

Chemical Science

rsc.li/chemical-science



ISSN 2041-6539



EDGE ARTICLE

Chuanliu Wu *et al.*

De novo design of constrained and sequence-independent peptide scaffolds with topologically-formidable disulfide connectivities

Cite this: *Chem. Sci.*, 2018, 9, 569

De novo design of constrained and sequence-independent peptide scaffolds with topologically-formidable disulfide connectivities†

Yiwu Zheng, Xiaoting Meng, Yaqi Wu, Yibing Zhao and Chuanliu Wu *

Disulfide-rich peptides are interesting scaffolds for drug design and discovery. However, peptide scaffolds constrained by disulfide bonds, either naturally occurring or computationally designed, have been suffering from the elusive (oxidative) folding behavior complying with Anfinsen's dogma, which strongly restricts their applicability in bioactive peptide design and discovery; because when primary peptide sequences are extensively manipulated, their disulfide connectivities might become scrambled. Here we present the design of cysteine/penicillamine (C/Pen)-mixed peptide frameworks that are capable of folding into specific regioisomers without dependence on primary amino acid sequences. Even certain folds that are considered to be topologically formidable can be generated in high yields. Currently, almost all disulfide-rich peptide scaffolds are vitally correlated to primary amino acid sequences, but ours are exceptional. These scaffolds should be of particular interest for further designing constrained peptides with new structures and functions, and more importantly, the ultimately designed peptides would not suffer from general oxidative folding problems.

Received 9th September 2017
Accepted 17th November 2017

DOI: 10.1039/c7sc03956e

rsc.li/chemical-science

Introduction

Constrained peptides—lying between larger biologics and small molecules in size, and presumably combining the advantages of both—have been a treasured chemical space for drug developments.^{1,2} Among the diverse pharmacologically active peptides constrained with covalent crosslinks, naturally occurring disulfide-rich peptides, including plant-derived cyclotides, antimicrobial defensins, and conotoxins from venoms of predatory marine snails, are most extensively explored.^{2–5} Novel bioactive disulfide-constrained peptides can be routinely developed by re-engineering these naturally occurring scaffolds using loop grafting and high-throughput sequence selection.^{6,7} However, there are limits for the structural variety of naturally occurring constrained peptides, which thus restricts the design of inhibitors to certain targets with a surface topology congenitally complementary to the constrained peptide scaffolds.⁸ A strategy to break through this limitation comes from the recent advances in *de novo* protein design, based on which disulfide-constrained peptide scaffolds with new structures can be created without reference to known structures.^{8,9} Moreover, the

computational methods enable the design of constrained peptide structures with non-canonical backbones through incorporating unnatural amino acids. However, for both the naturally occurring and *de novo* designed disulfide-rich peptides, the correct pairing of disulfide bonds and the formation of native structure is primarily driven by the amino acid sequences according to Anfinsen's dogma.^{2,4,10} Thus, existing peptide scaffolds are, in principle, not compatible with the design of sequence-randomized libraries, where the sequences are extensively manipulated and the disulfide pairing may become scrambled. Therefore, it is certainly worth developing constrained and sequence-independent peptide scaffolds with an inherently high propensity of forming correct disulfide connectivities. Such scaffolds should be ideal topologically-fixed backbones amenable to sequence randomization, and if combining with the computational design,⁸ disulfide-constrained peptides with new structures and functions would be created *de novo*.

Previous reports have demonstrated the unique effect of CXC (cysteine-*any*-cysteine) motifs and the incorporation of penicillamine (Pen) on the oxidative folding of either natural or synthetic peptides.^{11–13} However, the essentials of directing the folding of peptide into specific isomers with up to three disulfide bonds have not yet been grasped. We hypothesized that the interplay between CXC motifs and Pen residues in peptides can be rationally manipulated to direct the folding of C/Pen-rich peptides into specific folds with three or four disulfide bonds without dependence on amino acid sequences. Thus, in this work we describe our effort of *de novo* designing C/Pen-mixed

The MOE Key Laboratory of Spectrochemical Analysis and Instrumentation, State Key Laboratory of Physical Chemistry of Solid Surfaces, Department of Chemistry, College of Chemistry and Chemical Engineering, Xiamen University, Xiamen, 361005, P. R. China. E-mail: chlwu@xmu.edu.cn

