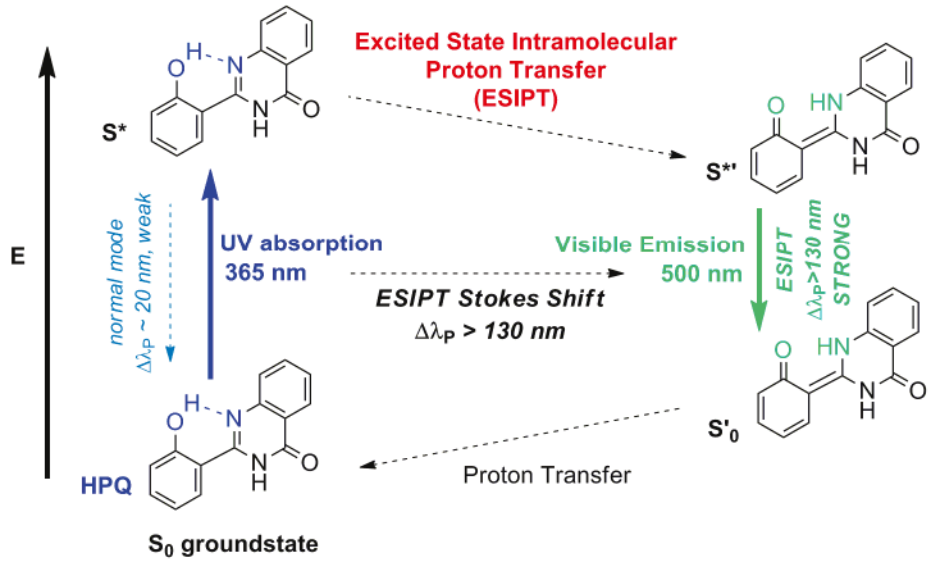


Figure I.4 – Energy diagram of the ESIPT process, as illustrated for fluorophore HPQ

The ESIPT mechanism dissipates excited-state energy non-radiatively by proton transfer from the excited phenol S* to give an excited ketone form S*', which then relaxes radiatively to the ketone groundstate S_0'. Since the groundstate ketone is higher in energy than the groundstate phenol, subsequent back-transfer of the hydrogen once more dissipates energy nonradiatively. The net result is a significant nonradiative dissipation of energy accompanying the radiative relaxation, ie. an exceptionally large Stokes shift. For **HPQs**, this shift is so large that it tends towards substantial NIR emission even when excited in the near-UV, although **HPQs** avoid the (thiol-reactive) quinonimine ring which is a standard feature of soluble phenolic red/NIR emitters. ESIPT is actually not such an uncommon mechanism: it is seen, although more weakly, by the characteristic blue-fluorescent spots on TLC in undergraduate experiments with salicylaldehyde, or during the twin esterification+ transesterification of aspirin to oil of wintergreen[57], illuminating at either 254 or 365 nm.[58]

The most obvious advantage for ESIPT-based probes is therefore (5) their enormous Stokes shifts, eg **HPQ** and **diClHPQ** show $\Delta\lambda_P$=135-175 nm, which greatly outstrips standard soluble phenolic fluorophores (15-50 nm, but usually less when towards the NIR; Figure I.5). This can promise an exceptionally low signal background for *in cellulo* or *in vivo* imaging.

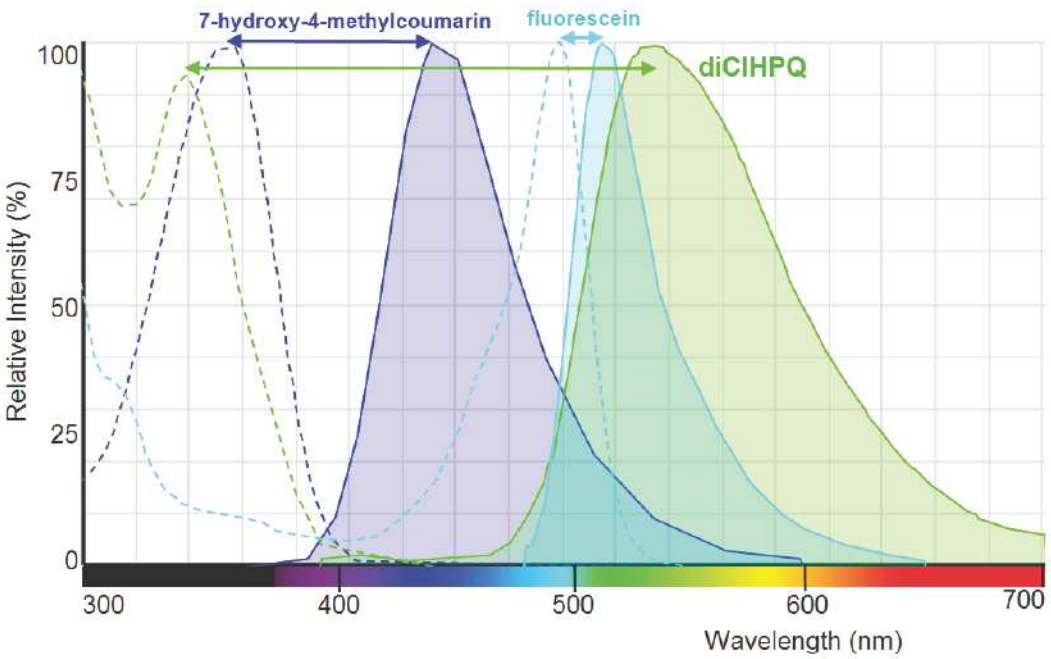


Figure I.5 – Excitation (dotted) and emission (shaded) spectra for solid-state ESIPT fluorophore diClHPQ (green) compared to non-ESIPT, soluble fluorophores 7-hydroxy-4-methylcoumarin (purple) and fluorescein (blue) shows the enormous Stokes shift obtainable with the HPQ system, and the extraordinary ~100 nm baseline separation between its excitation and emission spectra, as well as the long-wavelength emission tail (10% at 680 nm).[59]

If the ESIPT is highly favoured relative to the normal fluorescence pathway, then (6) very high quantum yields of the ESIPT fluorescence can be obtained. For example, Φ~0.36 for **diClHPQ**, whose phenol is acidified by the *para*-chlorine and *ortho*-carboxylic acid derivative, and where π-stacking in the crystal lattice aligns the imine nitrogen exactly in the plane to receive the hydrogen transfer. Therefore, strongly-absorbing systems like the **HPQs** are bright ESIPT fluorophores.

Yet best of all, ESIPT solves the problem of proagent signal quenching, finally allowing sensitive, truly off-ON probes: (7) ESIPT-fluorescence can be **mechanistically quenched** giving no ESIPT signal whatsoever from the proagent state, if the transferrable phenolic hydrogen is replaced with any masking group (eg. ester, glycoside, ether, phosphate, or in this work, carbamate). There is also a second, more subtle consideration: free ESIPT fluorophores dissolved in hydrogen-bond-acceptor solvents show no ESIPT fluorescence either.[53,54] This is because uninterrupted proton transfer is required for the ESIPT process to be favourable. Therefore, (8) precipitating an ESIPT fluorophore from aqueous media provides **a totally general microscopic method to activate large-Stokes-shift, ESIPT fluorescence in a macroscopically binary off-ON mode**. Now, just like tissue autofluorescence, the normal-mode fluorescence from ESIPT fluorophore cores (eg. 20 nm) cannot approach the Stokes shift of their ESIPT fluorescence, (eg 130-200 nm) so any residual normal-mode fluorescence in the ESIPT-quenched proagent may be avoided by using a large wavelength split in the fluorescence detection. The **HPQs** have the further advantage that their normal-mode fluorescence is also very weak - at least an order of magnitude less strong than eg. $p\text{TsO}^-$, which is not even considered a fluorophore (see section I.5.3.3). Therefore (9) designs which can ensure that nonenzymatic release and precipitation of the ESIPT phenol form is impossible, ie. probes involving a

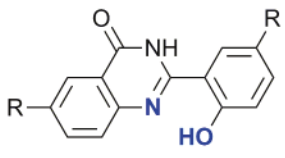
stable bond masking the phenol, which is only broken after enzyme activity, are therefore truly and completely, mechanistically off-ON both microscopically and macroscopically,[17] as is required for sensitive *in vivo* imaging.

This is not at all equivalent to the situation for covalent modification of standard fluorophores, even if some may give a high signal-to-control ratio. Totally mechanistically off-ON probes operate in a mathematically superior regime of attainable detection sensitivity, which far more easily permits powerful *in vivo* applications, and are subject to far less strict sensitivity constraints (section I.3.1).

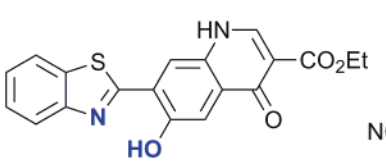
Note too, that an amount of probe greater than the fluorophore's solubility limit K_S must be processed before the fluorophore starts precipitating and signal is observed. This usually incurs a certain delay time. However, I show later that actually, (10) the reliability of enzyme quantification is much improved by adopting an off-ON precipitating design: the period when fluorophore accumulates without precipitation acts as a safety buffer ensuring reliable attribution of signal in a noisy, *in vivo* context. In one sample *in vivo*-style calculation, off-ON precipitating mode probes give a maximal error of 18% in quantifying enzyme activity: compare this to truly off-ON but soluble probes (200% max. error), or worse still, 50-fold relative increase probes (90,000% max. error; see section I.3.1).

In light of these ten selected advantages, the Hasserodt group considers that **precipitating off-ON ESIPT probes offer truly exceptional advantages for sensitive and reliable *in vivo* imaging**.

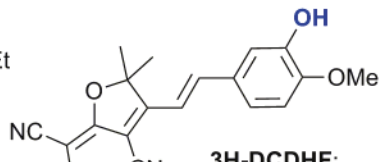
HPQ derivatives are not the only ESIPT fluorophores suitable for use in this manner. One example which was also selected was the highly insoluble red-fluorescent benzothiazole fluorophore **BTQC**, recently reported by Stefani *et al.*[60] as having an absorption maximum at 440 nm and an emission maximum at 630 nm; the enormous Stokes shift can once again be attributed to the ESIPT mechanism in the solid state. It was attractive compared to the **HPQs** because of its longer wavelength absorption/emission (approaching more 'cell/tissue-friendly' values) and its qualitatively greater insolubility (so it should give more rapid fluorescence response, cf. sections III.3.4 and III.5.5.2). Although its brightness qualitatively seemed far lower than that of the **HPQs**, this only somewhat compromises its sensitivity, as the Stokes shift is usually far more important (cf. section I.3.2). The critical factor was that **BTQC** synthesis requires several steps, the last of which has a yield of 5-10%, while the **HPQs** have cheap single-step synthesis: so **HPQs** were to be used in initial trials.



HPQ: R=H; λ=365/495 nm
diClHPQ (= Cl_2HPQ): R=Cl; λ=345/530 nm



BTQC: λ=440/630 nm



3H-DCDHF:
λ_{ex}=575 nm (+TPA)

Figure I.6 – Fluorophores considered for work towards mechanistically-quenched, off-ON precipitating probes.

Although my PhD work focussed on ESIPT fluorophores, these are not the only mechanistic possibilities for off-ON functionality. One such example of a different off-ON mechanism may be

shown by the red-fluorescent 3-hydroxylated dicyanomethylenedihydrofuran **3H-DCDHF**. Drs Chantal Andraud and Yann Bretonniere of the research group Chimie pour l'Optique indicated during a discussion that the bright and highly photostable[61] DCDHF push-pull motif gave two-photon absorption almost only in the precipitated form, and was in general strongly insoluble. I therefore proposed a derivative of this core with a phenol in a *meta*-position (ie. not involved in the crucial push-pull process) as a potential solubility-based off-ON fluorophore system, giving what proved to be the known structure **3H-DCDHF**[62] (Figure I.6). My consideration was that derivatising this phenol with a solubilising enzyme substrate construct (either charged, such as a peptide, or hydrophilic such as a glycosidic derivative) might prevent two-photon-fluorescence from a proagent form by keeping it in solution; after enzymatic processing and spacer immolation, the fluorophore could precipitate and two-photon fluorescence might therefore activate. This is a different mechanism for off-ON functionality than for the ESIPT probes I studied, but could show similar signal localisation advantages from precipitation. Also, since two-photon fluorescence imaging should have no upconverted background signal whatsoever, this strategy might also offer a high sensitivity off-ON approach to enzyme imaging. This and related DCDHF fluorophores are currently under further consideration for future work in the group.

To conclude, the intrinsic properties of solid-state ESIPT fluorophores address or resolve all of the chief problems facing probes based on soluble phenolic fluorophores, **except the intrinsic instability of the bond joining the spacer to the inherently good-leaving-group phenol**. This is especially critical for ESIPT fluorophores, because anchimeric hydrolytic assistance by the *ortho*-situated hydrogen acceptor's lone pair greatly lowers stability even relative to similarly inductively-activated compounds.[63,64] Therefore, standard phenolic probe designs cannot be simply transposed to the potentially more-sensitive ESIPT fluorophores without actually losing (greatly!) in sensitivity: **their instability has so far prevented the reliable exploitation of ESIPT fluorophores in any probes** other than the phosphatase probe ELF-97, where the monophosphate's negative charge stabilises the ester against hydrolysis.[55] Overcoming this instability requires the design of new, exceptionally stable auto-immolative spacers, which are still performant enough to give rapid post-triggering response.

1.4. Prior auto-immolative spacers for phenolic peptidase probes

As Meyer *et al*. wrote explicitly in 2010, there was no existing spacer that would present optimal solutions to all the requirements for constructing the desired peptidase-sensing phenolic probes:

> *"[we had to try to] find the best compromise between stability, self-reactivity upon enzyme triggering, and substrate scopes and limitations[...] to select an optimised linker for the construction of protease-sensitive fluorogenic probes fulfilling all requirements for* in cellulo *or* in vivo *imaging applications."*[12]

The lack of a clearly optimal or even successful spacer is echoed by Lavis *et al.*: "many of the existing auto-immolative linkers add undesirable size and synthetic complexity to the probe or cannot offer significant advantages to phenolic fluorophores, such as fluorescein."[4]

What is required – not just for this PhD, but urgently for the field of molecular imaging and molecular medicine in general – is a spacer or stacked spacer set which connects an N-terminal peptide substrate to a phenol, and which would give to the probe: (1) exceptional hydrolytic stability, (2) good enzymatic processing of its substrate, (3) rapid auto-immolation, (4) only nontoxic by-products, and (5) an easy, cheap and modular synthetic route suitable for targeting general exopeptidases with general phenolic drugs. A review of the literature can conclude that such spacers simply do not exist: prior-art spacers always involve a tradeoff favouring some of these qualities at the expense of others. Most of these points have already been raised, especially in sections 2.3-2.4 of the Introduction, but a quick summary is given here to reassemble that information:

Elimination spacers: PABA is the most-used conjugated elimination spacer (**11** and Introduction Figure 11), but does not have sufficient stability, especially with chlorinated phenols[30] such as the target fluorophore **diClHPQ**. PABA-type anilides show severely non-optimal enzymatic processing by native exopeptidases, as compared to aliphatic primary amides,[21] although their immolation kinetics are ~instantaneous and their synthesis is easy and general, at least for exopeptidases. The critical factor is the well-known toxicity of the azaquinone methide byproducts, which is entirely unacceptable for *in vivo* studies, and even *in cellulo* studies where the cells ought not to be killed.[34,65-68] Therefore PABA-type conjugated elimination spacers were rejected, as were any stacked spacers involving them, because these three problems were unlikely to be overcome simultaneously.

Cyclisation spacers: 1,2-diamines are the most developed of these spacers. Because N,N'-dimethylethylenediamine **13** is not recognised by native peptidases, N'-methylethylenediamine **12** has been the cyclisation spacer most widely used in biological probe design.[21,69] However, the kinetics of its cyclisation to release an attached agent are too slow for use *in vivo*: the substrate-dependent cyclisation halflife is around 5 h for phenols,[70] or 23 h for alcohols,[21] and this is more than enough time for the triggered probe to diffuse far from the site of enzymatic activity, invalidating the concept of an *in vivo* probe. Several other cyclisation spacers which either do not solve the cyclisation kinetics problem,[26,71] or else which achieve better kinetics but lose stability[72], were also rejected. For example, prior work in the Hasserodt group on phenol-releasing amine-terminal cyclisation spacers had focussed not on diamines but on γ-aminobutyric acid-derived spacers eg. **34b** and **34a**, but these were both too unstable for further development, and also synthetically delicate (see Figure I.31, sections I.5.6.1 and III.3.5.1).[73]

However, 1,2-diamine probes do have all the other advantages sought, ie. excellent **hydrolytic stability** and **enzymatic processing**, **nontoxic byproducts**, and **modular syntheses**. Therefore,

1,2-diamine-based designs were considered promising for further development if only the cyclisation kinetics of the diamine spacers could be adressed by the design of a suitable, novel spacer system.

2. A Modular, Robust, Practical off-ON Phenolic Probe System

2.1.1. Spacer design

Implementing a fluorophore-spacer bond both stable towards spontaneous hydrolysis to avoid generating a false positive signal, yet selectively cleavable upon enzymatic activation, has been a major challenge in prior work in the group, especially with electron-withdrawn, delocalised phenolic fluorophores.[63,73] Prof. Hasserodt's design for the current work therefore began by selecting a tertiary carbamyl link to maximise probe robustness, as tertiary carbamyl bonds are very stable in aqueous and biological media;[74,75] the carbamate is a general, modular linkage that can easily be synthesised at a late stage by adding a carbonyl equivalent, eg from phosgene or CDI. 1,2-diamine-based cyclisation systems were chosen as the spacers, because 5-membered cyclisation provides more rapid response than 6-membered cyclisation, and rapid response is a crucial parameter for eventual *in vivo* use.[70]

To create a 1,2-diamine system which would be enzymatically processed required the substrate to be attached to a primary amine site;[12,21] however the best known primary amine spacer *N'*-methyl-ethylenediamine (**12**) simply has too slow response kinetics (~5 hours cyclisation halftime).[70] A solution to this problem has been known for decades in other systems, though never yet applied to diamine spacers: cyclisation preorganisation. For example, the classic *gem*-dimethyl Thorpe-Ingold effect increases the cyclisation rate by orders of magnitude[76], attributed both to higher populations of the reactive rotamer, and to effectively lower loss of degrees of freedom upon cyclisation. This strategy has been notably used to great succes in the design of many trimethyl lock (**TML**) spacers, which show superlative cyclisation relative to their non-preorganised counterparts: up to 11 orders of magnitude rate enhancement.[77,78] This resulted in Prof. Hasserodt's choice of spacer **21** (Figure I.7).

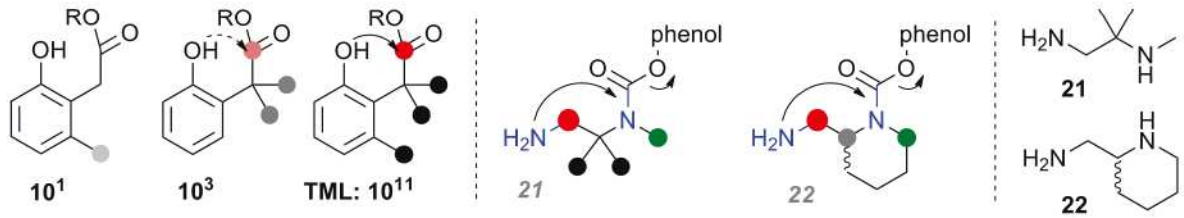


Figure I.7 – Cyclisation preorganisation, cf. the trimethyl lock (TML), applied to the design of dimethyl spacer 21; piperidyl spacer 22 has more favourable annelation preorganisation, also minimising steric demand at the β-carbon.

Conceiving further improvements spacers required some thought. Bridging the *gem*-methyl groups in a spiro ring would have enhanced the cyclisation rate but at the expense of a still bulkier β-carbon that would have risked slower enzymatic processing. A more elegant solution was to realise that the steric bulk of the two methyl groups is not *actually* needed. What *is* crucial, is that the reactive methylamino group (red circle on the image) should be kept as much as possible in the reactive rotamer conformation depicted: ie. oriented towards the carbamate, rather than pointing away from it in the extended conformation which is favoured for **12**; or alternatively, the molecule should lose fewer effective degrees of freedom upon cyclisation. The *gem*-dimethyl effect and its analogues are only one way of achieving this.

Annelation was therefore selected as an even better *kinetic* solution, since by locking the methylamino carbon *permanently* into the reactive conformation, the loss of degrees of freedom upon cyclisation is reduced further. Such preorganisation has already been seen eg. in probe **19**, where the annelating proline favours cyclisation over the linear case, even if it seems unclear that **19** really does cyclise as intended; annelation preorganisation is however confirmed in other settings outside probe development.[79,80] Then, annelation was elegantly combined in one cyclic group with the stability requirement for a tertiary carbamyl link, which minimises the *steric presence* at the β-carbon to the scissile bond. Prof Hasserodt's second design (**22**) therefore could perhaps have even more promise, both from the viewpoint of cyclisation kinetics and for enzymatic processing (Figure I.7).

2.1.2. Substrate choice

Leucine was chosen as a synthetically easy model peptide substrate (cheap monopeptide with only one protecting group, apolar). A model aminopeptidase enzyme target which efficiently cleaves it ("aminopeptidase M", EC 3.4.11.2, also known as alanyl peptidase, leucine aminopeptidase, etc) was cheaply available and could be used in *in vitro* tests. This substrate-target pair is not thought of great interest in biochemistry: despite claims that its limited correlation[81] to tumours is a justification for LAP probe development[82], it is actually unspecifically distributed in all tissues.

However, once the proof-of-concept was performed, more relevant substrates/peptidases were to be explored, to demonstrate the utility of this modular approach. ***β-alanine*** (3-aminopropanoic acid) was chosen as one such relevant substrate. Though a monopeptide, it is a non-natural amino acid and thus may have high specificity of enzymatic processing. It is of particular interest due to its cleavage by endogenous enzymes native to several key infectious bacteria, especially *pseudomonas* spp.[83]. Fluorogenic probes detecting β-alanyl aminopeptidase would be useful to rapidly and specifically detect *pseudomonas* infections, allowing for their earlier treatment and thus better patient outcomes.[15] As β-alanyl aminopeptidase was not commercially available, *in cellulo* tests of the probes' specificity for *pseudomonas* cultures were to be performed externally.

2.1.3. Assembled designs for probes

With the spacers and substrates chosen, and the fluorophores **HPQ**, **diClHPQ** and **BTQC** retained from the discussions of section I.1.3.4, it remains to draw up final designs and check that they responded to the design criteria. Sample probes from this chemical space are presented in Figure I.8. Note that the unusual insolubility of **HPQ** and several other ESIPT fluorophores requires special care to conceive soluble probes based around them (cf. probe requirements in Introduction section 2.1). The **HPQs** are however an inspired choice for the current approach, as the presence of the phenolic hydrogen and the overall planarity of the molecule are probably two of the major driving forces behind their insolubility, and both are disrupted by the phenol derivatisation that is used to block the possibility of fluorescence in the proagent. The final major insolubility factor is the dimerisation

between quinazolinone amides opposite the ESIPT side; however, by blocking the phenol with a spacer bearing only a single charge on the amine substrate, an increase in solubility by more than three orders of magnitude was immediately obtained, which was more than sufficient for the present work.

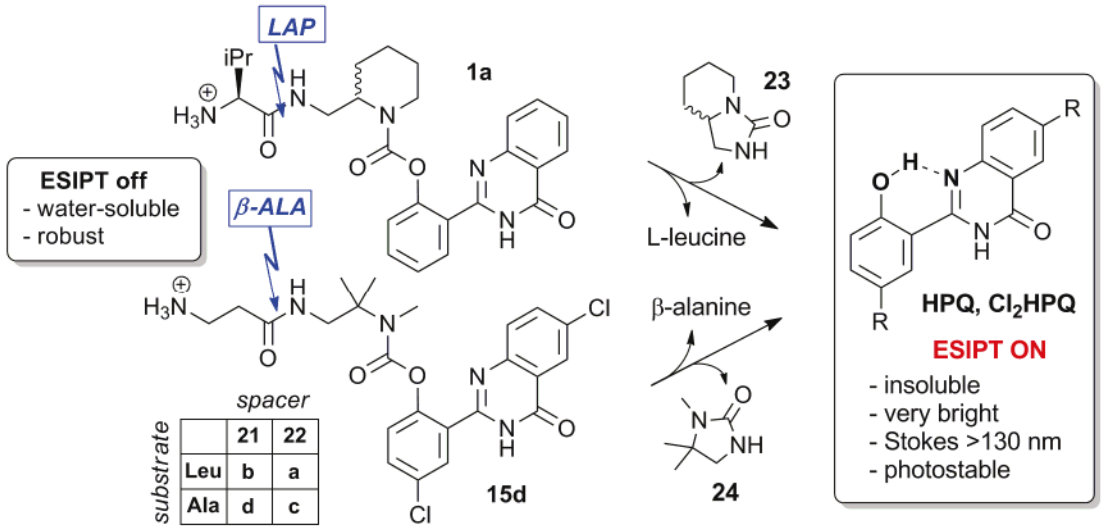


Figure I.8 – Sample probes drawn up from the design phase. β-ALA = β-alanyl aminopeptidase

The weakest bond in the probes is the tertiary carbamyl link, which is nevertheless known to have outstanding hydrolytic stability. The probes are small (≤500 Da); they bear a single charge on their peptidyl substrate at physiological pH; and the CLogP values are also acceptable (ChemDraw estimates 2.1 and 3.3 for probes **1a** and **15d** respectively): therefore the Lipinski guidelines for design of water-soluble yet membrane-permeable druglike agents are addessed.[84,85] Critically, the carbamyl bond in the proagent state replaces the phenolic hydrogen, so no ESIPT-mode fluorescence is possible, therefore the proagent is mechanistically *off* before activation. The preorganised yet sterically undemanding designs for the cyclisation spacers should ensure good enzymatic recognition and processing of the targeted peptide bonds, followed by rapid cyclisation to release the fluorophores. These should still be incapcable of ESIPT fluorescence in solution, but their low solubilities should ensure precipitation once they accumulate past their low K_S thresholds, ensuring functioning in the off-ON precipitating mode. The ESIPT-fluorescent precipitate should be excellently photostable, thus marking the site of enzymatic activity with the bright, large-Stokes-shift fluorophore. It was hoped that the ureas **23** and **24** would prove non-toxic, thus not poisoning the enzyme (as quinone methides may do), and allowing for true catalytic amplification of the signal.

The Hasserodt probe designs were therefore complete and seemed to satisfy the design criteria. Yet rather than proceed at once to synthesis and enzymatic validation, I wish now to take a 'detour' and report on theoretical analysis that I performed to understand and quantify several unusual aspects of the probe system, since this understanding will ultimately be required to design appropriate assays and to interpret the data collected from them.

3. Pre-Modelling of Probe Performance

The final probe designs are conceptually far removed from prior art in fluorogenic designs for peptidase detection, so it was not obvious just what degree of improvements in performance could be expected. I therefore constructed two models to evaluate the *mathematical* advantages in sensitivity for *in vivo* applications of using off-ON probes of high-Stokes-shift precipitating fluorophores, as compared to (1) relative change probes and off-ON soluble probes; and (2) lower-Stokes-shift probes. It was also unclear what problems I would face in data analysis, so (3) I also performed a preliminary analysis of what microscopic information the measurable fluorescence signal from a precipitated probe can be expected to give; this was later to be crucial for quantitative treatment of experimental data.

3.1. Detection sensitivity↔probe mode model

The Introduction presented a qualitative argument for the design and use of truly off-ON probes to achieve high sensitivity *in vivo* (section 1.2), and section I.1 presented additional qualitative arguments for the use of mechanistically off-ON *precipitating* probes for still better mathematical implementation of the high-sensitivity concept, as well as for practical reasons (signal localisation/bleeding, photostability, etc). However, I considered that it was important to provide a more quantitative argument for this unusual probe system. As precipitating fluorophores have not yet successfully been widely used in stable probes suitable for *in vivo* work[73], they are rather an unknown quantity. As a reviewer for our publication[86] asked, what sensitivity can be obtained ***in practice*** by the off-ON precipitating probe system with solubility constant K_S, compared to standard systems? The reviewer could as well have asked, "why not simply implement these new spacers in phenolic probes based on more classical soluble fluorophores such as fluorescein or coumarin? After all, these *in vitro* studies cannot fully benefit from the *in vivo* advantages of large Stokes shift and signal localisation that ESIPT fluorophores are more well-known for, but must instead suffer the extra problems from having to pass the K_S before generating signal, and especially from developing new assay conditions and data analysis procedures [see section I.5.1] rather than profiting from established ones..."

Detection sensitivity determination is a complex statistical procedure which has been thoroughly explored in signal processing physics, particularly for binary decision thresholds (a signal is present or else no signal is present)[87]. However I was unable to find a suitable approach for *reliably quantifying* enzymatic activity (a binary decision of presence/absence of a threshold enzyme activity is not sufficient), comparing different *modes of probe operation*. This is not surprising, since prior art probes are almost without exception relative change probes, whose sensitivity is most strongly limited by the relative change magnitude n, and this is easily compared (eg. "10× fluorescence increase"; see section I.1.3.2) so no more thought need be given to the matter. For truly off-ON mode probes, an "infinite fluorescence increase" n is clearly not a sensible sensitivity-indicating parameter.

Therefore I developed my own approach and model from scratch, to compare the sensitivities that can be expected from relative change, off-ON, and precipitating off-ON probes in a more satisfying, quantitative manner. I abandoned 'solving' an initial analytical model as it involved too many assumptions; instead I present a numerical model which illustrates the trends.

3.1.1. Philosophy of the sensitivity model

General Approach: This model intercompares the sensitivity of the three probe modes in a range of settings mimicking *selected* experimental parameters from *in vitro* to *in cellulo* and *in vivo* cases.

(1) Two crucial factors for *in vivo* sensitivity are ignored: (a) The model considered that the proagent is statically and homogenously distributed before enzyme processing, and no diffusion occurs after agent activation: ie. an *in vitro* situation (Figure I.3b). This avoids the most advantageous situation for off-ON and especially off-ON precipitating probes, ie. *in vivo* dynamically inhomogenous concentrations. Only the off-ON precipitating mode is unaffected by these conditions, therefore *only the off-ON precipitating mode can hope to reproduce these calculated sensitivities if these experiments are performed in practice in cellulo/in vivo*; the other modes will greatly lose in sensitivity (compare their loss of sensitivity in going from Figure I.3b→Figure I.3d). (b) The same intrinsic background B_{ACQ} (~autofluorescence) was used for both the precipitating probes (eg. **diClHPQ**) and the soluble ones (eg. fluorescein). Given their Stokes shifts (185 and 24 nm, respectively), *far* greater background is expected for such soluble probes *in vivo*; so once again, only the off-ON precipitating probe could hope to reproduce these sensitivities in practice.

(2) These *practical* models specifically considered the scenario when **enzyme activity is very low compared to the available probe concentration**, as this is a typical *in vivo* situation. When enzymatic activities used are instead *large* compared to probe concentration, detection sensitivity can be uncoupled from probe hydrolytic stability, by out-competing hydrolysis as the major signal-generating mechanism. In this author's opinion this is often exploited for *in vitro* tests by researchers *wishing* to report highly sensitive probes but who have in fact synthesised unstable ones. A possible example is Urano (JACS 2011)[40], whose *in vitro* assays used ~50 U/mL enzyme concentration – ie. 50 $mM.min^{-1}$ of standard substrate can be converted – with an initial probe concentration of only 1 µM. I do not see any scientific justification for failing to state the motivation behind the use of such weighted conditions, especially when the results are used to claim what can be expected *in vivo*.

(**3**) Data were to be **treated without subtracting 'background readings'**. An ideal experiment would leave the subject free to move about as normal and simply be measured hours or days after injection, giving a good signal-to-background ratio by inherent probe characteristics (eg. off-ON probe in Figure I.3c) rather than data processing. Background due to the **environment/acquisition conditions, B_{ACQ}**, may be difficult to subtract *in vivo* as it would require the subject essentially to remain immobile during the experiment, however it may be possible to reduce it. For an *in vitro* experiment, B_{ACQ} removal corresponds to subtracting the signal from a solution *with no probe added* (no-probe control),

so is universally performed; but it is important to remember that the *in vivo* situation is different. What is almost never possible to remove *in vivo* however, is the **background due to the unactivated probe, B_{PR}**. *In vivo*, a probe is continuously activated from the moment the injection starts, and diffusion / biodistribution of the probe is a highly complex process that cannot be simplified to a 'uniform probe background level'. So for an *in vitro* experiment to give a better estimate of maximum attainable sensitivity *in vivo*, the signal-to-background ratio reported should be the ratio of signal from an active cuvette, to the signal from a no-enzyme control cuvette at the same timepoint: the **signal to control ratio**. Only if B_{ACQ} subtraction is explicitly justified (eg. "strictly *in vitro* calculation only"), should the signal from a no-probe cuvette be subtracted from each. **The no-enzyme control signal should never be subtracted as if it were a blank** when claiming a probe for *in cellulo/in vivo* applications! Also, if a probe truly *is* off-ON, $B_{PR}=0$ so B_{PR}-subtraction *should* be unnecessary: this questions those studies which claim off-ON performance, yet use B_{PR}-subtraction.

General Conclusions: The primary considerations determining sensitivity *within* each mode will be shown to be probe stability, and intrinsic background signal. *Despite* the twin handicaps for showing the superiority of the off-ON precipitating mode (point **1**), its advantages over both other probe modes even *in vitro* are unambiguously clear. If those handicaps are then considered, the *unambiguous superiority*, and necessity, of the off-ON precipitating mode for translation to *in vivo* studies, is argued far more strongly still. This modelling provides the theoretical justification to my experimental conclusion that this probe mode, which the present PhD work has implemented for the first time for peptidase probes, does indeed deliver a new and mathematically superior regime of detection *ultra*sensitivity for both *in vitro* and *in vivo* imaging.

3.1.2. Setup of the numerical models

Intrinsic background signal B is split into two components: the **probe-concentration-independent intrinsic acquisition background B_{ACQ}**, and the **background from the proagent B_{PR}**: thus $\mathbf{B=B_{ACQ}+P_R \times B_{PR}}$, with a concentration P_R of probe in the pre-processed state. B_{ACQ} for fluorogenic probes mainly reflects how much autofluorescence is collected at $\lambda_{monitoring}$ after excitation at λ_{ex}; B_{PR} reflects just how "off" the *off*-state really is. The intrinsic background will be shown to be a **more vital** parameter for achieving high sensitivity *in vivo* than the probe's emission intensity, even though the dependency on B can be hidden by unrealistic *in vitro* tests (high enzyme activity). B was parametrised relative to the **signal emitting strength S** of one µM of activated probe, thus giving B/S; so the component B_{ACQ}/S measures the concentration of probe implied by a measured background signal; and B_{PR}/S measures signal emitted by 1 µM of proagent relative to that of 1µM of active agent.

Probe spontaneous hydrolysis H is the unimolecular rate constant for proagent hydrolysis (min^{-1}).

Enzymatic activity A and Michaelis-Menten constant K_M ($\mu mol_{probe} \cdot min^{-1} L^{-1}$ and µM, respectively): I assume a low *in vivo* enzyme concentration relative to the initial probe concentration,

and (crucially) a *competitive probe* which is processed efficiently even in the presence of the enzyme's native substrate, so I approximate the enzymatic processing with simple Michaelies-Menten kinetics.

Precipitating probe Solubility Constants K_S (μM) are based on experimentally determined values.

Note that each of these factors can be addressed separately in practice during probe/assay design. For **H**, the probe must be as stable as possible. Rather than developing ever more expensive detection setups, **B/S** can best and most cheaply be minimised with good design of the probe, so that there is minimal background collected with the signal output. For example, a large Stokes shift of the probe decreases the autofluorescence contribution to B_{ACQ} potentially by orders of magnitude (see section I.3.2), and a truly off-ON probe will have B_{PR}=0. While high S, ie a probe with a bright *collectable* signal, is also important, the crucial factor determining this will generally be the Stokes shift and *not* the traditional measure of total quantum yield, Φ (see section I.3.2). During *in vivo* work, $\mathbf{A/K_M}$ can only be addressed by using a probe design which ensures efficient and competitive enzymatic processing, but can be (questionably) adjusted *in vitro* by adding more purified enzyme; $\mathbf{K_S}$ is determined by the choice of fluorophore but likewise can be (questionably) adjusted *in vitro* by pre-spiking and cosolvents (see section I.3.1.6); yet both adjustments misrepresent what is possible *in vivo*.

Lastly, it is assumed that the enzymatic hydrolysis step is the only limiting step in the release of fluorophore, ie. the spacer response and precipitation times are negligeable. Especially at low enzyme activities, *with fast spacer cyclisation* (actually another relative novelty of the current PhD work; compare prior art spacer **12**), this was a robust assumption for this model.

3.1.3. Model formulation

Probe concentrations (μM): P_0: initial proagent; $P_R(t)$: remaining proagent at time t; $P_{TRUE}(t)$: running total for the concentration of probe enzymatically processed by time t (not including spontaneous hydrolysis);

→Background: $B=B_{ACQ}+P_R \times B_{PR}$ (RFU);

B_{ACQ} and B_{PR} are parametrised relative to S, so S is an arbitrary parameter (so for all probes I used S=10000 RFU·μM^{-1} on the Signal(t) graphs);

→Spontaneous hydrolysis rate = $H \times P_R$ (μM·min^{-1})

→Enzymatic rate = $A \times P_R/(K_M+P_R)$ (μM·min^{-1})

Probe concentration timecourses are calculated with small timesteps relative to rates A and H. The timesteps were between 0.1-0.25 min depending on the values of A and H used; before choosing the timestep for each run, I always confirmed that smaller timesteps did not give different model behaviour due to continuity effects. Rates A and H must be compensated for the size of timestep used; but to illustrate for a timestep size of one minute, the model runs:

→$P_R(t+1) = P_R(t) - A \times P_R(t)/[K_M+P_R(t)] - H \times P_R(t)$;

Signal(t) curves are then calculated: these are commonly reported, as the raw observables for enzymatic experiments (note, no-enzyme control data are not presented yet):

→Relative Change Soluble Signal(t) = $B_{ACQ} + P_R(t) \times B_{PR} + S \times [P_0 - P_R(t)]$

→off-ON Soluble Signal(t) = $B_{ACQ} + S \times [P_0 - P_R(t)]$

→off-ON Precipitating Signal(t) = IF $[P_0 - P_R(t) - K_S] > 0$ THEN $B_{ACQ} + S \times [P_0 - P_R(t) - K_S]$ ELSE B_{ACQ}

Signal(t)/Signal(t=0) ratios were then calculated. This is, unfortunately, a common way[40] of reporting *in vitro* signal behaviour: unfortunate because despite appearances, it **does not provide indications about the *reliability* of the signal data**. A better, *in vivo*-style calculation would instead be the ratio of enzymatic signal to *nonenzymatic control*; however this was considered in the next step:

→Error ratios E(t) = Soluble Signals(t)/$[S \times P_{TRUE}(t)]$:

These final curves at last present the *underlying reliability of the observed data*, by evaluating the correlation between the observed signal and the true amount of enzymatic processing that has occurred. It is no good having a large fluorescence signal, if it does not reflect the enzymatic mechanism one is trying to check! These are *not* experimentally observable, but only by reporting (and not simply subtracting!) no-enzyme control data, can these errors be calculated. They will reveal the S(t)/S(t=0) curve to be a totally uninformative method for reporting data or drawing conclusions.

Note that until the precipitating probe begins to precipitate, no calculation of E(t) is physically reasonable as the signal is unchanged from its value at t=0, so no calculation of E(t) is performed until precipitation starts. Once the fluorophore solubility threshold K_S has been crossed, for an observed signal corresponding to 3 µM of precipitated probe, I conclude that $(3+K_S)$ µM have been processed, so the two off-ON modes give similar results of signal generation, and identical results of error in correlating signal↔enzyme activity, *in this no-diffusion and identical-B_{ACQ} scenario, once the precipitating probe shows a signal*. In a realistic setting, once diffusion and B_{ACQ} are considered, precipitating probes have considerable further advantages.

3.1.4. Results of the numerical models

Several scenarios using 'realistic' values of the model's parameters were numerically calculated, to illustrate the behaviour of the three probes modes under the assumptions set out in section I.3.1.1. The values for A, K_M, H, P_0, S, and B_{ACQ} were kept identical for each mode to enable 'fair' comparison; recall though that this ignores likely trends of B_{ACQ}, and that the model is anyway intrinsically *unfair* towards the off-ON and especially the off-ON precipitating modes, by underestimating their relative *in vivo* sensitivity (section I.3.1.1). The Signal(t), Signal(t)/Signal(t=0) and Error(t) are graphed.

(1) *In vitro* HPQ Model: A=0.4, K_M=10, H=1×10^{-6}, P_0=100, B_{ACQ}/S=0.03, B_{PR}/S=0.05, K_S=6

These values were based closely based on my own *in vitro* experiments with HPQ probes **1a** and **1b** (K_S~7 μM, τ~4151 days→H~1×10^{-6} min^{-1}, K_M(**1a**)~15 μM, raw B_{ACQ}/S→0.03 μM for **1b** in run 2, Figure I.29). For the relative change probe, I here supposed a 20-fold signal enhancement, which is towards or beyond the upper limits seen for many relative change probes. Pleasingly, the numerical model returned results that even closely matched the absolute values of my experimental data. A and K_M here give enzymatic processing complete within 4 hours, the probe starts precipitating after 20 minutes (Figure I.9).

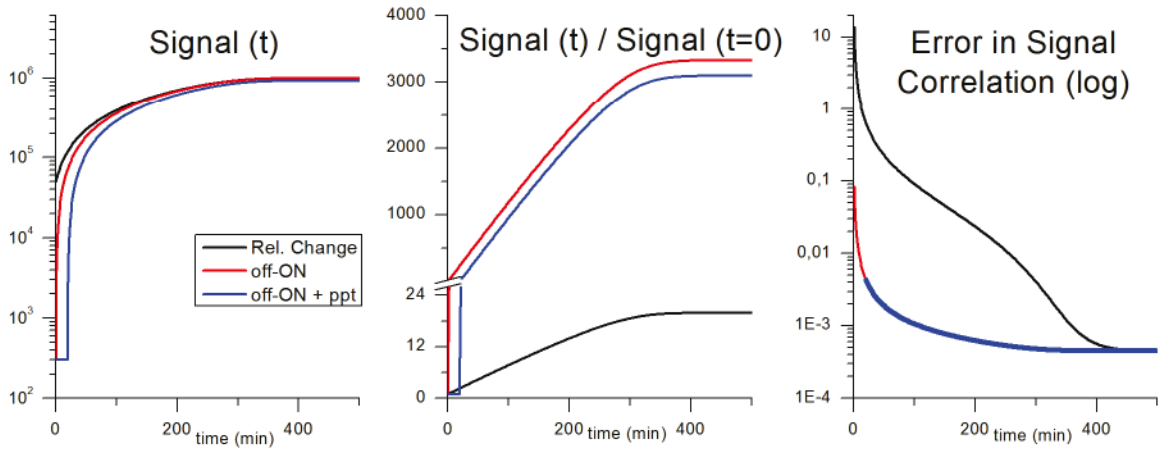


Figure I.9 – Modelling experiment 1, average values for all parameters (note the log scale of the signal timecourse). Despite outwardly similar signal curves, note the significantly lower error of correlating signal with enzyme activity for the off-ON probes, and their 120-fold greater signal enhancement.

The absolute signal observed is closely similar for all the probe modes. However, this model illustrates that the *absolute signal* is *not* the important characteristic of a probe: (1) The signal *enhancement* is *150-fold higher* for the off-ON mode probes than the relative change probe; but still more importantly, (2) the *error* of correlating observed signal with true enzyme activity is *two orders of magnitude lower* for the off-ON probes than than for the relative change probe towards the start of the reaction, although the error difference diminishes slowly over time.

Concluding that relative change probes would achieve acceptable correlation if assays were simply run for longer, is false for two reasons. (a) Mechanistically, the next example will show that even over a 24 hour *in vitro* experiment, the correlation error for the relative change probe can remain significantly greater (by eg. one order of magnitude), if enzymatic activity is low (or processing less favourable). Not even off-ON probes can always hope for low-error situations; if this order of magnitude difference turns out to lie between 20% (off-ON) and 200% (relative change) errors, then the relative change mode will be simply useless. (b) Practically, the relative change probe's Signal(t) or Signal(t)/Signal(t=0) seem to show plateau approaching by 200 minutes, and certainly near-plateau by 300 minutes. However, the correlation error at 200 minutes is still 30 times higher than for the off-ON mode; at 300 minutes the error is still 10 times higher. Therefore the relative change mode is a *far worse gauge of enzymatic activity*; and simply examining its experimentally-determinable signal does

not even indicate where it can hope to start providing more *reliable* quantification of enzymatic activity: a let-down for kinetic analysis.

The error curves also introduce a crucial aspect of the precipitating off-ON mode: it prevents unreliable signal assignment at the start of the reaction, because **as soon as a signal is observed (20 min), the maximum correlation error is already low** (0.4%); whereas in these comparatively favourable assay settings, the error at the start of the relative change probe's timecourse is 1400%, and that of the soluble off-ON probe is 8%. The relative change probe has to wait for 150 min before attaining the same reliability as the precipitating probe does at the first moment of signal generation!

Still, because of the relatively **large enzymatic activity** used in this model, **all probes eventually show acceptable plateau behaviour** (depending on the standards of the scientist), so if one is willing to sacrifice obtaining reliable kinetic data, then the relative change probe may be allowable.

(2) *In vitro* diClHPQ Model: A=0.004, K_M=10, H=2×10^{-5}, P_0=25, B_{ACQ}/S=0.005, B_{PR}/S=0.1, K_S=1.5

These values were chosen to simulate the use of **15a**, featuring the more insoluble (K_S~1.5 μM) but more unstable (τ~567 h→H~2×10^{-5} min^{-1}) **Cl_2HPQ** core. I used a **hundred-fold lower enzyme concentration**, and a lower probe concentration, examined over a longer time, under better signal collection conditions (lower B_{ACQ}, eg. relying on the very large Stokes shift of this solid-state fluorophore, and longer acquisition times): see Figure I.10. The ten-fold signal enhancement posited for a relative change probe is a standard figure[30].

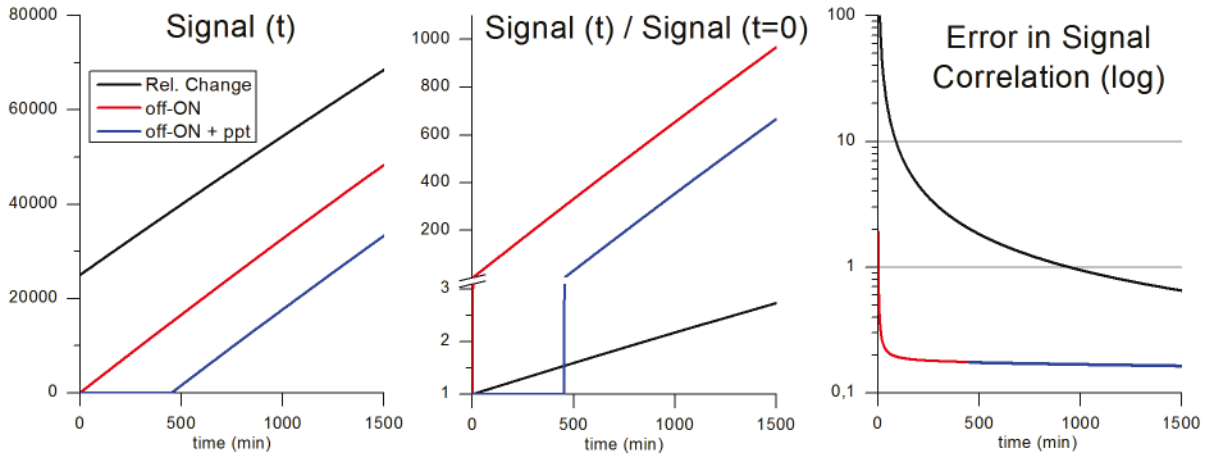


Figure I.10 - Modelling experiment 2, with low enzyme activity and lower-solubility Cl_2HPQ probe. The error of correlating signal with enzyme activity is drastically lower for the off-ON probes, which also have 300-fold greater signal enhancement. Note the maximal error values: 876 (rel. change), 1.93 (off-ON) and 0.175 (off-ON + ppt).

Although the *absolute* signal seen for the relative change probe is now significantly the largest, its signal enhancement is still a factor of 300 *less* than for the off-ON probes. Worse, the error in its signal assignment remains above 100% for the first 16 hours of the assay, and is only 67% even by the end of the reaction (24 h). For the precipitating probe, the **error in signal assignment is only 17%** as soon as precipitation starts, after 7.5 hours. This 'safety net' is of vital importance for reliably quantifying small enzymatic activity, since **with low enzymatic activity**, neither the signal nor the

signal enhancement curves of soluble fluorophores give *any clues* by their form about when it is reliable to start correlating signal seen with causative enzymatic activity (Figure I.10, note the maximal error values). Only the precipitating mode offers an indication, by its signal onset; and although the off-ON soluble mode does rapidly offers a low-error regime, this is not always the case, so simply waiting for an hour or two cannot be relied upon as a strategy for obtaining better reliability (see following).

(3) *In vivo* NIR Model: A=0.0002, K_M=10, H=1×10^{-6}, P_0=10, B_{ACQ}/S=0.1, B_{PR}/S=0.02, K_S=0.2

This model argues that even superlative relative change probes (here, a fifty-fold enhancement of signal) have generally insufficient performance for reliable and quantitative *in vivo* work with very low enzyme concentrations, even if a very low rate of spontaneous hydrolysis and a low signal background can be obtained (I consider that an *in vivo* study with a background signal corresponding to only 0.1 µM probe would be an exceptionally sensitive scenario; whether this truly can be obtained for lower-Stokes-shift soluble fluorophores is another a separate question). It shows however that off-ON probes, and *in particular* precipitating probes, may provide a robust solution to both quantitation and reliability, especially if a relatively low-K_S fluorophore is used (value based on estimate for NIR-emitting **BTQC**): see Figure I.11.

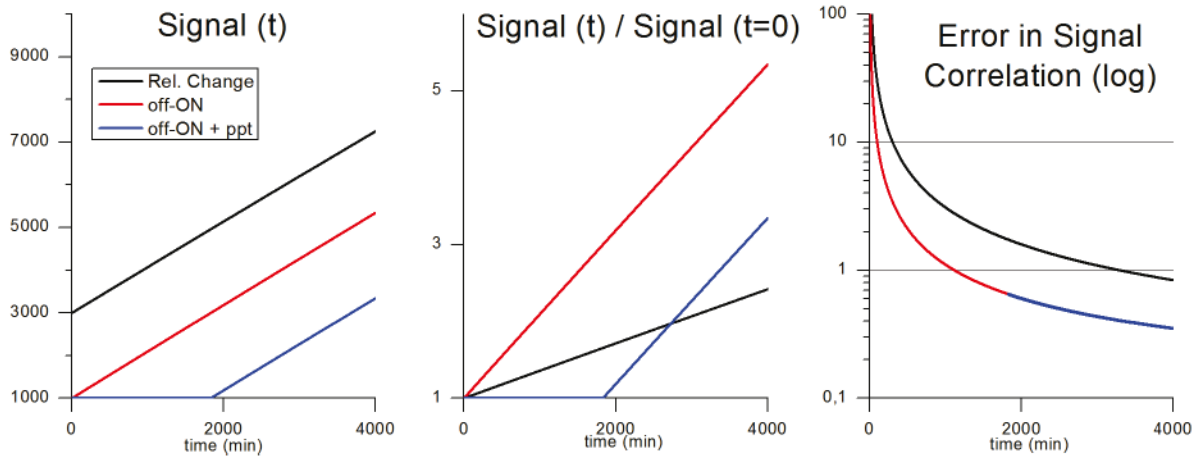


Figure I.11 - Modelling experiment 3, with very low, *in vivo*-style enzyme activity, supposing probes operating in the NIR with consequently low acquisition background.

Despite similar signal forms, but now also very similar absolute signal enhancements, the percentage error of correlating signal with enzyme activity is still drastically lower for the off-ON probes. Indeed, if the off-ON error timecourse is bleak (>100% error for the first 24 h), then that of the relative change probe is all but useless. Yet the off-ON precipiation threshold again shows when more reliable signal correlation begins, while neither soluble mode can indicate this based on their signal or signal enhancement timecourses. I conclude that with low enzyme activity, relative change mode probes are simply unsuitable for achieving high sensitivity, even if they can boast of otherwise impressive low intrinsic background, or high signal enhancement. By contrast, the precipitating off-ON mode gives excellent reliability of enzyme detection under these challenging conditions.

(4) Colony Assay Model: $A=1\times10^{-3}$, $K_M=10$, $H=1\times10^{-4}$, $P_0=30$, $B_{ACQ}/S=1$, $B_{PR}/S=0.05$, $K_S=1$

This experiment was intended to show the performance of the probe modes when applied to a different detection aim: deciding which cells in a colony express or do not express an enzyme of interest, in challenging conditions: **when the activity of that enzyme is actually lower than the background hydrolysis rate**. I set the initial rate of enzymatic processing as one quarter of the spontaneous hydrolysis rate, and compared samples *with* enzymatic activity to controls with spontaneous hydrolysis only. Effectively, this hydrolytic instability means that the *microscopically* off-ON soluble probe is *macroscopically* a relative change probe (see section I.2.1): however, the **precipitating off-ON probe retains a window of true macroscopically binary off-ON detection** (Figure I.12).

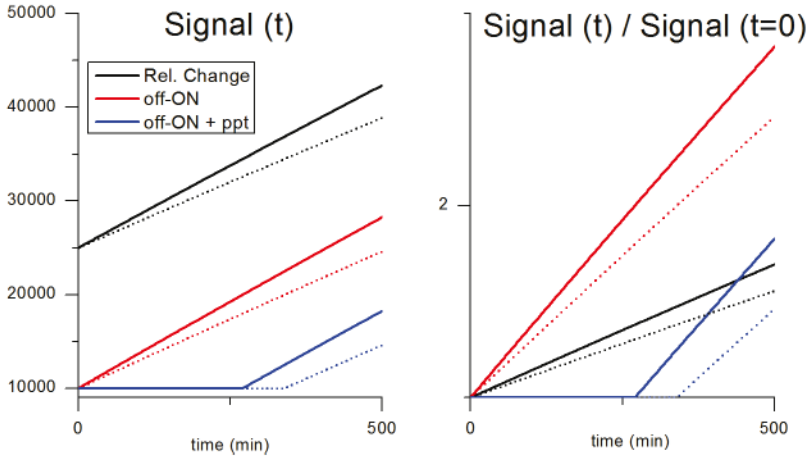


Figure I.12 - Modelling experiment 4, with spontaneous hydrolysis occurring four times faster than enzymatic processing (solid lines). Reliable discrimination against samples without enzyme at all (dotted lines) is only possible to a limited extent for the precipitating probe model, exploiting the binary detection window (275-325 min).

With significantly more signal being generated by nonspecific hydrolysis than by localised enzymatic activity, one would guess that detection would be impossible, and indeed, almost nothing differentiates between the enzymatic and non-enzymatic signals for the soluble probes of both modes. Therefore both soluble probes generate only a (low) relative change in signal, and so give unreliable detection.

However, the precipitating probe offers some sort of solution: the delay time before precipitation starts is significantly shorter for the enzymatic sample than for the hydrolysis-only sample. Therefore, at the moment when the first enzyme-positive colonies generate a signal, the other colonies will still show no signal whatsoever for a further hour. This is an ample **window to take a binary off/ON image of the active colonies** and so unambiguously distinguish them. This is not an ideal solution; and the colony must be regularly imaged to ensure that images are taken during this window: but it is by far the best option, and offers to differentiate between colonies with otherwise unimaginable binary sensitivity.

Note that the slower the rate of hydrolysis relative to enzymatic activity, the larger is this binary decision window: at limits **where the probes are very stable** (ie. this thesis' probe designs), the **binary decision window covers essentially the whole experimental timeline**. An implicit factor which is also crucial to keep this window as wide as possible is the **design of well-recognised**

spacers, which give high K_M values to their probes, as this drastically enhances the rate of enzymatic processing, with a given probe concentration, especially in a competitive *in vivo* situation.

This is the basis for my conclusions that ***in practice, a hydrolytically robust, molecularly off-ON precipitating design which is well processed enzymatically provides a unique, macroscopically off-ON, binary system for the reliable detection and quantitation of enzymatic activity in a range of settings from* in vitro *to* in vivo**. In my enzymatic experiments (section I.5), molecular off-ON behaviour was confirmed by measuring negligeable signal for the proagent state compared to high signal for the enzyme-activated agent; hydrolytic robustness was confirmed by monitoring the absence of spontaneous signal generation for more than 8 hours and by calculating halflives explicitly; and good enzymatic processing was confirmed by the rapid attainment of the signal plateau and by the explicit deconvolution and calculation of favourable enzyme kinetic data. The prediction from these modelling experiments, that these probes should therefore provide a ***macroscopically off-ON binary detection system***, was also amply confirmed in practice *in vitro* by comparing kinetic test behaviour.

3.1.5. Conclusions from the sensitivity model

The greatest *mathematical* sensitivity advantage of using a precipitating rather than soluble off-ON probe, is the possibility of much more confidently selecting samples with a small amount of enzymatic activity. When the enzymatic activity is on a comparable or lower level than the rate of degradation by spontaneous hydrolysis, *only* the precipitating mode allows for an observation window where the macroscopic signal response is still binary: *off* for the nonenzymatic case, and *ON* for the sample with an added enzymatic component.

The mathematical advantage of using either off-ON mode probe, as against even the most sensitive relative change probe (100-fold signal enhancement), is in their far lower error of signal↔enzyme correlation when the enzymatic activity is small. As *in vivo* imaging must monitor very small enzymatic activities, the conclusion is that off-ON probes are vastly superior for this application, and precipitating probes especially interesting from the mathematical point of view of reliability.

For precipitating probes, the moment of precipitation acts as a binary switch for macroscopic signal observation. During the delay time while probe is processed/hydrolysed but not yet precipitated, there is no signal, so there is no false assignment of enzyme activity; whereas the timepoint when the precipitating probe first shows a signal confidently marks the time when that observed signal can be correlated to an accumulated amount of enzymatically processed probe (error usually under 40% under my conditions). By contrast, off-ON soluble probes continuously generate signal and it is unclear from the signal generation curve exactly where the observed signal can reliably be correlated to enzymatic activity. The soluble probe error curve overlays the precipitating probe error curve at longer timescales once precipitation has begun, but if the signal↔enzyme correlation is performed at the very start of the reaction with soluble probes, the error can be very high (>1200%). It is the readily noticeable onset (precipitate appears) of this low-error regime which distinguishes the precipitating

mode, and acts like a 'safety catch' against false signal interpretation, which is fortunate as it is in exactly these low-enzyme-activity situations that the potential unreliability of the correlation becomes critical.

3.1.6. A crucial note on effective *solubility constants*

Recall that when the fluorophore solubility threshold K_S has been crossed, the two modes (off-ON soluble and off-ON insoluble) give similar results of signal generation, and identical results of error in correlating signal↔enzyme activity. Therefore the more the *effective* K_S tends to zero, the *earlier* the models converge, in terms of enzymatically processed probe concentration.

In vitro, **solubility constants for a given fluorophore are commonly juggled in two ways:**

(1) By **pre-spiking** the test solution with a (final) concentration X of the hydrolysed probe, only (K_S-X) of the applied probe needs to be hydrolysed before precipitation begins. Invitrogen cleverly takes full advantage of this[17,88], as they supply their ELF-97 solution with a small filter so that the **diClHPQ** which continuously precipitates can be filtered out just before using the solution; then only a small extra amount of probe hydrolysis will generate the precipitate and so signal is rapidly seen[17].

(2) **Using a cosolvent** (eg. diluting a DMSO solution of the probe which is also saturated in fluorophore) may also allow greater pre-spiking than would be possible by diluting a filtered, saturated aqueous solution of the probe+fluorophore, so achieving even lower effective K_S. Conversely, adding a pure cosolvent (as is often done with insoluble probes for *in vitro* characterisation[30,89]) substantially delays precipitation by solubilising the generated fluorophore, thus increasing the effective K_S.

Pre-spiking or cosolvents may be advantageously used to adapt the natural K_S of a probe to a given situation: eg., if the absolute enzyme activity is low, then the smaller the effective K_S should be to achieve rapid yet reliable detection (as long as the probe is hydrolytically stable). However, this depends on ensuring that reliable pre-spiking is possible. Both in the model and in my practical tests I avoided cosolvents and pre-spiking because in my consideration **pre-spiking/cosolvents misrepresent what is reliably possible *in vivo***: their use is restricted to *in vitro*, or tissue slices only: otherwise, even *in cellulo*, the different biodistribution kinetics of the probe and the pre-spiking fluorophore will become problematic and no uniform or predictable effective K_S can be achieved. Cosolvents may also have adverse effects on enzymatic processing and crystallisation pattern/kinetics.

Detailed considerations of K_S are directly relevant to designing the most sensitive assays, and obtaining the most extensible results. For the experiments I performed, with K_M fixed by the substrate-spacer unit and with enzymatic processing a rate-limiting factor for generating enough fluorophore to precipitate, the only way to ensure faster detection with a higher signal-to-control ratio, was to increase P_0 or lower K_S. Increasing P_0 is undesirable biologically (requires substantial probe injection, incurs toxicity risks) as well as mathematically: it poses a greater risk of passing into the strong-absorption domain, and so giving a poorer quantitative correlation of signal with enzymatic activity

(see section I.3.3.3). Instead, choosing a fluorophore with lower *intrinsic* K_S (**diClHPQ**) gave both earlier signal generation and far higher signal-to-control ratio than the corresponding **HPQ** probe, when both were used at a low concentration (10 μM) to avoid the strong-absorption regime (subsequently confirmed). In all realistic non-pre-spiking situations, **diClHPQ** is therefore the preferred of these two fluorophores. However if pre-spiking is possible, so that both probes only require the processing of the *same* concentration of probe to achieve a signal (eg. 0.5 μM before precipitation begins), then with the effective K_S equalised, the far greater **hydrolytic stability** of **HPQ**-based probes gives them more sensitive detection: they can be used over a longer time (or at higher concentration), without causing hydrolytic signal generation.

3.2. Sensitivity↔Stokes shift model

My next model considered the effect of the fluorophore's Stokes shift $\Delta\lambda_P$ on the sensitivity obtainable in an *in vivo* setting. Qualitatively, it is well known that the higher the Stokes shift, the better the sensitivity[1]. However, I considered that the ~quantitative dependency of sensitivity on Stokes shift was important to evaluate, since I did not find any explicit treatment in the literature, particularly one that would be appropriate for the present high-$\Delta\lambda_P$ probes. I was actually surprised by the magnitude of the dependency, which this model shows is roughly *exponential*, parametrised to the background autofluorescence Stokes shift $\Delta\lambda_A$. This will illustrate why even a mechanistically off-ON probe may be all but useless in an *in vivo* setting unless the fluorophore presents a sufficiently high Stokes shift. This is another powerful argument for using high-$\Delta\lambda_P$ **HPQ**-based fluorophores, rather than common *in vitro* workhorses such as fluorescein and resorufin, even if those more common fluorophores were to be entirely quenched in the proagent (which usually they are not, see section I.1.3).

3.2.1. Stokes shift considerations: introduction

For off-ON fluorogenic probes *in vivo* (B_{PR}=0) the main unavoidable background signal component in B_{ACQ} will be tissue autofluorescence. This has a complex excitation-dependent form[90], but a moderate Stokes shift $\Delta\lambda_A$ of around 30-45 nm. Even *in cellulo*, autofluorescence may complicate images to a lesser degree[91]. Autofluorescence usually has a 'normal' pattern, ie a tail of ~100 nm, similar in shape to the fluorescein spectrum shown below (Figure I.13); I take the form of this spectrum as a model for the form of *in vivo* autofluorescence, but adjusting its Stokes shift to $\Delta\lambda_A$=40 nm. In the model I assume that all fluorophores will likewise give this normal signal envelope, just with different $\Delta\lambda_P$.

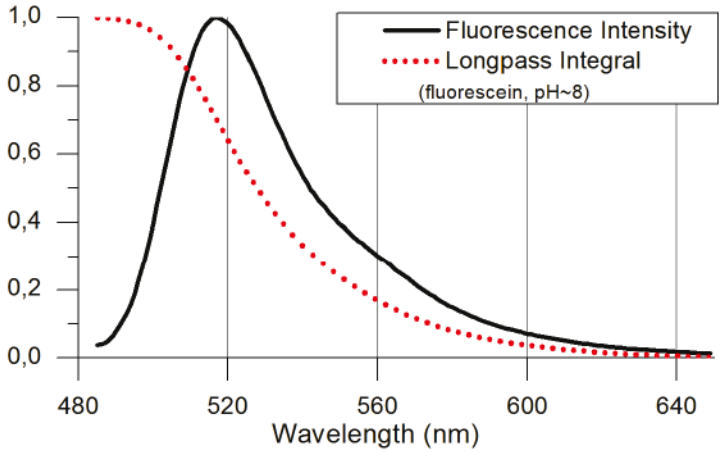


Figure I.13 – Black solid line: fluorescein emission spectra (493 nm excitation) at pH=8[92]; Stokes shift $\Delta\lambda_P$=24 nm. Red dotted line: longpass integral $\Phi_L(\lambda)$.

Detection sensitivity must pass two thresholds. First, the signal to background ratio *in the collected signal* must be above a *reliability threshold* so that the signal actually reflects enzymatic activity not background noise; secondly, the total *amount* of collected signal must be high enough to pass the *experimental detection threshold*, eg. background noise fluctuations, machine insensitivity. Optimising detection sensitivity requires addressing both points; high reliability is useless if no signal is actually observed above the background noise levels; equally, a large but unreliable signal serves no purpose.

(a) First consider a standard soluble fluorophore with low $\Delta\lambda_P$, eg. fluorescein itself ($\Delta\lambda_P$=24 nm), such that $\Delta\lambda_P<\Delta\lambda_A$. Because its signal intensity drops faster with wavelength than that of the autofluorescence, the best-sensitivity signal must be collected from a narrow window around the fluorophore's emission maximum $\lambda_{em,max}$ where the signal/autofluorescence ratio is most favourable. Therefore, the total signal *actually collected* [integral of signal(λ) in the observation window] will be small, ie. the experiment will give low *detection threshold* sensitivity, even if the fluorophore has a large *total* quantum yield Φ_{TOT} [integral of signal(λ) over the whole spectrum]. Worse, if only a small amount of the fluorophore is activated, there is a risk that the absolute amount of autofluorescence in this window is stronger than the activated probe's signal; therefore an enzymatic signal may actually be impossible to distinguish (low *reliability threshold*).

(b) Now consider a fluorophore with $\Delta\lambda_P=\Delta\lambda_A$. Whatever light is collected will therefore contain both probe signal and autofluorescence in a constant ratio, so the autofluorescence background B_{ACQ} will be relatively high. The overall sensitivity will depend entirely on the amount of probe that can be activated to drown out the background. Therefore reliability will be low when small amounts of probe are converted, as per the description of relative change probes (see section I.1.3.2); but at least the whole wavelength range of the emission signal can be collected (detection threshold).

(c) If however $\Delta\lambda_P>\Delta\lambda_A$, we may apply a longpass filter for a certain λ onwards, which blocks more of the shorter-wavelength autofluorescence than the longer-λ probe fluorescence (let $\Delta\lambda_P$-$\Delta\lambda_A$=**ΔStokes**). Therefore B_{ACQ} is much reduced, and we obtain a *much* higher **(λ-longpass)-signal-to-background**

ratio S*(λ) in the collected signal than in the previous cases (greater reliability), even though some of the total signal has been cut out. Sensitivity will then depend on passing the detection threshold; **$\Phi_L(\lambda)$**, the fraction of total possible probe signal which is longpass-collected from λ onwards, will measure of the probe's **signal intensity collected**. Therefore **S*(λ) considers *reliability*, while $\Phi_L(\lambda)$ treats *detection*; both must be as high as possible for good overall sensitivity**.

3.2.2. Stokes shift↔sensitivity model

Having assumed that probe and autofluorescence emission spectra are identical in shape, just shifted from each other by amount ΔStokes, I now focus on evaluating the dependency on ΔStokes of S*(λ) and $\Phi_L(\lambda)$ across the wavelength range λ. As there is no basis to decide on the *absolute* amount of probe activated relative to the autofluorescence background, the calculation assumes $\Phi_P=\Phi_A$, ie considering the situation where the total amount of autofluorescence emission is the same as the total amount of fluorophore emission (disregarding the amounts of each actually collected): if more, or less, mol of probe is activated, then sensitivity factor S*(λ) will be scaled accordingly (*linear* scaling only).

Translating these results to the case of different fluorophores activated to the same molecular amount is also important. This too requires *linearly* scaling the S*(λ) by their relative total signal intensities Φ_{TOT}. However, the dependency of S*(λ) on ΔStokes will be seen to be orders of magnitude beyond these potential linear scalings in Φ_{TOT}, since molecules which are considered even to be fluorophores already have good values of Φ_{TOT}, perhaps with variations of usually only a factor of ten or twenty, while S*(λ) varies by a factor of >1600 for the range of fluorophore ΔStokes values considered (ESIPT like **diClHPQ/BTQC**, vs normal like fluorescein/resorufin). Therefore by post-scaling the results appropriately, this model can provide *semiquantitative information* about the S*(λ) that may truly be observed with a range of fluorophores.

I therefore calculated $\Phi_L(\lambda)$ for different values of ΔStokes by longpass integration of the signal intensity, and S*(λ) by taking the instantaneous ratio of the probe $\Phi_L(\lambda)$ to the $\Phi_L(\lambda)$ of autofluorescence (Figure I.14). Note that for (a) ΔStokes<0, S*(λ)<<1 for all λ (cf. fluorescein); and for (b) ΔStokes=0, S*(λ)=1 for all λ. Only results from case (c) where ΔStokes>0 are shown.

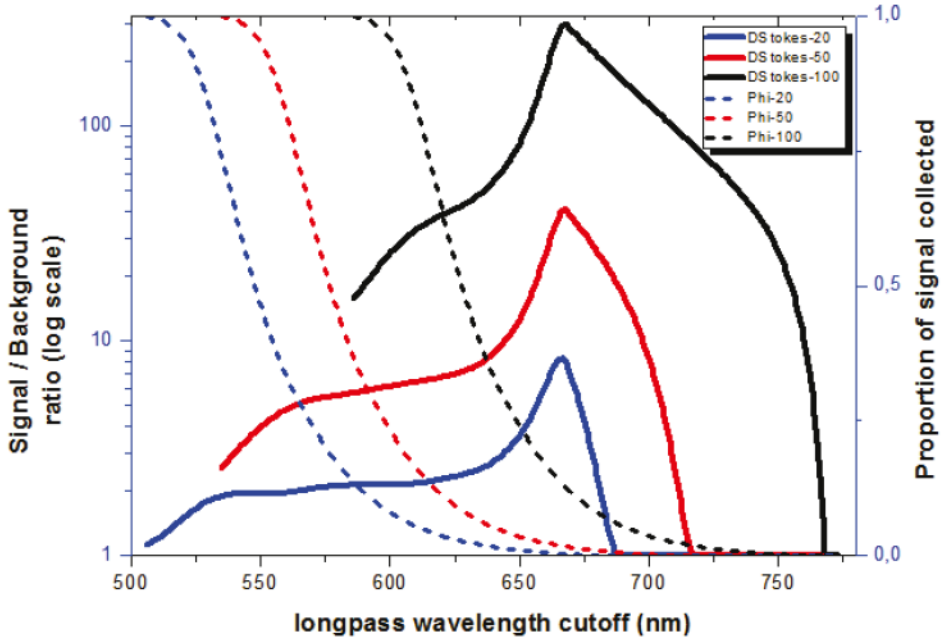


Figure I.14 – Signal-to-background S*(λ) (logarithmic left axis, solid lines) based on collected longpass intensities $\Phi_L(\lambda)$ (right axis, dotted lines) for three different ΔStokes values: 20, 50 and 100 nm. The empirical relationship for $S^*(\lambda_{max})$ is roughly $S^*(\lambda_{max})=k\times e^{(\Delta\lambda_P/\Delta\lambda_A)}$.

Firstly, S*(λ) does not increase indefinitely but is maximised, here at a wavelength λ_{max}=670 nm, which is set by the ***intrinsic background spectrum*** (autofluorescence + noise) and is independent of the probe's ΔStokes. This is a physical effect: λ_{max} is where the autofluorescence signal becomes so low that it is 'lost' in the noise, so longpass signal from the probe is superimposed purely on this background noise. This reflects the most sensitive possible detection scenario (I have included a constant level of noise in both probe and autofluorescence spectra).

Then, *reliability*, as per **$S^*(\lambda_{max})$, almost *exponentially* depends on the probe's stokes shift** (scale of S*(λ) is logarithmic), by shifting the Φ_L so that more signal can be collected around λ_{max}. The empirical relationship seen is around $S^*(\lambda_{max})=k\times e^{(\Delta\lambda_P/\Delta\lambda_A)}$ for this simplified model (I did not find any such considerations or treatment in the literature to compare it to). For example, S*(670) = 8 for ΔStokes = 20 nm (note: implies $\Delta\lambda_P$=60 nm); for ΔStokes = 50 nm, S*(670) = 40; and for ΔStokes = 150 nm (cf. **diClHPQ**, $\Delta\lambda_P$~185 nm) the value is 1600 (off scale so not shown on image).

3.2.3. Stokes shift+Sensitivity combined conclusions

Now reconsider the model of section I.3.1.2. The absolute value of the parameter **B_{ACQ}/S depends on both the Stokes shift and the fluorophore's brightness**. S/B_{ACQ} is inherently given by S*(λ), ie an exponential dependency on the Stokes shift. For intercomparison between different fluorophores, S must also be scaled linearly with the total collectable signal quantity, Φ_{TOT}, which is a common argument for using the brightest possible fluorophore. However, the far greater S*(λ) dependency shows that actually, **Stokes shift is much more important for detection sensitivity in a high-background, small-signal *in vivo* setting than is total fluorophore brightness.** Therefore, for *in*

vivo applications, it is far better to choose a fluorophore with large Stokes shift and thus gain exponentially in B_{ACQ}/S, rather than choosing large Φ and gaining only linearly.

Why might this calculation not be easily found in the literature? I believe that the reason is similar to that of why the advantages of (novel) *precipitating off-ON* probes, vs standard soluble relative change probes, are not explicitly treated either: using such probes has simply not been widely possible before. Prior probes face different (and more constrictive) limits on their sensitivity (*n*, diffusion, photobleaching, etc), and almost all have similar Stokes shifts anyway (<50 nm, max advantage ~factor of 3), so perhaps no-one has been overly interested in pursuing the Stokes shift to extremes (>180 nm, max advantage >factor of 1600). In the present case, the novelty is the use of *off-ON* fluorophores with Stokes shifts in an entirely different *regime* to that of soluble fluorophores, as applied to *ultrasensitive* detection of *small, specific enzyme activities* for *in vivo* use.

For example, consider that if the B_{PR} component of B is large relative to B_{ACQ} (relative change probes) then the dependency of the background signal on ΔStokes may be irrelevant; relative change probes benefit *less* from choosing fluorophores with high Stokes shifts than do off-ON probes. Since prior art for *in vivo* targeted peptidase probes, *entirely* concerns relative change probes, it seems logical that this prior work would therefore neglect a consideration which was not the largest limiting factor. Note however that the relative independency of relative change probe sensitivity from Stokes shift *in vivo* is not an advantage: in fact it reflects the inherently lower sensitivity that the relative change mode may attain. Only for truly *ultra*sensitive probes, such as the present work, do such factors start to be important, since the probes progress towards different theoretical and practical sensitivity limitations.

3.3. Signal from precipitating fluorophores

The macroscopic observable seen in enzymatic experiments is the fluorescence signal over time. This should be explicitly treated with standard signal processing algorithms to compare a probe's macroscopic signal behaviour with that from other systems under similar conditions and data treatment: since commonly as an end-user one would like to compare eg. the macroscopic signal-to-control ratios obtained with different probe systems, under the desired assay conditions. The more *in vivo*-realistic are the assay conditions/treatment, the better the conclusions will apply to other settings.

However, only a full, microscopic understanding of the system can yield the molecular information that is required for predictive and analytical description/comparison of the probe system's performance in different settings. As the present system is so far removed from prior art, I considered it vital to evaluate, in particular, the relative molecular kinetics of enzymatic processing, spacer cyclisation, precipitation, and spontaneous hydrolysis: which is a far greater challenge than simply reporting macroscopic data. Therefore it was crucial to reliably relate the macroscopic signal data with its microscopic cause (precipitated fluorophore): yet this is *not* as straightforward as it might seem.

Homogeneous systems (ie. using soluble fluorophores as is standard in cell biology, such as fluorescein) at low concentration obey the Beer-Lambert law, and only require that the volume of test solution contributing to the collected fluorescence be identical between different runs to be able to reliably compare data between them: this can be assured by fixing the cuvette size and shape. However, as **HPQ**s are precipitating fluorophores, I considered that there were two major complications with the observed fluorescence signal stemming from inhomogeneity, which I wished to consider so as to design the best experiments, and extract the most reliable data from them.

3.3.1. A practical problem of precipitating fluorophores ↔ wellplate choice

As we use a solid-state fluorophore which forms a solid deposit on the bottom of its container, measurements cannot be performed in a cuvette fluorimeter as for soluble fluorophores (sideways illumination), but must be performed so that the excitation beam hits the precipitate as much and as evenly as possible; the cheapest equipment for this is a top-down microplate reader. The first complication is then a practical one: if the precipitate concentrates in some areas, then illuminating a standard part of the well sometimes hits, and sometimes misses these concentrated areas (Figure I.15d), so the deviation between otherwise identical experiments and timepoints may be large.

This was observed in an unexpected way with the initially-tested wellplates, (500 µL NUNCLON 96-well plates): precipitated **HPQ** first appeared at the well rim (triple interface between the solution's surface, the well's hydrophobic plastic and the air; see Figure III.10), where the excitation beam does not strike. Therefore this initial precipitate did not generate signal data. This effect was significant: there was initially approx. 30 min signal onset delay time; however in identical biochemical conditions but on the final wellplate design, the delay was only <10 min! I speculate that this rim problem may arise because of evaporation at the surface, which results in a higher local concentration of **HPQ** than in the bulk, and because the wells are smooth-walled, so the dust on the water's surface provides the best initial nucleation sites, from where surface tension pulls the precipitate to the walls. Only after more **HPQ** had been formed did **HPQ** begin to accumulate on the bottom of the wells (see Figure III.10). When on the bottom of those plates, the fallen particles were shaken (by machine vibration) into the central (illuminated) portion of the well, because those plates have a convex bottom. Yet, even vigorous well-plate shaking regimes did not detach precipitate around the rims, nor did using a micropipette to aspire and release the well contents several times (eg. Figure I.15c, vs the ideal case of eg. Figure I.15b).

I then tested four other wellplate types with different surfaces, angles, depths and widths, each of which gave different behaviour. For example, this rim problem was present but to what seemed a lesser degree with Nunclon 110 µL 384-well plates, where I observed a slight rim, but a brighter central speck (the plates also had a convex bottom where the precipitate initially accumulates).

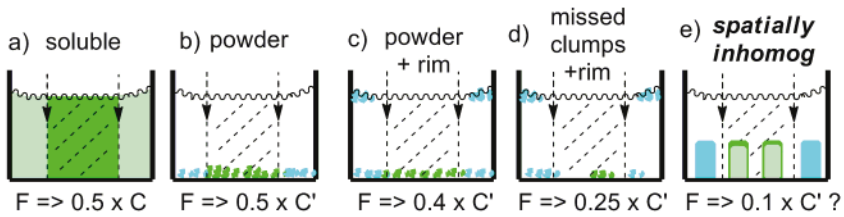


Figure I.15 – Different scenarios of spatial inhomogeneity discussed in the text. Green represents excited fluorophore; blue represents fluorophore missed by the excitation beam; grey represents unexcitable fluorophore.

This spatial inhomogeneity gives a practical problem which soluble probes do not have to consider: which shape, brand and size of wellplate for which application? For scanning experiments I will show that shallow flat-bottomed wellplates seemed vital for reasons of optical path, but only if the whole of the well-bottom+rim could be excited. For kinetic experiments where it was not possible to cover each well by scanning, but where rapid reading of the greatest possible proportion of precipitate was desired, I preferred wells that did not generate a rim, that were thin so the excitation beam could cover as much of the well bottom as possible, and that had a pointed bottom so the crucial initial deposition could be reliably acquired. These aspects of experimental design are addressed again in section I.5.1.2.

3.3.2. Theoretical problem ↔ spatial inhomogeneity modelling

The second complication is potentially more serious: there is a theoretical problem of spatial inhomogeneity, ie. that a fluorescence signal readout F cannot always be converted to a "concentration of precipitated probe" C_{PPT} by a Beer-Lambert law such as $F=kC_{PPT}$, if the fluorophore is a solid.

To apply the Beer-Lambert law *macroscopically* requires a uniform precipitate, so the irradiated area reliably correlates to a constant mole fraction of the total precipitated fluorophore. Explicitly, this means a powder with uniform (vertical) optical path lengths (effectively in mol; Figure I.15b). The optical path length must be identical for all points on the well bottom, at any point in time during the experiment. Then the scenario approaches that for a soluble fluorophore, except that the solid-state fluorophore is effectively concentrated into a small (solid) layer, not a large (solution) layer.

The path length increases the more probe is hydrolysed: so this already requires a flat well bottom. For example, if the bottom is convex, the optical path length is larger in the centre and smaller outside. Therefore the fluorescence signal per unit of fluorophore will decrease over time, since the signal is strongest from the surface layer of fluorophore if the fluorophore is a strong absorber, and the surface layer grows more slowly the bigger the conic radius (amount precipitated) becomes.

Generally however, the precipitate will not be a homogenous powder layer, even on a flat-bottomed plate, but may for example form clumped crystals. I therefore considered the influence of spatial inhomogeneity on the signal to conclude on how it could be justified to correlate fluorescence signal F with a precipitated concentration C_{PPT} resulting from enzymatic activity - as is indeed the goal of using a probe; only then could I consider how raw or processed data from these probes might be compared to results from soluble fluorophores. I also wanted to see how the results from different

fluorophores could be compared (they have different crystallinities: eg, **HPQ** has seemingly larger crystallinity than **diClHPQ**). Finding no suitable theoretical treatment, I constructed my own model.

Consider a sample of 'relatively large crystals,' of absorption coefficient ε, total volume M, uniform size and shape, parallel sides orientated vertically, and total cross-sectional area A in a well of area W. **A/W** is an effective "spatial inhomogeneity" factor: the closer (A/W) is to zero, the greater is the area-averaged spatial inhomogeneity of fluorophore optical path length. Compare these crystals' observed fluorescence signal to that collected from a fine powder also of volume M but distributed evenly across the bottom of the whole well, so A/W=1 (Figure I.16).

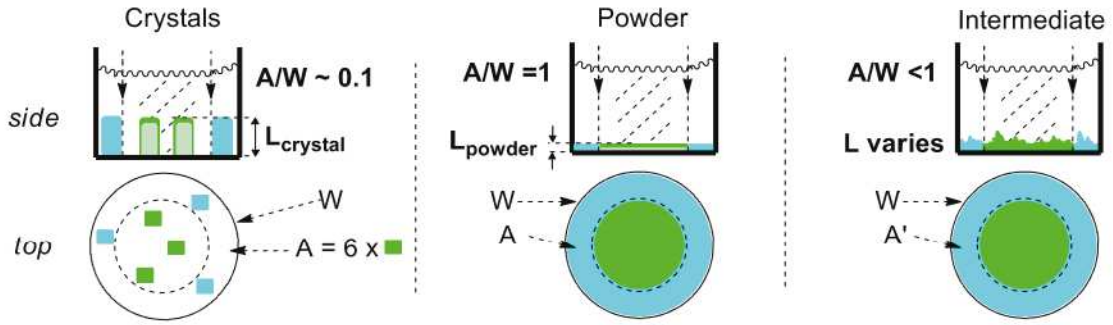


Figure I.16 – Parameters of the limiting case models for comparing homogenous (powder) vs. inhomogenous (crystals) signals, and the expected intermediate case of inhomogenously distributed clumped powder. Green represents excited fluorophore; blue represents fluorophore missed by the excitation beam; grey represents unexcitable fluorophore.

The fluorescence signals observed are proportional to the total cross-sectional area of the emitter times the integral of the pointwise excitation intensity throughout the emitter. As excitation intensity I(L) decreases the longer the optical path length by the Beer-Lambert law (note: applied on a microscopic, not the macroscopic, scale), the fluorescence signal will be proportional to the surface area times the integral from 0 to L of I(L), which has the form $\mathbf{I(L)=I_0\times10^{-\varepsilon L}}$. The total amounts of precipitate are the same, but the surface areas are different, so the path lengths L are also different (volumes of precipitate are equal: $WL_{powder} = AL_{crystal} = M$). Therefore the **signal observed for the powder is greater by a factor of** $\mathbf{(W/A)\times(1\text{-}10^{-\varepsilon M/W})/(1\text{-}10^{-\varepsilon M/A})}$. This is perhaps alluded to briefly by Invitrogen in their manual for the ELF-97 assay[88] (**diClHPQ** phosphatase probe), in saying that a fine fluorescent precipitate is best, and large grains are undesirable. Note that ε is always large for the present probes, because the absorption coefficients for both **diClHPQ** and **HPQ** are high (estimated equivalent to ~40,000 $M^{-1}cm^{-1}$ for a soluble probe). Also note that for intermediate inhomogeneity cases which cover approximately the entire well (A'=W), but do not have a homogenous distribution of optical path length L (clumped powder/crystals), an effective inhomogeneity A/W is still relevant: (1-A/W) is roughly the expectation value of the error ratio between L at any point and the average L.

Therefore I derived a simplistic mathematical prediction: as ε is large, then a linear correlation between the observed fluorescence signal F and concentration *truly* enzymatically processed and precipitated C_{PPT} will be an underestimate whenever the effective inhomogeneity A/W<1. If the precipitate is inhomogenously distributed, then the linear correlation will increasingly underestimate

the true C_{PPT} as the amount of precipitate M increases (ie. as signal increases). So, as all inhomogenous powder cases result in low A/W, a linear relationship would be considered unlikely for eg. samples with large crystals (A/W approaches 0), or wells with non-flat bottoms (A/W changes during the experiment, since the surface area depends on C_{PPT}). Note that this nonlinearity is *not* due to absorption of the fluorescence emission by neighbouring molecules: in the **HPQs**, the large Stokes shift ensures that no autoquenching takes place. Also, this problem arises even if any large crystals are *always* 'hit' by excitation light, ie the theoretical inhomogeneity problem is independent of the practical one described earlier.

The practical upshot of *this* theoretical problem is that I anticipated problems with linear signal↔concentration correlations in general, particularly at the smallest values (only a few crystals, so large inhomogeneity), but potentially at the highest values too (large M, so if inhomogenous, then bad correlation). I also planned that if the dependency of signal per unit precipitate appeared too strong, I would use more shallow layers of test solution, so that a thinner layer of precipitate would be generated for a given hydrolysed concentration (lower M, so less problems if inhomogenous).

3.3.3. Final problem and conclusions on the signal↔concentration relation

A third problem with the signal↔concentration relation arises even in homogenous situations. The strong absorption coefficient ε of the **HPQ** fluorophores begs the question: at what optical path length does absorption deviate into the 'strong' regime, so that the approximate linear relation $F=kC_{PPT}$ can no longer hold for reasons of strong absorption? For soluble systems (linear regime most reliably seen with absorptions $0.01<A<<1$), the dilution or cuvette length can easily be changed to measure stronger-absorbing samples accurately. However, for the precipitating system, this is not an option. The acquisition beamwidth limits the surface area that can be detected, and for practical reasons, the height of the solution in the well cannot be decreased below a certain amount without inviting large errors from differential evaporative effects. Also, the excitation power is limited both by the source, and by the fluorescence intensity that the detector can read before saturating: being restricted to low intensity excitation, we may quickly pass to the 'strong absorption' regime. As we approach the strong-absorption regime, the signal will increasingly deviate from the linear regime until eventually reaching a constant value even as more fluorophore is added: this limit is when all incoming light is absorbed by surface-layer fluorophore before the bottom layers can be excited. Alternatively, in a desirable weak absorption regime, the whole sample may be excited with approximately identical intensity irradiation: so doubling the sample amount will approximately double the signal.

This expected decrease in signal per μM of precipitate implied that unless one could be certain that the strong absorption regime was not reached, one ought avoid trying to fit the data to a linear response curve, but should expect a law of diminishing returns. This was exactly what was seen regardless of the wellplate choice (see section I.5.4): the hypothetical values of k for the relation $F=kC_{PPT}$ varied monotonously and with a range of a factor of two over the experimental conditions. However, as I

anticipated this effect (rather than dismissing it as anomalous), data treatment which enabled scientifically-justifiable conversion of the macroscopic signal to microscopic data about the precipitated concentration could be performed.

3.4. Pre-modelling conclusions

The best design for any *in vivo* probe must be hydrolytically robust, have as good enzymatic recognition of the substrate as possible, and release a fluorophore which absolutely must have as large a Stokes shift as possible. Once these conditions are satisfied, then a relatively good signal intensity should be designed. These factors influence spacer design (recognition), fluorophore choice (ΔStokes and S), and overall probe design (overall hydrolytic robustness).

Further to this, the *mathematically* most sensitive and reliable situation is the off-ON mode. The lower is the enzymatic activity one wishes to image, the more vital it is to pursue off-ON probes: even outstanding relative change probes with low background cannot compete with this mode for *in vivo* applications. Within the off-ON probe mode, precipitating fluorophores can give a mathematically far *more dependable* quantification of enzymatic activity based on the observed signal than can soluble ones. This is due to the 'safety net' of the K_S↔signal onset time; and this is especially important with very low enzymatic activity such as may be expected *in vivo*.

These modelling conclusions of course ignore the *practical* advantages discussed in section I.1, chiefly that for a precipitated probe, signal is not lost to excretion or diffusion but precisely marks the site of enzymatic activity, and its signal has extra photostability and lower environmental sensitivity. Furthermore, from a design point of view, the *conceptual* advantage of the precipitating design is that it ensures that the probe is not just *approximately* off-ON (eg. signal enhancement of 100-fold), but *truly*, binary, off-ON on both a microscopic and macroscopic level.

However, there are practical and theoretical complications with the design and interpretation of experiments involving precipitating fluorophores. Three considerations (practical wellplate effects, theoretical spatial inhomogeneity, strong absorption) were *all* needed to anticipate these pitfalls, and to design both experiments and data treatment so that microscopic information could be reliably obtained. For example, (1) initial kinetic experiments using the original wellplates had roughly tenfold lower sensitivity to low concentrations of probe than the final designs; (2) a linear correlation of observed signal with processed concentration would have resulted in greatly skewed kinetic analysis (eg. a factor of two difference in low processed concentrations – as larger figures are traditionally seen to be more reliable, these smaller values may otherwise have been overruled as erroneous, which would in turn have skewed all subsequent analysis eg. spontaneous hydrolysis / enzyme kinetics.)

Precipitating, mechanistically and macroscopically off-ON enzyme probes were previously restricted to the ELF-97 phosphatase probe system. This thesis attempted to obtain similar (or better) results with peptidase probes, by employing innovative molecular design to overcome the prior-art problems.

I also attempted to give what I believe to be the first explicit, fully-transparent and reliably-modelled mathematical bases for anticipating and comparing the *macroscopic* behaviour and possibilities of the precipitating off-ON design, in a range of common settings (*in vitro*, *in vivo*, colony assay), set against other classic probe modes (soluble relative change, and its limiting case, soluble off-ON).

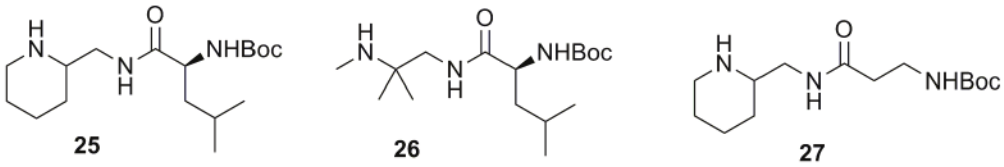
I will now report the experimental validation of the *mechanistically and macroscopically binary off-ON*, modular exopeptidase probe design, and I will show that it can promise **ultrasensitive detection and reliable quantification of enzymatic activity**, even in situations (such as low enzyme activity and high background *in vivo*) where current probe modes cannot hope to compete.

4. Synthesis of the Fluorogenic Probe System

4.1. Synthesis of substrate-spacer pairs

4.1.1. Access to spacers; choice of targets

Piperidyl spacer **22** was cheaply commercially available, and *gem*-dimethyl spacer **21** was delivered by Sean McKeon and Monica Vargas-Sanchez, following a known, high-yielding three-step procedure from acetone: acetone cyanohydrin formation[93], aminodehydroxylation with methylamine[94], and nitrile reduction by $LiAlH_4$[95]. Two initial substrate-spacer pairs were targeted: leucyl/piperidyl **25** and leucyl/dimethyl **26**, so probes could be tested in-house. The more efficient of the two spacers was to be used in the β-alanine probes; this proved to be the piperidyl spacer, giving design **27**. (Figure I.17).



25 **26** **27**

Figure I.17 – initial substrate-spacer pair targets 25 and 26, and a subsequent β-alanyl substrate-spacer

4.1.2. Protecting-group strategy

Commercial **22** was used in initial exploration of coupling strategies, saving the synthesised **21** for use once optimised coupling routes were found. The original synthesis proposed by Prof Hasserodt towards the substrate-spacer pairs planned to differentiate the primary and secondary amine sites with protecting groups, so substrate-spacer pair **25** could be formed regioselectively by peptide coupling to the free primary amine of protected species **29** (Figure I.18a). However in the imine-forming first step, I instead observed a large majority of the isomeric aminal forms **28b** (Figure I.18b).

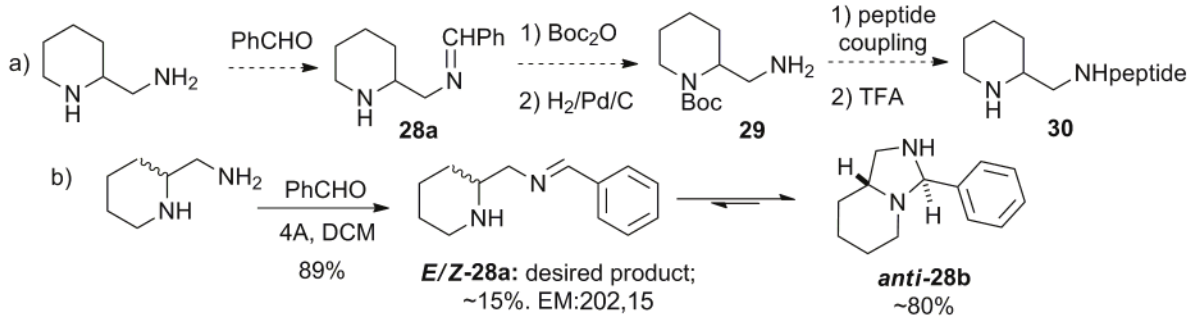


Figure I.18 – a) Prof Hasserodt's planned protecting-group strategy for differentiating the primary and secondary amine sites; b) in practice, imine 28a was preferentially present as preorganised bicyclic aminal 28b

Similar stabilisation of a bicyclic aminal by annelation can be seen in eg. Joshua[79], or Hetenyi[80] (who obtained a 9:86:5 ratio corresponding to forms ***E,Z*-28a : *anti*-28b : *syn*-28b** with an aromatic annelating ring). The ratios were solvent dependent, but the greatest proportion of **28a** was only ~15% (in chloroform, the ratio ***E*-28a**:***Z*-28a**:***syn*-28b**:***anti*-28b** was 15:0.5:82:2.5). Though synthetically disappointing, this result also indicated the success of the preorganised design for promoting

cyclisation. Trying to make use of this equilibrating mixture of isomers, a DCC-mediated coupling reaction (although with FmocLeuOH) was tested; yet no product whatsoever could be observed by DIMS and I abandoned this synthetic route.

4.1.3. Chemoselective acylations

B(OH)$_3$ method: I now sought to develop my own strategies for direct, regioselective peptide coupling to a primary amine in the presence of a secondary amine. Only one analogous reference[96] was found, using boric acid catalyst and heating in toluene with a Dean-Stark trap. I tested this procedure but replacing the Dean-Stark trap with activated 4Å molecular sieves; on the cyclisation-preorganised **22** however, this gave amidine **25b** by tandem acylation-intramolecular dehydration (again indicating the success of the preorganised design). Basic hydrolysis in a second step gave desired primary amide **25** regioselectively in 21% overall yield (Figure I.19).

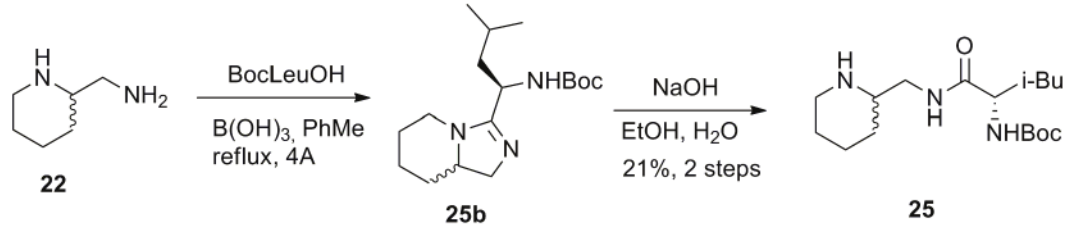


Figure I.19 - boric acid catalysed formation of amidine 25b under dehydrating conditions then subsequent regioselective hydrolysis allowed isolation of primary amide 25

β-alanyl/piperidyl substrate-spacer pair **27** was also synthesised by the B(OH)$_3$ method similarly to **25**, with hydrolysis of amidine **27b** giving **27** in 19% yield.

DCC method: Considering that kinetic control of acylation might favourise exclusive reaction with the primary amine site, I then attempted a low-temperature coupling procedure with a suitable activating agent. DCC-mediated coupling was tried with a short, low-temperature induction period in DCM (<5 min, <0 °C) reacting DCC with the protected amino acid to yield the *O*-acylisourea reactive intermediate; after the addition of the amine, the reaction was slowly warmed to room temperature and stirred for 24 h. Longer (10 min) or warmer (>5 °C) induction periods were found to give unacceptable yields of the unreactive *N*-acylurea rearrangement byproduct **25x**, even in this low-dielectric solvent; use of a salt-ice bath with only 2-3 minutes stirring before amine addition gave 10× better yields than obtained following a sample DCC protocol.[97] The several reaction products were separated by rather difficult and tedious column chromatography, giving **25** in yields up to 40-50%, as well as recovered yields of **25x** of up to 60% (Figure I.20).

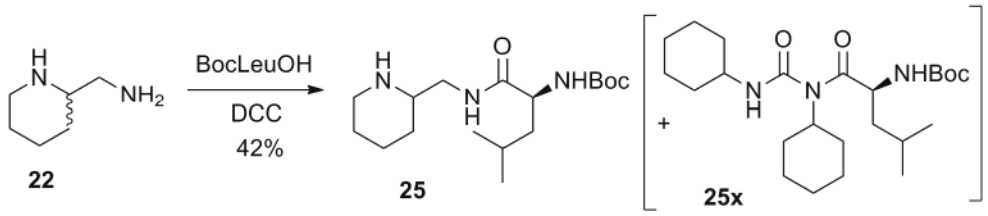


Figure I.20 – the DCC-coupling method gave primary leucine amide 25 in 40-50% yields, with variable yields of byproduct 25x depending on the induction conditions.

Modified DCC methods: The DCC-only method likely racemised[98] a small part of the chiral acid leucine (CD or chiral HPLC could be used to evaluate the percentage racemisation – or else, coupling a peptide containing at least one extra chiral centre, eg BocAla-Leu-OH, could be used to examine diastereoisomer formation by NMR). Alternative coupling methods were considered to reduce both the racemisation and the formation of byproduct **25x**. Pivaloyl-mediated coupling was imagined to probably give more racemisation[98], and the more expensive coupling reagent EDC was considered a waste as the acidic wash to remove it would also take the amine product (and possibly de-Boc the peptide; see acid/base extractions below). Biphasic reactions[99] or coupling reagents such as HOBt[100] seemed a likely choice. However, a DCC-mediated coupling with **HOBt additive**, as per Adrio *et al.*[101], actually gave a lower isolated yield (20%), and the HOBt was chromatographically inseperable from the product, although the impure mixture could be used without drawbacks in the next step. Various **acid/base extraction** procedures were tested, eg. extraction of the amines into aqueous acid to eliminate DCU, **25x** and BocLeuOH, then basification and back-extraction into DCM to eliminate **22** and residual HOBt. These reduced the isolated yields somewhat (eg. 30%) but allowed pad filtration separation rather than long columns. Use of FmocLeuOH was also examined, despite misgivings about the presence of a secondary amine, and DIMS did not indicate product formation. Replacing **22** by **21**, **26** was generated in 40% yield on the first run by the DCC method, though during acid/base extractions much of the product was de-Boc'd if the acidic phase pH was too low (pH<2). β-alanyl/piperidyl substrate-spacer pair **27** was synthesised by a milder acid/base extraction protocol with pH buffered at 2.5 (acid) and 12 (base) by phosphate, giving **27** in 30% yield after pad filtration.

Conclusions: The $B(OH)_3$ method is an atom-economical and environmentally-friendly 2-step procedure avoiding separations, so could be very useful for preparing large stocks of cheap and base-stable mono/di-peptides. It might be possible to perform the reaction under less forcing conditions (lower T, less $B(OH)_3$) and less dehydrating conditions (less MS) to yield amides such as **25x** and **26** directly, without forming the amidines. In its current form however, the $B(OH)_3$ method is first hot and dehydrating, and then very basic, so it did not seem promising for use with the polypeptide substrates which were the later targets of the project. At least initially, the more expensive and less "green" DCC method gave higher isolated yields, so this may be a better option especially for long or expensive peptides, as well as any which are unstable to the $B(OH)_3$ method. The longer is the peptide, the easier it should be to separate the products from residual HOBt (or other coupling

reagents[98]); also, the longer the peptide, the more costly it will be, so the more vital it will be to reduce racemisation and increase coupling yields by using such additives. As the spacer amines are far cheaper than the peptides, an excess of DCC and of amine could be used with impunity. For short peptides, where solubility of the product amides in both organic and aqueous media is possible depending on pH, the tedious column separation of the byproducts of the DCC method may be avoided by using careful acid/base extractions, so only pad filtration is required. Each method therefore has optimal applications.

4.2. Synthesis of three-component probes

4.2.1. ESIPT Fluorophores

Both **HPQ** and **diClHPQ** were desired for preliminary studies. An initial attempt to synthesise **HPQ** by the convenient iodine oxidation procedure reported by Bakavoli *et al.*[102] failed to give the product, but adopting the group's one-pot acid-catalysed coupling + DDQ oxidation[103] gave both **HPQ** and **diClHPQ** as powders in 80-85% yield, each NMR-spectroscopically pure after simple filtration and rinsing (Figure I.21). Each was strongly green-fluorescent under UV light, so that even small quantities were easy to detect (eg. by TLC in subsequent reactions). They both have *very* low solubility in all solvents tested, which posed some practical challenges in subsequent synthetic steps. A sample of **BTQC**[60] (Figure I.6) was delivered by Delphine Pitrat.

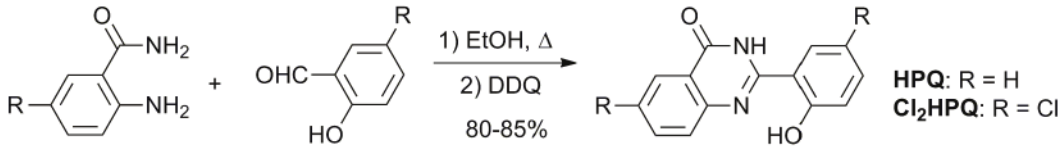


Figure I.21 – Synthesis of fluorophores HPQ and diClHPQ

4.2.2. Phosgenation coupling to synthesise protected probes

Carbamylation of the phenols **HPQ** and **diClHPQ** was planned by first forming their chloroformates, then reacting these *in situ* with the free secondary amine termini of the substrate-spacer pairs **25** and **26**. I preferred this order, rather than coupling the phenol to the spacer isocyanate, since the phenols are far cheaper than the substrate-spacer pairs and also easier to purify and recover, so they should be used in excess such that no amine is unreacted, ensuring minimum wastage. Additionally, the phenols are strongly insoluble unless derivatised; therefore it was thought better to perform the coupling when they were more solubilised (soluble chloroformates + amine), rather than risk slower heterogenous coupling reactions (solid phenol + isocyanate).

In a representative gramscale procedure, a DCM solution of solid, low-volatility triphosgene[104,105] was added to phenol **HPQ** (1 g, 1 eq) in dry DCM with 2 eq NEt_3 at 0 °C, so that the generated phosgene (4 eq) would stay firmly in DCM solution (phosgene bp 8 °C). After an hour, the volatiles were evaporated to a cold trap to leave **HPQ** chloroformate (**31a**) and $Et_3N{\cdot}HCl$, and the trapped phosgene carefully quenched into cold 2-aminoethanol solution. Then a DCM solution of amine **25**

(0.2 eq) and NEt_3 (3 eq) was injected and left to react at RT for an hour, before the reaction was quenched (see following paragraph) and the organics extracted and purified by column chromatography to yield protected probe **32a** as a colourless powder with pale blue fluorescence. Similarly, **HPQ** was coupled with leucyl substrate-spacer pair **26** to yield protected probe **32b** in 55% yield, and with β-alanyl substrate-spacer pair **27** to give **32c** in 47% yield; also, **diClHPQ** was coupled with **25** to give **32d** in 49% yield, and with **27** to give **32e** (Figure I.22).

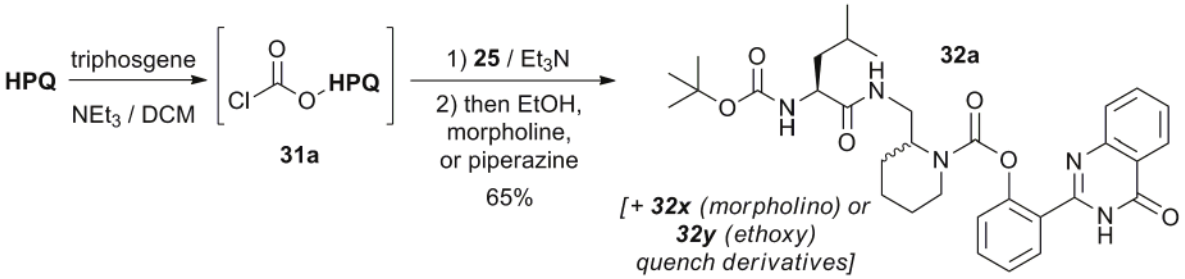


Figure I.22 – Synthesis of protected probes by phenol chloroformylation then reaction with substrate-spacer pair

The quench step had some influence on the purity of the product subsequently obtained. Briefly, whatever quenching agent was added, a substantial amount of **HPQ** was recovered during workup, which due to its low solubility but fine powder structure, resisted attempts to filter it off entirely. Column chromatography was also inefficient to purify samples as **HPQ** has a high R_f, though very low solubility, so is continuously obtained during the column (requiring a second column to eliminate it from the sample). Therefore I tested non-aqueous quenching agents. Ethanol gave the ethyl carbonate **32x** which ran faster than **32a**; so first morpholine (giving **32y**), then piperazine were tried as quench agents, to give easier separation by slower-running, soluble byproducts, which was found to be an effective strategy. The piperazine derivative was not recovered from the column; but **32y** could be recovered and this was initially considered important for stability testing of the carbamylated design (see section 5.2). It was considered best to quench with piperazine as around 9-10 spots were commongly observed by TLC in the crudes of each reaction, which necessitated somewhat careful separation, and the slower-running piperazide was more easily separable than either ethoxy or morpholino derivatives.

These phosgenation conditions were also applied to the fluorophore **BTQC** with amine **25**, however this failed to result in any products which could be identified by NMR, or isolated by chromatography, or any peaks that could be meaningfully interpreted on DIMS. The interpretation was that the seemingly far lower solubility of this fluorophore resulted in its lower reactivity, and also in correspondingly greater problems in identification or isolation of the products (Figure I.23).

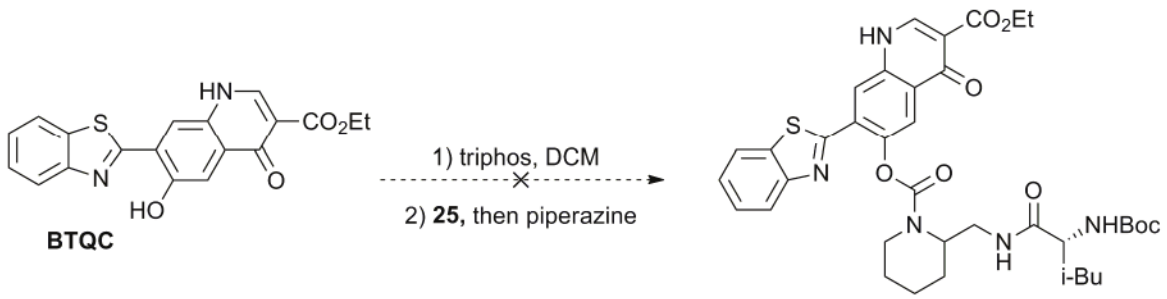


Figure I.23 – trial phosgenation/coupling of the fluorophore BTQC did not allow isolation of the desired product

4.2.3. Deprotection and isolation of the probes

The Boc protecting group on the N-terminus of the probes was removed by relatively anhydrous acid treatment: TFA in DCM at RT was found to give the best isolated yields and purity after chromatography (>90%), but HCl/OEt_2/DCM at RT, or *p*TsOH in EtOH at reflux, both also gave acceptable yields after chromatography (>77%). Chromatography was considered absolutely necessary, although the amount of probe which was hydrolysed at the carbamyl site during these deprotections was estimated <1%, because that amount of fluorophore would potentially interfere strongly with the enzymatic experiment, either by pre-spiking (see section 3.1.6) or by increasing the background signal (see section 3.1.2). The probes were therefore isolated in their TFA, *p*TsOH or HCl salt forms, which aided solubility; of these, the TFA salt was particularly water-soluble (>>2 mM). For the *p*TsOH salt, the *p*TsOH could be almost quantitatively stripped off by trituration with water, or the amine freebased into $CHCl_3$, if further exchange of the anion were desired. The final ensemble of leucyl probes for enzymatic testing in-house were therefore **1a**, **1b** and **15a**; these are shown, along with the β-alanyl probes **1c** and **15c** for external testing, in Figure I.24.

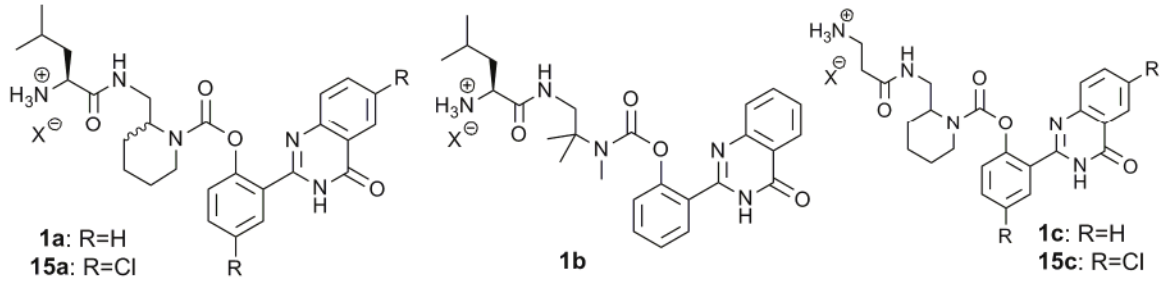


Figure I.24 – Probes synthesised for enzymatic testing

4.2.4. Resumé; and evaluation of the final synthetic route

The final synthesis of the probe system, as exemplified by **1a-1c**, **15a** and **15c**, is easy and cheap, requiring only three steps from commercial (**22**) or described (**21**) 1,2-diamine spacers (Figure I.25). The modularity of the synthetic pathway described, the orthogonality of the reactions it involves, and the exploration of strategies which may be more or less appropriate for regioselective coupling of different peptide substrates, provides a firm basis for the high-yielding synthesis of other probes or prodrugs for phenols or general alcohols, following this route.

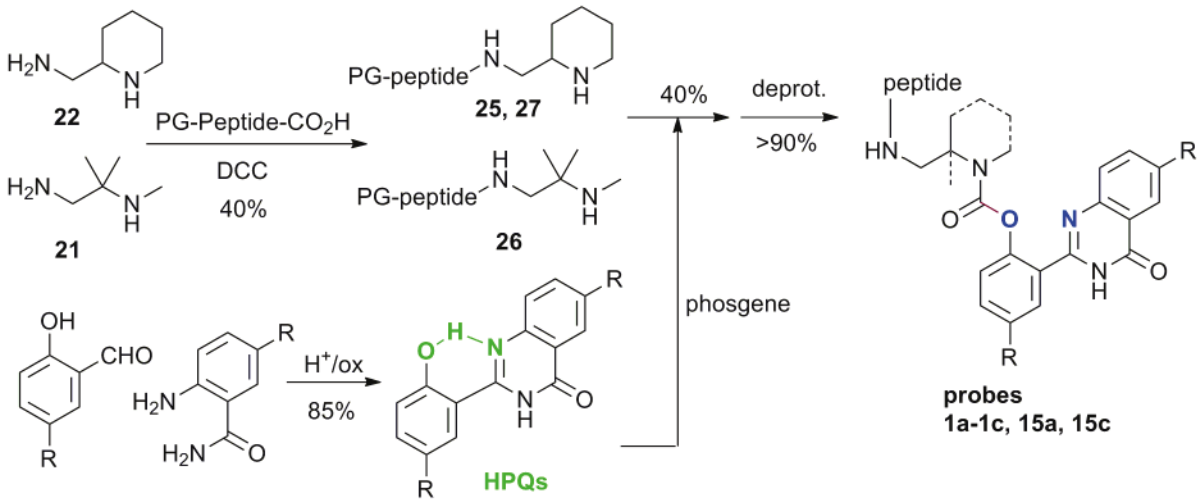


Figure I.25 – Synopsis of the final synthetic route to probes 1a-1c, 15a and 15c.

5. Enzymatic evaluation of the fluorogenic probe system

Preamble: Recall of Design Principles

Probes **1a**, **1b** and **15a** do not display any ESIPT fluorescence as their phenol is carbamylated, so no excited-state phenol-to-imine hydrogen transfer is possible. Because the ESIPT mechanism is responsible for the very large Stokes shift (>130 nm) in the very strong fluorescence spectra of **HPQ** and **diClHPQ**, the probes gave virtually no background fluorescence signal in the proagent state, since their normal fluorescence has small Stokes shift ($\Delta\lambda_P$~20 nm) and is anyway very weak, but collection was performed with a 145 nm wavelength split. The enzymatic cleavage of the probes' leucine residue exposes the amine termini of their spacers in intermediates **33a**, **33b** and **33c** respectively; these then cyclise spontaneously to generate ureas (**23** and **24**), releasing the ESIPT fluorophores **HPQ** and **diClHPQ** (Figure I.26). Once the solubility limit K_S for the fluorophore has been surpassed and nucleation has commenced, the precipitating fluorophores give a strong ESIPT fluorescence signal.

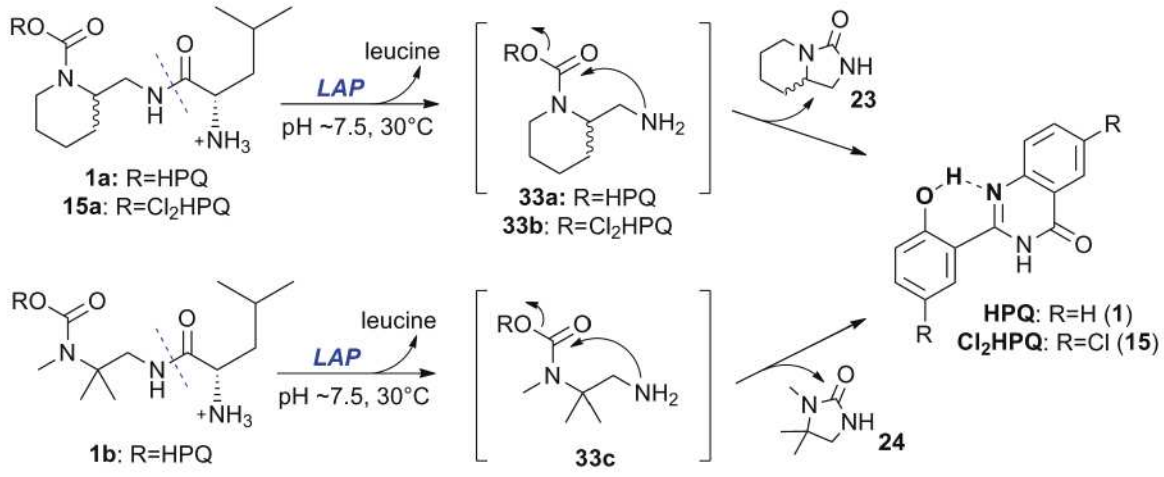


Figure I.26 - Enzymatic cleavage of the aminopeptidase-targeting substrates of probes 1a, 1b and 15a results in fast, spontaneous cyclisation of their spacers and liberation of the fluorophores HPQ and diClHPQ.

Salient Results

Kinetic studies of the probes incubated with leucine aminopeptidase validated the linker designs: shortly after enzyme addition, fluorescence signals began to rise to plateau values which were well correlated to the initial probe concentration, indicating complete conversion. The speed of attaining the plateau even with a low enzymatic activity in the well indicated that the probes were well tolerated by the enzyme (and can therefore hope for competitive processing *in vivo*), vindicating the design for low steric demand at the scissile bond, and the rapid response times (~instantaneous onset of signal and 30 min for signal plateau at 100 µM of **15a**), vindicated the design for efficient preorganisation.

Control experiments in the absence of enzyme, testing the generation of false positive signal by spontaneous hydrolysis, gave no detectable signal increase over 11 hours at up to 209 µM probe concentrations of **1a** and **1b**, while more electron-withdrawn **15a** did *begin* to generate a small spontaneous signal after around 10 hours at 100 µM, but these stabilities are truly exceptional for phenolic probe designs.

I performed detailed considerations of assay design and implementation, theoretical treatment of the data, then analysis of the calculated quantities, and I could finally deconvolute reliable and justified numerical determinations for the signal↔concentration relation $C_{PPT}(F)$, then the crucial parameter values of enzymatic processing (K_M and v_{max}), spacer cyclisation (k_{CYC}) and fluorophore precipitation kinetics (k_{PPT}, K_S, K^*_S and p). These numerical values supported the qualitative conclusions of rapid enzymatic processing and spacer response favourable for *in vivo* applications, and also enabled the construction of a numerical model which fully reproduces the experimentally-observed raw data.

The correlation between the macroscopic behaviour observed and that predicted by the physically reasonable microscopic parameters I determined enables me to conclude that I have been able to cover both theoretically and practically, every crucial step from theory to implementation, analysis and both practical and theoretical understanding, to validate and support this novel probe system as an *exceptionally robust, ultrasensitive design for* in vivo *aminopeptidase probes*.

5.1. Assay design and data processing

Assay design is a vital and non-obvious consideration. Here, the emphasis was to perform *in vitro* assays whose methods and results could as reliably as possible be extended to *in cellulo* and *in vivo* tests, and whose conclusions would reflect what results can really be achieved directly in practice in a standard biology lab, without optimisation of equipment, or complex procedures needing specific and high-level technician training or difficult and personally-controlled data processing.

The goals of the assays were split into two levels (see section III.3.2.7). (1) **Macroscopically**, the assays should use the collected signal to illustrate the sensitivity that can be obtained with the probes, and its time dependence; this highlights the standout features of the probe design (hydrolytic stability, photostability, rapid response). (2) **Microscopically**, the assays sought to model, understand, process, and deconvolute the obtained data to determine *all* the underlying molecular parameters of enzymatic processing, spacer cyclisation and precipitation, and ultimately to confirm this understanding by building a molecular mechanistic model that reproduces the macroscopically-observed behaviour. Only once these two levels of observation and understanding were attained could the results be reliably transposed to other situations, and therefore used to compare to prior literature systems as well as to predict the limiting-case behaviour of the current system.

