

Des protéines de synthèse taillées sur mesure pour investiguer le vivant

Résumé La montée en puissance de la synthèse chimique de protéines alliée à la diversité moléculaire offerte par la chimie organique en fait un outil inédit pour décrypter des phénomènes biologiques et, au-delà, pour accéder à une large palette de propriétés. Cet article décrit un état des lieux de ce domaine illustré par des réalisations récentes d'équipes françaises ainsi qu'une mise en perspective discutant les prochains verrous à lever.

Depuis la synthèse des premiers dipeptides par Curtius et Fischer [1], les protéines ont constitué un objectif fascinant pour des générations de chimistes. Cinquante ans de développements méthodologiques ont été nécessaires pour aboutir en 1954 à la synthèse d'une hormone bioactive de neuf acides aminés, l'ocytocine [2] (figure 1). En 1971, la synthèse d'une enzyme, la ribonucléase A (124 acides aminés), a été réalisée par Merrifield [3] en enchaînant 248 réactions sur un support insoluble. Ce tour de force a démontré la grande efficacité de la synthèse de peptides en phase solide (SPPS), qui est aujourd'hui un outil classique pour la synthèse automatisée de peptides composés de quelques dizaines d'acides aminés. Le dernier jalon marquant de cette épopée scientifique est l'invention du concept de « ligation chimique native », pour lequel des réactions chimiques sélectives sont mises en œuvre pour coupler des segments non protégés obtenus par SPPS ou biotechnologie.

La « native chemical ligation » (NCL) [4-5], découverte il y a vingt-cinq ans, reste de loin la réaction la plus utilisée pour la synthèse de protéines [6]. Elle est basée sur une réaction de trans-thioestérification entre un résidu cystéine N-terminal d'un peptide (rouge) avec un thioester C-terminal d'un autre peptide (bleu), suivi d'un réarrangement (transfert d'acyle intramoléculaire) pour obtenir la liaison amide native (figure 2A). Elle a permis l'accès à des cibles de plus de 300 acides aminés [7] et a ouvert une voie inédite pour la synthèse parfaitement contrôlée de protéines natives, ou modifiées spécifiquement au niveau de leur squelette ou de certaines chaînes latérales. En ce sens, l'approche « chimique »

de la synthèse de protéines vient compléter les méthodes biotechnologiques exploitant des systèmes vivants, avec l'avantage de pouvoir créer des outils sur mesure pour décrypter le vivant à une résolution atomique.

Aujourd'hui, l'accès par synthèse totale à des protéines arborant diverses structures, linéaires, cycliques ou branchées, est facilité par une riche palette d'outils chimiques qui sont venus élargir la portée de la réaction de NCL. D'importants développements méthodologiques ont porté sur la mise au point de substituts du résidu peptidique thioester, dont la synthèse est restée longtemps problématique. Durant la dernière décennie, les équipes françaises ont été particulièrement actives dans la conception d'amides ayant la capacité de se réarranger spontanément dans l'eau et dans des conditions très douces en thioester transitoire, et donc de jouer le rôle de donneur d'acyle dans des réactions de NCL (figure 2B) [8-10]. Ce processus a pu être considérablement favorisé par catalyse intramoléculaire ou effet Thorpe-Ingold, pour aboutir, par un jeu de déplacements d'équilibres, à des réactions de NCL efficaces et rapides.

