RESEARCH FRONT

CSIRO PUBLISHING

Aust. J. Chem. **2015**, *68*, 1365–1372
http://dx.doi.org/10.1071/CH15146

Full Paper

Propargyloxyproline Regio- and Stereoisomers for Click-Conjugation of Peptides: Synthesis and Application in Linear and Cyclic Peptides

Susan E. Northfield,[A] *Simon J. Mountford,*[A] *Jerome Wielens,*[A,B] *Mengjie Liu,*[A] *Lei Zhang,*[C] *Herbert Herzog,*[C] *Nicholas D. Holliday,*[D] *Martin J. Scanlon,*[A] *Michael W. Parker,*[B,E] *David K. Chalmers,*[A] *and Philip E. Thompson*[A,F]

[A]Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, 381 Royal Parade, Parkville, Vic. 3052, Australia.
[B]ACRF Rational Drug Discovery Centre, St Vincent's Institute of Medical Research, Fitzroy, Vic. 3065, Australia.
[C]Neuroscience Research Program, Garvan Institute of Medical Research, St Vincent's Hospital, Darlinghurst, NSW 2010, Australia.
[D]Institute of Cell Signalling, School of Life Sciences, University of Nottingham, Queen's Medical Centre, Nottingham NG7 2UH, UK.
[E]Department of Biochemistry and Molecular Biology, Bio21 Molecular Science and Biotechnology Institute, The University of Melbourne, Parkville, Vic. 3010, Australia.
[F]Corresponding author. Email: philip.thompson@monash.edu

The use of the click reaction for the introduction of conjugate groups, such as affinity or fluorescent labels, to a peptide for the study of peptide biochemistry and pharmacology is widespread. However, the nature and location of substituted 1,2,3-triazoles in peptide sequences may markedly affect conformation or binding as compared with native sequences. We have examined the preparation and application of propargyloxyproline (Pop) residues as a precursor to such peptide conjugates. Pop residues are available in a range of regio- and stereoisomers from hydroxyproline precursors and are readily prepared in Fmoc-protected form. They can be incorporated routinely in peptide synthesis and broadly retain the conformational properties of the parent proline containing peptides. This is exemplified by the preparation of biotin- and fluorophore-labelled peptides derived from linear and cyclic peptides.

Manuscript received: 26 March 2015.
Manuscript accepted: 15 May 2015.
Published online: 24 June 2015.

Introduction

The ability to readily label and/or conjugate peptides is an important facet of biological chemistry research.[1] Such conjugates can be used for tagging bioactive peptides with specific labels to track and identify binding targets, or they can be used to alter physicochemical properties for improved pharmacological activity.[2] The positions at which peptides can be usefully functionalised are critically dependent on what region of the peptide is necessary for activity. In linear peptides, conjugates can be added to either the N- or C-terminus, or from one of the amino acid side-chains, commonly a lysine or tyrosine, but these regions should not be part of the key pharmacophore. Conjugates can also be attached using a range of linking chains, varying in length and/or polarity, to distance them from the bioactive peptide. When dealing with head-to-tail cyclic peptides, the choice of where to introduce conjugates is more limited as there is no free terminus available at which to functionalise, and local structural changes can have a marked impact on cyclic peptide conformation. A position in the peptide must be identified where a conjugate could be incorporated without disrupting the binding or activity of the parent peptide sequence.

Proline provides unique conformational restraint as compared with the other natural amino acids, driven by its cyclic structure and the presence of a tertiary amide bond. Proline is also frequently reported as a tool to induce reverse-turns in cyclic peptides.[3] This function can place the proline residue in a conformation protruding away from the peptide binding/activity site, making proline an attractive residue at which to incorporate functionality (Fig. 1). Proline derivitisation, or 'editing' as it has recently been termed, has been shown to be a diverse approach to peptide derivatisation.[4] Zondlo et al. have described a range of diverse modifications to proline that allow for further derivatisation. This report on an array of substituted prolines notes the influence substitution might have on *cis–trans* isomerisation and even peptide conformation.

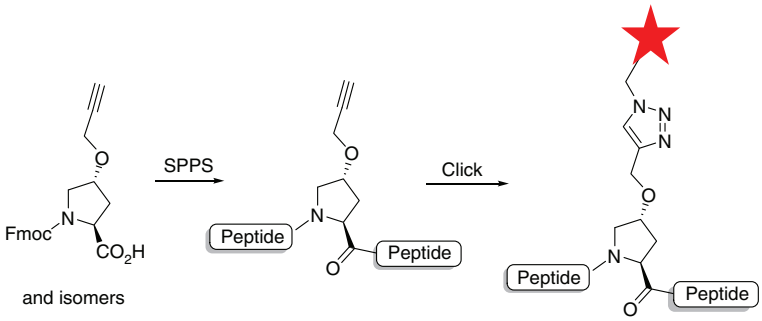


Fig. 1. 'Exposed' proline residues in cyclic or linear peptides can be modified to include 'handles' for side-chain conjugation. Proline residues appended with a propargyloxy handle are incorporated in place of proline in peptides. Labels (star) are introduced using standard conditions for copper-catalysed azide–alkyne conjugation (CuAAC) reactions (SPPS: solid phase peptide synthesis).

Similarly, we envisaged that introduction of a propargyloxy group yielding propargyloxyproline (Pop) residues might fulfil a similar role in the development of peptide conjugates undergoing click reactions. The click reaction is one of the most widely used methods of introducing conjugates into azido or alkyne-containing peptides.[5] The use of click reactions using Pop derivatives as substrates has included the synthesis of macrocyclic peptides[6] and also as a way of immobilising proline as an asymmetric catalyst for aldol reactions.[7] We anticipated using hydroxyproline stereoisomers as precursors, to gain access to a range of isomeric Pop-containing peptides. The resultant triazolylmethoxy conjugates would supply a useful spacer away from the peptide chain, minimising the impact of the conjugate on the peptide structure.

We report here on the synthesis of a range of Fmoc-protected Pop regio- and stereoisomers and their incorporation into two classes of peptide of interest in our laboratories. The first were derivatives of the Y_1 receptor antagonist peptide, BVD15, and the second were cyclic hexapeptides incorporating a Lys–Ile–Asp–Asn (KIDN) pharmacophore motif of lens epithelium-derived growth factor (LEDGF), a key protein for the activity of HIV integrase (IN).

We and others have had an on-going interest in the development of conjugates of peptides that bind to the Y_1 G-protein coupled receptors,[8] and we had reported modification of a proline residue by Ctp (Fig. 2) in the dimeric Y1 antagonist, 1229U91 but with significant loss of activity.[9] Another prominent starting point has been the 10-residue peptide Y_1 antagonist, BVD-15,[10] which has been more amenable to conjugation.[8,9] We decided to investigate the use of conjugates built around a variety of Pop isomers.

