Controlling pathways and kinetics in fueled chemical reaction networks

Anastasiia Sharko

UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE 222

UMR 7140 Chimie de la Matière Complexe

THÈSE présentée par :

Anastasiia SHARKO

soutenue le : **18 Novembre 2022**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Chimie

Controlling pathways and kinetics in fueled Chemical Reaction Networks

THÈSE dirigée par :

M HERMANS Thomas	Professeur, UMR 7140, Université de Strasbourg, France

RAPPORTEURS :

M ESTEVEZ-TORRES André	Directeur de recherches, UMR 8237, CNRS and Sorbonne Université, France
Mme KATSONIS Nathalie	Professeur, the Stratingh Institute of Chemistry of the University of Groningen, the Netherlands

AUTRES MEMBRES DU JURY :

M FALLER Peter	Professeur, UMR 7177, Université de Strasbourg, France

Table of Contents

Résumé

Cette thèse, intitulée "Contrôle des chemins et de la cinétique dans les réseaux de réaction chimique alimentés en fuel" est consacrée à l'exploration pratique des réseaux de réactions chimiques et aux facteurs qui influencent leur évolution.

Inspiré par la nature, le nouveau domaine de la chimie systémique étudie les systèmes complexes avec des caractéristiques supplémentaires comme le comportement dynamique, le mouvement moléculaire contrôlé chimiquement, les réseaux de réaction chimique dans les compartiments, l'auto-réplication, etc.[1] Ces systèmes fonctionnent dans des conditions de non-équilibre, ce qui signifie que le système reçoit un apport constant d'énergie et/ou de matière par l'ajout de réactifs à haute énergie ("fuels") ou par l'apport de lumière. Lorsque l'énergie se dissipe, ces systèmes présentent des fonctions émergentes et des propriétés structurelles/comportementales qui peuvent impliquer un niveau supplémentaire d'organisation et l'exécution d'un travail.[2] Il est important de noter qu'un tel état ne peut être maintenu que si le système dispose d'un fuel. Les réseaux de réaction chimique (RRC) sont un ensemble de réactions interdépendantes qui transforment les réactifs en produits de manière progressive. Les RRC sont omniprésents dans les systèmes biochimiques et remplissent diverses fonctions biologiques complexes telles que la signalisation, la synthèse des protéines et l'homéostasie. Si les RRC fonctionnent très loin de l'équilibre, ils peuvent donner lieu à des instabilités qui se traduisent par des comportements émergents complexes tels que les oscillations, la bistabilité, l'ultrasensibilité, l'amplification et la formation de motifs.[3]

Des approches mathématiques sont développées pour étudier la théorie des RRC et notamment la cinétique de tels systèmes. Les modèles décrivent la complexité du réseau, sa vitesse d'évolution, ses rendements, ses flux, etc. En général, ils ne tiennent pas compte de facteurs de complication supplémentaires comme l'organisation spatio-temporelle des éléments du système ou la diffusion, car les expressions deviennent plus difficiles à résoudre. Dans ce contexte, l'auto-assemblage est l'un des moyens d'isoler les composants d'un système complexe. Il peut conduire à une séparation temporelle de phase de certains des réactifs, ce qui peut les protéger des autres réactifs ou, au contraire, accélérer les processus grâce à l'effet multivalent ou à la catalyse coopérative. L'influence de tels événements sur la cinétique du système entier est souvent imprévisible et n'est pas triviale à décrire, même avec des études expérimentales supplémentaires. L'existence de modèles avec des hypothèses simplificatrices appropriées aide à développer l'intuition dans la compréhension des tendances dans les systèmes complexes ; cependant, ils sont encore loin d'être une représentation exacte de la réalité.

Dans le **Chapitre 1**, nous abordons nos travaux récemment publiés,[4] dans lesquels la polymérisation supramoléculaire se produit dès que de l'énergie (fuel) est fournie et est, de cette façon, le système maintenue hors équilibre. Nous examinons de près les changements de longueur et de distribution des polymères dans de telles conditions et recherchons les oscillations de ces paramètres. Ce chapitre passe en revue le comportement complexe des polymères supramoléculaires dans le régime de non-équilibre dissipatif et souligne l'importance des études cinétiques dans les systèmes complexes qui sont essentielles pour évoluer vers des systèmes encore plus sophistiqués.

Dans les **Chapitres 2 à 4**, nous utilisons une approche ascendante pour construire un réseau de réactions alimenté chimiquement qui contrôle la formation et l'auto-assemblage de peptides cycliques en nanotubes et leur durée de vie.

Les peptides cycliques (PC) sont largement présents dans la nature et présentent un intérêt pour les systèmes synthétiques en raison de leur structure unique. Les PC composés d'acides aminés D et L alternés sont planaires et s'empilent en structures tubulaires - nanotubes peptidiques cycliques (NPC).[5] Les NPC sont très robustes et stables grâce aux liaisons H coopératives, aux interactions électrostatiques et hydrophobes qui sont favorables entre les macrocycles planaires. L'accessibilité d'une large gamme d'acides aminés L/D naturels et synthétiques et les différentes manières d'obtenir des NPC ont conduit à l'application des nanotubes dans l'administration de médicaments, les dispositifs électroniques, les antimicrobiens, les canaux transmembranaires et autres.[6] Ces assemblages peptidiques tubulaires - qui forment des feuillets β antiparallèles par association anneau-anneau, souvent à l'aide d'un mécanisme d'auto-assemblage coopératif - comptent parmi les assemblages les plus stables et les plus robustes connus et sont notoirement difficiles à désassembler. Le développement d'un cycle de réaction peut conduire à la formation et à la dégradation sur demande des NPC et, par conséquent, faciliter leur application en tant que matériaux et produits pharmaceutiques.

La chimie choisie à cet effet est basée sur le couplage réversible d'amines et de thiols de fragments peptidiques montrée précédemment par Anslyn et ses collaborateurs.[7] Ils ont démontré l'utilisation d'un dérivé de l'acide de Meldrum (DAM) pour former une liaison covalente avec une amine primaire et un thiol en donnant un rendement quantitatif dans des conditions aqueuses à température ambiante. Ensuite, ils ont récupéré les substrats originaux en utilisant un autre composé nucléophile contenant un thiol et une amine libres ou deux thiols tels que la cystéine ou le dithiothréitol (DTT) pour éliminer le DAM des peptides, des réseaux de polymères et des petites molécules.[8,9] Avant de combiner ces deux réactions en un cycle, nous étudions leur cinétique en utilisant nos peptides et le processus d'auto-assemblage des PCs résultants. Nous avons conçu deux peptides : i) **P1** (Val_D-Phe-Arg_D-Tyr-Val_D-Trp-Arg_D-Cys-Gly), ii) et **P2** (Glu_D-Trp-Val_D-Cys-Arg_D). Les acides aminés hydrophobes (Val, Phe, Trp, Tyr) et hydrophiles (Arg, Glu) ont été choisis pour équilibrer les interactions entre les molécules de PC et faciliter la caractérisation instrumentale du mélange réactionnel et du produit.

Le **Chapitre 2** est entièrement consacré à la discussion de la cinétique de cyclisation et du mécanisme d'auto-assemblage. Nous utilisons les données CL-MSHR, DC et UV-Vis pour reconstruire le profil cinétique et l'ajuster ensuite en utilisant l'intégration numérique des équations différentielles. L'UV-Vis est également utilisé pour étudier le mécanisme de polymérisation supramoléculaire des NPC. Une fois la cyclisation terminée, la morphologie des NPC a été étudiée à l'aide de l'AFM, du SEM et de la microscopie confocale.

Dans le **Chapitre 3**, le processus de décyclisation est étudié sur la réaction de cyclisation brute et différents agents de décyclisation sont discutés. Un point difficile pour atteindre la réversibilité de la formation des NPC est de maitriser les taux de construction (c'est-à-dire la cyclisation et l'assemblage) et de déconstruction (c'est-à-dire la décyclisation et le désassemblage) du NPC au sein du système.[10] La réaction de décyclisation est basée sur la réaction d'un agent de décyclisation (AD) - par exemple, une molécule contenant un dithiol ou une amine-thiol, comme le 1,4-dithiothréitol (DTT) - avec la partie DAM de la molécule de PC. AD réagit également avec le DAM lui-même en tant que réaction secondaire. Les réactions

permettant la construction et la déconstruction lors du cycle de réaction sont toutes les deux dépendantes du pH de la même manière : l'augmentation du pH accélère les deux processus, la réaction DAM-AD étant toujours 2 à 3 fois plus rapide que la réaction DAM-peptide. Ce dernier point est problématique lorsqu'on essaie d'obtenir un rendement élevé de PC. Ainsi, nous concluons qu'il n'est pas possible de construire un cycle de réaction alimenté chimiquement avec des NPC transitoires en combinant simplement des peptides, DAM et AD dans un pot.

Une autre approche serait d'utiliser un "pré-fuel".[11] Ce composé est inerte vis-à-vis de la réaction de cyclisation et libérerait lentement le DA dans le système via un réseau de réaction indépendant. Cette approche est discutée dans le chapitre 4, où une façon de fermer le cycle de réaction et d'obtenir des NPC transitoires est présentée. Comme pré-fuel possible, nous avons étudié les échanges thiol-disulfure et thiol-thioester, les groupes de protection des amines et thiols labiles (imines, Fmoc, sels d'isothiouronium, thiolactones, thiazolidine), les réactifs de réduction du disulfure (TCEP, THPP, autres dérivés de la phosphine avec une moindre nucléophilie du phosphore, $NaBH_4$, dithionate de sodium), et la chimie de coordination des complexes labiles des métaux de transition avec des ligands contenant un thiol. Il ressort de cette étude que notre meilleur choix actuel de pré-fuel est un dérivé de la phosphine, AdaP, grâce auquel nous avons obtenu une formation retardée de DTT in situ à partir de son analogue oxydé DTTox. En combinant le DTTox, l'AdaP, le peptide linéaire et le DAM dans un rapport spécifique, nous avons obtenu des PCs transitoires avec une durée de vie variable de 30 à 80 heures.

D'autres études cinétiques sont nécessaires pour évaluer l'utilisation potentielle de ce pré-fuel dans le cycle de réaction, en réduisant au minimum les réactions secondaires et en maximisant le rendement des PCs. Notre travail montre une voie vers des PC transitoires hors équilibre, qui pourraient être utilisés dans des biomatériaux ou des réseaux de réaction plus complexes comme éléments constitutifs de matériaux synthétiques semblables à la vie.

Dans le **Chapitre 5**, nous étudions un réseau réactionnel plus large d'interconversion d'espèces non-assemblées. Les accepteurs de Michael (AM) ont été abondamment appliqués dans divers domaines, de la chimie des polymères aux bioconjugaisons. Une attention accrue a été portée sur les méthacrylates β-substitués pour leur capacité à réaliser une chimie de type trigger-to-release et trigger-to-reverse.[12] Ces processus visent la libération et la récupération contrôlées de nucléophiles plus faibles comme les amines, alors que les AM électrophiles sont consommés de manière irréversible une fois qu'ils ont réagi avec des nucléophiles plus forts comme les thiols. Notre objectif est de combler cette lacune et de développer un système qui permet la récupération de la fraction AM. Nous montrons que les AM à fonction sulfure peuvent être transformés en leurs analogues sulfoxyde et sulfone en utilisant des réactions d'oxydation. Ces espèces plus déficientes en électrons peuvent ensuite être attaquées par des amines, récupérant ainsi l'amine-AM initiale. De plus, nous montrons qu'en fonction de la force de l'oxydant employé, ce réseau de réaction peut se dérouler selon trois voies distinctes. Nous envisageons un large éventail d'applications potentielles pour cette nouvelle chimie, allant de la fonctionnalisation des protéines aux systèmes de libération réversible de cargaison en passant par les matériaux polymères transitoires.

Le **Chapitre 6** résume les RRC étudiés et, en guise de perspective, donne une idée de la manière dont les avantages des deux réseaux peuvent être utilisés plus avant.

Cette thèse montre différentes approches pour créer et étudier les réseaux de réaction chimique. Dans les premiers chapitres, nous explorons l'auto-assemblage de peptides cycliques en tant que partie du réseau afin d'obtenir la formation transitoire de RRC et de régler leur rendement et leur durée de vie. Dans le dernier chapitre expérimental, nous nous concentrons sur l'étude des différentes voies du réseau de réaction et sur la manière de les contrôler en utilisant différentes forces d'oxydation. Les études présentées sont essentielles dans le contexte des RRC programmables et biocompatibles qui peuvent facilement être interfacés avec d'autres processus biologiques et d'autres matériaux. Les connaissances acquises peuvent être utilisées pour concevoir des (bio)matériaux nouveaux et améliorés dotés de fonctionnalités sans précédent, tels que des tissus autonomes capables de leur propre métabolisme (limité) et entièrement compatibles avec les organismes humains/animaux, des systèmes de vie synthétiques capables d'interagir avec les systèmes naturels, ou des stations nanochimiques efficaces et sans déchets qui peuvent synthétiser localement des médicaments dans les organes endommagés et "réparer" les bugs dans les cellules défectueuses.[13] Une autre perspective très souhaitable à long terme de nos études est d'obtenir un tel degré de contrôle qu'il pourrait permettre de diriger les systèmes complexes vers des structures et des fonctions souhaitées parmi la quasi-totalité de celles qui sont possibles en dehors de l'équilibre.

Summary

This thesis, entitled "Controlling pathways and kinetics in fueled Chemical Reaction Networks" is dedicated to the practical exploration of chemical reaction networks and the factors that influence their evolution.

Inspired by nature, a newly emerged field of Systems Chemistry studies complex systems with additional features like dynamic behavior, chemically controlled molecular motion, chemical reaction networks within compartments, self-replication, etc.[1] These systems operate under nonequilibrium conditions, which means that the system is provided with a constant supply of energy and/or matter by adding high-energy reactants ('fuels') or shining light. Upon energy dissipation, such systems demonstrate emergent functions and structural/behavioral properties that can involve an additional level of organization and the performance of work.[2] Importantly, such a state can only be maintained while the fuel is available to the system. Chemical reaction networks (CRNs) are a set of interrelated reactions that convert reactants into products in a stepwise manner. CRNs are omnipresent in biochemical systems and serve various complex biological functions such as signaling, protein synthesis, and homeostasis. If CRNs operate considerably far from equilibrium, they can give rise to instabilities that result in complex emergent behavior such as oscillations, bistability, ultrasensitivity, amplification, and pattern formation.[3]

A CRNs theory develops mathematical approaches to study the kinetics of such systems. The models describe network complexity, its evolution rate, yields, fluxes, etc. Usually, it does not account for additional complicating factors like the spatiotemporal organization of the system's elements or diffusion because the expressions become more difficult to solve. Self-assembly is one of the ways to isolate components of a complex system. It can lead to temporal phase separation of some of the reactants, which can shield them from other reactants or, on the contrary, speed up the processes due to the multivalent effect or cooperative catalysis. The influence of such events on the kinetics of the whole system is often unpredictable and is not trivial to describe, even with additional experimental studies. The existence of models with appropriate simplifying assumptions aids in developing intuition in understanding trends in complex systems; however, they are still far from an accurate depiction of reality.

In **Chapter 1**, we discuss our recently published work,[4] where supramolecular polymerization occurs once energy (fuel) is supplied and is, in this way, maintained out of equilibrium. We are looking closely at the changes in polymer length and distribution under such conditions and searching for the oscillations of these parameters. The chapter reviews the complex behavior of supramolecular polymers in the dissipative non-equilibrium regime and highlights the importance of kinetics studies in complex systems that are essential for moving towards even more sophisticated systems.

In **Chapters 2–4** we use a bottom-up approach to build a chemically fueled reaction network that controls the formation and self-assembly of cyclic peptides into cyclic peptide nanotubes and their lifetime.

Cyclic peptides (CPs) are widely present in nature and are of interest in synthetic systems because of their unique structure. CPs composed of alternating D- and L-amino acids are planar and stack into tubular structures—cyclic peptide nanotubes (CPNs).[5] CPNs are highly robust and stable due to cooperative H-bonds, electrostatic, and hydrophobic interactions that are

favorable between the planar rings. The accessibility of a wide range of natural and synthetic D/L-amino acids and different ways to obtain CPNs has led to applications of nanotubes in drug delivery, electronic devices, antimicrobials, transmembrane channels, and others.[6] These tubular peptide assemblies—shown to form antiparallel β-sheets by ring-ring association, often using a cooperative self-assembly mechanism—are some of the most stable and robust assemblies known and are notoriously difficult to disassemble. Developing a reaction cycle can lead to the on-demand formation and degradation of CPNs and, therefore, facilitate their application as materials and pharmaceuticals.

The chemistry chosen for this purpose is based on the reversible ligation of amines and thiols of peptide fragments shown earlier by Anslyn and co-workers.[7] They have demonstrated the use of a Meldrum acid derivative (MAD) to form a covalent bond with a primary amine and a thiol giving a quantitative yield in aqueous conditions at room temperature. Next, they recovered original substrates by using another nucleophilic compound containing a free thiol and amine or two thiols such as cysteine or dithiothreitol (DTT) to scavenge MAD from peptides, polymer networks, and small molecules.[8,9] Before combining these two reactions into a cycle, we study their kinetics using our peptides and the self-assembly process of the resulting CPs. We designed two peptides: i) **P1** (Val_D-Phe-Arg_D-Tyr-Val_D-Trp-Arg_D-Cys-Gly), ii) and **P2** (Glu_D-Trp-Val_D-Cys-Arg_D). Hydrophobic (Val, Phe, Trp, Tyr) and hydrophilic amino acids (Arg, Glu) were chosen to balance interactions between CP molecules and facilitate instrumental characterization of the reaction mixture and the product.

Chapter 2 is entirely devoted to discussing cyclization kinetics and self-assembly mechanism. We use LC-HRMS, CD, and UV-Vis data to reconstruct the kinetic profile and then fit it using numerical integration of differential equations. UV-Vis is also used to study the supramolecular polymerization mechanism of CPNs. After completion of cyclization, the morphology of CPNs was studied using AFM, SEM, and confocal microscopy.

In **Chapter 3**, the decyclization process is studied on the crude cyclization reaction and different decyclization agents are discussed. A challenging point to achieve the reversibility of CPNs formation is tuning the forward (i.e., cyclization and assembly) and backward (i.e., decyclization and disassembly) rates.[10] The decyclization reaction is based on the reaction of a decyclizing agent (DA)—for example, a dithiol or amine-thiol containing molecule, like 1,4-dithiothreitol (DTT)—with the MAD part in the CP molecule. The DA also reacts with the forward fuel MAD itself as a side reaction. Both forward and backward reactions in the reaction cycle are pH-dependent in the same way: increasing pH speeds up both processes, where the MAD–DA reaction is always 2–3 times faster than MAD–peptide. The latter is problematic when trying to obtain a high yield of CP. We conclude, that it is impossible to build a chemically fueled reaction cycle with transient CPNs by simply combinig peptides, MAD and DA in one pot.

Another approach would be to use a 'pre-fuel.'[11] This compound is inert towards the cyclization reaction and would slowly liberate DA into the system via an independent reaction network. This approach is discussed in **Chapter 4**, where a way to close the reaction cycle and achieve transient CPNs is presented. As a possible pre-fuel, we have been looking into thiol-disulfide and thiol-thioester exchange, labile amine, and thiol protection groups (imines, Fmoc, isothiouronium salts, thiolactones, thiazolidine), disulfide reduction reactants (TCEP, THPP, other phosphine derivatives with lower nucleophilicity of phosphorous, $NaBH_4$, sodium

dithionate), and coordination chemistry of labile complexes of transition metals with thiol-containing ligands. Our current best choice of the pre-fuel is a phosphine derivative AdaP, by which we achieved a delayed formation of DTT in situ from its oxidized analog DTT_{ox}. Combining DTT_{ox}, AdaP, Linear peptide, and MAD at a specific ratio, we further obtained transient CPs with a variable lifetime of 30–80 hours.

Further kinetic studies are needed to evaluate the potential use of this pre-fuel in the reaction cycle, keeping side reactions to a minimum and maximizing the yield of CPs. Our work shows a pathway toward out-of-equilibrium transient CPs, which could be used in biomaterials or more complex reaction networks as building blocks of synthetic life-like materials.

In **Chapter 5**, we study a broader reaction network of interconverting non-assembling species. Michael acceptors (MAs) have been abundantly applied in various fields, from polymer chemistry to bioconjugations. Increased attention has been focused on β-substituted methacrylates for their ability to perform trigger-to-release and trigger-to-reverse chemistry.[12] These processes target the controlled release and recovery of weaker nucleophiles like amines, whereas electrophilic MAs are consumed irreversibly once reacted with stronger nucleophiles like thiols. We aim to close this gap and develop a system that enables the recovery of the MA moiety. We show that sulfide-functionalized MAs can be transformed into their sulfoxide and sulfone analogs by employing oxidation chemistry. These more electron-deficient species can then be attacked by amines, recovering the initial amine-MA. Furthermore, we show that depending on the strength of the employed oxidant, this reaction network can proceed via three distinct pathways. We envision a wide range of potential applications for this novel chemistry, ranging from protein functionalization to reversible cargo release systems to transient polymer materials.

Chapter 6 summarizes studied CRNs and, as an outlook, shares an idea of how the advantages of both networks can be used further.

This thesis shows different approaches to creating and studying chemical reaction networks. In the first chapters, we explore the self-assembly of cyclic peptides as part of the network aiming to achieve the transient formation of CPNs and tune their yield and lifetime. In the final experimental chapter, we focus on studying different pathways of the reaction network and how to control them using various oxidation strengths. The presented studies are essential in the context of programmable and biocompatible CRNs that can easily be interfaced with other biological processes and other materials. The gained insights can be used in designing novel and improved (bio)materials with unprecedented functionalities, such as self-sustained tissues capable of their own (limited) metabolism and fully compatible with human/animal organisms, synthetic life systems that are able to interact with the natural systems, or efficient waste-less nanochemical stations that can locally synthesize drugs in the damaged organs and ‘fix’ bugs in the malfunctioning cells.[13] Another very desirable long-term perspective of our studies is to gain a such degree of control that may allow one to steer the complex systems towards desired structures and functions out of almost infinite ones that are possible out of equilibrium.

References

1. Ashkenasy, G., Hermans, T. M., Otto, S. & Taylor, A. F. Systems chemistry. *Chem. Soc. Rev.* **46**, 2543–2554 (2017).
2. Ragazzon, G. & Prins, L. J. Energy consumption in chemical fuel-driven self-assembly. *Nature Nanotech* **13**, 882–889 (2018).
3. Epstein, I. R. & Pojman, J. A. *An Introduction to Nonlinear Chemical Dynamics: Oscillations, Waves, Patterns, and Chaos*. (Oxford University Press, 1998).
4. Sharko, A., Livitz, D., Piccoli, S. D., Bishop, K. & Hermans, T. Insights into chemically-fueled supramolecular polymers. (2022) doi:10.26434/chemrxiv-2022-jd607.
5. Ghadiri, M. R., Granja, J. R., Milligan, R. A., McRee, D. E. & Khazanovich, N. Self-assembling organic nanotubes based on a cyclic peptide architecture. *Nature* **366**, 324–327 (1993).
6. Hsieh, W.-H. & Liaw, J. Applications of cyclic peptide nanotubes (cPNTs). *J Food Drug Anal* **27**, 32–47 (2019).
7. Diehl, K. L. *et al.* Click and chemically triggered declick reactions through reversible amine and thiol coupling via a conjugate acceptor. *Nat. Chem.* **8**, 968–973 (2016).
8. Johnson, A. M. & Anslyn, E. V. Reversible Macrocyclization of Peptides with a Conjugate Acceptor. *Org. Lett.* **19**, 1654–1657 (2017).
9. Sun, X. *et al.* Chemically Triggered Synthesis, Remodeling, and Degradation of Soft Materials. *J. Am. Chem. Soc.* **142**, 3913–3922 (2020).
10. Wang, G. & Liu, S. Strategies to Construct a Chemical-Fuel-Driven Self-Assembly. *ChemSystemsChem* **2**, e1900046 (2020).
11. Singh, N., Lainer, B., Formon, G. J. M., De Piccoli, S. & Hermans, T. M. Re-programming Hydrogel Properties Using a Fuel-Driven Reaction Cycle. *J. Am. Chem. Soc.* **142**, 4083–4087 (2020).
12. Zhuang, J. *et al.* A programmable chemical switch based on triggerable Michael acceptors. *Chem. Sci.* **11**, 2103–2111 (2020).
13. Das, K., Gabrielli, L. & Prins, L. J. Chemically Fueled Self-Assembly in Biology and Chemistry. *Angew. Chem.* **133**, 20280–20303 (2021).

List of abbreviations

AFM	Atomic force microscopy
ACN	Acetonitrile
BME	β-mercaptoethanol
CD	Circular dichroism spectroscopy
CP	Cyclic peptide
CPN	Cyclic peptide nanotube
CRN	Chemical reaction network
CSTR	Continuously stirred tank reactor
Cys	Cystein
DA	Decyclization agent
DLS	Dynamic light scattering
DMF	Dimethylformamide
DMPS	Dimercaptopropanesulfonate
DMSA	Dimercaptosuccinic acid
DMSO	Dimethylsulfoxide
DTT	Dithiothreitol
eq.	equivalents
EM	Effective molarity
ESI	Electron spray ionization
LC-HRMS, HRMS	Liquid chromatography high resolution mass spectrometry
LP	Linear peptide
MA	Michael acceptor
MAD	Meldrum's acid derivative
MP	MAD-peptide product
MPM	MAD-peptide-MAD product
MSDTT	1,4-bis(methyldisulfaneyl)butane-2,3-diol
NCL	Native chemical ligation
NMR	Nuclear magnetic resonance
ODE	Ordinary differential equations
P1	Peptide 1
P2	Peptide 2
PCOOH	2-(Diphenylphosphino)benzoic acid

PS2K	Bis(p-sulfonatophenyl)phenylphosphine dihydrate dipotassium salt
PS3Na	Triphenylphosphine-3,3',3''-trisulfonic acid trisodium salt
PSCyc	2- (Dicyclohexylphosphino)benzenesulfonic acid
PSOMe	2-[Di(2- methoxyphenyl)phosphino]benzenesulfonic acid
SEM	Scanning electron microscopy
TCEP	Tris(2- carboxyethyl)phosphine
TGA	Thioglycolic acid
THPP	Tris(3-hydroxypropyl)phosphine
UV-Vis	Ultraviolet-visible spectroscopy

List of symbols

Abs	Absorbance	a.u.
m/z	Mass over charge ratio	-
Peak area fr.	Peak area fraction	-
k	Rate constant	s^{-1}, $M^{-1}s^{-1}$
t	Time	s, min, h
α	Degree of supramolecular polymerization	-
δ	Chemical shift	ppm
σ	Cooperativity factor	-

Chapter 1. General introduction to CRNs and self-assembly

Abstract
While chemical reaction networks (CRNs) are of importance in all forms of life and are foundational to metabolic pathways, chemists still have much to learn on how to design such networks and keep them away from equilibrium. In addition, understanding how CRNs can be used to control self-assembled structures, such as in the cell's cytoskeleton, would be invaluable to developing the next generation of biomimicking materials. In this chapter, we give an overview of nonequilibrium systems that combine CRNs with self-assembly developed in recent years and highlight the features of self-assembly powered by CRNs within these systems.

Parts of this chapter have been published:
Anastasiia Sharko, Dimitri Livitz, Serena De Piccoli, Kyle J. M. Bishop*, and Thomas M. Hermans*, "Insights into Chemically Fueled Supramolecular Polymers" *Chem. Rev.* **2022**, 122, *13*, 11759–11777.

1.1 Chemical reaction networks and inspiration

Diverse biological organisms, from the smallest unit of life (i.e., a cell), clonal colonies of quacking aspens, to the Great Barrier Reef, are all united by how they function—through chemical reaction networks (CRNs). In nature, CRNs define rules under which functions such as division, sensing, and adaptation are performed. A system of coupled CRNs enables signal transduction and information processing. Metabolic networks are incredibly intertwined systems of reactions that operate out of equilibrium, include thousands of species, and are directed to produce macromolecular building blocks from sources of carbon, nitrogen, oxygen, and sulfur (Figure 1.1).[1] They are driven by the input of energy in the form of chemical or light fuel, dissipated while the network continues to evolve. Depending on starting and concomitant conditions, the evolution happens via distinct pathways. Living cells are in a steady state that is highly dynamic: fuels enter the cell, waste leaves, and the overall composition does not fluctuate much. When the cell is perturbed by external triggers, regulatory mechanisms kick in to adjust the corresponding metabolic pathways bringing the system into a new steady state.[2]

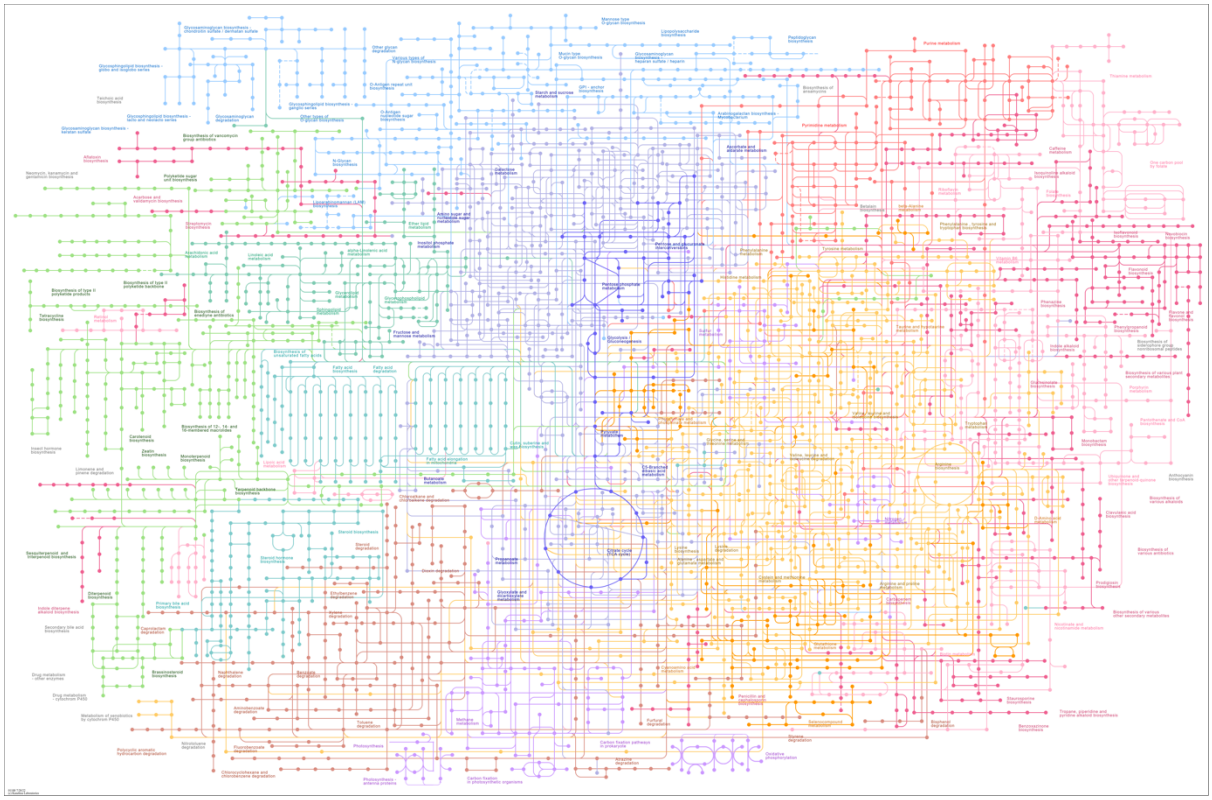


Figure 1.1 | Metabolic pathways in the human body. The Krebs cycle is in the center, connected to amino acid metabolism (yellowish region to the right), metabolism of nucleotides, co-factors, and vitamins (red-pink area to the right and top), carbohydrate metabolism (blue-purple area to the top), and lipid metabolism (green area to the left). Reprinted from ref. 1.

The complexity of these networks[3] has mesmerized and puzzled scientists since the early days of our existence. Nowadays, systems biology, biochemistry, and systems chemistry explore different approaches that could permit the reproduction and understanding of existing reaction networks and the creation of new artificial metabolic-like systems.[4] The missing part lies, however, somewhere in increasing complexity that transforms a set of chemical processes into the formation of biopolymers, vesicles, and, at some stage, cells.[5]

The first model systems studied in the context of understanding biological CRNs were inorganic reaction networks that demonstrated such features as oscillations (periodic or chaotic changes in components' concentrations) and pattern formations that were thought impossible in a homogeneous medium.[6] Intensively disputed problems of stability of these systems, entropy production, or emergent behavior led to the development of models based on non-linear chemical dynamics that could describe the regimes in which these systems exhibit 'unusual states'.[7] Nevertheless, it was clear that a bottom-up approach with the united efforts of physical chemists, biologists, chemical engineers, and mathematicians is required to practically deconstruct such systems into their functional components to approach possible answers to some of how and why questions. The goal of this approach is to localize and build *in vitro* the minimum CRN needed to reproduce a specific biological function using natural building blocks (enzymatic DNA bistable switches, transcriptional oscillator).[4]

1.1.1 Self-assembly and CRNs

A common denominator for inorganic and biological reaction networks capable of exotic behavior is self-organization.[8] It is often discussed in relation to emergence, as it presumes the appearance of non-linear effects (amplification, feedback loops) in the system powered by energy and interactions between the components.[9] In biology, self-organization is seen in morphogenesis (membrane and pattern formation), ants social superorganisms, or swarming (birds and fish flocking).[10–12] In chemistry, besides oscillating reactions, self-organization also includes autocatalytic networks, liquid crystals (switched with an electric field), and reaction-diffusion systems. Finally, it is strongly related to self-assembly, with often very vague borders between these two terms.[13] The main distinction is that self-organization is essentially a nonequilibrium process intrinsic to dynamic multistable systems.[14] The stability of self-organized structures is complemented by energy dissipation and material accumulation or loss. In contrast, self-assembly occurs toward thermodynamic equilibrium, it is an energy minimization process that brings together substrates (monomers) according to the rules encoded within their (molecular) structure.

In other words, self-assembled structures are usually not emergent in the sense that their properties (or level of order) can be derived from that of their components.[14] Emergent property is not a fundamental property of any individual system's component and appears on a bigger scale of the organization compared to the individual elements.[15] Despite this, self-assembly can occur in nonequilibrium (fueled) conditions—dissipative self-assembly.[16] A famous illustrative example is microtubules and actin filaments, cytoskeletal fibrous structures that are responsible for the cellular shape, its rigidity, or dynamicity. Microtubule growth is associated with the self-assembly of GTP-bound tubulin dimers, which does not require energy. Its disassembly, however, is energy-driven, as it requires a GTP-to-GDP hydrolysis reaction to trigger a conformational change of the dimer (Figure 1.2).[17] Microtubules originate from the microtuble organizing center (MTOC), located at the center of the cell; they spread to the cellular walls where they exert mechanical forces thus changing the MTOC position. This process is enabled by GTP-mediated growth and shrinkage of microtubules, a dynamic instability—a self-organization process powered by self-assembly.[18–21]

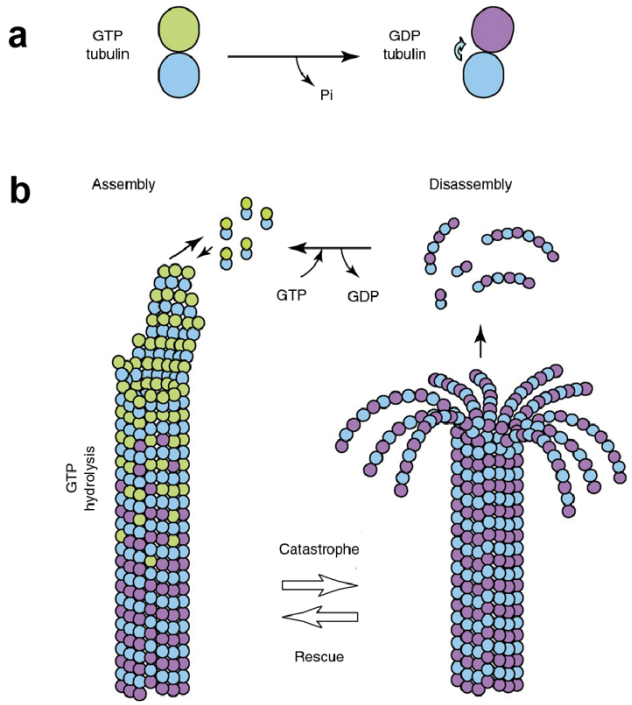


Figure 1.2 | Microtubule dynamics. (a) Hydrolysis of the GTP-tubulin dimer to GDP-tubulin with the change of conformation. (b) Microtubule growth from assembling GTP-tubulin dimers and shrinkage from disassembly of GDP-tubulin dimers. The processes are also known as rescue and catastrophe. Adapted from ref. 17.

The control over nucleation, their remarkable stiffness and self-repair of microtubules both *in vivo* and *in vitro* has motivated research toward artificial non-equilibrium systems that utilize self-assembly combined with chemical fuels. During the last decade, many new chemistries have appeared that explore the functions of CRN along with supramolecular polymerization and the formation of vesicles, colloids, and nanoparticles.[22–28] In these systems, the activation reaction transfers energy from a chemical fuel or light to the monomer, which becomes 'activated' and can then assemble. Once the fuel is exhausted, the monomer and/or its self-assembled structures 'deactivate' either by reacting with a second fuel (present from the beginning) or by spontaneous relaxation to the equilibrium. The activation and deactivation reactions are chosen to induce large changes in the non-covalent interactions between monomers such that self-assembly can be switched on and off on demand. The combination of activation and deactivation have given rise to 'transient self-assembly' systems with various interesting applications: timed drug release, temporary artery clamping, and transient catalysis.[29–36]

To understand how self-assembly is affected by thermodynamically unfavorable states, several models have been developed.[37–39] However, these models mostly consider only the dimerization of the monomer, which gives a good intuition about the system's behavior, but it is quite simplified as compared to the reported experimental systems.

1.2 Insights into chemically fueled supramolecular polymers

In collaboration with the group of Prof. Kyle Bishop, we have recently presented a new theoretical framework to describe dissipative self-assembly, using a range of increasingly complex models.[40] In this chapter, we will present a shortened version of the published review article.

The aim is to advance the previous models by introducing detailed mechanisms of supramolecular polymerization that occurs upon (chemical) activation and is inhibited by deactivation. With this, we describe how self-assembly is coupled to reaction networks and how they mutually affect each other. The model is built hierarchically: from the simplest case where there is (presumably) no coupling to the more advanced situations, where such coupling leads to oscillations in monomer concentration and supramolecular polymer length. The complexity of each model is defined by how many deactivation rate constants are considered, what mechanism of supramolecular polymerization is chosen (isodesmic or cooperative), and whether the system is closed or opened (batch vs. CSTR conditions). Specifically, deactivation can happen at the scale of the monomer k_{d1}, at the polymer end k_{d2}, or along the polymer chain k_{d3}. Whether the latter three rate constants are different from one another determines the complexity of the model, as we will see later on. The analysis has been limited to the one-dimensional supramolecular polymers; however, the literature examples considered with each model usually contain more complex structures.

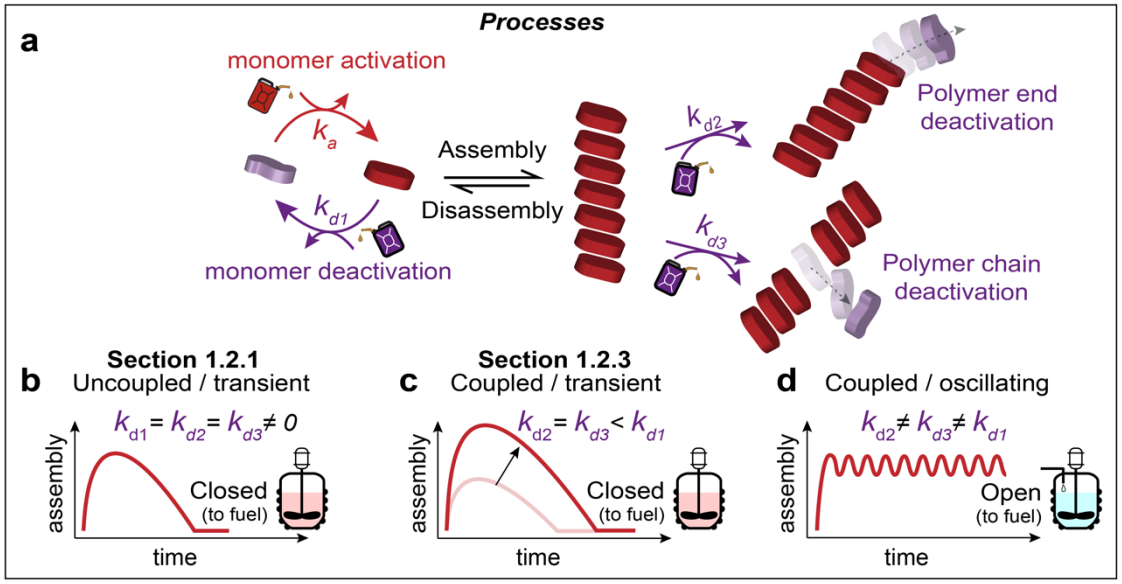


Figure 1.3 | Processes in transient systems with dissipative self-assembly. (a) Key processes and corresponding rate constants are used in the models. k_a—activation; k_{d1}—monomer deactivation in solution; k_{d2}—polymer deactivation deactivation; k_{d3}—polymer chain deactivation. (b) Transient activation uncoupled from self-assembly: see section 1.2.1. (c) Transient self-assembly with coupling: see section 1.2.3. (d) Oscillations in coupled dissipative self-assembly in an open system. Adapted from ref. 40.

1.2.1 Transient activation model

We questioned whether the typically transient self-assembly 'hump' (see Figure 1.3b) that is often reported in experimental papers was because all reported systems shared similarities in (de)activation and (dis)assembly kinetics, or perhaps the experimentalists tuned their reaction and assembly conditions to achieve such an interesting hump in time. To answer the latter question, we defined different regimes of chemically fueled self-assembly.

To capture our discussion quantitatively, we start with a simple uncoupled model (Appendix Section 3). Systems are uncoupled if: the activation/deactivation does not depend on the state of the assembly; consequently, the consumption of fuels and activation of monomers proceeds independently of whether or not the monomers assemble ($k_{d1} = k_{d2} = k_{d3}$). For the latter situation, we can write a 'Transient Activation' model, where deactivated monomer D is activated by reacting with chemical fuel F, leading to active monomer A, the latter of which can form self-assembled structures. A can be deactivated back to D by reaction with a second chemical fuel G. During this reaction cycle waste products P and P' are produced. As chemical equations, we can write:

$$\mathrm{D} + \mathrm{F} \xrightarrow{k_\mathrm{a}} \mathrm{A} + \mathrm{P}$$

$$\mathrm{A} + \mathrm{G} \xrightarrow{k_\mathrm{d}} \mathrm{D} + \mathrm{P}'$$

The concentrations of each of the species are given as italic lower case letters a, d, f, etc., and their initial values at $t = 0$ with subscript 0. We define the degree of activation as a/c_{tot}, where c_{tot} is the total concentration of the building block ($c_{tot} = a + d$), which stays constant at all times. To allow for generalization, we express the activation rate constant as $k_a c_{tot}$ and we divide all species concentrations by c_{tot}. The latter corresponds to setting $k_a \rightarrow 1$ and $c_{tot} \rightarrow 1$. Assuming a large excess of G—justified since, in many systems, the solvent governs deactivation—we can express $k_d' = k_d g_0$. It is easier to think about these simplifications in terms of typical times of activation $(k_a c_{tot})^{-1}$ and deactivation $1/k_d'$, both in units of seconds.

We can now recast the processes in the following rate equations:

$$\dot{a} = k_a f(c_{tot} - a) - k_d' a$$

$$\dot{f} = -k_a f(c_{tot} - a)$$

The typical behavior of this toy model shows four qualitatively different transient self-assembly regimes I–IV, defined by just two dimensionless parameters: i) $k_d'/k_a c_{tot}$: the ratio of activation of deactivation times, and ii) f_0/c_{tot}: the non-dimensional fuel concentration (blue, green, red, orange in Figure 1.4a). That means that for values below zero on the x-axis, the activation is faster than the deactivation, and vice versa for positive numbers. On the y-axis, one can read whether fuel is abundant (positive values) or sparse (negative values) relative to the monomer. For now, only consider the boxed numbers in Figure 1.4a, indicating uncoupled systems discussed further in Section 1.2.2. We will get back to the others in the next section (1.2.4). The exact details of how the chemical kinetics were extracted from each of the literature references are described in Appendix Section 6.

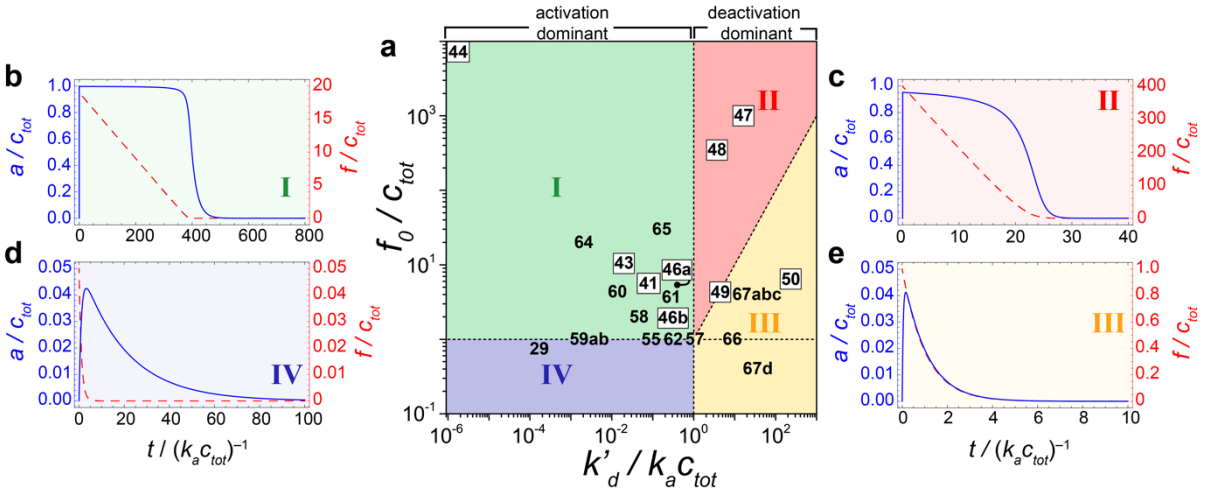


Figure 1.4 | Regimes in chemically fueled transient activation. (a) Dimensionless representation where four regimes can be distinguished (solid green, red, yellow, and blue areas), which smoothly transition from one to another. The numbers in the figure refer to the literature examples discussed in the text. Boxed numbers: uncoupled systems (section 1.2.2); normal numbers: coupled systems (section 1.2.4). Typical time traces (solid lines show activation a/c_{tot}, dashed lines are the dimensionless fuel concentration): (b) Regime I: fast and full activation, (c) Regime II: deactivation is fast, but since an excess of fuel is available, activation can still be high, (d) Regime III: very little fuel available, so activation is low, (e) Regime IV: fast deactivation, and not a lot of fuel leads to low activation and short transients. Time t in units $(k_a c_{tot})^{-1}$, fuel concentration (f/c_{tot}), f_0 is the fuel concentration at $t = 0$, timescale of activation over deactivation ($k_d'/k_a c_{tot}$). Adapted from ref. 40.

1.2.2 Literature examples

Using this simple 'Transient Activation' model, we mapped known literature examples onto a two-dimensional plot, which allows systems to be directly compared with one another.

1.2.2.1 Regime I

We start from regime I, which is the one that intuitively makes the most sense (to experimentalists). That is, it contains systems with an excess of fuel to the monomer, and a faster activation rate than deactivation. Starting from $t = 0$ (Figure 1.4b), the monomer quickly reaches a fully activated state ($a = c_{tot}$). A plateau is reached, where despite continuous deactivation, activation is maintained at the expense of continuous (pseudo-1st order) fuel consumption. After full fuel depletion, deactivation returns the system to the initial (deactivated/disassembled) state. Experimentally, a long plateau is not common, likely because the researchers build the conditions to achieve a 'hump' (monomer → assembly → monomer) within a reasonable experimental time scale or because the fuel itself degrades or reacts with a second fuel used for deactivation, thus not giving access to a plateau of the concentration of the activated monomer. This model assumes that the self-assembly of the activated monomer is faster than its deactivation, and deactivation happens prior to disassembly, triggering the latter.

An example by Adams and co-workers[41] shows a transient pH change followed by the assembly of a self-supporting gel in acidic conditions that redissolves when pH is basic (#41 in Figure 1.4a and Figure 1.5). A pH switch is implemented via (i) hydrolysis of urea by urease, releasing ammonia that increases the pH above the monomer pK_a, resulting in gelation, and (ii) hydrolysis of methyl formate producing formic acid that decreases the pH. Adjusting the amounts of urea, urease, and methyl formate, the conditions are found where pH increases much faster than decreases. The stiff gel is produced with 76% of the activated gelator. The

lifetime of the gels increases at higher gelator concentrations while amounts of urea, urease, and methyl formate are maintained the same. Repeated refueling delays the start of gelation, and gel's lifetime decreases. The authors reason this effect by waste accumulation, a common issue[42] in chemically fueled transient systems.

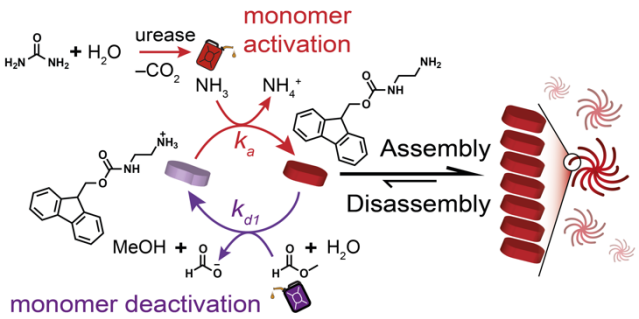


Figure 1.5 | Regime I, uncoupled. A system of Adams and co-workers[41] with a gelator that follows transient pH change. Reprinted from ref. 40.

Adams and co-workers reused the pH-control system for a different cycle[43], namely a gel–sol–gel instead of a sol–gel–sol transition, which is used in the annealing of hydrogels (#43 in Figure 1.4a). After rapid activation of a dipeptide-based hydrogelator (increase in pH corresponding to the dissolution of the gel), slower deactivation leads to the recovery of the original gel. The stiffness and morphology of the final hydrogel were shown to depend on the deactivation rate. Slower reaction kinetics resulted in a uniform and dense network of fibers, whereas the original gel contained mostly spherical assemblies. We cannot exclude the presence of coupling of self-assembly to the activation due to the slower diffusion of the fuel into the starting gel than the activation on the dissolved hydrogelator.

Panzarasa et al.[44,45] reported transient self-assembly of perylenediimide derivative following a programmable pH cycle (#44 in Figure 1.4a). The pH change from 5.5 to 10.5 was generated by the methylene glycol-sulfite 'clock' reaction. The mechanism relies on the reaction between formaldehyde and sulfite to produce hydroxymethanesulfonate and hydroxyl ions. Interestingly, by coupling the pH change with the hydrolysis of 1,3-propanesultone or δ-gluconolactone, they could achieve a transient pH change that could be repeated 10 times. The controlled variation in pH led to the aggregation of the perylenediimide derivative above its pK_a (6.5) promoted by π–π stacking and hydrophobic effects. During stirring, larger aggregates formed, eventually leading to precipitation. Upon hydrolysis of the sultone, the pH decreased to 4.5, causing protonation of tertiary amine groups of the perylenediimide derivative leading to immediate disassembly due to the charge repulsion.

Recently, George and co-workers[46] published a detailed study of fuel-controlled supramolecular polymerization (#46b in Figure 1.4a). Self-assembly of a charge-transfer complex consisting of tetrapotassium coronene and methyl viologen modified by benzaldehyde occurs after imine bond formation with 87% conversion. The cooperative mechanism and the living nature of polymerization were confirmed by seeded experiments. Polymer growth kinetics was controlled by varying amine ligands, their concentration, and pH, which delay imine formation and self-assembly. Slow ester hydrolysis was used to trigger disassembly: in the first example to lower the pH and promote imine degradation, and in the second example to truncate the amine tail, which rapidly reduces the size of the hydrophobic part of the monomer, breaking the hydrophobic contribution into polymer stabilization. In the latter case,

the ester moiety was embedded into the imine design, and various amounts of lipase were exploited to achieve the transient behavior of the system. Although we attributed this system to uncoupled examples, the authors also explored enzymatic (lipase) deactivation, where it is unclear if the self-assembly increases the deactivation rate due to multivalency (i.e., many cleavable bonds close to each other).

1.2.2.2 Regime II

Adjacent is regime II, where deactivation is faster than activation, but at sufficiently high fuel equivalents, the degree of activation (*a*) can still be very high (Figure 1.4c). Here, we find the system by Sorrenti et al.[47] using an enzymes-catalyzed reaction cycle with a perylene diimide substrate (#47 in Figure 1.4a). The monomer has two peptide sequences LRRASLG recognizable by a kinase (that phosphorylates the serine residues in the presence of ATP), and a phosphatase (that dephosphorylates the serines again). The phosphorylation promotes the self-assembly of the substrate and leads to a chiral inversion of the resulting supramolecular polymer. Dephosphorylation recovers the original substrate and its assemblies. When both enzymes and a shot of fuel ATP are added to the same batch reaction, transient self-assembly occures. The time scale of self-assembly is very fast compared to that of the enzymatic reactions (i.e., the system is uncoupled).

Besenius and co-workers[48] showed a unique example of a transient system where both reactions, responsible for monomer assembling and disassembling, are introduced by the same trigger (#48 in Figure 1.4a and Figure 1.6).

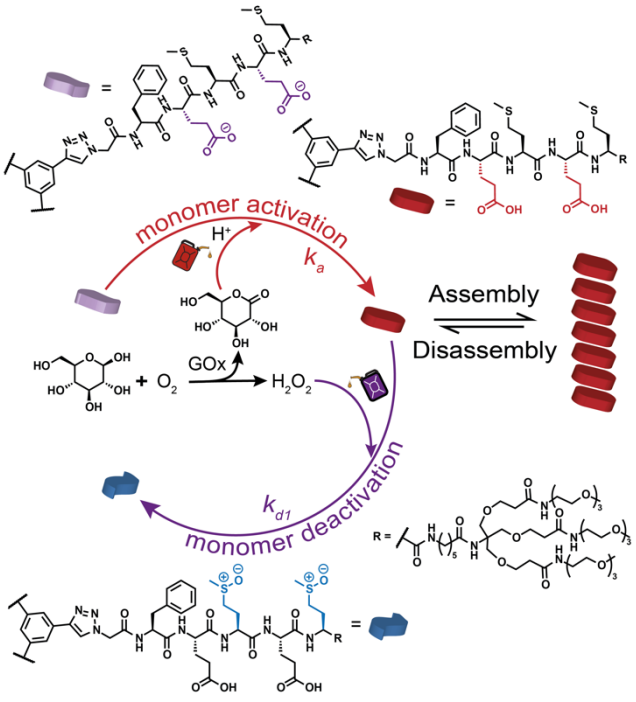


Figure 1.6 | Regime II, uncoupled. An example from Besenius and co-workers[48] with transient self-assembly driven by activation via pH change and by deactivation via oxidation. Two types of deactivated monomers exist (purple on the top left and blue on the bottom). Reprinted from ref. 40.

Oxidation of glucose catalyzed by glucose oxidase produces glucono-δ-lactone and hydrogen peroxide. Glucono-δ-lactone forms gluconic acid upon hydrolysis, decreasing the pH and leading to monomer protonation and self-assembly, yielding 99.8% activation. Increasing

amounts of hydrogen peroxide promote competing oxidation of monomer methionine residues, which reintroduces charge repulsion and causes disassembly. In Figure 3 of their paper, the increased amounts of glucose oxidase, i.e. increasing the amounts of both fuels, results in longer gel lifetimes and slower deactivation and disassembly. This can hint at the coupling of the deactivation and disassembly (see sections on coupled systems below), but this is to be confirmed.

1.2.2.3 Regime III

In regime III, levels of activation are lower as compared to regimes I and II, as the deactivation process is faster than activation, which limits the maximum degree of activation. Moreover, the fuel concentration is low and cannot compensate for fast deactivation (Figure 1.4e).

Recently, Quintard and co-workers[49] introduced a system that performs a sol–gel–sol transition (#49 in Figure 1.4a). The cycle starts with a solution of the anionic CO_2-adduct of O-tert-butyl-L-tyrosine (i.e., the carbamate form), with protonated DBU base as the counter-cation. Trichloroacetic acid (TCA) causes decarboxylation of carbamate and the formation of neutral O-tert-butyl-L-tyrosine, which gellates with chiral information preserved in bulk. By reabsorbing carbon dioxide from the air, the initial carbamate is restored, and the gel dissolves. The base DBU acts as a catalyst in the activation (promoting decomposition of TCA into CO_2 and chloroform) and deactivation (as a counter-ion for the anionic carbamate). At the same time, it is unclear if it can influence the self-assembly process. Due to the convenient waste removal of gaseous CO_2 and volatile chloroform, the system could go through 25 refueling cycles with minimal damping. We assigned the system to regime III because both decarboxylation and intermediate protonation were considered together as the activation process, which, when combined, are slower than deactivation.

In 2018, the group of George[50,51] presented ATP-fueled supramolecular polymerization of dipicolylethylenediamine-zinc (DPA-Zn) containing monomers (#50 in Figure 1.4a and Figure 1.7).

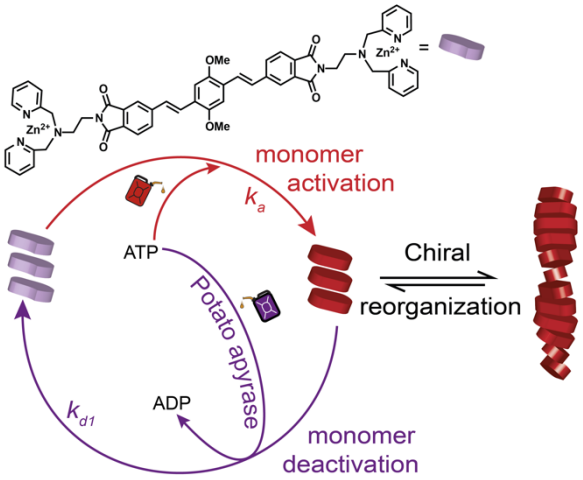


Figure 1.7 | Regime III, uncoupled. A system of George and co-workers[50] with transient chiral reorganisation of preassembled monomers triggered by interaction with ATP. Reprinted from ref. 40.

In the absence of fuel, the monomer is pre-assembled in slip-stacked aggregates. Multivalent binding of ATP to several DPA-Zn monomers organizes them into a left-handed helix, giving a 7% maximum activation. The potato apyrase enzyme catalyzes the hydrolysis of ATP

promoting mutual dislocation of monomers with subsequent depolymerization. The enzyme does not react with the unbound ADP that is present simultaneously with the complexed ATP, which suggests that self-assembly unlikely influences the kinetics of the deactivation process, and the system is uncoupled.

1.2.2.4 Regime IV

So far, no systems were found to belong in regime IV (Figure 1.4a), which has fast activation by low fuel amounts. According to the model, lack of fuel leads to low activation yields (Figure 1.4d), further decreasing when going deeper into more negative x and y values in Figure 1.4a. This means that this regime is unfavorable if a long-term transient assembled state is needed unless the assembling molecule has a very low critical aggregation concentration, where low activation (e.g., $a < 0.05$) would already result in significant assembly or gelation. Alternatively, the activation must be highly efficient, possibly with a very reactive fuel. Further, we will see, that coupling the reaction cycle to self-assembly (and thus protecting from deactivation) can be helpful. Alternatively, one could significantly increase the amount of activating fuel to push into regime I.

1.2.2.5 Other ways to obtain transient self-assembly

There are a few examples where mass transport phenomena are crucial to the transient self-assembly cycle. For example, Miravet and co-workers[52] used yeast-catalyzed hydrolysis of sucrose that gradually produces CO_2 *in situ,* causing acidification of the reaction mixture. Protonation of amphiphilic monomers triggers their assembly into fibrous networks. Decrease of the CO_2 content via exchange with the air over time increases the pH above the monomer pK_a resulting in gel dissolution due to electrostatic repulsion.

Recently, the group of Kim[53] investigated transient self-assembly, where the transient crystallization can be tuned by the ratio of acid and base in the solution. They used a non-assembling monomer (methyl orange) that is activated and assembles in acidic conditions due to electrostatic attraction. The latter is induced by CO_2 absorption into the solution and a decrease in pH. The authors also showed another approach to control the system, by passing gas mixtures through the monomer solution. Self-assembly of the monomer was achieved by passing a mixed flow of CO_2 (fuel) and argon (carrier gas) through the solution. After a defined time, the authors switch to a pure argon flow, which slowly degassed CO_2 from the solution, leading to recovery of the initial disassembled monomer. The latter approach, however, is akin to a traditional stimuli-responsive approach.

In 2015, Walther and co-workers[29] used urease to hydrolyze urea to increase the solution pH slowly (#29 in Figure 1.4a). Mixing an acidic buffer solution of urea with a basic urease solution resulted in a transient high–low–high pH cycle. Dissolving in this reaction mixture a pH-responsive peptide hydrogelator (Fmoc-LG-OH) resulted in transient gelation where self-assembly follows the pH change. If there is no co-assembly between charged and neutral peptides, the estimated activation degree is 94%. Interestingly, the time progression of pH was significantly slower in the presence of the peptide compared to a case without it, influencing both the activation and deactivation processes. Because peptides can be protonated, they increase the buffer capacity of the entire system (in addition to the used citrate buffer). This makes it difficult to return to the initial high pH state via the urea-urease reaction. The authors

observed this effect due to the almost equimolar amounts of peptide and buffer, whereas buffer is usually used in large excess to the reacting species. In this system, it is the chemical nature of the monomer that influences the activation and deactivation reactions. The study of Mondal et al.[54] on a similar system demonstrated a similar behavior when using 1 M HCl instead of the acidic buffer.
Mass transport or poor buffering can play a crucial role, but for clarity, we will not consider this in our models.

1.2.3 Coupled transient self-assembly model

As one could tell from the Transient Activation model, in regime II or IV it should be difficult to get appreciable amounts of activated monomer (i.e., high *a* values cannot be achieved). Thus, one would have to work at very high monomer concentrations to surpass the critical aggregation concentration. However, as will be discussed in this section, coupling chemical reactions to some form of self-assembly can distort the regimes, thus enabling transient self-assembly to occur. In contrast to the Transient Activation model, in this model, the deactivation is restricted in the assembled states, thus having very different kinetics depending on whether monomers are in solution or a polymer. To describe self-assembly, the isodesmic mechanism of supramolecular polymerization was used, to which the transient activation model was connected (see Appendix Section 4). Only the activated monomer can form polymers, whereas the deactivated monomer stays disassembled. Once an activated monomer is inside the polymer it cannot be deactivated unless it first dissociates from the polymer and is deactivated in the solution. In simpler words, the self-assembly protects the activated monomers from deactivation in solution. Under such conditions—even if the deactivation is much faster than the activation regime and fuel is not in great excess—there can still be many self-assembled species. We further describe examples where the coupling is evident or suspected, where we will not differentiate between isodesmic or cooperative polymerization mechanisms since often they are not specified or studied. We will discuss two different scenarios: 1) that of ‘catalysis’, when deactivation is accelerated by on-polymer catalysis (Section 1.2.4.1); and 2) that of ‘shielding’, where monomers are protected from deactivation once assembled (Section 1.2.4.2). As we will see later in the literature examples, there are cases where deactivation (by hydrolysis) is accelerated due to catalytic activity from proximal monomers in the supramolecular polymer structure. The result is that monomer activation is much more pronounced, and self-assembly can occur. At the same time, when self-assembly (of dimers or larger) occurs it results in rapid deactivation and disassembly. The second, more common scenario is when monomers can no longer be deactivated once self-assembled into supramolecular polymers, resulting in a shielding effect. Under such conditions—even if the deactivation is much faster than the activation regime and fuel is not in great excess—there can still be a large amount of self-assembled species.
The literature examples are still mapped onto the same Transient Activation model, where we use apparent activation and deactivation rates since most authors did not measure the chemical- and self-assembly kinetics separately. See Appendix Section 6 for a detailed description of how the kinetics was quantified.

1.2.4 Literature examples

1.2.4.1 Scenario 1: Self-assembly catalyzes deactivation ($k'_{d2} >> k'_{d1}$)

In 2018, Das and co-workers[55] introduced a dissipative self-assembly system based on an amphiphile molecule containing histidine (#55 in Figure 1.4a and Figure 1.8). In the presence of the activating fuel (EDC), the substrate self-assembles due to forming ester bonds with 4-nitrophenol. A self-supporting gel is formed within 2 minutes, redissolves with time, and the initial solution is recovered. This transient behavior can be explained by the presence of histidine, which is spatially close to the ester bond and induces a cooperative effect causing the hydrolysis of the ester, the restoration of the initial molecules, and the disruption of the gel. The same chemistry was later demonstrated in transient amyloid polymerization,[56] where the authors probed the electrical properties of the formed fibers. A histidine derivative, in the presence of EDC (fuel) and nitrophenol, allowed the transient formation of a self-supporting gel. Interestingly, by using phenol instead of nitrophenol, the gel lifetime was increased from hours to about 5 days. The increased hydrophobicity protects the supramolecular polymers from deactivation.

In 2019, the same group described transient metastable helical nanostructures[57] based on the previous system (#57 in Figure 1.4a). In this work, the self-assembly is driven by a stearoyl histidine and a nitrophenyl ester of stearoyl phenylalanine. The latter forms a gel due to the hydrophobic effect of the carbon tail and the phenyl ring in about 2 hours. No activating chemical fuel is required for this process. As for the previously described network, histidine catalyzes the hydrolysis of the ester bond, releasing stearoyl phenylalanine and nitrophenol. This causes the dissolution of the gel. The morphology of the system was analyzed by TEM, where the formation of helical ribbons was observed after 2 hours, and 90% of the population featured helical morphology in 4 hours. After 6 hours, the population of helices started to decrease, that is, about 15% hydrolysis of the ester bond. At 10 hours, all the helical nanostructures disappeared.

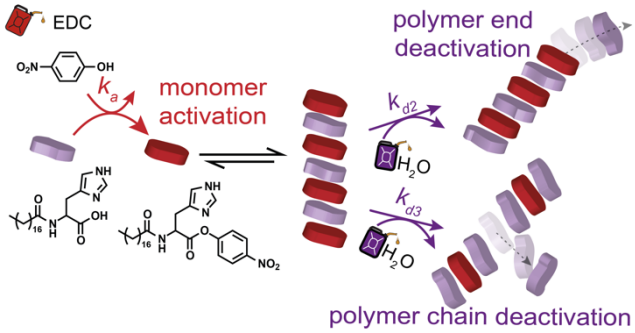


Figure 1.8 | Catalytic deactivation. The proximity of histidines accelerates hydrolysis in the systems of Das and co-workers.[55] Reprinted from ref. 40.

1.2.4.2 Scenario 2: Shielding of assembled monomers from deactivation ($k'_{d2} << k'_{d1}$)

Ulijn and co-workers[58] studied in 2015 a transient tripeptide system where the product of fast aspartame transacylation undergoes slow hydrolysis over 10 or 72 h depending on the coupled amino acid amide (#58 in Figure 1.4a and Figure 1.9). The formation of tripeptide is accompanied by the growth of supramolecular fibers and gelation that can be speeded up by additional sonication of the solution. Interestingly, the timescale of the activation process catalyzed by α-chymotrypsin is similar for L-phenylalanine (F) and L-tyrosine (Y) amides—

92% conversion is reached within 30 minutes—however, hydrolysis of the resulting tripeptide catalyzed by the same enzyme is 8 times slower for aspartame-phenylalanine (DFF-NH_2) than for aspartame-tyrosine (DFY-NH_2) product. Notably, both aspartame and tripeptide can hydrolyze with the formation of waste DF. The authors compared the behavior of several amino acid amides (F, Y, W, L, V, S, T) in the same reaction conditions. Only in cases F and Y the tripeptide product accumulate, and gelation occur. In contrast, with other amino acid amides, no evidence of tripeptide formation was observed, possibly due to its rapid hydrolysis. It suggested that nanofiber self-assembly in the case of tripeptides DFY-NH_2 and DFF-NH_2 prevented their enzymatic hydrolysis to DF, which means the system belongs to the coupled category. Upon refueling, the reaction cycle runs up to 3 times, reducing the conversion of tripeptide after each cycle due to the DF waste accumulation.

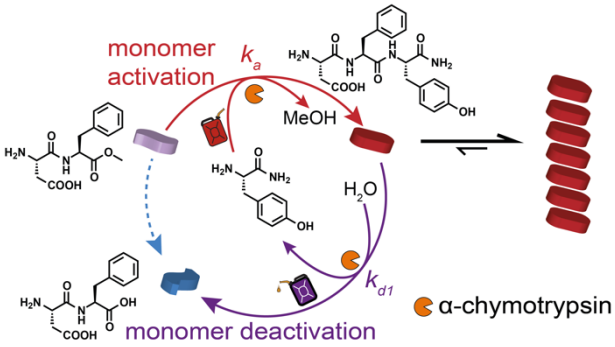


Figure 1.9 | Self-assembly shields monomers from deactivation. An example from Ulijn and co-workers.[58] Reprinted from ref. 40.

Earlier, in 2013, the group of Ulijn[59] showed a similar transient system based on naphthalene-dipeptide gelators (#59ab in Figure 1.4a). They pointed out that the sol-gel-sol conversion in transient systems will not be observed if the critical assembly concentration is lower than the equilibrium concentration of the system that it reaches after activation. In other words, the transient gel will reach equilibrium as a gel without solubilization. To confirm this, the authors compared kinetic and thermodynamic pathways from the transient product to the waste and vice versa. The stiffest gel showed the slowest deactivation kinetics and equilibrium hydrogelation in the end, whereas the least stiff gel formed transiently. Even though the system is coupled, we estimated the variety of its positions in regime I depending on the type of the activation fuel (amino acid amide) and the concentration of deactivation fuel (OH^-). At pH 10, deactivation proceeds faster than at pH 8, which reflects the position of the latter points further from the intersection of all regimes.

Finally, in 2018, Ulijn and co-workers[60] introduced a transient supramolecular chirality in their system (#60 in Figure 1.4a). α-Chymotrypsin catalyzes hydrolysis and amide bond formation of tyrosine methyl ester and prefers the L-amino acids compared with D-amino acids. Therefore, the naphthalene diimide monomer derivatized with L- and D-tyrosine has two unequal reaction sites targeted for activation. In the presence of glutamic acid (E), for example, the monomer forms a non-covalent complex that correlates with the reorganization from two-dimensional nanosheets into 1D left-helical fibers. Over time circular dichroism signal decreases, corresponding to decay of the complex, hydrolysis of methyl esters of the monomer, and the disintegration of nanofibers into small spherical aggregates. Depending on the amino acid used, transient helicity or delayed helicity could be achieved.

Ulijn's systems fall in regime I or on the border between regimes I and IV. Interestingly, self-assembly here enables the existence of transient behavior, slowing down the otherwise fast deactivation process. Without coupling, these systems will probably be deeply in regime III because of the fastest deactivation process.

Recently, Singh et al.[61] developed a new reaction cycle to achieve kinetic control over a gel-sol-gel conversion of an aldehyde-containing saccharide hydrogelator (#61 in Figure 1.4a). During the cycle, it can be reduced to its α-hydroxy sulfonate upon adding dithionite. The fuel induces the dissolution of the hydrogel fibers leading to a solution state. In the presence of an excess of formaldehyde, the initial substrate and, consequently, the initial hydrogel can be restored. However, a transient behavior could be obtained using an acid-catalyzed conversion of hexamethylenetetramine to formaldehyde mediated by glucono-δ-lactone (GdL). Then, formaldehyde is produced directly *in situ*, which means that by varying the concentration of GdL, the kinetics of the systems can be tuned. As such, at higher GdL concentrations, a fast gel-sol-gel conversion was observed, while at lower GdL concentrations, the plateau (solution state) is maintained for much longer. In this system, dithionite reacts with formaldehyde, making the transient disassembly difficult. Thus, to achieve it, the authors controlled the production of formaldehyde via the direct hydrolysis of a 'pre-fuel'. This cycle represents an example where an additional reaction pathway was implemented to delay deactivation because of the conflicting fuels, which allowed to achieve the transient behavior of the hydrogel.

A more detailed study of the coupling phenomenon in the transient systems was shown by Boekhoven and co-workers[62] in 2017 (#62 in Figure 1.4a). In their system where Fmoc protected amino acid or tripeptide undergoes activation and gelation upon conversion into anhydride, the rate of the deactivation, anhydride hydrolysis, is dependent on the activation yield, i.e. the more anhydride is produced, the stiffer is the gel and the lower is the hydrolysis rate. Authors observed that in case the assemblies were formed, the deactivation rate was two orders of magnitude lower than that of non-assembling or weakly assembling anhydrides. The yield of activation is strongly regulated by the amount of added activation carbodiimide fuel. When more than 0.3–0.5 eq of fuel is added, and anhydride yield is more than 2–3%, deceleration of hydrolysis two orders of magnitude, compared to non-assembling monomers, is observed. The reaction evolution with coupling can be fit using an additional 1st or 0th order rate constant representing delayed hydrolysis due to the self-assembly and is applicable above this certain threshold of aggregate concentration. The same group later studied how the molecular structure of the (peptide) monomer, the ionic strength of the solution, and the amount of added fuel (EDC) dictate whether self-assembly is transient, permanent, or non-occurring.[63] The work shows the factors to be considered while designing a transient chemically fueled self-assembling system, highlighting the influence of the solution conditions.

The system by Thordarson and co-workers[64] (#64 in Figure 1.4a) is found in regime I. The activation process is pH-based, which might suggest that the system is uncoupled. In acidic conditions, protonation of N,N'-dibenzoyl-L-cystine (DBC) triggers gelation. At pH 2.70, 88% of DBC is activated. The deactivation reaction—reduction of disulfide bonds by TCEP—is pH-independent. The authors show that the rate of activation can be higher (pH < 3) or lower (pH > 3). In the latter case, transient hydrogelation does not occur because the rate of disulfide reduction is higher than that of carboxylic acids' protonation, i.e. activation cannot outcompete deactivation. The authors observed that the lifetime of the transient gel is longer at lower pH,

due to the slower disulfide reduction rate. Since only the activation process is pH-dependent, it suggests that deactivation of DBC occurs only after its dissociation from the gel structures. In this way, self-assembly acts like a trap for the activated monomer protecting it from deactivation. The system was refueled 3 times, having an excess of TCEP and adding shots of DBC after each cycle was completed, during which no major influence of waste was observed. This example in the coupled regime shows that though, at first glance, conditions would be unfavorable to reach any transient behavior since the usual values of rate constant of disulfide bond reduction by TCEP are three orders of magnitude higher than the apparent rate constant DBC reduction in case of its gelation, also exceeding the activation rate constant of this system. However, fast self-assembly kinetically balances activation and deactivation, introducing a delay for the latter, which increases the lifetime of the activated state.

A similar system, which is also in regime I, is from Guan and co-workers (#65 in Figure 1.4a).[65] Their reaction cycle is based on thiol-disulfide oxidation-reduction reactions utilizing N-benzoyl-L-cysteine amide. Upon oxidation of cysteine thiol by hydrogen peroxide N,N'-dibenzoyl- L-cystine amide forms with 44% conversion that rapidly assembles in a self-supporting gel. Over time a disulfide reduction by dithiothreitol takes over, causing gel solubilization and retrieving the deactivated monomer. Compared to the Thordarson system, here, the monomer can be recovered to perform further reaction cycles after the addition of a shot of fuels. It resolves the recycling problem, which is important for material application, as, for some systems, deactivation of monomer leads to the formation of species that cannot be easily and quantitatively converted into a starting material. Again, by comparing an apparent deactivation rate constant with that of analogous reactions in the literature and considering the similarity of monomers in Thordarson and Guan systems, we assume the latter system is also coupled.

The first chemically fueled self-assembly system was developed in 2010 by van Esch and co-workers (#66 and 67 in Figure 1.4a).[66] They proposed the assembly of DBC by using a methylating agent. The latter neutralizes the total charges derived from the carboxylate groups resulting in the assembly of the molecule into fibers and macroscopic gelation. With time, the pH decrease leads to the disassembly of the fibers system and the recovery of the original molecule. In this system, the activation rate is higher than the deactivation rate, which means that fibers are continuously formed since fuel is present. In addition, in 2015,[67] the kinetics of the system were improved by the use of a more reactive methylating agent, which accelerates the (dis)assembly of the structure to a few hours rather than days.

Having discussed the different ways to obtain transient self-assembly of either uncoupled or coupled systems (section 1.2.2 and 1.2.4, respectively), we would like to highlight a final example moving toward a biological-like function. The group of Hamachi[68] demonstrated force generation in a propagating wave of supramolecular fibers whose growth and degradation are spatiotemporally controlled by non-interfering chemical stimuli. To demonstrate such behavior a solution of peptide hydrogelator together with a catalyst was prepared between two glass slides. Simultaneous addition of both activation and deactivation fuels at the one edge of the specimen caused the emergence of a wavefront of transient nanofibers that traveled to the opposite edge of the solution. The amount of fibrous material, its lifetime, and movement across the sample was shown to be highly dependent on reactants ratios. In the preliminary study of the system, besides full orthogonality of chemical triggers, authors also showed

activation to be 4 times faster than deactivation, which altogether makes a perfect system to develop a transient material and focus study on its properties. The authors quantitatively compared the generated force in their system to that of biological examples by measuring the speed of beads displacement that happens along the path of the gelation wave. Even though the persistence length of fibers was akin to that of actin the generated force was 150–550 times weaker than the stalling force of actin and microtubule polymerization. This observation was explained by the lack of directionality of artificial fibers and their stochastic entanglement compared to those in nature. The work brings the field of nonequilibrium systems to the next page showing the first steps toward the actual use of emergent properties and functionality on a material scale.

Finally, in the last and the most complex model, which includes fueled cooperative mechanism of supramolecular polymerization in an open system, we show the conditions under which oscillations in the length and concentration of supramolecular polymers can be observed and summarize the implications of these observations for the development of complex experimental systems. For a more detailed discussion, see ref. 40 and Appendix Section 5.

1.3 Conclusions

Artificial CRNs are becoming increasingly complex and approaching those found in nature. Growing not only by the number of intertwined reactions, they also include other processes mainly studied at equilibrium, such as self-assembly, trying to achieve emergent behavior. In recent years, many transient reaction cycles appeared, exploring different types of self-assembly and their role within the reaction networks. Supramolecular materials can now show intricate spatiotemporal behavior, a unique and new feature compared to stimuli-responsive materials.

In this chapter, we have discussed some examples of chemically fueled transient supramolecular polymers in the context of their rates of chemical (de)activation and (de)polymerization. The ‘uncoupled’ examples, where only the chemical (de)activation is considered, are described by ‘transient activation’ model. Coupling of self-assembly to the reaction networks has been reported to affect the apparent rates of all of the processes within the system, sometimes making self-assembly crucial to achieving any activation levels. To better understand such dependencies, it would be really helpful if the exact kinetics of chemical activation and deactivation were to be measured separately from those of assembly and disassembly, which we found not to be the case for most of the available literature examples. In a broader sense, any type of coupling of reactivity and self-assembly can lead to more interesting behavior, also beyond one-dimensional supramolecular polymers. Modeling of the coupling in the context of isodesmic polymerization mechanism shows that long-lived activation states are possible even in the regimes where they seem unfavorable. Taking into account the progress made so far, one can expect a range of exciting ways to couple activation and/or deactivation chemistries to self-assembly in the coming years. Lastly, we have provided a comprehensive model for supramolecular oscillations in chemically fueled supramolecular polymers. Oscillations, normally found in nonlinear systems, were discovered to be feasible in chemically fueled supramolecular polymers.

Overall, supramolecular Systems Chemists have unlocked the next level in complexity towards dissipative supramolecular polymers and further materials that can mimic that in nature. We anticipate that a deeper understanding and control over chemically fueled supramolecular polymers will be key to develop the next generation of interactive (bio)materials.

1.4 Outlook and aim of the thesis

A careful distinction between the terms (dissipative) self-assembly and self-organization is of importance to the field of systems chemistry as the implications direct the research towards its aims of understanding life and the creation of life-like systems. To move forward, it is not enough to trigger self-assembly out of equilibrium, the goal is to deepen the understanding of self-assembly coupled to CRN's with feedback loops, which can result in multistable systems with self-organization processes that could include self-replication, metabolism, and are compartmentalized when needed. For this, developing a detailed quantitative approach and a well-established characterization methodology is required, which can be used for interdisciplinary communication and training the newly entered the field.

In this thesis, we explore a careful bottom-up approach to CRN's with and without coupling to self-assembly. From **chapter 2** to **chapter 4**, we built a reaction cycle to follow fueled dissipative self-assembly of cyclic peptides. First, we separately study peptide cyclization and cyclic peptide self-assembly (**chapter 2**), followed by decyclization and disassembly (**chapter 3**). Then, we look for ways of combining cyclization and decyclization reactions under the same reaction conditions. Initially seemingly straightforward, the system required additional optimization of the deactivation reaction to not compete with cyclization (**chapter 4**). In these chapters, we explore Michael addition chemistry and use its versatility in nucleophile functionalization.

Chapter 5 presents a CRN that, contrary to previous chapters, exploits conjugate additions to recover electrophilic compounds in each cycle. Depending on the used oxidants, the system proceeds via different pathways distinguished by different system compositions. It poses an interest in studying varieties of its species under CSTR conditions and yet remains to be coupled to a self-assembling system.

Chapter 6 concludes this work with a summary of the main findings and gives a perspective of the learned lessons.

1.5 Authors contributions

Anastasiia Sharko extracted kinetic parameters from the literature examples to map them following the Transient activation model and, together with Serena De Piccoli, prepared summaries of the systems described in the literature. Dimitri Livitz, Kyle Bishop, and Thomas Hermans did the models' development and interpretation of the results.

1.6 Appendix

The appendix includes the entire Supporting Information of the published review.

1 Polymer Assembly & Disassembly

We consider the assembly of activated monomers A_1 into linear polymer chains A_n of length n and their disassembly through reactions of the form

$$\text{assembly/disassembly:} \quad \mathrm{A}_i + \mathrm{A}_j \underset{k_{ij}^-}{\overset{k_{ij}^+}{\rightleftharpoons}} \mathrm{A}_{i+j} \tag{1.1}$$

where k_{ij}^+ is the rate constant for assembling polymers of length i and j, and k_{ij}^- is the rate constant for the reverse disassembly process. In this section, we review some standard results on the equilibrium size distribution and the assembly kinetics for isodesmic and cooperative models of polymerization.

1.1 Equilibrium Size Distribution

The chemical potential of polymers of length n is approximated as

$$\mu_n = \mu_n^o + k_\mathrm{B}T \ln a_n \tag{1.2}$$

where μ_n^o is the standard part of the chemical potential (the free energy due to monomer-monomer interactions), $k_\mathrm{B}T$ is the thermal energy, and a_n is the polymer concentration.[1] At equilibrium, the chemical potential of monomers in polymers of length n is equal to that of monomers in polymers of length j such that

$$\frac{\mu_n}{n} = \frac{\mu_j}{j} \tag{1.3}$$

Setting $j = 1$, we obtain the following expression for the polymer concentration a_n in terms of the monomer concentration a_1

$$a_n = a_1^n \exp\left(-\frac{\mu_n^o - n\mu_1^o}{k_\mathrm{B}T}\right) \tag{1.4}$$

In the absence of activation and deactivation processes, the total concentration of activated monomer is constant and equal to m_1

$$m_1 = \sum_{n=1}^{\infty} n a_n \tag{1.5}$$

Given a model for the free energy of polymerization, $\mu_n^o - n\mu_1^o$, equations (1.4) and (1.5) fully specify the polymer concentrations a_n at equilibrium for $n = 1, 2, \ldots$.

1.2 Polymerization Kinetics

The polymerization reaction (1.1) proceeds with a rate

$$r_{ij} = k_{ij}^+ a_i a_j - k_{ij}^- a_{i+j} \tag{1.6}$$

where the rate constants are invariant to permutations of the i, j indices—that is, $k_{ij}^\pm = k_{ji}^\pm$. Moreover, the principle of detailed balance requires that the reaction rate vanish at equilibrium such that

$$\frac{k_{ij}^-}{k_{ij}^+} = \exp\left(\frac{\mu_{i+j}^o - \mu_i^o - \mu_j^o}{k_\mathrm{B}T}\right) = K_{ij} \tag{1.7}$$

where K_{ij} is the equilibrium dissociation constant. Summing over all such reactions, the polymer concentrations evolve in time as

$$\dot{a}_n = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \nu_{ij}^n r_{ij} \tag{1.8}$$

where ν_{ij}^n is the stoichiometric coefficient of species n in the reaction i, j. Substituting the rate law (1.6) and simplifying, the kinetic equations become

$$\dot{a}_n = \overbrace{-2a_n \sum_{j=1}^{\infty} k_{jn}^{+} a_j}^{\text{assembly losses}} + \overbrace{\sum_{j=1}^{n-1} k_{n-j,j}^{+} a_j a_{n-j}}^{\text{assembly gains}} - \overbrace{a_n \sum_{j=1}^{n-1} k_{n-j,j}^{-}}^{\text{disassembly losses}} + \overbrace{2\sum_{j=1}^{\infty} k_{nj}^{-} a_{j+n}}^{\text{disassembly gains}} \tag{1.9}$$

for $n = 1, 2, \ldots$ where the respective terms describe losses and gains due to polymer assembly and disassembly.

1.3 Isodesmic Polymerization Model

For isodesmic polymerization, the free energy of polymerization is modeled as

$$\frac{\mu_n^o - n\mu_1^o}{k_{\mathrm{B}}T} = (n-1)\varepsilon \quad \text{(isodesmic)} \tag{1.10}$$

where ε is the constant bond energy of each monomer-monomer bond (in units of $k_{\mathrm{B}}T$). Note that $\varepsilon < 0$ corresponds to attractive interactions between monomers thereby favoring assembly. Substituting this model into the detailed balance condition (1.7), we find that the dissociation constant is independent of polymer length, $K = e^{\varepsilon}$. We further assume that the rate constants for assembly and disassembly are also independent of length: $k_{ij}^{+} = k_{+}$ and $k_{ij}^{-} = k_{-}$.

Equilibrium Size Distribution

Substituting equation (1.10) into equation (1.4) for the equilibrium polymer concentrations, one finds an exponential distribution in polymer length n of the form

$$a_n = a_1^n e^{-(n-1)\varepsilon} = K\left(\frac{a_1}{K}\right)^n \tag{1.11}$$

where e^{ε} is identified as the dissociation constant K. Substituting this result into equation (1.5), the total monomer concentration m_1 is related to the free monomer concentration a_1 as

$$m_1 = \sum_{n=1}^{\infty} n a_n = \frac{a_1}{(1 - a_1/K)^2} \tag{1.12}$$

Solving for the the concentration of free monomer a_1, one finds

$$a_1 = \frac{K}{2m_1}\left(K + 2m_1 - \sqrt{K(K+4m_1)}\right) = \begin{cases} m_1 + \ldots & \text{for } m_1 \ll K \\ K - \sqrt{K^3/m_1} + \ldots & \text{for } m_1 \gg K \end{cases} \tag{1.13}$$

The dissociation constant K acts as a critical concentration, above which added monomers are incorporated into the polymer assemblies.

Equilibrium Moments

The p^{th} moment of the polymer size distribution is defined by the following sum over all polymer lengths

$$m_p = \sum_{n=1}^{\infty} n^p a_n \tag{1.14}$$

The zeroth moment m_0 describes the total concentration of polymer species of any length. The first moment m_1 describes the total concentration of activated monomer. The second moment m_2 is related to the dispersity $Đ$ as

$$Đ = \frac{m_2 m_0}{m_1^2} \tag{1.15}$$

At equilibrium, the first few moments can be expressed in terms of the total monomer concentration m_1 as

$$\begin{aligned} m_0 &= \frac{1}{2} K \left(-1 + \sqrt{1 + 4m_1/K} \right) \\ m_2 &= m_1 \sqrt{1 + 4m_1/K} \\ m_3 &= m_1 (1 + 6m_1/K) \end{aligned} \tag{1.16}$$

The dispersity varies from one for $m_1 \ll K$ to two for $m_1 \gg K$ as

$$Đ = \frac{2}{1 + (1 + 4m_1/K)^{-1/2}} \tag{1.17}$$

Polymerization Kinetics

With the assumption of constant rate constants for assembly and disassembly, the kinetic equations (1.9) governing the transient polymer concentrations simplify as

$$\dot{a}_n = -2k_+ a_n \sum_{j=1}^{\infty} a_j + k_+ \sum_{j=1}^{n-1} a_j a_{n-j} - k_-(n-1)a_n + 2k_- \sum_{j=1}^{\infty} a_{j+n} \tag{1.18}$$

for $n = 1, 2, \ldots$. At time $t = 0$, we will assume that there are no polymers in solution

$$a_n(0) = m_1 \delta_{1n} \tag{1.19}$$

where m_1 is the total monomer concentration. For the isodesmic model, this infinite system of equations can be solved exactly for the initial condition (1.19) by making use of evolution equations for the first two moments of the polymer distribution (Fig. 1).

Moment Evolution Equations

Dynamical equations for the moments m_p are obtained by summing equation (1.9) over all polymer lengths as

$$\dot{m}_p = \sum_{n=1}^{\infty} n^p \dot{a}_n \tag{1.20}$$

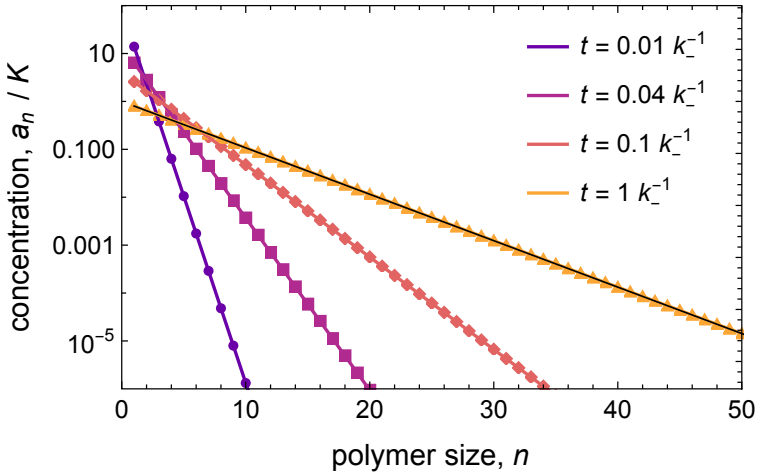


Figure 1: Transient polymer size distribution for $m_1/K = 20$. The markers show the concentrations a_n (units of $K = e^{\varepsilon}$) of polymers of size n for four times (units of k_-^{-1}) as obtained by numerical integration of equations (1.9). The solid curves show the predictions of equation (1.22) based on the zeroth moment m_0 obtained from (1.21); the black curve denotes the equilibrium distribution. The number average polymer length at equilibrium is 5.

The resulting evolution equations for the first few moments are

$$\begin{aligned} \dot{m}_0 &= -k_+ m_0^2 + k_-(m_1 - m_0) \\ \dot{m}_1 &= 0 \\ \dot{m}_2 &= 2k_+ m_1^2 + \tfrac{1}{3}k_-\,(m_1 - m_3) \\ \dot{m}_3 &= 6k_+ m_1 m_2 + \tfrac{1}{2}k_-\,(m_2 - m_4) \end{aligned} \tag{1.21}$$

subject to the initial condition, $m_p(0) = m_1$ for $p = 0, 1, \ldots$. As expected, the first moment—the total monomer concentration—does not change with time due to assembly or disassembly of monomers into polymers. Notably, the evolution of the zeroth moment evolves independently of the higher moments.

In general, the higher moments are governed by an infinite system of coupled equations. For this problem, however, there exist an exact solution for the transient polymer size distribution that depends only on the moments $m_0(t)$ and m_1

$$a_n(t) = \frac{m_0^2(t)}{m_1}\left(1 - \frac{m_0(t)}{m_1}\right)^{n-1} \tag{1.22}$$

From this exponential size distribution, the higher order moments can be expressed in terms of $m_0(t)$ and m_1—for example, the second moment is

$$m_2(t) = m_1\left(\frac{2m_1}{m_0(t)} - 1\right) \tag{1.23}$$

The dispersity evolves in time as as

$$Đ(t) = \frac{m_0(t)m_2(t)}{m_1^2} = \frac{2m_1 - m_0(t)}{m_1} \tag{1.24}$$

Equilibrium Relaxation Rate

The moment evolution equations can be used to determine the relevant time scale on which the system relaxes to its equilibrium state. In particular, we can linearize equation

(1.21) for the zeroth moment m_0 about its equilibrium value m_0^{eq} to determine the rate at which perturbations from equilibrium $\Delta m_0 = m_0 - m_0^{\text{eq}}$ decay in time

$$\Delta \dot{m}_0 = -\lambda_{\text{eq}} \Delta m_0 + O(\Delta m_0^2) \tag{1.25}$$

where $\lambda_{\text{eq}} = 2k_+ m_0^{\text{eq}} + k_-$ is the equilibrium relaxation rate. Substituting the equilibrium polymer concentration m_0^{eq} from equation (1.16), the relaxation rate can be expressed as

$$\lambda_{\text{eq}} = k_- \sqrt{1 + 4m_1/K} = \begin{cases} k_- & \text{for } m_1 \ll K \\ \sqrt{4k_+ k_- m_1} & \text{for } m_1 \gg K \end{cases} \tag{1.26}$$

where the approximate expressions correspond to weakly associating ($m_1 \ll K$) and strongly associating ($m_1 \gg K$) monomers.

1.4 Cooperative Polymerization Model

For cooperative polymerization, the free energy of polymerization is modeled as

$$\frac{\mu_n^o - n\mu_1^o}{k_{\text{B}}T} = \begin{cases} (n-1)\varepsilon_{\text{c}} & \text{for } n < n_{\text{c}} \\ (n-n_{\text{c}})\varepsilon + (n_{\text{c}}-1)\varepsilon_{\text{c}} & \text{for } n \geq n_{\text{c}} \end{cases} \quad \text{(cooperative)} \tag{1.27}$$

where n_{c} is the number of monomers in the critical nucleus, and ε_{c} is the monomer-monomer bond energy (scaled by $k_{\text{B}}T$) within subcritical polymers $A_{n<n_{\text{c}}}$. Cooperative (as opposed to anticooperative) polymerization requires that $\varepsilon_{\text{c}} > \varepsilon$ such that the average strength of monomer-monomer interactions increases with polymer length. For this model, the detailed balance condition (1.7) implies the following relation among the rate constants for assembly and disassembly

$$\ln\left(\frac{k_{ij}^-}{k_{ij}^+}\right) = \varepsilon_{\text{c}} + \frac{1}{2}(\varepsilon_{\text{c}} - \varepsilon)\big(|i-n_{\text{c}}| + |j-n_{\text{c}}| - |i+j-n_{\text{c}}| - n_{\text{c}}\big) \tag{1.28}$$

Figure 2 shows how the dissociation constant K_{ij} varies with the sizes i and j of the participating polymers. For short polymers ($i + j \leq n_{\text{c}}$), the dissociation constant is $K_{ij} = e^{\varepsilon_{\text{c}}}$; for long polymers ($i, j \geq n_{\text{c}}$), it falls to $K_{ij} = e^{\varepsilon_{\text{c}} - n_{\text{c}}(\varepsilon_{\text{c}} - \varepsilon)}$. These thermodynamic considerations do not specify the rate constants for assembly and disassembly but only their ratio as a function of polymer length. To describe the kinetics of cooperative polymerization, we make the simplifying assumption that the forward rate constants for assembly are constant and independent of polymer length, $k_{ij}^+ = k_+$; the rate constant for the reverse process of disassembly is then proportional to the dissociation constant, $k_{ij}^- = k_+ K_{ij}$.

Equilibrium Size Distribution

Substituting equation (1.27) into equation (1.4) for the equilibrium polymer concentrations, one finds a piecewise exponential distribution in polymer length n of the form

$$a_n = \begin{cases} a_1^n e^{-(n-1)\varepsilon_{\text{c}}} & \text{for } n < n_{\text{c}} \\ a_1^n e^{-(n-n_{\text{c}})\varepsilon - (n_{\text{c}}-1)\varepsilon_{\text{c}}} & \text{for } n \geq n_{\text{c}} \end{cases} = \begin{cases} K_{\text{c}}(a_1/K_{\text{c}})^n & \text{for } n < n_{\text{c}} \\ K_{\text{c}}(K/K_{\text{c}})^{n_{\text{c}}}(a_1/K)^n & \text{for } n \geq n_{\text{c}} \end{cases} \tag{1.29}$$

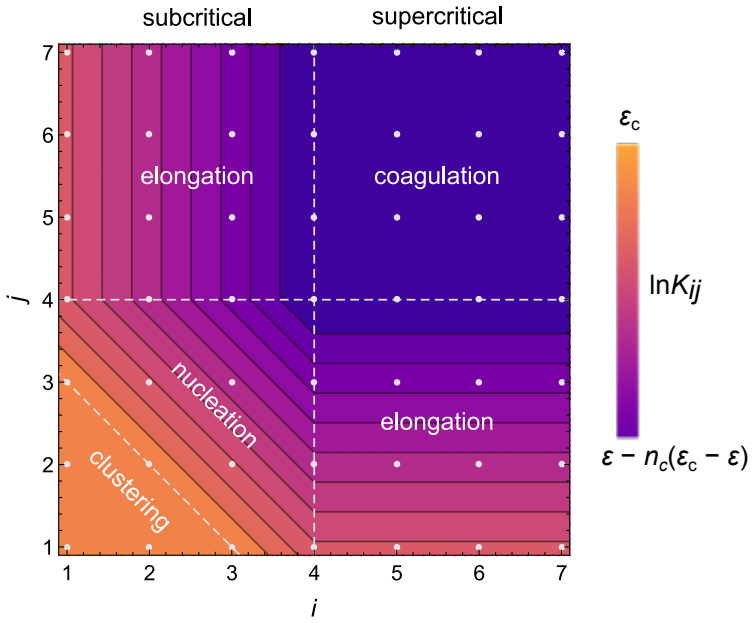


Figure 2: Log dissociation constant $\ln K_{ij}$ for the polymer A_{i+j} into fragments A_i and A_j in the cooperative model. Here, the critical nucleus size is $n_\mathrm{c} = 4$. The different assembly processes are distinguished by the the sizes i and j of the participating components: (i) *clustering* of monomers and subcritical polymers of length $n < n_\mathrm{c}$ to form subcritical polymers $(i, j, i + j < n_\mathrm{c})$, (ii) *nucleation* of supercritical polymers of length $n \geq n_\mathrm{c}$ from monomers and subcritical polymers $(i, j < n_\mathrm{c}$ and $i+j \geq n_\mathrm{c})$, (iii) *elongation* of these supercritical polymers by addition of monomers and subcritical polymers $(\min(i, j) < n_\mathrm{c}$ and $\max(i, j), i + j \geq n_\mathrm{c})$, and (iv) *coagulation* of two supercritical polymers to form one larger polymer $(i, j, i + j \geq n_\mathrm{c})$. Each assembly process is accompanied by a reverse disassembly process.

where $K = e^{\varepsilon}$ and $K_\mathrm{c} = e^{\varepsilon_\mathrm{c}}$ are dissociation constants. Substituting this result into equation (1.5), the total monomer concentration m_1 is related to the free monomer concentration a_1 as

$$m_1 = \frac{a_1^{n_\mathrm{c}+1}}{K_\mathrm{c}^{n_\mathrm{c}-1}} \left(\frac{a_1 n_\mathrm{c} - (n_\mathrm{c}+1)K_\mathrm{c}}{(a_1 - K_\mathrm{c})^2} - \frac{a_1 n_\mathrm{c} - (n_\mathrm{c}+1)K}{(a_1 - K)^2} \right) + \frac{a_1 K_c^2}{(a_1 - K_\mathrm{c})^2} \tag{1.30}$$

which simplifies to equation (1.12) when $K_\mathrm{c} = K$ in the absence of cooperativity. The solution to this equation for a_1 has the same limiting behaviors as equation (1.13) for isodesmic polymerization

$$a_1 = \begin{cases} m_1 + \ldots & \text{for } m_1 \ll K \\ K - K\sqrt{\frac{K}{m_1}\left(\frac{K}{K_\mathrm{c}}\right)^{n_\mathrm{c}-1}} + \ldots & \text{for } m_1 \gg K \end{cases} \tag{1.31}$$

The parameters K_c and n_c control the width of the transition region between these limits, with larger K_c and n_c producing sharper transitions (Fig. 3).

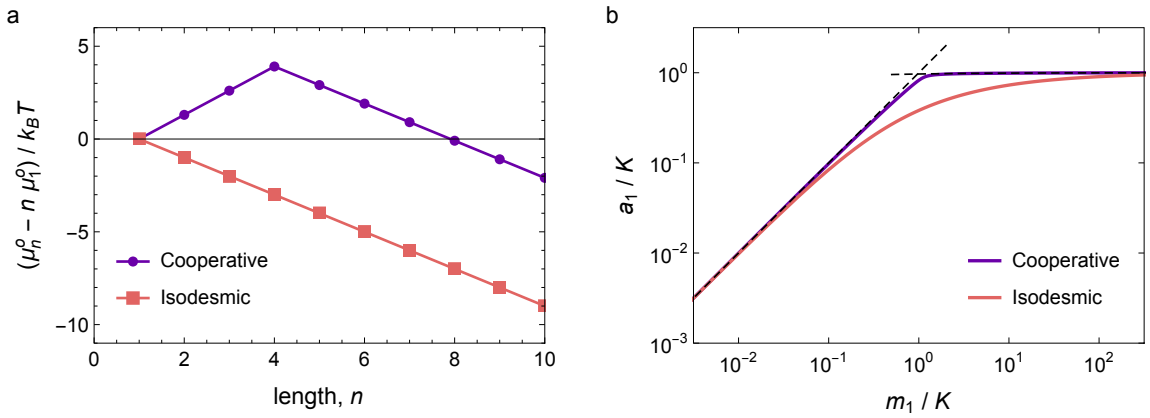


Figure 3: (a) Free energy of polymerization as a function of polymer length for the cooperative and isodesmic models with a common monomer-monomer bond energy of $\varepsilon = -1$. Here, the size of the critical nucleus is $n_c = 4$, and the bond energy within subcritical polymers is $\varepsilon_c = 1.30$. (b) Concentration of free monomers a_1 as a function of the total monomer concentration m_1 for the cooperative and isodesmic models. Concentrations are scaled by the dissociation constant $K = e^{\varepsilon}$. The parameters for the cooperative model are $n_c = 4$ and $K_c/K = e^{\varepsilon_c - \varepsilon} = 10$ as in (a). The dashed lines show the limiting behaviors of equation (1.31).

Polymerization Kinetics

For cooperative polymerization, the polymer concentrations are governed by the rate equations (1.9) with forward rate constants $k_{ij}^{+} = k_{+}$ and reverse rate constants

$$k_{ij}^{-} = k_{+} e^{\varepsilon} \begin{cases} e^{\varepsilon_c - \varepsilon} & \text{for } i, j, i+j < n_c \\ e^{-(i+j-n_c-1)(\varepsilon_c - \varepsilon)} & \text{for } i, j < n_c \le i+j \\ e^{-(j-1)(\varepsilon_c - \varepsilon)} & \text{for } j < n_c \le i, i+j \\ e^{-(i-1)(\varepsilon_c - \varepsilon)} & \text{for } i < n_c \le j, i+j \\ e^{-(n_c-1)(\varepsilon_c - \varepsilon)} & \text{for } n_c \le i, j, i+j \end{cases} \tag{1.32}$$

Here, the five expressions for the rate constants correspond to the five regions in Figure 2. For cooperative polymerization with $\varepsilon_c > \varepsilon$, shorter polymers disassemble faster than longer polymers. Note that we recover the same rate constants as the isodesmic model when $\varepsilon_c = \varepsilon$.

In general, the kinetic equations (1.9) cannot be solved analytically and are instead integrated numerically to determine the time evolution of the polymer concentrations $a_n(t)$. In the simplest approach, the infinite system of equations is truncated at some maximum polymer length $n_{\max}$, above which the polymer concentrations are set to zero

$$\dot{a}_n = -2k_{+}a_n \sum_{j=1}^{n_{\max}-n} a_j + k_{+} \sum_{j=1}^{n-1} a_j a_{n-j} - a_n \sum_{j=1}^{n-1} k_{n-j,j}^{-} + 2 \sum_{j=1}^{n_{\max}-n} k_{nj}^{-} a_{j+n} \tag{1.33}$$

for $n = 1, \ldots, n_{\max}$. The accuracy of this approach requires that the maximum polymer length be much larger than the number average polymer length at equilibrium. For strongly associating monomers ($m_1 \gg K$), this condition implies that

$$n_{\max} \gg \sqrt{\frac{m_1}{K} \left(\frac{K_c}{K}\right)^{n_c - 1}} \tag{1.34}$$

Figure 4 shows the transient polymer distribution computed in this way for parameters $n_c = 4$, $m_1/K = 10$, and $K_c/K = 10$. Note that the transient size distribution changes qualitatively with time as polymers nucleate, elongate, and coagulate en route to the final equilibrium distribution. Consequently, an exact low order description—for example, in terms of the first few moments of the distribution—is not possible for cooperative polymerization. Nevertheless, approximate schemes[2] are available that provide an efficient dynamical description under certain conditions (e.g., for $K \ll m_1 \ll K_c$).

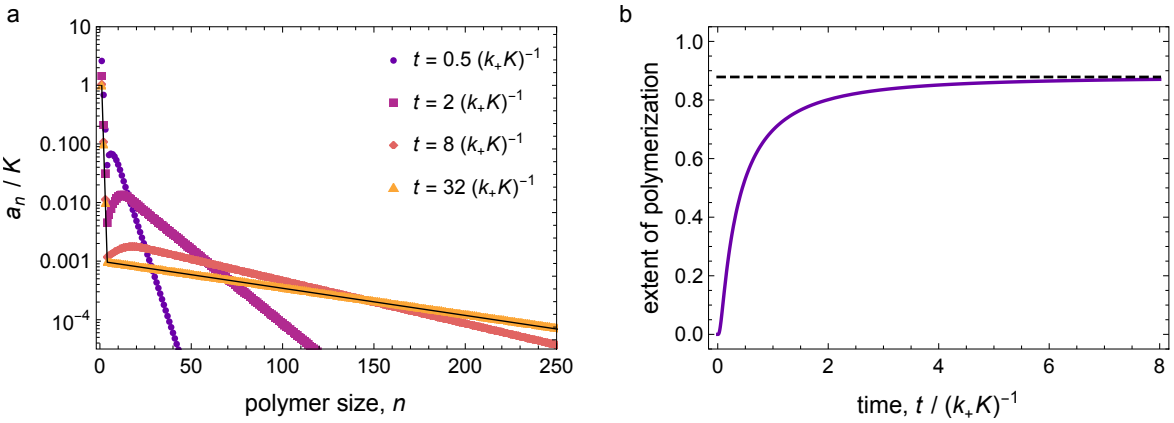


Figure 4: (a) Transient polymer concentration a_n (units of $K = e^{\varepsilon}$) as a function of size n as predicted by the cooperative model with parameters $n_c = 4$, $m_1/K = 10$, and $K_c/K = 10$. The markers denote the numerical solution of equations (1.33) with rate constants given by equation (1.32) and $n_{\max} = 500$. The solid curve shows the equilibrium distribution of equation (1.29). (b) Extent of polymerization vs. time corresponding to the cooperative polymerization process in (a). See also Supporting Video 1.

2 Model of Chemically Fueled Polymer Assembly

The processes of polymer assembly and disassembly (1.1) detailed in Section 1 can be driven from equilibrium by chemically fueled processes of monomer activation and deactivation:

$$\text{monomer activation:} \quad \mathrm{D} + \mathrm{F} \xrightarrow{k_\mathrm{a}} \mathrm{A}_1 + \text{products} \tag{2.1}$$

$$\text{monomer deactivation:} \quad \mathrm{A}_1 + \mathrm{G} \xrightarrow{k_\mathrm{d1}} \mathrm{D} + \text{products} \tag{2.2}$$

$$\text{polymer end deactivation:} \quad \mathrm{A}_{i+1} + \mathrm{G} \xrightarrow{k_\mathrm{d2}} \mathrm{A}_i + \mathrm{D} + \text{products} \tag{2.3}$$

$$\text{polymer chain deactivation:} \quad \mathrm{A}_{i+j+1} + \mathrm{G} \xrightarrow{k_\mathrm{d3}} \mathrm{A}_i + \mathrm{A}_j + \mathrm{D} + \text{products} \tag{2.4}$$

In the process of *monomer activation* (2.1), a deactivated monomer D reacts with an activating fuel F to produce the activated monomer A with a second order rate constant k_a. In general, this activation process is reversible with forward and reverse rate constants k_a^+ and k_a^- satisfying the detailed balance condition. Here, however, we will approximate the fueled reactions as irreversible; we denote the forward rate constant as $k_\mathrm{a} = k_\mathrm{a}^+$ and neglect the reverse reaction.