Des travaux récents portent également sur la conception de systèmes de capture qui permettent soit de faciliter l'assemblage des segments peptidiques (e.g. le système redox contrôlé SetCys [11], figure 2C), soit l'accès à des protéines ubiquitinées via un lien isopeptidique natif (e.g. l'auxiliaire de ligation Hdmb [12] greffé sur une lysine, figure 2C). Par ailleurs, des approches alternatives à la NCL ont été élaborées, exploitant la cycloaddition alcyne/azoture catalysée par le cuivre(II)

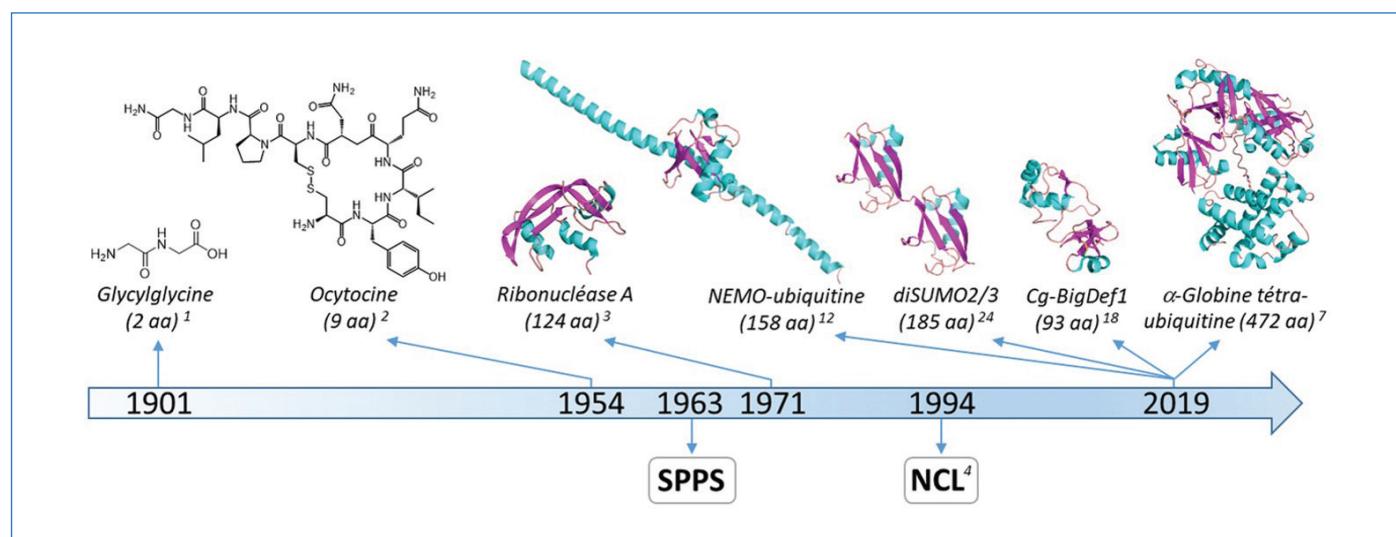
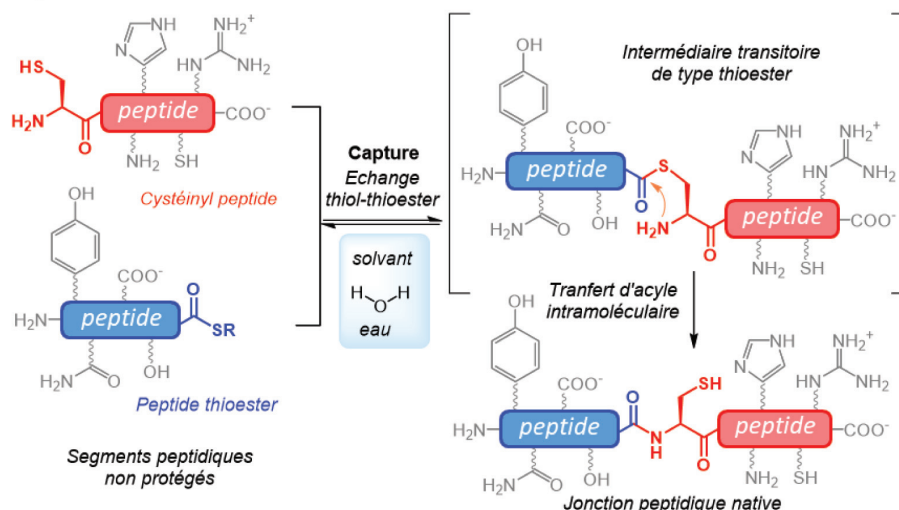
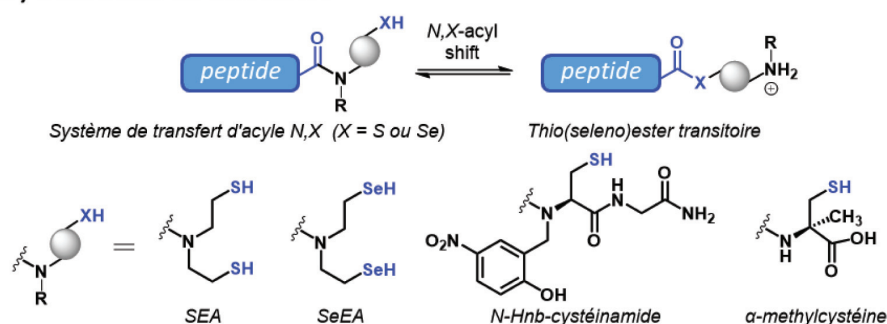


