Synthesis of bioactive and stabilized cyclic peptides by macrocyclization using C(sp^3)–H activation

Jian Tang, Yadong He, Hongfei Chen, Wangjian Sheng and Huan Wang

Cyclic peptides have attracted increasing attention in recent years due to their ability to inhibit protein–protein interactions. Current strategies to prepare cyclic peptides often rely on functional amino acid side chains or the incorporation of unnatural amino acids, thus limiting their structural diversity. Here, we describe the development of a highly versatile peptide macrocyclization strategy through a palladium-catalyzed C(sp^3)–H activation and the synthesis of cyclic peptides featuring unique hydrocarbon linkages between the β-carbon of amino acids and the aromatic side chains of Phe and Trp. We demonstrate that such peptides exhibit improved biological properties compared to their acyclic counterparts. Finally, we applied this method in the synthesis of the natural product celogentin C.

Introduction

Peptides and peptidomimetics are attracting increasing attention in both the chemical and pharmaceutical communities as a rich source of drug candidates and/or biological tools. Since the clinical employment of gramicidin S, hundreds of peptides have been identified from natural sources, and a number of them have advanced to the clinic. One major driving force for the growing interest in these substances is their potency to interrupt or control biological events mediated by protein–protein interactions (PPIs), which often involve binding contacts spread over a surface. However, native peptides generally suffer from poor pharmacological properties and lack of structural diversity. To address these limitations, various peptide modification strategies have been developed, among which macrocyclization is one of the most prominent methods. Compared to their acyclic counterparts, cyclic peptides generally have highly defined conformations, enhanced cell permeability, thermostability and resistance to proteolytic degradation.

Classic methods to generate cyclic peptides often follow a one-component strategy, including head-to-tail lactamization, internal disulfide formation or thioether formation, ring-closing olefin metathesis (RCM) and catalyzed cycloaddition of azides to alkynes (Fig. 1a). A number of two-component cyclization procedures rely on external stapling agents, which usually utilize nucleophiles from lysine and cysteine residues. As a representative example, macrocyclization between lysine residues through nitrogen arylation by a perfluoroaryl linker is reported by Pentelute and co-workers. Recently, several three-component macrocyclization methods were reported by Yudin and co-workers in the synthesis of aziridine- and oxadiazole-containing peptide macrocycles. Despite the success achieved in this field, limitations remain to be overcome. Peptides shorter than five amino acids are generally difficult to cyclize by lactamization and often lead to oligomerization. For RCM and azide/alkyne cycloaddition, incorporation of a pair of unnatural amino acids at specific sites is usually required, making the preparation of precursor peptides challenging. The use of external stapling agents helps expand the structural diversity of cyclic peptides; however,
only amino acids with reactive side chains, such as lysine and cysteine (relative abundance ≈ 1%), can be involved in these cyclization reactions. Such requirement excludes amino acids with chemically inert aliphatic side chains, which are highly abundant in naturally occurring peptides/proteins, as macrocyclization sites and thus significantly limits the diversity of peptidomimics generated from these protocols. Therefore, strategies that could activate and involve these “inert” amino acids (e.g. at their β positions) in macrocyclization are highly desirable.

C-C coupling assisted by metal catalysis through direct C–H activation has now become a fundamental strategy to introduce functionalities to inert C–H bonds and has found applications in the synthesis of cyclic peptide natural products. In our pursuit to generate peptidomimics with novel architecture for biological studies, we envisioned that peptide macrocyclization by activation of the β-carbon C–H bond of amino acids would help overcome the requirement for reactive amino acid side chains, since almost all amino acids contain a β-carbon (Gly as the only exception). Recently, Yu and co-workers reported their inspiring work in the Pd-catalyzed intermolecular arylation of a C(sp3)–H bond of N-terminal amino acid in oligopeptides without the installation of directing groups (Fig. 1b). We postulated that this method would allow peptide macrocyclization through intramolecular arylation with β-carbons of amino acids as stapling sites (Fig. 1c).

