Enzymatic Macrocyclization of 1,2,3-Triazole Peptide Mimetics

Emilia Oueis, Marcel Jaspars, Nicholas J. Westwood, and James H. Naismith*

Abstract: The macrocyclization of linear peptides is very often accompanied by significant improvements in their stability and biological activity. Many strategies are available for their chemical macrocyclization, however, enzyme-mediated methods remain of great interest in terms of synthetic utility. To date, known macrocyclization enzymes have been shown to be active on both peptide and protein substrates. Here we show that the macrocyclization enzyme of the cyanobactin family, PatGmac, is capable of macrocyclizing substrates with one, two, or three 1,4-substituted 1,2,3-triazole moieties. The introduction of non-peptidic scaffolds into macrocycles is highly desirable in tuning the activity and physical properties of peptidic macrocycles. We have isolated and fully characterized nine non-natural triazole-containing cyclic peptides, a further ten molecules are also synthesized. PatGmac has now been shown to be an effective and versatile tool for the ring closure by peptide bond formation.

Macrocycles in general and peptidic macrocycles in particular have recently gained more attention since they have been identified as potential sources for novel therapeutics.[1] These compounds, abundant in nature, are known to possess a wide range of biological activities.[1,2] They owe their attractiveness to having higher structural rigidity and better chemical stability, especially against hydrolysis and peptidase activity in the case of peptide-based macrocycles.[3] The increased molecular weight of these compounds is usually balanced by their increased specificity and selectivity towards the targets compared to “small” molecules.[4] Patellamides are natural cyclic octapeptides that belong to the large family of cyanobactins.[5] Their biosynthesis consists of the modification of a long precursor peptide via the action of different tailoring enzymes to achieve heterocyclization of cysteines, or three-
gained significant attention through the use of the so called stapled peptides, but in this case, the non-peptidic module tends to be chosen to lock the peptide in a specific conformation rather than being a key component for activity.

The ring closure of peptide macrocycles can be achieved either chemically or enzymatically with a number of distinct enzymes known including butelase, POP, Phe-soritis, recombinant asparaginyl endopeptidase (AEP), and thioesterase, and the macrocyclase enzyme (PatGmac) from the patellamide family. PatGmac recognizes a C-terminal AYD(G) motif located outside the core of the precursor peptide substrate and cleaves it to form a water shielded acyl-enzyme intermediate using its catalytic Asp-His-Ser triad. The intermediate is then attacked by the free N-terminal amine of the core peptide. The substrate must have a thiazoline, oxazoline or a proline immediately N-terminal to the (cleaved off) AYD(G) motif with no other restrictions on the rest of the core peptide. This very broad substrate range has made PatGmac an attractive tool for biotransformation (Figure 1). We decided to test its ability to macrocyclize substrates in which non-peptidic motifs are included in the sequence. We have chosen to investigate the incorporation of 1,4-substituted 1,2,3-triazole groups into macrocycles. The triazole ring can act as an isostere for the amine and its synthesis is facile. Further the triazole moiety has been incorporated in many biologically active compounds as it is thermodynamically and physiologically stable. We reasoned therefore this would be a suitable test of the versatility of PatGmac by systematically incorporating the triazole unit (Tz) at different positions of the precursor peptide while varying their number. We report herein the successful formation (catalyzed by PatGmac) of macrocycles containing one, two, and three 1,2,3-triazoles.

A series of 19 different linear precursor peptides based on the core sequence VGAGIGFP were prepared by standard solid phase peptide synthesis using the Fmoc strategy (see Supporting Information, SI). The linear peptides VGAGIGF-PAYD-NH2 consisted of the core sequence modified with one, two, or three non-consecutive triazoles and the PatGmac recognition motif AYD with an N-terminal amide. The triazoles were generated directly on solid phase using a copper(I) catalyzed azide-alkyne cycloaddition (CuAAC). Azido acids 1a,b and 1d were synthesized in one step by diazo transfer reaction (Scheme S3). Fmoc protected amino alkyne 2a-c were synthesized in five steps starting from the commercial Boc protected amino acid. First, the Weinreb amides were synthesized by a standard coupling reaction. The use of DIBAL-H for the amide reduction in our hands was found to yield very little conversion to the aldehyde. Separating the two steps with the reduction to aldehyde using LiAlH4 followed by the Bestmann-Öhira reagent in our hands was found to yield very little conversion to the aldehyde. Stabilization of the diol and protection step afforded the desired alkyne 2a-c in moderate yields overall (Scheme S4). The N-terminal azide was attached on the growing peptide by a standard coupling reaction then was subjected to the CuAAC reaction with the corresponding alkyne counterpart 2 on solid phase (Scheme 1; Scheme S1). Final precursor peptides P1-6 have a single triazole (Tz) at a different position of the core sequence (the number denoting the position), precursors P13-16, P24-26, P35-36 and P46 have two triazoles (i.e. at positions 1 and 3 of the peptide for P13; VTzGATzGIGFPAYD), and P135 and P246 have three triazoles (Table S1; Table S2). P0 has no triazoles and was synthesized to serve as a control for the macrocyclization reaction.