In our second application of the building blocks we focussed on a series of cyclic hexapeptides based upon the reverse turn motif of LEDGF, a protein that binds at the dimer interface of IN.[11] The interaction is essential for efficient integration of HIV DNA into the host chromosome, and consequently for successful viral replication.[12] These hexapeptides contain the tetrapeptide sequence Lys–Ile–Asp–Asn (KIDN) linked by a dipeptide scaffold comprised of one or two proline residues to support a reverse-turn pharmacophore at the tetrapeptide portion. In X-ray structures of peptide–IN complexes, the dipeptide scaffold projected away from the protein binding site and was not taking part in any binding interactions. This presented an opportunity for us to incorporate additional functionality to our peptides.

Results and Discussion

Synthesis of Fmoc-Pop Stereoisomers

Fmoc-protected Pop derivatives were prepared from their corresponding hydroxyproline isomers, including the commonly occurring amino acid (2*S*,4*R*)-hydroxyproline (or *trans*-4-hydroxy-L-proline). A selection of isomers were synthesised using a common synthetic strategy (Scheme 1). For example, (2*S*,4*R*)-hydroxyproline (**i**) was Boc-protected (**v**), and then treated with propargyl bromide under basic conditions to give the Boc-(2*S*,4*R*)-Pop (**ix**). Deprotection and then reprotection of Nα gives the key building block Fmoc-(2*S*,4*R*)-4-propargyloxyproline (**xiii**) in 25 % overall yield. Note that the preparation of **ix** was recently reported and the sensitivity of the alkylation to selected base and solvent conditions was highlighted.[13] The three step synthesis could be carried out without purification of intermediates, carrying through the crude products at each step. In the final step, Fmoc-OSu was limited to 1.0 molar equivalent, limiting the formation of Fmoc-β-Alanine-OH,[14] which had proved difficult to separate from **xiii**.

In the same manner, we prepared the *N*-Fmoc-protected *O*-propargyl derivatives of (2*R*,4*R*)-, [**xiv**, *cis*-4-hydroxy-D-proline] (2*S*,4*S*)- (**xv**, *cis*-4-hydroxy-L-proline), and (2*S*,3*R*)-hydroxyproline (**xvi**, *cis*-3-hydroxy-L-proline) stereoisomers from the corresponding building blocks giving a collection of Fmoc-protected amino acids on a gram scale for incorporation into peptides using standard solid phase peptide synthesis (SPPS) protocols (see Supplementary Material).

Synthesis of NPY Analogues

Three parent Y1 antagonists Lys^4-BVD15 (**1**), Arg^4-BVD15 (**2**), and a cyclic peptide, c(Glu^2,Dap^4)-BVD15, **3** and their Pop-containing analogues **4**–**9** were prepared. The synthesis of linear peptides **1**, **2**, and **4**–**7** was achieved using standard Fmoc-based SPPS protocols using Rink Amide resin. Cleavage using trifluoroacetic acid (TFA) yielded the products in good recovery and purity.[9] Peptides **3** and **8** were prepared by solid phase synthesis of the linear Fmoc-precursor and solution phase Glu to Dap cyclisation. Peptide **9** was prepared similarly but with an N-terminal 4-fluorbenzoic acid group. Cyclisation was carried out in DMF (1 mg mL^{-1} peptide) using PyClock (2 equiv.) coupling reagent and N-methylmorpholine (NMM) (12 equiv.) as activating base. We have previously reported the use of PyClock as a cyclisation reagent for the dimeric forms of the peptides, 1229U91,[9] but the formation of the dimeric product is suppressed by using NMM instead of diisopropylethylamine

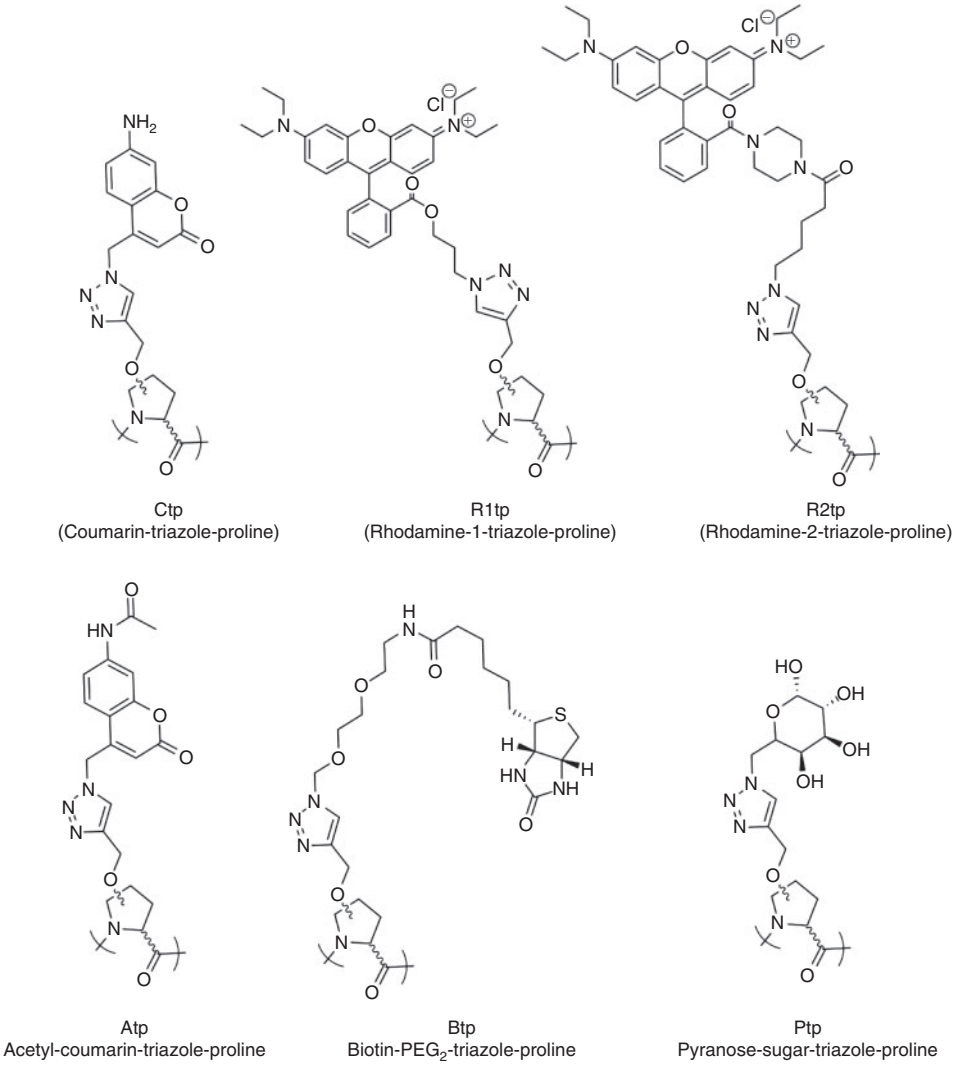
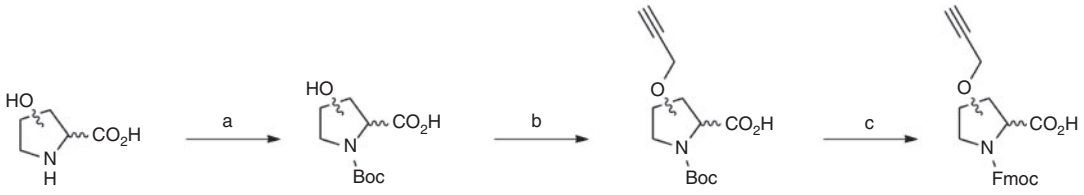