We consider three types of deactivation processes that depend on the context of the activated monomer involved. *Monomer deactivation* (2.2) describes the reaction of a free activated monomer A_1 and a deactivating fuel G to produce a free deactivated monomer D with forward rate constant k_d1. Alternatively, when an activated monomer is part of a polymer assembly, its chemically fueled deactivation may proceed at different rate—perhaps slower or faster than that of free monomer. We assume that the deactivated monomer does not form polymer assemblies. Consequently, the deactivation of a terminal monomer at either end of a polymer chain results in its dissociation via *polymer end deactivation* (2.3) with rate constant k_d2. The deactivation of other non-terminal monomers within polymer assemblies causes fragmentation via *polymer chain deactivation* (2.4) with rate constant k_d3.

2.1 Kinetic Equations

Within a well-mixed reactor, the concentrations of the activating and deactivating fuels evolve in time as

$$\dot{f} = k_\mathrm{f}(f_0 - f) - k_\mathrm{a} f d \tag{2.5}$$

$$\dot{g} = k_\mathrm{g}(g_0 - g) - k_\mathrm{d1} g a_1 - k_\mathrm{d2} g \sum_{j=2}^{\infty} 2a_j - k_\mathrm{d3} g \sum_{j=3}^{\infty} (j-2) a_j \tag{2.6}$$

Here, the rate constants k_f and k_g describe the rate of fuel exchange with reservoirs of concentrations f_0 and g_0, respectively. For a batch reactor closed to material exchange, these rate constants are zero: $k_\mathrm{f} = k_\mathrm{g} = 0$. For a continuously stirred tank reactor (CSTR), they equal the ratio between the volumetric flow rate Q exiting the reactor and the reactor volume V: $k_\mathrm{f} = k_\mathrm{g} = Q/V$. For a semi-batch reactor, the fuels are delivered at constant rates $k_\mathrm{f} f_0$ and $k_\mathrm{g} g_0$ even as the rates of fuel removal approach zero, $k_\mathrm{f} \to 0$ and $k_\mathrm{g} \to 0$. The general case of $k_\mathrm{f} \neq k_\mathrm{g}$ describes fuel delivery/removal from large reservoirs connected to the reactor by semi-permeable membranes.

The concentrations of deactivated monomers d, activated monomers a_1, and polymers a_n evolve in time as

$$\dot{d} = -k_{\mathrm{a}} f d + k_{\mathrm{d}1} g a_1 + k_{\mathrm{d}2} g \sum_{j=2}^{\infty} 2a_j + k_{\mathrm{d}3} g \sum_{j=3}^{\infty} (j-2) a_j \tag{2.7}$$

$$\dot{a}_1 = k_{\mathrm{a}} f d - k_{\mathrm{d}1} g a_1 + 2k_{\mathrm{d}2} g a_2 + 2k_{\mathrm{d}3} g \sum_{j=3}^{\infty} a_j - 2k_+ a_1 \sum_{j=1}^{\infty} a_j + 2 \sum_{j=1}^{\infty} k_{j1}^- a_{j+1} \tag{2.8}$$

$$\begin{aligned} \dot{a}_n = {} & -2k_{\mathrm{d}2} g (a_n - a_{n+1}) - k_{\mathrm{d}3} g (n-2) a_n + 2k_{\mathrm{d}3} g \sum_{j=2}^{\infty} a_{j+n} \\ & - 2k_+ a_n \sum_{j=1}^{\infty} a_j + k_+ \sum_{j=1}^{n-1} a_j a_{n-j} - a_n \sum_{j=1}^{n-1} k_{j,n-j}^- + 2 \sum_{j=1}^{\infty} k_{jn}^- a_{j+n} \end{aligned} \tag{2.9}$$

for $n > 1$. Here, the rate constants for polymer assembly are assumed constant $k_{ij}^+ = k_+$; those of the reverse disassembly process k_{ij}^- are either constant for the isodesmic model or given by equation (1.32) for the cooperative model. These equations assume that the reactor is closed to monomer exchange; consequently, the total monomer concentration—both activated and deactivated, assembled and disassembled—is at all times constant and equal to c_{tot}

$$c_{\mathrm{tot}} = d + \sum_{n=1}^{\infty} n a_n \tag{2.10}$$

Initially, at time $t = 0$, the concentrations are specified as

$$f(0) = f_0, \quad g(0) = g_0, \quad d(0) = c_{\mathrm{tot}}, \quad a_n(0) = 0 \tag{2.11}$$

where it is assumed that all monomers start from the deactivated state. Together, equations (2.5)—(2.11) govern the transient concentrations of activating fuel $f(t)$, deactivating fuel $g(t)$, deactivated monomer $d(t)$, activated monomer $a_1(t)$, and their polymer assemblies $a_n(t)$. The variables and parameters of this model are summarized in Tables 1 and 2, respectively.

Table 1: Model variables.

Symbol	Significance
t	time
$f(t)$	concentration of activating fuel F
$g(t)$	concentration of deactivating fuel G
$d(t)$	concentration of deactivated monomer D
$a_1(t)$	concentration of free activated monomer A_1
$a_n(t)$	concentration of activated n-mer A_n
$m_p(t) = \sum_{n=1}^{\infty} n^p a_n(t)$	p^{th} moment of the polymer number distribution
$m_p'(t) = \sum_{n=2}^{\infty} n^p a_n(t)$	p^{th} partial moment of the polymer number distribution

2.2 Numerical Solution

The kinetic model is solved numerically for polymers up to a maximum length $n_{\max}$. To ensure monomer conservation, all reactions that produce or consume polymers A_n for

Table 2: Model parameters.

Symbol	Significance
k_a	rate constant for activation
k_d1	rate constant for monomer deactivation
k_d2	rate constant for polymer end deactivation
k_d3	rate constant for polymer chain deactivation
k_+	rate constant for polymer assembly
k_{ij}^-	rate constant for polymer disassembly of A_{i+j} into A_i and A_j
n_c	size of the critical nucleus (cooperative model)
K_c	dissociation constant for sub-critical polymers $n < n_\mathrm{c}$ (cooperative model)
K	dissociation constant for super-critical polymers $n \geq n_\mathrm{c}$
k_f	mass transfer coefficient for activating fuel delivery
k_g	mass transfer coefficient for deactivating fuel delivery
f_0	initial and/or inlet concentration for activating fuel
g_0	initial and/or inlet concentration for deactivating fuel
c_tot	total monomer concentration

$n > n_\mathrm{max}$ are excluded from the kinetic equations. The maximum polymer length n_max is chosen as

$$n_\mathrm{max} = 5\sqrt{\frac{c_\mathrm{tot}}{K}\left(\frac{K_\mathrm{c}}{K}\right)^{n_\mathrm{c}-1}} \tag{2.12}$$

which is considerably longer than the average polymer length at equilibrium assuming all monomers are activated. Typically, we considered parameters (i.e., n_c, K_c, K, and c_tot) for which $n_\mathrm{max} < 1000$ to limit the computational burden of numerical integration. Under deactivating conditions, polymer assemblies may be considerably shorter than their equilibrium length, in which case it is possible to reduce n_max without loss of accuracy.

The model is integrated numerically in Python using the `pyodesys` package, which provides access via `pycvodes` to the integration routines from CVODES in the SUNDIALS suite.[3] We use a multistep method based on backward differentiation formula (BDF) designed to integrate stiff equations. The implicit BDF method requires a user supplied function for computing the jacobian matrix. We compute this matrix analytically using `pyodesys`, which relies on the 'sympy' library for symbolic mathematics. As compared to numerical estimates, the use of the analytical jacobian accelerates the speed of integration by two orders of magnitude.

3 Uncoupled Model ($k_{\mathrm{d1}} = k_{\mathrm{d2}} = k_{\mathrm{d3}}$)

When the rate constants for deactivation do not depend on the monomer context—free monomer vs. polymer ends vs. polymer chains—the processes of activation and deactivation decouple from those of assembly and dissassembly. Under these conditions ($k_{\mathrm{d1}} = k_{\mathrm{d2}} = k_{\mathrm{d3}}$), we can describe the transient activation and deactivation of all monomers without reference to the assembly/disassembly processes (1.1). We refer to this simplified model as the "uncoupled model".

3.1 Transient Activation & Deactivation

Kinetic Equations

Starting from the general kinetic equations (2.5)—(2.9), we consider the specific case of a closed batch reactor ($k_{\mathrm{f}} = k_{\mathrm{g}} = 0$) in which the three types of deactivation proceed with a common rate constant ($k_{\mathrm{d1}} = k_{\mathrm{d2}} = k_{\mathrm{d3}} = k_{\mathrm{d}}$). With this simplification, the governing equations simplify as

$$\begin{aligned} \dot{f} &= -k_{\mathrm{a}} f d \\ \dot{g} &= -k_{\mathrm{d}} g m_1 \\ \dot{d} &= -k_{\mathrm{a}} f d + k_{\mathrm{d}} g m_1 \\ \dot{m_1} &= k_{\mathrm{a}} f d - k_{\mathrm{d}} g m_1 \end{aligned} \tag{3.1}$$

where f, g, and d denote, respectively, the concentrations of activating fuel F, deactivating fuel G, and deactivated monomer D. The total concentration of activated monomer is denoted $m_1 = \sum_{n=1}^{\infty} n a_n$, which is unaffected by process of monomer assembly and disassembly. At time $t = 0$, the reactor is charged with specified concentrations

$$f(0) = f_0, \quad g(0) = g_0, \quad d(0) = c_{\mathrm{tot}}, \quad m_1(0) = 0 \tag{3.2}$$

From these initial conditions, equations (3.1) can be integrated to determine the transient concentrations.

Dynamical Invariants

As there are only two reactions (monomer activation and deactivation), the four differential equations are characterized by two invariant quantities that do not change in time. First, as monomers are not created or destroyed but only interconverted between activated and deactivated forms, the total monomer concentration c_{tot} is constant

$$d(t) + m_1(t) = c_{\mathrm{tot}} \tag{3.3}$$

Second, the excess concentration of deactivating fuel g_{ex} relative to that of activating fuel is related to the concentration of active monomers as

$$g(t) - f(t) - m_1(t) = g_0 - f_0 = g_{\mathrm{ex}} \tag{3.4}$$

Using these invariants, we can recast the above problem into two differential equations: one for the concentration of activating fuel $f(t)$ and another for the concentration of

activated monomers $m_1(t)$

$$\begin{aligned} \dot{f} &= -k_\mathrm{a} f(c_\mathrm{tot} - m_1) \\ \dot{m}_1 &= k_\mathrm{a} f(c_\mathrm{tot} - m_1) - k_\mathrm{d}(g_\mathrm{ex} + f + m_1)m_1 \end{aligned} \tag{3.5}$$

subject to initial conditions $f(0) = f_0$ and $m_1(0) = 0$.

Conditions of Interest

We focus our analysis of this system on conditions that lead to the transient activation and subsequent deactivation of the monomer species. Initially, monomer is present only in its deactivated form. At long times ($t \to \infty$), we require that the activating fuel F is completely consumed such that all monomer returns to the deactivated form—that is, $f(\infty) = 0$ and $m_1(\infty) = 0$. This condition requires that $g_\mathrm{ex} = g_0 - f_0 > 0$, such that there is an excess concentration of deactivating fuel relative to that of activating fuel. For simplicity, we consider the limit in which the deactivating fuel is present in large excess such that $g_\mathrm{ex} \gg f_0$. Under these conditions, the concentration of deactivating fuel is well approximated as constant $g(t) \approx g_0$. With these simplifying assumptions, the dynamical equations (3.5) governing the concentrations of activating fuel f and activated monomer m_1 are well approximated as

$$\begin{aligned} \dot{f} &= -k_\mathrm{a} f(c_\mathrm{tot} - m_1) \\ \dot{m}_1 &= k_\mathrm{a} f(c_\mathrm{tot} - m_1) - k'_\mathrm{d} m_1 \end{aligned} \tag{3.6}$$

where $k'_\mathrm{d} = k_\mathrm{d} g_0$ is a pseudo-first order rate constant for deactivation. The initial conditions are $f(0) = f_0$ and $m_1(0) = 0$.

Dimensionless Parameters

In analyzing equation (3.6), it is convenient to measure concentration in units of c_tot, the total monomer concentration, and time in units of $(k_\mathrm{a} c_\mathrm{tot})^{-1}$, the characteristic time scale for activation. This choice of units corresponds to setting $k_\mathrm{a} \to 1$ and $c_\mathrm{tot} \to 1$ in equation (3.6). The resulting model depends on two dimensionless parameters: (1) $k'_\mathrm{d}/k_\mathrm{a} c_\mathrm{tot}$, the ratio between the characteristic rates of activation and deactivation, and (2) f_0/c_tot, the ratio between the initial concentration of activating fuel and the total monomer concentration. Depending on the values of these two parameters, there exist four qualitatively different transient responses as detailed below (Figure 5).

3.2 Dynamical Regimes

Here, we provide a more detailed analysis of the four regimes (I, II, III, and IV) introduced in the main text. The order in which the regimes are described is guided more by mathematical convenience than by chemical intuition or experimental relevance.

Regime IV: $f_0 \ll c_\mathrm{tot}$ and $k'_\mathrm{d} \ll k_\mathrm{a} c_\mathrm{tot}$. In this regime, the fuel concentration is small relative to the monomer concentration; consequently, the monomer activation is small

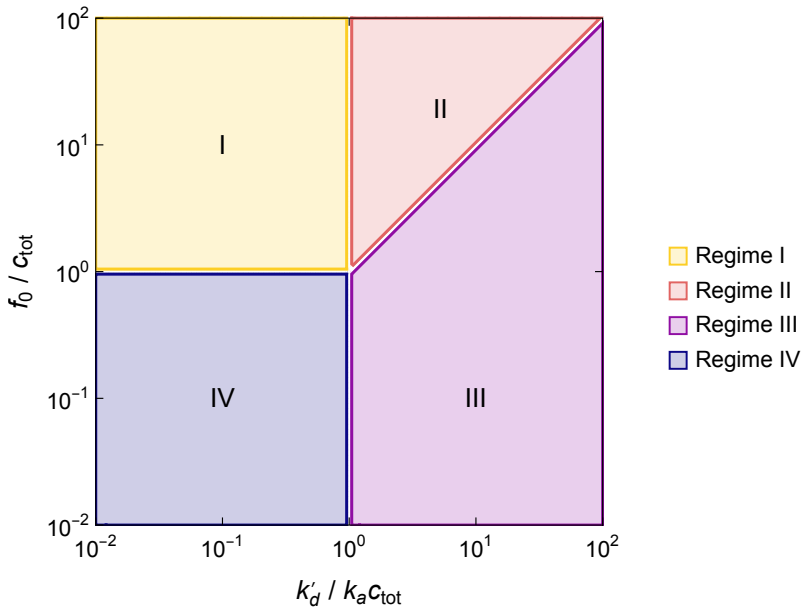


Figure 5: The dynamics of equations (3.6) is characterized by four distinct regimes depending on the magnitude of two dimensionless parameters: $k'_{\mathrm{d}}/k_{\mathrm{a}}c_{\mathrm{tot}}$, the ratio between the characteristic rates of activation and deactivation, and f_0/c_{tot}, the ratio between the concentration of deactivating fuel and the monomer concentration.

$m_1(t) \ll c_{\mathrm{tot}}$. We can therefore linearize equations (3.6) about $f = m_1 = 0$ and integrate the resulting equations to obtain the following approximate solution

$$\begin{aligned} f(t) &= f_0 e^{-k_{\mathrm{a}}c_{\mathrm{tot}}t} \\ m_1(t) &= \frac{k_{\mathrm{a}}c_{\mathrm{tot}}f_0}{k_{\mathrm{a}}c_{\mathrm{tot}} - k'_{\mathrm{d}}}\left(e^{-k_{\mathrm{a}}c_{\mathrm{tot}}t} - e^{-k'_{\mathrm{d}}t}\right) \end{aligned} \tag{3.7}$$

This solution exhibits a peak activation $m_1^\star$ at a time t^* given by

$$t^* = \frac{\ln(k_{\mathrm{a}}c_{\mathrm{tot}}/k'_{\mathrm{d}})}{k_{\mathrm{a}}c_{\mathrm{tot}} - k'_{\mathrm{d}}} \quad \text{and} \quad m_1^* = f_0\left(\frac{k'_{\mathrm{d}}}{k_{\mathrm{a}}c_{\mathrm{tot}}}\right)^{\frac{k'_{\mathrm{d}}}{k_{\mathrm{a}}c_{\mathrm{tot}}-k'_{\mathrm{d}}}} \tag{3.8}$$

For slow deactivation ($k'_{\mathrm{d}} \ll k_{\mathrm{a}}c_{\mathrm{tot}}$), the system achieves a peak activation of $m_1^\star \sim f_0$ at time $t^\star \sim 1/k_{\mathrm{a}}c_{\mathrm{tot}}$ (Figure 6, regime IV). Once the fuel is consumed, the monomer activation slowly decays at a rate k'_{d}.

Regime III: $f_0 \ll k'_{\mathbf{d}}/k_{\mathbf{a}}$ and $k'_{\mathbf{d}} \gg k_{\mathbf{a}}c_{\mathbf{tot}}$. The approximate solution (3.7) is also valid for fast deactivation ($k'_{\mathrm{d}} \gg k_{\mathrm{a}}c_{\mathrm{tot}}$) provided that the monomer concentration is small ($m_1 \ll c_{\mathrm{tot}}$). Under these conditions, the approximate solution (3.7) suggests that the peak activation occurs at time $t^\star \sim 1/k'_{\mathrm{d}}$ with a magnitude $m_1^\star \sim k_{\mathrm{a}}c_{\mathrm{tot}}f_0/k'_{\mathrm{d}}$. Consequently, the condition of small activation ($m_1 \ll c_{\mathrm{tot}}$) requires initial fuel concentrations $f_0 \ll k'_{\mathrm{d}}/k_{\mathrm{a}}$. Under these conditions, the monomer activation increases at an initial rate $k_{\mathrm{a}}f_0$ until reaching a quasi-steady state, at which monomer activation is balanced by monomer deactivation—that is, $k_{\mathrm{a}}c_{\mathrm{tot}}f \sim k'_{\mathrm{d}}a$ (Figure 6, regime III). This balance persists as the fuel is slowly consumed at a rate $k_{\mathrm{a}}c_{\mathrm{tot}}$ and all monomer returns to the deactivated form.

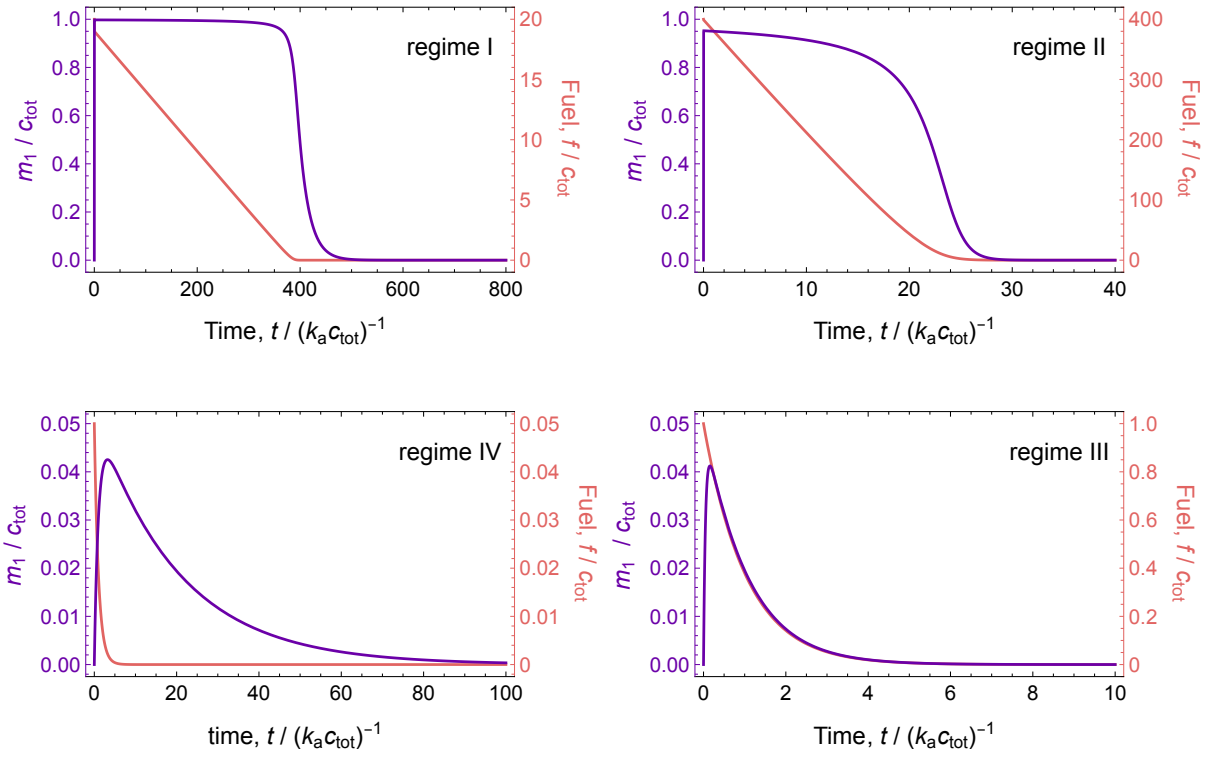


Figure 6: Concentrations of activated monomer m_1 and fuel f as a function of time obtained by numerical integration of equation (3.6) for different dimensionless parameters $f_0/c_{\rm tot}$ and $k'_{\rm d}/k_{\rm a}c_{\rm tot}$. Regime I: $f_0/c_{\rm tot} = 20 \gg 1$ and $k'_{\rm d}/k_{\rm a}c_{\rm tot} = 0.05 \ll 1$. The monomer activation rises quickly to unity and remains high for a time $f_0/k'_{\rm d}c_{\rm tot}$ during which the fuel is steadily consumed. Once the fuel is completely consumed, the activation decays slowly at a rate $k'_{\rm d}$. Regime II: $k_{\rm a}f_0/k'_{\rm d} = 20 \gg 1$ and $k'_{\rm d}/k_{\rm a}c_{\rm tot} = 20 \gg 1$. The monomer activation rises quickly to near unity and remains high for a time $f_0/k'_{\rm d}c_{\rm tot}$ during which the fuel is steadily consumed. Following this period, both the activation and fuel concentration decay together at a rate $k_a c_{\rm tot}$. Regime III: $k_{\rm a}f_0/k'_{\rm d} = 0.05 \ll 1$ and $k'_{\rm d}/k_{\rm a}c_{\rm tot} = 20 \gg 1$. The monomer concentration rises to $k_{\rm a}c_{\rm tot}f_0/k'_{\rm d}$ on a time scale of order $1/k'_{\rm d}$ before slowly decaying at a rate $k_{\rm a}c_{\rm tot}$ as the fuel is consumed. Regime IV: $f_0/c_{\rm tot} = 0.05 \ll 1$ and $k'_{\rm d}/k_{\rm a}c_{\rm tot} = 0.05 \ll 1$. The fuel is rapidly consumed on a time scale $1/k_{\rm a}c_{\rm tot}$, at which the monomer concentration reaches a maximum value of order f_0. Once the fuel is consumed, the monomer slowly deactivates at a rate $k'_{\rm d}$.

Regime I: $f_0 \gg c_{\mathbf{tot}}$ **and** $k'_{\mathbf{d}} \ll k_{\mathbf{a}} c_{\mathbf{tot}}$**.** When the initial fuel concentration is sufficiently high, nearly all the monomer is activated for some finite period of time before the fuel is consumed (Figure 6, regime I). Monomer activation increases to near unity on a time scale of order $1/k_{\mathrm{a}} f_0$, quickly consuming some fraction c_{tot}/f_0 of the available fuel. Once the monomer activation has reached its quasi-steady state where $\dot{m}_1 \approx 0$, the fuel is consumed at a constant rate of $k'_{\mathrm{d}} c_{\mathrm{tot}}$ over a time scale $f_0/k'_{\mathrm{d}} c_{\mathrm{tot}}$. Once the fuel is consumed, the monomer activation slowly decays to zero at a rate k'_{d}.

Regime II: $f_0 \gg k'_{\mathbf{d}}/k_{\mathbf{a}}$ **and** $k'_{\mathbf{d}} \gg k_{\mathbf{a}} c_{\mathbf{tot}}$**.** In this regime, the activation is maintained high for a finite period of time before decaying to zero along with the fuel concentration (Figure 6, regime II). Monomer activation increases to near unity on a time scale of order $1/k_{\mathrm{a}} f_0$ before reaching a quasi-steady state with $k_{\mathrm{a}} f(c_{\mathrm{tot}} - m_1) \approx k'_{\mathrm{d}} m_1$. The fuel concentration decays in time from the initial concentration f_0 to a small concentration $k'_{\mathrm{d}}/k_{\mathrm{a}}$ on a time scale $f_0/k'_{\mathrm{d}} c_{\mathrm{tot}}$. Finally, the activation decays to zero in concert with the remaining fuel on a time scale $1/k_{\mathrm{a}} c_{\mathrm{tot}}$, while maintaining the quasi-steady state activation.

4 Coupled Isodesmic Model ($k_{\mathrm{d1}} \neq k_{\mathrm{d2}} = k_{\mathrm{d3}}$)

4.1 Transient Activation & Deactivation

Kinetic Equations

Starting from the general kinetic equations (2.5)—(2.9), we consider the specific case of a closed batch reactor ($k_{\mathrm{f}} = k_{\mathrm{g}} = 0$) in which the rate constant of deactivation differs for free and assembled monomers ($k_{\mathrm{d1}} \neq k_{\mathrm{d2}} = k_{\mathrm{d3}}$). The simplified kinetic equations are

$$\begin{aligned}
\dot{f} &= -k_{\mathrm{a}} f d \\
\dot{g} &= -k_{\mathrm{d1}} g a_1 - k_{\mathrm{d2}} g \sum_{j=2}^{\infty} j a_j \\
\dot{d} &= -k_{\mathrm{a}} f d + k_{\mathrm{d1}} g a_1 + k_{\mathrm{d2}} g \sum_{j=2}^{\infty} j a_j \\
\dot{a}_1 &= k_{\mathrm{a}} f d - k_{\mathrm{d1}} g a_1 + 2k_{\mathrm{d2}} g \sum_{j=1}^{\infty} a_{j+1} - 2k_+ a_1 \sum_{j=1}^{\infty} a_j + 2k_- \sum_{j=1}^{\infty} a_{j+1} \\
\dot{a}_n &= -k_{\mathrm{d2}} g n a_n + 2k_{\mathrm{d2}} g \sum_{j=1}^{\infty} a_{j+n} - 2k_+ a_n \sum_{j=1}^{\infty} a_j + k_+ \sum_{j=1}^{n-1} a_j a_{n-j} \\
&\quad - k_-(n-1) a_n + 2k_- \sum_{j=1}^{\infty} a_{j+n}
\end{aligned} \tag{4.1}$$

subject to initial conditions

$$f(0) = f_0, \quad g(0) = g_0, \quad d(0) = c_{\mathrm{tot}}, \quad a_n(0) = 0 \tag{4.2}$$

As for isodesmic polymerization in the absence of chemical fuels (see Section 1.3), this model is characterized by a exponential distribution of polymer lengths n, which is fully described by its first two moments.

Partial Moment Evolution Equations

In order to couple the polymerization kinetics to the transient evolution of activated monomers, it is convenient to introduce the partial moments

$$m'_p = \sum_{n=2}^{\infty} n^p a_n \tag{4.3}$$

such that $m_p = a_1 + m'_p$ for $p = 0, 1, \ldots$. Multiplying the kinetic equations (4.1) for the activated species a_n by n^p and summing from $n = 2$ to infinity, one obtains the following

evolution equations

$$\begin{aligned}
\dot{f} &= -k_{\mathrm{a}} f d \\
\dot{g} &= -k_{\mathrm{d1}} g a_1 - k_{\mathrm{d2}} g m_1' \\
\dot{d} &= -k_{\mathrm{a}} f d + k_{\mathrm{d1}} g a_1 + k_{\mathrm{d2}} g m_1' \\
\dot{a}_1 &= k_{\mathrm{a}} f d - k_{\mathrm{d1}} g a_1 + 2k_{\mathrm{d2}} g m_0' - 2k_+ a_1 (a_1 + m_0') + 2k_- m_0' \\
\dot{m}_0' &= k_{\mathrm{d2}} g (m_1' - 4m_0') + k_+ (a_1^2 - m_0'^2) + k_- (m_1' - 3m_0') \\
\dot{m}_1' &= -k_{\mathrm{d2}} g (m_1' + 2m_0') + 2k_+ a_1 (a_1 + m_0') - 2k_- m_0'
\end{aligned} \tag{4.4}$$

subject to initial conditions

$$f(0) = f_0, \quad g(0) = g_0, \quad d(0) = c_{\mathrm{tot}}, \quad a_1(0) = m_0'(0) = m_1'(0) = 0 \tag{4.5}$$

These equations are characterized by two conserved quantities—namely, the total monomer concentration, $d + a_1 + m_1' = c_{\mathrm{tot}}$, and the excess fuel, $g - f - a_1 - m_1' = g_0 - f_0$, such that

$$\begin{aligned}
0 &= \dot{d} + \dot{a}_1 + \dot{m}_1' \\
0 &= \dot{g} - \dot{f} - \dot{a}_1 - \dot{m}_1'
\end{aligned} \tag{4.6}$$

Equations (4.4) can be integrated numerically to determine the coupled dynamics of monomer activation/deactivation and assembly/disassembly.

4.2 Example I: Polymers Inhibit Deactivation ($k_{\mathrm{d2}} = 0$)

To illustrate how monomer assembly/disassembly can influence the kinetics of monomer activation/deactivation, we first consider the limiting scenario in which assembled monomers are fully protected from chemically-fueled deactivation ($k_{\mathrm{d2}} = 0$). As a result of this protective effect, high levels of monomer activation can be achieved under otherwise unfavorable conditions—namely, those of Regime III in Section 3. As in that section, we will assume that the deactivating fuel is present in excess ($g_0 - f_0 \gg f_0$) such that monomer deactivation is characterized by the pseudo-first-order rate constant $k_{\mathrm{d1}}' = k_{\mathrm{d1}} g_0$. With these assumptions, equations (4.4) can be further simplified to obtain the following kinetic equations for the fuel concentration f, the monomer concentration $a = a_1$, and the partial moments m_0' and m_1'

$$\begin{aligned}
\dot{f} &= -k_{\mathrm{a}} f (c_{\mathrm{tot}} - a_1 - m_1') \\
\dot{a}_1 &= k_{\mathrm{a}} f (c_{\mathrm{tot}} - a_1 - m_1') - k_{\mathrm{d1}}' a_1 - 2k_+ a_1 (a_1 + m_0') + 2k_- m_0' \\
\dot{m}_0' &= k_+ (a_1^2 - m_0'^2) + k_- (m_1' - 3m_0') \\
\dot{m}_1' &= 2k_+ a_1 (a_1 + m_0') - 2k_- m_0'
\end{aligned} \tag{4.7}$$

Initially, all monomer is present in the deactivated form such that $a_1(0) = m_0'(0) = m_1'(0) = 0$; the fuel concentration is $f(0) = f_0$. The concentration of deactivated monomer is $d(t) = c_{\mathrm{tot}} - a_1(t) - m_1'(t)$. Note that these equations provide an exact description of this kinetic model involving polymers of all sizes $n \geq 1$.

Dimensionless Parameters

As in Section 3, it is convenient to measure concentrations in units of c_{tot} and time in units of $(k_{\mathrm{a}}c_{\mathrm{tot}})^{-1}$, which corresponds to setting $k_{\mathrm{a}} \to 1$ and $c_{\mathrm{tot}} \to 1$ in equation (4.7). The resulting model depends on four dimensionless parameters: (1) $k'_{\mathrm{d1}}/k_{\mathrm{a}}c_{\mathrm{tot}}$, the ratio between the characteristic rates of activation and deactivation, (2) f_0/c_{tot}, the ratio between the initial concentration of activating fuel and the total monomer concentration, (3) k_-/k_+c_{tot}, the ratio between the dissociation constant k_-/k_+ and the total monomer concentration, and (4) k_+/k_{a}, the ratio between the rate of assembly and activation. As a full exploration of this 4D space is beyond our present scope, we focus instead on a particular region of parameter space to illustrate how coupling between activation and assembly can lead to new behaviors. In particular, we consider Regime III of transient activation (Figure 5), in which deactivation is fast and fuel concentration is low. Here, the assembly of active monomers into polymer assemblies protects them against deactivation thereby fostering further assembly.

Regime III: $f_0 \ll k'_{\mathbf{d}}/k_{\mathbf{a}}$ and $k'_{\mathbf{d}} \gg k_{\mathbf{a}}c_{\mathbf{tot}}$

As detailed in Section 3, "uncoupled" systems in regime III are characterized by a peak concentration of active monomer m_1 (including both free a_1 and polymerized m'_1 monomer) of order $m_1^{\max} \sim k_{\mathrm{a}}c_{\mathrm{tot}}f_0/k'_{\mathrm{d}} \ll c_{\mathrm{tot}}$ where $k'_{\mathrm{d}} = k'_{\mathrm{d1}} = k'_{\mathrm{d2}}$ is the common rate of deactivation for free and polymerized monomer alike. The assembly of active monomers is expected to occur when this peak concentration exceeds the dissociation constant, $m_1^{\max} \gg k_-/k_+$. Additionally, the assembly rate $\sqrt{4k_+k_-m_1^{\max}}$ of equation (1.26) should be fast relative to that of monomer deactivation, which is controlled by fuel consumption at rate $k_{\mathrm{a}}c_{\mathrm{tot}}$ in this regime (Figure 6). These conditions for transient self-assembly can

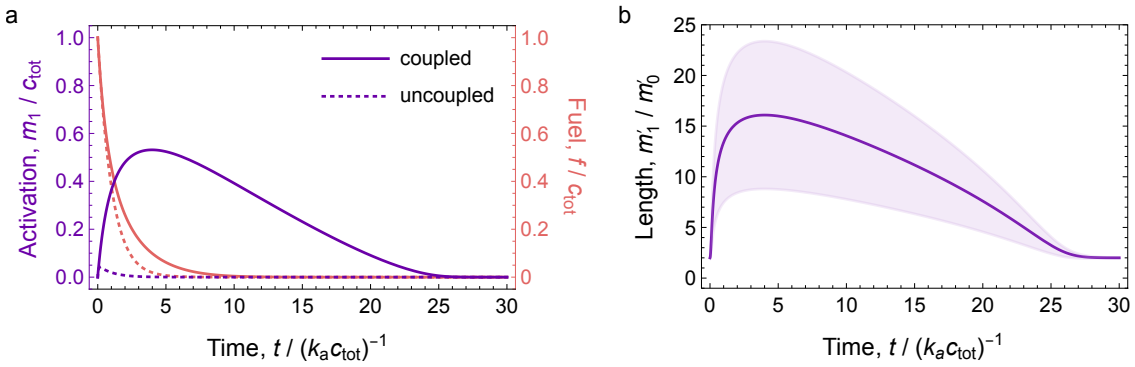


Figure 7: Transient self-assembly under conditions of fast deactivation and low fuel concentration—$k'_{\mathrm{d1}} = 20\,k_{\mathrm{a}}c_{\mathrm{tot}}$ and $f_0 = 1\,c_{\mathrm{tot}}$—corresponding to Regime III in Figure 6. (a) When activated monomers are deactivated at a rate independent of their assembly state ("uncoupled" model), the peak activation is of order $k_{\mathrm{a}}c_{\mathrm{tot}}f_0/k'_{\mathrm{d}}$ as in Figure 6. By contrast, when assembled monomers are protected from deactivation ("coupled" model), the total concentration of activated monomers $m_1(t)$ rises much higher—here, to the order of the initial fuel concentration f_0. The rate constants for assembly and disassembly are, respectively, $k_+ = 400\sqrt{5}\,k_{\mathrm{a}}$ and $k_- = \sqrt{5}\,k_{\mathrm{a}}c_{\mathrm{tot}}$, which satisfy the conditions summarized in equation (4.8). (b) The number average polymer length m'_1/m'_0 as a function of time for the coupled model. The shaded region corresponds to one standard deviation centered on the mean length.

be summarized as

$$\frac{k_-}{k_+c_{\rm tot}} \ll \left(\frac{f_0}{c_{\rm tot}}\right)\left(\frac{k_{\rm a}c_{\rm tot}}{k'_{\rm d}}\right) \quad \text{and} \quad \left(\frac{k_+}{k_{\rm a}}\right)^2 \gg \frac{1}{4}\left(\frac{k_+c_{\rm tot}}{k_-}\right)\left(\frac{k'_d}{k_{\rm a}c_{\rm tot}}\right)\left(\frac{c_{\rm tot}}{f_0}\right) \tag{4.8}$$

in terms of dimensionless parameters $k'_{\rm d}/k_{\rm a}c_{\rm tot}$, $f_0/c_{\rm tot}$, $k_-/k_+c_{\rm tot}$, and $k_+/k_{\rm a}$ introduced above. When these conditions are satisfied as well as those characteristic of regime III, the relatively small amounts of monomer that form are expected to assemble rapidly to form transient polymers, which disassemble when the fuel is consumed.

For "coupled" systems with $k_{\rm d2} = 0$, transient assembly under the above conditions acts to inhibit monomer deactivation thereby increasing the level of activation far beyond that expected for the "uncoupled" system. Figure 7 shows a specific example with parameters $k'_{\rm d1}/k_{\rm a}c_{\rm tot} = 20$ and $f_0/c_{\rm tot} = 1$ characteristic of regime III and identical to the conditions used in Figure 6. Consistent with equations (4.8), we consider conditions favorable to rapid self-assembly with $k_-/k_+c_{\rm tot} = 1/400$ and $k_+/k_{\rm a} = 400\sqrt{5}$. Under these conditions, the concentration of activated monomer m_1 rises more than ten times higher in the coupled model than in the uncoupled model as determined by the available fuel.

4.3 Example II: Polymers Enhance Deactivation ($k_{\rm d1} = 0$)

The local environment provided by the polymer need not inhibit chemical deactivation as in the previous example. The polymer environment can alternatively enhance the rates of deactivation—for example, due to strains induced in the assembled monomer or enrichment of deactivating fuel along the polymer chain (e.g., due to electrostatic interactions). Catalytic effects that accelerate the deactivation of polymerized monomer lead, not surprisingly, to shorter polymers; however, they can also facilitate transient activation under otherwise unfavorable conditions—namely, those of Regime III (Fig. 5).

To demonstrate this effect, we consider the limiting scenario in which monomer can only be deactivated inside polymer assemblies—in other words, there is no deactivation of free monomer, $k_{\rm d1} = 0$. With this simplification, equations (4.4) can be written as

$$\begin{aligned}
\dot{f} &= -k_{\rm a}f(c_{\rm tot} - a_1 - m'_1) \\
\dot{a}_1 &= k_{\rm a}f(c_{\rm tot} - a_1 - m'_1) + 2k'_{\rm d2}m'_0 - 2k_+a_1(a_1 + m'_0) + 2k_-m'_0 \\
\dot{m}'_0 &= k'_{\rm d2}(m'_1 - 4m'_0) + k_+(a_1^2 - m_0'^2) + k_-(m'_1 - 3m'_0) \\
\dot{m}'_1 &= -k'_{\rm d2}(m'_1 + 2m'_0) + 2k_+a_1(a_1 + m'_0) - 2k_-m'_0
\end{aligned} \tag{4.9}$$

where $k'_{\rm d2}$ is the pseudo-first order rate constant for monomer deactivation within the polymer. As above, all monomer is initially present in the deactivated form such that $a_1(0) = m'_0(0) = m'_1(0) = 0$; the fuel concentration is $f(0) = f_0$. The concentration of deactivated monomer is $d(t) = c_{\rm tot} - a_1(t) - m'_1(t)$.

Regime III: $f_0 \ll k'_{\rm d}/k_{\rm a}$ and $k'_{\rm d} \gg k_{\rm a}c_{\rm tot}$

In the absence of monomer deactivation ($k_{\rm d1} = 0$), the concentration of activated monomer is expected to rise in time with a characteristic of either rate $k_{\rm a}f_0$ (regime I) or $k_{\rm a}c_{\rm tot}$ (regime IV) depending on the fuel concentration. Once, however, the activated monomer begin to assemble, they deactivate rapidly under conditions $k'_{\rm d2} \gg k_{\rm a}c_{\rm tot}$ characteristic of regime III. To achieve higher levels of activation under these conditions, the rate of

assembly must be slower than the rate of activation. The assemblies that form catalyze rapid monomer deactivation thereby preventing the growth of long polymers, which may otherwise be favored.

Figure 8 shows a specific realization of this scenario with parameters $k'_{\mathrm{d2}}/k_{\mathrm{a}}c_{\mathrm{tot}} = 20$ and $f_0/c_{\mathrm{tot}} = 1$ characteristic of regime III. We consider a strongly associating polymer with $k_-/k_+c_{\mathrm{tot}} = 1/400$, identical to that of the previous example. The one remaining parameter k_+ characterizes the rate of assembly, which should be slower than that of activation to achieve the highest levels of activation. Here, we consider an intermediate scenario in which these rates are comparable $k_+ = k_{\mathrm{a}}$. Owing to the catalytic activity of the polymers to accelerate their own deactivation, the length of polymers that form is simply that of the dimers (Fig. 8b).

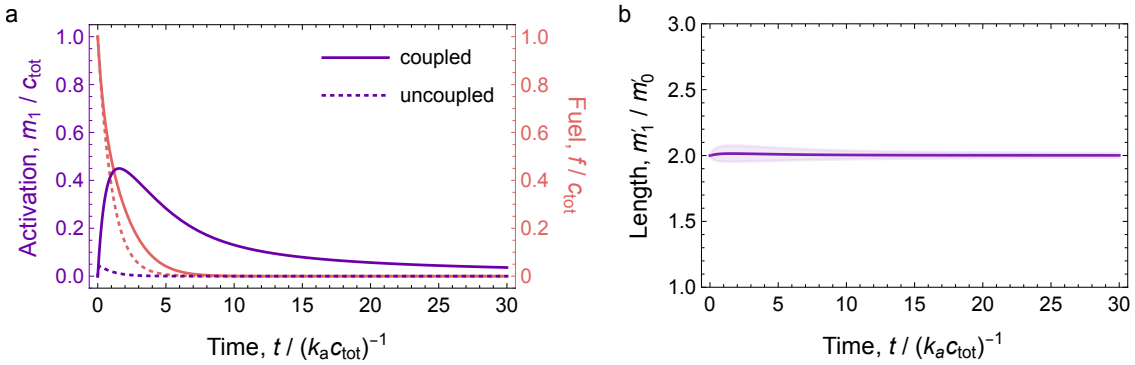


Figure 8: Transient self-assembly under conditions of fast deactivation and low fuel concentration—$k'_{\mathrm{d2}} = 20\,k_{\mathrm{a}}c_{\mathrm{tot}}$ and $f_0 = c_{\mathrm{tot}}$—corresponding to Regime III in Figure 6. (a) Comparison of the "coupled" ($k_{\mathrm{d1}} = 0$) and "uncoupled" ($k_{\mathrm{d1}} = k_{\mathrm{d2}}$) models showing the degree of activation and the fuel concentration as a function of time. In the coupled model, the rate constants for assembly and disassembly are, respectively, $k_+ = k_{\mathrm{a}}$ and $k_- = 0.0025\,k_{\mathrm{a}}c_{\mathrm{tot}}$. (b) The number average polymer length m'_1/m'_0 as a function of time for the coupled model. The shaded region corresponds to one standard deviation centered on the mean length.

5 Coupled Cooperative Model

In this Section, we investigate how chemically fuelled activation and deactivation of monomer units can couple to their cooperative polymerization to create sustained temporal oscillations in the concentration and size of polymer assemblies. We use the cooperative polymerization model of Section 1.4 together with the kinetic model of Section 2.1. Here, the concentrations of activating and deactivating fuel are maintained constant at all times: $f(t) = f_0$ and $g(t) = g_0$. The concentrations of deactivated monomer d, activated monomer a_1, and activated polymer a_n are governed by equations (2.7)—(2.9) with rate constants for disassembly k_{ij}^- given by equation (1.32). Starting from the initial conditions (2.11), the kinetic equations are integrated numerically to determine the transient distribution of polymer sizes (see Section 2.2). This model is characterized by nine parameters:

n_c	critical nucleus size
k_+	rate constant for polymer assembly
K	dissociation constant, e^{ε}
K_c	dissociation constant, $e^{\varepsilon_\mathrm{c}}$
c_tot	total monomer concentration
k'_a	pseudo-first order rate constant for monomer activation, $k_\mathrm{a} f_0$
k'_d1	pseudo-first order rate constant for monomer deactivation, $k_\mathrm{d1} g_0$
k'_d2	pseudo-first order rate constant for polymer end deactivation, $k_\mathrm{d2} g_0$
k'_d3	pseudo-first order rate constant for polymer chain deactivation, $k_\mathrm{d3} g_0$

We set $k_+ \to 1$ and $K \to 1$ in the model such that time is measured in units of $(k_+ K)^{-1}$ and concentration in units of K. To facilitate analysis, we set the critical nucleus size to $n_\mathrm{c} = 2$. With these simplifications, the (dimensionless) model is specified by the six remaining parameters listed above.

We first use numerical simulations to explore the six-dimensional parameter space in search of oscillatory dynamics (Section 5.1). We identify a single region of parameter space for which the kinetic model exhibits sustained oscillations. Within this region, we characterize the system's oscillatory dynamics to determine the dominant kinetic processes and explain the mechanism of polymer oscillations (Section 5.2). Building on this understanding, we develop a simplified 3-variable model based on the monomer concentration a_1 and the polymer moments m'_0 and m'_1 that reproduces the oscillatory dynamics (Section 5.3). Finally, we investigate the mechanism of autocatalytic growth via polymer elongation and chain deactivation to identify the characteristic size of the fastest growing polymer population (Section 5.4).

5.1 Searching Parameter Space for Oscillatory Dynamics

Given a nonlinear kinetic model, it is nontrivial to determine whether or not the model has one or more oscillatory solutions (so-called limit cycles) for some values of the model parameters and initial conditions. Moreover, for all but the most experienced human experts, our ability to intuit model conditions that produce oscillatory dynamics is quite limited—increasingly so as the complexity of the model grows. Fortunately, computer

simulations can be used to explore high dimensional parameter spaces in search of desired kinetic behaviors such as oscillations.

We use Latin hypercube sampling (LHS) to generate five million points sampled from the space of possible parameter values. As the model parameters are positive quantities spanning several orders of magnitude, we sample the logarithm of each parameter from a uniform distribution over a specified range. LHS is implemented using the surrogate modeling toolbox (SMT) package in Python. This sampling method is generally preferred over both random sampling and regular grids as the resulting parameter values are more evenly distributed—particularly when projected onto subspaces of lower dimension.[4]

For each sampled parameter value, we integrate the kinetic model from time $t = 0$ to $t_{\max}$ to determine the transient concentrations $d(t)$, $a_1(t)$, and $a_n(t)$ for $n = 2, \ldots, n_{\max}$ as described in Section 2.2. To identify oscillatory dynamics, we evaluate an appropriate statistic from the computed solution. Specifically, we count the number of local maxima in the concentration of activated monomers $a_1(t)$ using the `find_peak()` function in the SciPy library. A local maximum is considered a concentration "peak" when its height rises more than 10% above the mean value. The number of peaks N_{peaks} provides a convenient measure with which to identify oscillatory solutions with $N_{\text{peaks}} > 3$. To distinguish sustained oscillations (stable limit cycles) from transient oscillations (spiral attractors), we exclude solutions where the peak height changes by 1% or more between the last two peaks. We set the maximum time $t_{\max} = 10$ and polymer length $n_{\max} = 200$ to accelerate the computation, thereby prohibiting the discovery of slow oscillations involving longer polymers.

The goal of this heuristic search process is to identify "interesting" regions of parameter space where sustained oscillations are most likely to occur. These regions can then be investigated more closely to establish the existence of limit cycle(s) and to understand the chemical mechanisms that enable them. The absence of oscillations in other regions provides evidence against their existence but not proof owing to the finite density of sampled points and the limitations of the N_{peak} statistic.

Figure 9 shows the oscillation statistic N_{peak} for the five million simulations projected onto 2D planes formed by each pair of model parameters. While the vast majority of simulations show no evidence for oscillations ($N_{\text{peak}} < 3$; grey markers), there are some examples of damped oscillations (purple markers) and even fewer of sustained oscillations (magenta-orange-yellow). From the simulation data, we can identify several conditions on the model parameters that should be satisfied in order to achieve oscillatory dynamics. First, the total monomer concentration must be large ($c_{\text{tot}} \gg K$) such that activated monomers interact to form polymer assemblies. Second, the dissociation constant for dimers must be large ($K_{\text{c}} \gg K$) such that polymerization is highly cooperative. Third, the rate constants for monomer activation k'_{a}, polymer end deactivation k'_{d2} , and polymer chain deactivation k'_{d3} should be comparable to that of disassembly k_+K; sustained oscillations are not observed when these rate constants are much larger or smaller. Finally, the rate constant k'_{d1} for the deactivation free monomer should be much slower than that of the other processes (e.g., $k'_{\text{d1}} \ll k_+$).

Having identified parameter regions in which oscillations are most likely to occur, we repeat the analysis in smaller region with a higher sampling density to more clearly establish the conditions needed for sustained oscillations (Figure 10). Based on this exploration, there is a single region of parameter space where sustained oscillations are possible—particularly, that region centered on $c_{\text{tot}} \sim 10^4 K$, $k_{\text{d3}} \sim k_+$, $k_{\text{d2}} \sim 10^3 k_+$,

Source code and GUI are available here:
https://github.com/bishopgroup/ChemRevOscillator

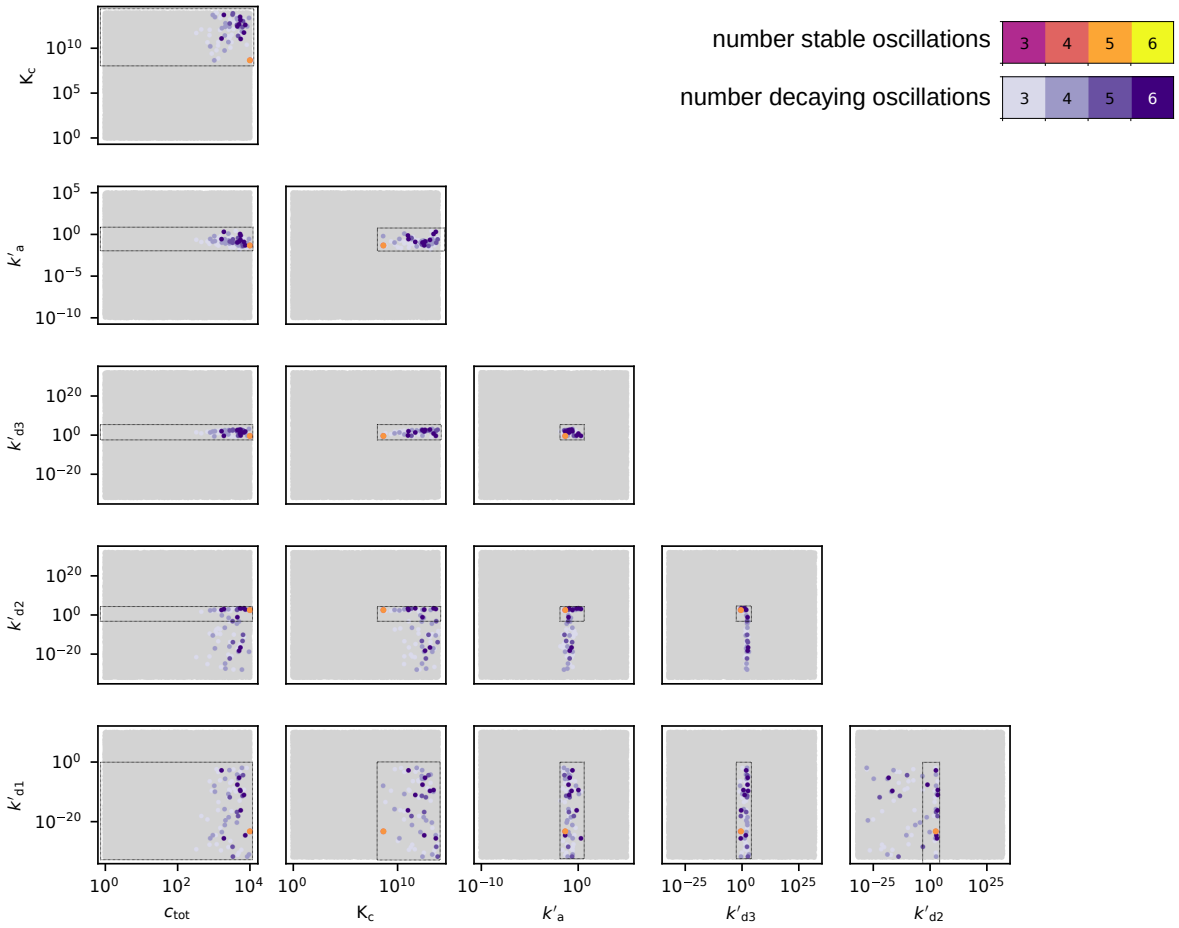


Figure 9: Summary of the oscillation statistic N_{peak} for five million (5×10^6) simulations with model parameters sampled from the full 6-dimensional parameter space. Each plot shows the five million points projected onto 2D subspaces formed by the 15 different parameter pairs. Most points (grey) exhibit less than three peaks in the transient monomer concentration $a_1(t)$. Few points (purple) show three or more peaks of decreasing height indicating damped oscillations. Only a small fraction of the five million points (magenta-orange-yellow) show sustained oscillations with three of more peaks that do not decay in time. Here, the critical nucleus size is set to $n_c = 2$; the dissociation constant K and the assembly rate constant k_+ are set to one which specifies the units for concentration and time used in the simulations.

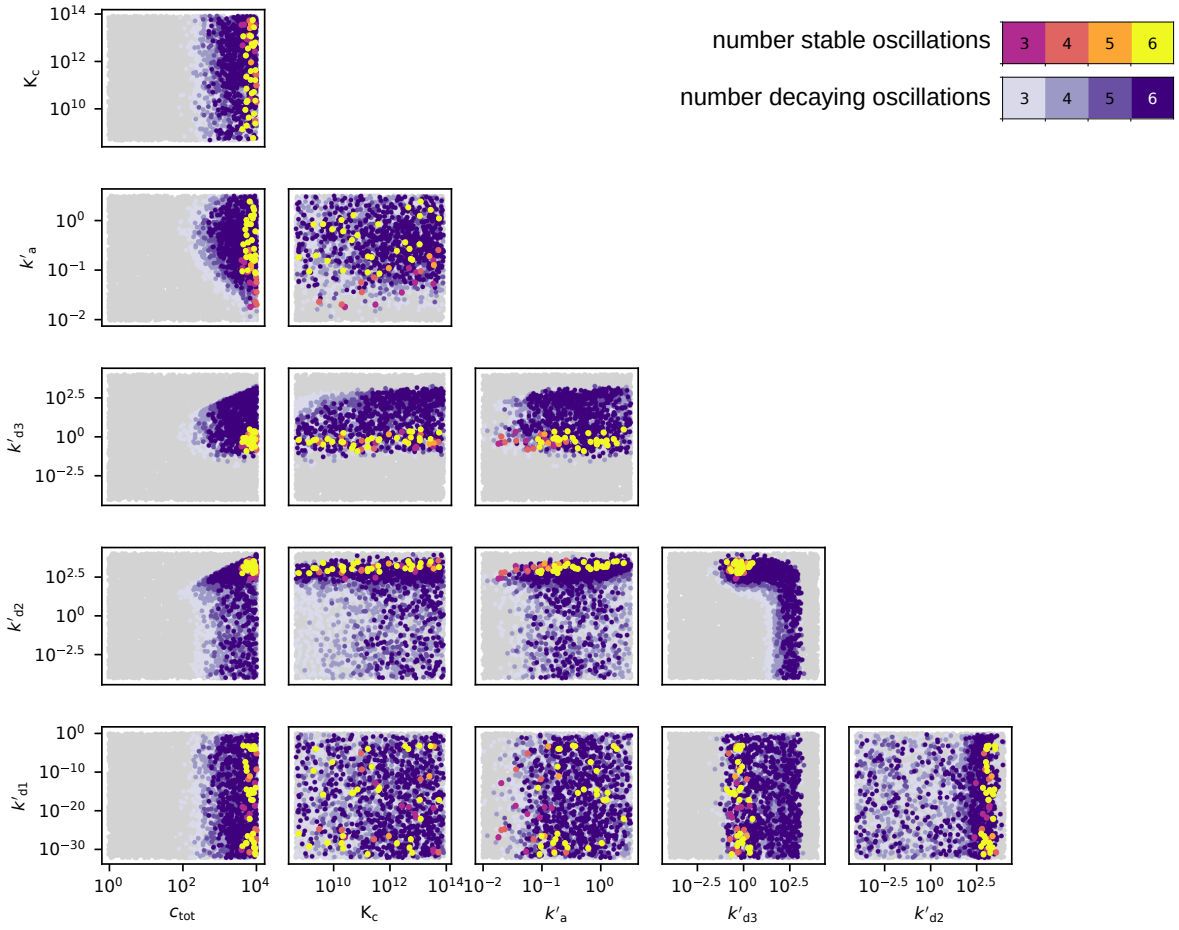


Figure 10: Summary of the oscillation statistic N_{peak} for one hundred thousand (1×10^5) simulations with model parameters sampled from a smaller 6-dimensional parameter space denoted by the dashed boxes in Figure 9.

$k_a \sim 0.2k_+$, and $k'_{d1} \ll k'_a$. Through brute force computation, this search process has made an significant discovery. It shows that the proposed model is capable of sustained oscillations, and it inform us about the conditions necessary for these dynamics. The computational search does not, however, explain how these oscillations occur or why they require the conditions that they do.

5.2 Understanding Sustained Oscillations

Our computational exploration of parameter space reveals a single region in which the system exhibits sustained oscillations in the species concentrations. Here, we investigate more closely a particular set of conditions—namely, $K_c = 10^{11}$, $c_{tot} = 10^4$, $k'_a = 0.2$, $k'_{d1} = 0$, $k'_{d2} = 10^3$, $k'_{d3} = 1$—to characterize the oscillations and explain their origins.

Figure 11 shows the oscillatory dynamics of the different monomer concentrations: deactivated $d(t)$, activated $a_1(t)$, and polymerized $m'_1(t) = \sum_{n=2}^{\infty} n a_n(t)$. During each oscillation cycle, there is a slow phase of monomer activation, during which deactivated monomers are converted into activated monomers. During this activation phase, polymers begin to grow and divide at an accelerating rate. As we will show, the combination of polymer elongation (less end deactivation) and chain deactivation create an autocatalytic process of growth and division fuelled by the activated monomer. Once ignited, this autocatalytic process rapidly consumes activated monomer causing a sharp decrease in its concentration $a_1(t)$. When this concentration falls below a critical value (to be determined), the competition between polymer elongation and end deactivation shifts in favor of disassembly. The remaining polymers are quickly destroyed by chemically fuelled deactivation processes, and the activation phase begins a new.

The plots in Figure 12 characterize the oscillatory dynamics of the polymer size distribution. The partial moments $m'_0(t)$ and $m'_1(t)$ exhibit periodic bursts in the number of polymer chains and the number of polymerized monomers, respectively (Fig. 12a). The ratio of these quantities provides a measure of the polymer length, $L(t) = m'_1(t)/m'_0(t)$,

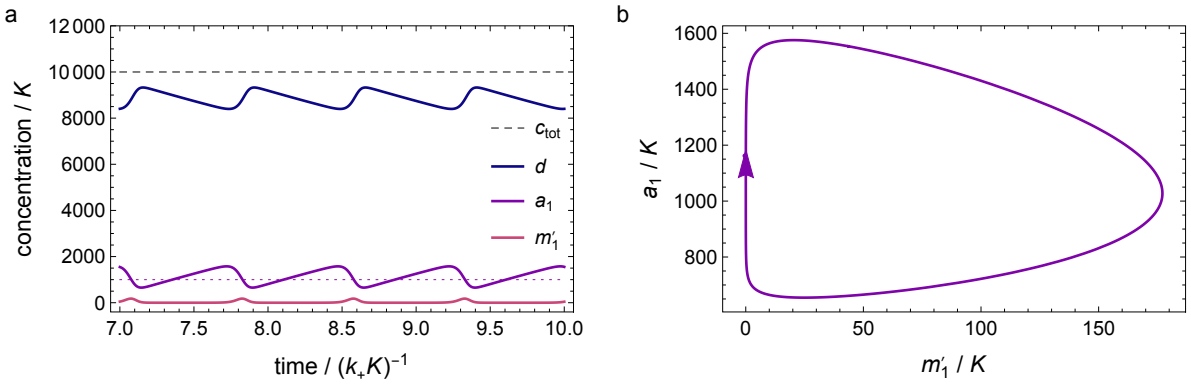


Figure 11: Oscillatory dynamics of the coupled cooperative model with parameters: $K_c = 10^{11}$, $c_{tot} = 10^4$, $k'_a = 0.2$, $k'_{d1} = 0$, $k'_{d2} = 10^3$, $k'_{d3} = 1$. (a) Concentrations of deactivated monomer $d(t)$, activated monomer $a_1(t)$, and polymerized monomer $m'_1 = \sum_{n=2}^{\infty} n a_n$ as a function of time showing sustained oscillations with a period of 0.751. (b) Stable limit cycle behavior projected on the a_1-m'_1 plane. Here, all concentrations are scaled by K; times and rates are scaled by $(k_+K)^{-1}$ and k_+K, respectively. See also Supporting Video 2.

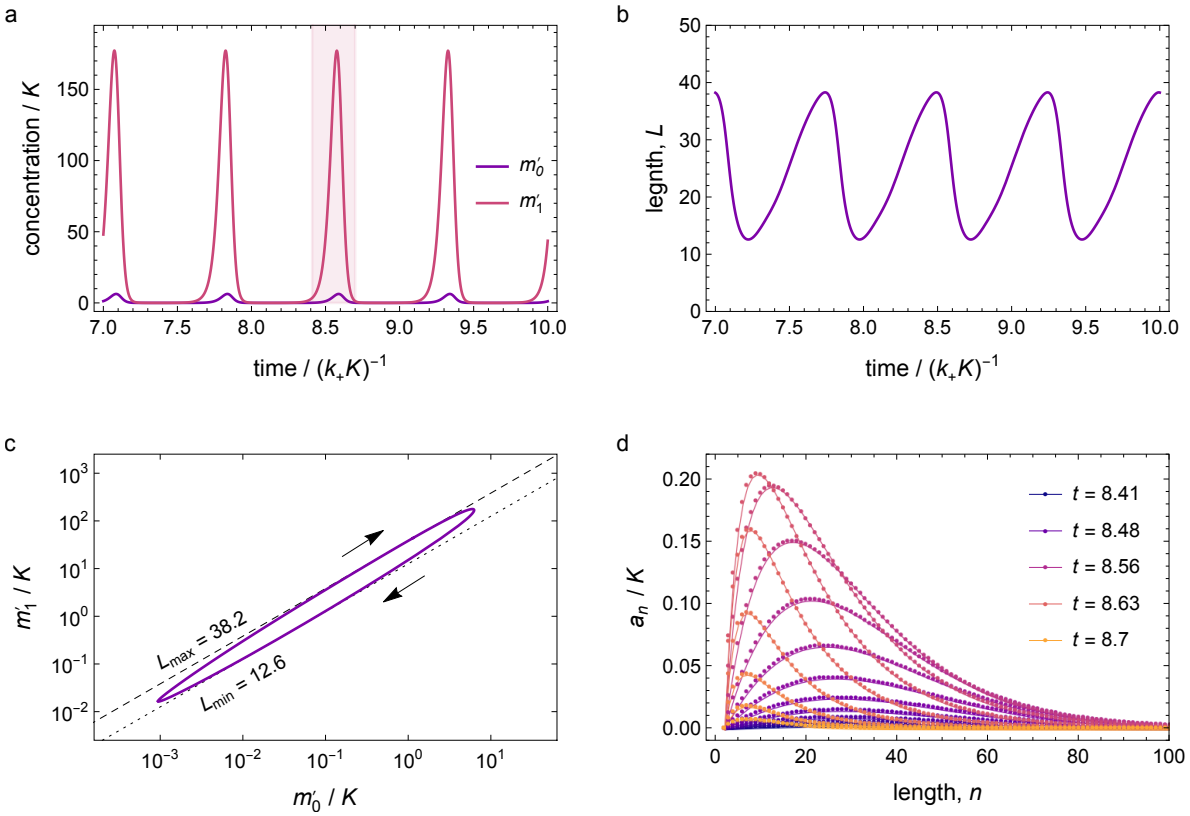


Figure 12: Oscillatory polymer dynamics in the coupled cooperative model with parameters: $K_c = 10^{11}$, $c_{tot} = 10^4$, $k'_a = 0.2$, $k'_{d1} = 0$, $k'_{d2} = 10^3$, $k'_{d3} = 1$—identical to those of Figure 11. (a) Concentrations of polymer chains $m'_0 = \sum_{n=2}^{\infty} a_n$ and polymerized monomer $m'_1 = \sum_{n=2}^{\infty} n a_n$ as a function of time showing sustained oscillations with repeated bursts of polymerization. (b) Mean polymer length $L = m'_1/m'_0$ as a function of time shows the repeated growth and decay of polymer chains during each oscillation cycle. (c) Stable limit cycle projected onto the m'_1-m'_0 plane showing polymer growth and decay. During the growth phase, polymers are characterized by a maximum length L_{max} set by the balance of polymer elongation and division by polymer chain deactivation. During the decay phase, polymers are characterized by a smaller length L_{min} as polymers disassemble by polymer end deactivation. (d) Polymer distribution $a_n(t)$ at regular time intervals during a single burst of polymer growth and decay corresponding to the shaded in (a). See also Supporting Video 2.

which oscillates in time over a range spanning $L_{min} = 12.6$ to $L_{max} = 38.2$ for the present parameters (Fig. 12b). Polymers are longest during the growth phase as the polymer concentration rises and shortest during the decay phase as the polymer concentration falls (Fig. 12c). Figure 12d shows the polymer size distribution $a_n(t)$ as a function of time as the polymer concentration rises and falls during a single oscillation cycle.

Moment Equations

The system dynamics can be described in terms of the partial moments, $m'_0 = \sum_{n=2}^{\infty} a_n$ and $m'_1 = \sum_{n=2}^{\infty} n a_n$, and the concentrations, a_1 and d, of activated and deactivated monomer free in solution. The concentration of deactivated monomer d changes in time as

$$\dot{d} = -k'_a d + k'_{d1} a_1 + 2k'_{d2} m'_0 + k'_{d3}(m'_1 - 2m'_0) \tag{5.1}$$

due to losses by monomer activation and to gains by monomer deactivation, polymer end deactivation, and polymer chain deactivation, respectively. Note that $2m_0'$ is the concentration of polymer ends, and $m_1' - 2m_0'$ is the concentration of non-terminal monomers with the polymer chains. Figure 13a shows the respective contributions of these four processes during a single oscillation cycle. The dominant contributions are losses due to monomer activation and gains due to polymer end deactivation. Monomer activation proceeds a steady rate throughout the oscillation cycle whereas polymer end deactivation occurs only during the transient burst of polymer growth and decay.

The concentration of activated monomer a_1 changes in time as

$$\begin{aligned}\dot{a}_1 = {} & k_a' d - k_{d1}' a_1 + 2k_{d2}' a_2 + 2k_{d3}'(m_0' - a_2) \\ & - 2k_+ a_1^2 + 2k_+ K_c a_2 - 2k_+ a_1 m_0' + 2k_+ K(m_0' - a_2)\end{aligned} \tag{5.2}$$

due to monomer activation gains, monomer deactivation losses, polymer end deactivation gains, polymer chain deactivation gains, nucleation losses, and de-nucleation gains, elongation losses, and de-elongation gains, respectively. During each oscillation cycle,

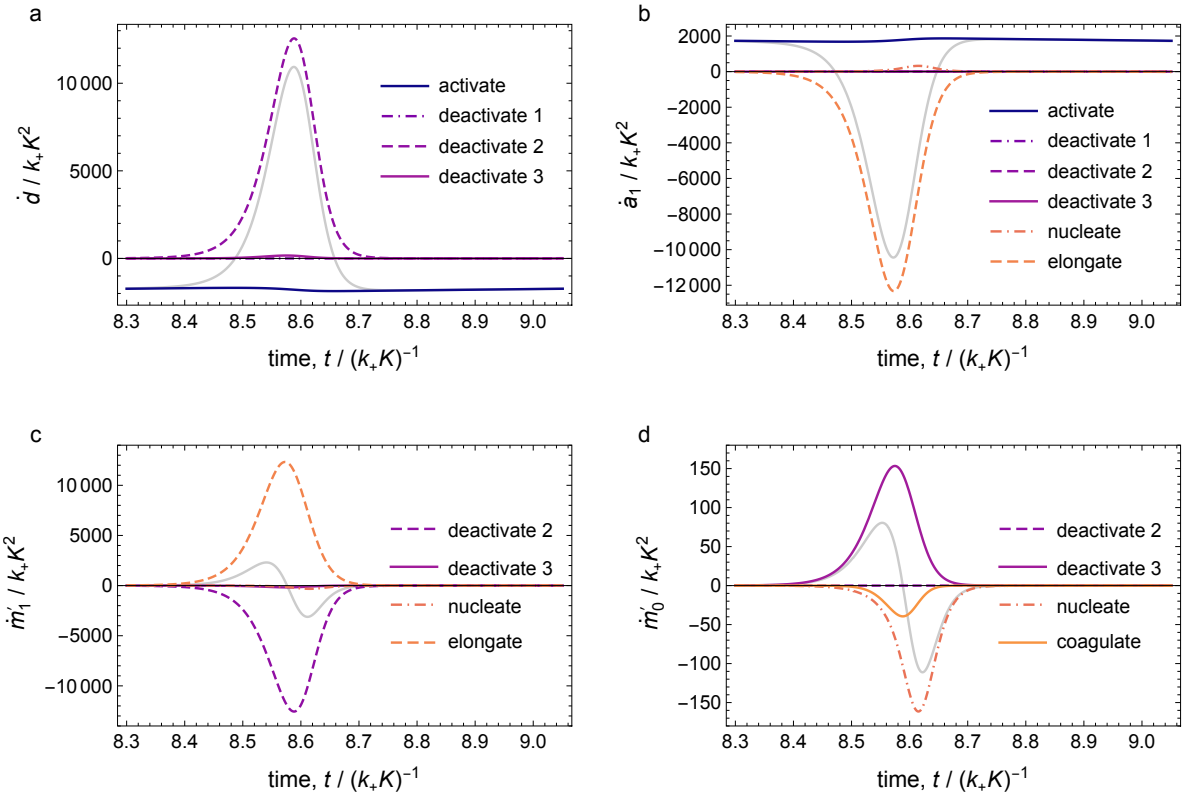


Figure 13: Rates of change in the species concentrations (grey curves) as a function of time highlighting the dominant contributions (colored curves) in equations (5.1)—(5.4) through a single oscillation cycle. The conditions are identical to those in Figures 11 and 12. (a) The deactivated monomer concentration d changes due primarily to gains by polymer end deactivation (denoted deactivate 2) and losses by monomer activation. (b) The activated monomer concentration a_1 changes due primarily to gains by monomer activation and losses by polymer elongation. (c) The polymerized monomer concentration m_1' changes due primarily to gains by polymer elongation and losses by polymer end deactivation. (d) The total polymer concentration m_0' changes due primarily to gains by polymer chain deactivation (denoted deactivate 3) and losses by de-nucleation and coagulation.

the dominant processes include the steady activation of deactivated monomers and the transient removal of monomers by elongation of the growing polymers (Fig. 13b). While comparably less important, polymer disassembly during the decay phase leads to transient monomer gains due to de-nucleation processes (Fig. 13b).

Multiplying the equation (2.9) for the polymer concentration a_n by the polymer length n and summing from $n = 2$ to infinity, one obtains the following rate equation for the concentration of polymerized monomer m_1'

$$\begin{aligned}\dot{m}_1' = &-2k_{\mathrm{d2}}'(a_2 + m_0') - k_{\mathrm{d3}}'(m_1' - 2a_2) \\ &+ 2k_+a_1^2 - 2k_+K_ca_2 + 2k_+a_1m_0' - 2k_+K(m_0' - a_2)\end{aligned} \tag{5.3}$$

This expression includes polymer end deactivation losses, polymer chain deactivation losses, nucleation gains, de-nucleation losses, elongation gains, and de-elongation losses, respectively. Note that equations (5.1)—(5.3) sum to zero, $\dot{d} + \dot{a}_1 + \dot{m}_1' = 0$, as required by monomer conservation. Consequently, the dominant processes governing the concentration of polymerized monomer m_1' are losses by polymer end deactivation and gains by polymer elongation. The system dynamics is characterized by a cyclic flux of monomers from the deactivated form to the activated form (via monomer activation) to the polymerized form (via polymer elongation) and back to the deactivated form (via polymer end deactivation). As discussed below, this cyclic flux is a necessary but insufficient condition for obtaining temporal oscillations in the species concentrations.

Similarly, we can sum equation (2.9) for the polymer concentration a_n from $n = 2$ to infinity to obtain the following rate equation for the total polymer concentration m_0'

$$\begin{aligned}\dot{m}_0' = &-2k_{\mathrm{d2}}'a_2 + k_{\mathrm{d3}}'(m_1' - 4m_0' + 2a_2) + k_+a_1^2 - k_+K_{\mathrm{c}}a_2 \\ &- k_+{m_0'}^2 + k_+\tfrac{K^2}{K_{\mathrm{c}}}(m_1' - 3m_0' + a_2)\end{aligned} \tag{5.4}$$

which includes polymer end deactivation losses, polymer chain deactivation gains, nucleation gains, de-nucleation losses, polymer coagulation losses, and polymer coagulation gains. There are three significant contributions to dynamics of the polymer number concentration m_0': gains by polymer chain deactivation as polymers divide in two, net losses by nucleation and de-nucleation as polymers disassemble, and, to a lesser extent, net losses by coagulation and de-coagulation as polymers assemble to form longer chains (Fig. 13d).

Notably, equations (5.1)–(5.4) for the concentrations of deactivated monomer d, activated monomer a_1, polymerized monomer m_1', and total polymer m_0' cannot be solved in closed form owing to their dependence on the dimer concentration a_2. Moreover, the rate equation for the dimer concentration depends in turn on the concentration of trimers a_3

$$\begin{aligned}\dot{a}_2 = &\,2k_{\mathrm{d2}}'(a_3 - a_2) + 2k_{\mathrm{d3}}'(m_0' - a_2 - a_3) + k_+a_1^2 - k_+K_{\mathrm{c}}a_2 \\ &- 2k_+a_1a_2 + 2k_+Ka_3 - 2k_+a_2m_0' + 2k_+\tfrac{K^2}{K_{\mathrm{c}}}(m_0' - a_2 - a_3)\end{aligned} \tag{5.5}$$

This rate equation describes polymer end deactivation gains, polymer chain deactivation gains, nucleation gains, de-nucleation losses, elongation losses, de-elongation gains, coagulation losses, and de-coagulation gains, respectively. To avoid an endless regress of equations, we must introduce some simplifying assumptions to derive an approximate system of equations governing the dynamics of the monomer concentrations d and a_1 and the polymer moments m_0' and m_1'.

5.3 Simplified Model of Polymer Oscillations

Guided by the dominant mechanisms identified in Figure 13, we can write the following approximate equations for the dynamics of the different monomer concentrations d, a_1, and m_1' as well as the polymer concentration m_0'

$$\begin{aligned} \dot{d} &= -k_{\mathrm{a}}' d + 2k_{\mathrm{d2}}' m_0' \\ \dot{a}_1 &= k_{\mathrm{a}}' d - 2k_+ a_1 m_0' \\ \dot{m}_1' &= -2k_{\mathrm{d2}}' m_0' + 2k_+ a_1 m_0' \\ \dot{m}_0' &= k_{\mathrm{d3}}'(m_1' - 4m_0') - F(a_1, m_1', m_0') \end{aligned} \tag{5.6}$$

Here, we have excluded all but the dominant contributions from the exact equations derived in the previous sections . The function $F(a_1, m_1', m_0')$ is introduced to describe the removal of polymers from the system, which occurs primarily by de-nucleation, nucleation, and coagulation processes (Fig. 13d)

$$F(a_1, m_1', m_0') \approx k_+ K_c a_2 - k_+ a_1^2 + k_+ {m_0'}^2 \tag{5.7}$$

To estimate the net rate of nucleation, we apply the quasi-steady-state approximation (QSSA) on the dimer concentration ($\dot{a}_2 = 0$) and substitute the result into equation (5.7) to obtain

$$F(a_1, m_1', m_0') \approx 2k_{\mathrm{d2}}' a_3 + 2k_{\mathrm{d3}}' m_0' + k_+ {m_0'}^2 \tag{5.8}$$

where we have again neglected all but the dominant contributions.

Finally, to close this approximation, we require an estimate of the trimer concentration a_3 in terms of the monomer concentration a_1 and the polymer moments m_0' and m_1'. To this end, we approximate the polymer length distribution by a continuous function that varies in time due only to changes in the moments $m_0'(t)$ and $m_1'(t)$ as

$$a_n(t) \approx \frac{m_0'(t)}{m_1'(t)/m_0'(t) - 2}\, G\left(\frac{n-2}{m_1'(t)/m_0'(t) - 2}\right) \tag{5.9}$$

The distribution $G(x)$ is assumed to have the following properties

$$\int_0^\infty G(x)dx = 1 \quad \text{and} \quad \int_0^\infty xG(x)dx = 1 \quad \text{and} \quad G(x) \to Cx \text{ as } x \to 0 \tag{5.10}$$

where the final condition implies that the dimer concentration is vanishingly small. The form of this continuous distribution ensures that the first two moments are equal to m_0' and m_1'. The continuum approximation is appropriate for long polymers with characteristic length $L = m_1'/m_0' \gg 2$. Setting $n = 3$ in equation (5.9), the trimer concentration is approximated as

$$a_3(t) \approx \frac{C m_0'(t)}{(m_1'(t)/m_0'(t) - 2)^2} \tag{5.11}$$

where C is a positive order one constant that characterizes the slope of the distribution $G(x)$ for small arguments.

With these preliminaries, the evolution of the polymer concentration m_0' in equation (5.6) is approximated as

$$\dot{m}_0' = k_{\mathrm{d3}}'(m_1' - 4m_0') - \frac{2k_{\mathrm{d2}}' C m_0'}{(m_1'/m_0' - 2)^2} - 2k_{\mathrm{d3}}' m_0' - k_+ {m_0'}^2 \tag{5.12}$$

For the value $C = 4.6$, this simplified system of equations reproduces the oscillatory dynamics with a period of $0.75(k_+ K)^{-1}$ as observed in the full numerical simulations (Fig. 14).

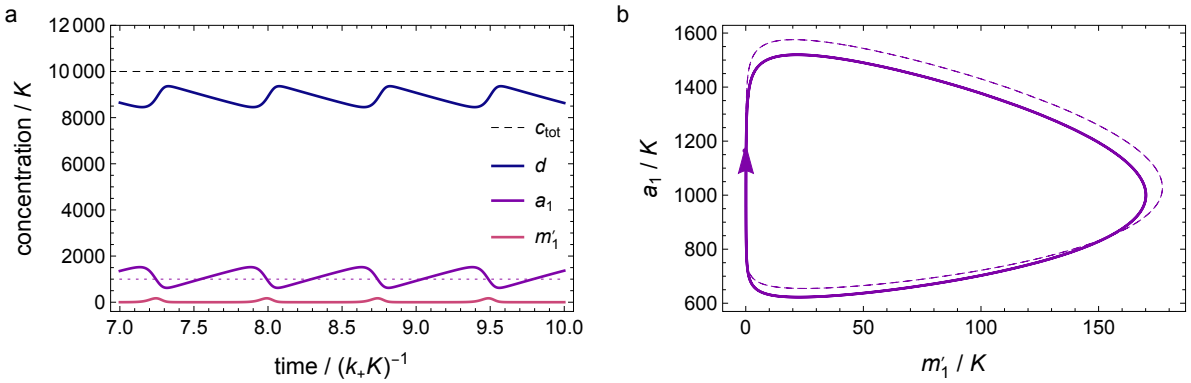


Figure 14: Oscillatory dynamics of the simplified model of equations (5.6) and (5.12) for the same parameters used in Figure 11. The constant C in equation (5.12) use to close the model is set to $C = 4.6$. (a) Concentrations of deactivated monomer $d(t)$, activated monomer $a_1(t)$, and polymerized monomer $m_1'(t)$ as a function of time showing sustained oscillations with a period of 0.75. (b) Stable limit cycle behavior of the simple model (solid curve) and the full model (dashed curve) projected on the a_1-m_1' plane. Concentrations are scaled by K; times and rates are scaled by $(k_+K)^{-1}$ and k_+K, respectively.

5.4 Autocatalysis in Polymer Growth and Division

Polymers form by an autocalytic process in which chains grow by monomer attachment at polymer ends (elongation) and divide by polymer chain deactivation. Chain division enhances growth through the creation of terminal monomers at the polymer ends; growth enhances division through the creation of non-terminal monomers susceptible to chemically fueled deactivation. Together, these processes drive a positive feedback loop that continues until activated monomers become sufficiently depleted, and the rate of elongation falls below that of polymer end deactivation.

Mathematically, this autocatalytic process can be quantified in the limit of small polymer concentrations ($m_0' \ll a_1$) by linearizing equation (2.9) about $a_n = 0$ for a specified monomer concentration a_1 and calculating the dominant eigenmode. For dimers with $n = 2$, the linearized equation is

$$\begin{aligned}\dot{a}_2 = &-2k_{\mathrm{d2}}'(a_2 - a_3) + 2k_{\mathrm{d3}}' \sum_{j=4}^{\infty} a_j + k_+ a_1^2 - k_+ K_{\mathrm{c}} a_2 \\ &- 2k_+ a_1 a_2 + 2k_+ K a_3 - \cancel{2k_+ a_2 \sum_{j=2}^{\infty} a_j} + 2k_+ \tfrac{K^2}{K_{\mathrm{c}}} \sum_{j=4}^{\infty} a_j\end{aligned} \tag{5.13}$$

which includes losses and gains by polymer end deactivation, gains by polymer chain deactivation, gains by nucleation, losses by de-nucleation, losses by elongation, gains by de-elongation, losses by coagulation (neglected), and gains by de-coagulation, respectively. The rate of coagulation losses are nonlinear in the polymer concentration and negligible for small concentrations. For cooperative polymerization with $K_{\mathrm{c}} \gg K$ (as needed for oscillations), gains due to de-coagulation are also negligible despite their linear dependence on the polymer concentration.

For longer chains with $n > 2$, the linearized equations governing the polymer concen-

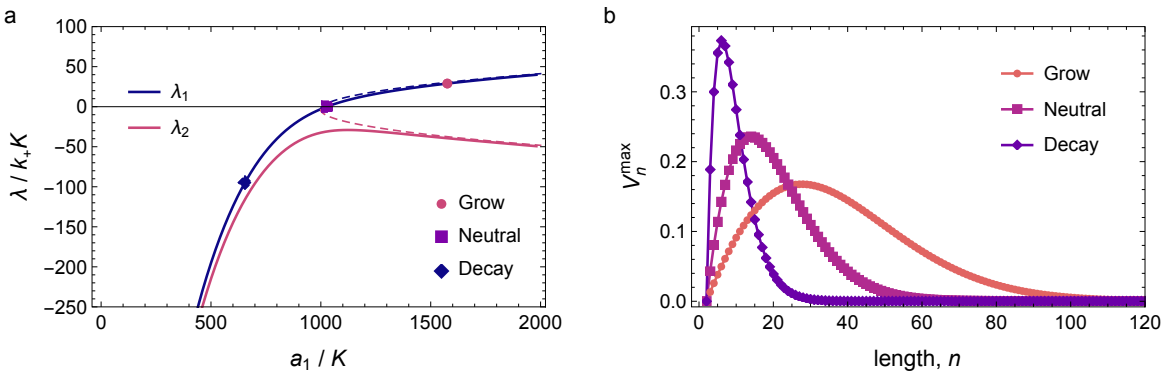


Figure 15: (a) Two largest eigenvalues of equation (5.14) as a function of the monomer concentration a_1 for parameters: $K_c = 10^{11}$, $c_{tot} = 10^4$, $k'_a = 0.2$, $k'_{d1} = 0$, $k'_{d2} = 10^3$, $k'_{d3} = 1$—identical to those of Figure 11. Above critical concentration—here, $a_1^* = 1027\,K$ (square marker)—the largest eigenvalue λ_1 is positive corresponding to exponential growth in the polymer concentration. (b) Eigenmodes V_n^{max} associated with the largest eigenvalue λ_1 at three different monomer concentrations: $a_1 > a_1^*$ (grow), $a_1 = a_1^*$ (neutral), and $a_1 < a_1^*$ (decay); the corresponding eigenvalues are denoted by the markers in (a).

trations a_n are

$$\begin{aligned}\dot{a}_n = &-2k'_{d2}(a_n - a_{n+1}) - k'_{d3}\left[(n-2)a_n - 2\sum_{j=2}^{\infty} a_{j+n}\right] - 2k_+a_1(a_n - a_{n-1}) \\ &- 2k_+K(a_n - a_{n+1}) - k_+\tfrac{K^2}{K_c}\left[(n-3)a_n - 2\sum_{j=2}^{\infty} a_{j+n}\right]\end{aligned} \tag{5.14}$$

which includes losses and gains by polymer end deactivation, polymer chain deactivation, elongation, de-elongation, and de-coagulation, respectively. Contributions due to coagulation (not shown) are neglected owing to their nonlinear dependence on the polymer concentration.

The linear system of equations (5.13)—(5.14) admits solutions of the form

$$a_n(t) = \sum_{j=1}^{\infty} w_j V_{nj} e^{\lambda_j t} \tag{5.15}$$

where V_{nj} is the n^{th} element of the j^{th} eigenvector and λ_j is the corresponding eigenvalue. The solution is a weighted average of these independent eigenmodes with weights w_j determined by the initial conditions. At long times, the solution is well approximated by the eigenmode with the largest eigenvalue λ_{max}, which describes the fastest exponential growth for $\lambda_{max} > 0$ or the slowest exponential decay for $\lambda_{max} < 0$. The corresponding eigenvector V_n^{max} describes the polymer size distribution for this dominant mode.

Figure 15a shows how the largest eigenvalue λ_{max} changes with the monomer concentration a_1 under conditions that produce oscillations. Below a critical monomer concentration, λ_{max} is negative, and the polymer concentrations decay to zero. Above this critical concentration, λ_{max} is positive, and the polymer concentrations grow exponentially with time. Here, the largest eigenvalue is computed numerically using a truncated system of equations that neglects contributions from polymers longer than n_{max}. Importantly, the

polymer size distribution is kinetically controlled: the distribution that grows the fastest is the one that is ultimately selected among the different eigenmodes. Figure 15b shows the eigenvectors $V_n^{\max}$ corresponding to the largest eigenvalues $\lambda_{\max}$. By convention, these vectors are normalized as $\sum_{n=2}^{\infty}(V_n^{\max})^2 = 1$.

Approximating Autocatalytic Growth

When the concentration of free monomer a_1 is sufficiently high (to be determined), we can use the simplified model of equation (5.6) to estimate the dominant eigenvalue and the average length of the corresponding mode. Linearizing those equations about $m_0' = m_1' = 0$, the approximate governing equations for the polymer moments become

$$\begin{bmatrix} \dot{m}_1' \\ \dot{m}_0' \end{bmatrix} = \begin{bmatrix} 0 & 2k_+a_1 - 2k_{\mathrm{d2}}' \\ k_{\mathrm{d3}}' & -6k_{\mathrm{d3}}' \end{bmatrix} \begin{bmatrix} m_1' \\ m_0' \end{bmatrix} \tag{5.16}$$

The two eigenvalues of this linear system are given by

$$\lambda_{\pm} = -3k_{\mathrm{d3}}' \pm \sqrt{k_{\mathrm{d3}}'(9k_{\mathrm{d3}}' + 2k_+a_1 - 2k_{\mathrm{d2}}')} \tag{5.17}$$

These eigenvalues are represented by the dashed lines in Figure 15a. The greater of the two eigenvalues is positive—indicating autocatalytic growth—when the monomer concentration is larger than some critical value, $a_1 > a_1^* = k_{\mathrm{d2}}'/k_+$. The corresponding eigenvector specifies the average polymer length m_1'/m_0' for the growing polymer

$$L_{\mathrm{grow}} = 3 + \sqrt{\frac{9k_{\mathrm{d3}}' + 2k_+a_1 - 2k_{\mathrm{d2}}'}{k_{\mathrm{d3}}'}} \tag{5.18}$$

For example, at the peak monomer concentration of $a_1 = 1575$ during each oscillation cycle, this length scale is $L_{\mathrm{grow}} = 37$, which is close to the observed value from Fig. 12c.

References for Sections 1–5

(1) Israelachvili, J. N., *Intermolecular and surface forces*; Academic press: 2011.

(2) Cohen, S. I.; Vendruscolo, M.; Welland, M. E.; Dobson, C. M.; Terentjev, E. M.; Knowles, T. P. *J. Chem. Phys.* **2011**, *135*, 065105.

(3) Hindmarsh, A. C.; Brown, P. N.; Grant, K. E.; Lee, S. L.; Serban, R.; Shumaker, D. E.; Woodward, C. S. *ACM Trans. Math. Software* **2005**, *31*, 363–396.

(4) McKay, M. D.; Beckman, R. J.; Conover, W. J. *Technometrics* **2000**, *42*, 55–61.

6 Analysis of literature examples

All figures and their captions were taken from the corresponding cited manuscripts.

Used notations:
f_0 = Concentration of fuel at time 0
c_{tot} = total concentration of monomer
k_a = apparent rate constant of activation
$k'_d = k_d \times g_0$ apparent deactivation rate constant is equal deactivation rate constant times deactivation fuel concentration.

Buffer pK_a values were taken from:
https://microscopy.berkeley.edu/resources/instruction/buffers.html

First-order integrated rate law[1]:

$$k = \frac{1}{t} \ln \frac{a_0}{a}$$

$$k = \frac{\ln 2}{t_{0.5}} = \frac{0.693}{t_{0.5}}$$

Second-order integrated rate law[1] ($A + B \longrightarrow C,\ A_0 = B_0$):

$$k = \frac{1}{t}\left(\frac{1}{a} - \frac{1}{a_0}\right)$$

$$k = \frac{1}{t_{0.5} a}$$

Second order integrated rate law[1] ($A + B \longrightarrow C,\ A_0 \neq B_0$):

$$k = \frac{1}{t(b_0 - a_0)} \ln \frac{a_0 b}{b_0 a}$$

Summary of the coordinates of the systems in Figure 2a of the main text

		Main text ref.	$\frac{f_0}{c_{tot}}$	$\frac{k_d g_0}{k_a c_{tot}}$
Adams	2019	41	5.64	0.08
Adams	2020	43	10.46	0.02
Panzarasa	2020	44	8000	0.0000013
George	2019	46A	5	0.3
		46B	2	0.1467
Hermans	2017	47	1000	16.21
Besenius	2017	48	346.971	3.69
Quintard	2021	49	4.26	4.4
George	2018	50	6.4	231.69
Walther	2015	29	0.738	0.0001659
Das	2018	55	1	0.08720
Das	2019	57	1	1.021
Ulijn	2015	58	2	0.1076

Ulijn	2013	59B	1	0.004462
		59A	1	0.002746
Ulijn	2018	60	4	0.02452
Hermans	2020	61	3.53	0.2535
Boekhoven	2017	62	1	0.317
Thordarson	2018	64	19.99	0.00256
Guan	2020	65	30	0.2267
van Esch	2015	67A	4	18.06
		67B	4	44.80
		67C	4	475.00
		67D	0.4	44.80
van Esch	2010	66	1	6.912

Adams 2019 — ref. 41

$f_0 = 0.02\ \text{M} \times 2 = 0.04\ \text{M}$, urea produces 2 equiv. of ammonia.

$$c_{tot} = \frac{2\ \text{mg/mL}}{282\ \text{g/mol}} = 0.007092\ \text{M}$$

Activation: increase in pH to 9.1-9.2 to trigger gelation; autocatalytic reaction between urease and urea that produces NH_3. We approximate the pH change using 1st order reaction kinetics and the corresponding expression for half-life. Activation half-life was determined from the data of urea-urease reaction in the presence of 1 on the Fig. 2a (closed square symbol – variation of pH) as the time needed to reach pH 8.5 from 8.0.

$t_{0.5} = 2\ \text{min} = 120\ \text{s}$

$$k_a = \frac{0.693}{120\ \text{s}} = 0.005775\ \text{s}^{-1}$$

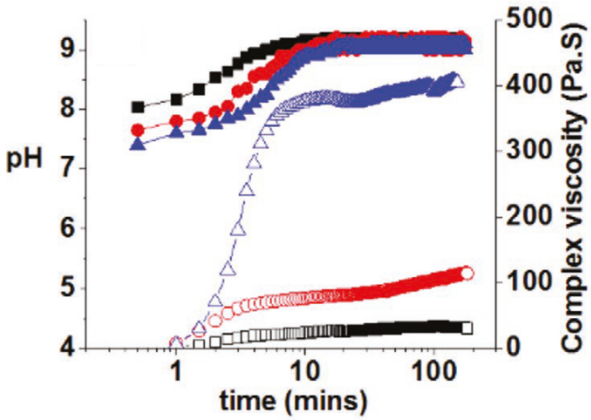


Figure 2a. Variation of pH (closed symbol), complex viscosity (open symbol) with time for 1 in presence of urea–urease reaction.

Deactivation: decrease in pH to induce disassembly, triggered by hydrolysis of methyl formate. The process is also describe using 1st order rate expressions. Deactivation half-life time was determined from the experiment that combined urea-urease reaction and hydrolysis of methyl formate without 1 (Fig. S3, black squares). It was estimated as time needed for pH to drop from 9.0 to 8.15.

$t_{0.5} = 25\ \text{min} = 1500\ \text{s}$

$$k_d = \frac{0.693}{1500\ \text{s}} = 0.000462\ \text{s}^{-1}$$

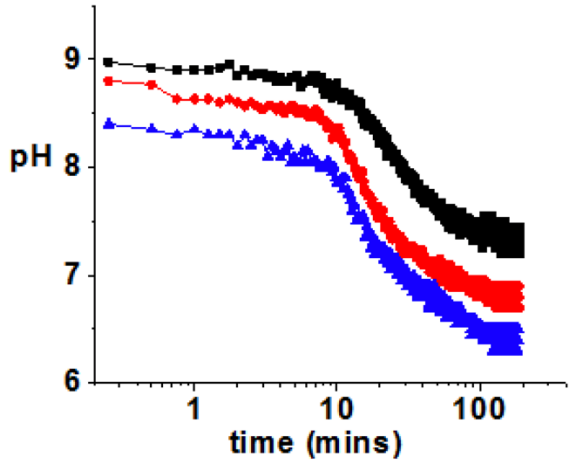


Figure S3a. Change in pH with time from the urease–urea reaction in absence of gelator 1. Initial reaction condition is: [urease] =0.2 mg/mL, [urea] = 0.02 M, volume of methyl formate = 25 μL (black), 50 μL (red) and 100 μL (blue).

$$\frac{f_0}{c_{tot}} = \frac{0.04\ \text{M}}{0.007092\ \text{M}} = 5.64$$

Expression $\frac{k_d g_0}{k_a c_{tot}}$ transforms into $\frac{k_d}{k_a}$ since we used 1st order equations.

$$\frac{k_d g_0}{k_a c_{tot}} \Longrightarrow \frac{k_d}{k_a} = \frac{0.000462\ \text{s}^{-1}}{0.005775\ \text{s}^{-1}} = 0.08$$

Adams 2020 — ref. 43

$f_0 = 0.02\ \text{M} \times 2 = 0.04\ \text{M}$, urea produces 2 equiv. of ammonia.

$c_{tot} = \frac{2\ \text{mg/mL}}{523.22\ \text{g/mol}} = 0.003822\ \text{M}$ of hydrogelator

Activation: increase of pH due to the production of ammonia in urea hydrolysis catalyzed by urease. Deactivation: decrease of pH due to hydrolysis of methyl formate. We approximate the pH change using 1st order reaction kinetics and the expression for half-life.

$$t_{0.5} = \frac{\ln 2}{k} = \frac{0.693}{k}$$

$t_{0.5}$ of activation and deactivation was determined from the Figure 2c of the paper as the time that is needed to change pH from minimum or maximum value to the half of the range: from 4.3 to 6.15 (activation, $t_{0.5} = 1.8\ \text{min} = 108\ \text{s}$) and from 8 to 7 (deactivation, $t_{0.5} = 100 - 10 = 90\ \text{min} = 5400\ \text{s}$).

$$k_a = \frac{0.693}{108\ \text{s}} = 0.00642\ \text{s}^{-1}$$

$$k_d = \frac{0.693}{5400\ \text{s}} = 0.000117\ \text{s}^{-1}$$

$$\frac{f_0}{c_{tot}} = \frac{0.04\ \text{M}}{0.003822\ \text{M}} = 10.46$$

Expression $\frac{k_d g_0}{k_a c_{tot}}$ transforms into $\frac{k_d}{k_a}$ since we used 1st order equations.

$$\frac{k_d g_0}{k_a c_{tot}} \Longrightarrow \frac{k_d}{k_a} = \frac{0.000117\ \mathrm{s}^{-1}}{0.00642\ \mathrm{s}^{-1}} = 0.02$$

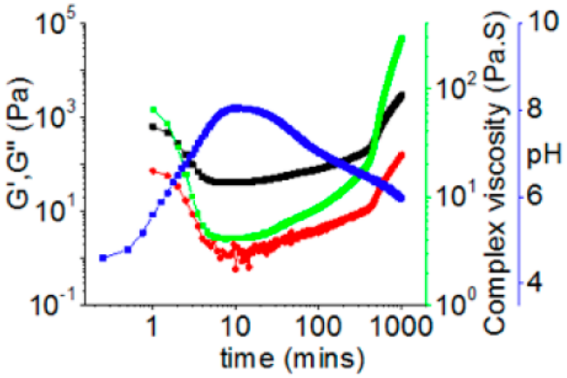


Figure 2c. Variation of pH (blue), G′ (black), G″ (red), and complex viscosity (green) with time for 1 from the enzymatic reaction in the presence of methyl formate.

Panzarasa 2020 — ref. 44

Conditions: 25 μM PDIamine, 200 mM MG, 100 mM PrS

Activation: The methylene glycol-sulfite clock reaction that increases the pH. The rate-determining step is formation of formaldehyde in formaldehyde-methylene glycol equilibrium with the rate constant $k_a = 10\ \mathrm{s}^{-1}$. (Rate constants for MG were taken from the literature[2]).

Deactivation: 1,3-Prpanesultone hydrolysis that decreases pH.

The deactivation rate constant was taken from the main text, p. 2: $k_d = 13.0 \times 10^{-6}\ \mathrm{s}^{-1}$.

$$\frac{f_0}{c_{tot}} = \frac{0.2\ \mathrm{M}}{0.000025\ \mathrm{M}} = 8000$$

Expression $\frac{k_d g_0}{k_a c_{tot}}$ transforms into $\frac{k_d}{k_a}$ since we used 1st order equations.

$$\frac{k_d g_0}{k_a c_{tot}} \Longrightarrow \frac{k_d}{k_a} = \frac{13.0 \times 10^{-6}\ \mathrm{s}^{-1}}{10\ \mathrm{s}^{-1}} = 13.0 \times 10^{-7}$$

George 2019 — ref. 46

Conditions: A) amine 8a 5 mM, aldehyde-containing monomer 1 mM. B) 8A-est 2 mM, aldehyde-containing monomer 1 mM.

Activation: an imine formation. The rate constant of activation is given in the main text, p.3:

$$k_a = 0.003\ \mathrm{s}^{-1}$$

Deactivation: A) Imine hydrolysis due to pH decrease after γ-BL lactone hydrolysis. Data from Fig S34. γ-BL 500 mM. Half-life of deactivation $t_{0.5} = 1100 - 330 = 770$ s. The deactivation constant was calculated using a pseudo-1st order expression for the half-life:

$k_d = \frac{0.693}{770\ \mathrm{s}} = 0.0009\ \mathrm{s}^{-1}$.

Expression $\frac{k_d g_0}{k_a c_{tot}}$ transforms into $\frac{k_d}{k_a}$ since we used 1st order equations.

$$\frac{k_d g_0}{k_a c_{tot}} \Longrightarrow \frac{k_d}{k_a} = \frac{0.0009\ \mathrm{s}^{-1}}{0.003\ \mathrm{s}^{-1}} = 0.3$$
$$\frac{f_0}{c_{tot}} = \frac{0.005\ \mathrm{M}}{0.001\ \mathrm{M}} = 5$$

B) Ester cleavage promoted by lipase. On Figure 5b the half-life for the experiment with 2.5 U lipase is $t_{0.5} = 2125\ \mathrm{s} - 550\ \mathrm{s} = 1575\ \mathrm{s}$. The deactivation constant was calculated using a pseudo-1st order expression for the half-life:
$k_d = \frac{0.693}{1575\ \mathrm{s}} = 0.00044\ \mathrm{s}^{-1}$.

$$\frac{k_d g_0}{k_a c_{tot}} \Longrightarrow \frac{k_d}{k_a} = \frac{0.00044\ \mathrm{s}^{-1}}{0.003\ \mathrm{s}^{-1}} = 0.147$$
$$\frac{f_0}{c_{tot}} = \frac{0.002\ \mathrm{M}}{0.001\ \mathrm{M}} = 2$$

Hermans 2017 — ref. 47

Conditions: 100 µM PDI, 1 mM ATP, 0.13 µM protein kinase A (PKA), 0.3 µM **λ**-protein phosphatase (**λ**PP).
Activation: phosphorylation of serine in PDI to p-PDI and 2p-PDI by PKA upon consumption of ATP. Waste: 2 molecules of ADP. Rate constant of phosphorylation by PKA in optimized conditions (see SI p. 21) is $k_a = 2.93\ \mathrm{s}^{-1}$.
Deactivation: cleavage of 2 phosphate groups by **λ**PP. Waste: phosphate. Rate constant of dephosphorylation by **λ**PP in optimized conditions (see SI p. 21) is $k_d = 47.5\ \mathrm{s}^{-1}$.

$$f_0 = 1 \times 10^{-3}\ \mathrm{M}$$
$$c_{tot} = 1 \times 10^{-6}\ \mathrm{M}$$
$$\frac{f_0}{c_{tot}} = 1000$$

Expression $\frac{k_d g_0}{k_a c_{tot}}$ transforms into $\frac{k_d}{k_a}$ since we used 1st order equations.

$$\frac{k_d g_0}{k_a c_{tot}} \Longrightarrow \frac{k_d}{k_a} = \frac{47.5\ \mathrm{s}^{-1}}{2.93\ \mathrm{s}^{-1}} = 16.2116$$

Notes:
The polymer is more assembled in the activated 2p-PDI state, however, it never disassembles to a monomeric state in the conditions of the transient experiment. The activation process seems to not perturb PDI assemblies or to trigger disassembly. The same holds for the deactivation process. The mechanism of the change of polymer helicity is not fully clear.

Besenius 2017 — ref. 48

Conditions: 7 mg/ml amphiphile (MW 8096 g/mol), 300 mM glucose, 650 µg glucose oxidase, 100 mM NaCl, pH 5.7.

Oxidation of glucose catalyzed by glucose oxidase produces gluconolactone that hydrolyzes decreasing pH (activation). Another product of this reaction is H_2O_2 that oxidizes methionine residue to sulfoxide (deactivation).
Activation half-life estimated from the Fig. S15 as the time that is needed to decrease pH from 5.7 to 4.25 (which is half of the experimental pH range (5.7-2.8)/(2+2.8).
To fully activate monomer 6 eq of acid is needed. The activation rate constant was calculated using pseudo-first order expression for half-life.

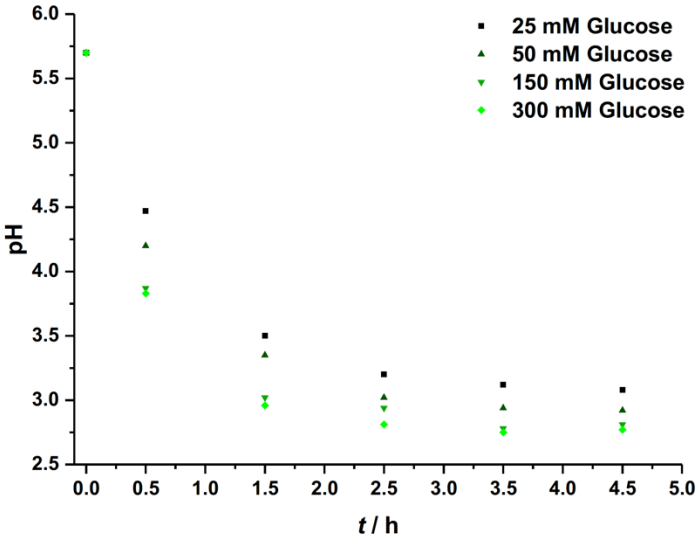


Figure S15. pH tracing of a solution of I (7 mg/mL) in ddH_2O with 260 μg/mL GOx and various concentrations of glucose at 40°C.

$$t_{0.5} = 18 \text{ min} = 1080 \text{ s}$$
$$k_a = \frac{0.693}{1080 \text{ s}} = 0.00064167 \text{ s}^{-1}$$

The yield of activation was determined using Henderson-Hasselbalch equation at pH 2.8 and pK_a 5.5 of the monomer to be roughly 99.8 %.
Deactivation:
Assuming that all glucose is converted, 300 mM of H_2O_2 is formed.
For 1 molecule of monomer 3 x 2 = 6 eq of peroxide is needed.
Second-order rate constant of methionine oxidation by H_2O_2 is $k_d = 7.9 \times 10^{-3}$ $M^{-1}s^{-1}$.[3]

$$\frac{f_0}{c_{tot}} = \frac{0.3 \text{ M}}{0.000864625 \text{ M}} = 346.971$$

Expression $\frac{k_d g_0}{k_a c_{tot}}$ transforms into $\frac{k_d g_0}{k_a}$ since we used 1st order equation for the activation rection.

$$\frac{k_d g_0}{k_a c_{tot}} \Longrightarrow \frac{k_d g_0}{k_a} = \frac{7.9 \times 10^{-3} \text{ M}^{-1}\text{s}^{-1} \times 0.3 \text{ M}}{0.00064167 \text{ s}^{-1}} = 3.69351$$

Quintard 2021 — ref. 49
Fuel and monomer concentrations were taken from Fig 6 description.

$$f_0 = \frac{35.8 \text{ μmol}}{409 \text{ μL}} = 0.0875 \text{ M}$$

$$c_{tot} = \frac{8.4\ \mu\text{mol}}{409\ \mu\text{L}} = 0.0205\ \text{M}$$

Activation: protonation of the gelator by TCA and TCA decarboxylation. Activation was taken as stages 1–2 on Figure 6. Assuming ECD signal to be directly proportional to the amount of activated monomer we determine that half of this amount was converted at 32.5 min, therefore, half-life was calculated as

$$t_{0.5} = 32.5\ \text{min} - 29.2\ \text{min} = 3.3\ \text{min} = 198\ \text{s}$$

$$k_a = \frac{0.693}{198\ \text{s}} = 0.0035\ \text{s}^{-1}$$

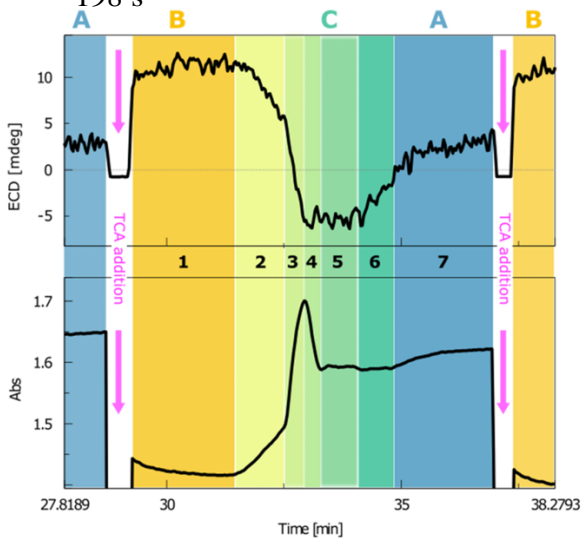


Figure 6. ECD (top) and UV (bottom) −COOH band intensity at 279 nm with respect to time during cycle 5. Conditions: 2 (8.4 μmol), DBU (18.8 μmol) in 0.4 mL of DMF at room temperature. Addition of 9 μL of TCA solution (C = 0.65 mg/μL, 35.8 μmol) in DMF.

Deactivation: formation of carbamate by reaction of the monomer and 1 equiv. CO_2. Strictly speaking, it is not fully correct to approximate this process according to the 1st order kinetics because at the latest stage of deactivation ratio of reactants is 1:1. However, due to the lack of data and taking into an account that deactivation can possibly start much earlier, i.e., when there is an excess of CO_2 due to the TCA decomposition, we proceeded with the 1st order rate equation. Half-life was determined from the same figure, and the rate constant of deactivation is

$$t_{0.5} = 0.75\ \text{min} = 45\ \text{s}$$

$$k_d = \frac{0.693}{45\ \text{s}} = 0.0154\ \text{s}^{-1}$$

$$\frac{f_0}{c_{tot}} = \frac{0.0875\ \text{M}}{0.0205\ \text{M}} = 4.26$$

Expression $\frac{k_d g_0}{k_a c_{tot}}$ transforms into $\frac{k_d}{k_a}$ since we used 1st order equations.

$$\frac{k_d g_0}{k_a c_{tot}} \Longrightarrow \frac{k_d}{k_a} = \frac{0.0154\ \text{s}^{-1}}{0.0035\ \text{s}^{-1}} = 4.40$$

Judging by turbidity measurements the self-assembly (corresponds to increase of turbidity) is faster than the disassembly (decrease of turbidity), even though activation appears to be slower than deactivation. We also cannot exclude that disassembly in this system is limited by deactivation. Interestingly, the system might be able to

operate even in the case of low activation levels as the substrate readily forms gels at 0.1 % wt/v in DMF[4] which is about 10 times lower of the used concentration.