5.1.1. In vitro *test considerations and motivation*

The practical goal of mimicking realistic conditions for tests covering *in vitro*, *in cellulo*, High-Throughput Screening, plate assays, and towards *in vivo* testing, implied designing a robust and scientifically sound experimental protocol.

Three important criteria for my test design were **(1)** low enzyme concentration, **(2)** minimal background manipulation, **(3)** using only standard equipment in strictly unoptimised fashion to collect

data (eg. standard filter sets and exposure times, simple procedures), with a statistically-defensible, generalised algorithm to perform data treatment.

Criteria (1) and (2) were addressed in section I.1.3. Standard equipment is an obvious requirement for robust practical implementation; although research groups may achieve outstanding sensitivities on home-built equipment[82]. Data processing is perhaps less obvious. Some imaging methods, such as MRI, require significant training before images can be reliably interpreted (years), because many data processing adjustments must be made depending on the specific results obtained in each experiment. Usually this involves at least selecting image 'brightness', 'contrast' and 'background' levels to highlight features of interest in a particular zone. Yet this results in a very qualitative interpretation of any features seen, as these operations, each with their own assumptions and errors, are overlaid on the data; this may be unavoidable with high- and variable-background methods. A far more satisfying scenario, especially for fluorescence imaging where a low background *can* be achieved by appropriate design, is to develop a standard data-processing algorithm to treat the raw data without requiring human intervention, still generating processed data which would enable quantitative conclusions about any features seen: as was in fact achieved. High-Throughput Screening and other automated biotech procedures naturally require this absolutely; though I suspect it is not performed nearly so often or rigorously in this field of academic probe design.

Very often in the chemical literature, perhaps especially in the highest-profile journals, these three criteria are willfully ignored in the pursuit of reporting the most impressive calculations of sensitivity, efficiency, speed, or specificity: calculations which are actually futile, as they misrepresent the true figures which will be obtained in the biological test settings (that are the ostensible intended applications of such papers). What is worse, **(4)** probe stability is usually not reported, and spontaneous signal generation control experiments are either not reported or else are used as 'background' readings: as introduced in section I.3.1, this masks the poor performance which is seen without fail once such probes are applied *in cellulo* or *in vivo*.[1,4] A more extensive commentary on these features is continued in section III.3.1.

5.1.2. In vitro *test conditions*

A low concentration of the target enzyme (~60 nM) was used, at 30°C, in Tris/NaCl buffer at pH~7.5, without cosolvents. The excitation/emission wavelengths were selected by 10 nm bandpass filters at 355 nm / 510 nm, ie. standard near-UV excitation, and standard GFP emission filters for biology (transferability of results). As we use a precipitating fluorophore which forms a solid deposit on the bottom of its container, measurements cannot be performed in a cuvette fluorimeter as for soluble fluorophores (sideways illumination), but must be performed so that the excitation beam hits the precipitate as much and as evenly as possible; the cheapest equipment for this is a top-down microplate reader (though a bottom-up fluorimeter would have been more convenient to stop evaporation). Full biochemical and physics parameters are detailed in sections III.3.2.1 and III.3.2.4.

Two unoptimised data collection protocols were devised, 'scan mode' where the beam scans over the wells pixel by pixel (reflecting *in cellulo*/plate/colony assays), and 'kinetic mode' where the beam is directed briefly only into the centre of each well (reflecting HTS or standard *in vitro* use). These were also thought to have different potential for quantitative reliability (sections III.3.2.2 – III.3.2.3).

Spatially-dependent evaporation of the test solutions was countered by scattering each data series over the wellplate (Figure III.11) and by topping up the solution levels over the 10-20 h assay times; signal crosstalk from neighbouring wells was avoided by surrounding each active well with water-only wells. The choice of wellplates was critical for assay success (see Figure III.10), and was different for scan and kinetic modes. Substrate stocks were carefully prepared from chromatographed samples, with minimum time between preparation and use to avoid pre-spiking which has been an unrealised problem in the group's previous work (section III.3.5.1). The time between enzyme addition and the first acquisition was controlled and minimised. A minimum of five replicates were used per data set.

5.1.3. Data processing

I wrote a multistep conditional data processing algorithm to ensure coherant and objective treatment and reporting of the data. Although this is a standard procedure in signal processing in physics/engineering, I have not found it applied explicitly to probe reports in academia (perhaps because they remove the possibility reporting the best features case-by-case). The five steps were (1) reduction of systemic variations due to detector drift, (2) removal of background due to scattering from the wellplate (*not* from control solutions containing anything other than pure water), (3) removal of outlier data points, which was a significant problem especially at high signal values, (4) selection of harmonious sets of at least three of the replicates to define a data set, (5) averaging the replicates to yield fluorescence signal over time data, F(t), for each probe concentration in that run (sections III.3.2.5 – III.3.2.6).

5.2. Results

I performed more than 120 hours of enzyme test acquisitions, optimising my sample preparation methods, experimental time, and confirming my results by repetition. I have selected two kinetic runs (run 1: **1a** and **1b**; run 2: **1b** and **15a**) and one scanning run (run 3: **1a** and **1b**) as the only runs for numerical analysis here, although results from other runs are depicted where convenient. The results from the potential *pseudomonas* probes are currently confidential and cannot be given here.

5.2.1. Visual results (photographs and scanning test image)

A sequence of photographs taken with a personal digital camera (Canon IXUS 110) of the activation of probe **1a** by leucine aminopeptidase (LAP) enzyme in pH~7.5 buffer at RT is shown in Figure I.27. The top left panel is a bright-field image, all other images are obtained from visible-spectrum fluorescence under 365 nm irradiation. The evolution of the fluorescence signal with enzyme activation can be seen (vials **2**† and **3**†), as well as the absence of spontaneous signal generation (other

vials). Only after around 3 weeks at 2 mM concentration did probe **1a** start to form a fluorescent precipitate, which gives an indication of the timescale needed, even at this high concentration, for spontaneous hydrolysis to be detectable.

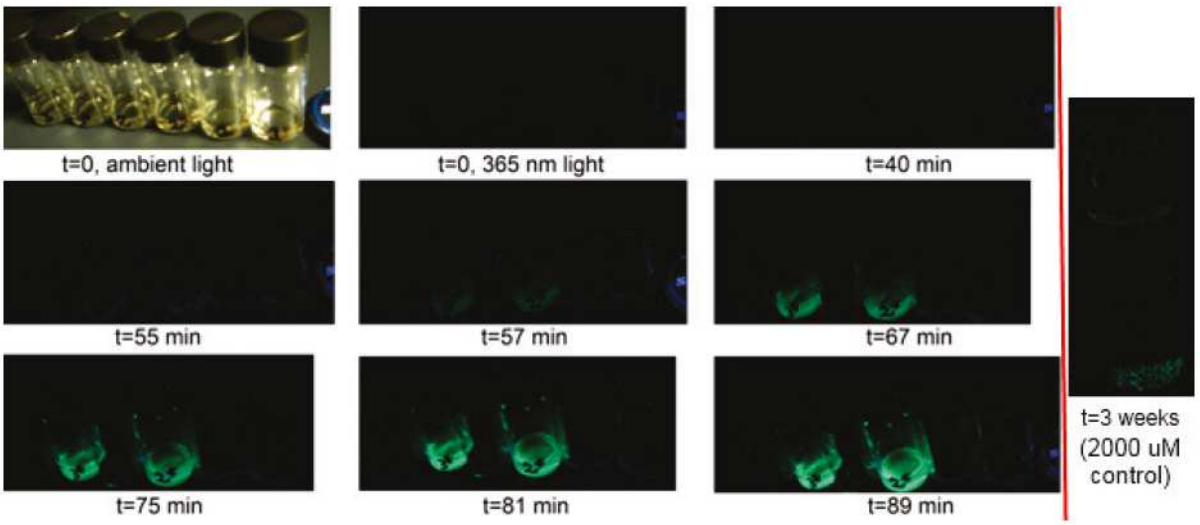


Figure I.27 – (left): Enzymatic activation of probe 1a by LAP. A fluorescence signal appears over time when the probe is activated by LAP (vials 2† and 3†), but zero fluorescence signal is seen in the absence of enzymatic activation (other vials). A blue-fluorescent marker is occasionally visible at the right of the wider images and gives a rough indication of a constant level of fluorescence for normalisation of the images' intensities. (right): After three weeks at RT when concentrated to 2000 μM, probe 1a showed the precipitation of some tiny crystals of HPQ.

Figure I.28 shows a detail of a greyscale fluorescence image (dark=no signal) of a microplate from a typical scanning-mode test of compound **1a** (after 40 h incubation). The three series of active wells (different concentrations of **1a**, with fixed concentration of LAP) illustrate the relationship between the observed signal and the initial probe concentration: wells with 209 μM (B2, D4, F6, H8) show more signal than with 80 μM (B8, D2, F4, H6), or 21 μM (B6, D8, F2, H4). The other wells are controls, either containing distilled water for signal background measurement (B4), or no-enzyme controls (*i.e.* with probe **1** but without enzyme: H2 at 209 μM; D6 and F8 at 80 μM).

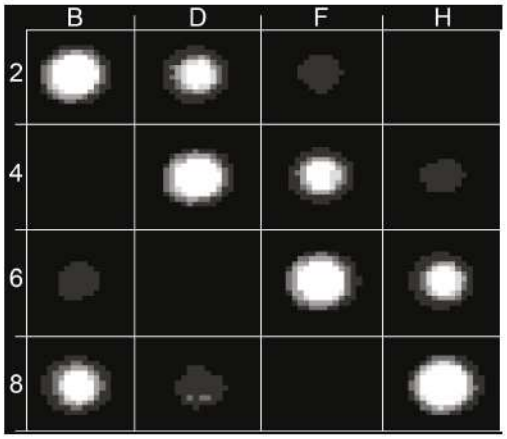


Figure I.28 - Detail of a scanning mode fluorescence image of a microplate, taken after 40 h incubation of probe 1a.

The probe is robust (there is no false positive fluorescence signal arising from spontaneous hydrolysis of the carbamate linkage) and in its pro-fluorescent state is truly 'off' (the unactivated probe wells are

indistinguishable from the no-substrate control even after 40 h). Note too the distribution of wells to combat evaporative effects and signal crosstalk (see section III.3.2.3).

5.2.2. Kinetic results

Description: After a certain delay time (signal onset time), all the probes began to generate enzymatic fluorescence signal; the higher the probe concentration, the shorter was the delay time. Then the fluorescence signal values rose rapidly to plateau levels (plateau time). As expected, the rise was more rapid the higher was the initial probe concentration (Figure I.29; note that the observed signals are not ten-fold- or twenty-fold-enhanced, but increase by up to *three orders of magnitude*, indicating great potential sensitivity). Similar observations were seen for both scan mode and kinetic mode.

Spacer kinetics: I interpreted the rapid rise to the plateaux as reflecting good enzymatic processing even at low substrate concentrations (as the probe was depleted). This indicates a satisfactorily low K_M, and predicts a probe which may therefore be competitively processed even in the presence of native substrate (*in vivo* aim), which was one crucial factor in the design of these low-steric-demand spacers. Both signal onset and plateau times were shorter for spacer **22** than **21**, which I interpreted as indicating both better overall substrate processing and faster subsequent cyclisation rate for the **22**-based probes; I later attempted to deconvolute the kinetics of these processes to support and quantify this conclusion (see section I.5.5). Importantly, comparing the absolute signal from **1a** and **1b** assayed under identical conditions, the plateau values reached were identical; I interpreted this result to mean that (a) both probes were ~100% processed. Furthermore the forms of the curves were similar (simple sigmoid shape); as **1a** has diastereoisomers (based on the spacer enantiomer) but **1b** does not, I concluded that (b) for *neither* stereoisomer of spacer **1a** was enzymatic recognition/processing a rate limiting factor in signal generation by this model enzyme.

Fluorophore choice: The plateau signals did not diminish in intensity despite long irradiation, reflecting the expected near-total resistance to photobleaching by the precipitated fluorophores. Both signal onset and plateau times were shorter for fluorophore **diClHPQ** than **HPQ**, which I interpreted as reflecting the lower solubility of **diClHPQ**; this interpretation was later supported by calculations of both K_S values and by confirmation that spontaneous hydrolysis was not a rate-determining factor in signal generation. Note that to enable intercomparison of the results between the probes with different fluorophores used in two different runs, the signals $F_X(t)$ on Figure I.29 have each been divided by the plateau signal for their fluorophore from incubation with 100 μM probe, which approximately normalises their intensity.

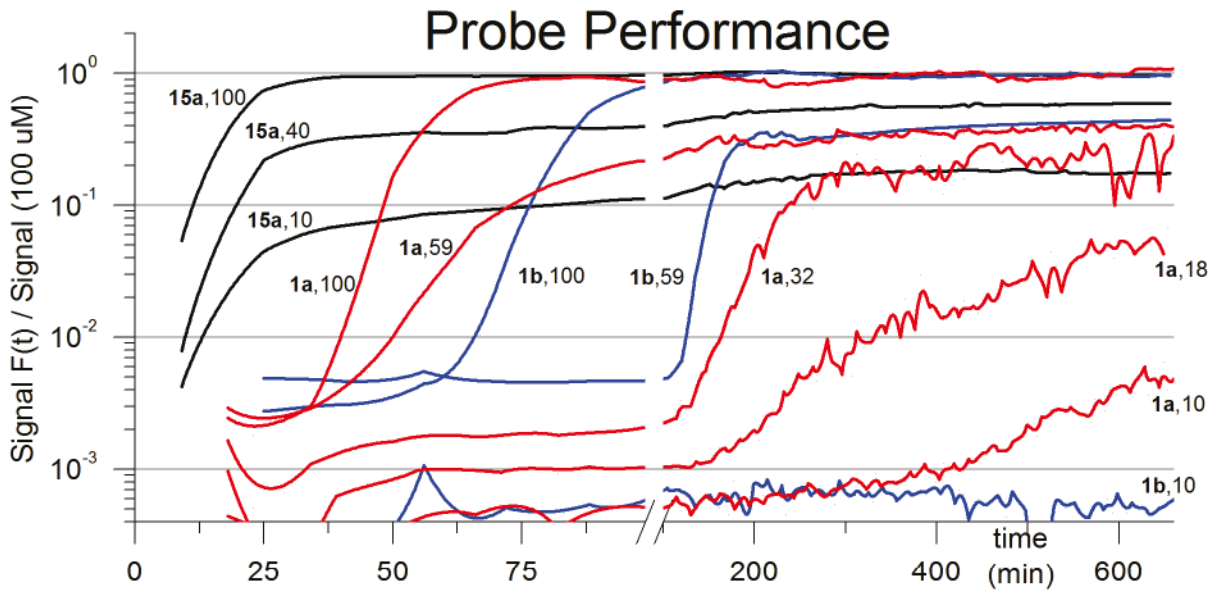


Figure I.29 – General results from kinetic tests with different concentrations of the probes, indicated in µM. Note the logarithmic vertical axis and the break in the horizontal axis.

Probe Stability: To ensure the probes were hydrolytically stable over the lifetime of the experiments, high concentrations of probe were monitored in the absence of enzyme to control for spontaneous signal generation. Visually, the results were pleasing: neither **1a** nor **1b** gave any measurable signal whatsoever even at 209 µM, over 40 h. By contrast, dichlorinated **15a** at 100 µM begins to give a signal due to spontaneous hydrolysis after about 9 hours (Figure I.30). Importantly for later analysis, the 100 µM controls for run 1 were prepared by dilution of 44 µL of substrate stock, and those for run 2 by dilution of 25 µL, to a final volume of 100 µL.

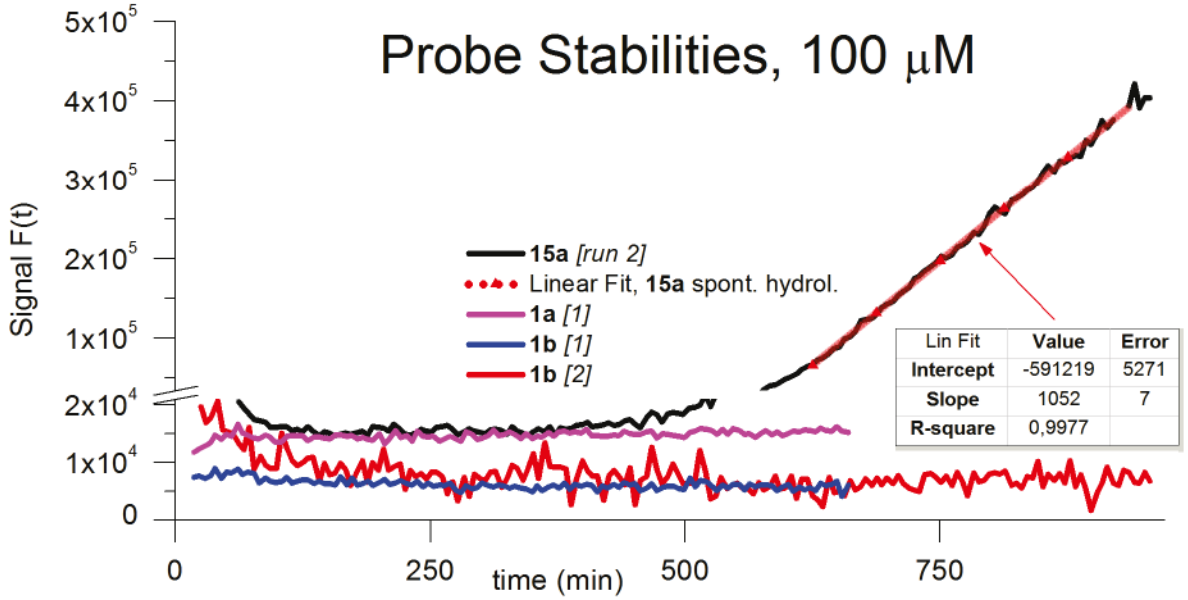


Figure I.30 – probe stabilities were high over the lifetime of the lifetime of the experiment; the hydrolysis signal seen for 15a coresponds to a halflife of 15 days. Data taken from kinetic experiments. Note the break in the vertical axis, which is rescaled by 5-fold after the break. Note too the variations in signal at the start of the experiment as the system stabilised (signals from run 2 decrease), whereas in run 1, pre-stabilisation had been performed (no change).

These observations should firstly be qualitatively compared with the group's previous, *non*chlorinated-HPQ peptidase probe structures **34a** and **34b** (Figure I.31) which at only 38 μM, were reported to begin spontaneously generating signal after 30 min (though, see analyses in section I.5.6.1 and especially section III.3.5.1).[73] However, the absence of observable spontaneous signal generation for **1a** and **1b** is not to say spontaneous hydrolysis was zero, but rather that the level of spontaneously hydrolysed probe did not reach the fluorophore solubility limit during this time. For example, when a solution of 2000 μM probe **1a** was left to stand in buffer, after 3 weeks a fluorescent precipitate had just begun to form. Dichlorinated **15a**'s lower hydrolytic stability than **1a** and **1b** was expected (similar effects have been seen for other electron-withdrawn chlorinated fluorophores[30]) but even so, it compares *very* favourably with the group's previous efforts with the less withdrawn fluorophore. Furthermore, in absolute terms, the spontaneous hydrolysis signal at the end of the experiment with **15a** is 4×10^5 RFU (16 h), whereas its enzymatic reaction plateaus at 2.5×10^7 RFU after only 40 min (Figure I.29): no confusion between the enzymatic and non-enzymatic scenarios is possible.

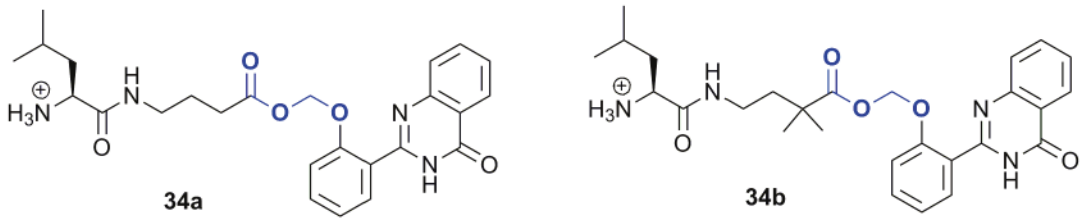


Figure I.31 – Prior art probes in the group: 34a and 34b; the labile acyloxymethyl group is highlighted

As an aside, the fluorescence signal from the no-enzyme control of **1b** was, pleasingly, concentration-independent: it was identical for 10 and 100 μM concentrations. The corresponding signals for **1a** were five times higher for its 100 μM control than for its 10 μM control, or equivalently, ten times higher than the comparable control **1b** from the same run (Figure I.30, run 1). One possible explanation, is that the probe counterion of **1a** used in this run was *p*-TsO^-, while the TFA counterion was used for **1b**. I observed in the solid state that the *p*-TsOH salt of **1a** gave a far stronger (yet strictly blue) fluorescence than its TFA salt, and I attributed this fluorescence to the *p*-TsO^- counterion. I therefore considered that the tosylate may have retained some of its fluorescence activity in solution, leading to the observed increase in the solution-state fluorescence with its increasing concentration. Note however the extremely low absolute values of the no-enzyme control wells' fluorescence, and crucially, the fact that their signals do not increase over time.

5.2.3. Scanning mode or kinetic mode?

Although much care was taken to design the scanning mode (cf. section III.3.2.3), after data treatment it did not appear to give substantially more reliable, more sensitive, or more quantitative data than the kinetic mode, which was far easier to implement and also gave data which were better suited to kinetic parameter determination (more data points per unit time; see for example Figure I.32).

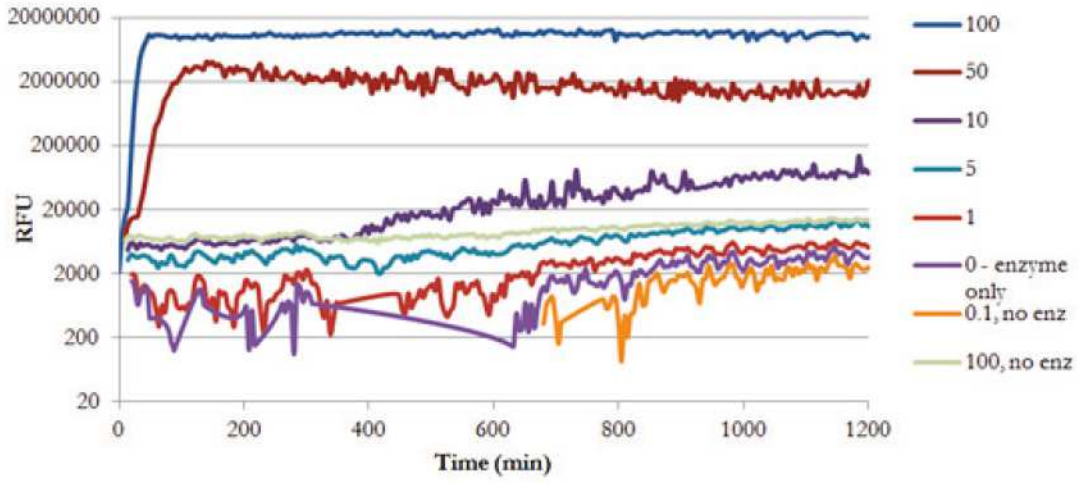


Figure I.32 – For HPQ-probe 1a, the lowest tested probe concentration that resulted in signal detection was 10 µM, and this required around six hours' incubation to be seen; 5 µM probe wells never gave signal (concentration below K_S). Pleasingly, the no-enzyme control of the 100 µM probe was indistinguishable from the no-signal wells (data shown are from run 3). All these signal curves reproduced the behaviour seen for the scanning mode tests, so the kinetic mode was the only one examined further.

I therefore restricted my analyses mainly to the kinetic mode measurements, although I will show that *both modes give equivalent results*, so the method I developed could be used equally well for both desired assays (cf. kinetic↔High-Throughput Screening, scanning↔Colony Assay; section I.5.1).

5.3. Enzyme test analyses: steady-state parameters

To reliably compare the probe performance accurately with other literature data, some rather non-obvious analyses must be made. There are two distinct ways that data must be analysed: (1) "steady-state parameters" are determined over long times and with no time-dependence errors, and include hydrolytic stabilities and signal-to-control ratios; (2) kinetic parameters are determined over short times and with relatively large time-dependent errors, including the Michaelis-Menten constants for the enzymatic processing of the substrate-spacer pairs, and the halflives of cyclisation of the spacer-fluorophore pairs. These two analysis modes are presented in separate sections; the lower-error steady-state parameters are calculated first.

5.3.1. Calculation of solubility constants K_S

The solubility thresholds K_S of both **HPQ** and **diClHPQ** need to be considered when correlating an observed signal with a processed probe concentration, because the quantity K_S is always 'missing' from the signal-generating precipitate (cf. the consistent display of a lag time before signal generation is observed in the kinetic experiments: see Figure I.29). Neither fluorophore had a reported solubility

constant I could directly use in these enzymatic tests, although the solubility of **diClHPQ** was indicated at "around 1-5 μM",[17] and it was thought to be more insoluble than **HPQ**.

Both phenols' solubilities are pH-dependent due to deprotonation, and at the time we did not know how strong the dependency would be. Therefore I tested probe performance in different buffer pH (6.7-8.3) to examine if that would have solubility effects. I also examined different buffers (Tris, Tris + EDTA, Tris + NaCl) to see mainly if they would affect nucleation or enzyme activity. No significant differences were seen in either sets of results, so I retained the pH~7.5, Tris+NaCl buffer for all subsequent work.

There were three methods I used to calculate K_S: extrapolation, and plateau fitting (both steady-state methods), and initial velocity (kinetic measure). These methods are detailed in section III.3.3.1; the end results I selected were K_S(**diClHPQ**) = 1.5 μM and K_S(**HPQ**) = 7 μM, based on the more reliable steady-state methods.

5.3.2. Spontaneous hydrolysis halflives

As introduced in section III.3.3.1, for **15a**, the hydrolysis calculation was relatively easy, considering the strongly linear spontaneous hydrolysis regime (R^2=0.9977 over >6 hours; Figure I.30), and returned τ(**15a**)=567 hours (24 days).

As no spontaneous hydrolysis was observed in either **1a** or **1b**, I instead calculated thoroughly reliable minimum halflives based on assuming the appearance of signal just after the end of the longest experiment, which gave τ(**1a**), τ(**1b**)> 427 h. However, based on more *ad hoc* experiments checked every few days over three weeks, I derived a more reasonable estimate for the hydrolysis halflife as τ(**1a**|**1b**)≈4151 days. These calculations are detailed in section III.3.3.2.

Knowing the hydrolysis halflives, the precipitating design now permits us to conclude reliably whether any small observed readings are due to spontaneous hydrolysis, or not (eg. machine noise, detector drift). To do so, choose three distinct time points, then knowing the solubility threshold, calculate the concentration that would have to have been hydrolysed at each timepoint to generate the observed signal (relative to zero), and see if the change in concentration can coherantly corespond to spontaneous hydrolysis. For example, take Figure III.15 and examine the increasing values of F(t) for the 209 μM stability test (no enzyme). The last three F(t) values are 54(210 min), 112(430 min), and 228(1260 min), implying *possible* hydrolysed concentrations of (K_S+0.02), (K_S+0.04), (K_S+0.08) μM respectively. But as the solutions were confirmed not to be pre-spiked (see sample preparation), then knowing K_S=7 μM, this proves that the cause of such apparent signal increase cannot be slow spontaneous probe hydrolysis, but implies that it is probably due to drift in the machine (or accumulation of dust, or other non-probe-related effects).

5.3.3. Probe sensitivity: signal-to-control ratios at plateau

The spacer design ensuring rapid enzymatic processing and fluorophore release by cyclisation, plus the use of a brightly fluorescent precipitating fluorophore with large inherent signal-to-background in an off-ON mode probe, ensured the rapid generation of a large, photostable signal. Uniting this large and photostable signal with the probes' hydrolytic stability, the precipitating off-ON mode, and the small inherent background due to the probes' large ΔStokes, I will now examine the outstanding **signal-to-control ratios** delivered by the probes.

The observed signals were qualitatively large relative to the background (Figure I.29), due to the use of an off-ON probe with large Stokes shift and emission yield, so could easily be distinguished. The fluorophore's photostability (plateau signals do not decrease over time), then enabled a coherant correlation of the signals at the plateaux of fluorescence with the maximum possible signal from those wells. Note that if photobleaching *had* affected the fluorophore, there would be no plateau since both the C_{PPT} and the resulting signal-to-control ratios would be time-dependent and continuously diminishing.

Section III.3.3.3 details the analysis of the observed signal plateaux, plus extensive consideration of the physical reasonableness of determining plateau values at all; abbreviated results and discussion are presented here.

After data processing, both scan and kinetic modes clearly showed huge sensitivity for the enzyme-active wells, and no false positive signals for the no-enzyme controls or the probes with concentration less than K_S, for probes **1a** and **1b**. As shown in Figure I.30, the 100 μM no-enzyme control of **15a** gave a signal due to spontaneous hydrolysis after 8 hours, but even so, sensitivity was very high. Representative results are presented in Table I.1 (kinetic mode, all probes). Recall that processing raw data A(t) into F(t) involves not only outlier removal but series consistency checking, which explains small variations in trends between the two. The very high signal-to-control ratios calculated for the plateaud signals (eg 3000:1), and even for signals that were not yet plateaud, indicates the extremely high sensitivity of these probes as a practical system for detecting aminopeptidase activity.

Kinetic mode, 11 h	**A_{PLAT} ($\times 10^5$ RFU)**	**F_{PLAT} ($\times 10^3$ RFU)**	**F_{PLAT}/F_{NoEnz} (Signal/Control), no t_0 subtract**	**Signal / Control, t_0 subtract**	**Comment**
RUN 1 (660 min):					
Distilled water	3.4	(0)	-	-	-
Enzyme without probe	3.4	0.8	-	-	-
1a, 378 µM	362	37517	728	7441	-
1b, 100 µM	112	8661	1467	∞	-
1a, 100 µM	101	9711	618	5712	-
1b, 59 µM	80	5631	954	∞	-
1a, 59 µM	55	3861	371	3976	-
1b, 32 µM	38	3749	635	∞	-
1a, 32 µM	30	2951	427	6009	-
***1b**, 18 µM*	*13*	*605*	*101*	∞	*no plateau*
***1a**, 18 µM*	*13*	*551*	*108*	*2274*	*no plateau*
***1b**, 10 µM*	*3.5*	*9.5*	*1.5*	∞	*no signal*
***1a**, 10 µM*	*4.8*	*47.6*	*12*	*476*	*no plateau*
1b, 100 µM, NoEnz	3.5	5.9	(1)	-	no signal
1a, 100 µM, NoEnz	3.5	15.7	(1)	-	no signal
1b, 10 µM, NoEnz	3.5	6.0	(1)	-	no signal
1a, 10 µM, NoEnz	3.5	4.1	(1)	-	no signal
RUN 2 (960 min):					
Distilled water	4.5	(0)	-	-	-
Enzyme without probe	4.5	8	-	-	-
15a, 100 µM	310	25818	1655 (64)	43000 (66)	value vs. t=300 (t=960)
1b, 100 µM	113	20613	2944	21000	-
15a, 40.5 µM	166	15009	962	25000	-
1b, 40.5 µM	38.5	9872	1400	10000	-
15a, 10 µM	56	4314	276	7190	-
1b, 10 µM	6.5	16.6	1.1	17	signal *just* starting (960)
15a, 100 µM, NoEnz, 200<t<400 min	4.7	15.6	(1)	-	no signal
***15a**, 100 µM, NoEnz, t=960 min*	*9.2*	*404*	*26*	*673*	*spont. hydrol. signal*
1b, 100 µM, NoEnz	4.6	7	(1)	-	no signal

Table I.1 (see section III.3.3.3 for details) – Raw and processed plateau signals from kinetic-mode experiments with all the probes (two separate runs, between which data are not directly comparable; the fluorophores of 15a (diClHPQ) and 1b (HPQ) also have different signal absorption/emission so they cannot be directly intercompared yet either).

Recall that literature data is almost universally t_0-subtracted: the no-enzyme control $F_N(t=0)$ values are subtracted from all F_{PLAT} values during calculation of "signal/control" ratios. For example, Urano *et al.*[40] subtract their *t*=0 fluorescence readings, so as to calculate a signal-to-time-zero value of only 76, after only 8 minutes with 5 U enzyme and only 1 µM probe (as I outlined in section III.3.1.1, I consider that this is because with their 5 hour hydrolysis halflife, 2% of their probe has already hydrolysed during this time, so their true signal-to-control ratio is less than 25, and dropping constantly). The present system's t_0-subtracted signal-to-control values range from 500 to 43000 after more than 11 hours of incubation and using 2 orders of magnitude less enzyme; this is unquestionably superior sensitivity compared even to the theoretical *in vitro* maxima which can be obtained with

literature-reported, soluble relative change probes (maximum possible value = *n*). Even expressly avoiding such t_0-subtraction during signal treatment, the sensitivities for the present probes are still far superior, which illustrates the sensitivity that these probes can achieve even over long timescales at *in vivo* conditions of enzyme concentration, without 'fudging' the signal processing.

Table I.2 presents sensitivity results obtained with the scanning mode, on probe **1a**, to illustrate the very close similarity of the results obtained by the two data acquisition methods, over different timescales and on different experimental runs. This confirms the practical utility of the probe system as a reliable method for ultra-sensitive detection of aminopeptidase activity across a range of (activity/probe-concentration) values, in different assay applications and with different data acquisition methods.

Probe 1a; Scan mode	**$A_X(t)$ (RFU)**		**$F_X(t)/B(t)$ (Signal)**		**$F_X(t)/F_{NoEnz}$; no t_0 sub (Signal/Control)**	
Well Series	**24 h**	**40 h**	**24 h**	**40 h**	**24 h**	**40 h**
Water	72839	73029	(0)	(0)	(0)	(0)
81 μM, NoEnz	78944	82526	0.08	0.13	(1)	(1)
209 μM, NoEnz	86302	94865	0.18	0.30	(1)	(1)
209 μM	4900858	5042099	66.3	68.0	368	226
81 μM	2415589	2462269	32.2	32.7	403	250
21 μM	511910	543580	6.03	6.44	est. 300	est. 161

Table I.2 – Raw and processed plateau signals from a scan-mode experiment with 1a. Brackets indicate that the data point was used to calibrate the values of the others in its column, so is not statistically significant.

5.4. Conversion of signals F(t) to concentrations $C_{PPT}(t)$

The quality of the correlation concentration↔signal is acceptable: for both fluorophores, increasing signal strongly reflects increasing precipitated concentration (cf. Table I.1, Table I.2). This is not obvious: an initial worry was that high fluorophore concentrations might reach the strong absorption limit and give no further increase of signal with probe concentration, therefore be unable to correlate an observed signal with a precipitated concentration. Alternatively, hit-and-miss excitation of highly spatially inhomogenous samples (see section I.3.3) might give such apparently incoherant results within and between data sets, that the prediction of concentration based on observed signal would be fraught with high error.

Simply examining the F(t) curves is insufficient to conclude on the true probe potential however. To analyse and compare data scientifically, the precipitated concentrations $C_{PPT}(t)$ corresponding to each signal level F(t) must be reliably estimated and **with a precipitated, strongly-absorbing fluorophore this is not obvious**. My analysis (section I.3.3) explicitly warns of potential nonlinearity of the F↔C_{PPT} relation as it changes from a weak absorption regime (quasilinear) to a strong absorption limit (asymptotic). Then, given that **F(t) spans three orders of magnitude in value** (Figure I.29), I considered that the simple linear conversion $F(t)=k_{LIN}C_{PPT}(t)$ with k_{LIN} determined by plateau values (high C_{PPT}) might **introduce significant error, especially at lower concentrations**.

Now, kinetic parameter analyses require accurate calculation of $\mathbf{m_{max}}$ (the maximum rate of precipitation, in $\mu M \cdot min^{-1}$) and $\mathbf{C_m}$ (the instantaneous concentration precipitated at the moment of m_{max}). Unfortunately, m_{max} is probably attained at about $(0.2\text{-}0.5)\times F_{plateau}$, as can be seen by crudely normalising signal data and plotting them on a linear scale (Figure I.33). Therefore one could guess that C_m was roughly between low-absorption and higher-absorption regimes, with different S↔C relationships, so a plateau-linear relation would strongly skew the calculation of $C_{PPT}(t)$ and thereby doubly skew the kinetic analyis. So I had to develop a more rigorous correlation than the linear relation, to monitor signal↔concentration in the range that I subsequently estimated at roughly $(0.1\text{-}0.35)\times C_{PPT}$; ie, approx. 4-30 µM.

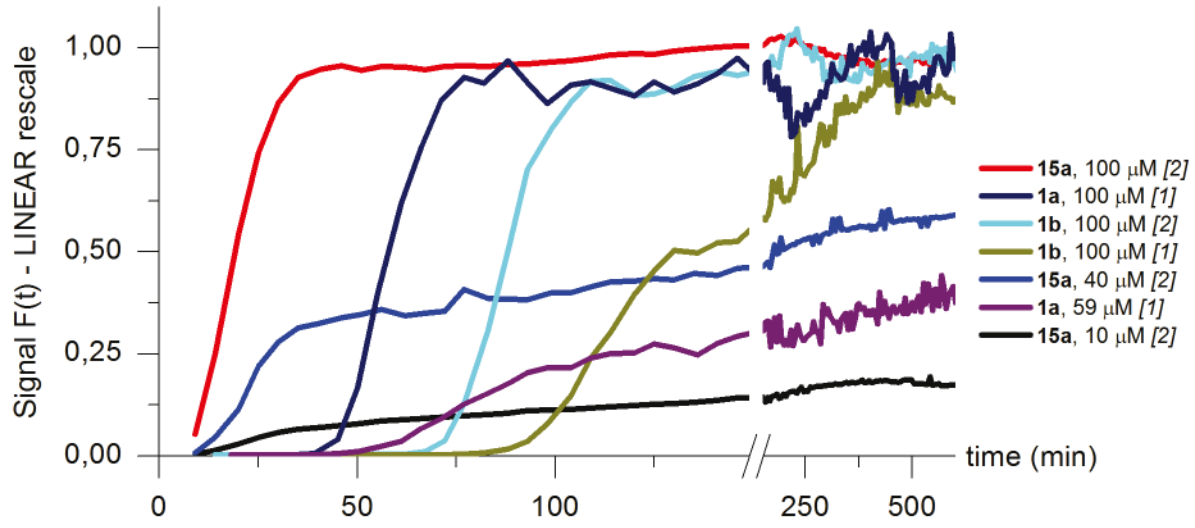


Figure I.33 – roughly normalised S(t) data show that the regions of greatest S(t) gradient are at about $(0.3\text{-}0.5)\times S_{plateau}$.

I normalised the only reliable steady-state signal↔concentration relations (at plateau) for fluorophore brightness, between runs, and optimised a concentration dependence function informed by my theoretical considerations of weak and strong absorption regimes and general signal↔concentration relations (see section I.3.3) to deliver a bimodal linear-exponential fit to the observed relations (full analysis in section III.3.3.4; plot in Figure I.34).

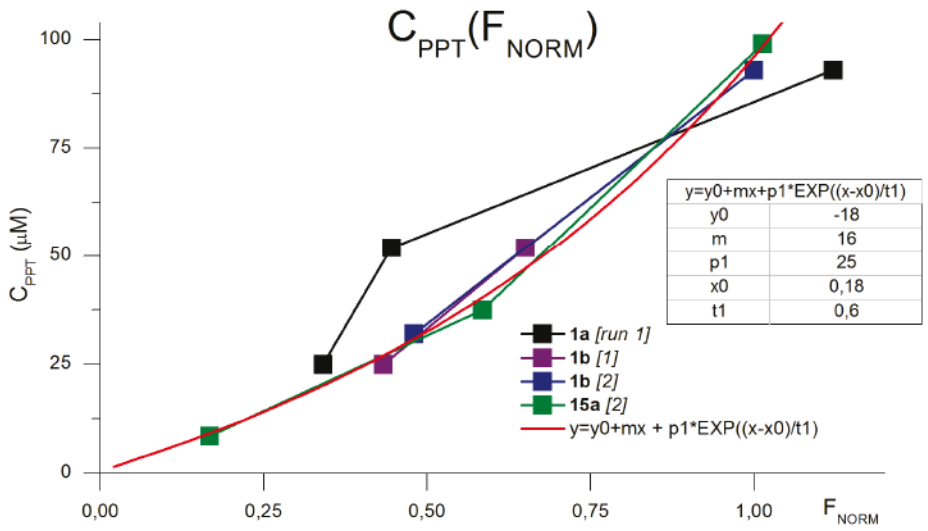


Figure I.34 – Concentration dependence of the ratio of plateau fluorescence signal to precipitated concentration, with my best-fit bimodal curve.

I could now use this linear-exponential relation to convert all the observed signals F(t) into precipitated concentrations of fluorophore over time $C_{PPT}(t)$. From there, I could *initially* model the concentration of released fluorophore as $[K_S+C_{PPT}](t)$, in an initial consideration of kinetic parameters.

5.5. *Enzyme test analyses: kinetic parameter determinations*

A much more robust kinetic analysis of the enzymatic processing parameters and spacer cyclisation kinetics could have been performed with a soluble fluorophore that gave an instant, linear response of fluorophore liberated↔signal produced. However, my involvement with the project was terminated before making soluble-fluorophore derivatives, so I instead performed more complex, but ultimately highly informative, analysis with the data for the precipitated fluorophore. To my knowledge this is the first explicit demonstration of a physically justified model for the deconvolution of the macroscopic signal response from a three-component probe which releases a precipitating fluorophore, to yield a justifiable evaluation of the microscopic kinetic parameters which in turn enables a numerical model which can reproduce the macroscopically observed multistep kinetic behaviour. Note that prior work in the group,[73] though molecularly similar in the experiment, had in the analysis probably not only failed to consider the signal↔concentration relation, but also was strongly suspected to have unreliable sample concentration (see section I.5.6.1), and did not provide explicit treatment of these non-obvious calculations nor determination of the multistep kinetic processes, which is why I claim this novelty for my work.

5.5.1. *Mechanistic assumptions for initial k_{CYC} determinations*

The three consecutive processes involved in signal generation are illustrated in Figure I.35, noting that the letters S,P,Q,R are dummy variables that can be adapted to fit any of the three probes. Explicitly these processes are: 1) enzymatic hydrolysis of the amide bond, ie. P⟶Q + leucine; 2) spacer

cyclisation releasing a urea and the fluorophore into solution, ie Q→R + urea; 3) when the solubility threshold K_S has been attained, precipitation of the amount S of fluorophore that is supersaturating the solution. Therefore, all three microscopic kinetic processes must be deconvoluted from the acquired macroscopic signal data, to achieve a reliable analysis of the kinetics of the reactions.

Because I reasoned that the cyclisation was under some circumstances the strongest rate-limiting step for generating the observed signal, I decided that I could at first simplify the observed kinetics to deconvolute reasonable lower-bound values of k_{CYC}, and then later use those values in a numerical model for the unsimplified data to give rough estimates for the convoluted enzymatic kinetic behaviour (section 5.5.2). Note that the enzymatic processing rate cannot under any assumptions be estimated from the activity (in units) reported by the supplier: these refer to an entirely different substrate under different conditions of pH and temperature (section III.3.2), and anyway they have a quoted standard deviation of 150% from the mean.

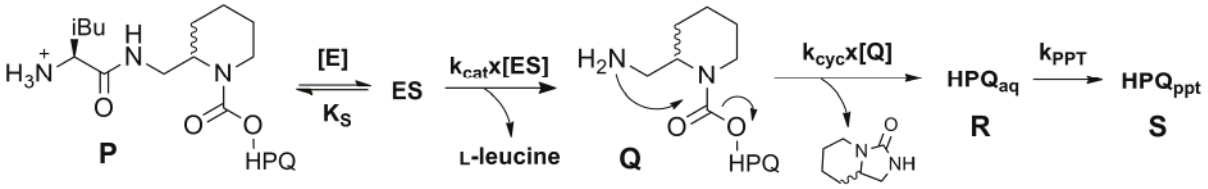


Figure I.35 – The steps involved in signal generation, as illustrated for probe 1a. An effective concentration of precipitated fluorophore [S] is generated by enzymatic hydrolysis of P to give intermediate Q which cyclises to release fluorophore into solution; once the fluorophore's accumulated concentration R is greater than K_S, precipitation to S is irreversible.

Recall that the signals for probes **1a** and **1b**, at the same concentration, reached the approximately the same fluorescence plateau values with similar-shaped signal response curves. This is expected if enzymatic hydrolysis is complete; therefore I concluded that the diastereomers of **1a** do not influence the enzymatic processing step. So, I then assumed, and justified based on calculations of the observed data, the following: (1) simple Michaelis-Menten kinetics for **P→Q**, (2) rate-determining cyclisation in the linear regions of maximum signal generation gradient, (3) the possibility of correcting data for supersaturation/precipitation kinetics by linear subtraction of a time t_{corr}, to enable initial lower-bound estimates for k_{CYC}. These assumptions required fairly involved mathematical/physical justifications, which I detail in full in sections III.3.4.1-III.3.4.3; notably, the logical optimisation of t_{corr} requires some careful thinking. The result of these operations is tabulated below (Table I.3); for full discussion of the required kinetic parameters, see III.3.4.

Kinetics: AFTER least t correction	t_{corr} (min)	t_{lag} (min)	$t_{C=0}$ (min)	t_{LIN} (min)	C_{LIN} (µM)	m_{max} (µM/min)	$k_{CYC,LB}$ (min^{-1})	t_{CYC} (min)	$t_{1/2}$ (min)
RUN 1:									
1a, 378 µM	17.5	11.5	22.5	27.5	50	10	0.0444	15.6	47.5
1a, 100 µM	30	9	22.5	25	10	4	0.0444	15.6	28
1a, 59 µM	9.1	29.9	22.9	61.9	13	0.33	0.0444	15.6	126.9
1b, 100 µM	28.8	37.2	61.2	83.2	27	1.2	0.0165	42.1	117.2
1b, 59 µM	37.7	55.3	60.3	87.3	12	0.45	0.0165	42	124.3
RUN 2:									
15a, 100 µM	0	0	10	14	16	4.1	0.099	7	22
15a, 40.5 µM	0	0	11	17	6	1	0.0909	7.6	33
1b, 100 µM	*56.4*	*10.6*	*19.6*	*28.6*	*30*	*3.5*	*0.0499*	*13.9*	*35*
1b, 40.5 µM	65	55	60	86	12	0.45	0.0169	41.1	102

Table I.3 - Values of the kinetic parameters t^*_{lag}, t^*_{LIN}, m_{max} and $t^*_{1/2}$, altered to fit the instantaneous precipitation kinetic model, and the cyclisation kinetics calculated from them, including upper-bound estimate for cyclisation halflife t_{CYC}.

Because this analysis forced the minimum physical value of t_{corr}, it gave a scientifically justifiable *lower bound* for k_{CYC}. This is in fact an advantage, because the original measure of the observed concentration halftime $t_{1/2}$ gives an indisputable *upper bound* for k_{CYC}. Therefore I have not only calculated a cyclisation rate value, but obtained a very strong confidence interval for it, which was necessary given the rather hair-raising data treatment that had to be performed since acquiring the original macroscopic raw signal data, A(t).

5.5.2. Enzyme test analysis 2: multistep, complete kinetic model

With physically reasonable starting estimates for k_{CYC} (and especially of the *ratios* between the k_{CYC} values) in hand, I returned to the non-t-adjusted concentration data to build a full numerical model for the concentration precipitated over time. As before, I assumed Michaelis-Menten kinetics for the enzymatic processing, and used the observed lower-bounds of $v_{MAX}=2\times t_{1/2}/P_0$ with the 378 µM dataset of **1a** and the 100 µM run 2 dataset of **1b**, ie. 2.9 $\mu M \cdot min^{-1}$ and 0.55 $\mu M \cdot min^{-1}$ for the piperidyl (**1a**, **15a**) and *gem*-dimethyl (**1b**) spacers respectively. Given the stability of the probes, I neglected the small rate contributions from spontaneous hydrolysis. I used the values of K_S determined previously as starting values. I also observed that from 378 µM to 100 µM, probe **1a** was operating in a large-P regime, therefore I could conclude that $K_M << 100$ µM as a starting value, and considered that the K_M for **1a** and **15a** should be near-identical, because of the identical structure of the piperidyl spacer-substrate pair at the 'business end' of the exopeptidase probe; and I expected a larger value for K_M for the dimethyl spacer of **1b**, due to its α-quaternary rather than α-tertiary structure with larger steric bulk near the scissile bond.

I could not directly use the model first developed in section I.3.1.2 because it neglects precipitation kinetics. As I was building a numerical rather than mechanistic model, aimed at evaluating whether the enzyme kinetic parameters could be roughly estimated while keeping reasonable model behaviour, the form of the precipitation rate was not a priority as long as the calculation was physically

reasonable and the results reproduced the observed data. Noting that the precipitation reaction is not unimolecular, especially when powders are precipitated rather than single crystals grown, and that precipitation kinetics often involve a rate equation as per *precipitation rate* = $k_{PPT}\times([R]-K_S)^p$ where p is from 3-5, I chose a value of p=3 which reproduced the observed behaviour satisfactorily.[106,107]

Using a short timestep (0.25 minutes) relative to the experimental timestep (5 min) gave an accurate enough approximation to enable stepwise integration (normalised to timestep size) as per:

$\rightarrow P(t+1) = P(t) - v_{max}\times P(t)/[K_M+P(t)]$

$\rightarrow Q(t+1) = Q(t) + v_{max}\times P(t)/[K_M+P(t)] - k_{CYC}\times Q(t)$

$\rightarrow \Delta precipitate = IF\ R(t)>K_S\ THEN\ k_{PPT}\times(R(t)-K_S)^3\ ELSE\ 0$

$\rightarrow R(t+1) = R(t) + k_{CYC}\times Q(t) - \Delta precipitate$

$\rightarrow S(t+1) = S(t) + \Delta precipitate$

By examining the generated curves S(t) and varying the kinetic parameters within sensible limits, I aimed to fit the following features of the experimentally observed signals: (1) the concentration halflives, (2) signal onset times, (3) the general form of the $C_{PPT}(t)$ curves (especially for higher concentration runs where an error in K_M would have less influence on the form of the results), (4) plateau values and (5) linear region gradients as closely as possible.

Note that especially, this kinetic model using *shared* parameters had to reproduce $C_{PPT}(t)$ behaviour observed for *different* probe data sets in *different* runs (potentially with different enzyme activities: see section III.3.2.7), with signal onset times ranging between <2 and >100 minutes (Figure I.36a, 100 μM vs panel c, 40 μM) and $C_{PPT}(t)$ to be fitted spanning 8-370 μM. These are diffcult conditions to fulfil, especially under the constraints of identical K_M and v_{max} for probes **15a** and **1a** on different runs. However there was no other method to ensure that the parameter fits remained intercomparable using only this data. The positive outcome to this unforgiving fit requirement was that signals would also be difficult to reproduce serendipitously, so even approximate fits ought to indicate useful molecular information. In the event, the fits obtained were remarkably good, both in form *and* in absolute value (Figure I.36).

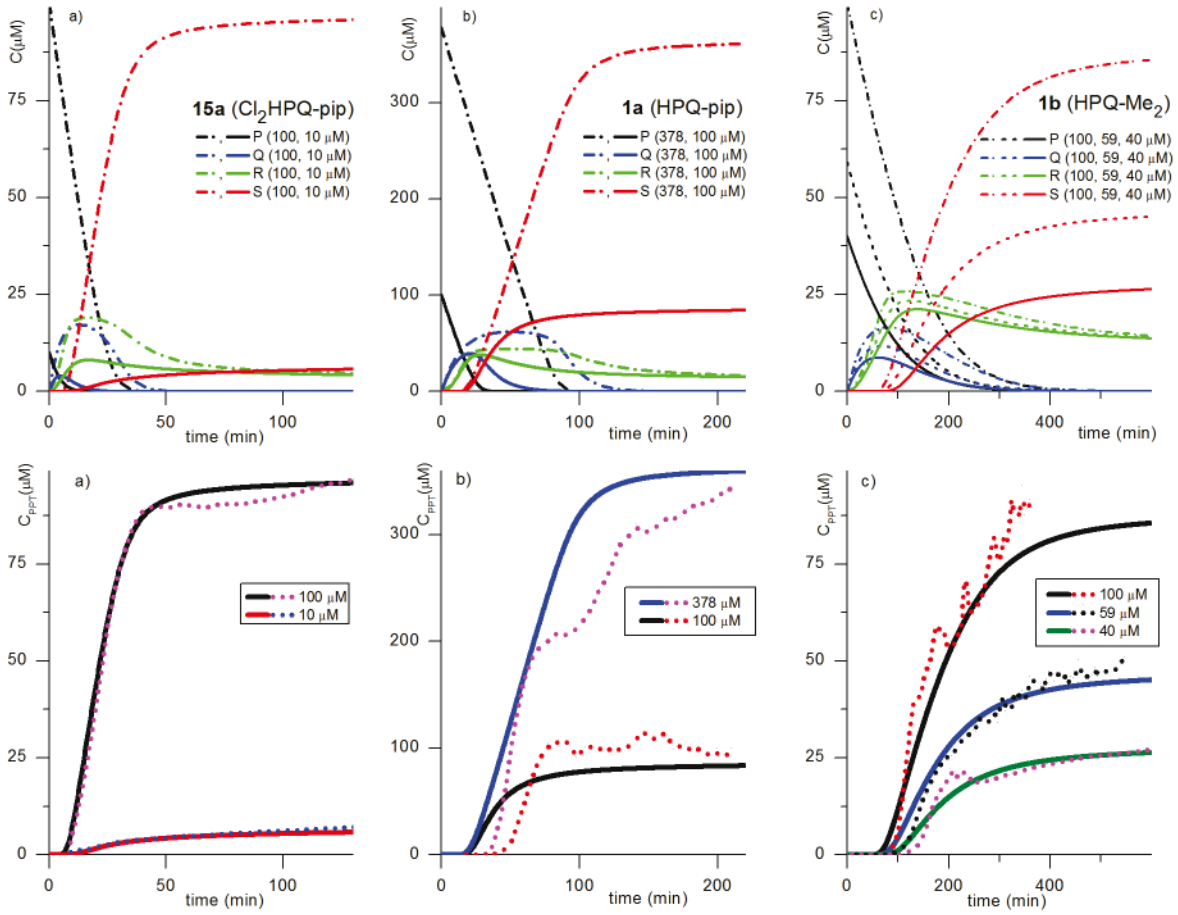


Figure I.36 – Upper panels: Modelled concentrations of P, Q, R and S for the three probes (a) 15a, (b) 1a and (c) 1b over time. Lower panels: comparison of the precipitated concentration fits S (solid lines) to the $C_{PPT}(t)$ curves calculated from experimental data (dotted lines). Multiple concentrations are shown for each probe. Experimental data curves are truncated upon first reaching their plateau values due to large scatter at high signal intensity.

Mathematically, the model had enough parameters that *unlikely* choices of one parameter could result in acceptable fits to the $C_{PPT}(t)$ calculated from the experimental data, if the other parameters compensated for its effects. Therefore I do not claim that its values are necessarily accurate. However, by enforcing the conditions that (a) all parameters be informed by the upper and lower bounds determined throughout the previous section, as well as (b) by the trends in the parameter bounds calculated above, and (c) by what was physically and mechanistically reasonable, I still obtained a fit which reproduced much of the desired behaviour with thoroughly physical values of the fitted parameters, which were similar to those indicated in the t-adjusted k_{CYC} analyses above.

Better still, this fit was difficult to reproduce by allowing small (+/- 30%) changes in the parameters: because the cyclisation step is rate-limiting but occurs between the two unknown rate processes and had a reliable starting value and confidence interval, the actual variations that could be produced by changing the other parameters were fairly small. Approximately speaking, for example, K_M and v_{max} determine the gradient of the linear region and to a lesser extent the onset time, so are more or less fixed especially by the observed signal gradient; k_{PPT} scales the data especially at higher signals.

Therefore the obtained macroscopic signal fit seems to represent a likely scenario of microscopic kinetic parameters. The final values used in the fits of Figure I.36 are tabulated in Table I.4.

	15a	**1a**	**1b**
K_M (μM)	15		60
v_{max} (μM·min^{-1})	5		1
k_{CYC} (min^{-1})	0.22	0.073	0.029
τ_{CYC} (min)	3.2	9.5	24
K_S fitted (μM)	2	10	
k_{PPT} (μM^{-2}·min^{-1})	7.5×10^{-4}	1.1×10^{-4}	

Table I.4 – Parameters for the full multistep model fits to the experimental data

Salient Features:

(1) K_M(**1a**|**15a**) < K_M(**1b**) and likewise v_{max}(**1a**|**15a**) < v_{max}(**1b**), each differing by a factor of 4-5. I believe this reflects better enzymatic recognition and steric access to the less crowded α-tertiary piperidyl spacer. Globally, the values would indicate rather favourable processing, especially of the piperidyl probes, and also correspond to reasonable values which are supported by the experimentally-determined lower bounds for v_{max}. Combining v_{max} and the well volume of 0.1 mL allows a calculation of the enzyme activity in the well: probes **1a** and **15a** experienced 5 μM·min^{-1}× 1×10^{-4} L/well = 0.5×10^{-3} μmol·min^{-1}/well conversion == 0.5 mU/well true enzyme activity.

(2) k_{CYC}(**15a**) > k_{CYC}(**1a**) > k_{CYC}(**1b**), each differing by a factor of 3. I believe this reflects both the better leaving-group capacity of the more electron-withdrawn **diClHPQ**, and the more favourable preorganisation of the piperidyl spacer which has fewer degrees of freedom to lose in cyclisation. The values are in all cases only double the lower-bound figures found during the t-corrective analysis, and more importantly, they preserve the original ratios found between those $k_{CYC,LB}$ values (0.099, 0.044, 0.016 respectively), so I consider them strongly supported. Values of the cyclisation halftime τ_{CYC} that these k_{CYC} imply are included for easier interpretation of just how fast the cyclisations are.

(3) The K_S values used are almost identical to those determined in steady-state analysis (1.5 and 7), which again reflects what is seen in the experiment, perhaps adding a small contribution for tolerated supersaturation. The values for k_{PPT} parameters are, pleasingly, inversely proportional to their solubility constants, which seems physically reasonable given their underlying meaning, as well as the qualitative observation of **diClHPQ**'s greater precipitation speed (ie lower crystallinity; section I.3.3). Note that the apparently small values of k_{PPT} in reality give quite favourable precipitation rates, because they multiply the cube of the supersaturation concentration. Therefore, these results finally do support the starting-point for my kinetic analyses, that the most strongly rate-limiting step in the overall kinetics is cyclisation, but that even so this has a very favourable halflife.

5.5.3. Conclusions from data analysis and modelling

I believe that while the fit parameters in Table I.4 are not overly precise, they do substantially reflect the true molecular underpinnings of the experiment. Their physical reasonableness coupled with the multistep microscopic model's good reproduction of the macroscopic experimental data, over wide

ranges of probe concentrations across different probe types with large ranges of onset times, argues that I have at last covered every crucial step from design, to implementation, analysis and understanding in this system:

- theoretical modelling of the detection behaviour and sensitivity that can be obtained with off-ON precipitating probes (sections I.3.1 – I.3.2) and the theoretical problems facing the analysis of their signal data (section I.3.3),

- rational design of reliable experiments avoiding the errors of prior work (sections I.5.1 and I.5.6.1);

- development and implementation of an involved signal treatment algorithm, for scientifically rigorous data processing (sections I.5.1.3 and III.3.2.5);

- extensive and non-obvious, but crucial, further treatment of the signal data to extract microscopic information: the nonlinear signal↔concentration relations, K_S values and halflives for spontaneous hydrolysis (sections I.5.3 – I.5.4);

- initial deconvolution of the experimental data to obtain good lower-bound analytical estimates for the most strongly-rate-limiting factor, k_{CYC} (section 5.5.1),

- implementation of a final multistep numerical model relying on the parameters identified as critical during these analyses, which generates fitted values in very close agreement both in trend and in absolute values with those obtained during piecewise analysis, and which are also physically reasonable and explicable (5.5.2). That the previous analyses, with their substantial simplifications, performed so closely to the full numeric model, also argues that their careful approach was well justified.

At last, this accumulated analysis will now permit some preliminary comparisons of these probe systems with prior literature designs (section I.5.6).

5.5.4. Future modifications to monitoring kinetics

The most crucial improvement to perform a reliable kinetic analysis would be using a soluble fluorophore, as mentioned in section I.5.5. HPLC analysis would potentially also help to discriminate between the different concentrations of the three soluble species (P, Q, R) over time, which would greatly improve the accuracy of the multiparameter kinetic fit.

5.6. Evaluation of probe performance

5.6.1. Comparison with prior art in the group

The most relevant and reliable comparison of the performance of the present precipitating peptidase probes would normally be with the team's prior art probes **34b** and **34a**, each for leucyl aminopeptidase and releasing the **HPQ** fluorophore. However, in section III.3.5.1, I show extensively that unfortunately, the data reported for these probes is complicated by several factors that raise doubts about the accuracy of the reported analyses. I have examined and recalibrated the reported data to

calculate what I believe to be a close estimate for the hydrolytic stability of those prior probes as τ(**34a**|**34b**)~13.3 h (no stability was previously calculated), which is orders of magnitude below the stability of the current HPQ-releasing probes (τ~100,000 h). This is to be expected for the relative stabilities of an acyloxy methyl ether of an electron-withdrawn, delocalised and anchimerically-assisting phenol, vs. the present probes' secondary carbamate of the same.

However, I concluded that no recalibration could allow me to compare enzymatic kinetic parameters K_M and v_{max}, because of the problems with the previous report's data and analysis. Notably, (a) their times of signal acquisition are estimated to be at least ten minutes shifted from the real start times; and worse, (b) the prior test solutions seem strongly to have been pre-spiked by free **HPQ** due to the sample preparation procedure employed. This pre-spiking entirely changes the kinetics of signal response, the concentrations which can be obtained, and the absolute values of the signal plateaux. Therefore in my consideration this invalidates simple comparison of the kinetic parameters which were reported for that system. Furthermore, (c) that prior work did not consider the strongly nonlinear signal↔concentration relation, and (d) did not detail its calculations of enzymatic quantities nor provide sufficient data to enable me to do so. My opinion is that the chief implicit assumptions behind their calculations were that the cyclisation of the spacers and the precipitation of the probes were near-instantaneous (unlikely), that signal↔concentration response was linear (contradicting what I have argued extensively above) and that 38 μM of probe were precipitated by plateau time (I recalculated a maximum of ~29 μM), and that neither pre-spiking nor start time needed consideration (which I argue strongly against).

I therefore concluded that the enzymatic quantities reported there are, unfortunately, not to be trusted. These reported enzymatic values were K_M(**34b**)=98 μM, v_{max}(**34b**)=3.1 μM·min^{-1}, at 25°C, 0.02 U·mL^{-1}, pH~7.4. As was seen during the fitting of the multiparameter model, convolution of hydrolysis with non-instantaneous cyclisation and precipitation kinetics would give a K_M skewed high, and a v_{max} skewed low. Comparing the probe structures, **34b**'s main difference is its secondary β-carbon, so by similar reasoning to section I.2.1.1, it ought to display even lower K_M than **1a**/**15a** (ie <15 μM) and even higher v_{max} (ie >> 5 μM·min^{-1} : both steric and temperature considerations). Therefore these reported quantities seem to support my analysis of their assumptions. As I consider these assumptions so great, and in any case the absolute difference between my predicted values and their reported ones so large, I consider that reliable comparison is impossible.

5.6.2. Comparisons with prior art in the literature

STABILITY: All the present probes synthesised were exceptionally stable, especially when compared to other electron-withdrawn, conjugated phenolic probes, and deliver exceptionally low control signals over time, despite not subtracting a time-zero background reading. The hydrolytic stability of the present probes is orders of magnitude better than the group's previous efforts: τ(**1a**|**1b**)>4151 days, whereas for the same **HPQ** fluorophore with the previous spacer design, τ(**34a**|**34b**)~13.3 h. Even

using the much more electron-withdrawn **diClHPQ**, τ(**15a**)=567 hours (24 days), which is almost two orders of magnitude better. The probes also compare outstandingly with other phenolic probes whose utility *in cellulo* and *in vivo* is often claimed: cf. Urano's glycosidase probe, τ~5 h (section III.3.1.1)[40]; fluorescein, Oregon Green or resorufin diacetates (τ~1.5 h, 6 min, and <2 min respectively[4]), or fluorescein diacetoxymethyl ether (τ~32 h)[4].

ENZYMATIC PROCESSING: While determination of K_M and v_{max} based on my triply convoluted data involves assumptions, I determined well-fitting and physically reasonable values that were in good absolute and trend agreement with my experimentally determined lower bounds for v_{max} and k_{CYC}, and a value of K_M that was well below the indicated upper bound. The fitted values for K_M and v_{max} were promising for eventual *in vivo* use of the probes (cf. the models in section 3.1.4), indicating very good, and therefore potentially competitive, enzymatic processing especially of the piperidyl spacer (K_M = 15 μM, and v_{max} = 5 μM·min^{-1} at 30°C, for an enzyme concentration giving a quoted enzymatic activity 50-200 μM·min^{-1} for *p*-nitroanilide at saturation and *at 37°C*), though the dimethyl spacer also performed acceptably (K_M = 60 μM, and v_{max} = 1 μM·min^{-1}). Especially, the low K_M values can be compared with those for other rapidly-processed probes but bearing superior leaving groups, such as fluorescein diacetate (K_M ~11 μM with porcine liver esterase)[4].

True comparisons of the k_{cat} are difficult because even anilides such as the enzyme's quoted standard substrate are chemically markedly different leaving groups, whereas aliphatic amines were not found with enough data to compare to. However, a characterisation of leucyl-AMC probe **5** under similar circumstances (T=25°C, pH~7.5) but adjusting the v_{max} for enzyme concentration would give K_M(**5**) = 40 μM, v_{max}(**5**) = 0.04 μM·min^{-1}. The apparently low v_{max} *could* indicate a very strong temperature dependency of this enzyme. However, an interesting observation by de Groot[21] is that using the biologically significant exopeptidase plasmin, his aliphatic amide based on an N'-methylethylenediamine spacer gave an order of magnitude faster enzymatic rate than a PABA anilide. This effect could partially explain the far better v_{max} for the present probes compared to previous designs such as **5**. Therefore especially if enzymatic concentration is at low and limiting *in vivo* levels, the greater processing speed of the current probe/prodrug system could be crucial in recommending the current spacer technology over the less biomimetic but much more developed PABA spacer (from the point of view of competitive *in vivo* rate, let alone toxicity), and in recommending this spacer's three-component probe design over two-component exopeptidase designs such as **5** (competitive rate + off-ON).

SPACER KINETICS AND FUNCTION: I expected very different values for k_{CYC}(**1a**) and k_{CYC}(**15a**), because of the markedly superior leaving-group character of **diClHPQ**; and also very different k_{CYC}(**1b**) because of the different preorganisation of the spacer. In the event, I obtained the following relations by my analytical t-correction model: run 1: k_{CYC}(**1a**) ≈ 3× k_{CYC}(**1b**), which I attributed to both the higher population of the reactive rotamer in the groundstate of the annelated

spacer and to the lower loss of degrees of freedom during cyclisation by the annelated spacer; run 1 vs run 2: k_{CYC}(**15a**) ≈ 2.5×k_{CYC}(**1a**), which I attributed to the better leaving group character of the *p*-chlorinated phenol. I was surprised by this result, as I was expecting *substantially* greater difference in kinetics depending on the leaving group: but it would therefore seem that the *cyclisation preorganisation is more important in achieving rapid response than the leaving-group character*. Recall that these values are in exactly the same ratios as for the lower-bound analysis, which supports their physicality. Absolute kinetic values: the piperidyl spacer gave a rate estimate of $k_{CYC}=7.3\times10^{-2}$ min^{-1}, while the *gem*-dimethyl spacer gave a rate of $k_{CYC}=2.9\times10^{-2}$ min^{-1}, for cyclisation liberating the **HPQ** fluorophore at 30°C; the k_{CYC} for **diClHPQ** liberation by the piperidyl spacer was 0.22 min^{-1}.

The best prior-art enzymatically processable 1,2-diamine spacer (**12**) has $k_{CYC}=2.3\times10^{-3}$ min^{-1}, and the non-enzymatically-processable and much more N-nucleophilic spacer **13** has only $k_{CYC}=1.9\times10^{-2}$ min^{-1}, each measured with a more favourable temperature of 37°C albeit with a less favourable *p*-methoxyphenolate leaving group[70]. I therefore conclude strongly that both the present preorganised spacers **21** and **22** give much more rapid response than the best prior art in stable aminopeptidase spacers and even for general N→OPh spacers, since there is *at least* a clear order of magnitude difference in the rates, which should probably hold true when comparable leaving groups are used at the same temperature because of the opposite weighting that the factors apply (recall too the comparatively low rate enhancement of ×2.5 produced by using the far more active leaving group **diClHPQ** rather than **HPQ**).

The present preorganised spacers therefore present markedly superior cyclisation rates and should therefore find immediate and wide-ranging applications wherever current diamine spacers are used. For an enzymatically-processable diamine spacer there can be no competition: these preorganised spacers give almost two orders of magnitude better response time than the prior art, which is critical for *in vivo* performance. For example, the N'-methylethylenediamine spacer of de Groot's paclitaxel prodrug retards its liberation strongly (τ~23 hours)[21], which is why although the PABA spacer gave an order of magnitude slower enzymatic processing (τ_{enz}~43 min with high *in vitro* concentrations, 200 μM prodrug and 100 μg/mL plasmin), he was forced to retain the PABA strategy (near-instantaneous fragmentation) as the most suitable one available for *in vitro* tests (note: high enzyme activity, uncompetitive setting). The current spacer technology clearly addresses this need to bridge the gap between well-processed and rapidly-responsive spacers, and therefore should allow *in vivo* applications at small enzyme activities.

SIGNAL RESPONSE: The lower precipitation threshold of the **diClHPQ** fluoro and its superior precipitation kinetics resulted in significantly earlier signal onset (<1 min even at 10 μM) than for the **HPQ** probes. This onset time is highly competitive with all reported probes, although the caveat is that because the **HPQ**-based system relies on precipitation, the initial concentration of probe cannot be

reduced much below 1.5×K_S or else signal onset is slow (cf. **1a** at 10 μM: with its higher solubility threshold, clear signal onset took around 6 hours). Once precipitation starts, the exceptional signal strength of the fluorophore ensures a strong and unmistakeable response that cleared the level of background noise fluctuations by a factor of around up to 20,000 even on these unoptimised acquitions (Table I.1); its photostability further ensures that despite repeated irradiations over many hours, this exceptionally high signal remains entirely undiminished over more than 16 hours. The brightness of these model fluorophores is entirely comparable to workhorses such as fluorescein and resorufin, but the **HPQs**' exceptional photostability (at least two orders of magnitude slower photobleaching)[17] is an unbeatable advantage for absolute signal levels.

SIGNAL-TO-CONTROL SENSITIVITY: The signal-to-control ratios for the present probes are absolutely enormous. Without even subtracting the t=0 fluorescence background signal, under unoptimised conditions, after 11-16 hours, one still obtains signal-to-control ratios of 618(**1a**), 2944(**1b**), and 1655(**15a**): consider that in the theoretical limit, the maximum signal-to-control ratio ever observable for relative change probes is *n*, the activation signal-enhancement factor, commonly with $5<n<50$. Performing t=0 subtraction, as is routinely done by almost all *in vitro* tests, results in even more impressive ratios for the current system: 5712(**1a**), 21000(**1b**), and 43000(**15a**), which should be contrasted with anything obtainable from relative change probes (or probes reported as off-ON but which are *not*, actually, truly off-ON): for example, Urano's ratio of ~25 after only 8 minutes and with 500-fold higher enzyme concentration (see section III.3.1.1). It should also be noted that the present signal-to-control values are static over time because the probe neither diffuses away nor photobleaches, both of which are common problems with soluble fluorophores, certainly over such long incubation/irradiation times. The **HPQs**' enormous Stokes shifts, as well as the probes' unique, macroscopically binary off-ON signal response, is responsible for this outstanding feature.

DETECTION SENSITIVITY: Research on fluorogenic probes commonly cites a 'detection limit' for the protein targeted, which is a function of the ratio between control and active fluorescence signals at a fixed time t. In this case, the situation is altogether different: the principle of a true binary off-ON system is to give an *arbitrarily low* detection limit for enzyme activity, by allowing longer incubation to compensate for lower activity without any false positive signal masking the results. The Hasserodt design provides the first demonstration of such a fluorogenic aminopeptidase probe. Therefore the concept of a detection limit should no longer be taken as an intrinsic limitation of this system, but rather as an expected value which can be optimised by probe design following this pattern. Despite this caveat, in section III.3.5.2, I calculate anyway **24-hour** detection thresholds for probe **1a** at 100 μM as **8×10^{-6} UmL^{-1}**, and **15a** at only 35 μM as **8×10^{-7} UmL^{-1}**, of leucine aminopeptidase under the present experimental conditions (eg. non-optimal temperature of 30°C not 38°C).

These can be *very* favourably compared with existing probes; and the possibilities for increasing sensitivity depend on designing still-lower-solubility fluorophores for use in yet-more hydrolytically-

stable probes, with ever lower inherent background signal due to still higher Stokes shifts. Then the sensitivity of the off-ON precipitating approach will *continue* to increase, and it is already well beyond what is feasible for any relative change probe, or any current soluble system (sections I.3.1 – I.3.2).

OTHER SALIENT POINTS: (a) The fluorophores' insolubility is key to allow localisation of the site of enzymatic activity, and buildup of signal without signal bleed or diffusion losses; (b) the precipitation mechanism also enforces true mechanistic off-ON behaviour which is *very* rare by electronic modulation alone, but which is crucial for achieving the high signal-to-control ratios that the current work so amply demonstrates (cf. section I.3.1); (c) the very high Stokes shift should allow these probes to retain almost all of their exceptional sensitivity when transferred to *in cellulo* studies and beyond, and preliminary results indicate that indeed they do (cf. section I.3.2). These points should be compared against the markedly poorer performance of many systems that appear to show good behaviour *in vitro*, but are useless already *in cellulo*, let alone *in vivo*, because they fail to fulfill these criteria.[4,17]

General Conclusions from Part I

The novel diamine spacers designed and implemented in this work have for the first time allowed the robust and practical use of a phenolic fluorophore in an off-ON system for the ultrasensitive detection of aminopeptidase activity. The excellent performance of the probe system can be attributed to both the spacer designs and to the precipitating, ESIPT fluorophore whose use has now been allowed.

Spacers: Many prodrugs[21] and auto-immolative polymers[108] and dendrimers[109] have relied on 1,2-diamines as auto-immolative spacers. These have been restricted to either *adapter* spacers (**13**), or *enzyme-recognising* spacers (**12**). These prior art spacers have (a) somewhat slower (**13**, 36 min halflife, 37 °C) or vastly slower (**12**, 5 h) post-triggering cyclisation halftimes[70] compared to the present spacers (**22**, 9 min with **HPQ**, and **21**, 23 min likewise; 30 °C); (b) the prior spacers probably[86] have no significant enzyme-recognition advantages (for **12**), and certainly **13** cannot be recognised at all[21,110], whereas the present spacers present excellent enzymatic recognition and good processing kinetics (**22**: K_M ~15 μM, v_{max} ~5 μM·min^{-1} with 0.1 U·mL^{-1} enzyme at 30°C); and (c) the prior spacers probably give no greater stability than the current spacers, which have halflives of up to >4000 days (**22** with the electron-withdrawn **HPQ**).

Therefore the present cyclisation-preorganised, low-steric-demand, N'-primary-amine spacers represent a **highly significant advantage** over both prior best designs in terms of the speed of post-triggering response (a vital end-user criterion). For enzyme-triggered applications, the improvement in response kinetics is by almost two orders of magnitude – a crucial factor for successful *in vivo* implementation. This enormous improvement is provided without sacrificing any measure of stability, and likely without losing much enzymatic triggering speed relative to **12**. The spacers are also cheap, and are easily regio-differentiated. Therefore this author cannot imagine situations where they cannot immediately be used to replace both the prior *adapter*, and especially *enzyme-recognising* diamines, without bringing significant advantages at no functional cost.

To give one example, these new diamine spacers are an excellent solution for a robust, yet quickly-processed and rapidly responsive amine-substrate-to-alcohol-release spacer system such as was unsuccessfully sought by eg. Vrudhula *et al.*, working on Paclitaxel prodrugs,[111] where "the rate-limiting cyclization of the linker ($t_{1/2}$ ~50-230 min [**13**]) still proved to be an issue that might compromise any possible *in vivo* applications of such a strategy."[112] It also supersedes the Hasserodt group's previous efforts in aminopeptidase probe designs.[73] Even outside probes and prodrugs, these diamines may find applications in eg. kinetic studies, as a better testbed than previously-used **13** for examining the auto-immolation of other upstream spacers[12] in stacked designs, since the current spacers' degradation on a significantly shorter timescale improves the reliability of kinetic analysis by deconvoluting the responses.

It is also hoped that the highly thermodynamically stable cyclic urea byproduct given off by these spacers may also present minimal interference with the enzyme's and the cell's normal functionality. This is in sharp contrast to the well-known toxicity[34] and alkylating[66] behaviour of 4-hydroxy- or 4-aminobenzylic auto-immolative spacer byproducts[113-115] – Papot[68] reports that quinone methides may even have around fourfold higher toxicity than the combination of a pair of chemotherapeutics (doxorubicin + MS275; although the additive effect is unclear). As avoiding quinone methide spacers is vital for diagnostics *in vivo*, these far less toxic diamine spacers - which may also give both kinetic advantages when targeting native aminopeptidases[21], and greater stability with electron-withdrawn phenols[30] - seem to be strongly supported as robust, general solutions for connecting amine terminal release to the downstream release of a general phenol or alcohol drug or fluorophore.

Fluorophores: For envisaged uses of these probes in Molecular Imaging *in vitro* or *in cellulo*, the remarkable insolubility of their ESIPT fluorophores should allow signal accumulation at the site of enzymatic activity over time without signal bleed, and continuous diffusion of an off-state probe into cells or tissues containing an active target enzyme may label them with increasing amounts of fluorescence over time. This fluorophore insolubility further complements their brightness and photostability, as a means of accumulating as strong a signal as possible at the site of enzymatic activity; for example, the **diClHPQ** fluorophore is already well-known to give subcellularly-resolved and –localised imaging of enzyme activity with the phosphatase probe ELF-97[17]. In addition, this insolubility may reduce their perturbation of the native/cellular chemistry.

In general conclusion: The excellent microscopic kinetic, and macroscopic signal/sensitivity parameters which I have determined in this PhD support hopes that this highly robust, modular, off-ON system will open up new perspectives for the ultrasensitive, subcellularly-resolved detection and quantification of the enzymatic activity of the vast and important class of peptidases.[1,28,116]

Future Directions

Ongoing research in the group on the present work's 3-component probe structures, especially by Maxime Prost, has already validated this probe system *in cellulo*, showing good cell penetration of the proagent (cf. considerations such as druglikeness) and the desired precipitation of the activated fluorophore in the cell interior. Crucially, the probes currently being explored bear different substrates of greater biological relevance, eg. tri/tetra/hexapeptides of interest for targeting classic cancer-associated peptidases. A disadvantage of the insoluble ESIPT fluorophores used in this PhD is the necessity of exciting them in the near UV (around 360 nm). For *in vivo* applications, choosing or designing a fluorophore with visible wavelength excitation and emission in the red or near IR, would be preferred to minimise tissue damage and maximise imaging depth[117]. Therefore current research in

the group is also aiming towards *in vivo* applications of similar probes bearing longer-wavelength off-ON fluorophores, for imaging in the biological transparency window as well as multiplexing (eg. novel ESIPT-type fluorophores similar to **BTQC**, as well as non-ESIPT cores such as **3H-DCDHF** derivatives). Any such longer-wavelength-excitable fluorophores must still maintain both insolubility and photostability, to profit from similar sensitivity as has been proven for the current probes.

To demonstrate the broader possibilities of the peptidase-triggered phenol-releasing design, I selected the powerful phenolic anticancer compound Combretastatin A-4[118] as a core phenol to be released from a 3-component prodrug. This choice has several advantages, chiefly: (1) combretastatin's clinical use was essentially blocked for almost 20 years due to its notorious insolubility/biounavailability, but the present 3-component design enforces both solubility and bioavailability, as well as allowing specific targeting (which the two currently available prodrugs do not); (2) combretastatin's antimitotic activity is entirely blocked if the phenol is substituted with a moderate-sized group, so a pharmacologically off-ON design should be ensured;[119,120] (3) combretastatin's electron-rich tetramethoxystilbene structure should ensure that the prodrug shows far greater spontaneous hydrolytic stability than even **1a**. I again chose leucine as a simple model substrate; leucine aminopeptidase plasmids are available and will be used as positive controls in the *in cellulo* testing. I also retained the efficient piperidyl spacer as the auto-immolative linker. The combretastatin core was synthesised and delivered by Jean-Christophe Mulatier, and I performed phosgenation coupling to give 84% isolation of the Boc-protected combretastatin prodrug. *In cellulo* proof-of-concept tests evaluating the correlation of toxicity and model enzyme expression are scheduled for mid-late June 2013. Prodrug strategies involving other active agents and different peptide sequences are also pursued by other group members.

Lastly, the design of 'more intelligent' spacers is also a team priority. For example, spacers which mimic one or even multiple post-scissile residues, but retain the key hydrolytic stability of the current designs, are already under evaluation. These hope to allow for targeting endopeptidases by very similar technology, and therefore to profit from the exceptional stability and signal response that these designs can give. Other spacers which may allow logic-gate functions, such as an OR-gate using a 2,6-bis(aminomethyl)piperidine spacer baited for two different peptidases, or even dendrimer-like adaptions (eg. piperazine-based spacers) are also clear avenues for further research.

Key publications – Part I

- **Probes - Design Requirements and State of the Art:** Boonacker E, Van Noorden CJF. Enzyme Cytochemical Techniques for Metabolic Mapping in Living Cells, with Special Reference to Proteolysis. *J. Histo. Cyto.* **2001**, *49*, 1473-1486 (DOI weblink: 10.1177/002215540104901201).
- **Cyclisation Spacer Use:** de Groot FMH, van Berkom LWA, Scheeren HW. Synthesis and Biological Evaluation of 2'-Carbamate-Linked and 2'-Carbonate-Linked Prodrugs of Paclitaxel: Selective Activation by the Tumor-Associated Protease Plasmin. *J. Med. Chem.* **2000**, 43, 3093-3102 (DOI weblink: 10.1021/jm0009078).
- **Probe Stability and General Use:** Lavis LD, Chao TY, Raines RT. Synthesis and Utility of Fluorogenic Acetoxymethyl Ethers. *Chem. Sci.* **2011**, 2, 521-530 (DOI weblink: 10.1039/c0sc00466a).
- **Requirements for Auto-Immolative Spacers:** Richard JA, Meyer Y, Jolivel V, Massonneau M, Dumeunier R, Vaudry D, Vaudry H, Renard PY, Romieu A. Latent Fluorophores Based on a Self-Immolative Linker Strategy and Suitable for Protease Sensing. *Bioconjugate Chem.* **2008**, 19, 1707-1718 (DOI weblink: 10.1021/bc8001997).
- **Publication of Selected Results:** Thorn-Seshold O, Vargas-Sanchez M, McKeon S, Hasserodt J. A Robust, High-Sensitivity Stealth Probe for Peptidases. *Chem. Commun.* **2012**, 48, 6253-6255 (DOI weblink: 10.1039/C2CC32227G; reproduced after Part III – Experimental).
- **Probes - State of the Art:** Wysocki LM, Lavis LD. Advances in the Chemistry of Small Molecule Fluorescent Probes. *Curr. Opin. Chem. Biol.* **2011**, 15, 752-759 (DOI weblink: 10.1016/j.cbpa.2011.10.013).

Bibliography – Part I

DOI are indicated where possible (and failing that, PubMed ID [PMID]), for fastest lookup.

[1] E. Boonacker, C. J. F. Van Noorden; *J. Histo. Cyto.* **2001**, *49*, 1473-1486 *(10.1177/002215540104901201)*.

[2] M. Baker; *Nature* **2010**, *463*, 977-980 *(10.1038/463977a)*.

[3] J.-P. Goddard, J.-L. Reymond; *Curr. Opin. Biotechnol.* **2004**, *15*, 314-322 *(10.1016/j.copbio.2004.06.008)*.

[4] L. D. Lavis, T. Y. Chao, R. T. Raines; *Chem. Sci.* **2011**, *2*, 521-530 *(10.1039/c0sc00466a)*.

[5] L. M. Wysocki, L. D. Lavis; *Curr. Opin. Chem. Biol.* **2011**, *15*, 752-759 *(10.1016/j.cbpa.2011.10.013)*.

[6] M. Ogris, D. Oupicky, C. Rome, J. Gravier, M. Morille, G. Divita, A.-L. Bolcato-Bellemin, V. Josserand, J.-L. Coll; in *Nanotechnology for Nucleic Acid Delivery*, (*Vol. 948*), Humana Press, **2013**, pp. 49-65.

[7] M. Keramidas, V. Josserand, J. J. Feige, J. L. Coll; *Mol. Imaging Biol.* **2012**, 1-6 *(10.1007/s11307-012-0595-6)*.

[8] A. Koenig, L. Herve, G. Gonon, V. Josserand, M. Berger, J. M. Dinten, J. Boutet, P. Peltie, J. L. Coll, P. Rizo; *J. Biomed. Opt.* **2010**, *15*, 016016 *(10.1117/1.3309738)*.

[9] J. Mütze, V. Iyer, John J. Macklin, J. Colonell, B. Karsh, Z. Petrášek, P. Schwille, Loren L. Looger, Luke D. Lavis, Timothy D. Harris; *Biophys. J.* **2012**, *102*, 934-944 *(10.1016/j.bpj.2011.12.056)*.

[10] J. V. Frangioni; *Curr. Opin. Chem. Biol.* **2003**, *7*, 626-634 *(10.1016/j.cbpa.2003.08.007)*.

[11] L. D. Lavis, R. T. Raines; *ACS Chem. Biol.* **2008**, *3*, 142-155 *(10.1021/cb700248m)*.

[12] Y. Meyer, J.-A. Richard, B. Delest, P. Noack, P.-Y. Renard, A. Romieu; *Org. Biomol. Chem.* **2010**, *8*, 1777-1780 *(10.1039/B926316K)*.

[13] K. Kominami, T. Nagai, T. Sawasaki, Y. Tsujimura, K. Yashima, Y. Sunaga, M. Tsuchimochi, J. Nishimura, K. Chiba, J. Nakabayashi, K. Koyamada, Y. Endo, H. Yokota, A. Miyawaki, N. Manabe, K. Sakamaki; *PLoS One* **2012**, *7*, e50218 *(10.1371/journal.pone.0050218)*.

[14] R. Weissleder, C. H. Tung, U. Mahmood, A. Bogdanov, Jr.; *Nat. Biotechnol.* **1999**, *17*, 375-378 *(10.1038/7933)*.

[15] A. V. Zaytsev, R. J. Anderson, A. Bedernjak, P. W. Groundwater, Y. Huang, J. D. Perry, S. Orenga, C. Roger-Dalbert, A. James; *Org. Biomol. Chem.* **2008**, *6*, 682-692 *(10.1039/B716978G)*.

[16] S. Izumi, Y. Urano, K. Hanaoka, T. Terai, T. Nagano; *J. Am. Chem. Soc.* **2009**, *131*, 10189-10200 *(10.1021/ja902511p)*.

[17] V. B. Paragas, J. A. Kramer, C. Fox, R. P. Haugland, V. L. Singer; *Journal of Microscopy* **2002**, *206*, 106-119 *(10.1046/j.1365-2818.2002.01017.x)*.

[18] A. T. Ali, C. B. Penny, J. E. Paiker, C. van Niekerk, A. Smit, W. F. Ferris, N. J. Crowther; *Clin. Chim. Acta* **2005**, *354*, 101-109 *(10.1016/j.cccn.2004.11.026)*.

[19] S. S. Chandran, K. A. Dickson, R. T. Raines; *J. Am. Chem. Soc.* **2005**, *127*, 1652-1653 *(10.1021/ja043736v)*.

[20] S. M. Hacker, N. Hardt, A. Buntru, D. Pagliarini, M. Mockel, T. U. Mayer, M. Scheffner, C. R. Hauck, A. Marx; *Chem. Sci.* **2013**, *4*, 1588-1596 *(10.1039/C3SC21916J)*.

[21] F. M. H. de Groot, L. W. A. van Berkom, H. W. Scheeren; *J. Med. Chem.* **2000**, *43*, 3093-3102 *(10.1021/jm0009078)*.

[22] S. K. Grant, J. G. Sklar, R. T. Cummings; *J. Biomol. Screen.* **2002**, *7*, 531-540 *(10.1177/1087057102238627)*.

[23] M. Shamis, H. N. Lode, D. Shabat; *J. Am. Chem. Soc.* **2004**, *126*, 1726-1731 *(10.1021/ja039052p)*.

[24] B. Renoux, T. Legigan, S. Bensalma, C. Chadeneau, J.-M. Muller, S. Papot; *Org. Biomol. Chem.* **2011**, *9*, 8459-8464 *(10.1039/C1OB06081C)*.

[25] G. B. Jones, C. F. Crasto, J. E. Mathews, L. Xie, M. O. Mitchell, A. El-Shafey, A. V. D'Amico, G. J. Bubley; *Bioorg. Med. Chem.* **2006**, *14*, 418-425 *(10.1016/j.bmc.2005.08.015)*.

[26] N.-H. Ho, R. Weissleder, C.-H. Tung; *Bioorg. Med. Chem. Lett.* **2006**, *16*, 2599-2602 *(10.1016/j.bmcl.2006.02.045)*.

[27] U. Mahmood, C. H. Tung, A. Bogdanov, Jr., R. Weissleder; *Radiology* **1999**, *213*, 866-870 *(PMID: 10580968)*.

[28] B. Law, C.-H. Tung; *Bioconjugate Chem.* **2009**, *20*, 1683-1695 *(10.1021/bc800500a)*.

[29] E. Matayoshi, G. Wang, G. Krafft, J. Erickson; *Science* **1990**, *247*, 954-958 *(10.1126/science.2106161)*.

[30] J. A. Richard, Y. Meyer, V. Jolivel, M. Massonneau, R. Dumeunier, D. Vaudry, H. Vaudry, P. Y. Renard, A. Romieu; *Bioconjugate Chem.* **2008**, *19*, 1707-1718 *(10.1021/bc8001997)*.

[31] Y. Yang, P. Babiak, J.-L. Reymond; *Org. Biomol. Chem.* **2006**, *4*, 1746-1754 *(10.1039/B601151A)*.

[32] O. Redy, E. Kisin-Finfer, E. Sella, D. Shabat; *Org. Biomol. Chem.* **2012**, *10*, 710-715 *(10.1039/C1OB06667F)*.

[33] J. H. Ryu, S. A. Kim, H. Koo, J. Y. Yhee, A. Lee, J. H. Na, I. Youn, K. Choi, I. C. Kwon, B.-S. Kim, K. Kim; *J. Mater. Chem.* **2011**, *21*, 17631-17634 *(10.1039/C1JM13064A)*.

[34] J. D. Sellars, M. Landrum, A. Congreve, D. P. Dixon, J. A. Mosely, A. Beeby, R. Edwards, P. G. Steel; *Org. Biomol. Chem.* **2010**, *8*, 1610-1618 *(10.1039/B920443A)*.

[35] N. Ziegler, J. Bätz, U. Zabel, M. J. Lohse, C. Hoffmann; *Bioorg. Med. Chem. Lett. 19*, 1048-1054 *(10.1016/j.bmc.2010.07.060)*.

[36] H. Yao, S. Jin; *Sensors* **2012**, *12*, 16759-16770 *(10.3390/s121216759)*.

[37] T. Förster; *Annalen der Physik* **1948**, *437*, 55-75 *(10.1002/andp.19484370105)*.

[38] M. Beija, C. A. Afonso, J. M. Martinho; *Chem. Soc. Rev.* **2009**, *38*, 2410-2433 *(10.1039/b901612k)*.

[39] M. Briggs, I. Bruce, J. Miller, C. Moody, A. Simmonds, E. Swann; *J. Chem. Soc., Perkin Trans. 1* **1997**, 1051-1058 *(10.1039/A605012C)*.

[40] M. Kamiya, D. Asanuma, E. Kuranaga, A. Takeishi, M. Sakabe, M. Miura, T. Nagano, Y. Urano; *J. Am. Chem. Soc.* **2011**, *133*, 12960-12963 *(10.1021/ja204781t)*.

[41] Y. Z. Yang, P. Babiak, J. L. Reymond; *Helv. Chim. Acta* **2006**, *89*, 404-415 *(10.1002/hlca.200690041)*.

[42] E. Leroy, N. Bensel, J.-L. Reymond; *Bioorg. Med. Chem. Lett.* **2003**, *13*, 2105-2108 *(10.1016/S0960-894X(03)00377-9)*.

[43] N. Bensel, M. T. Reymond, J.-L. Reymond; *Chem.–Eur. J* **2001**, *7*, 4604-4612 *(10.1002/1521-3765(20011105)7:21<4604::aid-chem4604>3.0.co;2-z)*.

[44] W. Kotka, Z. Ptaki, D. Antwoord; *BioChemPhys* **2009**, 82-83 *(10.1007/s13305-02f-6183-x)*.

[45] M. N. Levine, R. T. Raines; *Chem. Sci.* **2012**, *3*, 2412-2420 *(10.1039/C2SC20536J)*.

[46] L. Stryer, J. M. Berg, J. L. Tymoczko; *Biochemistry,* Freeman & Company Limited, W H **2002**.

[47] National Research Council; *Visualizing Chemistry: The Progress and Promise of Advanced Chemical Imaging,* The National Academies Press, Washington, D.C., **2006**.

[48] P. Hermann, J. Kotek, V. Kubicek, I. Lukes; *J. Chem. Soc., Dalton Trans.* **2008**, 3027-3047 *(10.1039/B719704G)*.

[49] K. Komatsu, Y. Urano, H. Kojima, T. Nagano; *J. Am. Chem. Soc.* **2007**, *129*, 13447-13454 *(10.1021/ja072432g)*.

[50] Y. Gabe, Y. Urano, K. Kikuchi, H. Kojima, T. Nagano; *J. Am. Chem. Soc.* **2004**, *126*, 3357-3367 *(10.1021/ja037944j)*.

[51] T. Komatsu, K. Kikuchi, H. Takakusa, K. Hanaoka, T. Ueno, M. Kamiya, Y. Urano, T. Nagano; *J. Am. Chem. Soc.* **2006**, *128*, 15946-15947 *(10.1021/ja0657307)*.

[52] G. Wyszecki; *J. Opt. Soc. Am.* **1963**, *53*, 1318-1319 *(10.1364/JOSA.53.001318)*.

[53] J. Zhao, S. Ji, Y. Chen, H. Guo, P. Yang; *Phys. Chem. Chem. Phys.* **2012**, *14*, 8803-8817 *(10.1039/C2CP23144A)*.

[54] S. M. Ormson, R. G. Brown; *Prog. React. Kinet.* **1994**, *19*, 45-91.

[55] Z. Huang, E. Terpetschnig, W. You, R. P. Haugland; *Anal. Biochem.* **1992**, *207*, 32-39 *(10.1016/0003-2697(92)90495-S)*.

[56] K. D. Larison, R. BreMiller, K. S. Wells, I. Clements, R. P. Haugland; *J. Histo. Cyto.* **1995**, *43*, 77-83 *(10.1177/43.1.7822768)*.

[57] O. Thorn-Seshold; in *Travaux Pratiques de Chimie Organique, L3*, (Eds.: N. Cheron, E. Payet, R. Ramozzi and O. Thorn-Seshold), ENS de Lyon, Lyon, **2010**, pp. 18-19.

[58] K. Stock, T. Bizjak, S. Lochbrunner; *Chem. Phys. Lett.* **2002**, *354*, 409-416 *(10.1016/S0009-2614(02)00152-5)*.

[59] Invitrogen; *Fluorescence SpectraViewer* (www.invitrogen.com).

[60] G. O. W. Lins, L. F. Campo, F. S. Rodembusch, V. Stefani; *Dyes Pigm.* **2010**, *84*, 114-120 *(10.1016/j.dyepig.2009.07.010)*.

[61] S. J. Lord, N. R. Conley, H.-l. D. Lee, R. Samuel, N. Liu, R. J. Twieg, W. E. Moerner; *J. Am. Chem. Soc.* **2008**, *130*, 9204-9205 *(10.1021/ja802883k)*.

[62] D. Villemin, L. Liao; *Synth. Commun.* **2001**, *31*, 1771-1780 *(10.1081/scc-100104407)*.

[63] M. Waibel, X.-b. Zhang, J. Hasserodt; *Synthesis* **2009**, 318-324 *(10.1055/s-0028-1083296)*.

[64] J.-J. Zamet, M. J. Haddadin, C. H. Issidorides; *J. Chem. Soc., Perkin Trans. 1* **1974**, 1687-1691 *(10.1039/P19740001687)*.

[65] S. E. Rokita, J. Yang, P. Pande, W. A. Greenberg; *J. Org. Chem.* **1997**, *62*, 3010-3012 *(10.1021/jo9700336)*.

[66] M. Freccero; *Mini-Rev. Org. Chem.* **2004**, *1*, 403-415 *(10.2174/1570193043403091)*.

[67] E. E. Weinert, R. Dondi, S. Colloredo-Melz, K. N. Frankenfield, C. H. Mitchell, M. Freccero, S. E. Rokita; *J. Am. Chem. Soc.* **2006**, *128*, 11940-11947 *(10.1021/ja062948k)*.

[68] M. Grinda, J. Clarhaut, I. Tranoy-Opalinski, B. Renoux, A. Monvoisin, L. Cronier, S. Papot; *ChemMedChem* **2012**, *6*, 2137-2141 *(10.1002/cmdc.201100355)*.

[69] F. Kratz, I. A. Muller, C. Ryppa, A. Warnecke; *ChemMedChem* **2008**, *3*, 20-53 *(10.1002/cmdc.200700159)*.

[70] W. S. Saari, J. E. Schwering, P. A. Lyle, S. J. Smith, E. L. Engelhardt; *J. Med. Chem.* **1990**, *33*, 97-101 *(10.1021/jm00163a016)*.

[71] H.-j. Jin, J. Lu, X. Wu; *Bioorg. Med. Chem. Lett.* **2012**, *20*, 3465-3469 *(10.1016/j.bmc.2012.04.012)*.

[72] Y. Meyer, J.-A. Richard, M. Massonneau, P.-Y. Renard, A. Romieu; *Org. Lett.* **2008**, *10*, 1517-1520 *(10.1021/ol800198f)*.

[73] X.-b. Zhang, M. Waibel, J. Hasserodt; *Chem.–Eur. J* **2010**, *16*, 792-795 *(10.1002/chem.200902412)*.

[74] I. Christenson; *Acta Chem. Scand.* **1964**, *18*, 904-922 *(10.3891/acta.chem.scand.18-0904)*.

[75] A. F. Hegarty, L. N. Frost, J. H. Coy; *J. Org. Chem.* **1974**, *39*, 1089-1093 *(10.1021/jo00922a016)*.

[76] M. A. Dewit, A. Beaton, E. R. Gillies; *J. Polym. Sci., Part A: Polym. Chem.* **2010**, *48*, 3977-3985 *(10.1002/pola.24180)*.

[77] S. Milstien, L. A. Cohen; *J. Am. Chem. Soc.* **1972**, *94*, 9158-9165 *(10.1021/ja00781a029)*.

[78] R. T. Borchardt, L. A. Cohen; *J. Am. Chem. Soc.* **1972**, *94*, 9166-9174 *(10.1021/ja00781a030)*.

[79] A. V. Joshua, J. R. Scott, S. M. Sondhi, R. G. Ball, J. W. Lown; *J. Org. Chem.* **1987**, *52*, 2447-2451 *(10.1021/jo00388a020)*.

[80] A. Hetenyi, T. A. Martinek, L. Lazar, Z. Zalan, F. Fulop; *J. Org. Chem.* **2003**, *68*, 5705-5712 *(10.1021/jo034417+)*.

[81] S. Mizutani, K. Shibata, F. Kikkawa, A. Hattori, M. Tsujimoto, M. Ishii, H. Kobayashi; *Expert. Opin. Ther. Targets* **2007**, *11*, 453-461 *(10.1517/14728222.11.4.453)*.

[82] Y. Urano, M. Sakabe, N. Kosaka, M. Ogawa, M. Mitsunaga, D. Asanuma, M. Kamiya, M. R. Young, T. Nagano, P. L. Choyke, H. Kobayashi; *Sci. Transl. Med.* **2011**, *3*, 110-119 *(10.1126/scitranslmed.3002823)*.

[83] P. Hechtman, C. R. Scriver, R. B. Middleton; *J. Bacteriol.* **1970**, *104*, 851-856 *(PMID: 5489438)*.

[84] C. A. Lipinski; *Drug Discovery Today: Technologies* **2004**, *1*, 337-341 *(10.1016/j.ddtec.2004.11.007)*.

[85] C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney; *Adv. Drug Delivery Rev.* **2001**, *46*, 3-26 *(10.1016/S0169-409X(00)00129-0)*.

[86] O. Thorn-Seshold, M. Vargas-Sanchez, S. McKeon, J. Hasserodt; *Chem. Commun.* **2012**, *48*, 6253-6255 *(10.1039/C2CC32227G)*.

[87] R. L. Dawe; in *Detection Threshold Modelling Explained*, Department of Defence Aeronautical and Maritime Research Laboratory, **1997**.

[88] Invitrogen; *Product Information Sheet, ELF-97 Endogenous Phosphatase Detection Kit (E6601)* **2004**.

[89] A. L. James, J. D. Perry, S. P. Stanforth; *J. Heterocycl. Chem.* **2006**, *43*, 515-517 *(10.1002/jhet.5570430241)*.

[90] J. Haringsma, G. N. J. Tytgat, H. Yano, H. Iishi, M. Tatsuta, T. Ogihara, H. Watanabe, N. Sato, N. Marcon, B. C. Wilson, R. W. Cline; *Gastrointestinal Endoscopy* **2001**, *53*, 642-650 *(10.1067/mge.2001.114419)*.

[91] G. Zlokarnik, P. A. Negulescu, T. E. Knapp, L. Mere, N. Burres, L. Feng, M. Whitney, K. Roemer, R. Y. Tsien; *Science* **1998**, *279*, 84-88 *(10.1126/science.279.5347.84)*.

[92] G. McNamara, A. Gupta, J. Reynaert, T. D. Coates, C. Boswell; *Cytometry. Part A : the journal of the International Society for Analytical Cytology* **2006**, *69*, 863-871 *(10.1002/cyto.a.20304)*.

[93] K. Faghihi, K. Zamani, A. Mirsamie, M. Reza Sangi; *Eur. Pol. J* **2003**, *39*, 247-254 *(10.1016/S0014-3057(02)00200-8)*.

[94] L. J. Exner, L. S. Luskin, P. L. de Benneville; *J. Am. Chem. Soc.* **1953**, *75*, 4841-4842 *(10.1021/ja01115a513)*.

[95] W. K. M. Chong, R. K. Duvadie, L. Li, Y. Yang; *Antiproliferative 2-(heteroaryl)-aminothiazole compounds and pharmaceutical compositions, and methods for their use (US20050038078)*, **2005**.

[96] R. K. Mylavarapu, K. Gcm, N. Kolla, R. Veeramalla, P. Koilkonda, A. Bhattacharya, R. Bandichhor; *Org. Process Res. Dev.* **2007**, *11*, 1065-1068 *(10.1021/op700098w)*.

[97] H. Alarabi, R. A. Bell, H. E. Howard-Lock, J. Kowanetz, C. J. L. Lock; *Can. J. Chem.* **1996**, *74*, 574-582 *(10.1139/v96-061)*.

[98] C. A. G. N. Montalbetti, V. Falque; *Tetrahedron* **2005**, *61*, 10827-10852 *(10.1016/j.tet.2005.08.031)*.

[99] G.-J. Ho, K. M. Emerson, D. J. Mathre, R. F. Shuman, E. J. J. Grabowski; *J. Org. Chem.* **1995**, *60*, 3569-3570 *(10.1021/jo00116a057)*.

[100] J. S. Davies, R. J. Thomas, M. K. Williams; *J. Chem. Soc., Chem. Commun.* **1975**, 76-77 *(10.1039/C39750000076)*.

[101] J. Adrio, C. Cuevas, I. Manzanares, M. M. Joullié; *J. Org. Chem.* **2007**, *72*, 5129-5138 *(10.1021/jo070412r)*.

[102] M. Bakavoli, A. Shiri, Z. Ebrahimpour, M. Rahimizadeh; *Chin. Chem. Lett.* **2008**, *19*, 1403-1406 *(10.1016/j.cclet.2008.07.016)*.

[103] M. Waibel; *Design and Synthesis of Molecules to Probe Peptidase Activity* (PhD Thesis), Ecole Normale Supérieure de Lyon, **2009**.

[104] H. Eckert, B. Forster; *Angew. Chem., Int. Ed. Engl.* **1987**, *26*, 894-895 *(10.1002/anie.198708941)*.

[105] L. Cotarca, H. Eckert; *Phosgenations - A Handbook,* Wiley-VCH, **2005**.

[106] J. Schott, O. S. Pokrovsky, E. H. Oelkers; *Reviews in Mineralogy and Geochemistry* **2009**, *70*, 207-258 *(10.2138/rmg.2009.70.6)*.

[107] W. P. Inskeep, P. R. Bloom; *Geochimica et Cosmochimica Acta* **1985**, *49*, 2165-2180 *(10.1016/0016-7037(85)90074-2)*.

[108] E. K. Y. Chen, R. A. McBride, E. R. Gillies; *Macromolecules* **2012**, *45*, 7364-7374 *(10.1021/ma301667c)*.

[109] D. Shabat, R. J. Amir, A. Gopin, N. Pessah, M. Shamis; *Chem.–Eur. J* **2004**, *10*, 2626-2634 *(10.1002/chem.200305715)*.

[110] T. Jenkins, A. Kolesnikov; *Controlled Release of Phenolic Opioids (WO2007140272)*, **2007**.

[111] V. M. Vrudhula, J. F. MacMaster, Z. Li, D. E. Kerr, P. D. Senter; *Bioorg. Med. Chem. Lett.* **2002**, *12*, 3591-3594 *(10.1016/S0960-894X(02)00784-9)*.

[112] V. Stella, S. Dhareshwar; in *Prodrugs: Challenges and Rewards*, Springer, New York, **2007**, pp. 731-799.

[113] M. Wakselman; *N. J. Chimie* **1983**, *7*, 439-447.

[114] P. L. Carl, P. K. Chakravarty, J. A. Katzenellenbogen; *J. Med. Chem.* **1981**, *24*, 479-480 *(10.1021/jm00137a001)*.

[115] S. Halazy, V. Berges, A. Ehrhard, C. Danzin; *Bioorg. Chem.* **1990**, *18*, 330-344 *(10.1016/0045-2068(90)90007-R)*.

[116] C. J. F. Van Noorden; *J. Histo. Cyto.* **2010**, *58*, 481-497 *(10.1369/jhc.2010.955518)*.

[117] R. Weissleder, B. D. Ross, A. Rehemtulla, S. S. Gambhir; *Molecular Imaging,* People's Medical Publishing House, Shelton, Connecticut, **2010**.

[118] G. R. Pettit, S. B. Singh, E. Hamel, C. M. Lin, D. S. Alberts, D. Garcia-Kendal; *Experientia* **1989**, *45*, 209-211 *(10.1007/bf01954881)*.

[119] G. R. Pettit, C. Temple, V. L. Narayanan, R. Varma, M. J. Simpson, M. R. Boyd, G. A. Rener, N. Bansal; *Anti-Cancer Drug Des.* **1995**, *10*, 299-309.

[120] K. Ohsumi, R. Nakagawa, Y. Fukuda, T. Hatanaka, Y. Morinaga, Y. Nihei, K. Ohishi, Y. Suga, Y. Akiyama, T. Tsuji; *J. Med. Chem.* **1998**, *41*, 3022-3032 *(10.1021/jm980101w)*.

cette page a été laissée vierge intentionellement

Part II

Towards spacers for magnetogenic MRI probes

Table of Contents: Part II

II. TOWARDS MRI DETECTION OF GLYCOSIDASES

1. State of the Art: Molecular Probes for MRI

1.1. Magnetic Resonance Imaging

1.1.1. General parameters

Magnetic Resonance Imaging (MRI) is a clinical soft-tissue imaging modality that has become firmly established over the last 30-40 years. MRI uses 3D gradients of magnetic fields to acquire water proton resonance intensity data, which are reconstructed into spatially-resolved images. A variety of pulse sequences are well-established for imaging both simple and complex *anatomical* features by weighting the images according to specific features. The simplest is spin-density weighting, which distinguishes between tissues based on water concentration; however relaxation-time weighting (both T_1 and T_2) is most commonly used to distinguish between soft tissues, since they usually have more significant differences between these relaxation times than between their water concentrations. More advanced clinical applications extend to flow angiography (imaging to show mobile vs static blood pools) and diffusion imaging (fluid leakage), as well as 'functional' MRI (fMRI) which shows blood oxygenation levels and hence is used to reveal brain activity.[1]

MRI has two extremely attractive qualities as an imaging modality. It has essentially *no depth limitations*, and it has *no intrinsic risks of harming the subject*, because of the very low absorption of the low-energy radiofrequency electromagnetic waves it relies upon. This distinguishes it sharply from PET and X-ray imaging (radiation-dependent dose limitations, similar or lower spatial resolution, lower contrast resolution), and from optical/fluorescence imaging (wavelength-dependent excitation power and depth limitations, but generally depth-limited). Because of the absence of a depth limitation, whole-human MRI scans have been practised for decades, which makes it an especially attractive modality for routine screening or diagnostics (no limit on the number of scans). It can also have extremely high resolution, currently down to ~10×10×10 μm voxels.[1-3]

Its chief drawback compared to other imaging modalities is its intrinsically low sensitivity.[4] The *signal* that can be measured is always only a tiny proportion of total potential water proton signal, because the energy gap between the two proton nuclear magnetic states is so small that they are in equilibrium with only a $\sim 10^{-5}$ differential of population. Furthermore, the native differences between relaxation times in different tissues may be small, so *contrast agents* are routinely applied to reveal physical and biochemical features not otherwise differentiable with native tissue imaging (in about a third of all MRI procedures).[5] This thesis chapter will be concerned with the design of new auto-immolative spacers for use in bio-activated contrast agents, which will have the potential to adapt MRI from an anatomical to a *molecular* imaging modality.

1.1.2. Contrast agents for T_1-weighted MRI

A brief introduction to the relevant effects concerning MRI contrast agents is given. A slightly more extended discussion is included in section III.4, but excellent and comprehensive reviews are available[1,6,7] to which the reader is directed for further information.

The T_1, or spin-lattice time constant, describes the timescale for the decay of the external-field-aligned (longitudinal) component of the net proton nuclear spin magnetic moment from its excited state to its equilibrium value. The T_2, "transversal" or spin-spin time constant describes the timescale for the loss of coherence of the orthogonal component due to dephasing between spin packets. Contrast agents may affect either the T_1 or T_2 (sometimes, both to a similar degree) of the water molecules that surround them, therefore pulse sequences that weight images based on these relaxation times are used to image the distribution of these contrast agents. Depending on how much the contrast agent alters the tissue relaxation times from their native state, this imaging may be more or less sensitive.[1]

The clinical contrast agents in current use are T_1-agents, paramagnetic metal coordination complexes which lower the T_1 of surrounding water, so they give a net increase in observed signal (faster relaxation during the acquisition period). The spin magnetic moment μ of the paramagnetic complex can interact with the nuclear spin moments of nearby excited water protons, reducing their T_1 and so increasing their relaxation speed by up to many orders of magnitude. Among the results of the generalised Solomon-Bloembergen-Morgan theory, which is widely used to describe these interactions, are two effects that will be discussed here. Firstly, the greater is μ, the greater is the relaxation induced in surrounding water molecules and therefore the greater is the signal enhancement; if the complex is not paramagnetic at all, then no relaxation will be observed. Secondly, the coordination of water molecules is important for *good* relaxivity r_1 ($=1/[c\times T_1']$, for concentration c and observed relaxation time T_1'): but relaxation can also be transmitted by second-sphere effects in coordination-saturated paramagnetic complexes. The average water coordination number is noted q.

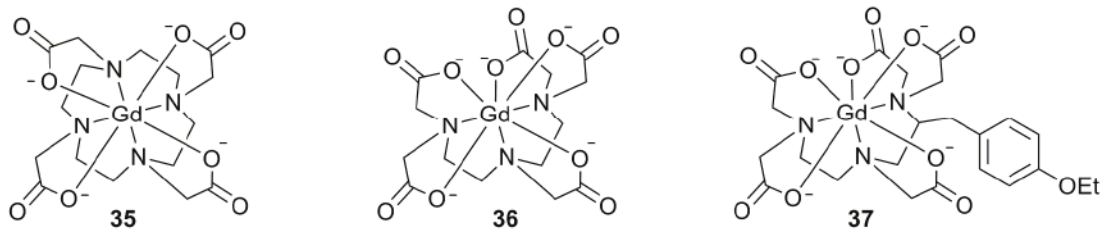


Figure II.1 – Classical, Gd^{3+}-based paramagnetic contrast agents Gd-DOTA (35), Gd-DTPA (36) and Gd-EOB-DTPA (37). The water molecule in the ninth coordination site is omitted for clarity.

Current clinical paramagnetic complexes (Figure II.1) are usually based on Gd^{3+}, as are most T_1-contrast agents in research, though other ions such as Fe(III)[8], Mn(II)[5] and Cr(III)[9] are also seen. These central metal ions are chosen primarily to maximise the μ of their complex and so allow for high-signal contrast agents: Gd^{3+} is the most outstanding ion for high contrast as it has seven unpaired *f* electrons; the others have fewer (eg. Mn(II) has 5 unpaired *d* electrons, etc).

In particular for Gd, its complexes are both macrocyclic (Gd-DOTA, **35** – a reference standard, with relaxivity r_1 = 2.44 $mM^{-1}s^{-1}$ at the relatively high field strength of 7T)[2] and non-macrocyclic (Gd-DTPA, **36**); the former are much preferred for toxicity reasons.[10,11] All are octa-coordinate, with the ninth coordination site occupied by water: while complexes with q=2 can show good relaxivity, they are not as kinetically inert and may more quickly release Gd^{3+}, which is a toxic heavy metal, so they have not been allowed for medical use.[1] Complexes with some selectivity for the biochemical environment of certain tissues have been developed (eg. "Primovist", Gd-EOB-DTPA, **37**). In summary, these contrast agents allow MRI with generally enhanced anatomical tissue contrast.[1]

1.1.3. Prior approaches to MRI for molecular imaging

Many efforts have been made to design molecular contrast agents whose relaxivity can be enzyme-modulated, to serve as molecular probes. One principal axis of research has examined RIME agents, where an enzymatic reaction provokes a change of binding affinity of the complex towards a biological protein (eg. albumin).[12] Binding increases the rotational correlation time of the complex (slower tumbling) and increases of relaxivity are seen, especially at low field strengths.[13] This is however an inherently a relative change approach to MRI probes;[14] its effects are also essentially limited to low magnetic field strengths and tail off at higher frequencies. By contrast, high field MRI is absolutely necessary for good imaging resolution and acquisitions on a practical timescale, and clinical machines are increasingly moving towards these higher frequencies.[1,15-17]

The second axis of molecularly-enzyme-responsive contrast agent research is the more relevant to the current work, and has principally focussed on Gd(III) complexes where q is modified upon enzymatic activity, such as the classic example **EgadMe** (Introduction Figure 8) which really launched the field in 1997[18]. **EgadMe**'s design asks that the targeted hydrolase should recognise and cleave a substrate which is essentially coordinated to the central Gd^{3+} ion, so that water may replace it and be relaxed by the complex. Whether this highly non-native substrate design really warrants the epithet "smart" is debatable; indeed, its authors found that an *in vitro* incubation time of 7 days was needed, despite a probe:enzyme *concentration ratio* of only 1000:1, in order to reach signal plateau. This was highly unpromising for *in vivo* applications; however this direct blocking strategy has several more serious molecular concerns, as other researchers trying to exploit this design have found.[19] Meade *et al*. later acknowledged this: "the *in vitro* behaviour of these agents does not reflect their *in vivo* activity:"[20] a polite way of rephrasing 'the effect forced *in vitro* cannot be reproduced *in vivo*.' The logical development was to use auto-immolative spacers to separate the complex from the substrate.[20,21] A range of functionally similar designs have since been reported, however the molecular mechanism which is altered upon enzymatic activity is still the same - the water coordination number q.

Regardless of the actual molecular design proposed, these prior agents are united by the fact that they all *inherently operate in a small relative change mode*. They are based on classical paramagnetic complexes, and like them *will remain essentially limited to anatomical imaging*, because they are

always paramagnetic and hence are always *substantially MRI-active* even if only by second-sphere effects.[1-3] Consider that in Part I, it has been argued that if any given probe functions in a relative change mode (eg. optimistically, ×20 enhancement even *in vivo*), the *in vivo* detection sensitivity is minimal. Discussion and modelling of the requirements for *in cellulo→in vivo* detection of enzyme activity have also shown the severe difficulties for proposed fluorogenic probes which face relatively large background signal (eg. which use low Stokes shift fluorophores). Yet the situation for enzyme-responsive MRI is so much more challenging that such considerations were published only one year after the first proposed responsive design, to suggest caution in following this approach:

> *"Although MRI is a powerful tool for noninvasively mapping biological structure and function… the method has* ***ubiquitous*** *limitations in the available signal-to-noise ratio. This limitation is especially true as the method is pushed to increasingly higher resolutions. A consequence is that the* ***ability to discriminate between tissues on the basis of differences in T_1 can be limited****… when exogenous contrast agents are utilized to enhance T_1 differences, relatively* ***large agent concentration… must be delivered****."*[4]

MRI has a far *higher* inherent background than fluorescence: MRI commonly requires ~mM of active contrast agent to pass above noise in *anatomical* imaging, compared to ~μM for fluorescence. Yet prior designs for enzyme-responsive MRI probes (all based on these classical paramagnetic agents) can offer only *much smaller* signal change (than even the more useless of fluorogenic probes): eg. maximum +50% on *total* sample activation for these probes. By the arguments of section I.3.1, this should show that relative change MRI probes will never be able to achieve sensitive detection of realistic amounts of enzyme activity *in vivo*, no matter what (small) effects are claimed *in vitro*.

Ahrens *et al.*[4] modelled the minimal amount of a standard enzyme-responsive contrast agent which must be activated by a target enzyme in one location, in order to pass a threshold just *differentiating* it from a second nonenzymatic location, within an optimised high-field small-embryo imaging setup (*much* more sensitive than whole-human imaging). They calculate a *minimal* threshold activation difference of 1400 μM. This is far beyond what one can hope for *in vivo* for a variety of reasons, principally because this would imply enormous enzyme activities: yet there are other problems which this calculation hides: proagent/agent signal diffusion/bioconcentration are entirely ignored (compare Figure I.3 panel b vs panel d!), agent hydrolytic stability is neglected, the form of enzyme kinetics (reaction slows near K_M) is ignored, etc. Therefore the *minimum amount of classical-design enzyme-responsive contrast agents which must be activated in order to differentiate biochemical features* (let alone quantify them) *is too high for any realistic use* in vivo.[22]

Therefore, the Bio-Organic Chemistry group pursues a different approach entirely. Molecular designs for Magnetogenesis - **truly off-on magnetogenic behaviour** - could solve the sensitivity problem, therefore allowing for detection of much lower quantities of activated probe (*threshold*), *and this* in a fashion that correlates to the underlying enzymatic activity (*reliability*; cf. section I.3).[23]

1.2. Magnetogenesis

1.2.1. Requirements, and selected advantages/issues

The Bio-Organic chemistry group has pursued truly magnetogenic designs since 2005[24]. The goal is to develop MRI-silent (diamagnetic) pro-agents, which switch to MRI-active (paramagnetic) contrast agents after triggering from a specific targeted enzyme, for high-sensitivity off-ON-mode imaging.[2,25]

The requirement that the proagent signal must be *totally and mechanistically* quenched implies that no permanently-paramagnetic metal ion can be used. The simplest way to design this is if the proagent complex is *diamagnetic*, then enzymatic activity will trigger a spin switch to the *paramagnetic* state. Recalling the argument of section I.3 that intrinsic background, and not signal brightness, is usually the limiting factor for *in vivo* sensitivity, we abandon the previous central metals, which were recommended only by their high μ (high potential relaxivity), and instead focus on metals whose magnetism can be switched off→ON by appropriate covalent modifications in their ligand.