In this work, we synthesized a variety of cyclic peptides containing unique Cβ–Ar crosslinks between the β-carbon of amino acids (e.g., Ala, Val) and the aromatic ring of Phe/Trp through a Pd-catalyzed C(sp3)–H activation process. This method is highly versatile and compatible with the standard Fmoc SPPS procedure, resulting in cyclic peptides of different sizes. The impact of peptide conformation on the efficiency of macrocyclization is also investigated. Moreover, we demonstrated that peptidomimics generated from this method possess improved stability compared to their linear peptide precursors. In biological studies, cyclic RGD peptide generated by this method exhibited potent binding affinity to integrin. We also successfully applied this methodology to the synthesis of a peptide natural product containing β-carbon-to-Trp indole linkages.

Results and discussion

Investigation of Pd-catalyzed macrocyclization

To apply the Pd-catalyzed β-carbon arylation method to peptide macrocyclization, we started our initial attempt on a tetrapeptide 1a with the sequence of Phth-Ala-Gly-Gly-(m-I-Phe)-OMe (Phth = phthaloyl, m-I-Phe = meta-iodophenylalanine), where the N- and C-termini of the peptide were protected by phthaloyl and methyl group, respectively (Fig. 2, Scheme S1†). This sequence was chosen because it contains minimal side chains and m-I-Phe provides direct access to peptides containing modified phenylalanine moieties. The installation of protecting groups was meant to minimize the undesired coordination to the Pd(u) catalyst. After optimization of reaction conditions (Table S1†), it was found that with 1a (0.05 mmol), Pd(OAc)2 (5 mol%) as the catalyst, AgOAc (1.2 equiv.) as the additive, in dichloroethane (DCE) at 100 °C, the macrocyclization proceeded smoothly, and the desired product 2a was isolated in 62% yield. The resulting cyclic peptide 2a was characterized by analytical HPLC, high-resolution mass spectrometry (HRMS) and NMR analysis (Fig. 2 and S1†). After cyclization, the Ala βH peak shifted significantly (Fig. 2a) in the 1H NMR spectrum, and the newly formed macrocyclic linkage was further assigned by a diagnostic heteronuclear multiple bond correlation (HMBC) between the Ala βH and the Phe C3 (Fig. 2b). No dimerization products were observed by either HPLC or MS analysis.

With optimized reaction conditions, we next evaluated the efficiency of this macrocyclization strategy on a variety of tetrapeptide sequences. The incorporation of aliphatic amino acids at position 2 (the second residue to the N-terminus), including Leu, Ile and β-branched Val (2b–2d), did not interrupt the cyclization process, and all afforded cyclic peptides in good isolated yields (53–58%, Table 1). In addition, the cross-linking occurred solely at the β-carbon of the N-terminal Ala,
Table 1  Sequences of precursor peptides and resulting cyclic peptides synthesized in this work