George 2018 — ref. 50

Conditions: 0.02 mM oligo(p-phenylenevinylene) derivative (donor) 1, functionalized with DPA–Zn phosphate fragments (acceptors), 6.4 eq. ATP, 0.2 units/mL of potato apyrase, CH_3CN/HEPES, 10/90, v/v, 30 ° C.

Activation: Yield of activation was calculated considering that 0.9 eq ATP is enough to saturate all the Zn positions and it corresponds to –410 mdeg of ellipticity (SI Fig7). In the transient system Fig.5b –28 mdeg was possible to reach. Hence, activation yield of transient experiment is 28/410 = 6.8 % — about 7 % assuming CD signal to be proportional to concentration.

The molecule 1 is activated selectively by ATP allosteric binding that triggers conformational change of 1 after which the polymer starts to grow. Half-life of activation is taken from the SI Fig. 12 (t_{50} Fluorescence, 1 eq ATP) and the activation constant was approximated to the second order:

$$t_{1/2} = 3.50\ \text{min} = 210\ \text{s}$$

$$k_a = \frac{1}{210\ \text{s} \times (0.000128\ \text{M} - 0.00002\ \text{M})} \ln \frac{0.00002\ \text{M} \times (0.000128\ \text{M} - 0.00001\ \text{M})}{0.000128\ \text{M} \times 0.00001\ \text{M}} = 26.9754\ \text{M}^{-1}\text{s}^{-1}$$

Deactivation: Enzyme potato apyrase hydrolyses ATP to ADP, AMP, which brings conformation of assembled monomer back to the initial state. The rate constant of deactivation was taken as a frequency of 0.2 units/mL apyrase:

$$k_d = 0.125\ \text{s}^{-1}$$

$$\frac{f_0}{c_{tot}} = \frac{0.000128\ \text{M}}{0.00002\ \text{M}} = 6.4$$

Expression $\frac{k_d g_0}{k_a c_{tot}}$ transforms into $\frac{k_d}{k_a c_{tot}}$ since we used 1st order equation for the deactivation rection.

$$\frac{k_d g_0}{k_a c_{tot}} \Longrightarrow \frac{k_d}{k_a c_{tot}} = \frac{0.125\ \text{s}^{-1}}{26.9754\ \text{M}^{-1}\text{s}^{-1} \times 0.00002\ \text{M}} = 231.6927$$

Walther 2015 — ref. 29

Fuel concentration was taken as the total citric acid concentration in the buffer:

$f_0 = 0.0108\ \text{M}$

Total monomer concentration was determined from the given weight percentage 0.6 % wt and molecular weight 410.46 g/mol.

$c_{tot} = 0.0146\ \text{M}$

Activation: decrease of pH of the reaction mixture as a result of mixing urea solution in acidic buffer and basic urease solution of the dipeptide.

The rate constant of activation was calculated according to the 2nd order rate equation since concentrations of fuel and monomer are similar. Time and concentrations were determined from Figure 3b, red dashed curve.

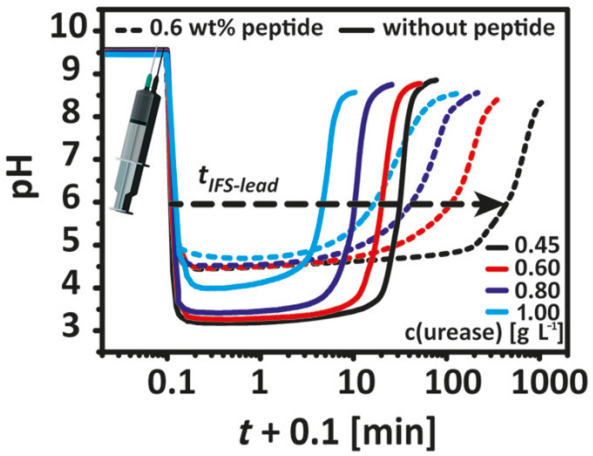


Figure 3b. Influence of the peptide on the pH–time profiles, as programmed by the use of different urease concentrations.

Time for pH to reach 6 from 9.5 is roughly 3.6 s. Concentration of inactive monomer (anionic form of the peptide) was calculated using Henderson-Hasselbalch equation and dipeptide pK_a 5.7.

$$\frac{[A^-]}{[HA]} = \frac{10^{6.0}}{10^{5.7}}; [A^-] + [HA] = 0.0146\ \text{M}$$
$$[A^-]_{\text{pH6}} = 0.0097\ \text{M}$$

Concentration of fuel was approximated considering ratios of citric acid to sodium acetate and of the peptide protonated to deprotonated forms in the buffer at pH 6 using Henderson-Hasselbalch equation:

$$\frac{[citrate]}{[citric\ acid]} = \frac{10^{6.0}}{10^{6.4}}; [citrate] + [citric\ acid] = 0.012\ \text{M}$$
$$[citric\ acid]_{\text{pH6}} = 0.0086\ \text{M}$$

Concentration of citric acid in the presence of peptide:

$$[citric\ acid]_{real} = [citric\ acid]_{\text{pH6}} - [HA]_{\text{pH6}} = 0.0086\ \text{M} - 0.0049\ \text{M} = 0.0037\ \text{M}$$

2nd order rate constant:

$$k_a = \frac{1}{t(f_0 - c_{tot})} \ln \frac{c_{tot}[citric\ acid]_{\text{pH6}}}{f_0[A^-]_{\text{pH6}}}$$
$$= \frac{1}{3.6\ \text{s}\ (0.0108\ \text{M} - 0.0146\ \text{M})} \ln \frac{0.0146\ \text{M} \times 0.0037\ \text{M}}{0.0108\ \text{M} \times 0.0097\ \text{M}} = 48.102\ \text{M}^{-1}\text{s}^{-1}$$

Deactivation: increase of pH due to the production of ammonia by hydrolysis of urea catalyzed with urease. To determine the rate constant the timepoint was taken again at pH 6 but as time that it takes to reach pH 6 from 4.5, which is roughly 99 min (5940 s). Approximated by 1st order equation the deactivation rate constant is:

$$k_d = \frac{0.693}{5940\ \text{s}} = 0.000116667\ \text{s}^{-1}$$

Finally:

$$\frac{f_0}{c_{tot}} = \frac{0.0108\ \text{M}}{0.0146\ \text{M}} = 0.7388$$

Expression $\frac{k_d g_0}{k_a c_{tot}}$ transforms into $\frac{k_d}{k_a c_{tot}}$ since we used 1st order equation for the deactivation rection.

$$\frac{k_d g_0}{k_a c_{tot}} \Longrightarrow \frac{k_d}{k_a c_{tot}} = \frac{0.000116667\ \mathrm{s^{-1}}}{48.102\ \mathrm{M^{-1}s^{-1}} \times 0.0146\ \mathrm{M}} = 0.00017$$

Das 2018 — ref. 55

Conditions: 4mM EDC, 12 mM C18H, 4 mM NP, 20 % v/v DMSO-water pH 6.5

$$f_0 = 0.004\ \mathrm{M} = c_{tot}$$

Activation: formation of ester C18H-NP out of acid C18H and nitrophenol NP. The process is rapid after addition of EDC ("Rapid conversion from acid to ester with conversion of about 1.5 mM occurred in the first 2 min"). The maximum conversion of ester reached was 38 % (at 2 min after EDC addition, Fig. 4f), therefore, the amount of acid available for esterification is $4\ \mathrm{mM} - 0.38 \times 4\ \mathrm{mM} = 2.48\ \mathrm{mM} = 0.00248\ \mathrm{M}$. The activation rate constant according to the 2nd order rate law (with equal concentration of reactants):

$$k_a = \frac{1}{120\ \mathrm{s}}\left(\frac{1}{0.00248\ \mathrm{M}} - \frac{1}{0.004\ \mathrm{M}}\right) = 1.2769\ \mathrm{M^{-1}s^{-1}}$$

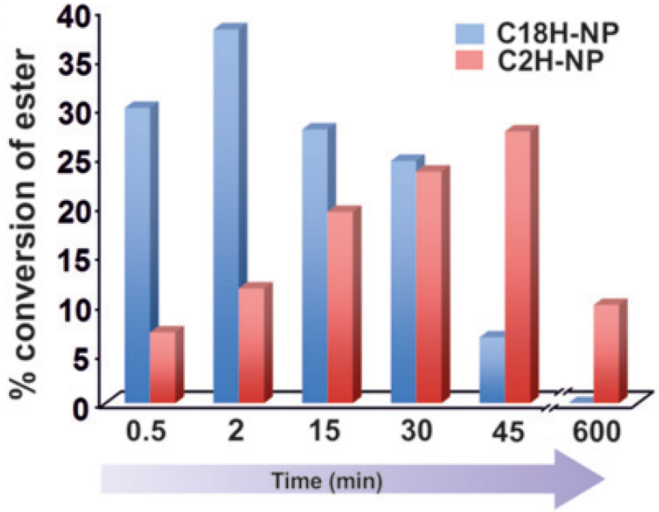


Figure 4f. Time-resolved conversion of C2HNP and C18H-NP.

Deactivation: hydrolysis of C18H-NP ester catalyzed by proximal histidine residue. The rate of hydrolysis was measured in different pH conditions (Fig.4e). At pH 6.5 the rate is about 31 μM $\mathrm{min^{-1}}$. For the pseudo-first order reaction we took data from the Figure 4f, where at 15 min only 29 % of ester was present:

$$\left[C18H_{NP}\right]_t = 0.29 \times 0.004\ \mathrm{M} = 0.00116\ \mathrm{M} = 1.16 \times 10^3 \mu\mathrm{M}$$

$$V = 31\ \mu\mathrm{M\ min^{-1}} = k_d\left[C18H_{NP}\right]_t = k_d \times 1.16 \times 10^3 \mu\mathrm{M}$$

$$k_d = \frac{31\ \mu\mathrm{M\ min^{-1}}}{1.16 \times 10^3 \mu\mathrm{M} \times 60} = 0.0004454\ \mathrm{s^{-1}}$$

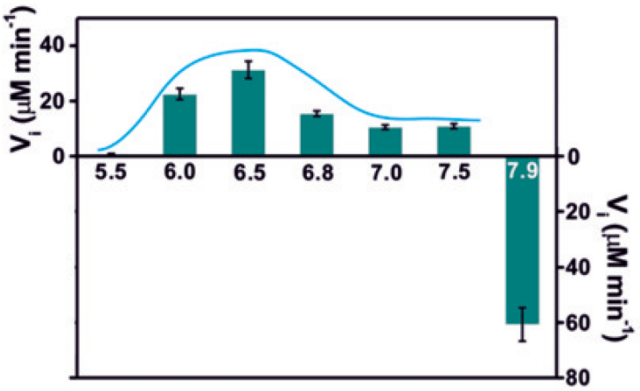


Figure 4e. Rate of hydrolysis of ester as a function pH 5.5 to 7.5 and the rate of formation of ester at pH 7.9.

$$\frac{f_0}{c_{tot}} = \frac{0.004\ \text{M}}{0.004\ \text{M}} = 1$$

Expression $\frac{k_d g_0}{k_a c_{tot}}$ transforms into $\frac{k_d}{k_a c_{tot}}$ since we used 1st order equation for the deactivation rection.

$$\frac{k_d g_0}{k_a c_{tot}} \Longrightarrow \frac{k_d}{k_a c_{tot}} = \frac{0.0004454\ \text{s}^{-1}}{1.2769\ \text{M}^{-1}\text{s}^{-1} \times 0.004\ \text{M}} = 0.0872051$$

Das 2019 — ref. 57

Conditions: SYS1 (30 mM H + 30 mM F-NP, 20 % v/v DMSO-water pH 6)

$$f_0 = 0.03\ \text{M} = c_{tot}$$

Activation: Mixing H and F-NP triggers the process of gelation that is described by increase of storage modulus G' to its highest value in about 2 h (text) or in 105 min (Fig. S1). Half-life of gelation for SYS1 was roughly determined from the Fig. S1 as a time that is needed to reach half of the highest value of storage modulus (purple squares data) and from Fig. 4a as a time that is needed to reach 15 a.u. of fluorescence intensity. The half-life was further used to calculate the 1st order activation rate constant:

$$t_{0.5} = 60\ \text{min} = 3600\ \text{s}$$

$$k_a = \frac{0.693}{3600\ \text{s}} = 0.0001925\ \text{s}^{-1}$$

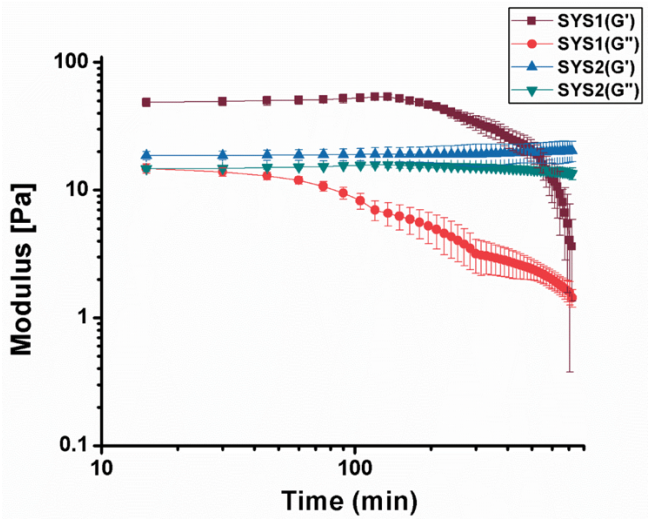


Figure S1. Change of storage and loss modulus with time for SYS1 and SYS2 for triplicates.

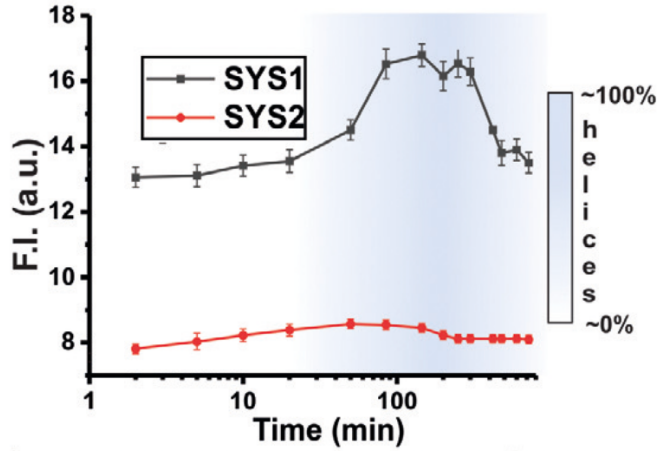


Figure 4a. Time-dependent fluorescence for SYS1 and SYS2. Error bars represent standard deviations of three independent measurements. The intensity band represents broadly the percentage of helices.

Right after mixing H and F-NP the gel consists mainly of fibers that over 4 h form transient helices.

Deactivation is defined as hydrolysis of F-NP ester via catalytic action of histidine residue of H. This process correlates with the change in gel morphology: helices disappear after 6-10 h from the start of experiment with formation of fibrillar and "other" morphologies.

Deactivation rate is referred to an accelerated catalytic rate given at the page 4 which is the rate of the rapid phase of the cooperative catalysis (Fig. 4b), the rate constant was calculated from the rate using the pseudo-first expression.

$$V = 5.9\ \mu\text{M s}^{-1} = k_d[F\_NP] = k_d \times 30 \times 10^3\ \mu\text{M}$$

$$k_d = \frac{5.9\ \mu\text{M s}^{-1}}{30 \times 10^3\ \mu\text{M}} = 0.000196667\ \text{s}^{-1}$$

$$\frac{f_0}{c_{tot}} = \frac{0.004\ \text{M}}{0.004\ \text{M}} = 1$$

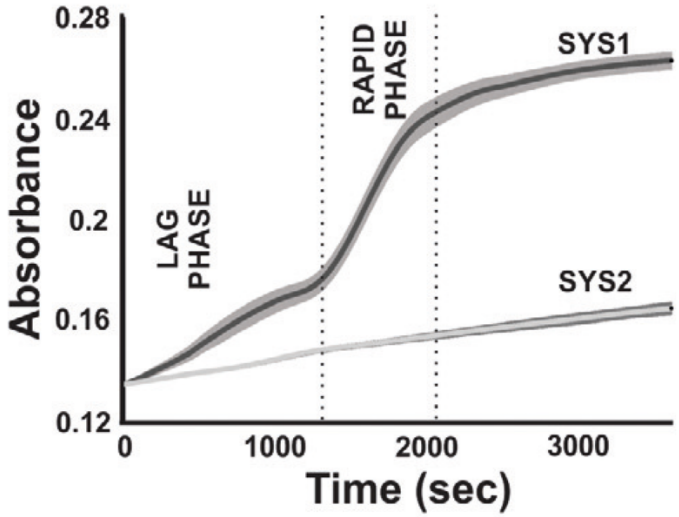


Figure 4b. Time-course kinetics of hydrolysis of F-NP for SYS1 and SYS2.

Expression $\frac{k_d g_0}{k_a c_{tot}}$ transforms into $\frac{k_d}{k_a}$ since we used 1st order equations.

$$\frac{k_d g_0}{k_a c_{tot}} \Longrightarrow \frac{k_d}{k_a} = \frac{0.000196667\ \mathrm{s}^{-1}}{0.0001925\ \ \mathrm{s}^{-1}} = 1.02165$$

Ulijn 2015 — ref. 58

Conditions: Aspartame (DF-OMe) 20 mM and Y-NH2 40 mM form DFY-NH_2 gelator, 1 mg chymotrypsin, 0.1 M pH 8.

Activation: hydrolysis of methyl ester and reaction with tyrosine amide DFY-NH_2. From the Figure 1b the highest conversion is 86 %. Activation time is taken from the main text, p. 2, 30 min (1800 s). Applying a second-order expression:

$$k_a = \frac{1}{1800\ \mathrm{s} \times (0.04\ \mathrm{M} - 0.02\ \mathrm{M})} \ln \frac{0.02\ \mathrm{M} \times (0.04\ \mathrm{M} - (0.02\ \mathrm{M} - 0.86 \times 0.02\ \mathrm{M}))}{0.04\ \mathrm{M} \times (0.02\ \mathrm{M} - 0.86 \times 0.02\ \mathrm{M})}$$
$$= 0.0525984\ \mathrm{M}^{-1}\mathrm{s}^{-1}$$

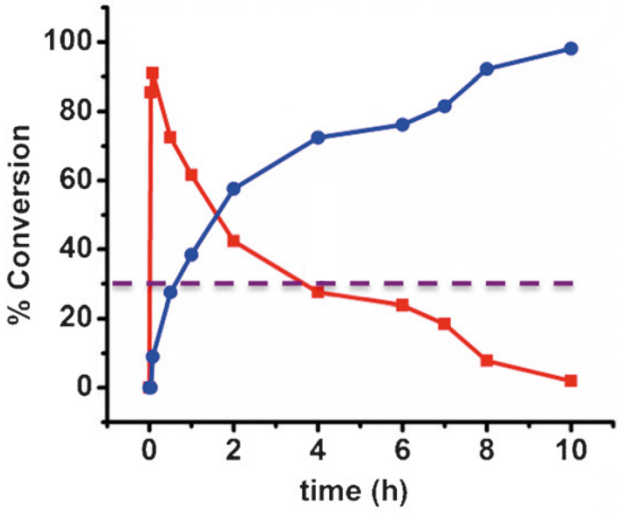


Figure 1b. HPLC analysis of 2b product formation and degradation with the red line representing DFY-NH_2 and the blue line representing DF (1).

Deactivation: hydrolysis of DFY-NH_2 to DF-OH.
The deactivation rate constant was determined using the half-life expression for the pseudo-first order. The data is taken from Figure 1b: $t_{0.5} = 1.7\ \text{h} = 6120\ \text{s}$, 50 % of monomer is present.

$$k_d = \frac{0.693}{6120\ \text{s}} = 0.000113235\ \text{s}^{-1}$$

$$\frac{f_0}{c_{tot}} = \frac{0.04\ \text{M}}{0.02\ \text{M}} = 2$$

Expression $\frac{k_d g_0}{k_a c_{tot}}$ transforms into $\frac{k_d}{k_a c_{tot}}$ since we used 1st order equation for the deactivation rection.

$$\frac{k_d g_0}{k_a c_{tot}} \Longrightarrow \frac{k_d}{k_a c_{tot}} = \frac{0.000113235\ \text{s}^{-1}}{0.0525984\ \text{M}^{-1}\text{s}^{-1} \times 0.02\ \text{M}} = 0.1076414$$

Ulijn 2013 — ref. 59
Conditions: Nap-Y-OMe 20 mM and Y-NH_2 20 mM form Nap-YY-NH_2 gelator, 1 mg chymotrypsin, pH 8 or 10.
Activation: hydrolysis of methyl ester and reaction with tyrosine amide.
The activation rate constant was calculated using the integrated form of second-order rate expression for the equal amounts of reactants and data from the text p. 2 and from Figure 1b:

$$k_a = \frac{1}{(2\ \text{min} \times 60)}\left(\frac{1}{(0.02\ \text{M} - 0.66 \times 0.02\ \text{M})} - \frac{1}{0.02\ \text{M}}\right) = 0.808824\ \text{M}^{-1}\text{s}^{-1}$$

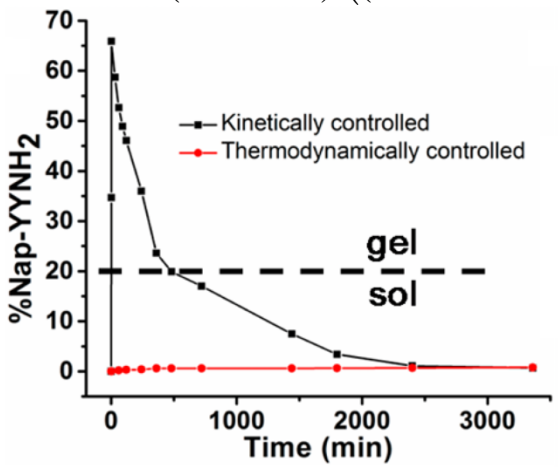


Figure 1b. Nonequilibrium assembly and hydrogelation. Time course of the kinetically controlled reaction (chymotrypsin catalyzed) of Nap-Y-OMe (black line) and thermodynamically controlled reaction (thermolysin catalyzed) of Nap-Y-OH (red line) with Y-NH_2 at pH 8.

Deactivation: hydrolysis of DFY-NH_2 to DF-OH.

The deactivation rate constant was determined using the half-life expression for the pseudo-first order.

$$k_d = \frac{0.693}{t_{0.5}}$$

The data was taken from Figure S7.

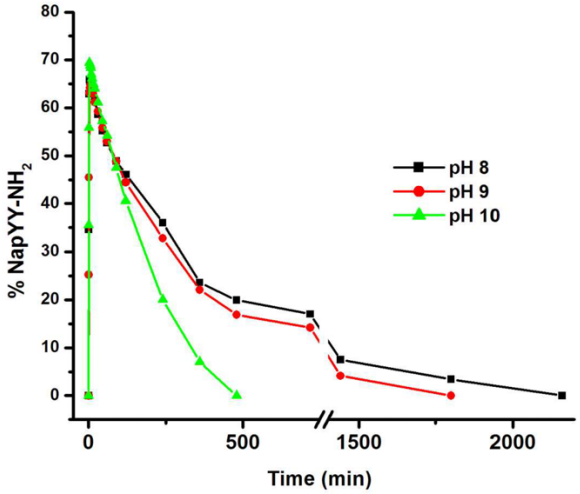


Figure S7. pH dependent change of reaction kinetics of the reaction between Nap-Y-OMe and Y-NH_2 in presence of 1 mg of chymotrypsin.

Reference in Figure 2a in the main text	pH	$t_{0.5}$	k_d, s^{-1}	$\frac{k_d}{k_a c_{tot}}$
102A	8	260 min (15600 s)	7.21875×10^{-5}	0.002746154
102B	10	160 min (9600 s)	4.44231×10^{-5}	0.0044625

Expression $\frac{k_d g_0}{k_a c_{tot}}$ transforms into $\frac{k_d}{k_a c_{tot}}$ since we used 1st order equation for the deactivation rection.

$$\frac{f_0}{c_{tot}} = \frac{0.02\ \mathrm{M}}{0.02\ \mathrm{M}} = 1$$

Ulijn 2018 — ref. 60

Conditions: 0.2 M of non-activated monomer 1, 0.08 M of E amide, 1 mg/mL of chymotrypsin in 10 % DMSO–90 % of 100 mM phosphate buffer (pH 8).
Activation: formation of E-OMe:OH, transient chirality switch.
Deactivation: hydrolysis of E-OMe:OH to E-OH:OH.
From SI Figure 16a. the rate constant of activation is calculated taking 85 % conversion of monomer at 150 min (9000 s). The k_d is calculated taking 76 % conversion of activated (100 %-(85 %–20 %)) at time (5 days – 360 min = 6840 min = 410400 s), where 360 min is the beginning of phase where deactivation overtakes activation. Applying a second-order expression:

$$k_a = \frac{1}{9000\ \text{s} \times (0.08\ \text{M} - 0.02\ \text{M})} \ln \frac{0.02\ \text{M} \times (0.08\ \text{M} - (0.02\ \text{M} - 0.85 \times 0.02\ \text{M}))}{0.08\ \text{M} \times (0.02\ \text{M} - 0.85 \times 0.02\ \text{M})}$$
$$= 0.00344241\ \text{M}^{-1}\text{s}^{-1}$$

The deactivation rate constant was determined using the half-life expression for the pseudo-first order.

$$k_d = \frac{0.693}{410400\ \text{s}} = 1.6886 \times 10^{-6}\ \text{s}^{-1}$$

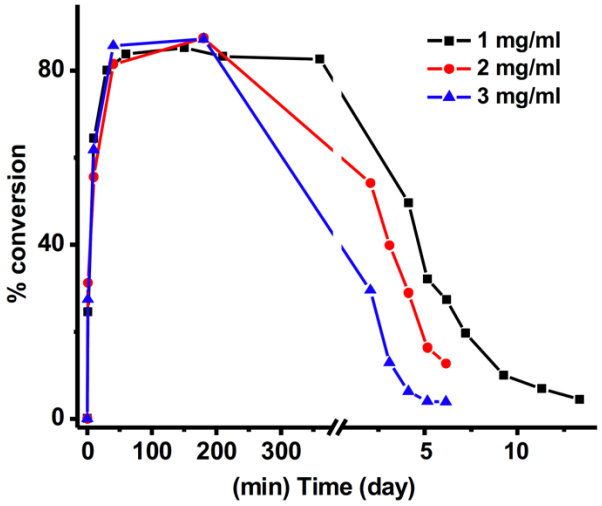


Figure S16a. Time dependent HPLC analysis of 1-E sample with varying concentration of enzyme (α-Chymotrypsin) showing the variation of E-OME:OH species.

$$\frac{f_0}{c_{tot}} = \frac{0.08\ \text{M}}{0.02\ \text{M}} = 4$$

Expression $\frac{k_d g_0}{k_a c_{tot}}$ transforms into $\frac{k_d}{k_a c_{tot}}$ since we used 1st order equation for the deactivation rection.

$$\frac{k_d g_0}{k_a c_{tot}} \Longrightarrow \frac{k_d}{k_a c_{tot}} = \frac{1.6886 \times 10^{-6}\ \text{s}^{-1}}{0.00344241\ \text{M}^{-1}\text{s}^{-1} \times 0.02\ \text{M}} = 0.02453$$

Hermans 2020 — ref. 61

Conditions: 54 mM SachCHO, 312 mM GdL, 190 mM HMTA, 191 mM DT, pH 8

Activation: formation of $SachSO_3^-$ from SachCHO. The half-life (60 s) and corresponding concentration of $SachSO_3^-$ (0.027 M) were determined from Fig. 1e:

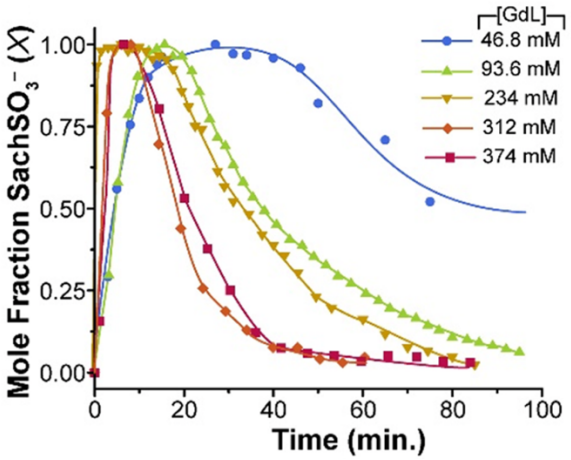


Figure 1e. Evolution of $SachSO_3^-$ in the transient cycles as seen by NMR. Lines to guide the eye. Initial thermal gel with [SachCHO] = 54 mM, and [HMTA] = 190 mM, [DT] = 191 mM in chemically fueled transient cycles.

$$k_a = \frac{1}{60\ \text{s} \times (0.191\ \text{M} - 0.054\ \text{M})} \ln \frac{0.054\ \text{M} \times (0.191\ \text{M} - 0.027\ \text{M})}{0.191\ \text{M} \times 0.027\ \text{M}} = 0.06578\ \text{M}^{-1}\text{s}^{-1}$$

Deactivation: formation of SachCHO from $SachSO_3^-$ and formaldehyde released from HMTA by GdL. 1 equiv. HMTA releases 6 equiv. HCHO, therefore $g_0 = 0.19\ \text{M} \times 6 = 1.14\ \text{M}$. The half-life was determined from the Fig. 1e as time needed to reach 0.5 mole fraction from its maximum value (data for 312 mM GdL):

$$t_{1/2} = 13\ \text{min} = 780\ \text{s}$$

$$k_d = \frac{1}{780\ \text{s} \times (1.14\ \text{M} - 0.054\ \text{M})} \ln \frac{0.054\ \text{M} \times (1.14\ \text{M} - 0.027\ \text{M})}{1.14\ \text{M} \times 0.027\ \text{M}} = 0.00078998\ \text{M}^{-1}\text{s}^{-1}$$

Finally,

$$f_0 = 0.191\ \text{M}$$

$$c_{tot} = 0.054\ \text{M}$$

$$\frac{f_0}{c_{tot}} = 3.537$$

$$\frac{k_d g_0}{k_a c_{tot}} = \frac{0.00078998\ \text{M}^{-1}\text{s}^{-1} \times 1.14\ \text{M}}{0.06578\ \text{M}^{-1}\text{s}^{-1} \times 0.054\ \text{M}} = 0.25352$$

Boekhoven 2017 — ref. 62

Conditions: 10 mM Fmoc-D, 10 mM EDC, pH 6 buffered water.

Activation: formation of anhydride upon reaction with EDC.

The activation rate constant was taken from SI Table 1 $k_a = 0.12\ \text{M}^{-1}\text{s}^{-1}$.

Deactivation: hydrolysis of Fmoc-D anhydride "was slowed down two orders of magnitude if assemblies were formed", i.e. self-assembly as a negative feedback on deactivation (molecules within aggregates are protected from water that causes hydrolysis).

The deactivation rate constant was taken from the SI Table 1 (calculated from k_5), see also SI p. 2:

$$k_d = 0.00038\ \text{s}^{-1}$$

$$\frac{f_0}{c_{tot}} = \frac{0.01\ \text{M}}{0.01\ \text{M}} = 1$$

Expression $\frac{k_d g_0}{k_a c_{tot}}$ transforms into $\frac{k_d}{k_a c_{tot}}$ since we used 1st order equation for the deactivation reaction.

$$\frac{k_d g_0}{k_a c_{tot}} \Longrightarrow \frac{k_d}{k_a c_{tot}} = \frac{0.00038\ \text{s}^{-1}}{0.12\ \text{M}^{-1}\text{s}^{-1} \times 0.01\ \text{M}} = 0.31667$$

Thordarson 2018 — ref. 64

Conditions: pH 2.70 (0.1 M citric acid/trisodium citrate buffer), 5 mM DBC (N, N'-dibenzoyl-L-cystine), 5 mM TCEP.

$$c_{tot} = 0.005\ \text{M}$$

Activation: protonation of DBC.

SI p.6, gelation time of 5 mM DBC in the presence of 5 mM TCEP (deactivation reactant) is 30 sec. Half-life of activation is roughly estimated to be

$$t_{0.5} = 4\ \text{s}$$

Concentration of citric acid was determined using Henderson-Hasselbalch equation and pK_a 6.40.

$$f_0 = 0.09998\ \text{M}$$

Activation rate constant was approximated using 1st order equation:

$$k_a = \frac{0.693}{4\ \text{s}} = 0.17325\ \text{s}^{-1}$$

Deactivation: reduction of disulfide bond of DBC using 5 mM TCEP, formation of N-benzoyl-L-cysteine.

Deactivation rate constant is given in the manuscript $k_d = 0.443 \times 10^{-3}\text{s}^{-1}$.

$$\frac{f_0}{c_{tot}} = \frac{0.09998\ \text{M}}{0.005\ \text{M}} = 19.996$$

Expression $\frac{k_d g_0}{k_a c_{tot}}$ transforms into $\frac{k_d}{k_a}$ since we used 1st order equations.

$$\frac{k_d g_0}{k_a c_{tot}} \Longrightarrow \frac{k_d}{k_a} = \frac{0.000443\ \text{s}^{-1}}{0.17325\ \text{s}^{-1}} = 0.00256$$

Authors indicate that deactivation of DBC can happen only after its dissociation from gel into solution which suggests that the system is coupled.

Guan 2020 — ref. 65

Conditions: 10 mM CSH, 200 mM DTT, 150 mM H_2O_2, pH 6.0.

Activation: oxidation of thiol in CSH (N-benzoyl-L-cysteine) by H_2O_2 to form disulfide in CSSH.

This process was studied separately from the reaction cycle using 0.005 M of H_2O_2 and 0.01 M of CSH (Figure S9).

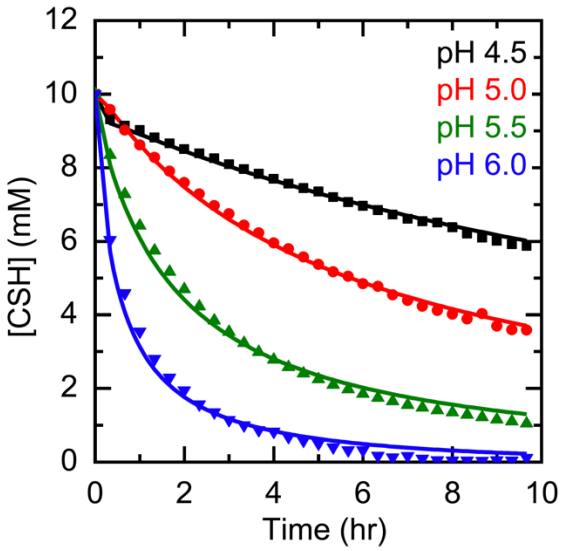


Figure S9. Oxidation of CSH by H_2O_2 ($[H_2O_2]_0$ = 5 mM, $[CSH]_0$ = 10 mM).

Oxidation of thiols by hydrogen peroxide is a 2nd order reaction[5]. Formation of disulfide from thiol is a multi-step process that goes via dominant intermediate sulfenic acid RSOH.[5] Rate-determining step is thiolate oxidation[6]:

$$RS^- + H_2O_2 \rightarrow RSOH + H_2O$$

Assuming that deprotonation of CSH is happening faster than CS– oxidation we limit formation of disulfide to two consecutive second-order steps:

1. Formation of sulfenic acid (see an equation above), k_1
2. $RSOH + RS^- \rightarrow RSSR + HO^-$; k_2

Solving linear differential equations for this system and applying $k_1 \ll k_2$ condition[1]:

$$\tau = \frac{1}{2}\left(\frac{1}{\beta} - 1\right), \alpha = \beta$$

Where $\alpha = \frac{[CSH]_t}{[CSH]_0}$; $\beta = \frac{[H_2O_2]_t}{[H_2O_2]_0}$; $\tau = [H_2O_2]_0 k_1 t$

Experimental values were taken from Figure S9, blue triangles data points (pH 6.0) at t = 2 h = 7200 s:

$$[H_2O_2]_t = [H_2O_2]_0 - \frac{[CSH]_0 - [CSH]_t}{2} = 0.005\ \text{M} - \frac{0.010\ \text{M} - 0.002\ \text{M}}{2} = 0.001\ \text{M}$$

$$\beta = \frac{0.001\ \text{M}}{0.005\ \text{M}} = 0.2$$

$$0.005\ \text{M} \times 7200\ \text{s} \times k_1 = \frac{1}{2}\left(\frac{1}{0.2} - 1\right)$$

$$k_1 = 0.055556\ \text{M}^{-1}\text{s}^{-1} = k_a$$

It is worth to mention that the calculated k_a is much smaller than rate constants of cysteine oxidation with hydrogen peroxide that can be found in literature. Depending on pH, the solvent system, concentration and temperature it can be much higher, e.g. $k = 14\ \text{M}^{-1}\text{s}^{-1}$.[7]

Deactivation: reduction of disulfide bonds in CSSC by DTT.

The rate constant was calculated using the data from Figure S11 (pH 6, blue triangles).

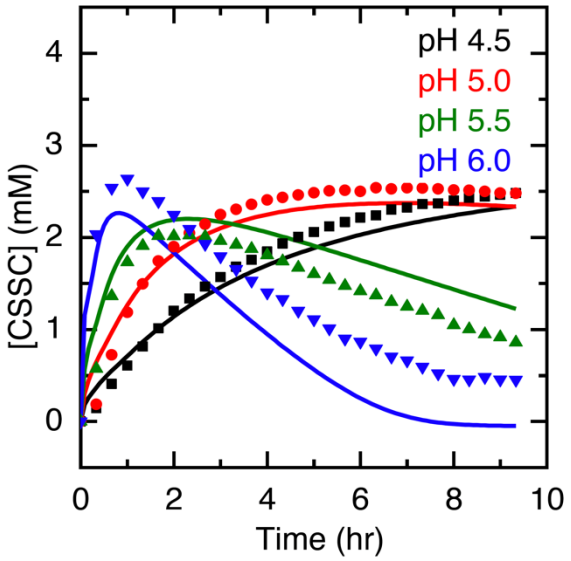


Figure S11. Full active material system ($[CSH]_0$ = 10 mM, $[DTT]_0$ = 200 mM, $[H_2O_2]_0$ = 150 mM).

Concentration of DTT that was present at the time 1.1 h (the time of the highest activation point) was calculated assuming that all hydrogen peroxide that did not react with CSH reacted with DTT and that thiol oxidation proceeds much faster than thiol reduction by DTT:

$$[DTT]_{1.1h} = [DTT]_0 - \left([H_2O_2]_0 - [H_2O_2]_{reacted\ with\ CSH}\right)$$
$$= 0.2\ \text{M} - (0.15\ \text{M} - 0.0026\ \text{M}) = 0.0526\ \text{M}$$

This concentration was taken as initial DTT concentration for the deactivation process. The second-order rate constant for the deactivation process at time t = 4.2 h − 1.1 h = 3.1 h = 11160 s can be written as:

$$k_d = \frac{1}{([DTT]_{1.1h} - [CSSC]_{1.1h}) \times t} \ln \frac{[CSSC]_{1.1h} \times [DTT]_t}{[DTT]_{1.1h} \times [CSSC]_t}$$
$$= \frac{1}{(0.0526\ \text{M} - 0.0026\ \text{M}) \times 11160\ \text{s}} \ln \frac{0.0026\ \text{M} \times (0.0526\ \text{M} - 0.0013\ \text{M})}{0.00526\ \text{M} \times 0.0013\ \text{M}}$$
$$= 0.0011974\ \text{M}^{-1}\text{s}^{-1}$$

Finally:

f_0 = 0.15 M, total hydrogen peroxide concentration

c_{tot} = 0.005 M, total concentration of assembling unit is half of the concentration of thiol.

$$\frac{f_0}{c_{tot}} = \frac{0.15\ \text{M}}{0.005\ \text{M}} = 30$$

$$\frac{k_d g_0}{k_a c_{tot}} = \frac{0.0011974\ \text{M}^{-1}\text{s}^{-1} \times 0.0526\ \text{M}}{0.055556\ \text{M}^{-1}\text{s}^{-1} \times 0.005\ \text{M}} = 0.22673$$

van Esch 2015 — ref. 67

Conditions: monomer 2a 50 mM, fuel DMS 200 mM.

Activation: methylation of carboxylic acid groups of the monomer by DMS.

Deactivation: hydrolysis of methyl easter resulting in fiber collapse.

Activation and deactivation rate constants for different pH were taken from the SI Table S3; we calculated system's coordinates accounting for the different conditions (pH, hydrolysis of DMS which results in only 10 % activation possible, see SI p.7):

Reference in Figure 2a in the main text	f_0, M	pH	k_a, $M^{-1}h^{-1}$	k_a, $M^{-1}h^{-1}$	g_0 = [OH], M	$\frac{k_d g_0}{k_a c_{tot}}$	$\frac{f_0}{c_{tot}}$
117A	0.2	9	3.1	2.80E+05	1E-05	18.06	4
117B	0.2	10	2.5	5.60E+04	1E-04	44.80	4
117C	0.2	11	1.6	3.80E+04	1E-03	475.00	4
117D	0.02 (only 10 % of DMS reacts with the monomer)	10	2.5	5.60E+04	1E-04	44.80	0.4

van Esch 2010 — ref. 66

Conditions: monomer DBC 0.1 M, fuel MeI 0.1 M, pH 7.

Activation: methylation of carboxylic acid groups of the monomer by MeI. The activation rate constant was estimated using data from the main text, p.2, (after 3 h 0.25 mM of monoester formed) and from the Figure 3a using second-order rate expression for the equal reactants concentrations:

$$k_a = \frac{1}{(3\text{ h} \times 3600)}\left(\frac{1}{0.1\text{ M} - 0.00025\text{ M}} - \frac{1}{0.1\text{ M}}\right) = 2.321 \times 10^{-6}\ \text{M}^{-1}\text{s}^{-1}$$

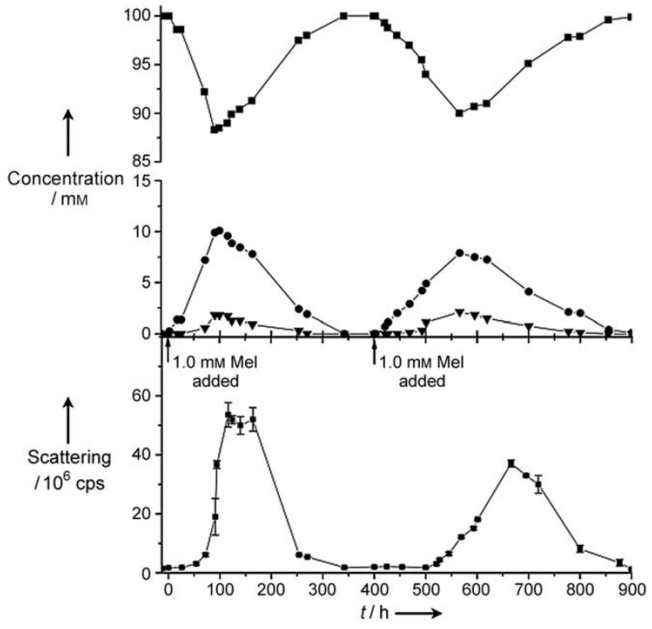


Figure 3a. Concentrations of DBC (squares), DBC-OMe (circles), and DBC-OMe_2 (triangles) as determined by HPLC–MS over time. After 3 h the formation of DBC-OMe was observed. After 100 h the maximum concentrations of DBC-OMe and DBC-OMe_2 were reached and the hydrolysis rate is higher than the methylation rate. Quantitative measurement of aggregation by dynamic light scattering (DLS; l=633 nm) in counts per second. Aggregation is observed after 70 h.

Deactivation: hydrolysis of methyl ester resulting in gel solubilization. Hydrolysis was approximated to the pseudo-first order conditions assuming no formation of demethylated product and no MeI hydrolysis.

$$t_{0.5} = 220\ \text{h} - 100\ \text{h} = 120\ \text{h} = 432000\ \text{s}$$

$$k_d = \frac{0.693}{432000\ \text{s}} = 1.604 \times 10^{-6}\ \text{s}^{-1}$$

Expression $\frac{k_d g_0}{k_a c_{tot}}$ transforms into $\frac{k_d}{k_a c_{tot}}$ since we used 1st order equation for the deactivation rection.

$$\frac{k_d g_0}{k_a c_{tot}} \Longrightarrow \frac{k_d}{k_a c_{tot}} = \frac{1.604 \times 10^{-6}\ \text{s}^{-1}}{2.321 \times 10^{-6}\ \text{M}^{-1}\text{s}^{-1} \times 0.1\ \text{M}} = 6.91267$$

$$\frac{f_0}{c_{tot}} = \frac{0.1}{0.1} = 1$$

References for Section 6

(1) Moore, J. W.; Pearson, R. G. *Kinetics and Mechanism*; John Wiley & Sons, 1981.

(2) Kovacs, K.; McIlwaine, R.; Gannon, K.; Taylor, A. F.; Scott, S. K. Complex Behavior in the Formaldehyde−Sulfite Reaction. *J. Phys. Chem. A* **2005**, *109* (1), 283–288. https://doi.org/10.1021/jp0464324.

(3) Sjöberg, B.; Foley, S.; Cardey, B.; Fromm, M.; Enescu, M. Methionine Oxidation by Hydrogen Peroxide in Peptides and Proteins: A Theoretical and Raman Spectroscopy Study. *Journal of Photochemistry and Photobiology B: Biology* **2018**, *188*, 95–99. https://doi.org/10.1016/j.jphotobiol.2018.09.009.

(4) Aykent, G.; Zeytun, C.; Marion, A.; Özçubukçu, S. Simple Tyrosine Derivatives Act as Low Molecular Weight Organogelators. *Sci Rep* **2019**, *9* (1), 4893. https://doi.org/10.1038/s41598-019-41142-z.

(5) Ashby, M. T.; Nagy, P. On the Kinetics and Mechanism of the Reaction of Cysteine and Hydrogen Peroxide in Aqueous Solution. *Journal of Pharmaceutical Sciences* **2006**, *95* (1), 15–18. https://doi.org/10.1002/jps.20521.

(6) Barton, J. P.; Packer, J. E.; Sims, R. J. Kinetics of the Reaction of Hydrogen Peroxide with Cysteine and Cysteamine. *J. Chem. Soc., Perkin Trans. 2* **1973**, No. 11, 1547. https://doi.org/10.1039/p29730001547.

(7) Luo, D.; Smith, S. W.; Anderson, B. D. Kinetics and Mechanism of the Reaction of Cysteine and Hydrogen Peroxide in Aqueous Solution. *Journal of Pharmaceutical Sciences* **2005**, *94* (2), 304–316. https://doi.org/10.1002/jps.20253.

1.7 References

1. KEGG pathway: Metabolic pathways - Reference pathway. https://www.genome.jp/kegg-bin/show_pathway?map01100.
2. David L. Nelson, Michael M. Cox. *Lehninger Principles Of Biochemistry*. (W. H. Freeman, 2004).
3. Feinberg, M. *Foundations of Chemical Reaction Network Theory*. vol. 202 (Springer International Publishing, 2019).
4. Roekel, H. W. H. van *et al.* Programmable chemical reaction networks: emulating regulatory functions in living cells using a bottom-up approach. *Chem. Soc. Rev.* **44**, 7465–7483 (2015).
5. Walker, S. I., Packard, N. & Cody, G. D. Re-conceptualizing the origins of life. *Philos Trans A Math Phys Eng Sci* **375**, 20160337 (2017).
6. Epstein, I. R. & Pojman, J. A. *An Introduction to Nonlinear Chemical Dynamics: Oscillations, Waves, Patterns, and Chaos*. (Oxford University Press, 1998).
7. Nicolis, G. & Prigogine, I. *Self-Organization in Nonequilibrium Systems: From Dissipative Structures to Order Through Fluctuations*. (Wiley, 1977).
8. Lehn, J.-M. Toward Self-Organization and Complex Matter. *Science* **295**, 2400–2403 (2002).
9. Hagan, M. F. & Baskaran, A. Emergent self-organization in active materials. *Current Opinion in Cell Biology* **38**, 74–80 (2016).
10. Camazine, S. *et al. Self-Organization in Biological Systems*. *Self-Organization in Biological Systems* (Princeton University Press, 2020). doi:10.1515/9780691212920.
11. Isaeva, V. V. Self-organization in biological systems. *Biol Bull Russ Acad Sci* **39**, 110–118 (2012).
12. Wedlich-Söldner, R. & Betz, T. Self-organization: the fundament of cell biology. *Philosophical Transactions of the Royal Society B: Biological Sciences* **373**, 20170103 (2018).
13. Lehn, J.-M. Towards Complex Matter: Supramolecular Chemistry and Self-organization*. *European Review* **17**, 263–280 (2009).
14. Halley, Julianne. D. & Winkler, D. A. Consistent concepts of self-organization and self-assembly. *Complexity* **14**, 10–17 (2008).
15. Mnif, M. & Müller-Schloer, C. Quantitative Emergence. in *Organic Computing — A Paradigm Shift for Complex Systems* (eds. Müller-Schloer, C., Schmeck, H. & Ungerer, T.) 39–52 (Springer, 2011). doi:10.1007/978-3-0348-0130-0_2.
16. Arango-Restrepo, A., Barragán, D. & Rubi, J. M. Self-assembling outside equilibrium: emergence of structures mediated by dissipation. *Phys. Chem. Chem. Phys.* **21**, 17475–17493 (2019).
17. Al-Bassam, J. & Chang, F. Regulation of microtubule dynamics by TOG-domain proteins XMAP215/Dis1 and CLASP. *Trends in Cell Biology* **21**, 604–614 (2011).
18. Holy, T. E., Dogterom, M., Yurke, B. & Leibler, S. Assembly and positioning of microtubule asters in microfabricated chambers. *Proceedings of the National Academy of Sciences* **94**, 6228–6231 (1997).
19. Laan, L. *et al.* Cortical Dynein Controls Microtubule Dynamics to Generate Pulling Forces that Position Microtubule Asters. *Cell* **148**, 502–514 (2012).
20. Dogterom, M. & Yurke, B. Microtubule Dynamics and the Positioning of Microtubule Organizing Centers. *Phys. Rev. Lett.* **81**, 485–488 (1998).
21. Brouhard, G. J. & Rice, L. M. Microtubule dynamics: an interplay of biochemistry and mechanics. *Nat Rev Mol Cell Biol* **19**, 451–463 (2018).
22. Del Giudice, D., Frateloreto, F., Sappino, C. & Di Stefano, S. Chemical Tools for the Temporal Control of Water Solution pH and Applications in Dissipative Systems. *European Journal of Organic Chemistry* **n/a**, (2022).
23. Leng, Z., Peng, F. & Hao, X. Chemical-Fuel-Driven Assembly in Macromolecular Science: Recent Advances and Challenges. *ChemPlusChem* **85**, 1190–1199 (2020).
24. Das, K., Gabrielli, L. & Prins, L. J. Chemically Fueled Self-Assembly in Biology and Chemistry. *Angew. Chem.* **133**, 20280–20303 (2021).
25. Singh, N., Formon, G. J. M., De Piccoli, S. & Hermans, T. M. Devising Synthetic Reaction Cycles for Dissipative Nonequilibrium Self-Assembly. *Advanced Materials* **32**, 1906834 (2020).
26. De, S. & Klajn, R. Dissipative Self-Assembly Driven by the Consumption of Chemical Fuels. *Adv. Mater.* **30**, 1706750 (2018).
27. Wang, Q., Qi, Z., Chen, M. & Qu, D.-H. Out-of-equilibrium supramolecular self-assembling systems driven by chemical fuel. *Aggregate* **2**, e110 (2021).
28. Wang, G. & Liu, S. Strategies to Construct a Chemical-Fuel-Driven Self-Assembly. *ChemSystemsChem* **2**, e1900046 (2020).

29. Heuser, T., Weyandt, E. & Walther, A. Biocatalytic Feedback-Driven Temporal Programming of Self-Regulating Peptide Hydrogels. *Angew. Chem. Int. Ed.* **54**, 13258–13262 (2015).
30. Bai, S. *et al.* Chemical reaction powered transient polymer hydrogels for controlled formation and free release of pharmaceutical crystals. *Chemical Engineering Journal* **414**, 128877 (2021).
31. Del Grosso, E., Amodio, A., Ragazzon, G., Prins, L. J. & Ricci, F. Dissipative Synthetic DNA-Based Receptors for the Transient Loading and Release of Molecular Cargo. *Angewandte Chemie* **130**, 10649–10653 (2018).
32. Chen, R., Neri, S. & Prins, L. J. Enhanced catalytic activity under non-equilibrium conditions. *Nat. Nanotechnol.* **15**, 868–874 (2020).
33. Del Grosso, E., Ragazzon, G., Prins, L. J. & Ricci, F. Fuel-Responsive Allosteric DNA-Based Aptamers for the Transient Release of ATP and Cocaine. *Angewandte Chemie* **131**, 5638–5642 (2019).
34. Fan, B., Zhang, K., Liu, Q. & Eelkema, R. Self-Healing Injectable Polymer Hydrogel via Dynamic Thiol-Alkynone Double Addition Cross-Links. *ACS Macro Lett.* **9**, 776–780 (2020).
35. Lewis, R. W., Klemm, B., Macchione, M. & Eelkema, R. Signal responsive transient coacervation in complex coacervate core micelles. (2021) doi:10.33774/chemrxiv-2021-l3jm1.
36. Solís Muñana, P. *et al.* Substrate-Induced Self-Assembly of Cooperative Catalysts. *Angewandte Chemie International Edition* **57**, 16469–16474 (2018).
37. Shandilya, E. & Maiti, S. Deconvolution of Transient Species in a Multivalent Fuel-Driven Multistep Assembly under Dissipative Conditions. *ChemSystemsChem* **2**, (2020).
38. Ragazzon, G. & Prins, L. J. Energy consumption in chemical fuel-driven self-assembly. *Nature Nanotech* **13**, 882–889 (2018).
39. Arango-Restrepo, A., Rubi, J. M. & Barragán, D. Understanding Gelation as a Nonequilibrium Self-Assembly Process. *J. Phys. Chem. B* **122**, 4937–4945 (2018).
40. Sharko, A., Livitz, D., Piccoli, S. D., Bishop, K. & Hermans, T. Insights into chemically-fueled supramolecular polymers. (2022) doi:10.26434/chemrxiv-2022-jd607.
41. Panja, S., Patterson, C. & Adams, D. J. Temporally-Programmed Transient Supramolecular Gels. *Macromol. Rapid Commun.* 1900251 (2019) doi:10.1002/marc.201900251.
42. Ashkenasy, G., Hermans, T. M., Otto, S. & Taylor, A. F. Systems chemistry. *Chem. Soc. Rev.* **46**, 2543–2554 (2017).
43. Panja, S., Fuentes-Caparrós, A. M., Cross, E. R., Cavalcanti, L. & Adams, D. J. Annealing Supramolecular Gels by a Reaction Relay. *Chem. Mater.* **32**, 5264–5271 (2020).
44. Panzarasa, G., Sai, T., Torzynski, A. L., Smith-Mannschott, K. & Dufresne, E. R. Supramolecular assembly by time-programmed acid autocatalysis. *Mol. Syst. Des. Eng.* **5**, 445–448 (2020).
45. Panzarasa, G., Torzynski, A. L., Sai, T., Smith-Mannschott, K. & Dufresne, E. R. Transient supramolecular assembly of a functional perylene diimide controlled by a programmable pH cycle. *Soft Matter* **16**, 591–594 (2020).
46. Jain, A., Dhiman, S., Dhayani, A., Vemula, P. K. & George, S. J. Chemical fuel-driven living and transient supramolecular polymerization. *Nat Commun* **10**, 450 (2019).
47. Sorrenti, A., Leira-Iglesias, J., Sato, A. & Hermans, T. M. Non-equilibrium steady states in supramolecular polymerization. *Nat Commun* **8**, 15899 (2017).
48. Spitzer, D., Rodrigues, L. L., Straßburger, D., Mezger, M. & Besenius, P. Tuneable Transient Thermogels Mediated by a pH- and Redox-Regulated Supramolecular Polymerization. *Angew. Chem. Int. Ed.* **56**, 15461–15465 (2017).
49. Olivieri, E., Quintard, G., Naubron, J.-V. & Quintard, A. Chemically Fueled Three-State Chiroptical Switching Supramolecular Gel with Temporal Control. *J. Am. Chem. Soc.* **143**, 12650–12657 (2021).
50. Mishra, A. *et al.* Biomimetic temporal self-assembly via fuel-driven controlled supramolecular polymerization. *Nat Commun* **9**, 1295 (2018).
51. Dhiman, S., Jain, A., Kumar, M. & George, S. J. Adenosine-Phosphate-Fueled, Temporally Programmed Supramolecular Polymers with Multiple Transient States. *J. Am. Chem. Soc.* **139**, 16568–16575 (2017).
52. Angulo-Pachón, C. A. & Miravet, J. F. Sucrose-fueled, energy dissipative, transient formation of molecular hydrogels mediated by yeast activity. *Chem. Commun.* **52**, 5398–5401 (2016).
53. Mukhopadhyay, R. D., Choi, S., Sen, S. K., Hwang, I. & Kim, K. Transient Self-assembly Processes Operated by Gaseous Fuels under Out-of-Equilibrium Conditions. *Chem. Asian J.* **15**, 4118–4123 (2020).

54. Mondal, S., Podder, D., Nandi, S. K., Roy Chowdhury, S. & Haldar, D. Acid-responsive fibrillation and urease-assisted defibrillation of phenylalanine: a transient supramolecular hydrogel. *Soft Matter* **16**, 10115–10121 (2020).
55. Bal, S., Das, K., Ahmed, S. & Das, D. Chemically Fueled Dissipative Self-Assembly that Exploits Cooperative Catalysis. *Angew. Chem. Int. Ed.* **58**, 244–247 (2019).
56. Bal, S., Ghosh, C., Ghosh, T., Vijayaraghavan, R. K. & Das, D. Non-Equilibrium Polymerization of Cross-β Amyloid Peptides for Temporal Control of Electronic Properties. *Angewandte Chemie International Edition* **59**, 13506–13510 (2020).
57. Afrose, S. P., Bal, S., Chatterjee, A., Das, K. & Das, D. Designed Negative Feedback from Transiently Formed Catalytic Nanostructures. *Angew. Chem. Int. Ed.* **58**, 15783–15787 (2019).
58. Pappas, C. G., Sasselli, I. R. & Ulijn, R. V. Biocatalytic Pathway Selection in Transient Tripeptide Nanostructures. *Angew. Chem. Int. Ed.* **54**, 8119–8123 (2015).
59. Debnath, S., Roy, S. & Ulijn, R. V. Peptide Nanofibers with Dynamic Instability through Nonequilibrium Biocatalytic Assembly. *J. Am. Chem. Soc.* **135**, 16789–16792 (2013).
60. Kumar, M. *et al.* Amino-acid-encoded biocatalytic self-assembly enables the formation of transient conducting nanostructures. *Nature Chem* **10**, 696–703 (2018).
61. Singh, N., Lainer, B., Formon, G. J. M., De Piccoli, S. & Hermans, T. M. Re-programming Hydrogel Properties Using a Fuel-Driven Reaction Cycle. *J. Am. Chem. Soc.* **142**, 4083–4087 (2020).
62. Tena-Solsona, M. *et al.* Non-equilibrium dissipative supramolecular materials with a tunable lifetime. *Nat Commun* **8**, 15895 (2017).
63. Dai, K. *et al.* Regulating Chemically Fueled Peptide Assemblies by Molecular Design. *J. Am. Chem. Soc.* **142**, 14142–14149 (2020).
64. Wojciechowski, J. P., Martin, A. D. & Thordarson, P. Kinetically Controlled Lifetimes in Redox-Responsive Transient Supramolecular Hydrogels. *J. Am. Chem. Soc.* **140**, 2869–2874 (2018).
65. Ogden, W. A. & Guan, Z. Redox Chemical-Fueled Dissipative Self-Assembly of Active Materials. *ChemSystemsChem* **2**, (2020).
66. Boekhoven, J. *et al.* Dissipative Self-Assembly of a Molecular Gelator by Using a Chemical Fuel. *Angewandte Chemie International Edition* **49**, 4825–4828 (2010).
67. Boekhoven, J., Hendriksen, W. E., Koper, G. J. M., Eelkema, R. & van Esch, J. H. Transient assembly of active materials fueled by a chemical reaction. *Science* **349**, 1075–1079 (2015).
68. Kubota, R. *et al.* Force generation by a propagating wave of supramolecular nanofibers. *Nat Commun* **11**, 3541 (2020).

Chapter 2. Cyclic peptides and their structures

Abstract

Cyclic peptides (CPs) and their nanomaterials are proven to have great potential and application in various fields of science. They are highly interesting because of their versatility, easy functionalization, and biocompatibility. In this chapter, we give a brief overview of the recent advances and challenges in the development of cyclic peptide nanotubes (CPNs). Robustness and propensity to stack make CPNs an excellent model system for studying the behavior of natural fibrous structures like microtubules and actin filaments. To bring the formation and destruction of CPNs out of equilibrium, we set out to develop a network of coupled reactions based on a reversible addition of nucleophiles to a Michael acceptor. First, the cyclization of different peptides is studied, and reaction conditions are optimized to yield mostly monocyclic peptides. The reaction kinetics are followed by HRMS and UV-Vis. The rate of cyclization is dependent on peptide length: shorter peptide cyclizes faster and gives less side products. Finally, the structure of formed CPs and their self-assembly mechanism is discussed.

2.1 General introduction

2.1.1 Cyclic peptides

Interest in cyclic peptides (CPs) stems from structural studies of naturally occurring polypeptides that are often cytotoxic, and because CPs can form interesting cavities that can selectively complex ions.[1] The majority of natural cyclic peptides originates in plants, fungi, and bacteria, where they help to defend the organism against external threats. Their length ranges from just two amino acid residues (2,5-diketopiperazines = cyclo(Gly-Gly)) to hundreds of amino acids in the chain.[2] The first discovered natural cyclic peptide, Gramicidin S (Figure 2.1, **1**), was obtained from bacteria and used as an antibiotic during the Second World War.[3] Nowadays, it is produced synthetically and used as a local antibacterial agent. Other famous examples are mammal hormones oxytocin and vasopressin, actively used as model compounds in industrial drug development.[4] Cyclic peptides and their derivatives (e.g., N-methylated, with unnatural amino acids, or with ester linkers instead of amide (Figure 2.1, **2**–**6**)) are valued for their antitumor, antimicrobial, antifungal, antibacterial, antiviral and anti-inflammatory properties, to name a few.[5–7]

A great advantage of CPs over linear peptides is enhanced cellular membrane permeability and resistance against enzymes. A closed circular structure reduces levels of conformational freedom and makes molecules rigid which can lead to an increase in stability and, consequently, biological activity.[8,9] Interestingly, an enhancement of cyclic peptides activity *in vivo* correlates with higher resistance to common degradation ways, such as exo- and endopeptidase digestion. Small cyclic peptides (5–6 residues) showed higher resistance against enzymes experiencing greater conformational restraints compared to larger analogs.[8] Additionally, the reduced degree of flexibility leads to improved receptor binding affinity, however, it is also seen by some research groups as a correlation and not as a direct causality.[9] Compared to their linear analogs, CPs are more stable under thermal and chemical stress.

A notable feature commonly found in natural cyclic peptides is polycyclic motifs, i.e., multiple sequential rings within the peptide molecule. The presence of several cycles limits the overall conformation of a molecule, defining its structure in solution and improving the peptide's binding properties.[10] Synthetic versions of polycyclic peptides are usually small, relatively easier to obtain and show better membrane penetration than large monocyclic proteins.[11]

There are many ways to synthesize cyclic peptides depending on their further application and this field continues to develop rapidly.[1,12–16] The most common methods to cyclize peptides include (i) closing the linear molecule end-to-end (N terminus to C terminus), (ii) end-to-backbone (N/C terminus to the thiol group of Cys, the hydroxyl group of Ser, Thr, Tyr, etc.), (iii) backbone-to-backbone (disulfide bridges, biaryl or biaryl-ether linkages) or (iv) more complicated arrangements and combinations of methods.[1] Synthesis of cyclic peptides can be challenging, especially depending on the aimed size of the macrocycle, since the larger the peptide needs to be cyclized, the higher the risk of side reactions due to the intermolecular reactivity.[17] For small peptides (< 7 residues), ways of cyclization should also be chosen carefully as the formation of the product is less favorable because of the higher impact of constraints.[18] Introducing turning structural elements into peptides can facilitate cyclization.[19]

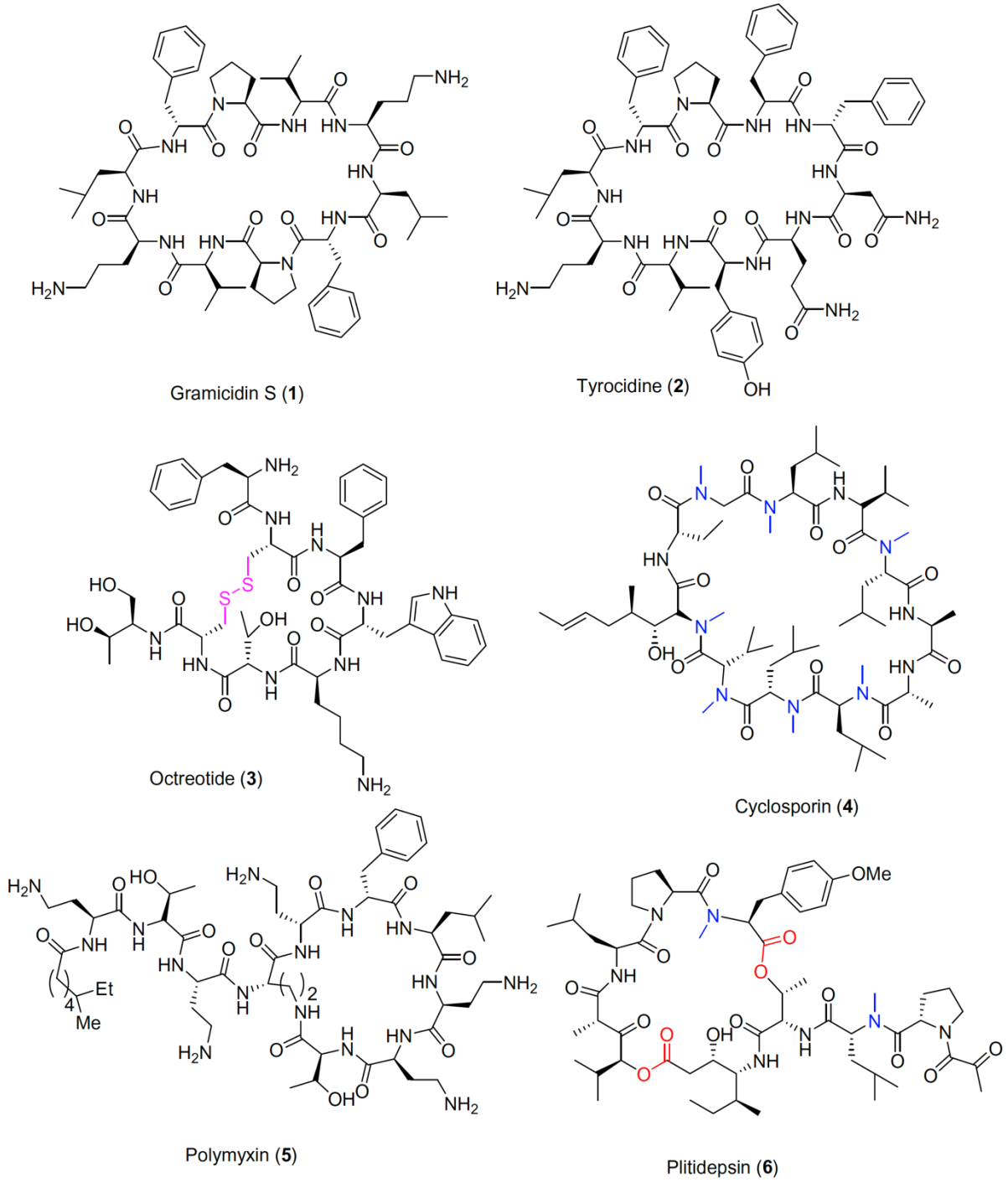


Figure 2.1 | Several natural and synthetic CPs drugs are available on the market. Reprinted from ref. 7.

The functionality of peptides is highly tunable even by changing only one of its amino acid residues. This can be useful in adjusting the properties of desired final medicine. Hence, substituting phenylalanine residue with tyrosine causes a complete loss of immunosuppressive activity of toxic mushroom metabolite antamanide.[20] Any change in peptide sequence or conformation of its amino acids is reflected in biological activity and self-assembly properties. Currently, extensive research and use of cyclic peptides and their derivatives or mimics are challenged by a high variety of properties that can dramatically change with the smallest structural alteration. More efforts are directed to make cyclic peptide isolation and structural

identification more accessible and simpler. The lack of comprehensive understanding of structure-properties relations, including the self-assembling process, may pose another obstacle to immediate access of these natural products to the market. It was shown that the number of key factors to investigate further is significant: pH, concentration, chain modifications, reaction time, and influence of mechanical agitation.[21,22]

Despite the challenges, unarguable advantages of cyclic peptides cause expansion of their application area in pharmacological and agricultural assessments, and toxicity studies.

2.1.2 Planar CPs and cyclic peptide nanotubes (CPNs)

Cyclic peptides that are composed of an even number of D/L-amino acids adopt an overall flat conformation where all the residues point out of the macrocycle, leaving solvent-filled central space within it, and carbonyl and N-H bonds of amide are oriented perpendicularly to the plane.[23] Planarity and directionality of residues favor interactions between cyclic peptide molecules via net H-bonds between amide groups causing macrocycles to stack on each other resulting in β-sheet tubular hollow structures—cyclic peptide nanotubes (CPNs).[24] First extensively studied by Ghadiri and Granja,[23] these structures can be additionally stabilized by interactions of peptide sidechains: π-π stacking or Coulombic attraction—which contribute to the remarkable robustness of each nanotube or bundles of nanotubes. The latter feature is so prominent, that CPNs are often compared to amyloid fibers, bone, and enamel, all being very stiff natural materials.[25] Their unique stability is due to the efficient balance between hydrophobic and hydrophilic forces and is beneficial for material and biomedical science.[26–29] CPNs are explored for their antiviral and antibacterial effects, potential in drug and gene delivery, as conductors in organic electronics, and ion transport channels, which can be reinforced by the properties of each CP itself (Figure 2.2).[30–32]

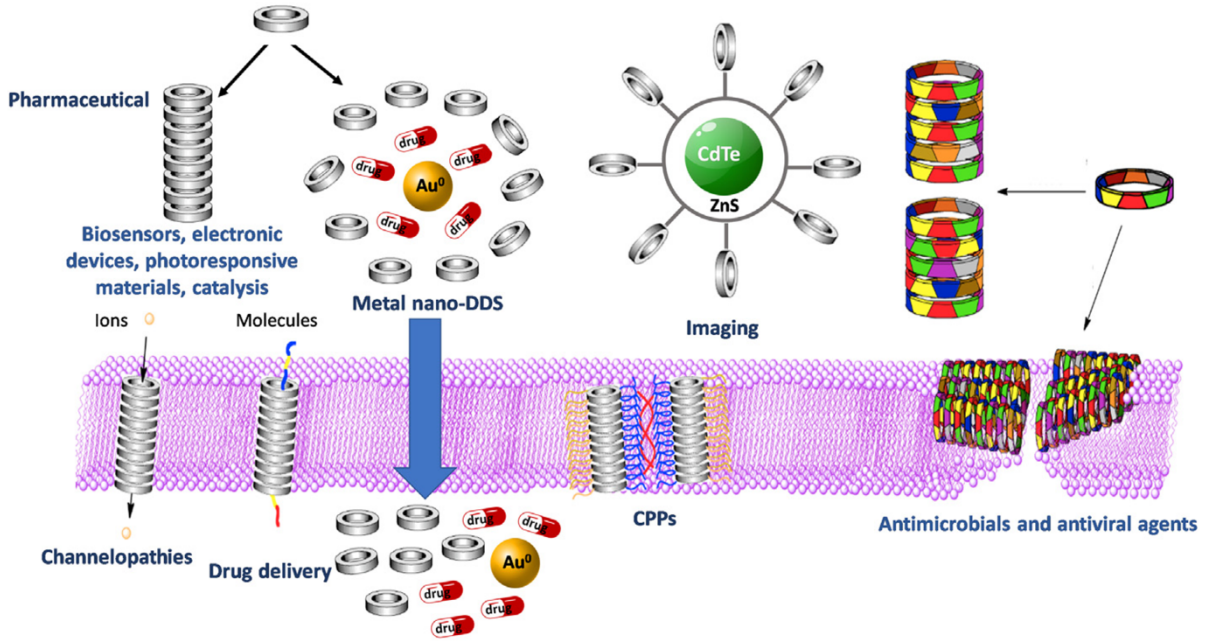


Figure 2.2 | Schematic illustration of some of the CPNs applications. Reprinted from ref. 1.

A drawback of great rigidity of CPNs is the lack of dynamics: once CPNs are formed, it is almost impossible to disassemble them without destroying the CPs that constitute them along

the way. Different approaches exist to influence CPNs stability without compromising on the desired properties, for example, by changing the CP backbone. The latter is usually done by including amino acid residues that influence H-bonding directly (N-Methylation, use of noncanonical amino acids) or indirectly (introducing bulky or pH-sensitive groups into side chains to promote CPNs disassembly).[33,34]

2.1.3 Types of CPs and self-assembly

There are a few types of CPs that form CPNs: cyclic D,L-α-peptides, cyclic β-peptides, and cyclic α-γ-peptides, cyclic δ-peptides, cyclic ε-peptides (Figure 2.3).[30] The first discovered cyclic D,L-α-peptide nanotubes consist of cyclic peptide antiparallel β-sheet structure, H-bonds between amides of neighboring peptides are bidirectional and are perpendicular to the cyclic peptide plane and parallel to the tube axis. The amino acid sequence influences the lateral packing of CPNs: hydrophobic residues promote inter-tube interactions and, hence, bundle formation, and hydrophilic residues disfavor it.[24] A delicate design of side-chain non-covalent interactions is required to stir the formation of 2D and 3D assemblies. This strategy was implemented by Montenegro and co-workers in designing CPNs sheets[35] and studying dependencies[36] of bulk structure on the amino acid sequence. The smallest CPs that can form CPNs consist of 4 amino acids, the biggest—of 12.

Cyclic β-peptides have similar assembled structures, however, amide-carbonyl H-bonds are unidirectional and not parallel to the tube axis, which is the consequence of parallel stacking and the anisotropic character of the tube.[37] Cyclic α-γ-peptides strongly differ from the previous two types and have α-α and γ-γ H-bonding interactions.[32] Further studies of the correlation between peptide structure and functions reveal an influence of the presence of noncanonical amino acids (β, γ, δ, ε) within the peptide sequence, which is mainly used to tailor the outer surface of CPNs. The interior of planar cyclic peptides, composed of α- and β-amino acids, is hydrophilic and cannot be changed via modification of residues but by using γ-amino acids.[38] This approach introduces hydrophilic functional groups inside CPNs, which is important for channel and drug delivery applications.[39] δ- and ε-peptides can form large cavities and require a robust side-chain modification to stabilize macrocycle structure and avoid intramolecular β-sheet formation.[40,41]

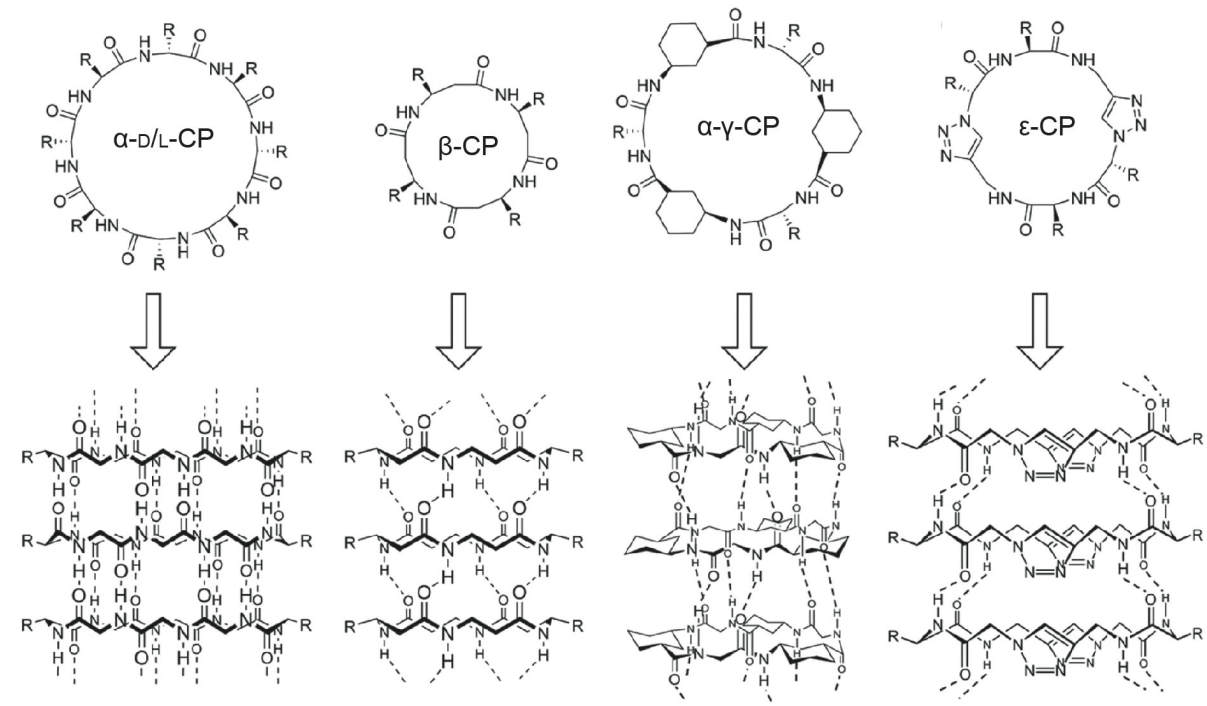


Figure 2.3 | Types of planar CPs and corresponding CPNs. Adapted from ref. 25.

The mechanism of cyclic peptides self-assembly is not studied in detail, but some studies indicate cooperativity of CPNs formation, which means the self-assembly proceeds according to a nucleation-elongation mechanism.[25,42] Control over CPs assembly can also be implemented by varying solvent polarity, ionic strength, and pH. From their discovery till now, CPNs studies focus on establishing structure-properties trends, expanding usage of CPNs cavity, developing multifunctional materials that benefit from CPNs robustness, that can decrease it if needed, and be applied in water-based systems.

It is important to note that despite all attractive functional features allowing broad application of CPNs, it is still unclear how to control the dispersity and uniformity of this material at an industrial level; scaled production of cyclic peptides is also restricted by the need for native chemical ligation techniques.[31,43] A possible resolution here might be research towards new simpler ways of peptide cyclization.

Controlling CPNs assembly and disassembly simplifies their use, making them controllable, tunable materials. Therefore, much attention is currently directed toward developing stimuli-responsive, dynamic CP materials.[44] Although having a chemical switch adds a great degree of control over CPs self-assembly, being able to directly pre-determine and program the lifetime of CPNs from the moment of their formation without the need for further intervention opens prospects for applications like timed drug release, transient catalysis, or responsive membrane channels.

2.1.4 Aza-Thiol-Michael addition reactions

Michael additions are reactions between a nucleophile and α, β-unsaturated carbonyl compounds.[45–47] Aza-Michael and thiol-Michael additions have largely contributed to the development of material chemistry.[48–52] These two types of conjugate additions are particularly interesting for biochemistry as they involve one of the most abundant nucleophiles in the living

organisms and can be done in a physiological medium.[53,54] The latter together with the fast, quantitative, and reversible nature of these additions, have made them useful for systems chemists,[55] particularly for dynamic combinatorial chemistry.[56,57] Michael additions are rare in living systems because of the high reactivity of Michael acceptors with the vital nucleophiles, amines and thiols, which can affect such functions as DNA repair or replication.[58,59] On the other side, this reactivity is used to modify enzymes aiming to adjust their functionality[60] or for drug development.[61–63] Also, when desired, the reactivity of Michael acceptors can be modulated to decrease the risks of unwanted interferences *in vivo*.[64–66] Because of their high efficiency, the aza- and thiol-Michael additions are often referred to as 'click'-reactions.[52,67] This is a big addition to the field of click chemistry,[68] as most click reactions are generally irreversible or reversible under severe conditions,[69–71] which is not the case, especially for thiol-Michael additions.[72,73]

A step forward was made by Anslyn and co-workers by introducing click-'declick' chemistry based on different Michael acceptors that can release the substrate upon reacting with a different nucleophile.[67,74] One of these conjugate acceptors, a Meldrum's acid derivative (MAD), can selectively and reversibly couple amine and thiol functional groups in aqueous conditions at neutral pH.[75] The same research group developed a library of different Michael acceptors,[76] but MAD has found a particularly wide application.[77–82]

The addition of an amine and thiol to MAD proceeds sequentially with the release of 2 equivalents of gaseous methyl mercaptan, which can be recognized by the pungent smell of the reaction mixture (Figure 2.4). The unique feature of MAD as a crosslinker is its selectivity because of the drop of electrophilicity upon the addition of the amine, which makes the conjugate acceptor inaccessible for other amine groups, and therefore enables the addition of thiol. The resulting amine-thiol Michael adduct can be 'de-clicked' by a chemical nucleophilic trigger containing a free thiol and alcohol or amine such as dithiothreitol (DTT), cysteine, or ethanedithiol to release the original unmodified amine and thiol quantitatively.[67]

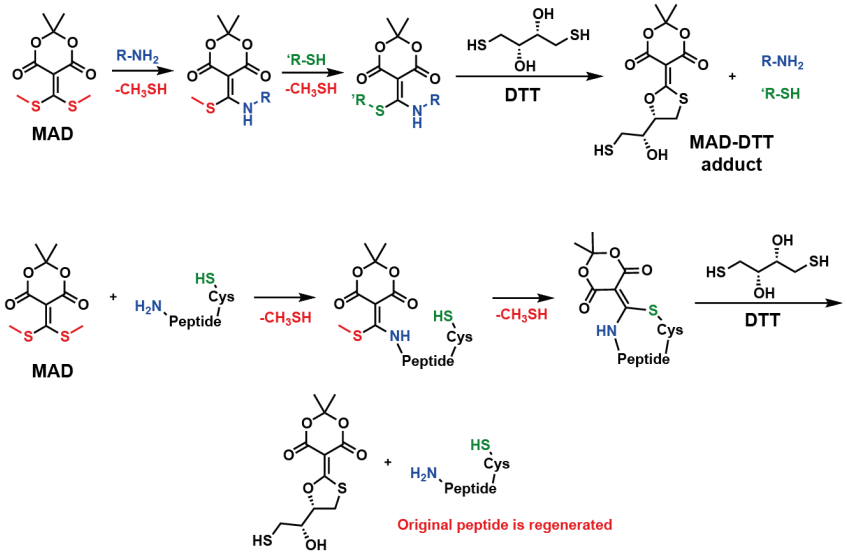


Figure 2.4 | Click-'declick' reaction of MAD with amine and thiol, with peptide. Adapted from ref. 83.

MAD chemistry was proven to work on a wide range of amine-thiol-containing substrates. The original concept was developed on a range of small molecules, each of them containing either amine or thiol groups.[75] Further, the use of these reactions was broadened and worked on

macromolecules—cysteine containing peptides with free N-terminus (Figure 2.4).[83] Finally, the same principle was applied to polymer networks to remodel its architecture by a sequence of click reactions and disulfide link formations.[84] All of the shown examples can be de-clicked to the starting components in a controlled fashion when desired. Owing to the mild reaction conditions and ease of use in a variety of applications, the method is predicted to have wide utility.

2.1.5 Aim of this chapter

Even though the peptide cyclization with MAD has already been shown in the literature, neither has it been demonstrated on D/L-peptides nor have the aspects of self-assembly of the resulting CPs been studied. The objective of this chapter is to study the cyclization kinetics of D/L-peptide containing cysteine with MAD (Figure 2.5), the structure properties of the obtained CPs, and compare it to literature data on CPs synthesized via conventional methods. Our greater goal is to achieve more dynamic, recyclable, but still rigid CPNs that can be accessed in an easy, straightforward way (chapters 2 and 3) and have a predefined lifetime (chapter 4). Development and exploration of the orthogonal reversible reactions that involve bio-abundant functional groups can unlock access to modelling and creating life-like materials and mimicking existing nature motives such as microtubules.

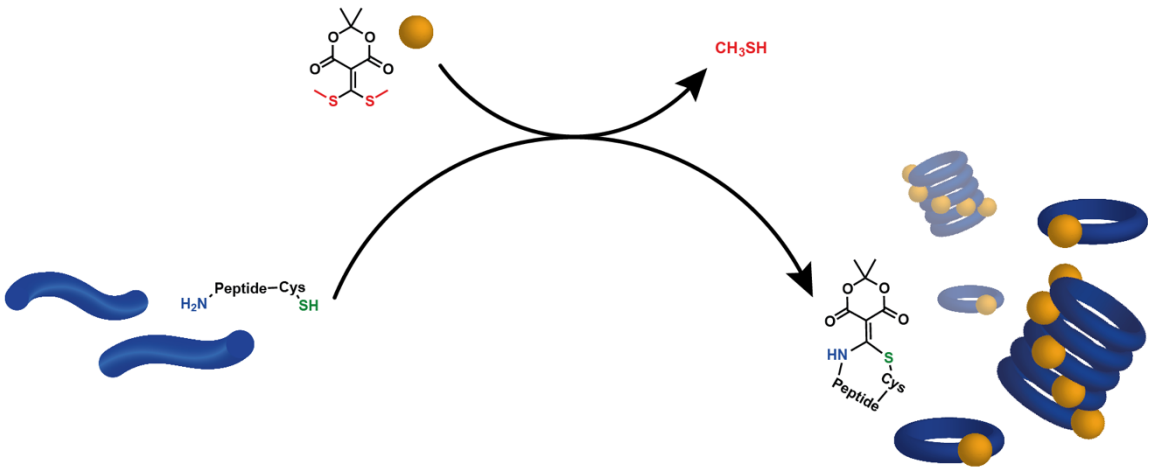


Figure 2.5 | General scheme of the reaction of D/L-peptides with MAD, leading to formation of CPNs.

2.2 Cyclization studies

2.2.1 Peptides design and MAD synthesis

We designed two peptides: i) **P1** (D-Val-Phe-D-Arg-Tyr-D-Val-Trp-D-Arg-Cys-Gly), ii) and **P2** (D-Glu-Trp-D-Val-Cys-D-Arg), see **LP** in Figure 2.7. These specific peptide lengths were chosen to have a distinct difference in cyclic peptide sizes, keeping an even number of amino acids of the opposite chirality, which is required to make a planar CP. The sequence was chosen with the idea of balancing the hydrophobic and hydrophilic character of the macromolecule, which would promote the stacking of CPs but at the same time remain reasonably soluble.[36] MAD was synthesized according to the literature procedure,[85] see Appendix Section 2.6.2.

2.2.2 Initial trials

We set out to compare the cyclization kinetics of both peptides using LC-HRMS and UV-Vis, aiming to compare the activation reaction rates and hence derive what the ideal deactivation rate should be (which will become important in Chapter 3).

The first trials were done to test suitable concentrations and pH to have a transparent clear starting peptide solution. It turned out that the best solvent composition includes DMF, buffer, and ACN. DMF helped peptide solubilization and ACN was added to ensure the solubility of MAD also at high concentrations. We tested phosphate buffer pH 6.0, 7.0, 8.0, and carbonate buffer pH 9.2. At pH 6.0 and 9.2, the reaction with MAD proceeded the slowest, as expected from the literature.[67] In the first case, the fraction of deprotonated amine and thiol was low; in the second case—an N-H group of the peptide-MAD product (**MP**, Figure 2.7) is deprotonated, which decreases the electrophilic nature of MAD, making it inaccessible for thiol.[67] The pH range 7–8 resulted in the most optimal reaction rate and cyclization yields, so we proceeded with pH 7.4. Further, similar experiments were done with varying relative ratios of peptide and MAD to find the conditions that give visible aggregates throughout the reaction. At that time, it seemed that 1.5 eq. of MAD with 1–2 mM peptide gave visually the most promising result: clear non-turbid solution before adding MAD that got turbid and slowly formed aggregates after MAD was added. Later, through HRMS studies, we found that the MAD-peptide product composition depends on the peptide length and is highly sensitive to the amount of MAD. But at that early stage of the research, the fast formation of large aggregates mattered the most because the idea was to obtain cyclic peptide nanotubes visible through a microscope and then optimize already developed chemistry to disassemble them.

For **P1** at 1.41 mM with 1.5 eq. MAD, we observed a gradual increase in turbidity of a solution after 20–25 min, followed by the growth of larger spherical structures that started to sink to the bottom of the vial, yielding fluffy colorless structures in 40–60 min that did not change further. The solution remained yellowish all the time. The AFM images of the crude reaction (after dilution and sonication) and the pure CP product look similar (Figure 2.6a–b, Figure 2.20a–f): objects 50–60 nm in diameter and 20–30 nm in height, sometimes merged into longer stick-like shapes In confocal microscopy, we observed the formation of globular aggregates that grew over time, reaching several μm in 3 hours and conglomerating into larger structures the next days (Figure 2.6c–d). Thioflavin-T was used to mark β-sheets. SEM imaging of the crude reaction after pre-treatment of the crude cyclization reaction mixture with a few eq. of cysteine showed spherical particles surrounded by fibrous structures (Figure 2.6e), which might be the CPNs. Products of **P2** with MAD did not yield any visible aggregates in the reaction with MAD, even at 5 mM of peptide concentration. In AFM, however, similar structures to that of **P1** were observed (Figure 2.20g–i). Despite the fact, that **CP1** or **MP1M** are positively charged molecules, they are more hydrophobic than the neutral **CP2** or **MP2M**, which can be the reason why **P2**-MAD products do not form visible aggregates at our range of concentrations.

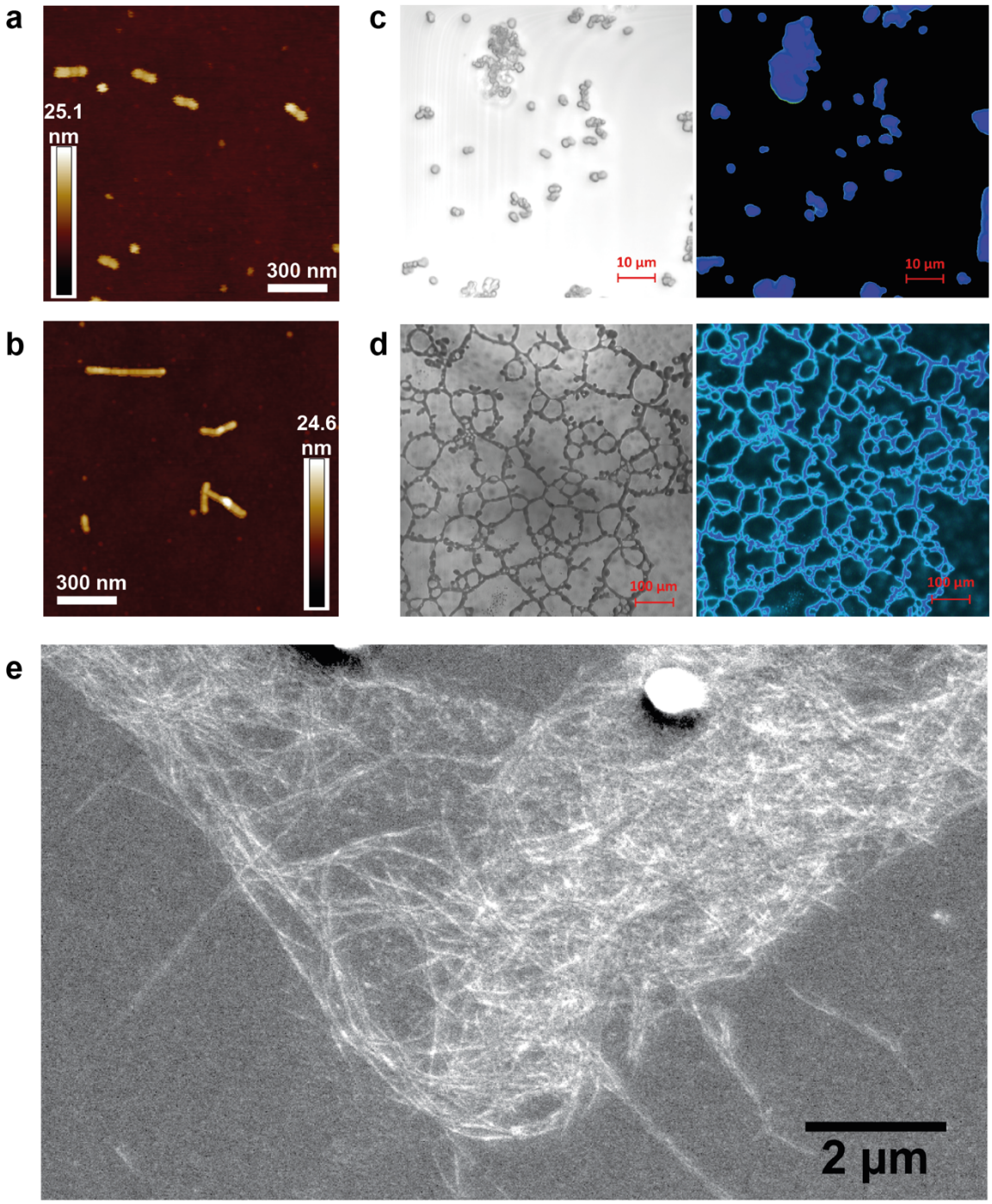


Figure 2.6 | Microscopy images of P1 cyclization reaction products. (a–b) AFM images prepared from 0.1 mM **CP1** solution. (c–d) Confocal microscopy images of **P1** products at 3-4 hours, (c) and 48 hours (d) of reaction with MAD in transmission (left) and fluorescent (right, using Thioflavin-t) regimes. All confocal images correspond to reaction solutions, not dried peptide-MAD products. (e) SEM image.

Once it became clear that peptides react with MAD and form the products that aggregate more than the starting compounds, the next questions arose: what products formed, at what rate, and what ratio of reactants is the best. We quickly realized that the most straightforward technique to track the reaction kinetics is LC-HRMS because it gives the most information about the evolution of different peptide species over time. With CD and UV-Vis spectroscopy, LC-HRMS was used the most in this project.

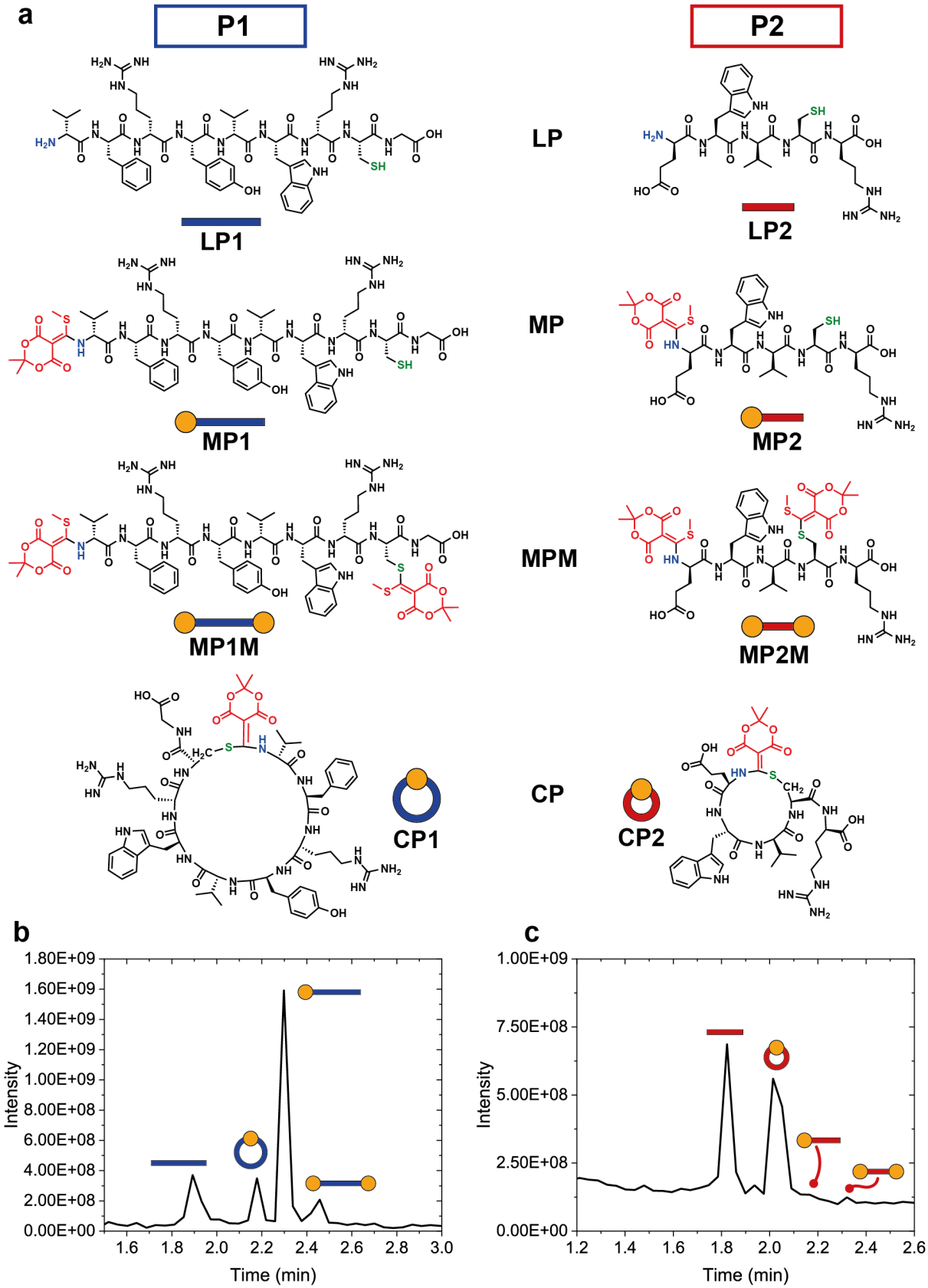


Figure 2.7 | Structures of the peptide species and LC-HRMS chromatogram of cyclization reaction. (a) Structures of **LP** and cyclization products of **P1** and **P2**. (b) LC-HRMS chromatogram of **P1** cyclization. (c) LC-HRMS chromatogram of **P2** cyclization. Both chromatograms are from the measurements taken around the first hour of the reaction.

2.2.3 LC-HRMS studies of cyclization kinetics

According to the mass spectra of the reaction mixture, the main products that formed in all the trials for both **P1** and **P2** were linear MAD-peptide (**MP**), linear MAD-peptide-MAD (**MPM**), and cyclic MAD-peptide (**CP**) products (Figure 2.7). Linear peptide (**LP**) disappeared completely within the first hour of the reaction. Keeping in mind the main goal of getting cyclic peptide nanotubes via reaction with MAD, we then focused on understanding (i) what species form visible aggregates, (ii) how to suppress the formation of **MPM** and favor **CP**, (iii) how to speed up the conversion of **MP** into **CP**, (iv) influence of aggregation on the formation rate of species.

First, we followed the cyclization reaction at high concentrations of peptide (1–1.41 mM) with a slight excess of MAD (1.5 eq.), where aggregation was observed for **P1** but not **P2**. **P1** cyclization yielded 30% **CP1**, 60% **MP1M**, and 10% **MP1**; **P2** cyclization—almost exclusively gave **CP2** (Figure 2.8a–b). As specified by Anslyn and co-workers,[83] cyclization of peptides becomes less efficient if the peptide is too short or too long. Typically, quantitative yields of **CP** were observed for peptides with a length of 4–6 amino acids between reactive sites. However, the cyclization of shorter peptides did not proceed cleanly in their case. Our results are somewhat in contrast to the reported by Anslyn. **P2** has only 2 amino acids between N-terminal amino acid and cysteine, but yielded more **CP2** compared to **CP1**, which has 6 amino acids. This can be explained by the types of peptides the group of Anslyn used: those that have less than 4 amino acids between NH_2 and cysteine also have 2–5 amino acids in the structure after cysteine, possibly causing sterical hindrance and decreasing accessibility of thiol to MAD. Moreover, their reaction conditions involve roughly 5–7 mM peptide, and the solvent mixture does not contain DMF, only ACN as co-solvent to buffer. This suggests that their linear peptides might be partially self-assembled, further influencing the cyclization yield, because assembled linear peptides would not be directly available to MAD.[86,87] In our trials, we used lower peptide concentrations and DMF to improve solubility. Also, both **P1** and **P2** have only one amino acid after cysteine, although the influence of the specific amino acid(s) in that position still needs to be studied.

Despite differences in peptide behavior in the reaction with MAD, the authors[83] pushed all cyclizations to high yields by sparging the reaction mixture with nitrogen to remove the methyl mercaptan byproduct. We also tried this approach to see if it could help favor **CP1** over **MP1** and **MP1M** and speed up the reaction for the **P2** system. Unfortunately, it did not improve the reaction outcome, although it did speed it up (compare panels a to c and b to d in Figure 2.8). Our next step was to decrease peptide concentration to favor cyclization over oligomerization, which can be rationalized using the concept of effective molarity (as we will go into later in this chapter).[88,89]

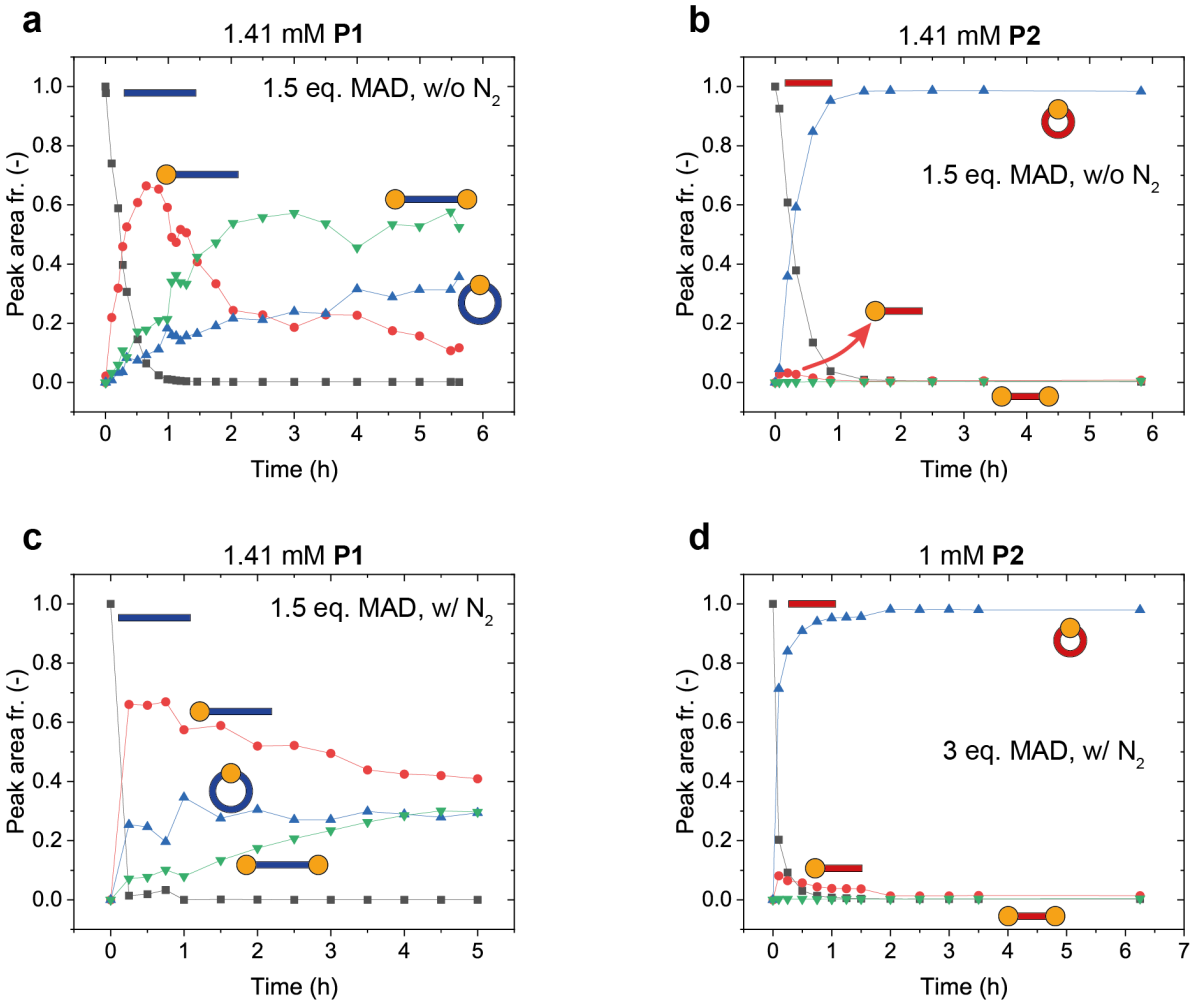


Figure 2.8 | Influence of nitrogen sparging on the kinetics of P1 and P2 cyclization reactions. (a) 1.41 mM **P1** + 1.5 eq. MAD without nitrogen sparging. (b) 1.41 mM **P2** + 1.5 eq. MAD without nitrogen sparging. (c) 1.41 mM **P1** + 1.5 eq. MAD with nitrogen sparging. (d) 1 mM **P2** + 3 eq. MAD with nitrogen sparging. All data were obtained from LC-HRMS studies.

The first notable difference between cyclization at 0.2 mM and 1.41 mM of **P1** was a visual absence of turbidity. The reaction mixture always remained clear and transparent. The product composition changed, improving **CP1** yield to 40% but still having an excess **MP1M** (50%) (Figure 2.9a). For **P2**, the reaction slowed down quite significantly (Figure 2.9b): at 0.1 mM, **LP2** disappeared after 4 h, whereas at 1.41 mM—after 1.5 h (compare Figure 2.8b and Figure 2.9b). The product composition remained the same with the exclusive formation of **CP2**. Interestingly, **MP2** was barely observed at 0.1 mM in contrast to the experiment at 1.41 mM, where the formation and decay of **MP2** were traceable (compare Figure 2.8b and Figure 2.9b). To further influence the product distribution of **P1**, we decreased the amount of MAD from 1.5 to 0.9 eq. This finally led to the quantitative formation of **CP1** over 1 day of the reaction, with already 80% after 6 h (Figure 2.9c). For **P2**, these conditions were decided to be suboptimal, as the reaction was slowed down too much (Figure 2.9d). We also checked the cyclization et 0.5 eq of MAD, but it did not bring any further improvements (Figure 2.21).

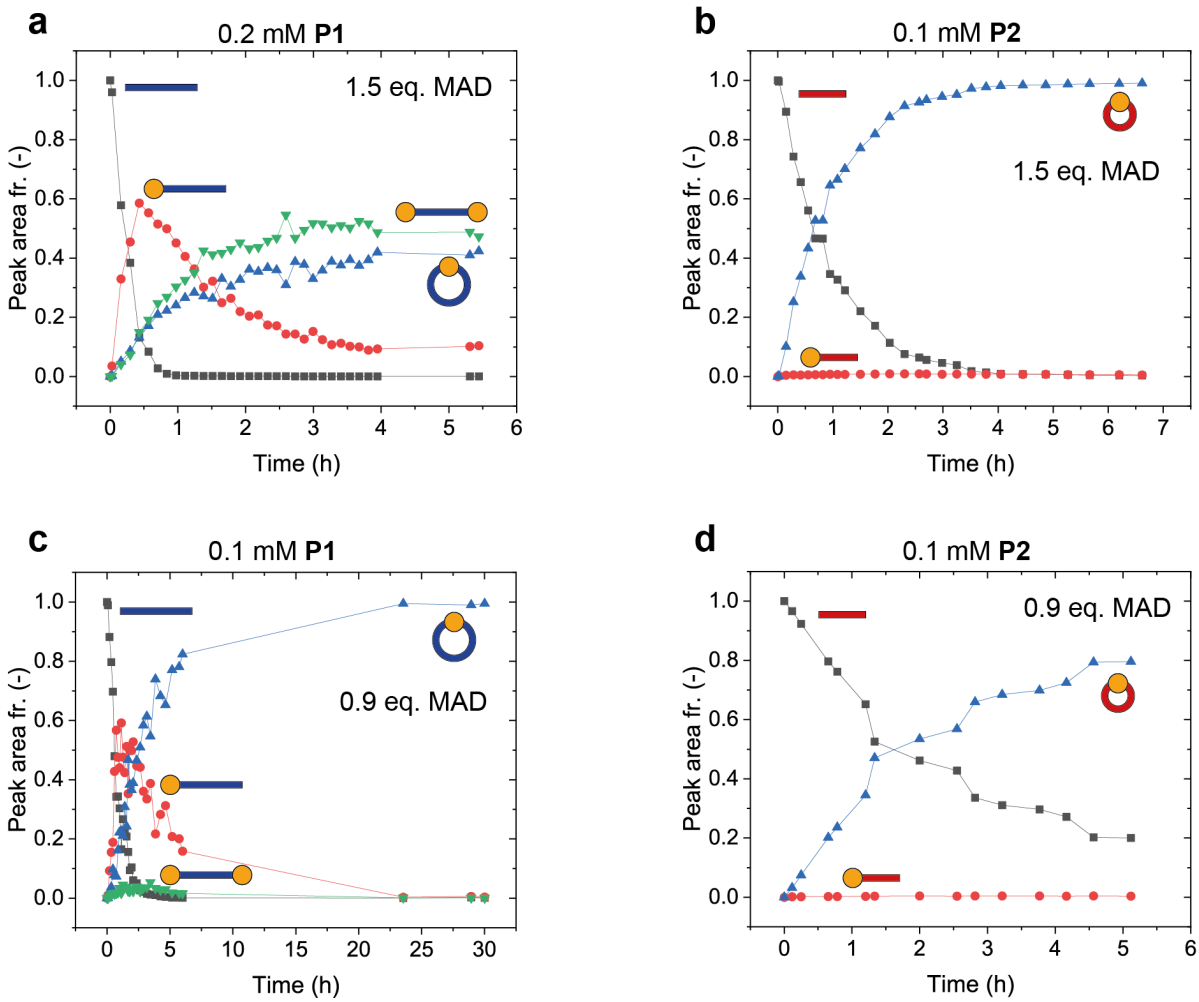


Figure 2.9 | Influence of concentration on the kinetics of P1 and P2 cyclization reactions. (a) 0.2 mM **P1** + 1.5 eq. MAD. (b) 0.1 mM **P2** + 1.5 eq. MAD. (c) 0.1 mM **P1** + 0.9 eq. MAD. (d) 0.1 mM **P2** + 0.9 eq. MAD. All data were obtained from LC-HRMS studies. Experiments are done without nitrogen sparging.

