Synthesis of L-Cyclic Tetrapeptides by Backbone Amide Activation CyClick Strategy

Rachel Wills[1], Victor Adebomi[1], Caroline Spancake[1], Ryan D. Cohen[2] and Monika Raj[*,1]

[1]Department of Chemistry, Emory University, Atlanta, Georgia 30322, USA

[2]Merck & Co. Inc., Rahway, New Jersey 07065, USA

Supporting Information Placeholder

ABSTRACT**:** Cyclic tetrapeptides exhibit high cellular permeability and a wide range of biological properties and thus have gained great interest in the field of medicinal chemistry. We synthesized highly strained 12-membered head to tail cyclic peptides with varying reactive amino acids and without the formation of any dimers or oligomers by the unique exclusively intramolecular CyClick chemistry. This occurs by a two-step process involving the low-energy formation of a 15 atom-containing cyclic imine, followed by a chemoselective ring contraction of the peptide backbone generating a highly strained 12 atom-containing cyclic tetrapeptide. The formation of a cyclic intermediate brings the electrophilic imine in close proximity of the N-terminal amide, thereby lowering the activation energy to synthesize strained cyclic tetrapeptides compared to conventional approaches. This reaction exhibited high substrate scope and generated head to tail cyclic tetrapeptides with varying amino acids at the N-terminus, showing chemoselectivity without the need for side group protection.

INTRODUCTION

A wide range of biological activities, high protease stability, and ability to selectively inhibit protein-protein interactions (PPIs) make cyclic peptides highly attractive in the field of medicinal chemistry.[1-6] Despite these significant advantages, there are severe challenges particularly associated with the synthesis of small cyclic peptides. Tetrapeptides, which contain 12 atoms in their backbone, are often generated in poor yields.[7-12] This is due to the high ring strain associated with additional transannular interactions between substituents of cyclic tetrapeptides.

Current methods for chemical synthesis of cyclic tetrapeptides in solution are limited due to the formation of dimers or oligomers [9,10] and the requirement of multiple turn inducers[11] that reduces overall yields and restricts chemical diversity of cyclic peptides. To avoid the formation of oligomers, cyclization reactions are carried out at very low concentrations (0.001-0.005 M)[10-12] leading to significantly slow reaction rates (~ 3 days).[11-12] Therefore, the synthesis of cyclic tetrapeptides comprising any of the regular L-amino acids is still considered highly challenging.

Herein, we describe the synthesis of head to tail cyclic tetrapeptides through an exclusively intramolecular CyClick strategy (Figure 1) that involves formation of an imine between a C-terminal aldehyde and the N-terminus of the linear pentapeptide followed by ring-closure with the amide backbone chain to form a lactam bridge and a 4-imidazolidinone moiety at the site of cyclization.[13-15] We propose that the ring-contraction of cyclic imines of pentapeptides by the CyClick approach is responsible for

decreasing the activation energy for the synthesis of cyclic tetrapeptides. Since the CyClick reaction is exclusively intramolecular, we hypothesized that cyclization can be carried out at high concentrations (25 mM) without the formation of any linear dimers or oligomers while concomitantly reducing the overall rate of the reaction.

MATERIALS AND METHODS

The materials, methods, and compound characterization are described in detail in Supporting Information. All environmental and safety precautions were followed according to the material safety data sheets of the chemicals utilized in this work.

