



Épigénétique et méthylation de l'ADN : développement d'outils pour la compréhension du mécanisme de méthylation de l'ADN impliquant UHRF1

Nicolas Barthes

► To cite this version:

Nicolas Barthes. Épigénétique et méthylation de l'ADN : développement d'outils pour la compréhension du mécanisme de méthylation de l'ADN impliquant UHRF1. Autre. Université Nice Sophia Antipolis, 2015. Français. NNT : 2015NICE4124 . tel-01430507

HAL Id: tel-01430507

<https://theses.hal.science/tel-01430507>

Submitted on 10 Jan 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**UNIVERSITÉ NICE SOPHIA ANTIPOLIS - UFR SCIENCES
ECOLE DOCTORALE DE SCIENCES FONDAMENTALES ET
APPLIQUÉES**

THÈSE

POUR OBTENIR LE TITRE DE :

DOCTEUR EN SCIENCES

DE L'UNIVERSITÉ NICE SOPHIA ANTIPOLIS

DISCIPLINE : CHIMIE

PRÉSENTÉE ET SOUTENUE PAR :

Nicolas P. F. BARTHES

**Epigénétique et méthylation de l'ADN :
Développement d'outils pour la compréhension du
mécanisme de méthylation de l'ADN impliquant UHRF1.**

THESE DIRIGÉE PAR LE Pr. Alain BURGER ET Co-ENCADREE PAR LE
Dr. Benoît Y. MICHEL

SOUTENUE LE 18 DECEMBRE 2015

Jury :

Pr. Mélanie ETHEVE-QUELQUEJEU	Professeur, Université Paris Descartes	Rapporteur
Dr. Jean-Jacques VASSEUR	Directeur de Recherche CNRS, Université de Montpellier II	Rapporteur
Pr. Yves MELY	Professeur, Université de Strasbourg	Examinateur
Dr. Benoît Y. MICHEL	Maître de conférences, Université Nice Sophia Antipolis	Co-encadrant de thèse
Pr. Alain BURGER	Professeur, Université Nice Sophia Antipolis	Directeur de thèse

“ *On ne fait jamais attention à ce qui a été fait : on ne voit que ce qu'il reste à faire.* **”** - Marie Curie

Remerciements

Le travail de thèse présenté dans ce manuscrit a été réalisé sur une période de trois ans au sein de l’Institut de Chimie de Nice (ICN), dirigé par le Dr. Elisabet Duñach et plus particulièrement dans l’équipe Molécules Bioactives (MB), administrée par le Dr. Rachid Benhida.

Je tiens ensuite à remercier mon directeur de thèse, le Pr. Alain Burger et mon co-encadrant le Dr. Benoît Y. Michel pour m’avoir accueilli et encadré au sein de leur groupe de recherche (Groupe Sondes Fluorescentes - GSF). Je tiens à leur exprimer toute ma reconnaissance pour leurs conseils, leur implication dans mes travaux de recherche, la confiance qu’ils m’ont accordée ainsi que pour l’autonomie scientifique qu’ils m’ont permis d’acquérir. Je remercie aussi le Fond pour la Recherche Médicale (FRM) pour m’avoir attribué une bourse de thèse durant ces trois années et pour sa contribution aux frais de fonctionnement, ainsi que l’Agence Nationale de la Recherche (ANR) pour avoir également contribué au financement de mes travaux de recherche.

J’exprime également toute ma gratitude à l’ensemble des membres du jury : le Pr. Mélanie Etheve-Quelquejeu, le Dr. Jean-Jacques Vasseur et le Pr. Yves Mély pour le temps qu’ils ont consacré à examiner et juger ce travail.

Je remercie l’ensemble des étudiants ayant contribué à ce projet de recherche en tant que technicien, stagiaires, doctorants ou post-doctorants : Dr. Krishna Gavvala (P-D), Marianna Sholokh (D), Vasyl Kilin (D), Dominique Bonhomme, Dr. Dmytro Dziuba (T), Dr. Iuliia Karpenko (M2) et Jérémy Auffret (M2).

Je n’oublie pas non plus nos collaborateurs à Strasbourg, Kiev, Nice, Valbonne et Novosibirsk :

- L’équipe du Pr. Yves Mély, pour les travaux effectués sur la méthylation de l’ADN et la rétro-transcription du VIH et pour leur contribution à la caractérisation des oligonucléotides fluorescents,
- Le Dr. Andrey S. Klymchenko et le Pr. Alexander P. Demchenko pour leurs expertises sur la chimie et la photophysique des 3-hydroxychromones lors de l’écriture des publications,
- Le Dr. Nadine Martinet pour le soutien qu’elle m’a apporté au cours de ces trois années et son expertise lors de l’écriture de ce manuscrit,
- L’équipe de spectrométrie de masse de l’IPMC (Anne Sophie Dabert-Gay et Delphine Debayle) pour la caractérisation spectroscopique des oligonucléotides,
- L’équipe du Pr. Olga S. Fedorova pour sa contribution lors de l’études de l’excision de base endommagée.

REMERCIEMENTS

Mes remerciements, pour leur bonne humeur et leur gentillesse, vont également aux autres membres du laboratoire niçois que j'ai côtoyés au cours de ces trois années et plus particulièrement : Anh, Caterina, Cécile, Coralie, Hella, Janah, Janwa, Nadia, Pascal, ainsi que les étudiants qui ne travaillent plus au laboratoire : Adeline, Camille, Hamid, Ilhem, Julien, Marie-Angélique et Mohsine. Vous avez tous contribué à rendre cette expérience encore plus agréable.

Enfin, merci à ma maman et mon papa pour votre amour et pour m'avoir encouragé et soutenu tout au long de cette période. Mamie j'espère que tu es fière de moi de là-haut, tu nous as quitté beaucoup trop tôt mais j'espère que tu partages quand même ce moment avec moi. Je te remercie également ma chérie pour ton soutien et tes encouragements. Merci encore à vous tous pour avoir toujours répondu présent quand j'en avais besoin.

Je prie tous ceux que j'oublie de me pardonner mais vous savez que vous avez évidemment votre place ici ...

Ces travaux de thèse ont été effectués à :

**Université Nice Sophia Antipolis - Faculté des Sciences
Institut de Chimie de Nice (ICN) - CNRS UMR 7272
28 Avenue Valrose - Parc Valrose
06100 - Nice Cedex 2**

TABLE DES MATIERES

Remerciements	3
Liste des Abréviations	11
Introduction Générale	15
CHAPITRE 1. Etat de l'art	19
I. Méthylation de l'ADN	19
A. Génétique & Epigénétique	19
1. Rappels sur les acides nucléiques.....	19
i. Constitution des acides nucléiques	19
ii. Structure de l'ADN.....	21
2. Le génome	23
i. Rappels	23
ii. De l'ADN aux protéines	24
iii. Génome et division cellulaire.....	26
3. Les marques épigénétiques	29
i. Modification des Histones / Modification de la chromatine.....	30
ii. Méthylation de l'ADN	31
iii. Autres modifications de l'ADN	34
iv. Les micro-ARNs	35
B. Rôle et mécanisme de la méthylation de l'ADN.....	37
1. Développement et Méthylation de l'ADN.....	37
i. Division Cellulaire	37
ii. Empreinte génomique parentale	38
ii. Influence de l'environnement sur le développement	40
2. Transmissions de la méthylation de l'ADN et couple DNMT1 / UHRF1(SRA).....	40
i. Les DNMTs	40
ii. DNMT1	41
iii. UHRF1 / SRA	43
C. Implication dans le cancer	46
1. Méthylation de l'ADN et cancers	46
2. DNMT1 / UHRF1 et cancers	48
II. Fluorescence	50
A. Rappels	50
1. Transitions électroniques	51
2. Diagramme de Jablonski	52
3. Avantages de la fluorescence	53
4. Facteurs jouant sur le signal de fluorescence	54
5. Grandeurs caractéristiques de l'émission de fluorescence	55
i. Déplacement de Stokes	56

TABLE DES MATIERES

ii. Rendement Quantique	56
iii. Brillance et Intensité de Fluorescence.....	56
iv. Durée de vie	57
6. Diversité des fluorophores.....	57
B. Sondes Fluorescentes pour le développement de capteurs.....	58
1. Paramètres exploitables.....	59
i. Mesure d'intensité à une ou deux longueurs d'onde	59
i) A une longueur d'onde (monobande).....	59
ii) A deux longueurs d'ondes (monobande ou duale)	60
a. Marquage Simple.....	60
b. Double Marquage	62
ii. Anisotropie	64
iii. Temps de vie	65
2. Applications à la chimie et à la biologie.....	67
i. Sondes « pH-sensibles ».....	67
ii. Sondes à ions.....	68
iii. Interaction et repliement de macromolécules.....	69
iv. Imagerie cellulaire	70
III. Les 3-hydroxychromones	71
A. Synthèse	71
1. Synthèse d'Auwers.....	71
2. Synthèse d'Allan-Robinson	72
3. Synthèse d'Algar-Flynn-Oyamado (AFO).....	73
4. Autres voies de synthèse	73
B. Caractéristiques Photophysiques	74
1. Présentation	74
i. ESIPT	74
ii. Les 3-hydroxychromones (3-HC)	76
2. Applications	80
i. Marquage des protéines / peptides	81
ii. Marquage des phospholipides / micelles	82
iii. Marquage de l'ADN	83
IV. Oligonucléotides	84
A. Synthèse des ODNs	84
1. Synthèse chimique moderne des ODNs	86
2. Fonctionnalisation des ODNs	88
i. Approche pré-synthétique	89
ii. Approche post-synthétique	89
B. Marquages Fluorescents de l'ADN	90
1. Diversité des Sondes	91

TABLE DES MATIERES

i. Marquage non covalent	91
i) Intercalants de l'ADN.....	91
ii) Interaction avec le petit sillon de l'ADN	92
ii. Marquage covalent	93
i) Bras espaceurs longs.....	93
ii) Polycycles (hétéro)aromatiques connectés au sucre	94
iii) Bras espaceurs court.....	99
2. Applications	101
i. Séquençage de l'ADN (Méthode Sanger)	101
ii. Détection de polymorphisme mononucléosidique	102
iii. Microenvironnement.....	103
CHAPITRE 2. Résultats et Discussions	107
Présentation du programme de recherche.....	107
A. Première stratégie de marquage (intra-duplexe)	113
PUBLICATION 1	113
B. Seconde stratégie de marquage (grand sillon).....	129
PUBLICATION 2	129
PUBLICATION 3.....	143
PUBLICATION 4	155
PUBLICATION 5.....	167
Conclusion Générale et Perspectives	177
CHAPITRE 3. Partie Expérimentale	183
SI PUBLICATION 1	185
SI PUBLICATION 2.....	189
SI PUBLICATION 3.....	199
SI PUBLICATION 4	221
SI PUBLICATION 5.....	237
ANNEXES	249
ANNEXES S1	251
ANNEXES S2	252
ANNEXES S3	255
PUBLICATION 6	255
ANNEXES S4	269
TAVAUX DE VULGARISATION	269
REFERENCES	279

Liste des Abréviations

2-AP – 2-aminopurine	BPT – <i>Back Proton Transfer</i>
3-HC – 3-hydroxychromone	BRCA1 – <i>Breast Cancer 1</i>
3-HF – 3-hydroxyflavone	BTT – 5-(Benzylthio)-1 <i>H</i> -tetrazole
3-HQ – 3-hydroxyquinolone	C – Cytosine
3MI – <i>3-methylsiloxanthopterin</i>	Cbz – Groupement Carboxybenzyl
5mC – <i>5-methylcytidine</i>	CD – <i>Circular Dichroism</i>
5hmC – <i>5-hydroxymethylcytidine</i>	CDKN2a – <i>Cyclin-Dependent Kinase inhibitor 2A</i>
5fC – <i>5-formylcytidine</i>	CEP – 2-cyanoéthyl <i>N,N'</i> -diisopropylphosphoramidite
5caC – <i>5-carboxylic acid cytidine</i>	CH ₃ CN – Acetonitrile
6mA – 6-méthyladénosine	cm – Centimètre
6-MAP – 4-amino-6-methyl-8-(2-deoxy-β-D-ribofuranosyl)-7(8H)-pteridine	CMC – Concentration Micellaire Critique
6MI – <i>6-methylsiloxanthopterin</i>	COSY – <i>COrelation SpectroscopY</i>
8vdA – 8-vinyladénosine	CPG – <i>Controlled Pore Glass</i>
A – Adenine	CpG – Dinucléotide Cytosine-phosphate-Guanosine
Å – Angstrom	Cys – Cystéine
ADN – Acide Désoxyribonucléique	DAC – <i>Dibasic Ammonium Citrate</i>
AFO – Algar-Flynn-Oyamada [réaction]	DAPI – 4',6-diamidino-2-phenylindole
Ago – Protéine argonaute	DAPK – <i>Death-Associated Protein Kinase</i>
ApT – Dinucléotide Adénosine-phosphate-Thymidine	DCM – Dichlorométhane
Ar – Substituant aromatique	ddNTP – 2',3'-didésoxynucléotide triphosphate
Arg – Arginine (= R)	DGCR8 – <i>Di George Critical Region 8</i>
ARN – Acide Ribonucléique	DHU – 5,6-dihydrouridine
ARNm – ARN messager	DIEA – Diisopropyl Ethyl Amine
ARNr – ARN ribosomal	DMAD – <i>Drosophila DNA 6mA demethylase</i>
ARNt – ARN de transfert	DMAP – 4-diméthylaminopyridine
ATP – Adénosine TriPhosphate	DMAP – (4-amino-2,6-dimethyl-8-(2-deoxy-β-D-ribofuranosyl)-7(8H)-pteridine
a.u. – <i>arbitrary units</i>	DMBAN – 4- <i>N,N'</i> -diméthylaminobenzonitrile
BDF – <i>Base-Discriminating Fluorescence</i>	DMF – Diméthylformamide
BE – Bromure d'éthidium	
BER – <i>Base Excision Repair</i>	
Bi – Brillance	
BPA – Bisphénol A	

LISTE DES ABREVIATIONS

DMSO – Diméthylsulfoxide	HPLC – <i>High-Performance Liquid Chromatography</i>
DMT – Groupement 4,4'-diméthoxytrityle	HPTS – <i>8-HydroxyPyrene-1,3,6-TriSulfonic acid</i>
DNMT – <i>DNA Methyltransferase</i>	HRMS – <i>High-Resolution Mass Spectrometry.</i>
dNTP – 2'-désoxynucléotide triphosphate	HSA – <i>Human Serum Albumine</i>
E – Energie	HSQC – <i>Heteronuclear Single Quantum Correlation</i>
EDTA – <i>EthyleneDinitrioloTetraacetic Acid</i>	I – intensité de fluorescence
ESI – Ionisation électrospray	IGF2 – <i>Insulin-like Growth Factor 2</i>
ESICT – <i>Excited State Intramolecular Charger Transfer</i>	IR – InfraRouge
ESIPT – <i>Exited State Intramolecular Proton Transfer</i>	J – Constante de couplage spin-spin
Et – Groupement Ethyle	JmjC – <i>JmjC domain-containing histone demethylase</i>
EtOAc – Acétate d'Ethyle	K – Lysine (Lys)
EtOH – Ethanol	kb – killobase (1 kb = 1 000 pb)
ETT – 5-(Ethylthio)-1 <i>H</i> -tetrazole	kDa – kiloDalton
eV – electronVolt	λ – Longueur d'onde
Excimère – <i>Excited Dimer</i>	(LSD)1 – <i>Lysine-Specific Demethylase</i>
Exciplexe – <i>Excited Complex</i>	LUMO – <i>Lowest Unoccupied Molecular Orbital</i>
FLIM – <i>Fluorescence-Lifetime Imaging Microscopy</i>	Lys – Lysine (= K)
FRET – <i>Fluorescence (Förster) Resonance Energy Transfer</i>	m5k – 5-méthylzébularine
Fur – Groupement Furyle	Mb – Mégabase (1 Mb = 1 000 000 pb)
G – Guanine	m-CPBA – Acide méta-chloroperbenzoïque
GFP – <i>Green Fluorescent Protein</i>	MeOH – Méthanol
GSTP1 – <i>Glutathione S-Transferase P1</i>	Met – Méthionine
H – Histone	METTL4 – <i>METHyITransferase Like 4</i>
HAT – <i>Histone Acetyl Transferase</i>	MGMT – <i>O⁶-MethylGuanine-DNA-MethylTransferase</i>
HCV – <i>Hepatitis C Virus</i>	miARN –microARN
HDAC – <i>Histones Deacetylases</i>	̄ – Nombre d'onde
HIV-1 – <i>Human Immunodeficiency Virus type 1</i>	N – Normalité d'une solution
HM – Hémi-méthylé (ADN)	N – Forme normale des 3-HC
HMBC – <i>Heteronuclear Multiple Bond Correlation</i>	N* – Forme normale excitée des 3-HC
HMT – <i>Histone MethylTransferases</i>	NBS – <i>N-Bromosuccinimide</i>
HOMO – <i>Highest Occupied Molecular Orbital.</i>	NCp7 – <i>HIV-1 NucleoCapsid protein</i>
	NM – Non méthylé (ADN)

LISTE DES ABREVIATIONS

nm – nanomètre	SKOV-3 – Cellules humaines de carcinome ovarien
NMAD-1 – <i>6mA Demethylase</i>	SNI-PEX – <i>Single-Nucleotide Incorporation-PEX</i>
NMR – <i>Nuclear Magnetic Resonance</i>	SNP – <i>Single Nucleotide Polymorphism</i>
nt – Nucleotides	SRA – <i>SET and RING finger-associated domain</i>
ODF – Oligodésoxyfluoroside	τ_S – temps de vie de fluorescence
ODN – Oligodésoxyribonucleotide	T – Thymine
ORN – Oligoribonucleotide	T – Forme tautomère des 3-HC
OVCAR-3 – Cellules humaines de cancer ovarien	T* – Forme tautomère excitée des 3-HC
pb – paire de base	TAR – <i>Trans-Activation Response element</i>
PBS – <i>Primer Binding Site</i>	TDG – <i>Thymine DNA Glycosylase</i>
PCNA – <i>Proliferating Cell Nuclear Antigen</i>	TdT – <i>Terminal deoxynucleotidyl Transferase</i>
PCR – <i>Polymerase Chain Reaction</i>	TET – <i>Ten-Eleven Translocation</i>
PDB – <i>Protein Data Bank</i>	TFIID – <i>Transcription Factor II D</i>
PEX – <i>Primer EXtension</i>	Th – Groupement Thiényle
Ph – Groupement Phényl	THF – Tétrahydrofurane
PHD – <i>Plant Homeo Domain</i>	Thr – Thréonine
Φ_F – Rendement quantique de fluorescence	TICT – <i>Twisted Intramolecular Charge Transfer</i>
ppm – Partie par million	TIPS – Groupement Triisopropylsilyl
Pyr – Pyridine	TLC – <i>Thin Layer Chromatography</i>
R – Arginine (Arg)	T_m – Température de fusion des ODN
RASSF1 – <i>Ras Association (RalGDS/AF-6) Domain Family Member 1</i>	TMS – Groupement Triméthylsilyl
RING – <i>Really Interesting New Gene</i>	TO – Thiazole Orange
RISC – <i>RNA-Induced Silencing Complex</i>	TOF – <i>Time of Flight</i>
RFTD – <i>Replication Foci Targetting Domain</i>	Tol – Groupement Toluoyl
RSA – <i>Rat Serum Albumine</i>	TRBP – <i>Transactivating ARN-Binding Protein</i>
rt – <i>room temperature</i>	Trp – Tryptophane
RX – Diffractions des rayons X	TTD – <i>cryptic Tandem Tudor Domain</i>
s – seconde	U – Uracile
SAM – <i>S-Adenosyl-L-méthionine (= AdoMet)</i>	UBL – <i>UBiquitin-Like</i>
SAH – <i>Adenosyl-L-homocystéine (= AdoHcy)</i>	UHRF1 – <i>Ubiquitin-like, containing PHD and RING Finger domains 1</i>
siARN – <i>small interfering ARN</i>	UV – UltraViolet
SIDA – Syndrome de l'Immunodéficience Acquise	Val – Valine

LISTE DES ABREVIATIONS

vazo® 67 – 2,2'-azobis(2-méthylbutyronitrile)

VIH – Virus de l'Immunodéficience Humaine

X – Halogène

xDNA – *Extended DNA*

yDNA – *Wild DNA*

YO – *Yellow Oxazole*



Introduction Générale

A l'heure actuelle, environ 25 millions de personnes sont atteintes de cancer dans le monde. On estime qu'un homme sur deux, et qu'une femme sur trois, se verront diagnostiqués un cancer avant l'âge de 85 ans et qu'à partir de 2030, la maladie touchera entre 20 et 26 millions de nouveaux individus par an. L'apparition de ces cancers est due à la succession et à l'accumulation de traumatismes causés à l'ADN. Ces altérations touchent plus particulièrement des gènes-clés contrôlant la division, la différenciation ou encore la mort cellulaire (apoptose). Ainsi, les cellules touchées deviennent « immortelles » et acquièrent la capacité de se diviser de manière infinie et anarchique, aboutissant à l'apparition de tumeurs.

Les altérations à l'origine de la carcinogenèse peuvent être liées à notre vieillissement, au contact avec certains agents infectieux (rétrovirus) ou certaines substances polluantes, aux rayons du soleil ainsi qu'à notre comportement, notre style de vie (tabagisme, alcoolisme, alimentation déséquilibrée ou manque d'activité physique). Ces différents facteurs de risques entraînent des modifications dites « épigénétiques » qui altèrent certaines voies de régulation de l'expression des gènes. Plus particulièrement, elles touchent les voies de méthylation de l'ADN, de méthylation et d'acétylation des histones ou encore l'expression des miARNs.

En temps normal, la méthylation de l'ADN garantit une parfaite régulation des phénomènes majeurs tels que le maintien de l'intégrité du génome, l'empreinte génomique parentale, l'expression tissu-spécifique des gènes, l'inactivation du chromosome X et le *silencing* des transposons et des rétrovirus. Cependant, dans le cas de cellules cancéreuses, une hypométhylation globale de l'ADN ainsi qu'une hyperméthylation des

promoteurs des gènes suppresseurs de tumeurs peuvent être observées. Une meilleure compréhension de ce mécanisme permettrait d'aider les chimistes médicinaux dans l'élaboration et la recherche de nouveaux inhibiteurs plus spécifiques que ceux actuellement présents sur le marché (Vidaza® et Dacogen®). *Notons que des dérégulations de la méthylation de l'ADN sont également mises en cause dans le développement de certains troubles neurodégénératifs et de maladies génétiques et métaboliques (Alzheimer, Huntington, ...).*

L'une des cibles privilégiées, impliquées dans ce mécanisme de méthylation de l'ADN, est le couple UHRF1 / DNMT1. L'action conjointe de ces deux protéines permet de réguler le profil de méthylation au cours de la division cellulaire. Cependant, des dérégulations de leur comportement ont été observées dans le cas de nombreuses lignées cancéreuses. Une clarification de leur mode d'interaction pourrait donc fournir la clé d'un traitement plus efficace contre le cancer.

Pour répondre à cette problématique, nous avons donc entrepris le développement d'oligonucléotides (ODN) marqués à l'aide d'un rapporteur fluorescent capable de sonder les changements environnementaux se produisant lors de ces interactions acides nucléiques / protéines. La fluorescence est la technique de choix de part sa grande sensibilité, sa polyvalence et son caractère non destructif. Néanmoins, cette méthode est limitée par la disponibilité de fluorophores capables de sonder avec une grande sensibilité et de manière spécifique et multiparamétrique les changements de conformation se produisant lors de la formation de ce genre de complexes nucléoprotéiques.

Une famille de fluorophores s'est avérée particulièrement prometteuse : celle des 3-hydroxychromones (3-HC). Ces sondes ratiométriques et multiparamétriques présentent une très grande sensibilité aux changements environnementaux grâce à leur capacité à se tautomériser à l'état excité, par transfert de proton intramoléculaire (ESIPT). Leur spectre d'émission présente deux bandes parfaitement résolues dont le rapport varie en fonction des changements environnants.

Ainsi, mes travaux de thèse se sont inscrits dans la lignée des recherches entreprises au sein du Groupe Sondes Fluorescentes depuis plusieurs années et qui consistent à utiliser ces analogues nucléosidiques fluorescents à base de motifs 3-hydroxychromones afin d'aider à l'élucidation de mécanismes nucléoprotéiques majeurs impliqués dans les cancers, le VIH ou la réparation de l'ADN.

Plus particulièrement, mes travaux de thèse s'inscrivent dans la cadre d'une collaboration entre notre équipe et celle dirigée par le Professeur Yves Mély de l'Université de Strasbourg.

Ce manuscrit de thèse est divisé en trois parties :

- Dans un premier chapitre bibliographique (**Etat de l'art**), des généralités sur l'ADN seront présentées, ainsi que sa composition, sa structure, ses fonctions principales et les modifications épigénétiques, plus particulièrement la méthylation de l'ADN, pouvant

affecter l'expression du génome. Ensuite, des rappels sur la spectroscopie de fluorescence et ses avantages (résolution spatiale et temporelle, sensibilité, rapidité de réponse et son côté non destructif) seront détaillés avant de nous intéresser aux motifs 3-hydroxychromones utilisés pour ces études. Enfin, quelques rappels sur les oligonucléotides (ODNs) et les différentes techniques de marquage fluorescent de l'ADN seront exposés.

- Le second chapitre (**Résultats et Discussions**) sera consacré à la présentation des résultats obtenus au cours de ces trois années. Il sera divisé en deux sous-parties, chacune traitant de l'une des deux stratégies de marquage développées.
- Enfin, le dernier chapitre (**Partie Expérimentale**) sera dédiée aux protocoles et aux analyses réalisées sur les différents composés synthétisés.

Note : Les structures de cristallographie aux rayons X ont été réalisées en utilisant le logiciel de modélisation gratuit UCSF Chimera[©], version 1.8.1.



CHAPITRE 1. Etat de l'art

I. Méthylation de l'ADN

A. Génétique & Epigénétique

L'expression du code génétique d'un individu est due à l'enchaînement des bases nucléiques le long de la molécule d'ADN, mais pas seulement. L'état moléculaire de l'ADN, et des entités interagissant avec lui, permet également d'influencer l'expression du génome. Ainsi, en plus du code génétique, dont quelques notions vont être rappelées ici, il existe un code dit « épigénétique » venant jouer ce rôle « *au-delà du génome* ».

Les notions concernant la constitution et le rôle biologique de l'ADN (transcription, traduction, mitose) vont tout d'abord être rappelées, avant de présenter les marques épigénétiques pouvant influencer son expression. Puis, dans une seconde partie, la méthylation de l'ADN sera passée au crible : son influence sur notre développement, le rôle que joue l'environnement sur elle et son mécanisme à l'échelle moléculaire. Enfin, son implication dans la genèse des cancers sera étudiée.

1. Rappels sur les acides nucléiques

i. Constitution des acides nucléiques

Les acides nucléiques sont des macromolécules biologiques, indispensables au développement de toute forme de vie animale, végétale, microbienne ou virale. Il existe

deux types d'assemblage nucléique : l'ADN, pour acide désoxyribonucléique et l'ARN, pour acide ribonucléique [cf. Figure 1 et 3].

- Le premier à pour rôle de stocker l'information génétique nécessaire au développement des différents constituants de la cellules. Il se localise dans le noyau de cette dernière.
- Le second, quant à lui, joue de multiples rôles. Il est, par exemple, impliqué dans le transport de l'information génétique du noyau vers le cytoplasme (ARNm) ; il peut également jouer le rôle de catalyseur à l'instar des enzymes : on parle alors de ribozymes. C'est le cas de l'ARN ribosomal (ARNr) qui, bien qu'associé à des protéines structurales du ribosome, assure le rôle catalytique¹ ; il peut également servir de transporteur : les ARNt vectorisent les acides aminés vers le ribosome au moment de la synthèse protéique (cf. I.A.2.ii).

Les acides nucléiques sont constitués de briques élémentaires appelées nucléosides [cf. Figure 1]. Ces unités monomériques appartiennent à deux familles bien distinctes : les purines d'une part (adenine (A) et guanine (G)) et les pyrimidines d'autre part (cytosine (C), thymine (T), et uracile (U)). La T est spécifique de l'ADN et est remplacée par l'uracile (U) dans l'ARN. Dans le cas de l'ADN, ces nucléosides sont constitués par l'assemblage d'un noyau sucre, le 2'-désoxyribose, sur lequel vient se fixer une des quatre bases, en liant l'azote en 1 des bases pyrimidiques ou 9 des bases puriques à la position anomérique (1') du sucre [cf. Figure 1]. Dans le cas de l'ARN, le noyau sucre mis en jeu est du ribofuranose [cf. Figure 1].

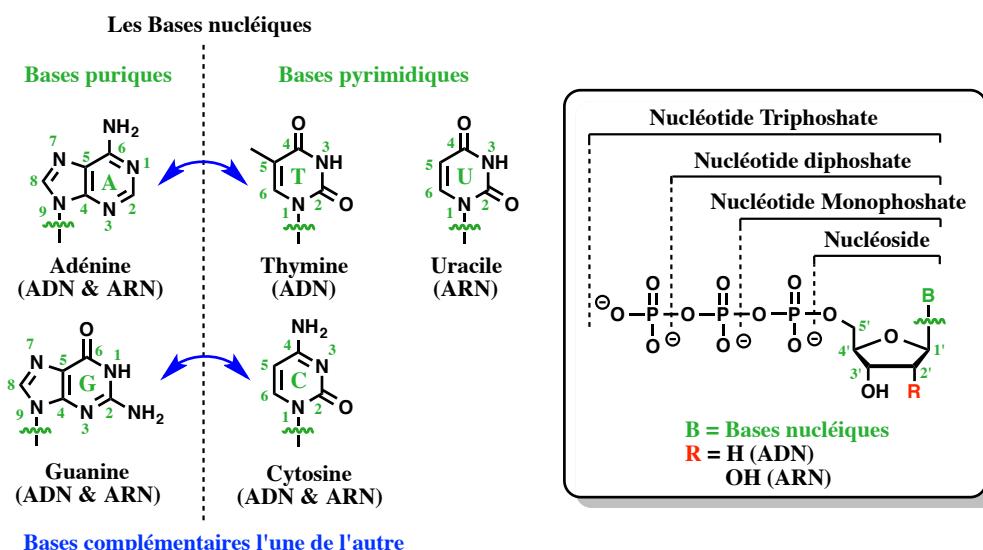


Figure 1. Constitution des bases nucléiques.

D'autres bases, résultant de modifications post-traductionnelles (cf. I.A.3.ii-iii), peuvent également être rencontrées de façon plus anecdotique au sein de l'ADN et de l'ARN. Par exemple, l'ADN présente un grand nombre de dérivés de la C et de la A (5mC, 5hmC, 5fC, 5caC et 6mA), impliqués dans des processus épigénétiques. L'ARN, peut lui aussi présenter des bases méthylées mais également d'autres bases telles que la pseudouridine,² la wytosine³ ou l'inosine, qui sont plus particulièrement rencontrées dans les ARNt [cf. Figure 2].

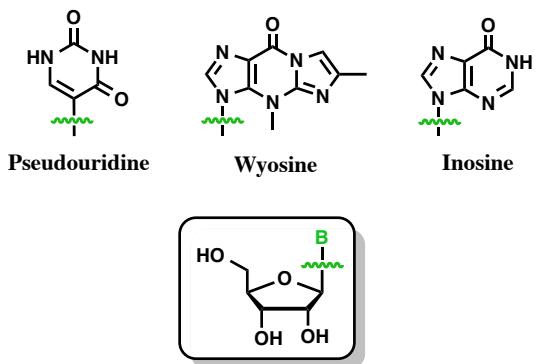


Figure 2. Autres bases nucléiques pouvant être rencontrées.

ii. Structure de l'ADN

L'ADN est un polymère formé par l'assemblage de nucléotides, à savoir des nucléosides présentant des groupes phosphates en 5' [cf. Figure 1]. Ils sont reliés entre eux de manière covalente par des ponts phosphodiesters, entre la position 5' du nucléoside n et la position 3' du nucléoside $n+1$ [cf. Figure 3A], afin de former un brin d'ADN : c'est la structure primaire de l'ADN [cf. Figure 3]. Le brin représenté en Figure 3 est conventionnellement noté d(ACGT). Cette nomenclature vise à nommer les acides nucléiques dans le sens 5' → 3'.⁴

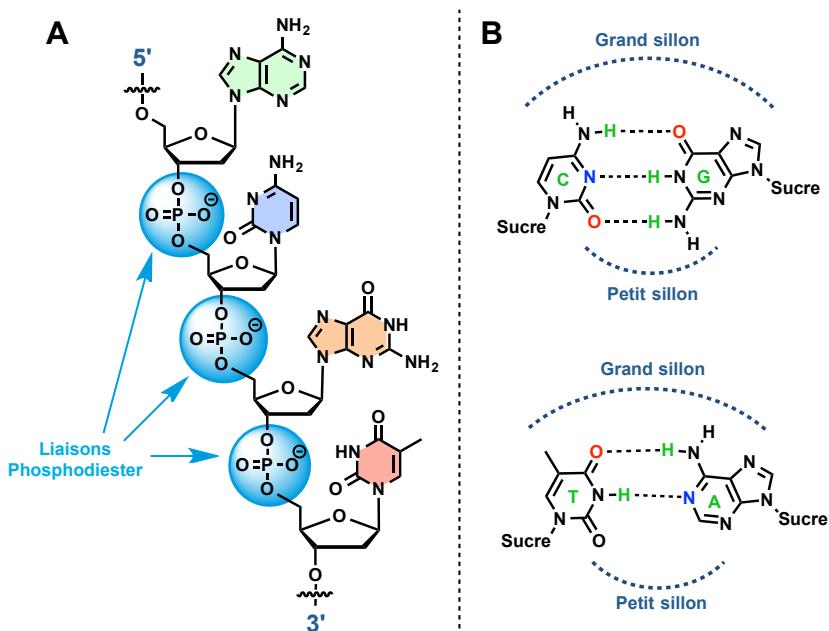


Figure 3. (A) Structure primaire de l'ADN et (B) liaisons Watson-Crick entre les paires de bases C-G & A-T.

Cependant, deux brins d'ADN sont nécessaires car il s'agit d'un système bicaténaire : c'est-à-dire constitué par l'assemblage de deux chaînes d'acides nucléiques, complémentaires l'une de l'autre [cf. Figure 3B]. Ces deux brins complémentaires s'apparentent entre eux de façon antiparallèle grâce à l'établissement de liaisons hydrogènes spécifiques. En effet, deux liaisons hydrogènes sont établies entre les bases nucléiques A & T, contre trois entre les bases C & G : on parle d'interactions de type Watson-Crick [cf. Figure 3B]. Les bases s'empilent ensuite les unes sur les autres - on parle d'empilement

(stacking) - sous la forme d'une hélice, afin de renforcer l'interaction inter-brin. Le double brin d'ADN ainsi formé adopte une structure hélicoïdale 'droite' (sens de rotation vers la droite), telle que l'ont décrit Watson et Crick en 1953 [cf. Figure 4 - forme B].⁵ Dans cette structure, les nucléobases sont orientées vers l'intérieur du système (cœur hydrophobe), et les groupements phosphates vers l'extérieur.

Il n'existe pas qu'une seule structure secondaire hélicoïdale de l'ADN. En effet, en fonction de l'enchaînement des acides nucléiques au sein de la séquence, des conditions de salinité ou du taux d'hydratation environnant, les deux brins d'ADN peuvent s'agencer de plusieurs façons et ainsi donner lieu à différents polymorphes. Citons parmi les plus répandus, ceux présentés dans la Figure 4, à savoir les polymorphes A, B et Z.⁶

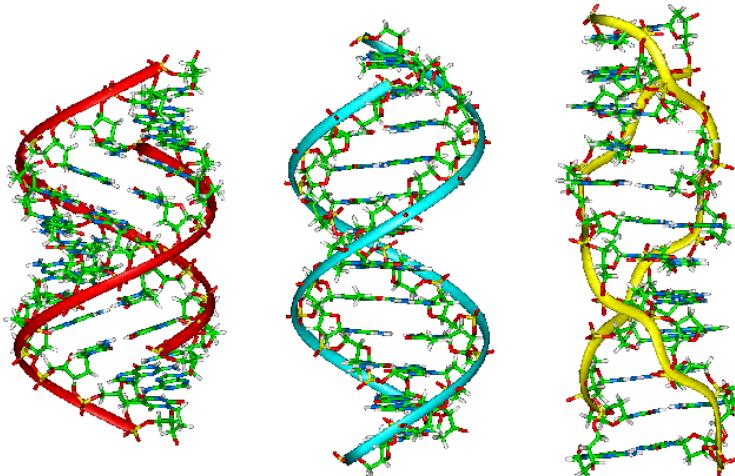


Figure 4. Différentes conformations de l'ADN : ADN A, ADN B et ADN Z.^{6b}

La conformation B [cf. Figure 4, bleu], est la plus couramment rencontrée dans les conditions physiologiques. Elle est caractérisée par un pas d'hélice de 10,5 paires de bases (pb) (35,7 Å). Les nucléosides sont tous de conformations C2' endo / anti [cf. Figure 5] et les bases nucléiques sont situées dans un plan perpendiculaire à l'axe de l'hélice. Enfin, deux sillons profonds, courent le long de la structure hélicoïdale : le grand et le petit sillons. Ils exposent vers l'extérieur du duplexe un profil de liaisons hydrogènes séquence-dépendant, pouvant être reconnu et lu par certaines protéines, comme les facteurs de transcription (cf. I.A.2.ii).⁷

Le polymorphe A [cf. Figure 4, rouge], quant à lui, est observé pour de fortes concentrations en ions Mg²⁺ ou lorsque le taux d'hydratation diminue. Il est caractérisé par une hélicité de 11 pb (28,05 Å) et cette fois-ci, l'ensemble des nucléosides sont de conformations C3' endo / anti [cf. Figure 5]. Comme pour l'ADN B, deux sillons peuvent être observés à sa surface, à ceci près que dans le cas du dérivé A, le grand sillon est plus étroit et profond alors que le petit sillon est plus large et moins profond.⁶

Contrairement aux deux autres polymorphes mentionnés précédemment, la forme Z [cf. Figure 4, jaune], pour zigzag, est structurée en hélice 'gauche'. Elle est observée pour des séquences riches en paires de bases G-C et lorsque la concentration saline est élevée (2,5 M NaCl ou 0,7 M MgCl₂). Cette structure résulte de l'alternance des

conformations *C3'* *endo* / *syn* et *C2'* *endo* / *anti* des bases G et C respectivement [cf. Figure 5].⁶ Il faut cette fois-ci 11,6 pb (42,92 Å) afin d'effectuer un tour complet autour de l'axe principal. Le grand sillon est complètement aplati à la surface et le petit sillon est, quant à lui, étroit et profond.

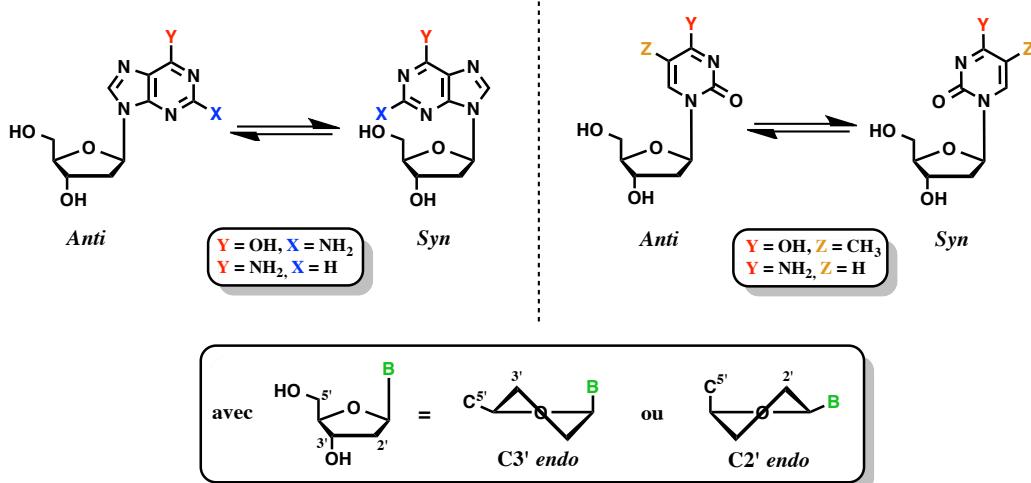


Figure 5. Conformation des nucléosides.

D'autres conformations de l'ADN peuvent également être rencontrées telles que les G-quadruplexes, les i-motifs, les tige-boucles (*hairpin*), les A-motifs ou encore les triplexes, mais ces dernières ne seront pas traitées dans ce manuscrit.⁸

2. Le génome

i. Rappels

Intéressons-nous à présent au rôle biologique de l'ADN. Il représente ce que l'on appelle le génome d'un être vivant, c'est-à-dire, l'ensemble du matériel génétique de ce dernier, qu'il s'agisse de virus, de bactéries, de plantes ou de toutes formes de vie animale. Deux types de génomes, se situant au niveau de deux organelles bien distinctes, peuvent être rencontrés :

- le génome nucléaire, localisé au niveau du noyau de la cellule. Le reste du manuscrit va exclusivement se focaliser sur lui.
- le génome mitochondrial (ou chloroplastique chez les plantes) qui se situe dans les mitochondries (ou les chloroplastes), organelles responsables de la respiration cellulaire et de la production d'énergie (ATP). La présence de ces génomes s'explique par le fait que, d'après la théorie endosymbiotique, les mitochondries et les chloroplastes seraient des bactéries ancestrales, assimilées il y a plusieurs milliards d'années, par d'autres bactéries primitives.^{9, 10}

Le génome nucléaire contient toutes les informations indispensables au développement et à la vie de l'individu et joue un rôle de première importance pour ce dernier. Sa taille varie en fonction des espèces. Les virus possèdent par exemple des génomes « courts » de quelques kb (kilobase) de long alors que chez les eucaryotes, il

peut atteindre plusieurs centaines de milliers de Mb (mégabase). Il est constitué de parties codantes, les gènes, qui sont transcrits en ARNm puis traduits en protéines (cf. I.A.2.ii.) ; ainsi que de parties non-codantes, à hauteur de 80-90 % (introns, promoteurs de gènes, ...).

Le nombre de gènes varie en fonction des espèces, du millier chez les bactéries à quelques dizaines de milliers chez les individus plus développés. L'homme, par exemple, possède entre 20 et 25 000 gènes, présents sur 46 chromosomes [cf. Figure 6].¹¹ Ils sont constitués de trois parties principales : la région promotrice, située en amont du gène (5') et qui initie son expression ; le corps du gène qui contient l'information génétique à exprimer ; et le site de terminaison, situé en aval du gène (3') et qui stoppe son expression.

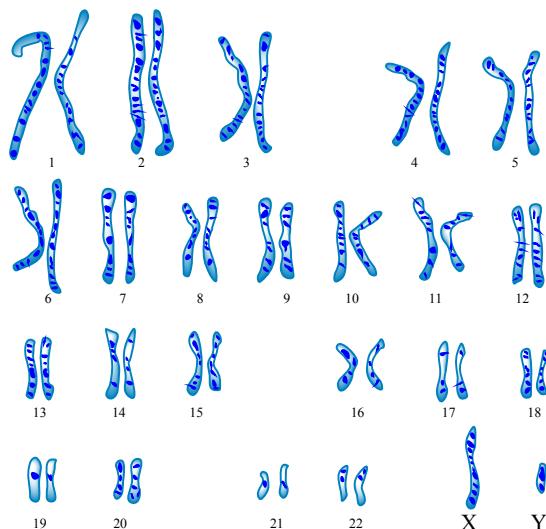


Figure 6. Représentation schématique des 23 paires de chromosomes humains.

Il est intéressant de noter que chez certains virus appelés ribovirus (grippe, rougeole, HCV, ébola, ...), le génome est porté par de l'ARN. Il existe également une sous-catégorie de virus, les rétrovirus (comme le SIDA), qui ont en plus comme particularité de rétro-transcrire cet ARN viral en ADN viral. Ce dernier est ensuite introduit au sein du génome (ADN) de l'organisme hôte afin de l'infecter et d'utiliser son appareillage cellulaire dans le but de se reproduire.

ii. De l'ADN aux protéines

L'ADN est une macromolécule essentielle, point de départ de la biosynthèse du matériel cellulaire, dont font par exemple parti les protéines. Alors que l'ADN se localise exclusivement dans le noyau de la cellule eucaryote, la synthèse des protéines, elle, n'est réalisée que dans son cytoplasme ; et l'ADN ne peut pas s'y déplacer. Afin de permettre la synthèse de ce biomatériel, des copies transitoires et simples brins de l'ADN sont donc produites grâce à « l'ARN polymérase II » [cf. Figure 7A et B]. On les appelle des ARN messagers (ARNm). Ils sont constitués de sous-parties codantes, les introns, et non-codantes, les exons et possèdent exactement les mêmes informations que les gènes desquels ils sont issus. Ils présentent l'avantage de pouvoir être exportés dans le

cytoplasme de la cellule, grâce à leurs petites tailles, et la plupart sont également transcrits en plusieurs exemplaires afin d'accélérer la synthèse protéique.¹²

Afin de générer ces brins d'ARNm, des facteurs de transcription (comme TFIID, TFIIDa et TFIIDb) se positionnent tout d'abord au niveau du promoteur du gène à transcrire. Ce complexe protéique prépare l'ADN à la liaison de l'ARN polymérase II et oriente par la même occasion son positionnement au début du gène. L'enzyme se déplace ensuite le long de la séquence d'ADN en ouvrant localement la double hélice afin de « lire » l'information génétique contenue sur le brin matrice. Au fur et à mesure de son avancée, elle génère le brin d'ARNm correspondant, complémentaire de ce brin matrice, par incorporation des différents nucléotides présents dans le milieu [cf. Figure 7A et B]. Une fois le site de terminaison de la transcription du gène atteint, le complexe ADN / ARNm / ARN polymérase II se dissocie en libérant les trois entités.¹²

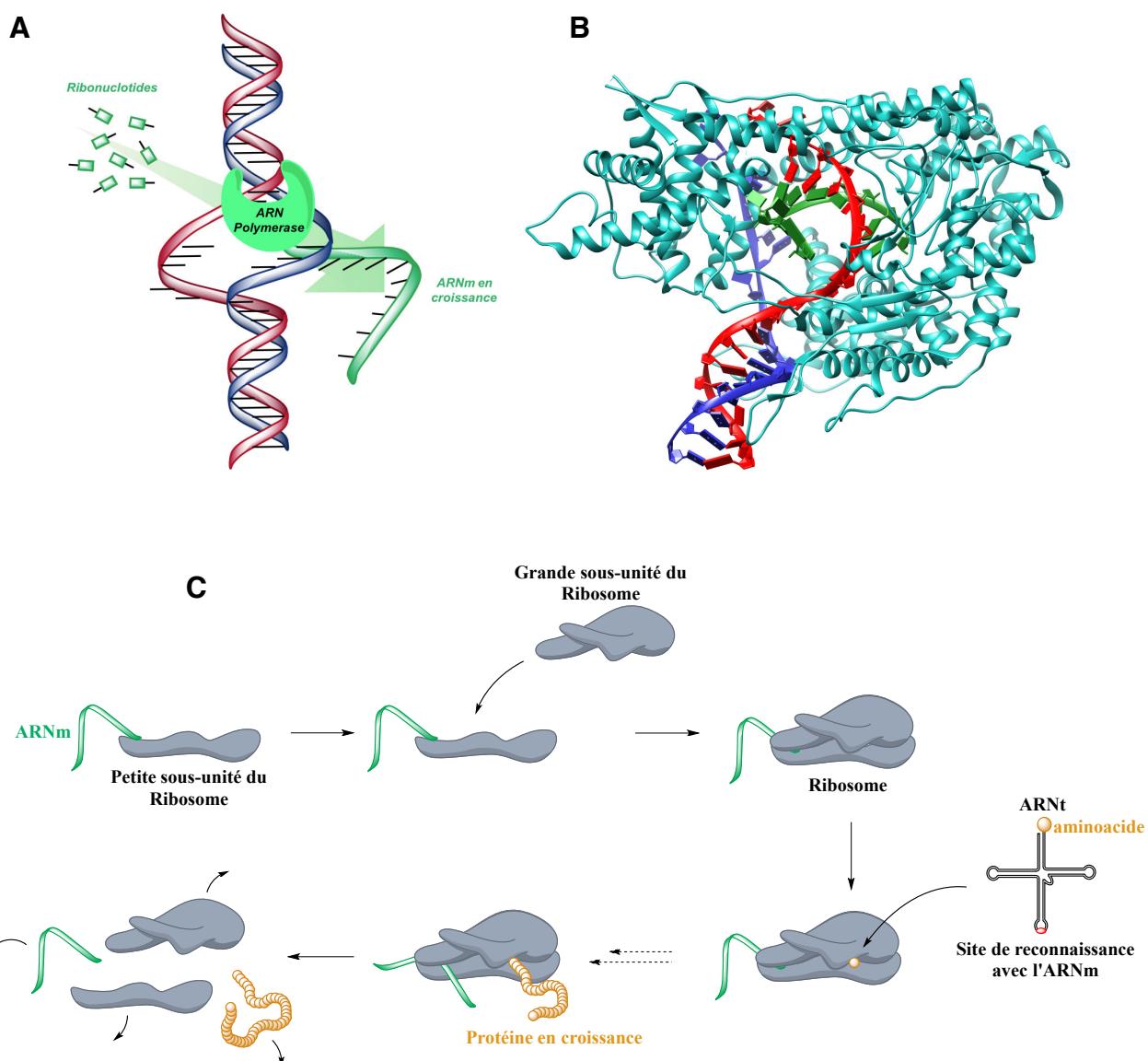


Figure 7. (A) Transcription de l'ADN en ARNm, (B) Structure obtenue après diffraction des rayons X (RX) d'une ARN polymérase et de l'ARNm en synthèse (vert) (PDB : [1MSW](#))^{12a} et (C) traduction de l'ARNm en protéine.

L'ARNm subit ensuite une étape de maturation au cours de laquelle les parties non-codantes, introns, sont éliminées : c'est l'épissage. Une fois arrivé à maturité, il peut alors transporter l'information génétique qu'il renferme du noyau vers le cytoplasme de la cellule où se trouve la machinerie cellulaire (ribosome). Une fois arrivé, les deux sous-unités du ribosome viennent interagir avec lui et c'est avec la formation de ce complexe que débute la phase de traduction [cf. Figure 7C].¹³ Elle consiste à convertir le message génétique transporté par l'ARNm en unités protéiques (protéines, enzymes).

L'ARNm, tout comme l'ADN, est constitué de ce que l'on appelle des codons. Un codon est une séquence de 3 nucléotides faisant référence à un acide aminé bien particulier [cf. Tableau 1 - ANNEXE S1]. L'enchaînement des codons les uns à la suite des autres détermine la structure primaire de la protéine à synthétiser.

Afin de produire ces protéines, le complexe ribosome / ARNm doit interagir avec différents transporteurs d'acides aminés, les ARN de transferts (ARNt). Ces entités, en forme de « L », présentent chacune un site de reconnaissance spécifique à un codon de l'ARNm : l'anti-codon. A chaque codon de l'ARNm correspond un anti-codon de l'ARNt, et donc un acide aminé de la protéine.

Au contact du codon-start (en vert dans le Tableau 1 - ANNEXE S1), la synthèse protéique débute. Le ribosome parcourt alors le brin d'ARNm de codon en codon. Pour chacun d'entre eux, un ARNt vient interagir avec la séquence d'ARNm (interaction anti-codon / codon) et l'acide aminé qu'il porte vient s'ajouter à la protéine en croissance dans une des cavité du ribosome. A la lecture du codon-stop (en rouge dans le Tableau 1 - ANNEXE S1), la synthèse s'arrête et le complexe ARNm / petite sous-unité du ribosome / grande sous-unité du ribosome / protéine se dissocie [cf. Figure 7C].

iii. Génome et division cellulaire

Notre organisme compte plusieurs milliers de milliards de cellules de formes et d'activités diverses (250 types différents entre les cellules de la peau, du foie, les neurones, ...) qui se renouvellent à un rythme effréné. Environ 1 milliard de nouvelles cellules se forment dans notre organisme chaque jour grâce, entre autres, à la division cellulaire [cf. Figures 8 & 9]. Il s'agit du phénomène au cours duquel une cellule « mère » se divise afin de donner deux cellules « filles », présentant chacune le même patrimoine génétique, la même spécialisation (cellule musculaire, hépatique, de la peau, ...) que la cellule mère.¹⁴

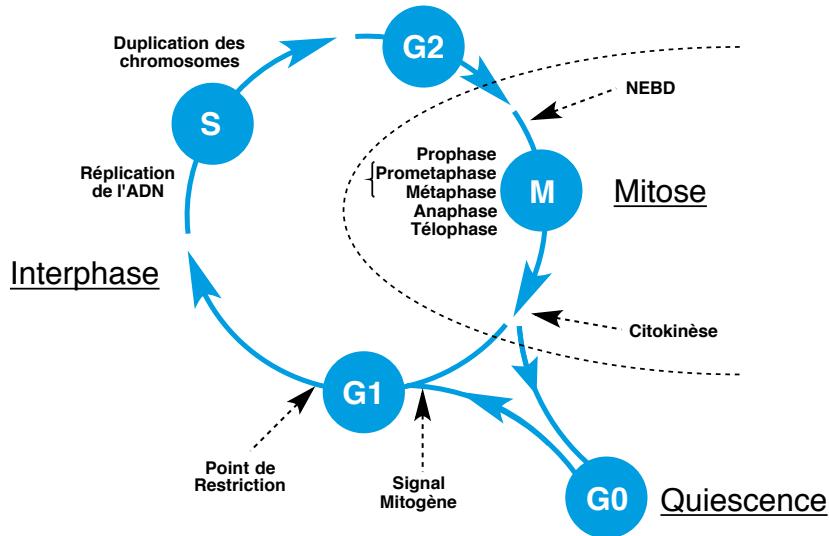


Figure 8. Représentation schématique du cycle de division cellulaire.

Ce cycle cellulaire est constitué de deux phases principales [cf. Figure 8]. La première, l'interphase, se fait sur trois étapes (G1, S et G2) et permet la croissance des cellules et la réplication de l'ADN. La seconde, la mitose, est sub-divisée en cinq étapes (prophase, prométaphase / métaphase, anaphase, télophase et citokinèse [cf. Figures 8 & 9]) et vise à former les deux cellules filles et à y répartir un exemplaire de chaque chromosome répliqué. Lorsque les cellules ont cessé de se diviser, elles quittent ce cycle de division et se mettent au repos : on dit qu'elles entrent en quiescence [cf. Figure 8].^{14b}

G1 est la première phase de croissance cellulaire. Elle correspond au point de départ de la division cellulaire sous l'influence d'hormones mitogènes. La phase S nous intéresse plus particulièrement ici, car elle correspond à la réplication de l'ADN [cf. Figure 10A]. La phase G2 est la deuxième phase de croissance cellulaire et précède la phase M de mitose cellulaire dont l'ensemble de ses cinq étapes permet, *in fine*, la formation des deux cellules filles.^{14b}

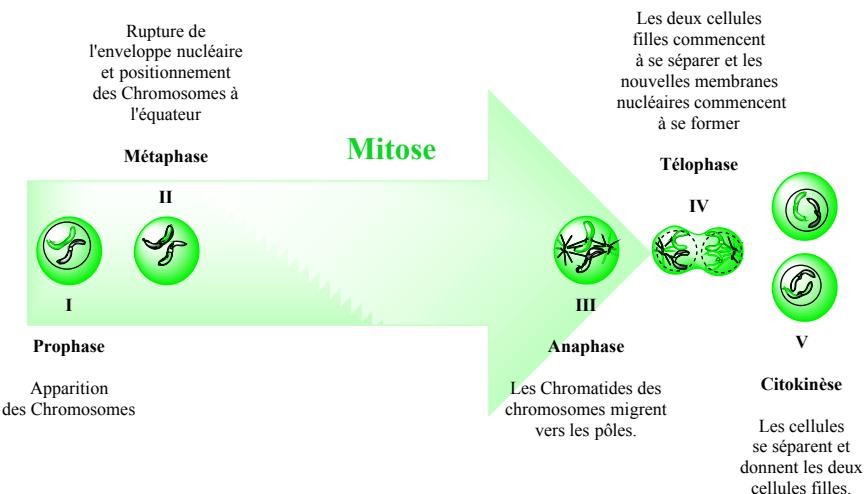


Figure 9. Représentation schématique de la Mitose.

La phase S de réplication de l'ADN est réalisée en trois temps : l'initiation, l'elongation et la terminaison. Elle débute à des points spécifiques de l'ADN appelés

'origine de réPLICATION' et met en jeu une trentaine de protéines différentes, dont des ADN hélicases [cf. Figure 10A] et des ADN polymérases [cf. Figure 10A et B]. Les hélicases ont la capacité d'ouvrir la double hélice de l'ADN afin de former des fourches de réPLICATION. Ceci permet aux polymérases III de venir dupliquer le matériel génétique contenu sur chacun des brins parents [cf. Figure 10A]. La réPLICATION des deux brins s'effectuant en parallèle, deux ADN polymérases III sont donc nécessaires, une sur chacun des brins parents. Afin de les maintenir sur leurs brins respectifs, d'autres entités protéiques sont mises en jeu : les « *sliding clamps* » [cf. Figure 10A et C].^{14d}

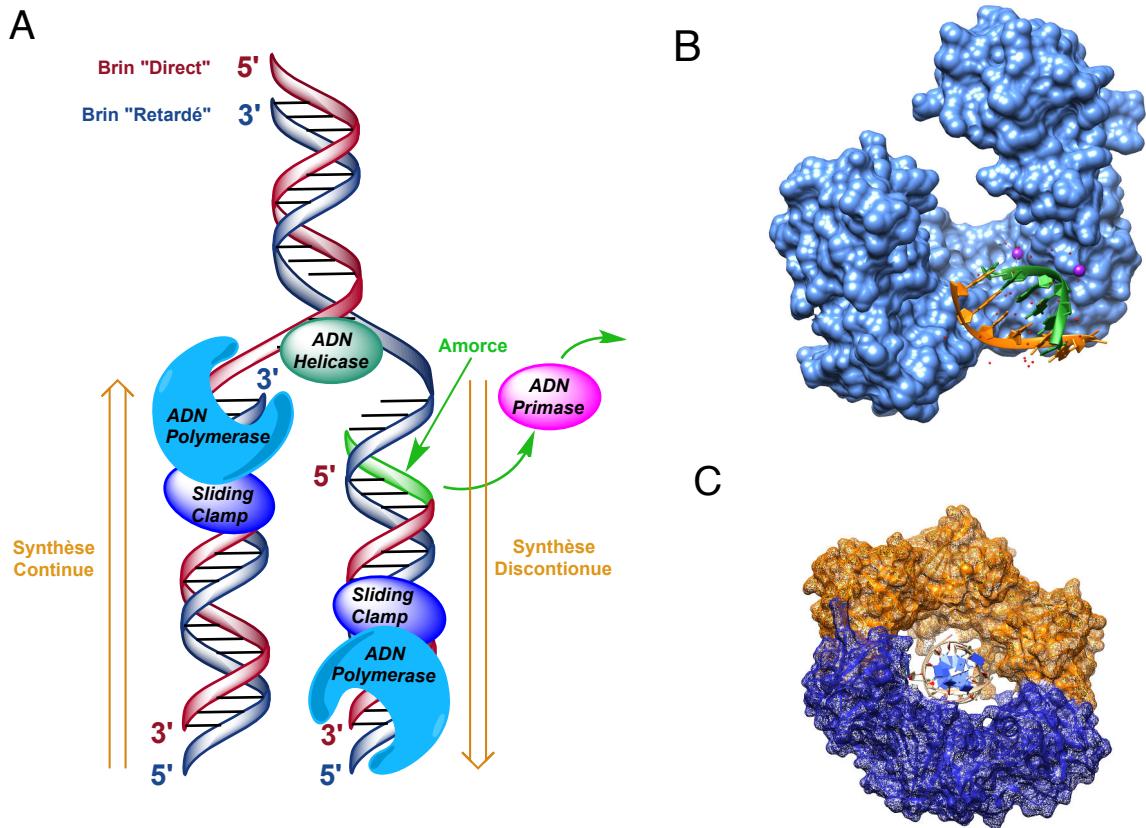


Figure 10. (A) Représentation schématique de la phase S - RéPLICATION de l'ADN et (B) structure RX d'un complexe ADN polymérase / double brin d'ADN (PDB : [7ICG](#))^{14a} et (C) d'un « *sliding clamp* » complexé à l'ADN double brin (PDB : [3BEP](#)).^{14c}

Au cours de la phase d'élongation, chaque ADN polymérase III progresse le long du brin parent auquel elle est liée, dans le sens 5' → 3'. Au fur et à mesure de leur avancée, elles lisent l'information contenu sur le brin parent et rajoutent les nucléotides en position 3' sur la séquence complémentaire en croissance. Cependant, les deux brins parents étant anti-parallèles, deux mécanismes différents ont lieu. Une des ADN polymérases III effectue la synthèse « continue » à partir du brin parent « direct », orienté 3' → 5' (rouge sur la Figure 10A), et l'autre effectue la synthèse « discontinue » à partir du second brin parent, le brin « retardé » (bleu sur la Figure 10A).^{14d}

La réPLICATION du brin direct est réalisée en continu alors que pour le brin retardé, elle s'effectue de manière séquentielle, grâce à l'intervention répétée de l'ADN primase. Les ADN polymérases III ne peuvent pas effectuer de réPLICATION sans la présence d'un

nucléotide hydroxylé en 3' (3'-OH). Des amores d'ARN, synthétisées par l'ARN polymérase et présentant ce nucléoside 3'-OH, sont donc positionnées toutes les 200 pb (eucaryotes), par l'ADN primase, le long du brin retardé, afin de faire progresser sa réPLICATION [cf. Figure 10A]. Ces unités de 200 pb de long sont appelées « les fragments d'Okazaki ».¹⁴ Une fois que l'un d'entre eux a été synthétisé, l'ADN polymérase I vient remplacer l'amorce d'ARN [cf. Figure 10 - en vert] par les 2'-désoxyribonucléotides correspondants et l'ADN ligase vient ensuite relier ce fragment nouvellement formé aux précédents, déjà liés entre eux. A ce moment là, une nouvelle amorce est apportée par l'ADN primase, et la synthèse du brin retardé peut continuer.^{14d}

Finalement, le processus s'achève lors de la « collision » entre deux fourches de réPLICATIONS ou lorsque chacune d'entre elles rencontre un signal de terminaison.

3. Les marques épigénétiques

Alors que le décryptage du génome nous donnait l'illusion de pouvoir expliquer le phénotype des individus, un nouveau code, le code épigénétique, est découvert.

Le terme « épigénèse » a été utilisé par Aristote (384–322 av. J.-C.) afin d'expliquer la formation progressive des organes au stade embryonnaire. Le terme « épigénétique » a, quant à lui, été défini plus récemment (1942) par Conrad Hal Waddington, un biologiste et embryologiste britannique.¹⁵ Il vient du grec « épi » (έπι) et « génétique » (γενετικός) signifiant « au-dessus du gène ». Ce terme caractérise l'ensemble des variations des caractères phénotypiques cellulaires et physiologiques non provoquées par des mutations de la séquence d'acides nucléiques (génotype), et pouvant être transmises à la descendance.

L'expression combinée du génome et de son épigénome semble être la clé à l'origine de notre phénotype. En effet, les modifications d'expression du génome, conditionnées par l'épigénome, ont une influence sur le développement des individus. Contrairement aux mutations, ces épi-modifications sont réversibles (un paramètre majeur pour l'approche thérapeutique) et se produisent au niveau de l'ADN ou des protéines l'entourant (histones). Elles modifient l'accessibilité aux gènes ; elles influencent leur expression.

Le code épigénétique est caractérisé par trois différents marqueurs, que des chercheurs du monde entier tentent de décrypter. Le tout premier touche les histones, des protéines permettant de réguler le compactage de la chromatine et donc l'accessibilité aux gènes. Le second, concerne les modifications chimiques des bases de l'ADN telles que leur méthylation. Enfin, un dernier marqueur concerne les ARNs non-codants, comme les micro-ARNs (miARN), qui ont eux aussi une influence sur la transcription des gènes et dont un certain nombre portent des cytosines méthylées pour de nouveaux rôles biologiques. Le manque d'outils pour étudier ce domaine fait stagner la compréhension de leur importance (cf.I.A.3.iv).

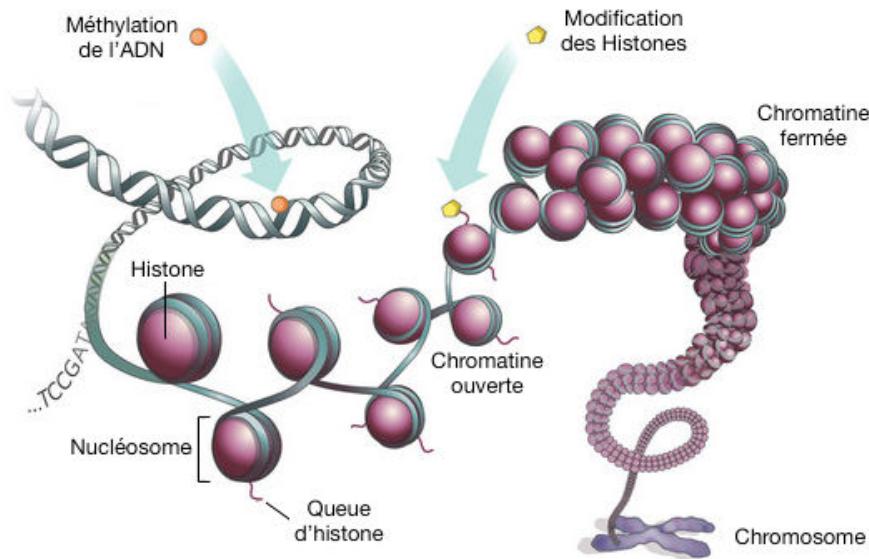


Figure 11. Marques épigénétiques (adapté de Marx *et al.*).¹⁶

i. Modification des Histones / Modification de la chromatine.

L'ADN est une macromolécule gigantesque. Pour donner un ordre de grandeur, si l'on met bout à bout le matériel génétique présent sur les 23 paires de chromosomes humain, notre génome mesure approximativement 2 m de long. Or, le noyau d'une cellule ne mesure que quelques micromètres (10 à 100 µm) de diamètre. Ainsi, afin de faire rentrer le génome (et du coup son épigénome) dans un espace aussi restreint, il est indispensable de le compacter.

Le compactage de l'ADN passe par les nucléosomes. Il s'agit d'un assemblage nucléoprotéique constitué d'un cœur protéique (les histones - H) et de l'ADN qui s'est enroulé autour (environ 147 pb) [cf. Figure 11 & 12]. Plus particulièrement, le cœur est un octamère formé d'un tétramère H3–H4 et de deux dimères H2A–H2B [cf. Figure 12A et B]. Les nucléosomes se retrouvent toutes les 157 – 240 pb, si bien que dix millions d'entre eux se trouvent dans le noyau de nos cellules.¹⁷

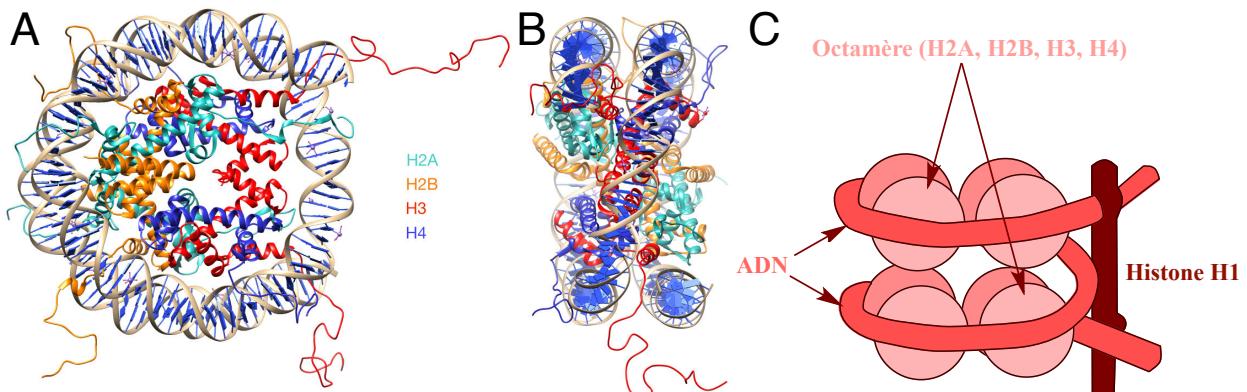


Figure 12. (A) Structure RX d'un nucléosome vue de face et (B) de profil – Résolution : 1,9 Å (PDB : 1KX5) et (C) représentation schématique de l'histone H1.

Dans un second temps, les nucléosomes se rapprochent les uns des autres afin de former des fibres de chromatine. Plus le taux de compactage est élevé (nucléosomes

« serrés ») et plus la fibre de chromatine est dense, on parle alors d'hétérochromatine [cf. Figure 11]. Afin de maximiser le compactage de cette dernière, une autre histone, l'histone H1 [cf. Figure 12C], vient se positionner aux points d'entrée et de sortie de l'ADN sur le nucléosome afin de « sceller », de manière plus efficace encore, le complexe nucléoprotéique.¹⁸ L'encombrement stérique important rend impossible l'accès à l'information génétique et donc l'hétérochromatine correspond à une répression génique. Quand le taux de compactage est plus faible (euchromatine), l'information génétique est accessible aux complexes enzymatiques (comme l'ARN polymérase). Cela se traduit par l'expression des gènes.¹⁹

Le taux de compactage de la chromatine est directement lié à la distance qui sépare les différents nucléosomes les uns des autres. Mais cette distance dépend elle-même des modifications potentiellement subies par les histones (méthylation, acétylation, ubiquitylation, ...) au niveau de leurs chaînes N-terminales (*visibles sur les Figures 12A et B*).²⁰ Ces modifications sont réversibles et touchent des acides aminés bien spécifiques tels que des lysines (K) ou des arginines (R) et bien d'autres. Les modifications les plus étudiées sont l'acétylation et la méthylation.^{19, 20}

Les enzymes responsables de ces modifications sont très spécifiques. Par exemple, les enzymes à l'origine de l'acétylation et de la désacétylation des histones sont respectivement HAT (*Histone Acetyl Transferase*) et HDAC (*Histone DeAcetylases*). De son côté, la méthylation est caractérisée par l'enzyme de méthylation HMT (*Histone MethylTransferases*) et par les enzymes de démethylation (LSD)1 (*Lysine-Specifique Demethylase*) ou JmjC (*JmjC domain-containing histone demethylase*).¹⁸ Il est à noter que, au-delà de la modification des histones, certaines de ces enzymes modifient également des protéines comme les facteurs de transcription.²¹

L'acétylation des histones est plus communément associée à l'activité transcriptionnelle alors que la désacétylation contribue à l'extinction de la transcription. En effet, l'acétylation d'une lysine masque la charge positive de son ammonium à la suite de la formation de l'amide. L'ammonium ne pouvant plus interagir avec les charges négatives des phosphates de l'ADN, le complexe nucléoprotéine se relâche en laissant un libre accès au gène à transcrire. De son côté, la méthylation des lysines K4, K36, et K79 de H3 est associée à la transcription des gènes alors que la méthylation de K9 et K27 de H3 et K20 de H4 est liée à l'effet inverse.^{19, 22}

ii. Méthylation de l'ADN

La méthylation de la C est la principale modification épigénétique touchant directement l'ADN. Ce processus permet l'ajout d'un motif CH₃ (méthyle) en position 5 du noyau pyrimidique (5mC) et ne modifie en rien la séquence d'ADN [cf. Figure 13]. Ce phénomène réversible est tout sauf aléatoire, il répond à des stimuli bien particuliers dont certains, évoqués ultérieurement [cf. ANNEXE S2], peuvent être extérieurs à l'organisme. Chez l'homme, 1 % du génome est occupé par des 5mC pour environ 21 % de C non modifiées (ou Cp - p = phosphate).²³

La synthèse de la 5mC est catalysée par des enzymes appelées DNMTs (*DNA MethylTransferases*). Il en existe au moins 5 : DNMT1, DNMT2, DNMT3a, DNMT3b et DNMT3L. Chacune joue un rôle bien spécifique qui sera présenté plus en détails dans les prochains paragraphes (cf. I.B.2.i).

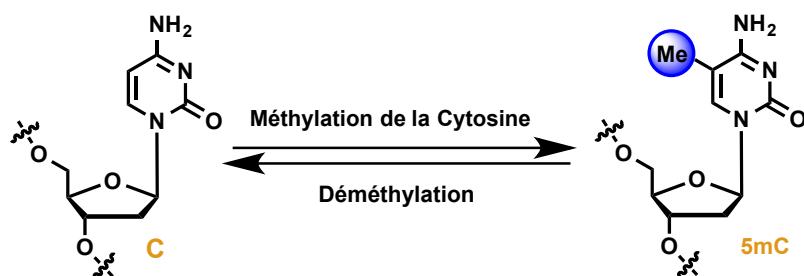


Figure 13. Schéma représentant la méthylation réversible de la cytosine.

Elles abritent, au sein de leur structure, une poche accommodant le cofacteur SAM (pour S-Adenosyl-L-méthionine - ou AdoMet) donneur du groupe méthyle [cf. Figure 14].⁸ Ce composé, obtenu suite à la combinaison d'ATP et de méthionine par la *Méthionine Adénosyltransférase*, présente un groupe méthyle hautement réactif. Les méthylases, comme les DNMTs, l'utilisent afin de méthylérer leurs substrats. Une fois le transfert de méthyle réalisé, le SAM laisse la place au SAH, la *S*-adénosyl-L-homocystéine (ou AdoHcy) [cf. Figure 14], qui peut alors être recyclé en SAM.

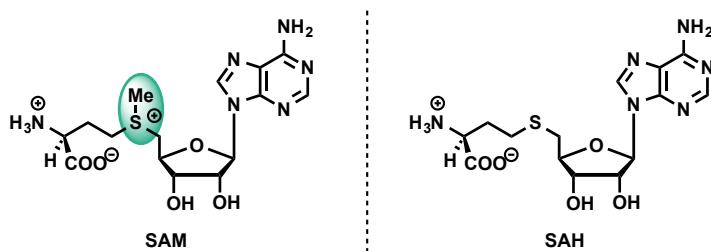


Figure 14. Structures de SAM et SAH.

Chez les mammifères, la 5mC se retrouve quasi exclusivement au niveau de dinucléotides CpG et plus rarement au niveau de trinucléotides CpXpG ou CpXpX (X = A, T ou C). Les dinucléotides sont répartis sur l'ensemble du génome ; 70 à 80 % d'entre eux sont C méthylés de manière constitutive chez l'homme. Le tiers restant correspond à des régions riches en CpG, dénommées îlots CpG. Ils contiennent environ 55 % de dinucléotides CpG et dans 50 % des cas, ils se trouvent au niveau de régions promotrices de gènes. La plupart du temps, ils sont déméthylés, sauf dans les tumeurs.^{24, 25}

Comme indiqué précédemment, les grand et petit sillons exposent, vers l'extérieur du duplexe d'ADN, un profil de liaisons hydrogènes particulier, indispensable à l'interaction avec certaines protéines (facteurs de transcriptions, par exemple).⁸ Des altérations de ce profil, comme la méthylation de l'ADN, influencent la capacité de ces entités à se lier ou non à l'ADN. En effet, les états actifs et réprimés des gènes sont corrélés au taux de méthylation des îlots CpG de leurs régions promotrices.²⁶ Il a par exemple été démontré en 1980 par l'équipe de Cedar, que la méthylation *in vitro* du promoteur du gène de la β -globuline, par l'intermédiaire d'une méthylase, empêchait sa transcription et d'ailleurs, le

traitement de certaines hémoglobinopathies, par des agents déméthylants, est actuellement promu.^{25a, 8}

La présence des groupes méthyles sur les îlots CpG des promoteurs géniques se traduit par une répression transcriptionnelle. Deux mécanismes peuvent être mis en cause :

- la présence des groupes méthyles provoque une gène stérique qui empêche la liaison des différentes entités de transcriptions. Ne pouvant pas être lu, le gène est donc réprimé.^{25c}
- les îlots CpG méthylés peuvent être reconnus par au moins trois familles de protéines capables de recruter des enzymes de modification des histones pour provoquer un remodelage de la chromatine. La méthylation de l'ADN s'accompagne donc d'une structure répressive de la chromatine.²⁶

Le profil de méthylation de l'ADN s'établit au tout début du développement car il est indispensable au développement complet et harmonieux de l'embryon.²⁷ Il est en effet directement impliqué dans la mise en place de « l'empreinte génomique »,²⁸ dont l'exemple le plus connu est l'inactivation d'un des chromosome X chez la femme.²⁹ En effet, cette empreinte permet l'expression mono-allélique d'un gène, ce qui permet de doubler le répertoire d'expression du génome. Dans certains cas, ces méthylations mono-alléliques sont transmises à l'embryon par un seul de ces deux parents : c'est l'empreinte parentale (cf. I.B.1.ii). La méthylation de l'ADN permet aussi de réguler la prolifération et la survie des cellules souches totipotentes (Dolly) et celle des cellules adultes permettant par exemple la croissance dendritique de nouveaux neurones.^{23, 30}

Notons que chez les bactéries, un autre marqueur de méthylation a été repéré : la 6-méthyl-2'-désoxyadénosine (6mA) [cf. Figure 15]. Elle permet de réguler la réPLICATION et la transcription bactérienne et facilite la réparation des mésappariements de l'ADN. Mais de récentes études ont démontré que cette 6mA était également présente chez de nombreux eucaryotes (plantes et champignons) et qu'elle pouvait également jouer un rôle épigénétique (drosophiles, vers et plantes). Chez *Chlamydomonas reinhardtii* (micro algue), elle est présente dans près de 14 000 gènes, soit 84 % du génome de cette algue. On la retrouve essentiellement au niveau de dinucléotides ApT et elle est fortement présente au niveau des promoteurs de gènes (contrairement à la 5mC, présente de façon prépondérante sur le reste du génome). Elle semble jouer un rôle quant au positionnement de l'ADN au niveau du nucléosome, et semble également intervenir sur la régulation de l'expression des gènes. Ces études soulèvent la question de la présence d'un tel marqueur au sein du génome des mammifères ou plus généralement des vertébrés, sachant que sa présence est avérée dans leur ARNm.³¹

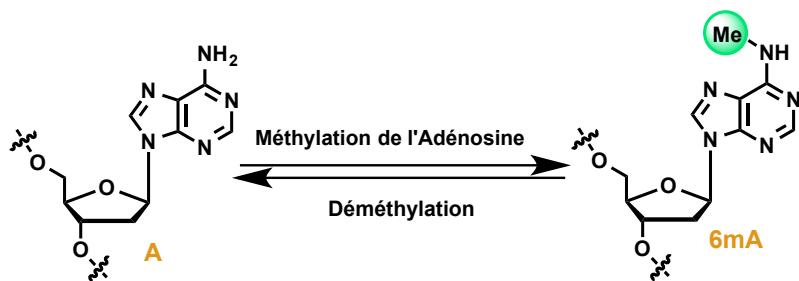


Figure 15. Schéma représentant la méthylation réversible de l'adénosine.

iii. Autres modifications de l'ADN

La méthylation de l'ADN s'accompagne également du phénomène inverse : la déméthylation. Cette déméthylation était, et est toujours, obligatoire puisque mesurée lors de la maturation de gamètes (cf. I.B.1.ii). Cependant, on ne savait pas à quoi l'attribuer. Alors que la méthylation est étroitement liée à la réPLICATION, la déméthylation, elle, ne semble pas dépendre du cycle cellulaire. Comme vu précédemment, l'hyperméthylation d'un promoteur de gène entraîne la réPRESSION de ce gène ; sa déméthylation a plutôt TENDANCE à l'ACTIVER.

La 5mC peut être déméthylée via deux processus bien distincts l'un passif et l'autre actif. En ce qui concerne le mécanisme passif, il semble être lié à un manque de maintien du profil de méthylation après la réPLICATION de l'ADN. Une perte d'activité de l'enzyme responsable de ce maintien, DNMT1, ou bien une absence (diminution) de complexes DNMT1 / ADN peuvent provoquer, au cours des mitoses successives, une déméthylation progressive de l'ADN.^{25b, 32}

La version active de ce mécanisme semble être un mécanisme de « nettoyage », permettant de réparer les erreurs de méthylation acquises au cours de la division cellulaire.³³ Plusieurs hypothèses ont été avancées quant à ce mécanisme. Les premières étapes, où l'oxydase TET (*Ten-Eleven Translocation*) oxyde de manière séquentielle la 5mC en 5caC (*5-carboxylic acid Cytosine*), en passant par les intermédiaires hydroxyméthyle (5hmC) et formyle (5fC), semble faire l'unanimité [cf. Figure 16].^{8, 31b, 34} Les avis divergent quant à l'étape finale, permettant le retour à la cytosine [cf. Figure 16] :

- (1) Une première hypothèse, largement acceptée, procède selon un mécanisme de réparation de l'ADN par excision de la base (BER : *Base Excision Repair*) qui implique une série d'enzymes. Elle débuterait par l'excision des bases 5fC et/ou 5caC par la TDG (*Thymine DNA Glycosylase*) afin de former un site abasique sur l'ADN en cours de « nettoyage ». ^{25b, 35}
- (2) Une seconde hypothèse traite de l'action d'une décarboxylase, encore inconnue, qui viendrait décarboxyler 5caC, afin de redonner directement la cytosine de départ. L'hypothèse d'une déformylase, qui agirait sur 5fC, a également été évoquée.^{25b, 8}
- (3) Enfin, une dernière hypothèse, uniquement démontrée *in vitro* pour l'instant, semble indiquer que DNMT1 serait capable de convertir 5hmC en C.^{25b}

Parmi ces analogues de cytosine, notons que 5hmC peut également être considérée comme un marqueur épigénétique à part entière, à l'instar de 5mC. Par exemple, dans le cas où le chromosome X est inactif, on constate une forte déplétion en 5hmC et lorsque le chromosome est réactivé, elle s'accumule de nouveau. De plus, un enrichissement en 5hmC au niveau de promoteurs de gènes, et simultanément une diminution en 5mC, semble permettre de réguler l'expression des gènes (grâce à la variation du rapport 5hmC / 5mC). Il a également été démontré que le taux de 5hmC varie énormément dans les tissus, comme c'est le cas pour 5mC. Elle semble aussi intervenir au cours du développement embryonnaire.^{23, 31b}

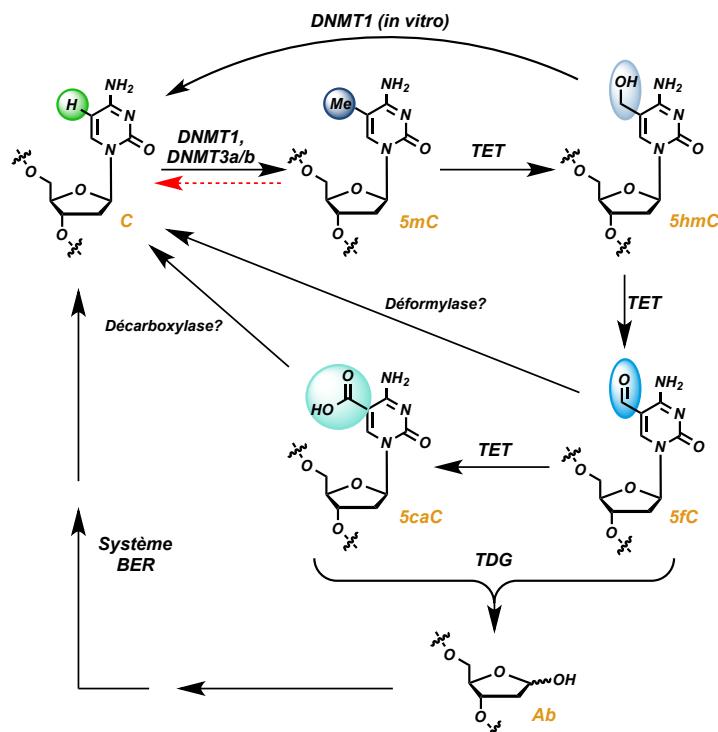


Figure 16. Déméthylation active et passive (flèche rouge) de la 5mC.

Enfin, en ce qui concerne la 6mA, les premières études démontrent l'existence d'enzymes permettant de la déméthyler activement : DMAD chez les drosophiles, NMAD-1 chez *Caenorhabditis elegans* et il semblerait qu'une enzyme similaire soit également présente chez l'homme, METTL4.^{31a} Cependant, aucun mécanisme de déméthylation passive de la 6mA n'a été encore observé.

iv. Les micro-ARNs

Les miARNs sont des portions d'ARN de 21 à 23 pb de long. Ils interviennent au cours du cycle cellulaire, lors de l'apoptose, du développement et de la différenciation cellulaire. Ils sont encodés par le génome et ont la capacité d'accélérer la dégradation des ARNm ou de réprimer la traduction protéique (cf. I.A.2.ii). Ils régulent ainsi, en association avec les protéines argonautes (Ago), l'expression de 60 % de notre génome. On dit que ce sont des répresseurs post-transcriptionnels.³⁶ Ces fragments d'ARN ont également la capacité d'influencer sur la structure de la chromatine, la méthylation de l'ADN et l'expression des gènes codants les facteurs de transcription.³⁷

Les miARNs sont transcrits par l'ARN polymérase II sous la forme de pré-miARN, qui sont ensuite clivés par les enzymes Drosha (*ribonuclease III de classe 2*) et DGCR8 (*Di George Critical Region 8*). Ces précurseurs ainsi obtenus mesurent 70 nt de long et adoptent une structure en tige boucle. Il sont ensuite transportés dans le cytoplasme, où l'enzyme Dicer (endoribonuclease) hydrolyse la tige-boucle afin de donner un double brin d'ARN [cf. Figure 17]. Son interaction avec une des protéines argonautes (*Ago1 à 4*) et la protéine TRBP (*transactivating ARN-binding protein*) donne le complexe RISC (*RNA-induced silencing complex*). Au cours de la formation de RISC, le miARN double brin laisse la place à un simple brin [cf. Figure 17].^{37b, 38}

Le complexe RISC peut alors se positionner au niveau de l'ARNm, par complémentarité de base miARN / ARNm. Si la protéine Ago2 est présente dans le complexe RISC et si les deux brins sont parfaitement complémentaires, la dégradation de l'ARNm cible est induite. En revanche, si l'une des autres protéines Ago (1, 3 ou 4) est présente dans le complexe et si la complémentarité miARN / ARNm est partielle (mésappariement), la traduction est réprimée [cf. Figure 17]. Cette reconnaissance partielle rend un seul miARN capable de cibler plus de 200 gènes.^{37, 38}

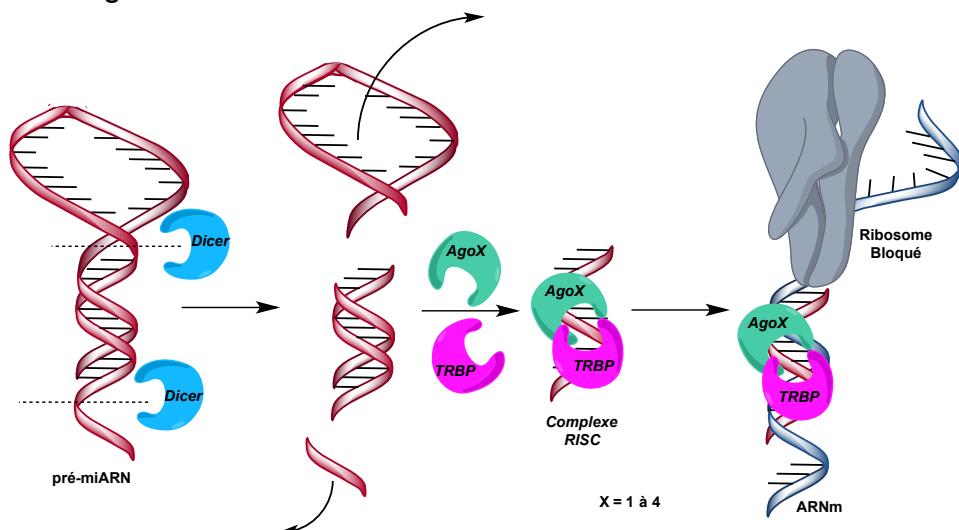


Figure 17. Formation du complexe RISC et mode d'action des miARN sur la traduction.

Les miARN jouant naturellement un rôle sur la prolifération cellulaire, des modifications de leur taux ont été associées au développement des cancers. En effet, la formation de clusters de miARN ou leur duplication ont été observées au niveau d'oncogènes (oncomiARN) ou de gènes suppresseurs de tumeurs. Ces modifications peuvent être provoquées par un dysfonctionnement des enzymes Dicer et Drosha ou par des modifications épigénétiques des promoteurs de gènes codants pour ces régulateurs nucléiques (méthylation de l'ADN et modifications de leurs histones).³⁷

L'ensemble des marques épigénétiques qui viennent d'être présentées agissent en synergie afin d'atteindre un but commun : soit exprimer, soit réprimer des gènes. Elles évoluent en fonction des facteurs environnementaux (nourriture, exposition au soleil, tabagisme, pollution, ...) auxquels un individu est exposé et peuvent même subir des

dérégulations, pouvant aboutir à différentes pathologies (cancers, Alzheimer, Huntington, ...). Ces modifications de leur mode de fonctionnement peuvent par la suite être transmises à la descendance (mitose et gamétogenèse).

La méthylation de l'ADN est une marque particulièrement intéressante car elle joue à la fois un rôle direct sur l'expression des gènes et un rôle indirect sur la structure de la chromatine et sur l'expression des miARNs. Elle semble donc être la « clé de voûte » de toute cette machinerie transcriptionnelle complexe. Il n'en reste pas moins que, bien qu'étant associées, méthylation de l'ADN et modifications des histones ne sont pas encore temporellement définies. En d'autres termes, qu'elle est celle qui commence, la première ou la seconde, pour éteindre les gènes ?

Intéressons-nous, à présent, au rôle biologique de la méthylation de l'ADN, aux effets que l'environnement peut avoir sur elle et à son mécanisme, à l'échelle moléculaire.

B. Rôle et mécanisme de la méthylation de l'ADN

1. Développement et Méthylation de l'ADN

i. Division Cellulaire

Comme nous l'avons déjà évoquée (cf. I.A.2.ii), la réPLICATION de l'ADN permet de dupliquer l'information génétique afin de donner deux duplexes d'ADN identiques, qui sont ensuite répartis entre les deux cellules filles [cf. Figure 10A]. La cellule mère présente un profil de méthylation symétrique du fait que ses deux brins soient méthylés. Cependant, au cours de la réPLICATION, les nucléotides utilisés pour produire les brins complémentaires des brins parents sont les dérivés dans leurs formes naturelles, à savoir qu'ils n'ont subi aucune modification chimique. Ainsi, une fois la réPLICATION achevée, les cellules filles nouvellement formées ne présentent qu'un seul brin d'ADN méthylé, à savoir le brin matrice provenant de la cellule mère [cf. Figure 18A] ; l'autre brin nouvellement synthétisé, lui, ne l'est pas : on parle donc d'ADN hémiméthylé. Afin de redonner un ADN symétrique, identique à celui de la cellule mère, DNMT1 vient méthylé le brin néo-formé, en se basant sur le modèle des 5mC présentes sur le brin parent. Les cellules filles vont donc ainsi posséder un ADN di-méthylé symétrique fidèlement reproduit selon le modèle de la cellule mère [cf. Figure 18B].^{23, 33}

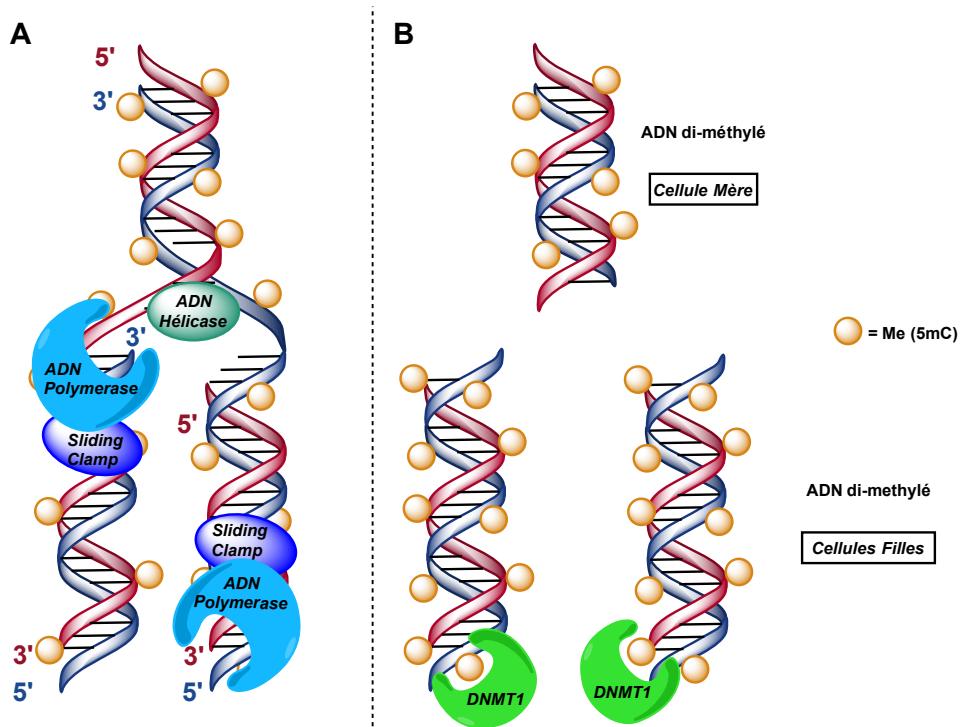


Figure 18. (A) Représentation simplifiée de la phase de réPLICATION de l'ADN et (B) de la conservation du profil de méthylation de l'ADN au cours de la mitose d'une cellule mère en deux cellules filles.

Cette étape de méthylation du nouveau brin est indispensable au bon fonctionnement de nos cellules. Nous avons vu précédemment que la déméthylation passive de l'ADN pouvait apparaître à cause d'un manque d'activité de cette enzyme ou à l'absence de complexe DNMT1 / ADN et progresser au fur et mesure des divisions cellulaires successives.

ii. Empreinte génomique parentale

Outre la mitose (cf. I.A.2.iii), un autre processus de division cellulaire a lieu lors de la production de gamètes (gamatogénèse), il s'agit de la méiose [cf. Figure 19]. Les gamètes sont les cellules reproductrices (spermatozoïdes et ovules), qui sont à l'origine du transfert d'information génétique des parents vers la descendance. Les gamètes ne présentent que la moitié du génome de chaque parent, soit 23 chromosomes. La fusion d'un ovule et d'un spermatozoïde donne lieu à une cellule-œuf qui présente à nouveau un code génétique complet (46 chromosomes), indispensable au développement du nouvel être. Ce dernier possède donc, dans son matériel génétique, 50% du patrimoine de chacun de ses parents.

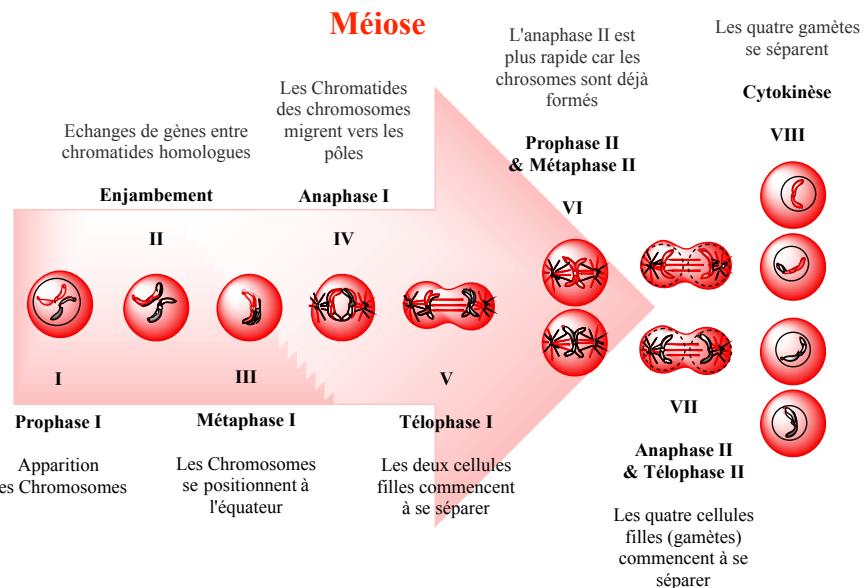


Figure 19. Schéma représentatif de la Méiose des gamètes.

La rencontre entre un génome mâle et un génome femelle est indispensable car ils sont complémentaires l'un de l'autre. Cependant ils présentent des différences fonctionnelles dues à une signature bien spécifique de chaque génome : c'est ce que l'on appelle « l'empreinte génomique parentale » ou encore « l'expression mono-allélique » des gènes. Par exemple, un même gène peut être « éteint » chez le père et « allumé » chez la mère. Cela implique donc que le gène transmis par le père ne sera pas exprimé chez l'enfant. Ce type de gène est ce que l'on appelle un « gène soumis à empreinte ».^{28, 39, 40}

L'empreinte génomique parentale se met en place au moment de la gamétogenèse. Pour ce faire, les cellules germinales primaires, obtenues après méioses, subissent d'abord une étape d'effacement au cours de laquelle les 5mC sont déméthylées. Ceci permet la création, dans un second temps, d'une nouvelle empreinte propre au sexe du futur embryon, mais toujours liée à l'empreinte parent [cf. Figure 20]. La Figure 20 montre l'évolution du taux de méthylation de l'ADN au cours de cette gamétogenèse et du développement embryonnaire. On peut constater que les gènes soumis à empreintes (rouge et bleu foncés) restent méthylés après la fécondation alors que les autres gènes (non-soumis à empreintes : rouge et bleu clair) se déméthyilent. Cette déméthylation semble avoir pour rôle de mettre fin aux phénomènes de régulations propres aux gamètes. Le nouveau profil d'expression génique de l'embryon peut alors se mettre en place et les tissus peuvent commencer leur différenciations (neurones, cellules musculaires, cardiaques, ...).^{28, 40} Les mécanismes précis qui sous-tendent la mise en place et la transmission des empreintes géniques, et donc la méthylation précise de certains gènes, sont inconnus.

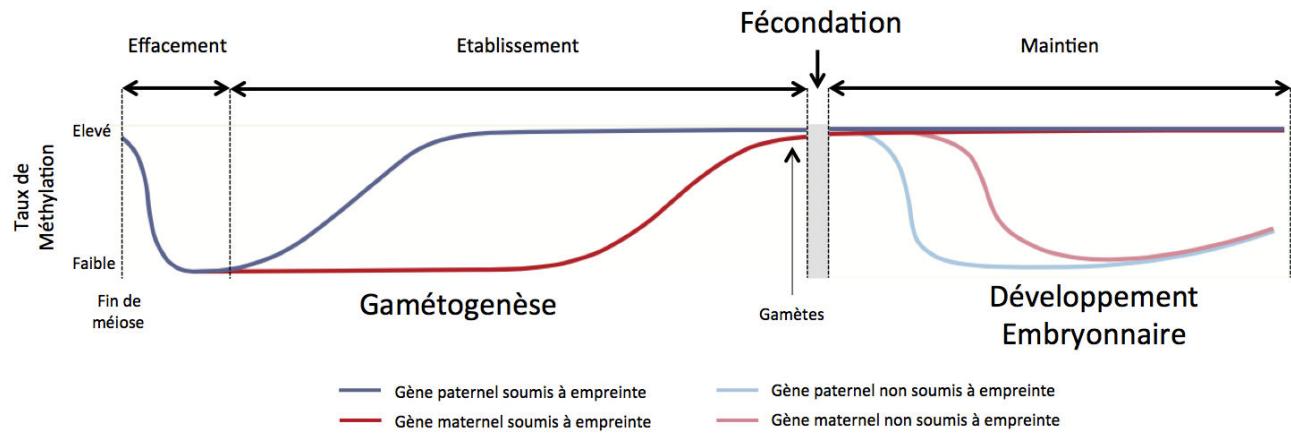


Figure 20. Evolution du taux de méthylation de l'ADN parental, des gamètes à l'embryon (adapté de Dandolo *et al.*).²⁸

Notre environnement peut influer sur notre génome, plus particulièrement sur notre épigénome, tout au long de notre vie. Cependant, c'est au cours de notre développement qu'il est le plus influent car il peut avoir un impact sur l'empreinte génomique parentale et les épimutations qu'il provoque peuvent être transmises à notre descendance [cf. ANNEXES S2].

ii. Influence de l'environnement sur le développement

cf. ANNEXES S2

2. Transmissions de la méthylation de l'ADN et couple DNMT1 / UHRF1(SRA)

i. Les DNMTs

La méthylation de l'ADN est catalysée par les différentes DNMTs évoquées précédemment. Ces enzymes possèdent toutes un rôle bien particulier. Par exemple, DNMT1 permet le maintien du profil de méthylation après réPLICATION du matériel génétique [cf. Figure 16].^{18, 41}

DNMT3a et DNMT3b sont chargées d'introduire les groupes méthyles sur des portions d'ADN naturellement non méthylées, comme dans le cas des cellules germinales ou lors de phénomènes pathologiques ; on parle alors de méthylation *de novo*. Par exemple, DNMT3a catalyse la méthylation d'ADNs non-méthylés avec 3 fois plus d'efficacité que dans le cas d'ADN hémiméthylés.¹⁹ Elle semble également pouvoir activer TDG, l'enzyme d'excision de 5fC et 5caC, qui en retour semble inhiber son activité méthyltransférase.^{25b} DNMT3L est une enzyme de régulation. Grâce à sa chaîne C-terminale, elle peut réguler l'activité des deux autres DNMT3.⁴² A noter également que DNMT1 peut, dans une moindre mesure, réaliser la méthylation *de novo*.⁴³

Enfin, l'enzyme DNMT2 présente une activité méthyltransférase plus faible que celle des autres DNMTs. Son rôle au niveau de l'ADN semble très limité, voire inexistant. En effet, l'inactivation du gène codant pour DNMT2 chez des souris n'entraîne aucune

variation du taux global de méthylation de leur ADN. Des études démontrent plutôt un rôle dans la méthylation d'ARNr, comme celui de l'ARNr-aspartyl (cytosine 38), ou encore dans l'activation des siARN.^{19, 25b}

ii. DNMT1

DNMT1 est une enzyme de 183 kDa, constituée de 1 616 acides aminés. Elle est omniprésente dans notre organisme ; fortement exprimée dans les tissus fœtaux, le cœur, les reins, le placenta, et moins exprimée dans la rate, les poumons, le cerveau, l'intestin grêle, le côlon, le foie et les muscles squelettiques. Quatre isoformes peuvent être observées : DNMT1a (la plus abondante) et DNMT1b sont naturellement présentes dans notre organisme ; DNTM1o et DNMTp ne sont exprimées que dans le cas des ovocytes et lors de la spermatogenèse, respectivement.^{25b, 42}

Comme nous l'avons déjà évoqué, DNMT1 est impliquée dans le maintien du profil de méthylation de l'ADN. En effet, elle est très peu abondante lorsque les cellules sont en quiescence (G0) [cf. Figure 8] et n'est induite que lors de l'entrée en phase S (réPLICATION de l'ADN). Plus particulièrement, son domaine RFTD (*Replication Foci Targetting Domain*) permet son recrutement au niveau de la fourche de réPLICATION.⁴⁴ De plus, une déplétion en DNMT1 n'entraîne aucune variation du taux de méthylation *de novo*, indiquant donc qu'elle n'intervient que faiblement dans ce processus mais bel et bien dans le maintien du profil de méthylation.^{25b, 42, 45}

Elle présente une affinité sélective pour l'ADN hémi-méthylé. Son activité méthyltransférase est en effet 30 à 40 fois plus importante pour les duplexes hémi-méthylés que pour les non-méthylés.^{25b} En fait, DNMT1 reconnaît la 5mC du brin complémentaire (rouge sur la Figure 21A) en positionnant son groupe méthyle à l'intérieur d'une poche hydrophobe se trouvant sur la boucle TRD [cf. Figure 21A]. Le duplexe d'ADN est maintenu en place grâce à cette boucle TRD et à une deuxième boucle : la boucle catalytique chez les souris [cf. Figure 21A] et la boucle CXXC chez l'homme. Dans ces deux cas, TRD et la deuxième boucle se positionnent respectivement au niveau des grands et petits sillons de l'ADN.

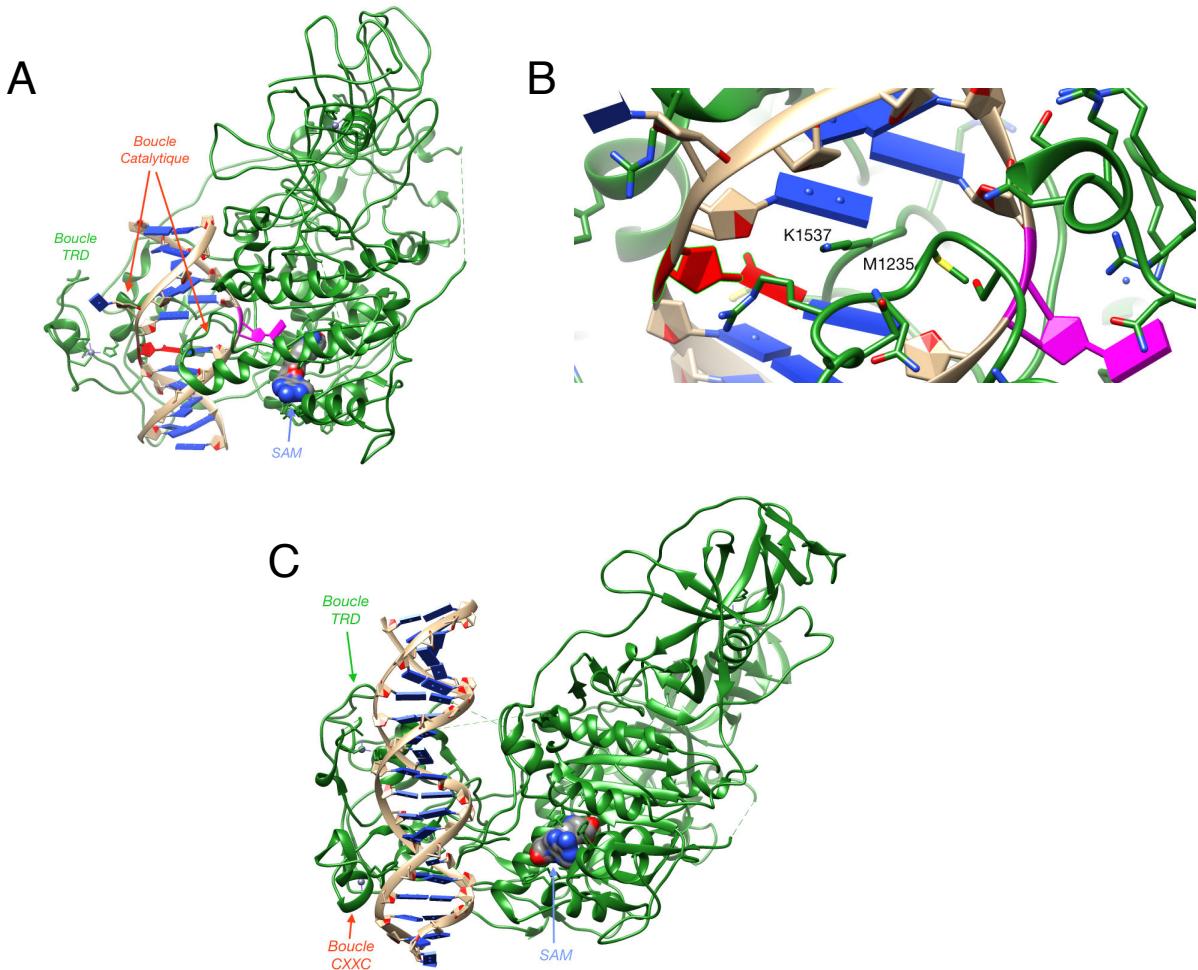


Figure 21. (A) Structure RX de la seule DNMT1 murine, co-cristallisée en présence d'un ADN hémiméthylé (PDB : [4DA4](#)), (B) Zoom sur le site de stabilisation de l'ouverture du duplexe et (C) Structure RX de la DNMT1 humaine liée à un double brin d'ADN non méthylé (PDB : [3PTA](#)).⁴⁶

Dans le cas de la méthylase murine (la seule ayant été co-cristallisée avec un ADN hémiméthylé), la cytosine à méthyler [cf. Figure 21A et B - en rose] est basculée (*flippée*) hors du duplexe d'ADN et se positionne parfaitement dans sa poche catalytique, au plus près de SAM [cf. Figure 23A]. Afin de stabiliser la position laissée vacante par le basculement (*flip*) de la base, les Met 1235 (boucle catalytique) et Lys 1537 (boucle TRD) viennent occulure l'ouverture du duplexe [cf. Figure 21B].⁴⁶ Ce mécanisme semble être identique chez toutes les enzymes homologues à DNMT1 puisque également observé dans le cas des méthyltransférases (M.Mpel) de champignons microscopiques (*Mycoplasma penetrans*).⁴⁷

Le mécanisme standard de formation de la 5mC [cf. Figure 22] commence tout d'abord par une addition de Michael acido-catalysée par un résidu acide carboxylique d'un acide glutamique du site actif de l'enzyme. Elle favorise l'attaque nucléophile du thiol de la cystéine 1229 de l'enzyme sur la position 6 de la cytosine.^{46a} Ensuite, l'attaque nucléophile du C5 de l'énamine de la nucléobase, sur le donneur de méthyle SAM, permet le transfert du groupe méthyle sur la base, tout en libérant SAH. Cette étape bénéficie d'une catalyse

basique, grâce à l'action du résidu carboxylate de l'acide glutamatique, qui augmente le caractère nucléophile de la base nucléique. Enfin, la β -élimination de la cystéine restaure l'aromaticité de la nucléobase et permet la dissociation du produit méthylé du site actif de l'enzyme.^{8, 42}

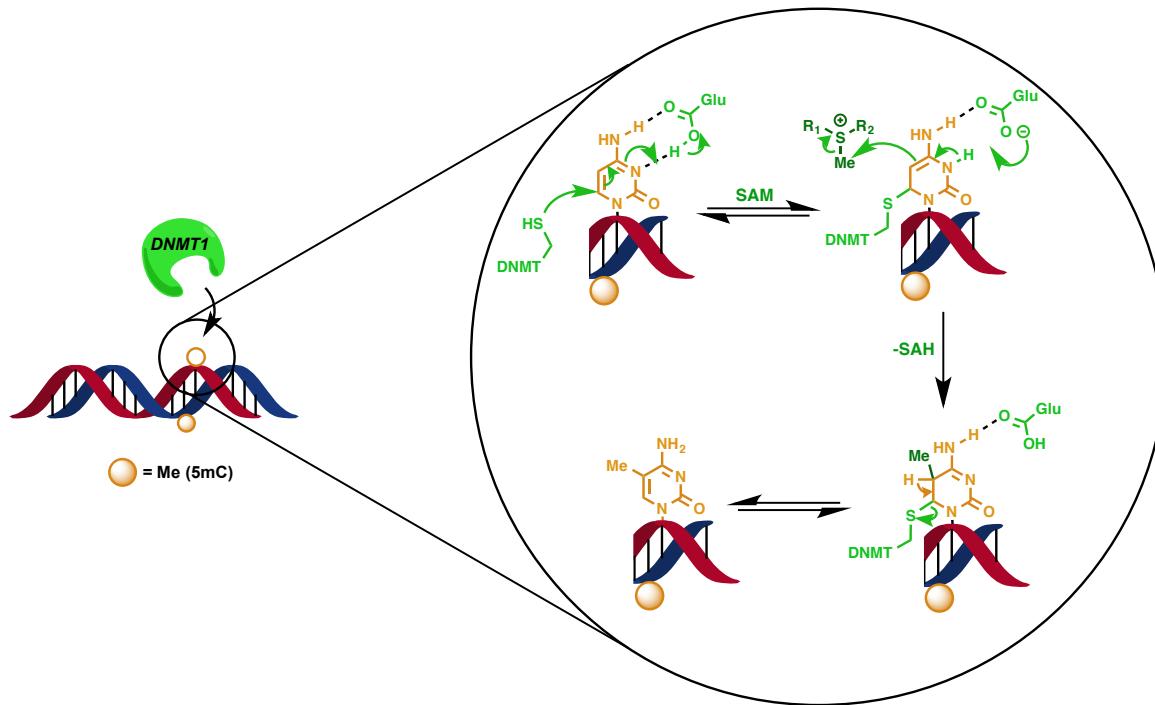


Figure 22. Mécanisme de méthylation impliquant DNMT1 et SAM.

Une fois la méthylation achevée, la cytosine reprend sa place au sein du duplexe et DNMT1 relâche son étreinte.

iii. UHRF1 / SRA

Le maintien du profil de méthylation assuré par DNMT1 se fait avec une très haute fidélité. Cependant, sa préférence pour l'ADN hémi-méthylé ne suffit pas, à elle seule, à expliquer cette fidélité. Elle repose en fait sur l'interaction de DNMT1 avec une autre protéine : UHRF1 (*Ubiquitin-like, containing PHD and RING finger domains 1* - appelée aussi ICBP90 chez l'homme ou Np95 chez la souris). Cette protéine possède un poids moléculaire apparent de 90 kDa (homme) avec cinq domaines bien distincts : le domaine UBL (pour *Ubiquitin-like*), le domaine TTD (pour *Cryptic Tandem Tudor Domain*), le domaine PHD (pour *Plant Homeodomain*), le domaine SRA (pour *Set and Ring Associated*) et le domaine RING (pour *Really Interesting New Gene*).⁴⁸

Le rôle qu'elle joue sur la conservation du profil de méthylation ne fait aucun doute. En effet, dans le cas où les cellules sont dépourvues de UHRF1, ce maintien est compromis. Elle a la capacité d'interagir avec DNMT1 - ainsi qu'avec les deux DNMT3a et b - afin de les guider sur le site à méthyler. A l'instar de DNMT1, elle reconnaît elle aussi l'ADN hémi-méthylé grâce à son domaine SRA [cf. Figure 23]. Plus particulièrement, le domaine SRA (210 résidus) reconnaît la 5mC du brin parent [cf. Figure 23, en rouge] et la fait basculer hors du duplexe d'ADN afin de la positionner dans son site de reconnaissance, un

tonneau β ainsi que l'ont montré différentes structures obtenues par diffraction des rayons X.^{42, 48, 49}

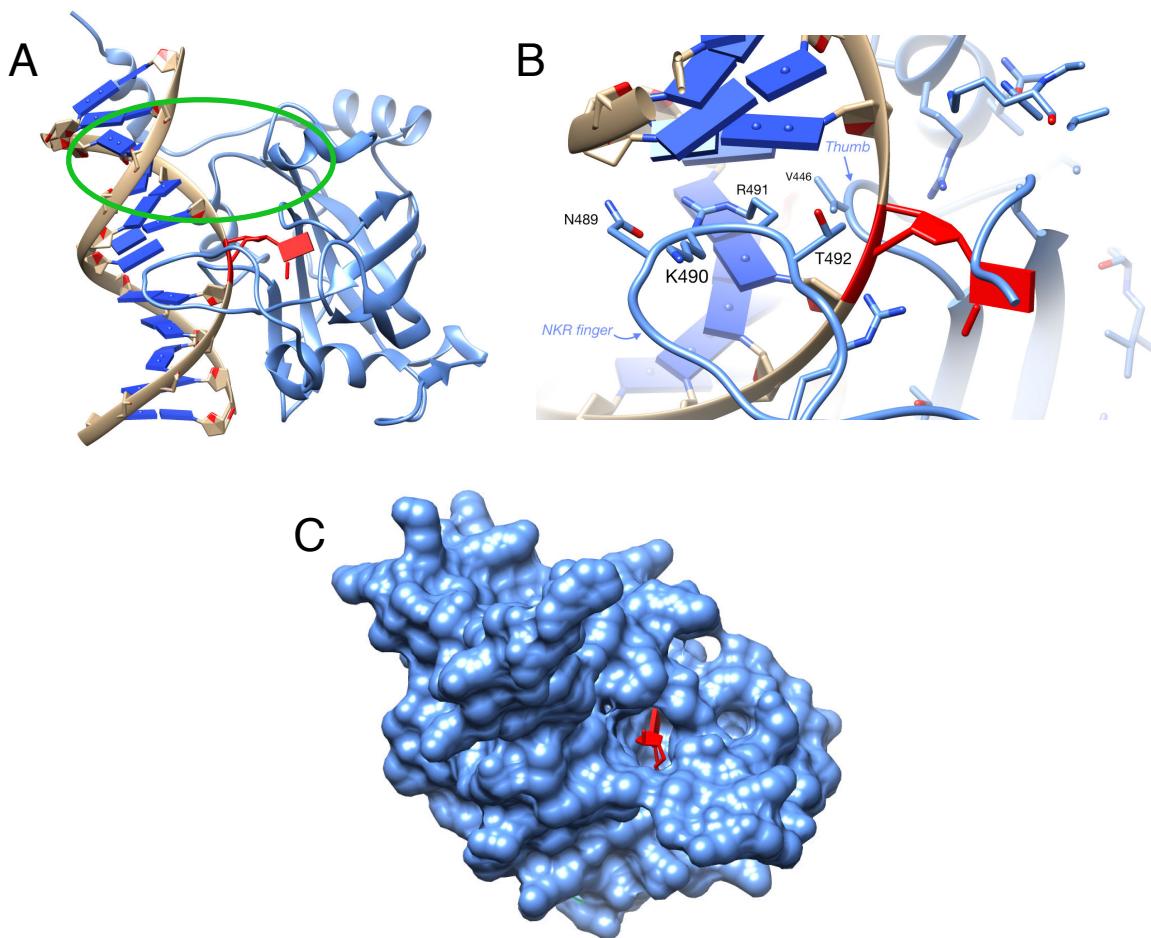


Figure 23. (A) Structure RX de SRA liée à un double brin d'ADN hémi-méthylé (PDB : [3CLZ](#)) ; (B) Zoom sur le site de stabilisation de l'ouverture du duplex et (C) Poche de fixation de la 5mC.^{48a}

Ce basculement est couramment rencontré chez les enzymes telles que les DNMTs (cf. I.2.2.ii) et les enzymes de réparation de l'ADN. Cependant, dans le cas de SRA il s'agit du premier exemple d'un domaine protéique, dépourvu d'activité catalytique, dont la liaison à l'ADN implique le basculement d'une base nucléique de la sorte. Dans le cas des enzymes, ce mécanisme de basculement permet un accès optimal à la cible afin de pouvoir la modifier. Dans le cas de SRA, il permet probablement un ancrage spécifique de UHRF1 sur l'ADN en évitant toute diffusion linéaire le long du duplex.⁵⁰

L'interaction avec l'ADN est renforcée par la présence de deux boucles qui, comme dans le cas de DNMT1, viennent maintenir le duplex en place : le *NKR finger* et le *thumb* [cf. Figure 23A et B]. Elles agissent comme l'index et le pouce d'une main, pendant que la 5mC se positionne au coeur du tonneau β .⁴⁸

Le *NKR finger* et le *thumb* interagissent respectivement avec le grand sillon et le petit sillon de l'ADN. Les acides aminés du *NKR finger* qui sont à proximité du duplex (Asn 489, Lys 490, Arg 491 et Thr 492) viennent stabiliser, grâce à des liaisons hydrogènes et du *stacking*, la position laissée vacante par le *basculement* de la 5mC [cf. Figure 23B]. En

plus de cela, la valine 446 du *thumb* vient effectuer une interaction de type Van der Walls, au travers de l'ouverture formée à l'intérieur du duplexe, avec la partie aliphatique de Arg 491 du *NKR finger*, pour stabiliser encore plus le système [cf. Figure 23B].⁴⁸

Ainsi UHRF1 permettrait de recruter DNMT1 et d'augmenter la fidélité dans la reproduction du patron de méthylation de l'ADN. Sur la base des résultats des différentes études réalisées, un mécanisme en quatre étapes a été proposé afin d'expliquer le rôle de UHRF1 [cf. Figure 24].^{48b} Par le biais de son domaine SRA, UHRF1 se lierait dans un premier temps à l'ADN, puis reconnaîtrait la 5mC du brin parent et la ferait basculer de sa position intra-hélice vers le site de liaison de la protéine. Dans une seconde étape, le complexe SRA / ADN permettrait de recruter DNMT1, grâce à l'interaction d'un second site de liaison, conduisant à un complexe ternaire transitoire. Dans une troisième étape, cette liaison de DNMT1 à UHRF1 déstabiliserait le complexe SRA / ADN, libérant l'ADN héméthylé, et facilitant l'accès à DNMT1. La liaison à l'ADN conduirait alors au basculement et à la liaison de la cytosine à méthyler au site actif de l'enzyme. Dans une dernière étape, une fois la méthylation réalisée, DNMT1 se dissocierait de l'ADN avec repositionnement à l'intérieur du duplexe de la 5mC nouvellement synthétisée.⁴² Cependant la façon dont UHRF1 et DNMT1 communiquent entre elles pour conduire à la méthylation de l'ADN et son mécanisme dans ces différentes étapes à l'échelle moléculaire restent encore à élucider.

Enfin le mécanisme serait plus complexe car d'autres partenaires protéiques seraient également impliqués. En particulier, le « *sliding clamp* » PCNA (*Proliferating Cell Nuclear Antigen*) permettrait d'augmenter la processivité du couple UHRF1 / DMT1.⁵¹

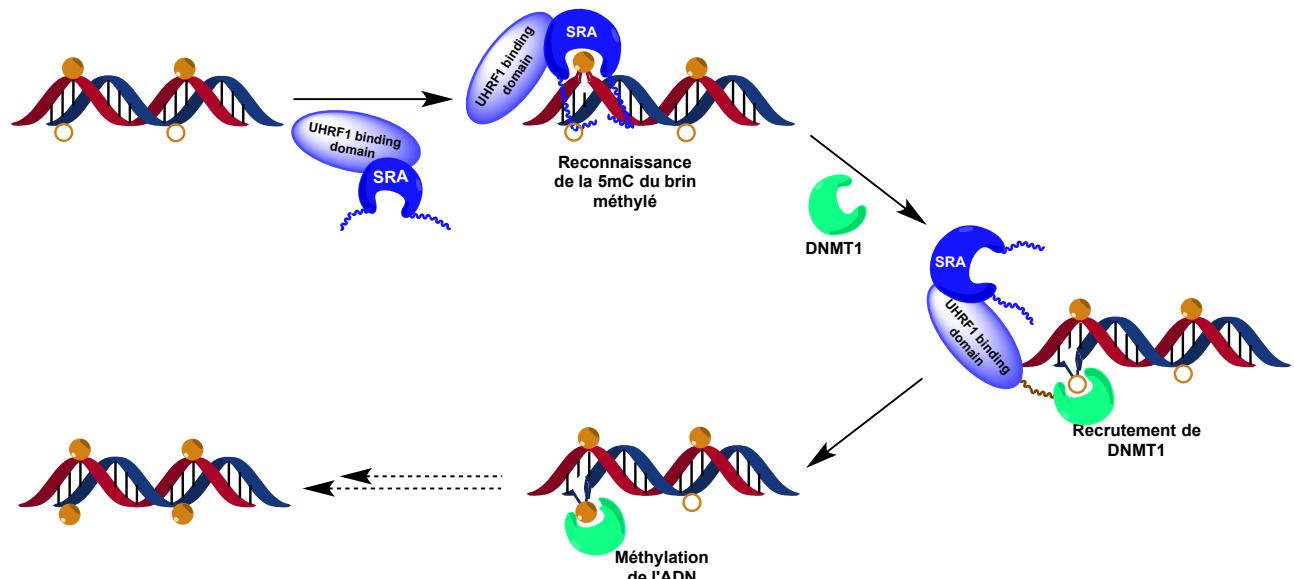


Figure 24. Illustration du mécanisme concerté supposé entre UHRF1(SRA) et DNMT1 (adapté de Arita *et al.*).^{48b}

C. Implication dans le cancer

Le cancer est une pathologie qui résulte de la perte de contrôle, par notre organisme, du cycle de vie de certaines cellules. Suite à une mutation ou une épimutation, ces cellules deviennent insensibles à l'ensemble des signaux visant à les réguler, comme les signaux mitogènes ou apoptotiques. Elles acquièrent alors la capacité à se multiplier de façon incontrôlée et infinie (anarchique) au sein du tissu où elles se trouvent, en d'autres termes, elles deviennent immortelles. Leur division incontrôlée aboutit finalement à la formation de tumeurs. Les tumeurs bénignes (verrues, grains de beauté) sont, comme leur nom l'indique, sans gravité. En revanche, les tumeurs malignes (ou cancers), elles, envahissent progressivement les tissus avoisinants en les détruisant. Afin de mieux proliférer, elle peuvent également déclencher l'angiogenèse ou même se répandre dans l'organisme sous forme de métastases. Il existe approximativement 200 types de cancer qui peuvent toucher tous les tissus du corps.^{52b}

Les cancers ne sont causés qu'à hauteur de 5-10 % par des anomalies génétiques alors que 90-95 % sont provoqués par notre mode de vie et notre environnement (tabac, alcool, pollution, infections, stress, obésité, ...).⁵³

1. Méthylation de l'ADN et cancers

Durant de nombreuses années, les chercheurs du monde entier ont considéré que la première cause à l'origine de la carcinogenèse était l'accumulation de mutations génétiques. Cependant, les différentes découvertes relatives aux marqueurs épigénétiques, tendent à prouver que des dérégulations de l'épigénome seraient à l'origine du développement de cancers.

Le profil de méthylation des régions promotrices de gène influence grandement l'expression des gènes. Dans la plupart des cas, à l'exception des gènes soumis à empreintes, les îlots CpG de ces régions promotrices (environ 55 %) sont non méthylés, contrairement aux dinucléotides CpG présents sur le reste du génome [cf. Figure 25A]. Les gènes liés à ces promoteurs sont naturellement transcrits. Cependant, dans le cas de cellules cancéreuses, certains gènes subissent des altérations de leur profil de méthylation et des phénomènes d'hyper- et d'hypométhylation sont alors observés [cf. Figure 25A & B]. L'expression génétique s'en trouve alors altérée.^{23, 24, 25, 26, 37b, 54}

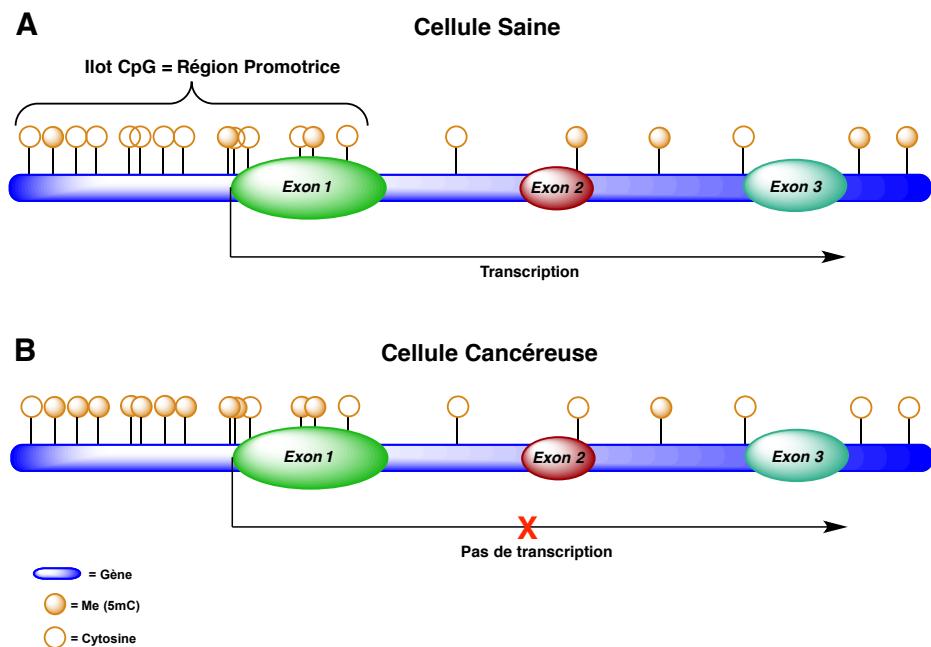


Figure 25. Illustration du profil de méthylation des îlots CpG (promoteurs de gène) dans le cas de (A) cellules saines et (B) cancéreuses.^{26b, 37b, 54b}

Les cellules cancéreuses présentent une hyperméthylation locale, essentiellement au niveau des îlots CpG (78 % des cas) [cf. Figure 25B]. Les gènes touchés sont impliqués dans la suppression de tumeurs (MLH1, CDKN2a,...), la réparation de l'ADN (BRCA1, MGMT) et la détoxification carcinogénique (GSTP1). Mais d'autres gènes, indispensables au cycle cellulaire ($p16^{INK4a}$, $p15^{INK4a}$, Rb, $p^{14}ARF$), à l'apoptose (DAPK, TMS1), ou encore à des phénomènes de résistance, de différenciation, d'angiogenèse et de métastase (E-cadherin), peuvent également être touchés.⁴³ Au total, ce ne sont pas moins de 500 promoteurs qui peuvent être méthylés dans le cas des cancers.⁵⁵

Cette hyperméthylation de la région promotrice [cf. Figure 25B] conduit à réprimer les gènes touchés, soit par réarrangement de la chromatine, sous l'action de protéines reconnaissant la 5mC (protéines MDB ou à doigts de zinc : ZBTB4, ZBTB38), soit par gène stérique.^{52b} Ne pouvant donc plus être exprimés, ces mécanismes de sécurité ne permettent plus de « contenir » les tumeurs en régulant le cycle cellulaire (division, apoptose), et la cellule devient immortelle / cancéreuse.^{26a, 52b}

Ces cellules cancéreuses présentent également une hypométhylation globale de leur ADN. Une perte de 20 à 60 % de CpG méthylées est, en effet, observée. Cette hypométhylation se retrouve dans bon nombre de cancers (sein, ORL, cerveau, ...) et tend à s'accroître au cours de l'évolution de la tumeur, de bénigne à maligne. Elle peut se produire au niveau des promoteurs de gène, du corps du gène ou sur des séquences répétées de l'ADN (50 % du génome).⁵⁶ Elle joue un rôle sur l'activation des oncogènes, naturellement réprimés (c-MYC, H-RAS, ...) ; sur le positionnement intranucléaire de la chromatine ; sur l'activation de retro-transposons, qui peuvent s'insérer dans les gènes afin de perturber leur expression ; ou encore sur l'instabilité chromosomique.^{52b} De plus, l'hypométhylation de certaines portions mineures et répétées de l'ADN peut

occasionnellement induire la transcription d'éléments rétroviraux endogènes, possédant un fort pouvoir oncogène. Enfin, les cancers peuvent également, dans certains cas, se développer sans hyperméthylation de leur gène ; une hypométhylation globale extrême, semble suffire.^{19, 26a, 48b, 56, 57}

Ces épimutations sont extrêmement variées et dépendent du patient et du type de cancer. Par exemple, le gène GSTP1 est hypermethylé dans 90 % des cancers de la prostate mais est non-méthylé dans le cas de la leucémie myéloïde aiguë.⁴³ Il a été démontré que, pour un grand nombre de cancers, ces altérations du profil de méthylation pouvaient être liées à une hyperactivité et/ou une surexpression de DNMT1^{58, 59} et de UHRF1.⁶⁰ Leurs rôles à toutes les deux vont donc être passés en revue dans les prochains paragraphes.

2. DNMT1 / UHRF1 et cancers

L'enzyme de maintien du profil de méthylation de l'ADN, DNMT1, semble jouer un rôle majeur dans la genèse des cancers. En effet, de nombreuses études tendent à montrer qu'elle est surexprimée dans bon nombre de cancers. Dans le cas du cancer du colon, elle présente une activité réduite : 5% en comparaison avec les cellules saines. Cependant, à cause de sa surexpression, l'impact de cette inactivation sur le profil de méthylation est minime.⁵⁸

Il a aussi été démontré qu'une surexpression en DNMT1 au sein de cellules NIH 3T3 de souris, entraînait l'apparition de tumeur gastro-intestinale.⁵⁸ En cas d'inactivation, le gène p21^{WAF1/CIP1} et le rétablissement de la mort cellulaire induite par p53.⁶¹ En ce qui concerne les cancers du rein et de la vessie, une déplétion en DNMT1 permet de réactiver les gènes RASSF1 et CDKN2A, respectivement.^{45, 62} Enfin, l'équipe de *Barton* a de son côté mis en avant le fait qu'une plus grande activité de DNMT1 était observée dans le cas du cancer colorectal et qu'ici, elle provenait probablement plus d'une hyperactivité que d'une surexpression de l'enzyme.⁵⁸

Cette surexpression n'est pas le seul facteur, à l'origine de la carcinogenèse. En effet, sa coopération avec les DNMT3a et b est amplifiée dans le cas de cellules cancéreuses. Prenons l'exemple des leucémies où DNMT1 et DNMT3a sont toutes les deux surexprimées et coopèrent dans le but d'hyperméthyler des îlots CpG.⁴² Citons également le cas du lymphome de Burkitt, où DNMT1 et DNMT3b, cette fois-ci, sont surexprimées. Cette étude de *Klumb* montre d'ailleurs que DNMT1 est essentielle à la prolifération cellulaire et à l'entretien du phénotype tumoral de ce lymphome.⁶³

De plus, de nombreuses études montrent que UHRF1, protéine responsable du recrutement de DNMT1, est surexprimée dans le cas de ces cancers (sein, rein, vessie, prostate, poumon, utérus, astrocytome, ...).^{60, 64} Elle semble avoir une influence sur le développement des cancers : prolifération des cellules cancéreuses et déclenchement de métastases.^{64c}

La déplétion en p53, inactif dans près de 50 % des cancers, pourrait être à l'origine de sa surexpression. En effet, dans le cas de cellules saines, ce gène suppresseur de tumeur régule l'expression de UHRF1 de manière indirecte, en désactivant le facteur de transcription E2F-1.⁶⁵

UHRF1 a par exemple été impliquée dans l'apparition d'une excroissance du foie chez le poisson zèbre. Sa surexpression déclenche une hypométhylation globale de l'ADN car elle empêche la parfaite localisation et stabilisation de DNMT1. Ceci aboutit finalement au contournement de la sénescence, normalement déclenchée par p53, et donc à la tumorigénèse (Carcinome Hépatocellulaire).^{59, 65, 66} Dans le cas du cancer des ovaires (lignées cellulaires SKOV-3 et OVCAR-3), l'utilisation de siARN, permettant de réguler l'expression de UHRF1, a permis de stopper la croissance des cellules cancéreuses et a déclenché leur apoptose.^{64a}

En plus de jouer un rôle important dans la carcinogenèse, UHRF1 semble également impliquée dans la toxoplasmose.⁶⁵

Enfin, il a également été démontré que la perturbation du complexe DNMT1 / PCNA / UHRF1 induit une hypométhylation de l'ADN. Cette perturbation agit comme un oncogène et provoque, plus particulièrement, l'apparition de tumeur du cerveau (astrocytes), du sein, du poumon et des cellules mésothéliales.⁶⁷

Ces entités protéiques sont indispensables au développement et à la survie des individus car elles permettent de maintenir un méthylome sain dans nos cellules. Cependant, à cause des effets de notre environnement, de notre style de vie plus particulièrement, l'activité et l'expression de ces protéines sont perturbées. Elles commencent à altérer le profil de méthylation de notre génome et induisent ainsi l'apparition de tumeurs.

Une meilleure compréhension du mécanisme de méthylation de l'ADN et des dérégulations qui l'affectent est indispensable d'un point de vue clinique. De telles informations permettraient de lutter plus efficacement contre les cancers, en développant de nouveaux principes actifs plus sélectifs, plus efficaces et moins toxiques.

Au total, l'épigénétique est basée sur au moins quatre paramètres clés. Le premier concerne la réversibilité des marques épigénétiques qui ouvrent le champ thérapeutique. Le second, affirme que les marques épigénétiques s'accumulent au fil de l'existence sous l'influence de l'environnement. Leur accumulation pourrait être prévenue par la nutrition pro-épigénétique et l'évitement de certains toxiques à effets épigénétiques. Le troisième, constate que la transmissibilité des marques épigénétiques à la descendance est un risque trop important pour être négligé. Le dernier paramètre met en avant le fait que l'épigénome n'est pas encodé dans le génome. Sa mise en évidence, sa mesure, sa détection nécessitent des outils spécifiques auxquels nous apportons notre contribution en développant des sondes fluorescentes adaptées.

De part sa sensibilité et son côté non destructif, la fluorescence est la technique de choix afin de pouvoir suivre l'interaction de ces entités à l'échelle moléculaire.

II. Fluorescence

Le terme luminescence vient du latin *lumen* (ou *luminis*) qui signifie lumière. Il fut pour la première fois introduit en 1888 par le physicien allemand *G. H. Wiedemann*.⁶⁸ Il s'agit d'une émission de lumière (rayonnement électromagnétique) se produisant lors du retour à l'état fondamental d'une substance se trouvant dans un état électronique excité.

Afin d'atteindre cet état excité, la substance doit au préalable être soumise à une source d'excitation. Il est possible de distinguer différents types de phénomènes de luminescence classés en fonction du mode d'excitation mis en jeu :

- la bioluminescence, où une réaction biochimique est la source d'excitation (ex : luciférase),
- la chimiluminescence, où une réaction chimique intervient afin d'exciter le système (ex : luminol),⁶⁹
- la photoluminescence, où l'excitation provient d'une source lumineuse,
- la mécanoluminescence, où une source mécanique est à l'origine de l'excitation (ex : tube cathodique),
- la cristalloluminescence, où l'énergie d'excitation provient de procédé de cristallisation,
- la thermoluminescence, où un changement de température génère l'énergie d'excitation,
- l'électroluminescence, où l'application d'un champ électrique provoque l'excitation,
- la radioluminescence, obtenue après irradiation RX ou rayonnement α , β , γ .

Photoluminescence est un terme général qui englobe à la fois la fluorescence et la phosphorescence.

A. Rappels

La fluorescence est un phénomène se produisant après interaction de la matière avec de la lumière. Si l'on se réfère à la dualité onde-corpuscule, une onde électromagnétique (lumière) présente les propriétés d'un phénomène ondulatoire, caractérisé par sa longueur d'onde (λ , nm) et son nombre d'onde (v , cm^{-1}), et d'une particule (photon) ayant une certaine énergie (E , eV) [cf. Figure 26]. Cette énergie est directement proportionnelle au nombre d'onde et inversement proportionnelle à la longueur d'onde.

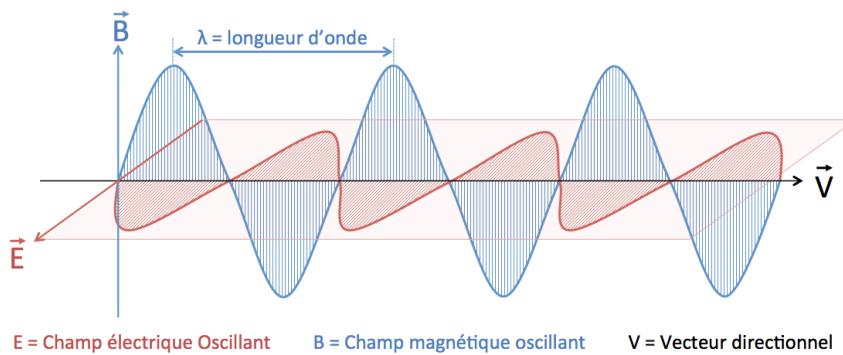


Figure 26. Spectre électromagnétique.

Une molécule capable d'absorber un photon est un chromophore. Une molécule qui est portée dans un état excité, après absorption de lumière, et qui revient spontanément à l'état fondamental, par émission de photons, suivant le processus de fluorescence, est un fluorophore.

Le spectre lumineux examiné en photochimie et photophysique s'étend généralement sur la gamme spectrale 100 nm (Ultraviolet - UV)-1000 nm (Infrarouge - IR) ; la portion visible s'étendant, elle, de 400 à 800 nm [cf. Figure 27].

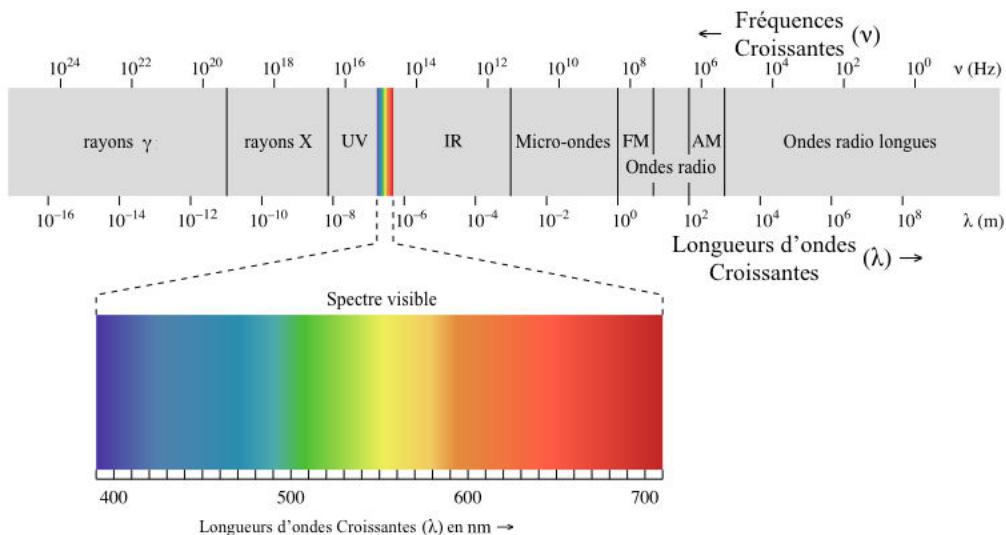


Figure 27. Spectre électromagnétique.

1. Transitions électroniques

Afin de mieux appréhender ces phénomènes d'absorption et d'émission de photons, intéressons nous au diagramme des orbitales moléculaires d'un composé fluorescent [cf. Figure 28]. A l'état fondamental S_0 , l'orbitale HOMO (*Highest Occupied Molecular Orbital*) contient deux électrons, respectivement de spin (S) +1/2 et -1/2, et la LUMO (*Lowest Unoccupied Molecular Orbital*) est vide. La multiplicité de spin M vaut 1 ($M = 2S+1$), on dit alors de cet état qu'il est singulet.

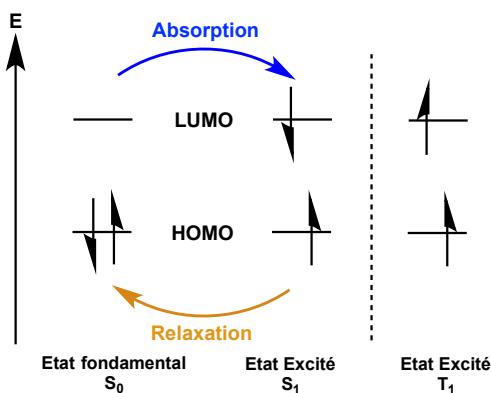


Figure 28. Diagramme électronique de l'état fondamental et des états excités singulet et triplet.

La différence d'énergie entre ces deux orbitales correspond, pour la majorité des composés organiques, à l'énergie de radiation UV - visible. Une exposition de cette molécule à des radiations UV proches (200-400 nm) ou à de la lumière visible (400-800 nm) peut alors provoquer l'absorption de photons et le transfert d'un électron de la HOMO vers la LUMO. Le passage d'un électron dans l'orbitale anti-liante est possible car extrêmement rapide ($\approx 10^{-15}$ s), beaucoup plus que les vibrations moléculaires (10^{-12} - 10^{-10} s). D'après le principe de Franck-Condon, « une transition électronique se produit sans changement des positions des noyaux dans l'entité moléculaire et son environnement ».⁶⁸ L'orientation de spin des électrons est conservée (+1/2 et -1/2) et la multiplicité de spin de cet état excité, appelé S_1 , est identique à celle de l'état fondamental S_0 .

Le retour à l'état fondamental, la relaxation, peut se faire selon trois voies différentes [cf. Figure 28 et 29]. Tout d'abord, le retour à S_0 peut se faire sans émission de photon : l'énergie absorbée diffuse. Dans le cas de la fluorescence, la relaxation s'effectue par émission de photons. Enfin, un dernier processus entraîne une conversion de l'état excité singulier S_1 vers un état excité triplet T_1 , où les deux électrons présentent des spins parallèles ($M = 3$). Cette conversion est suivie par une émission de photons avec inversion de spin : c'est ce que l'on appelle la phosphorescence [cf. Figure 29].

2. Diagramme de Jablonski

Le diagramme de *Jablonski* permet d'approfondir les explications évoquées avec le modèle des orbitales moléculaires. Il permet plus particulièrement, une meilleure visualisation des phénomènes d'absorption et de relaxation que peuvent subir les molécules fluorescentes [cf. Figure 29].

Les transitions entre états électroniques sont représentées par des flèches verticales afin d'illustrer la nature instantanée de ces phénomènes. Les traits horizontaux représentent les sous-niveaux vibrationnels associés aux différents états électroniques (S_0 , S_1 et T_1). Ils sont indispensables à la réalisation des phénomènes d'absorption et de relaxation.⁶⁸

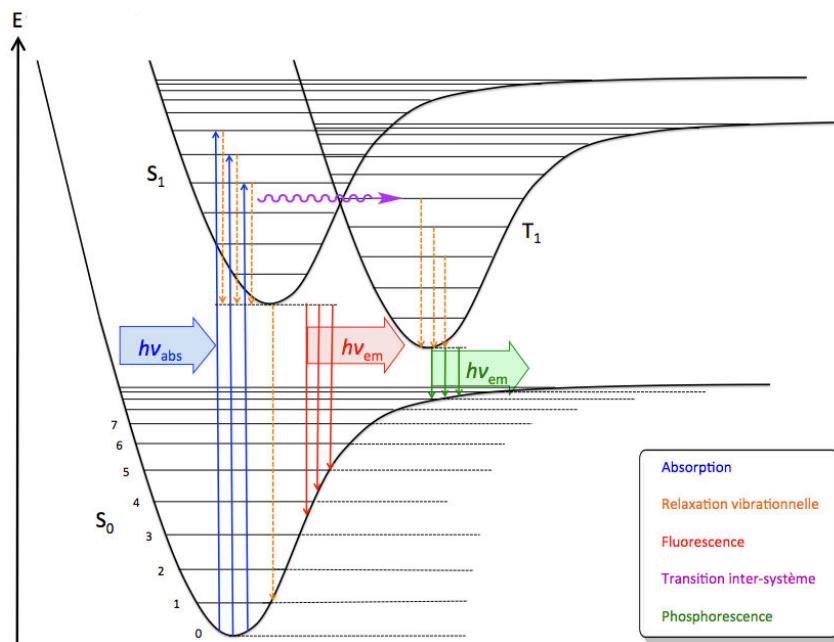


Figure 29. Diagramme de Jablonski.

A température ambiante, la plupart des molécules se trouvent dans le plus bas niveau vibrationnel de l'état électronique fondamental S_0 (loi de Boltzmann).⁶⁸

L'irradiation de la molécule par une onde électromagnétique provoque l'absorption de photons (*flèches verticales bleues* - $\approx 10^{-15}$ s), dans le sens $S_0 \rightarrow S_1$. Les électrons peuplent ainsi les différents niveaux vibrationnels de l'état électronique excité S_1 . L'ensemble des transitions possibles vers les différents niveaux excités (*flèches bleues*) déterminent le spectre d'absorption de la molécule. Après relaxation vibrationnelle (*flèches oranges* - 10^{-12} - 10^{-10} s), le composé se trouve dans son état excité de plus basse énergie.⁶⁸

Le retour de ce niveau S_1 relaxé à l'état fondamental S_0 , par émission de photons, traduit la fluorescence (*flèches verticales rouges*). Le temps de vie de l'état excité S_1 est de l'ordre de 10^{-10} - 10^{-7} s. En parallèle, l'énergie accumulée à l'état excité peut également être dissipée par transfert d'énergie vibrationnelle au milieu sous la forme de chaleur (collisions moléculaires) et/ou par d'autres processus non radiatifs.⁶⁸

La phosphorescence est le résultat d'émission de photons après transition inter-système (*flèche ondulée violette* - 10^{-10} - 10^{-8} s). Elle correspond à une transition non-radiative s'effectuant entre des états électroniques de multiplicités différentes ($S_1 \rightarrow T_1$). Une fois l'état triplet atteint et relaxé, le retour à l'état fondamental, par émission de phosphorescence, implique une inversion de spin de l'électron interdite ($T_1 \rightarrow S_0$), c'est pourquoi le temps de vie de l'état excité T_1 est de plusieurs ordres de grandeur supérieur à celui de S_1 (10^{-6} à 1 s).⁶⁸

3. Avantages de la fluorescence

La fluorescence et les techniques d'imagerie qui lui sont associées ont été mises à l'honneur à deux reprises au cours de ces dix dernières années par deux prix Nobel de

Chimie. *Osamu Shimomura, Martin Chalfie et Roger Y. Tsien* ont été récompensés en 2008 pour avoir mis en évidence la protéine GFP (*Green Fluorescent Protein*) et le développement de ses applications en particulier en imagerie cellulaire. Ce fut ensuite au tour d'*Eric Betzig, Stefan W. Hell et William E. Moerner* de recevoir cette distinction suprême en 2014, pour avoir réussi à améliorer considérablement la résolution spatiale et temporelle des procédés d'imagerie de fluorescence.⁷⁰

En plus de sa grande résolution spatiale et temporelle, la fluorescence présente aussi une sensibilité très importante. Les concentrations mises en jeu au cours de ces études sont extrêmement faibles (allant du μM au pM). Les instruments actuels permettent, également, la détection de photons émis par une seule molécule : on parle de spectroscopie de fluorescence en molécule unique (*single-molecule fluorescent spectroscopy*).⁷¹

De plus, la réponse est obtenue de manière extrêmement rapide. En effet, la transition $S_1 \rightarrow S_0$, et donc l'émission de photons, s'effectue avec une constante de vitesse de $10^{-10} - 10^{-8}$ s.

