New enzymatic approaches for the generation of peptide and protein conjugates

Can Araman

Gorlaeus Laboratories, Niels Bohrweg 2, 2333 CA Leiden, the Netherlands email: m.c.araman@lic.leidenuniv.nl

The field of protein conjugation is growing rapidly. Non-native handles, such as polyethylene glycol (PEG) or fluorophores, can easily be conjugated to proteins and peptides via amines or hetero-Michael type reactions of cysteines (Figure 1A).[1] Additionally, site-specific introduction of non-canonical amino acids (azidolysine/propargylglycine/cyclopropen-l-lysine, Figure 1B), which can undergo chemoselective reactions, to peptides or proteins either via SPPS or genetic manipulation (amber codon suppression or use of auxotrophic strains) are accepted as essential methodologies and result in great conjugation efficiencies *in vitro* and *in vivo*.[1]

Besides well-established chemical methods for site-specific labeling, such as native chemical ligation (NCL), copper-catalyzed- azide-alkyne-cycloaddtion (CuAAC), Staudinger ligation and inverse-electron-demand Diels-Alder cycloaddtition (IEDDA) (reviewed elsewhere [1]), enzymatic ligation methods asserted themselves as adequate substitutes for peptide conjugation. To date, more than 20 enyzmes, especially sortase A and transglutaminase, are widely used for protein conjugation reactions. [2] Thus, enzymatic conjugation is relevant more than ever in peptide/protein chemistry.

In this commentary, I will discuss two recent publications on innovative enzymatic approaches to generate peptide and protein conjugates; SnoopLigase (Buldun *et al. J. Am. Chem. Soc.* **2018**, 140, 3008−3018) [3] and formylglycine-generating enzymes, FGEs (Krueger *et al. Angew. Chem. Int. Ed.* **2018**, ASAP, 10.1002/anie.201803183). [4]

