



HAL
open science

Pedal to the Metal: The Homogeneous Catalysis of the Native Chemical Ligation Reaction

Vincent Diemer, Olga Firstova, Vangelis Agouridas, Oleg Melnyk

► **To cite this version:**

Vincent Diemer, Olga Firstova, Vangelis Agouridas, Oleg Melnyk. Pedal to the Metal: The Homogeneous Catalysis of the Native Chemical Ligation Reaction. *Chemistry - A European Journal*, 2022, 28 (16), 10.1002/chem.202104229 . hal-03834897

HAL Id: hal-03834897

<https://hal.science/hal-03834897>

Submitted on 31 Oct 2022

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

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

Pedal to the metal: the homogeneous catalysis of the native chemical ligation reaction

Vincent Diemer,^[a] Olga Firstova,^[a] and Vangelis Agouridas,^{*[b]} Oleg Melnyk^{*[a]}

a) Dr. V. Diemer, O. Firstova, Dr. O. Melnyk
Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, U1019 - UMR 9017 - CIL - Center for Infection and Immunity of Lille, F-59000 Lille, France
E-mail: oleg.melnyk@ibl.cnrs.fr

[b] Dr. V. Agouridas
Centrale Lille, F-59000 Lille, France
E-mail: vangelis.agouridas@ibl.cnrs.fr

Abstract: The native chemical ligation reaction of peptide thioesters with cysteinyl peptides is a pivotal chemical process in the production of native or modified peptides and proteins, and well beyond in the preparation of various biomolecule analogs and materials. To benefit from this reaction at its fullest and to access all the possible applications, the experimentalist needs to know the factors affecting its rate and how to control it. This concept article presents the fundamental principles underlying the rate of the native chemical ligation and its homogeneous catalysis by nucleophiles. It has been prepared to serve as a quick guide in the search for an appropriate catalyst.

Introduction

By enabling the concatenation of unprotected peptide segments in water under mild conditions, the Native Chemical Ligation (NCL) reaction has become one of the most valuable transformations for accessing synthetic or semisynthetic proteins (Figure 1A).^[1-4] The core principle of this process consists in opposing a peptide thioester and a cysteinyl (Cys) peptide whose chemoselective reaction results in the formation of a native peptide bond between the two segments and thus, in the formation of a larger and well-defined polypeptide. Since the advent of NCL in 1994, its synthetic scope has been significantly enriched, providing protein chemists with an eclectic repertoire of practical methods for accessing a tremendous diversity of protein targets.^{[5][6]} This gain in operational synthetic capacity is reflected in the advances made in numerous domains where NCL is used as a routine tool, including the fields of chemical biology and biomedical research (Figure 1B).^[7]

Considered as a pivotal transformation for the production of proteins, NCL has seen its mechanism come under scrutiny very shortly after its discovery.^[8] The apparent straightforward equation represented in Figure 1A in reality conceals a two-step process, which starts with a thiol-thioester exchange leading to a capture intermediate (Figure 2A, Step 1). Then, the latter rearranges through an intramolecular *S*-to-*N* acyl group transfer to yield the product with a native peptide bond to cysteine (Figure 2A, Step 2). In standard conditions, experimental evidence suggests that the reaction is kinetically governed by step 1 which proceeds through tetrahedral intermediate **1** (Figure 2A). Logically, the rate of step 1 is under the dependence of the reactivity of the thioester and Cys thiol groups. While the reactivity of the Cys thiol is mainly determined by the fraction of Cys thiolate species present at the working pH,^[9] the reactivity of the thioester partner largely depends on two factors: the nature of the C-terminal residue and the type of thiol leaving group installed (Figure 2B). Thanks to their straightforward synthetic accessibility and good stability, peptide alkyl thioesters remain the most appreciated acylating source used for protein semi- or total synthesis.^[2, 6] However, practically useful ligation rates cannot usually be attained with such thioesters at millimolar

concentration due to their tempered reactivity. As a consequence, the NCL with peptide alkyl thioesters needs to be promoted most of the time by addition of catalysts. The aim of this concept article is to examine the fundamental principles underlying the homogeneous catalysis of the NCL reaction and to guide the reader in the search for an appropriate catalyst.

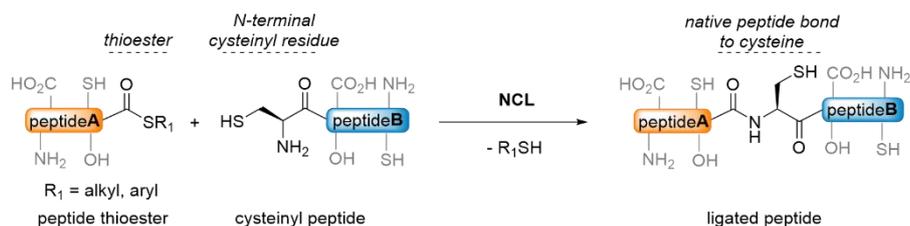
The guiding principles of NCL catalysis

Among the several modes of catalytic activation, covalent catalysis by nucleophiles^[10] provides the most significant rate enhancements for NCL. This approach consists in using a nucleophile additive (Nu) intended to react with the thioester partner for generating an activated acyl-Nu intermediate *in situ* (Figure 3A). During the reaction, the starting thioester and the acyl-Nu intermediate coexist in solution. Although the latter is usually formed in tiny amounts, its enhanced acylating capability makes it kinetically the main acylating species in the process. To achieve this goal, the nucleophile must be chosen carefully according to the nature of the starting thioester so that the energy barrier going from reactants to tetrahedral intermediate **3** is lower than that leading to tetrahedral intermediate **1** (Figure 3B). In practice, nucleophilic catalysis of NCL is usually carried out by the addition of various sulfur-, selenium-, nitrogen- or phosphorus-derived nucleophiles, which all present some benefits and drawbacks reviewed hereafter (Figure 3C).

