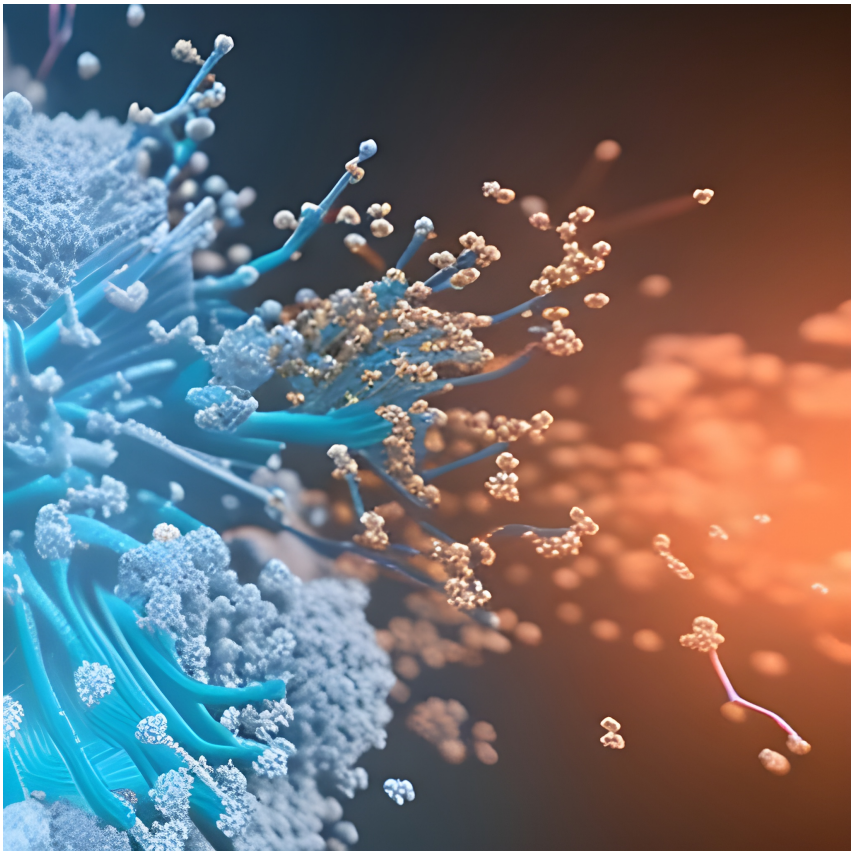
Comptes Rendus
de l'Académie des sciences

1878-1543 (electronic)

Chimie



Volume 26, Special Issue S3, 2023

Special issue / *Numéro spécial*

Breaking Barriers in Chemical Biology – Toulouse 2022 / *Briser les barrières entre la chimie et la biologie – Toulouse 2022*

Guest editors / *Rédacteurs en chef invités*

Marie Lopez, Elisabetta Mileo, Eric Defrancq, Agnes Delmas, Boris Vauzeilles, Dominique Guianvarch, Christophe Biot

Académie des sciences — Paris

INSTITUT DE FRANCE
Académie des sciences

CENTRE MERSENNE

Comptes Rendus

Chimie

Objective of the journal

Comptes Rendus Chimie is an international peer-reviewed electronic journal, covering all areas of chemistry.

It publishes special issues, original research articles, review articles, accounts, historical perspectives, pedagogical texts or conference proceedings, without length limit, in English or in French and in a format as flexible as necessary (figures, associated data, etc.).

Comptes Rendus Chimie is published since 2020 with the centre Mersenne pour l'édition scientifique ouverte (Mersenne Center for open scientific publishing), according to a virtuous Diamond Open Access policy, free for authors (no author processing charges nor publishing fees) as well as for readers (immediate and permanent open access).

About the journal

Comptes Rendus Chimie is published exclusively in electronic format.

All information on the journal, as well as the full text of all articles, is available on its website at: https://comptes-rendus.academie-sciences.fr/chimie/.

Author enquiries

For any inquiries about submitting a manuscript, please refer to the journal's website: https://comptes-rendus.academie-sciences.fr/chimie/.

Contact

Académie des sciences
23 quai de Conti
75006 Paris (France)
cr-chimie@academie-sciences.fr

Comptes Rendus
Chimie
2023, 26, nº S3

Contents / *Sommaire*

Comptes Rendus
Chimie
2023, Vol. 26, Special Issue S3, p. 1-4
https://doi.org/10.5802/crchim.323

Breaking Barriers in Chemical Biology – Toulouse 2022

Guest Editors

Marie Lopez obtained her PhD in chemo-enzymatic synthesis of oligosaccharides in 2007 under the supervision of Dr. H. Driguez and Dr. A. Buléon. In 2008, she joined Poulsen's group (GRIDD, Brisbane) as a post-doctoral fellow to work on glycosylated carbonic anhydrase inhibitors. In a second post-doctoral position, she worked on anti-tuberculosis agents in Montpellier before being recruited, in 2013, as a CNRS research scientist at the ETaC laboratory in Toulouse, directed by Paola B. Arimondo, to work on DNA methylation and its implication in cancers. In 2016, she moved to the IBMM in Montpellier, managing a research group working in chemistry, biochemistry and chemical biology to decipher and target epigenetic modifications in pathological contexts. Her projects consist in identifying new strategies to target epigenetic mechanism, i.e. multivalent, multifunctional and hybrid inhibitors, ProTaCs, peptide-based inhibitors and in developing affinity-based chemical probes to understand epigenetic mechanism in diverse context, i.e. cancer, infectious diseases, invertebrate models. In 2020 she was elected as a member of the scientific board of the *Chémobiologie* group of the *Société Chimique de France* (SCF-ChemBio). Since 2021 she has been a member of the scientific committee of the ChemBio GDR.

Elisabetta Mileo is a CNRS researcher at the laboratory "Bioenergetique et Ingenierie des Protéines" (BIP) in Marseille. She is a chemist and a spectroscopist with experience in the synthesis and use of nitroxide radicals as spin probes and spin labels in supramolecular chemistry and applied to structural biology. She obtained her PhD in 2010 at the University of Bologna (Italy) under the supervision of Professor Marco Lucarini in the field of self-assembled supramolecular architectures investigated by Electron Paramagnetic Resonance (EPR) spectroscopy. During her PhD, she joined the group of Professor A. E. Kaifer, during 6 months (Miami, Florida) investigating the self-assembly of resorcinarene capsules. After a postdoctoral fellow in the group of Professor S. Marque (ICR, Marseille, France, 2010) working on the synthesis of new nitroxide-based spin labels for selective Tyrosine labeling, she joined the group of Professor B. Guigliarelli in 2012 at the BIP laboratory (Marseille, France) and her scientific interests shifted towards the investigation of protein structural dynamics by Site-Directed Spin Labelling coupled to EPR spectroscopy (SDSL-EPR). In 2014, she obtained a full position as CNRS researcher (Marseille, France). Since 2014,

E. Mileo research activities are mainly focused on the investigation of protein structural dynamics, in particular in chaperones and other flexible proteins, with the objective of gaining information on how protein dynamics affects protein function. The originality of her work resides on the fact that protein investigation by SDSL-EPR is carried directly inside cells (in-cell EPR), in physiological conditions. Her studies are also aimed to the development of new tools and new spin labels to follow proteins "in action" by EPR directly into living cells.

After obtaining a PhD in Chemistry in 1989 at the University J. Fourier of Grenoble (UJF), **Eric Defrancq** realized a post-doctoral internship at the Institute of Chemistry of Neuchâtel (Switzerland) in the field of the chemistry of natural substances. In 1992, he was appointed lecturer at the UJF at the LEDSS and in 2004 he became a university professor. The common thread of his research is the Chemical-Biology. In this context, he carries out his research activities within the Department of Molecular Chemistry (DCM-UMR CNRS 5250) at the University of Grenoble Alpes and within the laboratory, he is the scientific coordinator of the Chemical Engineering of Nucleic Acids group. His research themes relate more particularly to the development of high-performance chemical methods for incorporating a motif into a nucleic acid (DNA or RNA) in order to provide it with new properties. In particular, in recent years, he has been interested in specific DNA structures: G-quadruplexes and i-motifs which constitute new biological targets. In this context, he has developed an original concept (named TASQ for Template Assisted Synthetic G-Quadruplex) to constrain the topology of these tetrameric DNAs. These biomolecular systems are then used for the design and study of new photo-activable metal complexes targeting G-quadruplex DNA, the production of specific antibodies for G-quadruplex DNA and for the study of proteins interacting with these DNA. He is also director of the Institute of Molecular Chemistry of Grenoble (ICMG – UAR 2602) and scientific advisor to Hcéres.

Agnès Delmas, a pharmacist by training, is Director of Research Emeritus at the CNRS. Her scientific career, between chemistry and biology, was pursued at the Centre of Molecular Biophysics (Orléans), where she has promoted and structured a team focused on the chemistry of synthetic proteins. After two decades at the head of the peptide synthesis research group, she joined the NMR thematic group, which she led until 2020. Her role in facilitating collaborative research between the synthesis and NMR groups has highlighted the thematic axis focused on the synthesis and structural study of naturally occurring peptides/proteins rich in disulfide bridges. In addition to her scientific work, Agnès Delmas was the director of the doctoral school, of the Physics and Chemistry of Life federation and President of Section 16 of the CNRS, always seeking to promote research in chemical biology.

An alumnus of the Ecole Normale Supérieure in Paris, **Boris Vauzeilles** prepared his PhD under the supervision of Professor Pierre Sinaÿ in the chemistry department of this institution. He then joined Professor Julius Rebek, Jr. at MIT (Cambridge, Massachussets), for a post-doctoral experience. The group soon moved to the Scripps Research Institute in La Jolla, where Julius Rebek was promoted to Director of the Skaggs Institute for Chemical Biology. Boris Vauzeilles then returned to France as a CNRS researcher (Institut de Chimie Moléculaire et des Matériaux d'Orsay, Université Paris-Saclay), and in 2012, in parallel to his research at Orsay, created a new team at the Institut de Chimie des Substances Naturelles à Gif-sur-Yvette, where he founded the Chemical Biology Department in 2015. He is also co-founder of a startup company, Diamidex, and received, with his group, the Prix La Recherche in 2015. His research is primarily focused on using synthetic chemistry to develop molecular tools to probe biological processes. Since 2020, he has been the director of the Institute of Chemistry of Natural Substances. He chairs the Chemical Biology Group of the French Chemical Society (SCF-ChemBio), and is on the Board of Directors of the International Chemical Biology Society.

Dominique Guianvarc'h received a formation in Biochemistry and Organic Chemistry from the University Pierre et Marie Curie in Paris. She received her PhD in Organic Chemistry in 2001 under the supervision of Professor Jean-Louis Fourrey at the Institut de Chimie des Substances Naturelles in Gif sur Yvette (France) on *C*-nucleosides and nucleic acids chemistry. She then pursued a post-doctoral training in the laboratory of Biophysics of Professor Claude Hélène at the Museum National d'Histoire Naturelle in Paris on the development of oligonucleotide-conjugated topoisomerase inhibitors. In 2003, she was recruited as Associate Professor in the laboratory of Biomolecules founded by Professor Andrée Marquet at the University Pierre et Marie Curie. In 2007, she developed her own group devoted to the study of methyltransferases enzymes with chemical tools. In 2017, she was promoted Professor at the University Paris Saclay at the Institut de Chimie Moléculaire et des Matériaux d'Orsay (France). She is group leader of a team in Bioorganic Chemistry with complementary skills in carbohydrate and nucleic acid chemistry, protein engineering and enzymology. Her research interests are in the areas of chemical biology including the development of chemical probes to understand different biological issues involving enzymes with potential applications in the field of health. In 2020, she was elected as board member and scientific secretary of the Chemical Biology Group within the French Chemical Society. Since 2021, she has been deputy director of the CNRS french research consortium in Chemical Biology (GDR-ChemBio). She is strongly involved in trainings with interdisciplinary programs and created a CNRS Thematic School of Chemical Biology in 2021.

Christophe Biot obtained his PhD in bioorganometallics in 1998 under the supervision of Pr. J. Brocard. Following several post-doctoral positions, he joined the University of Lille as an assistant professor in 2004. He is currently a Professor and leads the Chemical GlycoBiology (CheGB) team at the Unity of Structural and Functional Glycobiology. The team comprises scientists with diverse expertise ranging from organic synthesis to cell biology, and their research topics encompass a broad range of significant areas in chemical biology. Biot has published more than 130 research articles and 15 book chapters. Christophe Biot is the director of the GDR ChemBio.

Comptes Rendus
Chimie
2023, Vol. 26, Special Issue S3, p. 5-8
https://doi.org/10.5802/crchim.257

Foreword

Breaking Barriers in Chemical Biology – Toulouse 2022

Foreword—Breaking barriers in chemical biology: the innovative tour de force of the ChemBio GDR

Marie Lopez [*,a], Elisabetta Mileo [b], Eric Defrancq [c], Agnès Delmas [d], Boris Vauzeilles [e], Dominique Guianvarc'h [f] and Christophe Biot [*,g]

[a] CNRS-Université de Montpellier-ENSCM UMR 5247, Institut des Biomolécules Max Mousseron (IBMM), 34296 Montpellier, France

[b] Aix Marseille Univ, CNRS, BIP, Bioénergétique et Ingénierie des Protéines, IMM, Marseille, France

[c] UMR 5250, Département de Chimie Moléculaire, Université Grenoble Alpes, Grenoble, France

[d] CNRS UPR4301, Centre de biophysique Moléculaire (CBM), Orléans, France

[e] Université Paris-Saclay, CNRS, Institut de Chimie des Substances Naturelles, UPR 2301, 91198, Gif-sur-Yvette, France

[f] Université Paris-Saclay, CNRS, Institut de Chimie Moléculaire et des Matériaux d'Orsay, UMR 8182, 91405, Orsay, France

[g] Université de Lille, CNRS, UMR 8576, UGSF, Unité de Glycobiologie Structurale et Fonctionnelle (UGSF), F-59000, Lille, France

E-mails: marie.lopez@cnrs.fr (M. Lopez), emileo@imm.cnrs.fr (E. Mileo), eric.defrancq@univ-grenoble-alpes.fr (E. Defrancq), agnes.delmas@cnrs-orleans.fr (A. Delmas), boris.vauzeilles@cnrs.fr (B. Vauzeilles), dominique.guianvarch@universite-paris-saclay.fr (D. Guianvarc'h), christophe.biot@univ-lille.fr (C. Biot)

Abstract. Chemical Biology can be defined as the design and the development of molecular tools to decipher or modulate biological processes of interest so that they can be better understood, controlled or modified. It also consists in the observation and analysis of these molecular tools within, and in interaction with, their complex biological environment. Chemical Biology can then lead to highly valuable basic knowledge and be applied in diverse fields such as health and environment. Therefore, Chemical Biology has strong interactions with the design of therapeutic strategies, diagnosis, agrochemicals or Ecology.

Keywords. Interdisciplinary research, Molecular tools, Biological environment, Translational applications, Innovative strategies.

[*] Corresponding authors

1. ChemBio GDR genesis

In 2018, the Deputy Scientific Director of the National Centre for Scientific Research (CNRS) mandated us to organize a scientific event to assess the expertise of French scientific community in different research and application areas of Chemical Biology, as part of the scientific policy of the CNRS-Institute of Chemistry (INC) and the "Convergence Actions" program. The events, called the Scientific Days of Chemical Biology, were held in Lille in 2018 and then Nice in 2020. It was an opportunity for a prospective reflection, which resulted in January 2021 in the creation of the GDR 2095—ChemBio directed by Professor Christophe Biot (UGSF, Lille). The aim of the ChemBio GDR is to bring together scientists from the French Chemical Biology community, which includes about 80 laboratories and 600 members, making it the second largest GDR supported by the CNRS-INC. Aiming to support exchanges of scientific expertise and skills for chemists and biologists, ChemBio GDR is a fertile place for discussion leading to the emergence of new key scientific questions.

Just after its creation, the first ChemBio GDR scientific meeting was held in Grenoble in October 2021, and the second one took place in Toulouse in June 2022. The upcoming scientific meetings will be held in Strasbourg in June 2023 and in Bordeaux in 2024. In addition to these annual days dedicated to scientific exchange through oral, poster presentations, and round tables on various topics, the ChemBio GDR supports and is closely associated to the CNRS ChemBio thematic school directed by Professor Dominique Guianvarc'h (Paris Saclay University, ICMMO). This intensive one-week formation was first organized in November 2021 in Le Touquet. Due to the success of the first edition, a second edition, also supported by the ChemBio GDR, will be organized in November 2023.

At the same time, in September 2019, considering the importance and the impact of the French Chemical Biology, the Thematic Group of Chémobiologie (SCF-ChemBio) was created within the French Chemical Society (SCF). This group currently has more than 500 members and is presided by Dr. Boris Vauzeilles (DR CNRS, ICSN). SCF-ChemBio collaborates with ChemBioChem (Chemistry Europe, Wiley) and the International Chemical Biology Society (ICBS) with which two e-symposia were co-organised in April 2021 and July 2021. Besides, in 2021, SCF-ChemBio also joined the European Federation for Medicinal chemistry and Chemical biology (EFMC). In 2021, SFBBM (French Society for Biochemistry and Molecular Biology) has introduced a new session entitled "*Chémobiologie*" for its annual meeting that takes place every early July.

2. ChemBio GDR objectives

The ChemBio GDR has several objectives aimed at unifying Chemical Biology research in France and increasing its visibility and recognition (Box 1). One of the primary objectives is to connect researchers and teams across the country and establish a network that fosters the sharing of experiences, skills, and the launch of new research dynamics. GDRs are valuable coordination structures that have proven to be successful for several years.

Collaboration is another strength of the ChemBio GDR, and the complementary expertise and knowledge of participating researchers and teams contribute to the creation of fruitful collaborations and international projects. The ChemBio GDR is also involved in student training, including the organization of scientific workshops and on-site training through its platforms. The ChemBio GDR supports short-term scientific missions for PhD students between partner laboratories to acquire specific expertise and/or techniques.

The ChemBio GDR has a further objective of increasing awareness about technology transfer among researchers, especially those who are at the early stages of their careers.

Moreover, the ChemBio GDR also seeks to promote creativity and expand the borders of scientific culture by promoting interactions between national artists and ChemBio GDR researchers. This initiative encourages innovative methods for the progression of knowledge and subsequent transfer to society. By exploring the aesthetic dimensions of Chemical Biology and evaluating the outcome from the perspective of the non-scientific community, members can improve media diffusion practices and encourage a shift in practices based on scientific challenges towards the artistic community.

Box 1—ChemBio GDR is articulated around 3 thematic axes

(1) Chemical targeting and modulation, understanding of biological processes: By developing molecular tools that can scan and modulate biological processes of interest, researchers in Chemical Biology are able to better understand and control those processes for a variety of applications.
(2) Chemical tools and molecular approaches: Using a range of chemical tools and molecular approaches, researchers in Chemical Biology are able to design and develop new compounds and therapies for a wide range of applications, from drug discovery to agriculture.
(3) Physicochemical technologies: Advancements in physicochemical technologies enable researchers in Chemical Biology to better study the complex interactions between molecules and biological systems, leading to new discoveries and breakthroughs in a variety of fields.

3. 2nd ChemBio GDR scientific days

The 2nd ChemBio GDR Scientific Days took place on June 8th and 9th, 2022 at the University of Toulouse III. This 2022 edition was closely related to the annual Chemistry–Biology–Health day in Toulouse, held on June 10th, 2022, which demonstrated the impact of chemical biology on the study and treatment of human diseases. The event had 148 registrations and spanned two days and was the opportunity to attend three plenary conferences, 14 lectures, 11 flash-talks, and two poster sessions. This special edition of "Comptes Rendus Chimie" is an opportunity to share with the scientific community the multidisciplinary research of the speakers who presented at this 2nd ChemBio GDR Scientific Days.

In this special issue on French Chemical Biology, we bring together a collection of cutting-edge research articles and reviews that highlight recent advances and exciting new directions in the field. It should be emphasized that the selection for this special issue was limited to the speakers who presented their work in Toulouse. Therefore, only the articles/reviews from those speakers have been chosen for publication.

Conflicts of interest

Authors have no conflict of interest to declare.

Director
Christophe Biot (Université de Lille)
christophe.biot@univ-lille.fr

Deputy directors
Dominique Guianvarc'h (Université Paris Saclay)
Boris Vauzeilles (CNRS)

Axis 1—Chemical targeting and modulation, understanding of biological processes
Marie-Pierre Rols (Toulouse)
Hélène Munier-Lehmann (Paris)
Ludger Johannes (Paris)
Laurence Lafanechère (Grenoble)

Axis 2—Chemical tools and molecular approaches
Eric Defrancq (Grenoble)
Marie Lopez (Montpellier)
Sylvain Routier (Orléans)

Axis 3—Physicochemical technologies
Agnès Delmas (Orléans)
Carine Van-Heijenoort (Paris Saclay)
Vladimir Torbeev (Strasbourg)

Webmaster
Elisabetta Mileo (Marseille)
Jean-Marc Sache (Grenoble)
Web site
www.gdr.chemobiologie.cnrs.fr/

Comptes Rendus
Chimie
2023, Vol. 26, Special Issue S3, p. 9-16
https://doi.org/10.5802/crchim.247

Breaking Barriers in Chemical Biology – Toulouse 2022

Quantitative analysis of the effect of microtubule-targeting drugs on the microtubule cytoskeleton of breast cancer cells with different invasive properties

Sophie Michallet [a], Lauriane Bosc [a] and Laurence Lafanechère [*, a]

[a] Université Grenoble Alpes, INSERM U1209, CNRS UMR5309, Institute for Advanced Biosciences, Team Cytoskeleton Dynamics and Nuclear Functions, 38000 Grenoble, France

E-mails: Sophie.michallet@univ-grenoble-alpes.fr (S. Michallet), Lauriane.bosc@univ-grenoble-alpes.fr (L. Bosc), Laurence.lafanechere@univ-grenoble-alpes.fr (L. Lafanechère)

Abstract. The characterization of microtubule-targeting drugs at the cellular level is an essential step in the development of drugs targeting the microtubule network. To that aim, we have previously developed a quantitative cell-based assay easy to perform in microplates that requires only a luminescence reader and no microscopic analysis. Here, we show that this assay can be easily adapted to different breast cancer cell lines. An ideal application of this test could be the comparative analysis of the response of human tumor samples to different microtubule targeting drugs, to optimize therapeutic treatment.

Keywords. Taxanes, Vinca-alkaloids, Combretastatin-A4, Microtubules, Quantitative cell-based assay, Breast cancer.

Manuscript received 28 February 2023, revised 12 June 2023 and 19 July 2023, accepted 19 July 2023.

1. Introduction

Microtubules (MTs) are dynamic cytoskeletal polymers that are involved in intracellular trafficking, cell shape establishment, cell movements and the segregation of condensed chromosomes during mitosis. They are composed of α–β tubulin heterodimers and their polymerization exhibits non-equilibrium dynamics, characterized by periods of polymerization and of depolymerization.

The organization and stability of MTs is tightly regulated by numerous cellular factors such as XMAP215/Dis1/TOGp, MCAK, MAP4, end-binding proteins or Op18/stathmine [1–3]. Targeted perturbation of this finely tuned process constitutes a major therapeutic strategy. Drugs that interfere with tubulin and MTs are, indeed, key components of combination chemotherapies for the treatment of carcinomas [4–6]. Vinca-alkaloids and taxanes are among the microtubule targeting drugs (MTD) commonly used in cancer chemotherapy [4,7].

* Corresponding author

Drugs that interfere with the tubulin/MT system are roughly classified into MT stabilizing agents such as taxanes, and MT destabilizing agents, such as combretastatin and vinblastine. Regarding MT destabilizing agents, different binding sites located on the β-tubulin subunit have been identified for colchicine, vinca-alkaloids and maytansine [7]. The colchicine site is located at the intradimer interface [8–10]. It is a large site, mostly buried in the β-tubulin subunit. Besides colchicine, numerous drugs bind to this site, such as nocodazole or combretastatin, but no ligand is large enough to occupy the entire site [6]. Colchicine-site ligands destabilize MTs by preventing the curved-to-straight conformational transition within the α–β tubulin heterodimer [7,11]. The vinca site is located at the inter-dimer interface between two longitudinally aligned tubulin dimers. Vinca-site agents destabilize MTs by introducing a wedge at the interface between two longitudinally aligned tubulin dimers at the tip of MTs, or by stabilizing assembly-incompetent ring-like oligomers of tubulin [7,12]. On the contrary, taxanes bind to a pocket of β-tubulin located on the luminal side of microtubules [7,13,14].

Recently, we have described two original assays that allow the quantitative evaluation of the depolymerizing or stabilizing effect of MT targeting drugs (MTDs) on cellular MTs [15,16]. These assays are different from the biochemical tests classically performed *in vitro* on purified tubulin. They have the advantage of providing information on the effects of MTDs in a cellular context. They allow the quantification of the cell MT mass which is measured by immunoluminescence after the use of a particular lysis buffer, which eliminates free tubulin and preserves intact cellular MTs. These assays are simple to use since they are based on a luminescence reading. Thus, for these assays, only a microplate reader is required and no microscopic analysis is needed.

We have recently conducted a systematic comparative analysis of the effect of four well-characterized MTDs on the kinetics of *in vitro* tubulin assembly, on the quantity of cellular MTs in HeLa cells (a cervical cancer cell line) using our assay and on HeLa cell viability. We found that there was no significant correlation between the activity of the different drugs on tubulin assembly *in vitro* and their activity in cells, indicating that the effect of drugs on pure tubulin may be different from their effect on MTs in the cell context. In contrast, we observed a striking similarity between the profile of the viability curves and that of the curves measuring the stability of cellular MTs. This strongly suggests that the cytotoxic effect of the drugs is due to their depolymerizing effect on cellular MTs [15]. These results highlight the value of these cellular microtubule assays for predicting drug therapeutic activity.

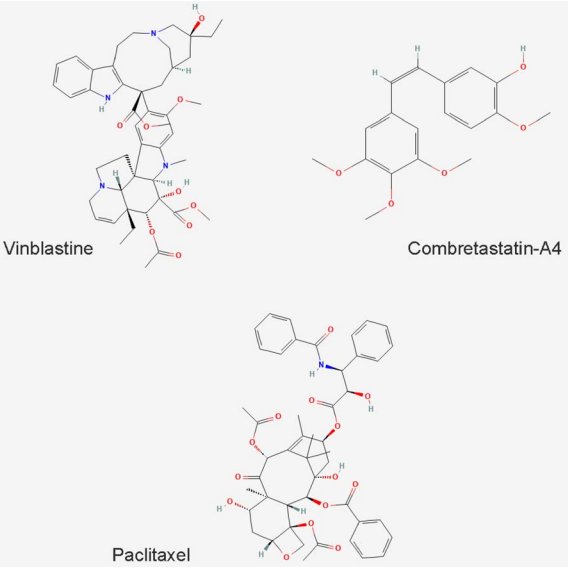


Figure 1. Chemical structure of the MTDs used in this study.

Here, we used these cellular MT assays to compare the effect of MTDs, used in cancer chemotherapy, on 3 human breast cell lines known to differ in invasiveness. The cell lines used were MCF-10A cells, isolated from human fibrocystic breast tissue [17], MCF-7, a poorly invasive cell line [18], and MDA-MB-231, an invasive breast cancer cell line [19] (Table 1).

The assayed drugs (Figure 1) are two MT depolymerizing drugs, vinblastine and combretastatin-A4 (CA-4) and a MT stabilizing drug, paclitaxel (PTX).

Vinblastine binds to the vinca site of tubulin, whereas CA-4 binds to the colchicine site and PTX to the taxane site.

Although these compounds all target microtubules, they differ in their applications: whereas vinblastine and PTX are used in antitumor chemotherapy [4,7,20], CA-4 is an antineoangiogenesis agent [21].

Table 1. Characteristics of the cell lines used in this study

Cell line	MCF-10A	MCF-7	MDA-MB-231
Disease	Fibrocystic disease	Adenocarcinoma	Adenocarcinoma
Origin of cells	Spontaneously immortalized cell line	Metastasis (pleural effusion)	Metastasis (pleural effusion)
Cell type	Epithelial cell	Epithelial cell	Epithelial cell
Phenotype	Normal-like	Luminal A	Claudin-low
ER	No	Yes	No
PR	No	Yes	No
HER2 amplification	No	No	No
Tumorigenic in mice	Non-tumorigenic	Yes, with estrogen supplementation	Yes
Appearance	Groups of adherent epithelial cells	Loosely attached three-dimensional clusters	Epithelial-like somewhat spindle-shaped
Model	Non-malignant breast cell line	Transformed estrogen responsive breast cancer cell line	Late-stage breast cancer

ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2.

2. Material and methods

2.1. *Chemical reagents and cells*

MCF-10A (CRL-10317), derived from fibrocystic breast tissue, and MCF-7 (HTB-22) and MDA-MB-231 (HTB-26) cells, derived from the pleural effusion of a metastatic breast adenocarcinoma, were obtained from the American Type Culture Collection (ATCC, Gainthersburg, MD, USA). Cell culture was performed according to the recommendations of the American Type Cell Culture. MCF-10A cells were grown in DMEM/F12 (Gibco 10565-018) supplemented with 5% horse serum (Gibco 16050-122), EGF (Peprotech AF-100-15) at 0.1 mg/ml, hydrocortisone (H0888) at 1 mg/ml, cholera toxin (C8052) at 1 mg/ml, insulin (I1882) at 10 mg/ml and 1% penicillin/streptomycin. MCF-7 cells were grown in MEM medium (Gibco 31095-029) supplemented with insulin at 0.1 mg/ml, 1% penicillin/streptomycin and 10% FBS. MDA-MB-231 cells were grown in DMEM Glutamax medium (Gibco 31966-021) and supplemented with 10% FBS and 1% penicillin/streptomycin. Cells were maintained in a humid incubator at 37 °C in 5% CO_2.

All media were purchased from Gibco Invitrogen (Carlsbad, CA, USA). All chemicals, except those for which it is specified, were purchased from Sigma-Aldrich (Saint-Quentin-Fallavier, France).

Combretastatin-A4 (C7744), vinblastine (V1377) and paclitaxel (T7402) were prepared at a 10 mM stock solution in Dimethyl sulfoxide (DMSO, #4540) aliquoted and stored at −20 °C.

2.1.1. *Quantitative assay of the effect of depolymerising agents on the cellular microtubule content*

Cells were seeded in 96-well microplates (#655086, Greiner bio One, Courtaboeuf, France) at the density of 5000 MCF-10A cells, 10,000 MCF-7 cells and 7500 MDA-MB-231 cells per well in 100 µl of complete medium and then incubated at 37 °C in 5% CO_2 for 24 h. Cells were then treated for 30 min at 37 °C with the compounds at concentrations ranging from 1 to 5000 nM (1 microplate per molecule, 1 concentration per column), with 0.1% DMSO used as positive control (6 wells per microplate). After medium aspiration, treated cells were permeabilized for 10 min using 100 µl per well of warmed (37 °C) OPT buffer (80 mM Pipes, 1 mM EGTA, 1 mM $MgCl_2$, 0.5% Triton X-100, and 10% glycerol, pH 6.8). Cells were fixed for 10 min at room temperature using 100 µl per well of 4% formaldehyde (Sigma Aldrich, #252549, Saint-Quentin-Fallavier, France) in PBS. Cells were washed 3 times in PBS (100 µl per well), then 50 µl of primary

anti-alpha-tubulin antibody (clone α3A1 [22], 1:5000 in PBS pH 7.4, 0.1% Tween-20 2% Bovine Serum Albumin (BSA)) was added for 45 min. Cells were washed twice again and secondary anti-mouse antibody coupled to HRP (1:2000 in PBS pH 7.4, 0.1% Tween-20 2% BSA, #715-035-150, Jackson Immuno-Research Laboratories, Cambridgeshire, UK) was added for 45 min. Then, cells were washed again with PBS and 100 µl of ECL substrate (#170-5061, Bio-Rad Laboratories Inc., USA) were injected in each well using the FLUOstar OPTIMA Microplate Reader (BMG Lab technology, Champagny-sur-Marne, France). The luminescent signal was read 5 min after ECL injection. IC50s, i.e., drug concentrations able to reduce the amount of cellular microtubules by half, were calculated for each independent experiment using GraphPad Prism software and are presented in the text as means ± SEM.

2.1.2. *Quantitative assay of the stabilizing effect of paclitaxel on cellular microtubule*

The same quantity of cells was seeded in microplates as for the quantitative assay of the effect of depolymerizing agents on the cellular microtubule. 24 h after seeding, cells were treated with paclitaxel at different concentrations (range 0–5000 nM). DMSO alone at 0.25% was used as positive controls (6 wells per microplate and per control), respectively. Then combretastatin-A4 was added to each well at 50 nM final concentration for 30 min. After medium aspiration, treated cells were permeabilized for 10 min using 100 µl per well of OPT buffer (80 mM Pipes, 1 mM EGTA, 1 mM $MgCl_2$, 0.5% Triton X-100, and 10% glycerol, pH 6.8) pre-warmed to 37 °C. After buffer aspiration, cells were fixed for 10 min at room temperature using 100 µl per well of 4% formaldehyde in PBS pH 7.2. Cells were washed 3 times with PBS 0.1% Tween-20 (150 µl per well), then 50 µl of α3A1 anti-tubulin antibody (1:5000 in PBS pH 7.4, 0.1% Tween-20 2% BSA) were added for 45 min. After washing of cells as described above, 50 µl of anti-mouse antibody coupled to HRP (1/2000 in PBS pH 7.4, 0.1% Tween-20 2% BSA) were added for 45 min. Then cells were washed again and, 50 µl of ECL Western blotting substrate (Pierce #32106) were added to each well and the luminescent signal was read after 5 min of incubation.

2.1.3. *Immunofluorescence*

Cells at a density of 30,000 cells per well for MCF-10A, 60,000 for MCF-7 and 50,000 for MDA-MB-231 were grown for 48 h on glass coverslips placed in a 24-well microplate. When cells reached 70% confluence, the medium was replaced with a fresh one supplemented with DMSO (0.005% or 0.01%) or the test compound at 50 nM for vinblastine and CA-4 or 1000 nM for PTX. After 30 min of incubation, cells were permeabilized in warm OPT buffer (80 mmol/L Pipes, 1 mol/L EGTA, 1 mol/L $MgCl_2$, 0.5% Triton X-100 and 10% Glycerol, pH 6.8) and fixed for 6 min in −20 °C methanol (Carlo ERBA SAS, #414855, Val-de-Reuil, France). After washing and saturation with a specific blocking buffer (3% BSA), 10% Goat serum (Gibco Invitrogen, #16210064, Carlsbad, CA, USA) in PBS, cells were incubated for 45 min at room temperature (RT) with anti-alpha-tubulin antibody (clone α3A1 [22] in blocking buffer). Cells were washed twice again and subsequently incubated with Alexa 488 conjugated anti-mouse antibody (1:500 in blocking buffer, #115-545-166, Jackson immune-research laboratory, Cambridgeshire, UK) for 30 min at RT. Coverslips were mounted on glass slides with Moviol 4–88.

Images were captured with a Zeiss AxioimagerM2 microscope equipped with the acquisition software AxioVision (Marly-le-Roi, France).

3. Results

3.1. *Effect of the depolymerizing agents Vinblastine and CA-4 on the MT contents of breast cancer cell lines*

We first compared the effect of different doses of vinblastine and CA-4 on cellular interphase MTs. After having determined a seeding density adapted to each cell line, we used our recently developed cell-based assay that quantifies intact MTs in cells [15,16]. The principle of this assay is based on the use of a particular lysis buffer, which eliminates free tubulin and preserves intact cellular MTs, which are then quantified by immunoluminescence. The detailed protocol is given in the methods section.

As shown on Figure 2, both compounds induced a dose-dependent depolymerization of cellular MTs, on the 3 cell lines. The IC50 for CA-4 was in the

same range for MCF-10A and MDA-MB-231, i.e., 6.6 ± 0.4 nM and 4.8 ± 0.3 respectively. This IC50 was found much higher in MCF-7 cells, i.e., 31.1 ± 7.9 nM. Similar results were obtained for vinblastine: the IC50 was 2.8 ± 0.4 for MCF-10A, 67.7 ± 15.4 for MCF-7 and 4.4 ± 0.4 for MDA-MB-231.

Interestingly, in MCF-10A and MDA-MB-231 cells, the sigmoid curve of CA-4 has a much steeper slope than that of vinblastine, indicating that a small variation of the concentration around the IC50 can lead to a complete depolymerization of the microtubule network. The effect of increasing concentrations of vinblastine is much more gradual on these cell lines.

On MCF-7 cells, both compounds behave similarly. On this cell line, unlike the other two cell lines, a complete depolymerization of the MT network was not observed, even at the highest doses (5000 nM) assayed.

3.2. *Effect of the stabilizing agent paclitaxel on the microtubule contents of breast cancer cell lines*

To measure the effect of PTX on cellular MTs, we used a variant of the test presented above for depolymerizing agents. The assay probes a PTX-induced resistance of the MT network to a depolymerization provoked by CA-4. CA-4 binds free tubulin dimers and prevents their incorporation into MTs, leading to a progressive loss of the polymerized MT network. Stabilized MTs with slow dynamics have reduced exchanges of their tubulin content with the free tubulin pool, and are thus less sensitive to CA-4-induced depolymerization [23]. After treatment with the drugs and before fixation, the same lysis buffer as described above is used in order to eliminate free tubulin and to preserve the cellular MTs, which are then quantified by immunoluminescence. The detailed protocol is given in the methods section. This assay is sensitive as it allows to detect a weak stabilizing effect, quantitative and statistically robust [16].

Figure 3 shows the dose-effect curves of the sensibility of the MT network of the different cell lines to the stabilizing effect of PTX against a depolymerization induced by 50 nM CA-4. This concentration of CA-4 is able to induce the depolymerization of the entire MT network in MCF-10A and MDA-MB-231 cells and only 50% of the network in MCF-7 cells (see blue curves in Figure 2, and ordinate at the origin in Figure 3). For MCF-10A and MDA-MB-231 cells, a similar dose of PTX, 246.9 ± 83.9 nM and 234.3 ± 32.5 nM respectively, is required to achieve 50% stabilization of the microtubule network. Regarding MCF-7, 16.3 ± 9.0 nM of PTX induced 50% of the effect. For the latter cell line, we tested the stabilizing effect of PTX against a depolymerization induced by 100 times more CA-4, i.e., 5000 nM (not shown). As shown in Figure 2, however, such a high concentration of CA-4 does not fully depolymerize the network. Under these conditions, a similar dose-dependent profile of stabilization by PTX is obtained, with 50% of the stabilizing effect achieved for a PTX concentration of 15.3 ± 6.9 nM.

3.3. *Immunofluorescence analysis of the effect of the different drugs on cellular MTs*

Using immunofluorescence, we checked that the luminescent values do reflect the state of cellular MTs. As shown in Figure 4, a decrease in the density of the MT network is observed in MCF-10A and MDA-MB-231 cells treated with 50 nM of vinblastine or CA-4, as well as typical microtubule bundle reorganization with 1 μM Paclitaxel treatment. Changes in the shape and density of the MT network are more difficult to observe in MCF-7 cells, which grow in clusters.

4. Discussion and conclusion

The development of agents targeting MTs remains an area of intense research. In this context, quantitative assessment of the compounds' effect on cellular MTs is essential. With this study, we showed that the quantitative assays we had initially developed using HeLa cells could be easily adapted, for other types of adherent cells by modifying only the density of cell seeding. This allows the investigation of the effect of MTDs on other types of cancer. However, this study highlighted a potential limitation concerning the cell types that can be studied. For example, we could not observe complete depolymerization of the microtubular network with two different drugs on MCF-7 cells. MCF-7 cells are known as drug-sensitive cells, which do not overexpress drug efflux pumps [24]. The reduced efficiency of depolymerizing drugs to act on the microtubule network is therefore not the

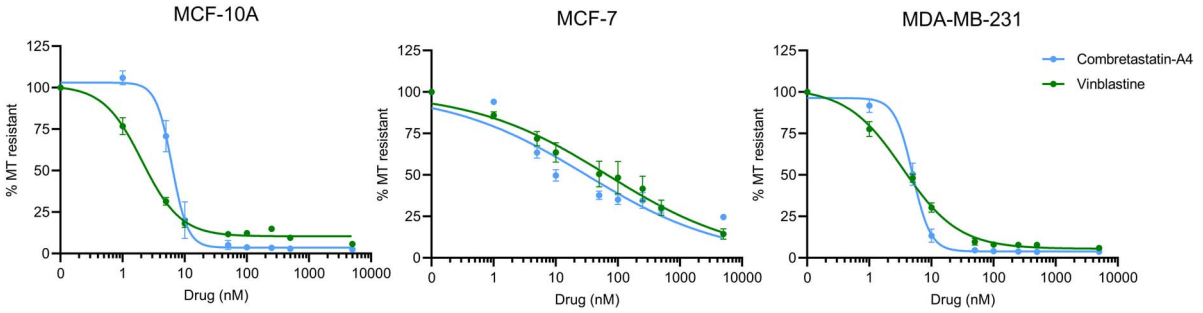


Figure 2. Comparison of the MT destabilizing effect of Vinblastine and CA-4. Different doses of compounds were applied to the different cells in microplates and their MT destabilizing effect was assessed after a 30 min incubation, using the luminescent assay, as described in the Materials and Methods section. Results are expressed as % of resistant MTs, with 100% corresponding to cells treated with DMSO only, without the depolymerizing agent. Datapoints are means ± SEM from three independent experiments.

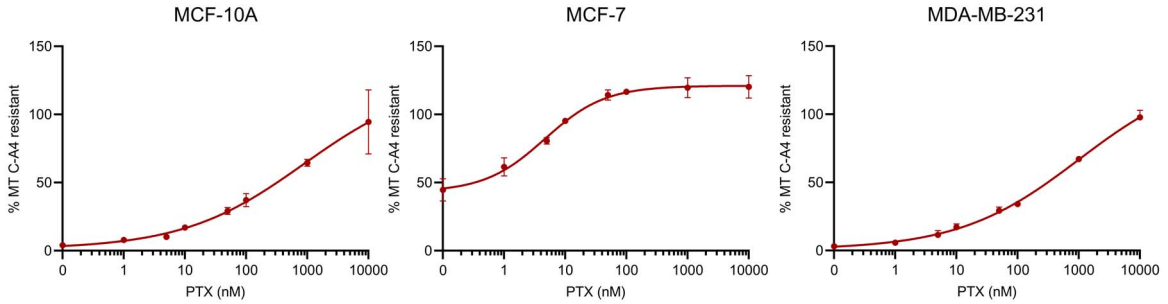


Figure 3. Comparative analysis of the effect of PTX on MT dynamics in breast cancer cells. Different doses of PTX were applied for 2 h to cells in microplates. Thirty minutes before the end of the PTX treatment 50 nM CA-4 was added. At the end of the PTX treatment, its MT stabilizing effect was assessed using the luminescent assay described in the material and methods section. The analyses were performed with a PTX concentration in 0.1% DMSO at the solubility limit for not inducing an additional toxicity due to DMSO. Results are expressed as % of MTs resistant to CA-4-induced depolymerization, with 100% corresponding to cells treated with DMSO without CA-4. Datapoints are means ± SEM from three independent experiments.

result of the activation of efflux pumps. One possible reason is that this is a result of poor penetration of the drugs, even when applied at high concentrations, due to the fact that the cells grow in clusters. Increasing the incubation time of the cells with the drugs, which implies additional methodological validations, could be a way to achieve complete MT depolymerization.

Despite this difficulty in achieving complete depolymerization of MTs in MCF-7 cells, this cell line appears to be particularly sensitive to the stabilizing action of PTX, as measured in our assay. Indeed, it requires about 15 times more PTX to stabilize 50% of the MT network for MCF-10A and MDA-MB-231 cells.

We observed that MCF-10A and MDA-MB-231 show similar sensitivity to MTDs, indicating that the invasiveness of MDA-MB-231 does not influence the drug sensitivity of this cell line. This cell line was established from a single sample of pleural effusion obtained from a 51-year-old woman, who had previously a right radical mastectomy for a poorly differentiated tumor. This patient received a systemic treatment with 5-FU and prednisone which were ineffective. Then a combined chemotherapy (cyclophosphamide, adriamycin, and amethopterin)

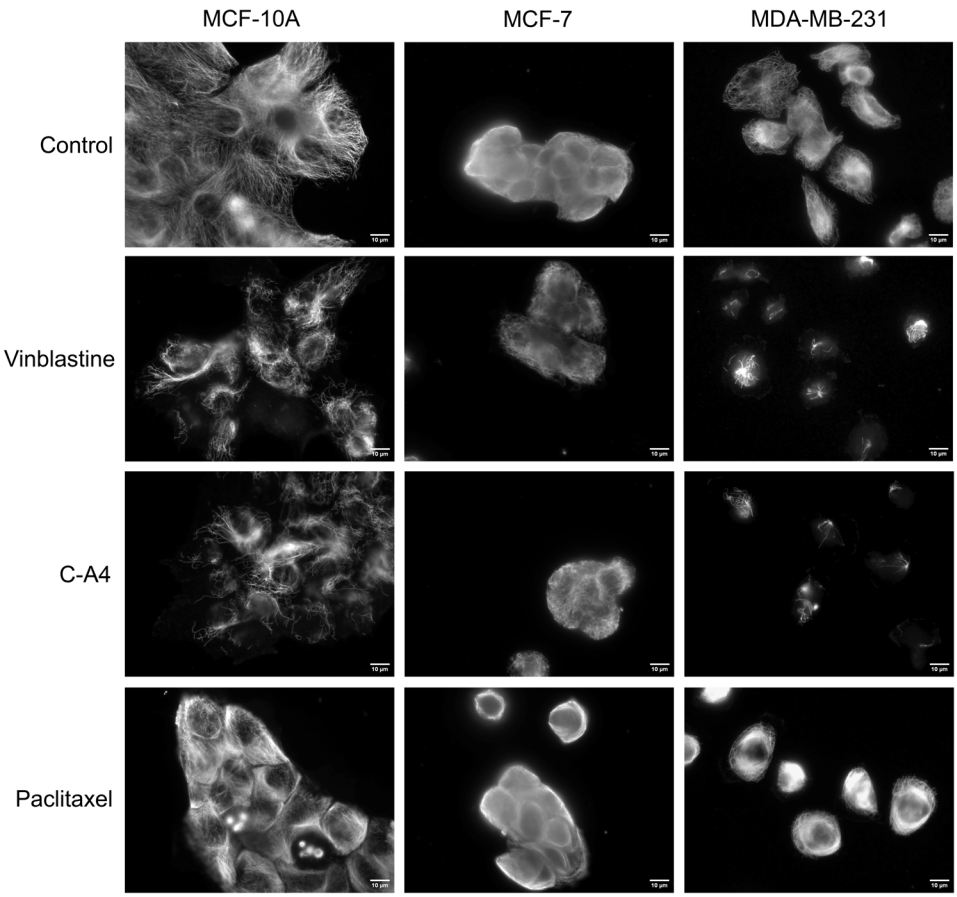


Figure 4. Immunofluorescence analysis of the effect of the different compounds on the MTs of the different breast cancer cells. Cells were incubated for 30 min with 50 nM Vinblastine, 50 nM CA-4, 2 h with 1 μM PTX or with DMSO (control). Cells were then permeabilized with OPT buffer, fixed, and processed for immunofluorescence using an anti-α-tubulin antibody. Scale bars, 10 μm.

was administered, with no success [19]. Interestingly, this patient has never been treated with drugs targeting MTs, and therefore could not develop resistance, explaining the sensitivity of MDA-MB-231 to MTDs.

It should be noted that the test duration is very short (2 h) and therefore has no detectable effect on cell viability. A limitation of the assay for measuring the effect of depolymerizing agents is that a decrease in luminescent signal can be observed not only when microtubules are depolymerized, but also if the agents tested have an effect on cell adhesion, thus reducing the number of stained cells. False positives of this kind can be identified by checking the effect of the drugs on the cells under the microscope. In this case, measuring the value of the luminescent signal in ratio to the number of residual adherent cells, as measured by nucleus labelling, e.g. with Hoechst, can provide information on the depolymerizing effect alone [25].

Here, we have shown that the quantitative analysis of the effect of MTDs can be easily performed on various cell lines. Ideally, this assay could be used not only for the development of new MTDs [26,27], but also for the comparative analysis of the response of human tumor samples to different drugs, in order to optimize therapeutic treatment.

Conflicts of interest

Authors have no conflict of interest to declare.

References

[1] R. Mohan, E. A. Katrukha, H. Doodhi, I. Smal, E. Meijering, L. C. Kapitein, M. O. Steinmetz, A. Akhmanova, J. R. Mcintosh, *Proc. Natl. Acad. Sci. USA*, 2013, **110**, 8900-8905.

[2] S. S. Andersen, *Trends Cell Biol.*, 2000, **10**, 261-267.

[3] A. Lieuvin, J. C. Labbé, M. Dorée, D. Job, *J. Cell Biol.*, 1994, **124**, 985-996.

[4] C. Dumontet, M. A. Jordan, *Nat. Rev. Drug Discov.*, 2010, **9**, 790-803.

[5] M. Kavallaris, *Nat. Rev. Cancer*, 2010, **10**, 194-204.

[6] M. J. Pérez-Pérez, E. M. Priego, O. Bueno, M. S. Martins, M. D. Canela, S. Liekens, *J. Med. Chem.*, 2016, **59**, 8685-8711.

[7] M. O. Steinmetz, A. E. Prota, *Trends Cell Biol.*, 2018, 776-792.

[8] K. H. Downing, E. Nogales, *Eur. Biophys. J*, 1998, **27**, 431-436.

[9] A. Giraudel, L. Lafanechere, M. Ronjat, J. Wehland, J. R. Garel, L. Wilson, D. Job, *Biochemistry*, 1998, **37**, 8724-8734.

[10] R. B. G. Ravelli, B. Gigant, P. A. Curmi, I. Jourdain, S. Lachkar, A. Sobel, M. Knossow, *Nature*, 2004, **428**, 198-202.

[11] A. Dorléans, B. Gigant, R. B. G. Ravelli, P. Mailliet, V. Mikol, M. Knossow, *Proc. Natl. Acad. Sci. USA*, 2009, **106**, 13775-13779.

[12] B. Gigant, C. Wang, R. B. Ravelli, F. Roussi, M. O. Steinmetz, P. A. Curmi, A. Sobel, M. Knossow, *Nature*, 2005, **435**, 519-522.

[13] A. E. Prota, K. Bargsten, D. Zurwerra, J. J. Field, J. F. Díaz, K.-H. Altmann, M. O. Steinmetz, *Science*, 2013, **339**, 587-590.

[14] T. Mühlethaler, D. Gioia, A. E. Prota, M. E. Sharpe, A. Cavalli, M. O. Steinmetz, *Angew. Chem. Int. Ed. Engl.*, 2021, **60**, 13331-13342.

[15] M. C. Laisne, S. Michallet, L. Lafanechère, *Cancers (Basel)*, 2021, **13**, article no. 5226.

[16] S. Ramirez-Rios, S. Michallet, L. Peris, C. Barette, C. Rabat, Y. Feng, M. O. Fauvarque, A. Andrieux, K. Sadoul, L. Lafanechère, *Front. Pharmacol.*, 2020, **11**, article no. 543.

[17] H. D. Soule, T. M. Maloney, S. R. Wolman, R. Brenz, J. Russo, R. J. Pauley, R. F. Jones, S. C. Brooks, C. M. McGrath, *Cancer Res.*, 1990, **50**, 6075-6086.

[18] H. D. Soule, J. Vazquez, A. Long, S. Albert, M. Brennan, *J. Natl. Cancer Inst.*, 1973, **51**, 1409-1416.

[19] R. Cailleau, R. Young, M. Olivé, W. J. Reeves, *J. Natl. Cancer Inst.*, 1974, **53**, 661-674.

[20] L. Lafanechère, *Front. Pharmacol.*, 2022, **13**, article no. 969183.

[21] J. Griggs, J. C. Metcalfe, R. Hesketh, *Lancet Oncol.*, 2001, **2**, 82-87.

[22] L. Peris, M. Thery, J. Fauré, Y. Saoudi, L. Lafanechère, J. K. Chilton, P. Gordon-Weeks, N. Galjart, M. Bornens, L. Wordeman, J. Wehland, A. Andrieux, D. Job, *J. Cell Biol.*, 2006, **174**, 839-849.

[23] R. Prudent, E. Vassal-Stermann, C.-H. Nguyen, C. Pillet, A. Martinez, C. Prunier, C. Barette, E. Soleilhac, O. Filhol, A. Beghin, G. Valdameri, S. Honoré, S. Aci-Sèche, D. Grierson, J. Antonipillai, R. Li, A. Di Pietro, C. Dumontet, D. Braguer, J.-C. Florent, S. Knapp, O. Bernard, L. Lafanechère, *Cancer Res.*, 2012, **72**, 4429-4439.

[24] D. Muthiah, R. Callaghan, *Eur. J. Pharmacol.*, 2017, **815**, 127-137.

[25] E. Vassal, C. Barette, X. Fonrose, R. Dupont, E. Sans-Soleilhac, L. Lafanechere, *J. Biomol. Screen*, 2006, **11**, 377-389.

[26] S. Pecnard, O. Provot, H. Levaique, J. Bignon, L. Askenatzis, F. Saller, D. Borgel, S. Michallet, M.-C. Laisne, L. Lafanechère, M. Alami, A. Hamze, *Eur. J. Med. Chem.*, 2021, **209**, article no. 112873.

[27] C. Hauguel, S. Ducellier, O. Provot, N. Ibrahim, D. Lamaa, C. Balcerowiak, B. Letribot, M. Nascimento, V. Blanchard, L. Askenatzis, H. Levaique, J. Bignon, F. Baschieri, C. Bauvais, G. Bollot, D. Renko, A. Deroussent, B. Prost, M.-C. Laisne, S. Michallet, L. Lafanechère, S. Papot, G. Montagnac, C. Tran, M. Alami, S. Apcher, A. Hamze, *Eur. J. Med. Chem.*, 2022, **240**, article no. 114573.

Comptes Rendus
Chimie
2023, Vol. 26, Special Issue S3, p. 17-38
https://doi.org/10.5802/crchim.246

Breaking Barriers in Chemical Biology – Toulouse 2022

Synthesis of trehalose-based chemical tools for the study of the mycobacterial membrane

Emilie Lesur [a], Paulin Rollando [a], Dominique Guianvarc'h [a] and Yann Bourdreux [*, a]

[a] Université Paris-Saclay, CNRS, Institut de Chimie Moléculaire et des Matériaux d'Orsay, UMR 8182, 91405, Orsay, France

E-mails: emilie.lesur@free.fr (E. Lesur), paulin.rollando@universite-paris-saclay.fr (P. Rollando), dominique.guianvarch@universite-paris-saclay.fr (D. Guianvarc'h), yann.bourdreux@universite-paris-saclay.fr (Y. Bourdreux)

Abstract. Corynebacteriales including the causative agent of many diseases such as tuberculosis are known to be extremely resistant against external stress as well as to antibiotic treatments which is believed to be related to the singular architecture of their mycomembrane. Over the last decades, both bioorthogonal chemical reporters and fluorescent probes for the metabolic labeling of bacterial cell glycans were developed including several trehalose-based probes to study the dynamics of mycomembrane components. This review presents an exhaustive view on the reported syntheses of trehalose-based probes enabling the study of the mycomembrane biogenesis.

Keywords. Chemical reporter, Metabolic labeling, Bioorthogonal probe, Trehalose, Mycomembrane, Corynebacteriales.

Funding. Agence Nationale de la Recherche (PTMyco, grant NO. ANR-22-CE44-0005-03).

Manuscript received 4 April 2023, revised 3 July 2023, accepted 19 July 2023.

1. Introduction

Corynebacteriales belong to *Actinobacteria* and include numerus bacteria responsible of human diseases such as leprosy (*Mycobacterium leprae*), diphtheria (*Corynebacterium diphtheriae*) and tuberculosis (*Mycobacterium tuberculosis*). Nowadays, the latter is still responsible for more than 10 million infections per year [1]. In contrast, some Corynebacteriales are non-pathogenic, such as *Corynebacterium glutamicum* which is involved in the industrial production of glutamate [2]. All bacteria of this group possess an atypical cell envelope exceptionally resistant against external stress and antibiotics, this resistance being attributed to the unique organization of their cell envelope. In addition to the plasma membrane, this envelope is composed of a cross-linked peptidoglycan layer (PG) bounded to an arabinogalactan (AG) complex itself connected to long fatty acids, called mycolic acids (MA) through an ester linkage. The esterified arabinogalatan constitutes the inner part of the outer membrane, the so called mycomembrane (Figure 1).

Mycolic acids (MA) are very long chain (C30–C90) α-branched and β-hydroxylated fatty acids, with an *anti*-relationship between the β-OH and the α-ramification. Beyond the mycoloylated AG (AGM), the outer leaflet of the mycomembrane is composed of non-covalently attached glycolipids esterified with mycolic acids. These glycolipids are mainly trehalose monomycolate (TMM) and trehalose dimycolate (TDM). After being biosynthesized in the cytoplasm, mycolic acids are esterified to trehalose by the polyketide synthase Pks13, giving rise to TMM, and then transferred across the inner membrane by a transporter of the MmpL3 family. TMM is then pro-

* Corresponding author

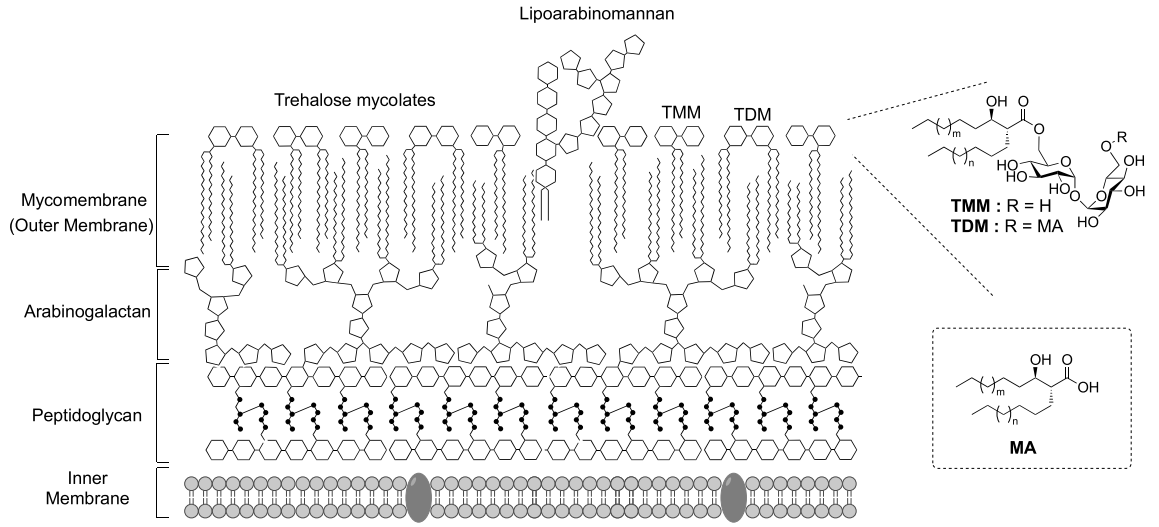


Figure 1. Schematic representation of the cell envelope of Corynebacteriales. (MA: mycolic acid; TDM: trehalose dimycolate; TMM: trehalose monomycolate.)

cessed by enzymes called mycoloyltransferases, and the mycolate part of TMM is transferred to different acceptors. Four mycoloyltransferases have been identified in *M. tuberculosis* (Ag85), six in *M. smegmatis* (Fbp) and six in *C. glutamicum* (Myt) [3]. TMM plays the role of a mycolate donor for several mycolate acceptors and the mycolate part can be transferred to (i) another TMM to form trehalose dimycolate (TDM); (ii) arabinogalactan; and (iii) channel-forming porins in some cases [4–6] (Figure 2).

Due to its importance for mycobacteria, the mycomembrane biogenesis is highly studied. Over the last decades, several chemical reporters have been developed to investigate biological processes in bacteria [7,8] and in this context, mycoloyltransferases have been widely targeted by trehalose- and TMM-based chemical reporters. Chemical labeling of the mycomembrane involves incorporation of chemical tools in the cell envelope and processing of these metabolic analogues by enzymes. The bacteria labeling can be performed in one or two steps, depending on the analogues used. It can be a detectable analogue of a mycomembrane metabolite, for example a mycomembrane precursor analogue conjugated with a fluorescent moiety. In such approach, the labeling is performed in one step allowing a direct mycomembrane imaging, but precautions should be taken in the design of such compounds since a bulky fluorophore could disrupt enzymatic processes. In another approach, the metabolic precursor analogues can be conjugated to a small bioorthogonal group, allowing the introduction of a detectable moiety in a distinct step using a bioorthogonal reaction with a fluorescent probe. Since the size of the bioorthogonal groups is small compared to fluorophores, this approach, called the bioorthogonal chemical reporter strategy, is supposed to be less disruptive for enzymes. From a metabolic point of view, the synthetic trehalose-based derivatives (Figure 2, dashed gray box at top) can be incorporated through the trehalose transporter LpqY-SugABC and then coupled to a mycolic acid and transferred into the mycomembrane, allowing specific labeling of the upper part of the mycomembrane, i.e., the TDM layer (Figure 2, red star). Alternatively, the TMM-based derivatives (Figure 2, dashed gray box at bottom), can be either directly inserted into the mycomembrane or processed by mycoloyltransferases, thus labeling both the inner and the outer part of the mycomembrane, AGM and TDM layers, respectively (Figure 2, green stars).

Trehalose is a non-reducing and C_2 symmetrical disaccharide composed of two glucose units with an $1,1,\alpha,\alpha$-linkage. It can be found in fungi, algae, insects, bacteria, and some invertebrates but not in mammalian. Trehalose has been extensively studied for its properties against stresses such as, for example, heat, oxidative stresses, or desiccation [9,

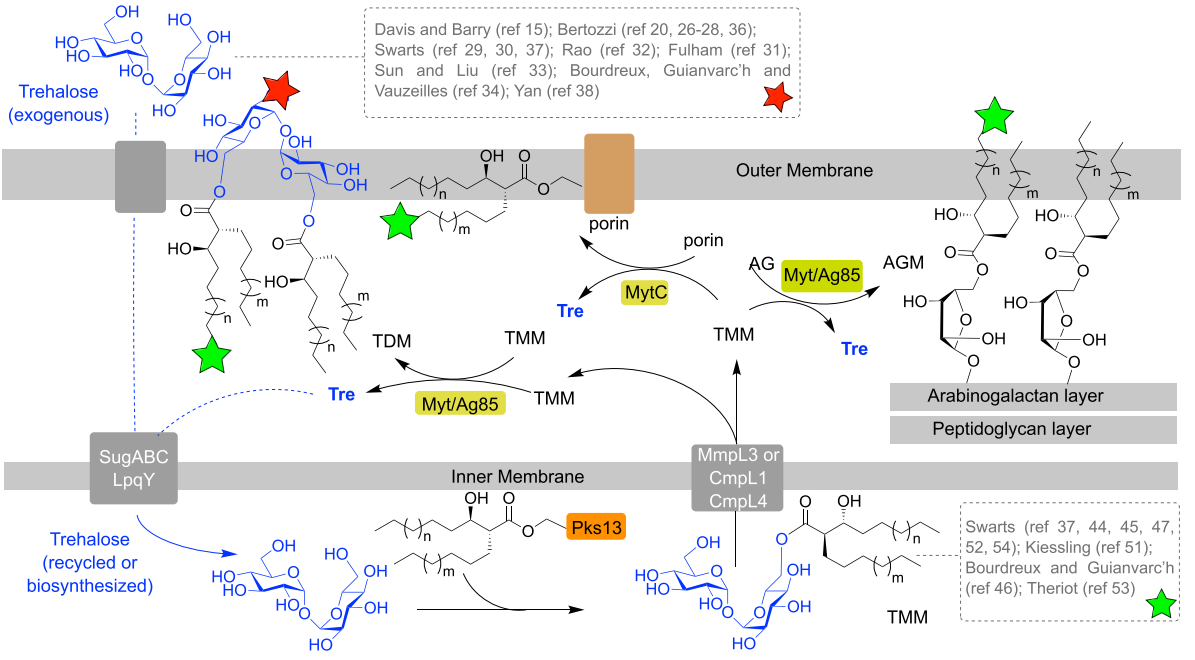


Figure 2. Schematic representation of the mycoloylation of various acceptors in the mycomembrane and labeling localization depending on the pathways used (green or red stars). (AG: arabinogalactan; AGM: mycoloylated arabinogalactan; Ag85: antigen 85; CmpL: corynebacterial membrane proteins large; MmpL: mycobacterial membrane proteins large; Myt: mycoloyltransferase; Pks: polyketide synthase; TMM: trehalose monomycolate; Tre: trehalose.)

10]. Due to its symmetry, and taking into account that many trehalose-containing glycoconjugates are unsymmetrical, the most important challenges in the synthesis of trehalose derivatives are its desymmetrization and its selective modification among its eight hydroxyl groups. Such transformations are clearly exposed in recent reviews and most of the syntheses rely on two different approaches [11–14]. The first approach relies on the formation of the trehalose core by an α,α-selective glycosylation, using selectively protected glucose derivatives. The second approach is based on the use of native trehalose as starting material and the selective post-modification of the eight hydroxyl groups. Both approaches are very challenging and were used for the selective synthesis of trehalose-based probes. In this review we will focus on the synthesis of trehalose- and TMM-based probes already used in bacteria for the study of the mycomembrane. We classified the probes in two categories: (i) the trehalose-based probes in which the detectable moiety or the bioorthogonal small functional group is directly linked to the trehalose core; and (ii) the TMM-based probes and analogues in which the detectable moiety or the bioorthogonal small functional group is attached at the end of the lipidic chain thus mimicking TMM. All the synthesized trehalose- and TMM-based probes, that have been successfully used so far for metabolic labeling experiments with diverse application, are presented in Figures 3 and 4 respectively.

2. Synthesis of trehalose-based probes

The synthesis of trehalose-based probes can be either straightforward or challenging, depending on the hydroxyl group of trehalose to be functionalized (i.e., primary or secondary alcohol). The first synthesis of a trehalose-based probe used in a cellular context was reported by Davis, Barry and co-workers with the fluorescein-containing trehalose **FITC-Tre** [15], bearing an extra methyl group at the anomeric position of one of the glucose units. One of the challenging steps of the preparation of **FITC-Tre** was the α,α-selective glycosylation. In this study

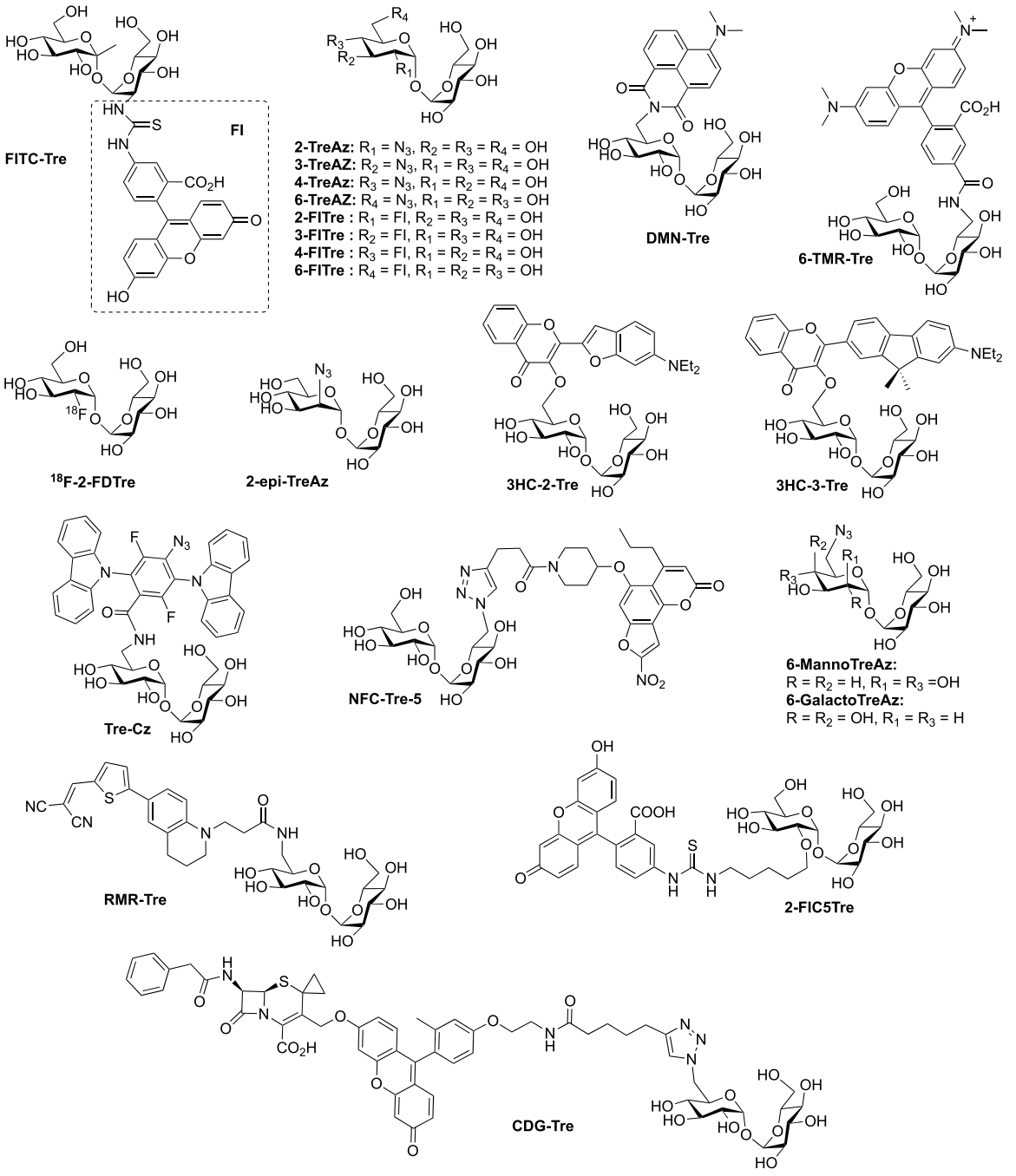


Figure 3. Structures of trehalose-based probes. (Az: azide; CDG: cephalosporinase-dependent green; Cz: carbazole; DMN: 4-(*N*,*N*-Dimethylamino)-1,8-naphthalimide; 2-FD: 2-deoxy-2-fluoro; FI: fluorescein; FITC: fluorescein-isothiocyanate; HC: hydroxychromone; NFC: nitrofuranyl calanolide; RMR: far-red molecular rotor; TMR: tetramethylrhodamine; Tre: trehalose.)

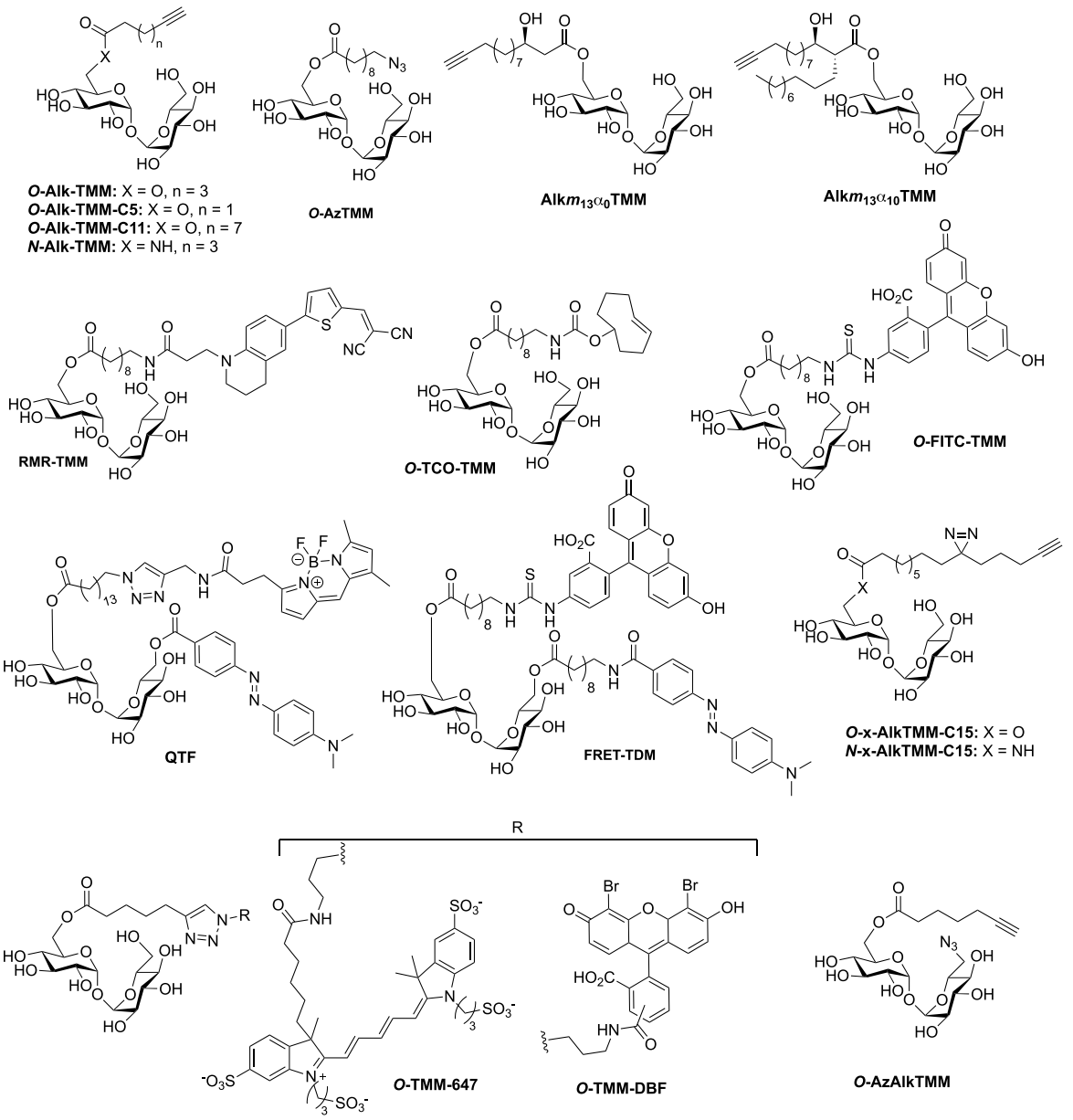
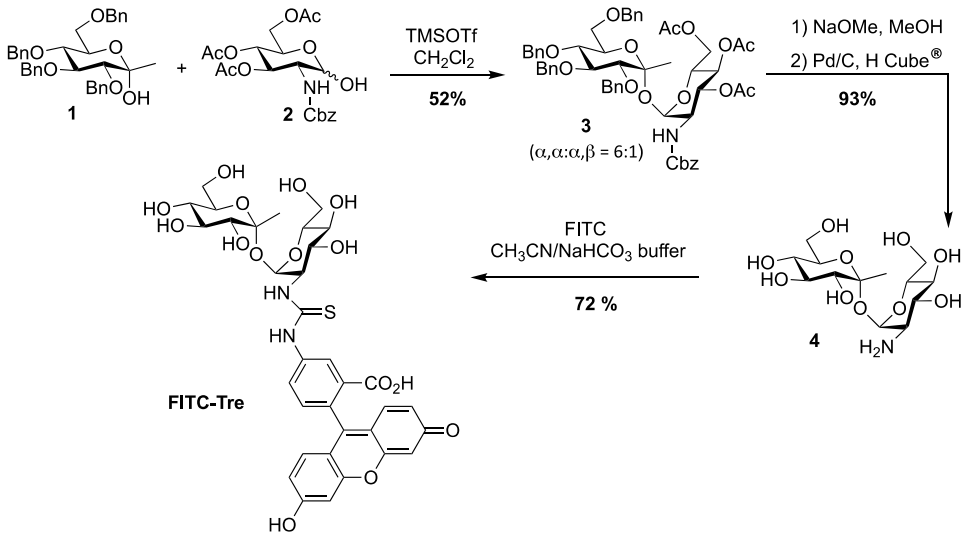


Figure 4. Structures of TMM-bases probes. (Alk: alkyne; Az: azide; DBF: dibromofluorescein; FRET: fluorescence resonance energy transfer; FITC: fluorescein-isothiocyanate; QTF: quencher-trehalose-fluorophore; RMR: far-red molecular rotor; TCO: *trans*-cyclooctene; TDM: trehalose dimycolate; TMM: trehalose monomycolate.)

the authors used the α-selective approach developed by Ikegami [16], starting from the ketoside **1** as glycosyl donor and the glucosamine derivative **2** as acceptor and governed by the extra methyl group of compound **1**. The glycosylation provided in one step the derivative **3** in 52% yield, functionalized with an equatorial amino group. Such glycosylations, in order to construct the trehalose core, are very challenging, and other studies were reported (see for example [17–19]). It should be noted that the synthesis



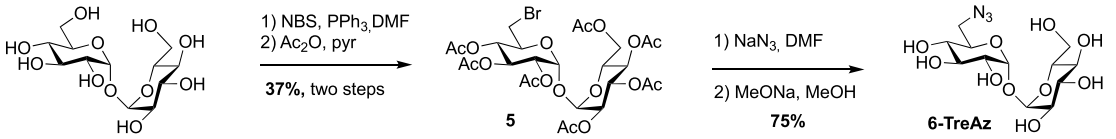
Scheme 1. Synthesis of **FITC-Tre** [15].

of amino-trehaloses starting from native trehalose is very challenging and usually takes many synthetic steps. Thus, even if the preparation of donor and acceptor **1** and **2** could be time consuming, this approach furnished in one glycosylation step the expected derivative **3**. After methanolysis of the acetate groups and hydrogenolysis of the benzyl carbamate, free derivative **4** was converted in **FITC-Tre** in high yield by treatment with fluorescein-isothiocyanate (FITC) (Scheme 1). Interestingly, this study revealed a strong Ag85 plasticity for trehalose processing and FITC-Tre was the first trehalose analogue used in metabolic labeling experiments on growing *Mtb* producing fluorescent bacteria.

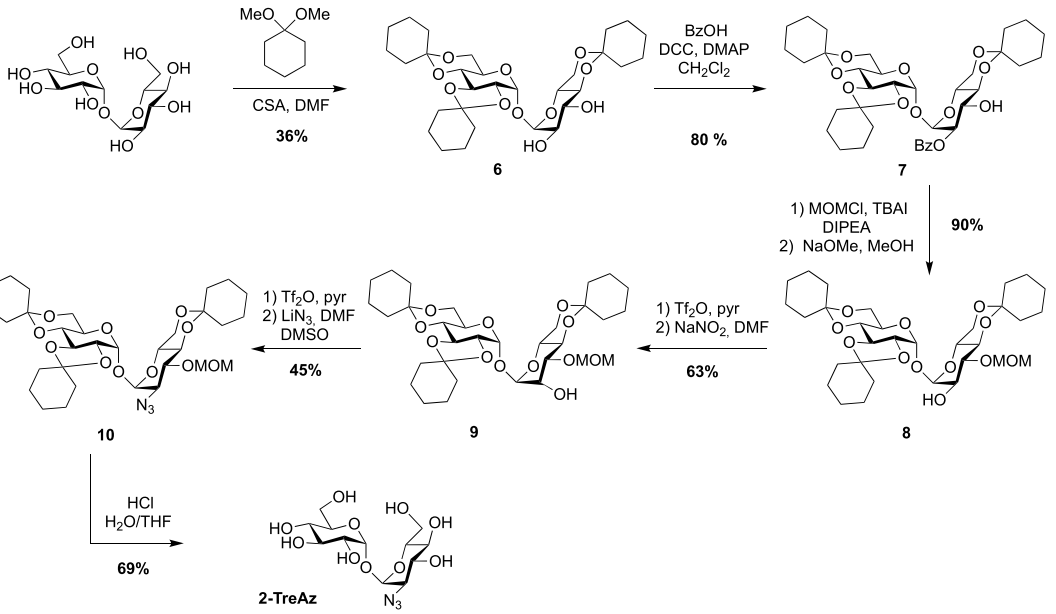
In 2012, Bertozzi published a mycomembrane labeling study using four azido-trehalose analogues referred to as **2**-, **3**-, **4**- and **6-TreAz** [20]. The latter[1] was prepared according to a synthesis reported by Hanessian [21]. The synthesis started with the mono-*O*-bromination of commercially available α-α-D-trehalose in the presence of *N*-bromosuccinimide (NBS) and triphenylphosphine. Due to the C_2 symmetry of trehalose, mono-*O*-bromination at C6 or C6′ gave a same and unique derivative. Desymmetrized derivative **5** was obtained after per-*O*-acetylation of the crude product, a step that facilitated its purification. **5** was then engaged in a nucleophilic substitution reaction in the presence of sodium azide followed by de-*O*-acetylation to obtain **6-TreAz** (Scheme 2).

The preparation of the azido-trehaloses functionalized through their secondary alcohol is more challenging because it needs robust methods enabling selective protection to differentiate the eight hydroxyl groups. Furthermore, the introduced azido group should be positioned in equatorial position. The synthesis of **2**- and **3-TreAz** started from the same desymmetrized intermediate **6** obtained according to a procedure reported by Wallace and Minnikin [22]. From this common intermediate different protection/deprotection steps were carried out to selectively isolate the 2- or 3-hydroxyl groups. In order to get **2-TreAz**, compound **6** was first protected on its C2 position *via* a selective esterification on the more reactive OH-2, leading to key compound **7** in good yield. The OH-3 was then protected using methoxymethyl chloride (MOMCl) and derivative **8** was obtained after de-*O*-benzoylation. This sequence resulted in the isolation of the 2-hydroxyl group, which was engaged on a double inversion process using the Lattrell–Dax nitrite-mediated inversion [23,24]. After activation of the hydroxyl group of **8** using triflic anhydride, compound **9** was obtained after treatment in the presence of $NaNO_2$. This sequence allowed the authors to isolate a protected trehalose with an axial hydroxyl on C2 posi-

[1] 6-TreAz was also prepared in the study of Davis, Barry and coworkers, see reference [15].



Scheme 2. Synthesis of **6-TreAz** [20,21].



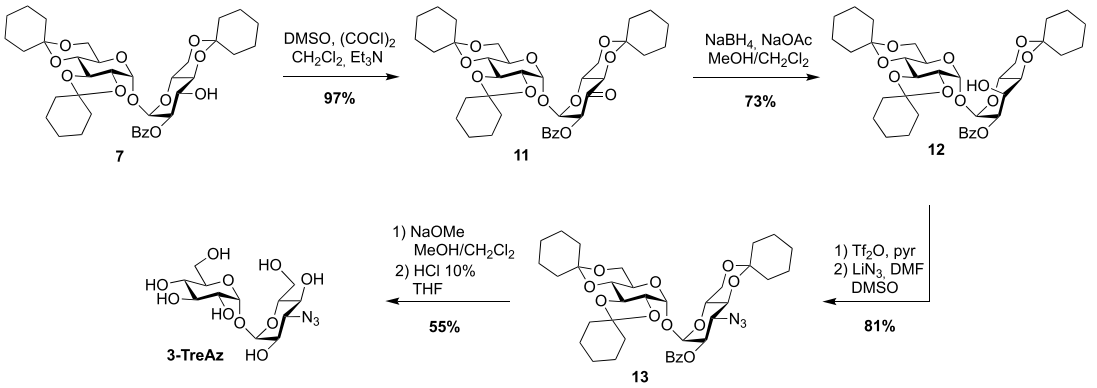
Scheme 3. Synthesis of **2-TreAz** [20].

tion. Compound **9** was then reacted with triflic anhydride followed by lithium azide, and the azido derivative **10** was obtained in 45% yield. The difficulty of this sequence could be the separation of the target compound which can be obtained as a mixture with a side product resulting from an elimination reaction. Finally, acidic hydrolysis afforded **2-TreAz** with 69% yield (Scheme 3).

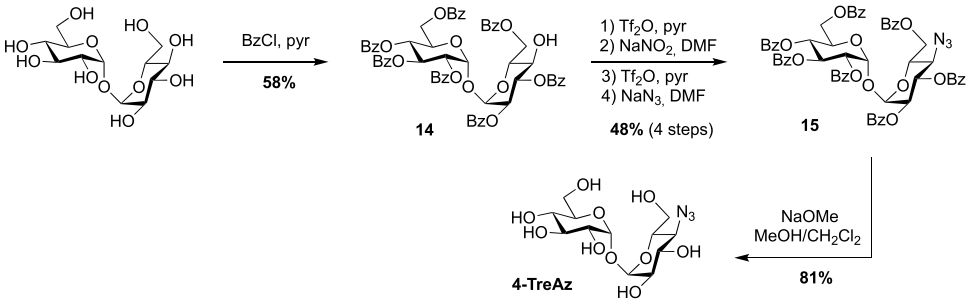
For the preparation of **3-TreAz**, Swern oxidation of key intermediate **7** led to **11** in excellent yield, and then derivative **12**, with an axial OH group at C3 was obtained through the selective reduction of the ketone function of **11**. Then, activation of the OH group of **12** using triflic anhydride followed by nucleophilic substitution of the triflate intermediate by lithium azide gave compound **13** in good yield. Finally, de-*O*-benzoylation and acidic hydrolysis led to expected **3-TreAz** (Scheme 4).

In the same study, the azido analogue, **4-TreAz** was prepared using the hepta-*O*-benzoate derivative **14** described by Nashed and coworkers in 1993 [25]. Indeed, treatment of trehalose by eight equivalents of benzoyl chloride led to the desymmetrized trehalose **14** in good yield. It should be noted that, in these conditions, two other symmetrical derivatives were also obtained, i.e., the per-*O*-benzoylated trehalose and the $2,2',3,3',6,6'$-hexa-*O*-benzoylated trehalose. The azide group was then incorporated through a double inversion using the same approach applied for **2-TreAz** with the Lattrell–Dax nitrite-mediated inversion. Finally, **4-TreAz** was obtained after de-*O*-benzoylation (Scheme 5).

All the four azido-trehalose analogues were used in metabolic labeling experiments in *M. smegmatis* strains using a two-step approach through the bioorthogonal chemical reporter strategy. The derivatives **2**- and **6-TreAz** were found to be the most efficient. Bertozzi next used these analogues as key precursors for the synthesis of new fluorescent trehaloses. Indeed, **TreAz** derivatives were subjected



Scheme 4. Synthesis of **3-TreAz** [20].