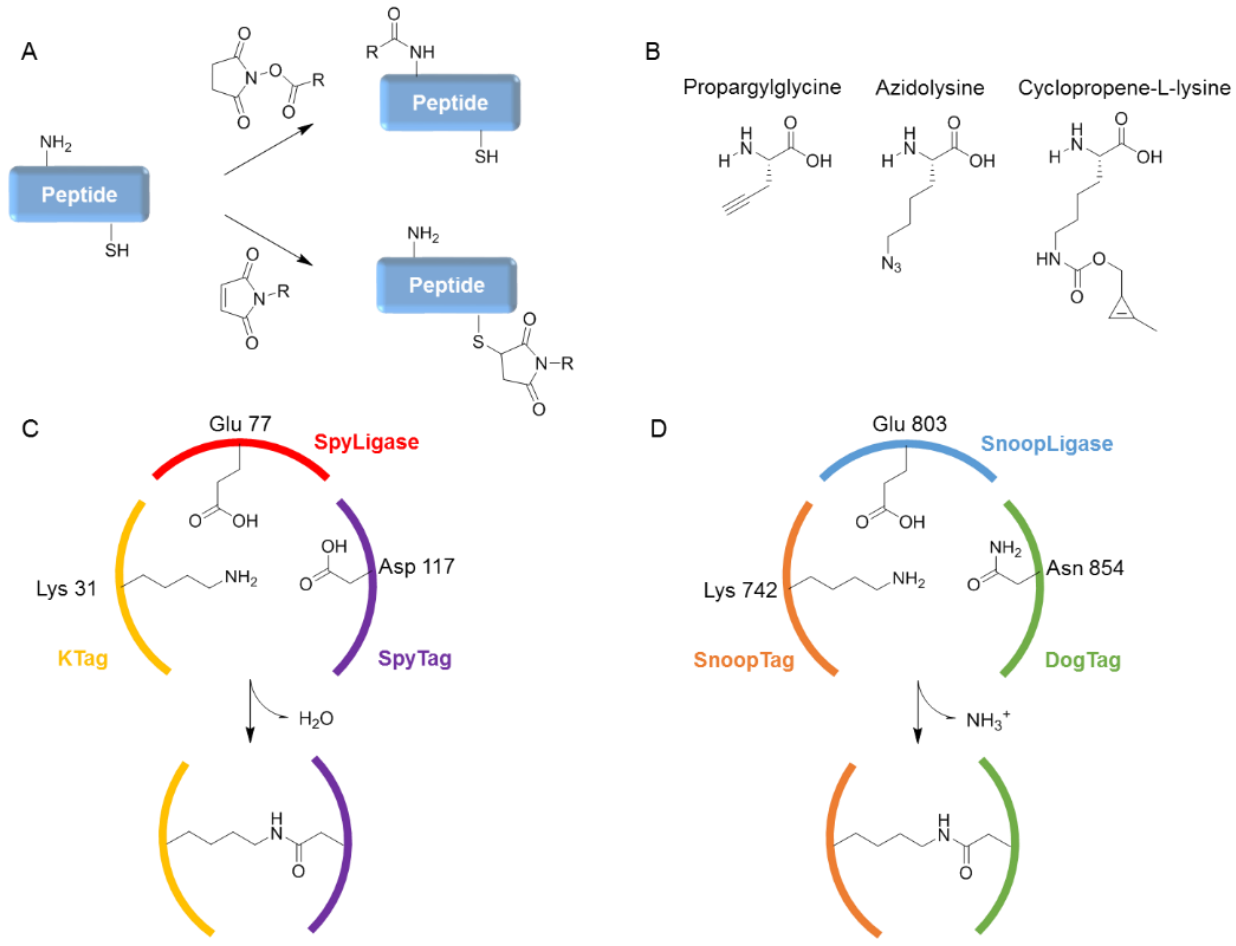

Figure 1. A) Lysine and cysteine modifications on proteins, B) Unnatural amino acids used in protein labeling, C) SpyLigase- mediated isopeptide bond formation, D) SnoopLigase-mediated isopeptide bond formation

SNOOPLIGASE

Already in 2014, the group of Mark Howarth reported a new enzyme for peptide conjugation, namely SpyLigase.[5] In their previous article, the researchers splitted FbaB from *Streptococcus pyogenes* into three segments, generating the so-called SpyTag and K-Tag as well as the ligase (Figure 1C). Next, they were able to promote isopeptide bond formation between the ε-NH_2 of a lysine on K-Tag and the carboxylic acid side chain of an aspartate of Spy-Tag conjugated to different proteins. Despite obtaining promising results, the conjugation efficiency could only be raised up to 50%. Hence, the researchers searched for other enzymes that form isopeptide bonds and found *Streptococcus pneumoniae* adhesin RrgA. The *C*-terminal part of RrgA can spontaneously form an isopeptide bond between a lysine and asparagine catalyzed via a glutamate residue. Therefore, they opted for the splitting of RrgA ligase (Figure 1D) and generating two tags termed SnoopTag (12 amino acids with crucial Lys), DogTag (23 amino acids with crucial Asn) as well as the ligase itself (104 amino acids comprising catalytical Glu). Furthermore, it was hypothesized that, using this enzyme, a thermodynamically more favorable isopeptide bond formation reaction over SpyLigase would occur, since ammonia and not water would be produced as a side-product (Figure 1C-D). The researchers introduced point mutations (two proline mutations at positions 808 and 837) to stabilize the split domain by introducing β-turns, which additionally resulted in enhanced enzymatic activity. A more drastic activity increase was achieved (~ 66-fold) when three further mutations at positions 737, 838 and 839, D→S, D→G and I→V respectively, were introduced to SnoopTag (SnoopTagJr). A proof-of-concept study was done via conjugating SnoopTagJr to an affibody against the growth factor receptor HER2 and DogTag to SUMO. The ligation yields for this reaction were reported to be quantitative (>95%).