<table>
<thead>
<tr>
<th>Entry</th>
<th>Linear peptide</th>
<th>Cyclic peptide</th>
<th>Isolated yield (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>Phth-Ala-Gly-Gly-m-I-Phe-OMe</td>
<td>2a</td>
<td>62</td>
</tr>
<tr>
<td>1b</td>
<td>Phth-Ala-Leu-Gly-m-I-Phe-OMe</td>
<td>2b</td>
<td>55</td>
</tr>
<tr>
<td>1c</td>
<td>Phth-Ala-Val-Gly-m-I-Phe-OMe</td>
<td>2c</td>
<td>58</td>
</tr>
<tr>
<td>1d</td>
<td>Phth-Ala-Ile-Gly-m-I-Phe-OMe</td>
<td>2d</td>
<td>53</td>
</tr>
<tr>
<td>1e</td>
<td>Phth-Ala-Phe-Gly-m-I-Phe-OMe</td>
<td>2e</td>
<td>58</td>
</tr>
<tr>
<td>1f</td>
<td>Phth-Ala-Gly-Gly-Ala-m-I-Phe-OMe</td>
<td>2f</td>
<td>62</td>
</tr>
<tr>
<td>1g</td>
<td>Phth-Aib-Gly-m-I-Phe-OMe</td>
<td>2g</td>
<td>52</td>
</tr>
<tr>
<td>1h</td>
<td>Phth-Phe-Gly-Gly-m-I-Phe-OMe</td>
<td>2h</td>
<td>19</td>
</tr>
<tr>
<td>1i</td>
<td>Phth-Ala-Ser(Bu)-Gly-m-I-Phe-OMe</td>
<td>2i</td>
<td>60</td>
</tr>
<tr>
<td>1j</td>
<td>Phth-Ala-Lys(Boc)-Gly-m-I-Phe-OMe</td>
<td>2j</td>
<td>65</td>
</tr>
<tr>
<td>1k</td>
<td>Phth-Ala-Asp(OMe)-Gly-m-I-Phe-OMe</td>
<td>2k</td>
<td>65</td>
</tr>
<tr>
<td>1l</td>
<td>Phth-Ala-Gly-m-I-Phe-Gly-Resin</td>
<td>2l</td>
<td>21</td>
</tr>
<tr>
<td>1m</td>
<td>Phth-Ala-Gly-Pro-m-I-Phe-OMe</td>
<td>2m</td>
<td>87</td>
</tr>
<tr>
<td>1n</td>
<td>Phth-Ala-Gly-Pro-Ala-m-I-Phe-OMe</td>
<td>2n</td>
<td>63</td>
</tr>
<tr>
<td>1o</td>
<td>Phth-Ala-Gly-Pro-Ala-p-I-Phe-OMe</td>
<td>2o</td>
<td>50</td>
</tr>
<tr>
<td>1p</td>
<td>Phth-Ala-Arg(Pbf)-Gly-Asp(OMe)-I-Phe-OMe</td>
<td>2p</td>
<td>23</td>
</tr>
</tbody>
</table>

highlighting the regioselectivity of this methodology (see ESI† for structural characterization). Phe-containing peptide 1e and pentapeptide 1f cyclized efficiently under these conditions as well (Table 1).

Next, we examined peptides with N-terminal residues other than Ala. Linear peptides 1g and 1h with N-terminal 2-aminooisobutyric acid (Aib) and Phe, respectively, were therefore synthesized and subjected to macrocyclization under standard conditions. Results showed that the C–H activation of βH (Aib) in peptide 1g was successful by yielding cyclic peptide 2g in 52% yield. The crosslink in 2g showed low diastereoselectivity (d.r. 60 : 40), as determined by 1H NMR and LC-MS analysis (Fig. S2a†). On the other hand, the activation of βH (Phe) in peptide 1h was much less efficient and resulted in a cyclization yield of only 19%, whereas a significant amount of deiodination product was observed by LC-MS, indicating the challenge of crosslinking at the βC of a bulky Phe residue. In contrast to diastereoselectivity observed in peptide 2g, NMR and LC-MS analysis indicated that the resulting cyclic peptide 2h was a single diastereomer (Fig. S2b†). The vicinal coupling between Hα and Hβ of Phe1 of the cyclic peptide 2h was determined to be 12.0 Hz by 1H NMR, suggesting an antiplanar arrangement of these two protons and therefore an R configuration of Phe1 Cβ in peptide 2h.

To explore the feasibility of performing peptide macrocyclization in solid phase, we first inserted Ser, Lys and Asp residues with side chains masked by standard Fmoc SPPS protecting groups (1i–1k). Results showed that these bulky protecting groups at position 2 were compatible with cyclization conditions without lowering the macrocyclization yields (Table 1). Encouraged by these results, using 2-chlorotriyl chloride resin and following standard Fmoc SPPS procedure, we synthesized peptide 1i with the sequence resin-Gly-(m-I-Phe)-Gly-Gly-Ala-Phth (Table 1 and Scheme S2†). On-resin macrocyclization was achieved with Pd catalyst in DMF at 100 °C for 16 h. The resulting cyclic peptide was cleaved from the resin by TFA and purified by RP-HPLC. NMR and HRMS analysis confirmed that cyclic peptide 2i was successfully synthesized; however, the isolation yield was lower than solution-phase reactions (Table 1). The synthesis of 2i was the first example in which I-Phe is not positioned at the C-terminus of the oligopeptide during cyclization, demonstrating the potential of such method in building in-sequence rings. Interestingly, when para- and ortho-I-Phe was incorporated at the C-terminus of tetrapeptides, none were cyclized successfully, and only deiodination products were detected by MS analysis (Table S2†). Overall, Pd-catalyzed coupling through C(sp³)–H activation was shown to be highly robust and efficient in the preparation of cyclic peptides with novel Cβ–Ar crosslinks (Scheme S3†).