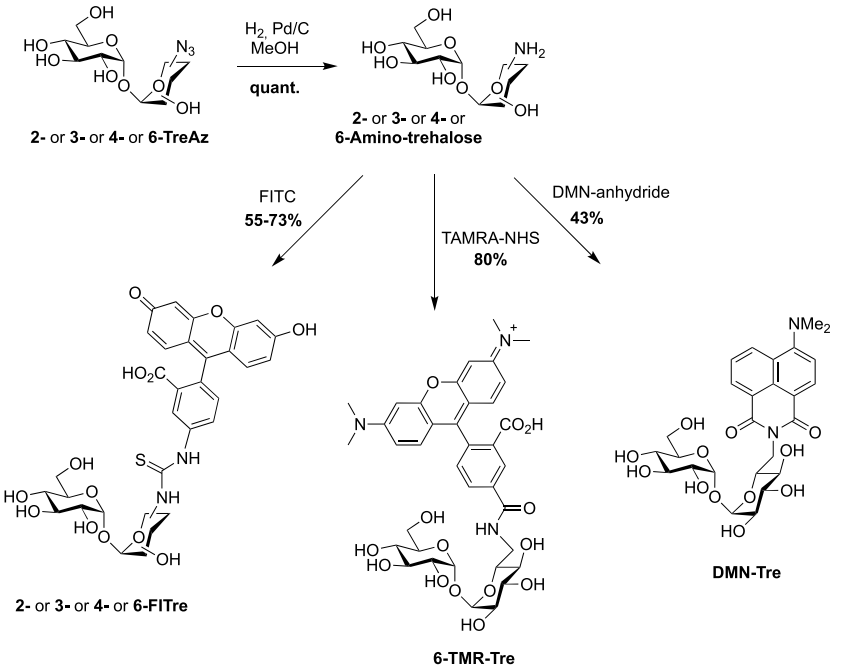
Scheme 5. Synthesis of **4-TreAz** [20].

to hydrogenolysis to give the corresponding amino-trehaloses. All the four obtained amino-trehaloses were then reacted with fluorescein isothiocyanate (FITC) to afford **2**-, **3**-, **4**- and **6-FITre** (Scheme 6) [26]. These fluorescein-modified trehalose probes were more efficient in metabolic labeling experiments on Mycobacteria compared to the first reported **FITC-Tre**. The same approach was performed to obtain the tetramethylrhodamine-trehalose **6-TMR-Tre** for super-resolution microscopy experiments [27]. Finally, in order to avoid the wash steps usually performed when using fluorescent probes, the solvatochromic dye **DMN-Tre** [28] was prepared starting from 6-amino-trehalose (Scheme 6). Indeed, solvatochromic tools show high shifts of fluorescence with solvent polarity modifications and are thus efficient for mycomembrane imaging without washing steps.

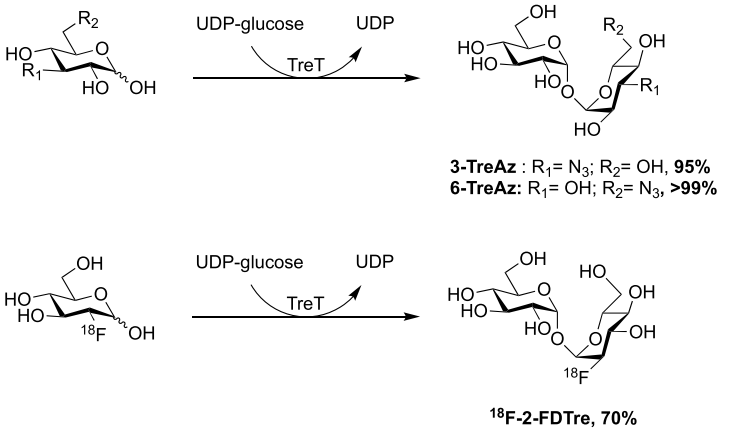
In 2014, the Swarts's group published a new route for the synthesis of azido trehalose derivatives using a chemoenzymatic approach starting from several azido-glucoses and uridine diphosphate-glucose (UDP-glucose) in presence of trehalose synthase TreT (Scheme 7) [29]. Under these conditions, access to **3**- and **6-TreAz** were improved in comparison to the firstly reported chemical syntheses. Unfortunately, this enzymatic approach failed for the 2- and 4-modified trehaloses. For the latter, the chemical synthesis remains still more efficient. Additionally, the same team also developed an access to the radioprobe **^{18}F-2-FDTre** using the same methodology [30].

In 2020, the Fulham group employed this enzymatic approach for the synthesis of 6-azido-modified derivatives with the use of 6-azido-mannose and 6-azido-galactose in place of 6-azido-glucose [31]. This resulted in the obtention of the non-symmetrical trehalose analogues **6-MannoTreAz** and **6-GalactoTreAz** (Scheme 8). These bioorthogonal trehalose analogues were used in *M. smegmatis* labeling experiments, highlighting a good substrate tolerance of mycoloyltransferases.

Still in 2020, the **6-TreAz** derivative was used, by Rao group, as a precursor for the preparation of the new fluorogenic probe **CDG-Tre** activated by the



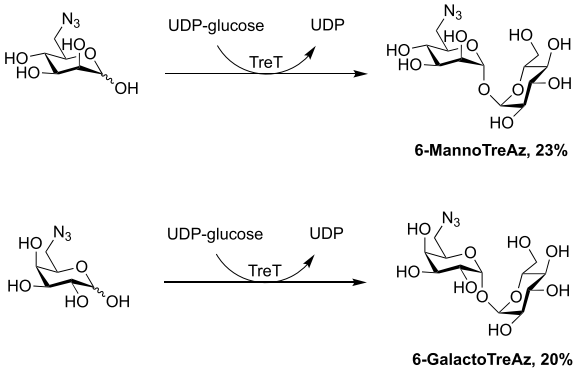
Scheme 6. Syntheses of **2**-, **3**-, **4**-, **6-FITre** [26]; **6-TMR-Tre** [27] and **DMN-Tre** [28]. (TAMRA-NHS: 5-Carboxy-tetramethylrhodamine *N*-succinimidyl ester.)



Scheme 7. Chemoenzymatic approaches for the synthesis of **3**- and **6-TreAz** [29] and **^{18}F-2-FDTre** [30].

β-lactamase BlaC expressed in *M. tuberculosis* [32]. In this approach, the fluorophore is quenched by the cephalosporin residue and the fluorescence is recovered after processing by the β-lactamase BlaC. The resulting fluorescent trehalose can then be processed by mycoloyltransferases and thus label the mycomembrane. Freshly prepared fluorescein derivative **16** and cephalosporin scaffold **17** were engaged in nucleophilic substitution followed by treatment under acidic conditions, leading to **18**. The crude was then directly coupled without purification to **6-TreAz** *via* copper-catalyzed azide-alkyne cycloaddition (CuAAC) (Scheme 9).

Later, Sun and Liu used **6-TreAz** as precursor for the synthesis of the additional fluorescent trehalose-based probe **NFC-Tre-5** [33]. This probe featured



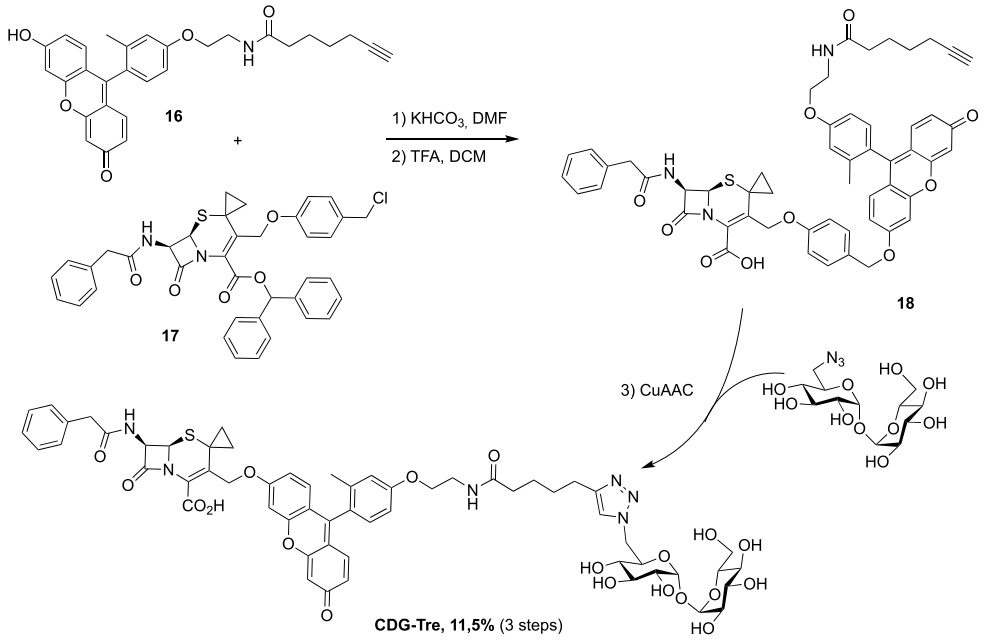
Scheme 8. Chemoenzymatic synthesis of **6-MannoTreAz** and **6-GalactoTreAz** [31].

an off-on fluorescent group activated by the nitroreductase Rv2466 allowing the specific labeling of single cells. **NFC-Tre-5** was obtained after CuAAC reaction between **6-TreAz** and an alkyne-modified nitrofuranyl calanolide scaffold (Scheme 10).

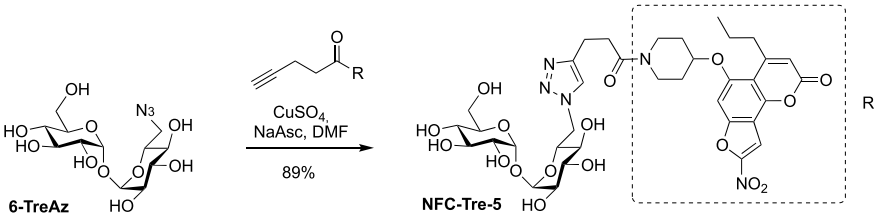
In 2022, Guianvarc'h, Vauzeilles and Bourdreux reported the synthesis of two new trehalose-based probes modified at C2 for the metabolic labeling of the mycomembrane [34] i.e., a native trehalose bearing a short ether linker functionalized with a fluorescein, **2-FIC5Tre**, and **2-epi-TreAz**, an epimer at C2 of the known **2-TreAz**. In the course of this study, they also proposed a new route for the synthesis of **2-TreAz**. The key step of these syntheses was a tandem one-pot protection and desymmetrization of trehalose mediated by $FeCl_3{\cdot}6H_2O$. In this study, the authors revisited a first published protocol [35] in order to get a protected trehalose with only a free OH group at C2 position. Briefly, α,α-D-trehalose was firstly per-*O*-silylated in pyridine and the resulting derivative **19** was engaged, in a tandem protection protocol mediated by $FeCl_3{\cdot}6H_2O$ in the presence of benzaldehyde and triethylsilane (Scheme 11). Under these conditions the desymmetrized and protected trehalose **20** was obtained in 40% yield. This key compound was used as starting material for the synthesis of the three derivatives **2-epi-TreAz**, **2-TreAz** and **2-FIC5Tre**. The tri-*O*-benzylated trehalose **20** was therefore converted into triflate **21** and then immediately engaged in a nucleophilic substitution reaction to introduce the azido moiety in presence of tetra-*n*-butylammonium azide ($TBAN_3$). This sequence allowed to rapidly obtain an axial azido moiety at C2 of trehalose. Challenging deprotections were then carried out using anhydrous iron(III) chloride in the presence of acetic anhydride followed by de-*O*-acetylation providing **2-epi-TreAz** in good yield (Scheme 12). The known **2-TreAz** was also synthesized during this study by performing the same double inversion protocol as the one used for its first synthesis by the group of Bertozzi. Triflate **21** was indeed treated with sodium nitrite and the resulting compound **23** with an axial OH group was then converted into **24** after an activation-nucleophilic substitution sequence using triflate anhydride in pyridine followed by a $TBAN_3$ treatment. Then, the same deprotection protocols were performed leading to **2-TreAz**. Finally, the authors also prepared a fluorescent probe starting from **20**. A small linker bearing an azido moiety was first introduced by alkylation of the hydroxyl group and the resulting compound **25** was subjected to hydrogenolysis. Then, the amine function of **26** was coupled to the fluorophore using a treatment with FITC leading to **2-FIC5Tre**. The three probes were successfully used in mycomembrane labeling experiments on *C. glutamicum*. **2-TreAz** and **2-epi-TreAz** appeared to have similar efficacy again indicating a good substrate tolerance of mycoloyltransferases.

Recently, Bertozzi published the synthesis of new solvatochromic trehalose-based probes featuring 3-hydroxychromone (3HC) dyes, **3HC-3-Tre** and **3HC-2-Tre**, the first displaying a 10-fold increase in fluorescence intensity compared to the previous report on **DMN-Tre** [36]. These fluorescent probes are known to have high fluorescence quantum yields and they proved to efficiently label *M. tuberculosis* cells within 10 min of probe treatment. For these syntheses, the authors took advantage of the known desymmetrized and monobrominated derivative **5** previously described by Hanessian. Compound **5** was engaged in nucleophilic substitution with the two 3-hydroxychromones **27** and **28**, providing mono-*O*-coupled derivatives **29** and **30**. **3HC-3-Tre** and **3HC-2-Tre** were finally obtained after de-*O*-acetylation with sodium methoxide (Scheme 13).

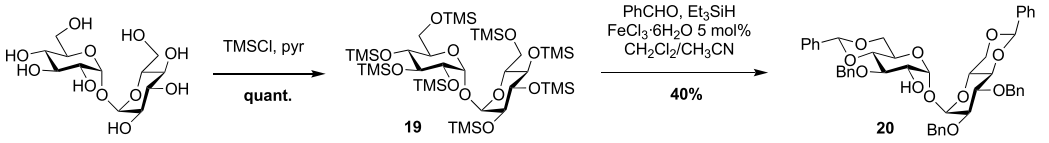
Very recently, Swarts's lab added a new fluorogenic trehalose analogue to the trehalose-based probes collection [37]. This probe, **RMR-Tre**, featured a molecular rotor allowing turn-on far-red fluorescence sensor upon interaction with constraint



Scheme 9. Synthesis of **CDG-Tre** probe targeting two mycobacterial enzymes [32].



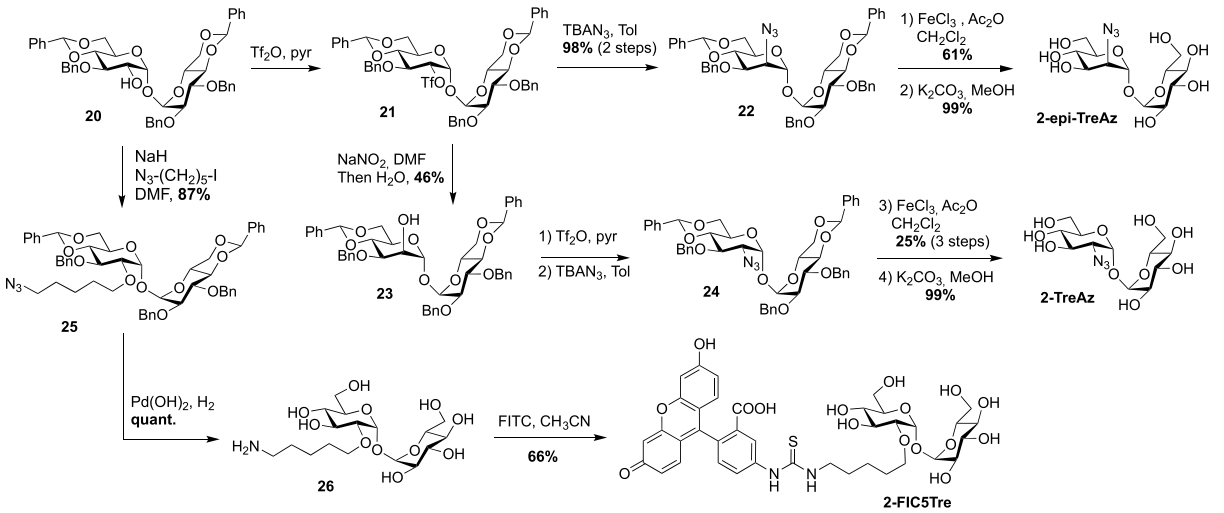
Scheme 10. Synthesis of **NFC-Tre-5** [33].



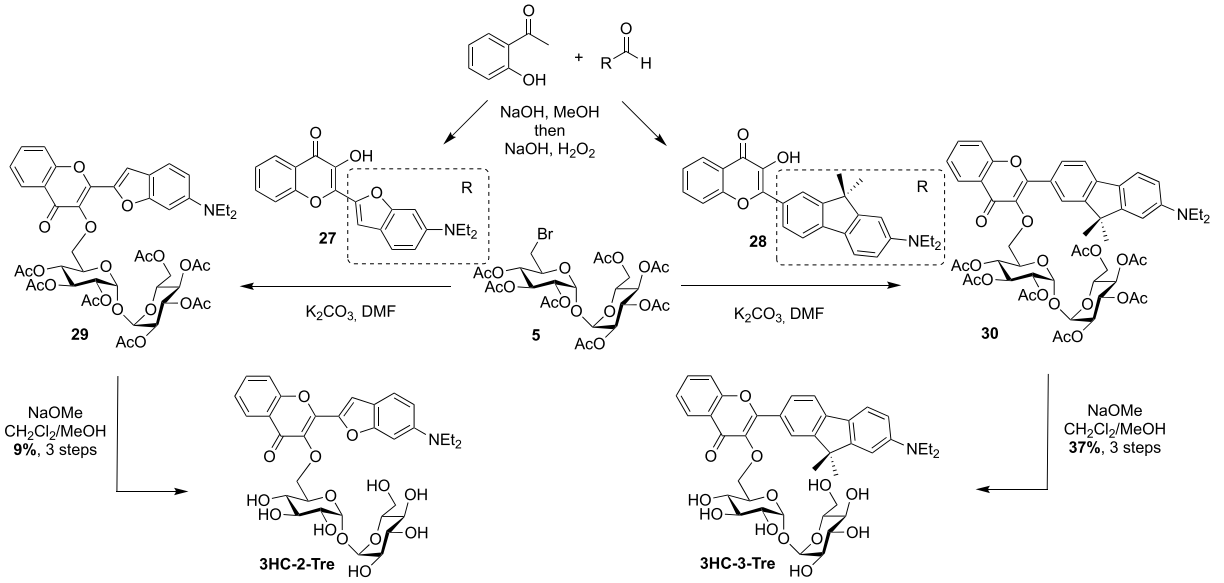
Scheme 11. $FeCl_3{\cdot}6H_2O$ mediated one-pot desymmetrization of trehalose [34,35].

environment such as the mycomembrane. Derivative **31**, obtained after Knoevenagel condensation between the known aldehyde **32** and malonitrile, was coupled, using N,N,N',N'-tetramethyl-*O*-(*N*-succinimidyl)uronium tetrafluoroborate (TSTU), to 6-amino trehalose **33** providing **RMR-Tre** in good yield (Scheme 14).

Finally, during the preparation of this review, Yan and coworkers reported the synthesis of a fluorescence turn-on probe **Tre-Cz** and its use for imaging mycobacteria after photoactivation of the aryl azide allowing the formation of the fluorescent product [38]. It was prepared starting from the known 6-amino-trehalose **33** and a *N*-hydroxysuccinimide(NHS)-functionalized carbazole derivative (Scheme 15).



Scheme 12. Syntheses of **2-epi-TreAz**; **2-TreAz** and **2-FIC5Tre** [34].

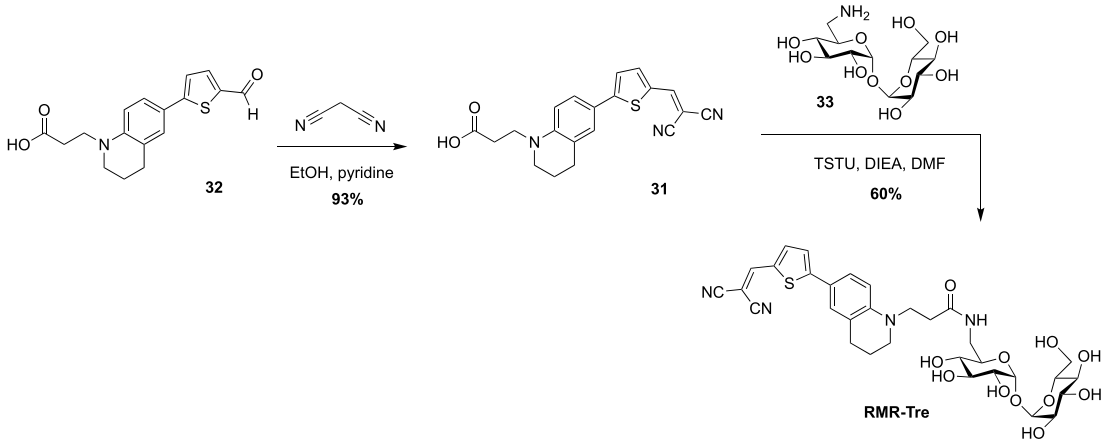


Scheme 13. Synthesis of **3HC-2-Tre** and **3HC-3-Tre** [36].

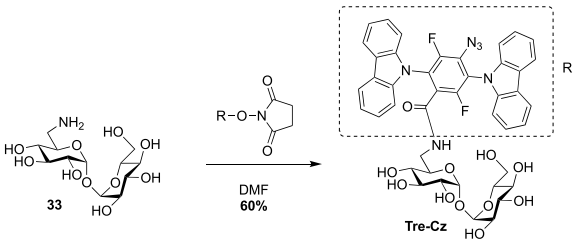
3. Synthesis of TMM-based probes

In parallel to trehalose-based probes, other useful chemical tools, i.e., acylated trehalose-based probes mimicking TMM, were synthesized to label the mycomembrane. In contrast to trehalose-based probes that are processed through the recycling trehalose pathway, TMM-based probes target mycoloyltransferases allowing the labeling of both the outer leaflet and the inner part of the mycomembrane (TDM and AGM layers, respectively).

From a synthetic point of view, different challenges have to be addressed for the preparation of TMM-based probes. The first one is the trehalose selective esterification of the lipidic part mimicking the mycolate chain, which relies on the difference of reactivity between primary and secondary alcohols. The most important challenge is the preparation of the mycolic acid pattern which implies (i) the control of the absolute configuration of its two stereogenic



Scheme 14. Synthesis of **RMR-Tre** [37].

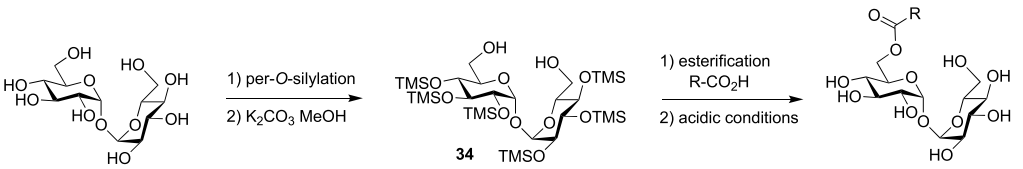


Scheme 15. Synthesis of **Tre-Cz** [38].

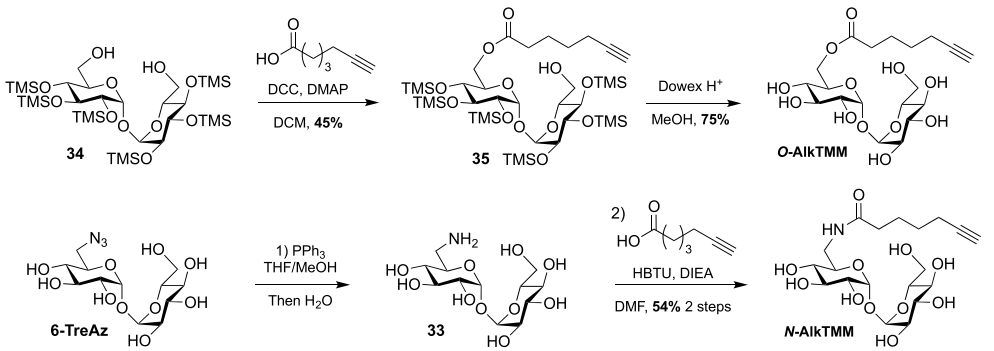
centers and (ii) the introduction of a bioorthogonal or a detectable tag on one of the long fatty acid chains. Some enantioselective syntheses of mycolic acids have already been reported in the literature (for a few examples see [39–41]), but the main difficulty for the preparation of such TMM-based probes is to make them compatible with a bioorthogonal or detectable moiety. Several teams have been interested in the synthesis of TMM-based chemical reporters and most of them reported simplified structures by replacing the complex native mycolic pattern by a simple acyl chain including the tag at the end. These probes were prepared using the same approach for the esterification of the lipidic chain of interest with trehalose. It relies on the per-*O*-silylation of trehalose followed by a controlled deprotection of the two primary positions leading to the symmetrical diol **34** according to a procedure reported by Toubiana [42]. Then, the deprotected derivative **34** is monoesterified in the presence of a coupling agent and the resulting ester is fully deprotected under acidic conditions (Scheme 16) [43].

In 2016, Swarts took advantage of this approach and published the first bioorthogonal analogue related to TMM using commercially available 6-heptynoic acid as simplified mycolic acid mimic [44]. This analogue was obtained very efficiently by reaction between compound **34** and 6-heptynoic acid in the presence of dicyclohexylcarbodiimide (DCC) and 4-dimethylaminopyridine (DMAP). The esterified intermediate **35** was then fully deprotected under acidic conditions leading to ***O*-AlkTMM** (Scheme 17). This simplified TMM analogue was successfully used for metabolic labeling experiments in *M. smegmatis*. During this study, 6-heptynoic acid was also used for the preparation of an amide-linked analogue. The known **6-TreAz** was first converted to compound **33** using Staudinger reduction and the resulting amine was coupled to 6-heptynoic acid in the presence of *O*-(benzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate (HBTU) affording ***N*-AlkTMM**. The main difference between these two probes is the fact that after being processed by mycoloyltransferases, ***O*-AlkTMM** can label both the inner and the outer leaflets of the mycomembrane whereas ***N*-AlkTMM** can only label the outer part (i.e., the TDM layer).

A few years later, the same team extended the range of TMM-based probes by publishing a collection of new analogues differing by the chain length, the bioorthogonal group or by the introduction of a fluorophore [45]. On the basis of their previous work, the authors prepared the two other



Scheme 16. Approach for selective synthesis of 6-*O*-acylated trehaloses.

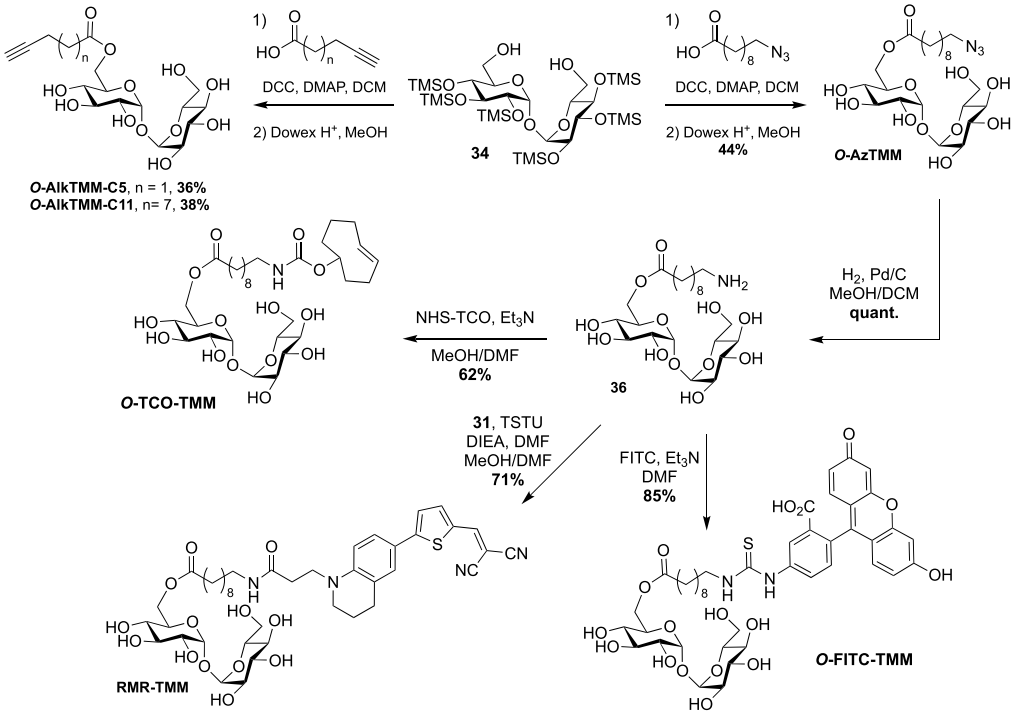


Scheme 17. Synthesis of ***N***- and ***O*-AlkTMM** [44].

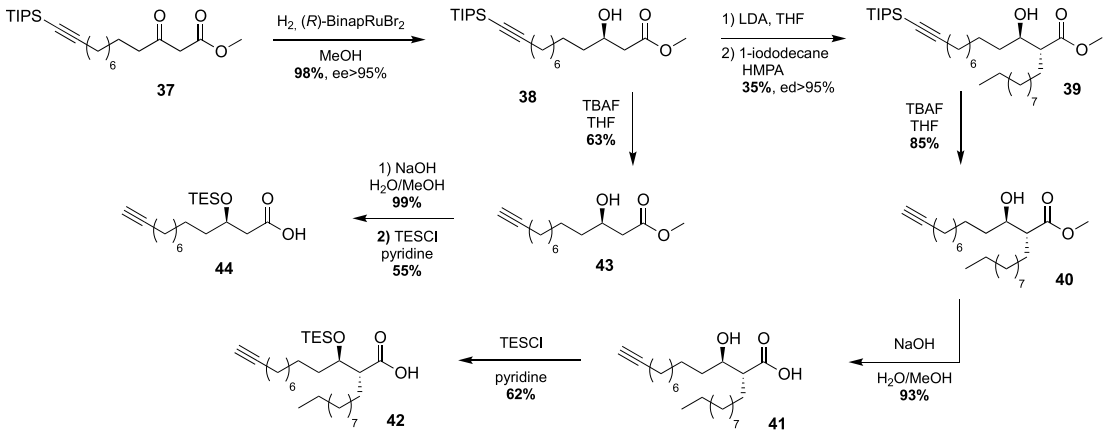
alkyne-TMM ***O*-AlkTMM-C5** and ***O*-AlkTMM-C11** with 5 and 11 carbon atoms on the lipidic moiety, respectively. Compound **34** was also esterified with 10-azido-decanoic acid providing ***O*-AzTMM**, an acylated trehalose with an azide as bioorthogonal group. This compound was finally converted to the corresponding amine allowing the coupling to a trans-cyclooctene *N*-hydroxysuccinimide ester (NHS-TCO) and to FITC, leading to ***O*-TCO-TMM** for tetrazine ligation and ***O*-FITC-TMM** respectively (Scheme 18). Very recently, the same lab expanded the far-red molecular rotor probe family by coupling compound **31** to intermediate **36** furnishing **RMR-TMM** (Scheme 18) [37].

These studies have provided an impressive molecular diversity with different bioorthogonal or fluorescent moieties attached to the end of the ester-linked mono-acyl chain, but none of these included the native pattern of mycolic acids which represents an important synthetic challenge. In this way, Guianvarc'h and Bourdreux reported in 2019 the first access to trehalose monomycolate bioorthogonal probes closely mimicking the complex structure of native TMM [46]. Both **Alk$m_{13}\alpha_0$TMM** and **Alk$m_{13}\alpha_{10}$TMM** analogues featured an OH group in β position with an (*R*) stereochemistry (see Scheme 20). Moreover, **Alk$m_{13}\alpha_{10}$TMM** also had the branched α-lipidic chain of mycolic acids, with an *anti*-relationship between the two α and β substituents. The β-ketoester **37**, obtained after a two-carbon Masamune homologation of the corresponding alkyne-protected carboxylic acid, was engaged in an enantioselective reduction in the presence of Noyori's catalyst (*R*)-BinapRuBr$_2$. Under these conditions, the triisopropylsilyl protecting group prevented the undesired concomitant reduction of the alkyne. Thus, the expected β-hydroxyester **38**, isolated in excellent yield and enantioselectivity, was then alkylated in the presence of 1-iododecane and hexamethylphosphoramide (HMPA) furnishing diastereoselectively **39** in a moderate yield. After *n*-tetrabutylammonium fluoride (TBAF) deprotection of the alkyne, the ester was saponified and the β-OH group was protected as triethylsilylether in anticipation of the future esterification with trehalose, providing carboxylic acid **42** (Scheme 19). During this study, the authors also prepared a bioorthogonal β-hydroxylated fatty acid lacking the α-side chain of mycolic acids. β-hydroxyester **38** was firstly desilylated, saponified and the β-OH group protected, leading to carboxylic acid **44**.

Finally, the two probes **Alk$m_{13}\alpha_{10}$TMM** and **Alk$m_{13}\alpha_0$TMM** were obtained using the same approach as previously described for the syntheses of the previous TMM analogues. Fatty acids **42** and



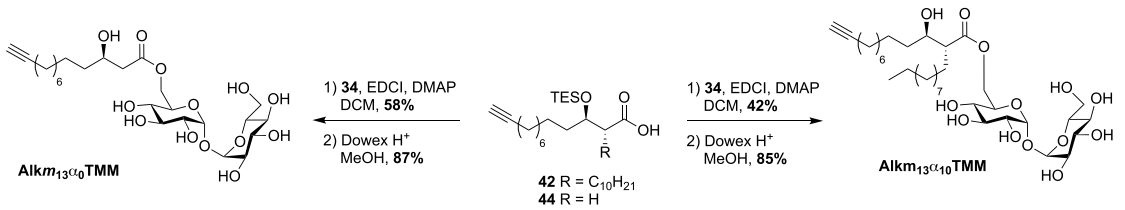
Scheme 18. Syntheses of ***O*-AzTMM**, ***O*-TCO-TMM**, ***O*-FITC-TMM** [45] and **RMR-TMM** [37].



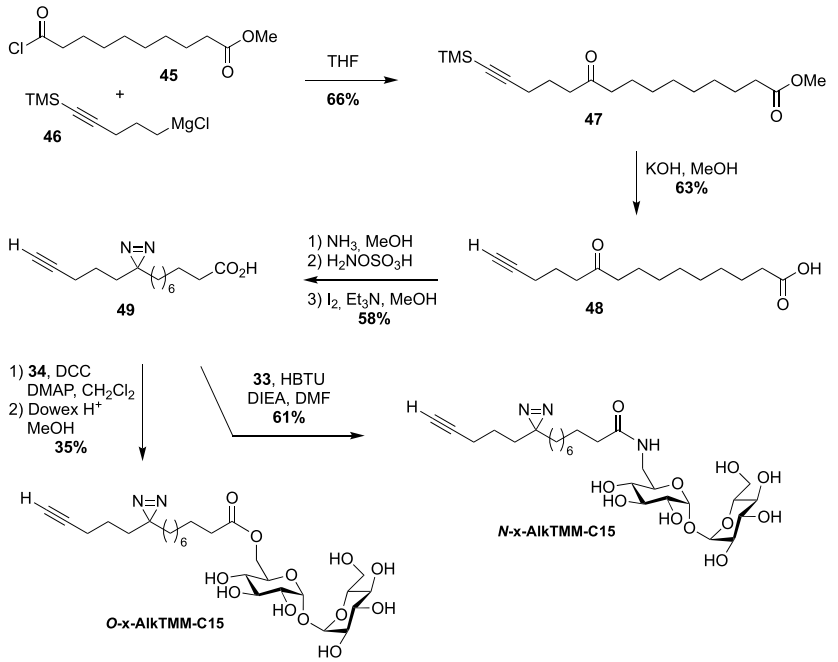
Scheme 19. Synthesis of the first mycolic acid with an alkyne bioorthogonal moiety [46].

44 were esterified to trehalose **34** in the presence of 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDCI) and the resulting intermediates were fully deprotected under acidic conditions, providing the native TMM-based probes (Scheme 20). These two tools proved to be highly efficient for the labeling of the *C. glutamicum* mycomembrane even at low concentration. They are supposed to avoid the substantial acylhydrolase activity observed with unbranched lipid and to improve the native transfer activity of mycolate to its acceptors.

In 2020, Swarts proposed the TMM-based probes, ***O*-x-AlkTMM-C15** and ***N*-x-AlkTMM-C15**, with an



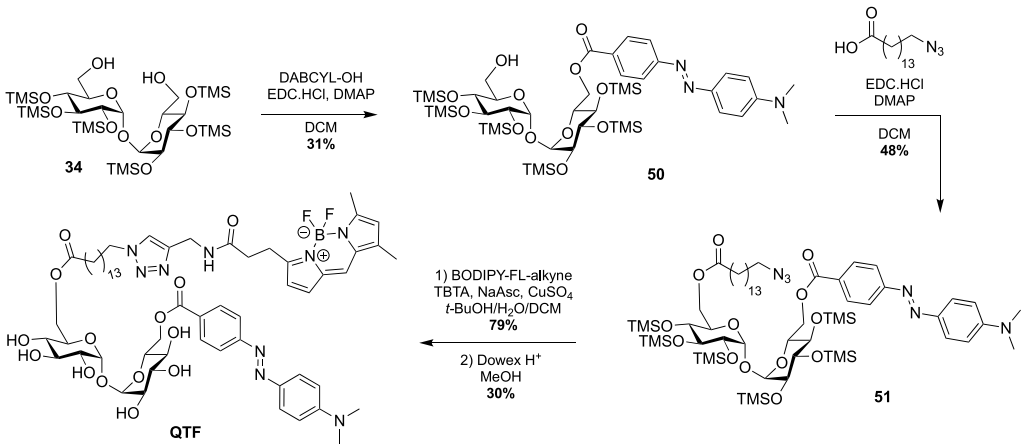
Scheme 20. Synthesis of **Alk*m*$_{13}\alpha_0$TMM** and **Alk*m*$_{13}\alpha_{10}$TMM** [46].



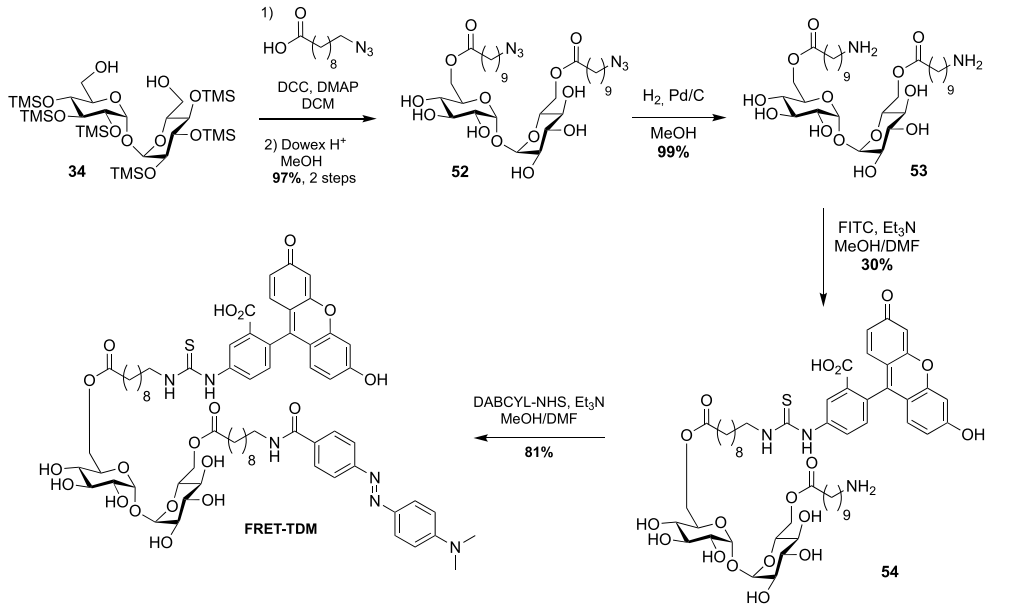
Scheme 21. Synthesis of ***O*-x-AlkTMM-C15** and ***N*-x-AlkTMM-C15** [47].

extra photoactivatable diazirine into the lipidic part in order to identify mycolate-protein interactions [47]. The synthesis of these tools relied on the preparation of the known bifunctional fatty acid **49** [48]. The latter was prepared by nucleophilic addition of the Grignard **46** to acyl chloride **45** and subsequent deprotections under alkaline conditions. The diazirine was then installed as previously reported [49] and coupled to trehalose derivatives **34** and **33** providing ***O*-x-AlkTMM-C15** and ***N*-x-AlkTMM-C15**, respectively (Scheme 21).

A few years ago, Kiessling reported an ingenious strategy based on an observed acylhydrolase activity of mycoloyltransferases in the presence of TMM analogues with unbranched lipids [50]. She designed the fluorogenic **QTF** that is structurally closer to TDM rather than TMM analogues [51]. Indeed, one primary alcohol of trehalose is esterified to an acyl chain bearing a fluorescent BODIPY-based-probe, whereas the second primary alcohol is linked to a fluorescence quencher. When **QTF** is processed by mycoloyltransferases, the unbranched lipid bearing the fluorophore moiety can promote hydrolysis of the acyl-enzyme intermediate thus restoring the fluorescence. **QTF** was prepared by two successive selective esterifications with 4-dimethylaminoazobenzene-4-carboxylic acid (DABCYL-OH) and 15-azidopentadecanoic acid, respectively. Di-*O*-esterified compound **51** was next



Scheme 22. Synthesis of **QTF** [51].



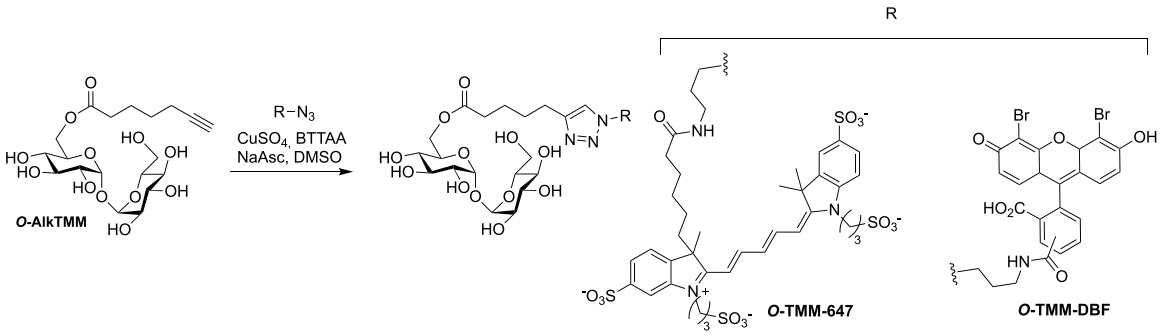
Scheme 23. Synthesis of **FRET-TDM** [52].

coupled to a fluorophore through CuAAC reaction leading, after deprotection, to **QTF** (Scheme 22).

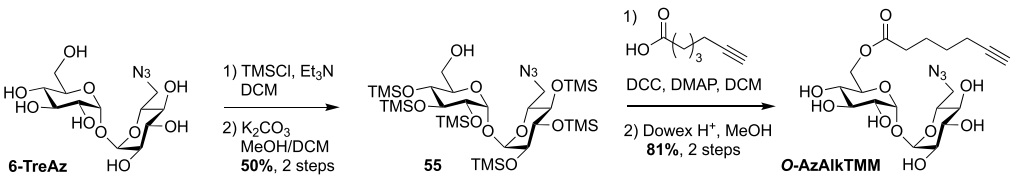
This strategy was next used by Swarts in a study targeting remodelling enzymes in mycobacteria, such as TDM hydrolase (Tdmh) [52]. The fluorogenic probe was obtained starting from **34** which was first fully esterified with 10-azidodecanoic acid in the presence of DCC and DMAP. The esterified derivative **52** was next converted to the diamine derivative **53** by reduction of the two azido groups. A fluorescein moiety was then introduced by treatment with FITC and the quencher of fluorescence was installed to provide the **FRET-TDM** probe (Scheme 23).

Following its description (see Scheme 17), ***O*-AlkTMM** was used as starting material for the preparation of the two additional fluorescent probes ***O*-TMM-647** and ***O*-TMM-DBF**. Indeed, Theriot reported in 2019 the use of these two probes, synthesized by CuAAC reaction between ***O*-AlkTMM** and two different azide fluorophores, in a study on *C. glutamicum* (Scheme 24) [53].

Finally, **6-TreAz** and ***O*-AlkTMM** were merged giving rise to the bifunctional reporter ***O*-AzAlkTMM** [54]. It was synthesized using the ap-



Scheme 24. Synthesis of ***O*-TMM-647**, ***O*-TMM-DBF** [53].



Scheme 25. Synthesis of ***O*-AzAlkTMM** [54].

proach used for most of the described TMM analogues as indicated in Scheme 16, starting from **6-TreAz**. Indeed **6-TreAz** was first per-*O*-silylated and then treated under mild basic conditions, leading to **55**. The free alcohol of **55** was then esterified with 6-heptynoic acid and the resulting compound was fully deprotected to afford the expected bifunctional chemical reporter ***O*-AzAlkTMM** (Scheme 25) with an azide moiety on the trehalose core, and an alkyne on the lipidic chain mimicking the mycolic acid. These two bioorthogonal groups can react successively with two different fluorophores. Furthermore, the azide group can label only the TDM layer of the mycomembrane, whereas the alkyne moiety can label both the TDM and AGM layers. This bifunctional probe was used to study different metabolic processes by targeting mycoloyltransferase activities as well as trehalose recycling process.

4. Conclusion

In recent years, many smart probes have been designed and successfully used in metabolic labeling experiments to explore the mycomembrane, which is specific to mycobacteria and crucial for their survival. They have provided new insights into the mycomembrane biogenesis as well as potential new diagnostic tools for the detection of *M. tuberculosis* [7,8]. A summary of the main features of these probes and their use is given in Tables 1 and 2 (see below). These trehalose-based probes, based on the selective trehalose functionalization, require important synthetic developments, especially in glycochemistry, in order to obtain more and more efficient tools, and we contributed to this exciting adventure. This review focused on the chemical challenges for the synthesis of the trehalose- and TMM-based probes described so far. Most of the reported trehalose-based analogues are modified at C6. Indeed, selective reaction on primary alcohols of trehalose is fast and convenient, and the C_2-symmetry of trehalose makes the two primary positions equivalent. Modifications at C2, C3 or C4 are more challenging and require more chemical steps to retain an equatorial position on glucose unit. However, several studies have shown a great tolerance of mycoloyltransferases. Indeed, the synthesis of trehalose-based probes bearing the tag on axial position of the carbohydrate ring could make faster the access to C2, C3 or C4 modified trehalose-based probes, by suppressing

Table 1. Recapitulative table on the principal features and uses of trehalose-based probes described so far

Classification	Probe name and reference[a]	Principal features
Fluorescein and tetramethylrhodamine *conjugated trehalose*	**FITC-Tre** [15]	• First trehalose-based fluorescent probe with a C1 methyl group at the anomeric position • Metabolic labeling experiments on growing *Mtb and* within infected macrophages • Revealed a strong Ag85 plasticity
	2-,3-,4-,6-FITre [26]	• First natural trehalose-based fluorescent probe • More efficient in labeling than the FITC-Tre (*Msmeg, Cglut*) • Applied to elucidate the mobility of labeled glycolipids in mycobacteria in several conditions
	2-FIC5Tre [34]	• Fluorophore attached to trehalose through an ether-linker • Enable efficient one step labeling of mycobacteria (*Cglut*)
	6-TMR-Tre [27]	• Mycomembrane visualization with super-resolution microscopy • Study of the mycomembrane and peptidoglycan dynamics using a dual metabolic labeling strategy (*Msmeg, Mm, Cglut*)
Azido-trehalose	**2-,3-,4-,6-TreAz** [20,29,34]	• First bioorthogonal trehalose for chemical reporter strategy • Study of trehalose metabolism in live mycobacteria (*Msmeg, Mtb*, BCG) • 2- and 6-TreAz more efficient
	6-MannoTreAz [31] **6-GalactoTreAz** [31]	• Good substrate tolerance toward mycoloyltransferases (*Msmeg*, BCG)
	2-epiTreAz [34]	• Good substrate tolerance toward mycoloyltransferases • Similar efficiency than 2-TreAz in labeling experiments (*Cglut*)
Trehalose conjugated to solvatochromic dye	**DMN-Tre** [28]	• Sensitive to changes in the polarity of the medium • 700-fold increase in fluorescence intensity in hydrophobic environment • Does not require sample washing • Strong potential to detect *Mtb* in patients' sputum (*Mtb, Msmeg, Mm*)
	3HC-3-Tre [36] **3HC-2-Tre** [36]	• 10-fold increase in fluorescence intensity compared to DMN-Tre • Efficiently label *Mtb* within 10 min of probe treatment (*Msmeg, Mtb, Cglut*)
	RMR-Tre [37]	• Rotor turn-on fluorophore, no-wash and fast detection of live cells • Bright far-red emission and low-background fluorescence detection • 100-fold enhancement in *Mtb* labeling compared to previous fluorogenic trehalose probes (*Msmeg, Mtb*)
Trehalose conjugated to off-on fluorescent dye activated by enzyme	**CDG-Tre** [32]	• Release of the caged fluorophore conjugated to trehalose by the β-lactamase BlaC (*Msmeg*, BCG) • low concentration for the labeling of phagocytosed live BCG within macrophages
	NFC-Tre-5 [33]	• Activated by the nitroreductase Rv2466 • Specific labeling of single cells (*Mmseg, Mtb*, BCG) • Strong potential for exploring the relationship between the host and pathogen
Trehalose conjugated to off-on fluorescent dye activated by light	**Tre-Cz** [38]	• Photoactivation and detection using a handheld UV lamp (*Msmeg, Mtb*) • Trehalose uptake pathway hijacked: intracellular incorporation through the LpqY-SugABC transporter
Trehalose functionalized with radioprobe	**^{18}F-2-FDTre** [30]	• Positron emission tomography (PET) imaging to allow *in vivo* imaging of trehalose metabolism (*Msmeg*)

[a] Structures of the probes are shown in Figure 3.

Table 2. Recapitulative table on the principal features and uses of TMM-based probes described so far

Classification	Probe name and reference[a]	Principal features
Bioorthogonal simplified analogs of TMM	***O*-AlkTMM** [44] ***N*-AlkTMM** [44]	• First bioorthogonal TMM analog with unbranched acyl chain, used in metabolic labeling experiments (*Msmeg, Cglut*) • *O*-AlkTMM processed by mycoloyltransferases and enabling selective *in situ* detection of the mycomembrane components in living mycobacteria • *N*-AlkTMM suitable to analyze the outer leaflet of the mycomembrane
	***O*-AlkTMM-C5** [45] ***O*-AlkTMM-C11** [45] ***O*-Az-TMM** [45] ***O*-TCO-TMM** [45]	• Expansion of the set of TMM reporter analogs (*Cglut, Msmeg*) • Rapid cell labeling (*Msmeg*) through tetrazine ligation
TMM mycolic acid-based bioorthogonal reporter	**Alk$m_{13}\alpha_0$TMM** [46] **Alk$m_{13}\alpha_{10}$TMM** [46]	• First bioorthogonal TMM featuring the complex mycolic acid moiety • Labeling of mycomembrane with a shorter labeling time and very low doses (*Cglut*)
Simplified analogs of TMM functionalized with a fluorescent dye	***O*-FITC-TMM** [45]	• First fluorescent TMM analog • Enable one step labeling of live cells using TMM analog (*Cglut, Msmeg*)
	RMR-TMM [37]	• Far-red molecular rotor probe • Enable no-wash and fast detection of live cells (*Msmeg, Mtb*)
	***O*-TMM-647** [53] ***O*-TMM-DBF** [53]	• Fluorescent TMM analog used to follow the assembly dynamics of the mycomembrane in live cells (*Cglut, Msmeg*)
Photoactivatable and bioorthogonal simplified analog of TMM	***O*-x-AlkTMMC15** [47] ***N*-x-AlkTMMC15** [47]	• First photoactivatable TMM analog enabling *in vivo* photo-cross-linking and click-chemistry-mediated analysis of mycolate-interacting proteins (*Msmeg, Cglut*)
Quencher-trehalose-fluorophore probe	**QTF** [51]	• First FRET trehalose mycolate-based probe (*Mtb, Msmeg, Cglut*) • Unbranched lipid bearing the fluorophore moiety to promote hydrolysis of the acyl-enzyme intermediate enabling real-time imaging in native cellular environment
	FRET-TDM [52]	• Unbranched TDM-based probe used to study remodeling enzymes such as Tdmh (TDM hydrolase) for *in vitro* and *in vivo* assays (*Msmeg, Mtb*) • Probe also activated by hydrolases from other bacteria species
Bifunctional TMM chemical reporter	***O*-AzAlkTMM** [54]	• First bifunctional TMM-based probe • Simultaneously marks mycomembrane biosynthesis by targeting mycoloyltransferases and subsequent trehalose recycling (*Msmeg, Mtb*)

[a]Structures of the probes are shown in Figure 4.

double inversion protocols required to preserve the absolute configuration at C2, C3 or C4. The second major chemical reporters are mimics of TMM. In this case the same approach is used for most of the probes: a selective esterification on the same protected trehalose followed by acidic deprotection. The main difficulty for the preparation of such TMM-based probes is to synthesize the bioorthogonal functionalized mycolic acid part of this glycolipid. Most of the reported tools are simplified fatty acyl chain without ramification and β-OH group, thus making their syntheses fast and efficient. In a recent work, our group reported the first synthesis of a bioorthogonal TMM-based probes with the natural pattern of mycolic acid which was an important synthetic challenge. Interestingly, it proved to be very efficient in labeling *C. glutamicum* and such TMM-based probes may help to elucidate the mycomembrane biogenesis at the molecular level. Notably, they should help to decipher the enzymatic

mechanism and specificity of mycoloyltransferases. Altogether, the molecular tools described here, in cooperation with biochemical approaches, could help to better understand the role of all mycoloyltransferases involved in the biogenesis of the mycomembrane, and may have a significant impact in the characterization of new targets to develop innovative therapeutic approaches.

Conflicts of interest

Authors have no conflict of interest to declare.

Acknowledgments

The authors thank the Ministère de l'enseignement supérieur et de la recherche for grants to EL and PR and for financial support. The authors also thank the Agence Nationale de la Recherche (PTMyco, grant NO. ANR-22-CE44-0005-03).

References

[1] World Health Organization (WHO), "Global tuberculosis report, 2021", 2021, https://www.who.int/publications/i/item/9789240037021.

[2] J. Becker, C. Wittmann, *Industrial Biotechnology*, John Wiley & Sons Ltd, 2017, 183-220 pages.

[3] N. Dautin, C. de Sousa-d'Auria, F. Constantinesco-Becker, C. Labarre, J. Oberto, I. Li de la Sierra-Gallay, C. Dietrich, H. Issa, C. Houssin, N. Bayan, *Biochim. Biophys. Acta Gen. Subj.*, 2017, **1861**, 3581-3592.

[4] C. Carel, J. Marcoux, V. Réat, J. Parra, G. Latgé, F. Laval, P. Demange, O. Burlet-Schiltz, A. Milon, M. Daffé, M. G. Tropis, M. A. M. Renault, *Proc. Natl. Acad. Sci. USA*, 2017, **114**, 4231-4236.

[5] E. Huc, X. Meniche, R. Benz, N. Bayan, A. Ghazi, M. Tropis, M. Daffé, *J. Biol. Chem.*, 2010, **285**, 21908-21912.

[6] H. Issa, E. Huc-Claustre, T. Reddad, N. Bonadé Bottino, M. Tropis, C. Houssin, M. Daffé, N. Bayan, N. Dautin, *PLoS ONE*, 2017, **12**, article no. e0171955.

[7] N. Banahene, H. W. Kavunja, B. M. Swarts, *Chem. Rev.*, 2022, **122**, 3336-3413.

[8] D. Guianvarc'h, Y. Bourdreux, C. Biot, B. Vauzeilles, in *Comprehensive Glycoscience* (J. J. Barchi, ed.), Elsevier, Oxford, 2nd ed., 2021, 303-328.

[9] A. D. Elbein, in *Advances in Carbohydrate Chemistry and Biochemistry* (R. S. Tipson, D. Horton, eds.), vol. 30, Academic Press, New York, 1974, 227-256.

[10] A. B. Richardsa, S. Krakowkab, L. B. Dexterc, H. Schmidd, A. P. M. Wolterbeeke, D. H. Waalkens-Berendsene, A. Shigoyukif, M. Kurimoto, *Food Chem. Toxicol.*, 2002, **40**, 871-898.

[11] A. A. Khan, B. L. Stocker, M. S. M. Timmer, *Carbohydr. Res.*, 2012, **356**, 25-36.

[12] V. A. Sarpe, S. S. Kulkarni, *Trends Carbohydr. Res.*, 2013, **5**, 8-33, https://www.trendscarbo.com/getf_shoppingcart.php?id=821577444.

[13] C.-H. Wu, C.-C. Wang, *Org. Biomol. Chem.*, 2014, **12**, 5558-5562.

[14] S. Jana, S. S. Kulkarni, *Org. Biomol. Chem.*, 2020, **18**, 2013-2037.

[15] K. M. Backus, H. l. Boshoff, C. S. Barry, O. Boutureira, M. K. Patel, F. D'Hooge, S. S. Lee, L. E. Via, K. Tahlan, C. E. Barry, B. G. Davis, *Nat. Chem. Biol.*, 2011, **7**, 228-235.

[16] R. Namme, T. Mitsugi, H. Takahashi, S. Ikegami, *Eur. J. Org. Chem.*, 2007, **22**, 3758-3764.

[17] M. R. Pratt, C. D. Leigh, C. R. Bertozzi, *Org. Lett.*, 2003, **5**, 3185-3188.

[18] C. D. Leigh, C. R. Bertozzi, *J. Org. Chem.*, 2008, **73**, 1008-1017.

[19] S. Izumi, Y. Kobayashi, Y. Takemoto, *Angew. Chem. Int. Ed.*, 2020, **59**, 14054-14059.

[20] B. M. Swarts, C. M. Holsclaw, J. C. Jewett, M. Alber, D. M. Fox, M. S. Siegrist, J. A. Leary, R. Kalscheuer, C. R. Bertozzi, *J. Am. Chem. Soc.*, 2012, **134**, 16123-16126.

[21] S. Hanessian, P. Lavallée, *J. Antibiot.*, 1972, **2**, 683-684.

[22] P. A. Wallace, D. E. Minnikin, *J. Chem. Soc., Chem. Commun.*, 1993, 1292-1293.

[23] R. Lattrell, G. Lohaus, *Justus Liebigs Ann. Chem.*, 1974, **1974**, 901-920.

[24] R. Albert, K. Dax, R. W. Link, A. E. Stütz, *Carbohydr. Res.*, 1983, **118**, C5-C6.

[25] R. W. Bassily, R. I. El-Sokkary, B. Azmy Silwanis, A. S. Nematalla, M. A. Nashed, *Carbohydr. Res.*, 1993, **239**, 197-207.

[26] F. P. Rodriguez-Rivera, X. Zhou, J. A. Theriot, C. R. Bertozzi, *J. Am. Chem. Soc.*, 2017, **139**, 3488-3495.

[27] F. P. Rodriguez-Rivera, X. Zhou, J. A. Theriot, C. R. Bertozzi, *Angew. Chem. Int. Ed.*, 2018, **57**, 5267-5272.

[28] M. Kamariza, P. Shieh, C. S. Ealand, J. S. Peters, B. Chu, F. P. Rodriguez-Rivera, M. R. B. Sait, W. V. Treuren, N. Martinson, R. Kalscheuer, B. D. Kana, C. R. Bertozzi, *Sci. Transl. Med.*, 2018, **10**, article no. eaam6310.

[29] B. L. Urbanek, D. C. Wing, K. S. Haislop, C. J. Hamel, R. Kalscheuer, P. J. Woodruff, B. M. Swarts, *ChemBioChem*, 2014, **15**, 2066-2070.

[30] S. Peña-Zalbidea, A. Y.-T. Huang, H. W. Kavunja, B. Salinas, M. Desco, C. Drake, P. J. Woodruff, J. J. Vaquero, B. M. Swarts, *Carbohydr. Res.*, 2019, **472**, 16-22.

[31] H. L. Parker, R. M. F. Tomás, C. M. Furze, C. S. Guy, E. Fulham, *Org. Biomol. Chem.*, 2020, **18**, 3607-3612.

[32] T. Dai, J. Xie, Q. Zhu, M. Kamariza, K. Jiang, C. R. Bertozzi, J. Rao, *J. Am. Chem. Soc.*, 2020, **142**, 15259-15264.

[33] X. Li, P. Geng, X. Hong, Z. Sun, G. Liu, *Chem. Commun.*, 2021, **57**, 13174-13177.

[34] M. Carlier, E. Lesur, A. Baron, A. Lémétais, K. Guitot, L. Roupnel, C. Dietrich, G. Doisneau, D. Urban, N. Bayan, J.-M. Beau, D. Guianvarc'h, B. Vauzeilles, Y. Bourdreux, *Org. Biomol. Chem.*, 2022, **20**, 1974-1981.

[35] Y. Bourdreux, A. Lemétais, D. Urban, J.-M. Beau, *Chem. Commun.*, 2011, **47**, 2146-2148.

[36] M. Kamariza, S. G. L. Keyser, A. Utz, B. D. Knapp, C. Ealand,

G. Ahn, C. J. Cambier, T. Chen, B. Kana, K. C. Huang, C. R. Bertozzi, *JACS Au*, 2021, **1**, 1368-1379.

[37] N. Banahene, D. M. Gepford, K. J. Biegas, D. H. Swanson, Y.-P. Hsu, B. A. Murphy, Z. E. Taylor, I. Lepori, M. S. Siegrist, A. Obregón-Henao, M. S. van Nieuwenhze, B. M. Swarts, *Angew. Chem. Int. Ed.*, 2023, **62**, article no. e202213563.

[38] S. H. Liyanage, N. G. Hasitha Raviranga, J. G. Ryan, S. S. Shell, O. Ramström, R. Kalscheuer, M. Yan, *JACS Au*, 2023, **3**, 1017-1028.

[39] P. L. van der Peet, C. Gunawan, S. Torigoe, S. Yamasaki, S. J. Williams, *Chem. Commun.*, 2015, **51**, 5100-5103.

[40] V. Ratovelomanana-Vidal, C. Girard, R. Touati, J. P. Tranchier, B. B. Hassine, J. P. Genêt, *Adv. Synth. Catal.*, 2003, **345**, 261-274.

[41] N. Tahiri, P. Fodran, D. Jayaraman, J. Buter, M. D. Witte, T. A. Ocampo, D. B. Moody, I. Van Rhijn, A. J. Minnaard, *Angew. Chem. Int. Ed.*, 2020, **59**, 7555-7560.

[42] R. Toubiana, B. C. Das, J. Defaye, B. Mompon, M.-J. Toubiana, *Carbohydr. Res.*, 1975, **44**, 308-312.

[43] V. A. Sarpe, S. S. Kulkarni, *J. Org. Chem.*, 2011, **76**, 6866-6870.

[44] H. N. Foley, J. A. Stewart, H. W. Kavunja, S. R. Rundell, B. M. Swarts, *Angew. Chem. Int. Ed.*, 2016, **55**, 2053-2057.

[45] T. J. Fiolek, N. Banahene, H. W. Kavunja, N. J. Holmes, A. K. Rylski, A. A. Pohane, M. S. Siegrist, B. M. Swarts, *ChemBioChem*, 2019, **20**, 1282-1291.

[46] E. Lesur, A. Baron, C. Dietrich, M. Buchotte, G. Doisneau, D. Urban, J.-M. Beau, N. Bayan, B. Vauzeilles, D. Guianvarc'h, Y. Bourdreux, *Chem. Commun.*, 2019, **55**, 13074-13077.

[47] H. W. Kavunja, K. J. Biegas, N. Banahene, J. A. Stewart, B. F. Piligan, J. M. Groenevelt, C. E. Sein, Y. S. Morita, M. Niederweis, M. S. Siegrist, B. M. Swarts, *J. Am. Chem. Soc.*, 2020, **142**, 7725-7731.

[48] P. Haberkant, R. Raijmakers, M. Wildwater, T. Sachsenheimer, B. Brügger, K. Maeda, M. Houweling, A.-C. Gavin, C. Schultz, G. van Meer, A. J. R. Heck, J. C. M. Holthuis, *Angew. Chem., Int. Ed.*, 2013, **52**, 4033-4038.

[49] C. Thiele, M. J. Hannah, F. Fahrenholz, W. B. Huttner, *Nat. Cell Biol.*, 2000, **2**, 42-49.

[50] C. S. Barry, K. M. Backus, C. E. Barry, B. G. Davis, *J. Am. Chem. Soc.*, 2011, **133**, 13232-13235.

[51] H. L. Hodges, R. A. Brown, J. A. Crooks, D. B. Weibel, L. L. Kiessling, *Proc. Natl. Acad. Sci. USA*, 2018, **115**, 5271-5276.

[52] N. J. Holmes, H. W. Kavunja, Y. Yang, B. D. Vannest, C. N. Ramsey, D. M. Gepford, N. Banahene, A. W. Poston, B. F. Piligian, D. R. Ronning, A. K. Ojha, B. M. Swarts, *ACS Omega*, 2019, **4**, 4348-4359.

[53] X. Zhou, F. P. Rodriguez-Rivera, H. C. Lim, J. C. Bell, T. G. Bernhardt, C. R. Bertozzi, J. A. Theriot, *Nat. Chem. Biol.*, 2019, **15**, 221-231.

[54] A. A. Pohane, D. J. Moore, I. Lepori, R. A. Gordon, T. O. Nathan, D. M. Gepford, H. W. Kavunja, I. V. Gaidhane, B. M. Swarts, M. S. Siegrist, *ACS Infect. Dis.*, 2022, **8**, 2223-2231.

Comptes Rendus
Chimie
2023, Vol. 26, Special Issue S3, p. 39-52
https://doi.org/10.5802/crchim.256

Breaking Barriers in Chemical Biology – Toulouse 2022

Chemical Biology of G-quadruplex and i-motif DNA: use of topologically constrained DNA

Jérôme Dejeu [a,b] **and Eric Defrancq** [*,a]

[a] Université Grenoble-Alpes, CNRS, UMR 5250, Département de Chimie Moléculaire (DCM), 38000 Grenoble, France

[b] SUPMICROTECH, Université de Franche-Comté, CNRS, Institut FEMTO-ST, F-25000 Besançon, France

E-mails: jerome.dejeu@femto-st.fr (J. Dejeu), eric.defrancq@univ-grenoble-alpes.fr (E. Defrancq)

Abstract. Tetrameric DNA structures such as G-quadruplex (**G4**) and i-motif (**i-DNA**) have attracted increasing interest in the last decades. They are indeed involved in many biological processes including translation regulation, pre-mRNA processing, mRNA targeting, telomere maintenance, etc. We have developed chemical tools named TASQ (Template-Assembled Synthetic Quadruplex) to address the following scientific goals: (i) identify unambiguous (i.e., affine and specific) **G4**- and **i-DNA**-interacting ligands, (ii) identify proteins interacting with those structures and determine their cellular relevance and (iii) select specific antibodies for **G4** and **i-DNA**. This review reports on our works over the past decade.

Keywords. G-quadruplex, i-motif, DNA, SPR, BLI, Oligonucleotide conjugate, Click chemistry.

Funding. The authors wish to acknowledge the institutions for providing financial support: Université Grenoble-Alpes, CNRS, Labex ARCANE (ANR-11-LABX-0003-01) and CBH-EUR-GS (ANR-17-EURE-0003), Agence Nationale de la Recherche (Quarpdiems grant ANR-12-BSV8-0008-04, G4-TopiPro grant ANR-16-CE11-0006-01, iCARE grant ANR-21-CE44-0005-02) and Région Auvergne-Rhône-Alpes.

Manuscript received 7 February 2023, revised and accepted 24 August 2023.

1. Introduction

The double-helical structure of DNA in which two antiparallel strands are held together through canonical A/T and G/C base pairing was established over half a century ago. However, the last decades brought accumulating evidence of the existence and biological relevance of four-stranded nucleic acid structures namely G-quadruplex (**G4**) and i-motif (**i-DNA**). **G4** structures could be formed from guanine rich sequences and consist in stacked tetrads of Hoogsteen hydrogen-bonded guanine nucleobases (i.e. G-tetrad or quartet, Figure 1A), connected by various loop-forming sequences, and stabilized through the coordination of physiologically abundant cations (Na^+, K^+) [1,2].

Bioinformatics studies suggested that the human genome contains around 370,000 sequences having the potential to form stable G-quadruplex structures (PQS) [3]. This was reevaluated using a novel algorithm and high-resolution sequencing-based method (termed "G4-seq"), which identified more than 700,000 PQS within the genome [4,5]. Interestingly, these putative **G4s** are not distributed randomly in the genome. Indeed a statistically significant enrichment of PQS was found in several relevant domains of the genome. DNA **G4**-forming sequences can be found in the telomeric region where their stabilization has been shown to inhibit activity of telomerase, which is over-expressed in

* Corresponding author

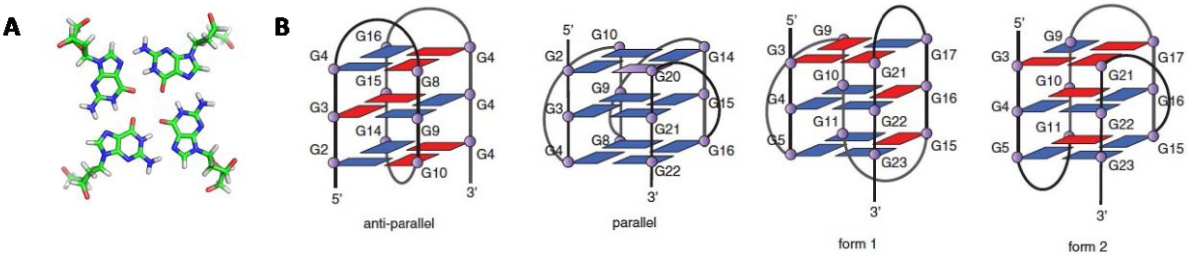


Figure 1. (A) G-tetrad (or G-quartet), (B) schematic representations of some G-quadruplex topologies with three G-tetrads from reference [2].

80% of cancer cells, thus evidencing their potential as anticancer drug targets [6,7]. However **G4** formation is not limited to the telomeric region: they are also over-represented in the promoter regions of a number of genes, including proto-oncogenes c-Myc, c-Kit, bcl-2 and KRAS [8]. Furthermore, the majority of the 250,000 human replication origins are close to **G4** motifs suggesting that the formation of stable **G4** structures participates in the initiation of replication [9]. It has been reported that certain pathologies or chronic diseases caused by cell dysfunction might involve the presence of **G4**. **G4** formation has been linked to genetic disorders (diabetes, fragile X disorder, Bloom syndrome), age-related degenerative illness (ALS, FTD) and cancer (telomere, MYC, Kit, BCL-2). **G4** formation has also been evidenced in the genomes of viruses suggesting functional significance [10]. Besides **G4-DNA**, G-rich RNA sequences are also prone to fold into stable **G4** architectures (**G4-RNA**) [11]. **G4-RNA**-forming sequences can be found in the 5′- and 3′-untranslated regions of many genes, and also in the open reading frame of some mRNAs [11]. To date, the formation of **G4-RNA** has been involved in several biological processes linked to RNA metabolism such as translation regulation, pre-mRNA processing, and mRNA targeting. Owing to the single-stranded nature of transcribed RNA, in vivo formation of **G4-RNA** is expected to occur more easily than **G4-DNA**. Strong arguments have been provided that argue in favor of the formation of DNA and RNA **G4** structures within cells, by using **G4**-specific antibodies [12], in vivo NMR [13], and binding-activated fluorescent **G4**-targeting ligands [14].

An essential feature of **G4** is their intrinsic polymorphic nature: numerous in vitro studies have revealed their susceptibility to adopt different topologies, which are in equilibrium. Indeed, depending on the length and composition of the sequence, as well as the environmental conditions (including the nature and concentration of metal cations, and local molecular crowding), a G-quadruplex-forming sequence can adopt different topologies in which the strands are in parallel or antiparallel conformations, with the co-existence of different types of loops (lateral, diagonal or propeller) with variable lengths (Figure 1B) [1].

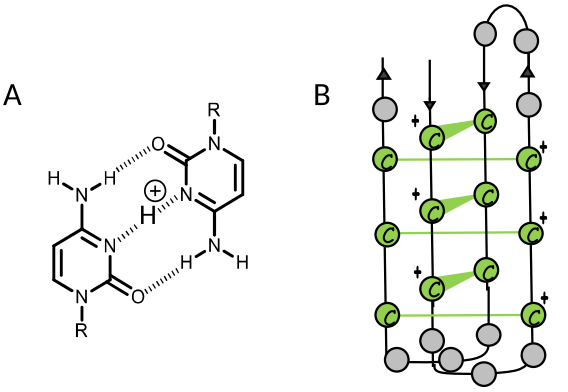


Figure 2. (A) (CH^+:C) base pairing, (B) schematic representation of **i-DNA**.

Cytosine, the complementary nucleobase of guanine, is also prone to assemble to form four-stranded structures named i-motif DNA (**i-DNA**) in which cytosines are intercalated via a stack of hemi-protonated (CH^+:C) base pairs (Figure 2) [15].

A major characteristic of **i-DNA** is the strong pH-dependency of its stability and formation. Indeed, **i-DNA** structures are typically observed in vitro at acidic pH, a particularity that has cast doubt on their existence in cell. However, two independent stud-

ies have recently demonstrated the stability of exogenous **i-DNA** structures in human cells through in cellulo NMR spectroscopy [16] as well as the presence of endogenous **i-DNA** in the nuclei of human cells through immunofluorescence using an antibody (*i-Mab*) raised against the i-motif [17].

The biological relevance of **i-DNA** has been less investigated mainly due to the scientific community's skepticism about the existence of i-motifs in vivo. A recent review from Brown and Kendrick provides some insight into the biological function of **i-DNA** structure [18]. In the context of bcl-2 oncogene, a transcriptional activation was reported, caused by the formation of stable i-motif through the interaction with IMC-48 compound, a cholestane derivative (selected from a screening of a NCI diversity set library) [19]. In contrast with bcl-2 oncogene, a transcriptional repression was reported with c-MYC i-motif [20]. Also, proteins such as hnRNP LL have been identified to interact with i-motif structures acting as an activating transcription factor [21].

The intrinsic polymorphism associated with the formation of **G4** and **i-DNA** as well as the pH-dependency of **i-DNA** stability represent severe bottlenecks for the studies of those tetrameric DNA structures. Indeed, the polymorphism could lead to intricate structural mixtures in solution that can complicate the rationalization of the relationships between **G4** or i-motif structures and recognition by proteins and ligands. Likewise, low-pH conditions used to induce the formation of **i-DNA** could lead to the protonation of many ligands (e.g., proteins), strongly increasing their non-specific nucleic acid binding. The design of chemical tools able to reduce the structural heterogeneity of **G4** and **i-DNA** as well as able to improve i-motif stability in physiological conditions is thus of high interest.

In this context, we developed some years ago an innovative concept that consists to constrain the accessible topologies of a G-quadruplex-forming sequence to a single one [22]. This strategy named TASQ for Template-Assembled Synthetic Quadruplex, is based on the use of a rigid cyclic peptide scaffold with two independently functionalizable faces, resulting from the orientation of the lysine side-chains. One face is dedicated to the anchoring of different oligonucleotide sequences to obtain the desired **G4** topology and a biotin residue is incorporated on the other side for attachment to streptavidin immobilized surfaces for various applications (Figure 3). This template concept allowed the formation of very stable G-quadruplex motifs in a unique conformation, in aqueous medium. It was next extended to the formation of constrained i-motif DNA.

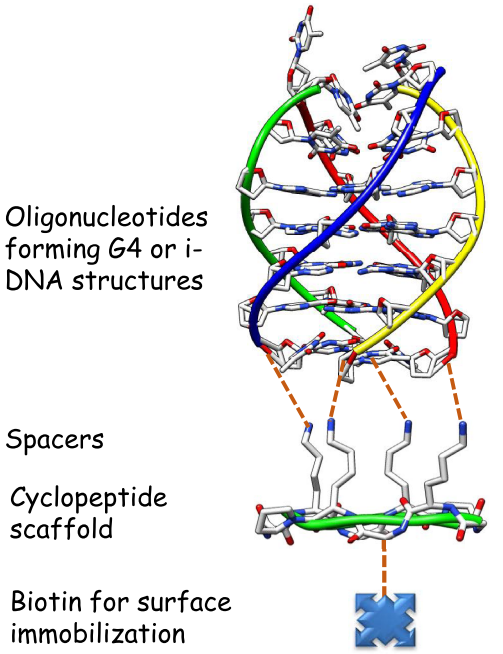


Figure 3. Schematic representation of the TASQ concept. The different oligonucleotide sequences forming the target structure are attached to the cyclopeptide scaffold through different ligation techniques (see below) on the top of the cyclopeptide and a biotin residue is incorporated in the lower face of the cyclopeptide.

In this article, we describe our contribution for the design of efficient chemical tools based on constrained nucleic acids for the study of **G4** and i-motif DNA. Different applications and perspectives of those chemical tools are then described.

2. Results and discussions

2.1. *Synthesis of the various constrained G-quadruplex and i-DNA systems*

The design of constrained DNA in the chemical biology research domain has already been investigated.

As an example, Escudier *et al.* have developed modified oligonucleotides in which the phosphodiester internucleotidic linkage is replaced with a dioxo-1,3,2-oxaza-phosphorinane moiety resulting in conformationally constrained nucleotides (CNA) [23]. In the case of G-quadruplex, the use of various templates to pre-organize G-quartet assemblies has also been described by different groups [24–27].

Our approach consisted in the use of cyclic peptide scaffolds based on the TASP concept (Template-Assembled Synthetic Proteins) for the design of folded proteins developed by Mutter in 1985 [28]. The chemoselectively addressable cyclic peptide template is the key intermediate as it exhibits two independent and chemically addressable domains, which allows the sequential and regioselective assembly of the different oligonucleotides forming the tetrameric nucleic acid target on one face, the other face serving for the attachment on surfaces. By using sequential ligation techniques we were able to prepare different G-quadruplex systems (Figure 4).

Conjugate **1** was the first constrained **G4** prepared [22]. It may mimic intermolecular-like G-quadruplexes. Four oligonucleotides derived from human telomeric sequence d(5′-TTAGGGT-3′) were attached onto the peptide scaffold by using oxime bond formation from 3′-aldehyde-containing oligonucleotides and the peptide scaffold bearing four aminooxy residues. By using CD melting studies, we demonstrated that the peptide template allows the formation of a very stable **G4** motif in a unique parallel conformation, in aqueous medium. Using the same oxime ligation (OL) method, we later synthetized the corresponding **RNA G4 2** from telomeric sequence (TERRA) and found again that the template allows stabilization of the desired parallel topology [29].

The synthesis of antiparallel topology required the use of two successive ligation reactions for the attachment of oligonucleotides at both 3′ and 5′ extremities. This was achieved through sequential oxime (OL) and Cu(I)-catalyzed azide–alkyne cycloaddition (CuAAc) reactions. The antiparallel topologies of G-quadruplex DNA from telomeric sequence **3** was obtained by the reaction of 3′-aldehyde, 5′-alkyne bis-functionalized oligonucleotides with the suitable cyclopeptide and was shown to exhibit high stability and reduced polymorphism [30].

Next, we expanded the TASQ approach to stabilize biologically relevant viral **G4** structures such as the one found in sequence (5′-TGGCCTGGGCGGGACTGGG-3′) derived from the LTR region of HIV-1 [31]. Unlike with telomeric **G4**-forming conjugates **1-3**, the site-specific attachment onto the cyclopeptide scaffold of the two G-rich DNA oligonucleotides with sequences 5′-TGGCCTGGGC-3′ and 5′-GGACTGGG-3′, respectively, mimicking the sequence from LTR region of HIV-1, implied the use of an additional orthogonal chemical ligation step along with OL and CuAAC. This was achieved using a SN2-thiol coupling reaction (TC). Conjugate **4** was thus prepared through successive conjugations, first with a functionalized oligonucleotide bearing an aldehyde at its 3′-end and an alkyne at its 5′-end, then with another one bearing a thiol function at its 5′-end and an alkyne at its 3′-end [32].

Lastly, the construction of conjugate **5**, a mimic of i-motif DNA formed from the telomeric sequence, was carried out. The synthesis of the i-motif structural mimic **5** was achieved via the stepwise assembly of peptide–DNA conjugates through four successive ligations with one OL, one TC and two CuAAC reactions [33]. The resulting conjugate **5** was found by CD to fold, at room temperature, into an i-motif structure which is stable at acidic and neutral pH. It may therefore be used to study, at physiologically relevant pH, the interaction of the i-motif with putative i-motif targeting ligands (i.e., small molecules or proteins).

2.2. *Applications of the constrained DNA chemical tools*

The different constrained G-quadruplex and i-motif DNA **1-5** were used to study previously described ligands or new molecules able to interact with the **G4** or **i-DNA** targets. The main objectives of those studies were to investigate the interactions of the ligands with different DNA targets and to assess their selectivity versus other DNA structures as well as versus different G-quadruplex topologies.

With these chemical tools in our hands, another goal was to identify, by classical capture methods, and characterize proteins which bind to a predetermined single **G4** topology, and study their interactions with the diverse structural motifs (i.e.,

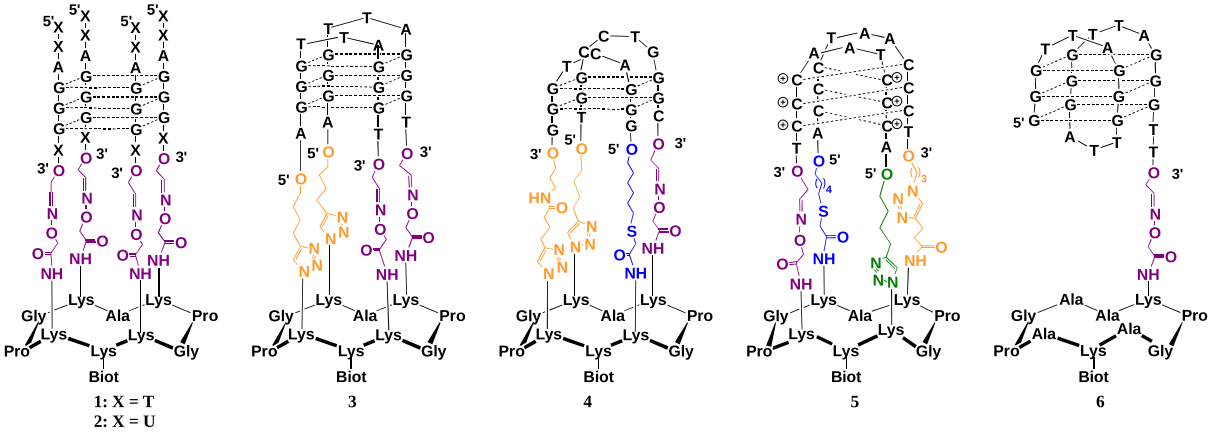


Figure 4. Structure of the different constrained tetrameric nucleic acid structures **1**–**5** and unconstrained control **6**.

loops, grooves, quartets) of the quadruplexes by comparing their binding properties to different defined constructs. This application is now extended to **i-DNA**.

We also envision to use the constrained G-quadruplex to produce and characterize antibodies for a given **G4** topology by using the above-described chemical tools. We believe that preventing the equilibrium between the different conformations that are associated with G-quadruplex-forming sequences will facilitate the production of specific antibodies. Again this will be extended to **i-DNA**.

2.3. *Use of constrained systems to study the interactions with ligands*

Most of the knowledge of the impact of **G4-DNA** secondary structures on cell metabolism resulted from the use of selective chemical probes that bind or modulate the formation of such structures [34]. A major challenge in G-quadruplex ligand synthesis is the development of compounds that are able to distinguish G-quadruplexes from duplex DNA and also discriminate between various **G4** topologies. A detailed picture of quadruplex structure is emerging from crystallographic and NMR studies, and together with computer modeling, it is possible to develop a rational approach toward the design and optimization of quadruplex-stabilizing compounds [35]. The desirable features of these stabilizing molecules are (i) a π-delocalized system that is able to stack on the face of a guanine quartet; (ii) a partial positive charge that lies in the center of the quartet, increasing stabilization by substituting for the cationic charge of the potassium or sodium that would normally occupy that site; and (iii) positively charged substituents that will interact with the grooves and loops of the quadruplex and the negatively charged backbone phosphates.

In contrast to **G4** ligands, relatively few molecules were reported to interact with **i-DNA**, and a controversy concerning their binding mode, affinity, and selectivity persists in the literature [36,37]. The main challenges in this regard are the strong pH-dependency, flexibility and polymorphism of **i-DNA**, introducing potential bias into screening methods. Indeed, low-pH conditions used to induce the formation of i-DNA lead to the protonation of many ligands, strongly increasing their non-specific nucleic acid binding.

To investigate the interactions of our constrained tetrameric nucleic acids with potential ligands, two optical techniques were used: surface plasmon resonance (SPR) and bio-layer interferometry (BLI). Those two label-free techniques are widely used to study the interactions of ligands (including proteins, nucleic acids, sugars, and small molecules) with analytes. The ligand is immobilized on the surface while the analyte is injected close to the surface via a micro-fluidic system for SPR or deposited in microplate for BLI. The sensorgram fittings provide the association and dissociation kinetic constants, and the responses obtained at the steady state (Req)

afford the equilibrium dissociation constant (K_D). Those two techniques display a number of advantages, including the non-use of special radioactive or fluorescent labeling of the molecules, the time efficiency, the use of very low quantity of materials associated with a high sensitivity, the access to a variety of commercial surface sensors and the possibility to assemble homemade sensors bearing specific chemical functionalities. Possible drawbacks of BLI/SPR techniques are related to the relatively high cost of such equipment as well as the requirement of a good expertise (i.e., to not over/mis-interpret the results).