Fe(II) is one outstanding candidate. Its truly diamagnetic, low-spin state (off) may be attained by a variety of complexes, eg. using an N_6 coordinating motif (eg. **2c**, Introduction Figure 4). Molecularly similar complexes, eg. N_5O configurations (**2a**) may give the paramagnetic high-spin state (ON). The high-spin state features four unpaired *d* electrons, resulting in a considerable μ, and the N_5 ligand allows for the presence of a water at the sixth coordination site (q=1), for efficient relaxation of the bulk solution. Therefore the model ON-state Fe^{2+} complex **2a** can offer a pleasingly large relaxivity of 1.29 $mM^{-1}s^{-1}$, which is only half that of **35** (both measured at 7T field).[2,26]

The off-ON possibility is not the only advantage of this approach. For Gd(III)-based contrast agents, the toxicity of free gadolinium is well documented.[7] As Molecular Imaging requires deep-tissue and cellular penetration of the probe, and therefore a long cellular residence time, even the most thermodynamically stable complexes of toxic ions may be too hazardous for use in *in vivo* cellular/ molecular imaging. Currently, Gd(III) agents are only approved for extracellular use, especially in the blood pool, and must be entirely evacuated within ~24 hours; this is unlikely to change.[7] In contrast, homeostatically-managed ions such as Fe may pose fewer problems if they are similarly leached out. Therefore Fe^{2+}-based contrast agents may not face the same limitations of dosage (which may compensate for their lower absolute relaxation) or ligand design which restrict the use and engineering of Gd^{3+} agents, as long as they address standard toxicity considerations for metal complexes.[27]

Achieving such a magnetogenic probe would give off-ON style sensitivity, a clear advantage over the current relative change probe designs (offering at best n~2, see section I.3.2). However, off-ON functioning is of course not the only sensitivity consideration (Introduction section 2). To give one example revelant to Part I, reliable *in vivo* work requires avoiding signal bleed, to ensure good localisation of the enzymatic activation (Figure I.3). Precipitation of a complex is not an option, but *increased cellular retention* may be. Consider eg. an Fe(II) proagent whose spacer bears a cell-

penetrating peptide (CPP)[28] and/or a group which charge-compensates the complex, for better passive membrane permeability. After enzymatic activation and spacer immolation, the agent will be significantly less membrane-permeable. Greater retention inside its activating cell would therefore mark the site of enzymatic activity similarly to precipitating probes. Yet given that no enzyme-triggered *magnetogenic* probe had yet been proven (the group published the first successful design in April 2013[29]), the goal of this work was in design and proof of concept rather than in refinement.

1.2.2. Design and relevant prior art

As per the discussions of modularity and probe tuning in Part I, the Hasserodt group considers that a 3-component architecture is most promising for the proof of concept of a bioactivated MRI probe: N_5 complex-*N_1 coordinating, auto-immolative spacer*-substrate. In this design, an enzyme should achieve good steric access to its substrate, then by processing it trigger the spacer immolation; then, the eliminated N_5 complex should coordinate water in the vacant coordination site, and so give the desired N_5O, high-spin, water-coordinating, activated contrast agent. The group has reported three families of N_5/N_6 Fe(II) complexes as candidates for the proagent/agent forms of an off-ON design, and more are under study (Figure II.2).

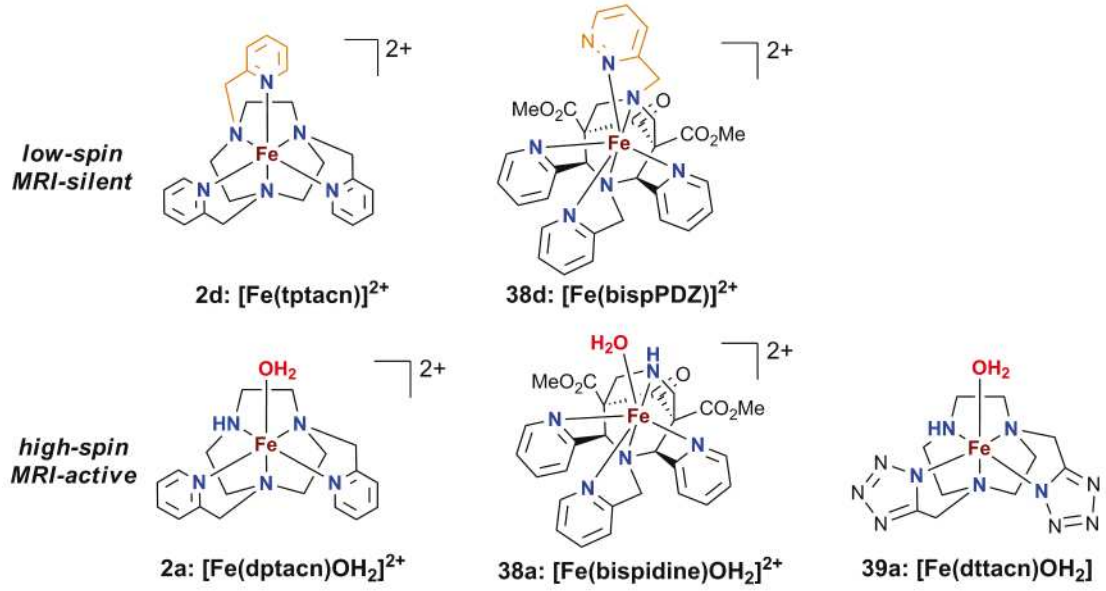


Figure II.2 – Structurally-related model lowspin/highspin pairs of Fe(II) complexes reported so far

2a and **2d**, based on the picolyl-substituted tacn ligand, were known in prior literature, and reported in the context of MRI by the group in 2008[2]. **38a** and **38d** are complexes with novel examples of the bispidine cage ligand, which were recently reported by the group[30]. Notably, **38d** and a closely-related example constitute the first truly low-spin Fe(II) complexes with the bispidine ligand system; the key to this is their low steric demand at the *ortho*-substituent to the sixth coordinating nitrogen (here, the nitrogen of a pyridazine ring): a simple pyridine ring is unacceptable due to steric clash in the more crowded complex. **39a** returns to the tacn system, but uses tetrazole coordinating rings to achieve electroneutrality in the complex. The exploration of such different complexes is motivated by

questions of solubility, charge compensation/osmolality, and especially, of ligand functionalisation for probe property tuning (eg. modifying the bispidine ligand on its carboxylate/ketone periphery by simple ester hydrolysis, or attaching solubilising groups/CPP).

To achieve a *biologically-activated* spin state transformation as outlined above (complexes type **2d** → complexes type **2a**), the sixth coordinating nitrogen must be decoordinated (group marked in orange in Figure II.2). Prior work in the group has attempted decoordination via the design of an N_1 coordinating synthon which would eliminate from a complex after enzymatic activation, leaving the N_5 complex (eg. **2a**) behind: a *coordinating, auto-immolative spacer*. Note that the restrictions on such a spacer are far more exacting than for eg. fluorogenic probes. Firstly, the macrocyclic amine cannot be derivatised (eg. as a carbamate, in the classic use of PABA spacers to eliminate an amine[31]) – forcing the low-spin state for the off-ON switch requires that the amine be free to coordinate in the proagent state. Secondly, although initial work desires only a proof-of-concept, the eventual biological applications of a system must always be kept in mind. Not only modularity must be assured, so that enzymes of true biological interest can be targeted, but especially the toxicity of the spacers must be considered, since MRI probes require such higher dosage than fluorescent probes that even spacers considered acceptable for fluorescent probes may be impossible to apply to MRI.

For an initial proof-of-concept approach however, phenolic glycosides of β-galactose on PHBA-type[21,31-34] auto-immolative spacers have been widely featured in prior enzyme-responsive MRI designs;[18,21] and the group's first approach was to exploit this same spacer with a coordinating group attached. Therefore dptacn-based ligand **40a** was designed; incubation with β-galactosidase cleaved the glycoside giving **40b** which as expected, quickly eliminated the amine to give the dptacn ligand plus a quinone methide byproduct **41**. However, once the Fe(II) complex of this ligand was examined (**2b**), although enzymatic cleavage of the glycoside was confirmed, the dptacn *complex* **2a** was not formed: ie, elimination of the quinone methide **41** was blocked by complexation[25] (Figure II.3).

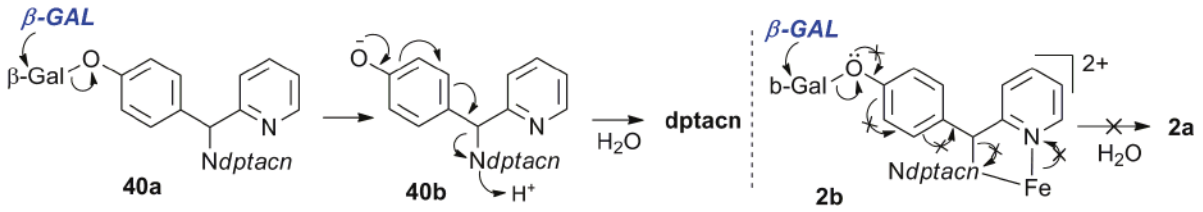


Figure II.3 – The initial PHBA-based design for a *coordinating auto-immolative* spacer was unsuccessful: although elimination occurs in the free ligand 40a, the corresponding Fe(II) complex 2b no longer fragments after enzymatic activation, so no spin state change was seen upon enzymatic activity.

My PhD work continued my M2 thesis work (see section III.4.1.2)[35] on the design and synthesis of suitable auto-immolative coordinating spacers for glycosidase substrates, which *would* be able to overcome the complexation stabilisation and therefore give biologically triggered magnetogenesis.

1.2.3. Addressing eliminative reactivity

There are several factors which would seem to contribute to this difference of reactivity between complexed and uncomplexed forms. Listed in order of increasing importance: (1) the complexation of the two potential fragments (ligand and quinone methide) to the metal may prevent their diffusion away from each other and promote back-reaction (similar to a tight-ion-pair effect), which is not a problem for free ligand **40b**. (2) As will be seen later in this report, the 2-pyridyl group is a rather powerful electron-attracting/delocalising group which can help to stabilise otherwise unfavourable tautomeric forms particularly in push-pull configuration (see section II.4.2.3). As coplanarity may be achieved in the transition state **40b**‡, the 2-pyridyl probably assists the elimination reaction by pulling π-electron density from the electron-rich phenolate into a well-conjugated π-system, which is even lower-energy than might be expected for *eg.* a phenyl ring. However, in **2b**‡, the 2-pyridyl is powerfully coordinated to the iron, locked in place by the 5-membered chelating ring with little flexibility in its movement[36]. Therefore ring coplanarity cannot be achieved and this advantage is lost (red arrow). (3) The eliminating dptacn nitrogen of **40b** may be protonated, which would assist its elimination (amine not amide): but in the complex **2b**‡, there is no free lone pair to accept a proton. Although the Lewis acidic Fe(II) must be considered, one could argue that its charge is shared over the ligand so it is probably a less effective activator of elimination. (4) Most seriously, achieving a favourable transition state of elimination requires a planar carbon with its nascent π-orbital aligned along the σ-bond to the leaving-group. The chelation in **2b**‡ fixes the angles in the chelate ring so that the nascent π-orbital cannot simultaneously align with the leaving group and be normal to the pyridyl axis, unless the pyridine-Fe coordinating bond is first broken (blue arrow; Figure II.4).[25]

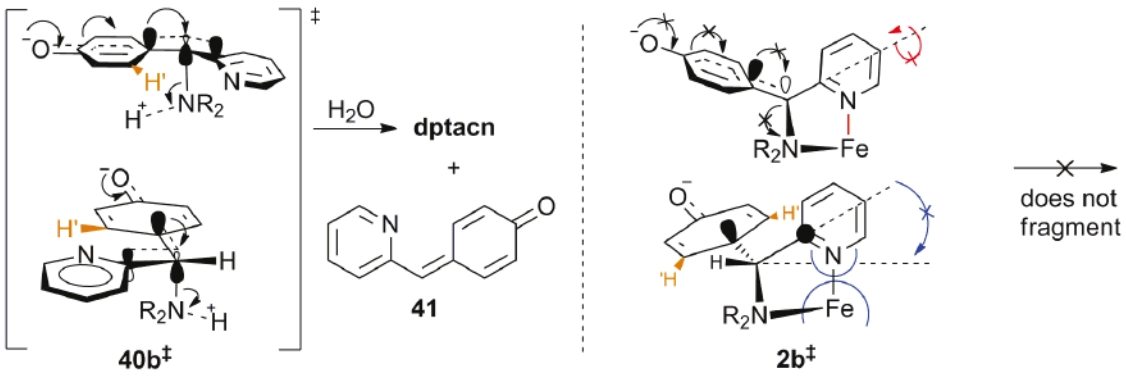


Figure II.4 – The elimination observed for 40b‡ may be favoured by protonation, conjugation, alignment and diffusion of the fragmentation products. But in the corresponding iron complex 2b‡, no fragmentation was observed, which may be due to the absence of these effects.

Although this prior design was ultimately unsuccessful, it was promising that the distal phenolic glycoside on the complex was well enzymatically processed; therefore distal aromatic glycosides were retained as targets for enzyme substrates. Also retained in the next designs for coordinating auto-immolative spacers was the imine-type coordinating nitrogen, needed to ensure a strong ligand field in the proagent (→low-spin, off). However, new molecular approaches were needed to overcome the

complexation stabilisation. Keeping within the design limits, the tethering effect cannot be avoided (eg. by replacing the critical pyridyl with a more coordination-labile group) since then the proagent may not be fully low-spin, and the off-ON advantage is lost. Therefore the reaction which removes the N_1 coordinating unit had to be improved. Proagent stability would have to be married with favoured post-triggering activation: so Prof. Hasserodt sought metastable species with thermodynamically favoured fragments accessible *via* low energy transition states, and adopted two molecular strategies:

(A) The *furan strategy* sought a more energetically-favourable elimination, retaining the same fragmentation strategy. (B) The *cyclic carbamate* strategy took an entirely different N_1-removing approach, with intrinsically less challenging thermodynamics. A previous strategy featuring elements of both approaches but without the principal advantages of either (pyridinone methides) had already proved unsuccessful during my M2 work (summarised in section III.4.1.2).[35]

This PhD chapter will now relate progress in these new auto-immolative spacers: furans (Part II.A) and cyclic carbamates (Part II.B).

Part II.A: Furanol-based Eliminative Spacer System

2. A Design for Elimination

The successful fragmentation of the phenylogous hemiaminal **40b**‡ first guided research towards other eliminative spacer constructions with a similar enzyme recognition advantage (substrate distal to the complex; vs. non-processed proximal designs eg. **EGadMe**), but with a more powerful elimination. The hope was that such a more powerful eliminative spacer could energetically overcome the chelation thermodynamic stabilisation and the lack of pyridyl ring coplanarity / amine protonation / ideal trigonal bipyramidal transition state geometry once complexed which were probably responsible for the non-fragmentation of **2b**‡. It was considered that the distance between substrate and complex given by its aromatic ring was good for proof-of-concept.

Initially we preferred to avoid stacked designs: for example, using a terminal 1,2-elimination spacer as the 'power source' spacer would require a second, stacked, aromatic spacer to attach a distal substrate in order to assure good glycosidase processing: this could become synthetically challenging (**42a**). We could also have attempted to modify the eliminative force by substitution motifs E affecting electron density, or by relieving steric strain of groups *G* upon elimination (**42b**, cf. the adamantyl quinone methide spacer seen in Introduction Figure 11). Yet we preferred to seek a new system entirely, and the initial search was for the simplest solution: a spacer that provided an aromatic hydroxyl attachment for the glycoside, and which would also give more powerful fragmentation (**43**; Figure II.5).

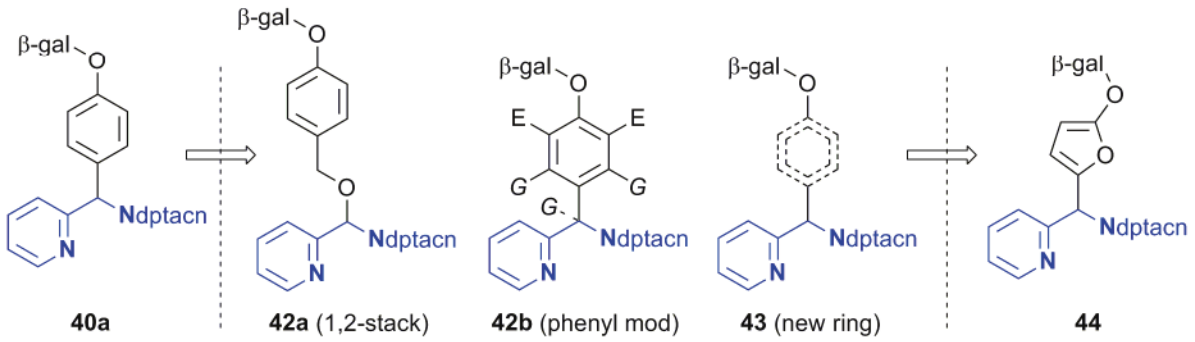


Figure II.5 – routes to attack the design problem for a more powerful eliminative spacer

2.1.1. Functional precedents for furan elimination

In my M2 work I had shown that the replacement of the spacer's phenyl ring by a pyridine (see section III.4.1.2) gave a non-fragmenting design, probably since it was too stabilised in the 'phenolate' form to show eliminative donation, even without a complexed metal.[35]. A more powerful spacer would, it was considered, have a 'phenolate' which was disfavoured relative to the donated form to such an extent that the hydroxyl was not the major species. This led Prof. Hasserodt to the γ-butenolide skeleton (sole natural tautomer of '2-furanol'), and so to design **44** where the butenolide is trapped in its furanol form. As well as its highly energetically favourable reversion to the butenolide, an alkoxyfuran spacer also has steric advantages over the quinone methide system in **40a**; the furan

system's smaller, 5-membered ring and especially the low steric presence of the ring oxygen, should increase the flexibility and steric freedom around the eliminating methine carbon: without actually breaking the iron-pyridine coordinative bond, this is the only way to favour a more trigonal bipyramidal eliminative transition state, so it seemed a doubly promising design.

2-furanols have never been used as auto-immolative eliminative spacers in probes. However, *γ-butenolides* have been the target of extensive research, particularly in pharmacological chemistry, and the elimination of 2-furanols to form butenolides is well attested[37] - both in synthetically useful reactions, as well as in disappointing reaction failures where the fragmentation was not desired; the alkoxyfuran's ketal acetal/enol ether nature makes them particularly prone to such degradations.[38-40] The literature extensively attests to the very energetic rewards of reverting to the highly favourable butenolide tautomer, and the tautomerisation-elimination reaction is even more favourable when it extends electron delocalisation in the product.[38,40,41] The reactions of Kraus and Sugimoto[41] and of D'Auria *et al.*[42] are most relevant immolative examples which illustrate this behaviour (Figure II.6).

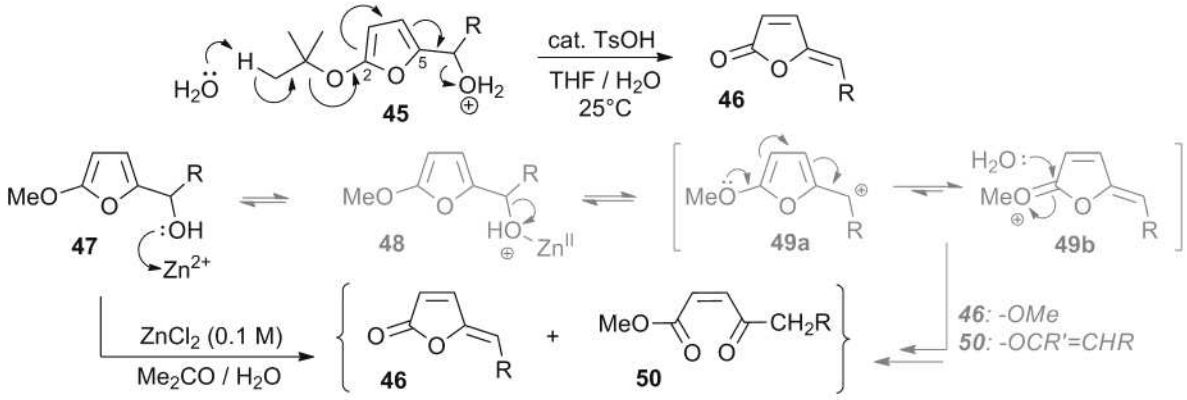


Figure II.6 – Literature examples showing 'furanol' eliminations giving γ-butenolides 46.

Kraus and Sugimoto show the fragmentation of **45** in "gentle" aqueous conditions to give γ-butenolides **46** (though the reaction is unlikely to be concerted, as was drawn). The kinetics are efficient, even for eliminating (essentially) hydroxide ion, and the reaction is more efficient with R=Ph (61%) than for non-aromatic groups, conjugated or not (R=*n*Pr: 45%; R=CH=CH-CH=CH-CH_3: 44%). Pleasingly, the relatively low-acidity conditions (cat. *p*TsOH) imply that furanols can relatively easily expel bad leaving-groups such as alkoxides.[41] D'Auria *et al.* show a similar fragmentation of a furanylogous hemiacetal, though with Zn^{2+} replacing cat. *p*TsOH, for pH ~6-7 fragmentation of 2-methoxy-5-carbinolfurans (**47**) to γ-butenolides (**46**), also isolating ring opening minor products **50**. Reactions of 15 min at 20°C give conversions of 85-90% (R = *n*-octyl, *n*-decyl and *n*-dodecyl).[42]

Both procedures illustrate the strong eliminative tendency of furanols in reasonable models for physiological media: certainly, this is far stronger than phenols, since the butenolides which are the furan equivalent of quinone methide byproducts (**46**), are isolable, rather stable pharmacophores. Hopefully, the pyridyl-coordinating auto-immolative spacer would also experience an even greater boost to reactivity/product stability than Kraus' phenyl example (delocalisation). As Kraus' example

essentially expels an alkoxide, hopefully the only partially charge-compensated amide (complex **2a**) may also be expelled favourably by furanol elimination. Furanols were therefore considered a promising basis for a new, more powerful eliminative spacer, to be applied as in eg. complex **44b**, giving potential off-ON enzyme-activated MRI contrast agents.

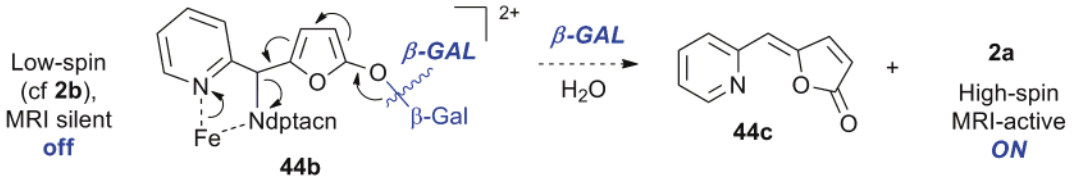


Figure II.7 – Design for off-ON enzyme-activatable MRI contrast agent 44b based on a powerful furan spacer

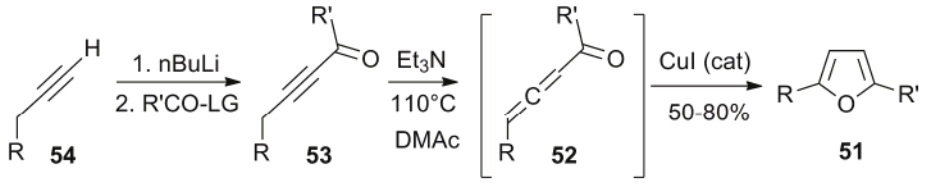
2.1.2. Molecular precedents

While the chemistry of γ-butenolides has been studied extensively, the chemistry of 2-alkoxy-5-functionalised furans is much less known: no suitable chemical trigger/enzyme-sensitive compounds are commercially available, easily accessible, or even reported. Synthetically, the converse side of the favoured furanol elimination/tautomerisation is that standard alcohol protections on a 'furanol', eg. by base-catalysed S_N2 reactions on the more easily obtained butenolide, are usually impossible:[37] attempts to use the "capricious" butenolide usually give C5-alkylation due to the favoured 2-furanone-5-carbanion tautomer.[43] The main exceptions are silyl groups, especially TMS, which can alkylate on the 2-O. Characteristically however, these products are extremely labile and serve as excellent methods for promoting more difficult C5-alkylations.[43-46] Notably for the present work, even the compound 5-hydroxy-2-furanone, which is *already trapped* in the "2-hydroxyfuran" form, resists O-glycosylation (tetraacetochloroglucose/Zn^{2+}, tetraacetoglucose/DCC both fail, Schmidt glycosylation low-yielding).[47] No precedents for the glycosylation on *untrapped* 2-furanone were available.

The O-alkyl groups which are seen in 5-alkyl-'2-furanol ethers' (~2000 hits in CAS) are therefore installed by other methods either on an existing furan ring (eg S_NAr) or before the furan is fully formed; these are almost exclusively non-functional, alkyl-only blocking groups. Phenyl is the most common (1300 hits),[48] then *tert*-butyl;[41] note that both of these groups *have no α-hydrogens and resist S_N2*, which will be seen to be important. Examples with methyl are also well-represented (250 hits) though they are much more sensitive than phenyl;[42,49-52] the few other reported 2-alkyloxy groups are usually long-chain *n*-alkoxy moieties.[53] Ethyl, isopropyl and benzyl groups, for example, are totally unreported (a sure sign that furanol ethers are avoided wherever possible); and certainly, no hydrolase substrate derivatives have been attempted (eg. aminofurans are notoriously unstable intermediates). With this general lack of molecular precedents, one literature report was distinguished from all others, as it provided hope for a general, permissive method to install sensitive substrates such as β-galactose onto the furan ring: a cycloisomerisation reaction of Kel'in and Gevorgyan[54].

2.1.3. Initial synthetic attack and selected points of caution

Kel'in and Gevorgyan[54] provided the only literature example of a 2-furanol-acetal, the THPO-furan **51b**. This is the only report of a furan which structurally mimics a 2-furyl glycoside, so their method seemed a prime method for initial exploration.



	R	R'	
a	OMe	nC_6H_{13}	78%
b	OTHP	*t*Bu	75%
c	nC_5H_{11}	Me	94%
d	H	nC_6H_{13}	63%

Figure II.8 – The route of Kel'in and Gevorgyan to furans 51 including THPO furan 51b

Kel'in and Gevorgyan's furans **51** were synthesised by Cu(I)-catalysed cycloisomerisation of allenes **52**, which could be formed *in situ* from the ketynes **53** by base-catalysed rearrangement; standard conditions were 1 mmol ketyne, 2 mmol Et_3N, 5 mol% CuI, at ~100°C in dry dimethylacetamide (DMAc, 2.2 mL) under Ar for 3–24 hours, giving yields of 2,5-disubstituted, and "5-monosubstituted" furans of 63–94% (only the product is symmetric, the distinction between 2 and 5-positions in the starting material is important). The ketynes themselves were easily be accessed from alkynes **54**. It is important to note the restricted set of substituents used in this paper (and others following similar synthetic routes[55]). No species **51** with R'=H were mentioned. In addition, the only "trapped 2-furanols" made by this route were a 2-methoxy-5-*n*-alkylfuran and a 2-tetrahydropyranoxy-5-*tert*-butylfuran. It is likely that the choice of these substituent pairs is no accident, but that the more sensitive THPO group *requires* a *tert*-butyl substituent opposite to achieve favourable cyclisation yield, while the more resistant methoxy group can be used with a simple straight-chain alkyl compound: this will be seen later in practice. In spite of these objections, no definite reasons to avoid this strategy were known, and as it offered the only good precedent it became the initial focus of synthetic endeavour (Figure II.8).

As applied to the present work, both THPO and βGalO groups were to be used to explore the syntheses and properties of functional trapped furanols, towards target **44b**. Since furanols were desired, either R or R' must be an alkoxy substrate group: clearly R is the only stable choice. As R is an alkoxy, a pre-existing allene moiety cannot be attached to it, but the allene must be formed from **53**. Therefore base-stable R and R' groups are required so that base-catalysed allene formation succeeds (either by the given mechanism or by a two-step reaction eg. with $NaNH_2$). Retrosynthesis from **53** might consider attaching the densely-functionalised alcohol R'CO-CC-CH_2OH to the desired substrate (eg. tetrabenzylbromogalactose), however these ketoalkynols are terribly unstable, decomposing above -20°C[56], so disconnecting the R'CO group as shown to start from alkynes **54** is the better choice. This requires an additionally base-resistant R group to survive the lithiation step.

Note that the densely-functionalised target **44b** presents further challenges than just installing the alkoxy substrate. The paucity of literature examples of furanol ethers is without doubt due not only to

the mechanistic challenges of synthesising them with more sensitive alkoxy groups, but also to the relative instability of the products once formed. 2-alkoxyfurans show a strong tendency to lose their protecting group and return to γ-butenolide form, and many are unstable even at low temperature or under nitrogen, even if 2,5-substituted furans are more easily handled.[53] However, the trapped furanol is reactive in a number of ways that phenyl or pyridyl analogues are not, and this will be seen to pose further synthetic problems downstream. For example, α-chlorination of a furan, such as a 2-furyl alcohol, may lead to total degradation,[57,58] although the presence of two α-substituents on the furan ring may slow the degradation and polymerisation experienced in those reactions.[59] Depending on their structure, furans and especially alkoxyfurans are known to be not only acid- but heat-sensitive (if 'heat' is defined as above eg. -40°C), to participate in Diels-Alder reactions[60], to be subject to ring opening reactions (eg. **50**) and solvolysis, etc.[42,46,61] Such complications will be addressed as they arise: in general, the reactivity problems will require both finding a method to install the alkoxy group, and doing so at the right stage so that minimal synthetic transformations afterwards are required.

Note that cyclisation of the alkynyl aldehydes directly to the 2-alkoxy-5-hydridofurans would provide an excellent synthetic route to the final targets: the 5-position could then be metallated and used either with 2-pyridinecarboxaldehyde or in a Petasis reaction, rapidly giving **44** (see **4**-**II** and **4**-**VII**, Figure II.25). The unattested alkylnyl aldehyde cyclisation was tried rather exhaustively during the M2 thesis[35] (details in section III.4.1.2). Over more than 20 cycloisomerisation trials, total decomposition was *always* observed, and my conclusion was that no alkynyl aldehyde cyclisation products would be isolated. This is perhaps unsurprising given the lability of the 2-unsubstituted alkoxyfuran structure as outlined above, and the successive rearrangements which can occur during the high-temperature allene cycloisomerisation mechanism. A later report by Gevorgyan *et al*. in 2003[62] showed that the cycloisomerisation of an alkynyl aldehyde *could* proceed, but for substrates which cannot be transposed to the present work: the only successful examples were with propargyl **thioether** analogues, with only **one α-hydrogen** [*not* with propargylic substrates, and *not* with substrates which can be sequentially deprotonated twice] and anyway these give the 3-substituted furans. Therefore at the very end of my M2 thesis I abandoned aldehyde cyclisation and attempted a ketone cycloisomerisation towards a 5-*t*-butyl-2-glycosyloxy-furan, similar to **51b** but examining whether the sugar would withstand the reaction conditions; a 25% yield was at once observed by crude NMR though was not isolated. So I began my PhD work with the aim to prove similar glycosyloxyfurans – a hitherto unreported type of aryl glycoside, which is something of a rarity in small-molecule chemistry in the 21st century – and to examine methods to functionalise them, towards the targets **44b**.

3. Synthesis of Furyl Spacers by Cycloisomerisation of Ketones

3.1. α-functionalisation strategy

As ketones appeared the only cycloisomerisable source of alkoxyfurans, the initial α-functionalisation strategy towards **44b** relied on the cyclisation of methyl propargyl ketones **56** to 2-alkoxy-5-methylfurans **57**, then methyl group activation to a reactive species **3-I**/**3-II**/**3-III** (Figure II.9).

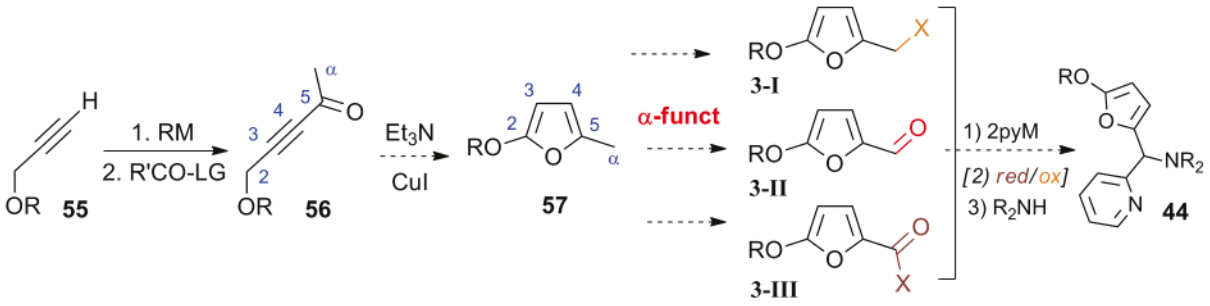


Figure II.9 – α-functionalisation strategy for targets 44b

3.1.1. Cycloisomerisations – common starting materials

A range of known propargyl ethers **55** were synthesised from propargyl alcohol by literature methods in usually good yields, and used in different roles throughout the upcoming syntheses (Figure II.10).

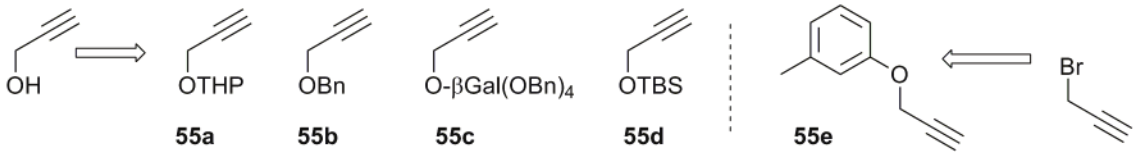


Figure II.10 - Syntheses of propargyl ether precursors on multi-gram scales

THP ether **55a** was synthesised with DHP and $POCl_3$ catalyst (Lingham *et al.*[63]; 78% yield), and was used as a cheaper model compound for the glycosyl targets. Benzyl propargyl ether **55b** was synthesised with BnBr and KOH in DMSO (modified from Trost and Xie[64] and Inamoto *et al.*[65]; 83%). 2,3,4,6-tetra-*O*-benzyl-D-galactoside **55c** was prepared in two ways: Fischer glycosylation then benzylation gave a mix of α/β anomers for a cheap reactivity model (Roy and Mukhopadhyay[66] then Mereyala *et al.*[67]; 40%); or the β-anomer was given stereoselectively from penta-O-acetyl-galactose via β-glycosylation, deacetylation and benzylation, involving more costly purification (Mereyala and Gurrala[68]; 29%). TBS ether **55d** was prepared with TBSCl/imidazole (White *et al.*[69]; 25%). Phenoxy ether **55e** was instead prepared from the phenol and propargyl bromide (cf. Ishii *et al.*[70]; 71%).

3.1.2. Ketone formation and cycloisomerisation

Acetylation of the propargyl ether acetylides **55** was relativly low-yielding with BuLi/DMAc or $NaNH_2/Ac_2O$ but proceded well with BuLi/Ac_2O giving alkynyl ketones **56**, though stoichiometry and order of addition were important for reducing yields of over-reacted **58**[71] and acetylated **59**

(Figure II.11). *t*BuLi was strictly required only for **55c**, the other acetylides were well formed with *n*BuLi.

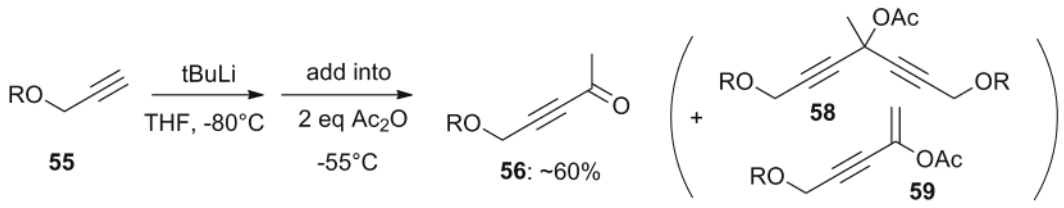


Figure II.11 – Acetylation of propargyl ethers 55 giving alkynyl ketones 56 for cycloisomerisation

The rather reactive[56] methyl propargyl ketones **56** were then submitted to cycloisomerisation. The crude reaction products required relatively fast workup as continual hydrolysis at the organic-aqueous interface was indicated, especially for the smaller substrates, and I observed continual generation of a black, robust-seeming interlayer which stabilised foaming, especially for glycoside **56c**.

Although the cycloisomerisation gave clean yields of the desired 2-THPO-5-*tert*-butylfuran in the original report[54], I isolated two major product types in my reactions. One type were the desired 2-alkoxy-5-methylfurans **57**, notably including **57c**, which therefore became what seems to be the first glycosylfuran to be known: no literature reports were found, and the only prior glycosylfuran of which I am aware is that whose distinctive NMR peaks were identified in the crude reaction products of the final reaction I performed during my M2 thesis (but which was not purified)[35]. The structures of both α and β anomers of **57c** were thoroughly characterised and reported in full (section III.4.2.2). As an entirely novel aryl glycoside, this structure is potentially of interest in a variety of applications in synthesis and pharmacology; purified **57c** seemed stable to air, but was not well-resistant to hydrolysis especially with trace acid when exposed to water.

The other product type had NMR indications of a nonaromatic system; and with the glycoside, two isomers of it were separated, each of which was confirmed as a β-galactoside, indicating a new diastereomeric centre on the proto-furanyl side. It did not show an allene peak in udeft ^{13}C-NMR[72] and submitting it to further heating with Et_3N/Cu(I) did not result in a change of NMR spectrum, so I ruled out the isolation of allene **3-IV** (which should continue to cycloisomerise). Since it showed three alkenic protons between 6.34-6.15 ppm plus two further unassigned protons at 4.46 and 4.17 ppm, I ruled out a ring-opened furan hydrolysis product (crotyl-type ester), supported by in-house IR which showed no carbonyl peaks but an unexplained weak band at 897 cm^{-1}. IR performed at an analysis centre instead indicated 4 carbonyl resonances and no weak band at 897 cm^{-1}. I assigned the enol ether acetal structure **60** anyway (Figure II.12), as it was supported by full NMR attributions of all resonances and HSQC and COSY confirmation, and postulated its fast hydrolytic/oxidative degradation during shipping as the reason for the difference between in-house and analysis centre IR spectra. The ring-opening hydrolysis of either **57** or **60** might generate the (amphiphilic?) species that

could have been responsible for the blackened foam interlayer (eg. a tetrabenzylsugar ketoester/ketoacid).

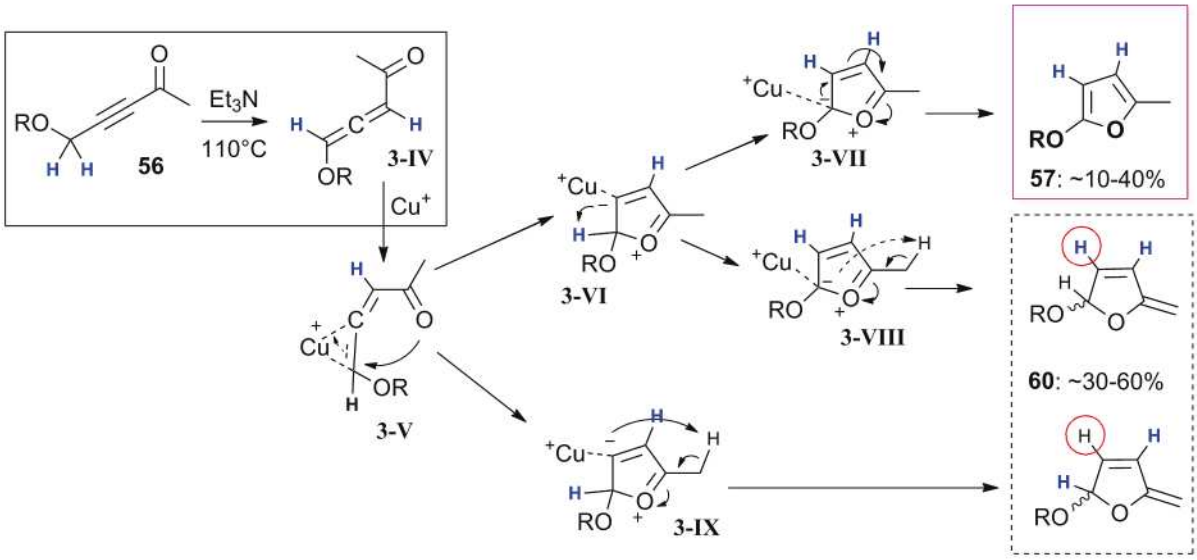


Figure II.12 – Synthesis of the target alkoxyfurans 57 in 10-40% yield, with cyclic enol ether-acetal byproducts 60. The mechanism of enol ether-acetal formation (via 3-VIII or via 3-IX) could be resolved by examining isotopic labelling in the products, with eg. a deuterated alkynone 56 (→either D or H at the circled position)

These reactions indicated that structures **60** were formed usually about equally with furans **57** in the crude mixtures, but **60** has apparently lower hydrolytic stability so was *usually* less recovered; the exception was with reactions towards the lower furans **57a** and **57b** where the furan was exceptionally volatile and easily lost almost entirely. Kel'in and Gevorgyan have not described such products **60**; this reactivity may be unique to 2-*alkoxy*-5-*methyl*furans, a substituent combination which they did not report on. For example, they report that methyl ketone **53c** with a 5-pentyl subsituent gave the furan in 94% yield (and presumably did not generate much, if any, of species **60**); and no such form **60** is possible for the THPO-furan which originally inspired this synthetic strategy (**53b**), since the *t*-butyl group prevents exocyclic elimination. I postulated two mechanisms of formation for **60** (Figure II.12); the internal-hydrogen-shift mechanism for the formation of **57** was established by Kel'in and Gevorgyan examining the products from deuterated reactants.[54] A deuterated propargyl ketone experiment could likewise determine whether the route for **60** is via **3-VIII** or **3-IX** or both, as they would give different deuteration patterns on the final product.

The high yield of enol ethers **60** represents a comparatively costly loss, but unfortunately they do seem inappropriate for adaptation into the desired synthetic strategy. A reaction of Joshi *et al.*[73] for tandem aromatisation/α-functionalisation of a 2-N-arylated-5-methylene enol ether acetal to 2-N-arylated-furfuryl alcohol with mCPBA was the only productive reaction using any similar substrate that was found searching CAS, but I strongly suspect that oxidative degradation would occur on these far more labile alkoxy substrates. I was therefore disappointed that **60** was not equilibrated with the furan when resubmitted to the original cycloisomerisation conditions as this could have increased the yield of the desired furan; either stronger base conditions (eg. LDA) or else palladium-catalysed allyl

isomerisation could potentially be tested (suggestion of Martin Ipuy) if this were at some stage desirable. In the meantime, the synthesis was continued using 2-galactosyloxy-5-methylfuran **57c**.

3.1.3. α-functionalisation trials and conclusion

With 2-alkoxy-5-methylfurans **57** in hand, I now examined possible methods for derivatising the methyl group. No required α-functionalisations were attested even for methylfurans (let alone alkoxyfurans), but a successful oxidation would have to be mild and nonacidic or else ring-opening would be certain[74]. I therefore trialled a radical halogenation in CCl_4, as a very apolar reaction medium which might avoid polar decomposition reactions; however this is not ideal since monohalogenation gives a species **3**-**I** which requires a *second* oxidative step later to attain the correct valency in **44** (cf. Figure II.9); and bis-halogenation would require potentially harsh conditions to convert the dihalide to the more useful furfural, so likely opening the ring. Indeed, polymerisation was indicated already during the halogenation, by the rapid, exclusive formation of black intractable tar, which has been seen before for furans exposed to even trace mineral acids.[38-40,57]

I therefore looked at two methods which might provide 2-alkoxy-5-furfurals **3**-**II** directly, and would not require oxidation state adjustment. The closest molecular precedents found in the literature were two methyl*benzo*furan oxidations to the benzofurfurals, though in the present case, without the stabilising effect of the annelating benzene, yields might be expected to be substantially lower. SeO_2 oxidation was trialled, based on a report of Ando *et al.*[75] where 2-methylbenzo[*b*]furan gave the aldehyde in 47% yield (the 3-methyl analogue gave 91% yield). Oxidation with freshly-prepared IBX[76] was also tried, based on the benzofuran oxidation of Nicolaou *et al.*[77] (50%): see Figure II.13.

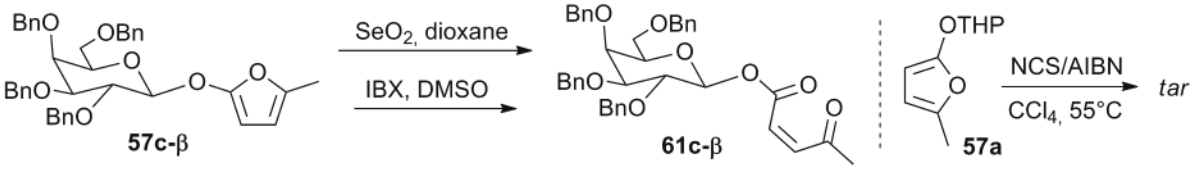


Figure II.13 – α-functionalisation trials of furans 57

These oxidations were examined in detail, however they gave only ring-opened Z-conjugated ketoester **61c**-**β** (though this in excellent purity even without separation). Simultaneously with this work, other reactions examining the direct formation of 2-(β-glycosyloxy)-5-furfurals **3**-**II** by S_NAr had yielded success with good yield (see section II.4.2.3). Therefore as the current targets **3**-**II** could be sourced by other means, I abandoned further attempts to functionalise the furan α-position; it was considered sufficiently rewarding to have been able to thoroughly characterise the glycosylfurans **57c**, where no glycosylfurans had been characterised before (Figure II.14).

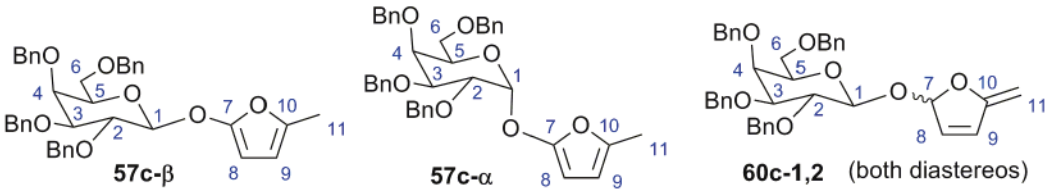


Figure II.14 – The first known examples of glycosylfurans 57c had been synthesised by the cycloisomerisation route. They and their isomeric enol ether acetal products 60c were thoroughly characterised.

3.2. Pyridyl preinstallation strategy

The cycloisomerisation route had not however exhausted its synthetic utility. It was not known whether furfurals **3-II** from S_NAr (section II.4.2.3) could reliably be reacted to produce **3-X** (Figure II.15): the furan, or the glycosidic bond, might for example undergo substitution or base-promoted decomposition, during reaction with metallated pyridine. Therefore I explored a cycloisomerisation route to produce *pyridyl-bearing* alkoxyfurans **3-XII**, as a means to circumvent S_NAr if needed.

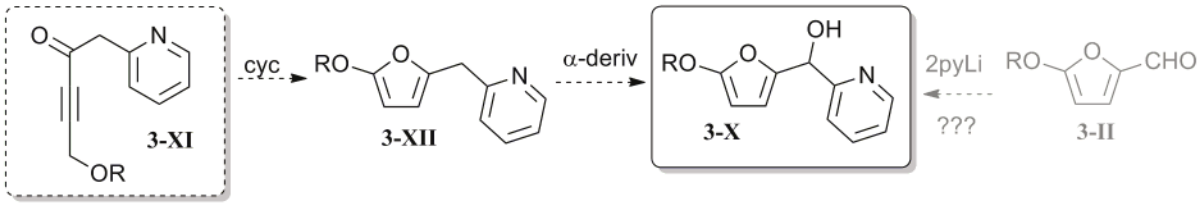


Figure II.15 – Pyridyl preinstallation pathway

One potential problem with this approach is that picolyl alkynones similar to **3-XI** have been reported by Natarajan *et al.*[78] to perform intramolecular cyclisation onto the alkyne terminal position under heating, giving quinazolin-2-ones. However, the reported substrates all had simple alkyl/aryl groups at the alkyne terminus. Alkyloxy groups as in **3-XI** were not examined, but the change in reactivity is potentially large. Additionally, it was hoped that with appropriate cyclisation conditions/catalyst, that preferential cyclisation to the furan could be achieved. Evaluation of whether this would be prohibitive or not, would have to await synthesis. Three routes to the key picolyl propargyl ketones **3-XI** were explored (Figure II.16). The two pathways with no changes of oxidation state would insert the ketone carbonyl between acetylides **55** and the picolyl group, by attack of either the metallated acetylide or the metallated picolyl group on the carboxylic acid derivative of its partner **3-XVI** or **3-XVII**, stopping the addition at the ketone state. The oxidative pathway would instead proceed by attack on the corresponding aldehyde (no over-reaction possible), and perform (hopefully favoured) oxidation of the resulting **3-XIV** to the conjugated ketone **3-XI**.

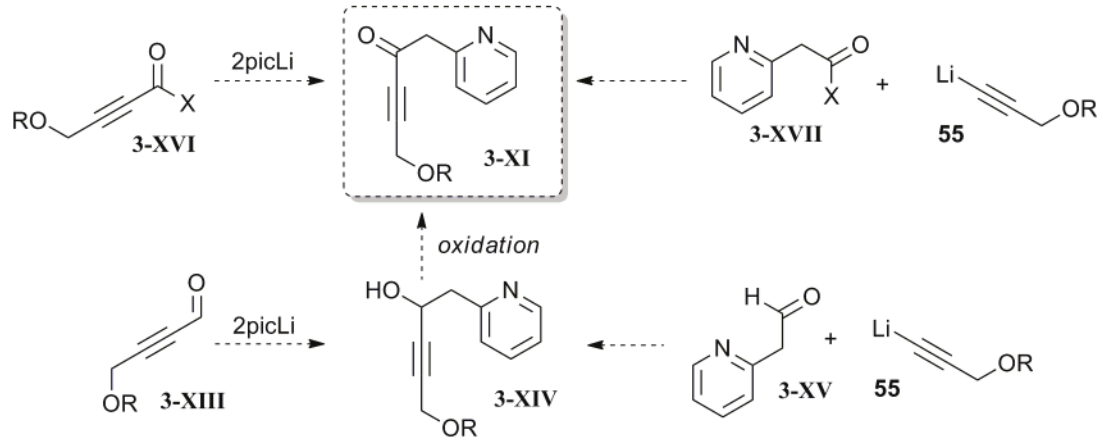


Figure II.16 – Routes identified towards key cyclisation precursor 3-XI

3.2.1. Oxidative pathway

The oxidative pathway was the first to be attempted since alkynyl aldehydes **3-XIII** were available from work during the M2 thesis, and the aldehyde ensures that only mono-addition is seen during coupling with the alkyllithium. THP ether **62a** had been obtained by an excellent literature protocol of Journet *et al.*[79], by which intern Claire Weisslinger (CW) also synthesised the unreported tetrabenzylgalactosyl analogue **62b**. The reaction of **62a** with 2-pyLi to the alcohol **63a** as per a protocol of Wang and Ding[80] (given in more detail than the original[81]) was uneventful, proceeding in 63% yield on a multi-gram scale (Figure II.17). The reaction of **62b** instead appeared prohibitively lower-yielding (<<15% in crude, no isolation: CW) so while that reaction was being retrialled (CW), I examined the reaction of 2-(2'-pyridyl)ethanal **3-XV** with **55** as a second approach.

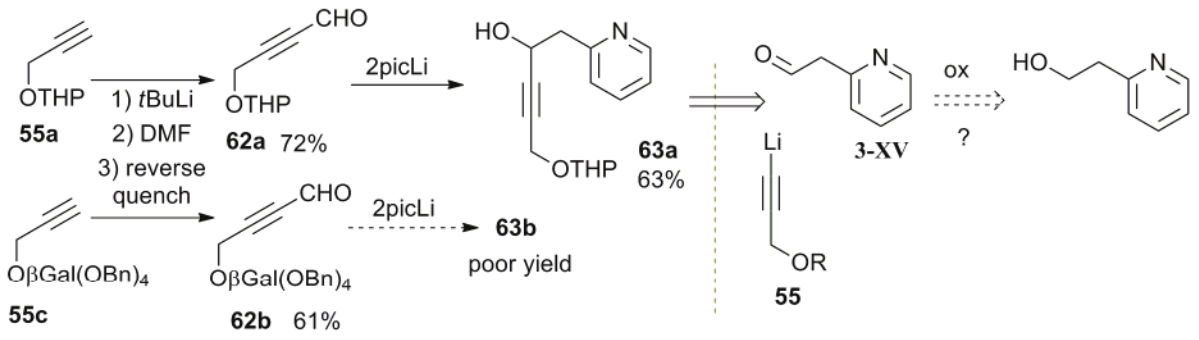


Figure II.17 – Initial steps in the first oxidative pathway, and consideration of the second possibility via 3-XV

Surprisingly, there *is no such aldehyde **3-XV** reported*. Searching CAS returned 7 papers and 3 patents citing this molecule (and none for its enol tautomer). Of these, 5 of the papers are misclassified (the real compound is 2-pyCHO); the sixth[82] does not give any details of how this unprecedented aldehyde was obtained – although they discuss its reaction product (molecule 25c in their notation) in some detail. A final report[83] showing Ce(IV) oxidation of alcohols to aldehydes confusingly lists 2-pyCH_2CH_2OH in both substrate *and product* columns of its synthetic report, and details a yield of

90% for this net zero transform, without further explanation; all the other substrates are correctly listed and also have ~90% yields. Two of the patents are again misclassified; the third[84] does report synthesising 2-(3-chloropyridin-2-yl)propanal by Swern oxidation (Example 38D in their notation). Their attribution of a non-delocalised system contrasts with what is reported for 2-pyridylacetones, which are usually in enol form[80,85,86], however only an experiment could examine this.

I therefore synthesised the Ce(IV) reagent, $(NO_3)_3CeBrO_3$[83], but no 2-pyCH_2CH_2OH oxidation products were isolated during separate reaction trials; in parallel, benzyl alcohol controls gave PhCHO in excellent yields, so I concluded that the protocol had been well-implemented but really was unsuitable for this substrate. I also attempted a Swern oxidation[87], but unlike the patent reporting 97% yield of the aldehyde, I obtained a complex mixture in the crude; although a small triplet aldehyde peak at 9.71 ppm which *might* be attributable to 2-pyCH_2CHO was seen, it was not isolated. I did confirm the novel *Z*-enolised-aldehyde thioacetal **64a** (2% recovered) which shows that the initial oxidation to the aldehyde/enol was successful. I also did observe fractions with the desired mass, MH^+=122 Th, but this was probably a minor product; NMR indicated that the majority products in those fractions were oxalyl ester **64b** (2% recovery) and the starting material (Figure II.18).

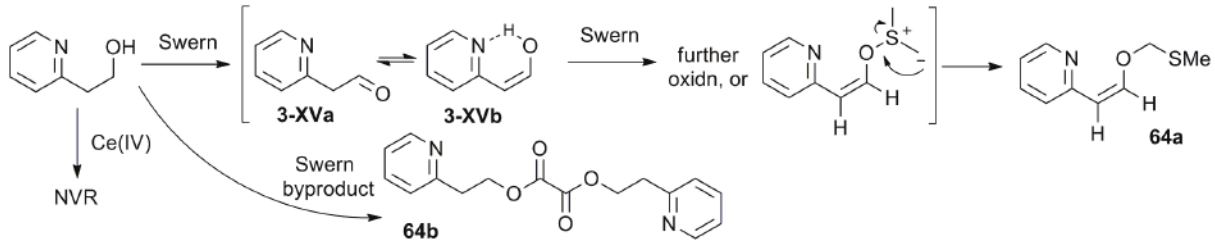


Figure II.18 – Oxidative attempts towards aldehyde 3-XV

A last attempt to isolate the novel aldehyde **3-XV** avoiding oxidative methods relied on formylation of 2-picolyllithium with DMF. A ~quantitative yield of dimethyl enamine **65a**[88] was initially obtained, possibly because the quench conditions were not sufficiently[79] acidic to expel the NMe_2 group from the intermediate. Acidic quench (H_3PO_4/KH_2PO_4) instead gave extensive degradation, with a small amount of the enamine as the only identifiable product, although **3-XV** could not be ruled out. Lastly, an *in situ* quench of the formylation with oxophilic TBSCl delivered E/Z-**65b** in 99% crude yield, though this degrades upon storing even at 4°C and is destroyed on silica. The silylated enol **65b** can probably be deprotected to **3-XV** (anhyd. TBAF) if this is desired out of interest: like the glycosylfurans, it would be another novel small-molecule heterocycle of (seemingly) useful structure which has been hitherto unreported (or misleadingly reported). However, at this moment, the final target ketones **3-XI** were obtained by the pathway via **3-XVI**, so this approach was dropped.

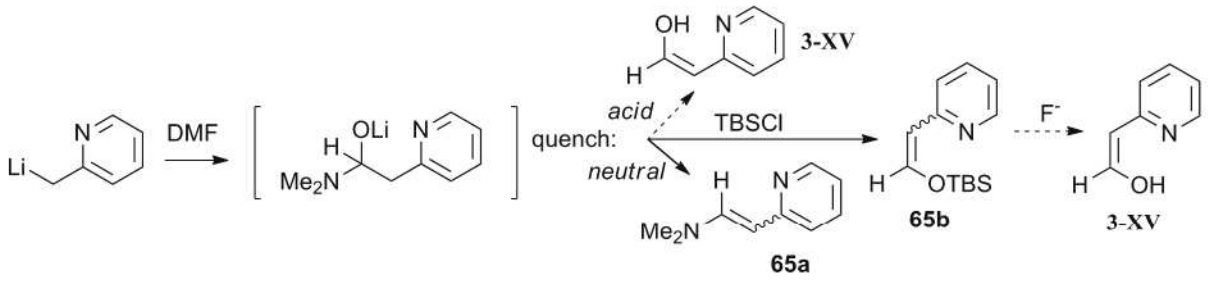


Figure II.19 – A trial towards 3-XV was stopped since its reactivity appeared unproductive for the desired synthesis

Concurrently, the oxidation of **63a** towards **3-XI** was being evaluated to see if the oxidative pathway was worth pursuing further (eg. involves finding correct synthesis conditions for **63c**). An initial trial used MnO_2 in DCM[89,90], the most relevant precedent (attested for a highly sensitive substrate[91]). No desired ketone could be evidenced, or any products identified conclusively; the desired product *mass* was however detected on LCMS and formation of a conjugated species was indicated as the product-mass peak had a significant absorption at 310 nm, whereas the starting material does not absorb there. I did not however assign the enol structure (**3-XI** enol) to this compound as the required alkenic protons were not visible, and the product remains unidentified (Figure II.20).

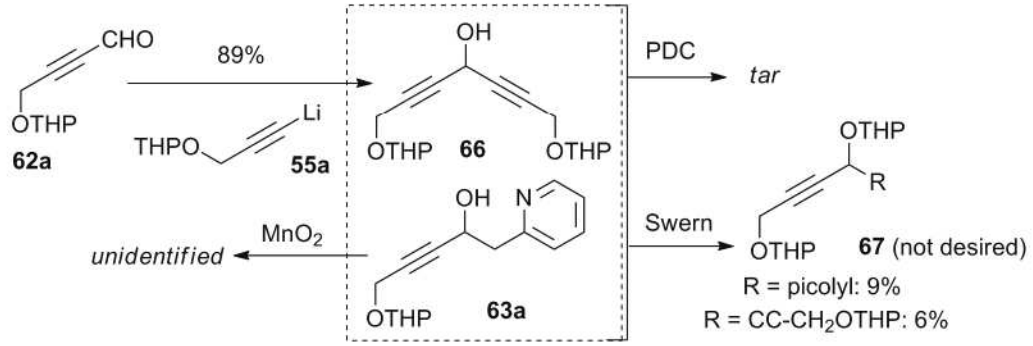


Figure II.20 – Final oxidative reaction failures with propargylmethanols 63a and 66 led to abandoning this route

The non-picolyl-bearing alkynol **66** was then synthesised as a test substrate to examine whether degradations could be attributed to the THP group or to the picolyl, so that if the THP protecting group was 'at fault' it could be exchanged. PDC oxidations[92] of both substrates **63a** and **66** were attempted, and both showed rapid tarring and no identifiable products. A Swern oxidation[85] was the next reaction attempted; both substrates **63a** (generated an *un*usually foul and pervasive stench) and **66** (acceptable smell) gave complex mixtures in which the only identifiable components were unreacted starting material, and small amounts of products **67** showing intermolecular THP group migration (Figure II.20). This does not rule out product formation and degradation: oxoalkynols are known to be extremely sensitive[56]; however, the oxidative route seemed less attractive than ever, and the non-oxidative pathways via **3-XVI** and **3-XVII** were targeted instead.

3.2.2. Non-oxidative pathway 1: acetylide attack on pyridylacetic 3-XVII

Experiences with the unusual reactivity of pyridyl ketones (especially, their undesired acidity, rather than desired electrophilicity) had suggested that strong-base acetylides might simply deprotonate pyridylacetic species **3**-**XVII** to stable intermediates, returning the starting materials after quench. Indeed, initial trials of acetylide addition onto commercial pyridylacetic ester **68a** gave no visible reaction. Nor did attempts at preparing its Weinreb amide **68b** by MeN(OMe)-MgCl attack on the ester succeed, though testing several protocols including one reported for 3-pyridylacetic acid[93,94]. Using nonbasic coupling, mediated by DCC as per the protocol of Baxendale for nicotinic acid[95] did deliver the apparently novel **68b** in 66%. However across several trials of lithium and magnesium acetylide additions with both **55a** and **55c**, no productive reaction was observed and typically only 2-25% of **68b** was recovered from among some highly fluorescent byproducts, the missing >75% of the pyridyl group could not be identified. One more adventurous reaction examined the relatively unused acid chloride hydrochloride **68c**.[96] As it had not been reported with acetylides and is comparatively little used in the literature, I first desired to check whether it would behave as for a traditional acid chloride. Attempting amide formation with piperidine, it instead underwent oxidative/decarboxylative/dimerisation rearrangements generating amusing fluorescent quinolizinone product **69**, and a second decarboxylation product whose structure was not satisfyingly determined but showed many features of compound **3**-**XVIII**. Upon successfully assigning the structure of **69** (from a decarboxylative dimerisation/rearrangement sequentially involving three molecules of acid chloride), it was decided that **68c** was unpromising for use with acetylides as they too might trigger similar decarboxylations and rearrangements (Figure II.21).

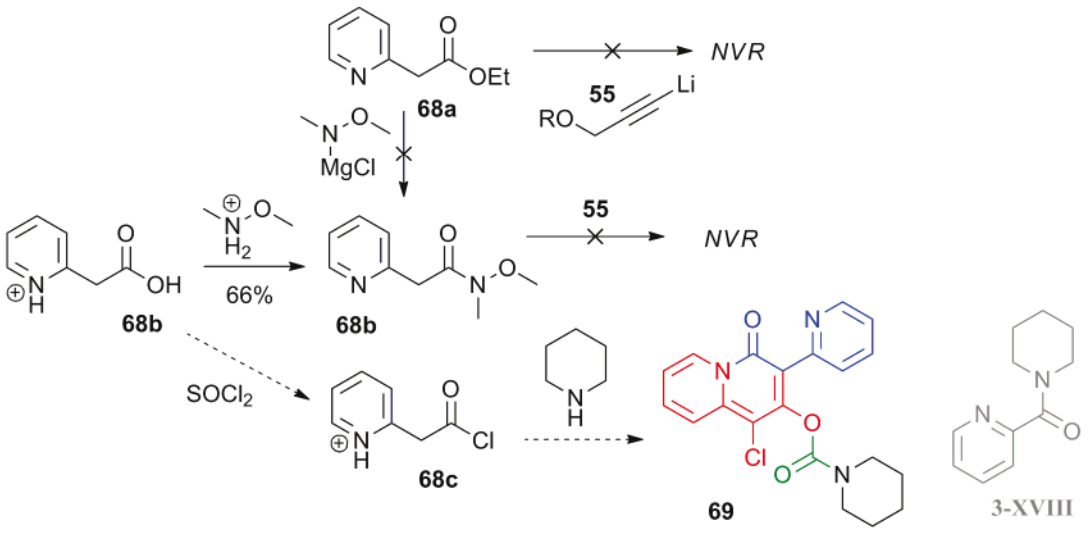


Figure II.21 – Failure of 2-pyridylacetic derivatives to give desirable behaviour for alkynyl ketone formation

Final reaction trials confirmed the general difficulties of the **3**-**XVII** pathway. Control substrates to attempt to evaluate the 2-pyridyl moiety's responsibility for initial reaction failures were first synthesised, choosing the known 3-pyridylacetic[93] Weinreb **70a**, prepared by DCC coupling (42%),

and phenylacetic Weinreb **70b**, from the acid chloride (87%)[97]. Then, a report from Raghavan *et al.*[98] suggested that acetylide addition to α-halogenated Weinreb amides can proceed without difficulty. Therefore, dibrominated Weinreb amide **71a** was sought, with the idea that halonium elimination should be much more challenging than proton elimination, so the enolisation process might be reduced; it was isolated along with the corresponding monohalide **71b** (its phenyl monobromo analogue **71c** was also obtained). These substrates were then submitted to reaction with acetylides **55a**. Phenyl Weinreb **70b** reacted acceptably (**72** obtained in 61% yield), confirming the possibility of this reaction protocol working; yet reaction was not observed with its α-halo derivative **71c**, and as pyridyl weinreb amides **70a**, **71a** and **71b** all failed to give reaction products, this route via **3-XVII** was abandoned (Figure II.22).

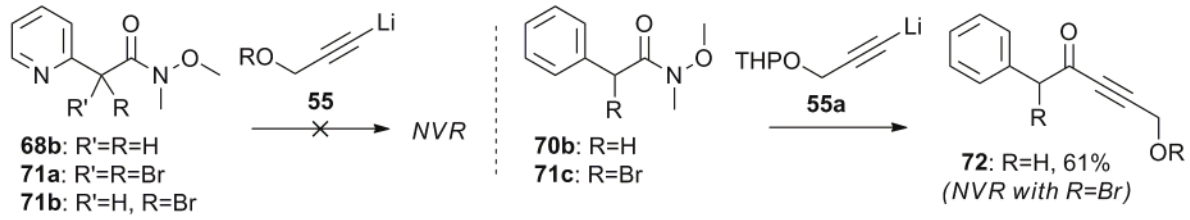


Figure II.22 – Pyridyl Weinreb amides failed to give the desired addition products with acetylides (3-pyridyl 70a not shown, likewise gave no reaction); phenylacetic Weinreb amide 70b did however give 72.

3.2.3. Non-oxidative pathway 2: picolyllithium attack on alkynyl ester 3-XVI

THP-bearing alkynyl ester **73** was prepared by addition of the acetylide of **55a** to methyl chloroformate (78%) but surprisingly, in my hands this reaction was at least initially unsuccessful with tetrabenzylgalactosyl **55c** (two trials). Proceeding with **73**, the addition of 2-picolyllithium at 0°C gave a complicated reaction mixture, with a distinct 'rodent' smell as of dessicated rat, containing multiple products, including mono-addition **74** (11% in crude, not isolated), quinolizinone **75** (10% in crude, 4% isolated), bis-addition **76** (6% isolated), and aldol dimer **77a** with its SiO_2-accelerated, extensively conjugated dehydration product **77b** (combined isolated yield 3%): Figure II.23.

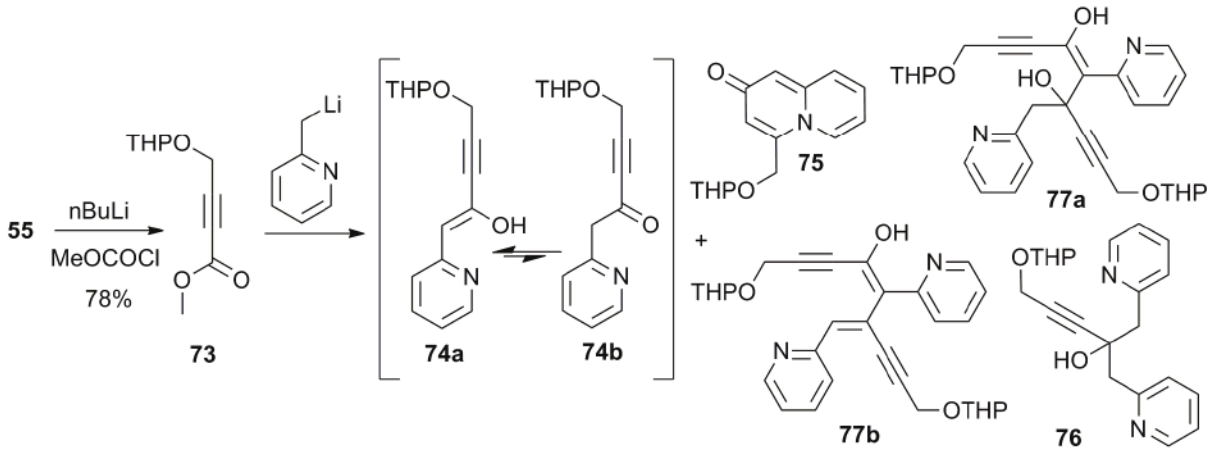


Figure II.23 – Complex initial reaction products from the alkynyl ester route

Unhappily, the desired picolyl alkynone **74** was not isolated after column. By the crude NMR, its preferential form in $CDCl_3$ was also shown to be enol **74a** (distinctive alkenyl singlet ^{1}H peak at 5.77 ppm) rather than the ketone **74b** (distinctive ketone ^{13}C peak at 183.5 ppm) though an accurate keto:enol ratio could not be determined as the extensive peak overlap prevented reliable spectral deconvolution. The crude mix was substantially composed of bis-addition products, which is as expected for ordinary ketone intermediates (more reactive than the starting material ester) but in this case somewhat surprising, since ketones **74b** are certainly deprotonatable (cf. non-reactivity of **70a**) and it had been hoped that this might prevent bis-addition; enolate attack products **77a**/**77b** were evidence of such deprotonation (and the absence of a corresponding ester-addition product is in line with the greater reactivity of the ketone).

To lower the bis-addition products, two approaches were taken. Firstly, my observations of the colour change during the reaction suggested that the consumption of 2-pyridyllithium took only seconds; longer reactions woud simply allow more time for the degradation of the product. Furthermore, to obtain the best yield of the mono-addition product, the product of the concentrations of the ketone and the acetylide should be minimal during the whole reaction, which is not ensured by either addition of the lithium to the ester, or addition of the ester to the lithium. However, continuous flow reactors[99] present an excellent solution. A cooled (0°C) continuous-flow reactor with a reaction time of around 10 seconds before cold quench/extraction was assembled out of standard lab glassware (see section III.4.3.3) and the ester reaction was re-run with good results – the crude contained 33% of **74**, along with 60% of unreacted **73**; only 3% of unwanted quinolizinone **75**, and low yields of undesired **76** (<5%) and **77a**/**77b** (<1%) were seen, as well as some unidentifiable pyridyl species. Although this indicates that a longer reaction time should be used (eg. 20-40 seconds), the amount of product synthesised was sufficient for further work. In parallel, the alkynyl Weinreb amide **78a** was prepared (56%); substantial conjugate addition of water to the alkyne giving the β-ketoamide byproduct was also seen (**78b**, 6%). **78a** gave even cleaner formation of the mono-addition product **74** than did **73** in the stopped-flow reaction (44% in crude, no bis products: Figure II.24).

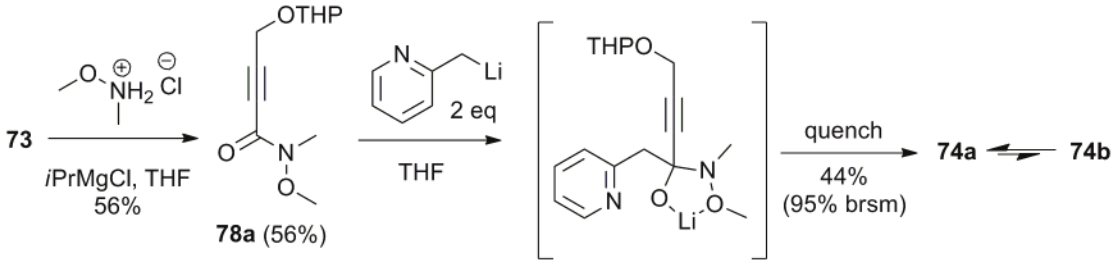


Figure II.24 – Weinreb amide 78a was used to obtain a cleaner yield of 74 than was possible using ester 73

Unfortunately, NMR of **74** showed its continual cyclisation-degradation to form **75** when left at room temperature (halftime est. hours) and even at 4°C, so working quickly was now essential since the lab was not equipped for low-temperature workup. A fast column of the freezer-stored crude was used to

purify some **74** immediately before launching cyclisation trials. A second approach to purification would be derivatisation in a form that blocks cyclisation (similarly to **65b**). Trapping the crude reaction products in the enol form by reaction with TBSCl was examined briefly both *in situ* and also after quench, taking reaction crude into TBSCl/DMF with imidazole or NaH, but in neither case was the silylated enol indicated, and this approach was abandoned.

3.2.4. Cycloisomerisations

Purified **74** was then quickly exposed to a variety of coinage metal catalysts in DCM, hoping to synthesise **3-XII** by preferential cyclisation to the furan. Given the strong tendency of **74** to cyclise to **75** spontaneously, forming the furan (should have distinctive shifts for furan H3 and H4 in ^{1}H-NMR, doublets at ~5.74 ppm and ~5.25 ppm) would not be easy. The choice of catalysts and conditions (DCM or THF, RT) was informed by recent papers from the Gevorgyan group discussing further developments in the cycloisomerisation of allenyl ketones to furans[100,101]. Unfortunately however, desired **3-XII** was never observed and all catalysts seemed to either destroy the samples, or allow/promote the cyclisation to quinolizinone **75**; an interesting exception was near-stoichiometric reaction with silver tosylate which gave a different reaction product, to which I assigned structure **75x**; this may be due to efficient coordination of the pyridyl nitrogen by the silver ion, which prevents N-nucleophilic cyclisation to **75**.

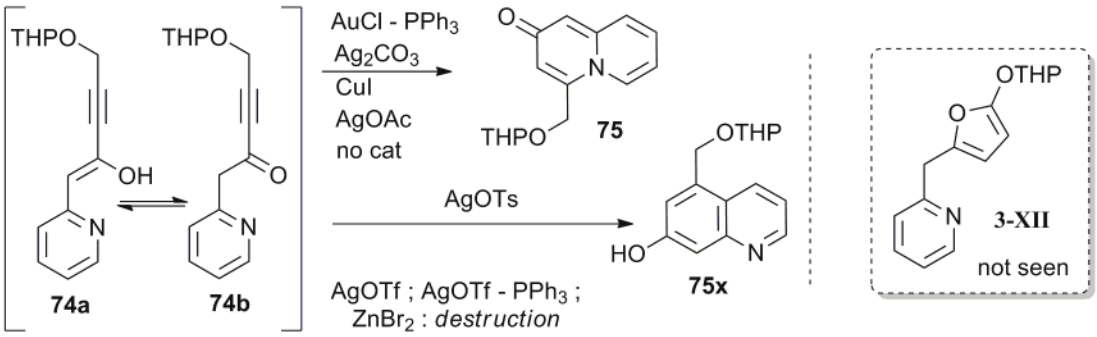


Figure II.25 – Attempted coinage-metal-induced cyclisations of 74 did not allow isolation of picolyl alkoxyfurans 3-XII

3.2.5. Synthetic conclusions from Part II.3 (cycloisomerisation)

The conclusions from this section were relatively clear: alkoxymethylfurans **57** are too sensitive to allow α-derivatising for subsequent pyridyl attack, but picolyl alkynyl enols such as **74** are too predisposed towards pyridyl-ring cyclisations to allow formation of alkoxypicolylfurans in ordinary conditions. Blocking the ring pyridine, or preforming the allene, could be ways to address this latter reactivity, but the chemistry developed in this section was already adventurous and time-consuming; further research into such unprecedented transforms was considered unlikely to give success efficiently. Synthetic efforts were therefore redirected towards a third method of installing the required alkoxy and picolyl groups on the furan nucleus: nucleophilic aromatic substitution.

4. Synthesis of Furyl Spacers by S_NAr

4.1. S_NAr precedents

The *S_NAr of an alkoxide on 2-halo-5-carbonylfurans* has been well-reported for primary or secondary alcoholates, usually attacking furoic acids[102] and esters.[48,53,103] The resultant esters are usually reduced directly to the alkoxy*carbinols* with DIBAL,[48] or saponified[53] to the alkoxyacids; the decarboxylation of the alkoxyfuroic acids is reported to be easy (Cu/quinoline/heat).[48,53,102,104] Nitrofurans sometimes give similar substitution, especially with phenoxides[105-108].

This S_NAr reaction provides a second approach towards installing an alkoxy moiety on the furan, as was proposed by Olivier Piva during the M2 thesis defence leading to this PhD work. However, these alkoxy groups are all derived from alcohols: no reports have addressed hemiacetals eg. sugars. However, a recent report of Morris *et al.*[109] was attractive. They used two sugar lactol alkoxides in a few S_N reactions including S_NAr on (activated) 2,4-dinitro-halobenzenes (~85%); the α/β-ratio can be influenced (at least for the 2-deoxylactols) by choice of conditions: KHMDS/THF/-78°C for α, NaH/dioxane/23°C for β.[109] Desiring β-alkoxyfurans, and with the difficulties of α-functionalisation and the impossibility of pyridyl preinstallation (sections II.3.1-II.3.2), the possibilities of applying such an S_NAr on a furan was examined in more detail (Figure II.25).

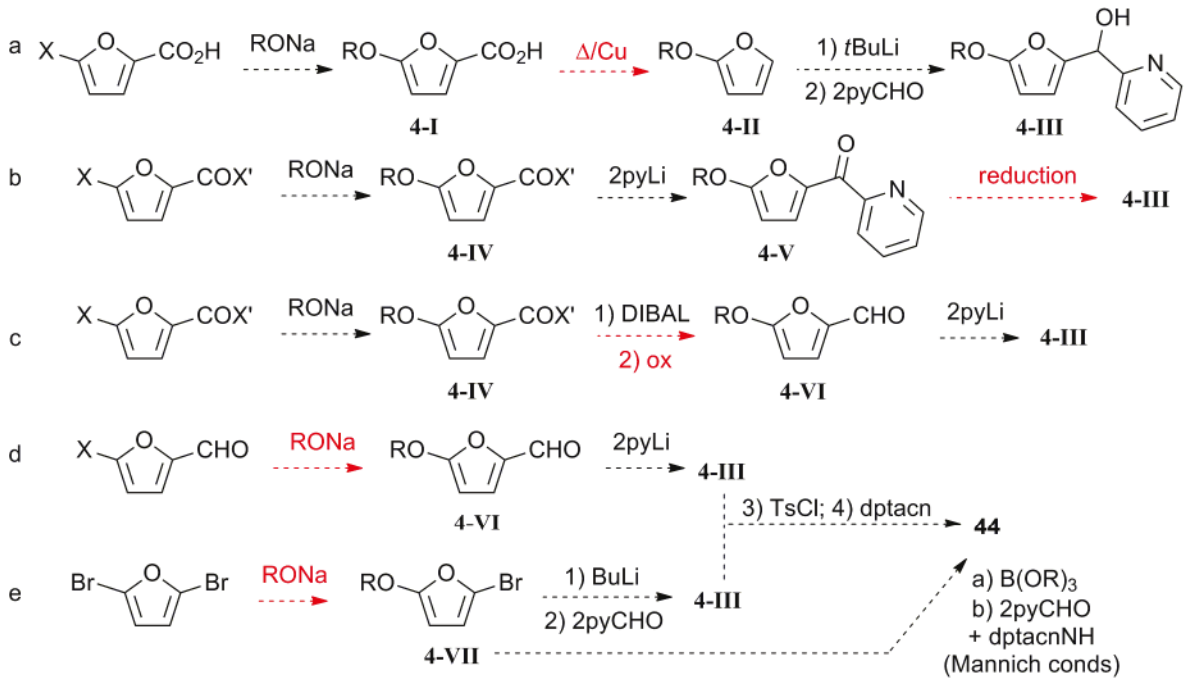


Figure II.26 – Precedents and plans for S_NAr approaches to alkoxyfuran creation and use

The α-carbon, between the furan and coordinative rings in the desired products **4-III**/**44**, is a *methine* group, ie. derived from an aldehyde. The *primary* published furan S_NAr methods could approach this in three ways, but which I ultimately considered relatively unpromising (Figure II.25). These were: (a) decarboxylated 2-alkoxy-furans **4-II** might be 5-lithiated with *t*BuLi for attack on 2-pyCHO

towards **4**-**III**; (b) the reaction of alkoxyesters/Weinreb amides[110] **4**-**IV** with 2-metallated-pyridine could form ketone **4**-**V**, for later reduction to **4**-**III**; (c) the alkoxycarbinols from DIBAL reduction of **4**-**IV** *might* be oxidised to the furfurals **4**-**VI**, for use with metallated pyridine giving **4**-**III**. These precedents and plans are summarised in Figure II.25.

These routes present potentially problematic pitfalls (marked in red on the figure). The likely instability of alkoxyfurans **4**-**II** has been discussed,[53] and their decarboxylative formation requires that the enzyme substrate in **4**-**I** resists these conditions (at a minimum, heating under pressure above 100°C). I especially mistrusted oxidation (c) of alkoxyfurans, since I had already seen them to be oxidation-sensitive in general, and they are known to be acid-sensitive (common byproduct of several oxidants). I also disliked reduction (b) as **4**-**V** is predicted to be an unusually stabilised ketone (related pyridylfurylmethanones form spontaneously from the alcohols under air[111]) so its reduction might need rather powerful conditions. Both oxidations and reductions would potentially involve substrate/protecting-group problems too. Unfortunately, the synthetic target **4**-**III** common to all routes is not only likely to oxidise back to **4**-**V**, but its later use will require forming its secondary electrophile analogue and substituting it onto the secondary amine dptacn to obtain final ligand product **44**. *If* standard conditions can be used, this reaction can be high-yielding eg. by chlorination, tosylation etc and then grafting[25]; however, if the alkoxyfurans proved to be at all acid-sensitive or to have strong eliminative tendencies then providing them with a leaving-group could be problematic (as was indeed observed, see section 4.2.3), even if mild conditions for establishing it are used (eg. N-chlorosuccinimide/SMe_2,[112] which has successfully been used on a 2-furyl alcohol, though a 2,5-dialkyl one[113]). Lastly, none of these methods is direct; and the entirely *general* lability of alkoxyfurans[51,52,102,114] led me to focus on more direct methods with fewer possibilities for degradation along the way.

One more direct method than these three approaches would be (d) exploiting the S_NAr reaction on the 2-nitro-5-furfurals which has been seen for *eg.* phenoxides[107,108], or potentially on the halofurfurals (no relevant examples). This would give the aldehyde oxidation state in **4**-**VI** directly, which might be reacted to **4**-**III** with eg. 2-metallated-pyridine[115]; Grignard reagents for example easily add to furfurals in standard conditions,[116,117] as do alkyllithiums,[111] although unsurprisingly no reactions involving 2-*alkoxy*-5-furfurals have been found.

A second direct method (e) would depend on whether 2,5-dihalofurans[118,119] are activated enough to give S_NAr (totally unattested). If so, the resultant alkoxyhalofurans **4**-**VII** would be ideal lithiation candidates (*n*BuLi) for reaction with aldehydes[120], eg. 2-pyCHO, towards **4**-**III**. This route also offers a softer alternative to passing by common intermediate **4**-**III**, since the bromofuran may likely be converted easily to the boronic acid for a Petasis reaction (boronic acid Mannich reaction).[121] This three-component procedure was recently improved *specifically* for 2-pyridinecarboxaldehyde and secondary amines[122], and gives species like **44** directly in high yields in one step (75-95%)[121-123],

without having to put a leaving-group at the α position to the furan.[124] The boronate used may be chosen for a balance of stability and reactivity[125,126]. Note that intermediate **4-II** from route (a) also allows the Petasis reaction, by metallation then quenching with a borate ester, but the decarboxylation step was considered probably prohibitive.

4.2. S_NAr syntheses

4.2.1. Alkoxides and substrates

Trial S_NAr reactions were commonly performed with commercial alcohols. BnOH could provide a conveniently monofunctional reactivity mimic for various PABA/PHBA-stacked substrates; 4-MeO-PhOH was examined as an electron-rich substrate mimic; and *p*NO_2BnOH could provide a method to demonstrate auto-immolation triggered by nitro group reduction (though achieving this orthogonally by chemical means for such structures is problematic, so NTR may be a better solution[29,127]). Differently-protected glycosidase substrates were also prepared in both simple and stacked-spacer configurations; they were intended both to examine and optimise syntheses (blocking groups), as well as eventually to provide a route towards final targets **44** (mildly- and orthogonally-deprotectable groups). These substrate syntheses are shown in Figure II.26.

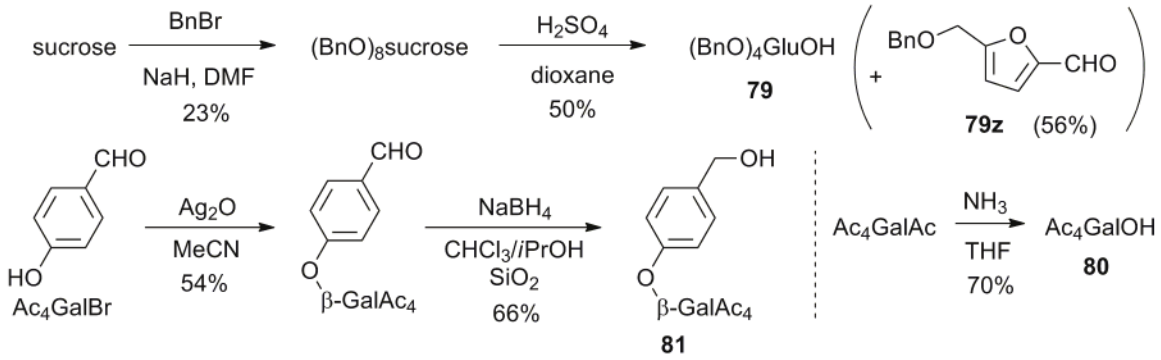


Figure II.27 – Successful syntheses of glycoside substrates for S_NAr reactions

Tetrabenzylated future β-glucosidase substrate **79** was synthesised on large scale with the cheap and convenient procedure of Yamanoi *et al.*[128] as a blocking-group example for evaluating direct S_NAr of the sugar onto the halofuran; the benzyl group is probably inappropriate for use in final probe syntheses, since its deprotection by hydrogenation may likely cleave the whole auto-immolative coordinating arm from the complex. High yields of a benzyloxy-bearing furfural were isolated from the reaction, but NMR showed it to be only the per-elimination byproduct from the fructose ring (**79z**); note that the stability of this species to acid is expected for a 5-*alkyl*-furfural, and cannot be expected for a 5-alkoxyfuran.

Grossmann *et al.* showed that standard acetyl deprotections destroy furanones, but developed a milder guanidine method that was suitable,[47] so as an initial deprotectable group I examined acetates. Tetraacetylated future β-GAL substrate **80** was cleanly delivered by selective anomeric deacetylation of galactose pentaacetate by a literature protocol[129]. Tetraacetylated future β-GAL stacked substrate

81 was synthesised by a 2-step literature method using Ag_2O-mediated coupling to 4-hydroxybenzaldehyde[130], as (cheaper) $BF_3{\cdot}OEt_2$-catalysed reaction of the galactose pentaacetate, and a reported single-step substitution onto PHBA[131] both gave intractable behaviour in my hands.

It was also highly desirable to obtain substrates with more base-stable removable protecting groups. The TBS group was initially considered the most promising non-ester protecting group, as its removal is probably orthogonal in mild conditions. I therefore made several efforts towards tetra-TBS derivatives in simple and stacked configurations, involving tetramethyl and tetrabenzyl analogues in protocol trials (Figure II.27).

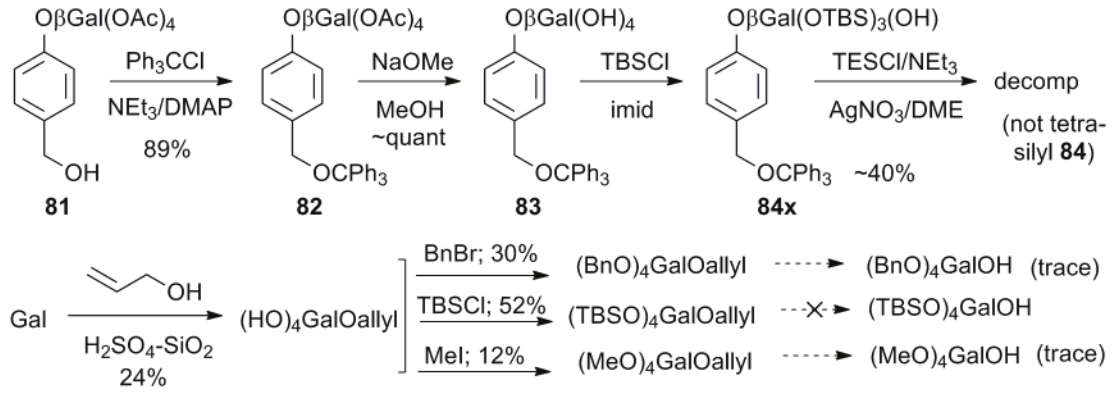


Figure II.28 – obtaining galactosyl species with different protecting groups than Ac was initially low-yielding

For the stacked spacers, with the β-configuration installed in **81**, the protecting groups were swapped towards the corresponding tetra-silylated species **84** by three-step alcohol protection, deacetylation, and silylation following a patent report[132]. However only three of the four silyl groups were reliably installed by this method (**84x**). An initial attempt to install the missing silyl group by forcing conditions[133] (and a more permissive TES silyl group) resulted instead in decomposition.

For the simple GalOH substrates, I hoped that without the PHBA glycoside of **84**, the silylation could be quantitative and without side-reactions, so I sought a smaller and orthogonal group for the anomeric protection. Allyl galactoside was cheaply formed by acid catalysis[66], and protected in a variety of ways so that I could find a reliable deallylation method to apply in the synthesis of desired $(TBSO)_4GalOH$. Thus tetrabenzylation[128], tetrasilylation[134], and tetramethylation[135] of the glycoside were performed though in relatively low yields; formation of the hindered tetrapivaloyl ester as a more base-stable alternative to the acetyl group (no deprotonatable α-hydrogens) was attempted with both standard and forcing conditions[136] but in my hands this was unsuccessful on the rather crowded substrate. Deallylations to the tetraprotected galactoses were examined repeatedly with a convenient procedure using NBS-acetone-water reported specifically for similar substrates[137], however that reaction could not be evidenced and decomposition products were instead obtained. Deallylation with $PdCl_2$/MeOH allowed isolation of tetramethylgalactose in 25% yield on a small scale (performed by intern Nils Aronssohn) but was unsuccessful for the desired TBS species. Deallylation was also attempted by the *t*BuLi procedure of Bailey *et al*.[138] giving deallylation for both

benzyl and methyl protecting groups but once again resulting in apparent destruction of the TBS analogue.

The methyl group was intended only as a stable blocking group to explore synthetic pathways as its removal requires harsh conditions which will likely decompose a furan; it gave no real synthetic advantages in the end so it was not pursued further: benzylated analogues were more easily accessible and should have similar base-stability. While no tetra-TBS substrate had yet been purified during these steps, these reactions had begun to consume more time than simply proceding with the established substrates **79**, **80** and **81** where possible, so further protecting-group chemistry was put on hold while the main focus on innovative chemistry (S_NAr reactions) was examined.

4.2.2. Routes (b), (c): halofuroic esters; route (e): dihalofuran

These routes were quickly evaluated to be less promising, so are covered briefly before the chosen halofurfural route is discussed in depth. Dibromofuran **85** is classically synthesised from stable bromofuroic acid by decarboxylation and then bromination,[104,139] but has more conveniently been synthesised by a recent direct bromination of furan with Br_2 – a reaction considered impossible for almost a century until it was retried in different solvents which prevent acidic polymerisations;[118,119] I synthesised **85** on gram-scale by this method. Its use in S_NAr is almost unattested; only one report (patent) was found in CAS, reacting it with a carboxylic acid in warm acetone using triethylamine base apparently giving 60% of mono-reaction product, but this report is perhaps doubtful or speculative.[140] Having also esterified 2-bromo-5-furoic acid to **86** in similarly large scale, I then chose to use the well-attested S_NAr conditions of Morris *et al*.[109] to test these furans' reactions with a range of potassium alkoxides including the benzylated sugar which was already giving success with 5-bromofurfural (see section 4.2.3). Reactions are summarised in Figure II.25.

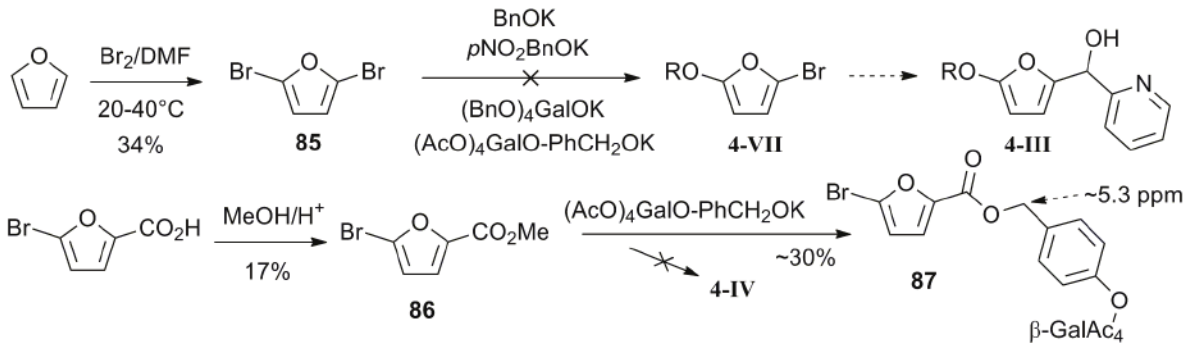


Figure II.29 – Initial trials of S_NAr on dibromofuran and methyl bromofuroate were unsuccessful and abandoned

The reaction of **86** appeared promising, as new furan species could be identified in the crude NMR by the shift of the H3/H4 doublets. Exhaustive separation of the products recovered the various sugar-bearing species almost quantitatively, however the major product was the transesterified **87**; comparable amounts of deacetylation products were also generated and no desired product **4-IV** was seen. Crude NMRs for the dihalofuran reactions were distinctly unpromising: no new furan H3-H4

doublets (desymmetrisation of the molecule after substitution) were visible, so I stopped this route, although it may be that simply heating the reaction could solve the reactivity problem.

These failures argue for changes of conditions; in both cases, adding an activator such as a silver salt may assist halogen reactivity although corresponding losses in the sugar species by oxidation may be expected. The S_NAr on **86** may also need to be favoured by a change of solvent and alkoxide metal ion. However, as the halofurfural route was already giving products **4**-**III**, I considered that there was no need to pursue these reactions further just yet, although **4**-**VII** remains a highly attractive target for future explorations because of the metallation/Petasis route that it allows.

4.2.3. Route (d): halofurfural

In contrast to the frustrations of the previous reactions, a trial of glycoside S_NAr on bromofurfural immediately gave the desired product, with yield depending on the base (KH: 60%; NaH: <10%; Figure II.29).

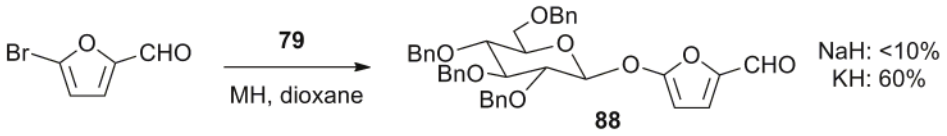


Figure II.30 – Successful first trials of S_NAr giving a furfural β-glycoside

Tetrabenzylglucose **79** reacted to give a mixture of at least 12 different products by TLC, but separation of the major product, protected furfural β-glucoside **88**, was relatively easy, the product was exhaustively confirmed, and the yield was sufficient to carry over. Adjusting the stoichiometry of the reactants still gave the desired product though it was shown that excess electrophile was preferable; in a parallel run with *p*NO_2BnOH under identical conditions, no nitrobenzyl product was observed. A first trial with tetraacetylgalactose **80** gave **89** albeit in low yield (10%); changing the reaction protocol gave varying yields of a species which I assigned as dimer structure **90**, which would indicate that the 2-furfural-5-alcoholate is also an excellent leaving-group (probably, rearrangements ensue). Returning to addition of ROK into the electrophile solution but performing this slowly also gave mostly the same species, which would indicate that **89** is more reactive than the bromofurfural (though if my structural assignment is correct, then reaction proceeds by a different mechanism: S_N2 not S_NAr); anchimeric assistance by the 2-acetyl group may activate the galactoside towards this substitution, explaining why dimerisation was not seen for the benzyl derivative. Following eactly the latter conditions but with 4-MeO-PhOK gave instead a 46% yield of the substituted furfural **91** (no dimerisation is possible), as an oil which degraded in the cold (very electron-rich). Consistently with the observation for *p*-NO_2BnOK, **81** did not give the desired reaction. Intern Nils Aronssohn had examined the use of 5-nitrofurfural with **79**, and concluded that it gave vastly reduced yield of **88** relative to 5-bromofurfural. These reactions are summarised in Figure II.30.

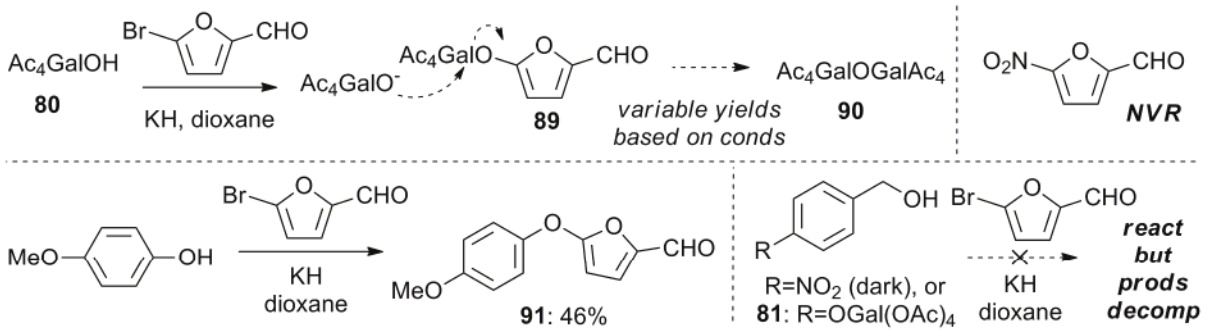


Figure II.31 – Subsequent S_NAr trials beginning examination of substrate scope

I found it surprising that the glycoside and the phenoxide would react, but not the benzoxides. Morris *et al*. had however performed a competition experiment showing the greatly enhanced reactivity of the lactol relative to a comparable alcohol and attributed this to the electronic repulsion β-effect,[109] and I considered that the phenoxide might be rendered similarly more nucleophilic in these conditions by the aryl π-system (hence being reactive). Therefore it was possible that the benzoxides might simply require more forcing conditions to do react (heat, silver salt, crown ether, etc). However, I was still concerned that the benzoxides had actually reacted but then either (1) been destroyed during workup, or (2) formed ether dimers as I had assigned in **90**.

Exhaustive purification and checking of the nitrobenzyl reaction confirmed that dimerisation was not observed; only one possible addition product could be identified on the crude NMR taken after SiO_2 filtration, and this was present at <1%. Its spectrum suggested a ring-opening rearrangement product as the aldehyde peak was absent but there was 1:1 integral correlation between a moiety bearing 2 alkenic protons at ~6.4 ppm with ^3J splitting (eg. ring-opened, oxidised acid form) and a nitrophenyl moiety, though the only peak I could assign as the putative nitrobenzylic CH_2 seemed too far upfield (3.02 ppm) to allow assignment as the ester (est >5.20 ppm) or indeed anything else that was likely.

I also reexamined the crude NMR of the reaction with **81**, but here I found no recovery of the starting benzyl alcohol (CH_2 at 4.63 ppm), with instead a quantitative shift of its CH_2 group *downfield* to 5.00 ppm, plus once more the 2 alkenic protons at ~6.4 ppm with ^3J splitting. This seems to further indicate a furan-ring-opened crotyl-type ester such as **4-VIIIb**. The precise structure of the original furfural moiety eludes me however (though no aldehyde is present); furan ring-opening mechanisms are known to be ferociously complex so this is perhaps not surprising[141,142]. Note that this shift is fairly close to that seen for what I assigned as the furoate **87** (5.27 ppm; Figure II.28) but the global peak patterns cannot be superimposed, which confirms that in this reaction, the bromofurfural was not simply tandem oxidised-esterified to give **87**.

Taken together, this evidence suggests that the benzoxides had indeed reacted by S_NAr, and then that the ring-opening had been provoked either during reaction or workup, eg. by ring-opening-hydrolysis. Further investigations will have to wait till May 2013, but as a first step I planned to install benzyl spacers with no α-hydrogens, eg. **4-IX**: these may give more hydrolysis-resistant products **4-X** (steric

repulsion of incoming water, more basic benzoxide), and also cannot be themselves deprotonated, if this is a concern for hydrogen shifts in the strongly basic reaction medium (Figure II.31).

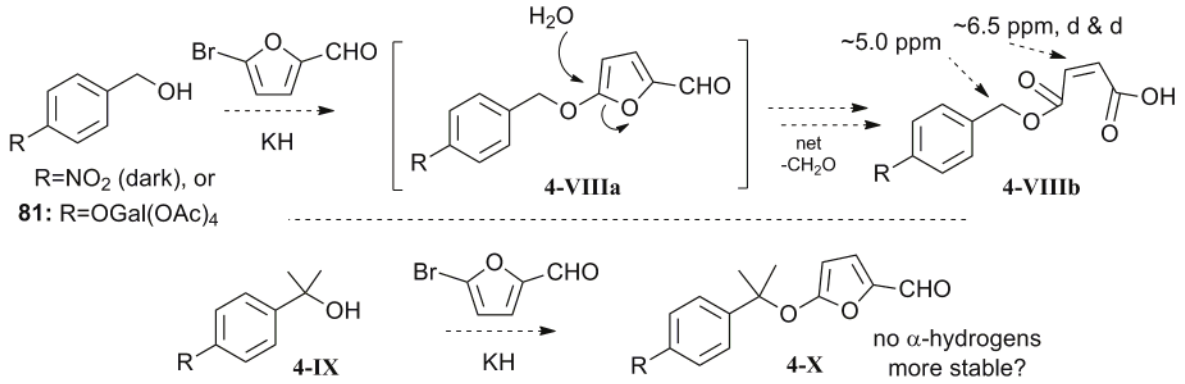


Figure II.32 – hypothesised degradation mechanism for benzoxide reactions, and new target PHBA analogue 4-IX

With experimentation time almost finished, **88** was then subjected to initial trials of reaction with metallated 2-pyridine. I ran the reaction with either stoichiometric or excess 2-pyridyllithium; when excess was used, I quenched the excess lithiate by addition of benzaldehyde before the next step. As fast oxidation of the furylpyridyl carbinol to ketone **4-XV** was feared[111], I quenched the lithiated alkoxide intermediate **4-XI** with tosyl chloride to give tosylate **4-XII**. I then attempted both isolation of **4-XII** and its *in situ* further reaction to **4-XIII** with amines such as morpholine and tacn, though no amine substitution products were isolated (the solvent conditions disfavour this reaction): Figure II.32.

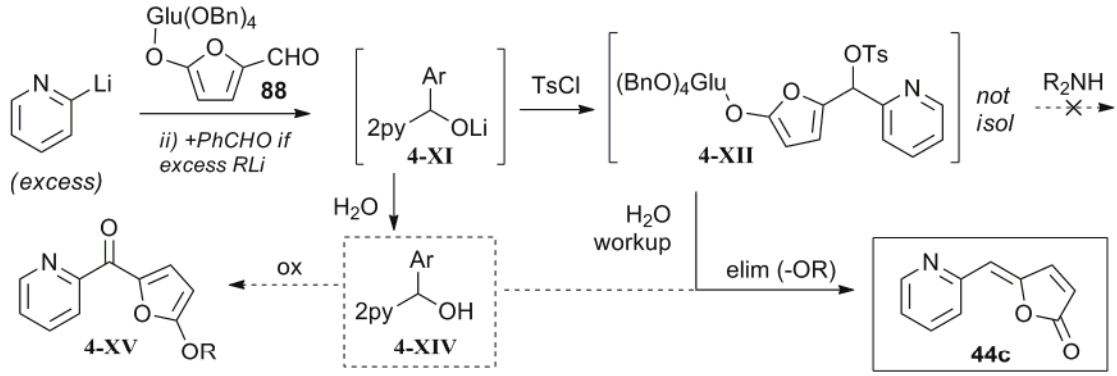


Figure II.33 – Multistep transform towards desired species 4-XIII

The results of these small-scale trials were promising for the currently upcoming experiments. Crude NMR showed complete disappearance of the aldehyde functionality of **88**. *After* workup and column, **79** was assigned as the major sugar-bearing product, and the major furan-bearing product was isolated and confirmed[143] as the pyridyl-bearing furan elimination product **44c**, indicating that at least reaction to the key intermediate **4-XI** had occurred. Reexamining the crude NMR, no evidence of either this elimination product or **79** was seen, indicating that without water and silica, either **4-XIV** or the more likely tosylate **4-XII** is stable in solution. Then PhCHO was used in a parallel reaction; as expected, 2-pyCH(OH)Ph (**4-XIV**, Ar=Ph) was isolated after column (since no elimination possible) which reinforces the eliminative interpretation. Oxidation of this species was not observed, suggesting that

perhaps isolation of the alcohol **4-XIV** (Ar=alkoxyfuryl) is possible. Returning to the initial crude NMR for **88** with tosylation, the spectral features for **4-XII**/**4-XIV** (α-(alkoxyfuryl)-α-2-pyridyl-CH-OR) could be identified, as a more satisfying indication of their solution stability before the workup procedures applied.

Therefore there is hope that further reactions of **4-XII**/**4-XIV** may give the desired tertiary amines **4-XIII**, if intermediate workup is avoided, eg. by evaporation of the solvent and replacement with an S_N2-favouring medium such as dry DMF or perhaps dry EtOH. One initial trial with a 1:1 DMF:THF mix (DMF solution of the amine added to THF solution of the tosylate) did seem however to indicate (base-induced?) elimination in the crude, showing **79** as the only sugar-bearing species. If the amine can be synthesised, it will hopefully prove (substantially) more stable to spontaneous elimination than the alcohol/tosylate, especially after complexation. However careful considerations of the S_N2 conditions will be needed to avoid competitive elimination, and it is still possible that the substitution onto a secondary amine may simply prove impossible.

5. Conclusions on the Alkoxy-Furyl Pyridyl Syntheses

5.1.1. General remarks

These synthetic efforts have delivered two routes giving the first reported glycosyl furans. The cycloisomerisation routes are of less practical interest to the present work than the S_NAr route, as furan functionalisation seems much easier by S_NAr, and since the picolyl keto/enols were not cyclised in the desired manner. The rather recalcitrant alkoxyfurans have shown considerable sensitivity, which may need to be improved eg. by steric blocking of the alkoxy α-position in a stacked spacer design, before they can be applied in probes. Their exceptional elimination tendency has been strongly indicated, which gives hope that their high-energy design will overcome the chelation stabilisation seen in the group's prior work. The 2-picolyl group has also been strongly indicated in the instability of the furan species, so if the 2-(5-alkoxy)furylpicolyl group proves simply too prone to elimination, then other coordinating heterocycles with less electron-attracting or delocalising character (oxazole, isoxazole, pyridazine) could present a solution which could add stability to the alkoxyfuran species. In fact, non-pyridine heterocycles like pyridazine are anyway a requirement for use in off-ON bispidine-based designs (cf. **38d**, Figure II.2): the flagpole hydrogen clash when a sixth coordinating arm bearing pyridine is used leads to an intermediate spin-state (ie. →only a relative change probe).

5.1.2. S_NAr route

Route (d) was convenient, fast and high-yielding for the installation of the glycosides directly on furfural giving unattested, functionalised-glycosylfurans **4-VI**. The stacked substrate-benzoxides were also installed, but faced especially severe hydrolytic problems, so more hindered benzoxides **4-IX** will be examined during upcoming work. This access to alkoxyfurfurals **4-VI** makes S_NAr a robust replacement for the cycloisomerisation/α-functionalisation route. The alkoxyfurfurals will be synthesised in large scale and carried through, attempting isolation of **4-XIV**, as well as carrying the pyridyl-bearing arm directly through to potentially more-stable **4-XIII**.

Route (e) to Petasis reaction precursor **4-VII** was not successful in initial trials, but would be convenient: especially if **4-XIV** is too unstable for reaction to **4-XIII**, then the Petasis reaction could provide an ideal orthogonal synthesis; so more activated conditions will be reexamined.

For glycosidase targeting, if dimerisation of the tetraacetylated glycoside (direct attachment) continues to be problematic, then a new mildly-removable protecting group will be needed; such a group may be needed anyway if deprotection of the acetates even under mild conditions also destroys the furan. For Petasis reaction, a protecting group stable to alkyllithiums will be needed; if pivaloyl and trialkylsilyl groups continue to be problematic, new groups (or combinations of protecting groups, eg. an acetonide and two mono-protecting groups) will have to be investigated.

5.1.3. Future strategies

Alternatively, a nitrofuran (or nitro-pyrrole/thiophene/imidazole, etc) is an obvious target for nitroreductase-triggered immolation by this architecture, since they cannot spontaneously eliminate or hydrolyse before reduction of the nitro to the amine, so are likely to be easier to install as a proof-of-concept. However this approach (like most NTR-targeting strategies) does not reflect what can be achieved with hydrolase-sensing spacers and therefore holds fewer rewards.

A second approach is to examine an easier elimination reaction altogether, allowing the use of lower-instability species. This Part II.A has addressed the elimination of the bond from the pivotal methine to the macrocyclic amine; and the group has already explored the elimination of the bond between the coordinative heterocycle and the pivotal methine[29]. The second part to this chapter will now explore the elimination of the coordinative bond itself – the lowest-energy member of the three critical bonds in the proagent design.

Part II.B: Cyclic Carbamate Tautomeric Shift Systems

6. Design of Cyclic Carbamate Tautomeric Spacers

Part A described research towards a new elimination spacer, whose powerful driving force of elimination was intended to allow elimination of the difficult spacer-macrocycle covalent bond, by overcoming the chelation stabilisation effects which were seen to block the use of quinone methides in that application. However, such a high eliminative driving force will likely incur not only thermodynamic metastability, but kinetic instability too, of both its synthetic intermediates and the final product: at least, such was amply seen for the sensitive furan species. Part B therefore followed a different strategy towards spin-switching spacers, based instead on the elimination of the coordinative bond *only*. This approach of Prof. Hasserodt's is based on the hope that this elimination would require a lower energetic driving force due to both the lower energy of the coordinative bond, and to the possibility of attaining the correct elimination geometry, and therefore be easier to implement.

A tautomerisation-based coordination/decoordination approach was designed, based on trapping then releasing a strongly basic moiety, as per complex **92**. This putative off-state proagent features a nitrogen which is trapped in the imine form by the presence of the distal enzyme substrate, and this nitrogen is predisposed to coordination by the favourable chelate size. Hopefully this should ensure the nitrogen is reliably coordinated as shown in **92a**, and therefore that the proagent is reliably low-spin (off). After **distal enzymatic cleavage**, **93b** is no longer trapped as the imine, and the more **basic** the moiety is, the more chance it will protonate giving **93c**. Once protonated, the **tautomerisation** energy gain of returning to the favoured, but *necessarily uncoordinated* carbamate tautomer (**93**) relative to the unfavoured, *coordinated* carbimidol (as in **93c**), provides the main driving force for elimination of the coordinative bond to give the ON state **93**: hence the design of an enzyme-triggered spin-switch off-ON MRI contrast agent (Figure II.33).

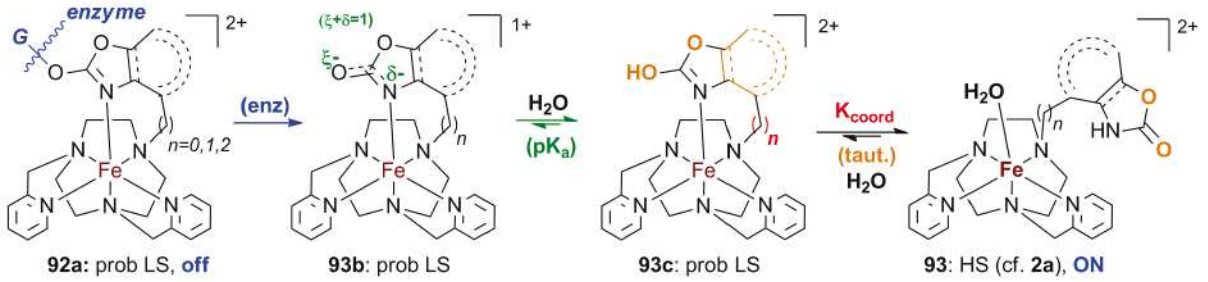


Figure II.34 – General design for cyclic carbamidate→carbamate enzyme-triggered coordinative switch. The three crucial equilibrium considerations are indicated in colour: basicity (green), tautomerisation (orange), chelation (red). The critical overall effect of the enzyme-activated off→ON spin-switching is indicated in blue.

The cyclic system ensures that the oxygens marked in orange *cannot* coordinate to the iron as their steric access is impossible. Preventing oxygen coordination was considered vital, both *a priori* and because of observations with a somewhat related iron complex investigated by Fayçal Touti. In that complex, the normally favoured immolation of a carbamic acid was prevented by oxygen chelation to

the metal, giving a complex that was *not* fully high-spin (ie. the complex was not really ON), *and* where water access to the metal was blocked (so even if the complex had been high-spin, relaxivity transfer would have been inefficient). Note that this is promising for the present design, as it indicates that spin quenching is rather easy (chelated oxygen almost suffices), therefore **92** with its chelatable *imine nitrogen* stands a good chance of being a reliably low-spin, off-state proagent (cf. **2d** or **38d**).

With the present cyclic *carbamate* tautomers, **93** is therefore *necessarily uncoordinated*: full deprotonation (returning to form **93b**) is a prerequisite of *carbamate* coordination, and full deprotonation is disfavoured by designing **high basicity**. The hope is that the protonated *carbimidol* form **93c** is more disfavoured by **tautomerisation** than it can be stabilised by chelation. However, preorganisation is known strongly to perturb molecular reactivity both kinetically (see essentially all of Part I) and thermodynamically. In the present setting, the stronger the **chelation** preorganisation, the more it would counteract the basicity and tautomerisation tendencies of the carbamates and so shift the equilibrium balance in favour of the undesired forms **93b**-**93c**. Therefore we had identified in total three critical factors (basicity, tautomerisation and chelation) which ought to determine the equilibrium balance between forms of **93**, and therefore the obtainable signal, *as long as* enzymatic activation of **92a** can be achieved. These three factors will be adressed by specific design features.

Chelation is perhaps the easiest effect to predict. Chelation could be destabilised both thermodynamically and, especially, kinetically by eg. designing a hydrophilic coordinating spacer which would not achieve optimal chelation sterics. The sterics could easiest be affected by changing the chelate cycle size, hence the general design for **92** allowed for n=0, 1 or 2 methylene units in the chelate ring, ie. **5-, 6- and 7-membered chelate scaffolds**; annelation of the coordinative ring (eg. oxazolone→benzoxazolone) also affects the chelation rigidity and hence kinetics.

pK_a: A **range of cyclic carbamates** were chosen to cover a range of highly basic pK_a values (Figure II.34). By sweeping this chemical space, these targets ought to give a corresponding range of basicities once in the complex: though note that the values given are *in DMSO*[144,145], and anyway will be strongly modulated by the presence of the iron. The range is important, since the optimal design will not simply be the *most basic* of the carbamates: proagent stability must also be ensured, and stability will probably be affected by pK_a, both of the carbamyl form as well as of the trapped imidate form. As these are unknown quantities, exploring the full range of structures was more desirable than targeting the most basic, even considering only the pK_a. Note that the benzoxazolone scaffold in particular provides a likely scaffold for rational tuning of pK_a (as well as ligand field strength if this is required) by changing the back-end groups G: eg. G=methoxy might show higher NH-pK_a than 'plain' benzoxazolone, or electron-withdrawing/delocalisation-extending groups might give lower pK_a.

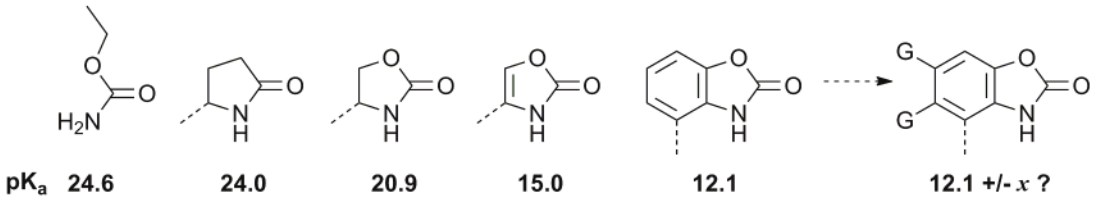


Figure II.35 – pK_a values for carbamates and related species, attachment point for chelation is indicated

Tautomerisation: To a greater extent than pK_a, the tautomerisation energy would have to balance stability with favouring elimination (a problem which has been seen extensively in the work on alkoxyfurans). Literature precedents were totally lacking to predict how this would best be designed for aqueous applications, so, a range of species would have to be experimentally verified. **Oxazolones, benzoxazolones and oxazolidinones** were retained: the imidate ester system provides such a highly energetically favourable tautomerisation that acyclic imidate esters are solvolysed instantaneously in methanol; and even some *endo*cyclic counterparts have a spontaneous hydrolysis halflife on the order of hours,[146] so greater stability had to be obtained. *Cyclic carbimidate* esters would probably have higher stability in water than simple imidate esters, with the cyclic design especially favouring stability, but this would have to be confirmed experimentally. More difficult to predict were the actual energies of tautomerisation; it seems no experimental determinations of the tautomer equilibria have been performed, although qualitatively the carbamate forms (cf **93**) are known to vastly predominate over the carbimidols (cf **93c**). The sole reference that was found for a numerical estimation of the tautomerisation energy in water concerned oxazolone itself, and concludes "[oxazolone] should be the dominant tautomer in all media because of the large negative value of ΔH_f in comparison with the hydroxy form [oxazolone carbimidol]."[147] The difference in ΔH_f that was calculated was fully 97 $kJ \cdot mol^{-1}$ – an enormous value that indeed ought to compensate the energy loss of decoordination (as we cannot assume a 7-coordinate intermediate with incoming water, which would lower the activation energy); this coordination bond is perhaps worth about 25 $kJ \cdot mol^{-1}$ for a *non-chelated* ligand, and is undoubtedly higher for the rigidified species, but precise estimates were not easy to make: less than twice as high, would certainly seem reasonable.[148]

Therefore the final targets retained were 5-, 6- and 7-membered-chelating oxazolones and oxazolidinones, plus (synthetic restriction) 6-membered benzoxazolones. The next chapter describes the practical research into these scaffolds, their functionalisation with enzyme substrates, and progress towards potential magnetogenic complexes **92**. **Note: Fearnley 2004**[149] provides an excellent reference for the synthesis, reactions and applications of oxazolones as well as touching on oxazolidinones.

7. Syntheses and Applications of Cyclic Carbamate Spacers

The reactions in this schema broadly follow a parallel structure for each scaffold, with four distinct phases: (1) scaffold synthesis, (2) O-alkylation, (3) electrophile formation, (4) grafting onto the ligand, towards complexes of type **92** (Figure II.35). Each scaffold's particular challenges are addressed separately within each phase.