Thiolates, a type of sulfur-based nucleophiles, have rapidly emerged as successful additives for the activation of peptide thioesters^[11] and are still the most widely deployed solution for enhancing NCL's rate today (Figure 4A). Their efficiency as catalysts can be rationalized according to the pioneering work of Jencks and coworkers, which allows to qualitatively appreciate the extent of activation of a thioester upon thiol-thioester exchanges.^[11] Indeed, the acylating power of the activated thioester, and thus the rate of NCL, largely depends on the acidity of the thiol catalyst R_2SH relative to that of the departing thiol R_1SH (Figure 4A). The use of a thiol additive R_2SH having a lower acid dissociation constant (pK_a) than that of R_1SH results in efficient catalysis because the thiolate R_2S^- is a better leaving group. However, in such a case, the formation of the activated thioester is a kinetically disfavored situation unless a very significant excess of catalyst is introduced in the reaction mixture as compensation.

A selection of thiols used either for thioester synthesis or NCL catalysis are listed by increasing acidity in Figure 4B (3-mercaptopropionic acid, MPA; sodium 2-mercaptoethyl sulfonate, MESNa^[12]; methyl thioglycolate, MTG^[13]; trifluoroethanethiol, TFET^[14]; 4-hydroxythiophenol^[15]; 3-mercaptophenyl acetic acid, MPAA^[15]; thiophenol^[8]). Despite the large diversity of commercially available thiols, just a handful of them are used as catalysts, one stringent selection criterion being their good to excellent water-solubility. Among these, MPAA has a pK_a well-below those of alkyl thiols

A) The Native Chemical Ligation (NCL) reaction



General considerations about NCL

- Highly **chemoselective** reaction, compatible with **unprotected peptide segments**
- Proceeds in **aqueous media**, at pH 6-8, at **millimolar** peptide concentration
- **Readily accessible peptide or protein thioesters** by chemical, biochemical or biological approaches
- Synthetic scope extended far beyond the use of classical peptide thioesters and cysteinyl peptides:
 - Sulfur containing amino acid surrogates
 - Selenocysteine and seleno amino acid surrogates
 - Selenoesters
 - Thio- and selenoester surrogates
 - ...

B) Aims & scope of NCL reaction

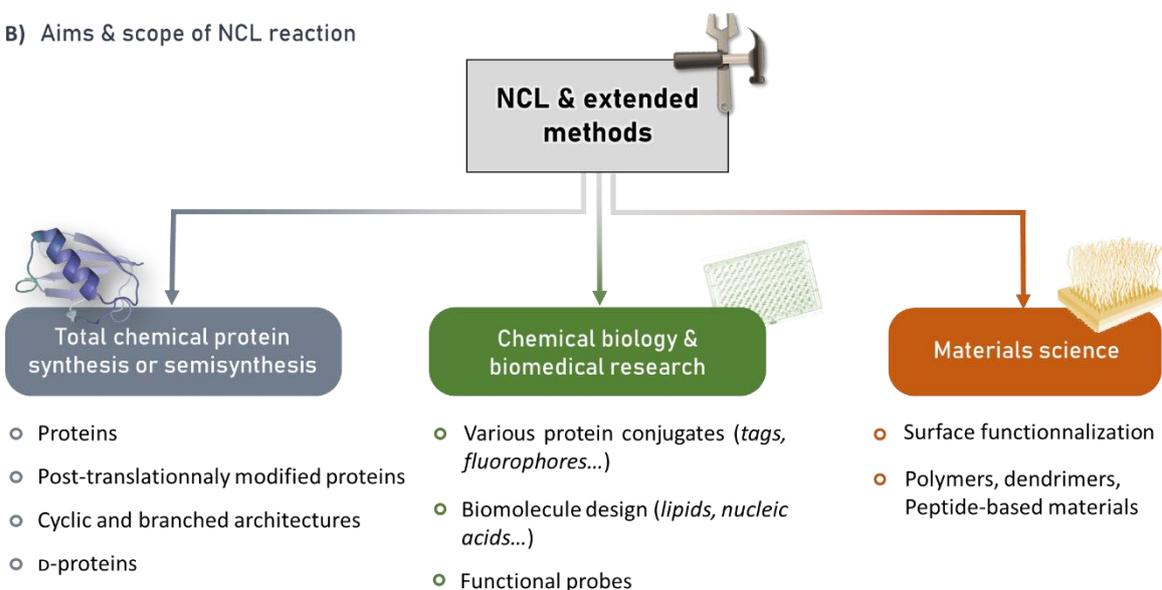


Figure 1. General considerations about the Native Chemical Ligation (NCL) reaction. A) Principle of the native chemical ligation reaction (functional groups on amino acid side chains are represented in gray on each peptide segment. For the sake of clarity, they will not be shown in the following figures). B) Most common applications of the NCL reaction. Various examples can be found in refs [3, 4, 7].

classically used for thioester synthesis, is odorless and highly water soluble at the working pH of NCL (≈ 7) due to its carboxylic acid moiety. Therefore, MPAA is by far the most widely applied catalyst in the field. In addition, it is important to note that thiol moieties harbored by the sidechain of internal Cys residues can in some cases offer a significant catalytic contribution owing to their mid-range pK_a .^[16]