2.3.1. *Study of the interaction with well-known G-quadruplex ligands*

A large number of G-quadruplex ligands have been reported in the literature and most of them interact with G-quadruplex DNA by π-stacking interactions with the external G-quartet of the quadruplex [38,39]. We have used some of them with the aim of verifying if the constrained G-quadruplex systems could act as efficient mimics before the investigation of unknown ligands. Moreover, we envisioned that our different G-quadruplex systems could afford some information about the mode of interaction. The following reported ligands TMPyP4 [40], MMQ1 [41], distamycin [42], FRHR [43], and DODC [44] were first studied (Figure 5). As anticipated, ligands displaying a π-stacking binding mode such as TMPyP4 showed a higher binding affinity for intermolecular-like G-quadruplex **1** due to the absence of loops which could prevent the interactions, whereas ligands with other binding modes (groove and/or loop binding) such as distamycin showed no significant difference in their binding affinities for the constrained quadruplex **1** and unconstrained control **6** [45]. In addition, the method has also provided information about the selectivity of ligands for G-quadruplex DNA over the duplex DNA through comparative studies with DNA hairpin duplex. Further studies with other well-known G-quadruplex ligands such as Phen-DC3 [46], PDS [47], BRACO-19 [48] and NMM [49] (Figure 5) were carried out next. The use of constrained or not constrained **G4** systems also allowed to obtain some information about the selectivity for the ligands with a single **G4** topology. Most of the described ligands do not show any **G4** topology preference, except NMM. We have demonstrated the high selectivity of NMM for the parallel **G4** structure with a dissociation constant at least ten times lower than those of other **G4** topologies as well as the ability of this ligand to shift the **G4** conformation from both the hybrid and antiparallel topologies toward the parallel structure [50].

The studies with well-known G-quadruplex ligands thus validated that the constrained systems are useful tools for investigating the interactions with G-quadruplexes. With those chemical tools in hands, we next investigated the interactions with various ligands designed by our collaborators.

2.3.2. *Study of the interaction of different families of G-quadruplex ligands based on metal complexes*

In the field of targeting G-quadruplex nucleic acids structures with small molecules, hundreds of ligands have been reported [38,39]. Most of them interact with G-quadruplex DNA by π-stacking interactions with the external G-quartet of the quadruplex. The design of metal complexes targeting G-quadruplex DNA has also attracted intense interest [51]. In comparison to organic compounds, metal complexes show many advantages, such as a net positive charge (i.e., able to increase the interactions with DNA), tunable geometry, and, most interestingly, some of them display potentially useful photochemical properties. In that context, we were interested in the design and study of different classes of G-quadruplex binders based on metal complexes including metal porphyrin derivatives, salophens, and ruthenium and iridium photoreactive complexes.

Porphyrin-based ligands. One of the first reported G-quadruplex ligands was the non-metalated porphyrin, meso-5,10,15,20-tetrakis(4-N-methylpyridiniumyl)-porphyrin (TMPyP4, Figure 5). The main disadvantage of TMPyP4 is its weak selectivity for **G4** versus all other DNA structure. In the aim of improving selectivity, Pratviel *et al.* have designed new porphyrin derivatives through the insertion of a metal ion (Ni^{2+}, Co^{3+}, or Mn^{3+}) into the porphyrin core (TMPyP4 series **8–10**) or by modification of the meso substituents R of the porphyrin with a phenyl-N-methylpyridinium

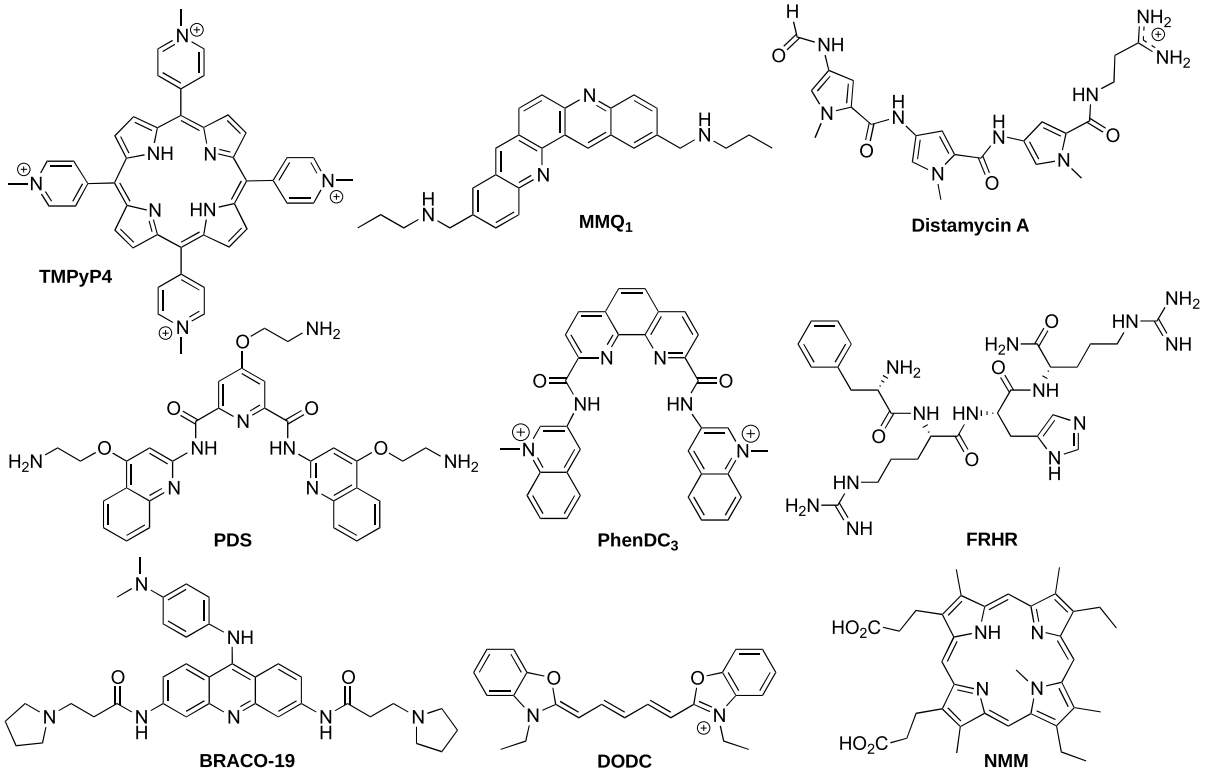


Figure 5. Structure of the different reported ligands.

group (TMPyP4-PP series **11–14**) and a guanidinium group (TMPyP4-PG **15–17**) as depicted in Figure 6 [52–54].

By using the constrained **G4** system **1**, we have demonstrated that the insertion of a metal ion into TMPyP4 (Ni, Co or Mn) or modification of the porphyrin meso substituents could drastically modify affinity and selectivity for the **G4**. For example TMPyP4 derivatives **9** and **10** with cobalt and manganese metal respectively, showed a high selectivity for G-quadruplexes **1** and **6** versus duplex DNA as no interaction occurs with duplex DNA, whereas for TMPyP4 derivative **8** with nickel metal as well as for parent TMPyP4 **7** an interaction for duplex DNA quite equivalent to that for G-quadruplex **6** was observed (Figure 7). That was because water/hydroxo as axial ligands on the cobalt and manganese derivatives could preclude the intercalation of the porphyrin moiety between the base pairs of duplex. For TMPyP4-PP **11–14** and TMPyP4-PG **15–17** series, the presence of bulky substituents should also prevent the intercalation between the base pairs of duplex DNA leading to a weak interaction of those compounds with duplex DNA in comparison with quadruplexes **1** and **6** (Figure 7).

Guanidinium phenyl porphyrin derivatives **15–17** exhibit moderate cytotoxicity toward cells in culture. Strikingly, the nickel porphyrin derivative **15** was able to displace hPOT1 sheltering protein from telomeres in human cells [53].

Ni-Salphen derivatives. Pioneering works by Neidle *et al.* revealed that salphen derivatives bind strongly to the human telomeric G-quadruplexes and inhibit the telomerase activity with EC_{50} of roughly 0.1 μM [55]. In order to study the impact on binding affinity of the length of the side-chains and their positions on the salphen scaffold, Thomas *et al.* have prepared new family of G-quadruplex binders based on the nickel(II) salphen platform (Figure 8). The side-chains are alkyl-imidazolium arms connected at para, ortho or meta positions of the phenol moieties [56,57]. The affinity for G-quadruplex DNA **1**

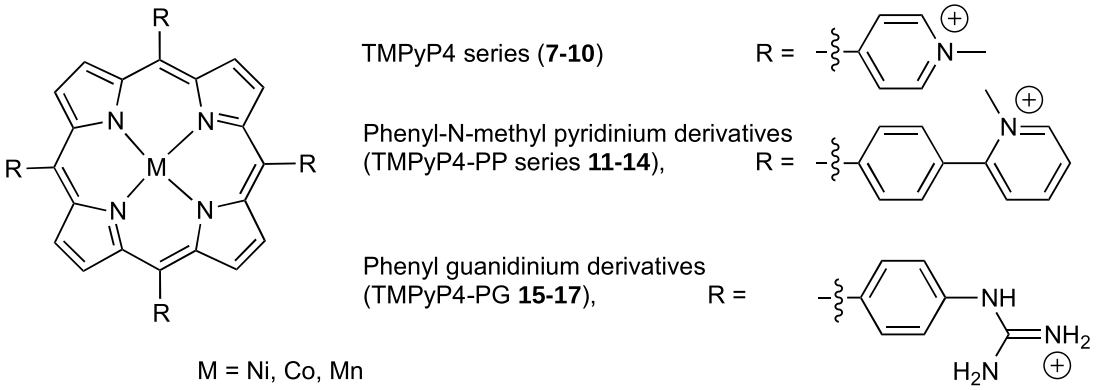


Figure 6. Structure of metalated TMPyP4 derivatives. The un-metalated ligands TMPyP4 **7**, TMPyPP4-PP **11** and TMPyP4-PG **15** correspond to those structures without M.

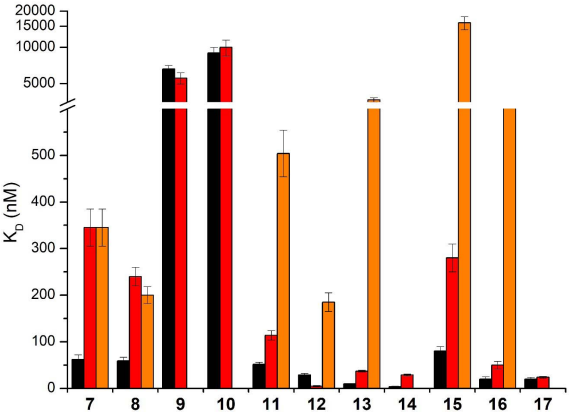


Figure 7. K_D values obtained with the different porphyrin derivatives **7**–**17** in interaction with G-quadruplexes **1** (black), unconstrained control **6** (red) and duplex DNA (orange). **7**: TMPyP4, **8**: Ni-TMPyP4, **9**: Co-TMPyP4, **10**: Mn-TMPyP4, **11**: TMPyP4-PP, **12**: Ni-TMPyP4-PP, **13**: Co-TMPyP4-PP, **14**: Mn-TMPyP4-PP, **15**: TMPyP4-PG, **16**: Ni-TMPyP4-PG, **17**: Mn-TMPyP4-PG. No histogram for duplex DNA means that the K_D value could not be determined due to too weak interaction. The reported values are the means of representative independent experiments, and the errors provided are standard deviations from the mean. Each experiment was repeated at least three times.

and **6** as well as the selectivity versus duplex DNA were evaluated by using SPR.

The different salphen derivatives **18–29** showed K_D values in the 0.1–2 μM range for both G-quadruplex **1** and **6**, which are in the range of those reported for related compounds interacting with the HTelo sequence [58] and most of them do not bind tightly to duplex DNA (Figure 9).

SPR studies with our systems **1** and **6** allowed to obtain some useful information. We observed that the shorter the side arms, the higher the affinity for **G4**. Furthermore the introduction of a third anchor on the diaminobenzene bridge also improved the affinity for **G4**. The difference of affinity for **1** versus **6** also suggested that the compounds interact both by π-stacking over the tetrad and electrostatic interactions in the grooves. From the SPR studies, we have concluded that salphen **21**, **25** and **29** were the optimal G-quadruplex binders: they were also the best salphen derivatives able to inhibit telomerase activity and in particular with an IC_{50} value measured from TRAP-G4 assays of 70 nM for **29**.

Ruthenium and iridium complexes. Ruthenium(II) and iridium metal complexes have been investigated as **G4**-ligands and some of them have shown good affinity and selectivity toward **G4s** [51,59–61]. In this context, Elias *et al.* have developed several ruthenium and iridium complexes which are able to target G-quadruplex DNA (Figure 10). The affinity for G-quadruplexes and the selectivity versus DNA duplexes were investigated using BLI and SPR (Figure 11).

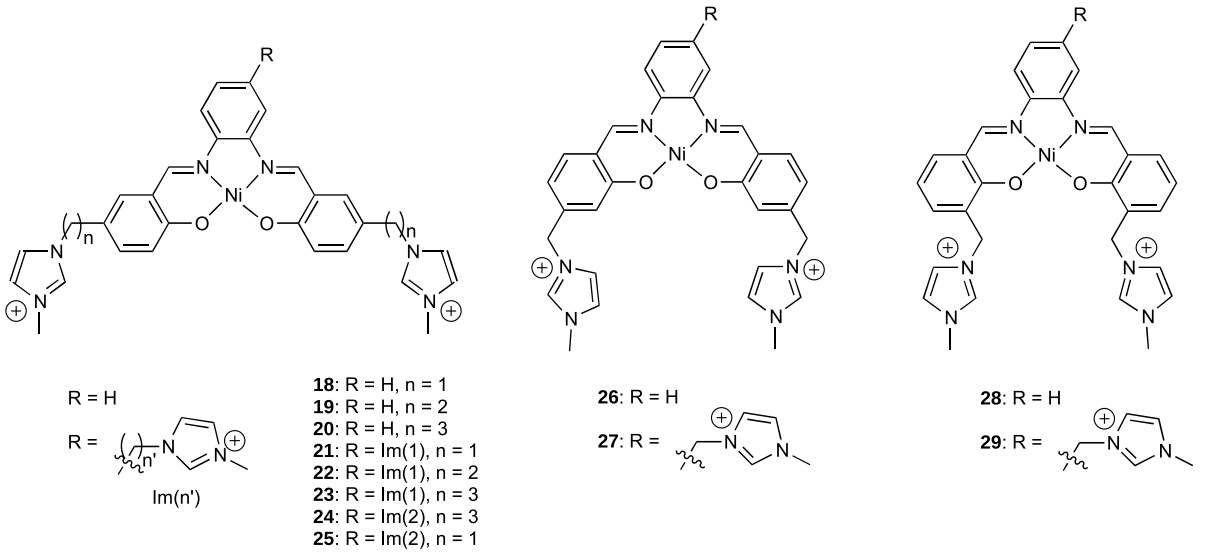


Figure 8. Structure of the different Ni-Salphen derivatives.

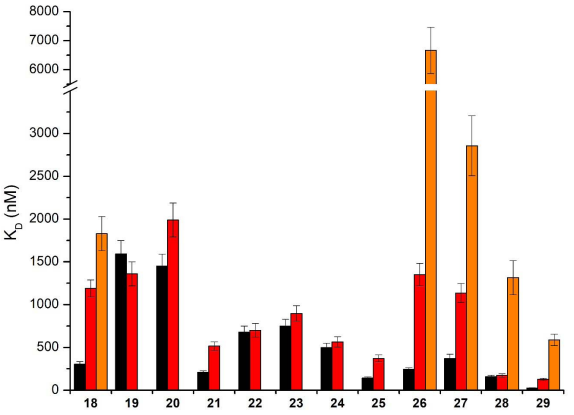


Figure 9. K_D values obtained with the Ni-salphen derivatives **18–29** in interaction with G-quadruplexes **1** (black), **6** (red) and duplex DNA (orange). No histogram for duplex DNA means that the K_D value could not be determined due to too weak interaction. The reported values are the means of representative independent experiments, and the errors provided are standard deviations from the mean. Each experiment was repeated at least three times.

A first series based on dph (dph = dipyrazino [2,3-a:2′,3′-h]phenazine) and bpph (bpph = benzo[a] pyrazino[2,3-h]phenazine) ancillary ligands (complexes **30–33**) were studied using G-quadruplex systems **1** and **6** [62,63]. It was shown that the removal of two non-chelating nitrogen atoms from the dph ligand in complex **30** to form the bpph ligand in complex **31** led to a huge impact on the interactions with **G4** and duplex DNA. Indeed, a decrease of the affinity toward **G4** structure **6** was observed (K_D = 9 μM and 32 μM for complexes **30** and **31**, respectively), while a weak affinity for duplex DNA could be measured with complex **31** (K_D = 102 μM) thus leading to a partial loss of selectivity. The dph analogues therefore appeared to be slightly more selective toward **G4** versus duplex DNA compared to their respective bpph analogues.

Another series was based on the CPIP (2-(4-chlorophenyl)-1H-imidazo[4,5-f][1,10]phenanthroline) and CPIPTAP ligands (complexes **34–37**). Investigation by BLI indicated that these complexes displayed a good affinity for G-quadruplex DNA (K_D around 1 and 3 μM for **36** and **35,** respectively, for **G4** system **3**) and selectivity over duplex DNA. It was also noticed that their affinities were higher for G-quadruplex structures **3** and **6** which contain TTA loops, than for parallel-stranded quadruplex **1**. This is consistent with interactions of the complexes with G-quadruplexes through mixed π-stacking with the guanine tetrad and further interactions with loops

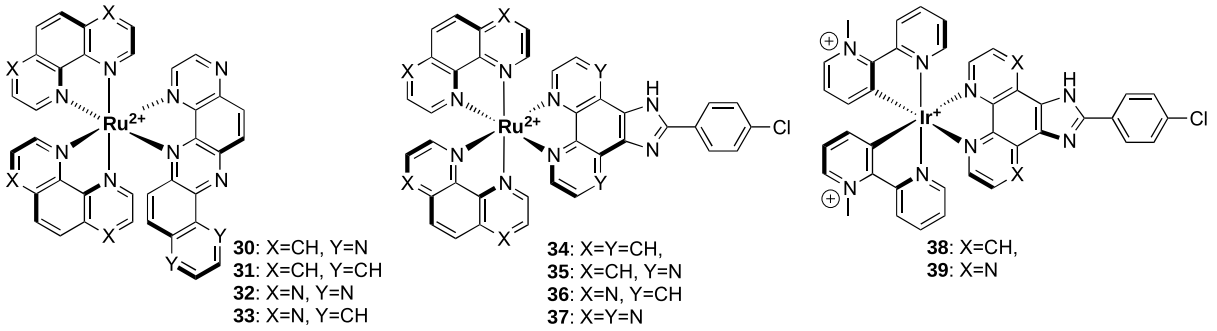


Figure 10. Structure of ruthenium and iridium complexes from Elias collaboration.

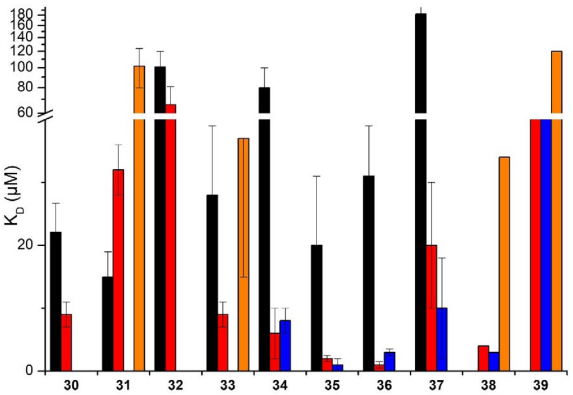


Figure 11. K_D values obtained with different ruthenium and iridium derivatives **30**–**39** in interaction with G-quadruplexes **1** (black), **3** (blue) and **6** (red) and duplex DNA (orange). No histogram for duplex DNA means that the K_D value could not be determined due to too weak interaction. The reported values are the means of representative independent experiments, and the errors provided are standard deviations from the mean. Each experiment was repeated at least three times.

and grooves. Due to their photophysical properties, these complexes are able to react with DNA through type II photoreaction (i.e., formation of singlet oxygen) or through photo-induced charge transfer (PET). Interestingly, complexes **34** and **36** elicited a dramatic photo-cytotoxic effect, as 100% mortality was obtained upon irradiation of U2OS osteosarcoma cells in their presence, whereas very low mortality was observed in the dark at the same drug concentration [64].

The interaction of iridium complexes **38** and **39** with G-quadruplex DNA was also investigated [65]. K_D values in the micromolar range was obtained for the G-quadruplex structures that fall within the range of those reported for similar ruthenium(II) complexes. However, a weak selectivity for the G-quadruplex structure versus duplex DNA was observed. This could be explained by the net positive charge of the Ir^{III} complexes in comparison with the Ru^{II} complexes that could favor non-specific ionic interactions with DNA.

2.3.3. *Study of the interaction with i-motif DNA*

As described below, the stability of **i-DNA** strongly depends on pH. Conditions typically used for screening of ligand candidates with native **i-DNA** sequences employ low pH which, at the same time, leads to protonation of the ligands and favors their non-specific interactions with nucleic acids. In that context, we have recently designed constrained **i-DNA 5** which shows an increased stability and invariability of the i-motif structure in a broad pH range, allowing to use this scaffold as an **i-DNA** substrate at physiologically relevant conditions [33]. This constrained i-DNA **5** was thus used to investigate the interaction between previously reported i-motif DNA ligands (Figures 5 and 12) and folded or unfolded **i-DNA** in acidic (pH 5.5) and near-neutral (pH 6.5) conditions by using BLI.

Very interestingly, we have observed that, despite several ligands such as macrocyclic bis-acridine (BisA) and pyridostatin (PDS) showing good affinities for the telomeric i-motif forming sequence, none of the ligands displayed selective interactions with the **i-DNA** structure nor were able to promote its

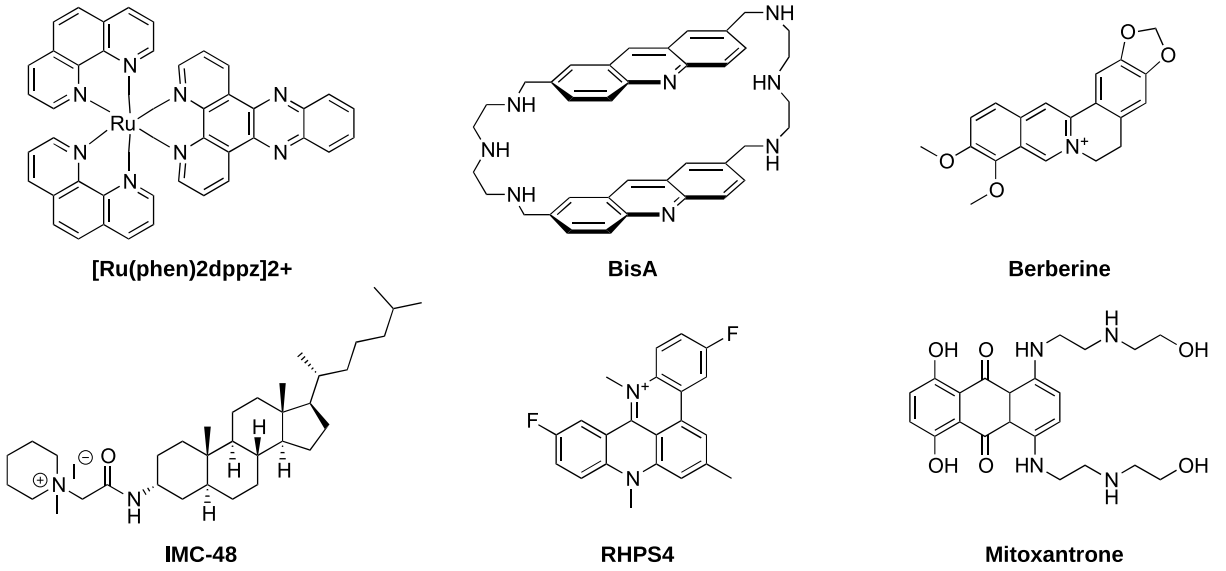


Figure 12. Structure of compounds used for the studies of i-motif interactions. **[Ru(Phen)$_2$dppz]**$^{2+}$ [66], **BisA** [67], **Berberine** [68], **IMC-48** [19], **RHPS4** [69] and **mitoxantrone** [70].

formation (Figure 13) [71]. More recently we have reported that **IMC-48**, although described as **i-DNA** ligand, is a very weak ligand of **i-DNA** as no quantifiable interaction or significant stabilization of i-motif structures could be observed, stimulating a quest for an alternative mechanism of its biological activity [72]. All together, these results further emphasize the need for efforts to identify specific i-motif ligands. In this context, the use of constrained **i-DNA 5**, which ensures an i-motif folding, represents an interesting alternative to identify unambiguous (i.e., affine and specific) i-DNA-interacting ligands.

2.4. *Use of constrained G4 for protein fishing*

Given the increasing roles of **G4** structures in cellular metabolism, extensive researches have been conducted in the last years to identify new **G4**-dependent mechanisms. Notably classical pull-down approaches identified hundreds of proteins associated to G-rich oligonucleotides forming **G4** structures [73]. However, in solution, G-rich single-stranded molecules are in equilibrium between unfolded and folded states, and thus numerous identified **G4** binding proteins are also able to recognize unfolded G-rich sequences [74]. In this context, we have used the constrained G-quadruplex **3** which folds into the single antiparallel topology.

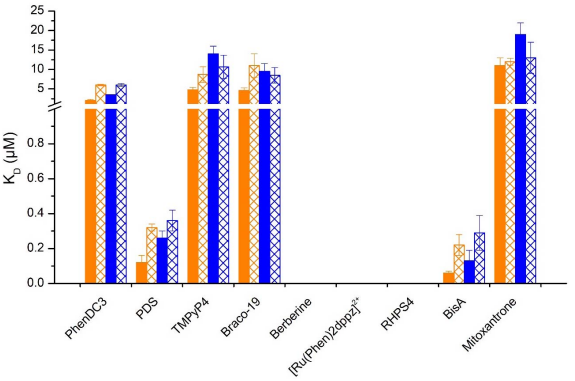


Figure 13. K_D values obtained with different ligands in interaction with I-motif **6** (blue), h-Telo sequence (orange) at pH 5.5 and 6.5 fill and sparse pattern. No histogram for duplex berberine, [Ru(Phen)$_2$dppz]$^{2+}$ and RHPS4 means that the K_D value could not be determined due to too weak interaction. The reported values are the means of representative independent experiments, and the errors provided are standard deviations from the mean. Each experiment was repeated at least three times.

Moreover, such locked **G4** displays a thermal stability significantly higher than unconstrained **G4**

which strongly reduces the possibility to form unfolded single-stranded sequences. We identified through affinity purifications coupled to mass spectrometry (MS)-based quantitative proteomics a set of human proteins associated to locked **G4** structures. Notably, this approach allowed us to identify NELF proteins as a new **G4**-interacting complex, leading us to investigate the impact of RNA-Pol II pausing mechanism in the response to **G4** stabilization by **G4** ligands [75].

3. Conclusion

The different constrained G-quadruplex systems have proved efficient tools for the identification of **G4** structure-specific synthetic ligands. They could also give some interesting information about the mode of interaction and the selectivity versus duplex and also versus **G4** topologies. The latter could be crucial to design more specific **G4** ligands associated with less off-target side effects. The constrained **i-DNA** represents also an interesting tool. Indeed, our recent study has demonstrated that all the molecules described so far as i-motif ligands, are not able to discriminate between folded and unfolded i-motif structures. The constrained **i-DNA** will thus be used to identify unambiguous (i.e., affine and specific) **i-DNA**-interacting ligands.

The pull-down strategy using constrained G-quadruplex DNA was proved efficient to identify proteins selective for a single **G4** topology. Future directions of our approach will concern the construction of constrained **G4** structures mimicking parallel **G4** topologies in order to refine the impact of loops on protein binding. We will also use constrained i-DNA **5** to identify proteins interacting with the non-canonical secondary structure i-motif.

Lastly, given the fact that constrained G-quadruplexes reduce the unfolding to single-stranded sequences, we envisioned to use system **3** for the selection of topologically specific **G4** antibody. In the same way, due to the lower sensitivity of constrained **i-DNA** to pH conditions, we will also use this system for the selection of specific i-motif antibodies.

Declaration of interests

The authors do not work for, advise, own shares in, or receive funds from any organization that could benefit from this article, and have declared no affiliations other than their research organizations.

Acknowledgements

The NanoBio-ICMG platforms (UAR 2607) are acknowledged for their support. We thank also our collaborators for providing the different samples of products: M. P. Teulade-Fichou and A. Grand from Institut Curie Orsay, B. Elias from Université catholique de Louvain, G. Pratviel from CNRS Toulouse, F. Thomas and O. Jarjayes from Université Grenoble-Alpes. The authors are also grateful to all the students who contributed to those works.

References

[1] C. K. Kwok, C. J. Merrick, *Trends Biotechnol.*, 2017, **35**, 997-1013.
[2] R. C. Monsen, J. O. Trent, J. B. Chaires, *Acc. Chem. Res.*, 2022, **55**, 3242-3252.
[3] J. L. Huppert, S. Balasubramanian, *Nucleic Acids Res.*, 2005, **33**, 2908-2916.
[4] V. S. Chambers, G. Marsico, J. M. Boutell, M. Di Antonio, G. P. Smith, S. Balasubramanian, *Nat. Biotechnol.*, 2015, **33**, 877-881.
[5] A. Bedrat, L. Lacroix, J.-L. Mergny, *Nucleic Acids Res.*, 2016, **44**, 1746-1759.
[6] D. Rhodes, H. J. Lipps, *Nucleic Acids Res.*, 2015, **43**, 8627-8637.
[7] S. Neidle, *J. Med. Chem.*, 2016, **59**, 5987-6011.
[8] M. Fry, in *Guanine Quartets* (L. Spindler, W. Fritzsche, eds.), Royal Society of Chemistry, Cambridge, 2013, 223-236.
[9] L. Chen, J. Dickerhoff, S. Sakai, D. Yang, *Acc. Chem. Res.*, 2022, **55**, 2628-2646.
[10] M. Metifiot, S. Amrane, S. Litvak, M. L. Andreola, *Nucleic Acids Res.*, 2014, **42**, 12352-12366.
[11] P. Agarwala, S. Pandey, S. Maiti, *Org. Biomol. Chem.*, 2015, **13**, 5570-5585.
[12] G. Biffi, D. Tannahill, J. McCafferty, S. Balasubramanian, *Nat. Chem.*, 2013, **5**, 182-186.
[13] G. F. Salgado, C. Cazenave, A. Kerkour, J. L. Mergny, *Chem. Sci.*, 2015, **6**, 3314-3320.
[14] A. Laguerre, K. Hukezalie, P. Winckler, F. Katranji, G. Chanteloup, M. Pirrotta, J. M. Perrier-Cornet, J. M. Wong, D. Monchaud, *J. Am. Chem. Soc.*, 2015, **137**, 8521-8525.
[15] K. Gehring, J. L. Leroy, M. Guéron, *Nature*, 1993, **363**, 561-565.
[16] S. Dzatko, M. Krafcikova, D. N. Korkut, R. Hänsel-Hertsch, T. Fessl, R. Fiala, T. Loja, J. L. Mergny, S. Foldynova-Trantirkova, L. Trantirek, *Angew. Chem. Int. Ed.*, 2018, **57**, 2165-2169.
[17] M. Zeraati, D. B. Langley, P. Schofield, A. L. Moye, R. Rouet, W. E. Hughes, T. M. Bryan, M. E. Dinger, D. Christ, *Nat. Chem.*, 2018, **10**, 631-637.
[18] S. L. Brown, S. Kendrick, *Pharmaceuticals*, 2021, **14**, article no. 96.

[19] S. Kendrick, H. J. Kang, M. P. Alam, M. M. Madathil, P. Agrawal, V. Gokhale, D. Yang, S. M. Hecht, L. H. Hurley, *J. Am. Chem. Soc.*, 2014, **136**, 4161-4171.

[20] C. Sutherland, Y. Cui, H. Mao, L. H. Hurley, *J. Am. Chem. Soc.*, 2016, **138**, 14138-14151.

[21] H. A. Assi, M. Garavis, C. Gonzalez, M. J. Damha, *Nucleic Acids Res.*, 2018, **46**, 8038-8056.

[22] P. Murat, D. Cressend, N. Spinelli, A. Van der Heyden, P. Labbé, P. Dumy, E. Defrancq, *ChemBioChem*, 2008, **9**, 2588-2591.

[23] C. Chardet, C. Payrastre, B. Gerland, J.-M. Escudier, *Molecules*, 2021, **26**, article no. 5925.

[24] F. W. Kotch, V. Sidorov, Y.-F. Lam, K. J. Kayser, H. Li, M. S. Kaucher, J. T. Davis, *J. Am. Chem. Soc.*, 2003, **125**, 15140-15150.

[25] M. Nikan, J. C. Sherman, *Angew. Chem. Int. Ed.*, 2008, **47**, 4900-4902.

[26] G. Oliviero, J. Amato, N. Borbone, S. D'Errico, A. Galeone, L. Mayol, S. Haider, O. Olubiyi, B. Hoorelbeke, J. Balzarini, G. Piccialli, *Chem. Commun.*, 2010, **46**, 8971-8973.

[27] R. Haudecoeur, L. Stefan, F. Denat, D. Monchaud, *J. Am. Chem. Soc.*, 2013, **135**, 550-553.

[28] M. Mutter, *Angew. Chem. Int. Ed.*, 1985, **24**, 639-653.

[29] L. Bonnat, J. Dejeu, H. Bonnet, B. Génnaro, O. Jarjayes, F. Thomas, T. Lavergne, E. Defrancq, *Chem. Eur. J.*, 2016, **22**, 3139-3147.

[30] R. Bonnet, T. Lavergne, B. Gennaro, N. Spinelli, E. Defrancq, *Chem. Commun.*, 2015, **51**, 4850-4853.

[31] S. Amrane, A. Kerkour, A. Bedrat, B. Vialet, M. L. Andreola, J. L. Mergny, *J. Am. Chem. Soc.*, 2014, **136**, 5249-5252.

[32] L. Bonnat, L. Bar, B. Gennaro, H. Bonnet, O. Jarjayes, F. Thomas, J. Dejeu, E. Defrancq, T. Lavergne, *Chem. Eur. J.*, 2017, **23**, 5602-5613.

[33] A. Devaux, L. Bonnat, T. Lavergne, E. Defrancq, *Org. Biomol. Chem.*, 2020, **18**, 6394-6406.

[34] J. Carvalho, J.-L. Mergny, G. F. Salgado, J. A. Queiroz, C. Cruz, *Trends Mol. Med.*, 2020, **26**, 848-861.

[35] D. V. Andreeva, A. S. Tikhomirov, A. E. Shchekotikhin, *Russ. Chem. Rev.*, 2021, **90**, 1-38.

[36] H. A. Day, P. Pavlou, Z. A. E. Waller, *Bioorg. Med. Chem.*, 2014, **22**, 4407-4418.

[37] S. S. Masoud, K. Nagasawa, *Chem. Pharm. Bull.*, 2018, **66**, 1091-1103.

[38] D. Monchaud, M.-P. Teulade-Fichou, *Org. Biomol. Chem.*, 2008, **6**, 627-636.

[39] S. Neidle, *Nat. Rev. Chem.*, 2017, **1**, article no. 41.

[40] R. T. Wheelhouse, D. Sun, H. Han, F. X. Han, L. H. Hurley, *J. Am. Chem. Soc.*, 1998, **120**, 3261-3262.

[41] C. Hounsou, L. Guittat, D. Monchaud, M. Jourdan, N. Saettel, J.-L. Mergny, M.-P. Teulade-Fichou, *ChemMedChem*, 2007, **2**, 655-666.

[42] L. Martino, A. Virno, B. Pagano, A. Virgilio, S. Di Micco, A. Galeone, C. Giancola, G. Bifulco, L. Mayol, A. Randazzo, *J. Am. Chem. Soc.*, 2007, **129**, 16048-16056.

[43] J. A. Schouten, S. Ladame, S. J. Mason, M. A. Cooper, S. Balasubramanian, *J. Am. Chem. Soc.*, 2003, **125**, 5594-5595.

[44] Q. Chen, I. D. Kuntz, R. H. Shafer, *Proc. Natl. Acad. Sci. USA*, 1996, **93**, 2635-2639.

[45] P. Murat, R. Bonnet, A. Van der Heyden, N. Spinelli, P. Labbé, D. Monchaud, M.-P. Teulade-Fichou, P. Dumy, E. Defrancq, *Chem. Eur. J.*, 2010, **16**, 6106-6114.

[46] A. De Cian, E. DeLemos, J.-L. Mergny, M.-P. Teulade-Fichou, D. Monchaud, *J. Am. Chem. Soc.*, 2007, **129**, 1856-1857.

[47] C. Granotier, G. Pennarun, L. Riou, F. Hoffschir, L. R. Gauthier, A. De Cian, D. Gomez, E. Mandine, J.-F. Riou, J.-L. Mergny, P. Mailliet, B. Dutrillaux, F. D. Boussin, *Nucleic Acids Res.*, 2005, **33**, 4182-4190.

[48] N. H. Campbell, G. N. Parkinson, A. P. Reszka, S. Neidle, *J. Am. Chem. Soc.*, 2008, **130**, 6722-6724.

[49] J. Ren, J. B. Chaires, *Biochemistry*, 1999, **38**, 16067-16075.

[50] M. Perenon, H. Bonnet, T. Lavergne, J. Dejeu, E. Defrancq, *Phys. Chem. Chem. Phys.*, 2020, **22**, 4158-4164.

[51] J. Jiang, T. Teunens, J. Tisaun, L. Denuit, C. Moucheron, *Molecules*, 2022, **27**, article no. 1541.

[52] J. Dejeu, T. Lavergne, J. Della Nora, E. Defrancq, G. Pratviel, *Inorg. Chim. Acta*, 2016, **452**, 98-103.

[53] L. Sabater, M.-L. Nicolau-Travers, A. De Rache, E. Prado, J. Dejeu, O. Bombarde, J. Lacroix, P. Calsou, E. Defrancq, J.-L. Mergny, D. Gomez, G. Pratviel, *J. Biol. Inorg. Chem.*, 2015, **20**, 729-738.

[54] L. Sabater, P.-J. Fang, C.-F. Chang, A. De Rache, E. Prado, J. Dejeu, A. Garofalo, J.-H. Lin, J.-L. Mergny, E. Defrancq, G. Pratviel, *Dalton Trans.*, 2015, **44**, 3701-3707.

[55] J. E. Reed, A. Arola-Arnal, S. Neidle, R. Vilar, *J. Am. Chem. Soc.*, 2006, **128**, 5992-5993.

[56] L. Lecarme, E. Prado, A. De Rache, M.-L. Nicolau-Travers, R. Bonnet, A. van Der Heyden, C. Philouze, D. Gomez, J.-L. Mergny, H. Jamet, E. Defrancq, O. Jarjayes, F. Thomas, *Inorg. Chem.*, 2014, **53**, 12519-12531.

[57] L. Lecarme, E. Prado, A. De Rache, M.-L. Nicolau-Travers, G. Gellon, J. Dejeu, T. Lavergne, H. Jamet, D. Gomez, J.-L. Mergny, E. Defrancq, O. Jarjayes, F. Thomas, *ChemMedChem*, 2016, **11**, 1133-1136.

[58] N. H. Campbell, N. H. Abd Karim, G. N. Parkinson, M. Gunaratnam, V. Petrucci, A. K. Todd, R. Vilar, S. Neidle, *J. Med. Chem.*, 2012, **55**, 209-222.

[59] L. Xu, X. Chen, J. Wu, J. Wang, L. Ji, H. Chao, *Chem. Eur. J.*, 2015, **21**, 4008-4020.

[60] K. J. Castor, K. L. Metera, U. M. Tefashe, C. J. Serpell, J. Mauzeroll, H. F. Sleiman, *Inorg. Chem.*, 2015, **54**, 6958-6967.

[61] S. A. Archer, A. Raza, F. Drçge, C. Robertson, A. J. Auty, D. Chekulaev, J. A. Weinstein, T. Keane, A. J. H. M. Meijer, J. W. Haycock, S. MacNeil, J. A. Thomas, *Chem. Sci.*, 2019, **10**, 3502-3513.

[62] G. Piraux, L. Bar, M. Abraham, T. Lavergne, H. Jamet, J. Dejeu, L. Marcelis, E. Defrancq, B. Elias, *Chem. Eur. J.*, 2017, **23**, 11872-11880.

[63] M. Gillard, G. Piraux, M. Daenen, M. Abraham, L. Troian-Gautier, L. Bar, H. Bonnet, F. Loiseau, H. Jamet, J. Dejeu, E. Defrancq, B. Elias, *Chem. Eur. J.*, 2022, **28**, article no. e202202251.

[64] J. Weynand, A. Dinan, M. Abraham, L. Marcelis, H. Jamet, A. Decottignies, J. Dejeu, E. Defrancq, B. Elias, *Chem. Eur. J.*, 2018, **24**, 19216-19227.

[65] J. Weynand, H. Bonnet, F. Loiseau, J.-L. Ravanat, J. Dejeu, E. Defrancq, B. Elias, *Chem. Eur. J.*, 2019, **25**, 12730-12739.

[66] B. J. Pages, S. P. Gurung, K. McQuaid, J. P. Hall, C. J. Cardin, J. A. Brazier, *Front. Chem.*, 2019, **7**, article no. 744.

[67] M. Jourdan, J. Garcia, J. Lhomme, M.-P. Teulade-Fichou, J.-P. Vigneron, J.-M. Lehn, *Biochemistry*, 1999, **38**, 14205-14213.

[68] L. Xu, S. Hong, N. Sun, K. Wang, L. Zhou, L. Jia, R. Pei, *Chem. Commun.*, 2016, **52**, 179-182.

[69] A. Pagano, N. Iaccarino, M. A. S. Abdelhamid, D. Brancaccio, E. U. Garzarella, A. Di Porzio, E. Novellino, Z. A. E. Waller, B. Pagano, J. Amato, A. Randazzo, *Front. Chem.*, 2018, **6**, article no. 281.

[70] E. P. Wright, H. A. Day, A. M. Ibrahim, J. Kumar, L. J. E. Boswell, C. Huguin, C. E. M. Stevenson, K. Pors, Z. A. E. Waller, *Sci. Rep.*, 2016, **6**, article no. 39456.

[71] H. Bonnet, M. Morel, A. Devaux, J. H. Boissieras, A. Granzhan, B. Elias, T. Lavergne, J. Dejeu, E. Defrancq, *Chem. Commun.*, 2022, **58**, 5116-5119.

[72] F. Berthiol, J. Boissieras, H. Bonnet, M. Pierrot, C. Philouze, J.-F. Poisson, A. Granzhan, J. Dejeu, E. Defrancq, *Molecules*, 2023, **28**, article no. 682.

[73] M. Vlasenok, O. Levchenko, D. Basmanov, D. Klinov, A. Varizhuk, G. Pozmogova, *Data Brief*, 2018, **18**, 348-359.

[74] V. Brazda, L. Haronikova, J. C. Liao, M. Fojta, *Int. J. Mol. Sci.*, 2014, **15**, 17493-17517.

[75] A. Pipier, A. Devaux, T. Lavergne, A. Adrait, Y. Couté, S. Britton, P. Calsou, J. F. Riou, E. Defrancq, D. Gomez, *Sci. Rep.*, 2021, **11**, article no. 13469.

Comptes Rendus
Chimie
2023, Vol. 26, Special Issue S3, p. 53-66
https://doi.org/10.5802/crchim.251

Breaking Barriers in Chemical Biology – Toulouse 2022

1.4 nm gold nanoparticle-antibody conjugates for in situ gold immunolabelling after transduction into living human cells

Nadja Groysbeck [a], Anne Marie Haeberlé [b], Stéphane Ory [b], Victor Hanss [c], Mikhael Eltsov [c], Patrick Schultz [c] and Guy Zuber [*, a]

[a] Université de Strasbourg - CNRS, UMR7242 Biotechnologie et Signalisation Cellulaire, 300 Bd Sébastien Brant, CS 10413, 67412 Illkirch, France

[b] Centre National de la Recherche Scientifique, Université de Strasbourg, Institut des Neurosciences Cellulaires et Intégratives, F-67000 Strasbourg, France

[c] Centre for Integrative Biology (CBI), Department of Integrated Structural Biology, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), 1 rue Laurent Fries, BP10142, F-67404 Illkirch Cedex, France

E-mails: nadja.groysbeck@gmail.com (N. Groysbeck), haeberle@inci-cnrs.unistra.fr (A. M. Haeberlé), ory@inci-cnrs.unistra.fr (S. Ory), hanssv@igbmc.fr (V. Hanss), eltsovm@igbmc.fr (M. Eltsov), pat@igbmc.fr (P. Schultz), zuber@unistra.fr (G. Zuber)

Abstract. Despite advances in Electron Microscopy (EM) that enable to image protein assemblies within vitreous sections of cells at nearly atomic resolution, labelling is still necessary to locate small proteins or rare complexes. Gold immunolabelling has been used for decades to localise specific proteins within cellular sections. However, current gold particle-antibody conjugates are not built with enough chemical precision to match the current resolution offered by cryo-EM methodology. Furthermore, as a close to native specimen state can only be achieved by strict preservation of a frozen hydrated state, it is required to deliver gold labelling agents into living cells prior to their vitrification. Several 1.4 nm gold nanoparticle-antibody conjugates were synthesised. Their abilities to bind to and label their corresponding epitopes within living cells after cytosolic delivery by electroporation are documented here.

Keywords. Gold nanoparticle, Antibody, Conjugate, Intracellular delivery, Gold immunolabelling.

Funding. This research was supported by the ITI Innovec (IdEx (ANR-10-IDEX-0002), SFRI (ANR-20-SFRI-0012)), the French Infrastructure for Integrated Structural Biology (FRISBI ANR-10-INBS-05), and the ITMO Cancer (ColorME, 22P096-00). NG received a PhD fellowship from the IdEX Unistra (Université de Strasbourg and Investissements d'Avenir). VH received a PhD fellowship from the CNRS.

Manuscript received 27 January 2023, revised 27 June 2023, accepted 24 July 2023.

[*] Corresponding author

1. Introduction

Understanding how multiple components of a human cell interact together to sustain the basic mechanisms of life remains mostly unknown but makes fast progress through the combination of multiple experimental approaches. Encyclopedia containing the human genome [1] and tissue-based maps of the human proteome [2] enable us to define gene expression profiles and subsequently the protein levels that compose each human cell type. Advanced quantitative methods (quantitative Polymerase Chain Reaction, mass spectrometry) estimate the abundance of mRNA and proteins per cell [3]. Progresses in microscopy allow imaging protein–protein interactions at resolution approaching (macro)molecule level for super resolution fluorescent microscopy and nanometric resolution for cryo-Electron Microscopy (EM) [4]. At nanometric dimensions, specimen preservation is crucial and dehydration or chemical cross-linking should be avoided to prevent structural reorganisation of cellular content. In this respect, eukaryotic cells, including mammalian ones, can be vitrified either by plunge freezing or by high-pressure freezing, to preserve a close-to-native frozen hydrated state, and sectioned below −140 °C for cryo-EM imaging [5]. Abundant proteins forming filaments or large assemblies such as ribosomes or nuclear pores are easily recognised in cryo-electron micrographs due to their characteristic shape [4,6]. However, less abundant biomolecular complexes or small proteins cannot be recognised unambiguously in electron tomograms and their identification relies on labelling tools. Electron-dense gold nanoparticles (AuNPs) conjugated to antibodies are popular tags for EM immunolabelling since their high electron scattering, round shape, and specific dimension provide sufficient contrast to identify the probe [7]. Colloidal AuNPs with size larger than 5 nm in diameter are readily detected even when the cellular sample is contrasted. They offer a large surface area onto which many proteins adsorb [8]. This ability has been exploited to make immunocolloids of popular usage [7]. However, size restricts penetration of those colloids in gel-like biological samples [9]. It has led to the development of gold nanoparticles with diameters below 2 nm and to modification of the gold-to-antibody conjugation because proteins do not adsorb tightly enough on the surface of gold particles below 2 nm [10]. Phosphine-coated 1.4 nm Nanogold® [11] and 0.8 nm ultrasmall AuNPs [12] have increased labelling efficiency but to the cost of an extra step consisting in increasing the size of the gold particles by silver deposition to reveal their positions in the cell sections. Due to the impermeability of the plasma membrane, labelling of specific intracellular targets is generally performed post-embedding on cell sections at temperatures above 0 °C [13]. This protocol is not consistent with high-resolution cryo-EM because vitrified cellular samples cannot be warmed up after sectioning. Therefore, the gold nanoparticles-antibody conjugates need to be delivered into living cells, prior to the vitrification step, using weakly perturbing transduction methods [14]. Ultra-small 0.8 nm AuNPs conjugated to antibodies directed against RNA polymerase II were shown to be delivered into living HeLa cells using a cationic lipid formulation, and the gold labels were detected within the nucleus by Scanning Electron Microscopy (SEM) after resin embedding and silver enhancement [15]. To develop gold particles of sizes directly visible in vitrified cell sections by cryo-EM, mercaptobenzoic acid (MBA)-coated AuNPs of defined sizes around 2 nm [16,17] appeared extremely promising [18,19]. We previously attempted to substitute MBA coating on AuNPs by thionitrobenzoate (TNB) [20] to increase ligand exchange kinetics [21]. During the synthesis of 1.4 nm AuNP with sodium borohydride, a fraction of nitro groups underwent reduction to amines, yielding AuNP with a mixed coordination layer consisting of TABs and TNBs (TAB-, TNB-AuNP) named AuZ. This mixed coverage was seen advantageous, in terms of ligand exchange with incoming thiols on the one hand, and colloidal stability on the other hand. Indeed, ligand exchange onto the AuNPs is rarely quantitative and preferentially occurs with the negatively charged TNBs, leaving zwitterionic TABs on the surface as putative protecting groups against strong associations to proteins [22].

In this study, we linked several commercial monoclonal antibodies (mAbs) to 1.4 nm AuZ gold nanoparticle to produce AuNP-mAb conjugates. First, we verified that these conjugates specifically bind to their intracellular targets using fixed and permeabilised cell lines in a classical gold immunolabelling procedure. We confirmed that shielding the gold surface of the AuNP-mAb conjugate with

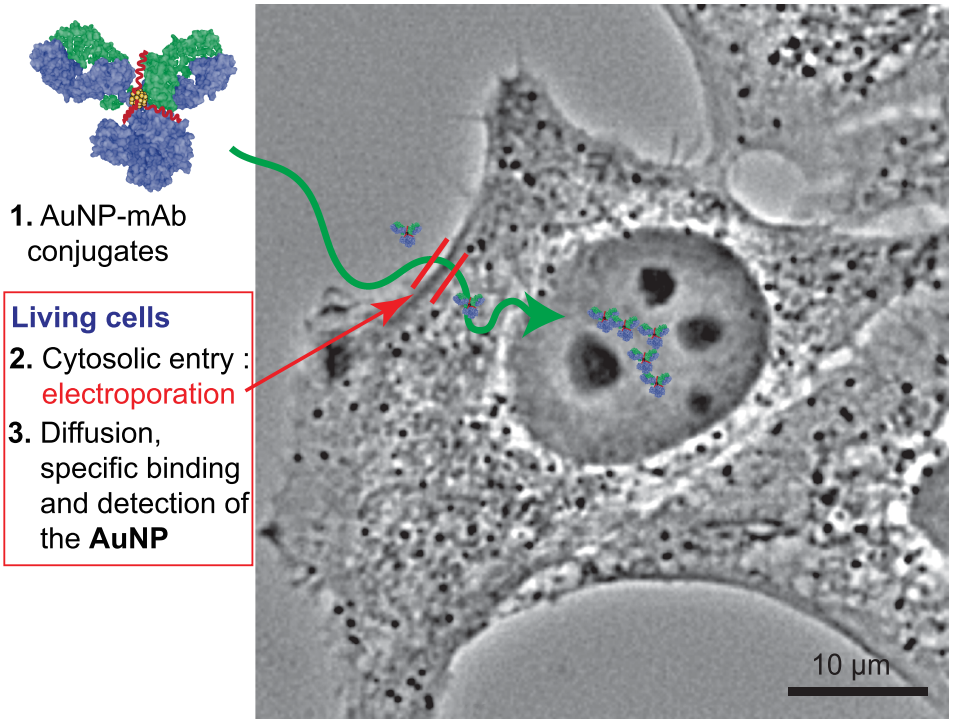


Figure 1. Illustration of gold immunolabelling in living cells. The strategy relies on the synthesis of specific AuNP-mAb conjugates and delivery procedure for the probes to diffuse, even into the cell nucleus, and bind to their specific targets.

2 kDa polyethyleneglycol drastically reduced unspecific adsorption of nanoparticles to extracellular constituents but also to intracellular ones [23,24]. The AuNP-mAb conjugates were then delivered inside living cells by electroporation [25,26] and the cells containing cytosolic AuNP-mAb conjugates were cultivated before monitoring the cell distribution of the probes (Figure 1). Our data demonstrate that the 1.4 nm AuNP-mAbs enter the cytoplasm where they circulate and associate to their targets. Accumulation of the probes in living cells was mAb-dependent. The accumulation patterns of the AuNP-mAb conjugates anti-GFP, anti-RPB1 (7G5, anti-RNA polymerase II) and anti-nucleoporins were analogous to the classical immunogold labelling ones indicating that the probes bind in cellulo to their targets without perturbing their function. Cytoplasmic-delivered anti-α-tubulin and anti-β-tubulin conjugates in contrast promoted intracytoplasmic aggregates, indicating that real-time binding of these conjugates perturbs microtubule dynamics.

2. Results and discussion

2.1. *Synthesis of gold nanoparticle-antibody conjugates*

The 1.4 nm AuZ used in this study has been previously described and comprehensively characterised [20]. To conjugate the mouse monoclonal antibodies to AuZ, we took advantage of the disulfide bonds present in the hinge region. These disulfide bonds are easily reduced to sulfhydryl groups using Tris(2-carboxyethyl)phosphine (TCEP) [28], thus creating a conjugation site for the 150 kDa mAb onto TAB-, TNB-AuNP at about 7 nm of the recognised antigen paratope (Figure 2) [29]. The various reaction intermediates can be monitored using a non-reducing SDS-PAGE experiment as illustrated by the conjugation of AuZ to the 7G5 mAb (Figure 2B). The 7G5 mAb targets the 52 times-repeated YSPTSPS sequence at the C-terminal domain of the largest subunit (RPB1) of mammalian RNA polymerase II (Pol II). TCEP at a concentration of 1 mM fully cleaved the disulfide bond holding the heavy

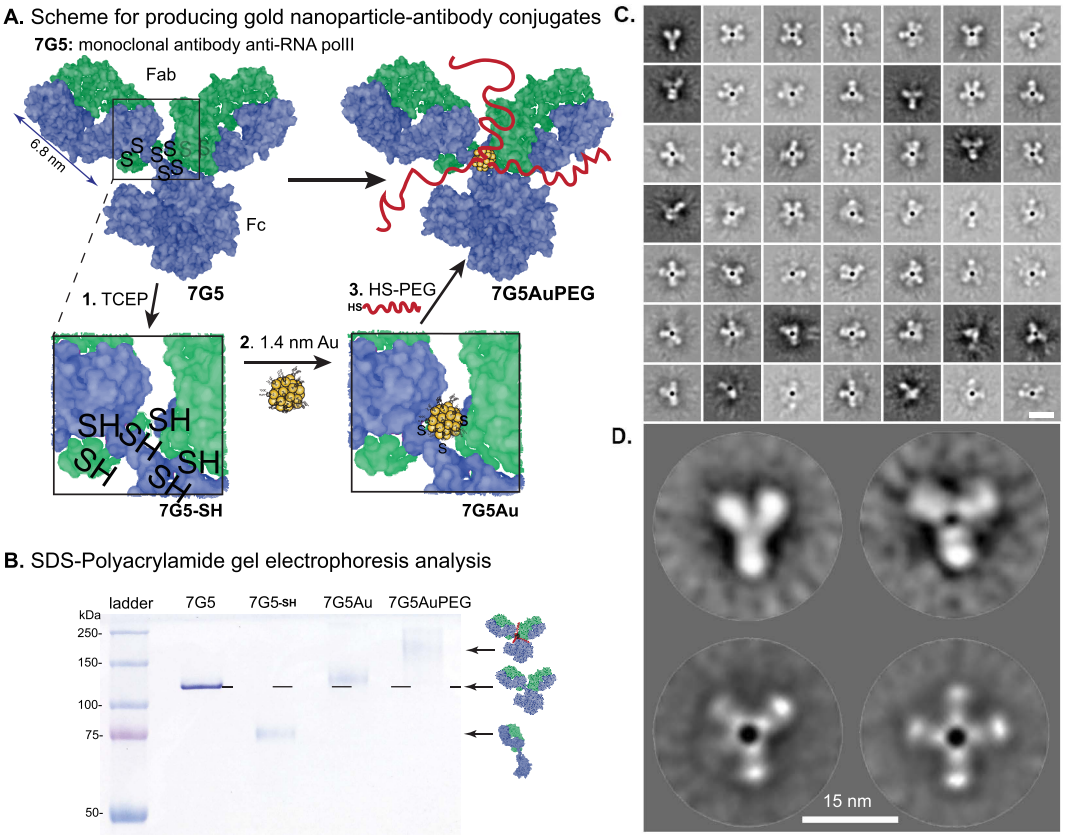


Figure 2. Synthesis and characterisation of the AuNP-antibody conjugate. (A) Disulfide bonds of the mouse monoclonal antibody targeting RPB1 of RNA polII (clone 7G5) are reduced at the hinge area with TCEP. The sulfhydryl groups at the hinge can then react with AuZ by thiolate exchange. After clearing the conjugate from excess AuZ, the remaining reacting thiolates on the AuNP surface are exchanged with thiolated 2 kDa polyethyleneglycol (PEG). (B) Non-reducing SDS-PAGE analysis of the reaction intermediates. Illustration of the antibody derived from protein data bank entry id 1IGY [27]. (C) Gallery of two-dimensional class averages from negatively stained 7G5AuPEG particles. (D) Representative class averages representing 2-D views of unlabelled (top) or gold-conjugated (bottom) IgG molecules. The bar represents 15 nm in C and D.

chains together as seen in the SDS-PAGE analysis (lane 7G5-SH). The mAb with an initial apparent molecular weight (MW) of 150 kDa was completely transformed into a 75 kDa protein corresponding to a single copy of the heavy and light chains. The reduced mAb was then reacted with a 1.5 molar excess of AuZ to yield a new component with an apparent MW slightly above 150 kDa (lane 7G5Au). Excess AuZ and the first released TNBs were removed by size exclusion chromatography. The purified 7G5Au was then reacted with an excess of thiolated 2 kDa polyethyleneglycol (PEG) to chemically neutralise the remaining reacting TNBs present on the gold particle surface [30] and to shield the gold surface from unspecific binding [8,23]. A material with a highly retarded electrophoretic mobility was observed (lane 7G5AuPEG), indicative of the shielding ability of PEG.

To assess the molecular organisation of the probe, the purified antibody-gold particle was visualised by transmission EM after negative staining with 2% uranyl acetate. Single particle analysis was performed to observe the position of the AuNP relative to the antibody, to determine the proportion

of Au-bound IgG molecules and to assess the integrity of the AuNP-bound antibody. The negatively stained protein moieties create a stain exclusion volume which diffuses less electrons than the surrounding medium. This property was used to select 80431 molecular images from 1125 micrographs that were aligned and clustered into 100 classes using Xmipp [31] (Figure 2C). After rejection of false positives, the remaining 36671 images were clustered into 100 classes. About 88% of the uranyl acetate-excluding materials were found associated with an electron-dense particle whose size is consistent with the input AuNP. The other images were found in classes representing characteristic views of non-conjugated IgG molecules such as 15 nm long, Y-shaped molecules (Figure 2D, top). The classes showing a dark electron-dense gold nanoparticle were heterogeneous and showed several extensions about 5 nm in size which could each correspond to Fab or Fc fragments. The 3-branched classes resemble the Y-shaped IgG molecule while the other classes are likely to represent two IgG molecules bound to the same AuNP particle with different rotational registers (Figure 2D, bottom). Strikingly, the electron-dense gold particle was always located at the centre of the multibranched structures, indicating that the AuNP associates with the hinge region of the IgG which exposes several thiolates following mild reduction step [32]. When the AuNP is bound to a single IgG molecule, the latter adopts a classical Y shape, suggesting that conjugation does not affect the molecular organisation of the antibody.

Binding of AuNP-mAb conjugates 7G5Au and 7G5AuPEG to RPB1 was first assayed using a classical immunocytochemistry procedure. HeLa cells cultivated on glass coverslips were fixed with 4% EM-grade paraformaldehyde for 20 min and their plasma membrane was permeabilised with 0.1% Triton X-100. The different 7G5 mAb conjugates were incubated with the permeabilised cells, excess probe was washed away and bound conjugates were detected with a secondary fluorescent goat anti-mouse antibody (Figure 3A). The 7G5 mAbs were found predominantly in the cell nucleus with reduced staining in the nucleolus, as expected for the nuclear RNA pol II enzyme. Enrichment of fluorescence was observed within foci in accordance with the hypothesis of transcription occurring mainly in factories or condensates containing several RNA pol II molecules [33]. A similar pattern was observed with the AuNP-mAb 7G5AuPEG component (Figure 3B). The intensity of the green fluorescence was nonetheless diminished as compared to the 7G5 mAbs signal. This feature might originate from the steric hindrance of PEGs [23] as well as from the fluorescence-quenching properties of AuNPs [34] leading to reduced association of the secondary Abs and dimming of the fluorescent signal, respectively. To map the distribution of gold nanoparticles, a similar immunolabelling experiment was performed using 7G5Au (Figure 3C) and 7G5AuPEG (Figure 3D), and a silver-enhancement procedure to reveal the AuNPs only (Supporting information, Figure S1 shows that the control HeLa cells remain unstained after incubation with the silver-enhancement solution). The 7G5Au probe showed a nuclear enrichment, but significant amounts of silver also stained the cytoplasm. Passivation of AuNPs with PEG 2 kDa diminished unspecific association to the cytoskeleton and provided a sharp nuclear gold immunolabelling in which dark spots are revealed. These dense spots might correspond to clustering of several RNA pol II molecules. The stealth property of PEGs surrounding AuZ was highly efficient since passivation of 7G5Au with the gold particle-stabilising CALNNG peptide [35] produced strong unspecific labelling of the cytoskeleton (Supporting information, Figure S2).

Commercial monoclonal antibodies (listed in Table 1) targeting RPB1 of RNA polII (7C2) [36], nucleoporins (Mab414), the trimethyl lysine 20 of histone H4 (6F8-D9), α-tubulin (B-5-1-2), β-tubulin (TU27) and the green fluorescent protein (GFP)(2A3) were first checked for their immunofluorescence specificity (Supporting information, Figure S3). They were then conjugated to AuZ and passivised with PEG using the same protocol as for 7G5 (Supporting information, Figure S4). Gold immunolabelling experiments on adherent cell lines led to images with precise localisation of targets using bright-field microscopy. The anti-nucleoporins conjugated to pegylated AuNPs (MAb414AuPEG) labelled the Nuclear Pore Complexes embedded within the nuclear envelope (Figure 4A). The post-translational modification of histone H4 at lysine 20 was also detected in the nucleus (Figure 4B). Microtubules were also clearly revealed in the cytoplasm using the anti-α-tubulin B-5-1-2AuPEG (Figure 4C) and the anti-β-tubulin TU27AuPEG (Figure 4D). The monoclonal anti-RPB1

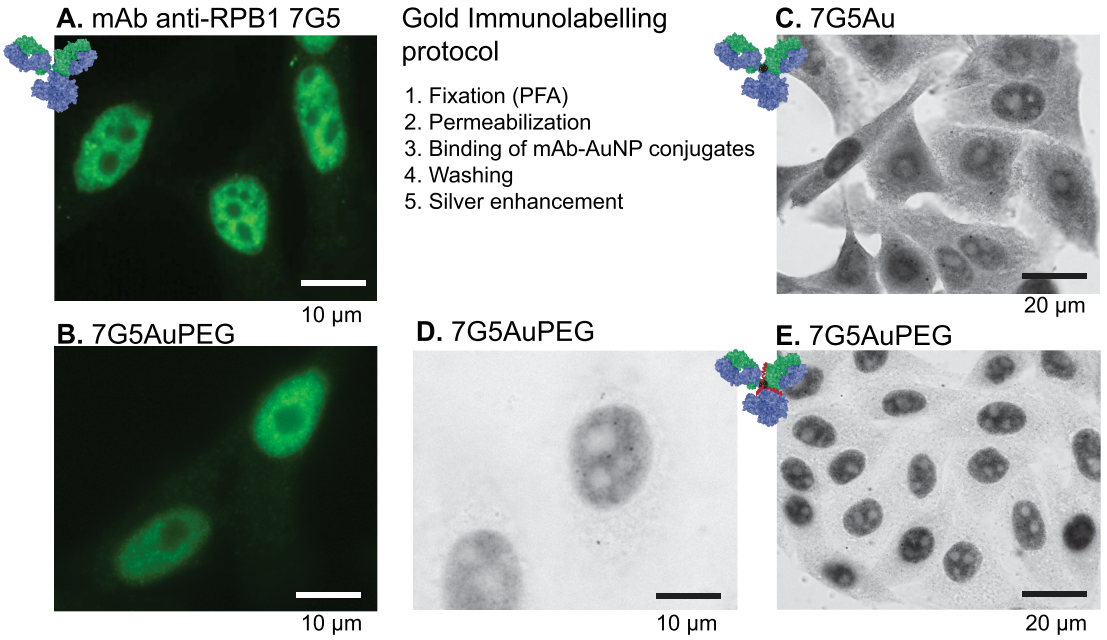


Figure 3. Immunofluorescence and gold immunolabelling analysis of the indicated monoclonal antibody anti-RPB1 (clone 7G5) derivatives on fixed and permeabilised HeLa cells. Images A and B: the mouse antibody was fluorescently detected with an AlexaFluo488-IgG anti-mouse conjugate. Images C, D and E: the gold particles were amplified with silver for detection by bright-field imaging. Conditions: fixation with 4% (v/v) PFA in 0.1 M phosphate buffer pH 7.6 for 20 min, permeabilisation with 0.1% (v/v) Triton X-100®, incubations with primary and secondary antibodies at 13 nM.

Table 1. List of antibodies used for conjugation

Name	Target	Source
7G5	RPB1 of RNA polII $[TSPSYSP]_3$ at C ter	IGBMC
7C2	RPB1 of RNA polII $[YSPTSPS]_3$ at C ter	IGBMC
Mab414	Nucleoporins (Nuclear pores)	Biolegend
6F8-D9	Histone H4 trimethyl Lysine20 (H4K20me3)	Biolegend
B-5-1-2	α-tubulin (C-t)	ThermoFisher
TU27	β-tubulin (C-t)	Biolegend
2A3	Green fluorescent protein (GFP)	IGBMC

7C2AuPEG was not as effective as 7G5AuPEG in terms of labelling sharpness (Figure 5A). Finally, the anti-GFP 2A3AuPEG [12,37] probe labelled extremely well the GFP-labelled histone H2B within the nucleus of stably transformed HeLa cells that express the GFP-H2B fusion (Figure 5C) but did not bind to wild-type HeLa devoid of GFP (Figure 5D). The importance of passivation by grafting PEG onto the gold surface to limit unspecific binding to cytoplasmic components was also confirmed for that antibody (Supporting information Figure S5).

Altogether, those first experiments demonstrated that 1.4 nm AuZ can be conjugated to several mAbs at the hinge region to produce effective gold labelling probes for classical immunocytochemistry procedures. Shielding the 1.4 nm gold nanoparticle with 2 kDa PEG was beneficial and considerably limited unspecific binding to cell components.

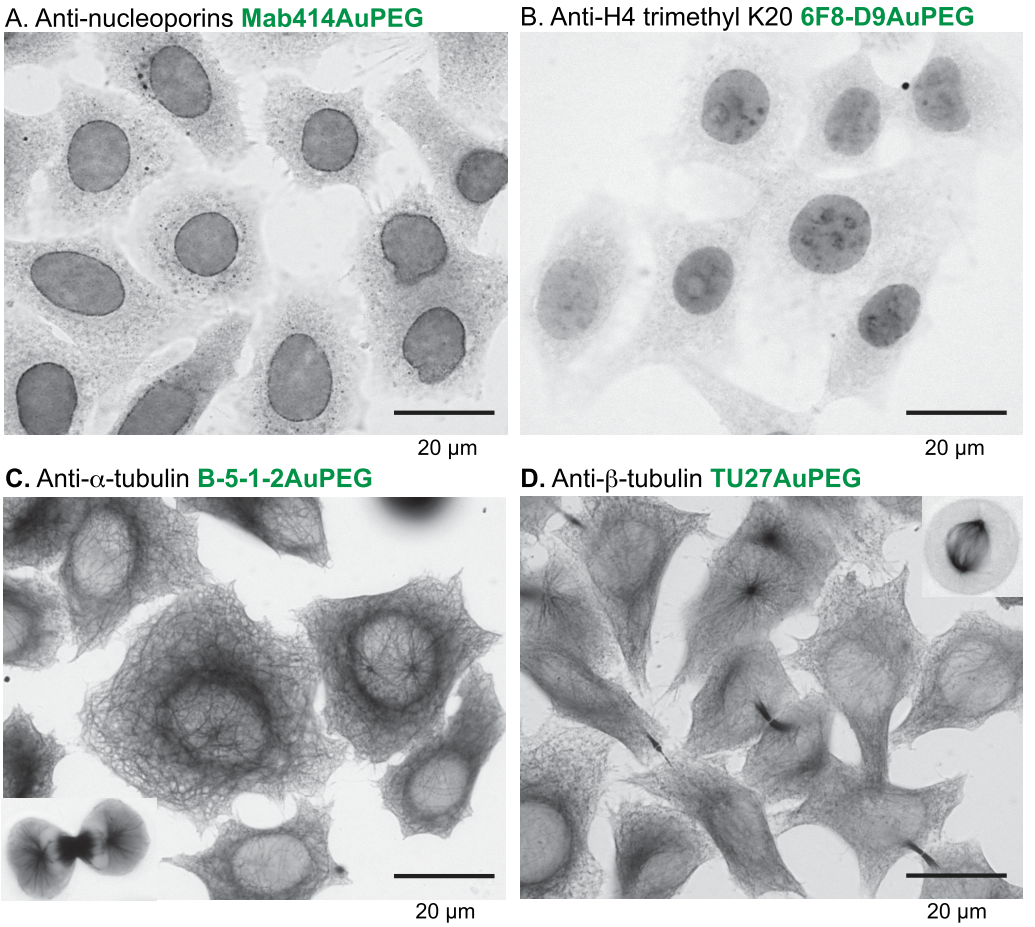


Figure 4. Gold immunolabelling analysis of the indicated monoclonal antibodies (mAbs) conjugated to 1.4 nm AuNP passivised with 2 kDa PEG. H2B-GFP HeLa cell line stably expresses the histone H2B fused to GFP. The gold particles were enhanced by silver for bright-field imaging.

2.2. *Fate of AuNP-mAb conjugates following cytosolic delivery into living HeLa cells*

The ability of the AuNP-mAb conjugates to diffuse inside living cells, bind to their targets and label them was assayed (Figure 6). To introduce the probes into the cytoplasm of living cells, transient holes in the plasma membrane of freshly trypsinised cells were created using an electroporation device and optimised pulses [26,38]. After the electric pulses and re-closure of the plasma membrane, HeLa cells containing cytoplasmic mAbs and AuNP-mAb conjugates were generally cultivated for 20 h to let them adhere onto glass coverslips. Cells were then processed for detection of antibodies and of AuNPs (Figure 6A). After delivery into the cytoplasm of HeLa cells expressing H2B-GFP, the unconjugated anti-GFP 2A3 accumulated into the cell nuclei and distributed similarly to the GFP, suggesting an effective binding of the antibody to its target (Figure 6B). Entry of 150 kDa Abs into the cell nucleus was unexpected since nuclear pore complexes usually prevent nuclear entry of proteins with a molecular weight above 60 kDa unless they are equipped with Nuclear Localisation Signals (NLS). The anti-GFP antibody is likely transported into the nucleus after binding to the NLS-containing GFP-H2B protein when it is synthetised in the cytoplasm by a piggyback mechanism [39]. When delivered into the cytoplasm of wild-type HeLa cells devoid of GFP-H2B, the

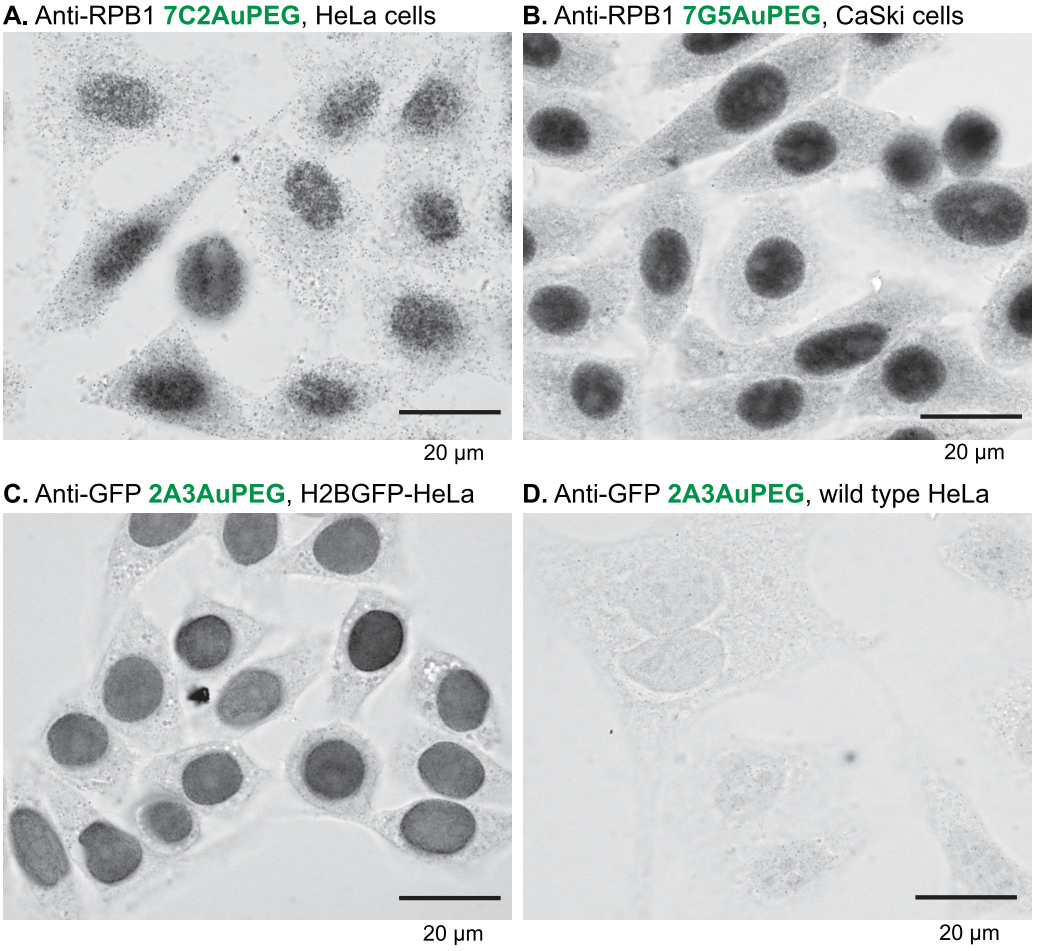


Figure 5. Gold immunolabelling analysis of the indicated monoclonal antibodies (mAbs) conjugated to 1.4 nm AuNP passivised with 2 kDa PEG. H2B-GFP HeLa cell line stably expresses the histone H2B fused to GFP. The gold particles were enhanced by silver for bright-field imaging.

anti-GFP 2A3AuPEG probe diffused in the cytoplasm but was excluded from the nucleus (Figure 6C). In contrast, the anti-GFP 2A3AuPEG delivered into the cytoplasm of HeLa cells expressing H2B-GFP accumulated in the nuclei in a fashion resembling 2A3 (Figure 6D), indicating that AuNP conjugation does not impair nuclear import of this specific mAb anti-GFP. The observed nuclear accumulation patterns relied on the presence of conjugates in the cytoplasm of living cells since electric pulses for electroporation were required (left-hand panels in Figure 6C and D). In cellulo labelling with anti-nucleoporins Mab414 and Mab414AuPEG yielded images (Figure 6E) which were highly reminiscent of the classical immunocytochemistry images (Figure 4A). Cytoplasmic-delivered anti-α-tubulin TU27AuPEG and anti-β-tubulin B-5-1-2AuPEG did not reveal an extensive microtubule network but rather accumulated as aggregates. This type of labelling and accumulation might be related to the dynamic nature of the microtubule network and to the electroporation procedure that involves a transient suspension of the cells using a trypsin treatment. Reshuffling of the microtubule network might take place during re-adhesion of the cells and after the electric pulses which are also known disruptors of microtubules [40,41]. The cytoplasm-delivered 7G5AuPEG and 7C2AuPEG also accumulated in the nucleus (Figure 6H) similarly to

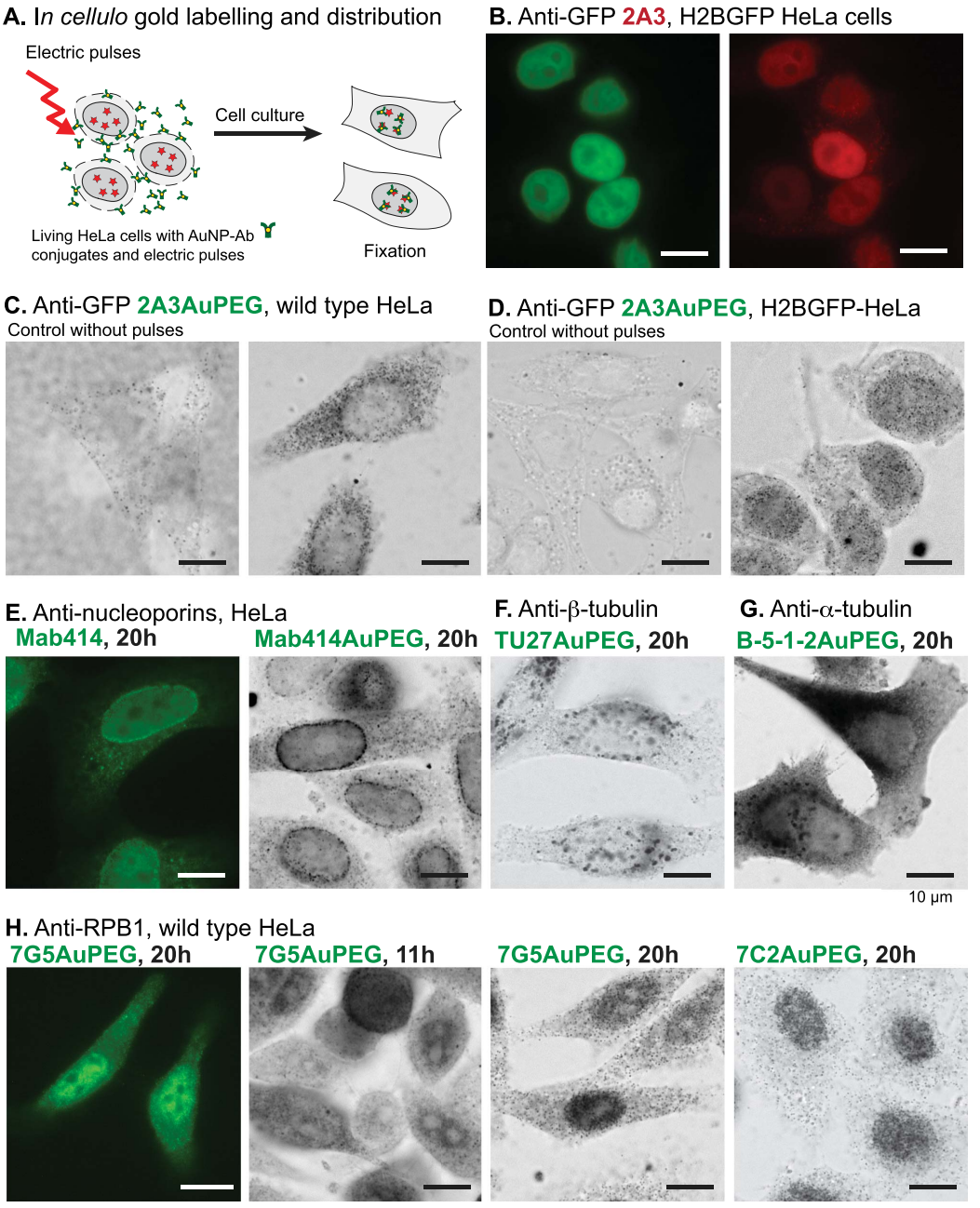


Figure 6. The AuNP-mAb conjugates (0.2 $\mu g \cdot \mu L^{-1}$) were delivered in HeLa cells by transient permeabilisation of the cell membrane using 3 electric pulses at 517 $V \cdot cm^{-1}$. Distribution of the mAbs and conjugates inside living cells was evaluated after culture during 20 h unless stated. Fixation: 4% PFA, 20 min. Permeabilisation: 0.1% Triton X-100. AuNPs were silver-enhanced. (A) Scheme of the procedure. (B) Images of the targeted H2B-GFP in HeLa cells (left) and of the delivered anti-GFP 2A3 (right). The mAb was detected using AlexaFluor568 goat anti-mouse (in red). (C, D) Detection of AuNPs by silver enhancement in cells incubated with 2A3AuPEG without application of electric pulses (left images) and with application of electric pulses. (E–H) Delivered mAbs were detected in cells using the antibody or the gold domain.

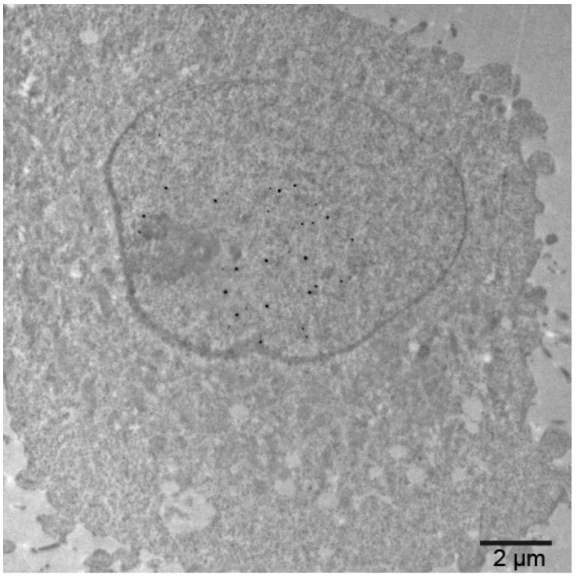


Figure 7. Electron micrograph of a HeLa cell section containing 7G5AuPEGs in the nucleus. Cells were incubated with 7G5AuPEG (0.2 $\mu g \cdot \mu L^{-1}$) and the plasma membrane was transiently permeabilised with 3 electric pulses at 517 $V \cdot cm^{-1}$. Cells were cultivated for 20 h, fixed with 2% glutaraldehyde and processed for classical TEM observation. Gold particles were enhanced with silver.

the unconjugated mAbs [38]. Similarly to the classical gold immunolabelling protocol, passivation of the AuNPs with PEG was also extremely useful for nuclear accumulation. Indeed, capping the AuNP of 7G5Au with CALNNG peptide did not lead to nuclear accumulation of 7G5AuC 24 h following electroporation (Figure S6). Finally, the nuclear accumulation appears not as with the anti-GFP conjugates. This property is likely related to the production of RNA pol II in the cytoplasm which is lower than that of H2B-GFP, limiting *de facto* the entry of Ab-bound RNA pol II [37,38]. The 7G5AuPEG labelling was further investigated by transmission electron microscopy (TEM) (Figure 7). Twenty hours after electroporation, cells were fixed, the gold signal was amplified by silver enhancement for better visualisation and thin sections of resin-embedded cells were observed by TEM. An image of a cell section displaying the nucleus and cytoplasm was recorded. The enhanced AuNPs were almost exclusively found in the cell nucleus, confirming that the RNA polII-driven accumulation of the corresponding AuNP-mAb occurred in this cell.

3. Conclusion

We hereby provide a synthesis protocol for preparing mAbs conjugated to 1.4 nm AuZ gold nanoparticles by taking advantage of the disulfide bonds which connect the hinge region of the antibody heavy chains. The reaction process can be easily monitored by non-reducing SDS-PAGE analysis which reveals the key reaction intermediates. Passivation of the gold surface with 2 kDa polyethyleneglycol provided probes with excellent gold immunolabelling properties such as solubility, absence of in vivo aggregation, good diffusion properties, and dramatically diminished the unspecific binding to cellular components. Moreover, these conjugates can be delivered into the cytoplasm of living cells by electroporation, where they diffuse and bind to their targets. The piggyback mechanism allowing nuclear entry of AuNP-mAb only upon binding to NLS-containing targets can be extremely useful as filter since unbound AuNP-mAbs remain in the cytoplasm. Our first proof of concept data demonstrated that AuNP-mAbs can be delivered into living cells, bind to their specific targets and accumulate at locations where their targets go. In principle, the anti-GFP antibody has the potential to become a rather generic gold immunolabelling probe. Standard immunolabelling procedure and silver enhancement will enable to monitor the labelling dynamic and efficiency, to adjust the amounts of the electroporated probes and the incubation time (and maybe other parameters) to ensure that the most specific labelling is observed later in cryo-EM. Specific monoclonal antibodies bind to a specific protein domain and then interfere with the functioning of their targets. After labelling with AuNPs and delivery inside living cells, these interfering probes might help decipher the function of protein/protein interactions and dynamics of exposure to Ab binding. Although the electroporation procedure is convenient and has demonstrated efficiency, the electric pulses are not neutral to cell physiology. Cells mostly recover but the cell cytoskeleton is perturbed. Alternatively, differentially disruptive liposomal formulations have been used as delivery solutions and can be employed as well.

4. Experimental section

4.1. *Materials*

The 2 kDa alpha-methoxy-omega-mercapto poly (ethylene glycol) (PEG-SH) was ordered from Iris Biotech. EM-grade paraformaldehyde (20% w/v solution) and glutaraldehyde (25% w/v solution) were purchased from Electron Microscopy Sciences. Other reagents were obtained from Sigma Aldrich, Carl Roth, VWR Chemicals, Euromedex or Honeywell. They were used without further purification unless stated otherwise. All solutions and buffers were made with water purified with a Millipore Q-POD apparatus. Precision Plus Protein Standard Dual Xtra (BioRad) was used as protein ladder for SDS-PAGE analysis. Monoclonal antibodies were purchased from commercial sources and were devoid of bovine serum albumin or gelatine. They were cleared from sodium azide or any other preservatives by size exclusion chromatography (Sephadex G-25) and eventually reconcentrated by ultracentrifugation.