Enfin, cette technique est non-invasive : elle ne détruit pas l'échantillon étudié, dans l'intervalle de temps de l'analyse.

4. Facteurs jouant sur le signal de fluorescence

L'émission de fluorescence d'un fluorophore peut fluctuer sous l'influence de plusieurs facteurs. Les processus inter- et intramoléculaires, se déroulant après l'absorption, peuvent modifier l'état excité du fluorophore (rendement quantique, durée de vie, intensité de fluorescence - cf. II.A.5.i-iv & II.B.1). A chaque processus correspond une voie de désexcitation [cf. Figure 30]. Ces différentes voies permettent d'obtenir des informations capitales sur l'environnement immédiat de la molécule (cf. II.B.1.). Certains de ces processus seront présentés plus en détails dans les prochains paragraphes.^{68a}

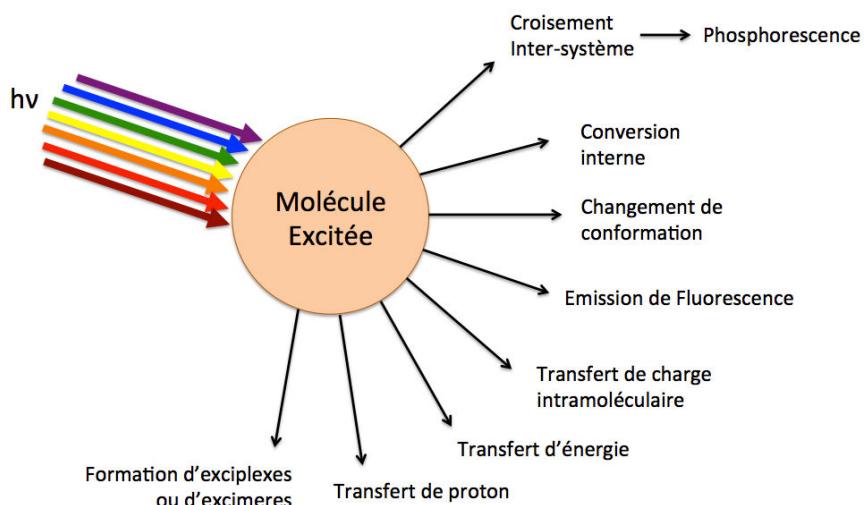


Figure 30. Processus suivant l'excitation d'un fluorophore.

Parmi ces processus :

- L'inhibition, ou *quenching*, est une perte du signal de fluorescence réversible suite à des interactions, à courte portée, avec des molécules environnantes (Q), y compris d'autres fluorophores.⁷²
- L'augmentation de température accroît l'agitation moléculaire et avec elle, le nombre de collisions entre molécules et les vibrations et rotations moléculaires. Les processus non-radiatifs deviennent alors beaucoup plus efficaces, au détriment de la fluorescence.⁶⁸
- Sur le même principe, l'augmentation de viscosité du milieu permet d'accroître le rendement quantique de fluorescence en figeant le système, limitant ainsi l'agitation moléculaire à l'origine de phénomènes de désexcitation non-radiative.
- A l'inverse des trois exemples précédents, le photoblanchissement, ou *photobleaching*, est irréversible. Il correspond à la destruction pure et simple du fluorophore à l'état excité. Il dépend de la stabilité du fluorophore dans son milieu, de la durée d'exposition et de l'intensité du faisceau incident.⁷¹

5. Grandeur caractéristiques de l'émission de fluorescence

Avant d'émettre des photons, tout fluorophore doit d'abord en absorber. Ainsi le maximum d'absorption et le coefficient d'absorption molaire ($L \cdot mol^{-1} \cdot cm^{-1}$), qui traduit la capacité du fluorophore à absorber des photons, sont les deux premières grandeurs spectroscopiques caractérisables.

Les autres grandeurs physiques liées aux sondes fluorescentes sont le maximum d'émission [cf. Figure 31], le déplacement de Stokes, le rendement quantique, la brillance et la durée de vie.

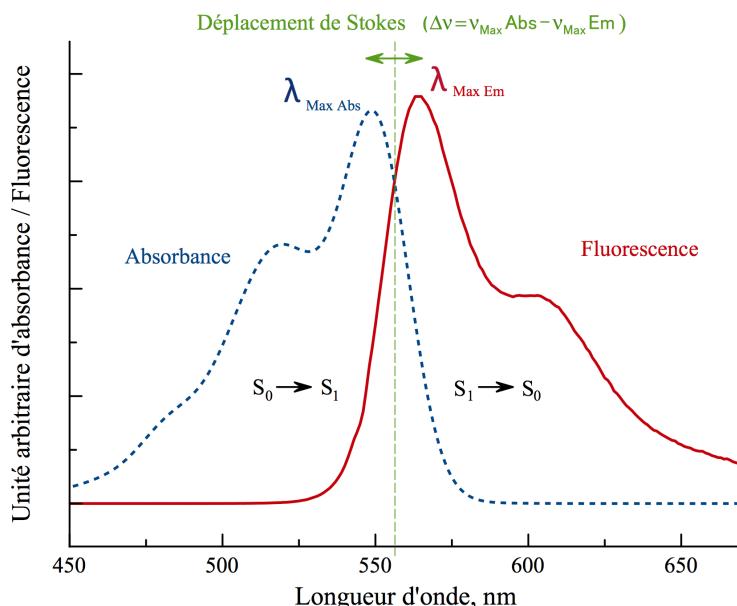


Figure 31. Illustration de spectres d'absorption et de fluorescence.

i. Déplacement de Stokes

Par rapport au spectre d'absorption, le spectre de fluorescence est décalé vers des longueurs d'ondes plus grandes (vers le rouge) [cf. Figure 31], en premier à cause de la perte d'énergie provoquée par les relaxations vibrationnelles [cf. Figure 29]. D'autres phénomènes peuvent augmenter cette différence pour certains type de fluorophores (comme le solvant). Cette différence entre les maxima d'absorption et d'émission s'appelle le déplacement de Stokes ($\Delta\nu$), ou *Stokes Shift (SS)* en anglais, et s'exprime en nombre d'onde [cf. Equation 1].⁶⁸

$$\Delta\nu = \nu_{Max}^{Abs} - \nu_{Max}^{Em} \quad (1)$$

Ce paramètre fournit des informations importantes sur les états excités. D'un point de vue pratique, la détection d'une espèce fluorescente est d'autant plus facile que le déplacement de Stokes est grand.

ii. Rendement Quantique

Le rendement quantique correspond au rapport entre le nombre de photons émis sur le nombre de photons absorbés ($0 \leq \Phi \leq 1$) [cf. Equation 2].⁶⁸

$$\Phi_F = \frac{N_{émis}}{N_{absorbés}} = \frac{k_r^S}{k_r^S + k_{nr}^S} \quad (2)$$

$N_{émis}$ est le nombre de photons émis, $N_{absorbés}$ correspond au nombre de photons absorbés, k_r^S est la constante de vitesse des phénomènes radiatifs et k_{nr}^S est la constante de vitesse de désexcitation non radiative.

Il est proportionnel à la durée de vie de l'état excité et est aussi fonction du rapport de la constante de vitesse radiative (k_r^S) sur la somme de toutes les vitesses (radiatives et non radiatives) (cf. II.A.5.iv.). Si la décroissance non-radiative est beaucoup moins importante que l'émission de fluorescence, le rendement quantique se rapproche de 1 (100 %).

iii. Brillance et Intensité de Fluorescence

La brillance (B_i) correspond au produit du coefficient d'absorption molaire (ε) du fluorophore et de son rendement quantique (Φ_F). Plus cette valeur est importante et plus le fluorophore est sensible [cf. Equation 3].⁷³

$$B_i = \varepsilon \times \phi_F \quad (3)$$

L'intensité de fluorescence représente, elle, la quantité de photons émis par unité de temps et de volume [cf. Equation 4]. Elle dépend de la brillance du fluorophore, de l'intensité du faisceau incident (I_0) ainsi que de la concentration de l'échantillon :

$$I_F = R \times \phi_F \times I_0 \times (c \times \varepsilon \times l) \quad (4)$$

I_F est l'intensité de fluorescence, R est une constante liée à l'appareil utilisé et I_0 l'intensité incidente (excitation).

Notons que l , la longueur de la cuve, mesure 1 cm.

Cette valeur étant directement proportionnelle à la concentration de l'échantillon (C), un fluorophore est d'autant plus sensible que de faibles concentrations en produits peuvent être détectées.

iv. Durée de vie

L'intensité de fluorescence est une grandeur exprimée par unité de temps. Elle décroît exponentiellement avec un temps caractéristique reflétant le temps de séjour moyen des molécules dans l'état excité S₁. Elle dépend donc d'un autre paramètre important : la durée de vie de l'état excité S₁ [cf. Equation 5].

$$\tau_s = \frac{1}{k_r^S + k_{nr}^S} \quad (5)$$

τ_s est le temps de vie de l'état excité (généralement de la picoseconde à quelques nanosecondes), k_r^S est la constante de vitesse des phénomènes radiatifs et k_{nr}^S correspond à la constante de vitesse de désexcitation non radiative.

Ce temps de déclin de fluorescence τ_s est une grandeur particulièrement importante pour un fluorophore. Il définit l'intervalle de temps pendant lequel un phénomène dynamique peut être observé (déclin de fluorescence) (cf. II.B.1.iii).^{68a}

6. Diversité des fluorophores

Il existe une grande variété de molécules fluorescentes. Dans la plupart des cas, il s'agit de composés aromatiques ou hétéroaromatiques présentant un système π étendu [cf. Figure 32].

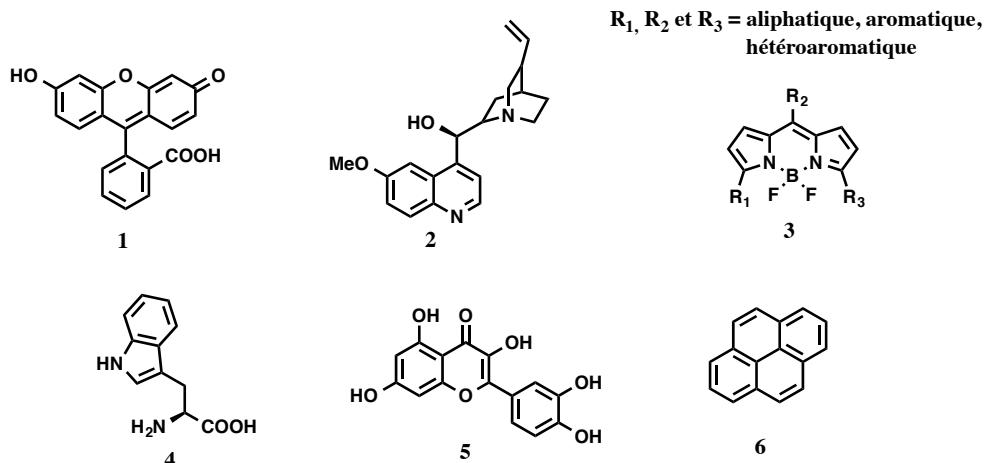


Figure 32. Quelques exemples de molécules fluorescentes.

La fluorescéine **1** fut synthétisée pour la première fois en 1871 par *Adolf Von Baeyer*. Elle et ses dérivés, dont la rhodamine, sont parmi les fluorophores les plus utilisés. En effet, ce composé de la famille des xanthènes présente des rendements quantiques élevés (même dans l'eau) ainsi qu'une forte solubilité en milieu aqueux (2 g/L).^{68b, 74}

Un autre fluorophore couramment rencontré est la quinine **2**. Elle fut étudiée pour la première fois par *Sir John Frederick William Herschel* en 1845. Elle présente une émission

de fluorescence bleue ($\lambda_{\text{Em}} = 450 \text{ nm}$). Elle est encore employée de nos jours comme produit de référence pour la détermination de rendements quantiques des nouveaux fluorophores (cf. II.A.5.ii.).^{68b}

Les BODIPY **3**, en fonctions des substituants (R_1 , R_2 et R_3) qu'ils portent, peuvent émettre sur une large gamme spectrale (510 - 763 nm).⁷⁵ Ils présentent des rendements quantiques élevés, proches de 1, des coefficients d'absorptions molaires de l'ordre de $80\,000 \text{ L.mol}^{-1}.\text{cm}^{-1}$, une grande sensibilité aux variations de polarité et de pH et des durées de vie élevées. Cependant, leurs déplacements de Stokes sont faibles.^{68b, 74}

Le tryptophane **4** est un des 20 acides aminés naturels présent dans nos protéines. Sa structure de type indole lui confère son caractère fluorescent ($\lambda_{\text{Abs}} = 280 \text{ nm}$; $\lambda_{\text{Em}} = 345 \text{ nm}$) et sa présence dans les assemblages protéiques en fait une sonde intrinsèque, permettant de les étudier, sans modifications préalables (cf. II.B.). Un décalage de sa bande de fluorescence permet de déterminer la nature structurée (décalage vers le bleu) ou non (décalage vers le rouge) d'une protéine, par exemple.^{68b}

La quercétine **5** est un flavonoïde naturel, produit par de nombreuses plantes. Cette molécule présente un large éventail d'activités biologiques⁷⁶ mais est également un fluorophore remarquable (cf. III. Les 3-hydroxychromones). Grâce à cette molécule, qui présente un déplacement de Stokes très important, *Kasha* décrivit pour la première fois le mécanisme de transfert de proton à l'état excité des 3-hydroxyflavones (ESIPT : *Excited State Intramolecular Proton Transfer* - cf. III.B.1.i).⁷⁷

Un grand nombre de composés polycycliques aromatiques sont utilisés comme fluorophores. Le pyrène **6**, par exemple, présente des rendements quantiques et une durée de vie importants (ex : 0,65 et 410 ns dans l'éthanol). Il est également extrêmement sensible aux changements de polarité du solvant.^{68b}

B. Sondes Fluorescentes pour le développement de capteurs

La fluorescence peut être utilisée pour deux raisons :

- Visualiser des entités biologiques ou moléculaires, et se soustraire ainsi au marquage radioactif. Une brillance très importante et un signal stable sont les premières qualités recherchées pour ce type de fluorophore.⁷⁸ La GFP (cf. II.A.3.), par exemple, est souvent employée pour observer l'expression de gènes par microscopie de fluorescence. Un autre exemple significatif, qui sera plus amplement présenté en Partie IV. Oligonucleotides, concerne le séquençage de l'ADN.
- Développer des capteurs. Un capteur est un dispositif sensible à un phénomène physique donné et transformant cette grandeur physique en signal exploitable. Pour un capteur fluorescent, l'interaction de l'analyte (la cible) avec le dispositif produit un complexe qui peut être détecté grâce au changement du signal de fluorescence et quantifié si la mesure est calibrée [cf. Figure 33].⁷⁸ Le dispositif peut-être de nature

extrêmement variée allant de la petite molécule organique à des assemblages supramoléculaires. Il en va de même de l'analyte qui va de l'atome à la cellule.

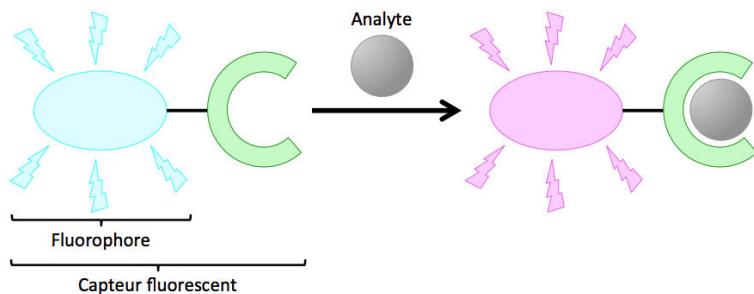


Figure 33. Représentation d'un capteur.

En raison de l'impact du milieu environnant sur l'émission de fluorescence, un grand nombre de capteurs sont aujourd'hui utilisés pour étudier des interactions biologiques et détecter et/ou doser différentes espèces chimiques (ions, protéines, ...). Les champs d'applications sont multiples et l'intensité de fluorescence, l'anisotropie ou le temps de vie sont par exemple les signaux de fluorescence qui peuvent être exploités.

Le développement d'un capteur nécessite l'emploi d'un groupe rapporteur fluorescent, ou sonde de fluorescence, pouvant être rangées en trois grandes classes :

- Les sondes intrinsèques (comme le tryptophane **4**), directement présentes sur la / les cible(s) protéiques à étudier.
- Les sondes extrinsèques liées à la cible à étudier. De nombreux procédés permettent de lier de manière covalente une sonde fluorescente sur une protéine ou des acides nucléiques (cf. IV.B.1.ii), par exemple,
- Les sondes extrinsèques qui s'associent à la cible via des interactions intermoléculaires (liaisons non covalentes). Certaines d'entre elles, interagissant avec l'ADN (DAPI, Hoechst, ...), seront présentées dans la Partie IV. Oligonucléotides (cf. IV.B.1.i).

Ici, nous allons plus particulièrement nous focaliser sur les paramètres de fluorescence qui peuvent être exploités à l'aide de ces sondes afin de développer un capteur, indépendamment de la classe à laquelle elles appartiennent.

1. Paramètres exploitables

- i. Mesure d'intensité à une ou deux longueurs d'onde
 - i) A une longueur d'onde (monobande)

Il s'agit de la technique de fluorescence la plus basique qui consiste à mesurer la variation de l'intensité de la bande d'émission. Les fluorophores concernés présentent un spectre de fluorescence avec un profil unique dont seule l'intensité varie en fonction des facteurs environnementaux. Ainsi, suivant les modifications environnantes (polarité, liaison avec une protéine, changement de conformation, viscosité, détection d'ions, ...), le spectre de fluorescence est soumis à des effets hypochromes (diminution de l'intensité) ou hyperchromes (augmentation de l'intensité). Dans le cas d'une détection par augmentation

de l'intensité de fluorescence, on dit qu'il s'agit d'une sonde « *Turn Off/On* » et dans le cas contraire, d'une sonde « *Turn On/Off* ». Ces effets peuvent être dus à des modifications structurales se produisant au niveau du senseur lui-même au moment de la détection, ou à des interactions intermoléculaires avec un inhibiteur de fluorescence, par exemple.⁷⁹

Cette façon de procéder présente cependant de nombreux inconvénients. En effet, l'intensité de fluorescence dépend de la concentration de l'échantillon et les conditions d'analyse doivent être parfaitement contrôlées et calibrées. La moindre erreur de préparation fausse toutes les mesures. De plus, des phénomènes parasites peuvent venir perturber l'analyse : l'intensité peut diminuer au cours du temps à cause de phénomènes extérieurs (photodégradation, présence d'inhibiteurs, fluctuation de température ou d'intensité de la lampe, ...). Ainsi, les fluctuations du signal de fluorescence observées ne sont pas nécessairement dues aux phénomènes étudiés mais peuvent provenir de ces facteurs parasites.

ii) A deux longueurs d'ondes (monobande ou duale)

Cette approche consiste à mesurer un rapport d'intensité de fluorescence à deux longueurs d'ondes. On dit donc de cette mesure qu'elle est ratiométrique. Elle peut tout d'abord être réalisée avec des dispositifs présentant un seul fluorophore dont le maximum de la bande d'émission varie (ESICT - *Excited State Intramolecular Charge Transfer*) ou capable de produire une émission duale grâce à des phénomènes intramoléculaires (ESIPT, TICT - *Twisted Intramolecular Charge Transfer*) ou intermoléculaires (transfert de proton par exemple). L'autre technique consiste à réaliser un double marquage soit avec deux sondes monobandes n'interagissant pas entre elles, dont une jouant le rôle d'étalement externe, soit avec deux sondes interagissant entre elles (FRET, excimère, exiplexe).

Les variations du rapport d'intensité de ces deux canaux d'émissions donnent des renseignements capitaux sur les changements environnementaux. Le fait de procéder ainsi permet de s'affranchir des contraintes propres à une détection monobande. En effet, le rapport d'intensité est indépendant de la concentration de l'échantillon, du photoblanchissement (si le marquage dépend d'un seul fluorophore, *vide infra*), des fluctuations d'intensité de la lampe, de la sensibilité du fluorimètre, ...^{68a}

a. Marquage Simple

- ESIPT

L'ESIPT sera plus amplement détaillé dans la Partie III - 3-hydroxyhromones (cf. III.B.1.i).

- ESICT

Les espèces soumises à un mécanisme ESICT présentent des groupements électrodonneurs (-NH₂, -NMe₂, -OMe, ...) et électroattracteurs (>C=O, -CN, ...) localisés de part et d'autre de la molécule, et reliés entre eux par un système insaturé. Ceci entraîne une polarisation électronique de la molécule. Lors de l'excitation du fluorophore, un mouvement électronique se produit d'une orbitale moléculaire vers une autre (cf. II.A. 1-2). Ainsi, dans le cas où le fluorophore est neutre, l'excitation de la molécule entraîne

une augmentation, souvent importante, de son moment dipolaire. Ce transfert de charge est favorisé et stabilisé en présence de solvants polaires. En effet, à l'état excité le donneur et l'accepteur d'électron deviennent plus forts. Ainsi, la polarisation électronique augmente de manière significative. Le grand moment dipolaire ainsi créé, interagit en particulier avec les dipôles des molécules de solvants polaires (stabilisation par le solvant), aboutissant à de forts déplacement de Stokes : l'état de transfert de charge intramoléculaire (ICT) relaxé est ainsi atteint [cf. Figure 34]. La bande de fluorescence se trouve alors élargie et décalée vers le rouge.^{68, 74} Les deux états, normal et ICT, étant en équilibre dynamique, un suivi ratiométrique des changements environnementaux peut être réalisé.^{68, 74}

Lorsque l'espèce est chargée à l'état fondamental, l'effet inverse est observé : les solvants polaires stabilisent plus l'état fondamental. Ainsi, la différence d'énergie entre les états fondamental et excité augmente dans les solvants polaires et l'on observe principalement un décalage du maximum d'absorption vers le bleu qui peut être suivi en fluorescence en mesurant la variation du maximum du spectre d'excitation.⁸⁰ Il faut cependant souligner que mesurer un rapport d'intensité à deux longueurs d'onde à partir du spectre d'excitation pour un suivi ratiométrique est moins pratique.

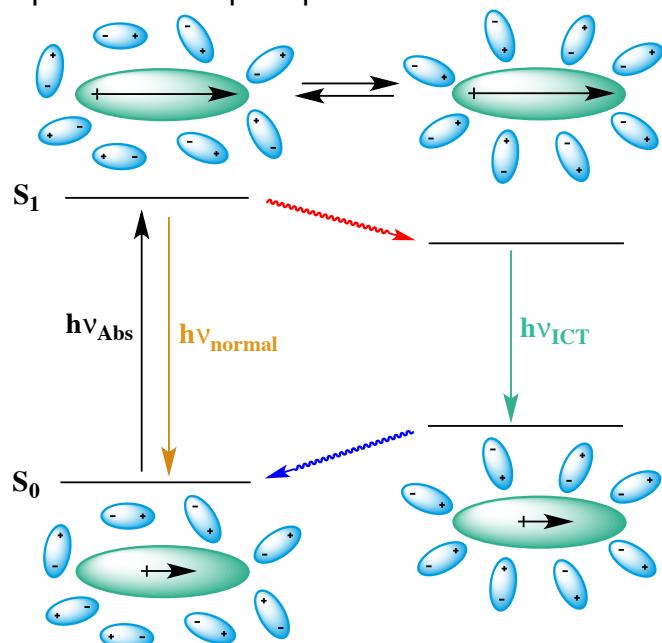


Figure 34. Diagramme d'énergie montrant la réorganisation des molécules de solvants et la stabilisation de l'état ICT par les dipôles du solvant.⁸¹

Dans certains cas, comme celui du DMBAN **7** (4-*N,N'*-diméthylaminobenzonitrile), la formation de l'état excité ICT nécessite la rotation d'un groupement du fluorophore, l'amine en l'occurrence : on parle d'état ICT twisté (TICT) [cf. Figure 35]. Les deux formes, plane et twistée, émettent à deux longueurs d'onde différentes (émission duale).⁶⁸

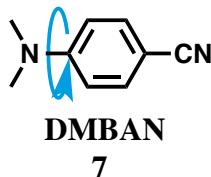


Figure 35. Structure du DMBAN.

b. Double Marquage

- FRET

Le FRET (pour *Fluorescence Resonance Energy Transfer* ou *Förster Resonance Energy Transfer*) est un transfert d'énergie non radiatif, résultant d'une interaction dipôle-dipôle (interactions coulombiennes) entre un fluorophore excité, le donneur, et une autre molécule à l'état fondamental, l'accepteur. Il se produit lorsque les états excités du donneur entrent en résonance avec ceux de l'accepteur. Pour ce faire, le spectre d'émission du donneur doit recouvrir, au moins partiellement, le spectre d'absorption de l'accepteur [cf. Figure 36].^{68b}

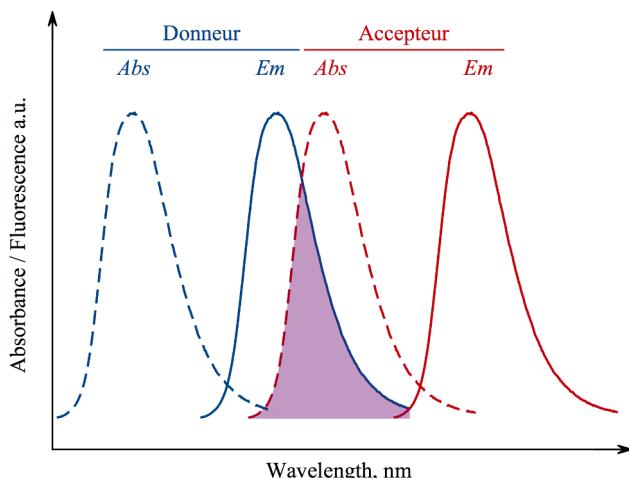


Figure 36. Illustration des spectres d'absorptions et d'émissions d'un couple FRET.

Le transfert d'énergie dépend de l'étendue du recouvrement spectral entre les deux entités mais également du rendement quantique du donneur, de l'orientation relative des deux espèces et de la distance qui les séparent. Cette distance, appelée distance de Förster, est caractéristique de chaque couple donneur-accepteur et correspond à la distance pour laquelle l'efficacité de transfert est de 50 %. Elle varie en général entre 10 et 100 Å (les distances de Förster les plus importantes sont obtenues avec des complexes de lanthanides).^{68b, 82}

Le FRET est couramment employé afin d'estimer la distance qui sépare les deux membres du couple. Par exemple, il est possible de suivre le repliement d'une protéine, doublement marquée, ou l'interaction entre deux macromolécules, l'une marquée avec le donneur, l'autre avec l'accepteur.^{68b}

- Excimères et Exciplexes

Comme pour le FRET, des processus intermoléculaires sont à l'origine de la formation des excimères (***Excited Dimer*** - homodimère) et des exciplexes (***Excited Complex*** - hétérodimère).

Le terme *excimère* a été introduit en 1960, par *Stevens et Hutton*. Ce sont des dimères de fluorophores qui n'existent qu'à l'état excité. Ils se forment par collision entre un fluorophore excité (M^*) et une autre molécule, chimiquement identique, restée à l'état fondamental (M). Les entités soumises à ce type d'interaction sont généralement des espèces polycycliques aromatiques telles que le pyrène, le pérylène ou le naphtalène.⁸³

Le spectre d'émission des excimères est très différent de celui des monomères ; Il est généralement plus large, décalé vers des longueurs d'ondes plus grandes et ne contient pas de structure vibrationnelle [cf. Figure 37].^{68a, 74}

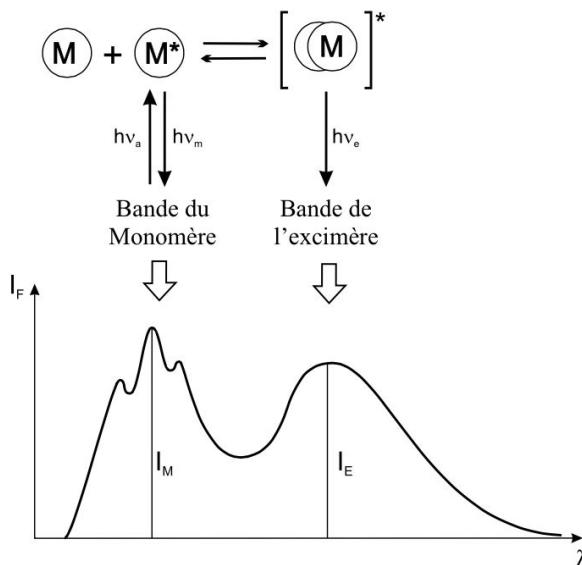


Figure 37. Exemple de spectre de fluorescence d'un monomère et de l'excimère correspondant.

Les exciplexes appartiennent à la même famille de composés, à ceci près qu'ils sont formés par la collision d'une molécule à l'état excité (M^*) avec une autre molécule non excitée, chimiquement différente (Q). Ce partenaire (Q) est généralement un donneur d'électrons, comme une amine. Ce type de processus a par exemple été observé lors de l'interaction entre l'anthracène et la *N,N'*-diéthylaniline, où un transfert d'électron de l'amine vers l'anthracène, à l'état excité, se produit.^{68a, 74}



La formation des excimères et des exciplexes nécessitent que les deux partenaires soient à proximité et correctement orientés. De plus, leur formation est réversible : après émission de photons, ils retournent à l'état fondamental où ils se dissoient. De nouveaux excimères et exciplexes peuvent alors se former.^{68a, 74}

- Etalon externe (fluorophores non couplés)

Une dernière solution permettant un dosage ratiométrique bimoléculaire consiste à mettre en présence une sonde, jouant le rôle de référence, et un capteur moléculaire, reliés entre eux ou non. Contrairement aux deux cas précédents, aucun processus intermoléculaire n'est mis en jeu (afin de limiter les signaux parasites). Ici, la référence fournit seulement un canal supplémentaire d'analyse (une deuxième bande de fluorescence). Deux cas de figure s'offrent au chimiste :

- la sonde de référence est insensible à la présence de la cible à analyser ou à tout autre perturbation. Elle joue alors le rôle d'étalon externe,⁸⁴
- la sonde de référence et le rapporteur présentent des sensibilités différentes une fois en présence de la cible (l'intensité de l'un augmente et celle de l'autre diminue ou les deux intensités augmentent / diminuent avec des cinétiques différentes).⁸⁵

Pour que le système soit optimal, les deux entités doivent pouvoir être excitées à la même longueur d'onde, émettre dans des domaines de fluorescence différents et posséder des intensités d'émission comparables.⁸⁶

Un des facteurs limitants importants de ce type d'approche peut cependant venir du fait que le photoblanchissement et/ou la dégradation chimique des deux entités peuvent être différents. Cela peut par conséquent faire varier le rapport d'intensité des deux bandes de manière irréversible.⁸⁵

ii. Anisotropie

L'anisotropie de fluorescence est une technique reposant sur l'excitation photosélective de fluorophores par de la lumière polarisée. Les fluorophores absorbent préférentiellement les photons dont les vecteurs électriques sont parallèles à leurs moments de transitions (par exemple, transition verticale - cf. II.A.2).^{68b}

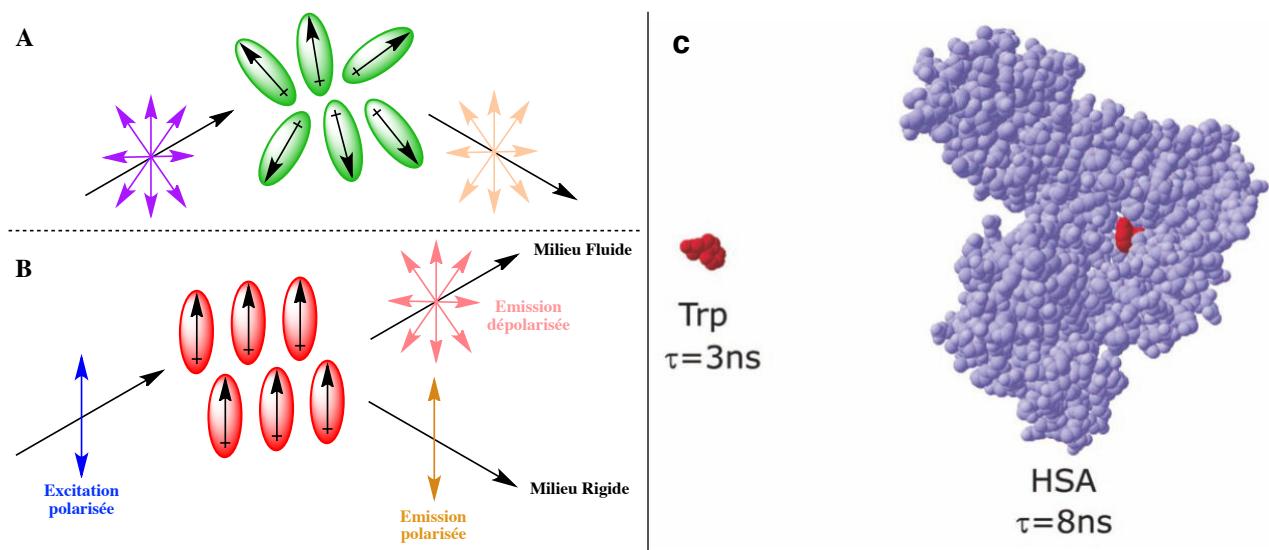


Figure 38. (A) Excitation et émission dépolarisée et (B) polarisée. (C) Exemple d'application : Trp (Tryptophane) = émission dépolarisée et complexe Trp-HSA (*Human Serum Albumin*) = émission polarisée.

Dans une solution isotrope, les moments de transition des fluorophores sont orientés de manière aléatoire [cf. Figure 38A]. Ces molécules absorbent donc des photons dont les vecteurs électriques sont orientés aléatoirement, eux aussi.

Lors de l'excitation avec de la lumière polarisée, seuls les fluorophores dont le moment de transition est parallèle au vecteur électrique d'excitation sont sélectivement excités : on parle de photosélection [cf. Figure 38B]. Cette technique est liée au temps de vie de fluorescence et à la vitesse de rotation de la molécule, et deux types d'émissions de fluorescence peuvent ainsi être observées :

- Si le système est fluide (solution, aucune gêne rotationnelle), la diffusion rotationnelle des molécules excitées (50 - 100 ps), pendant la durée de vie de l'état excité (\approx ns), provoque une dépolarisation de la lumière émise [cf. Figure 38B]. Plus généralement, les facteurs à l'origine de cette dépolarisation sont :
 - Des moments de transition d'absorption et d'émission non parallèles,
 - Des vibrations de torsion,
 - Des mouvements browniens,
 - Des transferts de l'énergie d'excitation à une autre molécule ayant une orientation différente.^{68, 73}
- Si le système est rigide (rotation du fluorophore entravée, basse température, viscosité élevée), ces facteurs sont empêchés / ralentis et les molécules sont considérées « immobiles » sur la période du temps de vie de l'état excité. Cela permet donc d'observer une émission de fluorescence également polarisée [cf. Figure 38B].^{68, 74}

L'enregistrement de l'anisotropie nécessite la mesure du rapport entre les intensités de fluorescence obtenues à deux polarisations différentes (verticale et horizontale).^{68, 74}

Ce type de mesures est couramment employé en biochimie. Elles fournissent des renseignements sur la mobilité moléculaire, leur taille, leur forme, leur souplesse ou encore sur la fluidité / rigidité du milieu [cf. Figure 38B et C]. Elles sont, par exemple, utilisées afin de mesurer des interactions entre protéines, la fluidité des membranes et pour des immunoessais sur de nombreuses substances [cf. Figure 38C].^{68, 74}

Par exemple, lorsque le fluorophore se trouve dans un milieu fluide, son anisotropie de fluorescence est proche de zéro, à cause des diffusions rotationnelles. Lors de l'interaction avec des macromolécules biologiques, la taille du système et/ou la viscosité locale augmente ; la vitesse de rotation diminue donc et l'on observe une augmentation de la polarisation et de l'anisotropie de fluorescence [cf. Figure 38C].⁷⁴

iii. Temps de vie

Suite à une courte impulsion excitatrice, un certain nombre de molécules passent à l'état excité S_1 . Une fois l'impulsion stoppée, ces molécules retournent ensuite spontanément à l'état fondamental S_0 en suivant une courbe exponentielle décroissante [cf. Figure 39]. Une même molécule peut présenter plusieurs temps de vie de fluorescence en fonction de son orientation et de son environnement proche.^{68a}

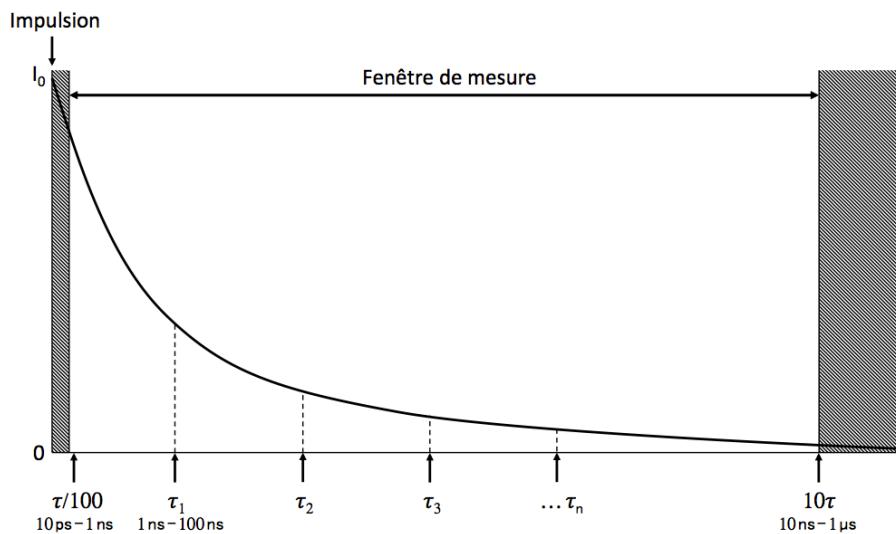


Figure 39. Déclin de l'intensité de fluorescence.

Ce type d'étude permet de suivre l'évolution temporelle de phénomènes dynamiques dont la vitesse d'exécution s'échelonne de $\tau/100$ à 10τ [cf. Equation 5]. Ce genre de considération est semblable à celle rencontrée lors de l'utilisation de radio-isotopes : la période radioactive doit être du même ordre de grandeur que l'âge de l'objet à dater.^{68a}

Le temps de vie est sensible aux changements du microenvironnement et une des applications importantes est l'imagerie par temps de vie de fluorescence (FLIM - *Fluorescence-Lifetime Imaging Microscopy*).⁸⁷ Cette technique utilise les différences de temps de vie de l'état excité du fluorophore comme un mécanisme de contraste pour l'imagerie [cf. Figure 40]. Les résultats obtenus sont indépendants de la concentration de la sonde.⁶⁸

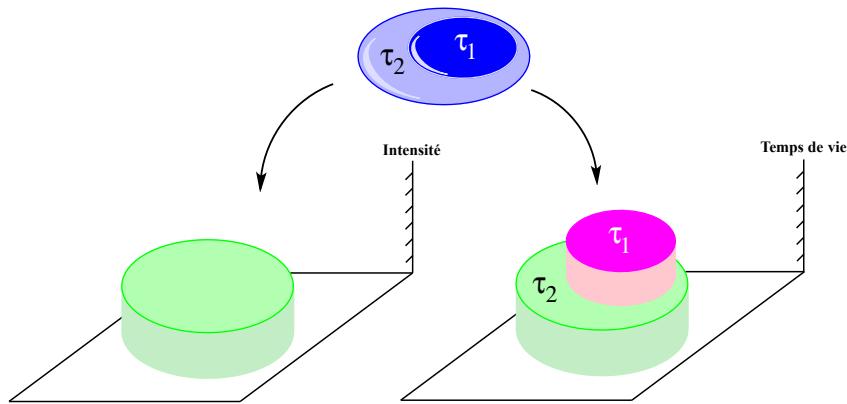


Figure 40. Concept de l'imagerie par temps de vie de fluorescence (FLIM).

Prenons l'exemple d'une cellule présentant deux régions distinctes, noyau et cytoplasme par exemple. Si l'intensité de fluorescence du fluorophore injecté dans la cellule est identique à l'état stationnaire dans chacune de ces régions 1 et 2 [cf. Figure 40], l'image obtenue ne permet pas de les distinguer. En revanche, si la durée de vie de la sonde est différente en fonction de la région (τ_1 et τ_2), il devient possible de les différencier. Plus particulièrement, ces mesures permettront de mettre en évidence la présence d'espèces ioniques, la liaison de la sonde à une macromolécule ou tout autres facteurs

environnementaux auxquels la sonde est sensible. C'est le même principe que celui de l'IRM, où la différence de relaxation des protons, en fonction du milieu, permet de différencier les tissus : le FLIM est en quelque sorte une méthode optique analogue à l'IRM.^{68b}

2. Applications à la chimie et à la biologie

La clé du succès de la fluorescence est le développement de sondes dont la structure permet de suivre des phénomènes spécifiques. Chaque type de sonde présente une sensibilité accrue à son environnement, d'une manière qui lui est propre. Les applications de ces sondes sont multiples :

- suivi du pH,
- dosage de composés (anions, cations, ...),
- détection d'interactions, de changement de conformation, ...,
- visualisation cellulaire.

i. Sondes « pH-sensibles »

Certaines sondes fluorescentes voient leurs intensités de fluorescence varier (effets hypso-, batho-, hyper- ou hypochrome) en fonction du pH environnant. Elles présentent une grande sensibilité et peuvent être utilisées afin de mesurer le pH à l'intérieur de cellules vivantes, par exemple.⁸⁸

En fonction de la nature de ces sondes, les variations de pH provoquent soit un transfert de proton photoinduit soit un changement de conformation de la molécule. On peut alors observer une augmentation / diminution du maximum d'émission de la sonde ou le décalage soit du maximum d'absorption, soit du maximum d'émission, soit des deux en même temps. Il est ainsi possible de réaliser un suivi monobande (variation d'intensité) [cf. Figure 41A], ou ratiométrique (décalage de bandes) [cf. Figure 41B], de l'évolution du pH. Afin de permettre une mesure optimale du pH, ces sondes doivent, bien évidemment, être utilisées dans des domaines de pH proches de leur pK_A .^{68a, 89}

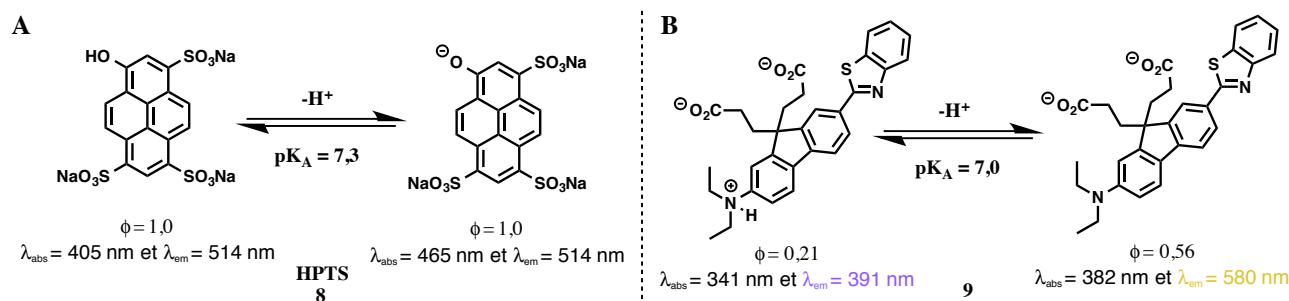


Figure 41. Exemple de sondes pH-sensible.

Ainsi, l'HPTS **8** (Acide 8-Hydroxypyrrène-1,3,6-Trisulfonique) [cf. Figure 41A], disponible commercialement, a été utilisé pour mesurer le pH cytoplasmique (6,8 - 7,4) ou celui d'organelles acides (4,5 - 6,0) de plusieurs lignées cellulaires. Alors que son maxima d'émission ne varie pas, ses maxima d'absorption, eux, fluctuent de 60 nm suivant que l'espèce soit protonée ou non [cf. Figure 41A]. De plus, les intensités d'absorption à 405 et

465 nm diminue et augmente, respectivement, quand le pH passe de 5 à 8. Exciter la molécule à 405 nm ou 465 nm revient donc à détecter une intensité de fluorescence décroissante ou croissante, respectivement (intensiométrie - monobande).^{89, 90}

De son côté, le dérivé fluorène **9** [cf. Figure 41B], également utilisé pour de l'imagerie cellulaire, présente des maxima d'absorption et d'émission qui varient respectivement de 41 et 189 nm en fonction du pH. La présence d'un point isobestique à 355 nm, permet d'exciter les deux espèces, protonée et non-protonée, simultanément afin de pouvoir réaliser une mesure ratiométrique du pH.^{89, 91}

ii. Sondes à ions

De nombreuses sondes ont la capacité de détecter des anions ou des cations, autres que le protons, afin de pouvoir les doser, par exemple. Les toutes premières sondes développées étaient monobandes et peu sélectives. Par exemple, l'oxine **10**, permettait de détecter des cations Cu(II), Al(III), In(III), ... [cf. Figure 42A] et la sonde SPQ **11** était souvent employée pour détecter des anions chlorures, même si elle répondait également à la présence d'I⁻, de Br⁻ et de SCN⁻ [cf. Figure 42B].^{68a}

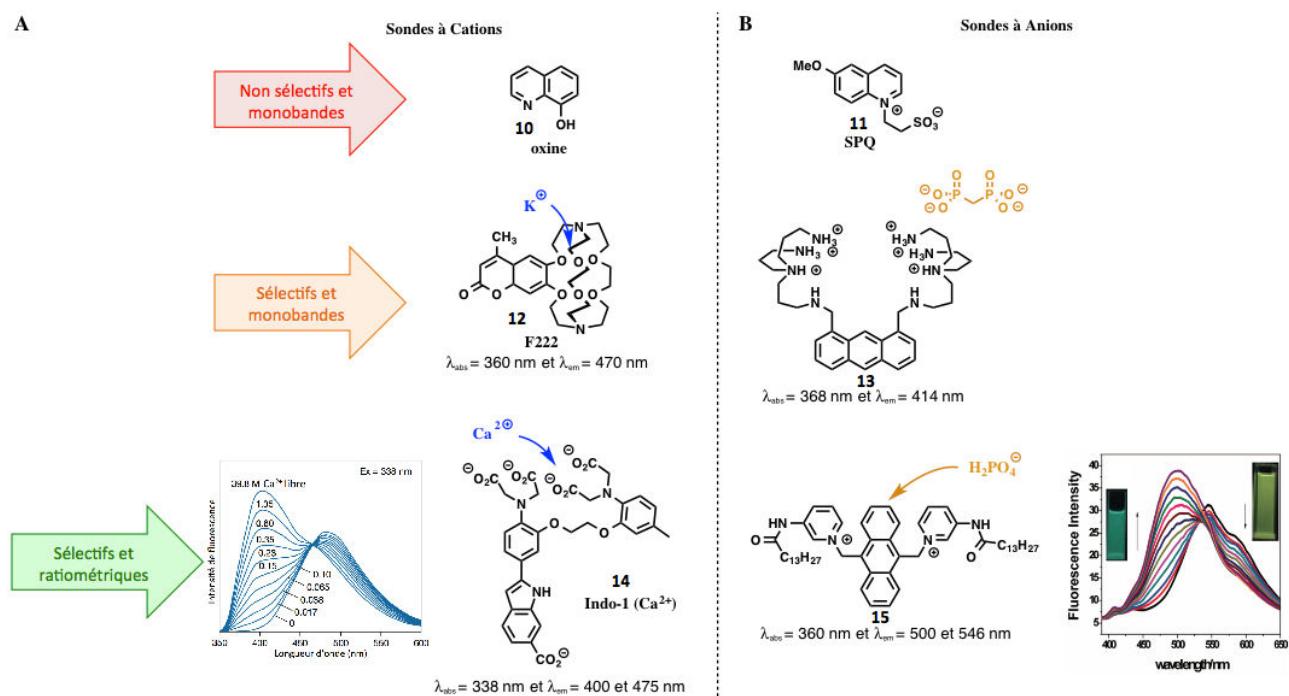


Figure 42. Exemple de sondes à cations et anions.^{84, 92}

Afin de développer des sondes plus spécifiques à un type d'ion donné, les chercheurs ont décidé de coupler des fluorophores à des « ionophores », des entités capables de complexer des ions, et ont appelés ces nouvelles molécules des « fluoroionophores » [cf. Figure 42]. Ces ionophores peuvent être de nature très variées : chélatants **13**, éther-couronnes, macrobicycles (cryptand **12**), calixarènes. Le principe de fonctionnement est simple. Un transfert d'électron inhibe la fluorescence lorsque le fluoroionophore est libre et la chélation de l'ion entraîne un changement structurel qui permet de « rallumer » la sonde.^{68, 93}

Cependant, ces fluoroionophores, bien que plus sélectifs, étaient toujours monobandes. Ainsi, les chercheurs ont mis au point de nouveaux analogues ratiométriques comme l'Indo-1 **14** et l'anthracène **15** [cf. Figure 42]. Utilisés *in vivo*, ils permettent de localiser et/ou de doser des ions endogènes tels que le calcium, le magnésium, le zinc, le sodium, ... A noter que l'Indo-1 appartient à la même famille que le Fura-2. Ils permettent un dosage ratiométrique intracellulaire du calcium, mais le Fura-2 n'est plus utilisé car la mesure est réalisée à partir du spectre d'excitation.⁹⁴

iii. Interaction et repliement de macromolécules

Les fluorophores présentant une sensibilité accrue aux changements de polarité sont dits solvatochromiques (cf. ESIPT et ESICT). Par exemple pour les sondes ESICT, un effet bathochrome ou hypsochrome, peut être observé sur leur spectre de fluorescence en fonction de la polarité environnante. Il est dû à la stabilisation de l'état excité du fluorophore par les molécules de solvant (augmentation du moment dipolaire à l'état excité - cf. II.B.1.i.ii.a.).^{68a, 95}

Le marquage protéique à l'aide de sondes solvatochromiques peut être utilisé afin de suivre leur repliement / ouverture (a), l'approche de leur ligand (c), la liaison avec une autre biomolécule (e et f) ou encore le passage d'une conformation active à une conformation inactive (d) [cf. Figure 43].^{79b}

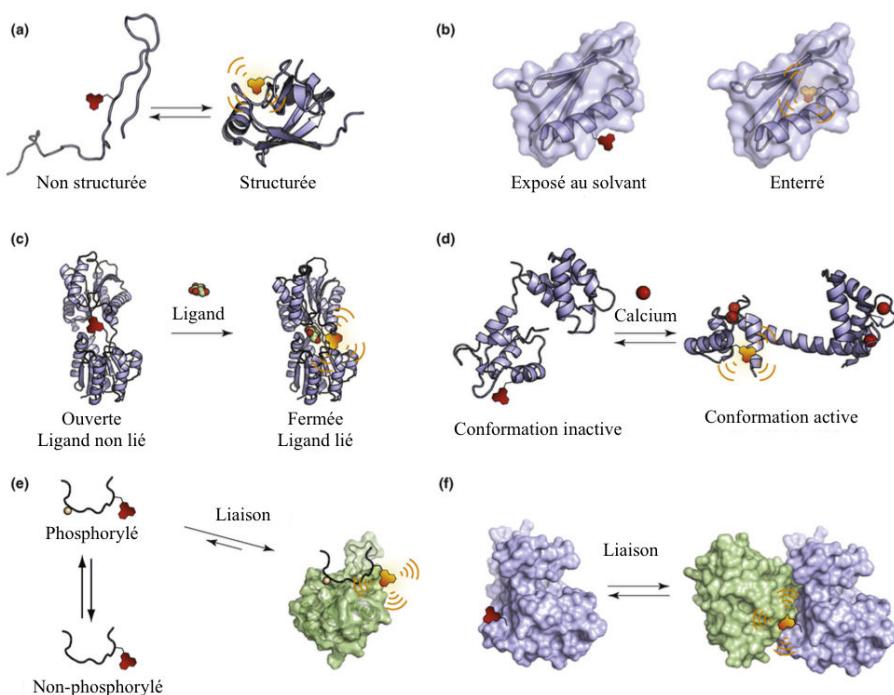


Figure 43. Applications possibles de sondes solvatochromiques (adapté d'*Imperiali et al.*).^{81b}

D'un autre côté, ces études de repliements (a et d) et d'interactions (e et f) peuvent également être réalisées par la méthode de FRET. Mis à part l'utilisation d'un fluorophore intrinsèque comme le tryptophane d'une protéine, un double marquage de la protéine (ou de l'acide nucléique) subissant le repliement, ou le marquage de chaque partenaire mis en jeu dans l'interaction, devient indispensable (cf. II.B.1.i.ii.b.).⁹⁶

iv. Imagerie cellulaire

La microscopie de fluorescence ou la cytométrie de flux sont des techniques qui permettent la visualisation des constituants présents dans nos cellules, qu'elles soient saines ou malades.⁹⁷ Ainsi, il existe des sondes capables de marquer la membrane cellulaire [cf. Figure 44C],⁹⁸ l'ADN (cf. Partie IV. Oligonucléotides) ainsi que diverses organelles [cf. Figure 44A et B].⁹⁹ Parmi ces organelles pouvant être visualisées par fluorescence, citons les mitochondries, les lysosomes, le réticulum endoplasmique, l'appareil de Golgi et le noyau, ainsi que plusieurs structures subcellulaires telles que les filaments d'actine ou encore les microtubules [cf. Figure 44A à C].

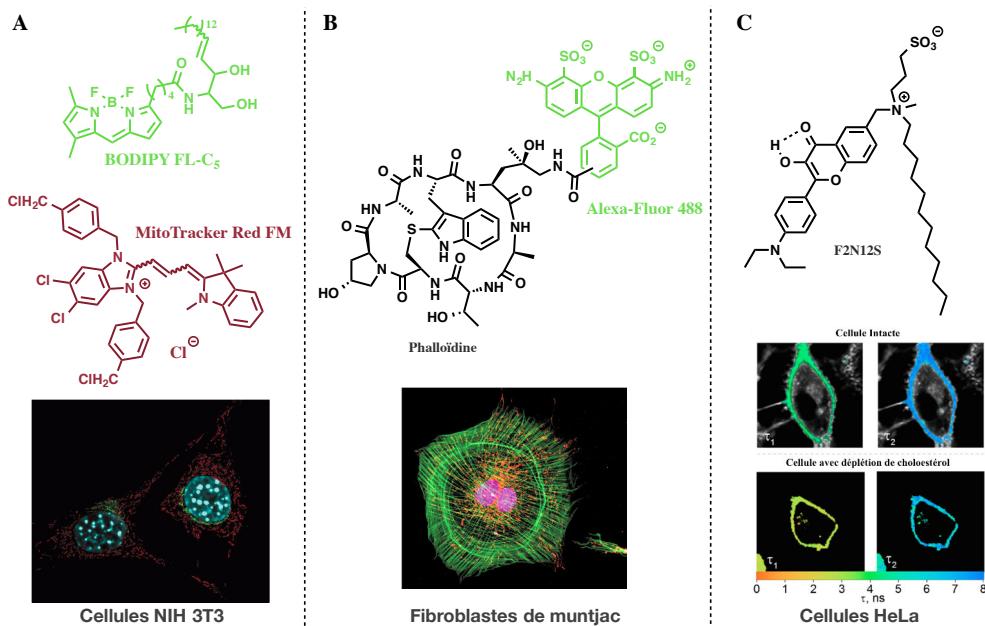


Figure 44. Sondes utilisées pour l'imagerie cellulaire.

Citons par exemple le BODIPY FL-C₅ [cf. Figure 44A] qui permet d'observer l'appareil de Golgi, le MitoTracker Red FM [cf. Figure 44A] qui fluoresce une fois à l'intérieur des mitochondries ou encore le DAPI et les colorants Hoechst qui interagissent avec l'ADN (cf. IV.B.1.i.ii). D'autres rapporteurs ont besoin d'être couplés à des molécules ciblant l'une ou l'autre de ces entités cellulaires. Par exemple, l'Alexa Fluor 488 [cf. Figure 44B] est couplé à la phalloïdine, une toxine produite par certains champignons et qui cible les filaments d'actine.⁹⁹ La chromone F2N12S [cf. Figure 44C] présente une chaîne aliphatique de 12 carbones de long et une tête hydrophile pour cibler et visualiser la membrane cellulaire. Ce composé a été utilisé afin de détecter une déplétion en cholestérol, composé indispensable à la rigidité de ces membranes, par un changement de durée de vie de fluorescence.⁹⁸

Nous allons à présent nous intéresser aux chromones et au mécanisme ESIPT auquel ces sondes ratiométriques sont soumises. Leur grande sensibilité aux changements environnementaux ainsi que leur grande versatilité en font les sondes de choix afin d'étudier les mécanismes intermoléculaires impliqués dans la méthylation de l'ADN.

III. Les 3-hydroxychromones

Les 3-hydroxychromones (3-HC) sont des composés hétérocycliques aromatiques dérivés de la 1-benzopyr-4-one (ou chromone) présentant un hydroxyle en position 3 de l'hétérocycle [cf. Figure 45]. Les motifs chromone et 3-HC sont présents dans les produits naturels de la famille des flavonoïdes, et de sa sous famille les flavonols. Ces composés sont des métabolites secondaires couramment rencontrés chez les plantes. Plus particulièrement, une flavone est un dérivé de chromone présentant nécessairement un groupement phényle en position 2. Les structures de la 1-benzopyr-4-one, de la flavone, de la 3-hydroxyflavone (3-HF) et de la 3-hydroxychromone (3-HC) sont représentées en Figure 45.

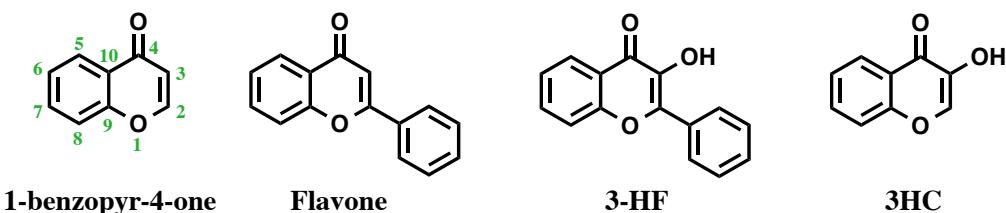


Figure 45. Structures des 1,4-benzopyrone, flavone, 3-HF et 3-HC.

Les flavonoïdes sont fortement répandus dans la nature, essentiellement dans les fruits et les légumes.¹⁰⁰ Ils présentent des activités bénéfiques sur notre santé, dans la prévention du cancer ou des maladies coronariennes (exemples : antioxydants, antitumoraux, anti-diabétique, ...). La quercétine est le dérivé de 3-HF ayant le plus fort potentiel.¹⁰¹

En plus de ces activités importantes, les composés 3-hydroxylés présentent également un intérêt pour leurs propriétés de fluorescence. En 1979, *Michael Kasha*¹⁰² découvrit que les flavonols pouvaient réaliser un transfert de proton intramoléculaire à l'état excité (ESIPT), à l'origine d'une émission de fluorescence duale (cf. III.B.1.i.). Les 3-HCs devinrent par la suite le modèle de prédilection pour l'étude de ce mécanisme ESIPT.

Dans les prochains paragraphes, différentes voies de synthèse des 3-HCs vont être passées en revue, ainsi que leurs caractéristiques photophysiques (ESIPT) et quelques applications.

A. Synthèse

Trois grandes voies de synthèses, développées au début du XX^{ème} siècle, permettent de construire le noyau des 3-HCs. Par ailleurs, certaines modifications de flavones ont été mises au point afin de pouvoir les hydroxyler et les convertir en flavonols.

1. Synthèse d'Auwers

Développée en 1908 par *Auwers* et *Müller*, cette réaction consiste à convertir un dérivé de coumarone **16** en aurone **17** puis en 3-HC **19**. Ainsi, une condensation en milieu acide du dérivé de coumarone **16** avec un aldéhyde aromatique donne l'aurone

correspondante (**17**). Cette dernière subit alors une réaction de dibromation, puis un traitement du dérivé dibromé **18** par du KOH, dans l'éthanol à reflux, permet finalement d'aboutir à la 3-HC souhaitée (**19**) [cf. Figure 46]. De nos jours, cette voie de synthèse n'est plus appliquée. Elle a été supplantée par les deux réactions suivantes.¹⁰³

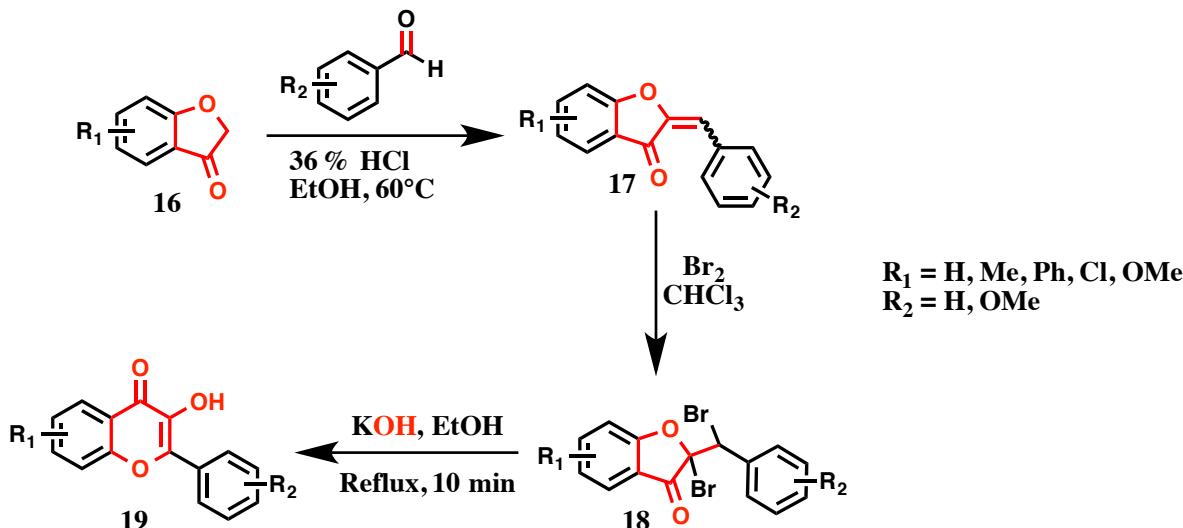


Figure 46. Synthèse d'Auwers.

2. Synthèse d'Allan-Robinson

En 1892, *Nagai* et *Tahara* firent réagir un phénol avec de l'anhydride acétique à reflux et en présence d'acétate de sodium. Ils obtinrent un composé hétérocyclique qui ne fut caractérisé que 10 ans plus tard (1901) par *Kostanecki* comme étant une chromone. Leurs travaux ont par la suite été complétés par *Allan* et *Robinson* en 1924 avec la synthèse de 2-alkyl- et 2-aryl-3-HC **21** [cf. Figure 47].¹⁰⁴

Cette dernière voie de synthèse, communément appelée condensation d'Allan-Robinson, consiste à faire réagir un dérivé de 2-alcoxy- ou 2-aryloxy-2'-hydroxyacétophénone **20** avec un anhydride aromatique (ou son chlorure d'acide correspondant) en milieu basique (amine). Elle se déroule à haute température et ne nécessite l'utilisation d'aucun oxydant.¹⁰⁴

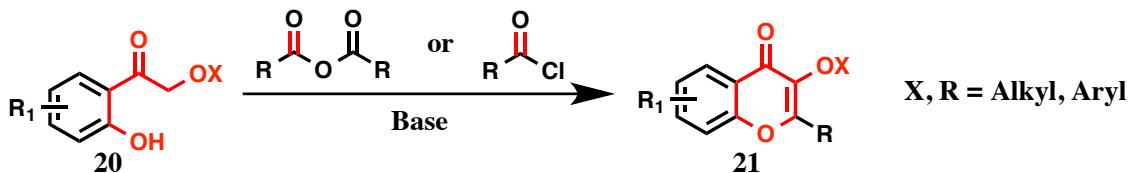


Figure 47. Condensation d'Allan-Robinson.

Elle a été essentiellement appliquée à la synthèse de dérivés naturels de flavonoïdes pour des applications thérapeutiques.

Notons enfin que la réaction de *Kostanecki* permet d'obtenir soit des chromones, soit des coumarines.¹⁰⁴

3. Synthèse d'Algar-Flynn-Oyamado (AFO)

Cette méthode fut pour la première fois reportée en 1934 simultanément par *Algar* et *Flynn* en Irlande et *Oyamada* au Japon.¹⁰⁵ Cette réaction d'Algar-Flynn-Oyamada (AFO) consiste tout d'abord à former une 2'-hydroxychalcone **23** par aldolisation-crotonisation entre un dérivé de 2'-hydroxyacétophénone **22** et un aldéhyde. Cette chalcone subit alors une cyclisation oxydante par traitement avec du peroxyde d'hydrogène en milieu alcalin afin d'être convertie en 3-HC. En fonction des substituants présents sur l'acétophénone et l'aldéhyde, les rendements en 3-HC finale **25** oscillent entre 15 et 70% [cf. Figure 48]. Notons également que l'aldéhyde peut présenter un hétérocycle aromatique.¹⁰⁵

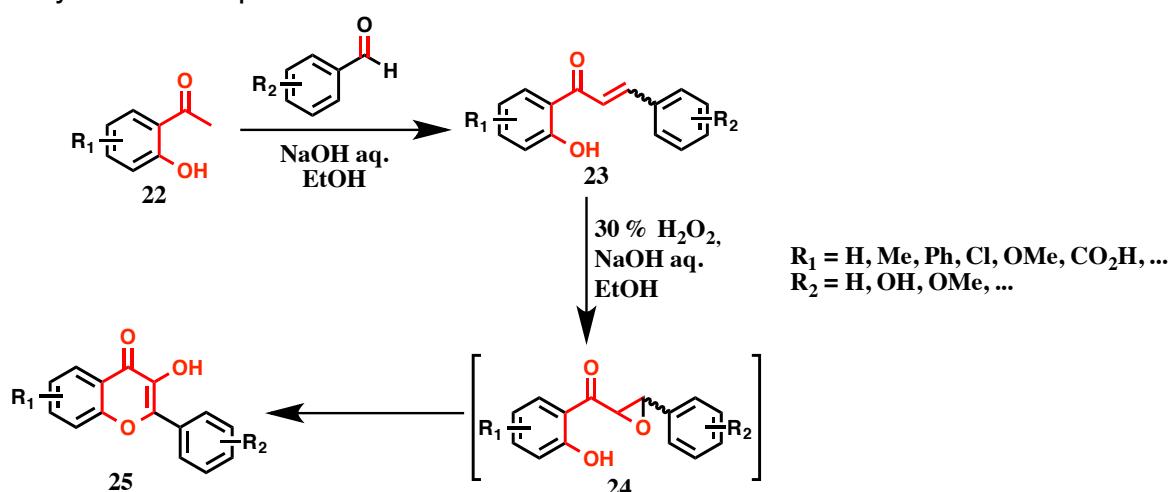


Figure 48. Réaction d'Agar-Flynn-Oyamada.

Cette réaction n'a subi que très peu de modifications depuis 1934. La plus intéressante d'entre-elles concerne la chalcone intermédiaire qui était initialement isolée puis purifiée. A présent, elle est générée *in situ* puis soumise, immédiatement après, à la cyclisation oxydante. Il est alors devenu possible d'obtenir rapidement, et en *un pot* (« *one pot* »), une large gamme de dérivés de 3-HCs, présentant différents substituants en position 2.¹⁰⁵

Enfin, notons que malgré son ancienneté elle est, encore de nos jours, la méthode de choix si l'on souhaite synthétiser des 3-HCs.¹⁰⁵

4. Autres voies de synthèse

L'obtention de dérivés 3-HCs peut également passer par une oxydation de la position 3 des chromones. Cette réaction a été réalisée suivant trois procédures différentes.

Le premier exemple implique une réaction tandem de lithiation - boration suivie par une oxydation de l'ester boronique en présence de peroxyde d'hydrogène. Cette méthodologie, appliquée à la synthèse totale du (-)-Episilvestrol et du (-)-Silvestrol, permet l'obtention du flavonol final avec un rendement global compris entre 60 et 92 % [cf. Figure 48].¹⁰⁶

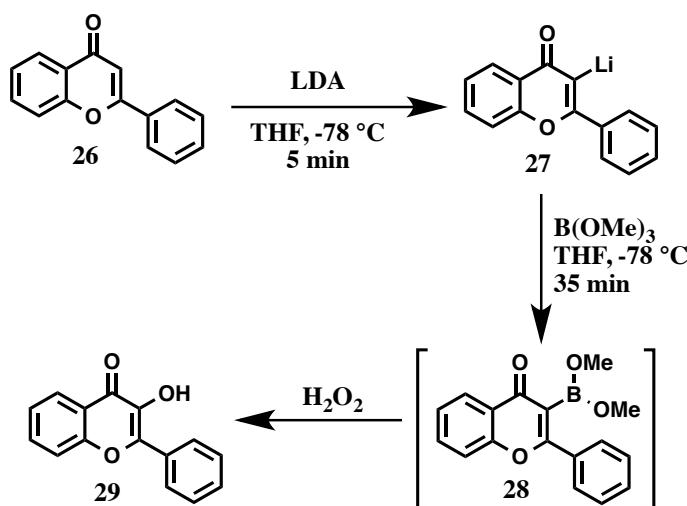


Figure 48. Lithiation, boration et oxydation de flavone en 3-HF.

Une autre approche consiste à utiliser un agent d'oxydation fort à base d'iode hypervalent. Par exemple, l'iodobenzène diacétate **31** fut utilisé par *Moriarty* et *Prakash* afin d'oxyder, en milieu basique, les flavones et chromones **30** [cf. Figure 49].¹⁰⁷

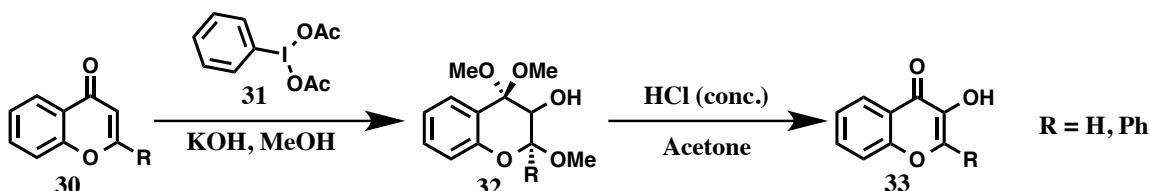


Figure 49. Iode Hypervalent afin d'oxyder les flavones / chromones en flavonols / 3-HCs.

Enfin, une dernière approche consiste à oxyder la flavone **30** en époxyflavone **34** qui subit, dans un second temps, une ouverture et réarrangement de l'époxyde en milieu acide afin de donner la 3-HC correspondante (**33**) [cf. Figure 50]. Typiquement, les conditions d'oxydation utilisées mettent en jeu du diméthyldioxirane ou de l'eau oxygénée ou du *m*-CPBA suivie par une ouverture concomitante de l'époxyde en milieu acide en présence d'acide *p*-toluenesulfonique ou du $NbCl_5$.¹⁰⁸

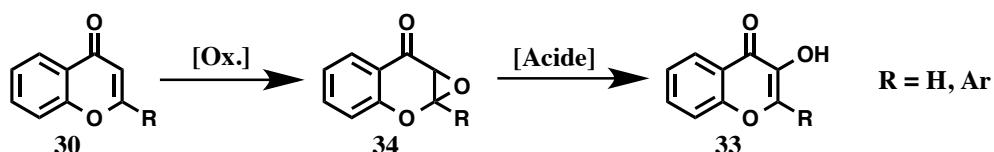


Figure 50. Epoxidation et ouverture de l'oxirane.

B. Caractéristiques Photophysiques

1. Présentation

i. ESIPT

L'ESIPT (*Excited-State Intramolecular Proton Transfer*) est un mécanisme de transfert de proton s'effectuant à l'état excité et pouvant être réversible. On le rencontre chez les fluorophores présentant des groupes donneur (hydroxyle, par exemple) et

accepteur (oxygène d'un carbonyle, par exemple) de liaison hydrogène à proximité l'un de l'autre. Lors de l'excitation, la redistribution de charges qui se produit entraîne une augmentation des caractères acide du proton et basique de l'accepteur, respectivement, et favorise ainsi l'ESIPT.⁷⁴

Ainsi, lorsque la forme normale N fondamentale est irradiée sélectivement, elle conduit à la forme normale excitée N* [cf. Figure 51]. Cette dernière peut ensuite soit revenir à l'état fondamental N (avec ou sans émission de photons), soit effectuer un transfert de proton intramoléculaire (ESIPT) afin de donner la forme tautomère excitée T*. Cette dernière subit alors, à son tour, une désexcitation (radiative ou non) qui conduit à la forme tautomère fondamentale métastable T. Enfin, un transfert de proton (BPT - *Back Proton Transfer*) permet le retour à la forme N [cf. Figure 51]. *Notons qu'il n'a jamais été observé de bande d'absorption de la forme T.*⁷⁴

Par conséquent, les deux formes excitées N* et T* permettent aux molécules soumises au mécanisme ESIPT d'émettre à deux longueurs d'onde différentes. Cette émission dualiste résulte du fait que chaque forme diffère par sa structure chimique et électronique, et donc qu'elle est plus ou moins stabilisée dans un solvant donné. Ainsi, chacune d'entre elles émet des photons avec une énergie différente, l'émission de la bande T* étant la plus décalée vers le rouge.

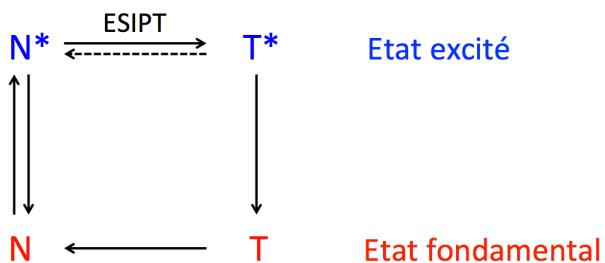


Figure 51. Schéma général du mécanisme ESIPT.

La première description du mécanisme ESIPT remonte à 1956 avec les travaux de Weller sur l'acide salicylique et l'ester méthylique correspondant. [cf. Figure 52A - *haut*].¹⁰⁹ Depuis, de nombreuses molécules pouvant réaliser ce transfert intramoléculaire ont été identifiées. Deux grandes familles peuvent être rencontrées ; celle présentant un transfert de proton vers :

- un atome d'oxygène [cf. Figure 52A],
- un atome d'azote [cf. Figure 52B].

*A noter que des transferts de proton vers des atomes de carbone peuvent également être rencontrés, de façon plus anecdotique, dans la littérature.*¹¹⁰

Les représentants de la première famille présentent un carbonyle comme accepteur. Il s'agit plus particulièrement de dérivés de l'acide salicylique, des 3-hydroxyflavones et 3-hydroxychromones ou encore des 3-hydroxyquinolones (3-HQ) [cf. Figure 52A]. Les molécules appartenant à la seconde famille sont des dérivés de la salicylideneaniline, des 2-(2'-hydroxyphenyl)benzimidazoles et 2-(2'-hydroxyphenyl)benzoxazoles [cf. Figure 52B].

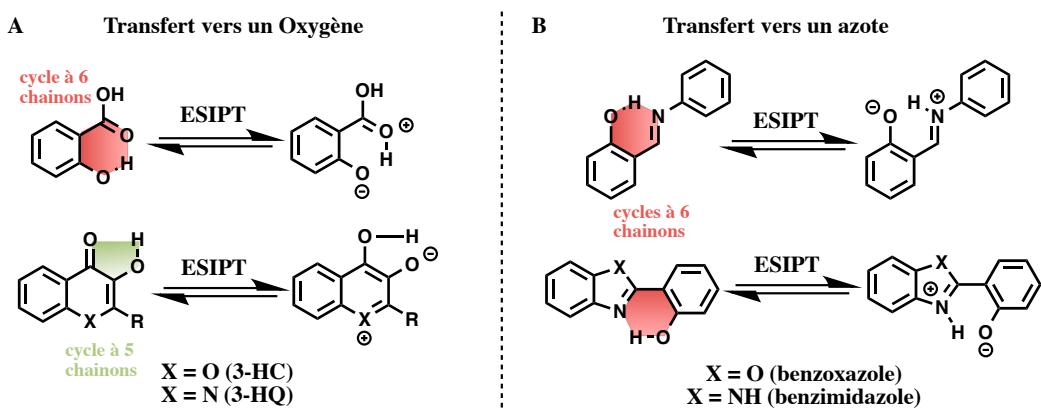


Figure 52. Exemple de molécules pouvant réaliser l'ESIPT.

La liaison hydrogène entre le donneur et l'accepteur forme dans la plupart des cas un cycle à 6 chaînons (cas optimal), à l'exception des dérivés 3-HF(C) et 3-HQ où un cycle à 5 chaînons est obtenu. Dans ce cas, la liaison H est plus longue que dans le cas des cycles à 6 chaînons ce qui conduit à un transfert de proton plus lent.

Le mécanisme ESIPT, et l'équilibre entre les deux formes, peuvent être fortement influencés par les interactions réalisées avec le solvant ainsi que par la structure du fluorophore, plus particulièrement par la nature électrodonneuse ou électroattractrice de ses substituants, comme nous allons pouvoir le constater dans les prochains paragraphes.

ii. Les 3-hydroxychromones (3-HC)

Les 3-hydroxychromones sont des sondes fluorescentes duales pouvant être excitées sélectivement dans des domaines où les biomolécules (protéines et acides nucléiques) sont transparentes ($\lambda > 320$ nm) et aux rendements quantiques qui peuvent être élevés. Comme indiqué précédemment, leur dualité provient d'un transfert de proton intramoléculaire à état excité (ESIPT), qui permet le passage de la forme normale excitée N^* vers la forme tautomère excitée T^* [cf. Figures 51 et 53]. Le donneur de proton (hydroxyle) et l'accepteur (carbonyle) sont respectivement positionnés sur les carbones 3 et 4 de l'hétérocycle.

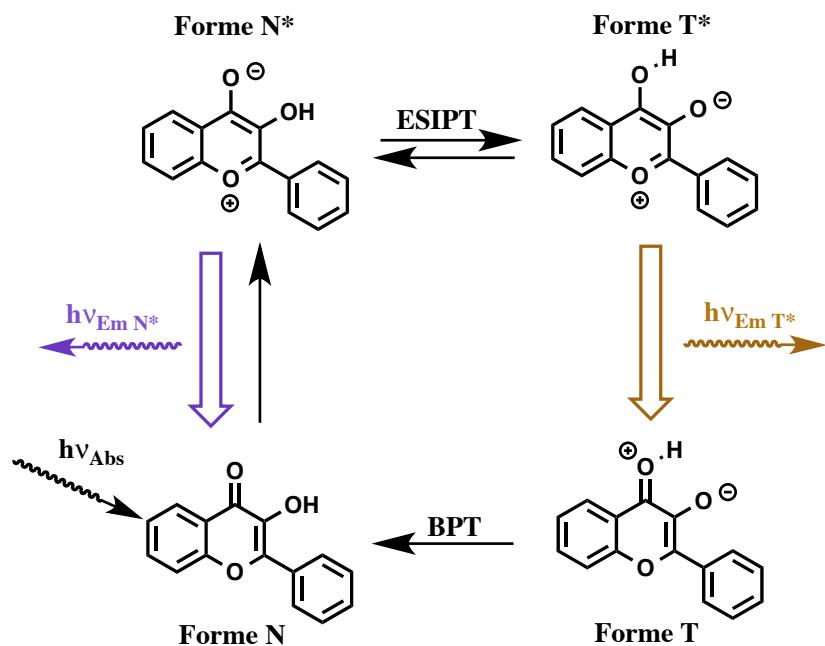


Figure 53. ESIPT des 3-HFs.

Le mécanisme ESIPT, auquel elles sont soumises, peut-être irréversible ou réversible et donc par conséquent la formation de la forme T^* peut-être sous contrôle cinétique ou thermodynamique, respectivement. Le type de mécanisme dépend du solvant et de la nature de la 3-HC et plus particulièrement de la présence ou non de substituants électrodonneurs ou électroattracteurs en position 2 ou 5 sur le noyau chromone [cf. Figure 54B et 55A].¹¹¹ La 3-HF [cf. Figure 54A] est l'exemple type d'une chromone qui suit un mécanisme de transfert irréversible, alors que la 4-diéthylamino-3-HF 36 [cf. Figure 54B et 56A], qui possède un substituant aryle très riche en électrons, est l'exemple type d'une chromone qui suit un mécanisme réversible. Après absorption de photons, la forme N^* possède un moment dipolaire plus large que celui des formes N et T^* , c'est donc la forme N^* qui sera la plus sensible à la polarité du solvant (*vide infra*). Notons que la forme T^* est également influencée par les effets du milieu mais dans une moindre mesure. Cette sensibilité de la forme N^* au milieu est d'autant plus importante que son moment dipolaire est élevé [cf. Figure 56B]. Ceci est particulièrement vrai pour la 4-diéthylamino-3-HF [cf. Figure 54B et 55B].^{109, 111e, 112, 113}

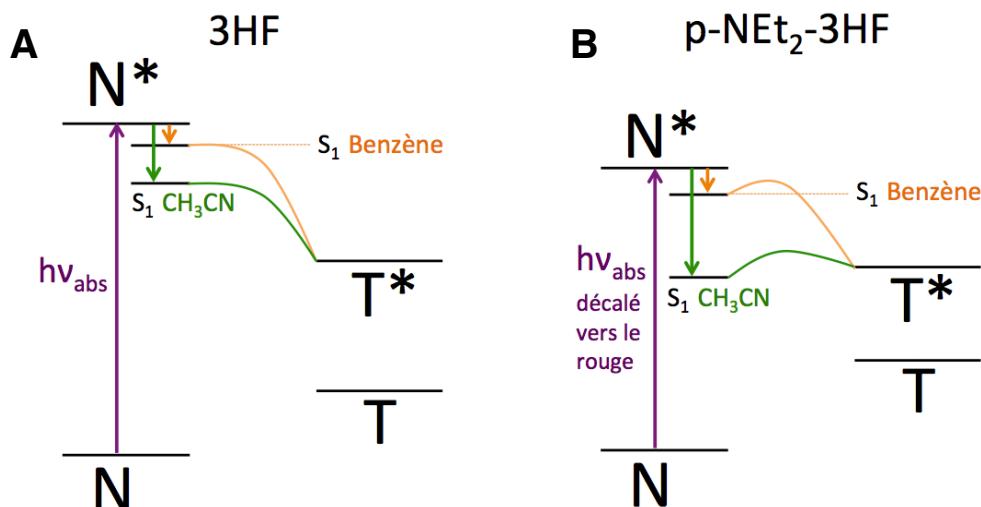


Figure 54. Représentation schématique et relative des différences entre les niveaux d'énergie des états excités de (A) la 3-HF et (B) du dérivé 4-diéthylamino-3-HF **36** dans des solvants aprotiques apolaires (benzène) et polaires (CH₃CN).

Prenons le cas de la 3-HF pour laquelle le mécanisme de transfert de proton est irréversible. Elle présente des cinétiques de transfert de proton très différentes quand elle est en présence de solvants aprotiques (Méthylcyclohexane, CH₃CN, ...) et protiques (H₂O, MeOH, EtOH, ...). La formation de liaisons hydrogènes entre la 3-HF et les solvants protiques semble être en grande partie responsable de ces différences. En effet, dans le cas des solvants aprotiques, la participation du solvant à la stabilisation de l'état excité (relaxation de solvant) est minime car l'augmentation du moment dipolaire de la 3-HF est faible lors de l'excitation.^{114e} Le niveau énergétique de l'état T* est donc bien plus bas que celui de la forme normale N* à l'état excité et la barrière énergétique entre les niveaux est assez faible (voire nulle) : le transfert de proton est donc très rapide.^{114c,e,f} En revanche, lorsque de fortes liaisons hydrogènes se forment entre le soluté et le solvant (35) [cf. Figure 55], comme en présence de solvants protiques, le transfert de proton intramoléculaire s'en trouve considérablement altéré.^{114e,f} Ceci est à l'origine d'un ralentissement de la vitesse du transfert de proton.^{114f} La barrière énergétique pour dissocier la liaison H solvant-soluté doit être franchie avant que l'ESIPT puisse se produire.^{114c,e} Ainsi, l'intensité de la bande N* augmente. Cependant, son intensité reste plus faible que celle de la bande T* qui demeure majoritaire [cf. Figure 54A].¹¹⁴ Notons que plus le pK_A du solvant est faible (proton plus acide) et plus le mécanisme ESIPT est inhibé (2,2,2-trifluoroéthanol vs éthanol).^{114c,j}

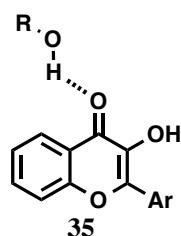


Figure 55. Inhibition du mécanisme ESIPT par compétition avec les donneurs de liaison hydrogène du milieu.^{114c,j}

Dans le cas de la 4-diéthylamino-3-HF **36**, le mécanisme de transfert de proton est réversible.^{114j} En effet, dans un solvant aprotique, comme l'acétonitrile, la barrière énergétique entre les niveaux N* et T* est assez faible (comme précédemment) mais en plus de cela, les états excités N* et T* ont des niveaux d'énergie très proches, grâce à une stabilisation plus importante de l'état excité due à un moment dipolaire accru par le transfert de charge.^{113c} De plus, le transfert de proton est ralenti, par rapport à la 3-HF, probablement à cause d'un mouvement de torsion entre le phényle et le groupement amino (TICT).^{114e,f} Le rapport d'intensité des deux bandes d'émission I_{N^*}/I_{T^*} est alors contrôlé par la distribution des populations de Boltzmann des niveaux N* et T* et la bande N* apparaît (elle devient même majoritaire dans CH₃CN, par exemple). D'un autre côté, la liaison hydrogène avec des solvants protiques (éthanol ou méthanol) entraîne la formation d'espèces solvatées N-H à l'état fondamental (en équilibre avec N). Elles peuvent être excitées directement afin de donner N-H* (en parallèle de la forme N non protonée qui donne N*) [cf. Figure 56B].^{114j} *Cet équilibre n'est pas observée à l'état excité.* L'espèce non solvatée N, présente un ESIPT rapide, empêchant ainsi toute solvatation une fois excitée. En revanche, pour l'espèce solvatée N-H*, un ralentissement du mécanisme ESIPT est observé.^{114j} Ceci peut s'expliquer par une stabilisation sélective de la forme normale (augmentation de la barrière énergétique entre les deux états excités) et un affaiblissement de la liaison hydrogène intramoléculaire (inhibition du transfert de proton et donc diminution de la population de la forme tautomère).^{114j} De plus, le complexe T-H*, entre la sonde et du solvant, est instable et se dissocie afin de redonner T*. La contribution de la forme tautomère dans l'émission duale diminue alors. Ainsi, l'intensité de la bande normale, celle de la forme N-H* (N* étant négligeable), augmente alors - bien plus que dans le cas de la 3-HF.^{112, 113c, 114f, 115}

Ainsi, la sensibilité des 3-HCs peut être modulée en fonction du milieu à sonder, par introduction de groupements électrodonneurs ou électroattracteurs en position 2 et/ou 5.

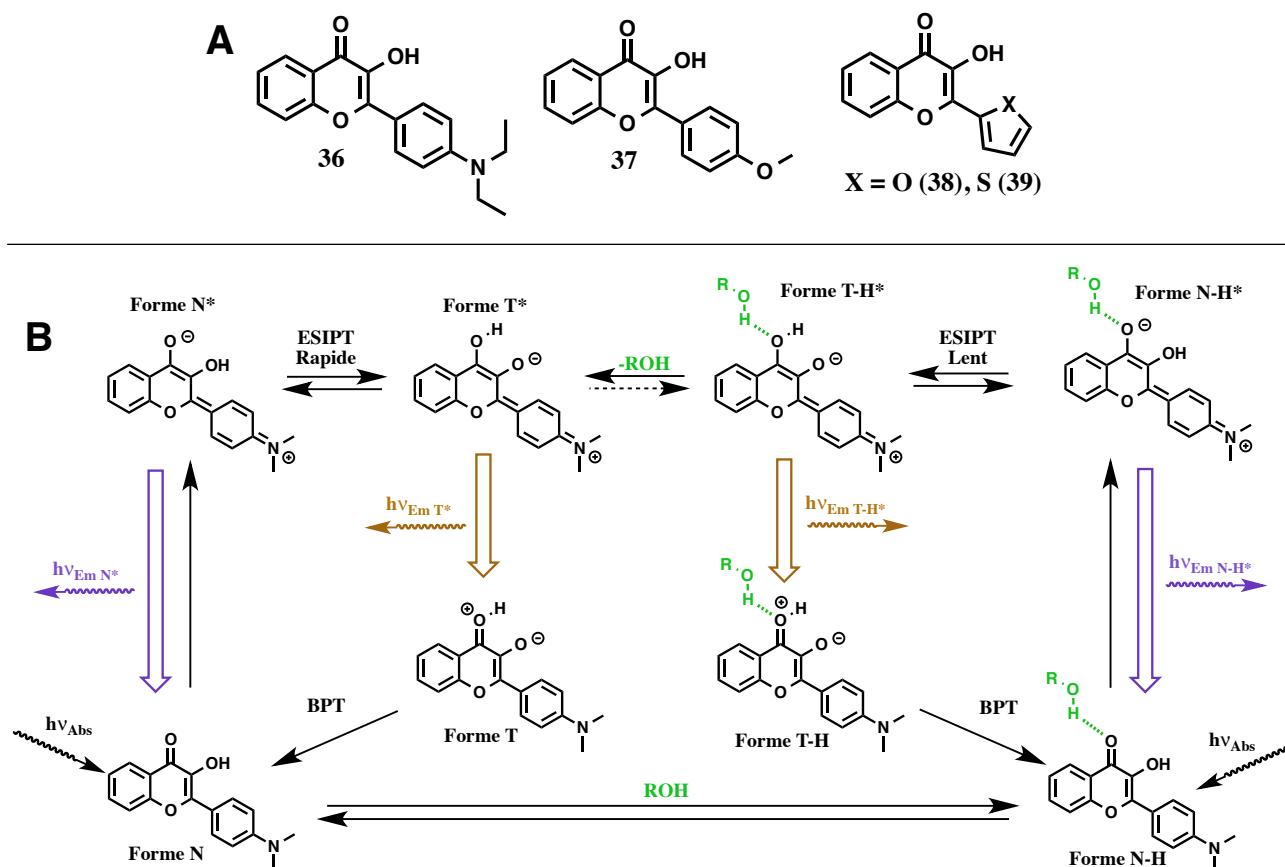


Figure 56. (A) Structure de trois dérivés de 3-HCs présentant divers groupements électrodonneurs en position 2 et (B) mécanisme ESIPT de la 4-diéthylamino-3-HF **36** au moment dipolaire important.^{112b-d, 114f,j 116}

La forme N^* étant plus sensible que la T^* aux changements environnementaux (interactions dipôle-dipôle, liaisons hydrogènes et donc hydratation ...), il devient possible de suivre l'évolution du milieu en mesurant le rapport d'intensité entre les deux bandes (I_{N^*}/I_{T^*}). Il s'agit d'un des canaux d'information les plus importants des sondes ESIPT, que l'on qualifie ainsi de sondes ratiométriques. Par ailleurs, d'autres paramètres tels que les maxima d'absorption et d'émission permettent également d'exploiter le caractère multiparamétrique de ces sondes, c'est-à-dire qu'elles sont capables de fournir simultanément des informations sur différents paramètres environnementaux.^{114j} Par exemple, avec la 4-diéthylamino-3-HF **36**, la mesure du déplacement de Stokes permet d'estimer la polarité du milieu.^{115c} Cette grande polyvalence a ainsi permis de les utiliser comme sondes fluorescentes pour le développement de capteurs, lors d'études en biologie cellulaire et moléculaire, comme nous allons pouvoir le constater dans les prochains paragraphes.

2. Applications

Les interactions macromoléculaires se produisant à l'intérieur d'une cellule sont des événements capitaux pour la survie de tout être vivant. Leur complexité rend cependant leur étude difficile. Il y a un besoin réel de sondes fluorescentes capables de sonder les changements environnementaux de manière site-spécifique et sensible. Dans ce contexte

le développement et l'utilisation de capteurs à base de 3-HCs a permis d'apporter des solutions innovantes. Elles ont par exemple été utilisées afin d'étudier le changement conformationnel de protéines, ou des interactions non-covalentes se déroulant au niveau de bicouches lipidiques, de membranes cellulaires, de protéines ou peptides et de l'ADN.¹¹⁷

i. Marquage des protéines / peptides

La première utilisation des 3-HCs afin d'étudier le comportement des protéines remonte à 1994 avec les travaux de *Kasha*.¹¹⁸ Il utilisa trois dérivés de 3-HF, dont le 4-diéthylamino-3-HF [cf. Figure 56A], capables d'interagir avec la RSA (*Rat Serum Albumine*), afin de déterminer son site de liaison. De telles études, par interaction non-covalentes avec le fluorophore, ont également permis de détecter deux sites de liaison de la 3-HF sur HSA (*Human Serum Albumine*) 2 ans plus tard.¹¹⁹ Cependant, au cours de ces travaux, les chercheurs se sont intéressés à l'intensité du signal de fluorescence et non au rapport d'intensité des deux bandes.

Il fut cependant utilisé par la suite afin de suivre l'interaction de protéines¹²⁰ ou de peptides¹²¹ avec d'autres biomolécules. Pour ce faire, le dérivé de chromone était lié de façon covalente, soit à une lysine ($-\text{NH}_2$) soit à une cystéine ($-\text{SH}$) d'une de ces entités protéiques. Prenons l'exemple illustré en Figure 57 où l' α 1-antitrypsine a été marquée avec le dérivé 2-furylchromone **40**, au niveau de sa cystéine 232, afin d'étudier son repliement induit par l'interaction avec l'élastase pancréatique. Le rapport I_{T^*} / I_{N^*} initialement de 1,18, chute d'environ 65 %, une fois les deux entités mises en présence l'une de l'autre. En d'autres termes, l'intensité de la bande T^* diminue et celle de la bande N^* augmente, suggérant une plus grande exposition à l'eau de la 3-HC après interaction [cf. Figure 57].^{117, 120b}

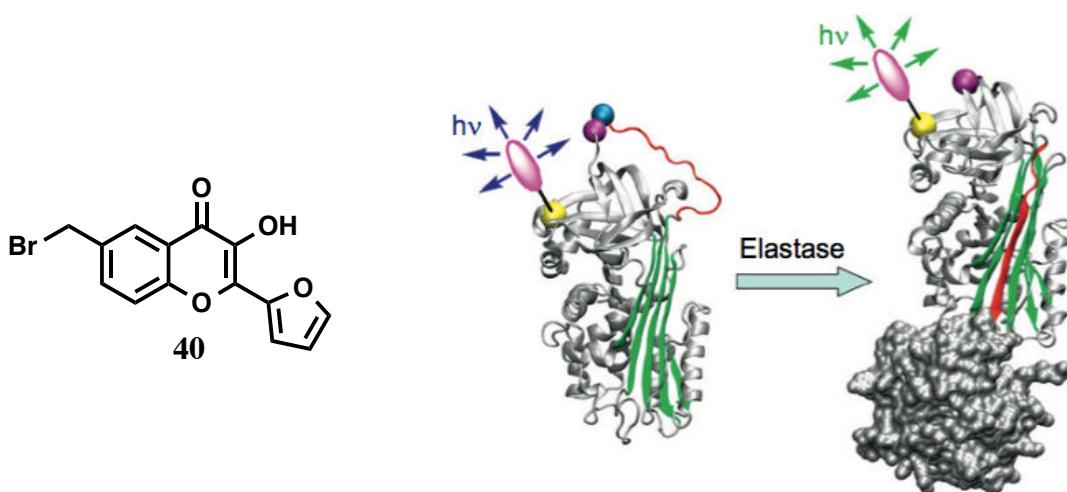


Figure 57. (A) Structure de la 3-HC utilisée pour marquer l' α 1-antitrypsine et (B) interaction entre l' α 1-antitrypsine marquée et l'élastase.^{119, 120b}

D'autres études sur des peptides marqués ont également été réalisées afin d'étudier leurs interactions avec des acides nucléiques,^{121a,d} des protéines^{121b-c} ou encore des membranes cellulaires.^{117, 121e-f} C'est par exemple le cas de la protéine de

nucleocapside (NCp7) du VIH-1, qui joue un rôle majeur en interagissant avec l'ARN viral et ses copies ADN. La partie *N*-terminale du peptide correspondant au domaine « doigt de zinc » a ainsi été marquée avec un dérivé de 3-HC afin d'étudier ses interactions avec des ODNs tels que cTAR (*Trans-Activation Response element*) ou (-)PBS (*Primer Binding Site*) par exemple. Une diminution du rapport d'intensité, soulignant une exposition moins importante à l'eau provoquée par l'interaction du peptide avec ces acides nucléiques, a ainsi pu être observée.^{121a}

ii. Marquage des phospholipides / micelles

La première étude d'environnement micellaire, à l'aide de 3-HF, remonte à 1991.¹²² L'hydrophobicité de la chromone la poussant à se localiser dans le cœur hydrophobe de la micelle, seule la forme tautomère T* y est observée (émission de fluorescence verte) et le suivie de son émission de fluorescence a permis de déterminer la concentration micellaire critique (CMC).¹²² Un autre dérivé de 3-HF a également permis de suivre la transition gel - crystal liquide dans des liposomes, par des études d'anisotropie de fluorescence en fonction de la température.¹²³ Enfin, de récentes études résolues en temps, ont permis de déterminer que le mécanisme ESIPT était plus lent dans le cas de micelles anioniques que dans le cas de micelles cationiques. En effet, le taux d'hydratation au niveau des micelles chargées négativement semble être plus important, ce qui a tendance à affaiblir la liaison hydrogène intramoléculaire.¹²⁴

De nouveaux dérivés de 3-HCs, présentant une structure de tensioactif [cf. Figures 44C et 58] ont également été développés afin d'étudier, de manière plus spécifique, les bicouches lipidiques des cellules.¹²⁵ Brièvement, les groupements chargés de la tête polaire créent un champ électrique local qui rend le mécanisme ESIPT très sensible à la constante diélectrique du milieu. Une telle modulation électrochromique de l'ESIPT permet à ces sondes d'être à la fois sensibles au champ électrique local (potentiel transmembranaire) et à l'hydratation. La plupart d'entre elles présentent des rendements quantiques corrects (0,05-0,17). *A noter que la présence d'une charge positive à proximité de la chromone permet d'augmenter le rendement de fluorescence de manière significative (0,3-0,7).*^{125b} Ce genre de sonde a par exemple été utilisé pour détecter la perte d'asymétrie de la membrane plasmatische dans les premières étapes de l'apoptose.^{125f}

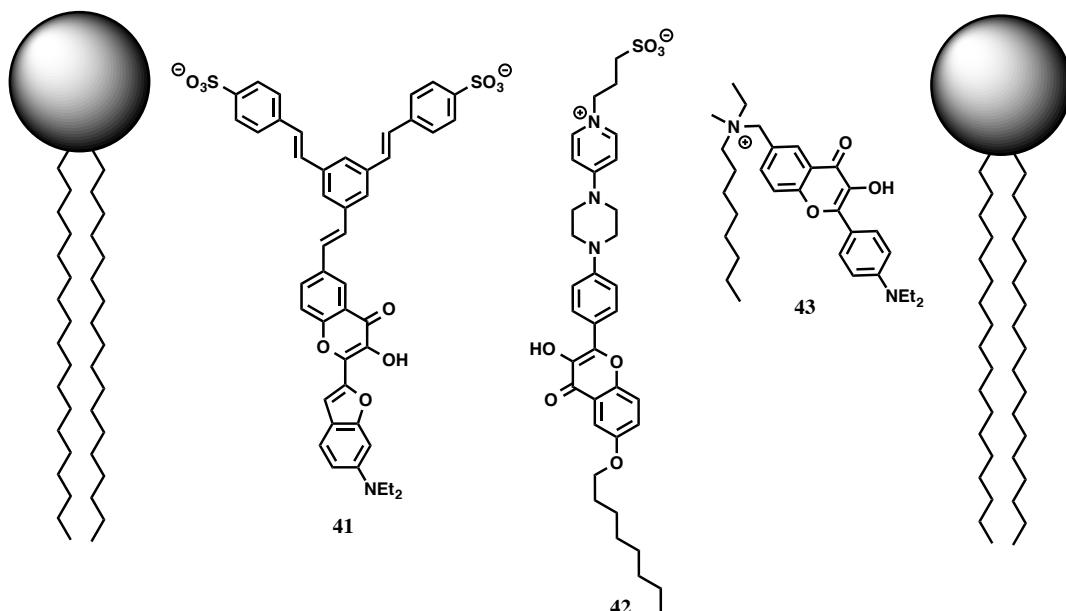


Figure 58. 3-HC présentant une structure de tensioactif.^{125h}

Comme autre exemple de ce type de tensioactifs, rappelons aussi le dérivé F2N12S [cf. Figure 44C], présenté précédemment, qui permet de détecter une déplétion en cholestérol grâce à des études de temps de vie.⁹⁸

iii. Marquage de l'ADN

Le marquage de l'ADN permet de suivre ses interactions avec d'autres macromolécules biologiques comme des protéines ou des enzymes et/ou d'étudier sa conformation. Le marquage des acides nucléiques constituant une part importante du manuscrit, seule une liste des différentes stratégies va être présentée ici. Elles seront plus amplement détaillée par la suite (cf. Partie IV. Oligonucléotides).¹²⁶

Comme dans le cas des protéines, l'étude par interaction non covalente de chromones avec l'ADN a initialement été réalisée. Par exemple, afin d'étudier le microenvironnement autour de simples brins et de duplexes d'ADN, la chromone a initialement été couplée à de la spermine **44**,^{126a} un polyamine naturel ayant une forte affinité pour l'ADN [cf. Figure 59].¹²⁷ D'autres travaux, plus récents, ont consisté à utiliser la capacité de la Fisétine **47** à interagir avec diverses conformations d'acides nucléiques (duplexes, i-motif, G-quadruplexes) afin de détecter des sites abasiques^{126c} ou de les différencier les unes des autres, par changement du rapport I_{N*} / I_{T*}.^{126h} Enfin, l'approche covalente, développée au sein du laboratoire, consiste à synthétiser des analogues nucléosidiques présentant un motif chromone et à les introduire dans la séquence d'ADN, de manière site-spécifique, afin de suivre des interactions ADN / protéines (cf. Partie IV.B.1.i.).^{126b,d-g}

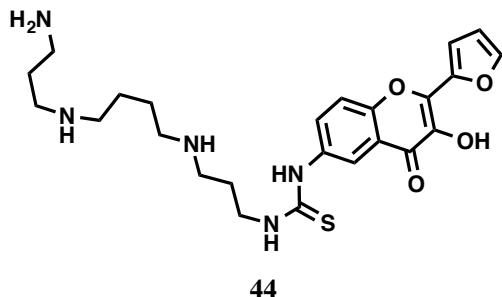


Figure 59. 3-HC couplé à de la spermine.^{126a}

Les différentes stratégies de marquages des acides nucléiques et leurs applications vont à présent être approfondies dans la prochaine partie IV. Oligonucléotides.

IV. Oligonucléotides

Afin de pouvoir jouer son rôle au sein de nos cellules, l'ADN doit interagir avec différentes protéines comme par exemple les histones qui compactent l'ADN, les ARN polymérases qui catalysent la production d'ARNm, ou encore les ADN Méthyltransférases qui méthylent l'ADN dans le but de masquer certains gènes... Pour pouvoir étudier ces interactions, il est indispensable de connaître la structure des entités mises en jeu ainsi que leur dynamique. La résonance magnétique nucléaire, la spectrométrie de masse, la cristallographie aux rayons X et la spectroscopie de fluorescence sont des techniques largement utilisées afin de caractériser ces assemblages complexes d'acides nucléiques et de protéines. Parmi celles-ci, la spectroscopie de fluorescence est une technique de choix pour suivre cette dynamique par sa sensibilité extrême qui permet de travailler à de très faibles concentrations, comparables à celles des milieux biologiques. La réalisation de telles études nécessite cependant de disposer de fluorophores aux propriétés photophysiques optimisées. Ils doivent, en effet, être capables de moduler leurs émissions de fluorescence en fonction des changements environnementaux s'opérant lors de la formation des complexes nucléoprotéiques, afin de permettre un suivi des interactions ADN / protéines qui en découlent. Pour réaliser de telles études, les chercheurs synthétisent des fragments d'ADN, ou oligonucléotides (ODN), ayant une forte affinité pour ces protéines, et les marquent à l'aide de rapporteurs fluorescents.

A. Synthèse des ODNs

La synthèse d'oligonucléotides peut être réalisée via deux approches bien distinctes : la voie enzymatique ou l'approche chimique, la plus employée. Ces techniques sont utilisées afin d'incorporer des nucléosides naturels et modifiés, tels que des rapporteurs fluorescents, au sein de brin d'ADN.

Chacune de ces approches présente des avantages et des inconvénients détaillés ci-après :

- La méthode enzymatique permet d'utiliser la grande spécificité et le pouvoir catalytique de ces biomolécules. Elle permet de générer des brins d'ADN long, comprenant un nombre de paires de bases de l'ordre de 10^4 .¹²⁸ Elle nécessite l'utilisation de nucléotides triphosphates (dNTP) qui peuvent présenter des nucléobases naturelles (A, T, C ou G), modifiées¹²⁹ ou artificielles.¹³⁰ Une contrainte majeure impose que les analogues de nucléotides, par exemple à bases modifiée ou artificielle, soient reconnus comme substrat de l'enzyme. D'un autre côté, la synthèse enzymatique est compatible avec un grand nombre de groupements hautement réactifs (azotures,¹³¹ aldéhydes¹³² ou encore accepteurs de Michael)¹³³ ne nécessitant pas l'utilisation de groupes protecteurs. A partir des bases naturelles, différentes positions peuvent être exploitées pour introduire un groupe rapporteur [cf. Figure 60]. Les points d'ancrage possibles sont pour les bases puriques la position 8, à la condition d'utiliser un bras espaceur flexible, la position 7 des deazapurines, le NH₂ de l'adénine (grand sillon), ou le NH₂ de la guanine (petit sillon) ; pour les bases pyrimidiques la position 5 et le NH₂ de la cytosine (grand sillon) [cf. Figure 60]. Enfin, notons qu'un grand nombre de techniques permettant de synthétiser ces ODNs existent (PCR - *Polymerase Chain Reaction*, PEX - *Primer Extension*, TdT - *Terminal deoxynucleotidyl Transferase* & SNI-PEX - *Single-Nucleotide Incorporation-PEX*). Elles se basent sur l'utilisation d'un brin matrice afin de générer son brin complémentaire (sauf TdT). Ainsi, l'incorporation de bases naturelles se fait de manière spécifique car se basant sur la présence, sur le brin matrice, de la nucléobase complémentaire. En revanche, l'incorporation site-spécifique de bases modifiées, analogues des bases naturelles, dans bons nombres de cas (PCR, PEX et TdT) n'est pas possible : par exemple un analogue de A* substituera toutes les A de la séquence en croissance. La synthèse chimique des ODNs ne présente pas cet inconvénient. A l'heure actuelle, une seule solution permet de marquer de façon site-spécifique un ODN, avec ces analogues de nucléobase : la méthode SNI-PEX.^{134d} Enfin, la création de couple de bases artificielles, permet d'étendre le répertoire de l'alphabet génétique et de proposer une autre solution et leur incorporation spécifique est possible. Pour cela, il est indispensable que le brin matrice présente la nucléobase artificielle complémentaire.¹³⁴

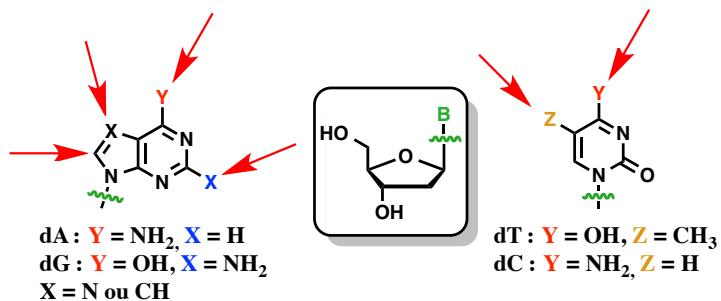


Figure 60. Positions de modification couramment employées (flèches rouges).

- La synthèse chimique est la méthode de choix afin d'incorporer des nucléotides présentant une base non-naturelle de façon site-spécifique. Elle permet d'introduire une plus large variété d'analogues de nucléotides que la méthode enzymatique. Bien qu'historiquement mise au point en solution,¹³⁵ la synthèse chimique des ODNs est

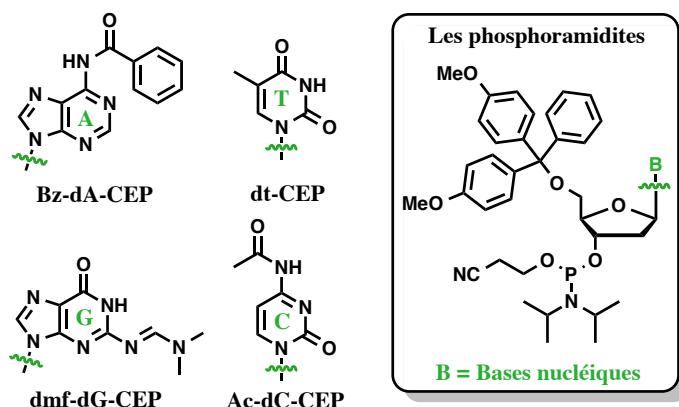
aujourd’hui réalisée sur support solide car c'est la technique la plus performante afin de produire des ODNs courts (inférieur à 200 nt). Elle présente de nombreux avantages. Le support solide permet d'utiliser un large excès de réactif afin d'augmenter la vitesse des réactions. Cet excès est facilement éliminé par des lavages une fois le couplage réalisé. Ainsi, seul l'ODN final est soumis à une étape de purification. Le processus est également automatisable. L'automate se charge d'introduire les différents nucléotides les uns à la suite des autres, suivant la séquence qui lui a été renseignée au préalable. L'inconvénient principal de la synthèse chimique vient du fait qu'elle est linéaire. Afin de contourner cela, les différentes étapes de la synthèse [cf. Figure 62] ont été optimisées afin d'obtenir les meilleurs rendements de couplage possible : de l'ordre de 98 - 99 % (chimie des phosphoramidites).

1. Synthèse chimique moderne des ODNs

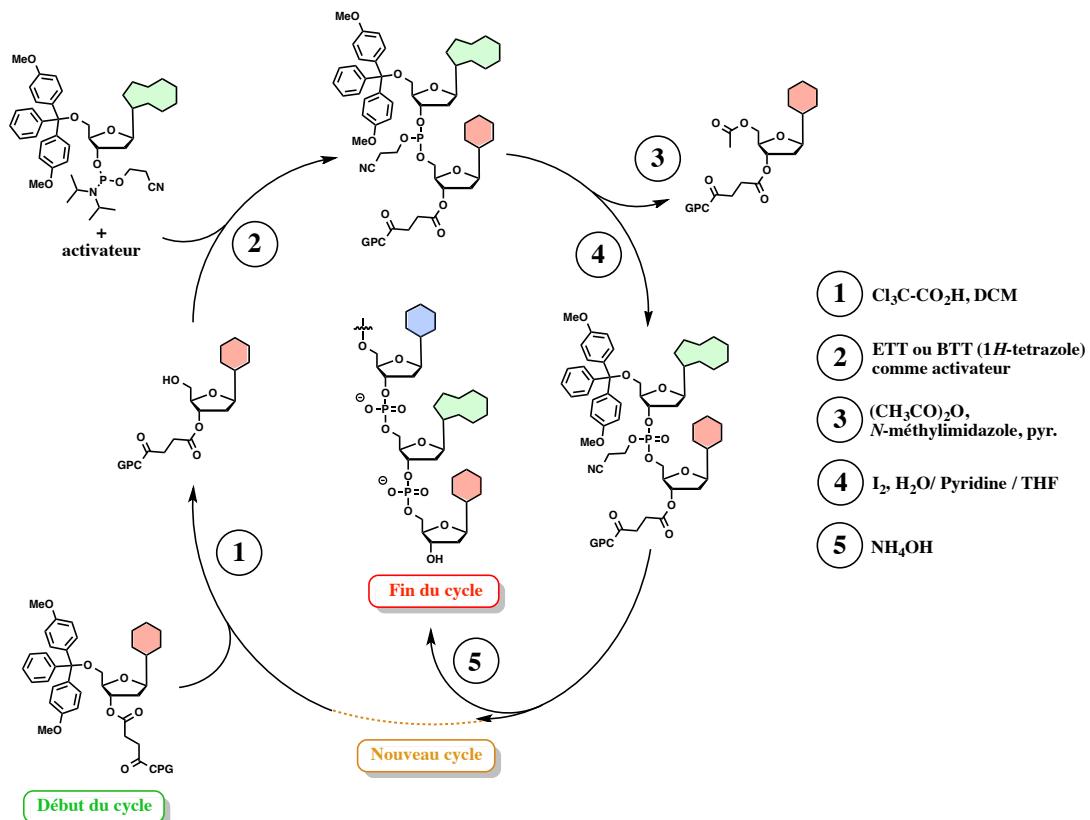
De nombreuses années ont été nécessaire avant de pouvoir synthétiser des ODNs de manière efficace. Depuis les tout premiers travaux d'*Alexander Todd* en 1955,¹³⁵ un grand nombre de techniques ont vu le jour. Parmi les plus importantes, citons les voies des phosphoramidites, des *H*-phosphonates,¹³⁶ des phosphodiesters¹³⁷ et des phosphotriesters¹³⁸. Chacune d'entre elle a initialement été développée en solution afin de mettre au point les conditions opératoires, avant d'être transposée sur support solide.¹³⁹

Aujourd’hui, seules les deux premières sont encore employées, les deux autres étant beaucoup trop contraignantes à mettre en oeuvre (rendements modérés, réactions parasites et problèmes de purification). Notons que la voie des *H*-phosphonates n'est qu'une alternative à la chimie des phosphoramidites, la méthode la plus couramment rencontrée.¹³⁶⁻¹³⁹

Cette chimie des phosphoramidites a été introduite en 1981 par *Beaucage et Caruthers*¹⁴⁰ qui ont reporté l'utilisation de ces dérivés du P(III) pour la synthèse d'ODN. Elle repose sur l'utilisation de nucléosides protégés en 5' par un groupement 4,4'-diméthoxytrityle (DMT), et en 3' par un groupement 2-cyanoéthyl *N,N*'diisopropylphosphoramidite (CEP) [cf. Figure 61]. Ces composés (phosphoramidites) sont extrêmement sensibles en présence d'espèces nucléophiles lorsqu'ils sont activés en milieu faiblement acide. Des conditions de préparation et de synthèse anhydres sont donc indispensables afin d'assurer les rendements de couplage les plus élevés et les amines primaires réactives des nucléobases doivent être protégées, afin d'éviter la synthèse d'ODNs branchés [cf. Figure 61].¹⁴¹

**Figure 61.** Structure des phosphoramidites.

McBride et Caruthers ¹⁴² établirent la méthodologie de synthèse des ODNs sur support solide que l'on connaît aujourd'hui [cf. Figure 62]. Elle se déroule sur un support solide CPG (Controlled-Pore Glass) ou polystyrène, dans la direction 3' → 5', à l'inverse de la synthèse enzymatique catalysée par l'ADN polymérase III (cf. I.A.2.iii).

**Figure 62.** Cycle de synthèse des oligonucléotides - Voie Phosphoramidite

Elle se compose de quatre étapes, qui s'enchaînent jusqu'à obtention de l'ODN de longueur souhaitée :

1. Déprotection (détirtilation) du DMT porté par le premier nucléotide, lié au support CPG par l'intermédiaire d'un bras succinique (acide trichloroacétique à 3 % dans du DCM).

2. Activation du phosphoramidite du nucléotide suivant (*5'-O-DMT* et *3'-O-CEP*), en présence d'un dérivé de *1-H-tétrazole* (ETT ou BTT)¹⁴³ puis couplage. L'hydroxyle 5', libéré à l'étape 1, attaque l'espèce activée afin de donner un dinucléotide, dont les deux oses sont reliés entre eux par un lien phosphite.
3. Coiffage (acétylation) des 5'-OH n'ayant pas réagit avec de l'anhydride acétique, activé par du *N*-méthylimidazole dans la pyridine, afin de masquer leur réactivité.
4. Oxydation du phosphore (III) du pont phosphite, en présence d'iode I₂ en milieu alcalin aqueux, afin de donner le phosphotriester.

A partir de là, le cycle recommence par la déprotection du DMT (1) présent sur le nucléotide récemment ajouté. Une fois tous les cycles achevés et la synthèse terminée, il est possible de conserver le DMT final, ou de le retirer par un traitement acide.

5. L'oligomère est finalement clivé de son support solide par traitement avec une solution aqueuse d'ammoniac à 30 %. Par la même occasion, les groupements protecteurs des différentes bases sont également déprotégés et le 2-cyanoéthyl du phosphate est clivé.^{141, 142}

Cette technique peut également être adaptée à la synthèse d'ARNs qui diffèrent de l'ADN par la présence, en position 2' du sucre, d'un groupement hydroxyle (cf. I.A.1.i). Ce dernier les rend plus instables que des brins d'ADN. Sa protection, par un groupement silylé par exemple, est recommandée afin de masquer sa réactivité.¹⁴⁴

Une fois l'ODN (ou ORN) obtenu, diverses méthodes s'offrent au chimiste afin de le purifier. La méthode la plus répandue est l'HPLC en phase inverse. Conserver le DMT terminal permet de rendre l'ODN plus lipophile et facilite sa purification. D'autres méthodes comme la chromatographie ionique ou l'électrophorèse, plus résolutives, peuvent aussi être employées. Elles nécessitent de travailler sans DMT et de dessaler l'ODN une fois purifié.

Le schéma de synthèse des ODNs illustré sur la Figure 62 n'est qu'un schéma général. De nombreuses variantes existent. En effet, Les conditions de déprotection présentées ici ne sont pas toujours adaptées à la synthèse des ARNs ou compatibles avec les modifications apportées aux nucléosides. Il y a quelques années de cela, la chimie PAC (ou « *UltraMILD* ») a par exemple été développée afin de permettre la déprotection des groupements protecteurs des nucléobases dans des conditions plus douces (0.05 M K₂CO₃ dans le méthanol). D'un autre côté, certains analogues, comme la 8-Aryl-dG, sont sensibles aux conditions acides utilisées pour déprotéger le DMT. Ainsi, le 5'-*O*-2,7-diméthylpixyle, encore plus labile que le DMT, a été mis au point.¹⁴⁵

2. Fonctionnalisation des ODNs

De nombreux chercheurs s'intéressent depuis des années à la fonctionnalisation des ODNs, afin de les utiliser comme capteurs (fluorescents, paramagnétiques, redox, ...), ou comme agents thérapeutiques ou diagnostiques. Dans cette optique, un grand nombre

de motifs divers et variés ont été incorporés dans des ODNs.^{134b,d} Deux techniques s'offrent à eux : l'approche ‘pré-synthétique’ et l’approche ‘post-synthétique’.

i. Approche pré-synthétique

L’approche ‘pré-synthétique’ consiste à synthétiser le phosphoramidite, le phosphonate ou le triphosphate de l’analogue de nucléoside afin de pouvoir l’incorporer au cours de la synthèse, chimique ou enzymatique, de l’ODN [cf. Figure 63]. La modification qui caractérise l’analogue de nucléoside peut être extrêmement variée, allant d’un atome (ex. S) à une molécule dont la taille est supérieure au nucléoside lui-même (ex : rhodamine, porphyrine, etc..). Elle peut être attachée au sucre, au phosphate, à la base, ou substituer cette dernière. Dans tous les cas, cette entité chimique doit être stable dans les conditions de synthèse afin d’éviter sa dégradation. La synthèse enzymatique permet dans certains cas de palier aux inconvénients de la synthèse chimique car elle est réalisée dans des conditions bien plus douces. Cependant elle est nettement moins modulaire quant à la nature de l’analogue de nucléotide qui peut être incorporé.

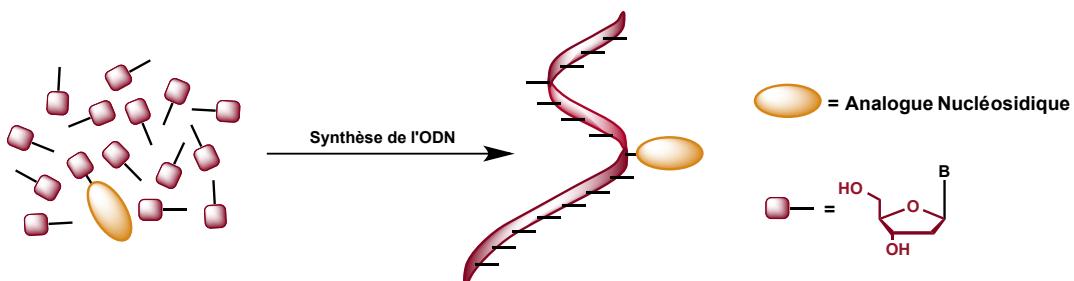


Figure 63. Représentation schématique de l’approche pré-synthétique.

ii. Approche post-synthétique

L’approche ‘post-synthétique’ consiste, elle, à synthétiser un ODN porteur d’un groupement réactif (amine,¹⁴⁶ thiol,¹⁴⁶ azoture,¹⁴⁷ alcyne,¹⁴⁸ halogène,^{149 & 150 ...}),^{134b,d, 151} qui servira de point d’accroche à la fonctionnalisation. L’analogue de nucléotide porteur de ce groupement réactif est tout d’abord incorporé dans le brin d’ADN au cours de sa synthèse. Ensuite, la séquence est engagée dans une réaction de couplage [cf. Figure 64] afin d’introduire l’entité chimique au niveau du point d’ancrage.

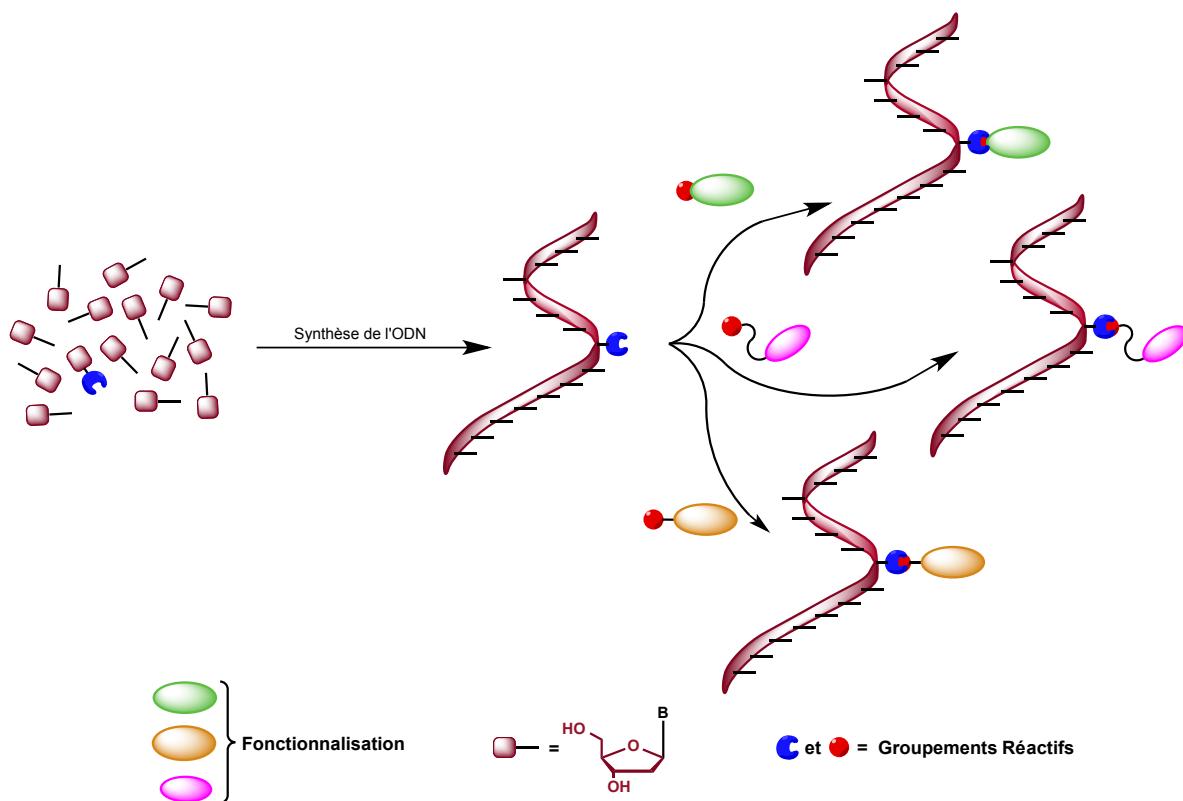


Figure 64. Représentation schématique de l'approche post-synthétique.

De nombreux travaux font référence à ce type d'approche dans la littérature car elles ont des avantages par rapport à l'approche ‘pré-synthétique’ :

- Introduction d'entités chimiques incompatibles avec une approche pré-synthétique (instabilité),
- Une grande modularité,
- Un nombre d'étapes de synthèse et de purifications réduit,
- De bons rendements de réaction,
- Une élimination de l'excès de réactif facilité par l'utilisation du support solide,
- et l'utilisation de moins de réactifs (20 équivalents de phosphoramidite en moyenne pour la méthode ‘pré-synthétique’ chimique).

B. Marquages Fluorescents de l'ADN

Le marquage de l'ADN par des sondes fluorescentes couvre un champ vaste d'applications, plus particulièrement en biologie et en biochimie. Il permet le séquençage de l'ADN à haut débit, de détecter l'ADN en imagerie cellulaire (DAPI),¹⁵² du polymorphisme nucléosidique simple base (SNP) ou des produits de PCR, d'étudier des mécanismes fondamentaux régis par des interactions acide nucléique / protéine et bien d'autres applications.

En 1971, *Daniels* a étudié les propriétés de fluorescence des nucléobases naturelles canoniques (A, G, C et U).¹⁵³ Leurs rendements quantiques extrêmement faibles (respectivement $2,6 \cdot 10^{-4}$, $3,0 \cdot 10^{-4}$, $0,8 \cdot 10^{-4}$ et $0,5 \cdot 10^{-4}$) ont conduit les chercheurs à se

tourner vers deux voies alternatives afin de rendre fluorescent un acide nucléique : le marquage non-covalent et le marquage covalent.

1. Diversité des Sondes

i. Marquage non covalent

Une grande variété d'espèces fluorescentes [cf. Figure 65 - liste non exhaustive] pouvant spontanément se lier à l'ADN ont été développées afin de pouvoir le rendre visible dans des conditions bien spécifiques. La plupart de ces sondes sont utilisées en biologie lors d'étude en microscopie, en cytométrie de flux, en PCR ou afin de révéler les molécules d'ADN après séparations sur gel d'électrophorèse.

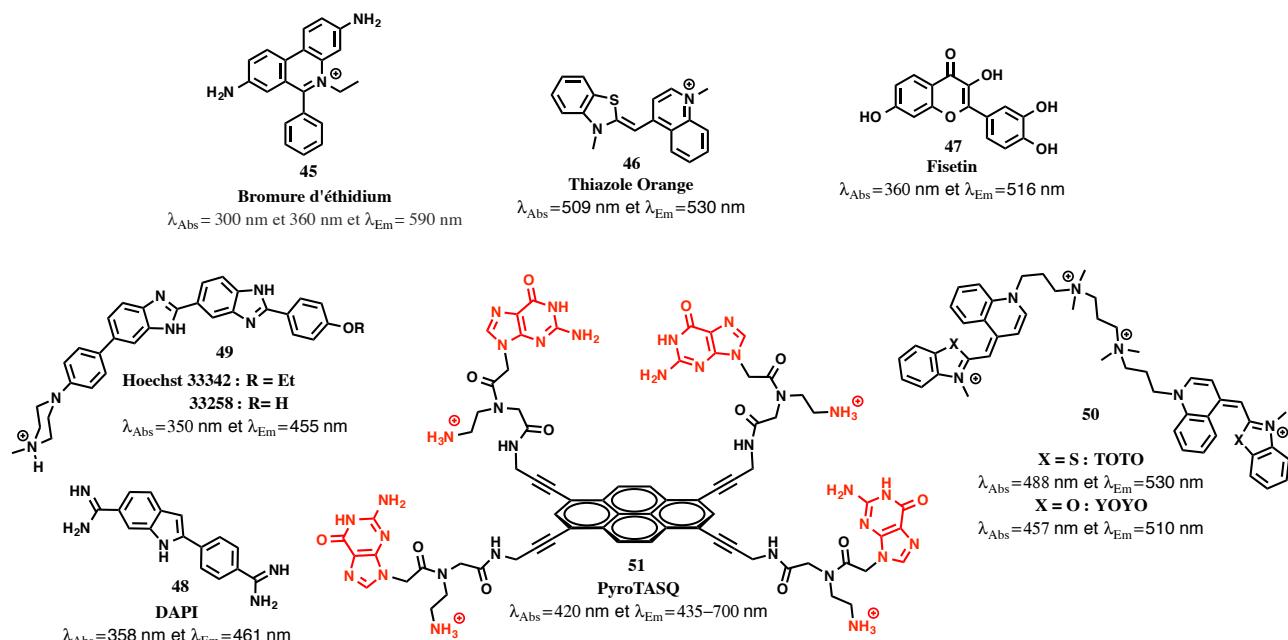


Figure 65. Structures de molécules marquant l'ADN de manière non covalente.

Ces composés peuvent interagir selon deux grands modes d'interaction avec l'ADN : l'intercalation entre les bases nucléiques et la liaison, principalement dans le petit sillon de l'ADN.

i) Intercalants de l'ADN

Parmi ces sondes, citons le bromure d'éthidium **45** (BE), l'un des intercalants les plus utilisés. Il permet de visualiser les brins d'ADN après séparation sur gel d'électrophorèse. Faiblement fluorescent dans l'eau, après intercalation dans le duplexe l'intensité de fluorescence et le temps de vie sont multipliés respectivement par 30 et quasiment 10.¹⁵⁴

Initialement synthétisée afin d'étudier les réticulocytes (globules rouges juvéniles), le thiazole orange **46** (TO) présente une forte affinité pour les acides nucléiques, tant pour l'ADN que l'ARN.¹⁵⁵ Les systèmes conjugués étendus comme le TO présentent des maximums d'absorption et d'émission fortement décalés vers le rouge ($\lambda_{\text{abs}} = 509 \text{ nm}$ et $\lambda_{\text{abs}} = 530 \text{ nm}$ pour **46**). Comme pour l'entité précédente, son rendement quantique

augmente lors de l'interaction sonde / acide nucléique, mais d'un facteur 3000 cette fois-ci.^{147f}

Sur le même principe, les fluorochromes TOTO et YOYO **50**, dimères de TO **46** et d'oxazole jaune (YO) respectivement, sont utilisés pour révéler l'ADN sur les gels d'électrophorèse. Ils présentent de bons rendements quantiques (0,52 et 0,34) et une affinité supérieure à la fois pour l'ADN et l'ARN.¹⁵⁶

De nouvelles sondes, récemment apparues, permettent, non plus de révéler l'ADN, mais d'étudier des phénomènes plus complexes, tels que le repliement des acides nucléiques.¹⁵⁷ Parmi ces composés, citons le PyroTASQ **51** développé par *Languette et al.*¹⁵⁸ Il s'agit d'un ligand fluorescent dit « intelligent », permettant de détecter et de suivre le repliement des G-quadruplexes dans la cellule. À l'état ouvert (tel que représenté sur la Figure 65), la proximité spatiale des guanines inhibe la fluorescence du noyau pyrène par transfert d'électron intramoléculaire photoinduit (iPET). Lors de la liaison avec un quadruplex, le mécanisme de transfert d'électrons est inhibé suite au repositionnement des guanines et leurs interactions avec le G-quadruplexe. C'est ainsi que le pyrène récupère son émission de fluorescence au contact du G-quartet, permettant ainsi sa détection et son suivi.

ii) Interaction avec le petit sillon de l'ADN

Le 4',6-diamidino-2-phenylindole **48** (DAPI), se lie au niveau de régions riches en paires de bases A-T de l'ADN, dans le petit sillon et avec un angle de 45° par rapport à l'axe principal du duplexe [cf. Figure 66 - PDB : [1D30](#)].¹⁵⁹ Ces maximums d'absorption et d'émission (358 et 461 nm) permettent de l'exciter dans des domaines où les bases naturelles sont transparentes. A l'instar du BE **45** ou du TO **46**, son rendement quantique est très faible dans l'eau et augmente de manière significative lors de sa liaison à l'ADN.¹⁶⁰ Le DAPI, ayant la capacité de passer au travers des membranes cellulaire et nucléaire, est essentiellement utilisé en microscopie de fluorescence. En se liant à l'ADN (nucléaire ou mitochondrial), il permet de le révéler sous la forme de points bleus fluorescents. Il peut également être utilisé simultanément en présence d'autres fluorophores afin de réaliser des études multicanaux de cellules.

Les dérivés de Hoechst **49** (33342 et 33258) jouent un rôle similaire au DAPI **48**. Ils sont moins toxiques que ce dernier et leur fluorescence est inhibée en présence d'iode.¹⁶¹

Tout comme le PyroTASQ **51**, la fisétine **47**, dérivé polyhydroxylé de la famille des 3-hydroxyflavones (cf. III.B.2.iii), permet de suivre le repliement des acides nucléiques. Son caractère ratiométrique a récemment permis d'étudier différentes conformations de l'ADN, dont des i-motifs et G-quadruplexes.¹⁶²

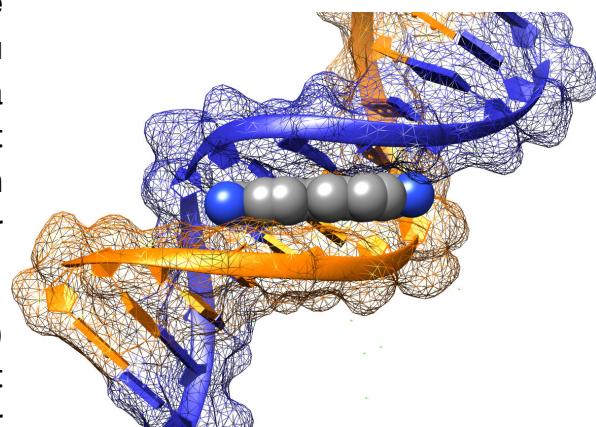


Figure 66. Le DAPI lié à un duplexe l'ADN.^{159b}

Afin d'augmenter la sélectivité de ces rapporteurs pour l'ADN, il est possible de les coupler à des entités affines pour les acides nucléiques, comme la spermine par exemple (cf. TOTO et YOYO 50 ou la chromone 44).^{126a, 156, 163} Cependant, n'étant pas liées de manière covalente, leur localisation au sein de la séquence n'est pas connu précisément. Dans un certain nombre de cas, ce genre de sonde peut même induire des distorsions au niveau de l'ADN.¹⁶⁴ Enfin, ce type de marquage est essentiellement utilisé pour visualiser les acides nucléiques, même si l'introduction de nouvelles sondes,^{158, 162} permet désormais d'étudier des évènements plus complexes, comme leur repliement.

ii. Marquage covalent

Un avantage notable du marquage covalent est qu'il présente une gamme plus large d'applications que le marquage non covalent.¹⁶⁵ De plus, la position d'ancrage de la sonde est connue avec précision, car incorporée lors de la synthèse de l'ODN de façon site-spécifique. Nous avons vu précédemment que le chimiste a à sa disposition de nombreuses techniques lui permettant de fonctionnaliser de manière covalente des ODNs, et donc d'introduire une entité fluorescente (cf. IV.A.2.). Schématiquement, trois modes d'attaches différents peuvent être distingués [cf. Figure 67]. Il est tout d'abord possible d'ancrer la sonde sur le squelette nucléosidique à l'aide d'un bras espaceur long, afin de lui offrir une certaine liberté de mouvement [cf. Figure 67A]. Ensuite, une nucléobase peut être remplacée par le rapporteur fluorescent, afin de pouvoir l'orienter vers l'intérieur du duplexe [cf. Figure 67B]. Enfin, le fluorophore peut être directement couplé à l'une des nucléobases, à l'aide d'un bras espaceur court, la plupart du temps insaturé, permettant le contrôle de son positionnement au niveau des sillons de l'ADN [cf. Figure 67C]. Des informations complémentaires peuvent être trouvées dans les revues de *Tor* et *Wilhelmsen* sur les différentes stratégies de marquage et la grande variété de fluorophores couplés à de l'ADN.

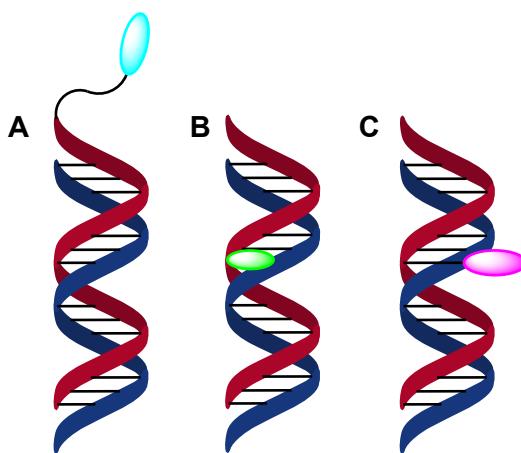


Figure 67. Sondes fluorescentes attachées à l'ADN par (A) un bras espaceur long, (B) substitution d'une nucléobase naturelle et (C) et un bras espaceur court et rigide.

i) Bras espaceurs longs

Le bras espaceur ‘long’, utilisé pour ancrer des sondes fluorescentes sur un ODN, est communément une chaîne flexible de 6 à 12 atomes (hydrocarbonée, polyéthylène-

glycol - PEG, ...). Les positions d'ancrage privilégiées sont les positions 3', 5' terminales de l'ODN et la nucléobase [cf. Figure 68].

Une grande diversité de fluorophores peuvent être attachés à l'extrémité du bras espaceur. Certains d'entre eux sont disponibles commercialement, comme les deux dérivés de la fluorescéine **52** et **53**, ou les Quasar **55** [cf. Figure 68].¹⁴¹

Citons également l'oligothiophènes **54**¹⁶⁶ ainsi que le BODIPY **56**,¹⁶⁷ pouvant tous deux être incorporés dans l'ODN en position 5'-terminale grâce à leurs phosphoramidites respectifs [cf. Figure 68].

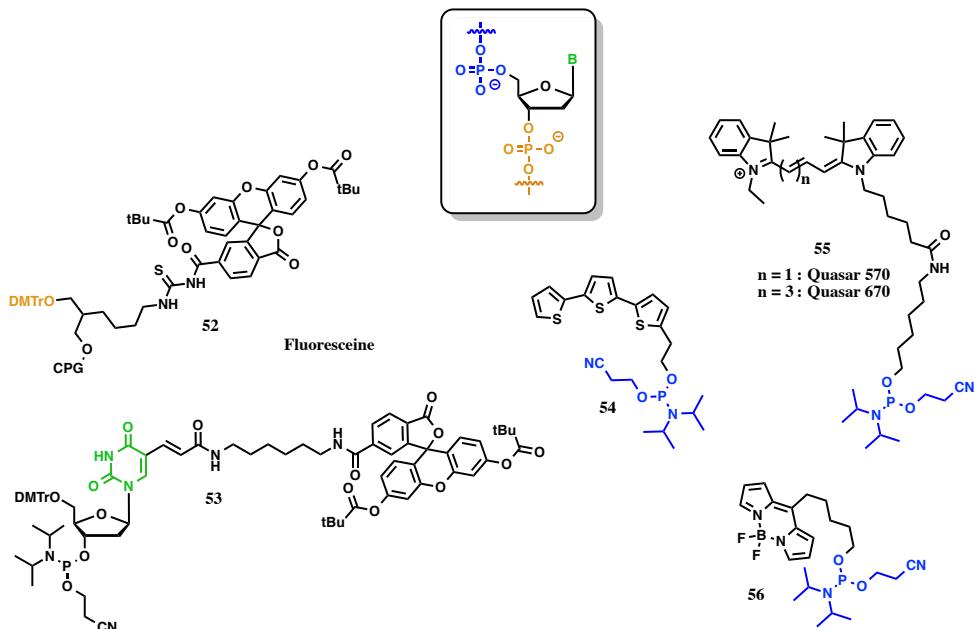


Figure 68. Exemples de sondes marquant l'ADN via un bras espaceur.

Ce type de marquage a trouvé un grand nombre d'applications comme pour le séquençage du génome, la recherche d'allèles par PCR ou de l'imagerie cellulaire. Cependant ce lien flexible confère au fluorophore une grande mobilité et génère une incertitude quant à sa position et à son orientation. De plus, le fluorophore, relativement volumineux, peut interagir avec l'ADN ou avec les protéines du milieu et interférer avec les systèmes étudiés, ce qui soulève des problèmes pour l'interprétation des données expérimentales.

ii) Polycycles (hétéro)aromatiques connectés au sucre

La capacité à s'intercaler dans l'ADN de certains fluorophores polyaromatiques a été exploitée afin de développer des capteurs à base d'ADN [cf. Figure 69]. En principe, ces sondes planes n'établissent pas de liaison hydrogène de type Watson-Crick avec les bases naturelles, mais s'intercalent entre les bases adjacentes grâce à des interactions enthalpiques (van der Waals et recouvrement π-π), et entropiques (hydrophobes) favorables. L'intercalant peut substituer une base naturelle ou être attaché au nucléoside. La dimension de ces fluorophores varie généralement entre celle d'une base naturelle et d'une paire de bases. Dans la plupart des cas, ils sont liés au carbone anomérique du

sucre, leur permettant ainsi de se positionner à l'intérieur du duplexe d'ADN. Ceci permet de connaître à la fois la localisation de la sonde et son orientation.^{165b, 168}

- Polycycles aromatiques hydrocarbonés

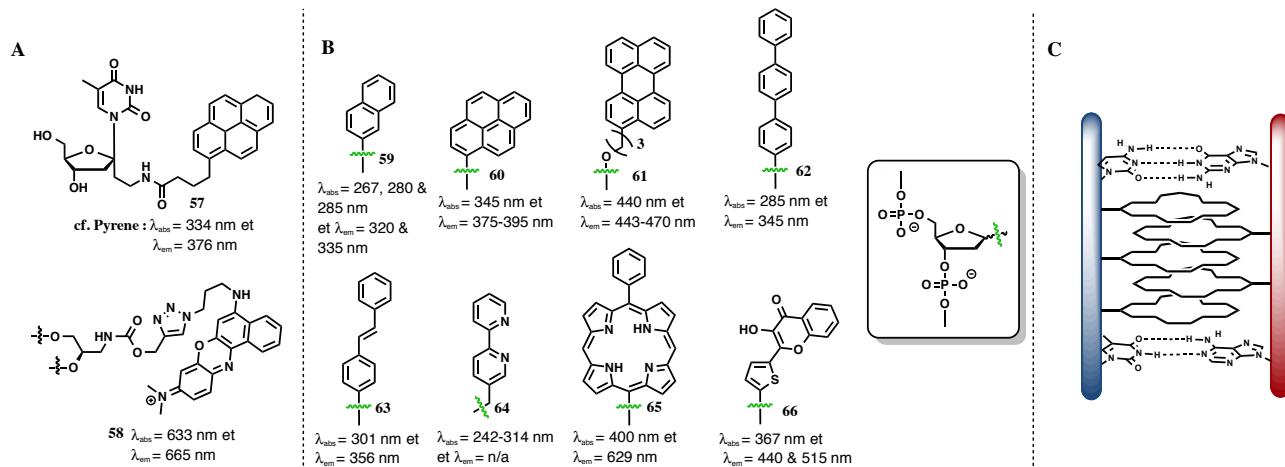


Figure 69. (A et B) Exemples de bases polycycliques aromatiques et (C) représentation schématique du positionnement de ces sondes dans un duplexe.¹⁶⁹

Un grand nombre de ces composés ont été synthétisés dont certains, comme les composés **57** et **58** par approche ‘post-synthétique’ [cf. Figure 69A].¹⁷⁰ Toutefois, la majorité de ces dérivés a nécessité une approche ‘pré-synthétique’ pour la fixation du fluorophore sur le désoxyribose.¹⁷¹ Le naphtalène **59** et le pyrène **60** ont été parmi les premiers composés de cette classe à être décrits, suivis par le perylène **61**, le triphényle **62** et le stylbène **63**.¹⁷² Ces composés présentent tous des maxima d’absorption décalés vers le rouge.

Kool développa, par chimie combinatoire, une librairie d’ODNs introduisant plusieurs de ces dérivés polycycliques aromatiques les uns à la suite des autres. Il les appela ODFs, pour « oligodeoxyfluorosides ». La proximité spatiale de ces analogues nucléosidiques [cf. Figure 69C] permettait la formation d’excimères ou exciplèxes (cf. II.B.1.i.ii.b). Les ODFs pouvaient ainsi être excités à une longueur d’onde unique (354 nm) et la couleur émise était spécifique à la séquence. Sélectionnés à partir d’une chimiothèque de 256 composés générée à partir de seulement quatre chromophores différents (4-(dicyanomethylene)-2-methyl-6(p-dimethylaminostyryl)-4H-pyran, pyrène, benzopyrène et perylène), une vingtaine d’entre eux permet ainsi de couvrir l’ensemble du spectre visible [cf. Figure 70]. Autre point important concernant ces ODFs, ils présentaient de larges déplacements de Stokes (jusqu’à 200 nm).^{172f, 173}

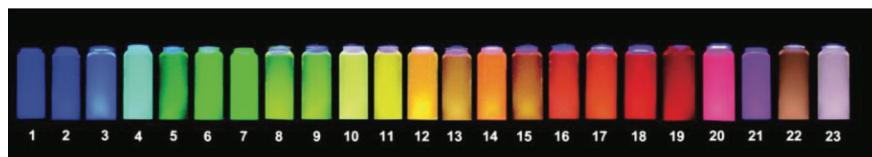


Figure 70. 23 différents ODFs recouvrant le spectre visible (excitation commune à 354 nm).

Cette chimie a été étendue à l’introduction de dérivés hétérocycliques tels les 2,2'-bipyridine (**64**) et porphyrine (**65**) [cf. Figure 69B].¹⁷⁴ Ces composés ont été nommés

« ligandosides ». A l'image des fluoroionophores (II.B.2.ii), ils ont la capacité de chélater des ions métalliques (Cu^+ , Pd^{2+} , Ag^+) afin, par exemple, de pouvoir les titrer. Enfin, la thiénylchromone **66**, synthétisée au laboratoire, est une sonde ratiométrique à émission duale capable de détecter des interactions ADN / protéine par un changement de couleur.^{126b-e}

Le comportement de ces composés, après incorporation, varie d'une espèce à l'autre. Le pyrène **60**, par exemple, est fortement éteint à cause des interactions avec les bases adjacentes. De son côté, la porphyrine **65** voit son rendement quantique peu fluctuer, quelles que soit les bases voisines. Enfin, le perylène **61** subit, lui, une légère augmentation de son rendement quantique, beaucoup plus marquée lorsque deux perylènes sont incorporés côte à côte.^{169, 175}

- Analogues de bases naturelles

Il a été démontré que des modifications structurales mineures des nucléobases naturelles permettait d'obtenir des nucléosides fortement émissifs. La plupart des dérivés de cette famille ont des tailles similaires à celles des bases naturelles et conservent la capacité à établir les mêmes liaisons Watson-Crick (ou Watson-Crick inverse) [cf. Figure 71]. La structure hélicoïdale de l'ADN-B (et/ou A) est en général peu ou pas perturbée.

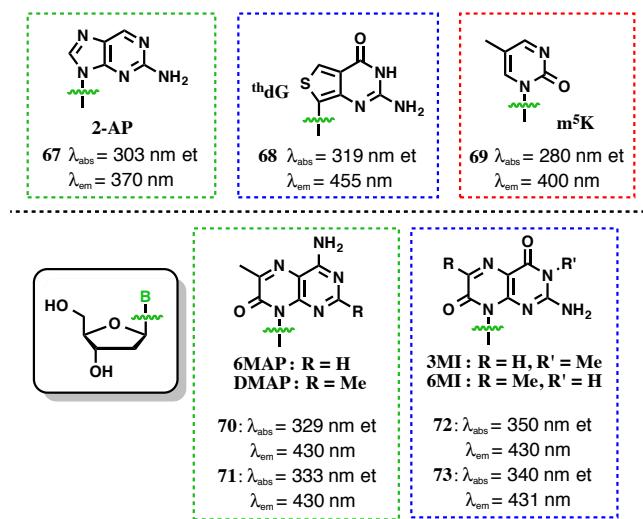


Figure 71. Divers analogues de bases nucléiques.