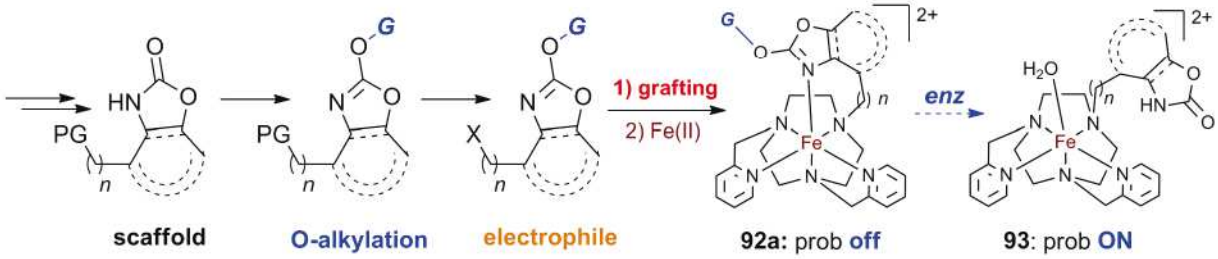


Figure II.36 – Four phases towards the synthesis of candidate bio-activated off→ON contrast design 92→93

7.1. Heterocycle scaffold syntheses

7.1.1. Oxazolone 5-membered scaffolds

Note that oxazolones with 4,5-substitution by nonfunctional stabilising groups such as phenyl rings[150] as well as 4,5-unsubstituted species[151,152] are well known. However their syntheses are comparatively simpler (especially by cycloadditions or photochemical rearrangements), and their subsequent stabilities higher, than those of *functionally*-substituted species (eg. esters), especially where the functional substituent must be at the 4-position.[149]

The oxazolones required for the 5-membered-chelating scaffold are known[149,153,154], and were synthesised following the route of Okonya *et al.*[155]: (1) by the method of Weinstock *et al.*[156], *n*-hexyl Grignard reagent was formed *in situ* and selectively mono-added to diethyl oxalate forming known hexyl α-ketoester **94**, and 2-phenylethyl α-ketoester **95** was bought; (2) bromination by the method of King *et al.*[157] afforded good yields of the known[155] β-bromo-α-ketoesters **96** and **97** in >80% yield; then AgOTf-induced condensation with methyl carbamate[158] affords known[155] 5-alkyl-4-ester-oxazolones **98** and **99** in >60% yield (Figure II.36). Use of the cheaper $AgNO_3$ to replace AgOTf unfortunately did not give any isolated desired cyclisation product. An attempt to follow this reaction path with benzylmagnesium bromide, towards R=Ph oxazolones, failed to produce the desired addition product, in accordance with a literature observation[156]. The β-bromo-α-ketoester with R=H was commercially available and was also condensed to give the known R=H, ester-functionalised oxazolone[155].

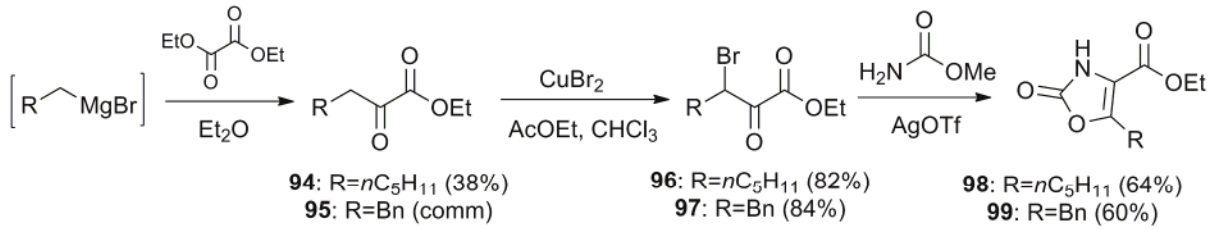


Figure II.37 - Route to simple oxazolone 5-member-chelate scaffolds 98 and 99

These esters **98** and **99** were desired for initial *O*-alkylation trials, as a potentially less-sensitive testbed for the alkylation reaction. Ultimately however, reducing the ester group after *O*-alkylation might pose protecting-group problems for the substrates used (eg. acetates on sugars). Therefore I immediately prepared alcohols **100** and (known) **101** using DIBAL, giving ~80% yield,[155] and protected them as TBS ethers **102** and **103** with TBSCl/imidazole in DMF in good yield (though initial attempts to form **102** in less forcing conditions led to reaction failure, whereas **103** was obtained both more easily and in better yield; Figure II.37). I also examined THP protection via $POCl_3$-catalysed reaction of DHP and **101** with stoichiometric pyridine to absorb acid and try to prevent decomposition, however no products were isolated from the reaction.

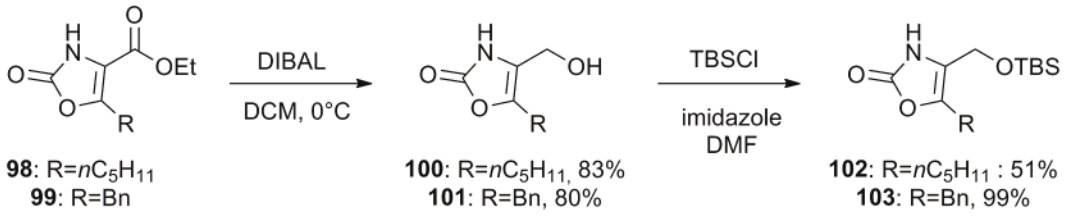


Figure II.38 – route to TBS ethers 102 and 103

Finally, a method of Dakin *et al.*[159] appeared to offer a fast route towards 4-ω-siloxylalkyl oxazolones, from cyclisation of the acyclic hydroxyketones. This route was initially examined with dihydroxyacetone (R=H), however the initial alcohol protection step was more complex than expected, which was attributed to the possibility of hydroxyketone dimerisation.[160-162] Eventually, six hours' refluxing in acetone was shown to crack the dimer enabling subsequent alcohol protection, but dimerisation was again favoured by concentration. This will be examined in more detail for the 6- and 7-membered chelate scaffolds; but since the alkyl-substituted oxazolones **102** and **103** were accessible in good yields, the dihydroxyacetone route was abandoned.

7.1.2. Oxazolone electrophilic derivatisation, part 1

In Part A, the pyridine-based coordinating moiety used in the work on furan spacers was known, from complex **2d**, to give a low-spin proagent. Yet for this work in Part B, no suitable model complexes have been reported which incorporate cyclic carbamate/carbimidates which can coordinate to a central Fe(II). Therefore we also wished to examine the putative high-spin complexes **93** directly: if their carbamates actually did coordinate (eg. **93b**/**93c**, instead of detaching as was planned) and the resultant species were low-spin or only weakly relaxing, then despite any enzyme activity, no good

relaxivity enhancement would be observed so this line of research would be a waste of time. Therefore I examined converting alcohols **100** and **101** into electrophiles for S_N2 onto the macrocyclic penta-coordinate target ligands, for a direct route to targets such as **93**.

No direct molecular precedents were found for halomethyl oxazolones or similarly leaving-group derivatised oxazolones, and even if several exist for eg. oxazoles[154], these have *clearly* different reactivity as they are not N-alkylated in the same way. Initial standard reactions gave no apparent success: (a) Iodination of **100** with I_2/PPh_3/imidazole in DCM as per the method of Kopach *et al.*[163] did not show any identifiable products by crude NMR, although TLC indicated the formation of two faster-running species and no remaining starting material; (b) similar chlorination of **100** as per Milbank *et al.*[164] with PPh_3/CCl_4 in DCM (similar to the oxazole procedure[154] which gave an easily isolable bromomethyloxazole) gave no identifiable species by DIMS, but though crude NMR now indicated conversion, TLC did not reveal any desired faster-running spots.

(c) Chlorination of **101** with $SOCl_2$/pyridine in DCM showed no starting material remaining by TLC, and crude NMR seemed very promising; ^{13}C-NMR notably showed an upfield shift in the 4-exocyclic carbon, from 55.9 to 34.1 ppm, consistent with conversion of the alcohol to chloride **104**. However TLC did not reveal any well-mobile spots and after chromatography the NMR of the isolated spot was uninterpretable. (d) Mesylation of **100** with MsCl / Et_3N in DCM as per the procedure of Al-Harrasi *et al.*[165] gave two closely-running spots on TLC but NMR indicated no mesylated product, and although DIMS(-) showed a possible (M-H) peak, no other interpretable peaks were present. (e) Tosylation of **101** with TsCl/*N*-methylimidazole in DCM gave total conversion to a single faster-running spot but chromatography failed to isolate any interpretable tosylate species (Figure II.38). Extensive reaction of electrophile-forming species with the ring nitrogen, possibly accompanied by ring opening degradations or intermolecular alcohol/ring nitrogen reactions of the desired product electrophiles, were therefore feared.

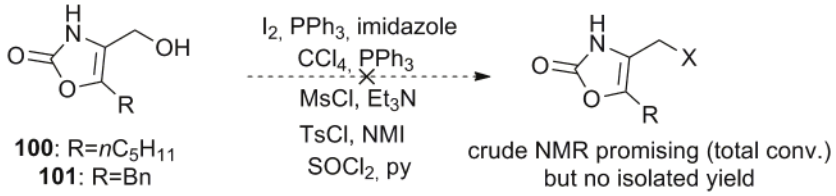


Figure II.39 – initial attempts to isolate grafting-ready electrophiles met with extensive failure

The key was found when re-examination of the NMR of mesylation and chlorination reactions of **100** revealed that the major crude product of both reactions was *identical*; fine examination of the tosylation of *O*-alkylated **142** (section II.7.2.2) then revealed that rather than the expected tosylated product, the chloride **144** was instead stoichiometrically recovered. It seems that the *non-alkylated* electrophiles are so reactive that not only is the *in situ* substitution by the initially expelled chloride ion quantitative, but these chlorides are themselves reactive enough that **104** stoichiometrically quaternised the (deficit of) triethylamine in its reaction, even at 0°C in DCM. Suggesting similar

reactivity, during the $SOCl_2$ chlorination, concentration to a crude oil appeared to give intermolecular alkylation even in the absence of base. These electrophilic species degrade on silica, during concentration, or with nucleophiles in general; I suspect that the free NH ensures greater reactivity than their *O*-alkylated carbimidate forms, given the comparative ease of isolating the chlorinated derivative of **144**, and the observable degradation of the electrophilic species upon concentration.

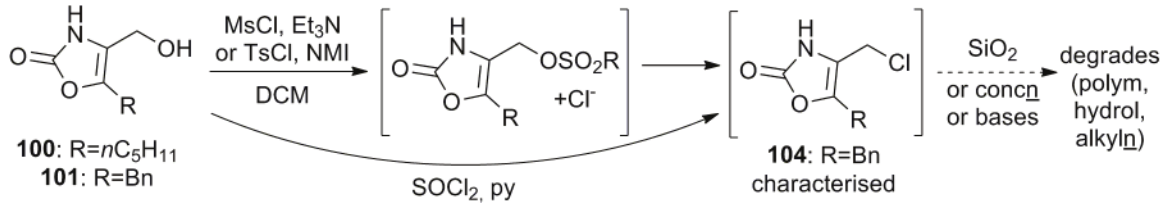


Figure II.40 – interpretation of the failure to isolate desired electrophilic species

However, by the time that this mechanism had been identified, access to the more tractable *O*-alkylated targets had been secured, so they could be used for all subsequent work. Therefore no resynthesis of electrophiles like **104** was made, although this could be performed eg. by reaction with a stoichiometric amount of MsCl in DCM with NMI or pyridine base in the cold (eg. 30 min at 0°C), then injection of the crude solution into a solution of the desired ligand in any solvent as or more polar than DCM; the $SOCl_2$ reaction likewise could easily be used as long as concentration was avoided.

7.1.3. Oxazolone 6- and 7-membered scaffolds

The synthesis of the 6- and 7-membered chelating oxazolones is not favoured by the diethyl oxalate route (section II.7.1.1): the activation provided by the α-ester in the 5-chelating synthesis is no longer present, and bromination of the ketone would anyway risk predominantly (6-membered) or statistically (7-membered) forming the undesired ester-proximal bromide. Dakin *et al.*[159] reported a synthesis for the 7-membered species, which I followed. TBDPS protection of ω-unsaturated alcohols with TBDPSCl/imidazole in DMF gave **105** and **106** in good yields (>90%). Oxidation to the primary hydroxyketone is then required. Dakin *et al.*[159] perform a high-yielding two-step procedure with OsO_4 then peroxotungstophosphate oxidations (90% overall). However to avoid the use of these reagents I followed a report by Bonini *et al.*[166] who perform $KMnO_4$ oxidation of similar terminal alkenes bearing β- and γ-silyloxy groups but with longer-chain substrates, obtaining the ketols in >70% yield. I obtained yields of **107** and **108** of 50-65%, which matched the data of Dakin *et al.*[159] However, the final one-pot cyclisation protocol of Dakin *et al.*[159], of phosgenation followed by addition of ammonia then H_2SO_4 acidification (yield reported 85%) was troublesome, and appeared to be complicated by extensive yields of intractable reaction products from dimerised ketol starting materials (recall dimerisation in dihydroxyacetone; section II.7.1.1). After multiple runs, desired oxazolones **109** and **110** were finally obtained at yields of only 7-12% brsm (Figure II.40), but on a scale suitable for further trials. It may be that the dimerisation highly depends on the conditions used for hydroxyketone isolation, especially if heated when concentrated.

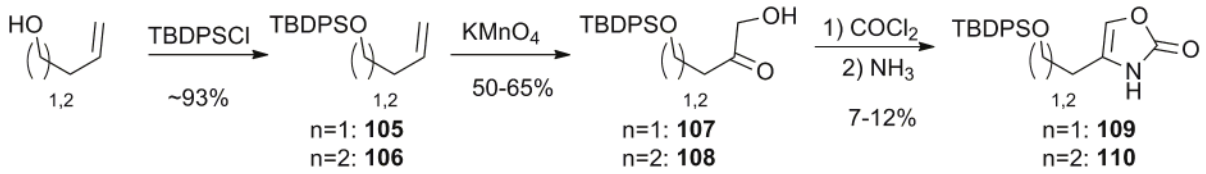


Figure II.41 – route to 6- and 7-membered chelate oxazolones 109 and 110

7.1.4. Benzoxazolone scaffolds

Surprisingly, appropriate 2-amino-3-carboxyphenol starting materials were prohibitively expensive, so attempts to synthesise some from cheaper species were made. In general these proved to be relatively resistant targets, and the benzoxazolones posed challenges for purification and isolation that perhaps should have been anticipated.[167] The cheapest commercial starting material was methoxy acid **111**; this was esterified to **112** with $SOCl_2$/MeOH in up to 80% yield. The methoxy group was cleaved to give the known compound[168] **113** using BBr_3/DCM[169] though purification and even characterisation of the products was problematic, and yields were around 30%, accompanied by variable amounts of ester cleavage. Interestingly, this compound is light- and air-sensitive,[168] which probably reduced yields further. Its corresponding, known[168], base-sensitive[170] benzoxazolone **114**, was then formed by reaction with phosgene as per the procedure of Fielder *et al*.[168] though in poorer yields (~20%; Figure II.41). Given the small amount of benzoxazolone thus synthesised, this route towards desired **115**, was paused while I examined other methods, and perfected the O-alkylation protocol.

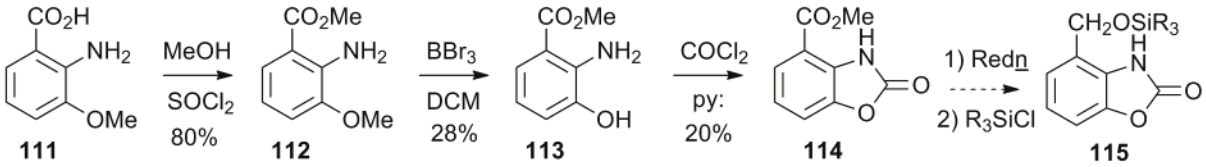


Figure II.42 – Initial synthesis of unsubstituted 6-membered benzoxazolone scaffold 114

While the demethylation reaction yield was anticipated to be poor, the phosgenation was surprising. To check the phosgenation protocol, I checked large-scale syntheses of benzoxazolone (**116**) from similar reactions of 2-aminophenol with phosgene. The isolated yields likewise were poor; a more specific reference shows that they depend heavily on the exact procedure used,[167] so they can probably be increased in the future. Still, these could provide a second source of **115** if the route via demethylation of **112** cannot be improved: directed *o*-lithiation of N-acylated[171] **117** can probably proceed without significant problems, and in good yield.

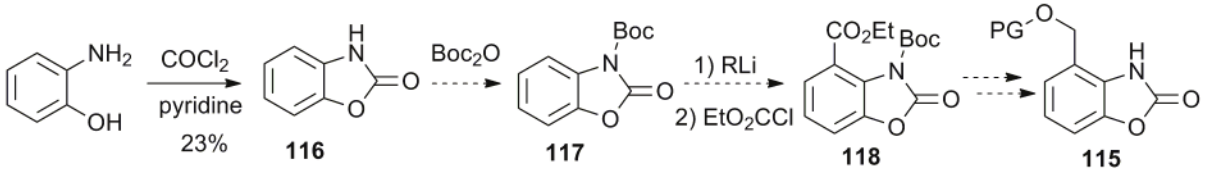


Figure II.43 – Large-scale route towards 115

For pK_a modulation, I also desired substituted benzoxazolones with inductive groups G (Figure II.34), such as **119**. Again, no appropriate *o*-aminophenols were directly available, but bis-methoxy **120** was surprisingly cheap. The Boyland-Sims[172-174] oxidation using $K_2S_2O_8$ seemed to be a likely direct method, although it is low-yielding and requires large-scale runs (literature reports cite only up to 10% usually for >10 g scale). The Boyland-Sims procedure reported by Toth *et al.*[175] was trialled but proved to give an intractably inseparable mixture of unknown products, with no isolation of desired aminophenol **121**. Therefore the current favoured alternative is aniline acylation, then catalytic $Pd(OAc)_2$-directed acetylation as per Wang *et al.*[176] (5 mol% Pd, 80% yield for a similar substrate), giving **122**. Total cleavage of the protecting groups, then acid-catalysed esterification as for **112**, followed by phosgenation, should provide a longer but much more amenable (and hopefully, generalisable) route to pK_a-modulated benzoxazolone derivatives of **119**.

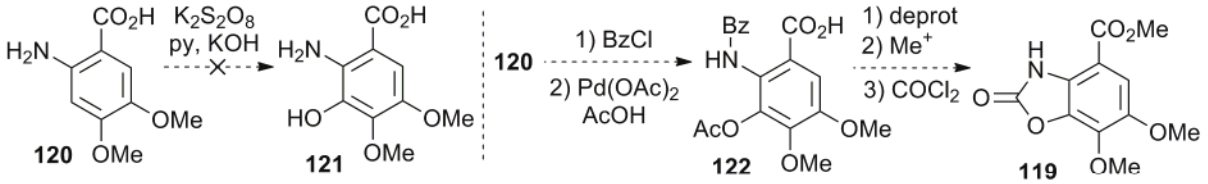


Figure II.44 – The current synthesis of pK_a-modulated benzoxazolone 119 is planned via 122

7.1.5. Oxazolidinone 5-membered scaffolds

Note that the oxazolidinones are chiral, although this is unlikely to be important in the current context. The 5-membered oxazolidinone ester **123** was accessed by a convenient reported chiral one-pot procedure by phosgenation of serine then esterification,[177] giving 52% yield without purifications. Its DIBAL reduction by similar conditions as for oxazolones **98** and **99** did not allow recovery of product **124** from the aqueous quench mixture, so as suggested by the patent of Celanire *et al.*[178], one pot silylation was attempted on the evaporation residue, but no yield of chiral **125** was obtained. However, racemic **124** was isolated in 88% yield by large-scale phosgenation of serinol as per Celanire *et al.*'s original procedure[178]. This was silylated similarly to the protocol for **101**, giving **125** in 72% yield after trituration and then precipitation at -80°C (Figure II.44).

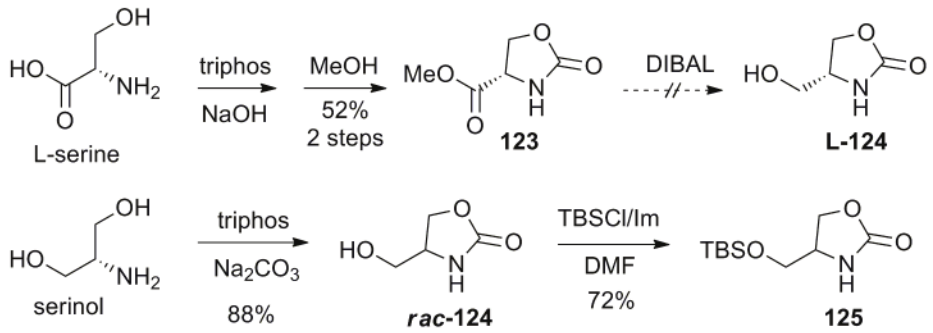


Figure II.45 – Syntheses of oxazolidinone 5-member-chelate precursor 125

Despite the failure to synthesise electrophilic derivatives of the oxazolones cleanly (section II.7.1.2), such species are attested for exactly this oxazolidinone, and following the simple published TsCl

procedure[178] I obtained a 31% yield of the known tosylate **126**, *after* silica column, with the ring-opened addition product **126x** isolated in ~10% yield. The tosylation yield appeared low until searching the literature found a report of the tosylation of the corresponding 6-chelating species, stating "tosylation… was difficult to achieve selectively because the oxazolidinone was also vulnerable to addition of this reagent. Consequently, the yield of the tosylation step was only 40%… We were unable to improve upon this yield or find an effective alternative to circumvent use of tosylate."[179] Another report exploiting this ring-opening states "Tosylation of the ring nitrogen… proved to be the best method for… ring opening."[180] The decomposition of the corresponding oxazolones in identical conditions was unfortunately again confirmed.

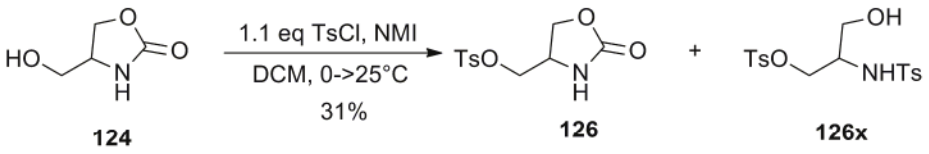


Figure II.46 – tosylation of 124 gave 126 as well as ring-opening byproduct 126x

7.1.6. Oxazolidinone 6- and 7-membered scaffolds

The syntheses were based on aspartic and glutamic acids; despite the basic conditions, there is a strong chance that chirality was mostly retained in the products. The synthetic approach was provided by Hou *et al.*[179], mirroring Lakanen *et al.*[180] (Figure II.46). Bis-esterification and carbamylation gave species **127a/b**, which were to be reduced to the bis-alcohols **128a/b** by $NaBH_4$. An initial effort using the methyl carbamate (R=Me) did not allow product recovery from the salt mix during this reaction, so the benzyl carbamate (R=Bn) was pursued. Reduction then gave excellent recovery of **128**, though one-pot cyclisation/silylation was no longer optimal as the benzyl alcohol liberated during cyclisation to **129a/b** was not removable, and the silylation to **130a/b** did not allow good recovery: 8/24% for compounds **130a/b**, with greater yield of the second perhaps because of the increased solubility of the longer-chain derivative in the organic phase. In the future these two steps, cyclisation and silylation, will be performed separately to increase yield.

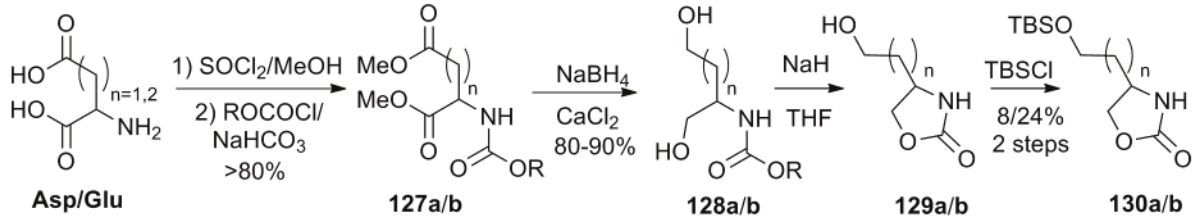


Figure II.47 – Synthesis of 6- and 7-membered chelate oxazolidinones 130

7.2. O-alkylation of carbamates to carbimidates

7.2.1. Problems of direct O-alkylation; initial S_NAr strategy

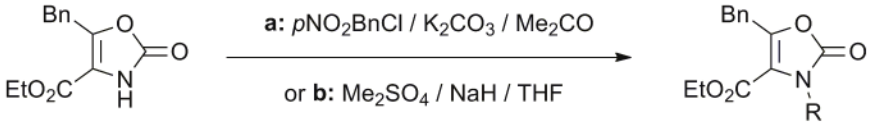
O-alkylating these cyclic carbamates with the desired substrate (initial target: galactose, eg. via tetraacetobromogalactose AcGalBr) presents a difficult regioselectivity problem: differentiating

between the ring nitrogen and exocyclic oxygen, both of which have negative and nucleophilic character: however, the ring nitrogen is the vastly preferred site for alkylation. Only one relevant O-alkylated oxazolidinone was found in the literature, Seki's benzylated intermediate[181]. Only very few O-alkylated benzoxazolones have been reported; unfortunately the most promising-seeming report found in CAS (O-alkylation with an acetochloro sugar giving 58% yield) was presented at a Russian congress without recorded experimental details and not followed by a publication;[182] the typical reactivity published is again on the ring nitrogen[183]. Oxazolones are the scaffold with the most O-alkylated examples in the literature, so their O-alkylation was studied in more detail.

Oxazolone O-alkylation has only been regioselectively achieved with highly oxophilic reagents such as diazomethane[184] and Meerwein salt[185] which are non-generalisable reagents. In general, the nitrogen site is a better and softer nucleophile so it is *exclusively* alkylated in standard basic conditions especially with polar solvents[186] though also in THF[187] (despite a patent claim to the contrary[188]):

> *"oxazolone possesses a significant negative character at the C_2 carbonyl oxygen... Trapping here with a range of electrophiles results in a rapid synthesis of 2-substituted oxazoles. These have mostly been limited to the formation of* ***simple*** *ethers,* ***or substitution reactions via the intermediate chloride*** *[halooxazole]... most of this chemistry has been carried out with [4,5-inertly-]substituted oxazolones [especially 4,5-diphenyloxazolone] or benzoxazolone... though* ***N-alkylation still predominates****."*[149]

N-alkylation by standard approaches was confirmed during my initial reaction trials (eg. Figure II.47) so direct electrophilic substitutions, eg. with AcGalBr, were abandoned.



99

131: R = *p*NO_2Bn
132: R = Me

Figure II.48 – Initial trials of direct S_N2 gave exclusively N-alkylation products (eg. 131 and 132)

The literature on O-alkylated oxazolones shows essentially a single general approach to O-alkylated derivatives: first, conversion to the 2-halooxazole, then S_NAr reaction with a metal alkoxide.[149] Halogens of interest are especially chlorine[189,190], as per the classic work of Gompper *et al.* in the 1950s using $POCl_3$ to chlorinate a variety of species; 4,5-dialkyloxazolones predominate[191], since chlorination may lead to decomposition of less stable species[191,192]. Other halogens require installation by different methods, and not directly from the oxazolone. Direct oxazolone bromination gives decomposition, so Ting *et al.*[193] install bromine via bromide substitution on an oxazolone 2-diazonium derivative, and replace it with a variety of nucleophiles. Perner *et al.*[194] install both bromine and chlorine by BuLi deprotonation of an otherwise inert oxazole; Flegeau *et al.*[195] similarly install iodine using LDA. These methods appear not to have been applied to oxazolidinones.

I swept a range of chlorination conditions with the ester/TBSO/HO-functionalised oxazolones/ oxazolidinones in 15 test reactions. Extensive decomposition was evidenced in all. All oxazolidinones, pentyl-oxazolones, and hydroxy/siloxy species gave intractable tars; only the benzyl oxazolone ester **99** partially resisted the chlorination conditions, giving the ester-bearing chlorooxazole **133** in up to only 30% yield (best conditions[189] after 5 trials, Figure II.48). Yet even disregarding the poor synthetic access, the synthetic utility of this intermediate is doubtful - in a S_NAr trial run with NaOMe/THF, neither desired O-methyl reaction product nor **133** were recoverable. Lastly, either the chlorooxazole or the O-alkylated oxazolone must resist subsequent ester reduction to create the grafting site for macrocycle attachment (cf. Figure II.35): unpromising signs.

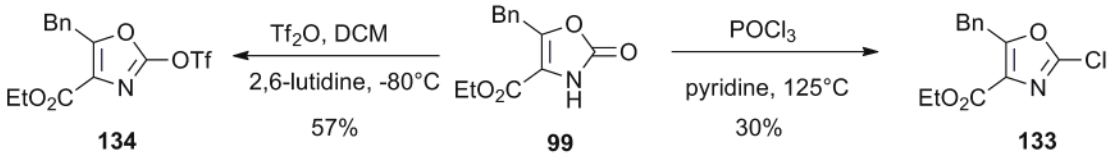


Figure II.49 – Desired S_NAr substrates 133 and 134 were accessed but were not useful

Triflate was also examined as a pseudohalide which Fearnley considers "the shape of things to come:" oxazolone triflatation is readily accomplished; and the triflate is routinely used in cross-coupling.[149] Examples include Vintonyak *et al.*[154] who synthesise the triflate from the oxazolone in 88% yield; Panek *et al.*[159,196] who demonstrate the that it may be synthesised and cross-coupled in the presence of an OTBDPS group; and others.[149,159,197] No precedents for evaluating its reactivity as a leaving group with standard heteroatom nucleophiles were found. In contrast to the frustrations of the chlorination, the first trial triflatation gave a 57% isolated yield of the triflate **134** after column (Figure II.48). A trial S_NAr reaction with **134** and NaOBn in dioxane however gave three reaction products: dibenzyl ether, the benzyl ester oxazolone, and regenerated oxazolone **99**; this highlights the good leaving-group character of the oxazolone unit. It is likely that the triflate sulfur is more reactive than the ring carbon for this incoming nucleophile, and that the outgoing benzyl triflate is then attacked by a second benzoxide to generate the ether. The S_NAr reaction was therefore abandoned.

7.2.2. Second direct O-alkylation strategy: initial proofs

I returned to the regioselectivity problem with two final clues: Seki's benzylated oxazolidinone, synthesised from a 4-phenyloxazolidinone-5-alkyl-ω-ester with $BnBr/Ag_2CO_3$ in hot PhMe in 83% yield,[181] and Fearnley's statement regarding oxazolones: "…N-alkylation… may prove less prevalent if non-polar solvents or 2-phase methods were to be employed."[149] I tested O-alkylations on both (4-unsubstituted) 5-ω-*siloxy*alkyl oxazolidinones, and 4-alkyl-5-ω-siloxyalkyl oxazolones. My initial electrophiles were BnBr, $p$$NO_2$BnBr, and tetraacetobromogalactose. Based on initial results, I adapted Seki's method[181] for longer reaction time, changing the stoichiometry and performing it in the dark, to obtain better yields of the desired O-alkylated products **135**, **136** and **137**. Undesired, N-alkylated products dominated for the oxazolidinones (**138** and **139**), but were minor for oxazolone (**140**).

Several other byproducts of ring opening, desilylation, desilylation/alkylation (**141**) and even redox were isolated. Moving away from the benzylic bromides, no products of any type were confirmed for galactosylation trials (Figure II.49).

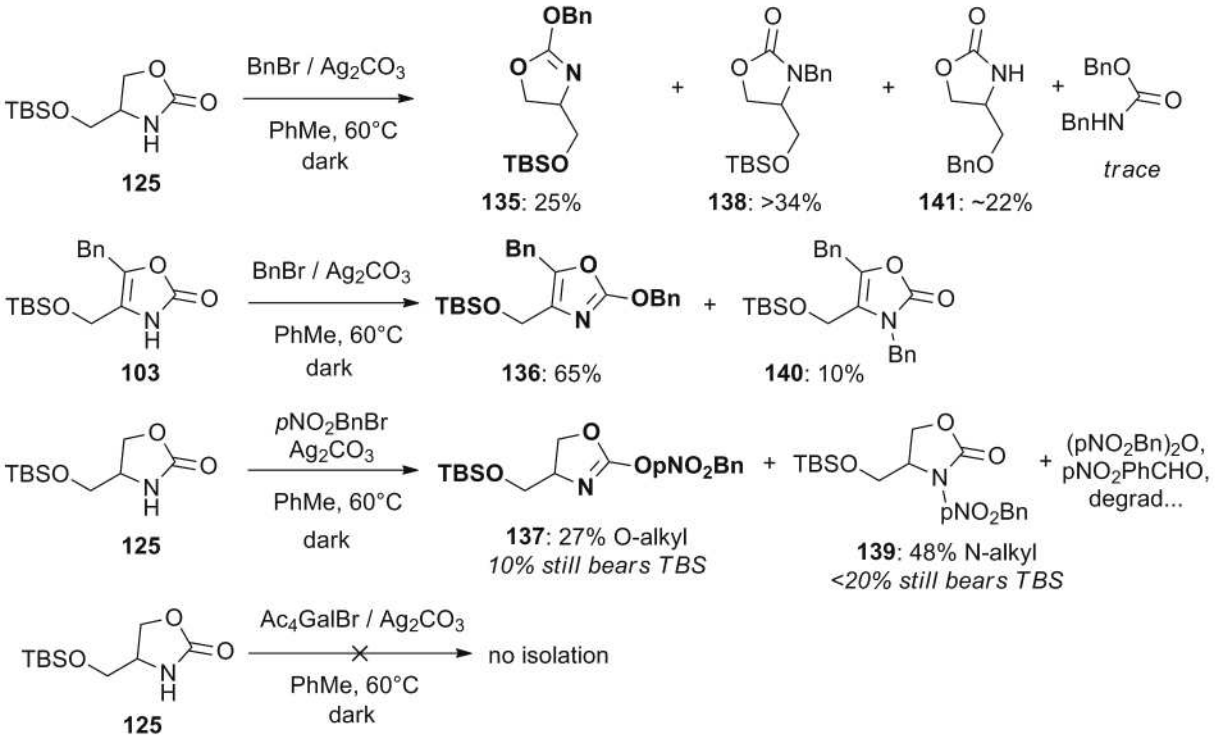


Figure II.50 – First trials of the final O-alkylation method

The yields were acceptable for pursuing initial electrophilic derivatives for alkylation onto model amines or ligand. First, desilylation was required. I wanted a non-aqueous desilylation procedure to avoid the possibility of O-alkyl hydrolysis (cf. behaviour observed for **134**) before stabilisation of the carbimidate could be effected by complexation. Kaburagi and Kishi[134] reported a convenient method which does not require purification afterwards: TBAF/THF followed by absorption of the TBA^+ and excess fluoride on $CaCO_3$ and Dowex resin. This was used with success on **136** (>95%) yielding alcohol **142**, however it resulted in extensive O-debenzylation of **135** and **137**, so for these compounds a different method of forming the leaving group was devised (section II.7.2.3) Pursuing alcohol **142**, reaction with TsCl was examined, towards desired tosylate **143**, however this was not isolated from the reaction. Reexamination of the products showed the chlorinative mechanism which has been introduced earlier (section II.7.1.2) and the chloride **144** was recovered although in low yields (25% after column). Note that this amenability contrasts with *non-O-alkylated* oxazolone electrophilic derivatisation, where only transient products were observed before decomposition (section II.7.1.2). This chloride was reacted with piperidine in EtOH, as a secondary amine model for the macrocyclic ligands of interest, giving high isolated yields of the ***model N-grafted, O-alkylated*** species **145** even on a small scale (80%; Figure II.50): a model for the ligand of **92**.

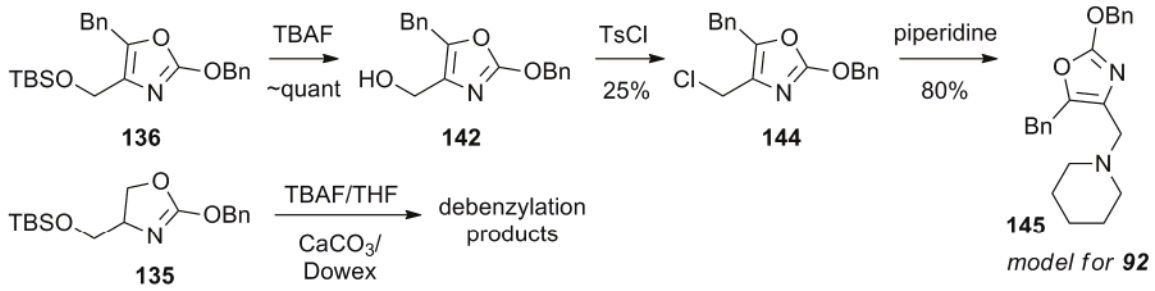


Figure II.51 – initial derivatisation/grafting trials towards 145 as a model compound for the ligand of 92

7.2.3. Second direct O-alkylation strategy: approach to substrates

To pursue enzymatically-activatable derivatives of species such as **145**, the enzyme substrate must be attached at the O-benzylated position. However the failure of galactosylation led me to pursue simple PHBA- and PABA-attached substrates, so that the ***benzyl bromide O-alkylation chemistry*** could be directly exploited. The PHBA/PABA then acts as the upstream spacer which should immolate after enzymatic activity, to give overall behaviour as shown in Figure II.35 (G = PABA-substrate or PHBA-substrate).

I therefore synthesised a range of PHBA/PABA bromides; most are known, though some rather convenient protocols were made up with success (**147**, **153**). The penicillen amidase (PGA) substrate was prepared from PABA acylation to **146** as per Nunez *et al.*[198], followed by bromination with HBr/AcOH to give a probably quantitative yield of unreported bromide **147**. The leucine aminopeptidase (LAP) substrate was prepared from Boc-leucine, forming the hydroxysuccinimide ester **148**[199] then reacting with PABA to give alcohol **149** following the procedure of Maekawa *et al.*[200], then brominating by CBr_4/PPh_3 to give the target bromide **150**. Synthesis of a second LAP substrate was begun, to exploit the probable enzyme-recognition advantage that has been discussed for the novel cyclisation spacer (fluorescent probes of Part I): leftover **148** was reacted with **22** in chloroform to give **25** in 35% yield but with superior separation, and chloroformylation of *para*hydroxybenzaldehyde then reaction with **25** gave **151**, for later reduction and bromination. A β-galactosidase substrate was prepared by anchimerically-assisted formation of β-galactoside **152** by BF_3-activated attack on the pentaacetate with *p*-cresol. This was radically brominated with NBS (1.15 eq) and AIBN thermal initiator giving bromide **153** in 82% yield. These reactions are summarised in Figure II.51.

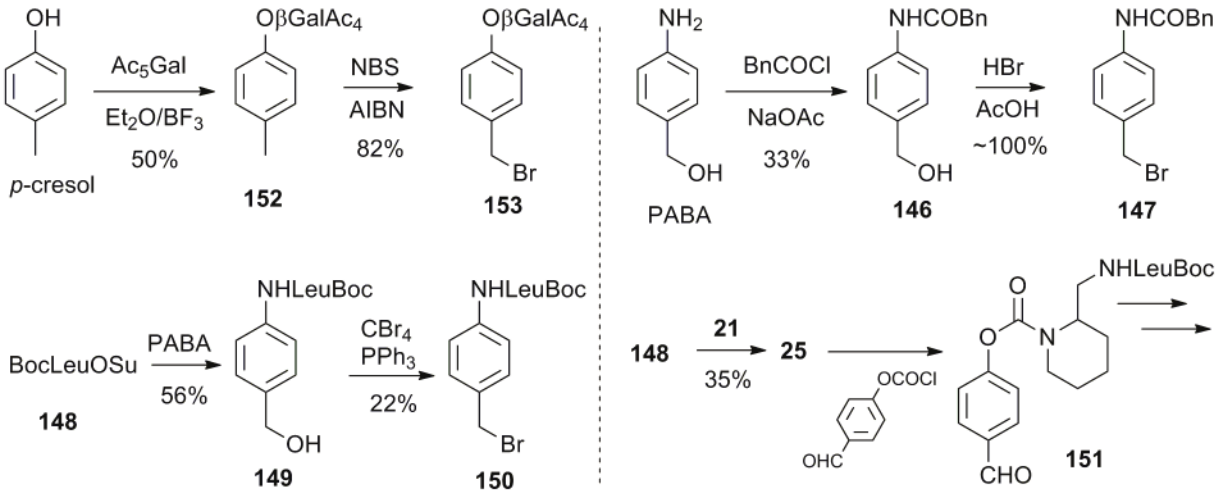


Figure II.52 – Syntheses of substrate-bearing benzylic bromides towards enzyme-responsive targets

7.2.4. Final O-alkylations

Several points from the initial O-alkylation trials led to a refined O-alkylation strategy which I considered a better option. Firstly, extensive desilylation was provoked during the initial O-alkylation reactions, especially of the oxazolidinones; but the next step was to be a desilylation anyway. As per von Platen's principle[201] I therefore designed conditions to specifically *perform* tandem ring-O-alkylation/exocyclic-desilylation. I also wanted to avoid exocyclic alkylation (such as in **141**), and ring-opening/dehydrative decompositions (which probably led to $(p\mathrm{NO_2Bn})_2\mathrm{O}$ formation in the reaction to **137**).

O-alkylation still requires a silver salt, but I required lower basicity in the medium, so that the acid liberated by O-alkylation would be able to give desilylation, and so the desilylated product would not risk undergoing exocyclic alkylation. Yet, if the solution acidity is too high, ring-opening decompositions are the major feature. I therefore trialled different mixtures of Ag_2O/ Ag_2CO_3,±K_2CO_3, as the source of base (in close excess), with amounts of AgOTf/AgOTs/AgOAc (varying amounts from 0.1 to ~10 eq relative to Ag_2O/Ag_2CO_3) as more-available electrophilic activators. The idea was that these relatively soluble species AgX should provoke O-alkylation by precipitating AgBr from the benzyl bromides; the liberated HX should then give desilylation, but be slowly absorbed into the much less soluble Ag_2CO_3; this regenerates the soluble activator and keeps acidity under control. I found that approximately 1:1 Ag_2O:AgOTs or Ag_2O:AgOTf gave identical behaviour: tandem O-alkylation/desilylation in good yields.

With little experimental time remaining, I applied this protocol to the syntheses of several O-alkylated targets; I used the enzyme substrate spacers elaborated above, as well as benzyl and *para*-nitrobenzyl bromides, sweeping the chemical space to see what reactions would work best (TableII.1).

Run	Carbamate	RBr	Base	Observations	Products/Yield
1	**103** (5-zolo)	**150**	Ag_2O:AgOTf 5:4	RBr decomp?	-
2	**103** (5-zolo)	**147**	Ag_2O:AgOTf 5:4	RBr decomp?	-
3	**102** (5-zolo)	*p*NO_2Bn	Ag_2O:AgOTf 7:10	mainly O-alkyl	acceptable **154**
4	**109** (6-zolo)	*p*NO_2Bn	Ag_2O:AgOTf 7:10	mixed N:O-alkyl mainly decomp	-
5	**130a** (6-lido)	*p*NO_2Bn	Ag_2O:AgOTf 4:5	3:1 N:O-alkyl	acceptable **155**
6	**125** (5-lido)	**153**	Ag_2O:AgOTf:AgOTs 2: 0.1 :3	decomp / redox?	-
7	**125** (5-lido)	Bn	Ag_2O:AgOTf:AgOTs 3: 0.1 :2	mainly O-alkyl	good **156**
8	**125** (5-lido)	*p*NO_2Bn	Ag_2O:AgOTf:AgOTs 1: 0.1 :0.1	mixed N:O-alkyl	acceptable **137**
9	**110** (7-zolo)	Bn	Ag_2O:AgOTs 1:2	exclusive O-alkyl much decomp	acceptable **158**
10	**130b** (7-lido)	Bn	Ag_2O:AgOTs 1:3	mixed N:O-alkyl	acceptable **159**

Table II.1 – Final sets of tandem O-alkylation/desilylations; zolo/lido = oxazolone/oxazolidinone, 5/6/7 = chelate size.

Several of the reactions gave desired O-alkylated, desilylated products directly, including the oxazolidinones (eg. **155**); this was pleasing as they had decomposed during the original desilylation procedure, so this one-pot synthesis gave a faster access to the desired species. In general, the yields were low, and as this was the last reaction I performed before ceasing experimentation several of the crudes are stored in the freezer; these will be reexamined in May 2013. Since the enzyme-sensitive substrates *appeared* to suffer degradation during these conditions by crude NMR, their reactions were not tackled yet; however, it has been seen on some other O-alkylations that extensive decomposition can hide signs of favourable reaction, so the answers will be known after the purifications are completed.

7.3. Towards enzyme-responsive ligands

145 had provided a good precedent towards the grafting of O-alkylated alcohols such as **142** onto a secondary amine. I therefore returned to **142**, attempting to improve the electrophile-forming step by using a one-pot procedure without isolation of the reactive intermediate. Wishing to form the electrophile quantitatively, but not to waste the expensive macrocyclic amine ligand, I launched a mesylation with triethylamine as the base and evaporated all unreacted materials after fifteen minutes, then added a solution of the macrocycle. After further reaction time, I observed the formation of a precipitate, and TLC on alumina showed the formation of a promising, faster-running spot; DIMS(+) likewise indicated the presence of desired product **160**, as well as the starting material **142**.

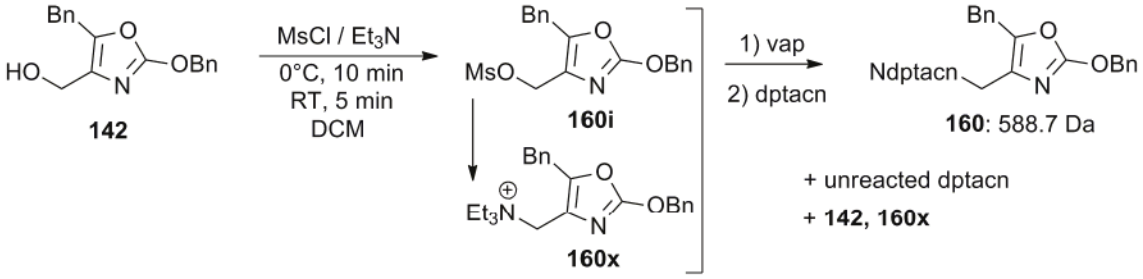


Figure II.53 – Second grafting reaction, using MsCl/NEt_3 instead of TsCl/N-butylimidazole, resulted in significant quaternary ammonium formation

An initial alumina column failed to separate the products entirely, however the DIMS(+) confirmed the fraction containing **160** and dptacn was returned. One complication was that mesylation also formed a substantial yield of the quaternised triethylammonium salt **160x**; in the future, reaction times and temperature will have to be diminished, or else, a weaker-nucleophile base such as DBU used instead. Perfect separation is required for reliable complexation studies since we aim to check the formation of an off-state compound, and even a small amount of ON-state impurity (from unreacted pentadentate ligand carried over) can interfere with measurements (cf. section I.4). Since my experimentation time had expired at this moment, I paused the purification/complexation, to begin in May 2013 when my experimentation time is again permitted.

General Conclusions from Part II

In the furyl spacer part, this PhD work has involved the first syntheses and characterisations of 2-furanol glycosides. Promising future avenues have been opened up for their use as high-energy auto-immolative spacers, which may be substantially more powerful alternatives to quinone methides, and which may permit the success of the off-ON MRI contrast agent design **44b**. The exploration of some of their chemistry including their elevated capacity to effect glycosylation has begun.

In the second part concerning cyclic carbamate spacers (oxazolones and oxazolidinones), this PhD work provides the first demonstration of an extensible method for preparing substrate-substituted functionalised cyclic carbamates, towards the synthesis of enzyme-activated magnetogenic complexes **92a**. The spin-switching possibility of the cyclic carbamate strategy remains to be shown: however the almost entirely unexplored procedures for (1) generalisable O-alkylation of these species with a future enzyme substrate, then (2) a tolerant route towards ligand-grafted O-alkylated species, have now been proven for model compounds, and the enzyme-responsive analogues are awaiting further reactions in the freezer.

In summary, the syntheses of the required model compounds based on both furan and cyclic carbamate spacer systems are still works in progress, and May 2013, when experimentation can begin again, is eagerly awaited. However, the proofs of concept of the novel designs and reactions which have already been achieved have broken the ground for the most unprecedented transforms. Together with the likely routes for further progress which have been identified, these comprehensive precedents on these novel functional scaffolds pave a clear way for exploiting them in the design of intelligent proagents and prodrugs.

Future Directions

Preliminary results have indicated that the hydrolytic stability of especially the initial generation of furan spacers may be low, but that the cyclic carbamates may also be somewhat affected. Therefore *stabilising strategies* such as steric crowding, intended to improve the hydrolytic stability of the initial spacers, are prime targets for investigation in the near future. Hopefully, such stabilised spacers will allow for the desired robust, off-ON-mode applications to magnetogenic probes, by both furan and carbamate strategies.

Also, now that glycosyl furans can be formed stereoselectively, and given the unusually labile glycosidic bond which they have shown, the *unstabilised* furan system will also be examined as a potential *glycosylation* reagent; applications in Diels-Alder reactions may also be of interest in creating complex, densely-functionalised structures.

Evaluating the enzymatic processing efficiency on the new glycosyloxyfurans, and evaluation/ demonstration of the unique reactivity that both the furans and the cyclic carbamates provide, is vital to capitalise fully on the synthetic advances that have been made during this PhD work so far, and hopefully both avenues will yield promising results over the next few months.

On the longer term also, this PhD work may open up new perspectives in the field beyond the applications of the molecular targets it has explored, but rather, inspired by its general approach. To cite just one example from this chapter, I consider that by requiring only the elimination of the weakest bond, the cyclic carbamate system provides an excellent design for a more refined approach to off-ON spin-switched complexes: the crucial effect is achieved by addressing the crucial bond. However, taking a longer perspective, it seems from the group's current experiences that the toxicity of Fe(II) complexes with a free coordination site may be prohibitive to their use at high concentration in routine clinical applications (cf. Fenton reactions and oxidative cycling). Yet one can hope that if this mode of toxicity continues to pose a problem, then the *approach* of the cyclic carbamate system points the way towards the best solution for future, *biocompatible* off-ON Fe(II) complexes. The idea that I consider, is that a change of *ligand field strength*, is all that is *really* required for a spin switching effect. The cyclic carbamates focus on the 'correct' crucial, coordinative bond, but a still more refined approach would be to retain it, only changing the *donor character* of the ligand atom. Even if the ON-state signal would be relatively small since second-sphere effects would be the only means of bulk relaxation (q=0), such a *coordinatively saturated complex* would address this principal toxicity concern, and therefore permit a much higher dose of the agent to be applied (or, any dose at all to be approved for clinical settings). The 'soft' cyclic carbamate system shows one (fairly powerful) method for achieving this change of coordinative strength; yet similar and still softer systems exploiting other 'latently poor donors' may rather easily be drawn up, and hopefully may see synthetic efforts in the future.

Key publications – Part II

- **Classical T_1 MRI Contrast Agents:** Hermann P, Kotek J, Kubicek V, Lukes I. Gadolinium(III) Complexes as MRI Contrast Agents: Ligand Design and Properties of the Complexes. *J. Chem. Soc., Dalton Trans.* **2008**, 3027-3047 (DOI weblink: 10.1039/B719704G).
- **Fe(II) spin-switch complexes, prior art:** Stavila V, Stortz Y, Franc C, Pitrat D, Maurin P, Hasserodt J. Effective Repression of the Fragmentation of a Hexadentate Ligand Bearing an Auto-Immolable Pendant Arm by Iron Coordination. *Eur. J. Inorg. Chem.* **2008**, 3943-3947 (DOI weblink: 10.1002/ejic.200800419).
- **2-Furanone System, General:** Rao YS. Recent Advances in the Chemistry of Unsaturated Lactones. *Chemical Reviews* **1976**, 76, 625-694 (DOI weblink: 10.1021/cr60303a004).
- **2-Furanol System:** O. Thorn-Seshold. *Towards new molecular candidates for an auto-immolative coordinating arm on macrocyclic ligands* (M2 Thesis), Ecole Normale Supérieure de Lyon, **2009**.
- **General Alkynyl Ketone Cycloisomerisation:** Kel'in AV, Gevorgyan V. Efficient Synthesis of 2-Mono- and 2,5-Disubstituted Furans Via the CuI-Catalyzed Cycloisomerization of Alkynyl Ketones. *J. Org. Chem.* **2002**, 67, 95-98 (DOI weblink: 10.1021/jo010832v).
- **Picolyl Alkylnyl Ketone Cycloisomerisation:** Natarajan SR, Chen M-H, Heller ST, Tynebor RM, Crawford EM, Minxiang C, Kaizheng H, Dong J, Hu B, Hao W, Chen S-H. Synthesis of the 2*H*-Quinolizin-2-one Scaffold via a Stepwise Acylation-Intramolecular Annulation Strategy. *Tetrahedron Lett.* **2006**, 47, 5063-5067 (DOI weblink: 10.1016/j.tetlet.2006.05.089).
- **Furan S_NAr:** Morris WJ, Shair MD. Stereoselective Synthesis of 2-Deoxy-β-Glycosides Using Anomeric O-Alkylation/Arylation. *Org. Lett.* **2008**, 11, 9-12 (DOI weblink: 10.1021/ol8022006).
- **Oxazolone Chemistry:** Fearnley SP. 2-(3*H*)-Oxazolone - A Simple Heterocycle with Manifold Potential. *Curr. Org. Chem.* **2004**, 8, 1289-1337 (DOI weblink: 10.2174/1385272043369971).

Bibliography – Part II

DOI are indicated where possible (and failing that, PubMed ID [PMID], NII Article ID [NAID] or ISI Web of Science index [WOS]), for fastest lookup.

[1] P. Hermann, J. Kotek, V. Kubicek, I. Lukes; *J. Chem. Soc., Dalton Trans.* **2008**, 3027-3047 *(10.1039/B719704G)*.

[2] V. Stavila, M. Allali, L. Canaple, Y. Stortz, C. Franc, P. Maurin, O. Beuf, O. Dufay, J. Samarut, M. Janier, J. Hasserodt; *New J. Chem.* **2008**, *32*, 428-435 *(10.1039/b715254j)*.

[3] J. Gong, J. P. Hornak; *Magn. Reson. Imag.* **1992**, *10*, 623-626 *(PMID:1501532)*.

[4] E. T. Ahrens, U. Rothbächer, R. E. Jacobs, S. E. Fraser; *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 8443-8448 *(PMID: 9671697)*.

[5] S. Aime, P. Anelli, M. Botta, M. Brocchetta, S. Canton, F. Fedeli, E. Gianolio, E. Terreno; *J. Biol. Inorg. Chem.* **2002**, *7*, 58-67 *(10.1007/s007750100265)*.

[6] A. Merbach, L. Helm, É. Tóth; *The Chemistry of Contrast Agents in Medical Magnetic Resonance Imaging*, John Wiley & Sons, Chichester (UK), **2013**.

[7] P. Caravan, J. J. Ellison, T. J. McMurry, R. B. Lauffer; *Chemical Reviews* **1999**, *99*, 2293-2352 *(10.1021/cr980440x)*.

[8] L. Thunus, R. Lejeune; *Coordination Chemistry Reviews* **1999**, *184*, 125-155 *(10.1016/S0010-8545(98)00206-9)*.

[9] V. M. Runge, M. A. Foster, J. A. Clanton, M. M. Jones, C. M. Lukehart, J. M. Hutchison, J. R. Mallard, F. W. Smith, C. L. Partain, A. E. James; *Radiology* **1984**, *152*, 123-126 *(PMID: 6427845)*.

[10] J.-M. Idée, M. Port, C. Corot; *Radiographics* **2009**, *29*, 2099-2099 *(PMID: 19926764)*.

[11] J. G. Penfield, R. F. Reilly; *Nat Clin Pract Neph* **2007**, *3*, 654-668 *(10.1038/ncpneph0660)*.

[12] R. B. Lauffer; *Magn. Reson. Med.* **1991**, *22*, 339-342; discussion 343-336 *(10.1002/mrm.1910220237)*.

[13] R. Hovland, A. J. Aasen, J. Klaveness; *Org. Biomol. Chem.* **2003**, *1*, 1707-1710 *(10.1039/b301944f)*.

[14] K. Hanaoka, K. Kikuchi, T. Terai, T. Komatsu, T. Nagano; *Chem.–Eur. J* **2008**, *14*, 987-995 *(10.1002/chem.200700785)*.

[15] A. L. Nivorozhkin, A. F. Kolodziej, P. Caravan, M. T. Greenfield, R. B. Lauffer, T. J. McMurry; *Angew. Chem., Int. Ed.* **2001**, *40*, 2903-2906 *(10.1002/1521-3773(20010803)40:15<2903::AID-ANIE2903>3.0.CO;2-N)*.

[16] E. L. Que, C. J. Chang; *Chem. Soc. Rev.* **2010**, *39*, 51-60 *(10.1039/b914348n)*.

[17] A. Y. Louie, M. M. Huber, E. T. Ahrens, U. Rothbacher, R. Moats, R. E. Jacobs, S. E. Fraser, T. J. Meade; *Nat. Biotechnol.* **2000**, *18*, 321-325 *(10.1038/73780)*.

[18] R. A. Moats, S. E. Fraser, T. J. Meade; *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 726-728 *(10.1002/anie.199707261)*.

[19] D. Bini, M. Gregori, U. Cosentino, G. Moro, A. Canales, A. Capitoli, J. Jiménez-Barbero, L. Cipolla; *Carbohydr. Res.* **2012**, *354*, 21-31 *(10.1016/j.carres.2012.03.002)*.

[20] J. A. Duimstra, F. J. Femia, T. J. Meade; *J. Am. Chem. Soc.* **2005**, *127*, 12847-12855 *(10.1021/ja042162r)*.

[21] T. Chauvin, P. Durand, M. Bernier, H. Meudal, B. T. Doan, F. Noury, B. Badet, J. C. Beloeil, E. Toth; *Angew. Chem., Int. Ed.* **2008**, *47*, 4370-4372 *(10.1002/anie.200800809)*.

[22] E. Gianolio, L. Maciocco, D. Imperio, G. B. Giovenzana, F. Simonelli, K. Abbas, G. Bisi, S. Aime; *Chem. Commun.* **2011**, *47*, 1539-1541 *(10.1039/c0cc03554h)*.

[23] J. Hasserodt; *New J. Chem.* **2012**, *36*, 1707-1712 *(10.1039/c2nj40209b)*.

[24] J. Hasserodt; *Contrast Agents for Magnetic Resonance Imaging (WO2005094903)*, **2005**.

[25] V. Stavila, Y. Stortz, C. Franc, D. Pitrat, P. Maurin, J. Hasserodt; *Eur. J. Inorg. Chem.* **2008**, 3943-3947 *(10.1002/ejic.200800419)*.

[26] D. Sappey-Marinier, O. Beuf, C. Billotey, E. Chereul, J. Dupuy, M. Jeandey, D. Grenier, J. Hasserodt, J. B. Langlois, C. Lartizien, W. Mai, C. Odet, J. Samarut, D. Vray, L. Zimmer, M. Janier; *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **2004**, *527*, 117-123 *(10.1016/j.nima.2004.03.092)*.

[27] W. S. Szulbinski; *Pol. J. Chem.* **2000**, *74*, 109-124 *(WOS: 000084644300009)*.

[28] M. J. Pittet, F. K. Swirski, F. Reynolds, L. Josephson, R. Weissleder; *Nat. Protoc.* **2006**, *1*, 73-79 *(10.1038/nprot.2006.11)*.

[29] F. Touti, P. Maurin, J. Hasserodt; *Angew. Chem., Int. Ed.* **2013**, *52*, 4654–4658 *(10.1002/anie.201208848)*.

[30] J. L. Kolanowski, E. Jeanneau, R. Steinhoff, J. Hasserodt; *Chem.–Eur. J* **2013**, n/a-n/a *(10.1002/chem.201300604)*.

[31] P. L. Carl, P. K. Chakravarty, J. A. Katzenellenbogen; *J. Med. Chem.* **1981**, *24*, 479-480 *(10.1021/jm00137a001)*.

[32] J. P. Krise, R. Oliyai; in *Prodrugs*, (*Vol. 5*), Springer, New York, **2007**, pp. 801-831.

[33] T. Rosenau, W. D. Habicher; *Chem. Pharm. Bull.* **1997**, *45*, 1080-1084 *(10.1248/cpb.45.1080)*.

[34] D. Shabat, R. J. Amir, A. Gopin, N. Pessah, M. Shamis; *Chem.–Eur. J* **2004**, *10*, 2626-2634 *(10.1002/chem.200305715)*.

[35] O. Thorn-Seshold; *Towards new molecular candidates for an auto-immolative coordinating arm on macrocyclic ligands* (M2 Thesis), Ecole Normale Supérieure de Lyon, **2009**.

[36] L. Christiansen, D. N. Hendrickson, H. Toftlund, S. R. Wilson, C. L. Xie; *Inorg. Chem.* **1986**, *25*, 2813-2818 *(10.1021/ic00236a031)*.

[37] Y. S. Rao; *Chemical Reviews* **1976**, *76*, 625-694 *(10.1021/cr60303a004)*.

[38] R. Antonioletti, M. D'Auria, A. De Mico, G. Piancatelli, A. Scettri; *Tetrahedron* **1984**, *40*, 3805-3808 *(10.1016/S0040-4020(01)88810-4)*.

[39] O. E. O. Hormi, J. H. Näsman; *Synth. Commun.* **1986**, *16*, 69-77 *(10.1080/00397918608057690)*.

[40] E. Negishi, M. Kotora; *Tetrahedron* **1997**, *53*, 6707-6738 *(10.1016/S0040-4020(97)00199-3)*.

[41] G. A. Kraus, H. Sugimoto; *J. Chem. Soc., Chem. Commun.* **1978**, 30-30 *(10.1039/C3978000030A)*.

[42] M. D'Auria, G. Piancatelli, A. Scettri; *Tetrahedron* **1980**, *36*, 3071-3074 *(10.1016/0040-4020(80)88034-3)*.

[43] C. W. Jefford, A. W. Sledeski, J.-C. Rossier, J. Boukouvalas; *Tetrahedron Lett.* **1990**, *31*, 5741-5744 *(10.1016/S0040-4039(00)97946-2)*.

[44] M. Asaoka, N. Yanagida, K. Ishibashi, H. Takei; *Tetrahedron Lett.* **1981**, *22*, 4269-4270 *(10.1016/S0040-4039(01)82929-4)*.

[45] H. Nagao, Y. Yamane, T. Mukaiyama; *Chem. Lett.* **2007**, *36*, 8-9 *(10.1246/cl.2007.8)*.

[46] J. Boukouvalas, F. Maltais, N. Lachance; *Tetrahedron Lett.* **1994**, *35*, 7897-7900 *(10.1016/0040-4039(94)80005-7)*.

[47] G. Grossmann, B. Jolivet, M. Bornand, U. Sequin, K. D. Spindler; *Synthesis* **2005**, 1543-1549 *(10.1055/s-2005-865356)*.

[48] S. D. Carter, D. Thetford, M. Voyle; *J. Chem. Soc.-Perkin Trans. 1* **1990**, 1231-1233 *(10.1039/P19900001231)*.

[49] A. F. Oleinik, K. Y. Novitskii, L. V. Brattseva, T. A. Gus'kova, G. N. Pershin, A. I. Kravchenko, V. A. Chernov; *Pharmaceutical Chemistry Journal* **1976**, *10*, 617-621 *(10.1007/BF00757689)*.

[50] G. F. D'Alelio, C. J. Williams, C. L. Wilson; *J. Org. Chem.* **1960**, *25*, 1028-1030 *(10.1021/jo01076a043)*.

[51] J. A. Cella; *J. Org. Chem.* **1988**, *53*, 2099-2103 *(10.1021/jo00244a048)*.

[52] J. E. Garst, G. L. Schmir; *J. Org. Chem.* **1974**, *39*, 2920-2923 *(10.1021/jo00933a021)*.

[53] D. G. Manly, E. D. Amstutz; *J. Org. Chem.* **1956**, *21*, 516-519 *(10.1021/jo01111a008)*.

[54] A. V. Kel'in, V. Gevorgyan; *J. Org. Chem.* **2002**, *67*, 95-98 *(10.1021/jo010832v)*.

[55] N. T. Patil, Y. Yamamoto; *Chemical Reviews* **2008**, *108*, 3395-3442 *(10.1021/cr050041j)*.

[56] E. Duranti, C. Balsamini; *Synthesis* **1974**, *1974*, 357-358 *(10.1055/s-1974-23320)*.

[57] H. Gilman, C. C. Vernon; *J. Am. Chem. Soc.* **1924**, *46*, 2576-2579 *(10.1021/ja01676a032)*.

[58] L. Pevzner; *Russian Journal of General Chemistry* **2006**, *76*, 621-625 *(10.1134/S1070363206040207)*.

[59] H. Gilman, G. F. Wright; *Chemical Reviews* **1932**, *11*, 323-367 *(10.1021/cr60040a002)*.

[60] R. Blanc, V. Heran, R. Rahmani, L. Commeiras, J.-L. Parrain; *Org. Biomol. Chem.* **2010**, *8*, 5490-5494 *(10.1039/c0ob00448k)*.

[61] G. Piancatelli, A. Scettri, M. D'Auria; *Tetrahedron Lett.* **1979**, *20*, 1507-1508 *(10.1016/S0040-4039(01)86191-8)*.

[62] J. T. Kim, A. V. Kel'in, V. Gevorgyan; *Angew. Chem., Int. Ed.* **2003**, *42*, 98-101 *(10.1002/anie.200390064)*.

[63] A. R. Lingham, T. J. Rook, H. M. Hugel; *Aust. J. Chem.* **2002**, *55*, 795-798 *(10.1071/ch02178)*.

[64] B. M. Trost, J. Xie; *J. Am. Chem. Soc.* **2006**, *128*, 6044-6045 *(10.1021/ja0602501)*.

[65] K. Inamoto, A. Yamamoto, K. Ohsawa, K. Hiroya, T. Sakamoto; *Chem. Pharm. Bull.* **2005**, *53*, 1502-1507 *(10.1248/cpb.53.1502)*.

[66] B. Roy, B. Mukhopadhyay; *Tetrahedron Lett.* **2007**, *48*, 3783-3787 *(10.1016/j.tetlet.2007.03.165)*.

[67] H. B. Mereyala, S. R. Gurrala, S. K. Mohan; *Tetrahedron* **1999**, *55*, 11331-11342 *(10.1016/S0040-4020(99)00631-6)*.

[68] H. B. Mereyala, S. R. Gurrala; *Carbohydr. Res.* **1998**, *307*, 351-354 *(10.1016/S0008-6215(97)10104-5)*.

[69] J. D. White, C. M. Lincoln, J. Yang, W. H. C. Martin, D. B. Chan; *J. Org. Chem.* **2008**, *73*, 4139-4150 *(10.1021/jo800335g)*.

[70] H. Ishii, T. Ishikawa, S. Takeda, S. Ueki, M. Suzuki, T. Harayama; *Chem. Pharm. Bull.* **1992**, *38*, 1775-1777 *(NAID: 110003628646)*.

[71] M. J. Piggott, D. Wege; *Aust. J. Chem.* **2000**, *53*, 749-754 *(10.1071/CH00057)*.

[72] M. Piotto, M. Bourdonneau, K. Elbayed, J.-M. Wieruszeski, G. Lippens; *Magn. Reson. Chem.* **2006**, *44*, 943-947 *(10.1002/mrc.1884)*.

[73] B. V. Joshi, C. B. Reese; *J. Chem. Soc., Perkin Trans. 1* **1992**, 441-443 *(10.1039/P19920000441)*.

[74] A. J. Moreno-Vargas, J. G. Fernández-Bolaños, J. Fuentes, I. Robina; *Tetrahedron: Asymmetry* **2001**, *12*, 3257-3266 *(10.1016/S0957-4166(01)00568-7)*.

[75] K. Ando, Y. Kawamura, Y. Akai, J.-i. Kunitomo, T. Yokomizo, M. Yamashita, S. Ohta, T. Ohishi, Y. Ohishi; *Org. Biomol. Chem.* **2008**, *6*, 296-307 *(10.1039/b710935k)*.

[76] M. Frigerio, M. Santagostino, S. Sputore; *J. Org. Chem.* **1999**, *64*, 4537-4538 *(10.1021/jo9824596)*.

[77] K. C. Nicolaou, T. Montagnon, P. S. Baran, Y. L. Zhong; *J. Am. Chem. Soc.* **2002**, *124*, 2245-2258 *(10.1021/ja012127+)*.

[78] S. R. Natarajan, M.-H. Chen, S. T. Heller, R. M. Tynebor, E. M. Crawford, C. Minxiang, H. Kaizheng, J. Dong, B. Hu, W. Hao, S.-H. Chen; *Tetrahedron Lett.* **2006**, *47*, 5063-5067 *(10.1016/j.tetlet.2006.05.089)*.

[79] M. Journet, D. Cai, L. M. DiMichele, R. D. Larsen; *Tetrahedron Lett.* **1998**, *39*, 6427-6428 *(10.1016/S0040-4039(98)01352-5)*.

[80] D. P. Wang, K. Ding; *Chem. Commun.* **2009**, 1891-1893 *(10.1039/b821212k)*.

[81] Y. Kurashina, H. Miyata, D.-i. Momose; *4H-quinolizin-4-one compounds exhibiting therapeutic activities (EP277755)*, **1988**.

[82] Z. Guo, Y.-F. Zhu, T. D. Gross, F. C. Tucci, Y. Gao, M. Moorjani, P. J. Connors, M. W. Rowbottom, Y. Chen, R. S. Struthers, Q. Xie, J. Saunders, G. Reinhart, T. K. Chen, A. L. K. Bonneville, Chen; *J. Med. Chem.* **2004**, *47*, 1259-1271 *(10.1021/jm030472z)*.

[83] F. Shirini, M. Esm-Hosseini, Z. Hejazi; *Synth. Commun.* **2005**, *35*, 2913 - 2919 *(10.1080/00397910500297362)*.

[84] A. Gomtsyan, R. G. Schmidt, E. K. Bayburt, J. F. Daanen, M. E. Kort; *TRPV1 Antagonists (WO2009055629)*, **2009**.

[85] T. Nicola, D. Schwarzrock, M. Keller, W. Eberbach; *Tetrahedron* **2001**, *57*, 1771-1777 *(10.1016/S0040-4020(01)00005-9)*.

[86] D.-W. Wang, X.-B. Wang, D.-S. Wang, S.-M. Lu, Y.-G. Zhou, Y.-X. Li; *J. Org. Chem.* **2009**, *74*, 2780-2787 *(10.1021/jo900073z)*.

[87] P. Y. S. Lam, Y. Ru, P. K. Jadhav, P. E. Aldrich, G. V. DeLucca, C. J. Eyermann, C.-H. Chang, G. Emmett, E. R. Holler, W. F. Daneker, L. Li, P. N. Confalone, R. J. McHugh, Q. Han, R. Li, J. A. Markwalder, S. P. Seitz, T. R. Sharpe, L. T. Bacheler, M. M. Rayner, R. M. Klabe, L. Shum, D. L. Winslow, D. M. Kornhauser, D. A. Jackson, S. Erickson-Viitanen, C. N. Hodge; *J. Med. Chem.* **1996**, *39*, 3514-3525 *(10.1021/jm9602571)*.

[88] R. P. Cassity, L. T. Taylor, J. F. Wolfe; *J. Org. Chem.* **1978**, *43*, 2286-2288 *(10.1021/jo00405a046)*.

[89] T. d. Haro, C. Nevado; *J. Am. Chem. Soc.* **2010**, *132*, 1512-1513 *(10.1021/ja909726h)*.

[90] C. Saccavini, C. Tedeschi, L. Maurette, C. Sui-Seng, C. Zou, M. Soleilhavoup, L. Vendier, R. Chauvin; *Chem.–Eur. J* **2007**, *13*, 4895-4913 *(10.1002/chem.200601191)*.

[91] S. Mukherjee, D. Kontokosta, A. Patil, S. Rallapalli, D. Lee; *J. Org. Chem.* **2009**, *74*, 9206-9209 *(10.1021/jo901950e)*.

[92] P. V. Ramachandran, A. V. Teodorovic, M. V. Rangaishenvi, H. C. Brown; *J. Org. Chem.* **1992**, *57*, 2379-2386 *(10.1021/jo00034a034)*.

[93] E. Alcalde, N. Mesquida, C. Alvarez-Rúa, R. Cuberes, J. Frigola, S. García-Granda; *Molecules* **2008**, *13*, 301-318 *(10.3390/molecules13020301)*.

[94] J. M. Williams, R. B. Jobson, N. Yasuda, G. Marchesini, U.-H. Dolling, E. J. J. Grabowski; *Tetrahedron Lett.* **1995**, *36*, 5461-5464 *(10.1016/0040-4039(95)01089-Z)*.

[95] I. R. Baxendale, G. Brusotti, M. Matsuoka, S. V. Ley; *J. Chem. Soc.-Perkin Trans. 1* **2002**, 143-154 *(10.1039/b109482n)*.

[96] R. J. Stedman, A. C. Swift, L. S. Miller, M. M. Dolan, J. R. E. Hoover; *J. Med. Chem.* **1967**, *10*, 363-366 *(10.1021/jm00315a017)*.

[97] U. Ghosh, D. Ganessunker, V. J. Sattigeri, K. E. Carlson, D. J. Mortensen, B. S. Katzenellenbogen, J. A. Katzenellenbogen; *Bioorg. Med. Chem.* **2003**, *11*, 629-657 *(10.1016/S0968-0896(02)00309-7)*.

[98] S. Raghavan, S. Mustafa, B. Sridhar; *J. Org. Chem.* **2009**, *74*, 4499-4507 *(10.1021/jo900569z)*.

[99] A. Nagaki, H. Kim, H. Usutani, C. Matsuo, J.-i. Yoshida; *Org. Biomol. Chem.* **2010**, *8*, 1212-1217 *(10.1039/b919325c)*.

[100] A. S. Dudnik, A. W. Sromek, M. Rubina, J. T. Kim, A. V. Kel'i, V. Gevorgyan; *J. Am. Chem. Soc.* **2008**, *130*, 1440-1452 *(10.1021/ja0773507)*.

[101] A. S. Dudnik, V. Gevorgyan; *Angew. Chem., Int. Ed.* **2007**, *46*, 5195-5197 *(10.1002/anie.200701128)*.

[102] R. A. Parker, T. Kariya, J. M. Grisar, V. Petrow; *J. Med. Chem.* **1977**, *20*, 781-791 *(10.1021/jm00216a009)*.

[103] R. J. Petfield, E. D. Amstutz; *J. Org. Chem.* **1954**, *19*, 1944-1946 *(10.1021/jo01377a011)*.

[104] R. Sornay, J. M. Meunier, P. Fournari; *Bull. Soc. Chim. Fr.* **1971**, 990-1000.

[105] M. C. Zaluski, M. Robba, M. Bonhomme; *Bull. Soc. Chim. Fr.* **1970**, 1838-1846.

[106] M. C. Zaluski, M. Robba, M. Bonhomme; *Bull. Soc. Chim. Fr.* **1970**, 1445-1450.

[107] A. Tanaka, T. Usui, M. Shimadzu; *Chem. Pharm. Bull.* **1980**, *28*, 2846-2849 *(NAID: 110003624113)*.

[108] C.-L. Chang; *Chem. Pharm. Bull.* **2009**, *57*, 550-556 *(10.1248/cpb.57.550)*.

[109] W. J. Morris, M. D. Shair; *Org. Lett.* **2008**, *11*, 9-12 *(10.1021/ol8022006)*.

[110] R. Tung, A. Morgan; *4-aminoquinazoline Derivatives and Methods of Use Thereof (US20110053964)*, **2011**.

[111] F. Sauter, P. Stanetty, W. Sittenthaler, R. Waditschatka; *Monatshefte für Chemie / Chemical Monthly* **1988**, *119*, 1427-1438 *(10.1007/bf00810286)*.

[112] E. J. Corey, C. U. Kim, M. Takeda; *Tetrahedron Lett.* **1972**, *13*, 4339-4342 *(10.1016/S0040-4039(01)94310-2)*.

[113] A. J. Moreno-Vargas, I. Robina, R. Demange, P. Vogel; *Helv. Chim. Acta* **2003**, *86*, 1894-1913 *(10.1002/hlca.200390152)*.

[114] J. Joule, K. Mills; *Heterocyclic Chemistry, (5th Ed),* John Wiley & Sons, Singapore, **2010**.

[115] F. Trécourt, G. Breton, V. Bonnet, F. Mongin, F. Marsais, G. Quéguiner; *Tetrahedron* **2000**, *56*, 1349-1360 *(10.1016/S0040-4020(00)00027-2)*.

[116] G. Piancatelli, A. Scettri, M. Dauria; *Tetrahedron Lett.* **1977**, 2199-2200 *(10.1016/S0040-4039(01)83720-5)*.

[117] L. Nie, J. Yong, H. Ba, H. A. Aisa; *Chin. J. Chem.* **2009**, *27*, 1716-1720 *(10.1002/cjoc.200990289)*.

[118] M. A. Keegstra, A. J. A. Klomp, L. Brandsma; *Synth. Commun.* **1990**, *20*, 3371-3374 *(10.1080/00397919008051573)*.

[119] P. Arsenyan, E. Paegle, S. Belyakov; *Tetrahedron Lett.* **2010**, *51*, 205-208 *(10.1016/j.tetlet.2009.10.133)*.

[120] N. O. Devarie-Baez, W. S. Kim, A. B. Smith, M. Xian; *Org Lett* **2009**, *11*, 1861-1864 *(10.1021/ol900434k)*.

[121] N. A. Petasis, A. Goodman, I. A. Zavialov; *Tetrahedron* **1997**, *53*, 16463-16470 *(10.1016/S0040-4020(97)01028-4)*.

[122] H. Mandai, K. Murota, T. Sakai; *Tetrahedron Lett.* **2010**, *51*, 4779-4782 *(10.1016/j.tetlet.2010.07.039)*.

[123] L. M. Harwood, G. S. Currie, M. G. B. Drew, R. W. A. Luke; *Chem. Commun.* **1996**, 1953-1954 *(10.1039/CC9960001953)*.

[124] W. J. Thompson, J. Gaudino; *J. Org. Chem.* **1984**, *49*, 5237-5243 *(10.1021/jo00200a045)*.

[125] X.-Q. Yu, T. Shirai, Y. Yamamoto, N. Miyaura; *Chem.–As. J.* **2011**, *6*, 932-937 *(10.1002/asia.201000589)*.

[126] T. Ishiyama, J. Takagi, Y. Yonekawa, J. F. Hartwig, N. Miyaura; *Adv. Syn. Cat.* **2003**, *345*, 1103-1106 *(10.1002/adsc.200303058)*.

[127] F. Touti, P. Maurin, L. Canaple, O. Beuf, J. Hasserodt; *Inorg. Chem.* **2012**, *51*, 31-33 *(10.1021/ic202268e)*.

[128] T. Yamanoi, N. Misawa, S. Matsuda, M. Watanabe; *Carbohydr. Res.* **2008**, *343*, 1366-1372 *(10.1016/j.carres.2008.03.024)*.

[129] J. Fiandor, M. T. Garcia-Lopez, F. G. De Las Heras, P. P. Mendez-Castrillan; *Synthesis* **1985**, *1985*, 1121-1123 *(10.1055/s-1985-31446)*.

[130] Y.-L. Leu, C.-S. Chen, Y.-J. Wu, J.-W. Chern; *J. Med. Chem.* **2008**, *51*, 1740-1746 *(10.1021/jm701151c)*.

[131] M. Sedlak, P. Drabina, E. Bilkova, P. Simunek, V. Buchta; *Bioorg. Med. Chem. Lett.* **2008**, *18*, 2952-2956 *(10.1016/j.bmcl.2008.03.065)*.

[132] F. Durrat, I. Texier-Nogues, V. Robert; *Substrats fluorescents saccharidiques, leur procédé de préparation et leurs utilisations (EP2033663A1)*, **2009**.

[133] S. Dong, L. A. Paquette; *J. Org. Chem.* **2005**, *70*, 1580-1596 *(10.1021/jo048071u)*.

[134] Y. Kaburagi, Y. Kishi; *Org. Lett.* **2007**, *9*, 723-726 *(10.1021/ol063113h)*.

[135] L. Bautista-Ibanez, K. Ramirez-Gualito, B. Quiroz-Garcia, A. Rojas-Aguilar, G. Cuevas; *J. Org. Chem.* **2008**, *73*, 849-857 *(10.1021/jo701926r)*.

[136] S. Mendonca, R. A. Laine; *Carbohydr. Res.* **2005**, *340*, 2055-2059 *(10.1016/j.carres.2005.05.016)*.

[137] R. Panchadhayee, A. K. Misra; *J. Carbohydr. Chem.* **2010**, *29*, 76 - 83 *(10.1080/07328301003664887)*.

[138] W. F. Bailey, M. D. England, M. J. Mealy, C. Thongsornkleeb, L. Teng; *Org. Lett.* **2000**, *2*, 489-491 *(10.1021/ol991342g)*.

[139] A. Carpita, R. Rossi, C. A. Veracini; *Tetrahedron* **1985**, *41*, 1919-1929 *(10.1016/S0040-4020(01)96555-X)*.

[140] N. Li; *Faming Zhuanli Shenqing* **2010**, 101885684.

[141] B.-L. Yin, Y.-L. Wu, J.-Q. Lai; *Eur. J. Org. Chem.* **2009**, *2009*, 2695-2699 *(10.1002/ejoc.200900129)*.

[142] J. Cardellach, C. Estopa, J. Font, M. Moreno-Mañas, R. M. Ortuño, F. Sanchez-Ferrando, S. Valle, L. Vilamajo; *Tetrahedron* **1982**, *38*, 2377-2394 *(10.1016/0040-4020(82)87016-6)*.

[143] S. Inack-Ngi, R. Rahmani, L. Commeiras, G. Chouraqui, J. Thibonnet, A. Duchêne, M. Abarbri, J.-L. Parrain; *Adv. Syn. Cat.* **2009**, *351*, 779-788 *(10.1002/adsc.200800757)*.

[144] F. G. Bordwell, H. E. Fried; *J. Org. Chem.* **1991**, *56*, 4218-4223 *(10.1021/jo00013a027)*.

[145] F. G. Bordwell; *Acc. Chem. Res.* **1988**, *21*, 456-463 *(10.1021/ar00156a004)*.

[146] J.-D. Charrier, S. J. Durrant, J. Studley, L. Lawes, P. Weber; *Bioorg. Med. Chem. Lett.* **2011**, *22*, 485-488 *(10.1016/j.bmcl.2011.10.102)*.

[147] M. M. Karelson, A. R. Katritzky, M. Szafran, M. C. Zerner; *J. Chem. Soc., Perkin Trans. 2* **1990**, 195-201 *(10.1039/p29900000195)*.

[148] G. Pilcher; *Pure Appl. Chem.* **1989**, *61*, 855-860 *(10.1351/pac198961050855)*.

[149] S. P. Fearnley; *Curr. Org. Chem.* **2004**, *8*, 1289-1337 *(10.2174/1385272043369971)*.

[150] H. Blaschke, E. Brunn, R. Huisgen, W. Mack; *Chem. Ber.* **1972**, *105*, 2841-2853 *(10.1002/cber.19721050909)*.

[151] S. P. Fearnley, C. Thongsornkleeb; *J. Org. Chem.* **2010**, *75*, 933-936 *(10.1021/jo902172r)*.

[152] F. C. Gaenzler, C. Guo, Y.-W. Zhang, M. E. Azab, M. A. I. Salem, D. P. Fan, M. B. Smith; *Tetrahedron* **2009**, *65*, 8781-8785 *(10.1016/j.tet.2009.08.066)*.

[153] J. H. Rothman; *J. Org. Chem.* **2008**, *74*, 925-928 *(10.1021/jo801910u)*.

[154] V. V. Vintonyak, B. Kunze, F. Sasse, M. E. Maier; *Chem.–Eur. J* **2008**, *14*, 11132-11140 *(10.1002/chem.200801398)*.

[155] J. F. Okonya, R. V. Hoffman, M. C. Johnson; *J. Org. Chem.* **2002**, *67*, 1102-1108 *(10.1021/jo010630z)*.

[156] L. M. Weinstock, R. B. Currie, A. V. Lovell; *Synth. Commun.* **1981**, *11*, 943-946 *(10.1080/00397918108065753)*.

[157] L. C. King, G. K. Ostrum; *J. Org. Chem.* **1964**, *29*, 3459-3461 *(10.1021/jo01035a003)*.

[158] R. V. Hoffman, M. C. Johnson, J. F. Okonya; *Tetrahedron Lett.* **1998**, *39*, 1283-1286 *(10.1016/S0040-4039(97)10854-1)*.

[159] L. A. Dakin, N. F. Langille, J. S. Panek; *J. Org. Chem.* **2002**, *67*, 6812-6815 *(10.1021/jo0204000)*.

[160] R. Sharma, J. Lee, S. Wang, G. W. A. Milne, N. E. Lewin, P. M. Blumberg, V. E. Marquez; *J. Med. Chem.* **1996**, *39*, 19-28 *(10.1021/jm950276v)*.

[161] J. Schröer, P. Welzel; *Tetrahedron* **1994**, *50*, 6839-6858 *(10.1016/S0040-4020(01)81337-5)*.

[162] J. Bálint, G. Egri, A. Kolbert, C. Dianóczky, E. Fogassy, L. Novák, L. Poppe; *Tetrahedron: Asymmetry* **1999**, *10*, 4017-4028 *(10.1016/S0957-4166(99)00415-2)*.

[163] M. E. Kopach, A. H. Fray, A. I. Meyers; *J. Am. Chem. Soc.* **1996**, *118*, 9876-9883 *(10.1021/ja961903o)*.

[164] J. B. J. Milbank, R. J. Stevenson, D. C. Ware, J. Y. C. Chang, M. Tercel, G. O. Ahn, W. R. Wilson, W. A. Denny; *J. Med. Chem.* **2009**, *52*, 6822-6834 *(10.1021/jm9008746)*.

[165] A. Al-Harrasi, F. Pfrengle, V. Prisyazhnyuk, S. Yekta, P. Kóóš, H.-U. Reissig; *Chem.–Eur. J* **2009**, *15*, 11632-11641 *(10.1002/chem.200900996)*.

[166] C. Bonini, L. Chiummiento, M. Funicello, P. Lupattelli, M. Pullez; *Eur. J. Org. Chem.* **2006**, *2006*, 80-83 *(10.1002/ejoc.200500737)*.

[167] J. Sam, J. L. Valentine; *J. Pharm. Sci.* **1969**, *58*, 1043-1054 *(10.1002/jps.2600580902)*.

[168] D. A. Fielder, F. W. Collins; *J. Nat. Prod.* **1995**, *58*, 456-458 *(10.1021/np50117a019)*.

[169] D. R. Chancellor, K. E. Davies, O. De Moor, C. R. Dorgan, P. D. Johnson, A. G. Lambert, D. Lawrence, C. Lecci, C. Maillol, P. J. Middleton, G. Nugent, S. v. D. Poignant, A. C. Potter, P. D. Price, R. J. Pye, R. Storer, J. M. Tinsley, R. van Well, R. Vickers, J. Vile, F. J. Wilkes, F. X. Wilson, S. P. Wren, G. M. Wynne; *J. Med. Chem.* **2011**, *54*, 3241-3250 *(10.1021/jm200135z)*.

[170] M. S. Singh, P. Singh, S. Singh; *Indian Journal of Chemistry - Section B* **2007**, *46*, 1666-1671.

[171] P. Carato, S. d. Yous, D. Sellier, J. H. Poupaert, N. Lebegue, P. Berthelot; *Tetrahedron* **2004**, *60*, 10321-10324 *(10.1016/j.tet.2004.08.070)*.

[172] E. J. Behrman; *J. Org. Chem.* **1992**, *57*, 2266-2270 *(10.1021/jo00034a016)*.

[173] E. Boyland, P. Sims; *J. Chem. Soc. (Res.)* **1954**, 980-985 *(10.1039/JR9540000980)*.

[174] E. Boyland, D. Manson, P. Sims; *J. Chem. Soc. (Res.)* **1953**, 3623-3628 *(10.1039/JR9530003623)*.

[175] J. E. Toth, J. Ray, W. J. Ehihardt; *Bioorg. Med. Chem. Lett.* **1992**, *2*, 1013-1014 *(10.1016/S0960-894X(00)80608-3)*.

[176] G.-W. Wang, T.-T. Yuan, X.-L. Wu; *J. Org. Chem.* **2008**, *73*, 4717-4720 *(10.1021/jo8003088)*.

[177] E. Falb, A. Nudelman, A. Hassner; *Synth. Commun.* **1993**, *23*, 2839-2844 *(10.1080/00397919308012605)*.

[178] S. Celanire, L. Quere, F. Denonne, L. Provins; *Compounds comprising a lactam or a lactam derivative moiety, processes for making them, and their uses (WO2007048595)*, **2007**.

[179] D.-R. Hou, J. H. Reibenspies, K. Burgess; *J. Org. Chem.* **2001**, *66*, 206-215 *(10.1021/jo001333h)*.

[180] J. R. Lakanen, A. E. Pegg, J. K. Coward; *J. Med. Chem.* **1995**, *38*, 2714-2727 *(10.1021/jm00014a023)*.

[181] M. Seki, K. Matsumoto; *Synthesis* **1999**, *1999*, 924-926 *(10.1055/s-1999-3495)*.

[182] *Kislorod i Serusoderzhashchie Geterotsikly (Moscow, Oct. 14-17 2003)*, Moscow, **2003**, pp. 319-323.

[183] H. Bravo, B. Weiss-Lopez; *Boletín de la Sociedad Chilena de Química* **1999**, *44*, 443-450 *(10.4067/S0366-16441999000400007)*.

[184] H. Zinner, H. Wigert; *Chem. Ber.* **1960**, *93*, 1331-1339 *(10.1002/cber.19600930614)*.

[185] S. E. Whitney, B. Rickborn; *J. Org. Chem.* **1991**, *56*, 3058-3063 *(10.1021/jo00009a025)*.

[186] A. Padwa, L. A. Cohen; *J. Org. Chem.* **1984**, *49*, 399-406 *(10.1021/jo00177a001)*.

[187] Y. Ni, R. M. Kassab, M. V. Chevliakov, J. Montgomery; *J. Am. Chem. Soc.* **2009**, *131*, 17714-17718 *(10.1021/ja907931u)*.

[188] J. Mauger, A. Nair, N. Ma, K. Bjergarde, B. Filoche-Romme, O. Angouillant-Boniface, S. Mignani; *1,4-dihydropyridine-fused heterocycles, process for preparing the same, use and compositions containing them (EP1746097)*, **2007**.

[189] T. Maekawa, N. Sakai, H. Tawada, K. Murase, M. Hazama, Y. Sugiyama, Y. Momose; *Chem. Pharm. Bull.* **2003**, *51*, 565-573 *(10.1248/cpb.51.565)*.

[190] M. Kameda, K. Kobayashi, H. Ito, H. Miyazoe, T. Tsujino, C. Nakama, H. Kawamoto, M. Ando, S. Ito, T. Suzuki, T. Kanno, T. Tanaka, Y. Tahara, T. Tani, S. Tanaka, S. Tokita, N. Sato; *Bioorg. Med. Chem. Lett.* **2009**, *19*, 4325-4329 *(10.1016/j.bmcl.2009.05.069)*.

[191] R. Gompper, F. Effenberger; *Angew. Chem.* **1958**, *70*, 628-628 *(10.1002/ange.19580702005)*.

[192] O. P. Shvaika, G. P. Klimisha; *Chem Heterocycl Compd* **1966**, *2*, 517-520 *(10.1007/bf00477508)*.

[193] P. Ting, R. Aslanian, M. A. Caplen, J. Cao, D. W.-S. Kim, H. Kim, R. Kuang, J. F. Lee, J. H. Schwerdt, H. Wu, G. Zhou, N. Zorn; *Inhibitors Of Diacylglycerol Acyltransferase (WO2010059606)*, **2010**.

[194] R. J. Perner, J. R. Koenig, S. DiDomenico, A. Gomtsyan, R. G. Schmidt, C.-H. Lee, M. C. Hsu, H. A. McDonald, D. M. Gauvin, S. Joshi, T. M. Turner, R. M. Reilly, P. R. Kym, M. E. Kort; *Bioorg. Med. Chem.* **2010**, *18*, 4821-4829 *(10.1016/j.bmc.2010.04.099)*.

[195] E. Ferrer Flegeau, M. E. Popkin, M. F. Greaney; *J. Org. Chem.* **2008**, *73*, 3303-3306 *(10.1021/jo800121y)*.

[196] N. F. Langille, L. A. Dakin, J. S. Panek; *Org. Lett.* **2002**, *4*, 2485-2488 *(10.1021/ol026099r)*.

[197] D. E. Paterson, F. K. Griffin, M.-L. Alcaraz, R. J. K. Taylor; *Eur. J. Org. Chem.* **2002**, 1323-1336 *(10.1002/1099-0690(200204)2002:7<1323::aid-ejoc1323>3.0.co;2-8)*.

[198] S. A. Nunez, K. Yeung, N. S. Fox, S. T. Phillips; *J. Org. Chem.* **2011**, *76*, 10099-10113 *(10.1021/jo2018763)*.

[199] S. V. Jadhav, A. Bandyopadhyay, S. N. Benke, S. M. Mali, H. N. Gopi; *Org. Biomol. Chem.* **2011**, *9*, 4182-4187 *(10.1039/C0OB01226B)*.

[200] K. Maekawa, K. Kubo, T. Igarashi, T. Sakurai; *Tetrahedron* **2005**, *61*, 11211-11224 *(10.1016/j.tet.2005.09.042)*.

[201] B. v. Platen; in *Modern Very High Pressure Techniques*, (Ed. R. Wentorf), Butterworths, Washington, **1962**, pp. 118-136.

Part III

Experimental

Table of Contents: Part III

III. EXPERIMENTAL

1. General Procedures

Reagents and Procedures: Dry tetrahydrofuran (THF), toluene and dichloromethane (DCM) were obtained by passing commercially available pre-dried formulations through activated alumina columns under argon. Dry DMF / DMAc were used as obtained from Sigma and Fluka (bottled under argon and over molecular sieves in septum-sealed bottles). Other dry solvents, unless noted otherwise, were obtained by standing commercial synthetic or HPLC grade solvents on thermally-activated molecular sieves (MS; 3Å or 4Å as appropriate, sieve activation by 24h heating at 315 °C). 'Rigorously dried' solvents were dried by at least two sequential MS dryings of at least one hour while stirring, unless otherwise stated. Moderately dried amines (piperidine, piperazine, triethylamine, pyridine) were obtained by standing on KOH pellets for at least 2 hours (piperazine was dissolved in warm pyridine). Unless stated otherwise, all reactions were performed with unpurified, undried, non-degassed solvents or analytical grade reagents, used as obtained, under closed air atmosphere. Petroleum ether was the 40-60°C fraction. Column chromatography was performed on Merck silica gel Si-60 (40-63 µm). Procedures and yields are unoptimized. Yields refer to isolated chromatographically and spectroscopically pure materials, unless otherwise stated.

Products already found in the literature* have their names *typeset in grey, not bolded, and are denoted with an asterisk*. All other products are to the best of my knowledge, original species.

Thin-layer chromatography (TLC) was run by default on 0.25 mm Merck silica gel plates (60, F-254), unless noted as Al_2O_3, which was run on the corresponding alumina plates. UV light (254 nm by default, 365 nm for fluorescent compounds when specified) was used as a visualising agent, and SiO_2-dispersed iodine (I_2) or standard TLC dips based on p-anisaldehyde (anis), Hanessian's cerium ammonium molybdate formulation (Han), basic $KMnO_4$ ($KMnO_4$), phosphomolybdic acid (PMA), Dragendorff's reagent (Drag), vanillin (van) and ninhydrin (nin) followed by heating where necessary were used as developing agents.

NMR: Standard NMR characterisation was by ^1H and ^{13}C 1D NMR spectra. Known compounds were checked against literature data and spectral analysis not given unless necessary. New compounds of special interest or unusual structure, and key intermediates, have NMR spectra fully assigned where possible, checked against 2D NMR (HSQC, COSY, HMBC). Selected spectra of important compounds are given in the section Key NMR Spectra (pages 265 ff.) All spectra were acquired on a Bruker DPX 200 (200 MHz & 50 MHz for ^1H and ^{13}C respectively) unless stated otherwise, in which case they were acquired on a Bruker AVANCE 500 (500 MHz & 125 MHz for ^1H and ^{13}C respectively) as indicated, at 300K. Where not indicated otherwise, the NMR solvent was $CDCl_3$.

Chemical shifts (δ) are reported in ppm with reference to residual non-perdeuterated solvent as an internal reference.[1] The following peak descriptions are used: singlet (s), doublet (d), triplet (t), quartet (q), multiplet (m), broad (br), quintet (quin), sextet (sext); apparent multiplicities (resolved by 2D experiments or determined by complete spectral assignment) are denoted by a tilde, eg. "doublet of doublets, appears as a triplet with apparent coupling constant $J = 3$ Hz" is denoted (dd~t, 3 Hz). Benzyloxy PhC$\underline{H_2}$OR hydrogens were sometimes seen as (dd) by ^{1}H-NMR, with $^2J \sim 11$ Hz and a separation of up to 200 Hz ("J_{eff}") between the individual hydrogens. Where of interest, such benzyloxy CH_2 groups are reported as (dd, 2H) with both J_{eff} and ^2J values to ensure the greatest portability of the NMR analysis to different field-strength machines, as only J_{eff} varied during acquisitions on different spectrometers. Their H-C-H couplings in $^1H^1H$ COSY (seen without exception) have however been omitted for clarity.

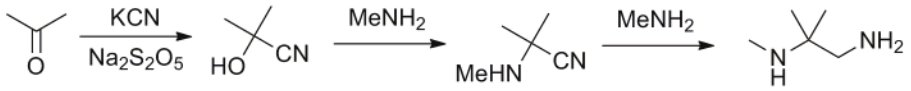
Mass Spectra: Unit mass measurements were performed on an AGILENT 1100 SL coupled LC-MS system with direct injection of the sample in ESI mode (DIMS, or simply, MS), with binary solvent mixtures of water-acetonitrile, with the water containing sodium/ammonium formate. Unless stated otherwise (because important for repeatability, chiefly for low-MW amines), solvent was 80:20 MeCN:H_2O, and two fragmentation voltages of 30 and 130 eV were used. Ion peaks from (positive/negative mode) are reported as (+/-) with units Th (m/z), thus 'MS(+): 328 @ 100, 227 @ 80' indicates ESI direct injection giving two positive ion peaks at m/z= 328 and 227 Th, with the peak at 227 Th being 80% of the height of the peak at 328 Th (isotopic peak patterns were sometimes useful to confirm molecular identity on DIMS spectra). Unless stated otherwise, all reported peaks in (+)-mode were $[MH]^+$ peaks, and all observed peaks in (-)-mode were $[M-H]^-$ peaks. HRMS was carried out by the Service Central d'Analyse du CNRS, Solaize, France, and is reported similarly. Compounds are commonly drawn with their exact mass in Da (EM) for easy comparison.

IR: Spectra were recorded on a Nexus Nicolet-Thermo-Fisher machine using an SiC source, KBr/Ge plates, resolution of 4 cm^{-1}, Happ-Genzel apodisation, a spectral window of 4000-500 cm^{-1} and 256 scans per spectrum; peaks are listed with units of cm^{-1} and with an arbitrary definition of peak strength based on the ratio R_p of the given peak height to the highest peak: strong [s] ($\sim R_p > 0.5$), medium [m] ($0.2 < \sim R_p < 0.5$), or weak [w] ($\sim R_p < 0.2$), as well as designations broad [br] (the peak is not sharply defined), or complex [c] (peak range contains multiple peaks/fine features). Only selected, characteristic or unexplained peaks are presented. In-house IR was prepared and acquired by Martine Simon; external IR analysis was carried out by the Service Central d'Analyse du CNRS, Solaize, France.

Experimental: Part I (Fluorogenic Probes)

2. Compounds of Part I

2.1. Spacers

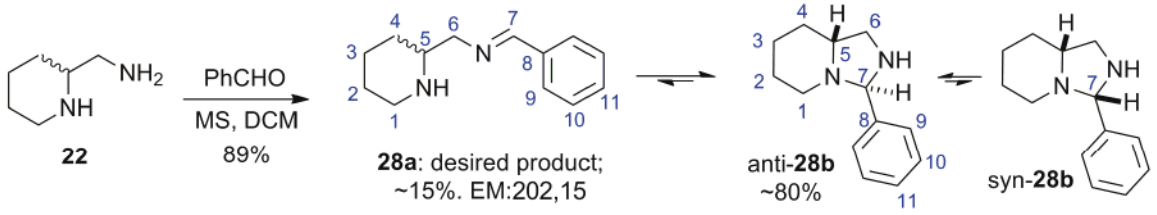


21: EM:102,12

N^2-2-dimethylpropane-1,2-diamine (21)*

The title compound was synthesised by Sean McKeon and Monica Vargas-Sanchez in three steps adapted from literature procedures. Following a procedure adapted from Faghihi *et al.*[2], acetone (5.8 g, 100 mmol) was added to a stirred aqueous solution of sodium metabisulfite (11 g, 58 mmol, 20 mL). An aqueous solution of KCN (6 g, 92 mmol, 20 mL) was added and upon completion of the reaction, the phases were separated and the upper phase was dried on Na_2SO_4 and filtered to give **acetone cyanohydrin** as a clear liquid (5.402 g, 69 %). Following a procedure adapted from Exner *et al.*[3], to a 40 wt% aqueous solution of methylamine (2.8 g, 36 mmol) cooled in an icebath was slowly added the acetone cyanohydrin (1 g, 12 mmol) such that the internal temperature did not exceed 15°C. After the addition was finished, the reaction mixture was stirred for a further 90 min, then extracted with Et_2O (3×10 mL). The combined organic phases were dried on Na_2SO_4, filtered, and concentrated under reduced pressure to give **2-methylamino-2-methyl-propionitrile** as a colourless liquid (0.700 g, 59 %). Following a procedure adapted from Chong *et al.*[4], 2-methylamino-2-methyl-propionitrile (1.960 g, 20 mmol) was added dropwise to a suspension of $LiAlH_4$ (1.520 g, 40 mmol) in Et_2O (2 mL) under argon. The reaction was stirred for 3 hours at RT, diluted with Et_2O (30 mL) and cooled in an icebath. Sat. K_2CO_3 was added dropwise until evolution of hydrogen ceased. Solid Na_2SO_4 was then added to the reaction mixture which was stirred for 10 min before the salts were filtered off on Celite and washed with large quantities of Et_2O. The filtrate was concentrated under reduced pressure to give **21** as a clear, somewhat volatile liquid (2.001 g, 98 %) that may be more conveniently stored, cold, as a solution in Et_2O. MS (ESI, positive mode): m/z = 103.1: $[MH]^+$. ^{1}H-NMR (200 MHz, $CDCl_3$): δ = 2.40 (s, 2H, CH_2), 2.14 (s, 3H, NHC$\underline{H}_3$), 1.80 (s, br, 3H, NH + NH_2), 0.87 (s, 6H, 2×CH_3) ppm. ^{13}C-NMR (50 MHz, $CDCl_3$): δ = 53.0 ($\underline{C}(Me)_2$), 50.2 (CH_2), 28.3 ($NHCH_3$), 23.0 (2×CH_3) ppm.

2.2. Protected spacer strategy



E,Z-N-benzylidene-1-(piperidin-2-yl)methanamine (28a) ↔ syn,anti-3-phenyloctahydroimidazo[1,5-a]pyridine (28b)

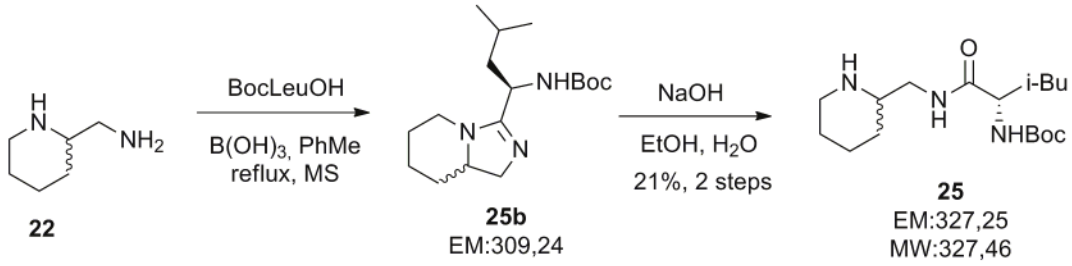
To a solution of commercial 2-aminomethylpiperidine **22** (1.46 g, 12.8 mmol) in dry DCM (4 mL) with activated 4Å molecular sieves (3 g) was added benzaldehyde (1.20 g, 11.3 mmol). The mixture was stirred for 1 h at RT, then the liquid phase filtered through Celite and the sieves rinsed twice with 1 mL DCM. The combined filtrates were evaporated under reduced pressure to yield a yellow oil (2.05 g, 89 %), which was a NMR-spectroscopically pure mixture of compounds ***E*-28a, *Z*-28a, *syn*-28b** and ***anti*-28b**. HRMS (TOF MS ESI$^+$) calcd for $[C_{13}H_{19}N_2]^+ = [MH^+]$: m/z 203.1543, found 203.1536. The ratio of compounds **28b** to compounds **28a** in a range of solvents was calculated by integration of the ^{1}H-NMR spectra, to see if the desired isomer **28a** could be favoured by an appropriate choice of the solvent, however no appropriate solvent was found.

Solvent	Ratio **28b**:**28a**
$CDCl_3$	5:1
CD_2Cl_2	6:1
acetone-d_6	13:2
DMSO-d_6	17:2
CD_3OD	13:2
D_2O	9:2

The following NMR characterisation on the mixture of isomers was performed in $CDCl_3$ (spectra reproduced on page 265 ff.); spectra were completely assigned for the majority species ***E*-28a** and ***syn*-28b**; at least one characteristic peak could be assigned for each of the minority species ***Z*-28a** and ***anti*-28b**. HSQC supports the spectral assignments. ***syn*-28b** (82 %): ^{1}H-NMR (500 MHz): δ = 7.38 (d, 7.8 Hz, 2H, 2×H9), 7.25 (~t, ~7.4Hz, 2H, 2×H10), 7.22-7.20 (m, 1H, H11), 3.91 (s, 1H, H7), 3.00 (dd, 9.2 Hz & 6.5 Hz, 1H, H6), 2.73 (~t, ~9.5 Hz, 1H, H6′), 2.66-2.65 (m, 1H, H1), 2.22-2.21 (m, 1H, H5), 1.85 (td, 11.2 Hz & 3.1 Hz, 1H, H1′), 1.79-1.70 (m, 2H, H3 + H4), 1.52-1.40 (m, 2H, 2×H2), 1.30-1.26 (m, 1H, H4'), 1.22-1.16 (dt, 12.8 Hz & 3.6 Hz, 1H, H3′) ppm. ^{13}C-NMR (125 MHz): δ = 141.0 (C8), 128.5 (2×C10), 128.3 (C11), 127.9 (2×C9), 82.7 (C7), 63.8 (C5), 50.6 (C6), 48.8 (C1), 29.1 (C4), 25.2 (C2), 24.2 (C3) ppm. ***E*-28a** (15 %): ^{1}H-NMR (500 MHz): δ = 8.22 (s, 1H, H7), 7.65-7.64 (m, 2H, 2×H9), 7.31-7.28 (m, 3H, 2×H10 + H11), 3.61 (dd, 11.8 Hz & 3.2 Hz, 1H, H6), 3.34 (dd, 11.6 & 8.5 Hz, 1H, H6′), 2.96-2.95 (m, 1H, H1), 2.80-2.79 (m, 1H, H5), 2.55 (td, 11.9 Hz & 2.3 Hz, H1′), 1.61 (d, 12.5 Hz, 1H, H4), 1.53-1.50 (m, 1H, H2), 1.40-1.33 (m, 1H, H2′), 1.31-1.26 (m, 2H, 2×H3), 1.18-1.16 (m, 1H, H4′) ppm. ^{13}C-NMR (125 MHz): δ = 162.2 (C7), 136.1 (C8), 130.7

(C11), 128.6 (2×C9), 128.2 (2×C10), 67.8 (C6), 56.9 (C5), 46.8 (C1), 30.8 (C4), 26.3 (C2), 24.7 (C3) ppm. The weak signals from the minor isomers were mostly overlapped with the major isomers preventing full spectral assignment. However characteristic peaks could be found in both spectra to support integration assignment of isomer proportions: ***anti*-28b** (2.5 %): ^{1}H-NMR (500 MHz): δ = 4.89 (s, 1H, H7) ppm. ^{13}C-NMR (125 MHz): δ = 141.5 (C8), 77.9 (C7) ppm; ***Z*-28a** (0.5 %): ^{1}H-NMR (500 MHz): δ = 8.00 (d, 7.9 Hz, 2H, 2×H9) ppm. ^{13}C-NMR (125 MHz): δ = 165.0 (C7), 129.6 (2×C9) ppm.

2.3. Substrate-spacer pairs



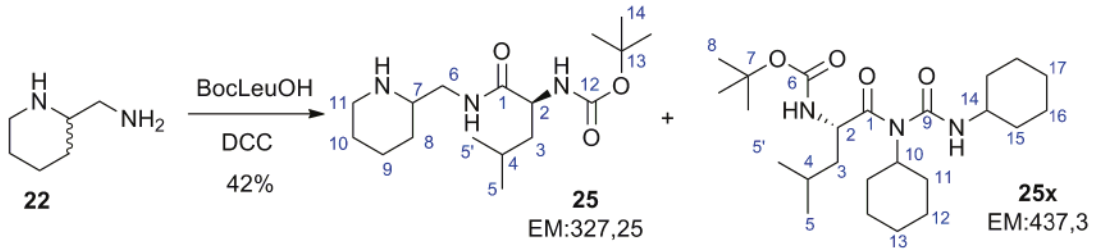
22

25b
EM:309,24

25
EM:327,25
MW:327,46

N′-tert-butoxycarbonyl-L-leucyl-(piperidin-2-ylmethyl)amide (25)

B(OH)$_3$ method: Following a procedure similar to that of Mylavarapu *et al.*[5], BocLeuOH (2.08 g, 8.99 mmol) was added to a flask containing dry toluene (60 mL) and activated 4Å molecular sieves (10.6 g). The mixture was heated to 60 °C until the acid had dissolved, then boric acid (67 mg, 1.1 mmol, 0.1 eq) was added, followed by 2-aminomethylpiperidine **22** (1.027 g, 9.00 mmol). The solution yellows and is stirred for 16h at 113 °C under a condenser with a closed air atmosphere. The mixture was then filtered, washed with sat. aq. Na_2CO_3 to eliminate the unreacted acid, and evaporated under reduced pressure yielding a crude red oil containing amidine **25b** (MS+ : m/z = 310 @ 100, 311 @ 20 : $[MH]^+$) as well as residual **25** (TLC, 3:3:1 EA:Cy:MeOH, ninhydrin). To the crude oil was added EtOH (30 mL), H_2O (20 mL) and NaOH (6 g, 250 mmol), and the solution was stirred overnight at RT. The basic solution was extracted with DCM (3×30 mL) and the combined organic extracts rinsed with brine (20 mL), dried over Na_2SO_4, filtered and evaporated to yield a crude yellow oil which was filtered through a silica pad using 1:1:0 → 1:1:1 EA:Cy:MeOH to give **25** as a yellow oil (630 mg, 1.92 mmol, 21 %), identical by NMR to that produced by the **DCC Method**.



DCC Method: Under closed air atmosphere, BocLeuOH (4.385 g, 19.0 mmol) was dissolved in dry DCM (50 mL) before cooling in an icebath. DCC (3.95 g, 19.1 mmol) was added and the mixture stirred for 5 minutes at 0 °C. 2-aminomethylpiperidine **22** (2.11 g, 18.5 mmol) was added and the icebath was removed. After 24 hours stirring the white solid DCU was filtered off on a porosity 3 glass frit and rinsed colourless with DCM. The yellow DCM phase was washed with sat. Na_2CO_3, then 0.5M phosphate buffer at pH>10, then dried with brine, Na_2SO_4 and evaporated to yield a crude yellow oil (6.6 g). The oil was left at 4 °C overnight, and the remaining DCU crystallised and was filtered off as previously. Crude NMR then indicates a mixture of **25** with N-acylurea byproduct **25x**. Column chromatography with a gradient of 1:1:0 → 1:1:1 EA:Cy:MeOH gave **25** (2.703 g, 42 %) as a colourless, crystalline solid. Compound **25** has R_f = 0.0 in 1:1 Cy:EA, R_f = 0.20 in 4:4:1 Cy:EA:MeOH, and is stained by $KMnO_4$ as well as ninhydrin (pink charring purple). HRMS (TOF MS ESI$^+$) calcd for $[C_{17}H_{34}N_3O_3]^+ = [MH^+]$: m/z 328.2595, found 328.2589. NMR indicates the presence of two rotamers or conformers in approximate ratio 2:1 for each of the two diastereoisomers (about C7), thus complicating the spectral analysis. HSQC, COSY and jmod spectra support the spectral assignment. **25:** ^{1}H-NMR (200 MHz): 7.09-6.81 ((s, br + s, br), 1H, NH_{amide}), 5.70-5.34 ((s, br + d, br, 7.6 Hz), 1H, $NH_{carbamate}$), 4.08-3.98 (m, 1H, H2), 3.31-3.15 (m, 1H, H6), 3.11-2.90 (m, 2H, H6′ + H7), 2.70-2.44 (m, 2H, 2×H11), 2.14 (s, br, 1H, shift varies strongly between spectra, NH_{amine}), 1.76-1.41 (m, 6H, 2×H3 + H4 + H8 + H9 + H10), 1.33 (s, 9H, 9×H14), 1.32-1.00 (m, 3H, H8′ + H9′ + H10′), 0.88 (d, 2.1Hz, 3H, 3×H5), 0.87 (d, 2.4Hz, 3H, 3×H5′) ppm. ^{13}C-NMR (50 MHz): Diastereoisomer 1: δ = 173.0 (C1), 155.7 (C12), 79.6 (C13), 55.8 (C7), 53.2 (C2), 46.5 (C11), 45.1 (C6), 41.4 (C3), 30.1 (C8), 28.3 (3×C14), 26.3 (C10), 24.7 (C4), 24.2 (C9), 22.9 (C5), 22.0 (C5′). Diastereoisomer 2: δ = 173.1 (C1), 155.8 (C12), 79.7 (C13), 55.5 (C7), 53.2 (C2), 46.5 (C11), 45.1 (C6), 41.3 (C3), 30.2 (C8), 28.3 (3×C14), 26.2 (C10), 24.7 (C4), 24.2 (C9), 22.9 (C5), 22.0 (C5′) ppm.

DCC Method + HOBt: HOBt monohydrate (426 mg, 2.78 mmol, 1 eq) was added to the mixture of BocLeuOH (608 mg, 2.62 mmol) and DCC (597 mg, 2.89 mmol) at 0 °C in DCM (15 mL) and the mixture stirred for 10 min before 2-aminomethylpiperidine **22** (350 mg, 3.07 mmol) was added and left stirring at RT for 24h. The reaction was worked up as for the **DCC Method** to give a crude yellow foam (830 mg). Two column chromatographic separations, first on DCM:MeOH gradient then on EA:Cy:MeOH gradient, did not separate the product **25** (173 mg, 0.53 mmol, 20 %) from residual HOBt (43 mg, 0.32 mmol). The mixture was used as is for the coupling to **31a** in the next step, as the

products of that reaction could easily be separated. TLC and NMR confirmed that the product was otherwise identical to that produced by the **DCC Method**.

1-(N'-tert-butoxycarbonyl-L-leucyl)-1,3-dicyclohexylurea (25x)

25x was isolated as a byproduct during DCC-mediated syntheses of **25** and **26** in variable yields, as a white powder, highly soluble in chlorinated solvent. **25x** has R_f = 0.91 in 195:5 DCM:MeOH, and is stained by $KMnO_4$ as well as ninhydrin (pink charring orange). MS (ESI, positive mode) : m/z 438,3@20, 439.3.3@5 : $[MH]^+$. HRMS (TOF MS ESI$^+$) calcd for $[C_{24}H_{43}N_3O_4Na]^+ = [MNa^+]$: *m/z* 460.3146, found 460.3125. ^{1}H-NMR (200 MHz): δ = 7.66 (d, br, 0.9H, $NH_{carbamate}$), 4.92 (d, 1H, H10), 4.39 (q, 1H, H14), 4.12-4.10 (m, 1H, H2), 3.67-3.64 (m, 1H, NH_{urea}), 1.40 (s, 9H, 9×H8), 2.00-1.42 + 1.36-1.05 (m, 23H, H2 + 2×H3 + 4×H11 + 4×H12 + 2×H13 + 4×H15 + 4×H16 + 2×H17), 0.88 (d+d, 6H, 3×H5 + 3×H5′) ppm. ^{13}C-NMR (50 MHz): δ = 172.1 (C1), 156.5 (C6), 153.6 (C9), 80.3 (C7), 54.8 (C10), 51.8 (C2), 50.1 (C14), 42.1 (C3), 33.5 (C15), 32.6 (C15′), 31.8 (C11), 31.6 (C11′), 29.3 (C3), 28.3 (3×C8), 26.1 (C16), 26.0 (C16′), 25.5 (C17), 25.4 (C13), 24.8 (C4), 24.6 (2×C12), 23.0 (C5), 21.9 (C5′) ppm.

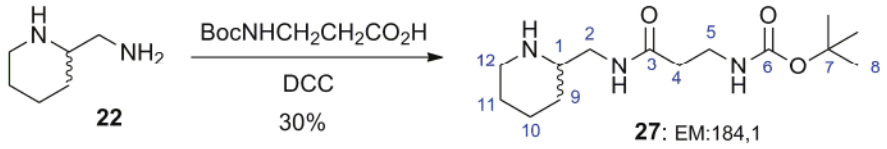
N′-tert-butoxycarbonyl-L-leucyl-(2-methyl-2-(methylamino)propyl)amide (26)

DCC Method + acid/base extractions, 1: Similar to the **DCC Method**, BocLeuOH (1.37 g, 5.92 mmol) was reacted with **21** (1.50 g of approx. 47 wt % solution in diethyl ether, ~6.9 mmol, >1 eq). However, after removing DCU by filtration, the amines were then extracted with aqueous HCl (2M, 3×30 mL); evaporation of the organic phase gave pure N-acylurea byproduct **25x** (1.60 g, 3.65 mmol). The clear, acidic, combined aqueous phases were then basified with KOH to pH > 12 and extracted with DCM (6×20 mL), and the combined organic phases were dried on Na_2SO_4, filtered, and evaporated under reduced pressure to yield a crude clear oil (0.54 g) containing product **26** (43 mol %, 52 wt %, 280 mg, 0.88 mmol) and a byproduct due to the loss of the Boc group (**26y**, 260 mg, 1.2 mmol). This mixture was filtered through a short silica column using a gradient of EA:Cy:MeOH (10:7:0→10:7:3) with which **26y** was immobile, and **26** was obtained as a colourless oil which solidified on standing (235 mg, 0.74 mmol, 13 %). **26** has R_f = 0.10 in 14:6:7 EA:Cy:MeOH and is stained by $KMnO_4$ as well as ninhydrin (purple charring black). MS(+) : m/z 316.3@100, 317.3@19, 318.3@3 $[MH]^+$. HRMS (TOF MS ESI$^+$) calcd for $[C_{16}H_{34}N_3O_3]^+ = [MH^+]$: *m/z* 316.2595, found 316.2599. HSQC, COSY and jmod spectra support the spectral assignment. ^{1}H-NMR

(500 MHz): 6.80 (s, br, 1H, NH_{amide}), 5.04 (~d, br, 7.7 Hz, 0.8H, $NH_{carbamate}$), 4.12-4.08 (m, 1H, H2), 3.26 (dd, 14.0 & 5.9 Hz, 1H, H9), 3.17 (dd, 13.3 & 4.9 Hz, 1H, H9′), 2.34 (s, 3H, 3×H12), 2.14 (s, br, 1H, NH_{amine}), 1.72-1.65 (m, 2H, H4 + H3), 1.54-1.49 (m, 1H, H3′), 1.46 (s, 9H, 9×H8), 1.11 (s, 6H, 6×H11), 0.96 (d, 3H, 3.7 Hz, 3×H5),0.95 (d, 3H, 3.6 Hz, 3×H5′) ppm. ^{13}C-NMR (125 MHz): δ = 172.9 (C1), 155.8 (C6), 80.0 (C7), 53.7 (C10), 53.5 (C2), 46.6 (C9), 41.4 (C3), 28.4 (3×C8), 28.3 (C12), 24.9 (C4), 24.6 (2×C11), 23.0 (C5), 22.2 (C5′) ppm.