When MAD is in excess (10–50 eq.) **P1** forms predominantly **MP1M**, but **P2** starts forming noticeable amounts of **MP2M** (10%) only at 50 eq. of MAD (Figure 2.10). This is useful for later experiments when cyclization and decyclization reactions are combined into a reaction cycle, as it gives an additional lever to control the transient lifetime of **CP** without producing more oligomeric **MPM** (in chapter 4).

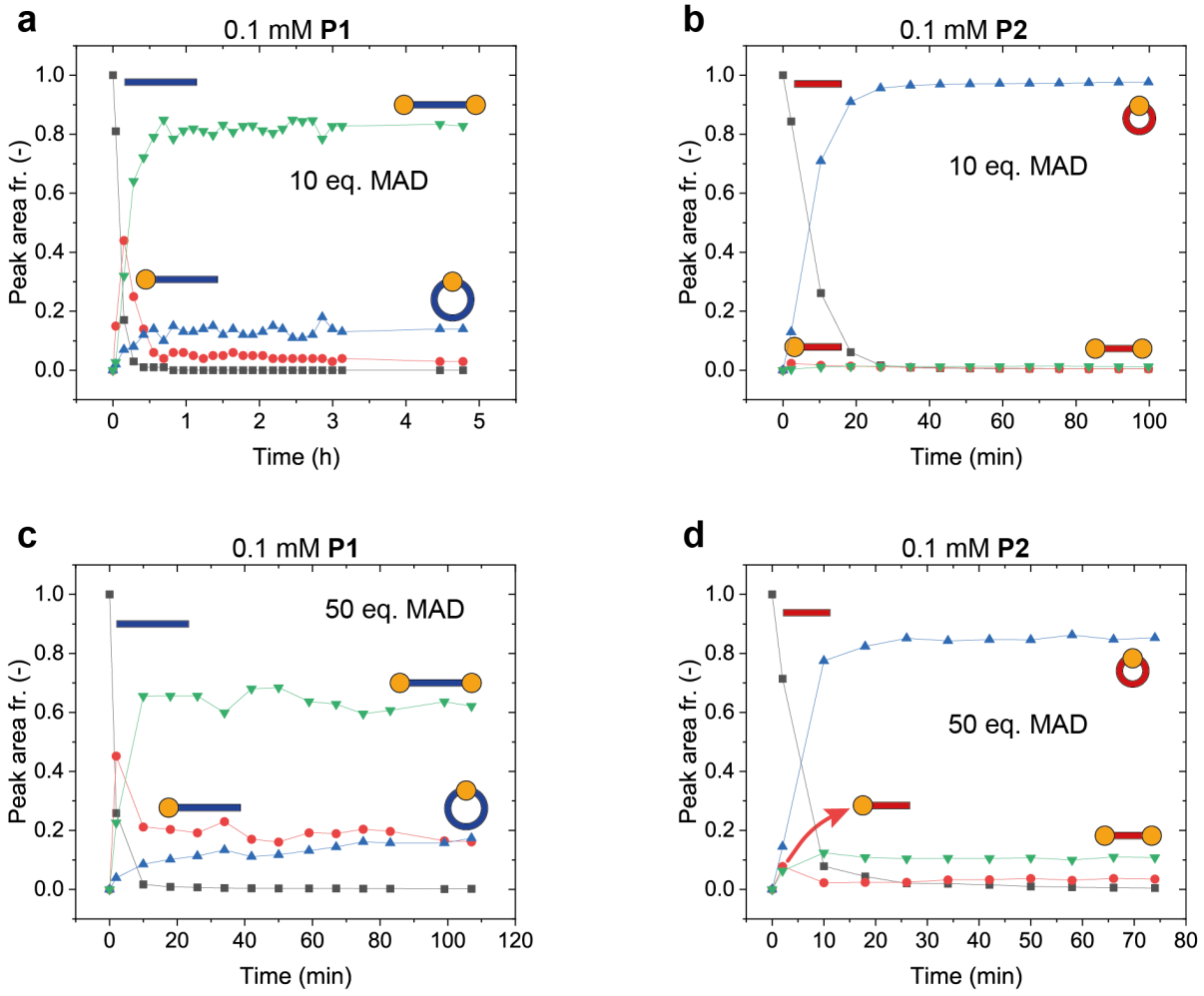


Figure 2.10 | Cyclization reaction of P1 and P2 with an excess of MAD. (a) 0.1 mM **P1** + 10 eq. MAD. (b) 0.1 mM **P2** + 10 eq. MAD. (c) 0.1 mM **P1** + 50 eq. MAD. (d) 0.1 mM **P2** + 50 eq. MAD. All data were obtained from LC-HRMS studies.

So far, we have mentioned only one type of oligomeric species, **MPM**. However, throughout the experiments, we observed the formation of longer chains both for **P1** and **P2**, e.g. linear **(MP)$_2$M**, **(MP)$_3$**, and **(MP)$_4$**, whose amount usually did not exceed 1–2%. Cyclic oligomeric species—CPs containing a few units of peptide and MAD—were also found, and their amounts, determined from their relative abundance of each species in LC-HRMS, varied depending on the MAD amount (Table 2.1). Particularly, the largest cyclic product for **P1** was found in the experiment with 0.9 eq MAD and contained 5 peptide-MAD monomers (i.e., **(CP1)$_5$**), whereas **P2** formed only a 2-meric ring with some traces of 3-mers apart from the main monocyclic product. Experiments with nitrogen sparging yielded almost exclusively monocyclic **CP** for both peptides. When methyl mercaptan is efficiently removed from the solution, it is withdrawn from the reaction equilibrium, which blocks the reshuffling of thiols at the MAD reaction center and, hence, decreases the probability of macrocycle formation. One could also see the reaction's sensitivity to the equilibrium by comparing MAD consumption in the reactions of **P1** and **P1-Met** with MAD, which can be tracked by UV-Vis. **P1-Met** is the peptide with the same amino acid sequence as **P1** but with methionine instead of cysteine (Figure 2.17b). Judging by the decrease of the MAD absorbance peak at 350 nm, it is possible to conclude that **P1** consumes more MAD than **P1-Met** (Figure 2.12b and Figure 2.22c). This

is to be expected, as **P1** in reaction with MAD forms **MP1**, which goes to **CP1** and **MP1M**, which in turn can self-assemble, creating a sink for the reaction and shifting the equilibrium to the products, whereas for **P1-Met**, the only possible product is **MP1-Met**. Finding a way to increase MAD selectivity towards cysteine thiol instead of methyl mercaptan would help to narrow CP product dispersity and increase the cyclization rate.

Table 2.1 | Composition of the cyclic product for P1 and P2 systems. Depending on the reaction conditions, the cyclic product included polycycles with a different number of peptide-MAD units (n).

P1, mM	MAD, eq.	Cyclic **(CP1)$_n$** product composition in %, n = number of peptide-MAD units in a cycle				
		1	**2**	**3**	**4**	**5**
1.41	1.5	64	26	8	2	–
1.41*	1.5	97	3	–	–	–
0.2	1.5	62	28	8	2	–
0.1	0.9	47	39	11	2	1
0.1	0.5	60	30	10	traces	–
0.1	50	69	25	6	–	–

P2, mM	MAD, eq.	Cyclic **(CP2)$_n$** product composition in %, n = number of peptide-MAD units in a cycle		
		1	**2**	**3**
1.41	1.5	81	19	traces
1.0*	3	99	1	traces
0.1	1.5	87	13	traces
0.1	0.9	86	14	traces
0.1	0.5	86	14	–
0.1	10	93	7	–
0.1	50	87	13	traces

*—experiments with N_2 sparging.

A useful concept to compare **P1** and **P2** and their differences in **CP** yields is effective molarity. Effective molarity (EM) measures how efficient the intramolecular reaction is compared to the intermolecular.[88,89] This concept is particularly important for understanding the catalytic activity of enzymes and the description of host-guest binding.[90] Effective molarity can be expressed as the ratio between the rate constants of intra- and intermolecular reactions (kinetically controlled reactions); as the ratio between the corresponding equilibrium constants (thermodynamically controlled reactions); as the ratio between equilibrium concentrations of cyclic and chain products (for reactions with low EM).[89,90]

$$EM = \frac{k_{\text{intra}}}{k_{\text{inter}}} = \frac{K_{\text{intra}}}{K_{\text{inter}}} = \frac{[\mathbf{CP}][\mathbf{MAD}]}{[\mathbf{MPM}]} \left(\frac{\text{mol}}{\text{L}}\right) \quad (2.1)$$

Knowing EM for cyclization reactions allows for predicting (and maximizing) the yield of cyclic products when competition with chain products is involved. EM in the example of the

peptide-MAD system can be imagined as the concentration of thiol required to make the **MP→MPM** reaction go at the observed rate of **MP→CP** reaction.

Effective molarity for cyclic peptides has been studied.[91] For pentapeptides, the formation of cyclic peptides through the new amide bond was compared to the growth of linear chains in the competitive reaction with another amino acid. The authors[91] conclude that for equal peptide lengths, the more flexible the linear peptide, the higher the yields of a cyclic peptide, and the higher the EM. They calculated EM values from the experimentally observed equilibrium product ratios and obtained values of 7–18 mM.[91] Adopting the same calculation for our system, see Equation (2.1), we obtain EM 0.10 ± 0.02 mM for **P1** and 78 ± 2 mM for **P2**. In this calculation method, all cyclic products (even those that are macrocyclic, $\mathbf{(CP)_n}$) were put in the numerator of Equation (2.1). However, effectively, $\mathbf{(CP)_n}$ should be counted as oligomers,[90] i.e., together with **MPM** species in the denominator. Accounting for the macrocycles $\mathbf{(CP)_n}$ formation as the increase of the oligomeric product (Table 2.1), EM becomes 0.045±0.009 mM for **P1** and 64±2 mM for **P2**. From these values, we can conclude that to favor the quantitative formation of **CP1** (or cyclic $\mathbf{(CP1)_1}$), the reaction should be done at 0.045 mM **P1** or lower, as according to Hunter and co-workers,[88] the formation of cyclic product is favored over a concentration window $\frac{1}{K^2 EM} < 2[\mathbf{MAD}] < EM$, where K is the association constant. To achieve the same for **P2,** we are already at the right concentration. At concentrations of MAD that are $2[\mathbf{MAD}] > EM$, formation of **MPM** is favored in the cyclization process. Indeed, in most of the **P1** cyclization experiments $2[MAD]$ is larger than EM of **P1,** which explains why the formation of **MP1M** is observed. For the **P2** system, this becomes the case only in the experiment 0.1 mM **P2** + 50 eq. MAD (Figure 2.10d), and as expected we observe the formation of **MP2M**.

2.2.4 Numerical methods for the cyclization reaction

Detailed kinetics analysis of a specific reaction typically involves the study of concentration versus time dependencies based on the certain rate expression(s). According to HRMS data, the cyclization reaction of **LP** with MAD proceeds via an intermediate **MP**, which then can transform into **CP** or react with another eq. of MAD to form **MPM**. This conversion can be summarized in the model:

$$\mathbf{LP} + \mathrm{MAD} \xrightarrow{k_1} \mathbf{MP} \xrightarrow{k_2} \mathbf{CP}$$

$$\mathbf{MP} + \mathrm{MAD} \xrightarrow{k_3} \mathbf{MPM}$$

Expressing the model via differential equations of mass balance:

$$\frac{d[\mathbf{LP}]}{dt} = -k_1[\mathbf{LP}][\mathrm{MAD}] \tag{2.2}$$

$$\frac{d[\mathrm{MAD}]}{dt} = -k_1[\mathbf{LP}][\mathrm{MAD}] - k_3[\mathbf{MP}][\mathrm{MAD}] \tag{2.3}$$

$$\frac{d[\mathbf{MP}]}{dt} = k_1[\mathbf{LP}][\mathrm{MAD}] - k_2[\mathbf{MP}] - k_3[\mathbf{MP}][\mathrm{MAD}] \tag{2.4}$$

$$\frac{d[\mathbf{CP}]}{dt} = k_2[\mathbf{MP}] \tag{2.5}$$

$$\frac{d[\mathbf{MPM}]}{dt} = k_3[\mathbf{MP}][\mathrm{MAD}] \tag{2.6}$$

To solve the system of (first order) ordinary differential equations (ODEs), scientists often prefer numerical integration to avoid massive derivations of analytical solutions. This approach is widely used to estimate rate constants of processes in the reaction network.[92–96] The set of ODEs for cyclization was solved in a Python environment to obtain the temporal evolution of peptide species fitted to the experimental data and the corresponding rate constants. All LC-HRMS data of cyclization reaction were subjected to the fitting algorithm (Figure 2.11), and all obtained rate constants were averaged and expressed with standard deviations (
Table 2.2).

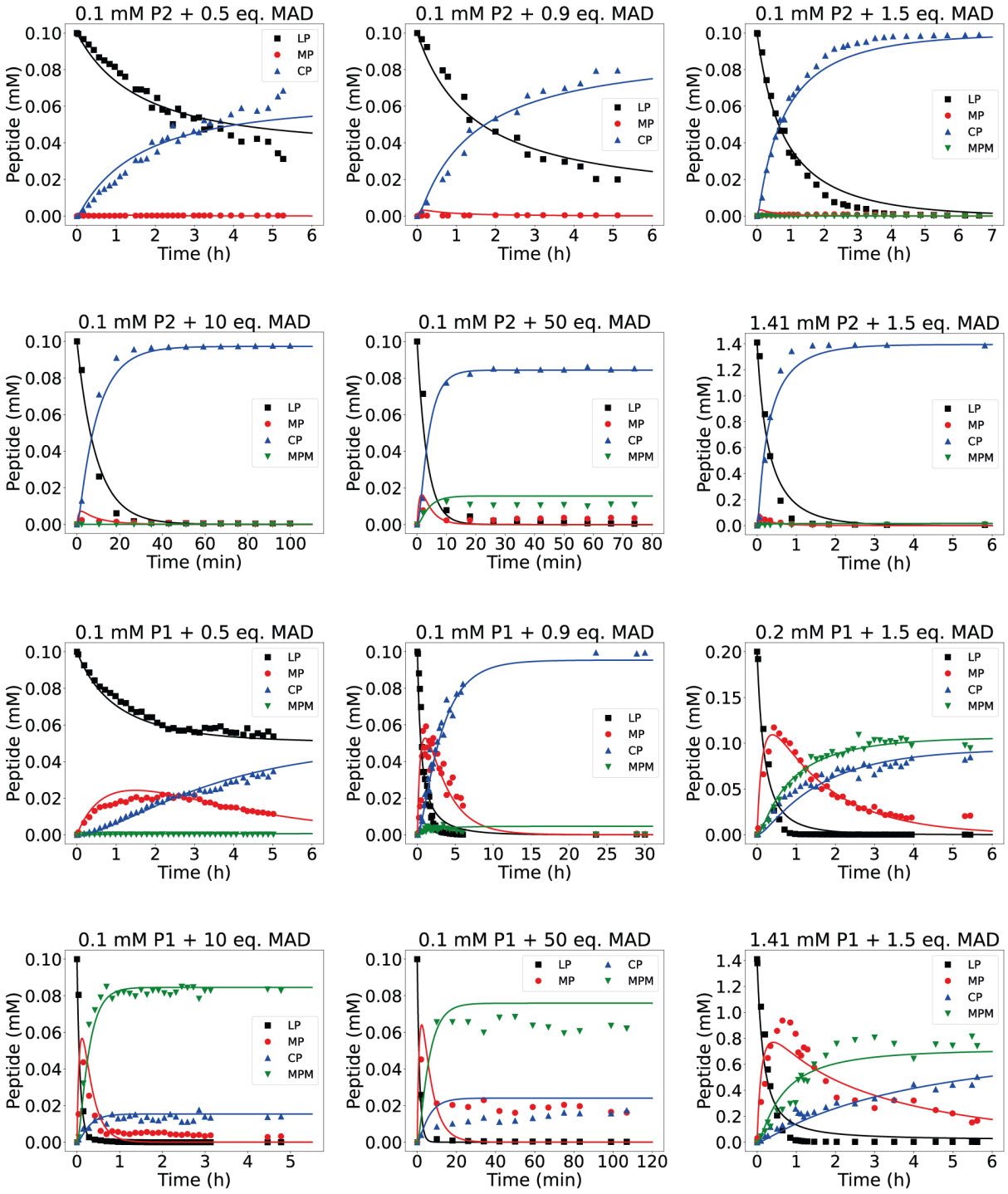


Figure 2.11 | The numerical fit of the LC-HRMS data based on ODEs of the cyclization model.

Table 2.2 | The ODEs model results. Mean fitted rate constants for all cyclization LC-HRMS experimental data and the standard deviations.

Peptide	k_1, $M^{-1}s^{-1}$	k_2, s^{-1}	k_3, $M^{-1}s^{-1}$
P1	3 ± 1	$(0.2 \pm 0.2) \times 10^{-3}$	0.6 ± 0.4
P2	1.6 ± 0.7	0.012 ± 0.008	0.5 ± 2

As one can see, the model does not always perfectly align with the experimental data, which may suggest that the model is oversimplified. From the data **P1** + 0.9 eq. MAD, one can see that **MP1M** decreases after 5 hours of the reaction. From the chemistry point of view, this can only happen if there is free thiol (e.g., methyl mercaptan) in the system that exchanges with the cysteine thiol of **MP1M**, liberating for the short moment **MP1**, which cyclizes into **CP1**. The model, however, does not account for the further possible conversion of **MP1M**. Furthermore, the self-assembly of any species is also not considered, neither the rate at which methyl mercaptan leaves the system nor the formation of poly-rings and poly-chains. Another problem with the fitting of some of the experiments is the lack of data points at the early stages of the reaction. This leads to poor model performance and huge differences in calculated rate constants from experiment to experiment. It is especially apparent in the experiments with the excess of MAD. Finally, experimental/measurement errors also affect the resulting fit. For instance, in the experiment 0.1 mM **P1** + 50 eq. MAD, **MP1** is not fully consumed at the later reaction times, which does not align with other experiments and with the theoretical hypothesis. Therefore, to obtain more precise values of rate constants, it is best to fit the averaged data from duplicates or triplicates of kinetic runs. Further details on the script are described in Appendix Section 2.6.4.2.

To further increase the goodness of fit, it might be worth expanding a set of ODEs by adding expressions that include methyl mercaptan and self-assembly of either of the peptides' species, as the latter may slow down the reaction. Moreover, although the waste product of the forward reaction leaves the system, it might be still useful to consider that the attack of amine on MAD is reversible, and the equilibrium is probably involved in each reaction step.

However, achieving a slightly better result in fitting should not be a goal in itself or bring an erroneous feeling of comfort, as the model, in the end, represents *only* our assumptions, but not necessarily what objectively happens.[97] The biggest use of this and other models in this thesis is to develop an intuition about the system and estimate its parameters. Compared to the sparse data on MAD reaction kinetics with other substrates,[78,98] the obtained rate constants are reasonable.

2.2.5 UV-Vis studies of cyclization kinetics

UV-Vis spectroscopy is a second suitable method to track the reaction of MAD with nucleophiles,[78,82,98] because MAD exhibits a strong absorption band at 350 nm, which differs from that of MAD-Nucleophile product. On the course of the reaction peak of MAD decreases, and a peak of the product appears, usually in the lower wavelength region. However, if one aims to elucidate reaction kinetics, the convenience of this method is reversely proportional to the structural complexity of the nucleophile. For example, following the reaction of MAD with

β-mercaptoethanol (BME) is straightforward,[78] as bands of reactants and products do not overlap, and there is one clear isosbestic point (Figure 2.12a). Plotting absorbance of MAD over time from different experiments gives access to the reaction order and rate constants. The reaction of MAD with DTT is already more complex[98] than that with BME as it proceeds via several steps, and the product is pH sensitive, which can cause a further change in the absorbance profile of the reaction mixture. The isosbestic point, in this case, is absent. Information extraction from such spectra already requires additional modeling techniques.[98] In our system, the typical absorbance profile of the reaction course of peptides with MAD, at first sight, looks like that reported for small molecules[82] (Figure 2.12b, Figure 2.13a–b).

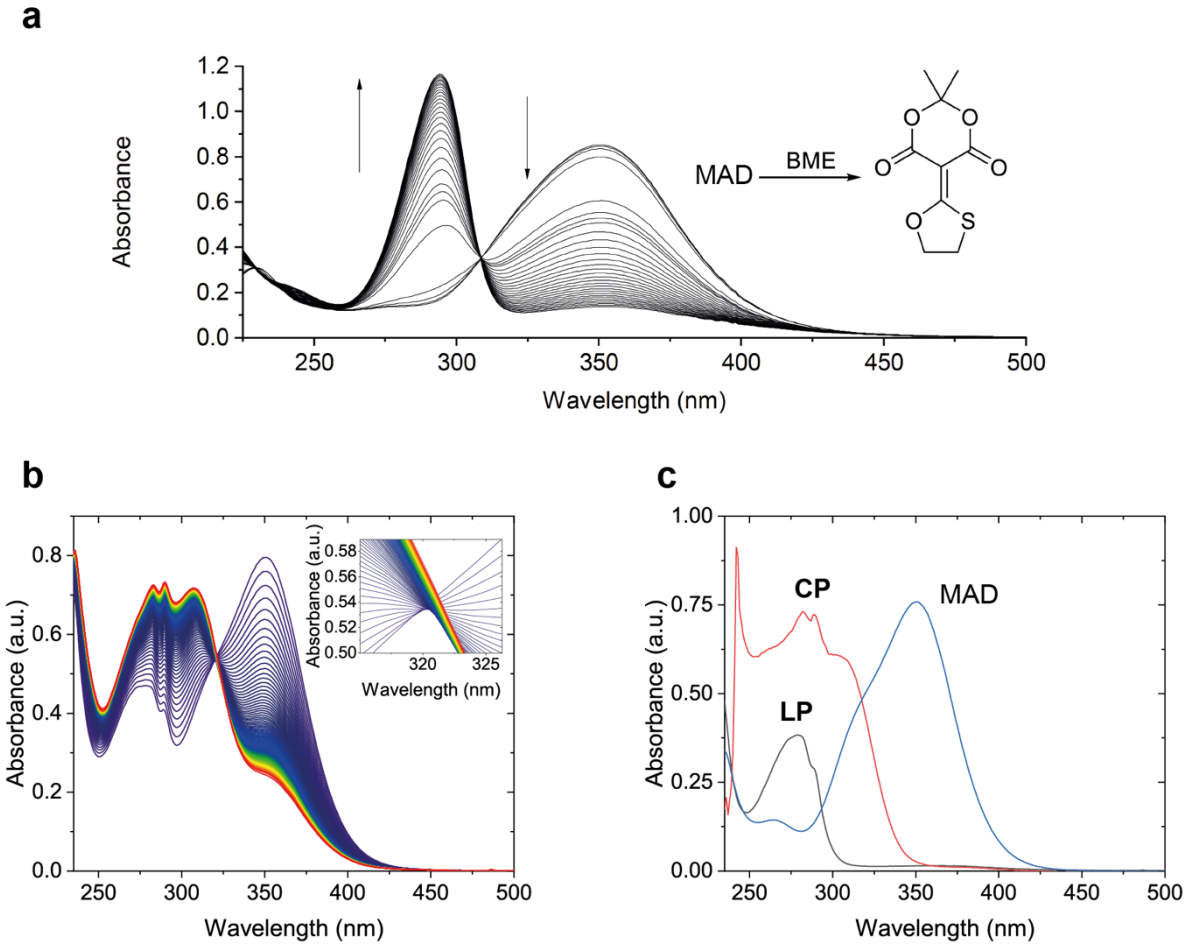


Figure 2.12 | Absorbance profile of MAD reaction. (a) with BME, adapted from ref. 78. (b) with **P1** (0.7 mM) and 1 eq. MAD; a zoomed absorbance region at 320–325 nm in the inset, representing shifted isosbestic point. (c) Absorption spectra of **LP1**, **CP1**, and MAD.

In contrast to these examples, several factors complicate the analysis of the cyclization kinetics via UV-Vis:

1) Overlap of absorption bands of all peptide species

P1 has three UV-active amino acid residues: tryptophan, tyrosine, and phenylalanine. They create a broad absorption band at 250–290 nm. An absorption of pure **P2** is similar (Figure 2.22a). The peptide chromophores remain unchanged during the reaction, and as a result, peptide intermediate and products have very similar absorption spectra (Figure 2.12c).

2) Stepwise mechanism of peptide reaction with MAD

According to the literature[83] and HRMS experiments, the addition of NH_2 to MAD occurs faster than of cysteine thiol, which causes the transient presence of **MP** (e.g., in Figure 2.9). Although in **P2** cyclization, conversion of **MP2** to **CP2** is fast enough, and **MP2** is rarely present in amounts more than 5%, for the **P1** system, lifetime and amounts of **MP1** are significant. The latter complicates further distinguishing reaction stages from UV-Vis data.

3) The absorbance peak of MAD (350 nm) overlaps with **MP1M**

In **MP1M**, MAD attached to the peptide via thiol will not have significantly different absorbance from MAD itself. We tested this hypothesis by tracking the reaction of N-acetylated **P1** with MAD (Figure 2.22b). In fact, for **P1** cyclization, the MAD peak never decreases fully to the baseline if ≥ 1 eq. of MAD is used (Figure 2.12b), and from HRMS, we know that 10–60% **MP1M** can form in these conditions. Decreasing MAD to 0.5 eq results in the full disappearance of the MAD peak (Figure 2.13a); correspondingly, no **MP1M** forms (Figure 2.21). For **P2** reaction with MAD, the MAD peak disappears fully (Figure 2.13b), which correlates with the absence of **MP2M** at any point of the reaction (Figure 2.9b–d).

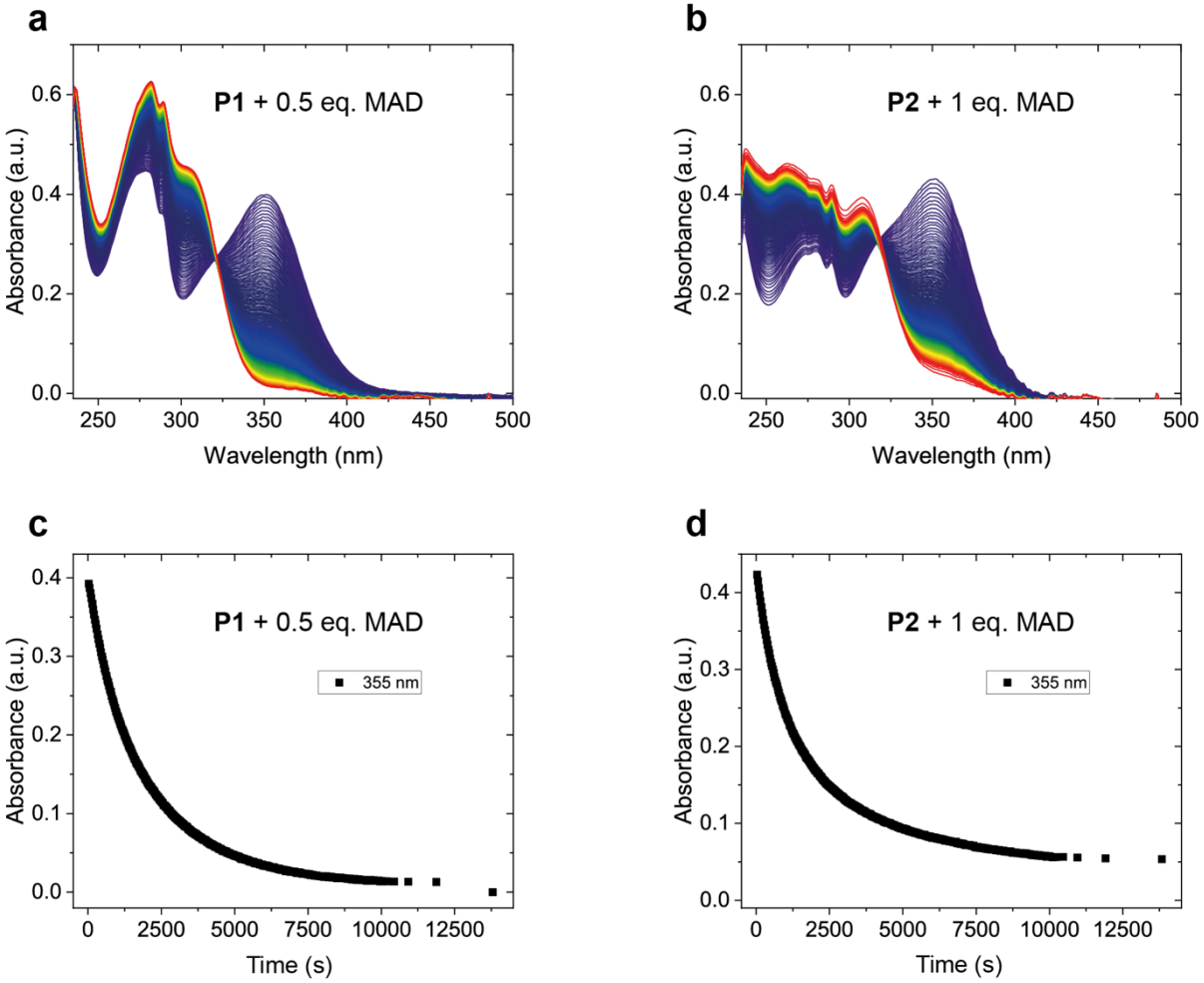


Figure 2.13 | Absorbance of MAD reaction with peptides. (a) In the reaction of 0.7 mM **P1** with 0.5 eq MAD, absorbance at 355 nm decreases almost fully to the baseline, which also happens in the reaction with **P2** (b). Tracking absorbance over time in these cases (c, d) gave insight into the rate of MAD consumption (see the main text and Appendix).

4) No clear isosbestic point

An isosbestic point is a wavelength at which absorbance remains constant during the process as molar absorption coefficients of reactants and products are equal at this wavelength.[92,99] The

absence of the isosbestic point can indicate that several processes are going on simultaneously, resulting in a non-linear change of concentrations of the reaction components that absorb in this spectral region.

Both for **P1** and **P2** reactions (Figure 2.12b inset, Figure 2.13b) with MAD, the isosbestic point at 321 nm starts shifting to the right after several min of reaction and becomes blurred. Since it occurs for both systems, it is reasonable to suppose that not just a complex reaction mechanism plays a role but also another common process for **CP** of both peptides—self-assembly. As shown later, depending on the solvent composition, **CP** is fully or partially assembled, which changes its absorbance spectrum (mainly relative intensities of peaks at 280 and 305 nm) (see Section 2.2.6). As the fraction of **CP** during the cyclization reaction increases, self-assembly becomes more dominant and can impact the overall spectrum of the reaction mixture.

Despite the difficulty of addressing these issues in one universal treatment method of UV-Vis spectra of cyclization reaction, we further demonstrate some approaches to get useful information from this data. The ultimate goal of finding a proper treatment of UV-Vis data was to obtain the rate characteristics of MAD consumption, which was tedious to follow by LC-HRMS, and to compare rates of **CP** and **LP** with that obtained from mass spectrometry kinetics. Because in the **P2** system, no **MPM** forms at low eq. of MAD, it was used the most for the initial evaluation of UV-Vis data. From the time-course measurement of 0.35 mM **P2** with 1 eq. MAD (Figure 2.13d), it was possible to estimate the apparent half-life of the reaction $t_{1/2} = 13$ min. The method of fractional-life periods[92] was applied here to determine the reaction order (see Appendix Section 2.6.5.2). However, values obtained from this method varied from 1.09 to 1.17, suggesting that a more complex expression for the reaction rate is needed. The same result was obtained for **P1** (Figure 2.13c; $t_{1/2} = 19$ min). Another way to estimate the reaction order of the simple stoichiometric reaction with equal initial concentrations is to plot the left part of Equation (2.7) over time and, by varying *n*, determine when the function is linear.[92] The slope of this graph would be equal to the rate constant (see Appendix Section 2.6.5.3).

$$\left(\frac{1}{n-1}\right)\left(\frac{1}{c^{n-1}} - \frac{1}{c_0^{n-1}}\right) = kt \qquad (2.7)$$

The reaction order obtained by this method was also fractional, 2.2, confirming the need for better rate expression (Figure 2.23). Since one may expect the reaction to be 2nd order, several methods of treating the UV-Vis data for second order reactions were also tried both for **P1** and **P2**. As expected, they did not give linear dependence that would allow extracting rate constants in a reproducible and unambiguous way. These methods are discussed in Appendix Section 2.6.5.4.

At that point, it became clear that using the decrease of absorbance of the MAD peak would not be useful for fitting the individual reaction rate of MAD, and we needed more data on how the change of the absorbance profile of the reaction mixture reflects the dynamics of separate components. For this, the principle of spectral decomposition was used. The absorption spectrum of a complex mixture is the linear combination of the individual component spectra:[92]

$$A^i = x_1 X_1^i + x_2 X_2^i + \cdots + x_n X_n^i \qquad (2.8)$$

or

$$\mathbf{A} = \mathbf{xX} \tag{2.9}$$

where n is a number of components, i is the wavelength, A^i is the absorbance of mixture at i-th wavelength, X^i_n is the absorbance of an individual component n at i-th wavelength, and **A**, **x**, **X** are corresponding matrices. Knowing individual spectra of all the components allows resolving the spectrum of the mixture by solving Equation (2.8) for all coefficients x_n and relating the latter to the concentrations of each component in the mixture.

Practically, the solution for **x** can be found using Solver of MS Excel. From an initial guess of the coefficients, **y**, a calculated spectrum of the reaction mixture, **B**, can be built. The solver optimizes **y** to approach **x** using the least-square method for decreasing the difference between **A** and **B** by minimizing the sum of squares of errors. Going through this procedure is relatively fast if a few spectra of a complex mixture need to be analyzed. However, the cyclization reaction can take up to 6 to 8 hours to equilibrate, which with frequent measurements, easily results in 300–500 absorption spectra. To simplify and speed up the analysis of this data, we used a Python script that runs the least-square optimization sequentially for each spectrum in batch and saves the final **y** in a separate file. This set of coefficients **y** is then related to the corresponding reaction times, after which the normalization relative to starting and final component concentrations is done (Appendix Section 2.6.5.5).

We used this method for the **P2** system with the lack of MAD or its moderate excess relative to the peptide because in these cases, we do not need to account for **MP2** or **MP2M**, which are difficult to isolate. From the resulting graph, for example, for 0.1 mM **P2** + 1.5 eq. of MAD (Figure 2.14a), it is possible to extract half-lives for each component and compare them to the corresponding data from LC-HRMS. For **LP2** and **CP2,** half-lives from UV-Vis are 24–26 min, close to HRMS—roughly 36 min (Figure 2.9b). The kinetics profile of MAD is consistent with that of **LP2**, as expected. A similar result gave the decomposition of the experiment 0.35 mM **P2** + 1 eq. MAD (Figure 2.22d). For the **P1** system, the method was used in experiments with the lack of MAD, where no **MP1M** is formed, but **MP1** was not considered during spectra decomposition. Nevertheless, the resulting half-live values obtained from UV-Vis treatment were comparable with that from LC-HRMS kinetics data. For the experiment 0.7 mM **P1** + 0.5 eq. MAD, the half-life of MAD and **P1** was 23 min (Figure 2.14b), and the **LP1** consumption started to plateau after 3.33 h, the same as in the HRMS experiment (Figure 2.21a). For a more precise study of peptide species, spectra of **MP1** and **MP1M** should be used as a part of the input data, although to estimate MAD consumption, the approach is already beneficial.

Overall, implementation of this method in not always convenient, as obtaining the absorption spectra of each of the individual components is not always feasible and, in most cases is a time-consuming task. Additionally, very often, the rate law of the reaction (or kinetics profiles of the components) is unknown. According to literature references, it is still possible, however, to elucidate the number of components, the components' individual spectra, and evolution of the concentration over time from a batch of reaction spectra. Such algorithms provide full spectra deconvolution based on matrix algebra and statistics. They include, for instance, Single Value Decomposition, Target Factor Analysis, Evolving Factor Analysis, Rank Annihilation Factor analysis, and Generalized Rank Annihilation Method.[100–104] The use of one of these methods could be considered in future studies CP-MAD reaction network.

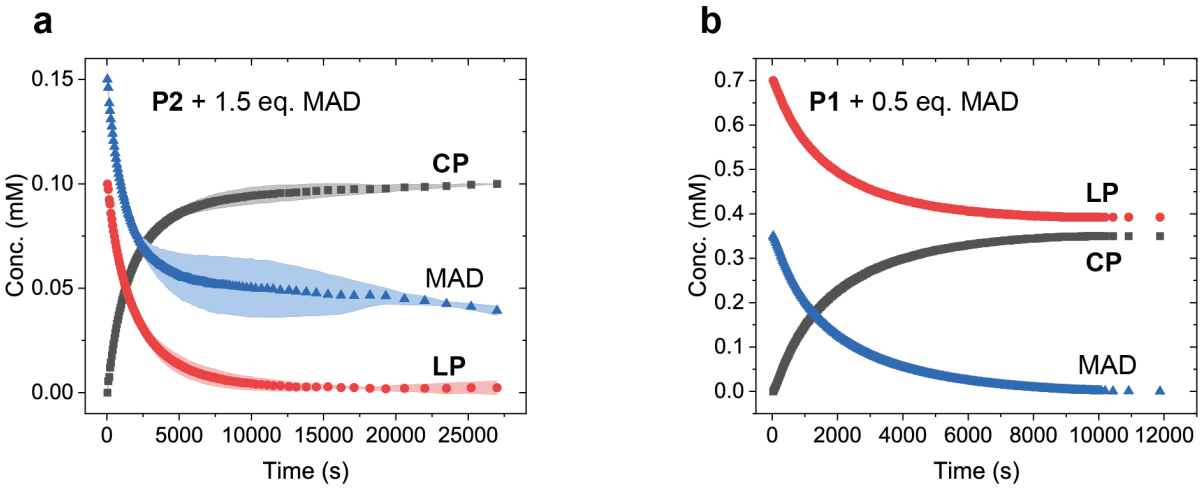


Figure 2.14 | Concentration profiles of CP, LP, and MAD obtained by UV-Vis spectra decomposition algorithm. (a) 0.1 mM **P2** + 1.5 eq. MAD measurement was done in triplicates; the shaded area represents the standard deviation. (b) 0.7 mM **P1** + 0.5 eq MAD measurement.

2.2.6 Self-assembly study by UV-Vis

We performed time-dependent UV-Vis measurements to follow an increase of optical density (O.D., which is due to scattering) over time at 400–600 nm, hoping to elucidate the kinetics of MAD-**P1** product formation and self-assembly as previously done in our group.[105] However, after an initial increase indicating an increasing amount of self-assembled species, there was a rapid decrease (Figure 2.15a). We found that the formed species are large enough to sediment and therefore be removed from the light path. We compared the solution composition of the identical HRMS experiment (Figure 2.8a) at the point when scattering in the UV-Vis experiment was increased (after 1 h of the reaction): 19% **CP**, 21% **MPM**, and 60% **MP**. The fast decrease of scattering was observed after 3 h of the reaction, which corresponded to 22% **CP**, 59% **MPM**, and 19% **MP**. Consequently, in any case, an attempt to derive the self-assembly mechanism of cyclic peptides in-situ from this method would not lead to a correct interpretation because, at any time of the reaction, there is a mixture of different MAD-peptide products. Also, for **P2**, the method is not suitable because the products remain soluble in the reaction mixture at least up to 5 mM.

Therefore, our next step was to synthesize and purify **CP1** and **CP2** (Appendix Section 2.6.3) and then study their supramolecular polymerization mechanisms. There are two main mechanisms to describe supramolecular polymerization: isodesmic and cooperative.[106] In the isodesmic, the process is described by one constant of the addition of the monomer to the polymeric chain regardless of the length of this chain. In the cooperative (nucleation-elongation), thermodynamically unfavored nucleation must occur before the more favored elongation process, which has its corresponding constants. Both mechanisms are easily distinguished by representing an intrinsic property of the molecule against concentration, temperature, or molar fraction of solvent. The mechanisms of self-assembly are usually derived either from temperature-dependent or solvent-denaturation measurements, which both explore the extent of supramolecular polymerization between fully assembled and fully disassembled states. From the plotted dependencies of these states over temperature or fraction of a good solvent, respectively, it can be obtained constants and energy parameters corresponding to the

polymerization process. A sigmoidal curve is characteristic of an isodesmic system, whereas a zero-slope region followed by a steep curve is assigned to a cooperative mechanism.

The temperature-dependent studies did not give us good results: the absorbance spectrum of **CP** before and after the heating-cooling cycle did not overlap, from which we concluded that **CP** is not temperature-stable. We proceeded with solvent-denaturation experiments.[107] In this model, the extent of supramolecular polymerization is modified by mixing a good solvent, where the monomer is fully dissociated, with a bad solvent, where self-assembly is favored and a monomer is fully assembled. It is assumed that the self-assembly process consists of sequential addition of monomers to the polymeric chain with a K_e. In the case of a cooperative mechanism, a K_n responsible for nucleation is also considered. Both constants are related by the cooperativity factor $\sigma = K_n/K_e$. If the value of $\sigma < 1$, it is a cooperative mechanism; if $\sigma = 1$, it is an isodesmic mechanism. K_n is related to the free Gibbs energy ($\Delta G^{0\prime}$) associated with the addition of the monomer to the chain by Equation (2.10):

$$K_n = \exp\left(\frac{\Delta G^{0\prime}}{RT}\right) \tag{2.10}$$

The model also takes into account the linear dependency of the free energy with the molar fraction of solvent *f*. Equation (2.11) describes this relationship, where ΔG^0 is the free energy that considers the addition of the monomer to the bad solvent, and the *m* parameter is a measure of the ability of the good solvent to associate with the monomer, thereby disrupting the supramolecular polymer.

$$\Delta G^{0\prime} = \Delta G^0 + mf \tag{2.11}$$

Thus, a solution of **CP2** in a good solvent (DMSO) was mixed with a solution of **CP2** in a bad solvent (phosphate buffer, pH 7.4) by keeping constant concentration. In Figure 2.15b, the degree of aggregation α (defined by the normalized ratio of the absorbances at 280 and 305 nm) is represented against a molar fraction *f* of DMSO for three different concentrations. The sigmoidal curves obtained are characteristic of the isodesmic mechanism. The thermodynamic values obtained after fitting to the solvent-denaturation model[107] (Table 2.3) confirm this conclusion.

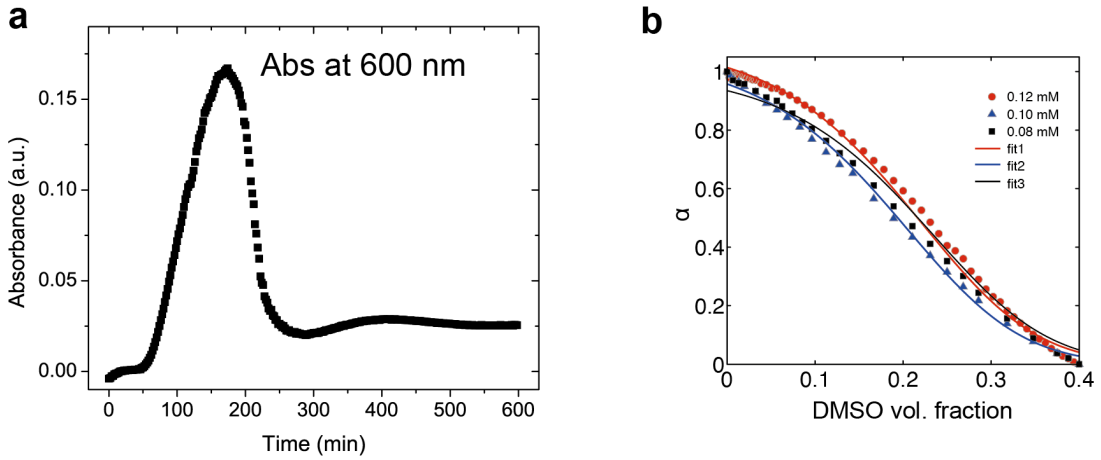


Figure 2.15 | Aggregation studies of P1 and P2 cyclization products. (a) Absorbance at 600 nm tracked over time by UV-Vis; 1.41 mM **P1** + 1.5 eq. MAD. A decrease in scattering correlates with sinking aggregates. (b) Plot of the degree of aggregation α of **CP2** versus the volume fraction of DMSO at different concentrations. The lines show the fitting to the solvent-denaturation model.

Table 2.3 | Thermodynamic data for the supramolecular polymerization of CP2 derived by the solvent-denaturation model.

Thermodynamic parameter	Value
ΔG^0	-30.67 ± 1.53 kJ mol^{-1}
m	32.2 ± 5.1
σ	1
K_e	293696.3 M^{-1}
K_n	293696.3 M^{-1}

2.3 CD studies of the supramolecular aggregates

It was particularly interesting to study **CP** and its formation by circular dichroism spectrometry (CD). Once we confirmed that the **CP** of **P1** and **P2** are partially self-assembled in our solvent mixture, the question became whether the rings are stacked with all residues being directly on top of each other or each ring is rotated relative to its neighbor. If such rotation systematically progresses helically throughout the stack, CD studies can reflect this macromolecular arrangement by the presence of Cotton effects (CE). The sign of CE indicates the character of the helix: a right-handed or *P*-helix results in positive CE at longer wavelengths and negative CE at shorter, and a left-handed or *M*-helix—in negative CE at longer wavelengths and positive CE at shorter.[108] The zero-crossing point (between two CE) generally corresponds to a maximum UV-Vis absorption peak, attributed to interacting chromophores.[109] For the **P1**/**P2**-MAD system, the spectra of pure **LP** did not show a strong CD signal, whereas each pure **CP** spectra had two CEs, characteristic of a left-handed helix (Figure 2.16a). The zero-crossing point for **CP1** was 295 nm, and for **CP2**—283 nm, corresponding to absorption bands of aromatic amino acid residues in peptide sequence: tryptophane, tyrosine, and phenylalanine. Following the kinetics of **P1**/**P2** with MAD, it is possible to track the appearance of CD bands (Figure 2.16b–c). Using peptide-MAD ratios that yield uniquely **CP**, we can see that half-lives from CD bands align with that from the corresponding HRMS experiments (Figure 2.16d-e, Figure 2.9b–c). Another curious feature of the CD of **P1**-MAD run is peaking molar ellipticity at 259–283 nm, specifically between 0 and 700 min (Figure 2.16e), which correlates with the lifetime of **MP** of **P1** (Figure 2.9c).

Generally, CD is also used to determine types of protein secondary structure: α-helix and β-sheet, and their relative amounts. It is done by comparing and deconvoluting bands in the region 190–240 nm. This spectral region is unfortunately inaccessible for the peptide-MAD system because of the interference with the solvent mixture that starts absorbing light at these wavelengths.

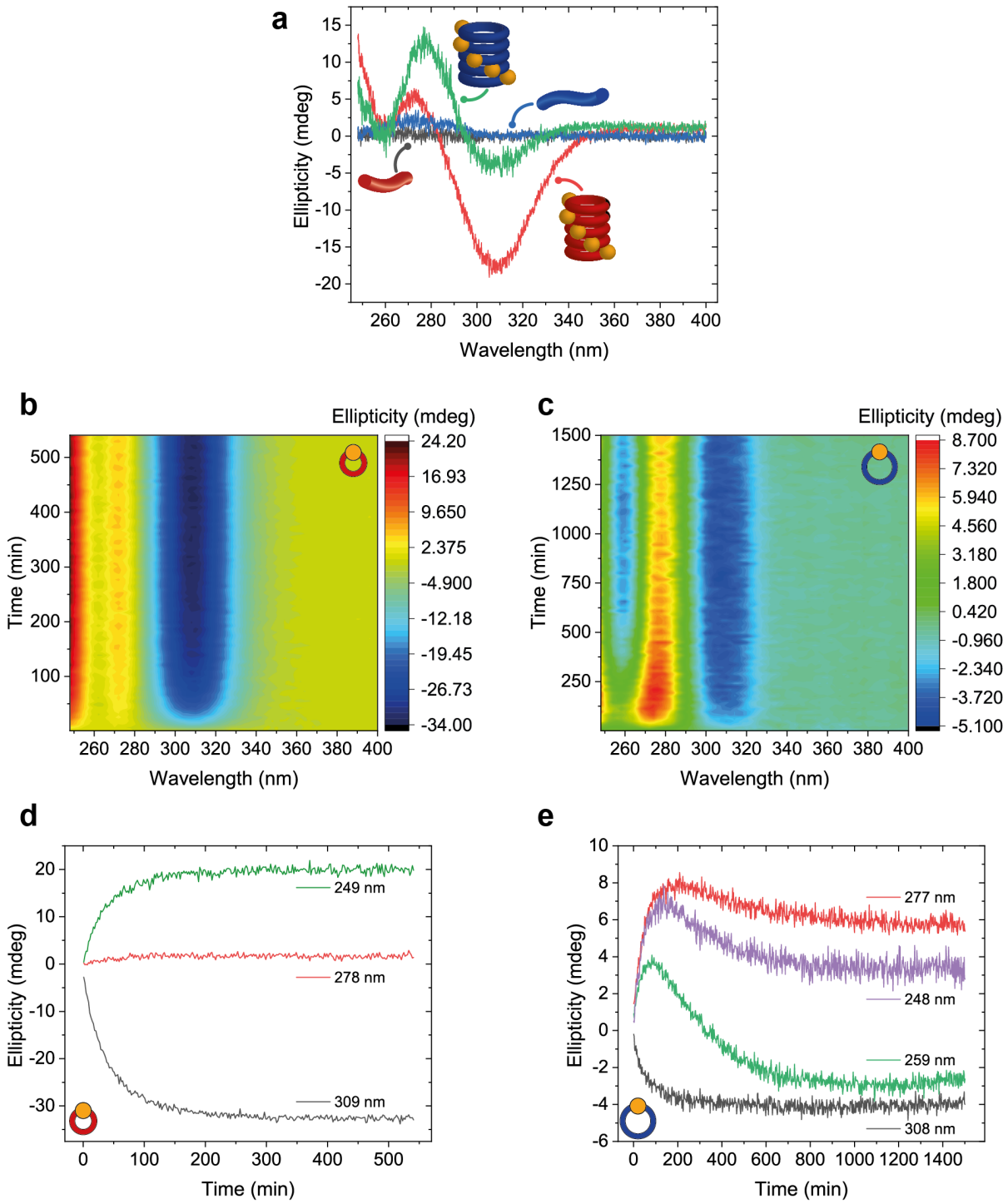


Figure 2.16 | CD data of P1 and P2 from LP to CP. (a) CD spectra of **LP1** and pure **CP1** (top, blue) and the same of **P2** (bottom, red) at 0.1 mM. (b) CD kinetics of 0.1 mM **P2** + 1.5 eq. MAD. (c) CD kinetics of 0.1 mM **P1** + 0.9 eq. MAD. (d) Ellipticity at selected wavelengths plotted over time for kinetics 0.1 mM **P2** + 1.5 eq. MAD. (e) Ellipticity at selected wavelengths plotted over time for kinetics 0.1 mM **P1** + 0.9 eq. MAD. All measurements were performed in a 1 cm quartz cuvette.

2.4 Conclusions

In this chapter, we achieved the cyclization of two different peptides **P1** and **P2** with a selective to amine and thiol Michael acceptor, MAD. The differences in the peptide lengths (and possibly amino acid composition) impact the rate of the reaction, the equilibrium product composition, the threshold of product precipitation, and the linear peptide solubility. Their effective molarities differ by three orders of magnitude. For **P1**, EM is 0.05 mM lower than the concentration at which the majority experiments were done, which explains why chain products were often more abundant than cyclic. Besides, the concentration of MAD for **P1** experiments is always higher than EM, which favors formation of oligomers. The reaction kinetics of both systems was assessed by both HRMS and UV-Vis and the results were commensurate and in reasonable with a simple ODE-based model. The reaction rate constants for the different processes are valuable for the following Chapters 3 and 4, where we describe peptide-decyclization and transient self-assembly, respectively.

Outlook

The equilibrium of the formed peptide products with the released methyl mercaptan at each reaction step influences the reaction rate and cyclic product composition. Removing the methyl mercaptan would lead to faster conversions and decrease macrocycle product speciation (favoring monocycles over larger macrocycles). Sparging the reaction mixture with nitrogen partially solves the problem but is impossible to implement at a small reaction scale (100-150 µL, at which all experiment without sparging were done) as it removes the solvent. Therefore, in preliminary studies, we tried to overcome the mercaptan problem chemically, by changing the structure of MAD. The idea was to have a different leaving group instead of methyl mercaptan that cannot re-attack MAD—bromide ion/HBr. We tested two new Michael acceptors in the reaction with amines and thiols in comparison with that of MAD, followed by UV-Vis (see Appendix Section 2.6.6.2.1 and 2.6.6.2.2). Compound **A** (malonate-derived Michael acceptor) reacted faster with thiol ($t_{1/2} = 25$ min) than with amine ($t_{1/2} = 66$ min). This would pose a drawback for the reaction with peptide because the formation of peptide-**A**-peptide products, connected by thiol, would be favored instead of desired **CP**. Compound **B** (MAD-Br) turned out to be unstable in our solvent mixture, probably due to hydrolysis. Unfortunately, both compounds were not ideal to substitute MAD in obtaining **CP**. Although, compound **B** showed promising results that can be used in a new reaction cycle in chloroform, based on different reactivity of **B** towards acidic/basic thiols.

2.5 Author contributions

Anastasiia Sharko performed all kinetics measurements, modeling in Python, experimental self-assembly studies, and imaging. Jorge Valera Gonzalez performed fitting of the data of supramolecular polymerization UV-Vis study and helped to interpret the results. Purification of CPs was done by Valentin Bauer at ISIS, University of Strasbourg. Nishant Singh came up with the idea to use MAD for CPNs formation and contributed to the discussions at the beginning of the project. Oleksandr Shyshov synthesized compounds A and B and helped with UV-Vis studies of their reaction with amines and thiols. Thomas Hermans supervised the research.

2.6 Appendix

2.6.1 Materials and methods

Chemicals and Materials. Chemicals and solvents were purchased from TCI chemicals, Sigma Aldrich, Fisher. **P1**, **P2**, **N-Ac-P1**, **P1-Met** were purchased from Ontores and Genscript. For the reaction system general solvent mixture composition was 1.6 % DMF, 16.7 % acetonitrile and 81.7 % buffer, that was reasoned by the peptide and MAD solubility. A VWR Model 110 pH meter was used to record the pH measurements during buffer preparations.

UV-Vis. UV experiments were performed on an Agilent CARY 8454-UV spectrometer in 1–2 mm quartz cuvettes. Wavelength range 235–600 nm or 190–1100 nm.

CD. CD spectra were recorded on a JASCO J-1500 or JASCO J-810 spectropolarimeters equipped with a Peltier system as temperature controller. Spectra were recorded in a 1 mm quartz cuvette.

HRMS. HRMS was performed with an Thermofisher Scientific UltiMate 3000 RSLCnano UHPLC System coupled with EMT Thermo OrbiTrap Mass analyzer and the type of ionization was ESI using Hypersil GOLD column, 50 x 2.1 mm, 1.9 μm.

Confocal microscopy. Images were acquired using a ZEISS LSM 710 confocal microscope. Thioflavin-T was used as a fluorescent marker (2 uL of sat. solution per 15 uL of 1.4 mM peptide).

AFM. The measurements were performed under ambient condition using a Bruker Veeco Dimension 3100 AFM microscope with Nanoscope IV controller, operating in tapping mode. Silica probes Bruker TESPA-V2 with a resonance frequency of ~360 kHz and a spring constant of ~42 Nm^{-1} were used. The samples were prepared by preparing a diluted solution of the reaction mixture with MQ water (10 nM–100 μM), followed by dip-coating, spin-coating or drop-casting on Si/SiO_2 wafer.

SEM. SEM images were taken using a FEI Quanta FEG 450 microscope. Using a 5.0 kV High Voltage and dwell time of 50 μs. The drop-casted sample of crude cyclization reaction was washed repeatedly with water to remove the salts and wastes, and sputtered with 10 to 15 nm of gold.

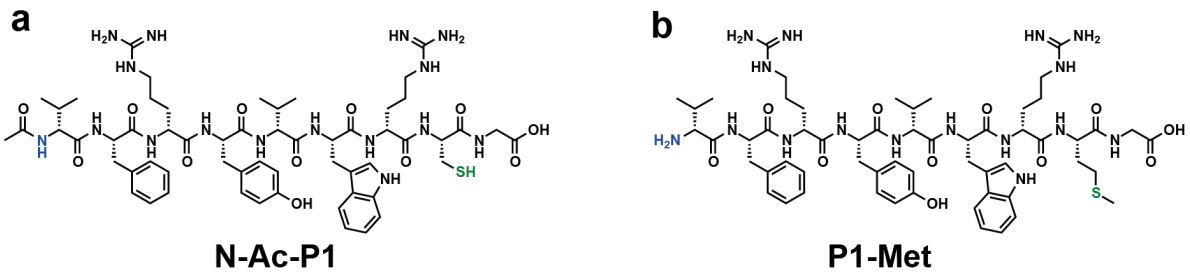


Figure 2.17 | Structures of N-Ac-P1 and P1-Met. These peptides were used to separately study reaction of peptide amine and thiol groups with MAD by UV-Vis and LC-HRMS.

2.6.2 Synthesis of MAD

Synthetic procedure of MAD is well-known and is thoroughly described in several literature sources.[85,110] We closely reproduced this procedure without optimization to obtain similar yields.

2.6.3 General synthesis of CP

Both **CP1** and **CP2** were synthesized according to the procedure described by Anslyn and co-workers,[67,83] but keeping the peptides at low concentrations to favor intramolecular cyclization. After 3 days, the reaction solutions were lyophilized, and the solid crude was purified by preparative HPLC, yielding fluffy colorless powder. The compounds were confirmed by LC-HRMS (**CP1** [M+H$^+$] 1337.6096, [M^{2+}] 669.3084; **CP2** [M+H$^+$] 843.3221) and ^{1}H NMR in deuterated DMSO.

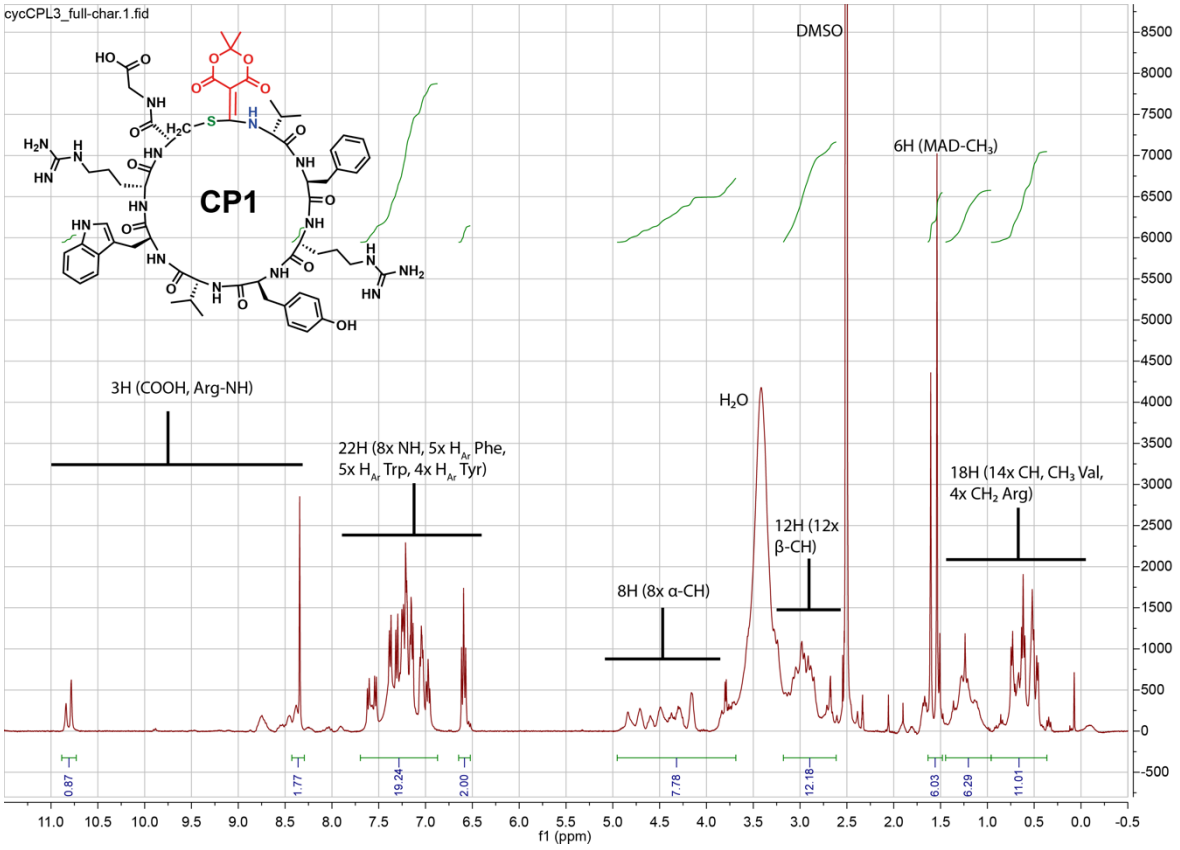


Figure 2.18 | ^{1}H NMR spectrum of CP1 in DMSO-d$_6$.

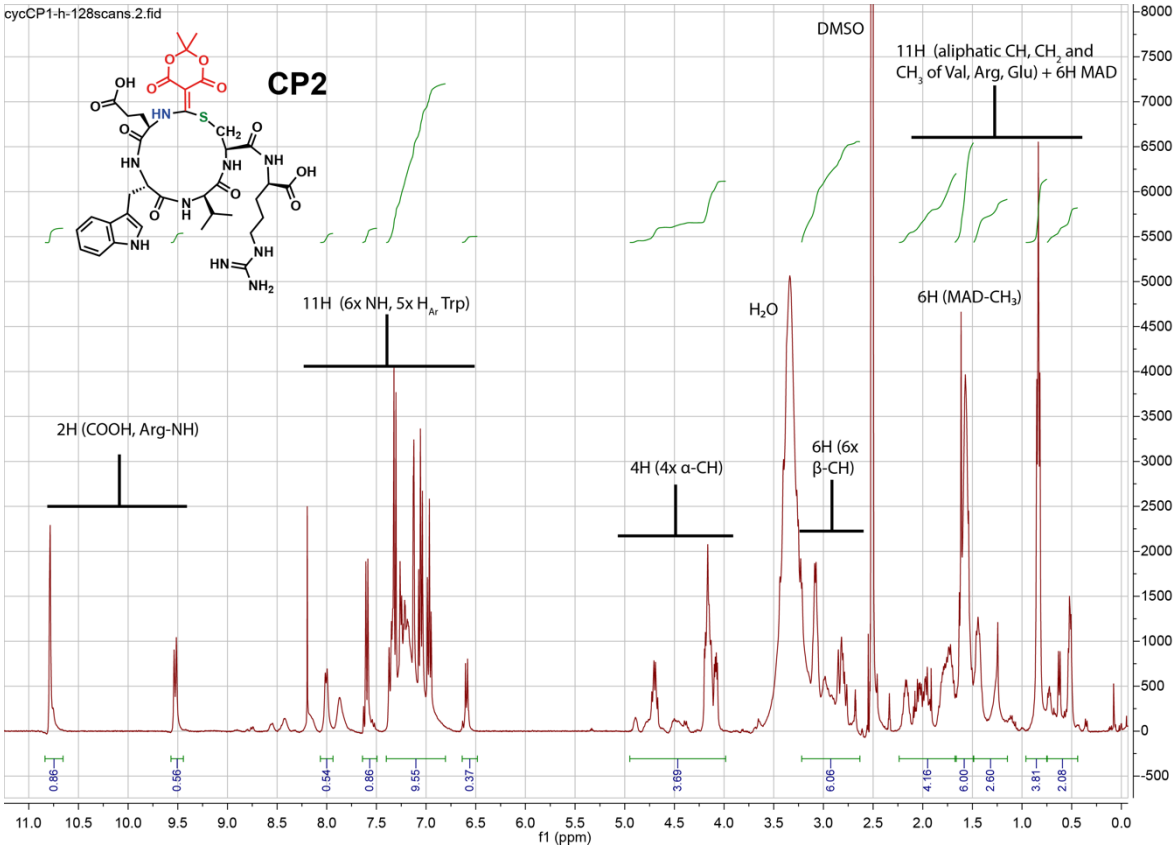


Figure 2.19 | ^{1}H NMR spectrum of CP2 in DMSO-d_6.

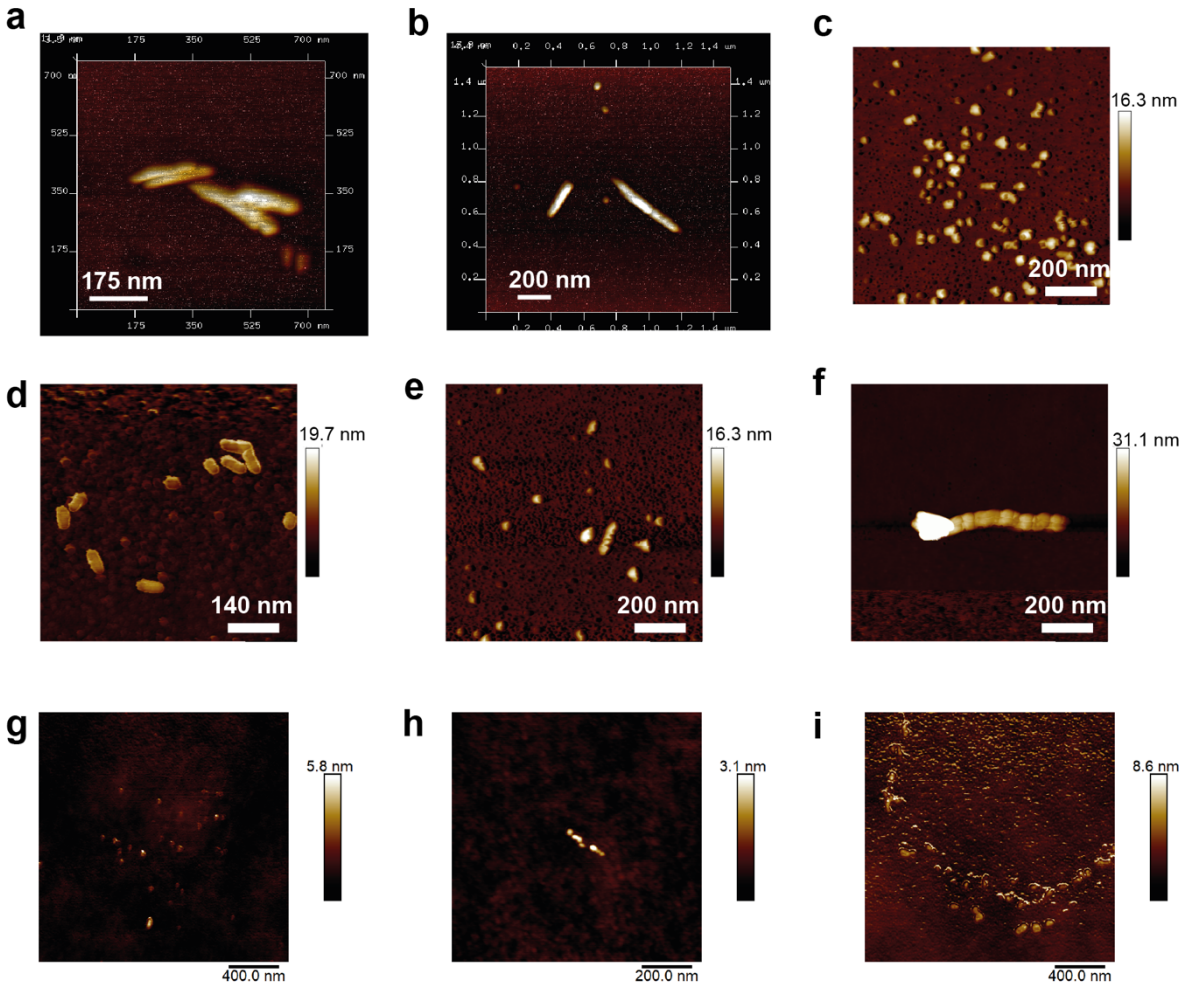


Figure 2.20 | AFM images of P1 and P2 cyclization products. (a–b) Sample taken from the crude reaction **P1** with MAD at pH 8 and diluted 10 times. Solution 0.14 mM was sonicated for 20 min and then spin-coated on Si/SiO_2 wafer. (c–f) 0.1 mM **CP1** drop-casted (c) or spin-coated (d–f). (g-i) 0.1 mM **P2** cyclization crude, dip-coated.

2.6.4 LC-HRMS studies

2.6.4.1 General measurement setup

For the kinetic plots, the relative concentration of each species was estimated as

$$\text{Peak area fr.} = \frac{\text{Peak area of specific species}}{\text{Sum of peak areas of all peptide species}}$$, and plotted over time. All of the reactions were usually done on 150 µL scale in the standard solvent mixture specified in the Materials and Methods. The reaction solution was prepared in a 1.5-mL LCMS glass vial with a screw cap from freshly prepared stock solutions of all reactants diluted to match the reaction concentration. The measurements usually started 4–6 min after reactants were mixed and were programmed for periodic sampling for several hours. All reactions were kept under normal conditions: 20°C (293.15 K), 1 atm.

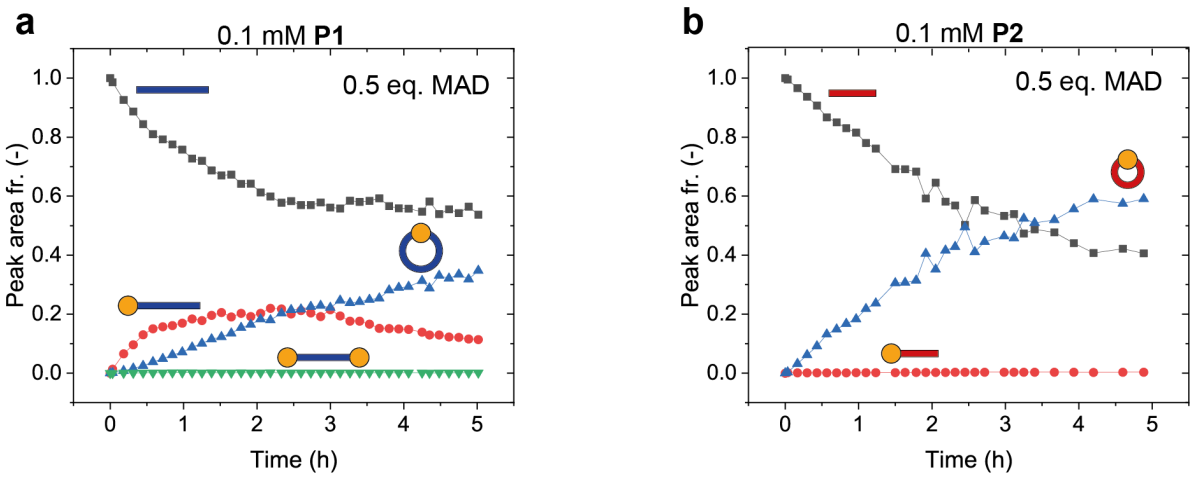


Figure 2.21 | Influence of low MAD amount on the kinetics of P1 and P2 cyclization reaction. (a) 0.1 mM **P1** + 0.5 eq. MAD. (b) 0.1 mM **P2** + 0.5 eq. MAD. All data were obtained from LC-HRMS studies.

2.6.4.2 Cyclization model code

The following code is exemplarily for getting rate constants from the experiment 0.1 mM **P1** + 10 eq. MAD. For every other experiment, parameters k_1, k_2, and k_3 were adjusted, along with starting concentrations and tvec (time domain on which the fitting occurs, usually a bit longer than experimental time).

Input [ ]:

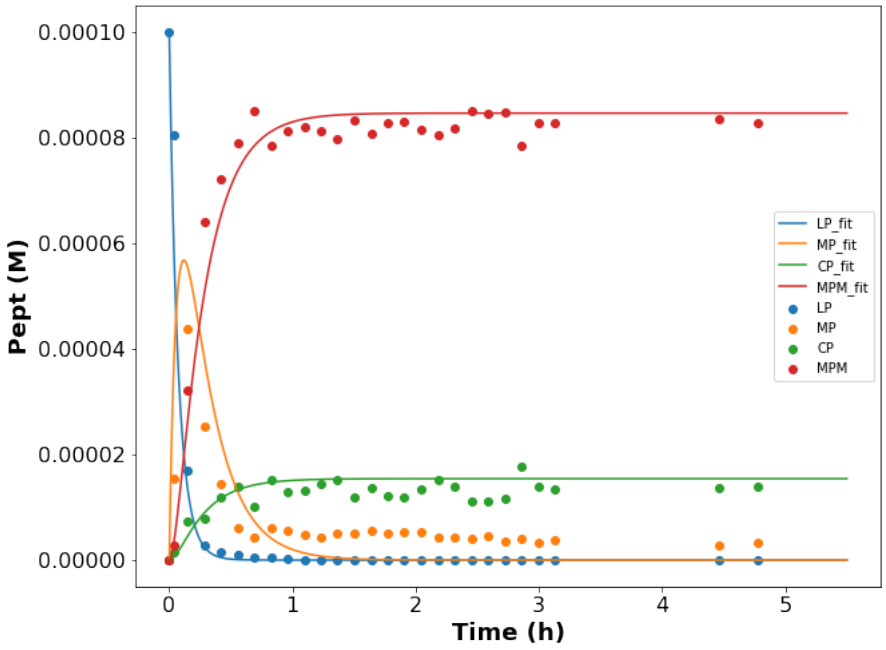
```
import pandas as pd
import numpy as np

#Here we import experimental data file:
p215 = pd.read_csv('/Users/anastasiia/Documents/2019_MAD-
CP_project/Documents/Python/Data/20210220-MS-0p1mM-cpl3-10eq-mad.csv', \
                   sep=',')

from symfit import parameters, variables, Fit, D, ODEModel, Parameter
import matplotlib.pyplot as plt

#needed for the correct figure output in Jupyter Notebook:
%matplotlib nbagg
%matplotlib inline

#Here we define variables (conc & time) and parameters (rate constants).
The latter require good starting values or/and limits (min, max) for a
better fit:
t, p, m, pm, cp, ol = variables('t, p, m, pm, cp, ol')
k1 = Parameter('k1', value=4*60*60) # min=1*60, max=2*60)
k2 = Parameter('k2', value=0.00019*60*60) #, min=0.009*60, max=0.045*60)
k3 = Parameter('k3', value=1.2*60*60) #, min=0.3*60, max=0.8*60)

#Import of data but more specifically from the file:
t_data=np.array(p215["time-h"], dtype=np.float64)
p_data=np.array(p215.LP, dtype=np.float64)
pm_data=np.array(p215.MP, dtype=np.float64)
```

```
cp_data=np.array(p215.CP, dtype=np.float64)
ol_data=np.array(p215.MPM, dtype=np.float64)

#Starting conditions (LP and MAD concentration, M):
t0=0.0
p0=0.0001
m0=0.001
pm0=0.0
cp0=0.0
ol0=0.0

#ODEs and the model itself:
model_dict = {D(p, t): -k1*p*m, \
              D(pm, t): k1*p*m - k2*pm - k3*pm*m, \
              D(cp, t): k2*pm, \
              D(ol, t): k3*pm*m, \
              D(m, t): -k1*p*m - k3*pm*m}


ode_model = ODEModel(model_dict, initial={t: t0, p: p0, m: m0, pm: pm0, cp:
cp0, ol:ol0})
print (ode_model)
fit = Fit(ode_model, t=t_data, p=p_data, pm=pm_data, cp=cp_data,
ol=ol_data, m=None) #here specify what experimental data we have and what
concentration profile we do not know (MAD)

fit_result = fit.execute()
print (fit_result)

tvec = np.linspace(0.0, 5.5, 10000)
cp, m, ol, p, pm = ode_model(t=tvec, **fit_result.params)

#Plotting results:
plt.figure(figsize=(10, 8))
plt.rcParams['xtick.labelsize'] = 16
plt.rcParams['ytick.labelsize'] = 16
plt.plot(tvec, p, label='LP_fit')
#plt.plot(tvec, m, label='mad_fit')
plt.plot(tvec, pm, label='MP_fit')
plt.plot(tvec, cp, label='CP_fit')
plt.plot(tvec, ol, label='MPM_fit')

plt.scatter(t_data, p_data, label='LP')
plt.scatter(t_data, pm_data, label='MP')
plt.scatter(t_data, cp_data, label='CP')
plt.scatter(t_data, ol_data, label='MPM')
plt.xlabel('Time (h)', fontsize=18, fontweight = 'bold')
plt.ylabel('Pept (M)', fontsize=18, fontweight = 'bold')
plt.legend(loc='best')
plt.show()
```

Output [ ]:

```
Derivative(cp, t; k1, k2, k3) = k2*pm
```

```
Derivative(m, t; k1, k2, k3) = -k1*m*p - k3*m*pm
Derivative(ol, t; k1, k2, k3) = k3*m*pm
Derivative(p, t; k1, k2, k3) = -k1*m*p
Derivative(pm, t; k1, k2, k3) = k1*m*p - k2*pm - k3*m*pm

Parameter Value        Standard Deviation
k1        1.440000e+04 2.666616e+03
k2        6.840000e-01 1.189716e-01
k3        4.320000e+03 5.485929e+02
Status message         Optimization terminated successfully.
Number of iterations   0
Objective              <symfit.core.objectives.LeastSquares object at
0x7ff34eccfb80>
Minimizer              <symfit.core.minimizers.BFGS object at
0x7ff34eccf190>

Goodness of fit qualifiers:
chi_squared            2.398340471834786e-09
objective_value        1.199170235917393e-09
r_squared              0.9239096187167792
```



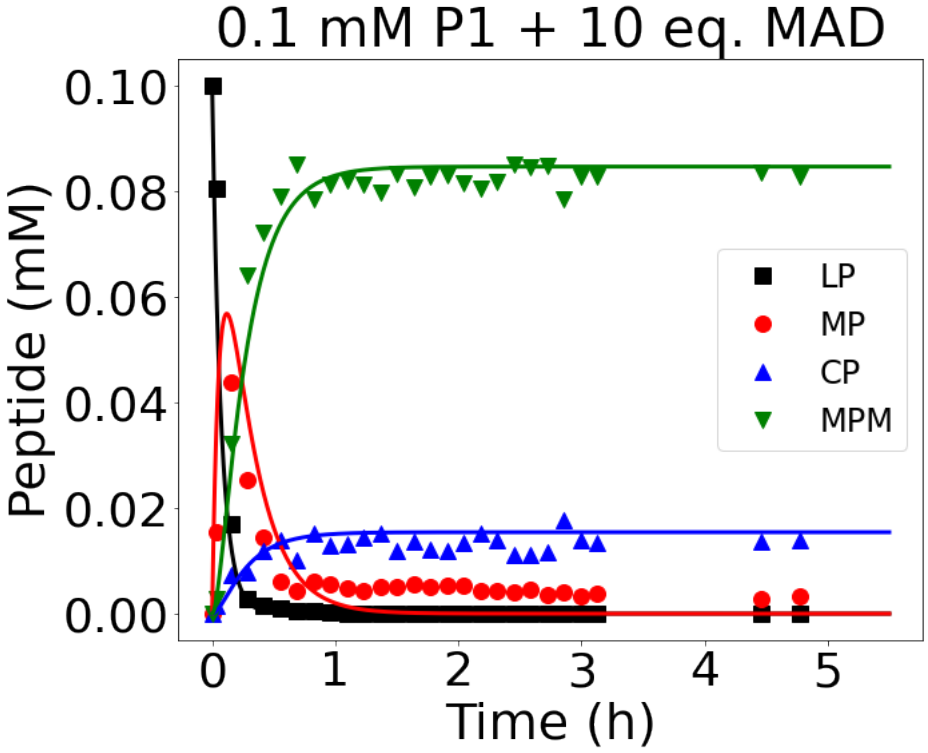
Additional formatting features (Figure 2.11):

Input [ ]:

```
fig, ax = plt.subplots(1,1, figsize=(10,8))

ax.plot(t_data, p_data*1000, marker='s', markersize=12, ls='',
color='black', label='LP')
ax.plot(t_data, pm_data*1000, marker='o', ls='', markersize=12,
color='red', label='MP')
ax.plot(t_data, cp_data*1000, marker='^', markersize=12, ls='',
color='blue', label='CP')
ax.plot(t_data, ol_data*1000, marker='v', markersize=12, ls='',
color='green', label='MPM')
```

```
ax.plot(tvec, p*1000, 'black', tvec, pm*1000, 'r', tvec, cp*1000, 'b',
tvec, ol*1000, 'g', linewidth = 3) #tvec, ol*1000, 'g', --> put this or
remove depending if you have MPM or not in your experimental data.

ax.set_xlabel('Time (h)', fontsize=38, fontweight = 'normal')
ax.set_ylabel('Peptide (mM)', fontsize=38, fontweight = 'normal')
plt.xticks(fontsize=36, fontweight = 'normal')
plt.yticks(fontsize=36, fontweight = 'normal')
ax.set_title('0.1 mM P1 + 10 eq. MAD', fontsize=40, fontweight = 'normal')
ax.legend(loc=0,fontsize=24)

plt.show()
fig.savefig('fit-0p1mM-P1-10eqMAD.svg', dpi=300, transparent=True)
```

Output [ ]:



2.6.4.3 How to get good starting values of the parameters: model with sliders

An initial guess of the parameters influences the fitting quality. The better the values, the closer the fit. To get an idea of the rate constants, the model with sliders is very useful, as it allows for adjustment of the concentrations and rate constants while immediately updating the resulting curve profiles. Combining it with the experimental data permits quickly estimating rate constants that one can use in the previous model for the final optimization.

Input [ ]:

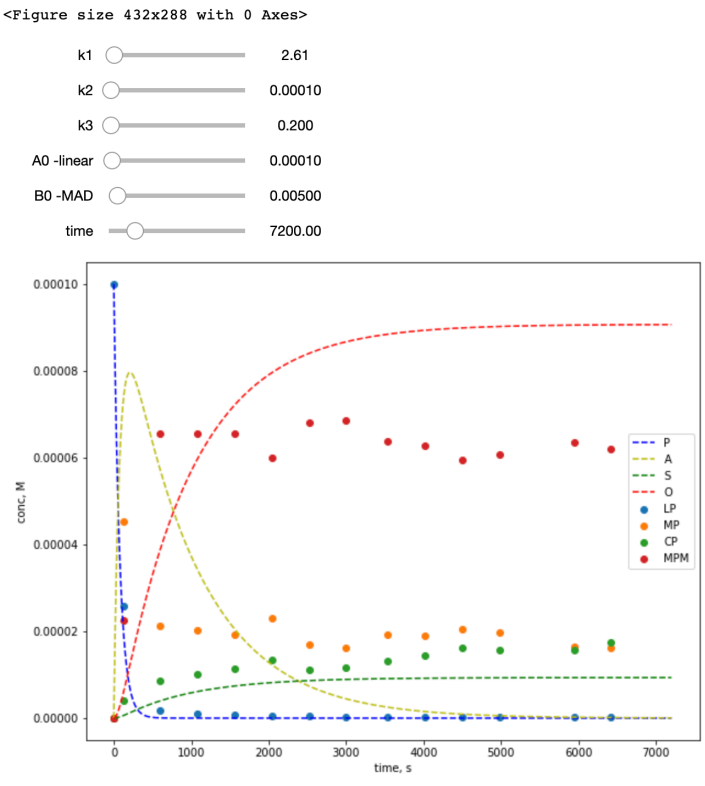
```
import numpy as np
import pandas as pd
from scipy.integrate import odeint
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D
```

```
import ipywidgets as widgets
from ipywidgets import fixed
from IPython.display import display

%matplotlib nbagg
%matplotlib inline

dat = pd.read_csv('/Users/anastasiia/Documents/2019_MAD-
CP_project/Documents/Python/Data/20210715-MS-0p1mM-cpl3-50eq-mad.csv', \
                  sep=',')
t_data=np.array(dat["time-min"], dtype=np.float64)
p_data=np.array(dat.LP, dtype=np.float64)
pm_data=np.array(dat.MP, dtype=np.float64)
cp_data=np.array(dat.CP, dtype=np.float64)
ol_data=np.array(dat.MPM, dtype=np.float64)

plt.figure()

def update_plot(k1, k2, k3, P, M, time):

    # initial condition
    C0 = [P,M,0,0,0]

    # function that returns Cout
    def model(C,t):
        P = C[0]
        M = C[1]
        A = C[2]
        S = C[3]
        O = C[4]
        dPdt = -k1*P*M
        dMdt = -k1*P*M - k3*A*M
        dAdt = k1*P*M - k3*A*M - k2*A
        dSdt = k2*A
        dOdt = k3*A*M
        Cout = [dPdt,dMdt,dAdt,dSdt,dOdt]
        return Cout


    # time points
    steps = 10000
    t = np.linspace(0,time,steps)

    # solve ODE
    Cint = odeint(model,C0,t)

    # simple plot results
    plt.figure(figsize=(10, 8))
    plt.clf() # clear the plot
    plt.plot(t,Cint[:,0],'b--',label=r'P')
    #plt.plot(t,Cint[:,1],'--',label=r'M')
    plt.plot(t,Cint[:,2],'y--',label=r'A')
    plt.plot(t,Cint[:,3],'g--',label=r'S')
```

```
    plt.plot(t,Cint[:,4],'r--',label=r'O')

    plt.scatter(t_data*60, p_data, label='LP') #*60 if exp. data is in min,
*3600 if in hrs.
    plt.scatter(t_data*60, pm_data, label='MP')
    plt.scatter(t_data*60, cp_data, label='CP')
    plt.scatter(t_data*60, ol_data, label='MPM')

    plt.ylabel('conc, M')
    plt.xlabel('time, s')
    plt.legend(loc='best')
    plt.show()

k1 = widgets.FloatSlider(min=0.01, max=100, value=1.6, step=0.01,
description='k1')
k2 = widgets.FloatSlider(min=0.00001, max=1, value=0.00011, step=0.00001,
description='k2', readout_format='.5f')
k3 = widgets.FloatSlider(min=0.01, max=100, value=0.25, step=0.005,
description='k3', readout_format='.3f')
P = widgets.FloatSlider(min=0, max=0.01, value=0.0001, step=0.0001,
description='A0 -linear',readout_format='.5f')
M = widgets.FloatSlider(min=0, max=0.1, value=0.005, step=0.0001,
description='B0 -MAD',readout_format='.5f')
time = widgets.FloatSlider(min=1, max=40000, value=7200, step=120,
description='time')
widgets.interact(update_plot, k1=k1, k2=k2, k3=k3, P=P, M=M, time=time)
```

Output [ ]:

```
<Figure size 432x288 with 0 Axes>
```




```
<function __main__.update_plot(k1, k2, k3, P, M, time)>
```

Both codes for the cyclization model were adapted and expanded by the author from the original code developed in Prof. Hermans' lab.

2.6.5 UV-Vis studies

2.6.5.1 General measurement setup

Stock solutions were prepared: 2 mM peptide with 2.5 eq. TCEP in the solvent mixture (1.7 % DMF, 16.7 % ACN, 81.6 % phosphate buffer pH 7.4, 0.1 M); 5 mM MAD in ACN. In a quartz cuvette (1-2 mm optical path), first, the solvent mixture was added, and the blank measurement was done. This blank spectrum is automatically subtracted from each subsequent measurement. Then peptide solution was added, the cuvette was shaken, after which the MAD was added, and the cuvette was thoroughly shaken again. The kinetics measurement started usually after 30–40 s after MAD had been added, which is accounted for in the further data treatment.

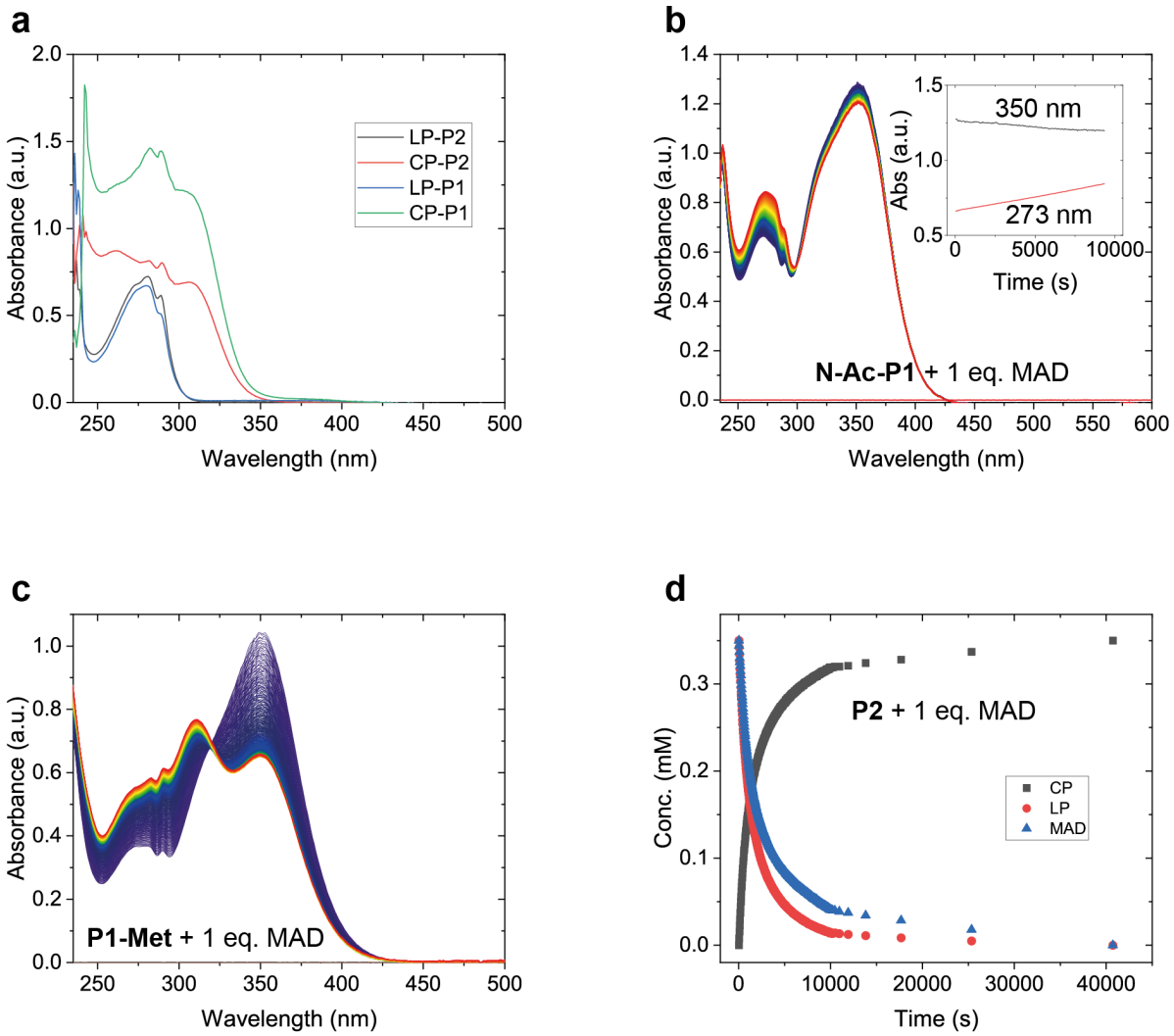


Figure 2.22 | Additional data on UV-Vis studies. (a) Linear, cyclic **P1** and **P2**. (b) 0.7 mM **N-Ac-P1** + 1 eq. MAD; in the inset absorbances of 350 and 273 nm over time. (c) 0.35 mM **P1-Met** + 1 eq. MAD reaction progress. (d) Spectra decomposition for the experiment 0.35 mM **P2** + 1 eq. MAD.

2.6.5.2 Fractional-life period method[92]

The method is used to determine the reaction order from the half-lives. If the rate expression is of the form

$$\frac{dx}{dt} = k(a-x)^n, \tag{2.12}$$

the half-life for $n = 1$ is

$$t_{1/2} = \frac{\ln 2}{k} \tag{2.13}$$

and for $n \neq 1$ is

$$t_{1/2} = \frac{2^{n-1}-1}{k(n-1)a^{n-1}}. \tag{2.14}$$

For two experiments that are done at different initial concentrations, it is possible to express the reaction order through the corresponding half-lives:

$$n = 1 + \frac{\log t'_{1/2} - \log t_{1/2}}{\log a - \log a'}. \tag{2.15}$$

Similarly, a single experiment can be used, but at two different time intervals:

$$n = 1 + \frac{\log[(t_2/t_1) - 1]}{\log[1/(1-y)]}, \tag{2.16}$$

where t_1 is the time corresponding to the concentration $a(1-y)$, t_2 is the time corresponding to the concentration $a(1-y)^2$, a is the initial concentration, and y is the reacted fraction.

Equation (2.16) was used to estimate whether the reaction of **P1** or **P2** with MAD can be precisely described by the second-order rate law. Instead of the concentration the absorbance value at 355 nm was taken at a corresponding time.

Determination of the reaction order from the experiments 0.35 mM **P2** + 1 eq. MAD (A, B); 0.35 mM **P1** + 1 eq. MAD (C):

	$a(1-y)$, a.u.	t_1, s	$a(1-y)^2$, a.u.	t_2, s	y	n
A)	0.33854	360	0.27069	775	0.20	1.08845
B)	0.31081	330	0.24389	760	0.22	1.17236
C)	0.33927	536	0.26950	11211	0.21	1.14579

2.6.5.3 General integrated rate expression for stoichiometric reaction for a single component

The fitting was performed on the data from UV-Vis kinetics experiment: 0.35 mM **P2** + 1 eq. MAD. According to Equation (2.7), $f(c, n)$ was plotted over time t. Instead of concentration, absorbance values at 355 nm were used. With the variation of n, the graph was changed, and linear fit was done with the Intercept = 0 (Figure 2.23a–d). The results were compared by the value of R^2 and by the analysis of residuals. For $n > 2$ the values of R^2 increase, however, remaining in the range 0.93–0.96. The values of the last eight time points severely deviate from the rest of the data. Removing them from the dataset and performing a linear fit gives the best result at n = 2.2 with $R^2 = 0.9999$. Despite the high value for the coefficient of determination, the plot of residuals as a function of the independent variable in all cases shows a nonrandom

distribution around the regression line (Figure 2.23e–f). This means that the data systematically deviates from the model, and thus the model does not adequately represent the data.[97]

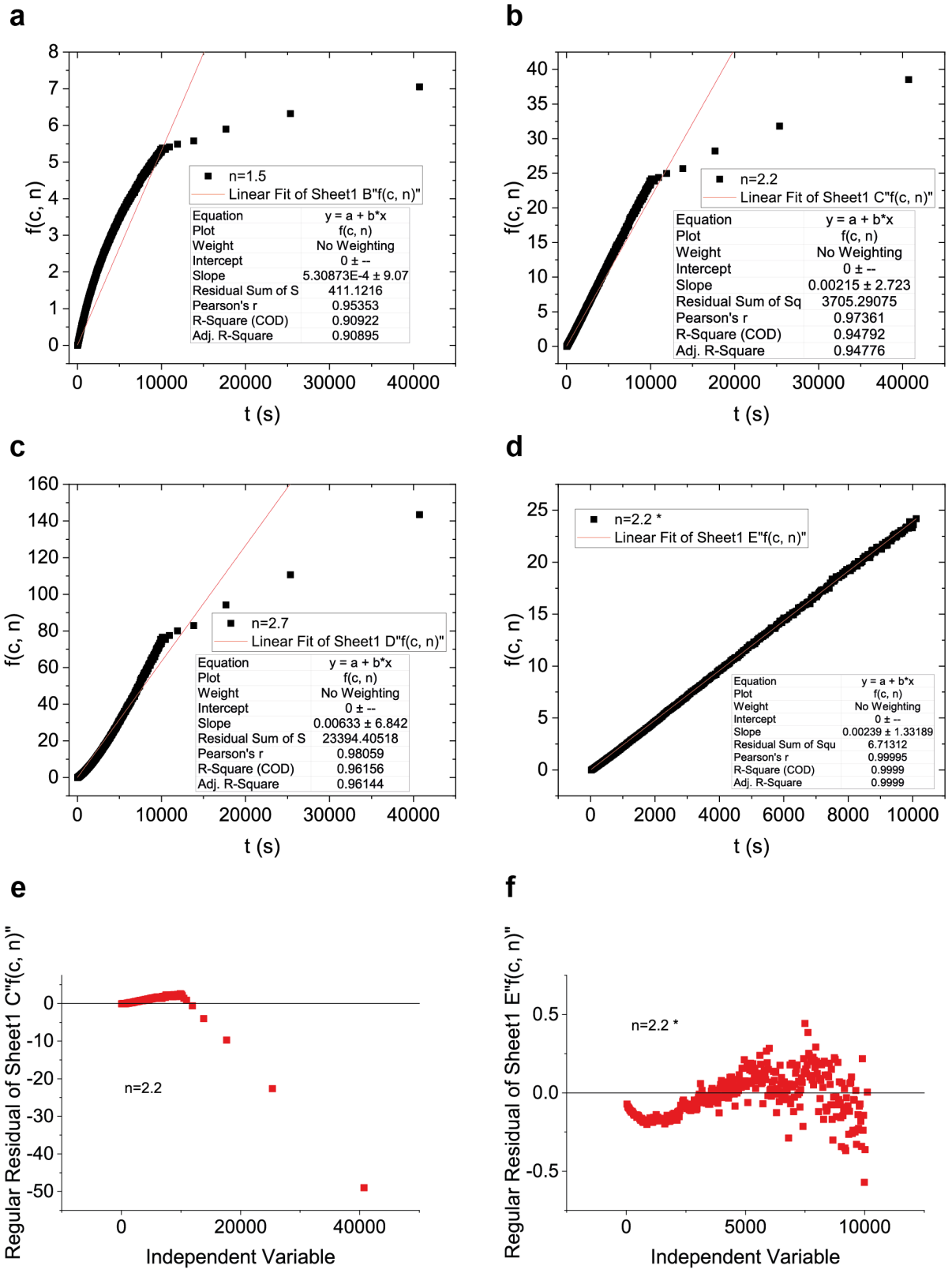


Figure 2.23 | Integrated general rate equation applied for P2 system with varied *n*, linear fit, and analysis of residuals. (a) n=1.5. (b) n=2.2. (c) n=2.7. (d) n= 2.2, fitting performed after the last 8 data points were removed. (e) A plot of the residuals for the linear fit for data at n=2.2. (f) A plot of the residuals for the linear fit for data without the last 8 points at n=2.2.

2.6.5.4 Selected methods to fit second-order reaction kinetics

2.6.5.4.1 Integrated rate law expressed via absorbance over time

Since the relation between absorbance and concentration is linear, it is possible to directly extract rate constants from the plot of function of absorbance over time. The same is applicable to any measured physical property that linearly changes with concentration. Moore and Pearson[92] explain in detail how to get kinetically useful expressions that are further used in the context of the chosen integrated rate law. For the reaction A + B → C, where *a* and *b* are initial concentrations of A and B respectively, $b > a$, and the stoichiometry is 1 to 1:

$$\frac{x}{a} = \frac{\lambda - \lambda_0}{\lambda_\infty - \lambda_0}, \tag{2.17}$$

$$\frac{a}{a - x} = \frac{\lambda_\infty - \lambda_0}{\lambda_\infty - \lambda}, \tag{2.18}$$

$$\frac{b}{b - x} = \frac{(b/a)(\lambda_\infty - \lambda_0)}{(b/a)(\lambda_\infty - \lambda_0) - (\lambda - \lambda_0)}. \tag{2.19}$$

Here, x is the extent of the reaction, $\frac{dx}{dt} = -\frac{d[B]}{dt} = -\frac{d[A]}{dt} = \frac{d[C]}{dt}$; λ_0 is the initial value of absorbance (at t_0 or as early as possible), λ_∞ is the final reading of the instrument (when the reaction is done), λ is the absorbance at time t.

We could express the rate of 2nd order reaction for one reactant (2.20) using (2.17) and (2.18), resulting in (2.21):

$$\frac{1}{a - x} - \frac{1}{a} = \frac{x}{a(a - x)} = kt \tag{2.20}$$

$$\frac{\lambda - \lambda_0}{a(\lambda_\infty - \lambda)} = kt \tag{2.21}$$

The rate of 2nd order reaction with two reactants (2.22) can also be expressed via (2.18) and (2.19) giving (2.23).

$$\frac{1}{a - b} \ln \frac{b(a - x)}{a(b - x)} = kt \tag{2.22}$$

$$\frac{1}{a - b} \ln \frac{(b/a)(\lambda_\infty - \lambda)}{(b/a)(\lambda_\infty - \lambda_0) - (\lambda - \lambda_0)} = kt \tag{2.23}$$

Equation (2.21) was used to fit the data from the experiments with equal initial concentrations of MAD and peptides, equation (2.23)—the data with different initial concentrations of MAD and peptides.

An important detail of this data treatment is the need of a proper final measurements reading λ_∞. Deviation of λ_∞ by a few percent results in a similar change of the rate constant value.[92] In some cases, a theoretical value can be used, for example, for the MAD-**P2** and absorbance at 355 nm or 375 nm at the equilibrium $\lambda_\infty \to 0$ (if all MAD is used up). Using theoretical value can help alleviate the effects of slow secondary processes if they later occur in the system, which for cyclization system is self-assembly.

Determination of the rate constant from fitting (2.21) to the function of absorbance over time from the experiment 0.35 mM **P2** + 1 eq. MAD (Figure 2.24):

Wavelength, nm	Baseline correction?	k, $M^{-1}s^{-1}$	R^2
355	no	1.1947	0.9999
	yes	1.1539	0.9999
375	no	1.2949	0.9998
	yes	1.2844	0.9997

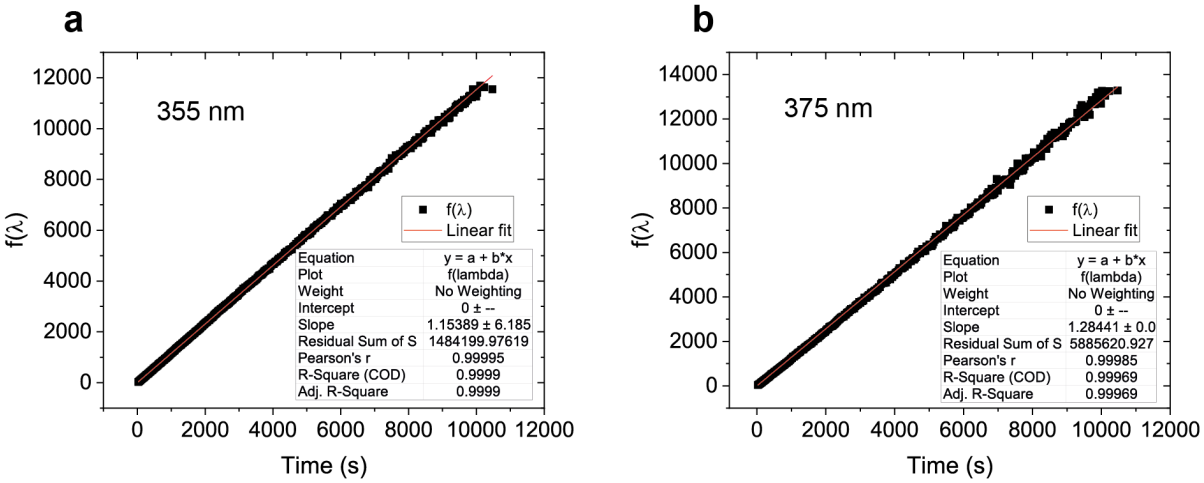


Figure 2.24 | 2nd order integrated rate function applied to the absorbance data of the P2 system with 1 eq. MAD. (a) Absorbance data at 355 nm. (b) Absorbance data at 375 nm. The rate constant is directly extracted from the slope.

Note on the data treatment

The last 6 points (out of 343 points) were removed from the data because of the huge deviation from the main trend. These last points were measured at longer experimental times, between 3 and 11 hours, when the changes to the spectra from the chemical reaction were minor. Other effects could influence the measurement: aggregation of peptide products, and instrumental error.

Baseline correction involves the additional correction of all spectra in the software Spectragryph called 'Simple baseline correction' that forces all spectra through zero at the desired wavelength (400, 440, 552, and 600 nm were tried, giving no significant differences in the resulting values of k). The absence of the baseline correction means that spectra were not treated additionally and were saved raw from the instrument software. In all cases, the spectrum of the solvent mixture was automatically subtracted from each spectrum of the reaction mixture.

For the **P1** system, in the experiment with 1 eq. MAD, the method showed worse compliance with the experimental data, which was expected as more peptide species co-exist at every time point, resulting in deviations of the decrease of MAD absorbance from the 2nd order rate function (Figure 2.25a).

Determination of the rate constant from fitting (2.21) to the function of absorbance over time from the experiment 0.35 mM **P1** + 1 eq. MAD:

Wavelength, nm	Baseline correction?	k, $M^{-1}s^{-1}$	R^2
355	no	1.4017	0.9979
375	no	1.3874	0.998

However, in the experiment with 0.5 eq. MAD, where the absorbance region of MAD is not influenced by the presence of **MPM**, the method showed less deviation from linearity (Figure 2.25b). In this case, the equation (2.23) was used to fit the experimental data.

Determination of the rate constant from fitting (2.23) to the function of absorbance over time from the experiment 0.7 mM **P1** + 0.5 eq. MAD:

Wavelength, nm	Baseline correction?	k, $M^{-1}s^{-1}$	R^2
355	no	0.9389	0.9997
	yes	1.016	0.9995
375	no	0.8827	0.9997
	yes	0.9739	0.9994

Note on the data treatment

No points were removed; the last measurement was at 3.03–3.83 h of the reaction (338–340 data points).

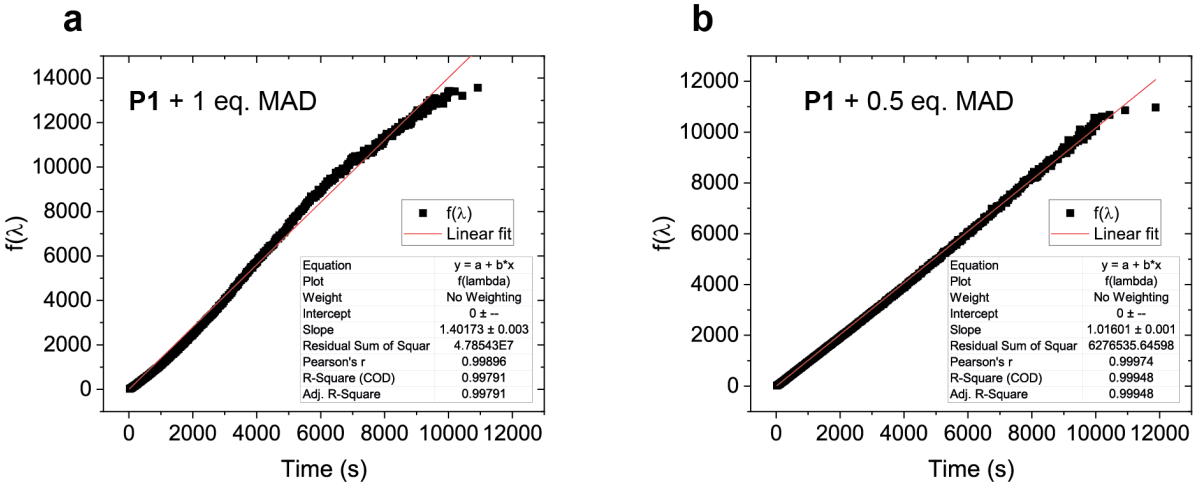


Figure 2.25 | 2nd order integrated rate function applied to the absorbance data of the P1 system. (a) Experiment with 1 eq. MAD. (b) Experiment with 0.5 eq MAD. In both cases absorbance data at 355 nm is taken. The rate constant is directly extracted from the slope.