The use of thiols, and more particularly of aryl thiols such as MPAA, can however prove problematic in some specific situations. This is especially the case when the NCL reaction needs to be combined with a subsequent Metal-Free Desulfurization (MFD) step in a one-pot process (Figure 5A).^[17] This popular approach was developed to extend the scope of NCL by converting the Cys junction residue to an alanine (Ala) residue, a significantly more represented amino acid in protein sequences. The mechanism of MFD is critically dependent on the formation of a Cys-derived thiyl radical intermediate, which must combine with the added phosphine to make the desulfurization process

successful. Aryl thiols are powerful thiyl radical scavengers and therefore kill the MFD reaction especially when used in large excess as discussed above. Consequently, their use is not compatible with one-pot NCL/MFD approaches unless the aryl thiol can be removed prior to adding the reagents for the MFD step. For example, the MPAA-derivative equipped with a hydrazide handle and designed by Brik and coll. was removed by using an aldehyde scavenger solid support (Figure 5B).^[18] Another option consists in using TFET as a catalyst, which can be removed by evaporation after ligation.^[14] Although being less powerful NCL catalysts than MPAA, alkyl thiols such as MTG^[13] are practicable MFD-compatible thiol additives. Imidazole and N-heterocyclic analogues thereof,^[19-21] i.e. sp^2 nitrogen-derived nucleophiles, provide an additional option in this regard. Indeed, they do not interfere with MFD reaction conditions and can recapitulate the catalytic properties of thiol additives through the formation of acyl imidazole type intermediates. They must however be introduced at much higher loadings to this end

because the equilibrated process depicted in Figure 5C favors the thioester species.

The last decade has witnessed the development of various thioester surrogates, which broaden the scope of classical

thioesters by permitting the design of novel peptide assembly strategies.^[22] The class of thioester precursors referred to as

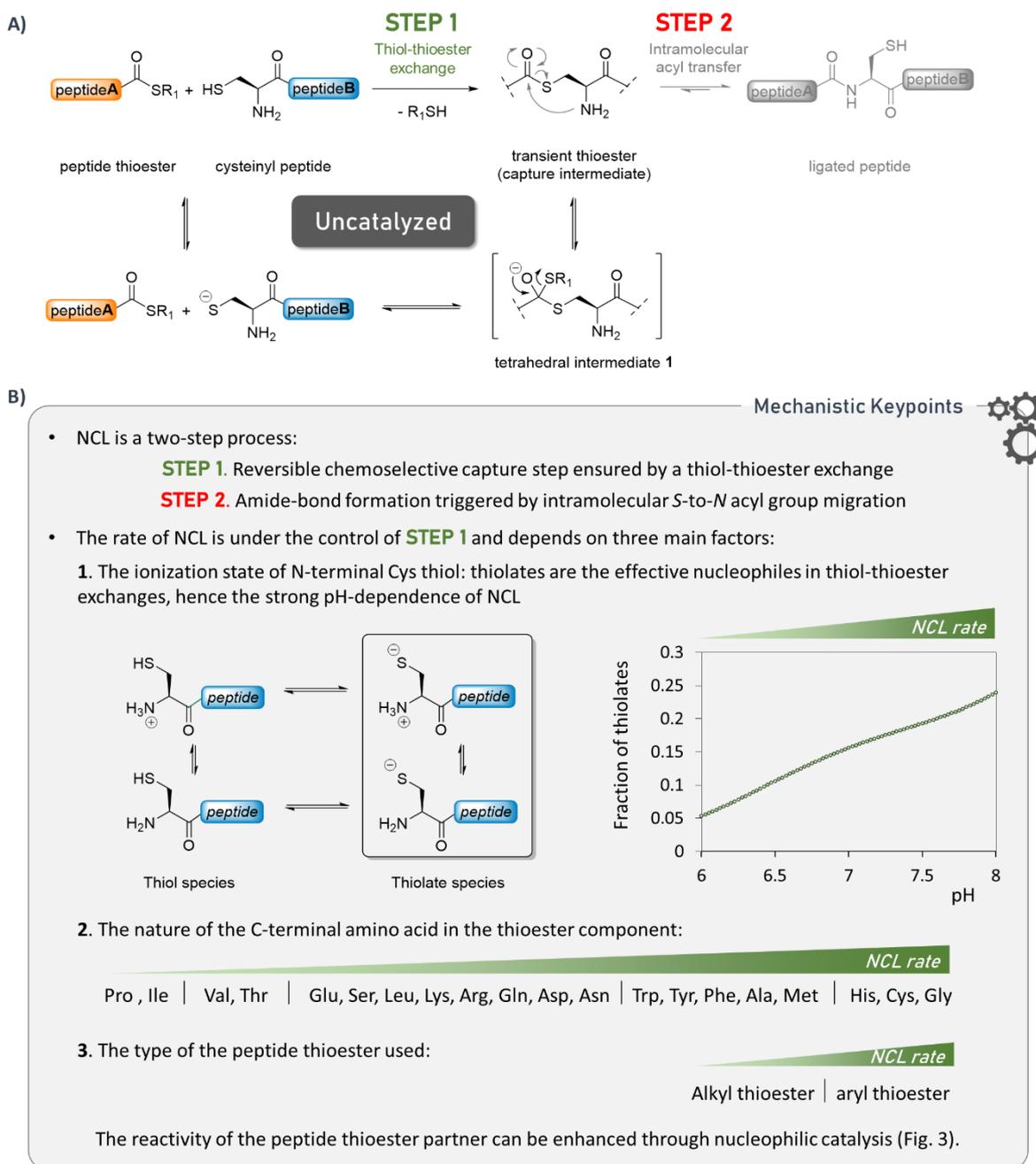


Figure 2. Insights into the mechanism of the uncatalyzed NCL reaction. A) NCL is a two-step process involving a thiol-thioester exchange for the capture step. B) NCL is usually under the kinetic control of the thiol-thioester exchange step and depends on three main factors: the ionization state of the N-terminal cysteine residue (the fraction of thiolate was calculated based on the model and the data in ref. ^[9]) and the environment as well as the nature of the thioester partner.