L-leucyl-(2-methyl-2-(methylamino)propyl)amide (26y)

This compound could be completely identified in the NMR spectra of crude **26**, but was not isolated after chromatography. The compound has R_f = 0.0 in 14:6:7 EA:Cy:MeOH and is stained by $KMnO_4$ as well as ninhydrin (red charring purple). HSQC, COSY and jmod spectra support the spectral assignment. ^{1}H-NMR (500 MHz): δ = 7.43 (s, br, 0.9H, NH_{amide}), 3.27 (dd, 9.5 & 4.3 Hz, 1H, H2), 3.02 (s, 2H, 2×H6), 2.17 (s, 3H, 3×H9), 1.62-1.50 (m, 2H, H3 + H4), 1.29-1.20 (m, 1H, H3′), 0.92 (s, 6H, 6×H8), 0.82 (d, 3H, 6.4Hz, 3×H5), 0.80 (d, 3H, 6.4Hz, 3×H5′) ppm. ^{13}C-NMR (125 MHz): 175.9 (C1), 53.3 (C7), 52.7 (C2), 46.4 (C6), 44.3 (C3), 28.6 (C9), 24.8 (C4), 24.7 (2×C8), 23.4 (C5), 21.4 (C5′) ppm.



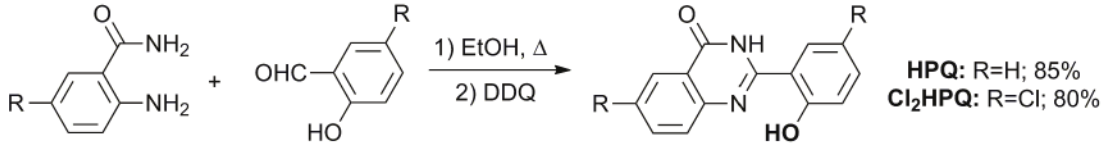
N-(Boc)-3-aminopropyl-(N'-methyl(piperidin-2-yl))amide (27)

27 could be obtained in 18% yield by the **B(OH)$_3$ Method**, or in 30% yield by a modified **DCC Method** using KH_2PO_4 extraction as follows:

DCC Method + acid/base extractions, 2: *N*-(Boc)-3-aminopropanoic acid (3.48 g,18.4 mmol) was reacted with DCC (3.79 g,18.4 mmol) in dry DCM (50 mL) in a salt-ice bath at -10 °C for five minutes. 2-aminomethylpiperidine **22** (2.08 g,18.2 mmol) was injected and the reaction left stirring at RT for 16h, before filtering off the DCU, washing with sat. aq. Na_2CO_3 (20 mL) and 1M NaOH (20 mL), back-extracting the basic aqueous phases with DCM (3×20 mL), then washing the combined organic phases with sat. aq. Na_2CO_3 (20 mL) and sat. aq. NaCl (20 mL) and drying on Na_2SO_4. Filtration and evaporation under reduced pressure gave a crude yellow oil (6.5 g) which was dissolved in DCM (15 mL). The amines were extracted into acidified potassium dihydrogenphosphate buffer (0.75M, pH~2.5, 2×25 mL) and the aqueous phases were washed with DCM and then basified with aqueous NaOH to pH~12. The basic aqueous phase was now extracted with DCM (2×30 mL) and the combined organic phases washed with sat. aq. NaCl (25 mL) and dried on $MgSO_4$, filtered and evaporated under reduced pressure to yield 2.29 g clear yellow oil, which was filtered through a silica

pad eluting with 1:1:0→1:1:1 EA:Cy:MeOH to give **27** as a white powder (1.555 g, 5.44 mmol, 30 %). This milder method increased the isolated yield relative to that seen for **26**, presumably by reducing Boc deprotection. **27** has R_f = 0 in 1:1 Cy:EA, R_f = 0.07 in 80:20 DCM:MeOH, and is stained by $KMnO_4$ as well as ninhydrin (red charring black). HRMS (TOF MS ESI^+) calcd for $[C_{14}H_{28}N_3O_3]^+ = [MH^+]$: m/z 286.2125, found 286.2123. ^{1}H-NMR (200 MHz): δ = 6.66 (s, br, 1H, NH_{amide}), 5.69 (~t, ~5.9 Hz, 1H, $NH_{carbamate}$), 3.36-3.20 (m, 3H, H1 + 2×H5), 3.07-2.94 (m, 2H, 2×H2), 2.60-2.46 (m, 2H, 2×H12), 2.38-2.26 (m, 3H, 2×H4 + NH_{amine}), 1.80-1.40 (m, 3H, H9 + H10 + H11), 1.36 (s, 9H, 9×H8), 1.32-0.92 (m, 3H, H9' + H10' + H11') ppm. ^{13}C-NMR (50 MHz): δ = 172.0 (C3), 156.2 (C6), 79.1 (C7), 56.0 (C1), 46.6 (C12), 45.1 (C2), 36.9 (C5), 36.4 (C4), 30.3 (C9), 28.4 (3×C8), 26.3 (C11), 24.3 (C10) ppm.

2.4. Fluorophores



2-(2-hydroxyphenyl)quinazolin-4(3H)-one (HPQ)*

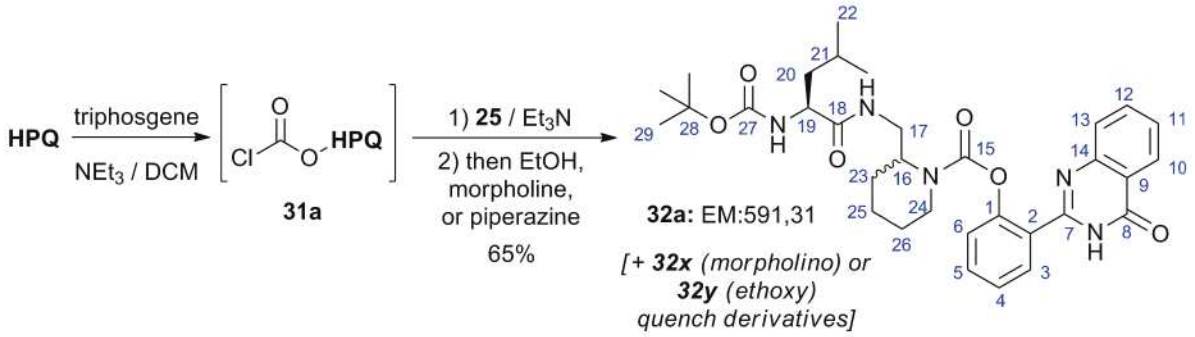
The following procedure was adapted from Waibel[6]. To a red solution of anthranilamide (1.60 g, 11.75 mmol) in absolute EtOH (30 mL) was added salicylaldehyde (1.44 g, 11.8 mmol) giving a yellow solution and a precipitate. The reaction was heated to 80 °C under a condenser open to the air for 30 min, then *p*-TsOH monohydrate (43 mg, 0.22 mmol) was added and the precipitate dissolved. Heating was continued for 1 h, then the solution was cooled to room temperature and DDQ (2.70 g, 11.8 mmol) was added. The mixture was stirred at RT, open to the air, overnight. The solid was filtered on a porosity 4 glass frit filter, air dried, rinsed twice with diethyl ether then dried at reduced pressure giving **HPQ** (2.4 g, 10 mmol, 86 %) as a beige powder showing strong green fluorescence under a 365 nm UV lamp. **HPQ** has R_f = 0.56 in 1:1 Cy:EA and is stained by $KMnO_4$ but not by ninhydrin. NMR in d_6-DMSO matched data of Baghbanzadeh *et al.*[7] (their compound 1b).

2-(4-chloro-2-hydroxyphenyl)-6-chloroquinazolin-4(3H)-one (diClHPQ)*

Cl_2HPQ was synthesised from 4-chloroanthranilamide (2.45 g, 14.4 mmol) and 4-chlorosalicylaldehyde (2.25 g, 14.4 mmol) similarly to **HPQ**. Due to its greater insolubility, the solid was washed exhaustively with 95% EtOH, *i*PrOH, iPr_2O, and finally pentane to obtain **diClHPQ** (3.53 g, 80 %) as a cream powder. ^{1}H-NMR (200 MHz, d_6-DMSO): δ = 13.4 (s, br, 1H), 12.65 (s, br, 1H), 8.27 (s, 1H), 8.06 (s, 1H), 7.84 (m, 2H), 7.47 (1H, d, 7.5 Hz), 7.01 (d, 1H, 8.6 Hz).

2.5. Phosgenation coupling reactions

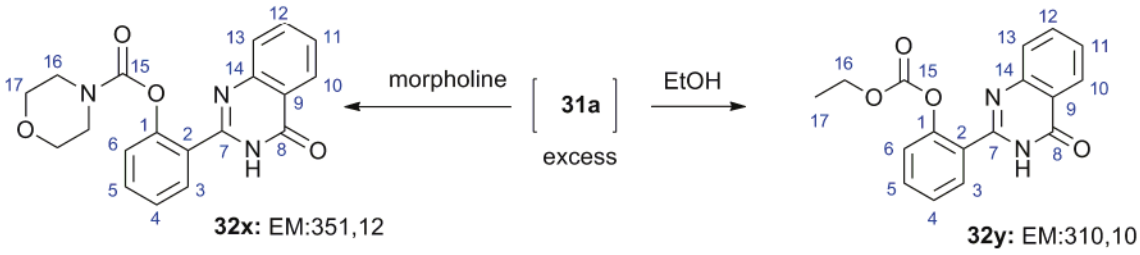
Caution : phosgene, b.p. 8 °C, is highly toxic and corrosive and can react violently with water or other nucleophiles especially if the reaction is in homogenous media. Reactions were kept cold to avoid boil-off of phosgene. Excess phosgene was caught apparently quantitatively during evaporation in a primary liquid nitrogen trap (a backup trap was employed but always found empty); it was destroyed when still cold by its dropwise addition to a vigorously stirred, cold mixture of 2-aminoethanol or piperidine (5 mL) and ethanol (30 mL) in dichloromethane (100 mL) in a well-ventilated hood.



2-(4-oxo-3,4-dihydroquinazolin-2-yl)phenyl 2-((N-(tert-butoxycarbonyl)-L-leucylamido)methyl)piperidine-1-carboxylate (32a)

To **HPQ** (1.068 g, 4.48 mmol, 1 eq) in a two-neck flask under argon was added CaH_2-distilled triethylamine stored on KOH (925 mg, 9.12 mmol, 2 eq), then dry DCM (10 mL), and the mixture cooled in an icebath. A solution of triphosgene (1.702 g, equiv. to 17.2 mmol phosgene, 4 eq) in dry DCM (8 mL) was injected with vigorous stirring, and the orange solution was stirred for 1 h at 0 °C, before all volatile components were removed at high vacuum with double liquid nitrogen traps (and a short-stem primary trap so as not to clog the system) to leave a beige solid containing HPQ chloroformate **31a** (not isolated). A solution of amine **25** (275 mg, 0.90 mmol, 0.2 eq) and CaH_2-distilled triethylamine stored on KOH (1.43 g, 14 mmol, 3 eq) in dry DCM (8 mL) was then injected at RT, forming an olive liquid which was stirred for two hours before **morpholine** (2.4 g, 27 mmol, 6 eq) in dry DCM (5 mL) was injected to quench the reaction. The reaction mixture was poured onto sat. Na_2CO_3, partitioned, and the aqueous phase extracted with DCM (2×15 mL). The combined organic layers were washed with sat. $NaHCO_3$, water, brine, dried on Na_2SO_4, filtered and evaporated under reduced pressure to give 2.0 g of crude yellow solid, fluorescing blue-green under 365 nm UV light. The crude solid was filtered on a porosity 4 glass frit, washing with ethanol (3×5 mL), to leave a residue of **HPQ** which fluoresces strongly green under 254 nm UV light. The filtrate was evaporated then separated by column chromatography with solid deposition (EA:Cy:MeOH 1:4:0 → 1:1:0 → 1:1:0.5) yielding **32a** as a colourless solid (345 mg, 0.58 mmol, 65 %), fluorescing blue under 365 nm UV light, as well as the slower-running morpholino derivative **32x**.

Quenching: In similar procedures, using dry ethanol in dry THF to quench the excess **31a** gave the faster-running ethoxy derivative **32y** instead of **32x**; instead, using dry piperazine in dry pyridine to quench the excess **31a** gave no isolated yield of a piperazine derivative, but separation of the components of the crude was much improved relative to that for the ethoxy and morpholino quenches, and **piperazine is the preferred quencher**. Compound **32a** has R_f = 0.21 in 1:1 EA:Cy, R_f = 0.73 in 2:2:1 EA:Cy:MeOH, and is stained by $KMnO_4$ as well as ninhydrin (orange charring pink). HRMS (TOF MS ESI$^+$) calcd for $[C_{32}H_{42}N_5O_6]^+ = [MH^+]$: m/z 592.3130, found 592.3123. Each of the two diastereoisomers of **32a** has an approximately 70:30 ratio of conformers, thus giving up to four resonances per nucleus and complicating NMR spectral analysis. The COSY spectrum supports the spectral assignment. ^{1}H-NMR (500 MHz, $CDCl_3$) : δ = 8.60 (s, br, 0.5H, 0.5×NH), 8.36-8.11 (m, 2H, H10 + H3), 7.89-7.78 (m, 2.5H, H11 + H12 + 0.5×NH), 7.58-7.47 (m, 2H, H5 + H13), 7.39-7.35 (m, 1H, H4), 7.31-7.15 (m, 1H, H6), 6.84 (s, br, 0.1H, 0.1×NH), 5.64 (s, br, 0.1H, 0.1×NH), 5.05-4.88 (d, br + d,br, 8.5 & 6.5 Hz, 1H, $NH_{carbamate}$), 4.71-4.66 (m, 1H, H16), 4.46-4.34 (m, 1H, H19), 4.20-3.94 (m, 1H, H17), 3.87-3.51 (m, 1H, H17′), 3.43-2.89 (m, 2H, 2×H24), 1.77-1.45 (m, 9H, 2×H20 + H21 + 2×H23 + 2×H25 + 2×H26), 1.09-0.84 (m, 15H, 6×H22 + 9×H29) ppm. ^{13}C-NMR (125 MHz, $CDCl_3$): δ = 174.6 + 174.0 (C18), 163.4 + 162.9 (C8), 156.1 + 155.7 (C27), 153.5 + 152.8 (C15), 151.5 + 151.2 (C7), 149.9 + 149.7 (C14), 149.6 + 149.4 (C1), 134.9 + 134.8 (C12), 132.1 + 131.9 (C5), 131.0 + 130.7 + 129.8 (C3), 128.3 + 128.0 (C11), 127.3 + 127.1 (C9), 126.8 + 126.6 (C13), 126.6 + 126.1 (C4), 126.3 + 126.1 (C10), 124.6 + 123.1 (C6), 121.1 + 120.9 (C2), 79.8 (C28), 54.4 + 52.6 (C16), 53.1 + 50.8 (C19), 43.6 + 41.6 (C20), 41.3 + 41.2 + 39.7 (C24), 39.9 + 39.8 + 38.8 (C17) , 29.4 + 27.1 + 26.9 (C23), 28.3 + 28.2 + 28.0 (3×C29), 25.6 + 25.3 (C26), 24.9 + 24.9 (C21), 23.0 + 22.3 (2×C22), 19.6 + 19.3 (C25) ppm.



The products from quenching excess **31a** with morpholine (giving **32x**, isolated off the column after **32a**) or dry EtOH (giving **32y**, isolated off the column before **32a**) could be isolated and characterised. The product from the piperazine quench was not isolated.

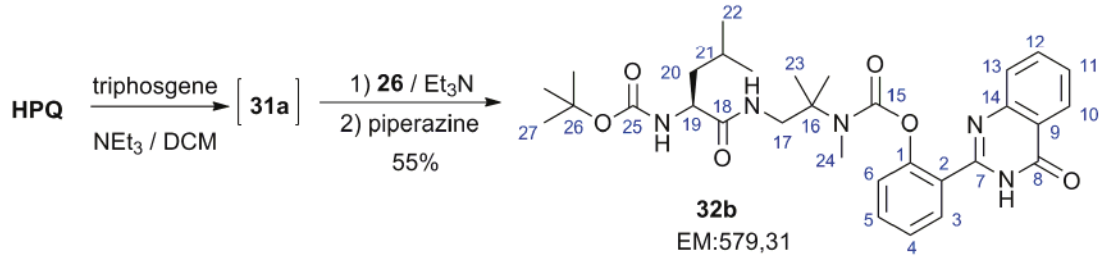
2-(4-oxo-3,4-dihydroquinazolin-2-yl)phenyl morpholine-4-carboxylate (32x)

Compound **32x** has R_f = 0.09 in 1:1 EA:Cy and is stained by $KMnO_4$ as well as ninhydrin (purple charring grey). HRMS (TOF MS ESI$^+$) calcd for $[C_{19}H_{18}N_3O_4]^+ = [MH^+]$: m/z 352.1292, found 352.1286. ^{1}H-NMR (200 MHz): δ = 10.8 (s, br, 1H, NH_{HPQ}), 8.27 (d, 7.8 Hz, 1H, H10), 7.88 (dd, 7.8 & 1.8 Hz, 1H, H3), 7.81-7.72 (m, 2H, H12 + H5), 7.56-7.44 (m, 2H, H11 + H13), 7.36 (td, 7.6 &

1.2 Hz, 1H, H4), 7.26 (dd, 8.1 & 1.1 Hz, 1H, H6), 3.67-3.52 (m, 6H, 4×H16 + 2×H17), 3.49-3.39 (m, 2H, 2×H17′) ppm. ^{13}C-NMR (50 MHz): δ = 162.5 (C8), 153.2 (C15), 150.5 (C7), 149.2 (C14), 149.0 (C1), 134.8 (C12), 132.1 (C5), 130.5 (C3), 127.8 (C11), 127.1 (C10), 127.0 (C9), 126.6 (C13), 126.3 (C4), 123.7 (C6), 121.1 (C2), 66.5 (2×C17), 45.0 (C16), 44.3 (C16′) ppm.

ethyl 2-(4-oxo-3,4-dihydroquinazolin-2-yl)phenyl carbonate (32y)

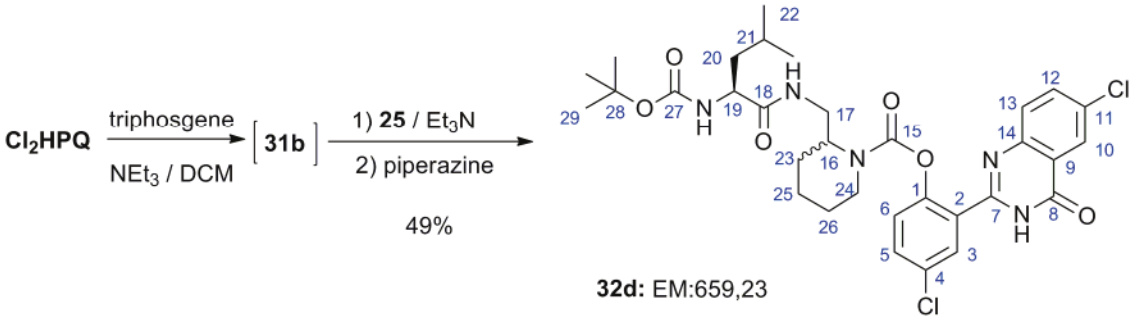
Compound **32y** has R_f = 0.45 in 1:1 EA:Cy, and is stained by $KMnO_4$ but not by ninhydrin. HRMS (TOF MS ESI$^+$) calcd for $[C_{17}H_{15}N_2O_4]^+$ = [MH$^+$]: m/z 311.1026, found 311.1023. The COSY spectrum supports the spectral assignment. ^{1}H-NMR (500 MHz): δ = 8.32 (d, 8.2 Hz, 1H, H10), 8.09 (dd, 7.7 & 1.7 Hz, 1H, H3), 7.83-7.78 (m, 2H, H12 + H13), 7.61 (td, 7.8 & 1.7 Hz, 1H, H5) , 7.54 (td, 7.1 & 1.6 Hz, 1H, H11), 7.48 (td, 7.6 & 1.0 Hz, 1H, H4), 7.40 (dd, 8.4 & 1.1 Hz, 1H, H6), 4.28 (q, 7.1 Hz, 2H, 2×H16), 1.32 (t, 7.1 Hz, 3H, 3×H17) ppm. ^{13}C-NMR (125 MHz): δ = 162.6 (C8), 153.0 (C15), 149.7 (C7), 149.2 (C14), 149.0 (C1), 134.9 (C12), 132.5 (C5), 130.7 (C3), 128.1 (C11), 127.3 (C10), 126.9 (C13), 126.6 (C4), 125.9 (C9), 123.3 (C6), 121.2 (C2), 65.6 (C16), 14.2 (C17) ppm.



2-(4-oxo-3,4-dihydroquinazolin-2-yl)phenyl N^1-(N′-tert-butoxycarbonyl-L-leucyl)-N^2-2-dimethyl-propane-1,2-diamino-N^2-carboxylate (32b)

32b was synthesised from **26** (198 mg, 0.627 mmol) and **HPQ** (149 mg, 0.628 mmol) similarly to **32a**, quenching excess **31a** with piperazine in dry pyridine. Part of the crude yellow oil (200 mg of 299 mg) was taken into DCM (2×1 mL) and the soluble portion decanted then chromatographed (Cy:EA 3:1 → 1:1) giving **32b** as a dense colourless powder (134 mg, 0.231 mmol, i.e. 55 % yield overall), fluorescing weakly blue under 365 nm UV light. Compound **32b** has R_f = 0.22 in 1:1 EA:Cy, R_f = 0.61 in 10:10:1 EA:Cy:MeOH and is stained by $KMnO_4$ as well as ninhydrin (pink charring orange). HRMS (TOF MS ESI$^+$) calcd for $[C_{31}H_{42}N_5O_6]^+$ = [MH$^+$]: *m/z* 580.3130, found 580.3106. The NMR spectra of **32b** suggest the presence of two rotameric forms in approximate ratio 4:1. HSQC and COSY spectra support the spectral assignment. ^{1}H-NMR (500 MHz): δ = 8.29 (d, 7.8 Hz, 1H, H10), 7.91 (d, 7.3 Hz, 1H, 0.8×H3),7.79 (d, 2.2H, H12 + H5 + 0.2×H3), 7.55-7.48 (m, 2H, H11 + H13), 7.39-7.35 (m, 1H, H4), 7.24 (d, 8.3 Hz, 1H, H6), 6.84 (s, br, 0.8H, NH_{amide}),6.18 (s, br, 0.1H, NH_{amide}), 5.60 (d,br, 6.8 Hz, 0.7H, $NH_{carbamate}$), 4.23-4.03 (m, 1H, H19), 3.81-3.37 (m, 2H, 2×H17), 3.05 (s, 3H, 3×H24), 1.80-1.49 (m, 3H, H21 + 2×H20), 1.34 (s, 6H, 6×H23), 1.23 (s, 9H, 9×H27),

0.95 (d, 6.6 Hz, 3H, 3×H22), 0.93 (d, 6.5 Hz, 3H, 3×H22′) ppm. ^{13}C-NMR (125 MHz): δ = 173.8 (C18), 162.5 (C8), 155.8 (C25), 154.8 (C15), 150.7 (C7), 149.4 (C14), 149.2 (C1), 134.9 (C12), 132.3 (C5), 130.7 (C3), 127.9 (C11), 127.6 (C9), 127.2 (C13), 126.6 (C10), 126.4 (C4), 123.8 (C6), 121.1 (C2), 79.5 (C26), 60.9 (C16), 53.6 (C19), 46.5 (C17), 41.2 (C20), 32.7 (C24), 28.3 (3×C27), 25.3 (C21), 25.0 (C23), 24.9 (C23′), 23.2 (C22), 22.0 (C22′) ppm.

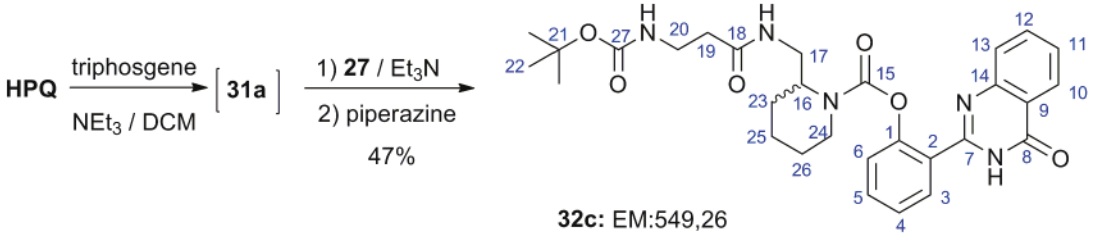


4-chloro-2-(6-chloro-4-oxo-3,4-dihydroquinazolin-2-yl)phenyl 2-((N′-tert-butoxycarbonyl-L-leucyl)methylamino)piperidine-1-carboxylate (32d)

32d was synthesised from **25** (249 mg, 0.762 mmol) and **diClHPQ** (260 mg, 0.846 mmol) similarly to **32b**. After column chromatography and trituration of the oily product fractions with cyclohexane, **32d** was recovered as a white powder (249 mg, 0.376 mmol, 49 %), pure by NMR but fluorescing faintly yellow-green under 365 nm UV light due to residual **diClHPQ** which could be removed by a second column (although this was not necessary for the next step) to give a colourless powder fluorescing faintly blue.

Compound **32d** has R_f = 0.40 in 1:1 EA:Cy, R_f = 0.85 in 10:10:1 EA:Cy:MeOH and is stained by $KMnO_4$ as well as ninhydrin (pink charring red). MS(+): 660 @ 30, 661 @ 10, 662 @ 23, 663 @ 7, 664 @ 5 ; $[MH]^+$. MS(-): 658 @ 100, 659 @ 20, 660 @ 75, 661 @ 25, 662 @ 10 ; $[M-H]^-$. HRMS (TOF MS ESI$^+$) calcd for $[C_{32}H_{39}N_5O_6Cl_2Na]^+ = [MNa^+]$: m/z 682.2170, found 682.2153. The NMR spectra of **32d** are complicated by the presence of two diastereoisomers, each with approximately 1:1 populations of conformers and/or rotamers, giving rise to multiple signals per nucleus. The udeft ^{13}C-NMR sequence[8] proved especially useful in resolving the multiple quaternary carbon peaks. HSQC, HMBC and COSY spectra support the spectral assignment. ^{1}H-NMR (500 MHz): δ = 11.50-11.00 (m, 0.9H, NH_{HPQ}), 8.44-8.35 (s, br, 0.2H, 0.2×NH_{amide},), 8.29+8.27+8.26+8.22 (s+s+s+s, 1H, H10), 8.04+7.98+7.96+7.86 (s+s+s,br + s,br, 1H, H3), 7.78-7.71 (m, 2.5H, H13 + H12 + 0.5×NH_{amide}), 7.67 (d, 7.0 Hz, 0. 25H, 0.25×H6), 7.44 (~d, 7.4 Hz, 1H, H5), 7.25 (d, 8.6 Hz, 0.5H, 0.5×H6), 7.09 (d, 8.6 Hz, 0.25H, 0.25×H6), 6.89 (s, br, 0.25H, 0.25×NH_{amide}), 5.36-4.95 (m, 0.9H, $NH_{carbamate}$), 4.64-4.10 (m, 1H overlapped, H16), 4.42-4.31 (m, 1H, H19), 4.12-2.85 (m, 4H, 2×H17+2×H24), 1.72–1.54 (m, 5H, 2×H25 + H21 + Hx'), 1.36 + 1.21 + 1.09 + 1.01 (s+s+s+s; 9H,

9×H29), 0.98-0.86 (m, 6H, 6×H22) ppm. ^{13}C-NMR (125 MHz): δ = 174.6 + 174.0 + 173.9 + 173.4 (C18), 162.2 + 162.0 + 161.9 + 161.2 (C8), 156.1 + 155.9 + 155.7 (C27), 154.0 + 152.7 + 152.3 + 151.7 (C15), 150.4 + 150.3 + 150.1 + 149.7 (C7), 148.3 + 148.1 + 148.1 + 148.0 (C1), 147.7 + 147.5 + 147.4 (C14), 135.3 + 135.2 (C12), 133.2 + 133.0 + 132.8 (C11), 132.1 + 131.7 + 131.6 + 131.6 (C5), 130.5 + 130.4 + 130.2 (C3), 131.4 + 129.9 (C4), 129.6 + 129.3 (C13), 129.1 + 129.0 + 128.1 (C9), 125.9 + 125.6 (C10), 125.9 + 124.9 + 124.4 + 123.9 (C6), 122.1 + 122.0 + 121.8 (C2), 80.1 + 79.8 (C28), 54.2 + 53.1 + 53.0 + 50.5 (C16), 52.6 + 52.3 + 52.0 (C19), 41.6 + 41.4 + 41.2 (C20), 41.0 + 40.4 + 39.8 + 39.7 + 39.0 + 38.7 (C17 + C24), 28.2 +27.9 + 27.8 (3×C29), 26.8 + 26.5 (C23), 25.5 + 25.3 + 25.2 (C26), 24.8 + 24.8 (C21), 23.1 + 22.8 + 22.2 + 21.7 (2×C22), 19.4 + 19.0 + 18.9 (C25) ppm.

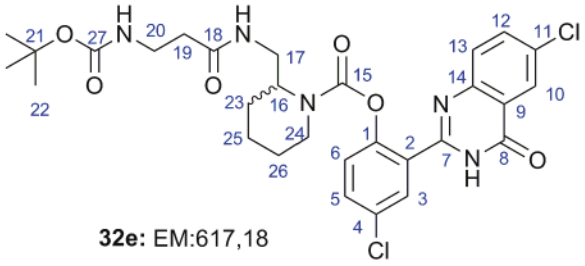


32c: EM:549,26

2-(4-oxo-3,4-dihydroquinazolin-2-yl)phenyl 2-((3-((tert-butoxycarbonyl)amino)propanamido)methyl)piperidine-1-carboxylate (32c)

32c was synthesised from **27** (895 mg, 3.13 mmol) and **HPQ** (1.00 g,4.19 mmol) similarly to **32b**, using Hünig's base instead of triethylamine. After column chromatography, **32c** was recovered as a white powder (802 mg, 1.45 mmol, 47 %), fluorescing weakly blue under 365 nm UV light. Compound **32c** has R_f = 0.18 in 20:20:1 EA:Cy:MeOH and is stained by $KMnO_4$ as well as ninhydrin (strongly green). HRMS (TOF MS ESI$^+$) calcd for $[C_{29}H_{36}N_5O_6]^+$ = [MH$^+$]: m/z 550.2660, found 550.2647. While there are no diastereomeric forms for **32c**, the presence of two conformers or rotamers in approximate ratio 3:2 gives rise to up to two signals per nucleus in the NMR spectra. HSQC and COSY spectra support the spectral assignment. ^{1}H-NMR (500 MHz): δ = 11.00-10.40 (s, br + s, br, 1H, NH_{HPQ}), 8.29 (d, 7.8 Hz, 1H, H10), 8.04 + 7.89 (each d, 7.0 Hz; 1H, H3), 7.83-7.78 (m, 2H, H11 + H12), 7.55-7.50 (m, 2H, H5 + H13), 7.37 (~t, ~7.8 Hz, 1H, H4), 7.31 + 7.21 (each d, ~7.3 Hz; 1H, H6), 7.15 + 6.76 (s, br + s, br; 1H, NH_{amide}), 5.44 + 5.31 (s, br + s, br; 1H, $NH_{carbamate}$), 4.56 + 4.38 (s, br + s, br; 1H, H16), 4.11 + 4.03 (each d, ~12.4 Hz; 1H; H24), 3.62-3.49 + 3.35-3.28 (m+m; 2H, 2×H17), 3.37-3.25 (m, 2H, 2×H20), 3.20 + 2.90 (each t, ~12.4 Hz; 1H, H24'), 2.42-2.22 (m, 2H, 2×H19), 1.40 (s, 9H, 9×H22), 1.77-1.60 (m, 2H, 2×H23), 1.72-1.62 + 1.52-1.42 (m + m; 2H, 2×H26), 1.72-1.48 (m, 2H, 2×H25). ^{13}C-NMR (125 MHz) : δ = 172.8 + 172.4 (C18), 162.6 + 162.4 (C8), 156.3 + 156.1 (C27), 154.9 + 152.9 (C15), 150.6 (C7), 149.3 (C14), 149.1 (C1), 135.0 (C12), 132.3 (C5), 131.1 + 130.7 (C3), 128.0 (C11), 127.5-127.3 (C9 + C13), 126.6 (C10), 126.6+126.0 (C4),

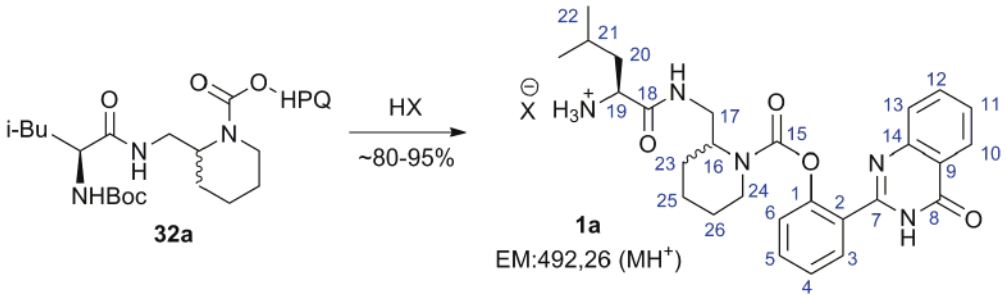
123.6 + 123.1 (C6), 121.1 (C2), 79.5 (C21), 52.0 (C16), 40.9 + 40.6 (C24), 39.4 (C17), 36.8 (C20), 36.4 + 35.9 (C19), 28.5 (3×C22), 26.9 + 26.3 (C23), 25.4 + 25.0 (C26), 19.2 + 18.8 (C25) ppm.



4-chloro-2-(6-chloro-4-oxo-3,4-dihydroquinazolin-2-yl)phenyl 2-((3-(tert-butoxycarbonylamino)propanamido)methyl)piperidine-1-carboxylate (32e)

32e was synthesised from **27** (264 mg, 0.93 mmol) and **diClHPQ** (765 mg, 2.49 mmol) similarly to **32c**. Compound **32e** has R_f = 0.09 in 1:1 EA:Cy, R_f = 0.91 in 1:1:1 EA:Cy:MeOH and is stained by $KMnO_4$ as well as ninhydrin (green charring grey). Much of the **diClHPQ** was removed by filtration on a porosity 3 frit using 6:2:1 Cy:DCM:AcOEt. MS (ESI, positive mode): m/z 618, 619, 620, 621 ; $[MH]^+$. MS (ESI, negative mode): m/z 616, 617, 618, 619; $[M-H]^-$. HRMS (TOF MS ESI^+) calcd for $[C_{29}H_{33}Cl_2N_5O_6Na]^+$ = $[MNa]^+$: m/z 640.1700, found 640.1702. The compound was transferred to Maxime Prost for purification and NMR analysis prior to sending for external tests.

2.6. Deprotection reactions



2-(4-oxo-3,4-dihydroquinazolin-2-yl)phenyl 2-((L-leucylamido)methyl)piperidine-1-carboxylate (1a)

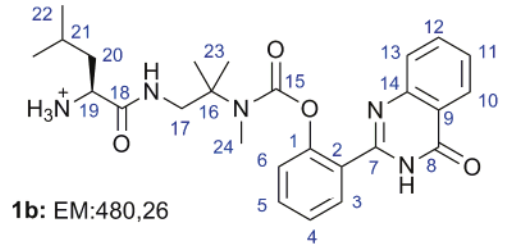
TFA method: To **32a** (15 mg, 0.026 mmol) was added dry DCM (1 mL) and TFA (200 mg, 1.8 mmol, ~7 eq). The reaction was stirred under argon at RT for 1 h, then high vacuum was gradually applied over an hour to remove all volatiles leaving a white powder (the TFA salt of **1a**, yield > 90 %). Short-path column chromatography using eluants of EA:Cy:MeOH could be used to purify the TFA salt of **1a** before enzymatic testing, eluting residual **HPQ** with 1:1:0, unreacted **32a** with 1:1:0.15, and pure **1a** trifluoroacetate with 1:1:1, as a white powder fluorescing weakly blue under 365 nm UV light. Alternatively, ***p*TsOH method**: To **32a** (174 mg, 0.294 mmol) were added *p*-TsOH.H_2O (61 mg, 0.318 mmol) and dry EtOH (10 mL), and the solution stirred at 80°C under a closed argon atmosphere for 8 h. TLC on 1:1:1 EA:Cy:MeOH, revealing with 254 and 365 nm UV and ninhydrin, was used to

monitor reaction progress. After concentration under reduced pressure, column chromatography using eluants of EA:Cy:MeOH was used to purify the tosylate salt of **1** before enzymatic testing, eluting residual **HPQ** with 1:1:0, unreacted **32a** with 1:1:0.15, and crude tosylate of **1a** with 1:1:1 (149 mg, containing a slight excess of *p*-TsOH, 80 %) as a colourless powder fluorescing blue under 365 nm UV light. Trituration of the crude tosylate of **1a** with water (1.5 mL per 100 mg) removes excess *p*-TsOH without dissolving much **1a**, leaving the pure tosylate salt of **1a** as a colourless powder. Methanol or ethanol can be used to dissolve the tosylate salt (~0.5 mL per 100 mg) before its dilution into water, without reprecipitation.

Free basing: $CDCl_3$ (0.5 mL per 15 mg) can be used to freebase the crude tosylate of **1a**, leaving a solid composed only of TsOH and a filtrate containing only the free amine of **1a**. This could be seen as a convenient method to change the counterion of **1a** quantitatively.

HRMS (TOF MS ESI$^+$) calcd for $[C_{27}H_{34}N_5O_4]^+ = [MH^+]$: *m/z* 492.2606, found 492.2588. The TFA salt of **1a** has $R_f = 0.20$ in 2:2:1 EA:Cy:MeOH, and is stained by $KMnO_4$ as well as ninhydrin (vivid orange). The TFA salt of **1a** was easily soluble in D_2O and with a suitable acquisition time (4 s / scan) good NMR spectra were obtained (spectra reproduced on page 267 ff.). The 2 diastereoisomers each appear to have an approximately 2:1 ratio of rotamers in this solvent. HSQC and COSY spectra support the spectral assignment. ^{1}H-NMR (D_2O, TFA counterion, 500 MHz): δ = 8.20 (d, ^3J=8.1 Hz, 1H, H10), 7.94 (~t, ^3J≈7.9 Hz; 1H, H12), 7.77-7.70 (m, 3H, H3 + H5 + H13), 7.67 (~t, ^3J≈7.5 Hz, 1H, H11), 7.48 (~t, ^3J≈7.4 Hz, 1H, H4), 7.36-7.32 (d+d, each ^3J=7.9 Hz; 1H, H6), 4.56-2.73 (m, 6H, H16 + 2×H17 + H19 + 2×H24), 1.66-1.55 (m, 1H, 1×H26), 1.55-1.34 + 1.08-1.00 (m+m, 8H, 2×H20 + H21 + 2×H23 + 2×H25 + H26'), 0.80-0.70 (m, 6×H22) ppm. ^{13}C-NMR (D_2O, TFA counterion, 125 MHz): δ = 170.5 + 170.3 (C18), 163.0 (q, $^3J_{C,F}$=35 Hz, TFA), 162.3 + 162.2 (C8), 154.1 (C15), 153.9 + 153.8 (C7), 148.4 + 148.2 + 148.1 (C14), 141.7 + 141.3 (C1), 137.1 + 137.1 (C12), 135.1 + 134.8 (C5), 130.2 + 130.2 (C3), 129.6 (C11), 127.0 (C10 + C4), 123.6 + 123.4 + 123.3 (C6), 122.4 + 122.3 (C9), 122.1 + 122.0 + 121.9 (C13), 119.2 (C2), 116.4 (q, $^2J_{C,F}$=294 Hz, TFA), 52.0 + 51.9 + 51.8 (C16), 51.6 + 51.5 + 51.4 + 50.9 (C19), 40.5 + 40.1 (C24), 40.0 + 39.9 (C20), 38.2 (C17), 25.8 + 25.6 (C23), 25.0 + 25.0 + 24.5 (C26), 23.9 + 23.9 + 23.8 (C21), 21.8 + 20.9 + 20.9 (2×C22), 18.1 + 18.0 (C25) ppm.

The ^{1}H- and ^{13}C-NMR spectra of the tosylate salt of **1** (in CD_3OD) were similar to that of its TFA salt (in D_2O), but without the peaks for the TFA anion in the ^{13}C-NMR spectrum, and with the following peaks for the tosylate counterion: ^{1}H-NMR (CD_3OD, tosylate counterion, 500 MHz): δ = 7.73 (d, ^3J=8.0 Hz, 2H), 7.26 (d, ^3J=7.8 Hz, 2H), 2.39 (s, 3H) ppm; ^{13}C-NMR (CD_3OD, tosylate counterion, 125 MHz): δ = 142.0, 140.6, 128.6 (×2), 126.0, 125.7 (×2).

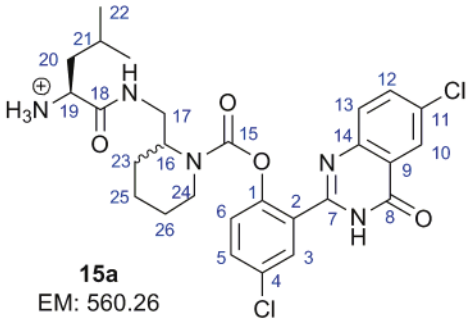


1b: EM:480,26

2-(4-oxo-3,4-dihydroquinazolin-2-yl)phenyl N^1-L-leucyl-N^2-2-dimethyl-propane-1,2-diamino-N^2-carboxylate (1b)

32b (56 mg, 0.097 mmol) was deprotected by TFA (1.5 mL) in dry DCM (1.5 mL) as per **1a**. The TFA salt of **1b** was obtained as a colourless crude powder (yield > 90 %). TLC revealed neither **HPQ** nor **32b**, but it was submitted to column chromatography anyway using a gradient of EA:Cy:MeOH (1:1:0→1:1:1) to be assured of purity before enzymatic testing, giving the TFA salt of **1b** (48 mg, 0.081 mmol, 84 %) as a dense colourless powder fluorescing faintly blue under 365 nm UV light. HRMS (TOF MS ESI$^+$) calcd for $[C_{26}H_{34}N_5O_4]^+ = [MH^+]$: *m/z* 480.2606, found 480.2594. The TFA salt of **1b** has $R_f = 0.0$ in 1:1:0 EA:Cy:MeOH, $R_f = 0.76$ in 1:1:2 EA:Cy:MeOH, and is stained by ninhydrin (vivid red).

NMR spectra are reproduced on page 269 ff.; the udeft ^{13}C-NMR sequence[8] proved necessary to observe the five quaternary carbon peaks marked with an asterisk (*). HSQC, HMBC and COSY spectra support the spectral assignment. **1b**: ^{1}H (D_2O, TFA counterion, 500 MHz): δ = 8.02 (d, ^3J=8.2 Hz, 1H, H10), 7.81-7.77 (m, 1H, H12), 7.64 (dd, ^3J=7.4 Hz & ^3J=1.7 Hz, H3), 7.60-7.56 (m, 2H, H5 + H13), 7.50-7.47 (m, 1H, H11), 7.40-7.37 (m, 1H, H4), 7.24 (d, 8.2 Hz, 1H, H6), 3.81 (t, ^3J=7.4 Hz, 1H, H19), 3.41 (d, br, ^3J=13.4 Hz, 1H, H17), 3.08-3.00 (m, 1H, H17'), 2.93 (s, 3H, 3×H24), 1.54-1.41 (m, 3H, 2×H20 + H21), 1.02-0.83 (s, br + s, br, 6H, 6×H23), 0.82 + 0.80 (each d, ^3J=4.6 Hz; 6H, 6×H22) ppm. ^{13}C-NMR (D_2O, TFA counterion, 125 MHz): δ = 170.4 (C18), 164.4 (*C8), 163.0 (q, $^3J_{C,F}$=35 Hz, TFA),154.5 (*C15), 151.6 (*C14), 148.3 (C1), 147.8 (*C7), 136.0 (C12), 132.9 (C5), 129.7 (C3), 128.1 (C11), 126.5 (C10), 126.2 (C13), 126.0 (*C9), 126.0 (C4), 123.3 (C6), 119.7 (C2), 116.4 (q, $^2J_{C,F}$=294 Hz, TFA), 59.4 (C16), 52.0 (C19), 45.9 (C17), 40.0 (C20), 32.1 (C24), 24.0 (C21), 23.8 (2×C23), 21.6 + 21.2 (2×C22) ppm.

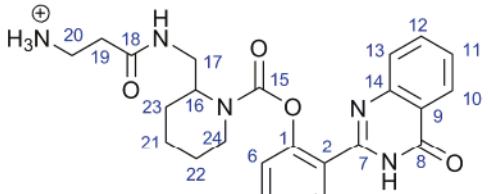


15a
EM: 560.26

L-leucyl 1-((4-chloro-2-(6-chloro-4-oxo-3,4-dihydroquinazolin-2-yl)phenoxy)carbonyl)piperidin-2-ylmethylamide (15a)

32d (150 mg, 0.23 mmol) was deprotected by $Et_2O{\cdot}HCl$ (2M, 3 mL) in DCM (3 mL) as per **1c**, yielding the crude hydrochloride salt of **15a** as a dense, colourless powder (120 mg, 0.20 mmol, 88 %) fluorescing blue-green under 365 nm UV light. Column purification eluted **HPQ** was eluted with Cy:EA:MeOH 1:1:0, then **15a** hydrochloride was eluted with 1:1:0.2 (82 mg, 0.15 mmol, 63%) as an off-white and hygroscopic powder fluorescing very weakly blue under 365 nm UV light. TLC: The hydrochloride salt of **24** has R_f = 0.0 in 1:1:0 Cy:EA:MeOH, R_f > 0.3 in 1:1:2 Cy:EA:MeOH. MS (ESI, positive mode): m/z = 560.2 @ 100,561.2 @ 32, 562.2 @ 65, 563.2 @ 17, 564.2 @ 10 : $[MH]^+$. MS (ESI, negative mode): m/z = 558.2, 559.2, 560.2, 561.2, 562.2 : $[M-H]^-$. HRMS (TOF MS ESI$^+$) calcd for $[C_{27}H_{32}Cl_2N_5O_4]^+ = [MH^+]$: m/z 560.1826, found 560.1814.

NMR spectra are reproduced on page 272 ff.; there were as expected 4 diastereoisomer-rotamer pairs by NMR, though not as well-resolved as the previous examples. ^{1}H-NMR (500 MHz, chloride counterion, CD_3OD): δ = 8.28 (d, 2.4 Hz, 1H, H10), 8.00 + 7.98 (d+d, each 2.4 Hz; 1H, H3), 7.94 (d, 2.4 Hz, 1H, H13), 7.88-7.84 (m, 1H, H12), 7.76 (dd, 8.7 Hz & 2.3 Hz, H5), 7.66-7.44 (m, 1H, H6), 4.71-2.91 (m, 6H, H16 + 2×H17 + H19 + 2×H24), 1.81-1.30 (m, 9H, 2×H20 + H21 + 2×H23 + 2×H24 + 2×H25), 1.04–0.85 (m, 6H, 6×H22) ppm. ^{13}C-NMR (chloride counterion, CD_3OD, 125 MHz, acquired by Maxime Prost): δ = 170.7 (C18), 162.0 (C8), 154.4 (C15), 154.1 (C7), 149.2 (C14), 145.3 (C1), 136.9 (C12), 135.0 (C4), 134.0 (C5), 132.4 (C11), 131.1 (C3), 128.2 (C9), 128.0 (C10), 127.0 (C6), 126.3 (C13), 123.0 (C2), 53.0 (C19), 52.7+52.3 (C16), 41.6 (C24), 41.2 (C20), 39.8+39.4+39.2 (C17), 27.1+26.9 (C23), 26.4+25.9 (C26), 25.4 (C21), 23.2+21.8 (2×C22), 19.8+19.7 (C25) ppm.



1c:
EM:450,21

3-oxo-3-(((1-((2-(4-oxo-3,4-dihydroquinazolin-2-yl)phenoxy)carbonyl)piperidin-2-yl)methyl)amino)propan-1-amine (1c)

HCl method: To **32c** (693 mg, 1.26 mmol) were added, under argon, dry DCM (10 mL) then Et_2O·HCl (2 M, 10 mL). After the initial evolution of gas had ceased (2 min) the reaction mixture was stirred at RT until completion as monitored by TLC (2 hours); sonication could be used to break up the clumps of precipitate that formed during the reaction. The mixture was then filtered and the colourless crystals washed with DCM (2×1 mL) then Et_2O (2×1 mL); the filtrate contained only contaminants. The reaction flask was then rinsed with dry methanol which was then poured onto the collected crystals (2×5 mL); the filtrate was concentrated to yield crude **1c** hydrochloride (474 mg, 0.98 mmol, 77 %), leaving some **HPQ** on the frit. The crude was deemed pure by NMR, but could be purified by column chromatography before enzyme tests. After solid deposition, using eluants of EA:Cy:MeOH, **HPQ** was eluted with 1:1:0, then **1c** hydrochloride was eluted with 1:1:1 (325 mg, 0.67 mmol, 53%) as an off-white and somewhat hygroscopic powder fluorescing very weakly blue under 365 nm UV light. Toluene azeotroping (3×5 mL) was used to partially dry the hydrochloride, which had $R_f \sim 0.0$ in 1:1 EA:Cy, $R_f = 0.05$ in 1:1:1 EA:Cy:MeOH and is stained by $KMnO_4$ as well as ninhydrin (green). MS (ESI, positive mode): m/z 450.2 @ 100, 451.2 @ 30, 452.2@4 : $[MH]^+$. MS (ESI, negative mode): m/z 448.2 @ 100, 449.2 @ 20, 450.2 @ 4 : $[M\text{-}H]^-$. HRMS (TOF MS ESI^+) calcd for $[C_{24}H_{28}N_5O_4]^+ = [MH^+]$: m/z 450.2136, found 450.2133. While there are no diastereomeric forms for **15**, the presence of two rotamers or conformers in approximate ratio 2:1 gives rise to up to two signals per nucleus in the NMR spectra. ^{1}H-NMR (CD_3OD, chloride counterion, 500 MHz): δ = 8.40 (dd, 7.8 Hz & 1.0 Hz, 1H, H10), 8.10 (~t, 7.7 Hz, 1H, H12), 7.96 (dd, 8.0 Hz & 1.0 Hz, 1H, H3), 7.92-7.87 (m, 2H, H5 + H13), 7.84 (~t, 7.6 Hz, 1H, H11), 7.62 (~t, 7.6 Hz, 1H, H4), 7.60-7.52 (m, 1H, H6), 4.74-2.45 (m, 9H, H16 + 2×H17 + 2×H19 + 2×H20 + 2×H24), 1.83-1.04 (m, 6H, 2×H21 + 2×H22 + 2×H23) ppm. ^{13}C-NMR (CD_3OD, chloride counterion, 125 MHz, acquired by Maxime Prost): δ = 171.2+171.1 (C18), 160.1+160.0 (C8), 155.8 (C7), 153.2+152.8 (C15), 149.3 (C14), 140.3+140.2 (C1), 136.5 (C12), 134.8 (C5), 130.3 (C3), 129.2 (C11), 127.2 (C13), 126.3 (C10), 123.5+123.3 (C4), 122.1+121.8 (C9), 121.0 (C6), 119.9 (C2), 51.5+51.3 (C16), 40.4+40.0 (C24), 38.4+38.0 (C17), 35.9 (C20), 31.7 (C19), 25.8 (C23), 25.2+24.8 (C22), 18.6+18.4 (C21).

3. Enzymatic Experiments

3.1. Philosophy of the enzyme tests

3.1.1. The importance of being in earnest

(expanded from section I.5.1.1): Four important aspects to my test design were **(1)** low enzyme concentration, **(2)** minimal background manipulation, **(3)** using only standard equipment in strictly unoptimised fashion to collect data (see section III.3.2.1), with a statistically-defensible, generalised algorithm to perform data treatment, and **(4)** explicitly determining probe stability with spontaneous signal generation control experiments, which were *not* to be subtracted from the active data, but rather to be taken as a minimum estimate of the observable background.

Even chemists who should (and do?) fully understand the importance of spontaneous signal generation for any practical applications, may neglect to explicitly report control data; others 'hide' control data on linear-scale graphs comparing active and control results, and leaving the reader to make a visual comparison of the two curves rather than providing explicit calculations. Whether this too is wilful omission, due to their full understanding that their probes cannot hope to offer the over-impressive results they claim for them when they are taken to *in vivo* settings, is a question which in my opinion should be asked more often by the reviewers who allow essentially fraudulent, or at the worst incompetent (as unrealistic) claims to enter the literature.

To select one recent example in a journal of good repute, consider Urano's 2011 paper in JACS[9]:

(1) 5 units per well of enzyme are used for *in vitro* assays, ie **50 U/mL** since the wells are presumably of volume 0.1 mL but this is unstated, and his cells are transfected to express ~**300 U/mL** (0.3 μU/cell), with the activities *as assayed* in their conditions. Their *in vitro* assay therefore has the potential to convert 50 μmol of substrate per minute per well, ie. **50 mM·min^{-1} probe conversion**, for an initial probe concentration of only 1 μM! Rather than being surprising that this assay therefore rapidly reaches half-maximal signal (43 seconds), it is more surprising that it does not do so very much sooner. The assay was also stopped after only 8 minutes. I consider it impossible that any enzyme of biological interest will be natively expressed at activity levels approaching even 1 U/mL, so even the test design already prevents any true estimation of *in vivo* performance.

By contrast, my *in vitro* fluorescence assays were monitored over 40 h at 30°C, and used only 0.1 U/mL in the well, with activity *as calculated for the standard substrate*, leucyl *p*-nitroanilide, and at 37°C. The true enzymatic activity should decrease sharply when the temperature is lowered to 30°C, and conversion is also likely to be substantially lower for the probes' amide leaving-group than for the small electron-withdrawn anilide; indeed I calculated an effective activity towards probes **1a** and **15a** of only **5 mU/mL**. Although this is ten thousand times lower than Urano's assay, I would still consider it a high activity as compared to *in vivo* situations: but at least it is explicitly stated, and approaching far more reasonable conditions.

(2) Urano's paper claims the **reading just after the addition of the fluorogenic probe as the reference 'background' signal in their reported data**. The claim of "up to 76-fold increase in fluorescence" is therefore misleading: in reality, it reports only on the ratio between absolute ON-state and off-state fluorescence levels, rather than indicating the fluorescence difference actually observable in the experiment, between the enzymatic and nonenzymatic reactions. In a real setting where the probe is dynamically altering its concentration in each tissue by diffusion while simulataneously being activated, no probe-at-time-zero subtraction is possible (see eg. section I.3.1) so this '76-fold increase' is simply impossible. In fact, the observed *in vitro* signal-to-control ratio after 8 minutes is actually less than 25 and dropping constantly (despite using 50 $U \cdot mL^{-1}$ enzyme and only 1 µM probe); eg. after 30 minutes the maximum signal-to-control ratio would be 14, and after 2 hours, only 3. As I will show, the likely signal-to-control ratio which could be observed in any more realistic settings of enzyme activity is almost 1: ie. the probe provides no information about enzyme activity at all.

Furthermore, as the Stokes shift of the probe is low (eg. Urano's 18 nm), then *in vivo*, the tissue autofluorescence will likely provide a large absolute background signal whose variations may well swamp any variation in signal due to probe activation. Low-Stokes-shift probes *cannot* claim that their *in vitro* background is any reflection of the *in vivo* case; and with small Stokes shifts, *in vivo* applications will *always* experience much reduced or even zero sensitivity, so the argument that this '76-fold increase' reflects favourably on *in vivo* applications is doubly flawed (see section I.3.2). By contrast, my *in vitro* fluorescence assays had a large enough Stokes shift to avoid most background increase due to autofluorescence (145 nm), and as per the *in vivo* case, subtract only the no-probe control signal to achieve a baseline reading; yet after 11 hours, their signal-to-control ratio is 3000.

(4) Urano's paper **does not report any no-enzyme control reaction**. A reason which would explain both this and the (unrealistically high) enzyme dose would be that the probe design (an electron-withdrawn conjugated-phenolic glycoside) was inherently unstable in water. Such instability would force assays to be performed with very high enzyme doses (so $[E][S]k_{enz} >> [S]k_{spont}$) and over short time scales; only then could enzymatic reactions plateau and be stopped before the rate of spontaneous hydrolysis outweighed the rate of enzymatic processing. Indeed, his assays are only reported over very short reaction timescales (8-30 min) with huge enzyme activity (50-300 $mM \cdot min^{-1}$ max conversion, 1-10 µM probe). Reading further, even the selection of information he reports actually makes any reliable calculation of probe stability by the reader almost impossible. However, considering the cell experiments, over an incubation period of 30 minutes, ROI inside "control" cells show a fluorescence increase from 27.7 to 30.5 fluorescence units (+2.8), and ROI of transfected cells show an increase from 34 to 73 (+39; includes spontaneous hydrolysis) – ie, a signal-to-control ratio of only 14 despite the high transfected enzyme activity, implying a spontaneous hydrolysis halflife of only 5 hours.

Therefore, were the cells to express a slightly more realistic 1 U/mL enzyme rather than 300 U/mL (note, this is still 200 times the acivity used in my *in vitro* tests), then the ratio of "active" to "control"

cells' fluorescence with his 10 µM probe assay would be [2.8 + (1/300)×(39-2.8)]/2.8, ie a ratio of only **1.04:1** (and dropping constantly). In my opinion these strong objections can hardly support their conclusion that "[this probe has] favorable characteristics for imaging in biological samples," even for transfected cells expressing vast quantities of enzyme relative to what can be expected in a native biological system. And almost certainly, the claims "it should be easy to extend this simple but effective design strategy, to develop fluorescent probes for a wide variety of target molecules, such as other glycosidases (e.g., sialidases/mannosidases), esterases..." are flawed in at least two major respects: (a) the glycosidases mentioned are unlikely to be used as reporters but are usually only present at tiny native activities far less than 1 U/mL, so by the above calculation cannot give any detection of enzyme activity (signal-to-control ~1:1); (b) the esterase probes he alludes to are shown in the Supporting Information to have *far higher* spontaneous signal generation than even the glycosides: an approximately 3:1 ratio of in-cell to extracellular fluorescence signal is shown for the esterase probe structure after only 30 minutes. Again, no data are presented to enable a proper calculation, but this would imply that these esterase probes have a spontaneous hydrolysis halflife of ~50 minutes: thus giving correspondingly far less sensitive enzyme detection than even the glycoside probes.

Note of course that reliable studies may also use huge enzymatic concentations *in vitro* (eg. Papot[10]: 0.13 µmol.mL^{-1} prodrug with 133 µmol^{-1}.min^{-1}.mL^{-1} enzymatic turnover for standard substrate) if stability has been otherwise proven, and the point of interest is to examine post-triggering spacer-degradation kinetics (in this case, halflife 9 hours) rather than proving good enzymatic processing kinetics (for this design, the substrate expressly was well removed from the payload).

3.2. Design of enzyme tests

3.2.1. Full equipment and physics parameters

Fluorescence measurements were performed on a Mithras LB 940 microplate fluorimeter controlled by MikroWin software. Excitation/emission wavelengths were selected by 10 nm bandpass filters at 355 nm / 510 nm, where **HPQ** and **diClHPQ** both give ~50 % of maximum fluorescence output. These filters correspond to standard near-UV excitation, and standard GFP emission filters, which was chosen for the greatest immediate applicability of the results to biology. Note however that their emission also extends well past 600 nm with sizeable intensity, and this far greater Stokes shift may allow even greater sensitivity (lower-B_{ACQ}; see section I.3). Experiments were launched after 20 minutes of dummy acquisitions, to stabilise the apparatus, otherwise the initial data is overlaid with nonphysical signal drift (possibly due to variations in lamp temperature or photomultiplier tube voltage; see eg. Figure III.4).

The choice of wellplate proved to have an influence on the form and reliability of the results. Circular- and square-well wellplates were both tested sucessfully; however, a wellplate that had a large volume and a convex bottom (500 µL, 96-well round-bottom plate, Nunclon) was seen to sometimes saturate

the detector in the geometric-scanning mode when reading central pixels (perhaps by concentrating particles into the centre of well during shaking), and therefore was avoided as unlikely to give the best quantitative results (consider too the effective spatial inhomogeneity factor A/W as outlined in section I.3.3). Therefore for 96-well plates in scan mode, shallower, flat-bottomed plates (220 µL, Corning) were preferred for the acquisition of semi-quantitative data especially with large signals obtained after long incubations or with higher probe concentrations (ie. their higher A/W compensates the high M, and the shallower solution contributes to lower optical path lengths at high signal: see section I.3.3).

Tests were conducted at 29-30°C since a temperature of 37°C too rapidly resulted in the evaporation of the test solutions, and the use of an optical film to stop this evaporation greatly lowered the sensitivity of the detection. The fluorimeter (used in long-dwell-time scan mode) also generates substantial heat which may melt an optical film. The evaporation also preferentially calls for periodic addition of water to the wells to maintain an approximate original volume. For unknown reasons (perhaps air currents from venting or cooling in the machine), wells in the lower left corner of the plate (eg. on a 96-well plate, D02-G05) were especially prone to evaporation; more logically, wells around the wellplate rim were also more prone to evaporation than wells in the centre, so the columns and rows of wells on the borders of the well plate were not used for testing and were filled with water to try and slow evaporation. Due to the spatially-dependent evaporative effects, wells from each set of replicates were spatially scattered across the well-plate to try to diminish skewing of the results (see Figure III.2). The maximal, 4 mm beamwidth setting was chosen to cover the greatest area of each well (well width = 4 mm for the 384-well microplates, 9 mm for the 96-well microplates); so to prevent any possible crosstalk of fluorescence signal from neighbouring wells, only wells lying on the intersection of every second row and every second column were used, with the other wells being filled with water which was intended also to try to slow evaporation (see Figure III.2).

All of these spatial/evaporative effects required careful design to give reliable assay results. However, they could have been avoided by using a clear-bottomed wellplate read from below, with an optical film or lid covering the solutions; however this reading mode was not possible for our apparatus.

Kinetic tests and scan-mode tests were then designed to resemble two desired applications of this probe technology, namely high-throughput screening and cell counting, respectively.

3.2.2. Kinetic mode setup

Kinetic data acquisition mode is the standard mode for *in vitro* assays with soluble fluorophores, and was conceived as an illustration of the probe's capabilities for High-Throughput Screening.

The excitation/emission beams are directed directly into the centre of each well, thereby reading only fluorescence from the centre of the well. Therefore I retained the largest possible excitation beam width (4 mm) centred on the well centre, but switched wellplates, choosing the smallest-width wells (384-well plates, vol 110 µL, Nunclone, 4 mm well diameter), to try to monitor fluorescence from the

whole well (note – as always, neighbouring wells were left blank to avoid crosstalk). Therefore it was considered possible that kinetic data could be considered semi-quantitative, despite the rim problem (see sections I.3.3 and I.5), and this was evaluated favourably later. Kinetic test acquisitions were performed by the "Repeated" option in the software, with microplate shaking for 10 seconds between each plate acquisition, giving one acquisition datapoint for each of the 77 wells used every 5.5 minutes. The acquisition time per well was 3 seconds, and a lamp power setting of 10000 arbitrary units (calibrated to a software internal standard) was used as this gave satisfactory signal without saturating the detector. The 77 series of fluorescence signal readings (in RFU) per well per time point could be acquired within a time range up to around 20 h, but acquisition was more reliable under 12 h due to evaporation of the well solutions.

3.2.3. Scanning mode setup

This data acquisition mode is a standard mode for *in cellulo* assays such as cell counting, and the test results also illustrate the probe's capabilities in that application.

To try to give the highest sensitivity towards the detection of small amounts of signal *e.g.* for evaluating probe stability and/or its detection threshold, a Geometric Scanning mode was set up whereby the excitation/emission beam was scanned over the entire well area including the rim, as well as covering some pixels outside the well boundaries to ensure complete coverage of the well. This could be advantageous not only to cover the initial rim, but also in case of situations *e.g.* where only one crystal of **HPQ** may have formed, in a corner of the well not usually reached by the Kinetic mode measurements (see section I.3.3 and Figure III.1).

The software's "Scan" mode was selected, typically with a 20×20 pixel acquisition grid set to completely cover the well of interest. Due to software requirements this actually implies setting the acquisition grid to cover also one well radius into the eight neighbouring wells, which in my experiments were filled with pure water to prevent false signal contributions (crosstalk); thus the true measurement grid lying directly over the well of interest is of approximately 10×10 pixels (Figure III.1). Therefore a false wellplate configuration file, with well diameters twice as large as the real size, was created (*OTS_CBO_trick.cfg*).

Due to the 4 mm beamwidth, the real data contained in each pixel overlaps with that of neighbouring pixels (typical settings : pixels nominally ~1×1 mm, but fluorimeter beamwidths ~4 mm), so the data of these scan-mode tests weight interior pixels more than exterior ones (more overlaps). Therefore it was of importance to try to consider how quantitative the results could be. Despite a simplistic treatment, they proved to give final signals which could be correlated reliably to the original concentration. Note too, that some pixels *centred* outside the real-signal area still show a fluorescence signal from overlap of the outer part of the pixel area with a portion of the well (Figure III.1), so only pixels centred in the wells were taken into account in analyses.

It was considered that a square inscribed *within* the true area of the circular wells was an adequate sample of the well, and that selecting this grid of 10×10 - 12×12 pixels was sufficient for reliable data acquisition (given a reasonable acquisition time and excitation intensity). Pixels where the detector had saturated were given the value of the average of their neighbours, thus giving a lower bound for the hypothetical reading. No more than 1/20 of the pixels per well were ever observed to saturate. A typical acquisition duration for a single scan of a plate (15 wells, 20×20 acquired pixels per well) was therefore around 30 minutes. Scanning measurements used 0.25 s acquisition per pixel and a lamp power setting of 30000 arbitrary units. Scans were conducted at specific time points (eg t=1 h, t=2 h, t=24 h etc.) during incubation, with a coverslide placed over the wellplate between readings to minimise evaporation. Data acquired could be visualised as a linear-scaled greyscale plot of fluorescence reading *vs*. pixel location on the well (Figure III.1).

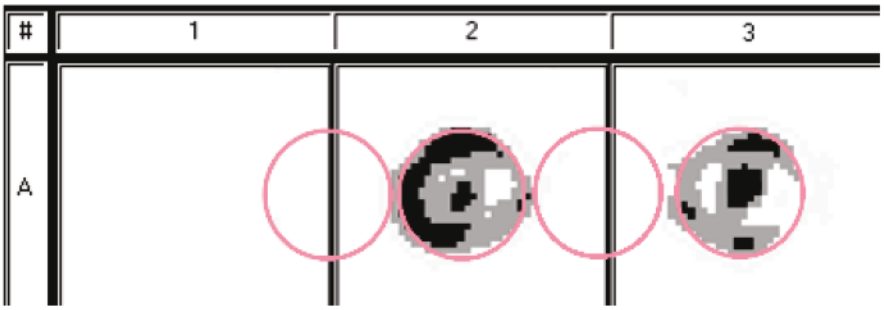


Figure III.1 - Scanning mode results visualised in black (high signal) on white (no signal), showing rim + centre deposits of fluorophore with the initially-used convex-bottomed Nunclon 96-well plates. The circular well boundaries are indicated in pink.

Note that each grid square focuses on one well, although neighbouring wells enter into the fluorescence images (well outlines marked in pink), due to the dummy renumbering. Gridsquares A2 and A3 illustrate the initial precipitation of **HPQ** around the cell rims, with subsequent precipitate then falling to the bottom (ie. centre, in a convex well; cf. section I.3.3). Figure III.1 thus demonstrates why geometric scanning was initially considered necessary for quantitative results, and why kinetic mode readings with this wellplate (9 mm wells, 4 mm beamwidth, so central pixels only are acquired) was not considered sufficient, as most of the rim signal would be missed. Note however that the resolution of the scanning experiments used in data analysis was less than that presented in Figure III.1; a typical scan delivered results like the following (Figure III.2).

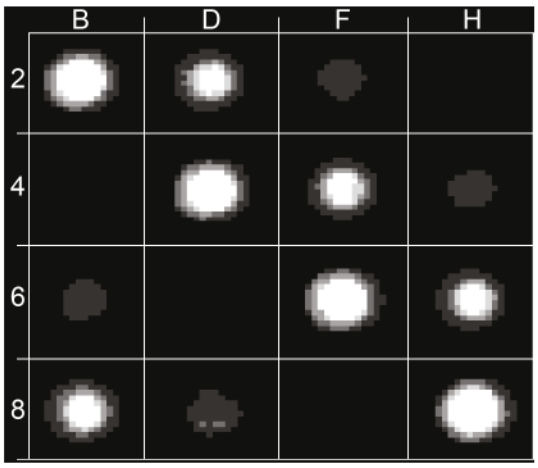


Figure III.2 – Typical scan mode results for a single timepoint visualised in greyscale. The grid has been re-numbered to represent the true wellplate configuration; note the series spatial distribution and also the 'skipping' of every second well (filled with water) to minimise crosstalk, cf. section 3.2.1. The four different data series (1a at 209 µM, 80 µM and 20 µM with enzyme, and a no-enzyme-control at 209 µM) can clearly be distinguished, even after 40 h.

3.2.4. Biochemical parameters and experimental procedure

To partially simulate physiological media, all tests were conducted at pH~7.5, buffered with Tris/Tris.HCl (25 mM) and containing 12 mM NaCl, hereafter referred to as simply pH~7.5 buffer. Stock solutions of the substrates were prepared fresh before each run by weighing precisely ~2 mg of the substrate salt, adding 0.2 mL MeOH to dissolve it, and diluting this stock with pH~7.5 buffer (with secondary dilutions when needed) such that the volume of substrate stock to be added to the well lay between 15-50 µL, and so that no individual dilution in the series was by a factor of more than 10 (for best accuracy). In a typical run, such prepared substrate stocks (from 70 – 400 µM) were **left standing for approximately 2-3 hours** at RT before the acquisition was launched, although the 378 µM data series for probe **1a** of run 1 was prepared from a solution that had been left standing for three weeks (illustrating a different practical aspect of the robustness of the probe system).

'Microsomal neutral leucine aminopeptidase from porcine kidney' (LAP, EC 3.4.11.1) was purchased from Sigma (Lot number 069K7356V) as a suspension in ammonium sulfate (3.5 M) containing $MgCl_2$ (10 mM), with 3.5 $mgmL^{-1}$ protein content and 10-40 U/mg protein reported for the standard substrate leucyl *p*-nitroanilide. LAP enzyme stock was prepared by diluting 10 µL of the enzyme suspension (stored at 4°C for less than 40 days before use) into 2 mL of the pH~7.5 buffer and shaking, then leaving at RT for at least 1 h. Incubation of the enzyme with $MgCl_2$ or $MnCl_2$ (described elsewhere as necessary to activate it from the dormant state[11]) was not observed to give an improvement of signal generation kinetics.

NaCl solution (25 µL), distilled water (11-30 µL), pH~7.5 buffer (0-40 µL) were then added as needed to each well such that both the total volume and the final buffer concentration were identical in all wells, irrespective of the volume of substrate stock to be added. The substrate stocks were then added (15.5-44 µL). The final step was the addition of **20 µL** of the enzyme stock to each of the active wells, so that in the kinetic tests the final concentration of enzyme was ~3.5 μgmL^{-1} (~60 nm[12]), with

approximately 10 mU of activity per well for the standard substrate leucyl *p*-nitroanilide, based on the activity quoted by the supplier. **This addition was performed rapidly and in the same order in which the wells were to be read**, giving approximately ± 2 minutes difference between the average real time elapsed after mixing and the fixed time as given by the microplate reader, to enable the best comparison of data between different wells. Importantly, the delay time between adding the first *substrate* stock and starting the first acquisition was always less than 30 min. After the addition of the enzyme, the contents of each well were homogenised by triple aspiration and expulsion of the contents, and then the plate was loaded into the fluorimeter and acquisition started. The time delay before the first acquisition (approximately 10-15 minutes) was factored in to the data during treatment, which is why the kinetic result signals have their first timepoint only at that moment.

Considering pre-spiking will be seen to be important when comparing the performance of my probes to the prior art, which was why I carefully chromatographed my probes before enzyme testing, why I insisted on a rapid preparation of the stocks and the well solutions from these chromatographed salts before each run, and why I chronometred the total substrate preparation times (<<200 min in total).

3.2.5. Data processing – kinetic mode

In each run, data were acquired from at least five replicates for each set of conditions used with enzyme-activated probes, or from at least three replicates for the controls (either no substrate or no enzyme). The parameters varied were substrate type, substrate concentration, and enzyme presence or absence. The raw data were obtained as series of fluorescence measurements in Random Fluorescence Units (RFU) over time. The raw data for each well are denoted $A_{XI}(t')$, where X indicates the set of conditions in that well, I indicates the number of the replicate within the set of wells with those same conditions, and t' is the number of the data point in the acquisition series, *e.g.* t'=1 for the first data point acquired, t'=2 for the next, etc. The series $A_{XI}(t')$ were first checked to confirm that no detector saturation had taken place. Then, three steps of data treatment were performed to give fluorescence signals $F_X(t)$ for each set of conditions X over real elapsed time *t*, with background signal subtraction, normalisation, outlier rejection and replicate selection performed as follows:

1) Inherent Background signal B_{ACQ} (due to scattering from the plate and the solution and noise in the system) was subtracted, and all data were normalised to compensate for temporal drift in detector sensitivity or lamp luminosity, permitting a more reliable quantitative comparison of the data over the long acquisition times of the experiment. The data from wells containing only *distilled water* (not even enzyme in buffer) were averaged to give the background signal B(t'), and an approximate normalisation was performed by multiplication of the series by the factor B(1)/B(t'). Thus the signals $G_{XI}(t')=[A_{XI}(t')-B(t')]\times B(1)/B(t')$ were obtained.

2) Outlier data points were then discarded from each series if they were suitably different from their neighbours. It was considered that such outliers (*e.g.* points differing in intensity by up to a factor of 3 from the mean of the two neighbouring points) were not necessarily due to true spikes in fluorescence

signal, but that some were due to detection faults, *e.g.* the presence of strongly scattering dust particles in the beam path, which could justify their deletion; others could have been due to the beam striking (or missing) an unusually dense clump of the precipitated fluorophore (see section I.3.3). Arbitrarily, 5% of the data points from each series were deleted as outliers, as determined by the following algorithm:

For each series $G_{X,I}(t')$, a series of expected values $\mu_{X,I}(t')$ was calculated as per $\mu_{X,I}(t') = [(3G_{X,I}(t'-3) + 2G_{X,I}(t'-2) + 2G_{X,I}(t'+2) + 3G_{X,I}(t'+3))/10]$. Data points further from the point of interest were more heavily weighted than nearer data points to compensate for the observed effect that such outliers often occurred in spikes of two to three data points, possibly due to *e.g.* the temporary settlement of a dust particle on the plate near the well being read. A weighted mean-squared difference function $\Delta_{X,I}(t') =[G_{X,I}(t') -\mu_{X,I}(t')]^2/[| \mu_{X,I}(t')| + |G_{X,I}(t')|]$ was then calculated, and the points $G_{X,I}(t')$ giving the greatest values of $\Delta_{X,I}(t')$ deleted one by one, with recalculation of $\mu_{X,I}(t')$ and $\Delta_{X,I}(t')$ between every deletion. If the outlier data points thus rejected (approximately 5-10 per series) corresponded to the region of fluorescence signal onset, they were reinstated, since this algorithm was unsuited for application to regions with an inherently non-linear signal. Treatment of the first and last three data points in a series required unidirectional calculation of $\mu_{X,I}(t')$.

3) Data series $G_{X,I}(t')$ within each set X of replicates were then compared to see whether they were **consistent**; series inconsistent with the others in their set were discarded in the following way. The mean $M_X(t')$ of all series in a set was calculated, and then the average value $Q^*_{X,I}$ of the mean-squared error function $Q_{X,I}(t') =|[G_{X,I}(t') - M_X(t')]^2/M_X(t')|$. To reject a series I, the ratio $R_{X,I} = [Q^*_{X,I}$ / (the average value of the last five points of $M_X(t')$)] had to be above an arbitrary threshold, chosen as 0.025. After having rejected the replicate with the largest $R_{X,I}$, the procedure was followed again (with prior recalculation of $M_X(t')$, $Q_{X,I}(t')$ and $R_{X,I}$) until either only three series remained in the set X, or until all series had $R_{X,I} \leq 0.025$. Then, data from corresponding time-values of these remaining replicates were averaged, and the dummy times t' converted to true elapsed times *t* since the addition of the enzyme, to give the fluorescence signals $F_X(t)$.

Although similar signal processing treatments are common in physics, I have not yet found them used explicitly in probe data analysis; one possible explanation for this absence is that they remove the possibility of 'touching up' data by visual inspection adjustments.

3.2.6. Data processing – scanning mode

RFU readings from the central 10×10 pixel grids were averaged to give a single RFU value per well per time point, and these values were filtered for consistency as per step **3)** in the kinetic test data treatment section, discarding down to a minimum of two replicates per series to obtain $A_X(t)$ as a raw RFU value of wells near the mean in their set. These were then corrected by subtraction of the instantaneous B_{ACQ} background reading B(t), determined using the pure-water-only wells as a

minimum estimate of background noise in the fluorescence acquisitions, yielding $S_X(t)$: using the letter S differentiates Scan data from kinetic data, which were denoted $F_X(t)$.

3.2.7. Choosing the important factors to vary

I performed kinetic and scanning experiments over 10-40 h, with probes **1a**, **1b** and **15a**, comparing their hydrolytic stability (no enzyme, high probe concentration) and kinetic response (with enzyme, varying probe concentration), to conclude on which spacer (piperidyl **22** or dimethyl **21**) gave the faster observed response (combines both enzymatic processing and cyclisation steps), and to compare the stability and kinetic performance differences that might be seen with the **HPQ** and **diClHPQ**-based probes at different concentrations.

I also ran negative control tests with morpholino derivative **32x** to see if the enzyme used actually nonspecifically hydrolysed this probe, which it should not. Pleasingly, no signal was observed during that experiment, so I conclude that the hydrolysis observed during the probe tests was the desired stepwise process: first, enzymatically catalysed peptide hydrolysis, then, cyclisation to expel the fluorophore.

It was considered that the most reliable comparisons of data would be between sets acquired in the same experiment, so in general I used one experiment to examine changes in only one or two factors. Replicates using similar conditions were used to prove the reliability of comparing results from two different acquisition sets. To enable comparisons between all data sets, I always used the same (approximate) concentration of the enzyme. Note that the enzyme is delivered in a somewhat inhomogenous suspension so the aliquots taken to prepare each run may contain different amounts of enzyme, even if the µL amount aliquotted can be well repeated (est. repeatability error ~3%).

3.3. Enzyme test analyses: steady-state parameters and $C_{PPT}(t)$

3.3.1. Calculation of solubility constants K_S

Extrapolation Method: Assuming the initial concentration of **diClHPQ** in my 100 µM no-enzyme control was strictly zero, the evaluation of the *effective* K_S was rather straightforward, because a spontaneous hydrolysis test of **15a** gave a clear, and **abruptly linear** start of signal in the 100 µM no-enzyme control reaction after ~500 minutes (R^2=0.9977 over >6 hours; Figure I.26), which I interpreted as resulting from rate-limiting, pseudo-first-order spontaneous hydrolysis of the carbamate. If precipitation from the supersaturated solution had been a significantly rate-affecting factor, the region of initial signal generation would not have shown an abrupt switch between the two regimes. The linear fit to the sponaneous hydrolysis was extrapolated to a y-intercept of -591×10^3 RFU, which should correspond to the signal expected from the 'missing' probe, ie to the *effective* K_S.

For the most reliable results in correlating this RFU reading with concentration, I initially considered the data for the 10.1 µM probe plateau in this experiment: because mechanistically I expected the best linearity $F_{PLAT}=k\times C_{PPT}$ in the weak absorption regime at low concentration (cf. section I.3). As its

plateau was at 4.4×10^6 RFU, I calculated that $K_S=10.1\times0.59/(0.59+4.4)=1.2$ μM (this method of calculation creates the lowest carry-on error in K_S due to non-linearity effects: see later). I then checked this estimate against the results from the 40.5 μM experiment in the same run: $F_{PLAT}=15.0\times10^6$ RFU, so $K_S=40.5\times0.59/(0.59+15.0)=1.5$ μM; a similar consideration from the 100 μM run ($F_{PLAT}=25.7\times10^6$ RFU) gives $K_S=100.5\times0.59/(0.59+25.7)=1.6$ μM. The agreement between these estimates was pleasing (although it was not necessary from a theoretical viewpoint, because the linear correlation of signal↔concentration is not actually true).

I then had to choose which value to trust most. Although I had greater confidence in the completion of attaining signal plateau with higher probe concentrations (see section III.3.3.3), what was actually much more important was considering the strong absorption regime. If this strong absorption regime is approached then the signal plateaux will be lower than expected for a linear relationship, so the estimated K_S will be greater than the real ones. In fact, initial calculations of the values of k from $F_{PLAT}=k\times C_{PPT}$, using $K_S=1.4$ μM as the average figure, give k=(0.494, 0.384, and $0.258)\times10^6$ RFU·M^{-1} for the three concentrations 10, 40 and 100 μM respectively, illustrating exactly this strong absorption effect, and highlighting why I chose the ratiometric calculation method for K_S, therefore I chose the *effective* $K_S=1.2$ μM (see section 3.3.4 for a more complete description of the nonlinear signal↔concentration relation).

A final consideration was that actually, this calculation method delivers not the true K_S but its *effective* counterpart (K_S-C_H), for a 'pre-spiking' concentration C_H of **diClHPQ** present at time t=0 in the experiment. To estimate the uncertainty in assuming (K_S-C_H)=K_S, I first considered that the 100 μM well no-enzyme control results were obtained by dilution of 25 μL of a stock solution (which had been prepared less than 3 hours before), to a final volume of 100 μL, with a delay of <30 min between dilution and signal acquisition (see section 3.2.4). Considering that the slope of this spontaneous hydrolysis regime was 1052 RFU/min, I used the initial value of $k=0.494\times10^6$ RFU·M^{-1} from the weakest-absorption regime (which is most relevant at small concentration) to convert this roughly to a spontaneous hydrolysis rate of about 0.6 μM probe per 5 hours, at a probe concentration of 100 μM. By calculating the decomposition at each step, I concluded that the likely maximum value of C_H was about 0.4 μM, so the *true* value of K_S was somewhat less than 1.6 μM. I therefore chose K_S(**diClHPQ**) = 1.5 μM.

Note that this consideration of pre-spiking makes little difference to the results here, because the probe was stable, and because the sample preparation was performed both rapidly and in a quantifiable manner. For unstable probes, as will be seen for **34a** and **34b** (section III.3.5.1) failure to consider pre-spiking completely distorts the results.

Initial Velocity Method: For **HPQ**, no spontaneous hydrolysis was observed in either **1a** or **1b**. Therefore I first examined the quasilinear signal generation regime, near the start of the observed signal in the kinetic tests with the enzyme present. Assuming that similar behaviour can be expected

here as for **15a**, then the spontaneous hydrolysis contribution will be not only strictly linear, but also negligeable. Now, assuming that (a) the enzymatic hydrolysis rate is approximately constant during the early period of the experiment (which is expected for high substrate loading with a well-processed substrate), and (b) both cyclisation and crystallisation are fast (ie. not limiting for signal generation), then the extrapolation of the $F_X(t)$ curves from their greatest gradients (at this linear region) to t=0 should likewise indicate the signal that would be expected from the 'missing' (non-precipitated) concentration of fluorophore which is originally generated before signal is seen. However, for **HPQ**, the results obtained by that method were very inconsistent and gave unrealistic estimates of K_S, eg. 50-100 µM. I therefore concluded that assumption (b) was wrong, and because I observed fast crystallisation with the similar fluorophore **diClHPQ**, I concluded that cyclisation *was* rate-limiting (see Figure III.8 and the extensive kinetic analysis in section I.5.5).

Plateau Fitting Method: Therefore I instead examined the scan mode results S_X(plateau) when the signals had stabilised, ie. all the probe had been hydrolysed and precipitated in the enzyme wells (eg. Figure III.3). Care was taken to cover the wellplate with a plastic slide between measurements (if the solutions go to dryness, then all the activated probe precipitates and the determination is inaccurate).

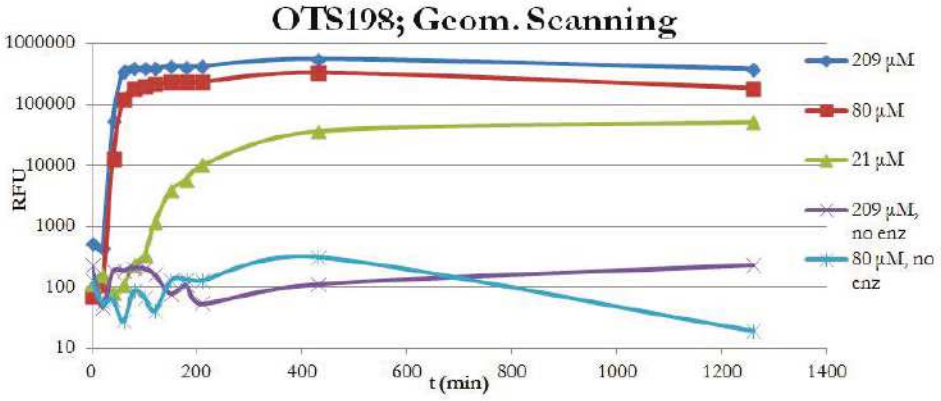


Figure III.3 - Representative scan mode results, SX(t), for probe 1a (note the logarithmic scale). Time-resolved scan data show that high concentration tests (80 µM-209 µM) reach their maximum signal plateau after 1 h and remain stable for the next 20 h, that a low-concentration test (20 µM) only approaches its signal plateau after around 7 h, and that high-concentration substrates without enzyme do not give rise to any appreciable signal over 20 h.

I considered that if the off-state really generates zero signal, then all significant signal should correspond to precipitated probe. Therefore the 'precipitated concentration' C_{PPT} at completion is given by $C_{PPT} = P_0 - K_S$; and if a linear correlation can be applied, then $S_X(\infty) = k \times C_{PPT}$. I minimised the error of fitting the observed signals to this relation, to calculate values of K_S within each experiment, while constraining K_S to physical values. For example, for **HPQ** this implied $10 > K_S > 5$, because 5 µM wells never gave a signal, while 10 µM wells of **1a** gave a signal after about 7 hours. As per section I.3.3, I gave most weight to the results of plateaud experiments with medium substrate concentrations, eg. 40-80 µM, to favour the most linearity in the $S \leftrightarrow C_{PPT}$ relation. From these fits I derived 6 µM$<K_S$(**HPQ**)$<$8 µM across three independent experiments and I chose K_S(**HPQ**) = 7 µM;

this also agrees well with an independent analysis of the behaviour of prior-art probes **34a** and **34b** (see section 3.5.1), so is therefore considered sufficiently reliable.

3.3.2. Hydrolysis halflives

As introduced in section III.3.3.1, for **15a**, the hydrolysis calculation was relatively easy, considering the strongly linear spontaneous hydrolysis regime (R^2=0.9977 over >6 hours; Figure I.26). I recalculated the linear estimation of signal↔concentration using K_S=1.5 μM, giving k=0.512×10^6 RFU·μM^{-1} by using the plateau signal from the lowest-concentration well (with therefore the most linear / weakest-absorption regime, and also, the concentration↔RFU relation which is most relevant at small concentrations). Then I converted the observed fluorescence increase over time for the 100 μM control dataset (+0.308×10^6 RFU from 550 to 850 min) to an amount 0.6 μM spontaneously hydrolysed in 300 min. Fitting C_0=98.5, C_{300}=97.9, t=300 min to a unimolecular decomposition model yields the hydrolysis halflife τ(**15a**)=567 hours (24 days).

As no spontaneous hydrolysis was observed in either **1a** or **1b**, I instead calculated thoroughly reliable minimum halflives by positing the appearance of a fluorescent signal just after the last timepoint of the longest kinetic monitoring of each probe at high concentration. The longest experiment was followed over 2093 minutes (=35 h) with neither **1a** nor **1b** giving a signal at 100 μM; a more concentrated, 21 h experiment at 209 μM in each confirmed this result. Given K_S=7 μM would therefore estimate the minimum halflives from the latter experiment as τ(**1a**), τ(**1b**)> 427 h.

Beyond this though, on 23/8/2011 I prepared two tubes containing **1a** at 2000 and at 801 μM and left them standing at RT for 3 weeks, and visually observed no fluorescence under UV. On 13/9/2011 I then used the 801 μM solution in an enzymatic experiment with a dilution factor of only 2, and I still observed a delay time before the onset of rapid signal generation, so conclude that there was minimal pre-spiking by hydrolysed probe. Compare this to the form of the results of pre-spiked probes **34a** and **34b** (Figure III.9), which exhibit absolutely minimal delay times. I therefore conclude that these vials likewise did not reach the solubility threshold, so I estimate a new lower-bound hydrolysis halflife (fitting C_0=2000, C_{21}=1993, t=21 days) of τ(**1a**)≈4151 days, and given their similar structure at the hydrolytically weakest carbamate bond, I would give a similar estimate for **1b**.

3.3.3. Probe sensitivity: signal-to-control ratios at plateau

Plateau values: Time-independent plateau values $F_{PLAT}(x)$ were obtained for signal-stabilised timecourses with initial probe concentrations x, by averaging their signals between t=500 min and t=660 min. Two chief sources of error in obtaining physically meaningful plateau values $F_{PLAT}(x)$ were anticipated:

(a) If the enzymatic reaction slowed substantially towards the end as substrate became scarce, the plateau values observed would be underestimates of the true plateau fluorescence values; however, based on the rapid attainment of the plateau, it seemed that the K_M was favourable and that therefore

the reaction could be considered practically complete. Then, fluorescence signals from probes **1a** and **1b** at identical concentrations show almost identical plateau values, suggesting that both diastereomers of the **22** spacer are enzymatically hydrolysed. Furthermore, both diastereomers of the **22** spacer appear to be hydrolysed at comparable rates, since otherwise the kinetic results (eg. Figures I.25 and I.27 for **1a**) would display two disctinct regions: first faster hydrolysis of one diastereomer, then slower action on the other, whereas the signal is in fact smoothly simple-sigmoidal. Therefore I conclude that the signal plateaux reached in the all the experiments can be assumed to indicate total activation of the initial quantities of the probes x.

(b) Continuous evaporation from the wells concentrates the test solutions; when the ratio K_S/P_0 was relatively high, this could significantly impact the effective x (the more evaporation, the more probe is precipitated out of the starting concentration, so (x–K_S) later leads to an underestimate of the signal-generating concentration). This was countered by periodic topping-up of the wells.

Before the enzyme-active wells are considered, first examine the raw signal from the distilled-water-only control wells in a typical kinetic experiment to see the expected size of variations due to noise and systematic variations (Run 1; Figure III.4). These show initial stabilisation (t<100 min), then smaller systematic variations in signal (period ~80 min, amplitude ~5000 RFU); by the data treatment algorithm outlined above, these were mostly compensated for in the generation of processed signals F(t) from the raw data A(t). The smaller-still variations due to true noise are superimposed on this background (amplitude ~500 RFU). Note that the bulk of the signal acquired is more like due to fluorescence of the wellplate plastic, rather than of any species in the distilled water.

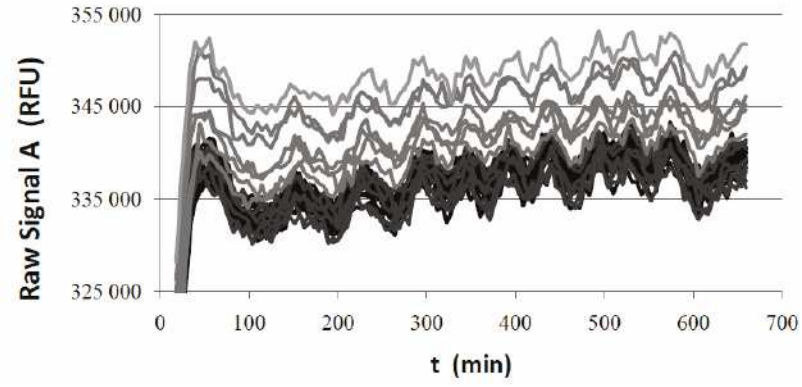


Figure III.4 – The raw fluorescence signals for 20 water-only blank wells illustrate the levels of correction needed in processing acquired signals A(t) into true fluorescence signals F(t).

After data processing, both scan and kinetic modes clearly showed huge sensitivity for the enzyme-active wells, and no false positive signals for the no-enzyme controls or the probes with concentration less than K_S, for probes **1a** and **1b**. As shown in Figure I.26, the 100 μM no-enzyme control of **15a** gave a signal due to spontaneous hydrolysis after 8 hours, but even so, sensitivity was very high. Representative results are presented in Table III.1 (kinetic mode, all probes). Recall that processing raw data A(t) into F(t) involves not only outlier removal but series consistency checking, which

explains small variations in trends between the two. The very high signal-to-control ratios calculated for the plateaud signals, and even for signals that were not yet plateaud, indicates the extremely high sensitivity of these probes as a practical system for detecting aminopeptidase activity.

Kinetic mode, 11 h	**A_{PLAT} ($\times 10^5$ RFU)**	**F_{PLAT} ($\times 10^3$ RFU)**	**F_{PLAT}/F_{NoEnz} (Signal/Control), no t_0 subtract**	**Signal / Control, t_0 *subtract***	**Comment**
RUN 1 (660 min):					
Distilled water	3.4	(0)	-	-	-
Enzyme without probe	3.4	0.8	-	-	-
1a, 378 µM	362	37517	728	7441	-
1b, 100 µM	112	8661	1467	∞	-
1a, 100 µM	101	9711	618	5712	-
1b, 59 µM	80	5631	954	∞	-
1a, 59 µM	55	3861	371	3976	-
1b, 32 µM	38	3749	635	∞	-
1a, 32 µM	30	2951	427	6009	-
***1b**, 18 µM*	*13*	*605*	*101*	∞	*no plateau*
***1a**, 18 µM*	*13*	*551*	*108*	*2274*	*no plateau*
***1b**, 10 µM*	*3.5*	*9.5*	*1.5*	∞	*no signal*
***1a**, 10 µM*	*4.8*	*47.6*	*12*	*476*	*no plateau*
1b, 100 µM, NoEnz	3.5	5.9	(1)	-	no signal
1a, 100 µM, NoEnz	3.5	15.7	(1)	-	no signal
1b, 10 µM, NoEnz	3.5	6.0	(1)	-	no signal
1a, 10 µM, NoEnz	3.5	4.1	(1)	-	no signal
RUN 2 (960 min):					
Distilled water	4.5	(0)	-	-	-
Enzyme without probe	4.5	8	-	-	-
15a, 100 µM	310	25818	1655 (64)	43000 (66)	t=300 (t=960)
1b, 100 µM	113	20613	2944	21000	-
15a, 40.5 µM	166	15009	962	25000	-
1b, 40.5 µM	38.5	9872	1400	10000	-
15a, 10 µM	56	4314	276	7190	-
1b, 10 µM	6.5	16.6	1.1	17	signal *just* starting (960)
15a, 100 µM, NoEnz, 200<t<400 min	4.7	15.6	(1)	-	no signal
15a, 100 µM, NoEnz, t=960 min	9.2	404	26	673	no signal
1b, 100 µM, NoEnz	4.6	7	(1)	-	no signal

Table III.1 (as presented in part I) – Raw and B_{ACQ}-subtracted signals from kinetic-mode experiments with all the probes (two separate runs, between which data are not comparable; recall too that the fluorophores of 15a (diClHPQ) and 1b (HPQ) have different signal absorption/emission so cannot be intercompared directly either.

In Run 1, given the considerations of the tosylate counterion (see section I.5.2.2), the appropriate concentration-dependent F_{NoEnz} figure for probe **1a** was calculated by linear fit based on the datapoints for F_{NoEnz}(10 µM) and F_{NoEnz}(100 µM), eg. estimated F_{NoEnz}(59 µM)=10400 RFU. Although this background was tiny in an absolute sense, its concentration dependence substantially reduced the sensitivity of that probe when computing *in vivo*-style sensitivity (no t_0 subtraction; cf. relative change

model), compared to the concentration-independent background seen for the probe **1b** which allowed higher sensitivities.

In Run 2, I used the 100 μM no-enzyme controls in calculating the ratios of signal to control for all probe concentrations. Note the enormous sensitivity that is given, even if I might have expected a small reduction in the no-enzyme control signal with decreased probe concentration for probe **15a** and therefore even higher sensitivities. Clearly, the limiting factor in detection sensitivity in these experiments is the attainment of the lowest possible *reading* in the proagent well; in my interpretation this has more to do with sample preparation (dust etc) than with any small residual fluorescence (clearly there was none above the noise level).

Recall that literature data is almost universally t_0-subtracted: the no-enzyme control $F_N(t=0)$ values are subtracted from all F_{PLAT} values during calculation of signal/control ratios. For example, Urano *et al.*[9] subtract their *t*=0 fluorescence readings, so as to calculate a signal-to-time-zero value of only 76, after only 8 minutes with 5 U enzyme and only 1 μM probe (as I outlined in section 3.1.1, I consider that this is because with their 5 hour hydrolysis halflife, 2% of their probe has already hydrolysed during this time, so their true signal-to-control ratio is less than 25, and dropping constantly). Therefore I also considered calculating t_0-subtracted values to compare my unoptimised results with literature values. In Run 1, I observed F_N(**1a**, 10 μM)=4000, F_N(**1a**, 100 μM)=14000, and F_N(**1b**)=8000 RFU (concentration independent). Therefore the calculation for **1a** could be done (controls now adjusted to F_{NoEnz}(10 μM)=100 RFU, F_{NoEnz}(100 μM)=1700 RFU, linear correction as before. However, for **1b**, subtracting the t=0 value would lead to negative fluorescence values for the control signal (ie. zero signal), which would imply infinite sensitivity. Therefore any signal observed above F_N(**1b**)=8000 RFU would give 'white-on-black' detection (entirely macroscopically off-ON, zero background); so I marked the corresponding signal-to-control values as ∞ to indicate this infinitely high sensitivity ratio. In Run 2, machine stabilisation took longer, so I chose the minimum stable no-enzyme control signals, F_N(**1b**)=6000 RFU and F_N(**15a**)=15000 RFU, over the whole range. As can be seen, my t_0-subtracted signal-to-control values range from 500 to 43000 after more than 11 hours of incubation and using 2 orders of magnitude less enzyme; this is unquestionably superior sensitivity compared to what can be obtained with literature-reported soluble relative change probes. Even expressly avoiding such t_0-subtraction during signal treatment, the sensitivities for the present probes are still far superior, which illustrates the sensitivity that these probes can achieve even over long timescales at *in vivo* conditions of enzyme concentration, without 'fudging' the signal processing (see sections I.3.1-I.3.2) .

Table III.2 presents sensitivity results obtained with the scanning mode, on probe **1a**, to illustrate the very close similarity of the results obtained by the two data acquisition methods, over different timescales and on different experimental runs. This confirms the practical utility of the probe system as a reliable method for ultra-sensitive detection of aminopeptidase activity across a range of (activity/probe-concentration) values, in different applications.

Probe 1a; Scan mode	**$A_X(t)$ (RFU)**		**$F_X(t)/B(t)$ (Signal)**		**$F_X(t)/F_{NoEnz}$; no t_0 sub (Signal/Control)**	
Well Series	**24 h**	**40 h**	**24 h**	**40 h**	**24 h**	**40 h**
Water	72839	73029	(0)	(0)	(0)	(0)
81 μM, NoEnz	78944	82526	0.08	0.13	(1)	(1)
209 μM, NoEnz	86302	94865	0.18	0.30	(1)	(1)
209 μM	4900858	5042099	66.3	68.0	368	226
81 μM	2415589	2462269	32.2	32.7	403	250
21 μM	511910	543580	6.03	6.44	est. 300	est. 161

Table III.2 (reproduced from Part I) – Raw and B_{ACQ}-subtracted signals from a scan-mode experiment with 1a. Brackets indicate that the data point was used to calibrate the values of others in its column, so is not statistically significant.

3.3.4. Conversion of signals F(t) to concentrations $C_{PPT}(t)$

Firstly, I examined the relation $F_{plateau} \leftrightarrow C_{plateau}$, which was **the only reliably-estimable signal↔concentration relation**. The F(t) data are scaled differently between different runs, and between different fluorophores in the same run; so I normalised them obtaining intercomparable values F_{NORM} by (1) dividing all fluorescence plateau values by the plateau for **1b** at 100 μM in the same run (both run 1 and run 2 used this probe at 100 μM) to normalise signals to a standard amount of **HPQ** emission, then (2) adjusting the for the different **HPQ** and **diClHPQ** emission intensities by multiplying **diClHPQ** intensities in run 2 by a factor of

$F_{PLAT}(\mathbf{HPQ}) \times C_{PPT}(\mathbf{HPQ}) / [F_{PLAT}(\mathbf{diClHPQ}) \times C_{PPT}(\mathbf{diClHPQ})] = 0.81$.

Then, combining both sets of experimental data I plot C_{PPT} against F_{NORM}: as expected, given my absorption/inhomogeneity considerations of section I.3, the dependency is nonlinear (Figure III.5).

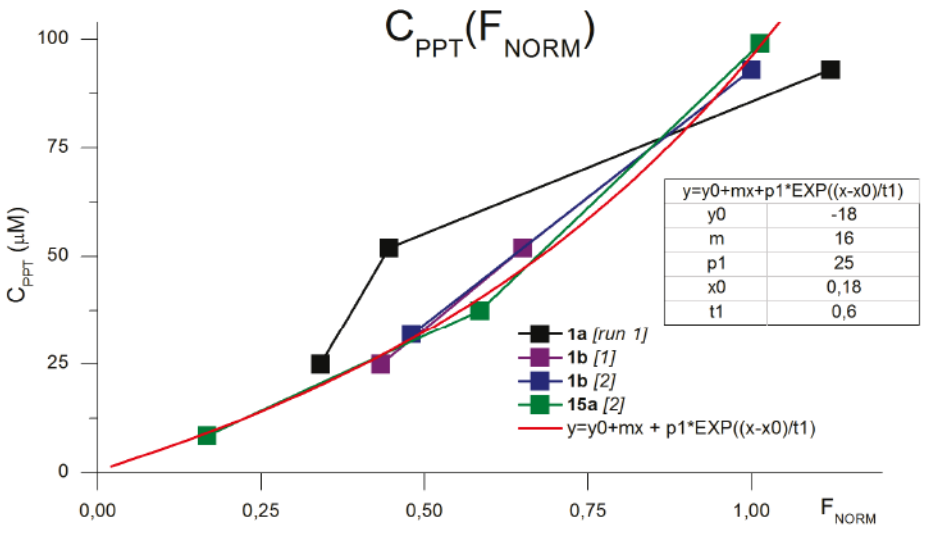


Figure III.5 (from part I) – Concentration dependence of the ratio of plateau fluorescence signal to precipitated concentration, with my best-fit bimodal curve (red).

Somewhat unexpectedly, the **diClHPQ** and **HPQ** fluorophores gave such similar results that I combined them for a single fit. I had actually expected a stronger nonlinearity of signal↔concentration response for the **diClHPQ** fluorophore than for **HPQ**, due to both its stronger

absorbance (ε), and to its probably smaller average particle size (thus lesser effective inhomogeneity and greater A/W, cf. section I.3). This probably indicates that both ε and A/W are so similar as makes no difference for these two very similar fluorophores.

Then, to fit this data, I expected a multimodal relation with two major features. Limit Region (1) would be the low-concentration, weak-absorption regime where k=C/F is concentration independent; recall that the spontaneous hydrolysis shown in Figure I.26 supports this because we expect a linear C(t) from hydrolysis at high probe concentrations, and a strongly linear signal is indeed seen at low concentrations of precipitated **diClHPQ** (est. up to 1.4 μM), so in this region C=kF. Intermediate Region (2) would be an intermediate absorption regime where k decreases with C_{PPT}, with an unknown dependency. Limit Region (3) would be the zone of the strong absorption limit, where no additional signal is seen despite precipitating more fluorophore, so the C(F) curve becomes vertical. Yet because k(**HPQ**@371 μM) is approximately the same as k(**HPQ**@93 μM), I conclude that this limit is not closely approached in the concentration domain of interest (<150 μM).

Therefore in Figure III.5 I have added a bimodal (linear+exponential) fit that (a) reproduces linearity at low concentration (<12 μM), (b) smoothly incorporates increasing nonlinearity as the concentration increases. Actually, predicting what dependency such a trendline should have (quadratic, exponential,...) was irrelevant in view of the experimental uncertainties and assumptions; a functional fit was what was important: the absolute values of the concentrations predicted are not actually as important as their trend. In performing my fit, I ignored the abnormally low k(C_{PPT}) series of **1a** at concentrations below 80 μM in run 1. Because for kinetic comparison I only need to model concentrations below 100 μM, I have not extended my fit beyond that. The signal↔concentration relation for **1a** at 371 μM is therefore calculated by self-normalisation, which is justified because the region of kinetic interest (C_{PPT}>40 μM) is at sufficiently high concentration that the low-signal correction is not needed. Using my fit equation, I now plot $C_{PROC}(t)= C_{PPT}(t)+K_S$ for the different probes once their signal-generating regions are attained (Figure III.6).

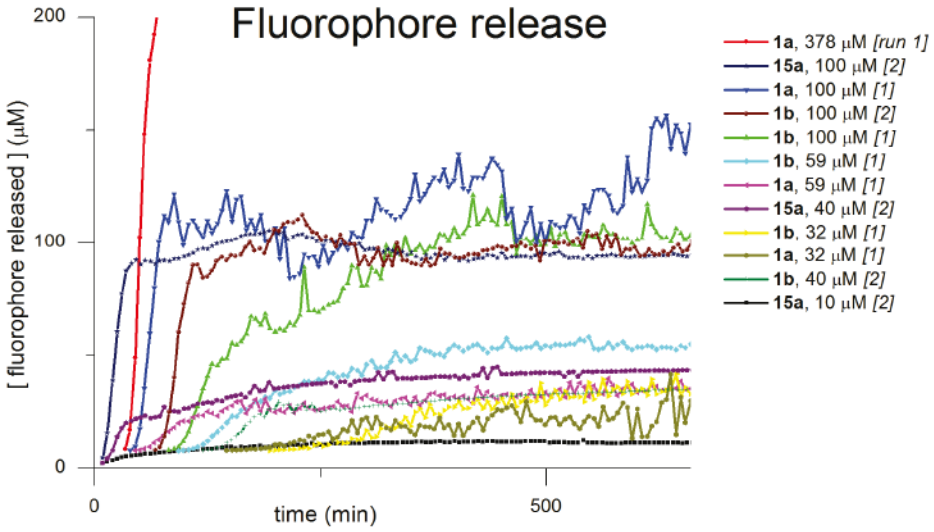


Figure III.6 – After much data processing and analysis, the probes' fluorophore release curves can be plotted for proper kinetic analysis. Raw data are used without smoothing, hence the large scatter superimposed on the true values.

Note that this plot is expected to conform roughly to the fluorescence signal obtainable with these substrate-spacer pairs but releasing a soluble fluorophore of comparable leaving-group capacity. Also note that these curves are based on the assumption that precipitation was an instantaneous process; the true curves for soluble fluorophores would therefore likely be *shifted to the left* (more rapid signal generation) by a factor compensating for the time of onset of precipitation (see section I.5.5).

3.4. Enzyme test analyses: kinetic parameter determinations

3.4.1. Kinetic analysis towards k_{CYC}

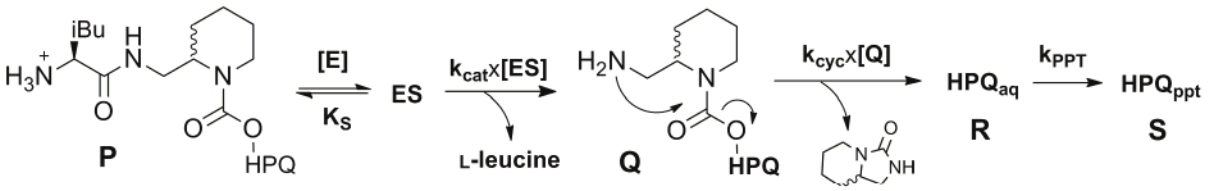


Figure III.7 (from part I) – The steps involved in signal generation, as illustrated for probe 1a. An effective concentration of precipitated fluorophore [S] is generated by enzymatic hydrolysis of P to give intermediate Q which cyclises to release fluorophore into solution; once the fluorophore's accumulated concentration R is greater than KS, precipitation to S is irreversible.

Recall that the signals for probes **1a** and **1b**, at the same concentration, reached the approximately the same fluorescence plateau values with similar-shaped signal response curves. This is expected if enzymatic hydrolysis is complete; therefore I concluded that the diastereomers of **1a** do not influence the enzymatic processing step. Therefore:

(1a) I assume simple monomodal Michaelis-Menten kinetics for **P→Q**, ie $-d[P]/dt = v_{MAX} \times [P]/(K_M+[P])$ for each of the spacers; I also assume that the cyclisation Q→R is unimolecular.

This also assumes simple enzyme kinetics (eg. no feedback inhibition of the enzyme); my preparation of the enzyme solution at least an hour before testing was also meant to ensure no 'enzyme activation time' was needed. Therefore both (-dP/dt) and the rate Q→R decrease monotonically with decreasing P or Q (respectively). Note that the exact timecourse solution to enzymatic processing of concentration P_0 by Michaelis-Menten kinetics results in a P(t) which is indistinguishable (~1%) from the linear approximation until $P(t)=4\times K_M$, and at $P(t)=2\times K_M$ the error is only 30%.[13]

I define $\mathbf{t_{LIN}}$ which is the time at which the [R+S](t) first reaches its maximum gradient m_{MAX}, with $C_{LIN} = [R+S](t_{LIN})$. I always observed that there were a minimum of 3 nearly collinear data points defining m_{max} (cf. linear regions marked in red on Figure III.8).

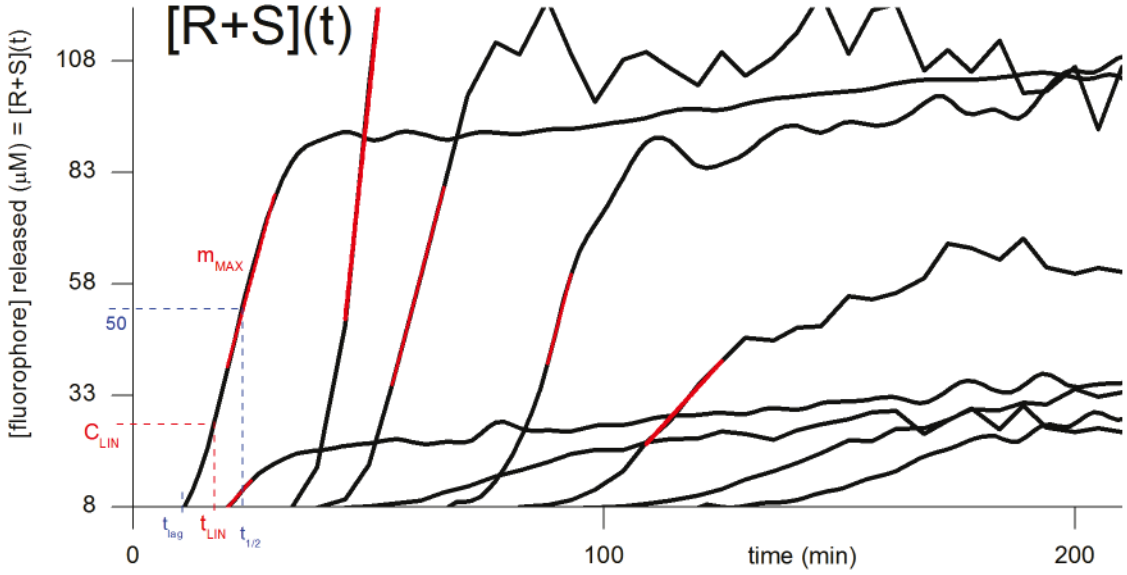


Figure III.8 – Magnified detail of Figure III.6, showing the initial stages of the kinetic experiments. Red lines indicate sample linear fits giving mmax as discussed below; sample kinetic parameters tlag, tLIN and t1/2 are shown for one curve.

(1b) I then correlated the linear region with the occurrence of a steady state concentration of a *single* species which is rate-limiting for the increase of signal. I had initially assumed that precipitation (R→S) was not, in general, strongly rate-limiting; although this assumption will be shown to be flawed when m_{max} is large, analysis will support the general conclusion that the strongest rate-limiting process is the cyclisation **Q→R**, and that deconvolution of the kinetic contributions P→Q and Q→R from the crystallisation kinetics of R→S will be possible by selecting this linear region.

I also define the signal onset time $\mathbf{t_{lag}}$ between the start of the experiment and the observation of a definite fluorescence signal (ie. K_S has been passed so [R+S](t) data are available), the time $\mathbf{t_{1/2}}$ when [R+S](t) reaches half of its maximum value, and the time $\mathbf{t_{C=0}}$ which is the t-intercept of the linear regression line. The determination of $t_{1/2}$ incurs much less t-error than trying to identify the moment when the plateau is reached, because the quasilinear region with large signal gradient and which contains the half-maximum concentration is easier to identify than is the plateau, which has relatively

large data vertical scattering superimposed on the signals, especially at high [R+S]. Note that $t_{1/2}$ also gives an upper bound for the *combined* halflife of (enzymatic processing + cyclisation), ie. a *large* upper bound for the individual halflives. These initial kinetic parameters are now tabulated.

Kinetics: before t correction	**t_{lag} (min)**	**$t_{C=0}$ (min)**	**t_{LIN} (min)**	**C_{LIN} (µM)**	**m_{max} (µM/min)**	**Q_{LB} (µM)**	**P_{UB} (µM)**	**t_{LB} (min)**	**t_{UB} (min)**	**$t_{1/2}$ (min)**
RUN 1:										
1a, 378 µM	29	40	45	50	10	(400)	(-70)	7.2	28.3	65
1a, 100 µM	39	52.5	55	10	4.0	(210)	(-120)	30	37.3	58
1a, 59 µM	39	32	71	13	0.33	10	36	-107.8	17.8	136
1a, 32 µM	104	nd	184	9	0.13	nd	nd	nd	nd	259
1b, 100 µM	66	90	112	27	1.2	(107)	(-34)	28.7	60.2	146
1b, 59 µM	93	98	125	12	0.45	44	3	-6.1	77.4	162
1b, 32 µM	168	nd	216	11	0.05	nd	nd	nd	nd	312
RUN 2:										
15a, 100 µM	0	10	14	16	4.1	41	43	-10.4	-0.4	22
15a, 40.5 µM	0	11	17	6	1.0	11	23	-23.5	-1.5	33
15a, 10 µM	0	nd	14	3	0.14	nd	nd	nd	nd	32
1b, 100 µM	67	76	85	30	3.5	(270)	(-200)	56.4	65	91
1b, 40.5 µM	120	125	151	12	0.45	(56)	(-28)	61	104.4	167

Table III.3 – Uncorrected values of the kinetic parameters t_{lag}, t_{LIN}, m_{max} and $t_{1/2}$, and the kinetic quantities calculated from them, based on the assumption of instantaneous precipitation. Gradients for probes 1b and 1a at 32 µM and probe 15a at 10 µM were considered too approximative to be relied upon for further calculations (nd = not determined), although that m is included to illustrate the trend.

I cannot explain why the gradients m_{MAX} should be so different between runs 1 and 2 for **1b** at 100 µM, except that the crystallisation kinetics may have been different for some reason; note however that both the signal onset time and C_{LIN} values are remarkably consistent. The observation that $t_{lag}<t_{C=0}$ for **1a** at 59 µM can be approximately taken to show that the underlying linear region with global maximum gradient (not of signal generation, but of fluorophore release) occurred before precipitation began.

(1c) I assumed that the kinetics of R⟶S can be treated as essentially instantaneous, *for the purpose of calculating kinetic parameters, once an appropriate correction for precipitation time is made.*

I justify this by assuming that both enzymatic processing and cyclisation are substantially slower than precipitation in the linear region of interest. I will show that this assumption gives relatively coherant results when the time dependence of the raw data is compensated for the accumulation of a *tolerated amount of supersaturation*, on top of the compensations for the solubility constant and for pre-spiking which have already been performed. This consideration of supersaturation, and an initial evaluation of its steady-state behaviour, has already been introduced in the discussion of Figure I.26 (recall the smooth change between no-signal and precipitated regimes, rather than the abrupt switchover expected if there is no supersaturation). The correction I will perform is the simplest one possible: linear subtraction of a time $\mathbf{t_{corr}}$ from the experimental times t. I considered that more complicated compensations would involve too many assumptions to be justified, but I justify this one as it delivers

a more coherant kinetic analysis as simply as possible. Notably, after t_{corr} subtraction, the values of k_{CYC} deconvoluted from the time-adjusted data can be reconciled with the time-adjusted value of $t_{1/2}$, which is a new upper bound for the (processing+cyclisation) coupled halflife.

(1d) For $P_0 >> (K_S, K_M)$, the linear signal region starting at t_{LIN} *should* correspond to a time when $P(t_{LIN})>2\times K_M$ if precipitation is instantaneous. This would enable *reasonably* accurate modelling of the enzymatic rate for $t<t_{LIN}$ as a constant, A.[13] Then, *in this region* $t<\sim t_{LIN}$ *only,* applying (1b) leads to $A=m_{max}$, and the equations

$\rightarrow dP(t)/dt \geq -A$, so $P(t)\leq P_0-t\times m_{max}$; so the upper bound for $P(t_{LIN})$ is $P_{UB}=(P_0-t_{LIN}\times m_{MAX})$ and the lower bound for $Q(t_{LIN})$ is $Q_{LB}=m_{max}\times t_{LIN}-C_{LIN}$.

$\rightarrow dQ(t)/dt \leq A - k_{CYC}\times Q(t)$, so $Q(t) \geq m_{max}/k_{CYC}\times(1-e^{-kt})$;

$\rightarrow d[R+S](t)/dt = k_{CYC}\times Q(t)$, so $[R+S](t) \geq m_{max}\times(t-1/k_{CYC})- m_{max}/k_{CYC}\times e^{-kt}$.

If $Q_{LB} > P_0-C_{LIN}$, this indicates that both cyclisation and precipitation, but not enzymatic processing, were the strongest rate-affecting factors in the linear region, and I assume that cyclisation was much more strongly rate-determining. Note that this applies to *all the high-concentration datasets*, except for **15a**. If $Q_{LB} \approx P_0-C_{LIN}$, this indicates that cyclisation was the strongest rate-limiting factor; note that this applies to the **1b** run at 59 μM. If $Q_{LB} << P_0-C_{LIN}$, this indicates that either or both enzymatic processing and cyclisation, but *not* precipitation, were rate-limiting; I assume that cyclisation was more strongly rate-determining (this is the case that applies to both datasets of **15a**).

I therefore conclude that for all the data acquired, the cyclisation step is the most strongly rate-affecting step; therefore the kinetic information that can be best determined is therefore the cyclisation halflife, not the enzymatic kinetics, although I will attempt to determine both nevertheless.

3.4.2. Supersaturation time correction: motivation

Performing deconvoluted data analysis, by assuming R→S is instantaneous, requires accounting for a supersaturation/precipitation timelag by subtracting a constant t_{corr} from t_{LIN}. To show this, consider that physically, the conditions on the two limits are, $P_{UB}>0$ and $0<Q_{LB}<P_0$: however, Table III.3 shows that this condition is not always met. In fact, t-correction is expected to be *most needed when the concentration is high and when the fluorophore is slow to crystallise* (if all other factors are equal), because higher probe concentrations will more rapidly generate R and so precipitation kinetics will become more rate-limiting in signal generation. Pleasingly, this is exactly what is observed: **1b** and **1a** at 59 μM in run 1 do not require correction (slower-crystallising **HPQ** fluorophore, but low concentration), and nor do either of the **15a** experiments in run 2 although one was at 100 μM (faster-crystallising **diClHPQ** fluorophore). This supported the idea that this way of altering the data really did target the removal of a component stemming from crystallisation kinetics, rather than simply being an arbitrary alteration without physical justification.

So I decided to subtract values of t_{corr} from the time parameters of the data, to try to remove as much as possible the influence of this supersaturation/nucleation time (which is not of intrinsic kinetic interest, whereas k_{CYC}, and potentially, the enzymatic processing kinetic parameters, are). I considered two facts about acceptable values of t_{corr}.

Fact 1: $t_{corr} < t_{lag}-(K_S/m_{max})$, and $t_{corr} < t_{LIN}-C_{LIN}/m_{max}$. These are upper bounds for t_{corr} based on the experimentally observed signals, and determined as upper bounds because cyclisation cannot be instantaneous. The two limits correspond to the calculated minimum time to accumulate the first observed signal, and to accumulate the signal observed at t_{LIN}. It was unknown which would be the more restrictive, but by calculation, it was always the solubility restriction. The upper bound for t_{corr} is tabulated as t_{UB}. In calculating t_{UB} for the probes **15a**, t_{lag} had to be extrapolated from the first two data points because the first data point acquired already showed signal.