2.6.5.4.2 Constant time interval method or Guggenheim time lag method

The method assumes that $f(A_1)$ varies linearly with $f(A_2)$ where A_2 is the concentration (or linearly related physical value, λ) of MAD measured at a constant time interval later than A_1.[111,112] In practice, the data is divided into two halves separated by $(t_2 - t_1)$, later absorbance data is subtracted from the first half. Then for the 2nd order reaction rate and equal reactants' concentrations, $(\lambda_1 - \lambda_2)$ is plotted over $(t_2\lambda_2 - t_1\lambda_1)$ according to equation (2.24), and the rate constant is extracted from the slope value:

$$\lambda_1 - \lambda_2 = (A_0)k(t_2\lambda_2 - t_1\lambda_1) - (A_0)\lambda_\infty(t_2 - t_1) \quad (2.24)$$

For non-equal MAD and peptide concentrations, $A_0 > B_0$, the equation (2.25) is used.

$$\frac{1}{\lambda_2} = \frac{\exp(k(B_0-A_0)(t_2-t_1))}{\lambda_1} + \frac{\exp(k(B_0-A_0)(t_2-t_1))-1}{B_0-A_0} \quad (2.25)$$

The graph of $1/\lambda_2$ over $1/\lambda_1$ is linear, the rate constant is obtained from the slope.[111]
The advantage of this method is no need for the precise λ_0 and λ_∞ (zero and infinite instrument reading), compared to the previous method.[113]
Application of this method to the cyclization reaction data gave the following results (Figure 2.26):

Experiment	k, $M^{-1}s^{-1}$	R^2	Equation
0.35 mM **P2** + 1 eq. MAD	0.999	0.9988	(2.i)
0.35 mM **P1** + 1 eq. MAD	1.137	0.9886	(2.ii)
0.7 mM **P1** + 0.5 eq. MAD	1.248	0.9994	(2.iii)

The absorbance values in all experiments were taken at 355 nm, with no additional post-measurement mathematical baseline correction. Noticeably, the treatment with smaller $(t_2 - t_1)$ interval gave better fitting results and more linear data trends.

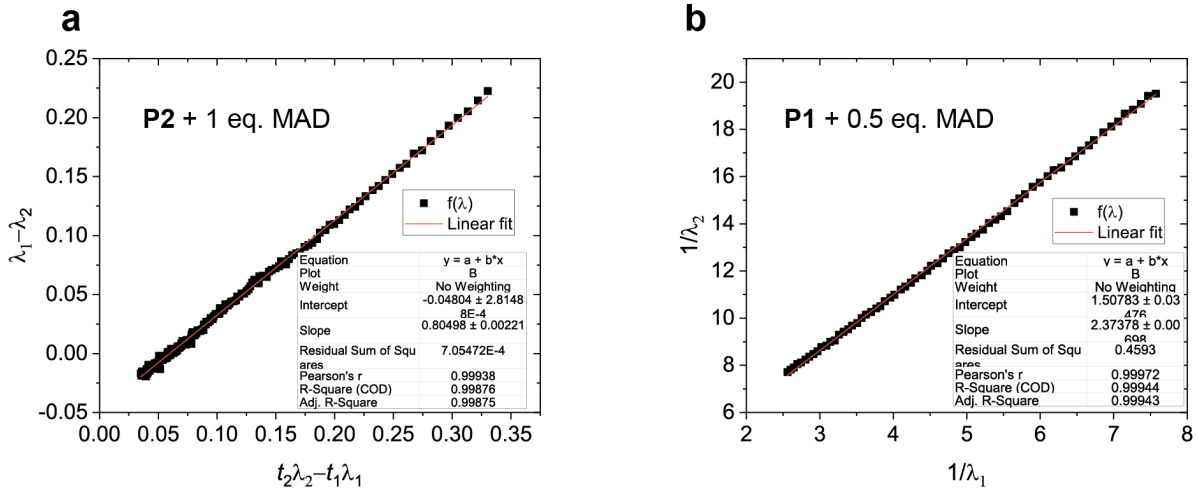


Figure 2.26 | Guggenheim method for fitting absorbance data of P1 and P2 systems. (a) 0.35 mM **P2** with 1 eq. MAD; time lag 4080 s. (b) 0.7 mM **P1** with 0.5 eq. MAD; time lag 2010 s.

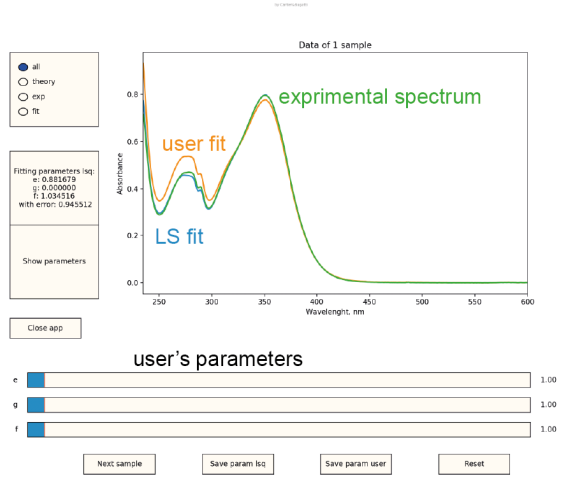
Although all these methods rely on a clean second-order rate law, which does not seem to be fully appropriate for the cyclization reaction of peptides with MAD, the obtained rate constants are of the same order of magnitude as those obtained by numerical integration of the HRMS kinetics data.

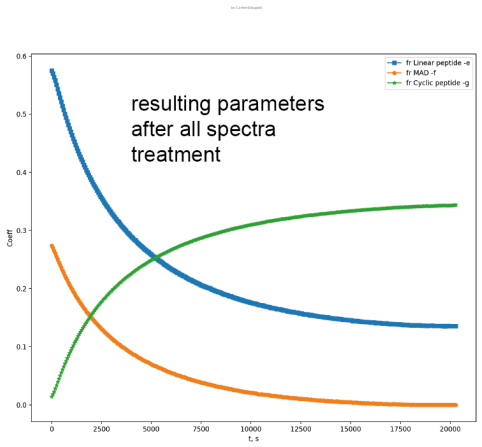
2.6.5.5 UV-Vis spectra decomposition script

The following code "DataFit.py" is general for the **P1** system, and the following treatment is exemplary for the experiment 0.7 mM **P1** + 0.5 eq. MAD. The code was written in Spyder (Python 3.8) within Anaconda Environment, the script is run from Terminal along with the input of initial parameters **y** and size of the data for treatment (number of columns in the Excel file with the experimental data "DATA1.xlsx"). The script also requires components' individual spectra in the file "DATA2.xlsx".

Example (in Terminal): python DataFit.py -e 1 -f 1 -g 1 -ncol 23
Here, the first guessed parameters are 1 1 1 and calculated for 23 samples

Output: a plot of experimental data, theoretical fit and least squares fit with printed parameter values, a plot of parameters over time.





Further data treatment

After parameters attributed to each component were extracted, the following normalization was applied to each parameter function depending on the component concentration. The normalization converts parameter values into concentration.

$$\mathbf{LP} = \frac{p - (p_\infty - 0.35)}{p_1 - (p_\infty - 0.35)} \times 0.7 \text{ mM} \tag{2.26}$$

$$\text{MAD} = \frac{p - p_\infty}{p_1 - p_\infty} \times 0.35 \text{ mM} \tag{2.27}$$

$$\mathbf{CP} = \frac{p - p_1}{p_\infty - p_1} \times 0.35 \text{ mM} \tag{2.28}$$

Here, p is the parameter at the time t, p_1 is the first parameter, p_∞ is the last. **LP**, MAD, and **CP** are concentrations of the components at time t. The resulting graph with concentrations plotted over time see in Figure 2.14b.

The script was developed in collaboration with Yelyzaveta Berezovska and Liudmyla Klochko, who created the script's interface and helped translate the algorithm of data treatment into Python programming language. The same approach with Solver in Excel was made with the help of Thomas Hermans.

DataFit.py:

```
import numpy as np
import matplotlib.pyplot as plt
from matplotlib.widgets import Slider, Button, RadioButtons
import pandas as pd
import argparse
from scipy.optimize import nnls
import sys

def f(a,b,c,x,y,z):
    return a*x+b*y+c*z
```

```
parser = argparse.ArgumentParser(description='members of TSP and MCube from
ICS')
parser.add_argument('-e=', '--e_name',help='first guess e
parameter',type=float)
parser.add_argument('-f=', '--f_name',help='first guess f
parameter',type=float)
parser.add_argument('-g=', '--g_name',help='first guess g
parameter',type=float)
parser.add_argument('-ncol=', '--ncol_name',help='number of samples to
show',type=int)
args=parser.parse_args()

axcolor = 'floralwhite'

e0 = args.e_name
f0 = args.f_name
g0 = args.g_name
number_of_comun_from_data = args.ncol_name
number_of_comun_from_data = number_of_comun_from_data - 1

a1 = pd.read_excel('DATA1.xlsx') #sep='\s*,\s*'
a2 = pd.read_excel('DATA2.xlsx')

name1 = a1.columns.to_numpy()
name2 = a2.columns.to_numpy()

wave = a1.loc[:,str(name1[0])].to_numpy()[1:]

E = a2.loc[:,str(name2[1])].to_numpy()[1:]
F = a2.loc[:,str(name2[2])].to_numpy()[1:]
G = a2.loc[:,str(name2[3])].to_numpy()[1:]

for ns in range(number_of_comun_from_data):
    fig, ax = plt.subplots(figsize=(12,10))

    plt.subplots_adjust(left=0.25, bottom=0.39)
    ax.set_xlabel('Wavelenght, nm')
    ax.set_ylabel('Absorbance')
    ax.set_title("Data of %d sample" %(ns+1))

    ads_theor = e0 * E + f0 * F + g0 * G
    ads_exper = a1.iloc[1:].to_numpy()[:,1:]

    coef = np.vstack((E,F,G)).astype(np.float64)
    #EE,FF,GG =
(np.linalg.lstsq(coef.T,ads_exper[:,ns].astype(np.float64),rcond=None)[0])
    EE,FF,GG = (nnls(coef.T,ads_exper[:,ns].astype(np.float64)))[0]

    dataset_fit, = plt.plot(wave, f(E,F,G,EE,FF,GG), lw=2)
    dataset_theory,  = plt.plot(wave, ads_theor, lw=2)
    dataset_experim, = plt.plot(wave, ads_exper[:,ns], lw=2)
```

```
ax.margins(x=0)

ax_e = plt.axes([0.055, 0.20, 0.85, 0.03], facecolor=axcolor)
ax_f = plt.axes([0.055, 0.10, 0.85, 0.03], facecolor=axcolor)
ax_g = plt.axes([0.055, 0.15, 0.85, 0.03], facecolor=axcolor)

se = Slider(ax_e, 'e',  0.0, 30.0,  valinit = e0,  valstep=0.001)
sf = Slider(ax_f, 'f',  0.0, 30.0,  valinit = f0,  valstep=0.001)
sg = Slider(ax_g, 'g',  0.0, 30.0,  valinit = g0,  valstep=0.001)

def update(val):
    e_coef = se.val
    f_coef = sf.val
    g_coef = sg.val
    dataset_theory.set_ydata(e_coef * E + f_coef * F + g_coef * G)
    fig.canvas.draw_idle()

se.on_changed(update)
sf.on_changed(update)
sg.on_changed(update)

resetax = plt.axes([0.75, 0.025, 0.12, 0.04])
button = Button(resetax, 'Reset', color=axcolor, hovercolor='0.975')

def reset(event):
    se.reset()
    sf.reset()
    sg.reset()
    dataset_experim.set_visible(True)
    dataset_theory.set_ydata(ads_theor)
    dataset_theory.set_visible(True)
    dataset_fit.set_visible(True)

button.on_clicked(reset)

rax = plt.axes([0.025, 0.73, 0.15, 0.15], facecolor=axcolor)
radio = RadioButtons(rax, ('all', 'theory', 'exp', 'fit'), active=0)

def showfunc(label):
    if(label == 'exp'):
        dataset_fit.set_visible(False)
        dataset_theory.set_visible(False)
        dataset_experim.set_visible(True)
    elif (label == 'theory'):
        dataset_fit.set_visible(False)
        dataset_theory.set_visible(True)
        dataset_experim.set_visible(True)
    elif(label == 'all'):
        dataset_fit.set_visible(True)
        dataset_theory.set_visible(True)
        dataset_experim.set_visible(True)
    elif(label == 'fit'):
        dataset_fit.set_visible(True)
```

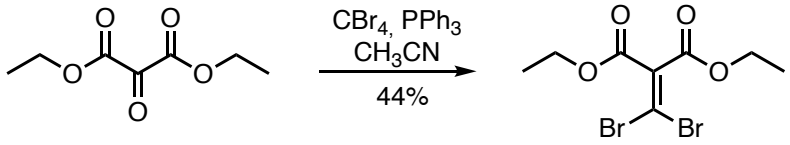
```
            dataset_experim.set_visible(True)
            dataset_theory.set_visible(False)
        fig.canvas.draw_idle()

    radio.on_clicked(showfunc)

    paramax = plt.axes([0.025, 0.38, 0.15, 0.15])
    button_param = Button(paramax, 'Show parameters', color=axcolor,
hovercolor='0.975')

    def printpar(event):
        e0_minim = se.val
        f0_minim = sf.val
        g0_minim = sg.val
        J_theory = e0_minim*E + f0_minim*F + g0_minim*G
        sumErr = np.sum(abs(ads_exper[:,0] - J_theory))
        textstrth = "Fitting from plot:\n e: %f \n g: %f \n f: %f \n with
error: %f \n Update param" %(e0_minim,g0_minim,f0_minim,sumErr)
        print(textstrth)
        button_param.label.set_text(textstrth)

    button_param.on_clicked(printpar)

    J_fit = EE*E + FF*F + GG*G
    fitErr = np.sum(abs(ads_exper[:,0] - J_fit))
    textstr = "Fitting parameters lsq:\n e: %f \n g: %f \n f: %f \n with
error: %f" %(EE,GG,FF,fitErr)
    textax = plt.axes([0.025, 0.53, 0.15, 0.15])
    button_text = Button(textax, textstr, color=axcolor,
hovercolor=axcolor)
    plt.suptitle('by Cartier&Bugatti', fontsize=5, color='grey' )
    print(textstr)

    printax = plt.axes([0.55, 0.025, 0.12, 0.04])
    buttonpr = Button(printax, 'Save param user', color=axcolor,
hovercolor='0.975')
    def saveparuser(event):
        with open('param_fit.dat','ab') as f:
            e0_minim = se.val
            f0_minim = sf.val
            g0_minim = sg.val
            J_theory = e0_minim*E + f0_minim*F + g0_minim*G
            sumErr = np.sum(abs(ads_exper[:,0] - J_theory))
            coef_fin = np.column_stack((se.val,sf.val,sg.val,sumErr))
            np.savetxt(f, coef_fin, delimiter='  ',header='e f g error for
%s sample' %(ns+1))
        f.close()
        plt.close()
    buttonpr.on_clicked(saveparuser)

    savex = plt.axes([0.35, 0.025, 0.12, 0.04])
    btn_save = Button(savex, 'Save param lsq',
color=axcolor,hovercolor='0.975')
```

```
    def savepar(event):
        with open('param_fit.dat','ab') as f:
            coef_fin = np.column_stack((EE,FF,GG,fitErr))
            np.savetxt(f, coef_fin, delimiter='  ',header='e f g error for
%s sample' %(ns+1))
        f.close()
        plt.close()
    btn_save.on_clicked(savepar)

    nextax = plt.axes([0.15, 0.025, 0.12, 0.04])
    btn_next = Button(nextax, 'Next sample',
color=axcolor,hovercolor='0.975')
    jump = False
    def nextsam(event):
        plt.close()
    btn_next.on_clicked(nextsam)

    closeax = plt.axes([0.025, 0.3, 0.12, 0.04])
    btn_close = Button(closeax, 'Close app', color=axcolor,
hovercolor='0.975')
    def closeapp(event):
        sys.exit()
    btn_close.on_clicked(closeapp)

    plt.show()

coef_fit = np.genfromtxt('param_fit.dat')
e=coef_fit[:,0]
f=coef_fit[:,1]
g=coef_fit[:,2]
#sumCoef=e+g
#c1 = e/sumCoef
#c2 = f/sumCoef
#c3 = g/sumCoef
dt = 60
fig, ax = plt.subplots(figsize=(12,10))
t = np.asarray([i*dt for i in range(e.shape[0])])

dat = np.column_stack((t,e,f,g))

plt.plot(t,e,label='fr Linear peptide -e',marker='s')
plt.plot(t,f,label='fr MAD -f',marker='p')
plt.plot(t,g,label='fr Cyclic peptide -g',marker='*')
plt.xlabel('t, s')
plt.ylabel('Coeff')
plt.legend()
plt.suptitle('by Cartier&Bugatti', fontsize=5, color='grey' )
savefax = plt.axes([0.75, 0.025, 0.12, 0.04])
btn_savef = Button(savefax, 'Save plot', color=axcolor,hovercolor='0.975')
def savepl(event):
    savefax.set_visible(False)
    plt.savefig('fraction.png')
    np.savetxt('conc.txt', dat, delimiter='  ',header='time c1 c2 c3')
```

```
#savefax.set_visible(False)
btn_savef.on_clicked(savepl)

plt.show()
```

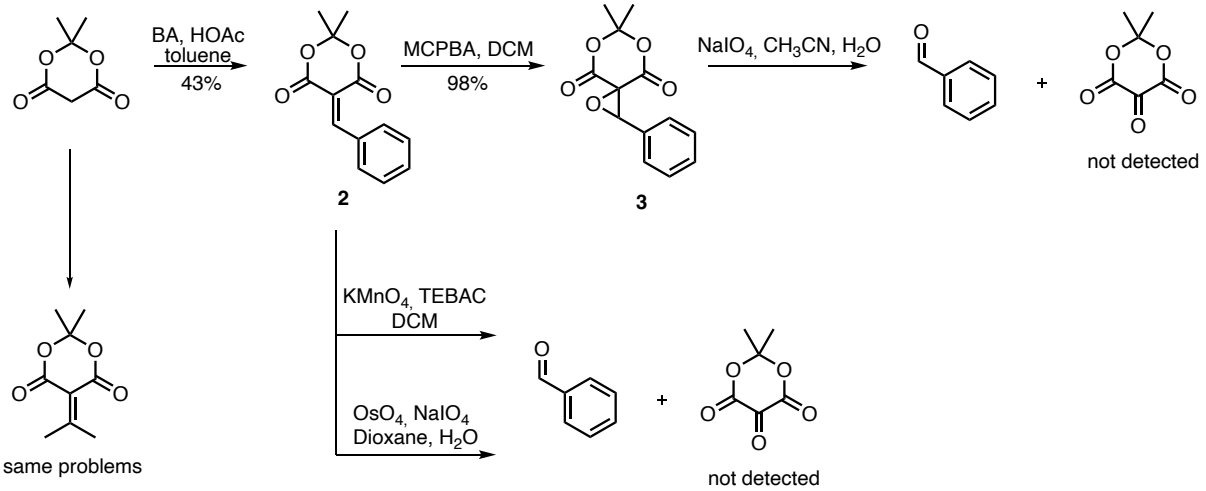
2.6.6 Compounds A and B as MAD alternatives

2.6.6.1 Synthesis of A and B

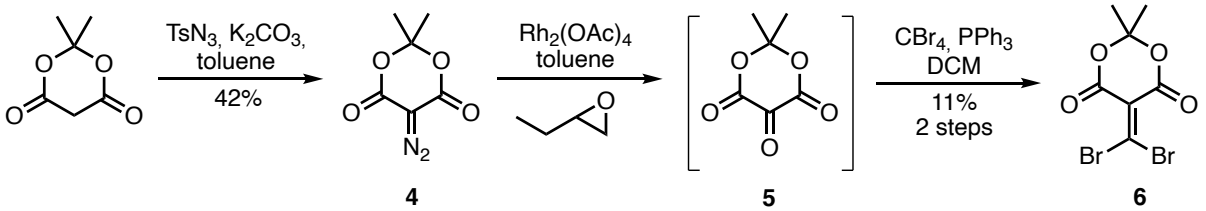


1

Synthesis of **A** (**1**) is described in literature: S. Akhil, A. Deepthi, *Tet. Lett.,* 55, 2014, 1890-1893.



The synthesis of **B** (MAD-Br, **6**) most conventionally can be performed starting from keto-Meldrum's acid derivative via Wittig reaction with CBr_4 and PPh_3. Unfortunately, the synthesis of keto-MA is not well described in the literature and usually required uncommon equipment e.g. ozone-generator. Therefore, we firstly tried the oxidation of alkene **2** using potassium permanganate or osmium tetraoxide, in this case the only detected product was benzaldehyde. We next tried to do oxidation of diols derived from epoxide **3** however the result was the same as in the previous attempt. We concluded that keto-MA, or possible reactive intermediates, might not be stable under these conditions. Next, we tried to perform direct oxidation of MA in ACN using sodium periodate. Although this method has been described for a large number of similar substrates in our attempt to oxidise MA we observed neither the product of oxidation nor its hydrated form. Finally, the suitable route appeared to be the reaction of diaso-MA **4** with epoxide.



4 **5** **6**

We were afraid that **5** might decompose during the work up or react with water forming hydrate therefore decided not to isolate it and proceeded to the next step. Conversion from **4** to **5** was monitored using NMR spectroscopy. In the final step, we performed a Wittig reaction to obtain **B** as a pale-yellow solid.

2.6.6.2 Some tests with A and B

2.6.6.2.1 Reaction trial with **A**

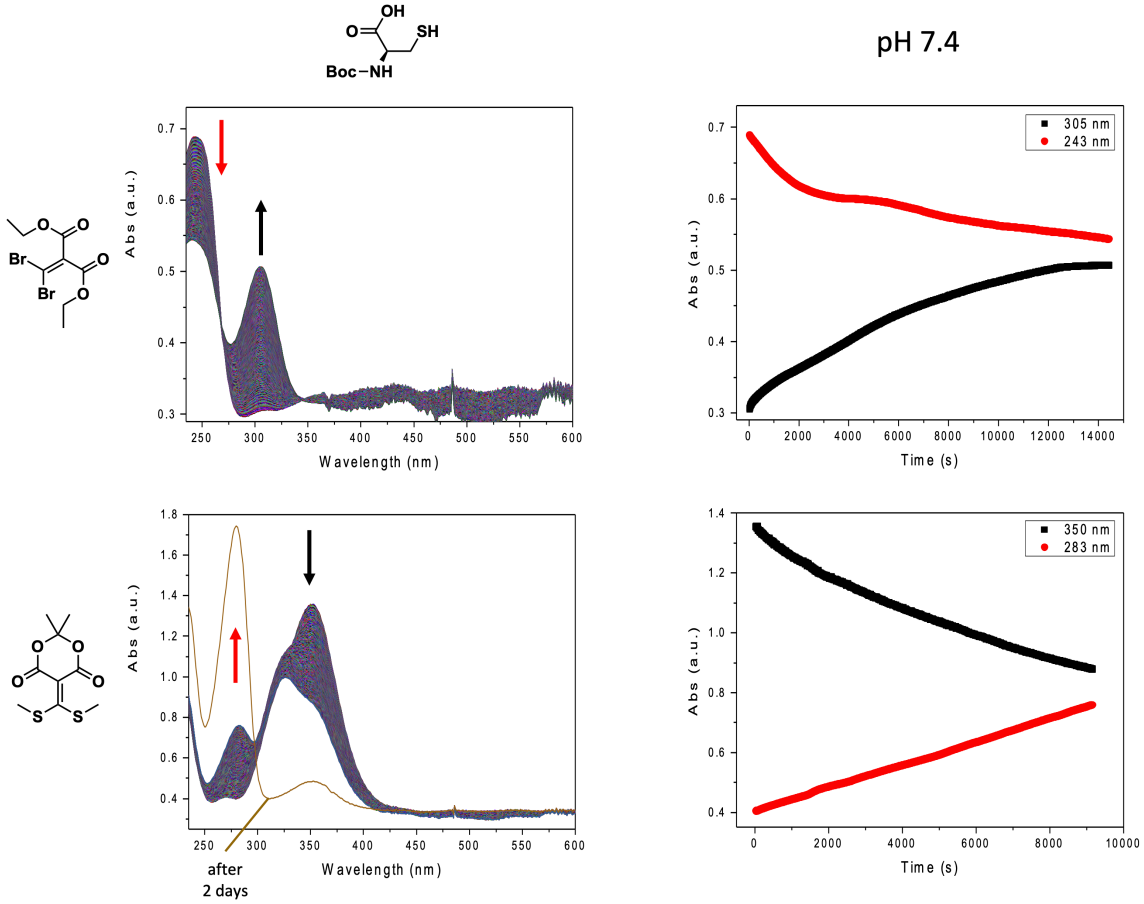


Figure 2.27 | UV-Vis kinetics comparison of the Boc-cysteine reaction with malonate-derived Michael acceptor (compound A) and MAD.

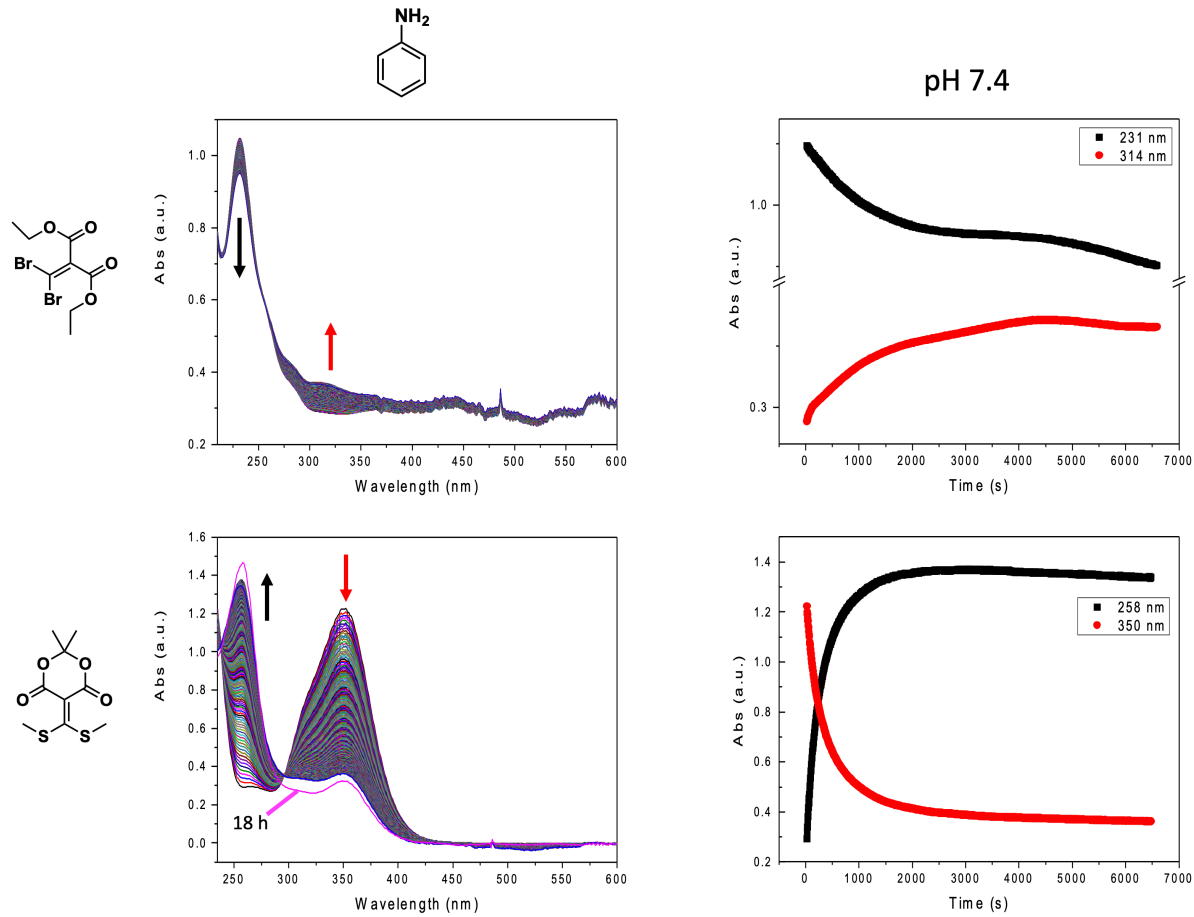
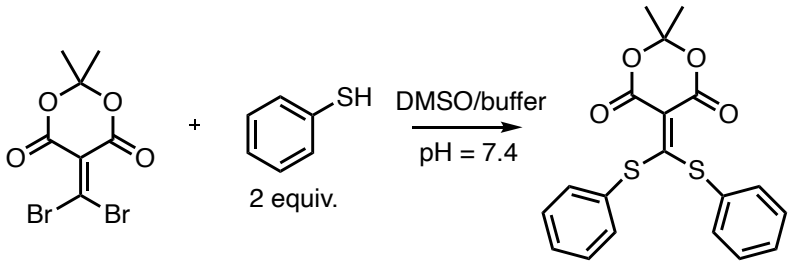


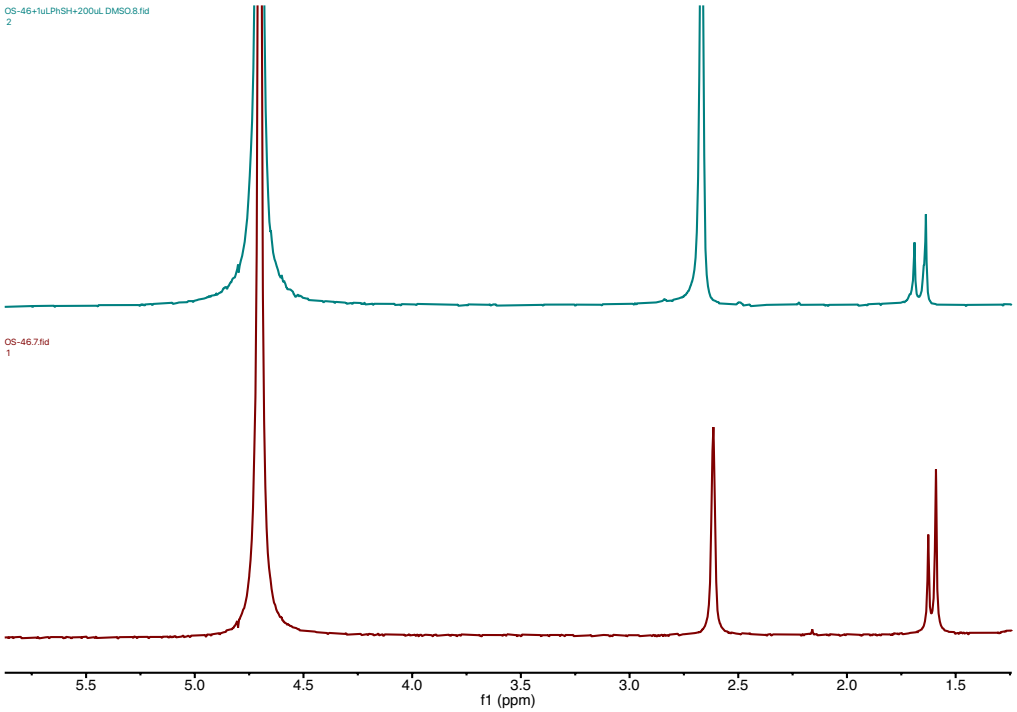
Figure 2.28 | UV-Vis kinetics comparison of the aniline reaction with malonate-derived Michael acceptor (compound A) and MAD.

2.6.6.2.2 Reaction trial with **B**

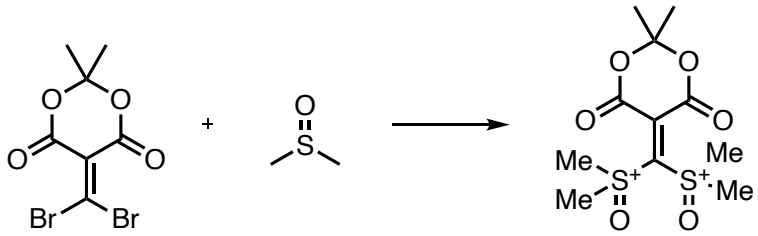
We studied by NMR the reaction of **B** with PhSH under typical conditions – DMSO/phosphate buffer (pH 7.4) 1:5 v/v – used to couple MAD with peptides.



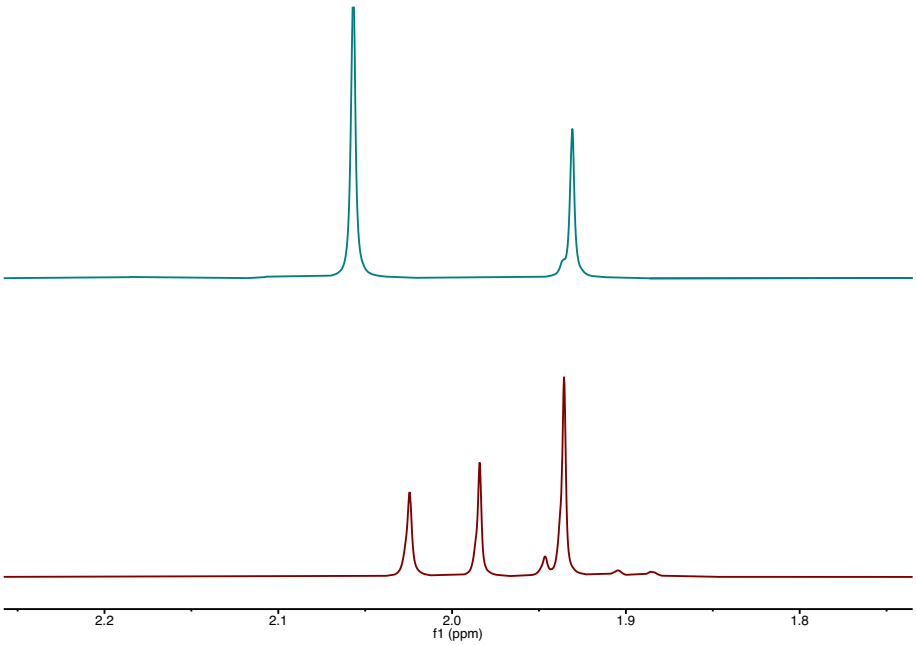
^{1}H NMR spectrum before and after addition on PhSH remained basically the same, under identical conditions MAD readily reacts with PhSH.



Possible explanation is that under these conditions something happens to **B.** Since it is known that DMSO can be relatively easily S-alkylated we assumed that **B** reacts with DMSO (or water) rendering inactive material.

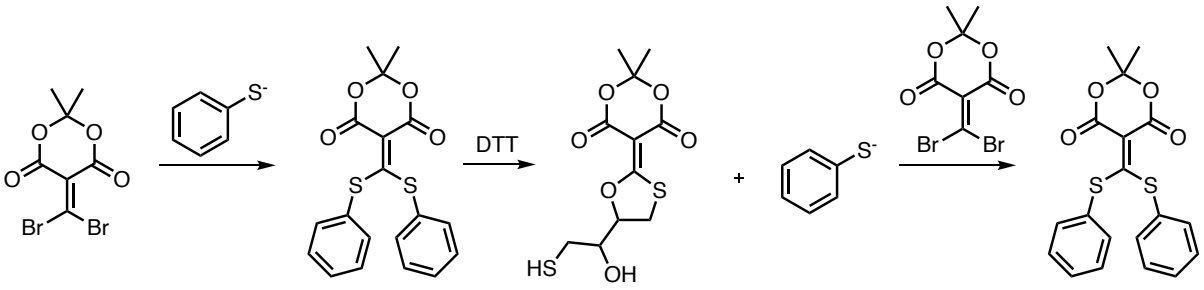


Similar experiment was carried out with mercaptoethanol, however this time with acetonitrile as a cosolvent. Interestingly in this case we observed dissapearence of a singlet at ca. 2.05 ppm and appearance of new singlets between 2.05 and 1.9 ppm when mercaptoethanol was added.



These results need to be studied in more detail.

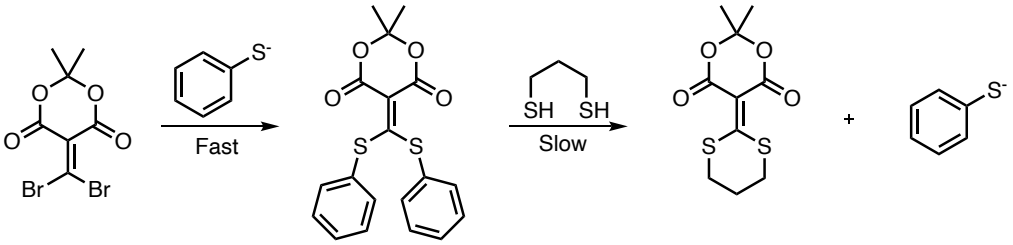
Alternatively, we decided to study reactivity of **B** in organic solvents – $CHCl_3$. Initial goal was to understand better chemical properties of **B**.



7 **8** **7**

To a $CDCl_3$ solution of **B** (1 equiv.) and TEA (4 equiv.) we added 2 equiv. of PhSH. **B** was almost completely consumed in ca. 10 minutes. To the obtained solution we added DTT. Again in less then 10 min. **7** was consumed and new singlets were formed that might correspond to two different adducts with DTT (5 and 7-membered rings). Next we added 1 more equivalent of **B**, again we observed the formation of **7,** however conversion was lower probably because part of TEA was protonated.

These results made us to wonder whether it is possible to use different acidities of aromatic and alyohatic thiols to achieve reaction cycles. The idea is that thiophenol, as more acidic species, in organic solvents in the presence of strong enough base we generate thiolate anion thus reacting faster with **B**.



2.7 References

1. Claro, B., Bastos, M. & Garcia-Fandino, R. Design and applications of cyclic peptides. in *Peptide Applications in Biomedicine, Biotechnology and Bioengineering* 87–129 (Elsevier, 2018). doi:10.1016/B978-0-08-100736-5.00004-1.
2. Daly, N., Wilson, D. & Craik, D. Cyclic peptides from plants and their promise in drug design. *Australian Biochemist* **43**, 7–9 (2012).
3. Gause, G. F. & Brazhnikova, M. G. Gramicidin S and its use in the treatment of infected wounds [3]. *Nature* vol. 154 703 Preprint at https://doi.org/10.1038/154703a0 (1944).
4. Granier, S. *et al.* A cyclic peptide mimicking the third intracellular loop of the V 2 vasopressin receptor inhibits signaling through its interaction with receptor dimer and G protein. *Journal of Biological Chemistry* **279**, 50904–50914 (2004).
5. Jing, X. & Jin, K. A gold mine for drug discovery: Strategies to develop cyclic peptides into therapies. *Med Res Rev* **40**, 753–810 (2020).
6. Shimizu, T., Ding, W. & Kameta, N. Soft-Matter Nanotubes: A Platform for Diverse Functions and Applications. *Chem. Rev.* **120**, 2347–2407 (2020).
7. Bajaj, K. Natural Bioactive Cyclic Peptides and Peptidomimetics. in *Studies in Natural Products Chemistry* vol. 62 343–376 (Elsevier, 2019).
8. Ramakrishnan, C., Paul, P. K. C. & Ramnarayan, K. Cyclic peptides - Small and big and their conformational aspects. *Journal of Biosciences* **8**, 239–251 (1985).
9. Roxin, Á. & Zheng, G. Flexible or fixed: A comparative review of linear and cyclic cancer-targeting peptides. *Future Medicinal Chemistry* vol. 4 1601–1618 Preprint at https://doi.org/10.4155/fmc.12.75 (2012).
10. Baeriswyl, V. & Heinis, C. Polycyclic Peptide Therapeutics. *ChemMedChem* **8**, 377–384 (2013).
11. Chen, S., Morales-Sanfrutos, J., Angelini, A., Cutting, B. & Heinis, C. Structurally Diverse Cyclisation Linkers Impose Different Backbone Conformations in Bicyclic Peptides. *ChemBioChem* **13**, 1032–1038 (2012).
12. Oppewal, T. R., Hekelaar, J. & Mayer, C. A phage-compatible strategy to access macrocyclic peptides featuring asymmetric molecular scaffolds as cyclization units. 43.
13. Liu, M. *et al.* Cyclization of several linear penta- and heptapeptides with different metal ions studied by CD spectroscopy*: Cyclization of peptides with metal ions. *The Journal of Peptide Research* **65**, 55–64 (2008).
14. Rivera, D. G., Ojeda-Carralero, G. M., Reguera, L. & Van der Eycken, E. V. Peptide macrocyclization by transition metal catalysis. *Chem. Soc. Rev.* **49**, 2039–2059 (2020).
15. Li, B. *et al.* Cooperative Stapling of Native Peptides at Lysine and Tyrosine or Arginine with Formaldehyde. *Angew. Chem. Int. Ed.* **60**, 6646–6652 (2021).
16. Bechtler, C. & Lamers, C. Macrocyclization strategies for cyclic peptides and peptidomimetics. *RSC Med. Chem.* **12**, 1325–1351 (2021).
17. White, C. J. & Yudin, A. K. Contemporary strategies for peptide macrocyclization. *Nature Chemistry* vol. 3 509–524 Preprint at https://doi.org/10.1038/nchem.1062 (2011).
18. Jiang, S., Li, Z., Ding, K. & Roller, P. Recent Progress of Synthetic Studies to Peptide and Peptidomimetic Cyclization. *Current Organic Chemistry* **12**, 1502–1542 (2008).
19. Gentilucci, L., De Marco, R. & Cerisoli, L. Chemical Modifications Designed to Improve Peptide Stability: Incorporation of Non-Natural Amino Acids, Pseudo-Peptide Bonds, and Cyclization. *Current Pharmaceutical Design* **16**, 3185–3203 (2010).
20. Wieland, T., Faulstich, H. & Fiume, L. Amatoxins, phallotoxins, phallolysin, and antamanide: The biologically active components of poisonous amanita mushroom. *Critical Reviews in Biochemistry and Molecular Biology* **5**, 185–260 (1978).
21. Kumaraswamy, P., Lakshmanan, R., Sethuraman, S. & Krishnan, U. M. Self-assembly of peptides: Influence of substrate, pH and medium on the formation of supramolecular assemblies. *Soft Matter* **7**, 2744–2754 (2011).
22. Sun, L. *et al.* Tunable synthesis of self-assembled cyclic peptide nanotubes and nanoparticles. *Soft Matter* **11**, 3822–3832 (2015).
23. Ghadiri, M. R., Granja, J. R., Milligan, R. A., McRee, D. E. & Khazanovich, N. Self-assembling organic nanotubes based on a cyclic peptide architecture. *Nature* **366**, 324–327 (1993).
24. Hartgerink, J. D., Granja, J. R., Milligan, R. A. & Ghadiri, M. R. Self-assembling peptide nanotubes. *Journal of the American Chemical Society* **118**, 43–50 (1996).

25. Song, Q. *et al.* Molecular Self-Assembly and Supramolecular Chemistry of Cyclic Peptides. *Chem. Rev.* **121**, 13936–13995 (2021).
26. Rubin, D. J. *et al.* Mechanical Reinforcement of Polymeric Fibers through Peptide Nanotube Incorporation. *Biomacromolecules* **14**, 3370–3375 (2013).
27. Shaikh, H. *et al.* Hydrogel and Organogel Formation by Hierarchical Self-Assembly of Cyclic Peptides Nanotubes. *Chemistry A European J* **24**, 19066–19074 (2018).
28. Méndez-Ardoy, A., Granja, J. R. & Montenegro, J. pH-Triggered self-assembly and hydrogelation of cyclic peptide nanotubes confined in water micro-droplets. *Nanoscale Horiz.* **3**, 391–396 (2018).
29. Méndez-Ardoy, A. *et al.* Spatially Controlled Supramolecular Polymerization of Peptide Nanotubes by Microfluidics. *Angewandte Chemie International Edition* **59**, 6902–6908 (2020).
30. Hsieh, W. H. & Liaw, J. Applications of cyclic peptide nanotubes (cPNTs). *Journal of Food and Drug Analysis* **27**, 32–47 (2019).
31. Claro, B., Bastos, M. & Garcia-Fandino, R. *Design and applications of cyclic peptides*. *Peptide Applications in Biomedicine, Biotechnology and Bioengineering* (Elsevier Ltd, 2018). doi:10.1016/B978-0-08-100736-5.00004-1.
32. Chapman, R., Danial, M., Koh, M. L., Jolliffe, K. A. & Perrier, S. Design and properties of functional nanotubes from the self-assembly of cyclic peptide templates. *Chemical Society Reviews* vol. 41 6023–6041 Preprint at https://doi.org/10.1039/c2cs35172b (2012).
33. Kobayashi, K., Granja, J. R. & Ghadiri, M. R. β-Sheet Peptide Architecture: Measuring the Relative Stability of Parallel vs. Antiparallel β-Sheets. *Angewandte Chemie International Edition in English* **34**, 95–98 (1995).
34. Rosenthal-Aizman, K., Svensson, G. & Undén, A. Self-Assembling Peptide Nanotubes from Enantiomeric Pairs of Cyclic Peptides with Alternating d and l Amino Acid Residues. *J. Am. Chem. Soc.* **126**, 3372–3373 (2004).
35. Insua, I. & Montenegro, J. 1D to 2D Self Assembly of Cyclic Peptides. *J. Am. Chem. Soc.* **142**, 300–307 (2020).
36. Díaz, S., Insua, I., Bhak, G. & Montenegro, J. Sequence Decoding of 1D to 2D Self-Assembling Cyclic Peptides. *Chemistry – A European Journal* **26**, 14765–14770 (2020).
37. Seebach, D. *et al.* Cyclo-beta-peptides: Structure and tubular stacking of cyclic tetramers of 3-aminobutanoic acid as determined from powder diffraction data. *Helvetica Chimica Acta* **80**, 173–182 (1997).
38. Brea, R. J., Castedo, L. & Granja, J. R. Large-diameter self-assembled dimers of α,γ-cyclic peptides, with the nanotubular solid-state structure of cyclo-[(L-Leu-D- MeN-γ-Acp)4-]·4CHCl2COOH. *Chemical Communications* 3267–3269 (2007) doi:10.1039/b703659k.
39. Amorín, M., Castedo, L. & Granja, J. R. New cyclic peptide assemblies with hydrophobic cavities: The structural and thermodynamic basis of a new class of peptide nanotubes. *Journal of the American Chemical Society* **125**, 2844–2845 (2003).
40. Bélanger, D., Tong, X., Soumaré, S., Dory, Y. L. & Zhao, Y. Cyclic Peptide–Polymer Complexes and Their Self-Assembly. *Chemistry – A European Journal* **15**, 4428–4436 (2009).
41. Leclair, S. *et al.* Micrometer-Sized Hexagonal Tubes Self-Assembled by a Cyclic Peptide in a Liquid Crystal. *Angewandte Chemie International Edition* **43**, 349–353 (2004).
42. Clark, T. D. *et al.* Cylindrical β-Sheet Peptide Assemblies. *J. Am. Chem. Soc.* **120**, 8949–8962 (1998).
43. Clark, R. J. *et al.* The Engineering of an Orally Active Conotoxin for the Treatment of Neuropathic Pain. *Angewandte Chemie International Edition* **49**, 6545–6548 (2010).
44. Cissé, N. & Kudernac, T. Light-Fuelled Self-Assembly of Cyclic Peptides into Supramolecular Tubules. *ChemSystemsChem* **2**, (2020).
45. Mather, B. D., Viswanathan, K., Miller, K. M. & Long, T. E. Michael addition reactions in macromolecular design for emerging technologies. *Progress in Polymer Science* **31**, 487–531 (2006).
46. Conjugate Additions. https://www2.chem.wisc.edu/areas/reich/chem842/_chem842-02-orgli%7B13%7D.htm.
47. Michael Addition. https://www.organic-chemistry.org/namedreactions/michael-addition.shtm.
48. Peyrton, J. & Avérous, L. Aza-Michael Reaction as a Greener, Safer, and More Sustainable Approach to Biobased Polyurethane Thermosets. *ACS Sustainable Chem. Eng.* **9**, 4872–4884 (2021).
49. Noordzij, G. J. & Wilsens, C. H. R. M. Cascade aza-Michael Addition-Cyclizations; Toward Renewable and Multifunctional Carboxylic Acids for Melt-Polycondensation. *Frontiers in Chemistry* **7**, (2019).
50. Noordzij, G. J. *et al.* The aza-Michael reaction: towards semi-crystalline polymers from renewable itaconic acid and diamines. *Polym. Chem.* **10**, 4049–4058 (2019).

51. Genest, A., Portinha, D., Fleury, E. & Ganachaud, F. The aza-Michael reaction as an alternative strategy to generate advanced silicon-based (macro)molecules and materials. *Progress in Polymer Science* **72**, 61–110 (2017).
52. Nair, D. P. *et al.* The Thiol-Michael Addition Click Reaction: A Powerful and Widely Used Tool in Materials Chemistry. *Chem. Mater.* **26**, 724–744 (2014).
53. Nair, D. P. *et al.* The Thiol-Michael addition click reaction: A powerful and widely used tool in materials chemistry. *Chemistry of Materials* vol. 26 724–744 Preprint at https://doi.org/10.1021/cm402180t (2014).
54. Folikumah, M. Y., Neffe, A. T., Behl, M. & Lendlein, A. Thiol Michael-Type Reactions of Optically Active Mercapto-Acids in Aqueous Medium. *MRS Advances* **4**, 2515–2525 (2019).
55. Reuther, J. F., Dahlhauser, S. D. & Anslyn, E. V. Tunable Orthogonal Reversible Covalent (TORC) Bonds: Dynamic Chemical Control over Molecular Assembly. *Angewandte Chemie International Edition* **58**, 74–85 (2019).
56. Shi, B. & Greaney, M. F. Reversible Michael addition of thiols as a new tool for dynamic combinatorial chemistry. *Chem. Commun.* 886–888 (2005) doi:10.1039/B414300K.
57. Dynamic Combinatorial Chemistry | Chemical Reviews. https://pubs.acs.org/doi/full/10.1021/cr020452p?casa_token=tY9jUMIEsDAAAAAA:aF57XZ3xkOuPrtihZ9-sCCwcg_Nqnvk7DSnSoKK2ZTBIzneHdN9y_U8AEKCMNP67ddtKA24xbT1lQBwb.
58. Reymond, J.-L. 7.2 Screening Methods for Enzymes. in *Comprehensive Chirality* (eds. Carreira, E. M. & Yamamoto, H.) 6–20 (Elsevier, 2012). doi:10.1016/B978-0-08-095167-6.00702-3.
59. Mulliner, D., Wondrousch, D. & Schüürmann, G. Predicting Michael-acceptor reactivity and toxicity through quantum chemical transition-state calculations. *Org. Biomol. Chem.* **9**, 8400–8412 (2011).
60. Maucher, I. V. *et al.* Michael acceptor containing drugs are a novel class of 5-lipoxygenase inhibitor targeting the surface cysteines C416 and C418. *Biochemical Pharmacology* **125**, 55–74 (2017).
61. Zheng, S. *et al.* Synthesis, Chemical Reactivity as Michael Acceptors, and Biological Potency of Monocyclic Cyanoenones, Novel and Highly Potent Anti-inflammatory and Cytoprotective Agents. *J. Med. Chem.* **55**, 4837–4846 (2012).
62. Santos, M. M. M. & Moreira, R. Michael Acceptors as Cysteine Protease Inhibitors. *Mini-Reviews in Medicinal Chemistry* **7**, 1040–1050.
63. Cabello, C. M. *et al.* The Cinnamon-derived Michael Acceptor Cinnamic Aldehyde Impairs Melanoma Cell Proliferation, Invasiveness, and Tumor Growth. *Free Radic Biol Med* **46**, 220–231 (2009).
64. Hearn, B. R. *et al.* Attenuation of the Reaction of Michael Acceptors with Biologically Important Nucleophiles. *Bioconjugate Chem.* **32**, 794–800 (2021).
65. Huang, S. *et al.* Determining Michael acceptor reactivity from kinetic, mechanistic, and computational analysis for the base-catalyzed thiol-Michael reaction. *Polym. Chem.* **12**, 3619–3628 (2021).
66. Zhuang, J. *et al.* A programmable chemical switch based on triggerable Michael acceptors. *Chem. Sci.* **11**, 2103–2111 (2020).
67. Diehl, K. L. *et al.* Click and chemically triggered declick reactions through reversible amine and thiol coupling via a conjugate acceptor. *Nat. Chem.* **8**, 968–973 (2016).
68. Fulton, D. A. Click chemistry gets reversible. *Nature Chem* **8**, 899–900 (2016).
69. Bielski, R. & Witczak, Z. Strategies for Coupling Molecular Units if Subsequent Decoupling Is Required. *Chem. Rev.* **113**, 2205–2243 (2013).
70. Boul, P. J., Reutenauer, P. & Lehn, J. M. Reversible Diels-Alder reactions for the generation of dynamic combinatorial libraries. *Organic Letters* **7**, 15–18 (2005).
71. Gandini, A., Silvestre, A. & Coelho, D. Reversible click chemistry at the service of macromolecular materials. Part 4: Diels-Alder non-linear polycondensations involving polyfunctional furan and maleimide monomers. *Polymer Chemistry* **4**, 1364–1371 (2013).
72. Krenske, E. H., Petter, R. C. & Houk, K. N. Kinetics and Thermodynamics of Reversible Thiol Additions to Mono- and Diactivated Michael Acceptors: Implications for the Design of Drugs That Bind Covalently to Cysteines. *J. Org. Chem.* **81**, 11726–11733 (2016).
73. Krishnan, S. *et al.* Design of Reversible, Cysteine-Targeted Michael Acceptors Guided by Kinetic and Computational Analysis. *J. Am. Chem. Soc.* **136**, 12624–12630 (2014).
74. Wu, T. *et al.* Chemically Triggered Click and Declick Reactions: Application in Synthesis and Degradation of Thermosetting Plastics. *ACS Macro Lett.* **10**, 1125–1131 (2021).
75. Diehl, K. L. *et al.* Click and chemically triggered declick reactions through reversible amine and thiol coupling via a conjugate acceptor. *Nature Chemistry* **8**, 968–973 (2016).

76. Zhong, Y., Xu, Y. & Anslyn, E. V. Studies of Reversible Conjugate Additions. *European Journal of Organic Chemistry* **2013**, 5017–5021 (2013).
77. Ishibashi, J. S. A. & Kalow, J. A. Vitrimeric Silicone Elastomers Enabled by Dynamic Meldrum's Acid-Derived Cross-Links. *ACS Macro Lett.* **7**, 482–486 (2018).
78. Sun, X. & Anslyn, E. V. An Auto-Inductive Cascade for the Optical Sensing of Thiols in Aqueous Media: Application in the Detection of a VX Nerve Agent Mimic. *Angew. Chem. Int. Ed.* **56**, 9522–9526 (2017).
79. Davis, G. J., Sofka, H. A. & Jewett, J. C. Highly Stable Meldrum's Acid Derivatives for Irreversible Aqueous Covalent Modification of Amines. *Org. Lett.* **22**, 2626–2629 (2020).
80. Lee, D.-H. *et al.* A self-degradable hydrogel sensor for a nerve agent tabun surrogate through a self-propagating cascade. *Cell Reports Physical Science* **2**, 100552 (2021).
81. Zhu, F. *et al.* Determination of biogenic amines in alcoholic beverages using a novel fluorogenic compound as derivatizing reagent. *RSC Advances* **11**, 19541–19550 (2021).
82. Sun, X. *et al.* Chemically Triggered Synthesis, Remodeling, and Degradation of Soft Materials. *J. Am. Chem. Soc.* **142**, 3913–3922 (2020).
83. Johnson, A. M. & Anslyn, E. V. Reversible Macrocyclization of Peptides with a Conjugate Acceptor. *Organic Letters* **19**, 1654–1657 (2017).
84. Sun, X. *et al.* Chemically Triggered Synthesis, Remodeling, and Degradation of Soft Materials. *Journal of the American Chemical Society* **142**, 3913–3922 (2020).
85. Ben Cheikh, A. *et al.* Synthesis of .alpha.-cyano carbonyl compounds by flash vacuum thermolysis of (alkylamino)methylene derivatives of Meldrum's acid. Evidence for facile 1,3-shifts of alkylamino and alkylthio groups in imidoylketene intermediates. *J. Org. Chem.* **56**, 970–975 (1991).
86. Hermans, T. M. Materials from a peptide soup. *Nature Nanotech* **11**, 920–921 (2016).
87. Debnath, S., Roy, S. & Ulijn, R. V. Peptide Nanofibers with Dynamic Instability through Nonequilibrium Biocatalytic Assembly. *J. Am. Chem. Soc.* **135**, 16789–16792 (2013).
88. Hunter, C. A. & Anderson, H. L. What is Cooperativity? *Angew. Chem. Int. Ed.* **48**, 7488–7499 (2009).
89. Di Stefano, S. & Ercolani, G. Equilibrium Effective Molarity As a Key Concept in Ring-Chain Equilibria, Dynamic Combinatorial Chemistry, Cooperativity and Self-assembly. in *Advances in Physical Organic Chemistry* vol. 50 1–76 (Elsevier, 2016).
90. Kirby, A. J. Effective Molarities for Intramolecular Reactions. in *Advances in Physical Organic Chemistry* vol. 17 183–278 (Elsevier, 1980).
91. Ahmed, Md. I., Harper, J. B. & Hunter, L. Incrementally increasing the length of a peptide backbone: effect on macrocyclisation efficiency. *Org. Biomol. Chem.* **12**, 4598 (2014).
92. Moore, J. W. & Pearson, R. G. *Kinetics and Mechanism*. (John Wiley & Sons, 1981).
93. Sorrenti, A., Leira-Iglesias, J., Sato, A. & Hermans, T. M. Non-equilibrium steady states in supramolecular polymerization. *Nat Commun* **8**, 15899 (2017).
94. van der Helm, M. P., de Beun, T. & Eelkema, R. On the use of catalysis to bias reaction pathways in out-of-equilibrium systems. *Chem. Sci.* **12**, 4484–4493 (2021).
95. Tena-Solsona, M. *et al.* Non-equilibrium dissipative supramolecular materials with a tunable lifetime. *Nat Commun* **8**, 15895 (2017).
96. Leira-Iglesias, J., Tassoni, A., Adachi, T., Stich, M. & Hermans, T. M. Oscillations, travelling fronts and patterns in a supramolecular system. *Nat. Nanotechnol.* **13**, 1021–1027 (2018).
97. Christopoulos, A. & Lew, M. J. Beyond Eyeballing: Fitting Models to Experimental Data. *Critical Reviews in Biochemistry and Molecular Biology* **35**, 359–391 (2000).
98. Meadows, M. K. *et al.* Mechanistic studies of a "Declick" reaction. *Chem. Sci.* **10**, 8817–8824 (2019).
99. Anslyn, E. V. & Dougherty, D. A. *Modern Physical Organic Chemistry*. (University Science Books, 2006).
100. Bijlsma, S., Boelens, H. F. M. & Smilde, A. K. Determination of Rate Constants in Second-Order Kinetics Using UV-Visible Spectroscopy. *Appl Spectrosc* **55**, 77–83 (2001).
101. Kałka, A. J. & Turek, A. M. Do Spectra Live in the Matrix? A Brief Tutorial on Applications of Factor Analysis to Resolving Spectral Datasets of Mixtures. *J Fluoresc* **31**, 1599–1616 (2021).
102. de Juan, A. & Tauler, R. Multivariate Curve Resolution-Alternating Least Squares for Spectroscopic Data. in *Data Handling in Science and Technology* vol. 30 5–51 (Elsevier, 2016).
103. Garrido, M., Rius, F. X. & Larrechi, M. S. Multivariate curve resolution–alternating least squares (MCR-ALS) applied to spectroscopic data from monitoring chemical reactions processes. *Anal Bioanal Chem* **390**, 2059–2066 (2008).

104. Frans, S. D. & Harris, J. M. Reiterative least-squares spectral resolution of organic acid/base mixtures. *Anal. Chem.* **56**, 466–470 (1984).
105. Singh, N., Lainer, B., Formon, G. J. M., De Piccoli, S. & Hermans, T. M. Re-programming Hydrogel Properties Using a Fuel-Driven Reaction Cycle. *J. Am. Chem. Soc.* **142**, 4083–4087 (2020).
106. De Greef, T. F. A. *et al.* Supramolecular Polymerization. *Chem. Rev.* **109**, 5687–5754 (2009).
107. Korevaar, P. A., Schaefer, C., de Greef, T. F. A. & Meijer, E. W. Controlling Chemical Self-Assembly by Solvent-Dependent Dynamics. *J. Am. Chem. Soc.* **134**, 13482–13491 (2012).
108. Cantekin, S., Balkenende, D. W. R., Smulders, M. M. J., Palmans, A. R. A. & Meijer, E. W. The effect of isotopic substitution on the chirality of a self-assembled helix. *Nature Chem* **3**, 42–46 (2011).
109. Harada, N. & Berova, N. 8.24 Spectroscopic Analysis: Exciton Circular Dichroism for Chiral Analysis. in *Comprehensive Chirality* (eds. Carreira, E. M. & Yamamoto, H.) 449–477 (Elsevier, 2012). doi:10.1016/B978-0-08-095167-6.00846-6.
110. Xian Huang & Chen Bang-Chi. Synthesis of bisalkylthiolydine derivatives of Meldrum's acid and barbituric acid. *Synthesis (Stuttg.)* 967–968 (1986).
111. Shank, N. E. Constant time interval method for analysis of reaction rate data. *Int. J. Chem. Kinet.* **5**, 577–582 (1973).
112. Zhou, Y. *et al.* Indoline Catalyzed Acylhydrazone/Oxime Condensation under Neutral Aqueous Conditions. *Org. Lett.* **22**, 6035–6040 (2020).
113. Tobey, S. W. Determining second order rate constants. *J. Chem. Educ.* **39**, 473 (1962).

Chapter 3. Decyclization and disassembly of CP

Abstract

So far, in the literature, CPs disassembly and decyclization have been studied separately and were done using different reactions or stimuli. Ring-opening reactions of CP are often irreversible, which limits their further use in materials. The reaction of MAD with cysteine-containing peptides leads to their cyclization, which can be reversed by adding DTT. The recovered peptide can go through the same stepwise process repeatedly. Here, we study the decyclization reaction on **P1** and **P2** systems, determine suitable decyclization agents, and discuss the possibilities of combining cyclization and decyclization processes into a reaction cycle.

3.1 Recovery of LP from CP and CPNs

3.1.1 Non-covalent disassembly of CPNs

As was mentioned in the previous chapters, there are very few examples of CPN disassembly. Most involve modification of the peptide backbone by adding pH-sensitive groups or by sterically manipulating the side chain environment of peptides to disfavor interactions between the macrocycles. In the first case, depending on the protonation state of those functional groups, attractive and repulsive forces between macrocycles are tuned to yield mostly assembled or molecularly dissolved CPs. In work from Kudernac and co-workers,[1] CPs tailored with glutamic acid residues form tubular assemblies at acidic pH and disassemble under basic conditions. The medium's acidity is controlled using a photoacid that can switch between its merocyanine (dark) and spiropyran (light) forms. Upon irradiation with light, conversion to the spiropyran form releases a proton and decreases the pH, which triggers CPNs formation once pH is below 5. The same principle for CPNs self-assembly control was also explored earlier.[2] There is also an inverse example from Perrier and co-workers,[3,4] where CPNs exist in basic medium and are fully disassembled at acidic pH. In a second example by Perrier and co-workers,[5] the influence of bulky substituents on the peptide self-assembly is illustrated. A host-guest interaction between the phenylalanine of the CP's side chain and cucurbit[7]uril results in an increase in molecular size that disrupts the intermolecular H-bonds of the nanotube and prevents efficient intermolecular interaction. There is also an example of photoswitchable cyclic peptide reorientation. Ghadiri and co-workers developed a system[6] containing two cyclic peptide molecules connected via an azo-benzene unit. Irradiation with UV light causes E→Z isomerization, which breaks intermolecular assemblies into intramolecular single-molecule units.

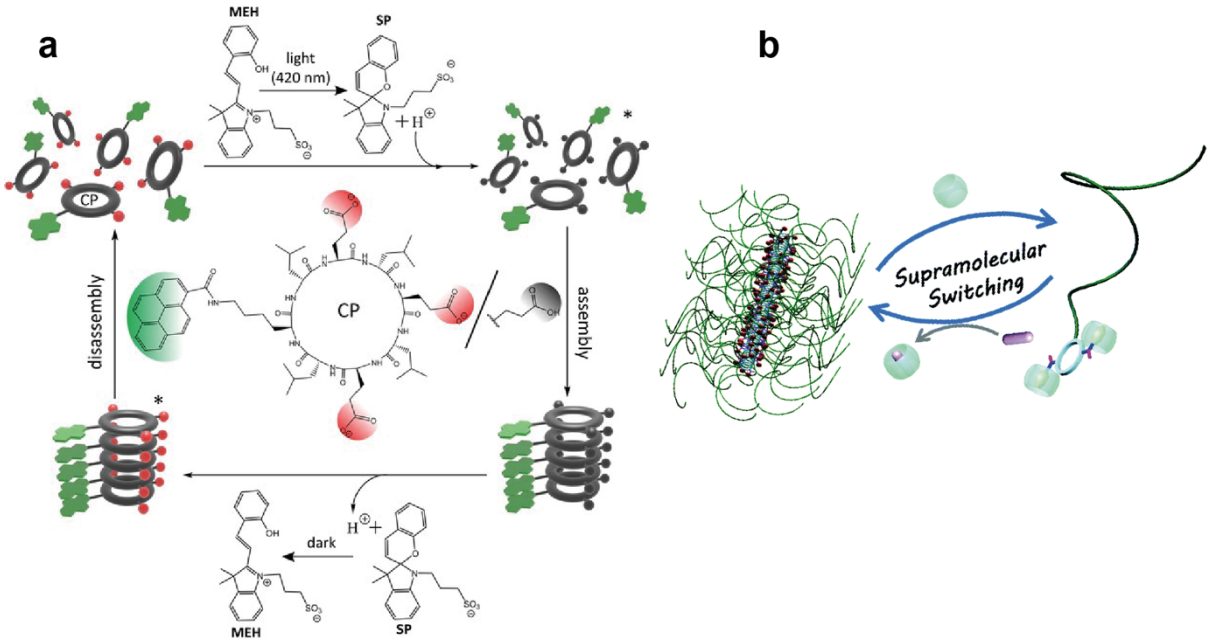


Figure 3.1 | Selected literature examples illustrating control over CPs (dis)assembly. (a) Control over electrostatic interactions between macrocycles realized via light-responsive pH switch. Reprinted from ref. 1. (b) Complexation of CPs' side chains by bulky CB[7] molecules induces CPs disassembly. The effect is reversible once a better guest for CB[7] is introduced. Adapted from ref. 5.

3.1.2 Covalent decyclization to induce CPN disassembly

Compared to the disassembly examples in the previous section, the number of covalent ring-opening techniques to destroy the CP motif is higher. Usually, it is studied separately from the disassembly of CPNs, as the motivation is different. Ring-opening strategies are used primarily to study the peptide sequence of natural products. Sequencing cyclic peptides to determine their structure is a tedious task, as regular methods, such as Edman degradation or tandem mass spectrometry, do not work well and require additional adjustment.[7] Therefore, various techniques have been developed to tackle this problem. They focus mainly on tailoring the structure of CP to enable the bond cleavage at a specific position in the ring when the trigger is applied (e.g., UV[8], cyanogen bromide[7,9,10], oxidant[11]). This often requires substantial synthetic procedures and modification of the peptide itself, which might limit its applications. Metabolic degradation of CPs has also been studied.[12] A common feature of these approaches is that the linear peptide cannot be recovered afterward, i.e., the methods are destructive.

To this end, the disassembly of CPNs and decomposition of CPs have not been studied as parts of the continuous process occurring in the context of the same reaction processes.

3.1.2 Deprotection and decoupling. MAD reaction's reversibility

Preparative organic synthesis focuses on straightforward, selective, fast, and robust ways to yield one product quantitatively. When it comes to cleaving bonds, one would usually expect protecting groups to be stable in a wide range of conditions which implies that the environment for its removal is very specific or even harsh if we think in terms of biocompatible chemistry.

The need for 'weaker' or more labile protecting groups appeared with interest in the time- or site-specific release of drugs that would require a different approach.[13,14] The more chemistry moves toward biology, the higher the demand for non-destructive, benign methods and techniques.

Although many ways exist to 'click' molecules together covalently, there are barely any options to cleave or 'de-click' bonds selectively and orthogonally in mild conditions. Such reactions are especially useful for bioconjugation studies because they allow creation of labile labels (radioactive, fluorescent) and studying biospecific interactions of receptors and ligands, especially when the binding molecule is unknown/uncharacterized, not to mention a wide application in material science.[15–19]

Depending on the purpose, one can distinguish between deprotection and decoupling. In deprotection, the original molecules are recovered after the crosslinker is removed. In decoupling, there is a 'sacrificial unit'—an extra functionality that, upon breaking (decoupling), will not necessarily recover the initial substrate.[20–22] Few linkers can do deprotection instead of decoupling in water-based systems at physiological pH, and even fewer are bio-orthogonal. Finding molecules that can, under mild conditions, selectively crosslink and then on-demand quantitatively release initial substrates unchanged is a desirable asset for bio-organic chemistry, biology, and now, systems chemistry.

To this end, the most famous types of cleavable bonds that work in water-based systems are disulfides and diazo linkers (cleavable via reduction), glycol groups (cleavable via oxidation), ester and sulfone groups (base-cleavable), acylhydrazones, carbamates and acetals (acid-labile), nitrophenyl groups (photocleavable).[15,23]

The Michael acceptor developed by Anslyn and co-workers,[24] MAD, is unique not just because of its selectivity but also because it can release the coupled amine and thiol unaltered—deprotect. Although not biorthogonal, it can link bio-relevant molecules and be used in bioassays *in vitro*.[24] The most efficient deprotection agents suggested by the authors are DL-dithiothreitol (DTT), L-cysteine (Cys), and ethanedithiol, which release amines and thiols from MAD-nucleophiles adducts efficiently (Figure 3.2). The researchers also tried other compounds. Namely, propanedithiol and cysteamine were also effective in liberating both amine and thiol, but the reaction proceeded slower than with the best decoupling agents. Thioglycolic acid (TGA) and β-mercaptoethanol (BME) only liberated thiol, but not amine, whereas ethylenediamine did not work. The authors concluded that the alcohol of BME and carboxylate of TGA are not sufficiently nucleophilic to react with the MAD-amine adduct.[24] The fact that BME did not perform similarly to DTT was to the authors' surprise as structurally, BME is half of DTT. It led them to the hypothesis that DTT reacts with MAD, first with its two thiols forming a seven-membered ring, which then rearranges to a more stable five-membered ring involving one of the alcohol groups (Figure 3.2, bottom left pathway).

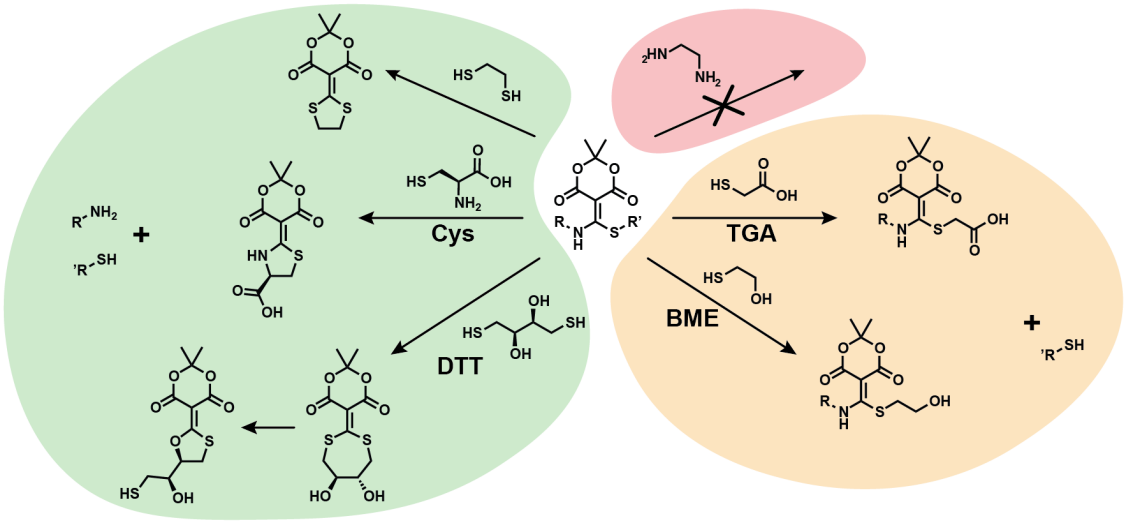


Figure 3.2 | Amine-thiol-MAD adduct reactivity towards different amine-thiol-containing molecules. In a green field (on the left), the most efficient declicking agents are presented: ethanedithiol, Cys, and DTT. They quantitatively remove amine and thiol from the MAD adduct. In an orange field (on the bottom right), two exemplary compounds are given that can only liberate thiol from amine-thiol-MAD adducts: TGA and BME. In a red field (on the top right), one exemplary compound, ethylenediamine, is given that does not react with amine-thiol-MAD adducts. Created according to the information in ref. 24.

However, it turned out to be a simplified explanation of a more complex mechanism studied in detail by the same group in 2019.[25] The 'de-clicking' process with DTT completes within several hours and involves several stages that contribute to the apparent (net) reaction rate: a second-order step, and two first-order steps of comparable rates. Each of these steps consists of individual processes, highlighted in the corresponding boxes in Figure 3.3 starting from **2Ph-**X. The rate-determining step of the overall reaction was found to be a combination of the last two steps (red and green boxes) with the formation of a seven-membered-ring product (6 in Figure 3.3) and its transformation into a tricyclic charged form (7 in Figure 3.3) and a five-membered-ring product (3 in Figure 3.3). The work demonstrates the use of a singular value decomposition (SVD) analysis for a detailed kinetics study of the complex reaction by UV-Vis. SVD is one of the techniques used to elucidate individual spectra of the components

present in the reaction mixture at each time, along with their kinetic profiles over time. This method can be useful for investigating details of UV-Vis data of the peptide-MAD reactions, as was mentioned in chapter 2. The authors highlight that the case study is done for the most complex reaction of the aniline-MAD adduct with DTT and that for other MAD adducts (with aliphatic amines), the mechanism of 'de-clicking' can be much simpler, as there are less elementary steps in the overall process. In these cases, UV-Vis spectra of the reaction evolution do not indicate the presence of several UV-active species that co-exist or interconvert in parallel processes.[25]

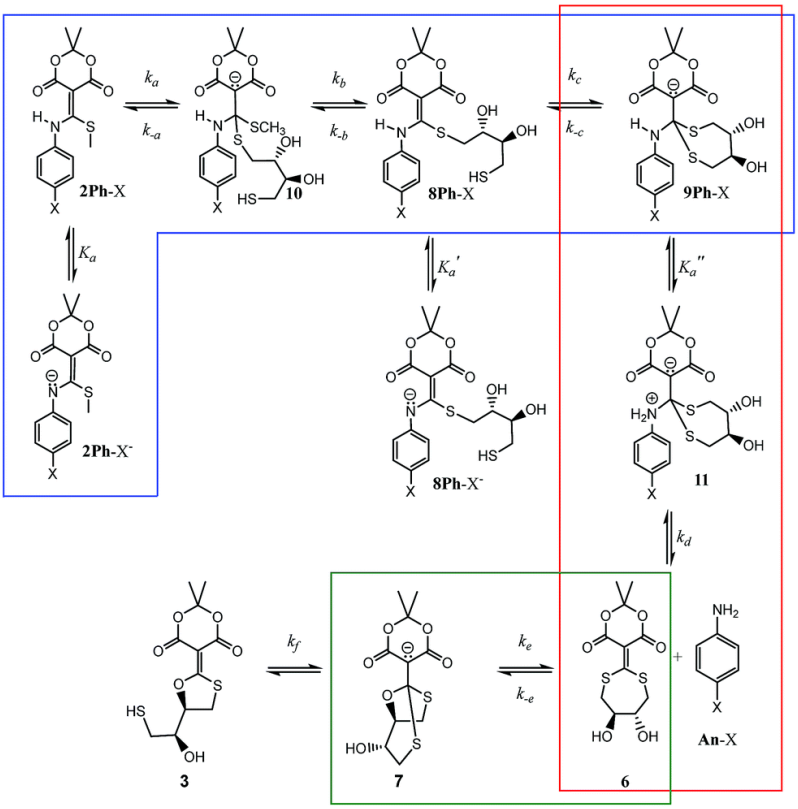


Figure 3.3 | Multistage mechanism of the reaction of DTT with aniline-MAD adduct. The reaction proceeds via three main stages that contain more elementary steps: one second-order (a blue box on the top) and two first-order reactions (in a red box on the right and a red box on the bottom). Reprinted from ref. 25.

For the peptide cyclization-decyclization process, Anslyn and co-workers also utilized DTT.[26] Complete release of the linear peptide was accomplished for all peptides except for those that had N-terminal cysteine. The authors did not study the kinetics of the decyclization in detail, neither did they explore other decyclization reactants apart from DTT.

In this chapter, we describe the decyclization process of the **P1**/**P2**-MAD system that was introduced in chapter 2 (Figure 3.4). The goal is compare the cyclization and decyclization rates, which are important to design an 'autonomous reaction cycle' that leads to transient self-assembly of CPNs (in chapter 4). We also cover several alternatives to DTT and their limitations.

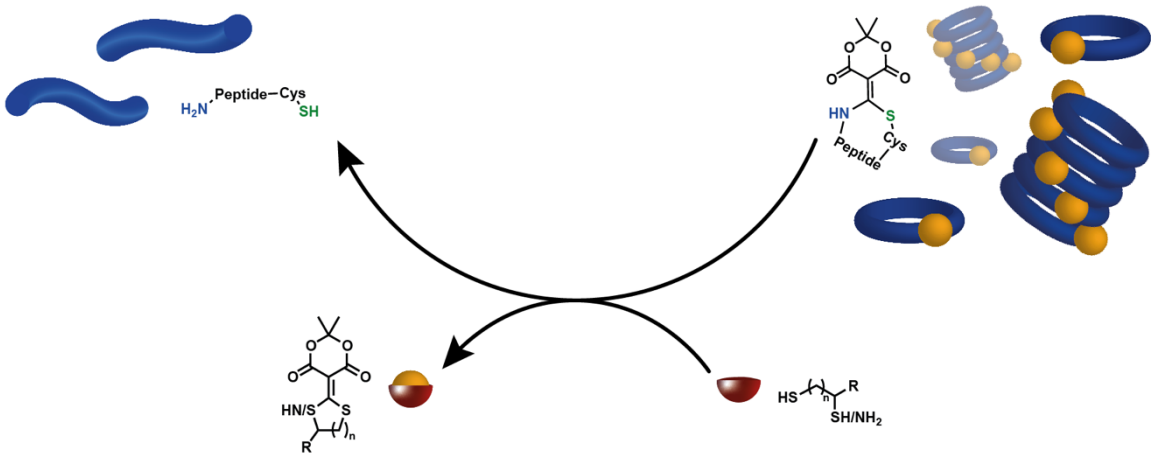
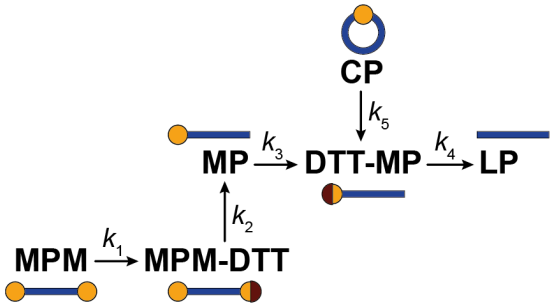


Figure 3.4 | General scheme of CPs decyclization with the recovery of the original linear peptide and yielding decyclization-agent-MAD adduct as a waste. A decyclization agent is indicated schematically as an amine-thiol- or dithiol-containing molecule with n = 1–2 without strict substituent R requirements.

3.2 Experimental studies of the decyclization reaction

3.2.1 Initial kinetic studies

In the first trials, we saw that adding an excess of L-cysteine (Cys) or DL-dithiothreitol (DTT) to the equilibrated cyclization reaction crude causes decolorization of the solution and decreases amounts of the visible aggregates. Next, we followed a reaction of the crude 1.41 mM **CP1** with 100 eq. DTT by LC-HRMS (Figure 3.5a). After 7 hours of reaction, the solution contained almost exclusively **LP1**. Several peptide species were found and followed over time: **LP1**, **MP1**, **CP1**, and **DTT-MP1**. The last one is the intermediate that forms once one thiol of DTT attacks the MAD part of CP, which results in a ring opening with the release of cysteine thiol of the peptide. In the scheme of Figure 3.3, this corresponds to the transition **2Ph**-X to **10** and the formation of **8Ph**-X. This step in the decyclization process occurs within the first hour; then, the amount of **DTT-MP1** declines as the linear peptide is released after the second thiol of DTT reacts with the MAD part. This is analogous to the conversion **8Ph**-X → **9Ph**-X → **11** → **6**+**An**-X in Figure 3.3. Taking all of this into account and if all peptide-MAD species can react with DTT, the corresponding apparent rate expressions for the decyclization process could be written ((3.1(3.6). This set of ODEs was solved numerically in a Python environment to estimate the rate constants. For a detailed description of the code, see Appendix Section 3.5.2.



$$\frac{d[\mathbf{MPM}]}{dt} = -k_1[\mathbf{MPM}][\mathrm{DTT}] \tag{3.1}$$

$$\frac{d[\mathbf{MPM\_DTT}]}{dt} = k_1[\mathbf{MPM}][\mathrm{DTT}] - k_2[\mathbf{MPM\_DTT}] \quad (3.2)$$

$$\frac{d[\mathbf{MP}]}{dt} = k_2[\mathbf{MPM\_DTT}] - k_3[\mathbf{MP}][\mathrm{DTT}] \quad (3.3)$$

$$\frac{d[\mathbf{DTT\_MP}]}{dt} = k_3[\mathbf{MP}][\mathrm{DTT}] + k_5[\mathbf{CP}][\mathrm{DTT}] - k_4[\mathbf{DTT\_MP}] \quad (3.4)$$

$$\frac{d[\mathbf{CP}]}{dt} = -k_5[\mathbf{CP}][\mathrm{DTT}] \quad (3.5)$$

$$\frac{d[\mathbf{LP}]}{dt} = k_4[\mathrm{DTT\_}\mathbf{MP}] \quad (3.6)$$

$$\frac{d[\mathrm{DTT}]}{dt} = -k_3[\mathbf{MP}][\mathrm{DTT}] - k_5[\mathbf{CP}][\mathrm{DTT}] - k_1[\mathbf{MPM}][\mathrm{DTT}] \quad (3.7)$$

Table 3.1 | Apparent rate constants of the decyclization process for the P1 system. The values and their standard deviations are obtained from the numerical integration of equations (3.1)–(3.7).

Rate constants	Values
k_1	$0.8 \pm 1000\ M^{-1}s^{-1}$
k_2	$1 \pm 2000\ s^{-1}$
k_3	$(4.5 \pm 0.8) \times 10^{-3}\ M^{-1}s^{-1}$
k_4	$(0.120 \pm 0.006) \times 10^{-3}\ s^{-1}$
k_5	$(0.73 \pm 0.07) \times 10^{-3}\ M^{-1}s^{-1}$

Values of k_1 and k_2 are not reliable as there was no experimental data on the temporal progress of **MP1M** and **MP1M-DTT**. These constants are, however, not rate limiting so the model does not suffer from this reliability issue. Although **MP1M** is present in the crude at the end of the **P1** cyclization reaction, it was not found already in the first measurement that was done several minutes after the reaction started (Figure 3.5a), which suggests that its reaction with DTT is the fastest in the system. The presence of **MP1M-DTT** is hypothesized to describe the consumption of **MP1M**, but its mass was also not found in the experimental data. Immediate consumption of **MP1M** can be seen in UV-Vis in Figure 3.5b. Once DTT is added to the cyclization crude, the residual absorbance at 350 nm (attributed to **MP1M** in the previous chapter) drops within the first minutes.

To design a reaction cycle where cyclization reaction leads to transient CPs and CPNs and decyclization reaction relaxes the system to equilibrium, one needs the cyclization process to be faster than the decyclization process. Alternatively, the decyclization reagent could be produced *in situ* (see chapter 4). Comparing the net rate of the cyclization process with that of decyclization according to the Transient activation model in chapter 1,[27] we can see that the **P1**-MAD-DTT system could work as a reaction cycle, as it falls into Regime I (see Appendix 3.5.1) together with most transient systems developed so far (chapter 1, Section 1.2.2.1). However, this would be the case only if there was no reaction between MAD and DTT, which occurs faster than the reaction between MAD and peptides.[26,28] Further, we compared different decyclization agents, aiming to find a molecule that would have a decreased reactivity towards free MAD compared to the peptide-MAD adducts.

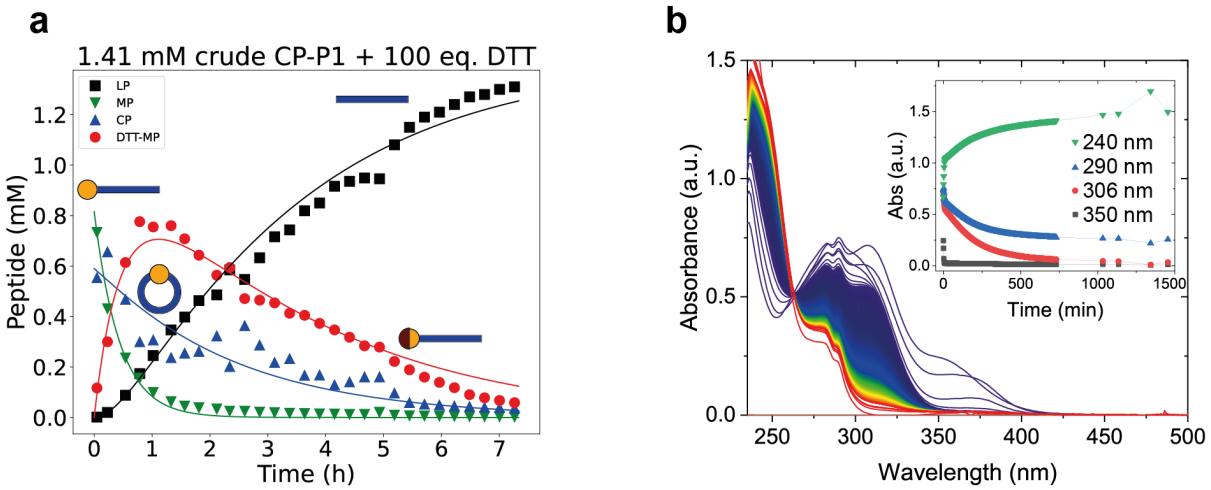


Figure 3.5 | LC-HRMS and UV-Vis data of the decyclization reaction of crude P1-CP with DTT. (a) The reaction of 1.41 mM **CP1** with 100 eq. DTT followed by LC-HRMS. Lines represent the fitting result according to the ODEs model. (b) The UV-Vis kinetics of 0.7 mM **CP1** with 30 eq. DTT. In the inset, the absorbance at different wavelengths is presented over time. 240 nm—absorbance of MAD-DTT product; 290 nm—absorbance of all peptide **P1** species; 306 nm—absorbance of **CP1** mainly; 350 nm—absorbance of MAD.

3.2.2 Comparison of decyclization agents (DAs)

We also tried to optimize the ratios of MAD, DTT and peptide, hoping to find such amounts of every reactant that would level out the impact of the MAD-DTT side reaction on the system. We found the minimum amount of DTT needed to fully revert **CP1** into **LP1** is 50 eq. (Figure 3.6a–b). It was done by sequential addition of 10 eq. DTT to the crude **CP1** solution and measuring the reaction mixture six hours after each addition. In Figure 3.6a, it is evident that an increase in the DTT amount causes the decrease of **CP1** and **MP1** fractions while **LP1** and **LP1-DTT** increase. **LP1-DTT** is a linear peptide connected to DTT via a disulfide bond. The minimum DTT amount needed was estimated from the plot of DTT eq. versus the amount of **CP1** is left after each DTT addition (Figure 3.6b). The intercept of the linear fit corresponds to 48±2 eq. DTT. Interestingly, the same experiment with Cys instead of DTT showed much lower efficiency of Cys as DA (Figure 3.6c). With the increasing amount of Cys, only the **MP1** fraction decreases with the increase of the fraction of **Cys-MP1** and **LP1**, while the fraction of **CP1** slightly varies.

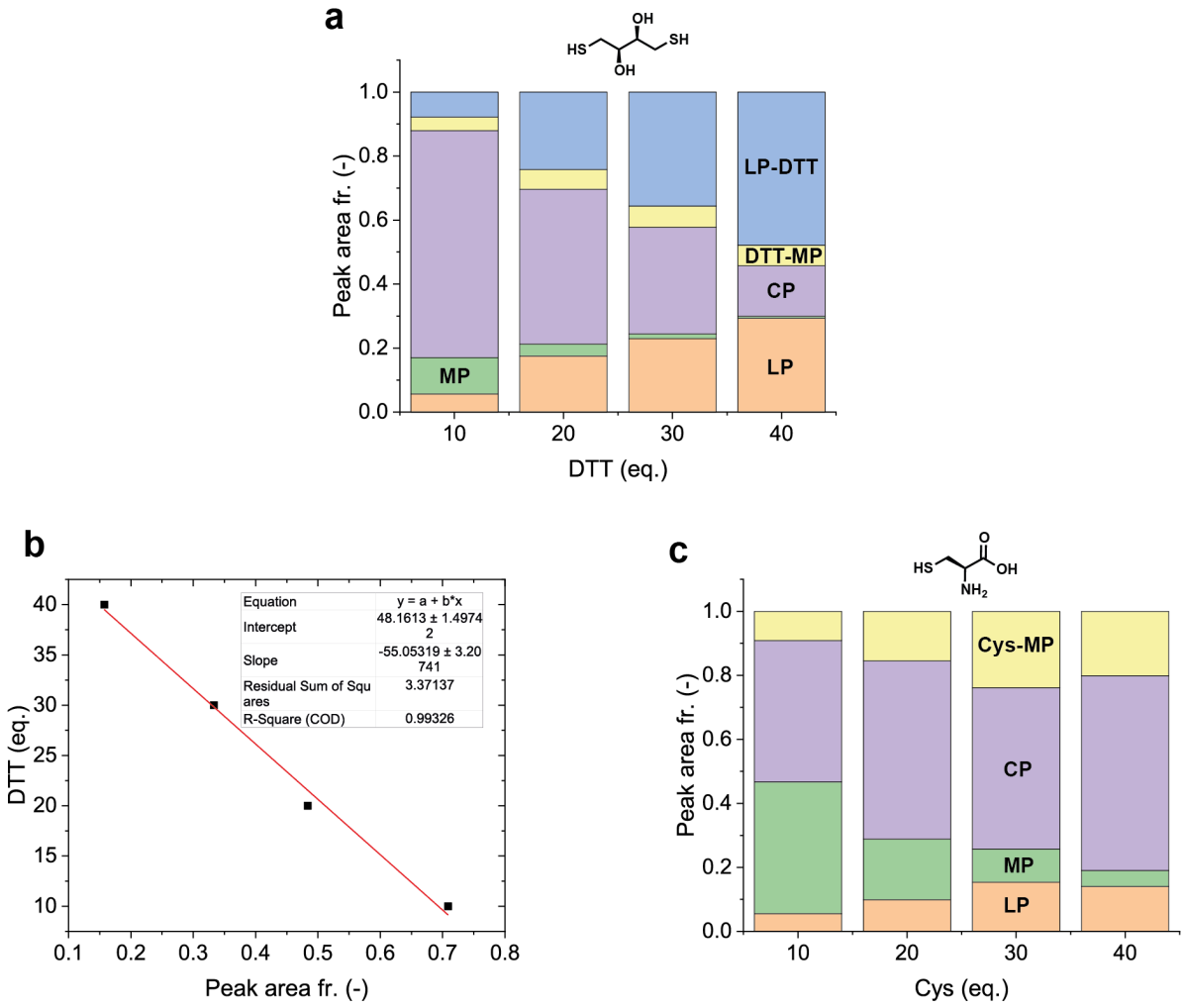


Figure 3.6 | Comparison of DTT and Cys efficiency in the reaction with crude CP1. (a) Comparison of the solution composition six hours after different amounts of DTT are added. (b) The plot of DTT amount over the amount of unreacted **CP1** to determine the minimum amount of DTT needed to achieve complete conversion of **CP1** into **LP1**. (c) Comparison of the solution composition six hours after different amounts of Cys are added.

The same studies were done for **P2** (Figure 3.7). Although Cys reacts fast with **CP2**, the recovery of **LP2** from **Cys-MP2** proceeds slower than in the same experiments with DTT. Using 100 eq. of Cys did not improve the recovery rate of LP; on the contrary, **Cys-MP2** was present in the system for longer times than with less Cys (Figure 3.7, compare panels c, d, and e).

We concluded that for both peptides, Cys is a much less efficient DA compared to DTT. Moreover, MAD reacts faster with Cys than with peptides, which makes it not suitable for designing the reaction cycle.

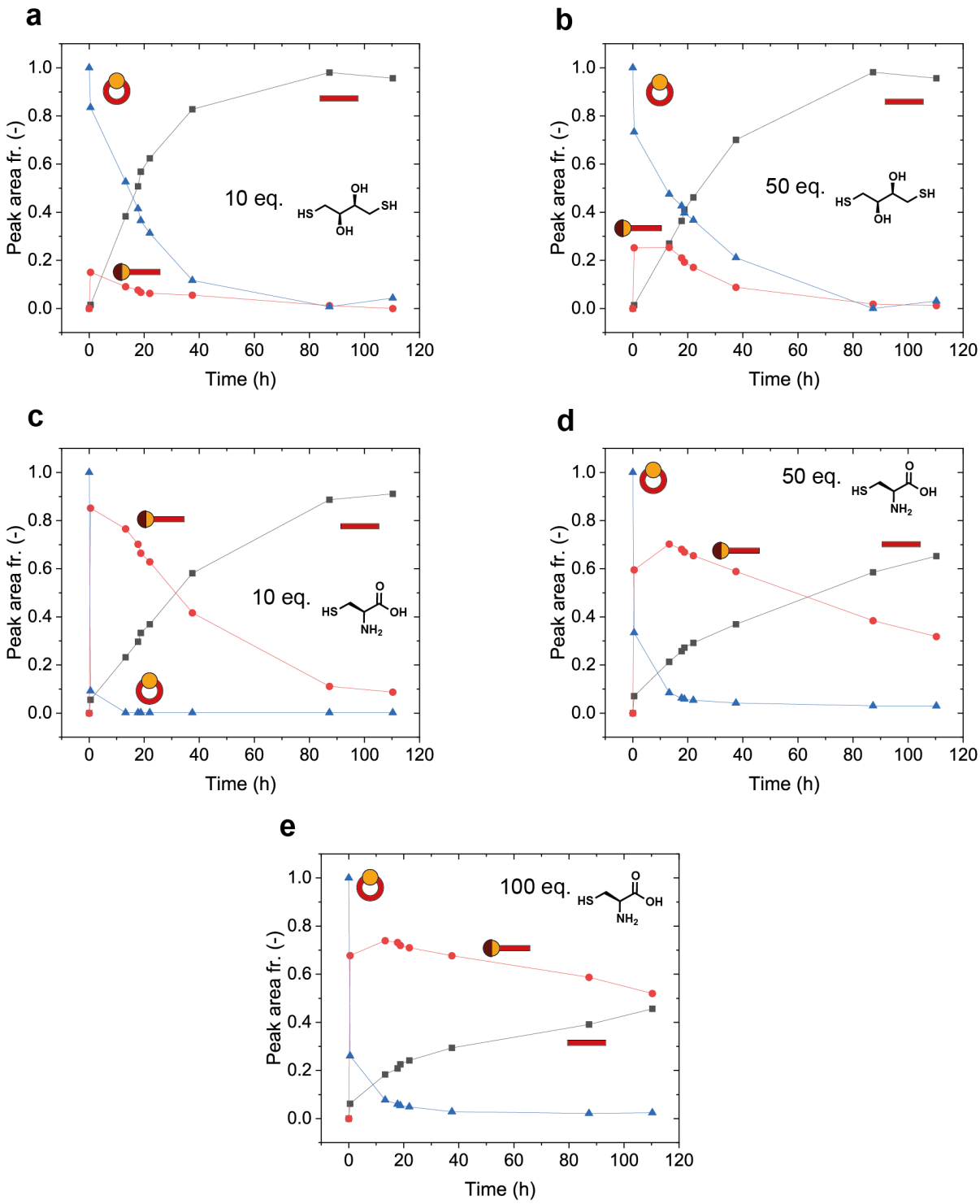


Figure 3.7 | Comparison of DTT and Cys for decyclization of the crude CP2. To a 1 mM solution of the crude P2-CP corresponding amounts of DTT and Cys were added. The reactions were monitored by LC-HRMS.

Next, we tested cysteamine (CA) in the reaction with MAD (Figure 3.8a) and found that it reacts much slower than DTT and Cys. The half-life of the reaction estimated from the decrease of the absorbance at 355 nm was 30 min (0.5 mM, equimolar amounts of CA and MAD). This is higher than half-lives of the reaction of MAD with **P1** and **P2** at a similar concentration (19 and 13 min, respectively). Consequently, CA could be a suitable DA to develop a reaction cycle with the peptide-MAD system. Unfortunately, testing CA in the reaction with the crude

CP2 revealed the same problem as with Cys: **CA-MP2** adduct recovers **LP2** extremely slowly (< 20% **LP2** in 116 hours, Figure 3.8b), which is contrary to our desire to recover **LP2** fully.

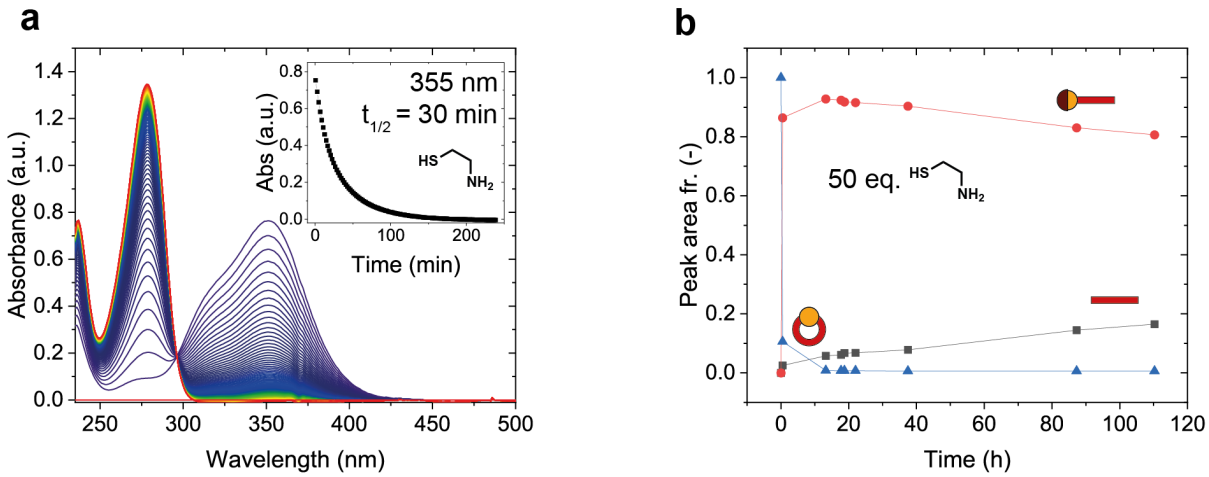


Figure 3.8 | Studies of CA in reactions with MAD and the crude CP2. (a) UV-Vis kinetics of 0.5 mM CA reaction with 1 eq. MAD. (b) LC-HRMS data of the reaction of CA with 1 mM crude **CP2** solution.

The following DA candidates were tested in the reaction with MAD: sodium dimercaptopropanesulfonate (DMPS, Figure 3.9a) and meso 2,3-dimercaptosuccinic acid (DMSA, Figure 3.9b) and compared to the DTT (Figure 3.9c, Figure 3.10a). DMPS, a water-soluble analog of ethanedithiol, reacted with MAD a bit slower than DTT but still too fast to use it directly for the peptide-MAD system in the reaction cycle. However, it should work as DA if added to the crude **CP**. DMS reacted with MAD extremely slowly, but when added in excess to the crude CP, resulted only in decolorization of the solution after one day with all aggregates (evidenced by visual turbidity) left intact for the following weeks.

Finally, we also checked whether, for our system, the conclusion made by Anslyn and co-workers holds that only small molecules containing two thiols or thiol and amine can remove MAD from CP efficiently (Figure 3.2). Thioglycolic acid, TGA, was tested in the reaction with the crude **CP2** (Appendix, Figure 3.10b–d). As expected, TGA managed to open the ring of **CP2** forming **TGA-MP2**, but no recovery of **LP2** was observed.

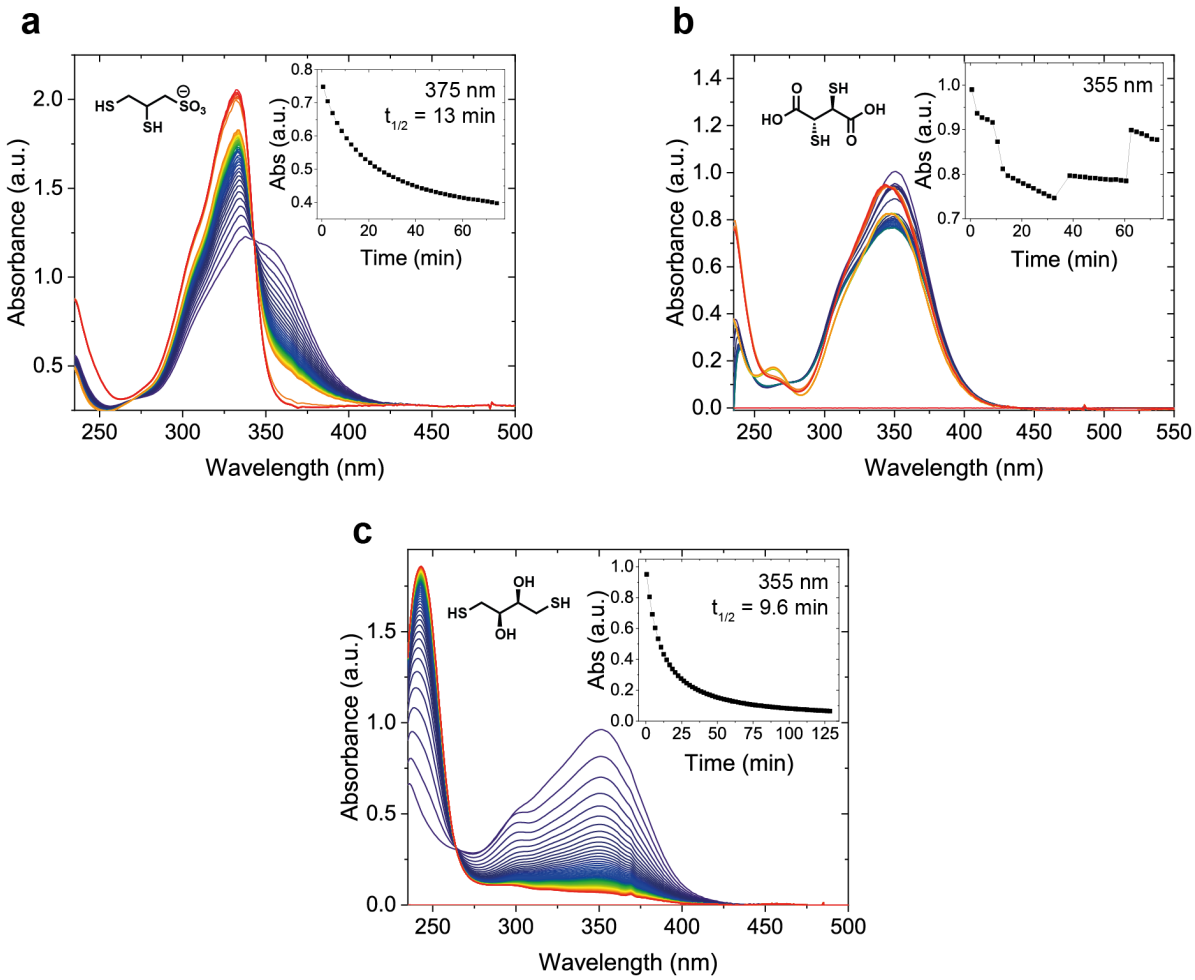


Figure 3.9 | UV-Vis studies of MAD reaction with different DA (1:1 mol. ratio). (a) DMPS, 0.5 mM. (b) DMSA, 0.5 mM. (c) DTT, 1 mM.

3.3 Conclusion and Outlook

The main task of this chapter was to understand which compounds are suitable for peptide-MAD system as efficient decyclization agents. DTT shows the most efficient linear peptide when used in 50-fold excess for **P1** and 10-fold excess for **P2** systems. Compared to Cys and CA, DTT does not form a long-lived adduct with **MP**, which makes it a better choice when the aim is to revert the system fully to the linear peptide and not just destroy CPs. The situation is similar for dimercaptopropanesulfonate (DMPS), which can efficiently proceed linear peptide from its cyclic analog. In case only a ring-opening reaction is needed, Cys and CA are sufficient already.

The initial idea was to find such decyclization agent that reacts with free MAD slower than peptide with MAD but is yet able to recover linear peptide quantitatively from **CP** and **MPM**. Finding such a compound would allow us to create a reaction cycle where **MP**, **CP**, **MPM**, and CPNs form transiently. Unfortunately, we did not find a compound meeting those criteria, which led us to the idea of using a 'pre-fuel' (see chapter 4)— an inactive form of the decyclization agent that is slowly activated and released in the reaction medium so cyclization can proceed unhampered, and decyclization follows later. In this way, the reaction of peptides

with MAD will be more favorable than mutual consumption of MAD and deactivating reactant. In the next chapter, we will look closely into developing such slowly releasing 'pre-fuels' and approach the final solution to achieve the desired transient CPNs.

3.4 Author contributions

Anastasiia Sharko performed all UV-Vis measurements and most of the LC-HRMS experiments, developed the kinetic expressions to model the decyclization process and fit the experimental data. Shamkhal Baybekov evaluated MS data for P2 system as a part of his M2 internship. Thomas Hermans supervised the research.

3.5 Appendix

Chemicals and Materials. Chemicals and solvents were purchased from TCI chemicals, Sigma Aldrich, Fisher. For the reaction system general solvent mixture composition was 1.6 % DMF, 16.7 % acetonitrile and 81.7 % buffer. A VWR Model 110 pH meter was used to record the pH measurements during buffer preparations.

UV-Vis. UV experiments were performed on an Agilent CARY 8454-UV spectrometer in 1–2 mm quartz cuvettes. Wavelength range 235–600 nm or 190–1100 nm.

LC-HRMS. HRMS was performed with an Thermofisher Scientific UltiMate 3000 RSLCnano UHPLC System coupled with EMT Thermo OrbiTrap Mass analyzer and the ESI ionization type. Hypersil GOLD column, 50 x 2.1 mm, 1.9 μm was used for LC part.

3.5.1 Transient activation model

The model was developed by Hermans and co-workers[27] to estimate whether the system of chemical reactions has transiently activated species based on the rates of the forward and backwards processes. Although this model is not designed to include self-assembly, for peptide-MAD system it also accounts formation of CPNs and their aggregates, as so do the apparent rate constants obtained via numerical integration.

The net cyclization rate was determined using the rate constants from Chapter 2.

$$\text{Cyc. R.} = k_1[\mathbf{MAD}] + k_2 \tag{3.8}$$

$$\text{DeCyc. R.} = (k_5[\mathbf{DTT}] + k_4) + (k_3[\mathbf{DTT}] + k_4) \tag{3.9}$$

$$\frac{k_d'}{k_a c_{tot}} = \frac{\text{DeCyc. R.}}{\text{Cyc. R.}} = 0.149;\ \frac{f_0}{c_{tot}} = 1.5 \tag{3.10}$$

f_0 is the MAD ('fuel') concentration, c_{tot} is the total peptide concentration. Calculations are done for the system 1.41 mM **CP1**, 1.5 eq. MAD, and (stepwise) 100 eq. DTT (Figure 3.5a). With the resulting parameter values, the **P1**-MAD system falls into Regime I with the dominant

activation over deactivation (see Figure 1.4 in chapter 1). This potentially means that high transient yields of CPNs would be possible to achieve, if MAD did not react with DTT.