N,S-acyl shift systems is particularly popular (Figure 6).^[23, 24] Including the *bis*(2-sulfanylethyl)amido (SEA^[25]) and the *N*-sulfanylethylaniide (SEAlide^[26]) devices, they consist of special amides that are able to rearrange *in situ* into thioesters through an *N,S*-acyl shift (Figure 6, STEP 1). Similar to classical thioesters, these transient thioesters can then be involved in an NCL-like process to provide ligated peptides (Figure 6A, STEPS 2 and 3; Figure 6B, STEP 2). Classical aryl thiol additives can promote the NCL-like part of the process but not the *N,S*-acyl shift, which is usually rate limiting and catalyzed by other mechanisms. This

specific feature of the *N,S*-acyl shift systems has prompted the development of dedicated catalysts. For example, the SEA group has benefited from the design of tailored selenol-based additives (Figure 6A).^[27, 28] Thanks to the low pK_a of the selenol group (pK_a ~ 5), the catalyst is still present in its nucleophilic selenolate form in the working acidic conditions that are favorable for the *N,S*-acyl shift rearrangement step (Figure 6A, STEP 1). Its reaction with the SEA thioester intermediate results in the *in situ* formation of a selenoester (Figure 6A, STEP 2), a type of acyl donor that proved particularly powerful under classical NCL conditions.^[29]

In another example, the rearrangement of the SEALide group to generate a thioester intermediate depends on the presence of phosphate.^[26] Otaka and coworkers have developed a dual catalyst featuring a phosphonic acid able to catalyze the

rearrangement of the SEALide device (Figure 6B, STEP 1) and an aryl thiol moiety for promoting the NCL-like part of the process (Figure 6B, STEP 2).^[30]

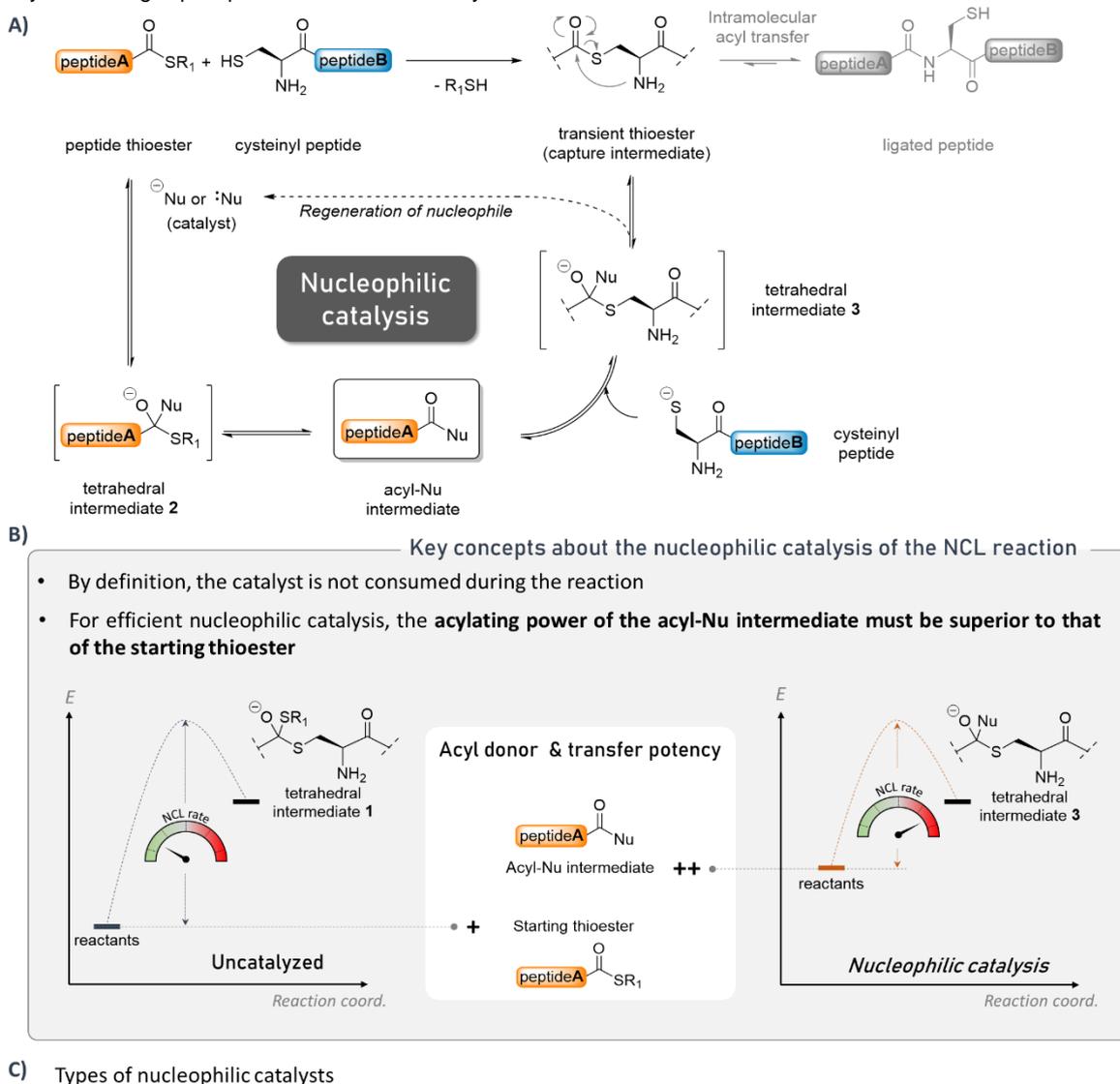
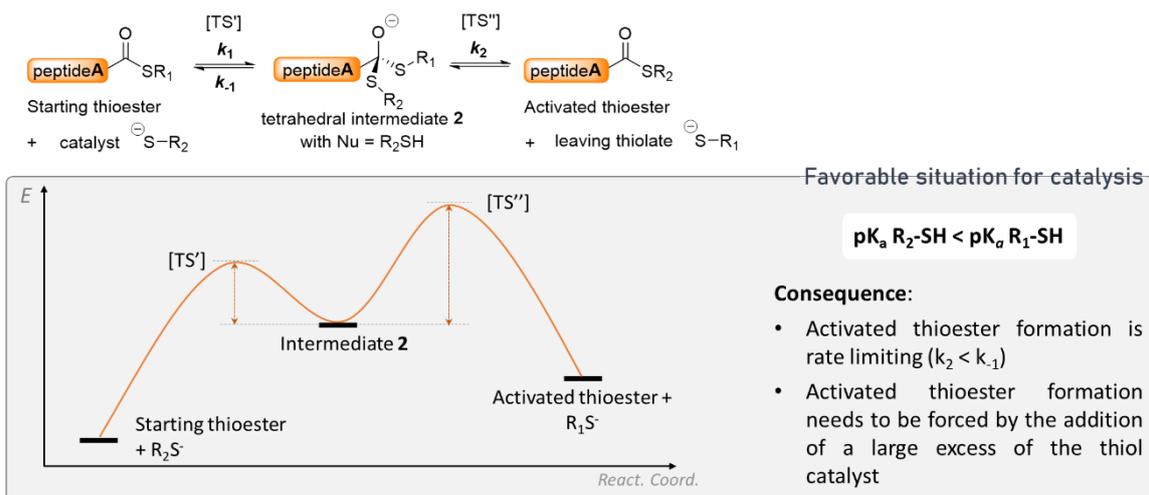