Fig. 2. Structures of labels incorporated in the NPY and LEDGF analogues.



i = 2*S*,4*R*
ii = 2*R*,4*R*
iii = 2*S*,4*S*
iv = 2*S*,3*R*

v = 2*S*,4*R*
vi = 2*R*,4*R*
vii = 2*S*,4*S*
viii = 2*S*,3*R*

ix = 2*S*,4*R*
x = 2*R*,4*R*
xi = 2*S*,4*S*
xii = 2*S*,3*R*

xiii = 2*S*,4*R*
xiv = 2*R*,4*R*
xv = 2*S*,4*S*
xvi = 2*S*,3*R*

Scheme 1. Synthetic scheme for the synthesis of Fmoc-propargyloxyproline residues **xiii**–**xvi**. (a) Boc-anhydride, triethylamine, MeOH reflux overnight; (b) NaH, DMF, propargyl bromide, 0°C, 2 h; (c) 1 : 1 triflouroacetic acid/dichloromethane, room temp 30 min, then Fmoc-OSu, dioxan, 0°C, pH 10, 60 min, room temperature overnight.

(DIPEA) as base. For peptides **3** and **8**, the N-terminal Fmoc group was removed and the peptides precipitated in diethyl ether. The recovered products were used directly for click chemistry and reverse phase (RP)-HPLC purification.

With the precursors in hand, conjugation reactions were performed. Peptides **4**–**7** were conjugated with 7-amino-4-azidomethylcoumarin to yield the products **10**–**15**. Peptide **4** was also conjugated with azido-functionalised rhodamine fluorophores to give **16** and **17** (Fig. 2 and Table 1).

Click reactions were carried out by one of two methods, depending on the solubility of the peptide and azide reagents involved. Reactions were performed at room temperature, in the presence of copper sulfate, sodium ascorbate, and stabilising ligand. Using a DMF solvent system, 1 mg mL^{-1} of peptide was

treated with a 4-fold excess of azide-conjugate including the copper-stablising ligand TBTA (tris[(1-benzyl-1*H*-1,2,3-triazol-4-yl)methyl]amine). Alternately, an aqueous phosphate buffer system could be applied.[15] In this method peptides at a concentration of 20 mg mL^{-1} in phosphate buffer (pH 7) were additionally treated with aminoguanidine hydrochloride and stabilising ligand THPTA (tris(3-hydroxypropyltriazolylmethyl)amine). Coupling reactions were very efficient, and the peptides were subsequently purified by RP-HPLC. While successfully prepared using crude Pop-containing peptides, the complexity of the product mixture suggests that purification of the precursor peptides is advisable.

Analysis of NPY Analogues

The labelled products of these studies were screened for Y_1 receptor affinity in competition binding studies using membrane preparations from Y_2Y_4 knockout mice as described previously.[9] The conjugates **10**–**15** showed dose-dependent competition with radiolabelled NPY for receptor binding and comparable to the parent sequences in most cases with 50 % inhibitory concentration (IC_{50}) values of between 0.6 and 6 nM (Table 1, Fig. S1 in the Supplementary Material). This showed that the receptor was relatively unaffected by the peptide substitutions irrespective of the position or chirality of the alkoxy substituent. Compound **15** was the exception with a relatively poor affinity, the combination of FBz group and coumarin (Fig. 2) both proving deleterious to affinity.

Two rhodamine derivatives **16** and **17** were also prepared, noting that these analogues could be potentially of use in receptor imaging using the fluorescence excitation properties of the rhodamine group. These were assayed in a recombinant Y_1-293TR cell system and compared again to the parent peptide **1**. The affinity of these conjugates proved comparable to **1** with IC_{50} values of 10 and 18 nM respectively (Table 1, Fig. S2 in the Supplementary Material). While still a reasonably strong affinity, we have developed superior conjugates through other routes, which will be reported elsewhere.

Synthesis of Head-to-Tail Cyclic LEDGF Mimics

The second series of peptides we chose to study using Pop-derived conjugates were designed to mimic the binding loop of LEDGF that binds to IN. Inhibitors of LEDGF–IN binding are postulated to be potential inhibitors of HIV DNA integration into host cells, and the binding loop of LEDGF comprises a tetrapeptide Lys–Ile–Asp–Asn motif.[12a,16] We had developed cyclic hexapeptides including turn-inducing dipeptide units of one or two proline residues. Crystal structures of these peptides bound to IN showed preservation of the tetrapeptide pharmacophore and that the Pro residues protruded away from the protein binding pocket, providing a potential conjugation point. While showing a modest binding affinity (K_d ~1 mM as measured by surface plasmon resonance (SPR) and HSQC NMR), these peptides show strong conformational mimicry of the native protein.

Pop residues were included in three cyclic peptides based upon three parent cyclic hexapeptides for which we had determined crystal structures: cyclo[Asn–D-Pro–Pro–Lys–Ile–Asp] **18**, cyclo[Asn–D-Val–Pro–Lys–Nle–Asp] **19**, cyclo[Asn–D-Val–Pro–Lys–D-Ile–Asp], and **20** (PDB ID: 3WNG and 3WNH) (Northfield et al., in preparation). In the first example, D-proline of **18** was replaced with *cis*-4-propargyloxy-D-proline (**xiv**) to

Table 1. Conjugated NPY-derived peptides, incorporating functionality using click chemistry
FBz = 4-Fluorobenzoyl. For other abbreviations see Fig. 2

