ChemComm
Chemical Communications

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: X. Mao, W. Li, S. Zhu, J. Zou, H. Tian, Y. Duan, Y. Wang, J. Fei and X. Wang, *Chem. Commun.*, 2020, DOI: 10.1039/D0CC02722G.

View Article Online
DOI: 10.1039/D0CC02722G

COMMUNICATION

Bifunctional Pyridoxal Derivatives as Efficient Bioorthogonal Reagents for Biomacromolecule Modifications

Received 00th January 20xx,
Accepted 00th January 20xx

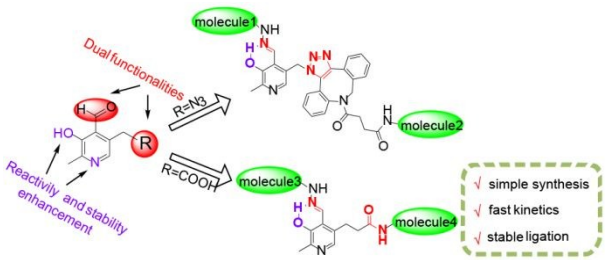
DOI: 10.1039/x0xx00000x

Xianxian Mao, Wei Li, Shiyu Zhu, Juan Zou, Hongyan Tian, Yuting Duan, Yuntao Wang, Jiayue Fei and Xiaojian Wang*

Two types of pyridoxal analogs, azido pyridoxal (PL-N_3) and carboxyl pyridoxal (PL-COOH), were developed as novel bifunctional bioorthogonal molecules. These molecules showed fast imine formation with hydrazinyl groups and stable covalent linkages via azido/carboxyl groups, and thus was of great use for site-specific peptide and protein modifications.

In the fields of chemical biology and biotechnology, a number of bioorthogonal tools have been developed to modify biological molecules such as nucleic acids, proteins and lipids to achieve the purpose of labeling and/or tracking of these molecules.[1] Among a large number of bioorthogonal reactions, the hydrazone/oxime ligation is one of the earliest and most widely used reactions.[2] Francis group has extensively studied the transamination reaction at the N-terminus of proteins with various compounds, which yielded a ketone or aldehyde group at that end of the protein for the further labeling via aldoxime condensation.[3] Kool et al reported the reaction kinetics between a variety of different carbonyl compounds and hydrazines, revealing the impact of functional groups at different substitution cites.[4]

The coenzyme pyridoxal-5'-phosphate (PLP) is a unique example of aldehydes that has been used for bioconjugation due to its fast kinetics and relatively stable imine products.[5] Hsein et al utilized PLP to label the amino dendrimer with imine formation.[6] Wang et al reported a simple two-step aqueous protocol which successfully modified different proteins with PLP through a phosphoramide linkage.[7] However, one major limitation of PLP is the poor stability of phosphoramide bond bridging PLP and amino compounds of interest, such as proteins and peptides, which makes PLP unsuitable for a variety of applications, especially ones involving acidic conditions.[8] Herein, we reported the design and synthesis of two types of PLP analogs (Scheme 1, Fig. 1a): azido pyridoxal (PL-N_3, **1**) and carboxyl pyridoxal (PL-COOH, **2** and **3**). Azido pyridoxal **1** was designed to react with alkynes, forming a stable covalent linkage and thus suitable for a wide range of crosslinking applications. PL-COOHs were designed with the carboxyl group bridged with linkers of different lengths, which replaces phosphate group of PLP for the more stable amide linkage, extending the application scope of such molecules. To the best of our knowledge, PL-COOH **3** was previously reported as an enzyme cofactor or protein antagonist, but never explored as a biorthogonal reagent.[9]



Scheme 1 Design strategy of PL analogs

The conversion from the phosphate of PLP to azide or carboxyl might have a negative effect on the activity, as it was reported that the phosphate group may catalyze the imine formation as an internal proton source.[5b, 10] However, we speculate that this might not be a major defect, because it was reported that the 5'-deoxypyridoxal had only a slightly weaker reactivity than PLP when reacting with glycine,[11] and PLP is one of the most reactive aldehyde reported.[4d] To further evaluate the design, DFT calculations were employed to estimate the reactivity of PL-N_3 and PL-COOHs, comparing to PLP, using methyl amine as a model reactant (see Supporting Information for calculation details). The analysis of rate-determine dehydration step yielded small differences of activation free energies among the four (0.06 kcal/mol for **1** vs PLP, 0.1 kcal/mol for **2** vs PLP and -0.53 kcal/mol for **3** vs PLP), indicating small effects of the substitutions on reactivity.[12]