Figure 3. Activation of the thioester partner through nucleophilic acyl substitution is key to NCL rate enhancement by *in situ* generation of more efficient acyl donors. A) The mechanism of the NCL reaction catalyzed by nucleophile additives involves the formation of an acyl-nucleophile intermediate. B) Nucleophilic catalysis operates on the condition that the acyl-Nu intermediate is a better acyl transfer reactant than the starting thioester. C) Types of nucleophilic catalysts.

A) Nucleophilic catalysis by thiol additives (Nu = R₂SH) leads to the formation of more reactive thioesters



B) Identification of an appropriate thiol catalyst depending on starting thioester's nature

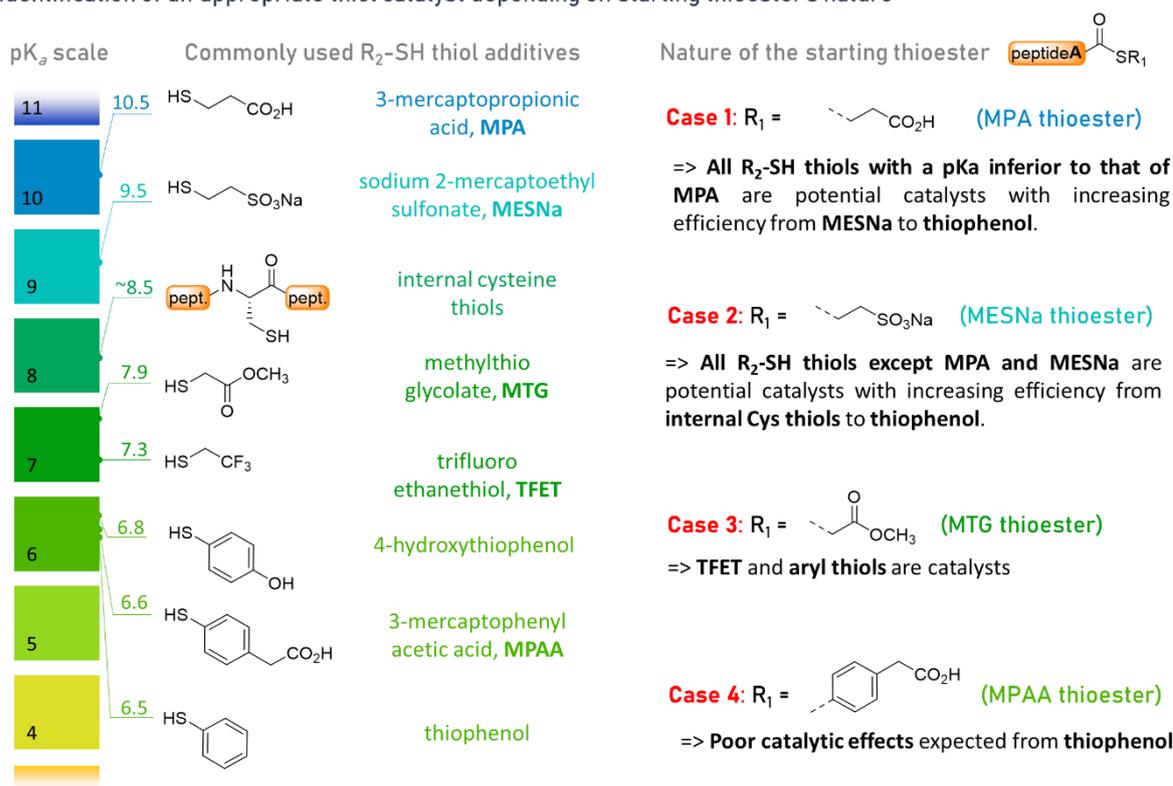
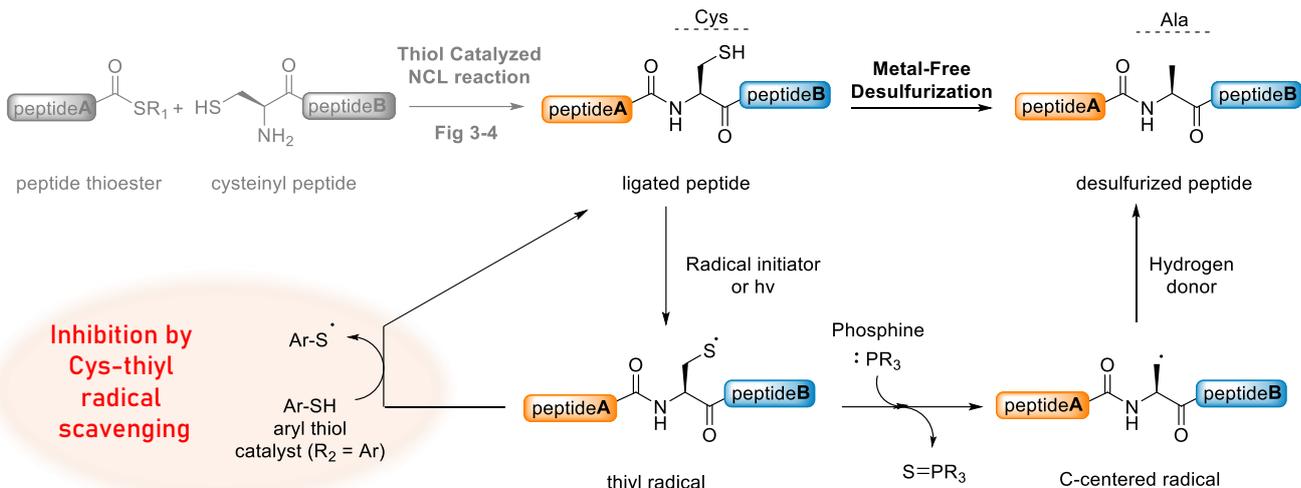