Peptide	Sequence	ESI-MS[A] [*m*/*z*]	IC_{50} [nM] Y_2Y_4 KO[E]
Precursor peptides			
1	Ile–Asn–Pro–Lys–Tyr–Arg–Leu–Arg–Tyr (Lys^4-BVD15)	611.7[B]	0.9
2	Ile–Asn–Pro–Arg–Tyr–Arg–Leu–Arg–Tyr (Arg^4-BVD15)	625.7[B]	1.3
3	Ile–cyc[Glu–Pro–Dap]–Tyr–Arg–Leu–Arg–Tyr (c[Glu^2,Dap^4]–BVD15)	589.2[B]	0.9
4	Ile–Asn–*trans*-4-L-Pop–Lys–Tyr–Arg–Leu–Arg–Tyr	638.6[B]	1.0
5	Ile–Asn–*trans*-4-L-Pop–Arg–Tyr–Arg–Leu–Arg–Tyr	652.7[B]	
6	Ile–Asn–*cis*-4-L-Pop–Lys–Tyr–Arg–Leu–Arg–Tyr	638.8[B]	
7	Ile–Asn–*cis*-3-L-Pop–Lys–Tyr–Arg–Leu–Arg–Tyr	638.8[B]	
8	Ile–cyc[Glu–*trans*-4-L-Pop–Dap]–Tyr–Arg–Leu–Arg–Tyr	616.3[B]	
9	FBz–Ile–cyc[Glu–*trans*-4-L-Pop–Dap]–Tyr–Arg–Leu–Arg-Tyr	677.4[B]	
Final products			
10	Ile–Asn–*trans*-4-L-Ctp–Lys–Tyr–Arg–Leu–Arg–Tyr	498.3[C]	0.6
11	Ile–Asn–*trans*-4-L-Ctp–Arg–Tyr–Arg–Leu–Arg–Tyr	507.8[C]	6.0
12	Ile–Asn–*cis*-4-L-Ctp–Lys–Tyr–Arg–Leu–Arg–Tyr	498.4[C]	1.9
13	Ile–Asn–*cis*-3-L-Ctp–Lys–Tyr–Arg–Leu–Arg–Tyr	498.3[C]	0.9
14	Ile–cyc[Glu–*trans*-4-L-Ctp–Dap]–Tyr–Arg–Leu–Arg–Tyr	724.3[B]	1.9
15	FBz–Ile–cyc[Glu–*trans*-4-L-Ctp–Dap]–Tyr–Arg–Leu–Arg–Tyr	785.5[B]	41
			IC_{50} [nM] Y_1-HEK293[F]
1			7.9
16	Ile–Asn–*trans*-4-L-R^{1}tp–Arg–Tyr–Arg–Leu–Arg–Tyr	685.7[D]	10
17	Ile–Asn–*trans*-4-L-R^{2}tp–Arg–Tyr–Arg–Leu–Arg–Tyr	649.0[D]	18

[A]ESI-MS = electrospray ionisation–mass spectrometry.
[B]ESI-MS base peak corresponds to $[M + 2H]^{2+}$.
[C]ESI-MS base peak corresponds to $[M + 3H]^{3+}$.
[D]ESI-MS base peak corresponds to $[M + TFA + 3H]^{3+}$ ($[M + 3H]^{3+}$ peaks were observed at lower intensity).
[E]Inhibition of ^{125}I-NPY (25 pM) binding to brain membrane homogenates.
[F]Inhibition of ^{125}I-PYY (15 pM) binding to Y_1 transfected 293TR cells.

give **21**, and in **19** and **20** L-proline was replaced with *trans*-4-propargyloxy-L-proline (**xiii**) to give **22** and **23** respectively (Table 2).

The synthesis of the peptides was achieved with Fmoc-based SPPS of side-chain protected linear precursors followed by solution phase cyclisation and then side-chain deprotection. The linear hexapeptide chains were prepared from Fmoc–Asp(OtBu)–chlorotrityl resin, followed by solution head-to-tail cyclisation of the linear peptides using diphenylphosphorylazide (DPPA) as the cyclisation reagent and finally side-chain cleavage. The sequences showed some propensity for racemisation in the cyclisation step, but the D-Asp-containing diastereomers were in general minor components and readily separated from the desired compounds. The recovered yields for cyclisation of Pop-containing peptides **21**–**23** ranged from 30 to 70 %, and were comparable to those obtained in the synthesis of the parent proline-containing sequences **18**–**20**. This demonstrated that substituting L-Pro for *trans*-4-L-Pop and D-Pro for *cis*-4-D-Pop did not have a detrimental effect on peptide cyclisation.

Table 2. Conjugated LEDGF-derived peptides, incorporating functionality using click chemistry

Peptide	Sequence	ESI-MS[A] [Da]
Precursors		
18	cyclic[Asn–D-Pro–Pro–Lys–Ile–Asp]	665.6
19	cyclic[Asn–D-Val–Pro–Lys–Nle–Asp]	667.5
20	cyclic[Asn–D-Val–Pro–Lys–D-Ile–Asp]	667.5
21	cyclic[Asn–*cis*-4-D-Pop–Pro–Lys–Ile–Asp]	719.5
22	cyclic[Asn–D-Val–*trans*-4-L-Pop–Lys–Nle–Asp]	721.6
23	cyclic[Asn–D-Val–*trans*-4-L-Pop–Lys–D-Ile–Asp]	721.6
Final products		
24	cyclic[Asn–*cis*-4-D-Ctp–Pro–Lys–Ile–Asp]	935.4
25	cyclic[Asn–*cis*-4-D-Atp–Pro–Lys–Ile–Asp]	977.4
26	cyclic[Asn–*cis*-4-D-Btp–Pro–Lys–Ile–Asp]	1119.6
27	cyclic[Asn–D-Val–*trans*–4-L-Actp–Lys–Nle–Asp]	979.5
28	cyclic[Asn–D-Val–*trans*-4-L-Btp–Lys–Nle–Asp]	1121.6
29	cyclic[Asn–D-Val–*trans*-4-L-Ptp–Lys–Nle–Asp]	926.6
30	cyclic[Asn–D-Val–*trans*-4-L-Btp–Lys–D-Ile–Asp]	1121.6

[A]ESI-MS peaks correspond to m/z: $[M + H]^+$.

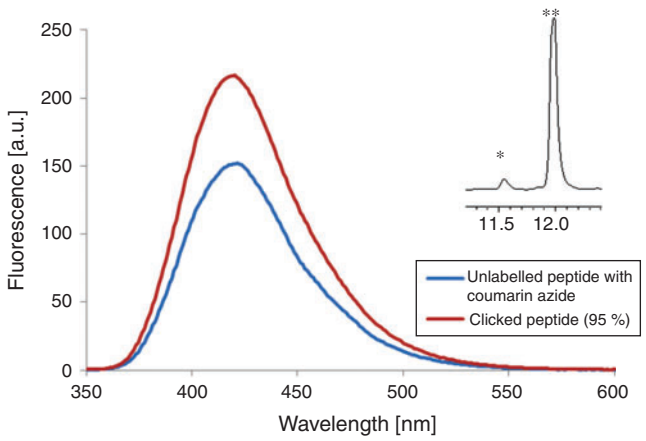


Fig. 3. Formation of labelled peptide **25** from **21**. (a) Reaction mixture in absence (blue) or presence (red) of catalyst after 48 h reaction time. (b) Reverse phase-HPLC trace of mixture showing **21** (*) and **25** (**).