To test the scope of their method, Buldun *et al.* attempted to generate four protein conjugates using SnoopLigase, with highly successful conjugation yields (>78%), even when DogTag was inserted into protein loops. Next, attention was drawn to the recycling of SnoopLigase and separation of product from the starting materials. To retain proper folding of the target conjugate and the ligase, the researchers developed an HPLC-free approach; the ligase was expressed with a biotin tag which enables immobilizing on streptavidin functionalized agarose beads. The conjugation product could be eluted via affinity based techniques under different conditions (glycine pH 2.0, 2 M imidazole or a competitor peptide with a high affinity to SnoopLigase). All three methods gave rise to quantitative recovery of the product. Moreover, the enzyme could be reused in at least 8 cycles for further ligation reactions.

Based on this data, the researchers aimed to generate a thermally stable cyclic enzyme. In order to do this, SnoopTagJr and DogTag were fused to *N*- and *C*-terminus of PhyC, an enzyme increasing the digestibility of phosphate and thus reducing environmental phosphate pollution. As expected, cyclized PhyC showed a greater thermal stability (resistance up to 100 °C) when compared to its linear variant (precipitation between 60-80 °C) and could be recycled for further use.

To summarize, Howarth and colleagues developed an elegant method to generate peptide/protein conjugates with simple purification protocols and quantitative conversion yields. Moreover, using this approach, thermal stability of enzymes could be increased via cyclization, with similar activity to that of native, linear enzymes.

FGE MEDIATED BIOCONJUGATION

Another promising enzymatic approach for peptide conjugation is the recently described "dual labeling via distinct FGEs" by Sewald and colleagues, in which they developed a new method for dual labeling of peptides using both FGE and its analogues prokaryotic (from *Methanosarcina mazei*) iron-sulfur protein AtsB.[5] Both enzymes can convert cysteines into formylglycines (FGly), which could undergo a variety of conjugation reactions, e.g. oxime ligations.[6] Interestingly, the consensus sequence for FGE is CXPXR, whereas AtsB prefers an alanine instead of the proline at the same position. Therefore, it was concluded that AtsB may generate FGly without converting the *N*-terminal cysteine in FGE-specific consensus sequence within the same peptide or protein.

To prove their hypothesis, Sewald and colleagues synthesized two peptides; a 15-mer **1** with the recognition sequence for AtsB and a 24-mer **2** comprising both consensus sequences (Figure 2), and treated them first with AtsB. As expected, after 80 min incubation, only one FGly could be detected and assigned to the AtsB specific motif CXAXR via MS^2 analysis. Subsequently, **2** was treated with FGE (80 min) to give rise to double FGly modification (Figure 2) with good conversion efficiency (~ 70%). With this proof of concept study, it was shown that subsequent dual labeling with AtsB and FGE is possible.

Based on this background, the researchers designed DARPin E01 protein constructs *C*-terminally fused to either AtsB or both AtsB and FGE recognition motifs. For all protein variants no alterations in secondary structure was observed. Next, Sewald and colleagues aimed for the dual labeling of DARPin derivatives. For this purpose, they incubated the proteins with AtsB to generate the first and site-specific aldehyde functionality. Subsequently they used a hydrazino-*iso*-Pictet-Spengler (HIPS) type conjugation to attach a hydrazine functionalized fluorophore to the aldehyde moiety (Figure 2).[7] The ligation reaction was analyzed via SDS-PAGE, in-gel fluorescence, and mass spectrometry. Following the ligation, anion exchange chromatography was utilized to purify the product. Upon purification, the second cysteine within the FGE recognition sequence was converted into an aldehyde via FGE incubation. The subsequent HIPS ligation gave rise to the conjugation of a PEG_{2k} molecule to the aldehyde residue.