Institute of Advanced Synthesis, School of Chemistry and Molecular Engineering, Nanjing Tech University, Nanjing 211816, China.
Email: ias_xjwang@njtech.edu.cn
Electronic Supplementary Information (ESI) available: Experimental details and supplementary figures.
See DOI: 10.1039/x0xx00000x

The designed compounds were then synthesized as shown in Fig. 1b. Briefly, starting from commercially available pyridoxine hydrochloride, 3 and 4' hydroxyl groups were first protected as the cyclic acetal **5** to prevent undesired side reactions.[13] The 5' hydroxyl group was then either converted into the azido group with DPPA through Mitsunobu reaction, or into the chloride which was further substituted by the cyanide group. The acetal protection was then cleaved with acid, and the cyanide was also hydrolyzed into carboxyl during the same process. Finally, the released 4' hydroxyl was oxidized into aldehyde with MnO_2, yielding azido pyridoxal **1** (PL-N_3) and carboxyl pyridoxal (PL-COOH) **2**. The PL-COOH **3** was prepared by oxidation of reported carboxyl pyridoxine derivative (see Supporting Information, Fig. S9).[14]

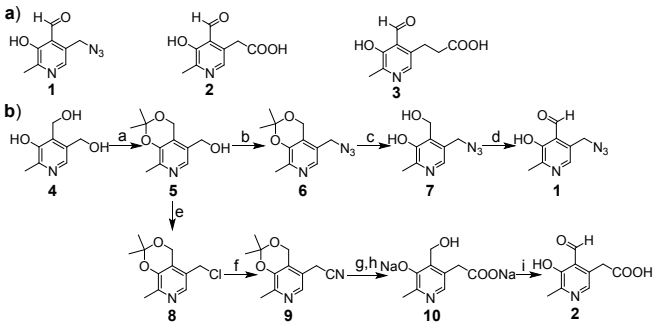


Fig. 1 (a) Structure of designed and synthesized **1** and **2**. (b) Synthesis of **1** and **2**. a) DMP, TsOH, acetone, RT, 20 h, 63%. b) DPPA, DBU, THF, RT, 4 h, 97%. c) 1M HCl, THF, reflux, 4 h, 95%. d) MnO_2, THF, RT, 6 h, 54%. e) $SOCl_2$, CH_2Cl_2, RT, 2 h, 91%. f) TMSCN, NaF, DMF, 69 °C, overnight, 94%. g) 6 M HCl, Reflux, 11 h. h) H_2O, 1 M NaOH, 40 °C, 97% over two steps. i) MnO_2, RT, overnight, 15%. DMP = 2,2-dimethoxypropane; TsOH = p-toluenesulfonic acid; DPPA = diphenylphosphoryl azide; DBU = 1,8-diazabicyclo[5.4.0]undec-7-ene; $SOCl_2$ = thionyl chloride; TMSCN = trimethylsilyl cyanide.

The reactivities of synthesized pyridoxal derivatives were then studied with 2-hydrazinopyridine (HP) as a model compound, comparing to the naturally occurring pyridoxal-5'-phosphate (PLP) and pyridoxal (PL).[15] Reactions were carried out in phosphate buffered saline (PBS, pH 7.4) and monitored with UV-VIS spectrometer, following the peak formation of the corresponding hydrazone.[7] The experimental data of PL-N_3 **1** and PL-COOH **3** fitted the rate equation for a reversible second-order reaction, and PL fitted the rate equation for a pseudo-first order reaction (see Supporting Information, Fig. S11,), while experiment with PL-COOH **2** yielded irreproducible results (data not shown).[16] As shown in Table 1, The reactivity of PL-N_3 and PL-COOH **3** was much higher than PL, since the intermolecular hemiacetal was broken after the substitution of the 5' hydroxyl group, which was consistent with previous report.[17] Two control bifunctional aldehydes, 2-(azidomethyl)benzaldehyde (2-AMBA) and 2-formyl benzoic acid (2-FBA), were also examined, yielding much slower kinetics. PL-N_3 and PL-COOH **3** were sightly less reactive but still comparable to PLP, and should be of great use for bioconjugation applications, because their hydrazone formation rates were among the fastest ones,[4d] and much higher than the most widely used Cu-free azide–alkyne cycloadditions (0.057 M^{-1} s^{-1} for dibenzocyclooctyne and 0.29 M^{-1} s^{-1} for aza-dibenzocyclooctyne).[18]

Table 1 The reaction rate (k_1) and equilibrium constant (K_{eq}) between aldehydes and HP.