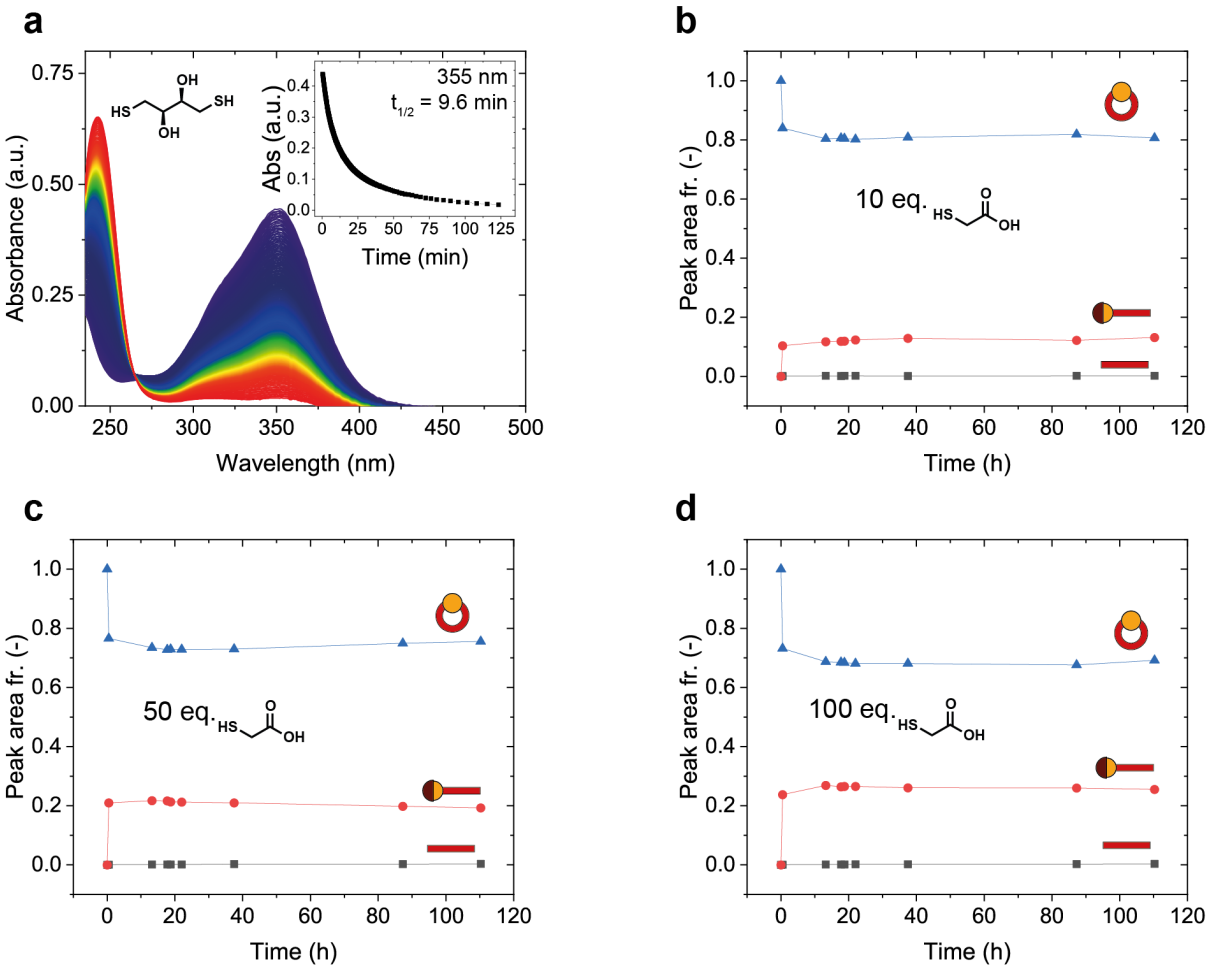


Figure 3.10 | Supplementary data of DA reactions with MAD and crude P2-CP. (a) UV-Vis data of 0.35 mM MAD-DTT reaction. (b–d) TGA reactions with 1 mM crude **CP2**, followed by LC-HRMS.

3.5.2 Python script

To fit the LC-HRMS data of the decyclization reaction, the following code was developed based on the one for the cyclization process. It was implemented in Jupyter Notebook in Anaconda environment.

Input [ ]:

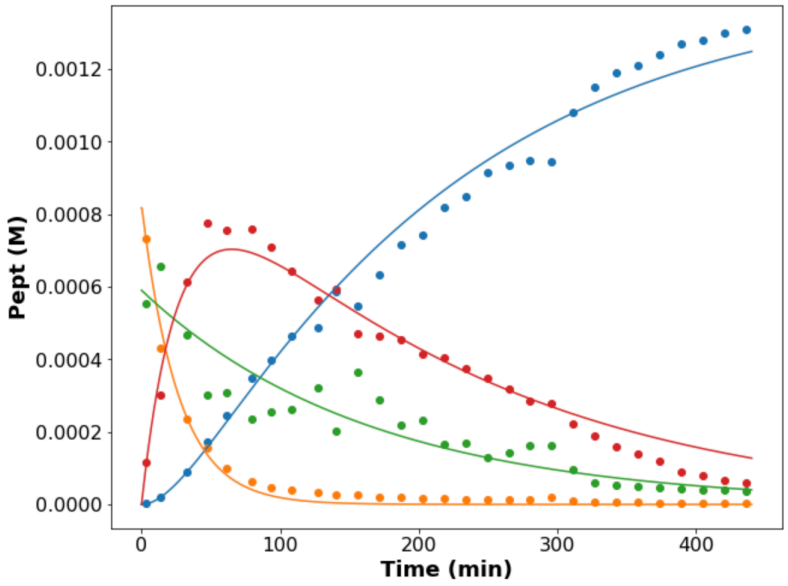
```
import numpy as np
from symfit import parameters,variables,Fit,D,ODEModel, Parameter
import matplotlib.pyplot as plt

%matplotlib nbagg
%matplotlib inline

# a=linear, ab=pept-mad, abc=mad-pept-mad, d=cyclic, e=pept-mad-dtt, e1=mad-
pept-mad-dtt
```

```
a, ab, abc, d, e, e1, g, t = variables('a, ab, abc, d, e, e1, g, t')
k1, k2, k3, k3a, k4 = parameters('k1, k2, k3, k3a, k4')
k1 = Parameter('k1', value=50) # min=1*60, max=2*60
k2 = Parameter('k2', value=60) #, min=0.009*60, max=0.045*60)
k3 = Parameter('k3', value=0.007) #, min=0.3*60, max=0.8*60)
k3a = Parameter('k3a', value=0.037) #, min=0.3*60, max=0.8*60)
k4 = Parameter('k4', value=0.0072) #, min=0.3*60, max=0.8*60)

t_data=np.array(dtt.time, dtype=np.float64)
a_data=np.array(dtt.linear, dtype=np.float64)
ab_data=np.array(dtt.arm, dtype=np.float64)
d_data=np.array(dtt.cyclic, dtype=np.float64)
e_data=np.array(dtt.armMADDTT, dtype=np.float64)

t0=0.0
a0= 0.0
ab0 =0.000816
d0 = 0.000590
e0=0.0
abc0 = 0.00001
e10 = 0.0
g0 = 0.141

model_dict = {D(a, t): k4*e, \
              D(ab, t): -k3a*ab*g + k2*e1, \
              D(d, t): -k3*d*g, \
              D(e, t): k3a*ab*g + k3*d*g - k4*e, \
              D(e1, t): k1*abc*g - k2*e1, \
              D(abc, t): -k1*abc*g, \
              D(g, t): -k1*abc*g - k3*d*g - k3a*ab*g}


ode_model = ODEModel(model_dict, initial={t: t0, a: a0, ab: ab0, abc: abc0,
d: d0, e: e0, e1:e10, g:g0})
print (ode_model)
fit = Fit(ode_model, t=t_data, a=a_data, ab=ab_data, abc=None, d=d_data,
e=e_data, e1=None, g=None)
fit_result = fit.execute()
print (fit_result)

tvec = np.linspace(0, 440, 100000)
a, ab, abc, d, e, e1, g = ode_model(t=tvec, **fit_result.params)

plt.figure(figsize=(10, 8))
plt.rcParams['xtick.labelsize'] = 16
plt.rcParams['ytick.labelsize'] = 16
plt.plot(tvec, a, label='Linear_fit')
plt.plot(tvec, ab, label='Linear-MAD_fit')
plt.plot(tvec, d, label='Cyclic_fit')
plt.plot(tvec, e, label='linear-MAD-DTT_fit')
#plt.plot(tvec, f, label='DTT_fit')
plt.scatter(t_data, a_data, label='Linear')
```

```
plt.scatter(t_data, ab_data, label='linear-MAD')
plt.scatter(t_data, d_data, label='Cyclic')
plt.scatter(t_data, e_data, label='linear-MAD-DTT')
#plt.ylim(-0.0001,0.001)
plt.xlabel('Time (min)', fontsize=18, fontweight = 'bold')
plt.ylabel('Pept (M)', fontsize=18, fontweight = 'bold')
#plt.legend(loc='best')

plt.show()
```

Output [ ]:

```
Derivative(a, t; k1, k2, k3, k3a, k4) = k4*e
Derivative(ab, t; k1, k2, k3, k3a, k4) = k2*e1 - k3a*ab*g
Derivative(abc, t; k1, k2, k3, k3a, k4) = -k1*abc*g
Derivative(d, t; k1, k2, k3, k3a, k4) = -k3*d*g
Derivative(e, t; k1, k2, k3, k3a, k4) = k3*d*g + k3a*ab*g - k4*e
Derivative(e1, t; k1, k2, k3, k3a, k4) = k1*abc*g - k2*e1
Derivative(g, t; k1, k2, k3, k3a, k4) = -k1*abc*g - k3*d*g - k3a*ab*g

Parameter Value        Standard Deviation
k1        5.000000e+01 6.022958e+04
k2        6.000000e+01 1.373386e+05
k3        4.361236e-02 4.020680e-03
k3a       2.696340e-01 4.821006e-02
k4        7.227089e-03 3.360819e-04
Status message         Optimization terminated successfully.
Number of iterations   44
Objective              <symfit.core.objectives.LeastSquares object at 0x7fb9e44d7640>
Minimizer              <symfit.core.minimizers.BFGS object at 0x7fb9e44d7f70>

Goodness of fit qualifiers:
chi_squared            2.7623709558383915e-07
objective_value        1.3811854779191958e-07
r_squared              0.9642940157837535
```



3.6 References

1. Cissé, N. & Kudernac, T. Light-Fuelled Self-Assembly of Cyclic Peptides into Supramolecular Tubules. *ChemSystemsChem* **2**, (2020).
2. Sun, L. *et al.* Tunable synthesis of self-assembled cyclic peptide nanotubes and nanoparticles. *Soft Matter* **11**, 3822–3832 (2015).
3. Catrouillet, S. *et al.* Tunable Length of Cyclic Peptide–Polymer Conjugate Self-Assemblies in Water. *ACS Macro Lett.* **5**, 1119–1123 (2016).
4. Larnaudie, S. C., Brendel, J. C., Jolliffe, K. A. & Perrier, S. pH-Responsive, Amphiphilic Core–Shell Supramolecular Polymer Brushes from Cyclic Peptide–Polymer Conjugates. *ACS Macro Lett.* **6**, 1347–1351 (2017).
5. Song, Q., Yang, J., Rho, J. Y. & Perrier, S. Supramolecular switching of the self-assembly of cyclic peptide–polymer conjugates via host–guest chemistry. *Chem. Commun.* **55**, 5291–5294 (2019).
6. Vollmer, M. S., Clark, T. D., Steinem, C. & Ghadiri, M. R. Photoswitchable Hydrogen-Bonding in Self-Organized Cylindrical Peptide Systems. *Angewandte Chemie International Edition* **38**, 1598–1601 (1999).
7. Sang Hoon Joo, Qing Xiao, Yun Ling, Bhaskar Gopishetty, and & Pei*, D. High-Throughput Sequence Determination of Cyclic Peptide Library Members by Partial Edman Degradation/Mass Spectrometry. *ACS Publications* https://pubs-acs-org.scd-rproxy.u-strasbg.fr/doi/full/10.1021/ja063722k (2006) doi:10.1021/ja063722k.
8. Liang, X., Vézina-Dawod, S., Bédard, F., Porte, K. & Biron, E. One-Pot Photochemical Ring-Opening/Cleavage Approach for the Synthesis and Decoding of Cyclic Peptide Libraries. *ACS Publications* https://pubs-acs-org.scd-rproxy.u-strasbg.fr/doi/full/10.1021/acs.orglett.6b00296 (2016) doi:10.1021/acs.orglett.6b00296.
9. Liang, X., Girard, A. & Biron, E. Practical Ring-Opening Strategy for the Sequence Determination of Cyclic Peptides from One-Bead-One-Compound Libraries. *ACS Comb. Sci.* **15**, 535–540 (2013).
10. Lee, K. J. & Lim, H.-S. Facile Method To Sequence Cyclic Peptides/Peptoids via One-Pot Ring-Opening/Cleavage Reaction. *Org. Lett.* **16**, 5710–5713 (2014).
11. Lee, J. H., Meyer, A. M. & Lim, H.-S. A simple strategy for the construction of combinatorial cyclic peptoid libraries. *Chem. Commun.* **46**, 8615–8617 (2010).
12. Kato, H., Tsuji, K. & Harada, K. Microbial degradation of cyclic peptides produced by bacteria. *J Antibiot* **62**, 181–190 (2009).
13. Ulrich, S. Growing Prospects of Dynamic Covalent Chemistry in Delivery Applications. *Acc. Chem. Res.* **52**, 510–519 (2019).
14. Ramström, O. & Lehn, J.-M. Drug discovery by dynamic combinatorial libraries. *Nat Rev Drug Discov* **1**, 26–36 (2002).
15. Hermanson, G. T. *Bioconjugate Techniques*. (Academic Press, 2013).
16. Min, Y., Axet, M. R. & Serp, P. Covalent Assemblies of Metal Nanoparticles—Strategies for Synthesis and Catalytic Applications. in *Recent Advances in Nanoparticle Catalysis* (eds. van Leeuwen, P. W. N. M. & Claver, C.) 129–197 (Springer International Publishing, 2020). doi:10.1007/978-3-030-45823-2_5.
17. Liu, B. & Thayumanavan, S. Substituent Effects on the pH Sensitivity of Acetals and Ketals and Their Correlation with Encapsulation Stability in Polymeric Nanogels. *J. Am. Chem. Soc.* **139**, 2306–2317 (2017).
18. Shigenaga, A., Yamamoto, J., Kohiki, T., Inokuma, T. & Otaka, A. Invention of stimulus-responsive peptide-bond-cleaving residue (Spr) and its application to chemical biology tools. *Journal of Peptide Science* **23**, 505–513 (2017).
19. Li, J., Jia, S. & Chen, P. R. Diels-Alder reaction–triggered bioorthogonal protein decaging in living cells. *Nat Chem Biol* **10**, 1003–1005 (2014).
20. Bielski, R. & Witczak, Z. Strategies for Coupling Molecular Units if Subsequent Decoupling Is Required. *Chem. Rev.* **113**, 2205–2243 (2013).
21. Bielski, R., Witczak, Z. J. & Mencer, D. Example of Sacrificial Unit Using Two Different Click Reactions in Coupling and Decoupling (CAD) Chemistry. in *Coupling and Decoupling of Diverse Molecular Units in Glycosciences* (eds. Witczak, Z. J. & Bielski, R.) 299–307 (Springer International Publishing, 2018). doi:10.1007/978-3-319-65587-1_13.
22. Cruz, C. M. *et al.* Vinyl Sulfonates: A Click Function for Coupling-and-Decoupling Chemistry and their Applications. *Advanced Synthesis & Catalysis* **358**, 3394–3413 (2016).
23. Maity, C., Trausel, F. & Eelkema, R. Selective activation of organocatalysts by specific signals. *Chem. Sci.* **9**, 5999–6005 (2018).

24.Diehl, K. L. *et al.* Click and chemically triggered declick reactions through reversible amine and thiol coupling via a conjugate acceptor. *Nat. Chem.* **8**, 968–973 (2016).
25.Meadows, M. K. *et al.* Mechanistic studies of a "Declick" reaction. *Chem. Sci.* **10**, 8817–8824 (2019).
26.Johnson, A. M. & Anslyn, E. V. Reversible Macrocyclization of Peptides with a Conjugate Acceptor. *Org. Lett.* **19**, 1654–1657 (2017).
27.Sharko, A., Livitz, D., De Piccoli, S., Bishop, K. J. M. & Hermans, T. M. Insights into Chemically Fueled Supramolecular Polymers. *Chem. Rev.* **122**, 11759–11777 (2022).
28.Sun, X. *et al.* Chemically Triggered Synthesis, Remodeling, and Degradation of Soft Materials. *J. Am. Chem. Soc.* **142**, 3913–3922 (2020).

Chapter 4. Transient cyclic peptide self-assembly in a chemical reaction network

Abstract

Although there are many peptide-based nonequilibrium systems, none show transient peptide cyclization and assembly. Our system is convenient for studying self-assembly coupling to the reaction network dynamics. Here we show how to combine cyclization and decyclization processes into the reaction cycle and omit the consumption of MAD by the decyclization agent (DA). Transient cyclic peptide nanotubes could be of use in biomaterials as drug carriers, able to release molecules in specific places in the body, where the environment would trigger the release and determine its half-life. In more complex artificial systems, the developed CRN will serve as a model where coupling of self-assembly to the (de)activating reactions can be studied in detail.

4.1 Transient peptide systems

Peptide substrates are especially attractive for building reaction cycles and studying fueled self-assembly because they resemble natural fibers; their chemical structure permits fine-tuning hydrophobic-hydrophilic properties by changing amino acid substituents.[1–8] Using specific peptide sequences allows using enzymes as catalysts to drive the system out of equilibrium, where the activated state promotes self-assembly of the peptide-based substrate and back to equilibrium.[9]

Peptide cyclization based on disulfide-thiol-thioester exchange in the autocatalytic reaction network has been shown by Semenov and co-workers.[10] In this study, peptide monomer, disulfide-bonded di-tripeptide with thioesters at C-termini, was left in a buffer in the presence of TCEP. Thioester hydrolyzed, releasing a thiol, which initiated the disulfide exchange that further resulted in the evolution of different species. The formation of cyclic peptides occurred via intermolecular disulfide bonds and cysteine-thioester exchange at C-terminus, followed by native chemical ligation (NCL). Macrocyclization through NCL in this system was terminal, i.e. no decyclization or recovery of the initial peptides occurred. The distribution of cycles of different sizes has been studied in batch conditions at equilibrium and at a steady state in a continuously stirred tank reactor (CSTR). In the experiments with a thiol-removing agent in CSTR (Figure 4.1a), the authors observed modification of cysteine residues of macrocycles, which rate was dependent on the macrocycle size. The work highlights the importance of such factors as ring strain and aggregation for the system where formation and degradation processes compete.[10] Thus, the more strained peptide cycles are more prone to be 'removed' (functionalized), whereas, in larger cycles, cysteine thiols are protected from functionalization by intramolecular disulfide bonds.[10] Similarly, aggregation or self-assembly of certain cycles can shield them from deactivation.

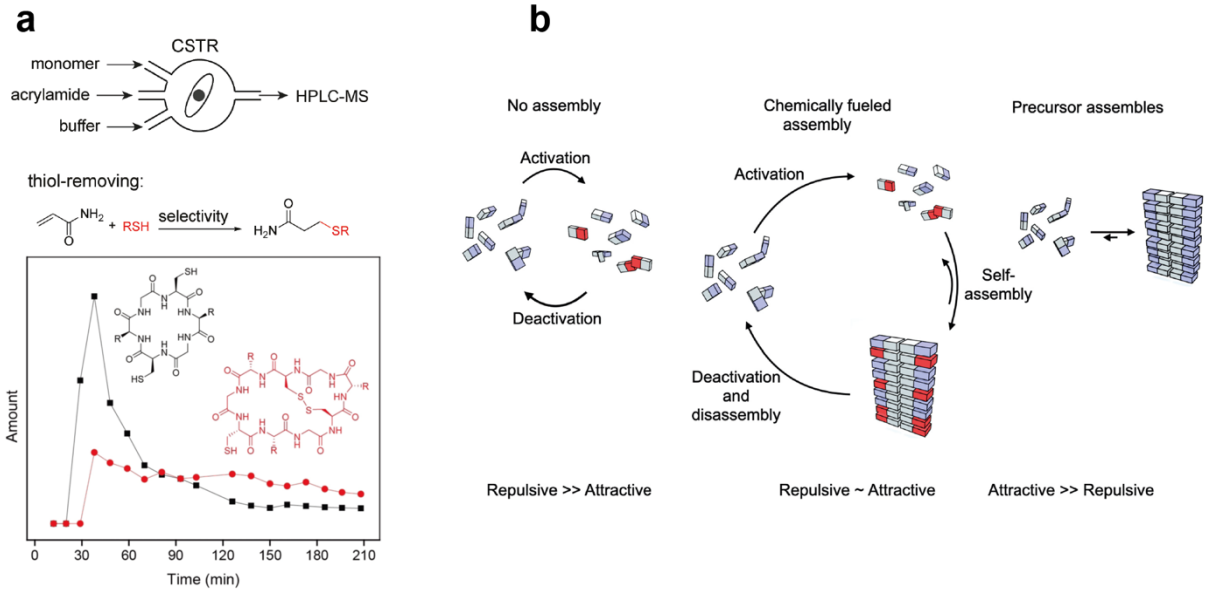


Figure 4.1 | Selected examples of out-of-equilibrium peptide-based systems. (a) Autocatalytic formation of cyclic peptides in CSTR. Adapted from ref. 10. (b) Balancing repulsive and attractive forces in the substrate, the transient self-assembling state can be tuned from a completely non-assembling into a permanently gelled. Adapted from ref. 8.

Boekhoven and co-workers[8] closely studied the influence of changing the structure of oligopeptide on the lifetime of the activated state (Figure 4.1b). Adding one hydrophobic amino

acid residue to the substrate increased the hydrophobic effect and resulted in transient gelation of the reaction solution, which did not produce any traceable activated substrate quantities before the structure was changed. The addition of one more substituent to the substrate resulted in gelation upon fueling that could not revert to the solution. The two main points of this study:[8] (i) coupling of self-assembly to the (de)activation reaction network impacts the activated state; (ii) the reaction cycle can be tuned not just by changing fuels concentration or medium conditions but by changing the substrate structure, namely by changing the ratio of repulsive-attractive interactions, which with peptides is a task with a well-developed strategy, as the needed combinations of the amino acids that promote gelations are established.[11,12]

For the peptide-MAD system, self-assembly will also impact the deactivation, slowing it down. The design of the reaction cycle should consider that the decyclization agent is longlived enough to drive cyclic peptides back to their linear analogs. In this chapter, our goal was to develop a reaction network to achieve transient formation and self-assembly of cyclic peptides and find the conditions to maximize a transient CP yield (Figure 4.2). To the best of our knowledge, so far, there is no other system that has studied out-of-equilibrium formation and aggregation of cyclic peptides followed by the recovery of the starting linear peptide.

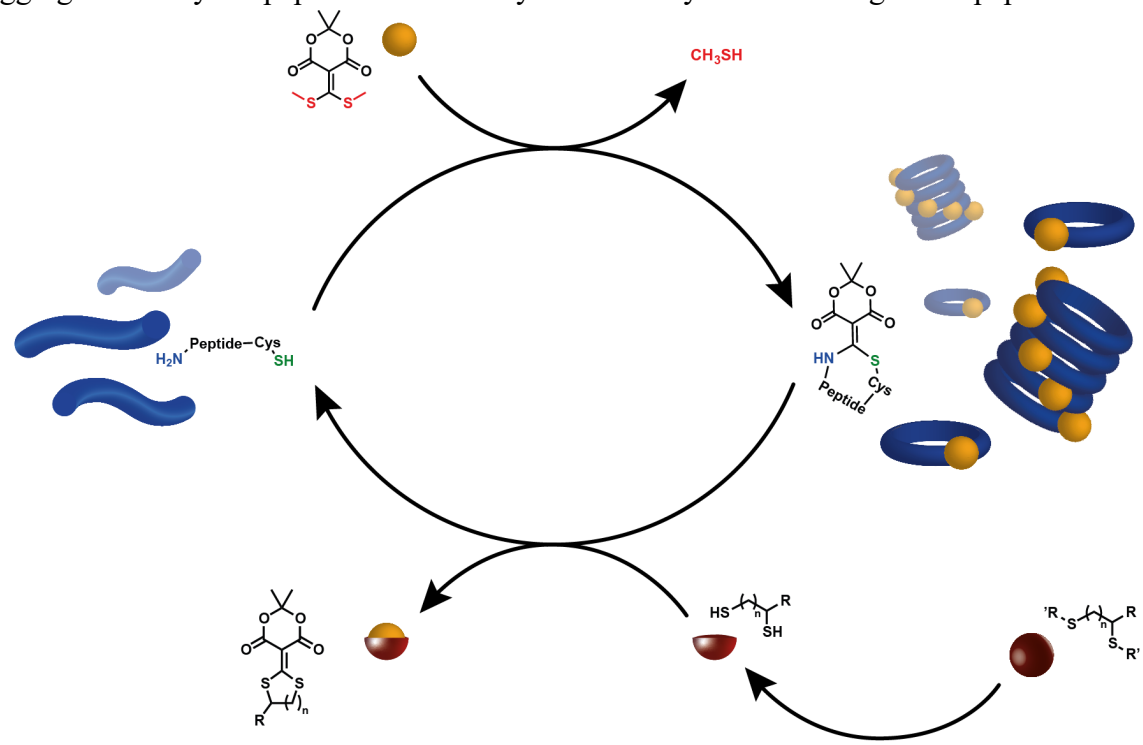


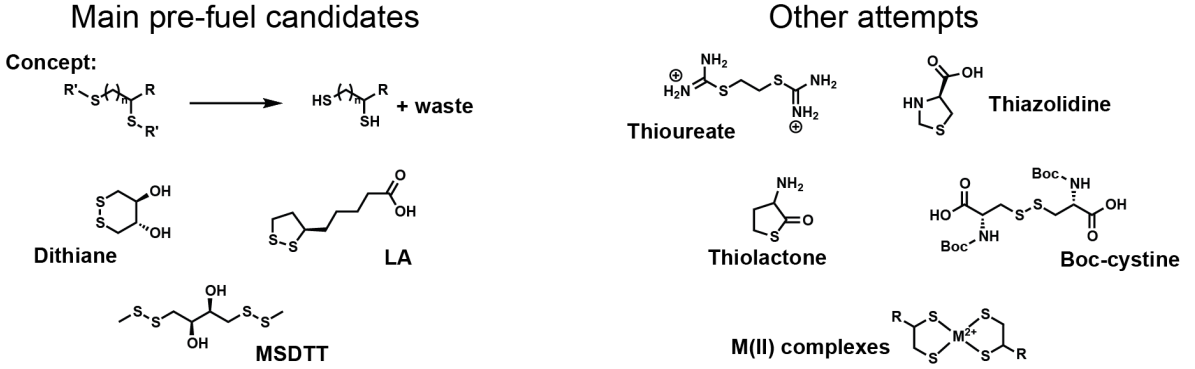
Figure 4.2 | The general scheme representing the peptide-MAD reaction network.

4.2 Conditions optimization for peptide-MAD reaction network

After studying both cyclization and decyclization reactions of **P1** and **P2** with MAD, we found that even though reaction rates of both processes allow for combining them into a reaction cycle, the reaction of MAD with the decyclization agent (DA) is much faster than with peptides. This led us to the idea of having a delayed release of DA from the inactive precursor via an additional reaction that would proceed in the system along with the cyclization process.

4.2.1 Pre-fuel concept

The use of a 'pre-fuel' was earlier demonstrated in our group in a reaction cycle where forward and backward processes also conflicted with each other.[13,14] To extend the lifetime of transient species, an additional reaction was introduced that slowly produced *in situ* an agent (or fuel) that drove the backward process.[13] A particular feature of the pre-fuel-to-fuel conversion in that system was acid catalysis. Changing the amount of catalyst affected not just the duration of the activated state but also the morphology of the formed transient fibers.[13] For the peptide-MAD system finding a suitable pre-fuel means finding a reaction in which the DA, like DTT or any other dithiol, is released after 1–5 hours of cyclization process, and the pre-fuel itself is inert toward peptides and MAD (Scheme 4.1).



Scheme 4.1 | Selected pre-fuels for the peptide-MAD system. The main candidates are dithiol-based. Other attempts also include amine-thiol-based molecules.

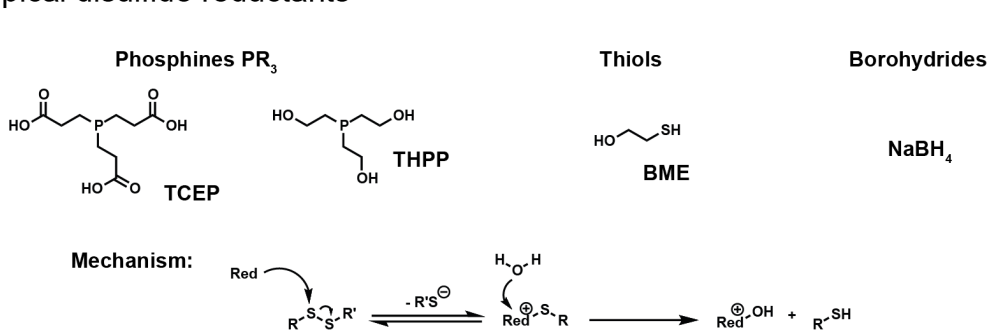
Tests with different candidates for DAs in chapter 3 concluded that compounds with two thiols are more efficient in recovering **LP** from **CP** than compounds with thiol and amine. Therefore, most of the pre-fuel research was focused on dithiol-based DAs, although some studies of amine-thiol pre-fuels were also done (Scheme 4.1, see Appendix 4.5.4).

Keeping in mind the structure of DTT, first, we reviewed thiol protecting groups[15] to find those that can be cleaved in mild aqueous conditions without any reactants that interfere with the peptide-MAD reaction. The options that fit all criteria were: disulfide, thiourea, and S-sulfonate R-SSO_3^- derivatives.[15] The latter one was discarded because of the challenge of putting this protection group on both thiols simultaneously. Thiourea derivative of ethanedithiol was obtained, and its cleavage via hydrolysis was studied (see Appendix 4.5.4.1). Unfortunately, the results were not reproducible. On the contrary, disulfide reduction reactions were well established in the biochemical literature with various suitable conditions.[16–19] The oxidized form of DTT is a cyclic disulfide, *trans*-4,5-dihydroxy-1,2-dithiane (Dithiane, Scheme 4.1), and α-lipoic acid (LA) were the first candidates considered as suitable pre-fuels. Further, a linear disulfide of DTT with methyl mercaptan was also studied. So, next, we focused on finding the right disulfide reductant.

4.2.2 Disulfide reduction

The reduction of disulfide is especially important in studying protein folding, which stimulated research in finding mild but efficient reductants. Disulfide bonds can be reduced using thiols (monothiols like BME, and dithiols like DTT), phosphines (TCEP (tris(2-carboxyethyl)phosphine) or THPP (tris(3-hydroxypropyl)phosphine)), sodium borohydride (Scheme 4.2), and electrochemically.[16,19,20] There are also procedures using sodium in liquid

ammonia and metallic zinc in acidic conditions.[20] Among many chemical methods to reduce disulfides, the most compatible reactants with the peptide-MAD seemed to be phosphines and borohydrides (Scheme 4.2) because thiol-based reductants would interfere with MAD, and metal-based reductions require low pH and release metal ions that could precipitate the peptides.



Scheme 4.2 | Reductants of disulfides used the most in aqueous systems and the general disulfide reduction mechanism. Adapted from ref. 20.

4.2.2.1 Borohydrides

Potassium or sodium borohydride is usually used in excess to reduce disulfides in proteins and then is quenched by acid or acetone once the reduction is complete.[19,21] This method is also used for thiol quantification and preparative synthesis.[20,22,23] To test whether this reductant is suitable to reduce pre-fuel into DA, first the reduction of LA was followed by NMR (Figure 4.3a). In equimolar amounts with $NaBH_4$, LA is reduced to its dithiol form, LA_{red}, in 2 hours (20 mM) with 98% yield. Over a longer time, partial oxidation LA_{red} to LA was observed (Figure 4.3a, decrease of the amount of LA_{red} after 5 hours and longer). The time of LA_{red} accumulation was found reasonable, so next the reaction of LA and sodium borohydride was tracked in the presence of MAD to see how the rate of LA_{red} production is coupled to its reaction with MAD, which is convenient to track by UV-Vis because of the characteristic change of MAD absorbance band. Before using $NaBH_4$ as a reductant of LA, we followed the reaction of MAD with LA_{red} that was reduced by TCEP to see how the absorption spectrum changes upon accumulation of MAD-LA product (Figure 4.3b). Over time, the absorbance increase at 350 nm was observed (Figure 4.3b, inset), indicating the formation of a product with the same peak position but higher molar absorptivity than that of MAD. When in the same measurement setup, $NaBH_4$ was used as a reductant instead of TCEP, the absorbance increase at 350 nm indicated the MAD-LA product formation did not happen (Figure 4.3c). Instead, during four times sequential addition of 1 eq. $NaBH_4$, little absorbance changes at 350 nm were observed (Figure 4.3c, inset). The addition of an excess of $NaBH_4$ (Figure 4.3c, inset, 100 eq. $NaBH_4$) resulted in a rapid decrease of MAD absorbance down to the baseline. Thus, despite the successful reduction of LA, the higher amounts of sodium borohydride correlated with a rapid decrease of the MAD peak at 350 nm, also when borohydride was left with MAD alone. We hypothesized that reduction of MAD occurred as $NaBH_4$ was shown to reduce conjugated esters, aldehydes, and ketones.[24–26] Another issue with using $NaBH_4$ was the practical difficulty of dosage of its stock solution. Because of the constant bubble formation that was especially

active when the stock solution was pipetted, it was impossible to know precisely the added amount of $NaBH_4$.

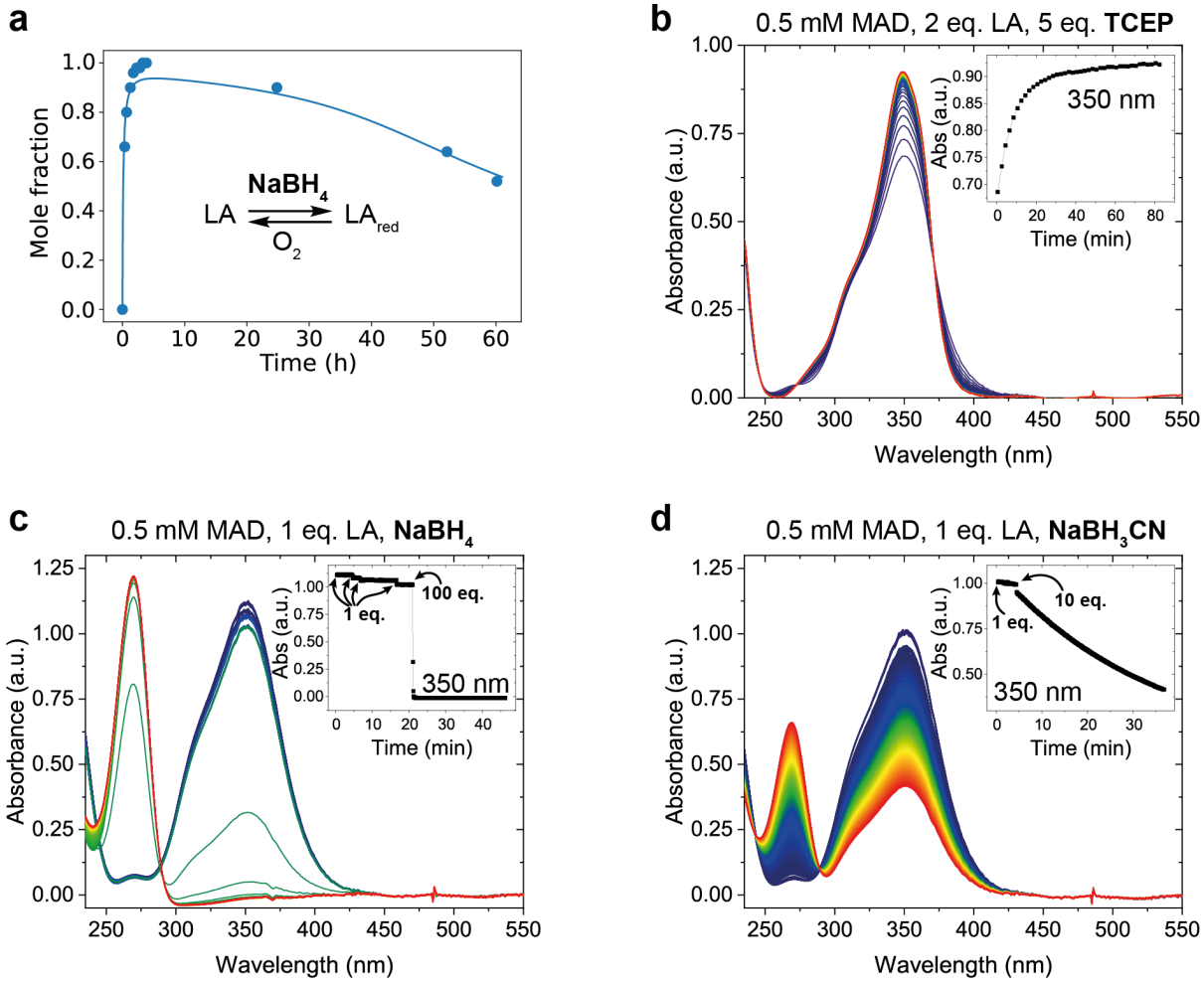


Figure 4.3 | Experiments with borohydrides as disulfide reductants. (a) The reaction of LA with $NaBH_4$ (1:1, 22 mM) was tracked by ^1H NMR in the deuterated buffer. (b) Reduction of LA by TCEP in the presence of MAD, tracked by UV-Vis (absorption). Inset: absorbance change at 350 nm over time, indicating the reaction between MAD and LA and MAD-LA product formation. (c) The reaction of MAD, LA, $NaBH_4$. Instead of increase of the absorbance at 350 nm, it decreases, signifying that MAD-LA product did not form, but MAD is consumed. (b) The reaction of MAD, LA, $NaBH_3CN$. The result is the same as with $NaBH_4$: MAD is consumed, but no MAD-LA product is formed. Inset: the decrease of the absorbance band at 350 nm when 1 eq. $NaBH_3CN$ was added, which went faster when another 10 eq. was added.

We also tried to use sodium cyanoborohydride, $NaBH_3CN$, which was shown to be a milder reductant than borohydride.[27,28] However, its effect on MAD was the same, but slower (Figure 4.3d, inset) and started already when 1 eq. of cyanoborohydride was added, and sped up upon the addition of another 10 eq.

Despite the fact that borohydrides seemed to react with MAD, we still attempted to use $NABH_4$ in the peptide-MAD system as a reductant. At a certain ratio of reactants (0.2 mM **P1**, 3 eq. MAD, 3 eq. LA, 10 eq. $NaBH_4$) throughout the reaction, we observed first a decrease of the absorbance at 350 nm (Figure 4.4, inset) and subsequently an increase, which can only happen if the MAD-LA product forms. An increase of the absorbance at 305 nm can be correlated with the MAD-peptide product (Figure 4.4, spectrum), although the appearance of a pronounced peak at 260–270 nm suggests that MAD is being consumed via a reaction with $NaBH_4$. Overall,

these results were obtained only under these conditions; an increased $NaBH_4$ amount led to the steep decrease of the MAD peak, and using fewer equivalents of borohydride did not result in any later increase of 350-nm absorbance (i.e. no MAD-LA product formed, MAD is consumed by peptide only).

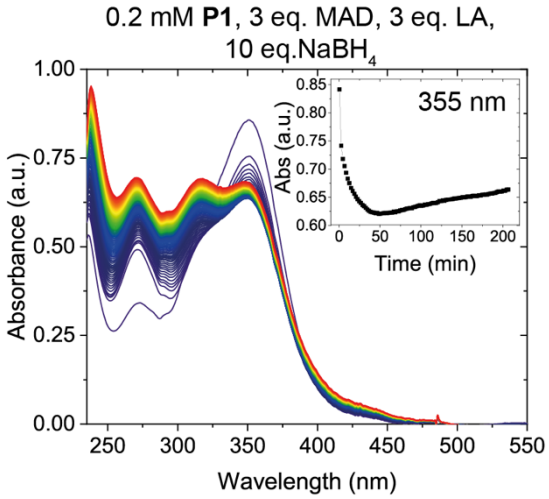
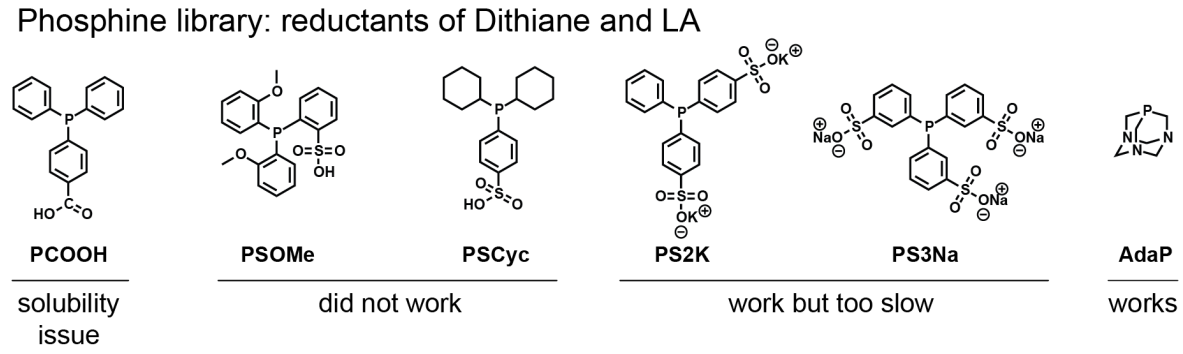


Figure 4.4 | One of the trials to use $NaBH_4$ as disulfide reductant in the peptide-MAD system, tracked by UV-Vis (absorption). Inset: decrease and increase of the absorbance at 350 nm, indicating the consumption of MAD by peptide and $NaBH_4$ in the first hour of the reaction, then the formation of MAD-LA product, signifying that LA_{red} formed and reacted either with free MAD or with MAD from the MAD-peptide products.

In conclusion, borohydrides were not suitable as disulfide reductants because of their reaction with MAD itself. So, we switched to testing other potential reductants that can slowly reduce LA or Dithiane.

4.2.2.2 Phosphines

Phosphine derivatives with various structures have been extensively used for disulfide reduction in organic and aqueous media.[29–34] The most commonly used in water, TCEP and THPP are extremely efficient and fast reductants. These two phosphine agents were tested first to reduce cyclic disulfides for the peptide-MAD system. To facilitate setup, the experiments were done in UV-Vis in the presence of MAD (e.g. see Figure 4.3b), as the rate of decreasing MAD absorbance during the reaction with the reduced dithiol is a good indicator to follow and compare to the rate of the peptide reaction with MAD. Again, we wanted to minimize the reaction of MAD with DA and were looking for a way to delay its release into the reaction mixture. TCEP and THPP were too efficient disulfide reductants for our purposes: Dithiane and LA were reduced within one minute when tested in the reaction mixture with MAD, rate of MAD consumption was not slower than in the reaction with DTT or reduced LA (LA_{red}).
Seeking to find a phosphine with a lower reduction rate, we selected a library of commercially available phosphines to be tested in a reaction with Dithiane or LA and MAD (Scheme 4.3). The aim was to have such a disulfide reduction rate that the reaction of MAD with formed DTT or LA_{red} would be at least twice slower than that of pure DTT/LA_{red} with MAD. The idea behind the phosphines' selection was that electron-poorer substituents would decrease the nucleophilicity of phosphorous, which would make it worse a reductant than TCEP and THPP.[34–36]



Scheme 4.3 | A library of different phosphines, chosen for the Dithiane and LA reduction in the peptide-MAD system. Under the structures and their labels, the summary of the experimental results is given.

All phosphines were screened in the reaction with equimolar amounts of Dithiane and LA (the concentration range of each reactant was 20-30 mM) first by ^{1}H NMR. The first tested phosphine, PCOOH (2-(Diphenylphosphino)benzoic acid), was not soluble in the solvent system used for the (de)cyclization reactions. Next, it was found that PSOMe (2-[Di(2-methoxyphenyl)phosphino]benzenesulfonic acid) and PSCyc (2-(Dicyclohexylphosphino)benzenesulfonic acid) did not yield any amount of LA_{red} during two days of measurement even when the ratio of phosphine to LA was changed to 10 versus to 1 initially. Since the ring in the molecular structure of LA is more strained than that in Dithiane and therefore should be easier to reduce, we did not test these phosphines in the reaction with Dithiane.

Next, PS2K (Bis(p-sulfonatophenyl)phenylphosphine dihydrate dipotassium salt) and PS3Na (Triphenylphosphine-3,3',3"-trisulfonic acid trisodium salt). At that point, another hypothesis was that linear disulfides could be more susceptible to reduction than cyclic ones. Indeed, with cysteamine disulfide and PS2K, a 25 % yield was achieved in the first hour of the reaction (Figure 4.5a), whereas with LA, no change occurred within this time.

The analog of DTT, where each thiol is connected via disulfide bonds with methyl mercaptan (MSDTT, Scheme 4.1), was synthesized and tracked in the presence of PS2K, as well as PSOMe. Conversion of MSDTT to Dithiane was observed (Appendix, Figure 4.11), which was independent of phosphine, as it was later confirmed by following only MSDTT in the solvent system. Interestingly, the disulfide of methyl mercaptan formed from MSDTT was reduced by PS2K within one day but not by PSOMe (Figure 4.12, Figure 4.13). We were surprised to find that PS3Na reduced Ditiane with 15% conversion in 3.5 hours (Figure 4.5b). A better performance was observed with AdaP which gave 70% DTT formation from Dithiane after the same time (Figure 4.5b). Based on this result, AdaP was chosen for the following experiments in the presence of MAD in UV-Vis.

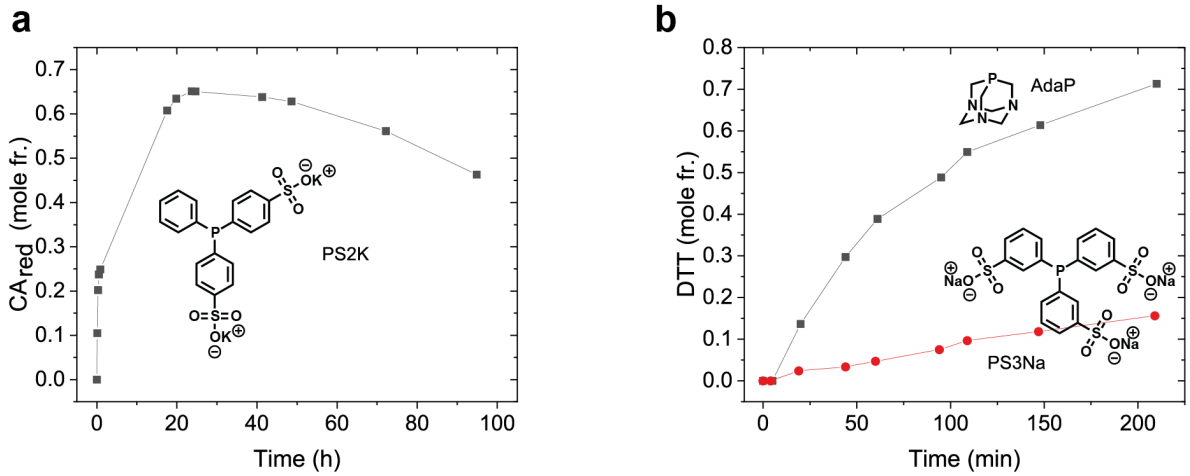


Figure 4.5 | Kinetics of Cysteine disulfide and Dithiane reduction followed by ^{1}H NMR. (a) Reduced cysteamine (CA_{red}) formation from the reduction by PS2K (equimolar amounts) with the subsequent oxidation CA_{red} back to CA disulfide. (b) Reduction of Dithiane by AdaP and PS3Na followed by the yield of DTT. Reactions are done in the solvent mixture with some amount of D_2O.

AdaP was tested as a Dithiane reductant in the presence of MAD (Figure 4.6a), and the change of the reaction mixture absorbance was followed over time by UV-Vis (Figure 4.6b). While the amount of MAD and Dithiane was fixed, the amount of AdaP was varied to understand how much is needed to slow down the rate of MAD-DTT product formation. From UV-Vis data, the usual absorbance at 350 nm (MAD) was tracked (Figure 4.6b). From these experiments, it became apparent that amounts of AdaP between 10 and 100 eq. with respect to MAD and Dithiane would be optimal to delay the release of DTT in the peptide-MAD system. With this, we proceeded to test these conditions with peptides in the context of the peptide-MAD environment.

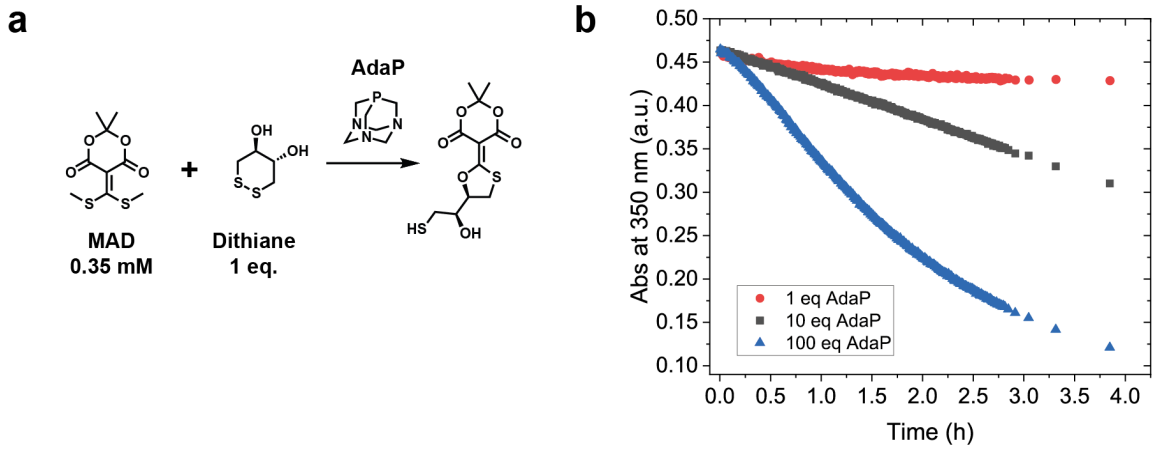


Figure 4.6 | Dithiane reduction by AdaP in the presence of MAD. (a) The reaction scheme indicating the quantities of each compound used in the UV-Vis experiment. (b) UV-Vis tracking of the absorbance at 350 nm in the experiments with a varied amount of AdaP.

4.2.3 Reaction cycle conditions

The first studies in the peptide-MAD system were done by LC-HRMS with a fixed excess of AdaP (100 eq., pseudo-first order conditions), fixed amounts of peptide **P1** and MAD, and a varying amount of Dithiane to understand the impact on the kinetics of CP formation and degradation. Following the amount of **CP1** over time (blue triangles in Figure 4.7a–c), it was possible to see that with 1 eq. of Dithiane (Figure 4.7a), **CP1** peaks at 70% after 20 hours of the reaction and then slowly decreases to 57-58% at 109 hours. With 10 eq. of Dithiane (Figure 4.7b), the decrease of **CP1** amount goes faster, but the transient yield of **CP1** (maximum amount formed during the reaction cycle) is also lower than with 1 eq. Dithiane and is 25%. With 50 eq. Dithiane (Figure 4.7c), the transient yield of **CP1** is only 10%, and only **LP1** is present in the system after 50 hours.

Next, the amount of AdaP was fixed at 50 eq. and the amount of Dithiane was varied again (Figure 4.7d–e). With 50 eq. Dithiane (Figure 4.7d), the transient yield of **CP1** was about 17%, which is 10% more than in the experiment with the same amount of Dithiane but with 100 eq. AdaP (Figure 4.7c). With 10 eq. Dithiane (Figure 4.7d), the transient yield of CP1 was 30%, which is 5% more than in the experiment with an equimolar amount of Dithiane and 100 eq. AdaP (Figure 4.7b).

Clearly, the best reactant ratio where the transient yield of **CP1** is maximized is around 10 eq. Dithiane and 50 eq. AdaP with respect to **P1** and MAD. However, maximizing the CP yield comes at the expense of the reaction time: the higher the yield, the longer it takes for the quantitative recovery of the linear peptide. Remarkably, no **MP1M** was observed in all these experiments. It gives the hope that increasing the amount of MAD to 2–5 eq. and keeping Dithiane and AdaP at excess might help improve the transient yield of **CP1** without extending the **LP1** recovery time.

Finally, we also tried to use PS2K and PS3Na as phosphines for Dithiane reduction in the **P1** reaction cycle (Appendix, Figure 4.14). However, the experiments showed no recovery of **LP1**. Therefore, we proceeded with AdaP as a reductant.

Dithiane was also tested with **P2**. The transient yield of **CP2** was similar (30%) in the experiments with 10 eq. Dithiane, 100 eq. AdaP (Figure 4.8a), and 50 eq. Dithiane, 50 eq. AdaP (Figure 4.8b). In the case of **P2**, there is a higher potential to increase the transient **CP2** yield by increasing the amount of MAD to 5–10 eq. because **MP2M** will not form in these conditions. The experiments with an increased amount of MAD will be the next step in optimizing **P1** and **P2** reaction cycles.

LA was tested as a pre-fuel so far only with **P2**. Comparing experiments with 50 eq. AdaP and 50 eq. Dithiane or LA (Figure 4.8b–c), it is possible to see that LA is slightly more efficient in removing **CP2** in these conditions: the transient yield of **CP2** with LA is 23%, whereas, with Dithiane, it is almost 30%. Also, at 9.5 hours of reaction, the **CP2** amount is 9% in the experiment with LA, and with Dithiane at this time, there is 15% of **CP2**. The factor that might influence this difference is that LA is a five-membered cyclic disulfide, whereas Dithiane has a six-membered ring. The LA ring is more strained than that of Dithiane and is easier to reduce, which results in a faster release of LA_{red} than DTT in the reaction system.

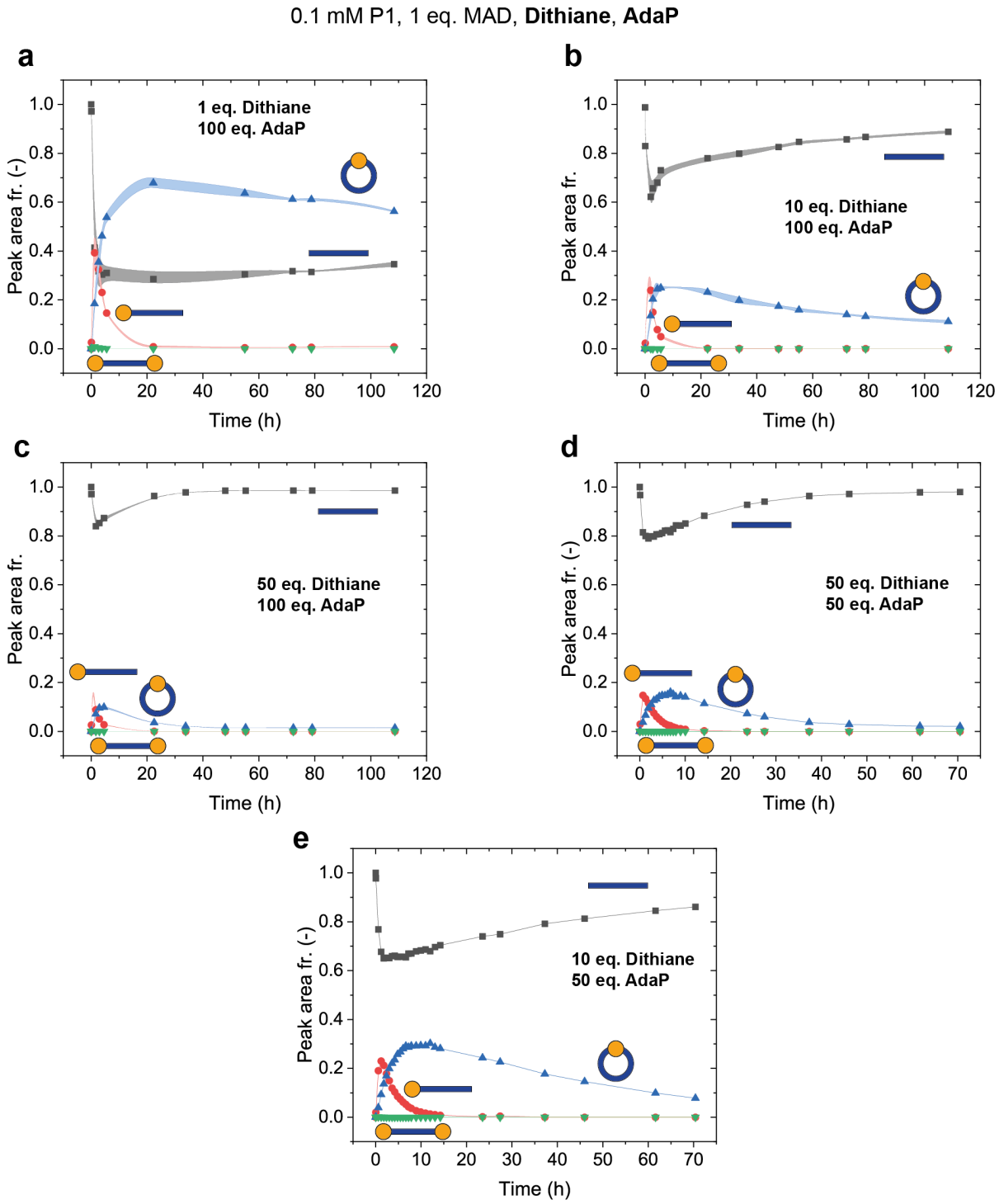


Figure 4.7 | Reaction cycle studies of the P1-MAD system with varying amounts of Dithiane and AdaP by LC-HRMS. The concentration of **P1** and MAD are kept the same for all experiments. Amounts of Dithiane and AdaP are indicated in each graph. The shaded area in panels a–c corresponds to standard deviations derived from duplicates.

Finally, the reaction cycle with **P2** and LA was also followed by UV-Vis (Figure 4.8d). The absorbance at 350 nm goes via an inflection point—a moment when the absorbance of forming MAD-LA overpowers the absorbance of MAD, which is seen as a first decrease followed by an increase in absorbance. Comparing the **P2** experiment with 50 eq. LA and 50 eq. AdaP (red circles in Figure 4.8d) with the identical experiment followed by LC-HRMS, it is interesting to notice that the inflection point in UV-Vis corresponds by time to the peak in **CP2** yield.

Further analysis of UV-Vis spectra of the reaction cycles in correspondence with LC-HRMS measurements at different conditions is needed to conclude whether the correlation of the inflection point with the highest transient amount of CP can be confirmed.

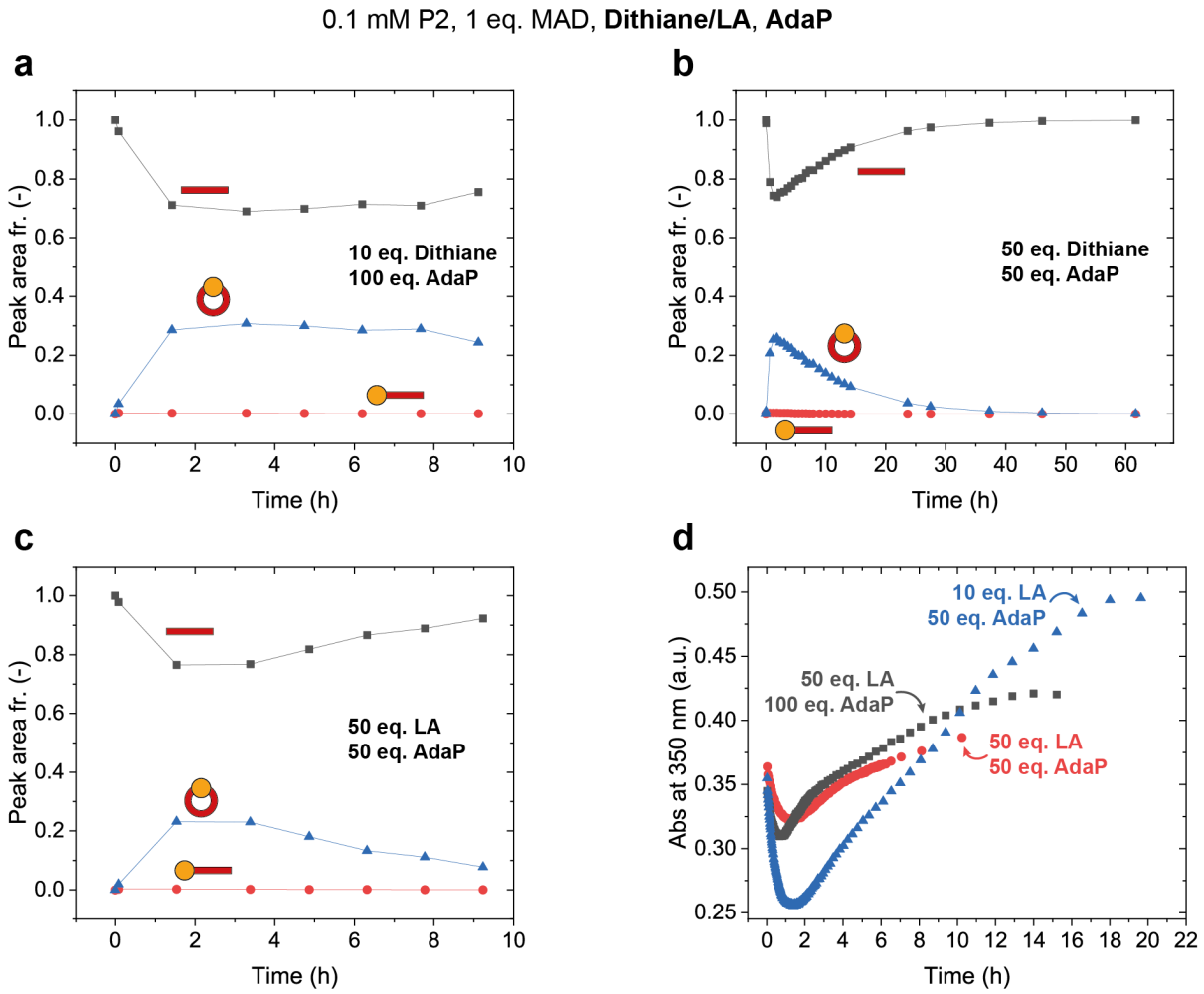


Figure 4.8 | Reaction cycle studies of the P2-MAD system with varying amounts of Dithiane or LA and AdaP. The concentration of **P2** and MAD are kept the same for all experiments. Amounts of Dithiane/LA and AdaP are indicated in each graph. (a–b) LC-HRMS kinetics of the reaction cycle with Dithiane. (c) LC-HRMS kinetics of the reaction cycle with LA. (d) UV-Vis studies tracked absorbance at 350 nm of the reaction cycle with different amounts of LA and AdaP.

So far, the only drawback of using AdaP as a reductant has been the formation of the **LP** functionalized with imine at the N-terminus, which happens for both peptides. These could occur if AdaP partially or completely hydrolyzes, releasing formaldehyde that reacts with amines to form imines. Thus, the total amount of linear peptide presented in Figure 4.7 and Figure 4.8 contained up to 60% of **LP** with N-terminal imine. This should be considered if several sequential reaction cycles are run because the presence of imine might slow down the **MP** formation from **LP** once the next batch of MAD is added.

4.3 Conclusions

The previous chapter concluded that it would not be possible to simply combine cyclization and decyclization processes in a reaction cycle with transient CPs because of the mutual consumption of MAD and DA. To overcome this issue, in this chapter, we introduced an additional reaction in the peptide-MAD system that slowly produces DA *in situ* from so-called pre-fuel giving time for the cyclization reaction to evolve and yield some **CP** before converting them back to **LP**. As the most suitable DA are dithiols, the most accessible form of pre-fuel was found to be disulfide, although other designs were considered. The oxidized form of DTT (Dithiane) and lipoic acid (LA) were chosen as pre-fuels. Disulfide reduction with well-known phosphine reductants is a fast and efficient reaction, which is the opposite of what we need. Therefore, different phosphines were tested to identify the least reactive ones to check their performance with Dithiane/LA, MAD, and peptides. AdaP was the phosphine derivative that yielded the highest transient **CP** yields (30% **CP1** and 27% **CP2**). At the same time, no significant difference has been observed depending on whether Dithiane or LA is used. Using this approach, we have shown the first CRN with the transient formation of CPs, where the linear peptide is recovered once the system reaches equilibrium.

Outlook

This work opens the perspective of creating transient peptide libraries, where different short linear peptides that form CP with several mixed peptide-MAD units are introduced into the reaction network (Figure 4.9a). In that case, CPs that self-assemble faster and/or are more stable might serve as a template for still forming CPs. This concept has been widely applied in dynamic combinatorial libraries.[37] The transient peptide assemblies have been shown based on carbodiimide-fueled chemistry by Boekhoven and co-workers.[38] Preliminary studies on cyclization reaction with short tripeptides (Cys-X-Cys) showed different kinetics and product distribution depending on the second amino acid X in the sequence (Figure 4.9b,d). the following reactions were performed in 0.1 M borate buffer (pH 9.2). We increased the concentration of peptides to 10 mM and reduced the molar ratio of peptide and MAD to 1:1. The increased concentration of building blocks accelerates the reaction, favoring the formation of large macrocycles. The equimolar amount of MAD reduces the possibility of forming (MP)nM linear products, which prevents cyclization. The formation of large rings in this system is indeed possible (Figure 4.9c). After optimization of conditions (reactant ratio, buffer, and concentration), we observed the formation of larger cycles which was not seen before in this system. Their formation was especially pronounced in the case of 1, which formed C_{12} cycles (see Figure 4.9c). In the case of 3 also, we could detect the C_2 product, which we could not observe earlier.

Future work will include tracking fueled formation, self-assembly, and degradation of CPs containing several peptide-MAD units. The correlation between cycle size, composition, self-assembling properties, and relative abundance will be studied in detail in comparison with non-fueled conditions.

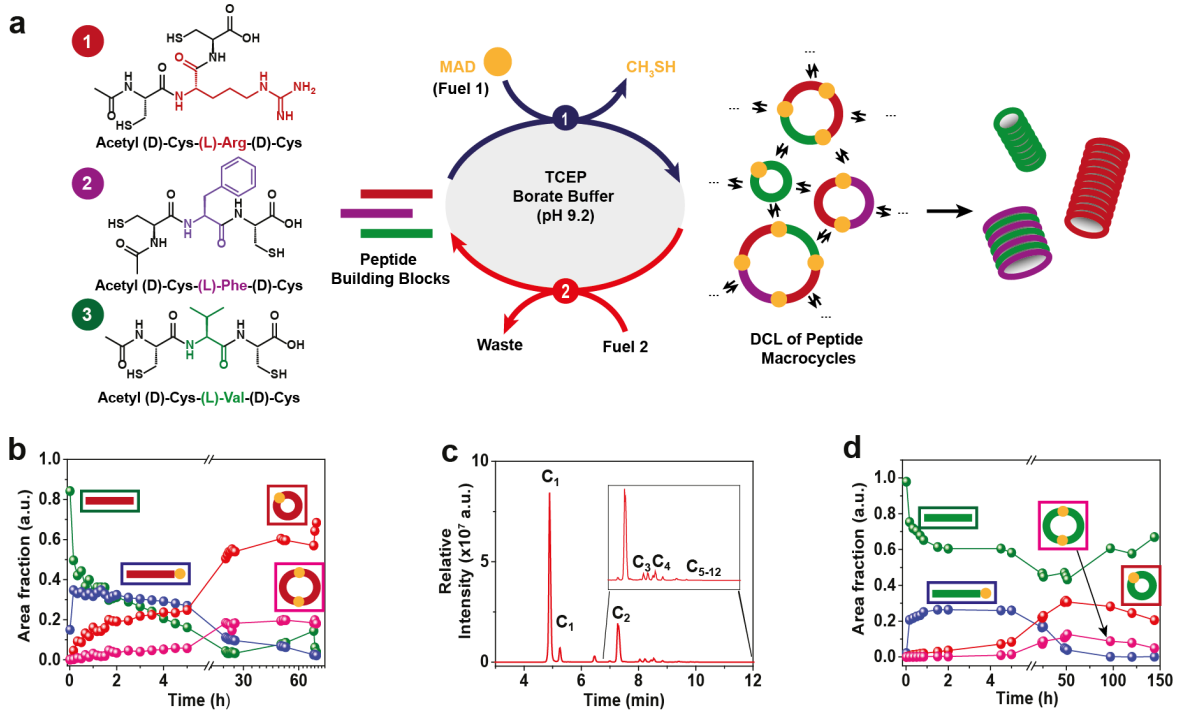
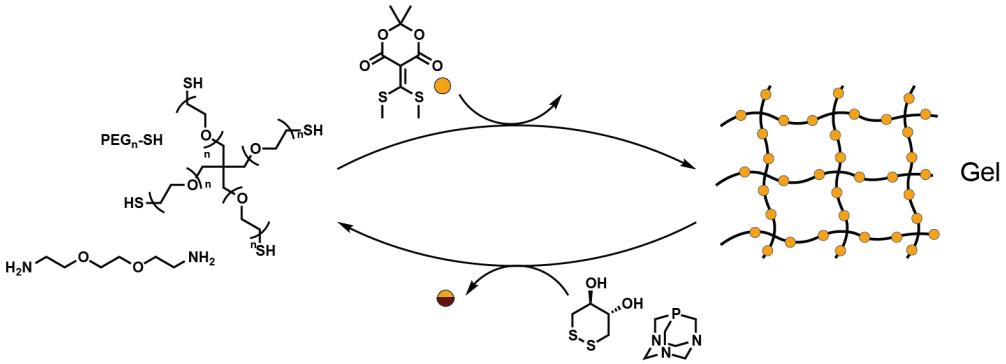


Figure 4.9 | Cyclization of tripeptides with MAD (1:1) at pH 9.2. Reaction progress of MAD-driven cyclization of (b) 1 (d) 3 was monitored using LC-HRMS. c) HPLC chromatogram of reaction 'b' showing separated peaks for cyclic peptides of varying sizes.

Further, as a part of different project, the proposed MAD-Dithiane-AdaP reaction network could possibly be implemented in PEG-based substrates instead of peptides, which would yield transient cross-linked material (Scheme 4.4). Overall, our approach seems versatile and applicable to a range of synthetic and, or biological systems.



Scheme 4.4 | A transient cross-linked network based on the MAD-Dithiane-AdaP chemistry.

4.4 Authors contributions

Anastasiia Sharko did all the measurements and pre-fuels testing. Oleksandr Shyshov synthesized ethanedithiol thiourea derivative and helped to test it. Benjamin Klemm, a Ph.D. student at TU Delft (Eelkema's group), suggested trying AdaP. Shana Shirin Valapra did studies on peptide libraries idea during her M2 internship. Thomas Hermans supervised the research.

4.5 Appendix

Chemicals and Materials. Chemicals and solvents were purchased from TCI chemicals, Sigma Aldrich, Fisher. For the reaction system, the general solvent mixture composition was 1.6 % DMF, 16.7 % acetonitrile, and 81.7 % buffer. A VWR Model 110 pH meter was used to record the pH measurements during buffer preparations.

UV-Vis. UV experiments were performed on an Agilent CARY 8454-UV spectrometer in 1–2 mm quartz cuvettes. Wavelength range 235–600 nm or 190–1100 nm.

LC-HRMS. HRMS was performed with a Thermofisher Scientific UltiMate 3000 RSLCnano UHPLC System coupled with EMT Thermo OrbiTrap Mass analyzer and the ESI ionization type. Hypersil GOLD column, 50 x 2.1 mm, 1.9 µm was used for the LC part.

4.5.1 MSDTT

MSDTT was synthesized according to the literature procedure.[39,40]

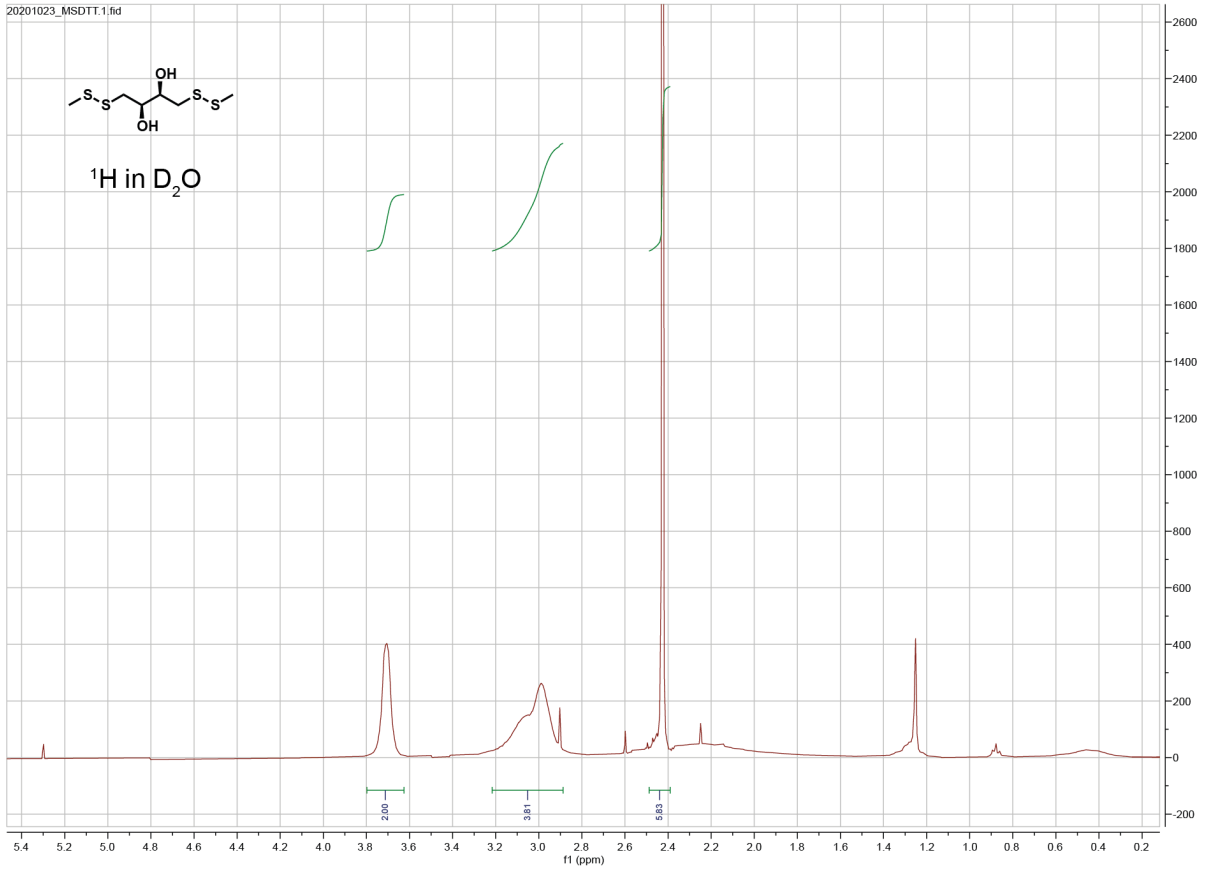


Figure 4.10 | ^{1}H NMR spectrum of MSDTT in D_2O.

4.5.2 MSDTT with phosphines

When kept in the buffered solvent mixture, MSDTT slowly converts into Dithiane with the formation of methyl mercaptan disulfide (MSSM, see Figure 4.11 and Figure 4.12).

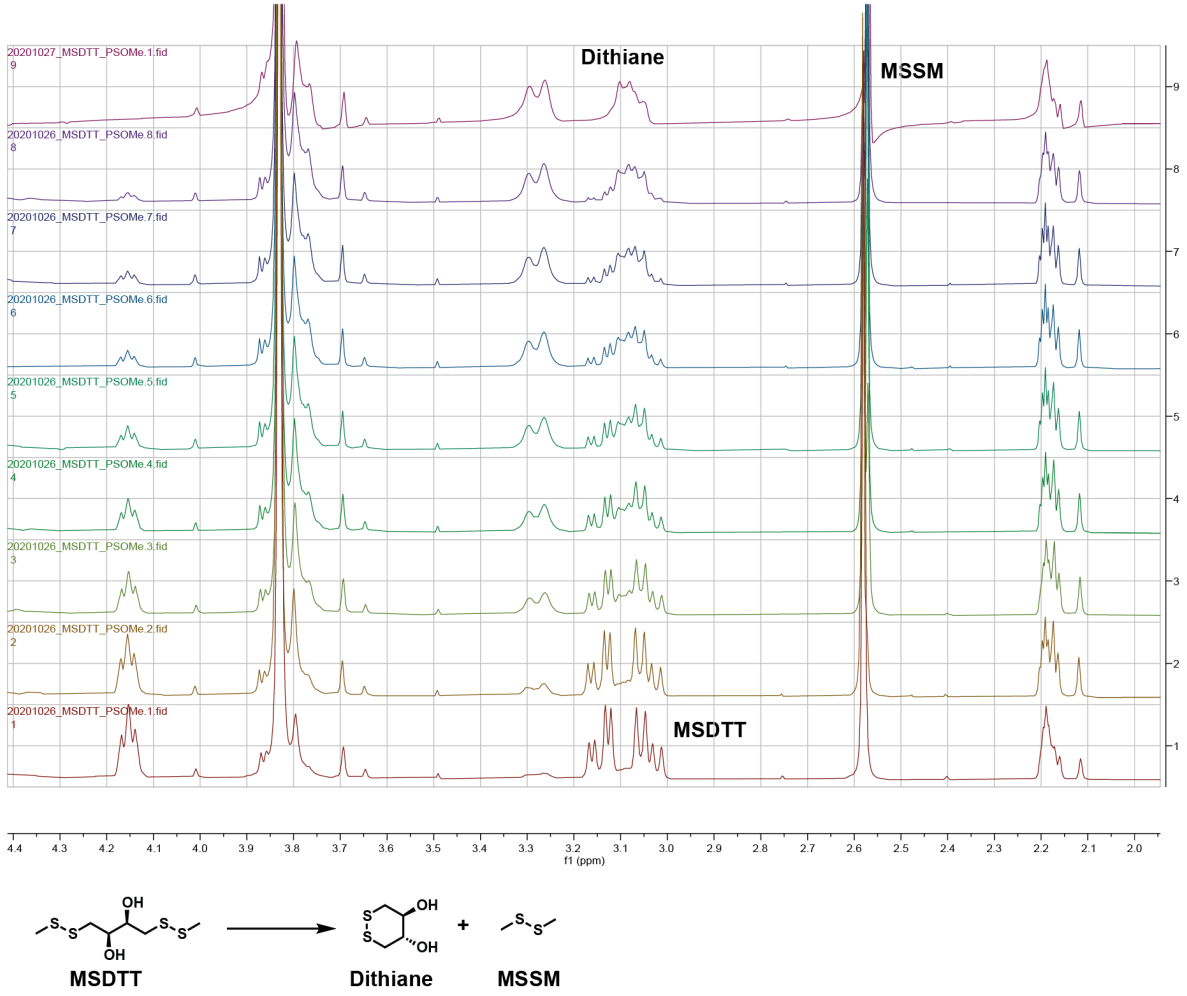


Figure 4.11 | MSDTT was followed in the presence of PSOMe by ^{1}H NMR.

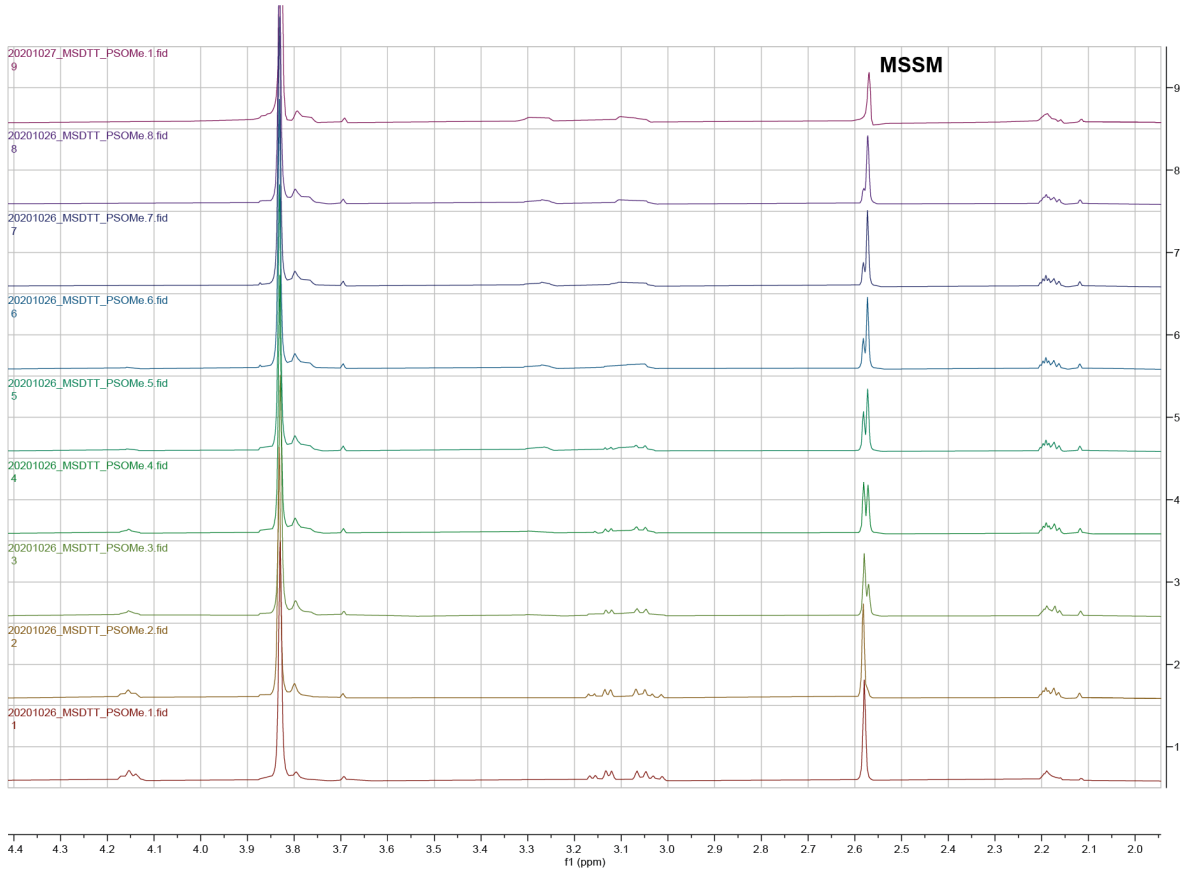


Figure 4.12 | MSSM formation in the experiment of MSDTT with PSOMe.

Interestingly, in the experiment with MSDTT and PS2K, the reduction of MSSM took place with the formation of free methyl mercaptan (MSH, Figure 4.13).

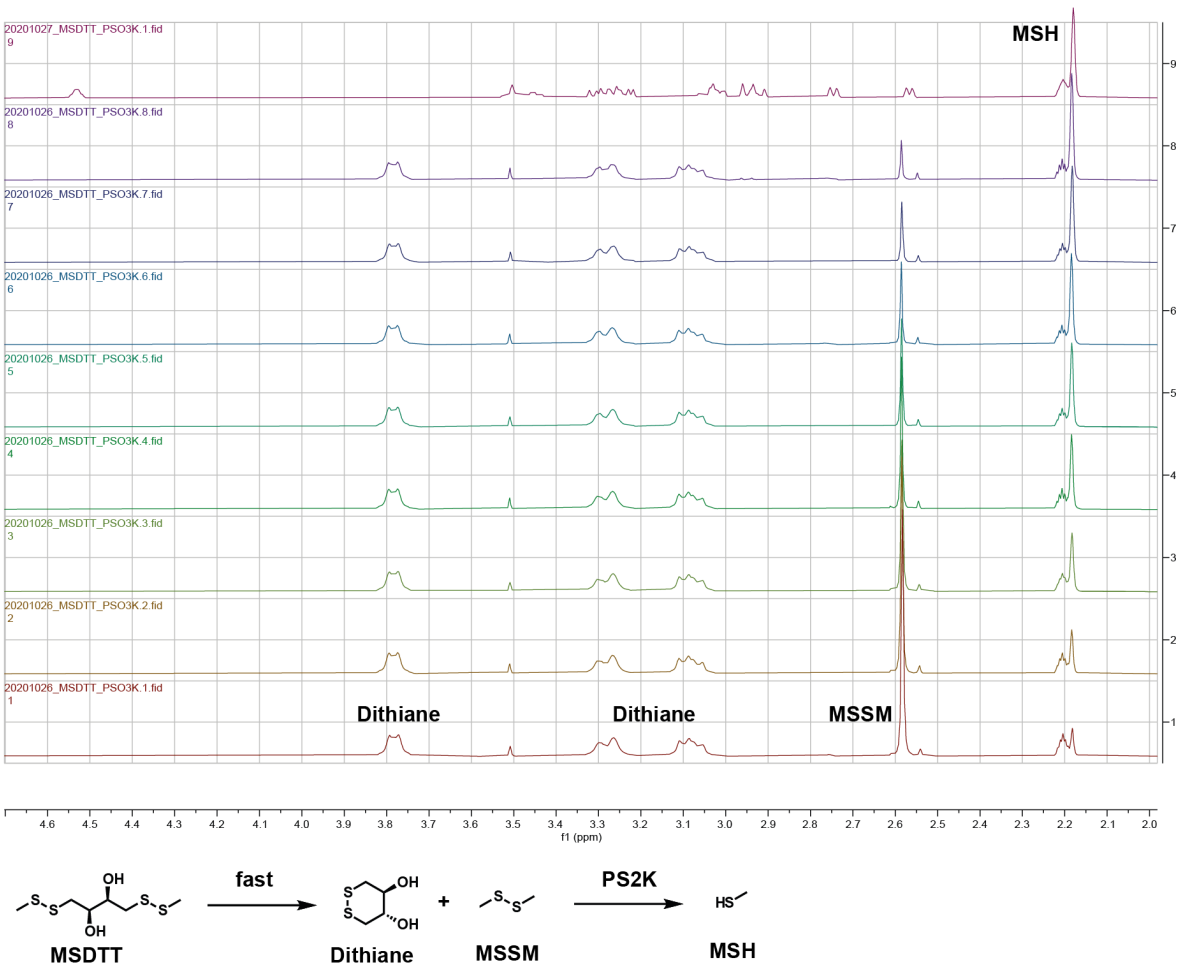


Figure 4.13 | MSDTT was followed in the presence of PS2K by ^{1}H NMR. The reduction of MSSM to MSH was observed.

4.5.3 P1 and aryl phosphines

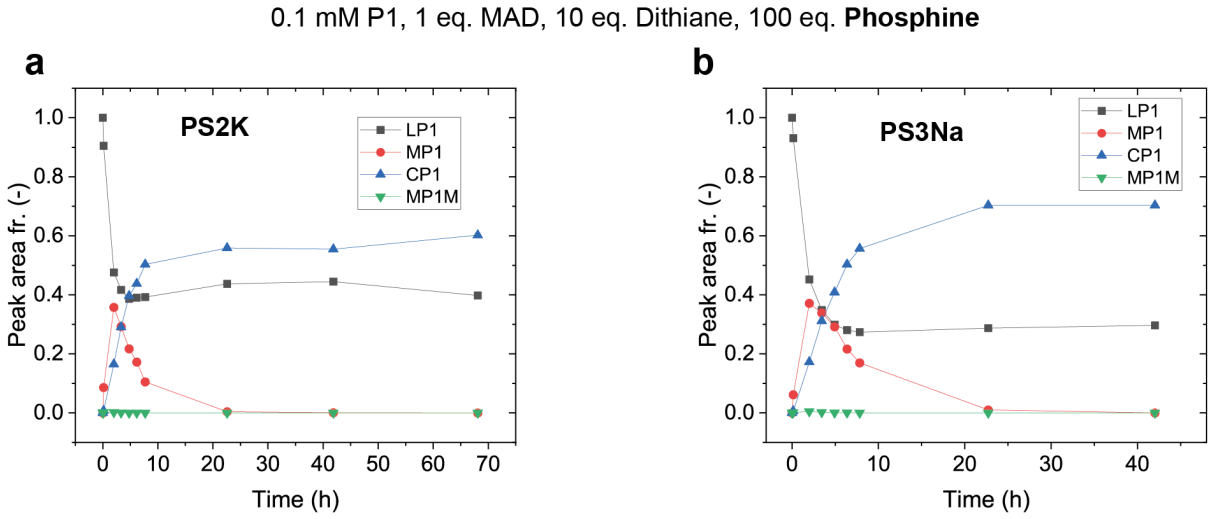


Figure 4.14 | Reaction cycle studies of the P1-MAD system with PS2K and PS3Na as Dithiane reductants by LC-HRMS.

The recovery of **LP1** was not observed in the presence of PS2K and PS3Na.

4.5.4 Other pre-fuel trials and ideas

This section contains other pre-fuels tried or considered that either did not work at all or had certain issues with the implementation into the peptide-MAD system.

4.5.4.1 Thiourea derivative of ethanedithiol

The idea to use thiourea derivative as pre-fuel appeared along with S-sulfonates and thioesters (Figure 4.15a). The latter were not considered further because of the possible exchange of thioesters of pre-fuel with cysteine of a peptide, which might hamper the cyclization process. A similar problem appears with S-sulfonates that can transfer the sulfonate group to the peptide cysteine. Therefore, we proceeded with thioureate.

The synthesis of the thiourea derivative of ethanethiol was done in one step (Figure 4.15b) according to a literature procedure.[41] The compound was tested in the reaction with MAD at different pH (Figure 4.15c–e) by tracking the absorbance change at 350 nm and 283 nm. The hydrolysis of thiourea and its reaction with MAD appeared to be very pH sensitive. Unfortunately, the results were also not always reproducible and seemed to be buffer-dependent, which made us quit this idea. For more detailed studies, synthesis of thiourea of a water-soluble derivative of ethanedithiol is advised to ensure solubility of MAD-DA product.

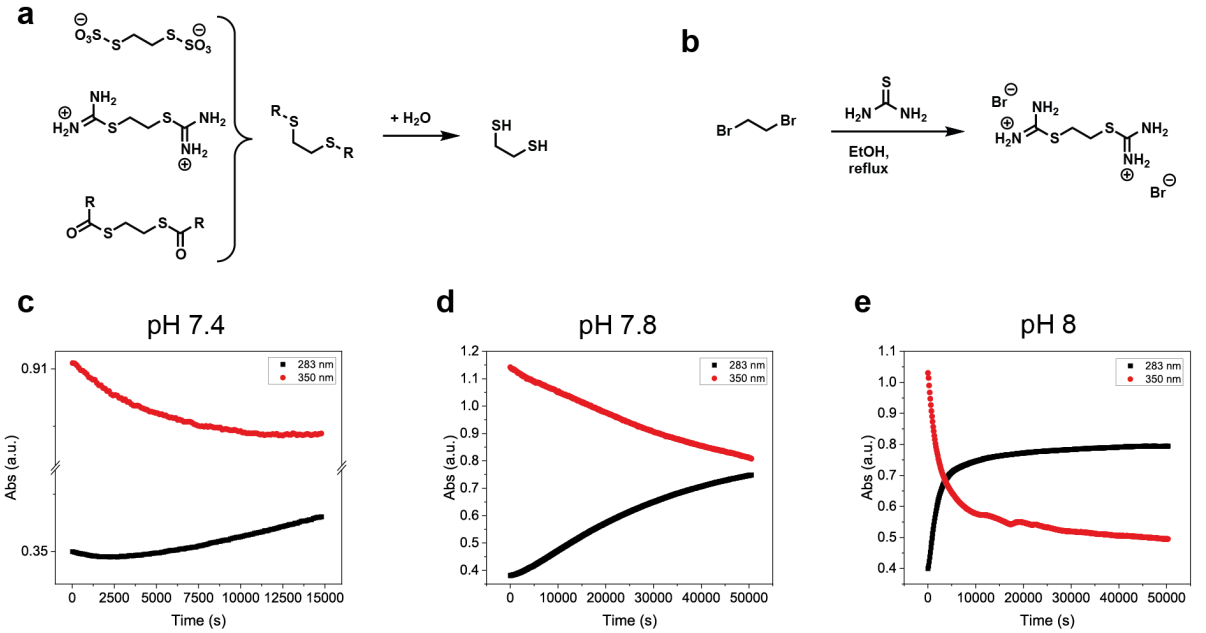
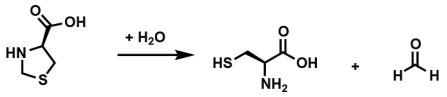


Figure 4.15 | Thioureate of ethanedithiol as a pre-fuel. (a) A scheme illustrating the concept of DA release by hydrolysis. (b) Synthesis of thioureate derivative of ethanedithiol. (c–e) UV-Vis time traces of the absorbance at 350 and 283 nm to study hydrolysis kinetics of thioureate in the presence of MAD.

4.5.4.2 Thiazolidine



Scheme 4.5 | The concept of the cysteine release from thiazolidine.

Thiazolidine of L-cysteine was chosen because it can release cysteine during hydrolysis (Scheme 4.5).[42] When this is the case in the presence of MAD, MAD is consumed through the reaction with released cysteine, and this process can be tracked by UV-Vis by the decrease of

absorbance at 350 nm. The rate of this reaction is much slower than that between peptides and MAD. If cysteine was able not just to open **CP** into **Cys-MP** (see chapter 3) but also recover LP, thiazolidine would make a suitable pre-fuel. Interestingly, when mixed simultaneously with MAD and **P1**, thiazolidine influences the formation of peptide-MAD products, which can be seen by the time trace of absorbance at 305 nm, which first goes up (correlating with the formation of peptide-MAD products), and then somewhat decreases. Since tracking the reaction by HRMS did not show a decrease in **CP1**, probably, a change in UV-Vis can correspond to the formation of **Cys-MP1**.

4.5.4.3 DL-Homocysteine thiolactone

Homocysteine thiolactone combines electrophilic and nucleophilic properties due to thioester and amine moieties. Susceptible to hydrolysis, the five-membered ring can open, releasing the thiol group (Figure 4.16a). Because of this property, this thiolactone was tested as a pre-fuel, although if homocysteine could recover LP, thiolactone would require additional protection for the amino group. Unfortunately, testing it stepwise with the crude solution of CP1 did not result in the disappearance of the peptide-MAD aggregates. However, the yellowish color of the crude CP1 solution disappeared after 40 minutes. After 3–5 days of observation, a new colorless crystalline precipitate was seen to form in all samples with homocysteine thiolactone: in the crude **CP1** sample, in a control vial with TCEP, and a vial with **LP1**. This was not observed in any of the samples without thiolactone. A possible explanation of this phenomenon lies in its high reactivity, which can result in self-polymerization when the amino group of one thiolactone molecule attacks the carbonyl of another.[43] This causes ring-opening and the formation of the product that can continue the same reaction with other thiolactone molecules (Figure 4.16b). Another disadvantage of thiolactones for the peptide-MAD system is their reactivity toward amino groups of peptides, which are used in the bioconjugation field to modify proteins and add thiol functional units (Figure 4.16c).[44]

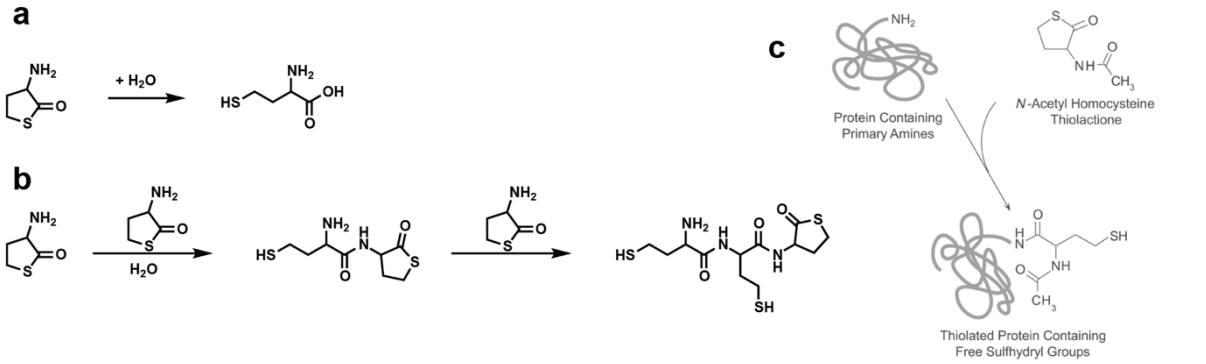
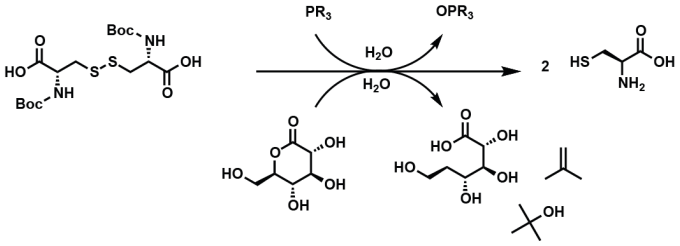


Figure 4.16 | Homocysteine thiolactone used as a pre-fuel. (a) Release of homocysteine via hydrolysis. (b) Self-polymerization of homocysteine thiolactone. (c) The reaction used for the modification of proteins and introduction of thiol-containing residues. Adapted from ref. 44.

4.5.4.4 Boc-protected cystine

Boc-protected cystine was tested along with previous amine-containing pre-fuels. Tert-butoxycarbonyl protecting group (Boc) on amine can be cleaved in acidic conditions. We tested whether gluconic acid produced from the hydrolysis of gluconolactone could remove Boc in

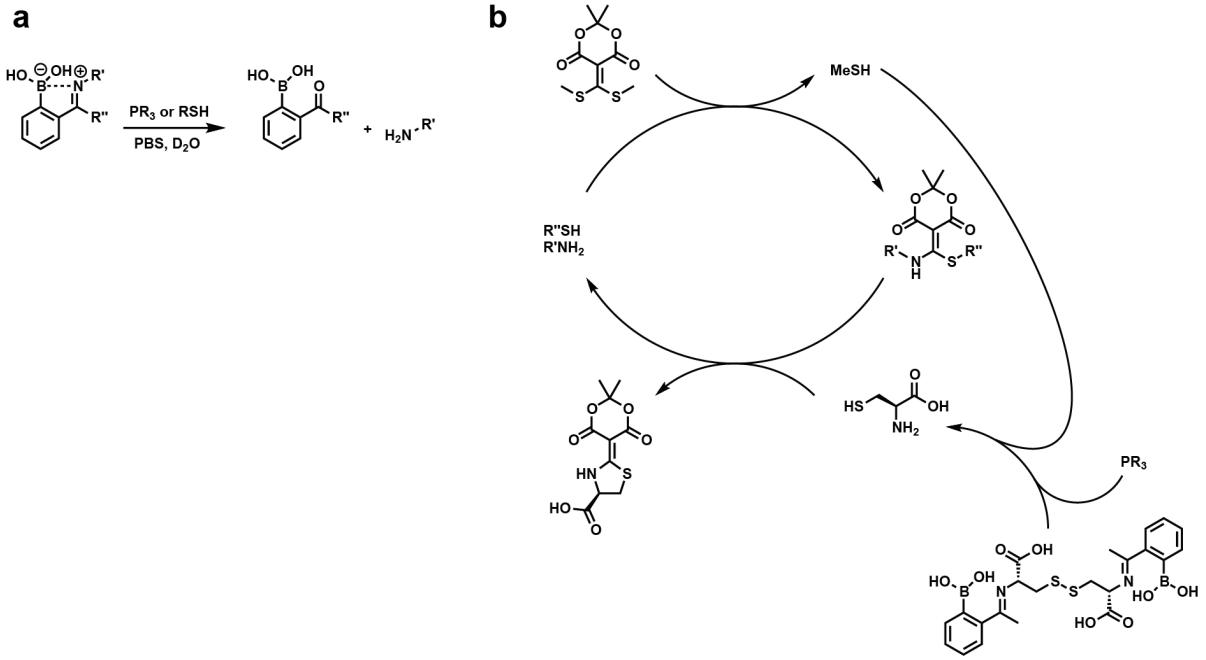
the conditions used for the peptide-MAD system (Scheme 4.6). The deprotection, however, did not occur, probably due to the buffered solution.



Scheme 4.6 | Boc-cystine deprotection idea to release cysteine.

4.5.4.5 Iminoboronates

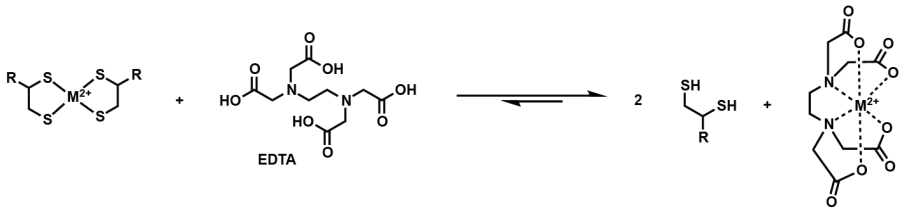
One example of pre-fuel that has not been tried but might be useful for developing a MAD-based reaction cycle with smaller molecules than **P1** or **P2** so that cysteine and similar amine-thiol-containing molecules can act as backward fuels. In the design of pre-fuel, in this case, protection for both amine and thiol is needed. While thiol can be in the form of disulfide, the amino group can be protected by iminoboronate substituent (Scheme 4.7a). The special feature of this protecting group is that its cleavage can be promoted by thiol and phosphines (TCEP and AdaP).[45] For the MAD-based reaction cycle, this means that methyl mercaptan formed in the forward reaction will initiate cleavage of the iminoboronate group from the pre-fuel, e.g. functionalized cystine. AdaP, present in the system, will speed up this process and, in parallel, will reduce disulfide (Scheme 4.7b). A combination of methyl mercaptan and AdaP actions will cause a release of cysteine, which will drive the system back to equilibrium.



Scheme 4.7 | Iminoborobates as an amine protecting group for the amine-containing DA. (a) Cleavage of iminoboronates. (b) A proposed reaction cycle with substrates that can be decoupled from MAD using cysteine as DA and iminoboronate-protected cystine as a pre-fuel.

4.5.4.6 Metal-chelated dithiols or phosphines

An idea to use a metal complex with a dithiol as a ligand was inspired by unrelated studies that show disulfide reduction by ascorbic acid catalyzed by copper ions.[46,47] (Note: a later study showed that copper presence does not influence disulfide reduction kinetics, and to reduce linear disulfide with ascorbic acid, at least 100-fold excess of it is needed to achieve 1% conversion in 0.5 h)[48,49] According to the pre-fuel idea, a chelated dithiol was supposed to be released in the presence of another, better ligand due to the slow ligands' exchange and formation of the new complex (Scheme 4.8). Similarly, a chelated TCEP, for example, would reduce disulfides slower or only upon release via exchange. Divalent metal (Zn, Cd, Ni, Cu) complexes with DTT and TCEP are described in the literature as well, as the impact of DTT complexation on its reduction of disulfide bonds in proteins.[50–53] Preliminary tests with DTT complexes with **P1** caused the immediate formation of an opaque colorless precipitate that could only redissolve upon adding an excess EDTA. The following experiments were not conducted. However, the presence of Zn(II) with DMPS (dimercaptopropanesulfonate) slowed down the rate of its reaction with MAD by factor two compared to the free DMPS. The presence of Zn(II) with TCEP did not impact the reduction of Dithiane and the subsequent reaction with MAD.



Scheme 4.8 | The concept of dithiol release from its metal complex is based on the slow ligand exchange reaction.

4.6 References

1. Mondal, S. & Haldar, D. A transient non-covalent hydrogel by a supramolecular gelator with dynamic covalent bonds. *New J. Chem.* **45**, 4773–4779 (2021).
2. Mondal, S., Podder, D., Nandi, S. K., Roy Chowdhury, S. & Haldar, D. Acid-responsive fibrillation and urease-assisted defibrillation of phenylalanine: a transient supramolecular hydrogel. *Soft Matter* **16**, 10115–10121 (2020).
3. Heuser, T., Weyandt, E. & Walther, A. Biocatalytic Feedback-Driven Temporal Programming of Self-Regulating Peptide Hydrogels. *Angew. Chem. Int. Ed.* **54**, 13258–13262 (2015).
4. Tena-Solsona, M. & Boekhoven, J. Dissipative Self-Assembly of Peptides. *Israel Journal of Chemistry* **59**, 898–905 (2019).
5. Dai, K., Tena-Solsona, M., Rodon-Fores, J., Bergmann, A. & Boekhoven, J. Morphological transitions in chemically fueled self-assembly. *Nanoscale* (2021) doi:10.1039/D1NR04954B.
6. Debnath, S., Roy, S. & Ulijn, R. V. Peptide Nanofibers with Dynamic Instability through Nonequilibrium Biocatalytic Assembly. *J. Am. Chem. Soc.* **135**, 16789–16792 (2013).
7. Sheehan, F. *et al.* Peptide-Based Supramolecular Systems Chemistry. *Chem. Rev.* (2021) doi:10.1021/acs.chemrev.1c00089.
8. Dai, K. *et al.* Regulating Chemically Fueled Peptide Assemblies by Molecular Design. *J. Am. Chem. Soc.* **142**, 14142–14149 (2020).
9. Sorrenti, A., Leira-Iglesias, J., Sato, A. & Hermans, T. M. Non-equilibrium steady states in supramolecular polymerization. *Nat Commun* **8**, 15899 (2017).
10. Miao, X. *et al.* Kinetic Selection in the Out-of-Equilibrium Autocatalytic Reaction Networks that Produce Macrocyclic Peptides. *Angewandte Chemie* **133**, 20529–20538 (2021).

11. Pappas, C. G. *et al.* Dynamic peptide libraries for the discovery of supramolecular nanomaterials. *Nature Nanotech* **11**, 960–967 (2016).
12. Frederix, P. W. J. M. *et al.* Exploring the sequence space for (tri-)peptide self-assembly to design and discover new hydrogels. *Nature Chem* **7**, 30–37 (2015).
13. Singh, N., Lainer, B., Formon, G. J. M., De Piccoli, S. & Hermans, T. M. Re-programming Hydrogel Properties Using a Fuel-Driven Reaction Cycle. *J. Am. Chem. Soc.* **142**, 4083–4087 (2020).
14. Singh, N., Lopéz-Acosta, Á., Formon, G. & Hermans, T. Chemically Fueled Self-sorted Hydrogels. (2021) doi:10.33774/chemrxiv-2021-n8lt5.
15. Protection for the Thiol Group. in *Greene's Protective Groups in Organic Synthesis* 837–894 (John Wiley & Sons, Ltd, 2014). doi:10.1002/9781118905074.ch06.
16. S. N. Mthembu, A. Sharma, F. Albericio, B. G. de la Torre. Breaking a Couple: Disulfide Reducing Agents. *ChemBioChem* **21**, 1947–1954 (2020).
17. Getz, E. B., Xiao, M., Chakrabarty, T., Cooke, R. & Selvin, P. R. A Comparison between the Sulfhydryl Reductants Tris(2-carboxyethyl)phosphine and Dithiothreitol for Use in Protein Biochemistry. *Analytical Biochemistry* **273**, 73–80 (1999).
18. Singh, R. & Whitesides, G. M. A reagent for reduction of disulfide bonds in proteins that reduces disulfide bonds faster than does dithiothreitol. *J. Org. Chem.* **56**, 2332–2337 (1991).
19. Jocelyn, P. C. Chemical reduction of disulfides. in *Methods in Enzymology* vol. 143 246–256 (Elsevier, 1987).
20. Winther, J. R. & Thorpe, C. Quantification of thiols and disulfides. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1840**, 838–846 (2014).
21. Duane Brown, W. Reduction of protein disulfide bonds by sodium borohydride. *Biochimica et Biophysica Acta* **44**, 365–367 (1960).
22. Stahl, C. R. & Siggia, Sidney. Determination of Organic Disulfides by Reduction with Sodium Borohydride. *Anal. Chem.* **29**, 154–155 (1957).
23. Eom, T. & Khan, A. Disulfides as mercapto-precursors in nucleophilic ring opening reaction of polymeric epoxides: establishing equimolar stoichiometric conditions in a thiol–epoxy 'click' reaction. *Chem. Commun.* **56**, 7419–7422 (2020).
24. Brown, M. S. & Rapoport, H. The Reduction of Esters with Sodium Borohydride1. *J. Org. Chem.* **28**, 3261–3263 (1963).
25. Johnson, M. R. & Rickborn, B. Sodium borohydride reduction of conjugated aldehydes and ketones. *ACS Publications* https://pubs.acs.org/doi/pdf/10.1021/jo00829a039 (2002) doi:10.1021/jo00829a039.
26. Aramini, A., Brinchi, L., Germani, R. & Savelli, G. Reductions of α,β-Unsaturated Ketones by NaBH4 or NaBH4 + CoCl2: Selectivity Control by Water or by Aqueous Micellar Solutions. *European Journal of Organic Chemistry* **2000**, 1793–1797 (2000).
27. Sodium Cyanoborohydride. http://www.commonorganicchemistry.com/Common_Reagents/Sodium_Cyanoborohydride/Sodium%20Cyanoborohydride.htm.
28. Sodium cyanoborohydride. https://www.organic-chemistry.org/chemicals/reductions/sodiumcyanoborohydride.shtm.
29. Humphrey, R. E. & Hawkins, J. M. Reduction of Aromatic Disulfides with Triphenylphosphine. *Anal. Chem.* **36**, 1812–1814 (1964).
30. Mizhiritskii, M., Shpernat, Y. & Ashkenazi, B. Methods for the reduction of disulfide bonds. (2004).
31. Sweetman, B. J. The Specificity of Certain Phosphine Derivatives as Reducing Agents for the Disulfide Bond in Wool Keratin. *Textile Research Journal* **36**, 1096–1101 (1966).
32. Cline, D. J. *et al.* New Water-Soluble Phosphines as Reductants of Peptide and Protein Disulfide Bonds: Reactivity and Membrane Permeability. *Biochemistry* **43**, 15195–15203 (2004).
33. Niemuth, N. J., Thompson, A. F., Crowe, M. E., Lieven, C. J. & Levin, L. A. Intracellular disulfide reduction by phosphine-borane complexes: Mechanism of action for neuroprotection. *Neurochemistry International* **99**, 24–32 (2016).
34. Humphrey, R. E. & Potter, J. L. Reduction of Disulfides with Tributylphosphine. *Anal. Chem.* **37**, 164–165 (1965).
35. Ayers, J. T. & Anderson, S. R. A Preparative Scale Reduction of Alkyl Disulfides with Tributyl Phosphine and Water. *Synthetic Communications* **29**, 351–358 (1999).
36. Organometallic HyperTextBook: Phosphine Complexes. http://www.ilpi.com/organomet/phosphine.html.