RESULTS AND DISCUSSION

Our first goal was to optimize the cyclization reaction with a peptide aldehyde using different bases. We synthesized C-terminal peptide aldehydes by the reaction scheme in Figure 2 involving coupling of Fmoc-Ala-CHO on free threonine on solid support, followed by successive residue attachment via Fmoc-SPPS (Figure 2, Supplementary Figure 1, Supporting Information).[13-16] The cleavage of a peptide from solid support using trifluoroacetic acid generated an aldehyde at the C-terminus. Our previous study showed very high cyclization efficiency (ca. 80-99%) with medium-sized head-to-side-chain macrocycles (20 atoms ring size) with varying amino acids at the N-terminus except with bulky residues such as Val (~ 37%). Therefore in this study, we proceeded with a peptide aldehyde VVGPFEY **1a** containing Val at the N-terminus for the head-to-side chain cyclization to evaluate the role of different bases in the efficiency of cyclization (Figure 3, Supplementary Figure 2, Supporting Information). We used bases with different pKa values[17,18] (DMAP(4-Dimethylaminopyridine), pKa ~9.6), (imidazole, pKa ~6.95), (pyrrolidine, pKa ~11.5) and (DBU (1,8-Diazabicyclo[5.4.0]undec-7-ene), pKa ~13.5) but did not observe any significant change in the amount of cyclization product (28-34%, Figure 3a, Supplementary Figure 2, Supporting Information). None of the bases led to the formation of any dimers or oligomers with a hexapeptide aldehyde VVGPFEY **1a,** thus further confirming the exclusive intramolecular nature of CyClick Chemistry. Monitoring formation of cyclic imine by the reaction between the N-terminus and C-terminal aldehyde is problematic using LCMS because the masses of the monocyclic imine and 4-imidazolidinone monoCyClick product are identical (Supplementary Figure 2, Supporting Information) Therefore, we distinguished the formation of CyClick cyclization product cyc-4-Imz-V(VGPFEY) **2a** with 4-imidazolidinone moiety at the site of cyclization as compared to the cyclic imine by adding sodium cyanoborohydride in the reaction for 16 h followed by analysis using LCMS. No change in the mass of CyClick peptide cyc-4-Imz-V(VGPFEY) **2a** was observed after incubation with sodium cyanoborohydride as compared to the to the cyclic imine which underwent reduction to generate reduced cyclic imine-product **2a'** that showed an increase in mass by +2. We analyzed all the reactions by HPLC and MS (Supplementary Figure 2, Supporting Information).

From base studies, we discovered that DMAP (7 equiv.) at room temperature gave the highest conversion to monocyclic 4-imidazolidinone product cyc-4-Imz-V(VGPFEY**) 2a** (33.8 %) and we moved to cyclizing head-to-tail linear peptide AGPFA **1b** under optimized conditions. We also carried out the head-to-tail cyclization in the absence of base but very a small amount of the cyclic product was observed thus confirming its

importance in this reaction (13%, Figure 3b, Supplementary Figure 2, Supporting Information). We observed (58%) conversion to head-to-tail monocyclic 4-imidazolidinone product cyc-4-Imz-A(GPFA) **2b** under the reaction conditions (Supplementary Figure 3, Supporting Information). To further increase the cyclization efficiency, we increased the amount of DMAP to 21 equiv. and observed a significant increase in the conversion to head-to-tail monocyclic product cyc-4-Imz-A(GPFA) **2b** (82%). The use of excess DMAP raises an issue in the HPLC data analysis. The large DMAP peak in HPLC chromatogram can overshadow peptide peaks, making purification and analysis challenging (Figure 3b, Supplementary Figure 3, Supporting Information). To avoid this problem, we switched to a volatile base, diisopropylethylamine (DIEA, pKa ~10.98), for the cyclization so that it can be removed before the analysis of a reaction mixture by HPLC (Figure 3b, Supplementary Figure 3, Supporting Information). The cyclization reaction of AGPFA **1b** with DIEA (21 equiv.) at 60 ºC proceeded smoothly and generated monocyclic product cyc-4-Imz-A(GPFA) **2b** (89%) without the formation of dimers, and DIEA was removed by rotary evaporator, resulting in a clean chromatogram for the reaction mixture (Figure 3b). Based on the high cyclization efficiency and facile HPLC analysis, we chose to use DIEA, 60 ºC for the synthesis of head-to-tail cyclic tetrapeptides. We confirmed the structure of monocyclic tetrapeptide cyc-4-Imz-A(GPFA) **2b** by NMR spectroscopy after synthesizing it on a large scale (81% conversion) from linear pentapeptide aldehyde AGPFA **1b (**Supplementary Figure 4, Supporting Information). Next, we conducted variable temperature (VT) NMR studies on **2b** to determine intramolecular hydrogen bonding in the cyclic product. Based on chemical shift temperature coefficients ($\Delta\delta_{NH}/\Delta T$ > -4.6 ppb/K),[19] both Ala and Gly amide NHs of **2b** are involved in intramolecular hydrogen bonds **(**Supplementary Figure 5, Supporting Information).