La 2-Aminopurine **67** (2-AP), la thiéno guanine **68** (thdG) et le composé **69** (5-méthylpyrimidin-2-one désoxyriboside - m⁵K) sont respectivement des analogues d'adénine, de guanine et de 5mC. Les modifications structurales subies par ces composés sont :

- une transposition du groupement amino pour **67**,¹⁷⁶
 - une substitution de l'imidazole par un thiényle pour la thdG **68**,¹⁷⁷
 - une élimination du groupement amino pour **69**.¹⁷⁸

Avec un maximum d'émission décalé vers le rouge, par rapport au domaine d'absorption des acides nucléiques, et un déplacement de Stokes significatif, ces composés peuvent

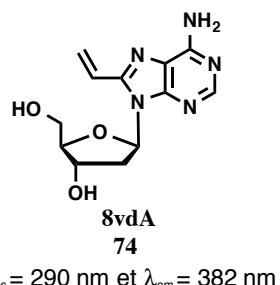
être excités sélectivement et facilement détectés. De plus, leurs rendements quantiques dans l'eau sont appréciables : $\Phi_{67} = 0,68$, $\Phi_{68} = 0,58$ et $\Phi_{69} = 0,13$.^{169, 175, 176, 177, 178, 179}

Les dérivés de ptéridines présentent également une grande homologie avec les bases puriques naturelles A (**70** et **71**) et G (**72** et **73**). Ils peuvent être impliqués dans les mêmes interactions hydrogènes que ces dernières. Leurs rendements quantiques sont très élevés (0,77–0,88 pour les dérivés de A et 0,39–0,48 pour ceux de G) et ils présentent également de forts déplacements de Stokes 75–120 nm (pour **71** et **73**).^{169, 175, 176}

La majorité des ces sondes est fortement éteinte une fois qu'elles sont incorporées dans des ODNs. Elles sont la plupart du temps utilisées comme sondes « *Turn Off/On* » (cf. II. 2.1.i.i.).^{175, 180}

- Extensions de conjugaison des bases naturelles

Outre ces changements structuraux, les propriétés photophysiques des bases naturelles peuvent également être améliorées en augmentant la conjugaison au niveau du noyau aromatique de la base. En effet, cette manière de procéder se traduit par un décalage des maxima d'absorption et d'émission vers le rouge et une augmentation, parfois très importante, du rendement quantique (par rapport aux bases naturelles).



$\lambda_{\text{abs}} = 290 \text{ nm}$ et $\lambda_{\text{em}} = 382 \text{ nm}$

Figure 72. Structure de la 8vdA.

Un exemple type est celui de la 8-vinyladénosine **74** (8vdA), synthétisée au sein de notre laboratoire [cf. Figure 72].¹⁸¹ En plus d'être un meilleur mime de l'adénine que la 2-AP **67**, elle est deux fois plus brillante que celle-ci. Cependant, elle présente des caractéristiques similaires à savoir une forte extinction de fluorescence une fois incorporée dans les ODNs, à cause de l'empilement avec les bases adjacentes.

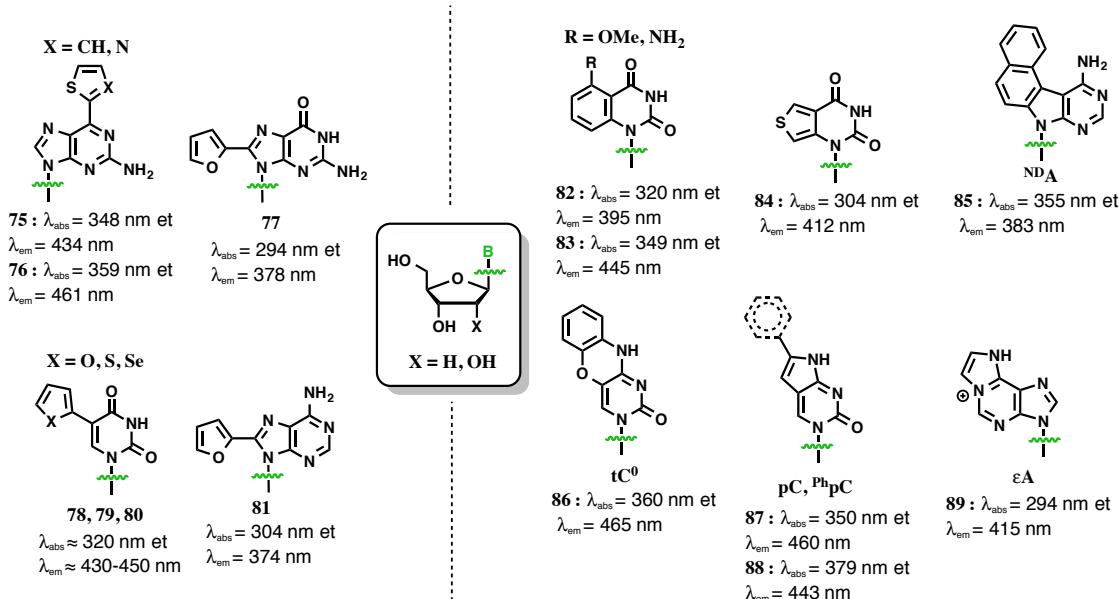


Figure 73. Analogues fluorescents présentant des noyaux (hétéro)aromatiques.

La conjugaison peut également être augmentée en attachant des cycles aromatiques à la base. Par exemple, des hétérocycles à 5 chaînons ont été couplés à la nucléobase en position 5 de l'uridine (**78–80**), 6 de la 2-AP (**75** et **76**) ou 8 des purines naturelles (**77**, **81**) [cf. Figure 73]. Les dérivés de pyrimidines sont faiblement fluorescents ($\Phi_{78} = 0,01$, $\Phi_{79} = 0,035$ et $\Phi_{80} = 0,014$).^{175, 182} En revanche, les purines **75–77** et **81** [cf. Figure 73], elles, donnent de bons rendements quantiques ($\Phi_{75} = 0,41$, $\Phi_{76} = 0,46$, $\Phi_{77} = 0,57$ et $\Phi_{81} = 0,69$).^{175, 183} Seules les 2 dérivés de 2-AP et les pyrimidines, présentant un plus grand intérêt pour leur découvreurs, furent introduites dans des ODNs. Et, comme dans bons nombres de cas précédents, leur émission de fluorescence a été inhibée une fois incorporées. *A noter que les dérivés de 2-AP (**75** et **76**) ont été développés par Hirao pour s'apparier sélectivement avec une base artificielle, car la gène stérique induite par l'hétérocycle à 5 chaînons empêche l'appariement avec une T (cf. IV.A.).*^{130c, 183a}

Notons également que des dérivés benzothiazole,¹⁸⁴ indoles,¹⁸⁵ ou encore benzothiophène¹⁸⁶ ont également été synthétisés par les équipes de Matsuda¹⁸⁴ et de Srivatsan.^{185, 186}

Une autre approche consiste à juxtaposer un cycle ou hétérocycle aromatique à la nucléobase, tels que des dérivés de phényle, thiophène ou anthracène (**82** à **85**) [cf. Figure 73].¹⁸⁷ D'autres exemples font état d'un hétérocycle obtenu après cyclisation au niveau de l'amine n°4 de la cytidine ou 6 de l'adénine [cf. Figure 73]. C'est le cas par exemple du tC⁰ **86** (et de son dérivé soufré tC - $\lambda_{\text{Abs}} = 375 \text{ nm}$ et $\lambda_{\text{Em}} = 505 \text{ nm}$),¹⁸⁸ des pyrrolo-dC **87**¹⁷⁶ et **88**¹⁸⁹ ou de l' ϵ A **89**.¹⁹⁰ Mise à part ^{NDA} **85**, ces composés présentent des rendements quantiques intéressants en version nucléosidique ($\approx 0,2$ pour **82**, **83**, **86–88**, $0,48$ pour **84** et $0,56$ pour **89**).¹⁷⁵ Cependant, seul tC ne subit aucune inhibition de fluorescence une fois incorporé dans un ODN.^{188b}

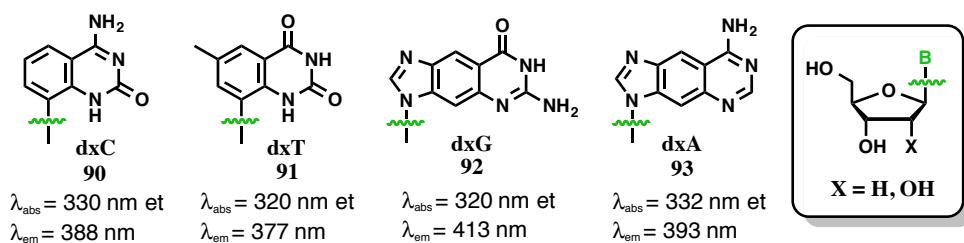


Figure 74. Extension de conjugaison des bases naturelles.

Enfin, une quatrième et dernière approche consiste à augmenter la taille de la base naturelle en insérant un noyau benzénique (ou naphtalène) entre le sucre et la base pyrimidique ou entre le noyau imidazole et le motif pyrimidine des bases puriques [cf. Figure 74].¹⁹¹

Les dérivés benzoN **90** à **93** ($N = A, T, C, G$) représentés en Figure 74 présentent des déplacement de Stokes importants (50 à 80 nm) et de bons rendements quantiques (0,3 à 0,6). Lors de leur incorporation dans des ODNs, ils donnent lieu à des paires de base plus larges que celles présentes dans l'ADN B classique, l'ajout du noyau benzénique augmentant la taille de la nucléobase de 2,4 Å. Kool appelle ce type d'ODN de l'ADNx (*xDNA* - pour ‘extended DNA’).^{169, 191b} Des duplexes hybrides *xN–N* (*xN–N* = *xA–T*, *xT–A*, *xC–G* et *xG–C*) ont ainsi été synthétisés et étudiés. Ces duplexes ADNx / ADNx et ADN / ADNx présentent une stabilité supérieure à celle des duplexes ADN / ADN conventionnels. Ceci peut facilement s’expliquer par un recouvrement supérieur des bases *dxN* grâce à la présence du noyau benzénique supplémentaire.^{169, 191} Cependant, leur fluorescence est plus ou moins inhibée une fois incorporés dans de l'ADN. Par exemple, les bases puriques inhibent leur émission de fluorescence de manière bien plus importante que les bases pyrimidiques (pour *dxA 93* : compris entre 0,05 et 0,2 respectivement dans l'ADN, contre 0,44 pour le monomère ; pour *dxT 91* : compris entre 0,01 et 0,33 dans l'ADN, contre 0,3 pour le monomère).¹⁹²

Ces travaux ont été par la suite complétés par l’élaboration d’autres analogues étendus. Ces composés présentent une autre géométrie, résultant du fait que le point d'accroche de la pyrimidine sur le noyau benzénique est différent. Ils ont conduit à des duplexes d'ADNy (*yDNA* - pour ‘wild DNA’).^{165a} Ces molécules démontrent des maxima d'émission encore plus décalés vers le rouge. Par exemple, *dyA* émet à 433 nm, soit 40 nm de plus que *dxA 93*.¹⁶⁹

Remarque : Lors de l’ajout d’un noyau naphtalène, la taille de la nucléobase augmente de 4,8 Å et l’ADN obtenu après incorporation de cette base est alors appelé ADNy. Seuls des exemples sur les bases pyrimidiques sont référencés.^{191b}

iii) Bras espaceurs court

Enfin, un chromophore ou un fluorophore peut être lié avec un bras espaceur court sur la nucléobase. Cette approche est en quelque sorte une combinaison des deux premières. En effet, le rapporteur est orienté vers l’extérieur du duplexe, comme avec un bras espaceur long. Cependant, sa petite taille et sa rigidité permettent de mieux contrôler

le positionnement et l'orientation du fluorophore, soit vers le grand sillon de l'ADN (cas le plus fréquent) soit vers le petit sillon. Comme dans l'approche précédente, la capacité à former des liaisons hydrogènes interbrins (Watson-Crick) est conservée. Si le lien est insaturé, le fluorophore (ou chromophore) et la base nucléique sont électroniquement couplés. Le bras espaceur insaturé est souvent un alcyne (plus rarement un alcène). Le couplage électronique entre la base nucléique et le rapporteur produit une nouvelle entité fluorescente dont les maxima d'absorption et d'émission sont décalées vers le rouge.¹⁹³

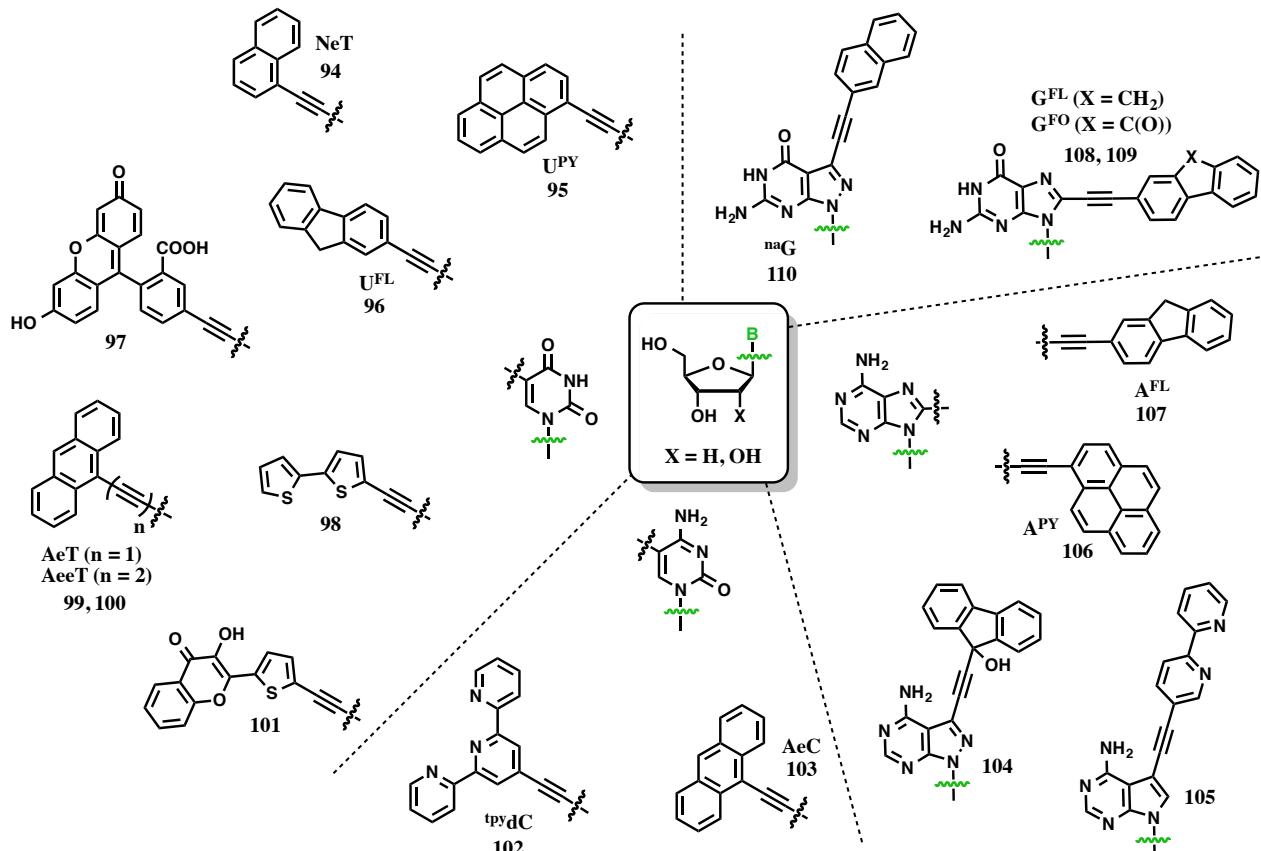


Figure 75. Sondes conjuguées aux bases naturelles.

Les positions de marquage utilisées sont le carbone n°5 pour les pyrimidines (**94** à **103**), le carbone 8 pour les purines naturelles (**106** à **109**) et le carbone 7 pour les dérivés 7-déazapurines (**104**, **105** et **110**) [cf. Figure 75]^{172a, 194} : Il est à noter que contrairement à la position 7, la modification en position 8 des purines va introduire une gêne stérique importante, défavorable à la formation d'une paire Watson-Crick stable dans l'ADN double brin. L'introduction en position 7 de la déazapurine est donc à privilégier lors de la synthèse d'analogues fluorescents devant conserver leur capacité à former des duplexes stables.

La Figure 75 illustre le grand nombre d'exemples de dérivés fluorescents synthétisés selon cette approche. Citons pour exemples, les motifs fluorescents issus du pyrène (**95**¹⁹⁵ et **106**¹⁹⁵), du fluorène (**96**,^{195b, 196} **107**,^{195b} **108–109**¹⁹⁷) et de l'anthracène (**99**, **100** et **103**).¹⁹⁸

La plupart des études dans la littérature font référence à des dérivés de l'uridine et de l'adénine, fonctionnalisés par des dérivés de pyrène **95** (U^{PY}) et **106** (A^{PY}). Le choix du

pyrène est essentiellement basé sur son fort rendement quantique. Par exemple U^{PY} **95** présente un rendement quantique de 0,81 dans le chloroforme et de 0,12 dans le méthanol, et son analogue adénosine A^{PY} **106** présente des rendements quantiques de 0,81 et 0,73, respectivement dans les mêmes solvants.¹⁹⁹

A noter aussi que des dérivés de la cytosine (**102** et **103**)²⁰⁰ et de la guanosine (**108–110**)^{70,201} ont également été préparés.

Le lien alcyne (ou alcène) permet d'éloigner le fluorophore de la nucléobase et, dans un duplexe parfaitement complémentaire, empêche l'empilement du fluorophore avec les bases adjacentes qui, dans bon nombre de cas, réduit le signal de fluorescence.¹⁹⁹

Toutes ces approches ont permis d'accéder à un ensemble d'analogues de nucléosides fluorescents très nombreux et d'une très grande diversité. Ils ont permis aux chimistes de marquer des nucléobases dans les ODNs afin d'en faire des rapporteurs fluorescents. Certaines applications de ces types de marquage vont à présent être passées en revue.

2. Applications

Les nucléosides fluorescents présentent un intérêt tout particulier en biologie moléculaire. Ils sont par exemple employés afin de déterminer l'enchaînement des acides nucléiques d'une séquence d'ADN donnée, de détecter des modifications de bases nucléiques (SNP) ou encore d'étudier des phénomènes cellulaires variés (changements de conformation, interactions intermoléculaires (petits ligands, protéines), ...).

Afin de pouvoir étudier ces derniers phénomènes, le fluorophore doit être incorporé de manière spécifique, au plus près du phénomène à étudier. De plus, le signal, et toutes les modifications l'affectant, doivent pourvoir être aisément observés afin de permettre un suivi optimal de l'objet étudié.

i. Séquençage de l'ADN (Méthode Sanger)

La méthode de *Frederick Sanger*, développée en 1977 permet de déterminer l'enchaînement des nucléotides dans une séquence d'ADN donnée.²⁰² Elle se déroule à la façon d'une synthèse enzymatique sélective et utilise, dans sa version améliorée, des nucléotides fluorescents à la place des nucléotides radioactifs initiaux.

A l'instar d'une PCR, la polymérisation de l'ADN est initiée par une amorce, complémentaire à une partie de la séquence d'ADN à séquencer. Les quatre 2'-désoxynucléotides naturels (dNTP - N = A, T, C, G - 0,5 mM), indispensables à la réaction, sont introduits dans le milieu. De plus, des dérivés dépourvus d'hydroxyle en 3' (2',3'-didésoxynucléotides ou ddNTP), permettant de bloquer l'élongation une fois incorporés, sont également ajoutés en faible concentration (0,005 mM). Ces ddNTP sont marqués avec des sondes fluorescentes émettant à des longueurs d'ondes différentes, afin de pouvoir les détecter sélectivement [cf. Figure 76A].

A chaque phase d'ajout d'un nucléotide, l'un des deux dérivés 2'-désoxy- ou 2',3'-didésoxynucléotide, est introduit de manière statistique dans la séquence. On obtient alors des séquences porteuses du dérivé naturel et d'autres terminées, par la présence du ddNTP fluorescent correspondant.²⁰² L'elongation se poursuit jusqu'à obtenir la séquence souhaitée. A ce moment là, le milieu contient autant de fragments terminées qu'il y avait de base N (N = A, T, C ou G) dans la séquence d'ADN analysée. Une séparation de ces fragments par électrophorèse (séparation en fonction de leur taille), couplée à un suivi par fluorescence, permet alors de déterminer l'enchaînement de ces bases N au sein de la séquence [cf. Figure 76B].²⁰²

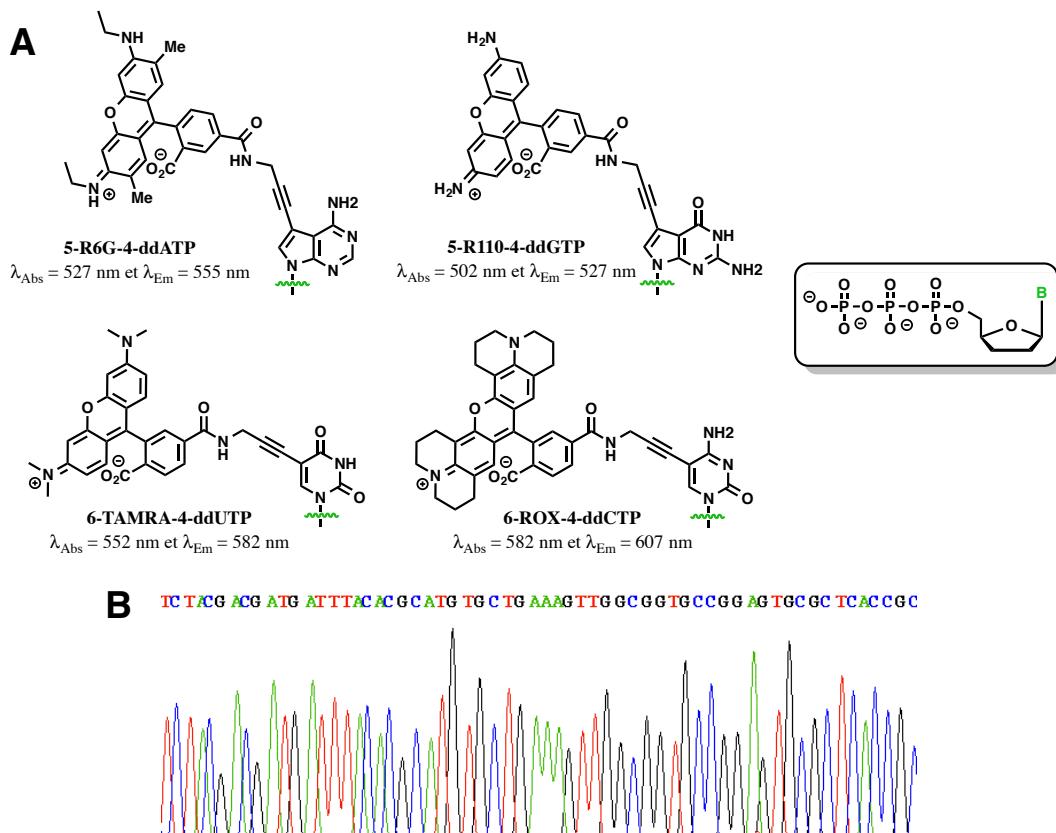


Figure 76. (A) Exemple de sondes pour marquer les ddXTP²⁰³ et (B) d'un électrophorégramme haut débit.

ii. Détection de polymorphisme mononucléosidique

Le polymorphisme mononucléosidique (ou SNP - ‘Single Nucleotide Polymorphism’) est une modification de la séquence d'ADN, survenant généralement au sein d'une population, pour laquelle un seul nucléotide (A, T, C ou G) diffère entre les membres de l'espèce biologique. Deux fragments d'ADN, provenant de deux individus et présentant un seul nucléotide différent (C remplacé par T, par exemple) sont appelés allèles. Les SNP peuvent se produire au niveau de régions codantes ou non codantes de gènes, ou de régions inter-géniques. Leur présence à l'intérieur d'une séquence codante ne modifie pas nécessairement l'enchaînement d'acides aminés de la protéine produite (cf. Tableau 1 - ANNEXE S1).

Ces SNP ont démontré un impact sur les différences dans les caractéristiques génétiques, la susceptibilité à la maladie et la réponse aux traitements médicamenteux. Un large

éventail de maladies humaines, par exemple l'anémie falciforme, la β -thalassémie et la fibrose kystique, résultent de SNP.²⁰⁴ Ces modifications sont très fréquentes (1 base sur 1000, chez l'homme).²⁰⁵

Leur détection a fait l'objet d'une attention toute particulière, en raison de leurs impacts sur la santé humaine. Afin de les identifier, deux approches principales existent. La première est une approche enzymatique. Elle met en jeu des ADN ligases, des ADN polymérases ou des nucléases (*TaqMan 5' Nuclease Assay*,²⁰⁶ *FLP endonuclease - FEN*). La seconde, elle, consiste à marquer des oligonucléotides complémentaires du domaine d'intérêt à l'aide de rapporteurs fluorescents (*Dynamic allele-specific hybridization - DASH* ; *Molecular Beacons*, puces à ADN ou *Base-Discriminating Fluorescence - BDF*). Si l'on prend l'exemple des BDF, tels qu'ils ont été définis par *Saito et al.*, les rapporteurs fluorescents sont généralement positionnés en face de la base à étudier. Après hybridation, le fluorophore exprime un profil d'émission qui varie en fonction du nucléotide présent sur le brin complémentaire, permettant ainsi d'identifier cette base.²⁰⁷

iii. Microenvironnement

Les nucléosides fluorescents sensibles aux changements environnementaux permettent d'étudier la dynamique, la structure et la fonction des acides nucléiques. Une fois incorporés, ils peuvent nous renseigner sur les changements conformationnels de l'ADN (ou de l'ARN) ou sur des interactions capitales dans lesquelles ils sont engagés.

La 2-AP **67** [cf. Figure 71] a par exemple été utilisée à de nombreuses reprises pour suivre la dénaturation de l'ADN ou ses changements de conformation (basculement de nucléobases hors du duplexe, présence de site abasique).²⁰⁸ D'autres sondes, telle que le UPY **95** [cf. Figure 75] positionné dans le grand sillon de l'ADN, ont également permis ce genre d'études. Il a par exemple été utilisé pour suivre les changement conformationnels ADN-B \rightarrow ADN-Z et 'tige boucle' \leftrightarrow duplexe.^{195b,c, 209}

Les interactions d'acides nucléiques avec des protéines / enzymes ou des petites molécules présentent un intérêt pharmaceutique majeur. Inhiber certaines de ces interactions ou confirmer le site de liaison d'autres, sont autant de but que se sont fixer les industries pharmaceutiques afin de développer de nouveaux principes actifs aux modes d'actions innovants. C'est pour cela que des fluorophores permettant de suivre ces interactions ont été mis au point.²¹⁰ Citons par exemple le cas de la pyrimidine **78**, utilisée afin d'étudier l'interaction de la protéine TAT avec l'ARN TAR du VIH (*Trans-Activation Response*).²¹¹ Cette même pyrimidine, son dérivé sélénisé **80** ou les ptéridines 3MI **72** et 6MI **73** [cf. Figure 71 & 73] ont également permis de suivre l'approche et la ligation d'antibiotiques aminoglycosides (paromomycine, tobramycine ou néomycine) sur l'ARN bactérien.^{182b,c, 212}

“ Le succès est d'avancer d'échec en échec sans perdre l'enthousiasme. ” - Winston Churchill

CHAPITRE 2. Résultats et Discussions

Présentation du programme de recherche

Beaucoup d'événements cellulaires fondamentaux mettent en jeu des machineries complexes d'acides nucléiques et de protéines. La méthylation de l'ADN est l'un de ces événements. Il s'agit d'une des principales modifications épigénétiques qui existent. Elle joue un rôle essentiel dans la régulation des processus majeurs tels que le maintien de l'intégrité du génome, l'empreinte génomique parentale ou encore l'expression tissu-spécifique des gènes. D'autre part, les altérations du profil de méthylation de l'ADN sont impliquées dans le développement de différentes pathologies telles que les cancers, les troubles neurodégénératifs et les maladies génétiques et métaboliques. La méthylation de l'ADN a récemment été reconnue comme l'une des principales cibles à inhiber afin de lutter contre le cancer. En effet, dans le cas de cellules cancéreuses, une hypométhylation globale de l'ADN ainsi qu'une hyperméthylation des promoteurs des gènes suppresseurs de tumeurs peuvent être observées. Bien que plusieurs inhibiteurs aient déjà été identifiés, comme la 5-azacytidine (Vidaza[®]) et la décitabine (Dacogen[®]),²¹³ leur manque de sélectivité, leur faible activité et leur toxicité limitent leur utilisation. Un examen plus approfondi de ce mécanisme permettrait d'aider les chimistes médicaux dans l'élaboration et la recherche de nouveaux inhibiteurs plus spécifiques à ce phénomène.

Actuellement, diverses hypothèses sont avancées quant à ce mécanisme de méthylation de l'ADN et méritent d'être confirmées ou clarifiées. L'une des cibles privilégiées est le couple UHRF1 / DNMT1. Ces deux protéines interviennent afin de réguler le profil de méthylation au cours de la division cellulaire qui produit deux cellules filles au matériel génétique identique à partir d'une cellule mère. Au cours de la réPLICATION de l'ADN, les

deux brins d'ADN de la cellule mère, méthylés de façon symétrique (sur chaque brins), se dissocient afin de servir de matrices pour synthétiser leurs brins complémentaires. Ceci permet d'obtenir deux cellules filles présentant chacune un duplexe d'ADN. Cependant, le brin néo-formé de ces cellules filles est dépourvu de méthyle sur les cytosines, la synthèse s'effectuant à l'aide des dNTP naturels ($N = A, T, C, G$) - on parle d'ADN hémiméthylé. Ainsi, afin de récupérer un profil de méthylation symétrique, identique à celui de la cellule mère, le couple UHRF1 / DNMT1 interviendrait au niveau du duplexe hémiméthylé afin de méthylé les cytosines du néobrin.^{23, 33} Cependant, la façon dont communiquent ces deux entités protéiques afin de jouer leurs rôles de régulatrices ainsi que leur mécanisme d'action restent encore à clarifier.

Un mécanisme en quatre étapes a été proposé afin d'expliquer le rôle de UHRF1 dans ce processus [cf. Figure 24 - Chapitre 1 et 1 - Chapitre 2].^{48b} Tout d'abord, elle se lierait à l'ADN hémiméthylé par le biais de son domaine SRA, puis ce dernier reconnaîtrait la 5mC du brin parent et la ferait basculer de sa position intra-duplexe vers son site de reconnaissance. Dans un second temps, le complexe formé avec l'ADN permettrait de recruter DNMT1, au niveau d'un second site de liaison de UHRF1, conduisant à un complexe ternaire transitoire. Dans une troisième étape, la déstabilisation du complexe nucléoprotéine, provoquée par l'approche de la méthylase, entraînerait le départ de UHRF1 et la libération de l'ADN hémiméthylé, facilitant ainsi l'accès à DNMT1. Cette dernière ferait alors basculer la cytosine à méthylé dans son site actif afin de la méthylé. Dans une dernière étape, une fois la méthylation réalisée, DNMT1 se dissocierait de l'ADN, permettant à la 5mC nouvellement synthétisée de se repositionner à l'intérieur du duplexe afin de s'apparier avec la G complémentaire qui lui fait face.^{48b}

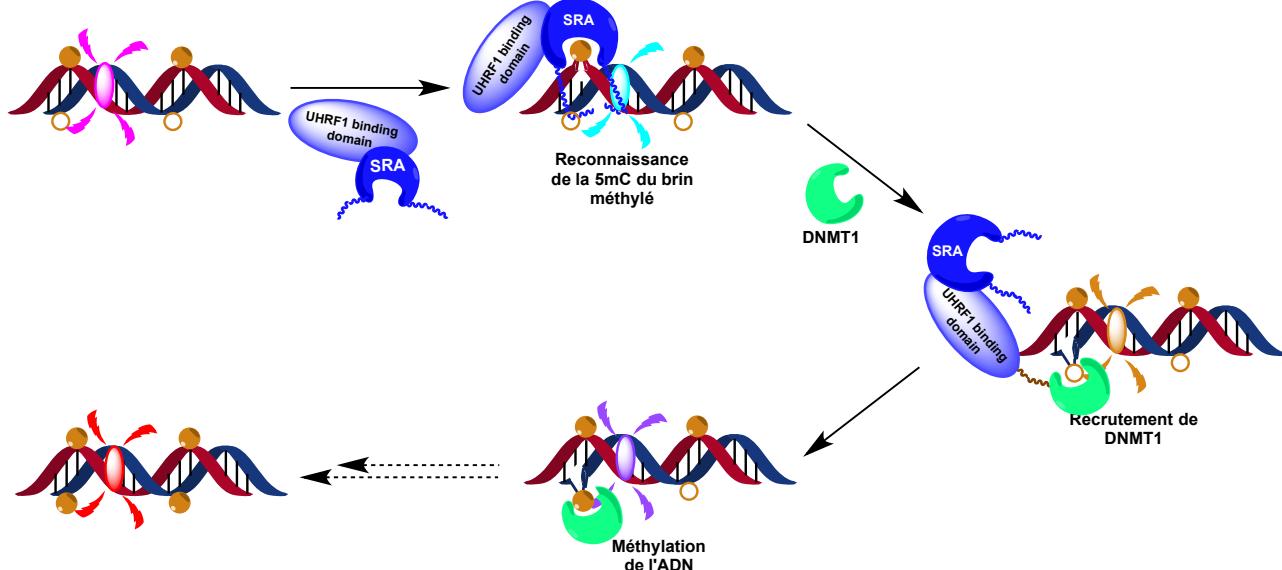


Figure 1. Hypothèse de travail pour l'étude du mécanisme de méthylation de l'ADN à l'aide d'une sonde sensible aux changements environnementaux.

Dans ce contexte, nous proposons de développer de nouveaux outils chimiques, des senseurs fluorescents, afin d'étudier la dynamique et le mécanisme de méthylation de l'ADN. Grâce à sa grande sensibilité, sa polyvalence et son caractère non destructif, la fluorescence est une technique de choix afin de comprendre, au niveau moléculaire,

l'assemblage de ces protéines, leur dynamique et leur fonction. Néanmoins, cette méthode est limitée par la disponibilité de fluorophores capables de sonder avec une grande sensibilité et de manière spécifique et multiparamétrique les changements de conformation de ces complexes nucléoprotéiques au cours de leurs interactions. Par conséquent, le développement de nouvelles sondes fluorescentes qui permettraient de surmonter ce problème fait l'objet de vives recherches dans le monde entier.

Parmi les sondes sensibles aux changements environnementaux, la famille des 3-hydroxychromones (3-HC) est une classe de fluorophores particulièrement prometteuse. En effet, ces sondes fluorescentes ratiométriques et multiparamétriques ont la capacité de se tautomériser à l'état excité, grâce à un transfert de proton intramoléculaire (ESIPT), ce qui leur permet de présenter sur leur spectre d'émission deux bandes parfaitement résolues dont le rapport d'intensité dépend des changements environnementaux (capacité du milieu à donner des liaisons hydrogène, polarité, constante diélectrique, ...).

Ainsi en partant de la séquence d'ADN double brin hémi-méthylée co-cristallisée avec le domaine SRA dont est issue la structure publiée en 2008,^{48a} nous nous sommes fixé pour objectif de synthétiser différentes séquences d'ADN marquées aux 3-HCs et de les utiliser afin de détecter et quantifier les interactions entre les différents partenaires protéiques (UHRF1 et DNMT1) et l'ADN ainsi que leur dynamique. Lorsque toutes ces interactions (URHF1 / ADN et ADN / DNMT1) auront été caractérisées [cf. Figure 1], il sera alors plus aisément de corroborer ou bien de corriger les hypothèses émises sur le mécanisme qui régit le ménage à trois (UHRF1 / ADN / DNMT1) des acteurs principaux de la méthylation des îlots CpG. Au cours de cette thèse, nous nous sommes focalisés sur l'étude de la première étape de ce mécanisme de méthylation de l'ADN, à savoir la reconnaissance du brin hémi-méthylé par le domaine SRA de UHRF1.

Pour atteindre ces objectifs, deux stratégies de marquage, complémentaires l'une de l'autre, ont été envisagées. La première d'entre elles repose sur la substitution d'une nucléobase naturelle (position anomérique) par le motif 2-thiényle-3-HC [cf. Figure 2A]. La synthèse et la caractérisation de cet analogue fluorescent de nucléoside a été mise au point par Marie Spadafora^{126b} et son incorporation et caractérisation dans des ODNs par Dmytro Dziuba^{126d} au cours de leurs thèses dans notre laboratoire. En collaboration avec l'équipe du Pr. Yves Mély (Université de Strasbourg), il a été démontré récemment que l'interaction de la nucléoprotéine du VIH (NCp7) avec sa cible ADN marquée par cette sonde se traduit par un changement du rapport d'intensité des deux bandes de fluorescence, attestant la validité du concept.^{126d} Cette approche devrait permettre de sonder les changements environnementaux à l'intérieur du duplexe d'ADN (changement conformationnel par exemple) [cf. Figure 2B]. Les propriétés de nucléobase universelle fluorescente de la 2-thiényle-3-HC, (possibilité de substituer n'importe quelle base, faible influence sur la stabilité et la conformation des duplexes et excellentes caractéristiques spectroscopiques : rendement quantique, émission duale) en font une candidate de choix pour ces études.^{126b,d}

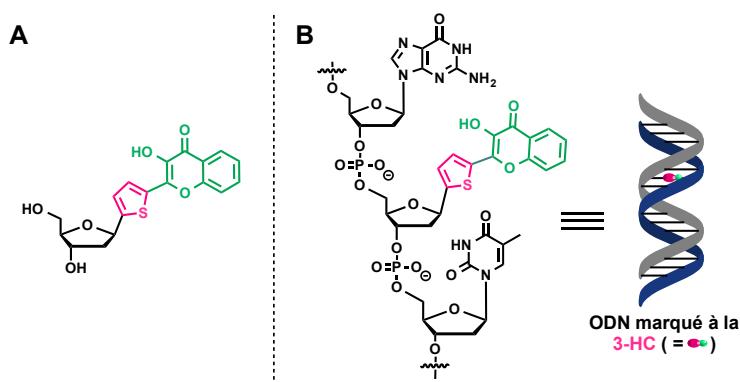


Figure 2. (A) Structure de la sonde de première stratégie et (B) schéma illustrant son incorporation dans un ODN et son positionnement à l'intérieur du duplexe.

L'autre stratégie fait intervenir, quant à elle, une modification d'une nucléobase naturelle par le biais d'un espaceur (alcyne) directement relié au motif fluorescent [cf. Figure 3]. Cette approche devrait permettre de localiser la copule fluorescente dans le grand sillon de l'ADN et ainsi de sonder les changements environnementaux à l'extérieur du duplexe d'ADN. L'accès aux analogues de nucléosides de ce type avait été initié par Marie Spadafora et Dmytro Dziuba. Mon second objectif a été de mettre au point les conditions de synthèse permettant d'accéder à ces dérivés de dT fonctionnalisés par le motif 3-HC fluorescents, sensibles aux changements environnementaux, et à les incorporer dans de l'ADN. Des travaux préliminaires ont d'abord été nécessaires afin de déterminer les propriétés photophysiques d'un tel rapporteur en partant dans un premier temps d'un composé modèle de dT, sur lequel le motif 3-HC était attaché en position 5 [cf. Figure 3A],^{126f} puis dans un deuxième temps en série 2'-désoxyuridine avec divers espaces aromatiques (phényle, furyl et thiényl - Figure 3B).^{126g} Dans un troisième temps, les copules fluorescentes démontrant les propriétés photophysiques les plus appropriées ont été introduites dans divers ODNs modèles, afin de pouvoir caractériser ses propriétés de fluorescence une fois incorporée dans l'ADN [cf. Figure 3C]. Par ailleurs le double brin d'ADN retenu pour l'étude des interactions du domaine SRA à l'ADN hémiméthylé comprend une dT dans la séquence non-méthylée. Cette dT est proche du site de liaison mais n'est pas impliquée dans les interactions qui s'établissent entre l'ADN et SRA (données cristallographiques).^{48a} Nous avons donc synthétisé des ODNs de séquence identique mais dont la dT a été substituée par nos nouvelles sondes afin de les utiliser pour les études du mécanisme de méthylation de l'ADN.

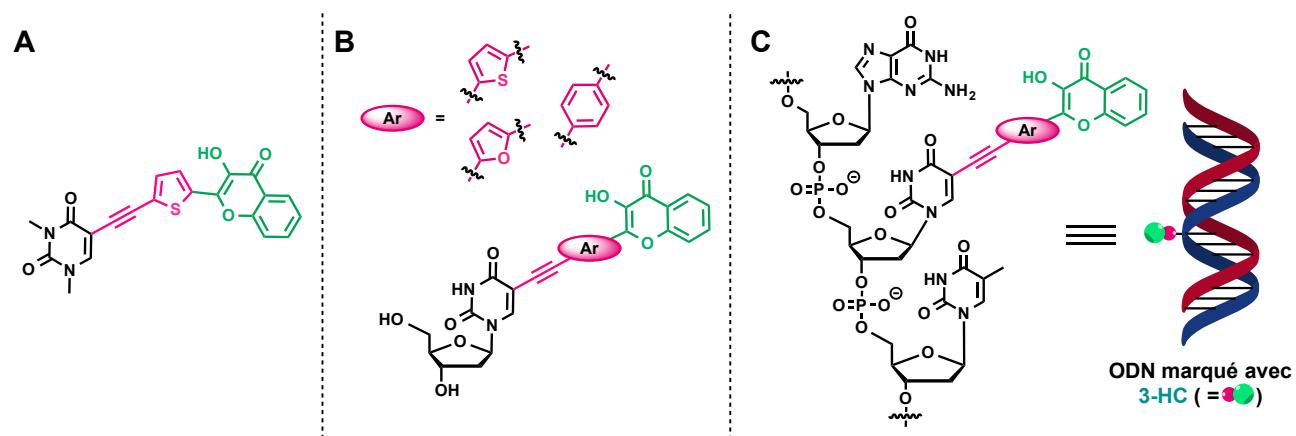


Figure 3. (A) Structure de la sonde de seconde stratégie diméthylée au niveau de la nculéobase, (B) analogues nucléosidiques correspondants et (C) schéma illustrant leur incorporation dans un ODN (positionnement dans le grand sillon).

Les deux approches développées et les résultats correspondants vont à présent être plus amplement développés dans ce Chapitre 2. Résultats et Discussions.

A. Première stratégie de marquage (intra-duplexe)

PUBLICATION 1

L'objectif principal du projet de thèse est le développement d'outils fluorescents pour l'étude du mécanisme de méthylation de l'ADN. Nous nous sommes tout d'abord intéressés à la première étape du mécanisme, à savoir l'approche du domaine SRA de UHRF1 sur l'ADN hémi-méthylé. Afin de pouvoir réaliser une telle étude, nous avions besoin de séquences d'ODN fluorescentes, marquées à la 3-HC, et présentant une forte affinité pour le domaine SRA. Le nucléoside 2-thiényl-3-HC (première stratégie) semblait être le fluorophore ratiométrique adapté afin de réaliser ce marquage et ces études car il présente des propriétés de base universelle, ce qui lui permet de substituer n'importe quelle nucléobase, avec un effet minime sur la stabilité du duplexe et une très grande sensibilité aux changements environnementaux.^{126b,d} De plus, la séquence sélectionnée pour l'étude correspond à celle ayant été co-cristallisée avec SRA en 2008.^{48a} En se basant sur la « cartographie » des interactions établies entre l'ODN modèle et le domaine SRA de UHRF1, obtenue par diffraction des rayons X [cf. Figure 4A],^{48a} nous avons donc déterminé des positions de bases azotées pouvant être substituées par la base fluorescente universelle [cf. Figure 4B].

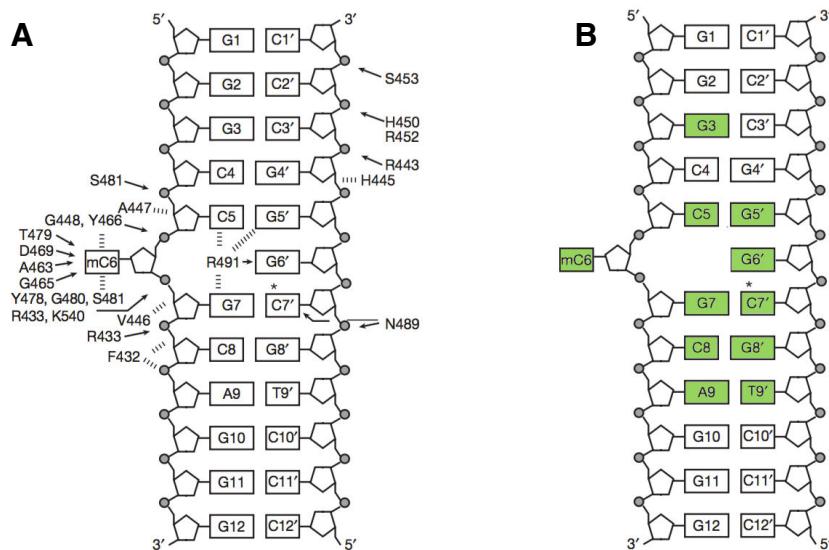


Figure 4. (A) Carte des interactions du domaine SRA avec l'ODN (les interactions de van der Waals sont représentées comme une série de lignes courtes et parallèles et les liaisons hydrogène comme des flèches) et (B) Positions sélectionnées pour le marquage à la 3-HC.

Le nucléoside fluorescent a ensuite été synthétisé à l'échelle du gramme, en 16 étapes, sous la forme d'un phosphoramidite, synthon indispensable afin de permettre son incorporation dans des ODNs. Pour y parvenir, le protocole de synthèse développé au laboratoire^{126b,d} a été appliqué [cf. Figure 5]. Cependant, afin de maximiser le rendement de synthèse, les étapes 9 → 10 et 10 → 11 [cf. Figure 5] ont été modifiées. En effet, les groupements protecteurs *tert*-butyl diphenyl silyles des alcools 3' et 5' du sucre

initialement utilisés, entraînaient des rendements de synthèse variables à cause de gênes stériques et de temps de déprotections très longs (4 à 5 jours), à l'origine d'une déprotection non négligeable du groupement 3-Cbz de la chromone. Ils ont donc été remplacés par des groupements moins volumineux et plus labiles (*tert*-butyl diméthyl silyles).

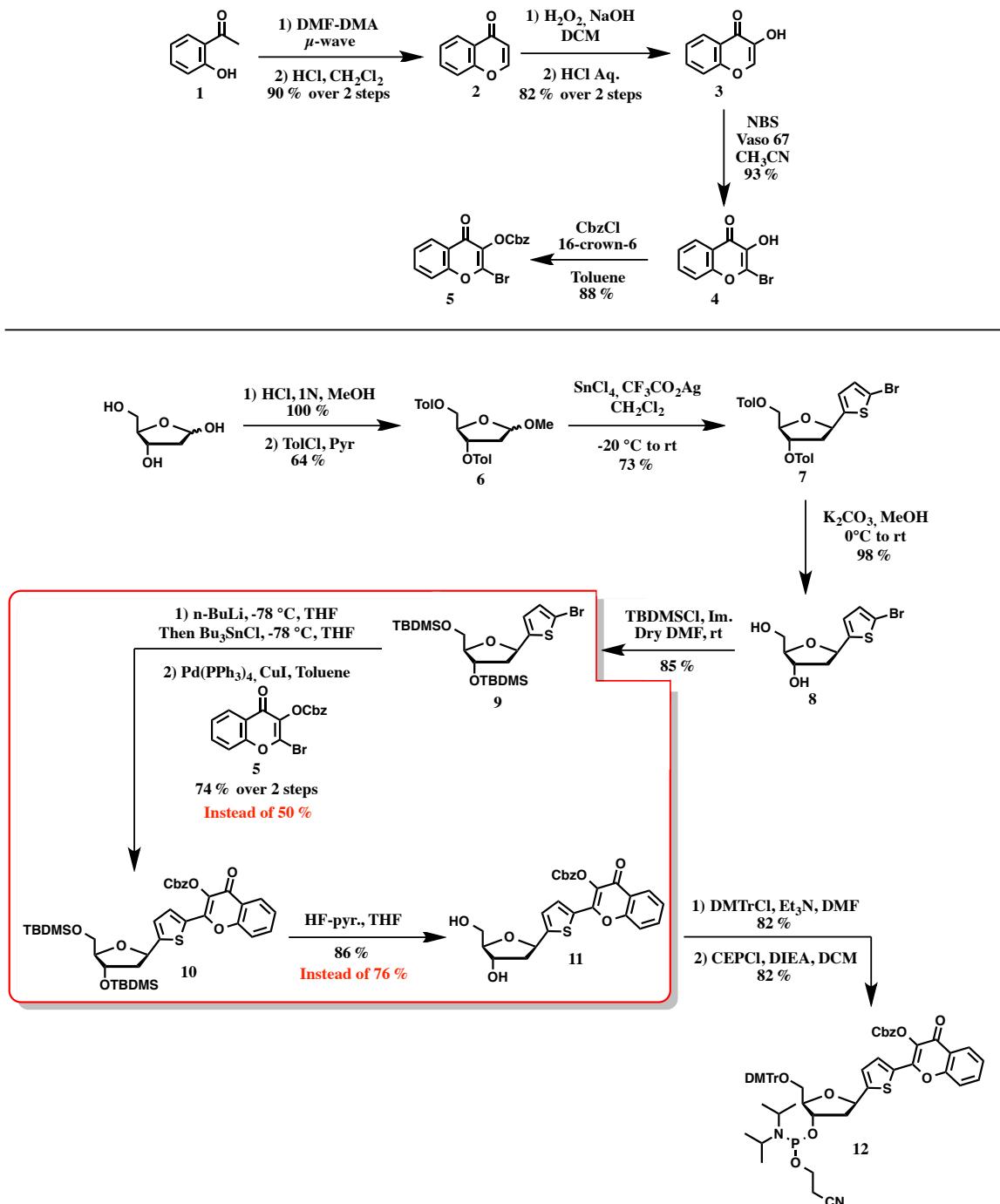


Figure 5. Synthèse du phosphoramidite de 1^{ère} génération (les étapes améliorées sont encadrées en rouge).

Le phosphoramidite ainsi obtenu a alors été engagé dans la préparation de 18 ODNs marqués à la 3-HC à proximité du site de reconnaissance du dinucléotide CpG central (brin méthylé, non-méthylé et brin complémentaire) [cf. Figure 4B]. Une phase de pré-sélection a ensuite permis d'isoler deux séquences particulièrement prometteuse

(marquées en C8 et G5' - Figure 1B), pour lesquelles le rapporteur perturbait à minima la conformation du duplexe et sa liaison avec l'entité protéique - conditions indispensables afin d'obtenir des résultats exploitables.

Enfin, les études de fluorescence à l'état stationnaire et résolue en temps ont été poursuivie avec ces deux séquences afin de préciser le mécanisme de reconnaissance de SRA pour sa cible ADN hémi-méthylé. Les variations du signal de fluorescence (rendement quantique et rapport I_{N^*}/I_{T^*}) ont ainsi permis de suivre l'approche du domaine SRA de UHRF1 ainsi que le basculement de la 5mC dans son site de reconnaissance. De plus, les études de *stopped-flow* ont permis de déterminer la cinétique (constantes de vitesse) du processus de basculement de la base. Un article a été rédigé pour être soumis à publication.

UHRF1-induced base flipping in methylated DNA can be monitored by using an environment sensitive fluorescent nucleobase

Vasyl Kilin^a, Nicolas Barthes^b, Benoît Michel^b, Christian Boudier, Valeriy Yashchuk, Marc Mousli, Christian Bronner^b, Alain Burger^b, Yves Mély^{a,}*

^a Laboratoire de Biophotonique et Pharmacologie, UMR 7213 CNRS, Université de Strasbourg, Faculté de pharmacie, 74 route du Rhin, 67401 Illkirch, France

^b Institut de Chimie de Nice, UMR 7272 CNRS, Université de Nice Sophia Antipolis, Parc Valrose, 06108 Nice Cedex 2, France

^c Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), INSERM U964 CNRS UMR 7104, Université de Strasbourg, Illkirch, France

^d Kiev National Taras Shevchenko University, Kiev, Ukraine

* Corresponding author: yves.mely@unistra.fr Tel: +33-368854263

Abstract

DNA methylation patterns, which play a key role in the regulation of gene activity, are mitotically inherited through postreplicative methylation of the newly synthesized DNA strand by the concerted action of the DNA methyltransferase 1 (DNMT1) and the ubiquitin-like with PHD and ring finger domains 1 (UHRF1) proteins. The first step of DNA methylation replication is thought to be the recognition and subsequent flipping of a methylated cytosine (mC) in the context of a mCpG dinucleotide by the Set and Ring Associated (SRA) binding domain of UHRF1. To monitor and investigate for the first time the kinetics of this base flipping, we used the fluorescent ratiometric nucleotide analog 2-thienyl-3-hydroxychromone (3HCnt). The probe was incorporated at various positions in a 12 base pair DNA, close to the central ⁶CpG⁷ recognition site containing either a non-modified C (NM duplex) or a mC (HM duplex). When the 3HCnt was at position 5' or 8, SRA interaction induced much larger spectroscopic changes (quantum yield and ratio of the two emission bands of 3HCnt) in HM duplexes as compared to NM duplexes. These large changes could be clearly associated to the flipping of methylcytosine, as they disappear with a G448D SRA point mutant unable to flip the mC. Furthermore, stopped-flow studies revealed that SRA interaction with HM duplexes was governed first by very fast kinetics, consistent with a “bind-and-slide” model, accompanied by a much slower kinetic component ($k = 10 \text{ s}^{-1}$) corresponding to mC6 base flipping. Dissociation experiments further revealed that the flipping back process was governed by a rate constant of 2.9 s^{-1} . Therefore, base flipping of mC6 was found to stabilize by four-fold the binding of SRA to HM as compared to NM duplexes and to considerably increase the lifetime of the SRA/HM duplex, which may be crucial for the recruitment of DNMT1.

Keywords

Conformational changes, DNA methylation, ESIPT, fluorescence, 3-hydroxychromone, SRA domain, UHRF1

Abbreviations

3HCnt, 2-thienyl-3-hydroxychromone; DNMT1, DNA methyltransferase 1; FRET, fluorescence resonance energy transfer; HM, hemi-methylated; NM, non-methylated; BM, bi-methylated; ODN, oligonucleotide; PHD, Plant Homeo Domain; QY, quantum yield; SRA, SET and RING-Associated; UHRF1, Ubiquitin-like containing PHD and RING Finger 1

Introduction

Epigenetics refers to the study of heritable phenotypic changes that occur without altering the DNA sequence. At the molecular level, major epigenetic marks include DNA methylation, post-translational modifications of histones, histone variants and nucleosome positioning¹⁻¹². DNA methylation is a heritable cytosine modification, realized by covalent addition of a methyl group to the fifth carbon of the cytosine base.

In mammalian cells, cytosine methylation mainly occurs in a symmetrical CpG context¹³, a fundamental unit of epigenetic inheritance, which supports the semi-conservative model for its duplication. In differentiated cells approximately 70-80 % of all CpG are methylated^{14,15}. This methylation is established and maintained by a family of enzymes called DNA methyltransferases (Dnmnts)¹⁶. One of the first steps of the reproduction of the DNA methylation profile involves recognition of hemi-methylated CpG sites (*i.e.*, only one DNA strand that is methylated) generated after DNA replication, which will be subsequently fully methylated through the action of DNMT1. However, the preference of DNMT1 for hemi-methylated CpG sites is insufficient to explain the high degree of fidelity to replicate the DNA methylation patterns. In this context, UHRF1 (Ubiquitin-like containing PHD and RING finger 1) plays a key role since it is the co-factor that guides DNMT1 to find its target on DNA. This is achieved by its ability, via its SRA (Set and Ring Associated) domain, to specifically recognize hemi-methylated DNA¹⁷⁻²⁴ and flip methylated cytosines.

Crystallographic studies carried out on the SRA domain gave interesting insights on the recognition of hemi-methylated CpG sites and the mechanism of methylcytosine flipping¹⁷⁻¹⁹. On binding of SRA to hemi-methylated DNA, methylcytosine is flipped out from the helix to a specific binding pocket. This binding pocket has several features that enable the SRA domain to selectively recognize methylcytosines. First, its size is too small to fit purines but contains a hemisphere of 2 Å radius, where a methyl group can exactly accommodate. Second methylcytosine is sandwiched through π -stacking interactions by two tyrosine residues at positions 466 and 478 (Fig 1A). Third, the methylcytosine is stabilized by two hydrogen bonds with an aspartic amino acid at the top of the binding pocket and a hydrogen bond with a

threonine¹⁷. Fourth, at the entry of the pocket, glycine 448 (Fig 1A) plays a crucial role since its mutation drastically decreases affinity of SRA for hemimethylated DNA¹⁷. Though X-ray crystallography was very useful to resolve the structure of a SRA/hemi-methylated DNA, understanding of the molecular mechanism how SRA reads DNA and promotes flipping of methylcytosine in diluted solutions requires other techniques. Among the possible techniques, fluorescence-based ones are particularly attractive, due to their high sensitivity, ease to use and multiparametric information. However, these methods rely on the availability of fluorescent probes that should respond sensitively and selectively to the molecular event of interest. In the case of nucleic acids, the most obvious strategy to obtain a selective response is to use fluorescent nucleobase analogues that could be incorporated at any desired position. However, monitoring base flipping is especially challenging due to several constraints. First, due to the highly constrained dimensions of the SRA binding pocket, it appears very difficult to find an appropriate fluorescent nucleobase able to substitute for methylcytosine. Moreover, substitution of a nucleobase close to the methylcytosine should not affect the binding of SRA, nor the stability or conformation of the labeled duplexes. Recently, 2-aminopurine (2-AP) was found to fulfill partly these criteria because, when incorporated at positions close to a methylcytosine, 2-AP was able to monitor the binding of SRA to methylated, hemi-methylated and fully methylated duplexes. However, it demonstrated insufficient sensitivity for rising to the challenge of reporting base flipping²⁵.

A prospective fluorescent nucleoside bearing 2-thienyl-3-hydroxychromone as a base surrogate was recently developed²⁶. This fluorescent nucleoside (3HCnt) features high environmental sensitivity, as a consequence of an excited-state intramolecular proton transfer (ESIPT) reaction^{27,28} Fig.1B, which results in two excited states (normal, N* and tautomer, T*) with well separated emission bands. These two forms being differently sensitive to the physico-chemical parameters of the environment, and notably to polarity and hydration²⁹, it is possible to sensitively monitor the annealing of the labeled sequences with complementary sequences or the binding of proteins through changes in the N*/T* ratio^{26,30}. Moreover, this fluorescent nucleobase behaves as a universal nucleobase that is able to substitute any

nucleobase with minimal effect on the stability of the duplexes.

In the present study, we labeled the 12-bp duplex used to solve the 3D structure of the SRA/duplex complex¹⁷ with the fluorescent 3HC nucleobase at several positions, close to the central "CpG"⁷ recognition site. This recognition site contained either a non-modified cytosine (NM duplex) or a methylated cytosine (HM duplex). When the 3HC nucleobase replaced G5' or C8, it was found to show a much larger change in quantum yield and N*/T* ratio on SRA binding to HM duplexes, as compared to the corresponding NM duplexes. This change in the spectroscopic properties could be clearly associated to the flipping of the methylcytosine, as the use of a SRA point mutant that is unable to flip the methylcytosine was found to provide similar spectroscopic changes in NM and HM duplexes. Finally, stopped-flow studies were used to characterize the kinetic rate constants of the base flipping process.

MATERIALS AND METHODS

Materials

The SRA domain of hUHRF1 (SRA, residues 408-643) was expressed in *Escherichia coli* BL21-pLysS (DE3) and purified as described previously³¹. The mutant G448D was produced as described elsewhere³². Unmodified oligonucleotides (DNAs) were synthesized and HPLC-purified by IBA GmbH Nucleic Acids Product Supply (Germany). 3HCnt and DNAs labeled by 3HCnt were produced as described previously²⁶. The sequence of the 12-bp duplex was 5'-GGGCCXGCAGGG-3'/5'-CCCTGCCGCC-3' with a single CpG site that was either nonmethylated (X = C) or hemimethylated (X = 5mC). The 3HCnt was selectively introduced at different positions (3, 5, 6, 7, 8, 9, 6' or 5') within the sequence described previously²⁶. Double stranded DNAs were prepared by mixing complementary strands in equal molar amounts in PBS 20mM, pH 7.5, NaCl 50mM, EDTA 1mM buffer and annealing them by heating to 90 °C for 5 min, followed by cooling down to room temperature. After this step, samples were kept on ice. The presence of EDTA was critical to avoid the formation of the anionic form of 3HCnt^{33,34}.

The lyophilized Calf thymus DNA (ctDNA) from Sigma Aldrich was reconstituted overnight at 10 °C and molar concentration of nucleotides was determined by measuring absorption at 260 nm using extinction coefficient of 6600 M⁻¹.cm⁻¹.

Absorption spectroscopy

Absorption spectra were recorded on a Cary 400 spectrophotometer (Varian). Extinction coefficients comprised between 86,500 and 135,020 M⁻¹.cm⁻¹ were used to determine the concentrations of the duplexes at 260 nm. Both diffusion and buffer absorbance were subtracted from the spectra. All experiments were performed at 20 °C in PBS 20 mM, pH 7.5, NaCl 50 mM, EDTA 1 mM, TCEP 2.5 mM buffer, in the presence of 0.04 % PEG 20,000 to prevent protein adsorption on the cuvette walls³⁵.

Melting curves of the duplexes (2 μM concentration of each strand) were recorded by following the temperature-dependence of the absorbance changes at 260 nm on a Cary 400 spectrophotometer (Varian) equipped with a Peltier thermostated cell holder. The optical pathlength of the cell was 1 cm. The temperature range was 20 – 80 °C, with a speed of heating of 0.5 °C/min. PBS 20 mM, pH 7.5, NaCl 50 mM, EDTA 1 mM buffer was used. The melting temperatures were extracted from melting curves as described elsewhere³⁶.

Steady-state fluorescence spectroscopy

Fluorescence emission spectra were recorded at 20 °C on a FluoroLog spectrofluorimeter (Jobin Yvon) equipped with a thermostated cell compartment. Excitation wavelength was set at 374 nm to excite selectively 3HCnt. Spectra were corrected for buffer fluorescence, lamp fluctuations, and detector spectral sensitivity. Quantum yields of the labeled duplexes in the absence or presence of the SRA protein were determined by using quinine sulfate (QY = 0.577 in 0.5 M H₂SO₄) as a reference³⁷. Measurements were performed by using SRA concentrations ensuring that at least 80 % of the 3HCnt-labeled duplexes are bound to SRA.

Anisotropy measurements were performed on the same instrument. Excitation wavelength was at 374 nm and emission was collected at 556 nm, which corresponded to the T* band emission. Anisotropy was calculated as I=(Ivv-Givh)/(Ivv+G2Ivh), where G is the correction factor. The anisotropy value was obtained by averaging 10 measurements.

The affinity of SRA wild type and G448D mutant to labeled DNAs was obtained by titrating a fixed amount of 3HCnt-labeled duplex by SRA and monitoring the fluorescence intensity and anisotropy signals, simultaneously. The affinity constants were determined by fitting the

fluorescence anisotropy changes to the following equation:

$$r = \frac{vRr_t - r_d(v-1)}{1+Rv-v} \quad (1)$$

where r and r_t are the anisotropy values at a given and a saturating SRA concentration respectively, r_d is the anisotropy in the absence of protein. R is the ratio of the quantum yields of the bound to free forms, K_a is the apparent affinity constant, v is fraction of bound SRA calculated as:

$$v = \frac{(K_a^{-1} + nL_t + P_t) - \sqrt{(K_a^{-1} + nL_t + P_t)^2 - 4n P_t L_t}}{2L_t} \quad (2)$$

where P_t and L_t are the concentrations of SRA and duplexes, and n is the number of SRA proteins bound per duplex^{38,39}.

Isothermal titration calorimetry (ITC)

Affinity of SRA to non-modified DNA was also determined by Isothermal titration calorimetry. The heat exchanged for the interaction of SRA with the non-modified methylated duplex was analyzed at 20 °C in PBS 20 mM, pH 7.5, NaCl 50 mM, EDTA 1 mM with a VP ITC microcalorimeter (Microcal, Northampton, MA, USA). Titration of SRA by the non-labeled methylated duplex was performed by monitoring under constant stirring (310 rpm) the thermal power generated by repeated injections of 4 µL aliquots (in 4 min) of 40 µM DNA contained in the syringe into a 6 µM SRA solution contained in the 1.42 mL cell compartment of the instrument. The total heat resulting from an injection of titrant was calculated as the integral versus time of the experimental signal. A control experiment in which DNA was titrated into the buffer alone was done to determine the heat of dilution. Instrument control, data acquisition, and analysis were done with the VPViewer and Origin software provided by the manufacturer.

Stopped flow spectroscopy

The reaction of SRA with the 3HCnt-labeled duplexes was monitored with time, using a stopped-flow apparatus (SFM-3, Bio-Logic, Claix, France), which allows rapid mixing of 100 µL of each reagent. The 3HCnt was excited at 365 nm, and its fluorescence intensity was recorded above 530 nm with a long-pass filter (Melles Griot, France). The data recording frequency was 5 kHz for the first 200 ms and 2 kHz between 200 ms and 1200 ms. The dead time of the set-up was 2 ms⁴⁰. The final concentration of labeled DNA was 0.3 µM and the concentration of SRA was adjusted to have 80% of DNA bound to protein. Blank experiments in which

the SRA was omitted were performed in the same conditions. Dissociation experiments were performed by adding an excess of ctDNA to pre-formed complexes of SRA with HM or NM duplexes labeled with 3HCnt. The final concentration of SRA was adjusted to have 50 % of DNA bound to protein. The condition of signal acquisition and experimental setup parameters were the same as for association measurements.

Data acquisition and processing were done with the Biokine software from the instrument manufacturer.

Results

Labeling position strategy and spectroscopic characterization of the labeled duplexes

Based on the reported structure of SRA in complex with a 12 bp duplex (Fig. 1A)¹⁷, we labeled the same duplex with 3HCnt (Fig 1B) at positions 5, 6, 7, 8, 9, 5' and 6'. These positions were selected due to their proximity to the methylcytosine (mC) at position 6 and their direct interaction with SRA. Positions 3 and 9, not in contact with SRA were also labeled. Moreover, corresponding 3HCnt-labeled NM duplexes with a cytosine at position 6 were also synthesized for systematic comparison with HM duplexes.

For all 3HCnt-labeled duplexes, we first determined their melting temperatures by monitoring their absorbance changes at 260 nm (Fig S1 Supplementary materials). With the exception of position 3, the Tm depression was found to be minimal (≤ 3 °C) (Table S1 Supplementary materials), confirming that though 3HCnt does not participate in base pairing, it destabilizes the duplexes less than a mismatch because of its large area of contact with the surrounding bases²⁶. Duplexes labeled at position 3 provide a melting curve of low amplitude (data not shown), suggesting that due to its proximity to the 5' end, the 3HCnt destabilizes the terminal base pairs, so that only limited hypochromicity is achieved. Nevertheless, at the working temperature of 20°C, the destabilization is limited, so that duplexes are the major form.

Inclusion of the 3HCnt into the duplexes was found to significantly shift the positions of the N* and T* bands, and strongly decrease the N*/T* ratios in comparison to the free 3HCnt (Fig. 1C). Both the spectral shift and the decrease in N*/T* ratio are consistent with stacking of the 3HCnt with its neighboring bases that protects it from the bulk

water²⁶. The quantum yields were found to be rather low, being between 0.5 and 2% (Table S2), probably due to the fact that at all positions, the 3HCnt has a G and/or C neighbors, which acts as an efficient fluorescence quencher²⁶. Importantly, the methylation of the cytosine at position 6 had negligible effect on the spectroscopic properties of 3HCnt at all used positions (Table S2).

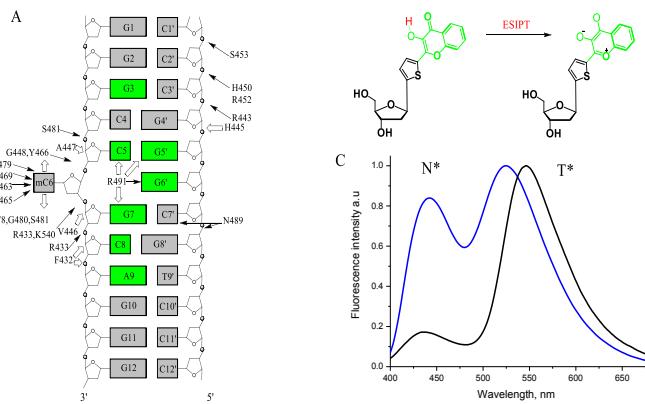


Figure 1. Structure of the used duplex and the 3HCnt. (A) Structure of the used duplex. The interactions of the duplex with SRA, as determined by X-ray crystallography¹⁷ are indicated by arrows. Hydrogen bonding and van der Waals interactions are indicated by black and white arrows, respectively. Positions substituted by 3HCnt are highlighted in green. (B) Structure of 3HCnt and ESIPT reaction. (C) Fluorescence spectra of 3HCnt either free (blue curve) or incorporated at position G5 in the duplex (black curve). Experiments were done in Phosphate buffer 20 mM, NaCl 50 mM, TCEP 2.5 mM, PEG 0.05%, pH 7.5, at T=20°C. The spectra were normalized at the maximum of the T* band.

Monitoring SRA-DNA interaction by steady-state spectroscopy measurements

To determine whether 3HCnt could be used to monitor the SRA-induced base flipping of the methylated cytosine at position 6 (m6C), we compared the SRA-induced changes in the emission spectra of the 3HCnt-labeled HM or NM duplexes. SRA was observed to increase the quantum yield and to decrease the N*/T* ratio of almost all the HM and NM duplexes, but the changes were strongly dependent on the position of 3HCnt (Fig. 2 and 3, and Table S2). The most interesting data were obtained for 3HCnt at positions 8 and 5' (Fig 2A and B). While SRA interaction with HM led to an increase of 3HCnt QY up to 2 and 2.7 times for positions 8 and 5', respectively, the corresponding increase was only of ~1.3 and 1.6 times, respectively in DNA-NM (Fig 2A and B).

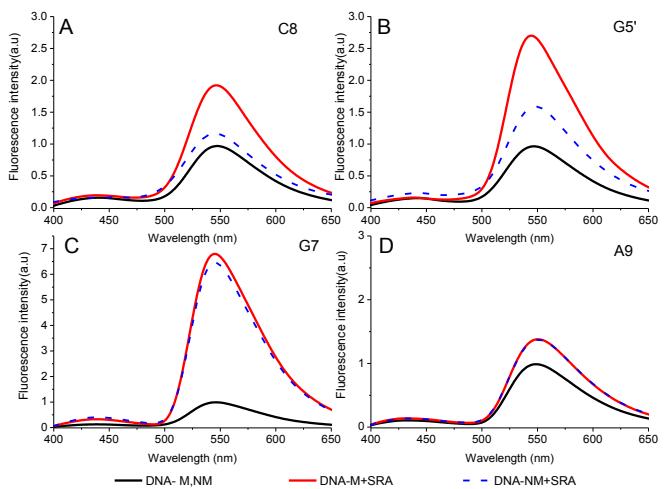


Figure 2. Fluorescence spectra of 3HCnt-labeled DNA at positions 8 (A), 5' (B), 7 (C) and A9 (D). Black lines - spectra of free labeled DNA (HM and NM are superimposable), blue lines – spectra of NM bound with SRA, red line - spectra of HM bound with SRA. Experiments were done in PBS buffer 20 mM, pH 7.5 NaCl 50 mM, TCEP 2.5 mM, PEG 0.05% at 20°C. Concentration of labeled duplexes was 1 μM, while the SRA concentration was adjusted to complex at least 80% of the DNA molecules.

In parallel, while 30 % and 69 % decreases in the N*/T* ratio were observed at positions 8 and 5' in HM respectively, the corresponding decreases were only of 7 % and 25 % in NM. As HM and NM differ only by a methyl group at the cytosine at position 6, it is tempting to conclude that the larger changes observed in HM may be related to the flipping of mC6.

The strongest changes in QY were observed at positions 5, 7 and 6' (Table S2), which are the immediate neighbors or the opposite nucleobase of C6 or mC6. At these positions, an up to 6.5-fold increase in QY could be observed. However, only marginal differences in QY, but also in N*/T* ratios, could be observed between HM and NM, suggesting that labeling at these positions does not allow to selectively monitor the flipping of mC6. Finally, at positions 3 and 9, only limited changes in QY and N*/T* ratios were observed on SRA binding, indicating that these positions were poorly sensitive to the interaction with SRA, as expected from the marginal interactions of the nucleobases at both positions with SRA in the crystal structure¹⁷.

Characterization of the binding of SRA to DNA

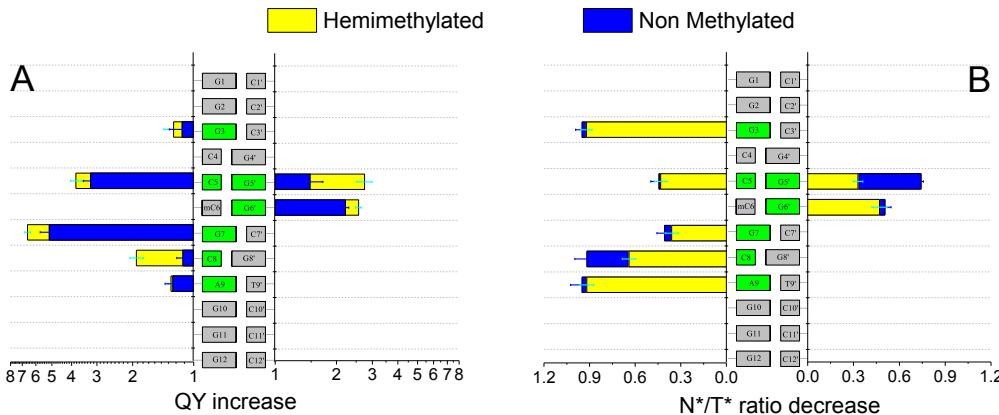


Figure 3. SRA-induced changes in quantum yield (A) and N*/T* ratio (B) of 3HCnt-labeled NM (blue bars) and HM (red bars). Experiments were performed as described in Figure 2. The QY increase in (A) shown in log scale corresponds to the ratio of the QY of 3HCnt in the DNA/SRA complex to that of 3HCnt in free DNA. The N*/T* ratio decrease in (B) corresponds to the ratio of the N*/T* ratio of 3HCnt in the DNA/SRA complex to that in the free DNA. Data are expressed as means +/- standard deviation for n = 5 independent experiments.

In a first step, we characterized the binding of SRA with non-labeled HM duplex by isothermal titration calorimetry (ITC). Titration of SRA by the duplex revealed that the reaction is exothermic (Fig. 4A top) and that the binding curve could be adequately fitted by a 1:1 binding model (Fig. 4A bottom), with an apparent affinity constant, $K_{app} = 1.4(\pm 0.3) \times 10^6 \text{ M}^{-1}$, in good agreement with the literature^{22,25}.

Next, we investigated the binding of SRA to 3HCnt-labeled duplexes by fluorescence anisotropy (Fig. 4B). A fixed concentration of labeled DNA was titrated by increasing concentrations of protein. The anisotropy changes as a function of the SRA/DNA molar ratio were fitted by equation (1) using a 1:1 binding model. The K_{app} values were found to depend on the 3HCnt position in the duplex (Figure 4C and Table S3). Noticeably, when 3HCnt was at positions 3, 8, 9, or 5', the K_{app} values obtained with the labeled HM duplex were close to the K_{app} value obtained by ITC with the non-labeled HM duplex. This indicates that at these positions, the 3HCnt does apparently not affect the binding of SRA to the duplex. This conclusion is further substantiated by the fact that at positions 8 and 5', the affinity of SRA for the HM duplex was ~5 and ~8 times higher than for the NM duplex (Figure 4D), in full agreement with previous reports on non-labeled sequences^{17,18,20-24}.

In contrast, 3HCnt at position 5 or 7 increased the affinity of the labeled HM duplex to SRA by 4 to 14 times as compared to the non-labeled HM duplex. As numerous contacts were evidenced between these positions and SRA in the SRA/DNA complex

¹⁷, it is likely that the complexes are stabilized through direct interaction between the 3HCnt and SRA amino acids. One consequence of this stabilization is that it precludes observing any binding difference between HM and NM duplexes, so that C6 and mC6 likely behave similarly in the two complexes. The similarity in the behavior of the nucleobase at position 6 is further highlighted by the similar SRA-induced spectroscopic changes in HM and NM duplexes labeled at position 5 or 7 (Figures 3A and B). At position 6', the 3HCnt also substantially stabilizes the complexes with SRA, but differences in the binding between DNA-HM and DNA-NM could still be observed.

Taken together, our data show that the 3HCnt at position 8 and 5' does not alter the binding of SRA to both HM and NM duplexes, and allows observing clear spectroscopic differences between the two duplexes on SRA binding. Thus, these positions were selected for further study.

Interaction of 3HCnt-labeled HM and NM duplexes with G448D SRA mutant

To further demonstrate that the differential spectroscopic response of HM and NM duplexes labeled at positions 5' and 8 to SRA binding may be connected with flipping of the mC6 nucleobase, we used a SRA mutant where the G448 residue has been replaced by a D residue. This G448D mutant was suggested to prevent the flipping of the mC6 residue through the steric hindrance induced by the mutation^{17,32}.

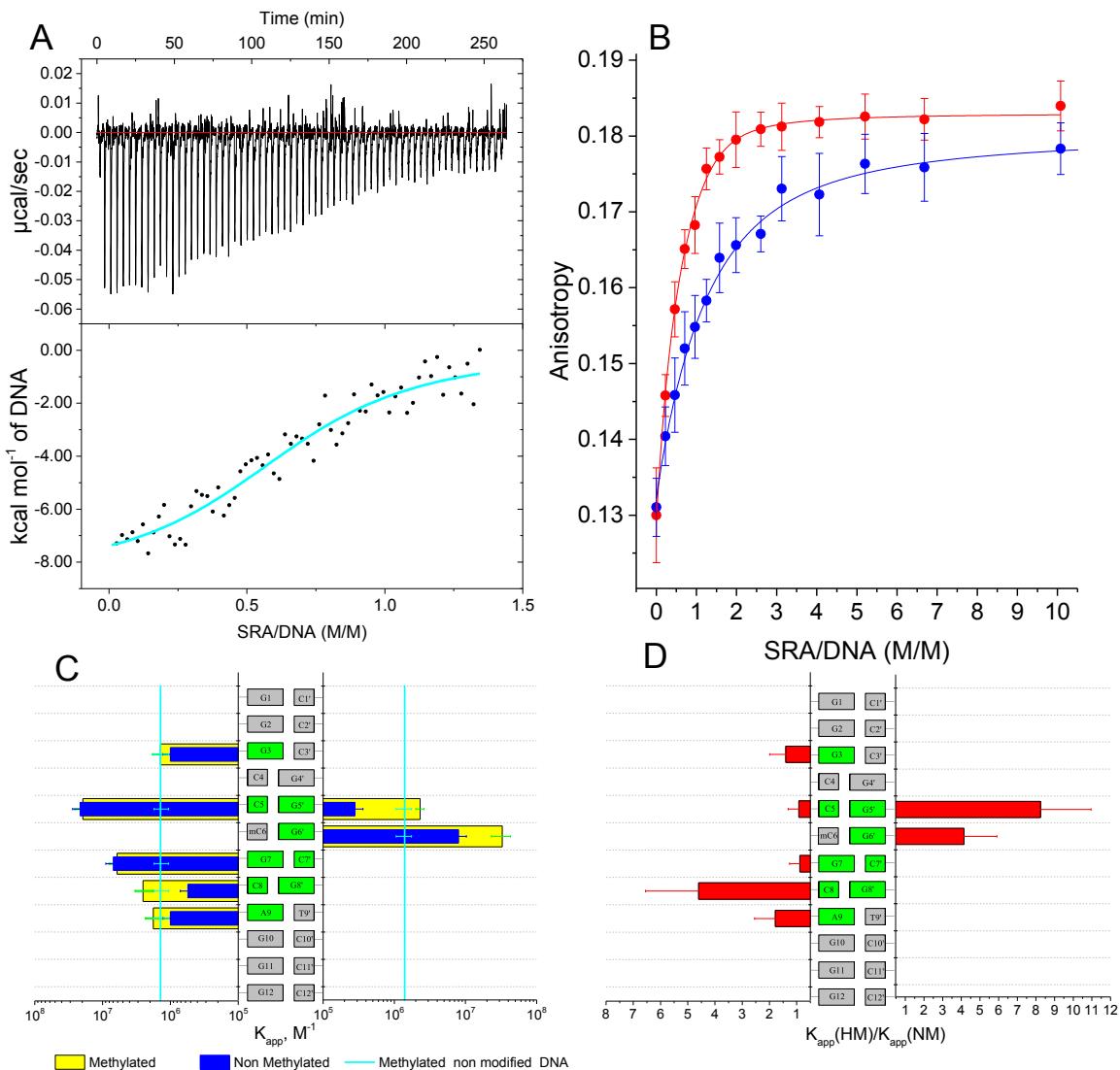


Figure 4. Binding of SRA to non-labeled HM duplex and to 3HCnt-labeled HM and NM duplexes. (A) Dependence of heat release on the time of injection (top) and the SRA/DNA molar ratio (bottom). The cyan line corresponds to the fit with a 1:1 binding model and the binding constant given in the text. B) Anisotropy titration curves for HM (red) and NM (blue) duplexes labeled by 3HCnt at position 5'. The concentration of duplexes was 2 μM in PBS buffer 20mM, pH 7.5 NaCl 50mM, TCEP 2.5mM, PEG 0.05% at 20°C. The lines correspond to the best fits of the experimental points to equation 1. Experimental points are expressed as means +/- standard deviation for n = 2 independent experiments. C) Plot of the affinity constants K_{app} for labeled HM (yellow) and NM (blue) duplexes. The cyan line describes the affinity of SRA to non-labeled HM duplex, as determined by ITC. D) Effect of cytosine 6 methylation on the affinity constants of SRA to the duplexes. At each position, the value of the ratio of the apparent affinity constant for the HM duplex to that of the NM duplex is given. Data in C) and D) are expressed as means +/- standard deviation for n = 3 independent experiments.

Interestingly, we noted a significant decrease of the affinity of the SRA mutant to the HM duplexes (Fig. 5A), in line with previous observations¹⁷. Moreover, the G448D mutation was also found to largely remove the preferential binding to HM duplexes, since the affinities of G448D to both HM and NM duplexes were found to be comparable, being on the level of those for the interaction of wild-type SRA with NM duplexes.

In sharp contrast to the wild-type SRA, the G448D SRA mutant induced only a limited increase in the QY (Fig. 5B) and N*/T* ratio (data not shown) of

the HM duplex labeled by 3HCnt at positions 8 or 5'. The changes in QY and N*/T* ratio observed on binding of the SRA mutant to the labeled HM duplexes were comparable to those observed on binding of both wild-type and mutant SRA to NM duplexes. These data thus confirm that the 3HCnt nucleobase at positions 8 and 5' can sense the SRA-induced flipping of mC6 and that the preferential affinity of SRA to HM duplexes relies on the ability of SRA to flip the mC6 residue into its binding pocket 17–19.

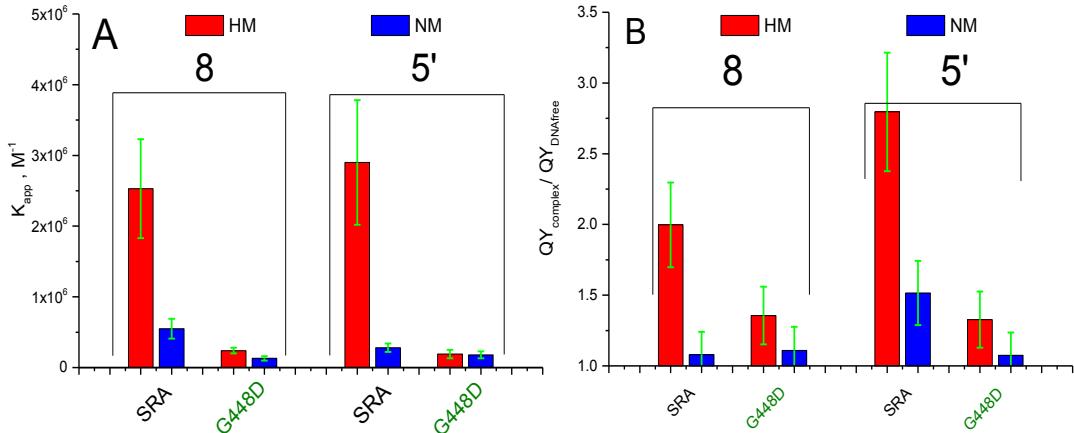


Figure 5. Comparison of the interaction of wild-type SRA and G448D SRA with HM and NM duplexes labeled with 3HCnt at positions 8 and 5'. (A) Affinity constants of wild-type and mutant SRA to the labeled NM and HM duplexes. (B) Changes in the QY of the labeled NM and HM duplexes on binding to wild-type and mutant SRA. The concentration of duplexes was 1 μ M. The concentration of the proteins was adjusted to ensure 80% of binding. Experimental conditions were as in Fig 3 and 4.

Kinetics of base flipping

To further confirm and characterize the SRA-induced flipping of mC6, we comparatively investigated by stopped-flow the interaction of

wild-type and mutant SRA with HM and NM duplexes labeled at positions 8 and 5'. The interaction kinetics were monitored through the changes in the fluorescence intensity of the T* band of 3HCnt.

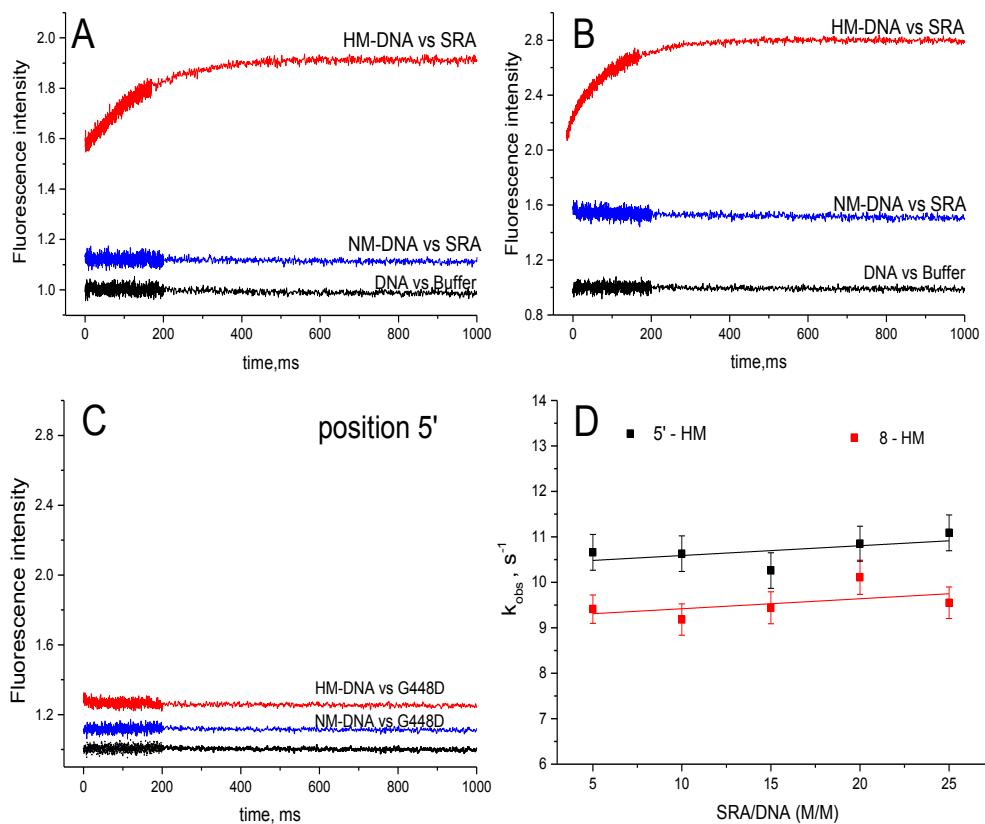


Figure 6. Interaction kinetics of wild-type SRA and G448D SRA mutant with HM and NM duplexes labeled at position 8 and 5', as monitored by stopped-flow technique. (A) Kinetic traces obtained with duplexes labeled by 3HCnt at position 8. The black curve corresponds to the control obtained with HM- or NM-duplexes mixed with the buffer. The blue and red solid lines are the kinetic traces obtained for the interaction of SRA with NM- and HM-duplexes, respectively. (B) Kinetic traces obtained with duplexes labeled by 3HCnt at position 5'. The color codes are as in (A). (C) Kinetic traces of the interaction of the G448D SRA mutant with HM and NM duplexes labeled by 3HCnt at position 5'. Color codes are as in (A). (D) Observed rate constants of the reaction of SRA with HM-duplexes labeled by 3HCnt at position 8 (red) and 5' (black) as a function of the ratio of SRA to DNA.

For NM duplexes, a non resolvable increase in fluorescence intensity was observed on mixing with SRA, indicating that most of the reaction was completed within the dead time (2 ms) of the instrument and thus, that the apparent bimolecular rate constant of the reaction is $> 10^9 \text{ M}^{-1} \cdot \text{s}^{-1}$ (Fig. 6A and B). This very fast kinetics with an apparent rate constant greater than the “diffusion-limited” one is typical of binding of proteins to nucleic acids^{41,42}. For the HM duplexes, a non resolvable increase in fluorescence intensity likely corresponding to the formation of the DNA/SRA complex was also observed. In sharp contrast to NM duplexes, the non resolvable fluorescence increase was followed by a much slower exponential increase of the fluorescence intensity that was completed in 0.5 s (Fig. 6 A and B). At the end of the process, the final fluorescence increase was fully consistent with the QY increase observed in steady-state spectroscopy (Fig. 2). This slow component, but not the initial non resolvable component, disappeared when the HM duplexes were mixed with the G448D mutant (Fig. C), strongly suggesting that this slow component may describe the kinetics of mC6 base flipping. To further confirm this attribution, we monitored the interaction kinetics of SRA with the 3HCnt-labeled HM duplexes, using SRA/DNA ratios varying from 5:1 to 25:1, and fitted the kinetic traces to single exponential functions. The observed rate constants were found to be independent on the concentration of SRA (Fig. 6D), indicating that the flipping reaction was first-order. This reaction order was fully expected as the flipping reaction can be envisioned as a conformational rearrangement that occurs after formation of the complex of SRA with the HM duplex.

Next, we performed dissociation experiments by adding an excess of calf thymus DNA (ctDNA) to pre-formed complexes of SRA with HM or NM duplexes labeled at the 5' position (Fig. 7). Labeling at the 5' position was preferred to the 8 position because it leads to larger changes in fluorescence intensity on binding to SRA (compare Fig. 6B with Fig. 6A). Addition of an excess of ctDNA to a preformed complex of SRA with 3HCnt-labeled NM duplex leads to a fast fluorescence decrease, likely related to the dissociation of the complex, with

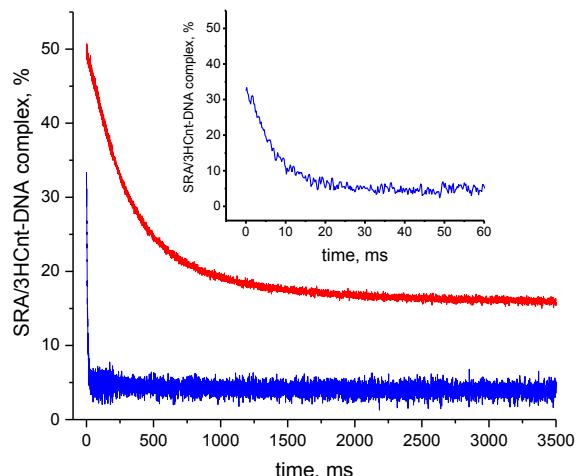


Figure 7. Dissociation kinetics of the complexes of SRA with HM (red) and NM (blue) duplexes labeled with 3HCnt at position 5'. The kinetics traces were recorded after addition of an excess of ctDNA to the complexes. The concentration of HM or NM duplexes was 0.3 μM . The SRA concentration was 1.8 and 3.6 μM for having 50% of complex with HM and NM duplexes, respectively. The concentration of ctDNA was 600 μM by molar concentration of nucleotides. For the kinetic traces, the measured fluorescence intensity was converted in percentage of 3HCnt-labeled duplexes bound to SRA. Inset: Highlight of the kinetic trace recorded with the NM-labeled duplex during the first 60 ms.

a rate constant of 150 s^{-1} . Under similar conditions, a much slower rate constant of 2.9 s^{-1} was observed when an excess of ctDNA (75-fold excess of ctDNA) was added to a preformed complex of SRA with 3HCnt-labeled HM duplex. This observed rate constant was found to be only weakly dependent on the ctDNA concentration (data not shown). Thus, the reaction associated with the slower reaction rate could be considered as being of first order, and likely corresponds to the flipping back of the mC6 into the HM duplex. Comparison of the dissociation curves obtained with HM and NM duplexes further suggests that the flipping back of the mC6 residue is the rate-limiting step and is immediately followed by the dissociation of the complex.

Taken together, the kinetic data clearly confirm that the 3HCnt at position 8 and 5' can monitor the SRA-induced flipping of the mC6 residue. Moreover, they allow us to determine for the first time the rate constants of the flipping reaction as well as of its back reaction.

Discussion

In the present study, we were able for the first time to monitor in diluted solution and real time the flipping of the mC6 residue of a HM duplex promoted by the SRA domain of UHRF1. This was achieved by inserting the environment-sensitive

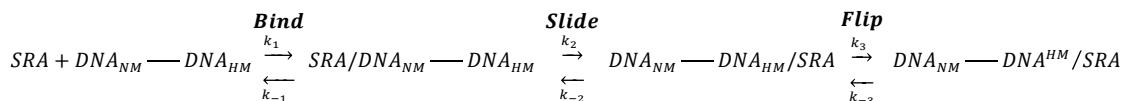
fluorescent nucleotide analogue 3HCnt at either position 8 or 5'. At these two positions, the 3HCnt probe marginally perturbed the binding of SRA and preserved the preferential binding to HM duplex over NM duplex (Figure 4). The flipping of the mC6 residue in the HM duplex labeled by 3HCnt at these positions was observed to strongly increase the emission intensity of the T* band of 3HCnt and decrease its N*/T* ratio (Figure 2). Monitoring by stopped-flow the 3HCnt signal changes occurring on interaction of the HM duplex with SRA further revealed that the kinetics of base flipping were much slower than the kinetics of SRA binding to its target (Figure 6). Finally, when the HM duplex was replaced by a NM duplex or the SRA domain by a G448D SRA mutant unable to perform flipping (Figure 5), the limited changes in 3HCnt spectroscopic properties as well as the suppression of the slow component in the kinetic traces clearly confirm that the observed spectroscopic changes upon binding were related to base flipping.

On a photophysical point of view, as the 3HCnt residue is thought to be highly quenched by stacking interactions with its neighbors G and C residues and as it undergoes an irreversible ESIPT reaction with a fast proton transfer rate constant⁴³ leading to rapid accumulation of the T* form, the observed spectroscopic changes likely result from a decrease of the stacking interactions with the T* form of the probe. These decreased stacking interactions could be rationalized from the 3D structure of the complex of SRA with the HM duplex¹⁷. In this complex, the Arg491 residue of SRA inserts in the hole left by the flipped mC6 residue in DNA and forms tight contacts with the C5 and G7 residues, as well as with the G5' and G6' residues. Thus, when 3HCnt is introduced at 8 position, the contacts of Arg491 with G7 likely reduces the dynamic stacking of 3HCnt by G7, explaining the increase in 3HCnt QY. Similarly, when 3HCnt is introduced at 5' position, the interaction of Arg491 with G6' likely restricts the dynamic quenching of 3HCnt by this residue.

This strong sensitivity of 3HCnt to mC6 base flipping appears quite unique, since in a previous report the 2-aminopurine probe was unable to sense the base flipping induced by UHRF1²⁵. Nevertheless, it should be stressed that the position of the 3HCnt in the duplex is critical to get relevant response to base flipping. Indeed, by labeling positions 5, 7 or 6', the 3HCnt was found to clearly increase the affinity for SRA and suppress the preferential binding to HM duplex, likely as a result of the multiple contacts of the nucleobases at these positions and the SRA amino acids¹⁷. The multiple contacts of 3HCnt with SRA may strongly restrict the dynamic quenching of 3HCnt by its neighbors, explaining the strong changes in the QY and N*/T* ratio at these positions for both HM and NM duplexes. Finally, the marginal changes in the spectroscopic properties of 3HCnt at position 3 and 9 may easily be rationalized by the limited interaction of SRA with the nucleobases at these positions¹⁷.

Using the 3HCnt label at position 8 or 5', we were able for the first time to measure the kinetics of interaction of SRA with its DNA target as well as the kinetics of base flipping. The kinetics of interaction were found to be faster than a "diffusion-limited" reaction, suggesting that the interaction proceeds through a two-step "bind-and slide" reaction mechanism (scheme 1), as classically described for protein/DNA interactions^{41,42,44}

In this mechanism, the protein binds nonspecifically to the duplex, and then slides to the CpG recognition site. As the sliding corresponds to a one-dimensional diffusion, this mechanism reduces the dimensionality of the search, explaining why apparent rate constants faster than diffusion-limited rate constants could be obtained. Due to the inability to kinetically resolve the binding reaction, we could not determine the values of k_1 , k_2 , k_{-1} and k_{-2} . In contrast, as the flipping



Scheme 1. Proposed reaction mechanism of SRA interaction with HM duplex. For the interaction with NM duplex, the third step is likely absent.

reaction was much slower, it could be represented as the third step in scheme 1. The kinetic rate constants were determined to be $k_3 = 10 \text{ s}^{-1}$ and $k_{-3} = 2.9 \text{ s}^{-1}$. From the ratio of these values, it can be seen that the flipping reaction and the accompanying conformational changes stabilize the complex by about 3.5-fold. This stabilization is in very good agreement with the four-fold increased affinity observed for HM over NM duplex²⁵. Importantly, the dissociation experiments with NM and HM duplexes (Fig. 7) indicate that the k_{-3} value appeared much slower than the k_2 and k_1 values. This suggests that the lifetime of the complex of UHRF1 bound to a CpG site in HM duplexes is much longer than in NM duplexes. This will obviously increase the time and the probability to recruit DNMT1, in order to duplicate DNA methylation profile. In contrast, the much shorter lifetime of SRA bound to NM duplexes is consistent with a reader role of SRA, able to slide rapidly along the DNA to scan for hemi-methylated CpG sites^{45,46}.

In conclusion, the environment-sensitive 3HCnt fluorescent nucleoside analogue appears as a unique tool to monitor with high sensitivity the SRA-induced flipping of methylcytosine and its dynamics. This sensitive tool will allow monitoring the cooperation of SRA with other domains of UHRF1, *e.g.*, the tandem tudor domain (TTD) that shows affinity for the methylated lysine at position 9 of histone H3⁴⁷. UHRF1 ensures the link between DNA methylation and histone methylation thanks to the readout capacities of both its SRA domain and TTD^{23,48,49}, which questions whether the different structural domains cooperate to regulate the affinity of each domain towards its target site. Consequently, there is a great need to investigate how the SRA domain functions to flip out the methylated cytosine in connection with other putative domains of the UHRF1 protein and/or of its partners, such as DNMT1. This work is currently under progress.

Acknowledgements

We thank the ANR (ANR-12-BS08-0003-02) and the FRM (DCM20111223038) for financial support and fellowships to V.K. and N.P.F.B.

References

1. Feinberg, A. P. & Tycko, B. The history of cancer epigenetics. *Nat. Rev. Cancer* **4**, 143–153 (2004).
2. Esteller, M. Epigenetics in cancer. *N. Engl. J. Med.* **358**, 1–12 (2008).
3. Jones, P. a. & Baylin, S. B. The Epigenomics of Cancer. *Cell* **128**, 683–692 (2007).
4. Laird, P. W. The power and the promise of DNA methylation markers. *Nat. Rev. Cancer* **3**, 253–266 (2003).
5. Ng, H. & Bird, A. DNA methylation and chromatin modification. *Genet. Dev.* **9**, 158–163 (1999).
6. Jenuwein, T. & Allis, C. D. Translating the histone code. *Science* **293**, 1074–1080 (2001).
7. Berger, S. L., Kouzarides, T., Shiekhattar, R. & Shilatifard, A. An operational definition of epigenetics. *Genes Dev.* **23**, 781–783 (2009).
8. Berger, S. L. The complex language of chromatin regulation during transcription. *Nature* **447**, 407–412 (2007).
9. Lorincz, M. C., Dickerson, D. R., Schmitt, M. & Groudine, M. Intragenic DNA methylation alters chromatin structure and elongation efficiency in mammalian cells. *Nat. Struct. Mol. Biol.* **11**, 1068–1075 (2004).
10. Kouzarides, T. Chromatin modifications and their function. *Cell* **128**, 693–705 (2007).
11. Li, B., Carey, M. & Workman, J. L. The role of chromatin during transcription. *Cell* **128**, 707–719 (2007).
12. Spruijt, C. G. & Vermeulen, M. P ERS P ECTI V E DNA methylation : old dog , new tricks ? *Nat. Publ. Gr.* **21**, 949–954 (2014).
13. Straussman, R. *et al.* Developmental programming of CpG island methylation profiles in the human genome. *Nat. Struct. Mol. Biol.* **16**, 564–571 (2009).
14. Bird, A. DNA methylation patterns and epigenetic memory DNA methylation patterns and epigenetic memory. *Genes Dev.* **16**, 6–21 (2002).
15. Jeltsch, A. Reading and writing DNA methylation. *Nature structural & molecular biology* **15**, 1003–1004 (2008).
16. Bestor, T. H. & Ingram, V. M. Two DNA methyltransferases from murine erythroleukemia cells: purification, sequence specificity, and mode of interaction with DNA. *Proc. Natl. Acad. Sci. U. S. A.* **80**, 5559–5563 (1983).
17. Avvakumov, G. V *et al.* Structural basis for recognition of hemi-methylated DNA by the SRA domain of human UHRF1. *Nature* **455**, 822–825 (2008).
18. Hashimoto, H. *et al.* The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix. *Nature* **455**, 826–829 (2008).
19. Arita, K., Ariyoshi, M., Tochio, H., Nakamura, Y. & Shirakawa, M. Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism. *Nature* **455**, 818–821 (2008).
20. Bostick, M. *et al.* UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* **317**, 1760–1764 (2007).
21. Arita, K., Ariyoshi, M., Tochio, H., Nakamura, Y. & Shirakawa, M. Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism. *Nature* **455**, 818–821 (2008).

22. Qian, C. *et al.* Structure and hemimethylated CpG binding of the SRA domain from human UHRF1. *J. Biol. Chem.* **283**, 34490–34494 (2008).
23. Rottach, A. *et al.* The multi-domain protein Np95 connects DNA methylation and histone modification. *Nucleic Acids Res.* **38**, 1796–1804 (2009).
24. Frauer, C. *et al.* Recognition of 5-hydroxymethylcytosine by the Uhrf1 SRA domain. *PLoS One* **6**, 1–8 (2011).
25. Greiner, V. J. *et al.* Site-Selective Monitoring of the Interaction of the SRA Domain of UHRF1 with Target DNA Sequences Labeled with 2-Aminopurine. *Biochemistry* **54**, 6012–6020 (2015).
26. Dziuba, D. *et al.* A universal nucleoside with strong two-band switchable fluorescence and sensitivity to the environment for investigating DNA interactions. *J. Am. Chem. Soc.* **134**, 10209–10213 (2012).
27. Klymchenko, A. S. & Demchenko, A. P. Electrochromic modulation of excited-state intramolecular proton transfer: the new principle in design of fluorescence sensors. *J. Am. Chem. Soc.* **124**, 12372–12379 (2002).
28. Kilin, V. *et al.* Fluorescence Lifetime Imaging of Membrane Lipid Order with a Ratiometric Fluorescent Probe. *Biophys. J.* **108**, 2521–2531 (2015).
29. Demchenko, A. P., Tang, K. C. & Chou, P. T. Excited-state proton coupled charge transfer modulated by molecular structure and media polarization. *Chem. Soc. Rev.* **42**, 1379–1408 (2013).
30. Kuznetsova, A. a. *et al.* New environment-sensitive multichannel DNA fluorescent label for investigation of the protein-DNA interactions. *PLoS One* **9**, e100007 (2014).
31. Delagoutte, B., Lallous, N., Birck, C., Oudet, P. & Samama, J. P. Expression, purification, crystallization and preliminary crystallographic study of the SRA domain of the human UHRF1 protein. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **64**, 922–925 (2008).
32. Achour, M. *et al.* Epigallocatechin-3-gallate up-regulates tumor suppressor gene expression via a reactive oxygen species-dependent down-regulation of UHRF1. *Biochem. Biophys. Res. Commun.* **430**, 208–212 (2013).
33. Tomin, V. I. & Jaworski, R. SPECTRAL PROPERTIES OF THE ANIONIC FORM OF THE MOLECULAR PROBE FET. *J. Polym. Sci. Part A: Polym. Chem.* **81**, 360–364 (2014).
34. Tomin, V. I. & Jaworski, R. Temporal characteristics of the fluorescence of the anionic form of 3-hydroxyflavone. *Opt. Spectrosc.* **103**, 952–957 (2007).
35. Vuilleumier, C., Maechling-Strasser, C., Gérard, D. & Mély, Y. Evidence and prevention of HIV-1 nucleocapsid protein adsorption onto fluorescence quartz cells. *Anal. Biochem.* **244**, 183–185 (1997).
36. Breslauer, K. J. in *Energetics of Biological Macromolecules* (ed. Michael L. Johnson, G. K. A. B. T.-M. in E.) **259**, 221–242 (Academic Press, 1995).
37. Spadafora, M. *et al.* Efficient Synthesis of Ratiometric Fluorescent Nucleosides Featuring 3-Hydroxychromone Nucleobases. *Tetrahedron* **65**, 7809–7816 (2009).
38. Rochel, N. *et al.* Common architecture of nuclear receptor heterodimers on DNA direct repeat elements with different spacings. *Nat Struct Mol Biol* **18**, 564–570 (2011).
39. Systems, I. S. on R. A. in D. D. & Anderson, J. M. Advances in drug delivery systems, 6: proceedings of the Sixth International Symposium on Recent Advances in Drug Delivery Systems, Salt Lake City, UT, U.S.A., February 21-24, 1993. (1994). at <<http://www.sciencedirect.com/science/book/9780444820273>>
40. Maillot, S. *et al.* Out-of-equilibrium biomolecular interactions monitored by picosecond fluorescence in microfluidic droplets. *Lab Chip* **14**, 1767–1774 (2014).
41. Von Hippel, P. H. & Berg, O. G. Facilitated target location in biological systems. *J. Biol. Chem.* **264**, 675–678 (1989).
42. Berg, O. G., Winter, R. B. & von Hippel, P. H. Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. *Biochemistry* **20**, 6929–6948 (1981).
43. Dziuba, D. *et al.* Rational design of a solvatochromic fluorescent uracil analogue with a dual-band ratiometric response based on 3-hydroxychromone. *Chemistry* **20**, 1998–2009 (2014).
44. Winter, R. B., Berg, O. G. & von Hippel, P. H. Diffusion-driven mechanisms of protein translocation on nucleic acids. 3. The Escherichia coli lac repressor–operator interaction: kinetic measurements and conclusions. *Biochemistry* **20**, 6961–6977 (1981).
45. Bronner, C., Fuhrmann, G., Chédin, F. L., Macaluso, M. & Dhe-Paganon, S. UHRF1 Links the Histone code and DNA Methylation to ensure Faithful Epigenetic Memory Inheritance. *Genet. Epigenet.* **2009**, 29–36 (2010).
46. Hashimoto, H., Vertino, P. M. & Cheng, X. Molecular coupling of DNA methylation and histone methylation. *Epigenomics* **2**, 657–669 (2010).
47. Rothbart, S. B. *et al.* Association of UHRF1 with methylated H3K9 directs the maintenance of DNA methylation. *Nat Struct Mol Biol* **19**, 1155–1160 (2012).
48. Bronner, C., Fuhrmann, G., Chédin, F. L., Macaluso, M. & Dhe-Paganon, S. UHRF1 links the histone code and DNA methylation to ensure faithful epigenetic memory inheritance. *Genet. Epigenetics* **2009**, 29–36 (2009).
49. Liu, X. *et al.* UHRF1 targets DNMT1 for DNA methylation through cooperative binding of hemi-methylated DNA and methylated H3K9. *Nat. Commun.* **4**, 1563 (2013).

B. Seconde stratégie de marquage (grand sillon)

PUBLICATION 2

En plus de la première stratégie de marquage qui permet de sonder l'intérieur du duplexe d'ADN, nous avons également envisagé la modification d'une nucléobase naturelle afin de pouvoir sonder l'intérieur du grand sillon de l'ADN. Pour ce faire, le rapporteur fluorescent devait être attaché par l'intermédiaire d'un bras espaceur court et rigide sur la nucléobase (en 5 de pyrimidine ou 7 de déazapurine). Le double brin d'ADN hémiméthylé sur lequel s'appuie notre étude du mécanisme de méthylation de l'ADN [cf. Figure 4A], présente sur son brin complémentaire dodécamère, une thymidine à proximité du site de liaison avec SRA. D'après la structure du complexe ADN / SRA obtenue par diffraction des rayons X, la dT n'est pas impliquée dans la liaison entre la protéine et le double brin d'ADN. L'introduction d'un analogue de dT couplé à la chromone sur cette position ne devrait pas perturber la liaison de SRA pour sa cible ADN. Ainsi, nous nous sommes consacrés à l'élaboration et à la synthèse d'un analogue fluorescent de dU, marqué à la 3-HC sur la position 5 de l'uracile via un bras espaceur acétylénique.

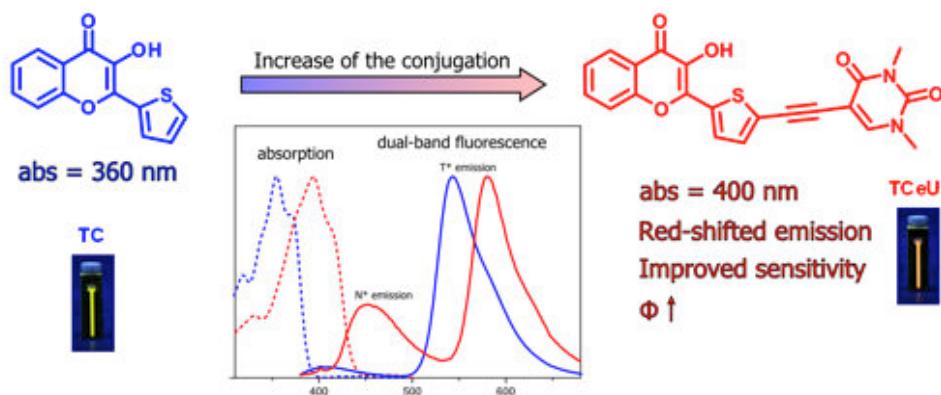


Figure 6. Modèle simplifié (TCU), diméthylé sur les azotes 1 et 3 de l'uridine.

Dans un premier temps, il était nécessaire de déterminer les propriétés photophysiques d'un tel rapporteur pour pouvoir définir les avantages qu'il apporterait. Nous avons donc décidé de synthétiser un analogue simplifié **TCeU**, diméthylé au niveau des azotes 1 et 3 de l'uracile [cf. Figure 6]. Les études de fluorescence à l'état stationnaire et résolues en temps ont permis de mettre en évidence un transfert de charge intramoléculaire entre l'azote 1 de l'uracile (donneur) et le carbonyle 4 de la chromone (accepteur), à l'origine d'un effet bathochrome important (déplacement des maxima d'absorption et d'émission vers le rouge de 40 nm). La sonde présente de bons rendements quantiques (de 15 à 25 %) ainsi qu'une variation du rapport d'intensité des deux bandes d'émission importante, sur une gamme de solvants protiques et aprotiques de polarités différentes plus large que celle déterminée pour la sonde de 1^{ère} génération. Enfin, elle possède un coefficient d'absorption molaire également bien supérieur ($35\ 000\ L.\text{mol}^{-1}.\text{cm}^{-1}$ au lieu de $13\ 000\ L.\text{mol}^{-1}.\text{cm}^{-1}$). Ces travaux firent l'objet d'un article, publié début Janvier 2014, dans

Chemistry – A European Journal. Ces résultats prometteurs nous ont conforté dans la poursuite de nos travaux pour cette stratégie de marquage.^{126f}

Fluorescent Probes

Rational Design of a Solvatochromic Fluorescent Uracil Analogue with a Dual-Band Ratiometric Response Based on 3-Hydroxychromone

Dmytro Dziuba,^[a, b] Iuliia A. Karpenko,^[a, c] Nicolas P. F. Barthes,^[a] Benoît Y. Michel,^[a] Andrey S. Klymchenko,^[d] Rachid Benhida,^[a] Alexander P. Demchenko,^[e] Yves Mély,^[d] and Alain Burger^{*[a]}

Dedicated to Professor J.-F. Biellmann

Abstract: Fluorescent nucleoside analogues with strong and informative responses to their local environment are in urgent need for DNA research. In this work, the design, synthesis and investigation of a new solvatochromic ratiometric fluorophore compiled from 3-hydroxychromones (3HCs) and uracil fragments are reported. 3HC dyes are a class of multi-parametric, environment-sensitive fluorophores providing a ratiometric response due to the presence of two well-resolved bands in their emission spectra. The synthesized conjugate demonstrates not only the preservation but also the improvement of these properties. The absorption and fluo-

rescence spectra are shifted to longer wavelengths together with an increase of brightness. Moreover, the two fluorescence bands are better resolved and provide ratiometric responses across a broader range of solvent polarities. To understand the photophysical properties of this new fluorophore, a series of model compounds were synthesized and comparatively investigated. The obtained data indicate that uracil and 3HC fragments of this derivative are coupled into an electronic conjugated system, which on excitation attains strong charge-transfer character. The developed fluorophore is a prospective label for nucleic acids.

Introduction

Fluorescent analogues of non-emissive natural nucleosides found broad applications in studying structure, dynamic prop-

erties and interactions of nucleic acids.^[1] Artificial fluorescent nucleosides can be site-selectively incorporated into nucleic acid strands by an enzymatic polymerization reaction or by chemical synthesis. As a result, a well-defined location and geometry of the fluorophore within the nucleic acid structure can be achieved together with the ability to follow its functional transformations.^[2] This opens many possibilities that cannot be realized with other types of covalent fluorescent labels. Advanced fluorescent nucleosides allow for the sensing of nucleic acid hybridization and conformational changes,^[3] typing of single-nucleotide polymorphisms (SNPs),^[4] and sensing the activity of DNA-binding proteins and enzymes.^[5,6] These results stimulate new efforts to synthesize advanced emissive nucleosides based on new mechanisms of response. Fluorescence offers a limited number of parameters for quantitative detection. Typically, the events occurring in the microenvironment of fluorophores may be monitored by changes in: 1) fluorescence intensity; 2) fluorescence anisotropy; 3) excited-state lifetime; and 4) position of the emission band maximum (emission color).^[7] Analytical assays based on fluorescent nucleosides commonly use one of these modalities. A classic example is 2-aminopurine (2AP), a constitutional isomer of adenine, changing emission intensity based on its local environment. 2AP fluoresces brightly in water ($\Phi=0.68$), whereas stacking interactions with neighboring nucleobases within the nucleic acid strands strongly quenches 2AP fluorescence.^[8] As a consequence, quantitative determination of fluorescence in-

[a] Dr. D. Dziuba, I. A. Karpenko, N. P. F. Barthes, Dr. B. Y. Michel, Dr. R. Benhida, Prof. Dr. A. Burger
Institut de Chimie de Nice, UMR 7272
Université de Nice Sophia Antipolis, CNRS
Parc Valrose, 06108 Nice cedex 2 (France)
E-mail: burger@unice.fr

[b] Dr. D. Dziuba
Present address: Institute of Organic Chemistry and Biochemistry
Academy of Sciences of the Czech Republic
Flemingovo nam. 2, CZ-16610, Prague 6 (Czech Republic)

[c] I. A. Karpenko
Present address: Laboratoire d'Innovation Thérapeutique
UMR 7200, Université de Strasbourg, CNRS, Faculté de Pharmacie
74 route du Rhin, 67401 Illkirch (France)

[d] Dr. A. S. Klymchenko, Prof. Dr. Y. Mély
Laboratoire de Biophotonique et Pharmacologie, UMR 7213
Faculté de Pharmacie, Université de Strasbourg, CNRS
74 Route du Rhin, 67401 Illkirch (France)

[e] Prof. Dr. A. P. Demchenko
A. V. Palladin Institute of Biochemistry
9 Leontovicha Street, Kiev 01030 (Ukraine)

 Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/chem.201303399> (including ^1H - ^{13}C NMR spectra of all compounds, representative ^1H - ^1H COSY, ^1H - ^{13}C HMQC and ^1H - ^{13}C HMBC spectra, Figure S1, and Tables S1 and S2 from steady-state fluorescent measurements).

tensity may be problematic because it depends on the concentration and is sensitive to unaccountable quenching effects. Application of color-changing nucleosides can only partially resolve this problem; typically they are molecules that possess a highly dipolar excited state and exhibit polarity-dependent shifts in their emission together with the intensity changes, which can be rather strong. Both parameters can monitor changes occurring in their microenvironment.^[9] Meanwhile, sensing based on fluorescent intensity or spectral shifts suffers from several limitations, related mainly to the problems of the calibration of the fluorescence response, which is commonly displayed in relative units. To overcome these limitations, an approach based on the λ -ratiometric sensing can be used.^[10] Based on recording of the fluorescence signal as the ratio of intensities at two or more wavelengths, this sensing has undergone rapid development during recent years. Such simple two-channel probing makes possible to record an intrinsically calibrated analytical signal that does not depend on local probe concentration or instrumentation settings. Analytical assays for nucleic acids based on λ -ratiometry are known. The vast majority of them exploit double-dye systems, such as fluorescence resonance energy transfer (FRET) pairs and formation of excimers.^[11] Nevertheless, this method requires double labeling to obtain informative signals. An attractive alternative using a single dye is still poorly explored because of the very limited number of dyes that can display this property and of the difficulties in their synthesis.^[10,12] In this context, the focus of our research is the synthesis of fluorescent nucleosides con-

taining single ratiometric emitters. With this aim, we concentrated our efforts on fluorescent nucleosides incorporating 3-hydroxychromone fluorophores as single-dye ratiometric reporters.

3-Hydroxychromones (3HCs) are a prospective class of fluorescent dyes offering exceptional solvatochromism together with dual-band fluorescence suitable for ratiometric detection. The two emission bands originate from a reaction occurring in the excited state: the excited-state intramolecular proton transfer (ESIPT, Figure 1).^[13] Upon excitation, 3HC molecules

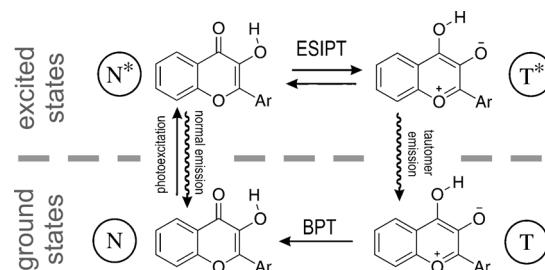


Figure 1. Origin of the dual emission of 3-hydroxychromones (ESIPT = excited-state intramolecular proton transfer; BPT = ground state back-proton transfer; N and T denote the normal and tautomer forms, respectively).

attain the normal excited form (N^*), which then undergoes an ESIPT reaction to produce the tautomer (T^*) excited form. Both states are highly emissive, the T^* band being typically 80–100 nm redshifted compared with the N^* band (Figure 1).

The dual emission of 3HCs is highly sensitive to the polarity of the surrounding medium.^[14] In particular, the ratio of the two emission bands changes with respect to the dielectric constant and H-bond donor ability of the environment.^[15] High values of dielectric constant and strong H-bond acidity of the solvent stabilize the N^* state and inhibit the ESIPT reaction, thus increasing the N^*/T^* intensity ratio. Therefore, the ratio of the two emission bands provides information about the microenvironment of the dye. In addition to the N^*/T^* ratio, information may be extracted from the relative positions of the two bands, so that 3HCs are multiparametric fluorescent probes.^[15] Over the past two decades, 3HCs have found applications in sensing polarity of normal^[16] and supercritical liquids,^[17] probing lipid membranes^[18] and proteins.^[19] However, applications of 3HC-based probes to nucleic acid sensing remain meager.

The work presented herein expands our earlier efforts in the use of 3HC for nucleic acid probing. We showed that the H-bond-sensitive fluorophore 2-furyl-3-hydroxychromone conjugated to a polyamine chain can distinguish single- from double-stranded DNA through its two-band ratiometric response.^[20] Subsequently, we synthesized a fluorescent nucleoside, which included the environment-sensitive 2-thienyl-3-hydroxychromone as a nucleobase mimic.^[21] Incorporation of the nucleoside into oligonucleotides induced only a small destabilization of the duplex structure. Most importantly, the 3HC nucleoside showed significant changes in its ratiometric response upon binding to a complementary strand or as the result of the conformational changes caused by the chaperone viral nu-

Abstract in Ukrainian:

Флуоресцентні аналоги нуклеозидів, що володіють високою чутливістю до навколошнього оточення, є важливими для досліджень ДНК. В цій роботі ми представляємо дизайн, синтез та вивчення нового сольватохромного ратіометричного флуорофору, що складається з фрагментів 3-гідроксихромону (3ГХ) та урацилу. 3-Гідроксихромони є класом мультипараметричних чутливих до оточення флуорофорів, що дають ратіометричний сигнал через наявність двох добре розділених смуг в спектрах емісії. Сполучки, описані в цій роботі, показують не тільки збереження, а й покращення ратіометричних властивостей. Спектри поглинання та флуоресценції зміщені в червону область, яскравість флуоресценції збільшилась. Більш того, дві смуги флуоресценції виявилися краще розділеними, а також дають ратіометричну відповідь у більш широкому інтервалі полярностей середовища. Для аналізу фотофізичних властивостей нового флуорофору, ми порівняли його з спеціально синтезованим рядом модельних сполучк. Отримані дані свідчать про те, що фрагменти урацилу та ЗГХ у флуорофорі утворюють єдину спряжену електронну систему, що зазнає переносу заряду у збудженному стані. Розроблений нами флуорофор є перспективним для дослідження нуклеїнових кислот.

cleocapsid protein.^[22] According to these data, 3HCs are prospective fluorophores for probing nucleic acids and their intermolecular interactions.

In this work, we developed a new 3HC-based ratiometric fluorophore with improved properties for the purpose of nucleic acid probing. We designed and synthesized a fluorophore composed of 3HC and uracil aromatic fragments assembled into a conjugated electronic system through an electron-conducting acetylenic linker and investigated its solvatochromic properties. The fluorophore showed improved fluorescence properties, including redshifted absorption and emission, higher brightness, and a new sensitivity profile to H-bond interactions. The dye constitutes a promising building block for incorporation into nucleic acids.

Results and Discussion

Design of the fluorophore

Design of new fluorophores with predefined photophysical characteristics is a great challenge. To reach this aim, the approach based on modification of known fluorescent cores appears particularly fruitful.^[23] For designing a new fluorophore based on the 3-hydroxychromone core for nucleic acid labeling, several requirements need to be met: 1) the target molecule should contain a 3HC moiety with an unprotected hydroxyl group to use ESIPT as the principal sensing mechanism; 2) the fluorophore should be connected with one of the natural nucleobases to allow its incorporation into DNA as a corresponding nucleoside; 3) the 3HC and nucleobase parts should be connected through a short rigid linker to ensure a defined localization of the fluorophore with respect to the nucleic acid main axis. Based on these requirements, the following chemical structure of the fluorophore **1** (named **TCeU**) was proposed.

Fluorophore **1** is composed of two fragments, namely 2-thienyl-3-hydroxychromone, which was previously shown to be a good ratiometric reporter sensitive to H bonding^[21] and uracil, as the simplest member of the genetic alphabet. 1,3-Dimethyl uracil was used to exclude any prototropic tautomerisation in the nucleobase part, which could trigger side reactions such as 5-endo cyclization during fragment assembly. An acetylenic linker was introduced between the 5' position of the thienyl ring and the 5 position of uracil. Such a modification of uracil is commonly used in the synthesis of fluorescent nucleosides with new photophysical characteristics.^[1b, 9] Moreover, the selected configuration of the nucleobase and chromone moieties is expected to allow an intramolecular charge-transfer (ICT) between 4-carbonyl of chromone and N-1 atom of

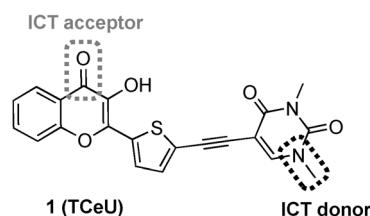


Figure 2. Proposed structure of the electronically connected 2-thienyl-3-hydroxychromone and uridine fragments.

uracil (Figure 2). ICT was shown to play an important role in modulating the environment-sensing properties of 3HC dyes.^[14a, 24] Furthermore, it has been shown for 2-thienyl-3-hydroxychromones that their 5'-position is highly sensitive to electron-donating groups.^[25]

To better understand the photophysics of this new compound, we compared it with two analogues, namely the parent 2-thienyl-3-hydroxychromone **2** (**TC**) and a **TC** bearing an acetylene moiety **3** (**TCe**; Figure 3).

In addition, we have synthesized and characterized for further comparison the methylated analogues of these three compounds **4**, **5** and **6** (named **MTC**, **MTCe**, and **MTCeU**, respectively) that are unable to undergo ESIPT.

Synthesis

Synthesis of the fluorophore **1 TCeU** and the model compounds is depicted in Scheme 1 and Scheme 2. Several synthetic strategies are reported to elaborate the 3HC scaffold. In our study, the Algar–Flynn–Oyamada (AFO) reaction was employed as the most facile and abundant reaction for constructing the 3-hydroxychromone core. Appreciable yields of the target chromones were achieved by using a “soft” one-pot protocol of the AFO reaction (Scheme 1).^[26] An important feature of this protocol is the possibility to obtain pure product with minimal manipulation. According to this procedure, 2-hydroxyacetophenone **7** was treated with thiophene-aldehydes **8** and **9** in ethanol in the presence of sodium hydroxide with subsequent addition of hydrogen peroxide and acidic work-up to give 3-hydroxychromones **2** (**TC**) and **10** in 50 and 48%

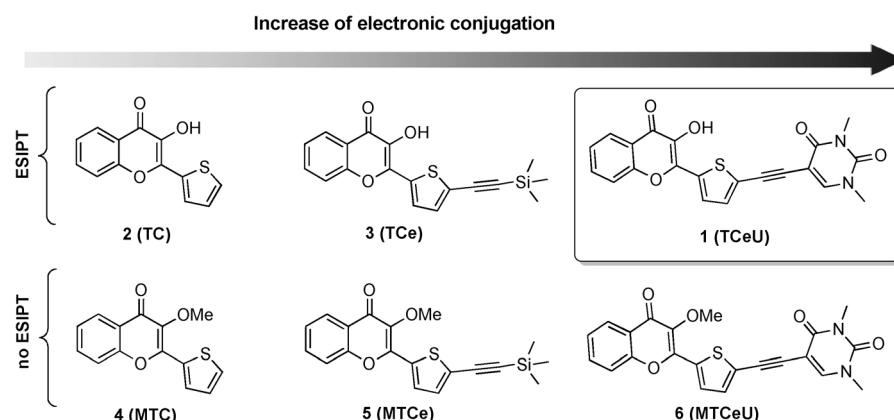
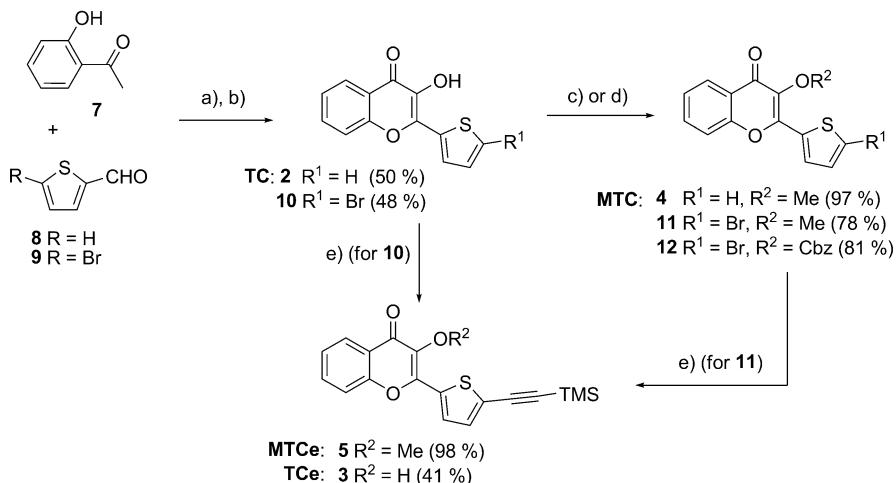
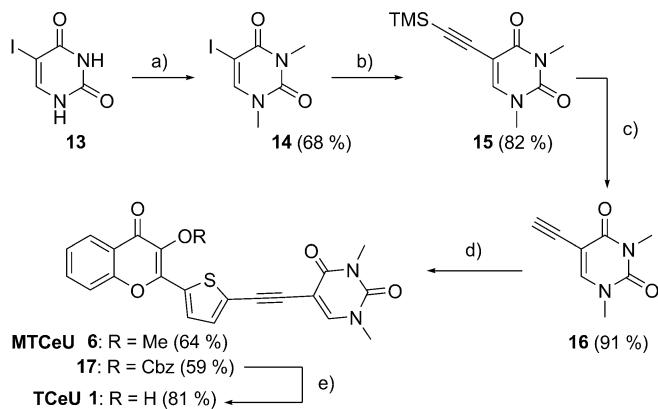


Figure 3. Structure of the designed fluorophore **TCeU** and model compounds.



Scheme 1. Synthesis of 2-thienyl-3-hydroxychromones. Reagents and conditions: a) NaOH, EtOH, RT, 48 h; b) H₂O₂ (30% aq.); c) Me₂SO₄, KOH, 18-crown-6, CH₂Cl₂/H₂O, RT, 1 h; d) Cbz-Cl, KOH, 18-crown-6, CH₂Cl₂/H₂O, RT, 1 h; e) [PdCl₂(PPh₃)₂], CuI, NEt₃, THF, reflux.



Scheme 2. Modular synthesis of 3HC-uracil fluorophores. Reagents and conditions: a) Me₂SO₄, NaOH, H₂O, 100°C, 2 h; b) TMS-acetylene, [PdCl₂(PPh₃)₂], CuI, DIPEA, THF, 55°C, 4 h; c) TBAF, THF, RT, 1 h; d) 11 or 12, [PdCl₂(PPh₃)₂], CuI, DIPEA, THF, 55°C; e) NH₄OH, MeOH, RT, 1 h.

yields, respectively.^[27] Derivative **4** (**MTC**) was obtained in 97% yield after methylation of **TC** by dimethyl sulfate under the phase-transfer catalysis (PTC) conditions reported previously.^[28] Compounds **2** (**TC**) and **4** (**MTC**) were used for spectroscopic characterizations. Bromo-chromone **10** was converted under the same conditions into the methylated derivative **11** needed for the synthesis of the model compound **MTCe**.

The synthesis of the designed fluorophore **TCeU** requires an appropriate protection for the 3-OH group of the chromone. We selected a carbonate-type protecting group, the benzyloxy-carbonyl group (known as Cbz or Z), which was successfully employed in our recent work for the synthesis of 3HC-containing nucleosides.^[22] The Cbz group was introduced onto bromo-chromone **10** by treatment with Cbz-Cl under PTC conditions. Sonogashira coupling with TMS-acetylene and compounds **10** and **11** resulted in moderate to excellent yields of

the model compounds **TCe** (**3**) and **MTCe** (**5**) required for the photophysical studies.

Concerning the preparation of the nucleobase fragment, 5-iodouracil **13** was first bis-methylated on its N-1 and N-3 positions as reported.^[29] A Sonogashira coupling in presence of TMS-acetylene provided the alkyne derivative **15** in good yields. Removal of TMS by treatment with tetrabutylammonium fluoride (TBAF) led to the required fragment for the final assembly. Sonogashira coupling between fluorescent moieties **11** and **12** and the nucleobase part **16** gave the desired dye **MTCeU** (**6**) and **17** in appreciable yields. The Cbz masking the 3-hydroxyl

group was cleaved by a simple treatment with aqueous ammonia in methanol at room temperature to afford the target fluorophore **TCeU** (**1**).

Spectroscopic properties of the designed fluorophore

UV/Vis absorption and fluorescence properties of the designed fluorophore as well as of the model compounds were investigated in a set of eleven solvents varying in polarity (Table 1).

Table 1. Spectroscopic properties of **TCeU** in different solvents.

Solvent	$E_T(30)^{[a]}$	$\sum \beta_2^H$ ^b	$\lambda_{abs}^{[c]}$	$\lambda_{N^*}^{[d]}$	$\lambda_{T^*}^{[e]}$	I_{N^*}/I_{T^*} ^f	$\Phi^{[g]}$
H ₂ O	63.1	0.35	401	500	—	—	0.08
MeOH	55.4	0.47	395	501	558	1.20	0.22
EtOH	51.9	0.48	397	488	574	0.63	0.18
BuOH	49.7	0.48	400	481	571	0.50	0.19
CH ₃ CN	45.6	0.32	392	472	576	0.54	0.15
DMSO	45.1	0.88	402	477	582	1.43	0.25
acetone	42.2	0.49	393	465	581	0.53	0.15
CH ₂ Cl ₂	40.7	0.05	397	458	577	0.31	0.17
EtOAc	38.1	0.45	394	451	579	0.36	0.15
THF	37.4	0.48	396	448	585	0.73	0.18
PhCH ₃	33.9	0.14	399	441	587	0.13	0.15

[a] Reichardt's empirical solvent polarity index.^[30] [b] H-bond basicity.^[15,31] [c] Position in nm of the absorption maximum. [d] Position in nm of the maximum of the normal emission band. [e] Position in nm of the maximum of the tautomer emission band. [f] Ratio of the intensities at the two emission maxima. [g] Fluorescence quantum yield measured using 3-hydroxyflavone in toluene ($\Phi = 0.29$) as a reference.^[32]

Four protic solvents (water, methanol, ethanol, *n*-butanol) and seven aprotic solvents (acetonitrile, dimethyl sulfoxide, acetone, ethyl acetate, THF, dichloromethane, toluene) were selected. According to their H-bond basicity ($\sum \beta_2^H$) values, three

of the aprotic solvents may be classified as neutral (toluene, dichloromethane, acetonitrile), whereas the other aprotic solvents are rather basic.^[15,31]

The compound **TCeU** exhibits a single absorption band centered at 392–401 nm. The positions of the maxima correlate poorly with the solvent polarity showing a little positive solvatochromism. In all solvents except for water, **TCeU** showed dual emission, in which the short-wavelength and long-wavelength bands can be attributed to the N* and T* states, respectively. The N*/T* ratio as well the position of the N* emission band were strongly solvent-dependent (Table 1, Figure 4).

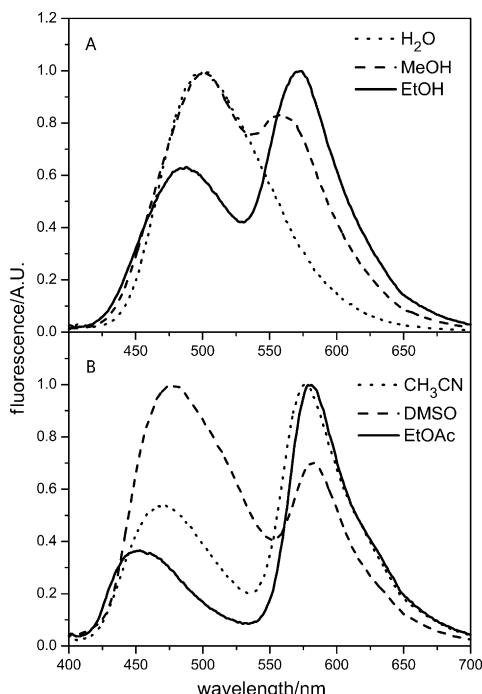


Figure 4. Fluorescence spectra of **TCeU** normalized at the maximal intensity in: A) protic, and B) aprotic solvents.

The N* band demonstrates positive solvatochromism, which is typical for this class of compounds.^[33] The position of the T* band maximum is blueshifted in protic solvents, as compared with aprotic solvents, probably due to the formation of a H bond between the phenoxide oxygen of **TCeU** in the excited tautomer state and the solvent, as it was proposed for 2-furyl-3HC.^[34] The Stokes shifts for the first and second emission bands are comprised between 2387–5356 and 7395–8027 cm⁻¹, respectively. The large Stokes shift obtained for the second emission band is characteristic for ESIPT probes. As a general trend, the N*/T* ratio increases with an increase in solvent polarity. Thus, the highest values of N*/T* ratio are observed for highly polar protic methanol and aprotic DMSO, whereas the lowest values are observed for the most apolar toluene. The fluorescence quantum yields are rather good in all solvents (0.15–0.25), except water (0.08).

Comparison of **TCeU** with model compounds

The absorption and fluorescence properties of **TC** and **TCe** were investigated and compared with those obtained for **TCeU** in the same set of solvents to get further insight regarding the effect of the electronic conjugation in the extended chromophore (Table 2). The absorption spectrum shifts to red

Table 2. Spectroscopic characteristics^[a] of **TC**, **TCe**, and **TCeU** in different solvents.

Compound	Solvent	λ_{abs}	λ_{N^*}	λ_{T^*}	$I_{\text{N}^*}/I_{\text{T}^*}$	Φ
TC	MeOH	359	424	540	0.72	0.07
	EtOH	360	424	542	0.34	0.08
	CH ₃ CN	354	414	540	0.07	0.08
	EtOAc	354	412	543	0.05	0.12
	CH ₂ Cl ₂	357	412	537	0.02	0.30
	PhCH ₃	358	412	546	0.01	0.24
TCe	MeOH	378	442	561	0.43	0.13
	EtOH	379	439	563	0.22	0.14
	CH ₃ CN	373	431	563	0.10	0.14
	EtOAc	374	430	566	0.12	0.15
	CH ₂ Cl ₂	379	432	561	0.05	0.19
	PhCH ₃	379	431	572	0.03	0.20
TCeU	MeOH	395	501	558	1.20	0.23
	EtOH	397	488	574	0.63	0.18
	CH ₃ CN	392	472	576	0.54	0.15
	EtOAc	394	451	579	0.36	0.15
	CH ₂ Cl ₂	397	458	577	0.31	0.17
	PhCH ₃	399	441	587	0.13	0.15

[a] See the footnotes [c]–[g] in Table 1 for the definitions of these parameters.

in the following sequence: **TC**→**TCe**→**TCeU** (Figure 5). Extinction coefficients increase upon extending the π-electronic system of the fluorophore. Values are 23500, 30500, and 35000 M⁻¹cm⁻¹ in methanol for **TC**, **TCe**, and **TCeU**, respectively.

This sequential redshift corresponds well to the gradual increase in the conjugation length of the fluorophores. The data also suggest that the chromone and uridine parts in **TCeU** are fused into a single conjugated electronic system as evidenced by the disappearance of the characteristic absorption band of

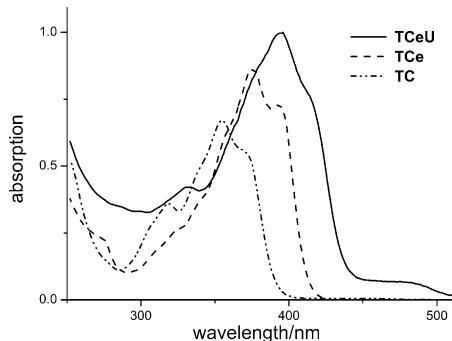


Figure 5. Absorption spectra of compounds **TC**, **TCe** and **TCeU** in ethyl acetate. Absorptions are proportional to their absorption coefficient.

uracil at about 260 nm (Figure 5). Interestingly, our designed fluorophore exhibits suitable absorption properties with an absorption band around 400 nm making possible its selective excitation in the presence of proteins and nucleic acids with conventional commercial lasers.

Similar to **TCeU**, the model compounds **TC** and **TCe** display dual emission, indicating that all three of them undergo ESIPT. The positions of the emission bands shift to the red with the increase in the fluorophore conjugation: **TC**–**TCe**–**TCeU**. For each solvent studied, the value of the N^*/T^* ratio of **TCeU** is much higher than those of **TC** and **TCe** (Table 2). Furthermore, the dual emission of these dyes depends differently on the solvent polarity.

In contrast to **TCeU**, compounds **TC** and **TCe** show only limited positive solvatochromism of the N^* band. Indeed, going from toluene to MeOH, the N^* band shifts to the red by only 12 nm for **TC** and **TCe**, whereas for **TCeU** the redshift reaches 60 nm. The T^* band of **TC** and **TCe** also does not reveal clear variations with solvent polarity. Remarkably, all three compounds showed an increase in the N^*/T^* ratio with increasing

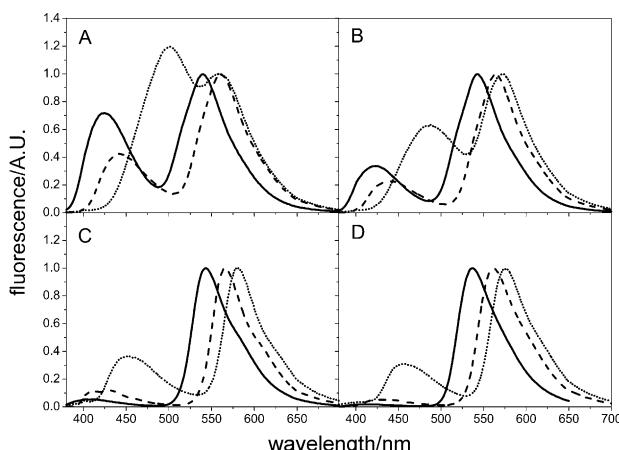


Figure 6. Normalized fluorescence spectra of **TC** (—), **TCE** (---), and **TCeU** (....) in: A) methanol, B) ethanol, C) ethyl acetate, and D) dichloromethane.

solvent polarity (Figure 6). However, the N^* band emission of **TCeU** is significant in aprotic solvents of different polarity, which is not the case of **TC** and **TCE**. Both compounds show only a marginal contribution of their N^* form to the emission spectra in apolar solvents, such as ethyl acetate and dichloromethane (Figure 6C and D).

Both the strong solvatochromism of the N^* band of **TCeU** and its high intensity in aprotic solvents are in line with a charge-transfer character of its N^* state. A very similar charge-transfer character of the N^* state was previously observed for 4'-(dialkylamino)-3-hydroxyflavone^[14a] and its analogues having a push-pull structure.^[35] It was hypothesized that in aprotic solvents, the strong interaction of the highly polar N^* state with the environment may stabilize it dielectrically and thus shifts the ESIPT equilibrium $N^*\rightleftharpoons T^*$ towards to the N^* state.^[36]

To further investigate the contribution of ICT to the photophysical behavior of the synthesized ESIPT dyes, their methylated analogues **MTC**, **MTCe**, and **MTCeU** were characterized in the same set of solvents (Supporting Information, Table S2). **MTC**, **MTCe**, and **MTCeU** showed comparable absorption coefficients to the parent non-methylated dyes and absorption bands centered at 337–341, 362–368, and 375–388 nm, respectively. When compared with the corresponding non-methylated chromones, the methylated analogues exhibit a significant blueshift of their absorption maximum. This significant blue-shift may be explained by a possible distortion of the planar structure of the fluorophores due to an increased steric hindrance between the methoxyl group and the thieryl substituent in positions 3 and 2 of the chromone ring.^[37]

As expected, no tautomer band in the emission spectra was observed due to blocked ESIPT. Notably, no significant blue-shift in the emission maxima was observed, when compared with the N^* band of non-methylated chromones. This indicates that the geometry of the relaxed excited state is close for both methylated and non-methylated compounds. The emission bands of the methylated 3HCs were centered at 426–450 nm for **MTC** and at 432–442 nm for **MTCe**, showing a poor dependence on solvent polarity. At the same time, compound **MTCeU** exhibited a strong solvent dependence of the position of its emission band, which ranges from 451 nm in ethyl acetate to 511 nm in water (Figure 7). Thus, the methylated com-

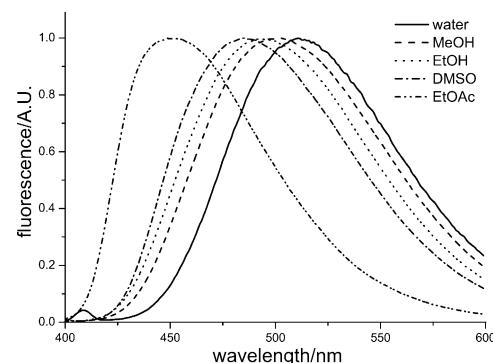


Figure 7. Normalized fluorescence spectra of **MTCeU** in representative solvents of different polarities.

pound **MTCeU**, similarly to its parent 3-hydroxychromone **TCeU**, shows a strong solvatochromism of its emission band, confirming our hypothesis on the charge-transfer character of the N^* state. Our data confirm the presence of an ICT state due to the interaction of the chromone and uracil moieties.

A significant increase in the fluorescence quantum yield was observed for **MTCeU** (10–50%) as compared with **MTC** (0.5–2%) and **MTCe** (2–20%). The high quantum yields of **MTCeU** together with its solvatochromic properties compare favorably with other color-changing nucleoside analogues^[38] and make this fluorophore a prospective candidate for incorporation in DNA and further investigations.

Table 3. Time-resolved decay parameters^[a] for **TC**, **TCE**, and **TCEU** in polar solvents.

Compound	Solvent	$\tau_1^{N^*}$ [ns] ^[a]	$\alpha_1^{N^*}$ ^[b]	$\tau_2^{N^*}$ [ns] ^[a]	$\alpha_2^{N^*}$ ^[b]	$\tau_1^{T^*}$ [ns] ^[a]	$\alpha_1^{T^*}$ ^[b]	$\tau_2^{T^*}$ [ns] ^[a]	$\alpha_2^{T^*}$
 2 (TC)	EtOH	0.10 ± 0.01	1.00	–	–	0.08 ± 0.01	-0.45	1.2 ± 0.1	0.55
	MeOH	0.11 ± 0.01	1.00	–	–	0.10 ± 0.01	-0.47	0.76 ± 0.05	0.53
 3 (TCE)	EtOH	0.11 ± 0.01	1.00	–	–	0.09 ± 0.03	-0.53	2.6 ± 0.9	0.47
	MeOH	0.13 ± 0.01	1.00	–	–	0.13 ± 0.02	-0.15	2.00 ± 0.01	0.85
 1 (TCEU)	EtOAc	0.12 ± 0.01	1.00	–	–	0.10 ± 0.01	-0.47	2.4 ± 0.2	0.53
	EtOH	0.22 ± 0.01	0.54	0.50 ± 0.01	0.46	0.30 ± 0.13	-0.12	1.9 ± 0.1	0.88
	MeOH	0.43 ± 0.01	0.79	0.91 ± 0.02	0.21	0.35 ± 0.08	-0.25	1.55 ± 0.1	0.75

[a] The short- (τ_1) and long- (τ_2) lived decay times for the N* and T* emission bands, respectively. [b] The corresponding pre-exponential coefficients for the N* and T* emission bands, respectively.

Time-resolved studies of ESIPT

Time-resolved investigation of the ESIPT reaction in compounds **TC**, **TCE**, and **TCEU** was performed to understand the dramatic differences in the sensitivity of these compounds to solvent polarity. Time-resolved measurements were conducted in polar solvents in which both the normal and tautomer fluorescence bands were well separated (Table 3). Except for **TCEU** in the two protic solvents, the decay of the N* band of the three compounds was characterized by a single short-lived component ($\tau_1^{N^*}$), whereas the T* band decay was characterized by a fast rise component ($\tau_1^{T^*}$) with a lifetime value close to that of $\tau_1^{N^*}$, followed by a slower decay component ($\tau_2^{T^*}$). A similar behavior was already reported for (2-furyl)-3-hydroxy-chromone^[39] and for 3-hydroxyquinolones,^[40] the aza-analogues of 3HCs, strongly suggesting an irreversible ESIPT mechanism that is kinetically controlled. In this mechanism, the $\tau_1^{N^*}$ and $\tau_1^{T^*}$ values describe the kinetics of the proton transfer, whereas the long-lived decay component of the T* state ($\tau_2^{T^*}$) describes the lifetime of the T* form. For **TCE**, the ESIPT is slightly slower than in **TC**. For both compounds, the short-lived lifetimes are slightly higher in methanol, confirming that the ESIPT process is slower in this more H-bond-donating solvent.^[39]

In the case of **TCEU**, the data indicate an apparent change in the ESIPT mechanism in polar protic solvents. Whereas the N* emission in ethyl acetate is characterized by monoexponential kinetics, the N* emission in EtOH and MeOH shows biexponential decay with commensurable amplitudes. The shorter component in these solvents still matches well with the corresponding rise component of the T* emission band, so it can be assigned to the ESIPT process. Interestingly, the value of this component drastically rises with the increase of the H-bond-donating ability of the solvent, indicating a slowing down of the ESIPT. The slowing down of ESIPT can be assigned to the formation of an intermolecular H bond between the 3-hydroxyl group of the dye with the solvent, which disrupts the intramolecular H bond with the carbonyl oxygen. Remarkably,

the ESIPT process is slower in **TCEU** in comparison with **TC** and **TCE** in the same solvents. Indeed, an increase in the charge-transfer character of the N* state in **TCEU** due to the electron-donating uracil group may further strengthen the intermolecular solvent-solute H bond, thus additionally slowing down the ESIPT reaction.^[39] The longer component in the N* emission, however, is not similar to the long-lived decay component in the T* emission, which one would expect in the case of reversible ESIPT.^[36]

Thus, the proton transfer in **TCEU** in polar protic solvents likely remains irreversible, whereas the longer component of the N* emission decay may be assigned to the fluorescence of a different form, probably the intermolecular H-bonded species that is generally non-emissive in 3HCs, but was shown to emit in 3-hydroxyquinolones^[40] and 4'-dialkylamino-3-hydroxyflavones.^[41] The presence of this new emissive species is consistent with the higher quantum yield measured for **TCEU** in methanol and ethanol as compared with **TC** and **TCE**.