37. Komáromy, D. *et al.* Self-Assembly Can Direct Dynamic Covalent Bond Formation toward Diversity or Specificity. *J. Am. Chem. Soc.* **139**, 6234–6241 (2017).
38. Kriebisch, C. M. E., Bergmann, A. M. & Boekhoven, J. Fuel-Driven Dynamic Combinatorial Libraries. *J. Am. Chem. Soc.* **143**, 7719–7725 (2021).
39. Synthesis of Donor-σ-Perylenebisimide-Acceptor Molecules Having PEG Swallowtails and Sulfur Anchors | The Journal of Organic Chemistry. https://pubs-acs-org.scd-rproxy.u-strasbg.fr/doi/full/10.1021/jo301701a.
40. Davis, B. G., Ward, S. J. & Rendle, P. M. Glycosyldisulfides: a new class of solution and solid phase glycosyl donors. *Chem. Commun.* 189–190 (2001) doi:10.1039/B008734N.
41. De Souza, M. C., Ohara, A. K., Araripe, D. R., Dias, P. D. C. V. & Kuwata, J. C. New Symmetric Diisopropylphosphoryl-S-Alkanedi Yldiisothioureas. Investigations on Their Synthesis. *Phosphorus, Sulfur, and Silicon and the Related Elements* **148**, 143–148 (1999).
42. Bermejo-Velasco, D., Nawale, G. N., Oommen, O. P., Hilborn, J. & Varghese, O. P. Thiazolidine chemistry revisited: a fast, efficient and stable click-type reaction at physiological pH. *Chem. Commun.* **54**, 12507–12510 (2018).
43. Chubarov, A. S. Homocysteine Thiolactone: Biology and Chemistry. *Encyclopedia* **1**, 445–459 (2021).
44. Hermanson, G. T. *Bioconjugate Techniques*. (Academic Press, 2013).
45. Liu, X. *et al.* Tris(2-carboxyethyl)phosphine promotes hydrolysis of iminoboronates. *Tetrahedron Letters* **58**, 3101–3106 (2017).
46. Baker, W. L. & Smiley, K. L. Possible explanation and implications of the reaction of ascorbic acid with some disulphide reagents. *Biochem J* **218**, 995–996 (1984).
47. Baker, W. L. Disulfide inhibition of copper-catalyzed oxidation of ascorbic acid: Spectrophotometric evidence for accumulation of a stable complex. *Archives of Biochemistry and Biophysics* **252**, 451–457 (1987).
48. Landino, L. M., Koumas, M. T., Mason, C. E. & Alston, J. A. Ascorbic acid reduction of microtubule protein disulfides and its relevance to protein S-nitrosylation assays. *Biochemical and Biophysical Research Communications* **340**, 347–352 (2006).
49. Giustarini, D., Dalle-Donne, I., Colombo, R., Milzani, A. & Rossi, R. Is ascorbate able to reduce disulfide bridges? A cautionary note. *Nitric Oxide* **19**, 252–258 (2008).
50. Krężel, A. *et al.* Coordination of heavy metals by dithiothreitol, a commonly used thiol group protectant. *Journal of Inorganic Biochemistry* **84**, 77–88 (2001).
51. Krężel, A., Latajka, R., Bujacz, G. D. & Bal, W. Coordination Properties of Tris(2-carboxyethyl)phosphine, a Newly Introduced Thiol Reductant, and Its Oxide. *Inorg. Chem.* **42**, 1994–2003 (2003).
52. Adamczyk, J., Bal, W. & Krężel, A. Coordination Properties of Dithiobutylamine (DTBA), a Newly Introduced Protein Disulfide Reducing Agent. *Inorg. Chem.* **54**, 596–606 (2015).
53. Podlahova, J., Kratochvil, B., Podlaha, J. & HaSek, J. Trigonal Bipyramidal Penta-aquazinc(11): Crystal Structure of Penta-aquazinc(i1) Bis(3,3',3"-phosphinetriyltripropionato)dizincate(II,II) Heptahydrate t. *J. CHEM. SOC. DALTON TRANS.* 4 (1985).

Chapter 5. Oxidation-driven recovery of a Michael acceptor in chemical reaction networks

Abstract

Michael acceptors (MAs) have been abundantly applied in various fields, from polymer chemistry to bioconjugations. Increased attention has been focused on β-substituted methacrylates for their ability to perform trigger-to-release and trigger-to-reverse chemistry. These processes are, however, irreversible in terms of recovering the initial MA species. Here we close this gap and develop a reaction cycle that is based on the recovery of a Michael accepting moiety. We show that by employing oxidation chemistry, sulfide-functionalized MAs can be transformed into sulfoxide- and sulfone-functionalized MAs. The latter two more electron-deficient species can then be attacked by amines, recovering the original amine-MA in the process. Furthermore, we show that depending on the strength of the used oxidant, this reaction network can proceed via three distinct pathways. We envision a wide range of applications for this chemistry, ranging from reversible cargo release systems to transient materials.

5.1 Introduction

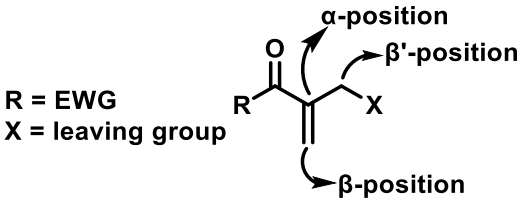
Unlike in nature, artificial chemical systems often strike compromises between efficiency and selectivity, making some properties such as pathway control and high recoveries in reaction networks hard to achieve or combine. In living systems however, one enzyme for example can be responsible for several independent and/or reversible processes. Glycolysis, a process of glucose conversion into pyruvate, is a metabolic pathway that consists of ten enzymatic reactions.[1] Some of these reactions are reversible and are used in gluconeogenesis in which a set of eleven enzymatic reactions generate glucose from lactate, glycerol, and glucogenic amino acids.[2] While the overall balance of forward glycolysis and backward gluconeogenesis is controlled by several regulatory enzymes, the reversible steps are governed by chemical equilibria with no or very little change in free energy compared to the irreversible steps.[2]

Redox reactions with $NADH/NAD^+$ are part of this metabolic pathway and assist, for example, in transport through the mitochondria membrane. Apart from these highly complex processes in nature, much simpler reversible reactions regulated by external conditions such as transaminations or transthioesterifications, to name a few, are also used in dynamic covalent chemistry.[3] These and similar types of reactions are the starting point in the search for reactions able to achieve transient activation of a substrate when combined into a chemical reaction network.[4]

One type of reaction that stands out are Michael/retro-Michael reactions. It is hard to overstate the importance of Michael addition and substitution reactions in a vast range of disciplines over the last decade. They are used in dynamic combinatorial libraries and predominantly comprise base-catalyzed thiol additions.[3] These reactions are especially interesting for developing complex bioorthogonal networks mimicking those in nature, as Michael additions can simultaneously involve the two most common nucleophiles in living systems—amines and thiols.[5] The addition of nucleophiles to Michael acceptors (MAs) has been widely used to functionalize molecules in synthetic organic chemistry, material chemistry as well as in biochemistry to modify proteins or release drugs, to name a few. Highlighting their versatility, Michael additions are considered to be 'click' reactions[6] because they are fast, quantitative and proceed at ambient conditions, both in organic and aqueous media.

An especially interesting group of molecules that—unlike common Michael acceptors—can not only undergo addition but also subsequent substitution reactions are β'-substituted MAs (Scheme 5.1). Interestingly, β'-substituted bromo-MAs have already been introduced in the 60s by the group of Richard Lawton, who continued to work on this chemistry for several decades. First, they were used as alkylating agents.[7] It was furthermore discovered that these double Michael additions with a good leaving group were reversible, which led to the term 'Equilibrium-transfer alkylating crosslink reagents' (ETACs) being established. This quickly resulted in their widespread application for crosslinking biomolecules such as antibodies and peptides.[8,9] PEGylation of proteins also became possible via this approach.[10] By extending the conjugation in the MA, these β'-substituted Michael acceptors were also applied as 'molecular yardsticks' and to synthesize macrocycles with high yields.[11] By making latent bis-sulfonylated MAs, it is possible to introduce a pH-switchable tag that can be applied in the presence of other MAs such as maleimides.[12] Popular derivatives of these β'-substituted have furthermore been used by the groups of Anslyn and Leigh for further applications in different fields.[6,13] A bis-bromo maleimide derivative has also been developed successfully and allows retention of the

MA functionality after two nucleophile addition-elimination cascades.[14] Different classes of MAs also exist: those that have been sulfinylated or sulfonylated in α- or β-position instead of in β'-position (Scheme 5.1). These molecules have been widely studied and reviewed.[15] However, their use cases are more limited as α-functionalized MAs lack a leaving group. In contrast, β-functionalized MAs are highly reactive, more prone to side reactions, and become sterically hindered and deactivated upon undergoing the first MA-addition-elimination cascade.



Scheme 5.1 | Structure of a general Michael acceptor with different substitution patterns in α-, β- and β'-position. The Michael acceptors used in this work possesses an electron-withdrawing group (EWG), as well as a leaving group (X) in β'-position.

Considering the successful applications of β'-substituted MAs in the past and their desirable ability to undergo highly efficient addition-substitution reactions, we chose to use β'-bromo and acetoxy-substituted MAs (Br-MAs, AcO-MAs) for this project. We focus specifically on amine and thiol additions to these conjugate acceptors in an aqueous medium, known as the 'trigger-to-release' process. If the addition product can subsequently react with a stronger nucleophile (thiol), recovering the weaker one (amine), the reaction is called the 'trigger-to-reverse' process. These processes can be combined in a chemical 'switch' to have a nucleophile that is first consumed during the cascade of reactions and afterward recovered.[5,6,16,17] An especially attractive feature of these MAs is that all substituents around the MA functionality influence its reactivity towards nucleophiles, which can be used for a fine-tuning of the compound to each specific application (Figure 5.1).[5] Manipulating the structure of the MA allows directing the addition to first proceed with amines, then with thiols (compounds 1–3, 5–10, Figure 5.1), or vice versa (compound 4, Figure 5.1), when both nucleophiles are present together.

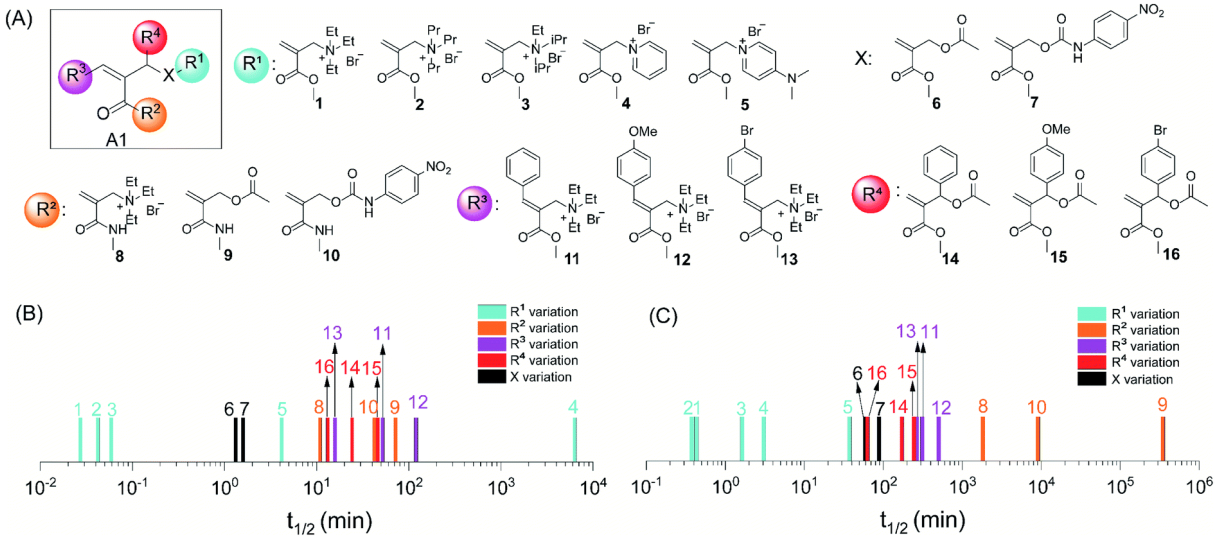


Figure 5.1 | Influence of MA structure (a) on its reactivity towards thiol (b) and amine (a). Half-lives of thiol and amine are presented correspondingly. Reprinted from ref. 5.

Despite all the advantages of MA derivatives for material functionalization, one of the pitfalls of this reaction cascade is that the thiol-functionalized MAs cannot be recovered due to the unfavorable leaving group properties of thiols. Therefore, to reuse the switch, an additional portion of the initial MA substrate is required. Hence, for the peptide-MAD system (chapters 2 – 4) to cycle again, one needs to add a new portion of MAD and the decyclization fuel. Or, in the example from Thayumanavan,[5] more bis-functionalized MA is needed to crosslink the 4-arm-PEG and obtain a gel again. To overcome these challenges, our idea was to develop a system where MA functional groups can be preserved and reused in subsequent cycles of material formation and degradation.

Here we show a CRN that enables the recovery of Michael acceptors via three distinct pathways, which can be selected by choice of oxidation reactants and starting conditions. Specifically, sulfide-MAs formed from amine-MAs are oxidized to convert them into more electron-deficient species, enabling the free amine to undergo a substitution reaction to recover the initial amine-MAs. The system operates in buffered aqueous solution under ambient conditions with β'-substituted Michael-accepting esters as these species are more electron-deficient and hence more reactive towards nucleophiles than amides.[5]

5.2 Experimental

The starting point of our CRNs is amine-MA **1** (Figure 5.2), which can be efficiently synthesized from amine and MA-ester with an excellent leaving group such as bromide or acetate in β'-position. These amine-MAs **1** can then be efficiently converted into sulfide-MA **2** (Figure 5.2), via the addition of thiol, releasing the amine. Traditionally, reverting to amine-MAs via re-addition of the amine is not possible, as the attack of a weaker nucleophile on sulfides is disfavored. Oxidation of sulfides to sulfoxides and sulfones increases the electrophilic character of the MA molecules, making them available for the reaction with weakly nucleophilic amines. The group of Bernardes showed that sulfone-MAs could functionalize proteins through the reaction with lysine residues.[18,19]

Choosing different oxidizers enables us to drive the CRN back to starting amine-MA **1** via three distinct pathways (Figure 5.2). With strong oxidizers (from NaClO to Oxone), oxidation of sulfide-MA **2** proceeds through sulfoxide-MA **3** to sulfone-MA **4** and subsequently enables the re-substitution reaction of amine to recover amine-MA **1** (Figure 5.2, left scheme). With weaker oxidants (KIO_4, $NaClO_2$), oxidation of sulfide-MA **2** can only go to sulfoxide-MA **3.** We found this species also to undergo substitution with amines—as shown for sulfone-MAs previously—to allow recovery of amine-MA **1** (Figure 5.2, central scheme). With even weaker oxidants (from KIO_3 to $KBrO_3$), sulfide oxidation does not happen. However, oxidation of the thiol released from an equilibrium between **2** and **1** shifts it toward the accumulation of amine-MA **1** (Figure 5.2, right scheme). While this reaction is typically not an equilibrium, using an especially electron-deficient thiol as a leaving group can make this conversion feasible under our conditions, albeit slow. The list of oxidants mentioned here is only exemplary, demonstrating the potential to influence the system by weaker and stronger oxidation reactants. Further, we discuss the oxidant-dependent evolution of the CRN in more detail.

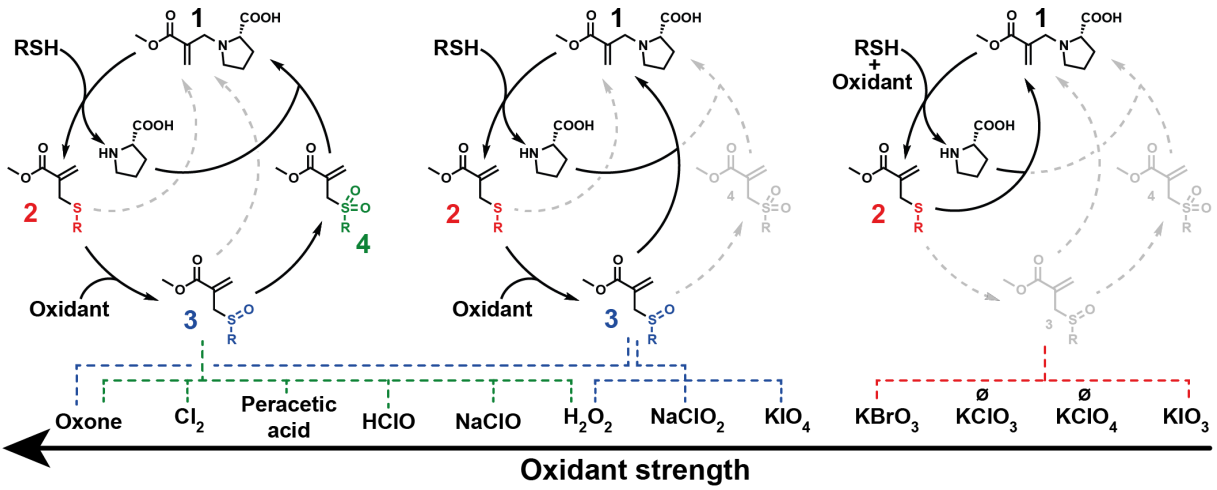


Figure 5.2 | Simplified CRN schemes; Depending on the oxidant strength, Proline-MA 1 recovery goes via different pathways. The left cycle (**Sulfone-MA cycle**) proceeds via Sulfone-MA **4** when stronger oxidants are used. The central cycle (**Sulfoxide-MA cycle**) proceeds via Sulfoxide-MA **3** when lower amounts of strong oxidants are used or equimolar amounts of milder oxidants. The right cycle (**Sulfide-MA cycle**) proceeds vis Sulfide-MA **2** only when weak oxidants are used.

Most of the reactions in this chapter involve Proline-MA **1** as amine-MA and 4-mercaptobenzoic acid (MBA) as a thiol. However, other amines, as well as thiols, were also investigated (see text further and Appendix). The syntheses of Proline-MA **1**, Sulfide-MA **2**, Sulfoxide-MA **3**, and Sulfone-MA **4** are described in the Appendix. For every network, we first study each step separately before combining them into one CRN and considering the whole system. In these experiments, a "stepwise-addition reaction" is defined as thiol being added to the reaction solution with Proline-MA **1** first, then, after a few minutes, oxidant is added; a "reaction cycle"—when thiol and oxidant are added simultaneously without a time delay.

5.2.1 Sulfone-MA cycle

We start from the most complex network (Figure 5.3a) where oxidation of Sulfide-MA **2** proceeds to the highest possible state—Sulfone-MA **4**. Sulfone contains sulfur in a higher +2 oxidation state compared to sulfoxide (0) and to reach it from the initial sulfide (sulfur is –2), one typically needs a strong oxidizer. We analyzed which oxidizers are capable of transforming sulfide-MAs (**2** in Figure 5.2) into sulfoxide- and sulfone-MAs (**3** and **4**) under the chosen conditions (pH 8). Commonly used oxidizers in literature to perform this task are Oxone (a complex potassium peroxomonosulfate-sulfate salt $2KHSO_5 \cdot KHSO_4 \cdot K_2SO_4$)[20] especially for acrylic substrates,[21,22] mCPBA,[23] hydrogen peroxide,[24–26] hypochlorites,[27] magnesium monoperoxophtalate (MMPP)[28] and potassium bromate.[29,30] We furthermore decided to test sodium chlorite and several other oxidants (potassium iodate, periodate, chlorate, and perchlorate). Oxone is one of the most popular literature choices for obtaining sulfones. It provides high yields, benign side products, and fast reaction times, which led us to move on with this oxidizer in the study of the sulfone CRN in Figure 5.3a. In phosphate buffer pH 8, oxidation of Sulfide-MA **2** with Oxone (2.5 eq.) proceeds stepwise via Sulfoxide-MA **3** (fast) to Sulfone-MA **4** within 1.5 h (Figure 5.3b).

We also studied the pH dependence of the oxidation at pH 7–8 but did not find significant differences. The pH, however, influences the addition of MBA to Proline-MA **1** and the

addition of proline to Sulfone-MA **4**, as for both reactions, there should be enough thiolate and unprotonated amine to perform an attack on the MA. In the abovementioned study of protein modifications,[18] the reaction between lysine residues and sulfone-MAs is selective due to the differences in pK_a values of amino groups depending on their surrounding within the protein structure. This suggests a strong pH dependence of the substitution of amines on sulfone-MAs, which was also discussed earlier by Anslyn.[6] We tested the addition of proline to Sulfone-MA **4** at pH 7–8 (Figure 5.3c, Appendix Figure 5.12), as this pH range was shown to be the most efficient. At pH 8, the reaction proceeds the fastest, and the equilibrium yield of Proline-MA **1** is the highest. Therefore, we decided to choose this pH for all following experiments. We avoided higher pH values as strongly basic conditions cause ester hydrolysis of Proline-MA **1**. The side product of the addition of proline to sulfone is a sulfinic acid or sulfinate (see **4→1**, Figure 5.3a), which we have shown to be in equilibrium with Sulfone-MA **4** (Figure 5.12).

Next, we set out to determine which nucleophiles can perform addition to sulfone-MA **4**. It has been shown that lysine,[18] as well as aliphatic alcohols,[31] can displace sulfinates from β'-substituted sulfone-MAs. We studied the reaction between Sulfone-MA **4** and 12 nucleophiles (Figure 5.3d and Appendix Figure 5.6–Figure 5.10): L-proline, morpholine, piperidine, (2-aminoethyl)trimethylammonium chloride hydrochloride (ATMA), 3-aminobenzoic acid, *N*-Methyl-aniline, *N*-Methyl-taurine, L-aspartic acid, *N*-Acetyl-L-lysine, phenol, pyrrole-2-carboxylic acid, and triethylamine (TEA). The reaction with morpholine was the fastest and gave the highest yield of Morpholine-MA (red circles in Figure 5.3d), but also resulted in fast decomposition, which caused approximately 40% loss over 17 h, likely due to hydrolysis and potentially double addition of sulfinate. Piperidine and ATMA showed a similar behavior: 50–59% maximum yield for Piperidine-MA and ATMA-MA respectively (blue triangle and green diamond in Figure 5.3d), with subsequent decomposition and loss of 29–30%. *N*-Methyl taurine gave a comparable result (Figure 5.9). In contrast to the latter three nucleophiles, proline showed almost no decomposition with an equilibrium yield of 57% (black squares in Figure 5.3d), which made it our primary candidate for further studies. The reaction of Sulfone-MA **4** with 3-aminobenzoic acid is the slowest compared to that with other nucleophiles (orange left-pointing triangle in Figure 5.3d), however, it is also the cleanest—no side products were observed, and the hydrolysis of the resulting MA product was minimal. We found that the reaction with aspartic acid and *N*-Acetyl-lysine (Figure 5.10) gives two main equilibrium products characteristic of conjugate additions of primary amines:[5] dialkylated and monoalkylated products. Hence, we chose not to proceed with primary amines. We were surprised to see that neither phenol reacted with Sulfone-MA **4** nor pyrrole, likely because they are not nucleophilic enough. TEA also did not react because according to Bernardes and co-workers,[18] the reaction between sulfone-MAs and amines proceeds via a hydrogen bond-assisted transition state which is only possible for primary and secondary amines.

With the conditions and rates for each step in hand, we explored the ability of the system to function as a reaction cycle where Proline-MA **1**, MBA, and Oxone are mixed simultaneously. In this case, all sulfur-containing MA species are transient. Since the addition of MBA to Proline-MA **1** is complete within a couple of minutes (Figure 5.3f), and so is the oxidation of thiol, we tried to find conditions where these competing reactions allow for some Sulfide-MA **2** formation which can then proceed through the cycle. The resulting recovered yield of Proline-MA **1** (Figure 5.3e) is rather low, and any significant improvement is difficult to achieve due

to the numerous side reactions that occur mainly due to Oxone's strong oxidizing power. For a more detailed discussion of the side reactions we could detect, please consult the 'Ultimate scheme' in the appendix (Scheme 5.2). The most problematic reaction is the oxidation of Proline-MA **1** to its *N*-oxide variant (Ox-Pro-MA in Figure 5.3e). While proline can recover Proline-MA **1** from Ox-Pro-MA (Figure 5.19), this pathway is hampered due to proline also being consumed by Oxone over time. We were surprised to see a slow accumulation of Sulfide-MA **2** throughout the reaction. This can be explained by the formation of thiol at the later stages of the CRN from the disproportionation of sulfenic acid[32,33] that forms from proline and Sulfoxide-MA **3** (see the product of **3** + proline in Figure 5.3a). Along with Sulfide-MA, other final products are Sulfone-MA **4,** Proline-MA **1,** and the analog of **1** with a hydrolyzed ester.

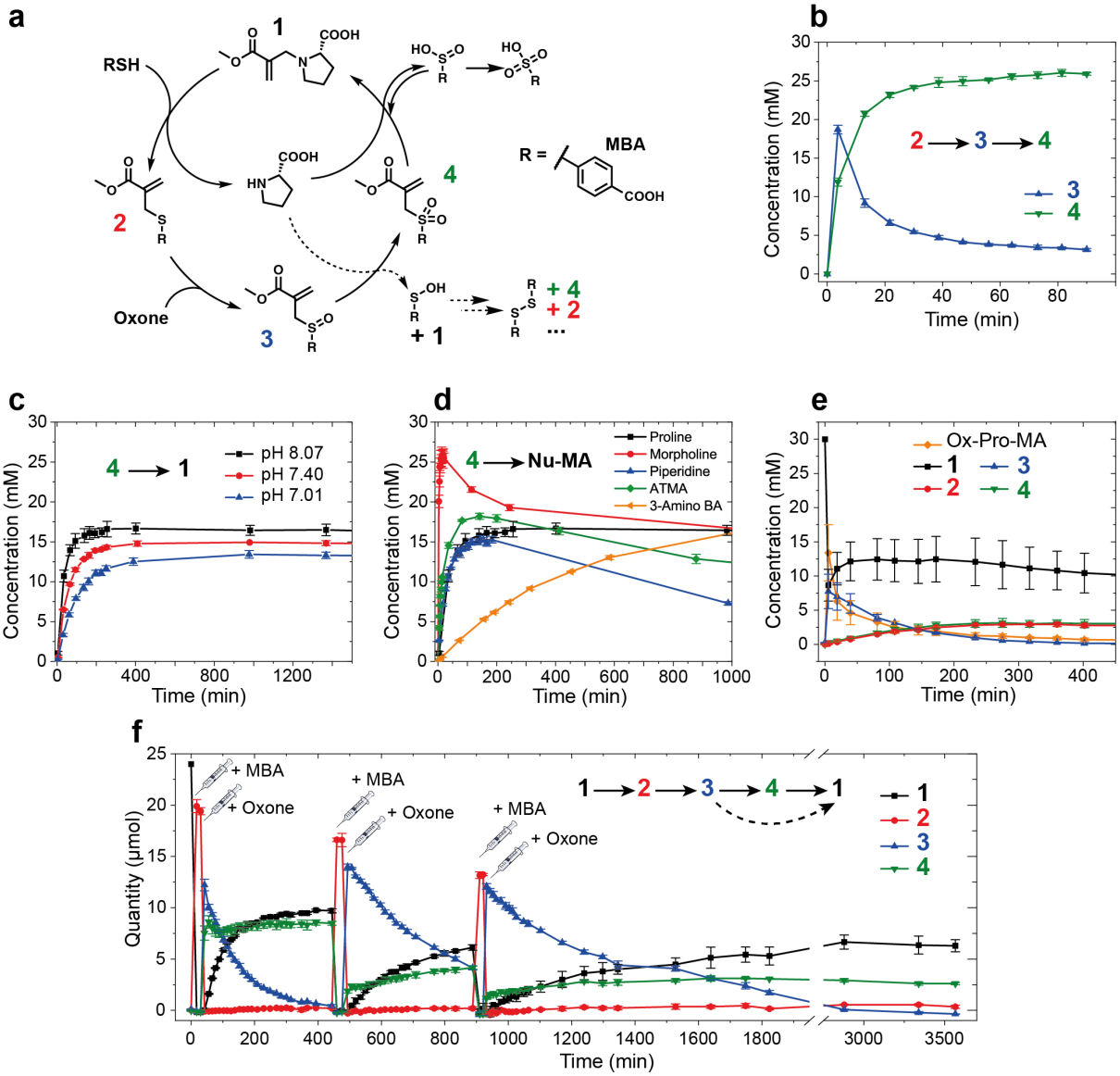


Figure 5.3 | Sulfone reaction network. a) The general scheme of the sulfone reaction network; b) Oxidation of Sulfide-MA **2** by Oxone at pH 8; c) Proline-MA **1** recovery from the reaction of Sulfone-MA **4** with proline at different pH; d) Nucleophile-MA (Nu-MA) recovery from the reaction of Sulfone-MA **4** with different nucleophiles at pH 8; e) Reaction cycle experiment, MBA and Oxone are added simultaneously to Proline-MA **1**; f) Stepwise addition of MBA and Oxone to Proline-MA **1**, the system is refueled two times after the first one. All kinetic measurements were performed in the same conditions: 10% DMF in 0.5 M phosphate buffer at 20 °C.

Next, we studied the system's evolution when MBA and Oxone were added in a stepwise manner a few minutes apart from each other (Figure 5.3f, a), giving time for Sulfide-MA **2** to form. The system was refueled twice after the Proline-MA **1** concentration started to plateau. Here, we achieve a 42% recovery (50% relative to sulfide-MA **2**) of Proline-MA **1** (black squares) in the first run, which dampens in the second run to 25% (37% relative to sulfide-MA **2**) with no further dampening observed after the third run. Contrary to our expectations, the **1→2** reaction is only about 83% in yield. The loss happens due to the formation of a double-addition product (see the Ultimate scheme Appendix, Scheme 5.2), which is to be expected if, in the earliest moments, there is a high local concentration of thiol.[5] In the stepwise addition network, there is almost no formation of Ox-Pro-MA, as at the time when Oxone is added, there is only Sulfide-MA **2** present, which consumes all the oxidant converting to Sulfoxide-MA **3** and Sulfone-MA **4** (Figure 5.3a). Looking closely at the evolution of the Sulfone-MA **4**, one can see that it maintains a steady state from the moment it is formed until all free proline is consumed to convert into Proline-MA **1,** and all oxidant is used up, reacting with Sulfoxide-MA **3**. As with many other systems (see chapter 1), we observe a waste accumulation over time (depending on pathway: disulfide, sulfonate, hydrolyzed or degraded MA species, sulfate), which likely slows down the kinetics of the CRN after each refuel. To summarize, we have shown that Oxone is an oxidant capable of closing the gap between thiol-Michael addition-substitution reactions and the recovery of the initial amine-MA by oxidizing the sulfide-MA **2** to yield more electron-deficient species that can undergo another addition-substitution reaction. Furthermore, this system can be refueled several times while still maintaining reasonable recovery yields.

Moving on from Oxone, in Figure 5.2, the sulfone network (left scheme) contains a few more oxidants that can also drive the oxidation of Sulfide-MA **2** to Sulfone-MA **4,** such as sodium hypochlorite (NaClO) or peracetic acid. Sodium hypochlorite can be protonated in water, forming HClO as well as disproportionate, releasing Cl_2. The latter two oxidants are stronger than NaClO, which is why we included them in the scheme—to point out that the oxidation process can be much more complicated than expected. Like with Oxone, the stepwise addition with NaClO proceeds relatively fast, giving however additional side-products. The reaction cycle with NaClO did not yield any Proline-MA **1** recovery due to extremely fast and quantitative Proline-MA **1** oxidation and other side reactions. We also explored several oxidants that can slowly release HClO in situ in a water-based solvent system, for example, sodium dichloroisocyanurate or 'pool disinfectant'.[34,35] The latter, however, before hydrolysis, converts MBA to disulfide via respective sulfenyl chloride and sulfenic acid, which outcompetes the reaction of MBA with amine-MA **1**.[36] Peracetic acid is the strongest organic oxidant of the peroxide type.[37] By indicating its place in the scheme, we refer to our experiments with similar oxidants, such as magnesium monoperoxophtalate (Figure 5.25) and its less soluble analog mCPBA, which turned out to give reasonable recoveries of Proline-MA **1** in stepwise addition runs.

Overall, for "cycle 1" we demonstrated a reaction network that enables the recovery of electrophilic amine-MA moieties through the oxidation of Sulfide-MA **2** to Sulfone-MA **4** with a range of different oxidants. The resulting mixture of products depends on the oxidant strength and how the CRN is initiated (via stepwise addition or as a reaction cycle via simultaneous addition of thiol and oxidant).

5.2.2 Sulfoxide-MA cycle

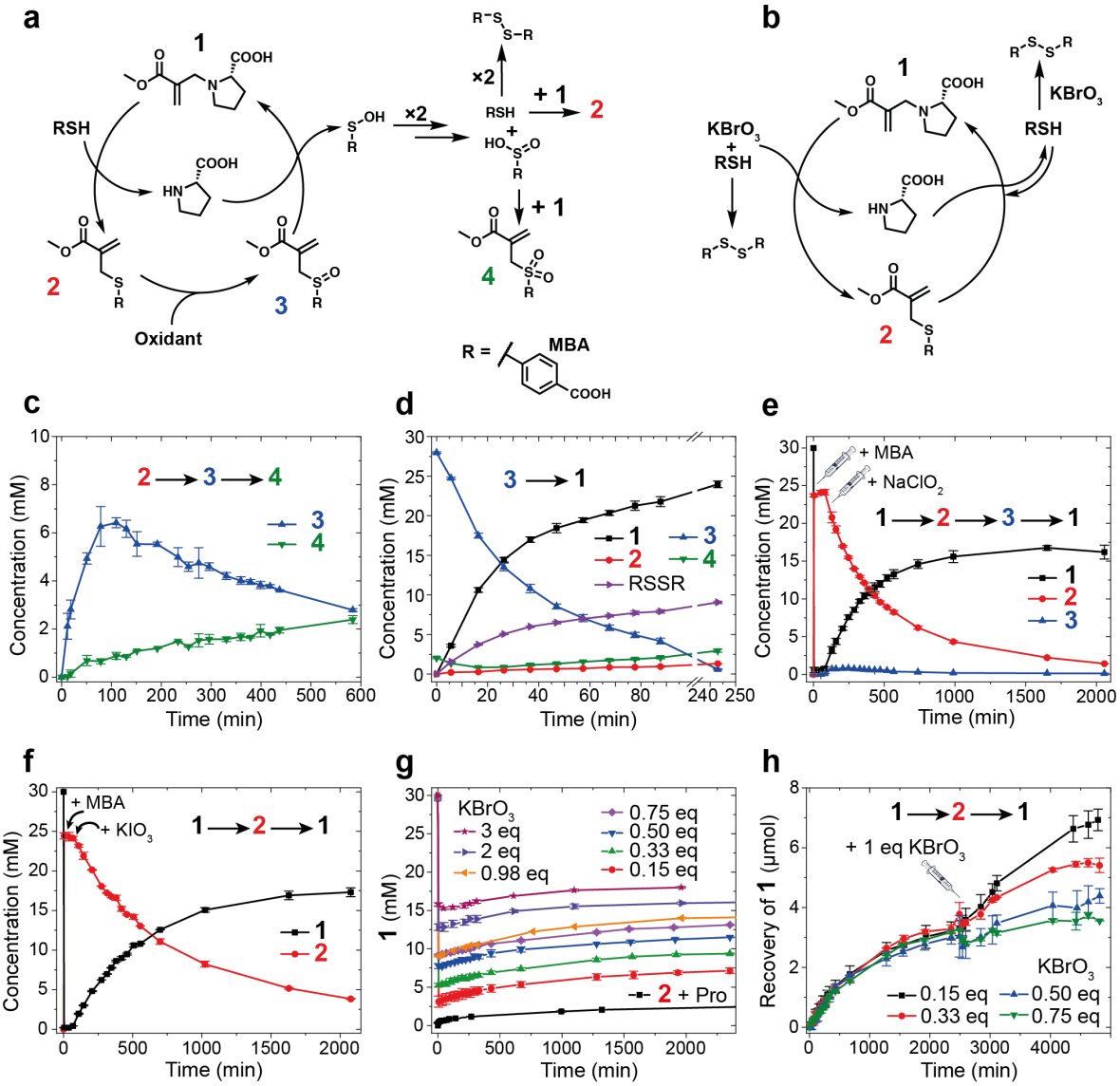


Figure 5.4 | Sulfoxide (a) and Sulfide (b) reaction networks. c) Oxidation of Sulfide-MA **2** with KIO_4; d) Sulfoxide-MA **3** + 1 eq. Proline; e) Stepwise addition of MBA and $NaClO_2$ to Proline-MA **1**; f) Stepwise addition of MBA and KIO_3 to Proline-MA **1**; g) Reaction cycle experiments with different amounts of $KBrO_3$; h) Differences in the recovery of Proline-MA **1** after addition of another 1 eq. of $KBrO_3$ to different reaction cycle experiments.

Oxidation of Sulfide-MA **2** with Oxone proceeds stepwise via Sulfoxide-MA **3** (Figure 5.3b). It is reasonable to question (i) whether there are oxidants that can drive oxidation only to sulfoxide and (ii) whether Sulfoxide-MA **3** is electrophilic enough for substitution by proline (Figure 5.4a). Literature reports the selective formation of sulfoxides upon treatment of sulfides with KIO_4,[38] Oxone (1 eq, see sulfoxide-MA **3** synthesis in Appendix), and H_2O_2 (in the presence of proline).[39] After testing several more oxidants, we found that $NaClO_2$, KIO_4, and reduced eq. of Oxone (see synthesis of sulfoxide-MA **3**) indeed yield mainly Sulfoxide-MA **3**, although further side reactions occur in aqueous solution (Figure 5.16, Figure 5.4c). Person and co-workers reported that oxidation of sulfides with H_2O_2 proceeds selectively to sulfoxides in the presence of proline.[39] Our results also align with this finding: oxidation of a pure Sulfide-MA **2** with H_2O_2 yields both Sulfoxide-MA **3** and Sulfone-MA **4**, whereas oxidation in the

presence of proline recovers Proline-MA **1** only via Sulfoxide-MA **3**. Separate studies of the reaction of Sulfoxide-MA **3** with proline (Figure 5.4d) not only surprised us with high Proline-MA **1** yield (84%) but also gave insights into the behavior of side products. The formation of significant amounts of MBA disulfide can be explained by looking closely at the sulfenic acid released through substitution with proline (Figure 5.4a). Sulfenic acids are highly reactive compounds which, in solution, quickly disproportionate to form thiols and sulfinic acids[32,33] along with other products in the presence of sulfides.[40] Sulfinic acid (or sulfinate) can then be further oxidized to sulfonate in the presence of excess oxidant or it can substitute proline in Proline-MA **1**, which would yield Sulfone-MA **4**. Similarly, thiol can either form disulfide or substitute proline in Proline-MA **1** and Sulfone-MA **4**, which would give rise to Sulfide-MA **2**. Even though the formation of disulfide prevails over other side products, we confirmed the presence of all these species via NMR and HRMS. Next, we studied stepwise addition experiments with each of these oxidants. With $NaClO_2$ for example (Figure 5.4e) we obtained a similar yield of Proline-MA **1** to that in the experiments with Oxone (Figure 5.3c, f). Notably, the transient yield of Sulfoxide-MA **3** was rather low, which can suggest that the conversion **3→1** is of comparable rate to that of **2→3**. To summarize, we have shown in this section, that by varying oxidant strength, a different cycle can be accessed. This "cycle 2" can, like "cycle 1", also achieve high recoveries of initial amine-MA through oxidation chemistry. However, the profile of side products differs strongly due to the release of highly reactive sulfenic acid as a waste species which can undergo a range of reaction cascades.

5.2.3 Sulfide-MA cycle

Next, we continued to study even weaker oxidants than those discuss previously. For several oxidants that we tested in the stepwise addition of MBA and oxidant ($KBrO_3$, $KClO_3$, $KClO_4$, and KIO_3, Figure 5.2), we found that they cannot oxidize Sulfide-MA **2** to either sulfoxide-MA **3** or sulfone-MA **4** (see SI, Figure 5.21, Figure 5.25). Stepwise addition experiments with these oxidants were compared to a 'blank' reaction of Sulfide-MA **2** with proline. The latter reaction was not expected to proceed to a noticeable extent since Sulfide-MA **2** is not electrophilic enough for proline to perform the substitution. To our surprise, even for these weak oxidants incapable of oxidizing sulfide-MA **2**, we observed recovery of Proline-MA **1** in the stepwise addition experiments over time (black squares in Figure 5.4g). Comparing this outcome to the stepwise addition of MBA and KIO_3 (black squares in Figure 5.4f), one can see a clear difference in formation rate and Proline-MA **1** yield. From this, we could conclude that there is another recovery pathway for Proline-MA **1** from Sulfide-MA **2** in the presence of an oxidant. Following the rate of disulfide formation in these experiments helped us to derive its likely mechanism (Figure 5.4b, right half of the scheme). Sulfide-MA **2** and proline can normally exist in equilibrium with free MBA and Proline-MA **1**. Commonly, amines—being worse nucleophiles—cannot re-attack sulfide-MAs.[5] However, MBA, being an aromatic, electron-poor thiol, makes it a sufficiently good leaving group and enables the sulfide- and proline-MAs to enter equilibrium with each other in the presence of free amine and thiol. However, once there is a weak oxidant in the system, the thiol is getting consumed to form disulfide, shifting in this way equilibrium toward Proline-MA **1**. The rate-limiting step, in this case, is the proline substitution of Sulfide-MA **2** since the rate of Proline-MA **1** recovery stays

the same if we change the amount of oxidant (Figure 5.4h). One oxidant that particularly stood out in these experiments was $KBrO_3$. Potassium bromate was also unable to oxidize sulfide-MA **2**. The reason for this finding is that its oxidation potential is much higher at low pH values.[41] A source of protons is required for the bromate to disproportionate into its more reactive forms. In stepwise addition experiments, this oxidant behaved similarly to KIO_3, however, once being added simultaneously with MBA to Proline-MA **1**, it did not consume all free thiol, like KIO_3, but allowed for the formation of Sulfide-MA **2** and was further able to scavenger MBA to yield disulfide (see the full scheme in Figure 5.4b and Figure 5.4g). We performed the same experiment with different amounts of $KBrO_3$ and found that the amounts of recovered Proline-MA **1** were the same (Figure 5.4h). The addition of another portion of bromate resulted in a further increase in Proline-MA **1** recovery. The final Proline-MA 1 yield was, in the end, higher for the experiments with a lower initial amount of bromate (black squares and red circles in Figure 5.4h) and lower for the experiments with a higher initial amount of bromate (triangles in Figure 5.4h). Hence, the more bromate is added in the beginning along with MBA, the more of the latter will be consumed, and the less Sulfide-MA **2** will form, releasing less proline. Less free proline results in lower or unchanged recoveries of Proline-MA **1** after an additional portion of the oxidant is supplied. This experiment also seems to suggest that bromate degrades in solution under our conditions over the reaction timescale. Therefore it may not be possible to reach quantitative recovery of Proline-MA **1**, as the lifetime of bromate appears to be shorter than the half-life of proline substitution of Sulfide-MA **2**, unless fresh shots of oxidant are added at later stages. Surprisingly, $KClO_3$ and $KClO_4$ being weaker oxidants than $KBrO_3$ but stronger than KIO_3 were not able to lead to any significant recovery of Proline-MA **1** (see SI Figure 5.25), potentially due to decreased oxidation potentials under our conditions or rapid degradation in solution.

5.3 Conclusion and outlook

Chemical reaction networks in nature usually are selective and specific for a certain function. Reaction networks that can proceed via different kinetic pathways are not common in artificial systems. Here we demonstrated a CRN that depending on the used oxidant, could evolve via three different routes, all yielding recovery of amine-functionalized Michael accepting species. Our objective is to utilize oxidation chemistry within CRNs that operate through amine/thiol 'trigger-to-release' Michael additions without loss of the MA moiety to allow for new possibilities in post-functionalization, catalyst release, and material applications. The steps in this CRN involve oxidation of sulfide MA to corresponding sulfoxide and sulfone, which corresponds to the increasing electrophilic nature of molecules, which enables the latter to react with an amine. This chemistry has not been used before to recover Michael acceptors.

The reaction system we have developed would allow access to switchable materials from a different perspective: the system's refueling is done by adding new portions of nucleophiles and not MAs, which can be more laborious to obtain or synthesize. To demonstrate this concept, we propose two follow-up projects that utilize the presented chemistry in a material scope.

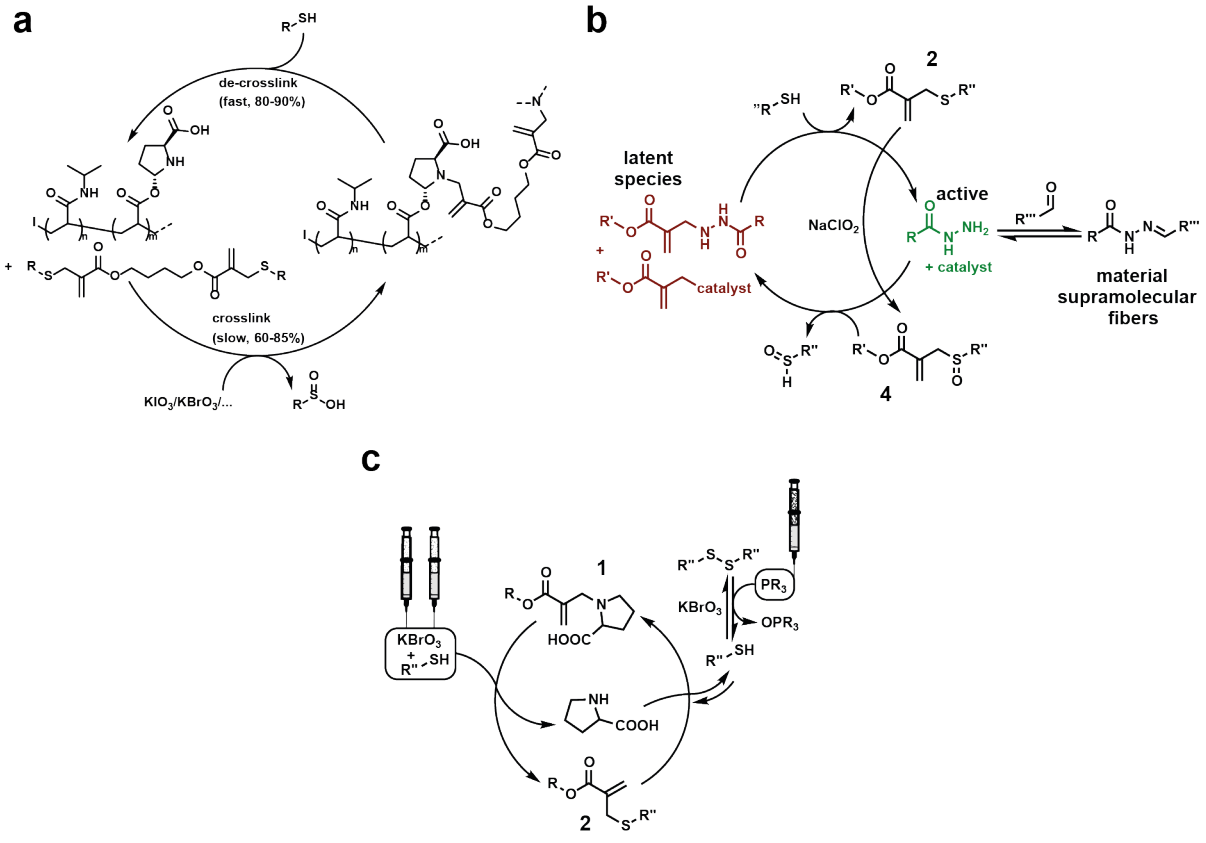


Figure 5.5 | Follow-up projects of the MA system. (a) Proline-functionalized polymers form a covalent polymer network in the presence of a bisfunctional sulfide-MA and a weak or medium-strength oxidant, to avoid side reactions of strong oxidants. Controlled degelation can be achieved by addition of a solution of monofunctional thiol. (b) Controlled hydrazone formation via triggered release of a catalyst and one of the building blocks, hydrazide. Hydrazone forms a gel, dissolved while oxidation of 2 to 3 also drives re-addition of hydrazide and catalyst to 3. (c) A flow setup to study the system's evolution under steady-state conditions. Syringes represent the syringe pumps where the flow of each reactant is controlled separately. Disulfide formed from 1-2-1 conversion can be recovered into thiol by the inflow of phosphine. This permits repeatedly running the network in the direction 2-1-2.

First, we propose the use of a proline-functionalized polymer, which is known from literature,[42] for sol-gel-sol transitions. By adding bis-functional Br-MAs, it should be possible to produce a covalently crosslinked polymer gel which can be disassembled in the presence of a monofunctional thiol and reformed with the addition of oxidants (Figure 5.5a). This would offer several advantages over previously explored systems,[5] namely the ability to repeatedly form and degrade the network, as well as minimize the waste species by reusing the relatively high-molecular weight MA crosslinkers.

Secondly, we propose that this new chemistry can be explored as a triggered catalyst release system to lead to gelation and degelation on demand. Supramolecular hydrazone gelations have been studied in depth previously.[43–45] However, they lack precise control over the formation kinetics and often present a one-way street, as degelation is difficult to achieve. By blocking both the catalyst as well as one of the building blocks (trishydrazide-gelator) (Figure 5.5b), the gelation can be postponed indefinitely and will only commence once a chemical signal, in this

case a thiol, is added. Upon addition of an oxidant, the sulfide-MA waste is oxidized and can now act as en electrophile again, blocking the catalyst as well as free trishydrazide-gelator, slowly pushing the equilibrium back towards a disassembled state. This would offer a high degree over the kinetics and stability of these hydrazine gels.

Finally, our current main objective of further studies is inspired by the reversible enzymatic reactions in the human body. The reaction cycle with MBA and potassium bromate (Figure 5.2, the right scheme) can be maintained in a steady state by a simultaneous influx of the thiol and the oxidant via syringe pumps (Figure 5.5c). A certain ratio of **1** and **2** is established depending on the starting conditions, flow rate, and thiol-to-oxidant ratio. According to the preliminary results, the main waste product of such an experiment is the disulfide of MBA, which is produced via direct oxidation of MBA by bromate and by removing thiol from Sulfide-MA-Proline. Finding a way to recover thiol from its disulfide would increase the amount of Sulfide-MA **2** in the system and, therefore, the amount of MBA that goes through the MA reaction cycle (Figure 5.5c). Moreover, when started from Sulfide-MA **2**, the reaction network with the waste recycling can be run in both directions repeatedly: **2** → **1** + disulfide → **2**. As was mentioned in chapter 4, phosphines are great agents for disulfide reduction. The setup for a forward-backward run of the reaction network will include a third syringe pump with the solution of water-soluble phosphine. Studying the system under the steady-state consumption of thiol, oxidant, and phosphine will give more information on how much thiol goes through the MA reaction cycle, how sensitive the system is to the perturbations of the **1**-**2** ratio, and the ratio of the injected reactants. The simplicity of the setup, availability of each reactant, complex dynamics of oxidation-reduction, and variety of possible thiol products when different oxidants are used make this system highly interesting and promising for a more detailed study under non-equilibrium conditions.

5.4 Author Contributions

Anastasiia Sharko and Benjamin Spitzbarth contributed equally to the idea development and all experiments including outlook trials. Stephen Eustace, a responsible person for the NMR facility of TU Delft, helped to establish a proper qNMR procedure. Thomas Hermans and Rienk Eelkema supervised the research.

5.5 Appendix

General Information

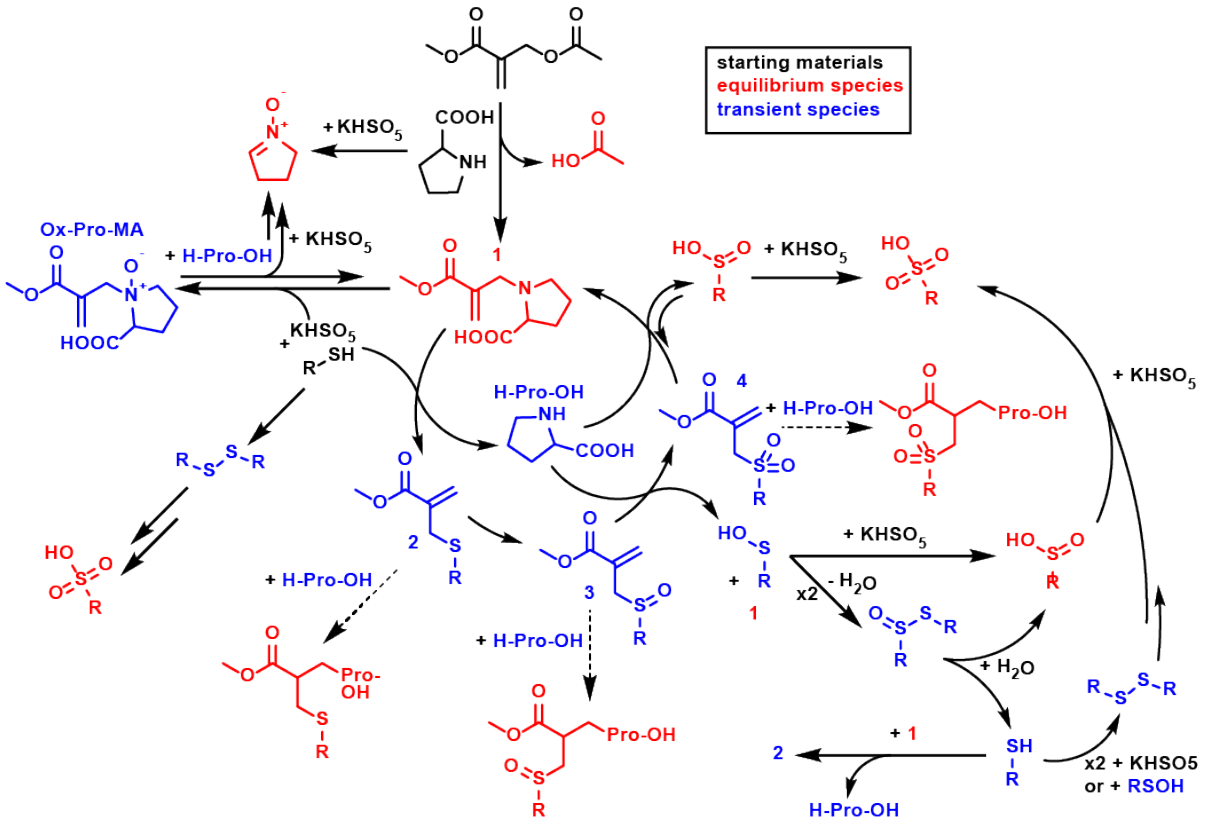
Methyl (2-hydroxymethyl)acrylate, L-Proline were purchased from Fluorochem. Oxone (potassium peroxymonosulfate) and sodium hypochlorite were purchased from TCI Europe. Acetonitrile, DMF, potassium hydrogen phosphate, potassium hydroxide, acetyl chloride, trimethylamine, anhydrous DCM, DSS, sodium chlorite, potassium iodate, potassium periodate, MMPP, SDCI, potassium chlorate and potassium perchlorate were purchased from Sigma Aldrich. DCM (technical grade) was purchased from VWR International. D_2O, MeOD, $CDCl_3$ and DMSO-d_6 were purchased from Eurisotop. Deionised (milliQ) water was made in

our laboratory. Unless stated otherwise, all chemicals were used as received. For water-free experiments, anhydrous solvents and pre-dried flasks were used.

NMR measurements

NMR spectra were recorded on an Agilent-400 MR DD2 (399.67 MHz) instrument. All measurements were taken at 298 K. For qNMR measurements, sealed ampoules (short NMR tubes with an outer diameter of 3 mm, purchased from VWR International) were prepared, containing a 10 mM solution of sodium trimethylsilylpropanesulfonate (DSS) in D_2O to ensure i) no potential reaction between standard and sample and ii) no presence of deuterated solvent in the sample to avoid a change in kinetics due to isotope effects. **All qNMR measurements were conducted in a 9/1 mixture of 0.5 M pH 8.0 potassium phosphate buffer and DMF unless stated otherwise.** For data analysis, the raw data was treated in Mestrenova V11.0. The 0.0 ppm signal of DSS was integrated and set to 1000 and the sample concentrations were derived accordingly, based on the known starting concentrations. To ensure full relaxation of all protons, T1 measurements (see Table S1) were performed and the scan time set to 5xT1 for the slowest relaxing signal in each experiment respectively. All kinetic studies were performed using PRESAT experiments with 8 scans and suppression of the H_2O peak.

Ultimate Scheme: sulfone CRN



Scheme 5.2 | Main reaction network of sulfone CRN, starting from proline-MA, followed by addition of thiol (MBA) and Oxone. All identified side reactions are shown. For clarity, starting materials are colored black, species that form transiently are colored blue and species that are present at the equilibrium after consumption of all reagents are colored red.

In the following, a chronological explanation of the reactions and side reactions we identified in the sulfone CRN (main text, Figure 1, left) follows. We chose to explain this CRN in detail, as it contains all identified reactions—the pathway and (side) reactions of the sulfoxide CRN (main text, Figure 1, center) and the sulfide CRN (main text, Figure 1, right) present a subset of the (side) reactions presented here.

The CRN starts by preformation of proline-MA **1** via addition and substitution of proline to AcO-MA, expelling acetate, which is inert under our conditions (see synthesis section for preparation of stock solution). (**Note:** an avoidance of Br-MA is crucial here due to the formation of hypobromite and bromine from the reaction of bromide with Oxone,[46] the latter of which we also confirmed visually.) Next, thiol (MBA) and Oxone are added, upon which several reactions can occur.

First, on the CRN's main route, thiol can react with proline-MA **1**, expelling proline as a transient species under formation of sulfide-MA **2**. Furthermore, all three species involved in this first step (proline-MA **1**, thiol and proline) can also react with Oxone in unwanted side reactions. This leads to:

i) Direct oxidation of proline, first to *N*-hydroxyproline,[47] and subsequently under decarboxylation to the corresponding nitrone as shown in literature.[48] We managed to identify the characteristic double bond signal of said nitrone (see Figure 5.17).

ii) Oxidation of proline-MA to oxidized proline-MA (Ox-Pro-MA in Scheme 5.2) which can degrade under double addition or re-form proline-MA **1** via substitution with proline under release of *N*-hydroxyproline (Figure 5.19).

iii) Oxidation of free thiol to form disulfide and subsequently higher oxidized species.

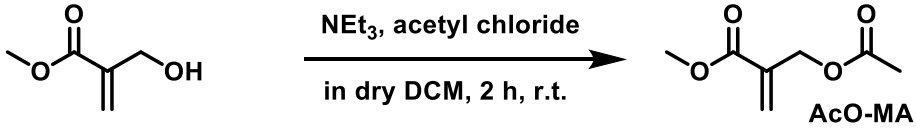
Continuing on the main route of the CRN, remaining Oxone can now oxidize sulfide-MA **2** to sulfoxide-MA **3** and sulfone-MA **4**. Oxone is a capable reagent for this task, as it can efficiently oxidize sulfides to sulfones, leaving the allylic double bond untouched; as previously shown in literature.[21] While Oxone can generally oxidize alkenes, this is only possible with electron-rich alkenes in the presence of ketones.[49] One of the side reactions that can occur at each stage of sulfide-MA **2**, sulfoxide-MA **3** and sulfone-MA **4** is the addition of proline under loss of double bond functionality. The double adduct species were identified by LCMS.

At the stage of sulfoxide-MA **3**, proline can now undergo an addition-substitution-reaction, expelling the corresponding sulfenic acid under generation of starting proline-MA **1**. This step is irreversible, as sulfenic acids are notorious for their high reactivity,[32] leading to their swift degradation. They can either be directly oxidized to sulfinates (and sulfonates subsequently), or dimerize under release of water to form thiosulfinates. These thiosulfinates can either comproportionate with sulfinates to form thiosulfonates, or, more likely under our conditions, hydrolyze (under sulfide-catalysis) to yield a thiol and a sulfinate, as shown in literature.[32,40] The thiol formed in this process can either be oxidized by free Oxone or react with proline-MA **1** to generate more sulfide-MA **2**. We managed to confirm these pathways by reacting sulfoxide-MA **3** with proline and identifying the proposed side products (Figure 5.4d).

At the last stage of the main route, sulfone-MA **4** is present and can directly react with proline to recover proline-MA **1** under generation of free sulfinate. This last step has been shown before in literature for the functionalization of lysine residues.[18] We propose that this reaction is an equilibrium (see Figure 5.12) and will only lead to high yields if the released sulfinate is further oxidized to inert sulfonate which presents the final waste species.

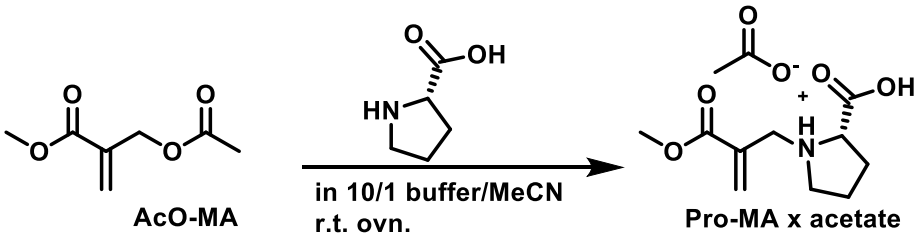
The cycle is closed and the recovered proline-MA **1** can now undergo the same reactions again upon addition of new portions of thiol and oxidant (Figure 5.3f).

Synthesis



Methyl (2-acetoxymethyl)acrylate (AcO-MA) was synthesized according to a literature procedure.[50] Briefly, 1.161 g (1.030 mL, 10.0 mmol, 1.0 eq.) methyl (2-hydroxymethyl)acrylate were dissolved in 20 mL anhydrous DCM under argon. The solution was cooled to 0 °C in an ice bath. 1.113 g (1.525 mL, 11.0 mmol, 1.1 eq.) triethylamine were added, followed by the dropwise addition of 0.785 g (0.714 mL, 10.0 mmol, 1.0 eq.) acetyl chloride. The ice bath was removed and a colorless precipitate formed within approximately one minute. After stirring at r.t. for two hours, the mixture was filtered and the filter was rinsed with additional DCM. The organic phase was washed with 20 mL water twice and with 20 mL brine once. The organic phase was dried over $MgSO_4$ and the solvent was removed under reduced pressure. The crude, yellow liquid was purified via a short silica plug (gradient 10/1 → 8/1 PE/EA) to yield 1.32 g (8.5 mmol, 85%) as a colorless oil.

^{1}H-NMR (400 MHz, $CDCl_3$): δ = 6.36 (s, 1H, C=C*H*$_2$), 5.84 (s, 1H, C=C*H*$_2$), 4.80 (s, 2H, C*H*$_2$OAc), 3.78 (s, 3H, OC*H*$_3$), 2.10 (s, 3H, O(O)CC*H*$_3$). The spectroscopic data was found to be in accordance with literature data.[50]



Note: This synthesis proved to be challenging. A synthesis starting from Br-MA was not feasible as the product would contain bromide as a counterion which can react with some oxidants, forming oxidized bromine species. The bromide counterion proved difficult to remove. Hence, a switch to AcO-MA was necessary to introduce the inert acetate counterion. Also, the use of buffer is crucial to avoid acidification from the release of acetic acid over time which would halt the reaction by protonating free proline. Furthermore, the pH MUST NOT exceed 8.0 as this will lead to swift hydrolysis of the methyl ester moiety. Hence, the use of other strong, sacrificial bases instead of buffer is also not feasible. The product was not isolated from the buffer salts and was used as a stock solution, the conctration of which was determined by addition of a known amount of pure 4-mercapto benzoic acid and comparing the signals of remaining Pro-MA and sulfide-MA.

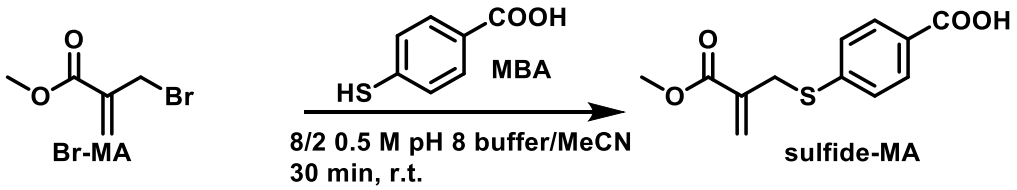
(2-(methoxycarbonyl)allyl)-L-proline (Pro-MA) was synthesized by dissolving 229.30 mg (2.0 mmol, 1.0 eq.) L-proline in a mixture of 10 mL 0.5 M pH 7.4 potassium phosphate buffer

and 1 mL acetonitrile. (**Note:** acetonitrile is not necessary, however it speeds up the reaction by aiding the dissolution of hydrophobic AcO-MA.) 647.6 mg (585.0 µL, 2.4 mmol, 1.2 eq.) AcO-MA were added dropwise. The mixture was vigorously stirred at room temperature overnight upon which the suspension cleared up. The excess AcO-MA was removed by extracting with 15 mL EtOAc twice. The acetonitrile was removed under reduced pressure and the water was removed by lyophilizing. The resulting colourless powder was dissolved in 9/1 0.5 M potassium phosphate buffer/DMF (pH dependent on experiment) to yield a 70 mM stock solution. The yield, determined by NMR, usually exceeded 90 % depending on the batch. **Note:** The following spectra were taken of a batch without acetate as counterion.

^{1}H-NMR (400 MHz, MeOD): δ = 6.56 (s, 1H, C=CH_2), 6.28 (s, 1H, C=CH_2), 4.20 – 4.08 (m, 2H, C_qCH_2N, **Note:** we found that the multiplicity of these CH_2 protons strongly depends on the solvent.), 3.93 (ddd, J = 9.6, 5.2, 1.7 Hz, 1H, NC*H*COOH), 3.83 (s, 3H, C(O)OCH_3), 3.72 (ddd, J = 11.8, 8.3, 5.1 Hz, 1H, NCH_2CH$_2$), 3.26 – 3.14 (m, 1H, NCH_2CH$_2$), 2.52 – 2.38 (m, 1H, NCH(COOH)CH_2), 2.22 – 2.04 (m, 2H, NCH(COOH)CH_2, NCH$_2$CH_2), 1.94 (dq, J = 13.2, 8.5 Hz, 1H, NCH$_2$CH_2).

^{13}C-NMR (101 MHz, MeOD): δ = 173.13 (*C*OOH), 167.05 (*C*OOMe), 135.58 (C_q=*C*H$_2$), 132.74 (*C*$_q$=CH$_2$), 70.64 (N*C*HCOOH), 56.61 (C_q*C*H$_2$N), 55.99 (N*C*H$_2$CH$_2$), 53.25 (C(O)O*C*H$_3$), 30.18 (NCH(COOH)*C*H$_2$), 24.64 (NCH$_2$*C*H$_2$).

ESI-LC/HRMS (m/z): calculated for [$C_{10}H_{16}NO_4$]$^+$: 214.1074, found: 214.1065; calculated for [$C_{10}H_{14}NO_4$]$^-$: 212.0928, found: 212.0923.

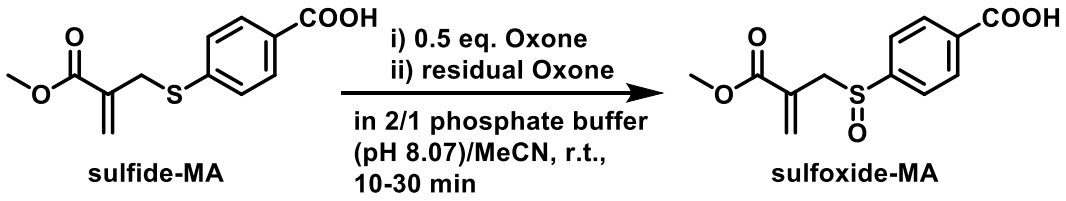


4-((2-(methoxycarbonyl)allyl)thio)benzoic acid (sulfide-MA) was synthesized by dissolving 154.18 mg (1.0 mmol, 1.0 eq.) 4-mercapto benzoic acid (MBA) in 10 mL of an 8/2 mixture of 0.5 M pH 8 potassium phosphate buffer and acetonitrile. 179.01 mg (119.34 µL, 1.0 mmol, 1.0 eq.) of Br-MA were added dropwise. After stirring at room temperature for 30 minutes, the acetonitrile was mostly removed under reduced pressure upon which the solution turned cloudy. (**Note:** Unlike sulfoxide- and sulfone-MA, sulfide-MA is not soluble in pure buffer even at elevated pH. Hence, it is important to add the reactants in a 1:1 ratio as precisely as possible as washing the aqueous phase will extract excess Br-MA as well as sulfide-MA even before acidification.) The aqueous phase was then diluted with approximately 10 mL water, acidified to pH 1 with approximately 5 mL 1M HCl and extracted with 20 mL EtOAc thrice. The combined organic phases were dried over $MgSO_4$ and the solvent was removed under reduced pressure to yield 227.1 mg (0.9 mmol, 90%) sulfide-MA as a pale-yellow powder.

^{1}H-NMR (400 MHz, MeOD): δ = 7.95 – 7.87 (m, 2H, C*H*C$_q$COOH), 7.41 – 7.35 (m, 2H, C*H*C$_q$S), 6.17 (s, 1H, C=CH_2), 5.76 (s, 1H, C=CH_2), 3.92 (s, 2H, CH_2S), 3.77 (s, 3H, C(O)OCH_3).

^{13}C-NMR (101 MHz, MeOD): δ = 169.39 (*C*OOH), 167.85 (*C*OOMe), 144.02 (*C*$_q$COOH), 137.52 (*C*$_q$S), 131.16 (*C*HC$_q$COOH), 129.33 (*C*$_q$=CH$_2$), 129.12 (*C*HC$_q$S), 127.74 (C$_q$=*C*H$_2$), 52.62 (*C*OOMe), 34.55 (*C*H$_2$S).

ESI-LC/HRMS (m/z): calculated for $[C_{12}H_{11}O_4S]^-$: 251.0384, found: 251.0382.

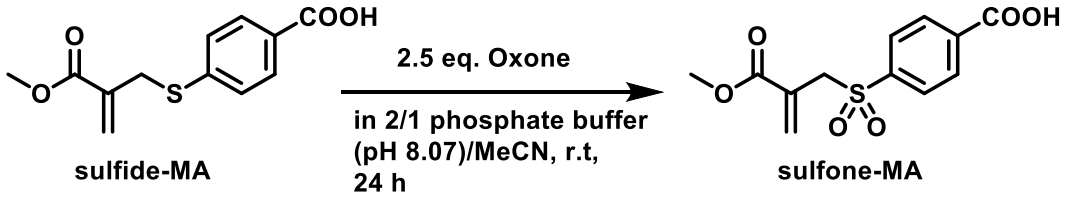


4-((2-(methoxycarbonyl)allyl)sulfinyl)benzoic acid (sulfoxide-MA) was synthesized by dissolving 47.58 mg (0.189 mmol, 1.0 eq.) sulfide-MA in 6 mL of a 2/1 mixture of 0.5 M pH 8.07 potassium phosphate buffer and acetonitrile. 28.97 mg (94.25 µmol, 0.5 eq.) Oxone was dissolved in 1 mL of water and added to the MA-sulfide solution. The solution was stirred at room temperature for 10 minutes. Upon NMR measurement, it was determined that 44.6 % of sulfide-MA had reacted to sulfoxide-MA. Hence, it was calculated that 35.2 mg Oxone were needed to achieve full conversion. This amount was weighed out, added as a solution in 1 mL water and it was stirred for a further 30 minutes. After completion, the reaction mixture was diluted with 15 mL water and acidified to pH 1 with approximately 5 mL 1 M HCl. Upon acidification, a thick, colorless precipitate formed. The milky suspension was extracted with 20 mL chloroform thrice. The combined organic layers were washed with 10 mL 1 M HCl and dried over $MgSO_4$. The solvent was removed under reduced pressure to yield 42.38 mg (0.158 mmol, 84%) of sulfoxide-MA as a colorless powder. **Note:** Unlike sulfide-MA and sulfone-MA, which we found to be stable over months at room temperature, sulfoxide-MA tends to degrade and needs to be stored in the fridge or preferably the freezer and be used up within several days to weeks.

^{1}H-NMR (400 MHz, MeOD): δ = 8.22 – 8.16 (m, 2H, C*H*C_qCOOH), 7.76 – 7.71 (m, 2H, C*H*C_qSO), 6.43 (d, J = 0.9 Hz, 1H, C=C*H*$_2$), 5.78 – 5.76 (m, 1H, C=C*H*$_2$), 4.02 (dd, J = 12.9, 1.0 Hz, 1H, C*H*$_2$SO, **Note:** we found that the multiplicity of these CH_2 protons strongly depends on the solvent.), 3.91 (dd, J = 12.9, 0.8 Hz, 1H, C*H*$_2$SO, **Note:** we found that the multiplicity of these CH_2 protons strongly depends on the solvent.), 3.62 (s, 3H, C(O)OC*H*$_3$).

^{13}C-NMR (101 MHz, MeOD): δ = 168.43 (*C*OOH), 167.17 (*C*OOMe), 148.22 (*C*$_q$COOH), 135.09 (*C*$_q$SO), 133.45 (C$_q$=*C*H$_2$), 131.36 (*C*HC$_q$COOH), 130.20 (*C*$_q$=CH$_2$), 125.79 (*C*HC$_q$SO), 59.08 (*C*H$_2$SO), 52.73 (C(O)O*C*H$_3$).

ESI-LC/HRMS (m/z): calculated for $[C_{12}H_{11}O_5S]^-$: 267.0333, found: 267.0333; calculated for $[C_{12}H_{13}O_5S]^+$: 269.0478, found: 269.0468.



4-((2-(methoxycarbonyl)allyl)sulfonyl)benzoic acid (sulfone-MA) was synthesized by dissolving 47.58 mg (0.189 mmol, 1.0 eq.) sulfide-MA in 6 mL of a 2/1 mixture of 0.5 M pH 8.07 potassium phosphate buffer and acetonitrile. 144.85 mg (0.471 mmol, 2.5 eq.) Oxone were dissolved in 1 mL of water and added to the MA-sulfide solution. A further 2 mL water were added to dissolve the Oxone fully. The solution was stirred at room temperature for

24 hours. After completion, the reaction mixture was diluted with 15 mL water and acidified to pH 1 with approximately 5 mL 1 M HCl. Upon acidification, a thick, colorless precipitate formed. The milky suspension was extracted with 20 mL chloroform thrice. The combined organic layers were washed with 10 mL 1 M HCl and dried over $MgSO_4$. The solvent was removed under reduced pressure to yield 40.68 mg (0.143 mmol, 76%) of sulfone-MA as a colorless powder.

^{1}H-NMR (400 MHz, MeOD): δ = 8.23 (d, J = 8.0, 2H, C*H*C_qCOOH), 7.96 (d, J = 8.0, 2H, C*H*$C_q$$SO_2$), 6.45 (s, 1H, C=C$H_2$), 5.85 (s, 1H, C=C$H_2$), 4.33 (s, 2H, C$H_2$$SO_2$), 3.60 (s, 3H, C(O)OC$H_3$).

^{13}C-NMR (101 MHz, MeOD): δ = 167.88 (*C*OOH), 166.80 (*C*OOMe), 143.40 (C_qCOOH), 137.19 ($C_q$$SO_2$), 134.41 ($C_q$=*C*$H_2$), 131.37 (*C*H$C_q$COOH), 130.59 ($C_q$=$CH_2$), 130.05 (*C*H$C_q$$SO_2$), 58.28 (*C*$H_2$$SO_2$), 52.83 (C(O)O*C*$H_3$).

ESI-LC/HRMS (m/z): calculated for $[C_{12}H_{11}O_6S]^-$: 283.0282, found: 283.0282; calculated for $[C_{12}H_{13}O_6S]^+$: 285.0427, found: 284.0417.

Supplementary experiments

T1 measurements

To allow for qNMR measurements to be conducted, we measured the T1 times for compounds to be quantified via inversion-recovery experiments. Measurements were performed under the same conditions (i.e. solvent, concentration and temperature) as in the kinetic studies. T1 times were extracted for the double bond signals of MA species and can be found tabulated below.

Table S1: Compounds that were used in qNMR, the signal of the nucleus that was used for quantification and its calculated T1 time.

Compound	Signal (ppm)	T1 time (s)
Sulfone-MA **4**	6.35	0.85
Sulfoxide-MA **3**	6.25	—
Pro-MA **1**	7.40	1.04
Sulfide-MA **2**	5.93	1.40
Morpholine-MA	6.22	1.09
Piperidine-MA	6.53	0.99
Lysine-MA	5.92	1.45
4-amino benzoic acid-MA	5.65	0.74
ATMA-MA	6.09	0.86
N-methyl taurine-MA	6.22	0.92
Aspartic acid-MA	6.27	0.82

Sulfone-MA + nucleophiles

Apart from the nucleophiles that can substitute on sulfone-MA discussed in the main text (Figure 5.3d), we tested several other nucleophiles. We found that phenol, tertiary amines such as trimethylamine and pyrroles such as pyrrole-2-carboxylic acid do not react with sulfone-MA. It has been in shown in literature that aliphatic alcohols can substitute sulfone MAs under certain conditions (in presence of K_2CO_3, see compound 5, condition iii. in reference).[31] Our rationale was that phenols, which are more acidic than aliphatic alcohols, should be able to act as nucleophiles under our conditions as well. However, no substitution took place (see Figure 5.6). Furthermore, in accordance with literature, we found that tertiary amines cannot substitute on sulfone-MA (Figure 5.7) due to the lack of the hydrogen-bonding motif.[18] Pyrrole-2-carboxylic acid can likely not participate in a substitution reaction due to the nitrogen's free electron pair being involved in aromaticity (Figure 5.8). We also tested the reaction of several other nucleophiles that successfully substituted on sulfone-MA. The conversion plots for these nucleophiles were left out in the main text for clarity reasons and the presence of some side reactions that will be discussed below.

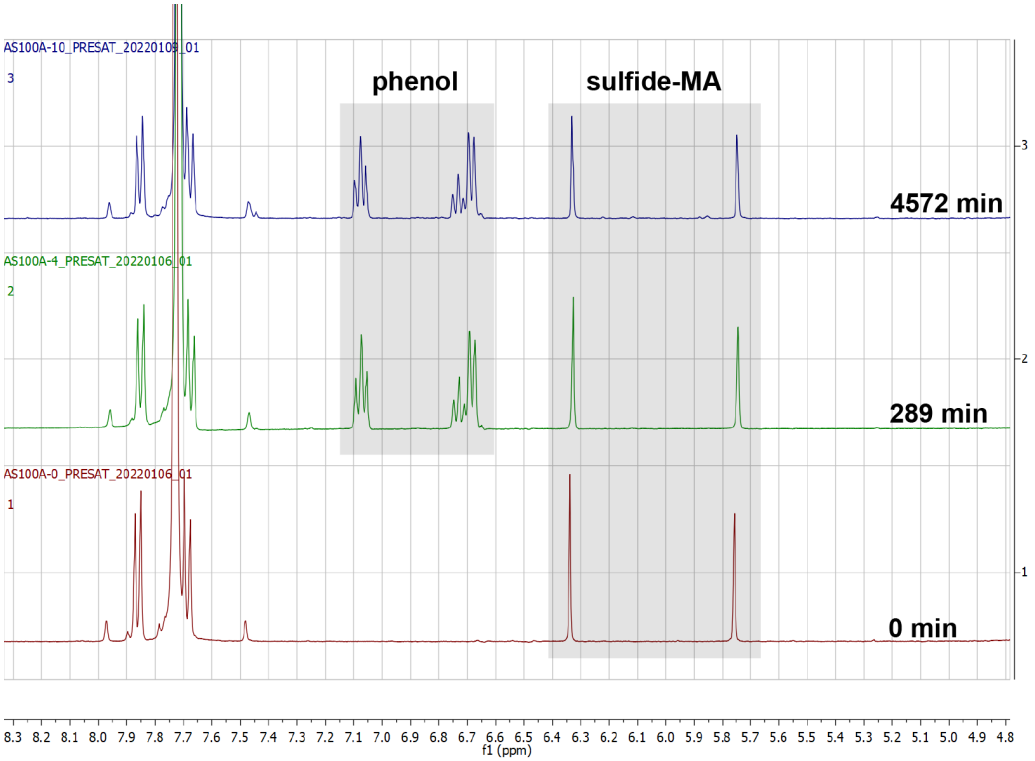


Figure 5.6 | ^{1}H-NMR cutout; bottom: sulfide-MA reference; middle and top: sulfide-MA after addition of 1.0 eq. phenol. No reaction could be observed. As stated in the general section, standard NMR conditions were used (i.e. 9/1 0.5 M pH 8.0 phosphate buffer/DMF), unless stated otherwise.

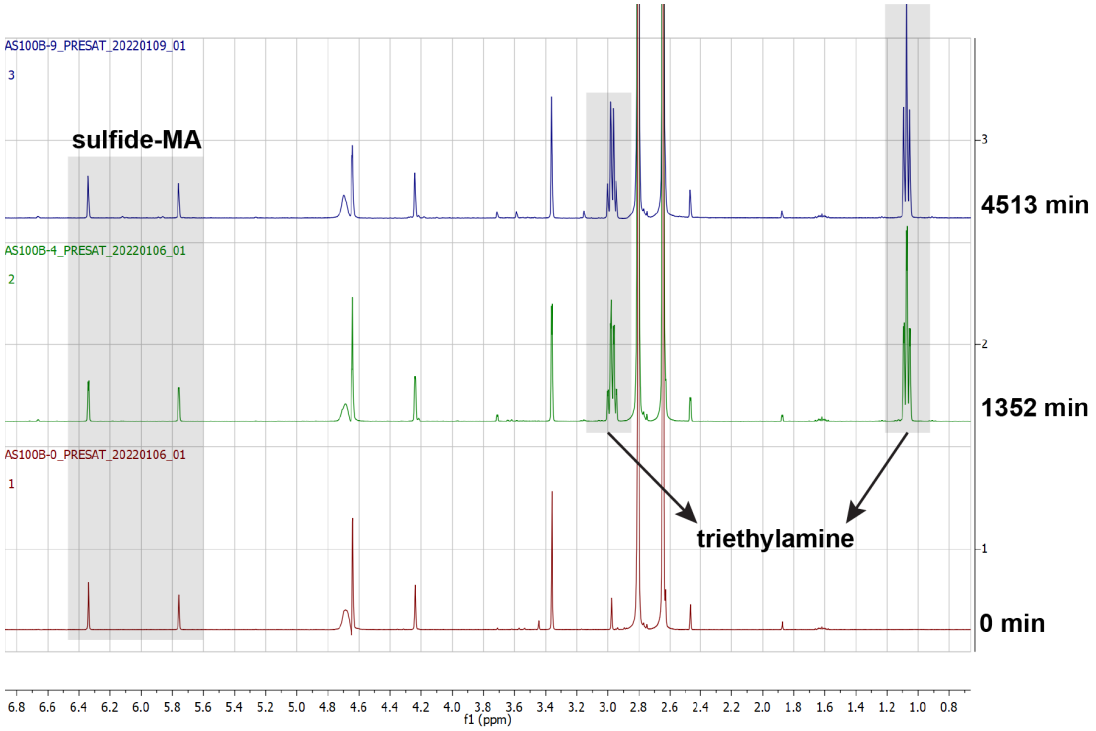


Figure 5.7 | ^{1}H-NMR cutout; bottom: sulfide-MA reference; middle and top: sulfide-MA after addition of 1.0 eq. triethylamine. No reaction could be observed.

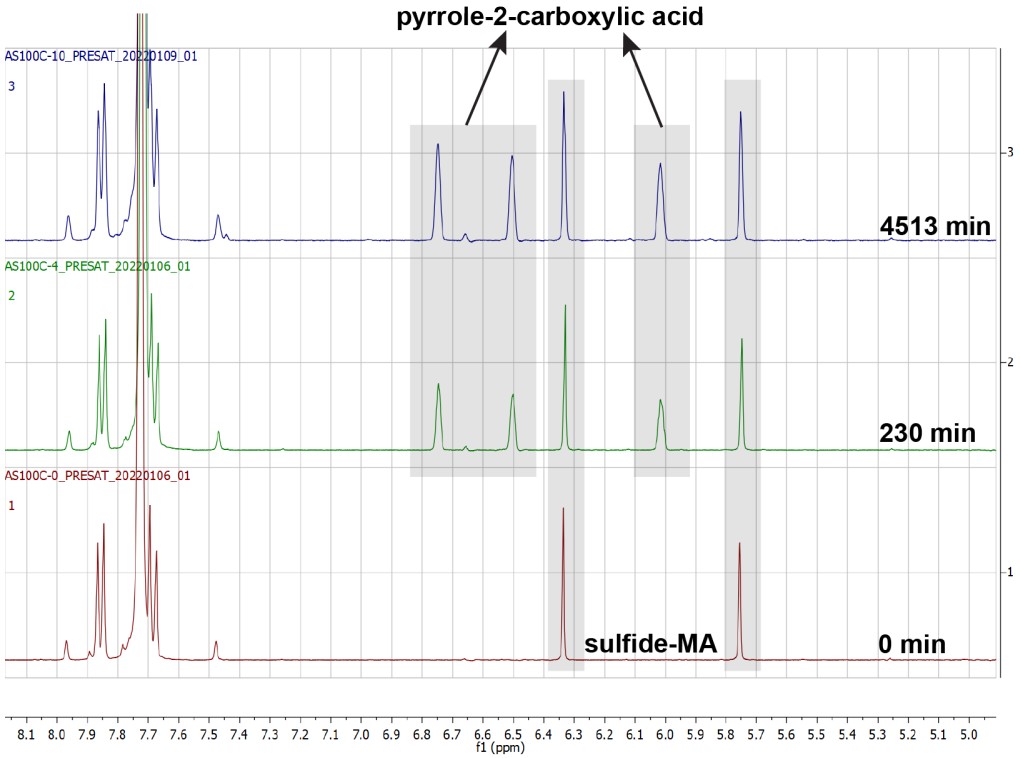


Figure 5.8 | ^{1}H-NMR cutout; bottom: sulfide-MA reference; middle and top: sulfide-MA after addition of 1.0 eq. pyrrole-2-carboxylic acid. No reaction could be observed.

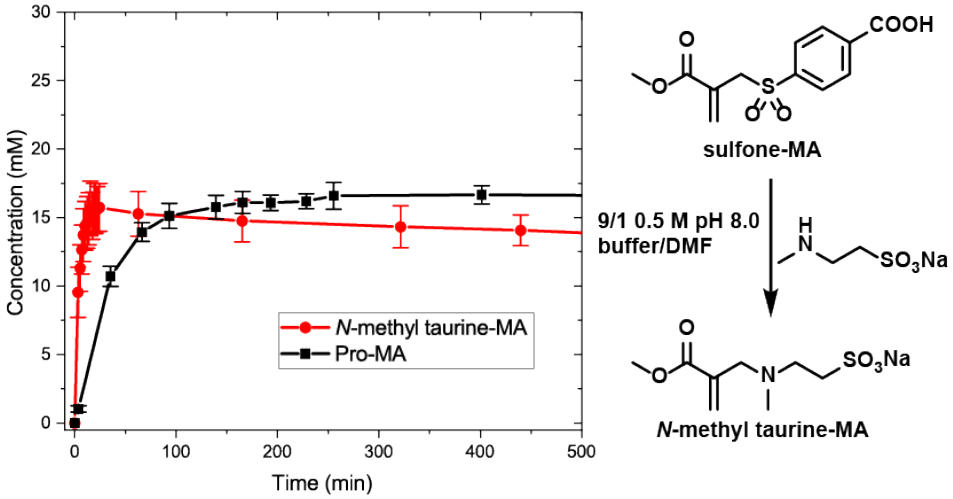


Figure 5.9 | Reaction of sulfone-MA with N-methyl taurine over time. The reaction with proline was plotted for comparison.

The reaction of sulfone-MA with the primary amines aspartic acid and *N*-acetyl lysine (Figure 5.10) showed the formation of double addition products which made these species not suitable. Interestingly, the third primary amine we tested, ATMA, did not show this double addition (main text figure).

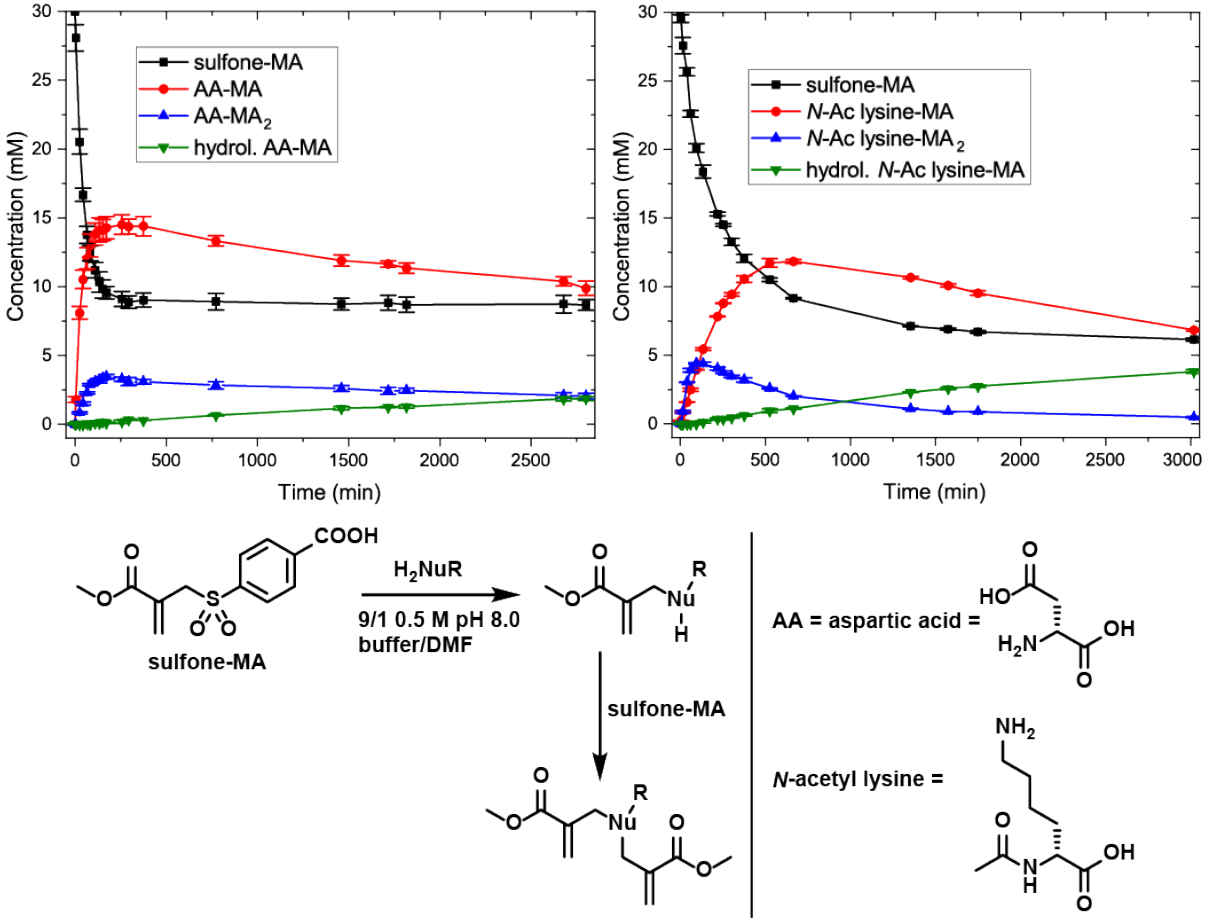


Figure 5.10 | Reaction of sulfone-MA with aspartic acid (left) and *N*-acetyl lysine (right). The graph shows the consumption of sulfone-MA over time, along with the formation of the mono- and bis-adducts over time. The green line shows the hydrolysis of the mono-adducts over time.

Apart from comparing the kinetics to choose which nucleophile would be best suited for our system, we also looked at any potential side reactions. The main side reaction we found (besides double addition of primary amines) was the hydrolysis of the substituted products (Figure 5.11). The comparatively slow hydrolysis of proline-MA was one of the points that made it a favorable candidate for our CRNs.

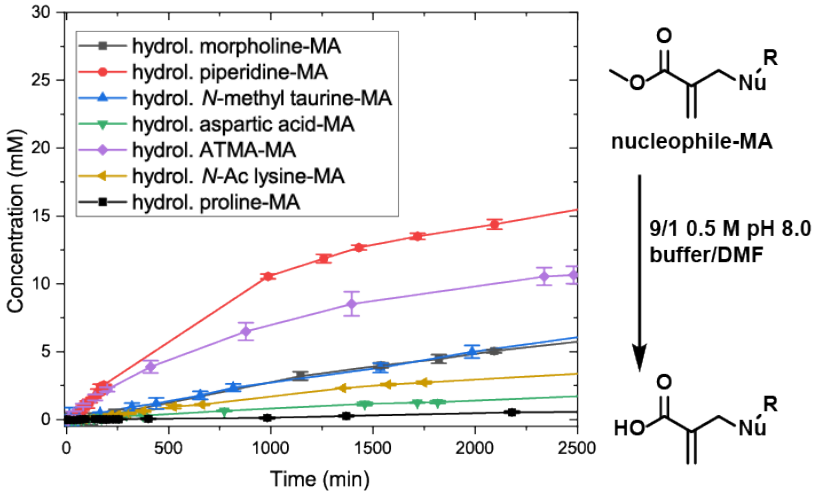


Figure 5.11 | Hydrolysis of nucleophile-MAs (products of sulfone-MA + nucleophile) over time. 3-amino benzoic acid is not shown as no hydrolysis product was detected over this time scale.

Sulfone-MA + additional proline: substitution reversibility and the fate of the sulfinate

As can be seen for the reaction of sulfone-MA + nucleophiles (Figure 5.3d, as well as Figure 5.9 and Figure 5.10), different nucleophiles give different yields of nucleophile-MA, yet some sulfone-MA always remains unreacted. We hypothesize that the substitution of sulfone-MA under release of the corresponding sulfinate salt is an equilibrium reaction. The sulfinate anion is nucleophilic enough to re-attack the nucleophile-MAs, displacing an amine. This would explain the lower yields with less nucleophilic amines (e.g. proline-MA vs. morpholine-MA). To further support this hypothesis, upon reaction of sulfone-MA with 1 eq. proline, a further 5 eq. of proline were added to examine how the system behaves. We indeed found an increase in yield of proline-MA, suggesting the presence of an equilibrium between sulfone-MA and nucleophile-MA. However, as demonstrated in Figure 5.12, an excess of nucleophile should be avoided, as the increase in yield is negligible compared to the loss of active MA moieties due to the double addition of nucleophiles under irreversible formation of double addition adducts.

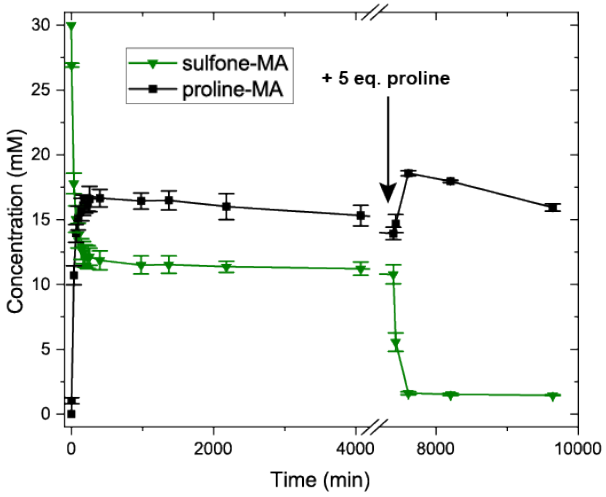
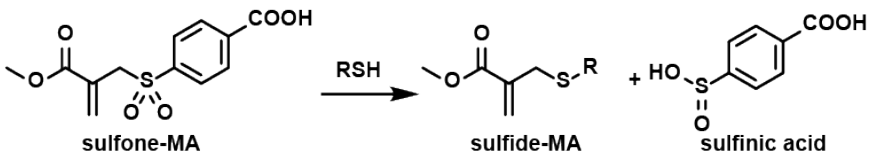


Figure 5.12 | Reaction of sulfone-MA with 1.0 eq. proline and addition of a further 5.0 eq proline.

Apart from these experiments, we managed to confirm the presence of the sulfinate of MBA via HRMS: found for $[M–H^+]^-$ 184.9907 (expected: 184.9914).
Furthermore, we found the sulfonate of MBA: found for $[M–H^+]^-$ 200.9858 (expected: 200.9863). It remains unclear, whether the sulfonate formed during analysis in the LC/HRMS, or whether oxidation to sulfonate occurred before analysis due to the presence of atmospheric oxygen. In general, sulfinates (and even more so sulfinic acids) are prone to oxidation; however, aromatic sulfinates have been found to be somewhat more stable than their aliphatic counterparts.[51] In general, we assume that the sulfinate of MBA is relatively stable in solution. Otherwise, in the reaction of sulfone-MA **4** with nucleophiles, more nucleophile-MA would form over time as the released sulfinate degrades to yield sulfonate which is not in equilibrium anymore.

Sulfone-MA + thiol(s)

As expected, we found that not only amines can displace sulfinates by undergoing a substitution reaction with sulfone-MA **2**, but also thiols—as they are the stronger nucleophiles—giving sulfide-MA and a sulfinate salt in the process (Scheme 5.3). This reaction is fast compared to the reaction with amines (i.e. full conversion within minutes). We chose an aromatic thiol (MBA, Figure 5.13) and an aliphatic thiol (3-mercapto propionic acid, Figure 5.14) to demonstrate this point.



Scheme 5.3: Reaction of sulfone-MA with a thiol (RSH) to yield sulfide-MA and sulfinic acid (effectively sulfinate at pH 8).

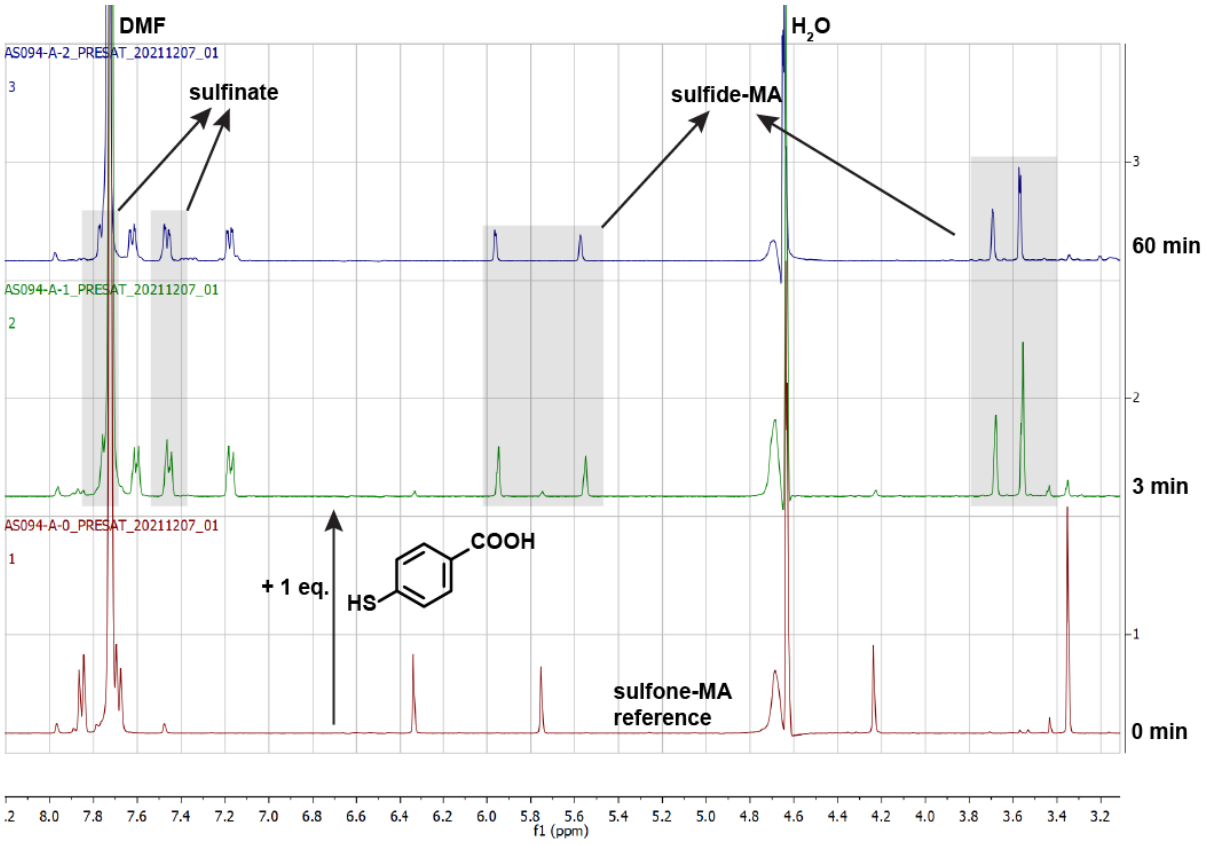


Figure 5.13 | Reaction of sulfone-MA 4 with 1.0 eq. of MBA to yield sulfide-MA 1 and sulfinate. Note: The downfield aromatic sulfinate signal overlaps with the DMF signal.

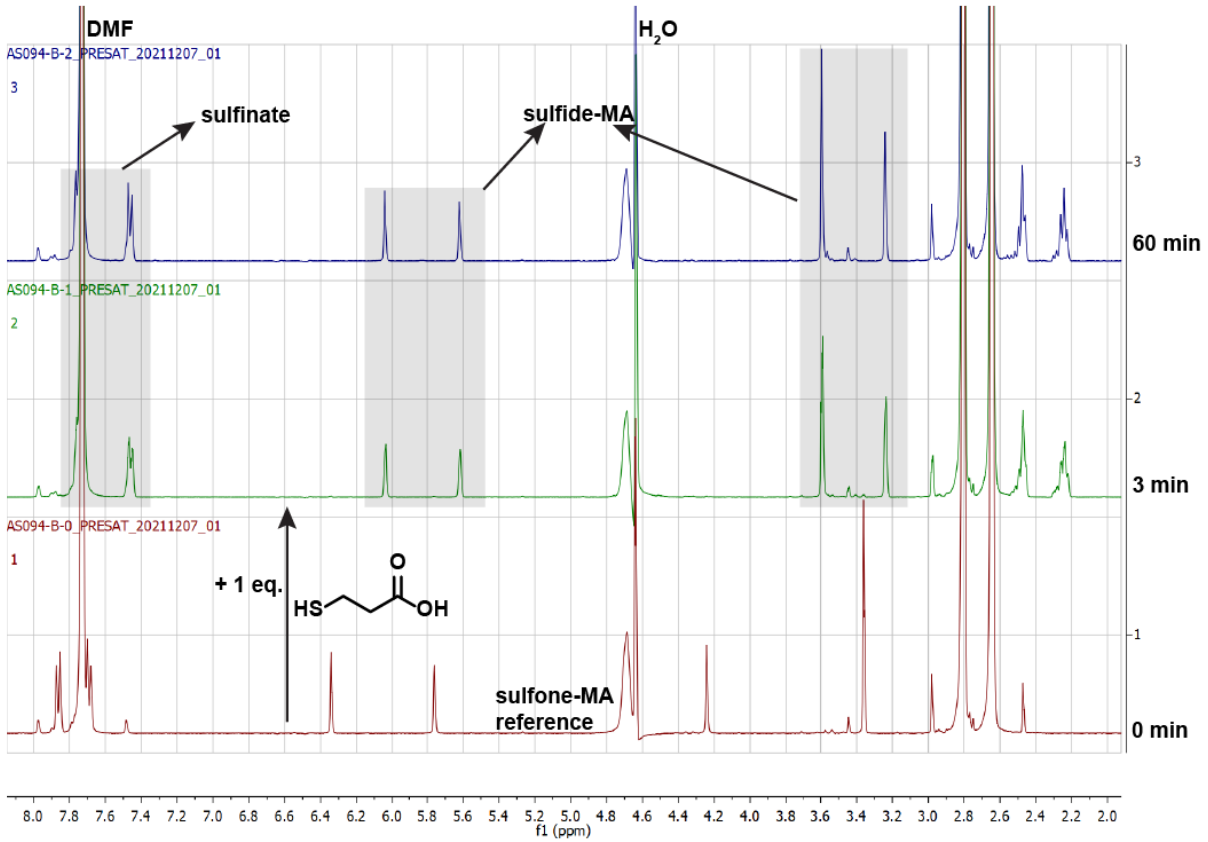


Figure 5.14 | Reaction of sulfone-MA 4 with 1.0 eq. of 3-mercapto propionic acid to yield sulfide-MA and sulfinate. Note: The downfield aromatic sulfinate signal overlaps with the DMF signal.

Sulfoxide-MA + thiol

Analogously to the reaction between sulfone-MA and thiols, the reaction between sulfoxide-MA and thiols can also proceed, yielding a sulfide-MA and a sulfenic acid which will further react according to Scheme 5.2 (see Figure 5.15). The same side products (i.e. sulfone-MA **4**, disulfide) can also be found from the reaction of sulfoxide-MA **3** with proline (Figure 5.4d).

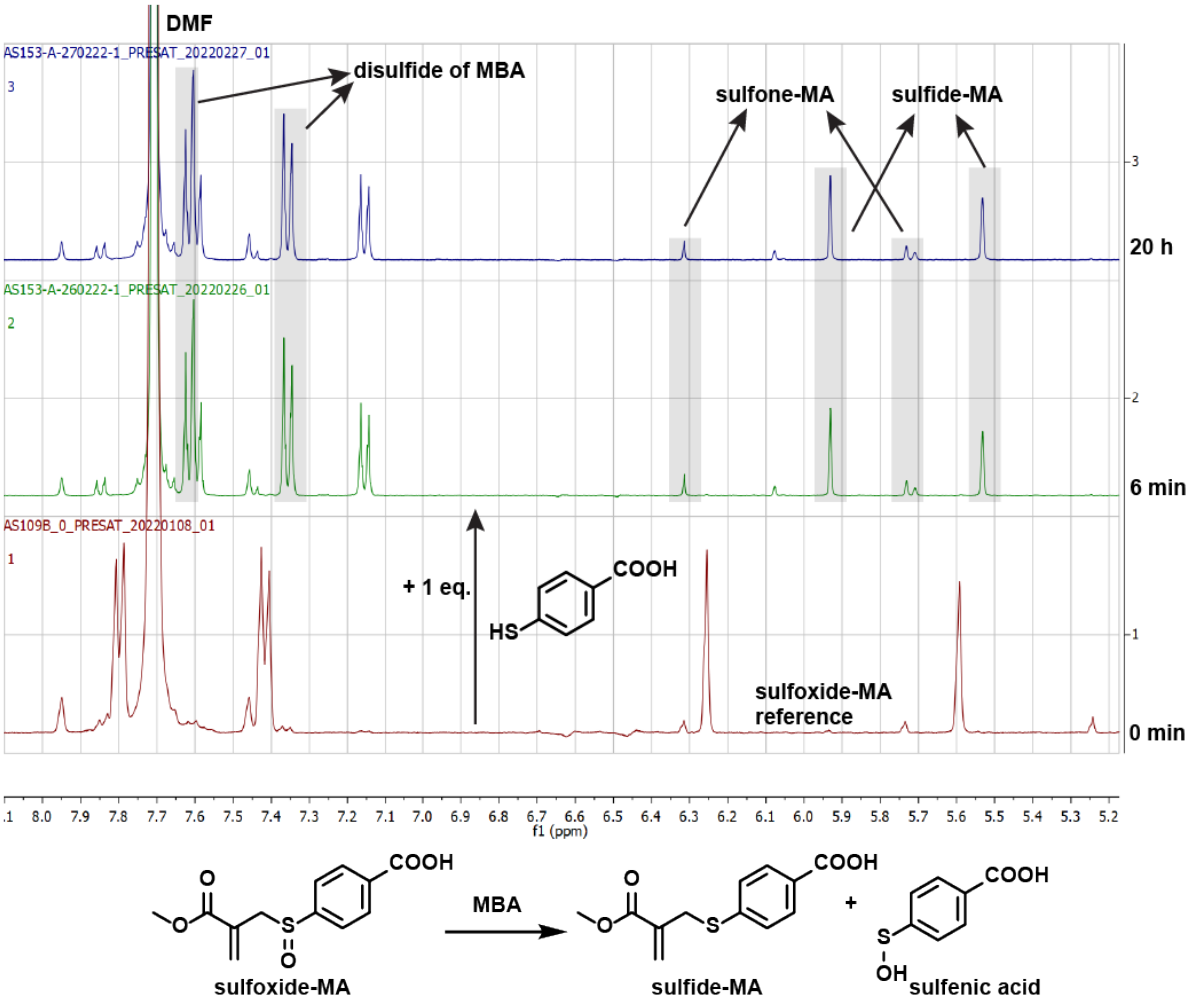


Figure 5.15 | Reaction of sulfoxide-MA 3 with 1.0 eq. MBA to yield sulfide-MA 1, along with the disulfide of MBA (from degradation of sulfenic acid) and small amounts of sulfone-MA 4 (from degradation of sulfenic acid and further reaction with sulfoxide-MA).

Sulfoxide-MA stability study

Unlike sulfide-MA **2** and sulfone-MA **4**, sulfoxide-MA **3** was neither stable in its pure form, nor in solution for prolonged periods of time. We found a comparatively swift degradation of sulfoxide-MA **3** in solution (9/1 buffer/DMF), likely via solvent-induced addition-substitution reaction, as we observed HO-MA as the main side product (attack of hydroxide and/or water on the highly polarized sulfoxide-MA species under release of sulfenic acid), along with the

common sulfenic acid degradation products mentioned previously (i.e. sulfide-MA **2**, sulfone-MA **4**, disulfide of MBA). The linear degradation of sulfoxide-MA **3** over time also suggests pseudo-first order kinetics which further supports attack of water as a plausible mechanism as water is present in large excess. Furthermore, this degradation (Figure 5.16) also explains why oxidants that should selectively proceed via the sulfoxide pathway (main text, Figure 1, center) also show the formation of minor amounts of sulfone-MA **4** over time—not due to over-oxidation but due to sulfoxide-MA degradation instead. This sulfone-MA **4** formation may however be somewhat suppressed in the presence of oxidants compared to the free degradation of sulfoxide-MA **3**, as oxidants can react with the sulfenic acid, reducing the formation of its degradation products that would otherwise form in a non-oxidizing environment such as in this experiment (Figure 5.16).