† Electronic supplementary information (ESI) available: Detailed experimental section, characterization of peptides, tryptic digestion HPLC and MS analysis of disulfide pairing in peptides (Fig. S1–S50). See DOI: 10.1039/c7sc03956e





Fig. 2 Amino acid sequences of peptides 1–5 (from N- to C-terminus). Chromatograms showing the oxidation of 1–5 (10 μ M) in 0.2 mM GSSG (100 mM phosphate buffer, pH 7.4) (black line: before the oxidation; red line: 2 h after oxidative folding). C-to-Pen substitutions favor the folding of several expected folds and disfavor the folding of many others (see Fig. 1a). Disulfide pairings were established by tryptic digestion HPLC and mass spectrometry (Fig. S1–S14[†]).

or inseparable mixtures). Tryptic digestion analysis of the oxidizing products using HPLC and mass spectrometry indicated, as expected, that cysteine residues only paired heterogeneously with Pen residues and the formation of ring-closing CXPen or PenXC was not observed. Thus, in principle our designs to these peptides (2–5) should enable the highly specific formation of a subset of isomers with disulfide connectivities fully covering all possible isomers that might form from the model peptide 1 (Fig. 1). However, further analysis of the folded isomers suggests that two desired folds with disulfide bonds in a 1–4, 2–5, 3–6 (*i.e.*, knotted or &) and 1–6, 2–5, 3–4 (*i.e.*, ladder or #) arrangement were not obtained, which can be considered as two of the most compact folds in topology forming from six-cysteine peptides.^{5,14} Other isomers can be obtained in different yields likely depending on their conformational entropies, with the least compact fold with a disulfide arrangement of 1–2, 3–4, 5–6 (*i.e.*, bead-like fold) most efficiently formed (Fig. 2; \sim 80% conversion). It is worth mentioning that the oxidative folding of these C/Pen-mixed peptides into specific isomers is certainly not driven by

primary sequences because all peptides are composed of primarily achiral glycine residues; and some other lysine and tryptophan residues are strategically inserted for facilitating the tryptic digestion analyses. Our strategy to favor the folding of several specific isomers and yet disfavor many other undesired ones without the recourse to manipulation of primary sequences should be of particular interest to *de novo* design of proteins, particularly considering that the position of CXC motif, the length of each peptide segment, and the cysteine residues for Pen substitution might all be manipulated arbitrarily.

Then, we developed a strategy to direct the oxidative folding of peptides into specific isomers with topologically-formidable disulfide connectivities, including the two missing folds not obtained from the folding of 2–5. This strategy takes advantage of the high propensity of disulfide pairing between two CXC motifs in peptide described previously.¹¹ Though the overall yield of the pairing of two CXC motifs in six-cysteine peptides is relatively low (\sim 5–20%),¹¹ here we demonstrate for the first time that the pairing of two CXPen and/or PenXC motifs in peptides (Fig. 1b) can achieve good to excellent yields (up to \sim 100% in some cases), which enables the highly specific formation of topologically-formidable folds in high yields (*e.g.*, the folds with a ladder or a knotted disulfide connectivity). Peptides 6–10 were designed and synthesized (Fig. 3), into which a pair of CXPen and/or PenXC motifs are incorporated adjacently (9 and 10), alternately (6), or distantly at the N- and C-terminal regions (7 and 8). For these peptides, the total number of isomers formed after oxidative folding can be reduced, in principle, to a minimum of three. Our results clearly show that oxidation of 6 yielded the expected fold with a knotted disulfide connectivity (6-&) in \sim 38% conversion as one of the three isomers (Fig. 3; totally \sim 100% conversion); 7 and 8 can be oxidized to the expected folds with the ladder disulfide connectivity (7-#) and a connectivity of 1–5, 2–6, 3–4 in \sim 100% and \sim 80% conversion, respectively (Fig. 3). In addition, the specific disulfide pairings can still be preserved in the oxidation of 9 and 10, though in these peptides the relative distance between the two pivotal motifs (*i.e.*, CXPen or PenXC) is obviously shorter compared to that in other peptides (6–8), a factor that can significantly disfavor the dimeric disulfide pairing of CXC motifs.¹¹ We can still obtain the expected folds in \sim 49% and \sim 85% conversion after the oxidative folding of 9 and 10, respectively, and importantly the fold with a connectivity of 1–6, 2–4, 3–5 from 9 are highly constrained in structure. For all of the peptides examined in this study, only \leq 3 isomers were detected after the folding, and the disfavored folds against the specified C/Pen pairing rules were not observed.