4.2. *Gold nanoparticle synthesis*

The 1.4 nm AuZ was synthesised and characterised in a previous publication [18–20]. The concentration of gold nanoparticles was determined by UV–vis spectrophotometry using the absorbance at 520 nm and an extinction coefficient ε_{520} of 2.7×10^5 $mol^{-1}{\cdot}L{\cdot}cm^{-1}$.

4.3. *Reduction of monoclonal antibody*

Before performing any conjugation, reduction of the monoclonal antibody at the hinge area was monitored by SDS-PAGE using a non-reducing loading buffer. The antibody (2 $\mu g{\cdot}\mu L^{-1}$ or 13.34 μM, 2 μL) was titrated with TCEP (final concentrations: 0.1, 0.25, 0.5, 1, and 2 mM). After 16 h incubation at 20 °C, the reaction mixtures were mixed with 30% (w/v) glycerol containing 0.1% (w/v) SDS and 0.05% (w/v) bromophenol blue (0.4 μL) and analysed by SDS-PAGE. The minimal TCEP concentration producing full disappearance of the band at about 150 kDa was selected for medium-scale synthesis.

4.4. *Synthesis of 7G5AuPEG*

The monoclonal antibody in PBS containing 1 mM EDTA (2 $\mu g{\cdot}\mu L^{-1}$ or 13.34 μM, 100 μL) was reduced by addition of 0.1 M TCEP, pH 7.0 (1 μL) to a final concentration of 1 μM. After a 12 h incubation at 20 °C, an aliquot was withdrawn for analysis and the mixture was then quickly added and mixed with Auz (25 μL of 80 μM solution). After overnight incubation at 20 °C, the crude mixture was purified by size exclusion chromatography (Biorad P100, 2 mL of resin, elution with PBS). Thiolated polyethyleneglycol 2000 (2.7 μL of a 20 mM aqueous concentration) was then added. The mixture was incubated at room temperature for 16 h and the crude mixture was purified by ultracentrifugation (7 cycles of 0.5 mL sterile PBS) using a 0.5 mL Microcon® centrifugal filter device with a 100 kDa cutoff (Merck, Molsheim). At the end of the process, the gold nanoparticle-antibody conjugate was recovered in 60 μL PBS and analysed by SDS-PAGE to estimate the concentration and conjugation efficiency. The other gold nanoparticles-antibody conjugates were similarly prepared.

4.5. *Electroporation protocol*

Transient permeabilisation of the plasma membrane of cells was conducted as previously described using the Neon® transfection system and three 10 ms pulses at 1550 V [20,26]. Typically, HeLa cells (2×10^5 cells in 10 μL of the Neon Buffer R) were mixed with each antibody (2.5 μL of 2 $\mu g{\cdot}\mu L^{-1}$ solution in PBS). After the electric pulses, cells were diluted in prewarmed antibiotic-free cell culture medium (1 mL). The cells were recovered by gentle centrifugation, suspended into the warmed antibiotic-free complete cell culture medium and plated in 24-wells plates for cell adherence and growth onto glass coverslips and left in a cell culture incubator, usually for 18 h.

4.6. *Cell culture*

Cells were cultured in a 37 °C humidified incubator supplied with 5% CO_2. HeLa, CaSki, and stably transformed H2B-GFP HeLa cells were maintained in Dulbecco's modified Eagle medium containing 2 mM L-glutamine, 10 mM HEPES buffer, pH 7.0, 10% heat-inactivated foetal bovine serum (FBS) and

50 μg/mL gentamycin. For immunocytochemistry, immunofluorescence and pre-embedding immuno-EM experiments, cells (25,000 cells/well) were seeded into 24-well plates in 0.5 mL cell culture medium in which 13 mm diameter glass coverslips were deposited. The cells were left to adhere overnight by culture in a 37 °C humidified incubator supplied with 5% CO_2.

4.7. *Immunocytochemistry for bright-field and fluorescence imaging*

Coverslip-adhered cells were fixed with 4% PFA in 100 mM Sorenson's buffer, pH 7.6 for 20 min at 20 °C. The coverslips were then washed with PBS (3 × 0.5 mL, 5 min) containing 50 mM glycine (0.5 mL, 20 min) and cell plasma membranes were permeabilised with 0.1% Triton X-100 in PBS (0.5 mL, 5 min). Coverslips were soaked in PBS containing 10% (w/v) BSA for 1 h, washed with 0.2% acetylated BSA (BSA-c) in PBS (2 × 0.5 mL, 5 min) and then incubated with 6 to 10 nM of the conjugates in 0.2% BSA-c containing 10% FCS, 0.5 mL for 1 h. Next, cells were washed with 0.2% BSA-c (2 × 0.5 mL, 5 min) and with 80 mM citrate buffer pH 6.7 containing 2% sucrose (3 × 0.5 mL, 5 min). Gold was revealed by silver enhancement according to a modified Danscher protocol [42] using silver acetate and propyl gallate [43]. For fluorescence imaging, the primary antibodies were labelled using secondary fluorescent anti-mouse conjugates (AlexaFluor568 goat anti-mouse, Invitrogen A11031 and AlexaFluor488 goat anti-mouse, Invitrogen A11029). Fluorescence observations were carried out using a Leica DM5500B microscope equipped with an HCX PL Apo 63 × 1.40 oil PH3CS objective and a Leica DFC350FX camera.

4.8. *Preparation of cells containing gold nanoparticle-antibody conjugates for bright-field microscopy imaging*

After the indicated time periods following application of the electric pulses, the adherent cells cultivated on glass coverslips were washed with PBS (2 × 1 mL). The cells were fixed with 4.0% paraformaldehyde (PFA) (1 mL) for 20 min or 2% glutaraldehyde in 100 mM Sorenson's buffer, pH 7.6 for 1 h. After removal of the fixative solution and 3 PBS washes (1 mL), the plasma membrane was permeabilised using 0.1% (w/v) saponin or Triton X-100 (1 mL, 15 min) in 0.1 M Sorenson's buffer pH 7.6. The phosphate-buffered solution was replaced by a 0.1 M citrate solution, pH 6.7 containing 2% sucrose (5 washes, 1 mL). Development of AuNPs [42] was then done in a dark room for 30 min using freshly prepared 6 mM silver acetate solution in 0.16 M sodium citrate, pH 6.7 containing 20% gum arabic and 2 mM propyl gallate [43]. The silver-promoted enhancement of AuNPs was stopped with 0.16 M sodium citrate solution, pH 6.7 (1 mL, several washes). The glass coverslips were finally mounted onto 3 × 1 inch microscope slides (knittelglass.com) using Fluoromount-G (Southern Biotech, Ref. 0100-01, batch I2819-WC79B). Observations were carried out using a Leica DM5500B microscope equipped with an HCX PL Apo 63 × 1.40 oil PH3CS objective and a Leica DFC350FX camera.

4.9. *Cellular specimen preparation for EM*

20 h after the electroporation treatment, the adherent cells were washed with warmed (37 °C) Sorenson's buffer (3 × 2 mL). The cells were then fixed with 2.0% glutaraldehyde (1 mL) for 24 h. They were then washed with TEM storage solution (5×, 1 mL). Enhancement of the gold particles with silver was performed using the SE-EM kit (Aurion, the Netherlands) according to the manufacturer's protocol. The cells were then contrasted with 0.5% osmium for 20 min followed by 2% uranyl acetate both diluted in water. The specimens were dehydrated in a solution containing increasing concentrations of ethanol and flat-embedded in Epon (Ladd Research Industries, Williston, Vermont, USA). The resin-embedded specimens were sectioned into 100 nm thick slices which were deposited on an electron microscope grid.

4.10. *TEM and image analysis*

Cellular specimens were imaged on a Hitachi H7500 transmission electron microscope (Hitachi High Technologies Corporation) equipped with an AMT Hamatsu digital camera (Hamatsu Photonics). For EM imaging of 7G5AuPEG we used nickel EM grids covered with a thick holey carbon film (Quantifoil R2/2). The grids were topped with a thin continuous

carbon layer and rendered hydrophilic by a 90s treatment in a Fischione 1070 plasma cleaner operating at 30% power with a gas mixture of 80% argon:20% oxygen. The 7G5AuPEG solution was diluted to a final concentration of 50 μg/mL in Buffer A (Tris 10 mM, pH 7.4, NaCl 50mM) and crosslinked for 30 s at a final glutaraldehyde concentration of 0.3%. Three μL of this solution were deposited on the glow-discharged EM grid and the particles were allowed to adsorb for 1 min. The grid was washed once on a drop of buffer A and negatively stained with a 2% uranyl acetate solution. The sample was then imaged in a Tecnai F20 (FEI) operating at 200 kV and equipped with a 2K CCD camera (Gatan). Micrograph montages were recorded at an image magnification of 50 kX with a 2.01 Å pixel size at −2 μm defocus using the SerialEM software [44]. For single particle analysis, micrographs were imported within the Scipion framework for image processing software [45] and Fourier band-passed filtered using the xmipp3 protocol [31]. Reference-free picking was performed using the eman2 protocol to generate a dataset of 80431 particles [46]. The particle set was aligned, masked, 2-D classified, cleaned to obtain a total of 36671 molecular images and classified again into 100 classes using xmipp3 protocols.

Declaration of interests

The authors do not work for, advise, own shares in, or receive funds from any organization that could benefit from this article, and have declared no affiliations other than their research organizations.

Supplementary data

Supporting information for this article is available on the journal's website under https://doi.org/10.5802/crchim.251 or from the author.

References

[1] The ENCODE Project Consortium, *Nature*, 2012, **488**, 57-74.

[2] M. Uhlen, L. Fagerberg, B. M. Hallstrom, C. Lindskog, P. Oksvold, A. Mardinoglu, A. Sivertsson, C. Kampf, E. Sjostedt, A. Asplund *et al.*, *Science*, 2015, **347**, article no. 1260419.

[3] B. Schwanhäusser, D. Busse, N. Li, G. Dittmar, J. Schuchhardt, J. Wolf, W. Chen, M. Selbach, *Nature*, 2011, **473**, 337-342.

[4] J. Mahamid, S. Pfeffer, M. Schaffer, E. Villa, R. Danev, L. K. Cuellar, F. Förster, A. A. Hyman, J. M. Plitzko, W. Baumeister, *Science*, 2016, **351**, 969-972.

[5] A. Al-Amoudi, J.-J. Chang, A. Leforestier, A. McDowall, L. M. Salamin, L. P. O. Norlén, K. Richter, N. S. Blanc, D. Studer, J. Dubochet, *EMBO J.*, 2004, **23**, 3583-3588.

[6] A. P. Schuller, M. Wojtynek, D. Mankus, M. Tatli, R. Kronenberg-Tenga, S. G. Regmi, P. V. Dip, A. K. R. Lytton-Jean, E. J. Brignole, M. Dasso *et al.*, *Nature*, 2021, **598**, 667-671.

[7] W. P. Faulk, G. M. Taylor, *Immunochemistry*, 1971, **8**, 1081-1083.

[8] M. P. Monopoli, C. Aberg, A. Salvati, K. A. Dawson, *Nat. Nanotechnol.*, 2012, **7**, 779-786.

[9] G. Mayer, M. Bendayan, *Progr. Histochem. Cytochem.*, 2001, **36**, 3-84.

[10] L. Boselli, E. Polo, V. Castagnola, K. A. Dawson, *Angew. Chem.*, 2017, **129**, 4279-4282.

[11] J. F. Hainfeld, *Science*, 1987, **236**, 450-453.

[12] J. M. Robinson, T. Takizawa, D. D. Vandre, *J. Histochem. Cytochem.*, 2000, **48**, 487-492.

[13] R. C. N. Melo, E. Morgan, R. Monahan-Earley, A. M. Dvorak, P. F. Weller, *Nat. Protoc.*, 2014, **9**, 2382-2394.

[14] M. Chiper, K. Niederreither, G. Zuber, *Adv. Healthc. Mater.*, 2018, **7**, article no. e1701040.

[15] I. Orlov, A. Schertel, G. Zuber, B. Klaholz, R. Drillien, E. Weiss, P. Schultz, D. Spehner, *Sci. Rep.*, 2015, **5**, article no. 8324.

[16] Y. Levi-Kalisman, P. D. Jadzinsky, N. Kalisman, H. Tsunoyama, T. Tsukuda, D. A. Bushnell, R. D. Kornberg, *J. Am. Chem. Soc.*, 2011, **133**, 2976-2982.

[17] M. Azubel, J. Koivisto, S. Malola, D. Bushnell, G. L. Hura, A. L. Koh, H. Tsunoyama, T. Tsukuda, M. Pettersson, H. Häkkinen *et al.*, *Science*, 2014, **345**, 909-912.

[18] M. Azubel, S. D. Carter, J. Weiszmann, J. Zhang, G. J. Jensen, Y. Li, R. D. Kornberg, *Elife*, 2019, **8**, article no. e43146.

[19] V. Postupalenko, D. Desplancq, I. Orlov, Y. Arntz, D. Spehner, Y. Mely, B. P. Klaholz, P. Schultz, E. Weiss, G. Zuber, *Angew. Chem. Int. Ed. Engl.*, 2015, **54**, 10583-10586.

[20] D. Desplancq, N. Groysbeck, M. Chiper, E. Weiss, B. Frisch, J.-M. Strub, S. Cianferani, S. Zafeiratos, E. Moeglin, X. Holy *et al.*, *ACS Appl. Nano Mater.*, 2018, **1**, 4236-4246.

[21] R. L. Donkers, Y. Song, R. W. Murray, *Langmuir*, 2004, **20**, 4703-4707.

[22] N. Zhan, G. Palui, M. Safi, X. Ji, H. Mattoussi, *J. Am. Chem. Soc.*, 2013, **135**, 13786-13795.

[23] J. Y. Wong, T. L. Kuhl, J. N. Israelachvili, N. Mullah, S. Zalipsky, *Science*, 1997, **275**, 820-822.

[24] N. Groysbeck, V. Hanss, M. Donzeau, J.-M. Strub, S. Cianferani, D. Spehner, M. Bahri, O. Ersen, M. Eltsov, P. Schultz *et al.*, *Small Methods*, 2023, **7**, article no. e2300098.

[25] H.-Y. Wang, C. Lu, *Biotechnol. Bioeng.*, 2008, **100**, 579-586.

[26] G. Freund, A.-P. Sibler, D. Desplancq, M. Oulad-Abdelghani, M. Vigneron, J. Gannon, M. H. Van Regenmortel, E. Weiss, *mAbs*, 2014, **5**, 518-522.

[27] L. J. Harris, E. Skaletsky, A. McPherson, *J. Mol. Biol.*, 1998, **275**, 861-872.

[28] A. Makaraviciute, C. D. Jackson, P. A. Millner, *J. Immunol. Methods*, 2016, **429**, 50-56.

[29] N. Groysbeck, A. Stoessel, M. Donzeau, E. C. da Silva, M. Lehmann, J.-M. Strub, S. Cianferani, K. Dembélé, G. Zuber, *Nanotechnology*, 2019, **30**, article no. 184005.

[30] R. Guo, Y. Song, G. Wang, R. W. Murray, *J. Am. Chem. Soc.*, 2005, **127**, 2752-2757.

[31] J. M. de la Rosa-Trevín, J. Otón, R. Marabini, A. Zaldívar, J. Vargas, J. M. Carazo, C. O. S. Sorzano, *J. Struct. Biol.*, 2013, **184**, 321-328.

[32] M. M. C. Sun, K. S. Beam, C. G. Cerveny, K. J. Hamblett, R. S. Blackmore, M. Y. Torgov, F. G. M. Handley, N. C. Ihle, P. D. Senter, S. C. Alley, *Bioconjug. Chem.*, 2005, **16**, 1282-1290.

[33] W.-K. Cho, N. Jayanth, B. P. English, T. Inoue, J. O. Andrews, W. Conway, J. B. Grimm, J.-H. Spille, L. D. Lavis, T. Lionnet *et al.*, *Elife*, 2016, **5**, article no. e13617.

[34] C. Xue, Y. Xue, L. Dai, A. Urbas, Q. Li, *Adv. Opt. Mater.*, 2013, **1**, 581-587.

[35] R. Lévy, N. T. K. Thanh, R. C. Doty, I. Hussain, R. J. Nichols, D. J. Schiffrin, M. Brust, D. G. Fernig, *J. Am. Chem. Soc.*, 2004, **126**, 10076-10084.

[36] S. Besse, M. Vigneron, E. Pichard, F. Puvion-Dutilleul, *Gene Expr.*, 1995, **4**, 143-161.

[37] N. Groysbeck, M. Donzeau, A. Stoessel, A.-M. Haeberle, S. Ory, D. Spehner, P. Schultz, O. Ersen, M. Bahri, D. Ihiawakrim *et al.*, *Nanoscale Adv.*, 2021, **3**, 6940-6948.

[38] S. Conic, D. Desplancq, A. Ferrand, V. Fischer, V. Heyer, B. Reina San Martin, J. Pontabry, M. Oulad-Abdelghani, N. K. Babu, G. D. Wright *et al.*, *J. Cell Biol.*, 2018, **217**, 1537-1552.

[39] N. Mosammaparast, K. R. Jackson, Y. Guo, C. J. Brame, J. Shabanowitz, D. F. Hunt, L. F. Pemberton, *J. Cell Biol.*, 2001, **153**, 251-262.

[40] J. Teissié, M. P. Rols, *Ann. N. Y. Acad. Sci.*, 1994, **720**, 98-110.

[41] C. Kanthou, S. Kranjc, G. Sersa, G. Tozer, A. Zupanic, M. Cemazar, *Mol. Cancer Ther.*, 2006, **5**, 3145-3152.

[42] G. Danscher, J. O. Nörgaard, *J. Histochem. Cytochem.*, 1983, **31**, 1394-1398.

[43] R. W. Burry, D. D. Vandre, D. M. Hayes, *J. Histochem. Cytochem.*, 1992, **40**, 1849-1856.

[44] D. N. Mastronarde, *J. Struct. Biol.*, 2005, **152**, 36-51.

[45] J. M. de la Rosa-Trevín, A. Quintana, L. del Cano, A. Zaldívar, I. Foche, J. Gutiérrez, J. Gómez-Blanco, J. Burguet-Castell, J. Cuenca-Alba, V. Abrishami *et al.*, *J. Struct. Biol.*, 2016, **195**, 93-99.

[46] G. Tang, L. Peng, P. R. Baldwin, D. S. Mann, W. Jiang, I. Rees, S. J. Ludtke, *J. Struct. Biol.*, 2007, **157**, 38-46.

Comptes Rendus
Chimie
2023, Vol. 26, Special Issue S3, p. 67-77
https://doi.org/10.5802/crchim.255

Breaking Barriers in Chemical Biology – Toulouse 2022

Reduced Schiff-base derivatives to stop reactive oxygen species production by the Cu(Aβ) species: a structure–activity relationship

Margot Lefèvre [a], Lielou Lantigner [a], Laura Andolfo [a, b], Corinne Vanucci-Bacqué [b], Eric Benoist [b], Charlène Esmieu [a], Florence Bedos-Belval [*, b] and Christelle Hureau [*, a]

[a] LCC-CNRS, Université de Toulouse, CNRS, Toulouse, France

[b] LSPCMIB, CNRS UMR 5068, Université Toulouse III-Paul Sabatier, 118 route de Narbonne, 31062 Toulouse cedex 9, France

E-mails: margot.lefevre@lcc-toulouse.fr (M. Lefèvre), lielou.lantigner@lcc-toulouse.fr (L. Lantigner), laura.andolfo21@gmail.com (L. Andolfo), corinne.bacque@univ-tlse3.fr (C. Vanucci-Bacqué), eric.benoist@univ-tlse3.fr (E. Benoist), charlene.esmieu@lcc-toulouse.fr (C. Esmieu), florence.bedos@univ-tlse3.fr (F. Bedos-Belval), christelle.hureau@lcc-toulouse.fr (C. Hureau)

Abstract. Aβ is the peptide involved in Alzheimer's disease. Its binding to the redox copper ions and the subsequent production of reactive oxygen species (ROS) contributing to the overall oxidative stress observed in the disease is among the deleterious effects Aβ may have. Here seven reduced Schiff-base ligands were studied for their ability to stop Cu(Aβ)-induced ROS production. The spectroscopic UV–vis and EPR characterizations of the Cu(II) complexes are reported as well. While all the ligands except one are able to stop Cu(Aβ)-induced ROS, only two maintain this ability in the presence of the endogenous Zn ions, due to kinetic competitiveness.

Keywords. Copper, Alzheimer, Peptide, 2N2O ligand, Oxidative stress.

Funding. C. H. acknowledges ERC StG 638712, C. E. and M. L. the ANR-20-CE07-0009 for financial support.

Manuscript received 5 April 2023, accepted 22 August 2023.

1. Introduction

Alzheimer's disease (AD) is an incapacitating disease that represents the major proportion of dementia cases in the elderly [1,2]. AD is a multifactorial pathology with genetic predispositions and multiple probable causes [3]. Among them, a first hypothesis is known as "the amyloid cascade" [4–6]. It is based on the detection of senile plaques consisting of aggregated amyloid-β (Aβ) peptides, where Aβ are peptides of approximately 40 amino acid residues in length. Therefore, the process from the monomeric peptides to the amyloid fibrils, so-called self-assembly or aggregation, is regarded as a key element in the development of the disease. Another hypothesis is the "metal hypothesis" [7–9]. It relies on a deregulation of metal ions levels, mainly copper (Cu) and zinc (Zn) ions. It is based on the following facts: (1) metal ions are exchanged with the synaptic cleft where Aβ peptides aggregate, (2) their respective affinities for the Aβ peptides match within their biological concentration, making possible their

* Corresponding authors

interaction with Aβ, (3) aberrant metal levels (up to mM to be compared to μM in the cerebrospinal fluid) are found in the senile plaques. In addition to their likely participation in the modulation of Aβ aggregation [10,11], and because they are redox-active, Cu ions can produce highly deleterious Reactive Oxygen Species (ROS) when bound to the Aβ peptide and thus can participate in the oxidative stress observed in AD [12]. Zn ions are present in much higher quantities than Cu ions in the synaptic cleft and, in contrast to Cu ions, they are redox silent.

Due to its redox ability and participation in oxidative stress, Aβ-bound Cu are thus considered as a therapeutic target of interest among others. We and many other groups have developed many copper-targeting strategies to overrule the deleterious effects of Cu (for recent reviews on that topic, see refs. [13–21] and references therein). Among the various properties, if Cu ions are considered as the target of choice of the intended ligands, they should possess a high Cu over Zn(II) thermodynamic selectivity, much higher than that of Aβ since there are about 10–100 fold more Zn(II) than Cu ions in the synaptic cleft [22,23].

In the present article, we aim to complete seminal works on Schiff base derivatives, used as Cu(II) ligands able to retrieve Cu(II) from Aβ, redox-silence it and be selective enough for Cu(II) versus Zn(II) to maintain such ability in the co-presence of stoichiometric amount of Zn(II). In a seminal paper, Storr and coworkers described the synthesis of glucose derivatives of reduced Schiff base ligands, such as $^{G}L_{Me}$ (Scheme 1), and the thorough characterization of the corresponding Cu(II) complexes [24]. Later on, with the objective of helping *in vitro* characterizations by increasing the solubility of the ligand, the $^{s}L_{Me}$ analogue was reported (Scheme 1). It showed the ability to remove Cu(II) from Aβ, to stop Cu(Aβ)-induced ROS formation, and to restore apo-like aggregation of Cu(Aβ) [25]. Then the $^{s}L_{Me}$ ligand was shown to have an appropriate selectivity to maintain its ability to stop Cu(Aβ)-induced ROS production in the presence of one equivalent of Zn(II) [23]. These ligands have higher affinity and Zn(II) over Cu(II) selectivity than Aβ and hence they maintain the ability to extract Cu from Aβ in the presence of one equiv. of Zn(II). The affinity for a ligand L and a metal ion M is defined as $K^L_M = (LM)/((L)(M))$ while the Cu(II) over Zn(II) selectivity for a ligand L as $S^L = K^L_{Cu}/K^L_{Zn}$.

Table 1. Apparent affinity values (for Cu(II) and Zn(II)) at pH 7.1 for the Aβ peptide and the two reference ligands

L	$\log(K^L_{Cu})$	$\log(K^L_{Zn})$	$\log(S^L)$	Ref.
Aβ	9.2	5.0	4.2	[26,27]
$^{G}L_{Me}$	12.1	4.6	7.5	[24]
$^{s}L_{Me}$	13.8	6.1	7.7	[25,27]

The affinity and selectivity values for $^{G/s}L_{Me}$ are reported along those for Aβ in Table 1.

Here, we report on the synthesis and characterizations of a series of ligands based on the two $^{G/s}L_{Me}$ previously described (Scheme 1). They will be noted $\mathscr{L}$ as a generic term. We aimed to improve the ligand design and several lines were followed: effect of the presence of (i) sulfonato groups on the phenol rings ($^{s}\mathscr{L}$ versus $\mathscr{L}$), (ii) methyl substituents on the amine functions ($\mathscr{L}_{Me}$ versus $\mathscr{L}$) and (iii) a (±)-trans-1,2-cyclohexyl as bridge ($\mathscr{L}'$ versus $\mathscr{L}$) inspired by previous works on similar scaffolds [28–30]. Rationales for such ligand alterations were (i) to demonstrate that $^{s}\mathscr{L}$ can be studied *in vitro* while $\mathscr{L}$ could be used for further *in vivo* applications, since the $^{s}\mathscr{L}$ would not have a correct drug profile, especially to cross the blood–brain barrier, a key step when AD is targeted, (ii) to draw a structure–activity relationship, linking Cu(II) first coordination sites and the arrest of Cu(Aβ)-induced ROS production including in the presence of Zn(II), and (iii) to question kinetic issues beyond the thermodynamic approach mostly described until now.

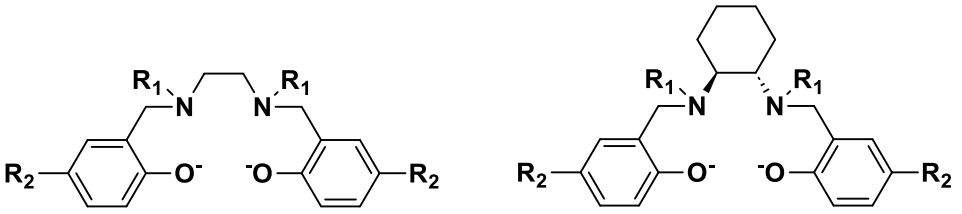
2. Results

2.1. *Ligands $\mathscr{L}$*

The synthesis of the ligands has been adapted from literature [31–36] and is described in the Supporting Information.

2.2. *Characterizations of Cu($\mathscr{L}$) complexes*

The Cu(II) complexes, formed in situ by the mixture of quasi-stoichiometric ratio between $\mathscr{L}$ and Cu(II) stock solutions (see Figure S1 for the determination of concentration of $\mathscr{L}$) were characterized by UV–vis and EPR spectroscopies. Their signatures are shown



R_1=R_2=H	L	R_1=R_2=H	L'
R_1=Me ; R_2=H	L_{Me}	R_1=Me ; R_2=H	L'_{Me}
R_1=H ; R_2=SO_3^-	^{S}L	R_1=H ; R_2=SO_3^-	$^{S}L'$
R_1=Me ; R_2=SO_3^-	$^{S}L_{Me}$		
R_1=Me ; R_2=O-Glucose	$^{G}L_{Me}$		

Scheme 1. Scheme of the various ligands $\mathscr{L}$ under study. Main variations are methylation of the amine (indicated by the *Me* subscript), sulfonation of the phenol arms (indicated by the *s* superscript), and introduction of a cyclohexyl on the ethylene bridge ($\mathscr{L}'$ series). $\mathscr{L}$ corresponds to the fully deprotonated form of any of the ligands. Charges are omitted for clarity.

Table 2. UV–vis and EPR parameters of the Cu($\mathscr{L}$) complexes

$\mathscr{L}$	UV–vis				EPR[a]		
	λ_{max}^{d-d} (nm)	ε^{d-d} ($M^{-1}\cdot cm^{-1}$)	λ_{max}^{LMCT} (nm)	ε^{LMCT} ($10^3\cdot M^{-1}\cdot cm^{-1}$)	$g_{//}$	$A_{//}$ ($10^{-4}\cdot cm^{-1}$)	$g_{\perp}$
L	**595**	232	**382**	1.24	2.23 ± 0.01	201 ± 2	2.05 ± 0.01
L_{Me}	**603**	264	**394**	1.17	2.24 ± 0.01	198 ± 2	2.06 ± 0.01
L′	**590**	194	**382**	1.18	2.23 ± 0.01	204 ± 2	2.05 ± 0.01
L'_{Me}	**577**	324	**398**	1.30	2.23 ± 0.01	206 ± 2	2.06 ± 0.01
^{S}L	**605**	251	**386**	1.26	2.23 ± 0.01	201 ± 2	2.05 ± 0.01
$^{S}L_{Me}$	**620**	254	**392**	1.23	2.24 ± 0.01	203 ± 2	2.06 ± 0.01
$^{S}L'$	**600**	264	**382**	1.35	2.23 ± 0.01	206 ± 2	2.06 ± 0.01

[a]The *g* values were calculated using the average position of the second and third hyperfine lines, while the hyperfine coupling values correspond to the field differences between the second and third lines to minimize second-order contributions to hyperfine splittings. ^{65}Cu isotope was used.

in Figure 1, panels A and B, while the corresponding parameters are listed in Table 2. All the Cu($\mathscr{L}$) complexes show similar features (note that the complexes are neutral with L, L′ and L_{Me}, L'_{Me} and di-anionic for ^{S}L, $^{S}L'$ and $^{S}L_{Me}$, but that charges will be omitted for clarity). In UV–vis spectra (Figure 1, panel A), d–d bands and phenolato-to-Cu(II) CT (Charge Transfer) transitions are observed near 600 nm and 380 nm, respectively. Some weak differences are observed according to $\mathscr{L}$. In contrast to the d–d band, the LMCT band is not affected by the presence of sulfonato groups on the ligand phenol moiety, whereas the N-methyl group affects it. It should also be noted that the presence of cyclohexyl instead of an ethylene bridge did not affect the LMCT band characteristics. In EPR spectroscopy (Figure 1, panel B), typical spectra of square-planar Cu(II) complexes are obtained, with hyperfine and *g*-value parameters in line with a 2N2O equatorial site, according to the Peisach an Blumberg correlation [37]. Again, some slight differences in the EPR parameters are observed between the various ligands, whereas the presence of two sul-

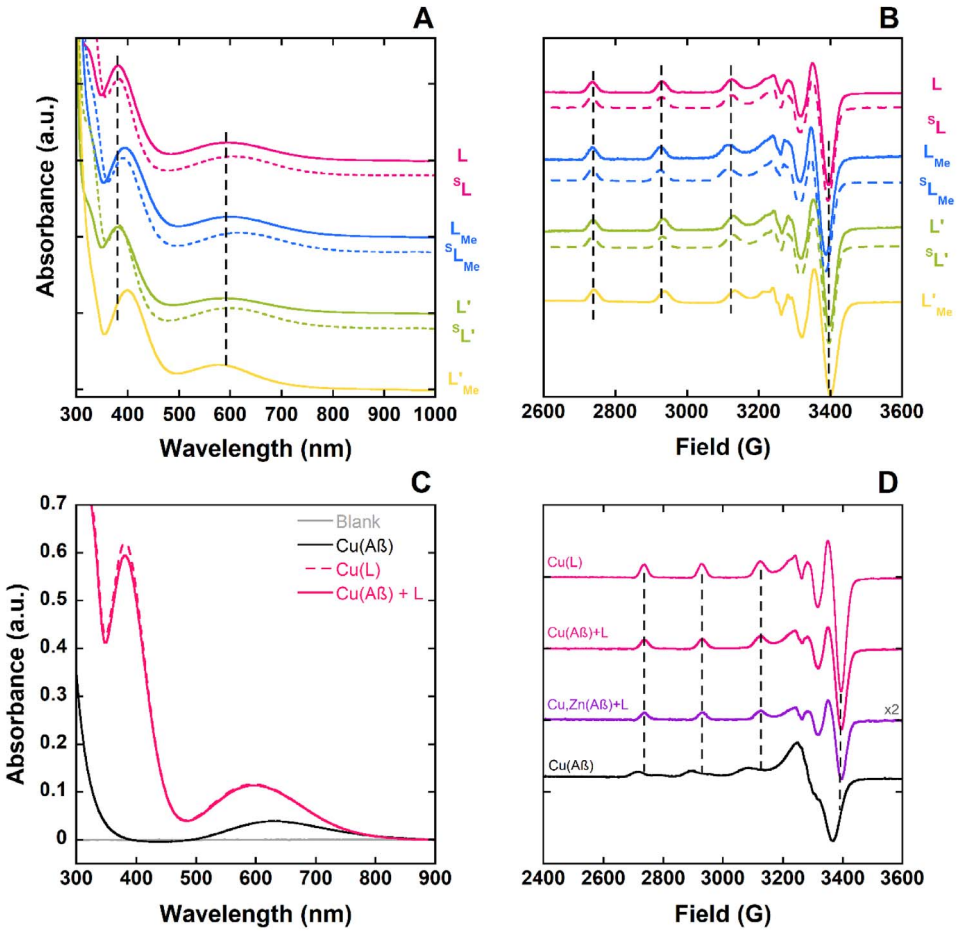


Figure 1. UV–vis (Panel A) and EPR (panel B) spectra of the Cu(II)($\mathscr{L}$) complexes. UV–vis (Panel C) and EPR (panel D) spectra of Cu(II) removal by L from Cu(Aβ) and Cu,Zn(Aβ) respectively. Experimental conditions: panels A and C: [Cu^{II}] = 500 μM, [L] = [Aβ] = [Zn^{II}] = 500 μM, [HEPES] = 100 mM, pH 7.4, T = 25 °C; panel B: [$^{65}Cu^{II}$] = 480 μM, [L, Aβ] = 500 μM; and panel D: [$^{65}Cu^{II}$] = 480 μM, [L, Aβ] = 500 μM or [$^{65}Cu^{II}$] = [Zn] = 190 μM, [L, Aβ] = 200 μM; panels B and D: [HEPES] = 50 mM, pH 7.4, 10% of glycerol as cryoprotectant, T = 120 K, ν ≈ 9.5 GHz, mod. ampl. = 5 G, microwave power: 5 mW. In panel D, EPR data corresponding to Cu,Zn(Aβ) + L have been multiplied by 2 in order to account for the 2.5-fold lower concentration used.

fonato groups on the ligands has virtually no impact on the Cu($\mathscr{L}$) EPR signatures. Again, as for UV–vis spectra, methylation of the secondary amine leads to the most significant difference with a decrease in the hyperfine coupling ($A_{//}$) values and an increase in the $g_{//}$ values, while grafting of the cyclohexyl on the ethylene diamine bridge has the opposite effect.

In brief, the spectroscopic characterizations indicate that the first coordination sphere of the complexes are identical in buffered solution, but that some minor second sphere structural changes occur.

2.3. *Cu(II) removal from Cu(Aβ)*

Next, we performed competition experiments, monitored by UV–vis and EPR spectroscopies (Figure 1, panels C and D in case of ligand L and Figures S2 and S3 for the other ligands), to check that the ligands $\mathscr{L}$ were able to remove Cu(II) from Cu(Aβ) including in

the presence of Zn. In other words, we check that $\mathscr{L}$ have higher Cu(II) affinity and Cu(II) over Zn(II) selectivity than Aβ. The spectroscopic signatures of Cu($\mathscr{L}$) were recovered during the competition experiments (Cu(Aβ) + $\mathscr{L}$ or Cu,Zn(Aβ) + $\mathscr{L}$). Thus all the ligands $\mathscr{L}$ are efficient in retrieving Cu from Cu(Aβ) regardless of the presence of one equiv. of Zn(II). Hence, the thermodynamic parameters of $\mathscr{L}$ are suitable for performing Cu(II) removal from Aβ including in the presence of one equiv. of Zn(II). However, during the course of the competition experiment, we noticed that the thermodynamic equilibrium is reached more slowly in the case of the ligand L'_{Me} (about 30 min versus less than 5 min for the others ligands, Figure S4).

2.4. *Effect of ligands $\mathscr{L}$ on Cu(Aβ)-induced ROS production*

Finally, Cu(Aβ)-induced ROS formation was evaluated using a very suitable and straightforward experiment, namely the ascorbate (Asc) consumption assay. This is an appropriate method to monitor ROS formation. Briefly, it consists in measuring the absorption of Asc (at 265 nm, $\varepsilon = 14{,}500\ M^{-1}{\cdot}cm^{-1}$) that is the reductant fueling the incomplete reduction of O_2 to $O_2^{\circ-}$, H_2O_2 and HO° [21], and that does not absorb once oxidized. It has been previously shown that Cu and Cu(Aβ)-induced Asc consumption mirrors the formation of H_2O_2 and HO° [38,39]. A ligand will be efficient in stopping Asc consumption if it can bind Cu(II) or remove it from Aβ, and then form a Cu(II) complex resistant to reduction by Asc, eventually interrupting the redox cycle of Cu and thus the production of ROS.

With Asc consumption assays, two distinct experiments can be performed, either by incubating the various chemical partners at play and then triggering the reaction by addition of Asc (named "experiments A", Figure S5 and Figure 2, panel D), or by adding ligand $\mathscr{L}$ during Cu(Aβ)-induced Asc consumption (named "experiments B", Figure 2, panels A to C). In experiments A, if the incubation time is long enough, the results indicate the ability of $\mathscr{L}$ to remove Cu from Cu(Aβ) including in the presence of increasing stoichiometry of Zn(II), thus documenting the thermodynamics of the reaction of Cu(II) extraction out of Aβ. This is a very straightforward way to evaluate whether the selectivity of a ligand *L* is appropriate [22,23]. More specifically, such a method is more suited to screen between various ligands than the individual determination of Cu(*L*) and Zn(*L*) formation that would also release the selectivity of the *L* (see introduction). In experiments B, additional kinetic parameters are involved (Scheme 2A) [40–43]. Indeed, during Asc consumption, Cu oscillates between the +I and +II redox states. To be efficient, the tested ligand thus has to be faster in removing Cu(II) from Aβ than the reduction of Cu(II)(Aβ) to Cu(I)(Aβ) by Asc. In the $\mathscr{L}$ series, we wanted to decipher the importance of Zn stoichiometry in both the thermodynamics (experiments "A") and kinetics (experiments "B") of Cu(II) removal. In the following, all the experiments were performed with 10 μM of Cu(II) ions, 1.2 equiv. of Aβ and/or $\mathscr{L}$, and 0, 1 or 10 equiv. of Zn(II). Note that for practical reasons, the experiments are run with a slight excess of Aβ and/or $\mathscr{L}$ to avoid any possibility of unbound Cu(II) ions and/or Aβ-bound Cu(II) complexes, respectively. Indeed, this would dramatically change the rate of Asc. consumption.

In the absence of Zn and after short incubation (300 s) (Figure S5, panel A), all the ligands except L'_{Me} are able to stop Cu(Aβ) induced Asc consumption. After a longer incubation time (>18 h), L'_{Me} becomes efficient (Figure S6). This indicates that Cu(II) removal from Aβ by L'_{Me} is slower than that with the other $\mathscr{L}$ ligands, but thermodynamically possible, in line with the competition experiments previously described. In addition to showing that all the ligands are able to extract Cu(II) from Aβ, these experiments confirm that all the Cu($\mathscr{L}$) complexes formed are resistant to reduction by ascorbate. In line with these first results, all the ligands, except L'_{Me} and to a lesser extent L′ and L_{Me}, are able to stop Cu(Aβ)-induced Asc consumption when added during the course of the experiments (Figure 2, panel A).

2.5. *Effect of ligands $\mathscr{L}$ on the Cu(Aβ)-induced ROS production in the presence of Zn(II)*

In the presence of Zn, the differences between the various ligands appear more clearly. This may be due to the additional competition reaction between Cu(II) or Zn(II) removal from Cu,Zn(Aβ) (Scheme 2, panel B). With a short incubation (300 s) (Figure S5,

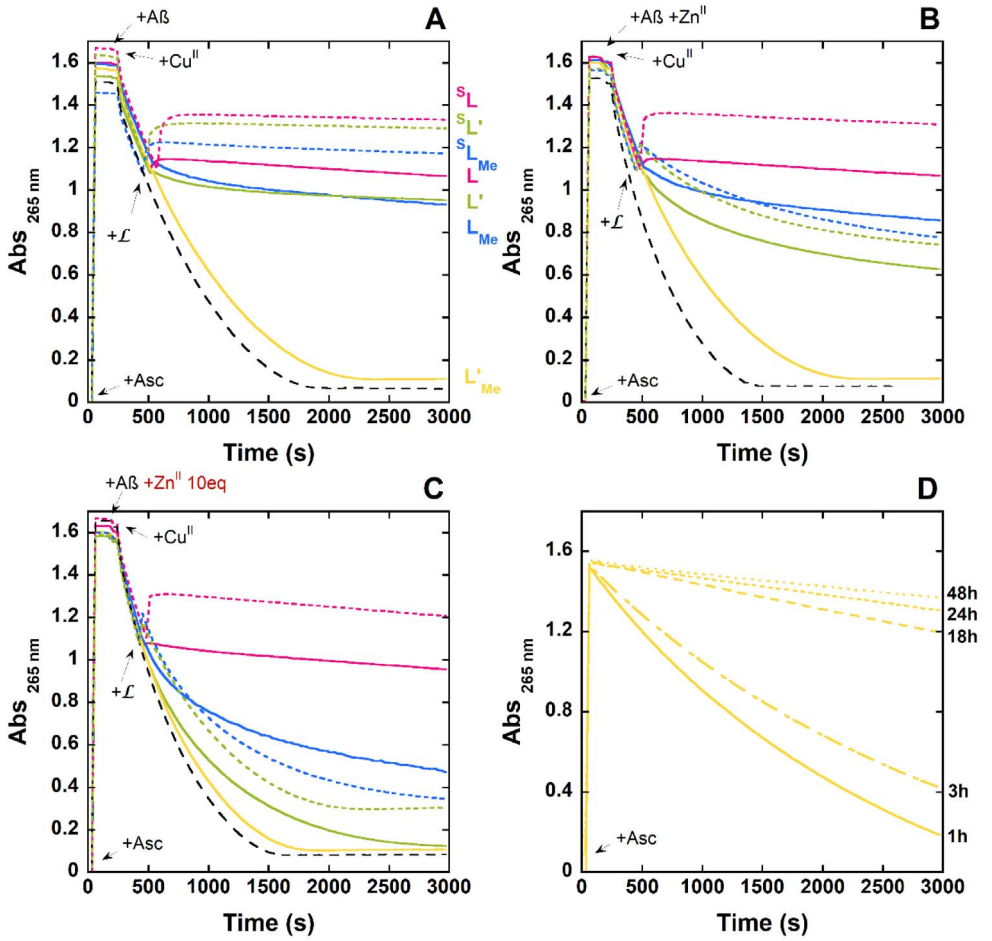
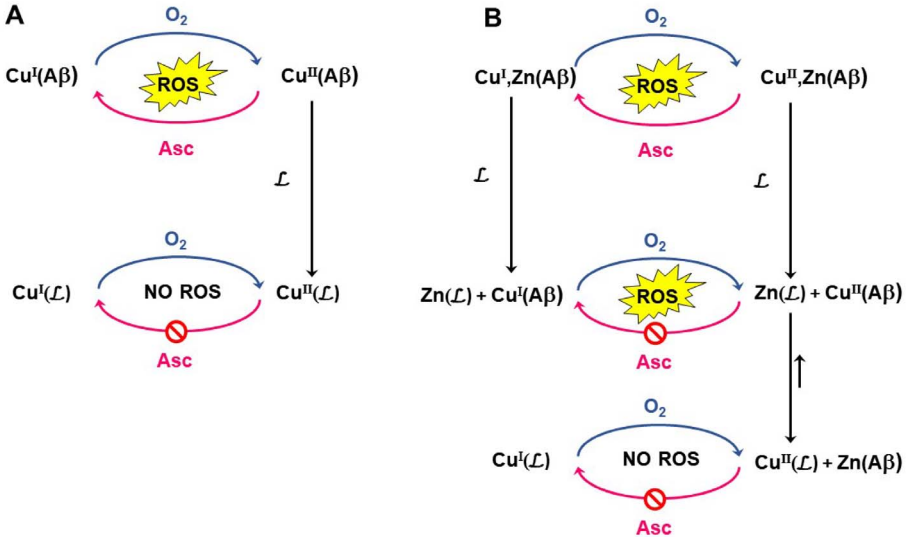


Figure 2. Ascorbate consumption is followed at 265 nm (with a baseline subtraction at 800 nm) induced by Cu(Aβ). Panels A to C: starting from a Cu(I)/Cu(II) mixture and in the presence of increasing Zn(II) stoichiometry. [$\mathscr{L}$] = [Aβ] = 12 μM, [Zn^{II}] = 0, 12 and 120 μM, [Cu^{II}] = 10 μM, [Asc] = 100 μM, [HEPES] = 100 mM, pH 7.4, T = 25 °C. Panel D: starting from Cu(II) and after increasing incubation times from 1 to 48 hours. [L'_{Me}] = [Aβ] = 12 μM, [Zn] = 120 μM, [Cu^{II}] = 10 μM, [Asc] = 100 μM, [HEPES] = 100 mM, pH 7.4, T = 25 °C. The arrows indicate the time at which the different components are added into the cuvette.

panel B) or when the ligand is added during Cu(Aβ)-induced Asc consumption (Figure 2, panel B), the trends are the same. The "efficiency trend" (i.e. the ability of the ligand to stop Cu(Aβ)-induced Asc consumption in the presence of Zn) is as follows: L ~ ^{s}L > L_{Me} ~ $^{s}L_{Me}$ $\gtrsim$ $^{s}L'$ $\gtrsim$ L′ ≫ L'_{Me}. With higher stoichiometry of Zn(II) (10 equiv.), the differences are more obvious and only the L and ^{s}L ligands maintain their ability to stop Cu(Aβ)-induced Asc consumption (Figure 2, panel C).

Hence, it appears here that Zn(II) induces a change in the ability of the ligands $\mathscr{L}$ to stop Cu(Aβ)-induced ROS production and that this effect is dependent on the ligands $\mathscr{L}$, with a dependence that is more clearly revealed in the presence of 10 equiv. of Zn(II). Two explanations can be proposed (i): when increasing the ratio of Zn(II), the thermodynamic equilibria at play change and the ligand becomes thermodynamically unable to remove Cu(II) from Aβ, or (ii) Zn(II) has a kinetic impact on Cu extraction from Aβ. This kinetic effect would be due to the faster formation of the Zn($\mathscr{L}$) complexes that have



Scheme 2. Mechanisms of arrest of Cu(Aβ)-induced ROS production by a Cu(II)-targeting ligand $\mathscr{L}$, in the absence (Panel A) and presence of Zn (panel B). In the absence of Zn(II), the reaction of Cu(II) extraction from Aβ is in competition with Cu(II)(Aβ) reduction by Asc. In the presence of Zn(II), an additional competition reaction occurs due to the possible formation of Zn($\mathscr{L}$) that will retard the formation of Cu($\mathscr{L}$).

further to swap their metallic center with Cu(Aβ) (Scheme 2, panel B). To discriminate between these two hypotheses, we also performed experiments A with a longer pre-incubation time of Cu(Aβ) + $\mathscr{L}$ in the presence of 10 of equiv. Zn(II). Several incubation times were tested: 1, 3, 18, 24, and 48 h (Figure 2, panel D and Figure S7). The results obtained show that with a sufficient incubation time, all the ligands prevent Cu(Aβ)-induced ROS production and thus that the overall effect of Zn(II) is kinetic. For all ligands, except L'_{Me} which requires at least 24 h of incubation, (Figure 2D), 1 h of incubation is sufficient to prevent Asc consumption (Figure S7).

3. Discussion

The various results obtained with the ligands $\mathscr{L}$, are compiled in Figure 3, where the rates of ascorbate consumption are reported.

3.1. *Effect of p-sulfonation on phenol rings*

The sulfonated version of three ligands were tested. Sulfonation has no significant effect on the properties of the parent ligand regarding Cu(Aβ) and Cu,Zn(Aβ)-induced ROS production. This makes possible the use of such water-soluble counterparts for *in vitro* investigations, while the parent ligands could be engaged in *in vivo* experiments.

3.2. *Structure–activity relationship*

3.2.1. *Ligands other than* L'_{Me}

Ligands other than L'_{Me} are all able to stop Cu(Aβ)-induced ROS production in the absence of Zn(II) when added in the course of Cu(Aβ)-induced Asc. consumption (Experiments "B", Figure 3, first set of columns). Besides, in the presence of increasing ratios of Zn(II) and with long enough incubation times, they all maintain this ability (Figure 3, fourth set of columns, Figures S5 and S7). Finally, when added during the course of Asc consumption, some of them become less and less efficient as the stoichiometry of Zn(II) is increased (Figure 3, second and third sets of columns).

To explain such observations, the chemical reactions shown in Scheme 3 are proposed. According to Scheme 3, the overall Zn(II)-induced slowdown of Cu(II) removal out of Aβ (and consequent effect on Cu(Aβ)-induced Asc consumption) could be due

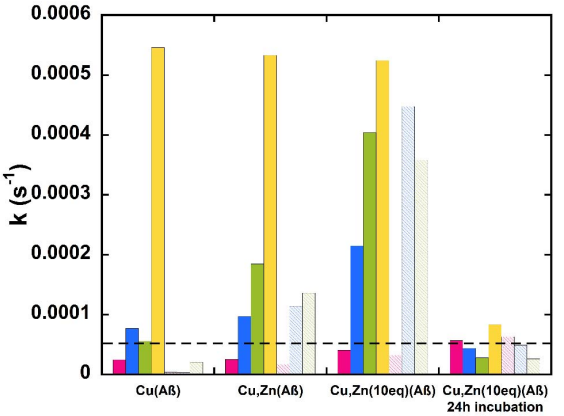


Figure 3. Rates of Cu(Aβ)-induced ascorbate consumption as a function of the ligand and equivalent of Zn(II), starting from Cu(I)/Cu(II) (corresponding to data in Figure 2, panels A–C) and from Cu(II) after 24 h of incubation (corresponding to data in Figure 2, panel D and Figure S7). L in pink, L_{Me} in blue, L′ in green, and L'_{Me} in yellow, corresponding sulfonated ligands are displayed in light colors. The rate of Asc consumption was determined by measuring the slope of the Asc consumption curves between the time (t) of the addition of $\mathscr{L}$ and $t+300$ s, approximating the curves as straight lines, and with [Asc] = 100 μM and [Cu] = 10 μM. The dotted line indicates the level of Asc consumption below which Asc consumption is considered to be similar to that of Asc alone in the buffer (due to Asc auto-oxidation).

to both thermodynamic and kinetic factors for each individual reaction at play, the determination of which is beyond the scope of the present study.

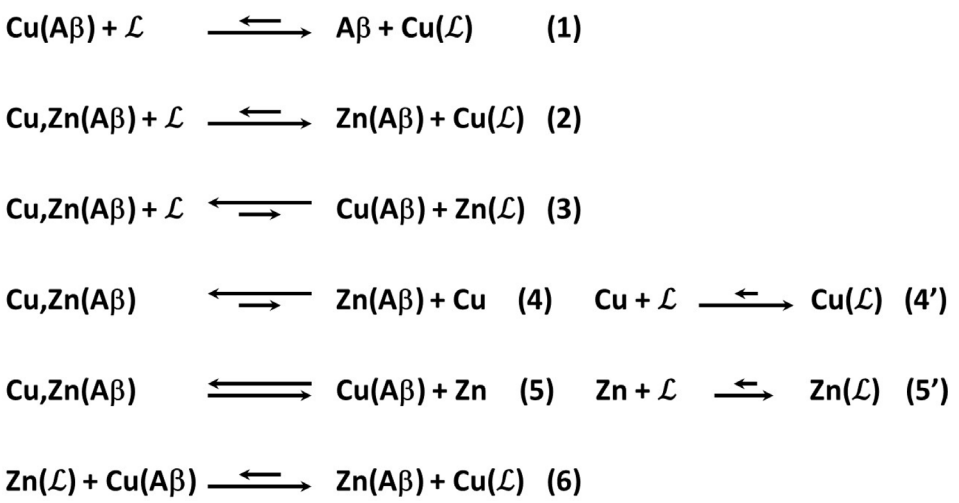
Reaction (1) corresponds to Cu(II) removal from Aβ by $\mathscr{L}$ in the absence of Zn(II). Then, in the presence of Zn(II), the addition of $\mathscr{L}$ to Cu,Zn(Aβ) leads to the two possible reactions (2) and (3) and the formation of Cu($\mathscr{L}$) or Zn($\mathscr{L}$), respectively. From a thermodynamic point of view, reaction (2) is much more favored than reaction (3) as probed by the experiments with long incubation times (Experiments "A"). However, reactions (2) and (3) are in kinetic competition, with reaction (3) being faster than reaction (2) because otherwise the presence of Zn(II) would have no strong effect on experiments "B". The progressive loss of the ability to stop Cu(Aβ)-induced Asc consumption with increasing levels of Zn(II) (Figure 3) thus mirrors the impact of Zn(II) on the overall rate of Cu(II) removal from Aβ. Such an effect may be due (i) to the kinetic competition of Zn(II) with Cu(II) for binding to $\mathscr{L}$ (reaction (2) versus reaction (3)) and/or (ii) to a modification in the rate of reaction (6) (Scheme 3).

With the series of ligands $\mathscr{L}$, Zn(II) induces an overall slowdown of Cu(II) removal. This contrasts with a previous study on peptide-based ligands [44], where an acceleration effect due to Zn-induced Cu expelling from Cu,Zn(Aβ) was reported (corresponding to the path shown by reactions (4) and (4′)). This indicates that reaction (4) is not predominant in the case of $\mathscr{L}$. Hence, it will not be considered in the rest of the discussion.

The best ligands able to resist up to 10 equiv. of Zn are L and ^{s}L, whereas ^{s}L′, L_{Me}, and $^{s}L_{Me}$ work correctly only in the presence of one equiv. of Zn(II). Hence, the cyclohexyl-grafted ethylene bridge and, to a lesser extent the N-methyl substituents, increase the level of Zn($\mathscr{L}$) versus Cu($\mathscr{L}$) at a given time (before the thermodynamic equilibrium is reached). The weaker effect of L′ on the arrest of Cu(Aβ)-induced Asc consumption may be in line with the reported *cis*-β arrangement (i.e. with the two phenolato moiety being in *cis*) of ligand L′ in $[Cu_2(HL')_2]^{2+}$ species [45] unobserved with other ligands, for which *trans* arrangements (i.e. all the endogenous atoms from $\mathscr{L}$ are in the equatorial plane) were reported [25,46–48]. We can indeed anticipate that such *cis*-β arrangement is more appropriate for Zn(II) versus Cu(II) binding in contrast to the *trans* arrangement which is more appropriate for Cu(II) versus Zn(II).

3.2.2. *Ligand* L'_{Me}

Among the series, ligand L'_{Me} holds a specific place with very slow kinetics of Cu extraction from Aβ both in the absence (reaction (1)) and presence of Zn(II) (reactions (2)–(6), Scheme 3). In the absence of Zn(II), this is shown in the experiments with a "short" (i.e. 300 s) incubation time (Figure S5, panel A). Indeed, since all the Cu(II) is not extracted from Aβ after such a short time, when Asc is added, its consumption is induced by the remaining Cu(Aβ). To complete the reaction of Cu extraction from Aβ, more than 18 h are required (Figure S6). In the presence of 10 equiv. of Zn(II), this is slightly worse (the



Scheme 3. Various possible reactions at play in Cu(II) removal out of Aβ by $\mathscr{L}$, including in the presence of one equiv. of Zn(II).

level of Asc. consumption lessening in the presence of 10 equiv. of Zn(II) after 24 h equals that in the absence of Zn(II) after 18 h, Figure 2, panel D to be compared with Figure S6). This is due to other reactions coming at play, namely reactions (2)–(6), as described previously for the other ligands. The specific properties of the L'_{Me} ligand may be linked to the combination of the presence of the 1,2-cyclohexyl bridge and the N-methyl group, which makes possible a *cis*-β arrangement of the ligand (i.e. with the two phenolato moiety being in *cis*), described in the solid-state for Cu(L'_{Me}) [49]. However this geometry is not recovered in the thermodynamically-stable complex formed in aqueous solution as probed by EPR and to a lesser extent UV–vis signatures. But it may exist transiently and slow down the Cu(II) complex formation (reaction (1)), favor and/or accelerate Zn(L'_{Me}) complex formation versus Cu(L'_{Me}) complex (reaction (3) versus (2)), and/or disfavor and/or slow down reaction (6) with respect to other ligands in the series. Because the difference between the ability to prevent Cu(Aβ)-induced ROS formation without Zn(II) (Figure S6) and 10 equiv. of Zn(II) (Figure 2D) is fairly weak, we can anticipate that reaction (1) has the most important contribution.

In addition, the trend $L'_{Me} < L' < L_{Me} \ll L$ in the ability to stop Cu(Aβ)-induced ROS formation in the presence of Zn(II) is reminiscent of recently published data using a series of pentadentate ligands (^{p}L) based on an ethylene bridge (possibly appended with a cyclo-hexyl moiety: ^{p}L′, with a N-propyl group: $^{p}L_{pro}$ and both: $^{p}L'_{pro}$) (Scheme S1). Once chelated with Mn(II), the ligand exchange with Cu(II) that corresponds to reaction (6) here, follows the order: $^{p}L'_{pro} < ^{p}L' < ^{p}L_{pro} < ^{p}L$ (Scheme S2) [28].

3.3. *Kinetic impact of Zn(II)*

In brief, the study of the series of ligands $\mathscr{L}$ reveals an impact of Zn(II) on Cu(II) extraction from Aβ that is apparent during the Asc consumption experiments and which is beyond only thermodynamic considerations. Indeed, if Cu,Zn(Aβ) and the ligands are incubated long enough, all the ligands can prevent Cu(Aβ)-induced ROS production. However, in the course of Asc. consumption (no incubation) or when the incubation is short (300 s), ligands $\mathscr{L}$ can be differentiated by their efficiency to stop Cu(Aβ)-induced ROS production in the presence of various Zn(II) levels. Their ability to stop Cu(Aβ)-induced ROS production strongly depends on subtle variations in the ligand scaffold because all the ligands studied here were built on the very same first coordination sphere. Thermodynamic selectivity is now a parameter considered in the design of Cu-targeting ligands in the context of AD [21,22]. In the present study, we also show that it is important to consider the effect of Zn(II) which can modify the rate of Cu(II)

extraction from Aβ by a ligand and can thus prevent the formation of the corresponding Cu(II) complex resistant to Asc. reduction and the associated arrest of Cu(Aβ)-induced ROS production.

3.4. *Perspectives*

Among the tested ligands, the simplest L appears to be the ideal candidate with respect to its ability to stop Cu(Aβ)-induced ROS formation in the presence of a biologically relevant ratio of the competing Zn(II) ion. Further studies on this ligand series will aim to relate the *in vitro* data obtained here with their ability to relieve the cellular toxicity induced by ROS produced by Cu(Aβ) with and without Zn, as reported for other ligands [50–54]. Since L is expected to be partly neutral at pH 7 based on potentiometric data reported for $^{G/S}L_{Me}$ [24,25], and thus to fulfill Lipinski's rules, it appears as a good candidate for blood–brain barrier penetration by passive diffusion and thus for further *in vivo* studies on AD animal models.

Declaration of interests

The authors do not work for, advise, own shares in, or receive funds from any organization that could benefit from this article, and have declared no affiliations other than their research organizations.

Acknowledgments

Sonia Mallet-Ladeira is acknowledged for providing the X-ray crystallographic data of related Cu(II) complexes.

Supplementary data

Supporting information for this article is available on the journal's website under https://doi.org/10.5802/crchim.255 or from the author.

References

[1] Alzheimer's Association, *Global Dementia Cases Forecasted to Triple by 2050*, 2021, https://aaic.alz.org/releases_2021/global-prevalence.asp.

[2] BrightFocus Foundation, "Alzheimer's Disease: Facts & Figures", 2021, https://www.brightfocus.org/alzheimers/article/alzheimers-disease-facts-figures.

[3] Z. Breijyeh, R. Karaman, *Molecules*, 2020, **25**, article no. 5789.

[4] G. B. Frisoni, D. Altomare, D. R. Thal, F. Ribaldi, R. van der Kant, R. Ossenkoppele, K. Blennow, J. Cummings, C. van Duijn, P. M. Nilsson, P.-Y. Dietrich, P. Scheltens, B. Dubois, *Nat. Rev. Neurosci.*, 2022, **23**, 53-66.

[5] J. Levin, J. Vöglein, Y. T. Quiroz, R. J. Bateman, V. Ghisays, F. Lopera, E. Mcdade, E. Reiman, P. N. Tariot, J. C. Morris, *Alzheimer's Dement.*, 2022, **18**, 2687-2698.

[6] D. J. Selkoe, J. Hardy, *EMBO Mol. Med.*, 2016, **8**, 595-608.

[7] L.-L. Chen, Y.-G. Fan, L.-X. Zhao, Q. Zhang, Z.-Y. Wang, *Bioorg. Chem.*, 2023, **131**, article no. 106301.

[8] C. Hureau, in *Alzheimer's Disease: Recent Findings in Pathophysiology, Diagnostic and Therapeutic Modalities* (T. Govindaraju, ed.), The Royal Society of Chemistry, 2022, 170-192.

[9] T. J. Huat, J. Camats-Perna, E. A. Newcombe, N. Valmas, M. Kitazawa, R. Medeiros, *J. Mol. Biol.*, 2019, **431**, 1843-1868.

[10] M. G. M. Weibull, S. Simonsen, C. R. Oksbjerg, M. K. Tiwari, L. Hemmingsen, *J. Biol. Inorg. Chem.*, 2019, **24**, 1197-1215.

[11] M. Rana, A. K. Sharma, *Metallomics*, 2019, **11**, 64-84.

[12] C. Cheignon, M. Tomas, D. Bonnefont-Rousselot, P. Faller, C. Hureau, F. Collin, *Redox Biol.*, 2018, **14**, 450-464.

[13] S. K. Singh, V. Balendra, A. A. Obaid, J. Esposto, M. A. Tikhonova, N. K. Gautam, B. Poeggeler, *Metallomics*, 2022, **14**, article no. mfac018.

[14] J. Han, Z. Du, M. H. Lim, *Acc. Chem. Res.*, 2021, **54**, 3930-3940.

[15] K. D. Fasae, A. O. Abolaji, T. R. Faloye, A. Y. Odunsi, B. O. Oyetayo, J. I. Enya, J. A. Rotimi, R. O. Akinyemi, A. J. Whitworth, M. Aschner, *J. Trace Elem. Med. Biol.*, 2021, **67**, article no. 126779.

[16] G. Gromadzka, B. Tarnacka, A. Flaga, A. Adamczyk, *Int. J. Mol. Sci.*, 2020, **21**, article no. 9259.

[17] H. W. Ejaz, W. Wang, M. Lang, *Int. J. Mol. Sci.*, 2020, **21**, article no. 7660.

[18] M. G. Savelieff, G. Nam, J. Kang, H. J. Lee, M. Lee, M. H. Lim, *Chem. Rev.*, 2019, **119**, 1221-1322.

[19] Y. Liu, M. Nguyen, A. Robert, B. Meunier, *Acc. Chem. Res.*, 2019, **52**, 2026-2035.

[20] C. Hureau, in *Encyclopedia of Inorganic and Bioinorganic Chemistry* (R. A. Scott, ed.), Wiley, 2019.

[21] C. Esmieu, D. Guettas, A. Conte-Daban, L. Sabater, P. Faller, C. Hureau, *Inorg. Chem.*, 2019, **58**, 13509-13527.

[22] E. Atrian-Blasco, A. Conte-Daban, C. Hureau, *Dalton Trans.*, 2017, **46**, 12750-12759.

[23] A. Conte-Daban, A. Day, P. Faller, C. Hureau, *Dalton Trans.*, 2016, **45**, 15671-15678.

[24] T. Storr, M. Merkel, G. X. Song-Zhao, L. E. Scott, D. E. Green, M. L. Bowen, K. H. Thompson, B. O. Patrick, H. J. Schugar, C. Orvig, *J. Am. Chem. Soc.*, 2007, **129**, 7453-7463.

[25] S. Noël, F. Perez, S. Ladeira, S. Sayen, E. Guillon, E. Gras, C. Hureau, *J. Inorg. Biochem.*, 2012, **117**, 322-325.

[26] T. Kowalik-Jankowska, M. Ruta, K. Wisniewska, L. Lankiewicz, *J. Inorg. Biochem.*, 2003, **95**, 270-282.

[27] S. Noël, S. Bustos, S. Sayen, E. Guillon, P. Faller, C. Hureau, *Metallomics*, 2014, **6**, 1220-1222.

[28] G. Schanne, M. Zoumpoulaki, G. Gazzah, A. Vincent, H. Preud'homme, R. Lobinski, S. Demignot, P. Seksik, N. Delsuc, C. Policar, *Oxid. Med. Cell. Longev.*, 2022, **2022**, article no. 3858122.

[29] E. M. Gale, I. P. Atanasova, F. Blasi, I. Ay, P. Caravan, *J. Am. Chem. Soc.*, 2015, **137**, 15548-15557.
[30] F. K. Kálmán, G. Tircsó, *Inorg. Chem.*, 2012, **51**, 10065-10067.
[31] X. Zhao, D. Zhang, R. Yu, S. Chen, D. Zhao, *Eur. J. Inorg. Chem.*, 2018, **2018**, 1185-1191.
[32] V. J. Lillo, J. Mansilla, J. M. Saá, *Angew. Chem. Int. Ed.*, 2016, **55**, 4312-4316.
[33] I. Correia, J. C. Pessoa, M. T. Duarte, M. F. M. da Piedade, T. Jackush, T. Kiss, M. M. C. A. Castro, C. F. G. C. Geraldes, F. Avecilla, *Eur. J. Inorg. Chem.*, 2005, **2005**, 732-744.
[34] K. Voronova, L. Homolya, A. Udvardy, A. C. Bényei, F. Joó, *ChemSusChem*, 2014, **7**, 2230-2239.
[35] Y. Yong-Kang, Y. Cai-Xia, H. Fang-Jun, *J. Coord. Chem.*, 2014, **67**, 2039-2047.
[36] H. Du, A. H. Velders, P. J. Dijkstra, J. Sun, Z. Zhong, X. Chen, J. Feijen, *Chem. Eur. J.*, 2009, **15**, 9836-9845.
[37] J. Peisach, W. E. Blumberg, *Arch. Biochem. Biophys.*, 1974, **165**, 691-708.
[38] B. Alies, I. Sasaki, O. Proux, S. Sayen, E. Guillon, P. Faller, C. Hureau, *Chem. Commun.*, 2013, **49**, 1214-1216.
[39] E. Atrian-Blasco, E. Cerrada, A. Conte-Daban, D. Testemale, P. Faller, M. Laguna, C. Hureau, *Metallomics*, 2015, **7**, 1229-1232.
[40] A. Conte-Daban, M. Beyler, R. Tripier, C. Hureau, *Chem. Eur. J.*, 2018, **24**, 13058-13058.
[41] A. Conte-Daban, M. Beyler, R. Tripier, C. Hureau, *Chem. Eur. J.*, 2018, **24**, 8447-8452.
[42] K. P. Malikidogo, M. Drommi, E. Atrián-Blasco, J. Hormann, N. Kulak, C. Esmieu, C. Hureau, *Chem. Eur. J.*, 2023, **29**, article no. e202203667.
[43] M. Lefèvre, K. P. Malikidogo, C. Esmieu, C. Hureau, *Molecules*, 2022, **27**, article no. 7903.
[44] P. S. Gonzalez, L. Mathieu, Emilie, P. Faller, C. Hureau, *Biomolecules*, 2022, **12**, article no. 1327.
[45] H. Hosseini-Monfared, S. Soleymani-Babadi, S. Sadighian, A. Pazio, K. Wozniak, M. Siczek, P. Mayer, *Transit. Met. Chem.*, 2015, **40**, 255-267.
[46] P. Adão, S. Barroso, F. Avecilla, M. C. Oliveira, J. C. Pessoa, *J. Organomet. Chem.*, 2014, **760**, 212-223.
[47] A. Hazari, C. Diaz, A. Ghosh, *Polyhedron*, 2018, **142**, 16-24.
[48] Y. Xie, Q. Liu, H. Jiang, J. Ni, *Eur. J. Inorg. Chem.*, 2003, **2003**, 4010-4016.
[49] W. Sun, E. Herdtweck, F. E. Kühn, *New J. Chem.*, 2005, **29**, 1577-1580.
[50] M. Okafor, P. Gonzalez, P. Ronot, I. El Masoudi, A. Boos, S. Ory, S. Chasserot-Golaz, S. Gasman, L. Raibaut, C. Hureau, N. Vitale, P. Faller, *Chem. Sci.*, 2022, **13**, 11829-11840.
[51] S. S. Hindo, A. M. Mancino, J. J. Braymer, Y. Liu, S. Vivekanandan, A. Ramamoorthy, M. H. Lim, *J. Am. Chem. Soc.*, 2009, **131**, 16663-16665.
[52] Y. Yang, T. Chen, S. Zhu, X. Gu, X. Jia, Y. Lu, L. Zhu, *Integr. Biol.*, 2015, **6**, 655-662.
[53] X. Hu, Q. Zhang, W. Wang, Z. Yuan, X. Zhu, B. Chen, X. Chen, *ACS Chem. Neurosci.*, 2016, **7**, 1255-1263.
[54] J.-S. Choi, J. J. Braymer, R. P. R. Nanga, A. Ramamoorthy, M. H. Lim, *Proc. Natl. Acad. Sci. USA*, 2010, **107**, 21990-21995.

Comptes Rendus
Chimie
2023, Vol. 26, Special Issue S3, p. 79-92
https://doi.org/10.5802/crchim.254

Breaking Barriers in Chemical Biology – Toulouse 2022

Synthesis and kinetic evaluation of analogs of (*E*)-4-amino-3-methylbut-2-en-1-yl diphosphate, a potent inhibitor of the IspH metalloenzyme

Benoît Eric Petit [a], Hannah Jobelius [a], Gabriella Ines Bianchino [a], Mélina Guérin [a], Franck Borel [b], Philippe Chaignon [*, a] and Myriam Seemann [*, a]

[a] Equipe Chimie Biologique et Applications Thérapeutiques, Institut de Chimie de Strasbourg, UMR 7177, Université de Strasbourg/CNRS, 4, rue Blaise Pascal, 67070 Strasbourg, France

[b] Univ. Grenoble Alpes, CEA, CNRS, IBS, F-38000 Grenoble, France

E-mails: p.chaignon@unistra.fr (P. Chaignon), mseemann@unistra.fr (M. Seemann)

Abstract. Our previous research revealed that (*E*)-4-amino-3-methylbut-2-en-1-yl diphosphate (AMBPP) is one of the best inhibitors of IspH, a [4Fe–4S]-dependent enzyme involved in the methylerythritol phosphate pathway that is a valuable target for the discovery of new antibacterial and antiparasitic drugs as it is absent in humans. AMBPP has substantial limitations for drug development due to its poor metabolic stability. Here, we investigate the replacement of the diphosphate moiety of AMBPP by more stable mimics: sulfonate, phosphonate or phosphinophosphonate. After synthesis of the derivatives, enzymatic assays demonstrated that none of these AMBPP analogs is an efficient IspH inhibitor.

Keywords. MEP pathway, IspH, LytB, Iron–sulfur clusters, Inhibitor, (*E*)-4-amino-3-methylbut-2-en-1-yl diphosphate, Anti-infectives.

Funding. Fondation Jean-Marie Lehn, the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie (Grant agreement no. 860816), Université franco-allemande, IdEx Unistra (ANR-10-IDEX-0002), SFRI-STRAT'US project (ANR-20-SFRI-0012).

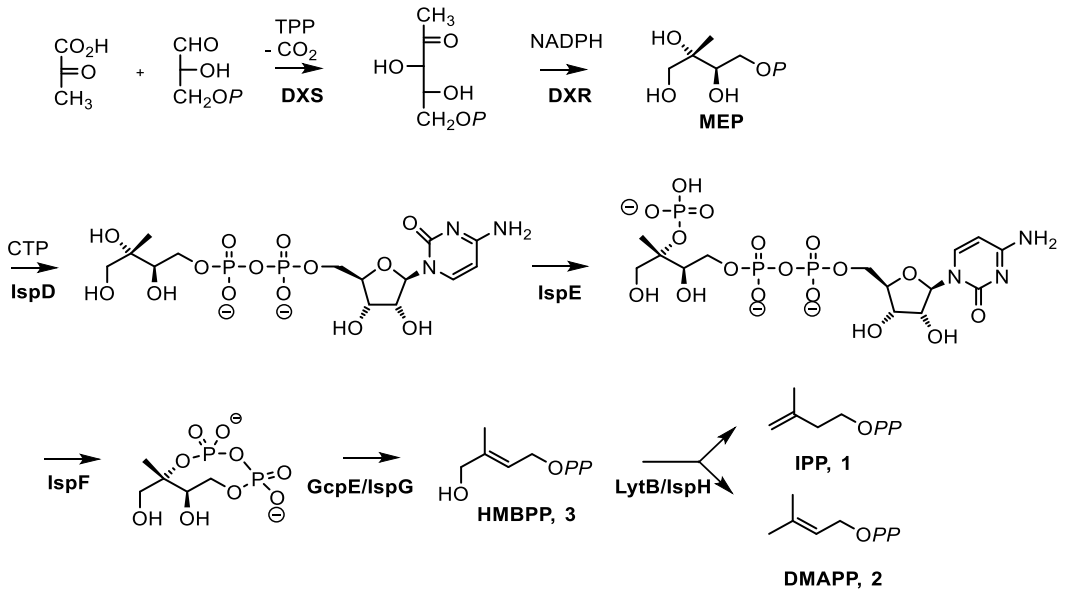
Manuscript received 7 April 2023, revised 25 June 2023, accepted 26 July 2023.

1. Introduction

Drug resistance is a constantly growing issue that poses a major challenge for the development of new drugs. In the context of antibiotic discovery, the situation is alarming, with antimicrobial resistant bacteria emerging and spreading all around the world [1]. Our ability to treat common infectious diseases is now compromised as some infections are already impossible to treat with the existing therapeutic repertoire. Antibiotic resistance not only increases mortality but also has a severe impact on medical expenses and hospitalization time.

The World Health Organization (WHO) published in 2017 a list of prioritization of bacteria to guide research and development of novel antibiotics [2]. In the latest WHO report dated December 9, 2022 [3], Dr. Tedros Adhanom Ghebreyesus, WHO Director-General, stated that "Antimicrobial resistance undermines modern medicine and puts millions of lives at risk". Given the severe menace caused by antibiotic-

* Corresponding authors



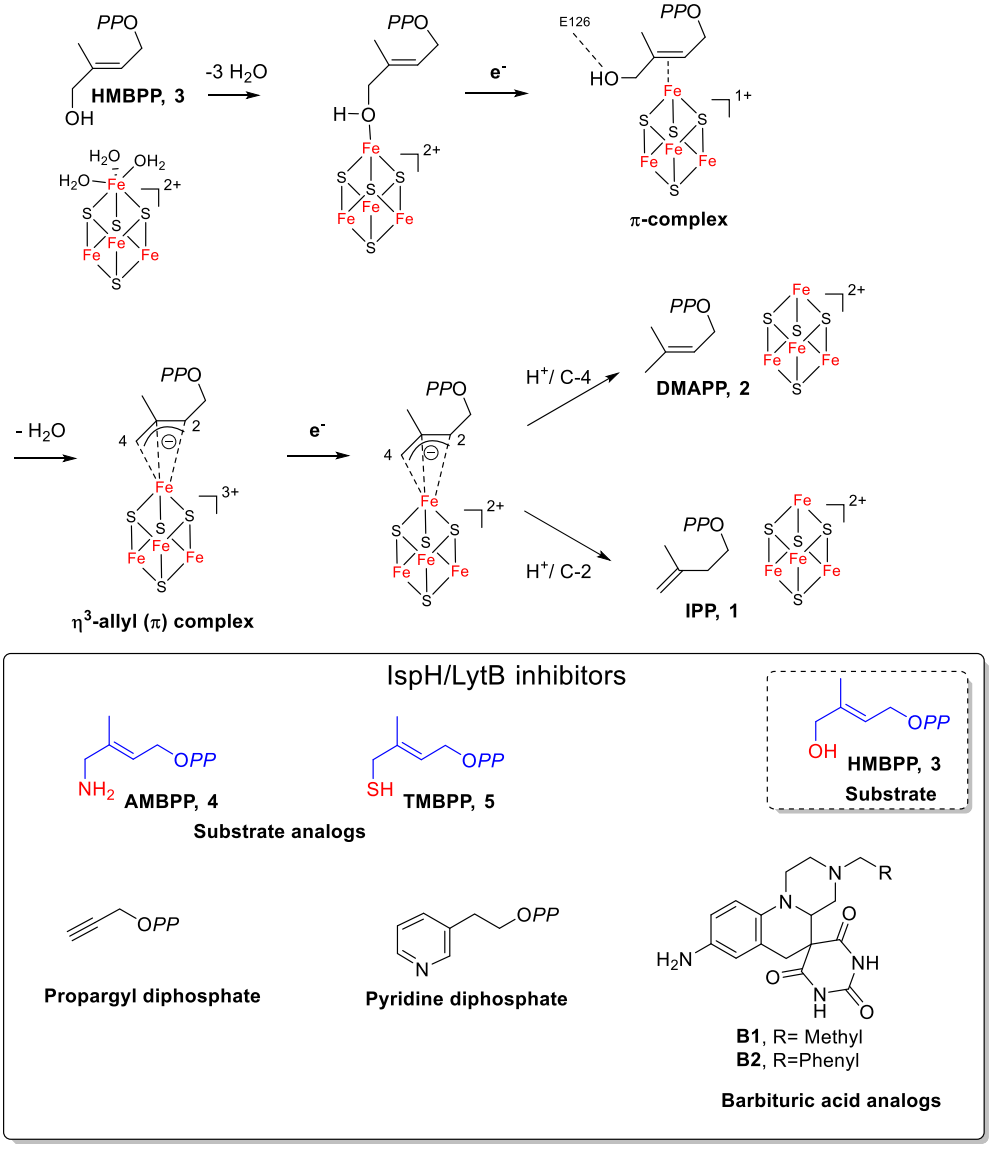
Scheme 1. The MEP pathway.

resistant bacteria, it is more than urgent to discover new antibacterial agents with new modes of action. In this context, the methylerythritol phosphate pathway (MEP, Scheme 1), responsible for the biosynthesis of the universal precursors of terpenoids in most bacteria, has emerged as an attractive target for drug development [4–7].

Terpenoids, also known as isoprenoids, represent the most diverse family of natural products, with over 55,000 known compounds. They are present in all living organisms and are involved in many important biological processes such as electron transport, cell wall biosynthesis, and protein prenylation [8,9]. Terpenoids are biosynthesized by the addition of one or more molecules of isopentenyl diphosphate (IPP, 1) to its isomer dimethylallyl diphosphate (DMAPP, 2, Scheme 1) [8]. Most pathogenic bacteria, including almost all that were prioritized by the WHO [2], use the MEP pathway (Scheme 1) for the production of IPP and DMAPP, whereas the biosynthesis of these building blocks relies exclusively on the mevalonate pathway in humans and animals [10]. This metabolic difference makes enzymes of the MEP pathway interesting targets for new antibacterial drug development that are expected to have no or reduced side effects in humans. Despite the hope sparked by the discovery of the MEP pathway, only fosmidomycin, an inhibitor of 1-deoxyxylulose-5-phosphate reductoisomerase (DXR), the second enzyme of the MEP pathway (Scheme 1), reached clinical trials as an antimalarial agent (*P. falciparum* is dependent on the MEP pathway) in combination with clindamycin and piperaquine [11]. This finding validates the MEP pathway as an innovative target for the development of new drugs [12]. However, new molecules need to enter the therapeutic pipeline to combat deadly bacterial infections.

Here, we focus our efforts on discovering new inhibitors of IspH, also called LytB, the last enzyme of the MEP pathway.

IspH contains an oxygen-sensitive $[4Fe–4S]^{2+}$ center that is essential for catalysis and converts (*E*)-4-hydroxy-3-methylbut-2-enyl diphosphate (HMBPP, **3**) into a mixture of IPP and DMAPP (Scheme 2). Mössbauer spectroscopy [13,14] and Nuclear Resonance Vibrational Spectroscopy studies highlighted that the $[4Fe–4S]^{2+}$ cluster of substrate-free IspH is particular, as one of its four iron sites is an Fe(II) atom in an octahedral coordination geometry, linked to three inorganic sulfur atoms of the iron–sulfur cluster and three water molecules (Scheme 2) [15]. This unusual Fe(II) coordination with three labile ligands is at the origin of the instability of the $[4Fe–4S]^{2+}$ cluster of IspH in the presence of oxygen. Indeed, the ox-



Scheme 2. IspH-catalyzed reaction and structures of potent inhibitors.

idation of this Fe(II) may trigger the decomposition of the prosthetic group. Consequently, IspH is only stable under anaerobic conditions, making this enzyme difficult to study and hence an underexplored target. No crystal structure of substrate-free IspH in its $[4Fe–4S]^{2+}$ form has been reported as the apical Fe(II) might dissociate during the crystallization process. The first X-ray structure of IspH described with an intact [4Fe–4S] cluster was obtained for the *E. coli* homolog in complex with HMBPP (Scheme 2) [16]. Since then, several other IspH structures harboring the [4Fe–4S] center in complex with ligands have been published (for a review see [17]).

The IspH mechanism is peculiar and involves bioinorganic and bioorganometallic intermediates (for reviews see [17–19]). It formally involves removal of the hydroxyl group, transfer of two electrons from the [4Fe–4S] cluster, and protonation of an intermediate allylic anion (Scheme 2). The binding of the OH group of HMBPP to the unique fourth iron site of the

$[4Fe–4S]^{2+}$ cluster, leading to the change in coordination geometry of this iron from octahedral to tetrahedral upon binding of the substrate (Scheme 2), was shown using Mössbauer spectroscopy [13] and is illustrated in the X-ray structure of the *E. coli* IspH–HMBPP complex [16]. After reduction of the first bioinorganic complex (Scheme 2), the OH group of the substrate undergoes a rotation to interact with E126, leading to a π-complex [20–23]. EPR/ENDOR investigations led to the characterization of a η^3-allyl (π) complex that forms after water elimination [24,25]. Further reduction of the paramagnetic η^3-allyl (π) complex followed by protonation at the *si* face of C-2 yields IPP, while protonation at C-4 yields DMAPP [26].

We and others further exploited the acquired knowledge of the IspH mechanism to design inhibitors [17]. In this context, we have already reported two molecular tools that were HMBPP analogs (Scheme 2) in which the OH group of HMBPP was replaced by an amino (AMBPP, **4**) or a thiol group (TMBPP, **5**). We had expected AMBPP and TMBPP to tightly bind to the IspH [4Fe–4S] cluster but not being capable of undergoing the elimination step. Enzymatic studies have revealed that these molecules are very potent inhibitors of IspH with K_i values in the nanomolar range: TMBPP is a tight-binding inhibitor (IC_{50} = 210 nM, K_i = 20 nM, *E. coli* IspH) and AMBPP is a slow-binding inhibitor (IC_{50} = 150 nM, K_i = 54 nM, *E. coli* IspH) [27]. These molecules remain the best IspH inhibitors known to date [17]. The mode of binding of these inhibitors to the apical iron of IspH [4Fe–4S] via the thiol or the amino function has been further confirmed [28–30]. In addition to these two inhibitors, two other potent inhibitors of *A. aeolicus* IspH were reported (Scheme 2): a propargyl diphosphate (IC_{50} = 6.7 μM) and a pyridine diphosphate (IC_{50} = 9.1 μM) [18]. However, all these inhibitors are diphosphate derivatives and therefore have substantial liabilities with respect to poor transport across bacterial membranes and inactivation upon hydrolysis by secreted phosphatases. Barbituric acid analogs, very different in structure compared to the substrate, were recently developed. They were found to be less potent IspH inhibitors (**B1**: IC_{50} = 22 μM, *P. aeruginosa* IspH; **B2**: IC_{50} = 23 μM, *E. coli* IspH, Scheme 2) [31].

All these results comforted us in our approach to optimize these substrate-based inhibitors. In this context, we report here the synthesis of analogs of AMBPP **4**, in which the diphosphate moiety was replaced by simple mimics, and the results of their biological evaluation on IspH.

2. Results and discussion

When preparing molecules harboring a diphosphate, this functional group is usually introduced at the end of the synthesis as its presence leads to a molecule almost insoluble in most common organic solvents, which limits further chemical transformations [27–34]. Moreover, the purification of diphosphorylated molecules is often tedious. Diphosphate entities are also prone to hydrolysis catalyzed by phosphatases excreted by bacteria, which would result in the inactivation of the diphosphate-containing inhibitors.

To avoid these issues, we investigated the replacement of the diphosphate moiety of the AMBPP inhibitor with more stable mimics. In this context, sulfonate, phosphonate or phosphinophosphonate were chosen. Sulfonate is used as isostere of phosphate as it has a tetrahedral shape similar to that of the phosphate group but is more acidic. The methylene phosphonate moiety is an isostere of phosphate with its phosphorus–carbon bond more stable towards hydrolysis compared to the phosphorus–oxygen bond of phosphate. These two moieties are shorter than diphosphate [35]. Finally, the phosphinophosphonate group is an isostere of the diphosphate group but less sensitive to hydrolysis. The structures of the corresponding AMBPP analogs that we investigated are displayed in Figure 1.

2.1. *Preliminary docking experiments*

Preliminary in silico docking and scoring experiments were performed using phosphonate **6**, sulfonate **7**, and phosphinophosphonate **8**. Due to the difference in pK_a values of these different entities and the fact that the activity of IspH is determined at pH = 8, docking experiments were performed at pH = 7 and at pH = 8.