The fact that tetrapeptides with para- or ortho-I-Phe could not be cyclized prompted us to investigate the principle underlying these observations. We reasoned that the cyclic tetrapeptides resulting from macrocyclization were highly constrained, and therefore, conformational organization of linear precursor peptides must allow two reactive ends to coordinate with the catalyst in close spatial proximity before ring closure. We then synthesized peptide 1m with a Pro residue at position 3, which introduced a conformational turn into the linear precursor tetrapeptide. To our delight, the cyclization of peptide 1m proceeded extremely well with almost full conversion as monitored by HPLC with an isolation yield of 87%, indicating the impact of peptide conformation on cyclization efficiency.
containing pentapeptide 1n also cyclized efficiently. Interestingly, pentapeptide 1o (Table 1), which contains a Pro residue and a para-1-Phe, cyclized smoothly with a yield of 50% (Scheme S3†). This result suggested that the failure of tetrapeptides containing of para- or ortho-1-Phe to cyclize was most likely due to conformational bias, instead of inherent reactivity. The insertion of a Pro residue, a conformational turn-inducing element, might relieve the constraint of the cyclic pentapeptide and facilitate Cβ–Ar crosslinking.

Biological evaluation of cyclic peptides

To evaluate the potential of the Cβ–Ar cyclized peptides for biological applications, we synthesized a novel cyclic RGD peptide 2p following this procedure (Fig. 3). We first examined the proteolytic stability of 2p in comparison with its linear counterpart 1p. Peptides 1p and 2p were incubated with endoprotease AspN, which hydrolyzes peptide bonds on the N-terminal side of Asp residue. Results showed that in the presence of AspN, cyclic peptide 2p remained intact for up to 5 h, whereas its linear counterpart 1p degraded quickly with a half-life of 50 min and was fully consumed in 5 h, as determined by HPLC and MS analysis (Fig. 3). Similar results were observed when peptides 1j and 2j, which were resulted from deprotected peptides 1f and 2f, were incubated with protease trypsin (Fig. 3), where peptide 1j had a half-life of only 21 min, and up to 95% cyclic peptide 2j remained intact after 5 h. These results demonstrate that the Cβ–Ar linkage in cyclic peptides significantly improved their stability against proteolytic degradation.

Next, we examined the binding of cyclic RGD peptide 2p to U87MG cells, which is a glioblastoma cell line overexpressing the αvβ3 integrin. The RGD sequence is well known for its selective binding to integrins when incorporated in proper cyclic structures.29 Fluorescin isothiocyanate (FITC) was then conjugated to the N-terminal amine of 2p to yield fluorescent-labelled 2p-FITC. Linear peptide 1p was also labelled with FITC as a direct comparison. As a positive control, we employed a commercially available cyclo-(RGDFK) and conjugated a FITC moiety onto the Lys amine. U87MG cells were then incubated with 1p-FITC, 2p-FITC and cyclo-(RGDFK)-FITC for 90 min and analyzed by confocal microscopy. As expected, cyclo-(RGDFK)-FITC elicited strong fluorescence staining of cells, whereas the linear peptide conjugate 1p-FITC yielded low fluorescence, suggesting its weak binding to integrins (Fig. 4a and b). Interestingly, cyclic peptide 2p-FITC caused slightly stronger fluorescence staining than cyclo-(RGDFK)-FITC, indicating that the cyclic structure introduced by a Cβ–Ar linkage in 2p significantly enhanced the integrin binding of the RGD sequence. These results nicely demonstrated the applicability of cyclic peptides with Cβ–Ar linkages in biological systems.