Figure 1 - La synthèse de peptides et de protéines a connu un essor considérable depuis le début du XX^e siècle, et notamment depuis 1994, avec la découverte de la NCL.

A) Principe de la ligation native



B) Substituts de thioesters



C) Systèmes de capture inspirés de la cystéine

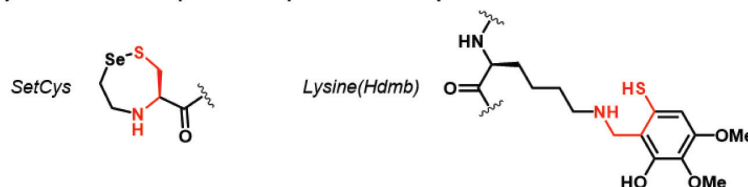


Figure 2 - A) Principe de la réaction de NCL montrant l'intermédiaire de capture thioester. B) Certains amides arborant sur l'azote un groupement β -mercaptoéthyle peuvent servir de donneur d'acyl *in situ* durant la NCL. C) Les systèmes de capture étendent la portée de la NCL en facilitant l'assemblage des segments peptidiques et l'accès à des structures cycliques ou branchées.

qui conduit à un 1,2,3-triazole, excellent mème de liaison amide [13-14]. Enfin, le développement de méthodes pour la synthèse supportée de protéines par assemblage itératif de segments peptidiques vient en complément des méthodes en solution [14-17]. Collectivement, ces techniques simplifient l'accès à des protéines complexes.

En combinant SPPS et techniques de ligation, les chimistes peuvent ainsi accéder à des protéines de structure native pour leur caractérisation biologique et structurale [18-21]. La synthèse d'une mini-protéine riche en ponts disulfure de la famille des « big défensines » (Cg-BigDef1, *figure 1*) a ainsi mis en évidence son activité bactéricide vis-à-vis de souches multi-résistantes [18]. L'étude a par ailleurs révélé le mode d'action original de ces composés, produits uniquement par des organismes invertébrés marins, et qui se caractérisent par la présence d'un domaine hydrophobe ancestral leur permettant d'être actifs en milieu salin. Les protéines arborant des modifications post-traductionnelles et difficilement accessibles par biotechnologie sont des cibles de choix pour les

chimistes (diSUMO2/3, NEMO-ubiquitine, *figure 1*) [12, 22-24]. Par exemple, la protéine NEMO joue un rôle central dans la réponse immunitaire innée. Son activation par liaison covalente de chaînes linéaires d'ubiquitine a été investiguée grâce à la synthèse d'un intermédiaire clé de ce processus, NEMO mono-ubiquitiné (*figure 1*). Cette étude a montré que l'ubiquitination linéaire de NEMO implique deux étapes distinctes d'amorçage et d'extension de chaîne, et a permis d'identifier l'enzyme responsable de l'extension dans les conditions physiologiques [12].