All precursor peptides P were subjected to the macrocyclization reaction conditions in the presence of PatGmac in bicine buffer at pH 8.1 and 37°C. The reaction progress was monitored by MALDI-MS and some of the resulting spectra are shown in Figure 2. All 19 precursor peptides converted to the desired macrocyclic structure and no significant linear product (AYD cleavage product) was detected in the reaction mixture. We observed that two triazole-containing substrates were slower to macrocyclize than single triazole-containing peptides and that both three triazole peptides were the slowest to convert. Nonetheless, because PatGmac is known to be a slow enzyme, the rate difference is insignificant (reactions were monitored 20 to 25 days for full conversion on large scale).

Nine of the precursor peptides, P0, P1, P2, P3, P4, P5, P6, P13 and P135, were reacted at a larger scale (8 μmol, 9–10 mg) in the presence of PatGmac under the previously stated conditions in order to provide sufficient amounts of the macrocyclic product for full characterization. Cyclic peptides CP were extracted and HPLC purified to afford moderate yields of high purity samples. Results are summarized in Scheme 1. Synthesis of triazole-containing cyclic peptides.
For some of these (CP3, CP13, CP135), two peaks on the HPLC had the expected mass of the cyclic peptide in the crude reaction mixture, as can be seen from the ratio column in Table 1 (SI, section VIII). When applicable, both peaks were collected and analyzed by HRMS and MS-MS fragmentation patterns and both were determined to be the desired cyclic peptide (Table S3). However, the NMR experiments were only conducted for the major isolated compound.

The MS-MS fragmentation data of the cyclic peptides gave a very complex series of ions consistent with their peptidic sequence, the newly formed valine to proline bond (macrocycle formation), and the triazole presence (SI, section VII).

NMR experiments were also conducted in order to further confirm their structures. The proton NMR data of the cyclic peptides showed at least two species present at different ratios depending on the compound. For CP0, CP1, CP2, CP3 CP4, CP6 and CP13 this was shown by Exchange Spectroscopy (EXSY) experiments to result from different conformers (SI, section VI). However, compound CP5 was found to have two distinct compounds in a 1:0.6 ratio. The NMR assignments of both species show they have the same peptidic sequence (cyclic VGAGITzGFP). Both HRMS and MS-MS fragmentation data confirm the presence of one mass/one sequence. This led us to believe that the two species are most likely epimers but we cannot exclude structurally rigid conformers unable to interconvert.[26]

CP135 has a very complex proton and 2D NMR spectra that showed at least four different species (with two other minor species). General assignment has been achieved for most proton peaks (SI, section V) however, the individual sequence for each was not determined. EXSY showed that two of them are related conformers and the other two are distinct compounds. Similarly to compound CP5, these species would most probably arise from epimers or from conformationally rigid variants.

Biotransformation is now widely used in organic chemistry for the synthesis of biologically relevant compounds. The PatGmac enzyme is unusual in that it operates on substrates and disposes of almost the entire recognition site during the reaction. PatGmac macrocyclizes a wide range of natural peptidic substrates and tolerates non-natural amino acids.[8a,b] From a chemical diversity point of view, the ability to mix amino acid residues and non-amino acid scaffolds within the backbone of macrocycles is highly desirable. We have now shown that it is possible to synthesize macrocycles with one, two, or three triazoles within a peptidic sequence. This led us to believe that the two species are most likely epimers but we cannot exclude structurally rigid conformers unable to interconvert.[26]

Table 1: PatGmac macrocyclization reaction products.

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Core peptide[a]</th>
<th>Ratio [%]</th>
<th>Yield [%][b]</th>
<th>Purity [%][c]</th>
</tr>
</thead>
<tbody>
<tr>
<td>CP0</td>
<td>VGAGIGFP</td>
<td>100</td>
<td>32</td>
<td>99</td>
</tr>
<tr>
<td>CP1</td>
<td>VTzGAGIGFP</td>
<td>100</td>
<td>43</td>
<td>99</td>
</tr>
<tr>
<td>CP2</td>
<td>VTzGAGIGFP</td>
<td>100</td>
<td>48</td>
<td>91</td>
</tr>
<tr>
<td>CP3</td>
<td>VTzGTzGIGFP</td>
<td>76/24</td>
<td>55</td>
<td>99</td>
</tr>
<tr>
<td>CP4</td>
<td>VGAGTzGFP</td>
<td>100</td>
<td>34</td>
<td>98</td>
</tr>
<tr>
<td>CP5</td>
<td>VGAGITzGFP</td>
<td>100</td>
<td>58</td>
<td>93</td>
</tr>
<tr>
<td>CP6</td>
<td>VGAGITzGFP</td>
<td>100</td>
<td>40</td>
<td>98</td>
</tr>
<tr>
<td>CP13</td>
<td>VTzGTzGIGFP</td>
<td>20/80</td>
<td>54</td>
<td>98</td>
</tr>
<tr>
<td>CP135</td>
<td>VTzGTzGItzGFP</td>
<td>15/85</td>
<td>39</td>
<td>97</td>
</tr>
</tbody>
</table>