Application in natural product synthesis

Finally, we applied this macrocyclization methodology to natural product synthesis. Several peptide natural products
containing Cβ–Ar linkages have been identified in recent years. For example, streptide is a recently characterized ribosomally synthesized and posttranslationally modified peptide (RiPP) featuring a C–C crosslink between Trp C5 and Lys Cβ (Fig. 5a).\textsuperscript{40} Celogentin C, a nonribosomal peptide possessing inhibitory activity against tubulin polymerization, contains a direct linkage of Trp C6 to Leu Cβ (Fig. 5a).\textsuperscript{41} Towards the total synthesis of celogentin C, Chen and co-workers employed a Pd-catalyzed strategy to construct the Leu1-Trp C–C bond with the assistance of a quinolone directing group.\textsuperscript{36,42} We postulated that our C–H activation macrocyclization strategy would allow the construction of the celogentin C-ring A from a precursor peptide containing 6-I-Trp and avoid the introduction and removal of directing groups (Fig. 5c). Accordingly, we synthesized a 6-I-Trp containing precursor peptide 1q (Table 1) and subjected it to macrocyclization.\textsuperscript{36} Gratifyingly, peptide 1q cyclized smoothly under coupling conditions, affording the desired cyclic peptide 2q in 32% isolated yield. Structural characterization of 2q was conducted by HPLC, HRMS and NMR analysis. Product 2q appeared as a sharp single peak in analytical HPLC, suggesting a single diastereoisomer (Fig. S3†). A correlation between Trp C6 and Leu1 Hβ was observed in HMBC analysis, indicating the direct Leu1-Cβ-to-Trp C6 linkage in 2q (Fig. S4†). The large vicinal coupling (J_{ab}(Leu1) = 12.0 Hz) and the absence of correlation in NOESY analysis between Hα and Hβ of Cβ-Leu1 strongly suggested an antiperiplanar arrangement of these two protons (Fig. S5†), meaning that the configuration of Leu1 Cβ in 2q is R and thus identical to the Leu2 Cβ of the natural product celogentin C.\textsuperscript{41} Further derivatizations of 2q yielded compound 2r, the ring A of celogentin C. The $^1$H NMR and $^{13}$C NMR spectra of compound 2r is in very good agreement with the same compound reported by Chen’s group, which further supports the correct configuration of Leu1 Cβ.\textsuperscript{36} This remarkable stereoselective 6-indolation of Leu Cβ in cyclic peptide 2q suggested that analogous to the palladium complex formed with directing groups,\textsuperscript{36} the amide bond N,N-bisected complex generated during peptide cyclization promoted the formation of the trans-palladacycle intermediate (Fig. S6†).\textsuperscript{37} In addition, the successful formation of the βC-Trp C6 linkage further expands the scope of this methodology to iodine-substituted indoles and β-branched amino acids.

Conclusions

In this work, we have developed a robust procedure to generate cyclic peptides with Cβ–Ar linkages through Pd-catalyzed C(sp$^3$)–H activation, which proceeds smoothly in both solution-phase and solid-phase peptide synthesis. This method allows a diverse set of amino acids to serve as peptide stapling sites and therefore could greatly expand the structural diversity of cyclic peptide libraries. This method has a broad scope of peptide substrates, including tetra- and pentapeptides, and it has been demonstrated as a powerful tool in natural product synthesis by introducing a Cβ–Ar bond with high regio- and
stereoselectivity. In addition, we demonstrated that a RGD cyclic peptide generated through this method has improved proteolytic stability and higher binding affinity to integrin than a commercially available cyclo-(RGDFK). The generation of natural products and peptidomimics of biological importance through this strategy is in progress in our laboratory.41

**Acknowledgements**

The authors acknowledge the financial support from the National Natural Science Foundation of China (21402091) and the Thousand Talents Program. We would like to thank Prof. Zhuangzhi Shi (NJU) for helpful discussion at the initial stage of the project and Mr Xiao Xiao for confocal microscopy analysis.

**Notes and references**


Note: while submitting this manuscript, Albericio et al. reported a similar strategy for the preparation of cyclic peptides through C–H activation (*Angew. Chem., Int. Ed.*, 2017, 56, 314–318). Whereas the disclosed data is in accord with our findings, yields are significantly improved under our cyclization conditions. Moreover, the biological properties of the cyclic peptides are evaluated for the first time along with faplication of this methodology in natural product synthesis.