Figure 4. NCL catalysis by thiol additives. A) Principle of NCL catalysis by thiols. B) Identification of an appropriate thiol catalyst depending on starting thioester's nature: as a rule of thumb, thiol catalysts being more acidic (from top to bottom on the pK_a scale) than the departing thiol have a catalytic effect. Examples are provided at the left of the figure.

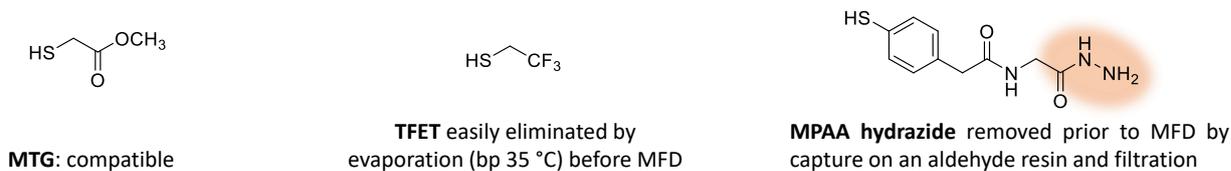
A) Compatibility of thiol additives with metal-free desulfurization (MFD)



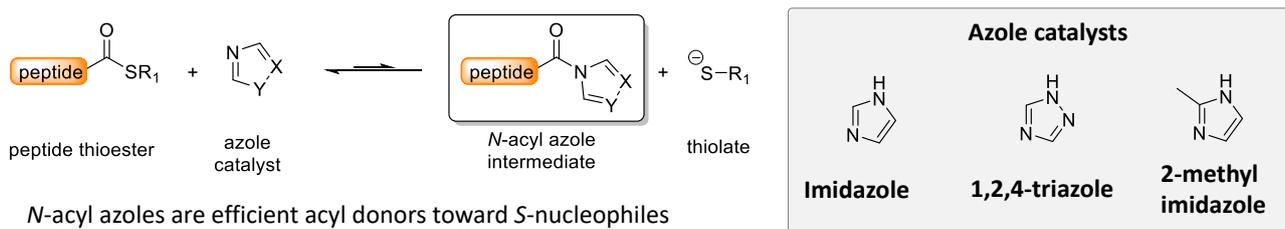
General considerations about metal-free desulfurization (MFD)

- MFD considerably extends the scope of NCL (Ala is 5.7 fold more abundant than Cys in proteins)
- NCL/MFD one-pot processes save time and yield by limiting work-up & purification steps
- Arylthiol catalysts (e.g. MPAA) inhibit MFD by neutralizing the intermediate Cys-thiyl radical

B) Which thiol catalyst to use for a one-pot NCL/MFD sequence?



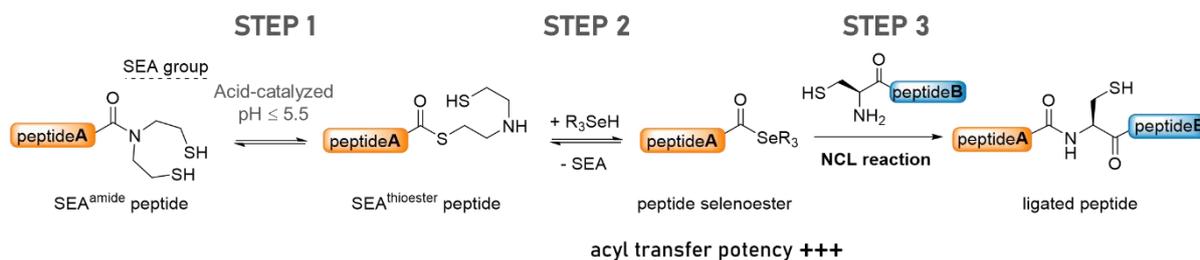
C) N_{sp}² based-catalysts: a powerful alternative to thiol catalysts for one-pot NCL/MFD approaches



- N-acyl azoles are efficient acyl donors toward S-nucleophiles
 - Process thermodynamically unfavored (shifted toward the formation of thioesters)
- => High concentrations of the azole catalyst are required to observe significant rate enhancements (> 2-5 M)

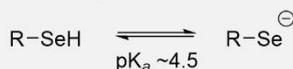
Figure 5. Compatibility of the NCL catalyst with one-pot post-ligation desulfurization methods. A) General mechanism of metal-free desulfurization (MFD) and its inhibition by aryl thiols. B) Some thiols used for one-pot NCL/MFD protocols. C). Nucleophilic catalysis by N_{sp}²-nucleophiles is compatible with one-pot NCL/MFD protocols.

