Assembling Branched and Macrocyclic Peptides on Proteins

Sven Ullrich,[a,+] Santhanalaxmi Kumaresan,[a,+] Marina G. Rahman,[a] Bishvanwesha Panda,[a] Richard Morewood,[a] and Christoph Nitsche*[,a]

Abstract: A two-step, biocompatible strategy enables site-specific generation of branched and macrocyclic peptide–protein conjugates. Surface-exposed cysteines on proteins are modified by a small bifunctional reagent at near-physiological pH, followed by cyanopyridine–aminothiol click reactions to create branched or macrocyclic peptide architectures. This method offers design strategies for next-generation protein therapeutics.

Inspired by the clinical success of classical antibodies,[1-3] the search for next-generation protein-derived therapeutics is advancing, exploring various types of engineered peptides and proteins.[4-14]

Protein grafting is an increasingly prominent approach to generate such a type of 'neobiologic' (**Fig. 1a**).[14, 15] High-affinity binding epitopes are stabilised in protein scaffolds when they are introduced by mutagenesis, which allows combination with genetically encoded peptide libraries to source these sequences.[16-18] The strategy, however, is often restricted to certain loops within the protein of interest and remains mainly limited to canonical amino acids.[14]

Bioconjugate chemistry, unlike sequence design, facilitates the attachment of molecules to proteins at any desired site, offering a highly customisable approach.[19] Since the process is post-translational, peptides used for conjugation are not limited to canonical amino acids, and can create branched structures extending beyond the primary polypeptide sequence. Strategies to attach shorter peptides to proteins, particularly to the amino acid side chains as opposed to the termini, remain scarce.[20-22]

Bridging both strategies, we aimed to create a method for the site-specific attachment of peptides on the surface of proteins (**Fig. 1**). We identified 4-fluoro-2,6-dicyanopyridine (FDCP) as an ideal small molecule for this, owing to its biocompatible dual reactivity, enabling nucleophilic aromatic substitution and the cyanopyridine-aminothiol click reaction with one small reagent of less than 150 Da.[23]

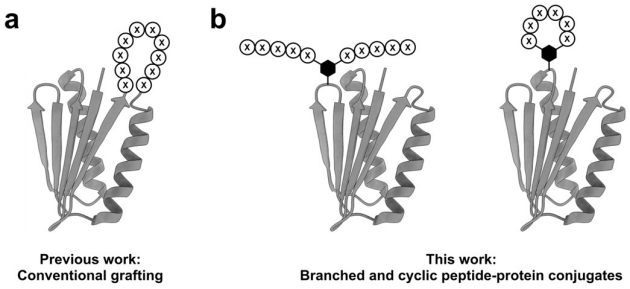

Fig. 1. (a) Conventional molecular grafting, inserting a peptide loop (white) in a protein (grey). (b) Generation of branched and cyclic peptide–protein conjugates facilitated by a small molecular adapter (black) as presented in this study.

We hypothesised that the reagent would enable the generation of branched and macrocyclic peptide–protein conjugates (**Fig. 1b**). Upon reaction, engineered cysteine residues on the protein surface form a peptide-reactive handle with our molecular adapter FDCP. The *in situ* generated dicyanopyridine (DCP) functionality allows for the conjugation of peptides containing 1,2-aminothiol motifs, selectively generating the desired peptide-displaying architectures.

To demonstrate labelling of proteins with FDCP, we modified the sequence of a green fluorescent protein[24] to contain a single surface exposed cysteine residue (GFP-1C, **Tab. S1**). This enabled FDCP to react site-selectively through nucleophilic aromatic substitution, generating a non-canonical dicyanopyridine amino acid with reactivity towards 1,2-aminothiols *in situ* (**Fig. 2a**). To maintain biocompatibility and prevent cysteine oxidation, all reactions were conducted in aqueous buffer containing reducing agent.

Exploring a wide range of conditions, GFP-1C was incubated with FDCP, trialling different pH environments, equivalents, incubation times and temperatures (**Fig. 2b**, **Tab. S5**). Reaction progress was monitored using intact protein mass spectrometry (**Fig. 2c**, **Fig. S8**). The synthesis of GFP-DCP was most efficient with 15 equivalents of FDCP, with the reaction generally requiring overnight incubation for completion. Room temperature and near-physiological pH promoted the formation of a single product with selective cysteine modification by FDCP.