We further proceeded with cyclization of random linear pentapeptide aldehydes XGPFA **1b-1j** with Ala at the C-terminus and varying amino acid residues at the N-terminus (X = A, G, P, S, Y, D, N, R, H) (Figure 4, Supplementary Figure 6 and Table 1, Supporting Information). The peptide aldehydes XGPFA **1b-1j** were dissolved in a DMF:water mixture (molar ratio 1:1) at the concentration of 3.3 mM along with DIEA under optimized conditions. Gratifyingly, all the peptide aldehydes XGPFA **1b-1j** cyclized smoothly and generated monocyclic tetrapeptides cyc-4-Imz-X(GPFA) (**2b-2j**) within 16 hours (Figure 4, Supplementary Figure 6 and Table 1, Supporting Information). Based on previous cyclization methods for making cyclic tetrapeptide, we recognized that there could be a possibility of the formation of cyclic dimers due to the high strain associated with small ring sizes. We carefully analyzed all the reactions by LC–MS and the desired monocyclic tetrapeptide products were exclusively observed which is in contrast to the other cyclic tetrapeptide formation strategies, which lead to the formation of cyclodimers.[7-9] The unprotected amino acid residues, including tyrosine, serine, asparagine, and aspartic acid, did not cause any apparent side reactions during the cyclization confirming the chemoselective nature of the CyClick Chemistry. In comparison, the conventional lactamization conditions (HATU, DIEA, DMF; DEPBT, DIEA, DMF; PyBOP, DIEA, DMF) have been reported to form linear dimers[20] along with modifications to the reactive side chains (Asp, Ser) of the amino acids[20]. We also attempted to synthesize a cyclic tetrapeptide from linear pentapeptide aldehyde AFGAA **1k** without the turn inducer. MonoCyClick product was not detected, but formation of

biCyClick dimers was confirmed by the reduction with sodium cyanoborohydride (Supplementary Figure 7, Supporting Information).

Since the CyClick method works in an exclusively intramolecular fashion, we attempted head-to-tail cyclization of linear peptide AGPFA **1b** at high concentrations (25 mM, 25 times of other methods reported in literature)[10] under the optimized conditions. The reaction resulted in the formation of monocyclic tetrapeptide cyc-4-Imz-A(GPFA) **2b** with high conversion (89.5%) in 16 h with minor monoCyClick to cyclodimer ratio of >17:1. These results indicate that the CyClick strategy with a two-step process involving the formation of the cyclic imine with a pentapeptide (15 atoms, low energy barrier) followed by contraction of the ring by the amide backbone to generate highly strained monocyclic tetrapeptides (12 atoms) can compensate for the high energy barrier in synthesizing head-to-tail cyclic tetrapeptides by conventional approaches.

CONCLUSIONS

In conclusion, we demonstrated the applicability of exclusively intramolecular CyClick strategy in cyclizing small peptides using a volatile base for easy analysis by HPLC. We first synthesized linear pentapeptides containing C-terminal aldehydes with varying reactive amino acids at the N-terminus. We showed the formation of the desired monocyclic tetrapeptide products with all the peptides without the formation of any byproducts due to the reaction at the reactive side chains and any linear dimers or oligomers. We confirmed the formation of stable 4-imidazolidinone cyclic tetrapeptide by NMR analysis and carried out VT-NMR studies to determine the intramolecular H-bonding pattern in cyclic tetrapeptides. We cyclized linear peptides at high concentrations and generated desired cyclic tetrapeptides in high conversion, which is in contrast to the traditional strategies. The results from these studies concluded that the high activation energy barrier associated with the head-to-tail cyclization of small tetrapeptides can be overcome by cyclic imine capture (15 atoms, less activation barrier) followed by chemoselective ring contraction due to intramolecular amide backbone addition to generate strained head-to-tail cyclic tetrapeptide (12 atoms). Such chemoselective ring contraction is not possible with linear imine due to the high-energy barrier thus our CyClick method provides a novel approach for generating small cyclic peptides without the formation of any dimers or oligomers. In our laboratory, we are exploring these small cyclic tetrapeptides as inhibitors for targeting protein-protein interactions responsible for diseases such as MDM2-P53 and nonstructural proteins (nsps) responsible for SARS-Cov2 infections.

ACKNOWLEDGEMENTS

This work was supported by grants (Grant No. CHE- 1752654 and CHE-2108774) from National Science foundation (NSF). Also, this work was supported by a grant (1R35GM133719-01) from the National Institute of Health (NIH) and Winship Pilot Grant by Emory.

CONFLICT OF INTEREST

There are no conflicts of interest to declare.

DATA AVAILABILITY STATEMENT

All data needed to evaluate the conclusions in the paper are present in the paper and/or in the Supporting Information.

AUTHOR INFORMATION

Corresponding Author

monika.raj@emory.edu

https://orcid.org/0000-0001-9636-2222

REFERENCES

[1] H. Chiba, H. Agematu, K. Dobashi, T. Yoshioka, *J. Antibiot.* **1999**, *52*, 700-709.