Correlation of the spectroscopic properties of **TCEU** with solvent polarity

In the prospect of using **TCEU** as an environment-sensitive fluorophore for DNA probing, it is important to establish the relationships between the photophysical parameters of the fluorophore (position of absorption/emission bands, and N*/T* ratio) and quantitative parameters of the environment polarity. In this work, we used the Reichardt's polarity scale $E_T(30)$,^[30] an empirical scale commonly used in the analysis of dye solvatochromism. It has also been established as a reliable scale for analyzing emission of 3HCs.^[16,35]

We observe that the positions of the absorption band and the T* emission band of **TCEU** show a poor correlation with $E_T(30)$. In contrast, a strong correlation is observed for the position of the N* band, which undergoes a bathochromic shift with an increase in solvent polarity (Figure 8). Obviously, this solvatochromism is caused by the strong dipole moment of the N* state, which possesses an ICT character. The strong sol-

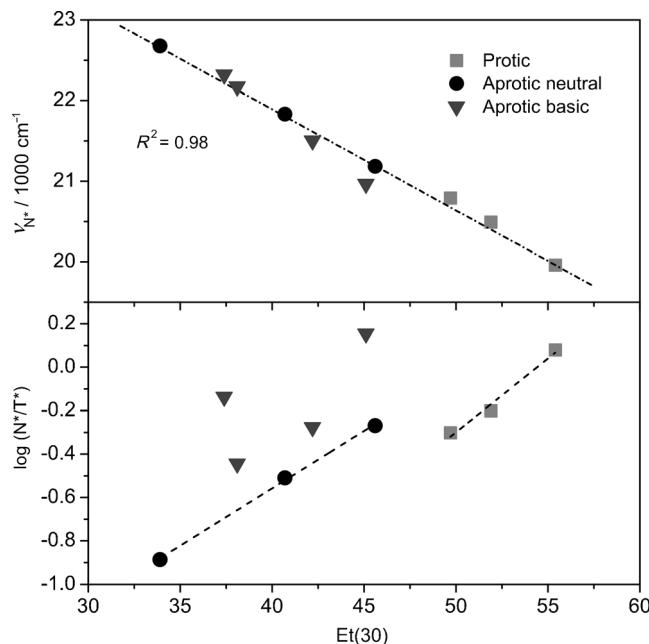


Figure 8. Dependence of the spectroscopic characteristics of **TCeU** on the solvent polarity index $E_T(30)$. Position of the N^* band (top) and logarithm of the N^*/T^* ratio (bottom) as a function of $E_T(30)$. The light-gray squares, black circles, and dark-gray triangles represent protic (MeOH , EtOH , BuOH), aprotic neutral (CH_3CN , CH_2Cl_2 , PhCH_3), and aprotic H-bond basic solvents (DMSO , acetone, EtOAc , THF), respectively.

vent-dependent variation in the N^*/T^* intensity ratio is the distinctive feature of 3HC fluorophores.^[15,35] In the case of **TCeU**, the N^*/T^* ratio increases with the solvent polarity both in protic and neutral aprotic solvents, showing a linear correlation between $\log(I_{N^*}/I_{T^*})$ and $E_T(30)$. Most importantly, the N^*/T^* ratio remains relatively high in aprotic media, so that the **TCeU** fluorophore is capable of distinguishing one apolar aprotic solvent from another; whereas, both the parent **TC** (Table 2, Figure 6) and its structural analogue 2-furyl-3-hydroxychromone^[20] show a uniformly low N^*/T^* ratio in these solvents. By this feature, **TCeU** is comparable to 4'-dialkylamino-3-hydroxyflavones that also possess an ICT nature of their normal excited state. However, the dependence of the N^*/T^* ratio of **TCeU** versus $E_T(30)$ shows some upward deviation for aprotic solvents that present strong H-bond-acceptor ability (basicity). The upward influence of basic solvents on the relative intensity of N^* band was observed even at trace concentration of a basic solvent originating from the stock solution; notably, this was observed only for fully conjugated **TCeU** (Supporting Information, Figure S1).

This deviation indicates that these aprotic solvents can hamper the ESIPT reaction in **TCeU**, probably by disrupting its intramolecular H bonding. This phenomenon is not usually observed for 3HC derivatives,^[15] but it is a common property of 3-hydroxyquinolones.^[40] A similar effect of basicity was also observed for several other 2-thienyl-3-hydroxychromones.^[25] The effect of solvent basicity on **TCeU** emission confirms the fact that unlike other 3HCs, its complex with disrupted H bonding is probably emissive, in line with its relatively high quantum

yields observed in H-bond-acceptor solvents (Table 1). Finally, the good correlation observed with all other solvents (Figure 8), indicates that the N^*/T^* ratio depends both on the polarity, H-bonding acidity and basicity of the solvent.

It should be noted that the strong variation of the dual emission of **TCeU** is observed in a broad range of solvent polarities, ranging from apolar toluene up to polar methanol. This is an important difference with its analogues **TC**, **TCe**, and the previously reported 2-furyl-3-hydroxychromone, which show strong variation of their dual emission mainly in polar protic solvents.^[39] The present new feature, resulting from the charge-transfer character of the N^* excited state, makes this new fluorophore prospective for incorporation into DNA, in which rather low polarities of internal regions can be found as well as the local environment of the major groove is expected to be relatively polar and comparable to methanol and therefore different to bulk water.^[42] The incorporation of **TCeU** into DNA and its further application for monitoring DNA interactions is under way and will be reported in due course.

Conclusion

A new fluorescent dye based on the fusion of 3HC and uracil moieties has been designed, synthesized, and investigated. The dye undergoes an ESIPT reaction and exhibits a dual-band emission in a set of solvents of varying polarity. The ratio of its two emission bands changes as a function of the solvent polarity across a broad range. The position of the N^* band correlates perfectly with the empirical solvent polarity scale $E_T(30)$. A strong correlation was found also between $\log(N^*/T^*)$ and $E_T(30)$ for protic and neutral aprotic solvents. Comparison with its structural analogues allowed for the demonstration of the electronic conjugation between 3HC and uracil heterocycles and the charge-transfer nature of its normal excited state that is responsible for the dependence of its N^*/T^* ratio on the solvent polarity. The fluorophore is a prospective candidate for incorporation into DNA as an environment-sensitive nucleobase mimic.

Experimental Section

General methods

All non-aqueous reactions involving water-sensitive reagents were performed in oven-dried glassware under argon using dry solvents. The synthetic intermediates were co-evaporated twice with toluene and dried in vacuo before use. All chemical reagents were obtained from commercial sources and were used as supplied. Anhydrous solvents were obtained according to standard procedures.^[43] The reactions were monitored by thin-layer chromatography (TLC, Merck silica gel 60 F254 plates) and visualized both by UV radiation (254 and 365 nm) and by spraying with phosphomolybdic acid in ethanol followed by a subsequent warming with a heat gun. Flash column chromatography^[44] was performed by using silica gel (40–63 µm). All NMR spectra (^1H , ^{13}C , ^2D) were recorded on 200 or 500 Bruker Advance Spectrometers (200 or 500 MHz). ^1H NMR (200 and 500 MHz), ^{13}C NMR (50 and 125 MHz, recorded with complete proton decoupling) spectra were obtained with samples dissolved in CDCl_3 , CD_3OD , or $[\text{D}_6]\text{DMSO}$, with the solvent signal used as an

internal reference: $\delta = 7.26$ for CHCl_3 , 3.31 for CD_2HOD , and 2.50 ppm for $(\text{CD}_3)(\text{CD}_2\text{H})\text{S}(\text{O})$ for ^1H NMR experiments, and $\delta = 77.0$ for CDCl_3 , 49.0 for CD_3OD , and 39.4 ppm for $(\text{CD}_3)_2\text{S}(\text{O})$ for ^{13}C NMR experiments.^[45] Chemical shifts (δ) are given in ppm to the nearest 0.01 (^1H) or 0.1 ppm (^{13}C). The coupling constants (J) are given in Hertz (Hz). The signals are reported as follows: (s=singlet, d=dou-blet, t=triplet, m=multiplet, br=broad). Assignments of ^1H and ^{13}C NMR signals were achieved with the help of D/H exchange, COSY, HMQC, HMBC experiments. Regular mass spectra (MS) were recorded on an Esquire 3000 Plus apparatus with ESI in both positive and negative mode. High-resolution mass spectrometry was conducted with a FINIGAN MAT 95 spectrometer with EI or ESI ionization techniques. Systematic flavone and nucleobase nomenclature is used below for the assignments of each spectrum. All solvents for absorption and fluorescence experiments were of spectroscopic grade. Absorption spectra were recorded on a Cary 4 spectrophotometer (Varian) using 1 cm quartz cells. Stock solutions were prepared using dimethyl sulfoxide, except for **TCeU** for which dioxane was used. The samples used for spectroscopic measurements contained approximately 0.1% v/v of the stock solvent. Fluorescence spectra were recorded on a FluoroLog 3.5 spectrofluorometer (Jobin Yvon, Horiba). Excitation wavelength was 360 nm. Time-resolved fluorescence measurements were performed with the time-correlated, single-photon counting technique using the excitation pulses at 315 nm provided by a pulse-picked frequency-tripled Ti:sapphire laser (Tsunami, Spectra Physics) pumped by a Millenia X laser (Spectra Physics). The fluorescence decays were collected at the magic angle (54.7°) of the emission polarizer to avoid any artifact due to vertically polarized excitation beam. The single-photon events were detected with a microchannel plate photomultiplier (Hamamatsu) coupled to a pulse preamplifier HFAC (Becker-Hickl) and recorded on a SPC-130 board (Becker-Hickl). Time-resolved data were analyzed by the maximum entropy method (Pulse 5 software).^[46] In all cases, the χ^2 values were close to 1 and the weighted residuals as well as their auto-correlation were distributed randomly around 0, indicating an optimal fit.

Synthesis

3-Hydroxy-2-(thiophen-2-yl)-chromen-4-one (2): A 5 N NaOH solution (6 mL) was added drop-wise to a stirred solution of o-hydroxyacetophenone **7** (1.16 mL, 9.63 mmol) and 2-thiophenecarbaldehyde **8** (900 μL , 9.63 mmol) in ethanol (50 mL). The reaction mixture was stirred 48 h at RT before the drop-wise addition of 30% aq. hydrogen peroxide solution (2 mL). The resulting mixture was stirred 7 h at RT, then poured into cold water (300 mL) and neutralized with diluted acetic acid. The resulting precipitate was filtered and thoroughly washed with water and hexane to provide the target derivative **2** (1.17 g, 50%; $\text{C}_{13}\text{H}_8\text{O}_3\text{S}$ (244.27)) as a yellow solid. ^1H NMR (200 MHz, $[\text{D}_6]\text{DMSO}$): $\delta = 7.30$ (dd, $^3J = 5.0$, $^3J = 3.8$ Hz, 1H, $\text{H}\beta\text{-thioph.}$), 7.46 (td, $^3J = 8.0$, $^4J = 1.2$ Hz, 1H, H6), 7.72 (dd, $^3J = 8.4$, $^4J = 1.2$ Hz, 1H, H8), 7.73–7.84 (m, 1H, H7), 7.91 (dd, $^3J = 5.0$, $^4J = 1.2$ Hz, 1H, $\text{H}\gamma\text{-thioph.}$), 7.97 (dd, $^3J = 3.8$, $^4J = 1.2$ Hz, 1H, $\text{H}\alpha\text{-thioph.}$), 8.10 (dd, $J = 8.0$, 1.2 Hz, 1H, H5), 10.28 ppm (s, 1H, OH); ^{13}C NMR (50 MHz, $[\text{D}_6]\text{DMSO}$): $\delta = 118.0$ (C8), 121.7 (C10), 124.5 (C6), 124.7 (C5), 127.7 (C β), 128.3 (C α), 130.9 (C γ), 132.3 (C2-Th), 133.6 (C7), 136.6 (C3), 143.1 (C2), 154.1 (C9), 171.9 ppm (C4); MS (ESI $^+$): m/z : 267.0 [$M+\text{Na}$] $^+$.

2-(5-Bromothiophen-2-yl)-3-hydroxy-chromen-4-one (10): A 5 N NaOH solution (6 mL) was added drop-wise to a stirred solution of o-hydroxyacetophenone **7** (1.16 mL, 9.63 mmol) and 5-bromo-2-thiophenecarbaldehyde **9** (1.14 mL, 9.63 mmol) in ethanol (50 mL) was. The reaction mixture was stirred 48 h at RT before a drop-

wise addition of 30% aq. hydrogen peroxide solution (2 mL). The resulting mixture was stirred 7 h at RT, then poured into cold water (300 mL) and neutralized with diluted acetic acid. The resulting precipitate was filtered and thoroughly washed with water and hexane to provide as a yellow solid the target derivative **10** (1.49 g, 48%; $\text{C}_{13}\text{H}_8\text{BrO}_3\text{S}$ (323.16)). ^1H NMR (200 MHz, CDCl_3): $\delta = 7.19$ (d, $^3J = 4.0$ Hz, 1H, $\text{H}\beta\text{-thioph.}$), 7.41 (td, $^3J = 8.0$, $^4J = 0.8$ Hz, 1H, H6), 7.54 (d, $^3J = 8.4$ Hz, 1H, H8), 7.71 (ddd, $^3J = 8.4$, $^3J = 8.0$, $^4J = 1.4$ Hz, 1H, H7), 7.73 (d, $^3J = 4.0$ Hz, 1H, $\text{H}\alpha\text{-thioph.}$), 8.22 ppm (dd, $^3J = 8.0$, $^4J = 1.4$ Hz, 1H, H5); ^{13}C NMR (50 MHz, $[\text{D}_6]\text{DMSO}$): $\delta = 112.2$ (Th-Br), 117.8 (C8), 121.6 (C α), 122.0 (C10), 122.9 (C6), 124.7 (C5), 129.3 (C β), 131.6 (C7), 136.1 (C2-Th), 141.6 (C3), 146.5 (C2), 153.3 (C9), 176.0 ppm (C4); MS (ESI $^+$): m/z : 344.9 [$M+\text{Na}$] $^+$.

3-Methoxy-2-(thiophen-2-yl)-chromen-4-one (4): A solution of 4.5 M KOH (2 mL), 18-crown-6 (5 mol%, 25 mg), and dimethyl sulfate (370 μL , 3.86 mmol) were added to a stirred suspension of **2** (472 mg, 1.93 mmol) in CH_2Cl_2 (8 mL). The reaction mixture became homogenous in 5 min and was stirred at RT for 1 h. After quenching by addition of H_2O (10 mL), the organic layer was extracted with CH_2Cl_2 (3 \times), dried over MgSO_4 , filtered, and the volatiles were removed in vacuo. The residue was purified by flash chromatography on silica gel eluted with cyclohexane/EtOAc mixture (10:1→4:1, v/v) to provide the desired compound **4** as a yellow solid (483 mg, 97%; $\text{C}_{14}\text{H}_{10}\text{O}_3\text{S}$ (258.29)). ^1H NMR (200 MHz, CDCl_3): $\delta = 4.06$ (s, 3H, CH_3), 7.19 (dd, $^3J = 4.0$, 5.0 Hz, 1H, $\text{H}\beta\text{-thioph.}$), 7.36 (td, $^3J = 8.0$, $^4J = 1.0$ Hz, H6), 7.50 (dd, $^3J = 8.4$, $^4J = 1.0$ Hz, 1H, H8), 7.60 (dd, $^3J = 5.0$, $^4J = 1.2$ Hz, 1H, $\text{H}\gamma\text{-thioph.}$), 7.65 (td, $^3J = 8.4$, $^4J = 1.6$ Hz, 1H, H7), 7.94 (dd, $^3J = 4.0$, $^4J = 1.2$ Hz, 1H, $\text{H}\alpha\text{-thioph.}$), 8.22 ppm (dd, $^3J = 8.0$, $^4J = 1.6$ Hz, 1H, H5); ^{13}C NMR (50 MHz, CDCl_3): $\delta = 59.6$ (OMe), 117.7 (C8), 124.2 (C10), 124.6 (C6), 125.6 (C5), 127.4 (C β), 129.5 (C α), 131.5 (C γ), 131.7 (C2-Th), 133.3 (C7), 138.6 (C3), 151.5 (C2), 154.7 (C9), 174.0 ppm (C4); HRMS (ESI $^+$): m/z calcd for $\text{C}_{14}\text{H}_{10}\text{O}_3\text{S}$: 259.0423 [$M+\text{H}$] $^+$; found 259.0424.

3-Methoxy-2-(5-bromothiophen-2-yl)-chromen-4-one (11): A solution of 4.5 M aq. KOH (5 mL), 18-crown-6 (5 mol%, 65 mg), and dimethyl sulfate (930 μL , 9.84 mmol) were added to a stirred suspension of **10** (1.59 g, 4.92 mmol) in CH_2Cl_2 (35 mL). The reaction mixture became homogenous in 5 min and was stirred at RT for 1 h. After quenching by addition of H_2O (40 mL), the organic layer was extracted with CH_2Cl_2 (3 \times), dried over MgSO_4 , filtered and the volatiles were removed in vacuo. The residue was purified by flash chromatography on silica gel eluted with cyclohexane/EtOAc mixture (10:1→4:1, v/v) to provide the desired compound **11** as a yellow solid (1.30 g, 78%; $\text{C}_{14}\text{H}_9\text{BrO}_3\text{S}$ (337.19)). ^1H NMR (200 MHz, CDCl_3): $\delta = 4.07$ (s, 3H, CH_3), 7.16 (d, $^3J = 3.8$ Hz, 1H, $\text{H}\beta\text{-thioph.}$), 7.38 (td, $^3J = 8.0$ Hz, $^4J = 1.0$ Hz, H6), 7.49 (dd, $^3J = 8.4$, $^4J = 1.0$ Hz, 1H, H8), 7.62–7.79 (m, 1H, H7), 7.66 (d, $^3J = 3.8$ Hz, 1H, $\text{H}\alpha\text{-thioph.}$), 8.22 ppm (dd, $^3J = 8.0$, $^4J = 1.6$ Hz, 1H, H5); ^{13}C NMR (50 MHz, CDCl_3): $\delta = 59.7$ (OMe), 117.7 (C8), 119.8 (Th-Br), 124.2 (C10), 124.7 (C6), 125.7 (C5), 129.2 (C α), 130.2 (C β), 132.7 (C2-Th), 133.5 (C7), 138.4 (C3), 150.2 (C2), 154.7 (C9), 173.8 ppm (C4); HRMS (ESI $^+$): m/z calcd for $\text{C}_{14}\text{H}_9\text{BrO}_3\text{S}$: 336.9529 [$M+\text{H}$] $^+$; found 336.9531.

3-Benzylloxycarbonyloxy-2-(5-bromothiophen-2-yl)-chromen-4-one (12): A solution of 4.5 M aq. KOH (9 mL), 18-crown-6 (5 mol%, 43 mg), and benzyl chloroformate (700 μL , 4.88 mmol) were added to a stirred suspension of **10** (1.05 g, 3.25 mmol) in CH_2Cl_2 (13 mL). The reaction mixture became homogenous in 5 min and was stirred at RT for 1 h. After quenching by addition of H_2O (10 mL), the organic layer was extracted with CH_2Cl_2 (3 \times), dried over MgSO_4 , filtered, and the volatiles were removed in vacuo. The residue was purified by flash chromatography on silica gel eluted with cyclohexane/EtOAc mixture (9:1→7:3, v/v) to provide the desired compound **12** as white crystals (1.20 g, 81%; $\text{C}_{21}\text{H}_{13}\text{BrO}_5\text{S}$ (457.29)).

$R_f = 0.53$ (cyclohexane/EtOAc = 7:3); ^1H NMR (200 MHz, CDCl_3): $\delta = 5.37$ (s, 2 H, CH_2), 7.16 (d, $^3J = 4.0$ Hz, 1 H, $\text{H}\beta$ -thioph.), 7.34–7.50 (m, 6 H, H_6 and Cbz), 7.52 (dd, $^3J = 8.4$, $^4J = 0.6$ Hz, 1 H, H8), 7.65 (d, $^3J = 4.0$ Hz, 1 H, $\text{H}\alpha$ -thioph.), 7.72 (td, $^3J = 8.4$, $^4J = 1.6$ Hz, 1 H, H7), 8.25 ppm (dd, $^3J = 8.0$, $^4J = 1.6$ Hz, 1 H, H5); ^{13}C NMR (50 MHz, CDCl_3): $\delta = 71.3$ ($\text{OC(O)CH}_2\text{Ph}$), 117.8 (C8), 120.4 (Th-Br), 123.6 (C10), 125.4 (C6), 126.1 (C5), 128.4 (o-C_{Ph}), 128.7 ($m\text{-C}_{\text{Ph}}$), 128.8 ($p\text{-C}_{\text{Ph}}$), 131.0 ($\text{C}\beta$), 131.0 (Ca), 131.5 (C2-Th), 132.0 (C3), 134.1 (C7), 134.5 ($i\text{-C}_{\text{Ph}}$), 150.3 (C2), 151.8 ($\text{OC(O)CH}_2\text{Ph}$), 155.0 (C9), 171.3 ppm (C4); HRMS (ESI $^+$): m/z calcd for $\text{C}_{21}\text{H}_{14}\text{BrO}_5\text{S}$: 456.9740 [$M+\text{H}]^+$; found: 456.9732.

3-Hydroxy-2-(5-((trimethylsilyl)ethynyl)thiophen-2-yl)-chromen-4-one (3): $[\text{PdCl}_2(\text{PPh}_3)_2]$ (3 mol %, 20 mg) and Cul (3 mol %, 5 mg) were added to a stirred solution of aryl bromide **10** (323 mg, 1 mmol), trimethylsilylacetylene (210 μL , 1.5 mmol) and *N,N*-diisopropylethylamine (DIPEA; 280 μL , 2 mmol) in THF (6 mL) under argon. The mixture was heated at reflux for 1 h, then allowed to cool down to RT. The volatiles were removed in vacuo and the residue was purified by flash chromatography on silica gel eluted with $\text{CH}_2\text{Cl}_2/\text{MeOH}$ (99:1 → 97:3, v/v) to provide the coupled derivative **3** as a yellow solid (142 mg, 41%; $\text{C}_{18}\text{H}_{16}\text{O}_3\text{SSi}$ (340.47)). ^1H NMR (500 MHz, $[\text{D}_6]\text{DMSO}$): $\delta = 0.27$ (s, 9 H, SiMe₃), 7.47 (dd, $^3J = 8.0$ Hz, $^3J = 7.5$ Hz, H6), 7.50 (d, $^3J = 4.0$ Hz, 1 H, $\text{H}\beta$ -thioph.), 7.72 (d, $^3J = 8.5$ Hz, 1 H, H8), 7.81 (ddd, $^3J = 8.5$, $^3J = 7.5$, $^4J = 1.5$ Hz, 1 H, H7), 7.85 (d, $^3J = 4.0$ Hz, 1 H, $\text{H}\alpha$ -thioph.), 8.10 (dd, $^3J = 8.0$, $^4J = 1.5$ Hz, 1 H, H5); 10.59 ppm (brs, 1 H, OH); ^{13}C NMR (125 MHz, $[\text{D}_6]\text{DMSO}$): $\delta = -0.4$ (SiMe₃), 97.2 (-C≡C-TMS), 101.9 (-C≡C-TMS), 118.0 (C8), 121.7 (C10), 124.5 (C6), 124.8 (C5), 125.3 (C2-Th), 127.8 (Ca), 133.2 (Cβ), 133.5 ($\text{Th-C}\equiv\text{C}$), 133.7 (C7), 137.5 (C3), 141.9 (C2), 154.1 (C9), 172.0 ppm (C4); HRMS (ESI $^+$): m/z calcd for $\text{C}_{18}\text{H}_{17}\text{O}_3\text{SSi}$: 341.0662 [$M+\text{H}]^+$; found: 341.0657.

3-Methoxy-2-(5-((trimethylsilyl)ethynyl)thiophen-2-yl)-chromen-4-one (5): $[\text{PdCl}_2(\text{PPh}_3)_2]$ (4 mol %, 45 mg) and Cul (4 mol %, 12 mg) were added to a stirred solution of aryl bromide **11** (536 mg, 1.6 mmol), trimethylsilylacetylene (340 μL , 2.4 mmol), and DIPEA (430 μL , 3.2 mmol) in THF (10 mL) under argon. The mixture was heated at reflux for 2 h, then allowed to cool down to RT. The volatiles were removed in vacuo and the residue was purified by flash chromatography on silica gel eluted with cyclohexane/EtOAc mixture (9:1 → 5:1, v/v) to provide the coupled derivative **5** as yellow solid (554 mg, 98%; $\text{C}_{19}\text{H}_{18}\text{O}_3\text{SSi}$ (354.49)). ^1H NMR (500 MHz, CDCl_3): $\delta = 0.28$ (s, 9 H, SiMe₃), 4.09 (s, 3 H, CH₃), 7.29 (d, $^3J = 4.0$ Hz, 1 H, $\text{H}\beta$ -thioph.), 7.39 (dd, $^3J = 8.0$, $^3J = 7.0$ Hz, H6), 7.50 (d, $^3J = 8.0$ Hz, 1 H, H8), 7.67 (ddd, $^3J = 8.0$, $^3J = 7.0$, $^4J = 1.2$ Hz, 1 H, H7), 7.79 (d, $^3J = 4.0$ Hz, 1 H, $\text{H}\alpha$ -thioph.), 8.24 ppm (dd, $^3J = 8.0$, $^4J = 1.2$ Hz, 1 H, H5); ^{13}C NMR (125 MHz, CDCl_3): $\delta = -0.2$ (SiMe₃), 59.8 (OMe), 96.9 (-C≡C-TMS), 102.7 (-C≡C-TMS), 117.7 (C8), 124.3 (C10), 124.8 (C6), 125.8 (C5), 128.9 (C2-Th), 129.0 (Ca), 132.3 ($\text{Th-C}\equiv\text{C}$), 132.6 (Cβ), 133.5 (C7), 139.0 (C3), 150.5 (C2), 154.8 (C9), 173.9 ppm (C4); HRMS (ESI $^+$): m/z calcd for $\text{C}_{19}\text{H}_{19}\text{O}_3\text{SSi}$: 355.0819 [$M+\text{H}]^+$; found: 355.0815.

5-Iodo-1,3-*N,N*-dimethyluracil (14): Dimethyl sulfate (870 μL , 9.2 mmol) was added drop-wise to a stirring slurry of 5-iodouracil **13** (1.00 g, 4.2 mmol) and NaOH (497 mg, 10.5 mmol) in H_2O (7 mL) at 0 °C. The mixture was heated under reflux for 2 h, cooled to RT, extracted with CH_2Cl_2 (3 ×), dried over MgSO_4 , filtered, and reduced in vacuo. The resulting insoluble solid after the evaporation was washed several times with cyclohexane and subsequently with toluene, ether, EtOAc, and cold CH_2Cl_2 . For each solvent, the solid particles were allowed to settle and the supernatant washings were taken out carefully with a syringe with a fine needle. The washings were collected and evaporated and the residual solid was washed as before. All the residues were combined and dried to obtain **14**

(763 mg, 68%; $\text{C}_6\text{H}_7\text{IN}_2\text{O}_2$ (266.04)) as an amorphous white solid. ^1H NMR (200 MHz, CDCl_3): $\delta = 3.41$ (s, 3 H, N^3CH_3), 3.42 (s, 3 H, N^1CH_3), 7.64 ppm (s, 1 H, H6); ^{13}C NMR (50 MHz, CDCl_3): $\delta = 29.4$ (N^3CH_3), 37.2 (N^1CH_3), 67.1 (C5), 147.3 (C6), 151.4 (C2), 160.4 ppm (C4); MS (ESI $^+$): m/z 288.9 [$M+\text{Na}]^+$.

5-(Trimethylsilyl)ethynyl-1,3-*N,N*-dimethyluracil (15): $[\text{PdCl}_2(\text{PPh}_3)_2]$ (6 mol %, 89 mg) and Cul (6 mol %, 24 mg) were added to a stirred solution of alkenyl iodide **14** (544 mg, 2.0 mmol), trimethylsilylacetylene (441 μL , 3.0 mmol), and DIPEA (1.80 mL, 10.0 mmol) in a 2:1 THF/DMF mixture (10 mL) under argon. The reaction mixture was warmed to 55 °C for 4 h. Then, the volatiles were removed in vacuo and the residue was purified by flash chromatography on silica gel eluted with toluene/EtOAc (10:1 → 4:1, v/v) to provide the TMS-alkyne derivative **15** (390 mg, 82%; $\text{C}_{11}\text{H}_{16}\text{N}_2\text{O}_2\text{Si}$ (236.34)) as a white foam. ^1H NMR (200 MHz, CDCl_3): $\delta = 0.17$ (s, 9 H, SiMe₃), 3.30 (s, 3 H, N^3CH_3), 3.38 (s, 3 H, N^1CH_3), 7.49 ppm (s, 1 H, H6); ^{13}C NMR (50 MHz, CDCl_3): $\delta = -0.3$ (SiMe₃), 28.1 (N^3CH_3), 37.2 (N^1CH_3), 95.6 (-C≡C-TMS), 98.7 (-C≡C-TMS), 98.7 (C5), 146.2 (C6), 150.7 (C2), 161.5 ppm (C4); MS (ESI $^+$): m/z : 259.1 [$M+\text{Na}]^+$.

5-Ethynyl-1,3-*N,N*-dimethyluracil (16): TBAF (1 M in THF, 931 μL , 0.93 mmol) was added drop-wise to a stirred solution of **15** (200 mg, 0.85 mmol) in THF (150 mL). The reaction mixture was stirred at RT for 1 h. Then, the volatiles were removed in vacuo and the residue was purified by flash chromatography on silica gel eluted with toluene/EtOAc (10:1 → 3:1, v/v) to provide the terminal alkyne **16** (126 mg, 91%; $\text{C}_8\text{H}_8\text{N}_2\text{O}_2$ (164.16)) as a white solid. ^1H NMR (500 MHz, CDCl_3): $\delta = 3.17$ (s, 1 H, -C≡CH), 3.37 (s, 3 H, N^3CH_3), 3.43 (s, 3 H, N^1CH_3), 7.50 ppm (s, 1 H, H6); ^{13}C NMR (125 MHz, CDCl_3): $\delta = 28.4$ (N^3CH_3), 37.4 (N^1CH_3), 74.9 (-C≡CH), 81.6 (-C≡CH), 98.0 (C5), 146.4 (C6), 150.9 (C2), 161.8 ppm (C4); MS (ESI $^+$): m/z : 187.0 [$M+\text{Na}]^+$.

5-[5-(3-Benzylloxycarbonyloxy-4-oxo-chromen-2-yl)-thiophen-2-yl-ethynyl]-1,3-*N,N*-dimethyluracil (17): $[\text{PdCl}_2(\text{PPh}_3)_2]$ (5 mol %, 19 mg) and Cul (5 mol %, 5 mg) were added to a stirred solution of aryl bromide **12** (236 mg, 0.50 mmol), **16** (92 mg, 0.55 mmol), and DIPEA (450 μL , 2.51 mmol) in THF (5 mL) under argon. The mixture was heated at reflux for 5 h, then allowed to cool down to RT. The volatiles were removed in vacuo and the residue was purified by flash chromatography on silica gel eluted with $\text{CH}_2\text{Cl}_2/\text{Et}_2\text{O}$ (9:1 → 6:1, v/v) to provide the coupled derivative **17** as an orange solid (160 mg, 59%; $\text{C}_{29}\text{H}_{20}\text{N}_2\text{O}_5\text{S}$ (540.5)). ^1H NMR (500 MHz, CDCl_3): $\delta = 3.39$ (s, 3 H, N^3CH_3), 3.47 (s, 3 H, N^1CH_3), 5.38 (s, 2 H, $\text{OC(O)OCH}_2\text{Ph}$), 7.27 (d, $^3J = 4.0$ Hz, 1 H, $\text{H}\beta$ -thioph.), 7.37 (d, $^3J = 7.0$ Hz, 1 H, $p\text{-H}_{\text{Ph}}$), 7.41 (t, $^3J = 7.0$ Hz, 2 H, $m\text{-H}_{\text{Ph}}$), 7.42 (m, 1 H, H6'), 7.48 (s, $^3J = 7.0$ Hz, 2 H, $o\text{-H}_{\text{Ph}}$), 7.53 (d, $^3J = 8.0$ Hz, 1 H, H8'); 7.57 (s, 1 H, H6), 7.71 (ddd, $^3J = 8.0$ Hz, $^3J = 7.0$ Hz, $^4J = 1.8$ Hz, 1 H, H7'); 7.76 (d, $^3J = 4.0$ Hz, 1 H, $\text{H}\alpha$ -thioph.), 8.23 ppm (dd, $^3J = 7.8$, $^4J = 1.8$ Hz, H5'); ^{13}C NMR (125 MHz, CDCl_3): $\delta = 28.4$ (N^3CH_3), 37.5 (N^1CH_3), 71.3 ($\text{OC(O)CH}_2\text{Ph}$), 85.6 ($\text{Th-C}\equiv\text{C-}$), 88.5 ($\text{Th-C}\equiv\text{C-}$), 98.3 (C5), 117.9 (C8'), 123.6 (C10'), 125.3 (C6'), 126.0 (C5'), 128.3 (o-C_{Ph}), 128.7 (m-C_{Ph}), 128.7 (p-C_{Ph}), 129.2 (C2'-Th-), 130.7 (Ca), 131.6 ($\text{Th-C}\equiv\text{C-}$), 131.9 (C3'), 132.8 (Cβ), 134.1 (C7'), 134.5 ($i\text{-C}_{\text{Ph}}$), 145.9 (C6), 150.6 (C2'), 150.7 (C2), 151.8 ($\text{OC(O)CH}_2\text{Ph}$), 155.0 (C9'), 161.1 (C4), 171.3 ppm (C4'); HRMS (ESI $^+$): m/z calcd for $\text{C}_{29}\text{H}_{20}\text{N}_2\text{O}_5\text{S}$: 563.0883 [$M+\text{Na}]^+$; found: 563.0872.

5-[5-(3-Hydroxy-4-oxo-chromen-2-yl)-thiophen-2-yl-ethynyl]-1,3-*N,N*-dimethyluracil (1): A solution of 33% aq. NH₄OH (870 μL , 22.2 mmol) was added drop-wise to a stirred solution of **17** (30 mg, 55.5 μmol) in CH_2Cl_2 (7 mL). The reaction mixture was stirred at room temperature for 1 h. The resulting insoluble solid after the evaporation was washed several times with cyclohexane and subsequently with toluene, ether, EtOAc, and CH_2Cl_2 . For each solvent, the solution was centrifuged for settling solid particles and

the supernatant washings were carefully taken out with a syringe with a fine needle. The washings were collected and evaporated and the residual solid was washed as before. All the residues were combined and dried to obtain **1** (18 mg, 81%; $C_{21}H_{14}N_2O_5S$ (406.4)) as a bright-yellow solid. 1H NMR (500 MHz, $CDCl_3$): δ = 3.42 (s, 3 H, N^3CH_3), 3.48 (s, 3 H, N^1CH_3), 6.94 (brs, 1 H, OH), 7.36 (d, 3J = 4.0 Hz, 1 H, $H\beta$ -thioph.), 7.42 (ddd, 3J = 7.0, 3J = 8.0, 4J = 1.5 Hz, 1 H, H'), 7.56 (d, 3J = 8.0 Hz, 1 H, $H8'$), 7.57 (s, 1 H, $H6$), 7.71 (ddd, 3J = 7.0, 3J = 8.0, 4J = 1.5 Hz, 1 H, $H7'$), 7.88 (d, 3J = 4.0 Hz, 1 H, $H\alpha$ -thioph.), 8.23 ppm (dd, 3J = 8.0 Hz, 4J = 1.5 Hz, 1 H, $H5'$); ^{13}C NMR (125 MHz, $CDCl_3$): δ = 28.5 (N^3CH_3), 37.5 (N^1CH_3), 86.3 (Th-C≡C-), 87.5 (Th-C≡C), 98.8 (C5), 118.1 (C8'), 121.0 (C10'), 124.7 (C6'), 125.5 (C5'), 126.8 (C2'-Th-), 129.1 (Ca), 133.1 (C7'), 133.8 (C β), 134.0 (C3'), 136.7 (Th-C≡C-), 141.6 (C2'), 145.5 (C6), 150.8 (C2), 155.1 (C9'), 161.3 (C4), 172.4 ppm (C4'); HRMS (ESI $^+$): m/z calcd for $C_{21}H_{15}N_2O_5S$: 407.0696 [M+H] $^+$; found: 407.0683.

5-[5-(3-Methoxy-4-oxo-chromen-2-yl)-thiophen-2-yl-ethynyl]-1-*N,N*-dimethyluracil (6**):** $[PdCl_2(PPh_3)_2]$ (6 mol%, 16 mg) and Cul (6 mol%, 4 mg) were added to a stirred solution of aryl bromide **11** (125 mg, 0.35 mmol), **16** (65 mg, 0.39 mmol), and DIPEA (415 μ L, 1.76 mmol) in THF (4 mL) under argon. The mixture was heated at reflux for 5 h, then allowed to cool down to RT. The resulting insoluble solid after the evaporation was washed several times with cyclohexane and subsequently with toluene, ether, EtOAc, and CH_2Cl_2 . For each solvent, the solution was centrifuged for settling solid particles and the supernatant washings were carefully taken out with a syringe with a fine needle. The washings were collected and evaporated and the residual solid was washed as before. All the residues were combined and dried to obtain **6** (95 mg, 64%; $C_{22}H_{16}N_2O_5S$ (420.4)) as a bright-yellow solid. 1H NMR (500 MHz, $[D_6]DMSO$): δ = 3.21 (s, 3 H, N^3CH_3), 3.36 (s, 3 H, N^1CH_3), 3.99 (s, 3 H, OCH_3), 7.50 (d, 3J = 4.0 Hz, 1 H, $H\beta$ -thioph.), 7.50 (ddd, 3J = 7.5, 3J = 8.0, 4J = 1.0 Hz, 1 H, $H6'$), 7.77 (d, 3J = 7.5 Hz, 1 H, $H8'$), 7.84 (td, 3J = 7.5, 4J = 1.5 Hz, 1 H, $H7'$), 7.98 (d, 3J = 4.0 Hz, 1 H, $H\alpha$ -thioph.), 8.09 (dd, 3J = 8.0, 4J = 1.5 Hz, 1 H, $H5'$), 8.36 ppm (s, 1 H, $H6$); ^{13}C NMR (125 MHz, $[D_6]DMSO$): δ = 27.8 (N^3CH_3), 36.8 (N^1CH_3), 59.4 (OCH_3), 84.1 (Th-C≡C-), 90.0 (Th-C≡C-), 95.2 (C5), 118.2 (C8'), 123.6 (C10'), 124.8 (C6'), 125.1 (C5'), 130.0 (Ca), 131.3 (C2'-Th-), 132.1 (C7'), 134.1 (C β), 138.2 (C3'), 148.8 (C6), 149.5 (Th-C≡C-), 150.4 (C2'), 154.3 (C9'), 160.9 (C2), 172.7 (C4), 186.8 ppm (C4'); HRMS (ESI $^+$): m/z calcd for $C_{22}H_{16}N_2NaO_5S$: 443.0672 [M+Na] $^+$; found: 443.0670.

Acknowledgements

We thank the ANR (ANR-12-BS08-0003-02) and the FRM (DCM20111223038) for financial support and a grant for N.B. We thank the French Embassy in Ukraine for the Master 2 grant to I.A.K. We are grateful to Moses Moustakim for critical reading of the manuscript.

Keywords: charge transfer · fluorescent probes · nucleobases · solvatochromism · time-resolved spectroscopy

- [1] a) D. P. Millar, *Curr. Opin. Struct. Biol.* **1996**, *6*, 322–326; for a review, see: b) R. W. Sinkeldam, N. J. Greco, Y. Tor, *Chem. Rev.* **2010**, *110*, 2579–2619; c) S. G. Srivatsan, A. A. Sawant, *Pure Appl. Chem.* **2011**, *83*, 213–232; d) L. M. Wilhelmsson, *Q. Rev. Biophys.* **2010**, *43*, 159–183.
- [2] a) M. J. Davies, A. Shah, I. J. Bruce, *Chem. Soc. Rev.* **2000**, *29*, 97–107; b) M. J. Rist, J. P. Marino, *Curr. Org. Chem.* **2002**, *6*, 775–793; c) J. N. Wilson, E. T. Kool, *Org. Biomol. Chem.* **2006**, *4*, 4265–4274.

- [3] a) N. B. Gaied, N. Glasser, N. Ramalanjaona, H. Beltz, P. Wolff, R. Marquet, A. Burger, Y. Mély, *Nucleic Acids Res.* **2005**, *33*, 1031–1039; b) M. E. Hawkins, F. M. Balis, *Nucleic Acids Res.* **2004**, *32*, e62–e62; c) A. Nadler, J. Strohmeier, U. Diederichsen, *Angew. Chem.* **2011**, *123*, 5504–5508; *Angew. Chem. Int. Ed.* **2011**, *50*, 5392–5396; d) S. P. Sau, P. J. Hrdlicka, *J. Org. Chem.* **2012**, *77*, 5–16; e) R. A. Tinsley, N. G. Walter, *RNA* **2006**, *12*, 522–529; f) Q. Xiao, R. T. Ranasinghe, A. M. P. Tang, T. Brown, *Tetrahedron* **2007**, *63*, 3483–3490.
- [4] a) J.-L. H. A. Duprey, Z.-Y. Zhao, D. M. Bassani, J. Manchester, J. S. Vyle, J. H. R. Tucker, *Chem. Commun.* **2011**, *47*, 6629–6631; b) K. Furukawa, M. Hattori, T. Ohki, Y. Kitade, Y. Ueno, *Bioorg. Med. Chem.* **2012**, *20*, 16–24; c) S. Kim, A. Misra, *Annual Rev. Biomed. Eng.* **2007**, *9*, 289–320; d) A. Okamoto, Y. Saito, I. Saito, *J. Photochem. Photobiol. C* **2005**, *6*, 108–122; e) J. H. Ryu, Y. J. Seo, G. T. Hwang, J. Y. Lee, B. H. Kim, *Tetrahedron* **2007**, *63*, 3538–3547; f) H. Zhang, M. Wang, Q. Gao, H. Qi, C. Zhang, *Talanta* **2011**, *84*, 771–776; g) Z.-Y. Zhao, M. San, J.-L. H. A. Duprey, J. R. Arrand, J. S. Vyle, J. H. R. Tucker, *Bioorg. Med. Chem. Lett.* **2012**, *22*, 129–132.
- [5] For reviews, see: a) N. Dai, E. T. Kool, *Chem. Soc. Rev.* **2011**, *40*, 5756–5770; b) M. Hawkins, *Cell Biochem. Biophys.* **2001**, *34*, 257–281.
- [6] Selected examples: a) T. Ono, S. Wang, C.-K. Koo, L. Engstrom, S. S. David, E. T. Kool, *Angew. Chem.* **2012**, *124*, 1721–1724; *Angew. Chem. Int. Ed.* **2012**, *51*, 1689–1692; b) J.-W. Jung, S. K. Edwards, E. T. Kool, *ChemBioChem* **2013**, *14*, 440–444; c) J. Riedl, P. Menova, R. Pohl, P. Orsag, M. Fojta, M. Hocek, *J. Org. Chem.* **2012**, *77*, 8287–8293.
- [7] A. P. Demchenko, *Introduction to Fluorescence Sensing*, Springer, Heidelberg, **2009**, p.612.
- [8] a) E. L. Rachofsky, E. Seibert, J. T. Stivers, R. Osman, J. B. A. Ross, *Biochemistry* **2001**, *40*, 957–967; b) J. Liang, S. Matsika, *J. Am. Chem. Soc.* **2011**, *133*, 6799–6808; c) J. Liang, Q. L. Nguyen, S. Matsika, *Photochem. Photobiol. Sci.* **2013**, *12*, 1387–1400.
- [9] Selected examples: a) N. J. Greco, Y. Tor, *J. Am. Chem. Soc.* **2005**, *127*, 10784–10785; b) N. J. Greco, Y. Tor, *Tetrahedron* **2007**, *63*, 3515–3527; c) A. Okamoto, K. Tainaka, Y. Fujiwara, *J. Org. Chem.* **2006**, *71*, 3592–3598; d) J. Riedl, R. Pohl, L. Rulíšek, M. Hocek, *J. Org. Chem.* **2012**, *77*, 1026–1044; e) Y. Saito, A. Suzuki, S. Ishioroshi, I. Saito, *Tetrahedron Lett.* **2011**, *52*, 4726–4729; f) D. Shin, R. W. Sinkeldam, Y. Tor, *J. Am. Chem. Soc.* **2011**, *133*, 14912–14915; g) A. A. Tanpure, S. G. Srivatsan, *Chem. Eur. J.* **2011**, *17*, 12820–12827; h) Y. Saito, A. Suzuki, Y. Okada, Y. Yamashita, N. Nemoto, I. Saito, *Chem. Commun.* **2013**, *49*, 5684–5686; i) T. Pesnot, L. M. Tedaldi, P. G. Jambrina, E. Rosta, G. K. Wagner, *Org. Biomol. Chem.* **2013**, *11*, 6357–6371.
- [10] A. P. Demchenko, *J. Fluoresc.* **2010**, *20*, 1099–1128.
- [11] a) J. Guo, J. Ju, N. J. Turro, *Anal. Bioanal. Chem.* **2012**, *402*, 3115–3125; b) D. M. Kolpashchikov, *Chem. Rev.* **2010**, *110*, 4709–4723.
- [12] a) A. P. Demchenko, *Trends Biotechnol.* **2005**, *23*, 456–460; b) A. P. Demchenko, *FEBS Lett.* **2006**, *580*, 2951–2957.
- [13] P. K. Sengupta, M. Kasha, *Chem. Phys. Lett.* **1979**, *68*, 382–385.
- [14] a) P. T. Chou, M. L. Martinez, J. H. Clements, *J. Phys. Chem.* **1993**, *97*, 2618–2622; b) T. C. Swinney, D. F. Kelley, *J. Chem. Phys.* **1993**, *99*, 211–221.
- [15] A. S. Klymchenko, A. P. Demchenko, *Phys. Chem. Chem. Phys.* **2003**, *5*, 461–468.
- [16] S. Ercelen, A. S. Klymchenko, A. P. Demchenko, *Anal. Chim. Acta* **2002**, *464*, 273–287.
- [17] N. Chattopadhyay, M. Barroso, C. Serpa, L. G. Arnaut, D. J. Formosinho, *Chem. Phys. Lett.* **2004**, *387*, 258–262.
- [18] a) A. S. Klymchenko, G. Duportail, Y. Mély, A. P. Demchenko, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 11219–11224; b) A. S. Klymchenko, Y. Mély, A. P. Demchenko, G. Duportail, *Biochim. Biophys. Acta Biomembr.* **2004**, *1665*, 6–19; c) V. V. Shynkar, A. S. Klymchenko, C. Kunzelmann, G. Duportail, C. D. Muller, A. P. Demchenko, J.-M. Freyssinet, Y. Mély, *J. Am. Chem. Soc.* **2007**, *129*, 2187–2193; d) S. Oncul, A. S. Klymchenko, O. A. Kucherak, A. P. Demchenko, S. Martin, M. Dontenwill, Y. Arntz, P. Didier, G. Duportail, Y. Mély, *Biochim. Biophys. Acta Biomembr.* **2010**, *1798*, 1436–1443; e) V. V. Shynkar, A. S. Klymchenko, G. Duportail, A. P. Demchenko, Y. Mély, *Biochim. Biophys. Acta Biomembr.* **2005**, *1712*, 128–136.
- [19] a) S. Ercelen, A. S. Klymchenko, A. P. Demchenko, *FEBS Lett.* **2003**, *538*, 25–28; b) A. S. Klymchenko, S. V. Avilov, A. P. Demchenko, *Anal. Biochem.* **2004**, *329*, 43–57; c) V. Y. Postupalenko, V. V. Shvadchak, G. Duportail, V. G. Pivovarenko, A. S. Klymchenko, Y. Mély, *Biochim. Biophys. Acta Biomembr.* **2011**, *1808*, 424–432; d) V. V. Shvadchak, A. S. Klym-

- chenko, H. de Rocquigny, Y. Mély, *Nucleic Acids Res.* **2009**, *37*, e25; e) A. V. Strizhak, V. Y. Postupalenko, V. V. Shvadchak, N. Morellet, E. Guittet, V. G. Pivovarenko, A. S. Klymchenko, Y. Mély, *Bioconjugate Chem.* **2012**, *23*, 2434–2443.
- [20] A. S. Klymchenko, V. V. Shvadchak, D. A. Yushchenko, N. Jain, Y. Mély, *J. Phys. Chem. B* **2008**, *112*, 12050–12055.
- [21] M. Spadafora, V. Y. Postupalenko, V. V. Shvadchak, A. S. Klymchenko, Y. Mély, A. Burger, R. Benhida, *Tetrahedron* **2009**, *65*, 7809–7816.
- [22] D. Dziuba, V. Y. Postupalenko, M. Spadafora, A. S. Klymchenko, V. Guérineau, Y. Mély, R. Benhida, A. Burger, *J. Am. Chem. Soc.* **2012**, *134*, 10209–10213.
- [23] L. M. Wysocki, L. D. Lavis, *Curr. Opin. Chem. Biol.* **2011**, *15*, 752–759.
- [24] a) S. M. Ormon, R. G. Brown, F. Vollmer, W. Rettig, *J. Photochem. Photobiol. A* **1994**, *81*, 65–72; b) A. Sytnik, D. Gormin, M. Kasha, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 11968–11972.
- [25] G. M'Baye, A. S. Klymchenko, D. A. Yushchenko, V. V. Shvadchak, T. Ozturk, Y. Mély, G. Duportail, *Photochem. Photobiol. Sci.* **2007**, *6*, 71–76.
- [26] T. Doussineau, M. Smaihi, S. Balme, J.-M. Janot, *ChemPhysChem* **2006**, *7*, 583–589.
- [27] When the present research was in progress, an independent report on facile one-pot synthesis of compounds **2** and **10** in 70 and 55% yield, respectively, was published, see: S. Gunduz, A. C. Goren, T. Ozturk, *Org. Lett.* **2012**, *14*, 1576–1579.
- [28] D. Dziuba, R. Benhida, A. Burger, *Synthesis* **2011**, *13*, 2159–2164.
- [29] a) K. L. Seley, S. Salim, L. Zhang, P. I. O'Daniel, *J. Org. Chem.* **2005**, *70*, 1612–1619; b) G. D. Allred, L. S. Liebeskind, *J. Am. Chem. Soc.* **1996**, *118*, 2748–2749.
- [30] C. Reichardt, *Chem. Rev.* **1994**, *94*, 2319–2358.
- [31] M. H. Abraham, *J. Phys. Org. Chem.* **1993**, *6*, 660–684.
- [32] A. S. Klymchenko, T. Ozturk, V. G. Pivovarenko, A. P. Demchenko, *Can. J. Chem.* **2001**, *79*, 358–363.
- [33] A. P. Demchenko, K.-C. Tang, P.-T. Chou, *Chem. Soc. Rev.* **2013**, *42*, 1379–1408.
- [34] a) C. A. Kenfack, A. S. Klymchenko, G. Duportail, A. Burger, Y. Mély, *Phys. Chem. Chem. Phys.* **2012**, *14*, 8910–8918; b) C.-C. Hsieh, C.-M. Jiang, P.-T. Chou, *Acc. Chem. Res.* **2010**, *43*, 1364–1374.
- [35] a) A. S. Klymchenko, V. G. Pivovarenko, T. Ozturk, A. P. Demchenko, *New J. Chem.* **2003**, *27*, 1336–1343; b) O. M. Zamotaiev, V. Y. Postupalenko, V. V. Shvadchak, V. G. Pivovarenko, A. S. Klymchenko, Y. Mély, *Bioconjugate Chem.* **2011**, *22*, 101–107.
- [36] V. V. Shynkar, Y. Mély, G. Duportail, E. Piémont, A. S. Klymchenko, A. P. Demchenko, *J. Phys. Chem. A* **2003**, *107*, 9522–9529.
- [37] O. A. Kucherak, L. Richert, Y. Mély, A. S. Klymchenko, *Phys. Chem. Chem. Phys.* **2012**, *14*, 2292–2300.
- [38] a) M. Weinberger, F. Berndt, R. Mahrwald, N. P. Ernsting, H.-A. Wagenknecht, *J. Org. Chem.* **2013**, *78*, 2589–2599; b) A. Fakhari, M. S. E. Rokita, *Chem. Commun.* **2011**, *47*, 4222–4224; c) J. Riedl, R. Pohl, N. P. Ernsting, P. Orsag, M. Fojtá, M. Hocek, *Chem. Sci.* **2012**, *3*, 2797–2806; d) K. Tainaka, K. Tanaka, S. Ikeda, N.-J. Nishiza, T. Unzai, Y. Fujiwara, I. Saito, A. Okamoto, *J. Am. Chem. Soc.* **2007**, *129*, 4776–4784.
- [39] R. Das, A. S. Klymchenko, G. Duportail, Y. Mély, *Photochem. Photobiol. Sci.* **2009**, *8*, 1583–1589.
- [40] a) M. D. Bilokin, V. V. Shvadchak, D. A. Yushchenko, G. Duportail, Y. Mély, V. G. Pivovarenko, *J. Fluoresc.* **2008**, *18*, 545–553; b) D. A. Yushchenko, V. V. Shvadchak, A. S. Klymchenko, G. Duportail, V. G. Pivovarenko, Y. Mély, *J. Phys. Chem. A* **2007**, *111*, 8986–8992.
- [41] V. V. Shynkar, A. S. Klymchenko, E. Piémont, A. P. Demchenko, Y. Mély, *J. Phys. Chem. A* **2004**, *108*, 8151–8159.
- [42] a) R. W. Sinkeldam, N. J. Greco, Y. Tor, *ChemBioChem* **2008**, *9*, 706–709; b) T. Kimura, K. Kawai, T. Majima, *Chem. Commun.* **2006**, 1542–1544; c) T. Kimura, K. Kawai, T. Majima, *Org. Lett.* **2005**, *7*, 5829–5832; d) V. R. Jadhav, D. A. Barawkar, K. N. Ganesh, *J. Phys. Chem. B* **1999**, *103*, 7383–7385; e) M. A. Young, B. Jayaram, D. L. Beveridge, *J. Phys. Chem. B* **1998**, *102*, 7666–7669; f) G. Lamm, G. R. Pack, *J. Phys. Chem. B* **1997**, *101*, 959–965.
- [43] W. L. F. Armarego, C. L. L. Chai, *Purification of Laboratory Chemicals*, 7th ed., Butterworth-Heinemann, Oxford, **2013**, p. 1024.
- [44] W. C. Still, M. Kahn, A. Mitra, *J. Org. Chem.* **1978**, *43*, 2923–2925.
- [45] H. E. Gottlieb, V. Kotlyar, A. Nudelman, *J. Org. Chem.* **1997**, *62*, 7512–7515.
- [46] a) A. K. Livesey, J.-C. Brochon, *Biophys. J.* **1987**, *52*, 693–706; b) J.-C. Brochon, *Methods Enzymol.* **1994**, *240*, 262–311.

Received: August 29, 2013

Published online on ■■■, 2013

PUBLICATION 3

Dans une seconde étape, nous nous sommes intéressés au couplage de la 3-hydroxychomone à la 2'-désoxyuridine, ce motif ayant pour vocation d'être introduit dans des brins d'ADN. Nous nous sommes focalisés sur 3 analogues nucléosidiques pour lesquels trois espaces aryles différents (phényle **7a**, furyle **7b** et thiényle **7c**) ont été utilisés afin d'étudier l'impact du lien arylique sur les propriétés de fluorescence de la sonde [cf. Figure 3].

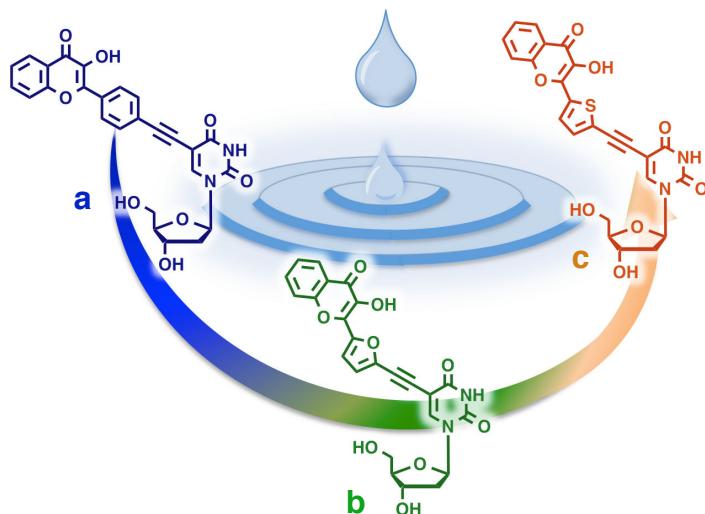


Figure 7. Analogues nucléosidiques thiényle, furyle et phényle.

Pour cela nous avons dans un premier temps essayé de reproduire le schéma de synthèse utilisé pour la préparation du composé modèle **TCeU** [cf. Figure 6]. Cependant, l'introduction de la copule fluorescente sur la nucléobase, par un couplage de Sonogashira, s'est révélée plus délicate que précédemment. En effet, la base uracile n'étant plus diméthylée, la réaction a conduit au produit recherché, contaminé par un sous-produit fluorescent difficilement séparable, probablement obtenu suite à une réaction de cyclisation *5-endo-dig*. Une étude approfondie a permis de déterminer que la protection de l'azote 3 de la dU par un groupement toluoyle permettait de solutionner le problème. Dans ces conditions, les analogues nucléosidiques comportant les trois liens aryles et couplés à la copule fluorescente ont été obtenus avec de bons rendements.

Les nucléosides déprotégés **7a,b** et **c** ont alors pu être caractérisés par spectroscopie d'absorption UV-visible et de fluorescence. Ces études ont permis de montrer que les caractéristiques photophysiques du composé **7c** [cf. Figure 7] étaient similaires à celles de son analogue diméthylé **TCeU** [cf. Figure 6]. Le composé furyle **7b** présente des propriétés photophysiques voisines avec cependant un maximum d'absorption faiblement décalé vers le bleu. En revanche, le dérivé phényle **7a** possède des propriétés similaires à la sonde de première génération. Il se distingue donc des composés **7b** et **7c** qui possèdent les propriétés les plus intéressantes. De plus, ces études ont permis de mettre en évidence une forte sensibilité de ces deux nucléosides fluorescents **7b** et **7c** au taux d'hydratation environnant. L'eau jouant un rôle clé dans le contrôle de la structure des acides nucléiques et leur fonction grâce aux interactions électrostatiques et liaisons

hydrogènes avec ces macromolécules, ils ont naturellement été retenus pour être incorporés dans des ODNs. Ces travaux firent l'objet d'un article, publié en Avril 2015, dans *RSC Advances*.^{126g}

ARTICLE

Cite this: DOI: 10.1039/x0xx00000x

Received 00th January 2012,
Accepted 00th January 2012

DOI: 10.1039/x0xx00000x

www.rsc.org/

Development of environmentally sensitive fluorescent and dual emissive deoxyuridine analogues

N. P. F. Barthes,^a I. A. Karpenko,^{a,b} D. Dziuba,^{a,c} M. Spadafora,^a J. Auffret,^a A. P. Demchenko,^d Y. Mély,^e R. Benhida,^a B. Y. Michel,^{*a} A. Burger^{*a}

Ratiometric and environment-sensitive fluorescent dyes present attractive advantages for sensing interactions in DNA research. Here, we report the rational design, synthesis, and photophysical characterization of 2-thienyl-, 2-furyl- and 2-phenyl-3-hydroxychromones bonded to the C-5 position of deoxyuridine. Since these two-color nucleosides were designed for incorporation into ODNs, we also investigated the sensitivity of the ratiometric response to hydration by using acetonitrile/water mixtures and neat solvents. The synthesized 2-thienyl and 2-furyl conjugates were found to exhibit more red-shifted absorption (by 31–36 nm) and emission (by 77–81 nm of N* band), two-fold increased molar absorption coefficients, and dramatically enhanced (by 3–4.5 times) fluorescence quantum yields. Demonstrating one manifold increase in brightness, they preserve the ability of exquisite ratiometric responses to solvent polarity and hydration. This makes the new fluorescent nucleoside analogues highly relevant for subsequent labeling of the major groove in nucleic acids and sensing their interactions.

Introduction

Fluorescence spectroscopy is a highly versatile tool widely applied in biomolecular researches.¹ Unfortunately, the very low intrinsic fluorescence of DNA allows only a very limited number of applications by this technique.² In contrast, the introduction of fluorescent probes provides a broad access to a variety of applications that include probing DNA hybridization,³ typing single-nucleotide polymorphism (SNP),⁴ and monitoring the dynamics of DNA/protein complexes,^{5,6} to name a few. One major approach to design a DNA-fluorescent sensor is to incorporate into DNA, a fluorescent nucleoside analogue as a fluorescent signal transducer.⁷ Fluorophores that exhibit extreme sensitivity to environmental changes and interactions are highly desirable to provide site-specific responses in sensing.

A large number of analogues of fluorescent nucleosides, such as the widely used 2-amino-purine label (2-AP), responds to environmental changes by a variation in the intensity of fluorescence.⁸ However, low sensitivity to the surrounding environment and low quantum yields in DNA duplexes limit the use of such intensiometric probes. These limitations encourage investigating the synthesis of advanced emissive nucleosides based on new mechanisms of response. Ratiometric sensing, also referred to as λ -ratiometric,⁹ is one such approach, and is based on recording a ratio of the intensities at two or more wavelengths. Compared to simple intensity sensing, ratiometric sensing is advantageous since it compensates for instrumental factors such as fluctuations in the intensity of the light source, and because the resulting intrinsically calibrated

analytical response does not depend on the concentration of the applied dye.¹⁰ To this point, two-channel probing was mainly based on double-labeling of nucleic acids as for the construction of fluorescence resonance energy transfer (FRET) and excimer pairs.^{4e,f,5j,6c} Difficulties and cost of synthesis obviously represent limitations for this approach. Alternatively, λ -ratiometric sensing can be performed with a single fluorophore⁹ using environment sensitive probes. One class of these probes comprises single-band solvatochromic dyes, which respond by a spectral shift due to an excited-state intramolecular charge transfer (ICT). However, these single λ -ratiometric emitters remain faintly explored in the field of DNA sensing¹¹ mainly due to the strong variations of the intensity of fluorescence that accompany the shifts. Dual emissive fluorophores emerge as a second class. These dyes exhibit two emission bands as a result of twisted excited-state intramolecular charge transfer (TICT)¹² or excited-state intramolecular proton transfer (ESIPT).¹³ Dual emissive dyes are very attractive because they offer facile and straightforward quantification through the ratio of their two bands. Nevertheless, they are still poorly exploited because of the very limited number of dyes that display such properties and the difficulties in their syntheses.

Among fluorophores with dual emission, 3-hydroxychromones (3HCs) appear particularly promising in terms of photophysics and applications as biosensor units. Due to the ESIPT reaction, 3HCs exhibit two excited states, the normal form (N*) and the tautomer form (T*), that provide two well separated emission bands (Fig. 1).¹⁴ 3HC fluorophores offer exceptional solvatochromism,¹⁵ since both an increase in

ARTICLE

the hydrogen bond donor strength and the dielectric constant of solvents inhibit the ESIPT reaction and thus, decrease the relative intensity of the T^* band.¹⁶ Therefore, changes in the polarity and the hydration of the microenvironment of 3HCs can be monitored by measuring the ratio of the intensities of the two emission bands (I_{N^*}/I_{T^*}).¹⁵ This unique behaviour of 3HCs was the cornerstone for the design of environment-sensitive fluorescent probes for a large range of applications. These include the characterization of supercritical fluids,¹⁷ micelles,¹⁸ phospholipids bilayers,¹⁹ plasma membranes of living cells,²⁰ and biomolecular interactions.^{21,22} Their applications as fluorescent nucleoside analogues to label oligonucleotides (ODNs) has only been recently explored.^{23,24}

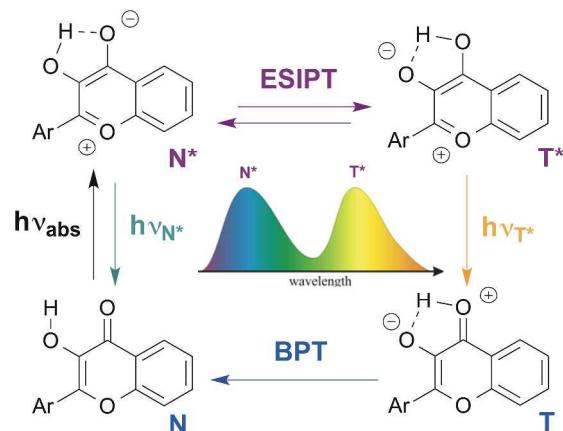


Fig. 1 Excited State Intramolecular Proton Transfer (ESIPT) reaction of 3-hydroxychromones: genesis of the dual emission. BPT: Back Proton Transfer.

The fluorescent C-nucleoside **1** using 2-thienyl-3-hydroxychromone moiety was first designed as a substitute for natural bases for internal labeling of ODNs and sensing (Fig. 2).^{25,26} As a next step, external labeling of DNA major groove was addressed by compiling 2-thienyl-3-hydroxychromone and uracil at the 5-position via an ethynyl bond.²⁶ The synthesized conjugate demonstrated that the nucleobase and 3HC fragments were coupled into an electronic conjugated system, which attained strong ICT character upon excitation.

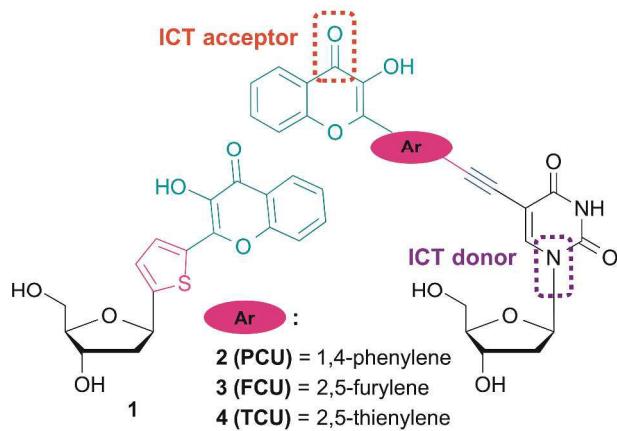


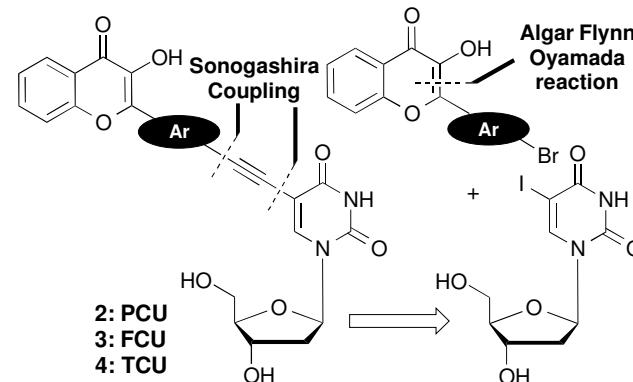
Fig. 2 Fluorescent nucleoside analogues incorporating 3-hydroxychromone fluorophore as a nucleobase surrogate **1**^{23,24} and natural base modifier **2-4**.²⁵

The ICT states are crucial in achieving the solvent-dependent two-band emission of 3HC dyes. The substitutions in the position 2 of the chromone demonstrates the strongest modulating power^{27a} that allows reaching a dynamic equilibrium between the populations of ICT and ESIPT states, from which the dual emission originates.^{27b} With the aim to label DNA and to tune the spectroscopic properties of the conjugates for sensing, reliable synthesis of 2'-deoxyuridine conjugates bearing aryl groups of different natures connected to the chromone was developed (Fig. 2). Herein, we report the synthesis and the photophysics of 2'-deoxyuridine coupled to 2-phenyl-, 2-furyl- and 2-thienyl-3-hydroxychromone referred as **PCU**, **FCU** and **TCU**, respectively (**2-4**).

Results and discussion

The 3HC deoxyuridines **PCU**, **FCU**, and **TCU** were furnished by a convergent synthetic strategy based on Sonogashira couplings and Algar-Flynn-Oyamada reaction (Scheme 1). Sonogashira reaction²⁸ is widely applied to couple aromatic scaffolds, including dyes such as pyrenes,^{12b,29} fluorenes,³⁰ naphthalenes and anthracenes,³¹ to deoxyuridine. This palladium-catalyzed cross-coupling proceeds usually by assembling a C-5 iodinated uracil moiety with an aromatic alkyne.³² Therefore, the ethynyl moiety is generally introduced to the aromatic molecule through a cross-coupling with TMS-acetylene, after which the alkyne is deprotected from its silyl group and directly coupled to the heterocyclic 5-iodo-2'-deoxyuridine derivative.

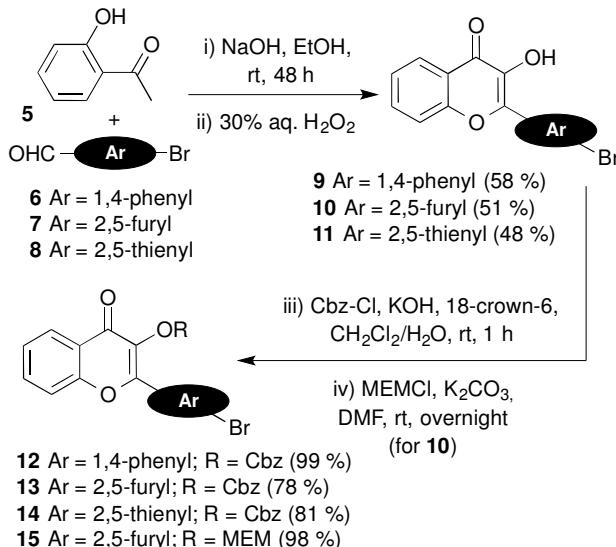
We applied this approach for the preparation of the target compounds. However, we faced difficulties in the elaboration of ethynylchromones (Scheme S1), and found that the Sonogashira coupling between our tested ethynylchromone and 5-ido-2'-deoxyuridine was sluggish and low yielding. This might be attributed to a competition between homo- and cross-couplings. Indeed, it has been known that homo-coupling is favored in the presence of electron-withdrawing groups on aryl alkynes.²⁸ So the 3HC substitution brings electron deficiency to the terminal alkyne, consequently increasing the competition of the homo- and cross-couplings to access the catalyst. To overcome these difficulties, we adopted an alternative approach based on reacting the electron-rich ethynyl-deoxyuridine derivative with the electron-poor bromo-arylchromones. The new suggested strategy established a more reliable synthetic access to the target compounds.



Scheme 1. Retrosynthetic analysis of the targeted two-color emission nucleosides **2-4**.

The preparation of the 3HC key intermediates **9-15** required for the subsequent coupling with the 5-ethynyldeoxyuridine

derivative is described in Scheme 2. The bromochromones **9–11** were obtained using the fast and versatile Algar-Flynn-Oyamada reaction.³³ Thus, starting from 2-hydroxyacetophenone **5**, an aldol condensation/dehydratation process with aldehydes **6–8** in the presence of sodium hydroxide in ethanol followed by an oxidative cyclization with aqueous hydrogen peroxide provided the chromones **9–11** (48–58 %).



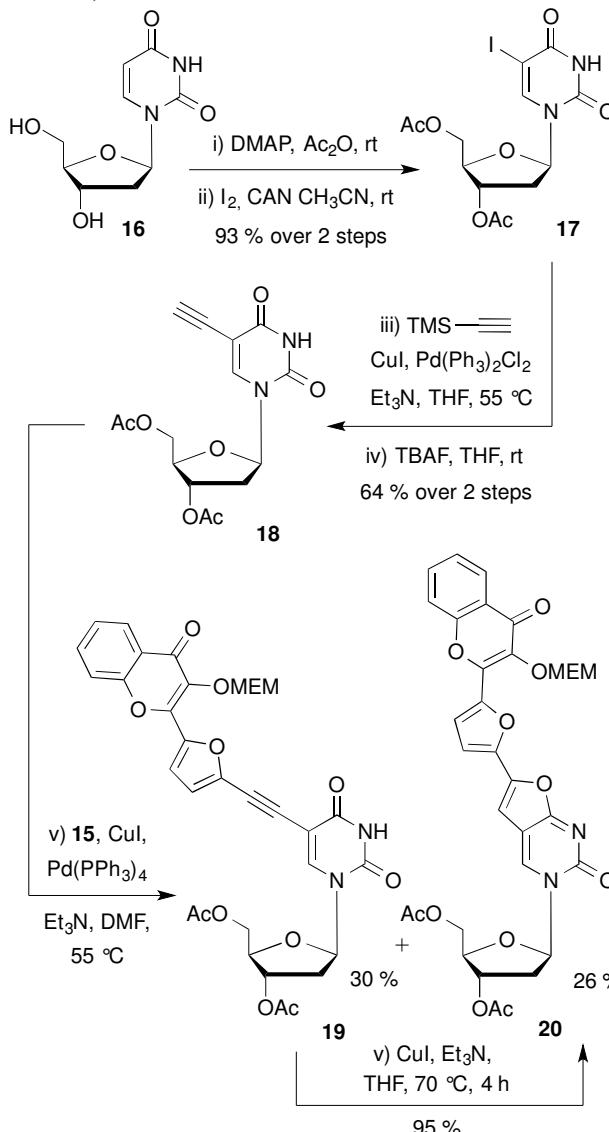
Scheme 2. Synthetic preparation of the fluorescent chromone-coupling partners.

The synthesis of the chromone-deoxyuridine conjugates required masking the reactive 3-OH group of the chromone. The benzyloxycarbonyl group (Cbz), which was successfully employed during the synthesis of 3HC-containing C-nucleosides and ODNs,²⁴ and the more robust methoxyethoxymethyl group (MEM)²⁵ were chosen as the protecting groups. The Cbz protection of the 3-OH group of chromones **9–11** was performed under phase-transfer catalysis conditions,³⁴ delivering compounds **12–14** in good yields. Protection with the MEM group was obtained by treatment of the furyl chromone **10** with potassium carbonate and MEMCl in DMF, producing compound **15**.

Next, we synthesized the 5-ethynyldeoxyuridine partner **18**, using common procedures (Scheme 3). The protection of deoxyuridine **16** before the iodination of the nucleobase at the C-5 position with cerium(IV) ammonium nitrate turned out to be more convenient and efficient.³⁵ Sonogashira coupling of the iodo intermediate **17** with TMS-acetylene followed by treatment with tetrabutylammonium fluoride furnished the terminal alkyne **18** in satisfactory yields (64% in 2 steps).[†]

Preliminary attempts of the final assembly were performed with the protected bromochromone **15** (Scheme 3). Different conditions were examined including the catalyst (Pd(PPh₃)₄ or PdCl₂(PPh₃)₂), the loading (5–20 mol%), and the solvent (toluene, dioxane, THF and DMF). All screenings converged to two major products: the desired coupling derivative **19** and the 5-*endo*-dig cyclization side-product **20**. Despite the use of DMF, which is reported to minimize the formation of the bicyclic furanopyrimidine,³⁶ the formation of the *endo* product could not be avoided. The ethynyl-coupled product **19** was difficult to purify and thus, was obtained in 30 % yield after silica gel and reverse phase chromatographies. NMR and MS

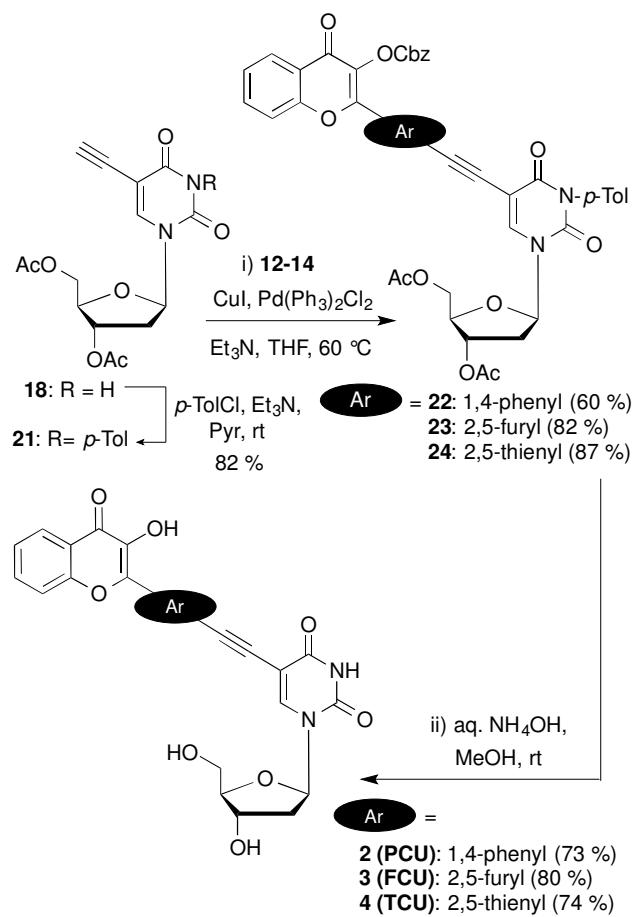
analysis,^{††} and chemical means supported the proposed structure of the *endo* product. As reported amongst others³⁷ by McGuigan et al.,^{38–40} after the Sonogashira coupling, an extended treatment with CuI/Et₃N in THF or MeOH at reflux for several hours enabled us to fully convert the coupled alkynyl derivative **17** into the 5-*endo* cyclized side product **18** (Scheme 3).



Scheme 3. Preliminary Sonogashira couplings on substrate **18** having no *N*³-protecting group on the uracil moiety.

Endo-cyclisation of 5-ethynyluridine derivatives is a well-known reaction that is favored by prolonged reaction time at elevated temperatures, high CuI loading, and the presence of electron-withdrawing groups on the coupling partner.⁴¹ Although the copper-free Sonogashira coupling was widely reported in the last decade,⁴² no product at all was formed when copper-free conditions were applied to the reaction of acetylenic and iodo partners as evidenced by TLC. Protection of the *N*³-imide with the base labile benzoyl group is known to prevent 5-*endo* cyclisation.⁴³ Consequently, we selected this type of protection, using *p*-toluoyl rather than benzoyl, since the former presents the advantage of displaying a characteristic

methyl singlet easy to detect by NMR. To set up the imide protecting group, mild conditions (*p*-toloyl chloride, Et₃N in pyridine) were employed to provide **21** (82%) as the building block for the final assembly of the two fragments (Scheme 4). Second Sonogashira coupling was then efficiently used for that purpose. Thus, when using a standard catalytic system (PdCl₂(Ph₃)₂, CuI, Et₃N in THF), the reaction proceeded cleanly to give the desired coupled derivatives **22–24**. It is noteworthy that the yields of C(sp)-C(sp²) cross coupling for the five-membered ring partners **23** and **24** (82 and 87%) are more satisfactory as compared to the one obtained with the phenyl moiety **22** (60%). The lower yield of the phenyl derivative could be attributed to the higher steric hindrance of the six membered ring and/or to the lower reactivity of the more electronically poor bromophenyl moiety with the electron withdrawing chromone. A final deprotection of the esters, carbonate and amide groups via a simple treatment with aqueous ammonia provided the three targeted conjugated fluorescent compounds **2** (PCU), **3** (FCU) and **4** (TCU), which were used for photophysical characterizations.



Scheme 4. Synthesis of dual emissive fluorescent nucleosides **2**, **3** and **4** bearing a 3HC scaffold.

Photophysical studies

Results

UV/Vis absorption and fluorescence properties of the three chromone-nucleoside analogues **PCU**, **FCU**, and **TCU** were investigated in methanol and compared with those of the

corresponding parent 2-phenyl, 2-furyl and 2-thienyl-3-hydroxychromones, **PC**, **FC**, and **TC** respectively (Table 1, Fig. S1).

These studies brought interesting results. In point of fact, the positions of the light absorption bands of the new derivatives were located in the same sequence as that of the parent 3HCs but demonstrated red-shifts by 25–36 nm, which could be explained by an electronic conjugation with uracil moiety (Fig. 3). Such conjugation should increase dramatically the light absorption cross-section leading to increase correspondingly the molar absorption coefficients. **PCU**, **FCU**, and **TCU** absorbed light about twice stronger, with molar absorption coefficients of about 27000, 37000 and 39000 M⁻¹ cm⁻¹. It must be noted that the shift of the absorption maxima to the red, and the increase of the molar absorption coefficients proceed in the following sequence: **PCU** < **FCU** < **TCU**. The more red-shifted absorption and emission of the thienyl and furyl derivatives as well as the increased absorption coefficients are in agreement with the higher electronic density of substituents in position 2 of the chromone ring, respectively. In our case, this may result in more efficient electronic coupling between the chromone and uracil electronic systems. The differences with the phenyl substituent can be accounted for the fact that the five-membered rings are more electron rich and allow to adopt a more planar conformation with respect to chromone for better coupling.⁴⁴ Thus, the novel conjugates display substantial shifts to the red of fluorescence spectra, for instance in methanol, the shifts of N* and T* bands are ca. 80 and 25 nm, respectively (Table 1, Fig. 3). The fluorescence quantum yields were also increased effectively, so that the brightness (defined commonly as the product of molar absorbance and quantum yield) could be increased roughly by one order of magnitude.

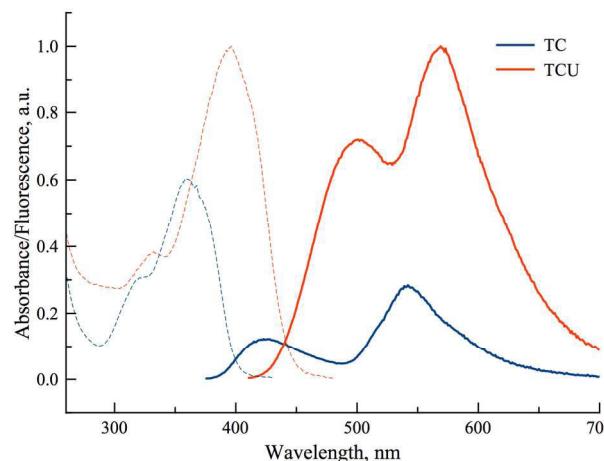


Fig. 3 Absorption and fluorescence emission spectra of **TC** (blue), and **TCU** (red) in methanol (excitation wavelength corresponds to the absorbance maximum of each compound). Absorptions and emissions are proportional to their absorption coefficients and quantum yields, respectively.

On the next step, the properties of **PCU**, **FCU** and **TCU** conjugates were investigated and compared in a set of solvents presenting a large range of polarities (Table 2). For these studies, five aprotic solvents (acetonitrile, acetone, chloroform, ethyl acetate, toluene) and four protic solvents (hexafluoropropan-2-ol (HFIP), H₂O, MeOH, EtOH) were selected.

Table 1. Spectroscopic comparison in methanol of parent 2-aryl-3-HCs (**PC**, **TC** and **TC**) and corresponding conjugates (**PCU**, **FCU** and **TCU**).

Dye	$\lambda_{\text{abs}}^{\text{a}}$	$\Delta\lambda_{\text{abs}}^{\text{b}}$	ε^{c}	$\varepsilon_d^{\square}/\varepsilon_d^{\square}$	$\lambda_{N^*}^{\text{e}}$	$\Delta\lambda_{N^*}^{\text{f}}$	Φ^{g}	$\Phi_h^{\square}/\Phi_d^{\square}$
PC ^{46,47}	343	-	13300	-	403	-	0.03	-
PCU	368	25	27000	2	500	97	0.30	10
FC ^{46,48}	356	-	16400	-	416	-	0.08	-
FCU	387	31	37000	2.2	497	81	0.24	3
TC ²⁶	359	-	23500	-	424	-	0.07	-
TCU	395	36	39000	1.7	501	77	0.32	4.5

Footnotes: a) Position of the absorption maximum in nm; b) difference of absorption maxima between the conjugate and the parent 3HC in nm; c) molar absorption coefficient in $M^{-1}\cdot cm^{-1}$; d) ratio of molar absorption coefficients between the conjugate and the parent 3HC; e) position of the normal N^* band maximum in nm; f) difference of emission maxima between the conjugate and the parent 3HC in nm; g) quantum yields determined using 3-hydroxyflavone in toluene ($\Phi=0.29$),⁴⁷ 4'-dimethylamino)-3-hydroxyflavone in ethanol ($\Phi=0.27$)⁴⁹ or quinine sulfate in 0.1 M H_2SO_4 ($\Phi=0.54$)⁵⁰ as standard references; h) ratio of quantum yields between the conjugate and the parent 3HC.

Table 2. Photophysical properties of **PCU**, **FCU** and **TCU** in different solvents.

Solvent	$E_T(30)^{\text{a}}$	$\lambda_{N^*}^{\text{c}}$			$\lambda_{T^*}^{\text{d}}$			$I_{N^*}/I_{T^*}^{\text{e}}$			Φ^{f}		
		PCU	FCU	TCU	PCU	FCU	TCU	PCU	FCU	TCU	PCU	FCU	TCU
HFIP	65.3	472	493	500	—	—	—	—	—	—	0.28	0.33	0.40
H_2O	63.1	480	499	502	525	—	—	0.74	—	—	0.05	0.13	0.06
MeOH	55.4	500	497	504	547	551	568	0.64	1.03	0.79	0.30	0.24	0.32
EtOH	51.9	489	488	501	552	566	575	0.33	0.56	0.44	0.26	0.22	0.23
CH_3CN	45.6	445	466	474	549	569	580	0.03	0.20	0.28	0.33	0.17	0.19
Acetone	42.2	455	462	469	553	574	583	0.04	0.17	0.29	0.22	0.16	0.25
$CHCl_3$	39.1	426	454	459	543	564	572	0.02	0.15	0.17	0.40	0.23	0.19
$EtOAc$	38.1	424	444	457	551	573	582	0.03	0.10	0.20	0.34	0.22	0.20

Footnotes: a) Reichardt's empirical solvent polarity index;⁴⁵ b) position of the absorption maximum; c) position of the N^* band maximum; d) position of the T^* band maximum; e) ratio of the two intensity maxima; f) quantum yields determined using 3-hydroxyflavone in toluene ($\Phi=0.29$),⁴⁷ 4'-dimethylamino)-3-hydroxyflavone in ethanol ($\Phi=0.27$)⁴⁹ or quinine sulfate in 0.1 M H_2SO_4 ($\Phi=0.54$) as references.⁵⁰

The Reichardt's polarity parameter $E_T(30)$ that accounts for the dielectric constant of the solvent and its H-bond donor ability was employed to scale the polarity of the solvents.⁴⁵ We observed that the positions of the absorption maxima were similar to the ones observed in MeOH (Table 1). Therefore, they correlate poorly with the solvent polarity displaying almost no solvatochromism, which is the evidence of an absence of charge-transfer character in their ground states.

PCU presented a dual emission in protic solvents, except HFIP. The short and long wavelength bands can be attributed to the N^* and T^* states. In contrast, it presented almost exclusively the T^* band in all aprotic solvents (Fig. 4). In all solvents, except water and HFIP, **FCU** and **TCU** exhibited the two emission bands (Fig. 4). The positive solvatochromism of the N^* band is typical for ESIPT probes with ICT character.^{26,27} The ICT character can be further estimated using the polarity scale of Lippert (Table S1 and Fig S5, S6).^{1a,51} Plotting the Stokes shifts as a function of the orientation polarizability of the aprotic solvents demonstrated linear fits with positive slopes for both compounds. These results are compatible with an increase of the dipole moments of the normal excited states of **FCU** and **TCU** after Franck-Condon light absorption. The position of their T^* band maxima appeared blue-shifted in protic solvents, as compared to aprotic ones. The blue shift can be explained by the formation of an H-bond between the 3-phenoxy oxygen in the T^* state and the donor proton in the solvent.^{52,53}

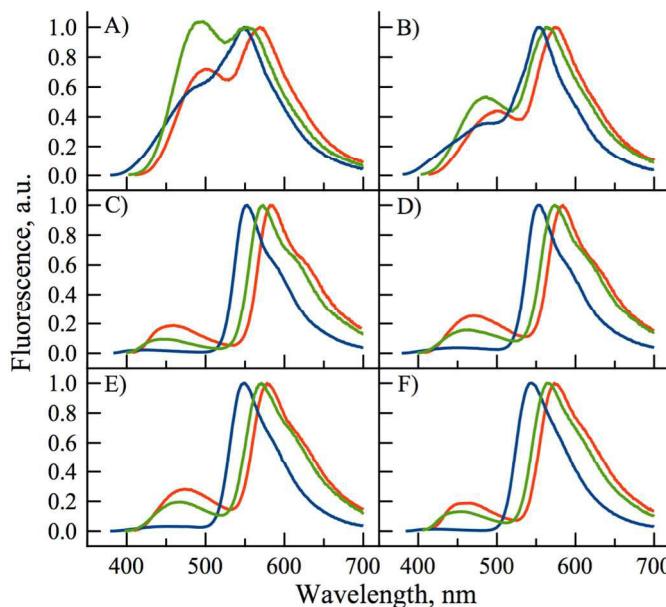


Fig. 4 Normalized fluorescence spectra of **PCU** (blue), **FCU** (green) and **TCU** (red) in: A) methanol, B) ethanol, C) ethyl acetate, D) acetone, E) acetonitrile and F) chloroform. All the spectra are normalized by T^* band (excitation wavelength corresponds to the absorbance maximum of each compound).

ARTICLE

All the three compounds displayed an increase in the I_{N^*}/I_{T^*} ratio by rising solvent polarity and H-bond donor strength (Table 2, Fig. 4). However, **FCU** and **TCU** were much more sensitive than **PCU**. As a general trend, the I_{N^*}/I_{T^*} ratio of **FCU** and **TCU** gradually increased when going from the most apolar aprotic solvent to the most polar protic solvent (Table 2, Fig. S2, S3, S4). It is worth indicating that the effect of H-bonding dominates the polarity effect, as that could be seen from the 3- to 5-fold larger value of the I_{N^*}/I_{T^*} ratio in methanol as compared to acetonitrile, though both solvents exhibit similar dielectric constants. Analysing the plots of Lippert further supports this interpretation. In protic solvents, the Stokes shifts were deviated to the red from the linear function indicating specific interactions between the protic solvent and the 3HC (Fig S5, S6). The deviation from the linear fit in protic solvents is known for donor-acceptor solvatochromic dyes.^{1a} The sensitivity of **FCU** and **TCU** is typical of the 3HCs exhibiting significant ICT characters, such as 2-(4-methoxy)phenyl-3HCs⁵⁴ and 2-(4-dialkylamino)phenyl-3HCs.^{16,55}

For **PCU**, the contribution of the N^* form was marginal in aprotic solvents hinting that **PCU** was mainly sensitive to protic solvents that donate protons for intermolecular H-bonding with the carbonyl at position 4 of 3HC. **PCU** behaves like the parent 3-hydroxyflavone,⁵⁶ but with a lower resolution of its dual emission.

The quantum yields of the three dyes were significantly enhanced for all the compounds in comparison with their parent analogs (Table 1). As a function of solvent, they range from 5 to 40 %. Quantum yields were in the range of 22 to 40 % in all the tested alcohols and 5 to 13 % in water. However, the quantum yields in HFIP, which is more polar and acidic than water,⁵⁷ are about 3 to 7 times larger than in water. Therefore, the reduced quantum yields in water are unlikely due to the H-bond strength between the donating water molecule and the accepting oxygen carbonyl of 3HC (Scheme 5). Reduction of quantum yields in water for solvatochromic dyes exhibiting strong ICT character is common. It was proposed that several water molecules act cooperatively to capture electrons from the excited-state species resulting in quenching of fluorescence.⁵⁸ Though being reduced in comparison with other solvents, the quantum yields of the three compounds are still significant in water making them attractive for sensing hydrated biological media. This is a major difference as compared to many other solvatochromic dyes, which are severely quenched in water.⁵⁹

To get further insight on the sensitivity of the photophysical properties of **PTU**, **FCU**, and **TCU** on hydration, emission spectra were recorded in mixtures of acetonitrile and water (Fig. 5). The same set of experiments was also performed in dioxane/water mixtures in order to determine the contribution from the organic solvent to the spectral changes (Fig. S7). A dual emission was evident in all the tested compositions for **PCU** and up to 70% and 90% water for **FCU** and **TCU**, respectively. In all cases, hydration increased the N^*/T^* ratio and shifted the T^* band position to the blue, as it was observed in neat protic solvents (Table 2). The ratios of the intensities correlated linearly with the water concentration demonstrating the absence of high-affinity specific hydration (Fig. 6).

These ratios increased gradually from 0.03, 0.2 and 0.27 in neat acetonitrile to 0.77, 1.4 and 1.15 in 100, 70 and 90% water for **PCU**, **FCU** and **TCU**, respectively (Fig. 5 & 6 and Table S2). To conclude, the sensitivity of the N^*/T^* ratio to water followed the order: **FCU** > **TCU** > **PCU** (slope comparison in Fig. 6). Comparable variations of the ratios of the intensities

were found when dioxane was used as an organic solvent (Table S2), confirming that the N^*/T^* ratio is mainly controlled by water (Fig. S8 & S9).

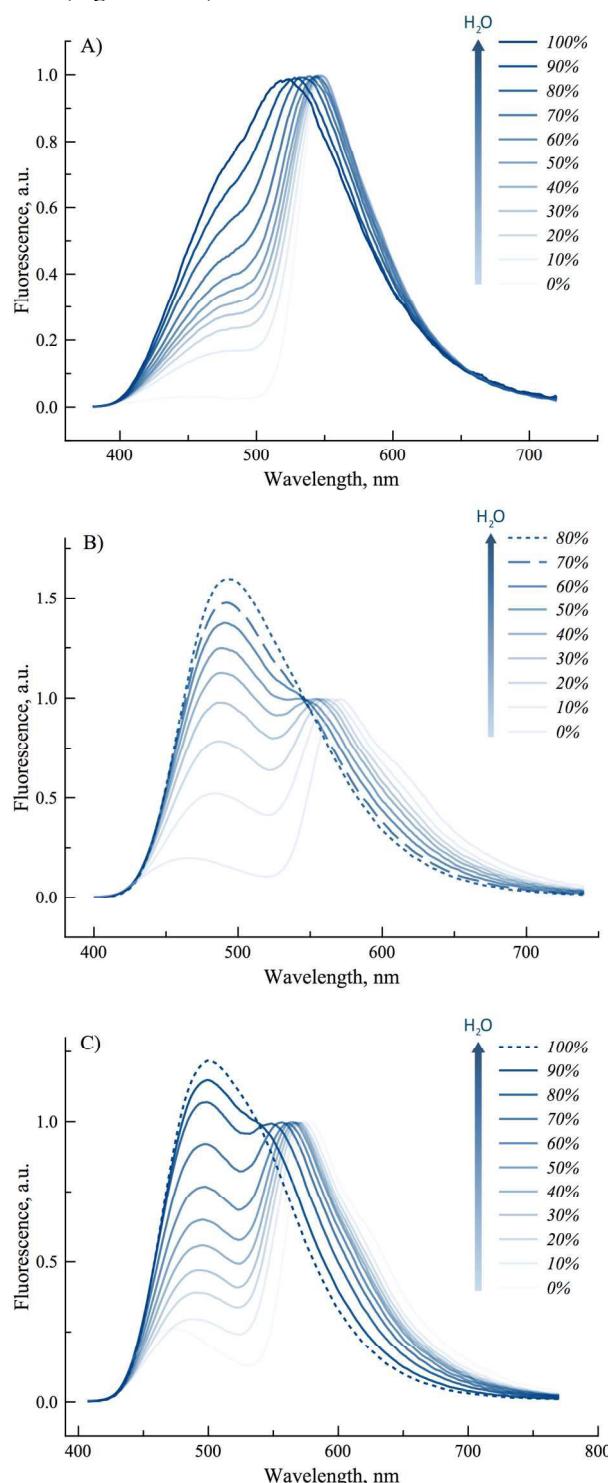


Fig. 5 Sensitivity of the emission spectra to hydration for A) **PCU**, B) **FCU** and C) **TCU**. Data were recorded in gradual mixtures of acetonitrile and water. Emission spectra were normalized by T^* band.

The remarkable sensitivity of 3HCs to hydration correlates

with previous studies carried out with derivatives of 3HCs for which a model was proposed.^{52,56,60} This latter involves specific H-bonding interactions of water with the electron rich oxygens of 4-carbonyl and phenoxide of 3HC (Scheme 5). As a result, the N* form exists in equilibrium with a protonated N*-H form, in which intermolecular H-bonding of the 4-carbonyl with water inhibits the ESIPT process.⁶¹ Thus, growing water concentration favors the N*-H form, and consequently, increases the N*/T* ratio. Moreover, hydration also favors H-bonding of the tautomer T* form, which results in its blue-shifted emission maximum.⁵²

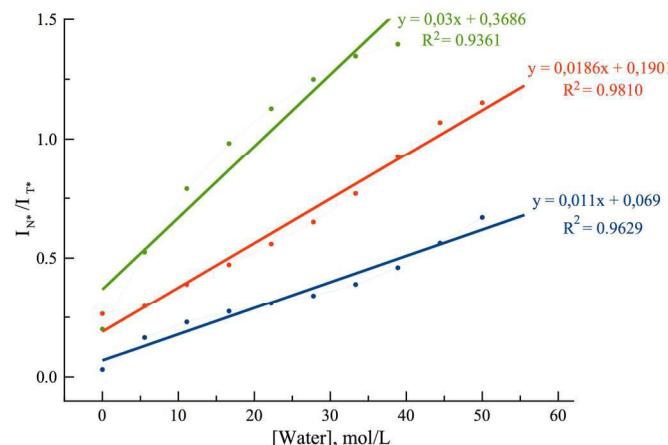
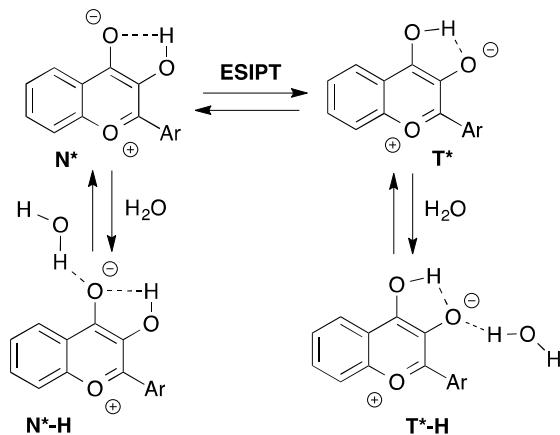


Fig. 6 Dependence of the ratiometric response (I_{N^*}/I_{T^*}) of PCU (blue), FCU (green), and TCU (orange) on the water concentration in acetonitrile. The ratios were extracted from the spectra in Fig. 4.



Scheme 5. Proposed mechanism for water sensing by 3HCs adapted from previous works.^{52,60}

Conclusions

We developed three dual emissive nucleoside derivatives based on the assembly of a 2-aryl-3HC dye and a 2'-deoxyuridine fragment, connected through a rigid and electron-conducting ethynyl linker. The three conjugates **PCU**, **FCU** and **TCU** differ from each other by the C-2 aromatic ring (phenyl, furyl and thiienyl) of the 3HC. Their photophysical properties were investigated in a wide range of solvents. The conjugation of furyl- and thiienyl-chromones with the uracil moiety resulted in

the most dramatic improvement of their spectroscopic properties.

In the present work, our results show the amplification of the brightness by about one order of magnitude due to two-fold increase of the absorbance and 4 times the enhancement of the fluorescence quantum yield, whilst retaining the most remarkable property of 3HC dyes to respond to polarity and hydration of their environment by changes in their dual emission.^{13,16} The enhancement of the quantum yield is quite unexpected, since an increase of the conjugation in the π -electronic systems generally results in an increase of the fluorescence quenching.⁶² Additionally, these derivatives keep significant quantum yields in water, which is not common among polarity-sensitive dyes^{59,63} but was observed for a series of 3HC derivatives.^{46,54} This makes our new compounds highly prospective for applications as fluorescent reporters in biological media. Since water plays a key role in controlling nucleic acid structure and function through H-bonding and electrostatic interactions, the exquisite ratiometric sensitivity of **FCU** and **TCU** to hydration makes them highly relevant for subsequent major groove labeling. Moreover, the ethynyl linker is thought to locate the dye in a precise position with respect to the DNA major axis, which is of key importance for further data interpretation of biomolecular interaction studies. Incorporation of these prospective candidates into oligonucleotides, as environmentally sensitive fluorescent (ESF) probes, is currently the focus of our ongoing researches and will be reported in due time.

Acknowledgements

We thank the ANR (ANR-12-BS08-0003-02), PACA region (DNAfix- 2014-02862 and 2014 07199) and the FRM (DCM20111223038) for financial support and a grant for N.B. We thank the French Government for the Master 2 grant to I.A.K.

Notes and references

- ^a Institut de Chimie de Nice, UMR 7272, Université de Nice Sophia Antipolis, CNRS, Parc Valrose, 06108 Nice Cedex 2, France.
- ^b Present address: Ecole Polytechnique Fédérale de Lausanne, Institute of Chemical Sciences and Engineering, Laboratory of Protein Engineering, BCH, Av. Forel 2, CH-1015 Lausanne, Switzerland.
- ^c Present address: Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo nam. 2, CZ-16610, Prague 6, Czech Republic.
- ^d A.V. Palladin Institute of Biochemistry, 9 Leontovicha Street, Kiev 01030, Ukraine.
- ^e Laboratoire de Biophotonique et Pharmacologie, UMR 7213, Faculté de Pharmacie, Université de Strasbourg, CNRS, 74 Route du Rhin, 67401 Illkirch, France.
- [†] An one-pot two-step procedure, consisting in Pd-catalyzed coupling followed the removal of the silyl group, was also attempted however a moderate yield of 64 % was obtained.
- ^{††} Concerning the furano[2,3-*d*]pyrimidin-2-one nucleoside **20**, NMR and MS analysis gave the spectroscopic characteristics consistent with a 5-*endo* cyclized derivative. Indeed, the disappearance in ¹H-NMR of the imide NH signal at 9.35 ppm was observed in favor of a new singlet at 6.92 ppm corresponding to the proton of the fused furan ring. Moreover in ¹³C-NMR, the loss of the two typical signals of the ethynyl carbons at

ARTICLE

83 and 87 ppm in favor of two new aromatic peaks respectively at 99 and 146 ppm bring the undeniable evidence of a furanopyrimidine scaffold. It is noteworthy that the deshielding (>10 ppm) of C-4 and C-5 of uracil fragment to respectively 107 ppm and 171 constitutes another proof of the proposed bicyclic aromatic structure.

Electronic Supplementary Information (ESI) available: the experimental protocols for the synthesis and characterization, the ^1H and ^{13}C NMR spectra of the compounds **2-4** and **9-24**, and the HPLC of the final products **2-4** can be consulted in the Supporting Information. See DOI: 10.1039/b000000x/

- 1 (a) J. R. Lakowicz, Principles of Fluorescence Spectroscopy, 3rd ed.; Springer: New York, 2006, p.954; (b) B. Valeur and M. N. Berberan-Santos, Molecular Fluorescence: Principles and Applications, 2nd ed.; Wiley-VCH: Weinheim, 2012, p.592; (c) A. P. Demchenko, Introduction to Fluorescence Sensing, Springer: Heidelberg, 2009, p.612; (d) A. P. Demchenko, Advanced Fluorescence Reporters in Chemistry and Biology III: Applications in Sensing and Imaging, Springer-Verlag: Berlin Heidelberg, 2011, p.352; (e) W. T. Mason and J. I. Gallin, Fluorescent and Luminescent Probes for Biological Activity, 2nd ed.; Academic Press, 1999, p.647; (f) L. Brand and M. Johnson, Methods in Enzymology: Fluorescence Spectroscopy, Vol. 450, Academic Press, 2008, p.400.
- 2 (a) M. Daniels and W. Hauswirth, *Science*, 1971, **171**, 675–677; (b) J. Pecourt, J. Peon and B. Kohler, *J. Am. Chem. Soc.*, 2000, **122**, 9348–9349; (c) E. Nir, K. Kleinermanns, A. L. Grace, M. S. de Vries, *J. Phys. Chem. A*, 2001, **105**, 5106–5110.
- 3 (a) J. Guo, J. Ju and N. J. Turro, *Anal. Bioanal. Chem.*, 2012, **402**, 3115–3125; (b) A. Nadler, J. Strohmeier and U. Diederichsen, *Angew. Chem. Int. Ed.*, 2011, **50**, 5392–5396; (c) M. E. Hawkins and F. M. Balis, *Nucl. Acids Res.*, 2004, **32**, e62–e62; (d) S. P. Sau and P. J. Hrdlicka, *J. Org. Chem.*, 2011, **77**, 5–16; (e) C. Holzhauser and H.-A. Wagenknecht, *Angew. Chem. Int. Ed.*, 2011, **50**, 7268–7272; (f) E. Socher, L. Bethge, A. Knoll, N. Jungnick, A. Herrmann and O. Seitz, *Angew. Chem. Int. Ed.*, 2008, **47**, 9555–9559.
- 4 (a) S. Kim and A. Misra, *Annual Rev. Biomed. Eng.*, 2007, **9**, 289–320; (b) K. Tainaka, K. Tanaka, S. Ikeda, K.-I. Nishiza, T. Unzai, Y. Fujiwara, A. I. Saito and A. Okamoto, *J. Am. Chem. Soc.*, 2007, **129**, 4776–4784; (c) Z.-Y. Zhao, M. San, J.-L. H. A. Duprey, J. R. Arrand, J. S. Vyle and J. H. R. Tucker, *Bioorg. Med. Chem. Lett.*, 2012, **22**, 129–132; (d) J.-L. H. A. Duprey, Z.-Y. Zhao, D. M. Bassani, J. Manchester, J. S. Vyle and J. H. R. Tucker, *Chem. Commun.*, 2011, **47**, 6629–6631; (e) K. Furukawa, M. Hattori, T. Ohki, Y. Kitamura, Y. Kitade and Y. Ueno, *Bioorg. Med. Chem.*, 2012, **20**, 16–24; (f) Zhang, M. Wang, Q. Gao, H. Qi and C. Zhang, *Talanta*, 2011, **84**, 771–776; (g) M. Hattori, T. Ohki, E. Yanase and Y. Ueno, *Bioorg. Med. Chem. Lett.*, 2012, **22**, 253–257; (h) D. W. Dodd and R. H. E. Hudson, *Mini-Reviews in Organic Chemistry*, 2009, **6**, 378–379; (i) M. L. Capobianco, A. Cazzato, S. Alesi and G. Barbarella, *Bioconjugate Chem.*, 2007, **19**, 171–177; (j) D. M. Kolpashchikov, *Chem. Rev.*, 2010, **110**, 4709–4723; (k) A. Okamoto, K. Tainaka and I. Saito, *J. Am. Chem. Soc.*, 2003, **125**, 4972–4973.
- 5 For reviews, see: (a) M. E. Hawkins, *Cell Biochem. Biophys.*, 2001, **34**, 257–281; (b) N. Dai and E. T. Kool, *Chem. Soc. Rev.*, 2011, **40**, 5756–5770; (c) Y. N. Teo and E. T. Kool, *Chem. Rev.*, 2012, **112**, 4221–4245.

6 Selected examples: (a) T. Ono, S. Wang, C.-K. Koo, L. Engstrom, S. S. David and E. T. Kool, *Angew. Chem. Int. Ed.*, 2012, **51**, 1689–1692; (b) J. W. Jung, S. K. Edwards and E. T. Kool, *ChemBioChem*, 2013, **14**, 440–444; (c) J. Riedl, P. Ménová, R. Pohl, P. Orság, M. Fojta and M. Hocek, *J. Org. Chem.*, 2012, **77**, 8287–8293; (d) J. Riedl, R. Pohl, N. P. Ernsting, P. Orság, and M. Fojta, *Chem. Sci.*, 2012, **3**, 2797–2806; (e) O. Köhler, D. V. Jarikote, I. Singh, V. S. Parmar, E. Weinhold and O. Seitz, *Pure Appl. Chem.*, 2005, **77**, 327–338.

7 (a) M. J. Davies, A. Shah and I. J. Bruce, *Chem. Soc. Rev.*, 2000, **29**, 97–107; (b) L. M. Wilhelmsson, *Q. Rev. Biophys.*, 2010, **43**, 159–183; (c) R. W. Sinkeldam, N. J. Greco and Y. Tor, *Chem. Rev.*, 2010; (d) M. Rist and J. Marino, *Curr. Org. Chem.*, 2002, **6**, 775–793; (e) J. N. Wilson and E. T. Kool, *Org. Biomol. Chem.*, 2006, **4**, 4265–4274; (f) A. A. Tanpure, M. G. Pawar and S. G. Srivatsan, *Isr. J. Chem.*, 2013, **53**, 366–378; (g) For RNA series, see: S. G. Srivatsan and A. A. Sawant, *Pure Appl. Chem.*, 2010, **83**, 213–232.

8 (a) N. J. Greco and Y. Tor, *J. Am. Chem. Soc.*, 2005, **127**, 10784–10785; (b) N. J. Greco and Y. Tor, *Tetrahedron*, 2007, **63**, 3515–3527; (c) D. Shin, R. W. Sinkeldam and Y. Tor, *J. Am. Chem. Soc.*, 2011, **133**, 14912–14915; (d) J. Riedl, R. Pohl, L. Rulíšek and M. Hocek, *J. Org. Chem.*, 2012, **77**, 1026–1044; (e) T. Pesnot, L. M. Tedaldi, P. G. Jambrina, E. Rosta and G. K. Wagner, *Org. Biomol. Chem.*, 2013, **11**, 6357–6371; (f) Y. Saito, A. Suzuki, Y. Okada, Y. Yamasaka, N. Nemoto and I. Saito, *Chem. Commun.*, 2013, **49**, 5684–5686; (g) Y. Saito, A. Suzuki, S. Ishioroshi and I. Saito, *Tetrahedron Letters*, 2011, **52**, 4726–4729; (h) A. A. Tanpure and S. G. Srivatsan, *Chem. Eur. J.*, 2011, **17**, 12820–12827; (i) A. Okamoto, K. Tainaka and Y. Fujiwara, *J. Org. Chem.*, 2006, **71**, 3592–3598; (j) N. Ben Gaiad, N. Glasser, N. Ramalanjaona, H. Beltz, P. Wolff, R. Marquet, A. Burger and Y. Mély, *Nucl. Acids Res.*, 2005, **33**, 1031–1039; (k) T. Kanamori, H. Ohzeki, Y. Masaki, A. Ohkubo, M. Takahashi, K. Tsuda, T. Ito, M. Shirouzu, K. Kuwasako, Y. Muto, M. Sekine and K. Seio, *ChemBioChem*, 2015, **16**, 167–176.

9 (a) A. P. Demchenko, *J. Fluoresc.*, 2010, **20**, 1099–1128; (b) A. P. Demchenko, *J. Mol. Struct.*, 2014, **1077**, 51–67.

10 For an example, see: L. Xu, M.-L. He, H.-B. Yang, and X. Qian, *Dalton Trans.*, 2013, **42**, 8218–8215

11 Selected examples: (a) P. A. Hopkins, R. W. Sinkeldam and Y. Tor, *Org. Lett.*, 2014, **16**, 5290–5293; (b) A. Suzuki, K. Kimura, S. Ishioroshi, I. Saito, N. Nemoto and Y. Saito, *Tetrahedron Lett.*, 2013, **54**, 2348–2352.

12 Selected example: (a) S. Sasaki, Y. Niko, A. S. Klymchenko and G.-I. Konishi, *Tetrahedron*, 2014, **70**, 7551–7559; *Applications of this concept to ODNs*, see: (b) A. Okamoto, K. Tainaka, K.-I. Nishiza and I. Saito, *J. Am. Chem. Soc.*, 2005, **127**, 13128–13129; (c) A. Suzuki, N. Nemoto, I. Saito and Y. Saito, *Org. Biomol. Chem.*, 2013, **12**, 660–666; (d) A. Suzuki, T. Yanaba, I. Saito and Y. Saito, *ChemBioChem*, 2014, **15**, 1638–1644; (e) H.-I. Un, C.-B. Huang, C. Huang, T. Jia, X.-L. Zhao, C.-H. Wang, L. Xu, and H.-B. Yang, *Org. Chem. Front.*, 2014, **1**, 1083–1090.

13 (a) A. P. Demchenko, *Trends Biotechnol.*, 2005, **23**, 456–460; (b) A. P. Demchenko, *FEBS Letters*, 2006, **580**, 2951–2957; (c) Z. Xu, L. Xu, J. Zhou, Y. Xu, W. Zhu, and X. Qian, *Chem. Commun.*, 2012, **48**, 10871–10873.

- 14 P. K. Sengupta and M. Kasha, *Chem. Phys. Lett.*, 1979, **68**, 382–385.
- 15 (a) P. T. Chou, M. L. Martinez and J. H. Clements, *J. Phys. Chem.*, 1993, **97**, 2618–2622; (b) T. C. Swinney and D. F. Kelley, *J. Chem. Phys.*, 1993, **99**, 211–221.
- 16 A. S. Klymchenko and A. P. Demchenko, *Phys. Chem. Chem. Phys.*, 2003, **5**, 461–468.
- 17 N. Chattopadhyay, M. Barroso, C. Serpa, L. G. Arnaut and S. J. Formosinho, *Chem. Phys. Lett.*, 2004, **387**, 258–262.
- 18 (a) M. Sarkar, J. G. Ray and P. K. Sengupta, *Spectrochim. Acta A*, 1996, **52**, 275–278; (b) A. S. Klymchenko and A. P. Demchenko, *Langmuir*, 2002, **18**, 5637–5639.
- 19 (a) A. S. Klymchenko, G. Duportail, Y. Mély and A. P. Demchenko, *Proc. Natl. Acad. Sci. USA*, 2003, **100**, 11219–11224; (b) A. S. Klymchenko, Y. Mély, A. P. Demchenko and G. Duportail, *Biochim. Biophys. Acta Biomembr.*, 2004, **1665**, 6–19; (c) V. V. Shynkar, A. S. Klymchenko, C. Kunzelmann, G. Duportail, C. D. Muller, A. P. Demchenko, J.-M. Freyssinet and Y. Mély, *J. Am. Chem. Soc.*, 2007, **129**, 2187–2193; (d) S. Oncul, A.S. Klymchenko, O.A. Kucherak, A. P. Demchenko, S. Martin, M. Dontenwill, Y. Arntz, P. Didier, G. Duportail and Y. Mély, *Biochim. Biophys. Acta Biomembr.*, 2010, **1798**, 1436–1443; (e) G. M'Baye, A. S. Klymchenko, D. A. Yushchenko, V. V. Shvadchak, T. Ozturk, Y. Mély and G. Duportail, *Photochem. Photobiol. Sci.*, 2007, **6**, 71–76.
- 20 V. V. Shynkar, A. S. Klymchenko, G. Duportail, A. P. Demchenko and Y. Mély, *Biochim. Biophys. Acta*, 2005, **1712**, 128–136.
- 21 (a) S. Ercelen, A. S. Klymchenko and A. P. Demchenko, *FEBS Lett.*, 2003, **538**, 25–28; (b) A. S. Klymchenko, S. V. Avilov, A. P. Demchenko, *Anal. Biochem.*, 2004, **329**, 43–57; (c) V. Y. Postupalenko, V. V. Shvadchak, G. Duportail, V. G. Pivovarenko, A. S. Klymchenko, Y. Mély, *Biochim. Biophys. Acta Biomembr.*, 2011, **1808**, 424–432; (d) V. V. Shvadchak, A. S. Klymchenko, H. de Rocquigny, Y. Mély, *Nucleic Acids Res.*, 2009, **37**, e25; (e) A. V. Strizhak, V. Y. Postupalenko, V. V. Shvadchak, N. Morellet, E. Guittet, V. G. Pivovarenko, A. S. Klymchenko, Y. Mély, *Bioconjugate Chem.*, 2012, **23**, 2434–2443.
- 22 A. S. Klymchenko, V. V. Shvadchak, D. A. Yushchenko, N. Jain and Y. Mély, *J. Phys. Chem. B*, 2008, **112**, 12050–12055.
- 23 B. Sengupta, S. M. Reilly, D. E. Davis, K. Harris, R. M. Wadkins, D. Ward, D. Gholar and C. Hampton, *J. Phys. Chem. B*, 2014, 10.1021/jp508599h.
- 24 (a) D. Dziuba, V. Y. Postupalenko, M. Spadafora, A. S. Klymchenko, V. Guérineau, Y. Mély, R. Benhida and A. Burger, *J. Am. Chem. Soc.*, 2012, **134**, 10209–10213; (b) A. A. Kuznetsova, N. A. Kuznetsov, Y. N. Vorobjev, N. P. F. Barthes, B. Y. Michel, A. Burger and O. S. Fedorova, *PLoS ONE*, 2014, **9**, e100007.
- 25 M. Spadafora, V. Y. Postupalenko, V. V. Shvadchak, A. S. Klymchenko, Y. Mély and A. Burger, *Tetrahedron*, 2009, **65**, 7809–7816.
- 26 D. Dziuba, I. A. Karpenko, N. P. F. Barthes, B. Y. Michel, A. S. Klymchenko, R. Benhida, A. P. Demchenko, Y. Mély and A. Burger, *Chem. Eur. J.*, 2014, **20**, 1998–2009.
- 27 (a) A. P. Demchenko, K.-C. Tang and P.-T. Chou, *Chem. Soc. Rev.*, 2013, **42**, 1379–1408; (b) V. I. Tomin, A. P. Demchenko, and P.-T. Chou, *J. Photochem. Photobiol. C: Photochem. Rev.*, 2015, **22**, 1–18.
- 28 R. Chinchilla and C. Najera, *Chem. Rev.*, 2007, **107**, 874–922.
- 29 (a) G. T. Hwang, Y. J. Seo and B. H. Kim, *Tetrahedron Lett.*, 2005, **46**, 1475–1477; (b) Y. Saito, K. Hanawa, K. Motegi, K. Omoto and A. Okamoto, *Tetrahedron Lett.*, 2005, **46**, 7605–7608; (c) Y. Saito, Y. Miyauchi, A. Okamoto and I. Saito, *Chem. Commun.*, 2004, 1704–1705; (d) A. Okamoto, K. Kanatani and I. Saito, *J. Am. Chem. Soc.*, 2004, **126**, 4820–4827.
- 30 J. H. Ryu, Y. J. Seo, G. T. Hwang, J. Y. Lee and B. H. Kim, *Tetrahedron*, 2007, **63**, 3538–3547.
- 31 Q. Xiao, R. T. Ranasinghe, A. M. P. Tang and T. Brown, *Tetrahedron*, 2007, **63**, 3483–3490.
- 32 L. A. Agrofoglio, I. Gilliaizeau and Y. Saito, *Chem. Rev.*, 2003, **103**, 1875–1916.
- 33 S. Gobbi, A. Rampa, A. Bisi, F. Belluti, L. Piazzì, P. Valenti, A. Caputo, A. Zampiron and M. Carrara, *J. Med. Chem.*, 2003, **46**, 3662–3669.
- 34 D. Dziuba, R. Benhida and A. Burger, *Synthesis*, 2011, 2159–2164.
- 35 (a) J.-I. Asakura and M. J. Robins, *Tetrahedron Lett.*, 1988, **29**, 2855–2858; (b) J. Asakura and M. J. Robins, *J. Org. Chem.*, 1990, **55**, 4928–4933.
- 36 (a) F. W. Hobbs Jr., *J. Org. Chem.*, 1989, **54**, 3420–3422; (b) M. J. Robins, R. S. Vinayak and S. G. Wood, *Tetrahedron Lett.*, 1990, **31**, 3731–3734.
- 37 (a) M. J. Robins and P. J. Barr, *Tetrahedron Lett.*, 1981, **22**, 421–424; (b) M. J. Robins and P. J. Barr, *J. Org. Chem.*, 1983, **48**, 1854–1862; (c) E. De Clercq, J. Descamps, J. Balzarini, J. Giziewicz, P. J. Barr and M. J. Robins, *J. Med. Chem.*, 1983, **26**, 661–666; (d) N. Esho, J.-P. Desaulniers, B. Davies, H. M. P. Chui, M. S. Rao, C. S. Chow, S. Szafert and R. Dembinski, *Bioorg. Med. Chem.*, 2005, **13**, 1231–1238.
- 38 From 5-ethynyluridines: (a) G. Luoni, C. McGuigan, G. Andrei, R. Snoeck, E. De Clercq and J. Balzarini, *Bioorg. Med. Chem. Lett.*, 2005, **15**, 3791–3796; (b) C. McGuigan, O. Bidet, M. Derudas, G. Andrei, R. Snoeck and J. Balzarini, *Bioorg. Med. Chem.*, 2009, **17**, 3025–3027.
- 39 From 5-iodouridines: (a) C. McGuigan, C. J. Yarnold and G. Jones, *J. Med. Chem.*, 1999, **42**, 4479–4484; (b) S. Srinivasan, C. McGuigan, G. Andrei and R. Snoeck, *Bioorg. Med. Chem. Lett.*, 2001, **11**, 391–393; (c) C. McGuigan, A. Brancale, G. Andrei and R. Snoeck, *Bioorg. Med. Chem. Lett.*, 2003, **13**, 4511–4513; (d) A. Brancale, C. McGuigan, G. Andrei and R. Snoeck, *Bioorg. Med. Chem. Lett.*, 2000, **10**, 1215–1217; (e) C. McGuigan, M. Derudas, M. Quintiliani, G. Andrei, R. Snoeck, G. Henson and J. Balzarini, *Bioorg. Med. Chem. Lett.*, 2009, **19**, 6264–6267; (f) A. Angell, C. McGuigan, L. Garcia Sevillano, R. Snoeck, G. Andrei, E. De Clercq and J. Balzarini, *Bioorg. Med. Chem. Lett.*, 2004, **14**, 2397–2399; (g) A. Brancale, C. McGuigan, B. Algain, P. Savy, R. Benhida, J. L. Fourrey, G. Andrei, R. Snoeck, E. De Clercq and J. Balzarini, *Bioorg. Med. Chem. Lett.*, 2001, **11**, 2507–2510.
- 40 For a review, see: (a) A. Brancale, S. Srinivasan, C. McGuigan, G. Andrei, R. Snoeck, E. De Clercq and J. Balzarini, *Antivir. Chem. Chemother.*, 2000, **11**, 383–393; (b) C. McGuigan, A. Brancale, H. Barucki, S. Srinivasan, G. Jones, R. Pathirana, S. Blewett, R. Alvarez, C. J. Yarnold, A. Carangio, A.; Velázquez, S.; Andrei, G.; Snoeck, R.; De Clercq, E.; Balzarini, J. *Drugs Fut.*, 2000, **25**, 1151–

ARTICLE

- 1161; (c) C. McGuigan, J. Balzarini, *Antiviral Res.*, 2006, **71**, 149–153.
- 41 (a) J. Haralambidis, M. Chai and G. W. Tregear, *Nucleic Acids Res.*, 1987, **15**, 4857–4876; (b) K. A. Cruickshank and D. L. Stockwell, *Tetrahedron Lett.*, 1988, **29**, 5221–5224; (c) G. T. Crisp and B. L. Flynn, *J. Org. Chem.*, 1993, **58**, 6614–6619; (d) M. J. Robins, K. Miranda, V. K. Rajwanshi, M. A. Peterson, G. Andrei, R. Snoeck, E. De Clercq and J. Balzarini, *J. Med. Chem.*, 2006, **49**, 391–398; (e) E. Petricci, M. Radi, F. Corelli and M. Botta, *Tetrahedron Lett.*, 2003, **44**, 9181–9184.
- 42 (a) N. E. Leadbeater and B. J. Tominack, *Tetrahedron Lett.*, 2003, **44**, 8653–8656; (b) J.-H. Kim, D.-H. Lee, B.-H. Jun and Y.-S. Lee, *Tetrahedron Lett.*, 2007, **48**, 7079–7084; (c) A. Tougerti, S. Negri and A. Jutand, *Chem. Eur. J.*, 2007, **13**, 666–676; (d) D. Gelman and S. L. Buchwald, *Angew. Chem. Int. Ed.*, 2003, **42**, 5993–5996; (e) O. R'kyek, N. Halland, A. Lindenschmidt, J. Alonso, P. Lindemann, M. Urmann and M. Nazaré, *Chem. Eur. J.*, 2010, **16**, 9986–9989; (f) P. Arsenyan, K. Rubina, J. Vasiljeva and S. Belyakov, *Tetrahedron Lett.*, 2013, **54**, 6524–6528. (g) P. Y. Choy, W. K. Chow, C. M. So, C. P. Lau and F. Y. Kwong, *Chem. Eur. J.*, 2010, **16**, 9982–9985.
- 43 E. S. Kumarasinghe, M. A. Peterson and M. J. Robins, *Tetrahedron Lett.*, 2000, **41**, 8741–8745.
- 44 A. S. Klymchenko, V. G. Pivovarenko and A. P. Demchenko, *Spectrochim. Acta Part A*, 2003, **59**, 787–792.
- 45 C. Reichardt, *Chem. Rev.*, 1994, **94**, 2319–2358.
- 46 A. S. Klymchenko and A. P. Demchenko, *New J. Chem.*, 2004, **28**, 687–692.
- 47 A. S. Klymchenko, T. Ozturk, V. G. Pivovarenko and A. P. Demchenko, *Can. J. Chem.*, 2001, **79**, 358–363.
- 48 C. Boudier, A. S. Klymchenko, Y. Mély and A. Follenius-Wund, *Photochem. Photobiol. Sci.*, 2009, **8**, 814–821.
- 49 (a) S. M. Ormson, R. G. Brown, F. Vollmer and W. Rettig, *J. Photochem. Photobiol. A: Chem.*, 1994, **81**, 65–72; (b) O. A. Kucherak, L. Richert, Y. Mély and A. S. Klymchenko, *Phys. Chem. Chem. Phys.*, 2012, **14**, 2292–2300.
- 50 W. H. Melhuish, *J. Phys. Chem.*, 1961, **65**, 229–235.
- 51 (a) E. Z. Lippert, *Z. Naturforsch. A*, 1955, **10**, 541–545; (b) N. Mataga, Y. Kaifu, and M. Koizumi, *Bull. Chem. Soc. Jpn.*, 1956, **29**, 465–470.
- 52 C. A. Kenfack, A. S. Klymchenko, G. Duportail, A. Burger and Y. Mély, *Phys. Chem. Chem. Phys.*, 2012, **14**, 8910–8918.
- 53 C.-C. Hsieh, C.-M. Jiang and P.-T. Chou, *Acc. Chem. Res.*, 2010, **43**, 1364–1374.
- 54 O. M. Zamotaiev, V. Y. Postupalenko, V. V. Shvadchak, V. G. Pivovarenko, A. S. Klymchenko and Y. Mély, *Bioconjugate Chem.*, 2010, **22**, 101–107.
- 55 (a) A. S. Klymchenko, V. G. Pivovarenko, T. Ozeturk and A. P. Demchenko, *New J. Chem.*, 2003, **27**, 1336–1343; (b) V. V. Shynkar, Y. Mely, G. Duportail, E. Piémont, A. S. Klymchenko and A. P. Demchenko, *J. Phys. Chem. A*, 2003, **107**, 9522–9529.
- 56 A. J. G. Strandjord and P. F. Barbara, *J. Phys. Chem.*, 1985, **89**, 2355–2361.
- 57 (a) K. Kipper, K. Herodes and I. Leito, *J. Chrom. A*, 2011, **1218**, 8175–8180; (b) K. Kipper, K. Herodes, I. Leito and L. Nei, *Analyst*, 2011, **136**, 4587–4594.
- 58 R. A. Moore, J. Lee and G. W. Robinson, *J. Phys. Chem.* 1985, **89**, 3648–3654.
- 59 (a) For instance, see 2,6-diaminopurine: D. C. Ward, E. Reich and L. Stryer, *J. Biol. Chem.*, 1969, **244**, 1228–1237; (b) For a recent example of a bright solvatochromic dye in water, see: Y. Niko, Y. Cho, S. Kawauchi and G.-I. Konishi, *RSC Adv.*, 2014, **4**, 36480–36484.
- 60 (a) V. G. Pivovarenko, O. M. Zamotaiev, V. V. Shvadchak, V. Y. Postupalenko, A. S. Klymchenko and Y. Mély, *J. Phys. Chem. A*, 2012, **116**, 3103–3109; (b) R. Das, A. S. Klymchenko, G. Duportail and Y. Mély, *Photochem. Photobiol. Sci.*, 2009, **8**, 1583–1589.
- 61 (a) V. V. Shynkar, A. S. Klymchenko, E. Piémont, A. P. Demchenko and Y. Mély, *J. Phys. Chem. A*, 2004, **108**, 8151–8159; (b) W. Caarls, M. S. Celej, A. P. Demchenko and T. M. Jovin, *J. Fluoresc.*, 2010, **20**, 181–190.
- 62 K. Rurack and U. Resch-Genger, *Chem. Soc. Rev.*, 2002, **31**, 116–127.
- 63 A. P. Demchenko, *Anal. Biochem.*, 2005, **343**, 1–22.

PUBLICATION 4

Une fois les nucléosides caractérisés, il nous a été possible de nous consacrer à la synthèse d'ODNs incorporant les analogues nucléosidiques sélectionnés. Pour atteindre cet objectif nous avons dans un premier temps travaillé avec un composé modèle, l'analogique 3-méthoxychromone (MTCU). Le groupement méthoxy porté par la chromone est stable et facile à manipuler. Il nous a ainsi permis d'optimiser les différentes étapes de la synthèse du phosphoramidite MTCU **8** [cf. Figure 8] jusqu'à son incorporation dans des séquences modèles. Ces travaux nous ont également permis de vérifier la stabilité de la sonde dans les conditions de déprotection classiques des ODNs (ammoniaque 30% - 12 heures).

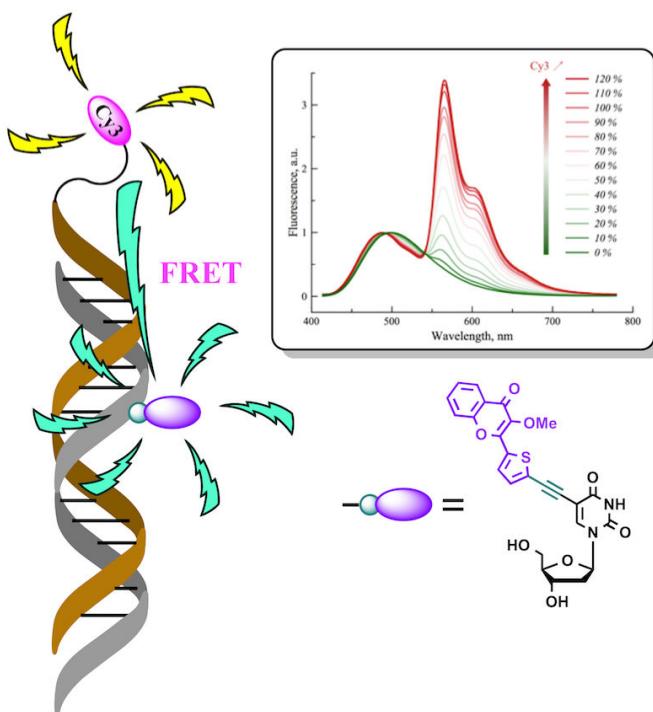


Figure 8. Dérivé Méthoxychromone (MTCU) et son utilisation dans un transfert d'énergie (FRET) avec la sonde Cy3 (*courbes normalisée sur la bande MTCU*).

Les différentes caractérisations photophysiques de ces ODNs ont permis de déterminer que MTCU présentait une absorption dans le domaine visible, un coefficient d'absorption molaire important, des rendements quantiques appréciables (multipliés par deux après hybridation : de 20-30 % dans les simples brins et de 40 à 50 % dans les duplexes) et une photostabilité très importante (duplexes appariés). Tous ces paramètres prometteurs ont permis de concevoir un couple FRET efficace, où notre fluorophore méga-Stokes agit comme donneur et où Cy3 ($\epsilon = 136\,000\text{ L}\cdot\text{mol}^{-1}\cdot\text{cm}^{-1}$; $\lambda_{\text{Abs}} = 547\text{ nm}$, $\lambda_{\text{Em}} = 570\text{ nm}$; $\Phi = 0,15$), attaché par un bras espaceur flexible en 5' sur le brin complémentaire, joue le rôle d'accepteur. Les études photophysiques ont également démontré une efficacité de transfert d'énergie importante et une réponse ratiométrique entraînant un net changement visuel de la couleur d'émission (du bleu ciel : simple brin MTCU au jaune : duplexe MTCU-Cy3). Ces travaux ont abouti à la rédaction d'un article qui sera soumis dans un journal à comité de lecture dans les prochains jours.

Design and development of a two-color emissive FRET pair based on a photostable fluorescent deoxyuridine donor presenting mega-Stokes shifts

Nicolas P. F. Barthes,^a Krishna Gavvala,^b Dmytro Dziuba,^a Dominique Bonhomme,^a Anne Sophie Dabert-Gay,^c Delphine Debayle,^c Benoît Y. Michel,^{*a} Yves Mély,^b Alain Burger^{*a}

a) Institut de Chimie de Nice, UMR 7272, Université de Nice Sophia Antipolis, CNRS, Parc Valrose, 06108 Nice Cedex 2, France

b) Laboratoire de Biophotonique et Pharmacologie, UMR 7213, Faculté de Pharmacie, Université de Strasbourg, CNRS, 74 Route du Rhin, 67401 Illkirch, France

c) Institut de Pharmacologie Moléculaire et Cellulaire, UMR 6097, Université de Nice Sophia Antipolis, 660 Route des Lucioles, 06560 Valbonne, France

ABSTRACT

We report the synthesis and the site-specific ODN incorporation of an emissive deoxyuridine analog electronically conjugated on its C5-position with a 3-methoxychromone moiety acting as a fluorophore. The corresponding labeled sequences exhibited remarkable enhanced photostabilities and improved quantum yields that were up to 2-fold larger after incorporation in DNA and hybridization. The resulting biosensor was successfully employed as a mega-Stokes shift donor in a FRET pair with a Cy3 acceptor displaying a yellow emission combined with a ratiometric response of wide amplitude.

INTRODUCTION

Applications of Förster Resonance Energy Transfer (FRET) have become an inescapable technique in biology and biophysics for determining structures, interactions and affinities between different partners from single molecules to cell imaging.¹ This process proved to be of particular great interest in DNA research field due to its large domain of applications as for instance in hybridization technology and DNA-DNA or DNA-protein interaction studies.² On the other hand, DNA offers a unique platform for the design and construction of optimized emissive donor and acceptor dye pairs since its structure and conformation can be easily controlled.³ One emerging and very attractive design involves for the donor, a dye exhibiting large Stokes shift (i.e. >100

nm). From a general point of view, such fluorophores, named mega-Stokes shift dyes, turn out to be very appealing for FRET applications since they prevent self-quenching, light scattering and thus reduce the signal background in imaging.⁴ The prototype of mega-Stokes shift dye is a push-pull dye. It exhibits a larger dipole in the excited state due to intramolecular charge transfer (ICT) and a bathochromic shift of the emission band with increasing solvent polarity. For biological applications, such a fluorophore should meet the following additional desired properties: absorption in the visible (>400 nm), high molar absorption coefficient (>30 000 M⁻¹.cm⁻¹), good fluorescence quantum yield (>40 %) and photostability. However, dyes that fulfil all these requirements are seldom encountered when incorporated in DNA.⁵ Indeed

insufficient quantum yields pose problems for fluorophores showing very large Stokes shifts because such dyes are usually quenched in aqueous media (e.g. pyridinium,⁵ 4-aminophthalimide,^{6a,b} biaryllic 5-aminobenzoaxazole,^{6c} DAN family members:^{6d,e} 6-(dimethylamino)-2-naphthalenecarboxamide,^{6d,e} 2-ethynyl-6-cyanonaphthalene,^{6f} 5-amidodansyl^{6g}).[‡]

Recently, we reported that the conjugate **1**, compiling a 3-methoxychromone electronically coupled to a dimethyluracil scaffold through a rigid ethynyl linker, absorbed strongly in the violet ($35.000 \text{ M}^{-1} \cdot \text{cm}^{-1}$), emitted light with large Stokes-shift ($\Delta\lambda = 100-135 \text{ nm}$) and exhibited appreciable quantum yields in the range of 10 to 53% in water and protic solvents.⁷ Anchoring the donor to the nucleobases through a short and rigid linker presents the advantage to control the orientation and position of the emissive dipole as well as the distance separating it from the acceptor.

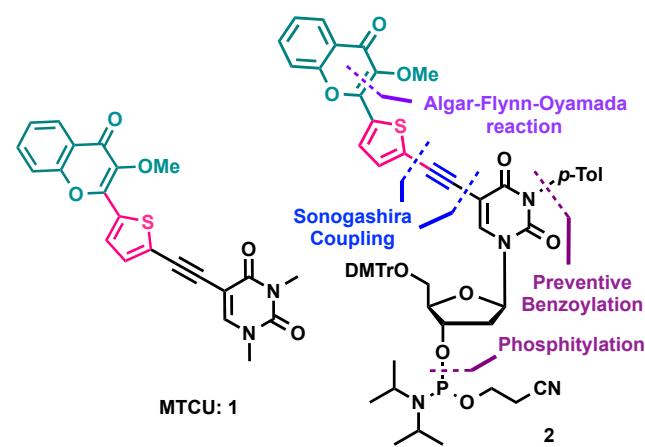


Figure 1. Fluorescent pyrimidine model **1** and retrosynthetic analysis of the targeted deoxyuridine analog incorporating methylated 3-hydroxychromone (3HC) fluorophore as a natural nucleobase modifier **2**.

In this context, our research focused on the development of an efficient FRET pair as a ratiometric emitter, for sensing interactions involving nucleic acids, by combining the fluorescent uracil

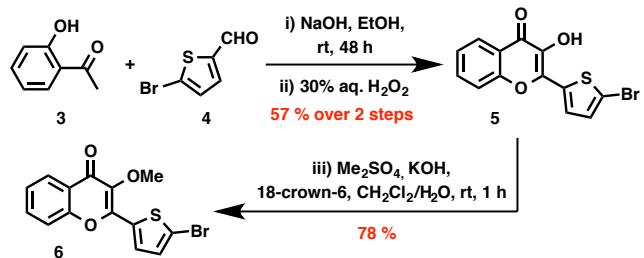
derivative as a mega-Stokes shift donor with a 5'-terminal Cy3-acceptor.⁸ Herein, we report the synthesis of the corresponding fluorescent ethynyldeoxyuridine amidite integrating the conjugated 3-methoxychromone fluorophore on its C5-position, its incorporation in DNA and characterization of the labeled DNA. We proved its site-specific incorporation in ODNs and demonstrated that the important properties desired for a mega Stokes shift dye are satisfied for FRET applications.⁹

RESULTS AND DISCUSSION

Retrosynthesis: The target fluorescent amidite **2** was obtained through a convergent strategy involving Sonogashira cross-couplings and Algar–Flynn–Oyamada reaction (Figure 1). As reported for the preparation of 3HC conjugated deoxyuridine nucleosidic series,¹⁰ the synthetic approach is based on an unusual final assembly. Indeed, face to competitive homo- and cross-couplings difficulties, it was found to be definitely more favorable in terms of yield and purification to couple the electron-rich 5-ethynyldeoxyuridine with the electron-poor bromothienylchromones. By this strategy, it provided an access to **2** on a gram-scale via a 12-step pathway.

Synthesis: The preparation of the 3HC fragment, required for the final cross-coupling between the fluorophore and the nucleoside moieties, is depicted in Scheme 1. It involves a smooth and efficient Algar-Flynn-Oyamada process consisting in a chalcone formation followed by an oxidative cyclization. For that purpose, starting from 2-hydroxy-acetophenone **3**, a Claisen-Schmidt condensation with carbaldehyde **4** in the presence of sodium hydroxide in ethanol followed by a treatment with aqueous hydrogen peroxide provided the 2-bromothienyl-3-hydroxychromone **5** in a satisfying over two-step yield (57%).¹¹ By using phase-transfer

conditions (PTC), the reactive 3-hydroxyl group was then converted as a methyl ether, facilitating significantly the purification stages of all the downstream synthesis.¹²

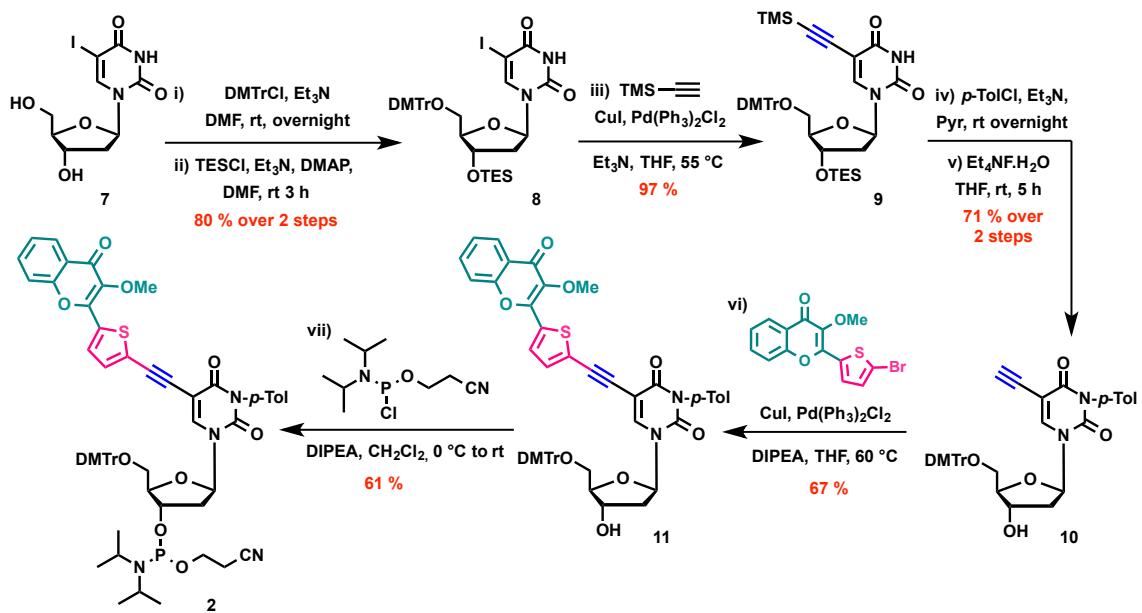


Scheme 1. Synthetic preparation of the methylated 3HC fluorescent coupling partner **6**.

Concerning the preparation of the nucleosidic fragment required for the final assembly, we started from the readily available 5-iodo-deoxyuridine **7** (Scheme 2).¹³ Its 5'-hydroxyl group was first protected as an acid-labile trityl ether whereas the 3'-OH was consecutively masked by a triethylsilyl ether, using standard conditions for both reactions. Sonogashira coupling of the iodo intermediate **8** with TMS-acetylene furnished the protected ethynyl derivative **9** in high yields. We recently reported that a preventive toluoylation of the *N*3-imide of uracil moiety avoids competitive 5-*endo*-dig-cyclization during the further Pd-catalyzed coupling.^{10,14} To set up this base labile imide protecting group, mild conditions (*p*-toluoyl chloride, Et_3N in pyridine) were employed.¹⁵ A subsequent treatment with the softly nucleophilic tetraethylammonium fluoride monohydrate enabled to concomitantly cleave the TMS and TES silyl groups providing the terminal alkyne **10** in satisfactory yields (71% over 2 steps, Scheme 2). A second Sonogashira reaction was then efficiently used for that purpose. Thus, when using a standard catalytic system ($\text{PdCl}_2(\text{PPh}_3)_2$, CuI , Et_3N in THF), the reaction proceeded cleanly to give the coupled derivative **11**.[#] The desired fluorescent amidite **2** was then obtained pure in acceptable yields under classical conditions. Fluorescent labeled

15-mer ODNs, incorporating site-selectively the amidite **2**, were efficiently obtained by solid-phase synthesis using standard procedures (SI). The sequences $d(\text{CGT TTT XMX TTT TGC})$, where **X** = **T**, **A**, **C** or **G**, were tagged with the conjugate **2** at the positions labeled **M**. The other ODNs required for the study were supplied from commercial sources. Complementary ODNs $d(\text{GCA AAA YZY AAA ACG})$, where **Y** and **Z** = **T**, **A**, **C** or **G**, were annotated in italics. $d(\text{Cy3GCA AAA YZY AAA ACG})$ was labeled with Cy3 at the terminal 5' position. Control experiments were conducted with wild-type ss strands $d(\text{CGT TTT XTX TTT TGC})$ as follows where **T** replaces **M**. MS, HPLC & UV analysis confirmed the veracity and purity of the ODNs (Figures S1 to S5; Table S1).

Photophysics: The labeled ODNs were chosen as model sequences to characterize the physical properties of the fluorophore **1** with respect to the nature of the flanking bases in single- (ss) and enabled to design and construct double-strands (ds) DNA (Table 1 & SI). The fluorescent emitter was located in the middle of the sequence rather than at the 5'-terminal positions of duplexes in order to avoid a sticky end association that could complicate the interpretation of the emission response.¹⁸ Melting temperature and CD experiments were first realized with the matched and mismatched labeled ds DNA, and the corresponding wild type sequences for comparison (Tables 1 and S2, Fig. S6 & S7). Thermal denaturation studies clearly demonstrate that the T emissive analogue preferentially base paired with A (Table S2, ΔT_m = +8.0 to +5.5 °C compared to mismatched pairs). Compared to the matched wild-type duplexes, the incorporation of **2** revealed to be only somewhat destabilizing (ΔT_m = -4.0 to -5.9 °C, Table S2, Figures S6 & S7) but less than a mismatch (ΔT_m = -8.2 to -10.5 °C, Table S2). Furthermore incorporation of the dye did not affect the secondary structure of the labeled DNA



Scheme 2. Synthesis of the modified deoxyuridine amidite **2** bearing a 3-methoxychromone scaffold as a fluorescent reporter.

Table 1. Spectroscopic data for labeled ODNs.^a

DNA	ss & ds samples	T _m (°C) ^b	λ _{abs} (nm) ^c	λ _{Em} (nm) ^d	Φ (%) ^e
1	TMT	-	395	500	30
2	TMT-AAA	45.5 (49.5)	392	490	50
3	AMA	-	399	498	21
4	AMA-TAT	41.4 (47.3)	392	491	47
5	CMC	-	398	500	27
6	CMC-GAG	46.5 (51.3)	396	497	7
7	GMG	-	396	497	2
8	GMG-CAC	50.1 (54.8)	396	495	1
9	AAA(Cy3)	-	548	565	18
10	TTT-AAA(Cy3)	51.9 (49.5)	548	564	20
11	TMT-AAA(Cy3)	47.5 (49.5)	393	564	-

a) Recorded at 20 °C with 2 μM of each ODN in buffer pH 7.0 (10 mM cacodylate buffer, 150 mM NaCl); b) Melting temperature (T_m) of the labeled matched duplex; T_m of the corresponding wild-type duplexes are listed in parentheses; ± 0.5°C; c) Position of the absorption maximum; d) Position of the emission maximum; e) Quantum yields determined by using quinine sulfate (QS) in 0.1 M HCl solution ($\lambda_{\text{ex}} = 350$ nm, $\Phi = 0.54$)¹⁶ and *p*-dimethylaminoflavone (dMAF) in EtOH ($\lambda_{\text{ex}} = 404$ nm, $\Phi = 0.27$)¹⁷ as standard references; ± 10% mean standard deviation.

duplexes as evidenced by CD spectra, which confirmed a typical B-form signature (Figure S8). The UV absorbance and fluorescence spectra of **1** in **DNA1-8** were recorded in phosphate buffer pH 7.0 (Figures S10, S12). **DNA1-8** showed absorption maxima in the range of 394-399 nm (Tables 1), indicating that the absorbance of the dye was marginally affected by the nature of the flanking bases in both ss and fully complementary ds DNA. Independently of the adjacent bases, the fluorescence emission maxima were located in the green region (ca. 500 nm), conferring to the dye incorporated in DNA the expected mega-Stokes shift (98-105 nm, Table 1). Remarkably, compared to the quantum yield reported for **1** in H₂O (10 %), the fluorophore displayed 2 to 5 fold enhanced quantum yields in **DNA1-5** (ca. 21-50 %). One hypothesis could be that rotational motion was reduced and therefore non-radiative dissipation of energy was lowered due to stacking and hydrophobic interactions of the fluorophore with the flanking bases.¹⁹ By contrast to **A**, **T** and **C** bases, the fluorophore is quenched by proximal **G**. **G** is a well

known and efficient quencher of fluorescent nucleoside analogues in ODNs.²⁰

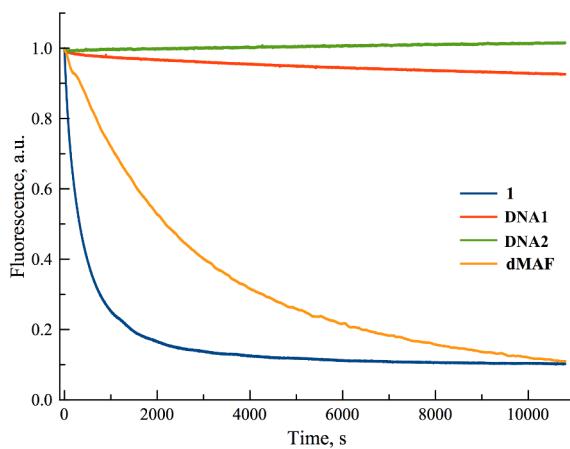


Figure 2. Photodegradation curves of fluorophore **1** in water (blue), ss **DNA1** (red) and ds **DNA2** (green) in buffer pH 7, and dMAF (yellow) in EtOH. Concentration of each sample was 2 μ M. Excitation was at 400 nm.