[a] *Research School of Chemistry, Australian National University, Canberra 2601 ACT, Australia. E-mail: christoph.nitsche@anu.edu.au.*

[+] These authors contributed equally: Sven Ullrich, Santhanalaxmi Kumaresan

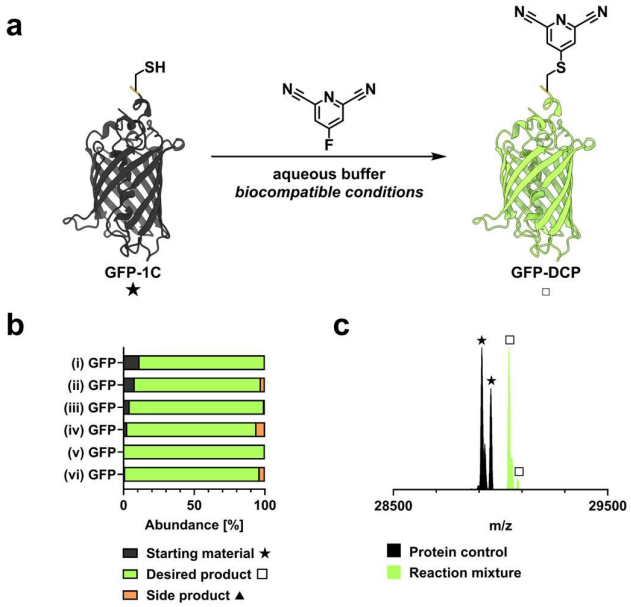

Fig. 2. (a) Reaction of FDCP with GFP-1C (black) yields GFP-DCP (green). (b) Percent abundance of protein starting material (black), desired product (green) and side product (orange) after exposure to FDCP under various conditions: (i) 50 mM Tris-HCl, pH 7.5, 2 mM TCEP, 15 eq. FDCP, 24 h, r.t.; (ii) as (i) with 300 mM NaCl; (iii) as (i) using 30 eq. FDCP; (iv) as (i) at 37 °C; (v) as (i) for 72 h, 4 °C; (vi) 50 mM HEPES-KOH, pH 8.0, with other components as (i). (c) Intact protein mass spectrometry overlay showing spectra of protein control (black) and reaction mixture (green) of condition (v). Peaks representing starting material (★) and desired product (□) are indicated, with a +42 Da adduct presumably deriving from post-translational modification during recombinant expression in *Escherichia coli*.

The reaction did generally not proceed in slightly acidic pH but showed tolerance to higher salt concentration in near physiological (7.5) or slightly basic (8.0) pH environments, suggesting compatibility with proteins requiring saline buffers for stabilisation. Notably, the reaction also proceeds at 4 °C, which is advantageous for temperature-sensitive proteins, albeit more slowly. While the tagging reaction showed promising conversion to the DCP-tagged protein under various conditions, 15 equivalents of FDCP for 24 h in near-physiological pH was deemed the most time- and material-efficient condition for FDCP-tagging with broad applicability.

The tagging method was successfully applied to other engineered proteins (**Tab. S1**), comprising Zika virus protease (ZiP-1C),[25] ubiquitin (Ubq-1C),[26] and the B1 domain of protein G (GB1-1C),[27] where similar reactivity trends with FDCP were observed and reactions to the modified protein proceeded efficiently (**Fig. 3**, **Fig. S9–S11**). After the desired products (GFP-DCP, ZiP-DCP, Ubq-DCP, GB1-DCP) were formed in sufficient quantity, excess FDCP was removed by buffer-exchange using centrifugal filters or spin desalting columns.

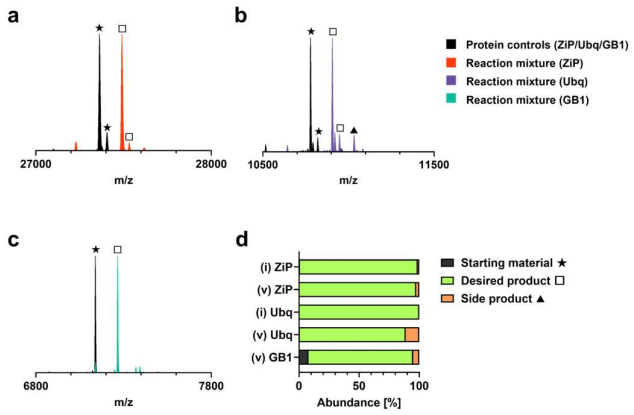

Fig. 3. (a–c) Intact protein mass spectrometry overlays showing spectra of protein control (black) and (a) ZiP (red), (b) Ubq (purple), and (c) GB1 (blue) reaction mixtures of condition (v). Peaks representing starting material (★) and desired product (□) are indicated, including +42 Da adducts. (d) Percent abundance of protein starting material (black), desired product (green) and side product (orange) after exposure to FDCP under the conditions (i) or (v).