[2] J. A. Aldrich, S. N. Senadheera, N. C. Ross, M. L. Ganno, S. O. Eans, J. P.

McLaughlin, *J. Nat. Prod.* **2013**, *76*, 433-438.

[3] Y. S. Ong, L. Gao, K. A. Kalesh, Z. Yu, J. Wang, C. Liu, Y. Li, H. Sun, S. S. Lee, *Curr Top Med Chem.* **2017**, *17*, 2302-2318.

[4] H. Zhang, S. Chen, *RSC Chem. Biol.* **2022**, *3*, 18-31.

[5] X. Jing, K. Jin, *Med. Res. Rev.* **2020**, *40*, 753-810.

[6] L. K. Buckton, M. N. Rahimi, S. R. McAlpine, *Chem. Eur. J.* **2020**, *27*, 1487-1513.

[7] V. Sarojini, A. J. Cameron, K. G. Varnava, W. A. Denny, G. Sanjayan, *Chem. Rev.* **2019**, *119*, 10318-10359.

[8] W. D. F. Meutermans, G. T. Bourne, S. W. Golding, D. A. Horton, M. R. Campitelli, D. Craik, M. Scanlon, M. L. Smythe, *Org. Lett.* **2003**, *5*, 2711-2714.

[9] M. El Haddadi, F. Cavelier, E. Vives, A. Azmani, J. Verducci, J. Martinez, *J. Pept. Sci.* **2000**, *6*, 560-570.

[10] C. T. T. Wong, H. Y. Lam, T. Song, G. Chen, X. Li, *Angew. Chem. Int. Ed.* **2013**, *52*, 10212-10215.

[11] K. A. Fairweather, N. Sayyadi, I. J. Luck, J. K. Clegg, K. A. Jolliffe, *Org. Lett.* **2010**, *12*, 3136-3139.

[12] D. Skropeta, K. A. Jolliffe, P. Turner, *J. Org. Chem.* **2004**, *69*, 8804-8809.

[13] V. Adebomi, R. D. Cohen, R. Wills, H. A. H. Chavers, G. E. Martin, M. Raj, *Angew. Chem. Int. Ed.* **2019**, *58*, 19073-19080.

[14] R. Wills, V. Adebomi, M. Raj, *ChemBioChem.* **2021**, *22*, 52-62.

[15] R. Wills, V. Adebomi, M. Raj, *Synlett* **2020**, *31*,1537-1542.

[16] A. Moulin, J. Martinez, J.-A. Fehrentz, *J. Pept. Sci.* **2007**, *13*, 1-15.

[17] R. Williams, pKa data compiled by R. Williams. [Online accessed on]. http://www.chem.wisc.edu/areas/organic/index-chem.htm. 2011. [06/12/2018].

[18] K. Kaupmees, A. Trummel, I. Leito, *Croat. Chem. Acta.* **2014**, *87*, 385-395.

[19] T. Cierpicki, J. Otlewski, *J.Biomol. NMR* **2001**, *21*, 249-261.

[20] C. Bechtler, C. Lamers, *RSC Med. Chem*. **2021**, *12,* 1325-1351.

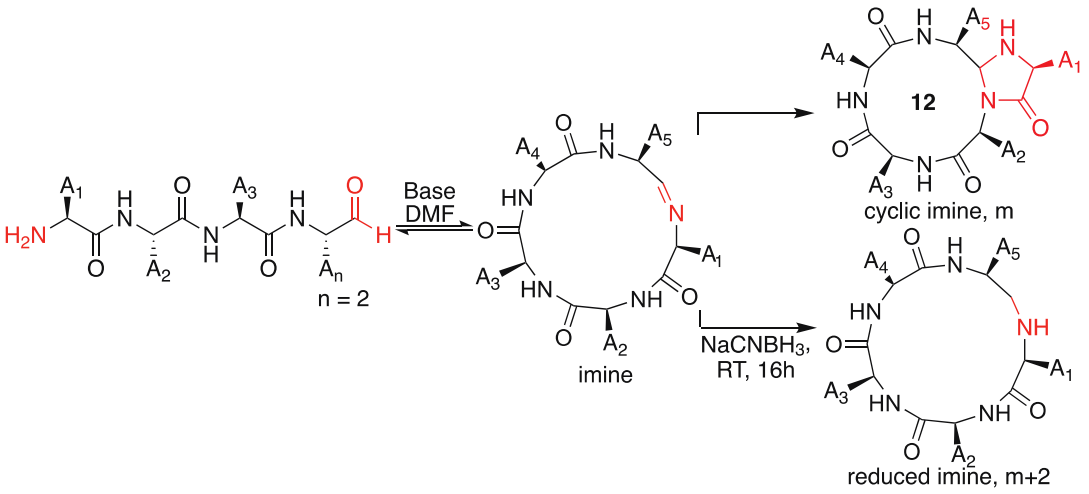


Figure 1. Scheme for the synthesis of cyclic tetrapeptides by CyClick chemistry and confirmation of CyClick 4-imidazolidinone product by reductive amination.