Experiments were carried out starting from the X-ray structure of IspH in complex with AMBPP (PDB: 3ZGL) [30]. Interestingly, the docking scores of the sulfonate analog **7** (D.S. = −6.93; −6.92, Figure 2) are slightly better than the docking scores of the parent

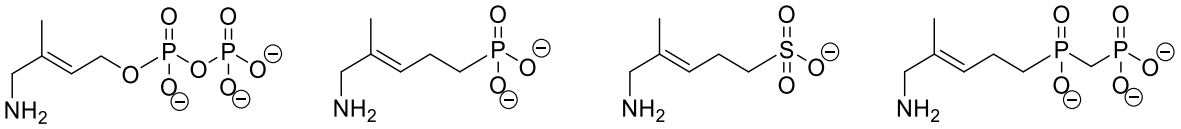


Figure 1. AMBPP and its phosphonate, sulfonate, and phosphinophosphonate analogs.

molecule AMBPP (D.S. = −6.75; −5.33, Figure 2). Replacement of the diphosphate by a simple phosphonate or a phosphinophosphonate leads to still acceptable docking scores (Figure 2). The corresponding docking poses revealed that compounds **6**, **7**, and **8** docked in the active site and placed their sulfonate, phosphonate or phosphinophosphonate group in the diphosphate binding pocket of AMBPP (Figure 2). The amino groups in these AMBPP analogs were also found close to the apical iron.

Encouraged by these indicators, compounds **6**, **7**, and **8** were prepared.

2.2. *Chemistry*

Target compounds were synthesized starting from dimethylallyl bromide **9**. The syntheses of **6** and **8** are outlined in Scheme 3.

2.2.1. *Synthesis of (E)-(5-amino-4-methylpent-3-en-1-yl) phosphonate **6***

Dimethylphosphonate **10** was synthesized by nucleophilic substitution of the bromine atom in dimethylallyl bromide **9** with deprotonated dimethyl methylphosphonate following the procedure reported by Wiemer [36]. Phosphonate **10** was selectively oxidized with selenium dioxide to yield the corresponding *E*-configurated aldehyde, which was subsequently reduced by $NaBH_4$ to alcohol **11** [37,38]. Displacement of the hydroxyl group in **11** with phthalimide under Mitsunobu conditions yielded **12**. Methyl phosphoesters and phthalimide were deprotected by treatment with bromotrimethylsilane and aqueous ammonia, respectively, to provide the corresponding amine **6**.

2.2.2. *Synthesis of (E)-(5-amino-4-methylpent-3-en-1-yl) phosphinophosphonate **8***

Phosphinophosphonate **16** could not be obtained by sequential activation of phosphonate **12** by oxalyl chloride to the corresponding phosphonic acid chloride and treatment of the latter by deprotonated dimethyl methylphosphonate. As an alternative, we used the procedure published by Wiemer and coworkers for the formation of **14** [39]. Subsequent allylic oxidation with selenium dioxide yielded **15**. Phosphinophosphonate **8** was then obtained following the same strategy used for phosphonate **6**.

The synthesis of sulfonate **7** is outlined in Scheme 4.

2.2.3. *Synthesis of (E)-5-amino-4-methylpent-3-ene-1-sulfonate **7***

The sulfonate analog **7** was prepared in three steps. Commercially available alkyl bromide **18** was selectively oxidized to alcohol **19**, as reported by Gaich and Mulzer [40]. A Mitsunobu reaction using phthalimide as the nucleophile allowed the synthesis of compound **20**. Bromide was substituted by treatment with sodium sulfite, and hydrolysis of the phthalimide by ammonia yielded **7**.

Biological experiments were further carried out to test the ability of **6**, **7**, **8** to act as IspH inhibitors.

2.3. *Biological evaluation*

E. coli IspH contains an oxygen-sensitive $[4Fe–4S]^{2+}$ cluster that is essential for catalysis and therefore needs to be handled in a glove box under a strictly inert (N_2) atmosphere. IspH converts HMBPP into a mixture of IPP and DMAPP in the presence of an external reduction system. In *E. coli*, the natural flavodoxin (FldA)/flavodoxin reductase (FpR1)/NADPH system plays this role [27,41]. Enzyme activity was therefore determined by monitoring NADPH consumption. The progress curve (Figure 3, navy blue) showed a sharp drop in the NADPH concentration in the first 7 min that was due to IspH catalysis under multiple turnover conditions. The resulting *E. coli* IspH activity was 990 $nmol{\cdot}min^{-1}{\cdot}mg^{-1}$, in agreement with previous reports [13,23,27]. It should

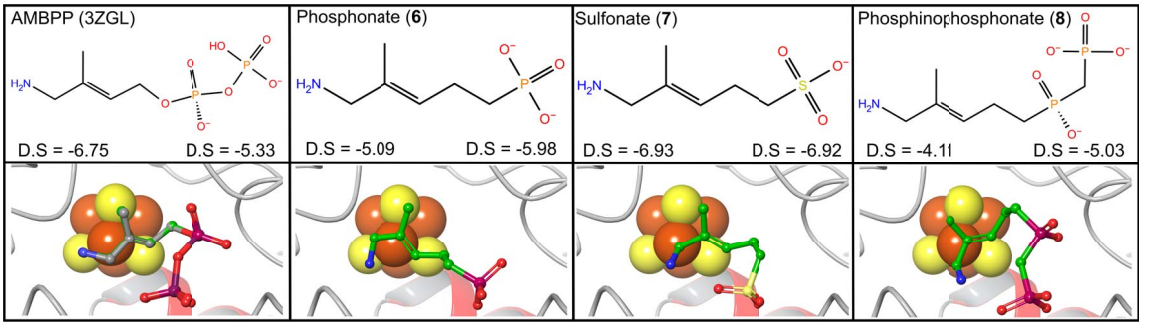
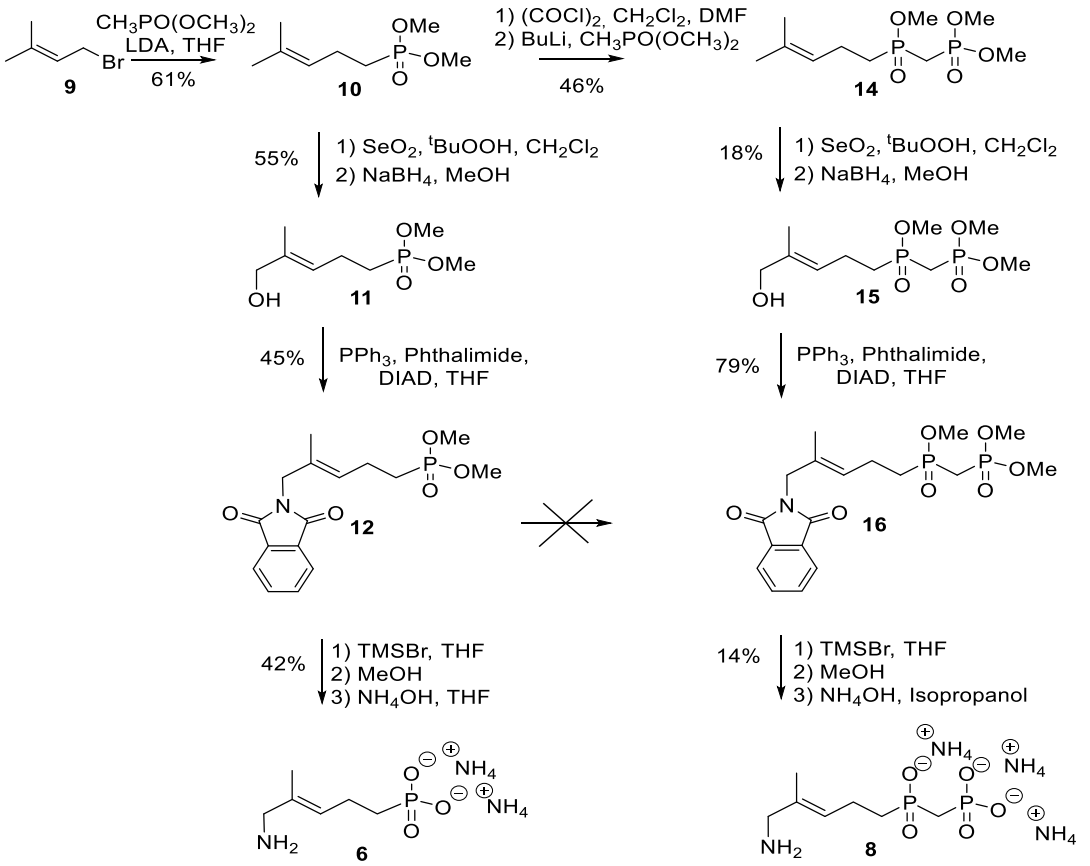


Figure 2. Docking experiments were performed using the X-ray structure of the *E. coli* IspH:AMBPP complex (PDB: 3ZGL). Upper panel: structural formula and docking score (D.S) at pH = 7 (left) and pH = 8 (right). Lower panel: docking poses of the compound in the IspH active site. Docked molecules are depicted with green carbon atoms. Docked AMBPP is superposed onto its corresponding crystallographic structure (gray carbon atoms). Iron and sulfur atoms are represented in orange and yellow, respectively.



Scheme 3. Synthesis of phosphonate **6** and phosphinophosphonate **8** analogs.

be noticed that the low-slope region observed after 7 min is due to spontaneous NADPH degradation at pH = 8.

We previously reported that AMBPP **4** is a slow-

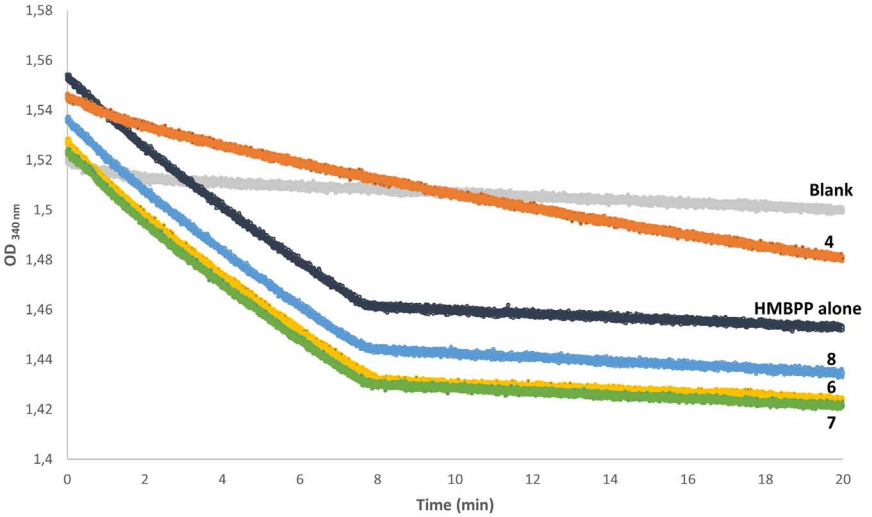
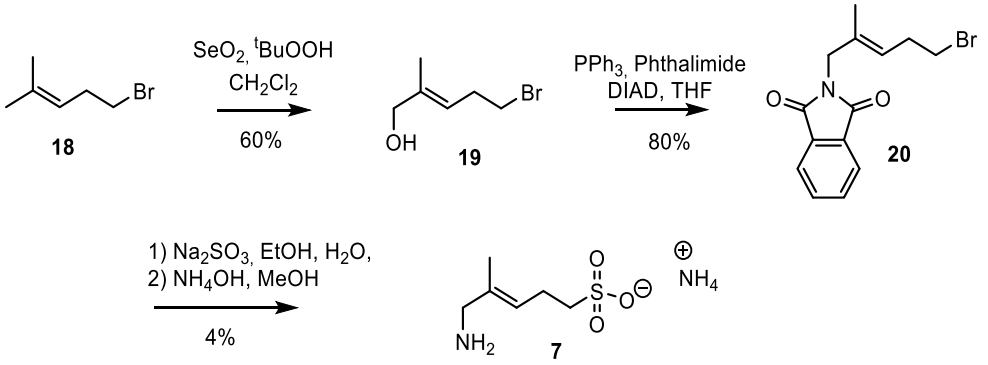


Figure 3. IspH enzymatic assays. Decrease in the absorbance of NADPH at 340 nm in the presence or absence of **6**, **7**, **8**, or AMBPP. Conditions: NADPH (2.2 mM), FldA (30 μM), FpR1 (17 μM), IspH (0.5 μM) in 50 mM Tris HCl buffer pH = 8 at 37 °C; HMBPP alone (dark blue) or in the presence of 1 mM of **6** (yellow), **7** (green), **8** (blue), or 25 μM AMBPP **4** (orange) that were preincubated with IspH for 15 min at 37 °C before initiating the reaction by addition of HMBPP (150 μM). Samples were prepared in a glove box, and additions were performed using a gastight syringe. A control sample (light gray) was prepared under the same conditions but replacing HMBPP by buffer.



Scheme 4. Synthesis of sulfonate analog **7**.

binding inhibitor of IspH under the conditions used in this assay and hypothesized that the slow-binding step might be due to the formation of the nonprotonated amine required for binding to the apical iron of the $[4Fe–4S]^{2+}$ cluster [27]. As **6**, **7**, and **8** also contain the amine function, this slow-binding behavior would also be expected for these compounds. As a consequence, we tested the ability of **6**, **7**, and **8** to promote IspH inhibition by first preincubating IspH with each of these compounds for 15 min, in order to favor the formation of the enzyme–inhibitor complex, and then initiate the IspH-catalyzed reaction by the addition of the HMBPP substrate. Progress curves recorded using **6**, **7**, or **8** at a concentration of 1 mM displayed the same initial slope as the curve recorded for IspH with HMBPP alone, indicating the same steady-state rates (Figure 3, yellow, green, blue). In contrast, the progress curve recorded under the same conditions for AMBPP at a concentration as low as 25 μM showed a drastic decrease in the IspH

reaction rate, indicative of an *E. coli* IspH activity of 263 nmol·min^{-1}·mg^{-1} that corresponds to 73% enzyme inhibition (Figure 3, orange). Together, these results reveal that replacement of the diphosphate in AMBPP with a phosphonate, a sulfonate, or a phosphinophosphonate, compromise the inhibition potential of the resulting derivatives **6**, **7**, and **8** towards IspH.

3. Conclusion

Three novel analogs of AMBPP, in which the diphosphate group was replaced by a sulfonate or a methylene phosphonate or a phosphinophosphonate, were synthesized and characterized by NMR spectroscopy and mass spectrometry. In contrast to the parent molecule AMBPP, which is one of the best two inhibitors known to date for *E. coli* IspH, a metalloenzyme containing an oxygen sensitive [4Fe–4S] cluster involved in the MEP pathway, these new molecules did not affect IspH activity. These results illustrate the essentiality of the diphosphate group of AMBPP, so far. The lack of inhibition potential of the sulfonate or methylene phosphonate analogs is most probably due to the size of these phosphate mimics that are shorter than diphosphate and might not completely fill the diphosphate binding pocket of IspH. In contrast, phosphinophosphonate has the same size as AMBPP and could undergo the same interactions as AMBPP with the surrounding amino acids of IspH. However, phosphinophosphonates are known to be less acidic than diphosphates. As a consequence, the phosphinophosphonate might retain a proton that might weaken some interactions within the active site.

Based on the knowledge gained from this study, new inhibitors derived from AMBPP or other promising IspH inhibitors need to be elaborated. Such optimization could consist in the use of other diphosphate isosteres such as difluoromethylphosphonates, difluoromethanediphosphonates or via structure-based fragment selection to find new scaffolds binding to the diphosphate pocket of IspH that do not rely on phosphate chemistry.

4. Experimental section

4.1. *Molecular docking*

In silico docking experiments were carried out using the Schrödinger suite 2020-4 (Schrödinger LLC, New York, NY, USA). The X-ray structure of IspH in complex with AMBPP (PDB: 3ZGL) was used for the studies. The protein structure was processed as previously described [30]. To generate the docking grid, we used AMBPP as the reference ligand and the protein model without water molecules. The binding region was defined by a square box centered on the inhibitor. Sizes that largely exceeded the volume of the binding site were used for both the enclosing (10 Å × 10 Å × 10 Å) and bounding box (20 Å × 20 Å × 20 Å). LigPrep was used for energy minimization, to generate the 3D structures of the compounds, and to produce the tautomers and the ionization states at pH = 7 and pH = 8. The docking study was performed using Glide's extra precision mode [42,43]. No constraints (such as hydrogen bond or atom position) were applied to guide the binding. Results of the in silico docking experiments were sorted according to the Glide docking score.

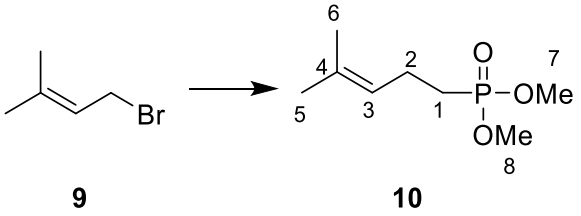
4.2. *Syntheses*

All reactions in nonaqueous solvents were conducted under an argon atmosphere with a magnetic stir bar. All reagents and solvents were purchased from commercial sources and used without further purification. Anhydrous CH_2Cl_2 and tetrahydrofuran (THF) were purchased (99.85%, water < 50 ppm). All other solvents were of HPLC grade. Reactions were monitored by thin layer chromatography (TLC) with silica gel 60-F254 plates. Flash column chromatography was performed using silica gel (0.04–0.063 mm, 230–400 mesh) under pressure. Yields refer to chromatographically and spectroscopically pure compounds.

NMR spectra were recorded on a 300- or 500-MHz spectrometer. All NMR spectra were measured in $CDCl_3$ or D_2O solutions and referenced, respectively, to the residual $CHCl_3$ signal (^{1}H, δ = 7.26 ppm; ^{13}C, δ = 77.16 ppm) or H_2O (δ = 4.79 ppm). For ^{31}P NMR spectroscopy, 85% phosphoric acid in D_2O was used as external reference (δ = −0.85 ppm). Chemical shifts and coupling constants are reported in ppm and Hz, respectively. High-resolution mass spectra were obtained using ESI-TOF.

4.2.1. *Synthesis of (E)-(5-amino-4-methylpent-3-en-1-yl) phosphonate* ***6***

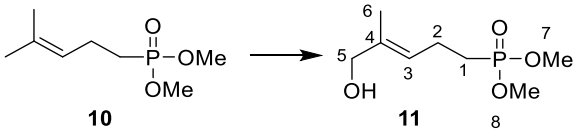
• Synthesis of dimethyl (4-methylpent-3-en-1-yl) phosphonate **10**



9 **10**

Dimethyl methylphosphonate (2.1 mL, 19.4 mmol, 1 eq.) was added dropwise at −78 °C to a stirred solution of LDA (2 M in THF, 10 mL, 20 mmol, 1 eq.) in dry THF (80 mL). The reaction mixture was stirred at −78 °C for 15 min before 1-bromo-3-methylbut-2-ene **9** (2.3 mL, 19.9 mmol, 1 eq.) was added dropwise. The reaction mixture was stirred at −78 °C for 30 min and then left to stand at 20 °C overnight. A saturated aqueous solution of NH_4Cl was then added, and the different layers were separated. The aqueous layer was extracted using diethyl ether. The combined organic layers were dried over Na_2SO_4, filtered, and the solvent was removed under reduced pressure. The crude product was purified by column chromatography on silica gel (ethyl acetate/petroleum ether, 8:2) yielding dimethyl (4-methylpent-3-en-1-yl) phosphonate **10** as an oil (2.26 g, 11.8 mmol, 61%, Rf (ethyl acetate/cyclohexane, 9:1) = 0.25). **^{1}H NMR** (500 MHz, $CDCl_3$): δ(ppm) = 1.61 (3H, s, H-6), 1.68 (3H, s, H-5), 1.70–1.82 (2H, m, H-1), 2.21–2.33 (2H, m, H-2), 3.73 (6H, d, J = 12.0 Hz, H-7 + H-8), 5.10 (1H, tq, J = 7.2 Hz, J = 1.5 Hz, H-3). **^{13}C NMR** (125 MHz, $CDCl_3$): δ(ppm) = 17.8 (C-6), 21.1 (d, J = 5.0 Hz, C-2), 25.1 (d, J = 137.0 Hz, C-1), 25.8 (C-5), 52.4 (d, J = 6.6 Hz, C-7 + C-8), 123.1 (d, J = 17.5 Hz, C-3), 133.1 (d, J = 2.5 Hz, C-4). **^{31}P NMR** (121 MHz, $CDCl_3$): δ(ppm) = 34.5.

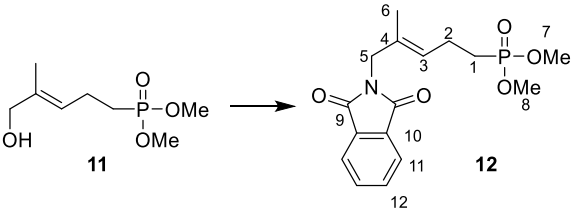
• Synthesis of dimethyl (*E*)-(5-hydroxy-4-methylpent-3-en-1-yl) phosphonate **11**



10 **11**

Dimethyl (4-methylpent-3-en-1-yl) phosphonate **10** (500 mg, 2.60 mmol, 1 eq.), SeO_2 (220 mg, 1.95 mmol, 0.75 eq.) and tert-butyl hydroperoxide (70% in water, 1.4 mL, 10.4 mmol, 4 eq.) were dissolved in CH_2Cl_2 (15 mL). The reaction mixture was stirred at 20 °C for 16 h before the reaction was quenched by the addition of a saturated aqueous solution of NaCl. The different layers were separated, and the aqueous layer was extracted with CH_2Cl_2. The combined organic layers were washed with an aqueous solution of $Na_2S_2O_3$, dried over Na_2SO_4, filtered, and the solvent was removed under reduced pressure. The resulting crude product was dissolved in methanol (7.5 mL), and $NaBH_4$ (200 mg, 5.2 mmol, 2 eq.) was added portionwise. The reaction mixture was stirred at 20 °C for 2 h before being quenched by the addition of a saturated aqueous solution of NH_4Cl. The resulting mixture was extracted with diethyl ether, and the combined organic layers were dried over Na_2SO_4, filtered, and the solvent was removed under reduced pressure. The crude product was purified by column chromatography on silica gel (MeOH/DCM, 4:96) yielding dimethyl (*E*)-(5-hydroxy-4-methylpent-3-en-1-yl) phosphonate **11** as an oil (298 mg, 1.43 mmol, 55%, Rf (ethyl acetate/cyclohexane, 9:1) = 0.11). **^{1}H NMR** (300 MHz, $CDCl_3$): δ(ppm) = 1.68 (3H, s, H-6), 1.77–1.84 (2H, m, H-1), 2.31–2.38 (2H, m, H-2), 3.74 (6H, d, J = 10.0 Hz, H-7 + H-8), 4.00 (2H, s, H-5), 5.42 (1H, tq, J = 7.0 Hz, J = 1.5 Hz, H-3). **^{13}C NMR** (125 MHz, $CDCl_3$): δ(ppm) = 13.8 (C-6), 20.8 (d, J = 5.0 Hz, C-2), 24.8 (s, J = 139.0 Hz, C-1), 52.5 (d, J = 6.28 Hz, C-7 + C-8), 68.6 (C-5), 124.2 (d, J = 16.25 Hz, C-3), 136.4 (d, J = 1.25 Hz, C-4). **^{31}P NMR** (121 MHz, $CDCl_3$): δ(ppm) = 34.29.

• Synthesis of dimethyl (*E*)-(5-(1,3-dioxoisoindolin-2-yl)-4-methylpent-3-en-1-yl) phosphonate **12**

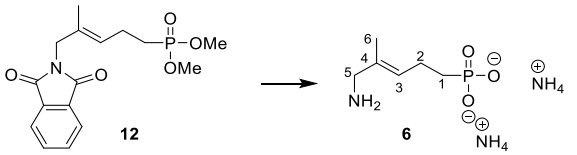


11 **12**

Dimethyl (*E*)-(5-hydroxy-4-methylpent-3-en-1-yl) phosphonate **11** (398 mg, 1.9 mmol, 1.1 eq.), PPh_3 (515 mg, 2.0 mmol, 1.1 eq.) and phthalimide (250 mg, 1.7 mmol, 1 eq.) were dissolved in dry

THF (12 mL). DIAD (430 µL, 2.0 mmol, 1.1 eq.) was added dropwise at 0 °C, and the reaction mixture was stirred at 0 °C for 30 min and then left to stand at 20 °C overnight. The reaction was quenched by the addition of MeOH (0.5 mL), and the solvent was evaporated under vacuum. The crude product was purified by column chromatography on silica gel (ethyl acetate) yielding **12** as a white solid (260 mg, 0.77 mmol, 45%, Rf (MeOH/DCM, 5:95) = 0.54). **^{1}H NMR** (300 MHz, $CDCl_3$): δ(ppm) = 1.66 (3H, s, H-6), 1.70–1.81 (2H, m, H-1), 2.24–2.36 (2H, m, H-2), 3.70 (6H, d *J* = 9.0 Hz, H-7 + H-8), 4.18 (2H, s, H-5), 5.31 (1H, tq, *J* = 8.7 Hz, *J* = 1.5 Hz, H-3), 7.70–7.74 (2H, m, H-11 or H-12), 7.81–7.86 (2H, m, H-11 or H-12). **^{13}C NMR** (125 MHz, $CDCl_3$): δ(ppm) = 14.8 (C-6), 20.9 (d, *J* = 4.5 Hz, C-2), 24.5 (d, *J* = 139.0 Hz, C-1), 44.8 (C-5), 52.4 (d, *J* = 6.5 Hz, C-7 + C-8), 123.5 (C-10), 125.7 (d, *J* = 16.9 Hz, C-3), 130.9 (d, *J* = 1.6 Hz, C-4), 132.1 (C-11 or C-12), 134.2 (C-11 or C-12), 168.3 (C-9). **^{31}P NMR** (121 MHz, $CDCl_3$): δ(ppm) = 33.94.

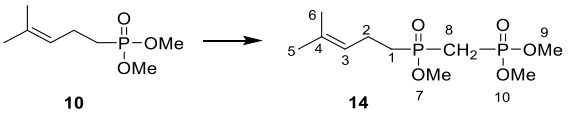
• Synthesis of (*E*)-(5-amino-4-methylpent-3-en-1-yl) phosphonate **6**



Me_3SiBr (180 µL, 1.68 mmol, 9 eq.) was added dropwise to a stirred solution of dimethyl (*E*)-(5-(1,3-dioxoisoindolin-2-yl)-4-methylpent-3-en-1-yl) phosphonate **12** (60 mg, 0.178 mmol, 1 eq.) in dry CH_2Cl_2 (1.5 mL) at 0 °C. The reaction mixture was stirred at 0 °C for 30 min and at 20 °C for 1 h. Methanol (1 mL) was added, and the reaction was further stirred for 1 h. The solvents were then evaporated under reduced pressure, and the resulting oil was dissolved in THF (4 mL). NH_4OH (1 mL) was added, and the reaction mixture was stirred overnight. Solvents were evaporated under reduced pressure, and the resulting mixture was purified by column chromatography on silica gel (Isopropanol/H_2O/NH_4OH, 7:1:2 to 5:3:2) yielding **6** as a white solid, which was dissolved in water, lyophilized, and further dried under high vacuum (13 mg, 0.074 mmol, 42%, Rf (Isopropanol/H_2O/NH_4OH, 6:2:2) = 0.20). **^{1}H NMR** (300 MHz, D_2O): δ(ppm) = 1.61–1.72 (2H, m, H-1), 1.75 (3H, s, H-6), 2.26–2.38 (2H, m, H-2), 3.53 (2H, s, H-5), 5.62 (1H, tq, *J* = 7.5 Hz, *J* = 1.5 Hz, H-3). **^{13}C NMR** (125 MHz, D_2O): δ(ppm) = 13.5 (C-6), 21.6 (d, *J* = 3.75 Hz, C-2), 27.2 (d, *J* = 132.5 Hz, C-1), 46.3 (C-5), 127.5 (C-4), 130.9 (d, *J* = 15.0 Hz, C-3). **^{31}P NMR** (121 MHz, D_2O): δ(ppm) = 25.32. High-resolution MS (ES-) *m/z*: $[M–H]^-$ ($C_6H_{13}NO_3P$) calculated 178.0638, found 178.0650.

4.2.2. *Synthesis of (E)-(((5-amino-4-methylpent-3-en-1-yl) oxidophosphoryl) methyl) phosphonate **8***

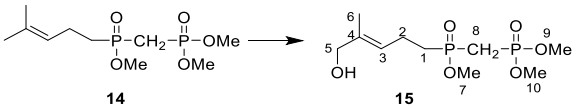
• Synthesis of dimethyl ((methoxy(4-methylpent-3-en-1-yl) phosphoryl) methyl) phosphonate **14**



Oxalyl chloride (3 mL, 35 mmol, 3 eq.) was added dropwise to a stirred solution of dimethyl (4-methylpent-3-en-1-yl) phosphonate **10** (2.05 g, 11.5 mmol, 1 eq.) and dry DMF (0.10 mL, cat.) in dry CH_2Cl_2 (60 mL) at 0 °C. The solution was stirred overnight at 20 °C. Solvents were then removed under reduced pressure, yielding a chlorinated compound, which was used without further purification. Dimethyl methylphosphonate (3.5 mL, 33 mmol, 2.9 eq.) was added dropwise to a stirred solution of BuLi (2.5 M in hexane, 12.5 mL, 31 mmol, 2.7 eq.) at −78 °C. The mixture was further stirred for 30 min before a solution of the chlorinated compound in CH_2Cl_2 (10 mL) was added dropwise. The reaction mixture was allowed to reach 20 °C while being stirred overnight. A saturated aqueous solution of NH_4Cl was then added, and the different layers were separated. The aqueous layer was extracted using CH_2Cl_2. The combined organic layers were dried over Na_2SO_4, filtered, and the solvent was removed under reduced pressure. The resulting mixture was purified by column chromatography on silica gel (EtOH/PE, 2:8) yielding **14** (1.50 g, 5.28 mmol, 46%, Rf (MeOH/DCM, 5:95) = 0.17) as yellowish oil. **^{1}H NMR** (300 MHz, $CDCl_3$): δ(ppm) = 1.63 (3H, s, H-5), 1.68 (3H, s, H-6), 1.92–2.02 (2H, m, H-1), 2.25–2.35 (2H, m, H-8), 2.33–2.46 (2H, m, H-2), 3.76 (3H, d, *J* = 11.1 Hz, H-7), 3.81 (6H, d, *J* = 11.4 Hz, H-9

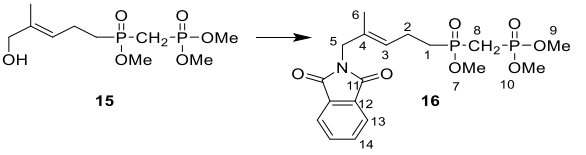
+ H-10), 5.13 (1H, tq, J = 7.2 Hz, J = 1.5 Hz, H-3). **^{13}C NMR** (125, $CDCl_3$): δ(ppm) = 17.8 (C-6), 20.4 (d, J = 4.12 Hz, C-2), 25.8 (C-5), 26.4 (dd, J = 134.2 Hz, J = 75.5 Hz, C-8), 29.5 (d, J = 96.6 Hz, C-1), 51.6 (d, J = 6.9 Hz, C-7), 53.2 (t, J = 6.7 Hz, C-9 + C-10), 122.9 (d, J = 15.6 Hz, C-3) and 133.4 (d, J = 1.4 Hz, C-4). **^{31}P NMR** (121 MHz, $CDCl_3$): δ(ppm) = 22.76 (d, J = 4.4 Hz), 48.34 (d, J = 4.5 Hz).

• Synthesis of dimethyl (*E*)-(((5-hydroxy-4-methylpent-3-en-1-yl) (methoxy)phosphoryl) methyl) phosphonate **15**



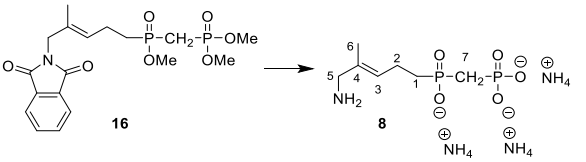
Dimethyl ((methoxy(4-methylpent-3-en-1-yl) phosphoryl) methyl) phosphonate **14** (911 mg, 3.2 mmol, 1 eq), SeO_2 (215 mg, 1.9 mmol, 0.6 eq.) and tert-butyl hydroperoxide (70% in water, 2.15 mL, 15.7 mmol, 5 eq.) were dissolved in CH_2Cl_2 (20 mL). The reaction mixture was stirred at 20 °C for 16 h before being quenched by the addition of a saturated aqueous solution of NaCl. The different layers were separated, and the aqueous layer was extracted with CH_2Cl_2. The combined organic layers were washed with an aqueous solution of $Na_2S_2O_3$, dried over Na_2SO_4, filtered, and the solvent was removed under reduced pressure. The resulting crude product was dissolved in methanol, and $NaBH_4$ (150 mg, 3.9 mmol, 1.2 eq.) was added at 0 °C. The reaction mixture was stirred at room temperature for 2 h before being quenched by the addition of a saturated aqueous solution of NH_4Cl (2 mL). The solvents were removed under reduced pressure, and the resulting mixture was purified over column chromatography on silica gel (MeOH/CH_2Cl_2, 8:92) yielding **15** as an oil (170 mg, 0.56 mmol, 18%, Rf (MeOH/DCM, 8:92) = 0.16). **^{1}H NMR** (500 MHz, $CDCl_3$): δ(ppm) = 1.69 (3H, s, H-6), 1.73 (bs, 1H, OH), 2.00–2.06 (2H, m, H-1), 2.35–2.45 (4H, m, H-2 + H-8), 3.77 (3H, d, J = 11.0 Hz, H-7), 3.81 (6H, dd, J = 11.0 Hz, J = 2.0 Hz, H-9 + H-10), 4.00 (2H, s, H-5), 5.45 (1H, tq, J = 7.5 Hz, J = 1.5 Hz, H-3). **^{13}C NMR** (125, $CDCl_3$): δ(ppm) = 13.7 (C-6), 19.9 (d, J = 4.2 Hz, C-2), 26.4 (dd, J = 134.5 Hz, J = 75.8 Hz, C-8), 29.0 (d, J = 97.6 Hz, C-1), 51.5 (d, J = 6.7 Hz, C-7), 53.1 (d, J = 6.5 Hz, C-9 + C-10), 68.4 (C-5), 123.8 (d, J = 14.1 Hz, C-3), 136.6 (d, J = 1.0 Hz, C-4). **^{31}P NMR** (121 MHz, $CDCl_3$): δ(ppm) = 23.07 (d, J = 4.2 Hz), 48.26 (d, J = 4.0 Hz).

• Synthesis of dimethyl (*E*)-(((5-(1,3-dioxoisoindolin-2-yl)-4-methylpent-3-en-1-yl) (methoxy)phosphoryl) methyl) phosphonate **16**



Dimethyl (*E*)-(((5-hydroxy-4-methylpent-3-en-1-yl) (methoxy)phosphoryl)methyl)phosphonate **15** (170 mg, 0.57 mmol, 1.2 eq.), PPh_3 (155 mg, 0.59 mmol, 1.2 eq.) and phthalimide (70 mg, 0.48 mmol, 1 eq.) were dissolved in dry THF. DIAD (130 µL, 0.64 mmol, 1.3 eq.) was added dropwise at 0 °C, and the reaction mixture was stirred at 0 °C for 30 min and left to stand at 20 °C for 2 h. The reaction was quenched by the addition of MeOH (500 µL), and the solvent was removed under reduced pressure. The crude product was purified by column chromatography on silica gel (MeOH/CH_2Cl_2, 8:92). A second column chromatography on silica gel (ethyl acetate) yielded **16** as a white oil (162 mg, 0.38 mmol, 79%, Rf (ethyl acetate) = 0.43). **^{1}H NMR** (500 MHz, $CDCl_3$): δ(ppm) = 1.67 (3H, s, H-6), 1.93–2.00 (2H, m, H-1), 2.30–2.42 (4H, m, H-2 + H-8), 3.73 (3H, d, J = 11 Hz, H-7), 3.78 (6H, d, J = 11 Hz, H-9 + H-10), 4.17 (2H, s, H-5), 5.35 (1H, tq, J = 7.5 Hz, J = 1.5 Hz, H-3), 7.70–7.72 (2H, m, H-13), 7.83–7.84 (2H, m, H-14). **^{13}C NMR** (125 MHz, $CDCl_3$): δ(ppm) = 14.8 (C-6), 20.1 (d, J = 16 Hz, C-2), 26.4 (dd, J = 134.1 Hz, J = 76.1 Hz, C-8), 29.0 (d, J = 97.4, C-1), 44.7 (C-5), 51.6 (d, J = 6.3 Hz, C-7), 53.2 (dd, J = 6.5 Hz, J = 3.1 Hz, C-9 + C-10), 123.4 (C-13), 125.5 (d, J = 15.5 Hz, C-3), 131.2 (d, J = 1.4 Hz, C-4), 132.1 (C-14), 134.1 (C-12), 168.3 (C-11). **^{31}P NMR** (121 MHz, $CDCl_3$): δ(ppm) = 22.96 (d, J = 4.0 Hz), 47.97 (d, J = 4.0 Hz).

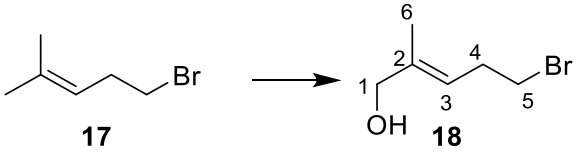
• (*E*)-(((5-amino-4-methylpent-3-en-1-yl) oxidophosphoryl) methyl) phosphonate **8**



Bromotrimethylsilane (340 µL, 3.2 mmol, 9.4 eq.) was added dropwise to a solution of dimethyl (*E*)-(((5-(1,3-dioxoisoindolin-2-yl)-4-methylpent-3-en-1-yl) (methoxy)phosphoryl) methyl) phosphonate **16** (146 mg, 0.34 mmol, 1 eq.) in dry CH_2Cl_2 (3 mL) while being stirred at 0 °C. The reaction mixture was further stirred at 0 °C for 30 min and then at 20 °C for 1 h. Methanol was added, and the mixture was further stirred for 1 h. Solvents were then removed under reduced pressure, and the resulting oil was dissolved in acetone. A mixture of isopropanol, water and ammonia (6/2/2) was added, and the mixture was stirred overnight. After evaporation of the solvents, the crude product was purified by column chromatography on silica gel (IPA/H_2O/NH_4OH, 6:2:2 to 5:3:2) yielding **8** as a white solid, which was dissolved in water, lyophilized, and further dried under high vacuum yielding **17** (12 mg, 0.046 mmol, 14%, Rf (IPA/H_2O/NH_4OH, 6:4:2) = 0.25). **^{1}H NMR** (300 MHz, D_2O): δ(ppm) = 1.69 (3H, s, H-6), 1.70–1.76 (2H, m, H-1), 1.98–2.06 (2H, m, H-7), 2.23–2.30 (2H, m, H-2), 3.46 (2H, s, H-5), 5.54 (1H, t, *J* = 7.0 Hz, H-3). **^{13}C NMR** (125 MHz, D_2O): δ(ppm) = 13.6 (C-6), 20.4 (d, *J* = 3.6 Hz, C-2), 29.9 (d, *J* = 95.0 Hz, C-1), 30.7 (dd, *J* = 119.8 Hz, *J* = 75.9 Hz, C-7), 46.2 (C-5), 127.5 (C-4), 130.8 (d, *J* = 14.4 Hz, C-3). **^{31}P NMR** (121 MHz, D_2O): δ(ppm) = 15.40 (d, *J* = 5.5 Hz), 38.13 (d, *J* = 5.5 Hz). High-resolution MS (ES-) *m/z*: $[M–H]^-$ ($C_7H_{16}NO_5P_2$) calculated 256.0509, found 256.0526.

4.2.3. *Synthesis of (E)-5-amino-4-methylpent-3-ene-1-sulfonate* ***7***

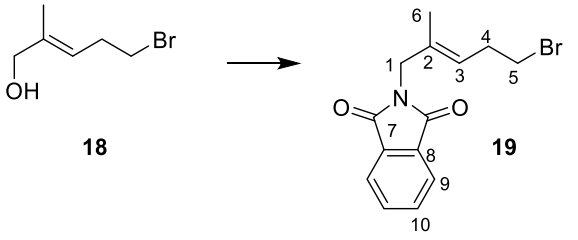
• (*E*)-5-bromo-2-methylpent-2-en-1-ol **18**



17 **18**

Tert-butyl hydroperoxide (70% in water, 1.7 mL, 12.6 mmol, 2 eq.) was added to a suspension of SeO_2 (340 mg, 3.01 mmol, 0.5 eq.) stirred at 0 °C. The reaction mixture was stirred at 0 °C for 5 min and at 20 °C for 30 min. 5-Bromo-2-methylpent-2-ene **17** (820 µL, 6.13 mmol, 1 eq.) was added dropwise at 0 °C. The reaction mixture was stirred overnight while slowly being allowed to warm up to 20 °C. It was then diluted with diethyl ether and washed twice with an aqueous solution of KOH (1M) then brine. The organic layer was dried over Na_2SO_4, filtered, and the solvent was removed under reduced pressure. The crude product was purified by column chromatography on silica gel (ethyl acetate/petroleum ether, 2:8) yielding **18** (660 mg, 3.69 mmol, 60%). **^{1}H NMR** (500 MHz, $CDCl_3$): δ(ppm) = 1.69 (3H, s, H-6), 2.63 (2H, dt, *J* = 7.0 Hz, 6.5 Hz, H-4), 3.39 (2H, t, *J* = 7.0 Hz, H-5), 4.03 (2H, s, H-1), 5.44 (1H, m, H-3). **^{13}C NMR** (125 MHz, $CDCl_3$): δ(ppm) = 14.0 (C-6), 31.3 (C-4), 32.6 (C-5), 68.5 (C-1), 122.1 (C-3), 138.1 (C-2).

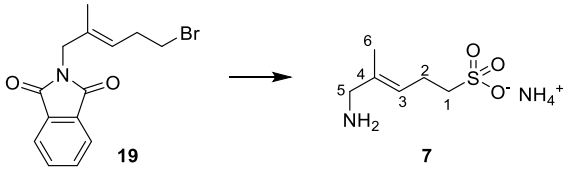
• (*E*)-2-(5-bromo-2-methylpent-2-en-1-yl) isoindoline-1,3-dione **19**



18 **19**

DIAD (400 µL, 2.0 mmol, 1.2 eq.) was added to a solution of phthalimide (300 mg, 2.0 mmol, 1.2 eq.), PPh_3 (530 mg, 2.0 mmol, 1.2 eq.) and **18** (300 mg, 1.68 mmol, 1 eq.) in dry THF while stirring at 0 °C. The reaction mixture was stirred for 3 h and simultaneously allowed to warm up to 20 °C. The reaction was quenched by the addition of MeOH (1 mL), and the solvent was removed under reduced pressure. The crude product was purified by column chromatography on silica gel (ethyl acetate/petroleum ether 1:9) yielding **19** as an oil (415 mg, 1.35 mmol, 80%, Rf (ethyl acetate/cyclohexane, 20:80) = 0.60). **^{1}H NMR** (300 MHz, $CDCl_3$): δ(ppm) = 1.68 (3H, s, H-6), 2.59 (2H, dt, *J* = 7.2 Hz, 6.3 Hz, H-4), 3.33 (2H, t, *J* = 7.2 Hz, H-5), 4.22 (2H, s, H-1), 5.37 (1H, m, H-3), 7.69–7.76 (2H, m, H-9), 7.83–7.89 (2H, m, H-10). **^{13}C NMR** (125 MHz, $CDCl_3$): δ(ppm) = 15.0 (C-6), 31.5 (C-4), 32.0 (C-5), 44.8 (C-1), 123.5 (C-9), 124.1 (C-3), 132.1 (C-8) 132.7 (C-2), 134.2 (C-10), 168.3 (C-7).

• (*E*)-5-amino-4-methylpent-3-ene-1-sulfonate **7**



19 **7**

A solution of sodium sulfite (360 mg, 2.86 mmol, 2.2 eq.) in water (5 mL) was added to a stirred solution of **19** (400 mg, 1.30 mmol, 1 eq.) in EtOH (5 mL). The reaction mixture was further stirred under reflux overnight. The suspension was filtered, and the filtrate was evaporated under reduced pressure. Nonpolar products were discarded by filtration on silica gel ($MeOH/CH_2Cl_2/NEt_3$, 2/8/0.2), and the most polar product was recovered and then dissolved in MeOH (2 mL) without further purification. NH_4OH (2 mL) was added, and the reaction mixture was stirred overnight. Solvents were removed under reduced pressure. Consecutive column chromatography (isopropanol/H_2O/NH_4OH, 8:0.8:2 then isopropanol/H_2O, 9:1) yielded **7** as a white solid, which was dissolved in water, lyophilized, and further dried under high vacuum (10 mg, 0.056 mmol, 4%, Rf (IPA/H_2O, 8:2) = 0.27). **^{1}H NMR** (300 MHz, D_2O): δ(ppm) = 1.76 (3H, s, H-6), 2.54 (2H, dt, J = 7.2 Hz, J = 6.2 Hz, H-2), 2.98 (2H, t, J = 7.8 Hz, H-1), 3.55 (2H, s, H-5), 5.60 (1H, m, H-3). **^{13}C NMR** (125 MHz, D_2O): δ(ppm) = 14.3 (C-6), 23.5 (C-2), 46.8 (C-1), 50.6 (C-5), 128.4 (C-3), 129.9 (C-4). High-resolution MS (ES-) *m/z*: $[M–H]^-$ ($C_6H_{12}NO_3S$) calculated 178.0543, found 178.0541.

4.3. *Biological experiments*

4.3.1. *IspH production*

Production and purification of *E. coli* IspH were performed as previously described [30]. Protein concentration was measured using the Bradford method with bovine serum albumin as a standard [44]. Iron was quantified according to Fish [45] and sulfide as described by Beinert [46]. UV/visible spectrum, iron and sulfur content of *E. coli* IspH were similar to those previously reported [13], in accordance with the presence of $[4Fe–4S]^{2+}$ cluster.

4.3.2. *Enzymatic assays*

Enzyme activity was determined by monitoring NADPH consumption in the presence of optimized concentrations of the reducing system under anaerobic conditions. A HMBPP solution (final concentration 150 μM) was added through a gas-tight syringe to a 0.1 cm light path cuvette prepared in an anaerobic glove box and containing NADPH (2.2 mM), FldA (30 μM), FpR1 (17 μM), IspH (0.5 μM), and either **6** or **7** or **8** or AMBPP in 50 mM Tris–HCl pH = 8 that had previously been incubated for 15 min at 37 °C. The reaction was monitored spectrophotometrically at 340 nm with a Cary 100 UV/visible spectrophotometer (Varian) maintained at 37 °C using a thermostat equipped with a Peltier element.

Declaration of interests

The authors do not work for, advise, own shares in, or receive funds from any organization that could benefit from this article, and have declared no affiliations other than their research organizations.

Funding

This work was funded by the Fondation Jean-Marie Lehn, the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No. 860816 and the "Université franco-allemande". This work of the Interdisciplinary Thematic Institute InnoVec, as part of the ITI program of the University of Strasbourg, CNRS and Inserm, was supported by IdEx Unistra (ANR-10-IDEX-0002) and by SFRI-STRAT'US project (ANR-20-SFRI-0012) under the framework of the French Investments for the Future Program.

Acknowledgments

We appreciate the help from the staff of the computing facility provided by the Commissariat à l'Energie Atomique et aux énergies renouvelables (CEA/DSV/GIPSI), Saclay, France, and the Centre de Calcul Recherche et Technologie (CEA/CCRT), Bruyères-le-Châtel, France.

Supplementary data

Supporting information for this article is available on the journal's website under https://doi.org/10.5802/crchim.254 or from the author.

References

[1] *Antimicrobial Resistance. Global Report on Surveillance*, World Health Organization, Geneva, Switzerland, 2014, https://www.who.int/publications/i/item/9789241564748.

[2] *Prioritization of Pathogens to Guide Discovery, Research and Development of New Antibiotics for Drug Resistant Bacterial Infections, Including Tuberculosis. Report No. WHO/EMP/IAU/2017.12*, World Health Organization, Geneva, Switzerland, 2017, https://www.who.int/publications/i/item/WHO-EMP-IAU-2017.12.

[3] *Report Signals Increasing Resistance to Antibiotics in Bacterial Infections in Humans and Need for Better Data*, World Health Organization, Geneva, Switzerland, 2022, https://www.who.int/news/item/09-12-2022-report-signals-increasing-resistance-to-antibiotics-in-bacterial-infections-in-humans-and-need-for-better-data.

[4] M. Rohmer, C. Grosdemange-Billiard, M. Seemann, D. Tritsch, *Curr. Opin. Investig. Drugs*, 2004, **5**, 154-162.

[5] T. Gräwert, M. Groll, A. Bacher, W. Eisenreich, *Cell Mol. Life Sci.*, 2011, **68**, 3797-3814.

[6] L. Zhao, W.-C. Chen, Y. Xiao, H.-W. Liu, P. Liu, *Annu. Rev. Biochem.*, 2013, **82**, 497-530.

[7] T. Masini, A. K. H. Hirsch, *J. Med. Chem.*, 2014, **57**, 9740-9763.

[8] M. Rohmer, *Nat. Prod. Rep*, 1999, **16**, 565-574.

[9] S. Heuston, M. Begley, C. G. M. Gahan, C. Hill, *Microbiology*, 2012, **158**, 1389-1401.

[10] K. Bloch, *Steroids*, 1992, **57**, 378-383.

[11] G. Mombo-Ngoma, J. Remppis, M. Sievers, R. Zoleko Manego, L. Endamne, L. Kabwende, L. Veletzky, T. T. Nguyen, M. Groger, F. Lötsch, J. Mischlinger, L. Flohr, J. Kim, C. Cattaneo, D. Hutchinson, S. Duparc, J. Moehrle, T. P. Velavan, B. Lell, M. Ramharter, A. A. Adegnika, B. Mordmüller, P. G. Kremsner, *Clin. Infect. Dis.*, 2018, **66**, 1823-1830.

[12] T. N. C. Wells, R. Hooft Van Huijsduijnen, W. C. Van Voorhis, *Nat. Rev. Drug. Discov.*, 2015, **14**, 424-442.

[13] M. Seemann, K. Janthawornpong, J. Schweizer, L. H. Böttger, A. Janoschka, A. Ahrens-Botzong, E. Ngouamegne Tambou, O. Rotthaus, A. X. Trautwein, M. Rohmer, V. J. Schünemann, *J. Am. Chem. Soc.*, 2009, **131**, 13184-13185.

[14] Y. Xiao, L. Chu, Y. Sanakis, P. Liu, *J. Am. Chem. Soc.*, 2009, **131**, 9931-9933.

[15] I. Faus, A. Reinhard, S. Rackwitz, J. A. Wolny, K. Schlage, H.-C. Wille, A. Chumakov, S. Krasutsky, P. Chaignon, C. D. Poulter, M. Seemann, V. Schünemann, *Angew. Chem. Int. Ed.*, 2015, **54**, 12584-12587.

[16] T. Grawert, I. Span, W. Eisenreich, F. Rohdich, J. Eppinger, A. Bacher, M. Groll, *Proc. Natl. Acad. Sci. USA*, 2010, **107**, 1077-1081.

[17] H. Jobelius, G. I. Bianchino, F. Borel, P. Chaignon, M. Seemann, *Molecules*, 2022, **27**, article no. 708.

[18] W. Wang, E. Oldfield, *Angew. Chem. Int. Ed.*, 2014, **53**, 4294-4310.

[19] T. Gräwert, I. Span, A. Bacher, M. Groll, *M. Angew. Chem. Int. Ed.*, 2010, **49**, 8802-8809.

[20] W. Wang, K. Wang, Y. L. Liu, J. H. No, J. Li, M. J. Nilges, E. Oldfield, *Proc. Natl. Acad. Sci. USA*, 2010, **107**, 4522-4527.

[21] C. A. Citron, N. L. Brock, P. Rabe, J. S. Dickschat, *Angew. Chem. Int. Ed.*, 2012, **51**, 4053-4057.

[22] I. Span, T. Grawert, A. Bacher, W. Eisenreich, *M. Groll. J. Mol. Biol.*, 2012, **416**, 1-9.

[23] P. Chaignon, B. E. Petit, B. Vincent, L. Allouche, M. Seemann, *Chem. Eur. J.*, 2020, **26**, 1032-1036.

[24] W. Wang, K. Wang, I. Span, J. Jauch, A. Bacher, M. Groll, E. Oldfield, *J. Am. Chem. Soc.*, 2012, **134**, 11225-11234.

[25] W. Xu, N. S. Lees, D. Hall, D. Welideniya, B. M. Hoffman, E. C. Duin, *Biochemistry*, 2012, **51**, 4835-4849.

[26] R. Laupitz, T. Gräwert, C. Rieder, F. Zepeck, A. Bacher, D. Arigoni, F. Rohdich, W. Eisenreich, *Chem. Biodivers.*, 2004, **1**, 1367-1376.

[27] K. Janthawornpong, S. Krasutsky, P. Chaignon, M. Rohmer, C. D. Poulter, M. Seemann. J, *Am. Chem. Soc.*, 2013, **135**, 1816-1822.

[28] A. Ahrens-Botzong, K. Janthawornpong, J. A. Wolny, E. N. Tambou, M. Rohmer, S. Krasutsky, C. D. Poulter, V. Schünemann, M. Seemann, *Angew. Chem. Int. Ed.*, 2011, **50**, 11976-11979.

[29] I. Span, K. Wang, W. Wang, J. Jauch, W. Eisenreich, A. Bacher, E. Oldfield, M. Groll, *Angew. Chem. Int. Ed.*, 2013, **52**, 2118-2121.

[30] F. Borel, E. Barbier, S. Krasutsky, K. Janthawornpong, P. Chaignon, C. D. Poulter, J. L. Ferrer, M. Seemann, *ChemBioChem*, 2017, **18**, 2137-2144.

[31] B. O'Dowd, S. Williams, H. Wang, J. H. No, G. Rao, W. Wang, J. A. McCammon, S. P. Cramer, E. Oldfield, *ChemBioChem*, 2017, **18**, 914-920.

[32] N. A. Heaps, C. D. Poulter, *J. Org. Chem.*, 2011, **76**, 1838-1843.

[33] V. J. Davisson, A. B. Woodside, T. R. Neal, K. E. Stremler, M. Muehlbacher, C. D. Poulter, *J. Org. Chem.*, 1986, **51**, 4768-4779.

[34] F. Cramer, W. Böhm, *Angew. Chem.*, 1959, **71**, 775.

[35] T. S. Elliott, A. Slowey, Y. Ye, S. J. Conway, *Med. Chem. Commun.*, 2012, **3**, 735-751.

[36] C. H. Hsiao, X. Lin, R. J. Barney, R. R. Shippy, J. Li, O. Vinogradov, D. F. Wiemer, A. J. Wiemer, *Chem. Biol.*, 2014, **21**, 945-954.

[37] D. Bartee, M. J. Wheadon, C. L. Freel Meyers, *J. Org. Chem.*, 2018, **83**, 9580-9591.

[38] B. J. Foust, M. M. Poe, N. A. Lentini, C. H. Hsiao, A. J. Wiemer, D. F. Wiemer, *ACS Med. Chem. Lett.*, 2017, **8**, 914-918.

[39] R. R. Shippy, X. Lin, S. S. Agabiti, J. Li, B. M. Zangari, B. J. Foust, M. M. Poe, C.-H. C. Hsiao, O. Vinogradova, D. F. Wiemer, A. J. Wiemer, *J. Med. Chem.*, 2017, **60**, 2373-2382.

[40] T. Gaich, J. Mulzer, *Org. Lett.*, 2010, **12**, 272-275.

[41] M. Seemann, M. Rohmer, *C. R. Chim.*, 2007, **10**, 748-755.

[42] R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis, P. S. Shenkin, *J. Med. Chem.*, 2004, **47**, 1739-1749.

[43] R. A. Friesner, R. B. Murphy, M. P. Repasky, L. L. Frye, J. R. Greenwood, T. A. Halgren, P. C. Sanschagrin, D. T. Mainz, *J. Med. Chem.*, 2006, **49**, 6177-6196.

[44] M. M. Bradford, *Anal. Biochem.*, 1976, **72**, 248-254.

[45] W. W. Fish, *Methods Enzymol.*, 1988, **158**, 357-364.

[46] H. Beinert, *Anal. Biochem.*, 1983, **131**, 373-378.

Comptes Rendus
Chimie
2023, Vol. 26, Special Issue S3, p. 93-103
https://doi.org/10.5802/crchim.252

Breaking Barriers in Chemical Biology – Toulouse 2022

Induced-volatolomics, a new research field in chemical biology

Estelle Blochouse [a], Rony Eid [a], Elsa Cannoni [a], Rémi Châtre [a], Claude Geffroy [a], Isabelle Opalinski [a], Sébastien Papot [a,b] and Pauline Poinot [*,a]

[a] University of Poitiers, UMR CNRS 7285, Institut de Chimie des Milieux et Matériaux de Poitiers (IC2MP), Equipe Labellisée Ligue Contre le Cancer, 4 rue Michel-Brunet, TSA 51106, 86073 Poitiers cedex 9, France

[b] Seekyo SA, 2 avenue Galilée, BP 30153, 86961 Futuroscope, France

E-mails: estelle.blochouse@univ-poitiers.fr (E. Blochouse), rony.eid@univ-poitiers.fr (R. Eid), elsa.cannoni@univ-poitiers.fr (E. Cannoni), remi.chatre@univ-poitiers.fr (R. Châtre), claude.geffroy@univ-poitiers.fr (C. Geffroy), isabelle.opalinski@univ-poitiers.fr (I. Opalinski), sebastien.papot@univ-poitiers.fr (S. Papot), pauline.poinot@univ-poitiers.fr (P. Poinot)

Abstract. Progresses in chemical biology have substantially deepened our understanding of biological pathways related to disease outcomes. Numerous optical probes activatable by disease-relevant enzymes have been designed and have shown great value for imaging crucial biochemical transformations specific to pathological processes. In comparison to their "always-on" counterparts, such stimuli-responsive probes stand out as promising candidates for exploring biological processes because of their high sensitivity, ingenious spatiotemporal resolution, ease of operation, and real-time and *in situ* imaging capacity. However, these imaging tools meet several bottlenecks related to either stability, background noise, limited penetration depth or systemic toxicity, pharmacokinetics, tissue distribution, and renal clearance.

In parallel, novel volatile organic compound (VOC)-based probes have been proposed under the emerging paradigm of induced-volatolomics. These probes are converted into exogenous VOCs in response to pathogen or eukaryote specific enzymatic stimuli. Once activated, VOC-based probes release volatile tracers that inform the biochemical processes arising in cells, liquid and solid biopsies, and even in entire organisms. After a brief presentation of recent developments in the field of induced-volatolomics, we discuss possible improvements of this new modality for the exploration of biological systems.

Keywords. VOC, Induced-volatolomics, Enzyme-responsive probes, Diagnosis, Chemical biology.

Funding. Local committees Vienne and Deux-Sèvres of the "Ligue Contre le Cancer" and "La Ligue nationale contre le Cancer". "Région Nouvelle Aquitaine". ANR through the Young Researcher grant "Volatolomix" (ANR-21-CE18-0009). CNRS through the "AAP Emergence" and FEDER.

Manuscript received 10 February 2023, revised 17 July 2023, accepted 26 July 2023.

[*] Corresponding author

1. Introduction

Each of us continuously emits more than a thousand volatile organic compounds (VOCs) that are responsible for our odor fingerprint. With the emergence of volatolomics, i.e., the discipline that investigates VOCs produced by living systems, the human volatolome has provided new insights into biological processes in real-time [1] (Figure 1).

Indeed, as evidenced by canine olfaction-based studies, volatile molecules can serve as chemical tracers for assessing health status in real time in a non-invasive way [2,4]. Along with the progress in analytical technologies, many research groups have attempted to find the perfect volatile marker for a given disease (Figure 1). Hence, viral (*Influenza* viruses, adenovirus, rhinovirus [5,6]), bacterial, parasitic, and fungal infections (tuberculosis [7], malaria [8], pneumonia [9,10]) were shown to be associated with the release of specific VOCs. Similarly, respiratory disorders, such as asthma [11], or acute respiratory distress syndrome [12] as well as metabolic disorders such as diabetes [13] could be diagnosed by the detection of alkanes and ketones in patients' breath. Similarly, a strong correlation was reported between cancers, Parkinson's and Alzheimer's diseases and the airborne release of hydrocarbons, alcohols, aldehydes, ketones, esters, nitriles, and aromatic compounds [14–20].

Despite these extensive studies, no endogenous volatile compound has been proven to be of sufficient diagnostic value so far. Such an issue is mainly explained by both interindividual variability [21,22] and individual daily variation [23]. Indeed, variables such as overall health status, diet, medication intake, exercise, and smoking habits induce qualitative and quantitative changes in patient volatolome that are unrelated to pathologies [24,25]. Moreover, although several studies have sought to demonstrate the biosynthesis patterns of endogenous VOCs, their metabolic pathways remain enigmatic, which adds to confusion [26–28]. Finally, given the very low concentration of VOCs in body fluids and gases (e.g., 1×10^{-11} M in blood and 0.02 ppb in breath [24]), more efficient sampling techniques and analysers are necessary to obtain reliable quantification [29–33]. In this respect, identification of rotational fingerprints with high accuracy, sensitivity, and selectivity is a challenge that can be addressed and automated using artificial intelligence and supervised or unsupervised machine learning algorithms. In very recent works, deep learning models were developed to provide breath volatolomics-based classifiers for the presence or stage different diseases [34–38]. These cutting-edge approaches demonstrate that machine learning methods could be the answer to make volatolomics a non-invasive clinical diagnostic tool for point-of-care applications [39].

Although volatolomics was booming in the early 2000s, another strategy, which relies on the monitoring of exogenous volatiles, was emerging. This novel paradigm, very recently called "induced-volatolomics" (Figure 1), is part of the chemical biology toolbox. It implies the use of off/on probes that can be converted into a volatile compound in response to a pathogen or metabolic-specific enzymatic stimulus [40–42]. In the absence of the activating process, the probe remains inactive. In the presence of the biochemical stimulus, the probe is turned on, thus releasing an exogenous volatile molecule. This targeted action mechanism significantly reduces variability issues that are met with the standard screening of endogenous VOCs. By correlating molecular processes with phenotypic states, these probes can provide valuable information (e.g. kinetics, localization, progress...) on the biochemical events that lead to the progression or regression of a pathology.

2. Contribution of induced-volatolomics to chemical biology

Chemical biology has generated many innovative tools for monitoring metabolic processes in cells [43]. Researchers in this field have developed numerous stimuli-responsive imaging probes with imaging tracers—from radioisotopes to small-molecule dyes, proteins, and carbon or metal nanomaterials—to drive and localize with high spatial resolution disease outcomes [44–53]. Such activity-based probes can now detect and localize a wide range of biomolecules with high resolution in cell culture, liquid or solid biopsies, and animal models [54–57]. However, the use of these stimuli-sensitive probes in mammalian models remains limited because of toxicity, stability, pharmacokinetics, and tissue distribution hurdles [58,59].

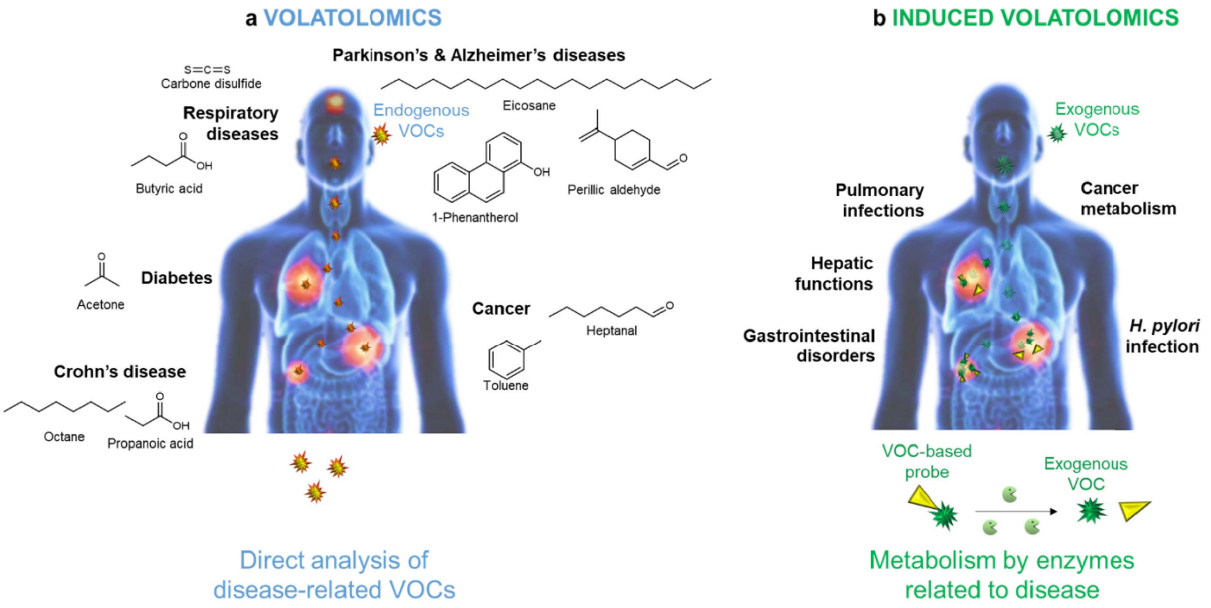


Figure 1. Volatolomics and induced-volatolomics paradigms. Volatolomics is the study of endogenous molecules naturally produced and released by living systems [1,2]. The concept of induced-volatolomics specifically monitors exogenous volatiles that come from the metabolisation of a probe [3].

In this context, a novel set of off/on probes has been designed to explore and understand, at the molecular level, the pathophysiological mechanisms arising in biological systems. These probes can be converted into a small exogenous volatile molecule by a disease- or pathogen-associated enzyme with optimal affinity. Once produced, the volatile tracer diffuses in the biological system and can be released either in animal breath [40,42,60] or in sample headspace (cell culture, liquid or tissue biopsy) [42, 61,62]. Although this action mechanism is common to all volatile-based probes, the nature of the exogenous volatile compounds released from this biochemical process has significantly evolved. The first probes were natural carbohydrates that could be metabolized into inorganic and very volatile compounds (e.g. H_2, CH_4 and CO_2) [63–65]. They were then succeeded by synthetic probes, among which a myriad of isotopic CO_2-based probes were designed for diagnosis purposes [66].

Among these probes, the ^{13}C-urea breath test (UBT) is the best standardized and most widely used test worldwide. UBT is prescribed when *Helicobacter pylori* (*H. pylori*) infection is suspected [40]. This gram-negative micro-aerophilic bacterium can be found on the luminal surface of the gastric epithelium. It is responsible for various upper gastrointestinal disorders, sometimes resulting in ulcers and cancers. In patients with *H. pylori* infection, the ^{13}C-urea probe is hydrolyzed by bacterial urease to NH_3 and $^{13}CO_2$. $^{13}CO_2$ is then excreted via the lungs and measured in the expired air. However, because 1.1% of the natural carbon atoms in the human body are present in the stable isotope ^{13}C form, high $^{13}CO_2$ background noise is observed in patients' breath. Therefore, and despite its high sensitivity and specificity (both around 97%) [67,68], false-negative and positive tests represent a clinical issue, especially for the 5% of samples near cut-off values. This issue has led researchers to design a new generation of volatile-based probes that can release a tracer absent from the pool of endogenous VOCs (Figure 2) [42,61,62, 69,70]. These emergent probes have very recently broadened the scope of "induced-volatolomics" offering the opportunity to explore a wide range of fundamental biological processes.

A set of VOC-based probes was designed to detect *in vitro* pathogenic strains such as *Pseudomonas aeruginosa* [62], *Salmonella* [69] and *Listeria monocytogenes* [70] (Figure 3). These probes are generally

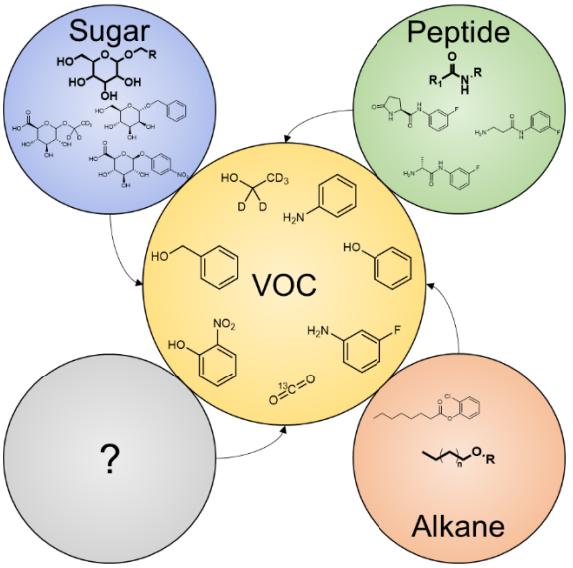


Figure 2. Exogenous VOCs released from VOC-based probes activation.

activated by hydrolytic enzymes such as peptidases or glycosidases to release an exogenous molecule (e.g. aniline, 3-fluoroaniline, 2-chlorophenol, phenol) in biological or food sample headspace. Very recently, Chan *et al.* [60] proposed bioorthogonal peptide-VOC reporters to probe the inflammatory response to respiratory lung infection (Figure 4). In their work, the authors designed a set of VOC-based probes consisting of chemically modified tetrapeptide substrates specific to both human and mouse serine protease neutrophil elastase. At the C-terminal end, an amine-containing VOC was attached via an amide bond. Hydrofluoroamines (HFAs; $CF_3(CF_2) \times CH_2NH_2$) were selected as exogenous VOC reporters because of their high volatility. VOC-based peptide probes were formulated into volatile-releasing Activity-Based Nanosensors (vABNs) by conjugation onto an eight-arm PEG nanocarrier in such a way that the VOCs were undetectable in a non-volatile state. vABNs were further delivered into the lungs via intratracheal instillation, and extracellular proteases produced during respiratory disease cleaved the surface-conjugated peptide substrates, thereby releasing the VOCs. VOCs were finally exhaled and breath samples were collected into a glass vial. VOCs concentrations were quantified using an MS detector (Figure 4).

In parallel, our group demonstrated the efficacy of a VOC-based probe to diagnose cancers and monitor their dynamics during chemotherapy [42]. Our labelled probe, D_5-ethyl-β-D-glucuronide, was designed to target, with high spatial resolution, extracellular β-glucuronidase accumulated selectively in the microenvironment of numerous solid tumors [71–73] (Figure 5). After intravenous administration at a dose as low as 50 $\mu g \cdot kg^{-1}$, D_5-ethyl-β-D-glucuronide was converted into D_5-ethanol and glucuronic acid after β-glucuronidase selective activation within the tumor. D_5-ethanol was then passed to the bloodstream, exhaled in the breath where it was trapped and finally analyzed. This strategy enabled unambiguous discrimination between healthy mice and mice bearing different solid tumors (cervix, mammary, and pulmonary; Figure 5b–d). Furthermore, this approach allowed for the first time to monitor either tumor growth or regression during cancer chemotherapy [74] (Figure 5e,f). Indeed, we showed that tumor regression observed in animals under chemotherapy was accompanied by a decrease in the D_5-ethanol amount in their breath. In contrast, in untreated mice, tumor growth was correlated with an increase in the D_5-ethanol concentration breathed out by the animals. Therefore, these results highlight that our β-glucuronidase-responsive VOC-based probe is a useful tool to follow tumor progression as well as tumor response to chemotherapy.

On the basis of our study conducted in mice, Owlstone Medical (Cambridge, UK), an industry leader in breath research and technical innovation, recently started testing the D_5-ethyl-β-D-glucuronide probe in clinic [75]. Their objective was to provide an early diagnosis test for lung cancer detection. After preclinical tests on lung tumor biopsy and mice bearing lung tumors, Owlstone initiated a phase Ia clinical trial. In this study, they administered D_5-ethyl-β-D-glucuronide to healthy individuals in single ascending dose assays to verify safety and background D_5-ethanol levels. They reported no side effects in these patients. In addition, they did not detect a D_5-ethanol signal in their breath, thereby proving that D_5-ethyl-β-D-glucuronide was not hydrolyzed by intracellular β-glucuronidase.

Despite these successes, induced-volatolomics approaches have challenges related to the selection

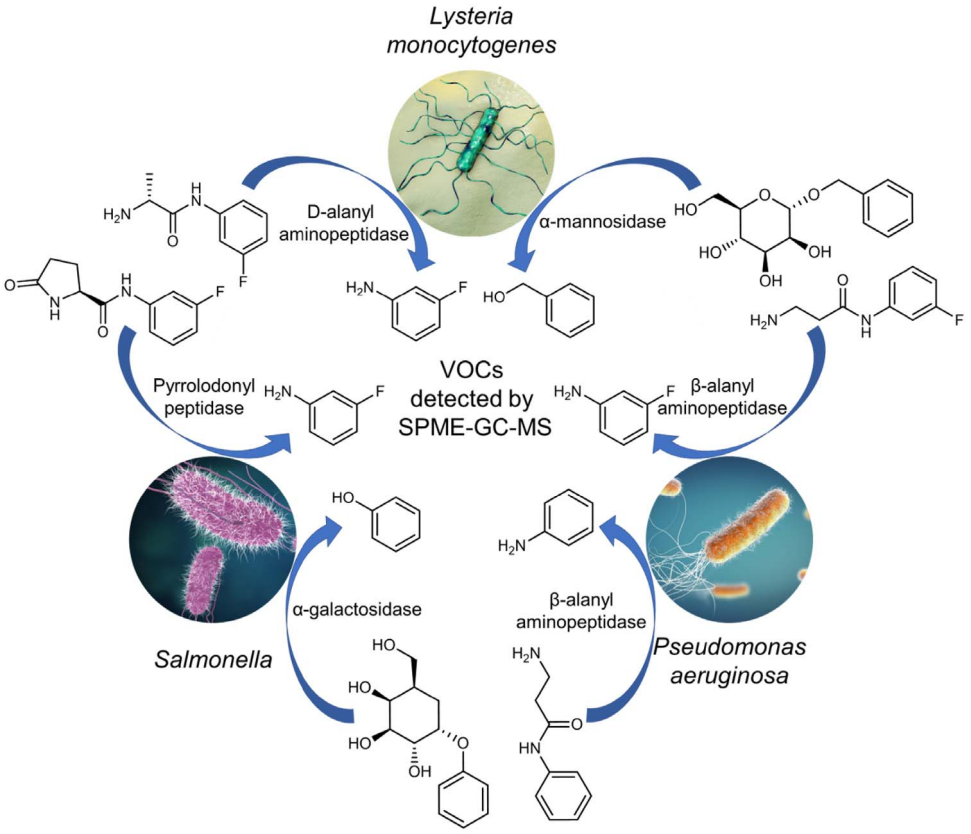


Figure 3. Induced-volatolomics for the *in vitro* detection of pathogens [61,62,69,70].

of the best enzymatic marker. Hence, we recently proposed to combine several VOC-based probes into a cocktail to simultaneously monitor the dysregulation of several tumor-associated enzymes in living mice or biopsies (Figure 6) [76]. Such an outcome is a great advantage in comparison with fluorescent probes, which cannot be used in cocktails because of imaging interferences. The cocktail was composed of four enzyme-responsive VOC-based probes, each targeting a particular glycosidase (β-galactosidase, α-L-fucosidase, β-glucuronidase, *N*-acetyl-β-D-glucosaminidase). Furthermore, in the presence of the corresponding enzyme, each probe led to the release of a specific ethanol isotope (D_2-, D_4-, D_5- and $^{13}CD_5$-ethanol), allowing the identification of the activating glycosidase. When the cocktail was injected into mice bearing tumors, two ethanol isotopes were detected in the animal breath (Figure 6a), highlighting the upregulation of β-glucuronidase and *N*-acetyl-β-D-glucosaminidase in the tumor microenvironment (Figure 6b). Applied *ex vivo* on tumor biopsies, our cocktail again indicated an increased catalytic activity of *N*-acetyl-β-D-glucosaminidase in tumors compared with that in healthy tissues (Figure 6c). Having identified this glycosidase as a potential target for cancer therapy, we designed an enzyme-responsive albumin-binding prodrug [77–80] of the potent monomethyl auristatin E programmed for the selective release of the drug into the tumor microenvironment. This tumor-activated therapy produced remarkable therapeutic efficacy on orthotopic triple-negative mammary xenografts implanted in mice, leading to the disappearance of tumors in 66% of treated animals (Figure 6d). Thus, this study demonstrated the potential of induced-volatolomics for the exploration of biological processes as well as the discovery of novel therapeutic strategies.

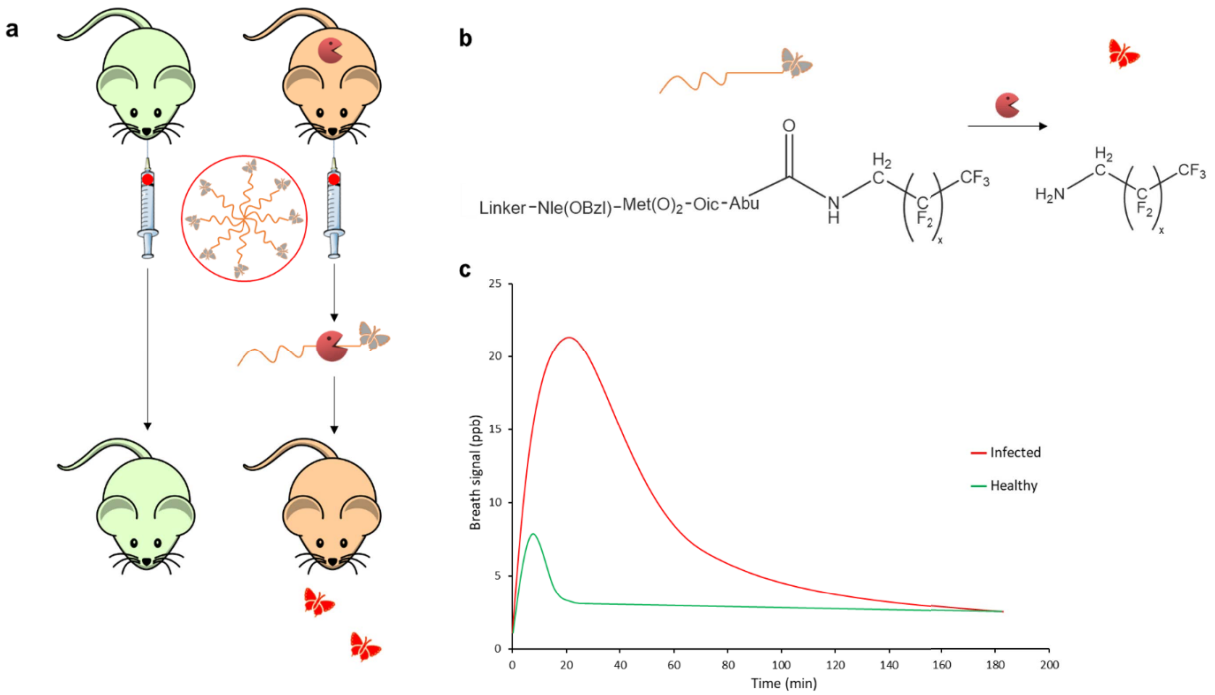


Figure 4. Volatile-releasing Activity-Based Nanosensors (vABNs) for the diagnosis of respiratory infections. (a) VOC-modified peptide substrates are formulated into vABNs by conjugation onto an eight-arm PEG nanocarrier and delivered into the lungs via intratracheal instillation. When attached to the vABN, VOC reporters are in an undetectable, non-volatile state (grey). Extracellular serine protease neutrophil elastase (NE) produced during respiratory disease cleaves the surface-conjugated peptide substrates, thereby releasing VOC reporters. Upon release from vABNs, reporters recover their characteristic mass and volatility (red) and are then exhaled. (b) Enzymatic cleavage by serine protease neutrophil elastase. (c) Breath signal after intrapulmonary delivery of HFA-releasing vABNs in healthy controls and mouse models of lung infection. Figures are reproduced from Chan *et al.* [60].

3. Discussion

The induced-volatolomics is an emerging area of science that could revolutionize our vision of analytical chemistry and strengthen the toolbox of chemical biology. Similar to enzyme-responsive probes, VOC-based probes provide critical information on enzymatic activity changes associated with disease appearance and progression both *in cellulo* and *in vivo* [42,60,62,75].

Whereas off/on optical probes deliver visual images allowing the localization of biochemical events in tissues or body, VOC-based probes can provide additional precision on their spatial distribution by discriminating catalytic events occurring inside versus outside cells [42,60]. Moreover, by using different isotopes of the same volatile tracer (i.e. labelled ethanol), VOC-based probes can be employed in cocktail to measure in real time and simultaneously several enzymatic activities in biological systems (organism or biopsy) [76]. This outcome is of particular interest compared with the usual colorimetric or fluorescent tracers. For instance, simultaneously probing several glycosidase activities through the conversion of the corresponding 4-methylumbelliferyl glycosides is not possible because all these probes release the same fluorescent tracer. While activatable dual and/or multimodal fluorescent probes [81–84] can be envisioned to highlight at least two catalytic activities, their development remains limited due to imaging interferences and/or high background noise. Moreover,

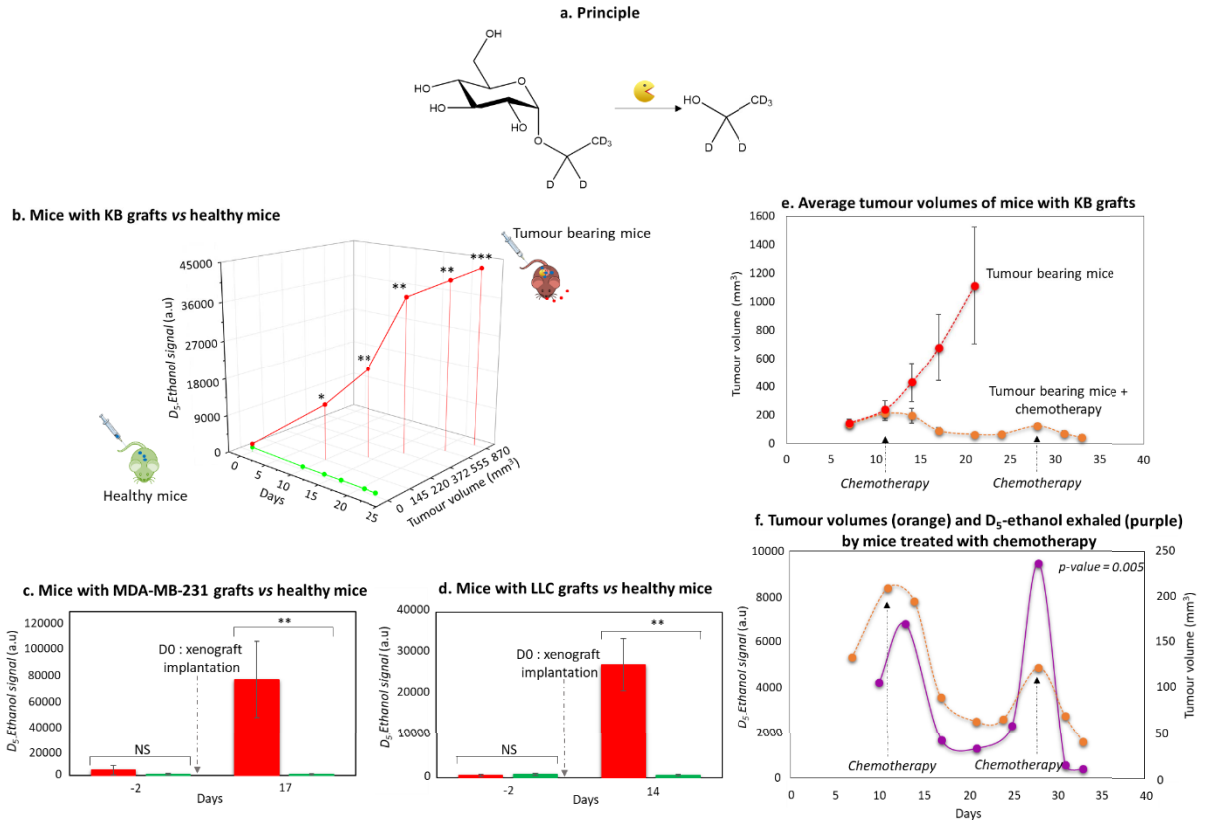


Figure 5. (a) Enzymatic reaction. (b–d) Efficacy of D_5-ethyl-β-D-glucuronide probe to diagnose solid tumors. KB: mouth epidermal carcinoma (n = 13 per group); MDA-MB-231: breast adenocarcinoma (n = 8 per group); LLC: Lewis Lung Carcinoma (n = 6 per group). $*$, $**$, $***$: Significant differences for p-value < 5%; NS: Non-significant. (e) Average tumor volume of mice with KB xenografts (n = 8 per group). (f) Spearman's correlation test (p-value = 0.005) relating the evolution of tumor volumes (orange dot line) and the evolution of D_5-ethanol exhaled (purple line) by KB mice treated with chemotherapy. Figure reproduced from Lange *et al.* [74].

differences in the sensitivity between certain imaging modalities (e.g. MRI and PET) remain an issue, which requires balancing the probe concentration necessary for each imaging technique. On the other hand, tracking the same volatile tracer, as we did with labelled ethanol-based probes [76], prevents analytical variability issues linked to volatile compound detection and quantification. Therefore, VOC-based probe cocktails could soon become very useful tools for the multimodal exploration of global enzymatic dysregulation related to various diseases.