A) Se-based nucleophilic catalysts for SEA-mediated ligation

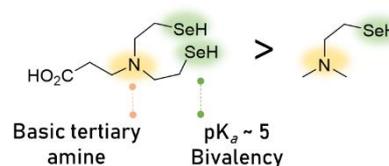


- Step 1 requires acidic pH to proceed at acceptable rate (pH 5.5)
- Common thiol additives are poor catalysts of thiol-thioester exchanges due to their pK_a well above the working pH

➔ Catalysis by selenols



Key features for Se-based catalysts ($R_3\text{SeH}$)



B) Dual phosphonic acid-arylthiol catalysts for SEALide-mediated ligation

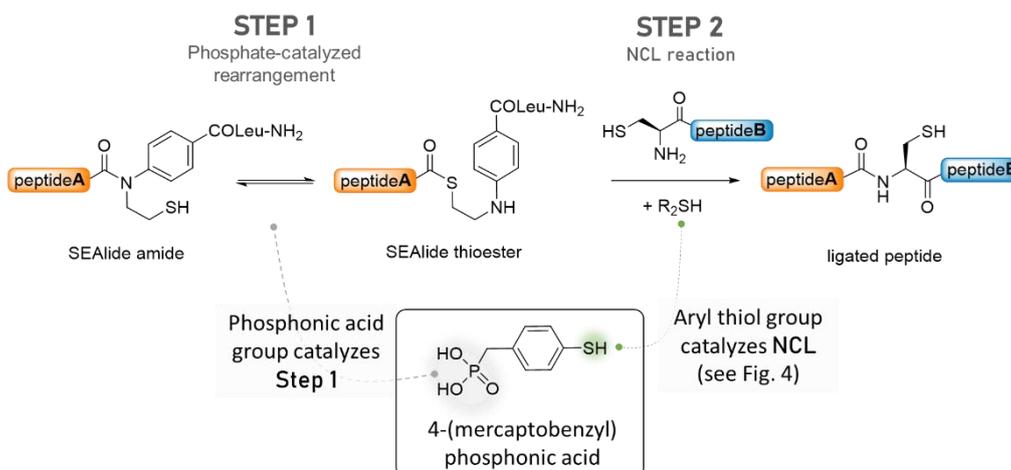


Figure 6. Catalysis of *N,S*-acyl shift-mediated ligations. A) Catalysis of SEA ligation by selenol compounds. B). Catalysis of SEALide-mediated ligation by a dual phosphonic acid-aryl thiol catalyst.

Summary and outlook

In conclusion, the rapid concatenation of peptide segments in water using NCL is central to protein chemical synthesis and many other applications of this chemoselective reaction. The catalysis of NCL and related reactions enables to save time, to force the formation of difficult junctions and to minimize the impact of protein aggregation and potential side reactions. Throughout the years, this quest for rate enhancement has been sustained by an in-depth understanding of reaction mechanisms based on experimental and theoretical approaches.

A limitation of existing catalysts is the large excesses used to achieve useful synthetic rates. Therefore, future research might see the emergence of novel catalysis approaches enabling the acceleration of ligation reactions while significantly reducing the heavy excesses of nucleophiles. A potential strategy can be to increase the effective concentration of the nucleophilic catalyst at the thioester site by incorporating it in the thiol handle. By

becoming intramolecular, the exchange process leading to the formation of the activated acyl-Nu species might not be limiting anymore. Such an approach using a Cys thiol as the internal nucleophile has been examined by Schmalisch and Seitz, who reported significant rate enhancements.^[31]

Non-covalent strategies for increasing the effective concentration of the nucleophilic catalyst at the thioester site remains to be developed. The concepts underpinning template-assisted NCL reactions might serve this purpose, albeit to the expense of increased synthetic work and reduced atom economy.^[2, 32, 33] Templated NCL has been developed for enabling NCL at very low reactant concentration such as in the labelling of proteins exposed to the cell surface.^[34] The method relies on the high affinity between pairs of nucleic acids or peptides for increasing the effective concentration of reactants and thus the ligation yields. Likewise, we can imagine that the biomolecular systems used so far for promoting reactant

encounter could be simplified in order to extend their application to catalysis.

Acknowledgements

This work was supported by the Centre National de la Recherche Scientifique (CNRS), the University of Lille Nord de France, Institut Pasteur de Lille.

Conflict of interest

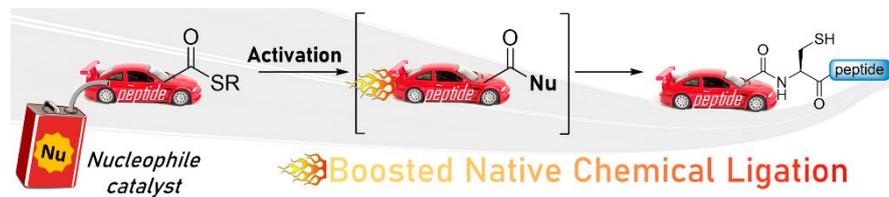
The authors declare no conflict of interest.

Keywords: Native Chemical Ligation • homogeneous catalysis • nucleophilic catalysis • thiol-thioester exchange • desulfurization • chemoselectivity

Chemical Synthesis of Proteins, Brik, A., Liu, L., Dawson, P., Eds. Wiley: 2021; pp 59-85.