La synthèse chimique donne aussi l'opportunité d'élaborer des systèmes moléculaires non natifs présentant de nouvelles fonctionnalités [25-27]. La synthèse de dérivés multivalents du facteur de croissance des hépatocytes a conduit à la conception d'analogues présentant un potentiel thérapeutique dans le domaine de la régénération tissulaire [25]. Dans un autre domaine, l'ingénierie de protéines amyloïdes ouvre la voie, via la synthèse d'analogues solubles, à une connaissance accrue de leur implication dans la maladie d'Alzheimer (*figure 3*) [26]. Enfin, l'introduction d'acides aminés modifiés

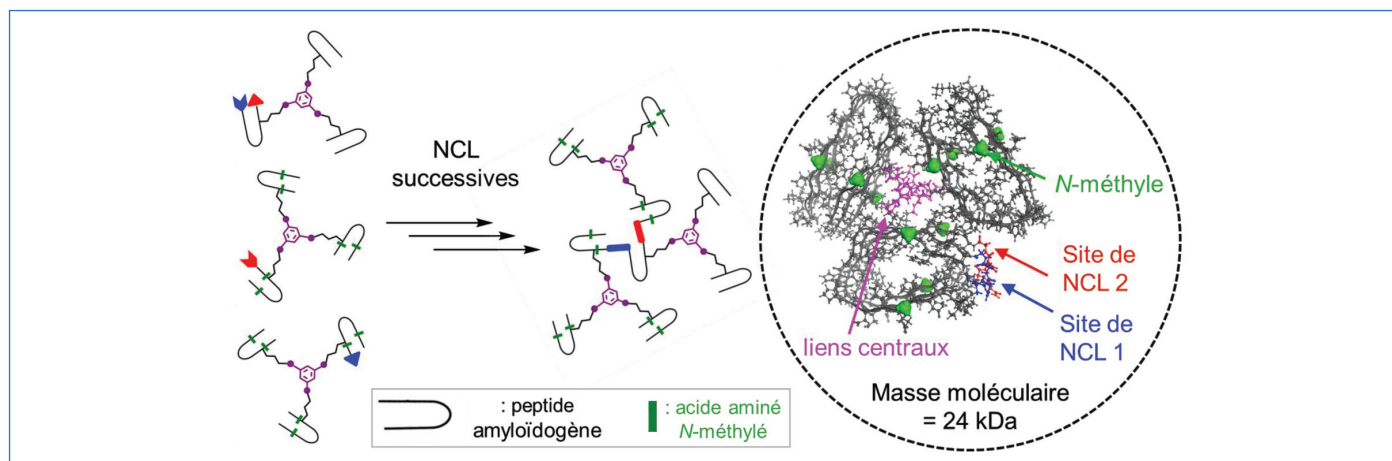


Figure 3 - Conception et synthèse d'un analogue soluble d'oligomère amyloïde. Neuf copies d'un bloc de construction peptidique amyloïdogène sont conjuguées via des liens centraux et des NCL successives. L'agrégation incontrôlée est empêchée par la *N*-méthylation d'acides aminés spécifiques.

dans une protéine intrinsèquement désordonnée a permis d'augmenter sa structuration et son affinité pour son partenaire [27], laissant espérer la conception de puissants inhibiteurs d'interactions protéine-protéine.

Les travaux discutés précédemment sont loin de représenter l'étendue des applications possibles de la synthèse totale des protéines. Potentiellement, toute la diversité fonctionnelle de la chimie organique peut être intégrée dans les protéines à des positions précises et ce, dans le but d'investiguer leurs fonctions ou de créer de nouveaux outils diagnostiques ou thérapeutiques. Il faut admettre que la synthèse totale de protéines de plus de 150 acides aminés nécessite la plupart du temps un réel investissement, même pour les spécialistes. Imaginer des méthodes d'assemblage plus rapides et efficaces constitue un défi et attire l'attention de nombreuses équipes dans le monde. Les équipes françaises participent naturellement à cette quête visant à rendre ces approches chimiques accessibles au plus grand nombre.