The Pop-substituted cyclic peptides were purified by semi-preparative RP-HPLC, then used as substrates for click reactions with a variety of azido-compounds: 7-amino-4-azidomethylcoumarin, 7-acetylamino-4-azidomethylcoumarin, Biotin-PEG_2-azide, and azido-6-deoxy-α-D-galactopyranose (Fig. 2, Table 2). All of the selected conjugates were successfully coupled to one or more of the cyclic peptides **21**–**23**, under standard conditions within 8 h using the same DMF click chemistry method described for the NPY analogue conjugation above. An interesting feature of the coupling of the 4-azidomethyl-7-acetamidocoumarin in the synthesis of peptide **25** was the increase in fluorescence with progression of the reaction over a 48 h period, while the spectrum of the same mixture in the absence of the catalyst was unchanged (Fig. 3). While modest, the ability to distinguish substrates from products by fluorescence might be useful in performing click reactions in more complex media.

Analysis of Cyclic LEDGF Mimics

With the conjugated peptides in hand we examined the effect the substitution had on the peptide conformation at the LEDGF binding site of IN. The peptides showed comparable, albeit weak affinity for IN to the parent hexapeptides, **18** and **19**, and we obtained crystal structures of respective labelled derivatives peptide **24** and peptide **28**, and the propargyloxy derivatives **21** and **22** bound to the core domain of IN.

The crystal structures all show well resolved density at the tetrapeptide sequence, allowing for comparison of the homologous series at the key pharmacophore (Fig. 4). The density of the prosthetic structures was poorly resolved suggesting that the labels are flexible and do not interact with the IN protein.

Peptides **18**, **21**, and **24** differ only in the presence of the substituent at the *cis*-4-position of the D-proline residue. Peptide **18** and the coumarin conjugate **24** show a close overlay of the tetrapetide pharmacophore. In peptide **21** however the lysine residue has moved substantially and cannot be said to be mimicking the native structure. In both cases, the proline motif has undergone some conformational change, although the

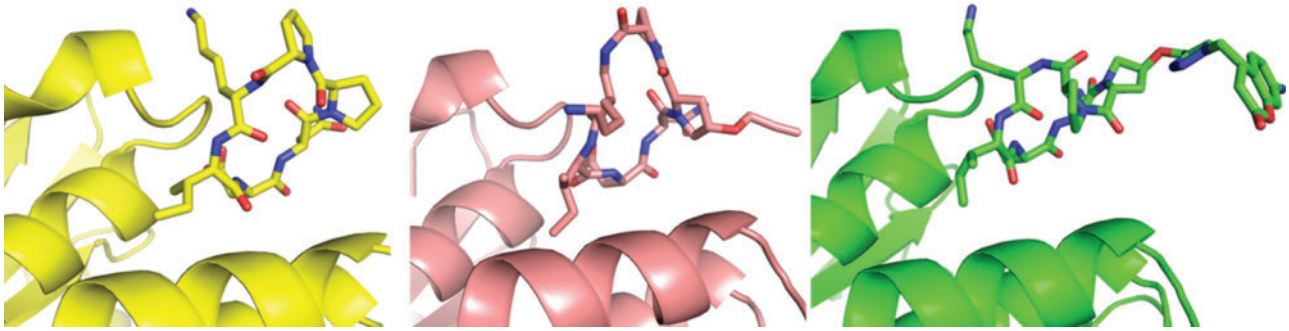


Fig. 4. Crystal structures of peptides **18**, **21**, and **24** (left to right) in complex with core domain of HIV integrase (IN).

resolution at those residues is not sufficient to identify the cause. In the case of the cyclic peptide structures **19**, **28**, and **22** again the conformation of the conjugate **28** more closely resembled the native Pro-containing peptide **19** than the Pop counterpart **22** (Fig. 3 in the Supplementary Material).

Conclusions

In commencing this work, we reasoned that the incorporation of (1*H*-1,2,3-triazol-4-yl)methoxy substituents on proline would be a successful strategy in order to produce conjugated versions of proline-containing peptides, especially as compared with the corresponding products from commonly used propargylglycine. Proline is a structurally rigid amino acid and so an additional substituent might not be expected to impact the native peptide conformation, and the linker group is also relatively remote from the peptide backbone. Except in the case of proline isomerases and proline specific proteases, proline does not generally play a direct role in ligand binding events and so can be a benign place to make a residue replacement. Furthermore, the use of varied stereochemistry or regiochemistry of substituents projecting from the proline ring allows conjugates to be directed away from the peptide pharmacophore or the target protein binding site, but may also influence levels of *cis* or *trans* amide conformers. The Pop stereoisomers can be readily obtained from commercially available starting materials and incorporated in peptide sequences using standard Fmoc-SPPS protocols. The pharmacological and biophysical data we have obtained in these two examples supports the approach for both small linear and cyclic peptides. Given the absence of N- or C-terminal residues, the ability to link through proline seems prospectively valuable in head-to-tail cyclic peptides especially. However, the data also provide the caveat that irrespective of the linking handle, the nature of the prosthetic group can have a marked effect on the binding affinity or conformation adopted by these conjugates.

Experimental

N^{α}-Fmoc-protected amino acids were purchased from Auspep and ChemImpex. Rink amide resin and *O*-(1*H*-6-chlorobenzotriazol-1-y1)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate (HCTU) were purchased from ChemImpex. Piperidine and TFA were purchased from Auspep. *N*,*N*-DIPEA, DMF, and dichloromethane (DCM), were purchased from Merck. Diphenylphosphorylazide (DPPA) and triisopropylsilane (TIPS) were purchased from Sigma–Aldrich. Fluorobenzoic acid was purchased from Alfa Aesar. 7-Amino-4-(azidomethyl)-2*H*-chromen-2-one[17] was a gift from Dr Bim Graham (Monash Institute of Pharmaceutical Sciences). The Rhodamine B derivatives were prepared in-house. All chemicals were used without further purification.