Fact 2: $t_{corr} > t_{LIN} - P_0 / m_{MAX}$. This is the lower bound for t_{corr} so that $P_{UB}>0$ and $Q_{LB} < P_0$, and is tabulated as t_{LB} (Table III.3).

3.4.3. Determination of k_{CYC} by supersaturation correction

Within each set of experiments with one probe in a run (eg. **1b** at 100 µM and at 59 µM in run 1), the largest of the lower bounds for t_{corr} (here, 28.7 min for 100 µM) was applied to its data set. Then its k_{CYC} was calculated, from the steady-state, large-P assumption, ie. $k_{CYC,LB} = m_{max}/Q(t_{LIN}-t_{corr})$.

Then the t_{corr} for the other data sets of the same probe was varied between its boundaries to achieve the best agreement with the initial k_{CYC}. If the k_{CYC} values did not converge within acceptable t_{corr} values, then the original t_{corr} was increased stepwise, repeating the convergence trials until the closest match was obtained across all data sets for that probe in the run.

Then the kinetic parameters t_{lag}, $t_{C=0}$ and $t_{1/2}$ were compared between the data sets to ensure they were consistent: ie, they increase as the initial probe concentration decreases (because of the lower ratio of the enzymatic rate to the concentration needing to be processed to achieve $t_{1/2}$). The exception to this would be if the enzyme were totally saturated with probe until $t_{1/2}$: then, decreasing the P_0 would decrease $t_{1/2}$, while leaving t_{lag} and $t_{C=0}$ unchanged within experimental error. This was *exactly* what was obtained for the highest probe concentration tested (comparing 378 and 100 µM of **1a**), which I took as a validation of my molecular and kinetic understanding of this system, as well as a validation of the physical justification for my supersaturation-only correction.

If the kinetic parameters were now physically reasonable, they were retained (otherwise, the original t_{corr} was again stepwise increased and the fitting process begun again). Final values retained are given in Table III.4. Note that because m_{max} was required to be known accurately to calculate the time corrections, the low-concentration data sets with low and uncertain m_{max} have been removed. Also, the inexplicably high gradient m_{max} seen for run 2 of **1b** at 100 µM would have forced fitting with t_{corr}=56.4 min and k_{CYC}=0.050. If however only its data set at 40.5 µM were considered, then applying

t_{corr}=65 min (≈t_{LB}) would result in matching exactly with the k_{CYC} calculated for that probe in run 1, so I eliminated the (**1b**, 100 μM, run 2) dataset from subsequent t_{corr} analysis. From the lower-bound estimates for k_{CYC}, upper-bound estimates for the cyclisation halflife $t_{CYC,UB}$ are calculated.

Kinetics: AFTER least t correction	**t_{corr} (min)**	**t_{lag} (min)**	**$t_{C=0}$ (min)**	**t_{LIN} (min)**	**C_{LIN} (μM)**	**m_{max} (μM/min)**	**$k_{CYC,LB}$ (min^{-1})**	**$t_{CYC,UB}$ (min)**	**$t_{1/2}$ (min)**
RUN 1:									
1a, 378 μM	17.5	11.5	22.5	27.5	50	10	0.0444	15.6	47.5
1a, 100 μM	30	9	22.5	25	10	4	0.0444	15.6	28
1a, 59 μM	9.1	29.9	22.9	61.9	13	0.33	0.0444	15.6	126.9
1b, 100 μM	28.8	37.2	61.2	83.2	27	1.2	0.0165	42.1	117.2
1b, 59 μM	37.7	55.3	60.3	87.3	12	0.45	0.0165	42	124.3
RUN 2:									
15a, 100 μM	0	0	10	14	16	4.1	0.099	7	22
15a, 40.5 μM	0	0	11	17	6	1	0.0909	7.6	33
1b, 100 μM	*56.4*	*10.6*	*19.6*	*28.6*	*30*	*3.5*	*0.0499*	*13.9*	*35*
1b, 40.5 μM	65	55	60	86	12	0.45	0.0169	41.1	102

Table III.4 (from part I) - Values of the kinetic parameters t^*_{lag}, t^*_{LIN}, m_{max} and $t^*_{1/2}$, altered to fit the instantaneous precipitation kinetic model, and the cyclisation kinetics calculated from them, including upper-bound estimate for cyclisation halflife t_{CYC}.

This procedure results in choosing the minimum possible t_{corr} values that give a mathematically coherant calculation of k_{CYC}, which is also the most inoffensive data manipulation that can be applied. Forcing the minimum of t_{corr} also results in a scientifically justifiable *lower bound* for k_{CYC}; that this is a lower bound is not only because k_{CYC} increases as t_{corr} increases, but mechanistically too because if the rates of P⟶Q or R⟶S are not effectively instantaneous after this linear t_{corr} adjustment, then their rates will contribute to the calculation of k_{CYC}. This is in fact an advantage, because the original measure of $t_{1/2}$ gives an indisputable *upper bound* for k_{CYC}. Therefore I have not only calculated a value, but obtained a very strong confidence interval for it, which was necessary given the rather hair-raising data treatment that has had to be performed since acquiring the original data, A(t).

Although I had rejected the (**1b**, 100 μM, run 2) dataset from this analysis due to its anomalous m_{max}, so direct intercomparison of **15a** and **1b** within run 2 was ruled out, importantly, run 1 generated coherant and reasonable data, and strictly compared the two spacers because the leaving-group phenol was identical.

I then used these data (lower-bound cyclisation and enzyme kinetics values, and the indications of their relative ratios) in a *full-fit multiparameter kinetic model* to determine best-fit values of *all* the relevant kinetic parameters: k_{CYC}, K_M, v_{max}, and precipitation kinetic quantities: see section I.5.5.2.

3.5. Evaluation of probe performance

3.5.1. Preparation for comparison to the group's previous designs[14]

Kinetic Parameters Cannot be Compared: The group's previous designs for precipitating-fluorophore probes, **34a** and **34b**, showed signal onset approximately 57 seconds after enzyme

addition, under apparently comparable experimental conditions (pH 7.4, 25°C, 38 µM probe, 20 $mU \cdot mL^{-1}$; Figure III.9). Below, I show that this is unrealistic for several reasons, and the signal data cannot be reliably compared. Although I hoped to correct partially for the previous work's heavy pre-spiking (see below) by adding a minimum of 168 s to the results, the previous data is still graphed at t=0. This cannot be physical, because it takes time to add the enzyme into each well and then launch the acquisition. In my experience, at least 5 s is needed per well to add enzyme and then another 15 s to launch acquisition, with multiple wells to be used for replicates. I cannot estimate how much time was taken, because the number of wells used is not specified. But I can be sure that what is represented as t=0 is at least t=60, if not t=500 (minimum time to launch acquisitions with about 40 wells used for replicates). This means that faithful comparison of *signal onset time* with this prior data is impossible, because a true time correction cannot be applied. Nor can I conclude on the accuracy of the K_M and v_{max} reported in this prior work (98 µM and 3 $\mu M \cdot min^{-1}$ respectively) because only kinetic data for one substrate concentration was graphed: seeing as no procedure for K_M determination was explained, I conclude that they were unaware of the pre-spiking of the solution (which not only changes the signal onset time greatly, but undoubtedly lowers the concentrations that were assumed, among other problems), and unaware of the non-linear signal↔concentration correlation, and probably equated the observed signal with instantaneous spacer immolation. Therefore I cannot correct any of their kinetic determinations if this is necessary, as I suspect.

Stability must be, but can be, substantially corrected: To compare these probes' characteristics of hydrolytic stability to those from the group's previous, **HPQ**-releasing peptidase probes **34a** and **34b** (Figure III.9)[14] was initially confusing. The previous probes' no-enzyme controls were shown to start generating spontaneous signal after 33 minutes from reported concentrations of 38 µM; so using my calculated K_S(**HPQ**)=7 µM, this would indicate that 7 µM out of 38 µM are hydrolysed in 33 min, predicting a halflife of 112 minutes. However, considering the subsequent linear increase in signal from only 0 to 10% of the plateau level of the enzymatic reaction (where C_{PPT}=31 µM) over a period of 120 min, this instead implied spontaneous hydrolysis of (0.1×31)=3 µM out of 31 µM probe remaining during these 120 min: ie, a halflife of 17 hours instead. As was seen in section III.3.3.1, this second procedure does not involve the assumption that there is no pre-spiking. The disagreement between these predictions was too severe to ignore, and will clarify the unusual signal generation that was seen for those previous probes.

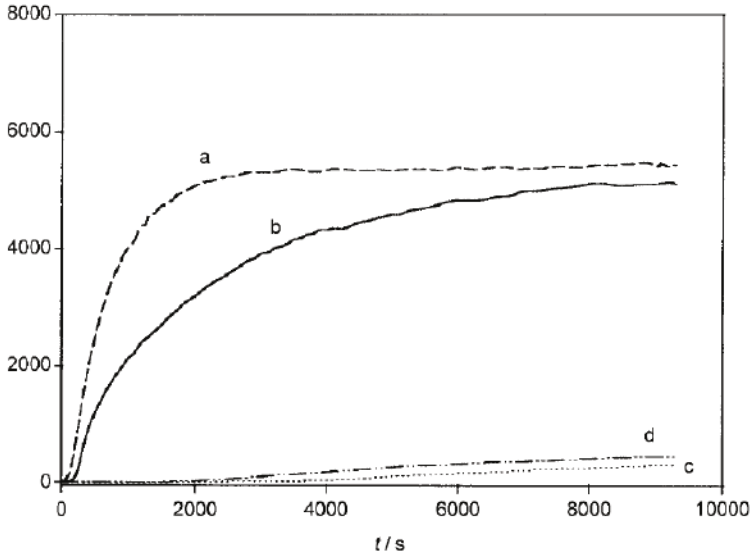


Figure III.9 – Enzyme testing data reprinted from the previous probe designs reported in the Hasserodt group[14]: HPQ-releasing peptidase probes 34a (b, d) and 34b (a, c) at 38 µM. Curves a and b with enzyme; curves c and d are no-enzyme controls.

My interpretation of the discrepancy is that the observed signals are trustworthy data, but the reported conditions are wrong, which leads to the disagreement in the halflife analysis: the substrate is not at 38 µM initially, but (a) the solution must have been effectively **pre-spiked** with **HPQ**, and (b) it is likely that the total initial concentration of (**HPQ** + probe) was substantially less than 38 µM, so the plateau signals do not correspond to 38 µM HPQ precipitated. These conditions would successfully reproduce several features of the observed signal kinetics which would otherwise be inexplicable under my off-ON precipitating model (see sections I.3 and I.5). Note that this evaluation does not intend to imply scientific dishonesty in the previous experiments, but instead highlights the extreme care that must be taken (and in the current work, was taken) with these unusual precipitating probes to obtain and report reliable results and conclusions.

I now recalculate for the prior work, the parameters which I require to compare stability, according to my interpretation of their data.

(1) **Raw data**: the plateau fluorescence values in that work are approximately 5540 RFU; the linear spontaneous signal regime starts at t=2000 s; the linear regime shows a linear increase of 520 RFU in 7000 seconds, qualitatively and mechanistically similar to the signal I observed in my own experiment studying spontaneous hydrolysis of **15a** (Figure I.26).

(2) **Facts**: Extending the linear regime to t=0 gives a y-intercept of -150 RFU; spontaneous hydrolysis is a strictly unimolecular proces with quasilinear behaviour at near-constant probe concentration, so this linear extrapolation is reliable, especially at low signal amounts (weak absorption regime). Therefore, this is unquestionably the hypothetical fluorescence that would be generated by the concentration of probe C_H which was spontaneously hydrolysed during the initial 33 min. At t=33 min, [**HPQ**]=K_S, so at t=0, [**HPQ**]$_0$=K_S-C_H.

(3) **Conditional calculation 1**: *Assuming that the plateau value attained really corresponds to (38-K_S)=31 μM precipitated*, then this implies that C_H=(150/5540)×31=0.84 μM, and using my calculated value for the K_S(**HPQ**) indicates that [**HPQ**]$_0$=6.2 μM.

(4) **Proposed explanation**: This pre-spiking may be a direct consequence of (i) the prior sample preparation procedure coupled with (ii) the hydrolytic instability of the prior design. The given sample prep procedure is: "a 38 μM solution of substrate... was prepared (sonication was used to speed up dissolution)... [to a] 95 μL aliquot of the substrate solution was [added a] 5 μL aliquot of concentrated leucyl aminopeptidase stock solution... and data acquisition was launched."[14] My interpretation is that the substrate solutions prepared were in fact constantly hydrolysing during sonication, and therefore the supernatant in the substrate stock was essentially saturated in **HPQ**. Spontaneous hydrolysis from 38 μM to **HPQ** saturation, ie (38-K_S)~31 μM, would only take 260 minutes, and there is no indication for how long sonication was performed. Furthermore, the dilution from 95 to 100 μL actually becomes significant: the total **HPQ** concentration in the well is really at most 36 μM. (Note that if the stock *were* saturated with **HPQ** but none had yet precipitated, then the solutions examined would have (95/100×K_S)=6.6 μM **HPQ** once the enzyme aliquot was added, which is remarkably close to the predicted value and argues strongly for the pre-spiking interpretation).

(4a) **Pre-saturation calculation 2a**: If my initially-calculated figure of [**HPQ**]$_0$=6.2 μM were correct, then I could conclude that the stock solution had [**HPQ**]$_{0,stock}$=(6.2/0.95)=6.5 μM, therefore [**34a**|**34b**]$_{0,well}$=(38-6.5)×0.95=30 μM, and I would estimate a sample preparation time of 240 min from addition of buffer to the start of the monitoring, to give this amount of hydrolysis: a not unreasonable figure. Recalculating the spontaneous hydrolysis rate revising the plateau precipitated concentration to (30+6.2-K_S)=29.2 μM↔5540 RFU, with +520 RFU hydrolysed in 7000 s from an initial concentration of 29.2 μM at the start of precipitation, gives a halflife of τ(**34a**|**34b**)~13.3 hours.

(4b) **Post-saturation calculation 2b**: I then wondered if any hydrolysed and precipitated **HPQ** would perhaps have been ignored during sample preparation as being simply undissolved starting material, and would have been avoided as only the supernatant was aliquotted into the well plates (although a quick check with a UV lamp would have allowed confirmation of whether any were present). I then calculated this through to see what effects it would have. Assuming dissolution for perhaps 360 minutes gives **HPQ** saturation as before (so [**HPQ**]$_{0,th}$=6.6 μM), and [**34a**|**34b**]$_0$=(38-8.25)×0.95=28.3 μM. Then the signal plateau corresponds to 35 μM precipitated, so reworking calculation 1 gives C_H=0.74 μM and [**HPQ**]$_{0,calc}$=6.3 μM, which is still in close agreement with the estimated saturation calculation. However, I have no basis for choosing this over (4a): I simply performed it to show that it also gives consistent results and therefore is a not unreasonable scenario.

(5) **Alternative explanation**: Similar pre-spiking could be observed by having used an impure probe sample heavily contaminated with **HPQ**.

(**6**) **Values retained for stability comparison to the present work**: I retain the spontaneous hydrolysis halflife I calculated in (4a), τ(**34a**|**34b**)~13.3 h. Considering the near-linear region of maximal signal response, which has gradient of 7 $RFU \cdot s^{-1}$, applying the conversion factor (29.2/5540)=5.27×10^{-3} $\mu M \cdot RFU^{-1}$, and calculating a correction for calculated pre-spiked concentration of 6.2 μM, suggests that (a vast underestimate of) at least an extra ($6.2/(5.27 \times 7 \times 10^{-3})$)=168 s must be added to the results' timescale, so the signal onset can be no earlier than 225 s. However, considering the form of the onset of signal seen with my **HPQ** probes (Figure III.8), and due to the convolution of heavily rate-limiting cyclisation kinetics, I considered that this correction factor was an underestimate of the true value by at least a factor of five or ten (see kinetic analysis, section 3.3.4).

3.5.2. Towards a practical enzymatic detection threshold

Research concerned with probes operating on an increase in fluorescence after enzymatic activation commonly cites a 'detection limit' for the protein targeted, which is a function of the ratio between control and active fluorescence signals at a fixed time t. In the present case, the situation is altogether different: the principle of a true binary off-ON system is to give an *arbitrarily low* detection limit for enzyme activity, by allowing longer incubation to compensate for lower activity without any false positive signal masking the results. Our system provides the first demonstration of such a fluorogenic aminopeptidase probe (see sensitivity models in section I.3.1).

Detection is therefore achieved as soon as the released fluorophore concentration (dependent on the enzyme activity used, the rate of substrate processing, initial probe concentration, and whether the solution is pre-spiked) passes the nucleation threshold specific to that fluorophore (dependant also on the temperature, pH, other dissolved species, etc) and the fluorescence passes the signal-to-noise threshold of the fluorimeter setup used (dependant on the acquisition background and acquisition sensitivity, the cleanliness of the setup or the autofluorescence of the sample, the purity of the probe sample used in the experiment, etc). Therefore a 'detection limit' will be a function of all these parameters, as well as a dynamic function of the time during the experiment. A better approach is therefore continuous monitoring, to be sure to finding the window of binary detection (as per Figure I.9) and ensuring ultrasensitive detection even of enzymatic hydrolysis rates below spontaneous hydrolysis rates. A fixed-time detection limit using this model aminopeptidase and model fluorophore in a model setting would not be strictly transferable once any of these parameters are changed, so will not necessarily provide useful extra numerical information to the reader interested in applying this novel, binary off-ON system to *eg.* prodrug design, let alone to probes constructed with other substrates, or using any other peptidase types or concentrations, types of fluorophore, temperature, pH, test media conditions, or detection protocol.

What *can* be transferred to any application, is that this off-ON precipitating system delivers, for the first time, a design for an *arbitrarily low detection limit* (in the context of probe design; of course, for applications to *eg.* prodrugs, such a figure is not meaningful) which will depend exclusively on the

nature of the application and implementation (choice of the parameters listed above, notably, fluorophore and test media conditions, protein activity, and probe concentration: see section I.5.1). In this sense, determining a detection limit for the model probe system I explored misses the greater, conceptual importance of the information already supplied.

However I give brief sample calculations as follows, keeping in mind that, while protein activity depends on a complex set of factors, I *already* chose the peptidase concentration to be among the lowest used in the literature I have seen, being for example three thousand times lower than that used in Urano's work[9]. The fact that the probes can generate signal in under a minute (**15a** extrapolation) with this low LAP enzyme activity already highlights the excellent sensitivity of our system; a sample 'detection limit' serves only to reinforce this conclusion, but should not form the exclusive basis of evaluation and comparison of the probe system with conceptually different, prior art systems.

Detection limits after ~24 h: Consider probe **1a** (K_S(HPQ) ~ 7 μM) at an initial concentration of 100 μM used in a 24-hour incubation with our described experimental parameters. No false positive background signal will arise, considering the halflife for spontaneous hydrolysis ($\tau_{1/2}\approx$4000 days, so it takes an est. 420 days to spontaneously hydrolyse 7 μM; see section III.3.3.2). Given the magnitude of K_S, we might arbitrarily require 12 μM to be processed to count as a highly reliable detection; macroscopically, this would deliver a fluorescence signal of *ten times* the *total* background fluorescence signal of a blank sample *without requiring the subtraction of any background at all*, but would give a signal:control ratio of easily >100 by subtracting a water-only blank, or potentially white-on-black 'infinite' sensitivity if t=0 subtraction were performed (cf. Table III.2, examine the signals from probes **1a** or **1b** at 18 μM at the end of the reaction, where ~5.5 μM has precipitated). The fitted v_{max} was 5 μM·min^{-1} with protein concentration of ~60 nm, so I approximate $v^* \approx 0.1$ μM·min^{-1} per 1 nm of peptidase. The detection limit requirement is 12 μM = $v^*\times24\times60\times$[E] at probe concentration >> K_M, therefore the probe's detection limit is [E]~**80 pM of leucine aminopeptidase**, in a noncompetitive setting. In terms of this peptidase's activity upon *this* substrate, this figure is equivalent to **8×10^{-6} U·mL^{-1} of leucine aminopeptidase.**

This is however a vast underestimation of the optimised detection limit of a probe system based on these spacers, principally because of (1) the relatively high K_S, (2) the non-optimal pH and temperature used to determine v^*, and which are therefore likely underestimates of the true v*. As to problem (1), using the lower-K_S probe **15a** at 35 μM (also high enough above K_M to ensure near-v_{max} behaviour) with similar requirement for 1.5×K_S hydrolysed leads however to around *ten* times lower detection threshold under identical cirumstances, because not only is the K_S is 5 times lower, but also because around 3% of the probe spontaneously hydrolyses during this period and therefore pre-spikes the test medium, so requiring only 0.5 μM to be enzymatically hydrolysed to first generate a signal, and only 1.2 μM to reach the threshold of 1.5×K_S, hence **8×10^{-7} UmL^{-1}**.

Experimental: Part II (Magnetogenic Probes)

4. Appendices to Part II

4.1.1. Quick recap: NMR, and MRI contrast agents

In a magnetic field B_0, the allowed projections I_Z of the spin magnetic moments of nuclei of spin I onto the field (which defines the Z axis) are quantised. For ^1H (I=½), two energy levels are allowed, separated by $\Delta E=\gamma B'$; the gyromagnetic ratio γ is a constant for all ^1H nuclei, and the local magnetic field B' is the external magnetic field B_0 modified by the chemical environment of each ^1H atom, (electronic de/shielding by neighbouring groups, etc). Thus for a known field strength, the frequency ν (= $\Delta E/h$) of the photons emitted during nuclear relaxation in a sample can be used to discriminate between the chemical environments and compounds of the emitting ^1H nuclei (NMR effect). Conversely, if the external magnetic field is applied with a 3D gradient, then these photons' frequencies can be correlated with their spatial location (MRI).[15]

The T_1 "longitudinal" or spin-lattice time constant describes the timescale for the decay of the external-field-aligned (longitudinal) component M_Z of the net nuclear spin magnetic moment vector M of a sample, from its excited state to its equilibrium value M_0. The T_2 "transversal" or spin-spin time constant describes the timescale for the loss of coherence of the vector projection of M onto the XY plane due to dephasing between spin packets of otherwise identical nuclei which are experiencing differently perturbed local magnetic fields (eg. due to inhomogeneities in the B_0 field or to intermolecular interactions). Contrast agents may affect either the T_1 or T_2 (or both) of surrounding water.[16] The generalised Solomon-Bloembergen-Morgan theory describes factors governing the relaxation of neighbouring water proton nuclei induced by paramagnetic complexes, notably the number of inner-sphere-coordinated water molecules (q) and their mean residence time (τ_M), the number of second-sphere water molecules (q_{SS}), and the complex's rotational time constant (τ_R).[15]

The closer a water molecule is to the paramagnetic metal ion, the more favourable can be its magnetic interaction with the ion's μ, whence the importance of q_{SS}; the interaction is especially favourable with directly-coordinated water molecules, hence the importance of q. As the contrast agent's concentration is much below that of water, the best transmission of the relaxing effect of the complex to bulk water requires coordinated water to interchange with bulk as rapidly as possible, while still residing long enough for favourable magnetic interaction, so τ_M is important for determining relaxivity. τ_R also affects the relaxivity by governing the timescale of the interaction between a proton's nucleus and the complex's μ-vector; in general, once τ_M is optimised, a larger τ_R (slow complex rotation) gives a larger relaxivity, especially at low fields (up to eg. 3T); however, the effect diminishes as the field strength increases and higher-field machines of the future (eg >5T) may show little to no benefit from τ_R tuning. Relaxivity (r_1) is calculated as the reciprocal of the product of the concentration c of the

paramagnetic species (in mM) with the measured water proton longitudinal time constant T_1' in its presence ($r_1=1/[T_1'\times c]$) and so has units $mM^{-1}s^{-1}$.[15,16]

4.1.2. M2 designs – pyridinylogous hemiaminal spacer, alkynyl aldehyde route

During my M2 work,[15] I had initially examined elimination spacer designs analogous to **44a** but which would have as much flexibility as possible at the critical methine position X (Figure III.10).

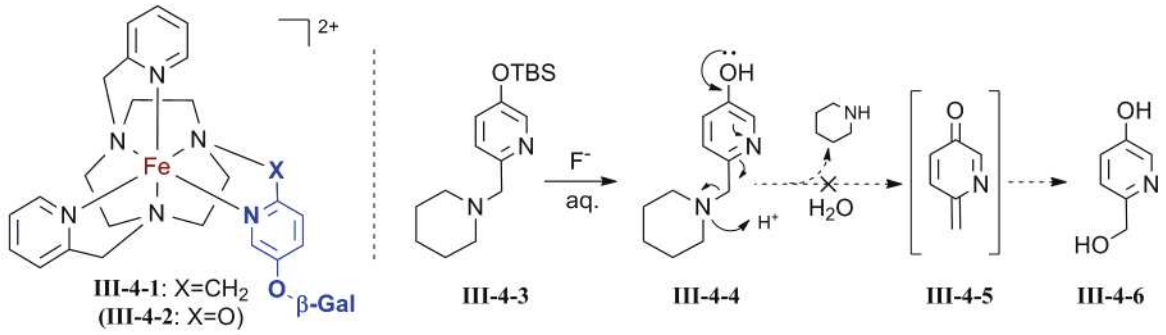


Figure III.10 - Pyridinyl spacer design and work from the M2 thesis

The pyridinylogous design **III-4-1** was tested by model compound **III-4-3**; this was deprotected in (somewhat basic) aqueous conditions but did not show fragmentation to piperidine and **III-4-6**. The basic media were not considered problematic, because although they reduce protonation of the piperidine, in the final complex **III-4-1**, protonation of the macrocyclic nitrogen is not possible during elimination anyway. Species **III-4-4** appeared stable in water for more than 2 months; as the coordination is expected to *increase* stability rather than promote elimination, this design was abandoned. A second, even more flexible and energetic possibility, the umpolung target **III-4-2**, was considered likely to be much too hydrolytically unstable, as well as synthetically challenging. Therefore the furan strategy towards targets **44** was begun (Figure III.11).

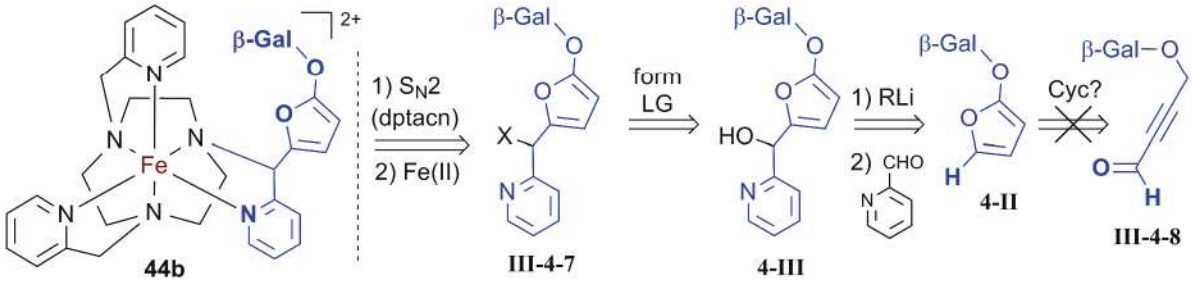


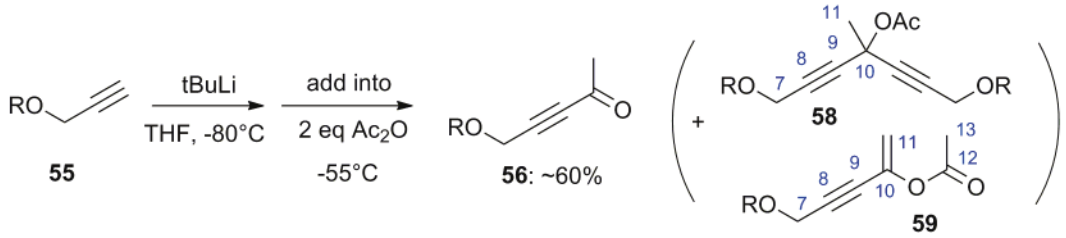
Figure III.11 – Alkynyl aldehyde cyclisation strategy from M2 work

Alkynyl aldehydes **III-4-8** were made by formylation of the propargyl ethers by the reverse quench strategy of Journet *et al.*[17]. Across extensive tests, attempts to cyclise them to compounds **4-II** (en route to **4-III** and eventually **44b**) all resulted in total decomposition, whereas a corresponding ketone cyclisation gave instead a 5-*t*Bu-2-alkoxyfuran on the first (and only) trial.

5. Compounds of Part II

5.1. Furyl cycloisomerisation – α-functionalisation

5.1.1. Acetylations



5-(2,3,4,6-tetrabenzyl-*β*-galactosyl)pent-3-yn-2-one (56c)

To **55c** (6.08 g, 10.5 mmol)[18,19] and 2,2'-bisquinoline as indicator (3 mg) in a vacuum-dried flask under argon was added THF (20 mL), then to the solution was added dropwise at -80 °C over five minutes, until the colour did not darken with further addition, *t*BuLi (8 mL, ~1.6 M, ~13 mmol) and the mixture stirred for 1 h. The brown solution was added dropwise via cannula into a solution of Ac_2O (3.08 g, 30.2 mmol) in THF (10 mL) at -55 °C. The reaction mixture becomes sun yellow during addition, and was stirred for one hour while warming to RT before quenching by addition into a rapidly stirred mixture of ice water (50 g) and ethyl acetate (100 mL). The aqueous phase was colourless after extraction with ethyl acetate (50 mL) and was further extracted with diethyl ether (2×40 mL) before the combined organic layers were washed with saturated $NaHCO_3$ (2×30 mL), H_2O (30 mL), and brine (2×30 mL), dried on sodium sulfate, filtered, and concentrated on the rotary evaporator to give 7.5g pale yellow oil. NMR of the crude mixture revealed **55c** and **56c** as the only detectable tetrabenzylgalactosyl compounds. Chromatography with 10:1 → 4:1 peth:EA eluant gave starting material **55c** (1.23 g, 2.12 mmol) and pale yellow oil **56c** (3.14 g, 5.05 mmol, 48%, or 61% brsm). No significant yield (<1% in crude) of undesired byproducts **58c** or **59c** was seen by this method. **TLC**: R_f = 0.31 (5:1 Cy:EA), R_f = 0.49 (3:1 Cy:EA); green (van), yellow (anis), positive (DNPH). **^{1}H-NMR ($CDCl_3$, 200 MHz)**: δ = 7.45-7.25 (m, 20H; 20×phenyl CH); 5.01-4.92 (m, 2H; H-1, H-3); 4.82-4.55 (m, 7H; H-2, H-4, H-5, CH_2-6, 1×benzyl CH_2); 4.46 (s, 2H, CH_2-7); 3.94-3.84 (m, 2H, benzyl CH_2); 3.61-3.55 (m, 4H, 2×benzyl CH_2); 2.31 (s, 3H, CH_3) ppm. **^{13}C-NMR ($CDCl_3$, 50 MHz)**: δ = 183.9 (C-10), 138.6, 138.5, 138.3, 137.9, 2×128.5, 128.4, 128.3, 128.2, 2×127.9, 127.7, 127.6, 102.2 (C-1), 86.9, 85.6, 82.2, 79.3, 75.2, 74.7, 73.7, 73.6, 73.5, 73.2, 68.8, 55.7 (C-7); 33.0 (C-11) ppm. **IR**: Characteristic bands: 2212[w] (*CC*-CO); 1680[s] (*CO*-CC). **(2,3,4,6-tetrabenzyl-galactosyl) bands**: 3088[w], 3063[w], 3030[w] (CH phenyl); 2915[w], 2868[w] (CH aliph); 1496[w], 1454[m] (C=C phenyl); 1150-1000 [s,br,c] (COC aliph); 731[s], 697[s] (CH). **HRMS (ESI-ToF)**: calcd for $[C_{39}H_{40}O_7Na]^+ = [MNa]^+$: 643.2672; found 643.2676.

4-methyl-1,7-bis(2,3,4,6-tetrabenzyl-β-galactosyl)hepta-2,5-diyn-4-yl acetate (58c)

Acetylation byproduct **58c** was obtained in variable yields usually under 5%, during the procedure for **56c**; its yield could be increased if the acetic anhydride were instead added to the acetylide solution (eg. 44%). TLC: R_f = 0.21 (5:1 Cy:EA); brown (anis). NMR: **^{1}H-NMR ($CDCl_3$, 200 MHz)**: δ = 7.44-7.24 (m, 40H), 5.03-4.95 (m, 4H), 4.79-4.62 (m, 10H), 4.48-4.45 (m, 8H), 3.95-3.86 (m, 4H), 3.64-3.55 (m, 8H), 1.96 (s, 3H), 1.91 (s, 3H) ppm. **^{13}C-NMR ($CDCl_3$, 50 MHz)**: δ = 168.0, 138.8, 138.5, 138.4, 137.8, 128.2, 128.2, 128.1, 128.0, 127.7, 127.6, 127.3, 127.3, **101.6, 101.5 (anomeric carbons)**, 84.0, 84.0, 81.9, 80.0, 79.1, 79.1, 74.8, 74.4, 74.3, 73.2, 73.1, 72.8, 72.8, 68.5, 63.4 (C10), 55.8, 30.3 (CH_3 Ac), 21.1 (C11) ppm. Salient IR band (distinguishes from **56c**): 1753[m] (CO_2R); (2,3,4,6-tetrabenzyl-galactosyl) bands (see **56c**). HRMS (ESI-ToF): Calc'd for $[C_{78}H_{80}O_{14}Na]^+$ = $[MNa]^+$: 1263.5446; found 1263.5481.

2-(2,3,4,6-tetrabenzyl-β-galactosyl)-5-methoxypent-1-en-3-yn-2-yl acetate (59c)

During the procedure for **56c**, beginning with 7.75 g **55c** (13.3 mmol) performing addition of the acetylide to the acetic anhydride, chromatography gave the faster-running **59c** (600 mg, 0.91 mmol, 8%) as an pale yellow oil which could be recrystallised from pentane/ethyl acetate to give colourless crystals. TLC: R_f = 0.24 (5:1 Cy:EA); yellow-brown (anis), blue (van). **^{1}H-NMR ($CDCl_3$, 500 MHz)**: δ = 7.41-7.39 (m, 2H), 7.36-7.27 (m, 18H) (20×phenyl CH); 5.27 (m, 1H, H-11b); 5.21 (d, 1H, ^2J = 1.7 Hz, H-11a); 4.83 (dd, 2H, J_{eff} = 106 Hz, ^2J = 11 Hz, benzyl CH_2); 4.78 (dd, 2H, J_{eff} = 160 Hz, ^2J = 11.6 Hz, benzyl CH_2); 4.74 (dd, 2H, J_{eff} = 27 Hz, ^2J = 11.2 Hz, benzyl CH_2); 4.57 (d, 1H, $J_{1\text{-}2}$ = 7.9 Hz, H-1); 4.55-4.48 (m, 2H, 2×H-7); 4.42 (dd, 2H, J_{eff} = 20 Hz, ^2J = 11.7 Hz, benzyl CH_2); 3.90 (d, 1H, $J_{4\text{-}3}$ = 2.6 Hz, H-4); 3.83 (dd, 1H, $J_{2\text{-}1,\,anti}$ = 7.4 Hz, $J_{2\text{-}3,\,anti}$ = 9.2 Hz, H-2); 3.61-3.57 (m, 1H, H-5); 3.58 (s, 2H, 2×H-6); 3.55 (dd, 1H, $J_{3\text{-}2,\,anti}$ = 9.7 Hz, $J_{3\text{-}4,\,gauche}$ = 2.8 Hz, H-3), 2.10 (s, 3H, CH_3) ppm. **^{13}C-NMR ($CDCl_3$, 125 MHz)**: δ = 168.4 (Ac); 138.8, 138.6, 138.5, 137.9 (4×phenyl C); 135.7 (C-10); 128.5, 128.4, 128.3, 128.3, 128.2, 127.9, 127.6, 127.6 (20× phenyl CH); 112.4 (C-11); 101.8 (C-1); 85.7 (C-8); 82.1 (C-3); 80.3 (C-9); 79.3 (C-2); 75.1 (benzyl CH_2); 74.5 (benzyl CH_2); 73.6 (benzyl CH_2); 73.5 (C-5); 73.4 (C-4); 73.2 (benzyl CH_2); 68.8 (C-6); 56.1 (C-7); 20.7 (Ac) ppm. **^{1}H^{1}H-COSY ($CDCl_3$, 500 MHz)**: δ↔δ = 7.44-7.27 (m, 20H, 20×phenyl CH) ↔ 7.44-7.27 (m, 20H, 20×phenyl CH); 5.27 (m, 1H, H-11b) ↔ 5.21 (d, 1H, H-11a); 4.57 (d, 1H, H-1) ↔ 3.83 (dd, 1H, H-2); 3.90 (d, 1H, H-4) ↔ 3.55 (dd, 1H, H-3); 3.83 (dd, 1H, H-2) ↔ 3.55 (dd, 1H, H-3) ppm. **^{1}H^{13}C-HSQC ($CDCl_3$, 500 MHz)**: δ→δ = 7.41-7.27 (m, 20H, 20×phenyl CH) → 128.5-127.6 (20×phenyl CH); 5.27 (m, 1H, H-11b) & 5.21 (d, 1H, H-11a) → 112.4 (C-11); 4.57 (d, 1H, H-1) → 101.8 (C-1); 4.78 (dd, 2H, benzyl CH_2) → 74.5 (benzyl CH_2); 4.83 (dd, 2H, benzyl CH_2) → 75.1 (benzyl CH_2); 4.74 (dd, 2H, benzyl CH_2) → 73.2 (benzyl CH_2); 4.55-4.48 (m, 2H, 2×H-7) → 56.1 (C-7); 4.42 (dd, 2H, benzyl CH_2) → 73.6 (benzyl CH_2); 3.90 (d, 1H, H-4) → 73.4 (C-4); 3.84 (dd, 1H, H-2) → 79.3 (C-2); 3.58 (s, 2H, 2×H-6) → 68.8 (C-6); 3.61-3.57 (m, 1H, H-5) → 73.5 (C-5); 3.55

(dd, 1H, H-3) → 82.1 (C-3); 2.10 (s, 3H, CH_3) → 20.8 (C-13) ppm. **Salient IR bands**: 2242[w] (CC); 1766-1710[m,br,c] (CO_2R); (2,3,4,6-tetrabenzyl-galactosyl) bands (see **56c**). **HRMS** (ESI-ToF): calcd for $[C_{41}H_{42}O_8Na]^+ = [MNa]^+$: 685.2777; found 685.2772.

5-tetrahydropyran-2-yloxy-pent-3-yn-2-one (56a)*[20]

Synthesised similarly to **56c** in 71% yield. **^{1}H-NMR ($CDCl_3$, 200 MHz)**: δ = 4.80-4.74 (m, 1H), 4.37 (s, 2H), 3.86-3.77 (m, 1H), 3.59-3.45 (m, 1H), 2.29 (s, 3H), 1.80-1.40 (m, 6H) ppm. **^{13}C-NMR ($CDCl_3$, 50 MHz)**: δ = 184.0, 97.3, 87.8, 85.3, 62.1, 53.9, 32.6, 30.2, 25.3, 18.9 ppm.

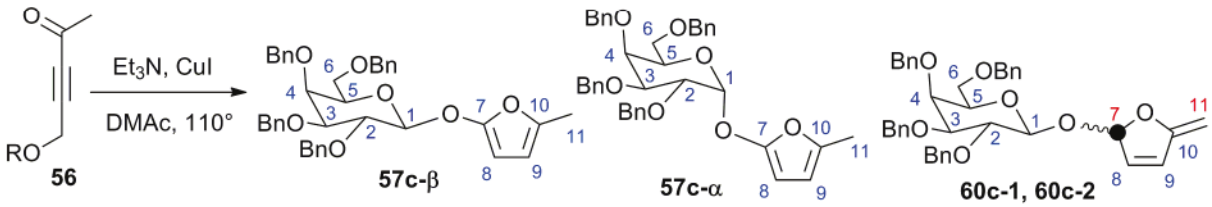
5-benzyloxy-pent-3-yn-2-one (56b)

Synthesised similarly to **56c** in 52% yield brsm. **^{1}H-NMR ($CDCl_3$, 125 MHz)**: δ = 7.38-7.24 (m, 5H), 4.57 (s, 2H), 4.27 (s, 2H), 2.31 (s, 3H) ppm. **^{13}C-NMR ($CDCl_3$, 50 MHz)**: δ = 183.6, 136.8, 128.4 (×2), 128.0 (×3), 87.2, 85.6, 71.9, 56.8, 32.4 ppm.

5-(*tert*-butyldimethylsilyloxy)-pent-3-yn-2-one (56d)

Synthesised similarly to **56c** in 25% yield. **^{1}H-NMR ($CDCl_3$, 200 MHz)**: δ = 4.41 (s, 2H), 2.29 (s, 3H), 0.86 (s, 9H), 0.09 (s, 6H) ppm. **^{13}C-NMR ($CDCl_3$, 50 MHz)**: δ = 183.9, 90.0, 84.9, 51.5, 32.5, 26.4, 25.7 (×3), -5.1 (×2) ppm.

5.1.2. Cycloisomerisations



2-(2,3,4,6-tetrabenzyl-β-galactosyl)-5-methylfuran (57c-β)

By the cycloisomerisation protocol of Kel'in and Gevorgyan,[21] to 1.09 g **56c** (1.75 mmol) in a vacuum-dried flask under argon was added CuI (87 mg, 0.45 mmol), DMAc (8 mL), then Et_3N stored over KOH (0.3 g, 3mmol); the resultant green solution was heated to 100 °C whereupon it darkens forming a black solution. It was found that the crude mixture presents an identical ^{1}H-NMR spectrum whether the reaction is performed in a sealed vial or in a two-neck flask under a condenser; this is probably only generally true for non-volatile products (not generalisable to highly volatile lower furans eg. **57a**). The reaction could be monitored by TLC (5:1 Cy:EA, anisaldehyde or vanillin developing solutions). After 16h neither starting material **56c** nor any unrecognised intermediate (eg. allene **3-IV**) could be detected in the crude mixture, and the reaction was cooled to RT and diluted with water (300 mL), then extracted with diethyl ether (6×50mL). The combined organic layers were washed with water (8×50 mL), 10% aq. LiCl (4×10mL), and brine (3×20 mL), dried on sodium

sulfate, filtred, and concentrated on the rotary evaporator to give 900 mg of black oil. A better procedure would involve simply evaporating the DMAc/Et_3N and performing a solid-deposition column at once, since hydrolytic losses during extraction could approach 50%, and losses on the column (~20%) were already difficult enough.

Chromatography with a gentle gradient from 100% → 10:1 → 5:1 peth:EA eluant gave **57c-β** (310 mg, 0.499 mmol) as a white powder (30%) which could be recrystallised from pentane/ethyl acetate or pentane/diethyl ether. Four other products (**57c-α**, **60c-1**, **60c-2** and **59c**) were isolated during chromatography; products running slower than **59c** (apparently including a partially debenzylated compound and two compounds in hemiacetal form after loss of the alkynyl moiety) were discarded. Overall recovery from this reaction was therefore 54%. All these furan/acetal species degrade if stored in $CDCl_3$ (timescale ~1 week) although the starting material **56c** is perfectly stable in this solution for a month (no spectral change). Samples should be stored as the solids in capped vials after high vacuum to remove water (to which they have a surprising affinity). Conveniently, the degradation products are relatively insoluble in Et_2O, so if a sample presents partial degradation, an ether rinse will catch the desired furans without requiring a second purification.

TLC: R_f = 0.41 (5:1 Cy:EA); green (van), green (anis). NMR spectra are reproduced on page 273 ff. **^{1}H-NMR ($CDCl_3$, 500 MHz)**: δ = 7.39-7.36 (m, 2H), 7.34-7.22 (m, 18H) (20×phenyl CH); 5.77 (m,1H, H-9); 5.34 (d, 1H, $J_{8\text{-}9}$ = 3.1 Hz, H-8); 4.88 (dd, 2H, J_{eff} = 90 Hz, ^2J = 10.7 Hz, benzyl CH_2); 4.87 (d, 1H, $J_{1\text{-}2,\ anti}$ = 7.6 Hz, H-1); 4.76 (dd, 2H, J_{eff} = 170 Hz, ^2J = 11.8 Hz, benzyl CH_2); 4.72 (dd, 2H, J_{eff} = 22.4 Hz, ^2J = 12.0 Hz, benzyl CH_2); 4.39 (dd, 2H, J_{eff} = 28.0 Hz, ^2J = 11.5 Hz, benzyl CH_2); 4.04 (dd, 1H, $J_{2\text{-}3,\ anti}$ = 9.3 Hz, $J_{2\text{-}1,\ anti}$ = 8.0 Hz, H-2); 3.89 (d, 1H, $J_{4\text{-}3,\ gauche}$ = 2.4 Hz, H-4); 3.60 (s, 2H, 2×H-6), 3.61-3.57 (m, 1H, H-5), 3.54 (dd, 1H, $J_{3\text{-}2,\ anti}$ = 9.5 Hz, $J_{3\text{-}4,\ gauche}$ = 2.6 Hz, H-3); 2.15 (s, 3H, CH_3). **^{13}C-NMR ($CDCl_3$, 125 MHz)**: δ = 157.0 (C-7); 143.0 (C-10); 138.5, 138.4, 138.3, 137.8 (4×phenyl C); 2×128.3, 2×128.2, 128.1, 127.8, 127.7, 127.6, 2×127.5 (20×phenyl CH); 106.2 (C-9); 104.0 (C-1); 86.7 (C-8); 81.8 (C-3); 78.8 (C-2); 75.3 (benzyl CH_2); 74.5 (benzyl CH_2); 74.0 (C-5); 73.5 (benzyl CH_2); 73.2 (C-4); 72.9 (benzyl CH_2); 68.5 (C-6); **13.5 (C-11)**. **^{1}H^{1}H-COSY ($CDCl_3$, 500 MHz)**: 7.39-7.22 (m, 20H, 20×phenyl CH) ↔ 7.39-7.22 (m, 20H, 20×phenyl CH); 5.77 (m,1H, H-9) ↔ 5.34 (d, 1H, H-8); 4.87 (d, 1H, H-1) ↔ 4.04 (dd, 1H, H-2); 4.04 (dd, 1H, H-2) ↔ 3.54 (dd, 1H, H-3); 3.89 (d, 1H, H-4) ↔ 3.54 (dd, 1H, H-3). **^{1}H^{13}C-HSQC ($CDCl_3$, 125 MHz)**: 7.39-7.22 (m, 20H, 20×phenyl CH) → 128.3-127.5 (20×phenyl CH); 5.77 (m,1H, H-9) → 106.2 (9); 5.34 (d, 1H, H-8) → 86.7 (C-8); 4.88 (dd, 2H, benzyl CH_2) → 75.3 (benzyl CH_2); 4.87 (d, 1H, H-1) → 104.0 (C-1); 4.76 (dd, 2H, benzyl CH_2) → 74.5 (benzyl CH_2); 4.72 (dd, 2H, benzyl CH_2) → 72.9 (benzyl CH_2); 4.39 (dd, 2H, benzyl CH_2) → 73.5 (benzyl CH_2); 4.04 (dd, 1H, H-2) → 78.8 (C-2); 3.89 (d, 1H, H-4) → 73.2 (C-4); 3.61-3.57 (m, 1H, H-5) → 74.0 (C-5); 3.60 (s, 2H, 2×H-6) → 68.5 (C-6); 3.54 (dd, 1H, H-3) → 81.8 (C-3); 2.15 (s, 3H, CH_3) → 13.5 (C-11). **Salient IR:** 1624[m], 1585[m] (C=C-C=$C_{furanyl}$: higher intensity than expected for a linear system due to rigidification in the heterocycle);

(2,3,4,6-tetrabenzyl-galactosyl) bands (see **56c**). **HRMS (ESI-ToF)**: Calc'd for $[C_{39}H_{40}O_7Na]^+$ = $[MNa]^+$: 643.2672; found 643.2697.

2-(2,3,4,6-tetrabenzyl-*α*-galactosyl)-5-methylfuran (57c-α)

During the procedure for **57c-β**, chromatography gave the faster-running **57c-α** (15 mg, 0.024 mmol) as a white powder (2%) which could be recrystallised from pentane/ethyl acetate or pentane/diethyl ether. It is thought that this product is more likely to have formed from an α-configured contaminant to the starting material **56c** rather than from epimerisation during reaction, although this could be checked with exhaustively purified β starting material. **TLC**: $R_f = 0.51$ (5:1 Cy:EA); green (anis). **^{1}H-NMR ($CDCl_3$, 500 MHz)**: δ = 7.39-7.25 (m, 20H, 20×phenyl CH); 5.79 (m, 1H, H-9); **5.36 (d, 1H, $J_{1\text{-}2,\ gauche}$ = 3.1 Hz, H-1, α-configuration)**; 5.29 (d, 1H, $J_{8\text{-}9}$ = 2.9 Hz, H-8); 4.78 (dd, 2H, J_{eff} = 191 Hz, ^2J = 11.5 Hz, benzyl CH_2); 4.90-4.74 (m = 2×dd, 4H, 2×benzyl CH_2); 4.43 (dd, 2H, J_{eff} = 33 Hz, ^2J = 11.6 Hz, benzyl CH_2); 4.18-4.12 (m, 2H, H-3 & H-5); 4.11-4.07 (m, 1H, H-2); 4.05 (s, 1H, H-4); 3.58-3.52 (m, 2H, 2×H-6); 2.19 (s, 3H, CH_3). **^{13}C-NMR ($CDCl_3$, 125 MHz)**: δ = 156.8 (C-7); 143.1 (C-10); 138.8, 138.7, 138.4, 138.0 (4×phenyl C); 2×128.5, 128.4, 128.3, 128.1, 127.9, 127.8, 127.7, 2×127.6 (20×phenyl CH); 106.3 (9); 99.4 (C-1); 86.5 (C-8); 78.8 (C-2); 75.9 (C-3); 75.1 (C-4); 75.0 (benzyl CH_2); 2×73.5 (3×benzyl CH_2); 70.7 (C-5); 68.7 (C-6); **13.5 (C-11)**. **^{1}H^{13}C-HSQC ($CDCl_3$, 125 MHz)**: 7.39-7.25 (m, 20H, 20×phenyl CH) → 128.4-127.6 (20×phenyl CH); 5.79 (m, 1H, H-9) → 106.3 (9); 5.36 (d, 1H, H-1) → 99.4 (C-1); 5.29 (d, 1H, H-8) → 86.5 (C-8); 4.78 (dd, 2H, benzyl CH_2) → 75.0 (benzyl CH_2); 4.90-4.74 (m = 2×dd, 4H, 2×benzyl CH_2) → 73.5 (2× benzyl CH_2); 4.43 (dd, 2H, benzyl CH_2) → 73.5 (benzyl CH_2); 4.18-4.12 (m, 2H, H-3 & H-5) → 75.9 (C-3) & 70.7 (C-5); 4.11-4.07 (m, 1H, H-2) → 78.8 (C-2); 4.05 (s, 1H, H-4) → 75.1 (C-4); 3.58-3.53 (m, 2H, 2×H-6) → 68.7 (C-6); 2.19 (s, 3H, CH_3) → 13.5 (C-11). **^{1}H^{1}H-COSY ($CDCl_3$, 500 MHz)**: 7.39-7.25 (m, 20H, 20×phenyl CH) ↔ 7.39-7.25 (m, 20H, 20×phenyl CH); 5.79 (m, 1H, H-9) ↔ 5.29 (d, 1H, H-8); 5.36 (d, 1H, H-1) ↔ 4.18-4.12 (m, 2H, H-3 & H-5); 4.18-4.12 (m, 2H, H-3 & H-5) ↔ 4.11-4.07 (m, 1H, H-2); 4.18-4.12 (m, 2H, H-3 & H-5) ↔ 3.58-3.52 (m, 2H, 2×H-6). **IR:** 1622[w], 1584[m] ($C{=}C\text{-}C{=}C_{furanyl}$: higher intensity than expected for a linear system due to rigidification in the heterocycle); (2,3,4,6-tetrabenzyl-galactosyl) bands (see **56c**). **HRMS (ESI-ToF)**: Calc'd for $[C_{39}H_{40}O_7Na]^+$ = $[MNa]^+$: 643.2672; found 643.2655.

2-(2,3,4,6-tetrabenzyl-β-galactosyl)-5-methylene-2,5-dihydrofuran (60c-1)

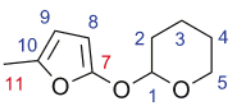
During the procedure for **57c-β**, chromatography gave the slower-running **60c-1** (100 mg, 0.16 mmol) as a powder (10%). This compound is thought to be a diastereoisomer of **60c-2** with a different C-7 configuration, based on the near-identical NMR spectra, although no crystals could be isolated for definitive structural determination. Like **60c-2**, it was observed to be unstable and HRMS could not be obtained before the compound degraded; likewise, IR spectra acquired rapidly in-house revealed

features which were no longer evident when the compound had been sent elsewhere for analysis. **TLC**: $R_f = 0.36$ (5:1 Cy:EA); green (anis), green (van). NMR spectra are reproduced on page 275 ff. **^{1}H-NMR ($CDCl_3$, 500 MHz)**: δ = 7.33-7.21 (m, 20H, 20×phenyl CH); 6.34 (s,1H, H-7); 6.31 (d, 1H, $J_{8\text{-}9}$ = 5.9 Hz, H-8); 6.15 (d, 1H, $J_{9\text{-}8}$ = 5.9 Hz, H-9); 4.82 (d overlapped, 1H, $J \sim 9$ Hz, H-1); 4.77 (dd, 2H, J_{eff} = 160 Hz, ^2J = 11.7 Hz, benzyl CH_2); 4.73 (dd, 2H, J_{eff} = 99 Hz, ^2J = 10 Hz, benzyl CH_2); 4.70 (dd, 2H, J_{eff} = 33.8 Hz, ^2J = 11.9 Hz, benzyl CH_2); **4.46 (s, 1H, H-11$_b$)**; 4.39 (dd, 2H, J_{eff} = 20.6 Hz, ^2J = 11.3 Hz, benzyl CH_2); **4.17 (s, 1H, H-11$_a$)**; 3.87 (d, 1H, $J_{4\text{-}3,\ gauche}$ = 2.7 Hz, H-4); 3.82 (apparent t, 1H, $J_{2\text{-}3,\ anti} \sim J_{2\text{-}1,\ anti} \sim 9$ Hz, H-2); 3.59-3.55 (m, 1H, H-5); 3.56 (s, 2H, 2×H-6); 3.53 (dd, 1H, $J_{3\text{-}2,\ anti}$ = 9.8 Hz, $J_{3\text{-}4,\ gauche}$ = 2.7 Hz, H-3). **^{13}C-NMR ($CDCl_3$, 50 MHz)**: δ = 162.5 (C-10); 2×138.7, 138.0 (4×C_{benzyl}); 131.7 (C-9); 129.0 (C-8); 128.7, 2×128.5, 128.3, 128.0, 127.9, 127.7 (20×phenyl CH); 105.4 (C-7); 100.8 (C-1); **84.6 (C-11)**; 82.2 (C-3); 79.3 (C-2); 75.3 (benzyl CH_2); 74.7 (benzyl CH_2); 73.7 (C-5); 2×73.6 (benzyl CH_2 and C-4); 73.4 (benzyl CH_2); 68.8 (C-6). **^{1}H^{1}H-COSY ($CDCl_3$, 500 MHz)**: 7.33-7.21 (m, 20H, 20×phenyl CH) ↔ 7.33-7.21 (m, 20H, 20×phenyl CH); 6.31 (d, 1H, H-8) ↔ 6.15 (d, 1H, H-9); 6.15 (d, 1H, H-9) ↔ 4.46 (s, 1H, H-11$_b$); 4.82 (d, 1H, H-1) ↔ 3.82 (apparent t, 1H, H-2); 4.46 (s, 1H, H-11$_b$) ↔ 4.17 (s, 1H, H-11$_a$), effective $^2J_{11\text{-}11}$ = 150 Hz; 3.87 (d, 1H, H-4) ↔ 3.53 (dd, 1H, H-3); 3.82 (apparent t, 1H, H-2) ↔ 3.53 (dd, 1H, H-3). **^{1}H^{13}C-HSQC ($CDCl_3$, 500 MHz)**: 7.33-7.21 (m, 20H, 20×phenyl CH) → 128.7-127.6 (20×phenyl CH); 6.34 (s,1H, H-7) → 105.4 (C-7); 6.31 (d, 1H, H-8) → 129.0 (C-8); 6.15 (d, 1H, H-9) → 131.7 (C-9); 4.77 (dd, 2H, benzyl CH_2) → 74.7 (benzyl CH_2); 4.73 (dd, 2H, benzyl CH_2)→ 75.3 (benzyl CH_2); 4.82 (d overlapped, 1H, H-1) → 100.8 (C-1); 4.70 (dd, 2H, benzyl CH_2) → 73.4 (benzyl CH_2); 4.39 (dd, 2H, benzyl CH_2) → 73.6 (benzyl CH_2); 4.46 (s, 1H, H-11$_b$) & 4.17 (s, 1H, H-11$_a$) → 84.6 (C-11); 3.87 (d, 1H, H-4) → 73.6 (C-4); 3.82 (apparent t, 1H, H-2) → 79.3 (C-2); 3.59-3.55 (m, 1H, H-5) → 73.7 (C-5); 3.53 (dd, 1H, H-3) → 82.2 (C-3); 3.56 (s, 2H, 2×H-6) → 68.8 (C-6). **Salient IR**: Similarly to **60c-2**, IR spectra acquired just after synthesis showed the desired vinylidene peak at 897[w] (=CH_2) and the absence of the strong furan peaks around 1600, finding only 1662[w] (C=C-C=CH_2). (2,3,4,6-tetrabenzyl-galactosyl) bands (see **56c**). Once sent for higher-resolution analysis, hitherto unobserved ketone-like peaks at 1791[w], 1760[w], 1722[w] and 1702[w] were observed (possibly hydrolysis/ring-opening byproducts).

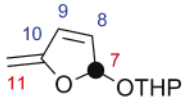
2-(2,3,4,6-tetrabenzyl-β-galactosyl)-5-methylene-2,5-dihydrofuran (60c-2)

During the procedure for **57c-β**, chromatography gave the slower-running **60c-2** (75 mg, 0.12 mmol) as a powder (8%). This compound is thought to be a diastereoisomer of **60c-1** with a different C-7 configuration, based on the near-identical NMR spectra, although no crystals could be isolated for definitive structural determination. Like **60c-1**, it was observed to be unstable and HRMS could not be obtained before the compound degraded; likewise, IR spectra acquired rapidly in-house revealed features which were no longer evident when the compound had been sent elsewhere for analysis.

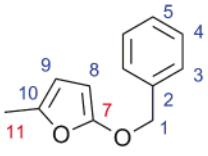
TLC: R_f = 0.30 (5:1 Cy:EA); green (anis), green (van). **^{1}H-NMR ($CDCl_3$, 500 MHz)**: δ = 7.44-7.40 (m, 2H), 7.37-7.25 (m, 18H) (20×phenyl CH); 6.34 (s, 1H, H-7); 6.31 (d, 1H, $J_{8\text{-}9}$ = 5.8 Hz, H-8); 6.16 (d, 1H, $J_{9\text{-}8}$ = 5.8 Hz, H-9); 4.88 (dd, 2H, J_{eff} = 89 Hz, ^2J = 11 Hz, benzyl CH_2); 4.79 (dd, 2H, J_{eff} = 171 Hz, ^2J = 12 Hz, benzyl CH_2); 4.73 (dd, 2H, J_{eff} = 23 Hz, ^2J = 11.8 Hz, benzyl CH_2); 4.54 (d, 1H, $J_{1\text{-}2,\ anti}$ = 7.7 Hz, H-1); **4.46 (s, 1H, H-11$_b$)**; 4.46 (dd, 2H, J_{eff} = 28 Hz, ^2J = 11.8 Hz, benzyl CH_2); **4.14 (s, 1H, H-11$_a$)**; 3.93 (dd, 1H, $J_{2\text{-}1,\ anti}$ = 7.8 Hz, $J_{2\text{-}3,\ anti}$ = 9.4 Hz, H-2); 3.90 (d, 1H, $J_{4\text{-}3,\ gauche}$ = 2.7 Hz, H-4); 3.63 (d, 2H, $J_{6\text{-}5}$ = 6.3 Hz, 2×H-6); 3.54-3.52 (m, 2H, H-3, H-5). **^{13}C-NMR ($CDCl_3$, 125 MHz)**: δ = 162.3 (C-10); 138.9, 138.7, 138.6, 138.0 (4×phenyl C); 131.8 (C-9); 128.9 (C-8); 128.5, 2×128.4, 2×128.3, 128.0, 127.9, 127.7, 127.6 (20×phenyl CH); 107.8 (C-7); 100.1 (C-1); 84.2 (C-11); 82.5 (C-3); 79.2 (C-2); 75.3 (benzyl CH_2); 74.6 (benzyl CH_2); 73.8 (C-5); 73.6 (benzyl CH_2); 73.4 (C-4); 73.1 (benzyl CH_2); 68.9 (C-6). **^{1}H^{1}H-COSY ($CDCl_3$, 500 MHz)**: 7.44-7.25 (m, 20H, 20×phenyl CH) ↔ 7.44-7.25 (m, 20H, 20×phenyl CH); 6.34 (s, 1H, H-7) ↔ 4.46 (s, 1H, H-11$_b$), weak; 6.34 (s, 1H, H-7) ↔ 4.14 (s, 1H, H-11$_a$), weak; 6.31 (d, 1H, H-8) ↔ 6.16 (d, 1H, H-9); 6.16 (d, 1H, H-9) ↔ 4.46 (s, 1H, H-11$_b$), weak; 4.54 (d, 1H, H-1) ↔ 3.93 (dd, 1H, H-2); 4.46 (s, 1H, H-11$_b$) ↔ 4.14 (s, 1H, H-11$_a$), effective $^2J_{11\text{-}11}$ = 166 Hz; 3.93 (dd, 1H, H-2) ↔ 3.54-3.52 (m, 2H, H-3 & H-5); 3.90 (d, 1H, H-4) ↔ 3.54-3.52 (m, 2H, H-3 & H-5). **^{1}H^{13}C-HSQC ($CDCl_3$, 500 MHz)**: 7.44-7.25 (m, 20H, 20×phenyl CH) → 128.5-127.6 (20×phenyl CH); 6.34 (s,1H, H-7) → 107.8 (C-7); 6.31 (d, 1H, H-8) → 128.9 (C-8); 6.16 (d, 1H, H-9) → 131.8 (C-9); 4.88 (dd, 2H, benzyl CH_2) → 75.3 (benzyl CH_2); 4.79 (dd, 2H, benzyl CH_2) → 74.6 (benzyl CH_2); 4.73 (dd, 2H, benzyl CH_2) → 73.1 (benzyl CH_2); 4.54 (d, 1H, H-1) → 100.1 (C-1); 4.46 (s, 1H, H-11$_b$) & 4.14 (s, 1H, H-11$_a$) → 84.2 (C-11); 4.46 (dd, 2H, benzyl CH_2) → 73.6 (benzyl CH_2); 3.93 (dd, 1H, H-2) → 79.2 (C-2); 3.90 (d, 1H, H-4) → 73.4 (C-4); 3.63 (d, 2H, 2×H-6) → 68.9 (C-6); 3.54-3.52 (m, 2H, H-3, H-5) → 82.5 (C-3) & 73.8 (C-5). A long ^{13}C acquisition of the region 160-240ppm was performed to check whether this compound might be an allene with a slowly-relaxing quaternary carbon (estimated shift 200-230 ppm); however the only peak found in this region was the C-10 peak at 162.3 ppm, in keeping with the methylene-dihydrofuran structural interpretation. **IR:** Similarly to **60c-1**, in-house analysis showed 910 [s] (=CH_2) and (2,3,4,6-tetrabenzyl-galactosyl) bands (see **56c**); however after sending for analysis, it was not longer visible and hitherto unobserved peaks at 1792[w], 1760[m], 1725[m] and 1702[m] were found (degradation).



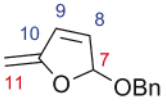
57a



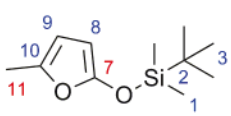
60a



57b



60b



57d

The following species were all synthesised similarly to **57c-β**, from the corresponding alkynyl ketones (200 mg scale). The products, ***especially the furans***, are very volatile (in contrast to the parent alkynyl

ketones) and probably azeotrope with cyclohexane; so pentane/diethyl ether can be used for better recovery from the column. In general they have similar properties of hydrolytic / chromatographic instability as the glycosyl analogues, and the furans are generally spontaneously oxidised with TLC revealing dips eg. PMA while the enol ether acetals usually require warming. The numbering system from **57c-β** has been retained to enable easier spectral comparison.

2-(tetrahydropyran-2-yloxy)-5-methylfuran (57a) and 2-(tetrahydropyran-2-yloxy)-5-methylene-2,5-dihydrofuran (60a)

From **56a**, **57a** (~45% in crude; R_f = 0.50 on 5:1 Cy:EA) and **60a** (~27% in crude; R_f = 0.41 on 5:1 Cy:EA) are returned, and are especially volatile, so yields after column are low and mostly depend on concentration conditions at the end (commonly ~10-20%).

57a: **^{1}H-NMR ($CDCl_3$, 200 MHz)**: δ = 5.80 (~dd, 3.1&1.1 Hz, 1H, H9), 5.32 (~t, 2.4 Hz, 1H, H1), 5.27 (d, 3.1 Hz, 1H, H8), 3.98 (td, 11.0&3.0 Hz, 1H, H5), 3.76-3.55 (m, 1H, H5), 2.19 (s, 3H, 3×H11), 2.00-1.40 (m, 6H, 2×H2+2×H3+2×H4) ppm. **^{13}C-NMR ($CDCl_3$, 50 MHz)**: δ = 157.4 (C7), 142.6 (C10), 106.3 (C9), 99.5 (C1), 85.4 (C8), 61.9 (C5), 29.8 (C2), 25.2 (C4), 18.1 (C3), 13.6 (C10), ppm.

60a: **^{1}H-NMR ($CDCl_3$, 200 MHz)**: δ = 6.35-6.30 (m, 2H, H7+H8), 6.14 (d, 6 Hz, 1H, H9), 5.08-5.03+4.89-4.85 (m, 1H total, H1, diastereos), 4.43-4.40 (m, 1H, $H11_a$), 4.15-4.07 (m, 1H, $H11_b$), 4.05-3.90 (m, 1H, H5), 3.67-3.51 (m, 1H, H5), 2.00-1.40 (m, 6H, 2×H2+2×H3+2×H4) ppm. **^{13}C-NMR ($CDCl_3$, 50 MHz)**: δ = 162.4 (C10), 132.2 (C8), 128.7 (C9), 104.7 (C7), 97.7 (C1), 84.0 (C11), 63.0 (C5), 30.5 (C2), 25.4 (C4), 19.5 (C3) ppm.

2-(benzyloxy)-5-methylfuran (57b) and 2-(benzyloxy)-5-methylene-2,5-dihydrofuran (60b)

From **56b**, **57b** (<~20% indicated in crude, not isolated; R_f = 0.55 on 5:1 Cy:EA) and **60b** (~60% in crude; 10% isolated; R_f = 0.49 on 5:1 Cy:EA) are obtained. The furan is exceptionally volatile.

60b: **^{1}H-NMR ($CDCl_3$, 200 MHz)**: δ = 7.40-7.29 (m, 5H), 6.40-6.37 (m, 2H, H7+H8), 6.16 (dd, 2.4&0.9 Hz, 1H, H9), 4.71 (dd, 34.3&11.7 Hz, 2H, 2×H1), 4.51 – 4.45 (m, 1H, H11), 4.16 (~s, 1H, H11), ppm. **^{13}C-NMR ($CDCl_3$, 50 MHz)**: δ = 162.4 (C10), 137.6 (C2), 131.5 (C9), 129.5 (C8), 128.5 (2×C4), 128.1 (C5), 127.9 (2×C3), 107.8 (C7), 84.0 (C11), 69.4 (C1) ppm.

57b: Spectra are as attributed in the crude mix: **^{1}H-NMR ($CDCl_3$, 200 MHz)**: δ = 7.38-7.30 (m, 5H), 5.31 (d, 1.6 Hz, 1H, H9), 5.24 (d, 1.6 Hz, 1H, H8), 4.65 (~d, 11.7 Hz, 2H, 2×H1), 2.18 (s, 3H, 3×H11) ppm. **^{13}C-NMR ($CDCl_3$, 50 MHz)**: δ = 160.0 (C7), 150.4 (C10), 134.8 (C2), 130.2 (2×C4), 128.5 (C5), 128.2 (2×C3), 120.0, 112.5 (C9), 88.7(C8), 72.0 (C1), 15.2 (C10) ppm.

2-(*tert*-burtyldimethylsilyloxy)-5-methylfuran (57d)

From **56d**, **57d** (50% in crude) and **60d** (25% in crude) were indicated and identified but not isolated; the provisional NMR attributions are given to show how the crude percentages were calculated.

57d: ^{1}H-NMR ($CDCl_3$, 200 MHz): δ = 5.74-5.68 (m, 1H, H9), 4.93-4.87 (m, 1H, H8), 2.10 (s, 3H, 3×H11), 0.93 (s, 9H, 9×H3), 0.15 (s, 6H, 6×H1) ppm. **60d: ^{1}H-NMR ($CDCl_3$, 200 MHz)**: δ = 6.28-6.14 (m, 2H, H7+H8), 6.07-6.02 (m, 1H, H9) 4.19-4.14 (m, 1H, H11), 4.05-4.00 (m, 1H, H11), 0.86 (s overlapped, 9H, 9×H3), 0.13+0.12 (2×s, overlapped, 6H, 6×H1) ppm.

5.1.3. α-oxidation procedures

(*Z*)-2-(2,3,4,6-tetrabenzyl-β-galactosyloxy) 4-oxopent-2-enoate (61-cβ)

In a representative procedure, based on the protocol of Ando *et al.*[22] for a benzofuran, to **57c-β** (46 mg, 0.07 mmol) were added dry dioxane (0.30 mL) and SeO_2 (30 mg, 0.27 mmol; caution: toxic) and the mixture heated to 80°C for 5 h. A brown-green solution over grey powder is produced. Pentane is added, the mixture is filtered on Celite/cotton pad and rinsed with Et_2O, giving 50 mg of yellow crude oil at high vacuum which was fairly pure by NMR save a small amount of what appeared to be 2,3,4,6-tetrabenzyl-β-galactose (the ester hydrolysis product); I assigned the reaction product as structure **61-cβ**. IBX oxidation as per the procedure of Nicolaou *et al.*[23] gave an identical product once the organic layer was washed several times with $NaHCO_3$. The yields for these reactions are estimated as practically quantitative, if the hydrolysis is avoided.

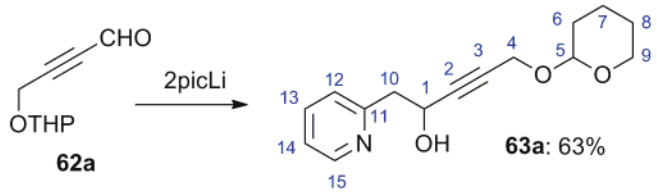
NMR spectra are reproduced on page 277 ff. **^{1}H-NMR ($CDCl_3$, 500 MHz)**: δ = 7.50-7.20 (m 20H), **6.49 (d, 12.2 Hz, 1H, H9)**, **5.95 (d, 12.1 Hz, 1H, H8)**, **5.63 (d, 8.0 Hz, 1H, H1 (anomeric ester))**, 5.00-4.30 (m, 8H), 4.15-3.90 (m, 2H), 3.66 – 3.41 (m, 4H), 2.34 (s, 3H). **^{13}C-NMR ($CDCl_3$, 125 MHz)**: δ = **201.8 (C10)**, 163.4 (C7), **144.6 (C9)**, 138.5+138.2+138.2+137.7 (4×C_{Bn}), 128.5+128.3+128.2+128.1+128.0+127.9+127.8+127.7+127.6+127.6 (20×CH_{Bn}), **122.9 (C8)**, 94.9 (C1), 82.4 (C3), 77.9(C2), 75.4+74.7 (2×CH_2 benzyl), 74.3 (C5), 73.6 (C4), 73.1+72.9 (2×CH_2 benzyl), 68.0 (C6), 29.9 (C11). Assigning the double bond as Z was performed with reference to literature examples, the most relevant of which is a thorough study by Cermolo *et al.*[24] who report a photochemical synthesis of a similar *Z*-oxopentenoate; "on contact with chromatographic adsorbents [SiO_2], [*cis*] compound […] isomerized into [trans], and the extent of the isomerization strictly depended on the time employed for the analysis." They give NMR data for the corresponding ω-*aldehyde*, not ω-methyl ketone, but the coupling constants of the protons should be similar: they give the *cis* form protons at 6.59 (d, ***11.4 Hz***) and 6.36 (dd, 11.4&7.4 Hz), and the *trans* form with protons at 6.68 (d, ***16.5 Hz***) and 6.97 (dd, 16.5&7.2 Hz), matching my *Z* interpretation. As I did not perform silica gel purification I could hope that my *Z*-assignment reflected the initial reaction product, which is mechanistically reasonable too.

5.2. Furyl cycloisomerisation – pyridyl preinstallation

5.2.1. Oxidative routes

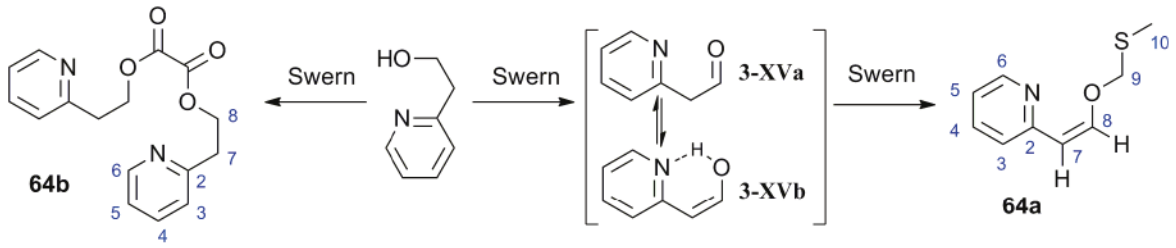
4-(2,3,4,6-tetra(benzyl)*β*-galactosyl)but-2-ynal (62b)

Synthesised and characterised by intern Claire Weisslinger from **55c**, following the formylation protocol from Journet *et al.*[17], but using *t*BuLi for deprotonation. **62b** is returned as a yellow oil in 61% yield (R_f = 0.28 on Cy:EA 4:1; Van). HRMS (TOF MS ES+) calcd for $C_{38}H_{38}O_7Na$ [MNa+]: m/z 629.2515, found 629.2542. NMR acquired by Claire Weisslinger. **^{1}H-NMR ($CDCl_3$, 200 MHz)**: δ = 9.19 (s, 1H), 7.56-7.25 (m, 20H), 5.09-4.45 (m, 11H), 4.01-3.82 (m, 2H), 3.74-3.46 (m, 4H) ppm. **^{13}C-NMR ($CDCl_3$, 50 MHz)**: δ = 176.2, 138.6, 138.5, 138.4, 137.8; 128.6, 128.4, 128.3, 128.2, 128.2, 128.1, 127.9, 127.7, 127.6 (total 20×C_{Ar}), 102.2, 91.6, 85.5, 82.1, 79.2, 75.3, 74.6, 73.7, 73.6, 73.3, 73.2, 68.7, 55.6 ppm.



1-(pyridin-2-yl)-5-(tetrahydro-2*H*-pyran-2-yloxy)pent-3-yn-2-ol (63a)

63% (1.81 g, 6.9 mmol) from a similar protocol to those for **78a** or as used by Wang and Ding[25], using 2-picoline (1020 mg, 11.0 mmol) and **62a** (2.19 g, 13 mmol, 1.2 eq) quenching with sat. NH_4Cl. Separation used 4:1→1:1 Cy:EA. Successively (2:1 Cy:EA), **62a** (R_f=0.62), then a THP-bearing alkene (allene?) of undetermined structure (R_f=0.34) then unreacted 2-picoline (R_f=0.21) are eluted, then **63a** (R_f=0.07) as a moderately stable yellow oil which browns on standing at 4°C and which appears to degrade on SiO_2. It stains a very distinctive blue with Dragendorff reagent, red with Van, and grey with Anis. NMR spectra are reproduced on page 279 ff. **^{1}H-NMR ($CDCl_3$, 200 MHz)**: δ = 8.40 (d, 4.6 Hz, 1H, H15), 7.56 (td, 7.7&1.5 Hz, 1H, H12), 7.21-7.00 (m, 2H, H13+H14), 5.59 (s br, 1H, OH), 4.81 (~t, 5.8 Hz, 1H, H5), 4.61 (~d, 2.6 Hz, 1H, H1), 4.15 (d, 1.2 Hz, 2H, 2×H4), 3.83-3.56 (m, 1H, H9), 3.42-3.35 (m, 1H, H9), 3.13-3.06 (m, 2H, H10), 1.83-1.36 (m, 6H, 2×H6+2×H7+2×H8) ppm. **^{13}C-NMR ($CDCl_3$, 50 MHz)**: δ = 158.2 (C11), 148.4 (C15), 136.6 (C12), 124.0 (C13), 121.7 (C14), 96.2 (C5), 86.5 (C2), 80.3 (C3), 61.7 (C9), 61.4 (C1), 53.9 (C4), 44.1 (C10), 30.0 (C6), 25.1 (C8), 18.8 (C7) ppm. The jmod spectrum was used to complete spectral assignment. HRMS (TOF MS ES^+) calcd for $C_{15}H_{20}O_3N$ [MH^+]: m/z 262.1443, found 262.1439.

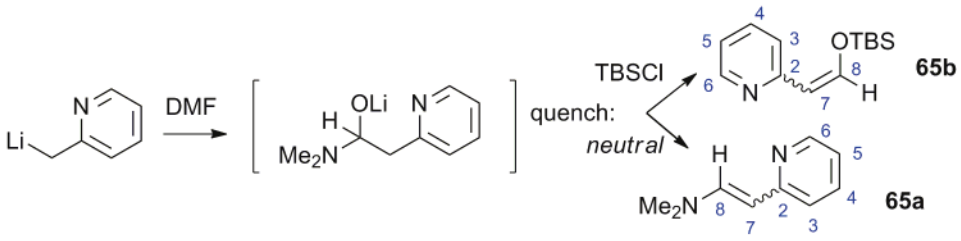


(Z)-2-(2-((methylthio)methoxy)vinyl)pyridine (64a) and bis(2-(pyridin-2-yl)ethyl) oxalate (64b)

Following the Swern protocol of Lam *et al.*[26], 2-(2-pyridyl)ethanol (2.038 g, 16.6 mmol) was engaged, and 1.8 g of red crude oil isolated. TLC (Dragendorff; PMA especially good for **64a**) is used to follow column separation. Column performed with 1:1 Cy:EA, delivering **64a** as a yellow oil (70 mg, 0.39 mmol, 2%; R_f =0.52 on 9:1 AcOEt:MeOH) and then **64b** as a yellow oil (130 mg, 0.43 mmol, 2.6%; R_f =0.07 on 9:1 AcOEt:MeOH).

64a: DIMS(+): 182 (MH^+). **^{1}H-NMR (CDCl$_3$, 200 MHz)**: δ = 8.44 (d, 4.8 Hz, 1H, H6), 7.55 (d, 12.6 Hz, 1H, H8), 7.55-7.48 (td overlapped, 1H, H3), 7.08 – 6.99 (m, 2H, H5+H4), 6.03 (d, 12.4 Hz, 1H, H7), 4.99 (s, 2H, 2×H9), 2.22 (s, 3H, 3×H10) ppm; I have assigned the Z-configuration from the coupling constant of the alkenic protons (similar to **61-cβ**).

64b: Unit Mass DIMS(+; only peaks below 300 Da collected): 152 (MH_2^{2+}+detector error?). **^{1}H-NMR (CDCl$_3$, 200 MHz)**: δ = 8.51 (d, 4.7 Hz, 2H, 2×H6), 7.64 – 7.54 (m, 2H, 2×H3), 7.23-7.08 (m, 4H, 2×H4+2×H5), 4.59 – 4.49 (m~td, 4H, 4×H8), 3.11 (t, 6.8 Hz, 4H, 4×H7).

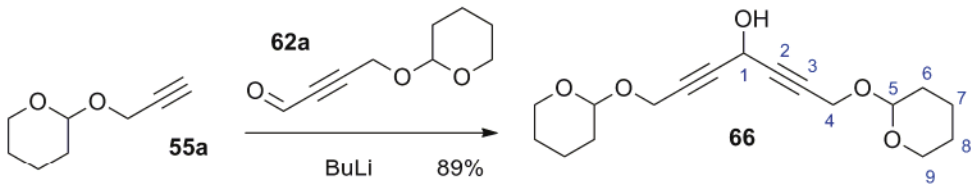


N,N-dimethyl-2-(pyridin-2-yl)ethenamine (65a)*

73% (601 mg, 4.1 mmol) from a similar protocol to that for **78a** or as used by Wang and Ding[25], using 2-picoline (520 mg, 5.59 mmol) and DMF (2.17 g, 30 mmol, 5 eq) but quenching into pure water, not buffer. If quenched into KH_2PO_4 (20 mL, 10%)+conc HCl (1 mL)+Et_2O (20 mL), this product is the only isolable product as the aldehyde appears to generate complex mixtures. HRMS (TOF MS ES$^+$) calcd for $C_9H_{14}N_2$ [MH^+]: m/z 149.1079, found 149.1080. NMR matches reference[27]. E/Z ratio was not assigned.

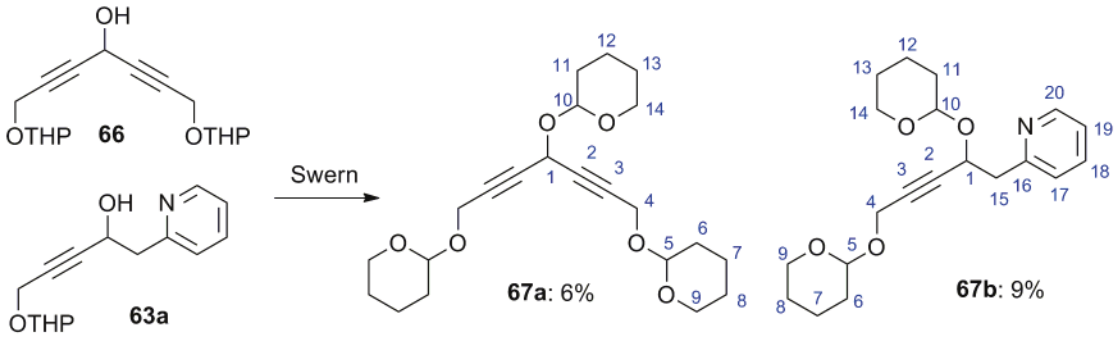
E/Z-2-(2-((tert-butyldimethylsilyl)oxy)vinyl)pyridine (65b)

Reaction run with 2-picoline (420 mg, 4.52 mmol) and DMF (927 mg, 12.7 mmol, 3 eq) similarly to the reaction for **78a** but at -90°C, and instead of quenching into aqueous media, a solution of TBSCl (1.95 g, 13 mmol) in THF (7 mL) was injected at -80°C. The colour of the reaction vanishes, and fuming occurs, during this quench. Extractive workup and high vacuum at room temperature gave a pale yellow crude oil (1.07 g) which was essentially pure **65b** (4.5 mmol, 99%) with almost no contaminants and one of the most beautifully-resolved NMR spectra of this PhD. The compound does degrade upon storage in the fridge at 4°C however and is destroyed after about a month. E/Z ratio assigned as 40/60. The compound is totally destroyed on silica gel chromatography. The NMR spectrum is reproduced on page 281 ff. **Z: ^{1}H-NMR (CDCl$_3$, 500 MHz)**: δ = 8.45 (d, 4.5 Hz, 1H, H6), 8.02 (dt, 8.2&0.9 Hz, 1H, H3), 7.55 (dd, 7.8&1.8 Hz, 1H, H4), 6.97-6.95 (m, 1H, H5), 6.58 (d, 6.7 Hz, 1H, H7, *cis*), 5.60 (d, 6.7 Hz, 1H, H8), 0.94 (s, 9H), 0.20 (s, 6H) ppm. **E: ^{1}H-NMR (CDCl$_3$, 500 MHz)**: δ = 8.39 (d, 4.8 Hz, 1H, H6), 7.56 (d, 11.8 Hz, 1H, H7, *trans*), 7.45 (td, 7.7&1.7 Hz, 1H, H5), 6.98-6.92 (m, 2H, H4+H3), 6.03 (d, 11.8 Hz, 1H, H8), 0.92 (s, 9H), 0.20 (s, 6H) ppm.



1,7-bis(tetrahydro-2*H*-pyran-2-yloxy)hepta-2,5-diyn-4-ol (66)

The procedure was similarly performed to that for **56c**, using propargyl ether **55a** (420 mg, 3 mmol) and aldehyde **62a** (420 mg, 2.5 mmol) with *n*BuLi (1.6 M, 2 mL, 3.2 mmol). Separation with 7:1→3:2 peth:EA gives **66** as a yellow viscous oil (687 mg, 2.23 mmol, 89%). HRMS (TOF MS ES$^+$) calcd for $C_{17}H_{24}O_5Na$ [MNa$^+$]: m/z 331.1521, found 331.1532. **^{1}H-NMR (CDCl$_3$, 200 MHz)**: δ = 5.08 (s, 1H, H1), 4.70 (s, 2H, 2×H5), 4.33 (s, 1H, OH), 4.24-4.12 (m ~q, 4H, 4×H4), 3.76-3.65 (m, 2H, 2×H9), 3.47-3.39 (m, 2H, 2×H9), 1.75-1.40 (m, 12H, 4×H6+4×H7+4×H8) ppm. **^{13}C-NMR (CDCl$_3$, 50 MHz)**: δ = 96.6 (2×C5), 83.1 (2×C2), 79.9 (2×C3), 61.7 (2×C9), 54.0 (2×C4), 51.6 (C1), 29.9 (2×C6), 25.1 (2×C8), 18.7 (2×C7) ppm.



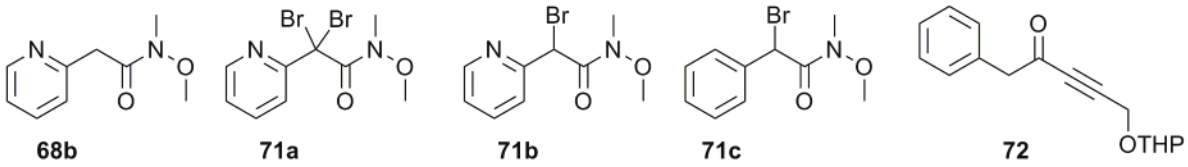
1,4,7-tris(tetrahydro-2*H*-pyran-2-yloxy)hepta-2,5-diyn-4-ol (67a)

Similarly to the Swern procedure giving **64a/b**, **67a** (27 mg, 0.07 mmol, 6%) was the only isolable product obtained from **66** (370 mg, 1.20 mmol). R_f=0.63 on 5:1 Cy:EA, stains brown with Anis. **^{1}H-NMR (CDCl$_3$, 200 MHz)**: δ = 5.30 (~s, 1H, H1), 4.91 (~s, 1H, H10), 4.78 (~s, 2H, 2×H5), 4.28 (~s, 4H, 4×H4), 3.90-3.70 (m, 3H, 2×H9+H14), 3.60-3.40 (m, 3H, 2×H9+H14), 1.85-1.60 (m, 18H, 4×H6+4×H7+4×H8+2×H11+2×H12+2×H13) ppm. **^{13}C-NMR (CDCl$_3$, 50 MHz)**: δ = 97.0 (C5), 96.9 (C5), 95.4 (C10), 81.5+81.4 (2×C2), 80.8(2×C3), 62.1+62.0 (2×C9+C10), 54.8+54.4+54.2 (2×C4+C1), 30.3+30.3+30.1 (2×C6+C11), 25.4 (2×C8+C13), 19.1 (2×C7), 18.9 (C13) ppm.

1-(pyridin-2-yl)-2,5-bis(tetrahydro-2*H*-pyran-2-yloxy)pent-3-yn-2-ol (67b)

Similarly to the Swern procedure giving **64a/b**, **67b** (116 mg, 0.35 mmol, 9%) was obtained as a pale yellow oil from Swern oxidation of **63a** (990 mg, 3.81 mmol). R_f=0.46 on 1:1 Cy:EA, stains purple with Van, strongly with Dragendorff. **^{1}H-NMR (CDCl$_3$, 200 MHz)**: δ = 8.50-8.43 (m, 1H, H20), 7.54 (td, 7.7&1.7 Hz, 1H, H19), 7.22 (d, 7.8 Hz, 1H, H7), 7.06 (dd, 6.9&5.5 Hz, 1H, H18), 4.93 (~s, 1H, H5), 4.80 (t, 7.0 Hz, 1H, H1), 4.65 (~s, 1H, H10), 4.18 (s, 2H, 2×H4), 3.81-3.61 (m, 1H, H9), 3.50-3.35 (m, 1H, H9), 3.29-2.99 (m, 4H, 2×H14+2×H15), 1.80-1.35 (m, 12H, 2×H6+2×H7+2×H8+2×H13+2×H12+2×H11) ppm. **^{13}C-NMR (CDCl$_3$, 50 MHz)**: δ = 157.4 (C16), 148.8 (C20), 136.0 (C17), 124.5 (C18), 121.5 (C19), 96.4+96.3 (C5, diast), 94.5 (C10), 84.1 (C2), 81.7 (C3), 64.1 (C1), 61.8 (C9), 60.9 (C14), 53.9 (C4), 44.2 (C15), 30.1 (C6), 30.0 (C11), 25.2 (C8+C13), 18.9 (C7), 18.4 (C12) ppm.

5.2.2. Oxidative pathway 1, via pyridylacetic acid derivatives



_N_-methoxy-_N_-methyl-2-(pyridin-2-yl)acetamide (68b)

Prepared by the protocol of Baxendale *et al*. for nicotinic acid[28], from 2-(2-pyridyl)acetic acid (4.302 g, 24.8 mmol), giving after chromatography (1:1:0→0:1:0→0:10:1 Cy:EA:MeOH) the

apparently unattested Weinreb **68b** as a yellow oil (725 mg, 4.01 mmol, 66%; R_f=0.41 on 9:1 EA:MeOH, stains without heating with $KMnO_4$ and also strongly with Dragendorff). **^{1}H-NMR ($CDCl_3$, 200 MHz)**: δ = 8.40 (d, 4.3 Hz, 1H), 7.50 (td, 7.7&1.8 Hz, 1H), 7.18 (d, 7.7 Hz, 1H), 7.02 (ddd, 7.7&2.6&0.7 Hz, 1H), 3.86 (s, 2H), 3.54 (s, 3H), 3.08 (s, 3H) ppm. **^{13}C-NMR ($CDCl_3$, 50 MHz)**: δ = 171.1, 155.3, 149.1, 136.2, 123.8, 121.6, 61.2, 41.7, 32.0 ppm. HRMS (TOF MS ES$^+$) calcd for $C_9H_{13}N_2O_2$ [MH$^+$]: m/z 181.0986, found 181.0977.

2,2-dibromo-*N*-methoxy-*N*-methyl-2-(pyridin-2-yl)acetamide (71a) and 2-bromo-*N*-methoxy-*N*-methyl-2-(pyridin-2-yl)acetamide (71b)

In a representative procedure, to **68b** (179 mg, 1.00 mmol) were added NBS (215 mg, 1.2 mmol) and AIBN (30 mg) under Ar. CCl_4 (15 mL) was added and the mixture heated to 90°C for 2 h. The reaction was concentrated, and the portion soluble in 2:1 Cy:EA was loaded onto column (remaining solid is mostly succinimide); chromatography with 2:1→1:1 Cy:EA separates **71a** (44 mg, 0.13 mmol, 13%; R_f=0.23 on 2:1 Cy:EA) and **71b** (196 mg, 0.76 mmol, 76%; R_f=0.12 on 2:1 Cy:EA). The yields can be adjusted according to the desired product mixture by changing the stoichiometry of NBS. Neither product is easily oxidised on TLC but both are clearly visible under UV.

71a: ^{1}H-NMR ($CDCl_3$, 200 MHz): δ = 8.49 (d, 4.8 Hz, 1H), 7.97 (td, 8.1&0.9 Hz, 1H), 7.81 (td, 7.8&1.8 Hz, 1H), 7.21 (ddd, 7.5&4.8&1.1 Hz, 1H), 3.27 (s, 3H), 3.13 (s, 3H) ppm. **^{13}C-NMR ($CDCl_3$, 50 MHz)**: δ = 165.3, 159.3, 148.5, 137.8, 123.4, 120.0, 63.4, **59.8**, 35.0 ppm. HRMS (TOF MS ES$^+$) calcd for $C_9H_{11}N_2O_2Br_2$ [MH$^+$]: m/z 336.9187, found 336.9202.

71b: ^{1}H-NMR ($CDCl_3$, 200 MHz): δ = 8.53 (d, 4.9 Hz, 1H), 7.85 (d, 8 Hz, 1H), 7.75 (td, 7.4&1.8 Hz, 1H), 7.25 (ddd, 7.3&4.9&1.2 Hz, 1H), 6.12 (s, 1H), 3.68 (s, 3H), 3.25 (s, 3H) ppm. **^{13}C-NMR ($CDCl_3$, 50 MHz)**: δ = 167.6, 155.9, 148.6, 137.2, 124.3, 123.4, 61.6, **45.3**, 32.8 ppm. HRMS (TOF MS ES$^+$) calcd for $C_9H_{12}N_2O_2Br$ [MH$^+$]: m/z 259.0082, found 259.0092.

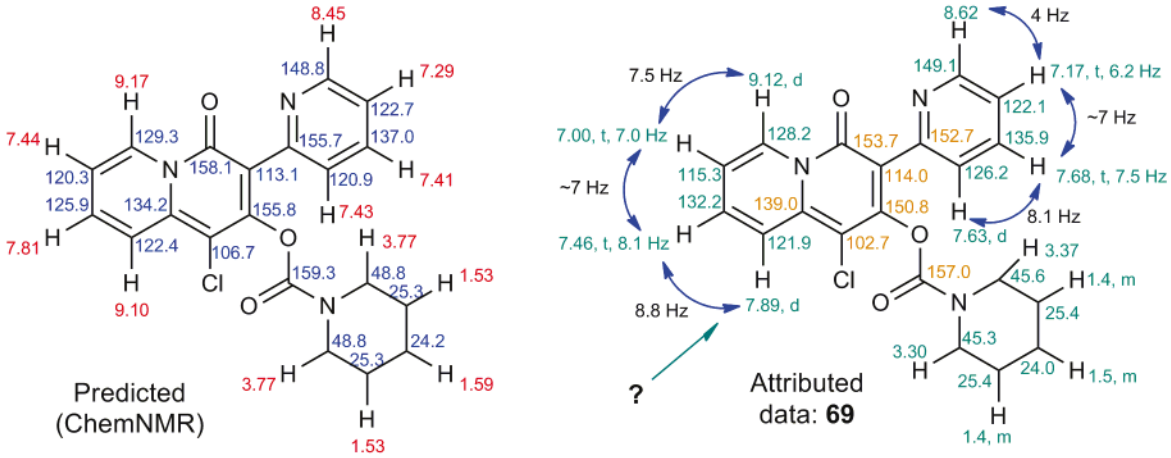
2-bromo-*N*-methoxy-*N*-methyl-2-phenylacetamide (71c)

Similarly to the procedure for **71a**/**71b**, **71c** (30 mg, 0.12 mmol, 17%; R_f=0.31 on 5:1 Cy:EA) was obtained as a minor product during radical dibromination (R_f=0.60 on 5:1 Cy:EA) of known **70b**, which had been synthesised by a literature procedure[29]. **^{1}H-NMR ($CDCl_3$, 200 MHz)**: δ = 7.60-7.53 (m, 2H), 7.41-7.30 (m, 3H), 5.99 (s, 1H), 3.60 (s, 3H), 3.23 (s, 3H) ppm. **^{13}C-NMR ($CDCl_3$, 50 MHz)**: δ = 168.4 (weak), 136.8, 129.1, 129.0 (×2), 128.9 (×2), 61.7, 45.4, 33.0 ppm. HRMS (TOF MS ES$^+$) calcd for $C_{10}H_{12}NO_2BrNa$ [MNa$^+$]: m/z 279.9949, found 279.9973.

1-phenyl-5-((tetrahydro-2H-pyran-2-yl)oxy)pent-3-yn-2-one (72)

Similarly to the procedure for **56c**, **70b** (96 mg, 0.54 mmol) was reacted with the lithium acetylide of **55a** (210 mg, 1.51 mmol) and quenched into KH_2PO_4/iPr$_2$O to give after workup, pad filtration (2:1

pentane:EA) and chromatography on 30:1→10:1 pentane:EA gradient, **72** as a colourless oil (86 mg, 0.33 mmol, 61%; R_f=0.13 on 20:1 Cy:EA). Note that quenching into NH_4Cl gives the rearrangement product as outlined by Choudhury *et al.*[30]; patents commonly cite citric acid as a quench buffer, but I find KH_2PO_4[17,31] an excellent general solution. Compound degrades on silica but a fast column can be run. **^{1}H-NMR ($CDCl_3$, 200 MHz)**: δ = 7.39-7.22 (m, 5H), 4.67-4.62 (m, 1H), 4.36 (s, 2H), 3.84 (s, 2H), 3.84-3.71 (m, 1H), 3.54-3.44 (m, 1H), 1.91-1.46 (m, 6H) ppm. **^{13}C-NMR ($CDCl_3$, 50 MHz)**: δ = 184.5, 132.8, 129.9(×2), 128.8(×2), 127.5, 96.9, 90.0, 84.7, 62.0, 53.7, 52.0, 30.1, 25.3, 18.9 ppm. HRMS (TOF MS ES$^+$) calcd for $C_{16}H_{18}O_3Na$ [MNa^+]: m/z 281.1154, found 281.1166.



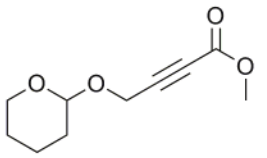
1-chloro-4-oxo-3-(pyridin-2-yl)-4*H*-quinolizin-2-yl piperidine-1-carboxylate (69)

Commercial 2-(2-pyridyl)acetic acid hydrochloride was crushed and dried under vacuum until a free-flowing powder was obtained. To this (737 mg, 4.24 mmol) was added $SOCl_2$ (15 mL) and the mixture heated to 60°C for 30 min (protocol of Momenteau *et al.* for 3-pyridyl analogue)[32]. Gas is evolved and mixing the bright brown-red solution can be problematic. The excess $SOCl_2$ was evaporated to a cold trap and quenched; then the residual red solid dissolved in $CHCl_3$. To this was added slowly piperidine (2.21 g, 30 mmol) in $CHCl_3$ (10 mL), giving much fuming and a deep red solution, which was cooled to 0°C. NaOH (20 mL, 1M, 0°C) was added and the organic layer washed with water and brine, dried on sodium sulfate, filtered and evaporated to give a black powder (690 mg). Two products are seen by TLC.

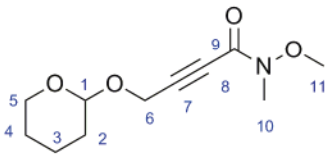
The first is a decarboxylation[33] compound whose structure could not satisfyingly be assigned, but whose structure contains an ordinary 2-pyridyl moiety, an ordinary N-piperidyl moiety, and exactly one equivalent of CO_2 (confirmed by HRMS, [$C_{11}H_{15}N_2O_2Na^+$]), yet presents the quaternary peak of the 'extra' carbon at 195.4 ppm in the ^{13}C spectrum which I could not explain by carbamate (~155), amide hydrate (~120 ppm) or N-hydroxyester (~170) structures. It has R_f = 0.21 on Cy:EA 2:1, fluoresces blue under UV, and is *not* the compound **3-XVIII** (literature reference for its NMR[34]). Yield ~150 mg, ie 0.75 mmol.

The second product fluoresces green under 254 nm UV indicating an extensive and unusual π-system, which I assign as **69** (R_f = 0.6 on MeOH:EA 1:9). HRMS (TOF MS ES$^+$) calcd for $[C_{20}H_{19}N_3O_3Cl]^+$ = [MH$^+$]: m/z 384.1115, found 384.1116. As this reaction product was so unexpected, the NMR data attribution is shown in full (C and H shifts, H multiplicity, and quinolizine COSY couplings, on the right-hand side), with the ChemNMR prediction on the left (red and blue) for comparison. The biggest difference with the prediction is the shift of hydrogen 8 (7.89 ppm, not prediction 9.10). HSQC confirms all the C↔H pairings. **^{1}H-NMR (CDCl$_3$, 500 MHz)**: δ = 9.12 (d, 7.5 Hz, 1H), 8.62 (d, 4 Hz, 1H), 7.89 (d, 8.8 Hz, 1H), 7.68 (t, 7.5 Hz, 1H), 7.63 (d, 8.1 Hz, 1H), 7.46 (t, 8.1 Hz, 1H), 7.17 (t, 6.2 Hz, 1H), 7.00 (t, 7.0 Hz, 1H), 3.40-3.20 (2×~s, 4H), 1.50-1.40 (m, 6H) ppm. **^{13}C-NMR (CDCl$_3$, 125 MHz)**: δ = 157.0, 153.7, 152.7, 150.8, 149.1, 139.0, 135.9, 132.2, 128.2, 126.2, 122.1, 121.9, 115.3, 114.0, 102.7, 45.6 (×2), 25.4 (×2), 24.0 ppm.

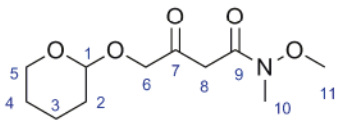
5.2.3. Non-oxidative pathway 2: continuous flow reaction; and cyclisation trials



73



78a



78b

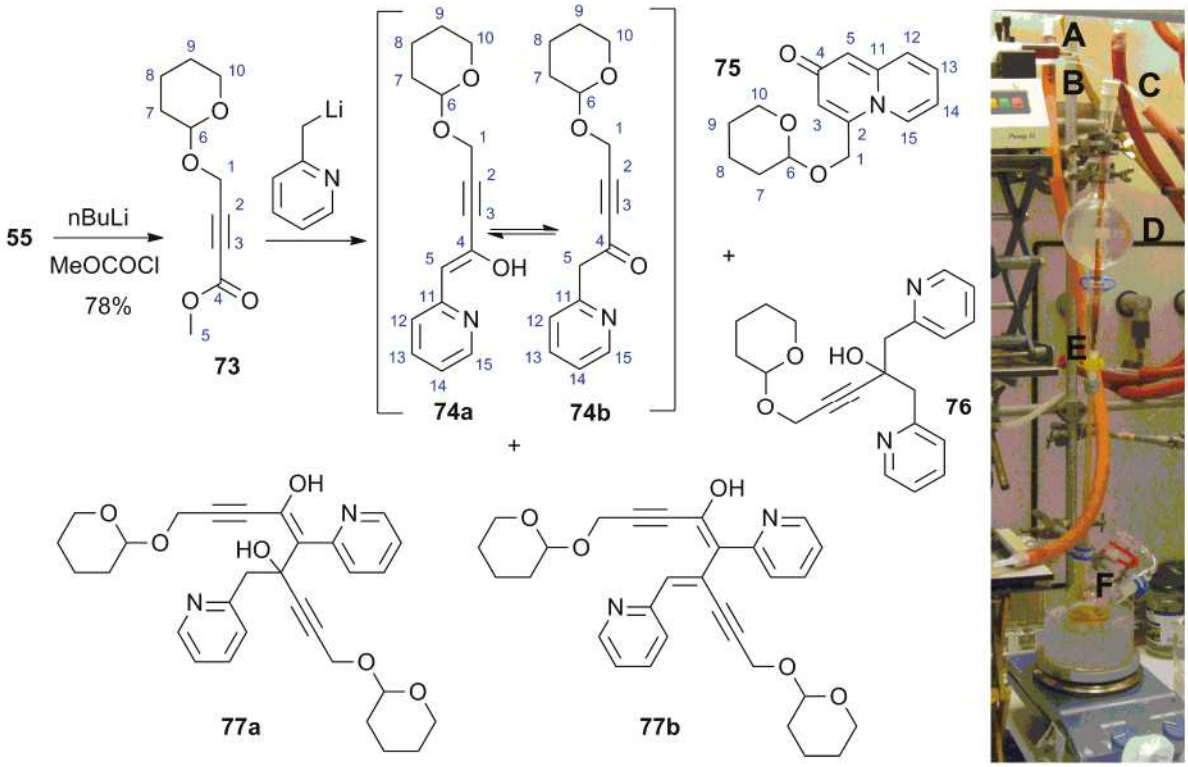
_N_-methoxy-_N_-methyl-4-(tetrahydro-2_H_-pyran-2-yloxy)but-2-ynamide (78a) and _N_-methoxy-_N_-methyl-3-oxo-4-(tetrahydro-2_H_-pyran-2-yloxy)butanamide (78b)

73 was synthesised in 78% yield from reaction of **55a** with *n*BuLi in THF at -65°C, then cannula addition into MeOCOCl (2 eq), and quenching into a rapidly-stirred 1:1 mix of Et_2O and phosphate buffer (0.33 M in HPO_4^{2-} and 0.24 M in $H_2PO_4^-$), separating the crude with 100:0→7:1 Cy:EA (R_f = 0.44 on Cy:EA 5:1), similarly to *eg.* **56c**. **^{13}C-NMR (CDCl$_3$, 50 MHz)** confirms the product; peaks are given for reference: δ = 153.4, 96.9, 83.8, 77.1, 61.7, 53.4, 52.5, 29.9, 25.1, 18.6 ppm. Replacement of MeOCOCl with diethyl carbonate resulted in ~25% yield of the corresponding ethyl ester. Replacement of **55a** by galactosyl **55c** keeping MeOCOCl resulted in no isolated yield although *t*BuLi was also used instead of *n*BuLi.

To **73** (450 mg, 2.12 mmol) and $H_2N(OMe)Me{\cdot}Cl$ (430 mg, 4.37 mmol) at -7°C under Ar were added THF (12 mL) and *i*PrMgCl (2M, 4.8 mL) dropwise over 2 minutes with the cooling bath in place. The mixture was warmed to RT with stirring for ten hours, quenched into sat. NH_4Cl, worked up as for **56c**, and **78a** isolated first (R_f = 0.45 on Cy:EA 1:1, chars pink with Van; 270 mg, 1.19 mmol, 56%), followed by **78b** (R_f = 0.25 on Cy:EA 5:1, chars yellow with Van; 33 mg, 0.13 mmol, 6%).

78a: ^{1}H-NMR (CDCl$_3$, 500 MHz): δ = 4.64-4.60 (m, 1H, H1), 4.24 (s, 2H, H6), 3.67-3.53 (m+s, 4H, 3×H11+H5), 3.37-3.28 (m, 1H, H5), 3.10 (s br, 3H, 3×H10), 1.82-1.32 (m, 6H, 2×H2+2×H3+2×H4) ppm. **^{13}C-NMR (CDCl$_3$, 50 MHz)**: δ = 153.4, 96.6, 87.4, 77.5 (br), 61.8 (br), 61.6, 53.5, 32.0 (br), 29.8, 25.0, 18.6 ppm; the quaternary carbons were very slow to relax in these acquisitions. DIMS(+): 228 (MH$^+$).

78b: ^{1}H-NMR (CDCl$_3$, 500 MHz): Only peaks for the major keto tautomer are given (keto:enol ~80:20): δ = 4.67 (~t, 3.3 Hz, 1H, H1), 4.25 (dd, 30&12.7 Hz, 2H, 2×H6), 3.91-3.72 (m, 1H, H5), (s, 3H), 3.55-3.42 (m, 1H, H5), 3.18 (s, 3H), 1.90-1.50 (m, 6H) ppm. **^{13}C-NMR (CDCl$_3$, 50 MHz)**: δ = 202.5, 167.9, 98.9, 72.1, 62.3, 61.2, 44.7, 32.0, 30.1, 25.1, 19.1ppm. **HRMS (ESI-ToF)**: Calc'd for $[C_{11}H_{19}O_5NNa]^+ = [MNa]^+$: 268.1161; found 268.1168.



1-(pyridin-2-yl)-5-(tetrahydro-2*H*-pyran-2-yloxy)pent-1-en-3-yn-2-ol (74a)

Static procedure: To dry THF (20 mL) under Ar at -80°C were added nBuLi (1.6 M in hexanes, 1 mL), then dropwise over 1 min 2pyMe (194 mg, 2.12 mmol) and the coloured solution stirred for 30 min at this temperature. Then a solution of **78a** (265 mg, 1.17 mmol) in dry THF (2 mL) was added at -110°C in one shot; the colour clears after 2-3 seconds. The temperature was returned to RT and the reaction quenched quickly into KH_2PO_4 (20 mL, 10%)+conc HCl (1 mL)+Et_2O (20 mL) at 0°C, and working rapidly, the aqueous layer was extracted twice with Et_2O once the pH was adjusted to pH~7, then the combined organic layers were washed with water and brine, dried on sodium sulfate, filtered and evaporated to obtain a crude oil containing **74a**:**78a** at mol ratio 1:1 (>95% crude yield brsm). An identical procedure but using ester **73** resulted in a distribution 1:3:0.25 **74a**:**73**:**76**, with varying amounts of the aldol products **77a**/**77b**.

Continuous flow procedure: Solutions of 2-pyridyllithium (syringe at A, 465 mg, 5 mmol) and ester **73** (at B, 1.5 g, 7.58 mmol) were injected continuously via syringe pump (concentrations must take

account of syringe cross-section to achieve correct stoichiometry, here ~1.5:1 RLi:Ester); the system is under positive pressure of Ar (inlet C). At injection the solutions enter a thin glass tube which is jacketed with ice-water (D) and are rather rapidly carried down to point E, where the tube is narrowed so that turbulent mixing occurs (note the considerable lightening of the solution at this point). After this mixing point the solutions drip down another section of tube (where the well-mixed reaction can continue to occur) until falling into strongly buffered quench mixture F, which is also under Ar blanket. After complete addition, solution F is extracted as per the above procedure. Total time elapsed during the reaction was only a few minutes, important for slowing degradative losses of **74a**; reaction time was ~10 s.

The product degrades on silica but a fast column is possible (Cy:EA gradient 20:1→2:1; R_f = 0.65 on 5:95 MeOH:AcOEt). The isolated yield depends strongly on the amount of time the product spends at warm conditions (RT), as **74a** degrades continuously by cyclisation to 75, with a halftime of hours at RT, days in the fridge at 4°C (days), and even degrades at -40°C (2 months). The product has a distinct rodent-like smell as of housemice or rats. **HRMS (ESI-ToF)**: Calc'd for $[C_{15}H_{18}O_3N]^+$ = $[MH]^+$: 260.1287, found 260.1291 – this mass almost certainly is that for its degradation product (shipping to the analysis centre takes some days), but as the degradation product is an isomer, it proves the mass of the starting material. HSQC/COSY was used to confirm **all** peak pairings. NMR spectra are reproduced on page 282 ff.

74a: **^{1}H-NMR ($CDCl_3$, 500 MHz)**: δ = 14.8 (s br, 1H, OH), 8.21 (d, 5.1 Hz, 1H, H15), 7.60 (~t, 8.2 Hz, 1H, H13), 6.98 (~t, 6.3 Hz, 1H, H14), 6.95 (d, 8.0 Hz, 1H, H12), 5.75 (s, 1H, H5), 4.84 (~t, 3.7 Hz, 1H, H6), 4.43 (d, 2.5 Hz, 2H, 2×H1), 3.84 (~td, 10.6&2.4 Hz, 1H, H10), 3.55-3.52 (m, 1H, H10), 1.85-1.50 (m, 6H, 2×H7+2×H8+2×H9) ppm. **^{13}C-NMR ($CDCl_3$, 125 MHz)**: δ = 157.6 (C4), 149.2 (C11), 144.0 (C15), 137.6 (C13), 121.5 (C12), 119.2 (C14), 103.7 (C5), 97.0 (C6), 85.9 (C2), 83.0 (C3), 62.1 (C7), 54.5 (C1), 30.3 (C7), 24.4 (C9), 19.1 (C8) ppm.

74b distinctive peaks (keto form; standard shifts for H/C at positions 7-10): **^{1}H-NMR ($CDCl_3$, 500 MHz)**: δ = 8.56 (1H, H15), 7.66 (1H, H13), 7.24 (1H, H12), 7.20 (1H, H14), 4.35 (2H, 2×H1), 4.07 (2H, 2×H5) ppm. **^{13}C-NMR ($CDCl_3$, 125 MHz)**: δ = **183.4 (C4)**, 166.3 (C11), 149.8 (C15), 136.9 (C13), 124.6 (C12), 122.5 (C14), 97.7 (C6), 90.1 (C3), 84.8 (C2), **54.1 (C5)**, 53.8 (C1) ppm.

4-((tetrahydro-2H-pyran-2-yloxy)methyl)-2H-quinolizin-2-one (75)

This compound is continually formed by cyclisation of **74a**. In the static procedure given for **74a** with reactant **73**, **75** was isolated as a bright yellow oil (R_f = 0.89 on 5:95 MeOH:AcOEt). NMR spectra are reproduced on page 284 ff. **^{1}H-NMR ($CDCl_3$, 500 MHz)**: δ = 8.06 (d, 7.2 Hz, 1H, H15), 7.18 (d, 9.1 Hz, 1H, H12), 7.10 (t, 7.6 Hz, 1H, H13), 6.90 (d, 2.8 Hz, 1H, H3), 6.63 (t, 6.9 Hz, 1H, H14), 6.53 (s, 1H, H5), 4.93+4.67 (each d, 13.4 Hz, 1H; 2×H1), 4.71 (~s, 1H, H6), 3.93-3.80 (m, 1H, H10), 3.58-3.50 (m, 1H, H10), 1.80-1.40 (m, 6H, 2×H7+2×H8+2×H9) ppm. **^{13}C-NMR ($CDCl_3$, 50 MHz)**: δ =

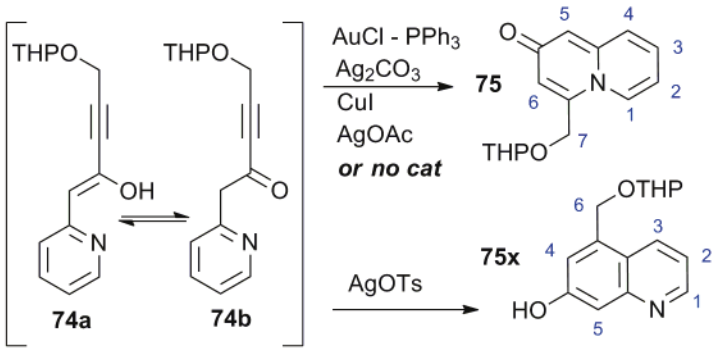
175.8 (C8), 145.2 (C2), 140.4 (C11), 128.7 (C15), 128.5 (C13), 124.8 (C12), 124.5 (C14), 112.5 (C3), 111.9 (C5), 98.5 (C6), 65.5 (C1), 63.2 (C10), 30.4 (C7), 25.2 (C9), 19.5 (C8) ppm. HRMS (TOF MS ES$^+$) calcd for $C_{15}H_{18}O_3N$ [MH$^+$]: m/z 260.1287, found 260.1279.

1-(pyridin-2-yl)-2-(pyridin-2-ylmethyl)-5-(tetrahydro-2H-pyran-2-yloxy)pent-3-yn-2-ol (76)

This compound is the bis-addition product of the reaction to **74a**. In the static procedure for **74a** with reactant **73**, **76** was isolated in up to 22% yield (R_f = 0.15 on 100% EA; R_f = 0.38 on 5:95 MeOH:AcOEt). **^{1}H-NMR ($CDCl_3$, 500 MHz)**: δ = 8.47 (d, 5 Hz, 2H), 7.60 (td, 7.6&1.6 Hz, 2H), 7.30 (d, 8 Hz, 2H), 7.15-7.12 (m, 2H), 4.37 (~s, 1H, THP), 4.05 (s, 2H, propargyl), 3.72-3.60 (m, 1H), 3.41-3.31 (m, 1H), 3.21 (s, 4H, 2×CH_2), 1.71-1.44 (m, 6H) ppm. **^{13}C-NMR ($CDCl_3$, 50 MHz)**: δ = 158.1, 158.0, 148.3 (×2), 136.3 (×2), 125.1 (×2), 121.7 (×2), 95.6, 88.3, 80.5, 70.5, 61.7, 53.6, 48.3, 48.2, 30.1, 25.2, 18.9 ppm. HRMS (TOF MS ES$^+$) calcd for $C_{21}H_{25}O_3N_2$ [MH$^+$]: m/z 353.1865, found 353.1871.

(4Z,6E)-5-(pyridin-2-yl)-6-(pyridin-2-ylmethylene)-1,9-bis(tetrahydro-2*H*-pyran-2-yloxy)nona-4-en-2,7-diyn-4-ol (77b)

This compound is the dehydration product of **77a**, and though barely visible in the crude products, it is the form isolated after chromatography. In the static procedure for **74a** with reactant **73**, **77b** was isolated in up to 10% yield (R_f = 0.40 on 100% EA). Even more than for **67a** and **66**, one is tempted to to wonder what heterocycles could be isolated from its coinage-metal catalysed cyclisation! **^{1}H-NMR ($CDCl_3$, 500 MHz)**: δ = 8.68 (dt, 4.6&0.9 Hz, 1H), 8.58-8.53 (m, 2H), 7.81-7.69 (m, 2H), 7.39 (d, 8 Hz, 1H), 7.29-7.21 (m, 3H), 4.53 (d, 13 Hz, 1H), 4.50 (t~s, 1H), 4.37 (t~s, 1H), 4.20 (d, 13 Hz, 1H), 4.12 (s, 2H), 3.75-3.61 (m, 2H), 3.44-3.32 (m, 2H), 1.90-1.46 (m, 12H) ppm. **^{13}C-NMR ($CDCl_3$, 50 MHz)**: δ = 158.2, 158.2, 156.3, 149.1, 146.0, 139.7, 136.8, 135.7, 135.5, 125.7, 125.1, 122.0, 121.9, 120.9, 120.2, 117.8, 98.0, 96.0, 93.3, 84.8, 66.4, 61.8, 61.6, 54.0, 30.3, 30.1, 25.3, 25.2, 19.0, 18.9 ppm. HRMS (TOF MS ES$^+$) calcd for $C_{30}H_{33}O_5N_2$ [MH$^+$]: m/z 501.2389, found 501.2393.

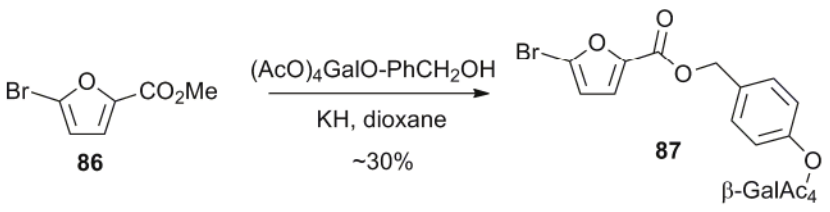


5-(((tetrahydro-2H-pyran-2-yl)oxy)methyl)quinolin-7-ol (75x)

Note: identical procedures but adopting the catalysts depicted (AuCl, Ag_2CO_3, CuI, AgOAc) resulted in formation of the (thermally-favoured) rearrangement product **75** instead. It is interesting that only this silver salt shows this activity: I suppose that the tosylate allows the silver to coordinate exceptionally well to the pyridine, blocking it from giving the nucleophilic addition product **75**. However, this reaction was only run once, and the products were identified on the basis of NMR, so there is limited rigour in this structural assignment. AgOTf gave degradation in the presence of PPh_3, but a clean product if no PPh_3 was added, whose structure could not however be assigned, but which was considered unlikely to be the desired furan.

To **74a** (5 mg) were added HPLC-grade DCM (2.5 mL), stirrer and catalyst (here, AgOTs; 8.7 mg, used in excess) and the reaction capped and stirred. After 24 hours, the solvent was evaporated and the crude taken into $CDCl_3$ for NMR analysis comparing the amount of **74a** and **75**, and checking if other pyridyl-bearing products were present, or any furans (two doublets, each 3 Hz, at about 5.7 and 5.3 ppm). With AgOTs I assigned 100% conversion to the structure **75x** based on: **^{1}H-NMR ($CDCl_3$, 500 MHz)**: δ = 8.22 (d, 7.3 Hz, 1H, H1), 7.35 (d, 8.9 Hz, 1H, H3), 7.11 (d, 2.2 Hz, 1H, H4), 6.90 (s, 1H, H5), 6.81 (~t, ~7 Hz, 1H, H2), 4.99+4.75 (each d, 13 Hz, 1H→2×H6), 4.72 (~m, 1H), 3.90-3.78 (m, 1H), 3.64-3.52 (m, 1H), 1.80-1.40 (m, 6H) ppm. DIMS(+): 260 (MH^+). The distinctively different aromatic peakshifts can be compared to those for **75**.