Photostability of fluorescent dyes is a major concern since it represents a severe limitation for their use in imaging applications and single molecule fluorescence spectroscopy. We therefore investigated the time-dependence photodegradation of the free dye **1** and its incorporated form in ss and matched ds **DNA1-2** (Figure 2) under continuous illumination of the sample. For comparison, the *p*-dimethylamino-3-hydroxyflavone (dMAF), a known unstable push-pull fluorophore was also studied.²¹ The curves were fitted to a monoexponential decay to extract the time constants of photodegradation. Dyes **1** and dMAF were found to be the less stable compounds ($t_{1/2} = 8$ and 56 min, respectively). Therefore the introduction of the methoxyl group at position 3 of chromone to give **1** did not prevent photodegradation as it was observed for dMAF. This result indicates that **1** follows a different photodegradation pathway than the methoxy derivative of dMAF.²² On the other hand, incorporation of the fluorophore in ss **DNA1** produced a dramatic enhancement of photostability

($t_{1/2} = 55$ h vs. 8 min).²³ Furthermore, when the dye was exposed to the major groove as in the ds matched **DNA2**, no degradation at all was noticed over a period of 3 hours under the photobleaching conditions used. This remarkable improvement of photostability of the fluorophore without oxygen depletion and adding additives may be due to a self-healing mechanism.²⁴ Herein, DNA may participate as a triplet state quencher producing less singlet oxygen and consequently less photo-oxidation. Less populated triplet state is also consistent with the observed improved quantum yields when the fluorophore was coupled to ss and ds **DNA1-2**.

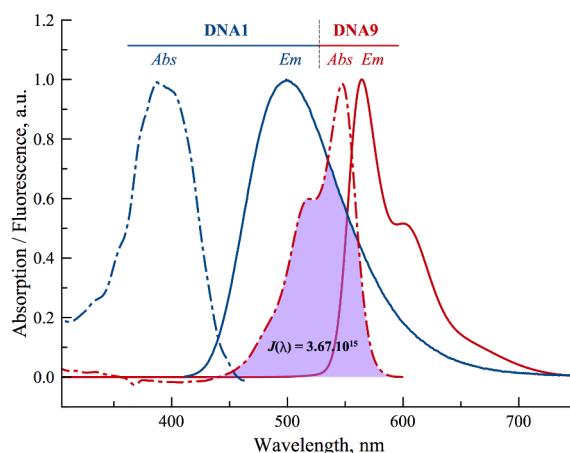


Figure 3. Absorption (dashed line) and emission (continuous line) spectra of ss **DNA1** (blue) and **DNA9** (red). Spectral overlap of the FRET emissive pair considered in this study was colored in purple. Absorption and emission of **DNA1** & **DNA9** were normalized on their respective maximum.

To sum up, our data demonstrate that the dye fulfils the desired requirements to be a prospective candidate as a mega-Stokes shift FRET donor.²⁰ Additionally, since its shifted emission ($\lambda_{em} = 490$ -498 nm) revealed to be fairly solvatofluorochromic independently of the flanking bases, it provides the opportunity to develop a ratiometric probe for sensing interactions via a radiative energy transfer. To verify this hypothesis, we selected the sequence

DNA1 since it gave after hybridization with its complementary strand the brightest combination of ss and ds DNA (Table 1, **DNA1-2**: $\Phi = 0.30$ and 0.50). The common Cy3 was selected as a FRET acceptor and was introduced in the complementary sequence at its 5'-terminal (**DNA9** - Figure S16). Cy3 was chosen because the large molar absorptivity of Cy3 ($\epsilon_{547\text{nm}} = 13.6 \times 10^4 \text{ m}^{-1} \cdot \text{cm}^{-1}$) and the significant spectral overlap integral ($J(l) = 3.7 \times 10^{15} \text{ m}^{-1} \cdot \text{cm}^{-1} \text{nm}^4$) between the donor emission and the acceptor absorption spectra are mandatory properties for efficient radiative energy transfer (Figure 3). Furthermore, its absorption band centered at 548 nm is well-separated from the donor's one ($\Delta\lambda = 156 \text{ nm}$) and the emission located in the yellow range (565 nm) is combined with an appreciable quantum yield ($\Phi = 0.20$).²⁶ Additionally, the fact that Cy3 is almost not absorbing ($\epsilon_{390 \text{ nm}} < 300 \text{ m}^{-1} \cdot \text{cm}^{-1}$) at the wavelength the 3HC donor will be excited, represents a considerable asset since it will dramatically increase the S/N ratio due to the fact that direct excitation of the acceptor will be negligible.

The control and FRET experiments were conducted with the single labeled ds **DNA10** and double-labeled ds **DNA11** obtained after annealing **DNA9** with the fully complementary wild type and **DNA1** strands, respectively. When **DNA10** was excited at 392 nm no emission was registered as shown in the Figure 4.

This result is consistent with the fact Cy3 does not absorb at this wavelength. By contrast, when the duplex was excited at the absorption maximum of Cy3 (548 nm) the characteristic emission spectrum of Cy3 was observed (Figures 4). When it was irradiated at one quarter of its absorption maximum at 490 nm, the shape of the emission did not change but the fluorescence intensity appeared to be one quarter of its maximum as expected (Figure 4). The evidence for efficient FRET was provided when the double-labeled **DNA11** was excited at the absorption

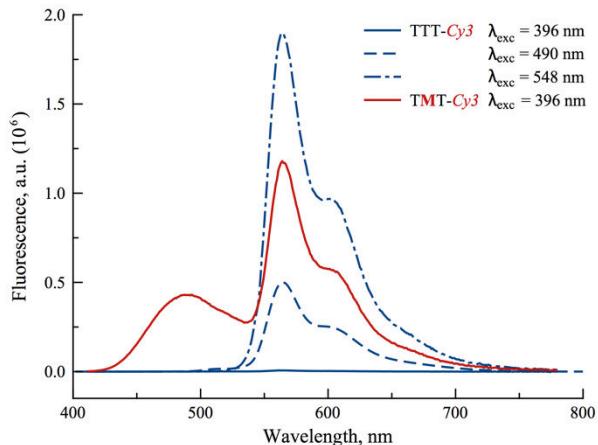


Figure 4. Emission spectra of Cy3 labeled **DNA10** (blue) and Cy3-**M** double-labeled **DNA11** (red) duplexes at different excitation wavelengths at pH 7.0.

maximum of the donor (393 nm). Indeed by comparison with the single labeled ds **DNA10** and **DNA2**, a marked increase (over a factor >100) of the Cy3 emission intensity at 564 nm together with a marked decrease of the emission intensity of the donor **M** at 490 nm were observed. The efficiency of energy transfer was found to be 92 % for **DNA11** (see SI, Eq1). Further support for FRET was furnished from the excitation spectra of the single and double-labeled **DNA10-11** (Figure S14). By collecting photons at 564 nm, only the spectrum registered from the double-labeled **DNA11** showed a band centered at 393 nm that perfectly fitted with the absorption spectrum of the donor (Figures S11 and S14). Since the donor emission locates in the blue edge and is well separated from the acceptor emission band, it allows easy ratiometric detection. Thus, by gradually increasing the concentration of **DNA9**, the energy transfer inherent with the formation of the duplex was probed by a ratiometric two-color response (Figure 5 & S13). As predicted, addition of 1 equivalent of **DNA9** was sufficient for completely inverting the intensity ratio, determined at the two emission maxima, from 0.7 to 3.4 (Figures 4 & S15^{##}). Indeed, a 10% excess of **DNA9** did not lead to further change in the ratio of intensities of the

two emission bands (saturation), thereby supporting that the observed FRET is a 1:1 stoichiometric process. Interestingly, the changes in emission color of **DNA11** from sky blue to yellow offer a clear visualization of hybridization useful for two-color imaging applications (Figure 5). The Förster distance R_o is an important physical data that characterizes a FRET pair and is informative at which distance FRET can be detected or not. Incomplete labeling is one source of error in calculating R_o ; but can be excluded because HPLC and MS analysis ascertained the purity and integrity of the labeled ss DNA and fluorescence titration supported a 1:1 stoichiometry for hybridization of the two strands (Figure 5 and SI). Another source of uncertainty is the orientational dipole factor κ^2 .²⁸ κ^2 of 2/3 is used for randomly oriented donor and acceptor partners. However, applying this value for calculating R_o with Cy3 poses question because it is known that a major fraction of Cy3 is not free but stacks with the bases at the end of the duplex.²⁹ Denaturation studies (T_m) agrees with the fact that Cy3 stacks since it conferred a 2.0 °C extra stabilization to the duplexes (compare single- **DNA10** and the double-labeled **DNA11** with the wild type and **DNA2**, respectively, see Table 1).^{20,27} To calculate R_o we envisioned the following two extreme situations. First with free Cy3, 55 Å for R_o was found when $\kappa^2 = 2/3$ was applied. Second with stacked Cy3, the distance separating the two dyes in **DNA11** was estimated to be about 27 Å for a 7 base pair separation in length based on a 3.4 Å rise per base pair of B-DNA and a stacked Cy3. CD confirmed the B-DNA conformation of **DNA11** (SI). When 27 Å and 92% for energy transfer efficiency were taken into account to calculate R_o , 41 Å was obtained (see SI, Eq5). Therefore R_o is likely within these two extreme values (41-55 Å: 25% variation). Nevertheless, it is in the range of Förster distances of FRET pairs composed of organic dyes, which are convenient for biological applications.^{1a}

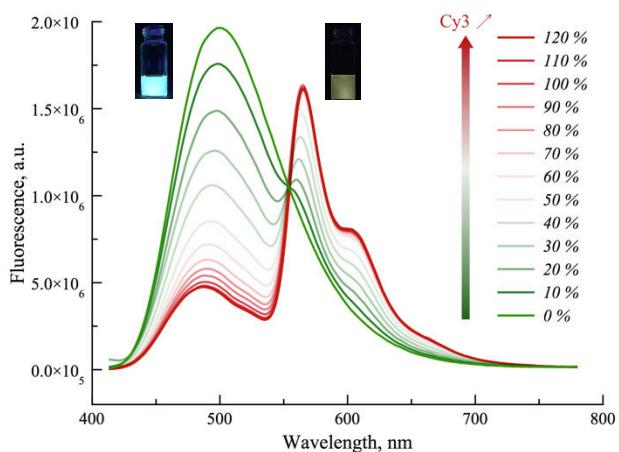


Figure 5. Fluorescence titration of **DNA1** with **DNA9** – (from 0 to 1.2 equivalents) monitored by ratiometric emission changes. Insets showed the color changes before (left: blue cell) and after addition of 1 eq. of **DNA9** (right: yellow cell).

Conclusion

In summary, we reported the convergent synthesis and spectroscopic characterizations of 15-mer sequences site-specifically labeled with a single-band emissive deoxyuridine analogue. Based on the 2-thienyl-3-methoxychromone scaffold tethered to the C-5 position of an ethynyl uracil, the fluorophore incorporated in DNA presents the exquisite qualities of displaying a visible absorption (≈ 400 nm), large absorption coefficient ($35.000\text{ M}^{-1}\cdot\text{cm}^{-1}$), emission of light with large Stokes shifts (≈ 100 nm), enhanced quantum yields (up to 50%) and remarkable improved photostability. Taken together, the promising parameters of the dye enabled to design and construct an efficient FRET pair where the mega-Stokes shift dye of the first strand acts as a donor whereas a 5'-Cy3 in the complementary strand plays the acceptor role. Fluorescence studies have demonstrated both a remarkable energy transfer efficiency and ratiometric response yielding to a net visual change in emission color upon hybridization switching its emission from cyan to yellow. Finally, since the deoxyuridine conjugate absorbs in the violet (ca. 400 nm) and the emission locates in the

green region (ca. 500 nm), its excitation using the common violet laser line and combination with a orange fluorescent protein³⁰ as the second FRET partner open prospective applications for two-color imaging in the near future.

MATERIAL AND METHODS

Supplementary data: experimental protocols and spectroscopic characterizations. This material is available free of charge via the Internet at <http://pubs.acs.org>."

NOTES

All authors have given approval to the final version of the manuscript.

‡ In water, strong ICT dyes usually follows that trend: the larger the Stokes shift, the lower the quantum yield.

†† Independently of the catalytic system being used, the sterically hindered Glaser coupling product was observed as a more polar spot by TLC (ca. 10-20 %).

††† S/N ratio detection of **DNA1** over the titration is ca. 7.

ACKNOWLEDGMENTS & FUNDING

We thank Andrey S. Klymchenko for valuable discussions.

We thank the ANR (ANR-12-BS08-0003-02), région PACA (DNAfix- 2014-02862 and 2014 07199) and the FRM (DCM20111223038) for financial support and a grant for N.P.F.B. We thank the French Government for the Master 2 grant to I.A.K. and Campus France for the Raman-Charpak fellowship to K.G.

REFERENCES

- (1) (a) Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*, 3rd ed.; Springer: New York, 2006; Chapter 13. (b) Valeur, B.; Berberan-Santos, M. N.; *Molecular Fluorescence: Principles and Application*, 2nd ed.; Wiley-VCH: Weinheim, 2012, p. 592. (c) Demchenko, A. P. *Introduction to Fluorescence Sensing*, Springer: Heidelberg, 2009, p. 612.
- (2) (a) Demchenko, A. P. *Advanced Fluorescence Reporters in Chemistry and Biology III: Applications in Sensing and Imaging*, Springer-Verlag: Berlin, Heidelberg, 2011, p. 352. (b) Mason, W. T.; Gallin, J. I. *Fluorescent and Luminescent Probes for Biological Activity*, 2nd ed.; Academic Press: London, 1999, p. 647. (c) Brand, L.; Johnson, M. *Methods in Enzymology: Fluorescence Spectroscopy*, Academic Press: San Diego, 2008, 450, p. 400.
- (3) (a) *General review on fluorescence techniques*: Ishikawa-Ankerhold, H. C.; Ankerhold, R.; Drummen, G. P. C. *Molecules*, **2012**, *17*, 4047. (b) *FRET applied to DNA, for a review*: Didenko, V. V. *Biotechniques* **2001**, *31*, 1106–1121; (c) Selected example on reversible RET: Bälter, M.; Hammarson, M.; Remón, P.; Li, S.; Gale, N., Brown, T.; Andréasson, J. *J. Am. Chem. Soc.* **2015**, *137*, 2444.
- (4) *FRET pair using a Mega-Stokes shift dye as a donor (not applied to DNA labeling)*: (a) Nagy, K.; Orbán, E.; Bosze, S.; Kele, P. *Chem.–Asian. J.* **2010**, *5*, 773. (b) Cserép, G. B.; Enyedi, K. N.; Demeter, A.; Mező, G.; Kele, P. *Chem.–Asian. J.* **2013**, *8*, 494. (c) Mizuno, T., Umezawa, K., Shindo, Y., Citterio, D., Oka, K., & Suzuki, K. *J. Fluoresc.* **2013**, *23*, 1007. (d) Cao, C.; Liu, X.; Qiao, Q.; Zhao, M.; Yin, W.; Mao, D. *Chem. Comm.* **2014**, *50*, 15811. (e) Moriarty, R. D.; Martin, A.; Adamson, K.; O'Reilly, E.; Mollard, P.; Forster, R. J.; Keyes, T. E. *J. Microscopy* **2014**, *253*, 204.
- (5) Ehrenschwender, T.; Varga, B. R.; Kele, P.; Wagenknecht H.-A. *Chem.–Asian J.* **2010**, *5*, 1761.
- (6) *Selected dyes examples presenting mega-Stokes shifts*: (a) Riedl, J.; Pohl, R.; Ernsting, N. P.; Orsag

- P.; Fojta, M. *Chem. Sci.* **2012**, *3*, 2797. (b) Weinberger, M.; Berndt, F.; Mahrwald, R.; Ernsting, N. P.; Wagenknecht, H.-A. *J. Org. Chem.* **2013**, *78*, 2589. (c) Riedl, J.; Pohl, R.; Rulišek, L.; Hocek, M. *J. Org. Chem.* **2012**, *77*, 1026. (d) Okamoto, A.; Tainaka, K.; Saito, I. *Bioconjugate Chem.* **2005**, *16*, 1105. (e) Kimura, T.; Kawai, K.; Majima, T. *Org. Lett.* **2005**, *7*, 5829. (f) Suzuki, A.; Kimura, K.; Ishioroshi, S.; Saito, I.; Nemoto, N.; Saito, Y. *Tetrahedron Lett.* **2013**, *54*, 2348. *No quantum yield provided:* (g) Jadhav, V. R.; Barawkar, D. A.; Ganesh, K. N. *J. Phys. Chem. B* **1999**, *103*, 7383.
- (7) Dziuba, D.; Karpenko, I. A.; Barthes, N. P. F.; Michel, B. Y.; Klymchenko, A. S.; Benhida, R.; Demchenko, A. P.; Mély, Y.; Burger, A. *Chem.–Eur. J.*, **2014**, *20*, 1998.
- (8) (a) Moreira, B. G.; You, Y.; Owczarzy, R. *Biophys. Chem.* **2015**, *198*, 36. (b) Iqbal, A.; Arslan, S.; Okumus, B.; Wilson, T. J.; Giraud, G.; Norman, D. G.; Ha, T.; Lilley, D. M. J. et al. (2008). *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 11176.
- (9) *Two fluorophore (radiative) FRET applied to nucleic acids; two-dye molecular beacon:* (a) Varghese, R.; Wagenknecht, H.-A. *Org. Biomol. Chem.* **2010**, *8*, 526. (b) Holzhauser, C.; Wagenknecht, H.-A. *Angew. Chem. Int. Ed.* **2011**, 7268. (c) Jockusch, S.; Martí, A. A.; Turro, N. J.; Li, Z.; Li, X.; Ju, J.; Stevens, N.; Akins, D. L. *Photochem. Photobiol. Sci.* **2006**, *5*, 493–498. (d) Socher, E.; Bethge, L.; Knoll, A.; Jungnick, N.; Herrmann, A.; Seitz, O. *Angew. Chem. Int. Ed.* **2008**, 9555. *Binary energy transfer* (e) Ikeda, S.; Kubota, T.; Wang, D. O.; Yanagisawa, H.; Yuki, M.; Okamoto, A. *Org. Biomol. Chem.* **2011**, *9*, 6598. *Detection of RNA–Small Molecule Binding:* (f) (1) Xie, Y.; Dix, A. V.; Tor, Y. *J. Am. Chem. Soc.* **2009**, *131*, 17605.
- (10) Barthes, N. P. F.; Karpenko, I. A.; Dziuba, D.; Spadafora, M.; Auffret, J.; Demchenko, A. P.; Mély, Y.; Benhida, R.; Michel, B. Y.; Burger, A. *RSC Adv.* **2015**, *5*, 33536.
- (11) Smith, M. A.; Neumann, R. M.; Webb, R. A. *J. Heterocycl. Chem.* **1968**, *5*, 425.
- (12) Dziuba, D.; Benhida, R.; Burger, A. *Synthesis* **2011**, 2159.
- (13) Asakura, J.-I.; Robins, M. J. *Tetrahedron Lett.* **1988**, *29*, 2855. (b) Asakura, J.-I.; Robins, M. J. *J. Org. Chem.* **1990**, *55*, 4928.
- (14) *From 5-ethynyluridines:* (a) G. Luoni, C. McGuigan, G. Andrei, R. Snoeck, E. De Clercq and J. Balzarini, *Bioorg. Med. Chem. Lett.*, **2005**, *15*, 3791–3796; (b) C. McGuigan, O. Bidet, M. Derudas, G. Andrei, R. Snoeck and J. Balzarini, *Bioorg. Med. Chem.*, **2009**, *17*, 3025.
- (15) (a) *General review including N3-protection:* Beaucage, S. L.; Iyer, R. P. *Tetrahedron*, **48**, 2223. (b) *N3-benzoylation:* Kumarasinghe, E. S.; Peterson, M. A.; Robins, M. J. *Tetrahedron Lett.* **2000**, *41*, 8741.
- (16) Melhuish, W. H. *J. Phys. Chem.* **1961**, *65*, 229.
- (17) Ormson, S. M.; Brown, R. G.; Vollmer, F.; Rettig, W. *J. Photochem. Photobiol., A* **1994**, *81*, 65.
- (18) Nesterova, I. V.; Erdem, S. S.; Pakhomov, S.; Hammer, R. P.; Soper, S. A. *J. Am. Chem. Soc.* **2009**, *131*, 2432.
- (19) (a) Dziuba, D.; Postupalenko, V. Y.; Spadafora, M.; Klymchenko, A. S.; Guérineau, V.; Mély, Y.; Benhida, R.; Burger, A. *J. Am. Chem. Soc.* **2012**, *134*, 10209. (b) Sinkeldam, R. W.; Marcus, P.; Uchenik, D.; Tor, Y. *ChemPhysChem* **2011**, *12*, 2260.
- (20) (a) Wilhelmsson, L. M. Q. *Rev. Biophys.* **2010**, *43*, 159. (b) Sinkeldam, R. W.; Greco, N. J.; Tor, Y. *Chem. Rev.* **2010**, *110*, 2579.
- (21) Kucherak, O. A.; Didier, P.; Mély, Y.; Klymchenko, A. S. *J. Phys. Chem. Lett.* **2010**, *1*, 616.
- (22) (a) Kucherak, O. A.; Richert, L.; Mély, Y.; Klymchenko, A. S. *Phys. Chem. Chem. Phys.* **2012**, *14*, 2292. (b) Sarkar, M.; Samanta, A. *J. Phys. Chem.. B* **2007**, *111*, 7027.
- (23) *PRODAN applied to DNA labeling:* (a) Tainaka, K.; Tanaka, T.; Ikeda, S.; Nishiza, K.-I.; Unzai, T.;

- Fujiwara, Y.; Saito, I.; Okamoto, A. *J. Am. Chem. Soc.* **2007**, 4776–4784. (b) Okamoto, A.; Tainaka, K.; Unzai, T.; Saito, I. *Tetrahedron* **2007**, 63, 3465–3470. (c) Kimura, T.; Kawai, K.; Majima, T. *Chem. Commun.* **2006**, 1542. Other contexts: (d) Marini, A.; Muñoz-Losa, A.; Biancardi, A.; Mennucci, B. *J. Phys. Chem. B* **2010**, 114, 17128. (e) Shynkar, V. V.; Klymchenko, A. S.; Piémont, E.; Demchenko, A. P.; Mély, Y. *J. Phys. Chem. A* **2004**, 108, 8151.
- (24) (a) Altman, R. B.; Terry, D. S.; Zhou, Z.; Zheng, Q.; Geggier, P.; Kolster, R. A.; Zhao, Y.; Javitch, J. A.; Warren, J. D.; Blanchard, S. C. *Nat. Meth.* **2011**, 9, 68. (b) Zheng, Q.; Juette, M. F.; Jockusch, S.; Wasserman, M. R.; Zhou, Z.; Altman, R. B.; Blanchard, S. C. *Chem. Soc. Rev.* **2014**, 43, 1044. (c) Tinnefeld, P.; Cordes, T. *Nat. Meth.* **2012**, 9, 426.
- (25) Berlier, J. E.; Rothe, A.; Buller, G.; Bradford, J.; Gray, D. R.; Filanoski, B. J.; Telford, W. G.; Yue, S.; Liu, J.; Cheung, C.-Y.; Chang, W.; Hirsch, J. D.; Beechem, J. M.; Haugland, R. P.; Haugland, R. P. *J. Histochem. Cytochem.* **2003**, 51, 1699.
- (26) Berlier, J. E.; Rothe, A.; Buller, G.; Bradford, J.; Gray, D. R.; Filanoski, B. J.; Telford, W. G.; Yue, S.; Liu, J.; Cheung, C.-Y.; Chang, W.; Hirsch, J. D.; Beechem, J. M.; Haugland, R. P.; Haugland, R. P. *J. Histochem. Cytochem.* **2003**, 51, 1699.
- (27) (a) Demchenko, A. P. *J. Fluoresc.* **2010**, 20, 1099. (b) Demchenko, A. P. *J. Mol. Struct.* **2014**, 1077, 51.
- (28) (a) Iqbal, A.; Arslan, S.; Okumus, B.; Wilson, T. J.; Giraud, G.; Norman, D. G.; Ha, T.; Lilley, D. M. J. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, 105, 11176. (b) Börjesson, K.; Preus, S.; El-Sagheer, A. H.; Brown, T.; Albinsson, B.; Wilhelmsson, L. Å. M. *J. Am. Chem. Soc.* **2009**, 131, 4288. (c) Teo, Y. N.; Kool, E. T. *Bioconjugate Chem.* **2009**, 20, 2371.
- (29) Iqbal, A.; Arslan, S.; Okumus, B.; Wilson, T. J.; Giraud, G.; Norman, D. G.; Ha, T.; Lilley, D. M. J. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, 105, 11176.
- (30) Shaner, N. C.; Lin, M. Z.; McKeown, M. R.; Steinbach, P. A.; Hazelwood, K. L.; Davidson, M. W.; Tsien, R. Y. *Nat. Meth.* **2008**, 5, 545.

PUBLICATION 5

Une fois les conditions mises au point avec le composé modèle MTCU **8**, nous nous sommes alors focalisés sur la synthèse d'ODNs incorporant le dérivé ratiométrique sélectionné **7c**. Ces conditions ont donc été appliquées à la synthèse du phosphoramidite correspondant qui a ensuite été incorporé dans des séquences d'ODNs dodécamère (cible de SRA) et pentadécamères (modèles), obtenues avec de bons rendements.

L'ensemble des séquences modèles synthétisées ont été analysées, par spectroscopie UV-visible (T_m) et Dichroïsme Circulaire, afin de déterminer l'influence de la 3-HC sur la stabilité du duplexe et la structure secondaire de ce dernier. Ensuite, les propriétés d'absorption et d'émission ont été étudiées sous différentes conditions sur les simples et doubles brins d'ADN afin de déterminer les maxima d'absorption et d'émission, le rapport d'intensité des 2 bandes d'émission et le rendement quantique. Ces données sont essentielles avant d'engager les études d'interactions entre la séquence dodécamère et le domaine SRA.

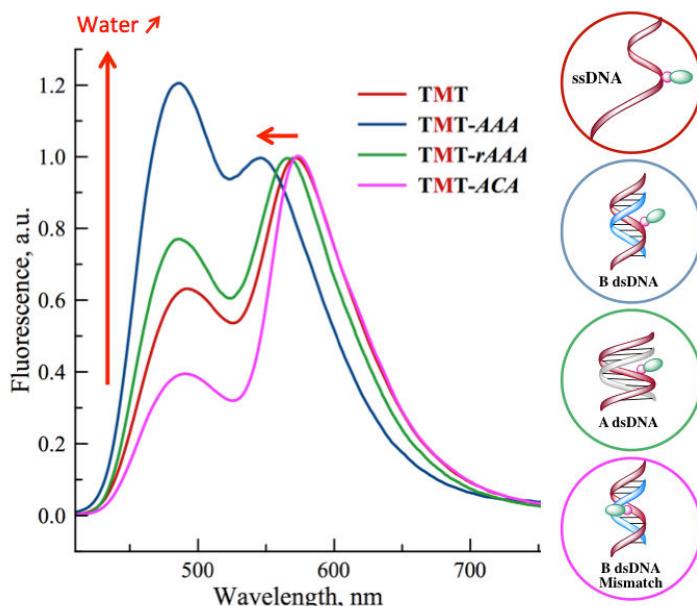


Figure 9. Sensibilité de la sonde **7c** aux changements de polarité environnante, après incorporation (courbes normalisées sur la bande T^*).

Nous avons dans un premier temps vérifié que cet analogue de dU gardait sa capacité à s'apparier sélectivement avec la dA et qu'il ne modifiait pas la conformation globale du duplexe. Nous avons également pu constater que le nucléotide présentait toujours une émission duale une fois incorporé et que cette émission était très sensible aux changements environnementaux. En effet, elle rendait possible la discrimination entre des duplexes appariés et mésappariés, ainsi qu'entre les conformations B- et A- de duplexes ADN / ADN et ADN / ARN, respectivement, par changement du rapport I_{N^*}/I_{T^*} , couplé à un décalage du maximum d'émission de la bande T^* . L'ensemble de ces données sont en accord avec l'hypothèse que nous avions formulée, à savoir qu'une sonde très sensible à l'hydratation pouvait être utilisée pour étudier des changements conformationnels de l'ADN. La grande sensibilité de ces sondes au taux d'hydratation devrait permettre de

proposer de nouveaux outils fluorescents pour détecter et étudier les interactions nucléoprotéiques. Ces travaux ont conduit à la rédaction d'un article qui a été soumis à publication.

Dual emissive analogue of deoxyuridine as a sensitive hydration-reporting probe for discriminating mismatched from matched DNAs and DNA/DNA and DNA/RNA duplexes

Nicolas P. F. Barthes,^{a†} Krishna Gavvala,^{b‡} Dmytro Dziuba,^a Dominique Bonhomme,^a Iuliia A. Karpenko,^a Anne S. Dabert-Gay,^c Delphine Debayle,^c Alexander P. Demchenko,^d Rachid Benhida,^a Benoît Y. Michel,^a Yves Mély,^b and Alain Burger^{*a}

Fluorescent nucleosides with high sensitivity to H-bonding are of great interest not only to investigate DNA hydration and DNA-protein interactions and dynamics, but also to develop screening assays. Here, we report on the incorporation of a dual emissive uridine analogue based on the 3-hydroxychromone scaffold (3HC) into DNA oligonucleotides of different sequences and the characterization of the spectroscopic properties of the labelled oligonucleotides in matched and mismatched duplexes and in duplexes of different forms. The probe exhibits two-colour emission and spectral λ -shift, which vary as a function of the local hydration of the 3HC. It provides the highest sensitivity together with calibration on the molecular level, which allow discriminating between matched and mismatched dsDNA, as well as between B- and A-forms of DNA/DNA and DNA/RNA duplexes. We suggest that the physical basis of such reporting is the change of hydration of fluorescent derivative of DNA base and use the most efficient probe for that.

Introduction

Novel fluorescent derivatives of DNA bases that report high sensitivity on hydration are of great interest for investigating DNA hydration and folding, DNA-protein interactions and dynamics, and developing screening assays. They are expected to emerge as new sensitive tools since water coating DNAs plays a key role in determining their folding, conformations and interactions. Indeed, water molecules bound to DNA display hydration patterns and physical properties that are different from bulk water and specific to each DNA form.^[1-12] On the other hand, the binding of proteins or drugs to nucleic acids is also well known to impact the water distribution in contact with DNA.^[4,13-15]

Environment-sensitive fluorophores exhibiting solvatochromic properties are important tools for fluorescence molecular sensing. Steady-state and time-resolved fluorescence spectra of solvatochromic dyes contain in particular the information about specific and non-specific interactions of the probe with the surrounding molecules and the rates of the media relaxation. Different solvatochromic dyes were investigated to interrogate the polarity of the minor and major grooves and different conformations of the DNA by monitoring the changes in the absorption and/or emission maxima.^[16-26] The dielectric constant and Lippert-Mataga scales^[27,28] were employed to estimate the polarity of the local environment of the probes. However, exploiting these scales to estimate the polarity in hydrated media is not fully satisfactory since they do not take into account the contribution of specific H-bonding interactions.^[17] Such interactions are particularly critical for probes having a conjugated donor–acceptor system (e.g. dansyl fluorophores: DAN),^[17,20,29] being in contact to water in the considered grooves. The usage of the Reichardt scale^[30] gives in general better correlations because this empirical scale accounts for the dielectric constant of the solvent and its H-bond donor ability. As a consequence, the introduction of a fluorescent

probe to sense the DNA grooves and displaying high sensitivity to H-bonding rather than to dipolar interactions was hitherto lacking.^[17]

An ideal hydration-sensitive probe should satisfy several additional demands. It should be incorporated in a way not perturbing the nucleic acid structure being located either in the major or minor groove of dsDNA. The H-bond sensor (usually an oxygen of a carbonyl, amide or sulfonamide group)^[17,20,29] should be located in a position that is sensitive to DNA transformations by shifting between H-bond free and H-bonded forms. To address all these requirements, we have recently reported a new dual emissive nucleoside **1** (Figure 1) highly sensitive to hydration.^[31,32] The new probe was engineered by connecting the 2-thienyl-3-hydroxychromone (3HC) fluorophore to the 5-position of deoxyuridine via an ethynyl bond to favour interactions with water molecules rather than to substitute a natural base by the dye.^[33,34] A rigid acetylenic linker was chosen to locate the dye in a precise position with respect to the DNA major axis with the aim to restrain undesired contact with H-bond donor and acceptor groups of DNA. Such control is of key importance for further data interpretation of hydration and biomolecular interaction studies.

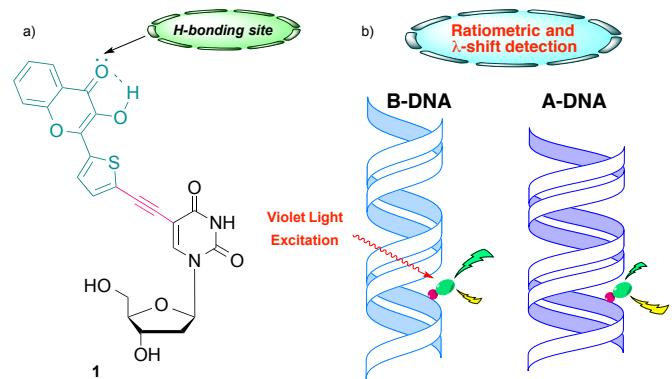
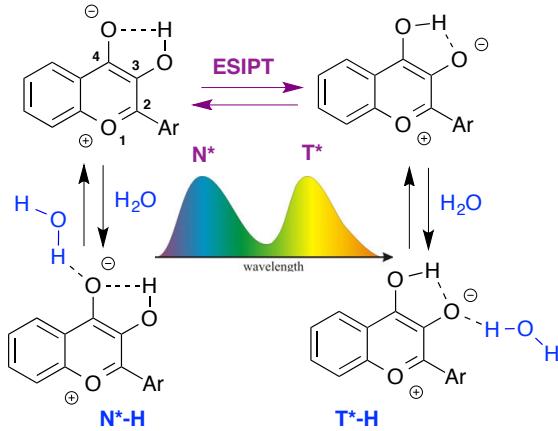


Figure 1. a) Structure of the dual emissive deoxyuridine conjugate **1** and b) Representative concept for the conformational study of hybrid duplexes based on the ratiometric detection of the local hydration changes.

Due to an excited state intramolecular proton transfer reaction (ESIPT),^[35] **1** can exist in two excited states, the normal (N^*) and the tautomer (T^*) forms, that emit light with different energies and provide two well-separated emission bands (Scheme 1).



Scheme 1. Proposed mechanism for water sensing by 3HC **1** adapted from.^[31] Excited states N^* and N^*-H , and T^* and T^*-H of the normal and tautomer forms are in equilibrium.^[36] N^*-H undergoes ESIPT reaction in a much slower time scale than N^* . N^* and N^*-H , and T^* and T^*-H generate the higher and lower energy emission bands, respectively referred as N^* and T^* on the spectrum for simplification.

In the excited normal state (N^*), the oxygen carbonyl at 4-position of 3HC becomes more basic, making this group subject to formation of H-bonds with protic solvents. In particular, **1** demonstrates high sensitivity to hydration since an increase in water concentration inhibits the ESIPT reaction and thus, increases the relative intensity of the N^* band (Figure 2).^[31]

Moreover, hydration also favours H-bonding of the tautomer forms, which results in the blue-shift of the T^* emission maximum as a consequence of the more stabilized ground state and the increased energy gap between the excited and ground states of the hydrated forms.^[37] Therefore, measuring the intensity ratios of the two emission bands (I_{N^*}/I_{T^*}) and also the λ -shifts of the T^* band allows monitoring, by two different channels of information, the hydration changes in the vicinity of the 3HC. Additionally, the

newly synthesized dye absorbs strongly in the violet ($\lambda_{\text{Abs}} \sim 400$ nm), exhibits large Stokes shifts (> 100 nm), keeps significant quantum yield in water (6%), which is unusual among polarity-sensitive dyes, and thus possesses useful properties for selective excitation and easier detection.^[38-41]

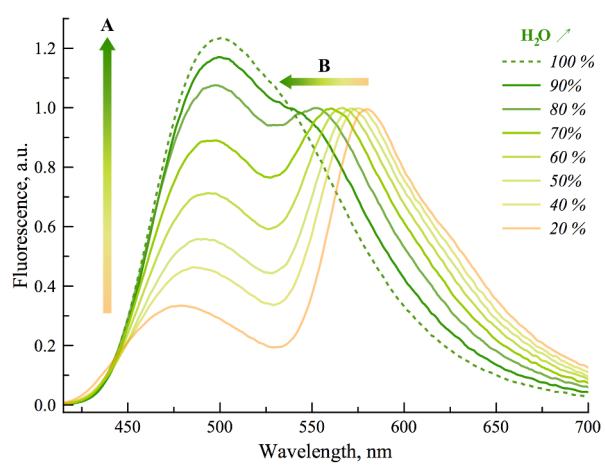
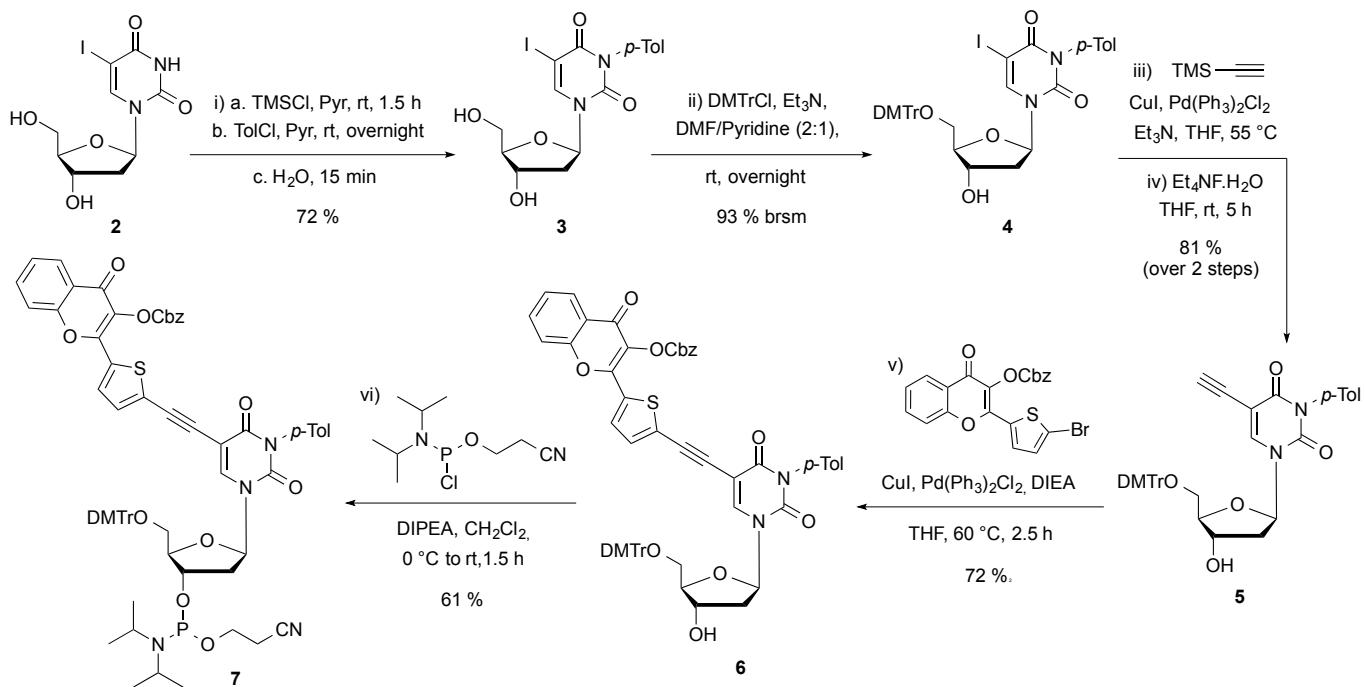


Figure 2. Two-band ratiometric and wavelength-shifting responses of conjugate **1** to hydration in water-dioxane mixtures.^[31] Growing water concentration results in the relative increase of the N^* band and the blue-shift of the maximum of the T^* band emission: A and B arrows, respectively. The fluorescence spectra are normalized by T^* band.

Successful synthesis of the conjugate **1** with such remarkable spectroscopic properties stimulated our efforts to obtain oligodeoxynucleotides (ODNs) with its incorporation. In this work, we describe the synthesis and photophysical properties of such labelled ODNs. We demonstrate that the incorporated nucleoside is able to probe the hydration of the major groove and to discriminate the environmental and conformational changes of DNA/DNA and DNA/RNA duplexes of different compositions (Figure 1b).



Scheme 2. Synthesis of the phosphoramidite of the conjugate 3HC-deoxyuridine **1**.

RESULTS AND DISCUSSIONS

Synthesis and photophysical properties of labelled DNA and DNA/DNA duplexes

The synthesis of the phosphoramidite **7** required for the preparation of the labelled ODNs is described in Scheme 2. The synthetic approach was adapted from the procedure developed for dye **1**.^[31] The protection of the 3-imido and 5'-hydroxyl groups of **2**, provided the 5-iododeoxyuridine derivative **4**. This intermediate was then engaged in a first Sonogashira coupling with trimethylsilylacetylene. Subsequent treatment with tetraethylammonium fluoride gave the terminal alkyne **5**. Concerning the final assembly, a second Sonogashira coupling with the bromothienyl chromone partner afforded the conjugate **6** of 3HC and deoxyuridine. Phosphitylation of **6** under classical conditions yielded phosphoramidite **7**, which was used for the solid-phase synthesis of the labelled 15-mer ODNs (SI). HPLC, UV and MS analysis ascertained the purity and integrity of the synthesized ODNs (Table S1, Figures S1, S2, S3 & S4).

The 15-mer d(CGT TTT **XMX** TTT TGC) sequences, where **X** = **T**, **A**, **C** and **G** and **M** = conjugate **1**, were chosen as model sequences to characterize the photophysical properties of **1** in different ODN contexts. The conjugate was positioned in the middle rather than at a terminal position of the duplexes in order to avoid structural disorders due to end effects. The sequences referred as **TMT**, **AMA**, **CMC** and **GMG** and differed by the residues flanking the modified nucleotide **1**. Control experiments were performed with wild-type strands of composition d(CGT TTT **XTX** TTT TGC). Each ODN was then annealed with

its complementary ODN (noted in italic) containing a natural base or an abasic site at the position opposite **1**, resulting in duplexes of different compositions (Table 1 and SI). The sequences of the complementary strands were *d(GCA AAA YXY AAA ACG)*, where **Y** = **A**, **T**, **C** or **G** and **X** is any of the natural bases (matched and mismatched pairs) or an abasic site (**Ab**). We first assessed whether the deoxyuridine derivative **1** selectively base pairs with **A** and does not affect the secondary structure of the duplexes by monitoring the temperature-induced absorbance changes at 260 nm and recording the CD spectra (Tables 1 & S2, Figures S5, S6 & S7).

Thermal denaturation studies showed that the most stable labelled duplexes were obtained with the fully complementary sequences. The labelled duplexes were slightly less stable than the corresponding wild-type matched and mismatched duplexes (ΔT_m = -1 to -4.9 °C) except for **C** opposite **1** (ΔT_m = 0.1 to 3.6 °C, Tables 1 & S2). Comparable observations were reported for other hydrophobic moieties, such as phenanthroline, anthracene, and pyrene, connected to **U** via an acetylenic linker.^[44-46] Remarkably, labelled duplexes containing an abasic site were significantly more stable than the corresponding controls (ΔT_m = +4.7 to +7 °C). The increase of T_m of the **M-Ab** duplexes suggests that the dye accommodates well the abasic lesion and stabilizes the duplex, likely through its large stacking area and hydrophobic character.^[47] Furthermore, analysis of the CD spectra showed that the label did not affect the B-helix conformation of the DNA/DNA duplexes (Figures 3 & S5).

Table 1. Spectroscopic data for conjugate **1** and labelled ODNs.

Samples ^{a,b}	T _m (°C) ^c	% H ₂ O ^d	λ _{abs} (nm) ^d	λ _{N*} (nm) ^e	λ _{T*} (nm) ^f	I _{N*} /I _{T*} ^g	Φ (%) ^h
1	100	395	500 ⁱ	-	-	-	6
	90	394	497	539	1.17	6	
	80	395	498	552	1.08	8	
	70	395	494	561	0.89	26	
	60	395	493	567	0.72	27	
	50	394	490	571	0.56	25	
TMT	-	405	493	572	0.63	13	
TMT-AAA	45.8	397	487	548	1.21	11	
TMT-ATA	37.9	404	491	573	0.53	12	
TMT-ACA	39.1	405	491	573	0.40	16	
TMT-AGA	38.1	406	490	573	0.57	6	
TMT-AAbA	39.2	408	492	573	0.47	11	
TMT-rAAA	38.8	402	487	566	0.78	20	
AMA	-	407	492	570	0.63	9	
AMA-TAT	42.4	400	488	552	1.21	8	
AMA-TTT	37.6	406	495	571	0.48	11	
AMA-TCT	39.2	407	493	574	0.47	11	
AMA-TGT	38.1	407	491	571	0.58	9	
AMA-TabT	36.3	408	497	573	0.67	13	
AMA-rUAU	37.1	406	490	564	0.85	15	
CMC	-	405	493	569	0.70	9	
CMC-GAG	48.1	394	487	548	1.00	9	
CMC-GTG	37.9	407	489	565	0.92	4	
CMC-GCG	42.6	412	492	569	0.83	3	
CMC-GGG	42.5	406	490	567	0.76	4	
CMC-GAbG	43.2	409	493	570	0.84	3	
GMG	-	408	491	573	0.66	3	
GMG-CAC	52.4	400	488	561	1.00	1	
GMG-CTC	45.6	407	490	572	0.60	2	
GMG-CCC	47.1	406	489	569	0.43	2	
GMG-CGC	47.3	407	488	571	0.84	1	
GMG-CAbC	44.5	409	490	573	0.60	2	

a) 2 μM of monomer **1** in water-dioxane mixtures; b) 2 μM of ODN in buffer pH 7.0 (20 mM Sodium phosphate, 100 mM NaCl, 1 mM EDTA); c) Melting temperature; ±0.5°C; d) Position of the absorption maximum; e) Position of the emission maximum of the normal N* band; f) Position of the emission maximum of the tautomer T* band; g) I_{N*/I_{T*}} corresponds to the ratio of the intensities of the two emission bands at their maxima; ±2.5% mean standard deviation; h) Quantum yields determined using quinine sulfate (QS) in 0.1 M HCl solution ($\lambda_{ex} = 350$ nm, $\Phi = 0.54$)^[42] and p-dimethylaminoflavone (dMAF) in EtOH ($\lambda_{ex} = 404$ nm, $\Phi = 0.27$)^[43] as references, ±10% mean standard deviation; i) In water, the N* and T* bands collapse in a single band.

In a next step, the UV absorbance and fluorescence spectra of **1** in the single and double stranded forms were recorded in phosphate buffer pH 7.0 (Figures 4, S8, S9, S10, S11 & S12). The spectroscopic characteristics of the labelled sequences together with the data of the monomer **1** in water-dioxane mixtures are reported in Table 1.

The fully complementary dsODNs showed absorption bands centred at 394 and 400 nm while the absorption maxima of the single strands and mismatched duplexes were in the range of 404-409 nm. The positions of the absorption maxima in the fully complementary dsODNs are close to the absorption maximum of the monomer **1** in water-dioxane mixtures suggesting that the dye is more accessible to water than in the single strands and

mismatched duplexes. By contrast to matched duplexes, the bathochromic shift of about 7-11 nm observed for the single strands and mismatched duplexes (Figure S8 and Table 1) suggests π-interactions of the hydrophobic dye with the surrounding bases. Similar effects were reported for the pyrenecarboxamide chromophore anchored at the surface of the major groove of DNA.^[26] A red-shift of the absorption maximum is also commonly observed with intercalating dyes. The dye in duplexes with an abasic site exhibited the largest red-shift, confirming its suitable intercalation in these sequences.

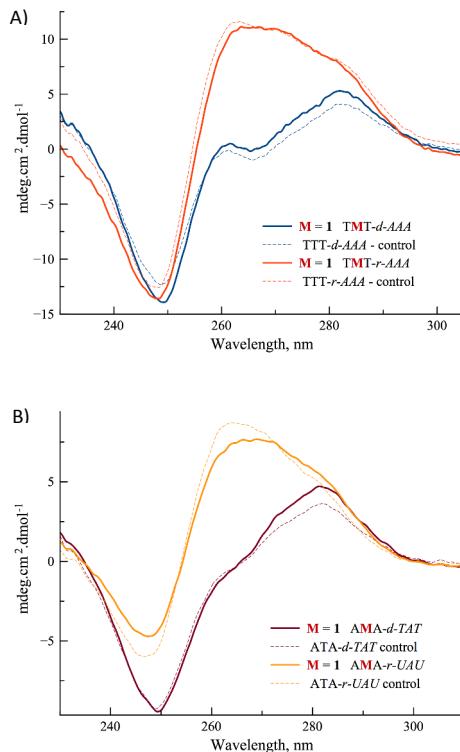


Figure 3. CD spectra of matched (A) **TMT-AAA**, **TMT-rAAA**, (B) **AMA-TAT**, **AMA-rUAU** and their native wild type duplexes recorded in pH 7.0 buffer (20 mM Sodium phosphate, 100 mM NaCl, 1 mM EDTA).

The fluorescence spectra showed two well-resolved bands of emission (Figures 4, S9, S10, S11 & S12). Comparison of the spectra revealed that the matched duplexes displayed the highest I_{N*/I_{T*}} ratio and the more blue-shifted T* band (Table 1, Figures 4, S9, S10, S11 & S12). For instance, for the matched **TMT-AAA** duplex, the intensity ratio I_{N*/I_{T*}} and emission maximum of the T* band were 1.21 and 548 nm, respectively; while for the **TMT** single strand and the mismatched duplexes, they were in the 0.4-0.63 and 572-573 nm ranges, respectively.

The intensity ratios and the maxima of the T* emission of the matched duplexes were close to the values obtained for **1** in 90 % water, while the corresponding values in the single strands and mismatched duplexes were close to the values observed for **1** in 50 % water (Table 1, Figures 2 and 4). Noticeably, **TMT** and **AMA** single and double strands gave very close results, indicating that the spectroscopic

properties of conjugate **1** are similarly influenced by T and A flanking bases (Table 1, Figures 4, S8, S9 & S10). The **CMC** and **GMG** sequences followed similar trend (Figures S11 & S12). Nevertheless, the intensity ratios of the fully matched duplexes were smaller than those obtained for the corresponding **TMT** and **AMA** double strands indicating that the dye is exposed in a slightly less hydrated environment (compare **1** with 1.21, similar to the values obtained for **1** in about 80 % water). All the **TMT** and **AMA** ODNs, and the **CMC** single and matched double strands gave similar or larger quantum yields as compared to that of the free dye **1** in water. The fluorescence quantum yield of the dye in these single- and double-stranded ODNs varied from 6 to 16 %. By contrast the mismatched duplexes of **CMC** and all the **GMG** ODNs gave lower quantum yields suggesting a direct contact quenching of **1** by **G** in these ODNs. Quenching of fluorescent nucleoside analogues in ODNs by **G** is common.^[48,49]

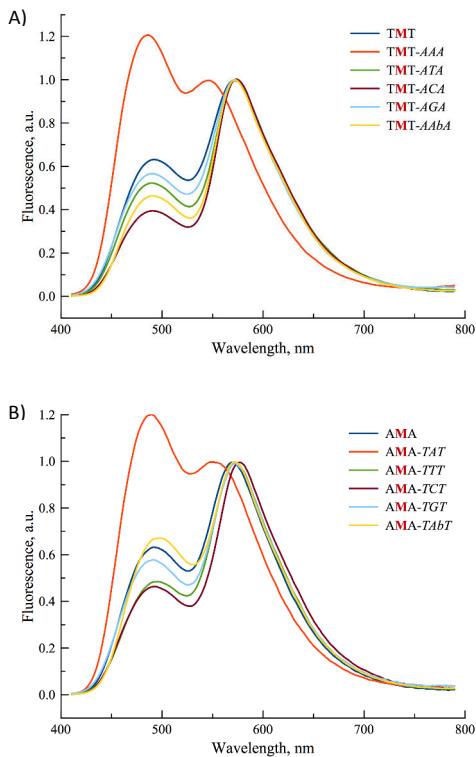


Figure 4. Fluorescence spectra of ss, matched (A) TMT & TMT-AXA, (B) AMA, AMA-TXT and their corresponding mismatched duplexes recorded in pH 7.0 buffer (20 mM Sodium phosphate, 100 mM NaCl, 1 mM EDTA) and normalized by T* band.

Taken together, our data indicate that the analysis of the two channels of information of the label allows easily distinguishing the matched duplexes from the single strands and mismatched duplexes. The I_{N^*}/I_{T^*} channel demonstrates superior sensitivity (up to 40 fold) to distinguish matched from mismatched duplexes with variation in I_{N^*}/I_{T^*} from 10 to 200 %[§] as compared to the 5 % variation of the shift of the T* band. In matched duplexes, the label is exposed to a highly hydrated

environment as compared to single strands and mismatched duplexes, where it is partially shielded from the water. Noticeably, the label is able to sense in the fully matched DNA the subtle change in hydration of the local environment of **G/C** and **A/T** base pairs.^[2] It distinguishes remarkably from the DAN (6-dimethylamino-2-acylnaphthalene) fluorophore, which was unable to detect such changes.^[24] The new label possesses also attractive features for A-allele detection. The ODNs with **A** and **T** flanking the label demonstrate the highest sensitivity for discriminating matched from mismatched bases.

Discrimination of DNA/DNA and DNA/RNA duplexes

Next, we examined the ability of **1** to monitor the transition from B- to A-form duplexes at constant ionic strength. Such transition is physiologically important as it is produced for instance in base pairing of DNA and RNA strands during transcription and it is thought to govern a number of protein/DNA interactions.^[6,50]

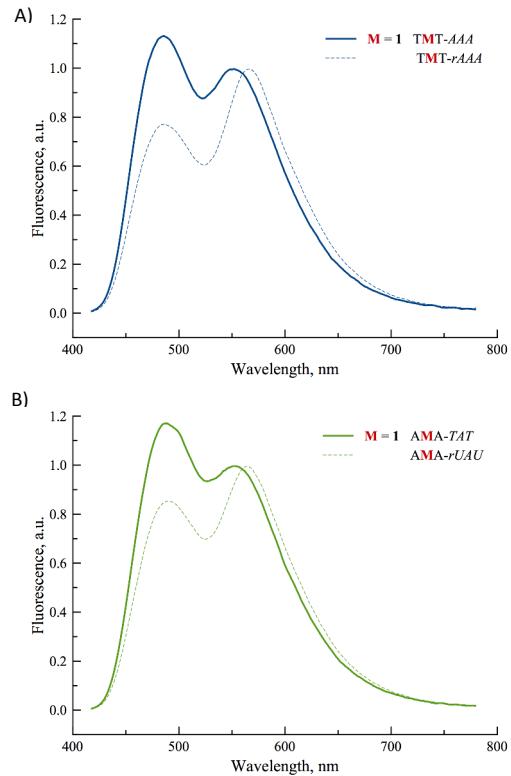


Figure 5. Fluorescence spectra of (A) TMT-AAA and TMT-rAAA (B) AMA-TAT and AMA-rUAU recorded in pH 7.0 buffer (20 mM Sodium phosphate, 100 mM NaCl, 1 mM EDTA) and normalized by T* band.

Thus, the **TMT** and **AMA** sequences were annealed with the complementary RNA strands *r(GCA AAA YAY AAA ACG)*, §§[51] where **Y** is **A** or **U**, and referred as **rAAA** and **rUAU**, respectively. Upon hybridization of RNA to DNA, a clear difference in the secondary structure of the obtained hybrid in comparison to the DNA/DNA duplex was evidenced from the CD spectra (Figure 3). A characteristic blue-shift of the

positive band, as well as an increase in its amplitude were observed, as expected.^[52] The impact of probe **1** on the stability of the hybrids was comparable to that observed for DNA/DNA duplexes (Table S2).

The absorption and fluorescence characteristics of the two duplexes are given in Table 1. The fluorescence spectra of the two DNA/RNA hybrids gave similar results but, revealed dramatic changes relative to the B-form duplexes (Figures 5 & S13). For **TMT-rAAA** for instance, the intensity ratio I_{N^*}/I_{T^*} and emission maximum of the T^* band were 0.78 and 566 nm, respectively. Comparison with the corresponding values of **TMT-AAA** gave a 1.6 fold ratio decrease and 16 nm red-shift, respectively.

It is noteworthy that the values of **TMT-rAAA** were comparable to the values obtained for dye **1** in 70% water (Table 1), indicating that the local environment of the dye was less hydrated in the DNA/RNA hybrid than in the DNA/DNA duplex. These data are consistent with the lower hydration of the narrower and deeper major groove of A-DNA in comparison to the major groove in B-DNA.^[24,53,54] They also illustrate that the proper positioning of the water-sensitive reporting carbonyl group of **1** affords a highly sensitive and straightforward method to discriminate B from A duplexes. Compared to the DAN fluorophore classically used for the same purpose,^[24] probe **1** demonstrates an about 11 fold-enhanced sensitivity to detect such transition with a 55 % variation in I_{N^*}/I_{T^*} for **1** as compared to 5% variation of the λ -shift of the emission maximum for DAN.

Conclusions

In this report, ODNs labelled with a new fluorescent conjugate of 3HC and deoxyuridine were synthesized and characterized. It was shown that the unique property of the designed 3HC derivative to provide high sensitive wavelength-ratiometric response to H-bonding is retained on its incorporation into oligonucleotides of different sequences. Monitoring the two channels of information (I_{N^*}/I_{T^*} ratio and λ -shift of T^*) allows estimating the local hydration and discriminating matched from mismatched DNA/DNA duplexes. Moreover, this conjugate enables also to detect the difference between the B and A forms of DNA/DNA and DNA/RNA duplexes, respectively, demonstrating ~ 10-fold increased sensitivity over the DAN fluorophore.^[24] The high sensitivity to hydration displayed in two information channels^[55] should open the access to a variety of applications in DNA research, sensing its interactions and drug design.

Acknowledgements

We thank the ANR (ANR-12-BS08-0003-02), PACA région (DNAfix-2014-02862 and 2014-07199) and the FRM (DCM20111223038) for financial support and a grant for N.P.F.B. We thank the French Government for the Master 2

grant to I.A.K. and Campus France for the Raman-Charpak fellowship to K.G.

Notes and references

- § Calculated by comparison with the largest and smallest intensity ratios of the mismatched duplexes, respectively.
- §§ For convenience, the more stable methoxy RNA was used for hybridization with DNA, since conformational effects induced by the methylation of the 2'-OH groups are negligible and hydration of the major groove is not altered.
- [1] P. Auffinger, E. Westhof, *Curr. Opin. Struct. Biol.* **1998**, *8*, 227–236.
- [2] M. Feig, B. M. Pettitt, *Biopolymers* **1998**, *48*, 199–209.
- [3] V. Makarov, B. M. Pettitt, M. Feig, *Acc. Chem. Res.* **2002**, *35*, 376–384.
- [4] B. Jayaram, T. Jain, *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 343–361.
- [5] D. Miyoshi, H. Karimata, N. Sugimoto, *J. Am. Chem. Soc.* **2006**, *128*, 7957–7963.
- [6] I. Brovchenko, A. Kruckau, A. Oleinikova, A. K. Mazur, *J. Am. Chem. Soc.* **2008**, *130*, 121–131.
- [7] S.-I. Nakano, D. Yamaguchi, H. Tateishi-Karimata, D. Miyoshi, N. Sugimoto, *Biophys. J.* **2012**, *102*, 2808–2817.
- [8] P. Ball, *Chem. Rev.* **2008**, *108*, 74–108.
- [9] J. Choi, T. Majima, *Chem. Soc. Rev.* **2011**, *40*, 5893–5909.
- [10] I. Son, Y. L. Shek, D. N. Dubins, T. V. Chalikian, *J. Am. Chem. Soc.* **2014**, *136*, 4040–4047.
- [11] T. Zubatiuk, O. Shishkin, L. Gorb, D. Hovorun, J. Leszczynski, *J. Phys. Chem. B* **2015**, *119*, 381–391.
- [12] K. E. Furse, S. A. Corcelli, *J. Phys. Chem. Lett.* **2010**, *1*, 1813–1820.
- [13] C. K. Reddy, A. Das, B. Jayaram, *J. Mol. Biol.* **2001**, *314*, 619–632.
- [14] P. L. Privalov, A. I. Dragan, C. Crane-Robinson, K. J. Breslauer, D. P. Remeta, C. A. S. A. Minetti, *J. Mol. Biol.* **2007**, *365*, 1–9.
- [15] S.-I. Nakano, D. Miyoshi, N. Sugimoto, *Chem. Rev.* **2014**, *114*, 2733–2758.
- [16] R. W. Sinkeldam, N. J. Greco, Y. Tor, *ChemBioChem* **2008**, *9*, 706–709.
- [17] R. W. Sinkeldam, Y. Tor, *Org. Biomol. Chem.* **2007**, *5*, 2523.
- [18] V. R. Jadhav, D. A. Barawkar, K. N. Ganesh, *J. Phys. Chem. B* **1999**, *103*, 7383–7385.
- [19] A. Okamoto, K. Tainaka, Y. Fujiwara, *J. Org. Chem.* **2006**, *71*, 3592–3598.
- [20] K. Tainaka, K. Tanaka, S. Ikeda, K.-I. Nishiza, T. Unzai, Y. Fujiwara, I. Saito, A. Okamoto, *J. Am. Chem. Soc.* **2007**, *129*, 4776–4784.
- [21] R. Jin, K. J. Breslauer, *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 8939–8942.
- [22] G.-Q. Tang, N. Tanaka, S. Kunugi, *Bull. Chem. Soc. Jpn.* **1999**, *72*, 1129–1137.
- [23] T. Kimura, K. Kawai, T. Majima, *Chem. Commun.* **2006**, 1542.
- [24] T. Kimura, K. Kawai, T. Majima, *Org. Lett.* **2005**, *7*, 5829–5832.

- [25] A. Okamoto, K. Tainaka, I. Saito, *Bioconjugate Chem.* **2005**, *16*, 1105–1111.
- [26] A. Okamoto, K. Kanatani, I. Saito, *J. Am. Chem. Soc.* **2004**, *126*, 4820–4827.
- [27] E. Z. Lippert, *Z. Naturforsch. A* **1955**, 541–545.
- [28] N. Mataga, Y. Kaifu, M. Koizumi, *Bull. Chem. Soc. Jpn* **1956**, *29*, 465–470.
- [29] J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*, **2006**.
- [30] C. Reichardt, *Chem. Rev.* **1994**, *94*, 2319–2358.
- [31] N. P. F. Barthes, I. A. Karpenko, D. Dziuba, M. Spadafora, J. Auffret, A. P. Demchenko, Y. Mely, R. Benhida, B. Y. Michel, A. Burger, *RSC Adv.* **2015**, *5*, 33536–33545.
- [32] D. Dziuba, I. A. Karpenko, N. P. F. Barthes, B. Y. Michel, A. S. Klymchenko, R. Benhida, A. P. Demchenko, Y. Mely, A. Burger, *Chem. Eur. J.* **2014**, *20*, 1998–2009.
- [33] M. Spadafora, V. Y. Postupalenko, V. V. Shvadchak, A. S. Klymchenko, Y. Mely, A. Burger, R. Benhida, *Tetrahedron* **2009**, *65*, 7809–7816.
- [34] D. Dziuba, V. Y. Postupalenko, M. Spadafora, A. S. Klymchenko, V. Guérineau, Y. Mely, R. Benhida, A. Burger, *J. Am. Chem. Soc.* **2012**, *134*, 10209–10213.
- [35] A. P. Demchenko, K.-C. Tang, K. C. Tang, P.-T. Chou, *Chem. Soc. Rev.* **2013**, *42*, 1379–1408.
- [36] V. V. Shynkar, A. S. Klymchenko, E. Piémont, A. P. Demchenko, Y. Mely, *J. Phys. Chem. A* **2004**, *108*, 8151–8159.
- [37] C. A. Kenfack, A. S. Klymchenko, G. Duportail, A. Burger, Y. Mely, *Phys. Chem. Chem. Phys.* **2012**, *14*, 8910–8918.
- [38] Y. Niko, Y. Cho, S. Kawauchi, G.-I. Konishi, *RSC Adv.* **2014**, *4*, 36480–36484.
- [39] A. S. Klymchenko, A. P. Demchenko, *New J. Chem.* **2004**, *28*, 687.
- [40] O. M. Zamotaiiev, V. Y. Postupalenko, V. V. Shvadchak, V. G. Pivovarenko, A. S. Klymchenko, Y. Mely, *Bioconjugate Chem.* **2011**, *22*, 101–107.
- [41] P. A. Hopkins, R. W. Sinkeldam, Y. Tor, *Org. Lett.* **2014**, *16*, 5290–5293.
- [42] W. H. Melhuish, *J. Phys. Chem.* **1961**, *65*, 229–235.
- [43] S. M. Ormson, R. G. Brown, F. Vollmer, W. Rettig, *J. Photochem. Photobiol., A* **1994**, *81*, 65–72.
- [44] D. J. Hurley, S. E. Seaman, J. C. Mazura, Y. Tor, *Org. Lett.* **2002**, *4*, 2305–2308.
- [45] Q. Xiao, R. T. Ranasinghe, A. M. P. Tang, T. Brown, *Tetrahedron* **2007**, *63*, 3483–3490.
- [46] C. Beyer, H.-A. Wagenknecht, *Chem. Commun.* **2010**, *46*, 2230.
- [47] N. J. Greco, Y. Tor, *J. Am. Chem. Soc.* **2005**, *127*, 10784–10785.
- [48] R. W. Sinkeldam, N. J. Greco, Y. Tor, *Chem. Rev.* **2010**, *110*, 2579–2619.
- [49] **2010**, 1–25.
- [50] M. Egli, N. Usman, A. Rich, *Biochemistry* **1993**, *32*, 3221–3237.
- [51] P. Auffinger, E. Westhof, *Angew. Chem. Int. Ed.* **2001**, *40*, 4648–4650.
- [52] M. Vorlíčková, I. Kejnovská, K. Bednářová, D. Renčíuk, J. Kypr, *Chirality* **2012**, *24*, 691–698.
- [53] W. Saenger, W. N. Hunter, O. Kennard, *Nature* **1986**, *324*, 385–388.
- [54] W. Saenger, *Principles of Nucleic Acid Structure*, Springer-Verlag, New-York, **1984**.
- [55] A. P. Demchenko, *J. Mol. Struct.* **2014**, *1077*, 51–67.

Conclusion Générale et Perspectives

Dans le cadre de ce projet de recherche, nous nous sommes intéressés à la synthèse d'ODNs fluorescents, incorporant un motif 3-hydroxychromone, et capables de sonder les changements environnementaux afin d'étudier les interactions ADN / Protéine. Plus particulièrement, nous nous sommes focalisés sur l'étude du mécanisme de méthylation de l'ADN, dont des dérégulations sont rencontrées dans le cas de nombreuses pathologies (cancers, maladie d'Alzheimer, de Huntington, ...). Notre objectif principal concernait l'étude du couple UHRF1 / DNMT1 qui intervient dans la régulation du profil de méthylation au cours de la division cellulaire. D'après l'hypothèse de travail énoncé au début de ce chapitre, l'ADN hémiméthylé (HM) serait spécifiquement reconnu par le domaine SRA de UHRF1 qui ferait alors basculer la 5mC du brin parent dans son site de reconnaissance. Ensuite, UHRF1 recruterait DNMT1 au niveau d'un second site de reconnaissance. L'approche de la méthylase déstabiliserait alors le complexe ternaire transitoire formé de DNMT1, de l'ADN et de UHRF1, provoquant le départ de cette dernière et laissant le champ libre à DNMT1 afin de méthylérer la cytosine du néobrin, formé lors de la réPLICATION de l'ADN. Une dernière étape consisterait pour DNMT1 à repositionner la 5mC, nouvellement synthétisée, dans le duplexe d'ADN et à se dissocier de l'assemblage nucléique.^{48b}

Afin de confirmer ou d'inflimer cette hypothèse, nous avons tout d'abord débuté notre travail par l'étude de la première étape de ce mécanisme, à savoir la reconnaissance du brin HM par le domaine SRA de UHRF1.

Deux stratégies différentes ont été explorées afin de marquer l'ADN. La première d'entre elles repose sur la substitution d'une nucléobase naturelle par un analogue de base fluorescente, afin de nous permettre de sonder l'intérieur du duplexe. Ce biocapteur s'est

Conclusion et Perspectives

révélé être un outil particulièrement sensible. En effet, lors de l'interaction du domaine SRA avec les séquences marquées sélectionnées, nous avons pu constater une grande variation du signal de fluorescence (rendement quantique et rapport d'intensité des deux bandes d'émission) entre le duplexe HM et le duplexe NM. Ces changements du signal semblent clairement associés à la détection du basculement de la 5mC. Afin de confirmer ces résultats, le mutant G448D de SRA, incapable de basculer la 5mC, a également été synthétisé et son interaction avec le duplexe analysée. Dans ce cas, de telles variations du signal ne sont plus observées. Enfin, des études de spectroscopie complémentaires, et en particulier de *stopped-flow*, ont permis de caractériser les constantes de vitesse du processus de basculement de la base.

En plus de cette stratégie de marquage « intra-duplex », nous souhaitions également développer une seconde approche, complémentaire de la première, qui nous permettrait d'analyser l'intérieur du grand sillon de l'ADN au cours de ces interactions avec des protéines. Nous avons donc entrepris la modification d'une base nucléique naturelle sur laquelle la copule fluorescente a été attachée par l'intermédiaire d'un bras espaceur acétylénique.

La séquence d'ADN hémi-méthylé sur laquelle est basée notre étude du mécanisme de méthylation de l'ADN, présente sur son brin complémentaire, une 2'-desoxythymidine à proximité du site de liaison avec SRA. La structure du complexe SRA / ADN obtenue par diffraction des rayons X montre que la dT n'est pas impliquée dans la liaison entre la protéine et le double brin d'ADN.^{48a} C'est donc naturellement vers cette base que notre choix s'est porté afin de mettre au point cette nouvelle stratégie de marquage.

Avant de débuter la synthèse des ODNs incorporant ce nouveau motif fluorescent, des études ont au préalable été réalisées afin de déterminer les avantages qu'apporterait un tel rapporteur. Nous avons donc tout d'abord étudié les propriétés photophysiques d'un analogue simplifié, diméthylée au niveau des azotes 3 et 5 de l'uridine^{126f} avant de nous intéresser à des analogues nucléosidiques incorporant trois espaces aryles différents (phényle, furyle et thiényle), dans le but d'étudier les effets de ce lien arylique.^{126g} Il en est ressorti qu'un transfert de charge intramoléculaire entre l'azote 1 (uracile) et le carbonyle 4 (chromone) était à l'origine d'un effet bathochrome important (40 nm) et que ces sondes présentaient également de bons rendements quantiques ainsi que des coefficients d'absorptions molaires trois fois supérieur à celui de la sonde de première génération (de l'ordre de 35 000 L.mol⁻¹.cm⁻¹). Enfin, ces études ont permis de mettre en évidence une forte sensibilité au taux d'hydratation environnant. Ces résultats étaient particulièrement encourageants étant donné que l'eau joue un rôle essentiel dans la structuration et l'activité des acides nucléiques.^{126f,g}

Une fois ces études préliminaires achevées, nous avons pu nous consacrer à la synthèse des ODNs fluorescents correspondants (seconde stratégie). Afin de mettre au point les différentes étapes de synthèse (du phosphoramidite jusqu'à son incorporation dans des séquences modèles), nous avons dans un premier temps travaillé avec un composé modèle, l'analogique 3-méthoxychromone (MTCU), car le groupement méthoxy porté par la chromone est stable et facile à manipuler. La stabilité de la sonde dans les conditions de

Conclusion et Perspectives

déprotection classiques des ODNs a également pu être vérifiée. Enfin, les études spectroscopiques réalisées ont démontré que ce composé présentait des propriétés intéressantes (absorption dans le visible, coefficient d'absorption molaire important, rendements quantiques appréciables multipliés par deux après hybridation) pour être utilisé comme donneur dans un couple FRET efficace avec l'accepteur Cy3.

Une fois la méthodologie mise au point avec MTCU, nous nous sommes focalisés sur le dérivé ratiométrique cible (thiophène) qui a alors été synthétisé et incorporé dans des ODNs. Ses caractéristiques photophysiques après incorporation et son influence sur la stabilité et la structure secondaire du duplexe ont ensuite été étudiées. Le nucléotide s'apparie sélectivement avec la dA et ne modifie pas la conformation globale du duplexe. De plus, il présente toujours une émission duale une fois incorporé et cette émission est très sensible aux changements environnementaux (hydratation). En effet, elle rend possible la discrimination entre des duplexes appariés et mésappariés, ainsi qu'entre les conformations B- et A- de duplexes ADN / ADN et ADN / ARN, respectivement, par changement du rapport I_{N^*}/I_{T^*} , couplé à un décalage du maximum d'émission de la bande T*. Bien que les études avec SRA n'aient pour l'instant pas débuté, ces résultats préliminaires obtenus sur les séquences modèles sont très encourageants pour la suite du projet.

L'ensemble de ces travaux de thèse a abouti à la publication de trois articles, dont deux traitant de la seconde stratégie de marquage^{126f,g} et un de l'identification des modifications de l'ADN et de la chromatine.^{25c} De plus, un autre papier a été soumis récemment pour publication et deux autres, faisant référence à l'étude de l'interaction SRA / ADN (première stratégie) et au couple FRET (seconde stratégie), sont actuellement en cours de rédaction. Enfin, plusieurs communications orales et par affiche ont été réalisées au cours de congrès nationaux et internationaux (SCF-PACA, RICT et COST).

Dans la continuité de ces travaux sur la seconde stratégie de marquage, nous avons également débuté la synthèse de dérivés fluorescents de la dC couplés à la 3-HC. Accéder à ce type d'analogues nous paraît particulièrement intéressant car le double brin d'ADN sur lequel repose nos études avec le domaine SRA comporte 11 dC.^{48a} De plus, nous serions en mesure de marquer n'importe quelle paire de base T*-A ou C*-G (* fait référence à la modification par la copule fluorescente).

En parallèle de ces travaux de thèse, nous avons également appliqué la première stratégie de marquage à l'étude de divers mécanismes impliqués dans la transcription inverse du VIH ou l'excision de base (cf. Annexes).

Ainsi, afin d'étudier le mécanisme de la transcription inverse du VIH impliqué dans la prolifération du virus, nous nous sommes intéressé à un fragment de 18 unités nucléosidiques, dPBS (séquence complémentaire ADN de la séquence *Primer Binding*

Conclusion et Perspectives

Site de l'ARN génomique du VIH), indispensable au virus lors de la rétro-transcription de l'ARN génomique du VIH en double brin d'ADN proviral. Le mécanisme de recrutement du brin complémentaire de (-)PBS, à savoir (+)PBS, lors de la transcription inverse a ainsi pu être étudié à l'échelle moléculaire. L'article relatif à cette étude est actuellement en cours de rédaction.

L'autre étude a concerné le mécanisme de réparation de l'ADN impliquant l'endonucléase VIII (*Nei*). Cette enzyme a la capacité de reconnaître et d'exciser spécifiquement des bases nucléiques endommagées, oxydées ou hydrogénées, en l'occurrence la 5,6-dihydrouridine (DHU). L'évolution des propriétés de fluorescence de notre rapporteur, positionné à proximité de cette base endommagée, nous a permis de suivre ce mécanisme étape par étape et de déterminer sa cinétique. Ces travaux ont abouti à la publication dans *PloS ONE* d'un article présenté en Annexes.^{126e}

“ **“** N'essaie pas ! Fais-le, ou ne le fais pas !
Il n'y a pas d'essai ! **”** **”** - Maître Yoda



CHAPITRE 3. Partie Expérimentale

SI PUBLICATION 1

Supporting materials

Characterization of UHRF1-induced base flipping in methylated DNA by using an environment sensitive fluorescent nucleobase

Vasyl Kilin^a, Nicolas Barthes^b, Benoît Michel^b, Christian Boudier, Valeriy Yashchuk, Marc Mousli,
Christian Bronner^b, Alain Burger^b, Yves Mély^{a,*}

^a Laboratoire de Biophotonique et Pharmacologie, UMR 7213 CNRS, Université de Strasbourg, Faculté de pharmacie, 74 route du Rhin, 67401 Illkirch, France

^b Institut de Chimie de Nice, UMR 7272 CNRS, Université de Nice Sophia Antipolis, Parc Valrose, 06108 Nice Cedex 2, France

^c Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), INSERM U964 CNRS UMR 7104, Université de Strasbourg, Illkirch, France

^d Kiev National Taras Shevchenko University, Kiev, Ukraine

* Corresponding author: yves.mely@unistra.fr Tel: +33-368854263

Thermal denaturation measurements

The melting curves were monitored by the absorbance changes at 260 nm and converted into a plot of α versus temperature, where α represents the fraction of strands in the duplex state. The melting temperatures were extracted by fitting the first derivative of these curves by a Gaussian (Table S1).

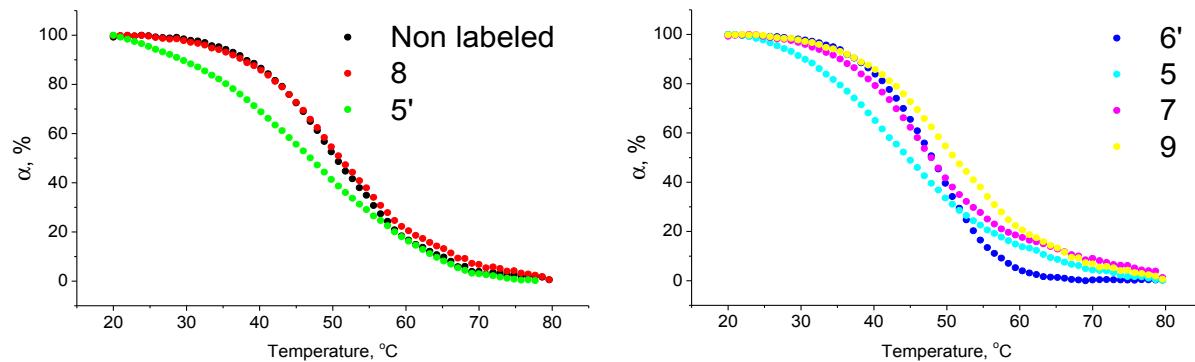


Figure S1. Thermal denaturation curves. The curves are expressed in fraction α of strands in the duplex state versus temperature for the non labeled duplex (A, black) and the duplexes labeled by 3HCnt at various positions (A and B). The concentration of duplexes was 2 μ M in 20 mM PBS buffer, 50 mM NaCl, pH = 7.4, EDTA 1 mM.

Table S1. Melting temperatures of 3HCnt-labeled duplexes

Position	Non labeled	5	7	8	9	5'	6'
Tm, °C	50	48	47	48	51	47	48
ΔTm , °C	-	2	2	2	-1	3	2

Quantum yield and N*/T* ratio measurements

Table S2. Quantum yield and N/T ratio of 3HCnt-labeled DNAs in the absence and the presence of SRA¹

Position	Methylation status	QY, % DNA free ²	QY, % SRA/DNA complex ²	$\frac{QY_{SRA/DNA}}{QY_{DNA}}$	N/T _{DNA} free	N/T _{SRA/DNA} complex	$\frac{N/T_{SRA/DNA}}{N/T_{DNA}}$
Free 3HC		5.5			0.84		
G3	HM	0.55	0.7	1.3	0.19	0.18	0.17
	NM	0.45	0.5	1.1	0.20	0.18	0.18
C5	HM	1.2	4.6	3.8	0.17	0.08	0.085
	NM	1.2	3.9	3.2	0.17	0.08	0.075
G7	HM	1.2	7.85	6.6	0.14	0.05	0.045

	NM	1.2	6.0	5.1	0.136	0.06	0.06
C8	HM	0.9	1.7	1.9	0.145	0.10	0.096
	NM	0.8	0.9	1.1	0.15	0.14	0.14
A9	HM	2.2	2.8	1.25	0.11	0.10	0.10
	NM	2.1	2.7	1.3	0.115	0.10	0.11
G5'	HM	0.5	1.4	2.7	0.19	0.06	0.06
	NM	0.5	0.7	1.45	0.19	0.14	0.14
G6'	HM	1.9	4.9	2.55	0.12	0.06	0.06
	NM	1.9	4.3	2.2	0.12	0.06	0.06

¹Values are average of 5 experiments. Standard deviations were equal or less than 20 %,

² Quantum yield for labeled DNA was calculated with free 2-thienyl-3-hydroxychromone in buffer as a reference.

Stoichiometry determination

Table S3. Apparent affinity constants of SRA and its G448D mutant for the 3HCnt-labeled DNAs

Position	Methylation status of C6	K _a SRA ($\times 10^{-6}$), M ⁻¹	$\frac{K_a \text{ labeled}}{K_a \text{ nonmodified}}$	K _a SRA-G448D ($\times 10^{-6}$), M ⁻¹
Non modified	HM	1.4	1.0	
G3	HM	1.4	0.98	
	NM	1.0		
C5	HM	20.0	14	
	NM	21		
G7	HM	6.1	4.3	
	NM	6.9		
C8	HM	2.5	1.8	0.24
	NM	0.55		0.13
A9	HM	1.8	1.25	
	NM	1.0		
G5'	HM	2.3	1.6	0.19
	NM	0.28		0.18
G6'	M	33	23	
	NM	7.9		

Values are average of 3 experiments. Standard deviations of values were equal or less than 20 %. K_a labeled/ K_a non labeled described the ratio of the apparent affinity constant of SRA for the labeled duplex as compared to the non labeled duplex.

CHEMISTRY

A EUROPEAN JOURNAL

Supporting Information

© Copyright Wiley-VCH Verlag GmbH & Co. KGaA, 69451 Weinheim, 2014

Rational Design of a Solvatochromic Fluorescent Uracil Analogue with a Dual-Band Ratiometric Response Based on 3-Hydroxychromone

Dmytro Dziuba,^[a, b] Iuliia A. Karpenko,^[a, c] Nicolas P. F. Barthes,^[a] Benoît Y. Michel,^[a] Andrey S. Klymchenko,^[d] Rachid Benhida,^[a] Alexander P. Demchenko,^[e] Yves Mély,^[d] and Alain Burger^{*[a]}

chem_201303399_sm_miscellaneous_information.pdf

Experimental section

1. Steady-state fluorescence measurements

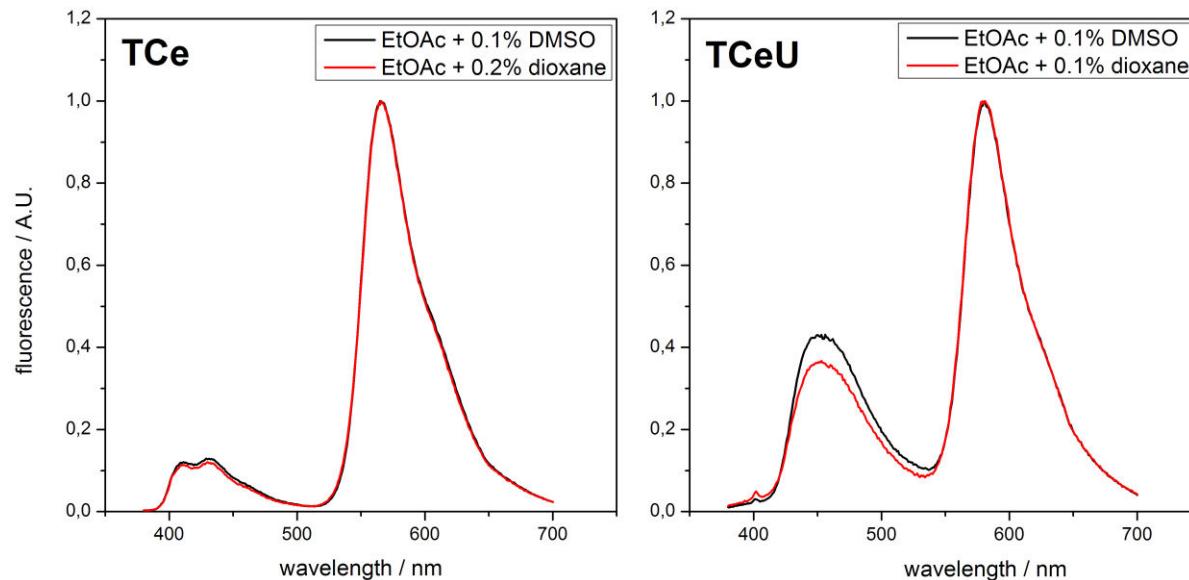


Figure S1. Influence of basic solvents on the relative intensity of N* band for **TCe** and **TCeU**.

Table S1. Spectroscopic properties of 3-hydroxychromones **TC**, **TCe**, and **TCeU**.

Solvent	$\lambda_{\text{abs}}^{\text{a}}$			$\lambda_{N^*}^{\text{b}}$			$\lambda_{T^*}^{\text{c}}$			$I_{N^*}/I_{T^*}^{\text{d}}$			QY ^e		
	TC	TCe	TCeU	TC	TCe	TCeU	TC	TCe	TCeU	TC	TCe	TCeU	TC	TCe	TCeU
H ₂ O	362	386	401	430	— ^f	500	509	— ^f	— ^g	1.80	— ^f	— ^g	0.04	< 0.01	0.08
MeOH	359	378	395	424	442	501	540	561	558	0.72	0.43	1.20	0.07	0.13	0.23
EtOH	360	379	397	424	439	488	542	563	574	0.34	0.22	0.63	0.08	0.14	0.18
BuOH	363	381	400	424	440	481	544	565	571	0.27	0.22	0.50	0.11	0.18	0.19
CH ₃ CN	354	373	392	414	431	472	540	563	576	0.07	0.10	0.54	0.08	0.14	0.15
DMSO	362	381	402	426	432	477	547	568	582	0.10	0.06	1.43	0.08	0.12	0.25
Acetone	354	374	393	412	413, 430	465	544	566	581	0.11	0.18	0.53	0.08	0.15	0.15
CH ₂ Cl ₂	357	379	397	412	432	458	537	561	577	0.02	0.05	0.31	0.30	0.19	0.17
EtOAc	354	374	394	412	411, 430	451	543	566	579	0.05	0.12	0.36	0.12	0.15	0.15
THF	357	377	396	415	413, 434	448	547	568	585	0.13	0.45	0.73	0.14	0.18	0.18
PhCH ₃	358	379	399	412	431	441, 467	546	572	587	0.01	0.03	0.13	0.24	0.20	0.15

Footnotes: a) position of the absorption maximum; b) position of the N* band maximum; c) position of the T* band maximum; d) the ratio of two intensity maxima; e) quantum yield determined using 3-hydroxyflavone in toluene ($\Phi = 0.29$) as a reference;¹ f) impossible to determine precisely due to low quantum yield; g) a single emission band was observed.

¹ A. S. Klymchenko,; T. Ozturk; V. G. Pivovarenko, A. P. Demchenko, *Can. J. Chem.* **2001**, 79, 358–363.

Table S2. Spectroscopic properties of 3-methoxychromones **MTC**, **MTCe** and **MTCeU**.

Solvent	$\lambda_{\text{abs}}^{\text{a}}$			$\lambda_{\text{em}}^{\text{b}}$			QY ^c		
	MTC	MTCe	MTCeU	MTC	MTCe	MTCeU	MTC	MTCe	MTCeU
H ₂ O	346	370	375	442	451	511	0.02	0.03	0.10
MeOH	341	366	387	434	441	501	0.01	0.13	0.49
EtOH	342	367	387	426	442	493	0.02	0.15	0.53
BuOH	344	368	388	434	442	489	0.03	0.18	0.45
CH ₃ CN	338	363	383	430, 444	442	466	0.01	0.03	0.43
DMSO	343	368	389	456	436	486	0.06	0.05	0.56
Acetone	337	362	384	447	449, 470	465	0.02	0.04	0.43
CH ₂ Cl ₂	341	366	386	426, 448	446	454	0.02	0.09	0.36
EtOAc	337	362	383	448	449, 474	451	0.02	0.04	0.34
THF	338	364	386	446, 546	448, 473	454	0.01	0.03	0.37
PhCH ₃	340	366	388	451, 546	449, 545	438, 457, 538	0.03	0.08	0.37

Footnotes: a) position of the absorption maximum; b) position of the emission maximum; c) yield determined using 3-hydroxyflavone in toluene ($\Phi = 0.29$) as a reference.¹

3. Colour figures of the manuscript

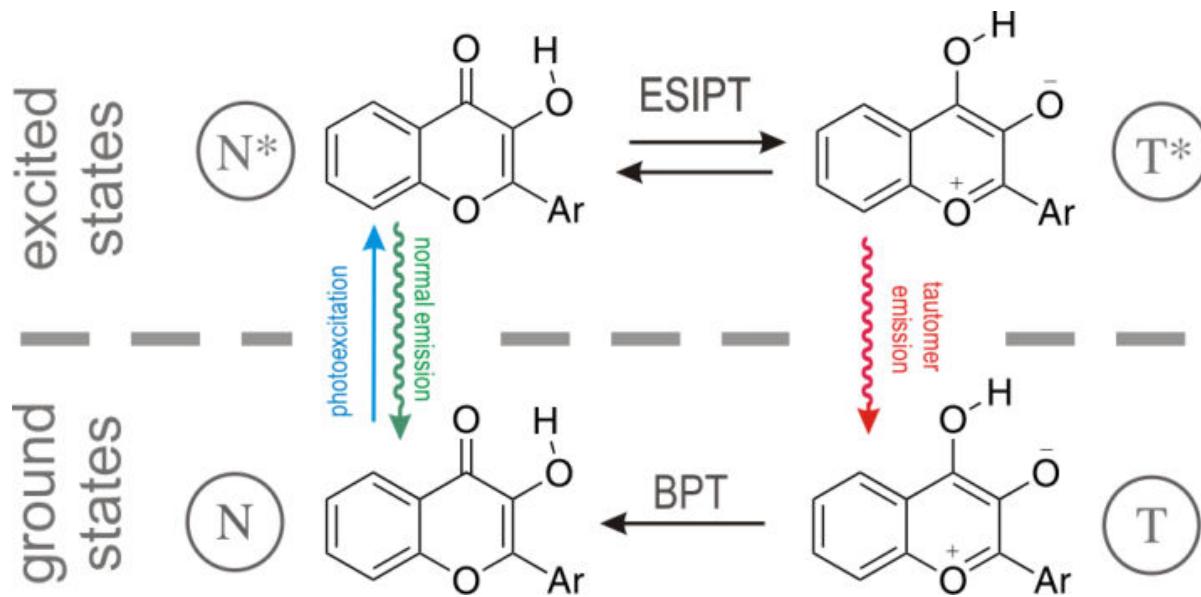


Figure 1. Origin of the dual emission of 3-hydroxychromones (ESIPT – excited state intramolecular proton transfer; BPT – ground state back proton transfer; N and T denote the normal and tautomer forms respectively).

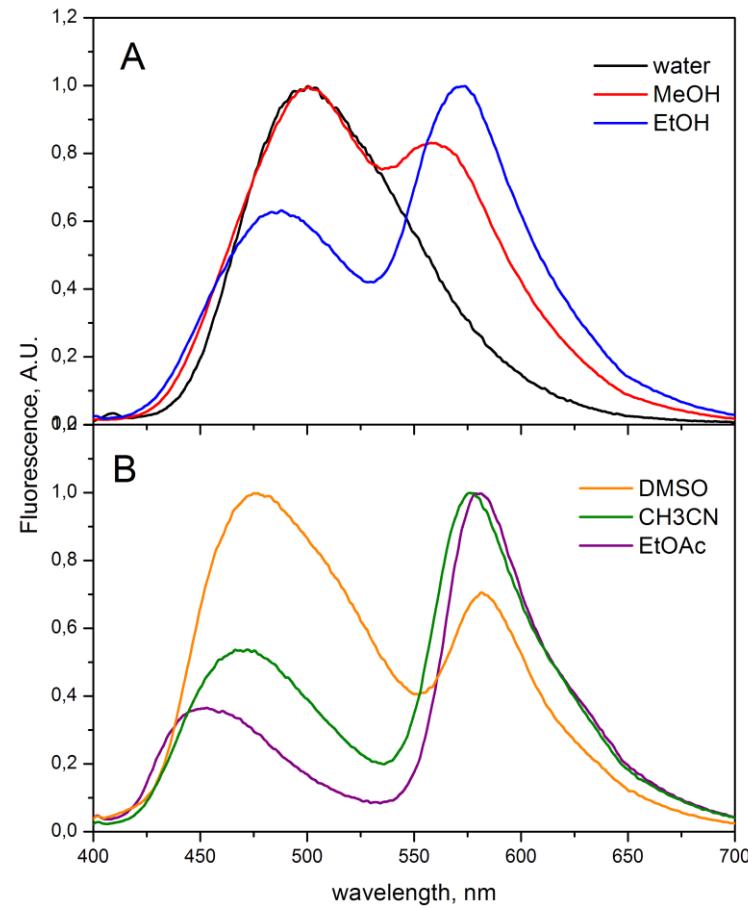


Figure 4. Fluorescence spectra of TCeU in protic (A) and aprotic (B) solvents.

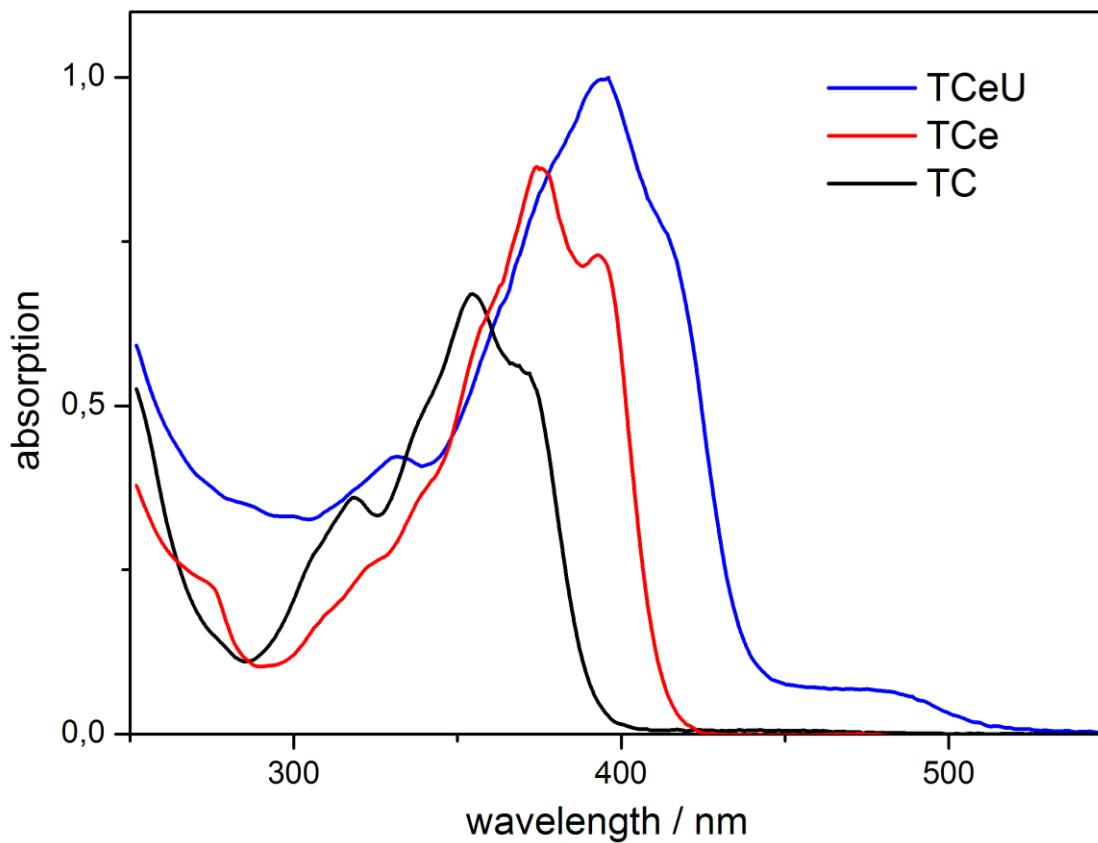


Figure 5. Absorption spectra of compounds **TC** (black line), **TCe** (red) and **TCeU** (blue) in ethyl acetate.

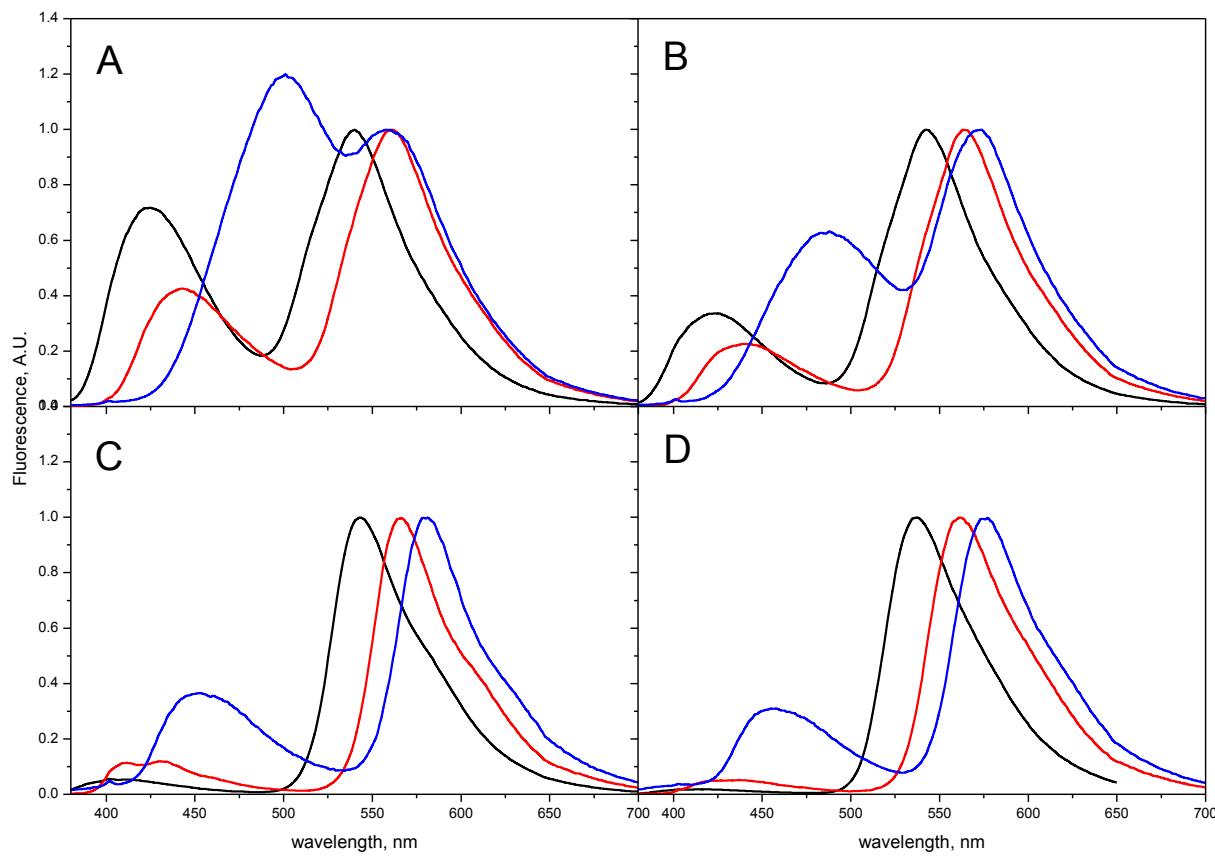


Figure 6. Normalized fluorescence spectra of **TC** (black line), **TCe** (red), and **TCeU** (blue) in methanol (A), ethanol (B), ethyl acetate (C) and dichloromethane (D).

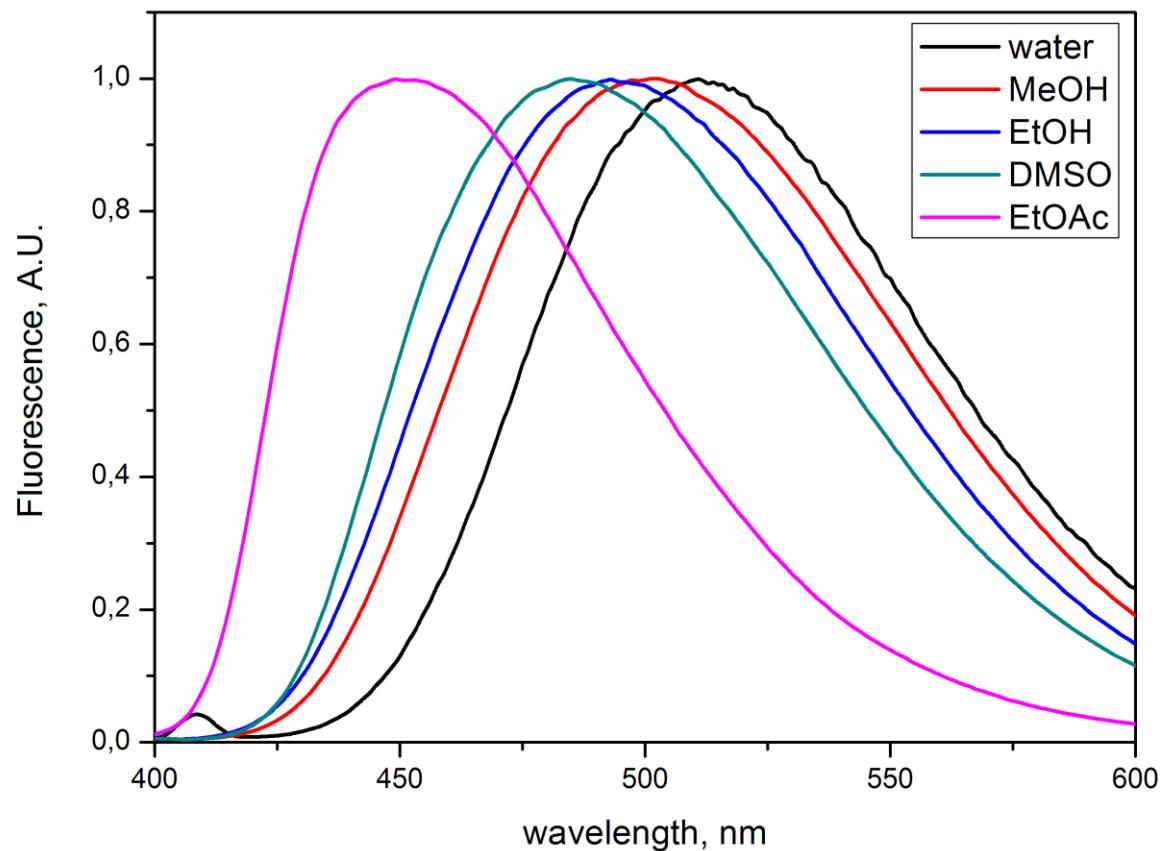


Figure 7. Normalized fluorescence spectra of MTCeU in representative solvents of different polarities.

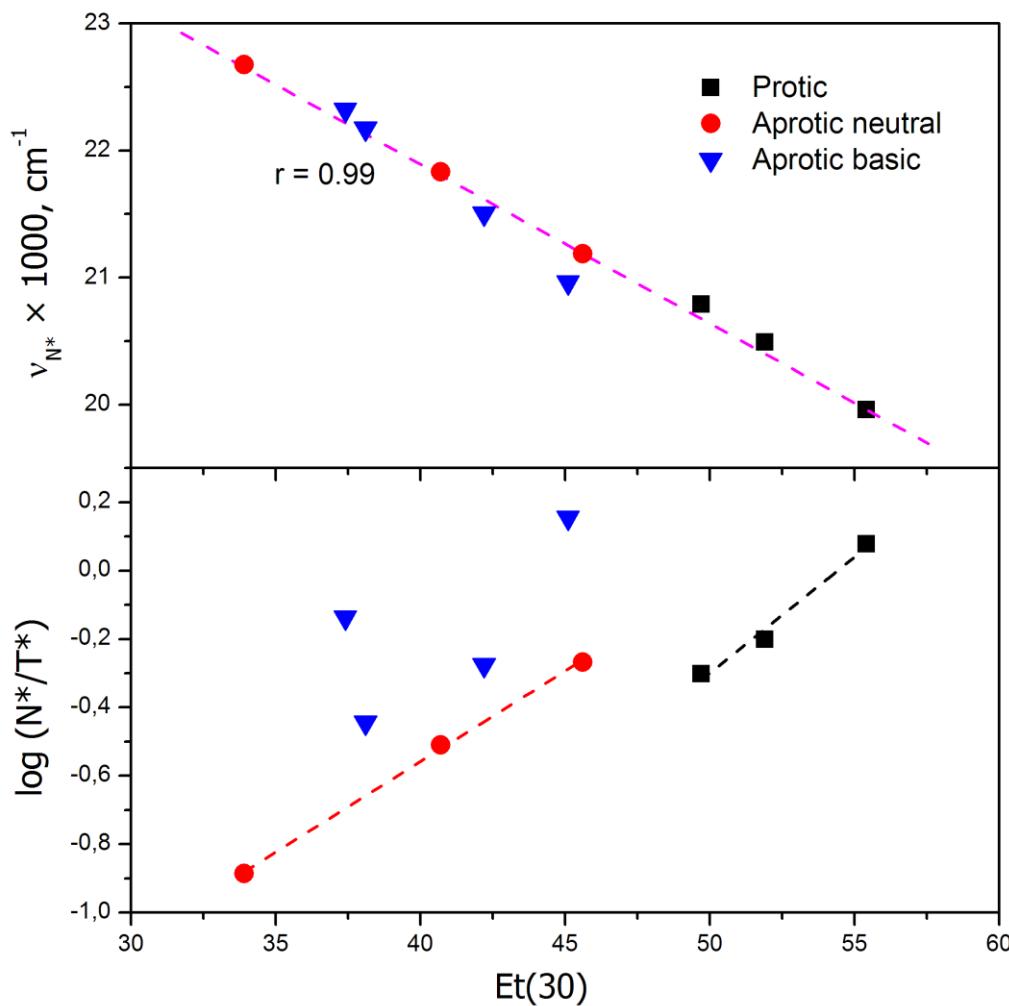


Figure 8. Dependence of the spectroscopic characteristics of **TCeU** on the solvent polarity index Et(30). Position of the N^* band (top) and logarithm of the N^*/T^* ratio (bottom) as a function of Et(30). The black squares, red circles and blue triangles represent protic (MeOH, EtOH, BuOH), aprotic neutral (CH_3CN , CH_2Cl_2 , PhCH_3) and aprotic H-bond basic solvents (DMSO, Acetone, EtOAc, THF), respectively.

SI PUBLICATION 3

Supporting Information To:

**Development of
environmentally sensitive fluorescent
and dual emissive deoxyuridine analogues**

Nicolas P. F. Barthes,^a Iuliia A. Karpenko,^{a,b} Dmytro Dziuba,^{a,c} Marie Spadafora,^a Jérémie Auffret,^a Alexander P. Demchenko,^d Yves Mély,^e Rachid Benhida,^a Benoît Y. Michel,*^a Alain Burger*^a

a) Institut de Chimie de Nice, UMR 7272, Université de Nice Sophia Antipolis, CNRS, Parc Valrose, 06108 Nice Cedex 2, France.

b) Present address: Ecole Polytechnique Fédérale de Lausanne, Institute of Chemical Sciences and Engineering, Laboratory of Protein Engineering, BCH, Av. Forel 2, CH-1015 Lausanne, Switzerland.

c) Present address: Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo nam. 2, CZ-16610, Prague 6, Czech Republic.

d) A.V. Palladin Institute of Biochemistry, 9 Leontovicha Street, Kiev 01030, Ukraine.

e) Laboratoire de Biophotonique et Pharmacologie, UMR 7213, Faculté de Pharmacie, Université de Strasbourg, CNRS, 74 Route du Rhin, 67401 Illkirch, France.

Corresponding author: burger@unice.fr

Experimental section

General methods:

All non-aqueous reactions involving water-sensitive reagents were performed in oven dried glassware under argon using dry solvents. The synthetic intermediates were coevaporated twice with toluene beforehand and dried in vacuo before use. All chemical reagents were obtained from commercial sources and were used as supplied. Anhydrous solvents were obtained according to standard procedures.¹ The reactions were monitored by thin-layer chromatography (TLC, Merck silica gel 60 F254 plates) and visualized both by UV radiation (254 & 365 nm) and by spraying with phosphomolybdic acid in ethanol followed by a subsequent warming with a heat gun. Column chromatography² was performed with flash silica gel (40–63 mm). Analytical HPLC were performed using a Waters™ 600 (600 pump, 600E Controller, 996 Photodiode Array Detector) and a Jasco LC-Net II / ADC apparatus. A RP C18 column (300 × 4.60 mm, 5 µm particle size, Luna® 100Å, Phenomenex®) was used with mobile phases: A: H₂O (0.1 % HCOOH) and B: CH₃CN (0.1 % HCOOH) at a flow rate of 0.5 mL/min (see SI page 109 for the control method). The NMR spectra (¹H, ¹³C, ²D; ¹H–¹³C COSY, ¹H–¹³C HMQC and ¹H–¹³C HMBC) were recorded on 200 or 500 Bruker Advance Spectrometers (200 or 500 MHz). ¹H NMR (200 and 500 MHz), ¹³C NMR (50 and 125 MHz, recorded with complete proton decoupling) spectra were obtained with samples dissolved in CDCl₃, CD₃OD, or DMSO-*d*⁶, with the solvent signals as internal references: 7.26 ppm for CHCl₃, 3.31 ppm for CD₂OD, 2.50 ppm for (CD₃)(CD₂H)S(O) for ¹H NMR experiments, and 77.0 ppm for CDCl₃, 49.0 ppm for CD₃OD, 39.4 ppm for (CD₃)₂S(O) for ¹³C NMR experiments.³ Chemical shifts (δ) are given in ppm to the nearest 0.01 (1H) or 0.1 ppm (¹³C). The coupling constants (J) are given in Hertz (Hz). The signals are reported as follows: (s=singlet, d=doublet, t=triplet, m=multiplet, br=broad). Assignments of ¹H and ¹³C NMR signals were achieved with the help of D/H exchange, COSY, DEPT, APT, HMQC, HSQC, HMBC experiments. Regular mass spectra (MS) were recorded on an Esquire 3000 Plus apparatus with ESI in both positive and negative mode. High-resolution mass spectrometry was conducted with a FINIGAN MAT 95 spectrometer with EI or ESI ionization techniques. All solvents for absorption and fluorescence experiments were of spectroscopic grade. Stock solutions were prepared using 1,4-dioxane. The samples used for spectroscopic measurements contained \approx 0.1% v/v of the stock solvent. The absorption spectra were recorded on a Cary 300 Scan spectrophotometer (Varian) using 1cm quartz cells at 20 °C. The fluorescence spectra were recorded on a FluoroMax 4.0 spectrofluorometer (Jobin Yvon, Horiba). They were recorded using excitation and emission slits of 2 nm and were corrected at excitation and emission. They were taken with absorbance less than 0.05 at 20 °C at the excitation wavelength mentioned in the corresponding experiments. The quantum yields were determined taking into account the change of the refractive index of the solvents.

¹ W. L. F. Armarego and C. L. L. Chai, *Purification of Laboratory Chemicals*, 7th ed.; Butterworth-Heinemann: Oxford, 2013.

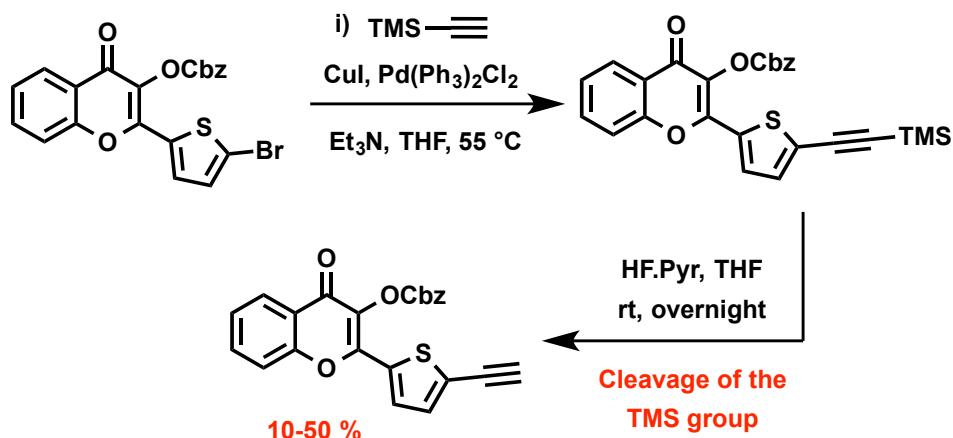
² W. C. Still, M. Kahn and A. Mitra, *J. Org. Chem.*, 1978, **43**, 2923–2925.

³ (a) H. E. Gottlieb, V. Kotlyar and A. Nudelman, *J. Org. Chem.*, 1997, **62**, 7512–7515; (b) G. R. Fulmer, A. Miller, N. H. Sherden and H. E. Gottlieb, *Organometallics*, 2010, **29**, 2176–2179.

Supporting Information

1. Additional Data

Preparation of ethynyl 3HCs



Scheme S1. Cleavage of the TMS group, exhibiting a lack of reproducibility due to a partial and concomitant removal of Cbz as well as a decomposition of the anionic form.

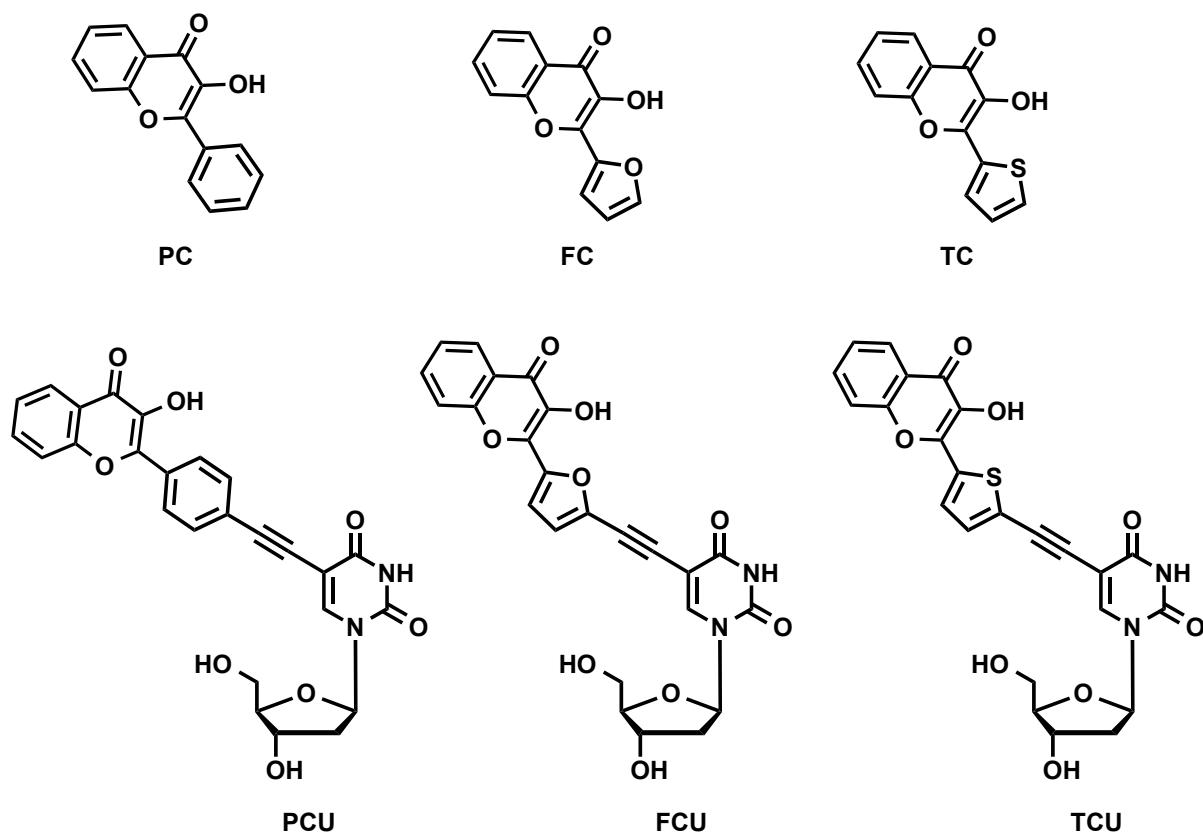


Figure S1. Considered 3-hydroxychromones for the spectroscopic comparison listed in Table 1: parent 2-aryl-3-HCs (**PC**, **TC** and **TC**) and corresponding nucleosidic conjugates (**PCU**, **FCU** and **TCU**) are depicted on the first and second rows, respectively.

2. Steady-state fluorescence measurements

PCU

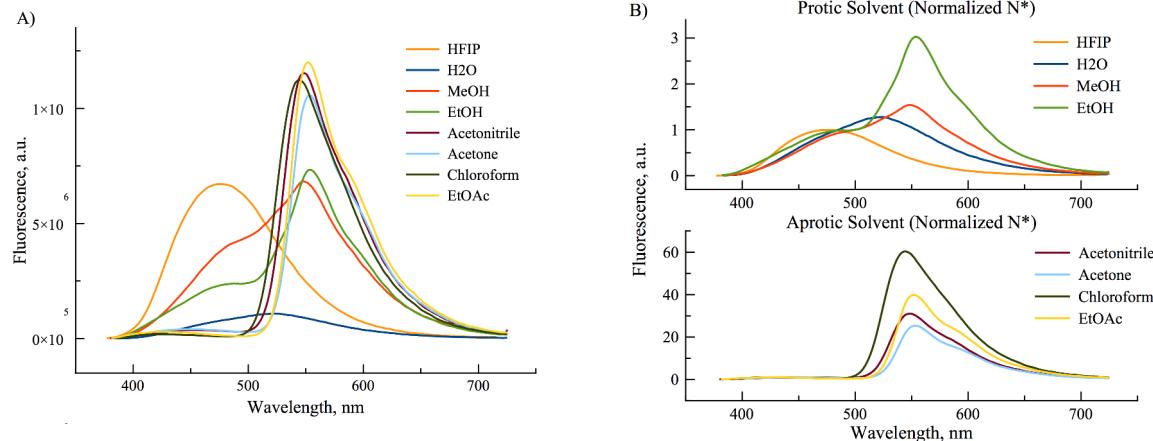


Fig. S2 A) Fluorescence spectra in the set of investigated solvents; B) Normalized emission spectra by N* band for protic (top) and aprotic (bottom) solvents.

FCU

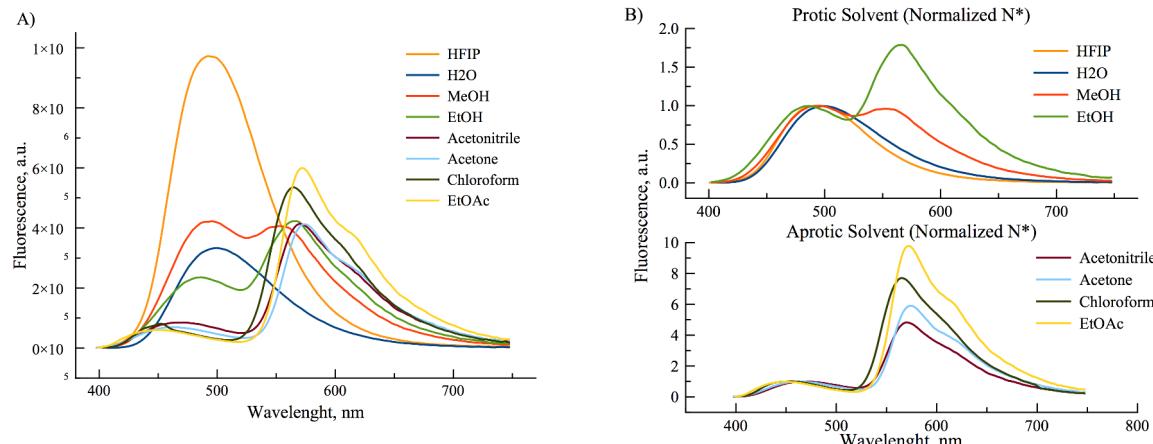


Fig. S3 A) Fluorescence spectra in the set of investigated solvents; B) Normalized emission spectra by N* band for protic (top) and aprotic (bottom) solvents.

TCU

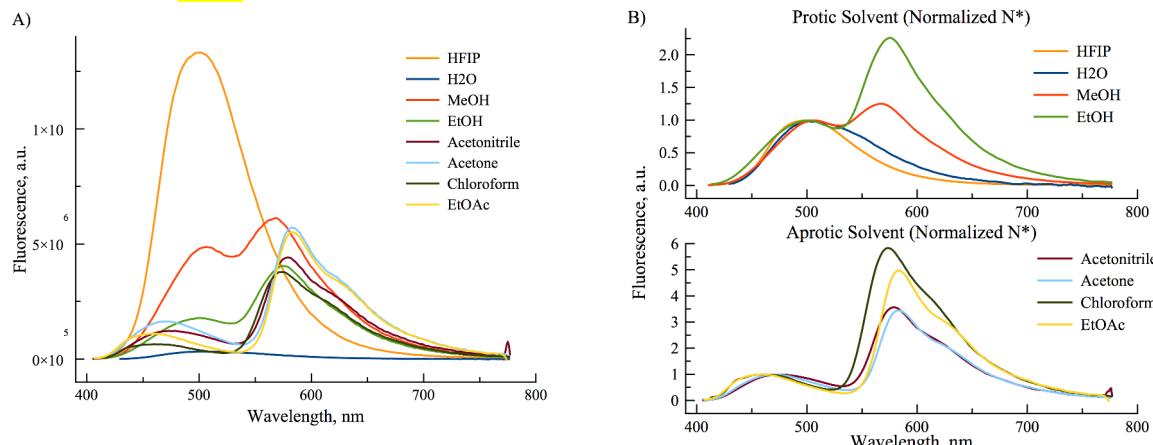


Fig. S4 A) Fluorescence spectra in the set of investigated solvents; B) Normalized emission spectra by N* band for protic (top) and aprotic (bottom) solvents.

Table S1: Lippert-Mataga data of **FCU** and **TCU**

Solvents	Δf^a	$\Delta v \text{ (cm}^{-1}\text{)}^b$	
		FCU	TCU
HFIP ^c	0.3118	5690	5253
H ₂ O	0.3201	5535	5396
MeOH	0.3087	5719	5475
EtOH	0.2888	5348	5166
CH ₃ CN	0.3055	4719	4544
Acetone	0.2842	4533	4188
CHCl ₃	0.1482	3251	3307
EtOAc	0.1998	4083	3789

Footnote: ^a Lippert's Parameter: $\Delta f = \left(\frac{\varepsilon - 1}{2\varepsilon + 1} - \frac{n^2 - 1}{2n^2 + 1} \right)$. In this equation, n is the refractive index and ε is the dielectric constant (values found in: R. C. Weast., M. J. Astle.; CRC Handbook of Chemistry and Physics, 60th ed.; CRC: Boca Raton, Florida, 1980, p:2528) ^b Stock shift values, ^c Value of ε found in Hong D. P., Hoshino M., Kuboi R., Goto Y. (1999) *J. Am. Chem. Soc.* **121**, 8427–8433.

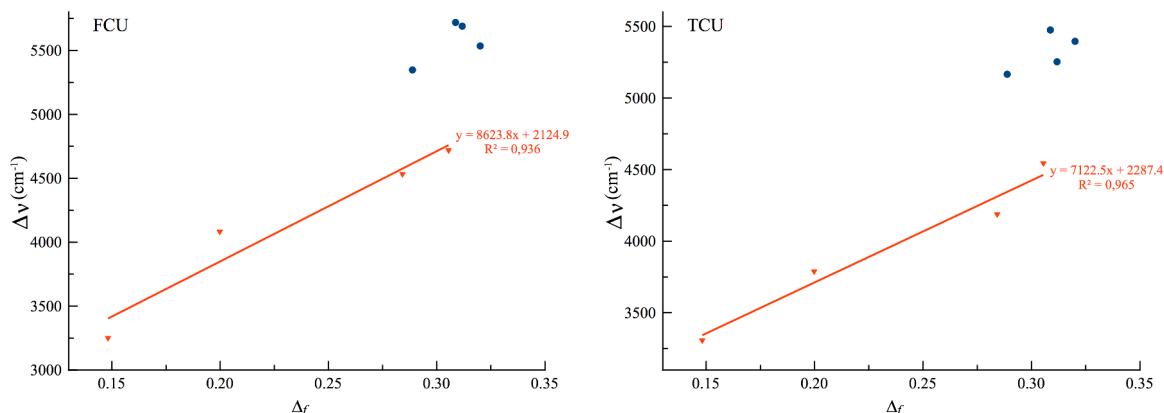


Fig. S5 & S6: Stokes shift vs. orientation polarizability function (Δf) for **FCU** and **TCU**. The linear plots shown were obtained with aprotic solvents according to the Lippert's equation (red triangles and red line). The positions of the points for protic solvents are shown in blue dots.

Lippert's Equation: $v_A - v_F = \frac{2}{hc} \Delta f \frac{(\mu_e - \mu_g)^2}{a^3}$.

In this equation, h ($= 6.6256 \times 10^{-27}$ ergs) is the Planck's constant, c ($= 2.9979 \times 10^{10}$ cm/s) is the speed of light, and a is the radius of the Onsager's cavity. v_A and v_F are the wavenumbers (cm^{-1}) of the absorption and emission, respectively, Δv is the Stokes shift (cm^{-1}), Δf is the Lippert's parameter, and μ_e and μ_g are the dipole moments of the excited and ground states, respectively.

Remark: Exploiting the Lippert's equation to get the difference of the dipole moments ($\mu_e - \mu_g$) needs to determine the Onsager radius. In the Lippert-Mataga model the fluorophore is treated as a solute occupying a spherical volume for which the radius corresponds to the Onsager radius. Measuring the distance at the two extremes of the fluorophore and dividing the distance by two determine the Onsager radius. In general this distance corresponds to the distance between the donor and acceptor groups of the fluorophore. Applying this approximation to **FCU** and **TCU** is not satisfactory because they are extended conjugated fluorophores with the acceptor group, the carbonyl group, which is not positioned at one extreme.

3. Hydration Study

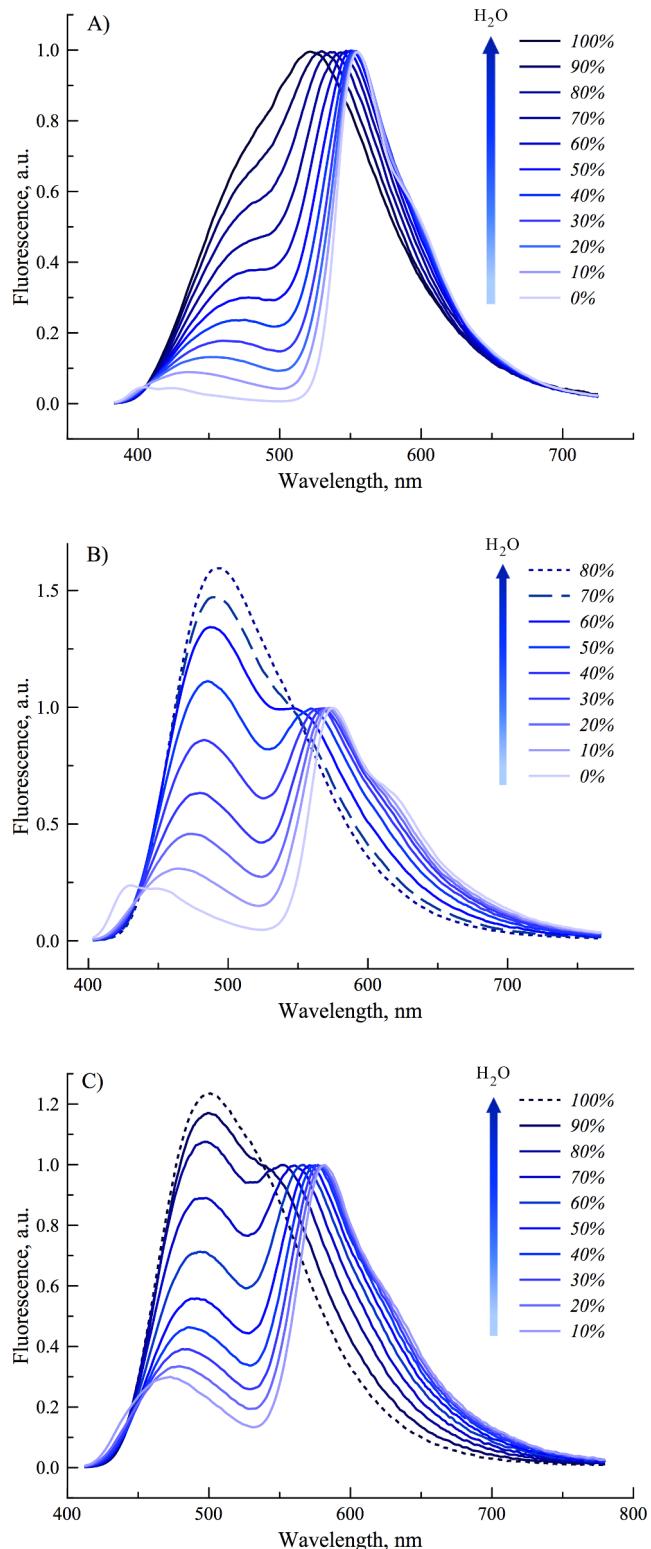


Fig. S7 Sensitivity of the emission spectra to hydration for A) PCU, B) FCU and C) TCU. Data were recorded in gradual mixtures of dioxane and water. Emission spectra were normalized by T* band.

Table S2. Spectroscopic properties of 3-hydroxychromone uridine analogs **PCU**, **FCU**, and **TCU** in dioxane / water and acetonitrile / water mixtures.

% Water	Acetonitrile								Dioxane									
	PCU			FCU			TCU		PCU			FCU			TCU			
	λ_{N^*} ^a	λ_{T^*} ^b	I_{N^*}/I_{T^*} ^c	λ_{N^*} ^a	λ_{T^*} ^b	I_{N^*}/I_{T^*} ^c	λ_{N^*} ^a	λ_{T^*} ^b	I_{N^*}/I_{T^*} ^c	λ_{N^*} ^a	λ_{T^*} ^b	I_{N^*}/I_{T^*} ^c	λ_{N^*} ^a	λ_{T^*} ^b	I_{N^*}/I_{T^*} ^c	λ_{N^*} ^a	λ_{T^*} ^b	I_{N^*}/I_{T^*} ^c
100	480	522	0.74	495	- ^d	- ^d	501	- ^d	- ^d	480	521	0.75	497	- ^d	- ^d	500	- ^d	- ^d
90	481	529	0.67	495	- ^d	- ^d	501	538	1.15	480	528	0.66	494	- ^d	- ^d	497	539	1.17
80	481	533	0.56	495	- ^d	- ^d	497	547	1.07	480	539	0.57	494	- ^d	- ^d	498	552	1.08
70	180	539	0.46	492	539	1.40	498	555	0.93	480	542	0.47	490	539	1.37	494	561	0.89
60	478	542	0.39	492	541	1.35	496	560	0.77	480	546	0.38	489	544	1.32	493	567	0.72
50	477	543	0.34	491	546	1.25	496	565	0.65	476	550	0.31	487	559	1.11	490	571	0.56
40	478	546	0.31	490	555	1.13	494	566	0.56	467	551	0.24	483	562	0.84	485	574	0.46
30	478	548	0.28	490	556	0.98	492	568	0.47	461	553	0.18	478	568	0.64	481	578	0.39
20	474	548	0.23	488	562	0.79	494	572	0.39	451	554	0.14	475	570	0.47	479	579	0.34
10	472	550	0.17	483	565	0.53	487	571	0.30	439	554	0.09	466	573	0.32	473	581	0.30
0	454	550	0.03	466	573	0.20	473	578	0.27	422	554	0.05	448	573	0.34	461 ^e	584	- ^e

Footnotes: a) position of the N* emission band maximum; b) position of the T* emission band maximum; c) ratio of two intensity maxima for the N* and T* emission bands;

d) single emission band was observed; e) unexplained additional emission band was observed at 438 nm.

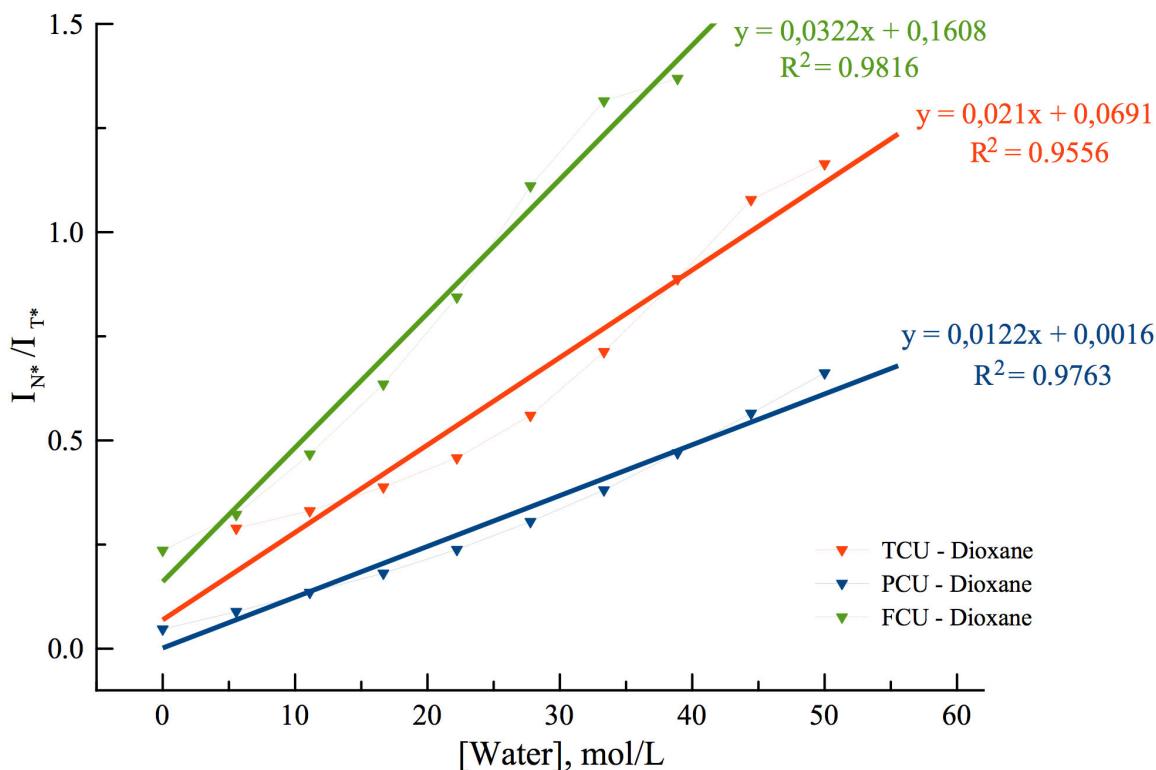


Fig. S8 Dependence of the ratiometric response (I_{N^*}/I_{T^*}) of **PCU** (blue), **FCU** (green) and **TCU** (orange) on the water concentration in dioxane. The ratios were extracted from the spectra in Fig. S4.

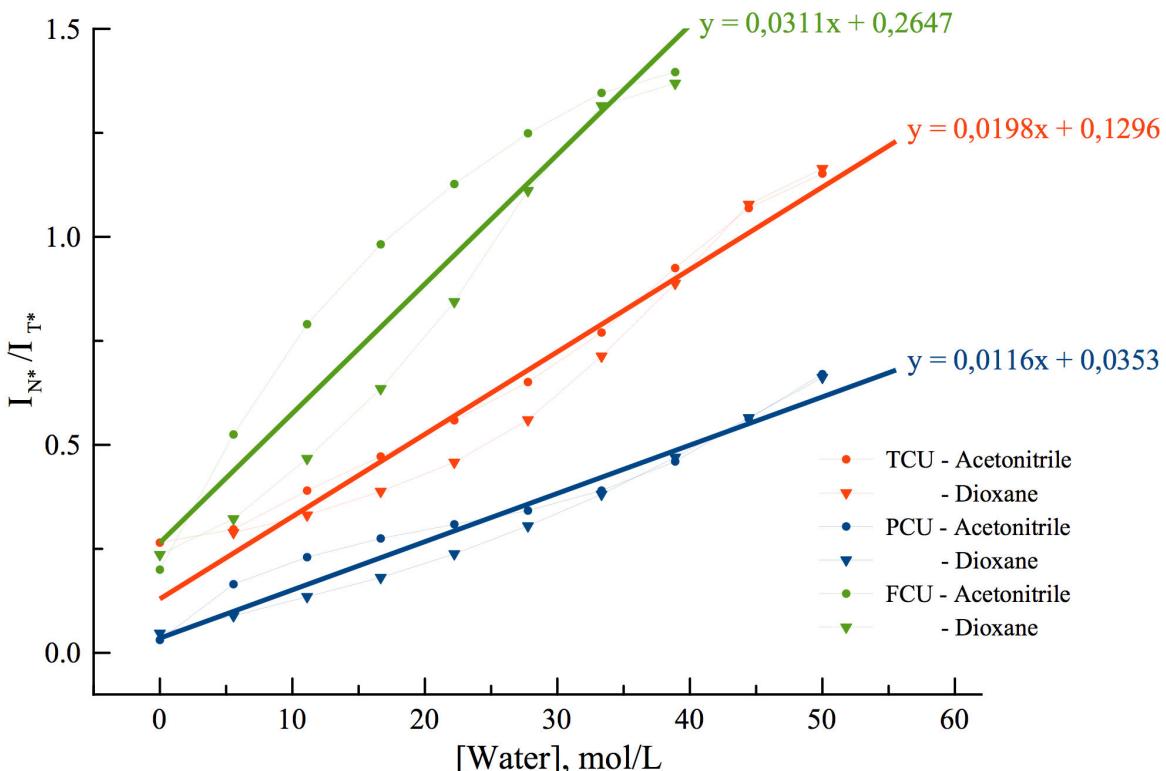
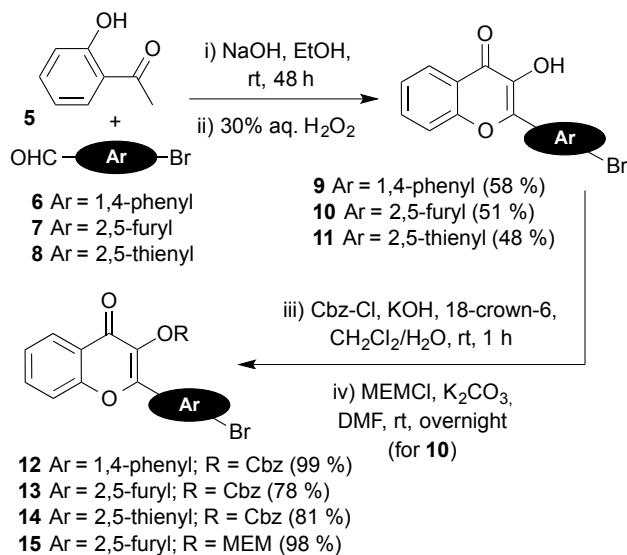


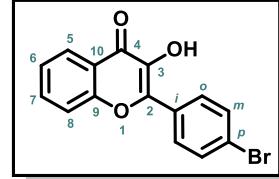
Fig. S9 Summary of Fig 5 & S5: correlation between the ratiometric response (I_{N^*}/I_{T^*}) and the water concentration in dioxane and acetonitrile.

4. Synthetic procedures

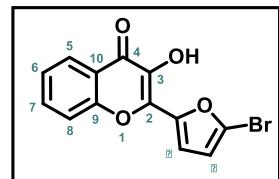
Preparation of the 3HC fluorophores



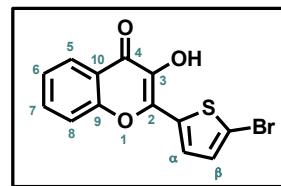
2-(4-Bromophenyl)-3-hydroxy-chromen-4-one (9): To a stirred solution of *o*-hydroxyacetophenone **5** (1.16 mL, 9.65 mmol) and 4-bromobenzaldehyde **6** (1.79 g, 9.65 mmol) in ethanol (50 mL) was dropwise added a 5 N NaOH solution (6 mL). The reaction mixture was stirred 48 h at rt before a dropwise addition of 30 % aq. hydrogen peroxide solution (2 mL). The resulting mixture was stirred 1 day at rt, and then poured into cold water (300 mL) and acidified with 1 N HCl to pH 3. The resulting precipitate was filtered and thoroughly washed with water and cyclohexane to provide as a yellow solid the target derivative **9** (1.78 g, 58 %). C₁₅H₉BrO₃ (317.13). R_f = 0.49 (cyclohexane/EA = 7:3) ¹H-NMR (CDCl₃, 200 MHz): 7.09 (s, 1H, OH), 7.43 (td, ³J=7.0 Hz, ⁴J=1.0 Hz, H₆), 7.59 (d, ³J=8.0 Hz, 1H, H₈), 7.66 (d, ³J=8.8 Hz, 1H, H_m-Ph), 7.73 (ddd, ³J=8.0, 7.0 Hz, ⁴J=1.4 Hz, 1H, H₇), 8.15 (d, ³J=8.8 Hz, 2H, H_o-Ph), 8.25 (dd, ³J=8.0 Hz, ⁴J=1.4 Hz, 1H, H₅); ¹³C-NMR (DMSO-d₆, 50 MHz): δ = 118.3 (C₈), 121.2 (C₁₀), 123.2 (C_p-Br), 124.5 (C₆), 124.7 (C₅), 129.4 (C_o-Ph), 130.5 (C_r-Ph), 131.5 (C_m-Ph), 133.7 (C₇), 139.4 (C₃), 143.9 (C₂), 154.4 (C₉), 172.9 (C₄); ¹³C-NMR (CDCl₃, 50 MHz): δ = 118.2 (C₈), 120.6 (C₁₀), 124.7 (C₆), 125.5 (C₅), 129.2 (C_m-Ph), 130.0 (C_p-Br), 131.9 (C_o-Ph), 133.8 (C₇), 138.5 (C₃), 143.8 (C_r-Ph), 155.3 (C₉), 157.3 (C₂), 173.4 (C₄); MS (ESI⁺, MeOH): *m/z*: 339.0, 341.0 [M+Na]⁺.



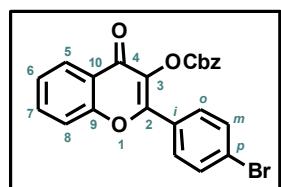
2-(5-Bromofur-2-yl)-3-hydroxy-chromen-4-one (10): To a stirred solution of *o*-hydroxyacetophenone **5** (3.0 mL, 24.9 mmol) and 5-bromo-2-furanecarbaldehyde (fufural) **7** (4.37 g, 24.9 mmol) in ethanol (45 mL) was dropwise added a 5 N NaOH solution (9 mL). The reaction mixture was stirred 48 h at rt before a dropwise addition of 30 % aq. hydrogen peroxide solution (17 mL). The resulting mixture was stirred 30 min at rt, and then poured into cold water (300 mL) and acidified with 1 N HCl to pH 3. The resulting precipitate was filtered and thoroughly washed with water and cyclohexane to provide as a yellow solid the target derivative **10** (3.90 g, 51 %). C₁₃H₇BrO₄ (307.10). R_f = 0.33 (CyHex/EA = 7:3); ¹H-NMR (DMSO-d₆, 200 MHz): δ = 6.89 (d, 1H, ³J=2.4 Hz, H_B-fur), 7.24 (d, 1H, ³J=2.4 Hz, H_A-fur), 7.45 (dd, 1H, ³J=6.8, 5.8 Hz, H₆), 7.64–7.84 (m, 2H, H₇, H₈), 8.09 (d, 1H, ³J=7.6 Hz, H₅), 10.23 (s, 1H, 3-OH); 6.74 (d, 1H, ³J=3.4 Hz, H_B-fur), 7.28–7.36 (m, 1H, H₆), 7.33 (d, 1H, ³J=3.4 Hz, H_A-fur), 7.58–7.66 (m, 2H, H₈, H₇), 8.05 (d, 1H, ³J=7.8 Hz, H₅); ¹³C-NMR (DMSO-d₆, 50 MHz): δ = 112.5 (C_A), 114.2 (C_B), 117.9 (C₈), 121.1 (C₁₀), 121.8 (Fur-Br), 123.0 (C₆), 124.8 (C₅), 131.6 (C₇), 137.8 (C₃), 149.5 (C₂-Fur), 149.6 (C₂), 153.4 (C₉), 176.2 (C₄); MS (ESI⁺, MeOH): *m/z*: 329.1, 331.0 [M+Na]⁺.



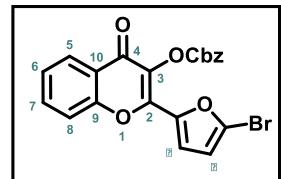
2-(5-Bromothien-2-yl)-3-hydroxy-chromen-4-one (11): To a stirred solution of *o*-hydroxyacetophenone **5** (1.16 mL, 9.65 mmol) and 5-bromo-2-thiophenecarbaldehyde **8** (1.15 mL, 9.65 mmol) in ethanol (50 mL) was dropwise added a 5 N NaOH solution (6 mL). The reaction mixture was stirred 48 h at rt before a dropwise addition of 30 % aq. hydrogen peroxide solution (2 mL). The resulting mixture was stirred 5 h at rt, and then poured into cold water (300 mL) and acidified with acetic acid to pH 4. The resulting precipitate was filtered and thoroughly washed with water and cyclohexane to provide as a yellow solid the target derivative **11** (1.79 g, 57 %). $C_{13}H_7BrO_3S$ (323.16). $R_f = 0.39$ (cyclohexane/EA = 7:3); 1H -NMR ($CDCl_3$, 200 MHz): 7.19 (d, $^3J=4.0$ Hz, 1H, H β -thioph.), 7.41 (td, $^3J=8.0$ Hz, $^4J=0.8$ Hz, 1H, H6), 7.54 (d, $^3J=8.4$ Hz, 1H, H8), 7.71 (ddd, $^3J=8.4$ Hz, $^3J=8.0$ Hz, $^4J=1.4$ Hz, 1H, H7), 7.73 (d, $^3J=4.0$ Hz, 1H, H α -thioph.), 8.22 (dd, $^3J=8.0$ Hz, $^4J=1.4$ Hz, 1H, H5); ^{13}C -NMR ($DMSO-d^6$, 50 MHz): $\delta = 112.2$ (Th-Br), 117.8 (C8), 121.6 (C α), 122.0 (C10), 122.9 (C6), 124.7 (C5), 129.3 (C β), 131.6 (C7), 136.1 (C2-Th), 141.6 (C3), 146.5 (C2), 153.3 (C9), 176.0 (C4); MS (ESI $^+$, MeOH): m/z : 345.0, 347.0 [M+Na] $^+$.



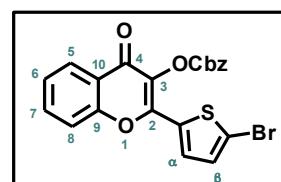
3-Benzoyloxycarbonyloxy-2-(4-bromophenyl)-chromen-4-one (12): To a stirred suspension of **9** (500 mg, 1.58 mmol) in a mixture of CH_2Cl_2 (8 mL) and toluene (8 mL) at 0 °C, were added 18-crown-6 (7 mol %, 30 mg), a 25 % w/w aq. NaOH solution (1.6 mL) and benzyl chloroformate (338 μ L, 2.37 mmol). The reaction mixture was stirred overnight at rt. After carefully neutralizing by addition of 1 N aq. HCl (10 mL), the organic layer was extracted with CH_2Cl_2 (3 x), washed with brine, dried over $MgSO_4$, filtered and the volatiles were removed *in vacuo*. The residue was purified by flash chromatography on silica gel eluted with toluene/Et₂O mixture (98:2 → 85:15, v/v) to provide the desired compound **12** as a yellow solid (705 mg, 99 %). $C_{23}H_{15}BrO_5$ (451.27). $R_f = 0.58$ (toluene/Et₂O = 9:1); 1H -NMR ($CDCl_3$, 200 MHz): $\delta = 5.17$ (s, 2H, CH₂), 7.22–7.33 (m, 6H, H₆ & Cbz), 7.40 (d, $^3J=8.4$ Hz, 1H, H8), 7.45 (d, $^3J=8.8$ Hz, 1H, H_m-Ph), 7.57 (ddd, $^3J=8.6$, 7.0 Hz, $^4J=1.6$ Hz, 1H, H7), 7.62 (d, $^3J=8.8$ Hz, 1H, H_o-Ph), 8.13 (dd, $^3J=8.0$ Hz, $^4J=1.4$ Hz, 1H, H5); ^{13}C -NMR ($CDCl_3$, 50 MHz): $\delta = 70.9$ (OC(O)CH₂Ph), 118.0 (C8), 123.4 (C10), 125.3 (C6), 125.9 (C5), 126.1 (C_p-Br), 128.1 (m-C_{CBz}), 128.3 (C_i-Ph), 128.5 (o-C_{CBz}), 128.6 (p-C_{CBz}), 129.7 (o-C_{Ph}), 131.9 (m-C_{Ph}), 133.8 (C3), 134.1 (C7), 134.3 (i-C_{CBz}), 152.0 (OC(O)CH₂Ph), 154.9 (C2), 155.2 (C9), 171.8 (C4); MS (ESI $^+$, MeOH) m/z : 451.2, 453.2 [M+H] $^+$, 473.1, 475.2 [M+Na] $^+$, 489.2, 491.1 [M+K] $^+$. HRMS (ESI $^+$): m/z calcd for $C_{23}H_{16}BrO_5$: 451.0176 [M+H] $^+$; found 451.0177.