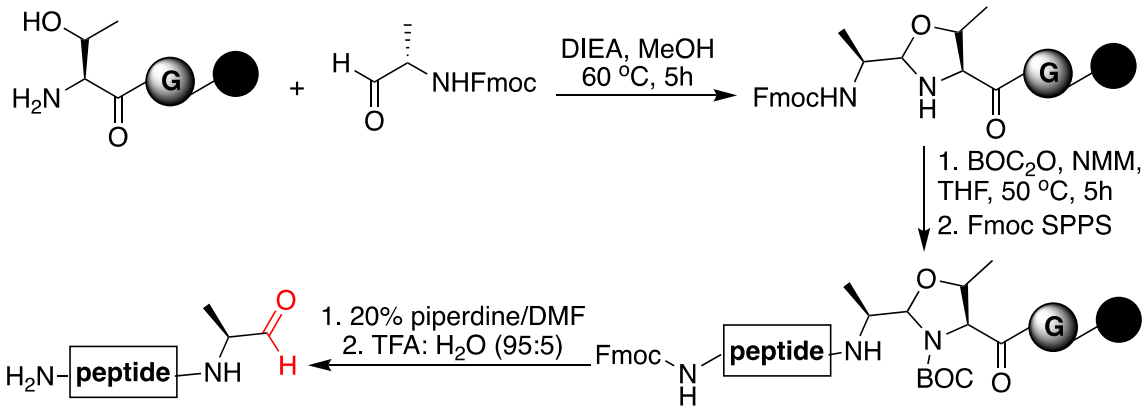
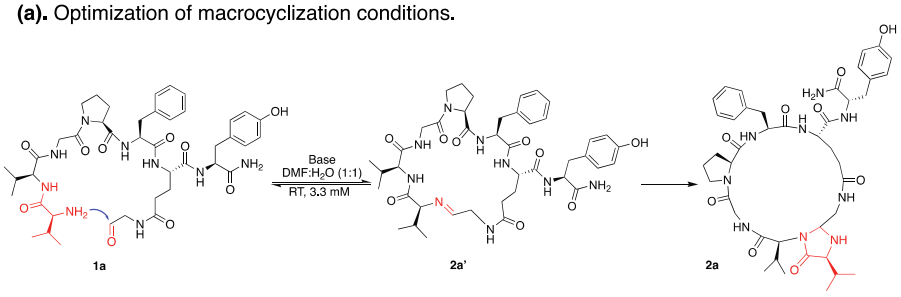


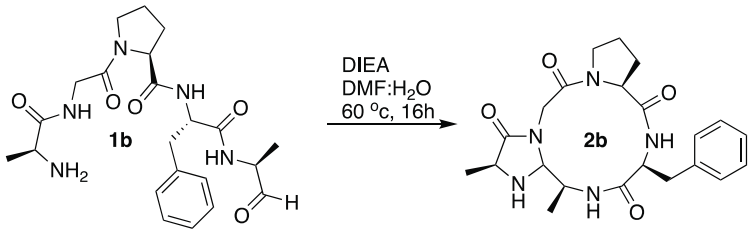
Figure 2. Synthetic scheme for the synthesis of C-terminal peptide aldehyde.

(a). Optimization of macrocyclization conditions.



Base (pKa)	Base (equiv.)	Linear (%)	mono-cyclic Imine (%)[a]	mono-CyClick **2a** (%)[a]
Imidazole (6.95)	7 equiv.	17.34	54.32	28.34
Pyrrolidine (11.27)	7 equiv.	44.17	36.00	30.48
DBU (13.5)	7 equiv.	29.88	42.44	27.68
DMAP (9.6)	7 equiv.	17.95	48.25	33.80

(b). Comparison of HPLC traces of cyclization reaction of **1b** to **2b** with DMAP and DIEA.