^{1}H NMR spectra were routinely recorded at 300 MHz using a 300 MHz Bruker Advance DPX-300 spectrometer or at 400 MHz using a 400 MHz Bruker Ultrashield–Advance III NMR spectrometer, with *TOPSPIN v2.1* software, at 298 K. ^{13}C NMR spectra were recorded at 101 MHz using a 400 MHz Bruker Ultrashield–Advance III NMR spectrometer, with *TOPSPIN v2.1* software, at 298 K. Liquid chromatography mass spectra were acquired on a Shimadzu 2020 LCMS system incorporating a photodiode array detector coupled directly into an electrospray ionisation source and a single quadrupole mass analyser. Standard RP-HPLC was carried out at room temperature employing a Phenomenex Luna C8 (100 × 2.0 mm internal diameter, I.D.) column eluting with a gradient of either 0–64 % acetonitrile (ACN) in 0.05 % aqueous TFA over 10 min or 0–100 % B over 15 min (Buffer B is 100 % ACN + 0.1 % TFA) at a flow rate of 0.2 mL min^{-1} unless stated otherwise. Mass spectra were obtained in positive mode with a scan range of *m*/*z* 200–2000. Semi-preparative RP-HPLC was performed using a Waters Associates liquid chromatography system (Model 600 controller and Waters 486 Tuneable Absorbance Detector) using a gradient of 0–64 % ACN in 0.1 % TFA over 20 min or 30 min at a flow rate of 10 mL min^{-1} on a Phenomenex Luna C8 100 Å, 10 µm (50 × 21.2 mm I.D.) or a Phenomenex Luna C8 100 Å, 10 µm (250 × 21.2 mm I.D.) column.

Chemical Synthesis

(2*S*,4R)-1-(Tert-butoxycarbonyl)-4-hydroxypyrrolidine-2-carboxylic Acid (Boc-trans-L-4-hydroxyproline, **v**)

To a stirred solution of *trans*-L-4-hydroxyproline **i** (2.0 g, 15.3 mmol) in MeOH (36.0 mL) was added Et_3N (4.0 mL, 28.7 mmol) and Boc anhydride (6.7 g, 30.5 mmol) and the reaction was refluxed for 3.5 h, cooled to room temperature, and stirred for 20 h. Solvent was removed under vacuum and the residue cooled to 0°C. Following the addition of NaH_2PO_4 (150 mg), the solution was acidified to pH 2 with 0.5 M HCl. The mixture was stirred at 0°C for 30 min before extracting the product with EtOAc (4 × 20 mL). The combined organic layers were dried with $MgSO_4$ and filtered. The solvent was removed under vacuum yielding **v** as a white foam (3.23 g, 14 mmol, 92 %)

δ_H (CD_3OD, 400 MHz) 4.40 (dd, *J* 5.5, 3.4, CH, 1H), 4.32 (dt, *J* 12.9, 8.0, CH, 1H), 3.54 (dt, *J* 11.4, 4.0, 0.5 × CH_2, 1H), 3.44 (dt, *J* 11.4, 1.9, 0.5 × CH_2, 1H), 2.27 (dddd, *J* 12.3, 7.7, 2.8, 1.8, 0.5 × CH_2, 1H), 2.06 (ddd, *J* 13.2, 8.6, 4.5, 0.5 × CH_2, 1H), 1.45 (s, Boc, 9H). δ_C (CD_3OD, 101 MHz) 176.75 and 176.37 (pair of rotamers, Cq), 156.54 and 156.02 (pair of rotamers, Cq), 81.72 and 81.42 (pair of rotamers, Cq), 70.68 and 70.06 (pair of rotamers, CH), 59.39 and 58.91 (pair of rotamers, CH), 55.85 and 55.51 (pair of rotamers, CH_2), 40.07 and 39.4 (pair of rotamers, CH_2), 28.71 and 28.53 (pair of rotamers, 3 × CH_3). *m*/*z* (LC-MS) 277.35 (100 %, $[M + 2Na]^+$).

(2*S*,4R)-1-(Tert-Butoxycarbonyl)-4-(prop-2-yn-1-yloxy) pyrrolidine-2-carboxylic Acid (Boc-trans-L-4-propargyloxyproline, ***ix***)

A solution of Boc-*trans*-L-hydroxyproline, **v** (2.80 g, 12.13 mmol) in dry DMF (30 mL) was added to a suspension of NaH (0.93 g, 38.75 mmol) in dry DMF (10 mL) under nitrogen at 0°C. After 15 min, 1.5 equivalents of propargyl bromide (80 % in toluene) was added dropwise to the reaction (1.68 mL, 18.85 mmol). The reaction was stirred at 0°C for 2 h and then quenched with H_2O and lyophilised in H_2O/ACN. The reaction was taken up in EtOAc and the pH adjusted to 2 with 10 % citric acid. The aqueous layer was extracted with EtOAc (3 × 20 mL). The combined organic layers were dried with $MgSO_4$ and filtered. Solvent was removed under vacuum to yield **ix** as a brown solid (2.97 g, 11.0 mmol, 91 %) which was directly carried on to the next step.

δ_H (CD_3OD, 400 MHz) 4.37–4.31 (m, CH, 1H), 4.31–4.21 (m, CH, 1H), 4.19 (d, *J* 2.4, CH_2, 2H), 3.64–3.57 (m, 0.5 × CH_2, 1H), 3.57–3.50 (m, 0.5 × CH_2, 1H), 2.94–2.77 (m, CH, 1H), 2.44 (tddd, *J* 14.3, 11.5, 3.0, 1.6, 0.5 × CH_2, 1H), 2.13– 2.04 (m, 0.5 × CH_2, 1H), 1.45 (s, Boc, 9H). δ_C (CD_3OD, 101 MHz) 178.33 and 175.54 (pair of rotamers, Cq), 156.04 and 155.93 (pair of rotamers, Cq), 81.64 and 80.9 (pair of rotamers, Cq),

79.31 (CH), 76.2 and 75.85 (pair of rotamers, CH), 75.04 (Cq), 57.90 and 57.87 (pair of rotamers, CH), 56.58 and 56.51 (pair of rotamers, CH_2), 51.93 and 51.18 (pair of rotamers, CH_2), 36.57 and 34.57 (pair of rotamers, CH_2), 28.45 and 28.33 (pair of rotamers, $3 \times CH_3$). *m/z* (LCMS) 315.35 (80 %, $[M + 2Na]^+$).

(2S,4R)-1-(((9H-Fluoren-9-yl)methoxy)carbonyl)-4-(prop-2-yn-1-yloxy)pyrrolidine-2-carboxylic Acid (Fmoc-trans-L-4-Propargyloxyproline-OH, ***xiii****)*

Boc-*trans*-L-propargyloxyproline **ix** (2.97 g, 11.01 mmol) was treated with 1 : 1 TFA/DCM (10 mL) at room temperature over 45 min and solvent removed under vacuum. The reaction was diluted with H_2O (10 mL) and adjusted to pH 9 with Na_2CO_3. To the reaction solution 1.4 equiv. of Fmoc-OSu (5.20 g, 16.18 mmol) in dioxane (22 mL) was added at 0°C and stirred for 1 h. The reaction was then brought to room temperature and stirred overnight. Dioxane was removed under vacuum and the reaction acidified to pH 3 with 1 M HCl. Product was extracted with EtOAc (3×20 mL), washed with brine, and dried with $MgSO_4$. Solvent was removed under vacuum to yield a yellow foam. Purification was achieved by flash chromatography (0–2 % MeOH in chloroform) yielding **xiii** as a white powder (1.31 g, 3.35 mmol, 30 %).