3-Benzoyloxycarbonyloxy-2-(5-bromofur-2-yl)-chromen-4-one (13): To a stirred suspension of **10** (600 mg, 1.95 mmol) in CH_2Cl_2 (8 mL) were added a 4.5 M aq. KOH solution (5 mL), 18-crown-6 (5 mol %, 26 mg) and benzyl chloroformate (420 μ L, 2.93 mmol). The reaction mixture became homogenous in 5 min and was stirred at rt for 1 h. After quenching by addition of H₂O (10 mL), the organic layer was extracted with CH_2Cl_2 (3 x), dried over $MgSO_4$, filtered and the volatiles were removed *in vacuo*. The residue was purified by flash chromatography on silica gel eluted with toluene/Et₂O mixture (99:1 → 92:8, v/v) to provide the desired compound **13** as a brown solid (674 mg, 78 %). $C_{21}H_{13}BrO_6$ (441.23). $R_f = 0.54$ (toluene/Et₂O = 9:1). 1H -NMR ($CDCl_3$, 200 MHz): $\delta = 5.35$ (s, 2H, CH₂), 6.53 (d, 1H, $^3J=3.8$ Hz, H β -fur), 7.12 (d, 1H, $^3J=3.8$ Hz, H α -fur), 7.37–7.48 (m, 6H, H₆ & Cbz), 7.57 (dd, $^3J=8.4$ Hz, $^4J=0.8$ Hz, 1H, H8), 7.72 (ddd, $^3J=8.6$, 7.0 Hz, $^4J=1.7$ Hz, 1H, H7), 8.26 (dd, $^3J=8.0$ Hz, $^4J=1.6$ Hz, 1H, H5); ^{13}C -NMR ($CDCl_3$, 50 MHz): $\delta = 71.2$ (OC(O)CH₂Ph), 114.6 (C β), 118.1 (C8), 118.9 (C α), 123.8 (C10), 125.4 (C6), 126.0 (C5), 127.6 (Fur-Br), 128.4 (o-C_{CBz}), 128.6 (m-C_{CBz}), 128.7 (p-C_{CBz}), 131.3 (C3), 134.1 (C7), 134.5 (i-C_{CBz}), 145.1 (C2-Fur), 146.8 (C2), 152.1 (OC(O)CH₂Ph), 154.9 (C9), 171.5 (C4); MS (ESI $^+$, MeOH) m/z : 441.2, 443.2 [M+H] $^+$, 463.2, 465.2 [M+Na] $^+$, 479.1, 481.1 [M+K] $^+$. HRMS (ESI $^+$): m/z calcd for $C_{21}H_{13}NaBrO_6$: 462.9788 [M+H] $^+$; found 462.9768.



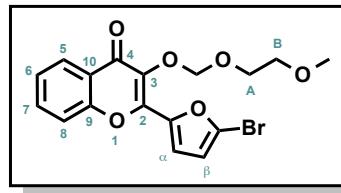
3-Benzoyloxycarbonyloxy-2-(5-bromothien-2-yl)-chromen-4-one (14): To a stirred suspension of **11** (1.05 g, 3.25 mmol) in CH_2Cl_2 (13 mL) were added a 4.5 M aq. KOH solution (9 mL), 18-crown-6 (5 mol %, 43 mg) and benzyl chloroformate (700 μ L, 4.88 mmol). The reaction mixture became homogenous in 5 min and was stirred at rt for 1 h. After quenching by addition of H₂O (10 mL), the organic layer was extracted with CH_2Cl_2 (3 x), dried over $MgSO_4$, filtered and the volatiles were removed *in vacuo*. The residue was purified by flash chromatography on silica gel eluted with cyclohexane/EA mixture (9:1 → 7:3, v/v) to provide the desired compound **14** as white crystals (1.20 g, 81 %). $C_{21}H_{13}BrO_5S$ (457.29). $R_f = 0.53$ (cyclohexane/EA = 7:3); 1H -NMR ($CDCl_3$, 200 MHz): $\delta = 5.37$ (s, 2H, CH₂), 7.16 (d, $^3J=4.0$ Hz, 1H, H β -thioph.), 7.34–7.50 (m, 6H, H₆ & Cbz), 7.52 (dd, $^3J=8.4$ Hz, $^4J=0.6$ Hz, 1H, H8), 7.65 (d, $^3J=4.0$ Hz, 1H, H α -thioph.), 7.72 (td, $^3J=8.4$ Hz, $^4J=1.6$



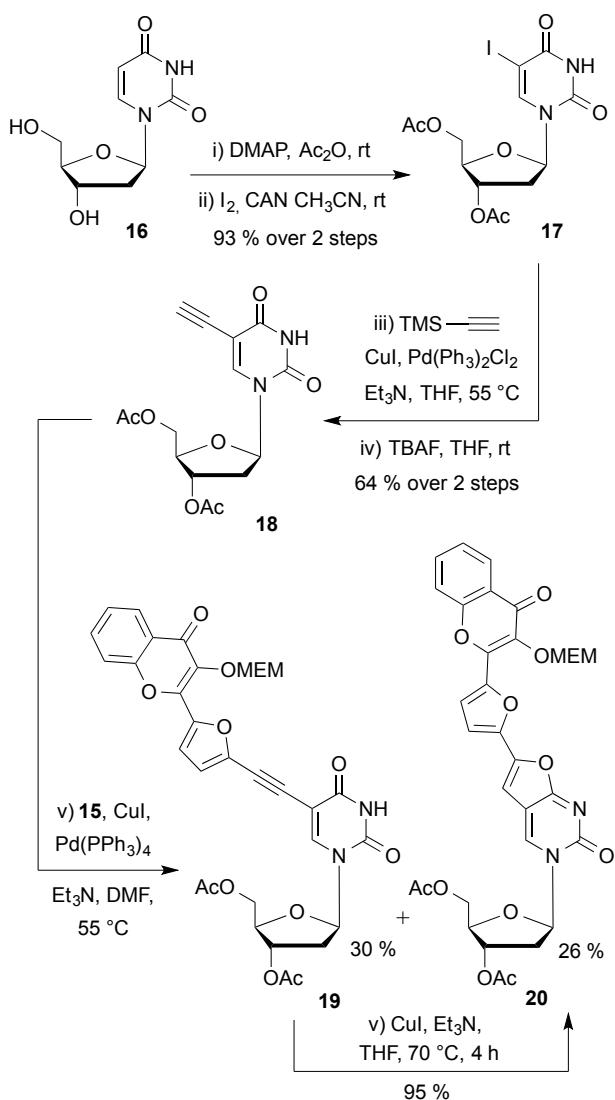
Hz, 1H, H7), 8.25 (dd, $^3J=8.0$ Hz, $^4J=1.6$ Hz, 1H, H5); ^{13}C -NMR (CDCl_3 , 50 MHz): $\delta = 71.3$ ($\text{OC(O)CH}_2\text{Ph}$), 117.8 (C8), 120.4 (*Th*-Br), 123.6 (C10), 125.4 (C6), 126.1 (C5), 128.4 (*o*-C_{Ph}), 128.7 (*m*-C_{Ph}), 128.8 (*p*-C_{Ph}), 131.0 (C β), 131.0 (C α), 131.5 (C2-*Th*), 132.0 (C3), 134.1 (C7), 134.5 (*i*-C_{Ph}), 150.3 (C2), 151.8 ($\text{OC(O)CH}_2\text{Ph}$), 155.0 (C9), 171.3 (C4); MS (ESI $^+$, MeOH/CH₂Cl₂): *m/z*: 479.5, 481.5 [M+Na] $^+$, 495.4, 497.4 [M+K] $^+$. HRMS (ESI $^+$): *m/z* calcd for C₂₁H₁₄BrO₅S: 456.9740 [M+H] $^+$; found 456.9737.

3-((2-methoxyethoxy)methoxy)-2-(5-bromofur-2-yl)-chromen-4-one

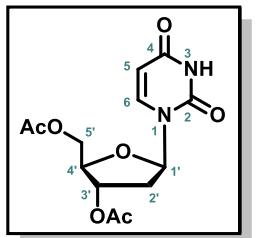
(15): To a stirred solution of **10** (500 mg, 1.63 mmol) in DMF (3 mL) at 0 °C, were sequentially added K₂CO₃ (451 mg, 3.26 mmol) and 2-methoxyethoxymethyl chloride (MEMCl, 372 μL , 3.26 mmol). The reaction mixture was stirred overnight at rt. After quenching by addition of H₂O (10 mL), the organic layer was extracted with CH₂Cl₂ (3 x), dried over MgSO₄, filtered and the volatiles were removed in vacuo. The residue was purified by flash chromatography on silica gel eluted with cyclohexane/EA mixture (19:1 → 1:1, v/v) to provide the desired compound **15** as a brown solid (630 mg, 98 %). C₁₇H₁₅BrO₆ (395.20). $R_f = 0.32$ (cyclohexane/EA = 7:3); ^1H -NMR (CDCl_3 , 200 MHz): $\delta = 3.28$ (s, 3H, OCH₃), 3.42–3.47 (m, 2H, CH₂B), 3.82–3.86 (m, 2H, CH₂A), 5.46 (s, 2H, OCH₂O), 6.52 (d, 1H, $^3J=3.6$ Hz, H β -fur.), 7.35 (d, 1H, $^3J=3.6$ Hz, H α -fur.), 7.30–7.38 (m, 1H, H6), 7.51 (dd, 1H, $^4J=0.9$ Hz, $^3J=8.4$ Hz, H8), 7.63 (ddd, 1H, $^4J=1.8$ Hz, $^3J=7.0$, 8.4 Hz, H7), 8.15 (dd, 1H, $^4J=1.6$ Hz, $^3J=8.0$ Hz, H5); ^{13}C -NMR (CDCl_3 , 50 MHz): $\delta = 59.0$ (OCH₃), 69.7 (CH₂A), 71.6 (CH₂B), 96.6 (OCH₂O), 114.5 (C β), 118.0 (C8), 118.9 (C α), 124.1 (C10), 124.9 (C6), 125.7 (C5), 126.2 (*Fur*-Br), 133.6 (C7), 135.9 (C3), 145.4 (C2-*Fur*), 146.1 (C2), 154.7 (C9), 173.8 (C4); MS (ESI $^+$, MeOH) *m/z*: 416.9, 418.9 [M+Na] $^+$. HRMS (ESI $^+$): *m/z* calcd for C₁₇H₁₆BrO₆: 395.0130 [M+H] $^+$; found 395.0127.



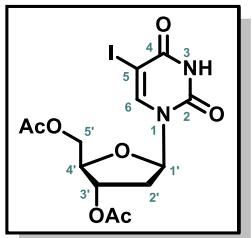
Preliminary approach for the final assembly via a Sonogashira coupling



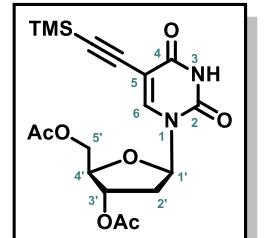
3',5'-Di-O-acetyl-2'-deoxyuridine: DMAP (0.16 mmol, 20 mg) was added to a stirred suspension of 2'-deoxyuridine **16** (1 mmol, 228 mg) in acetic anhydride (7.5 mL). The reaction mixture became homogenous in 6 h, and then the volatiles were removed in *vacuo*. The residue was purified by flash chromatography on silica gel eluted with EA/cyclohexane (4:1 → 19:1, v/v) to provide the desired product as a white resin (296 mg, 95 %). $C_{13}H_{16}N_2O_7$ (312.28). $R_f = 0.39$ (EA/cyclohexane = 9:1); 1H -NMR ($CDCl_3$, 200 MHz): $\delta = 2.10$ (s, 3H, 3'-CH₃), 2.11 (s, 3H, 5'-CH₃), 2.10–2.23 (m, 1H, H_{2'}), 2.53 (ddd, $^3J=14.3$, 5.7, 2.0 Hz, 1H, H_{2'}), 4.23–4.41 (m, 3H, H_{5'}, H_{4'}), 5.20 (dt, $^3J=6.6$, 2.1, 1.8 Hz, 1H, H_{3'}), 5.79 (d, $^3J=8.1$ Hz, 1H, H₅), 6.26 (dd, $^3J=8.4$, 5.6 Hz, 1H, H_{1'}), 7.49 (d, $^3J=8.1$ Hz, 1H, H₆), 9.66 (br s, 1H, NH); ^{13}C -NMR ($CDCl_3$, 50 MHz): $\delta = 20.7$ (CH₃), 20.8 (CH₃), 37.7 (C_{2'}), 63.7 (C_{5'}), 74.0 (C_{3'}), 82.2 (C_{4'}), 85.2 (C_{1'}), 102.9 (C₅), 138.8 (C₆), 150.3 (C₂), 163.4 (C₄), 170.2 (5'-C=O), 170.3 (3'-C=O); MS (ESI⁺, MeOH) m/z : 335.3 [M+Na]⁺.



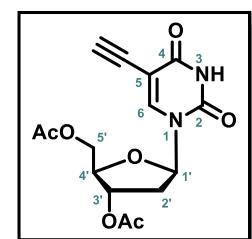
3',5'-Di-O-acetyl-5-iodo-2'-deoxyuridine (17): I₂ (3.46 mmol, 880 mg) and CAN (2.88 mmol, 1.58 g) were added to a stirred solution of the diacetate (5.77 mmol, 1.8 g) in acetonitrile (60 mL). The reaction mixture was refluxed for 1 h then cooled down to rt. The volatiles were removed *in vacuo* and the residue was dissolved in EA (60 mL). The organic solution was washed with water and brine, dried over MgSO₄, filtered and reduced under vacuum. The residue was purified by flash chromatography on silica gel eluted with cyclohexane/EA (1:1 → 0:1, v/v) to provide the desired product **17** as a white solid (2.48 g, 98 %). $C_{13}H_{15}IN_2O_7$ (438.17). $R_f = 0.61$ (EA/cyclohexane = 4:1); 1H -NMR ($CDCl_3$, 200 MHz): $\delta = 2.12$ (s, 3H, 3'-CH₃), 2.13–2.24 (m, 1H, H_{2'A}), 2.21 (s, 3H, 5'-CH₃), 2.54 (ddd, $^3J=14.3$, 5.6, 2.0 Hz, 1H, H_{2'B}), 4.30 (m, 1H, H_{4'}) 4.37 (ddd, $^3J=20.0$, 12.6, 1.7 Hz, 1H, H_{5'}), 5.23 (dt, $^3J=6.4$, 2.0 Hz, 1H, H_{3'}), 6.29 (dd, $^3J=8.2$, 5.6 Hz, 1H, H_{1'}), 7.97 (s, 1H, H₆), 8.81 (br s, 1H, NH); ^{13}C -NMR ($CDCl_3$, 50 MHz): $\delta = 20.8$ (3'-CH₃), 21.0 (5'-CH₃), 38.1 (C_{2'}), 63.7 (C_{5'}), 69.0 (C₅), 74.0 (C_{3'}), 82.5 (C_{4'}), 85.4 (C_{1'}), 143.8 (C₆), 150.0 (C₂), 160.0 (C₄), 170.1 (5'-C=O), 170.3 (3'-C=O); MS (ESI⁺, MeOH) m/z : 461.2 [M+Na]⁺, 899.2 [2M+Na]⁺.



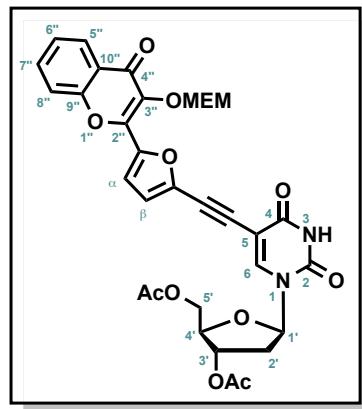
3',5'-Di-O-acetyl-5-trimethylsilylethynyl-2'-deoxyuridine: To a stirred solution of **17** (4.56 mmol, 2.00 g, previously azeotropically coevaporated with dry pyridine) in THF (23 mL) under argon, TMS-acetylene (2 eq., 9.13 mmol, 1.30 mL), Pd(PPh₃)₄ (6 mol %, 0.27 mmol, 316 mg), CuI (5 mol %, 0.23 mmol, 43 mg) and NEt₃ (5 eq, 22.82 mmol, 3.16 mL) were added. The reaction mixture was refluxed for 30 min then cooled down to rt, filtered through a Celite® 545 pad and the volatiles were evaporated *in vacuo*. The residue was purified by flash chromatography on silica gel eluted with cyclohexane/EA (1:0 → 1:1, v/v) to provide the desired compound as a beige solid (1.75 g, 4.29 mmol, 94 %). $C_{18}H_{24}N_2O_7Si$ (408.48). $R_f = 0.35$ (Cyclohexane/EA = 1:1). 1H -NMR ($CDCl_3$, 200 MHz): $\delta = 0.16$ (s, 9H, SiCH₃), 2.08 (s, 3H, 3'-CH₃), 2.08–2.24 (m, 1H, H_{2'A}), 2.14 (s, 3H, 5'-CH₃), 2.50 (ddd, $^3J=14.2$, 5.8, 2.2 Hz, 1H, H_{2'B}), 4.26 (t, $^3J=2.6$ Hz, 1H, H_{4'}), 4.32 (d, $^3J=2.6$ Hz, 2H, H_{5'}), 5.14–5.30 (m, 1H, H_{3'}), 6.29 (dd, $^3J=7.8$, 6.0 Hz, 1H, H_{1'}), 7.80 (s, 1H, H₆), 9.83 (s, 1H, NH); ^{13}C -NMR ($CDCl_3$, 50 MHz): $\delta = -0.3$ (SiMe₃), 20.7 (3'-CH₃), 20.8 (5'-CH₃), 38.1 (C_{2'}), 63.7 (C_{5'}), 73.9 (C_{3'}), 82.4 (C_{4'}), 85.2 (C_{1'}), 95.1 (C≡C-TMS), 99.6 (C≡C-TMS), 100.8 (C₅), 142.1 (C₆), 149.3 (C₂), 161.1 (C₄), 170.0 (5'-C=O), 170.3, (3'-C=O); MS (ESI⁺, MeOH) m/z : 430.7 [M+Na]⁺.



3',5'-Di-O-acetyl-5-ethynyl-2'-deoxyuridine (18): TBAF (4.1 mL, 1.0 M in THF) was added to a stirred solution of the silylated alkyne (4.29 mmol, 1.75 g) in THF (20 mL). After 20 min, the resulting mixture was concentrated *in vacuo* and the residue was purified by flash chromatography on silica gel eluted with cyclohexane/EA (1:1 → 1:9, v/v) to provide the desired compound **18** as a white solid (1.4 g, 4.16 mmol, 97 %). $C_{15}H_{16}N_2O_7$ (336.30). $R_f = 0.15$ (cyclohexane/EA = 1:1); 1H -NMR ($CDCl_3$, 200 MHz): $\delta = 2.12$ (s, 3H, 3'-CH₃), 2.17 (s, 3H, 5'-CH₃), 2.17–2.27 (m, 1H, H_{2'A}), 2.56 (ddd, $^2J=14.3$ Hz, $^3J=5.8$, 2.4 Hz, 1H, H_{2'B}), 3.20 (s, 1H, C≡CH), 4.21–4.48 (m, 3H, H_{4'}, H_{5'}), 5.17–5.28 (m, 1H, H_{3'}), 6.30 (dd, $^3J=7.8$, 5.9 Hz, 1H, H_{1'}), 7.91 (s, 1H, H₆); ^{13}C -NMR ($CDCl_3$, 50 MHz): $\delta = 20.8$ (CH₃), 20.8 (CH₃), 38.3 (C_{2'}), 63.7 (C_{5'}), 73.9 (C_{3'}), 74.5 (C≡CH), 82.2 (C≡CH), 82.6 (C_{4'}), 85.5 (C_{1'}), 99.6 (C₅), 142.7 (C₆), 149.1 (C₂), 161.1 (C₄), 170.1 (5'-C=O), 170.4 (3'-C=O); MS (ESI⁺, MeOH/DCM) m/z : 337.0 [M+H]⁺, 359.0 [M+Na]⁺.

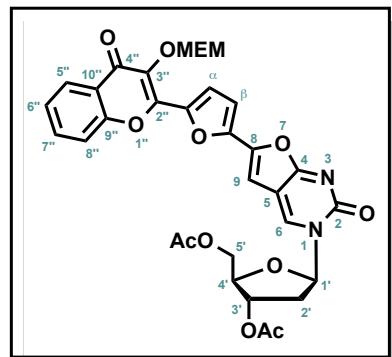


3',5'-Di-O-acetyl-5-(5-(2-methoxyethoxy)methoxy)-4-oxo-chromen-2-yl)fur-2-yl)ethynyl-2'-deoxyuridine (19**):** To a stirred solution of the chromone derivative **15** (1.0 mmol, 359 mg) and the nucleoside derivative **18** (306 mg, 0.91 mmol) in DMF (9 mL) under argon, NEt₃ (5 eq, 589 μL, 4.55 mmol), Pd(PPh₃)₄ (7 mol %), and CuI (7 mol %) were sequentially added. The mixture was warmed for 4 h at 60 °C and then cooled down to rt. The volatiles were removed *in vacuo* and the resulting crude was purified by flash chromatography on silica gel eluted with cyclohexane/EA (1:1 → 1:4, v/v) to give a mixture of two compounds which was further purified by reverse phase chromatography eluted with water/acetone, (2 : 3, v/v) to provide the desired compound **19** as a yellow solid (178 mg, 30 %) and the 5-*endo*-dig cyclized product **20** as an orange solid (154 mg, 26 %). C₃₂H₃₀N₂O₁₃ (650.59). R_f = 0.4 (cyclohexane/EA = 1:4). ¹H-NMR (CDCl₃, 200 MHz): δ = 2.12 (s, 3H, CH₃), 2.21–2.35 (m, 1H, H_{2'A}), 2.28 (s, 3H, CH₃), 2.60 (ddd, ²J=14.3 Hz, ³J=5.8, 2.6 Hz, 1H, H_{2'B}), 3.32 (s, 3H, OCH₃), 3.46–3.51 (m, 2H, CH_{2B}), 3.85–3.89 (m, 2H, CH_{2A}), 4.29–4.36 (m, 1H, H_{4'}), 4.36–4.44 (m, 2H, H_{5'}), 5.24–5.29 (m, 1H, H_{3'}), 5.52 (s, 2H, OCH₂O), 6.34 (dd, ³J=7.3, 6.1 Hz, 1H, H_{1'}), 6.86 (d, ³J=3.7 Hz, 1H, H_{β-fur.}), 7.39 (ddd, ³J=7.9, 6.8 Hz, ⁴J=1.4 Hz, 1H, H_{6''}), 7.46 (d, ³J=3.7 Hz, 1H, H_{α-fur.}), 7.61 (d, ³J=8.4 Hz, 1H, H_{8''}), 7.69 (ddd, ³J=8.4, 6.8 Hz, ⁴J=1.6 Hz, 1H, H_{7''}), 8.06 (s, 1H, H₆), 8.20 (dd, ³J=7.9 Hz, ⁴J=1.1 Hz, 1H, H_{5''}), 9.35 (br. s, 1H, NH); ¹³C-NMR (CDCl₃, 125 MHz): δ = 20.8 (C(O)CH₃), 20.9 (C(O)CH₃), 38.4 (C2'), 58.9 (OCH₃), 63.7 (C5'), 69.7 (CH_{2A}), 71.6 (CH_{2B}), 73.7 (C3'), 82.7 (C4'), 83.2 (Fur-C≡C-), 85.7 (C1'), 87.0 (Fur-C≡C-), 96.6 (OCH₂O), 99.6 (C5), 117.8 (C_α), 118.0 (C8''), 118.4 (C_β), 124.1 (C10''), 124.9 (C6''), 125.7 (C5''), 133.7 (C7''), 136.7 (C3''), 138.4 (Fur-C≡C-), 142.5 (C6), 145.0 (C2''-Fur), 146.8 (C2''), 148.9 (C2), 154.8 (C9''), 160.4 (C4), 170.2 (C=O), 170.3 (C=O), 173.9 (C4''); MS (ESI⁺, MeOH) m/z: 673.1 [M+Na]⁺. HRMS (ESI⁺): m/z calcd for C₃₂H₃₁N₂O₁₃: 651.1826 [M+H]⁺; found 651.1820.

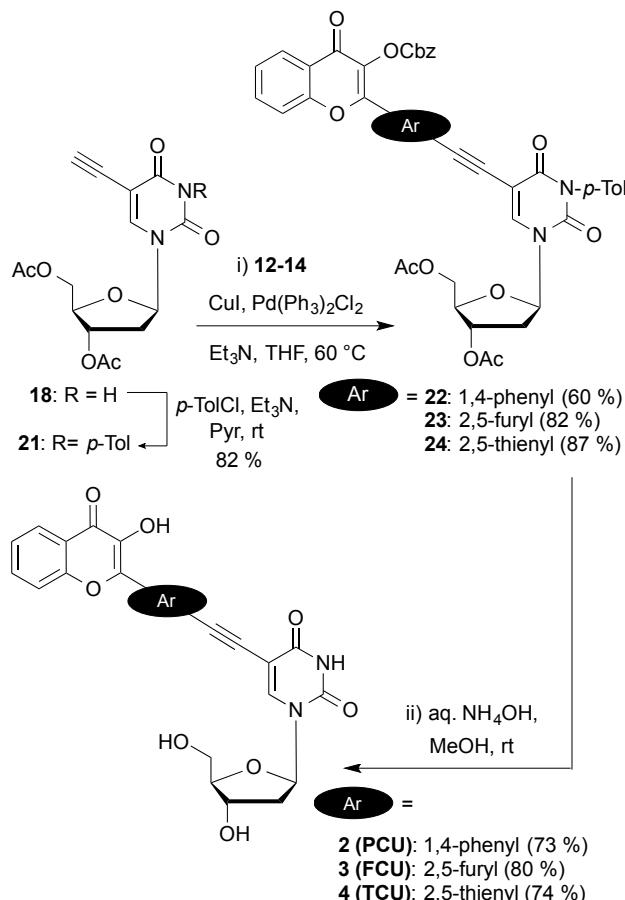


3-(2'-deoxy-3',5'-di-O-acetyl- β -D-ribofuranosyl-1')-6-[5-(3-(2-methoxyethoxy)methoxy)-4-oxo-chromen-2-yl]fur-2-yl]-2,3-dihydrofuro[2,3-*d*]pyrimidin-2-one (**20**):

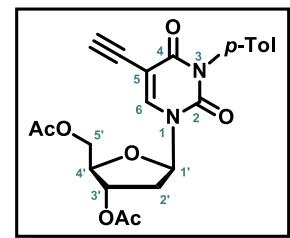
See above procedure since **20** was obtained as a side product. **20** can be derived from **19** via a 5-*endo*-dig-cyclization according to the protocol herein. To a stirred solution of the coupled derivative (400 mg, 0.615 mmol) in a mixture of THF/NEt₃ (15:13 mL), CuI (117 mg, 0.615 mmol) was portionwise added. The reaction mixture was stirred at 70 °C for 4 h and then cooled down to rt. The volatiles were removed *in vacuo*. The residue was purified by flash chromatography on silica gel eluted with cyclohexane/EA (1:1 → 1:4, v/v) to provide the cyclized compound **20** as a yellow solid (380 mg, 95 %). C₃₂H₃₀N₂O₁₃ (650.59). R_f = 0.33 (cyclohexane/EA = 1:19). ¹H-NMR (CDCl₃, 200 MHz): δ = 2.09 (s, 3H, CH₃), 2.09–2.17 (m, 1H, H_{2'}), 2.13 (s, 3H, CH₃), 3.01 (ddd, ²J=14.3 Hz, ³J=5.5, 2.2 Hz, 1H, H_{2'B}), 3.31 (s, 3H, OCH₃), 3.46–3.50 (m, 2H, CH_{2B}), 3.86–3.91 (m, 2H, CH_{2A}), 4.40–4.48 (m, 3H, H_{4'}, H_{5'}), 5.20–5.28 (m, 1H, H_{3'}), 5.54 (s, 2H, OCH₂O), 6.32 (dd, ³J=7.4, 5.6 Hz, 1H, H_{1'}), 6.92 (s, 1H, H₉), 7.04 (d, ³J=3.7 Hz, 1H, H_{β-fur.}), 7.41 (ddd, ³J=8.0, 7.0 Hz, ⁴J=1.1 Hz, 1H, H_{6''}), 7.55 (d, ³J=3.7 Hz, 1H, H_{α-fur.}), 7.58 (d, ³J=8.3 Hz, 1H, H_{8''}), 7.71 (ddd, ³J=8.3, 7.0 Hz, ⁴J=1.5 Hz, 1H, H_{7''}), 8.22 (dd, ³J=8.0 Hz, ⁴J=1.5 Hz, 1H, H_{5''}), 8.41 (s, 1H, H₆); ¹³C-NMR (CDCl₃, 125 MHz): δ = 20.9 (C(O)CH₃), 20.9 (C(O)CH₃), 39.4 (C2'), 59.0 (OCH₃), 63.7 (C5'), 69.7 (CH_{2A}), 71.6 (CH_{2B}), 74.0 (C3'), 83.5 (C4'), 88.8 (C1'), 96.8 (OCH₂O), 99.3 (C9), 107.4 (C5), 112.3 (C_α), 117.9 (C8''), 118.5 (C_β), 124.2 (C10''), 125.0 (C6''), 125.8 (C5''), 133.7 (C7''), 135.6 (Fur-C8), 136.9 (C3''), 145.7 (C2''-Fur), 146.0 (C6), 146.6 (C8), 146.8 (C2''), 154.2 (C2), 154.7 (C9''), 170.3 (C=O), 170.4 (C=O), 171.7 (C4), 173.8 (C4''); MS (ESI⁺, MeOH) m/z: 673.1 [M+Na]⁺. HRMS (ESI⁺): m/z calcd for C₃₂H₃₁N₂O₁₃: 651.1826 [M+H]⁺; found 651.1821.



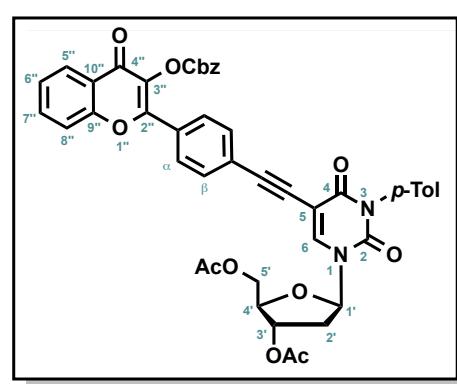
Preparation of the fluorescent emissive nucleosides 2, 3 and 4



3',5'-Di-O-acetyl-3-N-(4-methylbenzoyl)-5-ethynyl-2'-deoxyuridine (21): *p*-Toluoyl chloride (0.36 mmol, 0.05 mL) and $\text{NEt}(\text{iPr})_2$ (0.9 mmol, 0.15 mL) were added to a stirred solution of **18** (0.3 mmol, 100 mg) in pyridine (3 mL). The reaction mixture was stirred for 5 h at rt, and then the volatiles were reduced *in vacuo*. The residue was purified by flash chromatography on silica gel eluted with cyclohexane/EA (9:1 → 1:4, v/v) to provide the desired compound **21** as a beige powder (111 mg, 82 %). $\text{C}_{23}\text{H}_{22}\text{N}_2\text{O}_8$ (454.43). $R_f = 0.42$ (cyclohexane/EA = 1:1); $^1\text{H-NMR}$ (CDCl_3 , 200 MHz): $\delta = 2.08$ (s, 3H, CH_3), 2.18 (s, 3H, CH_3), 2.28 (dd, $^2J=14.4$ Hz, $^3J=7.4$ Hz, 1H, $\text{H}^{\prime\prime}\text{A}$), 2.43 (s, 3H, *p*- CH_3), 2.56 (ddd, $^2J=14.4$ Hz, $^3J=5.8$, 2.6 Hz, 1H, $\text{H}^{\prime\prime}\text{B}$), 3.20 (s, 1H, $\text{C}\equiv\text{CH}$), 4.28–4.32 (m, 1H, $\text{H}^{\prime\prime}\text{A}$), 4.32–4.45 (m, 1H, $\text{H}^{\prime\prime}\text{B}$), 5.24 (dt, $^3J=6.6$, 2.6 Hz, 1H, $\text{H}^{\prime\prime}\text{C}$), 6.29 (dd, $^3J=7.6$, 6.0 Hz, 1H, $\text{H}^{\prime\prime}\text{D}$), 7.28 (d, $^3J=8.2$ Hz, 2H, $\text{H}_{\text{o-Tol}}$), 7.80 (d, $^3J=8.2$ Hz, 2H, $\text{H}_{\text{o-Tol}}$), 7.99 (s, 1H, $\text{H}^{\prime\prime}\text{E}$); $^{13}\text{C-NMR}$ (CDCl_3 , 50 MHz): $\delta = 20.8$ ($\text{C}(\text{O})\text{CH}_3$), 20.8 ($\text{C}(\text{O})\text{CH}_3$), 21.9 ($\text{CH}_3\text{-Tol}$), 38.3 ($\text{C}^{\prime\prime}\text{A}$), 63.6 ($\text{C}^{\prime\prime}\text{B}$), 73.7 ($\text{C}^{\prime\prime}\text{C}$), 74.1 ($\text{C}^{\prime\prime}\text{D}$), 82.6 ($\text{C}^{\prime\prime}\text{E}$), 82.7 ($\text{C}^{\prime\prime}\text{F}$), 85.8 ($\text{C}^{\prime\prime}\text{G}$), 99.6 ($\text{C}^{\prime\prime}\text{H}$), 128.4 ($\text{C}^{\prime\prime}\text{I}$), 129.9 ($\text{C}^{\prime\prime}\text{J}$), 130.7 ($\text{C}^{\prime\prime}\text{K}$), 142.3 ($\text{C}^{\prime\prime}\text{L}$), 146.8 ($\text{C}^{\prime\prime}\text{M}$), 148.1 ($\text{C}^{\prime\prime}\text{N}$), 160.1 ($\text{C}^{\prime\prime}\text{O}$), 167.1 ($\text{C}=\text{O-Tol}$), 170.1 ($\text{C}=\text{O}$), 170.3 ($\text{C}=\text{O}$); MS (ESI^+ , MeOH/DCM) m/z : 454.9 [$\text{M}+\text{H}]^+$, 477.1 [$\text{M}+\text{Na}]^+$. HRMS (ESI^+): m/z calcd for $\text{C}_{23}\text{H}_{22}\text{N}_2\text{O}_8$: 455.1454 [$\text{M}+\text{H}]^+$; found 455.1450.



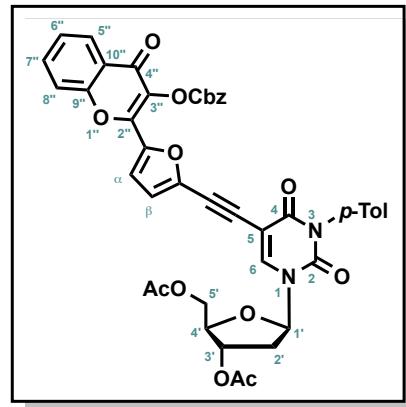
3',5'-Di-O-acetyl-4-(5-(3-(benzyloxycarbonyloxy)-4-oxo-chromen-2-yl)phenyl)ethynyl-3-N-(4-methylbenzoyl)-2'-deoxyuridine (22): To a stirred solution of the chromone derivative **12** (0.363 mmol, 172 mg) and the nucleoside derivative **21** (0.242 mmol, 110 mg) in THF (5.5 mL) under argon, NEt_3 (5 eq, 163 μL , 1.21 mmol), $\text{Pd}(\text{PPh}_3)_2\text{Cl}_2$ (7 mol %) and CuI (7 mol %) were sequentially added. The mixture was warmed for 2 h at 60°C then cooled down to rt. The volatiles were removed *in vacuo* and the resulting crude was purified by flash chromatography on silica gel eluted with toluene/EA (7:3 → 5:5, v/v) to provide the desired compound **22** as a yellow solid (120



mg, 60 %). $C_{46}H_{36}N_2O_{13}$ (824.78). R_f = 0.26 (toluene/EA = 4:1). 2.12 (s, 3H, CH₃), 2.22 (s, 3H, CH₃), 2.28–2.39 (dd, 2J =14.3 Hz, 3J =7.0 Hz, 1H, H_{2'A}), 2.44 (s, 3H, CH₃(Tol)), 2.61 (ddd, 2J =14.3 Hz, 3J =5.8, 2.4 Hz, 1H, H_{2'B}), 4.31–4.38 (m, 1H, H_{4'}), 4.40–4.47 (m, 2H, H_{5'}), 5.28 (s, 2H, CH₂), 5.28–5.33 (m, 1H, H_{3'}), 6.34 (dd, 3J =7.4, 5.8 Hz, 1H, H_{1'}), 7.32 (d, 3J =8.1 Hz, 2H, H_m-Tol), 7.36–7.41 (m, 5H, Cbz), 7.45 (ddd, 3J =8.0, 7.1 Hz, 4J =1.0 Hz, 1H, H_{6''}), 7.56 (d, 3J =8.5 Hz, 2H, H_m-Ph), 7.57 (d, 3J =8.5 Hz, 1H, H_{8''}), 7.74 (ddd, 3J =8.5, 7.1 Hz, 4J =1.6 Hz, 1H, H_{7''}), 7.85 (d, 3J =8.1 Hz, 2H, H_o-Tol), 7.88 (d, 3J =8.5 Hz, 2H, H_o-Ph), 8.04 (s, 1H, H₆), 8.28 (dd, 3J =8.0 Hz, 4J =1.6 Hz, 1H, H_{5''}); ^{13}C NMR (CDCl₃, 50 MHz): δ = 20.8 (C(O)CH₃), 20.9 (C(O)CH₃), 21.9 (CH₃-Tol), 38.4 (C2'), 63.7 (C5'), 71.1 (OC(O)CH₂Ph), 73.7 (C3'), 82.6 (Ph-C≡C-), 82.8 (C4'), 85.9 (C1'), 93.5 (Ph-C≡C-), 100.5 (C5), 118.1 (C8''), 123.6 (C10''), 125.4 (C_p-Ph), 125.4 (C6''), 126.1 (C5''), 128.3 (o-C_{CBz}), 128.4 (C_o-Ph), 128.4 (C_r-Tol), 128.6 (m-C_{CBz}), 128.8 (p-C_{CBz}), 129.6 (C2''-i-Ph), 130.0 (C_m-Tol), 130.7 (C_o-Tol), 131.7 (C_m-Ph), 134.1 (C3''), 134.2 (i-C_{CBz}), 134.4 (C7''), 141.3 (C6), 146.9 (C_p-Tol), 148.1 (C2), 152.2 (OC(O)CH₂Ph), 155.1 (C2''), 155.5.

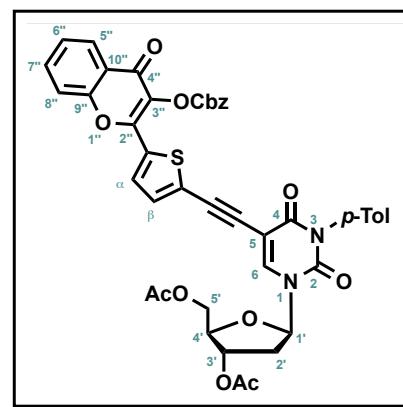
3',5'-Di-O-acetyl-5-(3-(benzyloxycarbonyloxy)-4-oxo-chromen-2-yl)fur-2-yl)ethynyl-3-N-(4-methylbenzoyl)-2'-deoxyuridine (23)

(23): To a stirred solution of the chromone derivative **13** (0.238 mmol, 110 mg) and the nucleoside derivative **21** (0.198 mmol, 90 mg) in THF (4.5 mL) under argon, NEt₃ (5 eq, 162 μ L, 0.99 mmol), Pd(PPh₃)₂Cl₂ (7 mol %), and CuI (7 mol %) were sequentially added. The mixture was warmed for 2 h at 60 °C and then cooled down to rt. The volatiles were removed *in vacuo* and the resulting crude was purified by flash chromatography on silica gel eluted with toluene/EA (95:5 → 7:3, v/v) to provide the desired compound **23** as a yellow solid (132 mg, 82 %). $C_{44}H_{34}N_2O_{14}$ (814.75). R_f = 0.26 (toluene/EA = 4:1). 1H -NMR (CDCl₃, 200 MHz): δ = 2.12 (s, 3H, CH₃), 2.24–2.38 (m, 1H, H_{2'A}), 2.28 (s, 3H, CH₃), 2.45 (s, 3H, CH₃(Tol)), 2.61 (ddd, 2J =14.4 Hz, 3J =5.9, 2.9 Hz, 1H, H_{2'B}), 4.30–4.38 (m, 1H, H_{4'}), 4.38–4.45 (m, 2H, H_{5'}), 5.25–5.31 (m, 1H, H_{3'}), 5.34 (s, 2H, CH₂), 6.33 (dd, 3J =7.2, 6.2 Hz, 1H, H_{1'}), 6.79 (d, 3J =3.8 Hz, 1H, H_B-thioph.), 7.16 (d, 3J =4.0 Hz, 1H, H_A-thioph.), 7.28 (d, 3J =8.2 Hz, 2H, H_m-Tol), 7.34–7.48 (m, 6H, H_{6''}, Cbz), 7.63 (d, 3J =7.8 Hz, 1H, H_{8''}), 7.74 (ddd, 3J =8.4, 6.8 Hz, 4J =1.4 Hz, 1H, H_{7''}), 7.84 (d, 3J =8.2 Hz, 2H, H_o-Tol), 8.13 (s, 1H, H₆), 8.26 (dd, 3J =7.9 Hz, 4J =1.5 Hz, 1H, H_{5''}); ^{13}C -NMR (CDCl₃, 50 MHz): δ = 20.8 (C(O)CH₃), 20.9 (C(O)CH₃), 21.9 (CH₃-Tol), 38.4 (C2'), 63.6 (C5'), 71.2 (OC(O)CH₂Ph), 73.5 (C3'), 82.8 (C4'), 83.2 (Fur-C≡C-), 86.0 (C1'), 87.0 (Fur-C≡C-), 99.4 (C5), 117.9 (C_A), 118.1 (C8''), 118.4 (C_B), 123.7 (C10''), 125.4 (C6''), 125.9 (C5''), 128.3 (o-C_{CBz}), 128.3 (C_r-Tol), 128.6 (m-C_{CBz}), 128.7 (p-C_{CBz}), 130.0 (C_m-Tol), 130.7 (C_o-Tol), 132.0 (C3''), 134.3 (C7''), 134.5 (i-C_{CBz}), 139.3 (C2''-Fur), 142.3 (C6), 143.7 (Fur-C≡C-), 146.8 (C2''), 147.0 (C_p-Tol), 148.0 (C2), 151.9 (OC(O)CH₂Ph), 155.0 (C9''), 159.4 (C4), 167.0 (C=O-Tol), 170.1 (C=O), 170.3 (C=O), 171.4 (C4''); MS (ESI⁺, MeOH) m/z : 815.5 [M+H]⁺, 837.4 [M+Na]⁺, 853.4 [M+K]⁺. HRMS (ESI⁺): m/z calcd for $C_{44}H_{34}N_2NaO_{14}$: 837.1902 [M+Na]⁺; found 837.1895.



3',5'-Di-O-acetyl-5-(3-(benzyloxycarbonyloxy)-4-oxo-chromen-2-yl)thien-2-yl)ethynyl-3-N-(4-methylbenzoyl)-2'-deoxyuridine (24)

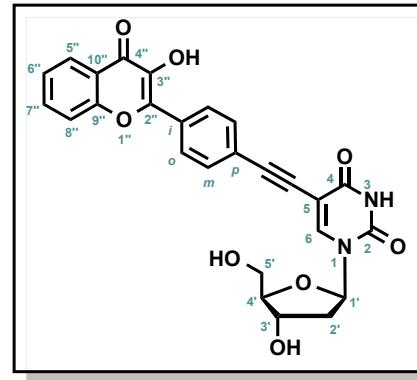
(24): To a stirred solution of the chromone derivative **14** (0.26 mmol, 120 mg) and the acetylenic nucleoside derivative **21** (0.24 mmol, 108 mg) in THF (2 mL) under argon, NEt₃ (5 eq, 162 μ L, 1.2 mmol), Pd(PPh₃)₂Cl₂ (7 mol %), and CuI (7 mol %) were sequentially added. The mixture was warmed for 2 h at 60 °C and then cooled down to rt. The volatiles were reduced *in vacuo* and the resulting crude was purified by flash chromatography on silica gel eluted with cyclohexane/EA (7:3 → 1:1, v/v) to provide the desired compound **24** as yellow solid (173 mg, 87 %). $C_{44}H_{34}N_2O_{13}S$ (830.81). R_f = 0.34 (cyclohexane/EA = 1:1); 1H -NMR (CDCl₃, 200 MHz): δ = 2.11 (s, 3H, CH₃), 2.19 (s, 3H, CH₃), 2.33 (ddd, 2J =14.4 Hz, 3J =7.4, 7.0 Hz, 1H, H_{2'A}), 2.44 (s, 3H, CH₃(Tol)), 2.61 (ddd, 2J =14.4 Hz, 3J =6.0, 2.8 Hz, 1H, H_{2'B}), 4.30–4.37 (m, 1H, H_{4'}), 4.37–4.45 (m, 2H, H_{5'}), 5.23–5.32 (m, 1H, H_{3'}), 5.37 (s, 2H, CH₂), 6.33 (dd, 3J =7.4, 6.0 Hz, 1H, H_{1'}), 7.27 (d, 3J =4.0 Hz, 1H, H_B-thioph.), 7.28 (d, 3J =8.2 Hz, 2H, H_m-Tol), 7.32–7.51 (m, 6H, H_{6''}, Cbz), 7.52 (d, 3J =8.5 Hz, 1H, H_{8''}), 7.72 (ddd, 3J =8.5, 7.0 Hz, 4J =1.6 Hz, 1H, H_{7''}), 7.77 (d, 3J =4.0 Hz, 1H, H_A-thioph.), 7.84 (d, 3J =8.2 Hz, 2H, H_o-Tol), 8.05 (s, 1H, H₆), 8.24 (dd, 3J =7.9 Hz, 4J =1.5 Hz, 1H, H_{5''}); ^{13}C -NMR (CDCl₃, 50 MHz): δ = 20.8 (C(O)CH₃), 20.9 (C(O)CH₃), 21.9 (CH₃-Tol), 38.4 (C2'), 63.6 (C5'), 71.4 (OC(O)CH₂Ph), 73.6 (C3'), 82.8 (C4'), 86.0 (C1'), 86.5 (Th-C≡C-), 87.4 (Th-C≡C-), 100.1 (C5), 117.9 (C8''), 123.6 (C10''), 125.4 (C6''), 126.0 (C5''), 128.4 (o-C_{CBz}),



128.4 (C_i -Tol), 128.7 (m -C_{CBz}), 128.8 (p -C_{CBz}), 130.0 (C_m -Tol), 130.7 (C_o -Tol), 130.7 ($C\alpha$), 132.0 ($C3''$), 132.1 (Th -C≡C-), 133.2 ($C\beta$), 134.2 ($C7''$), 134.4 (i -C_{CBz}), 141.7 ($C6$), 147.0 (C_p -Tol), 148.0 ($C2$), 150.4 ($C2''$), 151.8 ($O\bar{C}(O)CH_2Ph$), 155.0 ($C9''$), 159.6 ($C4$), 167.1 ($C=O$ -Tol), 170.0 ($C=O$), 170.3 ($C=O$), 171.3 ($C4''$); HRMS (ESI $^+$): m/z calcd for C₄₄H₃₅N₂O₁₃S: 831.1854 [M+H] $^+$; found 831.1854.

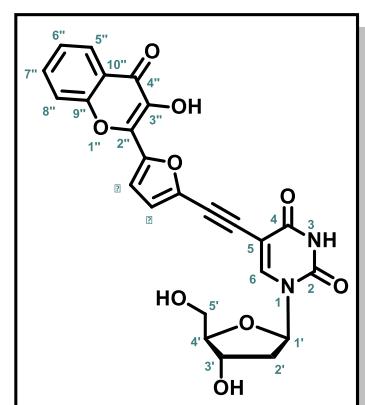
5-(5-(3-Hydroxy-4-oxo-chromen-2-yl)phenyl)ethynyl-2'-deoxyuridine (2: PCU):

To a stirred solution of **22** (0.024 mmol, 20 mg) in MeOH (1.4 mL) was dropwise added a 33 % ammonium hydroxide aq. solution (900 μ L). Protected from light, the reaction mixture was stirred at rt for 6 h and then acidified with AcOH (900 μ L). After addition of 6.5 mL of milliQ water, the resulting solution was stored 12 h at 0 °C. The heterogeneous mixture was centrifuged for settling solid particles and the supernatant was carefully taken out with a syringe with a fine needle. The precipitate was triturated with a solution of MeOH/H₂O (7:3) and after centrifugation, the supernatant was removed by suction. The final product was partially solubilized in water (300 μ L). The resulting solution was frozen and lyophilized to provide the desired compound **2** (PCU) as an orange solid (6 mg, 50 %). C₂₆H₂₀N₂O₈ (488.45). R_f = 0.65 (DCM/MeOH = 9:1). ¹H-NMR (DMSO- d^6 , 500 MHz): δ = 2.14–2.18 (m, 1H, H_{2'A}), 2.21 (dd, 2J =13.5 Hz, 3J =6.5 Hz, 1H, H_{2'B}), 3.60 (ddd, 2J =12.0 Hz, 3J =5.0, 3.5 Hz, 1H, H_{5'A}), 3.68 (ddd, 2J =12.0 Hz, 3J =5.0, 3.5 Hz, 1H, H_{5'B}), 3.81–3.83 (q, 3J =3.5 Hz, 1H, H_{4'}), 4.26–4.29 (m, 1H, H_{3'}), 5.19 (t, 3J =5.0 Hz, 1H, 5'-OH), 5.27 (d, 3J =4.0 Hz, 1H, 3'-OH), 6.14 (t, 3J =6.5 Hz, 1H, H_{1'}), 7.48 (td, 3J =8.0 Hz, 4J =1.5 Hz, 1H, H_{6''}), 7.66 (d, 3J =8.5 Hz, 2H, H_{m-Ph}), 7.79 (d, 3J =8.0 Hz, 1H, H_{8''}), 7.83 (td, 3J =8.0 Hz, 4J =1.5 Hz, 1H, H_{7''}), 8.13 (dd, 3J =8.0 Hz, 4J =1.5 Hz, 1H, H_{5''}), 8.29 (d, 3J =8.5 Hz, 2H, H_{o-Ph}), 8.45 (s, 1H, H₆), 9.87 (s, 1H, 3''-OH), 11.73 (s, 1H, NH); ¹³C-NMR (DMSO- d^6 , 125 MHz): δ = 40.1 (C_{2'}), 60.7 (C_{5'}), 69.8 (C_{3'}), 84.6 (Ph-C≡C-), 84.8 (C_{4'}), 87.5 (C_{1'}), 91.6 (Ph-C≡C-), 97.8 (C₅), 118.4 (C_{8''}), 121.2 (C_{10''}), 123.5 (C_{6''}), 124.5 (C_{i-Ph}), 124.7 (C_{5''}), 127.6 (C_{o-Ph}), 131.1 (C_{p-Ph}), 131.1 (C_{m-Ph}), 133.8 (C_{7''}), 139.5 (C_{3''}), 144.0 (C₂), 144.2 (C₆), 149.3 (C_{2''}), 154.5 (C_{9''}), 161.3 (C₄), 172.9 (C_{4''}); MS (ESI $^+$, MeOH) m/z : 489.4 [M+H] $^+$, 511.3 [M+Na] $^+$. HRMS (ESI $^+$): m/z calcd for C₂₆H₂₁N₂O₈: 489.1292 [M+H] $^+$; found 489.1294.



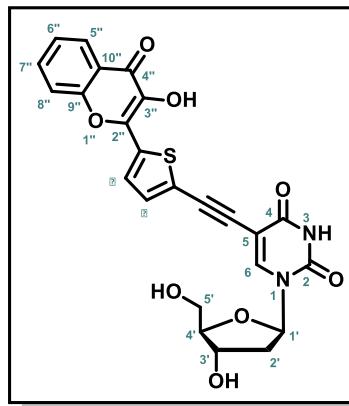
5-(5-(3-Hydroxy-4-oxo-chromen-2-yl)fur-2-yl)ethynyl-2'-deoxyuridine (3: FCU):

To a stirred solution of **23** (0.049 mmol, 40 mg) in MeOH (3 mL) was dropwise added a 33 % ammonium hydroxide aq. solution (1.8 mL). Protected from light, the reaction mixture was stirred at rt for 4 h and then acidified with AcOH (1.7 mL). After addition of 13 mL of milliQ water, the resulting solution was stored 12 h at 0 °C. The heterogeneous mixture was centrifuged for settling solid particles and the supernatant was carefully taken out with a syringe with a fine needle. The precipitate was triturated with a solution of MeOH/H₂O (7:3), and after centrifugation the supernatant was removed by suction. The final product was partially solubilized in water (300 μ L). The resulting solution was frozen and lyophilized to provide the desired compound **3** (FCU) as an orange solid (14 mg, 58 %). C₂₄H₁₈N₂O₉ (478.41). R_f = 0.42 (DCM/MeOH = 9:1). ¹H-NMR (DMSO- d^6 , 500 MHz): δ = 2.14–2.20 (m, 1H, H_{2'A}), 2.22 (dd, 2J =13.7 Hz, 3J =6.3 Hz, 1H, H_{2'B}), 3.59 (ddd, 2J =11.8 Hz, 3J =4.5, 3.5 Hz, 1H, H_{5'A}), 3.67 (ddd, 2J =11.8 Hz, 3J =4.5, 3.5 Hz, 1H, H_{5'B}), 3.82 (q, 3J =3.5 Hz, 1H, H_{4'}), 4.24–4.28 (m, 1H, H_{3'}), 5.18 (t, 3J =5.0 Hz, 1H, 5'-OH), 5.27 (d, 3J =4.5 Hz, 1H, 3'-OH), 6.13 (t, 3J =6.3 Hz, 1H, H_{1'}), 7.11 (d, 3J =3.5 Hz, 1H, H β -thioph.), 7.33 (d, 3J =3.5 Hz, 1H, H α -thioph.), 7.48 (ddd, 3J =8.0, 5.5 Hz, 4J =2.5 Hz, 1H, H_{6''}), 7.78–7.82 (m, 2H, H_{8''}, H_{7''}), 8.12 (dd, 3J =8.0 Hz, 4J =1.0 Hz, 1H, H_{5''}), 8.50 (s, 1H, H₆), 10.28 (s, 1H, 3''-OH), 11.80 (s, 1H, NH); ¹³C-NMR (DMSO- d^6 , 125 MHz): δ = 40.2 (C_{2'}), 60.7 (C_{5'}), 69.8 (C_{3'}), 81.5 (Fur-C≡C-), 85.0 (C_{1'}), 87.6 (C_{4'}), 89.1 (Fur-C≡C-), 96.8 (C₅), 116.2 (C α), 118.3 (C β), 118.4 (C_{8''}), 121.8 (C_{10''}), 124.7 (C_{6''}), 124.8 (C_{5''}), 133.7 (C_{7''}), 137.2 (C_{3''}), 138.0 (Fur-C≡C-), 138.2 (C_{2''}-Fur), 145.0 (C₆), 145.0 (C_{2''}), 149.2 (C₂), 154.1 (C_{9''}), 161.1 (C₄), 171.8 (C_{4''}); MS (ESI $^+$, MeOH) m/z : 479.1 [M+H] $^+$, 501.3 [M+Na] $^+$. HRMS (ESI $^+$): m/z calcd for C₂₄H₁₉N₂O₉: 479.1085 [M+H] $^+$; found 479.1084.



5-(5-(3-Hydroxy-4-oxo-chromen-2-yl)thien-2-yl)ethynyl-2'-deoxyuridine (4: TCU):

To a stirred suspension of **24** (0.042 mmol, 35 mg) in MeOH (3 mL) was dropwise added a 33 % ammonium hydroxide aq. solution (1.8 mL). Protected from light, the reaction mixture was stirred at rt for 3 h and then acidified with AcOH (1.8 mL). After addition of 12 mL of milliQ water, the resulting solution was stored 12 h at 0 °C. The heterogeneous mixture was centrifuged for settling solid particles and the supernatant was carefully taken out with a syringe with a fine needle. The precipitate was triturated with a solution of MeOH/H₂O (7:3), and after centrifugation the supernatant was removed by suction. This operation was repeated 3 times. The final product was partially solubilized in water (300 µL). The resulting solution was frozen and lyophilized to provide the desired compound **4** (TCU) as an orange solid (11 mg, 54 %). C₂₄H₁₈N₂O₈S (494.47). R_f = 0.50 (DCM/MeOH = 9:1). ¹H-NMR (DMSO-*d*⁶, 500 MHz): δ = 2.15 (ddd, ²J=13.5 Hz, ³J=6.5, 4.0 Hz, 1H, H2'_A), 2.21 (dd, ²J=13.5 Hz, ³J=6.5 Hz, 1H, H2'_B), 3.59 (ddd, ²J=12.0 Hz, ³J=4.5, 3.5 Hz, 1H, H5'_A), 3.66 (ddd, ²J=12.0 Hz, ³J=4.5, 3.5 Hz, 1H, H5'_B), 3.82 (q, ³J=3.5 Hz, 1H, H4'), 4.24–4.28 (m, 1H, H3'), 5.19 (t, ³J=5.0 Hz, 1H, 5'-OH), 5.27 (d, ³J=4.5 Hz, 1H, 3'-OH), 6.13 (t, ³J=6.5 Hz, 1H, H1'), 7.47 (ddd, ³J=8.0, 7.0 Hz, ⁴J=1.0 Hz, 1H, H6''), 7.48 (d, ³J=4.3 Hz, 1H, Hβ-thioph.), 7.73 (d, ³J=8.5 Hz, 1H, H8''), 7.81 (ddd, ³J=8.5, 7.0 Hz, ⁴J=1.5 Hz, 1H, H7''), 7.89 (d, ³J=4.3 Hz, 1H, Hα-thioph.), 8.10 (dd, ³J=8.0 Hz, ⁴J=1.5 Hz, 1H, H5''), 8.44 (s, 1H, H6), 10.67 (s, 1H, 3''-OH), 11.77 (s, 1H, NH); ¹³C-NMR (DMSO-*d*⁶, 125 MHz): δ = 40.1 (C2'), 60.7 (C5'), 69.8 (C3'), 84.7 (Th-C≡C-), 84.9 (C1'), 87.6 (C4'), 89.1 (Th-C≡C-), 97.6 (C5), 118.1 (C8''), 121.8 (C10''), 124.6 (C6''), 124.8 (C5''), 125.8 (C2''-Th), 128.0 (Cα), 132.4 (Cβ), 133.2 (Th-C≡C-), 133.7 (C7''), 137.3 (C3''), 142.1 (C2''), 144.3 (C6), 149.3 (C2), 154.1 (C9''), 161.1 (C4), 171.9 (C4''); MS (ESI⁺, MeOH) m/z: 495.3 [M+H]⁺, 517.2 [M+Na]⁺. HRMS (ESI[−]): m/z calcd for C₂₄H₁₉N₂O₈S: 493.0700 [M-H]⁻; found 493.0689.



Control Method

Control Method

Name Niko-060214-Free-Nucleotide-analytique
User Name Administrator
Date Modified 10/02/2014 09:17:43
Description
Method Time 61,0 [min]

Pump #1

Initial Condition

Pump Mode LPG1
Flow 0,500 [mL/min]
Max. Pressure 25,0 [MPa]
Min. Pressure 0,0 [MPa]
Solvents
A 90,0 [%] H₂O + 0,1% AF
B 10,0 [%] ACN + 0,1% AF
C 0,0 [%]
D 0,0 [%]

Time Program

#	Time [min]	Function	Value #1	Value #2	Value #3	Value #4
1	5,00	Composition [%]	90,0	10,0	0,0	0,0
2	40,00	Composition [%]	0,0	100,0	0,0	0,0
3	45,00	Composition [%]	0,0	100,0	0,0	0,0
4	55,00	Composition [%]	90,0	10,0	0,0	0,0
5	60,00	Composition [%]	90,0	10,0	0,0	0,0

Pump #1 Valve / Event

Initial Condition

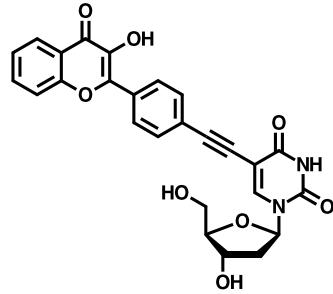
Valve #1 1
Event #1 Off
Event #2 Off
Event #3 Off

4WL-UV Detector

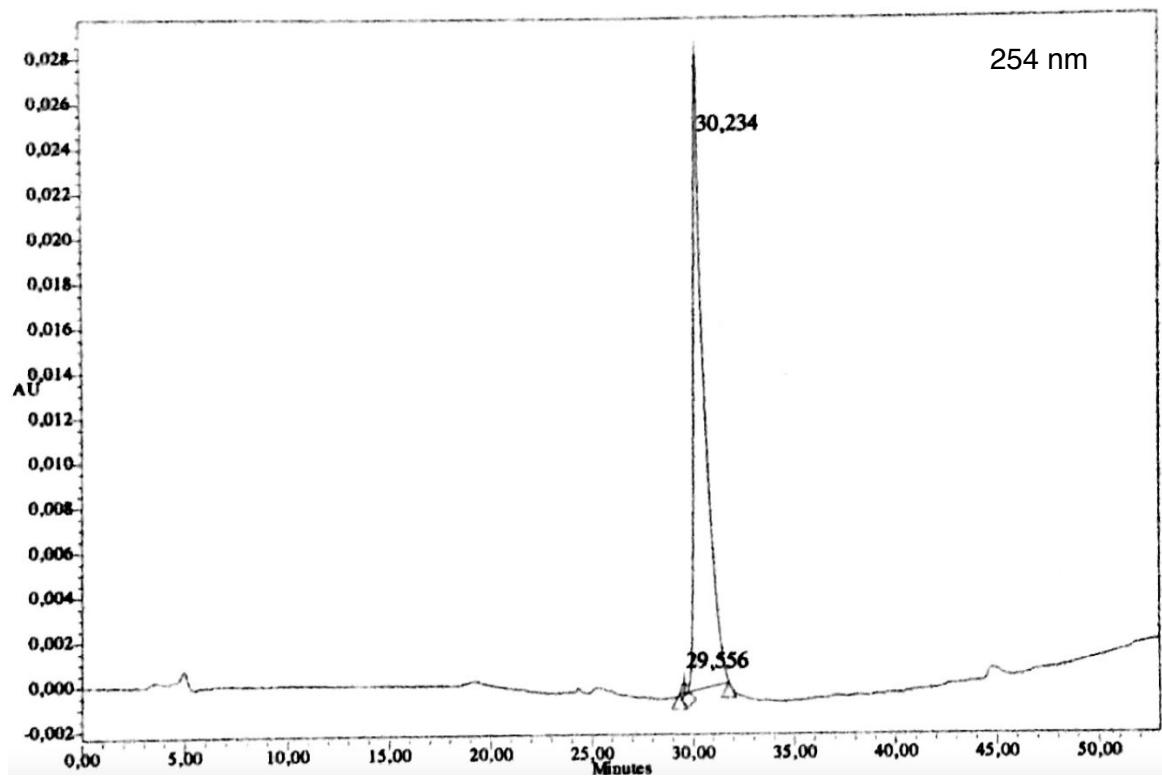
Initial Condition

Wavelength #1 254 [nm]
Wavelength #2 280 [nm]
Wavelength #3 360 [nm]
Wavelength #4 393 [nm]
Wavelength #5 (Ratio) 0 [nm]
Response STD
Autozero On

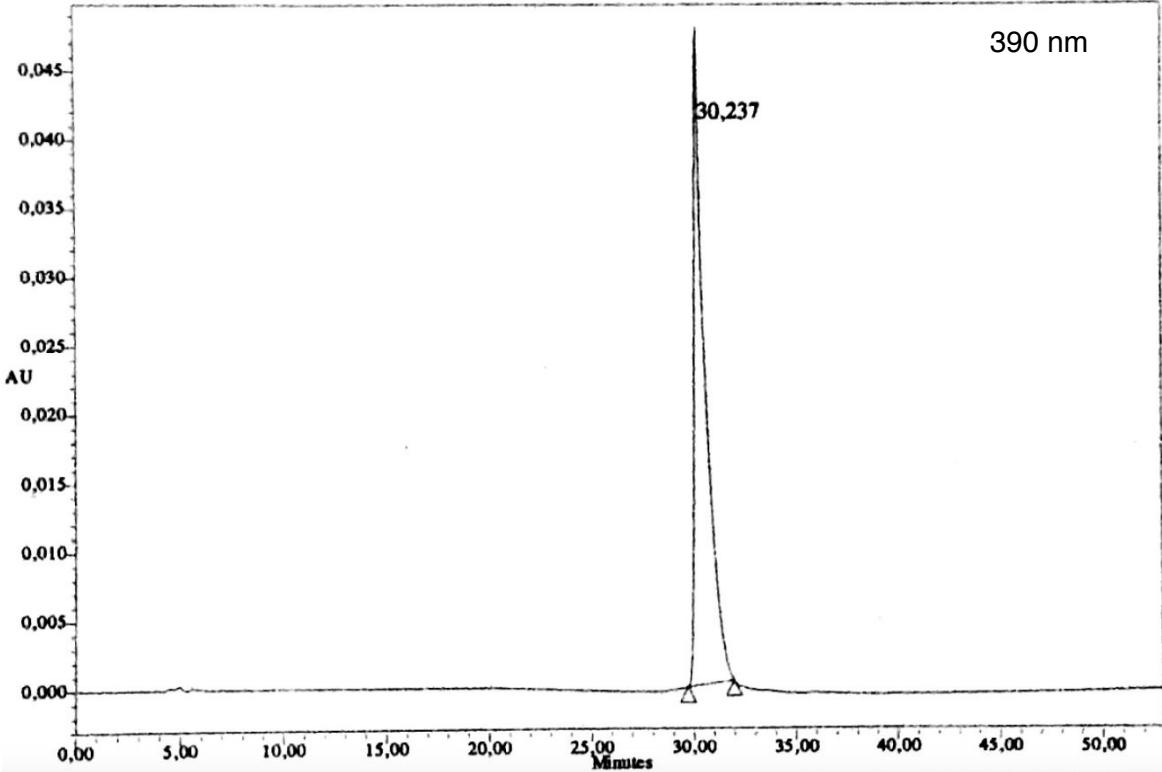
PCU



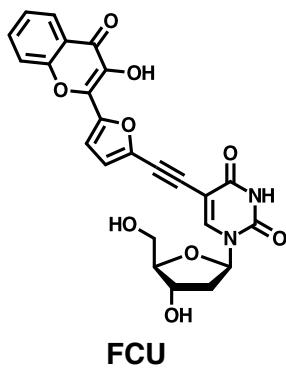
PCU



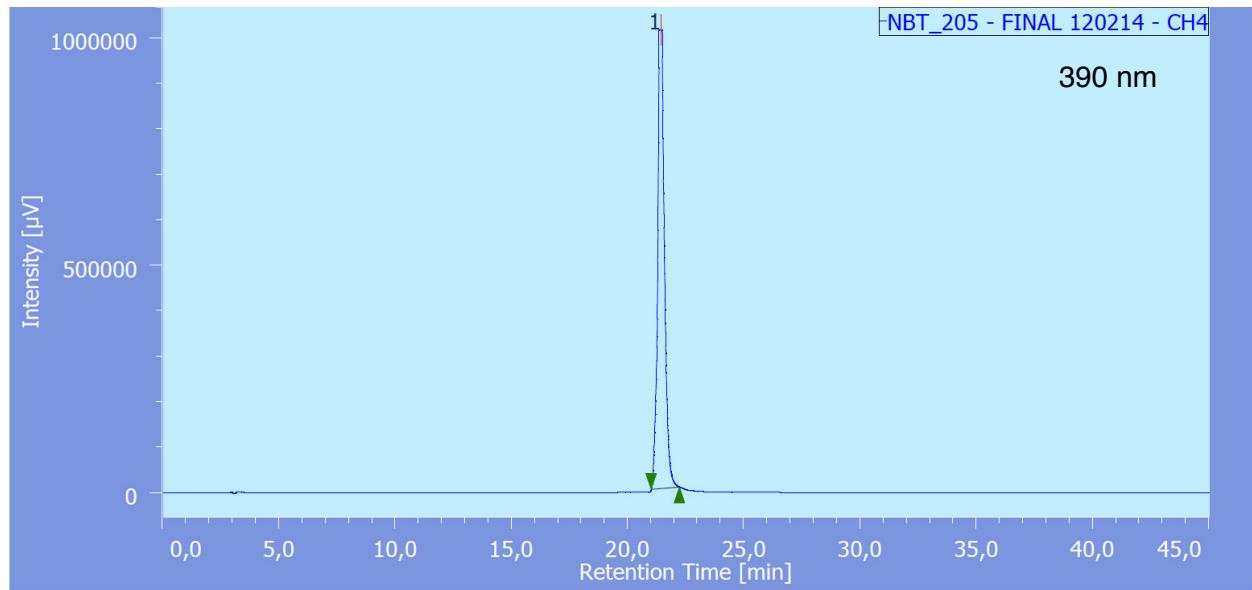
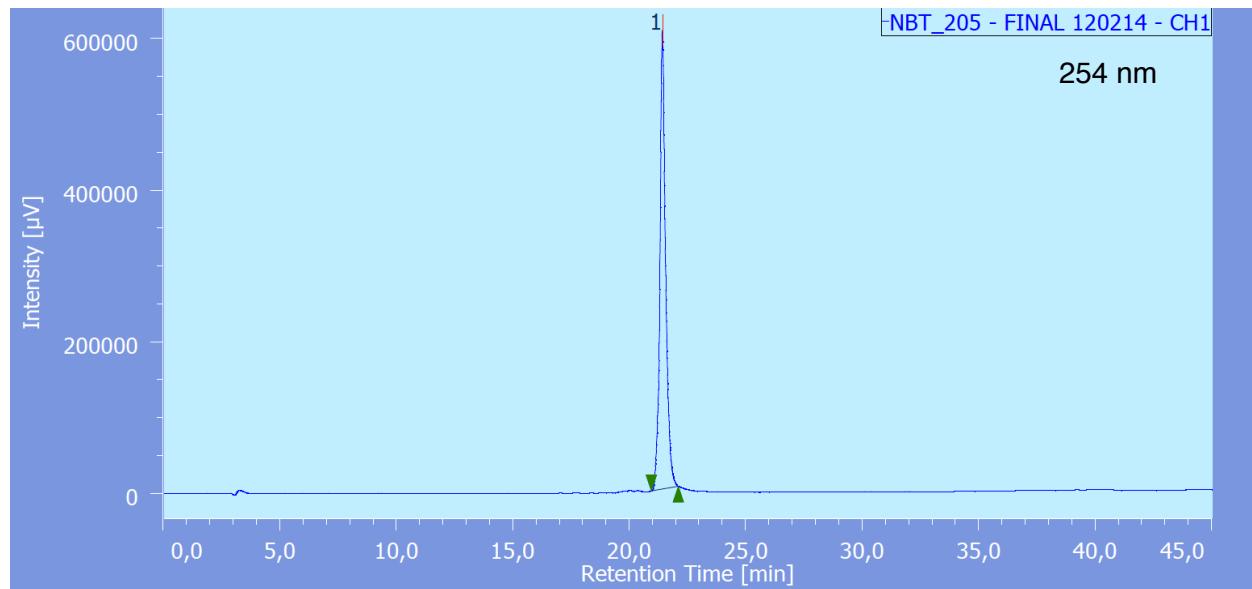
390 nm



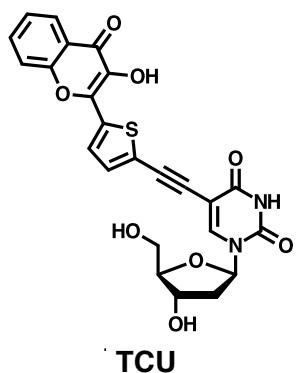
FCU



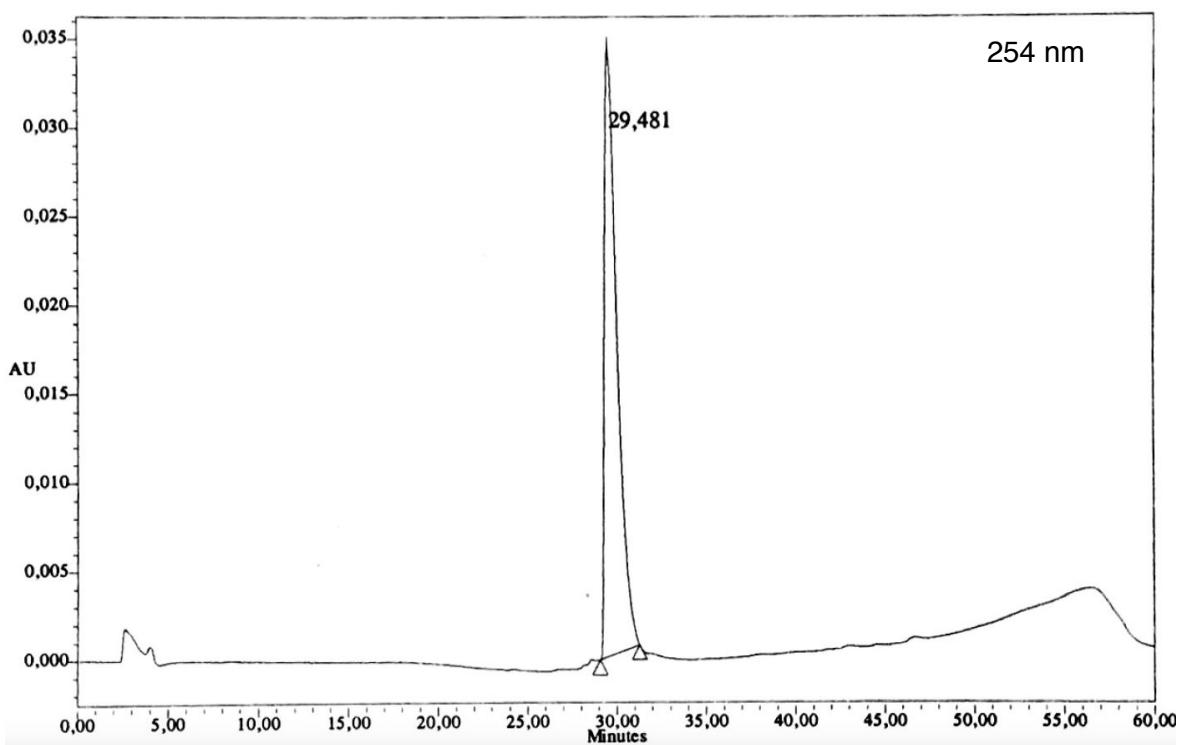
FCU



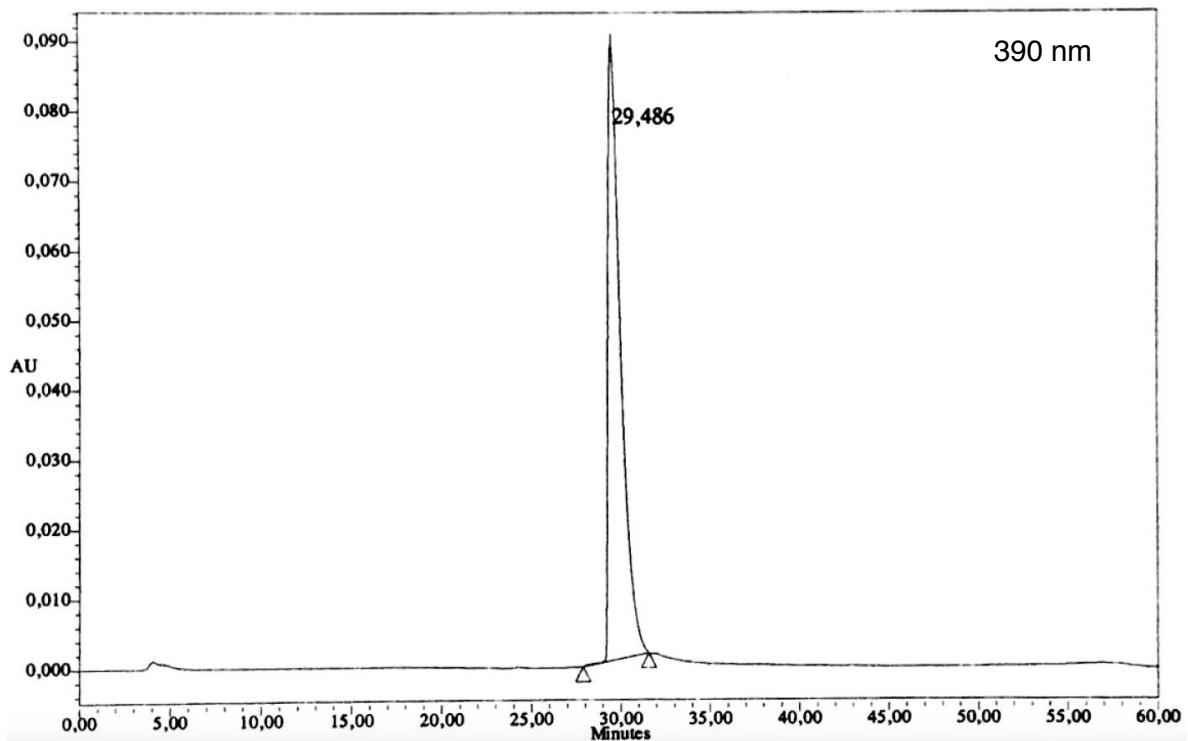
TCU



TCU



254 nm



390 nm

Supporting Information To:

Design and development of a two-color emissive FRET pair based on a photostable fluorescent deoxyuridine donor presenting mega-Stokes shifts

Nicolas P. F. Barthes,^a Krishna Gavvala,^b Dmytro Dziuba,^a Dominique Bonhomme,^a Anne Sophie Dabert-Gay,^c Delphine Debayle,^c Benoît Y. Michel,*^a Yves Mély,^b Alain Burger*^a

a) Institut de Chimie de Nice, UMR 7272, Université de Nice Sophia Antipolis, CNRS, Parc Valrose, 06108 Nice Cedex 2, France

b) Laboratoire de Biophotonique et Pharmacologie, UMR 7213, Faculté de Pharmacie, Université de Strasbourg, CNRS, 74 Route du Rhin, 67401 Illkirch, France

c) Institut de Pharmacologie Moléculaire et Cellulaire, UMR 6097, Université de Nice Sophia Antipolis, 660 Route des Lucioles, 06560 Valbonne, France

Corresponding authors: burger@unice.fr ; yves.mely@unistra.fr ; benoit.michel@unice.fr

Experimental section

General methods:

All reactions involving water-sensitive reagents were performed in oven-dried glassware under argon using dry solvents. The synthetic intermediates were co-evaporated twice with toluene beforehand and dried in *vacuo* before use. All chemical reagents were obtained from commercial sources (Aldrich, Acros, Ala Aesar) and were used as supplied. Anhydrous solvents were obtained according to standard procedures.¹ The reactions were monitored by thin-layer chromatography (TLC, Merck silica gel 60 F254 plates) and visualized both by UV radiation (254 & 365 nm) and by spraying with vanillin in ethanol containing H₂SO₄ followed by a subsequent warming with a heat gun. Column chromatography² was performed with flash silica gel (40–63 mm). All NMR spectra (¹H, ¹³C, ²D) were recorded on 200 or 500 Bruker Advance Spectrometers (200 or 500 MHz). ¹H NMR (200 and 500 MHz), ¹³C NMR (50 and 125 MHz, recorded with complete proton decoupling), ³¹P NMR (80 MHz, proton decoupling) spectra were obtained with samples dissolved in CDCl₃, CD₂Cl₂, CD₃OD, DMSO-*d*⁶, acetone-*d*₆, or CD₃CN with the residual solvent signals used as internal references: 7.26 ppm for CHCl₃, 5.32 ppm for CDHCl₂, 3.31 ppm for CD₂HOD, 2.50 ppm for (CD₃)(CD₂H)SO), 2.05 ppm for (CD₃)(CD₂H)C(O), 1.94 ppm for CD₂HCN for ¹H NMR experiments, and 77.0 ppm for CDCl₃, 53.8 ppm for CD₂Cl₂, 49.0 ppm for CD₃OD, 39.4 ppm for (CD₃)₂S(O), 30.8 ppm for (CD₃)₂C(O) for ¹³C NMR experiments.³ Chemical shifts (δ) are given in ppm to the nearest 0.01 (¹H) or 0.1 ppm (¹³C). The coupling constants (J) are given in Hertz (Hz). The signals are reported as follows: (s=singlet, d=doublet, t=triplet, m=multiplet, br=broad). Assignments of ¹H and ¹³C NMR signals were achieved with the help of D/H exchange, COSY, DEPT, APT, HMQC, HSQC, HMBC experiments. Regular mass spectra (MS) were recorded on an Esquire 3000 Plus apparatus with ESI in both positive and negative mode. High-resolution mass spectrometry was conducted with a FINIGAN MAT 95 spectrometer with EI or ESI ionization techniques. Supplementary data associated with this article, including the experimental protocols for the synthesis of intermediates **2** and **5–11**, the ¹H, ¹³C and (in part) ¹H–¹H COSY, ¹H–¹³C HMQC, ¹H–¹³C HSQC and ¹H–¹³C HMBC NMR spectra of all compounds, can be consulted in this *Supporting Information*. Systematic flavone and nucleobase nomenclatures are used below for the assignments of each spectrum. All solvents for absorption and fluorescence experiments were of spectroscopic grade. Absorption spectra were recorded on a Cary 4 spectrophotometer (Varian) using 1cm quartz cells. Stock solution of the MTCU model **1** was prepared using dioxane. The samples used for spectroscopic measurements contained \approx 0.1% v/v of the stock solvent. Fluorescence spectra were recorded on FluoroMax 4.0 spectrofluorometer (Jobin Yvon, Horiba). Excitation wavelength was used as mentioned in the corresponding experiments. Photostability studies were conducted in a 100 μ l fluorescence cell, excitation and emission slits were set to 4 nm. The concentrations of the samples were 2 μ M. The τ_s values were determined according to the equation $\ln(I_0/I) = k_d \tau_s$ (where k_d is the degradation constant, I_0 the initial irradiation intensity and I the intensity at time t).

¹ W. L. F. Armarego and C. L. L. Chai, *Purification of Laboratory Chemicals*, 7th ed.; Butterworth-Heinemann: Oxford, 2013.

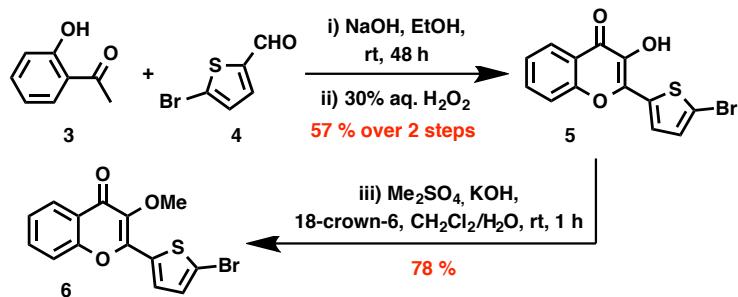
² W. C. Still, M. Kahn and A. Mitra, *J. Org. Chem.* **1978**, *43*, 2923–2925.

³ a) H. E. Gottlieb, V. Kotlyar and A. Nudelman, *J. Org. Chem.*, 1997, **62**, 7512–7515; (b) G. R. Fulmer, A. Miller, N. H. Sherden and H. E. Gottlieb, *Organometallics*, 2010, **29**, 2176–2179.

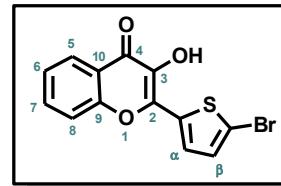
Supporting Information

1. Synthetic Procedures

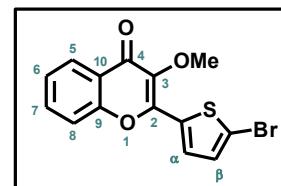
1.1 Preparation of the methylated 3HC fluorophore **6** (Scheme 1)



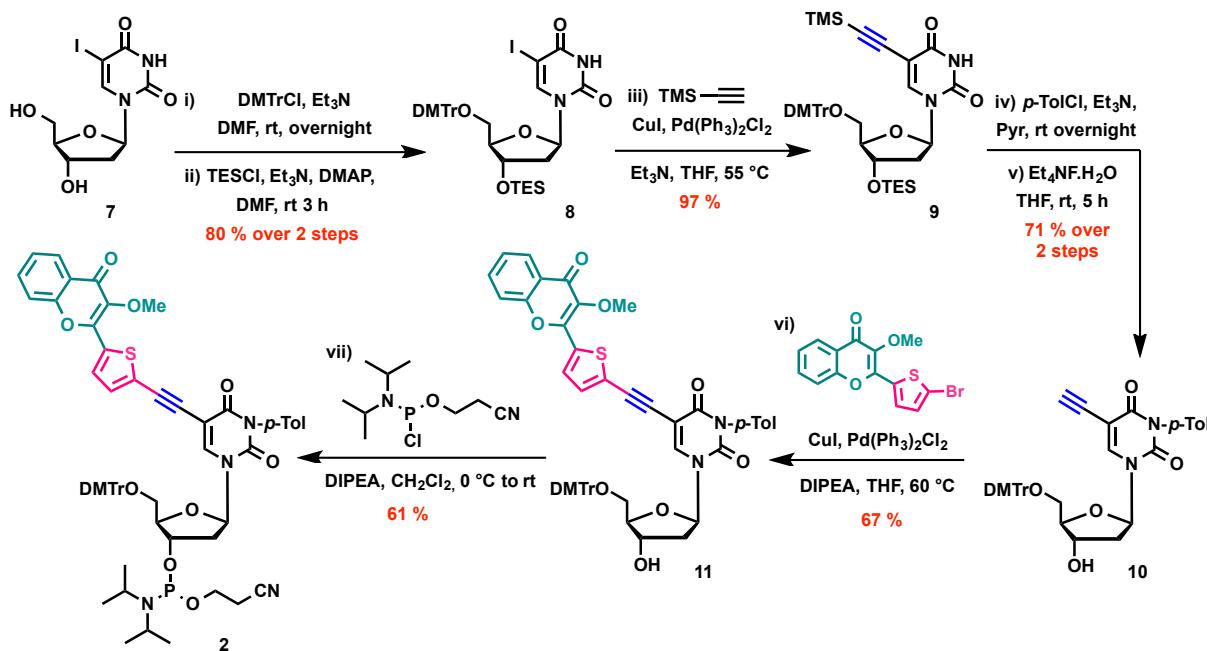
2-(5-Bromothien-2-yl)-3-hydroxy-chromen-4-one (5): To a stirred solution of *o*-hydroxyacetophenone **3** (1.16 mL, 9.65 mmol) and 5-bromo-2-thiophenecarbaldehyde **4** (1.15 mL, 9.65 mmol) in ethanol (50 mL) was dropwise added a 5 M NaOH solution (6 mL). The reaction mixture was stirred 48 h at rt before a dropwise addition of 30 % aq. hydrogen peroxide solution (2 mL). The resulting mixture was stirred 5 h at rt, and then poured into cold water (300 mL) and acidified with acetic acid to pH 4. The resulting precipitate was filtered and thoroughly washed with water and cyclohexane to provide as a yellow solid the chromone **5** (1.79 g, 57 %). C₁₃H₇BrO₃S (323.16). R_f = 0.39 (cyclohexane/EA = 7:3); ¹H-NMR (CDCl₃, 200 MHz): δ = 7.19 (d, ³J=4.0 Hz, 1H, H_β-thioph.), 7.41 (td, ³J=8.0 Hz, ⁴J=0.8 Hz, 1H, H₆), 7.54 (d, ³J=8.4 Hz, 1H, H₈), 7.71 (ddd, ³J=8.4 Hz, ³J=8.0 Hz, ⁴J=1.4 Hz, 1H, H₇), 7.73 (d, ³J=4.0 Hz, 1H, H_α-thioph.), 8.22 (dd, ³J=8.0 Hz, ⁴J=1.4 Hz, 1H, H₅); ¹³C-NMR (DMSO-*d*₆, 50 MHz): δ = 112.2 (*Th*-Br), 117.8 (C₈), 121.6 (C_α), 122.0 (C₁₀), 122.9 (C₆), 124.7 (C₅), 129.3 (C_β), 131.6 (C₇), 136.1 (C₂-*Th*), 141.6 (C₃), 146.5 (C₂), 153.3 (C₉), 176.0 (C₄); MS (ESI⁺, MeOH): m/z: 345.0, 347.0 [M+Na]⁺.



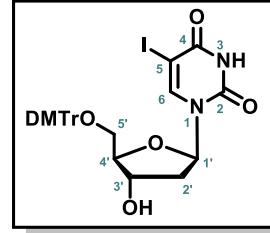
3-Methoxy-2-(5-bromothien-2-yl)-chromen-4-one (6): To a stirred suspension of **5** (1.59 g, 4.92 mmol) in CH₂Cl₂ (35 mL) were added a 4.5 M aq. KOH solution (5 mL), 18-crown-6 (65 mg, 5 mol%) and dimethyl sulfate (930 μL, 9.84 mmol). The reaction mixture became homogenous in 5 min and was stirred at rt for 1 h. After quenching by addition of H₂O (40 mL), the organic layer was extracted with CH₂Cl₂ (3 x), dried over MgSO₄, filtered and the volatiles were removed *in vacuo*. The residue was purified by flash chromatography on silica gel eluted with cyclohexane/EA mixture (10:1 → 4:1, v/v) to provide the methylated chromone **6** as a yellow solid (1.30 g, 78 %). C₁₄H₉BrO₃S (337.19). ¹H-NMR (CDCl₃, 200 MHz): δ = 4.07 (s, 3H, CH₃), 7.16 (d, ³J=3.8 Hz, 1H, H_β-thioph.), 7.38 (td, ³J=8.0 Hz, ⁴J=1.0 Hz, H₆), 7.49 (dd, ³J=8.4 Hz, ⁴J=1.0 Hz, 1H, H₈), 7.62–7.79 (m, 1H, H₇), 7.66 (d, ³J=3.8 Hz, 1H, H_α-thioph.), 8.22 (dd, ³J=8.0 Hz, ⁴J=1.6 Hz, 1H, H₅); ¹³C-NMR (CDCl₃, 50 MHz): δ = 59.7 (OMe), 117.7 (C₈), 119.8 (*Th*-Br), 124.2 (C₁₀), 124.7 (C₆), 125.7 (C₅), 129.2 (C_α), 130.2 (C_β), 132.7 (C₂-*Th*), 133.5 (C₇), 138.4 (C₃), 150.2 (C₂), 154.7 (C₉), 173.8 (C₄); HRMS (ESI⁺): m/z calcd for C₁₄H₁₀BrO₃S: 336.9529 [M+H]⁺; found 336.9531.



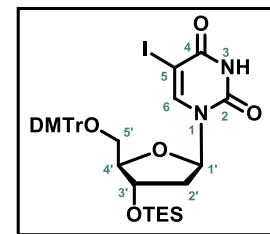
1.2 Preparation of the fluorescent amidite 2 (Scheme 2)



5'-O-(4,4'-Dimethoxytrityl)-5-iodo-2'-deoxyuridine (7b**)⁴:** To a stirred solution of 5-iodo-2'-deoxyuridine **7** (1.00 g, 2.82 mmol) in dry pyridine (20 mL) were added triethylamine (830 μ L, 5.64 mmol) and DMTrCl (1.00 g, 2.97 mmol). The reaction mixture was stirred at rt for 1 d and then quenched with H₂O. The organic phase was extracted with CH₂Cl₂ (3 x), dried over Na₂SO₄, filtered and the volatiles were removed in vacuo. The resulting crude was purified by flash chromatography on silica gel eluted with CH₂Cl₂/MeOH (1:0 \rightarrow 96:4, v/v) to provide the desired compound **7b** as a white solid (1.81 g, 98 %). C₃₀H₂₉IN₂O₇ (656.46). R_f = 0.41 (CH₂Cl₂/MeOH, = 95:5). ¹H-NMR (CD₂Cl₂, 200 MHz): δ = 2.27 (ddd, ²J=13.6 Hz, ³J=7.6, 6.0 Hz, 1H, H_{2'}A), 2.45 (ddd, ²J=13.6 Hz, ³J=5.8, 2.8 Hz, 1H, H_{2'}B), 3.35 (d, ³J=3.4 Hz, 2H, H_{5'}), 3.78 (s, 6H, OCH₃), 4.02–4.14 (m, 1H, H_{4'}), 4.47–4.60 (m, 1H, H_{3'}), 6.26 (dd, ³J=7.6, 5.8 Hz, 1H, H_{1'}), 6.86 (d, ³J=8.7 Hz, 4H, H_{m-PhOMe}), 7.20–7.32 (m, 3H, H_{m-Ph}, H_{p-Ph}), 7.34 (d, ³J=8.7 Hz, 4H, H_{o-PhOMe}), 7.40–7.46 (m, 2H, H_{o-Ph}), 8.09 (s, 1H, H₆), 8.99 (NH); ¹³C-NMR (CD₂Cl₂, 50 MHz): δ = 41.7 (C_{2'}), 55.7 (OMe), 64.0 (C_{5'}), 68.8 (C₅), 72.6 (C_{3'}), 86.1 (C_{1'}), 87.0 (C_{4'}), 87.3 (5'-O-C_{IV}), 113.7 (C_{m-PhOMe}), 127.4 (C_{p-Ph}), 128.4 (C_{m-Ph}), 128.5 (C_{o-Ph}), 130.4 (C_{o-PhOMe}), 130.5 (C_{o-PhOMe}), 135.8 (C_{i-PhOMe}), 136.0 (C_{i-PhOMe}), 144.9 (C₆), 145.0 (C_{i-Ph}), 150.6 (C₂), 159.2 (C_{p-PhOMe}), 160.8 (C₄); MS (ESI⁺, MeOH) m/z: 679.3 [M+Na]⁺, 695.3 [M+K]⁺.



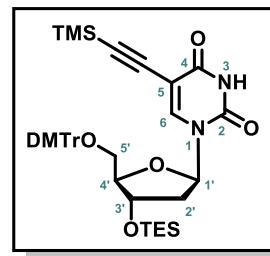
5'-O-(4,4'-Dimethoxytrityl)-5-iodo-3'-O-triethylsilyl-2'-deoxyuridine (8**):** To a stirred solution of **7b** (1.60 g, 2.19 mmol) in dry DMF (7 mL), previously cooled down to 0 °C, were sequentially added triethylamine (1.20 mL, 8.77 mmol), DMAP (81 mg, 0.66 mmol) and Et₃SiCl (441 μ L, 2.63 mmol). The reaction mixture was stirred at rt for 4 h and then carefully quenched with aq. NH₄Cl. The organic phase was extracted with CH₂Cl₂ (3 x), dried over MgSO₄, filtered and reduced in vacuo. The residue was purified by flash chromatography on silica gel eluted with toluene/Et₂O (95:5 \rightarrow 67:33, v/v) to provide the desired compound **8** as a light yellow foam (1.52 g, 90 %). C₃₆H₄₃IN₂O₇Si (770.73). R_f = 0.62 (toluene/Et₂O = 4:1). ¹H-NMR (CD₂Cl₂, 200 MHz): δ = 0.53 (q, ³J=7.8 Hz, 3H, CH₃), 0.89 (t, ³J=7.8 Hz, 2H, CH₂), 2.10–2.23 (m, 1H, H_{2'}A), 2.31–2.42 (m, 1H, H_{2'}B), 3.26 (dd, ²J=10.9 Hz, ³J=3.4 Hz, 1H, H_{5'}A), 3.36 (dd, ²J=10.9 Hz, ³J=2.9 Hz, 1H, H_{5'}B), 3.78 (s, 6H, OMe), 3.98–4.02 (m, 1H, H_{4'}), 4.39–4.44 (m, 1H, H_{3'}), 6.24 (dd, ³J=7.7, 5.7 Hz, 1H, H_{1'}), 6.85 (d, ³J=8.8 Hz, 4H, H_{m-PhOMe}), 7.12–7.28 (m, 3H, H_{m-Ph}, H_{p-Ph}), 7.34 (d, ³J=8.8 Hz, 4H, H_{o-PhOMe}), 7.40–7.46 (m, 2H, H_{o-Ph}), 8.14 (s, 1H, H₆), 8.72 (NH); ¹³C-NMR (CD₂Cl₂, 50 MHz): δ = 4.9 (CH₂), 6.8 (CH₃), 42.3 (C_{2'}), 55.6 (OMe), 63.6 (C_{5'}), 68.4 (C₅), 72.8 (C_{3'}), 86.2 (C_{1'}), 87.3 (5'-O-C_{IV}), 87.8 (C_{4'}), 113.7 (C_{m-PhOMe}), 127.3 (C_{p-Ph}), 128.4 (C_{o-Ph}), 128.4 (C_{m-Ph}), 130.4 (C_{o-PhOMe}), 130.5 (C_{o-PhOMe}), 135.8 (C_{i-PhOMe}), 135.9 (C_{i-PhOMe}), 144.9 (C₆), 145.0 (C_{i-Ph}), 150.2 (C₂), 159.2 (C_{p-PhOMe}), 160.5 (C₄); MS (ESI⁺, MeOH) m/z: 793.4 [M+Na]⁺, 809.3 [M+K]⁺. HRMS (ESI⁺): m/z calcd for C₃₆H₄₃IN₂NaO₇Si: 793.1776 [M+Na]⁺; found 793.1767.



⁴ Q. Xiao, R. T. Ranasinghe, A. M. P. Tang and T. Brown, *Tetrahedron*, 2007, **63**, 3483–3490.

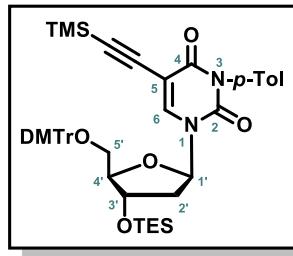
5'-O-(4,4'-Dimethoxytrityl)-5-trimethylsilylethylnyl-3'-O-triethylsilyl-2'-deoxyuridine (9**):**

To a stirred solution of **8** (810 mg, 0.946 mmol, previously azeotropically coevaporated with dry toluene) in DMF (10 mL) under argon, were sequentially added TMS-acetylene (203 μ L, 1.419 mmol, 1.5 eq), triethylamine (670 μ L, 4.729 mmol, 5 eq) and CuI (15 mg, 8 mol%)/PdCl₂(PPh₃)₂ (54 mg, 8 mol%) all together. The reaction mixture was warmed up to 60 °C for 3 h. The volatiles were removed *in vacuo* and the residue was purified by flash chromatography on silica gel eluted with toluene/Et₂O (95:5 → 66:34, v/v) to provide the desired compound **9** as a light yellow foam (683 mg, 97 %). C₄₁H₅₂N₂O₇Si₂ (741.03). R_f = 0.62 (pentane/Et₂O = 1:2). ¹H-NMR (CDCl₃, 200 MHz): δ = -0.03 (s, 9H, SiMe₃), 0.51 (q, ³J=7.8 Hz, 9H, CH₃), 0.88 (t, ³J=7.8 Hz, 6H, CH₂), 2.03–2.17 (m, 1H, H_{2'}A), 2.37 (ddd, ²J=13.0 Hz, ³J=5.6 Hz, ⁴J=1.8 Hz, 1H, H_{2'}B), 3.22 (dd, ²J=10.8 Hz, ³J=3.4 Hz, 1H, H_{5'}A), 3.45 (dd, ²J=10.8 Hz, ³J=2.7 Hz, 1H, H_{5'}B), 3.79 (s, 3H, OMe), 3.98–4.04 (m, 1H, H_{4'}), 4.26–4.34 (m, 1H, H_{3'}), 6.28 (dd, ³J=8.0, 5.4 Hz, 1H, H_{1'}), 6.85 (d, ³J=8.8 Hz, 4H, H_m-PhOMe), 7.15–7.32 (m, 3H, H_m-Ph, H_p-Ph), 7.35 (d, ³J=8.8 Hz, 4H, H_o-PhOMe), 7.40–7.49 (m, 2H, H_o-Ph), 8.11 (s, 1H, H_{6'}); ¹³C-NMR (CDCl₃, 50 MHz): δ = -0.4 (SiMe₃), 4.5 (SiCH₂), 6.6 (SiCH₂CH₃), 42.0 (C_{2'}), 55.2 (OMe), 63.2 (C_{5'}), 72.5 (C_{3'}), 86.0 (C_{1'}), 86.9 (5'-O-C_{IV}), 87.6 (C_{4'}), 94.8 (C≡C-TMS), 99.5 (C≡C-TMS), 100.3 (C₅), 113.3 (C_m-PhOMe), 126.9 (C_p-Ph), 127.9 (C_o-Ph), 128.0 (C_m-Ph), 129.9 (C_o-PhOMe), 135.4 (C_i-PhOMe), 135.6 (C_i-PhOMe), 142.8 (C₆), 144.4 (C_i-Ph), 149.0 (C₂), 158.5 (C_p-PhOMe), 161.3 (C₄); MS (ESI⁺, MeOH) *m/z*: 763.5 [M+Na]⁺, 779.5 [M+K]⁺. HRMS (ESI⁺): *m/z* calcd for C₄₁H₅₂N₂NaO₇Si₂: 763.3205 [M+Na]⁺; found 763.3192.



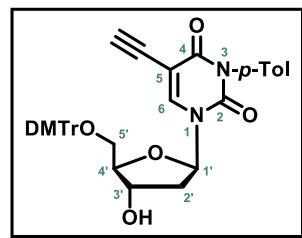
5'-O-(4,4'-Dimethoxytrityl)-3-N-(4-methylbenzoyl)-5-trimethylsilylethylnyl-3'-O-triethylsilyl-2'-deoxyuridine (9b**):**

To a stirred solution of **9** (1.190 g, 1.606 mmol) in dry acetonitrile (16 mL), previously cooled down to 0 °C, were sequentially added triethylamine (723 μ L, 5.139 mmol) and toluoyl chloride (297 μ L, 2.248 mmol). The reaction mixture was stirred at rt overnight and then carefully quenched with aq. NH₄Cl. The organic phase was extracted with CH₂Cl₂ (3 x), dried over MgSO₄, filtered and reduced *in vacuo*. The residue was purified by flash chromatography on silica gel eluted with toluene/Et₂O (99.5:0.5 → 96.5:3.5, v/v) to provide the desired compound **9b** as a beige foam (1.228 g, 89 %). C₄₉H₅₈N₂O₈Si₂ (859.16). R_f = 0.34 (toluene/Et₂O = 98:2). ¹H-NMR (CDCl₃, 200 MHz): δ = -0.00 (s, 9H, SiMe₃), 0.53 (q, ³J=7.8 Hz, 9H, CH₃), 0.89 (t, ³J=7.8 Hz, 6H, CH₂), 2.11–2.25 (m, 1H, H_{2'}A), 2.34–2.39 (m, 1H, H_{2'}B), 2.44 (s, 3H, p-CH₃), 3.27 (dd, ²J=10.8 Hz, ³J=3.6 Hz, 1H, H_{5'}A), 3.41 (dd, ²J=10.8 Hz, ³J=3.0 Hz, 1H, H_{5'}B), 3.79 (s, 3H, OMe), 4.00–4.09 (m, 1H, H_{4'}), 4.31–4.41 (m, 1H, H_{3'}), 6.23 (dd, ³J=8.0, 5.6 Hz, 1H, H_{1'}), 6.88 (d, ³J=8.8 Hz, 4H, H_m-PhOMe), 7.13–7.31 (m, 3H, H_m-Ph, H_p-Ph), 7.33 (d, ³J=8.2 Hz, ⁴J=0.6 Hz, 2H, H_m-Tol), 7.38 (dd, ³J=8.8 Hz, ⁴J=0.8 Hz, 4H, H_o-PhOMe), 7.49 (dd, ³J=8.4 Hz, ⁴J=1.4 Hz, 2H, H_o-Ph), 7.81 (d, ³J=8.2 Hz, 2H, H_o-Tol), 8.17 (s, 1H, H_{6'}); ¹³C-NMR (CDCl₃, 50 MHz): δ = -0.4 (SiMe₃), 4.9 (SiCH₂), 6.8 (SiCH₂CH₃), 22.0 (p-CH₃), 42.4 (C_{2'}), 55.6 (OMe), 63.6 (C_{5'}), 72.9 (C_{3'}), 86.8 (C_{1'}), 87.3 (5'-O-C_{IV}), 88.1 (C_{4'}), 95.1 (C≡C-TMS), 100.0 (C≡C-TMS), 100.5 (C₅), 113.6 (C_m-PhOMe), 127.3 (C_p-Ph), 128.3 (C_o-Ph), 128.4 (C_m-Ph), 129.0 (C_i-Tol), 130.3 (C_o-PhOMe), 130.4 (C_m-Tol), 130.9 (C_o-Tol), 135.9 (C_i-PhOMe), 142.9 (C₆), 145.0 (C_i-Ph), 147.5 (C_p-Tol), 148.6 (C₂), 159.1 (C_p-PhOMe), 160.9 (C₄), 168.2 (C≡O-Tol); MS (ESI⁺, MeOH) *m/z*: 881.6 [M+Na]⁺, 897.6 [M+K]⁺. HRMS (ESI⁺): *m/z* calcd for C₄₉H₅₈N₂NaO₈Si₂: 881.3624 [M+Na]⁺; found 881.3601.

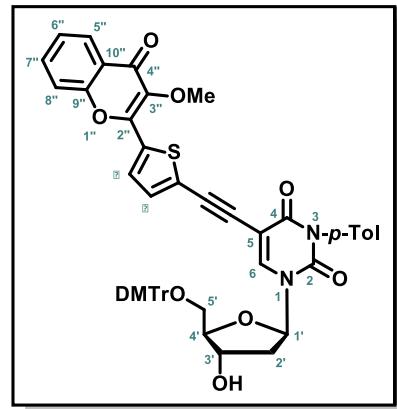


5'-O-(4,4'-Dimethoxytrityl)-3-N-(4-methylbenzoyl)-5-ethynyl-2'-deoxyuridine (10**):**

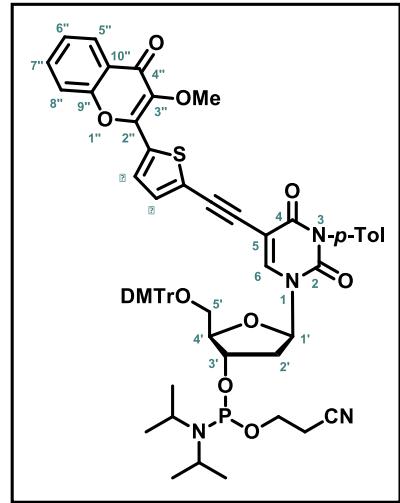
To a stirred solution of **9b** (1.90 g, 2.21 mmol) in THF (22 mL), previously cooled down to 0 °C, was portionwise added Et₄NF·H₂O (2.00 g, 13.27 mmol). The reaction mixture was stirred at rt for 1.5 h and then carefully quenched with aq. NH₄Cl. The organic phase was extracted with Et₂O (3 x), dried over MgSO₄, filtered and reduced *in vacuo*. The residue was purified by flash chromatography on silica gel eluted with toluene/Et₂O (93:7 → 1:1, v/v) to provide the desired compound **10** as a light yellow foam (1.23 g, 83 %). C₄₀H₃₆N₂O₈ (672.72). R_f = 0.22 (toluene/Et₂O = 7:2). ¹H-NMR (CD₂Cl₂, 200 MHz): δ = 2.25–2.43 (m, 1H, H_{2'}A), 2.44 (s, 3H, p-CH₃), 2.44–2.55 (m, 1H, H_{2'}B), 2.97 (s, 1H, C≡CH), 3.32 (dd, ²J=10.8 Hz, ³J=3.0 Hz, 1H, H_{5'}A), 3.41 (dd, ²J=10.8 Hz, ³J=3.4 Hz, 1H, H_{5'}B), 3.80 (s, 3H, OMe), 4.03–4.12 (m, 1H, H_{4'}), 4.52–4.62 (m, 1H, H_{3'}), 6.24 (dd, ³J=7.2, 6.0 Hz, 1H, H_{1'}), 6.89 (d, ³J=8.6 Hz, 4H, H_m-PhOMe), 7.20–7.31 (m, 3H, H_m-Ph, H_p-Ph), 7.33 (d, ³J=8.1 Hz, 2H, H_m-Tol), 7.39 (d, ³J=8.6 Hz, 4H, H_o-PhOMe), 7.48 (dd, ³J=8.2 Hz, ⁴J=1.4 Hz, 2H, H_o-Ph), 7.81 (d, ³J=8.1 Hz, 2H, H_o-Tol), 8.21 (s, 1H, H_{6'}); ¹³C-NMR (CD₂Cl₂, 50 MHz): δ = 22.0 (p-CH₃), 41.9 (C_{2'}), 55.6 (OMe), 63.8 (C_{5'}), 72.4 (C_{3'}), 74.4 (C≡CH), 82.4 (C≡CH), 86.6 (C_{1'}), 87.1 (C_{4'}), 87.5 (5'-O-C_{IV}), 99.3 (C₅), 113.7 (C_m-PhOMe), 127.3 (C_p-Ph), 128.3 (C_o-Ph), 128.5 (C_m-Ph), 128.8 (C_i-Tol), 130.4 (C_m-Tol), 130.4 (C_o-PhOMe), 130.9 (C_o-Tol), 135.7 (C_i-PhOMe), 136.0 (C_i-PhOMe), 144.0 (C₆), 145.0 (C_i-Ph), 147.7 (C_p-Tol), 148.6 (C₂), 159.2 (C_p-PhOMe), 161.0 (C₄), 168.1 (C≡O-Tol); MS (ESI⁺, MeOH) *m/z*: 695.5 [M+Na]⁺, 711.5 [M+K]⁺. HRMS (ESI⁺): *m/z* calcd for C₄₀H₃₆N₂NaO₈: 695.2364 [M+Na]⁺; found 695.2357.



5'-O-(4,4'-Dimethoxytrityl)-3-N-(4-methylbenzoyl)-5-(5-Methoxy-4-oxo-chromen-2-yl)thien-2-yl)ethynyl-2'-deoxyuridine (11**):** To a stirred solution of **10** (300 mg, 0.446 mmol) and **6** (206 mg, 0.580 mmol, 1.3 eq) in THF (10 mL) under argon, were sequentially added triethylamine (311 μ L, 2.230 mmol, 5 eq), and CuI (6 mg, 7 mol%)/PdCl₂(PPh₃)₂ (22 mg, 7 mol%) all together. The reaction mixture was warmed to 60 °C for 2 h. The volatiles were removed *in vacuo* and the residue was purified by flash chromatography on silica gel eluted with toluene/EA (93:7 → 3:2, v/v) to provide the desired compound **11** as a yellow foam (276 mg, 67 %). C₅₄H₄₄N₂O₁₁S (929.0). R_f = 0.38 (toluene/EA = 7:3). ¹H-NMR (CDCl₃, 200 MHz): δ = 2.39–2.49 (m, 1H, H_{2'A}), 2.43 (s, 3H, *p*-CH₃), 2.58 (ddd, ²J=13.6 Hz, ³J=5.8, 2.9 Hz, 1H, H_{2'B}), 3.37–3.51 (m, 2H, H_{5'}), 3.73 (s, 3H, OMe), 3.74 (s, 3H, OMe), 3.97 (s, 3H, 3"-OMe), 4.10–4.19 (m, 1H, H_{4'}), 4.59–4.71 (m, 1H, H_{3'}), 6.35 (dd, ³J=5.0, 4.0 Hz, 1H, H_{1'}), 6.82 (d, ³J=8.8 Hz, 2H, H_{m-PhOMe}), 6.84 (d, ³J=8.8 Hz, 2H, H_{m-PhOMe}), 6.87 (d, ³J=4.0 Hz, 1H, H β -thioph.), 7.15–7.35 (m, 6H, H_{6''}, H_{m-Tol}, H_{m-Ph}, H_{p-Ph}), 7.37 (d, ³J=8.8 Hz, 4H, H_{o-PhOMe}), 7.42–7.51 (m, 3H, H_{o-Ph}, H_{8''}), 7.63–7.72 (m, 1H, H_{7''}), 7.69 (d, ³J=4.0 Hz, 1H, H α -thioph.), 7.85 (d, ³J=8.2 Hz, 2H, H_{o-Tol}), 8.22 (dd, ³J=8.0 Hz, ⁴J=1.4 Hz, 1H, H_{5''}); 8.37 (s, 1H, H₆); ¹³C-NMR (CDCl₃, 50 MHz): δ = 21.9 (*p*-CH₃), 41.9 (C_{2'}), 55.2 (OMe), 59.7 (OMe), 63.4 (C_{5'}), 72.2 (C_{3'}), 86.4 (C_{1'}), 86.6 (Th-C≡C-), 86.9 (Th-C≡C-), 86.9 (C_{4'}), 87.2 (5'-O-C_{IV}), 99.8 (C₅), 113.4 (C_{m-PhOMe}), 117.7 (C_{8''}), 124.2 (C_{10''}), 124.8 (C_{6''}), 125.7 (C_{5''}), 127.1 (C_{p-Ph}), 127.9 (C_{o-Ph}), 128.1 (C_{m-Ph}), 128.1 (Th-C≡C-), 128.5 (C_{i-Tol}), 128.9 (C α), 129.8 (C_{m-Tol}), 130.0 (C_{o-PhOMe}), 130.7 (C_{o-Tol}), 132.4 (C_{2''-Th}), 132.6 (C β), 133.6 (C_{7''}), 135.3 (C_{i-PhOMe}), 135.5 (C_{i-PhOMe}), 138.8 (C_{3''}), 142.3 (C₆), 144.2 (C_{i-Ph}), 146.8 (C_{p-Tol}), 148.1 (C₂), 150.6 (C_{2''}), 154.7 (C_{9''}), 158.6 (C_{p-PhOMe}), 160.2 (C₄), 167.5 (C_{O-Tol}), 173.9 (C_{4''}); MS (ESI⁺, MeOH) *m/z*: 950.9 [M+Na]⁺, 966.9 [M+K]⁺. HRMS (ESI⁺): *m/z* calcd for C₅₄H₄₄N₂NaO₁₁S: 951.2558 [M+H]⁺; found 951.2531



5'-O-(4,4'-Dimethoxytrityl)-3-N-(4-methylbenzoyl)-5-(5-(3-Methoxy-4-oxo-chromen-2-yl)thien-2-yl)ethynyl-2'-deoxyuridine, 3'-(2-cyanoethyl)-N,N-diisopropyl-phosphoramidite (12**):** To a stirred solution of **11** (280 mg, 0.301 mmol, beforehand azeotropically coevaporated with dry toluene) in CH₂Cl₂ (2.3 mL), previously cooled down to 0 °C, were sequentially added DIPEA (212 μ L, 1.206 mmol, 4 eq) and 2-cyanoethyl-N,N-diisopropylchlorophosphoramidite (136 μ L, 0.603 mmol, 2 eq). The reaction mixture was stirred at rt for 1 h. The volatiles were removed *in vacuo* and the residue was purified by flash chromatography on silica gel eluted with toluene/EtOAc (95:5 → 65:35, v/v) to provide the desired compound **12** as a yellow foam (230 mg, 68 %). C₆₃H₆₁N₄O₁₂PS (1129.22). R_f = 0.48 (toluene/EA = 8:2). ¹H-NMR (CD₃CN, 200 MHz): δ = 1.06–1.19 (m, 12H, N(CH(CH₃)₂)₂), 2.45 (m, 3H, *p*-CH₃), 2.52–2.68 (m, 4H, H_{2'}, -CH₂CN), 3.35–3.43 (m, 2H, H_{5'}), 3.52–3.67 (m, 2H, N(CH(CH₃)₂)₂), 3.70 (s, 3H, OMe), 3.71 (s, 3H, OMe), 3.74–3.82 (m, 1H, POCH₂-), 3.93 (s, 3H, 3"-OMe), 4.16–4.23 (m, 1H, H_{4'}), 4.68–4.86 (m, 1H, H_{3'}), 6.13 (dd, ³J=7.2, 6.2 Hz, 1H, H_{1'}), 6.16 (dd, ³J=7.6, 7.4 Hz, 1H, H_{1'}), 6.84 (d, ³J=8.8 Hz, 2H, H_{m-PhOMe}), 6.86 (d, ³J=8.8 Hz, 2H, H_{m-PhOMe}), 6.96 (d, ³J=4.0 Hz, H β -thioph.), 7.14–7.31 (m, 3H, H_{m-Ph}, H_{p-Ph}), 7.35–7.43 (m, 3H, H_{m-Tol}, H_{6''}), 7.41 (d, ³J=8.8 Hz, 4H, H_{o-PhOMe}), 7.47–7.54 (m, 2H, H_{o-Ph}), 7.63 (d, ³J=7.6 Hz, H_{8''}), 7.76 (ddd, ³J=8.6, 7.0 Hz, ⁴J=1.6 Hz, 1H, H_{7''}), 7.80 (d, ³J=4.0 Hz, 1H, H α -thioph.), 7.94 (d, ³J=8.2 Hz, 2H, H_{o-Tol}), 8.20 (dd, ³J=8.0 Hz, ⁴J=1.6 Hz, 1H, H_{5''}); 8.38 (s, 1H, H₆), 8.40 (s, 1H, H₆); ³¹P-NMR (CD₃CN, 81 MHz): δ = 148.0, 148.1. HRMS (ESI⁺): *m/z* calcd for C₆₃H₆₁N₄NaO₁₂PS: 1151.3637 [M+H]⁺; found 1151.3665.



2. ODN synthesis, purification & physical characterizations

2.1 ODN synthesis and purification:

The ODN synthesis was performed on an Expedite 8900 DNA synthesizer (Applied Biosystem) using the “trityl off” mode and ultra-mild Pac phosphoramidite chemistry on a 0.2 μmol scale. Reagents and solvents, as well as dT, Ac-dC, Pac-dA, and dmF-dG or iPr-Pac-dG phosphoramidites were purchased from Link Technologies and Chemgenes. The standard DNA assembly protocol “DMT-off” was used except for the following modifications: 5-Ethylthio-1H-tetrazole (ETT) was used as activating agent; Pac-anhydride was used for capping; a longer coupling time (1200 s) was applied to the 3HC phosphoramidite. Non-labeled ODNs used as wild-type sequences were purchased from Microsynth AG. The ODNs were cleaved from the solid support and deprotected with concentrated aqueous ammonia at room temperature for 12 h. The ODNs were analyzed (0.5 mL/min) and purified (2.5 mL/min) by RP HPLC (HPLC apparatus: WatersTM 600 Controller with WatersTM 996 Photodiode Array Detector and Jasco LC-Net II / ADC apparatus. Columns: analytical, 300 × 4.60 mm, 5 μm particle size, Clarity® 100Å, Phenomenex®; semi-preparative, Clarity® 5μ Oligo-RP column 250 x 10 mm Phenomenex®). The following gradient system was used: 100 % A -(30 min) → 60 % A / 40 % B -(5 min) → 100 % B -(5 min) → 100 % A with A=Buffer pH 7.0 (1.9 L of MilliQ® water, 160 mL acetonitrile, 28 mL triethylamine, 12 mL of acetic acid) and B=0.2 CH₃CN:0.8 Buffer.

Fig S1. HPLC profile of DNA1 single strand ODN (254 nm left – 390 nm right)

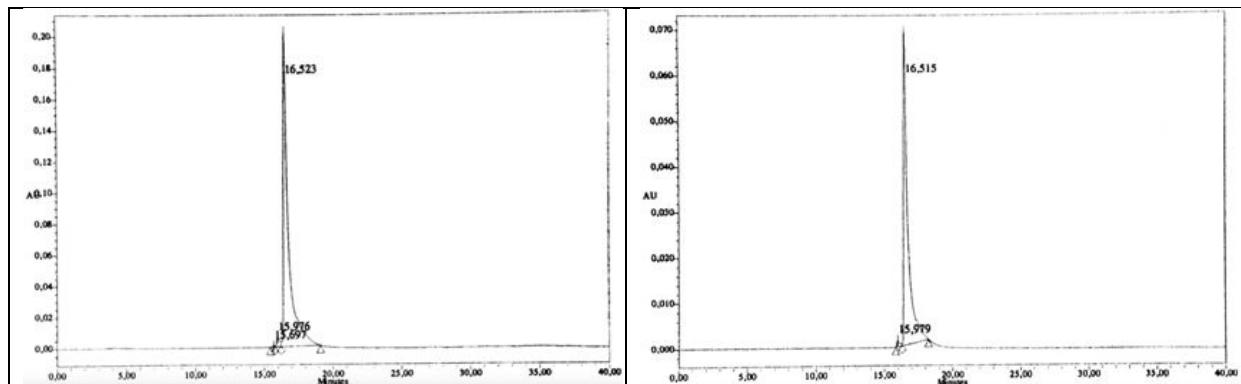


Fig S2. HPLC profile of DNA3 single strand ODN (254 nm left – 390 nm right)

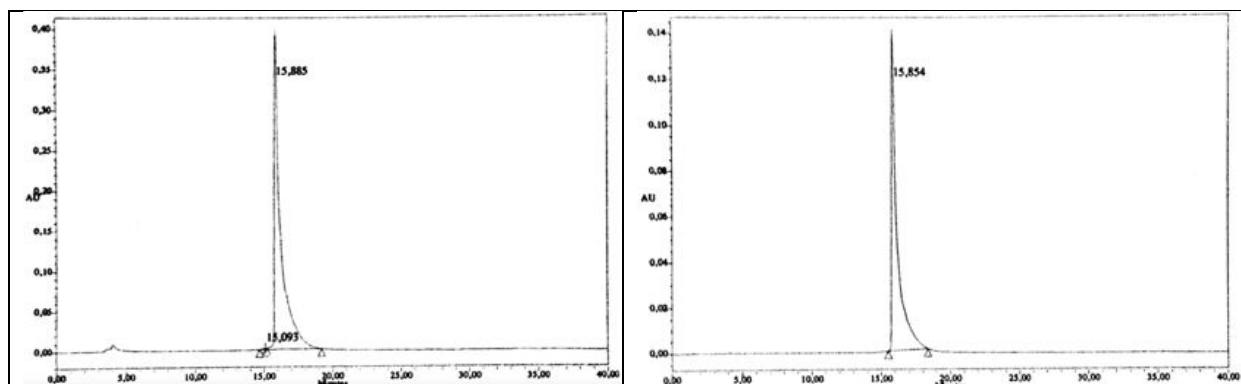


Fig S3. HPLC profile of DNA5 single strand ODN (254 nm left – 390 nm right)

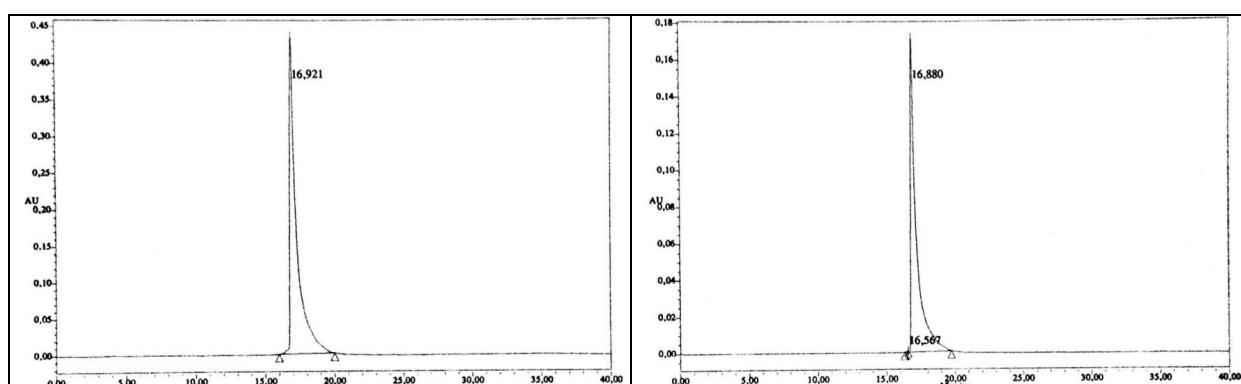


Fig S4. HPLC profile of DNA7 single strand ODN (254 nm left – 390 nm right)

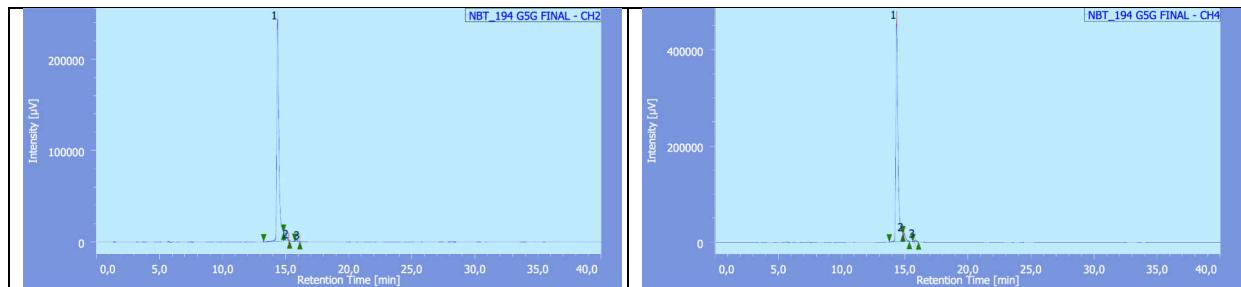
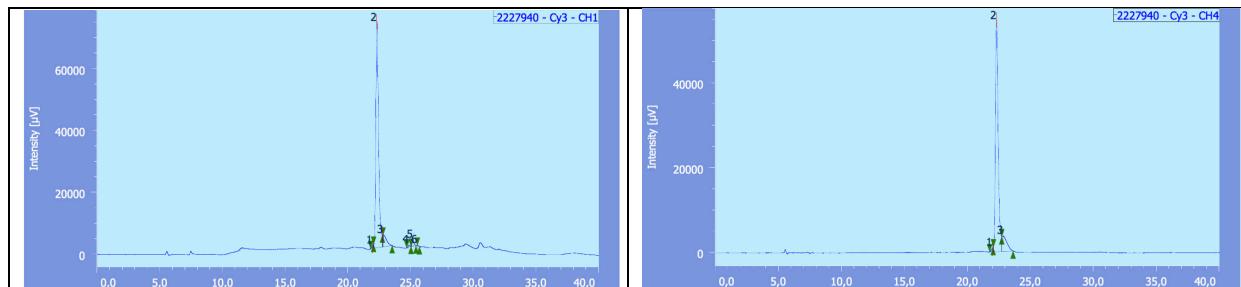


Fig S5. HPLC profile of DNA9 single strand ODN (254 nm left – 550 nm right)

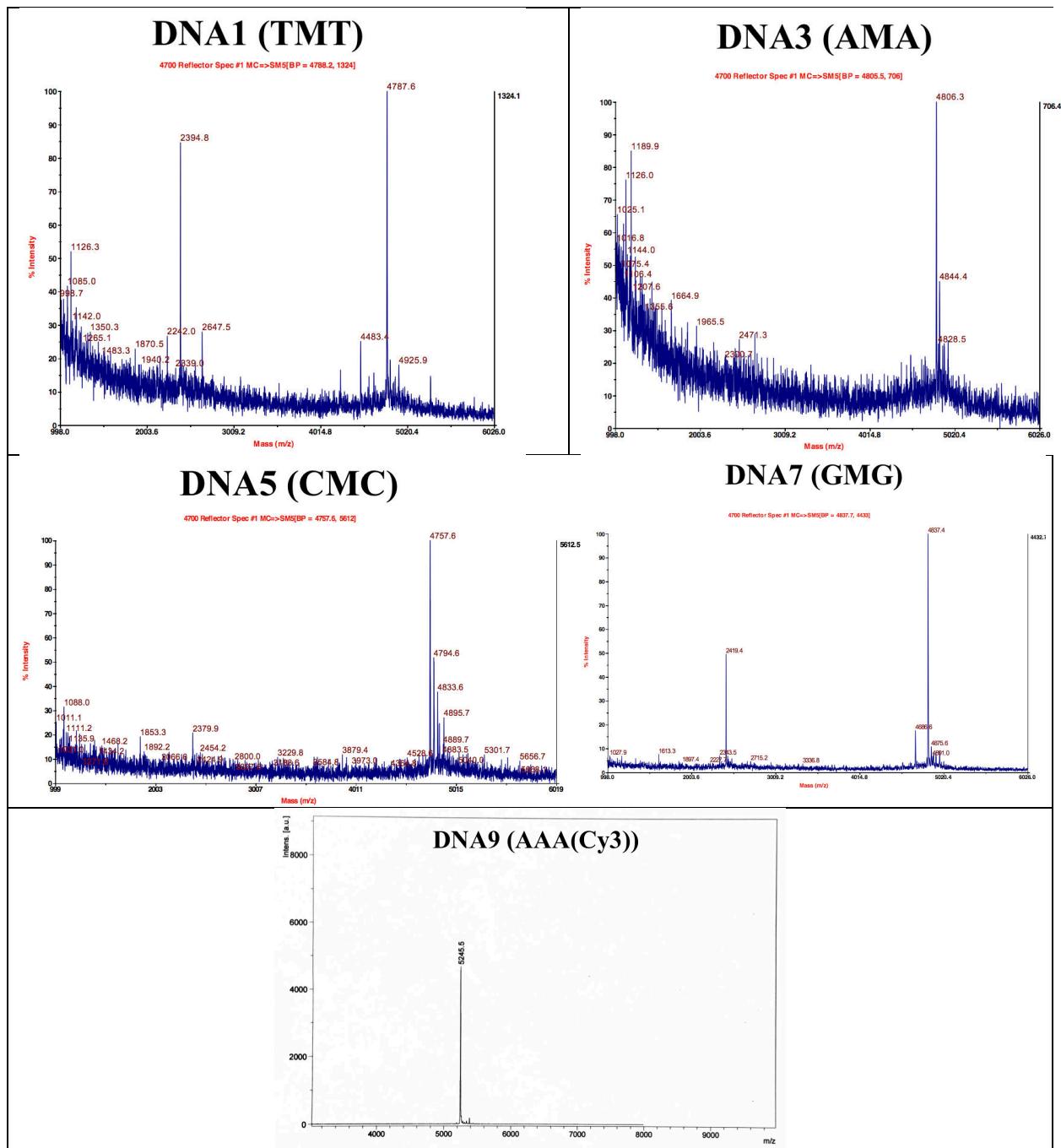


2.2 MALDI-TOF/TOF analysis of ODNs:

Dibasic Ammonium Citrate (DAC) (98% capillary GC) was obtained from Sigma. Acetonitrile of HPLC grade was purchased from Sigma-Aldrich. Ultrapure 3-Hydroxypicolinic Acid (3-HPA) MALDI matrix was purchased from Protea Biosciences. C4 pipette tips (Zip-Tip) were from Millipore. The samples (500 pmol) were diluted to 10 μ L of water and were desalting with a C4 pipette Tips (Zip-tip). The Zip-tip was activated before use with 2 x 5 μ L of water: CH₃CN (50:50) and 2 x 5 μ L of DAC (50 mg/ml diluted in water). The 10 μ L of the ODN solution were loaded on Zip-tip by drawing and expelling ten times. Next, the zip-tip was washed with 3 x 5 μ L of DAC (50 mg/mL) and 3 x 5 μ L of water. Elution was performed with 1.5 μ L of 3-HPA matrix (80 mg/mL, 50:50 CH₃CN:DAC) directly on MALDI plate. The ODN profile obtained in a ABSciex MALDI-TOF/TOF mass spectrometer in reflector mode with external calibration mixture (cal Mix 1+2 distributed by ABSciex). MALDI-TOF/TOF-MS analysis: MS spectra were recorded manually in a mass range of 500-6000 Da resulting from 400 laser shots of constant intensity fixed at 6200. Data were collected using 4000 series Explorer (AB SCIEX) experiments.

Table S1. Mass of the synthesized oligonucleotides.

DNA	Sequence	MALDI-TOF found (calcd) [M ⁺]
1	5'd-CGTTTTT M TTTTTGC-3'	4787.6 (4787.2)
3	5'd-CGTTTT A MATTTTGC-3'	4806.3 (4805.5)
5	5'd-CGTTTC M CTTTTGC-3'	4757.6 (4757.2)
7	5'd-CGTTTG M GTTTTGC-3'	4837.4 (4837.2)
9	5'd- Cy3 -GCAAAAAAAAAACG-3'	5245.5 (5233.5)



3. Spectroscopic characterizations

3.1 Denaturation studies and melting temperatures

The ODNs were analyzed in cacodylate buffer pH 7.0: 10 mM cacodylate, 150 mM NaCl.

Preparation of the single strand solution: the solution of the sample was prepared by mixing 400 µL of a stock solution of 20 mM cacodylate buffer solution pH 7.0, 80 µL of 1.5 M NaCl solution, 25 µL of 64 µM ssODN and 295 µL of MilliQ® water.

Preparation of the double strand solution: 400 µL of a stock solution of 20 mM cacodylate buffer solution pH 7.0, 80 µL of 1.5 M NaCl solution, 25 µL of 64 µM ODN1, 25 µL of 64 µM ODN2 and 270 µL of MilliQ® water.

3.2 Denaturation studies and melting temperatures

Melting curves were recorded by following the temperature-dependence of the absorbance changes of the sample in triplicate (2 µM concentration of each strand). Absorption spectra were recorded in a Peltier thermostated cell holder on a Cary 4 spectrophotometer (Varian). Wavelength for detection was 260 nm. The pathlength of cell was 1 cm. The temperature range for denaturation measurement was 5 – 80 °C. Speed of heating was 0.3 °C/min. Cacodylate buffer pH 7.0 was used (10 mM cacodylate buffer, 150 mM NaCl). The melting curves were converted into a plot of α versus temperature, where α represents

the fraction of single-strands in the duplex state. The melting temperatures were extracted from these curves after differentiation as described elsewhere.⁵

Table S2. Melting temperatures of the duplexes.^a

DNA	Duplexes	<i>Tm</i> (°C)		
		<i>labeled</i>	<i>Wild Type</i> ^a	<i>ΔTm</i> (°C) ^b
2	T M T-AAA	45.5	49.5 [50.2]	-4.0
	T M T-ATA	37.5	40.5 [42.2]	-3.0
	T M T-ACA	39.5	39 [42]	0.5
	T M T-AGA	37.5	40.7 [42.9]	-3.2
4	A M A-TAT	41.4	47.3 [47.4]	-5.9
	A M A-TTT	36.9	39.8 [40.6]	-2.9
	A M A-TCT	40.0	38.5 [39.3]	1.5
	A M A-TGT	35.9	42.7 [42.8]	-6.8
6	C M C-GAG	46.9	51.3 [53.7]	-4.4
	C M C-GTG	37.9	41.1 [46.7]	-3.2
	C M C-GCG	42.9	42.1 [45.3]	0.8
	C M C-GGG	41.9	46.5 [49.5]	-4.6
8	G M G-CAC	51.9	54.8 [55.1]	-2.9
	G M G-CTC	45.5	46.6 [47.5]	-1.1
	G M G-CCC	49.5	43.5 [44.8]	6.0
	G M G-CGC	46.5	49.7 [51]	-3.2
<hr/>				
DNA	Duplexes	<i>T_m</i> (°C)		
		<i>labeled</i>	<i>Wild Type</i> ^a	<i>ΔT_m</i> (°C) ^b
10	TTT-AAA(<i>Cy3</i>)	51.9	49.5 [50.2]	2.0
11	T M T-AAA(<i>Cy3</i>)	47.5	45.5 ^c	2.0

^a *T_m* of the corresponding duplex formed from unmodified ODNs and its theoretical values given in square brackets ; ^b *ΔT_m* refers to the difference of *T_m* between the labeled and wild type ODNs. ^c Comparison with the mono-labeled duplex.

Fig S6. Melting temperature curves of wild Type double strands

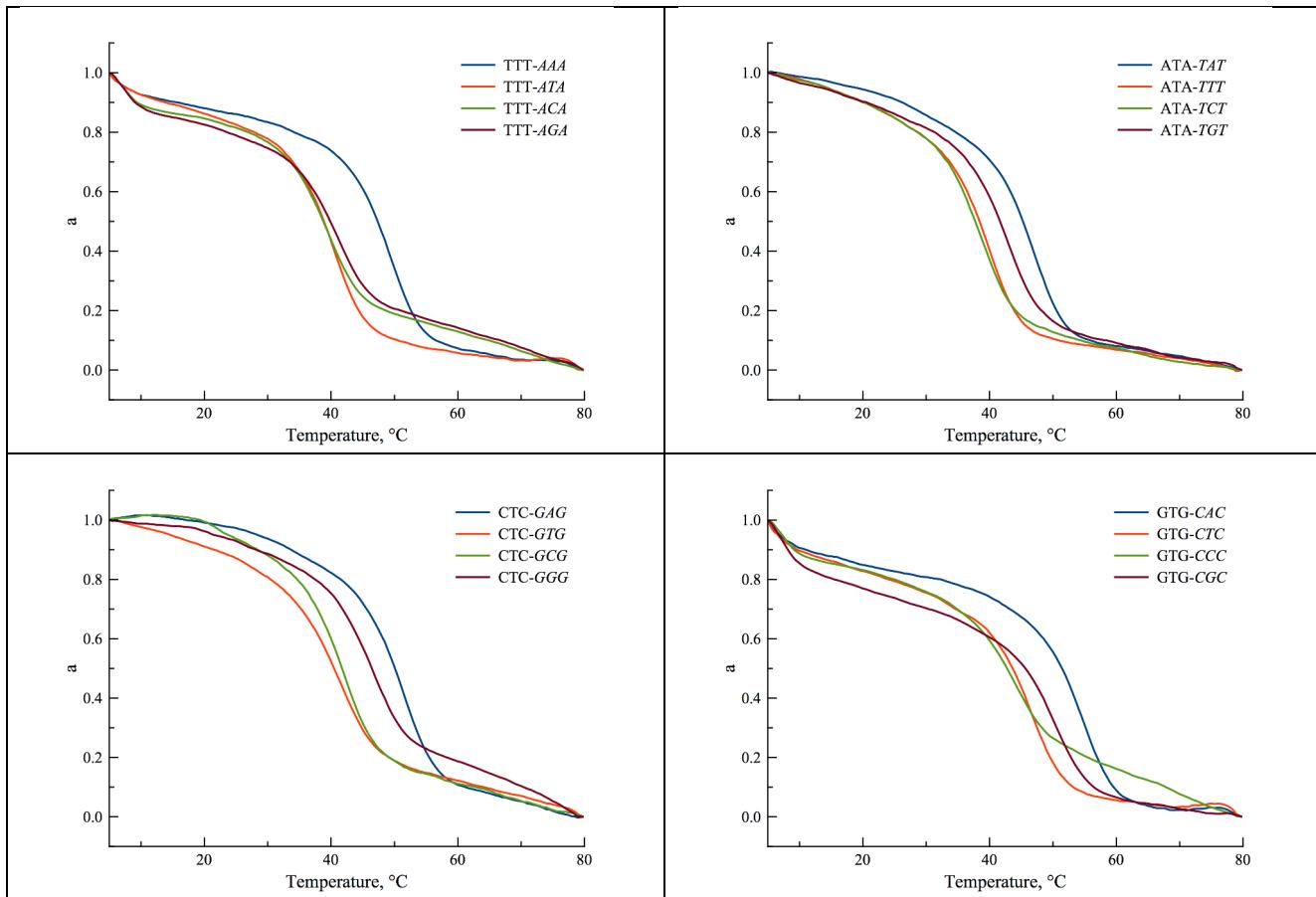
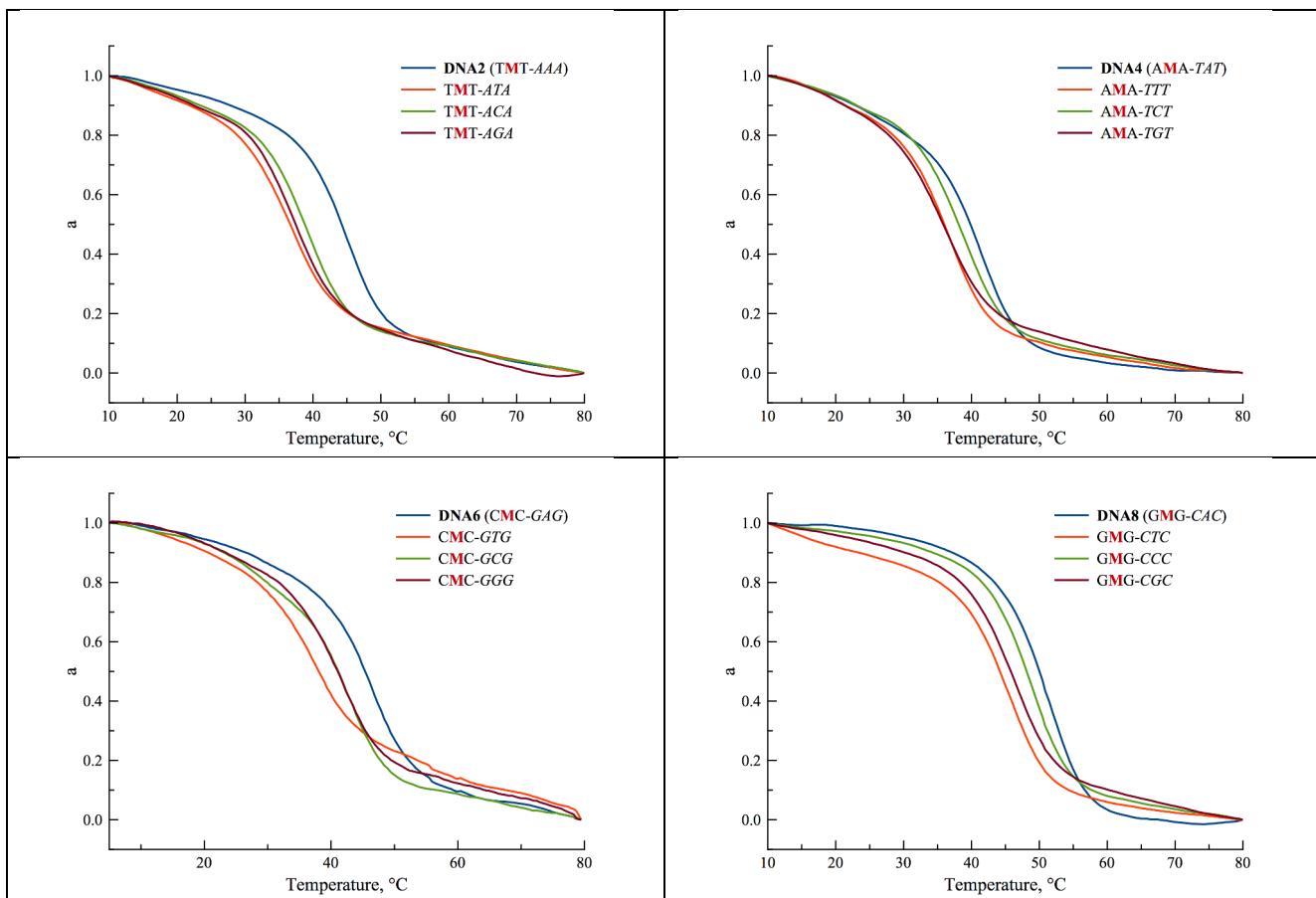


Fig S7. Melting temperature curves of labeled double strands



3.3 Circular dichroism

Circular dichroism spectra were recorded with 2 μ M solution of the canonical dsDNA and labeled dsDNA (3HC (**M**) opposite **A**, **T**) in buffer pH 7.0 (10 mM cacodylate buffer, 150 mM NaCl) at 25 °C on a Jasco J-810 spectropolarimeter. Two maxima were observed in CD spectra: one negative at 249 nm, the other one positive at 282 nm.

Fig S8. Representative CD spectra of wild type and labeled duplexes

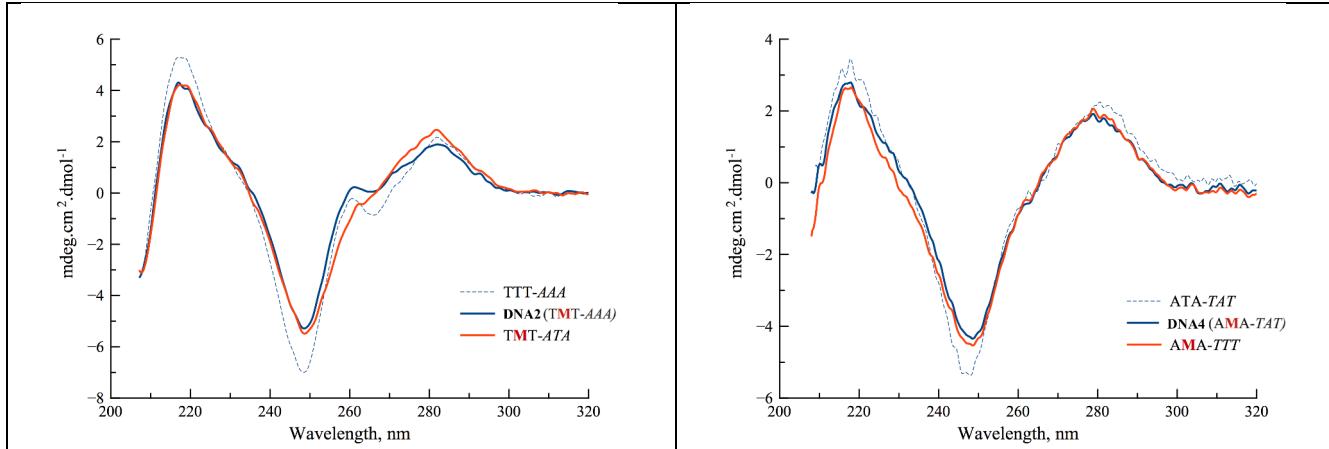
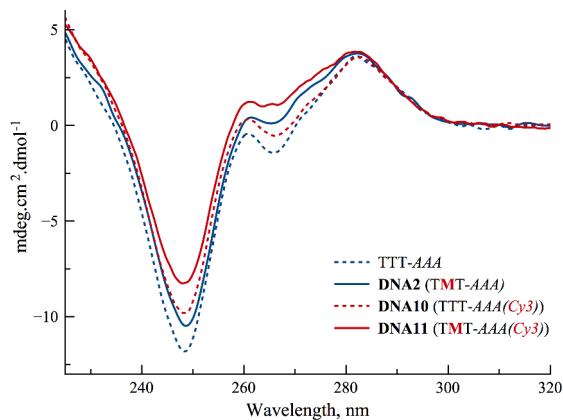


Fig S9. Representative CD spectra of wild type, mono- and double-labeled Cy3 duplexes



3.4 Absorbance and Fluorescence spectra

The absorption and fluorescence experiments were conducted in triplicate in pH 7.0 cacodylate buffer (10 mM cacodylate buffer, 150 mM NaCl, 1mM EDTA). The absorption spectra were recorded on a Cary 300 Scan spectrophotometer (Varian) using 1cm quartz cells at 20 °C. The fluorescence spectra were recorded on a FluoroMax 4.0 spectrofluorometer (Jobin Yvon, Horiba). They were recorded using excitation and emission slits of 2 nm and were corrected at excitation and emission. They were taken with absorbance of about 0.05 at 20 °C at the excitation wavelength mentioned in the corresponding experiments. The quantum yields were corrected according to the variation of the refractive index of the different solvents. FRET parameters (including energy transfer efficiency and rate constant, spectral overlap integral and Förster distance) were calculated according the definitions and the equations (*Eq1* to *Eq4*), herein.

- ❖ The efficiency of resonance energy transfer (*E* in %) was measured from the donor fluorescence quenching profile according the equation *Eq1*:

$$E = \left(1 - \frac{F_{DA}}{F_D} \right) \times 100 \quad (Eq1)$$

where F_D and F_{DA} are respectively, the relative fluorescence intensities in the absence and presence of the acceptor.

- ❖ Rate constant of energy transfer (k_T in s^{-1}) was calculated from the equation *Eq2*:

$$k_T = \frac{\left(\frac{F_D}{F_{DA}} - 1 \right)}{\tau_D} \quad (Eq2)$$

where τ_D is the lifetime of the donor

- The spectral overlap integral ($J(\lambda)$ in $M^{-1} \cdot cm^{-1} \cdot nm^4$) of the donor emission and the acceptor absorption were determined by using the equation Eq3:

$$J(\lambda) = \frac{\int_0^\infty F_D(\lambda) \epsilon_A(\lambda) \lambda^4 d\lambda}{\int_0^\infty F_D(\lambda) d\lambda} \quad (Eq3)$$

where $F_D(\lambda)$ is the fluorescence intensity of the donor in the wavelength range $[\lambda; \lambda + \Delta\lambda]$ and ϵ_A is the extinction coefficient of the acceptor at λ .