Base (pKa)	Base (equiv.)	Temperature	mono-CyClick **2b** (%)[a]
No Base	0 equiv.	RT	13.00
DMAP (9.6)	7 equiv.	RT	58.00
DMAP (9.6)	21 equiv.	RT	82.00
DIEA (9.6)	21 equiv.	RT	73.00
DIEA (9.6)	21 equiv.	60°C	89.00

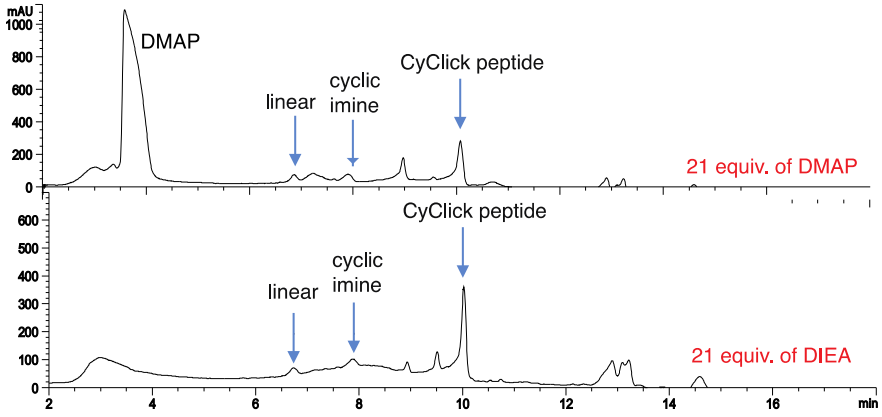
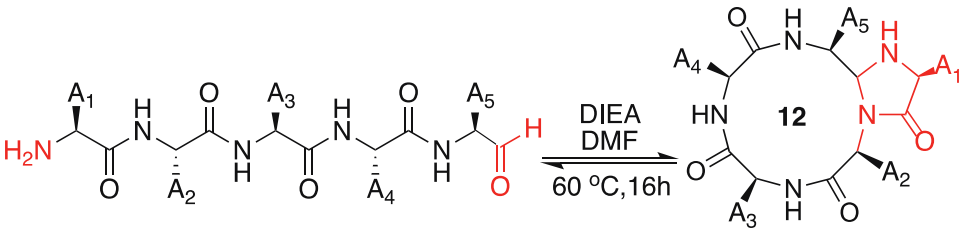


Figure 3. (a) Optimization of the cyclization by varying bases. [a]Conversion (%) is determined by sodium cyanoborohydride reduction followed by HPLC.Reaction conditions: 3.3 mM of peptide **1a** in DMF was treated with different bases and reaction was left stirring at room temperature for 16 h. (b) Comparison of the HPLC traces of the cyclization reaction with 21 equiv. of DMAP and DIEA. The use of volatile DIEA simplifies analysis of reaction mixtures.



linear sequence, **1** mono-Cyclick, **2**

linear sequence 1	mono-CyClick 2 (Conv.%)	m/z [Da]	monomer:dimer
AGPFA **1b**	cyc-4-Imz-A(GPFA) **2b** (89%)	428.2255	>99:00
GGPFA **1c**	cyc-4-Imz-G(GPFA) **2c** (73%)	414.1883	>99:00
PGPFA **1d**	cyc-4-Imz-P(GPFA) **2d** (57%)	454.2162	>99:00
SGPFA **1e**	cyc-4-Imz-S(GPFA) **2e** (21%)	444.1983	>99:00
YGPFA **1f**	cyc-4-Imz-Y(GPFA) **2f** (59%)	520.2243	>99:00
DGPFA **1g**	cyc-4-Imz-D(GPFA) **2g**(17.7%)	472.2198	>99:00
NGPFA **1h**	cyc-4-Imz-N(GPFA) **2h** (31%)	471.2367	>99:00
RGPFA **1i**	cyc-4-Imz-R(GPFA) **2i** (50%)	513.2929	>99:00
HGPFA **1j**	cyc-4-Imz-H(GPFA) **2j** (31%)	494.2510	>99:00

Figure 4. Scope of peptide cyclization for the synthesis of strained head-to-tail cyclic tetrapeptides. Reaction conditions: 3.3 mM of peptides **1b-1j** in H_2O:DMF were treated with DIEA (21 equiv.) and reaction mixtures were left for stirring at 60 °C for 16 h and analysis by HPLC to determine the (%) conversion of cyclic tetrapeptides **2b-2j.**