[1] E. Fischer, E. Fourneau, Ueber einige Derivate des Glykocolls, *Ber. Dtsch. Chem. Ges.*, **1901**, 34, p. 2868-877.
 [2] V. du Vigneaud, C. Ressler, J.M. Swan, W. Carleton, P.G. Katsoyannis, The synthesis of oxytocin, *J. Am. Chem. Soc.*, **1954**, 76, p. 3115-121.
 [3] B. Gutte, R.B. Merrifield, The synthesis of ribonuclease A, *J. Biol. Chem.*, **1971**, 246, p. 1922-941.
 [4] P.E. Dawson, T.W. Muir, I. Clark-Lewis, S.B. Kent, Synthesis of proteins by native chemical ligation, *Science*, **1994**, 266, p. 776-779.
 [5] V. Agouridas, O. El Mahdi, V. Diemer, M. Cargoët, J.-C.M. Monbaliu, O. Melnyk, Native chemical ligation and extended methods: mechanisms, catalysis, scope, and limitations, *Chem. Rev.*, **2019**, 119, p. 7328-443.
 [6] V. Agouridas, O. El Mahdi, M. Cargoët, O. Melnyk, A statistical view of protein chemical synthesis using NCL and extended methodologies, *Bioorg. Med. Chem.*, **2017**, 25, p. 4938-945.
 [7] H. Sun *et al.*, Diverse fate of ubiquitin chain moieties: the proximal is degraded with the target, and the distal protects the proximal from removal and recycles, *Proc. Natl. Acad. Sci. USA*, **2019**, 116, p. 7805-812.
 [8] N. Ollivier, J. Dheur, R. Mhidia, A. Blanpain, O. Melnyk, Bis(2-sulfanylethyl)amino native peptide ligation, *Org. Lett.*, **2010**, 12, p. 5238-241.
 [9] F. Burlina, G. Papageorgiou, C. Morris, P.D. White, J. Offer, In situ thioester formation for protein ligation using α -methylcysteine, *Chem. Sci.*, **2014**, 5, p. 766-770.
 [10] V.P. Terrier, H. Adihou, M. Arnould, A.F. Delmas, V. Aucagne, A straightforward method for automated Fmoc-based synthesis of bio-inspired peptide crypto-thioesters, *Chem. Sci.*, **2016**, 7, p. 339-345.
 [11] V. Diemer, N. Ollivier, B. Leclercq, H. Drobecq, J. Vicogne, V. Agouridas, O. Melnyk, A cysteine selenosulfide redox switch for protein chemical synthesis, *Nat. Commun.*, **2020**, 11, 2558.
 [12] F. Burlina *et al.*, Auxiliary-assisted chemical ubiquitylation of NEMO and linear extension by HOIP, *Commun. Chem.*, **2019**, 2, 111.
 [13] I.E. Valverde, F. Lecaille, G. Lalmanach, V. Aucagne, A.F. Delmas, Synthesis of a biologically active triazole-containing analogue of cystatin. A through successive peptidomimetic alkyne-azide ligations, *Angew. Chem. Int. Ed.*, **2012**, 51, p. 718-722.