The folds with their three disulfide bonds in a knotted arrangement are one of the most interesting scaffolds with exceptional stability and resistance to proteolytic degradation for peptide-based drug design.^{5,15} However, the *de novo* designed 6-& can only be obtained in a moderate yield (\sim 38%). After analyzing side products forming from the oxidation of 6 (Fig. 3), we realized that the low yield of 6-& is very likely caused by the relative easiness of forming one of the other two side-products that is obviously less compact in conformation; that





Fig. 3 Amino acid sequences of peptides 6–11 (from N- to C-terminus). Chromatograms showing the oxidation of 6–11 (10 μ M) in 0.2 mM GSSG (100 mM phosphate buffer, pH 7.4) (black line: before oxidation; red line: 2 h after oxidative folding). Topological drawings of the expected folds were given in bottom panels (red lines denote disulfide bonds; amino acid sequences can be read in a counter-clockwise direction). The disulfide connectivity of peaks a and b shown in the chromatograms (by products): 6-a = (1–6, 2–5, 3–4), 6-b = (1–4, 2–3, 5–6); 8-a = (1–5, 2–3, 4–6), 8-b = (1–4, 2–6, 3–5); 9-a = (1–5, 2–4, 3–6), 9-b = (1–2, 3–5, 4–6); 10-a = (1–6, 2–3, 4–5); 11-a = (1–4, 2–6, 3–5), 11-b = (1–3, 2–5, 4–6). Disulfide pairings were established by tryptic digestion HPLC and mass spectrometry (Fig. S15–S36[†]). Of note, a pair of conformers formed from the folding of 9, which have identical disulfide connectivity (1–6, 2–4, 3–5).

is, the fold with a disulfide connectivity of 1–4, 2–3, 5–6 (or termed as “tandem fold”). Accordingly, to further increase the yield of the knotted fold, we strategically switched the position of the CXPen and PenXC motif in 6 to design and synthesize a new peptide 11, which should retain the capability of forming a knotted fold (11- $\&$), but without hindered by forming the structurally-incompact tandem or bead-like folds, based on the specified disulfide pairing rules. Interestingly, after oxidation, the yield of 11- $\&$ was increased to 57%, obviously higher than that observed for 6 (Fig. 3). This result also demonstrated the feasibility of increasing the yield of desired folds through further optimizing the C/Pen pattern of peptides to disfavor competing folds.

We further demonstrated that both 7- $\#$ and 11- $\&$ are tolerant to the grafting of bioactive peptide sequences (*i.e.*, bioactive epitopes). While a yeast-selected sequence (PRPRGDNPLT; capable of binding to the cell-surface integrins) was grafted into the scaffolds (peptides 12 and 13; the other two peptide segments was shortened by two residues for the ease of synthesis),⁶ the yields of the oxidative folding into the desired isomers (12- $\#$ and 13- $\&$) were only reduced moderately (\sim 63% and \sim 70%, respectively; Fig. S37–S43[†]). In addition, both peptides can block the adhesion of U87 glioblastoma cells (with surface-expressed integrins¹⁶) to cell culture plates (Fig. 4), indicating the integrin-binding capability.