- Förster distance, corresponding to the distance where the resonance energy transfer is 50% efficient, was estimated according to the equation Eq4 et Eq5:

$$R_0 = 0.211 \sqrt[6]{(\kappa^2 n^{-4} Q_D J(\lambda))} \quad (Eq4)$$

$$R_0 = \sqrt[6]{\left(\frac{E}{1-E} \right) r^6} \quad (Eq5)$$

where κ^2 describes the relative orientation in space of the transition dipoles of the donor and acceptor (usually assumed to be equal to 2/3, which is appropriate for dynamic random averaging of the donor and acceptor); n is the refractive index of the medium (1.4 for biomolecules in aqueous solution) and Q_D is the quantum yield of the donor in the absence of acceptor.

Fig S10. UV spectra of labeled single and match double strands

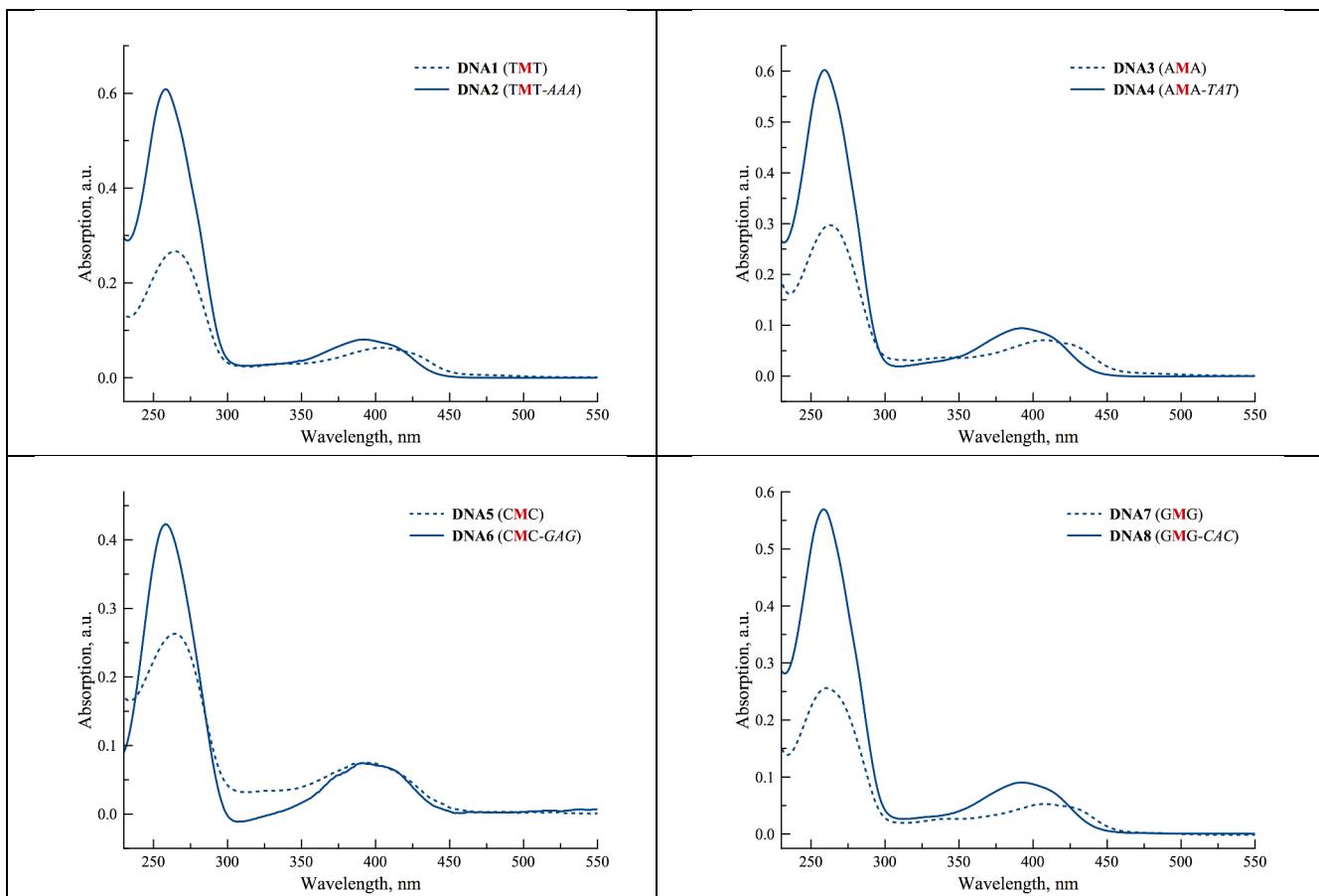


Fig S11. Absorbance spectra of wild type, mono- and double-labeled Cy3 duplexes

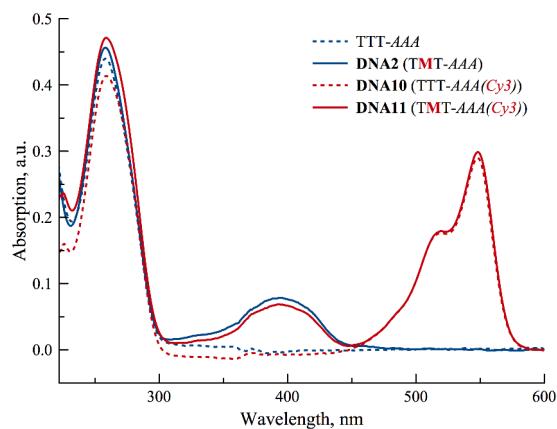


Fig S12. Fluorescence spectra of labeled single and match double strands

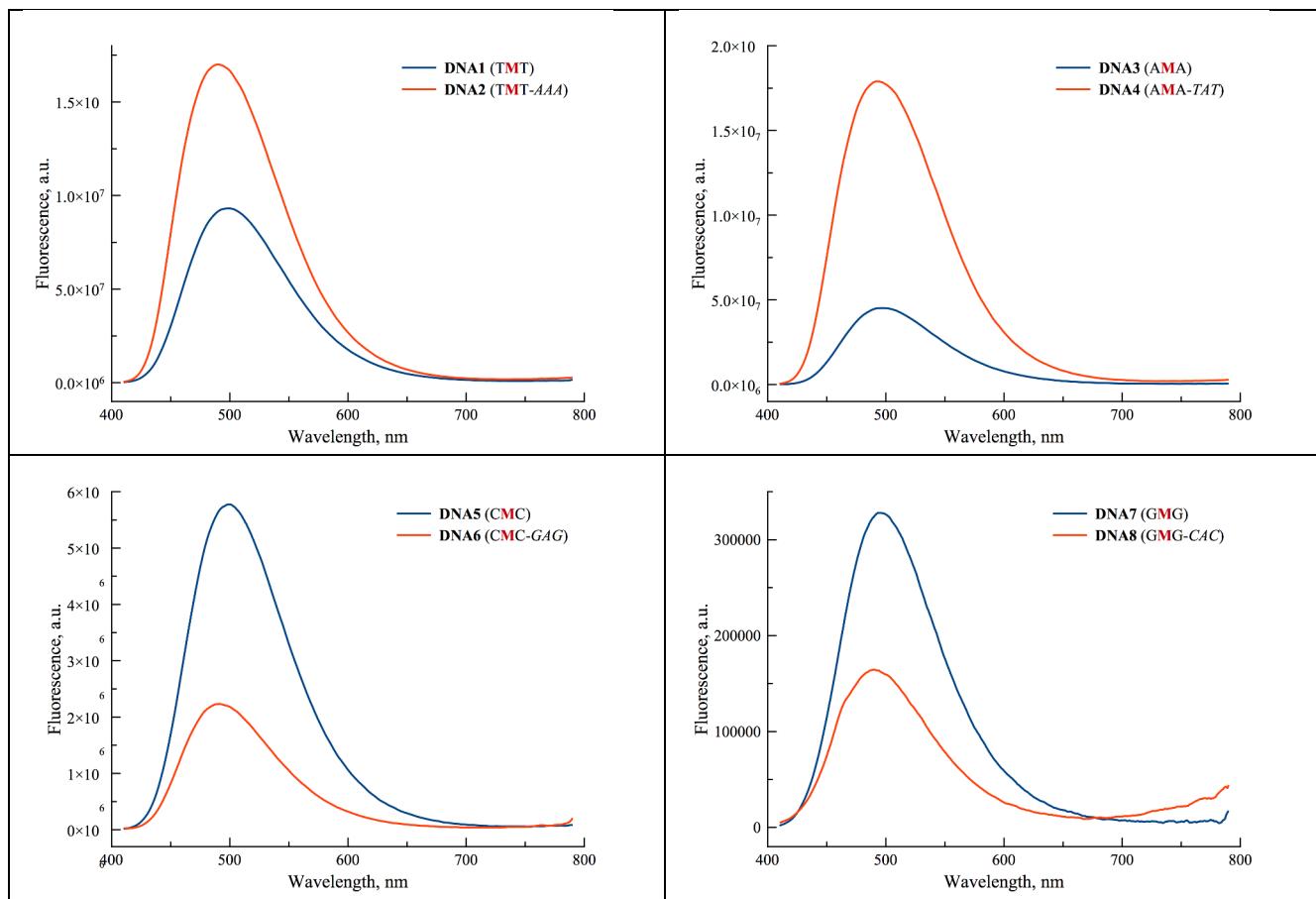


Fig S13. Fluorescence titration of DNA1 with DNA9 (normalized by the donor emission band on the right)

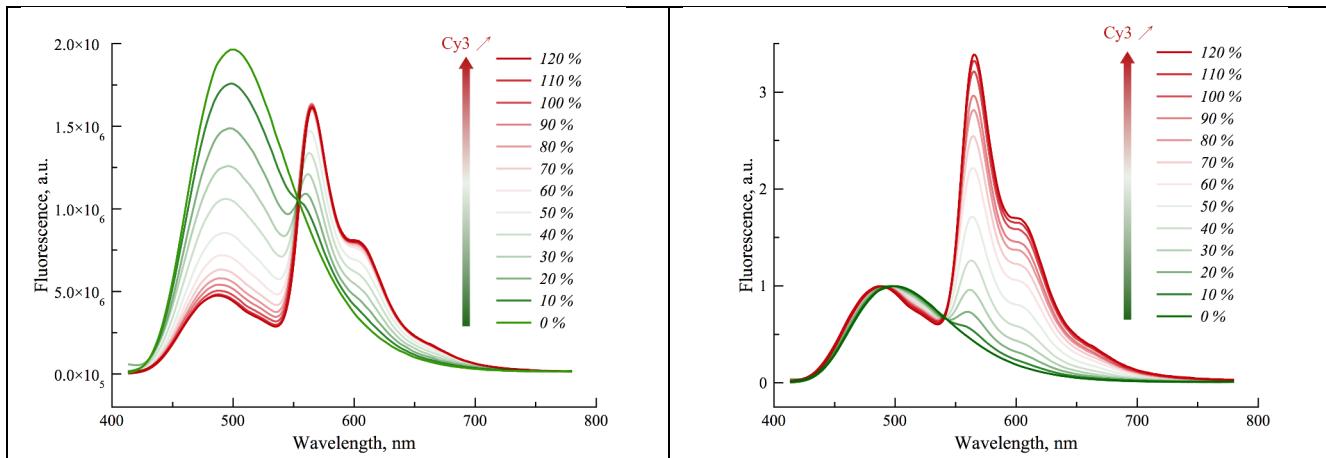


Fig S14. Excitation spectra of DNA10 and DNA11 collected at 490 nm (left) and at 564 nm (right)

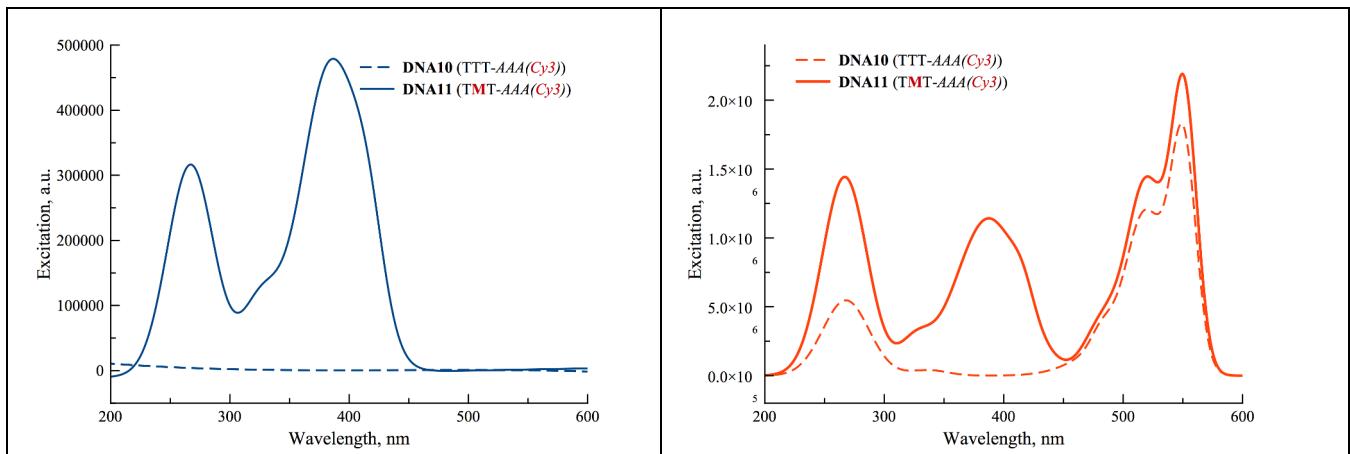
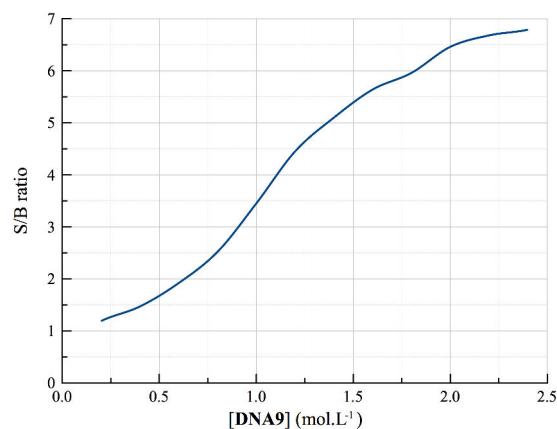
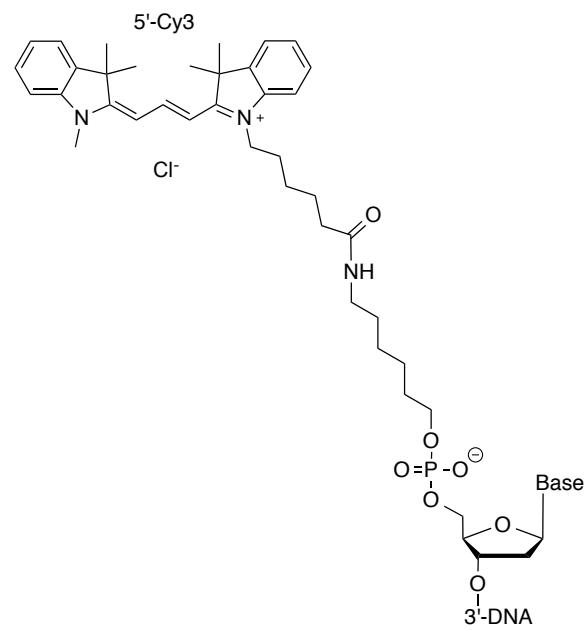


Fig S15. S/N ratio vs. [DNA9]^a



a) Plot of S/N ratio [$(I_{490i}/I_{564i})/(I_{490f}/I_{564f})$] vs. the concentration of DNA9, « i » and « f » indicate respectively without or with a stoichiometric addition of DNA9.

Fig S16. Structure of the 5'-end terminal Cy3 flexible linker used in this study



SI PUBLICATION 5

Supporting Information To:

Dual emissive analogue of deoxyuridine as a sensitive hydration-reporting probe for discriminating mismatched from matched DNAs and DNA/DNA and DNA/RNA duplexes

Nicolas P. F. Barthes,^{a†} Krishna Gavvala,^{b‡} Dmytro Dziuba,^a Dominique Bonhomme,^a Iuliia A. Karpenko,^a Anne Sophie Dabert-Gay,^c Delphine Debayle,^c Alexander P. Demchenko,^d Rachid Benhida,^a Benoît Y. Michel,^a Yves Mély,^{*b} Alain Burger^{*a}

a) Institut de Chimie de Nice, UMR 7272, Université de Nice Sophia Antipolis, CNRS, Parc Valrose, 06108 Nice Cedex 2, France.

b) Laboratoire de Biophotonique et Pharmacologie, UMR 7213, Faculté de Pharmacie, Université de Strasbourg, CNRS, 74 Route du Rhin, 67401 Illkirch, France

c) Institut de Pharmacologie Moléculaire et Cellulaire, UMR 6097, Université de Nice Sophia Antipolis, 660 Route des Lu-cioles, 06560 Valbonne, France.

d) A.V. Palladin Institute of Biochemistry, 9 Leontovicha Street, Kiev 01030, Ukraine.

Corresponding authors: burger@unice.fr; yves.mely@unistras.fr

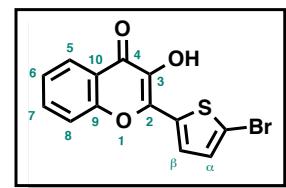
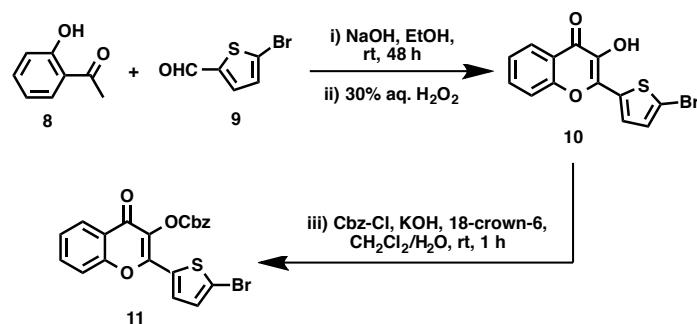
Experimental section

1. Synthetic Procedures

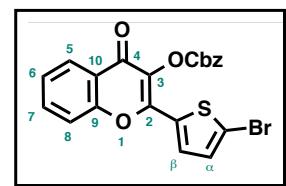
General methods:

All reactions involving water-sensitive reagents were performed in oven-dried glassware under argon using dry solvents. The synthetic intermediates were co-evaporated twice with toluene beforehand and dried in *vacuo* before use. All chemical reagents were obtained from commercial sources and were used as supplied. Anhydrous solvents were obtained according to standard procedures.[1] The reactions were monitored by thin-layer chromatography (TLC, Merck silica gel 60 F254 plates) and visualized both by UV radiation (254 & 360 nm) and by spraying with phosphomolybdic acid in ethanol followed by a subsequent warming with a heat gun. Column chromatography[2] was performed with flash silica gel (40–63 mm). All NMR spectra (^1H , ^{13}C , ^2D) were recorded on 200 or 500 Bruker Advance Spectrometers (200 or 500 MHz). ^1H NMR (200 and 500 MHz), ^{13}C NMR (50 and 125 MHz, recorded with complete proton decoupling) spectra were obtained with samples dissolved in CDCl_3 , CD_2Cl_2 , CD_3OD , $\text{DMSO}-d_6$, acetone- d_6 , CD_3CN with the solvent signals used as internal references: 7.26 ppm for CHCl_3 , 5.32 ppm for CDHCl_2 , 3.31 ppm for CD_2HOD , 2.50 ppm for $(\text{CD}_3)(\text{CD}_2\text{H})\text{S}(\text{O})$, 2.05 ppm for $(\text{CD}_3)(\text{CD}_2\text{H})\text{C}(\text{O})$, 1.94 ppm for CD_2HCN concerning ^1H NMR experiments, and 77.0 ppm for CDCl_3 , 53.8 ppm for CD_2Cl_2 , 49.0 ppm for CD_3OD , 39.4 ppm for $(\text{CD}_3)_2\text{S}(\text{O})$, 30.8 ppm for $(\text{CD}_3)_2\text{C}(\text{O})$ concerning ^{13}C NMR experiments.[3] Chemical shifts (δ) are given in ppm to the nearest 0.01 (1H) or 0.1 ppm (^{13}C). The coupling constants (J) are given in Hertz (Hz). The signals are reported as follows: (s=singlet, d=doublet, t=triplet, m=multiplet, br=broad). Assignments of ^1H and ^{13}C NMR signals were achieved with the help of D/H exchange, COSY, DEPT, APT, HMQC, HSQC, HMBC experiments. Regular mass spectra (MS) were recorded on an Esquire 3000 Plus apparatus with ESI in both positive and negative mode. High-resolution mass spectrometry was conducted with a FINIGAN MAT 95 spectrometer with EI or ESI ionization techniques. Supplementary data associated with this article: the experimental protocols for the synthesis of intermediates **3–7** and **10–11**, the ^1H NMR, ^{13}C NMR and (in part) ^1H - ^{13}C COSY, ^1H - ^{13}C HMQC and ^1H - ^{13}C HMBC spectra of all compounds are described herein. Systematic nomenclatures are used below for the assignments of the flavone and nucleoside.

1.1 Preparation of the 3HC fluorophores

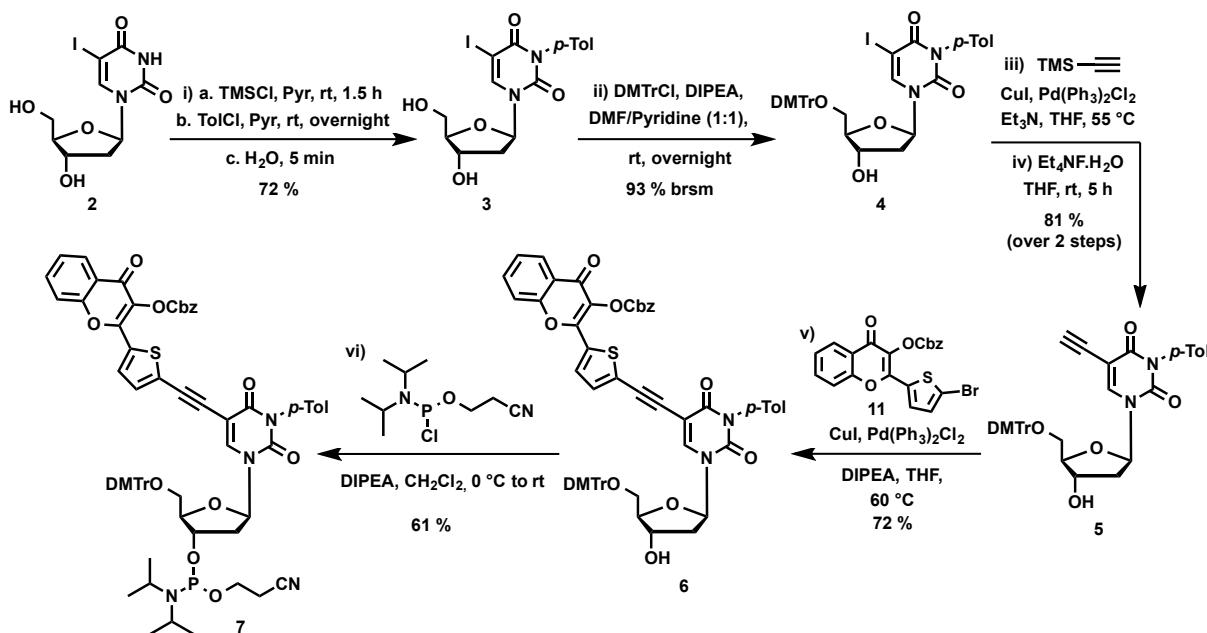


2-(5-Bromothien-2-yl)-3-hydroxy-chromen-4-one (10): To a stirred solution of *o*-hydroxyacetophenone **8** (1.16 mL, 9.65 mmol) and 5-bromo-2-thiophenecarbaldehyde **9** (1.15 mL, 9.65 mmol) in ethanol (50 mL) was dropwise added a 5 N NaOH solution (6 mL). The reaction mixture was stirred 48 h at rt before a dropwise addition of 30 % aq. hydrogen peroxide solution (2 mL). The resulting mixture was stirred 5 h at rt, and then poured into cold water (300 mL) and acidified with acetic acid to pH 4. The resulting precipitate was filtered and thoroughly washed with water and cyclohexane to provide as a yellow solid the chromone **10** (1.79 g, 57 %). $C_{11}\text{H}_7\text{BrO}_3\text{S}$ (323.16). $R_f=0.39$ (cyclohexane/EtOAc=7:3); ^1H -NMR (CDCl_3 , 200 MHz): δ = 8.22 (1H, dd, $^3J=8.0$ Hz, $^4J=1.4$ Hz, H5), 7.73 (1H, d, $^3J=4.0$ Hz, H β), 7.71 (1H, ddd, $^3J=8.4$ Hz, $^3J=8.0$ Hz, $^4J=1.4$ Hz, H7), 7.54 (1H, d, $^3J=8.4$ Hz, H8), 7.41 (1H, td, $^3J=8.0$ Hz, $^4J=0.8$ Hz, H6), 7.19 (1H, d, $^3J=4.0$ Hz, H α); ^{13}C -NMR (DMSO- d_6 , 50 MHz): δ = 176.0 (C4), 153.3 (C9), 146.5 (C2), 141.6 (C3), 136.1 (C2-*Th*), 131.6 (C7), 129.3 (Co), 124.7 (C5), 122.9 (C6), 122.0 (C10), 121.6 (C β), 117.8 (C8), 112.2 (*Th*-Br); MS (ESI $^+$, MeOH): m/z : 345.0, 347.0 [M+Na] $^+$.

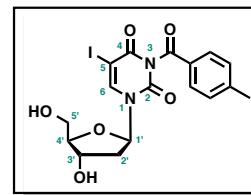


3-Benzoyloxycarbonyloxy-2-(5-bromothien-2-yl)-chromen-4-one (11): To a stirred suspension of **10** (1.05 g, 3.25 mmol) in CH_2Cl_2 (13 mL) were added a 4.5 M aq. KOH solution (9 mL), 18-crown-6 (5 mol %, 43 mg) and benzyl chloroformate (700 μL , 4.88 mmol). The reaction mixture became homogenous in 5 min and was stirred at rt for 1 h. After quenching by addition of H_2O (10 mL), the organic layer was extracted with CH_2Cl_2 (3x), dried over MgSO_4 , filtered and the volatiles were removed *in vacuo*. The residue was purified by flash chromatography on silica gel eluted with cyclohexane/EtOAc mixture (9:1 \rightarrow 7:3, v/v) to provide the desired compound **11** as white crystals (1.20 g, 81 %). $C_{21}\text{H}_{13}\text{BrO}_5\text{S}$ (457.29). $R_f=0.53$ (cyclohexane/EtOAc=7:3); ^1H -NMR (CDCl_3 , 200 MHz): δ = 8.25 (1H, dd, $^3J=8.0$ Hz, $^4J=1.6$ Hz, H5), 7.72 (1H, td, $^3J=8.4$ Hz, $^4J=1.6$ Hz, H7), 7.65 (1H, d, $^3J=4.0$ Hz, H β), 7.52 (1H, dd, $^3J=8.4$ Hz, $^4J=0.6$ Hz, H8), 7.34–7.50 (6H, m, H6 & Cbz), 7.16 (1H, d, $^3J=4.0$ Hz, H α), 5.37 (2H, s, CH_2); ^{13}C -NMR (CDCl_3 , 50 MHz): δ = 171.3 (C4), 155.0 (C9), 151.8 (OC(O)CH₂Ph), 150.3 (C2), 134.5 (i-C-Ph), 134.1 (C7), 132.0 (C3), 131.5 (C2-*Th*), 131.0 (C β), 131.0 (Co), 128.8 (*p*-C-Ph), 128.7 (*m*-C-Ph), 128.4 (*o*-C-Ph), 126.1 (C5), 125.4 (C6), 123.6 (C10), 120.4 (*Th*-Br), 117.8 (C8), 71.3 (OC(O)CH₂Ph); MS (ESI $^+$, MeOH/ CH_2Cl_2): m/z : 479.5, 481.5 [M+Na] $^+$, 495.4, 497.4 [M+K] $^+$. HRMS (ESI $^+$): m/z calcd for $\text{C}_{21}\text{H}_{14}\text{BrO}_5\text{S}$: 456.9740 [M+H] $^+$; found 456.9737.

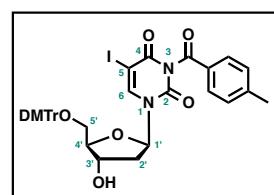
1.2 Preparation of the amidite 7



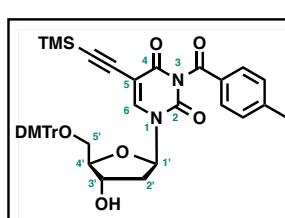
5-Iodo-3-N-(4-methylbenzoyl)-2'-deoxyuridine (3): To a stirred solution of 5-iodo-2'-deoxyuridine **2** (4.00 g, 11.3 mmol), in dry pyridine (14 mL), previously cooled down to 0 °C, was dropwise added Me₃SiCl (5.9 mL, 45.2 mmol, 4 eq). The reaction mixture was stirred at rt for 1.5 h. After completion by monitoring with TLC, the reaction mixture was cooled down to 0 °C and Et₃N (7.6 mL, 54.2 mmol, 4.8 eq) and *p*-toluoyl chloride (3.6 mL, 27.1 mmol, 2.4 eq) were sequentially added. The reaction mixture was stirred at rt overnight. H₂O was added and the mixture was stirred for 15 min. The organic phase was extracted with CH₂Cl₂ (2x) and EA (2x), dried over MgSO₄, filtered and reduced *in vacuo*. The residue was purified by flash chromatography on silica gel eluted with toluene/acetone (9:1 → 3:2, v/v) to provide the desired compound **3** as a white foam (3.85 g, 72 %). C₁₇H₁₇IN₂O₆ (472.2). *R*_f=0.55 (toluene/acetone=1:1). ¹H-NMR (acetone-*d*⁶, 200 MHz): δ = 8.80 (1H, s, H6), 7.94 (2H, d, ³J=8.2 Hz, *o*-H-Tol), 7.39 (2H, d, ³J=8.2 Hz, *m*-H-Tol), 6.24 (1H, t, ³J=6.3 Hz, H1'), 4.55 (1H, m, H3'), 4.01 (1H, m, H4'), 3.87 (2H, m, H5'), 2.44 (3H, s, *p*-Me-Tol), 2.41–2.34 (2H, m, H2'); ¹³C-NMR (acetone-*d*⁶, 50 MHz): δ = 168.9 (C(O)-Tol), 159.8 (C4), 149.9 (C2), 147.4 (*p*-C-Tol), 146.7 (C6), 131.4 (*o*-C-Tol), 130.7 (*m*-C-Tol), 129.7 (*i*-C-Tol), 88.9 (C4'), 86.8 (C1'), 71.4 (C3'), 67.3 (C5), 62.0 (C5'), 41.8 (C2'), 21.7 (*p*-Me-Tol); HRMS (ESI⁺): *m/z* calcd for C₁₇H₁₇IN₂NaO₆: 495.0029 [M+Na]⁺; found 495.0028.



5'-O-(4,4'-Dimethoxytrityl)-5-iodo-3-N-(4-methylbenzoyl)-2'-deoxyuridine (4): To a stirred solution of **3** (3.85 g, 8.2 mmol), in a mixture of dry DMF (15 mL) and pyridine (15 mL) previously cooled down to 0 °C were added sequentially DIPEA (3.7 mL, 40.8 mmol, 5 eq) and DMTrCl (8.29 g, 24.5 mmol, 3 eq). The reaction mixture was stirred at rt overnight. The volatiles were evaporated. The residue was purified by flash chromatography on silica gel eluted with toluene/acetone (93:7 → 82:18, v/v) to provide the recovery of starting material **3** as a white foam (2.01 g, 52 %) as well as the desired derivative **4** as a beige foam (3.02 g, 93 % brsm). C₃₈H₃₅IN₂O₈ (774.6). *R*_f=0.66 (toluene/acetone=7:3). ¹H-NMR (CD₂Cl₂, 200 MHz): δ = 8.22 (1H, s, H6), 7.79 (2H, d, ³J=7.3 Hz, *o*-H-Tol), 7.45 (2H, d, ³J=7.3 Hz, *m*-H-Tol), 7.36 (4H, d, ³J=8.6 Hz, *o*-H-PhOMe), 7.33–7.23 (5H, m, H-Ph), 6.87 (4H, d, ³J=8.6 Hz, *m*-H-PhOMe), 6.25 (1H, d, ³J=7.6, 5.9 Hz, H1'), 4.57 (1H, m, H3'), 4.07 (1H, dd, ³J=5.9, 3.1 Hz, H4'), 3.78 (6H, s, MeO), 3.37 (2H, d, ³J=3.2 Hz, H5'), 2.55–2.35 (5H, m, *p*-Me-Tol, H2'); HRMS (ESI⁺): *m/z* calcd for C₃₈H₃₅IN₂NaO₈: 797.1336 [M+Na]⁺; found 797.1311.

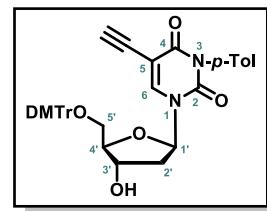


5'-O-(4,4'-Dimethoxytrityl)-5-trimethylsilylethynyl-3-N-(4-methylbenzoyl)-2'-deoxyuridine (4'): To a stirred solution of **4 (870 mg, 1 mmol) in dry THF (11 mL) under argon, were sequentially added TMS-acetylene (216 μL, 1.5 mmol, 1.5 eq), triethylamine (710 μL, 5 mmol, 5 eq) and CuI (8 mol %, 16 mg)/PdCl₂(PPh₃)₂ (8 mol %, 57 mg). The reaction mixture was warmed to 50 °C and stirred for 2 h. The mixture was filtered over a 545 Celite® (basic-washed) pad and the volatiles were removed *in vacuo*. The residue was purified by flash chromatography on silica gel eluted with toluene/EtOAc (93:7 → 55:45, v/v) to provide the desired compound **4'** as a beige foam (642 mg, 85 %). C₄₃H₄₄N₂O₈Si (744.9). *R*_f=0.53 (toluene/EtOAc=7:3). ¹H-NMR (CD₂Cl₂, 500 MHz): δ = 8.13 (1H, s, H6), 7.85 (2H, d, ³J=8.0 Hz, *o*-H-Tol), 7.52 (2H, d, ³J=7.8 Hz, *o*-H-Ph), 7.42 (5H, d, ³J=8.5 Hz, *o*-H-PhOMe), 7.39 (2H, t, ³J=7.8 Hz, *m*-H-Ph), 7.38 (2H, d, ³J=8.0 Hz, *m*-H-Tol), 7.29 (1H, t, ³J=7.8 Hz, *p*-H-Ph), 6.93 (4H, d, ³J=8.5 Hz, *m*-H-PhOMe), 6.27 (1H, dd, ³J=7.6 Hz, 5.8 Hz, H1'), 4.53 (1H, dd, ³J=6.0 Hz, 3.2 Hz, H3'), 4.12 (1H, dd, ³J=6.0 Hz, 3.2 Hz, H4'), 3.84 (6H, s, MeO), 3.43 (1H, dd, ²J=10.7 Hz, ³J=3.2 Hz, H5'_B), 3.38 (1H, dd, ²J=10.7 Hz, ³J=3.9 Hz, H5'_A), 2.52–2.49 (4H, m, H2'_A, *p*-Me-Tol), 2.31 (1H, m, H2'_B), 0.09 (9H, s, Si(CH₃)₃); ¹³C-NMR (CD₂Cl₂, 125 MHz): δ = 168.2 (C(O)-Tol), 160.9 (C4), 159.1 (*p*-C-PhOMe), 148.6 (C2), 147.5 (*p*-C-Tol), 145.0 (*i*-C-Ph), 142.8 (C6), 135.9 (*i*-C-PhOMe), 130.9 (*o*-C-Tol), 130.4 (*o*-C-PhOMe), 130.4 (*m*-C-Tol), 129.3 (*i*-C-Tol), 128.5 (*m*-C-Ph), 128.3 (*o*-C-Ph), 127.3 (*p*-C-Ph), 113.7 (*m*-C-PhOMe), 100.6 (C5), 100.2 (C≡CSi), 95.1 (C≡CSI).**

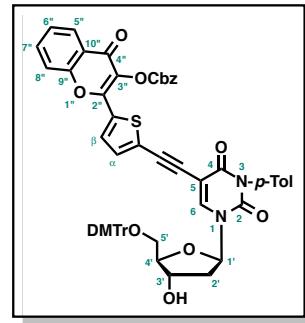


87.3 (C_{II} -O-5'), 87.0 (C4'), 86.5 (C1'), 72.6 (C3'), 63.8 (C5'), 55.6 (OMe), 41.8 (C2'), 22.0 (*p*-Me-Tol), 0.3 (Si(CH₃)₃); HRMS (ESI⁺): *m/z* calcd for C₄₃H₄₄N₂NaO₈Si: 767.2765 [M+Na]⁺; found 767.2742.

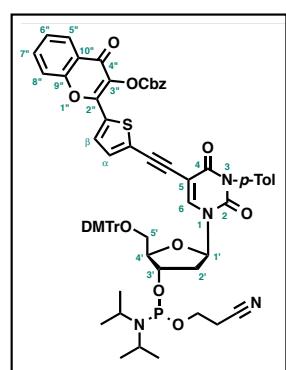
5'-O-(4,4'-Dimethoxytrityl)-3-N-(4-methylbenzoyl)-5-ethynyl-2'-deoxyuridine (5): To a stirred solution at 0 °C of the protected ethynyl derivative **4'** (0.725 mmol, 540 mg) in THF (7.5 mL) was portionwise added Et₄NFH₂O (4.35 mmol, 660 mg, 6 eq). The reaction mixture was stirred at rt for 1.5 h and then quenched with water. The organic layer was extracted with CH₂Cl₂ (3 x), dried over MgSO₄, filtered and reduced *in vacuo* to provide the desired compound **5** as a beige foam (468 mg, 96 %). The compound was sufficiently pure to be used in the next step without further purification. C₄₀H₃₆N₂O₈ (672.72). *R_f*=0.22 (toluene/Et₂O=7:2, v/v). ¹H-NMR (CD₂Cl₂, 200 MHz): δ = 8.21 (1H, s, H6), 7.81 (2H, d, ³J=8.1 Hz, *o*-H-Tol), 7.48 (2H, dd, ³J=8.2 Hz, ⁴J=1.4 Hz, *o*-H-Ph), 7.39 (4H, d, ³J=8.6 Hz, *o*-H-PhOMe), 7.33 (2H, d, ³J=8.1 Hz, *m*-H-Tol), 7.31–7.20 (3H, m, *m*-H-Ph, *p*-H-Ph), 6.89 (4H, d, ³J=8.6 Hz, *m*-H-PhOMe), 6.24 (1H, dd, ³J=7.2, 6.0 Hz, H1'), 4.62–4.52 (1H, m, H3'), 4.12–4.03 (1H, m, H4'), 3.80 (3H, s, OMe), 3.41 (1H, dd, ²J=10.8 Hz, ³J=3.4 Hz, H5'_B), 3.32 (1H, dd, ²J=10.8 Hz, ³J=3.0 Hz, H5'_A), 2.97 (1H, s, C≡CH), 2.55–2.44 (1H, m, H2'_B), 2.44 (3H, s, *p*-Me-Tol), 2.43–2.25 (1H, m, H2'_A); ¹³C-NMR (CD₂Cl₂, 50 MHz): δ = 168.1 (C(O)-Tol), 161.0 (C4), 159.2 (*p*-C-PhOMe), 148.6 (C2), 147.7 (*p*-C-Tol), 145.0 (*i*-C-Ph), 144.0 (C6), 136.0 (*i*-C-PhOMe), 135.7 (*i*-C-PhOMe), 130.9 (*o*-C-Tol), 130.4 (*o*-C-PhOMe), 130.4 (*m*-C-Tol), 128.8 (*i*-C-Tol), 128.5 (*m*-C-Ph), 128.3 (*o*-C-Ph), 127.3 (*p*-C-Ph), 113.7 (*m*-C-PhOMe), 99.3 (C5), 87.5 (C_{II} -O-5'), 87.1 (C4'), 86.6 (C1'), 82.4 (C≡CH), 74.4 (C≡CH), 72.4 (C3'), 63.8 (C5'), 55.6 (OMe), 41.9 (C2'), 22.0 (*p*-Me-Tol). MS (ESI⁺, MeOH) *m/z*: 695.5 [M+Na]⁺ 711.5 [M+K]⁺. HRMS (ESI⁺): *m/z* calcd for C₄₀H₃₆N₂NaO₈: 695.2364 [M+Na]⁺; found 695.2357.



5'-O-(4,4'-Dimethoxytrityl)-3-N-(4-methylbenzoyl)-5-(5-(3-benzylcarbonate-4-oxo-chromen-2-yl)thien-2-yl)ethynyl-2'-deoxyuridine (6): To a stirred solution of **5** (300 mg, 0.45 mmol, previously azeotropically coevaporated with dry toluene) and **11** (279 mg, 0.58 mmol, 1.3 eq) in THF (10 mL) under argon, were sequentially added triethylamine (311 μ L, 2.23 mmol, 5 eq), and CuI (7 mol %, 0.031 mmol, 6 mg.)/PdCl₂(PPh₃)₂ (7 mol %, 0.031 mmol, 22 mg) all together. The reaction mixture was warmed to 60 °C and stirred for 2 h. The volatiles were removed *in vacuo* and the residue was purified by flash chromatography on silica gel eluted with toluene/EtOAc (93:7 → 3:2, v/v) to provide the desired compound **6** as a yellow foam (339 mg, 72 %). C₆₁H₄₈N₂O₁₃S (1049.1). *R_f*=0.39 (toluene/EtOAc=7:3). ¹H-NMR (CDCl₃, 200 MHz): δ = 8.41 (1H, s, H6), 8.22 (1H, dd, ³J=8.0 Hz, ⁴J=1.6 Hz, H5''), 7.85 (2H, m, ³J=8.2 Hz, *o*-H-Tol), 7.71 (1H, ddd, ³J=8.3, 7.0 Hz, ⁴J=1.6 Hz, H7''), 7.62 (1H, d, ³J=4.1 Hz, H β), 7.51 (1H, d, ³J=8.3 Hz, H8''), 7.46–7.16 (17H, m, H-Ph-Cbz, H-Ph, *o*-H-PhOMe, H6'', *m*-H-Tol), 6.81 (4H, d, ³J=8.8 Hz, *m*-H-PhOMe), 6.74 (1H, d, ³J=4.1 Hz, H α), 6.34 (1H, dd, ³J=7.0 Hz, 6.0 Hz, H1'), 5.31 (2H, s, CH₂-Cbz), 4.67–4.58 (1H, m, H3'), 4.18–4.11 (1H, m, H4'), 3.72 (3H, s, OMe), 3.71 (3H, s, OMe), 3.49 (1H, dd, ²J=10.6 Hz, ³J=2.4 Hz, H5'_B), 3.37 (1H, dd, ²J=10.6 Hz, ³J=2.8 Hz, H5'_A), 2.64–2.39 (5H, m, H₂', *p*-Me-Tol); ¹³C-NMR (CDCl₃, 50 MHz): δ = 171.5 (C4''), 167.6 (C(O)-Tol), 160.2 (C4), 158.8 (*p*-C-PhOMe), 155.1 (C9''), 151.9 (C(O)-Cbz), 150.8 (C2''), 148.3 (C2), 147.0 (*p*-C-Tol), 144.4 (*i*-C-Ph), 142.8 (C6), 135.5 (*i*-C-PhOMe), 135.4 (*i*-C-PhOMe), 134.5 (C7''), 134.3 (*i*-C-Cbz), 133.4 (C β), 132.1 (C2''-Th), 131.8, 130.9 (*o*-C-Tol), 130.6 (C α), 130.1, 130.0, 129.2, 128.9, 128.8, 128.7, 128.4, 128.3, 128.2, 128.0, 127.3, 126.2 (C5''), 125.5 (C6''), 123.7 (C10''), 118.0 (C8''), 113.5 (*m*-C-PhOMe), 99.7 (C5), 87.6 (Th-C≡C-), 87.3 (C_{II} -O-5'), 87.1 (C4'), 86.6 (Th-C≡C-), 86.5 (C1'), 72.4 (C3'), 71.4 (CH₂-Cbz), 63.5 (C5'), 55.3 (MeO), 42.1 (C2'), 22.1 (*p*-Me-Tol); MS (ESI⁺, MeOH) *m/z*: 1070.7 [M+Na]⁺, 1086.8 [M+K]⁺. HRMS (ESI⁺): *m/z* calcd for C₆₁H₄₉N₂O₁₃S: 1049.2955 [M+H]⁺; found 1049.2920.



5'-O-(4,4'-Dimethoxytrityl)-3-N-(4-methylbenzoyl)-5-(5-(3-benzylcarbonate-4-oxo-chromen-2-yl)thien-2-yl)ethynyl-2'-deoxyuridine, 3'-(2-cyanoethyl)-N,N-diisopropyl-phosphoramidite (7): To a stirred solution of **6** (0.32 mmol, 339 mg, previously dried azeotropically by coevaporation with dry toluene) in DCM (2.3 mL) under argon and cooled down at 0 °C, were sequentially added DIPEA (1.1 mmol, 193 μ L, 3.4 eq) and 2-cyanoethyl-*N,N*-diisopropylchlorophosphoramidite (0.614 mmol, 138 μ L, 2 eq). The reaction mixture was stirred at rt for 2.5 h. The volatiles were removed *in vacuo* and the residue was purified by flash chromatography on silica gel eluted with toluene/EtOAc (95:5 → 60:40, v/v) to provide the desired compound **7** as a yellow foam (246 mg, 61 %). C₇₀H₆₅N₄O₁₄PS (1249.34). *R_f*=0.48 (toluene/EtOAc=8:2). ¹H-NMR (CD₃CN, 200 MHz): δ = 8.42 (1H, s, H6), 8.12 (1H, dd, ³J=8.0 Hz, ⁴J=1.7 Hz, H5''), 7.95 (2H, d, ³J=8.1 Hz, *o*-H-Tol), 7.82 (1H, ddd, ³J=8.8 Hz, ⁴J=1.7 Hz, H7''), 7.76 (1H, d, ³J=4.1 Hz, H β), 7.66 (1H, d, ³J=8.0 Hz, H8''), 7.53–7.32 (10H, m, H6'', H-Ph-Cbz, *o*-H-PhOMe), 7.29–7.14 (7H, m, H-Ph, *m*-H-Tol), 6.90–6.81 (5H, m, H α , *m*-H-PhOMe), 6.14 (1H, dd, ³J=13.0 Hz, 6.0 Hz, H1'H1'), 5.28 (2H, s, CH₂-Cbz), 4.86–4.68 (1H, m, H3'), 4.19 (1H, m, H4'), 3.82–3.73 (2H, m, CH₂-O), 3.68 (6H, s, MeO), 3.64–3.52 (2H, m, N-CH(CH₃)₂), 3.38 (2H, m, H5'), 2.69–2.52 (4H, m, H2', CH₂-CN), 2.45 (3H, s, *p*-Me-Tol), 1.19–1.06 (11H, m, N-CH(CH₃)₂); ³¹P-NMR (CD₃CN, 81 MHz): δ =148.1, 148.0; HRMS (ESI⁺): *m/z* calcd for C₇₀H₆₆N₄O₁₄PS: 1249.4034 [M+H]⁺; found 1149.4014.



2. ODNs synthesis, purification & physical characterization

2.1 ODNs synthesis and purification:

The ODN synthesis was performed on an Expedite 8900 DNA synthesizer (Applied Biosystem) using the “trityl off” mode and mild phosphoramidite chemistry on a 0.2 μ mol scale. Reagents and solvents, as well as dT, Ac-dC, Pac-dA, and dmF-dG phosphoramidites were purchased from Link Technologies. The standard DNA assembly protocol “DMT-off” was used except for the following modifications: 5-Ethylthio-1H-tetrazole (ETT) was used as activating agent; Pac-anhydride was used for capping; a longer coupling time (1200 s) was applied to the 3HC phosphoramidite. Non-labelled ODNs were purchased from Microsynth AG. The ODNs were cleaved from the solid support and deprotected with concentrated aqueous ammonia at room temperature for 12 h. The ODNs were analyzed (0.5 mL/min) and purified (2.5 mL/min) by RP HPLC (HPLC apparatus: WatersTM 600 Controller with WatersTM 996 Photodiode Array Detector and Jasco LC-Net II / ADC apparatus. Columns: analytical, 300 \times 4.60 mm, 5 μ m particle size, Clarity[®] 100Å, Phenomenex[®]; semi-preparative, Clarity[®] 5u Oligo-RP column 250 \times 10 mm Phenomenex[®]). The following gradient system was used: 100 % A –(30 min) \rightarrow 60 % A / 40 % B – (5 min) \rightarrow 100 % B –(5 min) \rightarrow 100 % A with A=Buffer pH 7.0 (1.9 L of MilliQ[®] water, 160 mL acetonitrile, 28 mL triethylamine, 12 mL of acetic acid) and B=0.2 CH₃CN:0.8 Buffer.

Fig S1. HPLC profile of TMT single strand ODN (254 nm left – 390 nm right):

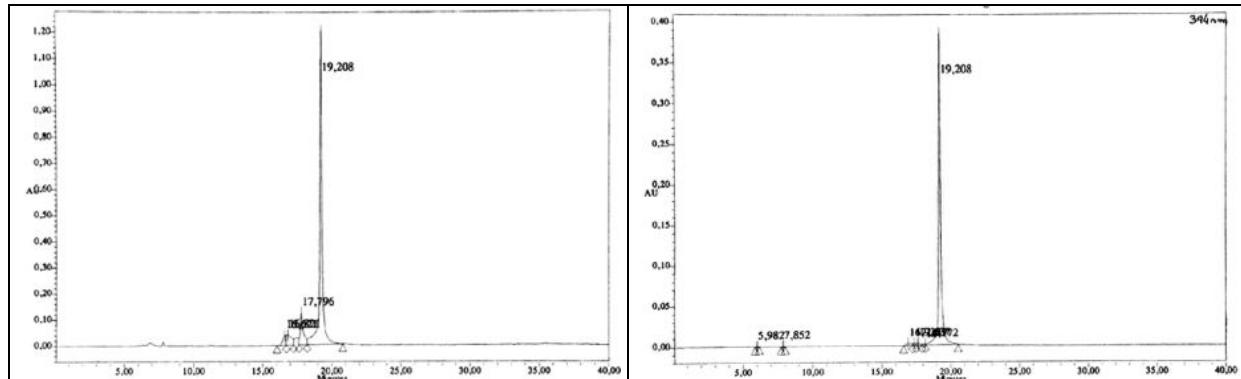


Fig S2. HPLC profile of AMA single strand ODN (254 nm left – 390 nm right):

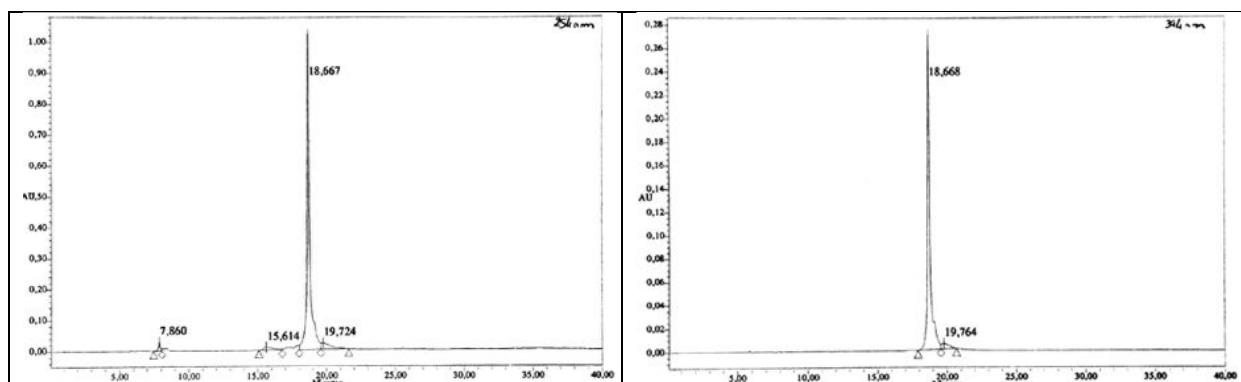


Fig S3. HPLC profile of CMC single strand ODN (254 nm left – 390 nm right):

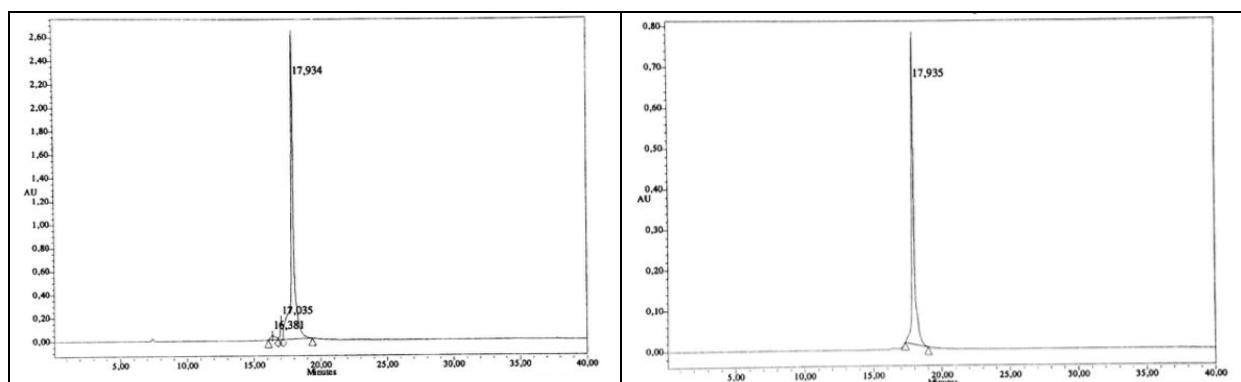
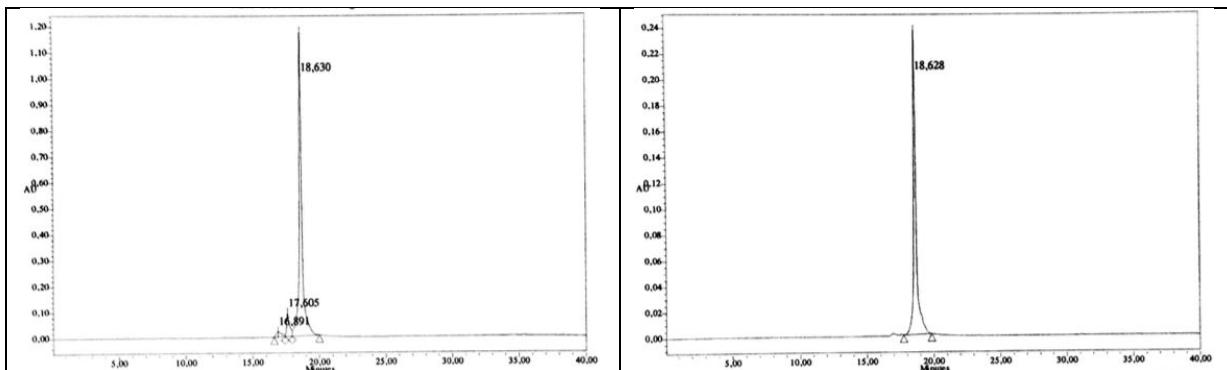


Fig S4. HPLC profile of GMG single strand ODN (254 nm left – 390 nm right):



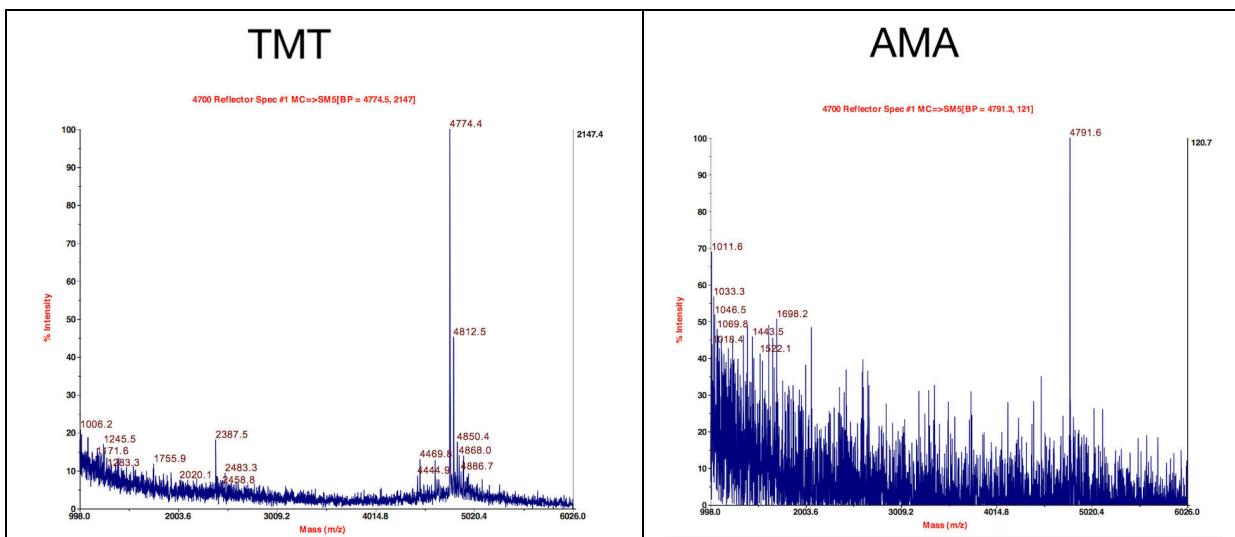
2.2 MALDI TOF/TOF analysis of ODNs:

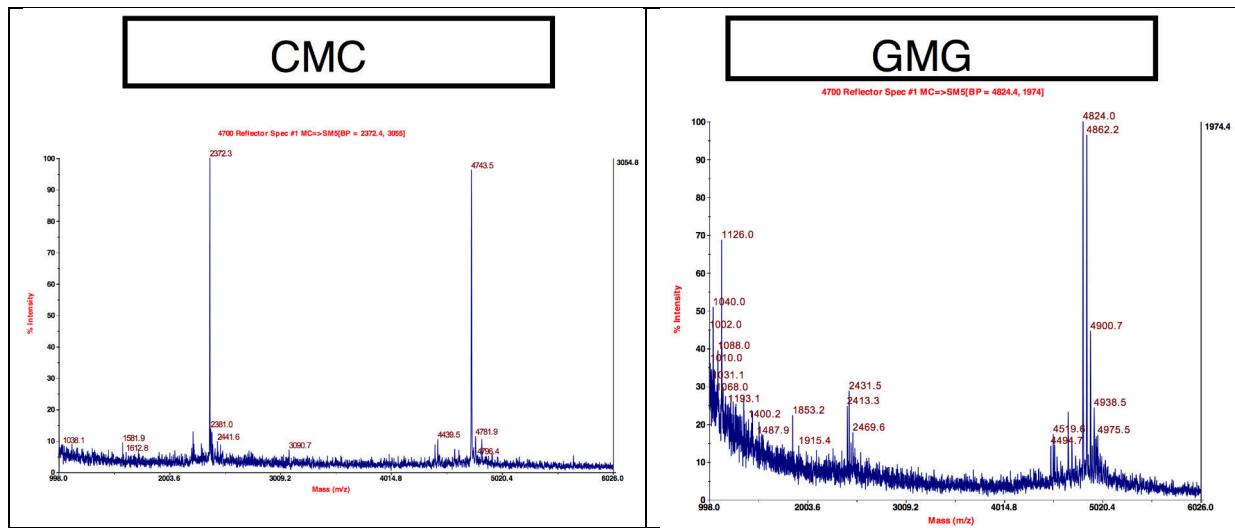
Dibasic Ammonium Citrate (DAC) (98% capillary GC) was obtained from Sigma. Acetonitrile of HPLC grade was purchased from VWR chemical. Ultrapure 3-Hydroxypicolinic Acid (3-HPA) MALDI matrix was purchase to Protea Biosciences. C4 pipette tips (Zip-Tip) were from Millipore.

The samples (500 pmol) were diluted to 10 μ L of water and were desalted with a C4 pipette Tips (Zip-tip). The Zip-tip was activated before use with 2 x 5 μ L of water: CH₃CN (50:50) and 2 x 5 μ L of DAC (50 mg/ml diluted in water). The 10 μ L of the ODN solution was loaded on Zip-tip by drawing and expelling ten times. Next the zip-tip was washed with 3 x 5 μ L of DAC (50 mg/mL) and 3 x 5 μ L of water. Elution was performed with 1.5 μ L of 3-HPA matrix (80 mg/mL, 50:50 CH₃CN:DAC) directly on MALDI plate. The ODN profile obtained in a ABSciex MALDI-TOF/TOF mass spectrometer in reflector mode with external calibration mixture (cal Mix 1+2 distributed by ABSciex). MALDI-TOF/TOF-MS analysis: MS spectra were recorded manually in a mass range of 500-6000 Da resulting from 400 laser shots of constant intensity fixed at 6200. Data were collected using 4000 series Explorer (AB SCIEX) experiments.

Table S1. Mass of the synthesized oligonucleotides.

ODN	Sequence	MALDI-TOF found (calcd) [M ⁺]
1	5'd-CGTTTTT M TTTTTGC-3'	4774.1 (4773.2)
2	5'd-CGTTTTA M TTTTGC-3'	4791.6 (4791.3)
3	5'd-CGTTTC M CCTTTGC-3'	4743.5 (4743.2)
4	5'd-CGTTTG M GTTTG-3'	4824.0 (4823.3)





3. Spectroscopic characterizations

3.1 Denaturation studies and melting temperatures

- ❖ ***Preparation of the samples:***

The ODNs were analyzed in triplicate cacodylate buffer pH 7.0 (10 mM cacodylate, 150 mM NaCl, 1 mM EDTA).

Preparation of the single strand solution: the solution of the sample was prepared by mixing 400 µL of a stock solution of 20 mM Cacodylate buffer solution pH 7.0, 80 µL of 1.5 M NaCl solution, 10 µL of 80 mM EDTA solution, 25 µL of 64 µM ssODN and 285 µL of MilliQ® water.

Preparation of the double strand solution: 400 µL of a stock solution of 20 mM Cacodylate buffer solution pH 7.0, 80 µL of 1.5 M NaCl solution, 10 µL of 80 mM EDTA solution, 25 µL of 64 µM ODN 1, 25 µL of 64 µM ODN 2 and 260 µL of MilliQ® water.

Melting curves were recorded by following the temperature-dependence of the absorbance changes at 260 nm of the sample (2 µM concentration of each strand). Absorption spectra were recorded in a Peltier thermostated cell holder on a Cary 4 spectrophotometer (Varian). The pathlength of cell was 1 cm. The temperature range for denaturation measurement was 5 – 80 °C. Speed of heating was 0.3 °C/min. The melting curves were converted into a plot of α versus temperature, where α represents the fraction of single-strands in the duplex state. The melting temperatures were extracted from these curves after differentiation as described elsewhere.[5]

Table S2. Melting temperatures of the duplexes.

Duplexes	T_m (°C)		
	M	Wild Type ^b	ΔT_m (°C) ^c
TMT·AAA	45.8	49.5 [50.2]	-3.7
TMT·ATA	37.9	40.5 [42.2]	-2.6
TMT·ACA	39.1	39.0 [42]	0.1
TMT·AGA	38.1	40.7 [42.9]	-2.6
TMT·AAbA	39.2	34.5	4.7
AMA·TAT	42.4	47.3 [47.4]	-4.9
AMA·TTT	37.6	39.8 [40.6]	-2.2
AMA·TCT	39.2	38.5 [39.3]	0.7
AMA·TGT	38.1	42.7 [42.8]	-4.6
AMA·TabT	36.3	30.5	5.8
CMC·GAG	48.1	51.3 [53.7]	-3.2
CMC·GTG	37.9	41.1 [46.7]	-3.2
CMC·GCG	42.6	42.1 [45.3]	0.5
CMC·GGG	42.5	46.5 [49.5]	-4.0
CMC·GAbG	43.2	37.2	6.0
GMG·CAC	52.4	54.8 [55.1]	-2.4

GMG·CTC	45.6	46.6 [47.5]	-1.0
GMG·CCC	47.1	43.5 [44.8]	3.6
GMG·CGC	47.3	49.7 [51.0]	-2.4
GMG·CAbC	44.5	37.5	7.0

Duplexes	T_m (°C)		
	M	Wild Type ^b	ΔT_m (°C) ^c
TMT·r-AAA	38.8	41.4	-2.6
AMA·r-UAU	37.1	40.9	-3.8

^a T_m of the corresponding duplex formed from unmodified ODNs and its theoretical values given in square brackets. ^b ΔT_m refers to the difference of T_m between the labelled and wild type ODNs.

Fig S5. Melting temperature curves of the wild type double strands:

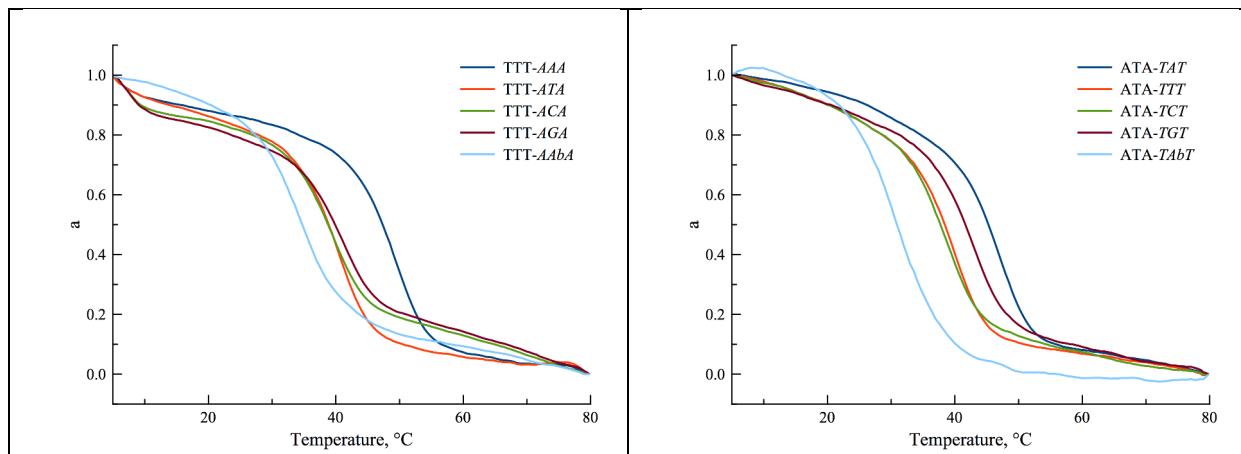
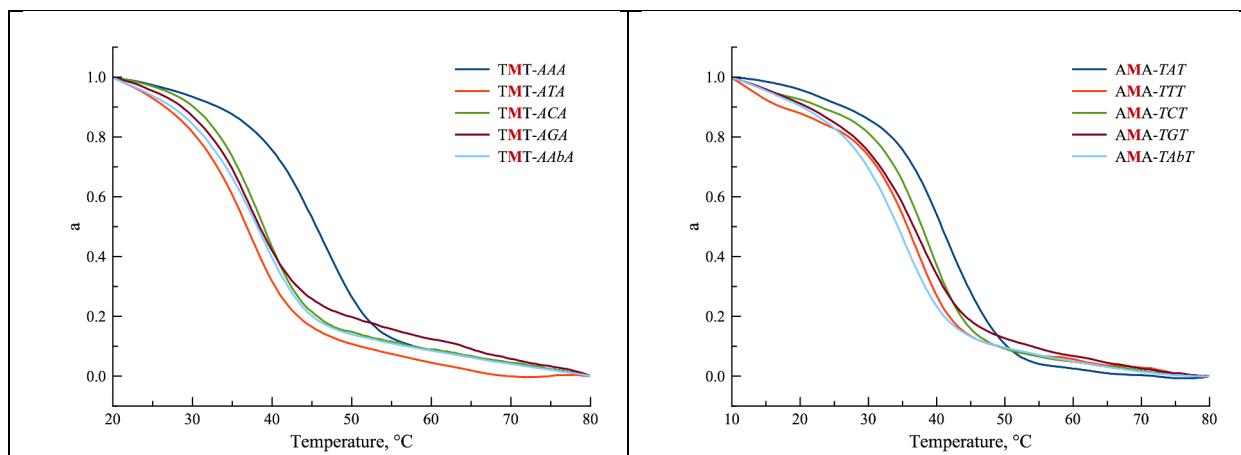
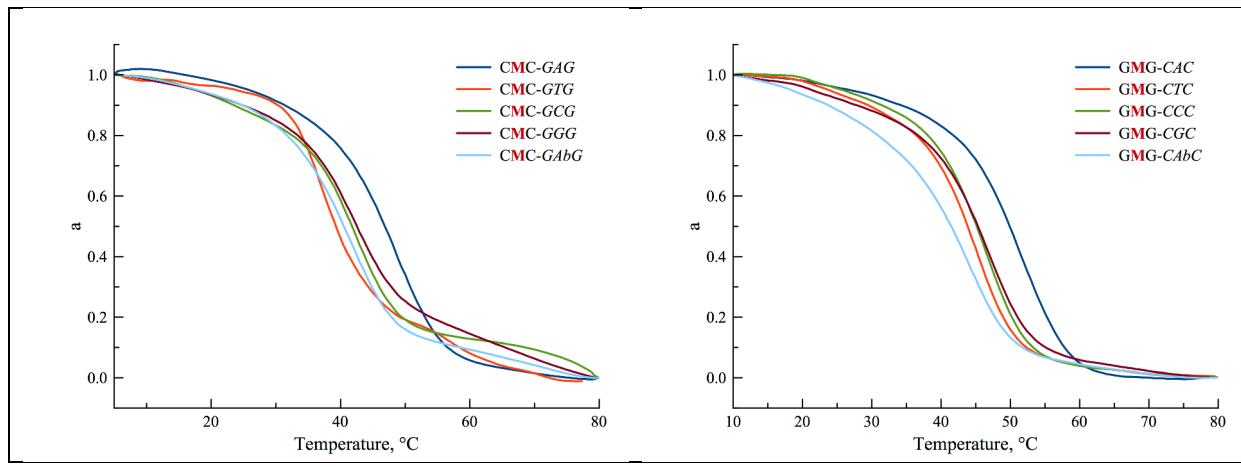


Fig S6. Melting temperature curves of labelled double strands:

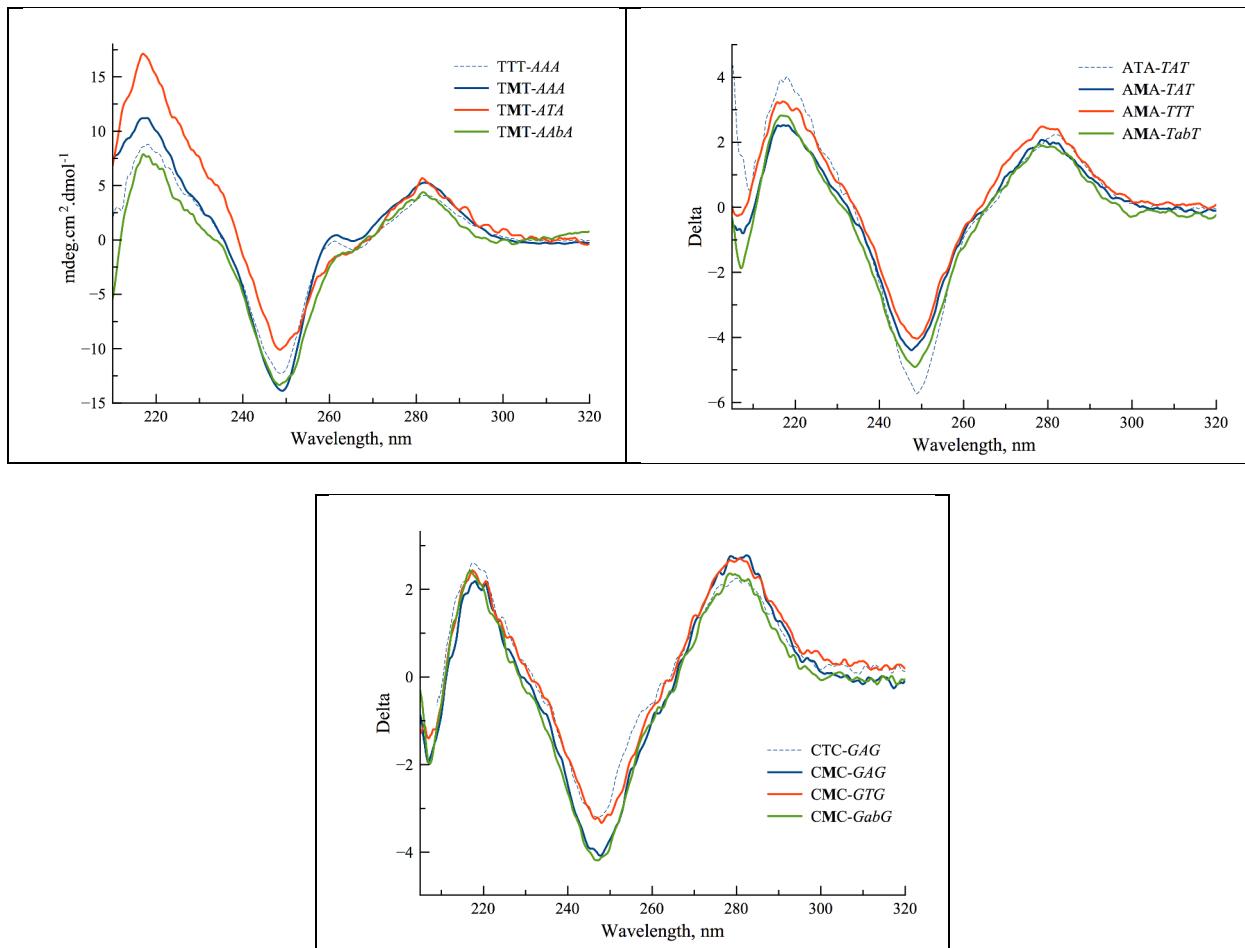




3.3. Circular dichroism

Circular dichroism spectra were recorded in duplicate with 2 μ M solution of the canonical dsDNA and labelled dsDNA (3HC (**M**) opposite **A**, **T** or **Ab**) in buffer pH 7.0 (20 mM Sodium phosphate, 100 mM NaCl, 1 mM EDTA) at 20 °C on a Jasco J-810 spectropolarimeter. Two maxima were observed in CD spectra: one negative at 249 nm and the other positive at 282 nm. Similar results were obtained by using an alternative buffer: 10 mM cacodylate buffer, 150 mM NaCl, 1 mM EDTA.

Fig S7. Representative CD spectra of wild type and labelled duplexes:



3.4. Absorption and Fluorescence spectra

The absorption and fluorescence experiments were realized in triplicate in pH 7.0 phosphate buffer (20 mM sodium phosphate, 100 mM NaCl, 1 mM EDTA). The absorption spectra were recorded on a Cary 300 Scan spectrophotometer (Varian) using 1cm quartz cells at 20 °C. The fluorescence spectra were recorded on a FluoroMax 4.0 spectrofluorometer (Jobin Yvon, Horiba) by using excitation and emission slits of 2 nm and were corrected at excitation and emission. They were taken with absorbance of about 0.05 at 20 °C at the excitation wavelength mentioned in the corresponding experiments.

The quantum yields were corrected according to the variation of the refractive index of the different solvents. Quantum yields were determined by using quinine sulfate (QS) in 0.1 M HCl solution ($\lambda_{ex} = 350$ nm, $\Phi = 0.54$) and *p*-dimethylaminoflavone (dMAF) in EtOH ($\lambda_{ex} = 404$ nm, $\Phi = 0.27$) as standard references.[6,7] Similar results were obtained by using an alternative buffer: 10 mM cacodylate buffer, 150 mM NaCl, 1 mM EDTA.

Fig S8. UV spectra of labelled single and double strands:

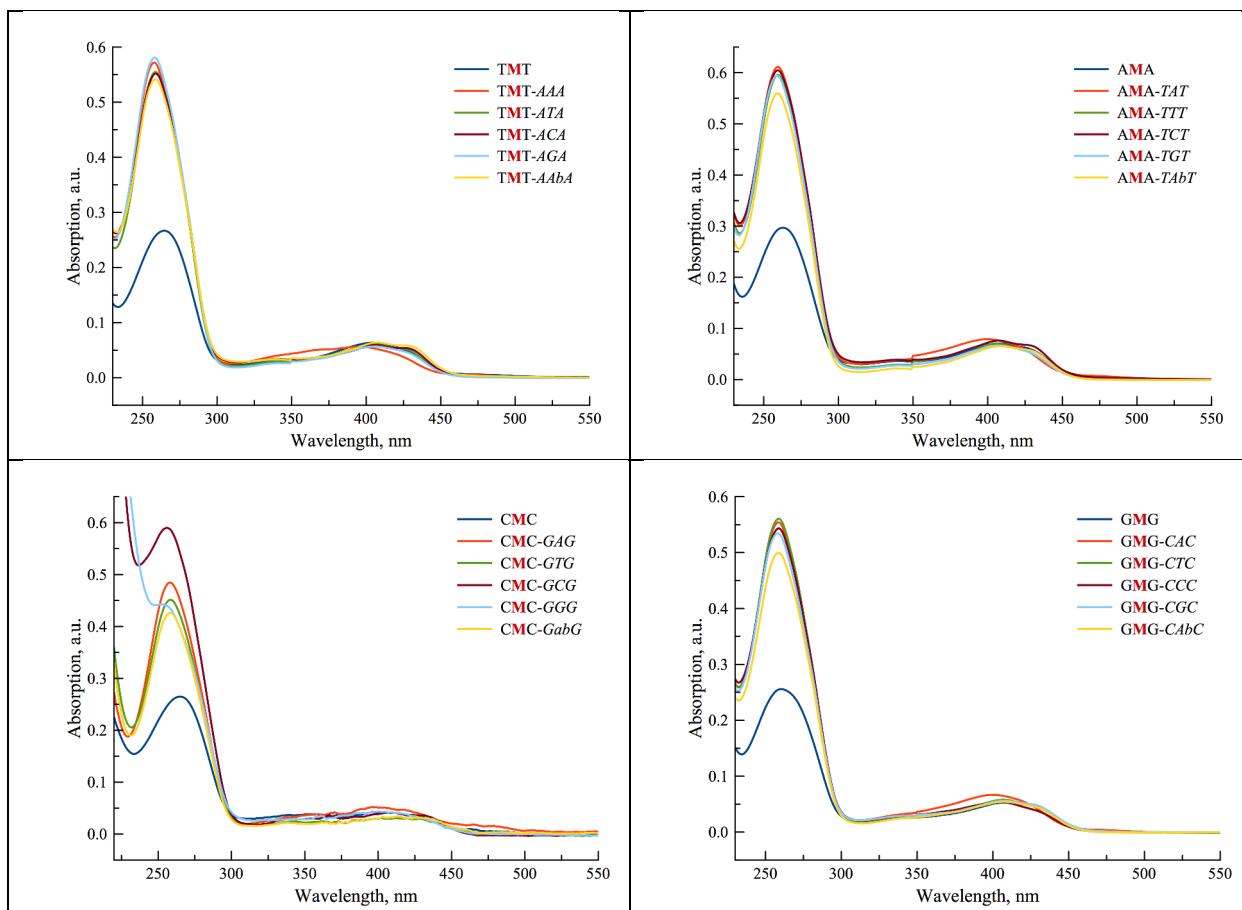


Fig S9. Fluorescence spectra of labelled single and double strands with TMT:

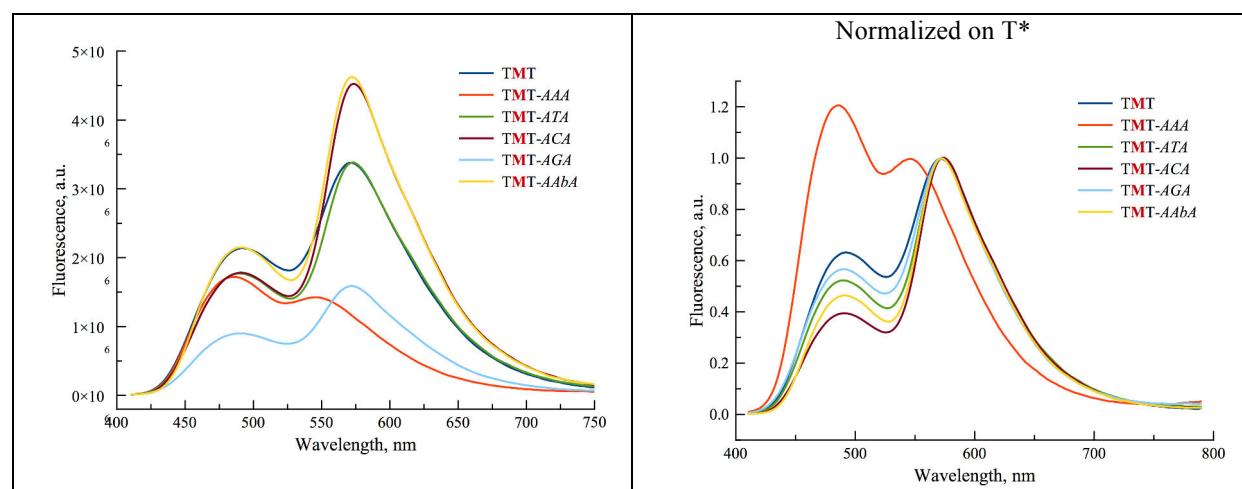


Fig S10. Fluorescence spectra of labelled single and double strands with AMA:

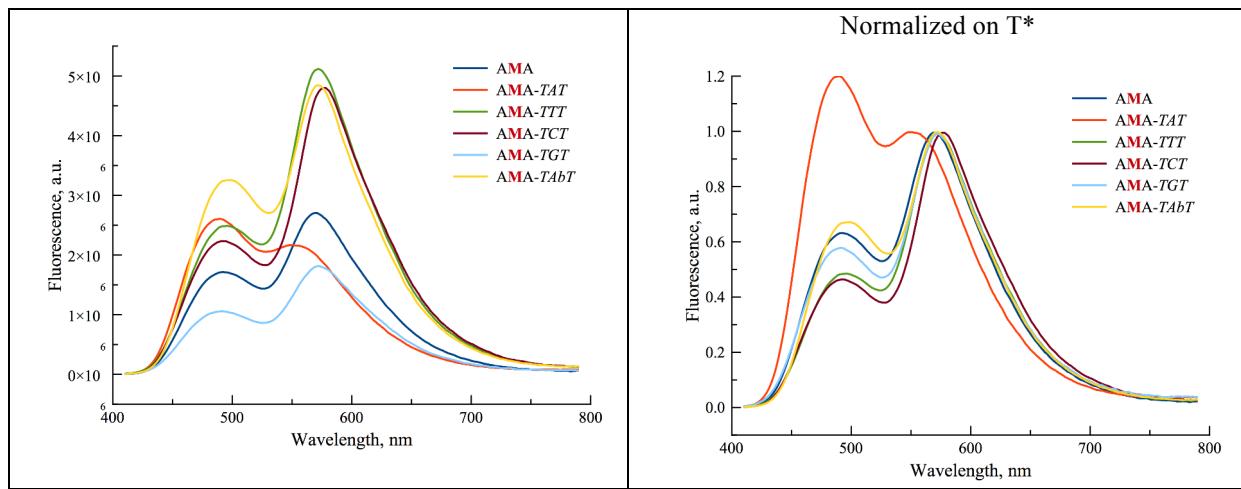


Fig S11. Fluorescence spectra of labelled single and double strands with CMC:

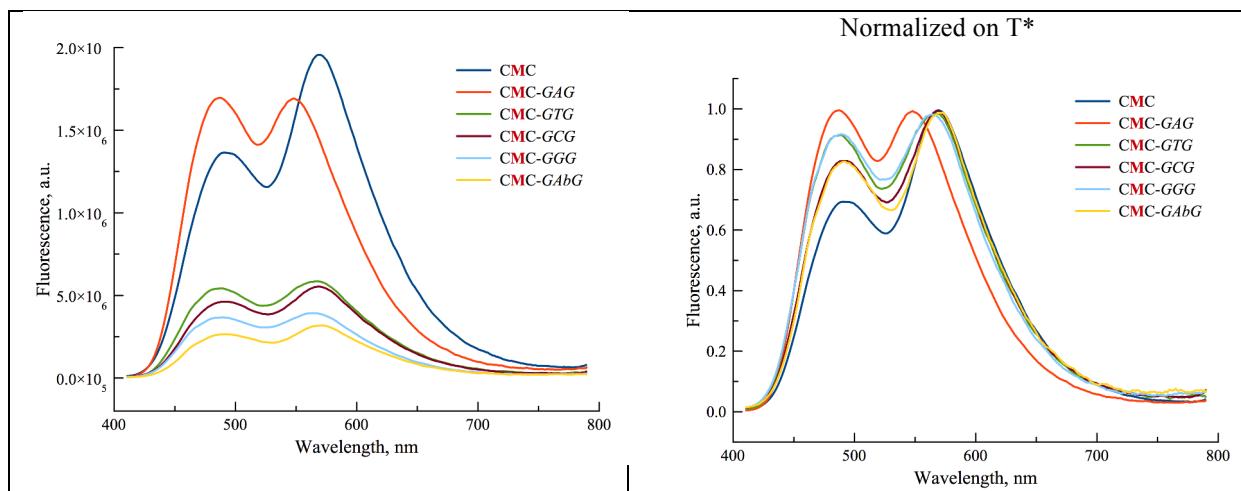


Fig S12. Fluorescence spectra of labelled single and double strands with GMG:

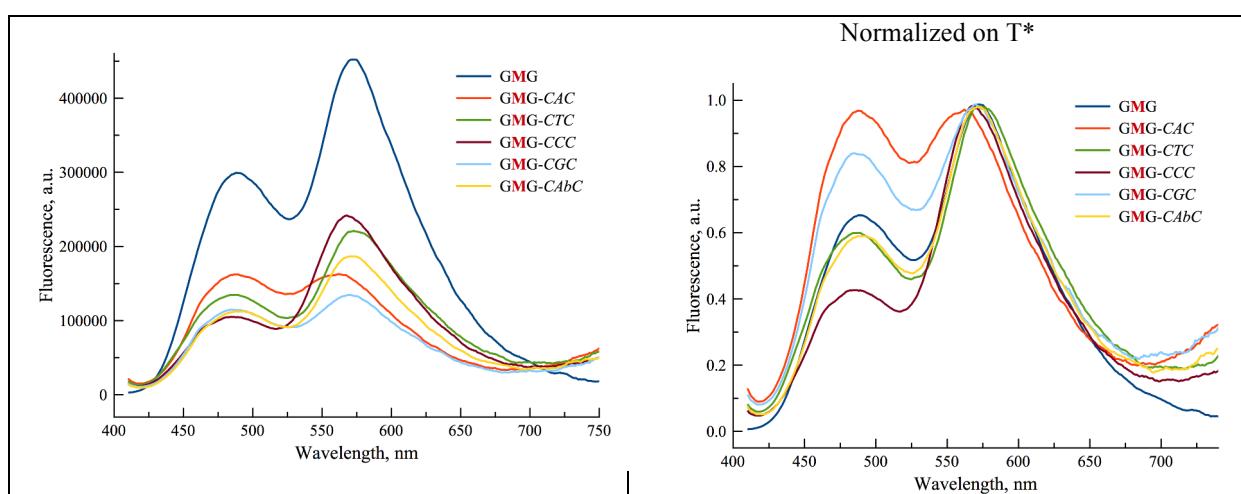
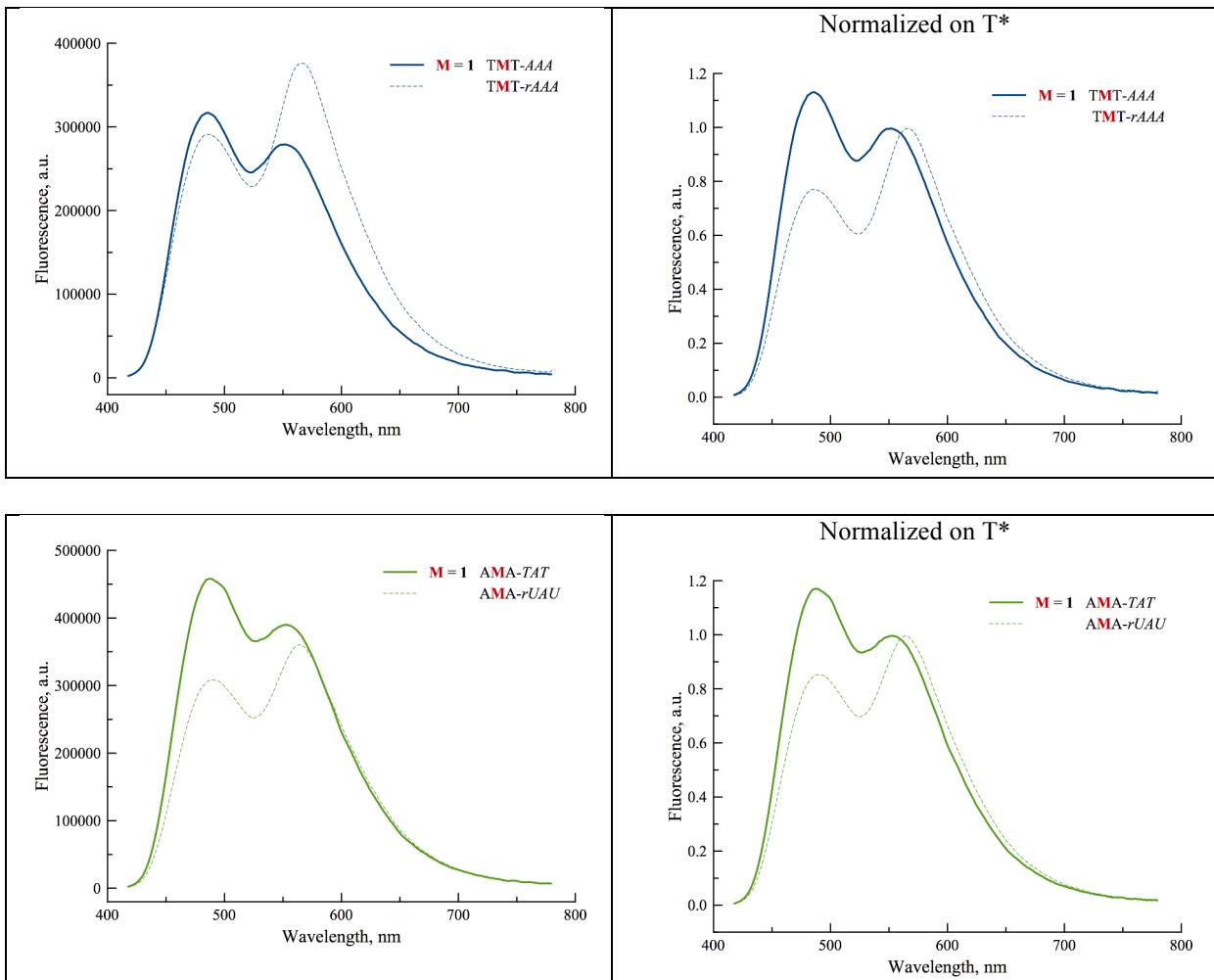


Fig S13. Fluorescence spectra of B DNA/DNA and A DNA/RNA:



References:

- [1] W. L. F. Armarego, C. L. L. Chai, *Purification of Laboratory Chemicals*, 7th ed.; Butterworth-Heinemann: Oxford, 2013.
- [2] W. C. Still, M. Kahn, A. Mitra, *J. Org. Chem.* **1978**, *43*, 2923-2925.
- [3] a) H. E. Gottlieb, V. Kotlyar, A. Nudelman, *J. Org. Chem.* **1997**, *62*, 7512-7515; b) G. R. Fulmer, A. Miller, N. H. Sherden, H. E. Gottlieb, *Organometallics* **2010**, *29*, 2176-2179.
- [4] N. P. F. Barthes, I. A. Karpenko, D. Dziuba, M. Spadafora, J. Auffret, A. P. Demchenko, Y. Mély, R. Benhida, B. Y. Michel, A. Burger, *RSC Adv.* **2015**, *5*, 33536-33545.
- [5] K. J. Breslauer, *Methods Enzymol.* **1995**, *259*, 221-242.
- [6] S. M. Ormson, R. G. Brown, F. Vollmer, W. Rettig, *J. Photochem. Photobiol. A* **1994**, *81*, 65-72.
- [7] W. H. Melhuish, *J. Phys. Chem.* **1961**, *65*, 229-235.



ANNEXES

ANNEXES

ANNEXES S1

Tableau 1. Traduction des codons en acides aminés.

		2ème Base									
		U		C		A		G			
1ère Base	U	UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys	U	
		UUC	Phe	UCC	Ser	UAC	Tyr	UGC	Cys	C	
		UUA	Leu	UCA	Ser	UAA	Stop	UGA	Stop / Sec / Trp	A	
		UUG	Leu / Start	UCG	Ser	UAG	Stop / Sec	UGG	Trp	G	
	C	CUU	Leu	CCU	Pro	CAU	His	CGU	Arg	U	
		CUC	Leu	CCC	Pro	CAC	His	CGC	Arg	C	
		CUA	Leu	CCA	Pro	CAA	Gln	CGA	Arg	A	
		CUG	Leu	CCG	Pro	CAG	Gln	CGG	Arg	G	3ème Base
	A	AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser	U	
		AUC	Ile	ACC	Thr	AAC	Asn	AGC	Ser	C	
		AUA	Ile	ACA	Thr	AAA	Lys	AGA	Arg	A	
		AUG	Met / Start	ACG	Thr	AAG	Lys	AGG	Arg	G	
	G	GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly	U	
		GUC	Val	GCC	Ala	GAC	Asp	GGC	Gly	C	
		GUA	Vla	GCA	Ala	GAA	Glu	GGA	Gly	A	
		GUG	Val / Start	GCG	Ala	GAG	Glu	GGG	Gly	G	

ANNEXES S2

ii. Influence de l'environnement sur le développement

L'épigénome est particulièrement sensible à la dérégulation au cours des différentes phases de notre évolution : gestation, développement néonatal, puberté et vieillesse.²¹⁴ En effet, il existe maintenant des preuves significatives que le risque de troubles ou maladies chroniques chez l'adulte résulte d'une exposition à des facteurs environnementaux lors des premières phases du développement. Il semble exister un lien étroit entre ce à quoi un individu est exposé *in utero* et les maladies qu'il développe à l'âge adulte.²¹⁵ Des modifications épigénétiques, comme la méthylation de l'ADN, semblent être impliquées dans ces phénomènes. Les perturbateurs endocriniens modifiant essentiellement la méthylation de l'ADN sont les exemples les plus frappants et les plus récents.²¹⁶

Notre régime alimentaire semble jouer un rôle particulièrement important sur notre épigénome. Par exemple, des femmes ayant été enceintes au moment de graves famines, ont donné naissance à des enfants présentant une prédisposition à développer plus tard un diabète de type 2 (hyperméthylation de IGF2 - *insulin-like growth factor*), des problèmes cardiovasculaires ou des troubles psychologiques (schizophrénie).²⁰

Des modifications liées à leur alimentation ont également été observées chez les souris *Agouti* présentées en Figure S1A. [*Agouti* est un terme qui permet de caractériser la couleur de la robe des mammifères. Cette nomenclature fait référence à un rongeur d'Amérique latine au pelage brun, l'*agouti*.] Le gène *Agouti* murin sauvage code pour un signal chimique permettant la production d'eumélanine noire ou de phéomélanine jaune. Une grande variation de couleur du pelage peut ainsi être obtenue, du jaune (gène agouti non méthylé) à l'*agouti* (gène agouti méthylé), suivant qu'il y ait une hyperméthylation des 2 allèles agouti, une monométhylation ou enfin une absence bi-allélique de méthylation du gène agouti. Chez ces animaux modèles, *Jirtle et al.* ont étudié l'influence, sur l'épigénome fœtal, de la consommation par la mère d'un œstrogène xénobiotique, le bisphénol A (BPA).²¹⁷ Ils ont observé que la prise de BPA 15 jours avant la période d'accouplement, puis tout au long de la gestation et de la lactation, augmentait le nombre de nouveaux-nés présentant une robe jaune et diminuait donc le nombre d'individus à la robe brune. Les mères ingérant du BPA présentaient un taux de méthylation de l'ADN moins important que les souris contrôles. L'exposition à ce composé chimique induit donc une hypométhylation de l'ADN et cette modification se répercute sur la coloration de la robe de leur progéniture.²¹⁴

D'autres études ont également été réalisées sur ces souris *Agouti*. La Figure S1B présente deux souris génétiquement identiques et de même âge. Pendant la grossesse, leur deux mères ont été nourries avec du bisphénol A mais la mère de la souris jaune (gauche) à suivi un régime alimentaire standard de souris alors que la mère de la souris agouti (droite) à reçu des compléments alimentaires (vitamine B12, choline, acide folique et bétaine : donneurs de méthyle). Le BPA a pour effet de déméthyler le gène agouti, donnant ainsi une couleur jaune à la robe. La consommation de ces agents donneurs de

ANNEXES

groupe méthyle permet d'atténuer, voire d'annuler, les effets du BPA ingéré par la mère. En effet, ces composés ont affecté l'activité de la DNMT en augmentant la quantité de SAM disponible pour la méthylation de l'ADN (ou en diminuant celle de l'inhibiteur SAH). Ceci a eu pour effet d'augmenter (de maintenir, ici) le taux de méthylation de l'ADN chez les embryons en développement.⁵²

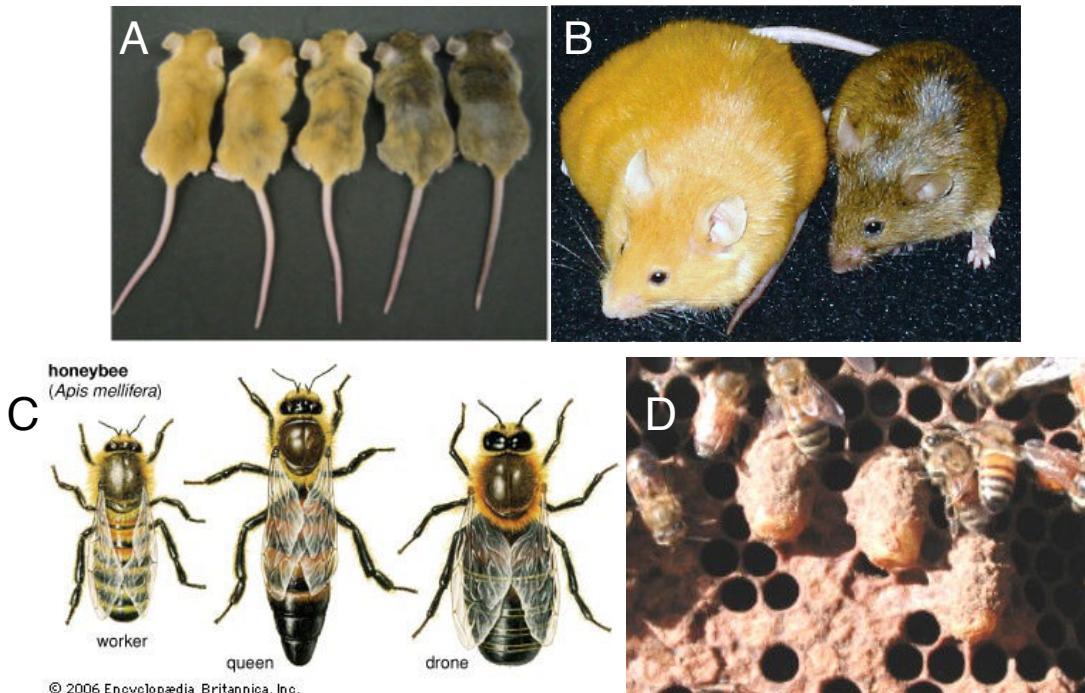


Figure S1. Influence de l'environnement.

Un autre exemple en lien avec la nutrition concerne les abeilles. Sur la Figure S1C sont représentés une ouvrière, une reine et un faux-bourdon (mâle). En fonction de sa nourriture, une même larve peut devenir soit une ouvrière soit une reine. En effet, si la larve est nourrie avec du miel, elle devient une ouvrière [cf. Figure S1D, *alvéoles à l'arrière plan*]. En revanche, si elle est nourrie avec de la gelée royale, c'est une reine qui se développe [cf. Figure S1D, *cellule royale au premier plan*]. Ceci est due à la présence d'un inhibiteur naturel de DNMT3 dans la gelée royale.²¹⁸

Mais la nourriture n'est pas le seul facteur pouvant influencer le développement des individus, citons par exemple la température. Chez tous les crocodiliens et un certain nombre de tortues et de lézards, la température au cours du mi-trimestre de l'embryogenèse détermine le sexe (mâle ou femelle) de l'individu.⁵⁷

Concernant le changement de sexe, les poissons possèdent un mécanisme tout aussi particulier. En effet, en fonction de la présence et du comportement de ses congénères un individu peut se transformer en un individu du sexe opposé.⁵⁷

Citons également comme autres facteurs l'exposition à des virus, au tabac, à l'alcool, à la pollution, à la prise d'antibiotiques, ...^{20, 214} Chez les jumeaux monozygotes humains, le profil de méthylation est identique à leur naissance. Cependant, au fur et à mesure qu'ils vieillissent, il évolue. Cette évolution est en réalité liée à l'accumulation de

ANNEXES

marques épigénétiques différentes, en lien avec les nombreux facteurs environnementaux auxquels chacun d'entre eux a été exposé.^{219, 42}

Ces modifications ne touchent pas que les animaux. Les plantes aussi peuvent subir les effets de l'environnement comme en attestent les Figures S2A et S2B. L'illustration montre une fleur (*Linaria vulgaris*) normale (S2B) et son analogue pélorique (S2A). [Une plante pélorique est une plante qui, suite à une anomalie, ne présente plus des fleurs zygomorphes (symétrie bilatérale : S2B) mais actinomorphes (symétrie radiale : S2A).]

Ces *Linaria* péloriques résultent d'une épimutation héréditaire, causée par une vaste méthylation du gène Lcyc, qui devient alors silencieux. Cette mutation épigénétique entraîne une modification du phénotype floral. La Figure S2C montre, de profil, l'évolution (de gauche à droite) entre l'espèce sauvage (gauche) et l'épimutant (droite) en passant par divers individus semi-péloriques.²²⁰



Figure S2. Effet de l'environnement sur des *Linaria vulgaris*.

Ainsi, la méthylation de l'ADN, et plus particulièrement l'influence qu'a l'environnement sur elle, lui permet de jouer un rôle majeur sur le développement, sur l'évolution du génome et par conséquent sur le phénotype des individus. De plus, ce nouveau paradigme met en avant le fait que notre style de vie semble être à l'origine de l'apparition de différentes pathologies et pourrait expliquer la plasticité du génome pour une meilleure adaptation à l'environnement (évolution).²²¹

ANNEXES S3

PUBLICATION 6

De nombreux phénomènes endommagent quotidiennement notre ADN (erreurs lors de la réPLICATION, rayonnements UV, radicaux libres, substances carcinogènes, ...). Fort heureusement, le monde du vivant dispose de systèmes capables de détecter et de réparer en continue ces anomalies avec une grande efficacité. La réparation de l'ADN par excision de base (BER - *Base Excision Repair*) est l'un de ces systèmes. Il consiste à éliminer une base ayant subi un dommage (oxydation, par exemple), puis à cliver le 2'-désoxyribose afin de pouvoir introduire un nucléotide non endommagé. *Tomas Lindahl, Paul Modrich et Aziz Sancar* ont d'ailleurs récemment (Mercredi 07 Octobre 2015) été honorés du prix Nobel de Chimie pour leurs recherches portant sur la réparation de l'ADN.²²²

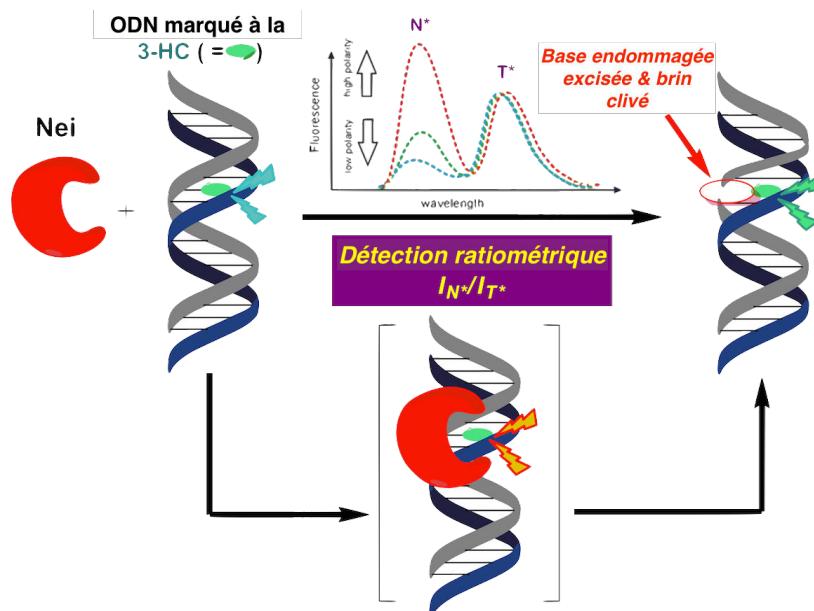


Figure S3. Détection de l'excision de base en présence de 3-HC.

C'est dans ce contexte que l'étude de l'excision d'une base endommagée, la 5,6-dihydrouridine (DHU), spécifiquement reconnue et éliminée par l'endonucléase VIII, a été réalisée, à l'aide de l'analogique nucléosidique de 3-hydroxychromone de première génération [cf. Figure S3]. Plus particulièrement, une comparaison avec d'autres rapporteurs commerciaux (2-AP **67**, tC° **86** et CPY) a été réalisée afin d'évaluer ses capacités à sonder les changements conformationnels de l'ADN qui ont lieu lors de l'excision d'une telle base endommagée.

Pour cela, des fragments d'ADN incorporant les différents analogues fluorescents de nucléosides ont été synthétisés et étudiés. Les fluorophores, incorporés soit en face, soit à proximité de la DHU endommagée, ont tous montré des sensibilités différentes aux changements conformationnels du duplexe d'ADN, induits par la liaison de l'enzyme. La sensibilité la plus élevée et les informations les plus détaillées ont cependant été obtenues en utilisant le fluorophore à motif 3-HC. A travers ces travaux, il a été démontré que cette

ANNEXES

base fluorescente artificielle est un outil particulièrement intéressant pour l'étude de mécanismes complexes mettant en jeu des partenaires protéiques et nucléiques. Ces travaux firent l'objet d'un article, publié en Juin 2014, dans *PloS ONE*.^{126e}

New Environment-Sensitive Multichannel DNA Fluorescent Label for Investigation of the Protein-DNA Interactions

Alexandra A. Kuznetsova^{1*}, Nikita A. Kuznetsov^{1*}, Yuri N. Vorobjev¹, Nicolas P. F. Barthes², Benoît Y. Michel², Alain Burger^{2*}, Olga S. Fedorova^{1*}

1 Siberian Branch of the Russian Academy of Sciences, Institute of Chemical Biology and Fundamental Medicine, Novosibirsk, Russia and Department of Natural Sciences, Novosibirsk State University, Novosibirsk, Russia, **2** Institut de Chimie de Nice, UMR 7272, Université de Nice Sophia Antipolis, CNRS, Nice, France

Abstract

Here, we report the study of a new multichannel DNA fluorescent base analogue 3-hydroxychromone (3HC) to evaluate its suitability as a fluorescent reporter probe of structural transitions during protein-DNA interactions and its comparison with the current commercially available 2-aminopurine (aPu), pyrrolocytosine (C^{PY}) and 1,3-diaza-2-oxophenoxazine (tC^O). For this purpose, fluorescent base analogues were incorporated into DNA helix on the opposite or on the 5'-side of the damaged nucleoside 5,6-dihydrouridine (DHU), which is specifically recognized and removed by Endonuclease VIII. These fluorophores demonstrated different sensitivities to the DNA helix conformational changes. The highest sensitivity and the most detailed information about the conformational changes of DNA induced by protein binding and processing were obtained using the 3HC probe. The application of this new artificial fluorescent DNA base is a very useful tool for the studies of complex mechanisms of protein-DNA interactions. Using 3HC biosensor, the kinetic mechanism of Endonuclease VIII action was specified.

Citation: Kuznetsova AA, Kuznetsov NA, Vorobjev YN, Barthes NPF, Michel BY, et al. (2014) New Environment-Sensitive Multichannel DNA Fluorescent Label for Investigation of the Protein-DNA Interactions. PLoS ONE 9(6): e100007. doi:10.1371/journal.pone.0100007

Editor: Vladimir N. Uversky, University of South Florida College of Medicine, United States of America

Received April 15, 2014; **Accepted** May 20, 2014; **Published** June 12, 2014

Copyright: © 2014 Kuznetsova et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. All data are included within the manuscript.

Funding: The study was supported by the Program of the Russian Academy of Sciences "Molecular & Cell Biology" (No 6.11), the Grants from Russian Foundation for Basic Research (12-04-00135, 14-04-00531, 13-04-00013 and 14-04-31174) and Russian Ministry of Education and Science (SS-1205.2014.4 and SP-4012.2013.4), the Grant from Russian Scientific Foundation (14-14-00063). The ANR (ANR-12-BS08-0003-02) and the FRM (DCM20111223038) are thanked for financial support and grant for N.P.F.B. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: fedorova@niboch.nsc.ru (OSF); burger@unice.fr (AB)

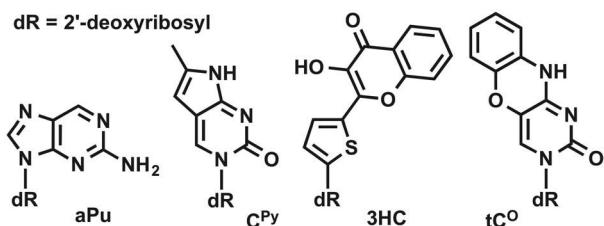
• These authors contributed equally to this work.

Introduction

Understanding at the molecular level the dynamics and functions of enzymes in interactions with their DNA targets is of main importance in biology and medicine. In this context, fluorescence spectroscopy is widely used in nucleic acids research to study the structure and the dynamics as well as the kinetics of protein-DNA interactions [1]. Different fluorescent analogues of DNA bases have been evaluated for examining of DNA conformational transitions [2–8]. Most of them are sensitive to the quenching of DNA nucleobases. The degree of quenching depends on the conformation of DNA and on the composition of the bases surrounding the fluorescent dye. The sensitivity to nucleobase quenching has therefore been exploited to study in real time the dynamic and function of nucleic acids. The most commonly used fluorescent base analogue is 2-aminopurine (aPu, Fig. 1). 2-Aminopurine can form stable base pairs with thymine but also moderately stable base pairs with cytosine [9]. The sensitivity of aPu fluorescence to the microenvironment has been utilized in the study of wide range protein-DNA interactions [10–14]. The cytosine analogue pyrrolocytosine (3-[β-D-2-ribofuranosyl]-6-methylpyrrolo[2,3-d]pyrimidin-2(3H)-one, C^{PY}, Fig. 1)

forms hydrogen bonds with guanine. In some studies, the DNA duplexes containing C^{PY} as a fluorescent reporter were used for the investigation of their bindings with some proteins [15–17]. Moreover, the tricyclic cytosine analogue as 1,3-diaza-2-oxophenoxazine (tC^O, Fig. 1) was reported as a fluorescent cytosine analogue and recently proposed for studying the physical properties of DNA and protein-DNA interactions [18–21].

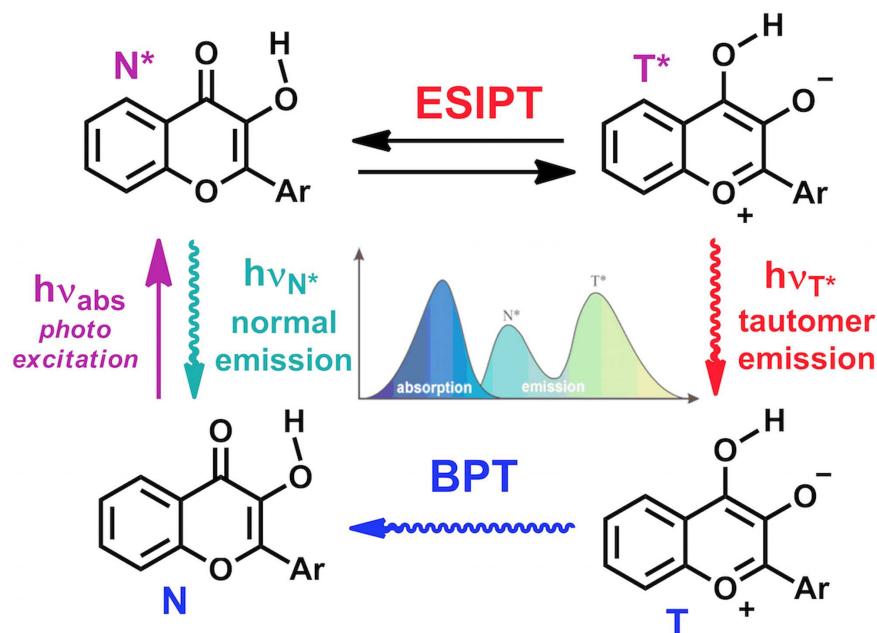
The design of new DNA-based fluorescent labels, with high site-specific responses to intermolecular interactions and sensitive to the environmental change by a different mechanism than quenching, attracts great attention of chemists. New probes are desirable to overcome the limited environment sensitivity of the dye when strongly quenched and to get further insight on the molecular mechanisms of their interactions by sensing new mechanisms of the fluorescent signal response. Recently, the interesting fluorescent properties of the base analogue 3-hydroxychromone (3HC, Fig. 1) have been reported [22,23]. This probe is up to 50-fold brighter than 2-aminopurine and can substitute any nucleobase in a duplex with minimal perturbation of the duplex structure. The total emission of 3HCs is sensitive to G, C and water mediated quenching, but not A and T. In addition, due to an excited state intramolecular proton transfer (ESIPT), 3HC

**Figure 1. Structures of the fluorescent base analogues.**

doi:10.1371/journal.pone.0100007.g001

fluorophores exhibit two excited states: the initially excited normal form (N^*) and the tautomeric (T^*) one; each form generates one well-resolved emission band (Fig. 2) [24]. The dual emission of 3HCs is highly sensitive to the polarity of the environment because an increase in the hydrogen bond donor strength and the dielectric constant of solvents inhibit the ESIPT reaction and thus, decrease the relative intensity of the T^* band [23,25,26]. Therefore the variation of dipole/dipole interactions and hydration of the microenvironment of 3HC can be monitored by measuring the intensity ratio of the two emission bands (I_{N^*}/I_{T^*}). Taken together, 3HC labeled oligodeoxynucleotides (ODNs) provide different channels of information: the intensiometric channel based on sensing quenching and the ratiometric channel based on sensing hydrogen bond donor strength and dipole/dipole interactions of 3HC microenvironment. An important advantage of the ratiometric dye over conventional single-band intensiometric dyes is that this ratio is independent from unaccountable quenching effects or instrument settings. In addition, 3HC based sensors distinguish from the majority of ratiometric sensors, which employed fluorescence resonance energy transfer (FRET) pairs or excimers and required double labeling. Therefore, 3HC labeled ODNs are prospective sensors for the studies of interactions with proteins [23].

The first objective of this work is a comparative study of the new 3HC fluorescent nucleobase moiety with current commercially available dyes (aPu, C^{Py}, tC^O) according to the following questions. Is the non-natural dye a suitable fluorescent base substitute for monitoring complex mechanism of DNA structural transitions during interaction with a protein [27]? Can the multi-channel probe complement the information obtained from conventional dyes? To address these questions, the Endonuclease VIII from *Escherichia coli* (Nei) was used as a model. Nei is a DNA repair enzyme that removes a wide range of damaged pyrimidine bases from DNA such as 5,6-dihydouracil (DHU), thymine glycol, 5,6-dihydrothymine, 5-hydroxypyrimidines, etc [28,29]. Based on crystal structures of the enzyme free and in complex with the target DNA [30,31], biochemical means [32,33] and stopped flow studies of the protein dynamics [34], a mechanism was proposed to explain how Nei process on the DNA substrate. The recognition of DNA lesions by Nei involves several conformational changes in both protein and DNA, such as DNA kinking, damaged base flipping out from DNA helix, insertion into the enzyme's active site and the intrusion of the enzyme loop (a triad: Gln-69, Leu-70, and Tyr-71) into the void created in DNA after eversion of the damaged base, damaged base excision and product release (Fig. 3). It was shown [29] that the excision of the damaged base by Nei proceeds through several chemical steps: N -glycosidic bond cleavage (Fig. 4, stage i), β -elimination (Fig. 4, stage ii) and δ -elimination (Fig. 4, stage iii) of the phosphates flanking the damaged nucleoside. Since the mechanism involves several conformational changes in both protein and DNA, monitoring the DNA dynamics in time should give further insight about the processing. However, preliminary studies using stopped flow kinetics and aPu labeled DNA with Nei failed to give such information because no significant fluorescence changes of aPu were measured [34]. Therefore, the second objective of this work is to fill this gap and complement the information collected following the protein dynamics.

**Figure 2. Origin of the dual emission of 3-hydroxychromones: the ESIPT reaction.** BPT denotes back proton transfer; N^* and T^* represent the normal and tautomeric emissive forms, respectively.
doi:10.1371/journal.pone.0100007.g002

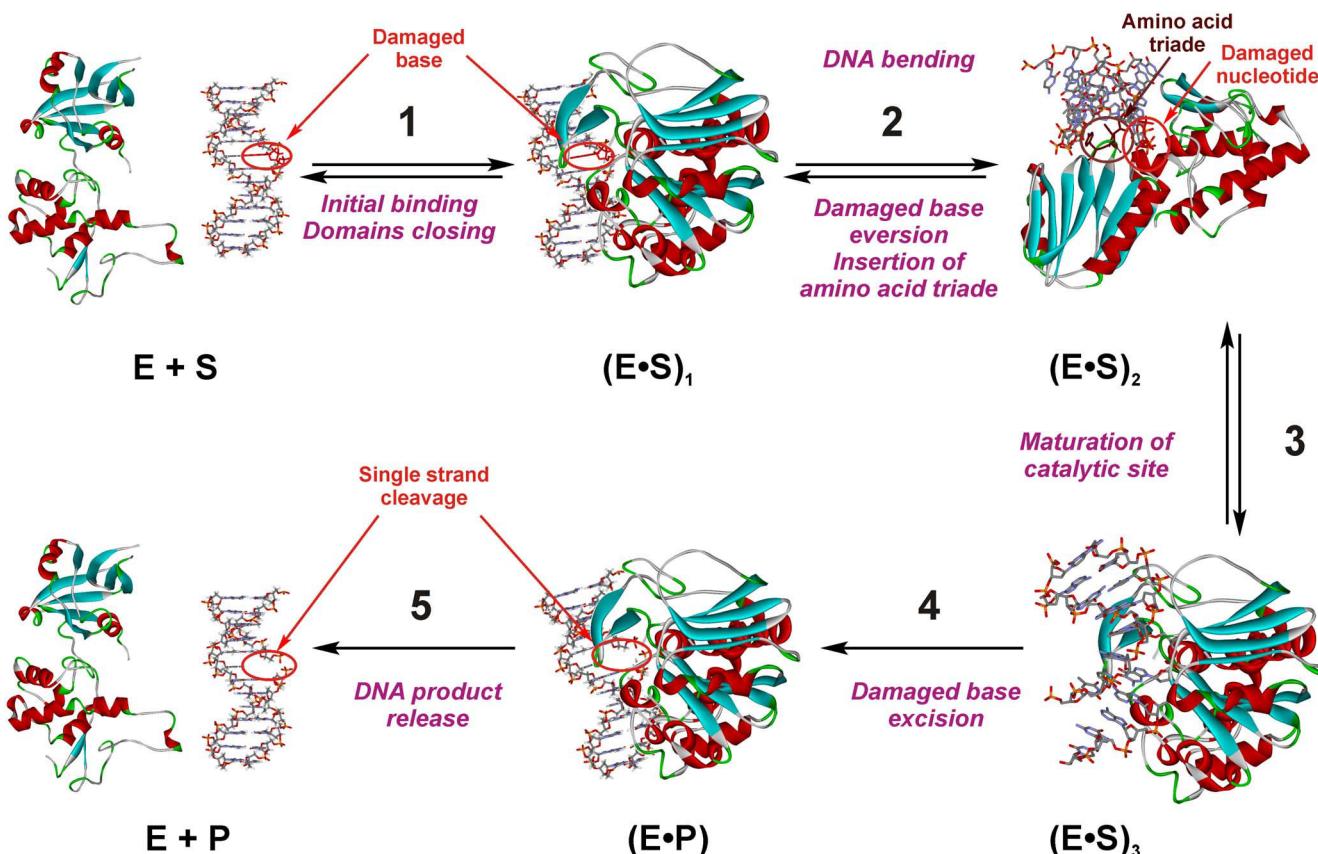


Figure 3. The proposed kinetic mechanism of Nei processing on the DNA substrate containing DHU [34]. Crystal structures of Nei (PDB ID 1Q39 and 1K3W) were used for schematic representation of the enzyme-DNA complexes.
doi:10.1371/journal.pone.0100007.g003

The overall structures of the free Nei and Nei covalently complexed with damaged DNA are depicted in Figure 5A [30,31]. As shown in Figure 5B, base placed on the 5'-side of the damaged nucleotide has the ability to interact only with Leu-70 (shortest distance is 3.7 Å). At the same time, the base placed opposite of damaged nucleotide interacts with Gln-69 (2.9 Å) and Tyr-71 (3.8 Å). Therefore, the bases placed opposite and on the 5'-side of the damaged nucleotide were selected for substitution by the fluorescent base analogues. The canonical and labeled ODNs used in this study are presented in Table 1. They contain a dihydrouracyl base (DHU), as a substrate for Nei,

Since the DNA is known to undergo conformational transitions upon binding by Nei, the stopped-flow kinetic assay was used to

measure the changes of fluorescence intensities of the various DNA base labels in the model DNA-substrates.

Materials and Methods

Oligonucleotides and Enzymes

Nei protein was purified as described [31]. The concentration of the active enzyme (~80%) was determined by trapping with NaBH₄ as described [34] using the ODN duplex containing a DHU/G pair; the reported concentrations of Nei are those of the active form. The ODNs (Table 1) were synthesized by established phosphoramidite methods on an ASM-700 synthesizer (BIOSSET Ltd., Novosibirsk, Russia) from phosphoramidites purchased from Glen Research (Sterling, VA) in the Laboratory of Bionanotech-

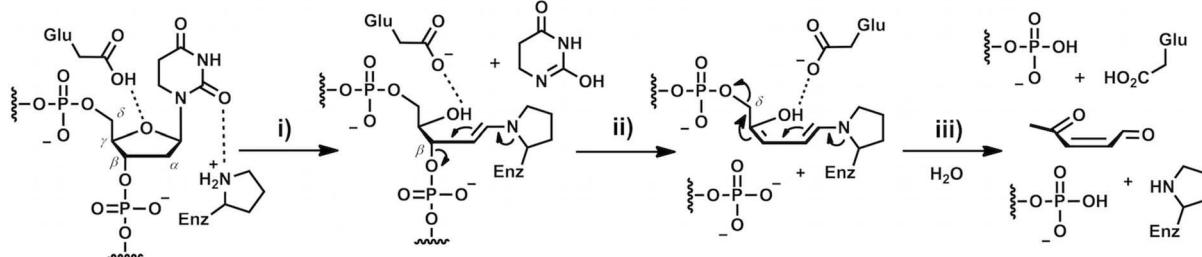


Figure 4. Main chemical steps of Nei catalysis. Step i: acid activation of the nucleobase and Schiff base formation; step ii: β -elimination of 3'-phosphate, step iii: δ -elimination of 5'-phosphate.
doi:10.1371/journal.pone.0100007.g004

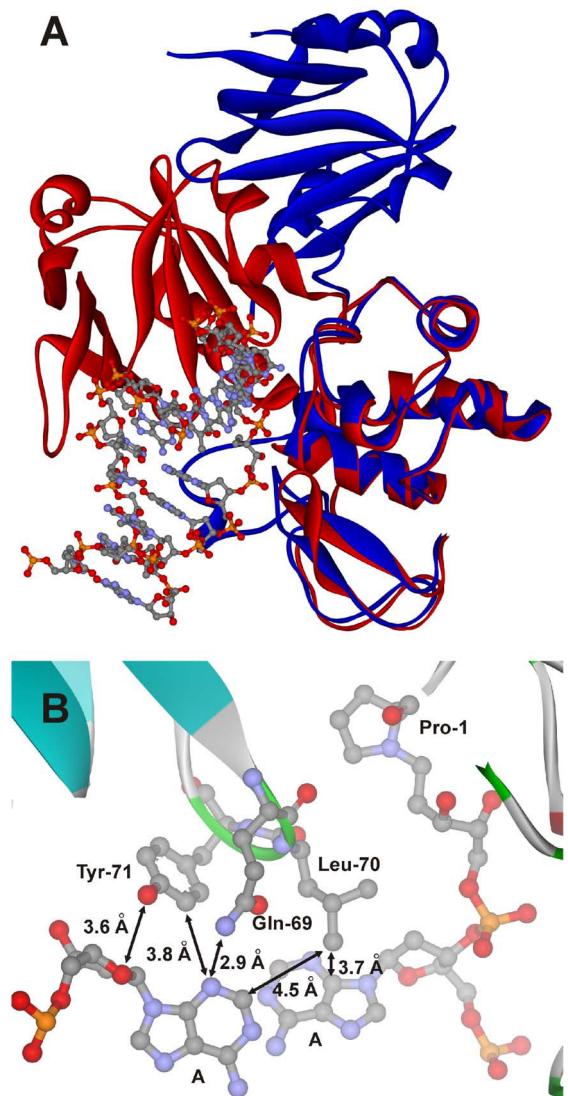


Figure 5. Overall structure of the Nei and close-up view of the active site. (A) The structures of the free Nei (blue, PDB ID 1Q39) and Nei covalently complexed with damaged DNA (red, PDB ID 1K3W). (B) The Gln-69, Leu-70, and Tyr-71 triad inserted into the duplex. The arrows indicate the distances between Tyr-71 C_ε2 or Gln-69 N_ε2 and N3 of the A base placed opposite of the damaged nucleotide (3.8 Å or 2.9 Å); between Tyr-71 O_η and O4' of the nucleotide (3.6 Å); between Leu-70 C_δ1 and C2 of the A base placed opposite of the damaged nucleotide (4.5 Å) or C4 of the A base placed on the 5'-side of the damaged nucleotide (3.7 Å).

doi:10.1371/journal.pone.0100007.g005

nology of ICBFM. 3HC phosphoramidite was synthesized as described previously [22,23]. Synthetic oligonucleotides were unloaded from the solid support with ammonium hydroxide according to manufacturer's protocols. Deprotected oligonucleotides were purified by HPLC. The purity of ODNs exceeded 98% as estimated by electrophoresis in 20% denaturing PAGE after staining with the Stains-All dye (Sigma-Aldrich). Concentrations of oligonucleotides were determined from their absorbance at 260 nm. ODN duplexes were prepared by annealing modified and complementary strands at a 1:1 molar ratio.

Stopped-flow Fluorescence Measurements

Stopped-flow measurements with fluorescence detection were essentially carried out as described [35,36]. A model SX.18MV stopped-flow spectrometer (Applied Photophysics Ltd, Leatherhead, UK) fitted with a 150 W Xe arc lamp and 2 mm path length optical cell was used. The dead time of the instrument was 1.4 ms. Experiments were performed at 10°C in the buffer containing 50 mM Tris-HCl (pH 7.5), 50 mM KCl, 9% (v/v) glycerol, 1 mM DTT, 1 to 3 μM Nei and 1 μM DNA substrates. The excitation wavelengths were 290, 310, 344, 360 and 375 nm for the Trp, aPu, C^{Py}, tC^O and 3HC fluorescent dyes, respectively. The emission was monitored using a 320 nm long pass wavelength filter for Trp, 370 nm long pass wavelength filter for aPu and C^{Py}, 395 nm long pass wavelength filter for tC^O and 3HC. Additionally, to separate detection of the second emission T* band of 3HC fluorophore, 495 nm long pass filter was used. Typically, each trace shown is the average of four or more traces obtained in individual experiments. Data obtained from the stopped-flow kinetic assays and pre-steady-state kinetic parameters for possible kinetic schemes were determined as described previously [34].

To calculate the intensity ratio (I_{N^*}/I_{T^*}) for each time point, the equation (1) was used.

$$I_{N^*}/I_{T^*} = (I_{(N^*+T^*)} - I_{T^*})/I_{T^*} \quad (1)$$

where $I_{(N^*+T^*)}$ is the total N*+T* fluorescence intensity (395 nm filter), I_{T^*} is the T* fluorescence intensity (495 nm filter) at the same time point of the kinetic curves.

PAGE Time-course Experiments

The reaction mixture included 2 μM Nei, 2 μM ODN substrate, 50 mM Tris-HCl (pH 7.5), 50 mM KCl, 9% (v/v) glycerol, and 1 mM DTT. All experiments were conducted at 10°C. The modified strands of the ODN substrates were 5'-end labeled using T4 polynucleotide kinase (New England Biolabs, Beverly, MA) and γ [32P]-ATP (4500 Ci/mol) (Radioizotop, Moscow, Russia) according to the manufacturer's protocol. The reactions were initiated by adding the enzyme. Aliquots (2 μL) of the reaction mixture were quenched with 3 μL of gel-loading dye

Table 1. ODN sequences used in this work.

Shorthand	Sequence
DHU/X	5'-CTCTC(DHU)CCTTCC-3'
X = G, aPu, C ^{Py} , tC ^O , 3HC	3'-GAGAG X GGAAGG-5'
Y-DHU	5'-CTCTY(DHU)CCTTCC-3'
Y = aPu, 3HC; Z = C	3'-GAGAZ Z GGAAGG-5'
Y = C ^{Py} ; Z = G	

doi:10.1371/journal.pone.0100007.t001

containing 7 M urea, and loaded on a 20% (w/v) polyacrylamide/7 M urea gel. The disappearance of substrate and the formation of product were analyzed by autoradiography and quantified by scanning densitometry using Gel-Pro Analyzer v4.0 software (Media Cybernetics, Bethesda, MD).

Molecular Modeling

As a complement to experimental studies, we performed molecular dynamics (MD) simulations using the BIOPASED program [37] with implicit Gaussian Shell solvation model extended for DNA [38]. Structures of eight DNA duplexes have been modeled by MS simulated annealing at 300 K. Comparison of conformational distortion of duplexes structures due to fluorescent nucleotide analog are performed at 50 K to elucidate structural differences between average structures. 3HC analog has the largest structural distortion from the ideal B-DNA \sim 1.2 Å and largest amplitude of thermal fluctuations at 300 K comparing to all other analogs investigated.

Results and Discussion

First, molecular dynamics (MD) simulation of the structures of all used DHU/X and Y-DHU model duplexes was performed to verify that the labels have minimal impact on the duplex structure (Fig. 6). It was shown that fluorescent labels did not disturb noticeably the regular structure of DNA double helix. The maximum deviation of DHU base in these duplexes is 1.2 Å, i.e. in the range of thermal fluctuation of the canonical DHU/G duplex.

Secondly, the duplex DHU/G was used as a control of the enzyme activity. The structural data reveal gross movements in the structure of Nei upon DNA binding, with the N-terminal and C-terminal domains “closing” by \sim 50° to assemble the enzyme’s active site (Fig. 5) [30,31]. The conformational changes of the enzyme during DNA-substrate binding, lesion recognition and excision were registered by the intrinsic fluorescence of the enzyme’s Trp residues [34]. As depicted in Figure 7A, complex fluorescence dynamics was observed as it was expected based on previous data [34]. The minimal kinetic scheme describing the observed changes of Trp fluorescence intensity was identical to scheme in Figure 3 and contained three equilibrium steps that characterized substrate binding followed by an irreversible chemical step and then an equilibrium step of product release. The rate constants of the elementary steps estimated according to this kinetic scheme are listed in Table 2. It should be noticed that formation of the first enzyme/substrate complex $(E \cdot S)_1^{Trp}$ proceeded during the first 50 ms. The maximum of the concentration of the second complex $(E \cdot S)_2^{Trp}$ was achieved at 0.6 s, whereas catalytically active complex $(E \cdot S)_3^{Trp}$ was formed at 5–10 s. The irreversible base excision and strand cleavage, as well as the product release are undistinguishable following Trp fluorescence and proceeded at times >20 s. The formation of E·P complex in the chemical step registered by Trp fluorescence changes is in good agreement with the accumulation of the reaction products detected by PAGE analysis (Fig. 7B). The first reversible step likely corresponds to the initial binding, which changes the shielding of the Nei Trp residues. The two remaining steps discernible by Trp fluorescence may reflect the insertion of

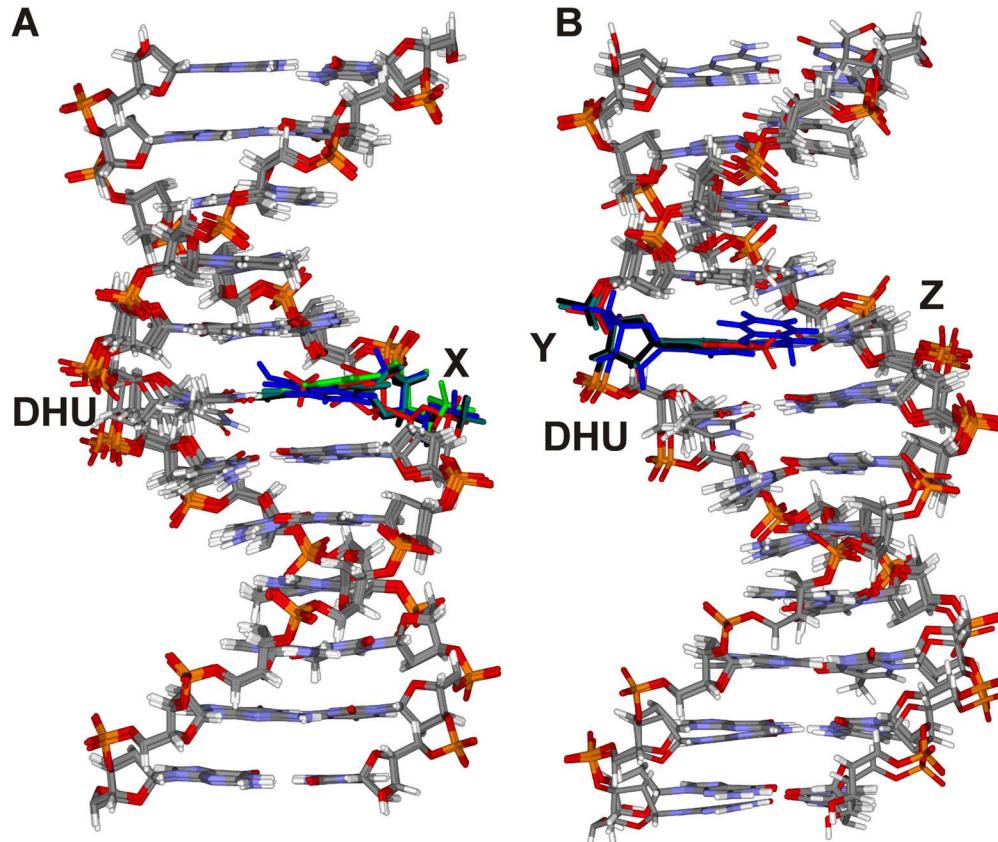


Figure 6. Comparison of the MD simulations of DNA duplexes. The MD simulation of DHU/X (A) and Y-DHU (B) model duplexes. The colors correspond to the duplexes containing aPu (black), CPY (red), tC^O (green), 3HC (blue), unlabeled duplex (dark green).
doi:10.1371/journal.pone.0100007.g006

amino acid triad into DNA, in exchange of the damaged base everted into the enzyme's active site, as well as movement of the zinc finger resulting in the fine-tuning of the enzyme active site. These data are consistent with the hypothesis that Nei maintains a stable conformation during the breakage of the covalent bonds in DNA (Fig. 3) [34]. It is noteworthy that the *N*-glycosidic bond cleavage is faster than the β -elimination process, making of this latter the rate-limiting step.

We then analyzed the DNA dynamics following the fluorescence intensity of the label (C^{Py} , aPu, tC^O or 3HC) in time. Previously, using C^{Py} opposite to the damaged base, significant changes of C^{Py} fluorescent intensity were registered during interaction of the Fpg (the structural homolog of Nei) with DNA [17]. According to the analogy with Fpg, we tried to register the DNA conformational changes in the course of interaction with Nei using the DHU/X model duplexes containing fluorescent labels (C^{Py} , aPu, tC^O or 3HC) opposite to the damaged base (Fig. 8A&B). It is important to notice that the processes of structural rearrangements in protein and DNA can occur with a time shift relative to each other, because of an intrinsic inertia of the DNA conformational and environmental changes induced by protein interactions (Fig. 8A&B). As shown in Figure 8C, aPu and C^{Py} did not affect the Nei activity and both tC^O and 3HC fluorescent base mimics retained substantial enzymatic activity [39].

The aPu fluorescence did not change at all (Fig. 8A&B). The absence of the aPu signal change denotes that the most widely used fluorescent base is unable in this example to sense the DNA conformational transitions occurring during protein-DNA interactions when the dye is placed opposite of the damaged base. As shown in Figures 9A&B, fluorescence of C^{Py} exhibited three different steps. On the basis of data (Fig. 9A), the minimal kinetic scheme (Fig. 10) was proposed for the interaction between Nei and DHU/ C^{Py} -substrate whereas the values for rate constants were reported in Table 2. Figure 9B illustrates the time-course of appearance and disappearance of the different intermediates as evidenced from C^{Py} fluorescence. The fast increase during 10 ms likely corresponds to the beginning of the first stage detected by Trp fluorescence. The gradual decrease of C^{Py} fluorescence between 10 ms and 10 s can be attributed to the end of the first stage and to the formation of complexes $(E \cdot S)_2^{Trp}$ and $(E \cdot S)_3^{Trp}$, which cannot be distinguished by C^{Py} fluorescence. The significant decrease of C^{Py} intensity after 90 s is likely connected to complex $(E \cdot P)^{Trp}$ and product release.

Using the tC^O dye, it was also possible to register the processes of three different DNA conformational transitions (Fig. 11A&B) during interaction with Nei but different from C^{Py} (Fig. 9A&B). The initial slight increase of the tC^O fluorescence during 0.6 s

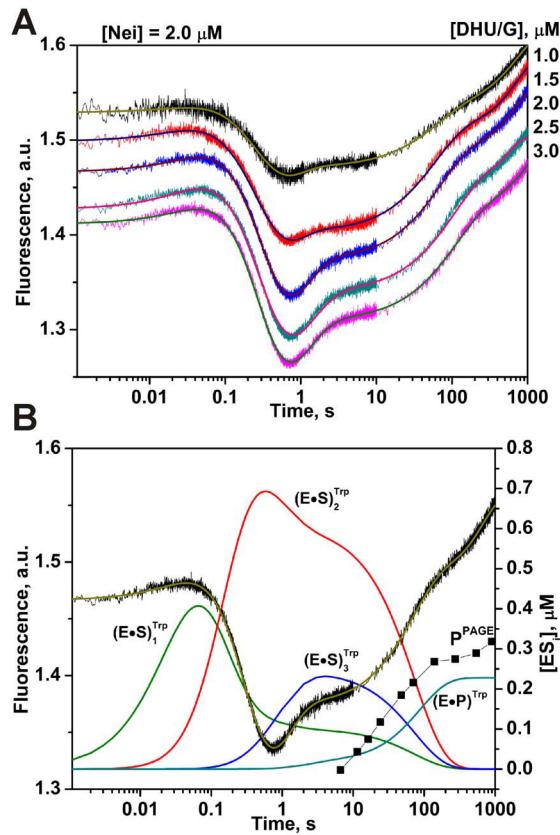


Figure 7. Interaction of Nei with DHU/G-substrate. (A) Changes in Trp fluorescence intensity during the interaction of Nei with DHU/G-substrate. Jagged traces display the experimental data whereas smooth curves correspond to the fit of the data to the kinetic model (Fig. 3). Concentrations of the DNA (μM) are shown next to the plots. (B) Time-course of appearance and disappearance of transient enzyme–substrate complexes during the cleavage of DHU/G by Nei. Modeling was done for 2 μM Nei and 2 μM substrate with the rate constants listed in Table 2. Accumulation of the reaction products detected by PAGE analysis (P^{PAAG}) is shown for comparison.
doi:10.1371/journal.pone.0100007.g007

characterizes the formation of $(E \cdot S)_1^{Trp}$ and $(E \cdot S)_2^{Trp}$ complexes in correspondence with Trp fluorescence data. The decreasing phase of the fluorescence intensity after 1 s up to 30 s is reasonably attributed to the $(E \cdot S)_3^{Trp}$ complex. The catalytic steps and dissociation of the complex of enzyme with the reaction product $(E \cdot P)^{Trp}$ marginally influence on the tC^O fluorescence. Only two

Table 2. The rate constants for interactions of Nei with DHU/G and DHU/X-substrates.

Constants	Trp	3HC	tC ^O	C^{Py}
$k_1, \text{M}^{-1}\text{s}^{-1}$	$(10.5 \pm 6.2) \times 10^6$	$(53 \pm 10) \times 10^6$	$(200 \pm 70) \times 10^6$	$(115 \pm 30) \times 10^6$
k_{-1}, s^{-1}	6.1 ± 3.4	69 ± 22	380 ± 150	246 ± 44
k_2, s^{-1}	9.5 ± 2.4	3.9 ± 1.7	0.46 ± 0.16	0.52 ± 0.11
k_{-2}, s^{-1}	1.6 ± 0.2	4.8 ± 1.5	0.005 ± 0.002	2.5 ± 0.2
k_3, s^{-1}	0.4 ± 0.1	0.30 ± 0.08	-	-
k_{-3}, s^{-1}	0.9 ± 0.2	0.4 ± 0.1	-	-
k_{cat}, s^{-1}	0.056 ± 0.027	0.014 ± 0.003	0.035 ± 0.015	0.040 ± 0.028
K_P, M	$(6.5 \pm 3.4) \times 10^{-6}$	$(0.16 \pm 0.08) \times 10^{-6}$	$(6.0 \pm 2.5) \times 10^{-6}$	$(3.6 \pm 2.1) \times 10^{-6}$

doi:10.1371/journal.pone.0100007.t002

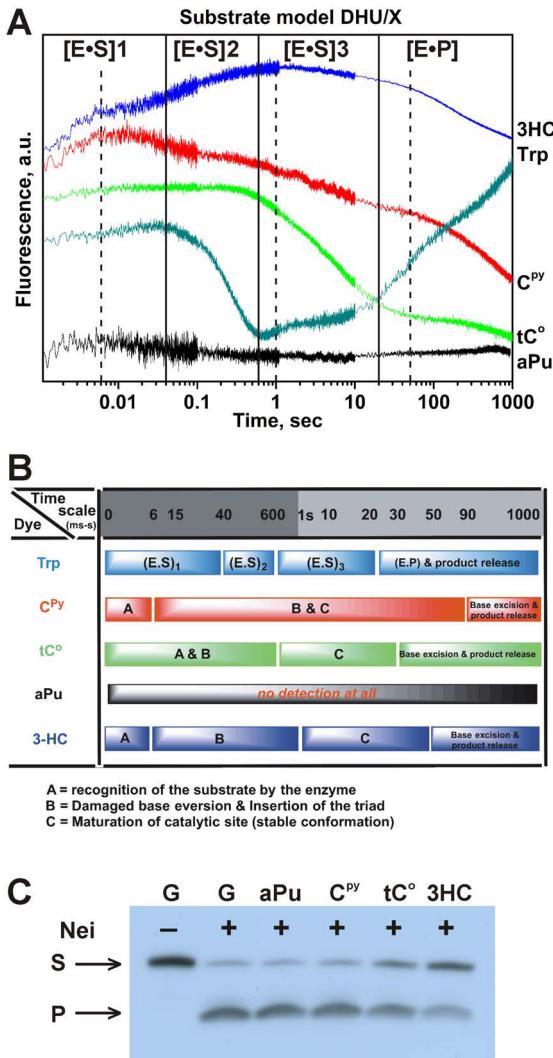


Figure 8. Interaction of Nei with DHU/X-substrates. (A) Changes in Trp, aPu, C^{Py}, tC° and 3HC fluorescence intensity during the interaction of 2 μM Nei and 1 μM DNA substrates, containing fluorescent nucleotides opposite the damaged nucleotide. The emission of the long-wavelength T* band (495 long wavelength pass filter) of 3HC dye is presented. The black vertical lines delimit the 4 kinetic steps of protein dynamics following Trp fluorescence. The dashed black vertical lines delimit the 4 kinetic steps of DNA dynamics following 3HC fluorescence. (B) Summary of the different steps detected by the different fluorescent reporters used in this study. (C) Effect of the fluorescent labels on the enzymatic activity of Nei. The concentrations of Nei and DNA substrates were 2 μM and 1 μM, respectively.

doi:10.1371/journal.pone.0100007.g008

binding equilibria were sufficient to describe conformational changes in DNA in minimal terms (Fig. 10). Table 2 lists the values of rate constants determined for this scheme.

By contrast to the C^{Py} and tC° labels, the kinetic curves obtained using 3HC label showed four steps (Fig. 12A&B) as were found by measurements of Trp fluorescence (Fig. 3). As for C^{Py} label, it demonstrates the fast initial increase of the fluorescence intensity, which is connected with the early formation of (E•S)₁^{3HC} complex. However, it is also able to differentiate the formation of a second DNA conformational change (complex (E•S)₂^{3HC}) as evidenced by the second intensity increase phase up to 1 s. The second phase is likely related to the end of the formation of

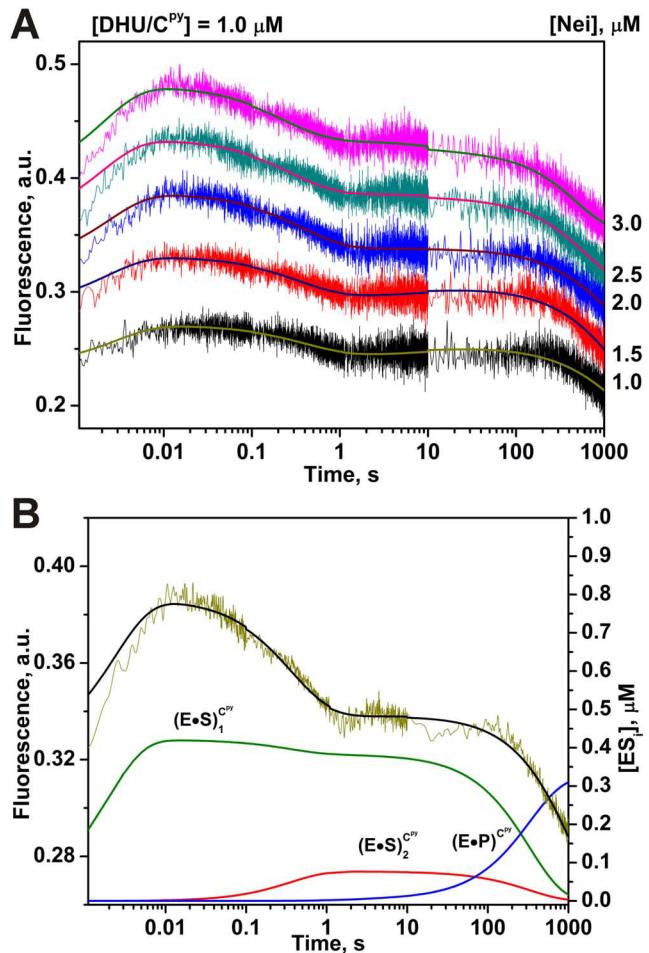


Figure 9. Interaction of Nei with DHU/CPY-substrate. (A) Changes in C^{Py} fluorescence intensity during the interaction of Nei with DHU/CPY-substrate. Jagged traces show the experimental data, smooth curves correspond to the fit of the data to the kinetic model (Fig. 10). Concentrations of the DNA are shown next to the plots. (B) Time-course of appearance and disappearance of transient enzyme–substrate complexes during the cleavage of DHU/CPY by Nei. Modeling was done for 2 μM Nei and 1 μM substrate with the rate constants listed in Table 2.

doi:10.1371/journal.pone.0100007.g009

(E•S)₁^{Trp} and to (E•S)₂^{Trp}. Between 1 and 50 s, the third step is accompanied by a small decrease in the fluorescence signal and leads to the formation of catalytically active complex (E•S)₃^{3HC}. The dissociation of enzyme-product complex (E•P)^{3HC} returns the fluorescence intensity to the initial level (Fig. 12A). Therefore, the difference in Trp and 3HC fluorescence changes during interactions is more informative compared to the other dyes, that allows correlating more precisely the fluorescence variations to the conformational changes of enzyme and DNA.