After obtaining various DCP-tagged proteins, we investigated the attachment of peptides *via* the cyanopyridine-aminothiol click reaction.[28] We therefore obtained seven synthetic peptides (**Tab. 1**) containing N-terminal cysteine, covering all major classes of amino acids, including non-canonical ones. Exposure of Ubq-1C (**Fig. 4**) and GB1-1C (**Fig. S14**) to excess peptides in aqueous reducing buffer at near-physiological pH resulted in the formation of novel branched peptide-protein architectures (**Fig. 4a**).

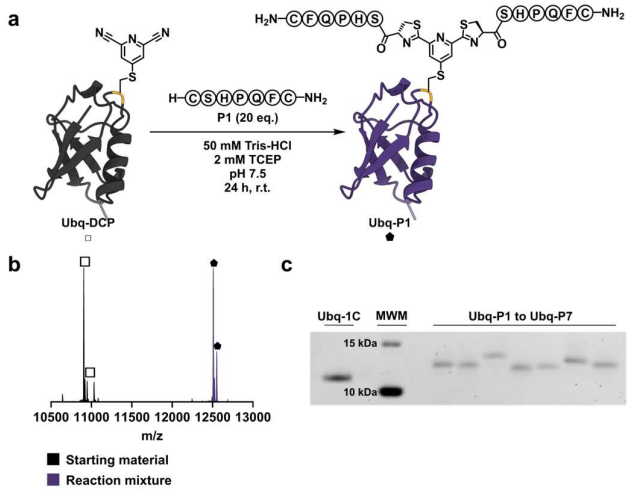

Fig. 4. (a) Reaction of Ubq-DCP (black) with N-terminal cysteine containing peptide P1 yields Ubq–P1 (purple). (b) Intact protein mass spectrometry overlay showing spectra of Ubq-DCP starting material (black) and reaction mixture producing Ubq–P1 (purple). Peaks representing Ubq-DCP (□) and branched peptide–protein conjugate Ubq-P1 (⬟) are indicated, including +42 Da adducts. (c) Excerpt of SDS-PAGE gel showing retention shift between Ubq-1C and Ubq–P1 to Ubq–P7, including molecular weight marker (MWM).

Tab. 1. Peptides and their sequences used in this study.

No.	Sequence
P1	H-Cys-Ser-His-Pro-Gln-Phe-Cys-NH_2
P2	H-Cys-Ala-Tyr-Thr-Asn-Cys-Gly-NH_2
P3	H-Cys-Gly-Lys-Arg-Lys-Ser-Cys-Phe-NH_2
P4	H-Cys-Ser-Asp-Glu-Val-Cys-Trp-NH_2
P5	H-Cys-His-Tyr-Leu-Cys-NH_2
P6	H-Cys-Gly-Ser-Gly-Tyr-Gly-Ser-Gly-Cys-NH_2
P7	H-(D)Cys-Pro-Pra-Ser-(D)Tyr-Cys-Ala-Lys-NH_2
P8	H-Dab(Cys)-Arg-Lys-Lys-Arg-Dab(Cys)-NH_2

Pra, L-propargylglycine; Dab, L-2,4-diaminobutyric acid.

Conjugates Ubq–P1 to Ubq–P7 and GB1–P1 to GB1–P7, respectively, contain two peptide units attached at the site of the former cysteine, forming spontaneously *via* the cyanopyridine–nitrile click reaction. The reactions with DCP-tagged proteins were highly selective and quantitative, with no observable side products in intact protein mass spectrometry (**Fig. 4b**, **Fig. S14**). Attachment of the peptides on Ubq-DCP could also be inferred from SDS-PAGE (**Fig. 4c**). Notably, peptides P1–P7 contain an additional cysteine residue in their sequence, which may be used for further bioconjugation. We demonstrate this with the addition of 1,3-bisbromomethylbenzene (BBMB), a frequently used peptide stapling agent, where the free cysteines appear partially cross-linked (**Fig. S10**). Additionally, due to the synthetic nature of the peptides, it is possible to introduce atypical stereochemistry and unnatural amino acids (P7) into the conjugate, which is challenging for mutagenesis-based grafting methods.