Compounds	K_1 ($M^{-1}s^{-1}$)	K_{eq} (M^{-1})
PL-N_3	2.6 ± 0.1	(2.1 ± 0.5) × 10^5
PL-COOH 3	5.8 ± 0.1	(1.2 ± 0.2) × 10^5
PLP	8.5 ± 0.1	(1.7 ± 0.1) × 10^5
PL	0.11 ± 0.01	N/D[a]
2-AMBA	0.15 ± 0.01	N/D[a]
2-FBA	0.21 ± 0.01	N/D[a]

[a]N/D: not determined

Thus, in order to study PL-N_3 mediated biomacromolecular modification, a hydrazine containing dye, fluorescein-5-thiosemicarbazide (FTZ) was first tested as a model molecule. PL-N_3 was incubated with cyclooctyne modified bovine serum albumin (BSA-FDIBO, **12**)[19] to obtain **13** (BSA-FDIBO-PL), which was then reacted with FTZ in PBS, yielding fluorescein-labled BSA (Fig. 2a). The reaction mixture was purified via ultracentrification and analyzed with UV-VIS spectroscopy. As shown in Fig. 2b, the absorption peak corresponding to PL-N_3 at 390 nm and FTZ at 490 nm was observed, and an average of 2.7 PL-N_3 and 0.85 FTZ molecules per BSA were estimated. The apparent low yield (31.5%) was mostly due to the undesired modification of PL-N_3 on BSA-FDIBO via imine formation or nonspecific absorption, which was also observed in the control experiment with BSA and PL-N_3 (see Supporting Information, Fig. S13a). Further analysis was conducted with sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE, Fig. 2c), and only BSA-PL-FTZ showed a strong fluorescent band (column 7), whereas the controls of unmodified BSA with FTZ (column 4) or BSA-FDIBO with FTZ (column 5) had negligible signal.

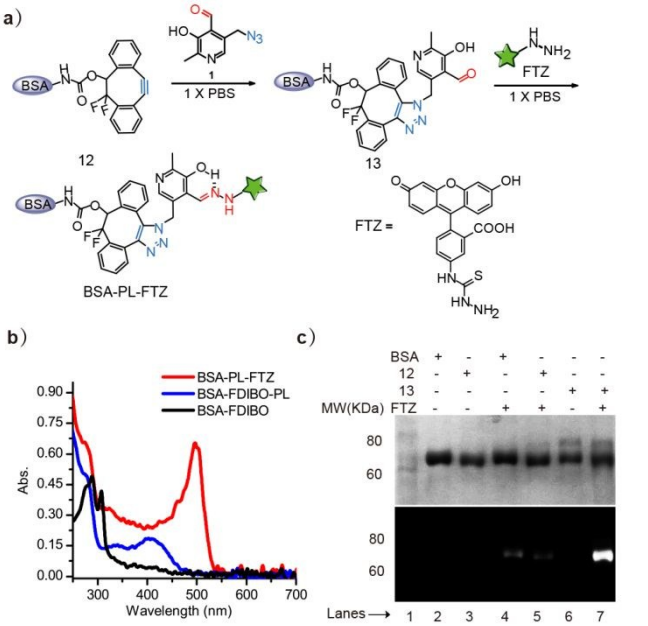


Fig. 2 (a) Modification of BSA with FTZ. (b) UV-vis absorption spectra of BSA-FDIBO (10.18 μM, black), BSA-FDIBO-PL (10.11 μM, blue), and BSA-PL-FTZ (10.07 μM, red). (c) SDS-PAGE visualized under UV light (bottom) and after Coomassie-blue stain (top). Well from left to right: protein ladder (each line: 80 kDa and 60 kDa, from top to bottom), BSA, BSA-FDIBO, BSA~FTZ, BSA-FDIBO~FTZ, BSA-FDIBO-PL and BSA-PL-FTZ from 2 to 7.