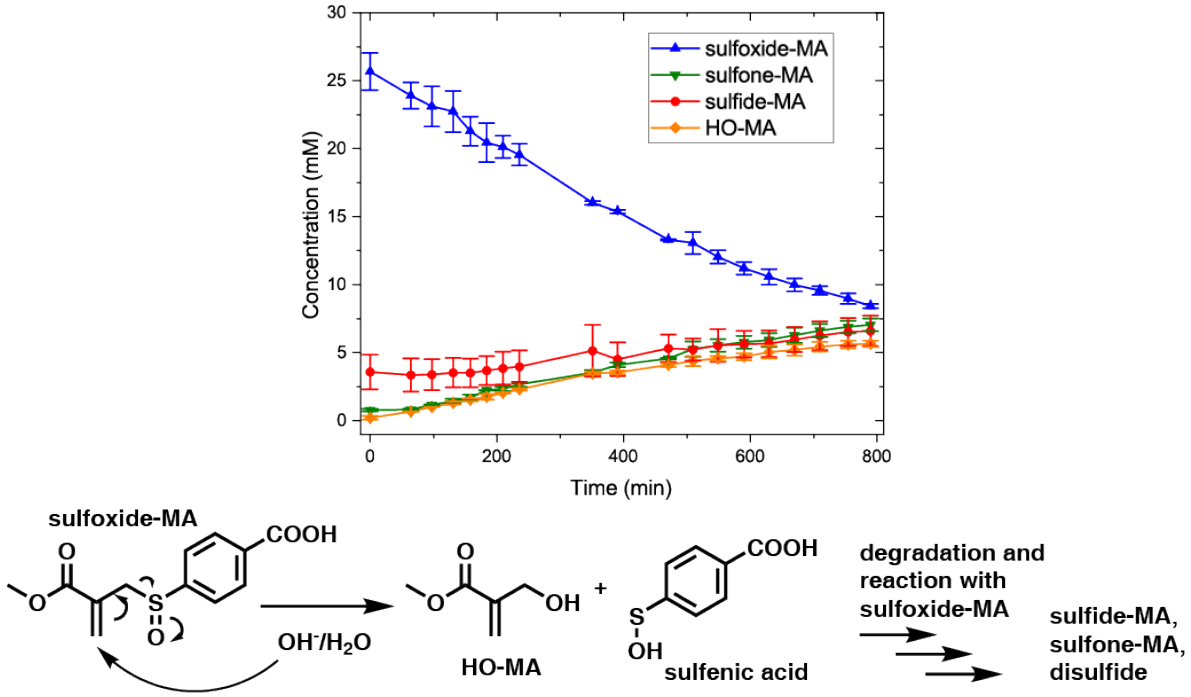


Figure 5.16 | Degradation of sulfoxide-MA 3 in 9/1 0.5 M pH 8.0 buffer/DMF. Unlike sulfide-MA 2 and sulfone-MA 4, sulfoxide-MA 3 is not stable in solution and degrades via solvent attack to yield HO-MA and the common side products of sulfenic acid degradation (refer to Scheme 5.2). Note: The initial presence of sulfide-MA 2 is due to incomplete oxidation of this batch of sulfoxide-MA 3. However, this plays no role for the degradation of sulfoxide-MA 3.

Oxone: reference reactions

To demonstrate potential side reactions of our sulfone CRN, we performed several reference reactions of some species with Oxone. Here, we show how proline as well as proline-MA **1** react with Oxone. Furthermore, we show that acetate (side product of proline-MA **1** formation) and the used solvent mixture (i.e. buffer/DMF) are compatible with Oxone.

First, we subjected proline to 2.5 eq. Oxone. It has been shown in literature that the oxidation of amino acids and subsequent decarboxylation to nitrones typically proceeds via *N*-hydroxylamino acids and a short-lived intermediate dihydroxylated species before decarboxylating.[52] It has also been showing that Oxone specifically is capable of producing a

nitrone from proline.[48] In agreement with literature data, we found the formation of the nitrone upon oxidation of proline. However, we were unable to achieve full conversion even with 2.5 eq. Oxone. Furthermore, we were unable to observe any potential intermediate species as the conversion occurred too fast to be tracked appropriately via ^{1}H-NMR (Figure 5.17).

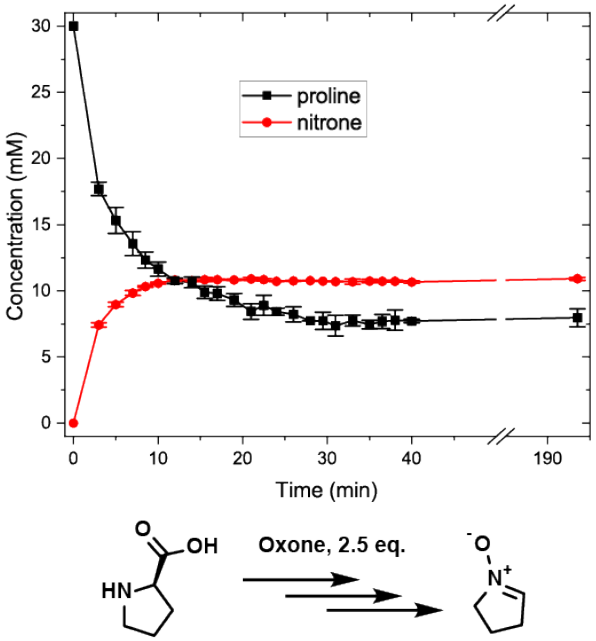


Figure 5.17 | Reaction of proline with 2.5 eq. Oxone. The formation of nitrone was observed, tracked by the appearance of the characteristic double bond signal at ~ 7 ppm.

Besides the NMR results, we also managed to confirm the nitrone via HRMS from the reaction of proline-MA **1** with Oxone as side product:
ESI-LC/HRMS (m/z): calculated for [$C_4H_8NO^+$] 86.0600, found: 86.0606.

Next, we subjected Pro-MA **1** to Oxone. It can be seen that a swift oxidation of Pro-MA **1** upon addition of Oxone takes place to form Ox-Pro-MA (see **Figure 5.18**, see also Scheme 5.2). We also managed to confirm this oxidized species via HRMS:
ESI-LC/HRMS (m/z): calculated for [$C_{10}H_{16}NO_5^+$] 230.1023, found: 230.1015.
In alignment with the mechanism proposed in literature,[52] a decarboxylation as in the oxidation of free proline cannot take place due to the tertiary nature of the amine in the proline-MA **1** species. Therefore, the oxidation stops at the stage of the quaternary oxidized Pro-MA.

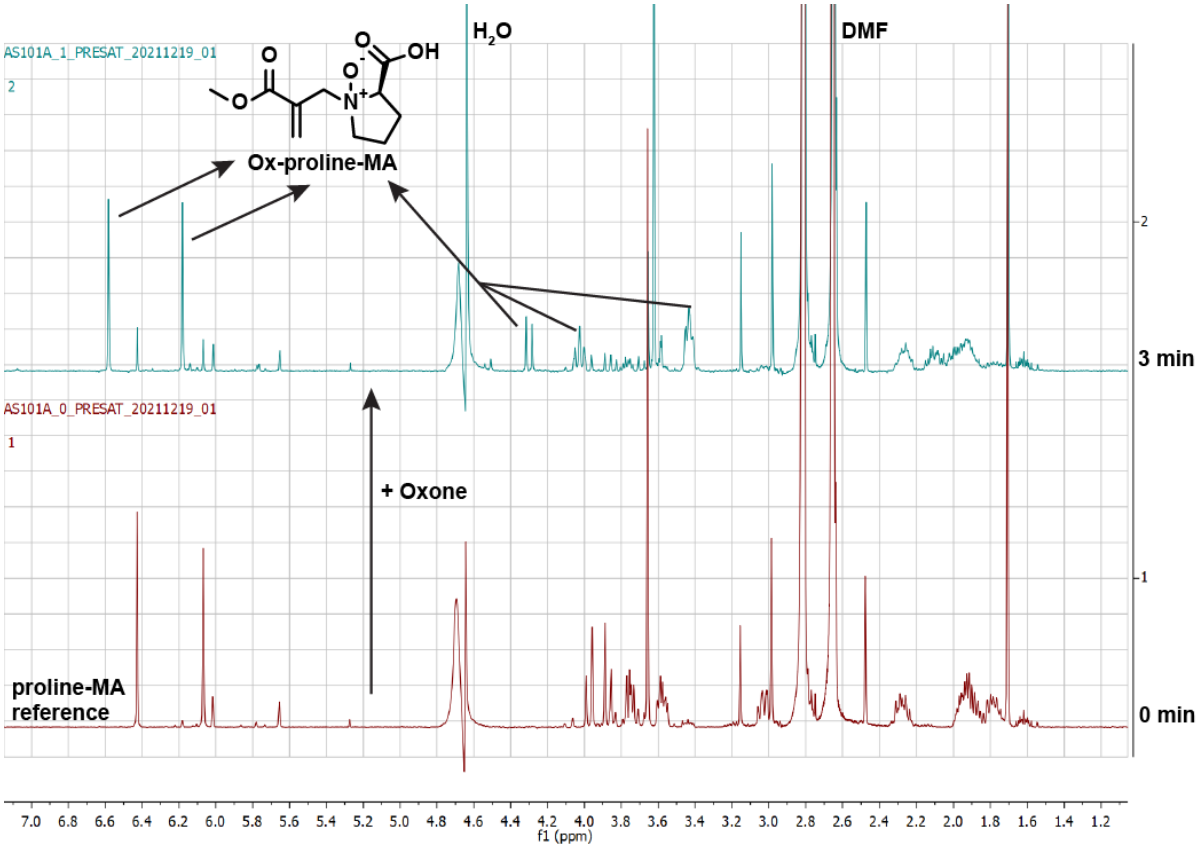


Figure 5.18 | Reaction of Pro-MA 1 with Oxone to yield oxidized Pro-MA.

Lastly, we also subjected sodium acetate, as well as our solvent mixture (9/1 0.5 M pH 8.0 potassium phosphate buffer/DMF) to Oxone. Neither of those experiments showed any changes over time, ruling out any potential cross-reactivity between acetate (side product of proline-MA **1** preparation) and Oxone as well as the used solvent and Oxone.

Recovery of Pro-MA 1 from Ox-Pro-MA

As shown in Scheme 5.2, we have added a backwards arrow leading back from oxidized Pro-MA to Pro-MA **1**. We have found that upon oxidation of Pro-MA **1**, the oxidized species becomes electron-deficient and hence proline can re-attack, leading to recovery of Pro-MA **1**. As can be seen in Figure 5.19, Pro-MA **1** quickly gets oxidized to form oxidized Pro-MA (compare Scheme 5.2). Further, this signal decreases in intensity and a new signal appears. We hypothesize that the oxidized Pro-MA undergoes an isomerization via nucleophilic attack of the amine-*N*-oxide on a second oxidized Pro-MA to yield the *O*-bonded form as the thermodynamic product.
Upon addition of proline, Pro-MA **1** gets recovered quickly from the oxidized Pro-MA species via nucleophilic attack and substitution of *N*-hydroxyproline. The reason why this recovery is less efficient when trying to run the Oxone CRN as a reaction cycle is the fact that proline, which is needed for the recovery of Pro-MA **1**, also gets consumed by the Oxone that is added to the reaction mixture.

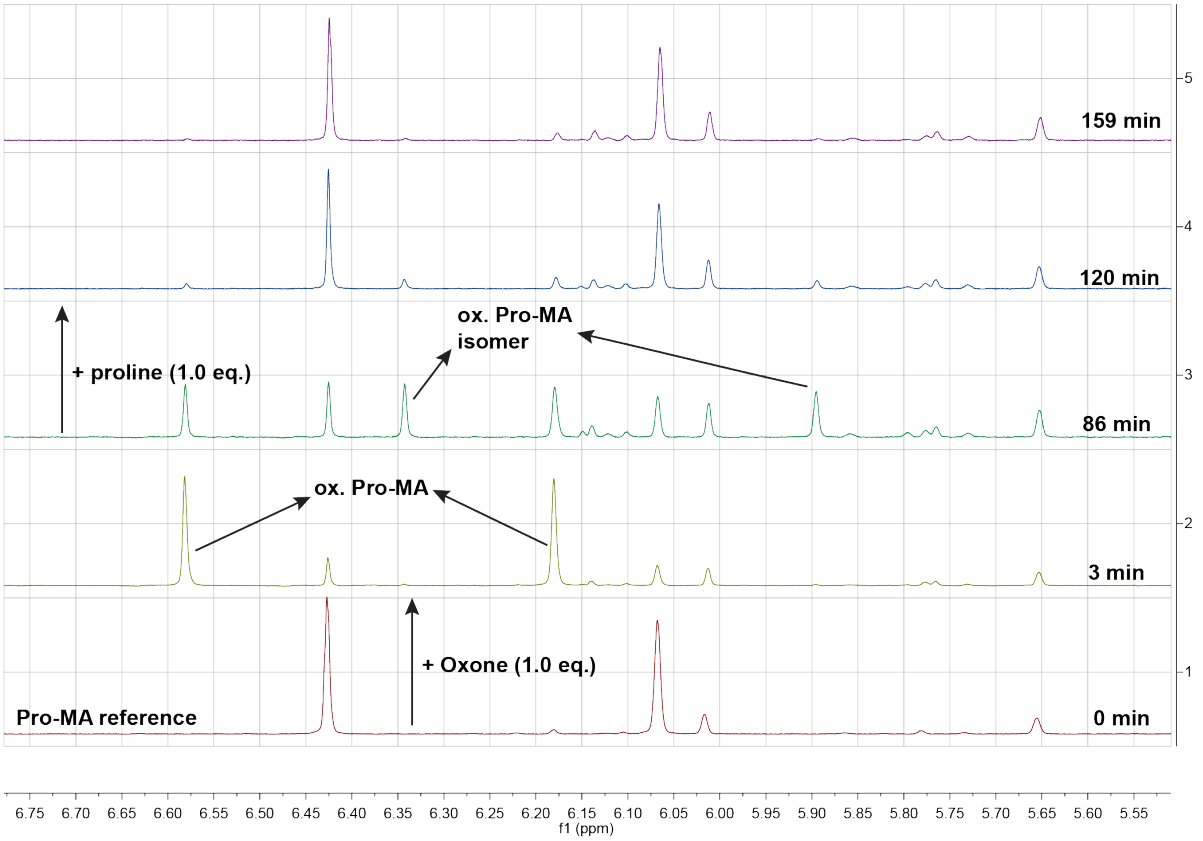


Figure 5.19 | Oxidation of Pro-MA 1 with Oxone and subsequent addition of proline to recover Pro-MA 1.

Addition reactions without substitution: Loss of double bond functionality

As can be seen in Scheme 5.2, one of the side reactions which causes a loss of double bond functionality is the addition of a nucleophile (amine, thiol) to any MA-species without a substitution (i.e. expelling of leaving group) taking place. This side reaction is difficult to track via NMR due to the absence of acrylic signals and strong overlap of the remaining signals. Hence, we resorted to HRMS to prove the formation of double adducts.

First, we investigated to behavior of Pro-MA **1** over time in the presence of 1.0 eq. free proline (Figure 5.20). We found a decrease of roughly 15 % in total MA-moiety concentration over 29 h when adding 1.0 eq. proline to a 30 mM solution of Pro-MA **1**. Quantifying the double adduct via NMR is challenging due to strong overlap and absence of acrylic signals. However, we managed to identify the double adduct via HRMS as the main side product:

ESI-LC/HRMS (m/z): calculated for $[C_{15}H_{24}N_2O_6]^+$ 329.1707, found: 329.1693.

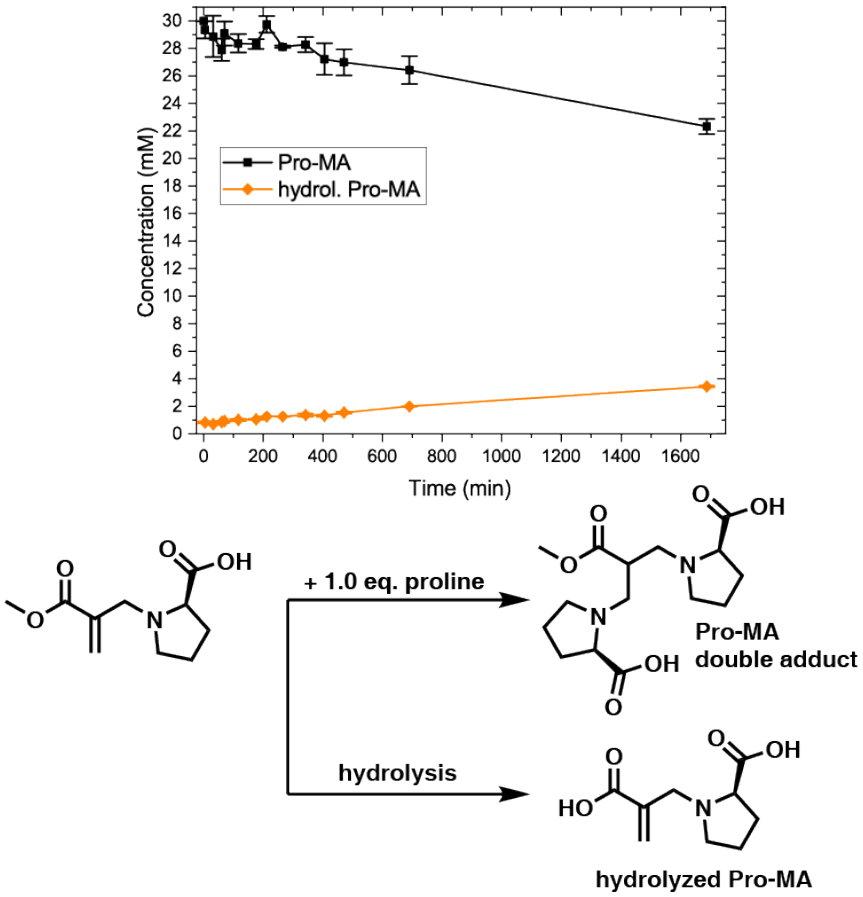


Figure 5.20 | Double addition of proline to Pro-MA 1, along with background hydrolysis. The overall concentration of double bond moieties (i.e. Pro-MA 1 + hydrol. Pro-MA) decreases from 30.0 mM to 25.8 mM.

Next, we tested the behavior of sulfoxide-MA **3** and sulfone-MA **4** in the presence of 2.0 eq. free proline. As the reaction of **3** with proline will also generate sulfide-MA **2** in the process, these experiments also serve to probe for the double addition to **2**. This also explains the non-quantitative yield of Pro-MA **1** + MBA in stepwise additions; a part (usually ~ 10 %) gets lost due to addition side reactions. Indeed, we managed to confirm all species proposed in Scheme 5.2 via HRMS:

- Addition of proline to sulfoxide-MA **3**:
  ESI-LC/HRMS (m/z): calculated for $[C_{17}H_{22}NO_7S]^+$ 384.1111, found: 384.1153.

- Addition of proline to sulfone-MA **4**:
  ESI-LC/HRMS (m/z): calculated for $[C_{17}H_{22}NO_8S]^+$ 400.1061, found: 400.1046; calculated for $[C_{17}H_{20}NO_8S]^-$ 398.0915, found: 398.0918.

- Addition of 4-mercapto benzoic acid (MBA) to sulfide-MA **2**:
  ESI-LC/HRMS (m/z): calculated for $[C_{19}H_{17}O_6S_2]^-$ 405.0472, found: 405.0475.

- Addition of MBA to sulfoxide-MA **3**:

ESI-LC/HRMS (m/z): calculated for $[C_{19}H_{19}O_7S_2]^+$ 423.0567, found: 423.0550; calculated for $[C_{19}H_{17}O_7S_2]^-$ 421.0421, found: 421.0425.

- Addition of MBA to sulfone-MA **4**:
  ESI-LC/HRMS (m/z): calculated for $[C_{19}H_{19}O_8S_2]^+$ 439.0516, found: 439.0502; calculated for $[C_{19}H_{17}O_8S_2]^-$ 437.0370, found: 437.0373.

Bromate: reference reactions

To further corroborate the way in which the third CRN proceeds (Figure 5.2, right), we reacted potassium bromate with sulfide-MA, with MBA and with proline. We found that unlike stronger oxidants, bromate is not capable to oxidize sulfide-MA, neither to sulfoxide-MA nor sulfone-MA (see Figure 5.21). This was expected, as bromate is typically not the active oxidizer, but requires acidic conditions and the presence of bromide ions to comproportionate into the more reactive bromine. At pH 8 and without bromide ions, these conditions are not given. Also, unlike Oxone, bromate does not react with proline (see Figure 5.22), avoiding the side reaction of forming oxidized Pro-MA as it is found with stronger oxidants. However, bromate efficiently oxidizes the thiol MBA to form the corresponding disulfide of MBA (see Figure 5.23).

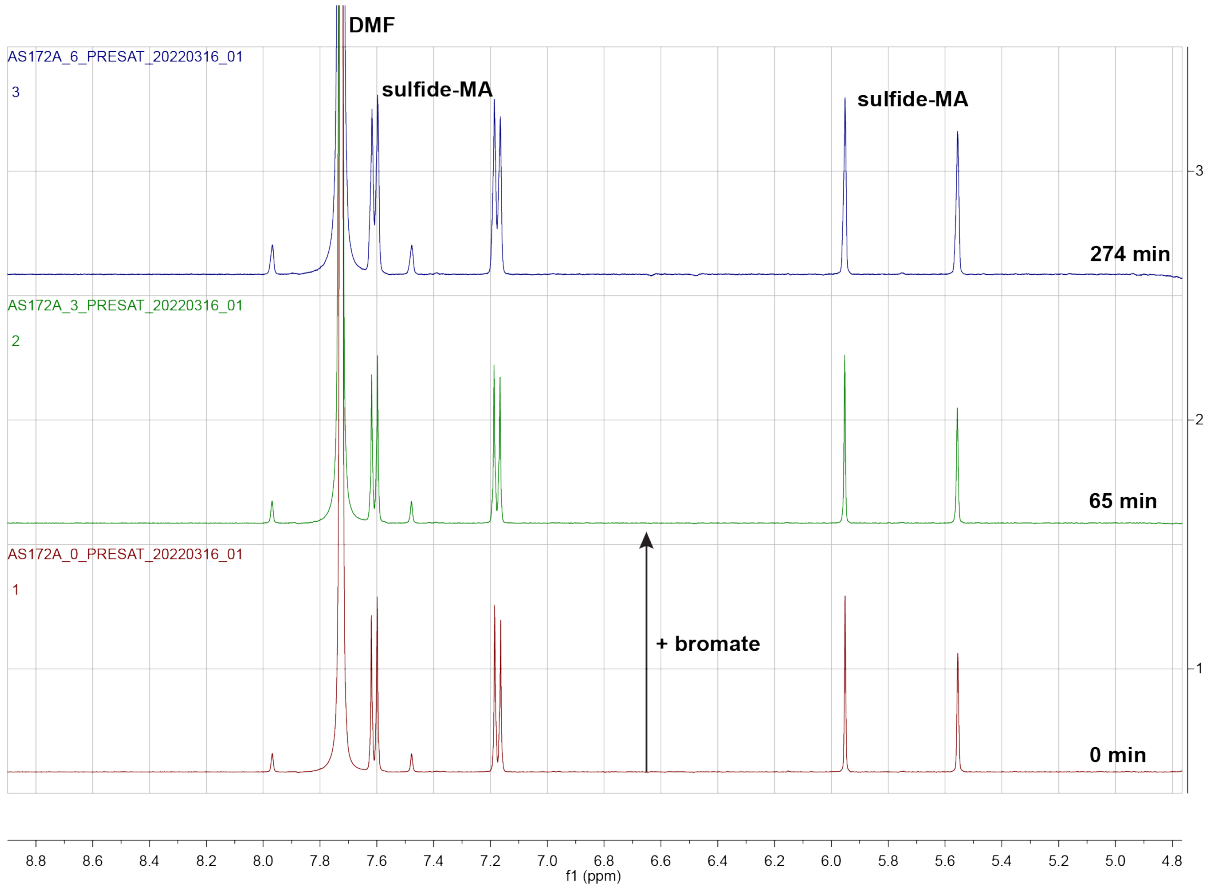


Figure 5.21 | sulfide-MA with 2.5 eq. potassium bromate. No oxidation of sulfide-MA could be observed at pH 8.0 over the observed timeframe.

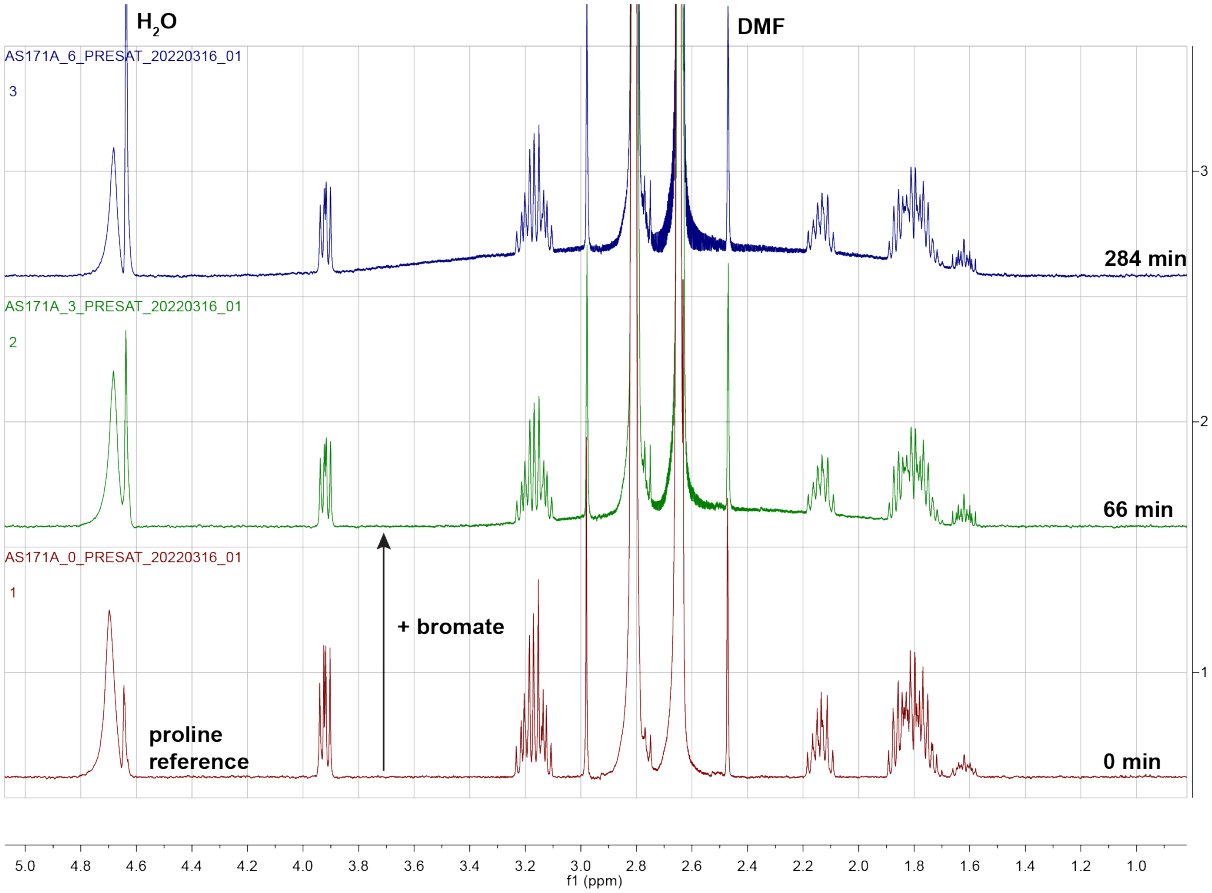


Figure 5.22 | proline with 2.5 eq. potassium bromate. No oxidation of proline could be observed at pH 8.0 over the observed timeframe.

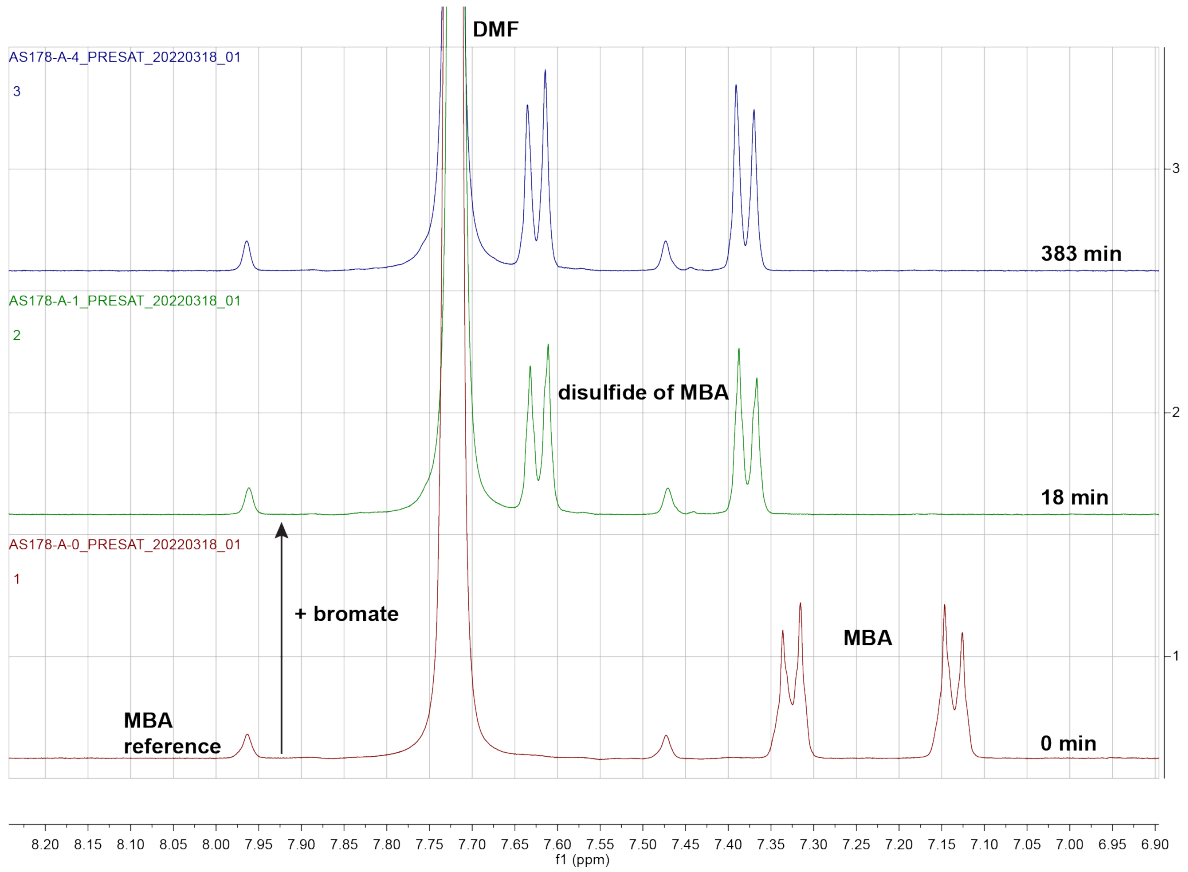


Figure 5.23 | MBA with 2.5 eq. potassium bromate. MBA was quickly oxidized to the corresponding disulfide of MBA.

Bromate RC with aliphatic thiol

After we managed to identify satisfying conditions for our bromate RC with our typically employed thiol MBA and proline, we tested the influence of using a different thiol. We started the cycle with proline, but using an aliphatic thiol – 3-mercaptopropionic acid – instead of MBA.

We found that with 3-mercaptopropionic acid (3MPrA), unlike with MBA, no significant recovery could be observed (see Figure 5.24). We hypothesize that this has to do with the acidity of the employed thiol. Under nucleophilic attack from proline, MBA is a much better leaving group due the high charge stabilization of its thiolate. The much lower acidity probably pushes the equilibrium between Pro-MA and 3MPrA-MA even further towards the side of the sulfide, strongly disfavoring the thiolate as a leaving group and hence not leading to any recovery of initial Pro-MA over the observed timescale.

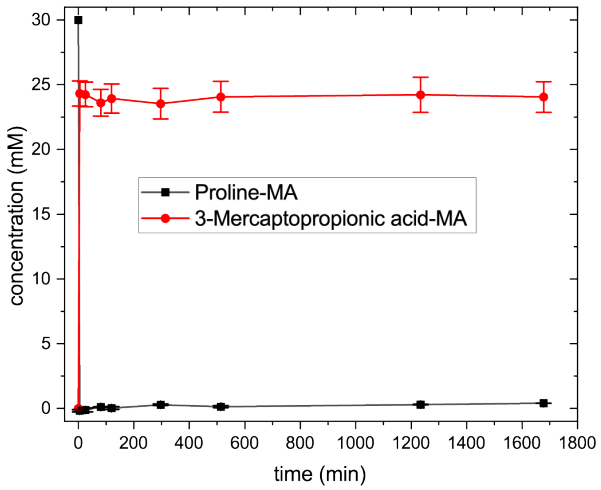


Figure 5.24 | simultaneous addition of 1.0 eq. of 3MPrA and 2.5 eq. potassium bromate to a 30 mM solution of Pro-MA. While the sulfide forms efficiently and fast as with MBA, no recovery of Pro-MA can be observed over time.

Stepwise addition: comparison of oxidants

To directly compare the efficiency of different oxidants of recovering Pro-MA over time, we plotted the recoveries of Pro-MA over time from stepwise addition experiments, adding 2.5 eq. of oxidant in each case. In each case, we started from the usual 30 mM Pro-MA stock solution, added 1.0 eq. MBA, followed by addition of a certain amount of oxidant (specified below).

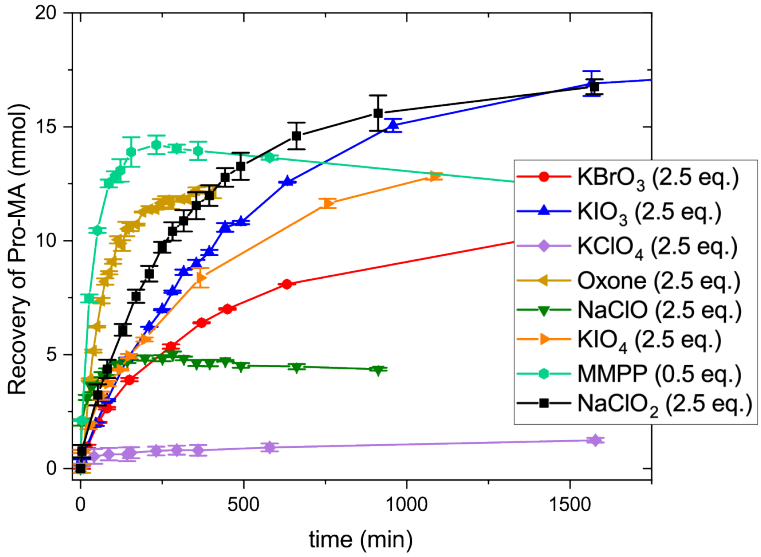


Figure 5.25 | Comparison of different oxidants in the recovery of Pro-MA upon addition of 1.0 eq. of MBA to a 30 mM Pro-MA stock solution, followed by addition of the oxidant.

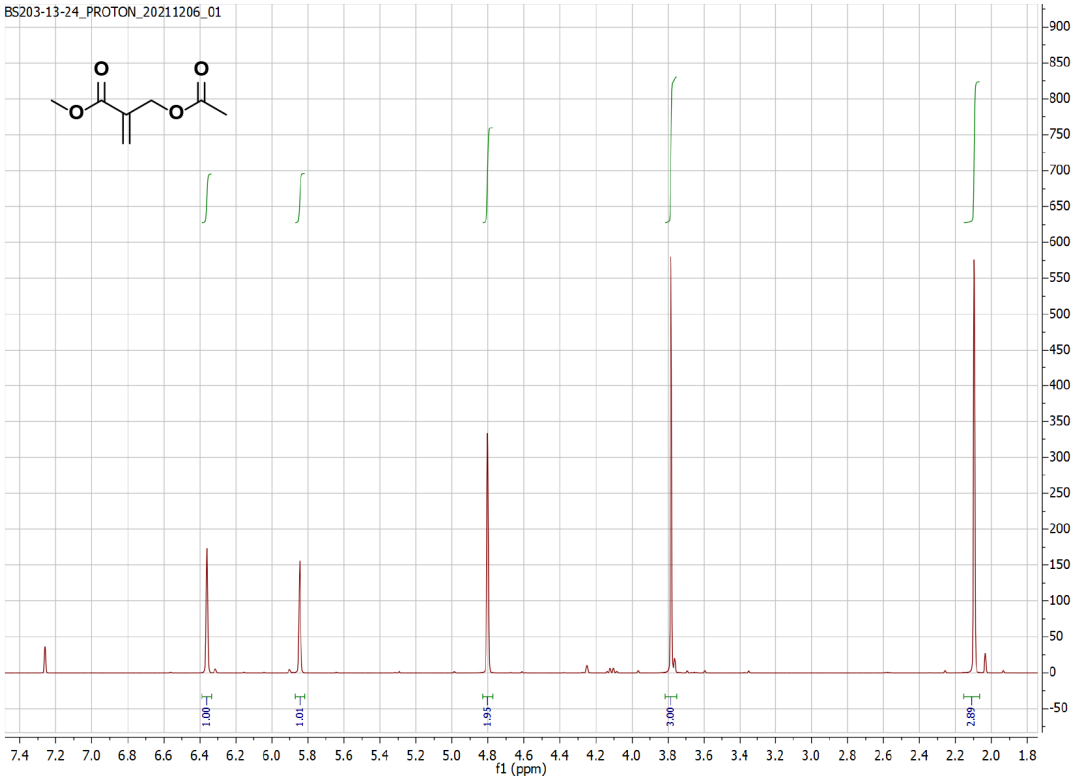
It can clearly be seen that the rates of recovery, the maximum level of recovery and the product stability over time all vary widely, depending on the oxidant. $KClO_4$ (violet line) and $KClO_3$ (omitted for clarity) do not work in recovering Pro-MA **1** from sulfide-MA **2**, likely due to their weak oxidation strength under the used conditions compared to the other oxidants used. $KBrO_3$ and KIO_3 recover Pro-MA **1** relatively slowly, as the recovery mechanism proceeds

purely via oxidation of free thiol from equilibrium. As the concentration of free thiol and the rate of equilibration limit the overall rate of recovery, a precise control of the recovery kinetics via variation of the oxidant quantity is not possible in the sulfide pathway. The oxidants that can oxidize sulfide-MA **2** to sulfoxide-MA **3** and/or sulfone-MA **4** lead to a more rapid recovery of Pro-MA **1**, as the recovery of **1** depends on the direct displacement of sulfenic and/or sulfinic acid by proline instead of release of MBA. As we have shown, Pro-MA can be recovered similarly fast from sulfoxide-MA and sulfone-MA.

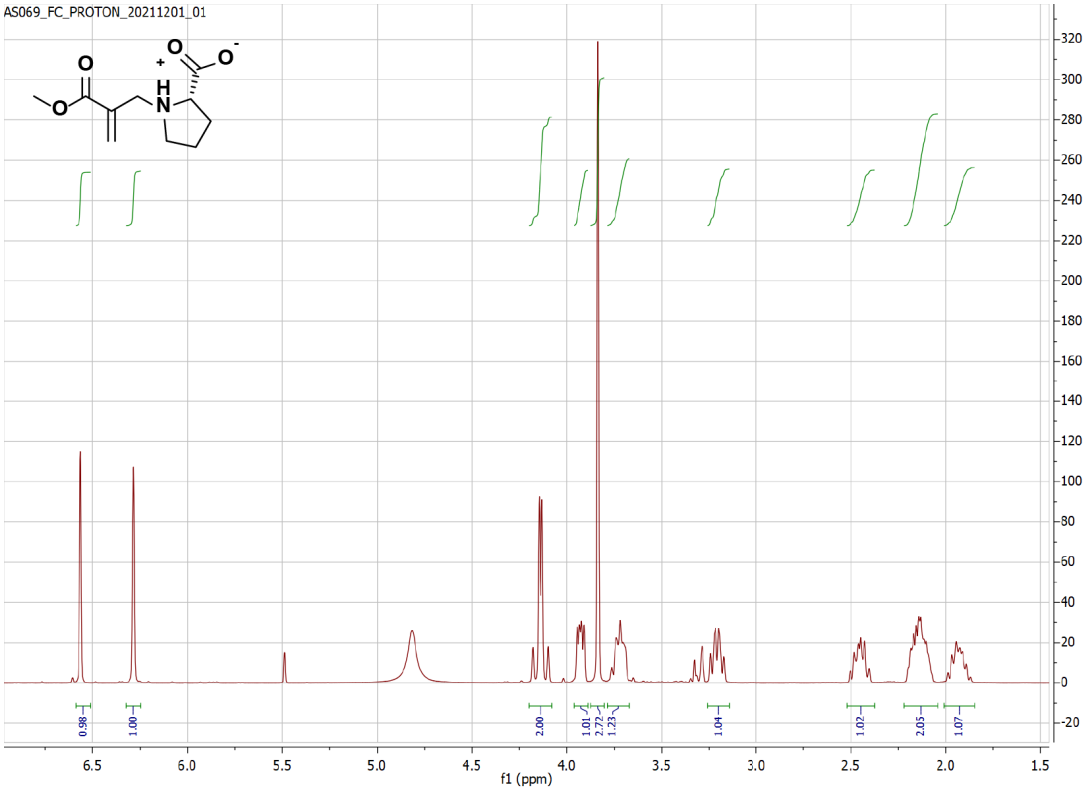
As formation of sulfoxide-MA is faster than that of sulfone-MA with strong oxidants, the amount of oxidant used is another crucial factor. A good example is MMPP. While adding 2.5 eq. (omitted for clarity) led to a whole range of side reactions, as well as low and slow recovery rates, the use of 0.5 eq. led to a fast recovery of Pro-MA. However, stability of recovered Pro-MA was low over time. Likewise, for sodium hypochlorite, we also observed side reactivity, low recovery rates and low stability of recovered Pro-MA over time. Oxone on the other hand led to a relatively fast recovery rate, high recovery level and improved stability over time.

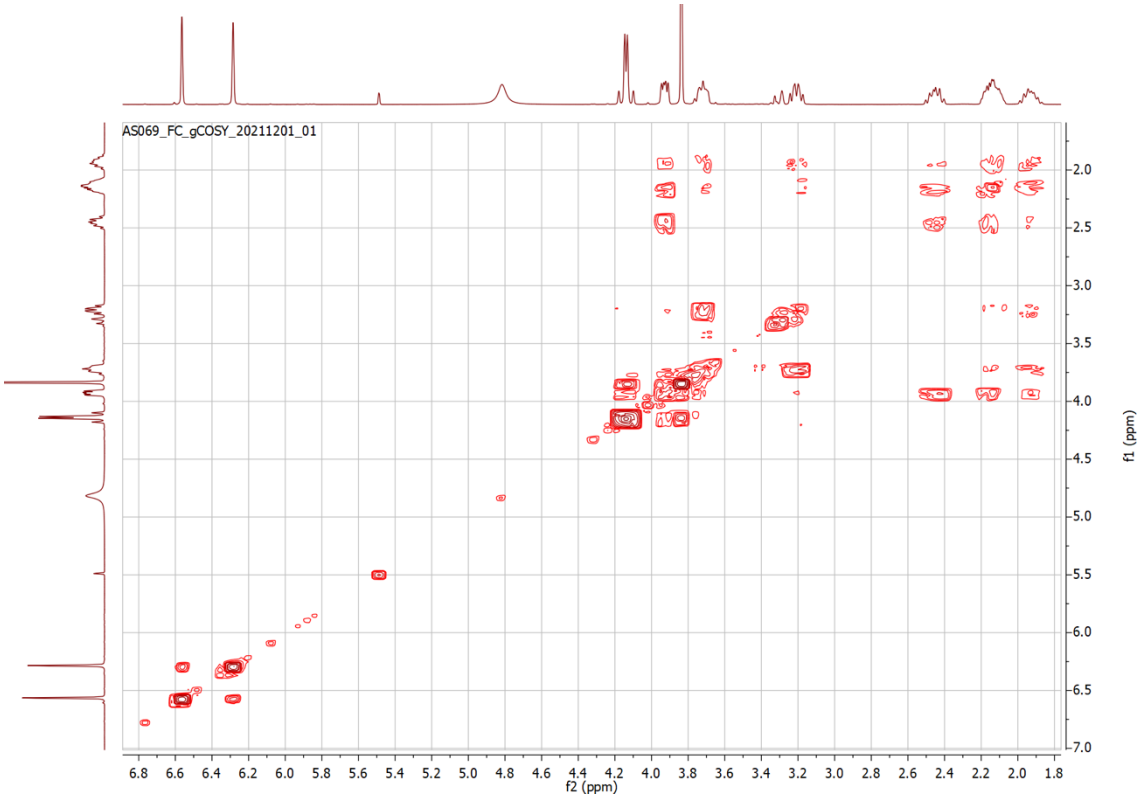
Hence, it can be concluded that by varying the applied oxidant, one has control over which pathway the recovery will take (i.e. slow sulfide pathway vs the much faster sulfoxide and sulfone pathways). Furthermore, the amount of strong oxidant (i.e. sulfone pathway) used, can kinetically bias the pathway; using stoichiometric or substochiometric amounts of oxidant can bias recovery mode more towards the sulfoxide pathway, while using an excess of oxidant will bias the recovery mode more towards the sulfone pathway. These factors strongly influence side products, recovery rates and kinetics of the recovery.

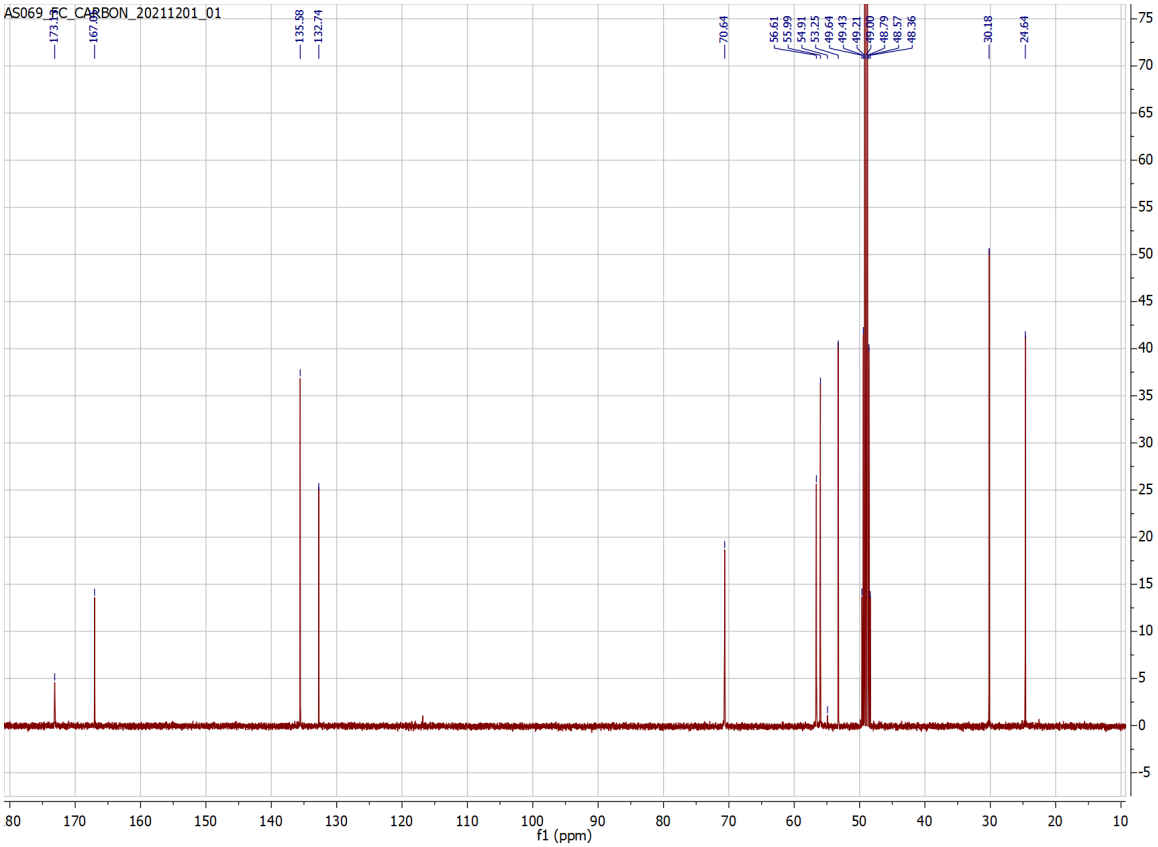
^{1}H-NMR spectrum of methyl (2-acetoxymethyl)acrylate (AcO-MA) in $CDCl_3$:

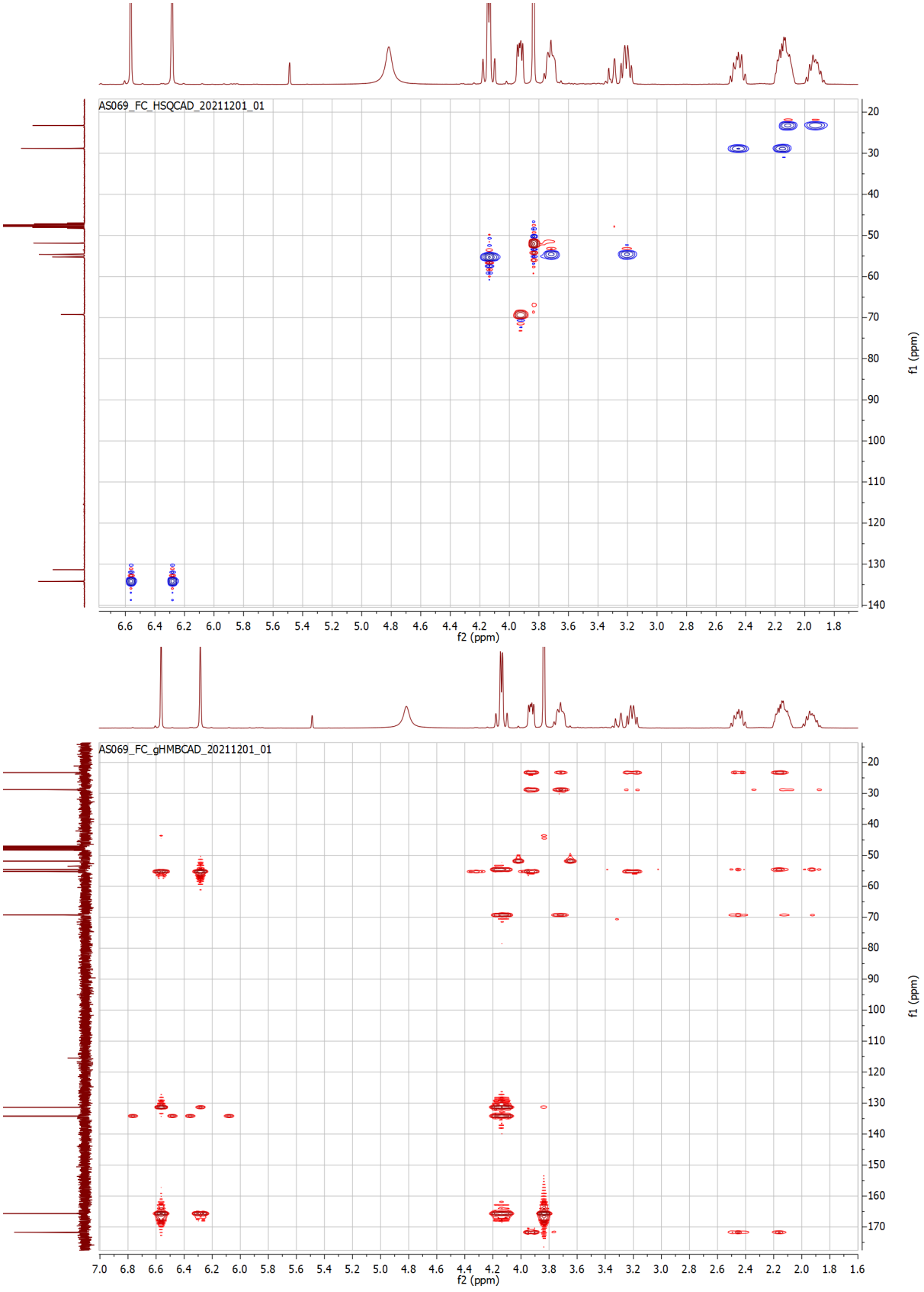


^{1}H-NMR, COSY, ^{13}C, HMBC and HSQC spectra of (2-(methoxycarbonyl)allyl)-L-proline (Pro-MA) in MeOD-d_4:

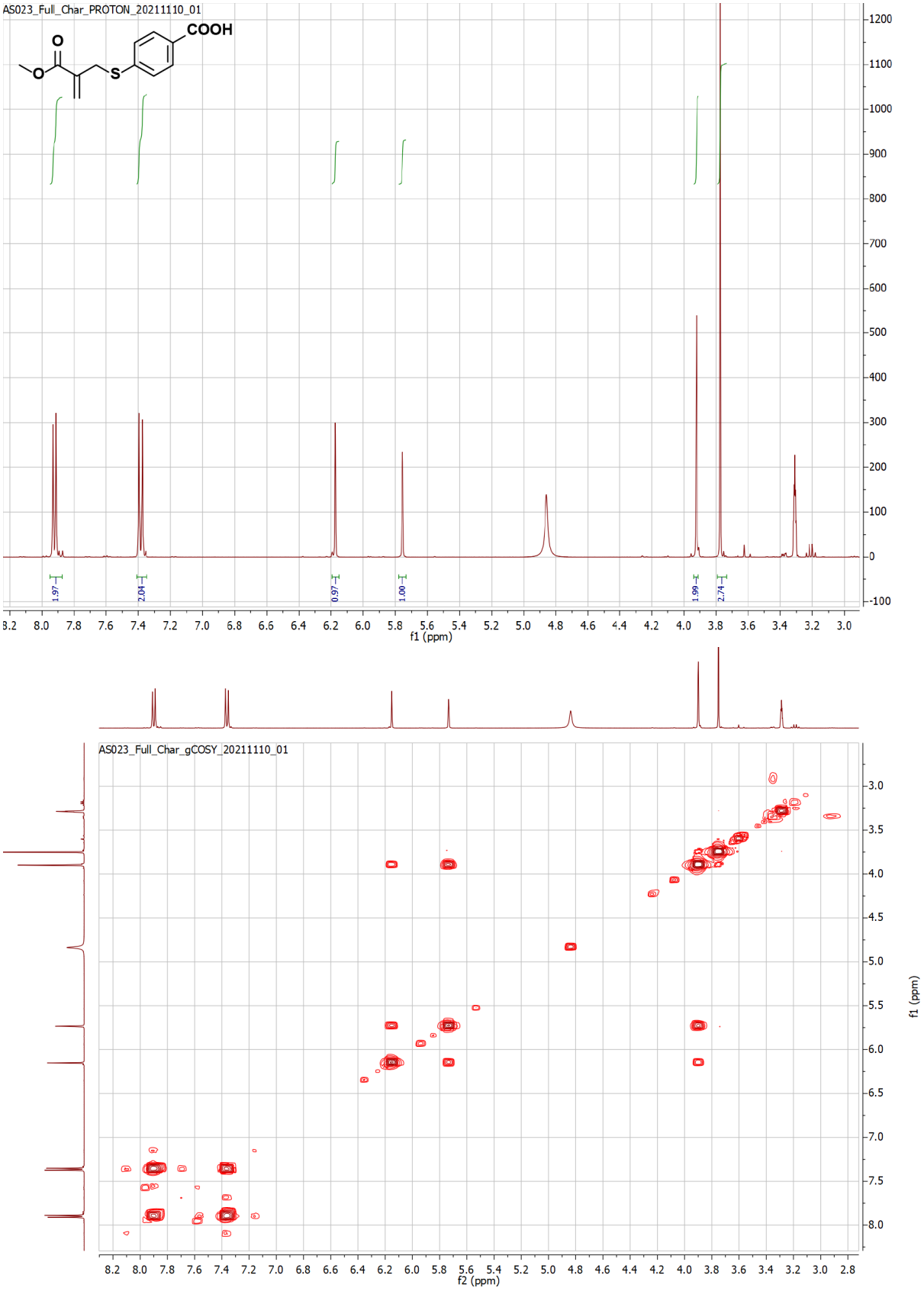


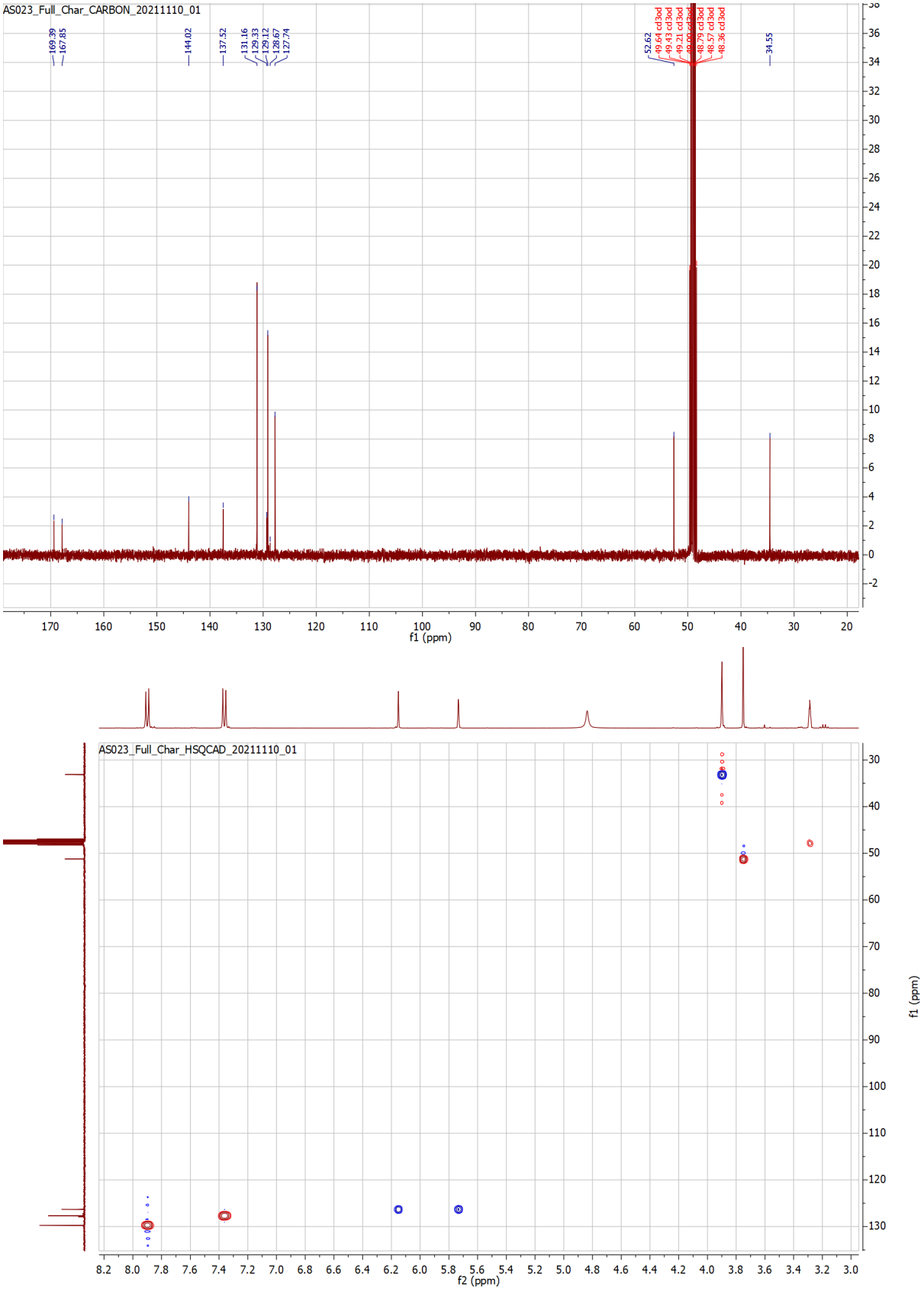

AS069_FC_gCOSY_20211201_01
f2 (ppm)
f1 (ppm)


AS069_FC_CARBON_20211201_01
173.12
167.85
135.58
132.74
70.64
56.61
55.99
54.91
53.25
49.64
49.43
49.21
49.00
48.79
48.57
48.36
30.18
24.64
f1 (ppm)

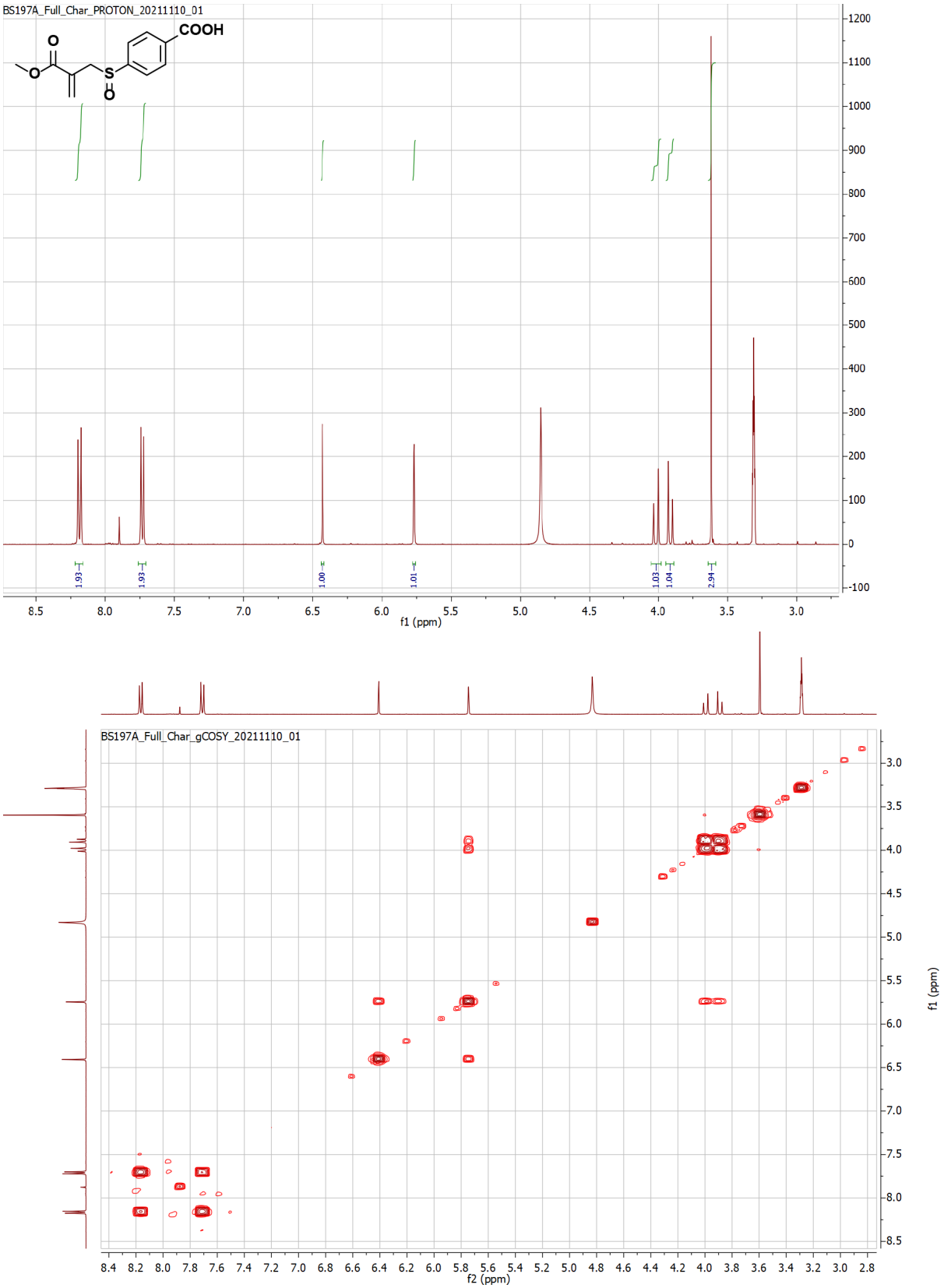


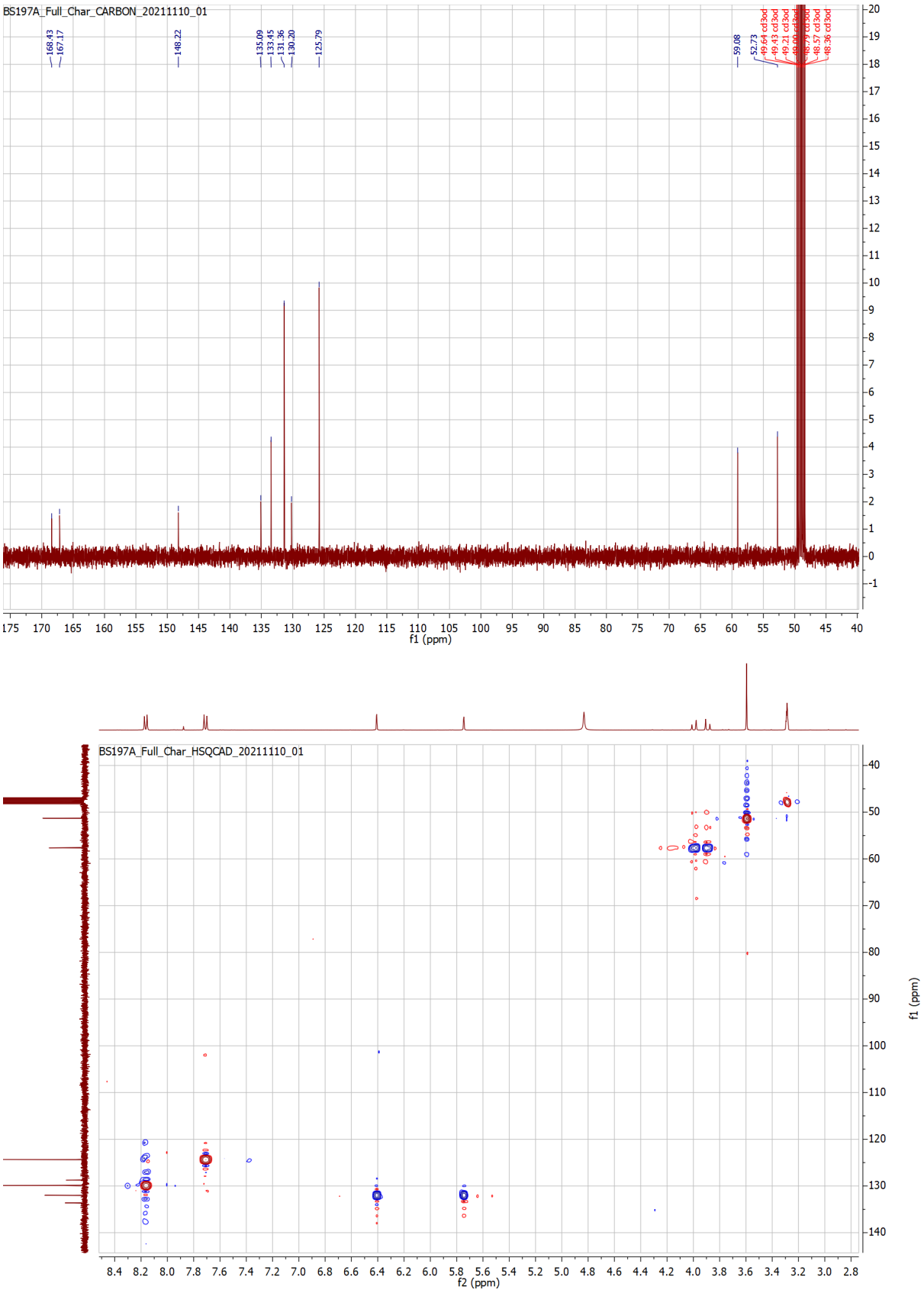
^{1}H-NMR, COSY, ^{13}C and HSQC spectra of 4-((2-(methoxycarbonyl)allyl)thio)benzoic acid (sulfide-MA) in MeOD-d_4:


AS023_Full_Char_PROTON_20211110_01
COOH
O
O
S
1.97
2.04
0.97
1.00
1.99
2.74
f1 (ppm)
AS023_Full_Char_gCOSY_20211110_01
f2 (ppm)
f1 (ppm)

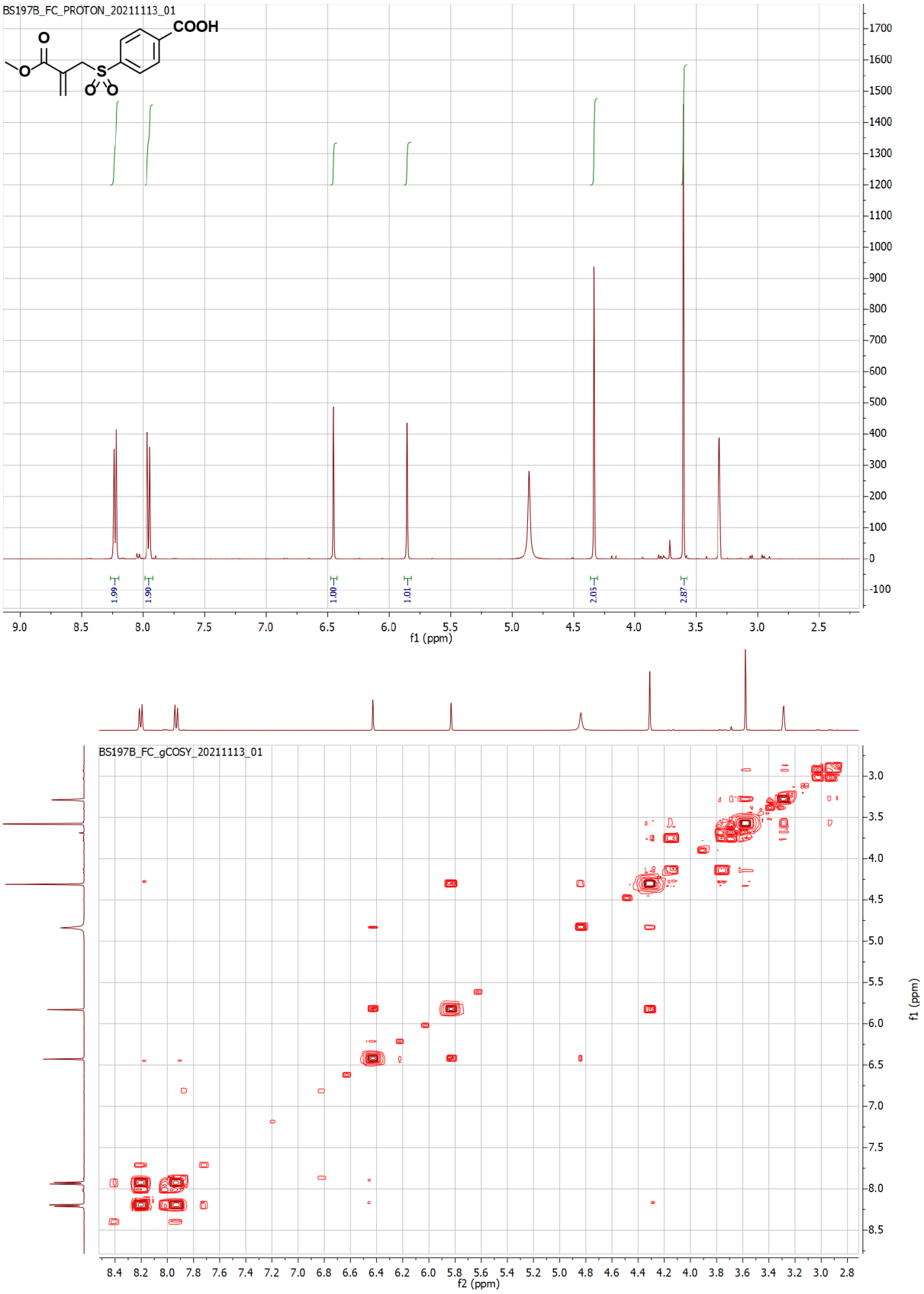

AS023_Full_Char_CARBON_20211110_01
169.39
167.85
144.02
137.52
131.16
129.33
129.12
128.67
127.74
52.62
49.64 cd3od
49.43 cd3od
49.21 cd3od
48.79 cd3od
48.57 cd3od
48.36 cd3od
34.55
f1 (ppm)
AS023_Full_Char_HSQCAD_20211110_01
f2 (ppm)
f1 (ppm)

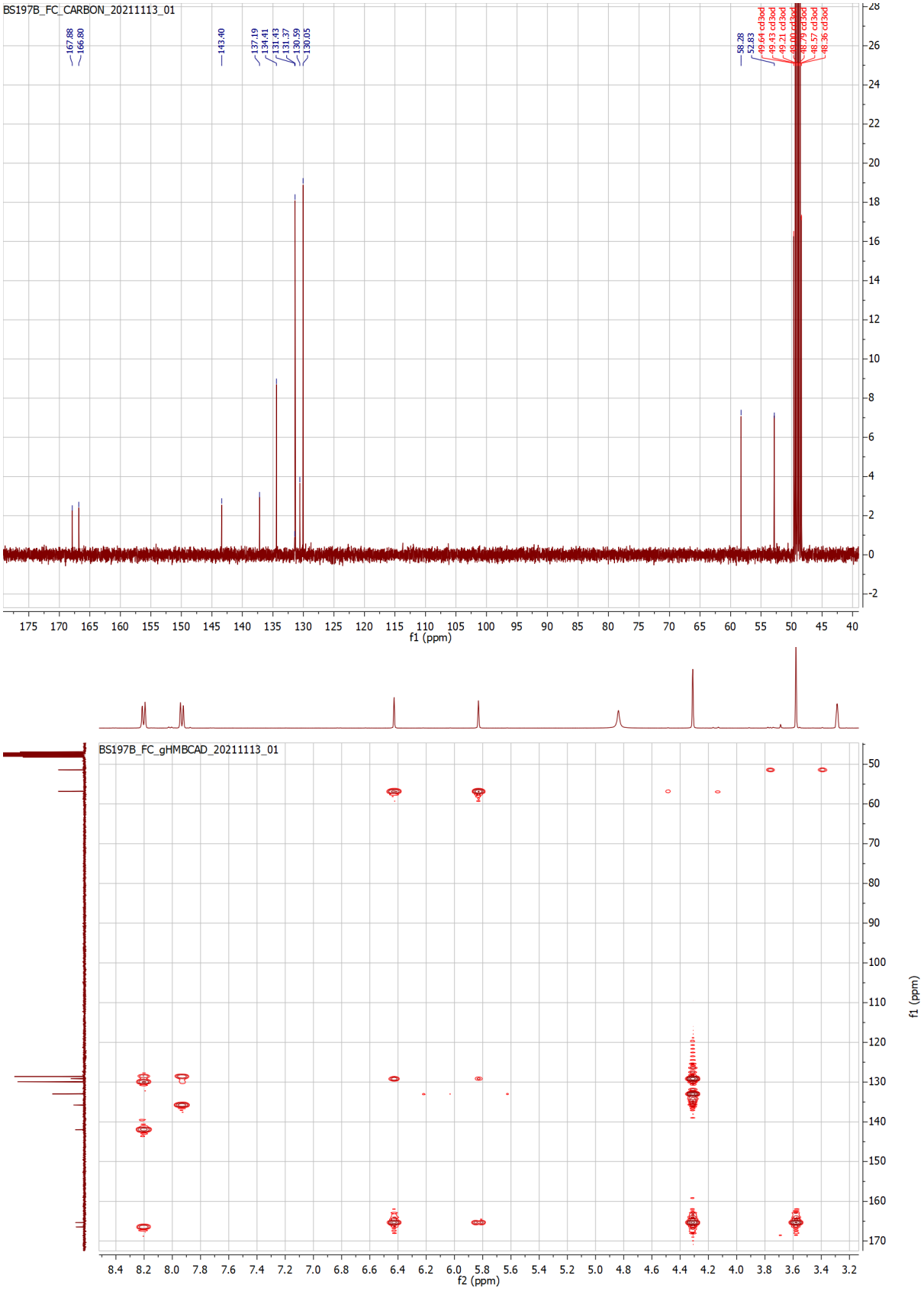
^{1}H-NMR, COSY, ^{13}C and HSQC spectra of 4-((2-(methoxycarbonyl)allyl)sulfinyl)benzoic acid (sulfoxide-MA) in MeOD-d_4:

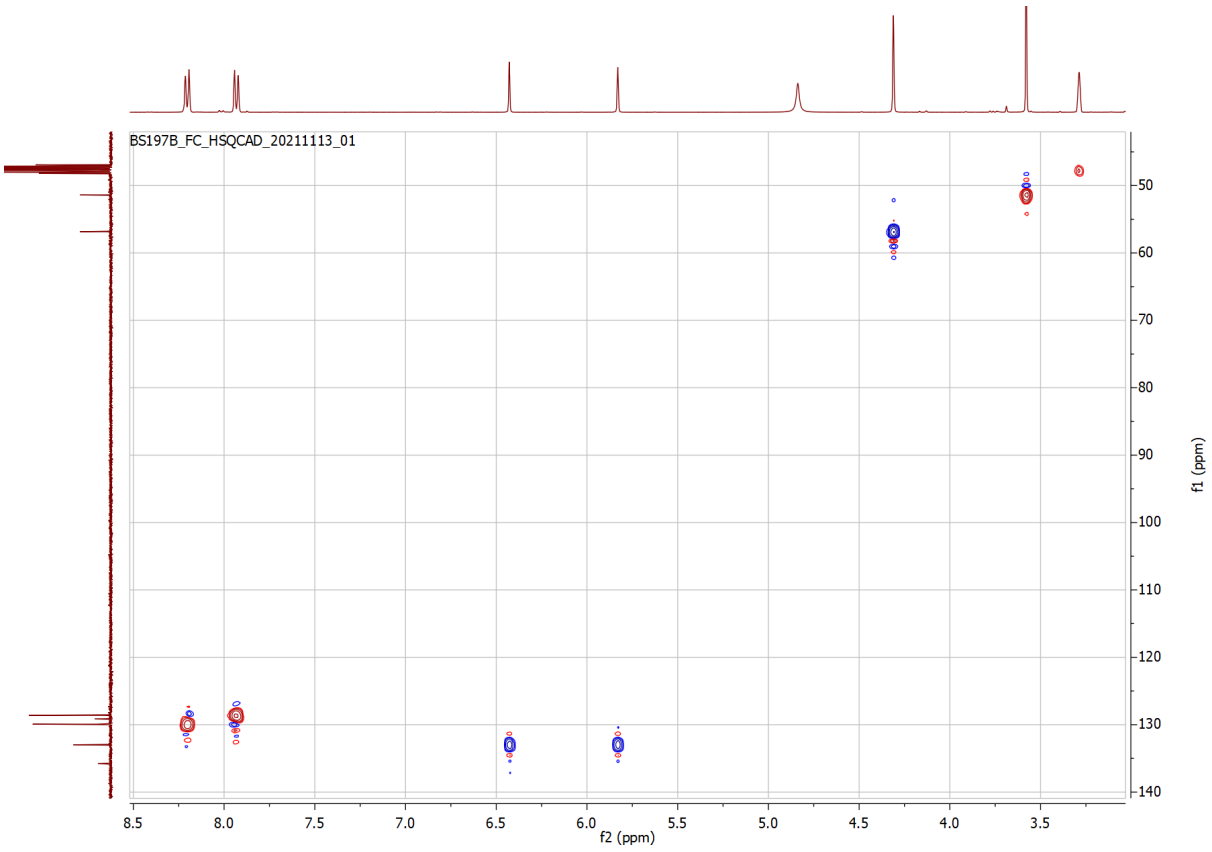



BS197A_Full_Char_CARBON_20211110_01
168.43
167.17
148.22
135.09
133.45
131.36
130.20
125.79
59.08
52.73
49.64 cd3od
49.43 cd3od
49.21 cd3od
48.79 cd3od
48.57 cd3od
48.36 cd3od
f1 (ppm)
BS197A_Full_Char_HSQCAD_20211110_01
f2 (ppm)
f1 (ppm)

^{1}H-NMR, COSY, ^{13}C, HMBC and HSQC spectra of 4-((2-(methoxycarbonyl)allyl)sulfonyl)benzoic acid (sulfone-MA) in MeOD-d_4:




BS197B_FC_CARBON_20211113_01
167.88
166.80
143.40
137.19
134.41
131.43
131.37
130.59
130.05
58.28
52.83
49.64 cd3od
49.43 cd3od
49.21 cd3od
49.00 cd3od
48.79 cd3od
48.57 cd3od
48.36 cd3od
f1 (ppm)
BS197B_FC_gHMBCAD_20211113_01
f2 (ppm)
f1 (ppm)



5.6 References

1. Berg, J. M., Tymoczo, J. L. & Stryer, L. *Biochemistry*. vol. 5 (W. H. Freeman, 1995).
2. Nelson, D. L. & Cox, M. M. *Principles of Biochemistry*. (W. H. Freeman, 2004).
3. Corbett, P. T. *et al.* Dynamic combinatorial chemistry. *Chem. Rev.* **106**, 3652–3711 (2006).
4. Singh, N., Formon, G. J. M., De Piccoli, S. & Hermans, T. M. Devising Synthetic Reaction Cycles for Dissipative Nonequilibrium Self-Assembly. *Adv. Mater.* **32**, (2020).
5. Zhuang, J. *et al.* Chemical Science A programmable chemical switch based on. 2103–2111 (2020) doi:10.1039/c9sc05841a.
6. Diehl, K. L. *et al.* Click and chemically triggered declick reactions through reversible amine and thiol coupling via a conjugate acceptor. *Nat. Chem.* **8**, 968–973 (2016).
7. Nelson, R. P. & Lawton, R. G. On the alpha,alpha' Annelation of Cyclic Ketones. *J. Am. Chem. Soc.* **88**, 3884–3885 (1966).
8. Liberatore, F. A. *et al.* Site-Directed Chemical Modification and Cross-Linking of a Monoclonal Antibody Using Equilibrium Transfer Alkylating Cross-Link Reagents. *Bioconjug. Chem.* **1**, 36–50 (1990).
9. Mitra, S. & Lawton, R. G. Reagents for the crosslinking of proteins by equilibrium transfer alkylation. *J. Am. Chem. Soc.* **101**, 3097–3110 (1979).
10. Brocchini, S. *et al.* PEGylation of native disulfide bonds in proteins. *Nat. Protoc.* **1**, 2241–2252 (2006).
11. Brocchini, S. J., Eberle, M. & Lawton, R. G. Molecular Yardsticks. Synthesis of Extended Equilibrium Transfer Aklylating Cross-link Reagents and Their Use in the Formation of Macrocycles. *J. Am. Chem. Soc.* **110**, 5211–5212 (1988).
12. Wang, T. *et al.* Cross-conjugation of DNA, proteins and peptides via a pH switch. *Chem. Sci.* **4**, 1889–1894 (2013).
13. Campaña, A. G., Leigh, D. A. & Lewandowska, U. One-dimensional random walk of a synthetic

small molecule toward a thermodynamic sink. *J. Am. Chem. Soc.* **135**, 8639–8645 (2013).
14. Jones, M. W. *et al.* Polymeric dibromomaleimides as extremely efficient disulfide bridging bioconjugation and pegylation agents. *J. Am. Chem. Soc.* **134**, 1847–1852 (2012).
15. Nenajdenko, V. G., Krasovskiy, A. L. & Balenkova, E. S. The chemistry of sulfinyl and sulfonyl enones. *Tetrahedron* **63**, 12481–12539 (2007).
16. Johnson, A. M. & Anslyn, E. V. Reversible Macrocyclization of Peptides with a Conjugate Acceptor. *Org. Lett.* **19**, 1654–1657 (2017).
17. Wu, T. *et al.* Chemically Triggered Click and Declick Reactions: Application in Synthesis and Degradation of Thermosetting Plastics. *ACS Macro Lett.* 1125–1131 (2021) doi:10.1021/acsmacrolett.1c00548.
18. Matos, M. J. *et al.* Chemo- and Regioselective Lysine Modification on Native Proteins. *J. Am. Chem. Soc.* **140**, 4004–4017 (2018).
19. Wang, T. *et al.* Water-soluble allyl sulfones for dual site-specific labelling of proteins and cyclic peptides. *Chem. Sci.* **7**, 3234–3239 (2016).
20. Wang, Y., Jiang, W. & Huo, C. One-pot synthesis of β-hydroxysulfones and its application in the preparation of anticancer drug bicalutamide. *J. Org. Chem.* **82**, 10628–10634 (2017).
21. Ligon, S. C. *et al.* Allyl sulfides and α-substituted acrylates as addition-fragmentation chain transfer agents for methacrylate polymer networks. *J. Polym. Sci. Part A Polym. Chem.* **54**, 394–406 (2016).
22. Batra, D. & Shea, K. J. Novel Trifunctional Building Blocks for Fluorescent Polymers. *Org. Lett.* **5**, 3895–3898 (2003).
23. Hussain, H. *et al.* meta-Chloroperbenzoic acid (mCPBA): a versatile reagent in organic synthesis. (2014) doi:10.1039/c3ra45702h.
24. Kaushik, R., Khan, I., Gupta, R. & Hussain, F. Synthesis of Na22[((RE)4(H2O)6Sb6O4)(SbW10O37)2(SbW8O31)2] [(RE)III = Y, Gd, Tb, Dy, Ho, Er, Tm, Yb]: Synthesis, characterization and catalytic application. *Inorg. Chem. Commun.* **122**, 108265 (2020).
25. Chu, J. W. & Trout, B. L. On the Mechanisms of Oxidation of Organic Sulfides by H2O 2 in Aqueous Solutions. *J. Am. Chem. Soc.* **126**, 900–908 (2004).
26. Liu, G. Z. *et al.* Stereoselective synthesis and anti-proliferative effects on prostate cancer evaluation of 5-substituted-3,4-diphenylfuran-2-ones. *Eur. J. Med. Chem.* **65**, 323–336 (2013).
27. Khurana, J. M., Panda, A. K., Ray, A. & Gogia, A. RAPID OXIDATION OF SULFIDES AND SULFOXIDES WITH SODIUM HYPOCHLORITE. *https://doi.org/10.1080/00304949609356529* **28**, 234–237 (2009).
28. Ali, M. H. & Bohnert, G. J. Chemoselective Oxidation of Sulfides to Sulfones with Magnesium Monoperoxyphthalate (MMPP) On Silica Gel Support in Methylene Chloride Solvent. *http://dx.doi.org/10.1080/00397919808004877* **28**, 2983–2998 (2006).
29. Shaabani, A., Bazgir, A., Soleimani, K. & Salehi, P. Solvent effects in the oxidation of sulfides with NaBrO 3/Mg(HSO4)2. *Synth. Commun.* **33**, 2935–2944 (2003).
30. Shaabani, A., Safaei, H. R. & Bagzir, A. Oxidation of Sulfides to Sulfoxides by NaBr03-NH4Cl in Aqueous Acetonitrile. *Iran. J. Chem. Chem. Eng.* **19**, 47–50 (2000).
31. Challenger, S., Derrick, A., Mason, C. P. & Silk, T. V. Stereoselective Synthesis of a Candoxatril Intermediate via Asymmetric Hydrogenation. *Tetrahedron Lett.* **40**, 2187–2190 (1999).
32. Nagy, P. & Winterbourn, C. C. Redox chemistry of biological thiols. *Adv. Mol. Toxicol.* **4**, 183–222 (2010).
33. Fan, B., Men, Y., Rossum, S. A. P., Li, G. & Eelkema, R. A Fuel-Driven Chemical Reaction Network Based on Conjugate Addition and Elimination Chemistry. *ChemSystemsChem* **2**, (2020).
34. Pinsky, M. L. & Hu, H. Evaluation of the Chloroisocyanurate Hydrolysis Constants. *Environ. Sci. Technol.* **15**, 423–430 (1981).
35. Pati, S. C. & Sarangi, C. Mechanism of Self-decomposition of Trichloroisocyanuric Acid in Acid Medium : A Kinetic Study. *Indian J. Chem.* **27**, 593–596 (1988).
36. Veisi, H., Sedrpoushan, A., Hemmati, S. & Kordestani, D. Trichloroisocyanuric Acid (TCCA) and N-Chlorosuccinimide (NCS) as Efficient Reagents for the Direct Oxidative Conversion of Thiols and Disulfides to Sulfonyl Chlorides. *http://dx.doi.org.tudelft.idm.oclc.org/10.1080/10426507.2011.621910* **187**, 769–775 (2012).
37. Zhang, C., Brown, P. J. B. & Hu, Z. Thermodynamic properties of an emerging chemical disinfectant, peracetic acid. *Sci. Total Environ.* **621**, 948–959 (2018).
38. Leonard, N. J. & Johnson, C. R. Periodate Oxidation of Sulfides to Sulfoxides. Scope of the Reaction.

J. Org. Chem. **27**, 282–284 (1962).
39. Reddy, K. R., Rajasekhar, C. V & Ravindra, A. L-Proline – H2O2 : A New Chemoselective Approach for Oxidation of Sulfides to Sulfoxides. *Synthetic* **36**, 3761–3766 (2006).
40. Kice, J. L., Venier, C. G., Large, G. B. & Heasley, L. Mechanisms of Reactions of Thiolsulfinates (Sulfenic Anhydrides). III. The Sulfide-Catalyzed Disproportionation of Aryl Thiosulfinates. *J. Am. Chem. Soc.* **91**, 2028–2035 (1969).
41. Kapungu, G. P. *et al.* Oxyhalogen-sulfur chemistry: Kinetics and mechanism of oxidation of captopril by acidified bromate and aqueous bromine. *J. Phys. Chem. A* **117**, 2704–2717 (2013).
42. Zhang, H., Han, M., Chen, T., Xu, L. & Yu, L. Poly(N-isopropylacrylamide-co-l-proline)-catalyzed Claisen-Schmidt and Knoevenagel condensations: Unexpected enhanced catalytic activity of the polymer catalyst. *RSC Adv.* **7**, 48214–48221 (2017).
43. Trausel, F. *et al.* Chemical signal activation of an organocatalyst enables control over soft material formation. *Nat. Commun.* **8**, 1–7 (2017).
44. Zhou, Y. *et al.* Indoline Catalyzed Acylhydrazone/Oxime Condensation under Neutral Aqueous Conditions. *Org. Lett.* **22**, 6035–6040 (2020).
45. Trausel, F., Fan, B., van Rossum, S. A. P., van Esch, J. H. & Eelkema, R. Aniline Catalysed Hydrazone Formation Reactions Show a Large Variation in Reaction Rates and Catalytic Effects. *Adv. Synth. Catal.* **360**, 2571–2576 (2018).
46. Koo, B. S., Lee, C. K. & Lee, K. J. Oxidation of benzyl alcohols with Oxone® and sodium bromide. *Synth. Commun.* **32**, 2115–2123 (2002).
47. Murahashi, S.-I. & Shiota, T. Short-Step Synthesis of Amino Acids And N-Hydroxyamino Acids From Amines. *Tetrahedron Lett.* **28**, 5241–5244 (1987).
48. Gella, C. *et al.* A metal-free general procedure for oxidation of secondary amines to nitrones. *J. Org. Chem.* **74**, 6365–6367 (2009).
49. Denmark, S. E., Forbes, D. C., Hays, D. S., DePue, J. S. & Wilde, R. G. Catalytic Epoxidation of Alkenes with Oxone. *J. Org. Chem.* **60**, 1391–1407 (1995).
50. White, D. H., Noble, A., Booker-Milburn, K. I. & Aggarwal, V. K. Diastereoselective Photoredox-Catalyzed [3 + 2] Cycloadditions of N-Sulfonyl Cyclopropylamines with Electron-Deficient Olefins. *Org. Lett.* **23**, 3038–3042 (2021).
51. Truce, W. E. & Roberts, F. E. A Convenient Synthesis of Aromatic and Aliphatic Sodium Sulfinates. *J. Org. Chem.* **28**, 593–594 (1963).
52. Murahashi, S.-I., Imada, Y. & Ohtake, H. Tungstate-Catalyzed Decarboxylative Oxidation of N-Alkyl-alpha-amino Acids: An Efficient Method for Regioselective Synthesis of Nitrones. *J. Org. Chem.* **59**, 6170–6172 (1994).

Chapter 6. Conclusion and Outlook

Self-assembly is an important aspect of life—all natural organisms have this attribute in their organization or activity. Not long ago, the question of self-assembly within a CRN arose, and the notion of dynamic self-assembly appeared.[1] It became important because, at that time, the more common way to see self-assembly was as the organized state of matter that exists at equilibrium with minimum Gibbs free energy and no entropy production. The realization that self-assembly can occur and be sustained under (chemical/light) energy flux, which is then dissipated along with the increase of entropy, attracted the attention of researchers in the field of dynamic processes.[2] It became clear that detailed studies in close collaboration between theoreticians and experimentalists are needed to rationalize how the systems built from dissipative self-assembly within chemical reaction networks (CRNs) function. Something that Nature does very well (after billions of years of evolution), we have been struggling to understand on a level that would move us closer to the rational creation and mindful use of complex systems' features, with the perspective of understanding what life is and how it works.[3]

From the first interest in complex phenomena like oscillations until the first semi-artificial fueled systems, it has been several decades. After a moonshot evolution of nonlinear dynamics theory, only in recent years have we witnessed experimental systems that have started advancing in the same direction toward complexity.

In **Chapter 1**, we discussed self-assembly in the context of CRNs. Studies on fueled supramolecular polymerization are relatively sparse. In our recent review,[4] we presented three models of increasing complexity: (i) the transient systems where (de)activation reactions do not influence the self-assembly of activated monomer and are not affected by it (i.e., uncoupled), (ii) transient systems with (de)activation reactions coupled to isodesmic supramolecular polymerization, (iii) oscillations in fueled coupled cooperative supramolecular polymerization.

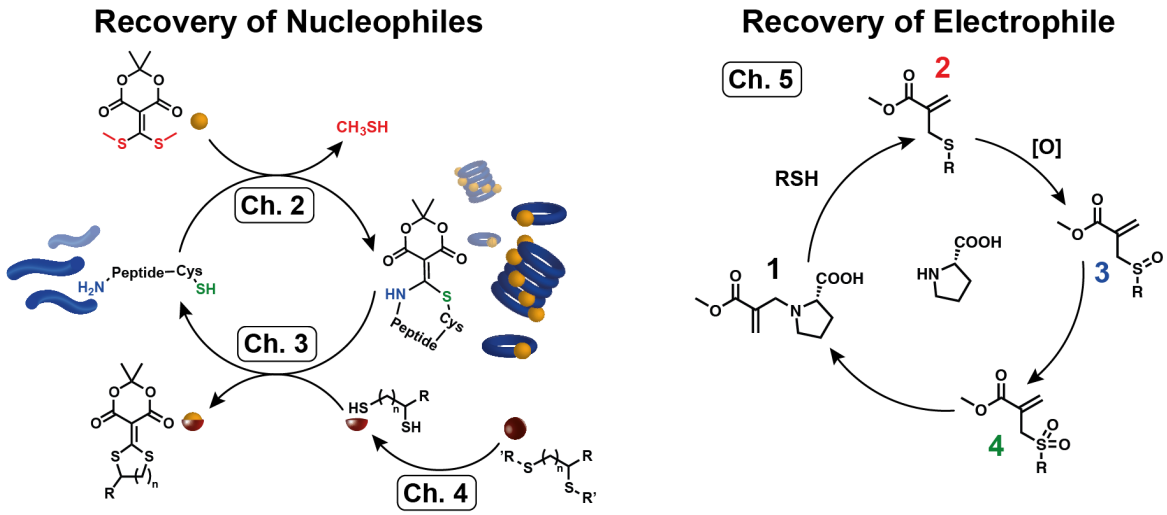


Figure 6.1 | Overview of the discussed systems and their parts. Chapters 2–4 described the step-by-step development of the reaction cycle with fueled self-assembly of CPNs. The recovery of peptide nucleophiles is demonstrated throughout the cycle. Chapter 5 shows oxidation-driven CRN with the recovery of an electrophile.

From this analysis, it becomes evident that self-assembly cannot be neglected in CRNs as it can affect the rates of all processes by speeding them up or slowing them down. Moreover, self-assembly properties should be treated as tools to design more complex artificial systems that can be compared to those in nature. This also means that there is a need for a detailed characterization methodology of such systems. We are still at the early stage of the development of this field, and many more exciting things are yet to come.

The goal of this thesis was to explore different features of CRNs (Figure 6.1, the left scheme). In **Chapters 2–4**, the transient self-assembly of cyclic peptides has been studied. In **Chapter 2**, we review the cyclic peptides' literature and identify the lack of studies on cyclic peptide decyclization and disassembly. We use the previously described Michael addition chemistry for peptides' (de)cyclization, where amine- and thiol-containing molecules react with Meldrum's acid derivative (MAD). [5] The advantage of this functionalization is that it is reversible, as MAD can be removed by reacting with stronger nucleophiles. The original amine and thiol are recovered in this case. The cyclization kinetics of two peptides of different lengths, **P1** and **P2**, with MAD is compared using LC-HRMS, UV-Vis, and CD. The cyclization model with rate expressions as ODEs is developed to fit the kinetics data. We found that a longer peptide (**P1**) has a lower rate of cyclization compared to that of a shorter peptide (**P2**), and the latter also produces less oligomeric side products. The mechanism of supramolecular polymerization is studied and is found to surprisingly be isodesmic.

In **Chapter 3**, we look in detail into the decyclization process and discuss why cyclization and decyclization reactions are difficult to combine into a reaction cycle even though, according to the kinetics of these reactions, transient activation should be possible for this system. We also search for different decyclization agents (DA) and determine the most suitable ones for our system. Although we find several good DAs, mainly dithiols like DTT, the reaction cycle is still not possible with the cyclization and decyclization reactions combined because of the cross-reactivity of MAD and DA.

In **Chapter 4**, we develop a way of a delayed release of DA into the reaction mixture by introducing a pre-fuel (an inert form of DA). We test different pre-fuels and reactions by which they can release DA and find one possible solution: disulfides as a pre-fuel can be reduced by a non-canonical reduction agent. With this result, we test the efficiency of the pre-fuel reaction to release DA within the peptide-MAD system and achieve transient formation and self-assembly of cyclic peptides.

In **Chapter 5**, we explore another CRN without coupling to self-assembly but with more interesting product speciation (Figure 6.1, the right scheme)—several transient compounds and a waste that can be potentially reused. Again, Michael acceptors are used, like in the previous chapters, but in this CRN, the goal is to recover the electrophile (Michael acceptor), rendering nucleophiles as waste. We found that sulfide oxidation, a critical reaction driving the subsequent recovery of the starting compound, can give a degree of control over the appearing species depending on the oxidant used. With the careful study of each of the steps in the CRN, we devise the Ultimate scheme that contains between 10 and 20 interconverting species. Further studies of this system will be done in CSTR conditions.

Finally, after studying conjugate addition chemistry from the position of recycling both nucleophiles and electrophiles, we propose a CRN that combines the advantages of both systems studied in this thesis. This new CRN (Figure 6.2) would permit the reversible

cyclization of peptides that do not contain cysteine, only lysine, or any other primary amine in the backbone, making it a versatile addition to the reactions with MAD proposed by Anslyn and co-workers.[6] The peptide A with a free N-terminus and amine in the backbone (e.g. lysine) in the presence of MAD, a mild oxidant, the disulfide pre-fuel, and a slow reductant would quickly react with MAD forming an adduct B (Figure 6.2). Since the second amine of the peptide cannot react with the MAD adduct due to its low nucleophilicity, oxidation will aid the second substitution by converting the sulfide of MAD B to its more electrophilic sulfoxide C. Now, the second amine will react with the MAD adduct C yielding cyclic peptide D. In parallel, a slow release of dithiol (DA) takes place so that D can be converted back to A.

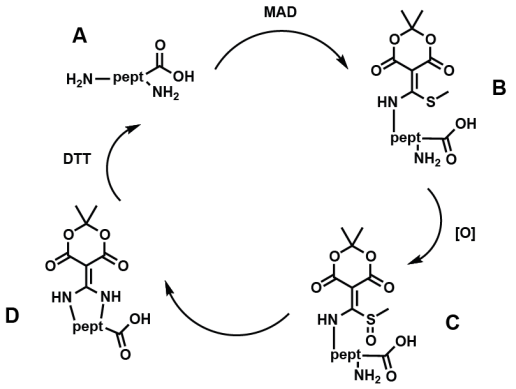


Figure 6.2 | A new CRN is proposed. It combines features of systems discussed in the thesis and allows reversibly forming cyclic peptides that do not contain cysteine.

Overall, the presented studies are essential in the context of programmable and biocompatible CRNs that can easily be interfaced with other biological processes and materials. The gained insights can be used in designing novel and improved (bio)materials with unprecedented functionalities, such as self-sustained tissues capable of their own (limited) metabolism and fully compatible with human/animal organisms, synthetic life systems that can interact with the natural systems, or efficient waste-less nano-chemical stations that can locally synthesize drugs in the damaged organs and 'fix' bugs in the malfunctioning cells. Another desirable long-term perspective of our studies is to gain such a degree of control which may allow one to steer the complex systems towards anticipated structures and functions out of almost infinite ones that are possible out of equilibrium.

References

1. Fialkowski, M. *et al.* Principles and Implementations of Dissipative (Dynamic) Self-Assembly. *J. Phys. Chem. B* **110**, 2482–2496 (2006).
2. Sorrenti, A., Leira-Iglesias, J., Markvoort, A. J., Greef, T. F. A. de & Hermans, T. M. Non-equilibrium supramolecular polymerization. *Chem. Soc. Rev.* **46**, 5476–5490 (2017).
3. Das, K., Gabrielli, L. & Prins, L. J. Chemically Fueled Self-Assembly in Biology and Chemistry. *Angew. Chem.* **133**, 20280–20303 (2021).
4. Sharko, A., Livitz, D., Piccoli, S. D., Bishop, K. & Hermans, T. Insights into chemically-fueled supramolecular polymers. (2022) doi:10.26434/chemrxiv-2022-jd607.
5. Diehl, K. L. *et al.* Click and chemically triggered declick reactions through reversible amine and thiol coupling via a conjugate acceptor. *Nat. Chem.* **8**, 968–973 (2016).
6. Johnson, A. M. & Anslyn, E. V. Reversible Macrocyclization of Peptides with a Conjugate Acceptor. *Org. Lett.* **19**, 1654–1657 (2017).

Acknowledgments

Dissertations highlight one person's contribution, but they are rarely the work of one. This thesis is not an exception. It is powered by many people's efforts and resources and would never become possible without them.
I thank Thomas Hermans, who inspired, supervised, and pushed this research. He gathered a group of exceptional people with funky, controversial, and fundamentally exciting science at its core. When Thomas allowed me to join his group, it became a genuinely transforming challenge with four years of precious personal and professional lessons. His confidence, directness, openness, and never-ending optimism gave me a great example (and I hope my future written works will be more positive ☺), especially when everything around is on fire.
I am grateful for his constant support, the chance to be on his team, and the great experience it turned out to be in the end.

I would like to thank André Estevez-Torres, Nathalie Katsonis, and Peter Faller for accepting to be part of my jury and for evaluating this thesis with their expertise.

I thank all current and former members of Hermans Lab:
Georges Formon and Michi Schicho, who were first my lab mentors during my master's internship, then advisors and allies in suffering later during the Ph.D., also, Michi for all shared sushi and mental support; Róża Szweda for incredible support, compassion, an example of ambition, power, and independence; Nishant Singh for giving harsh survival lessons and introducing me to climbing; Álvaro López-Acosta for being … amazing; Serena de Piccoli for a great collaboration on the review and all fun we had along the way; Alessandra Tassoni for her strength, that inspired and motivated me, for her guidance in all organizational stuff; Jennifer Ciesielski for professional reassurance and ban of the azide-based pre-fuel idea (which prevented some severe lab accidents, I imagine); Jorge Santos Valera González for fruitful discussions and withstanding my tears of stress and anger; Shana Shirin Valapra for a great boost of her part of the project; Marten Leendert Ploeger for advices on kinetics and sharing his attitude to our profession, which, I think, could be the healthiest possible; Tianheng Xu and Chuangyue Guo for the tea exchange session; Oleksandr Shyshov for Sundays biking trips; Shamkhal Baybekov and Mohammad Morsali for supporting the tea-therapy-crew; Lucas Giacchetti for his help with French and spreading atmosphere of calmness and zen; Caroline Thébault, Thomas Biellmann, and Vincent Marichez for all their help with translation, understanding start-up and its management; Bruno Lainer for occasional polish-soviet-random jokes; Shuntaro Amano, Guillermo Monreal Santiago, Ana-Mari Fuentes Capparós, Florencia Sacarelli, Ramarani Sethy, Prabhu Dhasaiyan, Chunfeng Chen, Takuji Adachi, Arvind Dev, Mohammadreza Zolala, Daniel Iglesias Asperilla for all scientific and personal support, involvement, and collaboration.

I thank Cyril Antheaume, Jean-Louis Schmitt, and Hava Aksoy for assistance with HRMS measurements and for trusting me with the extensive use of this instrument on the weekends and nights.
J'aimerais remercier le personnel permanent de l'ISIS, de l'ILB, et de l'ED222, administratif et logistique de créer des conditions de travail optimales et harmonieuses. Je remercie Annia Bertrand et Soumia Hnini pour toute l'aide qu'elles m'ont apportée en matière de bureaucratie, d'organisation administrative et de survie aux règles de l'université.

I thank Rienk Eelkema for his guidance, help, and empathy; Benjamin Klemm, Irene Piergentilli, Reece Lewis, and all members of the Eelkema group for their advice and hospitality at TU Delft♡. Thanks to Hendrik Hubbe, Ardeshir Roshanasan, and Georgy Filonenko for fun, deep discussions, and Greek/snack/coffee breaks.

Je remercie Pierre et sa grande famille, mes voisins, pour le soutien mental, le temps passé ensemble et les dîners communs. Спасибо Оле Машкиной, Алесе и Алине за поддержку, сырники, беседы за чаем, походы в бассейн и Дайну.

Спасибо Лизе за периодическое выгуливание меня, сладкое, разговоры о жизни и консультации с питоном. Спасибо Жанне за созвоны, доверие и обмены музыкой. Спасибо Ане за совместные обеды, взаимные жалобы, соционику и кофе.

Спасибо Дмитриевне, которая на моей стороне с начала начал. Она в курсе специфики. Благодарю свою семью (в т. ч. всех кошек), и особенно бабушку Надю, которая является для меня примером одной из самых сильных женщин, которых я знаю.

Благодарю папу за поддержку, участие и присутствие, которые дают мне опору. Твои советы, идеи, критика и подход к жизни вдохновляют и направляют меня на моём пути. Спасибо, что ты у меня есть.
Математику обязательно продолжаем!:)

Ich danke Gerti und Heinz für die freundliche Aufnahme in ihr Zuhause, für ihre Unterstützung und ihr Verständnis, für das Konzept einer glücklichen Familie, für Kefir, unbegrenzte Badezeit und Geduld mit meinem Deutsch.

I thank Benni, my Sonnenlicht in science and life, for sharing with me the ups and downs, for everything between the supermarket in Rehovot and these last months, which were particularly tough and beautiful. I wouldn't make it without you.

Anastasiia SHARKO

Controlling pathways and kinetics in fueled Chemical Reaction Networks

Résumé

Cette thèse vise à étudier les réseaux de réactions chimiques artificielles, les moyens de les contrôler et les facteurs qui influencent leur évolution.

Tout d'abord, nous montrons un cycle de réaction chimique qui ligature de manière réversible les amines et les thiols de fragments peptidiques, ce qui conduit à des peptides cycliques, suivis par la formation de nanotubes pendant un temps prédéterminé. À la fin, les nanotubes peptidiques sont ramenés aux fragments peptidiques d'origine. Notre travail montre une voie vers des peptides cycliques transitoires hors équilibre, qui pourraient être utiles dans les biomatériaux ou dans des réseaux de réaction plus complexes.

Deuxièmement, nous étudions un réseau de réaction plus large d'espèces interconvergentes non assemblées. Nous montrons qu'en utilisant la chimie de l'oxydation, les accepteurs de Michael (AM) à fonction sulfure peuvent être transformés en AM à fonction sulfoxyde et sulfone, qui récupèrent alors une amine-AM initiale dans le processus. En fonction de l'oxydant employé, ce réseau de réaction peut se dérouler selon trois voies distinctes. Nous envisageons un large éventail d'applications pour cette nouvelle chimie, de la fonctionnalisation des protéines aux systèmes de libération des cargaisons.

Mots clés : réseaux de réactions chimiques, peptides cycliques, auto-assemblage, chimie des systèmes, accepteurs de Michael.

Résumé en anglais

This thesis aims to study artificial chemical reaction networks, the ways to control them, and the factors that influence their evolution.

First, we show a chemically fueled reaction cycle that reversibly ligates amines and thiols of peptide fragments, which leads to cyclic peptides, followed by the formation of nanotubes for a predetermined time. In the end, the peptide nanotubes are reverted to the original peptide fragments. Our work shows a pathway toward out-of-equilibrium transient cyclic peptides, which could be of use in biomaterials or in more complex reaction networks.

Second, we are studying a broader reaction network of interconverting non-assembling species. We show that by employing oxidation chemistry, sulfide-functionalized Michael acceptors (MAs) can be transformed into sulfoxide- and sulfone-functionalized MAs, which then recover an initial amine-MA in the process. Depending on the employed oxidant, this reaction network can proceed via three distinct pathways. We envision a wide range of applications for this novel chemistry from protein functionalization to cargo release systems.

Keywords: chemical reaction networks, cyclic peptides, self-assembly, systems chemistry, Michael acceptors