Comparison of the individual rate constants (Table 2) for processes detected by different fluorophores (Trp, C^{Py}, tC° and 3HC) has shown that DNA faster responds to the initial binding. The rate constant k_1 , which characterizes DNA conformational change, is at least 5 times larger than the same rate constant for enzyme conformational change. Probably, this is due to the fact that the change in intensity of Trp residues occurs only with significant conformational changes of the protein, such as the domains “closing”. The rate constants k_2 of the second binding step detected by Trp and 3HC differ insignificantly (in 2.4 times),

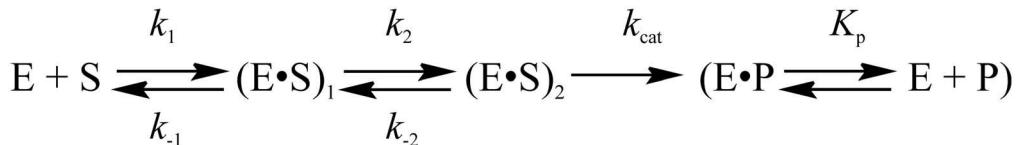


Figure 10. Kinetic mechanism of Nei processing of the DHU/C^{PY}- and DHU/tC^O-substrates.
doi:10.1371/journal.pone.0100007.g010

suggesting that both fluorophores characterize the sequentially coordinated processes, which can be DNA bending, DHU base eversion and insertion of the amino acids of enzyme into the DNA duplex. Interesting to note that rate constants k_3 of the third binding step detected by Trp and 3HC and rate constants k_2 of the second binding step detected by C^{PY} and tC^O have close values, indicating that rearrangements of the active site to achieve the catalytic structure requires mutual simultaneous conformational changes of enzyme and DNA. The rate constants k_{cat} of the catalytic step are a little reduced (~1.5 times) for DHU/C^{PY}- and

DHU/tC^O-substrates in comparison with DHU/G-substrate. For DHU/3HC-substrate this value differs from DHU/G-substrate by a factor 4 that is in accordance with the qualitative data obtained for PAGE analysis of the reaction product (Fig. 8C).

To get further insight about the mechanism of Nei interacting with DNA, we examined in time the intensity ratio changes of the dual emissive 3HC labels. Single and double stranded labeled ODNs displayed the two well-resolved emission bands of 3HC, with the short- and long-wavelength maxima (N* and T*) centered at 430 nm and 540 nm, respectively (Fig. 13A). For both single

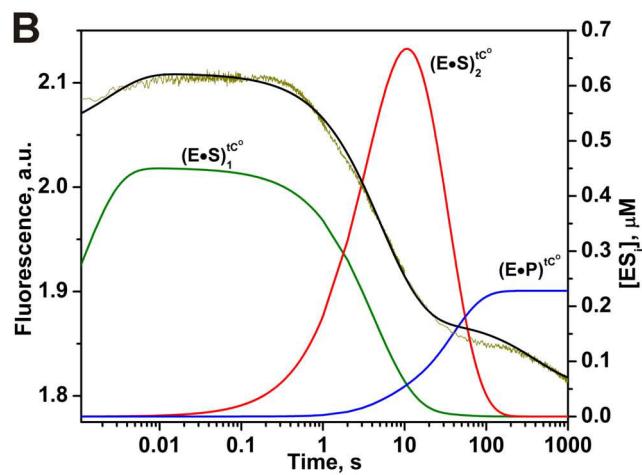
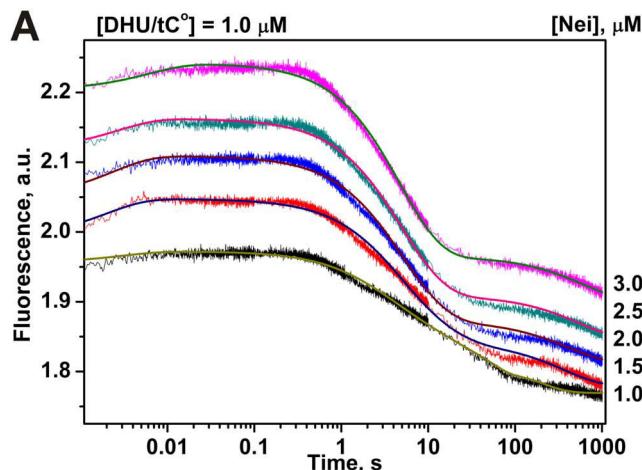


Figure 11. Interaction of Nei with DHU/tC^O-substrate. (A) Changes in tC^O fluorescence intensity during the interaction of Nei with DHU/tC^O-substrate. Jagged traces show the experimental data, smooth curves correspond to the fit of the data to the kinetic model (Fig. 10). Concentrations of the DNA are shown next to the plots. (B) Time-course of appearance and disappearance of transient enzyme-substrate complexes during the cleavage of DHU/tC^O by Nei. Modeling was done for 2 μM Nei and 1 μM substrate with the rate constants listed in Table 2.
doi:10.1371/journal.pone.0100007.g011

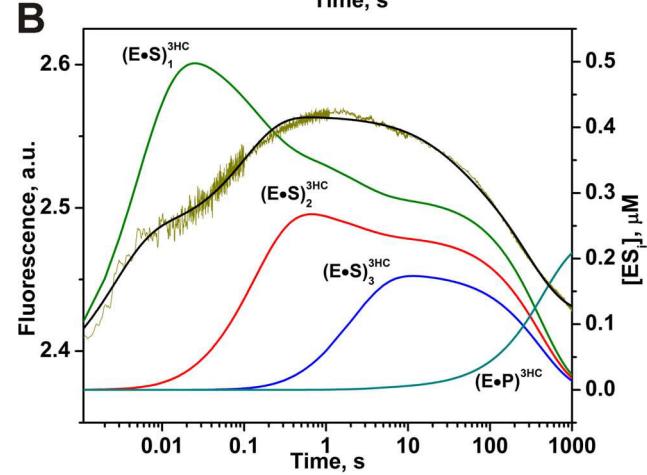
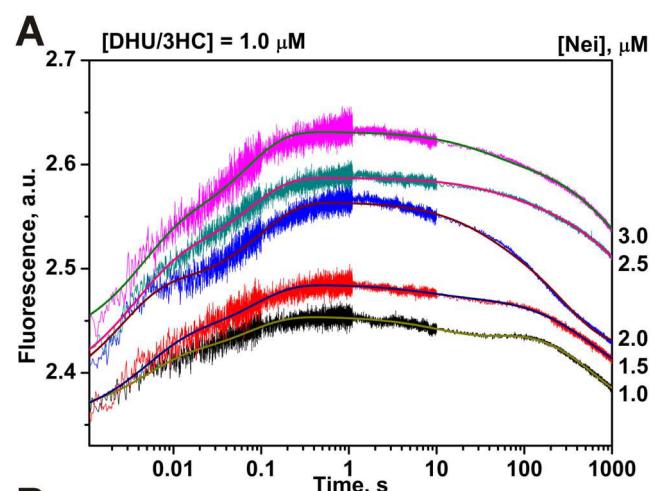


Figure 12. Interaction of Nei with DHU/3HC-substrate. (A) Changes in 3HC fluorescence intensity during the interaction of Nei with DHU/3HC-substrate. Jagged traces show the experimental data, smooth curves correspond to the fit of the data to the kinetic model (Fig. 3). Concentrations of the DNA are shown next to the plots. (B) Time-course of appearance and disappearance of transient enzyme-substrate complexes during the cleavage of DHU/3HC by Nei. Modeling was done for 2 μM Nei and 1 μM substrate with the rate constants listed in Table 2.
doi:10.1371/journal.pone.0100007.g012

and double stranded ODNs, the intensity ratio of the two bands (I_{N^*}/I_{T^*}) was close to 0.3. These results are consistent with the previously reported data and indicate that the environment of the labeling site is mainly aprotic and have medium polarity [23]. The different long pass filters for the 3HC dye were tested in order to distinguish the changes of the total emission corresponding to both short- and long-wavelength bands (N^* & T^* -395 nm long pass filter) from the only emission of the long-wavelength band (495 nm long pass filter) (Fig. 13B).

As described above four discernible stages were observed on kinetic curves using 3HC (Fig. 12A&B). Moreover, the amplitudes and the slopes of fluorescence intensity changes for the various stages are different for both channels indicating a change of the intensity ratio of the two emissive bands in time (Fig. 13B). Since the total N^*+T^* (395 nm long pass filter) and the T^* (495 nm long pass filter) emissions were collected separately, the kinetic curves of the intensity ratio I_{N^*}/I_{T^*} were calculated.

The first stage of the DHU/3HC duplex binding led to a decrease of the intensity ratio up to 6 ms (Fig. 13B). This stage

could be associated with the fast formation of the encountered nonspecific Nei-DNA complex because of binding to protein should reduce DNA structural fluctuations and increased shielding from water molecules. This interpretation is supported by the fact that the hydration and hydrogen bond network of DNA are dramatically affected by the protein interactions [40]. Then, intensity ratio I_{N^*}/I_{T^*} did not change in the time interval between 6 ms and 20 ms. Afterwards, a second decrease occurred up to 1.0 s indicating that the microenvironment of the dye is further shielded from water molecules and exhibited weaker dipole/dipole interactions. The decreasing phase of I_{N^*}/I_{T^*} ratio can be associated with the DNA bending, DHU base eversion and subsequent amino acids insertion, that is in the agreement with Trp data. The I_{N^*}/I_{T^*} ratio exhibited only marginal 3HC fluorescence change between 1 and 50 s. This stage should correspond to the formation of the productive complex. The last stage for the cleavable DHU/3HC substrate showed increase of intensity ratio I_{N^*}/I_{T^*} up to the initial value indicating the returning of the microenvironment of 3HC of the DNA that is more water exposed (Fig. 13B). The last step is therefore compatible with the dissociation of the enzyme-product complex and the release of the DNA product.

To develop the comparison of the fluorophores the model Y-DHU duplexes with different location of the fluorescent labels were used (Table 1). In this model system, the fluorescent dye (aPu, C^{PY}, 3HC) is placed between T and damaged base DHU, which should be flipped out and removed upon the interaction with Nei. The DNA dynamics following in time the fluorescence intensity of the label was monitored (Fig. 14A).

The interaction of Nei with aPu-DHU substrate presented no fluorescence changes up to 100 s despite the fact that the enzyme cleaves this substrate with good efficiency (Fig. 14B). Again aPu fluorescent label is not sensitive for binding by Nei whereas in our previous works related with hOGG1 and Fpg enzymes this fluorescent base was a very informative probe for the study of enzyme-DNA interactions [12–14,41,42]. The increase in the aPu fluorescence intensity at times >100 s can be the result of the transition to the states with poorer stacking and/or quenching abilities of the neighbor bases as a consequence of the strand cleavage and dissociation of the enzyme-product complex (Fig. 14). The binding of Nei with C^{PY}-DHU substrate displayed the slight increase of the C^{PY} fluorescence up to 10 s (Fig. 14). In contrast to the DNA substrates containing aPu and C^{PY}, the duplex 3HC-DHU containing the 3HC fluorophore showed at least four discernible stages as observed with the DHU/3HC model (Fig. 14&15). However, the PAGE analysis of the product accumulation showed that C^{PY} and 3HC labels deprive the enzyme of its catalytic activity (Fig. 14B) by contrast to aPu. In the case of C^{PY} or 3HC, the introduction of the sterically more demanding labels on the 5'-side of DHU in the DNA duplex may inhibit the formation of the catalytically competent state of enzyme (*vide infra*).

In the case of 3HC-DHU duplex, the fast change of the intensity ratio again occurs up to 10 ms. It was clearly discernable when analyzing the intensity ratio channel indicating about formation of the nonspecific Nei-DNA complex (compare Fig. 14&15). A second phase takes place in the time interval between 0.01 s and 0.2 s. It can be suggested that this stage corresponds to the insertion of the amino acids into DNA in exchange of the everted damaged base. The next stage was reached at 1 s, indicating an additional change of the dye environment. Between 1 and 70 s, the next step is accompanied by a marginal change in the fluorescence signal. The last phase was observed above 70 s by a net change of the fluorescence signal

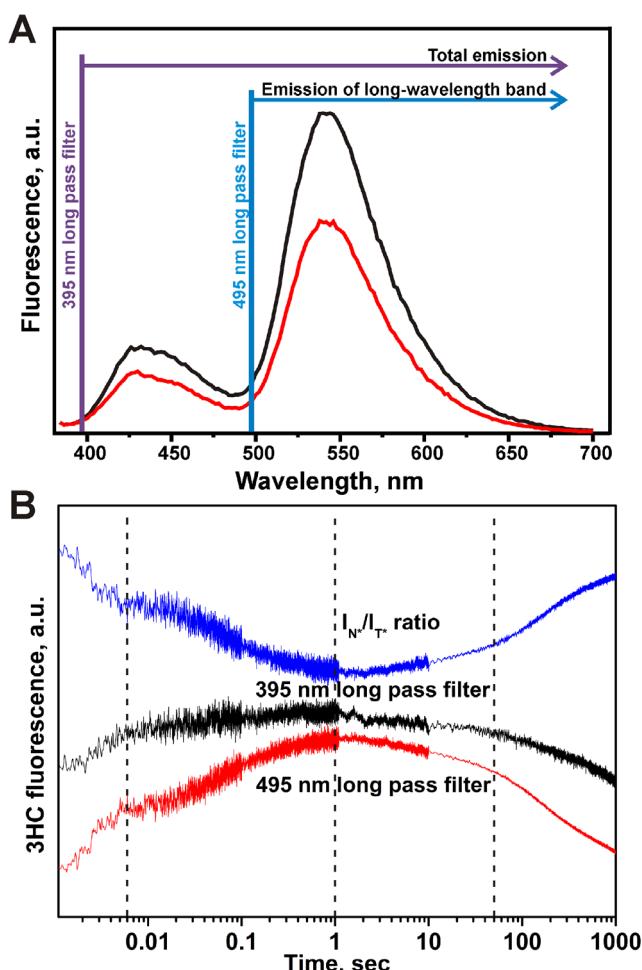


Figure 13. Fluorescent properties of 3HC dye. (A) Steady-state emission spectra of 3HC (2 μ M) in 3HC-DHU single- (black) and 3HC-DHU/GGC double-stranded (red) states. (B) Intensity changes of total (N^*+T^*) and long-wavelength T^* emissions of 3HC and N^*/T^* ratio during the interaction of Nei with DHU/3HC substrate. The concentrations of Nei and DNA substrates were 2 μ M and 1 μ M, respectively. The dashed black vertical lines delimit the 4 kinetic steps of DNA dynamics following 3HC fluorescence intensity ratio.
doi:10.1371/journal.pone.0100007.g013

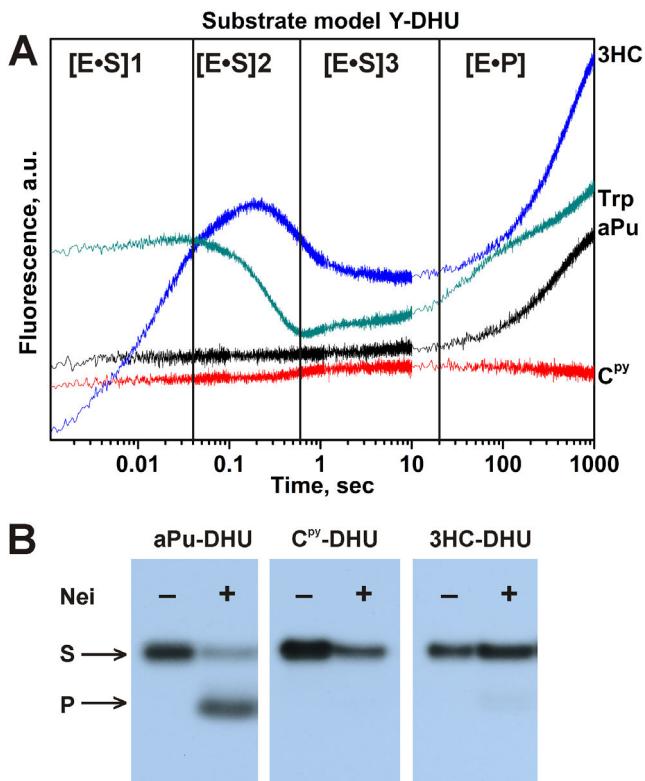


Figure 14. Interaction of Nei with Y-DHU-substrates. The changes in Trp, aPu, C^{Py} and 3HC fluorescence intensity during the interaction of Nei with DNA substrates, containing fluorescent nucleotides on the 5'-side of the damaged nucleotide. The emission of the long-wavelength T* band (495 nm long pass filter) of 3HC dye is presented. The black vertical lines delimit the 4 kinetic steps of protein dynamics following Trp fluorescence. (B) The effect of the fluorescent labels on the enzymatic activity of Nei. The concentrations of Nei and DNA substrate were 2 μ M and 1 μ M, respectively.
doi:10.1371/journal.pone.0100007.g014

(Fig. 14&15). Kinetic curves for uncleavable 3HC-DHU-substrate was evidenced by a decrease in the I_{N^*}/I_{T^*} ratio up to 1000 s (Fig. 15) indicating the attempts of the enzyme to further disturbance of the duplex structure. Indeed, for a productive enzyme-substrate complex, an increase of the I_{N^*}/I_{T^*} ratio should be observed as a consequence of the dye more exposed to water in the final product (CTCT3HC). Therefore the decrease in the I_{N^*}/I_{T^*} ratio for 3HC-DHU is indicative of the formation of a non-productive complex as supported by PAGE analysis of the product accumulation (Figures 14B&15).

Conclusions

To sum up, monitoring of the Trp and 3HC fluorescence (Figures 7A&B, 12A&B and 13B) is complementary and allows following the protein and DNA dynamics, respectively. It gives a detailed picture of the mutual sequential conformational adjustment of enzyme and DNA substrate up to the formation of an active complex and release of the product. The initial Trp fluorescence growth, which is expressed in an increase of the 3HC fluorescence intensity up to 6 ms for DHU/3HC, probably characterizes the fast process of the enzyme domains “closing” and formation of the first nonspecific complex. The DNA conformation and environment changes are reflected by the increase of the quantum yield of 3HC label, probably due to

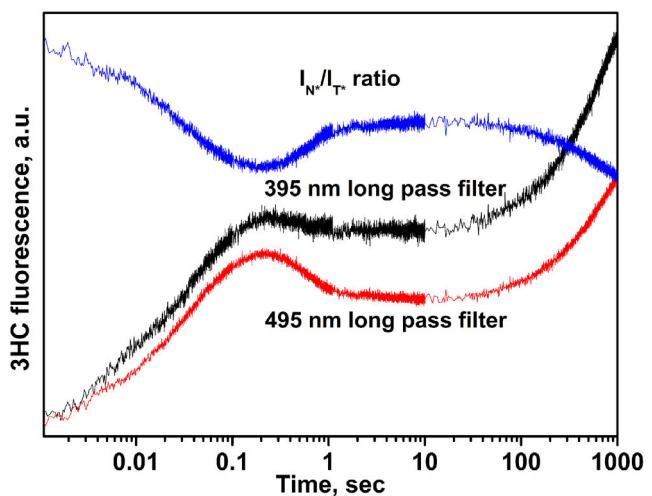


Figure 15. Stopped-flow fluorescence traces for interactions of Nei with 3HC-DHU-substrate. Intensity changes of total (N^*+T^*) and long-wavelength T^* emissions of 3HC and N^*/T^* ratio during the interaction of Nei with 3HC-DHU-substrate. The concentrations of Nei and DNA substrate were 2 μ M and 1 μ M, respectively.
doi:10.1371/journal.pone.0100007.g015

reduced neighboring G, C and water-mediated quenching and reduced DNA structural fluctuations [16]. On the other hand, they are characterized by a decrease of the intensity ratio I_{N^*}/I_{T^*} , indicating that the dye is more shielded from water molecules and that dipole/dipole interactions are weaker (Figure 13B). The analysis of fluorescence curves showed that the first change of the Trp fluorescence intensity (up to 0.6 s) consists of two 3HC fluorescently discernible conformational changes in DNA (the plateau between 0.006 s and 0.02 s and an intensity increase up to 1.0 s). Probably, these phases of 3HC fluorescence are corresponding to the DNA bending, DHU base eversion and subsequent insertion of the amino acid triad Gln-69, Leu-70, and Tyr-71 into the DNA duplex. These processes lead to increase in hydrophobicity of the medium near the 3HC resulting in the decrease of the 3HC fluorescence intensity ratio. The next growth of the Trp fluorescence signal (in the time interval from 1 s to 10 s) results in a marginal change of the 3HC fluorescence (DHU/3HC-substrate). In this time range, the catalytic complex should be formed as the consequence of the fine-tuning of the Nei active site structure. The last change in the Trp and 3HC fluorescence (DHU/3HC-substrate) intensities at times >10 s can be the result of the formation and the dissociation of the enzyme-product complex exposing the dye to a more hydrated environment of the DNA product. The increase of the I_{N^*}/I_{T^*} ratio of the 3HC label supports this idea. On the other hand, the 3HC-DHU duplex exhibited also at least four discernible stages but finished to a non-productive enzyme-substrate complex as supported when analyzing the intensity ratio channel and the product accumulation on PAGE. Non-productive enzyme-substrate complex was also observed with the C^{Py}-DHU substrate indicating that position 5' to DHU is sensitive to substitution with labels and critical to catalysis.

Thereby, comparison of the one-band emissive fluorescent nucleobase analogues (aPu, C^{Py} and tC^O) with the new dual emission label 3HC showed that, the usage of the multichannel 3HC label gave us the possibility to specify more exactly the nature of steps registered previously by Trp fluorescence changes (Fig. 3) [34]. Therefore, it was possible to propose a new kinetic mechanism which includes three reversible steps followed by an

irreversible step and DNA product release, corresponding to: (1) the fast initial DNA binding and formation of nonspecific Nei•DNA complex, (2) DNA bending at lesion site, damaged base eversion from the duplex and insertion of amino acids of the enzyme (Gln-69, Leu-70, and Tyr-71) into the DNA void, (3) adjustment of the active site to catalytically competent state, (4) the hydrolysis of the *N*-glycosidic bond, the elimination of the 3'-phosphate (β -elimination) followed by the release of the 5'-phosphate (δ -elimination), and the last stage (5) that fits with dissociation of the complex between DNA-product and enzyme.

References

1. Lakowicz JR (2006) Principles of fluorescence spectroscopy. New York: Springer.
2. Sinkelldam RW, Greco NJ, Tor Y (2010) Fluorescent analogs of biomolecular building blocks: design, properties, and applications. *Chem Rev* 110: 2579–2619.
3. Wilhelmsson LM (2010) Fluorescent nucleic acid base analogues. *Q Rev Biophys* 43: 159–183.
4. Kim KT, Kim HW, Moon D, Rhee YM, Kim BH (2013) (DNS)C: a fluorescent, environmentally sensitive cytidine derivative for the direct detection of GGG triad sequences. *Org Biomol Chem* 11: 5605–5614.
5. Suzuki A, Takahashi N, Okada Y, Saito I, Nemoto N, et al. (2013) Naphthalene-based environmentally sensitive fluorescent 8-substituted 2'-deoxyadenosines: application to DNA detection. *Bioorg Med Chem Lett* 23: 886–892.
6. Pawar MG, Nuthanakanti A, Srivatsan SG (2013) Heavy Atom Containing Fluorescent Ribonucleoside Analog Probe for the Fluorescence Detection of RNA-Ligand Binding. *Bioconjug Chem*.
7. Pawar MG, Srivatsan SG (2013) Environment-responsive fluorescent nucleoside analogue probe for studying oligonucleotide dynamics in a model cell-like compartment. *J Phys Chem B* 117: 14273–14282.
8. Segal M, Yavin E, Kafri P, Shav-Tal Y, Fischer B (2013) Detection of mRNA of the cyclin D1 breast cancer marker by a novel duplex-DNA probe. *J Med Chem* 56: 4860–4869.
9. Sowers LC, Boulard Y, Fazakerley GV (2000) Multiple structures for the 2'-aminopurine-cytosine mispair. *Biochemistry* 39: 7613–7620.
10. Dunlap CA, Tsai MD (2002) Use of 2-aminopurine and tryptophan fluorescence as probes in kinetic analyses of DNA polymerase β . *Biochemistry* 41: 11226–11235.
11. Stivers JT, Pankiewicz KW, Watanabe KA (1999) Kinetic mechanism of damage site recognition and uracil flipping by *Escherichia coli* uracil DNA glycosylase. *Biochemistry* 38: 952–963.
12. Kuznetsov NA, Koval VV, Zharkov DO, Vorobiev YN, Nevinsky GA, et al. (2007) Kinetic basis of lesion specificity and opposite-base specificity of *Escherichia coli* formamidopyrimidine-DNA glycosylase. *Biochemistry* 46: 424–435.
13. Kuznetsov NA, Koval VV, Nevinsky GA, Douglas KT, Zharkov DO, et al. (2007) Kinetic conformational analysis of human 8-oxoguanine-DNA glycosylase. *J Biol Chem* 282: 1029–1038.
14. Kuznetsova AA, Kuznetsov NA, Ishchenko AA, Saparbaev MK, Fedorova OS (2014) Step-by-Step Mechanism of DNA Damage Recognition by Human 8-Oxoguanine DNA Glycosylase. *Biochim Biophys Acta* 1840: 387–395.
15. Zang H, Fang Q, Pegg AE, Guengerich FP (2005) Kinetic analysis of steps in the repair of damaged DNA by human O6-alkylguanine-DNA alkyltransferase. *J Biol Chem* 280: 30873–30881.
16. Yang K, Stanley RJ (2008) The extent of DNA deformation in DNA photolysis-substrate complexes: a solution state fluorescence study. *Photochem Photobiol* 84: 741–749.
17. Kuznetsov NA, Vorobjev YN, Krasnoperov LN, Fedorova OS (2012) Thermodynamics of the multi-stage DNA lesion recognition and repair by formamidopyrimidine-DNA glycosylase using pyrrolocytosine fluorescence-stopped-flow pre-steady-state kinetics. *Nucleic Acids Res* 40: 7384–7392.
18. Sandin P, Stengel G, Ljungdahl T, Borjesson K, Macao B, et al. (2009) Highly efficient incorporation of the fluorescent nucleotide analogs tC and tCO by Klenow fragment. *Nucleic Acids Res* 37: 3924–3933.
19. Stengel G, Purse BW, Wilhelmsson LM, Urban M, Kuchta RD (2009) Ambivalent incorporation of the fluorescent cytosine analogues tC and tCo by human DNA polymerase alpha and Klenow fragment. *Biochemistry* 48: 7547–7555.
20. Stengel G, Urban M, Purse BW, Kuchta RD (2009) High density labeling of polymerase chain reaction products with the fluorescent base analogue tCo. *Anal Chem* 81: 9079–9085.
21. Rodgers BJ, Elsharif NA, Vashisht N, Mingus MM, Mulvahill MA, et al. (2014) Functionalized tricyclic cytosine analogues provide nucleoside fluorophores with improved photophysical properties and a range of solvent sensitivities. *Chemistry* 20: 2010–2015.
22. Spadafora M, Postupalenko VY, Shvadchak VV, Klymchenko AS, Mely Y, et al. (2009) Efficient Synthesis of Ratiometric Fluorescent Nucleosides Featuring 3'-Hydroxychromone Nucleobases. *Tetrahedron* 65: 7809–7816.
23. Dziuba D, Postupalenko VY, Spadafora M, Klymchenko AS, Guerneau V, et al. (2012) A universal nucleoside with strong two-band switchable fluorescence and sensitivity to the environment for investigating DNA interactions. *J Am Chem Soc* 134: 10209–10213.
24. Demchenko AP (2006) Visualization and sensing of intermolecular interactions with two-color fluorescent probes. *FEBS Lett* 580: 2951–2957.
25. Shynkar VV, Klymchenko AS, Kunzelmann C, Duportail G, Muller CD, et al. (2007) Fluorescent biomembrane probe for ratiometric detection of apoptosis. *J Am Chem Soc* 129: 2187–2193.
26. Das R, Klymchenko AS, Duportail G, Mely Y (2009) Unusually slow proton transfer dynamics of a 3-hydroxychromone dye in protic solvents. *Photochem Photobiol Sci* 8: 1583–1589.
27. Ono T, Wang S, Koo CK, Engstrom L, David SS, et al. (2012) Direct fluorescence monitoring of DNA base excision repair. *Angew Chem Int Ed Engl* 51: 1689–1692.
28. Melamede RJ, Hatahet Z, Kow YW, Ide H, Wallace SS (1994) Isolation and characterization of endonuclease VIII from *Escherichia coli*. *Biochemistry* 33: 1255–1264.
29. Jiang D, Hatahet Z, Melamede RJ, Kow YW, Wallace SS (1997) Characterization of *Escherichia coli* endonuclease VIII. *J Biol Chem* 272: 32230–32239.
30. Golan G, Zharkov DO, Feinberg H, Fernandes AS, Zaika EI, et al. (2005) Structure of the uncomplexed DNA repair enzyme Endonuclease VIII indicates significant interdomain flexibility. *Nucleic Acids Res* 33: 5006–5016.
31. Zharkov DO, Golan G, Gilboa R, Fernandes AS, Gerchman SE, et al. (2002) Structural analysis of an *Escherichia coli* Endonuclease VIII covalent reaction intermediate. *EMBO J* 21: 789–800.
32. Burgess S, Jaruga P, Dodson ML, Dizdaroglu M, Lloyd RS (2002) Determination of active site residues in *Escherichia coli* Endonuclease VIII. *J Biol Chem* 277: 2938–2944.
33. Kropachev KY, Zharkov DO, Grollman AP (2006) Catalytic mechanism of *Escherichia coli* Endonuclease VIII: roles of the intercalation loop and the zinc finger. *Biochemistry* 45: 12039–12049.
34. Kuznetsov NA, Koval VV, Zharkov DO, Fedorova OS (2012) Conformational dynamics of the interaction of *Escherichia coli* endonuclease VIII with DNA substrates. *DNA Repair (Amst)* 11: 884–891.
35. Koval VV, Kuznetsov NA, Zharkov DO, Ishchenko AA, Douglas KT, et al. (2004) Pre-steady-state kinetics shows differences in processing of various DNA lesions by *Escherichia coli* formamidopyrimidine-DNA glycosylase. *Nucleic Acids Res* 32: 926–935.
36. Kuznetsov NA, Koval VV, Zharkov DO, Nevinsky GA, Douglas KT, et al. (2005) Kinetics of substrate recognition and cleavage by human 8-oxoguanine-DNA glycosylase. *Nucleic Acids Res* 33: 3919–3931.
37. Popov AV, Vorob'ev Iu N (2010) GUI-BioPASED program for molecular dynamics modelling of biopolymers with a graphical user interface. *Mol Biol (Mosk)* 44: 735–742.
38. Kanazhevskaya LY, Koval VV, Vorobjev YN, Fedorova OS (2012) Conformational dynamics of abasic DNA upon interactions with AP endonuclease 1 revealed by stopped-flow fluorescence analysis. *Biochemistry* 51: 1306–1321.
39. Liu W, Shin D, Tor Y, Cooperman BS (2013) Monitoring translation with modified mRNAs strategically labeled with isomeric fluorescent guanosine mimetics. *ACS Chem Biol* 8: 2017–2023.
40. Reddy CK, Das A, Jayaram B (2001) Do water molecules mediate protein-DNA recognition? *J Mol Biol* 314: 619–632.
41. Koval VV, Kuznetsov NA, Ishchenko AA, Saparbaev MK, Fedorova OS (2010) Real-time studies of conformational dynamics of the repair enzyme *E. coli* formamidopyrimidine-DNA glycosylase and its DNA complexes during catalytic cycle. *Mutat Res* 685: 3–10.
42. Kuznetsov NA, Zharkov DO, Koval VV, Buckle M, Fedorova OS (2009) Reversible Chemical Step and Rate-Limiting Enzyme Regeneration in the Reaction Catalyzed by Formamidopyrimidine-DNA Glycosylase. *Biochemistry* 48: 11335–11343.

Acknowledgments

We thank Alexander V. Popov for assistance in molecular modeling. We thank Pr. Jean-François Gal for careful reading of the manuscript and fruitful discussion.

Author Contributions

Conceived and designed the experiments: NAK AB OSF. Performed the experiments: AAK NAK YNV. Analyzed the data: NAK AB OSF. Contributed reagents/materials/analysis tools: NPFB BYM AB OSF. Contributed to the writing of the manuscript: AAK NAK NPFB BYM AB OSF.

ANNEXES

ANNEXES

ANNEXES S4

TRAVAUX DE VULGARISATION

ANNEXES

Grands défis technologiques

Détection des modifications de la méthylation de l'ADN et de la chromatine

Benoît Y. Michel^a, Muriel Cuendet^b, Nicolas P. F. Barthes^a, Alain Burger^a, Philippe Bertrand^c, Nadine Martinet^{a,*}

RÉSUMÉ

Les marques épigénétiques non codées dans la séquence de l'ADN sont impliquées aussi bien dans le développement normal que pathologique des êtres vivants parce qu'elles modulent l'expression génique. Transmissibles à la descendance mais réversibles, elles ont généré un immense intérêt thérapeutique et des traitements épigénétiques sont déjà prescrits pour traiter les syndromes myélodysplasiques et certains cancers cutanés. D'autres sont en essais cliniques pendant que des méthodes de plus en plus fines de détections/mesures de ces marques sont constamment améliorées. Pour qu'elles donnent enfin lieu à des dosages cliniques, des essais multicentriques sont réalisés. Ils ont fait la preuve d'une mauvaise reproductibilité inter-laboratoires même si la répétabilité intra-laboratoire est acquise. Cependant des références se mettent en place et demain l'épigénomique sera bientôt analysé comme l'est, aujourd'hui, le génome. Nous décrivons les technologies les plus utilisées permettant de détecter deux marques épigénétiques : la méthylation de la cytosine de l'ADN et la méthylation des histones, ceci chez l'homme et en excluant les méthodes à base de séquençage et de puces traitées par ailleurs.

Méthylation – ADN – histone – détection.

1. La méthylation de la cytosine de l'ADN

La méthylation de la cytosine est la marque épigénétique de l'ADN la plus étudiée chez l'humain [1]. Elle a même retrouvé une seconde vitalité depuis la découverte des oxydases de la famille TET qui oxydent séquentiellement

a Institut de chimie de Nice

Université de Nice Sophia-Antipolis – CNRS UMR 7272
Parc Valrose
06108 Nice cedex 2

b École Section des sciences pharmaceutiques

Université de Genève – Université de Lausanne
30, quai Ernest Ansermet
CH-1211 Genève – Suisse

c Institut de chimie des milieux et matériaux de Poitiers

CNRS UMR 7582
4, rue Michel Brunet – TSA 51106 – B27
86073 Poitiers cedex 9

* Correspondance

nadine.martinet@inserm.fr

article reçu le 1^{er} décembre, accepté le 21 décembre 2014
© 2015 – Elsevier Masson SAS – Tous droits réservés.

SUMMARY

DNA and histone methylation detections

DNA and histone methylation are becoming more and more important to understand the regulation of gene expression. Their detection and measures are leaving the research domain to enter the clinical practice. We describe here the main methods which are used to date excluding sequencing and arrays.

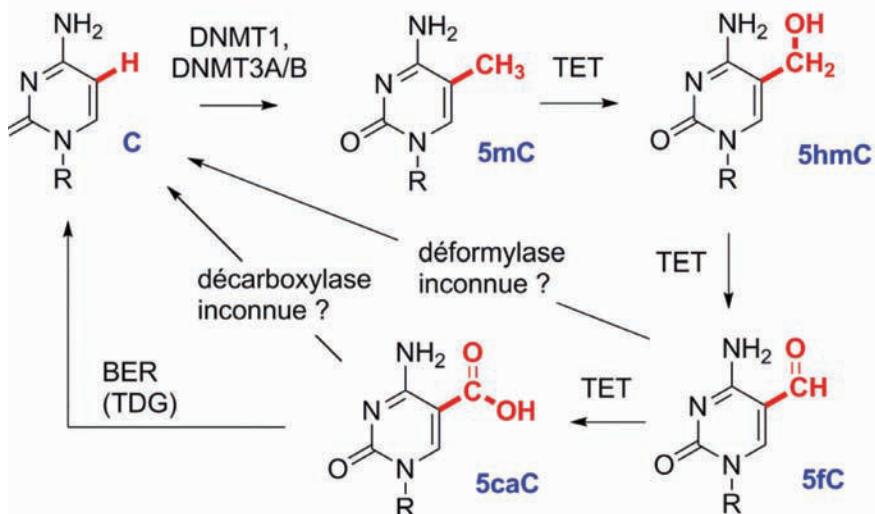
Methylation – DNA – histone – detection.

Abréviations

5mC	: 5-méthylcytosine	LC-MS	: liquid-chromatography-mass spectrometry
5hmC	: 5-hydroxyméthylcytosine	LINES	: long interspersed elements ou retroposons
5fC	: 5-formylcytosine	MBD	: methyl-CpG binding domain
5caC	: 5-carboxylcytosine	MeDIP	: methylated dna immunoprecipitation
ADN	: acide désoxyribonucléique	MGMT	: O-6-methylguanine-DNA methyltransferase
BER	: base excision repair	MS	: mass spectrometry
COBRA	: combined bisulfite restriction analysis	MS-PCR	: methylation-specific PCR
CpG	: désoxycytidine phosphate désoxyguanosine	MS-snupe	: methylation-sensitive single-nucleotide primer extension
DNMT	: DNA methyltransferase	PCR	: polymerase chain reaction
dNTP	: désoxynucléoside triphosphate	RAPD	: random amplified polymorphic DNA
ELISA	: enzyme-linked immuno-sorbent assay	SILAC	: stable isotope labeling by amino acids in cell culture
FRET	: Förster resonance energy transfer	TDG	: thymine-DNA glycosylase
HILIC	: hydrophilic interaction chromatography	TET	: ten-eleven translocation methylocytosine dioxygenase
HP1β	: heterochromatin protein 1		
HPLC	: chromatographie liquide haute performance		
HRM	: high resolving melting		

Dossier scientifique

Figure 1 – Modification de la cytosine chez l'homme.



DNMT: DNA methyltransferase; BER: base excision repair; TDG: thymine-DNA glycosylase, R = 2'-désoxyribosyl.

la nucléobase 5-méthylcytosine (5mC) en la 5-hydroxy-méthylcytosine (5hmC), la 5-formylcytosine (5fC) et la 5-carboxylcytosine (5caC) [2] (**figure 1**).

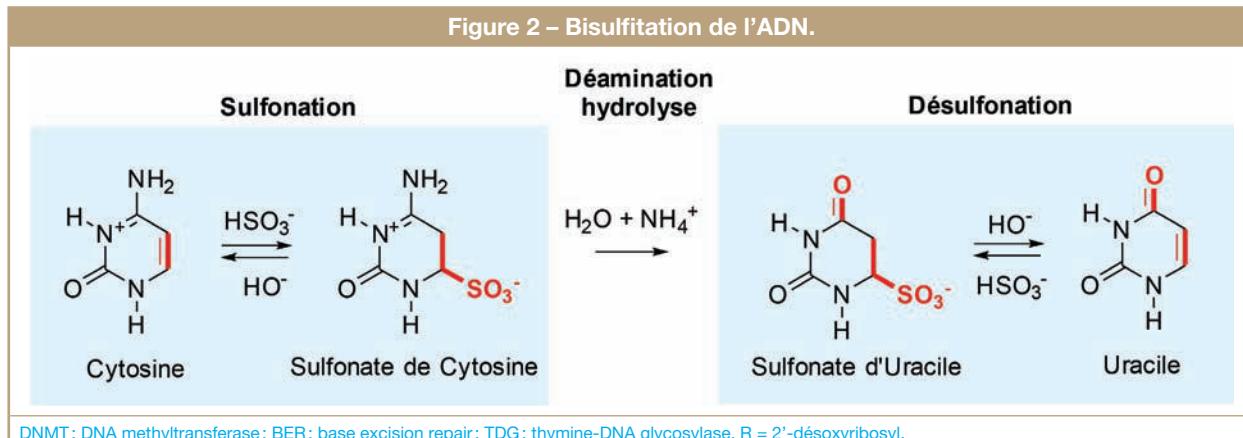
La méthylation de la cytosine, qui résulte de l'action des DNMTs [3], affecte son cycle pyrimidique sur son carbone C5. Elle est le plus souvent trouvée dans des séquences CpG et dans des régions génomiques riches en ces séquences appelées « îlots CpG ». L'addition covalente d'un groupement méthyle sur la cytosine, en créant un encombrement stérique, gêne l'attachement de la machinerie transcriptionnelle et réprime alors, le plus souvent, la transcription des gènes concernés. Dans un génome normal, entre 70-80 % des îlots CpG sont méthylés. Mais les DNMTs peuvent également méthylérer des cytosines hors îlots CpG ; comme c'est le cas dans les cellules souches embryonnaires, les cellules souches nerveuses des adolescents et les cellules corticales des adultes. Dans d'autres tissus, certaines régions géniques sont partiellement méthylées et sont alors associées à une réplication plus tardive. Enfin, il existe des déserts « sans méthylation de l'ADN » qui caractérisent les gènes du développement embryonnaire. Ils sont plutôt marqués par des méthylations (me) des lysines (K) des histones 3 (H3) comme H3K4me3 ou H3K27me3, régulant positivement ou négativement la transcription génique [3, 4].

La méthylation de l'ADN joue un rôle physiologique crucial pour la mise en place de l'empreinte génique ou expression mono allélique des gènes parentaux [3] intervient pour faire varier le répertoire immunologique [3], et empêcher la transcription des éléments répétés de type Alu et LINES [3]. Dans un cadre pathologique, elle conditionne l'expression phénotypique des maladies avec extension de polynucléotides comme l'ataxie de Friedreich et la maladie d'Huntington [5] et éteint les gènes suppresseurs de tumeurs dont elle affecte les promoteurs géniques [3]. C'est sans doute d'ailleurs la première anomalie qui conduit

à la transformation maligne d'une cellule. À ce jour, un seul biomarqueur de méthylation de l'ADN est utilisé en clinique. Il sert à détecter la forme méthylée du promoteur du gène MGMT, qui prédit la réponse au temozolamide, des malades traités pour un GBM [6]. Cette méthylation aberrante affecte une courte séquence d'ADN et elle peut donc être recherchée même dans des tissus fixés où l'ADN est dégradé. Tous les autres biomarqueurs de méthylation de l'ADN n'ont en effet ni démontré leur spécificité ni leur sensibilité bien que de nombreux travaux aient mis en évidence un phénotype « hyper méthylateur » chez certains cancéreux [6].

La sensibilité de diagnostic d'un tel biomarqueur correspond à la proportion d'individus malades qui sont testés positifs pour lui. Sa spécificité de diagnostic tient compte du nombre d'individus normaux qui sont positifs pour ce même marqueur. Enfin, la sensibilité analytique correspond, elle, à la plus petite quantité d'ADN MGMT méthylé détectée au sein d'un ADN et d'une technique donnée. Un biomarqueur idéal possède une bonne sensibilité dans une gamme de valeur étendue. Il est robuste, utilise des réactifs simples et est adapté au plus haut débit possible. Historiquement, la détection de la méthylation de l'ADN repose sur trois principes : (1) elle peut être détectée directement parce qu'elle crée ou fait disparaître un site de restriction enzymatique de l'ADN ; (2) l'ADN méthylé peut être sélectionné/enrichi par immuno précipitation avec des anticorps anti 5mC (ou anti 5hmC) voire des anticorps dirigés contre des protéines reconnaissant spécifiquement l'ADN méthylé ; (3) la bisulfitation de l'ADN permet de « fixer » l'état de méthylation de l'ADN et le rend donc analysable par des techniques bas débit comme la PCR de méthylation spécifique ou MS-PCR et aussi par des technologies haut débit comme le séquençage massif ou les puces dédiées. La détection directe de la 5mC reste encore du domaine de la recherche fondamentale soit parce qu'elle demande un appareillage sophistiqué (spectromètre de masse) pour

Figure 2 – Bisulfitation de l'ADN.



DNMT: DNA methyltransferase; BER: base excision repair; TDG: thymine-DNA glycosylase, R = 2'-désoxyribosyl.

une mesure globale du % de 5mC et de 5hmC présentes dans un ADN donné [7, 8], soit parce qu'elle est encore expérimentale quand elle utilise des nanopores capables de caractériser l'état, *in situ*, de chaque cytosine d'une séquence d'ADN traversant ce nanopore [9, 10].

La bisulfitation de l'ADN permet de transformer, par désamination hydrolytique, une cytosine en uracile (**figure 2**). La 5mC est aussi désaminée en thymine par le bisulfite de sodium mais si lentement que la méthodologie reste valide. Lors de sa bisulfitation, l'ADN est d'abord dénaturé à chaud, en simple brin, pour exposer ses bases à l'action du bisulfite de sodium. La dénaturation est répétée en cycles itératifs dans un thermocycleur pendant quelques secondes. Chaque dénaturation est suivie d'une exposition au bisulfite à chaud (50 °C) pendant 30 s. L'ensemble du procédé, réalisé dans un thermocycleur programmant environ 40 fois ces deux cycles, dure environ 4 h. Une dénaturation alcaline à température ambiante suivie d'une bisulfitation longue à 50 °C reste une alternative possible. Une dénaturation insuffisante donnera naissance, lors de l'analyse par MS-PCR, à une surestimation de la quantité de 5mC due à une bisulfitation insuffisante des cytosines. Le bisulfite désamine toutes les cytosines d'un simple brin d'ADN en créant d'abord un intermédiaire : le 5,6-dihydrocytosine-6 sodium sulfonate (qu'il faudra plus tard désulfonater sinon il sera reconnu pendant la MS-PCR comme une 5mC). L'utilisation d'une solution 6 M de NaI va permettre à la fois de changer le pH de la solution d'ADN contenant le bisulfite acide (des ions sulfites) (pH 5) et de faciliter l'adsorption de l'ADN bisulfité (sulfonaté) sur la silice. L'ADN ainsi adsorbé est ensuite lavé par de l'éthanol pour se débarrasser des sels puis il subit une désulfonation *in situ*, sur la silice, par traitement à la soude avant d'être élué de la silice avec de l'eau. La séquence de l'ADN bisulfité est donc, *in fine*, changée ; ses deux brins ne sont plus complémentaires. Des amorces de PCR complémentaires de l'ADN cible peuvent donc permettre leur amplification par PCR ; à savoir, pour la forme méthylée le brin complémentaire contient une guanine pour s'apparier avec la 5mC, et pour la forme non méthylée, une adénine pour l'uracile = thymine. L'interrogation, par la MS-PCR de chacune des cytosines d'une région de l'ADN donnée est presque toujours possible mais la réponse est alors de type détection sans réelle quantification.

Pourtant compte tenu de l'hétérogénéité cellulaire au sein d'une tumeur ou d'un tissu, une différence de méthylation minime peut tout changer.

Au terme de sa bisulfitation, la dégradation de l'ADN par le bisulfite est constante et atteint 90 % après seulement 1 h de traitement. Elle peut être visualisée après électrophorèse en gel d'agarose contenant du bromure d'éthidium. Pour une meilleure visualisation, le gel est préalablement refroidi dans la glace pendant 30 min. avant irradiation sur table UV. L'ADN bisulfité apparaît avec une longueur maximale d'environ 500 bases. Il est aussi mesurable comme du simple brin d'ADN à 240 nm avec un spectrophotomètre. Pendant le procédé de bisulfitation, les rendements n'atteignent généralement jamais plus de 50 % de l'ADN initial du fait de l'inefficacité de la chromatographie d'adsorption sur silice d'un ADN court et simple brin.

De nombreux facteurs affectent la réaction de bisulfitation de l'ADN. Sulfonation et désamination ont besoin d'un pH acide alors que la désulfonation a besoin d'un milieu alcalin. L'intermédiaire sulfoné est instable et sa formation dépend du pH du milieu réactif, de sa température et de la quantité de bisulfite de sodium utilisée. L'incubation de l'ADN à des températures élevées augmente les dégradations induites par le traitement chimique [11]. Pour en diminuer les effets néfastes, certains ajoutent de l'urée et des anti-oxydants divers mais sans amélioration majeure. Le plus difficile pour l'ensemble du procédé reste d'avoir un vrai contrôle d'une bisulfitation complète ; en effet seul le séquençage prouve l'efficacité d'une bisulfitation. Cette dernière ne permet pas de distinguer la 5hmC de la 5mC. En revanche, des modifications, soit par oxydation chimique ou bien alors par glycosylation enzymatique, permettent cependant d'identifier [12].

La technique de détection de la 5mC la plus sensible analytiquement (environ 0,01 %) et la plus sensible à un locus spécifique est la MS-PCR [13]. Elle utilise deux paires d'amorces complémentaires de la version méthylée (M) et de la version non-méthylée (U) de l'ADN bisulfité considéré. Idéalement, ces amorces encadrent la région à interroger et ne contiennent pas de CG mais ce n'est pas toujours possible compte tenu de la séquence de cet ADN. La technique est facile à mettre en œuvre, économique et spécifique de la séquence d'ADN sondée. La détection des amplicons peut se faire soit par

Dossier scientifique

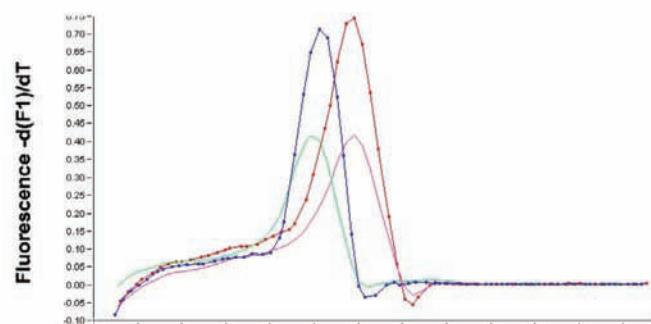
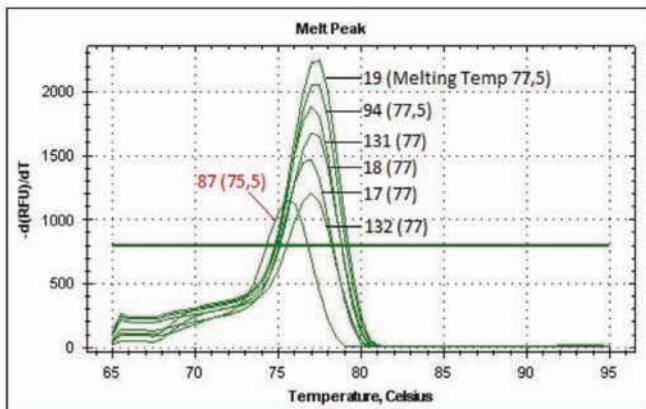
électrophorèse en gel d'agarose, soit par MS-PCR quantitative (qMS-PCR) où en présence de SYBR Green, les amplicons sont détectés lors de leurs fusions thermiques (relargage et capture du SYBR Green lors de la fusion puis de la reconstitution des amplicons). Après bisulfitation, les amplicons M conservant le dinucléotide CG sont reconnus par une amorce GC alors que les amplicons U contenant le dinucléotide UG sont reconnus avec une amorce AG. Les amplicons U contiennent en plus des T des U et sont donc plus riches en ces bases que les amplicons M. Leur température de fusion (T_m) est moins élevée que la température de fusion des amplicons M. Seuls quelques sites CG dans environ 200 pb d'ADN sont interrogables par chaque qMS-PCR : c'est la limite de la technologie. L'existence de méthylation mono-allélique de l'ADN rend la lecture des courbes de fusion encore plus difficile, avec des amplicons de T_m intermédiaires mal résolus (figure 3). Lors des mises au point des qMS-PCR, on tentera autant que possible de choisir des amores ne comportant pas d'unité CG, ou à défaut, en incorporant une ou deux maximum, mais dans ce cas-ci des essais à différentes températures d'appariement des amores sont alors nécessaires. Seules celles générant des amplicons spécifiques devront être utilisées ; des températures élevées vont favoriser les amplifications des amplicons M, c'est-à-dire biaiser la MS-PCR alors que des températures trop basses aboutiront à un manque de spécificité. L'usage de références standardisées comme celles créées par le bureau des standards US devrait permettre, à

l'avenir, une meilleure reproductibilité inter laboratoires [14]. Pour améliorer la lecture des courbes de dénaturation thermique des amplicons après qMS-PCR, il existe une méthode dite HRM [15] qui utilise le SYTO 9 ; un fluorophore utilisable à plus forte dose pour mieux saturer les amplicons sans inhiber la PCR, comme le ferait le SYBR Green à de fortes concentrations. L'usage du SYTO 9 augmente le signal de fluorescence pendant la fusion des amplicons. Il est alors analysé par un logiciel dédié permettant une fusion lente de tous les amplicons et une représentation graphique spécifique des signaux obtenus. La méthode permet alors de détecter entre 0,1 et 1,0 % d'ADN méthylé dans un ADN donné.

Dans une autre variante de qMS-PCR, plus onéreuse : Methylight, des sondes TaqMan reconnaissent chaque type d'amplicons U et M. La quantification des amplicons devient alors possible grâce à l'utilisation d'étalons.

La conception des amores de MS-PCR est simplifiée par l'usage d'un logiciel dédié : MethPrimer, disponible en ligne à l'adresse suivante : <http://www.urogene.org/methprimer/>. Le logiciel est capable de bisulfiter *in silico* (figure 4) une séquence d'ADN donnée. Il est également possible de localiser un îlot CpG dans cette séquence. Il propose aussi des paires d'amores de MS-PCR encadrant la région choisie. Il peut également fournir des amores de séquençage d'ADN bisulfité voire des amores pour une PCR dite de restriction spécifique. Cette dernière recherche le ou les nouveaux sites de restriction

Figure 3 – nMS-PCR avec courbes de fusion des amplicons retinoic acid receptor β , à gauche avec une seule paire d'amores universelles et à droite avec deux paires d'amores différentes U et M.



L'ADN bisulfité provient de cellules MCF7 traitées par des molécules dont on évalue le pouvoir démethylant (19, 94, 131, 18, et 132) avec comme contrôle U : ADN des cellules traitées avec 20 μ M de 5-azacytidine pendant 72 h. Les températures de fusion des amplicons générés sont indiquées entre parenthèses à gauche. Sur le panneau de droite : en vert et en bleu, la 5-azacytidine (amplicons U) et en rouge et en violet l'ADN contrôle des cellules MCF7 cultivées en milieu seul (amplicons M).

Figure 4 – Séquence de l'exon 2 du gène RAR β , bisulfitée avec le logiciel MethPrimer.

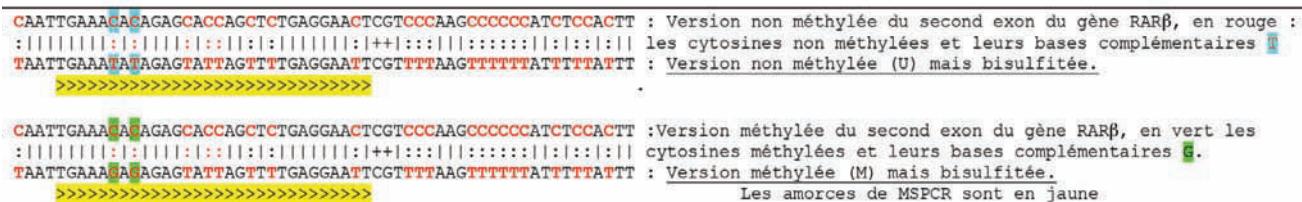


Tableau I – Récapitulatif de certaines méthodes de détection de la 5mC et 5hmC.		
Analyse cytosine par cytosine	Avantages	Inconvénients
MS-PCR (SYBR Green) sensibilité: 10^{-5} à 10^{-2}	Peu d'ADN bisulfité, facile et haut débit	Site par site mais biais de la MS-PCR
MethylLight (sondes Taqman spécifiques)	Peu d'ADN bisulfité, facile et haut débit	Onéreuse, site par site et biais de la MS-PCR
HRM (SYTO 5 + courbe de fusion des amplicons)	Peu d'ADN bisulfité, facile et moyen débit	Aspécifique si plusieurs C dans les amplicons
Cobra (restriction d'ADN bisulfité amplifié)	Facile, bas débit et peu onéreuse	Quelques CpG (amorces encadrant la région)
Clivage spécifique de base + spectrométrie de masse	Plus d'ADN et plus long	Appareillage onéreux
HeavyMethyl	Facile et moyen débit	Beaucoup d'oligonucléotides bloqueurs
MS-SNP extension	Peu d'ADN	Radioactivité
Analyse globale	Avantages	Inconvénients
LC-MS	Sensibilité de 0,01 % : 5mC et 50 % 5hmC	Au moins 3 µg d'ADN clivé (en %)
Immuno-histochimie	Simple (idem pour 5hmC)	Dénaturation de l'ADN nécessaire
Elisa	Simple (idem pour 5hmC)	Au moins 3 µg d'ADN clivé

enzymatique créés après modification de la séquence générée par la bisulfitation. Des résultats graphiques illuminant la séquence ciblée sont également obtenus [16]. Un forum de discussion en ligne permet aux internautes de poser des questions techniques à la communauté des étudiants de la question.

1.1. COBRA [17]

La digestion d'un amplicon, généré à partir d'un ADN bisulfité, par certaines enzymes de restriction est utilisée pour distinguer un ADN méthylique d'un ADN non-méthylique. En effet, la séquence de l'ADN bisulfité peut permettre soit le maintien soit l'apparition d'un nouveau site de restriction enzymatique dans ces amplicons. Ceux-ci sont alors séparés par électrophorèse sur gel d'agarose ou d'acrylamide suivie d'une hybridation quantitative. Malheureusement, de nombreuses séquences d'amplicons ne sont pas envisageables via cette technologie du fait de leur séquence. En plus du biais inhérent à toute PCR, la quantification peut être gênée par la formation d'hétéroduplexes entre des brins d'ADN qui contiennent un site de restriction et les brins mal bisulfités. Le coût de cette technologie reste peu élevé même si elle requiert de nombreuses heures de manipulation. *MspI* et *ApeKI* sont des enzymes de restriction insensibles à la méthylation de la cytosine. L'usage de *MspI* seule couvre 10 % des îlots CpGs des cellules humaines alors que la double digestion en couvre 20 %.

1.2. MS-snup [18]

À l'origine, cette technologie fut développée pour rechercher des mutations ponctuelles de l'ADN. Elle a été ensuite étendue à l'étude des 5mC. Le produit généré est isolé par électrophorèse puis hybridé à une amorce interne qui se termine immédiatement en 5' du brin d'ADN à étudier. L'amorce est ensuite allongée avec une ADN polymérase en présence de dNTP contenant de la dCTP ou dTTP marquées avec du ^{32}P . La quantité relative des amplicons sera mesurée par un système d'analyse radiographique. Plusieurs amorces sont utilisables pour un multiplexage qui permet d'augmenter le débit analytique. Le choix

des amorces non complémentaires de la séquence sauvage est primordial. La technique est chronophage et a l'inconvénient de l'usage de la radioactivité. Elle reste délicate à appliquer pour l'analyse de régions denses en îlots CpG. La technologie SNaPshot d'Applied Biosystems permet de se passer du marquage radioactif. Une autre variante utilise l'HPLC pour séparer et quantifier les amplicons obtenus.

1.3. Methylation sensitive-random amplified polymorphic DNA-polymerase chain [19]

Cette technique est basée sur l'amplification sélective de séquences génomiques qui sont flanquées, par chance, de séquences reconnues par une amorce d'amplification choisie arbitrairement. Lorsque l'on compare de l'ADN sauvage et de l'ADN bisulfité on observera des amplicons différents qui vont indiquer la présence de 5mC. Une alternative consiste à hybrider les amplicons sur une puce dédiée.

1.4. HeavyMethyl [20]

Cette technique utilise des oligonucléotides capables de bloquer l'amplification de la forme non-méthylique de l'ADN bisulfité d'un site à étudier, avec une sonde TaqMan, pour identifier les amplicons M. Cette technique simple à mettre en œuvre a été très utilisée (*tableau I*).

1.5. Détection de la 5mC *in situ*

L'anticorps monoclonal 33D3 qui reconnaît la 5mC est bien caractérisé mais d'autres ont été mis au point pour les autres dérivés de la cytosine. L'immuno-histochimie permettant la détection des 5mC est très peu utilisée car ce processus pose problème lors de la dénaturation de l'ADN *in situ* nécessaire à la fixation des tissus, processus détruisant la chromatine et affectant aussi l'ADN. Des sondes fluorescentes (green fluorescent protein) pour l'observation 3D des cellules vivantes ont été fabriquées en construisant un transgène fluorescent (domaine MDB fusionné à un domaine de localisation nucléaire de la même protéine), reconnaissant l'ADN méthylique. Les animaux transgéniques ainsi créés ont permis d'obtenir

Dossier scientifique

des images dynamiques de mise en place de la méthylation de l'ADN dans l'embryon. La technologie a été améliorée récemment, mais reste encore du domaine de la recherche [21].

1.6. Enrichissement en ADN méthylé [22]

L'immuno-précipitation MeDIP utilisant l'anticorps monoclonal 33D3 et la capture de l'ADN méthylé par des anticorps anti protéines spécialisées capables de reconnaître et de lier spécifiquement les protéines possédant des domaines MBD (technologie MBDCap (capture)) sont les techniques les plus économiques et les plus efficaces pour détecter la méthylation globale ou localisée de l'ADN par rapport à des références. L'ADN méthylé ainsi enrichi peut servir pour des PCR ou des hybridations sur puces dédiées.

1.7. Détection directe de la 5mC

La détection directe de la 5mC est possible pour des nucléotides en solution par ELISA, par chromatographie sur couche mince (CCM), par HPLC, par spectrométrie de masse et par ampérométrie de nananopore. Cette dernière technologie n'est toutefois pas encore utilisable en routine. L'ELISA est facile à mettre en œuvre et donne une bonne détection globale du taux de 5mC présent dans un ADN donné, cependant sa spécificité pour les seules 5mC et non pas les 5hmC n'est pas démontrée. Les autres techniques ne sont pas adaptées au haut débit.

1.8. La LC-MS/MS [23, 24]

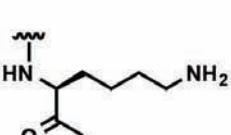
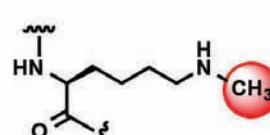
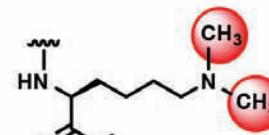
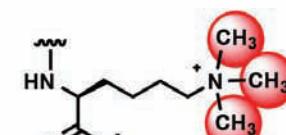
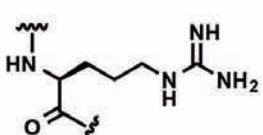
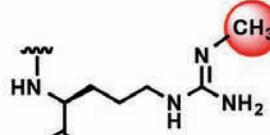
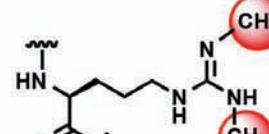
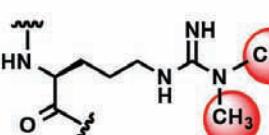
Des essais permettant de détecter simultanément les quantités globales de 5mC et la 5hmC dans des hydrolysats d'ADN, en utilisant des références marquées avec un isotope non radioactif comme étalons externes et l'extraction en phase solide permettent d'atteindre des limites de détection de 0,01 et 0,02 pg.

L'hydrolyse de l'ADN peut être enzymatique mais également purement chimique (88 % d'acide formique à 140 °C). Après chromatographie des hydrolysats d'ADN par HILIC sur une colonne bridged ethylene hybrid, la cytidine et ses dérivés sont analysés par spectrométrie de masse. Les limites de détection, sans radioactivité, sont de 1 pg/mL pour la C, 45 pg/mL pour la 5mC et 57 pg/mL pour la 5hmC. La variabilité inter mesure varie entre 1,87 % et 4,84 %. Des références commerciales calibrent les expériences et la répétabilité inter laboratoires. Malheureusement, le coût de l'appareillage confine ces techniques à la recherche.

2. Détection de la méthylation des histones [25-31]

Les lysines méthyltransférases (environ une cinquantaine d'enzymes humaines) catalysent la méthylation des résidus de lysine en utilisant la S-adénosyl-L-méthionine comme donneur de groupement méthyle (*tableau II*). L'amine de la chaîne latérale de la lysine peut être mono-, di- ou tri-méthylée.

Tableau II – Méthylation des histones (en haut les méthylations de la lysine et en bas celles de l'arginine).

Non-méthylée	Mono-méthylée	Bi-méthylée	Tri-méthylée
			
		 Symétrique	 Asymétrique

Cette modification affecte surtout les lysines des histones H3 (K4, K9, K27 et K36) et H4 (K20), mais aussi celles d'autres protéines comme p53 en lui conférant une meilleure stabilité. En 2012, et en l'absence de techniques d'enrichissement spécifiques, 169 lysines méthylées dans 72 protéines humaines différentes furent répertoriées. Les résidus d'arginine des histones mais aussi ceux d'autres protéines sont aussi mono- ou di-méthylés au niveau de l'azote de la fonction guanidinium et ce de manière symétrique ou asymétrique par les arginines méthyltransférases : 263 résidus d'arginines méthylés dans 92 protéines différentes ont été identifiées en 2012 [29]. Des anticorps très spécifiques des marques de méthylation des histones sont très souvent utilisés en Western blot et en immuno-précipitation, mais il faut souligner que leur spécificité reste discutée si bien que des méthodes d'enrichissement comme pour l'ADN méthylé sont aussi employées en utilisant des protéines spécialisées. Par exemple, les protéines à chromodomains reconnaissent les lysines di- et triméthylées des histones H3, alors que les protéines à domaine Tudor répété reconnaissent les lysines mono-, di- et tri-méthylées. La protéine de l'hétéro chromatine 1 β ou HP1 β possède un chromo domaine qui reconnaît spécifiquement et lie avec une forte affinité : H3K9me3. Ce domaine a servi d'accroche pour identifier un interactome humain de 109 protéines pour HP1 β dont 29 présentent 40 résidus lysine différemment méthylés. Pour déterminer la présence de ces marques de méthylation, les techniques utilisant des anticorps dédiés et le Western blot restent très employées mais d'autres, basées sur la spectrométrie de masse, ont été également développées. Du fait du coût des appareillages nécessaires pour ces dernières, elles restent des outils dédiés à la recherche. Il est en plus possible de marquer les acides aminés donneurs de groupes méthyles par des isotopes

stables, comme le deutérium, pour détecter ou mesurer alors la méthylation des protéines de cellules cultivées dans des conditions précises (SILAC).

Du fait de leur pH basique, les histones sont habituellement extraites des tissus en utilisant des conditions acides puis séparées par électrophorèse en deux dimensions avant d'être identifiées puis excisées pour être digérées par des protéases avant analyse par spectrométrie de masse. Les peptides résultants sont ensuite séparés par chromatographie en phase inverse sur colonne de silice greffée (C18) et sont directement injectés dans un spectromètre de masse qui mesure le rapport masse/charge (*m/z*) [31]. Il est aussi possible de raffiner la technologie en isolant les ions les plus nombreux pour générer des fragments analysables en MS/MS.

Récemment des sondes FRET permettant de rendre visible, *in vivo*, l'acétylation de H4 ont été développées dans des cellules transgéniques voire dans des animaux transgéniques afin de suivre le devenir de cellules souches épidermiques. Cependant, ces techniques restent encore du domaine de la recherche.

3. Conclusion

Nous avons ainsi présenté les principales technologies qui permettent de détecter la méthylation de l'ADN et des histones chez l'homme. Ces techniques nécessitent encore bien des progrès avant de devenir des outils cliniques mais il est clair qu'elles nous permettent déjà de mieux comprendre déjà l'expression génique.

Déclaration d'intérêts : les auteurs déclarent ne pas avoir de conflits d'intérêts en relation avec cet article.

Ce travail a reçu le soutien de l'ANR (12-BS08-0003-02) et de la FRM (DCM20111223038).

Références

- [1] Jones PA. The DNA methylation paradox. *Trends Genet* 1999;15(1):34-7.
- [2] Kriaucionis S., Tahiliani M. Expanding the epigenetic landscape: novel modifications of cytosine in genomic DNA. *Cold Spring Harb Perspect Biol* 2014;6(10):a018630.
- [3] Martinet N, Michel BY, et al. Small molecules DNA methyltransferases inhibitors. *Med Chem Commun* 2012;3:263-73.
- [4] Liyanage VR, Jarmasz JS, et al. DNA modifications: function and applications in normal and disease States. *Biology (Basel)* 2014;3(4):670-723.
- [5] Lee J, Jang SJ, et al. Presence of 5-methylcytosine in CpNpG trinucleotides in the human genome. *Genomics* 2010;96(2):67-72.
- [6] Jain S, Wojdacz TK, et al. Challenges for the application of DNA methylation biomarkers in molecular diagnostic testing for cancer. *Expert Rev Mol Diagn* 2013;13(3):283-94.
- [7] Le T, Kim KP, et al. A sensitive mass spectrometry method for simultaneous quantification of DNA methylation and hydroxymethylation levels in biological samples. *Anal Biochem* 2011;412(2):203-9.
- [8] Hu CW, Chen JL, et al. Trace analysis of methylated and hydroxymethylated cytosines in DNA by isotope-dilution LC-MS/MS: first evidence of DNA methylation in *Caenorhabditis elegans*. *Biochem J* 2014;465:39-47.
- [9] Laszlo AH, Derrington IM, et al. Detection and mapping of 5-methylcytosine and 5-hydroxymethylcytosine with nanopore MspA. *Proc Natl Acad Sci U S A* 2013;110(47):18904-9.
- [10] Koo KM, Wee EJ, et al. Microdevices for detecting locus-specific DNA methylation at CpG resolution. *Biosens Bioelectron* 2014;56:278-85.
- [11] Piperi C, Farmaki E, et al. DNA methylation signature analysis: how easy is it to perform? *J Biomol Tech* 2008;19(5):281-4.
- [12] Kriukiene E, Liutkeviciute Z, et al. 5-Hydroxymethylcytosine – the elusive epigenetic mark in mammalian DNA. *Chem Soc Rev* 2012;41(21):6916-30.
- [13] Herman JG, Graff JR, et al. Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. *Proc Natl Acad Sci U S A* 1996;93(18): 821-6.
- [14] Burke DG, Griffiths K, et al. Accurate measurement of DNA methylation that is traceable to the international system of units. *Anal Chem* 2009;81(17):7294-7301.
- [15] Wojdacz TK. Methylation-sensitive high-resolution melting in the context of legislative requirements for validation of analytical procedures for diagnostic applications. *Expert Rev Mol Diagn* 2012;12(1):39-47.
- [16] Li LC, Dahiya R. MethPrimer: designing primers for methylation PCRs. *Bioinformatics* 2002;18(11):1427-31.

- [17] Xiong Z, Laird PW. COBRA: a sensitive and quantitative DNA methylation assay. *Nucleic Acids Res* 1997;25(12):2532-4.
- [18] Gonzalgo ML, Jones PA. Quantitative methylation analysis using methylation-sensitive single-nucleotide primer extension (Ms-SNuPE). *Methods* 2002;27(2):128-33.
- [19] Singh KP. Screening of DNA methylation changes by methylation-sensitive random amplified polymorphic DNA-polymerase chain reaction (MS-RAPD-PCR). *Methods Mol Biol* 2014;1105:71-81.
- [20] Distler J. Quantification of methylated DNA by HeavyMethyl duplex PCR. *Methods Mol Biol* 2009;507: 339-346.
- [21] Yamagata K. DNA methylation profiling using live-cell imaging. *Methods* 2010;52(3):259-66.
- [22] Hsu HK, Weng YI, et al. Detection of DNA methylation by MeDIP and MBDCap assays: an overview of techniques. *Methods Mol Biol* 2014;1105:61-70.
- [23] Zhang L, Zhou K, et al. Simultaneous determination of global DNA methylation and hydroxymethylation levels by hydrophilic interaction liquid chromatography-tandem mass spectrometry. *J Biomol Screen* 2012;17(7):877-84.
- [24] Ehrich M, Nelson MR, et al. Quantitative high-throughput analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry. *Proc Natl Acad Sci U S A* 2005;102(44):15785-90.
- [25] Peters AH, Kubicek S, et al. Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. *Mol Cell* 2003;12(6):1577-89.
- [26] Sasaki K, Ito A, et al. Development of live-cell imaging probes for monitoring histone modifications. *Bioorg Med Chem* 2012;20(6):1887-92.
- [27] Qian J, Lu L, et al. Development of multiple cell-based assays for the detection of histone H3 Lys27 trimethylation (H3K27me3). *Assay Drug Dev Technol* 2013;11(7):449-56.
- [28] Carlson SM, Gozani O. Emerging technologies to map the protein methylome. *J Mol Biol* 2014;426(20):3350-62.
- [29] Lin S, Wein S, et al. Stable-isotope-labeled histone peptide library for histone post-translational modification and variant quantification by mass spectrometry. *Mol Cell Proteomics* 2014;13(9):2450-66.
- [30] Soldi M, Bremang M, et al. Biochemical systems approaches for the analysis of histone modification readout. *Biochim Biophys Acta* 2014;1839(8):657-68.
- [31] Zhang K, Li L, et al. Comparative analysis of histone H3 and H4 post-translational modifications of esophageal squamous cell carcinoma with different invasive capabilities. *J Proteomics* 2014;112C: 180-9.



REFERENCES

REFERENCES

- ¹ a) Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., Altman, S., *Cell*, **1983**, *35*, 849–857; b) Nissen, P., Hansen, J., Ban, N., Moore, P. B., Steitz, T. A., *Science*, **2000**, *289*, 920–930.
- ² a) Cohn, W. E., *J. Biol. Chem.*, **1960**, *235*, 1488–1498; b) Charette, M., Gray, M. W., *IUBMB Life*, **2000**, *49*, 341–351; c) T. Hamma, A. R. Ferré-D'Amaré, *Chem. Biol.*, **2006**, *13*, 1125–1135.
- ³ Itaya, T., Kanai, T., Sawada, T., *Chem. Pharm. Bull.*, **2002**, *50*, 547–548,
- ⁴ Ho, P. S., Carter, M., *DNA Replication-Current Advances*. **2011**, Chap 1.
- ⁵ Watson, J. D., Crick, F. H. C., *Nature*, **1953**, *171*, 737–738
- ⁶ a) Belmont, P., Constant, J.-F., Demeunynck, M., *Chem. Soc. Rev.*, **2001**, *30*, 70–81; b) https://commons.wikimedia.org/wiki/File:A-B-Z-DNA_Side_View.png.
- ⁷ Booth, M. J., Raiber, E.-A., Balasubramanian, S., *Chem. Rev.*, **2015**, *115*, 2240–2254.
- ⁸ Choi, J., Majima, T., *Chem. Soc. Rev.*, **2011**, *40*, 5893–5909.
- ⁹ Margulis, L., *Symp. Soc. Exp. Biol.*, **1975**, *29*, 21–38.
- ¹⁰ Cronk, Q. C., *Nat. Rev. Genet.*, **2001**, *2*, 607–619.
- ¹¹ Stein, L. D., *Nature*, **2004**, *431*, 915–916.
- ¹² a) Yin, Y. W., Steitz, T. A., *Science*, **2002**, *298*, 1387–1395; b) Clancy, S., *Nature Education*, **2008**, *1*, 41.
- ¹³ a) https://www.weizmann.ac.il/sb/faculty_pages/Yonath/; b) <https://www.youtube.com/watch?v=Jml8CFBWcDs>
- ¹⁴ a) Pelletier, H., Sawaya, M. R., Wolfle, W., Wilson, S. H., Kraut, J., *Biochemistry*, **1996**, *35*, 12762–12777; b) Meijer, L., *Oncologie*, **2003**, *5*, 311–326; c) Georgescu, R. E., Kim, S.-S., Yurieva, O., Kuriyan, J., Kong, X.-P., O'Donnell, M., *Cell*, **2008**, *132*, 43–54; d) <http://www.medical-institution.com/dna-replication-animation/>
- ¹⁵ Waddington, C. H., *Nature*, **1942**, *150*, 563–565.
- ¹⁶ Marx, V., *Nature*, **2012**, *491*, 143–147.
- ¹⁷ Davey, C. A., Sargent, D. F., Luger, K., Maeder, A. W., Richmond, T. J., *J. Mol. Biol.*, **2002**, *319*, 1097–1113.
- ¹⁸ Thoma, F., Koller, T., Klug, A., *J. Cell Biol.*, **1979**, *83*, 403–427.
- ¹⁹ Brooks, W. H., Le Dantec, C., Pers, J.-O., Youinou, P., Renaudineau, Y., *J. Autoimmun.*, **2010**, *34*, J207–J219.
- ²⁰ Le Dantec, C., Chevailler, A., Renaudineau, Y., *Revue francophone des laboratoires*, **2015**, *2015*, 71–75.
- ²¹ Liu, L., Souto, J., Liao, W., Jiang, Y., Li, Y., Nishinakamura, R., Huang, S., Rosengart, T., Yang, V. W., Schuster, M., Ma, Y., Yang, J., *J. Biol. Chem.*, **2013**, *288*, 34719–34728.
- ²² Ray-Gallet, D., Gérard, A., Polo, S., Almouzni, G., *Médecine Sciences*, **2005**, *21*, 384–389.
- ²³ a) Sun, W., Guan, M., Li, X., *Stem Cells Dev.*, **2014**, *23*, 923–930; b) Jeltsch, A., Jurkowska, R. Z., *Trends Biochem. Sci.*, **2014**, *39*, 310–318.
- ²⁴ a) Weber, M., Davies, J. J., Wittig, D., Oakeley, E. J., Haase, M., Lam, W. L., Schübeler, D., *Nat. Genet.*, **2005**, *37*, 853–862; b) Auclair, G., Weber, M., *Biochimie*, **2012**, *94*, 2202–2211.
- ²⁵ a) Yisraeli, J., Frank, D., Razin, A., Cedar, H., *Proc. Natl. Acad. Sci.*, **1988**, *85*, 4638–4642; b) Cartron, P.-F., Pacaud, R., Salbert, G., *Revue francophone des laboratoires*, **2015**, *2015*, 37–48; c) Michel, B. Y., Cuendet, M., Barthes, N. P. F., Burger, A., Bertrand, P., Martinet, N., *Revue francophone des laboratoires*, **2015**, *2015*, 55–62.
- ²⁶ a) Laird, P. W., Jaenisch, R., *Hum. Mol. Genet.*, **2003**, *3*, 1487–1495.; b) Deltour, S., Chopin, V., and Leprince, D., *Médecine Sciences*, **2005**, *21*, 405–411; c) Rottach, A., Leonhardt, H., Spada, F., *J. Cell. Biochem.*, **2009**, *108*, 43–51.
- ²⁷ Li, E., Bestor, T. H., Jaenisch, R., *Cell*, **1992**, *69*, 915–926.
- ²⁸ Gabory, A., Dandolo, L., *Médecine Sciences*, **2005**, *21*, 390–395.
- ²⁹ Jaenisch, R., Bird, A., *Nat. Genet.*, **2003**, *33*, 245–254.
- ³⁰ Shiota, K., Yanagimachi, R., *Differentiation*, **2002**, *69*, 162–166.

REFERENCES

- ³¹ a) Ratel, D., Ravanat, J.-L., Charles, M.-P., Platet, N., Breuillaud, L., Lunardi, J., Berger, F., Wion, D., *FEBS Lett.*, **2006**, *580*, 3179–3184; b) López, C. M. R., Lloyd, A. J., Leonard, K., Wilkinson, M. J., *Anal. Chem.*, **2012**, *84*, 7336–7342; c) Huang, W., Xiong, J., Yang, Y., Liu, S.-M., Yuan, B.-F., Feng, Y.-Q., *RSC Advances*, **2015**, *5*, 64046–64054; d) Greer, E. L., Blanco, M. A., Gu, L., Sendinc, E., Liu, J., Aristizábal-Corrales, D., Hsu, C.-H., Aravind, L., He, C., Shi, Y., *Cell*, **2015**, *161*, 868–878; e) Summerer, D., *Angew. Chem. Int. Ed.*, **2015**, *in press*. DOI: 10.1002/anie.201504594.
- ³² Chentsova, A., Kapourani, E., Giannis, A., *Beilstein J. Org. Chem.*, **2014**, *10*, 7–11.
- ³³ Jones, P. A., Liang, G., *Nat. Rev. Genet.*, **2009**, *10*, 805–811.
- ³⁴ Williams, K., Christensen, J., Helin, K., *EMBO Rep.*, **2012**, *13*, 28–35.
- ³⁵ Maiti, A., Michelson, A. Z., Armwood, C. J., Lee, J. K., Drohat, A. C., *J. Am. Chem. Soc.*, **2013**, *135*, 15813–15822.
- ³⁶ Renaudineau, Y., *Revue francophone des laboratoires*, **2015**, *2015*, 35–36.
- ³⁷ a) Savagner, F., Le Pennec, S., Rivalin, R., Eyer, J., *Revue francophone des laboratoires*, **2015**, *2015*, 49–54; b) Lopez-Serra, P., Esteller, M., *Oncogene*, **2012**, *31*, 1609–1622.
- ³⁸ <http://www.nature.com/nrg/multimedia/rnai/animation/index.html>
- ³⁹ Daxinger, L., Whitelaw, E., *Nat. Rev. Genet.*, **2012**, *13*, 153–162.
- ⁴⁰ Reik, W., Walter, J., *Nat. Rev. Genet.*, **2001**, *2*, 21–32.
- ⁴¹ Bird, A., *Genes Dev.*, **2002**, *16*, 6–21.
- ⁴² Martinet, N., Michel, B. Y., Bertrand, P., Benhida, R., *MedChemComm*, **2012**, *3*, 263–273.
- ⁴³ Das, P. M., Singal, R., *J. Clin. Oncol.*, **2004**, *22*, 4632–4642.
- ⁴⁴ Howell, C. Y., Bestor, T. H., Ding, F., Latham, K. E., Mertineit, C., Trasler, J. M., Chaillet, J. R., *Cell*, **2001**, *104*, 829–838.
- ⁴⁵ Robert, M. F., Morin, S., Beaulieu, N., Gauthier, F., Chute, I. C., *Nat. Genet.*, **2003**, *33*, 61–65.
- ⁴⁶ a) Song, J., Teplova, M., Ishibe-Murakami, S., Patel, D. J., *Science*, **2012**, *335*, 709–712; b) Song, J., Rechkoblit, O., Bestor, T. H., Patel, D. J., *Science*, **2011**, *331*, 1036–1040.
- ⁴⁷ Wojciechowski, M., Czapinska, H., Bochtler, M., *Proc. Natl. Acad. Sci.*, **2013**, *110*, 105–110.
- ⁴⁸ a) Avvakumov, G. V., Walker, J. R., Xue, S., Li, Y., Duan, S., Bronner, C., Arrowsmith, C. H., Dhe-Paganon, S., *Nature*, **2008**, *455*, 822–825; b) Bronner, C., Achour, M., Arima, Y., Chataigneau, T., Saya, H., Schini-Kerth, V. B., *Pharmacol. Ther.*, **2007**, *115*, 419–434.
- ⁴⁹ a) Bostick, M., Kim, J. K., Esteve, P. O., Clark, A., Pradhan, S., Jacobsen, S. E., *Science*, **2007**, *317*, 1760–1764; b) Arita, K., Ariyoshi, M., Tochio, H., Nakamura, Y., Shirakawa, M., *Nature*, **2008**, *455*, 818–821; c) Hashimoto, H., Horton, J. R., Zhang, X., Bostick, M., Jacobsen, S. E., Cheng, X., *Nature*, **2008**, *455*, 826–829.
- ⁵⁰ Hashimoto, H., Horton, J. R., Zhang, X., Cheng, X., *Epigenetics*, **2009**, *4*, 8–14.
- ⁵¹ Iida, T., Suetake, I., Tajima, S., Morioka, H., Ohta, S., Obuse, C., Tsurimoto, T., *Genes Cells*, **2002**, *7*, 997–1007.
- ⁵² a) Wolff, G. L., Kodell, R. L., Moore, S. R., Cooney, C. A., *FASEB J.*, **1998**, *12*, 949–957; b) Laget, S., Defossez, P. A., *Medecine Sciences*, **2008**, *24*, 724–730.
- ⁵³ Anand, P., Kunnumakkara, A. B., Kunnumakkara, A. B., Sundaram, C., Harikumar, K. B., Tharakan, S. T., Lai, O. S., Sung, B., Aggarwal, B. B., *Pharm. Res.*, **2008**, *25*, 2097–2116.
- ⁵⁴ a) Szyf, M., *Drug Resistance Updates*, **2003**, *6*, 341–353; b) Kristensen, L. S., Nielsen, H. M., Hansen, L. L., *Eur. J. Pharmacol.*, **2009**, *625*, 131–142.
- ⁵⁵ Fernandez, A. F., Assenov, Y., Martin-Subero, J. I., Balint, B., Siebert, R., Taniguchi, H., Yamamoto, H., Hidalgo, M., Tan, A. C., Galm, O., Ferrer, I., Sanchez-Cespedes, M., Villanueva, A., Carmona, J., Sanchez-Mut, J. V., Berdasco, M., Moreno, V., Capella, G., Monk, D., Ballestar, E., Ropero, S., Martinez, R., Sanchez-Carbayo, M., Prosper, F., Agirre, X., Fraga, M. F., Grana, O., Perez-Jurado, L., Mora, J., Puig, S., Prat, J., Badimon, L., Puca, A. A., Meltzer, S. J., Lengauer, T., Bridgewater, J., Bock, C., Esteller, M., *Genome research*, **2012**, *22*, 407–419.
- ⁵⁶ Ehrlich, M., Lacey, M., *DNA Hypomethylation and Hemimethylation in Cancer*, in *Epigenetic Alterations in Oncogenesis* (Karpf, A. R., Ed.), Springer : New York, pp 31–56, **2012**.

REFERENCES

- ⁵⁷ Crews, D., *Evol. Dev.*, **2003**, *5*, 50–55.
- ⁵⁸ Jones, P. A., Baylin, S. B., *Nat. Rev. Genet.*, **2002**, *3*, 415–428.
- ⁵⁹ a) Furst, A. L., Barton, J. K., *Chemistry & Biology*, **2015**, *22*, 938–945; b) Nuñez, N. N., Manlove, A. H., David, S. S., *Chemistry & Biology*, **2015**, *22*, 810–811.
- ⁶⁰ Mudbhary, R., Hoshida, Y., Chernyavskaya, Y., Jacob, V., Villanueva, A., Fiel, M. I., Chen, X., Kojima, K., Thung, S., Bronson, R. T., Lachenmayer, A., Revill, K., Alsinet, C., Sachidanandam, R., Desai, A., SenBanerjee, S., Ukomadu, C., Llovet, J. M., Sadler, K. C., *Cancer Cell*, **2014**, *25*, 196–209.
- ⁶¹ Rhee, I., Bachman, K. E., Ben Ho Park, Jair, K.-W., Yen, R.-W. C., Schuebel, K. E., Cui, H., Feinberg, A. P., Lengauer, C., Kinzler, K. W., Baylin, S. B., Vogelstein, B., *Nature*, **2002**, *416*, 552–556.
- ⁶² Fournel, M., Sapieha, P., Beaulieu, N., Besterman, J. M., MacLeod, A. R. *J. Biol. Chem.*, **1999**, *274*, 24250–24256.
- ⁶³ Robaina, M. C., Mazzoccoli, L., Arruda, V. O., Reis, F. R. de S., Apa, A. G., de Rezende, L. M. M., Klumb, C. E., *Exp. Mol. Pathol.*, **2015**, *98*, 200–207.
- ⁶⁴ a) Yan, F., Wang, X., Shao, L., Ge, M., Hu, X., *Tumour Biol.*, **2015**, *in press*. DOI: 10.1007/s13277-015-3638-1; b) Zhu, M., Xu, Y., Ge, M., Gui, Z., Yan, F. *Cancer Sci.*, **2015**, *106*, 833–839; c) Zhou, L., Shang, Y., Jin, Z., Zhang, W., Lv, C., Zhao, X., Liu, Y., Li, N., Liang, J., *Cancer Biol. Ther.*, **2015**, *16*, 1241–1251.
- ⁶⁵ Unoki, M., Brunet, J., Mousli, M., *Biochem. Pharmacol.*, **2009**, *78*, 1279–1288.
- ⁶⁶ n/a, *Cancer Discov.*, **2014**, *4*, OF9–OF9.
- ⁶⁷ a) Hervouet, E., Lalier, L., Debien, E., Cheray, M., Geairon, A., *Nature Precedings*, **2008**. hdl:10101/npre.2008.2509.1; b) Hervouet, E., Lalier, L., Debien, E., Cheray, M., Geairon, A., Rogniaux, H., Loussouarn, D., Martin, S. A., Vallette, F. M., Cartron, P.-F., *PLoS ONE*, **2010**, *5*, e11333; c) Pacaud, R., Brocard, E., Lalier, L., Hervouet, E., Vallette, F. M., Cartron, P.-F., *Sci. Rep.*, **2014**, *4*, 4230–4238.
- ⁶⁸ a) Valeur, B. *Invitation à la fluorescence moléculaire*, First ed.; De Boeck : Bruxelles, 2004; b) Lakowicz, J. R. *Principles of fluorescence spectroscopy*, Third ed.; Springer : New York, 2006.
- ⁶⁹ Specht, D. H. W., *Dtsch. Z. ges. gerichtl. Med.*, **1937**, *28*, 225–234.
- ⁷⁰ <http://www.sciencesetavenir.fr/fondamental/20141008.OBS1457/nobel-de-chimie-un-grand-pas-dans-l-observation-des-phenomenes-biologiques.html>
- ⁷¹ Ambrose, W. P., Goodwin, P. M., Jett, J. H., Van Orden, A., Werner, J. H., Keller, R. A., *Chem. Rev.*, **1999**, *99*, 2929–2956.
- ⁷² <https://www.thermofisher.com/fr/fr/home/references/molecular-probes-the-handbook/introduction-to-fluorescence-techniques.html>
- ⁷³ Hopkins, P. A., Sinkeldam, R. W., Tor, Y., *Org. Lett.*, **2014**, *16*, 5290–5293.
- ⁷⁴ Demchenko, A. P. *Introduction to Fluorescence Sensing*, Second ed.; Springer, 2008.
- ⁷⁵ a) Sozmen, F., Kolemen, S., Kumada, H. O., Ono, M., Saji, H., *RSC Advances*, **2014**, *4*, 51032–51037.; b) Caltagirone, C., Arca, M., Falchi, A. M., Lippolis, V., Meli, V., Monduzzi, M., Nylander, T., Rosa, A., Schmidt, J., Talmon, Y., Murgia, S., *RSC Advances*, **2015**, *5*, 23443–23449; c) Cheng, C., Gao, N., Yu, C., Wang, Z., Wang, J., Hao, E., Wei, Y., Mu, X., Tian, Y., Ran, C., Jiao, L. *Org. Lett.*, **2015**, *17*, 278–281.
- ⁷⁶ a) Zhou, J., Wang, L., Wang, J., Tang, N., *Transition Metal Chemistry*, **2001**, *26*, 57–63.; b) Bravo, A., Anacona, J. R., *Transition Metal Chemistry*, **2001**, *26*, 20–23; c) Gatto, M. T., Falcocchio, S., Grippa, E., Mazzanti, G., Battinelli, N., Nicolosi, G., Lambusta, D., Saso, L., *Bioorg. Med. Chem.*, **2002**, *10*, 269–272; d) Wang, R. E., Kao, J. L. F., Hilliard, C. A., Pandita, R. K., Roti Roti, J. L., Hunt, C. R., Taylor, J.-S., *J. Med. Chem.*, **2009**, *52*, 1912–1921; e) Huang, H., Jia, Q., Ma, J., Qin, G., Chen, Y., Xi, Y., Lin, L., Zhu, W., Ding, J., Jiang, H., Liu, H., *Eur. J. Med. Chem.*, **2009**, *44*, 1982–1988.
- ⁷⁷ a) Sengupta, P. K., Kasha, M., *Chem. Phys. Lett.*, **1979**, *68*, 382–385; b) Simkovitch, R., Huppert, D., *J. Phys. Chem. B*, **2015**, *119*, 10244–10251.
- ⁷⁸ Lemke, E. A., Schultz, C., *Nat. Chem. Biol.*, **2011**, *7*, 480–483.
- ⁷⁹ a) Zhuang, X., Ha, T., Kim, H. D., Centner, T., Labeit, S., Chu, S., *Proc. Natl. Acad. Sci.*, **2000**, *97*, 14241–14244; b) Ballmoos, von, C., Dimroth, P., *Anal. Biochem.*, **2004**, *335*, 334–337; c) Kong, X., Su, F., Zhang, L., Yaron, J., Lee, F., Shi, Z., Tian, Y., Meldrum, D. R., *Angew. Chem. Int. Ed.*, **2015**, *in press*. DOI: 10.1002/anie.201506038.

REFERENCES

- ⁸⁰ Reichardt, C., Welton, T., *Solvents and Solvent Effects in Organic Chemistry*, 4th ed. John Wiley & Sons, Weinheim, Germany, **2001**.
- ⁸¹ a) Reichardt, C., Welton, T., *Solvents and Solvent Effects in Organic Chemistry*, Fourth Ed., John Wiley & Sons : Weinheim - Germany, 718 pages, **2011**; b) Loving, G. S., Sainlos, M., Imperiali, B., *Trends in Biotechnology*, **2010**, *28*, 73–83.
- ⁸² a) Van der Meer, B. W., *Reviews in Molecular Biotechnology*, **2002**, *82*, 181–196; b) Berney, C., Danuser, G., *Biophys. J.*, **2003**, *84*, 3992–4010.
- ⁸³ a) Birks, J. B., *Rep. Prog. Phys.*, **1975**, *38*, 903–974; b) Conibear, P. B., Bagshaw, C. R., Fajer, P. G., Kovács, M., Málnási-Csizmadia, A., *Nature Structural & Molecular Biology*, **2003**, *10*, 831–835; c) Akimitsu Okamoto, Tomohisa Ichiba, A., Saito, I., *J. Am. Chem. Soc.*, **2004**, *126*, 8364–8365; d) Nakagawa, K., Numata, Y., Ishino, H., Tanaka, D., Kobayashi, T., Tokunaga, E., *J. Phys. Chem. C*, **2013**, *117*, 11449–11455.
- ⁸⁴ a) (1) Jayaraman, S., Biwersi, J., Verkman, A. S., *Am. J. Physiol. - Cell Physiol.*, **1999**, *276*, C747–C757; b) Tolosa, L., Ge, X., Rao, G., *Anal. Biochem.*, **2003**, *314*, 199–205; c) Xudong Ge, Leah Tolosa, A., Rao, G., *Anal. Chem.*, **2004**, *76*, 1403–1410.
- ⁸⁵ a) Carolyn C Woodroffe, Annie C Won, A., Lippard, S. J., *Inorg. chem.*, **2005**, *44*, 3112–3120; b) Wang, Z., Palacios, M. A., Zyryanov, G., Anzenbacher, P., *Chemistry*, **2008**, *14*, 8540–8546; c) Sun, S., Qiao, B., Jiang, N., Wang, J., Zhang, S., Peng, X., *Org. Lett.*, **2014**, *16*, 1132–1135; d) Yang, Z., He, Y., Lee, J. H., Chae, W.-S., Ren, W. X., Lee, J. H., Kang, C., Kim, J. S., *Chem. Commun.*, **2014**, *50*, 11672–11675; e) Chen, Y., Zhu, C., Cen, J., Bai, Y., He, W., Guo, Z., *Chem. Sci.*, **2015**, *6*, 3187–3194; f) Hou, J.-T., Li, K., Yang, J., Yu, K.-K., Liao, Y.-X., Ran, Y.-Z., Liu, Y.-H., Zhou, X.-D., Yu, X.-Q., *Chem. Commun.*, **2015**, *51*, 6781–6784.
- ⁸⁶ Demchenko, A. P., *J. Fluoresc.*, **2010**, *20*, 1099–1128.
- ⁸⁷ Gerritsen, H. C., Asselbergs, M. A. H., Agronskaia, A. V., Van Sark, W. G. J. H. M., *J. Microsc.*, **2002**, *206*, 218–224; b) Cen, L.-C., Lloyd III, W. R., Chang, C. W., Sud, D., Mycek, M.-A., *Methods in cell biology*, **2007**, *114*, 457–488.
- ⁸⁸ <https://www.thermofisher.com/fr/fr/home/references/molecular-probes-the-handbook/ph-indicators/overview-of-ph-indicators.html>
- ⁸⁹ Han, J., Burgess, K., *Chem. Rev.*, **2010**, *110*, 2709–2728.
- ⁹⁰ Zhujun, Z., Seitz, W. R., *Anal. Chim. Acta*, **1984**, *160*, 47–55.
- ⁹¹ Yao, S., Schafer-Hales, K. J., Belfield, K. D., *Org. Lett.*, **2007**, *9*, 5645–5648.
- ⁹² Lösel, R., Wehling, M., *Nature Reviews Molecular Cell Biology*, **2003**, *4*, 46–55.
- ⁹³ de Silva, A. P., Gunaratne, H. Q. N., Gunnlaugsson, T., Huxley, A. J. M., McCoy, C. P., Rademacher, J. T., Rice, T. E., *Chem. Rev.*, **1997**, *97*, 1515–1566.
- ⁹⁴ a) <https://www.thermofisher.com/content/dam/LifeTech/global/technical-reference-library/Molecular%20Probes%20Handbook/chapter-pdfs/Ch-19-Ca-Mg-Zn-Ion-Indicators.pdf?icid=WE216841>; b) Ballmoos, von, C., Dimroth, P., *Anal. Biochem.*, **2004**, *335*, 334–337.
- ⁹⁵ Reichardt, C., *Chem. Rev.*, **1994**, *94*, 2319–2358.
- ⁹⁶ a) Kenworthy, A. K., *Methods*, **2001**, *24*, 289–296; b) Heyduk, T., *Current opinion in biotechnology*, **2002**, *13*, 292–296; c) Bae, S., Kim, Y., Kim, D., Kim, K. K., Kim, Y.-G., Hohn, S., *J. Phys. Chem. B*, **2013**, *117*, 13866–13871.
- ⁹⁷ Day, R. N., Davidson, M. W., *Chem. Soc. Rev.*, **2009**, *38*, 2887–36.
- ⁹⁸ <https://www.thermofisher.com/content/dam/LifeTech/global/technical-reference-library/Molecular%20Probes%20Handbook/chapter-pdfs/Ch-13-Lipid-Membrane-Probes.pdf?icid=WE216841>
- ⁹⁹ <https://www.thermofisher.com/content/dam/LifeTech/global/technical-reference-library/Molecular%20Probes%20Handbook/chapter-pdfs/Ch-12-Probes-for-Organelles.pdf?icid=WE216841>
- ¹⁰⁰ <http://www.nal.usda.gov/fnic/foodcomp/Data/Flav/flav.pdf>
- ¹⁰¹ a) Zhou, J., Wang, L., Wang, J., Tang, N., *Transition Metal Chemistry*, **2001**, *26*, 57–63; b) Blasiak, J., Arabski, M., Pertynski, T., Malecka-Panas, E., Wozniak, K., Drzewoski, J., *Cell Biology and Toxicology*, **2002**, *18*, 279–288; c) Shukla, R., Barve, V., Padhye, S., Bhonde, R., *Bioorg. Med. Chem. Lett.*, **2004**, *14*, 4961–4965; d) Peng, W.-J., Han, X.-W., Yu, B., *Chin. J. Chem.*, **2006**, *24*, 1154–1162.
- ¹⁰² Sengupta, P. K., Kasha, M., *Chem. Phys. Lett.*, **1979**, *68*, 382–385.

REFERENCES

- ¹⁰³ Li, J.-J., and Corey, E. J., *Name Reactions in Heterocyclic Chemistry*. John Wiley & Sons, Inc. : Hoboken, NJ, USA, p262–265, **2005**.
- ¹⁰⁴ Li, J.-J., and Corey, E. J., *Name Reactions in Heterocyclic Chemistry*. John Wiley & Sons, Inc. : Hoboken, NJ, USA, p521–535, **2005**.
- ¹⁰⁵ Li, J.-J., and Corey, E. J., *Name Reactions in Heterocyclic Chemistry*. John Wiley & Sons, Inc. : Hoboken, NJ, USA, p496–503, **2005**.
- ¹⁰⁶ a) Costa, A. M. B. S. R. C. S., Dean, F. M., Jones, M. A., Smith, D. A., Varma, R. S., *J. Chem. Soc., Chem. Commun.*, **1980**, 1980, 1224–1226; b) Costa, A. M. B. S. R. C. S., Dean, F. M., Jones, M. A., Varma, R. S., *J. Chem. Soc. Perkin Trans. 1*, **1985**, 1985, 799–808; c) Sous, El, M., Khoo, M. L., Holloway, G., Owen, D., Scammells, P. J., Rizzacasa, M. A., *Angew. Chem. Int. Ed.*, **2007**, 46, 7835–7838.
- ¹⁰⁷ a) Moriarty, R. M., Prakash, O. O., *J. Heterocyclic Chem.*, **1985**, 22, 583–584.; b) Prakash, O. M., Pahuja, S., Tanwar, M. P., *Indian J. Chem. Sect. B: Org. Chem., Incl. Med. Chem.*, **1994**, 33, 272–273.
- ¹⁰⁸ a) Constantino, M. G., Júnior, V. L., Da Silva, G. V. J., *J. Heterocyclic Chem.*, **2003**, 40, 369–371; b) Maloney, D. J., Hecht, S. M., *Org. Lett.*, **2005**, 7, 1097–1099; c) Smith, J. A., Maloney, D. J., Hecht, S. M., Lannigan, D. A. D., *Bioorg. Med. Chem.*, **2007**, 15, 5018–5034.
- ¹⁰⁹ Paudel, S., Nandhikonda, P., Heagy, M. D., *J. Fluoresc.*, **2009**, 19, 681–691.
- ¹¹⁰ a) Isaks, M., Yates, K., Kalanderopoulos, P., *J. Am. Chem. Soc.*, **1984**, 106, 2728–2730; b) Kalanderopoulos, P., Yates, K., *J. Am. Chem. Soc.*, **1986**, 108, 6290–6295; c) Lukeman, M., Wan, P., *J. Am. Chem. Soc.*, **2002**, 124, 9458–9464.
- ¹¹¹ a) Chou, P.-T., Martinez, M. L., Clements, J. H., *J. Phys. Chem.*, **1993**, 97, 2618–2622; b) Klymchenko, A. S., Demchenko, A. P., Ozturk, T., Pivovarenko, V. G., *Can. J. Chem.*, **2001**, 79, 358–363; c) Klymchenko, A. S., Pivovarenko, V. G., Ozturk, T., Demchenko, A. P., *New J. Chem.*, **2003**, 27, 1336–1343; d) Demchenko, A. P., Klymchenko, A. S., *New J. Chem.*, **2004**, 28, 687–692.
- ¹¹² M'Baye, G., Klymchenko, A. S., Yushchenko, D. A., Shvadchak, V. V., Ozturk, T., Mely, Y., Duportail, G., *Photochem. Photobiol. Sci.*, **2007**, 6, 71–76.
- ¹¹³ a) Cheng, Y.-M., Pu, S.-C., Yu, Y.-C., Chou, P.-T., Huang, C.-H., Chen, C.-T., Li, T.-H., Hu, W.-P., *J. Phys. Chem. C*, **2005**, 109, 11696–11706; b) Das, R., Klymchenko, A. S., Duportail, G., and Mely, Y., *Photochem. Photobiol. Sci.*, **2009**, 8, 1583–1589.; c) Kenfack, C. A., Klymchenko, A. S., Duportail, G., Burger, A., Mely, Y., *Phys. Chem. Chem. Phys.*, **2012**, 14, 8910–8918.
- ¹¹⁴ a) McMorrow, D., Kasha, M., *J. Am. Chem. Soc.*, **1983**, 105, 5133–5134; b) McMorrow, D., Kasha, M., *J. Phys. Chem.*, **1984**, 88, 2235–2243; c) Strandjord, A. J. G., Barbara, P. F., *J. Phys. Chem.*, **1985**, 89, 2355–2361; d) Schwartz, B. J., Peteanu, L. A., Harris, C. B., *J. Phys. Chem.*, **1992**, 96, 3591–3598; e) Swinney, T. C., Kelley, D. F., *J. Chem. Phys.*, **1993**, 99, 211–221; f) Ameer-Beg, S., Ormsom, S. M., Brown, R. G., Matousek, P., Towrie, M., Nibbering, E. T. J., Foggi, P., Neuwahl, F. V. R., *J. Phys. Chem. C*, **2001**, 105, 3709–3718; g) Demchenko, A. P., Ercelen, S., *Polish J. Chem.*, **2002**, 76, 1287–1299; h) Demchenko, A. P., Bader, A. N., Pivovarenko, V. G., Ariese, F., Gooijer, C., *Spectrochim. Acta A Mol. Biomol. Spectrosc.*, **2003**, 59, 1593–1603; i) Tomin, V. I., Smolarczyk, G., *Opt. Spectrosc.*, **2008**, 104, 832–837; j) Shynkar, V. V., Klymchenko, A. S., Piémont, E., Demchenko, A. P., Mely, Y. *J. Phys. Chem. A*, **2004**, 108, 8151–8159; k) Dereka, B., Letrun, R., Svechkarev, D. A., Rosspeintner, A., Vauthay, E., *J. Phys. Chem. B*, **2015**, 119, 2434–2443; l) Tomin, V. I., Demchenko, A. P., Chou, P.-T., *J. Photochem. Photobiol. C: Photochem. Rev.*, **2015**, 22, 1–18.
- ¹¹⁵ a) Klymchenko, A. S., Ozturk, T., Pivovarenko, V. G., Demchenko, A. P., *Tet. Lett.*, **2001**, 42, 7967–7970.; b) Demchenko, A. P., Klymchenko, A. S., Pivovarenko, V. G., Ercelen, S., Duportail, G., Mely, Y., *J. Fluoresc.*, **2003**, 13, 291–295; c) Klymchenko, A. S., Demchenko, A. P., *Phys. Chem. Chem. Phys.*, **2003**, 5, 461–468; d) Klymchenko, A. S., Mely, Y., *Tet. Lett.*, **2004**, 45, 8391–8394.
- ¹¹⁶ Yesylevskyy, S. O., Klymchenko, A. S., Demchenko, A. P., *Journal of Molecular Structure: THEOCHEM*, **2005**, 755, 229–239.
- ¹¹⁷ Klymchenko, A. S., Mely, Y., *Prog. Mol. Biol. Transl. Sci.*, **2013**, 113, 35–58.
- ¹¹⁸ Sytnik, A., Gormin, D., Kasha, M., *Proc. Natl. Acad. Sci.*, **1994**, 91, 11968–11972.
- ¹¹⁹ Sytnik, A., Litvinyuk, I., *Proc. Natl. Acad. Sci.*, **1996**, 93, 12959–12963.
- ¹²⁰ a) Demchenko, A. P., Klymchenko, A. S., Avilov, S. V., *Anal. Biochem.*, **2004**, 329, 43–57.; b) Boudier, C., Klymchenko, A. S., Mely, Y., Follenius-Wund, A., *Photochem. Photobiol. Sci.*, **2009**, 8, 814–821.

REFERENCES

- ¹²¹ a) Shvadchak, V. V., Klymchenko, A. S., de Rocquigny, H., Mely, Y., *Nucleic Acids Research*, **2009**, 37, e25–e25; b) Choulier, L., Shvadchak, V. V., Naidoo, A., Klymchenko, A. S., Mely, Y., Altschuh, D., *Anal. Biochem.*, **2010**, 401, 188–195; c) Zamotaiev, O. M., Postupalenko, V. Y., Shvadchak, V. V., Pivovarenko, V. G., Klymchenko, A. S., Mely, Y., *Bioconjug. Chem.*, **2011**, 22, 101–107; d) Strizhak, A. V., Postupalenko, V. Y., Shvadchak, V. V., Morellet, N., Guittet, E., Pivovarenko, V. G., Klymchenko, A. S., Mely, Y., *Bioconjug. Chem.*, **2012**, 23, 2434–2443; e) Postupalenko, V. Y., Zamotaiev, O. M., Shvadchak, V. V., Strizhak, A. V., Pivovarenko, V. G., Klymchenko, A. S., Mely, Y., *Bioconjug. Chem.*, **2013**, 24, 1998–2007; f) Zamotaiev, O. M., Postupalenko, V. Y., Shvadchak, V. V., Pivovarenko, V. G., Klymchenko, A. S., Mely, Y., *Org. Biomol. Chem.*, **2014**, 12, 7036–7044.
- ¹²² Sarkar, M., Sengupta, P. K., *Chem. Phys. Lett.*, **1991**, 179, 68–72.
- ¹²³ a) Guharay, J., Chaudhuri, R., Chakrabarti, A., Sengupta, P. K., *Spectrochim. Acta A Mol. Biomol. Spectrosc.*, **1997**, 53, 457–462; b) Dennison, S. M., Guharay, J., Sengupta, P. K., *Spectrochim. Acta A Mol. Biomol. Spectrosc.*, **1999**, 55, 1127–1132.
- ¹²⁴ Das, R., Duportail, G., Richert, L., Klymchenko, A. S., Mely, Y., *Langmuir*, **2012**, 28, 7147–7159.
- ¹²⁵ a) Klymchenko, A. S., Duportail, G., Ozturk, T., Pivovarenko, V. G., Mely, Y., Demchenko, A. P., *Chemistry & Biology*, **2002**, 9, 1199–1208; b) Klymchenko, A. S., Demchenko, A. P., *J. Am. Chem. Soc.*, **2002**, 124, 12372–12379; c) Duportail, G., Klymchenko, A., Mely, Y., Demchenko, A. P., **2002**, *J. Fluoresc.* 12, 181–185; d) Klymchenko, A. S., Duportail, G., Mely, Y., Demchenko, A. P., *Proc. Natl. Acad. Sci.*, **2003**, 100, 11219–11224; e) Shynkar, V. V., Demchenko, A. P., Klymchenko, A. S., Duportail, G., Mely, Y., *Biochim. Biophys. Acta*, **2005**, 1712, 128–136.; f) Demchenko, A. P., Shynkar, V. V., Klymchenko, A. S., Kunzelmann, C., Duportail, G., Muller, C. D., Freyssinet, J.-M., Mely, Y., *J. Am. Chem. Soc.*, **2007**, 129, 2187–2193; g) Das, R., Klymchenko, A. S., Duportail, G., Mely, Y., *J. Phys. Chem. B*, **2008**, 112, 11929–11935; h) Demchenko, A. P., Mely, Y., Duportail, G., Klymchenko, A. S., *Biophys. J.*, **2009**, 96, 3461–3470; g) Kilin, V., Glushonkov, O., Herdly, L., Klymchenko, A. S., Richert, L., Mely, Y., *Biophys. J.*, **2015**, 108, 2521–2531.
- ¹²⁶ a) Klymchenko, A. S., Shvadchak, V. V., Yushchenko, D. A., Jain, N., Mely, Y., *J. Phys. Chem. B*, **2008**, 112, 12050–12055; b) Spadafora, M., Postupalenko, V. Y., Shvadchak, V. V., Klymchenko, A. S., Mely, Y., Burger, A., Benhida, R., *Tetrahedron*, **2009**, 65, 7809–7816; c) Xu, S., Shao, Y., Ma, K., Cui, Q., Liu, G., Wu, F., Li, M., *Analyst*, **2011**, 136, 4480–4485; d) Dziuba, D., Postupalenko, V. Y., Spadafora, M., Klymchenko, A. S., Guérineau, V., Mely, Y., Benhida, R., Burger, A., *J. Am. Chem. Soc.*, **2012**, 134, 10209–10213; e) Kuznetsova, A. A., Kuznetsov, N. A., Vorobjev, Y. N., Barthes, N. P. F., Michel, B. Y., Burger, A., Fedorova, O. S., *PLoS ONE*, **2014**, 9, e100007; f) Dziuba, D., Karpenko, I. A., Barthes, N. P. F., Michel, B. Y., Klymchenko, A. S., Benhida, R., Demchenko, A. P., Mely, Y., Burger, A., *Chemistry*, **2014**, 20, 1998–2009; g) Barthes, N. P. F., Karpenko, I. A., Dziuba, D., Spadafora, M., Auffret, J., Demchenko, A. P., Mely, Y., Benhida, R., Michel, B. Y., Burger, A., *RSC Advances*, **2015**, 5, 33536–33545; h) Sengupta, B., Reilly, S. M., Davis, D. E., Harris, K., Wadkins, R. M., Ward, D., Gholar, D., Hampton, C., *J. Phys. Chem. B*, **2015**, 119, 2546–2556.
- ¹²⁷ Tabor, H., *Biochemistry*, **1962**, 1, 496–501.
- ¹²⁸ Dziuba, D., Pohl, R., Hocek, M., *Bioconjug. Chem.*, **2014**, 25, 1984–1995.
- ¹²⁹ Stengel, G., Urban, M., Purse, B. W., Kuchta, R. D., *Anal. Chem.*, **2010**, 82, 1082–1089.
- ¹³⁰ a) Wojciechowski, F., Leumann, C. J., *Chem. Soc. Rev.*, **2011**, 40, 5669–5679; b) Weinberger, M., Berndt, F., Mahrwald, R., Ernsting, N. P., Wagenknecht, H.-A., *J. Org. Chem.*, **2013**, 78, 2589–2599; c) Kimoto, M., Hikida, Y., Hirao, I., *Isr. J. Chem.*, **2013**, 53, 450–468.
- ¹³¹ Weisbrod, S. H., Marx, A., *Chem. Commun.*, **2007**, 2007, 1828–1830.
- ¹³² a) Raindlova, V., Pohl, R., Šanda, M., Hocek, M., *Angew. Chem., Int. Ed.*, **2010**, 49, 1064–1066; b) Raindlova, V., Pohl, R., Hocek, M., *Chem. Eur. J.*, **2012**, 18, 4080–4087.
- ¹³³ a) Dadova, J., Orsaág, P., Pohl, R., Braždová, M., Fojta, M., Hocek, M., *Angew. Chem., Int. Ed.*, **2013**, 52, 10515–10518; b) Ingale, S. A., Mei, H., Leonard, P., Seela, F., *J. Org. Chem.*, **2013**, 78, 11271–11282.
- ¹³⁴ a) Bartlett, J. M. S., Stirling, D., A short history of the polymerase chain reaction, 2nd ed.; *Methods Mol. Biol.* 226, 3–6; b) Hocek, M., Fojta, M., *Chem. Soc. Rev.*, **2011**, 40, 5802–5814; c) Horáková, P., Macíčková-Cahová, H., Pivoňková, H., Spaček, J., Havran, L., Hocek, M., Fojta, M., *Org. Biomol. Chem.*, **2011**, 9, 1366–1371; d) Ménová, P., Cahová, H., Plucnara, M., Havran, L., Fojta, M., Hocek, M., *Chem. Commun.*, **2013**, 49, 4652–4654; e) Hocek, M., *J. Org. Chem.*, **2014**, 79, 9914–9921.
- ¹³⁵ a) Michelson, A. M., Todd, A. R., *J. Chem. Soc.*, **1955**, 0, 2632–2638; b) Hall, R. H., Todd, A., Webb, R. F., *J. Chem. Soc.*, **1957**, 0, 3291–3296.

REFERENCES

- ¹³⁶ a) Garegg, P. J., Lindh, I., Regberg, T., Stawinski, J., Strömberg, R., Henrichson, C., *Tet. Lett.*, **1986**, 27, 4051–4054; b) Froehler, B. C., Ng, P. G., Matteucci, M. D., *Nucleic acids Research*, **1986**, 14, 5399–5407; c) Sergueev, D. S., Shaw, B. R., *J. Am. Chem. Soc.*, **1998**, 120, 9417–9427; d) Uehara, S., Hiura, S., Higashida, R., Oka, N., Wada, T., *J. Org. Chem.*, **2014**, 79, 3465–3472; e) Warr, G. W., Karam, J. D., Chao, L., *Methods in Nucleic Acids Research*, CRC Press : Boca Raton, Florida, 432 pages, **1990**; f) Zon, G., *New J. Chem.*, **2010**, 34, 795–804.
- ¹³⁷ a) Khorana, H. G., Tener, G. M., Moffatt, J. G., Pol, E. H., *Chem. Ind.*, **1956**, 1523; b) Khorana, H. G., Razzell, W. E., Gilham, P. T., Tener, G. M., Pol, E. H., *J. Am. Chem. Soc.*, **1957**, 79, 1002–1003.
- ¹³⁸ a) Letsinger, R. L., Mahadevan, V., *J. Am. Chem. Soc.*, **1965**, 87, 3526–3527; b) Eckstein, F., Rizk, I., *Angew. Chem. Int. Ed.*, **1967**, 6, 695–696; c) Reese, C. B., Saffhill, R., *Chem. Commun.*, **1968**, 13, 767–768.
- ¹³⁹ Reese, C. B., *Org. Biomol. Chem.*, **2005**, 3, 3851–3868.
- ¹⁴⁰ Beauchage, S. L., Caruthers, M. H., *Tet. Lett.*, **1981**, 22, 1859–1862.
- ¹⁴¹ Link Technologies, *Guidebook for the Synthesis of Oligonucleotides*, **2015-16**.
- ¹⁴² McBride, L. J., Caruthers, M. H., *Tet. Lett.*, **1983**, 24, 245–248.
- ¹⁴³ Berner, S., Mühlegger, K., Seliger, H., *Nucleic acids Research*, **1989**, 17, 853–864.
- ¹⁴⁴ Scaringe, S.A., Francklyn, C., Usman, N., *Nucleic Acids Research*, **1990**, 18, 5433–5441.
- ¹⁴⁵ a) Surzhikov, S. A., Timofeev, E. N., Chernov, B. K., Golova J. B., Mirzabekov, A. D., *Nucleic Acids Research*, **2000**, 28, e29; b) Gillet, L. C. J., Alzeer, J., Schärer, O. D. *Nucleic Acids Research*, **2005**, 33, 1961–1969.; c) Sproviero, M., Rankin, K. M., Witham, A. A., Manderville, R. A., *J. Org. Chem.*, **2014**, 79, 692–699.
- ¹⁴⁶ Weisbrod, S. H., Marx, A., *Chem. Commun.*, **2008**, 44, 5675–5685.
- ¹⁴⁷ a) Rostovtsev, V. V., Green, L. G., Fokin, V. V., Sharpless, K. B., *Angew. Chem. Int. Ed. Engl.*, **2002**, 41, 2596–2599; b) Seela, F., Sirivolu, V. R., *Chemistry & Biodiversity*, **2006**, 3, 509–514; c) Burley, G. A., Gierlich, J., Mofid, M. R., Nir, H., Tal, S., Eichen, Y., Carell, T., *J. Am. Chem. Soc.*, **2006**, 128, 1398–1399; d) Johannes Gierlich, Glenn A Burley, Philipp M E Gramlich, Hammond, D. M., Carell, T., *Org. Lett.*, **2006**, 8, 3639–3642; e) Philipp M E Gramlich, Christian T Wirges, Johannes Gierlich, A., Carell, T., *Org. Lett.*, **2007**, 10, 249–251; f) Best, M. D., *Biochemistry*, **2009**, 48, 6571–6584; g) Krishna, H., Caruthers, M. H., *J. Am. Chem. Soc.*, **2012**, 134, 11618–11631; h) Wenge, U., Wagenknecht, H.-A., *Synthesis*, **2011**, 03, , 502–508; i) Wenge, U., Ehrenschwender, T., Wagenknecht, H.-A., *Bioconjug. Chem.*, **2013**, 24, 301–304; j) Wellner, C., Wagenknecht, H.-A., *Org. Lett.*, **2014**, 16, 1692–1695; k) Qu, D., Zhou, L., Wang, W., Wang, Z., Wang, G., Chi, W., Zhang, B., *Anal. Biochem.*, **2013**, 434, 128–135; l) Pujari, S. S., Ingale, S. A., Seela, F., *Bioconjug. Chem.*, **2014**, 25, 1855–1870; m) Fornich, M. A., Kvach, M. V., Navakouski, M. J., Weise, C., Baranovsky, A. V., Korshun, V. A., Shmanai, V. V., *Org. Lett.*, **2014**, 16, 4590–4593.
- ¹⁴⁸ a) Sonogashira, K., Tohda, Y., Hagihara, N., *Tet. Lett.*, **1975**, 16, 4467–4470; b) Khan, S. I., Grinstaff, M. W., *J. Am. Chem. Soc.*, **1999**, 121, 4704–4705; c) Chinchilla, R., Najera, C., *Chem. Rev.*, **2007**, 107, 874–922; d) Rist, M., Amann, N., Wagenknecht, H.-A., *Eur. J. Org. Chem.*, **2003**, 13, 2498–2504; e) Schiemann, O., Piton, N., Mu, Y., Stock, G., Engels, J. W., Prisner, T. F., *J. Am. Chem. Soc.*, **2004**, 126, 5722–5729; d) Kottysch, T., Ahlborn, C., Brotzel, F., Richert, C., *Chemistry*, **2004**, 10, 4017–4028; f) Filichev, V. V., Pedersen, E. B., *J. Am. Chem. Soc.*, **2005**, 127, 14849–14858; g) Filichev, V. V., Astakhova, I. V., Malakhov, A. D., Korshun, V. A., Pedersen, E. B., *Chemistry*, **2008**, 14, 9968–9980; h) Bomholt, N., Filichev, V. V., Pedersen, E. B., *Helv. Chem. Acta*, **2009**, 92, 716–730; i) Romainczyk, O., Endeward, B., Prisner, T. F., Engels, J. W., *Mol. Biosyst.*, **2011**, 7, 1050–1052; j) Hocek, M., Fojta, M., *Org. Biomol. Chem.*, **2008**, 6, 2233–2241; k) Ingale, S. A., Mei, H., Leonard, P., Seela, F., *J. Org. Chem.*, **2013**, 78, 11271–11282; l) Riedl, J., Pohl, R., Rulíšek, L., Hocek, M., *J. Org. Chem.*, **2012**, 77, 1026–1044.
- ¹⁴⁹ a) Miyaura, N., Yamada, K., Suzuki, A., *Tet. Lett.*, **1979**, 20, 3437–3440.; b) Miyaura, N., Suzuki, A., *J. Chem. Soc., Chem. Commun.* **1979**, 19, 866–867; c) Omumi, A., Beach, D. G., Baker, M., Gabryelski, W., Manderville, R. A., *J. Am. Chem. Soc.*, **2011**, 133, 42–50; d) Lercher, L., McGouran, J. F., Kessler, B. M., Schofield, C. J., Davis, B. G., *Angew. Chem. Int. Ed.*, **2013**, 52, 10553–10558.
- ¹⁵⁰ a) Milstein, D., Stille, J. K., *J. Am. Chem. Soc.*, **1978**, 100, 3636–3638.; b) Kosugi, M., Sasazawa, K., Shimizu, Y., Migita, T., *Chem. Lett.*, **1977**, 3, 301–302; c) Wicke, L., Engels, J. W., *Bioconjug. Chem.*, **2012**, 23, 627–642; d) Wicke, L., Engels, J. W., *Tetrahedron*, **2014**, 70, 327–333; e) Krause, A., Hertl, A., Muttach, F., Jäschke, A., *Chemistry*, **2014**, 20, 16613–16619.

REFERENCES

- ¹⁵¹ a) Tarasow, T. M., Tarasow, S. L., Eaton, B. E., *Nature*, **1997**, *389*, 54–57; b) Weisbrod, S. H., Marx, A., *Chem. Commun.*, **2007**, *18*, 1828–1830; c) Wang, Y., Vera, C. I. R., Lin, Q., *Org. Lett.*, **2007**, *9*, 4155–4158; d) Weisbrod, S. H., Baccaro, A., Marx, A., *Nucleic Acids Symp. Ser.*, **2008**, *52*, 383–384; e) Gramlich, P. M. E., Warncke, S., Gierlich, J., Carell, T., *Angew. Chem. Int. Ed.*, **2008**, *47*, 3442–3444; f) Borsenberger, V., Howorka, S., *Nucleic acids Research*, **2009**, *37*, 1477–1485; g) Angelov, T., Guainazzi, A., Schärer, O. D., *Org. Lett.*, **2009**, *11*, 661–664; h) Schoch, J., Wiessler, M., Jäschke, A. J. Am. Chem. Soc., **2010**, *132*, 8846–8847; i) Raindlová, V., Pohl, R., Sanda, M., Hocek, M., *Angew. Chem. Int. Ed.*, **2010**, *49*, 1064–1066; j) Kielkowski, P., Macíčková-Cahová, H., Pohl, R., Hocek, M., *Angew. Chem. Int. Ed.*, **2011**, *50*, 8727–8730; k) Ramil, C. P., Lin, Q., *Chem. Commun.*, **2013**, *49*, 11007–11022; l) Arndt, S., Wagenknecht, H.-A., *Angew. Chem. Int. Ed.*, **2014**, *53*, 14580–14582; m) Pyka, A. M., Domnick, C., Braun, F., Kath-Schorr, S., *Bioconjug. Chem.*, **2014**, *25*, 1438–1443; n) Vanikova, Z., Hocek, M., *Angew. Chem. Int. Ed.*, **2014**, *53*, 6734–6737.
- ¹⁵² Kapuscinski, J., *Biotech. Histochem.*, **1995**, *70*, 220–233.
- ¹⁵³ Daniels, M., Hauswirth, W., *Science*, **1971**, *171*, 675–677.
- ¹⁵⁴ a) Lakowicz, J. R., (**2006**) Principles of Fluorescence Spectroscopy 3rd ed., pp 1–961. Springer Science +Business Media, LLC, New York; b) LePecq, J. B., Paoletti, C., *J. Mol. Biol.*, **1967**, *27*, 87–106.
- ¹⁵⁵ a) Kienast, J., Schmitz, G., *Blood*, **1990**, *75*, 116–121; b) Tse, W. C., Boger, D. L., *Acc. Chem. Res.*, **2004**, *37*, 61–69
- ¹⁵⁶ a) Hirons, G. T., Fawcett, J. J., Crissman, H. A., *Cytometry*, **1994**, *15*, 129–140; b) Kim, S.-H., *Functional Dyes*, First Ed., Elsevier : Amsterdam, 650 pages, **2006**.
- ¹⁵⁷ a) Vummidi, B. R., Alzeer, J., Luedtke, N. W., *ChemBioChem*, **2013**, *14*, 540–558; b) Ohnmacht, S. A., Neidle, S., *Bioorg. Med. Chem. Lett.*, **2014**, *24*, 2602–2612.
- ¹⁵⁸ Laguerre, A., Stefan, L., Larrouy, M., Genest, D., Novotna, J., Pirrotta, M., Monchaud, D., *J. Am. Chem. Soc.*, **2014**, *136*, 12406–12414.
- ¹⁵⁹ a) Eriksson, S., Kim, S. K., Kubista, M., Nordén, B., *Biochemistry*, **1993**, *32*, 2987–2998; b) Larsen, T. A., Goodsell, D. S., Cascio, D., Grzeskowiak, K., Dickerson, R. E., *J. Biomol. Struct. Dyn.*, **1989**, *7*, 477–491.
- ¹⁶⁰ Barcellona, M. L., Cardiel, G., Gratton, E., *Biochem. Biophys. Res. Commun.*, **1990**, *170*, 270–280.
- ¹⁶¹ Latt, S. A., Stetten, G., Juergens, L. A., Willard, H. F., Scher, C. D., *J. Histochem. Cytochem.*, **1975**, *23*, 493–505.
- ¹⁶² Sengupta, B., Reilly, S. M., Davis, D. E., Harris, K., Wadkins, R. M., Ward, D., Gholar, D., Hampton, C., *J. Phys. Chem. B*, **2015**, *119*, 2546–2556.
- ¹⁶³ Granzhan, A., Kotera, N., Teulade-Fichou, M.-P., *Chem. Soc. Rev.*, **2014**, *43*, 3630–36.
- ¹⁶⁴ Brauns, E. B., Murphy, C. J., Berg, M. A., *J. Am. Chem. Soc.*, **1998**, *120*, 2449–2456.
- ¹⁶⁵ a) Dodd, D. W., Hudson, R. H. E., *Mini-Reviews in Organic Chemistry*, **2009**, *6*, 378–391; b) Teo, Y. N., Kool, E. T., *Chem. Rev.*, **2012**, *112*, 4221–4245.
- ¹⁶⁶ Capobianco, M. L., Naldi, M., Zambianchi, M., Barbarella, G., *Tet. Lett.*, **2005**, *46*, 8181–8184.
- ¹⁶⁷ Tram, K., Twohig, D., Yan, H., *Nucleosides Nucleotides Nucleic Acids*, **2011**, *30*, 1–11.
- ¹⁶⁸ Bandy, T. J., Brewer, A., Burns, J. R., Marth, G., Nguyen, T., Stulz, E., *Chem. Soc. Rev.*, **2011**, *40*, 138–148.
- ¹⁶⁹ Wilson, J. N., Kool, E. T., *Org. Biomol. Chem.*, **2006**, *4*, 4265–4274.
- ¹⁷⁰ a) Grünefeld, P., Richert, C., *J. Org. Chem.*, **2004**, *69*, 7543–7551; b) Lachmann, D., Berndl, S., Wolfbeis, O. S., Wagenknecht, H.-A., *Beilstein J. Org. Chem.*, **2010**, *6*, 1–7.
- ¹⁷¹ Stambaský, J., Hocek, M., Kocovský, P., *Chem. Rev.*, **2009**, *109*, 6729–6764.
- ¹⁷² a) Ren, R. X.-F., Chaudhuri, N. C., Paris, P. L., Rumney, S., Kool, E. T., *J. Am. Chem. Soc.*, **1996**, *118*, 7671–7678; b) Chaudhuri, N. C., Ren, R. X.-F., Kool, E. T., *Synlett*, **1997**, *4*, 341–347; c) Strässler, C., Davis, N. E., Kool, E. T., *Helv. Chim. Acta*, **1999**, *82*, 2160–2171; d) Gao, J., Strässler, C., Tahmassebi, D., Kool, E. T., *J. Am. Chem. Soc.*, **2002**, *124*, 11590–11591; e) Aubert, Y., Asseline, U., *Org. Biomol. Chem.*, **2004**, *2*, 3496–3503; f) Wilson, J. N., Gao, J., Kool, E. T., *Tetrahedron*, **2007**, *63*, 3427–3433; g) Grigorenko, N. A., Leumann, C. J., *Chemistry*, **2009**, *15*, 639–645.
- ¹⁷³ a) Cuppoletti, A., Cho, Y., Park, J.-S., Strässler, C., Kool, E. T., *Bioconjug. Chem.*, **2005**, *16*, 528–534; b) Teo, Y. N., Wilson, J. N., Kool, E. T., *J. Am. Chem. Soc.*, **2009**, *131*, 3923–3933.
- ¹⁷⁴ Morales-Rojas, H., Kool, E. T., *Org. Lett.*, **2002**, *4*, 4377–4380.

REFERENCES

- ¹⁷⁵ Sinkeldam, R. W., Greco, N. J., Tor, Y. *Chem. Rev.*, **2010**, *110*, 2579–2619.
- ¹⁷⁶ a) Ward, D. C., Reich, E., Stryer, L., *J. Biol. Chem.*, **1969**, *244*, 1228–1237.; b) Srivatsan, S. G., Sawant, A. A., *Pure Appl. Chem.*, **2011**, *83*, 213–232.
- ¹⁷⁷ Park, S., Otomo, H., Zheng, L., Sugiyama, H., *Chem. Commun.*, **2014**, *50*, 1573–1575.
- ¹⁷⁸ McCoy, L. S., Shin, D., Tor, Y., *J. Am. Chem. Soc.*, **2014**, *136*, 15176–15184.
- ¹⁷⁹ Jessen, H. J., Fendrich, W., Meier, C., *Eur. J. Org. Chem.*, **2006**, *2006*, 924–931.
- ¹⁸⁰ Schlitt, K. M., Millen, A. L., Wetmore, S. D., Manderville, R. A., *Org. Biomol. Chem.*, **2011**, *9*, 1565–1571.
- ¹⁸¹ Ben Gaiel, N., Glasser, N., Ramalanjaona, N., Beltz, H., Wolff, P., Marquet, R., Burger, A., Mely, Y., *Nucleic acids Research*, **2005**, *33*, 1031–1039.
- ¹⁸² a) Greco, N. J., Tor, Y., *J. Am. Chem. Soc.*, **2005**, *127*, 10784–10785; b) Srivatsan, S. G., Tor, Y., *J. Am. Chem. Soc.*, **2007**, *129*, 2044–2053; c) Pawar, M. G., Nuthanakanti, A., Srivatsan, S. G., *Bioconjug. Chem.*, **2013**, *24*, 1367–1377.
- ¹⁸³ a) Mitsui, T., Kimoto, M., Kawai, R., Yokoyama, S., Hirao, I., *Tetrahedron*, **2007**, *63*, 3528–3537; b) Greco, N. J., Tor, Y., *Tetrahedron*, **2007**, *63*, 3515–3527.
- ¹⁸⁴ Hirose, W., Sato, K., Matsuda, A., *Nucleic Acids Symp. Ser.*, **2009**, *53*, 135–136.
- ¹⁸⁵ Tanpure, A. A., Srivatsan, S. G., *Chemistry*, **2011**, *17*, 12820–12827.
- ¹⁸⁶ Pawar, M. G., Srivatsan, S. G., *J. Phys. Chem. B*, **2013**, *117*, 14273–14282.
- ¹⁸⁷ a) Okamoto, A., Tanaka, K., Fukuta, T., Saito, I., *ChemBioChem*, **2004**, *5*, 958–963; b) Srivatsan, S. G., Weizman, H., Tor, Y., *Org. Biomol. Chem.*, **2008**, *6*, 1334–1338.; c) Xie, Y., Dix, A. V., Tor, Y., *J. Am. Chem. Soc.*, **2009**, *131*, 17605–17614; d) Xie, Y., Maxson, T., Tor, Y., *J. Am. Chem. Soc.*, **2010**, *132*, 11896–11897.
- ¹⁸⁸ a) Lin, K.-Y., Jones, R. J., Matteucci, M., *J. Am. Chem. Soc.*, **1995**, *117*, 3873–3874; b) Wilhelmsson, L. M., Holmén, A., Lincoln, P., Nielsen, P. E., Nordén, B., *J. Am. Chem. Soc.*, **2001**, *123*, 2434–2435; c) Sandin, P., Börjesson, K., Li, H., Mårtensson, J., Brown, T., Wilhelmsson, L. M., Albinsson, B., *Nucleic Acids Research*, **2008**, *36*, 157–167; c) Sandin, P., Stengel, G., Ljungdahl, T., Börjesson, K., Macao, B., Wilhelmsson, L. M., *Nucleic Acids Research*, **2009**, *37*, 3924–3933.
- ¹⁸⁹ a) Inoue, H., Imura, A., Ohtsuka, E., *Nippon Kagaku Kaishi*, **1987**, *7*, 1214–1220; b) Wahba, A. S., Esmaeili, A., Damha, M. J., Hudson, R. H. E., *Nucleic acids Research*, **2010**, *38*, 1048–1056.
- ¹⁹⁰ Sechrist, J. A., Barrio, J. R., Leonard, N. J., *Science*, **1972**, *175*, 646–647.
- ¹⁹¹ a) Scopes, D. I., Leonard, N. J., *Science*, **1977**, *195*, 296–298; b) Leonard, N. J., *Acc. Chem. Res.*, **1982**, *15*, 128–135; c) Krueger, A. T., Lu, H., Lee, A. H. F., Kool, E. T., *Acc. Chem. Res.*, **2007**, *40*, 141–150.
- ¹⁹² Gao, J., Liu, H., Kool, E. T., *J. Am. Chem. Soc.*, **2004**, *126*, 11826–11831.
- ¹⁹³ Mayer, E., Valis, L., Wagner, C., Rist, M., Amann, N., Wagenknecht, H.-A., *ChemBioChem*, **2004**, *5*, 865–868.
- ¹⁹⁴ a) Malinovskii, V. L., Wenger, D., Häner, R., *Chem. Soc. Rev.*, **2010**, *39*, 410–422; b) Morihiko, K., Hasegawa, O., Mori, S., Tsunoda, S., Obika, S., *Org. Biomol. Chem.*, **2015**, *13*, 5209–5214.
- ¹⁹⁵ a) Seo, Y. J., Hwang, G. T., Kim, B. H., *Nucleic Acids Symp. Ser.*, **2005**, *49*, 135–136; b) Seo, Y. J., Ryu, J. H., Kim, B. H., *Org. Lett.*, **2005**, *7*, 4931–4933; c) Seo, Y. J., Kim, B. H., *Chem. Commun.*, **2006**, *2*, 150–152.
- ¹⁹⁶ Ryu, J. H., Seo, Y. J., Hwang, G. T., Lee, J. Y., Kim, B. H., *Tetrahedron*, **2007**, *63*, 3538–3547.
- ¹⁹⁷ Kim, M. J., Seo, Y., Hwang, G. T., *RSC Advances*, **2014**, *4*, 12012–12017.
- ¹⁹⁸ Xiao, Q., Ranasinghe, R. T., Tang, A. M. P., Brown, T., *Tetrahedron*, **2007**, *63*, 3483–3490.
- ¹⁹⁹ Hwang, G. T., Seo, Y. J., Kim, B. H., *Tet. Lett.*, **2005**, *46*, 1475–1477.
- ²⁰⁰ Sessler, J. L., Sathiosatham, M., Brown, C. T., Rhodes, T. A., Wiederrecht, G., *J. Am. Chem. Soc.*, **2001**, *123*, 3655–3660.
- ²⁰¹ a) Saito, Y., Suzuki, A., Okada, Y., Yamasaka, Y., Nemoto, N., Saito, I., *Chem. Commun.*, **2013**, *49*, 5684–5686; b) Suzuki, A., Nemoto, N., Saito, I., Saito, Y., *Org. Biomol. Chem.*, **2013**, *12*, 660–666.
- ²⁰² a) Sanger, F., Coulson, A. R., *J. Mol. Biol.*, **1975**, *94*, 441–448; b) Sanger, F., Nicklen, S., Coulson, A. R., *Proc. Natl. Acad. Sci.*, **1977**, *74*, 5463–5467.

REFERENCES

- ²⁰³ Middendorf, L. R., Humphrey, P. G., Narayanan, N., Roemer, S. C., *Sequencing Technology, in Biotechnology*, Second Ed., Wiley-VCH Verlag GmbH : Weinheim - Germany, p. 193–226, **2001**.
- ²⁰⁴ a) Ingram, V. M., *Nature*, **1956**, *178*, 792–794; b) Chang, J. C., Kan, Y. W., *Proc. Natl. Acad. Sci.*, **1979**, *76*, 2886–2889; c) Hamosh, A., King, T. M., Rosenstein, B. J., Corey, M., Levison, H., Durie, P., Tsui, L. C., McIntosh, I., Keston, M., Brock, D. J., *Am. J. Hum. Genet.*, **1992**, *51*, 245–250; d) Duan, R., Pak, C., Jin, P., *Human Molecular Genetics*, **2007**, *16*, 1124–1131.
- ²⁰⁵ a) Norton, N., Williams, N. M., Williams, H. J., Spurlock, G., Kirov, G., Morris, D. W., Hoogendoorn, B., Owen, M. J., O'Donovan, M. C., *Hum. Genet.*, **2002**, *110*, 471–478; b) Chen, X., Sullivan, P. F., *Pharmacogenomics J.*, **2003**, *3*, 77–96; c) Campbell, N., Reece, J., *Biologie*, Seventh Ed., Pearson Education, 438 pages, **2007**; d) <https://www.thermofisher.com/fr/fr/home/life-science/sequencing/dna-sequencing/snp-genotyping-variant-detection-sequencing/snp-genotyping-fragment-analysis.html>
- ²⁰⁶ Jensen, J. S., Bjornelius, E., Dohn, B., Lidbrink, P., *J. Clin. Microbiol.*, **2004**, *42*, 683–692.
- ²⁰⁷ a) Syvänen, A.-C., *Nat. Rev. Genet.*, **2001**, *2*, 930–942; b) Takatsu, K., Yokomaku, T., Kurata, S., Kanagawa, T., *Nucleic Acids Research*, **2004**, *32*, e156–e156; c) Kolpashchikov, D. M., *Chem. Rev.*, **2010**, *110*, 4709–4723; d) <http://www.atdbio.com/content/20/Sequencing-forensic-analysis-and-genetic-analysis#Single-nucleotide-polymorphisms-SNPs>
- ²⁰⁸ a) Xu, D., Evans, K. O., Nordlund, T. M., *Biochemistry*, **1994**, *33*, 9592–9599; b) Holz, B., Weinhold, E., Klimasauskas, S., Serva, S., *Nucleic Acids Research*, **1998**, *26*, 1076–1083; c) Stivers, J. T., *Nucleic Acids Research*, **1998**, *26*, 3837–3844; d) Rachofsky, E. L., Seibert, E., Stivers, J. T., Osman, R., Ross, J. B., *Biochemistry*, **2001**, *40*, 957–967; e) Hariharan, C., Bloom, L. B., Helquist, S. A., Kool, E. T., Reha-Krantz, L. J., *Biochemistry*, **2006**, *45*, 2836–2844; f) Jones, A. C., Neely, R. K., *Quarterly Reviews of Biophysics*, **2015**, *48*, 244–279.
- ²⁰⁹ a) Choi, J., Kim, S., Tachikawa, T., Fujitsuka, M., Majima, T., *J. Am. Chem. Soc.*, **2011**, *133*, 16146–16153; b) Bae, S., Son, H., Kim, Y.-G., Hohng, S., *Phys. Chem. Chem. Phys.*, **2013**, *15*, 15829–15832; c) Bae, S., Kim, Y., Kim, D., Kim, K. K., Kim, Y.-G., Hohng, S., *J. Phys. Chem. B*, **2013**, *117*, 13866–13871; d) Wang, F., Liu, X., Willner, I., *Angew. Chem. Int. Ed.*, **2015**, *54*, 1098–1129.
- ²¹⁰ a) Riedl, J., Ménová, P., Pohl, R., Orság, P., Fojta, M., Hocek, M., *J. Org. Chem.*, **2012**, *77*, 8287–8293; b) Moreno, A., Knee, J., Mukerji, I., *Biochemistry*, **2012**, *51*, 6847–6859; c) Machado, I., Ozalp, V. C., Rezabal, E., Schäfer, T., *Chemistry*, **2014**, *20*, 11820–11825.
- ²¹¹ Srivatsan, S. G., Tor, Y., *Tetrahedron*, **2007**, *63*, 3601–3607.
- ²¹² Parsons, J., Hermann, T., *Tetrahedron*, **2007**, *63*, 3548–3552.
- ²¹³ Stresemann, C., Lyko, F., *Int. J. Cancer*, **2008**, *123*, 8–13.
- ²¹⁴ Dolinoy, D. C., Huang, D., Jirtle, R. L., *Proc. Natl. Acad. Sci.*, **2007**, *104*, 13056–13061.
- ²¹⁵ Almouzni, G., Altucci, L., Amati, B., Ashley, N., Baulcombe, D., Beaujean, N., Bock, C., Bongcam-Rudloff, E., Bousquet, J., Braun, S., Bressac-de Paillerets, B., Bussemakers, M., Clarke, L., Conesa, A., Estivill, X., Fazeli, A., Grgurević, N., Gut, I., Heijmans, B. T., Hermouet, S., Houwing-Duistermaat, J., Iacobucci, I., Ilaš, J., Kandimalla, R., Krauss-Etschmann, S., Lasko, P., Lehmann, S., Lindroth, A., Majdič, G., Marcotte, E., Martinelli, G., Martinet, N., Meyer, E., Miceli, C., Mills, K., Moreno-Villanueva, M., Morvan, G., Nickel, D., Niesler, B., Nowacki, M., Nowak, J., Ossowski, S., Pelizzola, M., Pochet, R., Potočnik, U., Radwanska, M., Raes, J., Rattray, M., Robinson, M. D., Roelen, B., Sauer, S., Schinzer, D., Slagboom, E., Spector, T., Stunnenberg, H. G., Tiligada, E., Torres-Padilla, M.-E., Tsonaka, R., Van Soom, A., Vidaković, M., Widschwendter, M., *BMC Genomics*, **2014**, *15*, 487–493.
- ²¹⁶ Christodoulou, M. S., Thomas, A., Poulain, S., Vidaković, M., Lahtela-Kakkonen, M., Matulis, D., Bertrand, P., Bartova, E., Blanquart, C., Mikros, E., Fokialakis, N., Passarella, D., Benhida, R., Martinet, N., *Med. Chem. Commun.*, **2014**, *5*, 1804–1820.
- ²¹⁷ Takeshita, A., Koibuchi, N., Oka, J., Taguchi, M., Shishiba, Y., Ozawa, Y., *Eur. J. Endocrinol.*, **2001**, *145*, 513–517.
- ²¹⁸ Kucharski, R., Maleszka, J., Foret, S., Maleszka, R., *Science*, **2008**, *319*, 1827–1830.
- ²¹⁹ Fraga, M. F., Ballestar, E., Paz, M. F., Ropero, S., Setien, F., Ballestar, M. L., Heine-Suñer, D., Cigudosa, J. C., Urioste, M., Benitez, J., Boix-Chornet, M., Sanchez-Aguilera, A., Ling, C., Carlsson, E., Poulsen, P., Vaag, A., Stephan, Z., Spector, T. D., Wu, Y.-Z., Plass, C., Esteller, M., *Proc. Natl. Acad. Sci.*, **2005**, *102*, 10604–10609.

REFERENCES

- ²²⁰ a) Cubas, P., Vincent, C., Coen, E., *Nature*, **1999**, *401*, 157–161; b) Cronk, Q. C., *Nat. Rev. Genet.*, **2001**, *2*, 607–619.
- ²²¹ Kronholm, I., Collins, S., *Mol. Ecol.*, **2015**, *in press*. DOI: 10.1111/mec.13296.
- ²²² www.nobelprize.org/nobel_prizes/chemistry/laureates/2015/popular-chemistryprize2015.pdf

Résumé : La méthylation de l'ADN est une des marques épigénétiques majeures qui intervient dans la régulation de processus physiologiques importants (ex : empreinte génomique parentale ou expression tissus-spécifique de gène). Par ailleurs la méthylation de l'ADN a été reconnue comme une cible majeur pour lutter contre le cancer car son dysfonctionnement est impliqué dans le développement de pathologies comme les cancers. Bien que de nombreux inhibiteurs aient été identifiés, leur manque de sélectivité, leur faible activité et leur toxicité limitent leur développement. Une meilleure compréhension de ce mécanisme permettrait aux chimistes médicinaux de développer de nouveaux inhibiteurs plus spécifiques.

Actuellement, un grand nombre d'hypothèses ont été avancées concernant la méthylation de l'ADN et méritent d'être confirmées ou clarifiées. Dans ce contexte, nous avons décidé de développer de nouveaux analogues nucléosidiques fluorescents et ratiométriques présentant un motif 3-hydroxychromones (3-HC), afin d'étudier la dynamique et le mécanisme de méthylation de la cytosine impliquant UHRF1 et DNMT1. Nous avons tout d'abord décidé de nous focaliser sur la première étape du mécanisme, la reconnaissance du duplexe hémiméthylé par le domaine SRA de UHRF1.

Deux stratégies de marquage de l'ADN ont ainsi été explorées. La première a consisté à remplacer une base azotée naturelle par le fluorophore, afin de sonder l'intérieur du duplexe. Ce biocapteur, obtenu après une synthèse convergente en 16 étapes, a ainsi pu être utilisé afin de détecter l'approche du domaine SRA de UHRF1 ainsi que le basculement de la cytosine méthylée dans son site de reconnaissance. De plus, des études en écoulement bloqué ont permis de déterminer les constantes de vitesses du processus de basculement de la base. La deuxième approche repose, quant à elle, sur la modification d'une nucléobase naturelle sur laquelle la sonde ratiométrique est attachée à l'aide d'un bras espaceur insaturé rigide, lui permettant ainsi de sonder les interactions au niveau du grand sillon de l'ADN. Cette sonde, obtenue après 8 étapes de synthèse, s'est avérée particulièrement intéressante afin de sonder le taux d'hydratation dans le grand sillon de l'ADN, ouvrant la porte à des applications prometteuses pour l'étude du mécanisme de méthylation de l'ADN.

Mots clés : Méthylation de l'ADN, interaction ADN/protéine, sondes fluorescentes ratiométriques, 3-hydroxychromone, nucléosides.

Abstract: DNA methylation is one of the major epigenetic modification and plays an important role in the regulation of major processes (ex: genomic imprinting or tissue-specific expression of genes). On the other hand, alterations of the DNA methylation pattern are involved in the development of pathologies like cancers. DNA methylation was recently recognized as one of the major target to inhibit and to fight cancer. Although several inhibitors have already been identified, their lack of selectivity, poor activity and toxicity limit their development. A deeper insight of the DNA methylation mechanism should help medicinal chemists in the design and research of more specific inhibitors of DNA methylation.

Currently, miscellaneous hypotheses are advanced about the DNA methylation mechanism and deserve to be confirmed or clarified. In this context, we decided to develop new ratiometric fluorescent nucleoside analogs, incorporating 3-hydroxychromone (3-HC) moiety, to study the dynamics and mechanism of cytosine methylation involving UHRF1 and DNMT1. We first decided to focus on the first step of the mechanism, the recognition of the hemi-methylated duplex by the SRA domain of UHRF1.

Two different strategies of labeling DNA were explored. In the first one, a natural nucleobase was substituted by the fluorescent dye in order to probe inside the duplex. This biosensor, obtained after a convergent synthesis of 16 steps, allows to detect and monitor the binding of the SRA domain and flipping of the 5-methylcytosine from the duplex into its recognizing site. Moreover, *stopped-flow* studies were used to characterize the kinetic rate constants of the base flipping process. The second approach is based on a modification of a natural nucleobase on which, through a rigid unsaturated linker, the ratiometric probe was incorporated in order to sense the interactions in the major groove. This probe, obtained after a 8-step synthesis, gave particularly interesting and promising results on probing hydration rate of the major groove, suggesting upcoming applications to study the DNA methylation mechanism.

Keywords: DNA Methylation, nucleic acid/protein interactions, ratiometric fluorescent probe, 3-hydroxychromone, nucleosides.