δ_H (CD_3OD, 400 MHz) 7.8 (t, *J* 7.5, 2H), 7.63 (td, *J* 7.5, 2.4, 2H), 7.39 (td, *J* 7.4, 4.0, 2H), 7.35–7.28 (m, 2H), 4.46–4.18 (m, 6H), 4.15 (dd, *J* 4.7, 2.4, 1H), 3.64–3.5 (m, 2H), 2.89 (t, *J* 2.4, 1H), 2.57–2.4 (m, 1H), 2.22–2.05 (m, 1H). δ_C (CD_3OD, 101 MHz) 175.98 and 175.75 (pair of rotamers, Cq), 156.71 and 156.62 (pair of rotamers, Cq), 145.31, 145.29, 145.12, 145.05 (rotamers, Cq), 142.64, 142.61, 142.56, 142.49 (rotamers, Cq), 128.88 (CH), 128.25 (CH), 126.28, 125.25, 126.16, 126.15 (rotamers, CH), 121.03 and 120.98 (pair of rotamers, CH), 80.58 and 80.57 (pair of rotamers, Cq), 77.92 and 77.15 (pair of rotamers, CH), 76.27 and 77.26 (pair of rotamers, CH), 69.32 and 68.75 (pair of rotamers, CH_2), 59.26 and 59.01 (pair of rotamers, CH), 57.25 and 57.21 (pair of rotamers, CH_2), 53.21 and 52.78 (pair of rotamers, CH_2), 48.39 and 48.33 (pair of rotamers, CH), 37.6 and 36.6 (pair of rotamers, CH_2). *m/z* (LCMS) 392.30 (100 %, $[M + H]^+$). HRMS *m/z* 392.1494; $C_{23}H_{22}NO_5^+$ $[M + H]^+$ requires 392.1492.

Compounds **xiv**, **xv**, and **xvi** were prepared in the same manner. Full details are provided in the Supplementary Material.

BVD15 analogues **1**–**9** were prepared as previously reported. Peptide syntheses were performed on Rink amide resin (0.3–0.7 mequiv. g^{-1}, 100–200 mesh, 0.1 mmol scale) using conventional Fmoc-based solid phase peptide synthesis. Fmoc-protected amino acids in 3-fold molar excess were coupled using DMF as solvent, a 6-fold molar excess of DIPEA in DMF (70 mL L^{-1}) with a 3-fold molar excess of HCTU as the activating agent for 50 min. Fmoc deprotection was carried out by treatment with 20 % piperidine in DMF for 10 min.

Peptide cleavage from the resin was performed using a cocktail containing TFA/TIPS/DMB (92.5 : 2.5 : 5 %; DMB = 1,3-dimethoxybenzene) for 3 h.[18] The cleavage mixture was filtered, concentrated by a stream of nitrogen, precipitated by cold diethyl ether, and centrifuged. The resulting crude product was dissolved by water/ACN (1 : 1) and lyophilised overnight.

The click reactions to prepare peptides **10**–**15** involved dissolving the corresponding peptide–alkyne **5**–**9** (1 equiv.) in H_2O and adding a solution of the azidocoumarin[17] (4 equiv.) in DMF to give a 1 : 3 ratio of H_2O to DMF. Copper sulfate (10 equiv.), TBTA (10 equiv.), and sodium ascorbate (10 equiv.) were then added and the reaction mixed for 3 h. Peptides **16** and **17** were prepared in the same fashion but using the appropriate azido-substituted rhodamine B derivatives. Peptides were purified by reverse-phase preparative HPLC. Purity of fractions was assessed using electrospray ionisation-mass spectrometry (ESI-MS) (Table 1) and analytical HPLC (Fig. S5 in the Supplementary Material).

Synthesis of Cyclic LEDGF Analogues

Cyclic peptides **18**–**23** were synthesised on 2-chlorotrityl chloride (2CTC) resin on a 0.1 mequiv. scale. Couplings were performed using 3 equiv. of Fmoc-protected amino acid, 3 equiv. of HCTU, and 6 equiv. of DIPEA in DMF (0.1 M in amino acid) for 50 min. Fmoc deprotection was carried out with 30 % (v/v) piperidine in DMF (2×5 min). After each coupling and deprotection step, the resin was washed six times with DMF. Peptides were cleaved from the resin using 1 % (v/v) TFA in DCM. Head-to-tail cyclisation of side-chain protected peptide was performed in DMF (4 mM final concentration of peptide) with 3 equiv. of DPPA and 4 equiv. of DIPEA. Following removal of the solvent, side-chain protecting groups were removed in 95 : 5 TFA/TIPS. After cyclisation, the peptides were purified by reverse-phase preparative HPLC. Purity of fractions was assessed using ESI-MS (Table 2) and analytical HPLC.

Synthesis of Peptides ***24–30***

A solution of the Pop-containing peptide (1 mg mL^{-1} in H_2O) was treated with a 4-fold excess of the azido derivative (1 mg mL^{-1} in DMF). One equivalent of sodium ascorbate (1 mg mL^{-1} in H_2O), one equivalent of TBTA (1 mg mL^{-1} in DMF), and one equivalent of copper sulfate (1 mg mL^{-1} in H_2O) were subsequently added to the reaction. The reaction was left at room temperature and progression monitored by LCMS. When no remaining unlabelled peptide was observed, the reaction was diluted in 1 : 1 ACN/H_2O and lyophilised before purification by RP-HPLC. Purity of fractions was assessed using ESI-MS (Table 2) and analytical HPLC (Fig. S5 in the Supplementary Material).

In the case of peptide **29**, the click reaction was performed using 1,2 : 3,4-di-*O*-isopropylidne-6-azido-6-deoxy-α-D-galactopyranose.[19] The resultant acetonide (*m/z* 1006.7, $[M + H]^+$) was deprotected by treatment with 90 % TFA overnight, diluted in 1 : 1 ACN/H_2O, and lyophilised before purification by RP-HPLC to yield the free galactopyranose **29**.

Competition Binding Studies

Competition binding assays were carried out as described previously.[9] In brief, receptor binding assays to measure Y_1R affinity of the ligands **10**–**15** (described below) were performed on crude membranes prepared from the brains of Y_2R- and Y_4R-deficient mice ($Y_2-/-Y_4-/-$), where Y_1R accounts for the majority of remaining Y receptors. Peptides **16** and **17** were assayed using 293TR Y_1 receptor GFP membranes.

For mouse brain preparations, equal volumes (25 µL) of non-radioactive ligands and ^{125}I-human polypeptide YY (^{125}I-hPYY, 2200 Ci $mmol^{-1}$; PerkinElmer Life Science Products, Boston, MA, USA) were added into each assay. The final concentration of ^{125}I-hPYY in the assay was 25 pM. The binding of ^{125}I-hPYY competed with Y_1R ligands of interest at increasing concentrations ranging from 10^{-12} to 10^{-6} M over 2 h. Non-radioactive human PYY (Auspep, Parkville, Vic., Australia) at 10^{-6} M was used as the non-specific binding control.