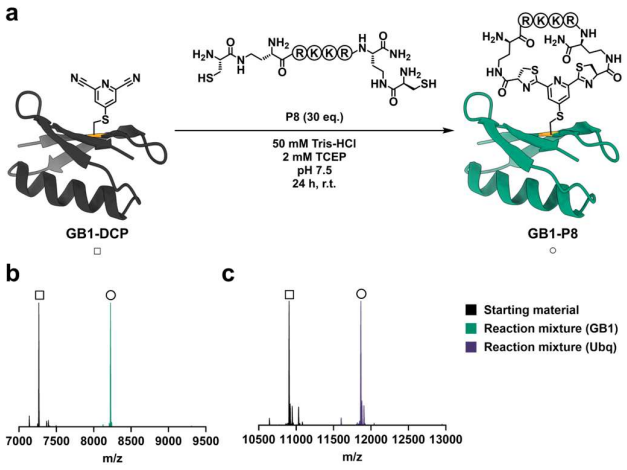

Fig. 5. (a) Reaction of GB1-DCP (black) with peptide P8, containing two 1,2-aminothiols, yields GB1–P8 (blue). (b–c) Intact protein mass spectrometry overlays showing spectra of protein-DCP starting materials (black) and reaction mixtures producing (b) GB1–P8 (blue) and (c) Ubq–P8 (purple). Peaks representing protein-DCP (□) and cyclic peptide–protein conjugates (○) are indicated.

We further showcased the versatility of DCP-tagged proteins by selectively synthesising macrocyclic peptide-protein conjugates (**Fig. 5a**). A peptide containing two 1,2-aminothiol functionalities (P8) in the form of a non-canonical pseudo-N-terminal cysteine amino acid[29] was used to achieve cyclisation of linear peptide P8 on the protein surface of both Ubq-DCP and GB1-DCP (**Fig. 5b** & **5c**). The reaction proceeded in the same aqueous environment as the bioconjugation of linear peptides in presence of excess of the peptide P8, resulting in near-quantitative conversion to the two respective macrocyclic peptide–protein conjugates. The selective generation of cyclic peptides using the cyanopyridine–aminothiol reaction directly on proteins represents a unique synthetic approach that may provide an alternative to lasso-grafting. While the loops of grafted proteins emulate the constraint of a cyclic peptides, the presented method generates a true cyclic peptide on the protein surface.

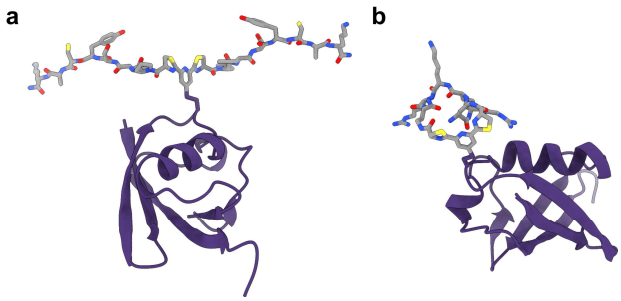

Fig. 6. Computer-generated visualisations of (a) branched Ubq–P1 and (b) cyclic Ubq–P8, highlighting the relative size difference between the attached peptide and protein and the connectivity at the former cysteine site.

This study establishes a versatile platform for the site-specific conjugation of peptides onto proteins using FDCP. Our method enables selective connection between side-chain cysteines and N-terminal cysteines in peptides and proteins. While accommodating non-canonical modifications, our strategy does not depend on them, allowing for the direct conjugation of natural peptide and protein substrates.

The successful synthesis of cyclic and branched peptide–protein conjugates (**Fig. 6**) poses an alternative to lasso-grafting techniques to incorporate or display a peptide on a protein scaffold. Pairing this concept with bioconjugate chemistry generated custom peptide-protein conjugates with potential applications in therapeutic or diagnostic settings. The presented method lays a strong foundation for the development of next-generation protein therapeutics with advanced branched and cyclic architectures.