Moreover, the usefulness of VOC-based tracers could cross the boundaries of chemical biology because they are complementary to conventional "omics" approaches. In these fields of research, continuous efforts are being made to design strategies that could provide fundamental insights into enzymatic markers. Within this framework, proteomics has held a central position with the development of multiplex workflows involving antibody-based strategies, electrophoresis strategies, mass-spectrometry analysis, or *in vitro* enzyme assays [85–87]. Although these methodologies showed high specificity, they do not cleanly notify on the fundamental understanding of the roles of these enzymes in their physiological environment. Indeed, several key physical and chemical features of cells are ignored when separating protein markers from their

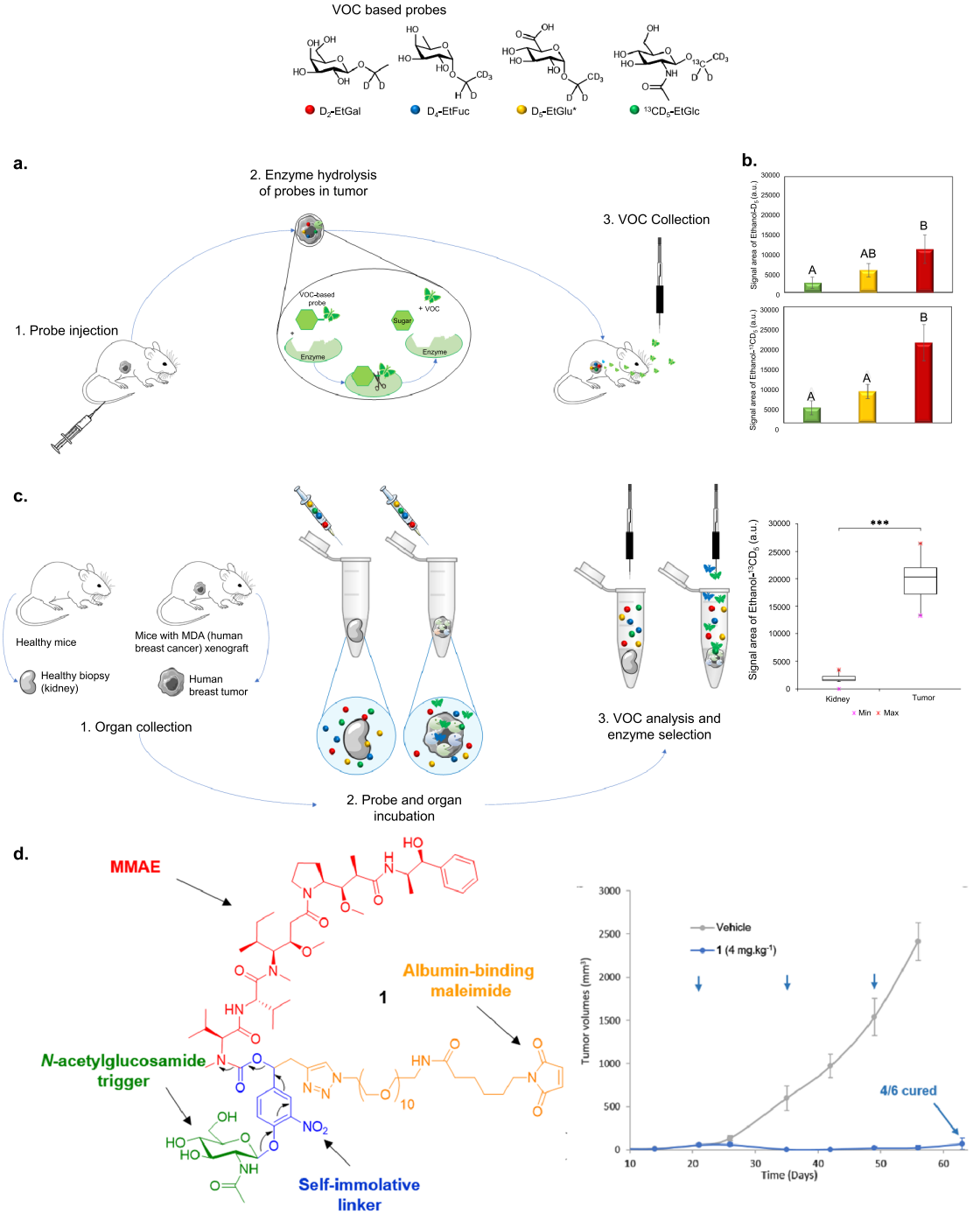


Figure 6. Cocktail of VOC-based probes for the design of tumor activated therapy. To highlight the dysregulation of glycosidases activity in tumors, a cocktail of VOC-based probes (D_2-EtGal, D_4-EtFuc, D_5-EtGlu and $^{13}CD_5$-EtGlcNAc) has been designed. (a) After administration to tumor-bearing mice, the probes can be selectively hydrolyzed in the tumor microenvironment by the corresponding glycosidases to release D_2, D_4, D_5 or $^{13}CD_5$-ethanol isotopes. When released into tumors, labelled VOCs pass through the bloodstream and are exhaled in the breath. (b) Analysis of VOCs signals allows the identification of the most active enzymes, β-glucuronidase and N-acetyl-β-D-glucosaminidase, in the microenvironment of human cervical tumors.

Figure 6. (**cont.**) Amount of D_5-ethanol and $^{13}CD_5$-ethanol exhaled by athymic nude healthy mice (green bars; $n = 6$); amount of D_5-ethanol and $^{13}CD_5$-ethanol exhaled by mice with KB subcutaneous xenograft 12 days after tumor implantation (orange bars, $n = 6$); amount of D_5-ethanol and $^{13}CD_5$-ethanol exhaled by mice with KB subcutaneous xenograft 15 days after tumor implantation (red bars, $n = 6$). For each isotope, significant differences in terms of ethanol isotope amounts were tested by ANOVA (confidence interval of 95%). Significant differences are indicated by the different letters. (c) *Ex vivo* protocol for the detection of exoglycosidases on solid biopsies (healthy kidneys and MDA-MB-231 grafts). The four ethanol isotopes were trapped in the sample headspace and analyzed. VOCs signals tended to show high catalytic activity of N-acetyl-β-D-glucosaminidase in tumor tissues. (d) Structure of the β-GlcNAc-responsive albumin-binding prodrug 1. MDA-MB-231 tumor growth inhibition under therapy with vehicle and 1. Each point shows mean ± s.e.m. from 6 tumor volumes. Figure reproduced from Châtre *et al.* [76].

actual biological contexts. Hence, they may sometimes fail because of the complexity of biological systems that could hamper the purification, detection, and sequencing of low-concentration markers. Thus, induced-volatolomics could provide major advances in terms of sensitivity and normalisation because the biological matrix should not interfere with the detection of exogenous volatile tracers. Furthermore, novel probes targeting other enzyme classes, such as esterases and oxidases, could be developed. By combining these probes with those targeting proteases and glycosidases, several enzymatic processes occurring in different cell compartments (e.g. membrane, cytosol, nucleus, bacteria periplasm and wall) could be integrated to better understand complex biochemical networks. This would certainly enhance our knowledge of enzyme interactions when a biological system is evolving towards diseases. Overall, induced-volatolomics could address fundamental and unsolved questions in biology.

In summary, the induced-volatolomics paradigm has great potential for fundamental understanding of how living cells function [58]. VOC-based probes represent a new generation of tools in the field of chemical biology, which would probably initiate scientific bridges between various disciplines and promote translational science and culture.

Declaration of interests

The authors do not work for, advise, own shares in, or receive funds from any organization that could benefit from this article, and have declared no affiliations other than their research organizations.

Funding

The authors thank the local committees Vienne and Deux-Sèvres of the “Ligue Contre le Cancer” and “La Ligue nationale contre le Cancer”, and the “Région Nouvelle Aquitaine” for financial supports. This work was also funded by the ANR through the Young Researcher grant “Volatolomix” (ANR-21-CE18-0009). We also thank the CNRS, which supplied the project through the “AAP Emergence” and FEDER.

References

[1] S. Giannoukos, A. Agapiou, B. Brkić, S. Taylor, *J. Chromatogr. B*, 2019, **1105**, 136-147.
[2] W. Hu, W. Wu, Y. Jian, H. Haick, G. Zhang, Y. Qian, M. Yuan, M. Yao, *Nano Res.*, 2022, **15**, 8185-8213.
[3] F. Djago, J. Lange, P. Poinot, *Nat. Rev. Chem.*, 2021, **5**, 183-196.
[4] H. Jin, J. Yu, S. Lin, S. Gao, H. Yang, H. Haick, C. Hua, S. Deng, T. Yang, Y. Liu, W. Shen, X. Zhang, X. Zhang, S. Shan, T. Ren, L. Wang, W. Cheung, W. Kam, J. Miao, D. Chen, D. Cui, *ACS Nano*, 2020, **14**, 15517-15532.
[5] S. Traxler, G. Barkowsky, R. Saß, A.-C. Klemenz, N. Patenge, B. Kreikemeyer, J. K. Schubert, W. Miekisch, *Sci. Rep.*, 2019, **9**, article no. 18894.
[6] G. Purcaro, C. A. Rees, W. F. Wieland-Alter, M. J. Schneider, X. Wang, P.-H. Stefanuto, P. F. Wright, R. I. Enelow, J. E. Hill, *J. Breath Res.*, 2018, **12**, article no. 026015.
[7] G. Walzl, R. McNerney, N. du Plessis, M. Bates, T. D. McHugh, N. N. Chegou, A. Zumla, *Lancet Infect. Dis.*, 2018, **18**, e199-e210.
[8] C. M. De Moraes, C. Wanjiku, N. M. Stanczyk, H. Pulido, J. W. Sims, H. S. Betz, A. F. Read, B. Torto, M. C. Mescher, *Proc. Natl. Acad. Sci. USA*, 2018, **115**, 5780-5785.
[9] M. P. P. Van Oort, S. De Bruin, H. Weda, H. H. Knobel, J. M. Schultz, D. L. Bos, On Behalf of the MARS Consortium, *Int. J. Mol. Sci.*, 2017, **18**, article no. 449.
[10] J. E. Belizário, J. Faintuch, M. G. Malpartida, *Front. Cell. Infect. Microbiol.*, 2021, **10**, article no. 564194.

[11] A. H. Neerincx, S. J. H. Vijverberg, L. D. J. Bos, P. Brinkman, M. P. van der Schee, R. de Vries, P. J. Sterk, A.-H. Maitland-van der Zee, *Pediatr. Pulmonol.*, 2017, **52**, 1616-1627.
[12] L. D. J. Bos, *Ann. Transl. Med.*, 2018, **6**, article no. 33.
[13] F. Usman, J. O. Dennis, A. Y. Ahmed, F. Meriaudeau, O. B. Ayodele, A. A. S. Rabih, *IEEE Access*, 2019, **7**, 5963-5974.
[14] D. K. Trivedi, E. Sinclair, Y. Xu, D. Sarkar, C. Walton-Doyle, C. Liscio, P. Banks, J. Milne, M. Silverdale, T. Kunath, R. Goodacre, P. Barran, *ACS Cent. Sci.*, 2019, **5**, 599-606.
[15] J. Zhou, Z.-A. Huang, U. Kumar, D. D. Y. Chen, *Anal. Chim. Acta*, 2017, **996**, 1-9.
[16] X. Sun, K. Shao, T. Wang, *Anal. Bioanal. Chem.*, 2016, **408**, 2759-2780.
[17] J. Giró Benet, M. Seo, M. Khine, J. Gumà Padró, A. Pardo Martnez, F. Kurdahi, *Sci. Rep.*, 2022, **12**, article no. 14873.
[18] S. N. Rai, S. Das, J. Pan, D. C. Mishra, X.-A. Fu, *PLoS One*, 2022, **17**, article no. e0277431.
[19] H.-C. Lau, J.-B. Yu, H.-W. Lee, J. S. Huh, J. O. Lim, *Sensors*, 2017, **17**, article no. 1783.
[20] G. A. Kwong, S. Ghosh, L. Gamboa, C. Patriotis, S. Srivastava, S. N. Bhatia, *Nat. Rev. Cancer*, 2021, **21**, 655-668.
[21] L. Dormont, J.-M. Bessière, A. Cohuet, *J. Chem. Ecol.*, 2013, **39**, 569-578.
[22] D. J. Penn, E. Oberzaucher, K. Grammer, G. Fischer, H. A. Soini, D. Wiesler, M. V. Novotny, S. J. Dixon, Y. Xu, R. G. Brereton, *J. R. Soc. Interface*, 2007, **4**, 331-340.
[23] H. AL-Kateb, B. de Lacy Costello, N. Ratcliffe, *J. Breath Res.*, 2013, **7**, article no. 036004.
[24] P. Mochalski, J. King, M. Klieber, K. Unterkofler, H. Hinterhuber, M. Baumann, A. Amann, *Analyst*, 2013, **138**, 2134-2145.
[25] A. Mitra, S. Choi, P. R. Boshier, A. Razumovskaya-Hough, I. Belluomo, P. Spanel, G. B. Hanna, *Metabolites*, 2022, **12**, article no. 824.
[26] S. Janfaza, B. Khorsand, M. Nikkhah, J. Zahiri, *Biol. Methods Protoc.*, 2019, **4**, 1-11.
[27] M. Serasanambati, Y. Y. Broza, A. Marmur, H. Haick, *iScience*, 2019, **11**, 178-188.
[28] Z. Jia, H. Zhang, C. N. Ong, A. Patra, Y. Lu, C. T. Lim, T. Venkatesan, *ACS Omega*, 2018, **3**, 5131-5140.
[29] A. T. Güntner, S. Abegg, K. Königstein, P. A. Gerber, A. Schmidt-Trucksäss, S. E. Pratsinis, *ACS Sens.*, 2019, **4**, 268-280.
[30] H. Lan, K. Hartonen, M.-L. Riekkola, *TrAC Trends Analyt. Chem.*, 2020, **126**, article no. 115873.
[31] O. Lawal, W. M. Ahmed, T. M. E. Nijsen, R. Goodacre, S. J. Fowler, *Metabolomics*, 2017, **13**, article no. 110.
[32] P. Shende, J. Vaidya, Y. A. Kulkarni, R. S. Gaud, *J. Control. Release*, 2017, **268**, 282-295.
[33] H. C. Ates, C. Dincer, *Nat. Rev. Bioeng.*, 2023, **1**, 80-82.
[34] J. N. Thomas, J. Roopkumar, T. Patel, *PLoS One*, 2021, **16**, article no. e0260098.
[35] M. A. Chowdhury, T. E. Rice, M. A. Oehlschlaeger, *Appl. Sci.*, 2022, **12**, article no. 8447.
[36] M. Arora, S. C. Zambrzycki, J. M. Levy, A. Esper, J. K. Frediani, C. L. Quave, F. M. Fernández, R. Kamaleswaran, *Metabolites*, 2022, **12**, article no. 232.
[37] M. Wieczorek, A. Weston, M. Ledenko, J. N. Thomas, R. Carter, T. Patel, *Front. Med.*, 2022, **9**, article no. 992703.
[38] Y. Huang, I.-J. Doh, E. Bae, *Sensors*, 2021, **21**, article no. 3923.
[39] P. Rajpurkar, E. Chen, O. Banerjee, E. J. Topol, *Nat. Med.*, 2022, **28**, 31-38.
[40] D. Y. Graham, M. Miftahussurur, *J. Adv. Res.*, 2018, **13**, 51-57.
[41] E. Gaude, M. K. Nakhleh, S. Patassini, J. Boschmans, M. Allsworth, B. Boyle, M. P. van der Schee, *J. Breath Res.*, 2019, **13**, article no. 032001.
[42] J. Lange, B. Eddhif, M. Tarighi, T. Garandeau, E. Péraudeau, J. Clarhaut, B. Renoux, S. Papot, P. Poinot, *Angew. Chem. Int. Ed.*, 2019, **58**, 17563-17566.
[43] M. Bucci, C. Goodman, T. L. Sheppard, *Nat. Chem. Biol.*, 2010, **6**, 847-854.
[44] W. Chyan, R. T. Raines, *ACS Chem. Biol.*, 2018, **13**, 1810-1823.
[45] L. Wu, Z. Armstrong, S. P. Schröder, C. de Boer, M. Artola, J. M. Aerts, H. S. Overkleeft, G. J. Davies, *Curr. Opin. Chem. Biol.*, 2019, **53**, 25-36.
[46] H. Fang, B. Peng, S. Y. Ong, Q. Wu, L. Li, S. Q. Yao, *Chem. Sci.*, 2021, **12**, 8288-8310.
[47] Y. Kuriki, T. Yoshioka, M. Kamiya, T. Komatsu, H. Takamaru, K. Fujita, H. Iwaki, A. Nanjo, Y. Akagi, K. Takeshita, H. Hino, R. Hino, R. Kojima, T. Ueno, K. Hanaoka, S. Abe, Y. Saito, J. Nakajima, Y. Urano, *Chem. Sci.*, 2022, **13**, 4474-4481.
[48] H. Singh, K. Tiwari, R. Tiwari, S. K. Pramanik, A. Das, *Chem. Rev.*, 2019, **119**, 11718-11760.
[49] Y. Wang, J. Weng, X. Wen, Y. Hu, D. Ye, *Biomater. Sci.*, 2021, **9**, 406-421.
[50] S. Shingo, K. Toru, W. Rikiya, Z. Yi, I. Taiki, K. Mitsuyasu, N. Hidehiko, U. Takaaki, O. Takuji, H. Kazufumi, N. Hiroyuki, U. Yasuteru, *Sci. Adv.*, 2020, **6**, article no. eaay0888.
[51] R. K. Muir, M. Guerra, M. M. Bogyo, *ACS Chem. Biol.*, 2022, **17**, 281-291.
[52] S. M. McKenna, E. M. Fay, J. F. McGouran, *ACS Chem. Biol.*, 2021, **16**, 2719-2730.
[53] I. Amarsy, S. Papot, G. Gasser, *Angew. Chem. Int. Ed.*, 2022, **61**, article no. e202205900.
[54] S. Sinharay, E. A. Randtke, C. M. Howison, N. A. Ignatenko, M. D. Pagel, *Mol. Imaging Biol.*, 2018, **20**, 240-248.
[55] L. Galas, T. Gallavardin, M. Bénard, A. Lehner, D. Schapman, A. Lebon, H. Komuro, P. Lerouge, S. Leleu, X. Franck, *Chemosensors*, 2018, **6**, article no. 40.
[56] E. G. Stoddard, R. F. Volk, J. P. Carson, C. M. Ljungberg, T. A. Murphree, J. N. Smith, N. C. Sadler, A. K. Shukla, C. Ansong, A. T. Wright, *J. Proteome Res.*, 2018, **17**, 2623-2634.
[57] S. Sabale, P. Kandesar, V. Jadhav, R. Komorek, R. K. Motkuri, X.-Y. Yu, *Biomater. Sci.*, 2017, **5**, 2212-2225.
[58] E. A. Lemke, C. Schultz, *Nat. Chem. Biol.*, 2011, **7**, 480-483.
[59] N. K. Devaraj, *ACS Cent. Sci.*, 2018, **4**, 952-959.
[60] L. W. Chan, M. N. Anahtar, T.-H. Ong, K. E. Hern, R. R. Kunz, S. N. Bhatia, *Nat. Nanotechnol.*, 2020, **15**, 792-800.
[61] E. Tait, S. P. Stanforth, S. Reed, J. D. Perry, J. R. Dean, *RSC Adv.*, 2015, **5**, 15494-15499.
[62] R. Thompson, D. Stephenson, H. E. Sykes, J. D. Perry, S. P. Stanforth, J. R. Dean, *RSC Adv.*, 2020, **10**, 10634-10645.
[63] M. Simrén, P.-O. Stotzer, *Gut*, 2006, **55**, 297-303.
[64] R. M. Harvie, C. J. Tuck, M. Schultz, *JGH Open*, 2020, **4**, 198-205.
[65] F. Gao, M. Wang, X. Zhang, J. Zhang, Y. Xue, H. Wan, P. Wang, *Anal. Methods*, 2018, **10**, 4329-4338.
[66] A. S. Modak, *J. Breath Res.*, 2013, **7**, article no. 037103.
[67] F. Gomollón, J. A. Ducons, S. Santolaria, I. Lera Omiste,

R. Guirao, M. Ferrero, M. Montoro, *Dig. Liver Dis.*, 2003, **35**, 612-618.

[68] S. Som, A. Maity, G. D. Banik, C. Ghosh, S. Chaudhuri, S. B. Daschakraborty, S. Ghosh, M. Pradhan, *Anal. Bioanal. Chem.*, 2014, **406**, 5405-5412.

[69] N. H. O. Bahroun, J. D. Perry, S. P. Stanforth, J. R. Dean, *Anal. Chim. Acta*, 2018, **1028**, 121-130.

[70] C. Taylor, F. Lough, S. P. Stanforth, E. C. Schwalbe, I. A. Fowlis, J. R. Dean, *Anal. Bioanal. Chem.*, 2017, **409**, 4247-4256.

[71] K. Bosslet, J. Czech, D. Hoffmann, *Tumor Target.*, 1995, **1**, 45-50.

[72] Y. Jin, X. Tian, L. Jin, Y. Cui, T. Liu, Z. Yu, X. Huo, J. Cui, C. Sun, C. Wang, J. Ning, B. Zhang, L. Feng, X. Ma, *Anal. Chem.*, 2018, **90**, 3276-3283.

[73] P. Awolade, N. Cele, N. Kerru, L. Gummidi, E. Oluwakemi, P. Singh, *Eur. J. Med. Chem.*, 2020, **187**, article no. 111921.

[74] B. Renoux, F. Raes, T. Legigan, E. Péraudeau, B. Eddhif, P. Poinot, I. Tranoy-Opalinski, J. Alsarraf, O. Koniev, S. Kolodych, S. Lerondel, A. Le Pape, J. Clarhaut, S. Papot, *Chem. Sci.*, 2017, **8**, 3427-3433.

[75] C. F. Labuschagne, R. Smith, N. Kumar, M. Allsworth, B. Boyle, S. Janes, P. Crosbie, R. Rintoul, *J. Clin. Oncol.*, 2022, **40**, 2569-2569.

[76] R. Châtre, E. Blochouse, F. Djago, M. Tarighi, B. Renoux, J. Sobilo, A. Le Pape, J. Clarhaut, I. Opalinski, C. Geffroy-Rodier, W. Tuo, S. Papot, P. Poinot, *Chem. Sci.*, 2023, **14**, 4697-4703.

[77] T. Legigan, J. Clarhaut, B. Renoux, I. Tranoy-Opalinski, A. Monvoisin, J.-M. Berjeaud, F. Guilhot, S. Papot, *J. Med. Chem.*, 2012, **55**, 4516-4520.

[78] B. Renoux, F. Raes, T. Legigan, E. Péraudeau, B. Eddhif, P. Poinot, I. Tranoy-Opalinski, J. Alsarraf, O. Koniev, S. Kolodych, S. Lerondel, A. Le Pape, J. Clarhaut, S. Papot, *Chem. Sci.*, 2017, **8**, 3427-3433.

[79] B. Renoux, L. Fangous, C. Hötten, E. Péraudeau, B. Eddhif, P. Poinot, J. Clarhaut, S. Papot, *Med. Chem. Commun.*, 2018, **9**, 2068-2071.

[80] R. Châtre, J. Lange, E. Péraudeau, P. Poinot, S. Lerondel, A. Le Pape, J. Clarhaut, B. Renoux, S. Papot, *J. Control. Release*, 2020, **327**, 19-25.

[81] Y. Wang, Y. Hu, D. Ye, *Angew. Chem. Int. Ed.*, 2022, **61**, article no. e202209512.

[82] J. M. Tam, L. Josephson, A. R. Pilozzi, X. Huang, *Molecules*, 2020, **25**, article no. 274.

[83] J. Tam, A. Pilozzi, U. Mahmood, X. Huang, *Int. J. Mol. Sci.*, 2020, **21**, article no. 3068.

[84] Y. Cheng, R. M. Borum, A. E. Clark, Z. Jin, C. Moore, P. Fajtová, A. J. O'Donoghue, A. F. Carlin, J. V. Jokerst, *Angew. Chem. Int. Ed.*, 2022, **61**, article no. e202113617.

[85] A. M. Sadaghiani, S. H. Verhelst, M. Bogyo, *Curr. Opin. Chem. Biol.*, 2007, **11**, 20-28.

[86] Y. Zhang, B. R. Fonslow, B. Shan, M.-C. Baek, J. R. I. Yates, *Chem. Rev.*, 2013, **113**, 2343-2394.

[87] R. M. Miller, L. M. Smith, *Analyst*, 2023, **148**, 475-486.

Comptes Rendus
Chimie
2023, Vol. 26, Special Issue S3, p. 105-117
https://doi.org/10.5802/crchim.261

Breaking Barriers in Chemical Biology – Toulouse 2022

Functionalized oligonucleotides, synthetic catalysts as enzyme mimics

Crystalle Chardet [a], Sandra Serres [a], Corinne Payrastre [a], Jean-Marc Escudier [a] and Béatrice Gerland [*, a]

[a] Laboratoire de Synthèse et Physico-Chimie de Molécules d'Intérêt Biologique, UMR CNRS 5068, Université Paul Sabatier, 118 Route de Narbonne, CEDEX 9, 31062 Toulouse, France

E-mail: beatrice.gerland@univ-tlse3.fr (B. Gerland)

Abstract. This review describes the chemical rationale behind the modification of deoxyribonucleotides, which are later incorporated through an enzymatic or automated supported process, resulting in the synthesis of functionalized oligonucleotides (FuON). These additional handles (located on the base or the sugar moiety) were carefully chosen to enhance the rather limited repertoire of DNA to develop FuONs that display enzyme-like properties such as catalysis and recognition. We will exemplify the development of FuON either presenting RNA-cleaving activity or acting as peroxidase or protease mimics and discuss the impact of the introduced modifications.

Keywords. Functionalized oligonucleotides, Enzyme mimics, Chemical modifications.

Funding. This work was supported by a grant from the Agence Nationale de la Recherche, ANR-18-CE07-0023, to BG.

Manuscript received 10 February 2023, revised 20 July 2023, accepted 11 September 2023.

1. Introduction

In Nature, enzymes are proteins that ensure the efficiency and selectivity of complex chemical transformations under physiological aqueous conditions. The key feature for catalysis is an appropriate macromolecular 3D folding to bring in close proximity catalytic partners (amino acid side-chains, cofactors…) while discriminating between potential substrates and assuring the final specificity outcome. Nucleic acids also display 3D complex structures with the unmatched advantages of a programable folding and dynamic renaturation processes (unlike proteins). Until the beginning of the 80s and the discovery of RNA acting as an enzyme with self-splicing properties (coining the term RNAzyme), the role of nucleic acids was thought to be only devoted to the storage and transfer of genetic information [1, 2]. DNA was not even primarily considered, assuming that an inert chemical nature would provide an evolutionary advantage by ensuring the absence of undesired alterations of genetic instructions. The major hurdles to overcome were four canonical nucleotides with limited functionalities (mostly involved into Watson and Crick base pairing) and no hydroxyl group at the 2′-position of the sugar. It took another decade to demonstrate that DNAzymes, single-stranded deoxyoligonucleotides (ODN) with no in vivo counterparts, were also capable of catalytic activities that could match those of enzymes [3,4]. Both aptamers (oligonucleotides able to bind a selected target but with no catalytic properties) and DNAzymes could be selected in vitro by the iterative and powerful SELEX method [5,6], relying on the use of unmodified nucleosides 5′-triphosphate (dNTP). These nucleotides are substrates of (mutant) DNA

* Corresponding author

polymerases that can be integrated during the DNA amplification step of the polymerase chain reaction (PCR). It is wisely assumed that a library composed of 10^{15} sequences (obtained from a randomized region made of 25 nucleotides) is sufficient to identify a sequence of interest by exploring a sufficiently diverse chemical space [7]. To achieve such catalytic properties, DNAzymes were able to compensate for the absence of the 2′-hydroxyl group (in comparison with RNA counterparts) by a less restrictive sugar puckering, leading to complex architectures [8–10] or unusual secondary structures such as triplex structures [11]. DNA secondary structures adopt local conformations that could influence the pKa values of functional groups in the same magnitude as that observed in enzymes [12]. Taking these remarks into consideration, it has been experimentally confirmed that DNAzymes could be as efficient as their RNA counterparts in the Diels–Alder reaction [13]. Moreover, DNA displays intrinsic qualities in terms of fast and effective automated supported synthesis (SPS) combined with increased chemical stability compared with RNA. Nucleic-based biomimetic catalysis could definitively benefit from these former qualities combined with the DNA aptamer property and capacity to be selected owing to in vitro biomolecular methodologies.

However, it rapidly emerged that modifications of the native DNAzyme backbone were required when moving toward biological and/or clinical applications. Switching to a phosphorothioate linkage combined to a 3′-inverted nucleotide was proven to be efficient in avoiding nuclease degradation, relying on the solutions already developed for the antisens approach [14]. Most of the work was devoted to structural modifications of the sugar moiety, alongside the classical 2′-*O*-methyl modification. For example, DNAzymes based on XNAs (for Xenobiotic Nucleic Acid) [15,16] and in particular on FANAs (2′-Fluoro-Arabino Nucleic Acid) were demonstrated to be less sensitive to nuclease digestion while maintaining catalytic RNA-ligating [17] or -cleaving activities [18,19]. FANAzymes present the additional advantage of being obtained by in vitro evolution using non-engineered DNA polymerases. It is also of great interest to mention the ligase-catalyzed oligonucleotide polymerization (LOOPER) method developed by Hili's laboratory [20]. Pentanucleotides anticodons encode modifications through functionalization of the 5′-end nucleotide with chemical scaffolds varying from alkyl chains and phenyl rings to organic functions while maintaining a 3′-end variable dinucleotide sequence for degeneracy. The combination of SELEX and LOOPER processes allows the selection of functional aptamers against human alpha-thrombin bearing diverse chemical groups with high binding affinity and not relying on the usual G-quadruplex structuration. Although this approach was concentrated to improve aptamer affinities, it also emphasizes the added value of increasing the DNA chemical repertoire toward its functional role in terms of recognition and ultimately biomimetic catalysis. We decided to narrow the scope of the review to covalently modified monomers (either dNTP or phosphoramidites) that lead to functionalized (deoxy)oligonucleotides (FuON) with catalytic enhancement purposes to mimic biochemical reactions. We will then first describe FuON with enzyme-like properties that catalyze reactions on nucleic acids substrates, such as the extensively studied DNA catalyzed-RNA cleavage. The focus will then be shifted to a broader reactivity scope as described by the development of FuON, which displays peroxidase-mimicking activity and DNA-based proteases.

As the chemical strategies to introduce modifications have been reviewed in the literature by numerous groups [21,22] and ourselves [23], we will only present the functionalized monomers of interest in charge of the catalytic activities.

2. Functionalized oligonucleotides as ribonuclease A mimics

Screening for new DNA catalysts candidates in the general context of degrading viral RNA in a sequence-specific fashion was inspired by the ribonuclease A (RNase A) responsible for cleaving the phosphodiester bond. The target, being a RNA substrate, presents the tremendous advantage of bypassing the recognition hurdle, as DNA is indeed complementary to RNA. RNase A catalyzes the cleavage of the P-5′-*O*-bond to give two RNA products bearing a 5′-OH termini and a 2′,3′-cyclic phosphodiester on the 3′-side of a pyrimidine nucleotide, which is further hydrolyzed to a 3′-monophosphate [24]. During the

first step of the 2′-*O*-trans-phosphorylation, the imidazole ring of Histidine 12 (His12) acts as a general base to increase the nucleophilic character of the 2′-hydroxyl, allowing intramolecular in-line attack on the adjacent phosphorus atom. On the other hand, the leaving ability of the 5′-oxygen is increased because of the protonation provided by the imidazolium side chain of His19.

Considering the key role of two histidine side chains in the general acid–base catalytic mechanism, the first modified DNAzyme selected to cleave RNA was named "16.2-11" and decorated with imidazole groups [25]. On a suitably protected 5-(3-aminopropenyl)-2′-deoxyuridine, the 4-imidazoleacrylic acid moiety was linked through an amide coupling to the primary amine before converting the resulting nucleoside either to its corresponding 5′-triphosphate **1** (for the SELEX process) or phosphoramidite **2**, for the solid-phase synthesis of the most active DNAzymes (Figure 1).

Introduction of modifications at the C5 position on pyrimidines (or at the N7 or C8 positions on the purine base) does not impair recognition by DNA polymerases during the enzymatic polymerization of dNTP inherent to the SELEX process. Moreover, Watson and Crick base pairing is also maintained during duplex formation, which is a key feature for efficient RNA recognition. The overall structure of such functionalized DNAzymes was described by a catalytic core containing four imidazole residues located close to the cleavage site and two binding arms complementary to the RNA substrate. Despite displaying high efficiency ($k_{obs} < 1.5$ min^{-1}), sequence specificity, and turnover (the DNA binding arms being short enough to denature the duplex structure once the RNA strand was cleaved), the presence of divalent metal ions (in particular Zn^{2+}) was required. This metal ion dependency could be explained by Zn^{2+} ions either playing a structural role in folding the DNAzyme in an active conformation or being part of the mechanism, activating the 2′-hydroxyl or a molecule of water (or both). Although encouraging, this result indicated that imidazole modifications were not sufficient to bypass the need for metal cations catalysis.

Lysine 31 (Lys31) was also proposed to be involved in transition state stabilization, and the lack of a cationic function needed for the general acid–base mechanism might be the main reason for M^{2+}-dependent catalysis. This rationale developed by Perrin's laboratory was confirmed by the selection of the first independent DNAzyme 9_{25}-11 bearing both imidazole and ammonium functionalities with $k_{obs} <$ 0.2 min^{-1} [26,27]. This approach relied on two dNTP **3** and **4** to introduce the different modifications on a purine or pyrimidine base, respectively (Figure 2).

First, the substitution of the bromine atom by histamine onto 8-bromo-2-deoxyriboadenosine-5′-monophosphate and subsequent conversion led to the corresponding (8-(2-(4-imidazolyl)ethylamino)-2′-deoxyriboadenosine-5′-triphosphate ($dA^{Im}TP$) **3**. Along with commercial 5-(3-aminoallyl)-2′-deoxyribouridine-5′-triphosphate ($dU^{aa}TP$) **4**, these compounds were used for in vitro selection. 9_{25}-11 was only able to cleave a single ribonucleotide bond embedded into a DNA substrate (Figure 3, left). To evaluate its *trans*-catalytic properties, M^{2+}-independent 9_{25}-11 presenting a minimal yet active structure (represented in italic, Figure 3, left) was synthesized de novo by SPS from the corresponding phosphoramidites of **3** and **4**. Kinetic studies indicated a k_{cat} of $\sim 1.5 \times 10^{-2}$ min^{-1}, a 3-fold decrease compared with the previously observed *cis* cleavage. However, a high substrate specificity against the target d(GCGTGCC)rCd(GTCTGTT) was observed as a drastic reduced cleavage for degenerate sequences was measured.

In contrast, subsequent work by Williams *et al.* led to the cleavage of substrates made only of RNA [28] (Figure 3, middle). They relied on previously described nucleotide **1** to introduce imidazole moieties and designed 7-aminopropynyl modified 7-deaza-dATP **5** (Figure 2). They observed that the introduction of modifications at this position on the adenine base was less destabilizing within a duplex, facilitating the PCR step, combined with a better tolerance of DNA polymerases toward a more rigid linker such as the propenyl group [30]. After a palladium-catalyzed coupling reaction between 7-deaza-7-iodo desoxyadenosine and *N*-propynyltrifluoracetamide followed by a deprotection step, the resulting nucleoside was converted into its corresponding dATP **5**. This DNAzyme clone 32 did present a M^{2+}-independent *cis* catalytic activity on a 12nt target RNA sequence with a weak k_{obs} of 0.07 min^{-1}, still a 10^5 rate improvement compared to the uncatalyzed RNA cleavage. The cleavage was site-specific, with the scissile bond

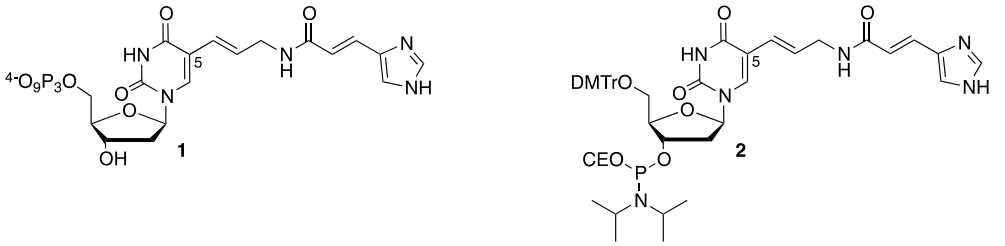


Figure 1. C5-Functionalized nucleotides bearing an imidazole moiety (CE = cyanoethyl).

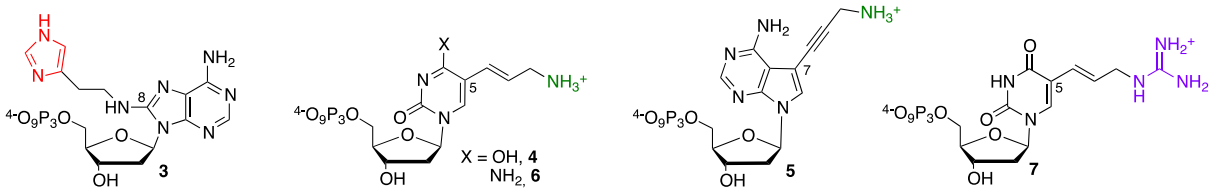


Figure 2. Functionalized nucleotides used for the selection of RNase A mimics.

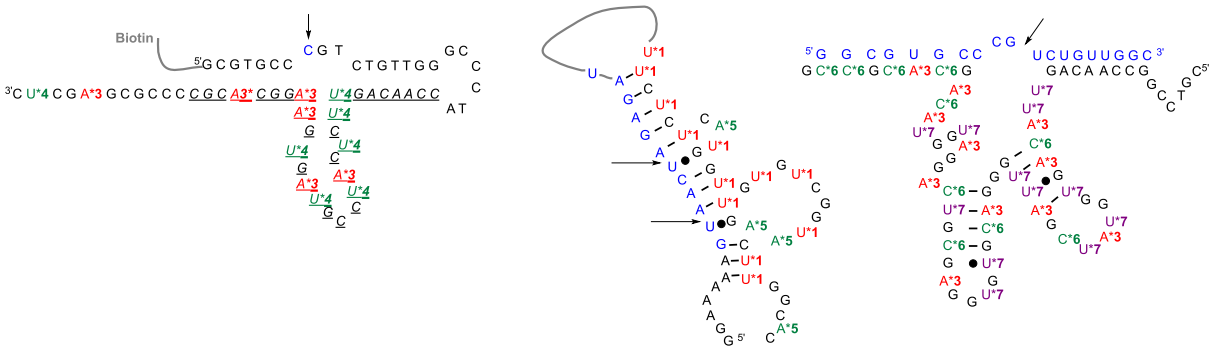


Figure 3. Predicted secondary structures of (left) DNAzyme 9_{25}-11 [26] (in italic the mimimal catalytic core [27]) with a single ribophosphodiester substrate; (middle) DNAzyme clone 32 with an all-RNA target [28]; (right) DNAzyme 7-38-22t with an all-RNA target [29]. (Blue letters: ribonucleotides; black letters: deoxyribonucleotides; star colored letters: modified nucleotides; black arrows indicate the cleavage site.)

located opposite the loop region presenting both imidazole and amino residues. The authors of those studies concluded on the unknown precise role of those appended functionalities, as they could either participate directly in the catalytic mechanism or be involved in the folding DNAzyme process leading to an active structuration, thus advocating for further investigations. Building on those first discoveries, other functionalized RNA-cleaving DNAzymes were developed as thoroughly reviewed by Hollenstein [31] and others [32,33]. The long-term research work from Perrin's laboratory was rewarded by the design of an efficient DNAzyme able to mimic RNase A catalysis in the absence of divalent metal cations. During selection, the authors combined the already developed 8-histaminyl-deoxyadenosine ($dA^{im}TP$) **3**, the 5-aminoallyl-deoxycytidine ($dC^{aa}TP$) **6**, an analog to $dU^{aa}TP$, with 5-guanidinoallyl-deoxyuridine ($dU^{ga}TP$) **7**, presenting a cationic group known to increase thermostability by electrostatic complementarity with the phosphodiester negative charge (Figure 2).

DNAzyme 7-38-22 was found to cleave a 19nt long all-RNA substrate in the absence of M^{2+} ions

and with multiple turnovers, displaying a rate constant of 4.9 min^{-1} for self-cleavage [29]. Designed to easily allow the evaluation of *trans*-catalysis, the DNazyme was able to cleave (k_{cat} = 0.27 min^{-1} with multiple turn-overs) with a 20-fold selectivity the UG bond of a 19nt all-RNA substrate embedded into a four bases bulge that was located (as usually observed) in front of a fully modified catalytic core (Figure 3, right). They then demonstrated that the presence of an imidazole ring, which is thought to be required as a key player for acid–base catalysis, could be bypassed by solely introducing both cationic amine and guanidine functions to select self-cleaving DNAzymes with higher activities than some imidazole-decorated DNAzymes [34].

Site-selective modification of the catalytic core of 10-23 DNAzyme was explored through SPS by introducing functionalized 2′-deoxy(deaza)adenosine phosphoramidites with imidazole, guanidinium, or amino group either at the C-6, C-7 or C-8 onto the purine base. The "chemical evolution" of the A9 or A12 position confirmed the necessity of maintaining the Watson and Crick base pairing (as expected), whereas other modifications did not improve cleaving activity [35]. Other phosphoramidites bearing different linkers with a terminal amino group at the 2′-position were synthesized to evaluate the impact of the modifications of the known conserved positions of the 8-17 DNAzyme. The catalytic effects of a Ca^{2+}-mediated reaction were found to be limited [36].

Apart from the progress that could arise from the optimization of SELEX conditions, there is also room for chemical input in designing new functionalized nucleotides with different linkers or even functionalities, considering that no clear rationale has emerged to pinpoint the beneficial positions to modify a given DNAzyme.

As indicated, RNA cleavage mediated by DNA usually requires the presence of bivalent cations. Consequently, DNAzymes have been selected for metal detection and biosensing purposes; however, when unmodified, metal selectivity remains low [37]. The introduction of additional coordination sites via an organic ligand onto the ODN could improve such selectivity. This approach was recently described using a glycil-histidine functionalized tertiary amine introduced onto a RNA-cleaving DNAzyme, which enhanced selectivity toward the desired divalent metallic ions [38]. Oxidation also requires a metallic center in charge of the catalytic function. A recent example from Martell's group demonstrated the use of a DNA double helix as a scaffold to spatially preorganize two co-catalysts (4-carboxy-TEMPO and 4,4′-dicarboxy-2,2′-bipyridine) involved in synergistic Cu-TEMPO alcohol oxidation [39]. Conjugated through an amide link at the 5′-end or 3′-end, respectively, of a matched DNA duplex, both are positioned at an optimal distance, as confirmed by the 70-fold activity enhancement in catalyst turnover number compared with the unscaffolded oxidation reaction; however, no specificity for the various tested alcohol substrates was observed. We decided to focus on the field of peroxidase-like mimics because studies have exemplified the intrinsic DNA properties in terms of structuration, recognition, and catalysis, which are combined in functionalized oligonucleotides purposefully designed to recognize their substrates and perform the subsequent catalytic reaction.

3. Functionalized oligonucleotides as horseradish peroxidase and phosphatase mimics

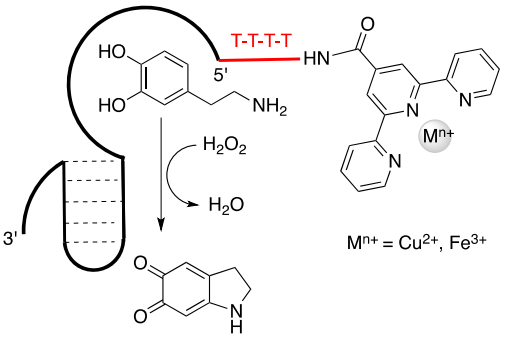
Hemin/G-quadruplex structures were well-known to act as horseradish peroxidase (HRP)-mimicking catalysts [40]. These unmodified short single-stranded oligonucleotides present sub-micromolar affinities toward the hemin metallofactor they were selected against. The DNA backbone is organized in a highly ordered G-quadruplex structure, acting as an apoprotein mimic by interacting in a supramolecular fashion through π–π stacking interactions with the hemin in charge of the enhanced (NADH) peroxidase activities [41]. Most are used for biosensing applications as the oxidation of substrates, either ABTS (2,2′-azinobis(3-ethylbenzothiazoline-6-sulfonic acid) by H_2O_2 to $ABTS^{\bullet+}$ or luminol induces a color change or generates chemoluminescence, respectively [42].

Willner's team developed an elegant concept based initially originated from combining a DNAzyme (in charge of the catalytic activity) with an aptamer with a short TATA sequence between the two nucleic parts. In this case, the aptamer is recognized by hydrogen bonding its substrate to oxidize, acting as a binding pocket. They anticipated that close proximity between partners would be

beneficial, inspired by RNA-cleaving DNAzymes in which the catalyst and RNA target substrate are kept in close proximity through Watson and Crick base pairing. Those "nucleoaptazymes" (as called) mimic the main features in the active site of native enzymes, namely catalysis and recognition. To the hemin/G-quadruplex horseradish peroxidase-mimicking DNAzyme was tethered one aptamer selected against L-DOPA to perform H_2O_2-mediated oxidation, leading to the chromophoric aminochrome with a 20-fold enhanced activity for the modified DNAzyme [43]. One of the key features is the facilitated release of the product after oxidation due to its decreased affinity toward the aptamer pocket, allowing a possible multiple turnover, as observed in native enzymes. A modest enantioselectivity was also observed for the oxidation of L-DOPA against its D-enantiomer, which was correlated with differences in affinities. However, the development of efficient unmodified nucleoaptazymes, as defined by Willner's team, faced two issues [44]. First, even close proximity between partners is not sufficient to ensure an optimum (catalytic) conformation. Second, despite a relatively furnished catalog of DNAzymes and aptamer sequences, finding relevant DNAzyme–aptamer couples that can perform a designed reaction is not trivial.

To overcome this former limitation, the same team proposed a new family of functionalized nucleoaptazymes by combining the dopamine binding aptamer (DBA) sequence while substituting the DNAzyme part by a metallic ligand such as Fe^{3+} or Cu^{2+}-terpyridine derivatives to perform the oxidation of L-dopamine, mimicking the catechol oxidase activity (Scheme 1) [45]. These oligonucleotides were functionalized by introducing a single metal complex ligand to chelate cupric or ferric ions in the right oxidation state because the native enzyme presents an oxidized Cu^{II}–Cu^{II} met state center with histidine side chains as proximal ligands in both metal binding sites [46]. Those FuON were synthetized through a convertible approach by reacting the 5′- or 3′-amino DBA sequence with the corresponding 4′-carboxylic acid of 2,2′:6′,2″-terpyridine, resulting in the formation of a stable amide bond between both partners.

They demonstrated the beneficial presence of a covalent flexible linker made of four thymidines (T4) between the metallic part and the aptamer binding site because oxidation does not occur efficiently when both sites are separated.



Scheme 1. Schematic representation of a nucleoaptazyme constituted by a Dopamine Binding Aptamer linked to a Cu^{2+}-terpyridine moiety at its 5′-end by a 4T sequence.

The best nucleoaptazymes present the T4 linker at the 5′-end with a k_{cat} of 40×10^{-4} s^{-1} and 267×10^{-4} s^{-1} for the Cu^{2+}- and Fe^{3+}-terpyridine complexes, respectively. This result highlights the impact on catalysis of the nature of the metal ion within the complex. As expected, the spatial organization adopted by the tethered partners also played a key role in the catalytic enhancement, as confirmed by the molecular dynamic simulations. An application of this concept was described for the bioconjugation of tyrosine derivatives using *N*-methyl luminol. The impact of this aptamer-assisted catalysis could be demonstrated by the 12-fold enhancement observed when the aptamer sequence binds to the tyrosine substrate [47].

The presence of metal ions is also crucial in the phosphatase mimics field because the native mechanism relies on the presence of a Zn^{2+} dinuclear site. The Zn^{2+} ions coordinate and activate the hydroxyl side chain of the serine that attacks the phosphorus atom of the monophosphoester bond while maintaining the right geometry for substrate binding through a hydrogen bonding network [48,49]. Several teams have investigated the design of these challenging DNAzymes as phosphatase mimics. Unfunctionalized DNAzymes were previously demonstrated by Silverman's team to be either in charge of the phosphorylation of the tyrosine side chain of a peptide substrate or able to catalyze the reverse process, i.e., the dephosphorylation step with multi-

ple turnovers and a k_{obs} of 0.19 min^{-1} [50,51]. Building on the nucleoaptazyme concept, Willner's team also designed phosphatase mimics with a bis-Zn^{2+}-pyridyl-salen-type complex tethered to an ATP aptamer [52]. Relying again on amine bond formation, an ATP aptamer (pdb 1AW4) bearing an amino function at the 5′- or 3′-end was covalently linked to a bis-Zn^{2+}-pyridyl-salen-type complex. In this case, an optimal distance composed of a dithymidine linker at the 3′-end between the two partners is sufficient to ensure both flexibility and cooperative proximity. As the promising measured k_{cat} (7 min^{-1}) shows a 10^3 times lower efficiency than the native enzyme, it appeared that the future FuON would still have room for improvement.

4. Functionalized oligonucleotides as serine protease mimics

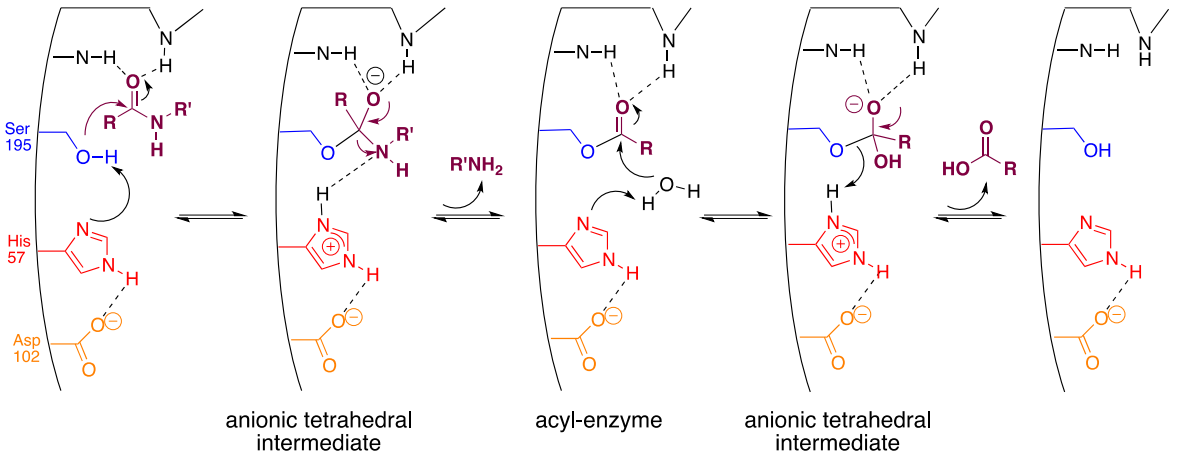
As illustrated with the last described biomimetic approach and for asymmetric chemical reactions, most of the literature on DNA-based catalysis relies on oligonucleotides using metallic co-factors [53]. Alongside some recent developments for Ribonuclease A mimics, other examples based only on organic co-factors in a protein-like fashion were developed for the DNA-based proteases field. We now describe the synthetic efforts in this area led by the functionalization of oligonucleotides with amino acid-like functions to hydrolyze the amide bond.

α-Chymotrypsin is one of the most representative members of the serine proteases family that uses three amino acids constitutive of its catalytic triad, serine 195 (Ser195), histidine 57 (His57), and aspartate 102 (Asp102), in a well-known cooperative mechanism (Scheme 2) [54].

Acting as a general base catalyst, the proximal imidazole ring from His57 enhances the nucleophilic character of the hydroxyl of Ser195, which is consequently sufficiently reactive to attack the carbonyl center of the peptide bond. The resulting anionic tetrahedral intermediate is stabilized in the so-called oxyanion hole by hydrogen bonds originating from the NH of the peptide backbone. It then collapses into a covalent acyl-enzyme complex, liberating the *N*-part of the peptide. The resulting (ester) acyl-enzyme is now more easily hydrolyzed than the previous amide bond. Thanks to the general acid catalysis of His57, under its protonated form stabilized by Asp 102, the active serine is finally regenerated following the liberation of the *C*-part of the peptide. Although the protease mimics field has been active for decades, no synthetic scaffolds have been able to match the enzyme efficiency [55]. Enzymatic catalysis and chemical hydrolysis are strictly two different processes. The first process introduces a two-step sequence to lower energy barriers through covalent acyl enzyme intermediates. The former is impaired by chemical factors such as amide bond stabilization by resonance, the combined poor nucleophily of water, and the poor ability of the amine to function as a good leaving group. Diverging from small organic catalysts or micellar systems, DNA was indeed expected to be a suitable scaffold because the nucleic part of the ribosome performs the reverse reaction, i.e., peptide bond formation [56]. Moreover, with the difference of peptide-based mimics that folding could neither be accurately predicted nor undergo any renaturation process, DNA has the unmatched ability to (re)fold into any designed secondary structures that could mimic the active site of enzymes. Finally, as stated earlier, phosphoramidite-based ON synthesis allows the precise introduction of amino acid residues alongside the nucleic backbone to reproduce the putative spatial organization of the catalytic triad.

Most of the efforts were concentrated on the SELEX approach, exploring de facto an expanded chemical space to screen for a functional nucleic catalyst. The RNAzyme field, which relies only on unmodified NTP, remains underexplored [57]. Dr. Hollenstein's work was aimed at the functionalization of dNTP, such as modified 7-deaza-2′-deoxyadenosine bearing at the 7-position a propargylamido-histamine moiety ($dA^{Hs}TP$) **8** to act as a histidine surrogate (Figure 4). Two 5-modified-2′-deoxyribopyrimidine-5′-triphosphates, either bearing a valeric acid moiety in the cytosine series ($dC^{Val}TP$) **9** or a pentynol arm in the uracil series ($dU^{POH}TP$) **10**, completed the set of functionalized dNTP, being mimics of aspartate and serine side chains, respectively [58].

They were obtained using a Sonogashira cross-coupling step starting from their commercial 5-iodo-deoxynucleoside counterparts, followed by the necessary protection/deprotection steps to introduce the 5′-triphosphate group required for the SE-



Scheme 2. Accepted mechanism for chymotrypsin-like serine proteases.

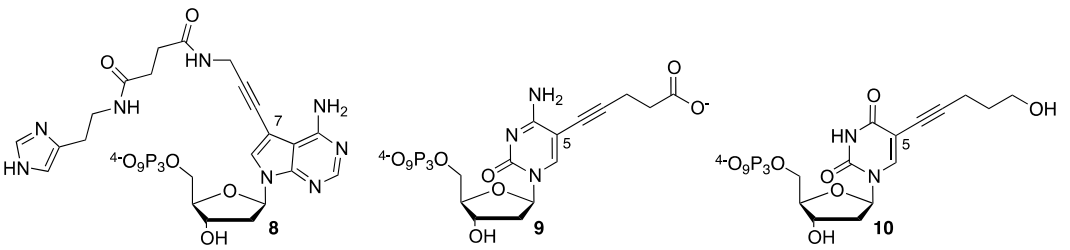
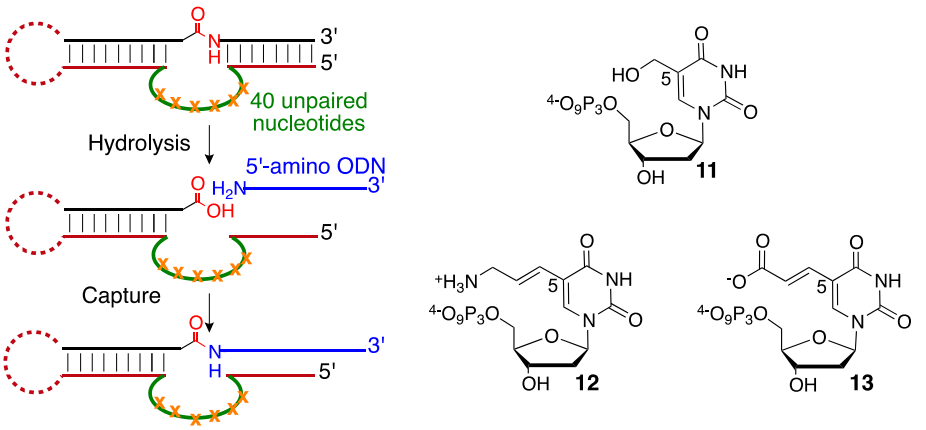


Figure 4. Functionalized dNTP with imidazole, carboxylate, or hydroxyl functions.

LEX process. Although the modified dNTPs were well tolerated by DNA polymerases, as demonstrated during primer extension experiments, they were less accepted during the PCR amplification step. Unfortunately, no SELEX attempt in the search for a DNAzyme with amidolytic properties was described.

The key player in the field was Silverman's group [59], which has devoted important efforts to developing DNA-based protease mimics and overcoming some pitfalls associated with the SELEX process. First, they were urged to design a sophisticated selection and "product capture" process to avoid the selection of DNA-cleaving DNAzyme, as previously observed [60]. By embedding the amide bond into a nucleic substrate facing the 40-nucleotide random loop region, they reported in 2013 the selection of unmodified DNAzymes performing esters hydrolysis (k_{obs} = 0.54 h^{-1}) leading to a moderate 240-fold rate enhancement compared with basal hydrolysis in the required presence of divalent cations such as Mg^{2+} [61]. Eventually, although a k_{obs} of 3.1 h^{-1} was measured for aromatic amide hydrolysis, no DNAzymes could be selected for cleaving the amide bond of aliphatic substrates. They then postulated that a better nucleophile than water might be required, such as an amino group. They confirmed their hypothesis in a subsequent publication starting from commercial 5-hydroxymethyl-, 5-aminoallyl- or (*E*)-5-(2-carboxyvinyl)-2′-deoxyuridine-5′-triphosphates **11**, **12**, and **13** (Scheme 3).

After rounds of SELEX optimization using one C5-substituted dUTP at the time, they were able to isolate the most potent DNAzymes decorated with hydroxyl groups with hydrolysis activity (k_{obs} = 0.1–0.2 h^{-1}) on an aliphatic amide embedded between two DNA fragments [62]. The two sequences leading to the most potent DNAzymes were re-synthetized with no modifications for control purposes. Surprisingly, they still presented substantial catalytic activity, although such unmodified se-



Scheme 3. Capture process (left) and C5-functionalized deoxyuridine 5′-triphosphates (right).

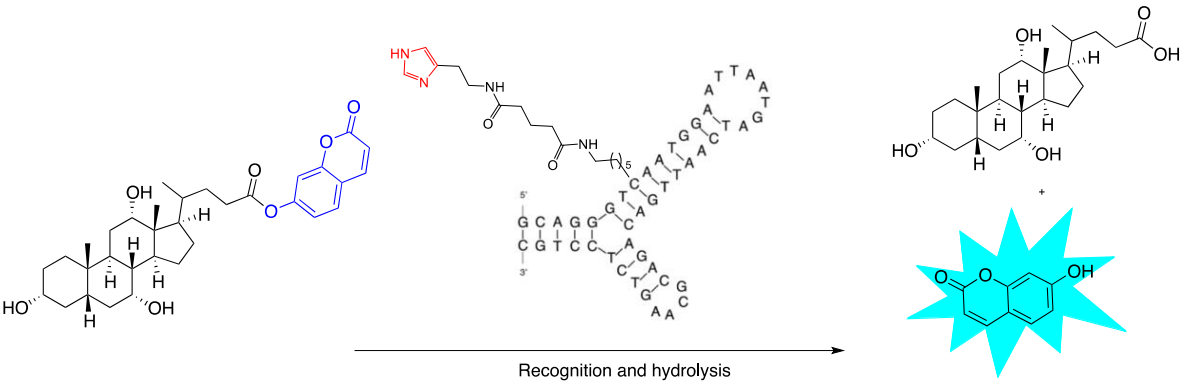
quences did not arise during the previous selections. This twist of events highlights the fact that SELEX is undoubtedly a powerful method, but it remains a tedious task prone to selection bias.

The complementary bottom-up approach is the SPS of functionalized ON based on the phosphoramidite strategy. Even if the chemical space explored cannot be as extensive as the one offered by SELEX, DNA-based structures mimicking the active site of the serine protease could be designed through careful and deliberate choice of the positions to be modified.

Attempts were made to design esterase mimics by solely using an imidazole pendant as the nucleophilic catalyst. 5-Ethylnyl-imidazole-deoxyuridine-3′-phosphoramidites were incorporated during SPS and organized within duplexes, leading to structures bearing up to three imidazole-modified positions. However, none of the duplexes displayed catalytic activity above that of imidazole, which was used as a buffer control. Relying also solely on imidazole modification, Gorin's team used an approach similar to the "nucleoaptazyme" strategy developed by Willner's laboratory [63]. They designed DNA-conjugated small-molecule catalysts (DCats) by functionalizing with an imidazole group two cholic acid aptamers folded in a three-way junction (3WJ) secondary structure (Scheme 4). At different positions alongside each arm of the junction, commercially available 2′-deoxycytidine or thymidine 3′-phosphoramidites bearing a C6-amino-linker at the C5 position were introduced and then conjugated with disuccinimidyl glutarate to histamine.

Up to ten DCats were obtained by modulating the position of the modification alongside the 3WJ to evaluate their ability to cleave after binding cholic acid-derived umbelliferone esters and the possible impact on the recognition process. This substrate was chosen because the cholic acid part is supposed to be recognized at the structure core, while the fluorogenic release of the umbelliferone allows the monitoring of the reaction evolution. The best DCat displayed a 100-fold enhancement esterase rate compared with the free equimolar imidazole but only when it was correctly folded, as corroborated by a turn on-off experiment with a toehold displacement.

However, only Madder's team developed 2′-functionalized deoxyuridine phosphoramidites **14**, **15**, and **16** bearing amino acid side chain-like residues such as suitably protected carboxylic acid, imidazole, and alcohol functions, respectively, which were tethered to the oligonucleotide via an amide bond (Figure 5) [64]. After solid-phase synthesis, up to three modifications could be introduced by strand, and the rigid but programable duplex structure was used to organize the residues along the major groove of the helix to obtain up to six different sequences, which were then combined into their corresponding duplexes [65]. Analysis of the thermal denaturation studies indicated that each introduction has a negative impact of −5 °C on the thermal value, although one mismatched duplex displayed a ΔTm of −5.3 °C (compared to an expected ΔTm of ≈−15 °C), compensating and largely overcoming the destabilization



Scheme 4. Schematic representation of the "DCat" approach based on a cholic acid aptamer linked to an imidazole moiety.

effect. This result could be explained by the fact that the imidazole and carboxylate moieties, facing each other, were able to interact and stabilize the structures through electrostatic interactions. Despite this chemical effort, no proteolytic activity has been reported.

As the structuration properties of DNA have not been exploited to their full potential, we decided to build a 3 way junction (3WJ) from three oligonucleotides partially complementary to each other, bearing at its core the three functional groups involved in the catalytic triad [66].

First, we established our proof of concept by relying on a versatile convertible approach by introducing organic functionalities not at the usual 2′- or 5-position but at the underestimated 5′-position. As the 2′- or C5-modifications direct the chemical handle toward the major groove within a B-DNA duplex, we demonstrated that, depending on the stereochemistry of the newly generated 5′-*C*-stereogenic center, the modification could point either toward the solvent or the minor groove when the absolute configuration of the 5′-carbon is *R*-assigned or *S*-assigned, respectively [67]. Second, we decided to organize our potential DNA-based proteases into a flexible secondary structure such as a 3WJ. This assembly, bearing three unpaired thymidines at its core, was expected to bring the three added modifications in close proximity with sufficient flexibility to avoid steric hindrance. After the efficient incorporation of the convertible 5′-*C*-(*S*)-propargyl thymidine phosphoramidite **17** into three different oligonucleotides, each strand underwent a post-synthetic copper(I)-catalyzed alkyne-azide cycloaddition (CuAAC) "click" reaction with the three azide derivatives bearing either a carboxylic, alcohol, or imidazole functions (Scheme 5). The FuON were then assembled in an equimolar mixture, forming a DNA 3WJ with a triply modified core.

The 3WJ assembly was confirmed by the presence of only one slow migrated band during non-denaturing polyacrylamide gel electrophoresis analyses and was also corroborated by the characteristic B-DNA signature obtained by circular dichroïsm spectroscopy. Thermal denaturation data were obtained to evaluate the contribution of each amino acid side chain-like residue to the overall stability of the secondary structure. As expected in the context of a polyanionic DNA backbone and different pKa values for each residue, the presence of a carboxylate moiety was modestly destabilizing (−1 °C), neutral for the alcohol moiety, and slightly stabilizing for the (cationic) imidazole residue (+1 °C). Taken together, unmodified, triply propargyl- or ^{Ser,His,Asp}3WJs displayed the same Tm value of 43 °C, indicating that the size of the three added cycloaddition products did not destabilize the resulting secondary structure. This decorated three-way junction ultimately failed to provide any amide bond hydrolysis. This could be explained by the lack of cooperative interactions between the three appended functionalities, as an alleged over-stabilization (that could originate from a close proximity of the functional groups) was not observed. Moreover, the stabilization of the anionic tetrahedral intermediate in the proximity of the negatively charged DNA backbone must be con-

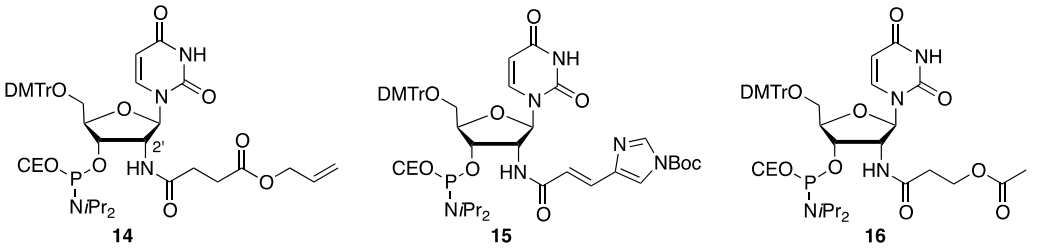
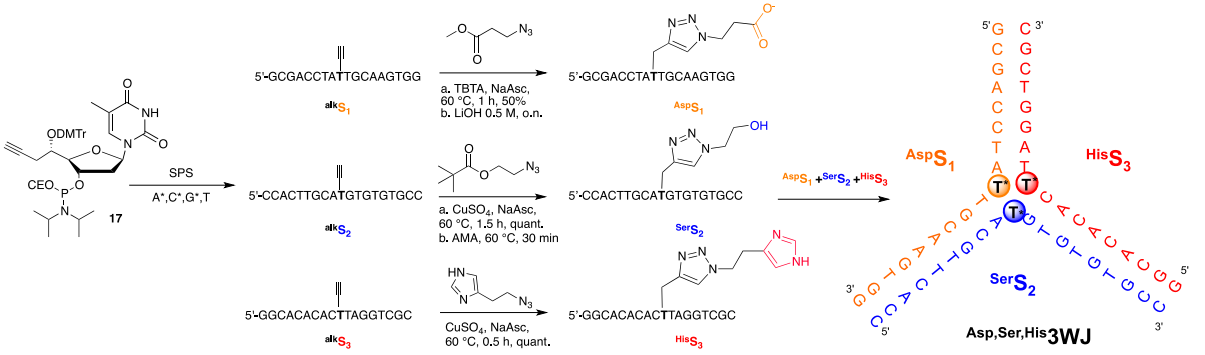


Figure 5. 2′-functionalized nucleotides bearing amino acid side chain-like residues and further introduced in oligonucleotides organized within a duplex structure.



Scheme 5. 5′-convertible nucleotide for the synthesis and functionalization of a 3WJ structure with amino acid side chain mimics at its core.

sidered. We are investigating new synthetic options to ensure cooperative interactions between the three functional groups to design a functional DNA-based protease that can cleave the amide bond.

5. Conclusion

Despite the exquisite evolutionary optimization of DNA to perform reactions thought to be only devoted to enzymes, organic chemists came into play to increase the DNA catalytic repertoire. As unmodified DNAzymes were first developed to cleave the phosphodiester bond of RNA, mimicking the catalytic activity of Ribonucleoase A, chemists rapidly took inspiration from the active site of such enzyme to functionalize oligonucleotides with organic groups reminiscent of the side chain of the amino acids involved in catalysis. Although a reduced catalytic efficiency was observed on non-nucleic acid substrates, FuON were also modified with metal complex ligands to confer either peroxidase or phosphatase properties or organic functions to mimic the amidolytic cleavage of the peptide bond, as exemplified in this review. Moreover, the recognition properties of DNA were also exploited to develop DNA catalysts capable of recognizing their target and performing the desired reaction in the so-called nucleoaptazyme approach.

Despite these promising results, more challenges lie in front of (bio)chemists in terms of matching the efficiency and selectivity of enzymes, with DNA catalysts being less potent than their enzymatic counterparts. Within the nucleoaptazyme approach, improving the binding properties would consequently increase the effective concentration of the substrate, leading to more efficient catalysis. Moreover, guidance could also be provided by in silico design [68] and algorithms [69] to rationally design functional FuON bearing, at the right positions, catalytic modifications or even to reach new stable secondary structures. Finally, advances must be made to improve each step of the process by developing new SELEX protocols and engineering

new enzymes to overcome the limitations of functionalized nucleotides that are sometimes poorly recognized by polymerases. Evolution could also be directed toward reactions not yet known to be catalyzed by DNA, for example, by modulating the modification status of RNA, as already explored with RNA-catalyzed RNA bio-orthogonal labeling reactions [70]. However, the catalog of enzymatic reactions remains underexploited, as exemplified by the challenging site-specific cleavage of DNA by DNAzymes [60,71]. Organic chemistry is associated with this ongoing optimization process with recent developments toward the expansion of the genetic alphabet [72], original orthogonal bioconjugation [73], or revisiting the solid phase synthesis process using P(V) chemistry to give ODN with optically pure phosphorothioate linkages [74]. Finally, starting from a single oligonucleotide displaying catalytic properties, recent advances in their supramolecular assembly in 3D organizations such as DNA tetrahedral [75] or origami nanostructures [76] or their capacity to mimic the enzymatic cascade process [77] will undoubtedly widen the scope and success of their applications in the foreseeable future.

Declaration of interests

The authors do not work for, advise, own shares in, or receive funds from any organization that could benefit from this article, and have declared no affiliations other than their research organizations.

References

[1] T. R. Cech, *Biosci. Rep.*, 1990, **10**, 239-261.
[2] S. Altman, *Biosci. Rep.*, 1990, **10**, 317-337.
[3] R. R. Breaker, G. F. Joyce, *Chem. Biol.*, 1994, **1**, 223-229.
[4] S. W. J. Santoro, F. Gerald, *Proc. Natl. Acad. Sci. USA*, 1997, **94**, 4262-4266.
[5] C. Tuerk, L. Gold, *Science*, 1990, **249**, 505-510.
[6] A. D. Ellington, J. W. Szostak, *Nature*, 1990, **346**, 818-822.
[7] R. Micura, C. Höbartner, *Chem. Soc. Rev.*, 2020, **49**, 7331-7353.
[8] J. Borggräfe, J. Victor, H. Rosenbach, A. Viegas, C. G. W. Gertzen, C. Wuebben, H. Kovacs, M. Gopalswamy, D. Riesner, G. Steger *et al.*, *Nature*, 2022, **601**, 144-149.
[9] H. Liu, X. Yu, Y. Chen, J. Zhang, B. Wu, L. Zheng, P. Haruehanroengra, R. Wang, S. Li, J. Lin *et al.*, *Nat. Commun.*, 2017, **8**, article no. 2006.
[10] A. Ponce-Salvatierra, K. Wawrzyniak-Turek, U. Steuerwald, C. Höbartner, V. Pena, *Nature*, 2016, **529**, 231-234.
[11] K. Svehlova, O. Lukšan, M. Jakubec, E. A. Curtis, *Angew. Chem. Int. Ed.*, 2022, **61**, article no. e202109347.
[12] J. L. Wilcox, P. C. Bevilacqua, *J. Am. Chem. Soc.*, 2013, **135**, 7390-7393.
[13] M. Chandra, S. K. Silverman, *J. Am. Chem. Soc.*, 2008, **130**, 2936-2937.
[14] F. R. Ortigão, H. Rösch, H. Selter, A. Fröhlich, A. Lorenz, M. Montenarh, H. Seliger, *Antisense Res. Dev.*, 1992, **2**, 129-146.
[15] K. Duffy, S. Arangundy-Franklin, P. Holliger, *BMC Biol.*, 2020, **18**, article no. 112.
[16] Y. Wang, Y. Wang, D. Song, X. Sun, Z. Li, J.-Y. Chen, H. Yu, *Nat. Chem.*, 2022, **14**, 350-359.
[17] A. I. Taylor, V. B. Pinheiro, M. J. Smola, A. S. Morgunov, S. Peak-Chew, C. Cozens, K. M. Weeks, P. Herdewijn, P. Holliger, *Nature*, 2015, **518**, 427-430.
[18] Y. Wang, A. K. Ngor, A. Nikoomanzar, J. C. Chaput, *Nat. Commun.*, 2018, **9**, article no. 5067.
[19] A. I. Taylor, C. J. K. Wan, M. J. Donde, S.-Y. Peak-Chew, P. Holliger, *Nat. Chem.*, 2022, **14**, 1295-1305.
[20] D. Kong, M. Movahedi, Y. Mahdavi-Amiri, W. Yeung, T. Tiburcio, D. Chen, R. Hili, *ACS Synth. Biol.*, 2020, **9**, 43-52.
[21] Y. Singh, P. Murat, E. Defrancq, *Chem. Soc. Rev.*, 2010, **39**, 2054-2070.
[22] S. Diafa, M. Hollenstein, *Molecules*, 2015, **20**, article no. 16643.
[23] C. Chardet, C. Payrastre, B. Gerland, J.-M. Escudier, *Molecules*, 2021, **26**, article no. 5925.
[24] R. T. Raines, *Chem. Rev.*, 1998, **98**, 1045-1066.
[25] S. W. Santoro, G. F. Joyce, K. Sakthivel, S. Gramatikova, C. F. Barbas, *J. Am. Chem. Soc.*, 2000, **122**, 2433-2439.
[26] D. M. Perrin, T. Garestier, C. Hélène, *J. Am. Chem. Soc.*, 2001, **123**, 1556-1563.
[27] L. Lermer, Y. Roupioz, R. Ting, D. M. Perrin, *J. Am. Chem. Soc.*, 2002, **124**, 9960-9961.
[28] A. V. Sidorov, J. A. Grasby, D. M. Williams, *Nucleic Acids Res.*, 2004, **32**, 1591-1601.
[29] Y. Wang, E. Liu, C. H. Lam, D. M. Perrin, *Chem. Sci.*, 2018, **9**, 1813-1821.
[30] T. Gourlain, A. Sidorov, N. Mignet, S. J. Thorpe, S. E. Lee, J. A. Grasby, D. M. Williams, *Nucleic Acids Res.*, 2001, **29**, 1898-1905.
[31] M. Hollenstein, *Curr. Opin. Chem. Biol.*, 2019, **52**, 93-101.
[32] P.-J. J. Huang, J. Liu, *ChemistryOpen*, 2020, **9**, 1046-1059.
[33] Y. Jiao, Y. Shang, N. Li, B. Ding, *iScience*, 2022, **25**, article no. 104018.
[34] S. Paul, A. Wong, L. T. Liu, D. M. Perrin, *ChemBioChem*, 2022, **23**, article no. e202100600.
[35] S. Du, Y. Li, Z. Chai, W. Shi, J. He, *RSC Adv.*, 2020, **10**, 19067-19075.
[36] W. Zhang, Y. Li, S. Du, Z. Chai, J. He, *Bioorg. Med. Chem. Lett.*, 2021, **48**, article no. 128234.
[37] R. J. Lake, Z. Yang, J. Zhang, Y. Lu, *Acc. Chem. Res.*, 2019, **52**, 3275-3286.
[38] P.-J. J. Huang, D. de Rochambeau, H. F. Sleiman, J. Liu, *Angew. Chem. Int. Ed.*, 2020, **59**, 3573-3577.
[39] E. B. Pimentel, T. M. Peters-Clarke, J. J. Coon, J. D. Martell, *J. Am. Chem. Soc.*, 2021, **143**, 21402-21409.
[40] P. Travascio, Y. Li, D. Sen, *Chem. Biol.*, 1998, **5**, 505-517.
[41] E. Golub, R. Freeman, I. Willner, *Angew. Chem. Int. Ed.*, 2011, **50**, 11710-11714.

[42] H. Peng, A. M. Newbigging, Z. Wang, J. Tao, W. Deng, X. C. Le, H. Zhang, *Anal. Chem.*, 2018, **90**, 190-207.
[43] E. Golub, H. B. Albada, W.-C. Liao, Y. Biniuri, I. Willner, *J. Am. Chem. Soc.*, 2016, **138**, 164-172.
[44] M. Vázquez-González, Z. Zhou, Y. Biniuri, B. Willner, I. Willner, *Biochemistry*, 2021, **60**, 956-965.
[45] Y. Biniuri, B. Albada, M. Wolff, E. Golub, D. Gelman, I. Willner, *ACS Catal.*, 2018, **8**, 1802-1809.
[46] C. Gerdemann, C. Eicken, B. Krebs, *Acc. Chem. Res.*, 2002, **35**, 183-191.
[47] S. Wintermans, J. F. Keijzer, M. Dros, H. Zuilhof, B. Albada, *ChemCatChem*, 2021, **13**, 4618-4624.
[48] K. M. Holtz, E. R. Kantrowitz, *FEBS Lett.*, 1999, **462**, 7-11.
[49] Y. Shi, *Cell*, 2009, **139**, 468-484.
[50] S. M. Walsh, A. Sachdeva, S. K. Silverman, *J. Am. Chem. Soc.*, 2013, **135**, 14928-14931.
[51] J. Chandrasekar, S. K. Silverman, *Proc. Natl. Acad. Sci.*, 2013, **110**, 5315-5320.
[52] Y. Biniuri, Z. Shpilt, B. Albada, M. Vazquez-Gonzalez, M. Wolff, C. Hazan, E. Golub, D. Gelman, I. Willner, *ChemBioChem*, 2020, **21**, 53-58.
[53] J. H. Yum, H. Sugiyama, S. Park, *Chem. Rec.*, 2022, **22**, article no. e202100333.
[54] L. Hedstrom, *Chem. Rev.*, 2002, **102**, 4501-4524.
[55] M. D. Nothling, Z. Xiao, A. Bhaskaran, M. T. Blyth, C. W. Bennett, M. L. Coote, L. A. Connal, *ACS Catal.*, 2019, **9**, 168-187.
[56] D. A. Hiller, V. Singh, M. Zhong, S. A. Strobel, *Nature*, 2011, **476**, 236-239.
[57] S. Ameta, A. Jaschke, *Chem. Sci.*, 2013, **4**, 957-964.
[58] M. Hollenstein, *Org. Biomol. Chem.*, 2013, **11**, 5162-5172.
[59] S. K. Silverman, *Trends Biochem. Sci.*, 2016, **41**, 595-609.
[60] M. Chandra, A. Sachdeva, S. K. Silverman, *Nat. Chem. Biol.*, 2009, **5**, 718-720.
[61] B. M. Brandsen, A. R. Hesser, M. A. Castner, M. Chandra, S. K. Silverman, *J. Am. Chem. Soc.*, 2013, **135**, 16014-16017.
[62] C. Zhou, J. L. Avins, P. C. Klauser, B. M. Brandsen, Y. Lee, S. K. Silverman, *J. Am. Chem. Soc.*, 2016, **138**, 2106-2109.
[63] M. Flanagan, A. E. Arguello, D. E. Colman, J. Kim, J. N. Krejci, S. Liu, Y. Yao, Y. Zhang, D. J. Gorin, *Chem. Sci.*, 2018, **9**, 2105-2112.
[64] M. Catry, A. Madder, *Molecules*, 2007, **12**, 114-129.
[65] M. A. Catry, V. Gheerardijn, A. Madder, *Bioorg. Chem.*, 2010, **38**, 92-97.
[66] C. Addamiano, B. Gerland, C. Payrastre, J. M. Escudier, *Molecules*, 2016, **21**, article no. 1082.
[67] V. Banuls, J.-M. Escudier, C. Zedde, C. Claparols, B. Donnadieu, H. Plaisancié, *Eur. J. Org. Chem.*, 2001, 4693-4700.
[68] M. Hendling, I. Barišić, *Comput. Struct. Biotechnol. J.*, 2019, **17**, 1056-1065.
[69] R. Mohammadi-Arani, F. Javadi-Zarnaghi, P. Boccaletto, J. M. Bujnicki, A. Ponce-Salvatierra, *Nucleic Acids Res.*, 2022, **50**, W261-W265.
[70] M. Ghaem Maghami, S. Dey, A.-K. Lenz, C. Höbartner, *Angew. Chem. Int. Ed.*, 2020, **59**, 9335-9339.
[71] V. Dhamodharan, S. Kobori, Y. Yokobayashi, *ACS Chem. Biol.*, 2017, **12**, 2940-2945.
[72] C. A. Jerome, S. Hoshika, K. M. Bradley, S. A. Benner, E. Biondi, *Proc. Natl. Acad. Sci.*, 2022, **119**, article no. e2208261119.
[73] F. Liu, H. Wang, S. Li, G. A. L. Bare, X. Chen, C. Wang, J. E. Moses, P. Wu, K. B. Sharpless, *Angew. Chem. Int. Ed.*, 2019, **58**, 8029-8033.
[74] K. W. Knouse, D. T. Flood, J. C. Vantourout, M. A. Schmidt, I. M. McDonald, M. D. Eastgate, P. S. Baran, *ACS Cent. Sci.*, 2021, **7**, 1473-1485.
[75] H. Guan, S. Yang, C. Zheng, L. Zhu, S. Sun, M. Guo, X. Hu, X. Huang, L. Wang, Z. Shen, *Talanta*, 2021, **233**, article no. 122543.
[76] J. Wang, L. Yue, Z. Li, J. Zhang, H. Tian, I. Willner, *Nat. Commun.*, 2019, **10**, article no. 4963.
[77] Y. Wu, J. Li, K. Quan, X. Meng, X. Yang, J. Huang, K. Wang, *Chem. Commun.*, 2020, **56**, 10163-10166.

Comptes Rendus
Chimie
2023, Vol. 26, Special Issue S3, p. 119-137
https://doi.org/10.5802/crchim.272

Breaking Barriers in Chemical Biology – Toulouse 2022

Structural characterization of stem cell factors Oct4, Sox2, Nanog and Esrrb disordered domains, and a method to detect phospho-dependent binding partners

Chafiaa Bouguechtouli [a], Rania Ghouil [a], Ania Alik [a], Florent Dingli [b], Damarys Loew [b] and Francois-Xavier Theillet [*, a]

[a] Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198, Gif-sur-Yvette, France

[b] Institut Curie, PSL Research University, Centre de Recherche, CurieCoreTech Spectrométrie de Masse Protéomique, Paris cedex 05, France

Current addresses: Structural Motility, Institut Curie, Paris Université Sciences et Lettres, Sorbonne Université, CNRS UMR144, 75005 Paris, France (C. Bouguechtouli), Université de Paris, Institut Cochin, CNRS UMR8104, INSERM U1016, Paris, France (A. Alik)

E-mail: francois-xavier.theillet@cnrs.fr (F.-X. Theillet)

Abstract. The combined expression of a handful of pluripotency transcription factors (PluriTFs) in somatic cells can generate induced pluripotent stem cells (iPSCs). Here, we report the structural characterization of disordered regions contained in four important PluriTFs, namely Oct4, Sox2, Nanog and Esrrb. Moreover, many post-translational modifications (PTMs) have been detected on PluriTFs, whose roles are not yet characterized. To help in their study, we also present a method (i) to produce well-characterized phosphorylation states of PluriTFs, using NMR analysis, and (ii) to use them for pull-downs in stem cell extracts analyzed by quantitative proteomics to detect potential Sox2 binders.

Keywords. Pluripotency transcription factors, Intrinsically disordered proteins, Post-translational modifications, Kinases, NMR, Proteomics, Quantitative mass spectrometry.

Funding. CNRS and CEA-Saclay, French Infrastructure for Integrated Structural Biology (FRISBI, https://frisbi.eu/, grant number ANR-10-INSB-05-01) and French National Research Agency (ANR; research grants ANR-14-ACHN-0015 and ANR-20-CE92-0013), IR INFRANALYTICS FR2054, "Région Ile-de-France" and "Fondation pour la Recherche Médicale".

Manuscript received 3 March 2023, revised 9 November 2023, accepted 10 November 2023.

1. Introduction

The possibility of reprogramming somatic cells to an induced pluripotency state was revealed in the 2000s, giving great hopes in the fields of Biology and Medicine [1–3]. Induced pluripotent stem cells (iPSCs) and embryonic stem cells (ESCs) are characterized by the active state of a pluripotency network, whose core comprises the pluripotency transcription factors (PluriTFs) Oct4, Sox2, Nanog and Esrrb (OSNE). These bind to enhancer sequences and thus activate or repress, or even "bookmark" during mitosis, a wealth of genes related to pluripotency or

* Corresponding author

cell differentiation [4–9]. Consistently, their misregulation correlates with cancer malignancy and stemness [10–14].

Comprehensive structural descriptions of OSNE are still missing to the best of our knowledge. The folded DNA-binding domains (DBDs) of OSNE have been structurally characterized in complex with their DNA target sequences [15–19] together with the ligand-binding domain (LBD) of Esrrb [20]. Recent studies have depicted even splendid structures of Oct4's and Sox2's DBDs bound to nucleosomes, hence deciphering their "pioneer factor" abilities [21–29]. Another structure of Sox2 bound to the importin Impα3 has also been published, showing how its two Nuclear Localization Sequences (NLSs) flanking the DBD are involved in Sox2 nuclear import [30]. The other segments of OSNE have been predicted to be intrinsically disordered regions of proteins (IDRs) [31], i.e., they should have no stable tertiary fold when isolated [32–36].

These IDRs appear to have important roles in binding partners involved in epigenetic reprogramming, chromatin reorganization, and in recruiting transcription or repression machineries [5,37–41]. These functions are poorly understood, and, to the best of our knowledge, no experimental characterization of the structural behavior of these regions in N- and C-terminal of DBDs has been released yet. Recent studies have shown that C-terminal regions of Oct4 and Sox2 are important for their reprogramming capacities [42,43], notably by contributing to the engagement in molecular phase-separated condensates with the Mediator complex [44]. More generally, the activating or repressive activities of IDRs of transcription factors (TFs) have been scarcely studied at the structural level: these segments are thought to contain hydrophobic patches flanked by acidic amino acids, which favors DNA-binding specificity, phase separation, and low-specificity interactions, notably with the Mediator subunit Med15 [44–53]; more specific interactions have been described in some cases [45,54,55].

Post-translational modifications (PTMs) add a layer of complexity by often regulating IDRs' interactions [32,33,56,57] and notably TFs' activity [58–61]. PTMs are classical carriers of cell signaling by regulating the stability and the interactions of proteins. An increasing number of PTMs have been described on OSNE's IDRs in the recent years [5,37,38,62–77], notably phosphorylation by cyclin-dependent kinases (CDKs) [66,78–87] or by mitogen-activated protein kinases (MAPKs) [83,88,89], or their complementary Ser/Thr O-GlcNAcylation by OGT [65,90–96].