- [24] Melnyk, O., Agouridas, V., *Curr. Opin. Chem. Biol.* **2014**, *22*, 137-145.
- [25] Ollivier, N., Dheur, J., Mhida, R., Blanpain, A., Melnyk, O., *Org. Lett.* **2010**, *12*, 5238-5241.
- [26] Sato, K., Shigenaga, A., Tsuji, K., Tsuda, S., Sumikawa, Y., Sakamoto, K., Otaka, A., *ChemBioChem* **2011**, *12*, 1840-4.
- [27] Kerdraon, F., Bogard, G., Snella, B., Drobecq, H., Pichavant, M., Agouridas, V., Melnyk, O., *Molecules* **2021**, *26*.
- [28] Cargoët, M., Diemer, V., Snella, B., Desmet, R., Blanpain, A., Drobecq, H., Agouridas, V., Melnyk, O., *J. Org. Chem.* **2018**, *83*, 12584-12594.
- [29] Durek, T., Alewood, P. F., *Angew. Chem. Int. Ed.* **2011**, *50*, 12042-5.
- [30] Eto, M., Naruse, N., Morimoto, K., Yamaoka, K., Sato, K., Tsuji, K., Inokuma, T., Shigenaga, A., Otaka, A., *Org. Lett.* **2016**, *18*, 4416-9.
- [31] Schmalisch, J., Seitz, O., *Chem. Commun.* **2015**, *51*, 7554-7557.
- [32] Seitz, O., *J. Pept. Sci.* **2019**, *25*, e3198.
- [33] Giesler, R. J., Erickson, P. W., Kay, M. S., *Curr. Opin. Chem. Biol.* **2020**, *58*, 37-44.
- [34] Reinhardt, U., Lotze, J., Zernia, S., Morl, K., Beck-Sickinger, A. G., Seitz, O., *Angew. Chem. Int. Ed.* **2014**, *53*, 10237-41.
- [1] Dawson, P. E., Muir, T. W., Clark-Lewis, I., Kent, S. B. H., *Science* **1994**, *266*, 776-779.
- [2] Agouridas, V., El Mahdi, O., Diemer, V., Cargoët, M., Monbaliu, J.-C. M., Melnyk, O., *Chem. Rev.* **2019**, *12*, 7328-7443.
- [3] Conibear, A. C., Watson, E. E., Payne, R. J., Becker, C. F. W., *Chem. Soc. Rev.* **2018**, *47*, 9046-9068.
- [4] Brik, A., Dawson, P. E., Liu, L. (Eds), *Total Chemical Synthesis of Proteins*, Wiley-VCH, Weinheim, **2021**.
- [5] Protein Chemical Synthesis database (PCS-db). <http://pcs-db.fr>, 2017 (accessed 20 July 2021)
- [6] Agouridas, V., El Mahdi, O., Cargoët, M., Melnyk, O., *Bioorg. Med. Chem.* **2017**, *25*, 4938-4945.
- [7] Agouridas, V., El Mahdi, O., Melnyk, O., *J. Med. Chem.* **2020**, *63*, 15140-15152.
- [8] Dawson, P. E., Churchill, M. J., Ghadiri, M. R., Kent, S. B. H., *J. Am. Chem. Soc.* **1997**, *119*, 4325-4329.
- [9] Maguire, O. R., Zhu, J., Brittain, W. D. G., Hudson, A. S., Cobb, S. L., O'Donoghue, A. C., *Chem. Commun.* **2020**, *56*, 6114-6117.
- [10] van der Helm, M. P., Klemm, B., Eelkema, R., *Nat. Rev. Chem.* **2019**, *3*, 491-508.
- [11] Hupe, D. J., Jencks, W. P., *J. Am. Chem. Soc.* **1977**, *99*, 451-464.
- [12] Evans, T. C., Jr., Benner, J., Xu, M. Q., *Protein Sci.* **1998**, *7*, 2256-64.
- [13] Huang, Y.-C., Chen, C.-C., Gao, S., Wang, Y.-H., Xiao, H., Wang, F., Tian, C.-L., Li, Y.-M., *Chem. Eur. J.* **2016**, *22*, 7623-7628.
- [14] Thompson, R. E., Liu, X., Alonso-Garcia, N., Pereira, P. J., Jolliffe, K. A., Payne, R. J., *J. Am. Chem. Soc.* **2014**, *136*, 8161-4.
- [15] Johnson, E. C., Kent, S. B. H., *J. Am. Chem. Soc.* **2006**, *128*, 6640-6.
- [16] Tsuda, S., Yoshiya, T., Mochizuki, M., Nishiuchi, Y., *Org. Lett.* **2015**, *17*, 1806-1809.
- [17] Wan, Q., Danishefsky, S. J., *Angew. Chem. Int. Ed.* **2007**, *46*, 9248-52.
- [18] Moyal, T., Hemantha, H. P., Siman, P., Refua, M., Brik, A., *Chem. Sci.* **2013**, *4*, 2496-2501.
- [19] Sakamoto, K., Tsuda, S., Nishio, H., Yoshiya, T., *Chem. Commun.* **2017**, *53*, 12236-12239.
- [20] Sakamoto, K., Tsuda, S., Mochizuki, M., Nohara, Y., Nishio, H., Yoshiya, T., *Chem. Eur. J.* **2016**, *22*, 17940-17944.
- [21] Chisholm, T. S., Clayton, D., Dowman, L. J., Sayers, J., Payne, R. J., *J. Am. Chem. Soc.* **2018**, *29*, 9020-9024.
- [22] Raibaut, L., Ollivier, N., Melnyk, O., *Chem. Soc. Rev.* **2012**, *41*, 7001-7015.
- [23] Diemer, V., Bouchenna, J., Kerdraon, F., Agouridas, V., Melnyk, O., N,S- and N,Se-Acyl Transfer Devices in Protein Synthesis. In *Total*

Entry for the Table of Contents



Nucleophilic catalysis is key to the acceleration of the native chemical ligation (NCL) reaction and thus, to its application for protein synthesis, functionalization and utilization in chemical biology. The present concept article discusses the fundamental principles of NCL rate and catalysis with emphasis on thiol nucleophiles.

Institute and/or researcher Twitter usernames: [@Melnyk_Group](#); [@PcsDb](#)