CN acknowledges the Australian Research Council for a Future Fellowship (FT220100010) and a Discovery Project (DP230100079). SU is supported by a Feodor Lynen Research Fellowship (Alexander von Humboldt Foundation). The authors thank Dr. Mahawaththa and Prof. Otting (Australian National University) for helpful discussions.

Conflicts of interest

There are no conflicts to declare.

Data availability

The data supporting this article have been included as part of the Supplementary Information.

References

1 R.-M. Lu, Y.-C. Hwang, I. J. Liu, C.-C. Lee, H.-Z. Tsai, H.-J. Li and H.-C. Wu, *J. Biomed. Sci.*, 2020, **27**, 1.

2 S. B. Ebrahimi and D. Samanta, *Nat. Commun.*, 2023, **14**, 2411.

3 S. Silpa, X. Janet Bertilla and S. Rupachandra, in *Protein-based Therapeutics*, 2023, DOI: 10.1007/978-981-19-8249-1_11, ch. 11, pp. 297–324.

4 D. G. Johns, L.-C. Campeau, P. Banka, A. Bautmans, T. Bueters, E. Bianchi, D. Branca, P. G. Bulger, I. Crevecoeur, F.-X. Ding, R. M. Garbaccio, E. D. Guetschow, Y. Guo, S. N. Ha, J. M. Johnston, H. Josien, E. A. Kauh, K. A. Koeplinger, J. T. Kuethe, E. Lai, C. L. Lanning, A. Y. H. Lee, L. Li, A. G. Nair, E. A. O'Neill, S. A. Stoch, D. A. Thaisrivongs, T. J. Tucker, P. Vachal, K. van Dyck, F. P. Vanhoutte, B. Volckaert, D. G. Wolford, A. Xu, T. Zhao, D. Zhou, S. Zhou, X. Zhu, H. J. Zokian, A. M. Walji and H. B. Wood, *Circulation*, 2023, **148**, 144–158.

5 V. Guerlavais, T. K. Sawyer, L. Carvajal, Y. S. Chang, B. Graves, J.-G. Ren, D. Sutton, K. A. Olson, K. Packman, K. Darlak, C. Elkin, E. Feyfant, K. Kesavan, P. Gangurde, L. T. Vassilev, H. M. Nash, V. Vukovic, M. Aivado and D. A. Annis, *J. Med. Chem.*, 2023, **66**, 9401–9417.

6 C. Lebon, S. Grossmann, G. Mann, F. Lindner, A. Koide, S. Koide, A. Diepold and O. Hantschel, *Cell Commun. Signal.*, 2024, **22**, 500.

7 P. Ghosh, L. J. Davies and C. Nitsche, *Angew. Chem. Int. Ed.*, 2024, DOI: 10.1002/anie.202419455, e202419455.

8 B. D. Ellenbroek, J. P. Kahler, D. Arella, C. Lin, W. Jespers, E. A. K. Züger, M. Drukker and S. J. Pomplun, *Angew. Chem. Int. Ed.*, 2024, DOI: 10.1002/anie.202416082, e202416082.

9 W.-H. Chen, A. Hajduczki, E. J. Martinez, H. Bai, H. Matz, T. M. Hill, E. Lewitus, W. C. Chang, L. Dawit, C. E. Peterson, P. A. Rees, A. B. Ajayi, E. S. Golub, I. Swafford, V. Dussupt, S. David, S. V. Mayer, S. Soman, C. Kuklis, C. Corbitt, J. King, M. Choe, R. S. Sankhala, P. V. Thomas, M. Zemil, L. Wieczorek, T. Hart, D. Duso, L. Kummer, L. Yan, S. L. Sterling, E. D. Laing, C. C. Broder, J. K. Williams, E. Davidson, B. J. Doranz, S. J. Krebs, V. R. Polonis, D. Paquin-Proulx, M. Rolland, W. W. Reiley, G. D. Gromowski, K. Modjarrad, H. Dooley and M. G. Joyce, *Nat. Commun.*, 2023, **14**, 580.