In order to prompt future studies on this topic, we describe here a feasibility study (i) for producing well-characterized samples made of post-translationally modified IDRs of PluriTFs and (ii) to use these as baits in pull-down assays for detecting PTMs' related binding partners. Hence, we characterized some of the phosphorylation reactions of Esrrb and Sox2 by p38α/β, Erk2 and Cdk1/2. Then, we showed that biotinylated chimera of Sox2 and Esrrb coupled to an AviTag peptide could be attached to streptavidin-coated beads. Finally, we loaded truncated segments of the C-terminal IDR of Sox2 (phosphorylated or not) on these beads, and exposed them to extracts of mouse ESCs (mESCs) in pull-down assays, which we analyzed using quantitative mass spectrometry-based proteomics. Among the quantified (phospho-)Sox2 binders, we verified the phospho-dependent interaction between the proline cis-trans isomerase Pin1 and Sox2 using NMR spectroscopy.

2. Material and methods

2.1. *Production of recombinant fragments of Oct4, Sox2, Nanog and Esrrb*

We used human protein sequences, unless specified. Codon-optimized (for expression in *Escherichia coli*) genes coding for human Oct4(aa1–145) and Oct4(aa286–380) were synthesized in the context of larger genes coding for Tev–Oct4(aa1–145)–Tev–GB1 and Tev–Oct4(aa286–380)–Tev–GB1 by Genscript and cloned into pET-41a(+) vector between SacII and HindIII restriction sites, hence permitting the expression of GST–His6–Tev1–Oct4(aa1–145)–Tev2–GB1 and GST–His6–Tev1–Oct4(aa286–380)–Tev2–GB1; Tev1 and Tev2 are the heptapeptide ENLYFQG cleavage site of the TEV protease, Tev2 is separated by GAGGAGG from GB1 (T2Q variant of the immunoglobulin binding domain B1 of the protein G from group G Streptococcus [97,98]). The C-terminal GB1 tag was added to avoid any C-terminal proteolysis of the IDR of interest during the expression and the first purification steps; we did not test constructs without this supplementary folded domain,

whose necessity for the stability of the IDR is thus not proven.

The same rationale (cDNA synthesis, cloning, vectors, chimera constructs) was used for producing Nanog(aa154–305), Nanog(aa154–215), Nanog(aa154–272), Nanog(aa154–305_C185A-C227A-C243A-C251A), Nanog(aa154–272_C185A-C227A-C243A-C251A), and a very similar rationale (chimera constructs missing the C-terminal Tev2–GB1) for Sox2(aa1–42), Sox2(aa115–317_C265A), Sox2(aa115–187), Sox2(aa115–236), Sox2(aa115–282_C265A), Esrrb(aa1–102_C12A-C72A-C91A), Esrrb(aa1–102_C12A-C91A), Nanog(aa1–85) (this latter was cloned in the MfeI/HindIII restriction sites from pET-41a(+)).

The recombinant production and the purification of the protein constructs followed the procedures described previously [99], using the soluble fraction of bacterial lysates, except for the constructs containing the Sox2 C-terminal fragments. These latter constructs were recovered from the insoluble fractions of the lysates and resolubilized in 8 M urea; these were submitted to a His-tag purification in urea, and the last size-exclusion chromatography (SEC) had to be carried out in 2 M urea, which avoided clogging of the column and permitted obtaining regular elution peak widths (these were otherwise extremely broad, up to 100 mL for the longest Sox2(aa115–317_C265A) construct). The samples were concentrated and stored at −20 °C, and thawed just before the NMR experiments. The Sox2 samples containing 2 M urea were submitted to 2–3 cycles of concentration/dilution in Hepes at 20 mM, NaCl at 75 mM to generate samples in urea at 0.25 or 0.125 M.

All purification steps were carefully carried out at 4 °C; protein eluates from every purification step were immediately supplemented with protease inhibitors (EDTA-free cOmplete, Roche) (together with DTT at 10 mM for cysteine-containing protein constructs), before being submitted to a concentration preparing the next purification step.

Chimera constructs of Sox2's and Esrrb's IDR fragments containing a 15-mer peptide AviTag GLNDIFEAQKIEWHE were produced using procedures similar to those described earlier for OSNE constructs. The construct Sox2(aa234–317)–AviTag–His6 was soluble and did not require to be purified in urea.

More details about the production of OSNE peptides are given in the Supplementary Material.

2.2. *Production of the biotin ligase BirA and specific biotinylation of the AviTag–peptide chimera*

The biotin ligase BirA was produced using recombinant production in *E. coli* BL21(DE3)Star transformed with a pET21-a(+) plasmid containing a gene coding for BirA cloned at EcoRI and HindIII restriction sites. pET21a-BirA was a gift from Alice Ting (Addgene plasmid #20857) [100]. The expression was carried out overnight at 20 °C in a Luria–Bertani culture medium. The construct contained a His6 tag in C-terminal and was purified using a two-step purification procedure including a His-trap followed by a SEC. Details about the production of BirA are given in the Supplementary Material.

The biotinylation was executed using a rationale inspired by a published protocol [101], at room temperature during 90 min, in samples containing the AviTag–chimera of interest at 100 μM and BirA at 0.7 μM in a buffer containing ATP at 2 mM, biotin at 600 μM, $MgCl_2$ at 5 mM, DTT at 1 mM, HEPES at 50 mM, NaCl at 150 mM, protease inhibitors (final concentration 1×, EDTA-free cOmplete, Roche), at pH 7.0. To remove some possible proteolyzed peptides and BirA, the biotinylated constructs were purified using a SEC in a column (Superdex 16/60 75 pg, Cytiva) preequilibrated with a buffer containing phosphate at 20 mM, NaCl at 150 mM at pH 7.4 (buffer called thereafter Phosphate Buffer Saline, PBS). The eluted fractions of interest were concentrated and stored at −20 °C.

2.3. *Assignment of NMR signals from OSNE fragments and structural propensities*

The assignment strategy was the same as in previous reports from our laboratory [99]. The ^{15}N relaxation data were recorded and analyzed according to the methods described in previous reports [102]. Details are given in the Supplementary Material.

Disorder prediction was calculated using the ODINPred website (https://st-protein.chem.au.dk/odinpred) [103]. Experimental secondary structure propensities of unmodified OSNE peptides were obtained using the neighbor-corrected structural propensity calculator ncSCP [104,105] (http://www.protein-nmr.org/, https://st-protein02.chem.au.dk/ncSPC/) from the experimentally determined, DSS-

referenced Cα and Cβ chemical shifts as input, with a correction for Gly–Pro motifs (−0.77 ppm instead of −2.0 ppm) [106]. We also used the δ2D method to get requested verifications of the experimental secondary structure propensities [107]. Some signals were too weak in 3D spectra from Sox2(aa115–317_C265A) recorded at 950 MHz, and their chemical shifts were not defined. In these cases, chemical shifts from 3D spectra of Sox2(115-236) or His6–AviTag–Sox2(aa234–317_C265A) were used to complete the lists of chemical shifts used to calculate the chemical shift propensities shown in Figure 2.

$^{1}H_{N}/^{15}N/^{13}Ca/^{13}Cb/^{13}CO$ NMR assignments of OSNE peptides, together with the corresponding experimental details, have been deposited in the Biological Magnetic Resonance Data Bank (BMRB) with accession numbers 51534 (Sox2_aa1–42), 51717 (Esrrb_aa1–102), 51756 (Oct4_aa1–145), 51758 (Oct4_aa286–360), 51782 (His6–AviTag–Sox2_aa234–317_C265A), and 51780 (Nanog_aa1–85).

2.4. *NMR monitoring of phosphorylation reactions and production of phosphorylated peptides*

We performed the phosphorylation kinetics presented in Figure 4a using commercial recombinant kinases GST–p38β at 10 μg/mL (Sigma-Aldrich, ref. B4437), GST–Erk2 at 20 μg/mL (Sigma-Aldrich, ref. E1283), GST–Cdk1/CyclinA2 at 20 μg/mL (Sigma-Aldrich, stock ref. C0244), and GST–Cdk2/CyclinA2 at 20 μg/mL (Sigma-Aldrich, ref. C0495). Then, we used kinases produced in-house in *E. coli*, using plasmids containing optimized genes coding for p38α(aa1–360, full-length) and Erk2(aa8–360); these were produced, activated, and purified in house as described previously [99]; in-house p38α was used at 40 μg/mL for the experiments shown in Figure 4. Indeed, the limited activities and high costs of commercial kinases motivated us to develop in-house capacities in kinase production. p38α was the most accessible to produce among the MAPKs and CDKs; we produced it and activated it using recombinant MKK6.

Phosphorylation reactions were carried out using ^{15}N-labeled IDRs at 50 μM, in Hepes 20 mM, NaCl 50 mM, DTT or TCEP at 4 mM, ATP 1.5 mM, $MgCl_2$ at 5 mM, protease inhibitors (Roche), 7.5% D_2O, pH 6.8 at 25 °C in 100 μL using 3 mm diameter Shigemi tubes. We monitored the phosphorylation kinetics by recording time series of $^{1}H–^{15}N$ SOFAST-HMQC spectra at 600 or 700 MHz, and by quantifying the NMR signal intensities of the disappearing unphospho- and appearing phospho-residues. We applied the methods that we described in earlier publications [108–111]. More details are given in the Supplementary Material.

2.5. *Pull-down assays*

The mESCs extracts were obtained from mESCs cultured in the conditions previously described [112]. Homogeneous extracts were obtained using DNA shearing by sonication in the presence of benzonase, as described by Gingras and colleagues [113].

The pull-down assays were executed using 25 μL of streptavidin-coated magnetic beads (Magbeads streptavidine, Genscript) loaded with 1 nmol of the biotinylated bait-peptides of interest. These were incubated for one hour at room temperature with mESCs extracts, washed in PBS and eluted using a 2× Laemmli buffer. Details are given in the Supplementary Material.

2.6. *Mass spectrometry-based proteomics analysis of pull-down assays*

The pull-down samples were treated on-beads by trypsin/LysC (Promega). The resulting peptides were loaded and separated on a C18 column for online liquid chromatography performed with an RSLCnano system (Ultimate 3000, Thermo Scientific) coupled to an Orbitrap Fusion Tribrid mass spectrometer (Thermo Scientific). Maximum allowed mass deviation was set to 10 ppm for monoisotopic precursor ions and 0.6 Da for MS/MS peaks. The resulting files were further processed using myProMS v3.9.3 (https://github.com/bioinfo-pf-curie/myproms; Poullet et al. [114]). False-discovery rate (FDR) was calculated using Percolator [115] and was set to 1% at the peptide level for the whole study. Label-free quantification was performed using extracted-ion chromatograms (XICs) of peptides, computed with MassChroQ [116] v.2.2.1. The complete details are given in the Supplementary Material.

2.7. *Recombinant production of Pin1 and NMR analysis of its interaction with Sox2 or phospho-Sox2*

The plasmid containing the gene coding for the Pin1–WW domain was a kind gift from Isabelle Landrieu. The production was executed according to the previously published protocol [117]. NMR analysis of the binding with phospho-Sox2(aa115–240) was performed at 283 K and pH 7.0 with the GST–Pin1–WW construct and ^{15}N-labeled Sox2(aa115–240) mixed in stoichiometric proportions, either at 50 or 10 μM for non-phospho- and phospho-Sox2, respectively. The details on the NMR acquisition, processing, and analysis are given in the Supplementary Information.

3. Results

3.1. *Structural characterization of the N- and C-terminal regions of Oct4*

We produced and purified protein constructs containing the fragments of human Oct4(aa1–145) and Oct4(aa286–360), which were both predicted to be mostly disordered (Supplementary material 3.2.3). The 2D ^{1}H–^{15}N HSQC NMR spectra showed cross-peaks in the region where random coil peptides resonances are usually found (Figure 1). We assigned the backbone NMR signals of $^{1}H_N$, ^{15}N, $^{13}C\alpha$, $^{13}C\beta$, ^{13}CO for both segments, which permitted to calculate experimentally derived secondary structure propensities (Figures 1c, f, Supplementary Figure S4). We did not find any sign of a stable secondary structure, the highest α-helical propensities reaching about 25% in short stretches of about 5 consecutive amino acids. Hence, we verified experimentally that these N- and C-terminal fragments of human Oct4 are IDRs. We noticed that one motif RTWLSF (aa33–38) generates cross-peaks out of the random coil area in the 2D ^{1}H–^{15}N HSQC, and its chemical shifts reveal the strongest α-helical propensity (about 25% according to two distinct software products, see Figures 1c, f, Supplementary Figure S4). This might typically correspond to an interaction site: IDRs' binding motifs often adopt secondary structure in complexes, whose formation is energetically favored by local conformational preferences for the bound structure in the free state [57,106,118].

We can highlight the fact that Oct4's IDRs contain a high density of prolines, which are not directly detectable in the present $^{1}H_N$-detected experiments, even though most of the $^{13}C\alpha$, $^{13}C\beta$, and ^{13}CO resonances were characterized via HNCAB and HNCO experiments. We have shown previously that the ^{13}C-detected experiments $^{13}C\alpha$ ^{13}CO permitted observing all these Pro residues in Oct4(aa1–145) [99], whose chemical shifts were those of random coil peptides.

3.2. *Structural characterization of the N- and C-terminal regions of Sox2*

We produced and purified peptide fragments of human Sox2, namely Sox2(aa1–42), Sox2(aa115–187), Sox2(aa115–236), Sox2(aa115–282), Sox2(aa234–317_C265A), and Sox2(aa115–317_C265A). We also produced and purified chimera peptides His6–AviTag–Sox2(aa115–240) and His6–AviTag–Sox2(aa234–317_C265A). These were all predicted to be disordered (Supplementary material 3.3.3).

We had solubility issues with all of them but Sox2(aa1–42), Sox2(aa234–317_C265A), and His6–AviTag–Sox2(aa234–317_C265A). We had to recover these troublesome peptides from the insoluble fraction of the bacterial extract after overexpression at 37 °C. We even had to carry out our final SEC purification step in a buffer containing urea at 2 M (at 4 °C) for Sox2(aa115–236), Sox2(aa115–282), His6–AviTag–Sox2(aa115–240), and Sox2(aa115–317_C265A). The assignments of these latter constructs were achieved in 0.25–0.5 M urea, after executing 2 to 3 concentration–dilution steps. Aggregates were forming during the acquisition, which made the assignment rather painful. This behavior correlated with liquid–liquid phase separation (LLPS) propensities (Supplementary Figure S5), which we observed a few months before such a behavior was reported by Young and collaborators [44]. The assignment of Sox2(aa115–317_C265A) was possible only at 950 MHz with the help of the previously assigned smaller fragments Sox2(aa115–236) and His6–AviTag–Sox2(aa234–317_C265A). Some stretches of amino acids were particularly difficult to observe in 3D spectra, e.g., the region aa160–185, because of an apparent fast T2 relaxation. We may investigate these phenomena in later reports.

We observed cross-peaks in the 2D ^{1}H–^{15}N HSQC NMR spectra that were all resonating in the spectral region of random coil peptides' resonances

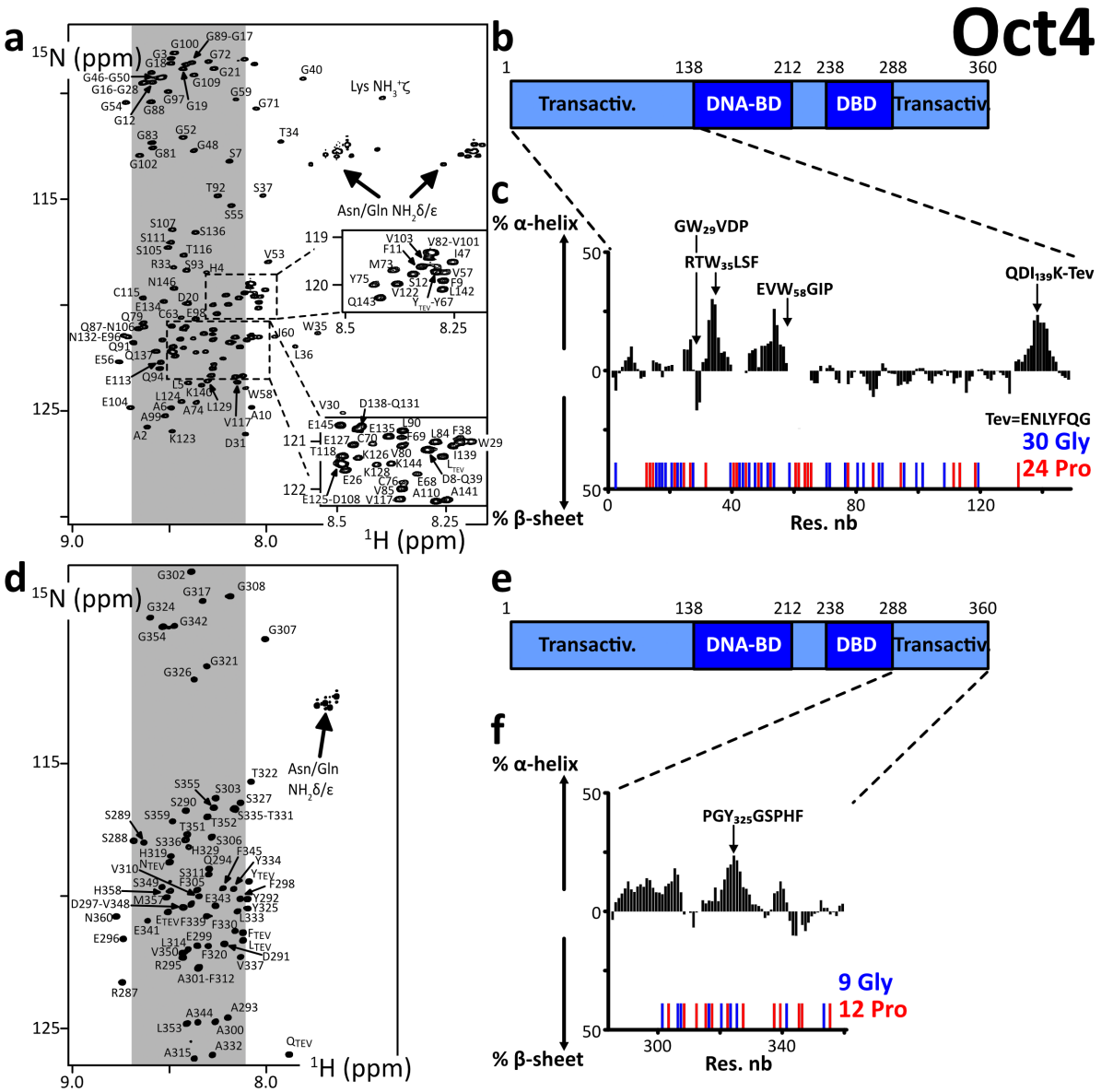


Figure 1. (a,d) 2D ^{1}H–^{15}N HSQC spectra of the N- and C-terminal IDRs of human Oct4, the labels indicating the assignments; the grey areas show the spectral regions where random coil amino acids resonate usually; (b,e) primary structures of Oct4; dark and light colors indicate the folded and disordered domains, respectively; blue and red sticks indicate the positions of Glycines and Prolines, respectively; (c,f) secondary structure propensities calculated from the experimental chemical shifts of the peptide backbone Cα and Cβ, using the ncSPC algorithm [104,105].

(Figure 2). The spectra of the short fragments of the C-terminal region of Sox2 were exactly overlapping with those of Sox2(aa115–317_C265A) (Supplementary Figure S1). This shows that all these fragments have very similar local conformational behaviors, a phenomenon regularly observed with IDRs. We assigned the backbone NMR signals of $^{1}H_N$, ^{15}N, $^{13}C\alpha$, $^{13}C\beta$, ^{13}CO for Sox2(aa1–42), Sox2(aa115–236), His6–AviTag–Sox2(aa234–317_C265A) and partially for Sox2(aa115–317_C265A). We aggregated the lists

of chemical shifts of the C-terminal fragments and used them to calculate the experimental secondary structure propensities (Figures 2c, f and Supplementary Figure S4). This confirmed the absence of any stable secondary structure elements in Sox2 N- and C-terminal region. The C-terminal region is poorly soluble below 0.25 M urea; this should not affect a stable fold, so we can affirm that these regions of Sox2 are experimentally proven IDRs.

3.3. *Structural characterization of the N-terminal region of Nanog*

We produced and purified the N-terminal peptide fragment of human Nanog(aa1–85). All our attempts to purify C-terminal regions of Nanog failed, even after alanine-mutation of cysteines in Nanog(aa154–305), Nanog(aa154–272), and Nanog(aa154–215). We managed to resolubilize our construct GST–His6–Tev–Nanog(aa154–305) from the insoluble fraction of the bacterial extract, to partially purify it and cleave it using the TEV protease. However, the resulting Nanog(aa154–305) peptide was barely soluble in a detergent (NP-40 at 2% v/v), and not in high-salt buffers, or not even in the presence of urea at 4 M. The 10 tryptophane residues are probably playing a role in this behavior, in the context of a primary structure containing not enough hydrophobic amino acids favoring stable folds.

The cross-peaks of Nanog(aa1–85) in the 2D ^{1}H–^{15}N HSQC NMR spectrum were all in the spectral region of random coil peptides' resonances (Figure 3). The assignment of the backbone NMR signals of $^{1}H_N$, ^{15}N, $^{13}C\alpha$, $^{13}C\beta$, ^{13}CO permitted calculating experimental secondary structure propensities, which were low through the whole peptide (Figure 3b, Supplementary Figure S4). Thus, we confirmed that this Nanog N-terminal is an IDR.

3.4. *Structural characterization of the N-terminal region of Esrrb*

We produced and purified the N-terminal fragment of human Esrrb(aa1–102), which was predicted to be disordered (Supplementary material 3.5.3), in the alanine-mutated versions Esrrb(aa1–102_C12A-C72A-C92A) and Esrrb(aa1–102_C12A-C92A). This was a strategic choice to attenuate the formation of disulfide bonds; the wild-type N-terminal fragments might however be workable too. Mutating cysteines permitted working in more comfortable conditions and maintaining our construct monomeric for longer periods of time in the next phosphorylation and biotinylation experiments. Cysteines are indeed highly solvent-accessible in IDRs and they are consequently difficult to keep in their thiol, non-disulfide forms, even in the presence of fresh DTT or TCEP at neutral pH. We also produced chimera constructs Esrrb(aa1–102_C12A-C72A-C92A)–AviTag–His6 and Esrrb(aa1–102_C12A-C72A)–AviTag–His6.

All the 2D ^{1}H–^{15}N HSQC NMR spectra revealed cross-peaks in the spectral region of random coil peptides (Figure 3). These spectra are overlapping to a large extent, confirming the weak influence of the mutations of cysteines: the mutation Cys72Ala has almost no consequences on the chemical shifts, below 0.05 ppm even for the neighboring amino acids (Supplementary Figure S2a,b); the mutation Cys91Ala has more impact, with chemical shifts perturbations of about 0.1 ppm for the next 5 amino acids (Supplementary Figure S2c,d), which is at least partially due to the fact that the Ala substitution favors an increase in local α-helicity (about 25%, see Supplementary Figure S2e,f). This N-terminal fragment of human Esrrb is thus an IDR, according to the calculated secondary structure propensities (Figure 3f, Supplementary Figures S2e,f, S4).

3.5. *Phosphorylation of Esrrb and Sox2 by p38α, Erk2, Cdk1/2 as monitored by NMR spectroscopy*

We reported recently the site-specific phosphorylation kinetics of Oct4 by p38α using ^{13}C-direct NMR detection [99]. Here, we used more standard ^{1}H-detected/^{15}N-filtered experiments to rapidly characterize the site preferences of MAPKs and CDKs on Esrrb and Sox2, which we thought to use as baits for performing phospho-dependent pull-down assays (see below).

To start with, we used commercial aliquots of MAPKs, namely p38β and Erk2, and CDKs, namely Cdk1/CyclinA2 and Cdk2/CyclinA2, on Esrrb(aa1–102). We observed the progressive phosphorylation of its three Ser–Pro motifs (at Ser22, Ser34 and Ser58)

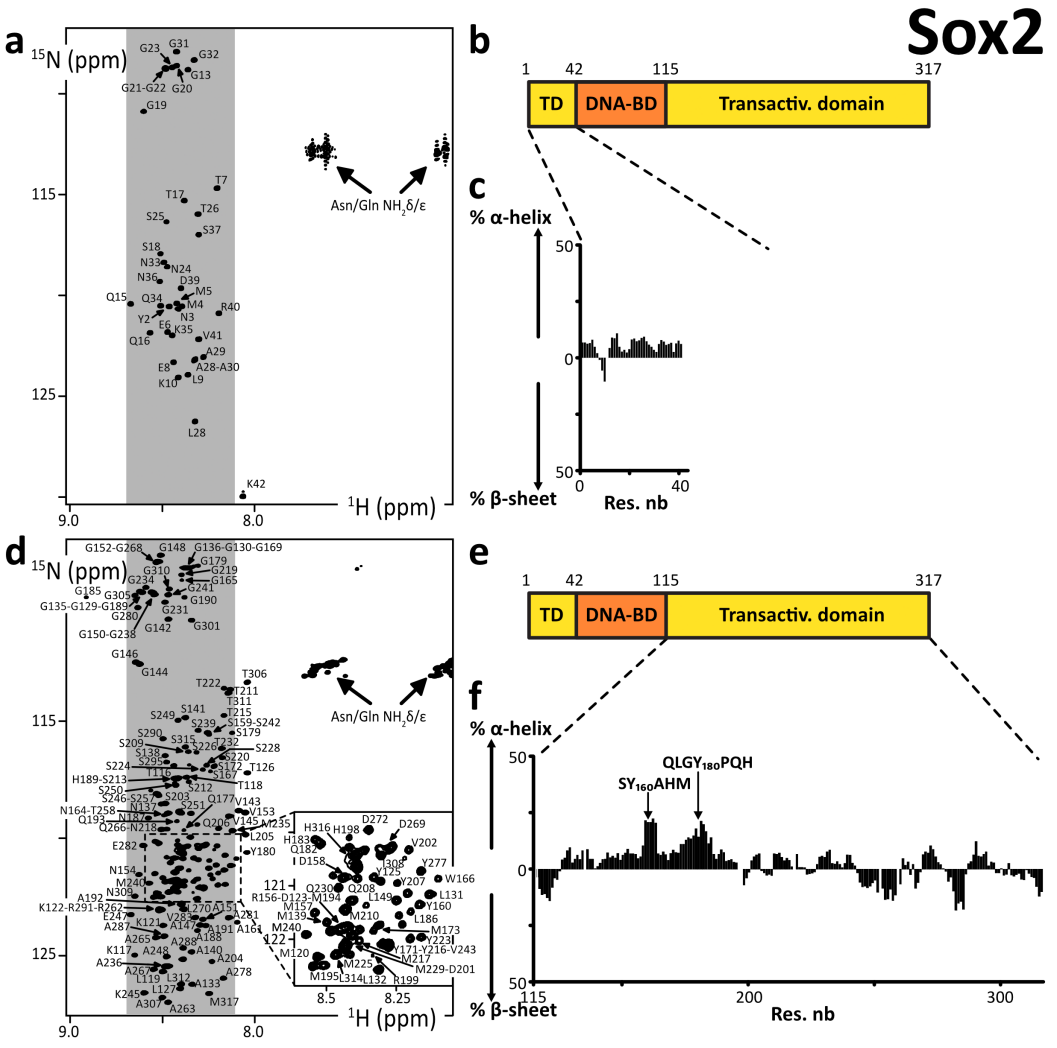


Figure 2. (a,d) 2D ^{1}H–^{15}N HSQC spectra of the N- and C-terminal IDRs of human Sox2, the labels indicating the assignments; the grey areas show the spectral regions where random coil amino acids resonate usually; (b,e) primary structures of human Sox2; dark and light colors indicate the folded and disordered domains, respectively; (c,f) secondary structure propensities calculated from the experimental chemical shifts of the peptide backbone Cα and Cβ, using the ncSPC algorithm [104,105].

in agreement with the consensus motifs of these kinases [119,120]. The kinetics reveal a classical distributive mechanism, where phosphorylation sites are processed independently according to their respective k_{cat} and K_M. Ser22 is the preferred target in all cases, while Ser34 is the least processed by CDKs, if at all: the commercial CDKs are poorly active in our hands, which we have verified with a number of other targets for years in the laboratory; this makes it difficult to distinguish between sites that are only mildly disfavored or those that are more stringently ignored by CDKs in NMR-monitored assays.

Then, we used a potent home-made activated p38α on His6–AviTag–Sox2(aa115–240) and His6–AviTag–Sox2(aa234–317). It phosphorylated all the Ser/Thr–Pro motifs of these two peptides, and also T306 in a $PGT_{306}AI$ context, which shows a favorable proline in position −2 [121], and a less common S212 in a $MTS_{212}SQ$ context.

Hence, we were able to generate AviTag–IDR

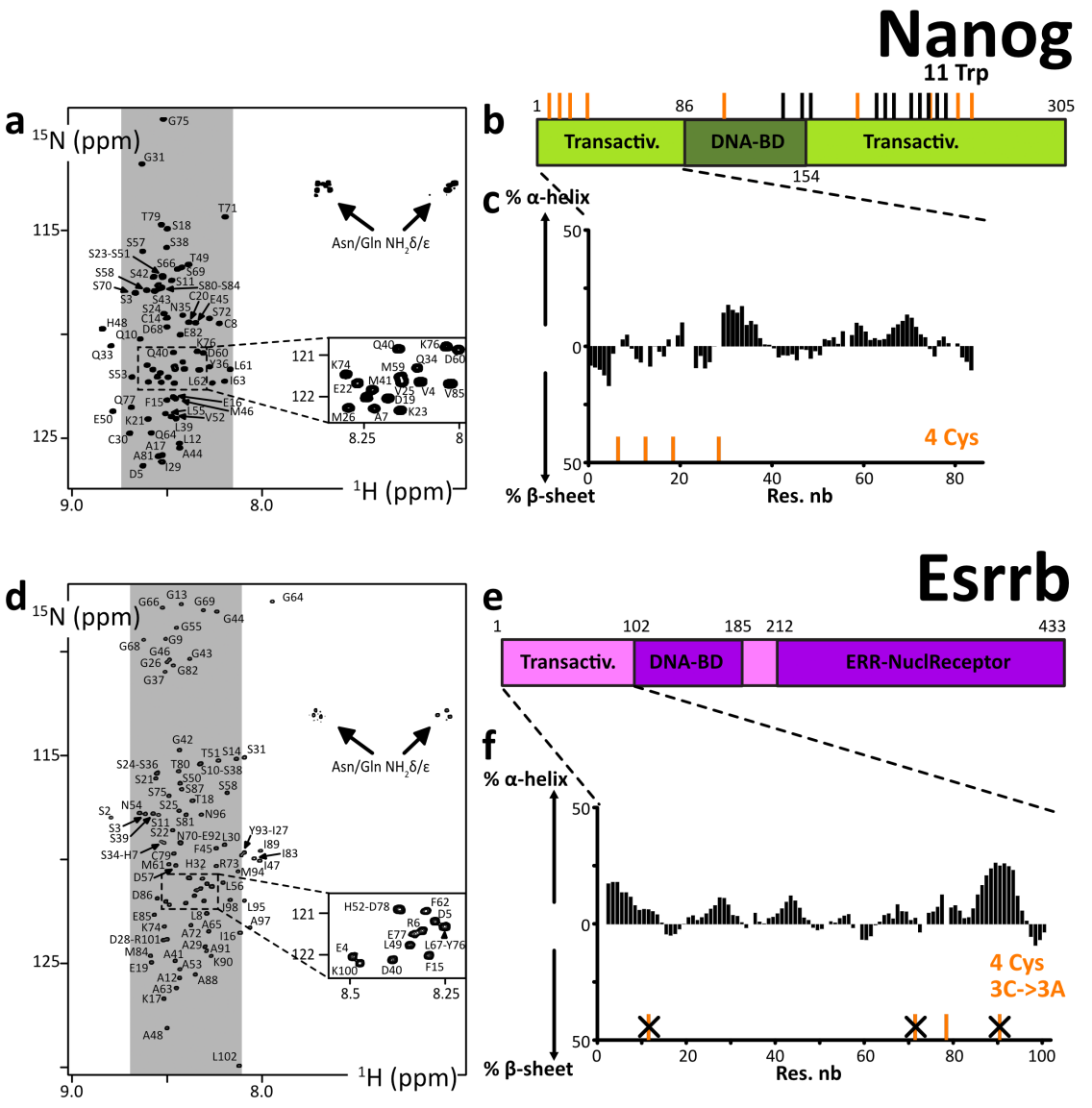


Figure 3. (a,d) 2D ^{1}H–^{15}N HSQC spectra of the N-terminal IDRs of human Nanog and Esrrb, the labels indicating the assignments; the grey areas show the spectral regions where random coil amino acids resonate usually; (b,e) primary structures of human Nanog and Esrrb; dark and light colors indicate the folded and disordered domains, respectively; orange and black sticks indicate the positions of cysteines and tryptophanes, respectively; (c,f) secondary structure propensities calculated from the experimental chemical shifts of the peptide backbone Cα and Cβ, using the ncSPC algorithm [104,105].

chimera in well-defined phosphorylated states. To produce phosphorylated ^{14}N-AviTag–IDR dedicated to pull-down assays, we executed the same protocol on ^{14}N-peptides, while NMR-monitoring in parallel "identical" but ^{15}N-labeled peptides. This allowed us to produce a well-defined phosphorylation state of the ^{14}N-AviTag–IDR for the next experiments.

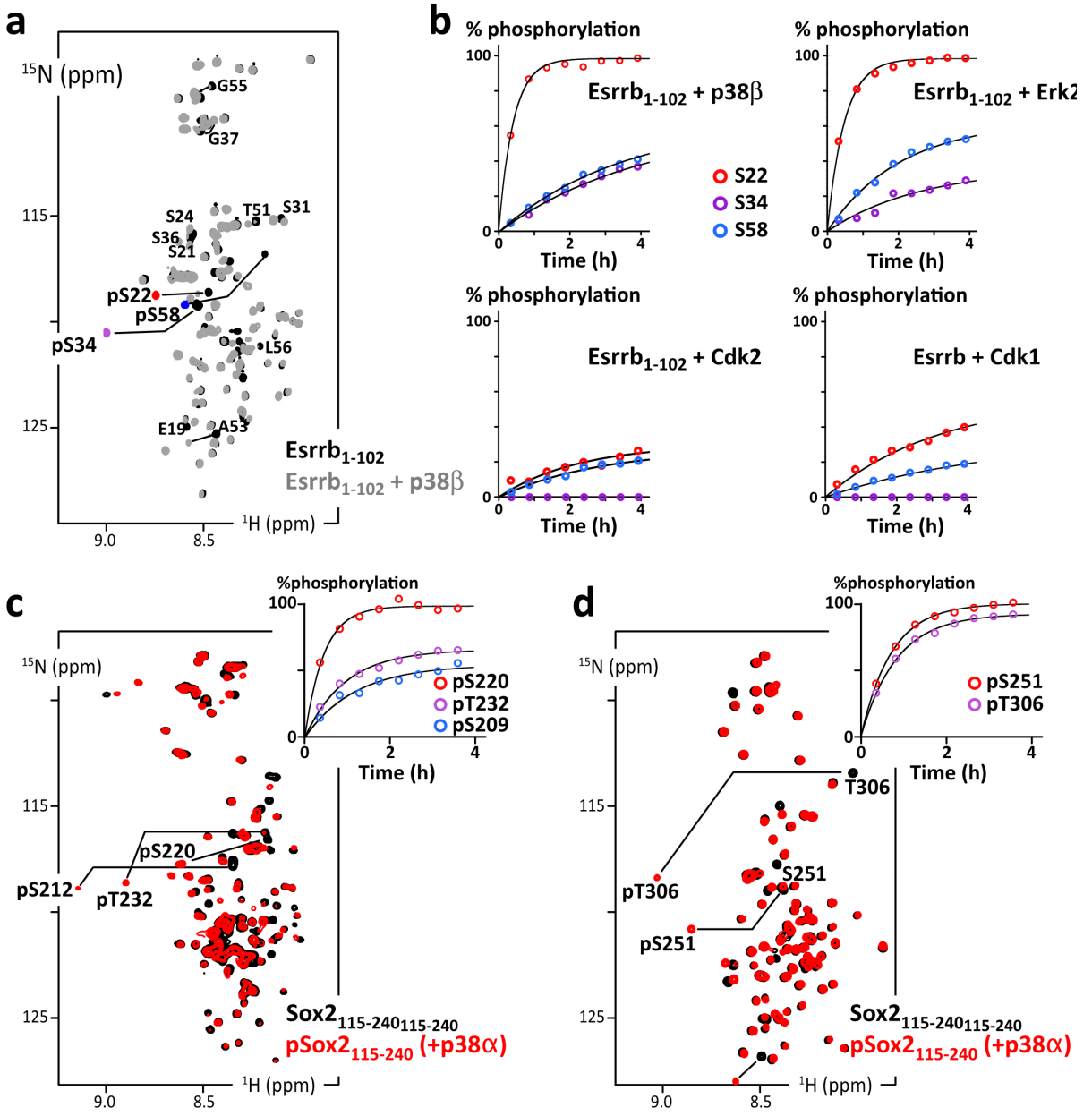


Figure 4. (a) Overlay of 2D ^{1}H–^{15}N HSQC spectra of Esrrb(aa1–102_C12A-C72A-C91A) before (black) and after (grey and phosphosites colored in red/purple/blue) phosphorylation by p38β; (b) residue specific time courses of the phosphorylation of Esrrb(aa1–102) executed by commercial kinases p38β, Erk2, Cdk2 or Cdk1, as measured in time series of 2D ^{1}H–^{15}N SOFAST-HMQC spectra recorded during the reaction; (c) overlay of 2D ^{1}H–^{15}N HSQC spectra of AviTag-Sox2(aa115–240) before (black) and after (red) phosphorylation by p38α; the inset at the top-right shows the residue specific phosphorylation build-up curves, as measured in time series of 2D ^{1}H–^{15}N SOFAST-HMQC spectra; (d) same as (c) with AviTag–Sox2(aa234–317_C265A).

3.6. *Structural characterization of the AviTag–peptide chimera and its biotinylated versions*

We aimed at detecting new partners of OSNE using pull-down assays. We thought to use chimera containing GST at the N-terminus, which appeared as a convenient approach: vectors integrating a GST-coding DNA sequence for overexpression in *E. coli* are available and of common use; glutathione-coated beads are also accessible and permit efficient and specific binding of GST-containing chimera peptides. However, we were unsatisfied by the performances of the method: GST binding to glutathione-coated beads is slow at low temperature (necessary to avoid IDR proteolysis); moreover, it appeared that GST–IDRs chimera are hampered by the IDRs' "molecular cloud" and are even weaker and slower to bind to the beads. Our attempts to bind GST–IDRs to the beads were thus resulting in poor yields, which were not very reproducible. In the context of our aims, i.e., to establish a method allowing quantitative detection of IDRs' binding partners, this unsatisfying lack of reproducibility was only foreboding supplementary variable parameters.

Thus, we decided to switch to another strategy: the use of the specifically biotinylated 15-mer peptide tag called AviTag [101]. This is efficiently and specifically biotinylated by biotin ligase BirA (Figure 5d), which permits high-affinity binding to streptavidin-coated beads. We designed AviTag–IDR chimera, with the AviTag in N-terminal position for Sox2's IDR constructs, and in C-terminal for Esrrb's IDR constructs. We characterized the AviTag and its impact on the IDRs of interest using NMR: the AviTag is unfolded and it does not affect the Sox2 and Esrrb fragments, according to the observed negligible chemical shift perturbations (Figure 5)—the GST Tag provoked also only very weak chemical shift changes on Esrrb(aa1–102) (Supplementary Figure S3). The biotinylated AviTag peptide appears to be slightly less mobile than the common IDPs on the ps–ns timescale, according to the heteronuclear ^{15}N–{^{1}H} NOEs (Figure 5b).

We observed that the biotinylation of the AviTag provokes weak but distinguishable chemical shift perturbations for the close neighbors of the biotinylated lysine, but had no effect on the peptides of interest (Figure 5c). It generated also the appearance of a HN-ester NMR signal, similar to that of acetylated lysine [108,122]. Hence, we could quantify and monitor the reaction advancement using NMR, and determine the incubation time that was necessary and sufficient to obtain a complete biotinylation of our chimera AviTag–IDRs (see Section 2.2). This was one among many optimization steps permitting the production of sufficient quantities of intact IDRs for the pull-down assays.

Next, we tested the binding of the biotinylated AviTag–IDRs on streptavidin-coated beads. This produced very satisfying results, i.e., stoichiometric, specific binding in one hour with no leakage (Supplementary Figure S6). This approach was thus selected for the pull-down assays.

3.7. *Use of AviTag–Sox2 for pull-down assays from mESCs extracts*

We prepared the four peptides AviTag–Sox2(aa115–240) and AviTag–Sox2(aa234–317) in their non-phosphorylated and phosphorylated versions, using p38α to execute the phosphorylation reactions (Figure 6a). These peptides were also biotinylated, and later bound to streptavidin-coated beads, which we used as baits for pull-down assays in extracts from mESCs (Figure 6b). Importantly, an additional SEC was carried out between every step to discard proteolyzed peptides, the enzymes (kinases of BirA) and their contaminants. We performed the pull-down assays with the four samples in parallel with the same cell extract, in duplicate, and then analyzed the bound fractions using quantitative LC-MS/MS analysis (see Supplementary material 1.6 for full description). Hence, we could detect and evaluate the relative quantities of proteins retained by the four AviTag–peptides (Figure 6). On paper, this presents the important advantage of removing false-positive binders, which can interact unspecifically with the streptavidin-coated beads.

We present here results that should be interpreted carefully: we produced only duplicates for every condition, using one single cell extract. To deliver trustful information, the common standards in the field recommend 3 to 5 replicates. Here, we considered only proteins with at least 2 distinct peptides across the 2 replicates (Figure 6c). We observed a two-fold change or more (FC > 2) of some TFs pulled out by AviTag–Sox2(aa115–240), among which the PluriTFs Oct4 and Klf5 are significantly enriched

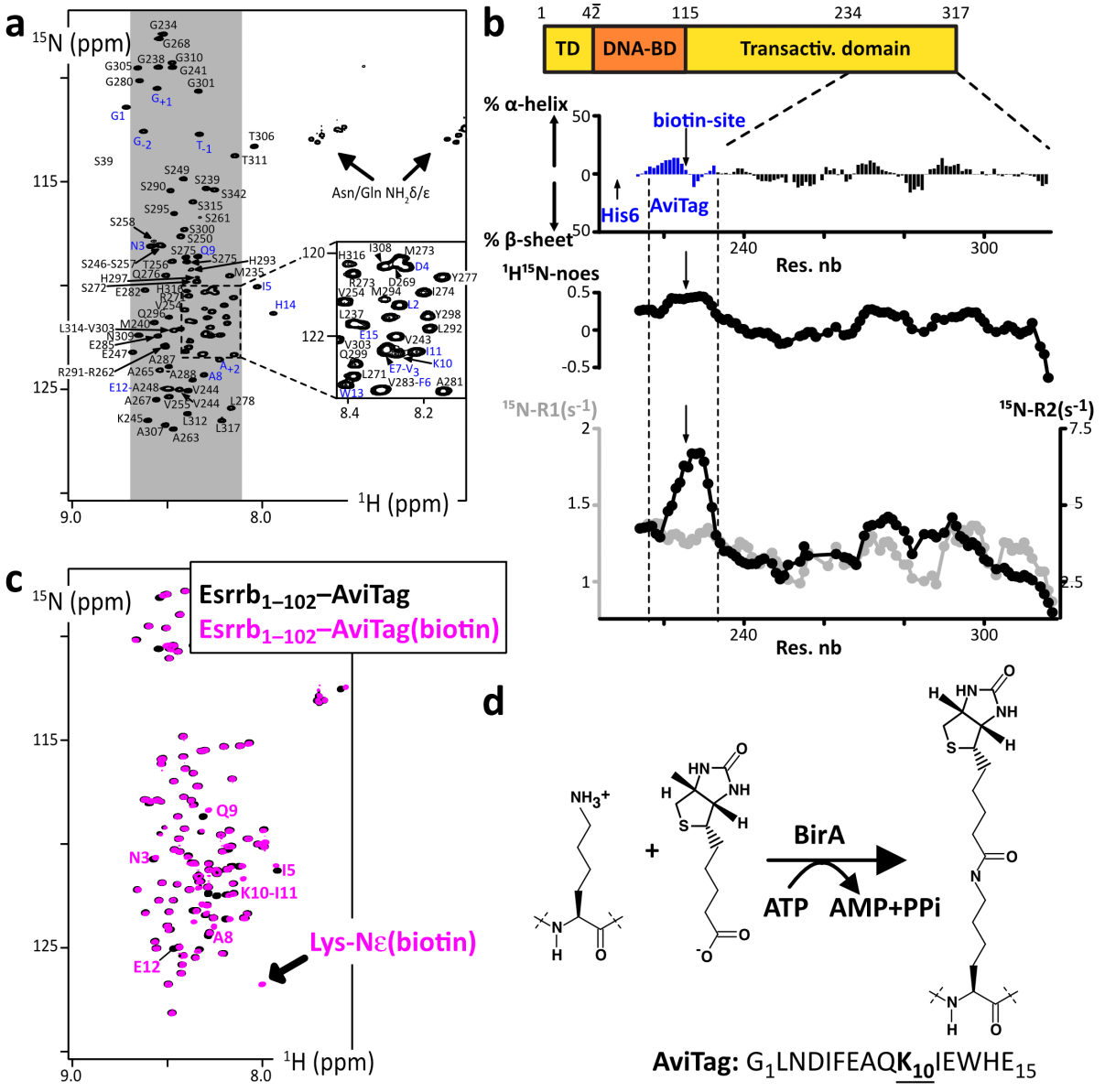


Figure 5. (a) 2D ^{1}H–^{15}N HSQC spectrum of ^{15}N-His6-AviTag-Sox2(aa234–317_C265A), the blue labels indicated the assigned signals from the AviTag residues; (b) secondary structure propensities calculated from the experimental chemical shifts of the peptide backbone Cα and Cβ, using the ncSPC algorithm [104,105]; the residue specific ^{15}N–{^{1}H} NOEs, ^{15}N-R1 (grey) and ^{15}N-R2 (black) measured at 600 MHz are shown below (the profiles show values averaged over three consecutive residues); (c) overlay of 2D ^{1}H–^{15}N HSQC spectra of the Esrrb(aa1–102_C12A-C72A-C91A)–AviTag–His6 before (black) et after (magenta) biotinylation by BirA; the NMR signals from the residues neighboring the biotinylation site are indicated, which permit the quantification of the biotinylated population; (d) scheme of the reaction of AviTag biotinylation executed by the ATP-dependent BirA.

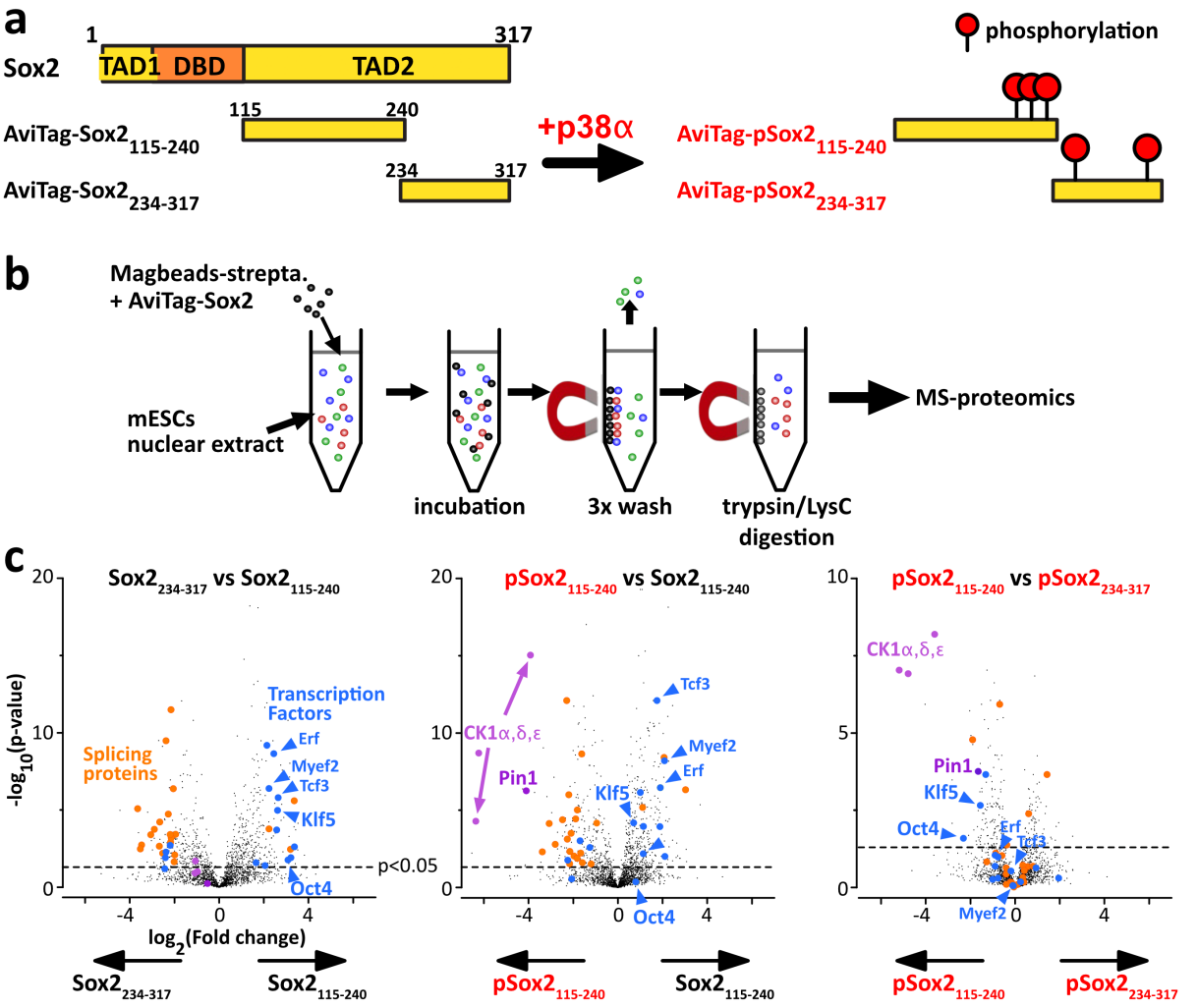


Figure 6. Differential interactomics of Sox2 constructs upon p38α phosphorylation. (a) We have produced unmodified/phospho-Sox2 truncations carrying N-ter biotinylation on the AviTag, and later attached on streptavidin-magnetic beads. (b) We have generated mESCs nuclear extracts for pull-down assays using our Sox2 constructs as baits. (c) The volcano plots of the log2 ratios, showing a quantitative analysis of the proteins present at the end of the pull-down assays; the dashed line indicates the threshold of p-value < 0.05; we highlighted interesting partners in: blue: transcription factors associating with Sox2(aa115–240); magenta: phospho-dependent partners of Sox2(aa115–240) and or Sox2(aa234–317). Pull-downs have been performed in duplicates, using 15 million cells per sample (extract protein conc.: 5 mg/mL), and 1 nmol of bait protein. Experimental conditions may be improved (higher number of replicates, cells, washing conditions, ...).

(>4 peptides, p-value < 0.02). Also, we found out that pSox2(aa115–240) was pulling out the three isoforms of CK1 (>3 peptides, p-values < 10^{-5}) and the proline isomerase Pin1 (>2 peptides, p-values < 2×10^{-4}). To confirm the value of the method and of the detected interactions, one should at least test them using an orthogonal approach, e.g., NMR spectroscopy of purified proteins.

3.8. *NMR characterization of the interaction between Pin1 and phospho-Sox2*

We decided to test the interaction between pSox2(aa115–240) and Pin1. We recorded ^{1}H–^{15}N NMR spectra of ^{15}N-labeled Sox2(aa115–240) or pSox2(aa115–240) alone or in the presence of the Pin1–WW domain (natural abundance peptide, i.e., 0.6% ^{15}N, 99.4% ^{14}N, hence "NMR-invisible" in ^{15}N-filtered experiments). We observed localized losses in signal intensities for the residues neighboring the three pSox2(aa115–240) phosphosites when mixed with Pin1 (Figure 7b); in contrast, no significant differences showed up in the spectra obtained with non-phosphorylated Sox2(aa115–240) in the absence or presence of Pin1–WW (Figure 7a). Hence, these signal losses are the typical signs of a position-specific interaction between an IDR and a folded protein in the intermediate or slow NMR timescale (μs–s), which corresponds to submicromolar affinities for this type of molecules. Interactions with this range of affinity can therefore be detected using the presented pull-down assay approach.

4. Discussion

The structural biochemistry analysis reported here can be applied to a large list of transcription factors (TFs). These are essential actors of cell signaling: they are key elements for inducing or maintaining pluripotency or differentiation, for cell proliferation or cell cycle arrest, by activating or repressing gene transcription [61,123]. About 90% of the ~1600 TFs are predicted to contain large disordered segments (>30 consecutive amino acids), which is particularly true for PluriTFs [31,124]. A correlation exists actually between TFs and predicted IDRs in all kingdoms of life [45,125,126]. The IDRs of these TFs recruit transcription co-factors or the transcription machinery, which is still not very well characterized in detail [44–55].

Indeed, the fine understanding of TFs' interactions via their IDRs appears to be hampered by the nature of these interactions, which are characterized by weak affinities and multivalency. Moreover, possible redundancy between TFs can emerge from coacervation. Post-translational modifications (PTMs), which can switch on or off IDRs' interactions, are a supplementary source of confusion when searching for binding partners. Among the difficulties, we should also mention the basic biochemistry issues: IDRs are difficult to produce and manipulate in vitro, because they are prone to degradation or aggregation. Here, we have tried to demonstrate the feasibility and interest of some biochemical and spectroscopic approaches to better characterize IDRs of TFs, their phosphorylation and the associated binding partners.

Using a residue specific NMR analysis, we have shown experimentally that the pluripotency TFs OSNE contained IDRs. None of these regions do show any strong secondary structure propensity, which often reveals functional binding sites. Faint ~20% helicities were detected by two independent algorithms (ncIDP and δ2D [104,105,107]) on 5–6 amino acid stretches, which are indicated on Figures 1 and 2 for Oct4 and Sox2. Altogether, the analyzed peptides show local conformational behaviors close to that of a random coil, without any obvious structural elements except the Oct4 stretch between amino acids 33 and 38. Finally, we shall highlight the high propensity for liquid–liquid phase separation of the fragments of Sox2 containing the region aa115–236.

Like other TFs, pluripotency TFs OSNE are post-translationally modified (see Section 1), notably by CDKs and MAPKs [66,78–88]. These two classes of kinases are fundamental actors in all aspects of eukaryotic cellular life, and understanding their activity and regulation in pluripotency or differentiation is of high significance. Interesting questions are still pending: what is the phosphorylation status of OSNE's IDRs in pluripotent cells, what is the impact on their interaction networks, and how does it affect pluripotency or differentiation? The inhibition of MAPK Erk signaling is necessary to maintain pluripotency in the standard culture conditions of ESCs and iPSCs [66,82–84], while these cells show a high CDK activity [8,86,87]; these two kinases family have the same core consensus sites (Ser/Thr–Pro) motifs, which are abundant in OSNE's IDRs and whose phosphorylation has apparently consequences for initiating differentiation [73, 78,80,83,85]. We have shown that we could produce well-defined phosphorylation status of these peptides, using recombinant kinases and NMR analysis, which makes it possible to study their interactions in vitro. We have also demonstrated our capacity to use these peptides as baits in pull-down assays for detecting potential new binding partners.

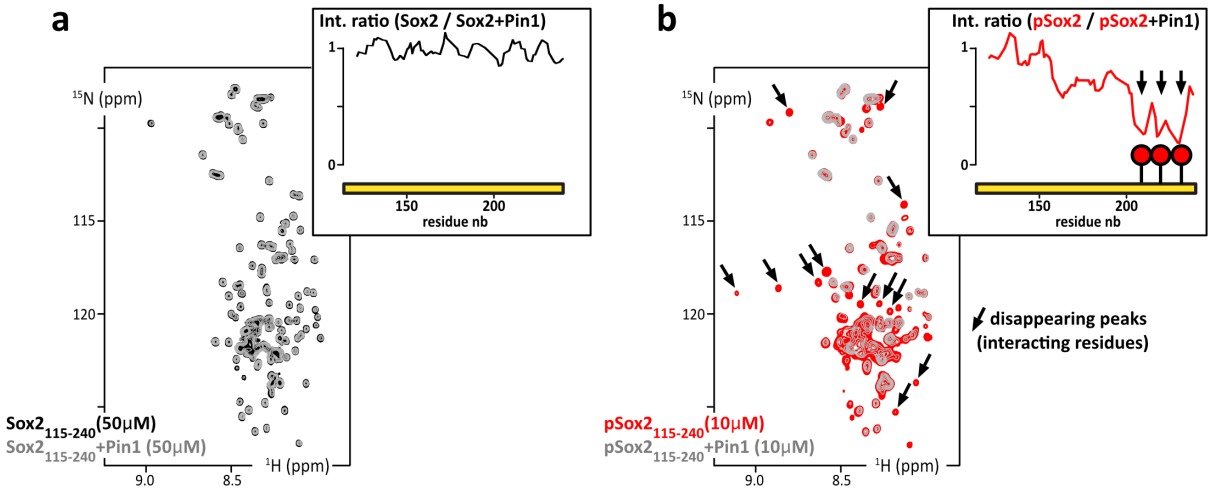


Figure 7. (a) Overlay of 2D ^{1}H–^{15}N HSQC spectra of ^{15}N-Sox2(aa115–240) alone at 50 µM (black) or mixed with Pin1 in isotopic natural abundance and in stoichiometric amounts (grey); inset up-right: residue specific NMR signal intensity ratios as measured in the two HSQC spectra. (b) Overlay of 2D ^{1}H–^{15}N HSQC spectra of ^{15}N-phosphoSox2(aa115–240) alone at 10 µM (red) or mixed with the Pin1–WW domain in stoichiometric amounts (grey); insets, up-right: residue specific NMR signal intensity ratios as measured in the two HSQC spectra (in the absence/presence of Pin1–WW domain).

However, the detected interactions between phospho-Sox2(aa115–240)–pS212–pS220–pT232 and Pin1 or CK1 correspond to widespread, degenerate interactions, whose biological significance may be questionable [127–129]. This is one of the major drawbacks in the field of IDRs' studies: they participate in multiple, degenerate and transient interactions of weak affinities, which can be easily released during the washes of our pull-down assays. In this regard, the "proximity labeling" approaches (BioID, APEX and their derivates) appeared recently to be quite adapted to transient interactions: these methods, developed in the last ten years, use chimera constructs containing enzymes that transfer chemical groups to their intracellular neighbors, which can be later identified by mass spectrometry [130–133]. IDRs are very flexible, solvent-exposed and establish a lot of poorly specific transient interactions. This was raising concerns about the possible production of many false positives if one used proximity labeling methods to detect IDRs' binding partners. This has been partially confirmed by a recent study, but this bias appears to be limited [134]. Interactomes of 109 TFs have actually been described using Bio-ID and affinity purification MS, showing the complementarity between proximity labeling and the pull-down approach proposed here [135]. Yeast two-hybrid, which can detect ~20 µM affinity interactions, and novel phage display approaches will also help in this task [136–139].

Another difficulty in studying IDRs of TFs is their propensity to coacervate [44–55]. Here, we have tried to use Sox2 as a bait in pull-down assays, a protein that has been later recognized to favor liquid–liquid phase separation [44,140]. We met this difficulty during the production steps, which forced us to purify most of the Sox2 constructs in urea at 2 M. We could straightforwardly observe liquid–liquid phase separation of Sox2(aa115–317) at 4 µM using differential interference contrast (DIC) microscopy in the presence of Ficoll-70 (Supplementary Figure S5), but also progressive aggregation and low solubility thresholds while working with our purified samples. These are clear limiting factors for sample production and NMR characterization, which will hamper a number of other studies on IDRs of TFs. This might also affect the results of pull-down assays: we noticed an enrichment in TFs in the samples obtained from pull-downs using Sox2(aa115–240) as a bait, which has a much higher coacervation propensity than Sox2(aa234–317). Is it possible to generate local surface liquid–liquid

phase separation on the surface of streptavidin-coated beads? This might be at the same time a blessing and a curse for future studies, by helping the formation of biologically significant assemblies, or by favoring unspecific, non-native macromolecules interactions.

A final bottleneck in the studies of these IDRs is the capacity to produce post-translationally modified samples. The commercial enzymes are not very well adapted to our NMR studies, because of the required quantities. Here, we have used in-house production of kinase p38α. Since we carried out the present work, we have developed our capacities in producing activated Erk2, Cdk2/CyclinA1 and Cdk1/CyclinB1. These will be part of our future studies. TFs are indeed quite adapted to NMR investigations: they are 300 to 500 residues long and contain large IDRs (~100 amino acids) separating small folded domains (also ~100 amino acids) [31,45]. Their structural characterization would permit understanding a number of cell-signaling mechanisms at the atomic scale, and possibly identifying new therapeutic targets, even though this class of proteins is notoriously difficult to inhibit [60,141,142].

5. Conclusion

We have applied NMR techniques to carry out a primary analysis of the pluripotency transcription factors Oct4, Sox2, Nanog and Esrrb, in particular of their intrinsically disordered regions. We have shown experimentally that they did not adopt a stable fold when isolated, and that we were able to conduct a residue-specific analysis. This relies on the delicate production and purification of these peptides, which are prone to proteolysis and aggregation; producing them in a well-defined PTM status was an even more arduous challenge. We have demonstrated the feasibility of these tasks using recombinant kinases and NMR analysis. We have also evaluated the usefulness of such protein constructs as baits in pull-down assays to detect new binding partners of IDRs. These characterizations and the associated methods provide firm basis for future investigations on transcription factors. The proposed experimental scheme is thus a promising methodology that still needs to be developed and to prove its merits in revealing novel and significant interactions.

Declaration of interests

The authors do not work for, advise, own shares in, or receive funds from any organization that could benefit from this article, and have declared no affiliations other than their research organizations.

Funding

This work was supported by the CNRS and the CEA-Saclay, by the French Infrastructure for Integrated Structural Biology (https://frisbi.eu/, grant number ANR-10-INSB-05-01, Acronym FRISBI) and by the French National Research Agency (ANR; research grants ANR-14-ACHN-0015 and ANR-20-CE92-0013). Financial support from the IR INFRANALYTICS FR2054 for conducting the research is also gratefully acknowledged. This work was also supported by grants from the “Région Ile-de-France” and “Fondation pour la Recherche Médicale” (DL and LSMP).

Acknowledgements

We thank Thaleia Papadopoulou, Amandine Molliex and Navarro Pablo for providing mESCs extracts and for fruitful discussions. We thank Nadia Izadi-Pruneyre for providing the bench space necessary to carry out the pull-down experiments with fresh mESCs extracts. We thank Romain LeBars and the microscopy facility of I2BC Imagerie-Gif for the preliminary tests on liquid–liquid phase separation. We thank the IR-RMN staff (now part of Infranalytics), notably Nelly Morellet, François Giraud and Ewen Lescop, for their reactivity, their support, and their long-standing, efficient care of the 950 MHz spectrometer. We thank Marie Sorin, Baptiste Nguyen and Benjamin Bacri, who contributed to this study during their Master1-Master2 internships.

Supplementary data

Supporting information for this article is available on the journal’s website under https://doi.org/10.5802/crchim.272 or from the author.

References

[1] H. Inoue, N. Nagata, H. Kurokawa, S. Yamanaka, *EMBO J.*, 2014, **33**, 409-417.

[2] Y. Shi, H. Inoue, J. C. Wu, S. Yamanaka, *Nat. Rev. Drug Discov.*, 2017, **16**, 115-130.
[3] R. G. Rowe, G. Q. Daley, *Nat. Rev. Genet.*, 2019, **20**, 377-388.
[4] W. Deng, E. C. Jacobson, A. J. Collier, K. Plath, *Curr. Opin. Genet. Dev.*, 2021, **70**, 89-96.
[5] T. W. Theunissen, R. Jaenisch, *Stem Cell*, 2014, **14**, 720-734.
[6] M. Li, J. C. I. Belmonte, *Nat. Rev. Genet.*, 2017, **18**, 180-191.
[7] C. Chronis, P. Fiziev, B. Papp, S. Butz, G. Bonora, S. Sabri, J. Ernst, K. Plath, *Cell*, 2017, **168**, 442-459, e20.
[8] N. Festuccia, I. Gonzalez, N. Owens, P. Navarro, *Development*, 2017, **144**, 3633-3645.
[9] I. Gonzalez, A. Molliex, P. Navarro, *Curr. Opin. Cell Biol.*, 2021, **69**, 41-47.
[10] P. Mu, Z. Zhang, M. Benelli, W. R. Karthaus, E. Hoover, C.-C. Chen, J. Wongvipat, S.-Y. Ku, D. Gao, Z. Cao, N. Shah, E. J. Adams, W. Abida, P. A. Watson, D. Prandi, C.-H. Huang, E. de Stanchina, S. W. Lowe, L. Ellis, H. Beltran, M. A. Rubin, D. W. Goodrich, F. Demichelis, C. L. Sawyers, *Science*, 2017, **355**, 84-88.
[11] A. C. Hepburn, R. E. Steele, R. Veeratterapillay, L. Wilson, E. E. Kounatidou, A. Barnard, P. Berry, J. R. Cassidy, M. Moad, A. El-Sherif, L. Gaughan, I. G. Mills, C. N. Robson, R. Heer, *Oncogene*, 2019, **38**, 4412-4424.
[12] S. Mirzaei, M. D. A. Paskeh, M. Entezari, S. reza Mirmazloomi, A. Hassanpoor, M. Aboutalebi, S. Rezaei, E. S. Hejazi, A. Kakavand, H. Heidari, S. Salimimoghadam, A. Taheriazam, M. Hashemi, S. Samarghandian, *Biomed. Pharmacother.*, 2022, **156**, article no. 113860.
[13] E.-H. Ervin, R. French, C.-H. Chang, S. Pauklin, *Semin. Cancer Biol.*, 2022, **87**, 48-83.
[14] A. Chaudhary, S. S. Raza, R. Haque, *Semin. Cancer Biol.*, 2023, **88**, 123-137.
[15] D. Esch, J. Vahokoski, M. R. Groves, V. Pogenberg, V. Cojocaru, H. vom Bruch, D. Han, H. C. A. Drexler, M. J. Araúzo-Bravo, C. K. L. Ng, R. Jauch, M. Wilmanns, H. R. Schöler, *Nat. Cell Biol.*, 2013, **15**, 295-301.
[16] A. Reményi, K. Lins, L. J. Nissen, R. Reinbold, H. R. Schöler, M. Wilmanns, *Genes Dev.*, 2003, **17**, 2048-2059.
[17] M. D. Gearhart, S. M. A. Holmbeck, R. M. Evans, H. J. Dyson, P. E. Wright, *J. Mol. Biol.*, 2003, **327**, 819-832.
[18] R. Jauch, C. K. L. Ng, K. S. Saikatendu, R. C. Stevens, P. R. Kolatkar, *J. Mol. Biol.*, 2008, **376**, 758-770.
[19] Y. Hayashi, L. Caboni, D. Das, F. Yumoto, T. Clayton, M. C. Deller, P. Nguyen, C. L. Farr, H.-J. Chiu, M. D. Miller, M.-A. Elsliger, A. M. Deacon, A. Godzik, S. A. Lesley, K. Tomoda, B. R. Conklin, I. A. Wilson, S. Yamanaka, R. J. Fletterick, *Proc. Natl. Acad. Sci. USA*, 2015, **112**, 4666-4671.
[20] B. Yao, S. Zhang, Y. Wei, S. Tian, Z. Lu, L. Jin, Y. He, W. Xie, Y. Li, *J. Mol. Biol.*, 2020, **432**, 5460-5472.
[21] S. O. Dodonova, F. Zhu, C. Dienemann, J. Taipale, P. Cramer, *Nature*, 2020, **580**, 669-672.
[22] A. K. Michael, R. S. Grand, L. Isbel, S. Cavadini, Z. Kozicka, G. Kempf, R. D. Bunker, A. D. Schenk, A. Graff-Meyer, G. R. Pathare, J. Weiss, S. Matsumoto, L. Burger, D. Schübeler, N. H. Thomä, *Science*, 2020, **368**, 1460-1465.
[23] K. Echigoya, M. Koyama, L. Negishi, Y. Takizawa, Y. Mizukami, H. Shimabayashi, A. Kuroda, H. Kurumizaka, *Sci. Rep.*, 2020, **10**, article no. 11832.
[24] G. A. Roberts, B. Ozkan, I. Gachulincová, M. R. O'Dwyer, E. Hall-Ponsele, M. Saxena, P. J. Robinson, A. Soufi, *Nat. Cell Biol.*, 2021, **23**, 834-845.
[25] E. Morgunova, J. Taipale, *Curr. Opin. Struct. Biol.*, 2021, **71**, 171-179.
[26] W. Kagawa, H. Kurumizaka, *Curr. Opin. Struct. Biol.*, 2021, **71**, 59-64.
[27] E. Luzete-Monteiro, K. S. Zaret, *Curr. Opin. Struct. Biol.*, 2022, **75**, article no. 102425.
[28] B. D. Sunkel, B. Z. Stanton, *IScience*, 2021, **24**, article no. 103132.
[29] F. C. M. Gadea, E. N. Nikolova, *J. Mol. Biol.*, 2023, **435**, article no. 167916.
[30] B. Jagga, M. Edwards, M. Pagin, K. M. Wagstaff, D. Aragão, N. Roman, J. D. Nanson, S. R. Raidal, N. Dominado, M. Stewart, D. A. Jans, G. R. Hime, S. K. Nicolis, C. F. Basler, J. K. Forwood, *Nat. Commun.*, 2021, **12**, article no. 28.
[31] B. Xue, C. J. Oldfield, Y.-Y. Van, A. K. Dunker, V. N. Uversky, *Mol. BioSyst.*, 2012, **8**, 134-150.
[32] M. M. Babu, *Biochem. Soc. Trans.*, 2016, **44**, 1185-1200.
[33] P. E. Wright, H. J. Dyson, *Nat. Struct. Mol. Biol.*, 2015, **16**, 18-29.
[34] D. Piovesan, M. Necci, N. Escobedo, A. M. Monzon, A. Hatos, I. Mičetić, F. Quaglia, L. Paladin, P. Ramasamy, Z. Dosztanyi, W. F. Vranken, N. E. Davey, G. Parisi, M. Fuxreiter, S. C. E. Tosatto, *Nucleic Acids Res.*, 2021, **49**, D361-D367.
[35] P. Kulkarni, S. Bhattacharya, S. Achuthan, A. Behal, M. K. Jolly, S. Kotnala, A. Mohanty, G. Rangarajan, R. Salgia, V. Uversky, *Chem. Rev.*, 2022, **122**, 6614-6633.
[36] D. Piovesan, A. Del Conte, D. Clementel, A. M. Monzon, M. Bevilacqua, M. C. Aspromonte, J. A. Iserte, F. E. Orti, C. Marino-Buslje, S. C. E. Tosatto, *Nucleic Acids Res.*, 2023, **51**, D438-D444.
[37] Y. Buganim, D. A. Faddah, R. Jaenisch, *Nat. Struct. Mol. Biol.*, 2013, **14**, 427-439.
[38] A. Rizzino, *Stem Cells*, 2013, **31**, 1033-1039.
[39] S. Jerabek, F. Merino, H. R. Schöler, V. Cojocaru, *Biochim. Biophys. Acta Gene Regul. Mech.*, 2014, **1839**, 138-154.
[40] A. Saunders, F. Faiola, J. Wang, *Stem Cells*, 2013, **31**, 1227-1236.
[41] S. E. Bondos, A. K. Dunker, V. N. Uversky, *Cell Commun. Signal*, 2022, **20**, article no. 20.
[42] K.-P. Kim, Y. Wu, J. Yoon, K. Adachi, G. Wu, S. Velychko, C. M. MacCarthy, B. Shin, A. Röpke, M. J. Arauzo-Bravo, M. Stehling, D. W. Han, Y. Gao, J. Kim, S. Gao, H. R. Schöler, *Sci. Adv.*, 2020, **6**, article no. eaaz7364.
[43] I. Aksoy, R. Jauch, V. Eras, W. A. Chng, J. Chen, U. Divakar, C. K. L. Ng, P. R. Kolatkar, L. W. Stanton, *Stem Cells*, 2013, **31**, 2632-2646.
[44] A. Boija, I. A. Klein, B. R. Sabari, A. Dall'Agnese, E. L. Coffey, A. V. Zamudio, C. H. Li, K. Shrinivas, J. C. Manteiga, N. M. Hannett, B. J. Abraham, L. K. Afeyan, Y. E. Guo, J. K. Rimel, C. B. Fant, J. Schuijers, T. I. Lee, D. J. Taatjes, R. A. Young, *Cell*, 2018, **175**, 1842-1855.
[45] L. Staby, C. O'Shea, M. Willemoës, F. Theisen, B. B. Kragelund, K. Skriver, *Biochem. J.*, 2017, **474**, 2509-2532.
[46] C. N. Ravarani, T. Y. Erkina, G. De Baets, D. C. Dudman, A. M. Erkine, M. M. Babu, *Mol. Syst. Biol.*, 2018, **14**, article no. e8190-14.

[47] S. Brodsky, T. Jana, K. Mittelman, M. Chapal, D. K. Kumar, M. Carmi, N. Barkai, *Mol. Cell*, 2020, **79**, 459-471, e4.
[48] A. Erijman, L. Kozlowski, S. Sohrabi-Jahromi, J. Fishburn, L. Warfield, J. Schreiber, W. S. Noble, J. Söding, S. Hahn, *Mol. Cell*, 2020, **78**, 890-902, e6.
[49] G. Næs, J. O. Storesund, P. Udayakumar, M. Ledsaak, O. S. Gabrielsen, *FEBS Open Bio*, 2020, **10**, 2329-2342.
[50] A. L. Sanborn, B. T. Yeh, J. T. Feigerle, C. V. Hao, R. J. Townshend, E. Lieberman Aiden, R. O. Dror, R. D. Kornberg, *ELife*, 2021, **10**, article no. e68068.
[51] L. M. Tuttle, D. Pacheco, L. Warfield, D. B. Wilburn, S. Hahn, R. E. Klevit, *Nat. Commun.*, 2021, **12**, article no. 2220.
[52] L. F. Soto, Z. Li, C. S. Santoso, A. Berenson, I. Ho, V. X. Shen, S. Yuan, J. I. Fuxman Bass, *Mol. Cell*, 2022, **82**, 514-526.
[53] M. V. Staller, E. Ramirez, S. R. Kotha, A. S. Holehouse, R. V. Pappu, B. A. Cohen, *Cell Syst.*, 2022, **13**, 334-345, e5.
[54] B. Bourgeois, T. Gui, D. Hoogeboom, H. G. Hocking, G. Richter, E. Spreitzer, M. Viertler, K. Richter, T. Madl, B. M. T. Burgering, *Cell Rep.*, 2021, **36**, article no. 109446.
[55] K. Teilum, J. G. Olsen, B. B. Kragelund, *Biochem. J.*, 2021, **478**, 2035-2050.
[56] F.-X. Theillet, A. Binolfi, T. Frembgen-Kesner, K. Hingorani, M. Sarkar, C. Kyne, C. Li, P. B. Crowley, L. Gierasch, G. J. Pielak, A. H. Elcock, A. Gershenson, P. Selenko, *Chem. Rev.*, 2014, **114**, 6661-6714.
[57] N. E. Davey, *Curr. Opin. Struct. Biol.*, 2019, **56**, 155-163.
[58] T. M. Filtz, W. K. Vogel, M. Leid, *Trends Pharmacol. Sci.*, 2014, **35**, 76-85.
[59] D. Han, M. Huang, T. Wang, Z. Li, Y. Chen, C. Liu, Z. Lei, X. Chu, *Cell Death Dis.*, 2019, **10**, article no. 290.
[60] M. Qian, F. Yan, T. Yuan, B. Yang, Q. He, H. Zhu, *Drug Discov. Today*, 2020, **25**, 1502-1512.
[61] P. Weidemüller, M. Kholmatov, E. Petsalaki, J. B. Zaugg, *Proteomics*, 2021, **21**, article no. 2000034.
[62] N. Cai, M. Li, J. Qu, G.-H. Liu, J. C. Izpisua Belmonte, *J. Mol. Cell Biol.*, 2012, **4**, 262-265.
[63] L. Fang, L. Zhang, W. Wei, X. Jin, P. Wang, Y. Tong, J. Li, J. X. Du, J. Wong, *Mol. Cell*, 2014, **55**, 537-551.
[64] D. S. Yoon, Y. Choi, Y. Jang, M. Lee, W. J. Choi, S.-H. Kim, J. W. Lee, *Stem Cells*, 2014, **32**, 3219-3231.
[65] H. Jang, T. W. Kim, S. Yoon, S.-Y. Choi, T.-W. Kang, S.-Y. Kim, Y.-W. Kwon, E.-J. Cho, H.-D. Youn, *Stem Cell*, 2012, **11**, 62-74.
[66] J. Brumbaugh, Z. Hou, J. D. Russell, S. E. Howden, P. Yu, A. R. Ledvina, J. J. Coon, J. A. Thomson, *Proc. Natl. Acad. Sci. USA*, 2012, **109**, 7162-7168.
[67] S. Dan, B. Kang, X. Duan, Y.-J. Wang, *Biochem. Biophys. Res. Commun.*, 2015, **456**, 714-720.
[68] Y. Cho, H. G. Kang, S.-J. Kim, S. Lee, S. Jee, S. G. Ahn, M. J. Kang, J. S. Song, J.-Y. Chung, E. C. Yi, K.-H. Chun, *Cell Death Differ.*, 2018, **25**, 1781-1795.
[69] C. A. C. Williams, A. Soufi, S. M. Pollard, *Semin. Cancer Biol.*, 2019, **67**, 30-38.
[70] X. Abulaiti, H. Zhang, A. Wang, N. Li, Y. Li, C. Wang, X. Du, L. Li, *Stem Cell Rep.*, 2017, **9**, 1630-1641.
[71] D. K. Kim, B. Song, S. Han, H. Jang, S.-H. Bae, H. Y. Kim, S.-H. Lee, S. Lee, J. K. Kim, H.-S. Kim, K.-M. Hong, B. I. Lee, H.-D. Youn, S.-Y. Kim, S. W. Kang, H. Jang, *Cancers*, 2020, **12**, article no. 2601.
[72] N. P. Mullin, J. Varghese, D. Colby, J. M. Richardson, G. M. Findlay, I. Chambers, *FEBS Lett.*, 2021, **595**, 14-25.
[73] K. T. G. Rigbolt, T. A. Prokhorova, V. Akimov, J. Henningsen, P. T. Johansen, I. Kratchmarova, M. Kassem, M. Mann, J. V. Olsen, B. Blagoev, *Sci. Signal*, 2011, **4**, article no. rs3.
[74] Y. Kamachi, H. Kondoh, *Development*, 2013, **140**, 4129-4144.
[75] J. Shin, T. W. Kim, H. Kim, H. J. Kim, M. Y. Suh, S. Lee, H.-T. Lee, S. Kwak, S.-E. Lee, J.-H. Lee, H. Jang, E.-J. Cho, H.-D. Youn, *Elife*, 2016, **5**, article no. e10877.
[76] P. N. Malak, B. Dannenmann, A. Hirth, O. C. Rothfuss, K. Schulze-Osthoff, *Cell Cycle*, 2015, **14**, 3748-3754.
[77] T. Schaefer, C. Lengerke, *Oncogene*, 2020, **39**, 278-292.
[78] J. Ouyang, W. Yu, J. Liu, N. Zhang, L. Florens, J. Chen, H. Liu, M. Washburn, D. Pei, T. Xie, *J. Biol. Chem.*, 2015, **290**, 22782-22794.
[79] S. Lim, A. Bhinge, S. Bragado Alonso, I. Aksoy, J. Aprea, C. F. Cheok, F. Calegari, L. W. Stanton, P. Kaldis, *Mol. Cell Biol.*, 2017, **37**, article no. e00201-17–24.
[80] H. J. Kim, J. Shin, S. Lee, T. W. Kim, H. Jang, M. Y. Suh, J.-H. Kim, I.-Y. Hwang, D. S. Hwang, E.-J. Cho, H.-D. Youn, *Nucleic Acids Res.*, 2018, **46**, 6544-6560.
[81] M. Moretto-Zita, H. Jin, Z. Shen, T. Zhao, S. P. Briggs, Y. Xu, *Proc. Natl. Acad. Sci. USA*, 2010, **107**, 13312-13317.
[82] S.-H. Kim, M.-O. Kim, Y.-Y. Cho, K. Yao, D. J. Kim, C.-H. Jeong, D. H. Yu, K. B. Bae, E.-J. Cho, S. K. Jung, M. H. Lee, H. Chen, J. Y. Kim, A. M. Bode, Z. Dong, *Stem Cell Res.*, 2014, **13**, 1-11.
[83] J. Brumbaugh, J. D. Russell, P. Yu, M. S. Westphall, J. J. Coon, J. A. Thomson, *Stem Cell Rep.*, 2014, **2**, 18-25.
[84] A. Saunders, D. Li, F. Faiola, X. Huang, M. Fidalgo, D. Guallar, J. Ding, F. Yang, Y. Xu, H. Zhou, J. Wang, *Stem Cell Rep.*, 2017, **8**, 1115-1123.
[85] L. Liu, W. Michowski, H. Inuzuka, K. Shimizu, N. T. Nihira, J. M. Chick, N. Li, Y. Geng, A. Y. Meng, A. Ordureau, A. Kołodziejczyk, K. L. Ligon, R. T. Bronson, K. Polyak, J. W. Harper, S. P. Gygi, W. Wei, P. Sicinski, *Nat. Cell Biol.*, 2017, **19**, 177-188.
[86] L. Liu, W. Michowski, A. Kolodziejczyk, P. Sicinski, *Nat. Cell Biol.*, 2019, **21**, 1060-1067.
[87] S. Jirawatnotai, S. Dalton, M. Wattanapanitch, *Semin. Cell Dev. Biol.*, 2020, **107**, 63-71.
[88] R. Spelat, F. Ferro, F. Curcio, *J. Biol. Chem.*, 2012, **287**, 38279-38288.
[89] K. B. Bae, D. H. Yu, K. Y. Lee, K. Yao, J. Ryu, D. Y. Lim, T. A. Zykova, M.-O. Kim, A. M. Bode, Z. Dong, *Stem Cell Rep.*, 2017, **9**, 2050-2064.
[90] Y. Hao, X. Fan, Y. Shi, C. Zhang, D. Sun, K. Qin, W. Qin, W. Zhou, X. Chen, *Nat. Commun.*, 2019, **10**, 1-13.
[91] N. S. Sharma, V. K. Gupta, P. Dauer, K. Kesh, R. Hadad, B. Giri, A. Chandra, V. Dudeja, C. Slawson, S. Banerjee, S. M. Vickers, A. Saluja, S. Banerjee, *Theranostics*, 2019, **9**, 3410-3424.
[92] D. K. Kim, J.-S. Lee, E. Y. Lee, H. Jang, S. Han, H. Y. Kim, I.-Y. Hwang, J.-W. Choi, H. M. Shin, H. J. You, H.-D. Youn, H. Jang, *Exp. Mol. Med.*, 2021, **53**, 1759-1768.
[93] S. Constable, J.-M. Lim, K. Vaidyanathan, L. Wells, *Glycobiology*, 2017, **27**, 927-937.
[94] T. Miura, S. Nishihara, *Trends Glycosci. Glycotechnol.*, 2019, **31**, E69-E75.
[95] L. Ciraku, E. M. Esquea, M. J. Reginato, *Cell. Signal*, 2022, **90**, article no. 110201.