Using membranes from the 293TR Y_1 receptor-sfGFP cell competition binding assays were performed for 90 min at 21°C in buffer (25 mM HEPES, 2.5 mM $CaCl_2$, 1.0 mM $MgCl_2$, 0.1 % bovine serum albumin, 0.1 mg mL^{-1} bacitracin; pH 7.4), increasing concentrations of unlabelled ligands (10^{-12} to 10^{-6} M, duplicate), and [^{125}I]PYY (15 pM). Non-specific binding in these experiments comprised less than 5 % of total counts, and was subtracted from the data.

In both sets of data, IC_{50} values were calculated from displacement curves (repeated 2–4 times for each peptide, fitted using non-linear least-squares regression in *GraphPad Prism 5.01* (Graphpad software, San Diego, CA, USA).

X-Ray Crystallography

Crystal structures of the cyclic hexapeptides bound to IN were determined as previously described.[20] The coordinates of the four $IN_{CORE4H123}$/cyclic LEDGF peptide complexes have been deposited in the protein database (PDB) with the accession numbers 4Y1C and 4Y1D.

Supplementary Material

Detailed synthesis procedures as well as additional supplementary figures showing dose–response curves for Y_1R binding by peptides **1**–**4** and **10**–**17**, structures of peptides **19**, **22**, and **28** complexed with IN, NMR data for Pop derivatives **v**–**xvi**, and RP-HPLC traces of peptides **10**–**17** and **24**–**30** are available on the Journal's website.

Acknowledgements

This work was supported by CRC for Biomedical Imaging Development (CRC-BID), Australia, and the Australian Research Council (ARC) Linkage project (LP0775192). This research was partly undertaken on the MX2 beamline at the Australian Synchrotron, Victoria, Australia and the authors thank the beamline staff for their assistance. Funding from the Victorian Government Operational Infrastructure Support Scheme to St Vincent's Institute is gratefully acknowledged. M.W.P. is a National Health and Medical Research Council of Australia Research Fellow. SEN and ML were supported by an Australian Post-graduate scholarship.

References

[1] R. J. Pieters, D. T. S. Rijkers, R. M. J. Liskamp, *QSAR Comb. Sci.* **2007**, *26*, 1181. doi:10.1002/QSAR.200740075

[2] C. Hein, X.-M. Liu, D. Wang, *Pharm. Res.* **2008**, *25*, 2216. doi:10.1007/S11095-008-9616-1

[3] (a) M. MacDonald, J. Aube, *Curr. Org. Chem.* **2001**, *5*, 417. doi:10.2174/1385272013375517
(b) C. M. Deber, B. Brodsky, A. Rath, in *Encyclopedia of Life Sciences* **2010** (John Wiley & Sons, Ltd: Chichester).

[4] A. K. Pandey, D. Naduthambi, K. M. Thomas, N. J. Zondlo, *J. Am. Chem. Soc.* **2013**, *135*, 4333. doi:10.1021/JA3109664

[5] M. Meldal, C. W. Tornøe, *Chem. Rev.* **2008**, *108*, 2952. doi:10.1021/CR0783479

[6] G. Chouhan, K. James, *Org. Lett.* **2011**, *13*, 2754. doi:10.1021/OL200861F

[7] D. Font, C. Jimeno, M. A. Pericàs, *Org. Lett.* **2006**, *8*, 4653. doi:10.1021/OL061964J

[8] B. Guérin, V. Dumulon-Perreault, M.-C. Tremblay, S. Ait-Mohand, P. Fournier, C. Dubuc, S. Authiera, F. Bénard, *Bioorg. Med. Chem. Lett.* **2010**, *20*, 950. doi:10.1016/J.BMCL.2009.12.068

[9] S. J. Mountford, M. Liu, L. Zhang, M. Groenen, H. Herzog, N. D. Holliday, P. E. Thompson, *Org. Biomol. Chem.* **2014**, *12*, 3271. doi:10.1039/C4OB00176A

[10] M. Liu, S. J. Mountford, L. Zhang, I. C. Lee, H. Herzog, P. E. Thompson, *Int. J. Pept. Res. Ther.* **2013**, *19*, 33. doi:10.1007/S10989-012-9330-Z

[11] A. Engelman, P. Cherepanov, *PLoS Pathog.* **2008**, *4*, e1000046. doi:10.1371/JOURNAL.PPAT.1000046

[12] (a) W. Thys, K. Busschots, M. McNeely, A. Voet, F. Christ, Z. Debyser, *HIV Ther.* **2009**, *3*, 171. doi:10.2217/17584310.3.2.171
(b) S. Hare, P. Cherepanov, *Viruses* **2009**, *1*, 780. doi:10.3390/V1030780

[13] V. Mihali, F. Foschi, M. Penso, G. Pozzi, *Eur. J. Org. Chem.* **2014**, *2014*, 5351. doi:10.1002/EJOC.201402429

[14] M. Obkircher, C. Stähelin, F. Dick, *J. Pept. Sci.* **2008**, *14*, 763. doi:10.1002/PSC.1001

[15] V. Hong, S. I. Presolski, C. Ma, M. G. Finn, *Angew. Chem. Int. Ed.* **2009**, *48*, 9879. doi:10.1002/ANIE.200905087

[16] Z. Hayouka, M. Hurevich, A. Levin, H. Benyamini, A. Iosub, M. Maes, D. E. Shalev, A. Loyter, C. Gilon, A. Friedler, *Bioorg. Med. Chem.* **2010**, *18*, 8388. doi:10.1016/J.BMC.2010.09.046

[17] M. A. Kamaruddin, P. Ung, M. I. Hossain, B. Jarasrassamee, W. O'Malley, P. Thompson, D. Scanlon, H.-C. Cheng, B. Graham, *Bioorg. Med. Chem. Lett.* **2011**, *21*, 329. doi:10.1016/J.BMCL.2010.11.005

[18] P. Stathopoulos, S. Papas, V. Tsikaris, *J. Pept. Sci.* **2006**, *12*, 227. doi:10.1002/PSC.706

[19] S. B. Ferreira, A. C. R. Sodero, M. F. C. Cardoso, E. S. Lima, C. R. Kaiser, F. P. Silva Jr, V. F. Ferreira, *J. Med. Chem.* **2010**, *53*, 2364. doi:10.1021/JM901265H

[20] J. Wielens, S. J. Headey, J. J. Deadman, D. I. Rhodes, G. T. Le, M. W. Parker, D. K. Chalmers, M. J. Scanlon, *ChemMedChem* **2011**, *6*, 258. doi:10.1002/CMDC.201000483