Thus far, the tolerance of C/Pen-bearing peptide scaffolds to sequence randomizations has never been evaluated. The following designs and experiments will unambiguously demonstrate that even the peptide scaffolds (or folds) with the topologically-formidable disulfide connectivities (*i.e.*, 7- $\#$ and

12: WGCKPenGKGCPRPRGDNPLTPenGKGCCKPenGW
 13: WGPenKCGKGCPRPRGDNPLTCKPenGKGPenGW

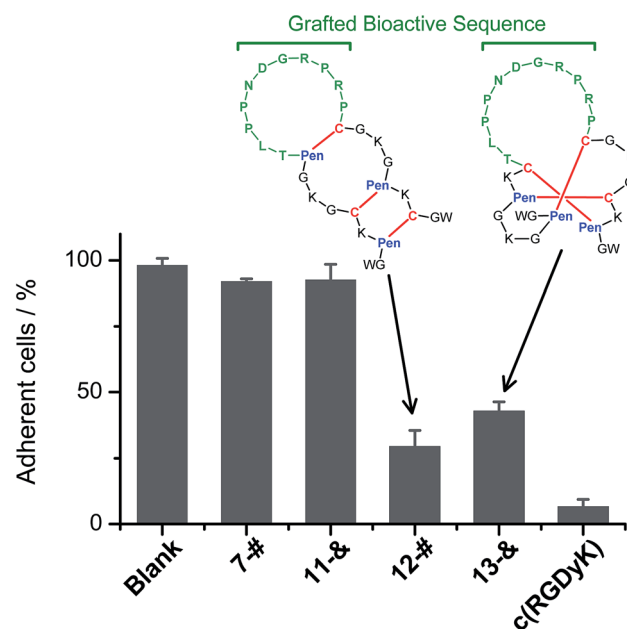


Fig. 4 Amino acid sequences of peptides 12 and 13 (from N- to C-terminus); bioactive sequence grafted into the scaffolds was highlighted in green. The inhibition of U87MG cell adhesion by the peptides (7- $\#$, 11- $\&$, 12- $\#$, 13- $\&$, and c(RGDyK); 50 μ M/3 h) determined by MTT assays. 7- $\#$ and 11- $\&$ are negative controls (without the grafted bioactive sequence), which exhibit a negligible ability of inhibiting cell adhesion; c(RGDyK) is commercially available cyclic RGD, which was served as a positive control. The error bars represent the standard deviation of the mean ($n = 3$).



including Energy Chemical (Shanghai, China), Sigma-Aldrich (Shanghai, China), Alfa Aesar (Beijing, China), and TCI (Shanghai, China).

Synthesis of peptides

All peptides were synthesized using solid-phase peptide synthesis (SPPS) and native chemical ligation methods according to reported procedures (detailed steps were described in the ESI, Pages S3–S5†).¹⁸ The peptides were purified by analytical and semi-preparative RP-HPLC and characterized using mass spectrometry (mass spectra were given in the ESI, Pages S93–S98†).

Oxidative folding of peptides

Fully reduced peptides were dissolved in 100 mM phosphate buffer (pH 7.4) containing 0.2 mM GSSG to achieve a concentration of $\sim 10 \mu\text{M}$. Two hours later, aliquots were taken into an empty tube and were immediately treated with 10% HPO_3 to quench the oxidative reactions. The samples were then analyzed using an analytical HPLC (HPLC gradient for the analysis was given in the ESI†).

Analysis of disulfide pairing in peptides

Certain amounts of oxidized peptides (isolated by HPLC, $\sim 30 \mu\text{M}$) were dissolved in 90 μL phosphate buffer (100 mM, pH = 7.4), which were then digested by the addition of 10 μL aqueous solution of trypsin (1 mg mL^{-1}) at room temperature (~ 0.5 h). The digested fragments were then analyzed by HPLC and MS.

Integrin-dependent cell adhesion assays

Cell adhesion assays were evaluated by the MTT assay reported previously in our lab.¹³ U87MG cells were plated into 24-well plates at an initial cell density of 80 000 cells per well and grown overnight at 37 °C, 5% CO_2 . After 24 h, the medium was removed and the peptides were added to each well in 300 μL DMEM contained FBS, incubated at 37 °C, 5% CO_2 for 3 h. After that, the medium was removed and cells were washed three times with DMEM to remove the detached cells. Then 300 μL fresh DMEM and 30 μL MTT (50 mg mL^{-1}) was added to each well. The cells were incubated for 4 h at 37 °C in culture hood. After that, remove the supernate and add 300 μL MTT solvent (DMSO). Cover with tinfoil and agitate cells on orbital shaker for 15 min. The absorbance was then measured at 490 nm using a plate reader (PerkinElmer Enspire®). The obtained absorbance was blank-corrected (blank: DMEM + MTT, no cells) and the cell viability in percent was calculated according to the following equation:

$$\text{Cell viability} = (\text{OD}_{490,\text{sample}}/\text{OD}_{490,\text{control}}) \times 100\%$$

where $\text{OD}_{490,\text{sample}}$ and $\text{OD}_{490,\text{control}}$ represent the optical density of the cells treated with peptides and that of the cells only treated with DMEM, respectively.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We would like to acknowledge the financial support from the National Basic Research Program of China (2014CB932004), the National Natural Science Foundation of China (21375110 and 21475109), the Program for Changjiang Scholars and Innovative Research Team in University (IRT13036), and the Foundation for Innovative Research Groups of the National Natural Science Foundation of China (21521004).

References

- 1 C. J. White and A. K. Yudin, *Nat. Chem.*, 2011, **3**, 509–524; T. A. Hill, N. E. Shepherd, F. Diness and D. P. Fairlie, *Angew. Chem., Int. Ed.*, 2014, **53**, 13020–13041; V. Azzarito, K. Long, N. S. Murphy and A. J. Wilson, *Nat. Chem.*, 2013, **5**, 161–173; C. Heinis, T. Rutherford, S. Freund and G. Winter, *Nat. Chem. Biol.*, 2009, **5**, 502–507; L. D. Walensky, A. L. Kung, I. Escher, T. J. Malia, S. Barbuto, R. D. Wright, G. Wagner, G. L. Verdine and S. J. Korsmeyer, *Science*, 2004, **305**, 1466–1470; J. D. Osteen, V. Herzig, J. Gilchrist, J. J. Emrick, C. Zhang, X. Wang, J. Castro, S. Garcia-Caraballo, L. Grundy, G. Y. Rychkov, A. D. Weyer, Z. Dekan, E. A. B. Undheim, P. Alewood, C. L. Stucky, S. M. Brierley, A. I. Basbaum, F. Bosmans, G. F. King and D. Julius, *Nature*, 2016, **534**, 494–499.
- 2 K. B. Akondi, M. Muttenthaler, S. Dutertre, Q. Kaas, D. J. Craik, R. J. Lewis and P. F. Alewood, *Chem. Rev.*, 2014, **114**, 5815–5847.
- 3 M. E. Klotman and T. L. Chang, *Nat. Rev. Immunol.*, 2006, **6**, 447–456; Y. Guo, D.-M. Sun, F.-L. Wang, Y. He, L. Liu and C.-L. Tian, *Angew. Chem., Int. Ed.*, 2015, **54**, 14276–14281; A. J. Wommack, J. J. Ziarek, J. Tomaras, H. R. Chileveru, Y. Zhang, G. Wagner and E. M. Nolan, *J. Am. Chem. Soc.*, 2014, **136**, 13494–13497.
- 4 M. Gongora-Benitez, J. Tulla-Puche and F. Albericio, *Chem. Rev.*, 2014, **114**, 901–926.
- 5 N. L. Daly and D. J. Craik, *Curr. Opin. Chem. Biol.*, 2011, **15**, 362–368.
- 6 N. Cox, J. R. Kintzing, M. Smith, G. A. Grant and J. R. Cochran, *Angew. Chem., Int. Ed.*, 2016, **55**, 9894–9897.
- 7 S. Chen, I. R. Rebollo, S. A. Buth, J. Morales-Sanfrutos, J. Touati, P. G. Leiman and C. Heinis, *J. Am. Chem. Soc.*, 2013, **135**, 6562–6569; C. K. Wang, C. W. Gruber, M. Cemazar, C. Siatskas, P. Tagore, N. Payne, G. Sun, S. Wang, C. C. Bernard and D. J. Craik, *ACS Chem. Biol.*, 2014, **9**, 156–163.
- 8 G. Bhardwaj, V. K. Mulligan, C. D. Bahl, J. M. Gilmore, P. J. Harvey, O. Cheneval, G. W. Buchko, S. V. S. R. K. Pulavarti, Q. Kaas, A. Eletsy, P.-S. Huang, W. A. Johnsen, P. Greisen Jr, G. J. Rocklin, Y. Song, T. W. Linsky, A. Watkins, S. A. Rettie, X. Xu, L. P. Carter, R. Bonneau, J. M. Olson, E. Coutasias, C. E. Correnti,