A major challenge of biomolecule conjugation is to make site-specific modification, which often involve unnatural amino acids and nucleic acids. Inspired by the reported work from Li et al,[20] a sortase A (SrtA) mediated reaction protocol was employed to specifically modify a protein's C-terminal into an azide functionality with PL-N_3. As shown in Fig. 3a, enhanced green fluorescent protein (eGFP) containing a SrtA recognition sequence was expressed and hydrozinolyzed by SrtA.[20a] The obtained eGFP hydrazide (eGFP-$NHNH_2$) was then reacted with PL-N_3 to afford azide-modified eGFP (eGFP-PL-N_3, **14**). To confirm the successful conjugation of PL-N_3, a DIBAC modified 5-(2-aminoethylamino)-1-naphthalenesulfonic acid (DIBAC-EDANS, **15**) was reacted with eGFP-PL-N_3 to afford eGFP-PL-EDANS. As shown in Fig. 3b, the emission spectrum of DIBAC-EDANS partly overlaps with the excitation spectrum of eGFP, making them a good pair for fluorescence resonance energy transfer (FRET). Upon excitation at 335 nm, an increase of fluorescence signal was observed at 508 nm and approximately 50% of the fluorescent intensity was recovered from FRET (Fig. 3c), which was not observed by just mixing DIBAC-EDANS and eGFP in the same cuvette, proving the successful conjugation between eGFP-PL-N_3 and DIBAC-EDANS.

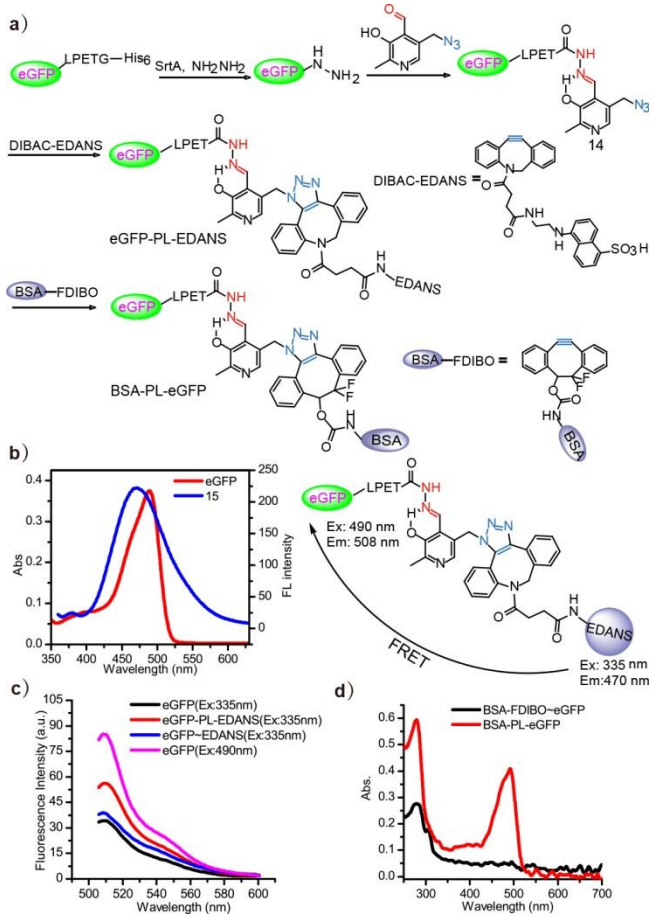


Fig. 3 (a) Modification of eGFP with DIBAC-EDANS and conjugation between eGFP and BSA with PL-N_3. (b) UV/Vis absorption of eGFP (6.6 μM, red) in 1 X PBS and emission spectra of DIBAC-EDANS **15** (0.73 μM, λ_{ex} = 335 nm, blue) in methanol. (c) Fluorescence emission spectrum of eGFP (100 nM, black, λ_{ex} = 335 nm), eGFP-PL-EDANS (100 nM, red, λ_{ex} = 335 nm), eGFP~EDANS (100 nM, blue, λ_{ex} = 335 nm), eGFP (100 nM, purple, λ_{ex} = 490 nm). (d) UV-vis absorption spectra of BSA-FDIBO~eGFP (5.23 μM, black), and BSA-PL-eGFP (7.30 μM, red).