[14] V. Aucagne, I.E. Valverde, P. Marceau, M. Galibert, N. Dendane, A.F. Delmas, Towards the simplification of protein synthesis: iterative solid-supported ligations with concomitant purifications, *Angew. Chem. Int. Ed.*, **2012**, 51, p. 11320-324.
 [15] L. Raibaut, H. Adihou, R. Desmet, A.F. Delmas, V. Aucagne, O. Melnyk, Highly efficient solid phase synthesis of large polypeptides by iterative ligations of bis(2-sulfanylethyl)amido (SEA) peptide segments, *Chem. Sci.*, **2013**, 4, p. 4061-066.
 [16] N. Ollivier *et al.*, A simple and traceless solid phase method simplifies the assembly of large peptides and the access to challenging proteins, *Chem. Sci.*, **2017**, 8, p. 5362-370.
 [17] S.A. Abboud, M. Amoura, B. Renoux, S. Papot, V. Pillier, V. Aucagne, Enzyme-cleavable linkers for protein chemical synthesis through solid-phase ligations, *Angew. Chem. Int. Ed.*, **2021**, 60, p. 18612-618.
 [18] K. Loth *et al.*, The ancestral N-terminal domain of big defensins drives bacterially triggered assembly into antimicrobial nanonets, *mBio*, **2019**, 10, e01821.
 [19] G. Martinez *et al.*, Spermaurin, an La1-like peptide from the venom of the scorpion *Scorpio Maurus Palmatus* improves sperm motility and fertilization in different mammalian species, *Mol. Hum. Reprod.*, **2016**, 23, p. 116-131.
 [20] C. Zoukian *et al.*, Synthesis by native chemical ligation and characterization of the scorpion toxin AmmTx3, *Bioorg. Med. Chem.*, **2019**, 27, p. 247-253.
 [21] J. Giribaldi *et al.*, Synthesis, structure and biological activity of CIA and CIB, Two α -conotoxins from the predation-evoked venom of *Conus catus*, *Toxins*, **2018**, 10, 222.
 [22] S. Sanulli *et al.*, Jarid2 methylation via the PRC2 complex regulates H3K27me3 deposition during cell differentiation, *Mol. Cell*, **2015**, 5, p. 769-783.
 [23] J. Bouchenna, M. Sénéchal, H. Drobecq, N. Stankovic-Valentin, J. Vicogne, O. Melnyk, The Role of the conserved SUMO-2/3 cysteine residue on domain structure investigated using protein chemical synthesis, *Bioconjugate Chem.*, **2019**, 30, p. 2684-696.
 [24] J. Bouchenna, M. Sénéchal, H. Drobecq, J. Vicogne, O. Melnyk, Total chemical synthesis of all SUMO-2/3 dimer combinations, *Bioconjugate Chem.*, **2019**, 30, p. 2967-973.
 [25] C. Simonneau *et al.*, Semi-synthesis of a HGF/SF kringle one (K1) domain scaffold generates a potent in vivo MET receptor agonist, *Chem. Sci.*, **2015**, 6, p. 2110-121.
 [26] R. Boehringer, B. Kieffer, V. Torbeev, Total chemical synthesis and biophysical properties of a designed soluble 24 kDa amyloid analogue, *Chem. Sci.*, **2018**, 9, p. 5594-599.
 [27] V. Bauer *et al.*, Conformational editing of intrinsically disordered protein by α -methylation, *Chem. Sci.*, **2021**, 12, p. 1080-089.

Vincent AUCAGNE^a, Fabienne BURLINA^b, Oleg MELNYK^c, directeurs de recherche CNRS, et Vladimir TORBEEV^d, professeur des universités.

^aCentre de Biophysique Moléculaire, UPR CNRS 4301, Orléans.

^bLaboratoire des BioMolécules, UMR 7203, Sorbonne Université, ENS, PSL University, CNRS, Paris.

^cUniversité de Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, U1019 - UMR 9017, Centre d'infection et d'immunité de Lille (CILL), Lille.

^dÉcole Supérieure de Biotechnologie de Strasbourg, UMR CNRS 7242 Biotechnologie et signalisation cellulaire, Université de Strasbourg.

* vincent.aucagne@cnrs-orleans.fr ; fabienne.burlina@sorbonne-universite.fr ; oleg.melnyk@ibl.cnrs.fr ; torbeev@unistra.fr