10 A. Roy, L. Shi, A. Chang, X. Dong, A. Fernandez, J. C. Kraft, J. Li, V. Q. Le, R. V. Winegar, G. M. Cherf, D. Slocum, P. D. Poulson, G. E. Casper, M. L. Vallecillo-Zúniga, J. C. Valdoz, M. C. Miranda, H. Bai, Y. Kipnis, A. Olshefsky, T. Priya, L. Carter, R. Ravichandran, C. M. Chow, M. R. Johnson, S. Cheng, M. Smith, C. Overed-Sayer, D. K. Finch, D. Lowe, A. K. Bera, G. Matute-Bello, T. P. Birkland, F. DiMaio, G. Raghu, J. R. Cochran, L. J. Stewart, M. G. Campbell, P. M. Van Ry, T. Springer and D. Baker, *Nat. Commun.*, 2023, **14**, 5660.

11 G. H. Hutchins, S. Kiehstaller, P. Poc, A. H. Lewis, J. Oh, R. Sadighi, N. M. Pearce, M. Ibrahim, I. Drienovská, A. M. Rijs, S. Neubacher, S. Hennig and T. N. Grossmann, *Chem*, 2024, **10**, 615–627.

12 S. J. de Veer, A. M. White and D. J. Craik, *Angew. Chem. Int. Ed.*, 2020, **60**, 8050–8071.

13 D. J. Craik and J. Du, *Curr. Opin. Chem. Biol.*, 2017, **38**, 8–16.

14 M. Imai, K. Colas and H. Suga, *ChemPlusChem*, 2024, **89**, e202400152.

15 A. A. Komar, *Molecules*, 2023, **28**, 2383.

16 Y. Komatsu, N. Terasaka, K. Sakai, E. Mihara, R. Wakabayashi, K. Matsumoto, D. Hilvert, J. Takagi and H. Suga, *iScience*, 2021, **24**, 103302.

17 E. Mihara, S. Watanabe, N. K. Bashiruddin, N. Nakamura, K. Matoba, Y. Sano, R. Maini, Y. Yin, K. Sakai, T. Arimori, K. Matsumoto, H. Suga and J. Takagi, *Nat. Commun.*, 2021, **12**, 1543.

18 K. Sakai, N. Sugano-Nakamura, E. Mihara, N. M. Rojas-Chaverra, S. Watanabe, H. Sato, R. Imamura, D. C.-C. Voon, I. Sakai, C. Yamasaki, C. Tateno, M. Shibata, H. Suga, J. Takagi and K. Matsumoto, *Nat. Biomed. Eng.*, 2022, **7**, 164–176.

19 P. Chauhan, R. V, M. Kumar, R. Molla, S. D. Mishra, S. Basa and V. Rai, *Chem. Soc. Rev.*, 2024, **53**, 380–449.

20 A. Okon, J. Yang, J. B. Giancola, O. J. Molina, J. Sayers, K. M. Cheah, Y. Li, E. R. Strieter and R. T. Raines, *Bioconj. Chem.*, 2024, **35**, 954–962.

21 A. C. Conibear, E. E. Watson, R. J. Payne and C. F. W. Becker, *Chem. Soc. Rev.*, 2018, **47**, 9046–9068.

22 H. Hojo, *Curr. Opin. Struct. Biol.*, 2014, **26**, 16–23.

23 S. Ullrich, J. George, A. E. Coram, R. Morewood and C. Nitsche, *Angew. Chem. Int. Ed.*, 2022, **61**, e202208400.

24 S. Nagasundarapandian, L. Merkel, N. Budisa, R. Govindan, N. Ayyadurai, S. Sriram, H. Yun and S. G. Lee, *ChemBioChem*, 2010, **11**, 2521–2524.

25 W. Becker, L. A. Adams, B. Graham, G. E. Wagner, K. Zangger, G. Otting and C. Nitsche, *J. Biomol. NMR*, 2018, **70**, 211–218.

26 M. C. Mahawaththa, H. W. Orton, I. Adekoya, T. Huber, G. Otting and C. Nitsche, *Chem. Commun.*, 2022, **58**, 701–704.

27 V. Vitali, K. Ackermann, G. Hagelueken and B. E. Bode, *Appl. Magn. Reson.*, 2023, **55**, 187–205.

28 C. Nitsche, *Synlett*, 2024, **35**, 1067–1071.

29 R. Morewood and C. Nitsche, *Chem. Sci.*, 2021, **12**, 669–674.