To expand the application of PL-N_3, conjuation between two macromolecules was examined. The BSA-FDIBO was reacted with eGFP-PL-N_3 to prepare BSA-PL-eGFP (Fig. 3a). The reaction mixture was purified via ultracentrification and analyzed with UV-VIS spectroscopy. As shown in Fig. 3d, when both eGFP-PL-N_3 and BSA-FDIBO were presented in the reaction, the absorption peak corresponding to eGFP at 485 nm was observed, and an average of 1.3 eGFP molecules per BSA was estimated. The control experiment was performed with BSA-FDIBO and plain eGFP, and minimal, if any, eGFP absorption was observed.

One key limitation of PLP application was the lack of stability of the phosphoramide bond formed between the phosphate group of PLP and the amino group of biomolecules of interest. Thus, we took advantage of the reaction between the carboxyl pyridoxal PL-COOH **3** and the amino residues for the modification of polypeptides at defined sites. A model peptide with a sequence of RGDYKGGKG was synthesized via standard Fmoc solid-phase chemistry on Rink Amide MBHA resin (Fig. 4a). The two lysine residues were installed with Fmoc-Lys(Boc)-OH (*Y*K*G*) or Fmoc-Lys(Mtt)-OH (G*K*G). The Mtt group was then selectively removed and the corresponding Lys residue was modified with PL-COOH **3** to yield peptide Ac-RGDYKGGK(PL)G-NH_2 (RGDPL) via a stable amide linkage. As a comparison, PLP failed to modify the peptide through the same procedure, due to the unstable phosphoramide linkage, and only unmodified model peptide (peptiRGD) was observed on the liquid chromatography (LC) analysis (Fig. 4b). The obtained RGDPL was able to further react with FTZ, and 1.5 eq FTZ was enough to convert 96.9% RGDPL into modified product (HPLC yield), despite the possible competeing reaction between the free amino group of the other Lys residue and the installed aldehyde group (Fig. 4c). The high conversion rate, stable product formation, and good bioorthogonality further proved the usefulness of PL-COOH **3** as a click reagent.

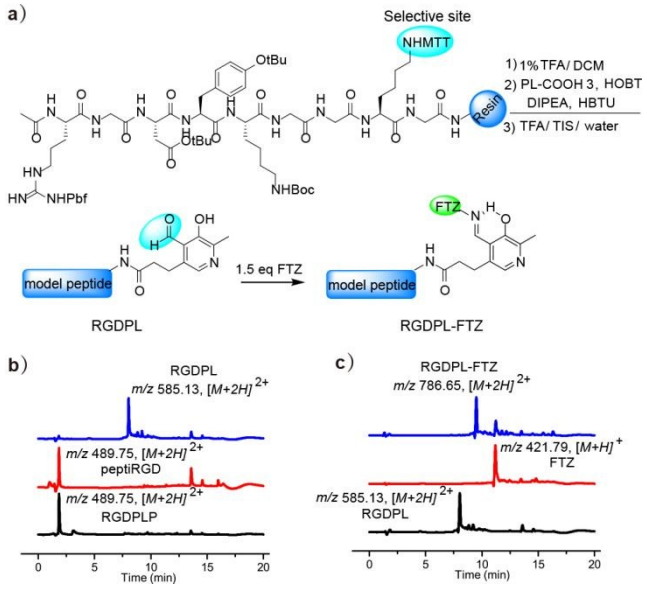


Fig. 4 (a) Modification of peptiRGD with PL-COOH **3** and RGDPL with FTZ. (b) Liquid chromatogram of peptiRGD with or withour modification: peptiRGD (red, 1.84 min), RGDPL (blue, 8.02 min) and RGDPLP (black, 1.85 min). (c) Liquid chromatogram of modification of RGDPL with FTZ: FTZ (red, 11.17 min), RGDPL (black, 8.02 min) and RGDPL-FTZ (blue, 9.49 min).