- T. Szyperski, D. J. Craik and D. Baker, *Nature*, 2016, **538**, 329–335.
- 9 P.-S. Huang, S. E. Boyken and D. Baker, *Nature*, 2016, **537**, 320–327.
- 10 B. S. Mainathambika and J. C. Bardwell, *Annu. Rev. Cell Dev. Biol.*, 2008, **24**, 211–235; D. J. Craik, *Nat. Chem.*, 2012, **4**, 600–602; N. Metanis and D. Hilvert, *Angew. Chem., Int. Ed.*, 2012, **51**, 5585–5588; A. Walewska, M.-M. Zhang, J. J. Skalicky, D. Yoshikami, B. M. Olivera and G. Bulaj, *Angew. Chem., Int. Ed.*, 2009, **48**, 2221–2224.
- 11 C. L. Wu, J. C. Leroux and M. A. Gauthier, *Nat. Chem.*, 2012, **4**, 1045–1050.
- 12 Y. W. Zheng, L. X. Zhai, Y. B. Zhao and C. L. Wu, *J. Am. Chem. Soc.*, 2015, **137**, 15094–15097; J. T. Hunt, V. G. Lee, E. C. K. Liu, S. Moreland, D. McMullen, M. L. Webb and M. Bolgar, *Int. J. Pept. Protein Res.*, 1993, **42**, 249–258; M. Tabachnick, H. N. Eisen and B. Levine, *Nature*, 1954, **175**, 701–702; K. Ramalingam and G. H. Snyder, *Biochemistry*, 1993, **32**, 11155–11161; H. Tamaoki, R. Miura, M. Kusunoki, Y. Kyogoku, Y. Kobayashi and L. Moroder, *Protein Eng.*, 1998, **11**, 649–659.
- 13 Y. W. Zheng, Z. R. Li, J. Ren, W. D. Liu, Y. Q. Wu, Y. B. Zhao and C. L. Wu, *Chem. Sci.*, 2017, **8**, 2547–2552.
- 14 A. C. Conibear and D. J. Craik, *Angew. Chem., Int. Ed.*, 2014, **53**, 10612–10623.
- 15 A. S. Ravipati, S. T. Henriques, A. G. Poth, Q. Kaas, C. K. Wang, M. L. Colgrave and D. J. Craik, *ACS Chem. Biol.*, 2015, **10**, 2491–2500; C. T. T. Wong, D. K. Rowlands, C.-H. Wong, T. W. C. Lo, G. K. T. Nguyen, H.-Y. Li and J. P. Tam, *Angew. Chem., Int. Ed.*, 2012, **51**, 5620–5624; B. H. Kwan, E. F. Zhu, A. Tzeng, H. R. Sugito, A. A. Eltahir, B. Ma, M. K. Delaney, P. A. Murphy, M. J. Kauke, A. Angelini, N. Momin, N. K. Mehta, A. M. Maragh, R. O. Hynes, G. Dranoff, J. R. Cochran and K. D. Wittrup, *J. Exp. Med.*, 2017, **214**, 1679–1690.
- 16 A. P. Silverman, A. M. Levin, J. L. Lahti and J. R. Cochran, *J. Mol. Biol.*, 2009, **385**, 1064–1075.
- 17 H. Heguri, T. Yamamoto and Y. Tezuka, *Angew. Chem., Int. Ed.*, 2015, **54**, 8688–8692; B. V. K. J. Schmidt, N. Fechner, J. Falkenhagen and J.-F. Lutz, *Nat. Chem.*, 2011, **3**, 234–238; J.-F. Lutz, J.-M. Lehn, E. W. Meijer and K. Matyjaszewski, *Nat. Rev. Mater.*, 2016, **1**, 1–14; T. Terashima, T. Mes, T. F. A. De Greef, M. A. J. Gillissen, P. Besenius, A. R. A. Palmans and E. W. Meijer, *J. Am. Chem. Soc.*, 2011, **133**, 4742–4745.
- 18 G. M. Fang, J. X. Wang and L. Liu, *Angew. Chem., Int. Ed.*, 2012, **51**, 10347–10350.