Tables 1 and S3 (see Scheme 1 or SI, section I for structures). For some of these (CP3, CP13, CP135), two peaks on the HPLC had the expected mass of the cyclic peptide in the crude reaction mixture, as can be seen from the ratio column in Table 1 (SI, section VIII). When applicable, both peaks were collected and analyzed by HRMS and MS-MS fragmentation patterns and both were determined to be the desired cyclic peptide (Table S3). However, the NMR experiments were only conducted for the major isolated compound.

The MS-MS fragmentation data of the cyclic peptides gave a very complex series of ions consistent with their peptidic sequence, the newly formed valine to proline bond (macrocycle formation), and the triazole presence (SI, section VII).

NMR experiments were also conducted in order to further confirm their structures. The proton NMR data of the cyclic peptides showed at least two species present at different ratios depending on the compound. For CP0, CP1, CP2, CP3 CP4, CP6 and CP13 this was shown by Exchange Spectroscopy (EXSY) experiments to result from different conformers (SI, section VI). However, compound CP5 was found to have two distinct compounds in a 1:0.6 ratio. The NMR assignments of both species show they have the same peptidic sequence (cyclic VGAGITzGFP). Both HRMS and MS-MS fragmentation data confirm the presence of one mass/one sequence. This led us to believe that the two species are most likely epimers but we cannot exclude structurally rigid conformers unable to interconvert.[26]

CP135 has a very complex proton and 2D NMR spectra that showed at least four different species (with two other minor species). General assignment has been achieved for most proton peaks (SI, section V) however, the individual sequence for each was not determined. EXSY showed that two of them are related conformers and the other two are distinct compounds. Similarly to compound CP5, these species would most probably arise from epimers or from conformationally rigid variants.

Biotransformation is now widely used in organic chemistry for the synthesis of biologically relevant compounds. The PatGmac enzyme is unusual in that it operates on substrates and disposes of almost the entire recognition site during the reaction. PatGmac macrocyclizes a wide range of natural peptidic substrates and tolerates non-natural amino acids.[8a,b] From a chemical diversity point of view, the ability to mix amino acid residues and non-amino acid scaffolds within the backbone of macrocycles is highly desirable. We have now shown that it is possible to synthesize macrocycles with one, two, or three triazoles within a peptidic sequence. Thus PatGmac is able to process peptide hybrid substrates extending its utility. The ability to macrocyclise such hybrid molecules highlights the scope for creating highly diverse libraries through split and pool approaches.

Acknowledgements

This work was supported by the European Research Council (339367), UK Biotechnology and Biological Sciences

Figure 2. MALDI-MS traces of the macrocyclization reactions of representative sequences with a) control peptide; VGAGIGFP, 20 days; b) VTzGAGIGFP, 1 triazole, 21 days; c) VTzGTzGIGFP, 2 triazoles, 23 days; d) VTzGTzGItzGFP, 3 triazoles, 25 days.
Research Council (K015508/1), The Wellcome Trust (Triple-TOF 5600 mass spectrometer (094476), the MALDI TOF-TOF Analyser (079272A1A), 700 NMR) and the EPSRC UK National Mass Spectrometry Facility at Swansea University. J.H.N. is a Royal Society Wolfson Merit Award holder and 1000 talent scholar at Sichuan University.

**Keywords:** biotransformation · cyanobactin · cyclic peptides · peptidomimetics · triazole

**How to cite:** Angew. Chem. Int. Ed. 2016, 55, 5842–5845

**Angew. Chem.** 2016, 128, 5936–5939

---


[17] In the rest of the paper, the 1,4-substituted 1,2,3-triazole will be referred to as triazole.


[21] The replacement of an amide bond by a triazole is represented by a “Tz” between the corresponding two amino acids in the sequence.


[23] Reagent 1e is commercially available. Reagent 2e is easily prepared from commercial propargylamine.


[25] The molecule number represents the position of the triazole (Tz) moiety/moieties in the core peptide (VTzG is position 1, GTzF is position 6).


Received: February 13, 2016
Published online: April 5, 2016