[96] J. Ma, C. Hou, C. Wu, *Chem. Rev.*, 2022, **122**, 15822-15864.
[97] A. M. Gronenborn, D. R. Filpula, N. Z. Essig, A. Achari, M. Whitlow, P. T. Wingfield, G. M. Clore, *Nature*, 1991, **253**, 657-661.
[98] C. K. Smith, J. M. Withka, L. Regan, *Biochemistry*, 1994, **33**, 5510-5517.
[99] A. Alik, C. Bouguechtouli, M. Julien, W. Bermel, R. Ghouil, S. Zinn-Justin, F.-X. Theillet, *Angew. Chem. Int. Ed.*, 2020, **59**, 10411-10415.
[100] M. Howarth, K. Takao, Y. Hayashi, A. Y. Ting, *Proc. Natl. Acad. Sci. USA*, 2005, **102**, 7583-7588.
[101] M. Fairhead, M. Howarth, *Methods Mol. Biol.*, 2015, **1266**, 171-184.
[102] F.-X. Theillet, A. Binolfi, B. Bekei, A. Martorana, H. M. Rose, M. Stuiver, S. Verzini, D. Lorenz, M. van Rossum, D. Goldfarb, P. Selenko, *Nature*, 2016, **530**, 45-50.
[103] R. Dass, F. A. A. Mulder, J. T. Nielsen, *Sci. Rep.*, 2020, **10**, article no. 14780.
[104] K. Tamiola, F. A. A. Mulder, *Biochem. Soc. Trans.*, 2012, **40**, 1014-1020.
[105] J. T. Nielsen, F. A. A. Mulder, *J. Biomol. NMR*, 2018, **70**, 141-165.
[106] W. Borcherds, F.-X. Theillet, A. Katzer, A. Finzel, K. M. Mishall, A. T. Powell, H. Wu, W. Manieri, C. Dieterich, P. Selenko, A. Loewer, G. W. Daughdrill, *Nat. Chem. Biol.*, 2014, **10**, 1000-1002.
[107] C. Camilloni, A. De Simone, W. F. Vranken, M. Vendruscolo, *Biochemistry*, 2012, **51**, 2224-2231.
[108] F.-X. Theillet, C. Smet-Nocca, S. Liokatis, R. Thongwichian, J. Kosten, M.-K. Yoon, R. W. Kriwacki, I. Landrieu, G. Lippens, P. Selenko, *J. Biomol. NMR*, 2012, **54**, 217-236.
[109] F.-X. Theillet, H. M. Rose, S. Liokatis, A. Binolfi, R. Thongwichian, M. Stuiver, P. Selenko, *Nat. Protoc.*, 2013, **8**, 1416-1432.
[110] A. Mylona, F.-X. Theillet, C. Foster, T. M. Cheng, F. Miralles, P. A. Bates, P. Selenko, R. Treisman, *Science*, 2016, **354**, 233-237.
[111] M. Julien, C. Bouguechtouli, A. Alik, R. Ghouil, S. Zinn-Justin, F.-X. Theillet, in *Intrinsically Disordered Proteins: Methods and Protocols* (B. B. Kragelund, K. Skriver, eds.), Springer US, New York, 2020, 793-817.
[112] N. Festuccia, N. Owens, A. Chervova, A. Dubois, P. Navarro, *Development*, 2021, **148**, article no. dev199604.
[113] J.-P. Lambert, M. Tucholska, T. Pawson, A.-C. Gingras, *J. Proteomics*, 2014, **100**, 55-59.
[114] P. Poullet, S. Carpentier, E. Barillot, *Proteomics*, 2007, **7**, 2553-2556.
[115] M. The, M. J. MacCoss, W. S. Noble, L. Käll, *J. Am. Soc. Mass Spectrom*, 2016, **27**, 1719-1727.
[116] B. Valot, O. Langella, E. Nano, M. Zivy, *Proteomics*, 2011, **11**, 3572-3577.
[117] C. Smet-Nocca, H. Launay, J.-M. Wieruszeski, G. Lippens, I. Landrieu, *J. Biomol. NMR*, 2013, **55**, 323-337.
[118] S. Elkjær, A. D. Due, L. F. Christensen, F. F. Theisen, L. Staby, B. B. Kragelund, K. Skriver, *Commun. Biol.*, 2023, **6**, article no. 63.
[119] W. Peti, R. Page, *Protein Sci.*, 2013, **22**, 1698-1710.
[120] D. J. Wood, J. A. Endicott, *Open Biol.*, 2018, **8**, article no. 180112.
[121] D. L. Sheridan, Y. Kong, S. A. Parker, K. N. Dalby, B. E. Turk, *J. Biol. Chem.*, 2008, **283**, 19511-19520.
[122] S. Liokatis, A. Stützer, S. J. Elsässer, F.-X. Theillet, R. Klingberg, B. van Rossum, D. Schwarzer, C. D. Allis, W. Fischle, P. Selenko, *Nat. Struct. Mol. Biol.*, 2012, **19**, 819-823.
[123] S. A. Lambert, A. Jolma, L. F. Campitelli, P. K. Das, Y. Yin, M. Albu, X. Chen, J. Taipale, T. R. Hughes, M. T. Weirauch, *Cell*, 2018, **172**, 650-665.
[124] J. Liu, N. B. Perumal, C. J. Oldfield, E. W. Su, V. N. Uversky, A. K. Dunker, *Biochemistry*, 2006, **45**, 6873-6888.
[125] I. Yruela, C. J. Oldfield, K. J. Niklas, A. K. Dunker, *Genome Biol. Evol.*, 2017, **9**, 1248-1265.
[126] K. Cermakova, H. C. Hodges, *Trends Biochem. Sci.*, 2023, **48**, 477-490.
[127] X. Z. Zhou, K. P. Lu, *Nat. Struct. Mol. Biol.*, 2016, **16**, 463-478.
[128] Y. Chen, Y. Wu, H. Yang, X. Li, M. Jie, C. Hu, Y. Wu, S. Yang, Y. Yang, *Cell Death Dis.*, 2018, **9**, article no. 883.
[129] J. Gebel, M. Tuppi, A. Chaikuad, K. Hötte, M. Schröder, L. Schulz, F. Lohr, N. Gutfreund, F. Finke, E. Henrich, J. Mezhyrova, R. Lehnert, F. Pampaloni, G. Hummer, E. H. K. Stelzer, S. Knapp, V. Dötsch, *Nat. Chem. Biol.*, 2020, **16**, 1078-1086.
[130] W. Qin, K. F. Cho, P. E. Cavanagh, A. Y. Ting, *Nat. Methods*, 2021, **18**, 133-143.
[131] J.-P. Lambert, M. Tucholska, C. Go, J. D. R. Knight, A.-C. Gingras, *J. Proteomics*, 2015, **118**, 81-94.
[132] A.-C. Gingras, K. T. Abe, B. Raught, *Curr. Opin. Chem. Biol.*, 2019, **48**, 44-54.
[133] X. Liu, K. Salokas, R. G. Weldatsadik, L. Gawriyski, M. Varjosalo, *Nat. Protoc.*, 2020, **15**, 3182-3211.
[134] D.-P. Minde, M. Ramakrishna, K. S. Lilley, *Commun. Biol.*, 2020, **3**, article no. 38.
[135] H. Göös, M. Kinnunen, K. Salokas, Z. Tan, X. Liu, L. Yadav, Q. Zhang, G.-H. Wei, M. Varjosalo, *Nat. Commun.*, 2022, **13**, article no. 766.
[136] C. P. Wigington, J. Roy, N. P. Damle, V. K. Yadav, C. Blikstad, E. Resch, C. J. Wong, D. R. Mackay, J. T. Wang, I. Krystkowiak, D. A. Bradburn, E. Tsekitsidou, S. H. Hong, M. A. Kaderali, S.-L. Xu, T. Stearns, A.-C. Gingras, K. S. Ullman, Y. Ivarsson, N. E. Davey, M. S. Cyert, *Mol. Cell*, 2020, **79**, 342-358, e12.
[137] Y. Ueki, T. Kruse, M. B. Weisser, G. N. Sundell, M. S. Y. Larsen, B. L. Mendez, N. P. Jenkins, D. H. Garvanska, L. Cressey, G. Zhang, N. Davey, G. Montoya, Y. Ivarsson, A. N. Kettenbach, J. Nilsson, *Mol. Cell*, 2019, **76**, 953-964, e6.
[138] C. Benz, M. Ali, I. Krystkowiak, L. Simonetti, A. Sayadi, F. Mihalic, J. Kliche, E. Andersson, P. Jemth, N. E. Davey, Y. Ivarsson, *Mol. Syst. Biol.*, 2022, **18**, article no. e10584.
[139] N. E. Davey, L. Simonetti, Y. Ivarsson, *Trends Biochem. Sci.*, 2022, **47**, 547-548.
[140] G. Krainer, T. J. Welsh, J. A. Joseph, J. R. Espinosa, S. Wittmann, E. de Csilléry, A. Sridhar, Z. Toprakcioglu, G. Gudiškytė, M. A. Czekalska, W. E. Arter, J. Guillén-Boixet, T. M. Franzmann, S. Qamar, P. S. George-Hyslop, A. A. Hyman, R. Collepardo-Guevara, S. Alberti, T. P. J. Knowles, *Nat. Commun.*, 2021, **12**, article no. 1085.
[141] K. Tsafou, P. B. Tiwari, J. D. Forman-Kay, S. J. Metallo, J. A. Toretsky, *J. Mol. Biol.*, 2018, **430**, 2321-2341.
[142] A. Chen, A. N. Koehler, *Trends Mol. Med.*, 2020, **26**, 508-518.

Comptes Rendus
Chimie
2023, Vol. 26, Special Issue S3, p. 139-155
https://doi.org/10.5802/crchim.309

Breaking Barriers in Chemical Biology – Toulouse 2022

Release of biologically relevant substrates from prodrugs and nanocarriers using X- and γ-rays as trigger signals

Amit Kumar [a] **and Peter István Dalko** [*, a]

[a] Laboratoire de Chimie et Biochimie Pharmacologiques et Toxicologiques, CNRS, Université Paris Cité, Paris, France

E-mail: peter.dalko@parisdescartes.fr (P. I. Dalko)

Abstract. X- and γ-ray-excited theranostic techniques are increasingly being developed for biomedical applications, as they overcome the limitations of light penetration and tissue attenuation. Beyond the historic opportunities that X-rays have opened up for medical imaging and cancer therapy, this modality is increasingly used for drug delivery from pro-drugs and nanomaterials. This brief review covers several strategies that have been developed for biomedical applications

Keywords. Redox probes, Smart nanomaterials, Radiolysis, Reactive oxygen species (ROS), Combination therapy.

Manuscript received 17 March 2023, revised 28 November 2023 and 10 January 2024, accepted 3 April 2024.

1. Introduction

Biomedical applications of X-rays were and still are closely linked to cancer imaging and therapy. Radiotherapy (RT) is an extremely important first-line treatment for cancer therapy which, unlike other treatments such as chemotherapy or immunotherapy which affect the whole body, is generally a local treatment. Despite the spectacular progress made in the field of X-ray sensitization and improvement, radiotherapy as a monotherapy generally fails to eliminate the entire tumor, as cancer cells can undergo DNA repair and regrow [1,2]. For this reason, current research is gradually shifting from monotherapy to multimodal X-ray therapy, combining two or more therapeutic modalities, producing synchronous/synergistic anticancer effects [3,4]. Increasing synergistic effect also optimizes several shortcomings of drug-based therapies by improving systemic effect, nonspecific biodistribution, poor bioavailability and limited circulation time, which can lead to severe side effects and low therapeutic index. Chemoradiotherapy has been widely used clinically, offering substantial advantages in terms of treatment and prognosis. Research on the simultaneous administration of drugs and the application of therapeutic beams dates back several decades [5, 6] and is used to treat certain stages of head and neck cancer, lung cancer, and gynecological cancers. Now the field has reached a next-level maturity to experiment with locally activable prodrugs and drug carriers, thus limiting the general chemotoxicity and efficiency of the treatment. Stimuli-responsive biodegradable nanoagents, which can not only deliver and control chemotherapies, but also attenuate hypoxia, or play as antimitotic agent to enhance chemoradiotherapy, have enormous potential. Key to the use of these smart materials is the presence of an interplay that is able to switch from one state to another, triggering the release of the covalently or

* Corresponding author

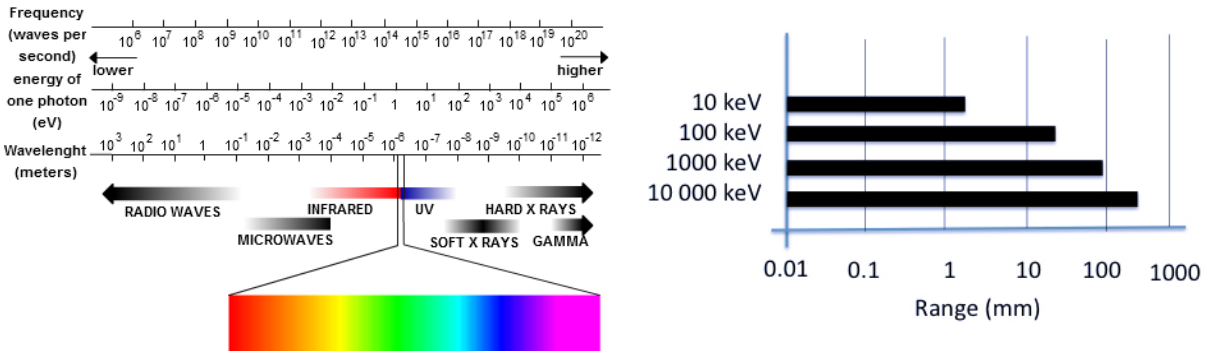


Figure 1. Electromagnetic map and modeled soft-tissue penetration capacity (calculated values).

non-covalently attached compound. Intelligent drug delivery systems can be designed around chemical functions or fragments of molecules that modify their structural or functional properties in engineered materials in response to specific external or internal factors [7,8]. These dynamic properties can be activated by chemical/biochemical, electrical, magnetic, electromagnetic/optical, ultrasonic, dielectric, redox or piezoelectric signals, some of internal (biological) origin, others of external origin. These interplay units, which are small chemical compounds, may undergo isomerization, rearrangement or cleavage of chemical bonds, such as UV-sensitive *azobenzene* and *o*-nitrobenzyl derivatives, or coumarin esters, leading to direct or indirect drug release. Some of these molecules were historically developed for neuroscience where these "caged" compounds revolutionized functional neurophysiology from the late 1970s. They are now commonly used to design prodrugs and light-controlled drug delivery systems, although their application in therapy is limited [9–11]. Light activation has inherent shortcomings. While the tissue penetration capacity of visible and near-infrared (NIR) light is superior to the more often used UV activation, their penetration in deep tissues does not extend beyond a few hundred micrometers using intensities below cell toxicity. The main obstacles are strong absorption and diffraction in turbid tissues. Fortunately, longer and especially shorter wavelengths than UV/Vis/NIR penetrate tissues better (Figure 1). As longer waves access vibrational and translational excitations, thus have not enough energy to break a chemical bond, shorter waves can provide sufficient energy to induce structural changes at the molecular level. As organic compounds have no electronic transitions in this energy window, they are transparent. Several ingenious strategies have recently been developed to translate short-wave photons into a trigger signal applicable to organic compounds that will be discussed in this review [12–19]. This strategy also opens up new perspectives for personalized medical care, bringing the dream of tailor-made medical treatments for individuals closer to reality.

1.1. *On the nature of X- and γ-rays*

Although X- and γ-rays are both electromagnetic radiation (photons), there is no consensus on how to precisely differentiate between them. They differ mainly in their origin: γ-rays are produced during the nuclear decay of atomic nuclei from the atomic core and particle annihilation, while X-rays are produced when the velocity of electrons (or other charged particles) slows down, and the energy difference between these two kinetic states is dispersed as high-energy photons by Bremsstrahlung and characteristic radiation [20,21]. X-rays are generally generated in a vacuum tube, where electrons emitted by the cathode, accelerated under high voltage, collide with a tungsten metal anode. X- and γ-rays are ionizing radiations but less so than alpha and beta particles. The energy ranges of X- and γ-rays overlap considerably, with X-rays essentially covering the "lower" part of the spectrum with energies above a few tens of eV, typically between 120 eV and 120 keV (thus the shortest ultraviolet spectrum) and γ-ray energies extending up to a few tens of MeV. In other words,

X-rays (generally) have longer wavelengths than γ-rays, mainly ranging from 0.01 to 10 nm. Both γ- and X-rays have a high penetration power in soft tissue, whose extent depends on their respective wavelengths. Penetration is greater at higher energies but also depends on the tissue type. The impact of tissue thickness on X-ray penetration is dramatic. Small increases in tissue thickness led to large attenuations in the intensity of the incident beam. Because of the critical importance of penetration into (biological) media, highly accurate transport equations were developed from the early 1950s and are currently used to estimate the exposure of biological tissues to radiation [22]. Indeed, an utmost important value in biomedical applications is the dose absorbed by the living body. The gray (symbol: Gy) is the unit of ionizing radiation dose that measures the energy deposited/absorbed (in joule per kilogram of matter).

Radiotherapy is used in cancer therapy as ionizing radiation can break down the DNA of cancer cells more effectively than that of healthy cells. This is made possible by the natural repair mechanism, which is more effective for normal cells than for cancer cells, which function in an intrinsically dysfunctional way. The upper limit of a "safe dose" has been defined as a total dose of 55–60 Gy administered to a focal field with fractions of 1.8–2 Gy per day. This means that the total irradiation dose is delivered in small pieces to be gentler on normal tissues and to allow normal cells to repair DNA damage after each treatment. The radiation threshold depends sensitively on the tissue/organ/type of cancer/general condition of the patient and is assessed individually for each patient. An example of a dose-response relationship is illustrated in the treatment of stage III non-small cell lung cancer, where the administration of a total dose of 40 Gy is applied in fractionated treatment (4 Gy/day for five days, followed by a two-week break and a further dose of 4 Gy/day for five days), or the administration of a total dose of 40 Gy, 50 Gy or 60 Gy is applied in continuous treatment (2 Gy/day) [23]. For this reason, drug delivery probes for simultaneous activation must meet the threshold of 2 to 4 Gy, meaning that they must release the maximum payload within this dose for biomedical use.

How can this high-energy radiation be used for prodrug activation and for drug delivery? Three conceptually different approaches have been developed, based either (i) on transformations mediated by energy or by an electron transfer-promoted change in the electronic structure (thus in the stability) of the carrier; (ii) on the incorporation of ROS-sensitive functions into the drug delivery matrix that react chemically and undergo fragmentation; (iii) on the upregulation of enzymes that have recently been explored for selective prodrug activation, but the scope of such activation is beyond the scope of this study [24].

1.2. *Drug release triggered by energy- or electron transfer-mediated structural changes*

Organic probes are transparent to X-rays, as organic materials do not exhibit absorption transitions in this high-energy spectrum. Consequently, irradiation can be used for sterilization to eliminate microbial growth and prepare pyrogen-free samples without altering the structure of the biomedical sample or equipment [25,26]. To be able to explore X/γ-rays as a trigger, an intermediate transduction mechanism is needed, which bridges the gap between high-energy electromagnetic radiation and what organic compounds can "understand" to react. Fortunately, there are a few such interactions. Noteworthily not necessarily the "simplest" solutions are the most effective. For example, the light produced by Cerenkov radiation, which is emitted when a charged particle passes through a dielectric medium at a speed greater than the phase velocity of light in that medium, would be useful as a local light trigger signal if the intensity of the UV light generated by this phenomenon could be increased [27]. Similarly, parametric down-conversion (an instantaneous nonlinear optical process that converts a higher-energy photon into a lower-energy photon pair, in accordance with the law of conservation of energy and the law of conservation of momentum) would be an interesting solution, although it cannot yet be used due to the lack of efficient generation of idler photons at optical wavelengths [28]. The conversion of X-rays to UV can be achieved more easily using scintillators (see above) [29,30]. These materials play an important role in X-ray-induced photodynamic therapy (PDT), enabling low doses of radiation to be applied, as do a variety of well-designed irradiation modes and intelligent strategies for modifying the tumor microenvironment [31]. This activation

modality generates $^1O_2$ and ROS similarly to conventional photosensitizers in PDT under ultraviolet (UV) or visible light (scintillator-derived strategies will be discussed in the paragraph on redox/electron transfer reactions).

The most frequently used electron/energy transfer-related activation strategies are based on inelastic scattering, which involves a loss of energy from the incident primary photon: the shorter wavelength primary photon is transformed to a longer wavelength secondary photon. Inelastic scattering occurs via a number of mechanisms. The energy is transferred to the sample, generating a series of useful signals that can be exploited in a number of analytical methods and chemical transformations. The interaction of X- and γ-rays with matter is described by the photoelectric effect, Compton scattering, and pair production. The photoelectric effect is a phenomenon whereby electrically charged particles are released from or within a material when it absorbs electromagnetic radiation (Figure 2). The effect is often defined as the ejection of electrons from an atom when light falls on it, and called the Auger–Meitner effect at higher frequencies, such as X- and γ-rays. The effect consists in the occupation of a cavity in the inner shell of an atom, accompanied by the emission of an electron from the same atom. When an electron is removed from the atom, leaving a vacancy, an electron of a higher energy level can fall into the vacancy, resulting in a release of energy (Figure 2). Although this energy is most often released in the form of an emitted photon, it can also be transferred to another electron, which is ejected from the atom; this second ejected electron is called an Auger electron. Since an electron vacancy initially produced by the incident photon leads to two new electron vacancies, which in turn can produce the same number, a complete cascade of electrons can be produced according to tabulated probabilities. The atom is finally left in a state of multiple ionizations. The Auger effect (nonradiative relaxation) is the predominant mode of relaxation in light target atoms such as C, N, O ($Z < 20$) or for L-layer ionizations.

The radiative mode corresponds to the emission of fluorescence radiation, whose spectrum of discontinuous lines is characteristic of the material. In heavy materials ($Z > 50$) and for K-layer vacancies, fluorescence is the main relaxation mode.

Compton scattering results from the collision of

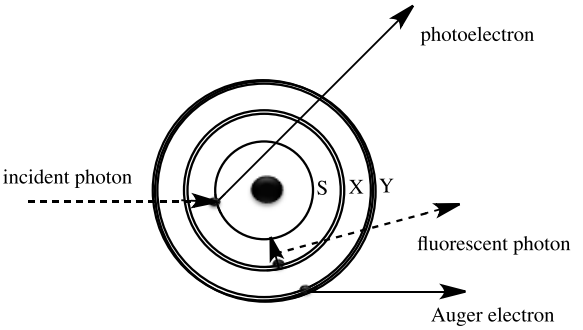


Figure 2. The photoelectric effect is the consequence of inelastic scattering between the photon and the electron, which generates a nonradiative relaxation (Auger effect) and a radiative relaxation in the target atom.

a photon with an electron at rest: this is the combination of an incident X-ray photon and a weakly bound electron in the atom, whose binding energy is negligible compared to the incident energy. During this interaction, the incident X-ray photon disappears: part of its energy is given to the peripheral electron, which is then ejected: this is known as a Compton electron; if the other part of its energy is transmitted in the form of a new X-ray photon, this is known as a scattered or recoil photon. At the energies used in radiology, the Compton effect is favored in low-density, water-equivalent (low Z) media (fat, soft tissue, etc.). The Compton effect also dominates the interaction processes of photons with energies between 0.2 and a few MeV, whatever the atomic number of the medium through which they pass. The probability of occurrence of these scatterings decreases as the energy of the incident photon increases. The photoelectric effect and Compton scattering can be enhanced by the presence of a captive sensor and can also be exploited by Förster-type energy transfer, which can take place between the sensor (often derived from a metal) and the organic effector, resulting in a change in the electronic configuration of the organic probe.

Pair production only occurs at high energies and is statistically rare under normal biomedical conditions. Through this effect, the γ-ray is transformed into matter in the form of a pair of negatively and positively charged particles (electron and positron). Since an electron has a rest mass equiva-

lent to 0.511 MeV of energy, a minimum γ energy of 1.02 MeV is required to produce this pair. Any excess energy of the γ-ray producing the pair is given to the electron–positron pair in the form of kinetic energy. In most cases, the positron will undergo annihilation by reaction with an electron in the detector material, creating two γ photons of 0.511 MeV each.

Photoelectric absorption dominates at low energies, followed by Compton scattering, then pair production as energy increases. Absorption of very high-energy photons leads to nuclear decay. The above processes (with the exception of photodisintegration) all result in the production of electrons (or their antimatter equivalent, positrons) and lower-energy X-rays, which undergo further absorption and scattering. In addition, these heavy nanoparticles can selectively scatter and/or absorb high-energy radiation [32–34], leading to energy transfer to photosensitizers, often porphyrin derivatives, metal complexes, scintillators or quantum dots that are sufficiently stable under radiative conditions [35].

1.3. *Chemical transformation of ROS-sensitive functions*

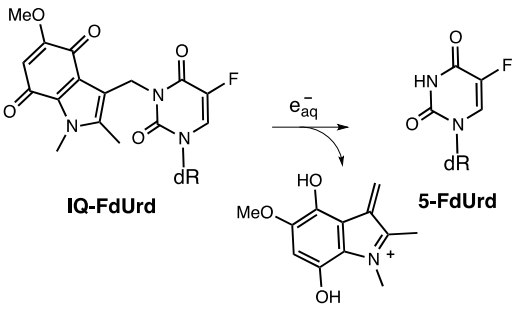
Materials sensitive to reactive oxygen species (ROS) can form covalent bonds with the transiently generated intermediates formed by the ionizing radiation. This can be facilitated by the presence of electron donor shuttles, or simply the presence of large quantities of water, ubiquitous in living organisms, which produce a large variety of reactive intermediates by irradiation [36]. When exposed to light or ionizing radiation, water undergoes a series of transformations involving electronic excitations, ionization of solvated species and the formation of free radicals and solvated electrons [37–42]. The radiolysis of water generates thus various reactive particles, in which the $^{\bullet}$OH and the $e^{-}_{(aq)}$ are the dominant products. These processes occur, for example, in (photo)electrochemical cells [43,44], biological molecules [45], as well as in atmospheric water [46]. Many of these are governed, at least in part, by the nature of electron transfer and binding energies in aqueous solutions. Among the many reactive species formed during this reaction are hydrated electrons and ROS [47] such as peroxides, singlet oxygen and oxyradicals, which are considered the chemically active component of radiation therapy. The HO$^{\bullet}$ radical (hydroxyl radical) is the most abundant radical, which can then react with surrounding tissues to create organic radicals (usually alkyl radicals). Historically, iodine and gold-based nanoparticles (AuNPs) were the first radiosensitizers due to their high X-ray absorption, synthetic versatility and unique chemical, electronic and optical properties [48]. The radiosensitization dates back to the pioneering work of Hainfeld, who demonstrated that gold nanoparticles behave as effective radioactive agents, since tumor-bearing animals survive longer when treated with radiotherapy after intravenous injection of ultra-small gold nanoparticles [49]. The energy of the initial photon is thus transferred to electrons, which create ionization that produces *in fine* ROS in the presence of water and oxygen, leading to significant chemical and biological effects such as DNA degradation. The range of radiosensitizers now extends to other heavy elements with high atomic numbers, such as hafnium, gadolinium or bismuth atoms/nanoparticles [34,50–53] which are well-known radiosensitizers, capable of amplifying radiation doses in tumor tissue [54–56].

As mentioned in the previous paragraph, γ/X-ray activation enables "second-generation" photodynamic therapy (PDT), which relies essentially on ROS generation and is driven by the activation of a photosensitizer (PS), followed by the transfer of excitation energy to molecular oxygen to result in the generation of cytotoxic ROS. Noteworthy, the metal–photosensitizer combination is superior to the use of metal particles alone for ROS generation in terms of biological efficacy, and has a wider reach than the more conventional UV- and IR-triggered ROS generation, as it overcomes the penetration limits of UV or IR into tissue [16,57,58].

Fragmentation of ROS-sensitive bonds, such as sulfide/disulfide or selenide/diselenide bonds, or other chemical functions, as well as changes in the mechanical properties (rigidity) of the nanomaterial under the influence of radiation, can lead to the release of the sequestered ligand.

2. Activation of prodrugs or nanocarriers by electron or energy transfer

The prodrug strategy that was developed steadily in the second half of the last century to improve how



Scheme 1. Activation of indolequinone-tethered 5-fluorodeoxyuridine by electron transfer and subsequent drug release.

the drug is absorbed, distributed, metabolized, and excreted in the living body involves masking one of the key parts (usually polar functions) of a biologically active compound, necessary for the activity, and restoring it when the masking group is cleaved [59–61]. This strategy was also developed in neurosciences, physiology, and molecular biology by using light as trigger signal, making it possible to control drug activation in space and time. By analogy with light, tissue-penetrating X-ray can be used under particular conditions.

2.1. *Electroreduction of indolequinones*

Historically, the first report on X- and γ-ray molecular fragmentation by electron transfer was reported by the Tanabe group using tethered 5-fluorouracil (5-FU) and 5-fluorodeoxyuridine (5-FdUrd) with 2-oxoalkyl groups covalently linked to radiation-sensitive quinone derivatives enabling the release of the antitumor agents 5-FU and 5-FdUrd, respectively (Scheme 1) [62,63].

Mechanistic studies have shown that these prodrugs undergo reduction by capturing hydrated electrons ($e^-_{(aq)}$) generated by the radiolysis of water. Optimization of the anchoring group structure has yielded indolequinones that can release the covalently attached ligand by bioreduction or under radiolytic conditions [64–67]. This transformation was presented as the first prototype transformation for the release of complex structures from encapsulated compounds sensitive to irradiation [68].

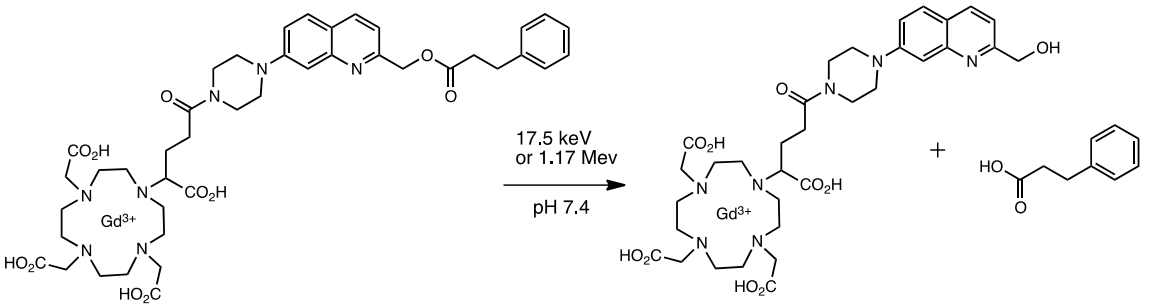
2.2. *Intramolecular electron or energy transfer from heavy elements*

The Gd-sensitized aminoquinolone probe was found to undergo radiolytic fragmentation with a covalently attached model dihydrocinnamate as substrate [4, 69]. Although the precise mechanism of the fragmentation has not been established, electron transfer or electronic coupling between the irradiated Gd complex and the quinolinium probe has been postulated. Radiolysis of the covalently attached probe by 17.5 keV and 1.17 MeV X-rays resulted in the liberation of the dihydrocinnamate model (Scheme 2).

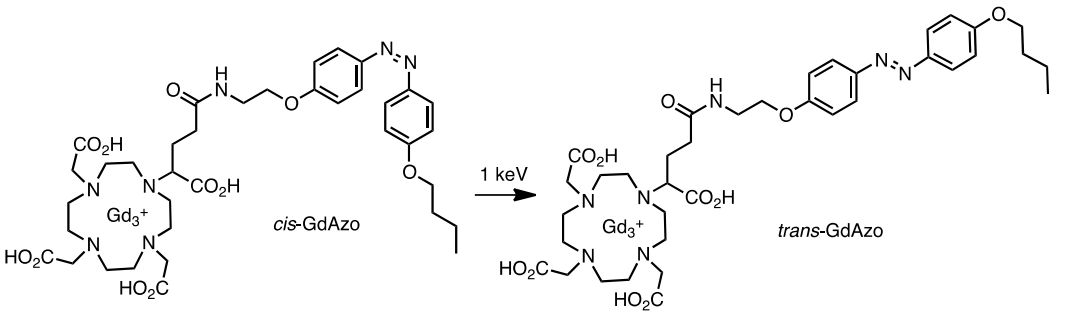
In a somewhat analog entry, a Gd chelate and azobenzene were combined in a radio-switch (Scheme 3) [70]. Azobenzene was isomerized by low-dose X-rays, offering the possibility of implementing trigger systems based on radiation penetrating the isomerization.

The activation efficiency (i.e., part of the isomerized (*trans*) compound compared to the starting (*cis*) material) was 33% and 69% at irradiation doses of 2 and 20 Gy, respectively, by using 1 keV source (~1.24 nm). No compounds other than *trans*-GdAzo were observed spectrophotometrically. This molecular activation based on a radio-switch seems promising and opens the way to new types of applications such as actionable tools in radiotherapy.

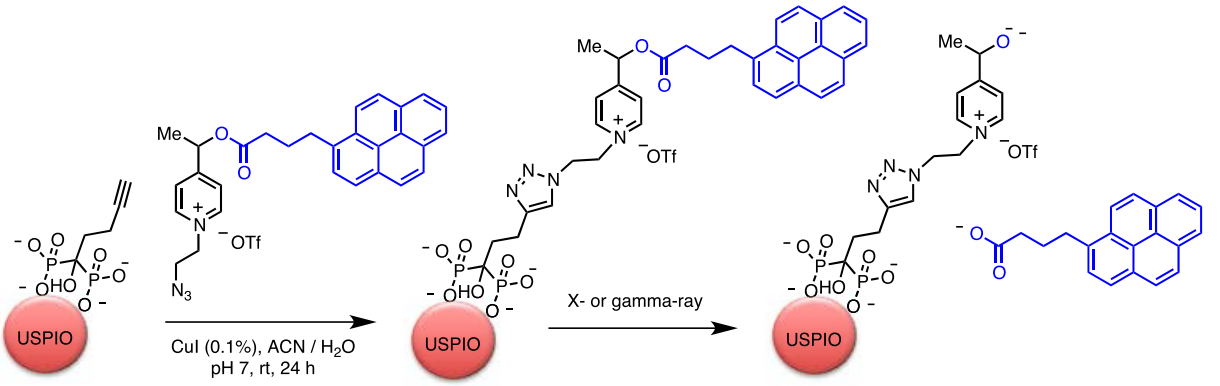
Solvated electrons and transiently formed radical species generated by X/γ-rays have been exploited in the fragmentation of sensitive redox probes derived from picolinium and immobilized on the surface of ultra-small superparamagnetic iron oxide nanoparticles (USPIO) [71]. The iron-derived nanoparticle enabled Auger/Compton electron generation upon X-ray or γ irradiation, as well as monitoring of probe biodistribution by T2-weighted MRI and by the appearance of fluorescence from the released probe (Scheme 4). The functionalized particles showed high colloidal stability at physiological pH. The picolinium group effectively quenched the fluorescence of the covalently attached pyrene reporter in the starting material, probably via Förster-type quenching and dark complex formation. Probe activation was validated under pulsed radiation and also by using a conventional Cs-137 source. Although fragmentation was observed under both conditions, pulsed sources proved considerably more effective: 82% of the fluorescence of the pyrene reporter was recovered upon



Scheme 2. Radiolysis of gadolinium-bound quinoline cage with release of dihydrocinnamate ligand.



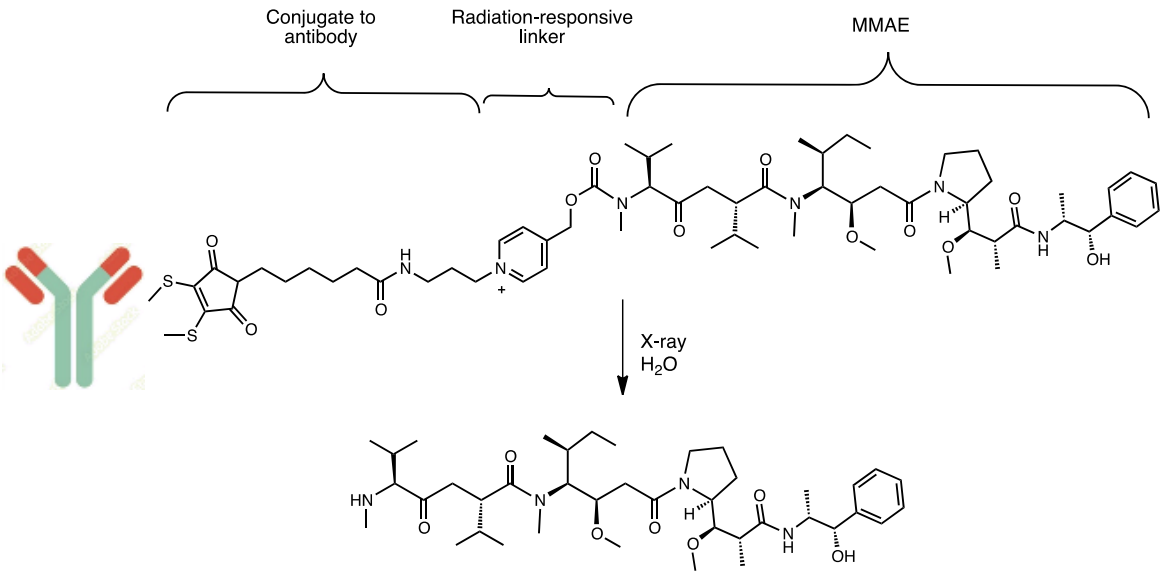
Scheme 3. Activation of *cis*-GdAzo to trans-GdAzo using a 1 keV source.



Scheme 4. Activation of pyrenebutyric acid immobilized by USPIO by X- or γ-ray irradiation.

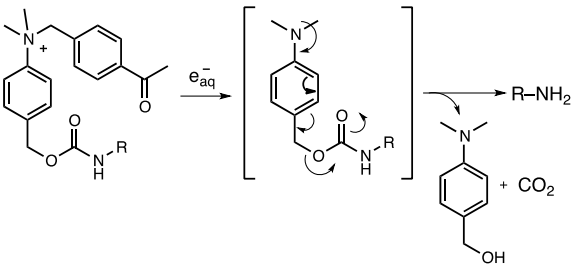
30 Gy irradiation, compared with 27% of the fluorescence under conventional conditions.

Several redox-sensitive compounds derived from picoline, quinolinine and picolinium were examined for X-ray activation [72]. 4-Hydroxymethylene pyridinium was selected as the most efficient probes for releasing client molecules in a dose gradient ranging from 0 to 60 Gy (Scheme 5). The cytotoxic payload (monomethyl auristatin E (MMAE)) was attached via a carbamate linker to the probe. A humanized antibody (sibrotuzumab) was attached to the release complex for better targeting (NAPC-ADC) (Scheme 5). The therapeutic efficacy of the radiation-induced activation was then assessed in vivo. Mice were irradiated with 3 × 4 Gy of X-rays on days 2, 4 and 6 post-injection. It should be noted that



Scheme 5. The NAPC-ADC conjugate undergo fragmentation by X-ray induced electron-transfer reaction in vivo to release a highly toxic MMAE.

the overall 12 Gy of X-ray dose was supposed to trigger the complete release of MMAE. Tumor growth inhibition in the NAPC-ADC + X-ray group was better: mean tumor size was around 85 ± 55 mm^3 at 21 days, whereas tumor size in the control (PBS only) and radiation-only groups was up to 1387 ± 463 and 433 ± 120 mm^3, respectively ($n = 5$). Administration of NAPC-ADC (5 mg·kg^{-1}) alone had no significant impact on tumor growth, suggesting that radiation-induced payload release contributed to tumor treatment.



Scheme 6. Radiolytic fragmentation of the quaternary ammonium group.

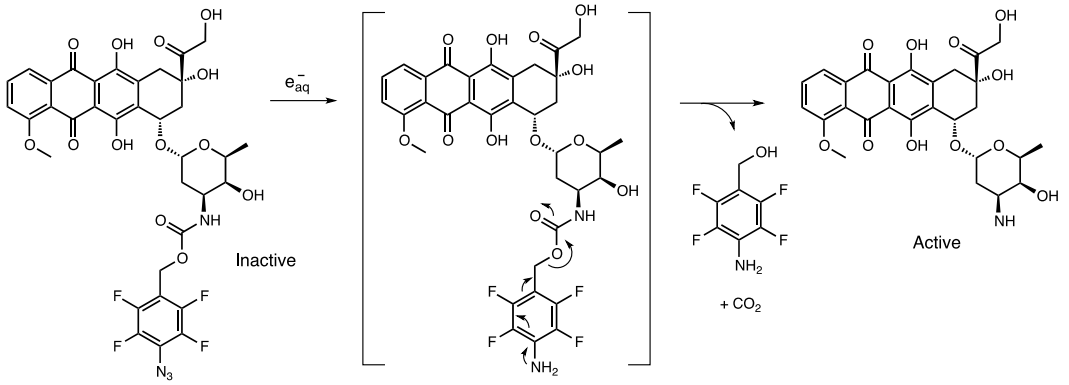
2.3. *Fragmentation of the quaternary ammonium group*

Similar to the C–N bond cleavage of quaternary organoammonium salts under photochemical conditions, quaternary ammoniums can be converted to tertiary amines under X-ray radiative conditions [73]. Fragmentation alters the polarization of the central benzyl group (i.e., through the transformation of an electron-withdrawing group (EWG) into an electron-donating substituent), triggering the 1,6-fragmentation of the benzyl-bound carbamate, thereby releasing the ligand (Scheme 6) [74].

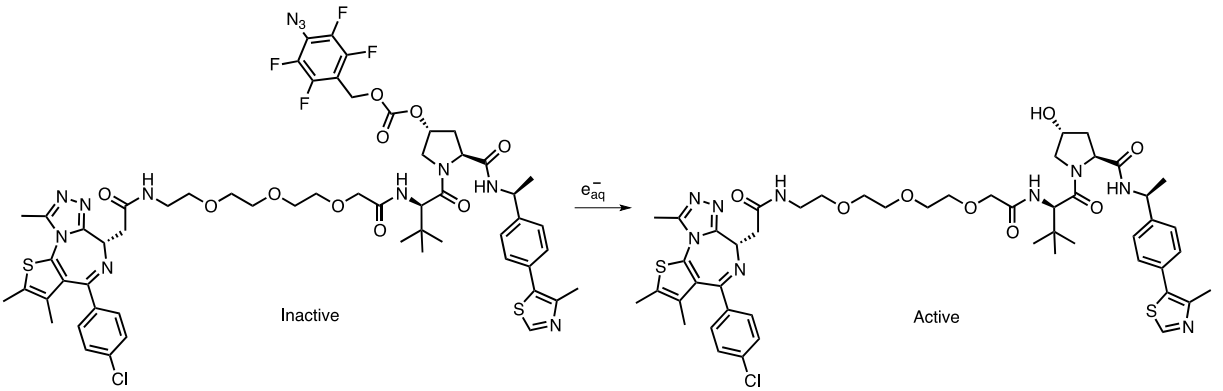
Carfilzomib, an anticancer drug with a morpholine motif, was released as a model substrate to demonstrate the feasibility of the in vitro strategy of radiation-induced prodrug release. Radiation-induced cleavage was demonstrated in living cells and in tumor-bearing mice, but high doses (60 Gy) were required to achieve satisfactory conversions.

3. Reduction of azides to amines

Prodrugs of pazopanib and doxorubicin encapsulated in sulfonyl azide and phenyl azide were released at irradiation doses of up to 60 Gy (Scheme 7) [75]. The driving force behind the transformation is the high reducing potential of e^- generated by the radiolysis of water, with a standard electrode potential of −2.77 V. Although the exact reduction mechanism is



Scheme 7. Radiolytic activation of doxorubicin.



Scheme 8. Activation of VHL E3 ligase by removing the tetrafluoroazidobenzyl carbonate protecting group by X-ray.

unclear, it appears to follow a similar pathway to the photolysis of sulfonyl azides in alcohols and water to generate sulfonamides via the formation of sulfonylamido radicals or nitrenes [76].

In a similar strategy, a tetrafluorobenzyl azide cage was used to mask the polar functions of the VHL E3 ligase via a carbonate linker. Continuous irradiation up to 60 Gy resulted in conversion of the azide group to an amine and triggered ligand release (Scheme 8) [77].

4. Activation of *N*-oxide prodrugs

The high electrode potential of electrons generated by local radiolysis of water can also effectively reduce nitrogen oxides [78]. The caging strategy involves oxidation of the target drug to *N*-oxide prodrugs, generating drugs whose biological activity is silenced. Radiolytic activation has been demonstrated from NO-imiquimod, NO-ampiroxicam, NO-pranoprofen, NO-loratadine, and NO-camptothecin by X-ray-induced electron transfer using radiation up to 60 Gy (Figure 3).

5. Chemical transformation of ROS-sensitive functions

5.1. *Oxidation of thioethers*

Nanovesicles were assembled from ROS-sensitive poly(propylene sulfide)-poly(ethylene glycol) (PPS-PEG) amphiphilic polymers and hydrophobic Au nanoparticles (Au NPs) linked to X-ray-labile linoleic

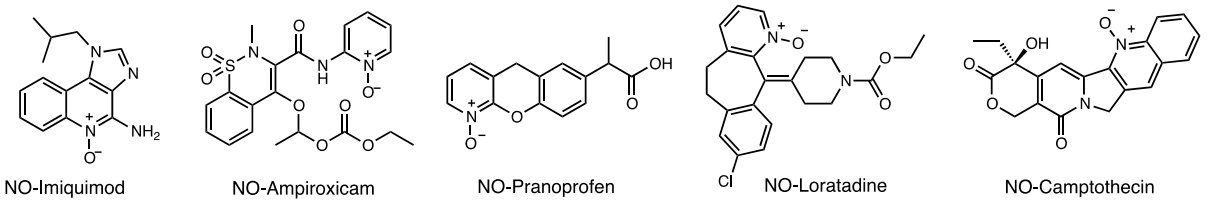


Figure 3. *N*-Oxide prodrugs for radiolytic activation.

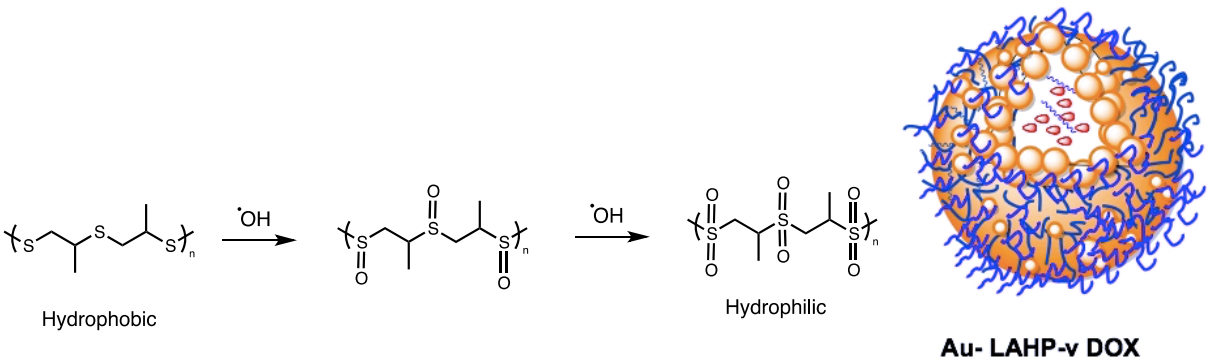


Figure 4. Nanoparticles, made by the co-assembly of the hydrophobic Au–LAHP NPs, amphiphilic oxidation-sensitive PPS–PEG polymers and hydrophilic DOX were activated by in situ-formed hydroxyl radicals (•OH). The oxidation of the thioether chain to sulfoxide and sulphone transformed the hydrophobic chain to hydrophilic, leading to vesicle degradation and the release of the drug.

acid hydroperoxide (LAHP) (Figure 4) [79]. The nanomaterial exhibited burst release of encapsulated doxorubicin (DOX) by irradiation at 8 Gy, up to 46.7%. Remarkably, drug release from Au–LAHP–vDOX incubated in H_2O_2 progressively increased to 76.7% 24 h after irradiation, while late drug release in PBS was minimal. In contrast, drug release from Au–LA–vDOX showed little or no response to X-ray irradiation. Irradiation (8 Gy) in a mouse subcutaneous tumor model significantly retarded tumor growth in the presence of Au–LA–vDOX and inhibited tumor growth even after irradiation, which could be attributed to the slow release of DOX from vesicles.

In an analogous study, radiation-generated ROS were the trigger signal for the activation of doxorubicin (DOX)-loaded nanomicelles derived from poly(propylene sulfide) (PPS) and hyaluronic acid (HA) [57]. HA–PPS@DOX nanoparticles exhibited prolonged circulation times of up to 24 h and showed decent accumulation in tumors. When oxidation by X-ray irradiation (8 Gy) converted the lipophilic thioether chain to polar sulfoxides, the ROS-sensitive HA–PPS@DOX NMs disintegrated and released the cytotoxic cargo.

5.2. *Fragmentation of disulfide and diselenide bonds*

Disulfide and diselenide are not particularly sensitive to direct X-rays without the synergistic effect of a high concentration of ROS. The Tanabe group devised a DNA amphiphile (DAM) composed of hydrophilic oligodeoxynucleotides linked to hydrophobic alkyl chains by disulfide bonds [80]. When DAM molecules are exposed to X-ray irradiation, selective fragmentation of the disulfide bonds leads to dissociation of the aggregates and release of the drug.

Radiolytic oxidation (5 Gy) resulted in the cleavage of the cross-linked polysulfide shell of the cysteine-decorated poly-G4.5 PAMAM dendrimer (a “4th generation” branched polyamidoamine

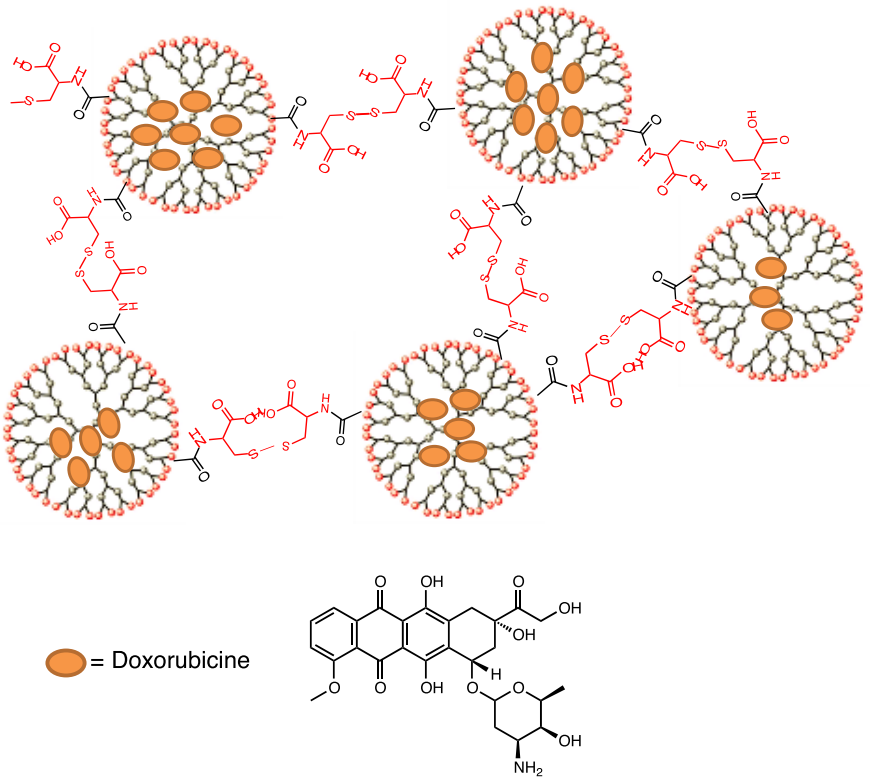


Figure 5. Cross-linked polysulfide shell of cysteine-decorated poly-G4.5 PAMAM dendrimer.

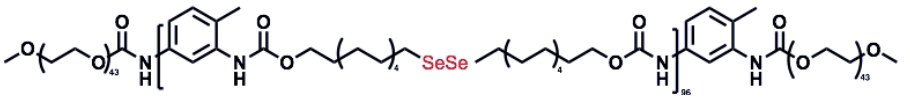


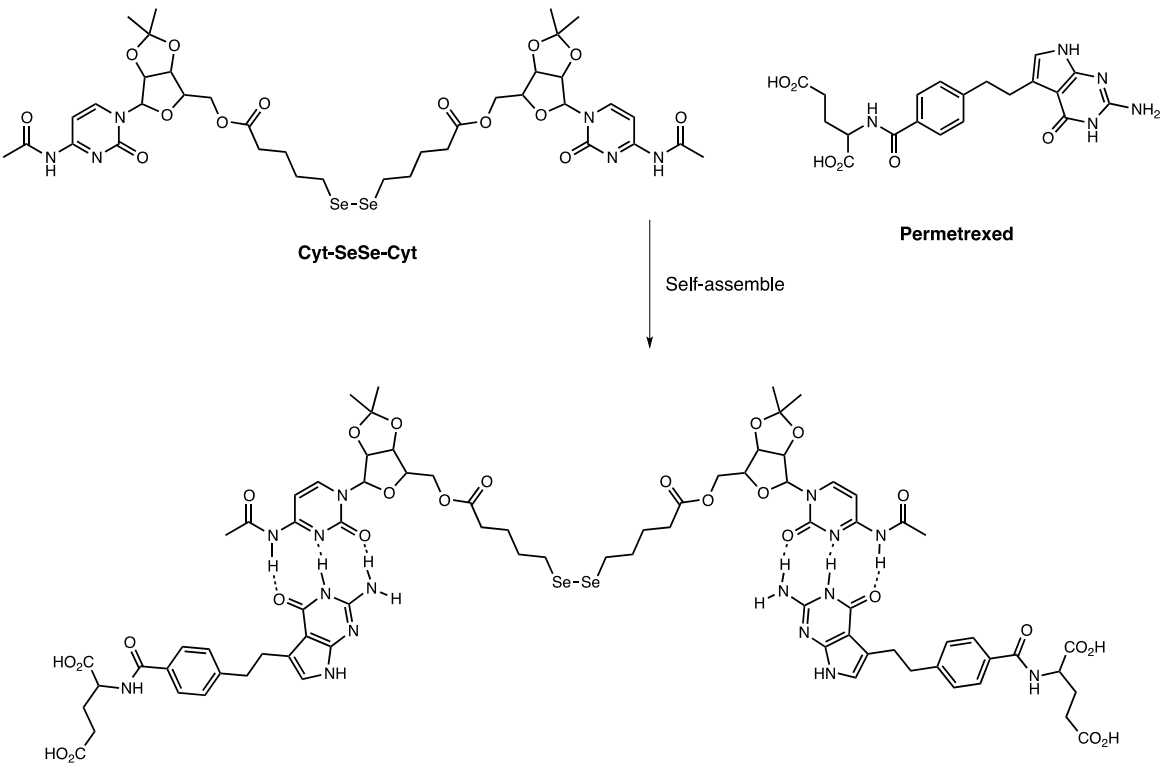
Figure 6. Diselenide-linked micellar aggregates were cleaved in the presence of ROS, generated by γ-rays (5 Gy).

polymer containing 128 surface carboxylate groups), allowing release of the anticancer drug (Figure 5) [81].

Ma *et al.* used PEG–PUSeSe–PEG, a radiation-sensitive diblock copolymer, to construct a nanoscale drug delivery system (Figure 6) [82]. On exposure to γ-rays (5 Gy), the generated ROS oxidized the diselenide groups by forming selenic acid and triggering the fragmentation of the PEG–PUSeSe–PEG micellar aggregates/release of the encapsulated drug (PUSeSe stands for polyurethane diselenide hydrophobic block). Unfortunately the PEG–PUSeSe–PEG polymer showed cytotoxicity to HepG2 cells.

Immunotherapy can be combined with radiotherapy and chemotherapy, providing a new strategy for cancer treatment. A combination of diselenide and pemetrexed has been developed to combine NK cell-based cancer immunotherapy with radiotherapy and chemotherapy in a single system [83]. The assemblies were prepared by co-assembly between pemetrexed and cytosine-containing diselenide via hydrogen bonds (Scheme 9). ROS generated by γ-radiation (5 Gy) not only cleaved the diselenide bonds (converted to seleninic acid), thereby suppressing human leukocyte antigen E (HLA-E) expression in cancer cells and activating the immune response of NK cells, but also released pemetrexed.

The Chen group has developed a strategy for the targeted delivery of doxorubicin (DOX) to tumor sites using mesoporous organosilica (MON) nanoparticles coated with a cancer cell membrane, which



Scheme 9. Assembly of the diselenide–pemetrexed block.

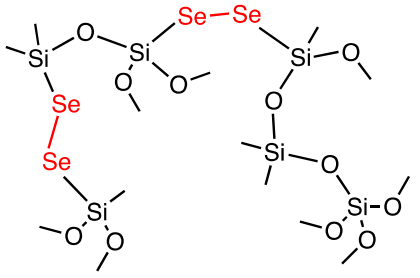


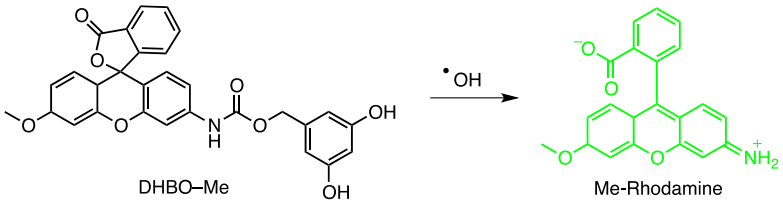
Figure 7. ROS-sensitized mesoporous organosilica nanoparticles (MONs).

can be activated by X-rays (Figure 7) [84]. This approach involved loading DOX and a PD-L1 inhibitor into the pores of MONs bridged by diselenic linkers. MONs were then coated with cancer cell-derived membrane fragments to produce anti-PD-L1 CM@MON@DOX +. Low-dose X-ray irradiation (less than 1 Gy) resulted in the fragmentation of the diselenic bond, disrupting the structure of the MONs, and releasing DOX and the PD-L1 inhibitor, resulting in DOX-mediated immunogenic cell death at the tumor site [85].

Diselenide-derived block copolymers have been used to load doxorubicin for precision chemotherapy [86]. ROS generated by X-ray irradiation (2 Gy) promoted release of the encapsulated drug under in vivo conditions in tumor-bearing mice, which produced a more pronounced overall antitumor effect than the other test groups.

5.3. *Dihydroxybenzyl functions sensitive to oxidation*

The Co-60 source has been tested in the radiative fragmentation of several carbamate-bound photosensitive compounds [87]. Best fragmentation kinetics were observed with resorcinol derivatives. The hypothesis of radiation-induced *ortho*- or *para*-hydroxylation of the aryl ring was put forward as the



Scheme 10. Fragmentation of rhodamine DHBO–Me to Me-Rhodamine (green) by γ-ray irradiation.

operative fragmentation mechanism, followed by a 1,4 or 1,6 elimination reaction releasing the covalently attached ligands (Scheme 10). Radiation activation has been studied in cell experiments as well as in tumor-bearing mice, using the ligand MMAE as an anticancer agent with a single-dose activation of 4 Gy. Details of oncological relevance were not provided, however.

5.4. *Release of small molecules*

5.4.1. *NO synthesis triggered by ROS*

The mixing of hafnium tetrachloride ($HfCl_4$) and 2-nitroimidazol in a microemulsion containing hexamethylene, hexyl alcohol, and a surfactant (triton) resulted in the formation of a unique type of nanoscale coordination polymer, HNP NPs [88]. NPs were coated with DOPA (sodium salt of 1,2-dioleoyl-sn-glycero-3-phosphate) and then mixed with 1,2-dihexadecanoyl-sn-glycero-3-phosphocholine (DPPC), cholesterol and 1,2-distearoyl-sn-glycero-3-phosphoethanolamine-N-[methoxy(poly(ethylene glycol))-5000] (DSPE–PEG5000). The nanomaterial was shown to act both as a radiosensitizer and as an immune adjuvant for NO donors (2-nIm), which released NO in a controlled manner under X-ray irradiation.

A template-free method was devised for the synthesis of Bi–SH nanoparticles (Bi–SH NPs), which were then coated with SiO_2 to improve aqueous solubility, and finally functionalized with the NO precursor S-nitrosothiol (Bi–SNO NPs) [89]. The particular advantage of this system is the local release of NO: the short half-life of NO and its high sensitivity to biological substances considerably limit the development of NO-based therapeutic platforms for clinical application. The release of NO by X-rays (5 Gy) has been validated in rodents.

5.4.2. *Carbon monoxide release triggered by X-rays*

A multifunctional X-ray-activatable theranostic nanoplatform has been developed based on Ce-doped $LiLuF_4$ scintillant and UV-sensitive $Mn_2(CO)_{10}$ for the controlled generation of CO and manganese dioxide (MnO_2) in tumors [90]. Under X-ray irradiation, radioluminescent UV light from SCNPs triggered the photolysis of $Mn_2(CO)_{10}$ by luminescent resonance energy transfer (LRET). In this process, the Mn–Mn and Mn–CO bonds of $Mn_2(CO)_{10}$ were cleaved photochemically to release CO forming MnO_2 as a by-product. In vitro and in vivo experiments demonstrated the potent antitumor activity and low systemic toxicity of this nanomedicine.

5.4.3. *Peroxynitrite generation*

Peroxynitrite ($ONOO^{\bullet-}$), a reaction product derived from nitric oxide (NO) and superoxide ($O_2^{\bullet-}$), is a powerful oxidizing and nitrating agent that modulates complex biological processes and promotes cell death. A multifunctional X-ray-controlled $ONOO^{\bullet-}$ generation platform based on scintillating nanoparticles (SCNPs) and UV-sensitive NO donors, Roussin black salt, has been reported [91]. In this system, Ce-doped $LiLuF_4$ acts as a radiosensitizer by promoting ROS, including $O_2^{\bullet-}$, under X-ray irradiation, and it also converts X-rays to UV light to activate Roussin's black salt to release NO. The simultaneous release of NO and $O_2^{\bullet-}$ ensures efficient X-ray-controlled production of $ONOO^{\bullet-}$ in tumors.

6. Miscellaneous

Doxorubicin (DOX), gold nanoparticles (AuNPs), DNA-coated AuNPs (DNA–AuNPs), and DOX conjugated to DNA–AuNPs (DOX–DNA–AuNPs) were

synthesized and incubated with MCF-7 breast cancer cells [92]. Strong internalization was observed with only residual leaching of immobilized doxorubicin. The extinction efficiency of DNA strand-attached DOX on AuNPs was estimated at 99.7%. Under the influence of X-rays, the DOX–DNA–AuNP DNA strands were cleaved, probably by the presence of $^{\bullet}$OH radicals, and the DOX molecules were removed from the AuNP surface. The drug-carrying nanoparticles became more toxic as the X-ray dose increased, from $-1\% \pm 15\%$ at 2 Gy to $+19\% \pm 13\%$ at 5 Gy and $+33\% \pm 21\%$ at 10 Gy. Noteworthily the observed toxicity was greater than the sum of the radiation toxicity and the toxicity of the drug-carrying nanoparticles. Given the amount of AuNPs in the cell (≈30 wt%), the small increase in toxicity observed seems to suggest that these nanoparticles remained in the cytoplasm, as this amount of AuNPs in the cell would cause significant damage only if the increase in X-ray absorption was taken into account. It should be noted that this report is the first demonstration of increased cytotoxicity due to X-ray-triggered release of chemotherapeutic drugs from nanoparticles in cells.

ROS can oxidize unsaturated lipids and destabilize liposomal membranes, releasing the encapsulated cargo. The Deng laboratory devised a drug delivery system using sensitized liposomes containing 3–5 nm gold nanoparticles (GNPs) and the photosensitizer verteporfin (VP), and X-rays to trigger liposome rupture [93]. This system takes advantage of the intensity of the radiation and the ability of liposomes to adapt to the environment. This system takes advantage of the radiation intensification effect of GNPs to amplify the production of singlet oxygen ($^1O_2$) and other ROS from verteporfin under X-ray irradiation, triggering thus the rupture of liposomes. The results indicate that liposomes loaded with GNPs and VP exhibit higher rates of $^1O_2$ production than those loaded with GNPs or VP individually. The antitumor efficacy of this drug delivery system was validated in vitro and in vivo, demonstrating that LipoDOX, a drug delivery system based on DOX-loaded liposomes, displayed superior tumor inhibition efficacy following X-ray irradiation.

The liposome was prepared from dioleoylglycerophosphocholine (DOPC) and dioctadecenoyltrimethylammonium-propane (DOTAP) [93]. Cellular uptake of the nanomaterial was facilitated by its positive charge at neutral pH. In vitro gene inhibition by X-ray activation of the liposome was demonstrated using an antisense oligonucleotide for a pituitary receptor specific for adenylate cyclase-activating polypeptide (PACAP) in rats. The system was also tested in X-ray-triggered chemotherapy, using doxorubicin in liposomes on human HCT 116 colorectal cancer cells.

It has been shown that low-dose X-rays (1–4 Gy at 6 MeV) can trigger the release of doxorubicin (DOX) from liposomes co-loaded with gold nanoparticles and photosensitizers [94]. ROS produced by gold nanoparticles and photosensitizers under X-ray irradiation destabilized the lipid membrane and led to the release of encapsulated DOX.

In a similar system, the anticancer drug paclitaxel (PTX) and radioluminescent $CaWO_4$ (CWO) nanoparticles were coencapsulated with poly(ethylene glycol)-poly(lactic acid) (PEG–PLA) block copolymers ("PEG–PLA/CWO/PTX NPs") [94]. Upon exposure to X-rays, the PEG–PLA coating layer degraded and PTX molecules were released. It should be noted that radioluminescent CWO nanoparticles emit UV-A/blue light (350–525 nm) under X-ray excitation, which also produced radioenhancement/radio-sensitization effects in tumor tissues.

GFP reporters were attached to silica-coated gold nanoparticles (AuNP@ SiO_2) by DNA strands as linkers [95]. The DNA strand is broken by reaction with ROS, releasing the GFP reporters. Although proteins also react with ROS at high rates, the low concentrations of ROS rendered this damage (less than one reaction per protein) undetectable in this work. The authors hypothesized that the efficiency of AuNP@SiO_2 release could potentially be improved if the thickness of the SiO_2 layer was reduced and DNA strand breaks could be chemically enhanced. This work has shown that radiation activation can be compatible with the delivery/activation of therapeutic proteins in biological systems.

Alkyne-modified dextran hydrogel (500 kDa) was irradiated with γ-rays (14.4 kGy), resulting in densification of cross-linking by the formation of reactive terminal alkyne radicals, generated directly by irradiation or indirectly by reaction products from radiolysis of water [96]. Polymerization controlled hydrogel rigidity, contraction, release, and fusion. Hydrogel stiffness increased twofold for hydrogels swollen with

H_2O and threefold for hydrogels swollen with phosphate buffer after γ irradiation, due to the increase in the cross-linking density.

7. Conclusion and outlooks

Tissue-penetrating X- and γ-rays have the potential to selectively reach specific tissues. Simultaneous release of organic compounds with the therapeutic beam presents a conceptually different approach compared to standard protocols that combine radiotherapy and chemotherapy. This approach may allow further reducing radiation doses and systemic cytotoxicity compared to current protocols. Since organic compounds are transparent at this wavelength, and direct irradiation is inefficient, the search for different transduction mechanisms has led to many ingenious drug delivery strategies. To overcome the difficulties associated with the biocompatibility, degradability and cellular toxicity of several nanocarriers, strategic solutions are still needed. One of the main challenges is to minimize the X-ray doses required to trigger cargo release close to unity. Even though the release efficiency is one of the most important parameters, in most reports the quantitative analysis of the drug release is neglected claiming technical difficulties. This should be certainly improved in the near future.

The X-ray activation method is likely to be applied clinically in the near future, provided that issues of toxicity (of nanomaterials) and efficacy (enabling the necessary activation doses to be reduced) are optimized. Also, the nanoparticle formulation can serve as a simultaneous multifunctional platform for chemo-, thermo-and immuno-radiotherapy; it can not only enhance tumor irradiation responses during RT but also leverage combination therapy to achieve a synergistic effect. We believe that this mode of activation is still in its infancy, while opening up the potential of nuclear medicine to the control of smart materials. Future developments can be extended to proton and carbon ion beams, although they have a very narrow Bragg peak, meaning that they do not scatter along the radiation path and reach the biological region without energy loss [97]. This young field is developing rapidly and, with access to increasingly powerful and safe experimental methods and equipment, promises a breakthrough in research and clinical applications.

Declaration of interests

The authors do not work for, advise, own shares in, or receive funds from any organization that could benefit from this article, and have declared no affiliations other than their research organizations.

References

[1] Y. Pan, W. Tang, W. Fan, J. Zhang, X. Chen, *Chem. Soc. Rev.*, 2022, **51**, 9759-9830.

[2] M. Dizdaroglu, *Mutat. Res./Rev. Mutat. Res.*, 2015, **763**, 212-245.

[3] G. D. Wang, H. T. Nguyen, H. Chen, P. B. Cox, L. Wang, K. Nagata, Z. Hao, A. Wang, Z. Li, J. Xie, *Theranostics*, 2016, **6**, 2295-2305.

[4] W. Cao, Y. Gu, M. Meineck, H. Xu, *Chem. Asian J.*, 2014, **9**, 48-57.

[5] D. F. O'brien, K. A. Mcgovern, B. Bondurant, R. M. Sutherland, "Radiation sensitive liposomes", 2002, WO2001039744A3, https://patents.google.com/patent/WO2001039744A3.

[6] D. E. Hallahan, "X-ray guided drug delivery", 2000, US6159443A, https://patents.google.com/patent/US6159443A/en.

[7] D. Liu, F. Yang, F. Xiong, N. Gu, *Theranostics*, 2016, **6**, 1306-1323.

[8] S. M. Mirvakili, R. Langer, *Nat. Electron.*, 2021, **4**, 464-477.

[9] M. J. Hansen, W. A. Velema, M. M. Lerch, W. Szymanski, B. L. Feringa, *Chem. Soc. Rev.*, 2015, **44**, 3358-3377.

[10] K. Glusac, *Nat. Chem.*, 2016, **8**, 734-735.

[11] L. Beauté, N. McClenaghan, S. Lecommandoux, *Adv. Drug Deliv. Rev.*, 2019, **138**, 148-166.

[12] G. Song, L. Cheng, Y. Chao, K. Yang, Z. Liu, *Adv. Mater.*, 2017, **29**, article no. 1700996.

[13] H. Liu, J. Zhao, Y. Xue, J. Zhang, H. Bai, S. Pan, B. Peng, L. Li, N. H. Voelcker, *Angew. Chem. Int. Ed.*, 2023, **62**, article no. e202306100.

[14] X. Li, H. Sun, Y. Lu, L. Xing, *Med*, 2022, **3**, 600-602.

[15] W. Fan, N. Lu, Z. Shen, W. Tang, B. Shen, Z. Cui, L. Shan, Z. Yang, Z. Wang, O. Jacobson, Z. Zhou, Y. Liu, P. Hu, W. Yang, J. Song, Y. Zhang, L. Zhang, N. M. Khashab, M. A. Aronova, G. Lu, X. Chen, *Nat. Commun.*, 2019, **10**, article no. 1241.

[16] S.-H. Lim, C.-H. Li, Y.-I. Jeong, W.-Y. Jang, J.-M. Choi, S. Jung, *Int. J. Nanomed.*, 2019, **14**, 8861-8874.

[17] J. Chen, H. Dong, L. Bai, L. Li, S. Chen, X. Tian, Y. Pan, *J. Mater. Chem. B*, 2022, **10**, 1328-1342.

[18] X. Chen, J. Song, X. Chen, H. Yang, *Chem. Soc. Rev.*, 2019, **48**, 3073-3101.

[19] W. Fan, W. Tang, J. Lau, Z. Shen, J. Xie, J. Shi, X. Chen, *Adv. Mater.*, 2019, **31**, article no. 1806381.

[20] N. Lee, S. H. Choi, T. Hyeon, *Adv. Mater.*, 2013, **25**, 2641-2660.

[21] H. Lusic, M. W. Grinstaff, *Chem. Rev.*, 2013, **113**, 1641-1666.

[22] L. V. Spencer, U. Fano, *Phys. Rev.*, 1951, **81**, 464-466.

[23] C. A. Perez, M. Bauer, S. Edelstein, B. W. Gillespie, R. Birch, *Int. J. Radiat. Oncol. Biol. Phys.*, 1986, **12**, 539-547.

[24] R. J. Tuieng, S. H. Cartmell, C. C. Kirwan, M. J. Sherratt, *Cells*, 2021, **10**, article no. 3041.

[25] K. G. Desai, H. J. Park, *Drug Deliv.*, 2006, **13**, 39-50.
[26] F. Sakar, A. Y. Özer, S. Erdogan, M. Ekizoglu, D. Kart, M. Özalp, S. Colak, Y. Zencir, *Pharmaceut. Dev. Technol.*, 2017, **22**, 775-784.
[27] N. Kotagiri, G. P. Sudlow, W. J. Akers, S. Achilefu, *Nat. Nanotechnol.*, 2015, **10**, 370-379.
[28] A. Schori, C. Bömer, D. Borodin, S. P. Collins, B. Detlefs, M. Moretti Sala, S. Yudovich, S. Shwartz, *Phys. Rev. Lett.*, 2017, **119**, article no. 253902.
[29] K. Yang, Y. Yang, D. Sun, S. Li, X. Song, H. Yang, *Sci. China Mater.*, 2023, **66**, 4090-4099.
[30] Y. Dou, Y. Liu, F. Zhao, Y. Guo, X. Li, M. Wu, J. Chang, C. Yu, *Theranostics*, 2018, **8**, 5870-5889.
[31] L. He, X. Yu, W. Li, *ACS Nano*, 2022, **16**, 19691-19721.
[32] D. Kwatra, A. Venugopal, S. Anant, *Transl. Cancer Res.*, 2013, **2**, 330-342.
[33] S. Jain, D. G. Hirst, J. M. O'Sullivan, *Br. J. Radiol.*, 2012, **85**, 101-113.
[34] V. Apanasevich, V. Avramenko, P. Lukyanov, A. Lagureva, A. Polkovnikova, K. Lukyanenko, V. Kustov, V. Temchenko, I. Agafonova, I. Pankratov, L. Stebunov, S. Bratskaya, *Cancer Oncol. Res.*, 2014, **2**, 17-20.
[35] D. Wang, L. Niu, Z.-Y. Qiao, D.-B. Cheng, J. Wang, Y. Zhong, F. Bai, H. Wang, H. Fan, *ACS Nano*, 2018, **12**, 3796-3803.
[36] I. Unger, R. Seidel, S. Thürmer, M. N. Pohl, E. F. Aziz, L. S. Cederbaum, E. Muchová, P. Slavíček, B. Winter, N. V. Kryzhevoi, *Nat. Chem.*, 2017, **9**, 708-714.
[37] B. Abel, U. Buck, A. L. Sobolewski, W. Domcke, *Phys. Chem. Chem. Phys.*, 2011, **14**, 22-34.
[38] J. M. Herbert, M. P. Coons, *Annu. Rev. Phys. Chem.*, 2017, **68**, 447-472.
[39] O. Marsalek, F. Uhlig, J. VandeVondele, P. Jungwirth, *Acc. Chem. Res.*, 2012, **45**, 23-32.
[40] F. Uhlig, O. Marsalek, P. Jungwirth, *J. Phys. Chem. Lett.*, 2012, **3**, 3071-3075.
[41] M. Boero, M. Parrinello, K. Terakura, T. Ikeshoji, C. C. Liew, *Phys. Rev. Lett.*, 2003, **90**, article no. 226403.
[42] A. Kumar, J. A. Walker, D. M. Bartels, M. D. Sevilla, *J. Phys. Chem. A*, 2015, **119**, 9148-9159.
[43] J. R. McKone, N. S. Lewis, H. B. Gray, *Chem. Mater.*, 2014, **26**, 407-414.
[44] F. Ambrosio, G. Miceli, A. Pasquarello, *J. Phys. Chem. Lett.*, 2017, **8**, 2055-2059.
[45] T. A. Pham, Y. Ping, G. Galli, *Nat. Mater.*, 2017, **16**, 401-408.
[46] E. Alizadeh, L. Sanche, *Chem. Rev.*, 2012, **112**, 5578-5602.
[47] Y. Ding, Q. Pan, W. Gao, Y. Pu, K. Luo, B. He, *Biomater. Sci.*, 2023, **11**, 1182-1214.
[48] N. Goswami, Z. Luo, X. Yuan, D. T. Leong, J. Xie, *Mater. Horiz.*, 2017, **4**, 817-831.
[49] J. F. Hainfeld, D. N. Slatkin, H. M. Smilowitz, *Phys. Med. Biol.*, 2004, **49**, N309-N315.
[50] Y. Liu, P. Zhang, F. Li, X. Jin, J. Li, W. Chen, Q. Li, *Theranostics*, 2018, **8**, 1824-1849.
[51] K. T. Butterworth, J. A. Coulter, S. Jain, J. Forker, S. J. McMahon, G. Schettino, K. M. Prise, F. J. Currell, D. G. Hirst, *Nanotechnology*, 2010, **21**, article no. 295101.
[52] S. Her, D. A. Jaffray, C. Allen, *Adv. Drug Deliv. Rev.*, 2017, **109**, 84-101.
[53] N. Brown, P. Rocchi, L. Carmès, R. Guthier, M. Iyer, L. Seban, T. Morris, S. Bennett, M. Lavelle, J. Penailillo, R. Carrasco, C. Williams, E. Huynh, Z. Han, E. Kaza, T. Doussineau, S. M. Toprani, X. Qin, Z. D. Nagel, K. A. Sarosiek, A. Hagège, S. Dufort, G. Bort, F. Lux, O. Tillement, R. Berbeco, *Theranostics*, 2023, **13**, 4711-4729.
[54] S. H. Cho, *Phys. Med. Biol.*, 2005, **50**, N163-N173.
[55] D. R. Cooper, D. Bekah, J. L. Nadeau, *Front. Chem.*, 2014, **2**, article no. 86.
[56] J. F. Hainfeld, D. N. Slatkin, H. M. Smilowitz, *Phys. Med. Biol.*, 2004, **49**, article no. N309.
[57] Y. Yu, Z. Feng, J. Liu, X. Hou, X. Zhou, J. Gao, W. Wang, Y. Zhang, G. Li, J. Liu, *ACS Omega*, 2021, **6**, 19445-19457.
[58] Y. Liu, W. Chen, S. Wang, A. G. Joly, *Appl. Phys. Lett.*, 2008, **92**, article no. 043901.
[59] R. Weinstain, T. Slanina, D. Kand, P. Klán, *Chem. Rev.*, 2020, **120**, 13135-13272.
[60] G. Bort, T. Gallavardin, D. Ogden, P. I. Dalko, *Angew. Chem. Int, Ed.*, 2013, **52**, 4526-4537.
[61] M. Klausen, M. Blanchard-Desce, *J. Photochem. Photobiol. C: Photochem. Rev.*, 2021, **48**, article no. 100423.
[62] Y. Shibamoto, Y. Tachi, K. Tanabe, H. Hatta, S.-I. Nishimoto, *Int. J. Radiat. Oncol. Biol. Phys.*, 2004, **58**, 397-402.
[63] K. Tanabe, M. Sugiura, T. Ito, S. Nishimoto, *Bioorg. Med. Chem.*, 2012, **20**, 5164-5168.
[64] S. A. Everett, E. Swann, M. A. Naylor, M. R. L. Stratford, K. B. Patel, N. Tian, R. G. Newman, B. Vojnovic, C. J. Moody, P. Wardman, *Biochem. Pharmacol.*, 2002, **63**, 1629-1639.
[65] M. A. Naylor, E. Swann, S. A. Everett, M. Jaffar, J. Nolan, N. Robertson, S. D. Lockyer, K. B. Patel, M. F. Dennis, M. R. L. Stratford, P. Wardman, G. E. Adams, C. J. Moody, I. J. Stratford, *J. Med. Chem.*, 1998, **41**, 2720-2731.
[66] E. Swann, P. Barraja, A. M. Oberlander, W. T. Gardipee, A. R. Hudnott, H. D. Beall, C. J. Moody, *J. Med. Chem.*, 2001, **44**, 3311-3319.
[67] M. Hernick, R. F. Borch, *J. Med. Chem.*, 2003, **46**, 148-154.
[68] K. Tanabe, Y. Makimura, Y. Tachi, A. Imagawa-Sato, S. Nishimoto, *Bioorg. Med. Chem. Lett.*, 2005, **15**, 2321-2324.
[69] M. Petit, G. Bort, B.-T. Doan, C. Sicard, D. Ogden, D. Scherman, C. Ferroud, P. I. Dalko, *Angew. Chem. Int. Ed.*, 2011, **50**, 9708-9711.
[70] A. Guesdon-Vennerie, P. Couvreur, F. Ali, F. Pouzoulet, C. Roulin, I. Martínez-Rovira, G. Bernadat, F.-X. Legrand, C. Bourgaux, C. L. Mazars, S. Marco, S. Trépout, S. Mura, S. Mériaux, G. Bort, *Nat. Commun.*, 2022, **13**, article no. 4102.
[71] A. Barosi, P. Dunkel, E. Guénin, Y. Lalatonne, P. Zeitoun, I. Fitton, C. Journé, A. Bravin, A. Maruani, H. Dhimane, L. Motte, P. I. Dalko, *RSC Adv.*, 2020, **10**, 3366-3370.
[72] Q. Fu, Z. Gu, S. Shen, Y. Bai, X. Wang, M. Xu, P. Sun, J. Chen, D. Li, Z. Liu, *Nat. Chem.*, 2024, Published Online: 2024-04-01.
[73] L.-L. Liao, G.-M. Cao, J.-H. Ye, G.-Q. Sun, W.-J. Zhou, Y.-Y. Gui, S.-S. Yan, G. Shen, D.-G. Yu, *J. Am. Chem. Soc.*, 2018, **140**, 17338-17342.
[74] Z. Guo, H. Hong, Y. Zheng, Z. Wang, Z. Ding, Q. Fu, Z. Liu, *Angew. Chem. Int. Ed.*, 2022, **61**, article no. e202205014.
[75] J. Geng, Y. Zhang, Q. Gao, K. Neumann, H. Dong, H. Porter, M. Potter, H. Ren, D. Argyle, M. Bradley, *Nat. Chem.*, 2021, **13**, 805-810.
[76] Y. Yang, G. Deng, Y. Lu, Q. Liu, M. Abe, X. Zeng, *J. Phys. Chem. A*, 2019, **123**, 9311-9320.

[77] C. Yang, Y. Yang, Y. Li, Q. Ni, J. Li, *J. Am. Chem. Soc.*, 2023, **145**, 385-391.
[78] Z. Ding, Z. Guo, Y. Zheng, Z. Wang, Q. Fu, Z. Liu, *J. Am. Chem. Soc.*, 2022, **144**, 9458-9464.
[79] Z. Zhou, A. Chan, Z. Wang, X. Huang, G. Yu, O. Jacobson, S. Wang, Y. Liu, L. Shan, Y. Dai, Z. Shen, L. Lin, W. Chen, X. Chen, *Angew. Chem. Int. Ed.*, 2018, **57**, 8463-8467.
[80] K. Tanabe, T. Asada, T. Ito, S. Nishimoto, *Bioconjugate Chem.*, 2012, **23**, 1909-1914.
[81] S. Wu, H. Chou, C. Yuh, S. L. Mekuria, Y. Kao, H. Tsai, *Adv. Sci.*, 2018, **5**, article no. 1700339.
[82] N. Ma, H. Xu, L. An, J. Li, Z. Sun, X. Zhang, *Langmuir*, 2011, **27**, 5874-5878.
[83] T. Li, S. Pan, S. Gao, W. Xiang, C. Sun, W. Cao, H. Xu, *Angew. Chem. Int. Ed.*, 2020, **59**, 2700-2704.
[84] D. Shao, F. Zhang, F. Chen, X. Zheng, H. Hu, C. Yang, Z. Tu, Z. Wang, Z. Chang, J. Lu, T. Li, Y. Zhang, L. Chen, K. W. Leong, W. Dong, *Adv. Mater.*, 2020, **32**, article no. 2004385.
[85] D. V. Krysko, A. D. Garg, A. Kaczmarek, O. Krysko, P. Agostinis, P. Vandenabeele, *Nat. Rev. Cancer*, 2012, **12**, 860-875.
[86] L. Zhang, S. Zhang, J. Xu, Y. Li, J. He, Y. Yang, T. Huynh, P. Ni, G. Duan, Z. Yang, R. Zhou, *ACS Appl. Mater. Interfaces*, 2020, **12**, 43398-43407.
[87] Q. Fu, H. Li, D. Duan, C. Wang, S. Shen, H. Ma, Z. Liu, *Angew. Chem. Int. Ed.*, 2020, **59**, 21546-21552.
[88] N. Liu, J. Zhu, W. Zhu, L. Chen, M. Li, J. Shen, M. Chen, Y. Wu, F. Pan, Z. Deng, Y. Liu, G. Yang, Z. Liu, Q. Chen, Y. Yang, *Adv. Mater.*, 2023, **35**, article no. 2302220.
[89] F. Zhang, S. Liu, N. Zhang, Y. Kuang, W. Li, S. Gai, F. He, A. Gulzar, P. Yang, *Nanoscale*, 2020, **12**, 19293-19307.
[90] Z. Du, X. Wang, X. Zhang, Z. Gu, X. Fu, S. Gan, T. Fu, S. Xie, W. Tan, *Angew. Chem. Int. Ed.*, 2023, **62**, article no. e202302525.
[91] Z. Du, X. Zhang, Z. Guo, J. Xie, X. Dong, S. Zhu, J. Du, Z. Gu, Y. Zhao, *Adv. Mater.*, 2018, **30**, article no. 1804046.
[92] Z. B. Starkewolf, L. Miyachi, J. Wong, T. Guo, *Chem. Commun.*, 2013, **49**, 2545-2547.
[93] W. Deng, W. Chen, S. Clement, A. Guller, Z. Zhao, A. Engel, E. M. Goldys, *Nat. Commun.*, 2018, **9**, article no. 2713.
[94] K. Sarkar, S. E. Torregrossa-Allen, B. D. Elzey, S. Narayanan, M. P. Langer, G. A. Durm, Y.-Y. Won, *Mol. Pharm.*, 2022, **19**, 2776-2794.
[95] M. Su, K. G. Guggenheim, J. Lien, J. B. Siegel, T. Guo, *ACS Appl, Mater. Interfaces*, 2018, **10**, 31860-31864.
[96] T. G. Brevé, H. Liu, A. G. Denkova, R. Eelkema, *Macro Mater. Eng.*, 2022, **307**, article no. 2100623.
[97] Y. Lee, R. Okayasu, *Int. J. Part. Ther.*, 2018, **5**, 114-121.

Comptes Rendus

Chimie

Objectif de la revue

Les *Comptes Rendus Chimie* sont une revue électronique évaluée par les pairs de niveau international, qui couvre l'ensemble des domaines des sciences chimiques.

Ils publient des numéros thématiques, des articles originaux de recherche, des articles de synthèse, des mises au point, des mises en perspective historiques, des textes à visée pédagogique, ou encore des actes de colloque, en anglais ou en français, sans limite de longueur et dans un format aussi souple que possible (figures, données associées, etc.).

Depuis 2020, les *Comptes Rendus Chimie* sont publiés avec le centre Mersenne pour l'édition scientifique ouverte, selon une politique vertueuse de libre accès diamant, gratuit pour les auteurs (pas de frais de publication) comme pour les lecteurs (accès libre, immédiat et pérenne).

À propos de la revue

Les *Comptes Rendus Chimie* sont exclusivement publiés au format électronique.

Toutes les informations sur la revue, ainsi que le texte intégral de l'ensemble des articles, sont disponibles sur son site internet, à l'adresse https://comptes-rendus.academie-sciences.fr/chimie/.

Informations pour les auteurs

Pour toute question relative à la soumission d'un manuscrit, merci de consulter le site internet de la revue : https://comptes-rendus.academie-sciences.fr/chimie/.

Contact

Académie des sciences
23 quai de Conti
75006 Paris (France)
cr-chimie@academie-sciences.fr

Comptes Rendus
de l'Académie des sciences

Chimie

Volume 26, nº S3, 2023

Special issue / *Numéro spécial*

Breaking Barriers in Chemical Biology – Toulouse 2022 / *Briser les barrières entre la chimie et la biologie – Toulouse 2022*

Guest editors / *Rédacteurs en chef invités*

Marie Lopez (CNRS-Univ. Montpellier-ENSCM, IBMM, Montpellier, France), Elisabetta Mileo (Aix-Marseille Univ, CNRS, BIP, IMM, Marseille, France), Eric Defrancq (Univ. Grenoble-Alpes-CNRS, DCM, Grenoble, France), Agnes Delmas (CNRS, CBM, Orléans, France), Boris Vauzeilles (CNRS-Univ. Paris-Saclay, ICSN, Gif-sur-Yvette, France), Dominique Guianvarch (CNRS-Univ. Paris-Saclay, ICMMO, Orsay, France) and Christophe Biot (CNRS-Univ. Lille, UGSF, Lille, France)

Cover illustration / *Illustration de couverture*

When Artificial Intelligence Unites Chemistry and Biology: A Generated Illustration.

Contents / *Sommaire*

1878-1543 (electronic)

Académie des sciences — Paris

Comptes Rendus
de l'Académie des sciences

Chimie

Volume 26, nº S3, 2023