In conclusion, we have reported two novel bifunctional bioorthogonal molecules: azido pyridoxal (PL-N_3) and carboxyl pyridoxal (PL-COOH **3**). These pyridoxal analogs were easy to prepare, and had fast reaction kinetics to form stable ligation products. Azido pyridoxal (PL-N_3), a molecule containing dual bioorthogonal functionalities, showed unique properties to mediate biomacromolecular crossliking with stable triazole linkage and hydrazone formation, utilizing its azido and formyl group respectively. The carboxyl pyridoxal (PL-COOH **3**) was a useful subsitution of PLP for modification of amino groups which are abundant in biomolecules, especially peptides and proteins. These results suggest that these pyridoxal analogs are promising molecules for bioorthogonal reactions and for biomacromolecule modifications and would be of great use for a variety of applications.

This work was supported by grants from the National Natural Science Foundation of China (21708019), Natural Science Foundation of Jiangsu (BK20170987), Jiangsu province "Double Plan" and start-up fund at Nanjing Tech University (39837135). We thank Prof. Lei Liu and Prof. Yiming Li (Tsinghua University) for providing the plasmids of SrtA and eGFP. We are grateful to the High Performance Computing Center of Nanjing Tech University for supporting the computational resources.

Conflicts of interest

There are no conflicts to declare.

Notes and references

1 (a) A. A. Poloukhtine, N. E. Mbua, M. A. Wolfert, G. J. Boons and V. V. Popik, *J. Am. Chem. Soc.*, 2009, **131**, 15769; (b) C. R. Becer, R. Hoogenboom and U. S. Schubert, *Angew. Chem. Int. Ed.*, 2009, **48**, 4900; (c) S. T. Laughlin and C. R. Bertozzi, *Proc. Natl. Acad. Sci.*, 2009, **106**, 12; (d) H. Kolb, M. G. Finn and K. Barry Sharpless, *Angew. Chem. Int. Ed.*, 2001, **40**, 2004; (e) C. P. Ramil and Q. Lin, *Chem. Commun.*, 2013, **49**, 11007.

2 (a) I. Chen, M. Howarth, W. Lin and A. Y. Ting, *Nat. Methods*, 2005, **2**, 99; (b) I. S. Carrico, B. L. Carlson and C. R. Bertozzi, *Nat. Chem. Biol.*, 2007, **3**, 321; (c) T. P. King, S. W. Zhao and T. Lam, *Biochemistry*, 1986, **25**, 5774; (d) Lara K. Mahal, Kevin J. Yarema and C. R. Bertozzi, *Science*, 1997, **276**, 1125; (e) Y. Zeng, T. N. Ramya, A. Dirksen, P. E. Dawson and J. C. Paulson, *Nat. Methods*, 2009, **6**, 207; (f) P. Agarwal, R. Kudirka, A. E. Albers, R. M. Barfield, G. W. de Hart, P. M. Drake, L. C. Jones and D. Rabuka, *Bioconjug. Chem.*, 2013, **24**, 846.

3 (a) J. M. Gilmore, R. A. Scheck, A. P. Esser Kahn, N. Joshi and M. B. Francis, *Angew. Chem. Int. Ed.*, 2006, **47**, 7788; (b) M. Dedeo, K. Duderstadt, J. M. Berger and M. B. Francis, *Nano letters*, 2009, **10**, 181; (c) L. S. M. Witus, Troy,Thuronyi, Benjamin W.Esser-Kahn, Aaron P., R. A. Scheck, A. T. Iavarone and M. B. Francis, *J. Am. Chem. Soc.*, 2010, **132**, 16812; (d) L. S. Witus and M. B. Francis, *Acc. Chem. Res.*, 2011, **44**, 774; (e) L. S. Witus, C. Netirojjanakul, K. S. Palla, E. M. Muehl, C. H. Weng, A. T. Iavarone and M. B. Francis, *J. Am. Chem. Soc.*, 2013, **135**, 17223.

4 (a) P. Crisalli and E. T. Kool, *Org. Lett.*, 2013, **15**, 1646; (b) D. K. Kolmel and E. T. Kool, *Chem. Rev.*, 2017, **117**, 10358; (c) E. T. Kool, P. Crisalli and K. M. Chan, *Org. Lett.*, 2014, **16**, 1454; (d) E. T. Kool, D. H. Park and P. Crisalli, *J. Am. Chem. Soc.*, 2013, **135**, 17663.