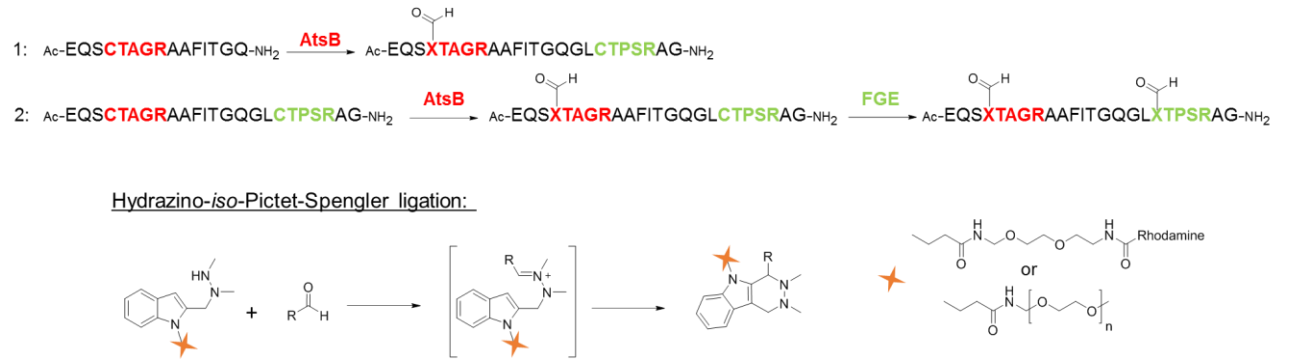

Figure 2. AtsB and FGE mediated HIPS ligation strategy

DARPin E01 is a natural binder for the epidermal growth factor receptor (EGFR). To test whether the binding efficiency of DARPin is retained after FGE mediated conjugation, cell binding experiments (FACS and confocal microscopy) at the surface of EGFR rich A431 cells were conducted. Indeed, neither the genetic changes on the *C*-terminus nor HIPS-mediated fluorophor-labelling of DARPin E01 derivatives did alter the binding efficiency to EGFR. Thus, this method showed a significant bio-orthogonal potential even though the reactions were not conducted under bio-orthogonal conditions (in a living system, in the presence of other nucleophiles/reaction partners, etc.).

Hence, the researchers reported the first utilization of a dual and orthogonal FGly labeling strategy. It would be interesting to further explore the bio-orthogonality of this method, by performing FGly conversion in mammalian cells overexpressing certain proteins. In any event, the method described herein is elegant and simple to perform, enabling a wider use by the protein conjugation community.

In summary, two promising enzymatic peptide/protein conjugation methods are here discussed. High selectivity and high conjugation yields are among the most important advantages of enzymatic peptide conjugation. However, the stoichiometric amount of enzymes required, the need for sophisticated purification steps after conversion, as well as the use of artificial and long tags are major limitation factors for these strategies. Buldun *et al.* provide an elegant solution to probable purification problems via attaching a biotin-tag to the ligase, whereas Sewald and coworkers use a maximum amount of six amino acids at the conjugation site. In doing so, minimum to no changes were observed in structural and biophysical properties of the proteins used for conjugation.

In my opinion, both methods will have a significant impact on enzyme mediated peptide and protein-conjugation and provide effective solutions to known limitations of existing methods.

REFERENCES

1. N. Krall, F. P. da Cruz, O. Boutureira, G. J. Bernardes, *Nat. Chem.*, 2016, **8**, 103-113. PMID: 26791892
2. E. M. Milczek, *Chem. Rev.*, 2018, **118**, 119-141. PMID: 28627171
3. C. M. Buldun, J. X. Jean, M. R. Bedford, M. Howarth, *J. Am. Chem. Soc.*, 2018, **140**, 3008-3018. PMID: 29402082
4. T. Krüger, S. Weiland, G. Falck, M. Gerlach, M. Boschanski, S. Alam, K. M. Müller, T. Dierks, and N.Sewald, *Angew. Chem. Int. Ed.*, 2018, **Accepted Author Manuscript**, doi:10.1002/anie.201803183. PMID: 29579347
5. J. O. Fierer, G. Veggiani, M. Howarth, *Proc. Nati. Acad. Sci.* 2014, **111**, E1176-E1181. PMID: 24639550
6. U. Sébastien, B. Didier, M. Alberto, R. Olivier, D. Pascal, *Chem. Eur. J.*, 2014, **20**, 34-41. PMID: 24302514
7. P. Agarwal, R. Kudirka, A. E. Albers, R. M. Barfield, G. W. de Hart, P. M. Drake, L. C. Jones, D. Rabuka, *Bioconj. Chem.* 2013, **24**, 846-851. PMID: 23731037