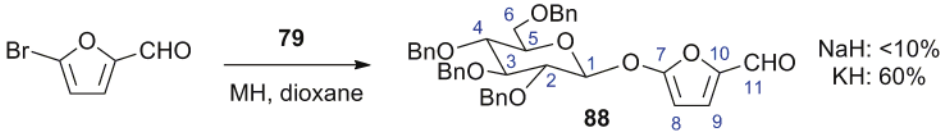
5.3. Furyl S_NAr



4-((tetra-O-acetyl)β-galactosyloxy)benzyl 5-bromofuran-2-carboxylate (87)

Similarly to the protocol of Morris *et al.*[35], to 4-((tetra-O-acetyl)β-galactosyloxy)benzyl alcohol (126 mg, 0.28 mmol) under Ar was added washed KH (25 mg, 0.57 mmol, 2 eq), and the flask capped. Dry dioxane (3 mL) was injected and the light tan solution was stirred at RT (10 min) before a solution of **86** (169 mg, 0.78 mmol, 3 eq) in dry dioxane (3 mL) was injected; the solution was stirred

overnight. To achieve an anhyrous quench, AcOH (43%wt in Et_2O, 55 mg, 0.4 mmol) was injected, SiO_2 added and the volatiles evaporated. Chromatography with solid deposition on 100:0→0:100 peth:EA returned excess **86** (R_f = 0.23 on 10:1 Cy:EA, chars red with Van) then **87** (45 mg in pure fraction, 0.07 mmol) as a colourless oil (R_f = 0.51 on 1:1 Cy:EA, Van). NMR confirms a 1:1 product of the reaction, with furan H3 and H4 intact, but without the methyl group, leading to structural assignment. The NMR spectrum is reproduced on page 286. **^{1}H-NMR** (200 MHz, $CDCl_3$): δ = 7.45-7.30 (m, 2H), 7.15 (d, 3.1 Hz, 1H, furan), 7.07-6.93 (m, 2H), 6.40 (d, 3.4 Hz, 1H, furan), 5.64-5.42 (m, 2H), 5.27 (s, 2H, OCH2), 5.18-5.03 (m, 2.5H), 4.40-4.00 (m, 3.5H), 2.17+2.08+2.08+2.03 (s+s+s+s, 4×3H) ppm. The furan peaks can be compared to the starting material **86**, showing that the ring shielding has barely changed during reaction: **86** **^{1}H-NMR** (200 MHz, $CDCl_3$): δ = 7.17 (d, 3.6 Hz, 1H), 6.50 (d, 3.8 Hz, 1H), 3.94 (s, 3H, methyl) ppm.

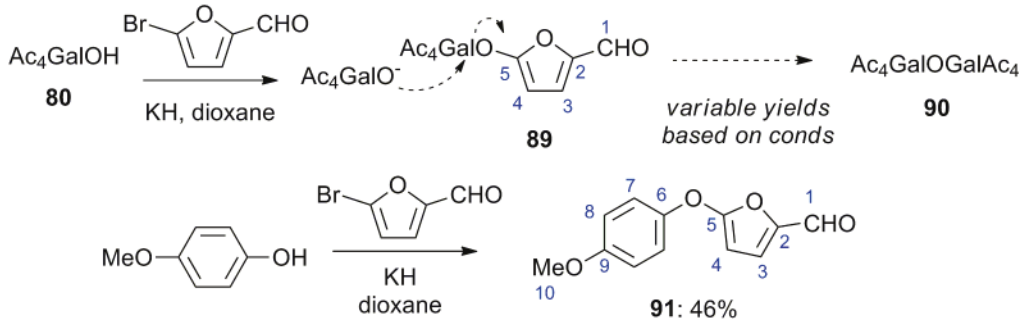


2-furfural-5-(tetra-O-benzyl)-β-glucoside (88)

Tetrabenzylated glucose **79** was synthesised as per Yamanoi *et al.*[36]. Similarly to the procedure for **87**, **79** (154 mg, 0.29 mmol) was reacted with washed KH (30 mg,0.75 mmol) in dry dioxane (2.5 mL) under Ar at RT for 10 min, giving a darkened solution. 5-bromofurfural (135 mg, 0.77 mmol) in dry dioxane (860 µL) was injected all at once and the reaction was stirred overnight. TLC was used to monitor the reaction (2:1 Cy:EA; Van: 5-bromofurfural: R_f = 0.64, barely any staining; **79**: R_f = 0.41 charring blue-green; **88**: R_f = 0.45 turning yellow almost without heating due to facile oxidation of the alkoxyfuran; anisaldehyde gave unreliable colouration.) The reaction was quenched into water (15 mL), brine (10 mL) added, and the mixture rapidly extracted with EtOAc (3×15 mL) giving dark red combined organic layers which were washed with brine, dried on Na_2SO_4 and filtered, and concentrated to a dark crude oil (197 mg). Chromatography with 5:1→2.5:1 Cy:EA recovers β-anomer **88** as a pale yellow oil which crystallises slowly (111 mg, 0.174 mmol, 60%). The alkoxyfuran had an ~1:4 ratio of two forms which I assigned as the α/β anomers, and tentatively concluded that the ratio was 1:4 α:β for this system (as per the β-favoured examples of Morris *et al.*[35]) based on the anomeric CH at 5.10 ppm (H, *J*=6 Hz) / 102 ppm (C) being the major form, and the minor anomeric form being at 5.45 ppm (H, no clear splitting) / 98.4 ppm (C). Use of NaH under identical conditions gave under 10% conversion as determined by crude NMR.

COSY and HSQC spectra were used to assign all peaks. NMR spectra are reproduced on page 287 ff. **^{1}H-NMR** (500 MHz, $CDCl_3$): δ = 9.44 (s, 1H, H11), 7.30-7.20 (m, 18H, 18×CH_{benzyl}), 7.12-7.07 (m, 2H, CH_{benzyl} + **d, 3.7 Hz, 1H, H9**), 5.64+5.55 (each d, 3.8 Hz; ratio ~3:1, total 1H, H8), 5.45+5.09 (m+~d, 5.9 Hz, 1:4 ratio, total 1H, H1), 4.93-4.34 (m, 8H, 4×$CH_{2,\ benzyl}$), 4.06 – 3.51 (m,

6H, 2×H6 + H4 + H5 + H3 + H2) ppm. In **^{13}C-NMR** only the peaks for the major anomer are given: (125 MHz, $CDCl_3$): δ = 175.4 (C11), 162.7 (C7), 144.2 (C10), 138.0 + 137.6 + 137.6 + 137.4 (4×C_{Bn}), 128.4-127.6 (20×CH_{Bn} + C9), 102.2 (C1), 89.1 (C8), 84.0 (C2), 81.0 (C3), 76.9 (C4), 75.5(×2) + 74.9 + 74.9 + 73.3 (4×$CH_{2,\ benzyl}$ + C5), 68.2 (C6) ppm.

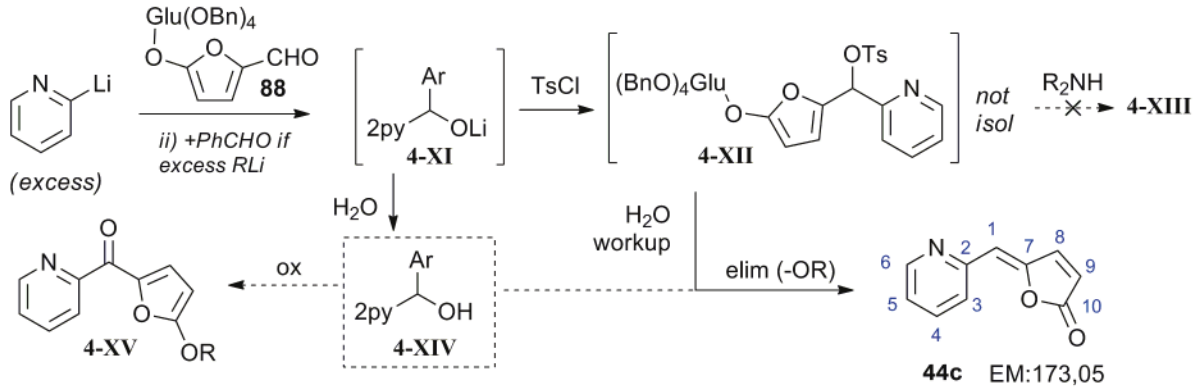


Reaction of 80 to 5-(tetra-O-acetylgalactosyloxy)-2-furfural (89) and/or bis(tetra-O-acetylgalactosyl) ether (90)

Tetraacetylgalactose **80** was cleanly delivered by by selective anomeric deacetylation of galactose pentaacetate by a literature protocol[37]. Similarly to the procedure for **88**, **80** (360 mg, 1.03 mmol) was reacted with KH (78 mg washed, 2 mmol) and 5-bromofurfural (475 mg, 2.71 mmol). The blackened mixture was quenched into a rapidly-stirred mixture of 1:1:1 Cy:EtOAc:$NaHCO_3$ and extracted with EA; the combined organic layers were washed with pH~7 phosphate buffer, brine, dried on sodium sulfate, filtered, and concentrated. TLC (1:1 Cy:EA, $KMnO_4$) shows three spots: bromofurfural (R_f = 0.69), product **89** (R_f = 0.51) and **80** (R_f = 0.38) [note that **90** has R_f = 0.56 in this eluant]. Crude NMR shows an approximately 1:7 ratio of **89**:**80** (yield **89** ~13% non-brsm) with no **90**. Alternatively, when the bromofurfural solution was added dropwise into the ROK, **89**:**90**:**80** were recovered at about 1:1:1 (yield **89** ~5% non-brsm). Alternatively, when ROK was added dropwise into the bromofurfural, recovery of **89**:**90**:**80** was around 0:1:10. *As these reactions were not considered successful, only simple characterisation data are given*: **89**: **^{1}H-NMR** (500 MHz, $CDCl_3$): δ = 9.37 (s, 1H), **7.19 (d, 3.7 Hz, 1H, H4)**, **5.72 (d, 3.8 Hz, 1H, H3)**, 5.31 (~d, 7.7 Hz, 1H, $H_{anomeric}$), 4.19-4.05 (m, 6H), 2.14-1.98 (4×s, 12H, 4×CH_3) ppm. For **90**, note that the presence of up to 4 anomer pairs brings spectral complications; however, the products appeared to be one of two type only, I suspect both are anchimerically-assisted β-glycosylations, of an incoming nucleophile that is itself more likely to be β than α; the major products' peaks are in bold: **^{1}H-NMR** (500 MHz, $CDCl_3$): δ = **6.40**+5.73 (s, 1.5H + d, *8.3 Hz*, 0.5H; 2H anomerics), **5.52**+5.46 (d, 10.8 Hz, 1.5H + d, 2.9 Hz, 0.5H; 2H), **5.40**-**5.31**+5.11 (m, 3.5H + dd, 10.4&3.3 Hz, 0.5H; 2H), **4.38**+4.10 (t, 6.6 Hz, 1.5H+m, overlapped, 0.5H; 2H), 4.25-4.02 (m, 4H, exocyclic), **2.19**+2.15+**2.07**+**2.05**+**2.03**+2.01 (6 overlapping singlets; 24H; acetyls) ppm. **^{13}C-NMR** (125 MHz, $CDCl_3$): δ = 170.4+170.3+170.2+170.2+170.0+169.9+169.4+169.0 (8×acetyl CO_2); 92.2, **89.7** (anomeric CHs), 71.7, 70.9, **68.8**, 67.8, **67.4**, **67.4**, 66.8, **66.5**, **61.3**, 61.0, **20.9**, 20.8, 20.6-20.6 (3×acetyl CH_3 overlapping) ppm.

5-(4-methoxy)phenoxy-2-furfural (91)

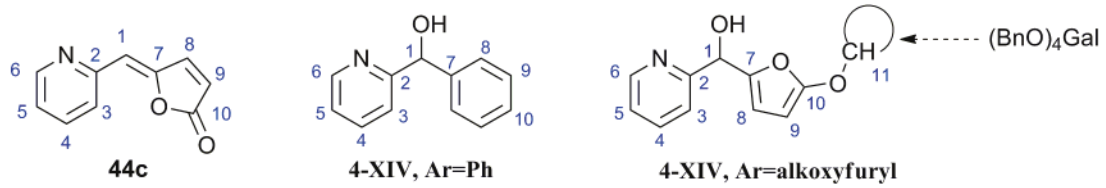
4-methoxyphenol was dried on molecular sieves in Et_2O, then similarly to the procedure for **88**, 226 mg (1.82 mmol) was reacted with KH (73 mg, 1.82 mmol, 1 eq) and 5-bromo-2-furfural (670 mg, 3.76 mmol, 2 eq) in dioxane (total vol. 13 mL. Gas evolution during the KH reaction was rapid. The solution stayed pale for a long time, in contrast to reactions with sugar substitutions that darken rapidly. The reaction was quenched into phosphate buffer (pH~6.8, 0.35 M, 15 mL), and worked up as for **88** to give a black crude oil (700 mg), separated with 5:1→2:1 Cy:EA yielding first bromofurfural, then the phenol, then **91** as colourless crystals (186 mg, 0.84 mmol, 46%; $R_f = 0.53$ on 1:1 Cy:EA, Haness) which rapidly brown under air, and which become a black tar after several days in the fridge. **^{1}H-NMR** (500 MHz, $CDCl_3$): $\delta = 9.16$ (s, 1H, H1), **7.09 (d, overlaps, 1H, H3)**, 7.00-6.90 (m, 2H, 2×H7), 6.80-6.67 (m, 2H, 2×H8), **5.29 (d, 3.8 Hz, 1H, H4)**, 3.63 (s, 3H, 3×H10) ppm. **^{13}C-NMR** (125 MHz, $CDCl_3$): $\delta = 175.61$ (C1), 164.1 (C5), 157.51 (C2), 150.5 (C9), 144.5 (C6), 120.6 (2×C7), 115.07 (2×C8), 114.9 (C3), 88.9 (C4), 55.7 (C10) ppm. The C3 peak seems quite low; if I have not mis-assigned it, this may be due to electron-donation from the phenol (cf. shift in glycoside **88**).



5-(2'-pyridylmethylene)-2-furanone (44c)*

Representative reaction without excess 2pyLi (this run carried out by intern Nils Aronssohn): to 2-bromopyridine (43 mg, 0.27 mmol) in dry THF (3 mL) under Ar at -80°C was added *n*BuLi (1.6M in hexanes, 0.19 mL, 0.30 mmol). The pink solution was stirred for 15 min then injected dropwise into a solution of **88** (198 mg, 0.31 mmol) in dry THF (2 mL) at -80°C, then stir for 2 h, allowing the temperature to rise to 0°C. A solution of TsCl (200 mg, 0.94 mmol) in dry THF (2 mL) was injected and the reaction stirred for a further 2 h. The reaction was quenched with degassed water (80 mL) still under Ar, and extracted with EtOAc (3×25 mL). The combined organic phases were dried on sodium sulfate, filtered, and concentrated to a brown oil (89 mg) which was chromatographed on silica (10:1→2:1 Cy:EA) to give **44c** (~30 mg with minor contaminants; ~0.17 mmol, ~65%). DIMS(+): 174 (MH^+). The following NMR data were confirmed to a reference giving an orthogonal method of synthesis[38]: **^{1}H-NMR** (500 MHz, $CDCl_3$): δ = 8.80 (d, 5.5 Hz, 1H, H9), 8.62 (dd, 4.8&1.7 Hz, 1H,

H6), 7.68 (td, 7.7&1.9 Hz, 1H, H4), 7.28 (d, 3.5 Hz, 1H, H3) 7.19 (ddd, 7.6&4.8&1.1 Hz, 1H, H5), 6.58 (d, 1.5 Hz, 1H, H1, couples weakly to H8), 6.34 (dd, 5.5, 1.8 Hz, 1H, H8, couples weakly to H1). **^{13}C-NMR** (125 MHz, $CDCl_3$): δ = 169.5 (C10), 153.1 (C2, C7), 150.12 (C6), 143.8 (C8), 136.8 (C4), 126.0 (C3), 123.0 (C5), 122.0 (C9), **112.9 (C1, -CH=)** ppm.



44c **4-XIV, Ar=Ph** **4-XIV, Ar=alkoxyfuryl**

2-pyridyl-phenylmethanol (4-XIV, Ar=Ph)*

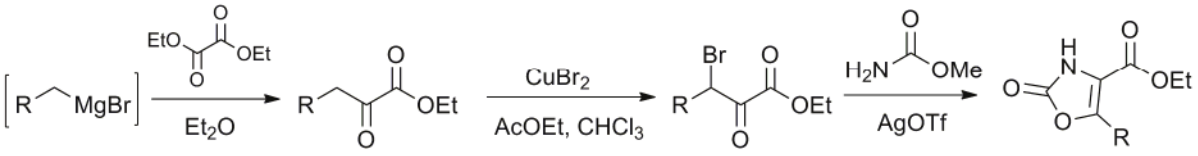
Applying the same procedure as for **44c** but using benzaldehyde allowed confirmation of (non-eliminated) 2-pyridyl-phenylmethanol. DIMS(+) = 186 (MH^+). **^{1}H-NMR** (500 MHz, $CDCl_3$): δ = 8.54 (~d, 4.6 Hz, 1H, pyH6), 7.62 (~t, 7.9 Hz, 1H, H4), 7.38-7.25 (m, overlapped, est. 5H), 7.23 (m, 1H, H3), 7.17 (m, 1H, H5), **5.78 (s, 1H, CHOH)**. **^{13}C-NMR** (125 MHz, $CDCl_3$): δ = 160.7 (C2), 147.7 (C6), 143.0 (C7), 138.1 (C4), 128.5 (C8×2), 127.9 (C10), 127.0 (C8×2), 122.3 (C3), 121.3 (C5), **73.5 (C1, CHOH)** ppm.

Spectral features in crude NMR assigned to 4-XIV, Ar=alkoxyfuryl

Salient ^{1}H-NMR (200 MHz, $CDCl_3$): No elimination product **44c** doublets at 8.80, 6.58, 6.34 ppm; no peak for the anomeric proton of **79** in the region ~5.3 ppm. Distinct peaks: δ = 8.60 (~d, 5.1 Hz, 1H, H6), 7.62 (~t, 7.9 Hz, 1H), 6.44 (d, 9.5 Hz, 1H, H1), 6.09 (d, 3.8 Hz, 1H, H8), 4.75 (d, overlap, ~4.5 Hz, est 1H, H9).

5.4. Oxazolones and oxazolidinones

5.4.1. 5-chelating oxazolone scaffolds



94: R=nC_5H_{11} (38%)
95: R=Bn (comm)

96: R=nC_5H_{11} (82%)
97: R=Bn (84%)

98: R=nC_5H_{11} (64%)
99: R=Bn (60%)

ethyl 3-bromo-2-oxooctanoate (96)*

Bromination was performed by the method of King *et al.*[39,40]: ethyl 2-oxooctanoate **94** (4.3 g, 23.2 mmol, freshly synthesised according to Weinstock *et al.*[41] in 38% yield) was dissolved in $CHCl_3$ (120 mL) and EtOAc (200 mL) and $CuBr_2$ (15.6 g, 70 mmol, 3 eq) added. The rapidly-stirred mixture

was heated capped under condenser at 86°C overnight, with a wide air-outlet needle to void the evolved HBr into an efficient fume hood, though a base trap could also be used. The resultant mixture of grey solids and green organic solution was filtered over 15 cm silica and concentrated to yield a yellow crude oil. The crude was chromatographed on 20:0→20:2 Cy:EA gradient to yield **96** as a mid-yellow oil (5.02 g, 18.9 mmol, 82%). **TLC**: $R_f = 0.39$ (10:1 Cy:EA, $KMnO_4$). NMR agrees with reference.[42]

ethyl 3-bromo-2-oxo-4-phenylbutanoate (97)*

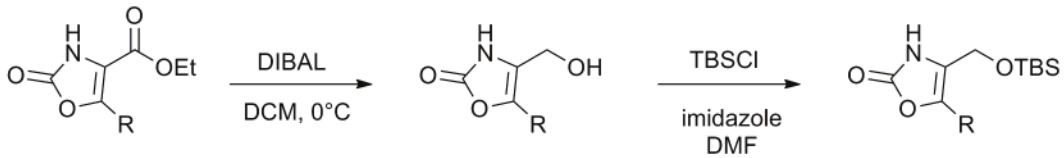
Commercial ethyl 2-oxo-4-phenylbutanoate **95** (9.9 g, 48 mmol) was treated similarly to **96** to yield pale yellow **97** (11.3 g, 40 mmol, 84%). **TLC**: $R_f = 0.18$ (20:1 pent:EA, $KMnO_4$ or DNPH). NMR agrees with reference.[42]

5-pentyl-4-carboethoxy-4-oxazolin-2-one (98)*

According to the method of Hoffman, Okonya *et al.*[42,43], to **96** (3.85 g, 14.5 mmol) was added dry toluene (250 mL), methyl carbamate (5.67 g, 75.6 mmol, 5 eq), AgOTf (3.71 g, 14.45 mmol, 1 eq) and catalytic *p*TsOH monohydrate (347 mg, 0.13 eq). The mixture was refluxed under Ar overnight at 124°C to yield a yellow turbid mixture which was poured carefully into a separating funnel, and after rinsing the residues with AcOEt (100) the organic layer was washed with water (4×50 mL), brine (4×30 mL), dried on Na_2SO_4, filtered and evaporated to a yellow crude oil which solidified on standing (3.75 g). The crude was carefully chromatographed on 9:1 DCM:EA to run off several faster impurities before giving **98** (2.10 g, 9.24 mmol, 64%) as a beautifully crystallising colourless solid tinged with pale yellow. **TLC**: $R_f = 0.42$ (9:1 DCM:EA, Haness, DNPH or UV). **DIMS**: (+) 228 (MH^+), 250 (MNa^+), 477 (M_2Na^+); (-) 226 (M-H). NMR agrees with reference.[42]

5-benzyl-4-carboethoxy-4-oxazolin-2-one (99)*

Similarly to **98**, **97** (11.3 g, 40 mmol) was reacted with methyl carbamate (16 g, 213 mmol), AgOTf (9.97 g, 38.8 mmol) and *p*TsOH (0.796 g, 0.1 eq) in ordinary toluene. Chromatography on 9:1 DCM:EA yielded **99** as beautifully fluffy white crystals (5.9 g, 60%). NMR agrees with reference.[42]



98: R=nC_5H_{11}
99: R=Bn

100: R=nC_5H_{11}, 83%
101: R=Bn, 80%

102: R=nC_5H_{11}, 51%
103: R=Bn, 99%

5-pentyl-4-hydroxymethyl-4-oxazolin-2-one (100)

Following the procedure of Okonya *et al.*[42], to **98** (1452 mg, 6.38 mmol) under Ar at 0°C were added dry DCM (16 mL) then, dropwise, DIBAL (1 M in Cy, 25 mL, 25 mmol, 4 eq). The mixture was stirred between -10°C and 0°C for four hours, then poured into vigorously stirred cold HCl (2 M, 110 mL, 0°C). The aqueous layer was extracted with EtOAc (2×40 mL), and the combined organic layers washed with brine (2×40 mL), dried on Na_2SO_4 and filtered and concentrated. The yellowed powder was then triturated with 3:1 pentane:EA (6×10 mL; note: this gives better behaviour than Cy:EA) to yield **100** as a colourless fluffy solid (975 mg, 5.26 mmol, 83%) from which it was rather difficult to remove traces of water. The compound is rather responsive to static electricity. **DIMS**: (+) 186 (MH^+), 208 (MNa^+), 371 (M_2H^+); (-) 184 (M-H). **^{1}H-NMR** (200 MHz, $CDCl_3$): δ = 9.25 (s br, 0.9H, NH), 4.37 (s, 2H), 3.55 (s br, 1H, OH), 2.39 (t, 7.3 Hz, 2H), 1.65-1.45 (m, 2H), 1.35-1.20 (m, 4H), 0.88 (t, 6.6 Hz, 3H) ppm. **^{13}C-NMR** (50 MHz, $CDCl_3$): δ = 156.1, 138.7, 120.2, 53.4, 31.1, 27.2, 24.4, 22.3, 14.0 ppm.

5-benzyl-4-hydroxymethyl-4-oxazolin-2-one (101)*

Similarly to **100**, **101** was produced as fluffy colourless crystals in 80% yield (1.44 g scale) from DIBAL reduction of **99**. NMR agrees with reference.[42]

5-pentyl-4-((*tert*-butyldimethylsilyloxy)methyl)-oxazol-2(3H)-one (102)

According to the procedure for **103**, **100** (145 mg, 0.78 mmol) was reacted with imidazole (632 mg, 9.3 mmol) and TBSCl (430 mg, 2.84 mmol) in DMF (6 mL). The crude was chromatographed on 10:1→1:1 Cy:EA giving **102** as a pale yellow oil (118 mg, 0.40 mmol, 51%). The compound is unusually responsive to static electricity. **^{1}H-NMR** (500 MHz, $CDCl_3$): δ = 9.23 (s, 1H), 4.41 (s, 2H), 2.41 (t, 7.4 Hz, 2H), 1.62-1.55 (m, 2H), 1.37-1.27 (m, 4H), 0.92 (s, 9H), 0.91 (t, 7.2 Hz, 3H), 0.13 (s, 6H) ppm. **^{13}C-NMR** (125 MHz, $CDCl_3$): δ = 157.0, 137.5, 120.1, 54.6, 31.2, 27.3, 25.9 (×3), 24.6, 22.4, 18.4, 14.0, -5.2 (×2) ppm.

5-benzyl-4-((*tert*-butyldimethylsilyloxy)methyl)-oxazol-2(3H)-one (103)

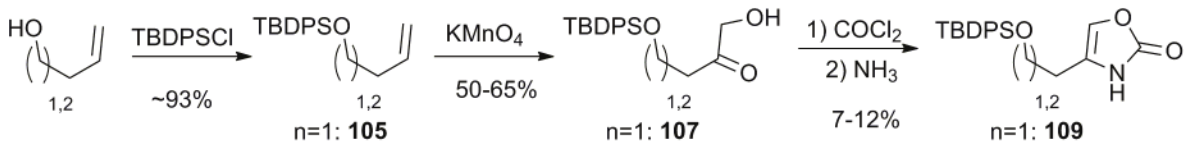
To **101** (292 mg, 1.45 mmol) was added imidazole (737 mg, 10.8 mmol, 6 eq), TBSCl (666 mg, 4.3 mmol, 3 eq), and ordinary DMF (8 mL) under argon. The solution was stirred overnight at 35°C, evaporated at 75°C to 2 g and then dissolved into EA and washed with water. The aqueous layer was

back-extracted with EA (4×10 mL), the combined organic layers were washed with water (2×10 mL), brine (2×10 mL), dried over Na_2SO_4, filtered and concentrated under high vacuum to a brown-yellow oily solid (465 mg, 1.45 mmol, 99%) which was shown to be pure **103** by NMR. Subsequent runs performed later on the same starting material delivered yields of between 75-90%, and *required* further purification, which was performed by trituratation with cold (eg -20°C - -70°C) pentane with a drop of $CHCl_3$ added, yielding colourless crystals, though removing a substantial amount of product which was also pure enough for further work. The silyl group seems to be more labile than is desirable, resulting in a distinctively sharp odour. The compound should not be excessively heated, nor heated under reduced pressure, or decomposition becomes extensive: this is possibly a result of residual water which was difficult to remove; but cyclohexane azeotroping, followed by pentane azeotroping of the residual cyclohexane if desired (not required), effected this. The compound is unusually responsive to static electricity. **^{1}H-NMR** (500 MHz, $CDCl_3$): δ = 7.47 (s br, 1H, NH), 7.32-7.20 (m, 5H), 4.39 (s, 2H), 3.72 (s, 2H), 0.83 (s, 9H), 0.05 (s, 6H) ppm. **^{13}C-NMR** (125 MHz, $CDCl_3$): δ = 155.6, 136.1, 135.2, 128.9 (×2), 128.7 (×2), 127.2, 120.5, 54.5, 31.3, 25.9 (×3), 18.3, -5.3 (×2) ppm.

dihydroxyacetone dimerisation/cracking*

Commercial dihydroxyacetone dimer has a complex ^{1}H-NMR spectrum in the region 3.8-3.4 ppm composed of overlapping pairs of doublets with J=11.7 Hz, centred at 3.50 and 3.60 ppm, plus other peaks; a J=11.7 Hz doublet at 4.13 ppm is also seen, as well as a distinctive singlet at 4.31 ppm. Commercial dihydroxyacetone dimer (1.51 g) was refluxed in technical acetone (10 mL), in which it is insoluble, under air.[44] After six hours, the suspension has dissolved indicating cracking of the dimer to the monomeric product, confirmed by HSQC. **^{1}H-NMR** (200 MHz, $CDCl_3$): δ = 4.34 (d, 5.3 Hz, 4H, CH_2), 4.03 (t, 5.3 Hz, 2H, OH) ppm. **^{13}C-NMR** (125 MHz, $CDCl_3$): δ = 211.0, 66.0 (×2) ppm. This monomer should be used without too much concentration/heating, or the dimer may reform.

5.4.2. 6- and 7-chelating oxazolone scaffolds



Dakin *et al.*[45] have reported the syntheses and characterisations of **105**, **107** and **109**, and **106** is known, however the compounds **108** and **110** are unreported.

4-((tert-butyldiphenylsilyl)oxy)-1-hydroxybutan-2-one (108)

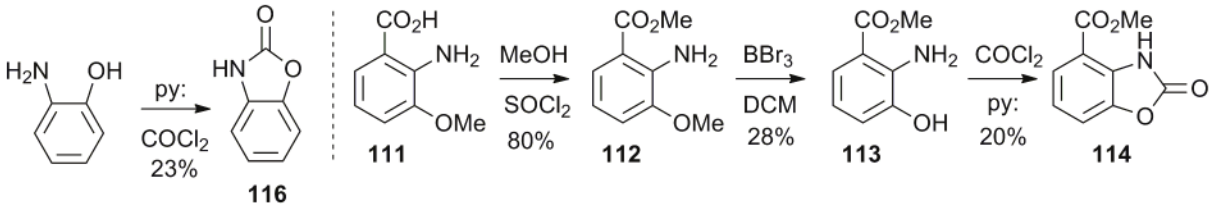
According to the method of Bonini *et al.*[46], a solution of $KMnO_4$ (8.42 g, 53.3 mmol) in water (34 mL) and acetone (102 mL) was prepared. To but-3-enyloxy-tert-butyl-diphenyl-silane[47] (10.58 g, 34.1 mmol) was added acetone (265 mL), water (60 mL), and AcOH (10 mL), then the $KMnO_4$

solution added dropwise over four hours. Stirring was continued for 3 hours until the last of the red tinge disappeared, then the flask was left on its side to sediment MnO_2. The supernatant was carefully decanted for filtration through sand and Celite. The MnO_2 was rinsed with Et_2O (3×40 mL) which also aids filtration. After evaporating most of the acetone and ether, sat. $NaHCO_3$ was added cautiously to quench AcOH, and the aqueous phase extracted with Et_2O (4×40 mL). The combined organic extracts were washed with water and brine, and dried on SiO_2/K_2CO_3 and filtered and concentrated. NMR revealed 55% product ketol, 15% starting material, 1% aldehyde, and approx. 30% cleavage and hydrolysis products. Chromatography on 100:0→8:1 Cy:EA ran off the byproducts before returning **108** as a colourless oil (5.65 g, 16.5 mmol, 50%). **TLC**: $R_f = 0.15$ (10:1 Cy:EA; chars red with vanillin). **^{1}H-NMR** (200 MHz, $CDCl_3$): $\delta = 7.73\text{-}7.55$ (m, 4H), 7.50-7.31 (m, 6H), 4.32 (~s, 2H), 3.96 (td, 6.1 & 2.5 Hz, 2H), 2.60 (td, 6.0 & 2.1 Hz, 2H), 1.04 (~s, 9H) ppm. **^{13}C-NMR** (50 MHz, $CDCl_3$): δ = 208.9, 135.6 (×4), 133.2 (×2), 130.0 (×2), 127.9 (×4), 69.3, 59.6, 31.6, 26.9 (×3), 19.2 ppm. The doubling of some peaks in the ^{1}H spectrum was interpreted as partial dimerisation of the hydroxyketone, probably during concentration.

4-(2-((tert-butyldiphenylsilyl)oxy)ethyl)oxazol-2(3H)-one (110)

A sample of **108** (3.21 g, 9.3 mmol) was twice azeotroped with toluene to remove residual water. The NMR revealed an anhydrous sample, but clearly showed doubling of the peaks (cf. dimerisation above). According to the method of Dakin *et al.*[45], to this were added benzene (17 mL; curiously, replacement with toluene had not given desired product), N,N-dimethylaniline (3 mL), then phosgene (5.5 mL, 2 M in toluene, 11 mmol) added dropwise at 0°C and the solution stirred for 35 min. Ammonia (40 mL, 30%, aqueous) was added carefully, and the solution stirred for 20 minutes, then a further dose (10 mL, 20%, aqueous) added and the solution stirred for a further ten minutes. Saturated KH_2PO_4 (80 mL) was added, then conc H_2SO_4 until pH~3 was obtained, and EtOAc (10 mL) was immediately added to secure the products out of the aqueous phase, then the solution returned to RT. The aqueous phase was twice more extracted with EtOAc (25 mL), and the combined organic extracts were washed with water and brine, and dried on $MgSO_4$ and filtered and concentrated. The crude contained multiple spots and was carefully chromatographed with 100:0→1.75:1 Cy:EA, revealing with vanillin. Starting material **108** (350 mg, 1.04 mmol) was first recovered, then product **110** was obtained as a yellow-tinged oil (375 mg, 1.02 mmol, 12% brsm). Much of the species present in the mixture seemed to be TBDPS cleavage products, which contrasts sharply with the published procedure where an 85% yield was obtained. **TLC**: $R_f = 0.50$ (1:1 Cy:EA; chars red-brown with vanillin, UV-active, and clearly hydrophobic). **^{1}H-NMR** (200 MHz, $CDCl_3$): $\delta = 8.71$ (s br, 1H), 7.73-7.55 (m, 4H), 7.46-7.35 (m, 6H), 6.66 (m, 1H), 3.82 (t, 6.8 Hz, 2H), 2.54 (td, 6.5 & 0.9 Hz, 2H), 1.06 (s, 9H) ppm. **^{13}C-NMR** (50 MHz, $CDCl_3$): δ = 157.0, 135.6 (×4), 133.0 (×2), 130.1 (×2), 128.0 (×4), 125.1, 62.0, 60.6, 27.3, 27.0 (×3), 19.2 ppm.

5.4.3. Benzoxazolone scaffolds



Benzoxazolone **114** was formed from **113** exactly as per the protocol of Fielder *et al*.[48]

Benzoxazolone (116)*

Similarly to Fielder *et al*.[48], to 2-aminophenol (1.47 g, 13.5 mmol) was added dry pyridine (20 mL) then phosgene (20% in PhMe, 10 mL, 19 mmol) in the dark. Vigorous heating resulted, and the bright red solution was stirred for 24 h. The remaining phosgene and any possible uncyclised chloroformate/isocyanide/chloroformamide species were quenched with dry EtOH (3 mL) and the volatiles evaporated. The crude paste was treated with charcoal suspension in acetone and filtered and evaporated (30 mbar, 60°C) to give 3.6 g of browned solid only faintly smelling of pyridine. The crude solid was chromatographed with 10:1→1:1 Cy:EA, revealing with $KMnO_4$, to provide **116** as a white solid (425 mg, 3.14 mmol, 23%). TLC: $R_f = 0.23$ (1:1 Cy:EA; UV). ^{1}H-NMR (200 MHz, $CDCl_3$): $\delta =$ 10.03 (s br), 7.26 (d, 7.8 Hz, 1H), 7.24-7.14 (m, 4H) ppm. ^{13}C-NMR (50 MHz, $CDCl_3$): $\delta =$ 156.6, 144.0, 129.6, 124.4, 122.9, 110.4, 110.3 ppm. The main byproduct isolated was the ethoxy carbamate, indicating incomplete reaction, however this was not successfully converted to the benzoxazolone in a trial reaction heating with Cs_2O_3 in DMF.

methyl 2-amino-3-methoxybenzoate (112)*

To commercial 2-amino-3-methoxybenzoic acid **111** (290 mg, 1.73 mmol) in dry MeOH (20 mL) was added slowly with cooling $SOCl_2$ (1.06 g, 6 eq), then the mixture refluxed overnight at 82°C under a KOH guard. NaOH/K_2HPO_4 were added to give pH~8 and the methanol mostly evaporated. The **112** as a colourless solid (250 mg, 1.4 mmol, 80%). **^{1}H-NMR** (200 MHz, $CDCl_3$): $\delta = 7.6$-7.4 (m, 1H), 6.9-6.75 (m, 1H), 6.65-6.45 (m, 1H), 3.86 (s, 3H) ppm. **DIMS(+)**: 182 (MH^+). **TLC**: $R_f = 0.45$ (1:1 Cy:EA, $KMnO_4$). A similar transform effected by cat. H_2SO_4 in MeOH gave only 30% yield.

methyl 2-amino-3-hydroxybenzoate (113)*

The known title compound[48] was prepared by a different procedure to the reference. To **112** (250 mg, 1.4 mmol) were added dry DCM (6 mL), then dropwise, BBr_3 (1M in DCM, 4.2 mL, 3 eq) at 0°C. The mixture was stirred for 22 h, then quenched in an icebath into 2M NaOH (6 mL), to which was added $NaHCO_3$ and phosphate buffers to achieve pH~8. The aqueous phase was extracted into DCM (4×15 mL), then the combined organic layers washed with brine and dried on Na_2SO_4, filtered and

concentrated to yield **113** as a yellow solid (63 mg, 0.38 mmol, 28%) whose NMR matches the reference[49]. **TLC**: R_f = 0.3 (10:1 Cy:EA, UV).

5.4.4. 5-chelating oxazolidinone scaffold

Ester **123**[50] and alcohol **124**[51] were synthesised by the reported methods. That alcohol synthesis is far better than questionable, four-step approach of Hwang *et al.*[52] where a crucial step for chirality retention relies on drawing a serinol left-to-right (paper since appears retracted).

4-(((tert-butyldimethylsilyl)oxy)methyl)oxazolidin-2-one (125)

As per the method for **103**, in one run, **125** was obtained from TBSCl (5.76 g, 2.6 eq) silylation of **124** (1.696 g, 14.4 mmol) in DMF (50 mL) with imidazole (5.8 g), with trituration, then scratching recrystallisation from pentane at -80°C giving a first pure crop of colourless crystals (1.92 g, 57%) which were difficult to dry; a second crop of less pure crystals (~250 mg product) could be obtained with difficulty. In a second run following the same procedure the yield was improved to 72% by better recrystallisation, still using extreme cold to separate the TBS-bearing reaction products. These display complex temperature-dependent solubility as they form a ternary mixture with the desired product and the solvent, but heating and cooling at about 20°C/min between 40°C and -80°C usually allows good accumulation of the flocculent crystals which are pure product (the TBS byproducts are liquid-phase only). The compound is unusually responsive to static electricity. It stains bright red on TLC with Van. **^{1}H-NMR** (500 MHz, $CDCl_3$): δ = 5.95 (s br, 1H), 4.43 (t, 8.7 Hz, 1H), 4.16 (dd, 8.7&4.1 Hz, 1H), 3.95-3.87 (m, 1H), 3.60 (d, 5.5 Hz, 2H), 0.88 (s, 9H), 0.06 (s, 6H) ppm. **^{13}C-NMR** (125 MHz, $CDCl_3$): δ = 160.0, 67.2, 64.9, 53.8, 25.9 (×3), 18.3, -5.4 (×2) ppm.

5.4.5. 6- and 7-chelating oxazolidinone scaffolds

Similar to the route of Hou *et al.*[53], aspartic and glutamic acid bis-esterifications ($SOCl_2$/MeOH) and carbamylations (BnOCOCl, $NaHCO_3$) by standard methods gave species **127a/b** in 87/90% yield on 24 mmol scale for 2 steps. Reduction to diols by $NaBH_4/CaCl_2$ gave 94/78% yields of **128a/b**, carried over to the next steps.

4-(2-((tert-butyldimethylsilyl)oxy)ethyl)oxazolidin-2-one (130a)

The first step was performed as per Hou *et al.*[53]. To bis-alcohol **128a** (4.78 g, 20 mmol) were added at 0°C under argon dry THF (50 mL) and NaH (60% in oil, 1.90 g, 47 mmol, 2.4 eq), and the mixture stirred vigorously overnight warming to RT. Some solid mass had formed; the reaction was now refluxed at 84°C for 7 hours. DIMS(+) confirms the cyclised product is present (154 Th, MNa^+). Rather than isolating the product **129a**, after cooling to RT, imidazole (461 mg, 0.15 eq) and TBSCl (7.15 g, 48 mmol, 1.1 eq relative to RONa) were added and the reaction stirred overnight. The reaction was then quenched into phosphate buffer (pH~7, 0.4 M, 40 mL), and most of the THF evaporated from the now biphasic mixture. Extraction with ethyl acetate gave a pale yellow crude oil (9.1 g) which NMR indicated contained <4 mmol of the desired product **130a**, however total evaporation and

multiple extractions on the ~10 g residue of the aqueous phase, using a variety of conditions, obtained less than 200 mg of product along with a variety of contaminants so the original crude was retained, chromatographed on 100:0→1:1 Cy/EA gradient, giving **130a** as a pale yellow oil (390 mg, 1.52 mmol, 7.6%) with R_f = 0.57 (1:1 Cy:EA, Haness). **^{1}H-NMR** (200 MHz, $CDCl_3$): δ = 6.12 (s br, 1H), 4.47 (t, 7.5 Hz, 1H), 4.04 (q, 7.4 Hz, 2H) 3.70 (~t, 5.6 Hz, 2H), 1.90-1.68 (m, 2H), 0.87 (s, 9H), 0.04 (s, 6H) ppm. **^{13}C-NMR** (50 MHz, $CDCl_3$): δ = 159.9, 70.8, 65.2, 51.9, 37.7, 26.0 (×3), 18.2, -5.4 (×2) ppm.

4-(3-((tert-butyldimethylsilyl)oxy)propyl)oxazolidin-2-one (130b)

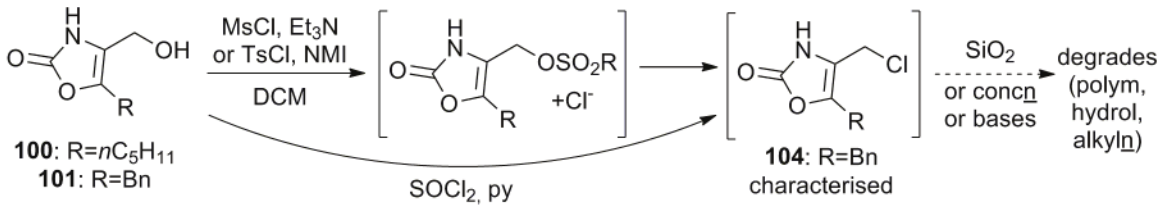
Similarly to the procedure for **130a**, bis-alcohol **128b** (4.10 g, 16.2 mmol) was cyclised with NaH (60% in oil, 1.54 g, 38 mmol, 2.4 eq), then reacted with TBSCl (6.14 g, 41 mmol, 2.5 eq) and imidazole (402 mg, 0.15 eq), and extracted to give a crude oil (9.2 g) which was similarly chromatographed affording **130b** as a pale oil (1.02 g, 3.9 mmol, 24%) with R_f = 0.39 (1:1 Cy:EA, Haness). **^{1}H-NMR** (200 MHz, $CDCl_3$): δ = 6.03 (s br, 1H), 4.48 (t, 8.2 Hz, 1H), 4.04 (q, 8.1 Hz, 2H) 3.70 (~t, 5.5 Hz, 2H), 1.77-1.50 (m, 4H), 0.88 (s, 9H), 0.04 (s, 6H) ppm. **^{13}C-NMR** (50 MHz, $CDCl_3$): δ = 160.1, 70.5, 62.6, 52.7, 32.5, 28.7, 26.0 (×3), 18.4, -5.2 (×2) ppm.

5.4.6. Non-O-alkylated electrophilic derivatisations

Oxazolidinone tosylate **126**[51] was synthesised by the reported method.

3-hydroxy-2-(4-methylphenylsulfonamido)propyl 4-tosylate (126x)

126x was isolated as the ring-opening byproduct of the reaction giving tosylate **126**[51], in approx. 10% yield after column. **TLC**: R_f = 0.55 (5:7 Cy:EA, $KMnO_4$). **^{1}H-NMR** (500 MHz, $CDCl_3$): δ = 7.71 (~d, 8.1 Hz, 4H), 7.32 (d, 8.3 Hz, 2H), 7.26 (d, 8.2 Hz, 2H), 5.50 (d, 8.0 Hz, 1H, NH), 4.02 (dd, 10.2&4.8 Hz, 1H), 3.94 (dd, 10.3&6.3 Hz, 1H), 3.65-3.59 (m, 1H), 3.52-3.41 (m, 2H=1H+1H, overlap), 2.56 (s bar, 1H, OH), 2.32 (s, 3H), 2.40 (s, 3H) ppm. **^{13}C-NMR** (125 MHz, $CDCl_3$): δ = 145.5, 143.9, 137.2, 132.2, 130.2 (×2), 130.0 (×2), 128.1 (×2), 127.2 (×2), 68.1, 60.8, 53.4, 21.8, 21.7 ppm.

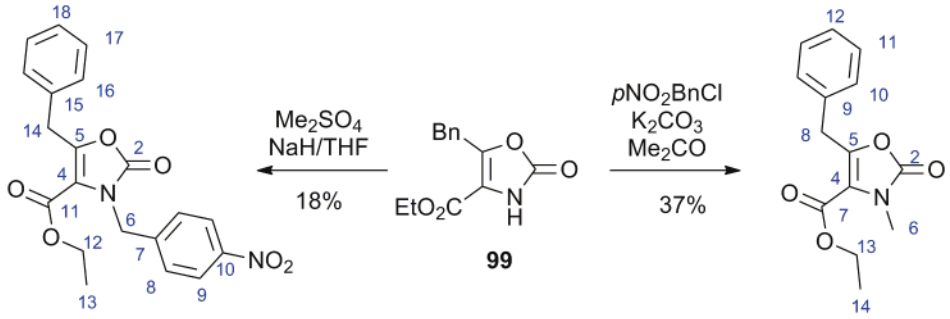


5-benzyl-4-chloromethyl-4-oxazolin-2-one 104

This compound was not isolated, although it was observed in crude products by a number of oxazolone chlorination methods ($SOCl_2$, CCl_4/PPh_3, MsCl with internal replacement of OMs, etc). One such procedure is given: to **101** (167 mg, 0.81 mmol) were added dry DCM (7 mL), dry pyridine

(130 mg, 1.68 mmol, 2 eq), then dropwise $SOCl_2$ (250 mg, 2.1 mmol, 2.6 eq). The insoluble starting material dissolves during the reaction. After the addition of 1.1 mmol $SOCl_2$ the reaction suddenly turns red (recall the starting material is difficult to dry, so this may represent 1:1 reaction). The reaction was stirred at RT for 2 h, then quenched into phosphate buffer (pH~12, 0.35 M), extracted into AcOEt, washed with water, brine, and dried on sodium sulfate, then concentrated to a black crude (165 mg, crystalline reflections) which was short-filtered on SiO_2 using AcOEt (2 cm in a pasteur pipette) to give a brown oil (160 mg). The TLC showed no starting material present, but no migrating spots either, even with 3:1 AcOEt:MeOH. DIMS did not reveal any expected peaks. Crude NMR was however promising and fairly clean, although most peaks were broad, and attributions were confirmed by HSQC; NMR spectra are reproduced on page 289 ff. Peak broadness might be due to intermolecular alkylation liberating acid, in the absence of base and during concentration. ^{1}H-NMR (500 MHz, $CDCl_3$): δ = 9.56 (s br, 1H, NH or possible OH cf. self-alkylation theory), 7.8-7.0 (m, 5H, overintegrates, benzyl), 4.4-4.0 (m, 2H, CH_2Cl), 3.75 (s, 2H, benzyl) ppm. ^{13}C-NMR (125 MHz, $CDCl_3$): δ = 156.5, 138.4, 135.4, 129.0 (×2), 128.8 (×2), 127.4, **118.5 (C4 upshifted rel. to starting material)**, **34.1 (CH_2Cl much upshifted)**, 31.2 ppm. However, chromatography did not yield any of the desired compound.

5.4.7. Initial trials: oxazolone and oxazolidinone O-alkylations



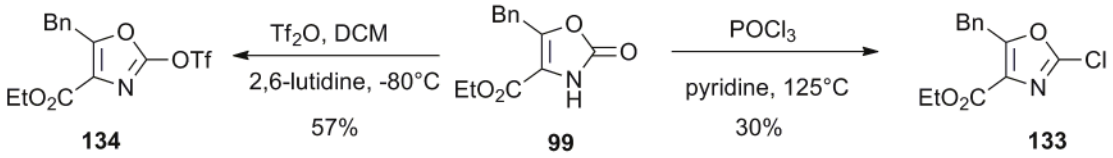
ethyl 5-benzyl-3-*para*-nitrobenzyl-2-oxo-2,3-dihydrooxazole-4-carboxylate (131)

A similar method as follows but using 3 times as much base, in MeCN, and adding 18-crown-6 (2 eq) appeared to give no isolable yield of alkylated products, but to my surprise gave chlorinated product **133** in 20% yield. To **99** (108 mg, 0.44 mmol) were added acetone (6 mL), K_2CO_3 (289 mg, 2 mmol) and *p*-NO_2BnCl (93 mg, 0.54 mmol) and the mixture stirred for 12 h. The crude was partitioned between H_2O and EtOAc, the aqueous phase extracted with EtOAc, the combined organic phases washed with water, brine, and dried on sodium sulfate, then concentrated and chromatographed carefully on 100:0→1:1 Cy:EA giving **131** as a clear oil (30 mg, 0.08 mmol, 18%) with R_f=0.50 on 5:2 Cy:EA (Haness). DIMS(+): 383 ($[MH]^+$); DIMS(-) 381 ($[M-H]^-$). Features confirming **N-alkylation** to a tertiary carbamate, not to the O-alkylated carbimidate, are given in bold. Salient IR: **1718[s] (carbamyl)**, 1774[s] (ester). ^{1}H-NMR (500 MHz, $CDCl_3$): δ = 8.18 (d, 8.7 Hz, 2H, 2×H9), 7.50 (d, 8.7 Hz, 2H, 2×H8), 7.30-7.20 (m, 5H, 2×H16, 2×H17, H18), **5.17 (s, 2H, *N*-alkyl=2×H6)**,

4.31 (q, 7.2 Hz, 2H, 2×H12), 4.13 (s, 2H, CH_2, 2×H8), 3.46 (s, 2H, 2×H14), 1.32 (t, 7.2 Hz, 3H, 3×H13) ppm.

ethyl 5-benzyl-3-methyl-2-oxo-2,3-dihydrooxazole-4-carboxylate (132)

A similar method using MeI gave no observable yield of O-alkylated product by TLC. To **99** (89 mg, 0.36 mmol) under Ar were added dry THF (6 mL), then NaH (60% in oil, 80 mg, 2 mmol). After 10 min, effervescence had ceased, and dimethyl sulfate (47 mg, 0.37 mmol) was added and the mixture stirred for 12 h. The crude was partitioned between H_2O and EtOAc, the aqueous phase extracted with EtOAc, the combined organic phases washed with water, brine, and dried on sodium sulfate, then concentrated and chromatographed carefully on 100:0:0→1:1:0→1:1:0.1→0:10:1 Cy:EA:EtOH, giving **132** as a clear oil (28 mg, 0.11 mmol, 30%) and recovered **99** (17 mg, 0.07 mmol, so yield brsm = 37%). **132** has R_f=0.48 on 7:5 Cy:EA (Haness), whereas **99** has R_f=0.28. DIMS(+): 262 ($[MH]^+$); DIMS(-) 260 ($[M-H]^-$). Features confirming **N-alkylation** to a tertiary carbamate, not to the O-alkylated carbimidate, are given in bold. Salient IR: **1722[s] (carbamyl)**, 1776[s] (ester). ^{1}H-NMR (500 MHz, $CDCl_3$): δ = 7.40-7.20 (m, 5H, 2×H10, 2×H11, H12), 4.43 (q, 7.3 Hz, 2H, 2×H13), 4.17 (s, 2H, CH_2, 2×H8), **3.46 (s, 3H, *N*-methyl=3×H6)**, 1.44 (t, 7.3 Hz, 3H, 3×H14) ppm. ^{13}C-NMR (125 MHz, $CDCl_3$): δ = **158.9 (C2)**, 153.6 (C7), 148.2 (C9), 135.5 (C5), 128.9+128.9 (2×C10 + 2×C11), 127.4 (C12), 114.4 (C4), 61.9 (C13), **32.5 (*N*-methyl=C6)**, 30.7 (C8), 14.4 (C14) ppm.



ethyl 5-benzyl-2-chlorooxazole-4-carboxylate (133)

A similar method using triethylamine seemed less efficient. A similar method applied to **98** resulted in extensive degradation. Similar methods applied to alcohols or silyl ethers **103**, **103** and **100** resulted in decomposition. Replacement with or tandem use of PCl_5 in benzene[54] with lutidine base did not give isolated products either. Based on the method of Maekawa[55], to **99** (247 mg, 1.00 mmol) were added under Ar, dry pyridine (100 mg, 1.3 mmol) then $POCl_3$ (850 mg, 5.4 mmol) at RT, then the mixture heated to 125°C for 1.5 h. The crude was quenched into phosphate buffer (pH~12, 0.35 M), extracted into AcOEt, and the combined organic layers washed with water, brine, and dried on sodium sulfate, filtered then concentrated and chromatographed on 10:1→1:1 Cy:EA giving **133** as a pale yellow oil (85 mg, 0.32 mmol, 30%) and starting material (0.13 mmol, so yield brsm = 37%). R_f=0.54 on 5:2 Cy:EA (UV). ^{1}H-NMR (500 MHz, $CDCl_3$): δ = 7.36-7.27 (m, 5H, benzyl), 4.43 (q, 7.3 Hz, 2H, ethyl),

4.39 (s, 2H, benzyl), 1.43 (t, 7.3 Hz, 3H, ethyl) ppm. ^{13}C-NMR (125 MHz, $CDCl_3$): δ = 161.1, **159.7**, 146.0, 135.4, 129.2, 128.9 (×2), 128.8 (×2), 127.4, 61.5, 32.1, 14.3 ppm.

ethyl 5-benzyl-2-triflyloxazolone-4-carboxylate (134)

The experimental procedure is adopted from Flegeau *et al.*[56] for a different substrate. To **99** (76 mg, 0.71 mmol) was added to a dry flask with stirrer, and the atmosphere replaced with argon. Dry DCM (5 mL) and 2,6-lutidine dried for 40 h on 3A MS (211 mg, 2 mmol, 2.7 eq) were added, and the reaction cooled to -80°C. Triflic anhydride (300 mg, 1.06 mmol, 1.5 eq) was added dropwise and the reaction stirred for one hour before quenching with phosphate buffer (pH~6.8, 0.35 M). The aqueous layer was extracted with DCM, and the combined organic layers washed with brine, dried on sodium sulfate, filtered then concentrated *in the cold* to a pale red oil (287 mg). The fastest-running spot was recovered from a short, fast column on 100:0→3:1 Cy:EA gradient, as TLC shows partial decomposition on SiO_2; R_f=0.61 on 5:2 Cy:EA (UV). **134** was obtained as a colourless solid (155 mg, 0.40 mmol, 57%). ^{1}H-NMR (500 MHz, $CDCl_3$): δ = 7.35-7.25 (m, 5H, benzyl), 4.42 (q, 7.1 Hz, 2H, ethyl), 4.38 (s, 2H, benzyl), 1.41 (t, 7.0 Hz, 3H, ethyl) ppm. ^{13}C-NMR (125 MHz, $CDCl_3$): δ = 160.8, **157.5**, 148.4, 134.9, 129.1 (×2), 128.8 (×2), 128.1, 127.6, 118.5 (q, 322 Hz, CF_3), 61.5, 32.1, 14.3 ppm.

Trial S_NAr with triflate 134

To NaH (74 mg, 60% in oil, 1.85 mmol) under Ar were added dry dioxane (2.5 mL), BnOH (206 mg, 1.94 mmol), and a solution of **134** (155 mg, 0.40 mmol) in dioxane (3 mL). The reaction was stirred at RT for 2.5 h, then quenched into a mixture of DCM (5 mL) and KH_2PO_4 (0.35 M, pH~3.5, 3 mL). The organic layer was washed with water and brine, dried on sodium sulfate, filtered, concentrated and chromatographed with 100:0:0→1:1:0→0:15:1 Cy:EA:MeOH, giving dibenzyl ether (21 mg, 0.11 mmol), and a 1:0.7 mol mixture of ethyl ester **99** (R_f=0.5) and the corresponding benzyl ester **99Bn** (R_f=0.6; 1:1 Cy:EA). **99Bn**: ^{1}H-NMR (200 MHz, $CDCl_3$): δ = 9.1 (s br, 1H), 7.45-7.15 (m, 10H, benzyls), 5.33 (s, 2H), 4.12 (s, 2H) ppm. ^{13}C-NMR Salient Peaks (overlapped aromatics) (50 MHz, $CDCl_3$): δ = 158.7, 154.1, 148.6, 135.5, 134.8, 128.9 (×4) 128.7 (×2), 127.4, 114.4, 67.0, 32.1 ppm.

5.4.8. Substrate-spacer pairs

Phenylacetyl-*para*-aminobenzyl bromide (147)

Phenylacetyl-*para*-aminobenzyl alcohol **146** was synthesised as per Nunez *et al.*[57]. Similarly to Senter *et al.*[58], to **146** (135 mg, 0.56 mmol) under Ar was added dry DCM (10 mL), and to this suspension was added dropwise HBr (48% in AcOH, 4 mL). The solid dissolved after the addition of 0.5 mL of HBr solution. The solution was stirred with a large-bore needle outlet to the Ar line for venting HBr. After two hours the solution was evaporated trap-to-trap to give an approximately

quantitative yield of unreported bromide **147** which was used directly. ^{1}H-NMR (500 MHz, $CDCl_3$): δ = 7.48-7.30 (m, 9H), 7.13 (s br, 1H), 4.50 (s, 2H), 3.80 (s, 2H) ppm. ^{13}C-NMR (125 MHz, $CDCl_3$): δ = 169.2, 137.8, 134.3, 133.9, 129.9 (×2), 129.7 (×2), 129.5 (×2), 128.0, 120.1 (×2), 45.1, 33.4 ppm.

Boc-leucyl-*para*-aminobenzyl bromide (150)

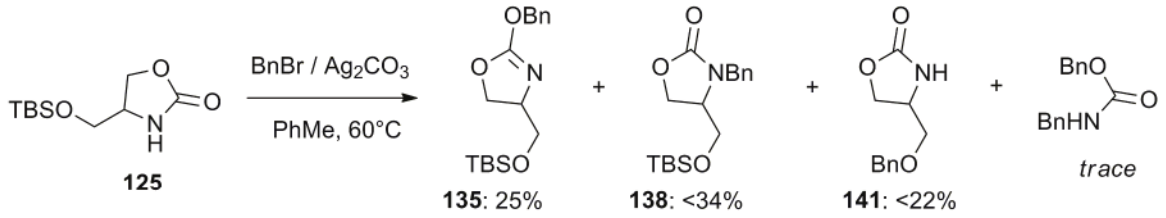
N-hydroxysuccinimide ester **148** was synthesised by the known method[59] then reacted with PABA (153 mg, 1.24 mmol) in $CHCl_3$ (5 mL) at RT for 2 hours, as per the protocol of Maekawa *et al.*[60]. After evaporation of the volatiles, rinsing with water and cyclohexane, the residue was chromatographed on 5:1→1:1 Cy:EA revealing with $KMnO_4$ to give alcohol **149** as a pale yellow oil (150 mg, 0.42 mmol, 44% for two steps; R_f = 0.51 on 2.5:1 Cy:EA). ^{1}H-NMR (200 MHz, $CDCl_3$): δ = 9.35 (s br, 1H), 7.35-7.20 (d, 8.4 Hz, 2H), 7.00-6.85 (d, 8.0 Hz, 2H), 5.70 (d, 8.2 Hz, 1H), 4.40 (s, 2H), 4.40-4.30 (m, 1H), 3.80 (s br, 1H), 1.80-1.40 (m, 3H), 1.30 (s, 9H), 0.85 (~t, 5.1 Hz, 6H) ppm. To this **149** (150 mg, 0.42 mmol) were added CBr_4 (265 mg, 0.73 mmol), DCM (6 mL), and 3A molecular sieves to dry residual water for 5 min before the solution was cooled to 0°C. PPh_3 (246 mg, 0.93 mmol) was dissolved in DCM (1.5 mL) and added to the solution, which yellows. After 1 hour stirring at RT, the solution was filtered and evaporated; crude NMR showed an approx. 1:1 mixture of alcohol (CH_2OH @ 4.46 ppm) and (CH_2Br @ 4.30 ppm); this was separated on a quick column as the bromide hydrolyses on silica. Chromatography on 10:1→2.5:1 Cy:EA gave **150** as a colourless oil (39 mg, 0.09 mmol, 22% ==27% brsm; R_f = 0.63 on 2.5:1 Cy:EA) and recovered **149** (24 mg, 0.07 mmol).

para-(tetraaceto-β-galactosyloxy)benzyl bromide (153)*

β-galactoside **152** was synthesised in 50% yield as follows. To galactose pentaaceate (12 g, 31 mmol) was added *p*-cresol (11.5 g, 107 mmol, 3.5 eq) which had been dried in solution on molecular sieves, and DCM (250 mL). The solution was cooled to 0°C and $BF_3{\cdot}OEt_2$ (17.2 g, 122 mmol) was added and the solution was stirred warming to RT overnight, quenched into KOH/Na_2CO_3 (better method: add carbonate solution), then DCM evaporated from the biphasic mixture, Et_2O added, and the organic layer washed repeatedly with KOH solution till the aqueous layers were not coloured. The organic layer was washed with brine, dried on sodium sulfate, filtered, concentrated, and chromatographed to give **152** as a colourless powder (6.9 g, 15.7 mmol, 50%). This was dissolved into $CHCl_3$ and dried on molecular sieves for 24 h, then NBS (2.87 g, 16.1 mmol, 1.15 eq) and AIBN (230 mg) were added and the solution heated to 60°C. After 1.5 h and 3 h, two further portions of AIBN (230 mg) were added, then all the solids had dissolved, and the reaction had attained the most favourable distribution (20:3:1 product:dibromo:starting material). The reaction was cooled, and colourless crystals were filtered off, then the solution concentrated and chromatographed with 10:1→2.5:1 Cy:EA revealing with Van to give known **153** as a colourless, hygroscopic foam which was triturated with pentane to give a brittle

solid (7.13 g, 11.7 mmol, 82%; R_f = 0.33 on 5:2 Cy:EA) containing trace residual dibromo (R_f = 0.48 on 5:2 Cy:EA) and nonbromo (R_f = 0.21 on 5:2 Cy:EA) compunds.

5.4.9. Final methods: O-alkylations of oxazolones and oxazolidinones

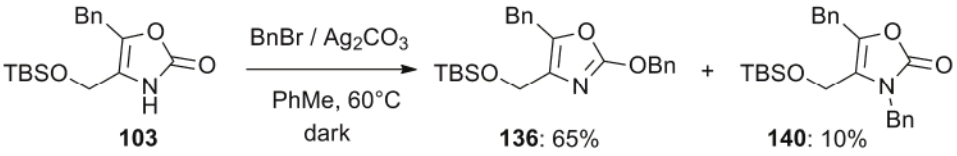


2-(benzyloxy)-4-(((tert-butyldimethylsilyl)oxy)methyl)-4,5-dihydrooxazole (135)

In one example, adapted from the method of Seki and Matsumoto,[61] **125** (818 mg, 3.54 mmol) was reacted, capped, under air, with BnBr (735 mg, 4.30 mmol) and Ag_2CO_3 (940 mg, 3.41 mmol) in dry PhMe (10 mL) at 60°C for 14 h in the dark, obtaining a green-yellow solid under a clear solution. This solution was loaded as is onto a silica column, the solid washed with AcOEt and that solution added, then chromatographed carefully on 100:0→1:1 peth:EA collecting all spots. On TLC with 7:2 Cy:EA revealing with Van these are sequentially, dibenzyl carbamate (R_f = 0.68, trace, chars red), **135** (R_f = 0.41, chars grey-green), **138+141** (R_f = 0.32 inseparable, chars grey), and **125** (R_f = 0.03); none of the products were easily oxidised by Van, but $KMnO_4$/UV could be used satisfactorily. **135** was recovered as a colourless oil (284 mg, 0.88 mmol, 25%) and assigned the O-benzylated structure on the basis of the following bolded NMR shifts (NMR spectra are reproduced on page 291 ff.): ^{1}H-NMR (500 MHz, $CDCl_3$): δ = 7.44-7.35 (m, 5H), **5.27 (s, 2H, OBn)**, 4.49-4.40 (m, 2H), 4.21-4.16 (m, 1H), 3.81 (dd, 10.5&3.5 Hz, 1H), 3.56 (dd, 10.2&6.8 Hz, 1H), 0.94 (s, 9H), 0.11 (s+s, 2×3H) ppm. ^{13}C-NMR (125 MHz, $CDCl_3$): δ = **163.8 (C2)**, 135.4, 128.6 (×2), 128.5, 128.1 (×2), **72.4 (OBn)**, 71.6, 65.5, 64.9, 25.9 (×3), 18.4, -5.2 (×2) ppm. Benzyl carbamate was attributed based on ^{1}H-NMR (200 MHz, $CDCl_3$): δ = 7.41-7.28 (m, 10H), 5.18 (s, 2H), 4.57 (s, 2H) ppm. **125** could be recovered in up to 15% yield.

3-benzyl-4-(((tert-butyldimethylsilyl)oxy)methyl)oxazolidin-2-one (138);
4-((benzyloxy)methyl)oxazolidin-2-one (141)

138+141 was obtained as an inseparable mixture during one such procedure for **135** as a colourless oil; in the following procedures the mixture was discarded. By NMR the yields were calculated as **138**: 34% brsm; **141**: 22% brsm. **138** was assigned the N-benzylated structure on the basis of the following bolded NMR shift, with the other attributions performed as best as possible: ^{1}H-NMR (500 MHz, $CDCl_3$): δ = 7.42-7.27 (m, 5H), **4.71 (s, 2H, NBn)**, 4.21-4.05 (m, 2H), 3.73-3.55 (m, 3H), 0.93 (s, 9H), 0.10 (s, 6H); **141** was assigned its structure based on the following attributions: ^{1}H-NMR (500 MHz, $CDCl_3$): δ = 7.42-7.27 (m, 5H), **4.83 (d, 15 Hz, 2H, CH_2OBn)**, 4.31 (~t, 18 Hz, 2H), 4.21-4.10 (md, 15 Hz, 1H), 3.85-3.75 (m, 1H), 3.73-3.65 (m, 1H).

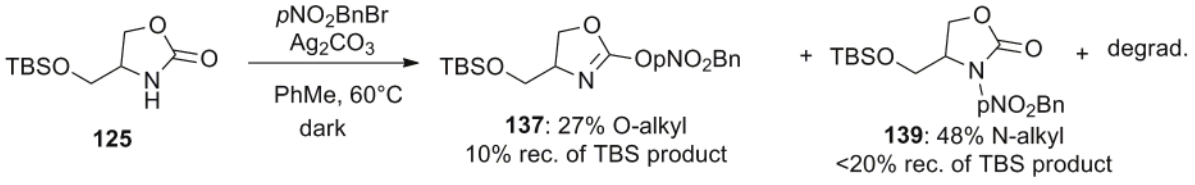


5-benzyl-2-(benzyloxy)-4-(((tert-butyldimethylsilyl)oxy)methyl)oxazole (136)

Similarly to the procedure for **135**, **103** (281 mg, 0.88 mmol) was reacted with BnBr (222 mg, 1.30 mmol) and Ag_2CO_3 (254 mg, 0.92 mmol) in dry PhMe (4 mL). TLC (Cy:EA eluents) was easily revealed with vanillin; main products were **136** (R_f = 0.39 on 10:1, R_f = 0.74 on 5:2, chars red) and **140** (R_f = 0.12 on 10:1, chars red). After reaction, SiO_2 was added and the toluene was carefully evaporated. Column with a solid deposition of the total contents gave **136** as a colourless oil (236 mg, 0.57 mmol, 65%), which was assigned the O-benzylated structure on the basis of the following bolded NMR shifts (NMR spectra are reproduced on page 293 ff.): ^{1}H-NMR (200 MHz, $CDCl_3$): δ = 7.45-7.18 (m, 10H), **5.36 (s, 2H, O-benzyl)**, 4.62 (s, 2H), 3.98 (s, 2H), 0.96 (s, 9H), 0.16 (s, 6H) ppm. ^{13}C-NMR (125 MHz, $CDCl_3$): δ = **160.5**, **141.8**, 137.8, 135.1, **133.5**, 128.7+128.7 (×6), 128.4 (×2), 126.7, **72.6 (O-benzyl)**, 58.5, 31.2, 26.1 (×3), 18.6, -5.0 (×2) ppm.

3,5-dibenzyl-4-(((tert-butyldimethylsilyl)oxy)methyl)oxazol-2(3H)-one (140)

140 was obtained during the procedure for **136** as a pale yellow oil (37 mg, 0.09 mmol, 10%). DIMS(+): 410 (MH^+). It was assigned the N-benzylated structure on the basis of the following bolded NMR shifts: ^{1}H-NMR (500 MHz, $CDCl_3$): δ = 7.45-7.20 (m, 10H), **4.93 (s, 2H, N-benzyl)**, 4.28 (s, 2H), 3.79 (s, 2H), 0.94 (s, 9H), 0.07 (s, 6H) ppm. ^{13}C-NMR (125 MHz, $CDCl_3$): δ = **155.9**, **136.7**, 136.3, 135.8, 128.9 (×2), 128.8 (×2), 128.6 (×2), 128.0, 127.3 (×2), 127.1, **122.0**, 53.8, **45.7 (N-benzyl)**, 31.0, 25.9 (×3), 18.3, -5.3 (×2) ppm.



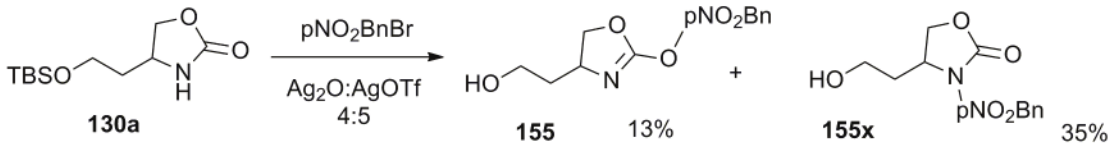
2-(3-*para*nitrobenzyloxy)-4-(((tert-butyldimethylsilyl)oxy)methyl)-4,5-dihydrooxazole (137)

Similarly to the procedure for **135**, **125** (501 mg, 2.17 mmol) was reacted with *p*NO_2BnBr (606 mg, 2.80 mmol) and Ag_2CO_3 (632 mg, 2.29 mmol) in dry PhMe (7 mL). TLC (1:1 Cy:EA) was only well-revealed by UV. Major products were **137** (R_f = 0.42) and **139** (R_f = 0.30) as well as significant isolated amounts of di(*para*nitrobenzyl) ether (12 mg, 0.04 mmol, 2% isol) and *para*nitrobenzaldehyde (18 mg, 0.12 mmol, 5% isol), starting material, and significant non-isolated

amounts of the three major desilylated products. **137** was obtained as a colourless oil (80 mg, 0.22 mmol, 10%), which was assigned the O-benzylated structure on the basis of the following bolded NMR shift: ^{1}H-NMR (500 MHz, $CDCl_3$): δ = 8.25 (d, 8.8 Hz, 2H), 7.58 (d, 8.6 Hz, 2H), **5.36 (s, 2H, OBn)**, 4.49 (~t, 8.6 Hz, 1H), 4.43 (dd, 8.3&6.5 Hz, 1H), 4.20-4.15 (m, 1H), 3.77 (dd, 9.8&3.7 Hz, 1H), 3.60 (dd, 9.8&5.9 Hz, 1H), 0.90 (s, 9H), 0.09+0.08 (s+s, 2×3H) ppm. ^{13}C-NMR (125 MHz, $CDCl_3$): δ = **163.4**, 147.9, 142.7, 128.1 (×2), 123.9 (×2), 71.8 (OBn), 70.6, 65.2, 64.9, 25.9 (×3), -5.2, -5.2 ppm. Di(*para*nitrobenzyl) ether was assigned based on ^{1}H-NMR (200 MHz, $CDCl_3$): δ = 8.21 (d, 8.8 Hz, 2H), 7.55 (d, 8.6 Hz, 2H), 4.72 (s, 2H); *para*-nitrobenzaldehyde was assigned based on ^{1}H-NMR (200 MHz, $CDCl_3$): δ = 10.16 (s, 1H), 8.40 (d, 8.7 Hz, 2H), 8.05 (d, 8.5 Hz, 2H).

3-*para*nitrobenzyl-4-(((tert-butyldimethylsilyl)oxy)methyl)oxazolidin-2-one (139)

139 was obtained during the procedure for **137** as a pale yellow oil (350 mg with inseperable contaminants, calc. 0.5 mmol, 25%). It was assigned the N-benzylated structure on the basis of the following bolded NMR shift: ^{1}H-NMR (200 MHz, $CDCl_3$): δ = **4.81 (s, 2H, NBn)** ppm; other ^{1}H-NMR features overlapped with standard shifts for similar structures (see **137**) in the contaminants.



2-(2-((4-nitrobenzyl)oxy)-4,5-dihydrooxazol-4-yl)ethanol (155)

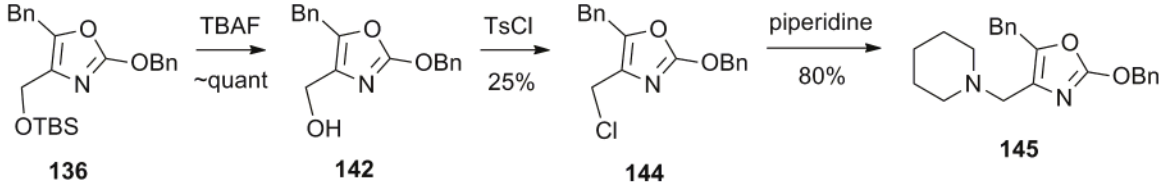
In a representative procedure, **130a** (1.52 mmol) was reacted with *p*NO_2BnBr (847 mg, 3.90 mmol) in the presence of AgOTf (516 mg, 2.02 mmol) and Ag_2O (423 mg, 1.84 mmol) in toluene (2 mL) at 65°C for 6h in the dark. A clear, pale solution over a grey solid was formed. Solid deposition chromatography as for **135** yields **155** (55 mg, 0.20 mmol, 13%; R_f = 0.62 on 1:2:0.02 Cy:EA:MeOH, grey with Van) before N-alkylated byproduct **155x**, as well as di-*p*-nitrobenzyl ether (72 mg, 0.25 mmol). DIMS(+): 267 (MH^+), DIMS(-): 265 $(M-H)^-$. NMR spectra are reproduced on page269 297 ff. **^{1}H-NMR** (500 MHz, $CDCl_3$): δ = 8.27-8.20 (m, 2H), 7.55-7.45 (m, 2H), **5.22 (s, 2H, OBn*p*NO_2)**, 4.05-3.92 (m, 1H), 3.87-3.80 (m, 2H), 3.75-3.61 (m, 2H), 2.35-2.25 (m, 1H), 1.91-1.83 (m, 1H) ppm. **^{13}C-NMR** (125 MHz, $CDCl_3$): δ = 155.4, 147.7, 144.0, 128.3 (×2), 123.9 (×2), 73.6 (O-*p*-NO_2Bn), 66.9, 65.3, 52.1, 33.3 ppm.

4-(2-hydroxyethyl)-3-(4-nitrobenzyl)oxazolidin-2-one (155x)

N-alkylated **155x** was isolated as a colourless oil (150 mg, 35%) during the procedure for O-alkylated **155**. TLC: R_f = 0.40 (1:2:0.02 Cy:EA:MeOH, grey with Van). DIMS(+): 267 (MH^+). ^{1}H-NMR (200 MHz, CDCN: product insoluble in $CDCl_3$): δ = 8.18 (d, 8.7 Hz, 2H), 7.55 (d, 8.3 Hz, 2H), 6.44

(s br, 1H), **4.61 (s, 2H, NBn*p*NO$_2$)**, 4.53 (t, 8.3 Hz, 1H), 4.20-4.05 (m, 2H), 3.65-3.55 (m, 2H), 1.88 (td~q, 6.1 Hz, 2H) ppm. ^{13}C-NMR (50 MHz, CDCN): δ = 160.4, 147.4, 146.6, 128.0 (×2), 123.6 (×2), 71.4, 70.9, 67.4, 51.1, 34.7ppm.

5.4.10. O-alkylated oxazolone/oxazolidinone electrophiles and grafting



136 **142** **144** **145**

(5-benzyl-2-(benzyloxy)oxazol-4-yl)methanol (142)

Following the TBAF protocol of Kaburagi[62], to **136** (233 mg, 0.57 mmol) was added TBAF (1M in THF, 2.5 mL) and the orange-red solution stirred under air for 4 h. $CaCO_3$ (580 mg) was added, then Dowex resin (50X8-100 acid form, 1.72 g), then the mixture filtered through a short, wide layer of sand/Celite rinsing with dry MeOH (2×2.5 mL); the solution was evaporated to give a quantitative yield of **142** (168 mg, 0.57 mmol) as a pale orange oil containing some (inoffensive) residual TBA^+ salts (76 mg). **^{1}H-NMR** (200 MHz, $CDCl_3$): δ = 7.35-7.15 (m, 10H), 5.27 (s, 2H), 4.46 (s, 2H), 3.88 (s, 2H) ppm. **^{13}C-NMR** (125 MHz, MeOD): δ = 160.7, 142.6, 137.5, 133.5, 128.5 (×2), 128.4 (×2), 128.3 (×2), 128.1 (×2), 126.5, 72.6, **55.3 (CH_2OH)**, 30.2 ppm. DIMS(+): 296 (MH^+).

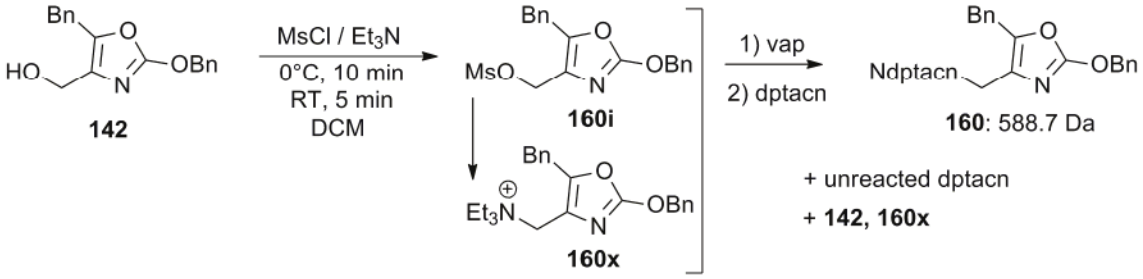
(5-benzyl-2-(benzyloxy)oxazol-4-yl)chloromethane (144)

To **142** (70 mg, 0.24 mmol) were added at 0°C dry DCM (10 mL), N-butylimidazole (101 mg, 0.81 mmol), and TsCl (74 mg, 0.39 mmol). The orange solution of starting material immediately turns pale yellow upon addition of the TsCl. The reaction was stirred warming to RT for 12 h; DIMS(-) indicated tosylate **OTS254B** (448 Th, $[M-H]^-$) and DIMS(+) indicated bis-tosylation (655 Th). Workup as for **126** followed by column chromatography on 100:0→0:1 Cy:EA followed by MeOH/NEt_3/AcOEt rinsing returned **144** as a colourless oil (19 mg, 0.06 mmol, 25% == 30% brsm; R_f = 0.57 on 5:2 Cy:EA, chars red with Van) and unreacted **142** (12 mg, 0.04 mmol, 16%). DIMS(+): 314@100, 316@25; MH^+. **^{1}H-NMR** (500 MHz, $CDCl_3$): δ = 7.48-7.25 (m, 10H), 5.42 (s, 2H), 4.45 (s, 2H), 3.97 (s, 2H) ppm. **^{13}C-NMR** (125 MHz, MeOD): δ = 160.7, 143.3, 136.4, 134.8, 131.0, 128.9 (×2), 128.9, 128.8 (×2), 128.7 (×2), 128.4 (×2), 127.2, 73.0, **37.5 (CH_2Cl)**, 31.2 ppm.

(5-benzyl-2-(benzyloxy)oxazol-4-yl)-(1-piperidyl)methane (145)

To **144** (19 mg, 0.06 mmol) were added dry EtOH (10 mL) and piperidine (242 mg, 3 mmol) and the mixture stirred at RT overnight. The volatiles were removed and the residue chromatographed on 100:0→1:1 peth:EA to remove contaminants, then the eluant changed to 45:45:5:2→40:40:10:10

peth:EA:MeOH:NEt_3 yielding pure **145** as a colourless oil (15 mg, 0.05 mmol, 80%; R_f = 0.45 on 1:1:0.2 Cy:EA:MeOH, chars brown with Van). DIMS(+): 363.2 (MH^+). NMR spectra are reproduced on page 295 ff. **^{1}H-NMR** (500 MHz, $CDCl_3$): δ = 7.45-7.15 (m, 10H), 5.40 (s, 2H), 3.96 (s, 2H), **3.37 (s, 2H, oxazolo-CH_2-piperidyl)**, 2.47 (~t, 5.4 Hz, 4H), 1.65-1.61 (m, 4H), 1.49-1.46 (m, 2H) ppm. **^{13}C-NMR** (125 MHz, $CDCl_3$): δ = 160.4, 142.7, 137.8, 135.4, 131.4, 128.6 (×2), 128.5 (×2), 128.5 (×2), 128.2 (×2), 126.6, 72.7, 54.6 (×2), 54.5, **46.4 (oxazolo-CH_2-piperidyl)**, 31.0, 25.9 (×2), 24.4 ppm.



1-(5-benzyl-2-(benzyloxy)oxazol-4-yl)methyl-4,7-dipicolyl-1,4,7-triazacyclononane (160)

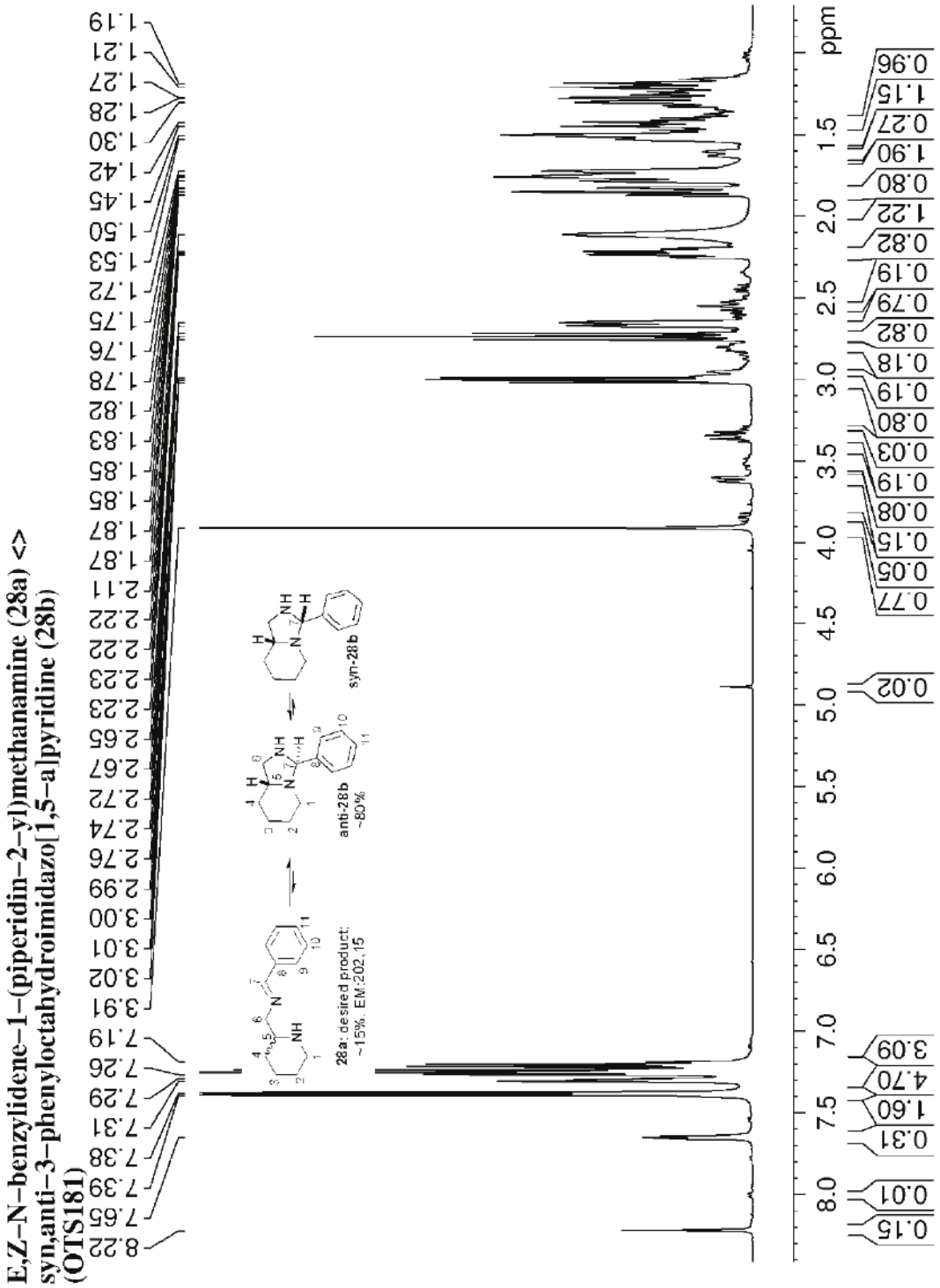
Note: a better procedure would use DBU instead of Et_3N. 1,4-dipicolyl-1,4,7-triazacyclononane (dptacn, aged and blackened, 112 mg, ≤0.38 mmol) was stirred with TFA (4 mL) for 1 h at RT to decompose carbamic acids, which are often cited as a storage-degradation product, before high-vacuum evaporation of the volatiles and trituration with peth. To **142** (168 mg, 0.57 mmol) were added at 0°C dry DCM (4 mL), KOH-dried Et_3N (0.3 mL), and then in one shot, a solution of MsCl (88 mg, 0.77 mmol) in DCM (1 mL). The solution was stirred for 10 min at 0°C, 5 min at RT, then evaporated. To the dptacn was added dry Et_3N (0.3 mL), dry EtOH (8 mL), and KI (30 mg); then the mesylation residue was quickly washed in with EtOH (2×2 mL) and the solution stirred overnight. Alumina TLC (5:5:1 EA:Cy:MeOH) shows dptacn (immobile), but also the formation of a faster-running yellow spot (R_f = 0.12). DIMS(+): 296 (**142**H^+), 379 (**160x**$^+$), 589 (**160**H^+), 611 (**160**Na^+). The volatiles were evaporated and a short column on basic alumina run with 100:0→90:16 $CHCl_3$:MeOH. The pure-$CHCl_3$ fractions recover regenerated/unreacted **142** (64 mg, 0.22 mmol, 38%). Polarity 90:10 eluant recovered a fraction of 84 mg of green-brown oil containing an estimated ~2:3 mixture of dptacn and **160**. **Salient ^{1}H-NMR** (500 MHz, $CDCl_3$): δ = 5.30 (s, 2H, O-benzyl), 4.50 (s, 2H), 4.21 (s, 2H), 3.83 (s, 4H), 3.15-3.10 (m, 4H, dptacn), 2.98-2.82 (m, 8H, dptacn). Repurification will be performed once experimentation restarts.

Key NMR Spectra

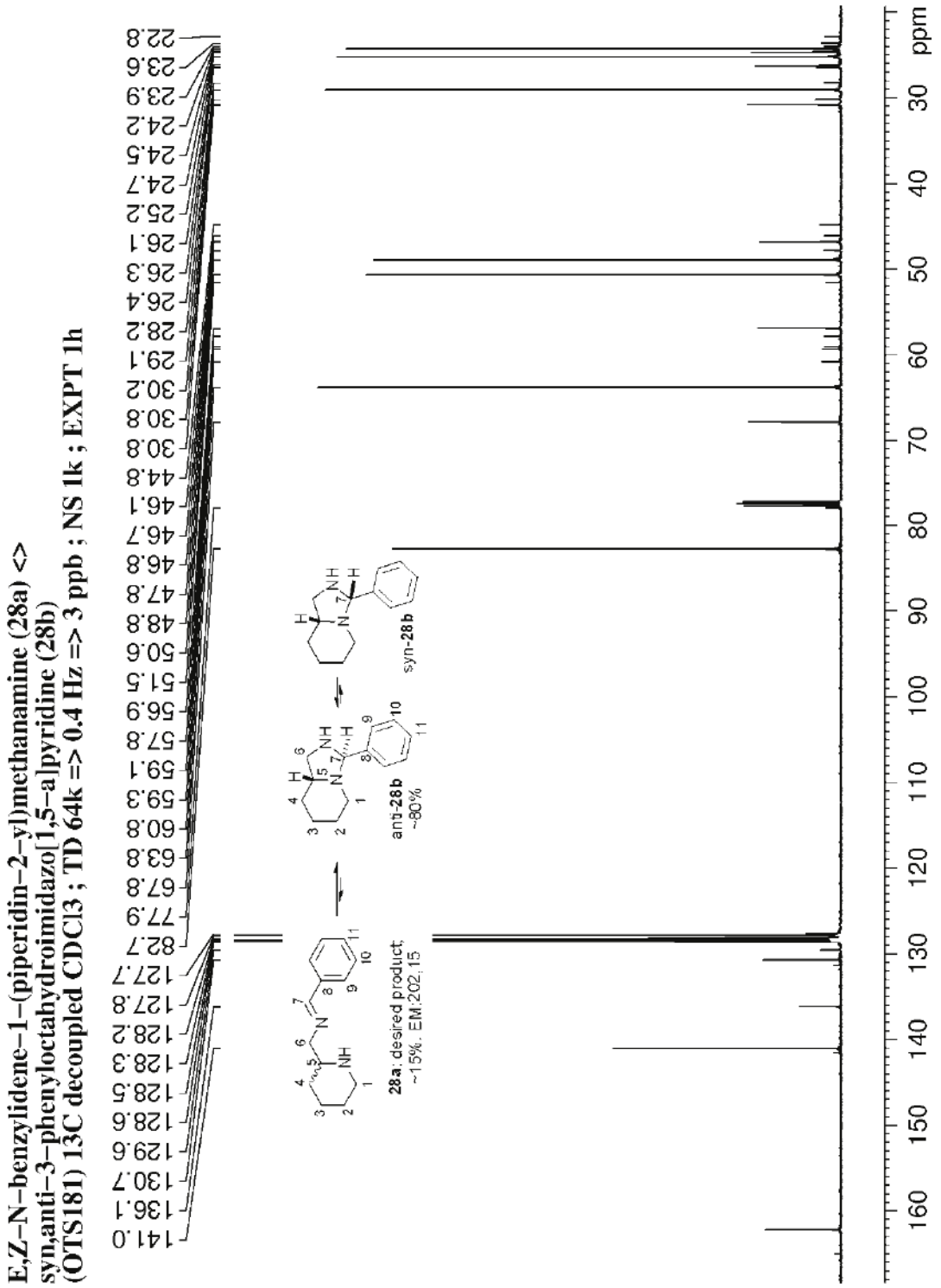
Key spectra are given in the order in which the molecules were presented in the text.

Spacer protecting group strategy

Equilibrating mixture **syn/anti-28b↔E/Z-28a**: ^{1}H-NMR spectrum (500 MHz, $CDCl_3$)

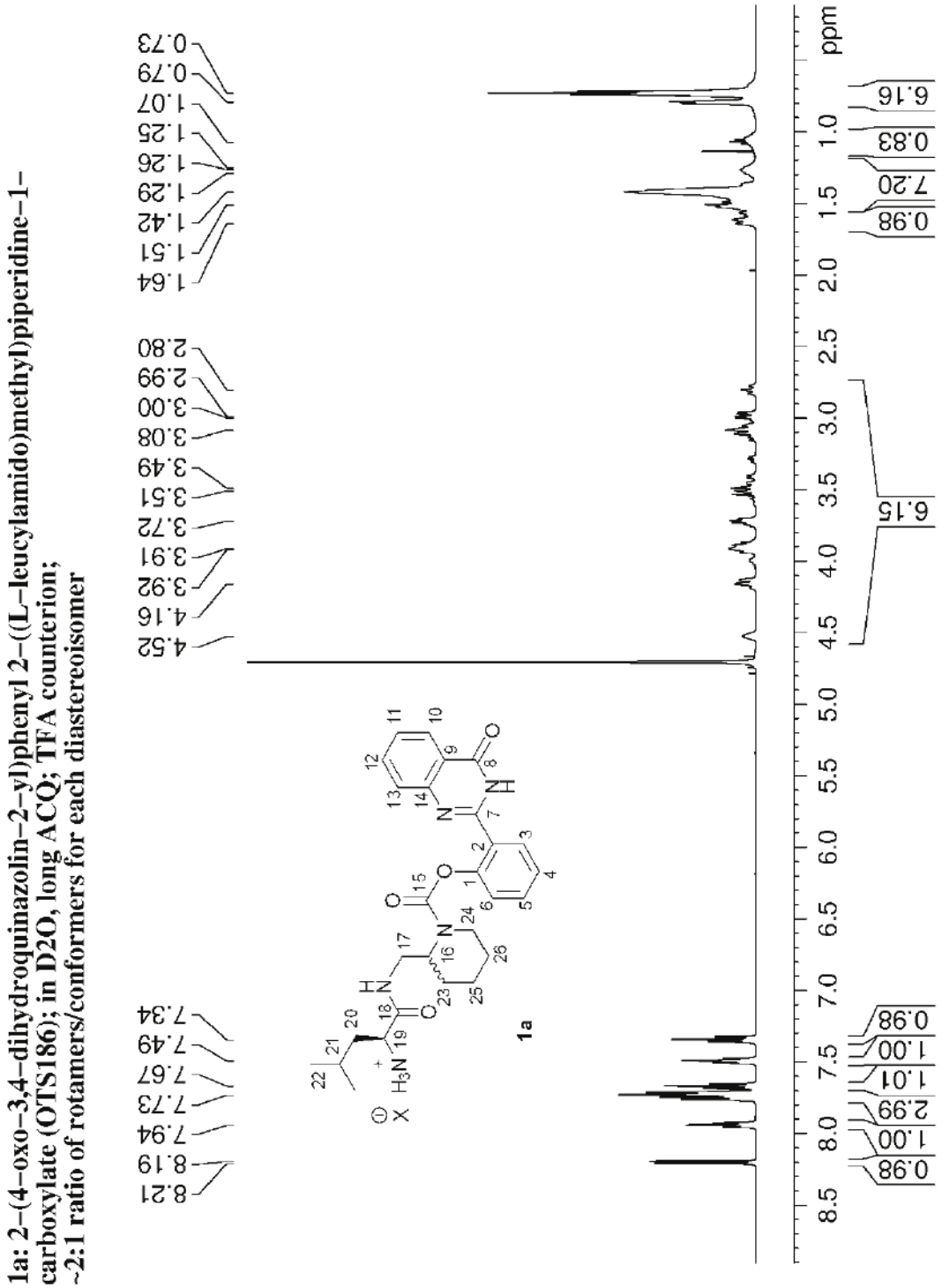


Equilibrating mixture **syn/anti-28b↔E/Z-28a**: ^{13}C-NMR spectrum (125 MHz, $CDCl_3$)

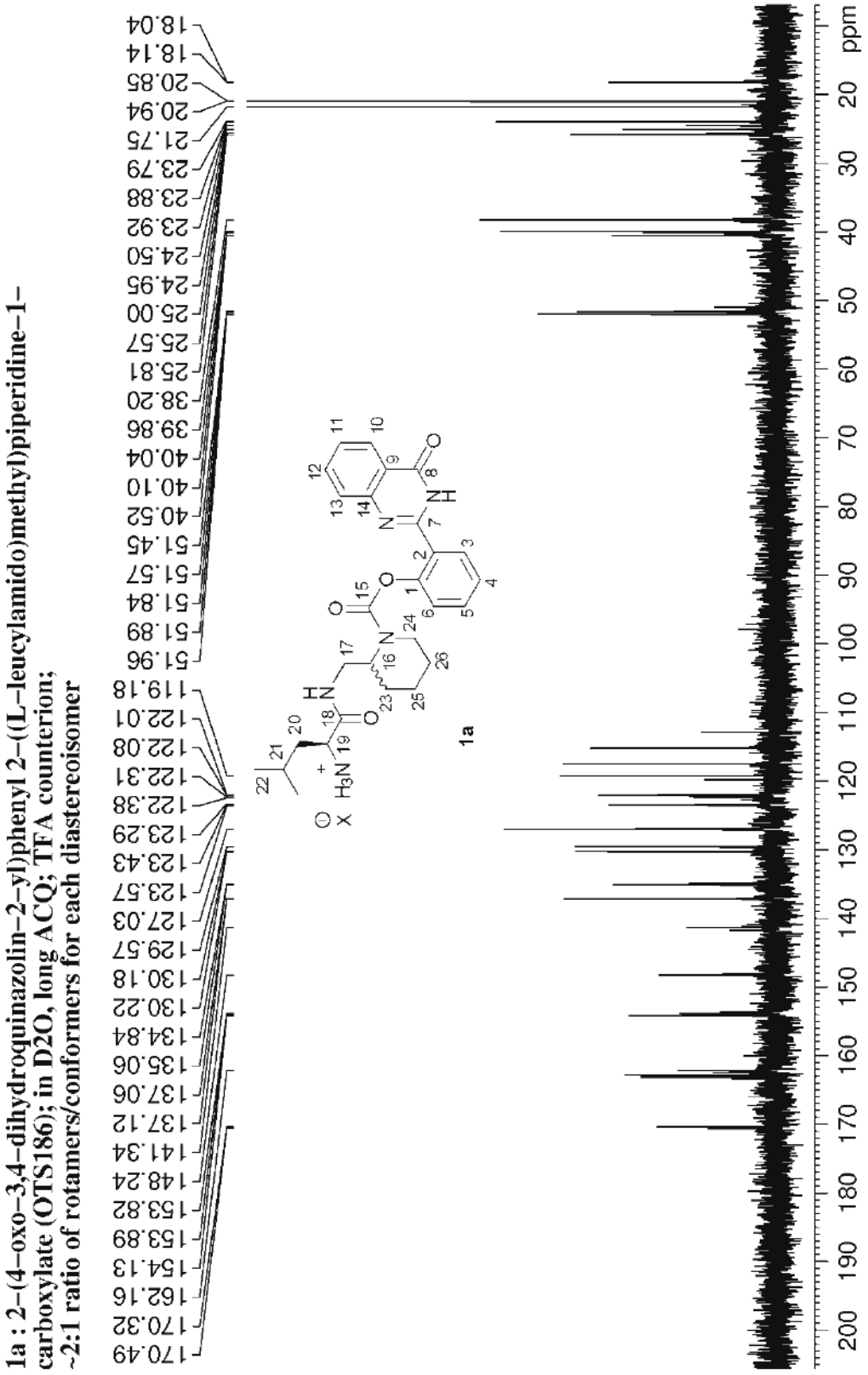


Fluorogenic probes tested in vitro

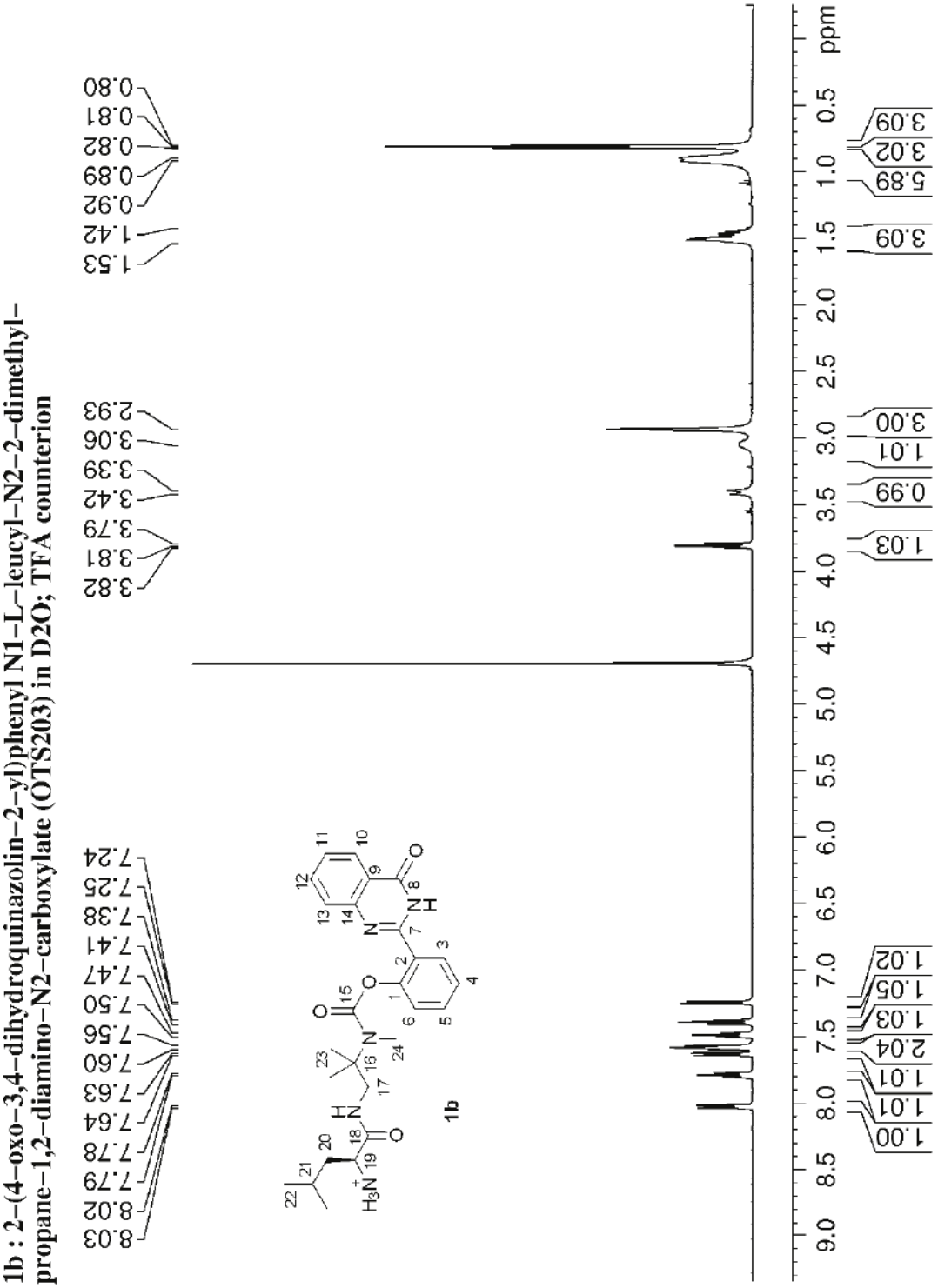
1a: ^{1}H-NMR spectrum (500 MHz, D_2O, TFA counterion)



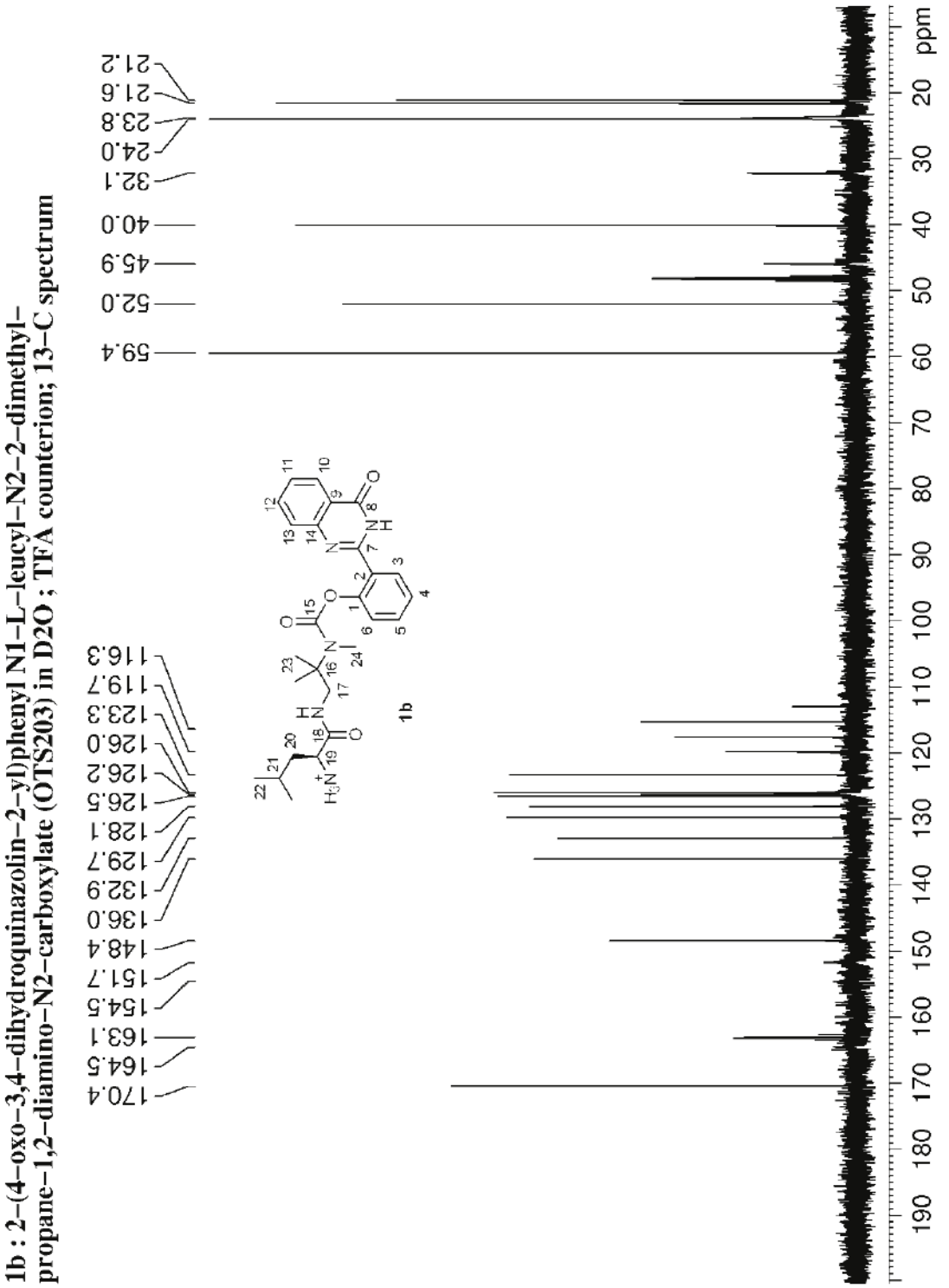
1a: ^{13}C-NMR spectrum (125 MHz, D_2O, TFA counterion)



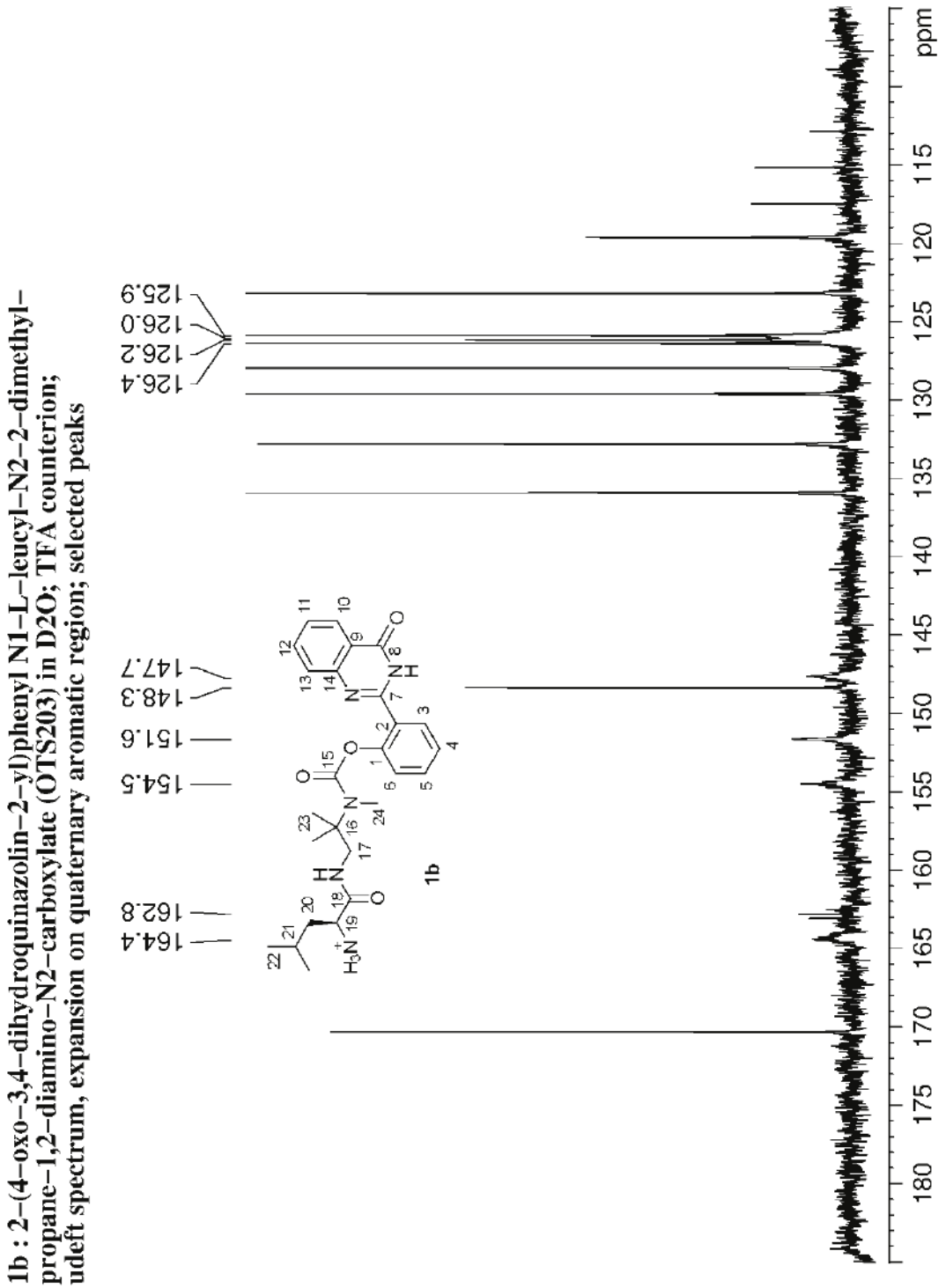
1b: ^{1}H-NMR spectrum (500 MHz, D_2O, TFA counterion)



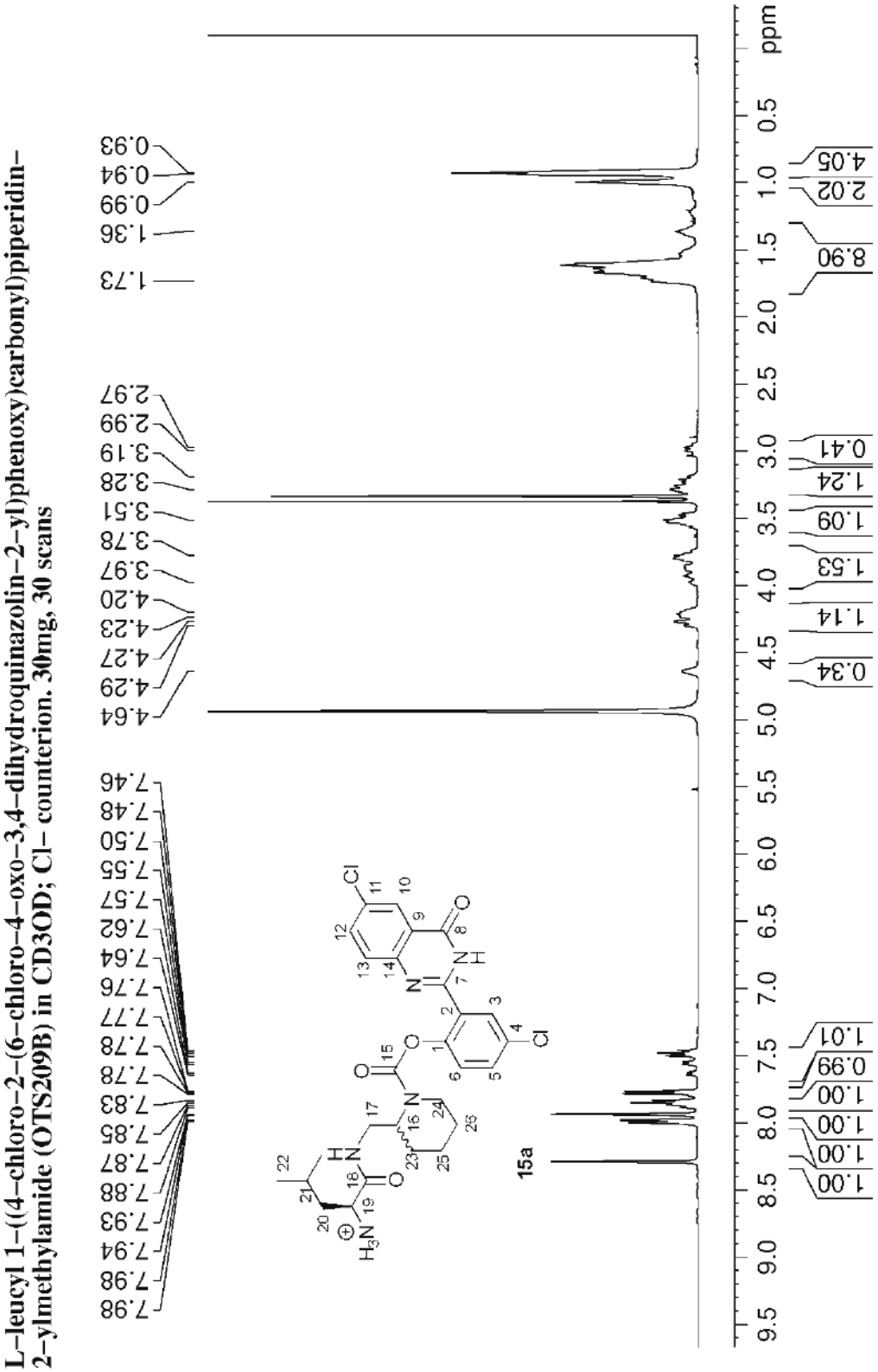
1b: ^{13}C-NMR spectrum (125 MHz, D_2O, TFA counterion)



1b: udeft ^{13}C-NMR, aromatic spectral window (125 MHz, D_2O, TFA counterion). The five problematic quaternary carbon peaks from 165-147 ppm are peak-picked.