5 (a) R. G. Wiegand, *J. Am. Chem. Soc.*, 1956, **78**, 5307; (b) M. A. G. del Vado, J. Donoso, F. Muñoz, G. R. Echevarria and F. G. Blanco, *J. Chem. Soc. Perkin Trans.*, 1987, 445; (c) G. R. Echevarría-Gorostidi, A. Basagoitia, E. Pizarro, R. Goldsmid, J. G. S. Blanco and F. G. Blanco, *HeIv. Chim. Acta*, 1998, **81**, 837; (d) G. R. Echevarría, A. Basagoitia, J. G. Santos and F. García Blanco, *J. Mol. Catal. A-Chem.*, 2000, **160**, 209; (e) B. Vilanova, J. M. Gallardo, C. Caldes, M. Adrover, J. Ortega-Castro, F. Munoz and J. Donoso, *J. Phys. Chem. A*, 2012, **116**, 1897.

6 K. C. Hsien, H. T. Chen, Y. C. Chen, Y. L. Chen, C. Y. Lu and C. L. Kao, *Org. Lett.*, 2009, **11**, 3526.

7 X. Wang and J. W. Canary, *Bioconjug. Chem.*, 2012, **23**, 2329.

8 A. W. Garrison and C. E. Boozer, *J. Am. Chem. Soc.*, 1968, **90**, 3486.

9 (a) E. Groman, Y. Z. Huang, T. Watanabe and E. E. SNELL, *Proc. Natl. Acad. Sci.*, **69**, 3297; (b) K. Y. Jung, J. H. Cho, J. S. Lee, H. J. Kim and Y. C. Kim, *Bioorg. Med. Chem.*, 2013, **21**, 2643; (c) N. Katunuma, A. Matsui, T. Inubushi, E. Murata, H. Kakegawa, Y. Ohba, D. Turk, V. Turk, Y. Tada and T. Asao, *Biochem. Biophys. Res. Commun.*, 2000, **267**, 850.

10 M. A. Vázquez, F. Munoz, J. Donoso and F. Blanco, García, *Amino Acids*, 1992, **3**, 81.

11 M. A. Vazquez, G. Echevarria, F. Munoz, J. Donoso and F. G. Blanco, *J. Chem. Soc. Perkin Trans.*, 1989, **2**, 1617.

12 B. Vilanova, J. M. Gallardo, C. Caldés, M. Adrover, J. Ortega-Castro, F. Muñoz and J. Donoso, *J. Phys. Chem. A*, 2012, **116**, 1897.

13 A. Hoegl, M. B. Nodwell, V. C. Kirsch, N. C. Bach, M. Pfanzelt, M. Stahl, S. Schneider and S. A. Sieber, *Nat. Chem.*, 2018, **10**, 1234.

14 C. Iwata and D. E. Metzler, *J. Heterocycl. Chem.*, 1967, **4**, 319.

15 (a) S. Egli, M. G. Nussbaumer, V. Balasubramanian, M. Chami, N. Bruns, C. Palivan and W. Meier, *J. Am. Chem. Soc.*, 2011, **133**, 4476; (b) S. Fredriksson, W. Dixon, H. Ji, A. C. Koong, M. Mindrinos and R. W. Davis, *Nat. Methods*, 2007, **4**, 327.

16 D. Anouk, D. Sjoerd, T. M. Hackeng and P. E. Dawson, *J. Am. Chem. Soc.*, 2006, **128**, 15602.

17 E. H. Cordes and W. P. Jencks, *Biochemistry*, 1962, **1**, 773.

18 (a) N. E. Mbua, J. Guo, M. A. Wolfert, R. Steet and G. J. Boons, *Chembiochem*, 2011, **12**, 1912; (b) M. F. Debets, S. S. van Berkel, S. Schoffelen, F. P. Rutjes, J. C. van Hest and F. L. van Delft, *Chem. Commun.*, 2010, **46**, 97.

19 W. Li, J. Zou, S. Zhu, X. Mao, H. Tian and X. Wang, *Chem. Eur. J.*, 2019, **25**, 10328.

20 (a) Y. M. Li, Y. T. Li, M. Pan, X. Q. Kong, Y. C. Huang, Z. Y. Hong and L. Liu, *Angew. Chem. Int. Ed.*, 2014, **53**, 2198; (b) Y. Xu, L. Xu, Y. Xia, C. J. Guan, Q. X. Guo, Y. Fu, C. Wang and Y. M. Li, *Chem. Commun.*, 2015, **51**, 13189.