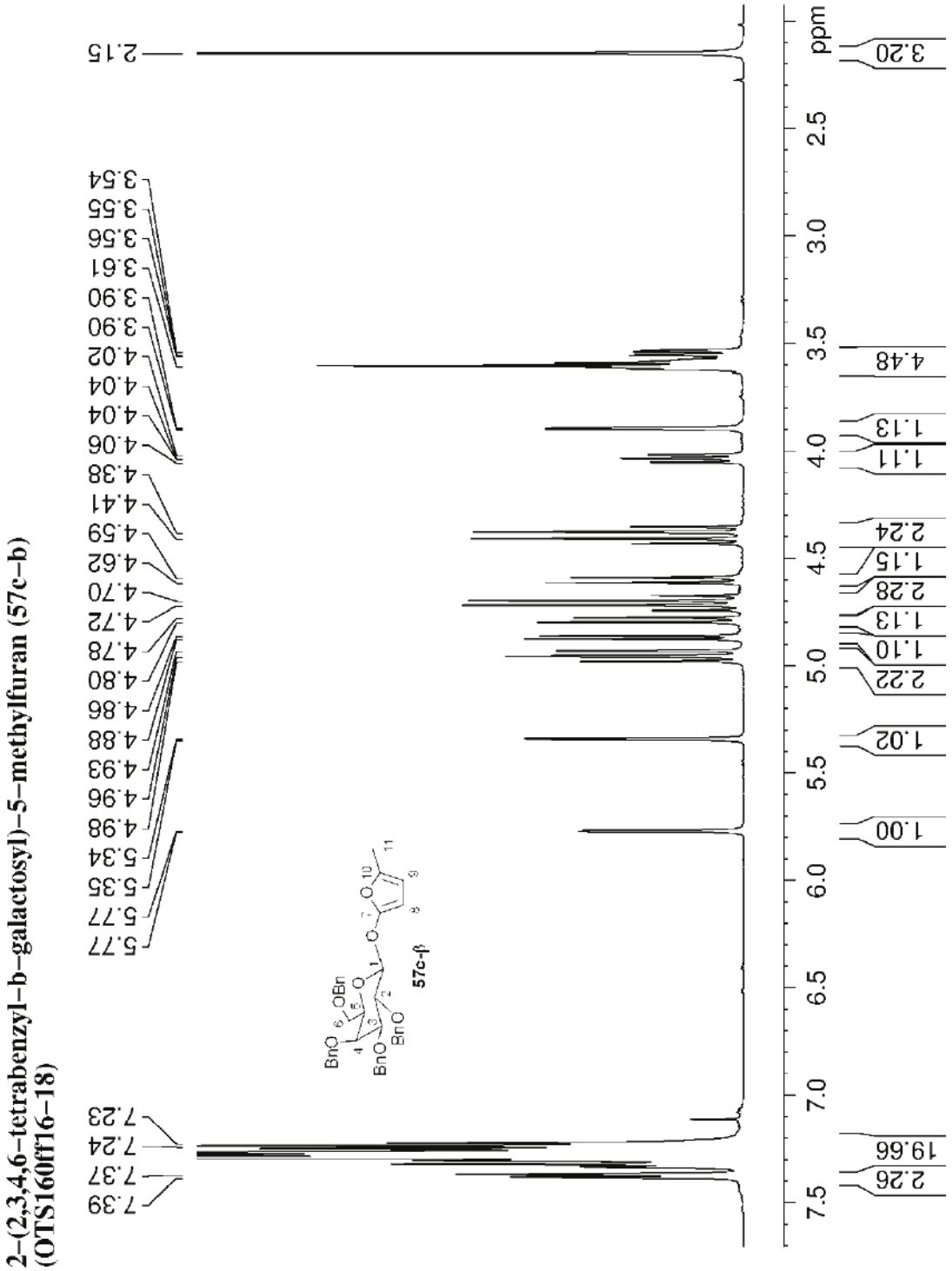


15a: ^{1}H-NMR spectrum (500 MHz, MeOD, chloride counterion)

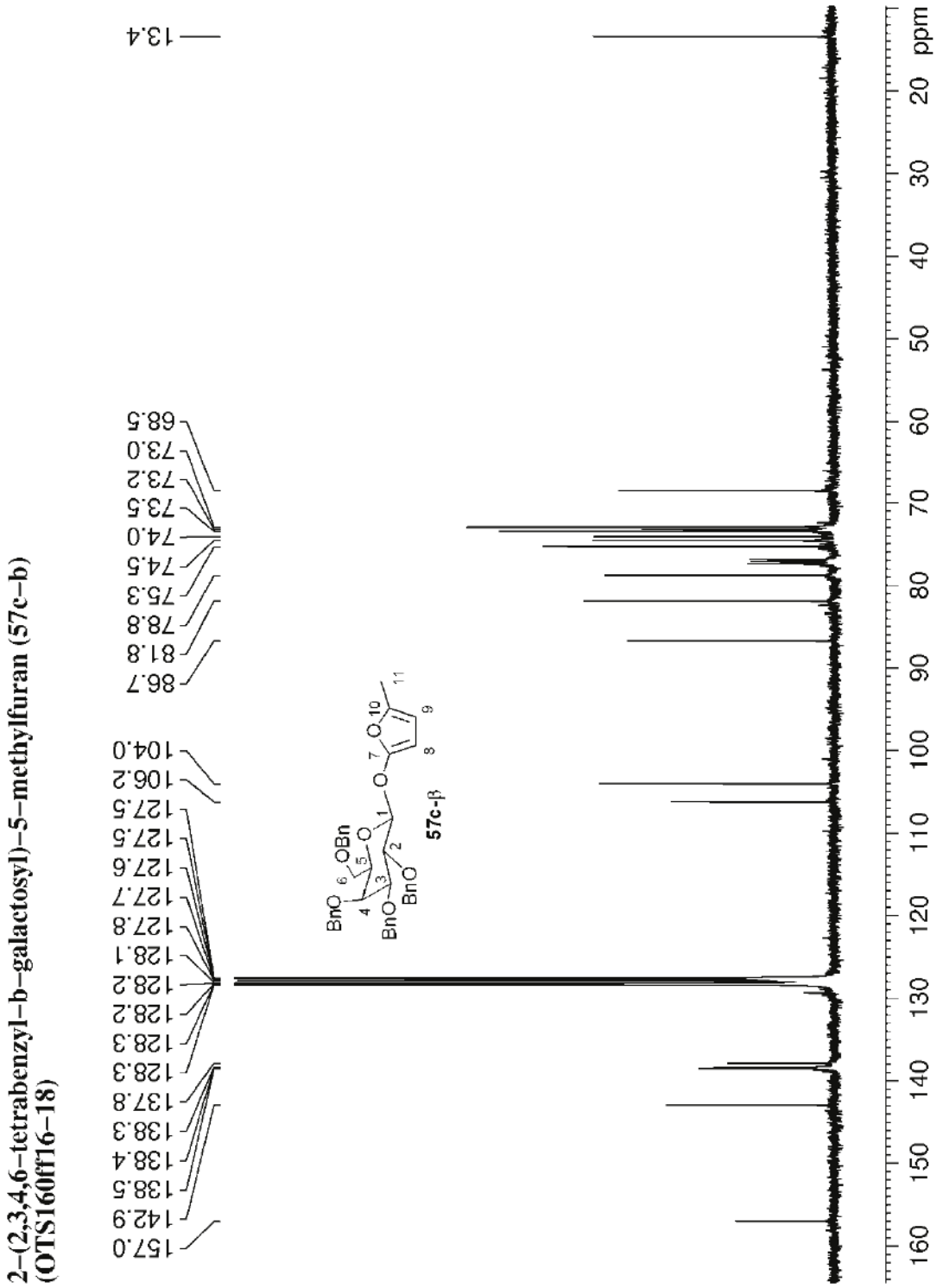


Methylfuranol Glycosides from Alkynyl Cycloisomerisation

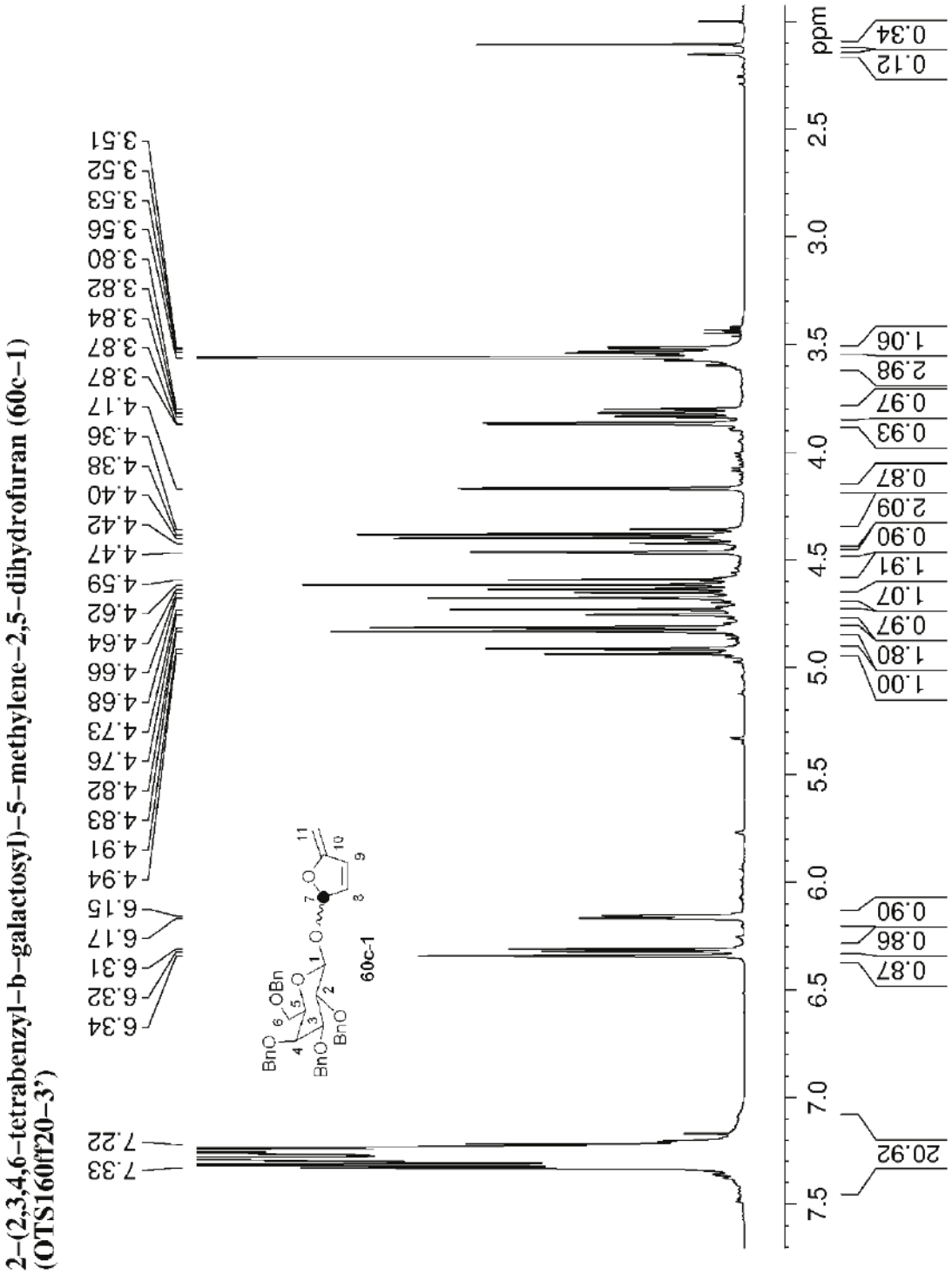
57c-β: ^{1}H-NMR spectrum (500 MHz, $CDCl_3$) (the anomeric **57c-α** presents such similar features that only the β form's spectra are shown)



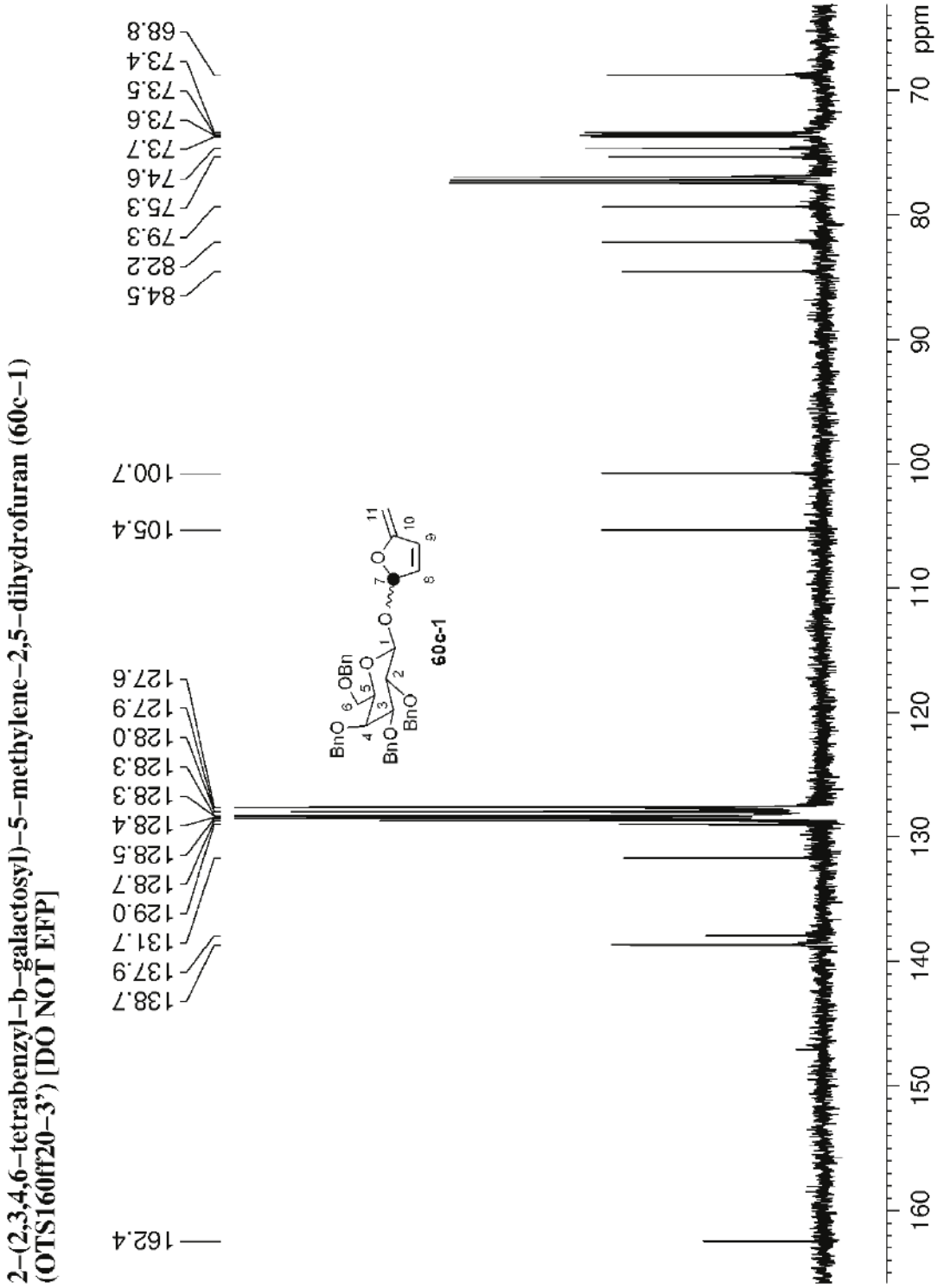
57c-β: ^{13}C-NMR spectrum (125 MHz, $CDCl_3$)



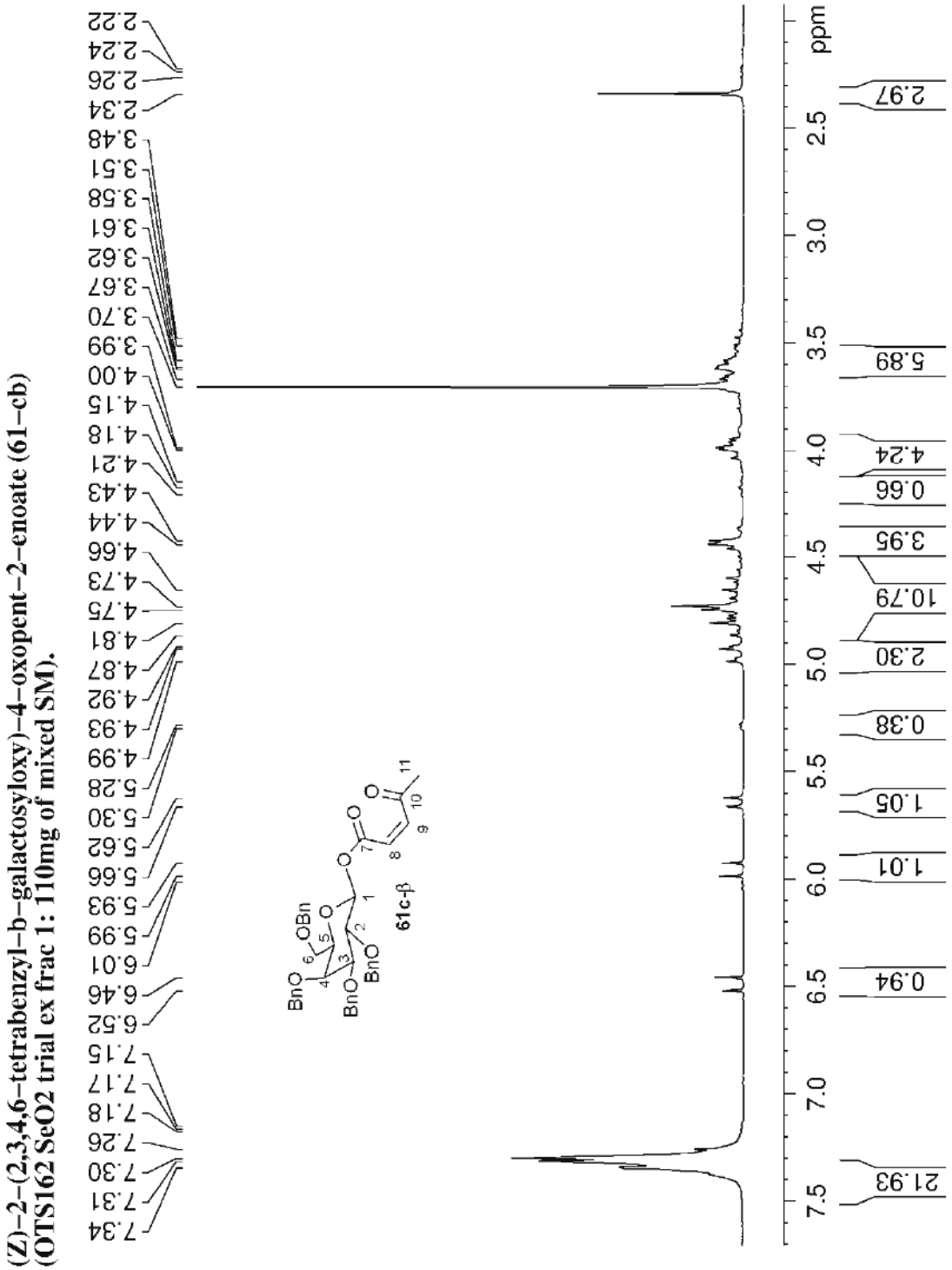
60c-1: ^{1}H-NMR spectrum (500 MHz, $CDCl_3$) (the epimeric **60c-2** presents such similar features that only this epimer's spectra are shown)



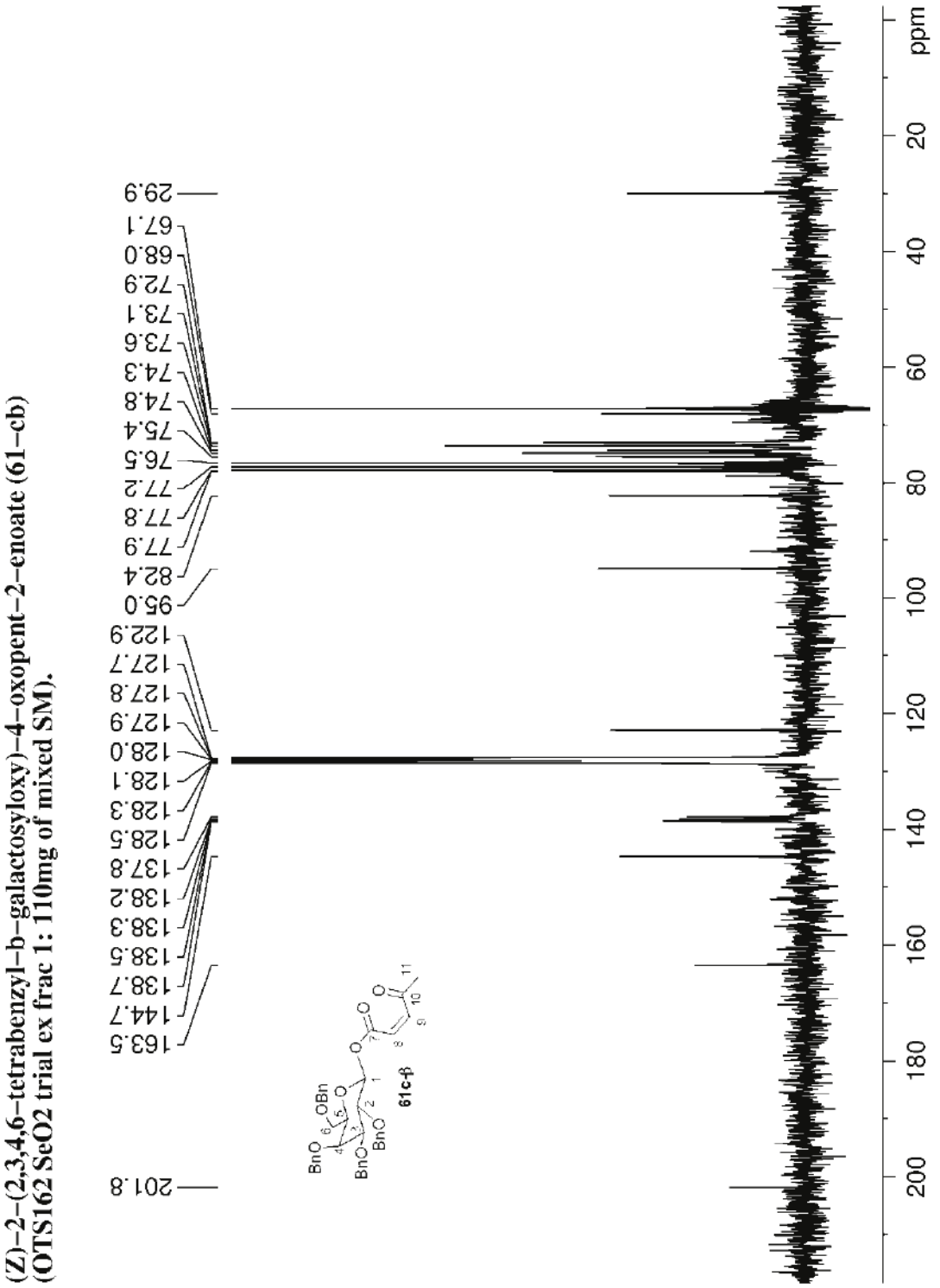
60c-1: ^{13}C-NMR spectrum (125 MHz, $CDCl_3$)



61-cβ: ^{1}H-NMR spectrum (500 MHz, $CDCl_3$)

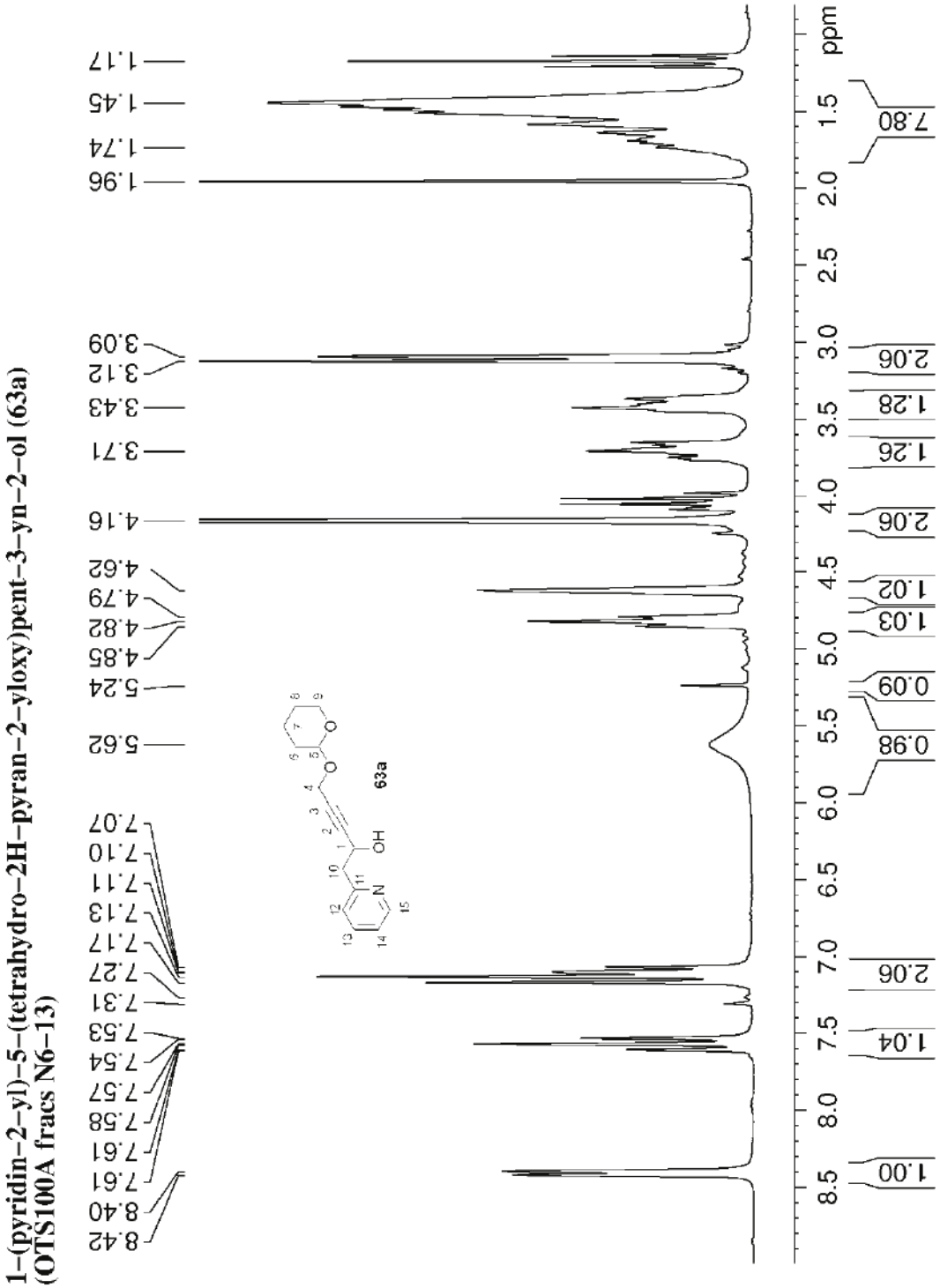


61-cβ: ^{13}C-NMR spectrum (125 MHz, $CDCl_3$)

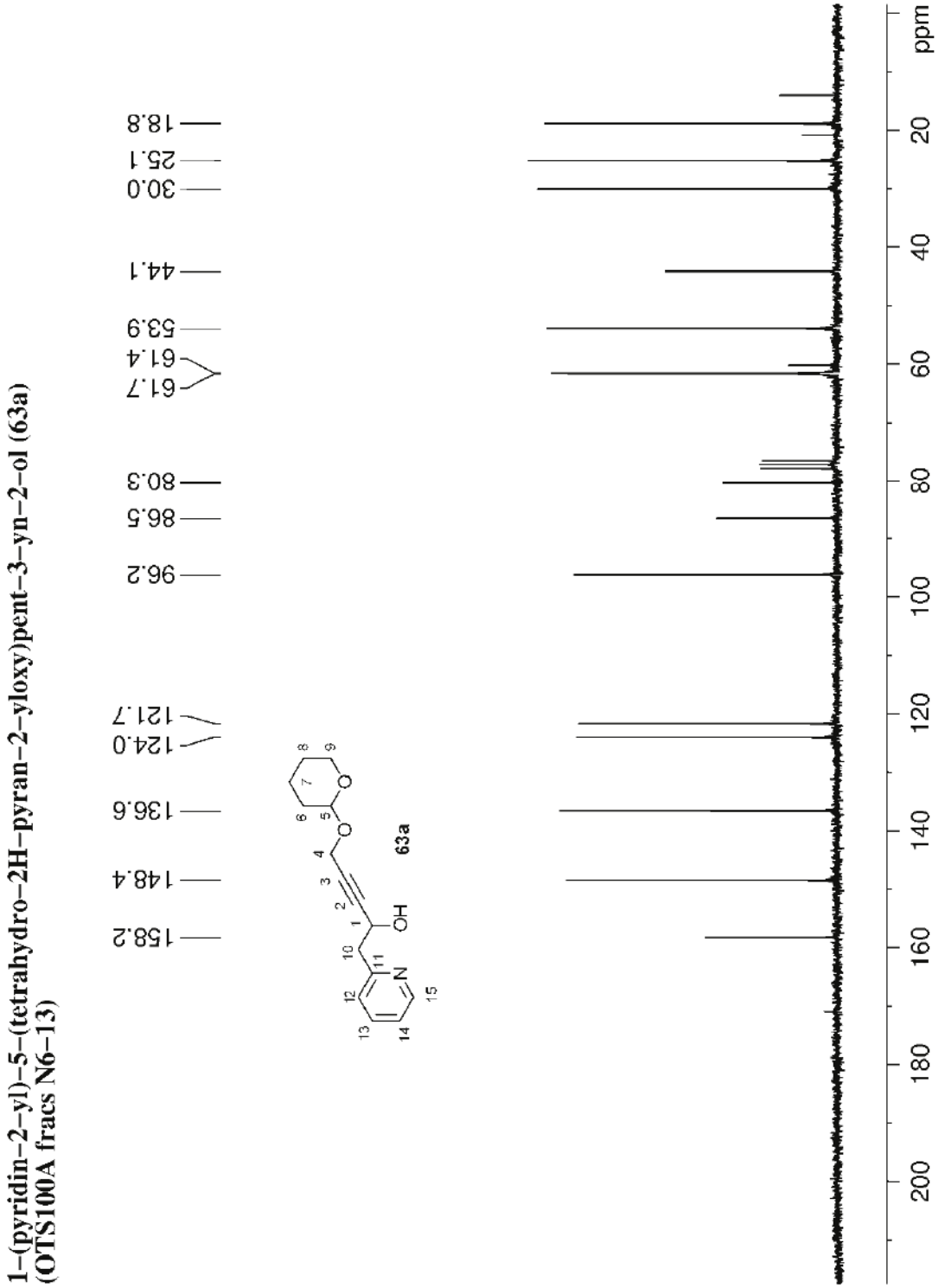


Picolyl Alkynyl Cycloisomerisation Route

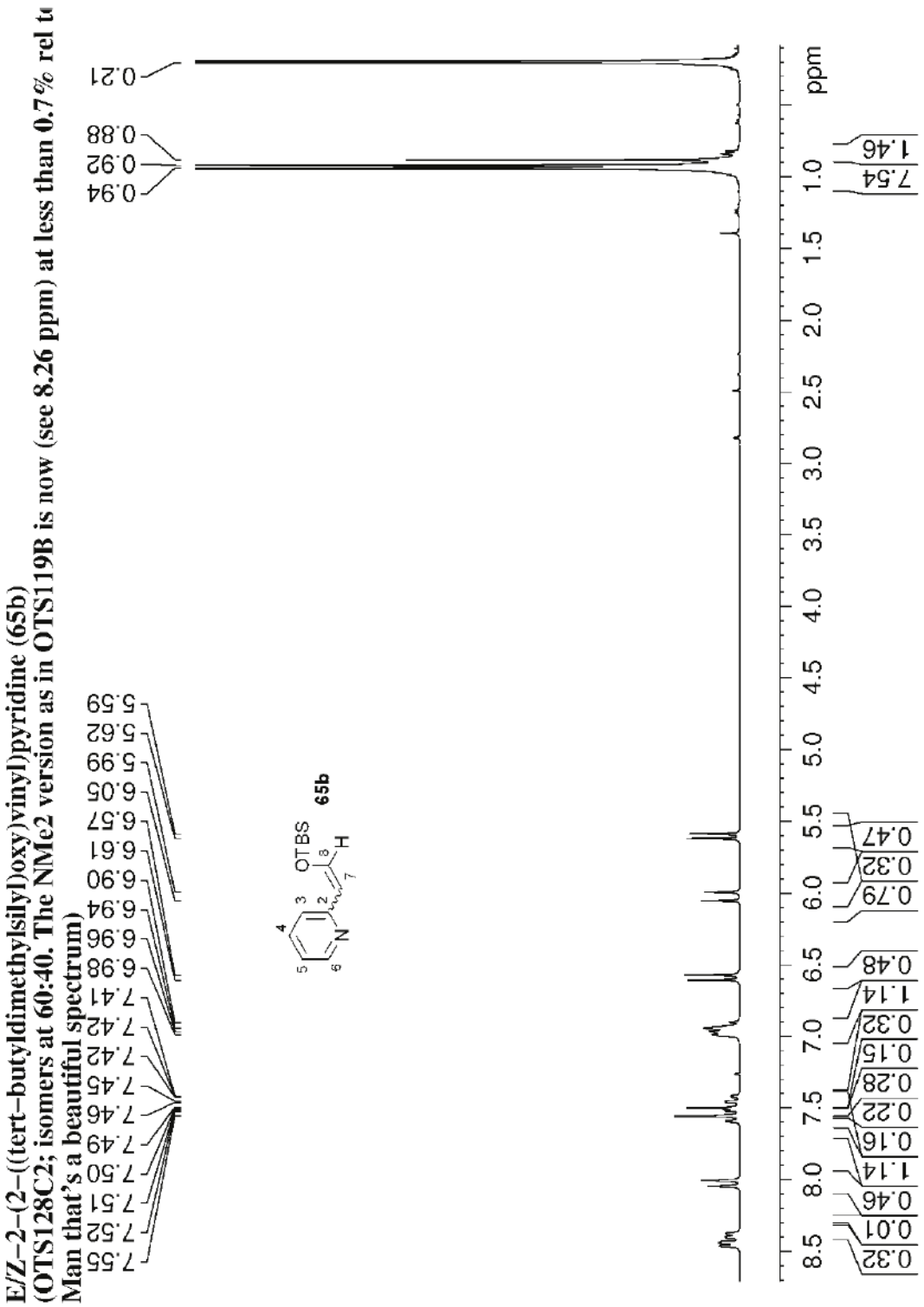
63a: ^{1}H-NMR spectrum (500 MHz, $CDCl_3$)



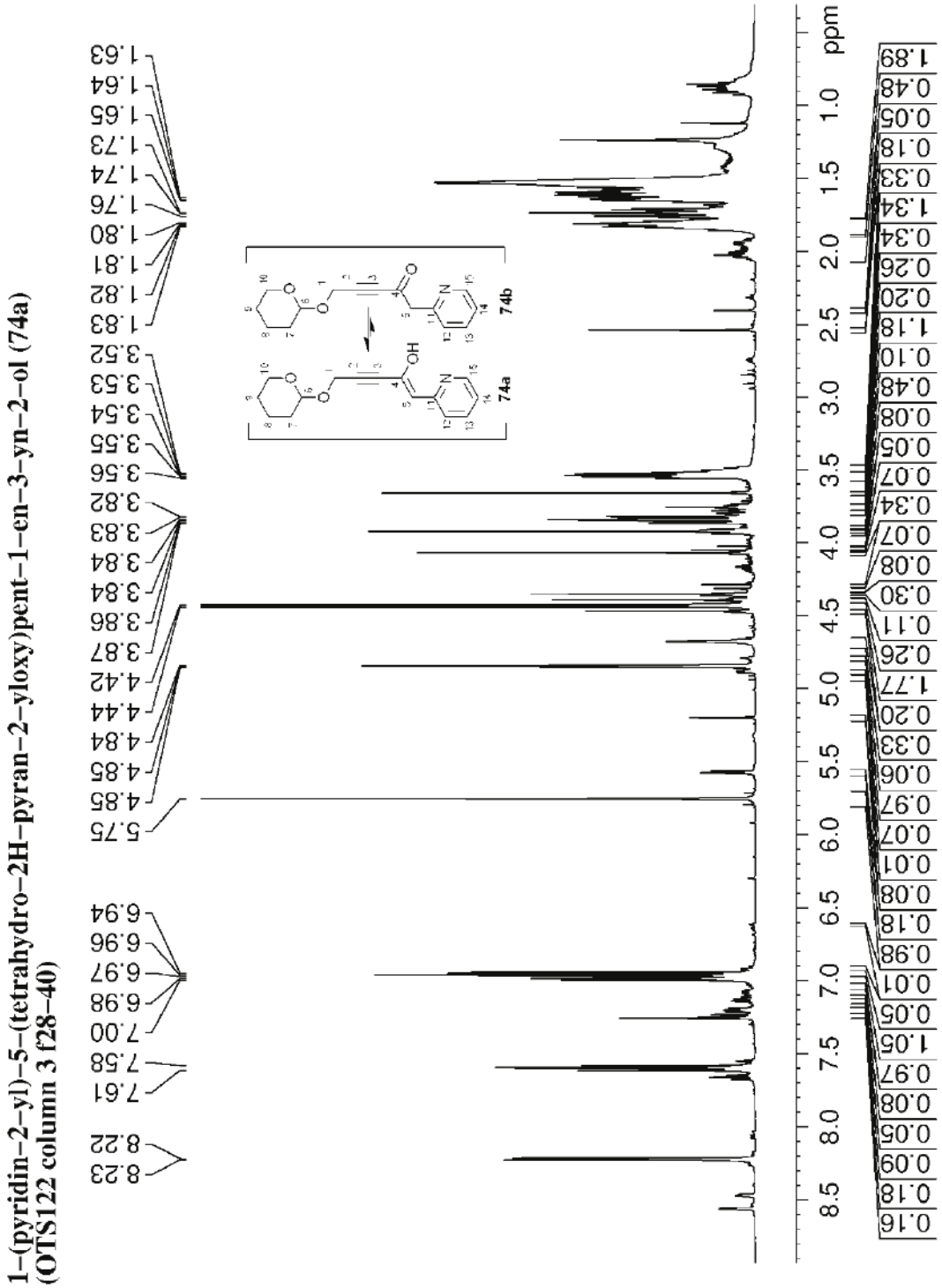
63a: ^{13}C-NMR spectrum (125 MHz, $CDCl_3$)



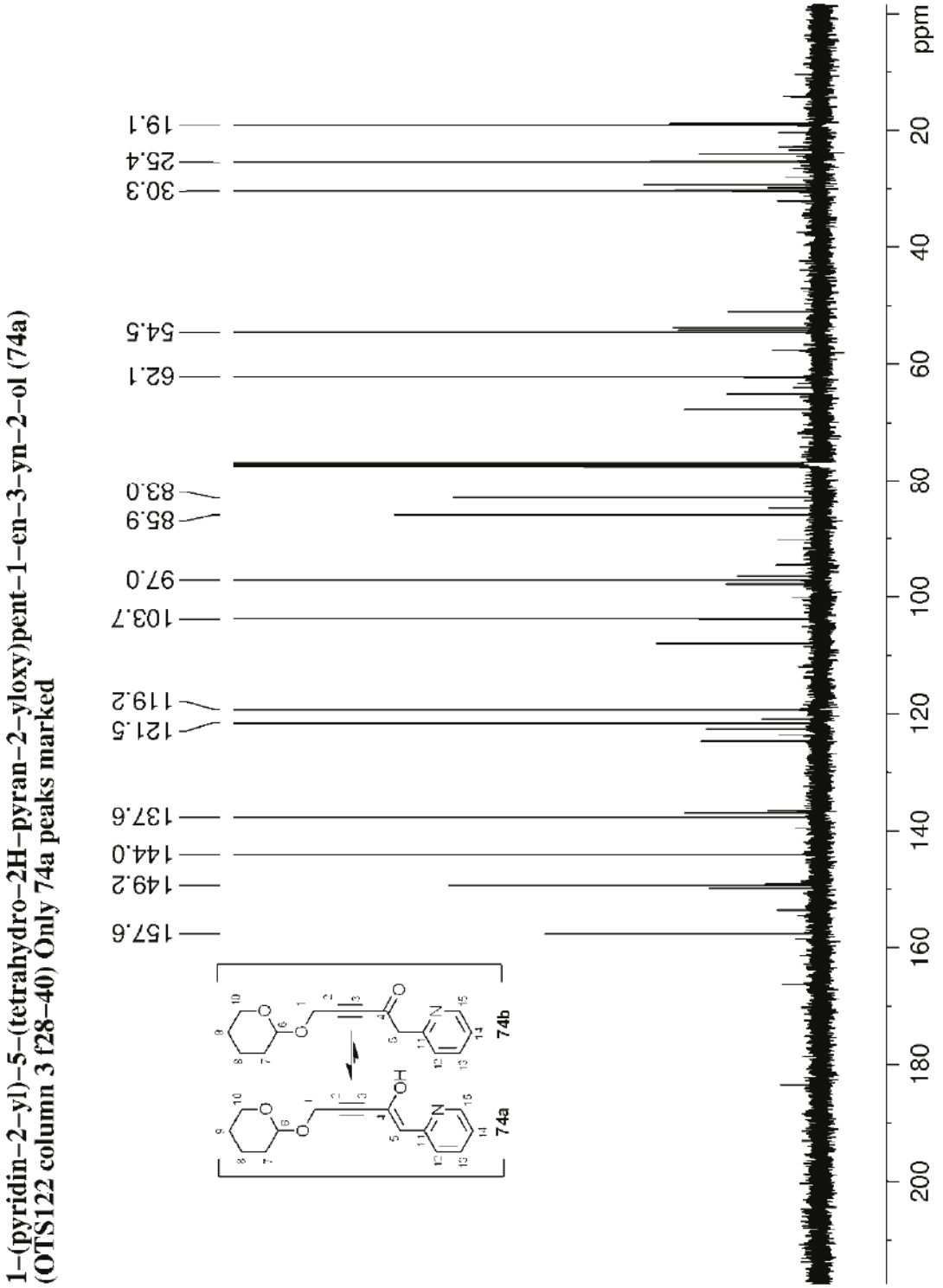
E/Z-65b: ^{1}H-NMR spectrum (500 MHz, $CDCl_3$, crude)



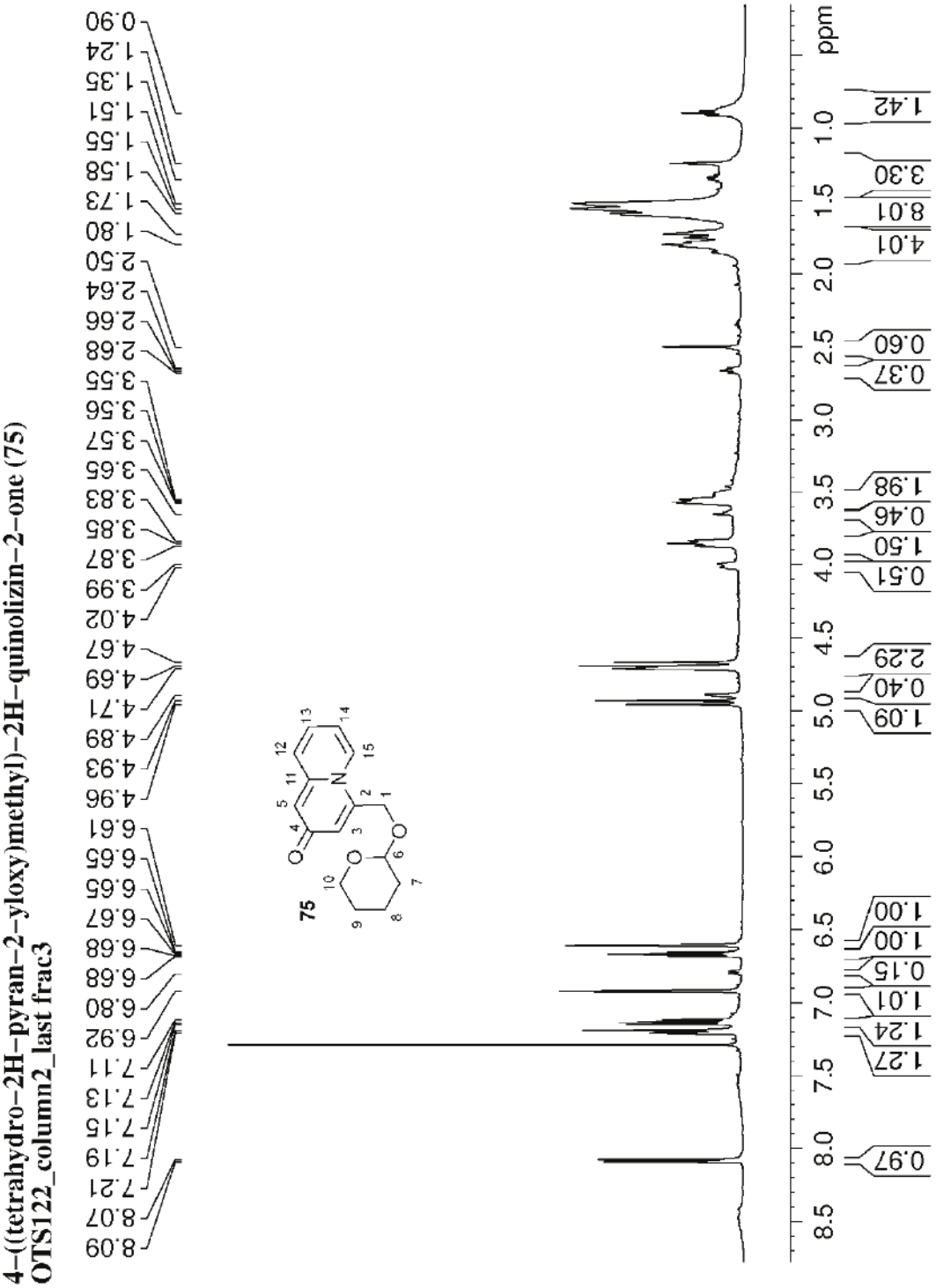
74a: ^{1}H-NMR spectrum (500 MHz, $CDCl_3$). Note the minor peaks from the alkynone form **74b**.



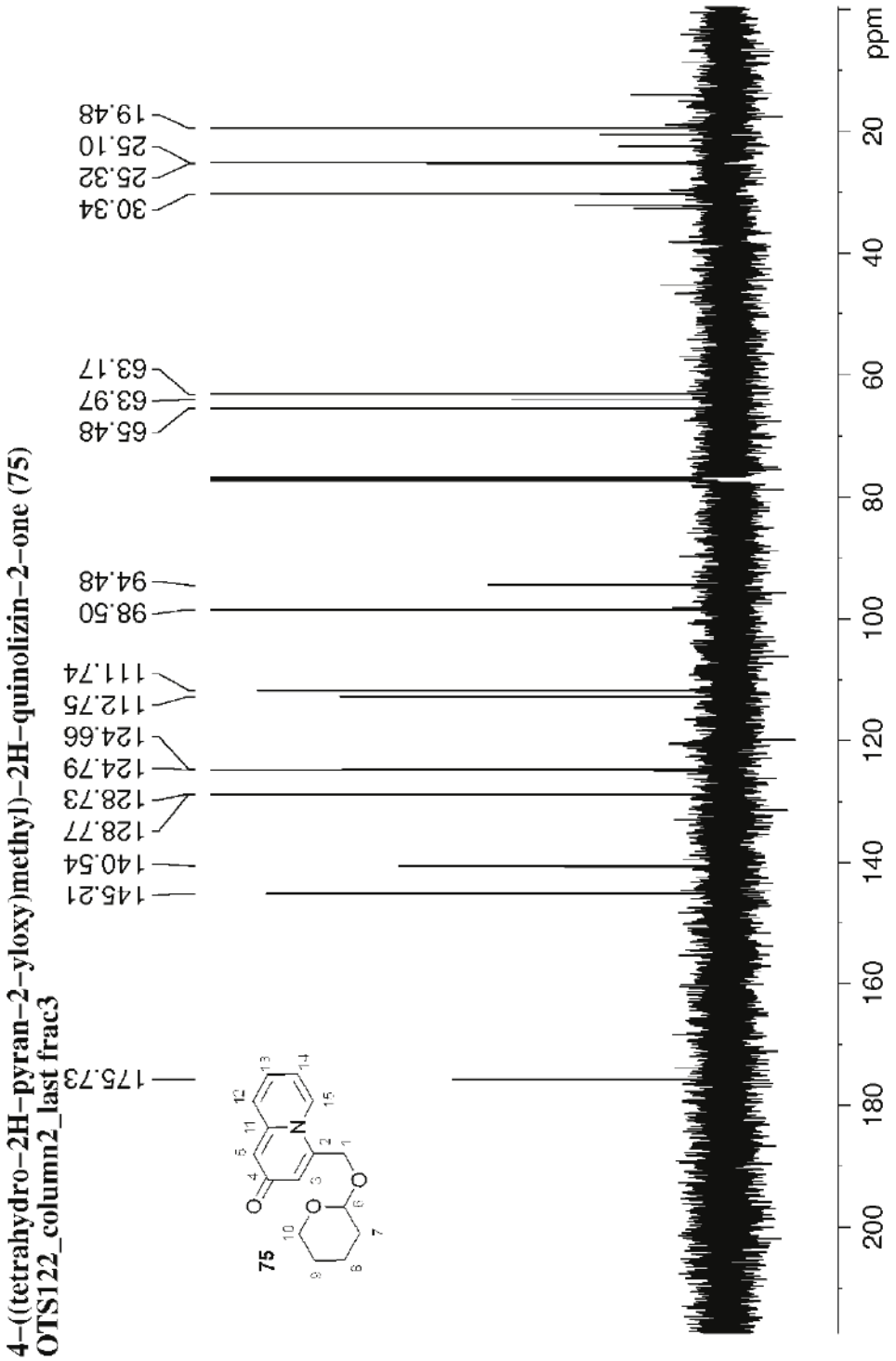
74a: ^{13}C-NMR spectrum (125 MHz, $CDCl_3$). Note the minor peaks from the alkynone form **74b**.



75: ^{1}H-NMR spectrum (500 MHz, $CDCl_3$)

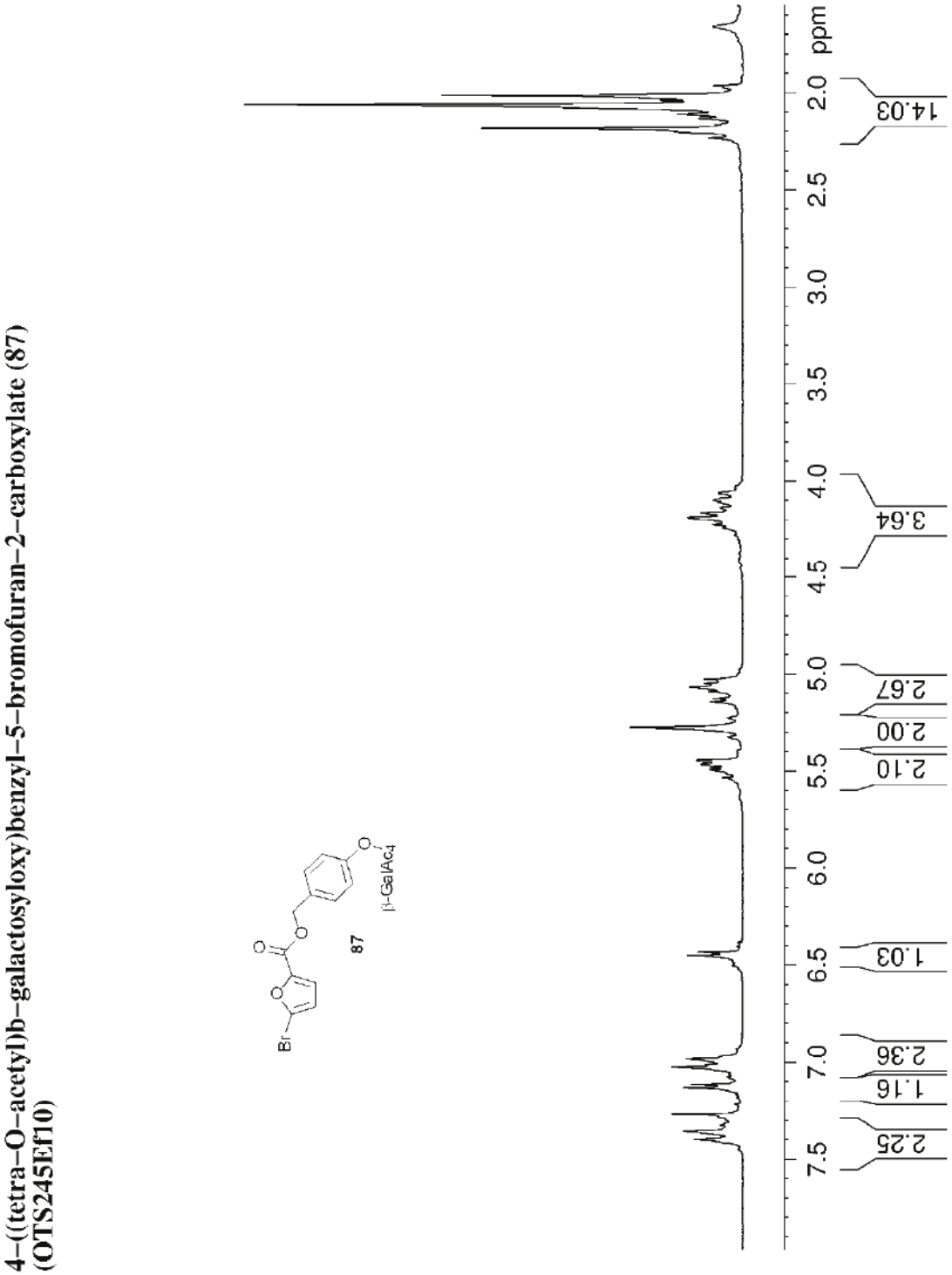


75: ^{13}C-NMR spectrum (125 MHz, $CDCl_3$)

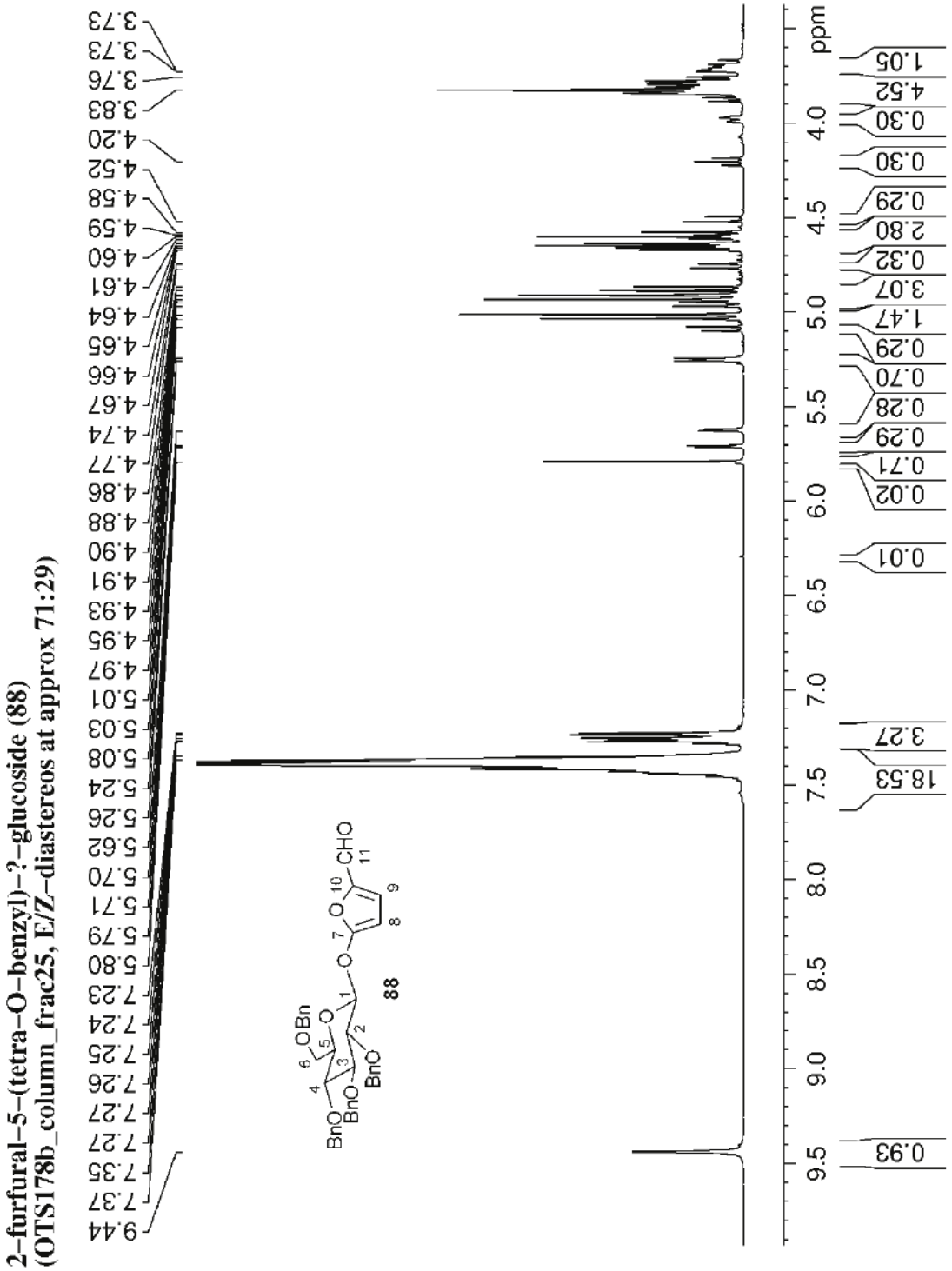


Furan S_NAr Route

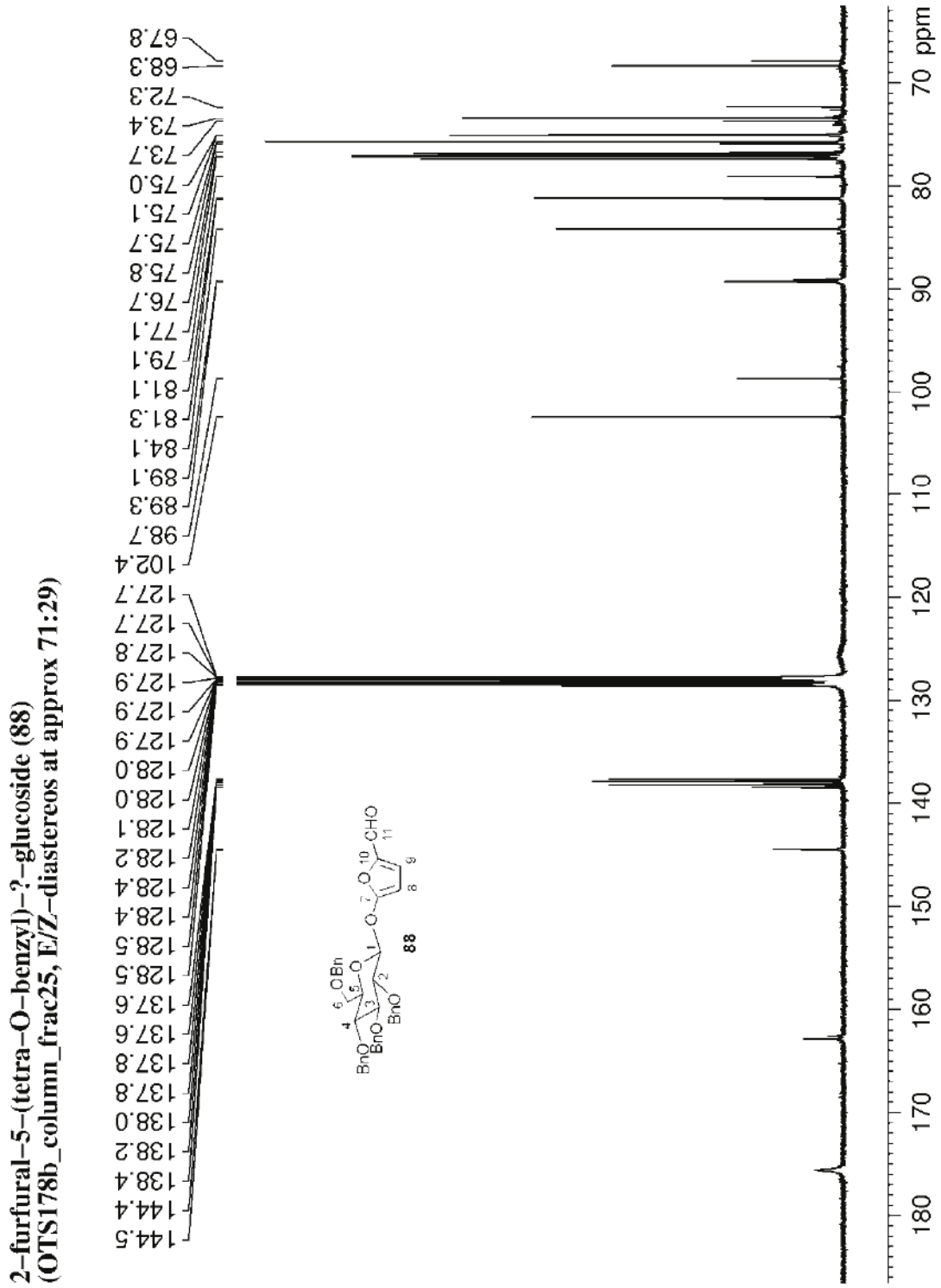
87: ^{1}H-NMR spectrum (500 MHz, $CDCl_3$)



88: ^{1}H-NMR spectrum (500 MHz, $CDCl_3$)

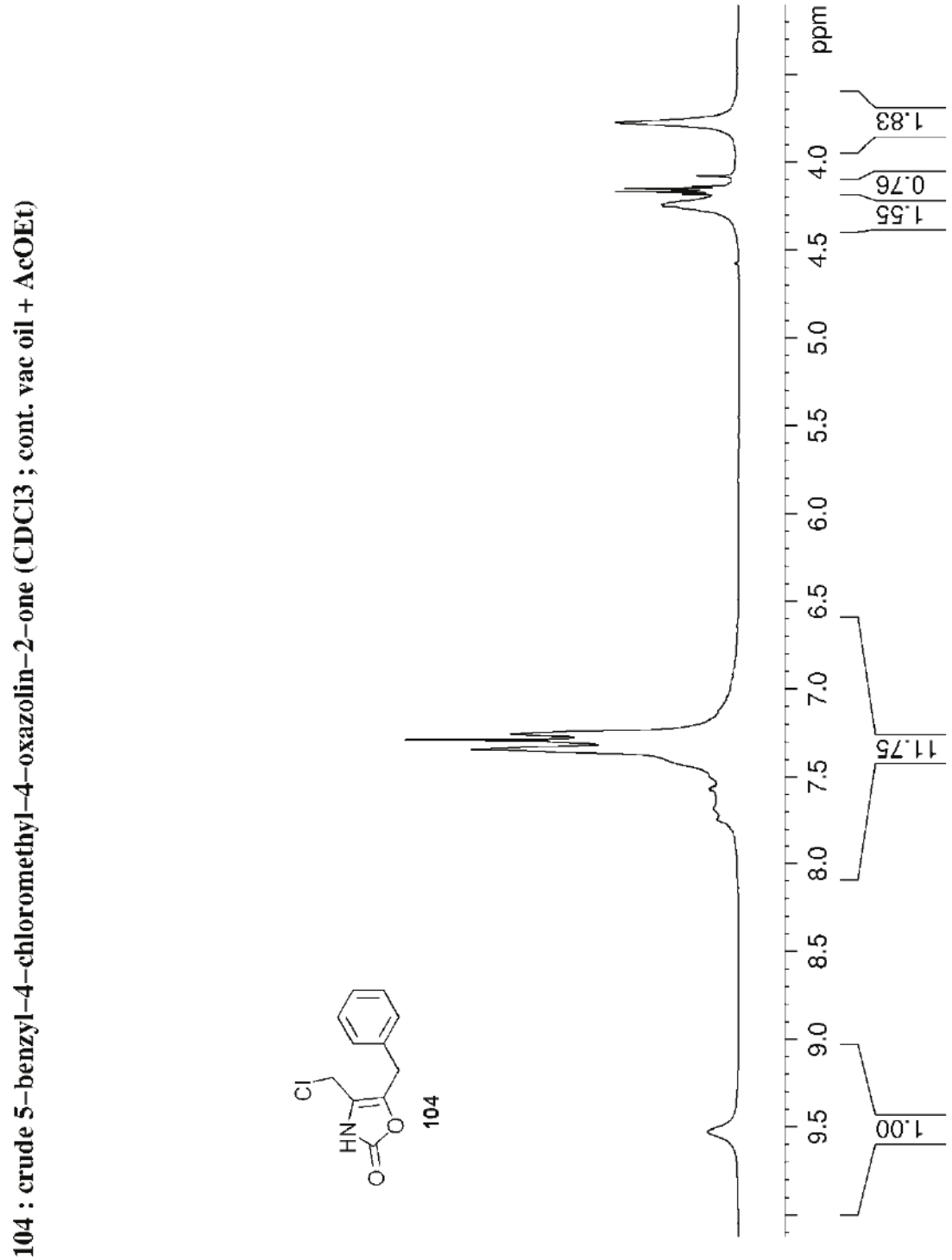


88: ^{13}C-NMR spectrum (125 MHz, $CDCl_3$)

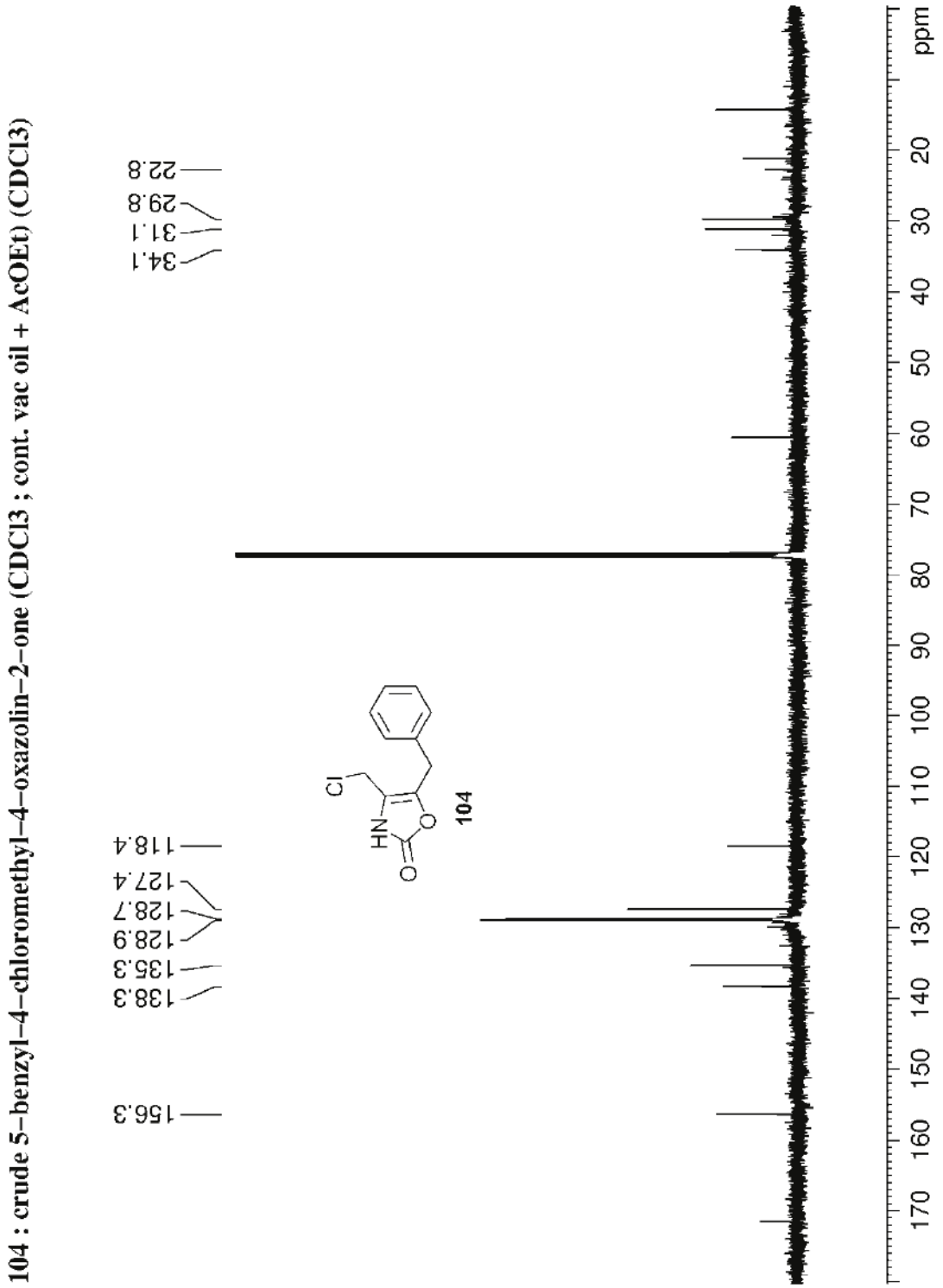


Cyclic Carbamates

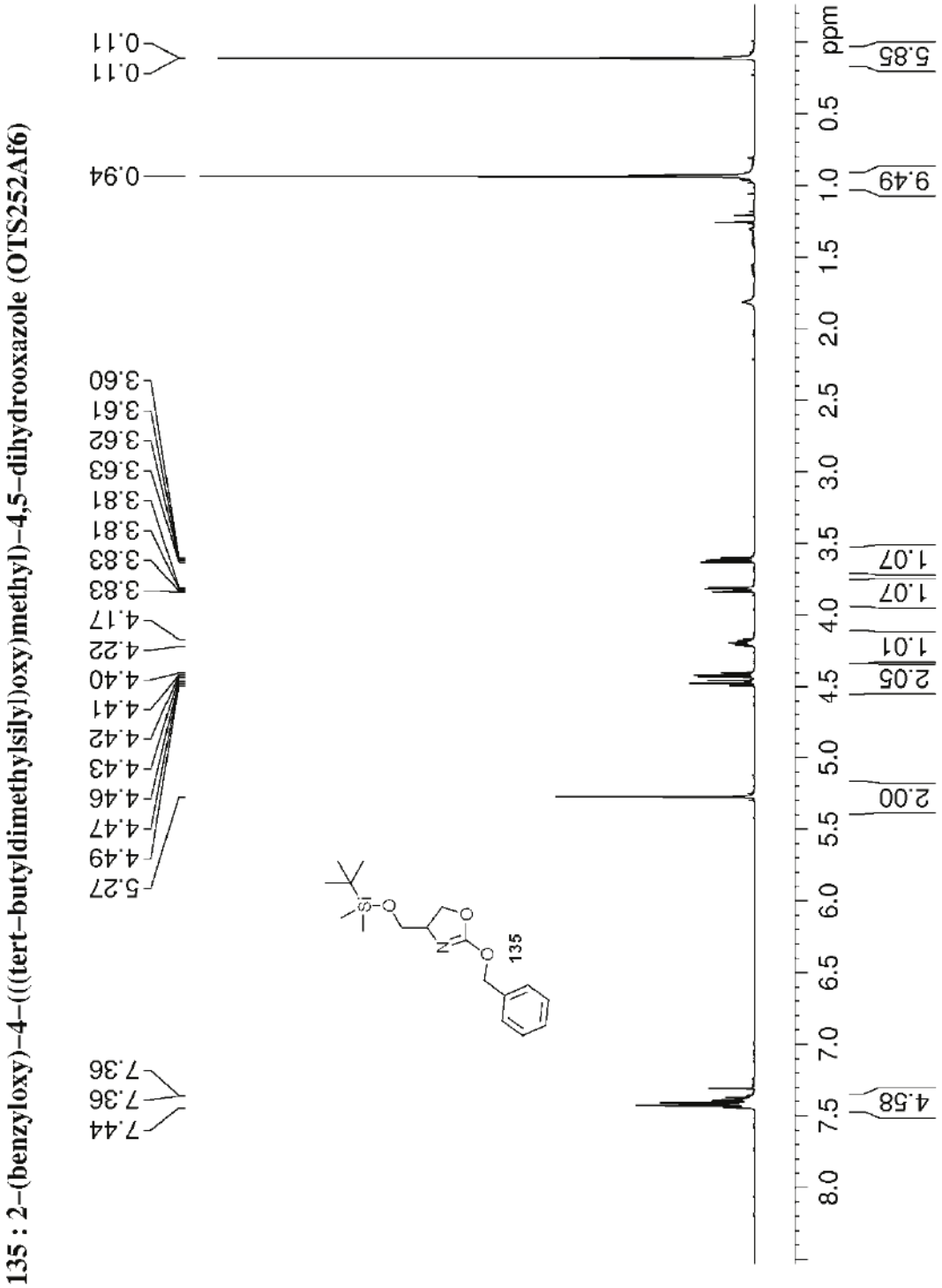
104: ^{1}H-NMR spectrum (500 MHz, $CDCl_3$)



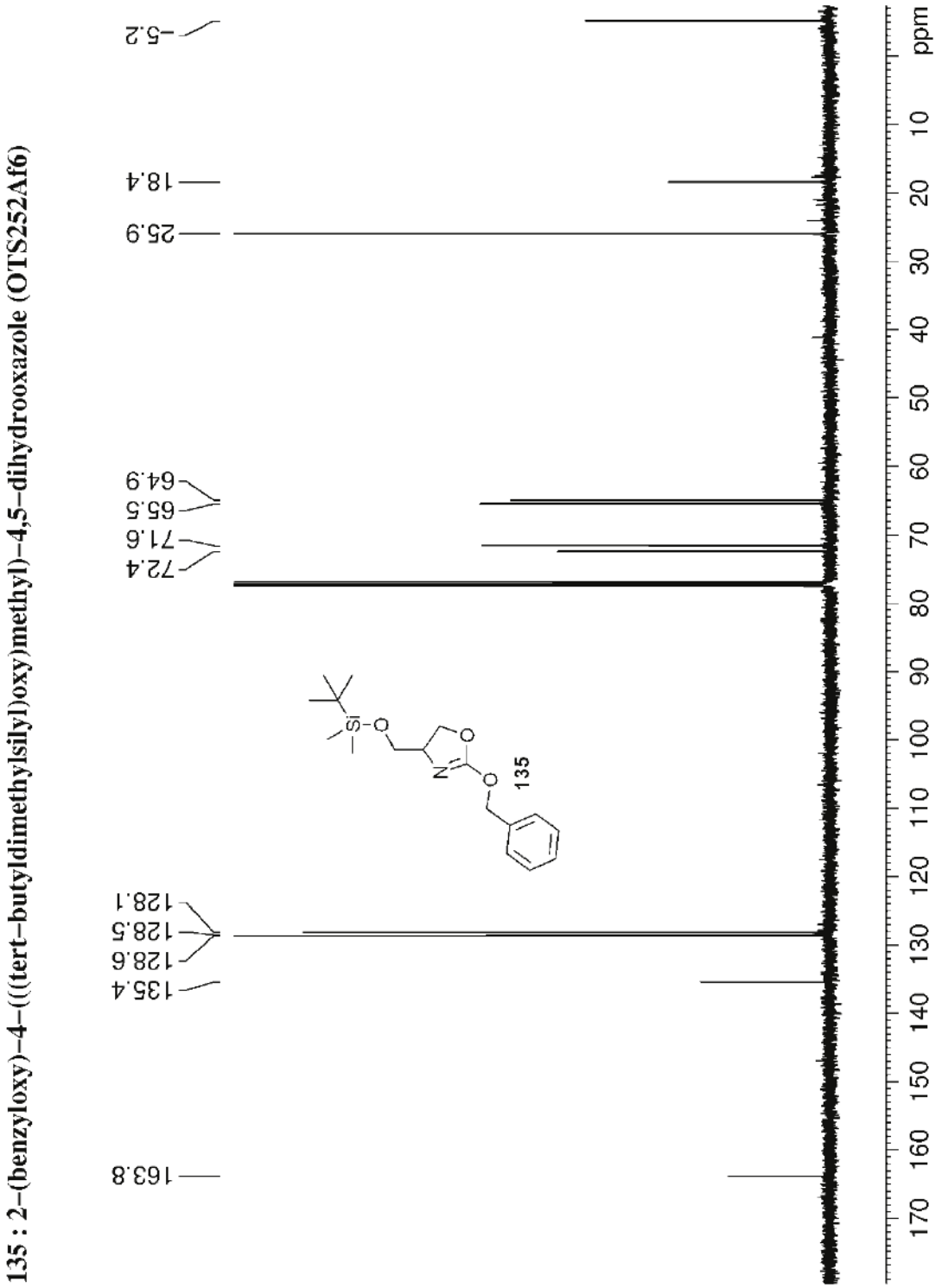
104: ^{13}C-NMR spectrum (125 MHz, $CDCl_3$)



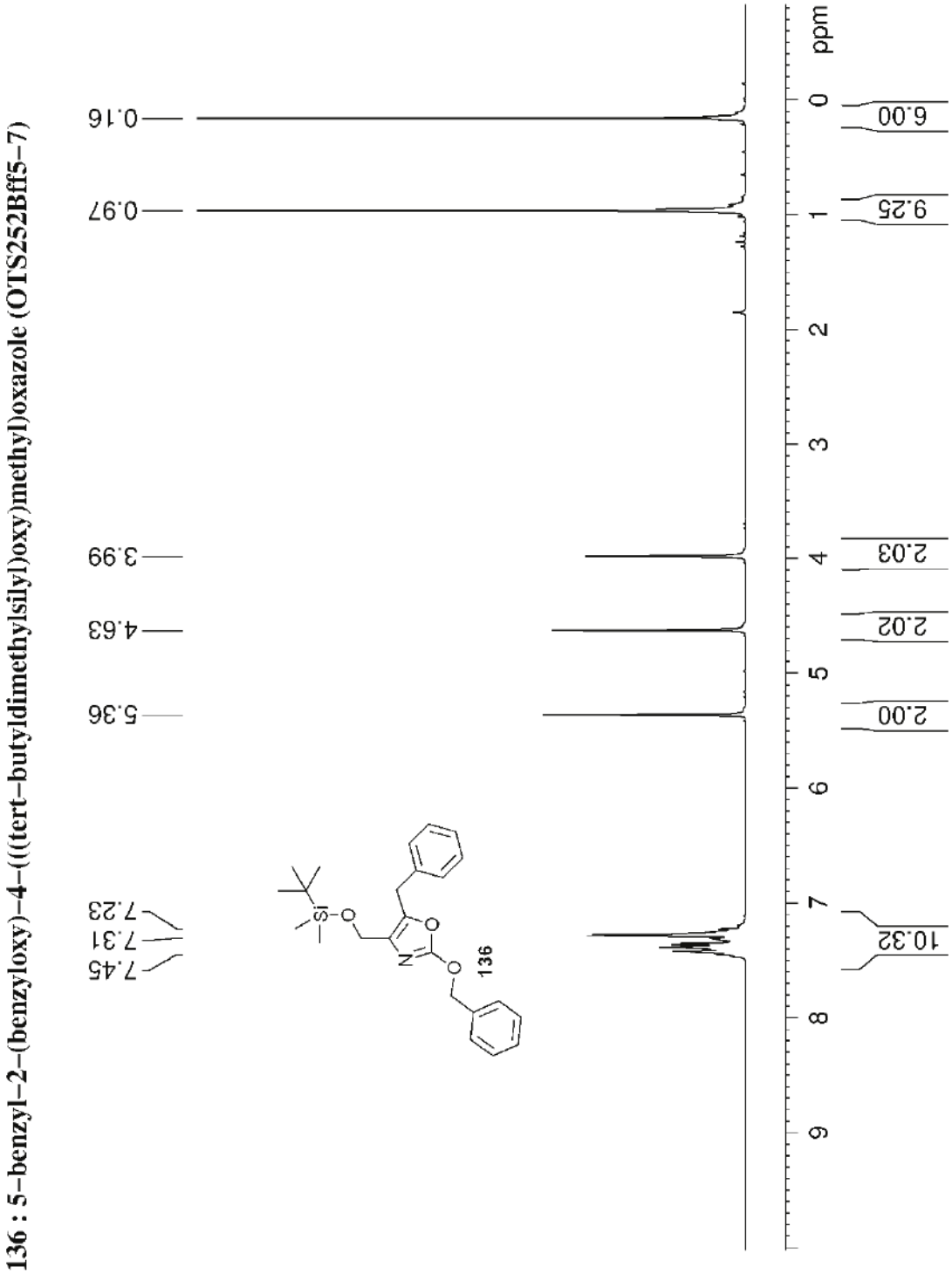
135: ^{1}H-NMR spectrum (500 MHz, $CDCl_3$)



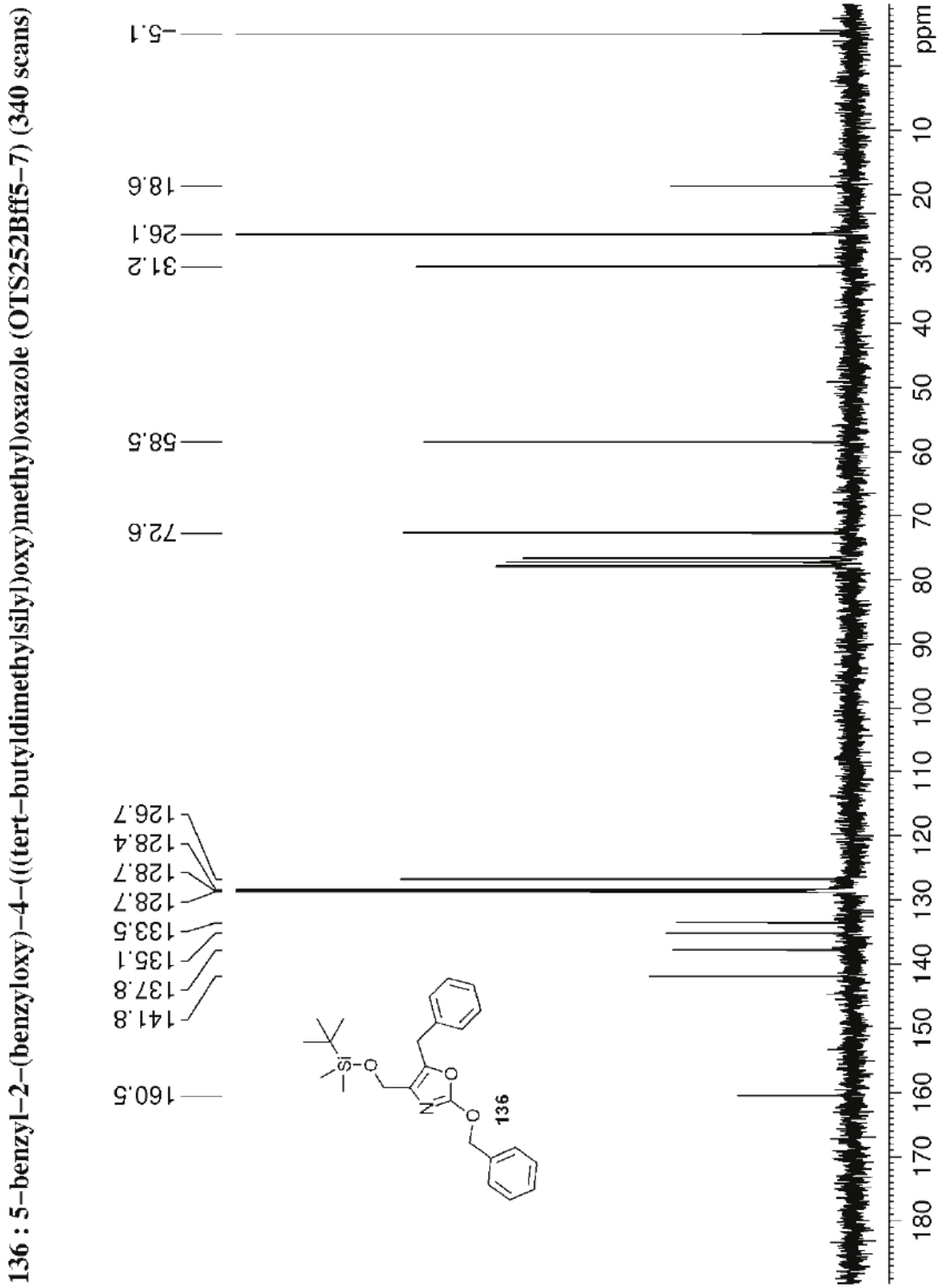
135: ^{13}C-NMR spectrum (125 MHz, $CDCl_3$)



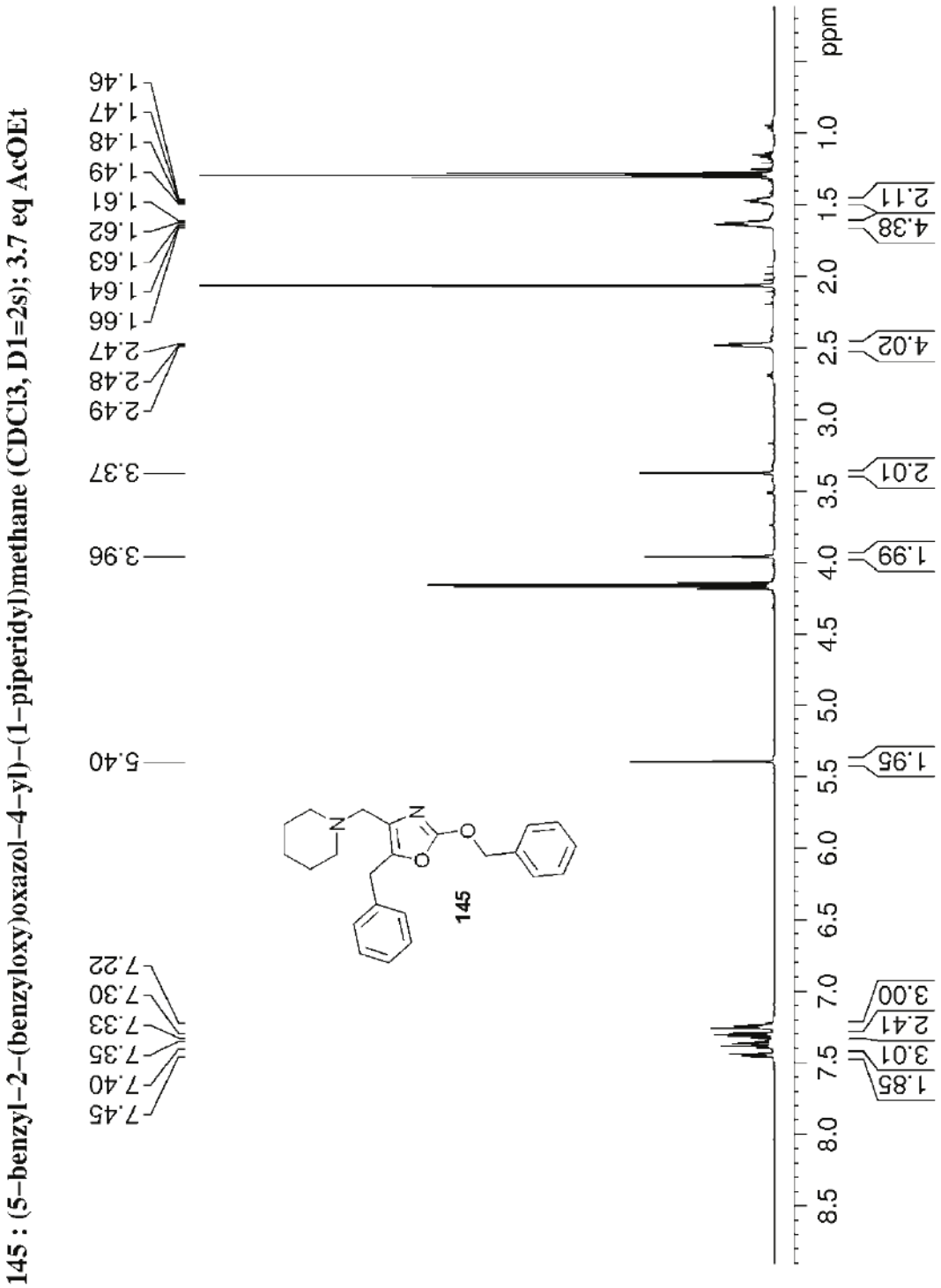
136: ^{1}H-NMR spectrum (500 MHz, $CDCl_3$)



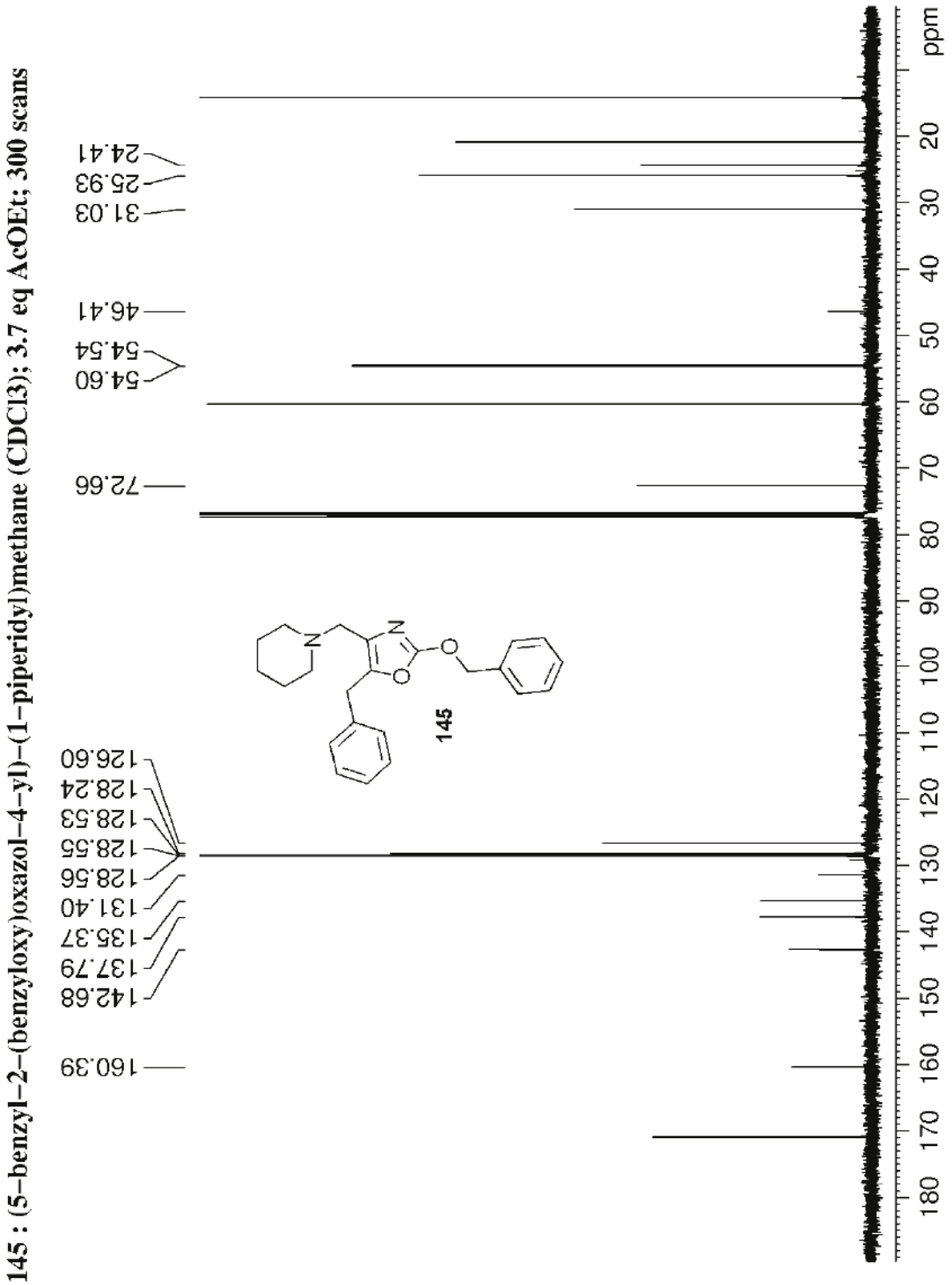
136: ^{13}C-NMR spectrum (125 MHz, $CDCl_3$)



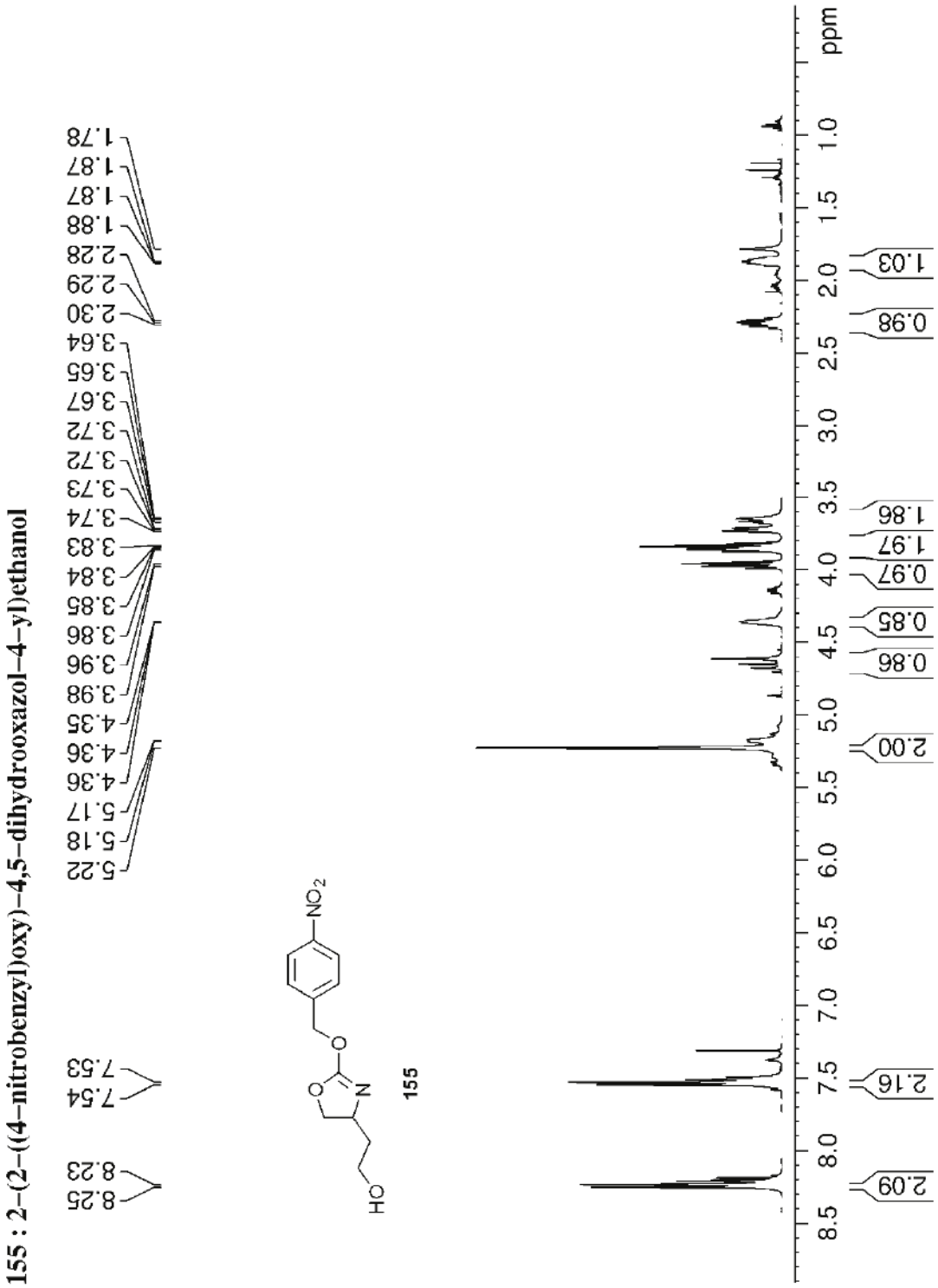
145: ^{1}H-NMR spectrum (500 MHz, $CDCl_3$)



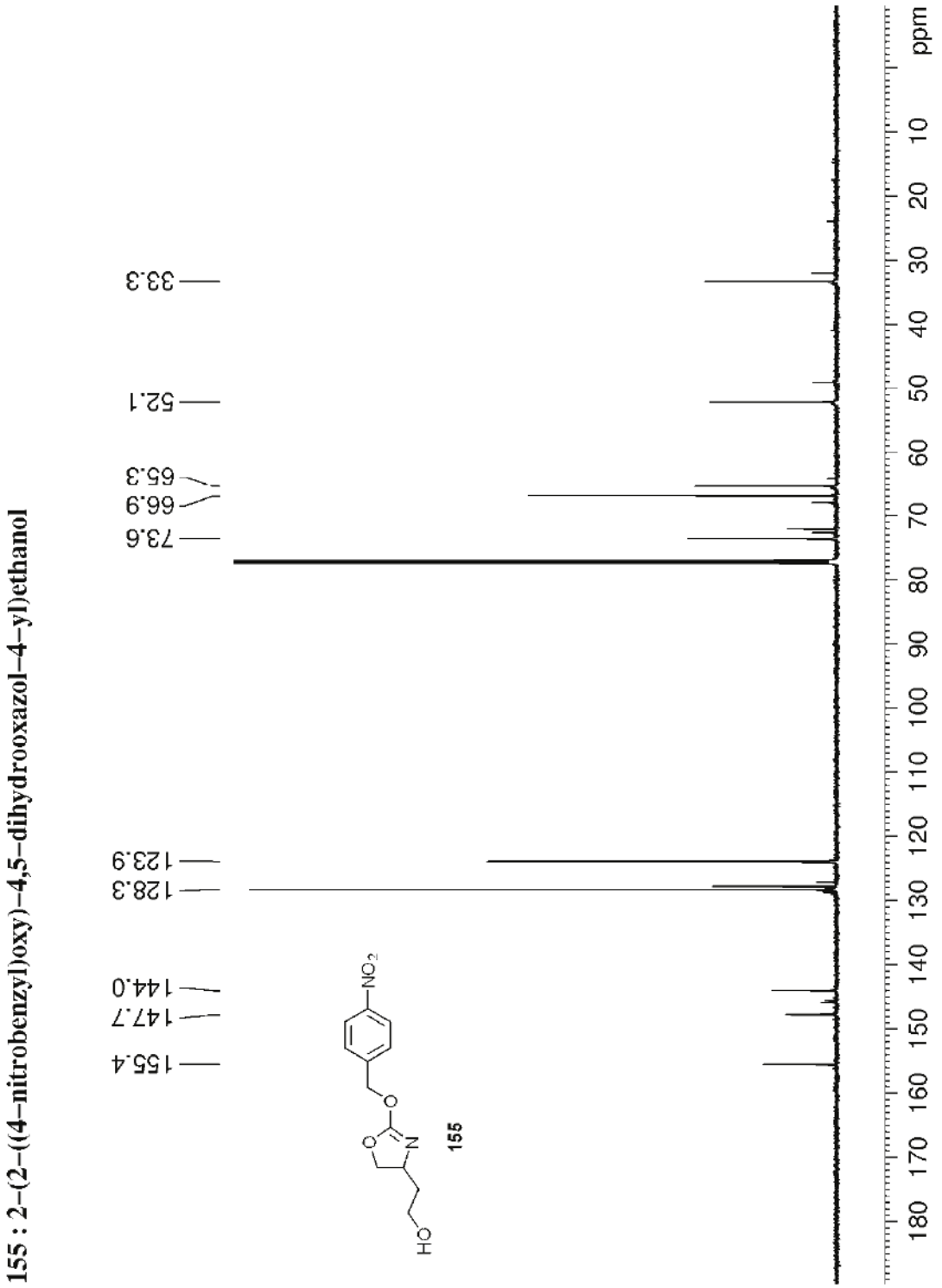
145: ^{13}C-NMR spectrum (125 MHz, $CDCl_3$)



155: ^{1}H-NMR spectrum (500 MHz, $CDCl_3$)



155: ^{13}C-NMR spectrum (125 MHz, $CDCl_3$)



Experimental Bibliography

DOI are indicated where possible (and failing that, PubMed ID [PMID]), for fastest lookup.

[1] H. E. Gottlieb, V. Kotlyar, A. Nudelman; *J. Org. Chem.* **1997**, *62*, 7512-7515 *(10.1021/jo971176v)*.

[2] K. Faghihi, K. Zamani, A. Mirsamie, M. Reza Sangi; *Eur. Pol. J* **2003**, *39*, 247-254 *(10.1016/S0014-3057(02)00200-8)*.

[3] L. J. Exner, L. S. Luskin, P. L. de Benneville; *J. Am. Chem. Soc.* **1953**, *75*, 4841-4842 *(10.1021/ja01115a513)*.

[4] W. K. M. Chong, R. K. Duvadie, L. Li, Y. Yang; *Antiproliferative 2-(heteroaryl)-aminothiazole compounds and pharmaceutical compositions, and methods for their use (US20050038078)*, **2005**.

[5] R. K. Mylavarapu, K. Gcm, N. Kolla, R. Veeramalla, P. Koilkonda, A. Bhattacharya, R. Bandichhor; *Org. Process Res. Dev.* **2007**, *11*, 1065-1068 *(10.1021/op700098w)*.

[6] M. Waibel; *Design and Synthesis of Molecules to Probe Peptidase Activity* (PhD Thesis), Ecole Normale Supérieure de Lyon, **2009**.

[7] M. Baghbanzadeh, M. Dabiri, P. Salehi; *Heterocycles* **2008**, *75*, 2809-2815.

[8] M. Piotto, M. Bourdonneau, K. Elbayed, J.-M. Wieruszeski, G. Lippens; *Magn. Reson. Chem.* **2006**, *44*, 943-947 *(10.1002/mrc.1884)*.

[9] M. Kamiya, D. Asanuma, E. Kuranaga, A. Takeishi, M. Sakabe, M. Miura, T. Nagano, Y. Urano; *J. Am. Chem. Soc.* **2011**, *133*, 12960-12963 *(10.1021/ja204781t)*.

[10] F. Hamon, B. Renoux, C. Chadéneau, J.-M. Muller, S. Papot; *Eur. J. Med. Chem.* **2010**, *45*, 1678-1682 *(10.1016/j.ejmech.2009.12.067)*.

[11] R. E. Beattie, D. J. Guthrie, D. T. Elmore, C. H. Williams, B. Walker; *Biochem. J.* **1987**, *242*, 281-283 *(PMID: 3593241)*.

[12] O. Vincent-Fiquet, J. C. Rogez, R. Plaquet; *Biochimie* **1984**, *66*, 171-174 *(PMID: 6733155)*.

[13] C. T. Goudar, S. K. Harris, M. J. McInerney, J. M. Suflita; *Journal of Microbiological Methods* **2004**, *59*, 317-326 *(10.1016/j.mimet.2004.06.013)*.

[14] X.-b. Zhang, M. Waibel, J. Hasserodt; *Chem.–Eur. J* **2010**, *16*, 792-795 *(10.1002/chem.200902412)*.

[15] O. Thorn-Seshold; *Towards new molecular candidates for an auto-immolative coordinating arm on macrocyclic ligands* (M2 Thesis), Ecole Normale Supérieure de Lyon, **2009**.

[16] P. Hermann, J. Kotek, V. Kubicek, I. Lukes; *J. Chem. Soc., Dalton Trans.* **2008**, 3027-3047 *(10.1039/B719704G)*.

[17] M. Journet, D. Cai, L. M. DiMichele, R. D. Larsen; *Tetrahedron Lett.* **1998**, *39*, 6427-6428 *(10.1016/S0040-4039(98)01352-5)*.

[18] H. B. Mereyala, S. R. Gurrala, S. K. Mohan; *Tetrahedron* **1999**, *55*, 11331-11342 *(10.1016/S0040-4020(99)00631-6)*.

[19] H. B. Mereyala, S. R. Gurrala; *Carbohydr. Res.* **1998**, *307*, 351-354 *(10.1016/S0008-6215(97)10104-5)*.

[20] E. Duranti, C. Balsamini; *Synthesis* **1974**, *1974*, 357-358 *(10.1055/s-1974-23320)*.

[21] A. V. Kel'in, V. Gevorgyan; *J. Org. Chem.* **2002**, *67*, 95-98 *(10.1021/jo010832v)*.

[22] K. Ando, Y. Kawamura, Y. Akai, J.-i. Kunitomo, T. Yokomizo, M. Yamashita, S. Ohta, T. Ohishi, Y. Ohishi; *Org. Biomol. Chem.* **2008**, *6*, 296-307 *(10.1039/b710935k)*.

[23] K. C. Nicolaou, T. Montagnon, P. S. Baran, Y. L. Zhong; *J. Am. Chem. Soc.* **2002**, *124*, 2245-2258 *(10.1021/ja012127+)*.

[24] F. Cermola, M. R. Iesce, G. Buonerba; *J. Org. Chem.* **2005**, *70*, 6503-6505 *(10.1021/jo0504159)*.

[25] D. P. Wang, K. Ding; *Chem. Commun.* **2009**, 1891-1893 *(10.1039/b821212k)*.

[26] P. Y. S. Lam, Y. Ru, P. K. Jadhav, P. E. Aldrich, G. V. DeLucca, C. J. Eyermann, C.-H. Chang, G. Emmett, E. R. Holler, W. F. Daneker, L. Li, P. N. Confalone, R. J. McHugh, Q. Han, R. Li, J. A. Markwalder, S. P. Seitz, T. R. Sharpe, L. T. Bacheler, M. M. Rayner, R. M. Klabe, L. Shum, D. L. Winslow, D. M. Kornhauser, D. A. Jackson, S. Erickson-Viitanen, C. N. Hodge; *J. Med. Chem.* **1996**, *39*, 3514-3525 *(10.1021/jm9602571)*.

[27] R. P. Cassity, L. T. Taylor, J. F. Wolfe; *J. Org. Chem.* **1978**, *43*, 2286-2288 *(10.1021/jo00405a046)*.

[28] I. R. Baxendale, G. Brusotti, M. Matsuoka, S. V. Ley; *J. Chem. Soc.-Perkin Trans. 1* **2002**, 143-154 *(10.1039/b109482n)*.

[29] U. Ghosh, D. Ganessunker, V. J. Sattigeri, K. E. Carlson, D. J. Mortensen, B. S. Katzenellenbogen, J. A. Katzenellenbogen; *Bioorg. Med. Chem.* **2003**, *11*, 629-657 *(10.1016/S0968-0896(02)00309-7)*.

[30] A. Choudhury, M. Breslav, J. S. Grimm, T. Xiao, D. Xu, K. L. Sorgi; *Tetrahedron Lett.* **2007**, *48*, 3069-3072 *(10.1016/j.tetlet.2007.02.115)*.

[31] M. Klein, M. Zabel, G. Bernhardt, B. König; *J. Org. Chem.* **2003**, *68*, 9379-9383 *(10.1021/jo035250n)*.

[32] M. Momenteau, J. Mispelter, B. Loock, J.-M. Lhoste; *J. Chem. Soc., Perkin Trans. 1* **1985**, 221-231 *(10.1039/p19850000221)*.

[33] K. Schank; *Chem. Ber.* **1969**, *102*, 383-387 *(10.1002/cber.19691020202)*.

[34] A. Takács, B. Jakab, A. Petz, L. Kollár; *Tetrahedron* **2007**, *63*, 10372-10378 *(10.1016/j.tet.2007.07.017)*.

[35] W. J. Morris, M. D. Shair; *Org. Lett.* **2008**, *11*, 9-12 *(10.1021/ol8022006)*.

[36] T. Yamanoi, N. Misawa, S. Matsuda, M. Watanabe; *Carbohydr. Res.* **2008**, *343*, 1366-1372 *(10.1016/j.carres.2008.03.024)*.

[37] J. Fiandor, M. T. Garcia-Lopez, F. G. De Las Heras, P. P. Mendez-Castrillan; *Synthesis* **1985**, *1985*, 1121-1123 *(10.1055/s-1985-31446)*.

[38] S. Inack-Ngi, R. Rahmani, L. Commeiras, G. Chouraqui, J. Thibonnet, A. Duchêne, M. Abarbri, J.-L. Parrain; *Adv. Syn. Cat.* **2009**, *351*, 779-788 *(10.1002/adsc.200800757)*.

[39] L. C. King, G. K. Ostrum; *J. Org. Chem.* **1964**, *29*, 3459-3461 *(10.1021/jo01035a003)*.

[40] M. A. Kinsella, V. J. Kalish, S. M. Weinreb; *J. Org. Chem.* **1990**, *55*, 105-111 *(10.1021/jo00288a022)*.

[41] L. M. Weinstock, R. B. Currie, A. V. Lovell; *Synth. Commun.* **1981**, *11*, 943-946 *(10.1080/00397918108065753)*.

[42] J. F. Okonya, R. V. Hoffman, M. C. Johnson; *J. Org. Chem.* **2002**, *67*, 1102-1108 *(10.1021/jo010630z)*.

[43] R. V. Hoffman, M. C. Johnson, J. F. Okonya; *Tetrahedron Lett.* **1998**, *39*, 1283-1286 *(10.1016/S0040-4039(97)10854-1)*.

[44] M. Majewski, D. M. Gleave, P. Nowak; *Can. J. Chem.-Rev. Can. Chim.* **1995**, *73*, 1616-1626 *(10.1139/v95-201)*.

[45] L. A. Dakin, N. F. Langille, J. S. Panek; *J. Org. Chem.* **2002**, *67*, 6812-6815 *(10.1021/jo0204000)*.

[46] C. Bonini, L. Chiummiento, M. Funicello, P. Lupattelli, M. Pullez; *Eur. J. Org. Chem.* **2006**, *2006*, 80-83 *(10.1002/ejoc.200500737)*.

[47] J. Waser, B. Gaspar, H. Nambu, E. M. Carreira; *J. Am. Chem. Soc.* **2006**, *128*, 11693-11712 *(10.1021/ja062355+)*.

[48] D. A. Fielder, F. W. Collins; *J. Nat. Prod.* **1995**, *58*, 456-458 *(10.1021/np50117a019)*.

[49] D. R. Chancellor, K. E. Davies, O. De Moor, C. R. Dorgan, P. D. Johnson, A. G. Lambert, D. Lawrence, C. Lecci, C. Maillol, P. J. Middleton, G. Nugent, S. v. D. Poignant, A. C. Potter, P. D. Price, R. J. Pye, R. Storer, J. M. Tinsley, R. van Well, R. Vickers, J. Vile, F. J. Wilkes, F. X. Wilson, S. P. Wren, G. M. Wynne; *J. Med. Chem.* **2011**, *54*, 3241-3250 *(10.1021/jm200135z)*.

[50] E. Falb, A. Nudelman, A. Hassner; *Synth. Commun.* **1993**, *23*, 2839-2844 *(10.1080/00397919308012605)*.

[51] S. Celanire, L. Quere, F. Denonne, L. Provins; *Compounds comprising a lactam or a lactam derivative moiety, processes for making them, and their uses (WO2007048595)*, **2007**.

[52] J.-M. Hwang, S.-H. Yeom, K.-Y. Jung; *J. Ind. Eng. Chem.* **2007**, *13*, 474-479.

[53] D.-R. Hou, J. H. Reibenspies, K. Burgess; *J. Org. Chem.* **2001**, *66*, 206-215 *(10.1021/jo001333h)*.

[54] E. Fischer, L. Ach; *Ber. D. Ch. Ges.* **1895**, *28*, 3135-3143 *(10.1002/cber.189502803156)*.

[55] T. Maekawa, N. Sakai, H. Tawada, K. Murase, M. Hazama, Y. Sugiyama, Y. Momose; *Chem. Pharm. Bull.* **2003**, *51*, 565-573 *(10.1248/cpb.51.565)*.

[56] E. Ferrer Flegeau, M. E. Popkin, M. F. Greaney; *Org. Lett.* **2006**, *8*, 2495-2498 *(10.1021/ol060591j)*.

[57] S. A. Nunez, K. Yeung, N. S. Fox, S. T. Phillips; *J. Org. Chem.* **2011**, *76*, 10099-10113 *(10.1021/jo2018763)*.

[58] P. D. Senter, S. Doronina, B. Toki; *Drug Conjugates and their use for treating cancer, an autoimmune disease or an infectious disease (WO2004010957)*, **2004**.

[59] S. V. Jadhav, A. Bandyopadhyay, S. N. Benke, S. M. Mali, H. N. Gopi; *Org. Biomol. Chem.* **2011**, *9*, 4182-4187 *(10.1039/C0OB01226B)*.

[60] K. Maekawa, K. Kubo, T. Igarashi, T. Sakurai; *Tetrahedron* **2005**, *61*, 11211-11224 *(10.1016/j.tet.2005.09.042)*.

[61] M. Seki, K. Matsumoto; *Synthesis* **1999**, *1999*, 924-926 *(10.1055/s-1999-3495)*.

[62] Y. Kaburagi, Y. Kishi; *Org. Lett.* **2007**, *9*, 723-726 *(10.1021/ol063113h)*.

Appendices

Appendices

Publications from this PhD

Part I: A patent (J. Hasserodt, O. Thorn-Seshold; *Substrat de Peptidase Fluorogene* [FR1158732/ WO2013045854], **2011**) and a paper (O. Thorn-Seshold, M. Vargas-Sanchez, S. McKeon and J. Hasserodt; *Chem. Commun.* **2012**, *48*, 6253-6255) have been published based on the work in Part I.

The Chem. Commun. paper is reproduced on the following pages; the full patent text is now in the public domain (as of April 2013) but is not reproduced here.

A paper including some of my further synthetic/enzymatic results with the probes of Part I, as well as much original work by Maxime Prost who took over the fluorogenic probes project in 2011, is planned for preparation beginning in June 2013 (M. Prost, O. Thorn-Seshold, L. Canaple, J. Hasserodt *et al.*). The combretastatin prodrug scheduled for *in cellulo* tests in June should enter into a separate medicinal chemistry publication. A final paper communicating some of the results of the modelling experiments of Part I may be prepared separately at a later date, for a general concept paper.

Part II: Further experiments are being performed (May-July 2013) to conclude on the feasibility (↔aqueous stability) of the furyl and especially carbimidyl systems. If the systems are deemed too unstable for use as spacers in off-ON probes, then they will be published separately for their synthetic chemistry content. However, if they allow off-ON MRI probes in aqueous media, their (later) publication as such can hope to command more attention.

oliver.torvald.thorn-seshold@ens-lyon.fr; ☎ +33 47272 8396; website.

Oliver Thorn-Seshold, M.Sc

Studies at: Ecole Normale Supérieure de Lyon, France (ENS-Lyon); University of Sydney, Australia (USYD); Göteborgs Universitätet, Sweden (GU).

Date of Birth: 5 June 1985, Sydney. ***Citizenships:*** Australian, Swedish, British.

Summary

2009-present: PhD, Ecole Normale Superieure de Lyon (France); Supervisor: Prof. J. Hasserodt. PhD in organic synthesis and chemical biology: *Development of Auto-Immolative Spacers for Probes of Enzyme Activity.*

2008: Master 2; ENS-Lyon: Graduated placed 5th. Internship in synthesis for bio-organic chemistry.

2007: Honours; University of Sydney (Australia): Edna Maude Goulston prize in Organic Chemistry. Research under Prof. R. Gilbert (physical chemistry) and A. Prof. A. Gray-Weale (theoretical chemistry).

2004-2006: Bachelor; USYD and Göteborgs Universitätet (Sweden): Double major in chemistry. Exchange to GU for theoretical chemistry and optics (2006). Walter Burfitt Chemistry Scholarship.

2002–2003: International Chemistry Olympiads: Member of the Australian teams at the 34th and 35th International Olympiads, winning Bronze (111th/228; 2002) and Silver (34th/232; 2003) medals.

1998–2003: Sydney Grammar School: After-hours research with Dr Michael Bishop in inorganic chemistry (2000-2002) leading to first coauthorships. Australian Student Prizes (2002 and 2003).

Publications / Patent

[6] O. Thorn-Seshold, M. Vargas-Sanchez, S. McKeon, J. Hasserodt; *Chem. Comm.* **2012**, 48, 6253: *A Robust, High-Sensitivity Stealth Probe for Peptidases*.

[5] J. Hasserodt, O. Thorn-Seshold, *Fluorogenic Peptidase Substrate* (French Patent FR1158732), **2011**.

[4] D. Konkolewicz, O. Thorn-Seshold and A. Gray-Weale; *J. Chem. Phys.* **2008**, 129, 054901.

[3] M. M. Bishop, L. Lindoy, M. McPartlin, A. Parkin, O. Thorn-Seshold, P. Turner; *Polyhedron* **2007**, 26, 415.

[2] M. M. Bishop, S. J. Coles, A. H. W. Lee, L. F. Lindoy, A. Parkin, O. Thorn-Seshold, P. Turner; *Supramolecular Chemistry* **2005**, 17, (7), 567.

[1] M. M. Bishop, L. F. Lindoy, O. Thorn-Seshold, R. O. Piltz, P. Turner; *J. Het. Chem.* **2001**, 38, (6), 1377.

Fellowships and Awards

French Regional Academic Research Council grant ARC1-Santé, and ENS-Lyon Research Support Fund grant (2013), for my independant-initiative project in photo-controlled chemotherapy, partnering with groups Trauner (Munich), Jalinot (Lyon), Vollmar (Munich), and Coll (Grenoble).

French Ministry of Research PhD Scholarship with Teaching Fellowship (2009–2012). Ampere Masters Scholarship (2008).

Edna Maude Goulston prize for Organic Chemistry; Honours thesis graded High Distinction (2007). University of Sydney Scholarship with Merit. (2004-2007).

Walter Burfitt Chemistry Scholarship N° 1 and Hush Prize (2006; USYD prizes for general and theoretical 3rd year chemistry, respectively). Chemistry Exchange Scholarship (2006). Royal Australian Chemical Institute's Student Chemistry Prize (2006). Physics Scholarship N° 1 (2004).

Silver medal (34th/232) at the 35th International Chemistry Olympiad (Athens, 2003). Bronze medal (111th/228) at the 34th IChO, (Groningen, 2002). Headmaster's Exhibition philosophy prize (SGS, 2003). Australian Student Prizes (Higher School Certificate examination 2003, and IChO 2002).

Skills

Scientific: Solid experience in many methods and technologies of chemistry, particularly organic synthesis. Particular experience with heterocycles, polyamines, and substrate-drug conjugation strategies. At ease with analysis and statistics, model building and programming (FORTRAN).

Research Methods: Highly independent, innovative and motivated approach to research. Already have experience in devising projects, assembling groups of collaborators across borders, writing and obtaining grants (eg. Academic Research Council and ENS Research Fund grants in 2013). Successful in sustained, high-intensity, complex research. Effective in teams, motivating group efforts, and leading junior researchers.

Actively engaged in transparent discussion with all group members, regardless of seniority, to achieve better research across all projects.

Concerned to minimise the cost, waste and environmental impact of syntheses. Aiming for routes that are *simple and modular*, for easy application outside ideal settings (*eg* in developing countries).

Languages: Native English speaker with excellent quality of written expression. Fluent in French and Swedish; beginner level in German. Broad cultural experience. At ease learning foreign languages (also studied Latin, Italian, Old Norse, Old English).

Enabling Technology: Mastery of Windows and Mac plus their usual programs. Experienced with Unix systems. At ease in x/html (created the trilingual Hasserodt Group website in 2010).

Career Objective

I wish to design simple and modular chemical methods for the conception of bio-compatible smart probes, and especially for targeted/bio-activated prodrugs. I expect multimodal, bio-orthogonal, cascade/autocatalytic, and multiply-gated strategies will favour robust *in vivo* applications to real-world problems. I aim to lead my own chemical biology research group in academia.

Other Academic and Teaching Experience

2009–2011: Teaching Fellowship at the ENS-Lyon. Responsible for the practical course in 3rd year organic chemistry (48 h / yr); assistant for the Agrégation Teaching Diploma course (20 h / yr). Supervisor of students Nils Aronssohn (3rd year, 2011), Claire Weisslinger (3rd year, 2010) and Andreas Brunner (2nd year of Masters, 2009) during their internships.

2007: Writing and delivery of a short course of unusual topics in the "Special Studies" chemistry group (the top students in 1st year chemistry at USYD); practical supervisior in 1st year chemistry.

2004–2007: Tutor in 1st-3rd year chemistry, and 1st-2nd year biochemistry, physics and maths at USYD. Tutor for IChO preparation at SGS. Internships in organic synthesis with Prof. Mal McLeod (2005) and Prof. Mark Coster (2004), and with Prof. Tim Schmidt in organic / analytical chemistry (2006).

2002–2003: Set up and ran an interest society for Viking culture and Old Norse language.

Other interests

Chamber music: cellist in the Chestnut Quartet and Paryon piano trio, performing in Strasbourg (2012) and in the Milko Kelemen festival (Croatia, 2010), as well as locally several times a year.

Sports: hiking, marathon/trail/raid, road cycling.

Referees / Reference Letters

[1] Prof. Jens Hasserodt, PhD/Masters supervisor (2008-2012), ENS-Lyon: jens.hasserodt@ens-lyon.fr

[2] Prof. Bob Gilbert, Honours supervisor (2007), University of Queensland: b.gilbert@uq.edu.au

[3] Prof. Emeritus Leif Holmlid and Prof. Jürgen Gräfenstein, Exchange supervisors (2006), GU.

cette page a été laissée vierge intentionellement

Bornons ici cette carrière :
Les longs ouvrages me font peur.
Loin d'épuiser une matière,
On n'en doit prendre que la fleur.

Il serait temps que je reprenne
Un peu de forces et d'haleine
Pour vaquer à d'autres projets :
Chimie, ce tyran de ma vie,
Veut que je change de sujets ;
Il faut contenter son envie.

Partons pour la Lumière ; partons envers le soufre ;
Les vents voudront qu'en Allemagne je m'engouffre.
J'y consens : sans doute ma veine
En sa faveur s'échauffera.
Heureux si ce travail n'est point la dernière peine
Que CBO me causera !

clin d'oeil à La Fontaine

cette page a été laissée vierge intentionellement

Développement d'espaceurs auto-effondrables pour des sondes d'activité enzymatique

Cette thèse traite de la conception et mise en œuvre d'espaceurs auto-effondrables novateurs pour une utilisation dans des sondes d'activité enzymatique *in vivo*.

La première partie détaille la synthèse et la validation *in vitro* d'espaceurs cyclisant, couplant l'activité d'un aminopeptidase à la libération d'un phénol. Les sondes fluorogènes modulaires basées sur ces espaceurs 1,2-diamine sont très robustes (demi-vie > 560 h), mais sont rapidement enzymatiquement hydrolysées, et puis relâchent rapidement (demi-temps ~ 3 min) un fluorophore ESIPT insoluble et exceptionnellement photostable.

Ces sondes ont une excellente sensibilité (rapport signal:contrôle > 3000:1), et fournissent la première démonstration d'un système macroscopiquement binaire *éteint–ALLUMÉ*, pour la libération de phénols sous activité d'aminopeptidases. Ce système pourrait permettre de faire de l'imagerie moléculaire ultra-sensible d'une gamme d'exopeptidases. Ces espaceurs pourraient également servir dans des sondes comportant d'autres fluorophores phénoliques, dans des promédicaments de phénols/alcools activés par des peptidases spécifiques (thérapies ciblées), ou comme adaptateurs chimiques en générale.

La deuxième partie détaille les synthèses de deux familles d'espaceurs tautomérisant/éliminant pour utilisation dans des sondes magnétogènes d'activité de glycosidases. Les premières architectures substrat-espaceur basées sur des 2-furanols et des carbimidates cycliques ont été explorées. Notamment, des glycosides 2-furanoliques ont été abordés comme espaceurs énergétiques alternatifs aux quinone méthydes, et des carbimidates ont été explorés comme espaceurs pour ligands modèles des sondes promagnétiques.

Development of Auto-Immolative Spacers for Probes of Enzyme Activity

This thesis concerns the design and implementation of novel auto-immolative spacers for use in probes for enzymatic activity *in vivo*.

The first part relates the development and *in vitro* validation of cyclisation spacers which couple the action of an aminopeptidase to the release of a phenol. The modular three-component fluorogenic probes based on these 1,2-diamine spacers are very robust (halflife > 560 h), but are also rapidly enzymatically processed, and quickly (halftime ~3 min) release an exceptionally photostable, insoluble ESIPT fluorophore.

The probes have excellent detection sensitivity relative to current methods (signal to control ratio > 3000:1), and provide the first demonstration of a macroscopically binary *off–ON* system for phenol-releasing probes of aminopeptidase activity. The probe system may allow the exceptionally sensitive, ESIPT-based molecular imaging of a range of exopeptidases. The spacers may also be applied in off-ON peptidase probes of other phenolic fluorophores, to peptidase-specific phenol/alcohol prodrugs for targeted therapy, or more generally in chemical adapter technologies.

In the second part, two novel families of auto-immolative elimination/tautomerisation spacers were designed for use in three-component off-ON magnetogenic probes sensing glycosidase activity. The first known substrate-spacer designs based on 2-hydroxyfurans and on carbimidates were explored. Notably, 2-furanol glycosides were synthesised in pursuit of high-energy alternatives to quinone methides, and a general method for preparing model carbimidate-bearing ligands for pro-magnetic probes was elaborated.

Oliver Thorn-Seshold, 05-VI-2013