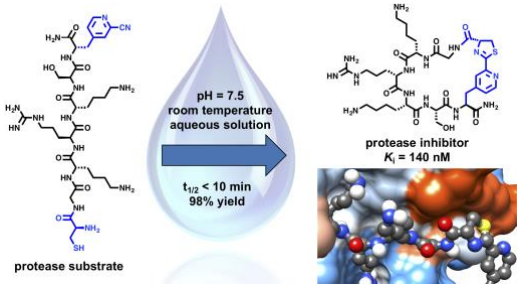
Biocompatible macrocyclization between cysteine and 2-cyanopyridine generates stable peptide inhibitors

Christoph Nitsche,*,† Hideki Onagi,† Jun-Ping Quek,‡ Gottfried Otting,† Dahai Luo,‡ Thomas Huber*,†

† Research School of Chemistry, Australian National University, Canberra, ACT 2601, Australia

‡ Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore 636921, Singapore

Supporting Information Placeholder



ABSTRACT: Peptides featuring an N-terminal cysteine residue and the unnatural amino acid 3-(2-cyano-4-pyridyl) alanine (Cpa) cyclize spontaneously in aqueous solution at neutral pH. Cpa is readily available and easily introduced into peptides using standard solid-phase peptide synthesis. The reaction is orthogonal to all proteinogenic amino acids, including cysteine residues that are not at the N-terminus. A substrate peptide of the Zika virus NS2B-NS3 protease cyclized in this way produced an inhibitor of high affinity and proteolytic stability.

Macrocyclic peptides fill an underexplored area of chemical space between small molecule therapies and larger antibodies. They constitute promising drug candidates and valuable tools in chemical biology.[1] Cyclization often enhances metabolic stability by greater resistance towards proteolysis and can promote biological uptake across cell membranes.[2] Currently available chemical strategies for macrocyclization, however, are not trivial and often require heavy metal catalysis, organic solvents, purification procedures, or other conditions that are bioincompatible.[3]

The click reaction between 1,2-aminothiols (as in an N-terminal cysteine residue) and activated nitriles is a yet underexplored chemical ligation strategy to enable conjugation and macrocyclization under biocompatible conditions,[4] facilitating applications from nanoparticle design to imaging techniques.[5] The reaction is a two-step condensation that involves an initial nucleophilic attack of the sulfhydryl group at the nitrile carbon to form a thioimidate intermediate, followed by condensation to the final dihydrothiazole with release of ammonia (Scheme 1a). In the present study we employ this chemistry to selectively cyclize peptides between an N-terminal cysteine residue and the unnatural amino acid L-3-(2-cyano-4-pyridyl)alanine (Cpa). Unlike less activated nitriles (e.g. benzonitrile), cyanopyridines react rapidly with cysteine.[4a-c] In this study, we investigated the 2-cyano-4-pyridyl substitution pattern for peptide macrocyclization. We developed a straightforward method to synthesize Fmoc-protected Cpa (**12**) in high yield from commercially available Fmoc-L-3-(4-pyridyl)alanine (~$40/g) by oxidation to the N-oxide intermediate (**11**) followed by regioselective cyanation in a total yield of 58% over two steps (Scheme 1b).[6]

Scheme 1. (a) Condensation reaction between 2-cyanopyridine and N-terminal cysteine to 2-(pyridin-2-yl)-4,5-dihydrothiazole. (b) Two-step synthesis of Fmoc-L-3-(2-cyano-4-pyridyl)alanine (Fmoc-Cpa, **12**) from commercially available Fmoc-L-3-(4-pyridyl)alanine. (c) Solid-phase peptide synthesis (SPPS) of peptides containing Cpa and their subsequent condensation with N-terminal cysteine to macrocyclic peptides in aqueous solution.

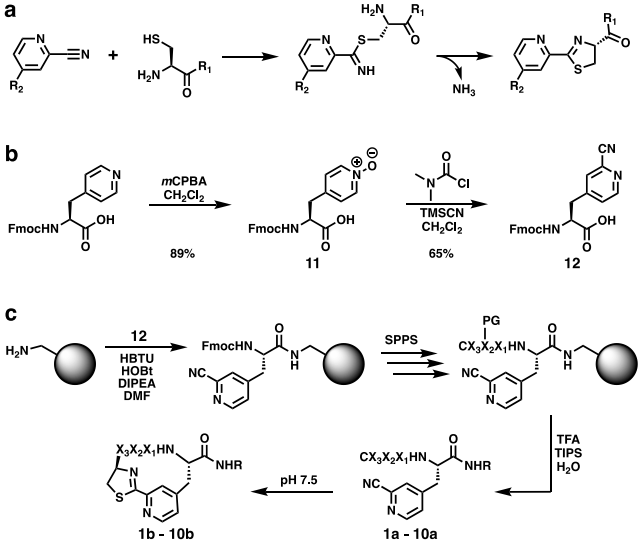
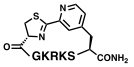
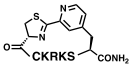
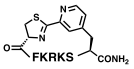
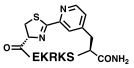
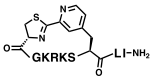
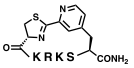
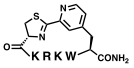
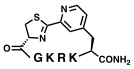
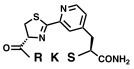
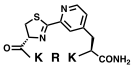


Table 1. Sequences and cyclization parameters of investigated peptides.

linear peptide[a]		cyclic peptide		cyclization parameters[c]	
cpd.	sequence[b]	cpd.	sequence	yield (%)[d]	half-time (min)[e]
1a	CGKRKS<u>Cpa</u>-NH_2	**1b**		98	8.3 ± 0.2
2a	CCKRKS<u>Cpa</u>-NH_2	**2b**		90	2.5 ± 0.1
3a	CFKRKS<u>Cpa</u>-NH_2	**3b**		97	3.4 ± 0.2
4a	CEKRKS<u>Cpa</u>-NH_2	**4b**		93	3.3 ± 0.1
5a	CGKRKS<u>Cpa</u>LI-NH_2	**5b**		96	4.4 ± 0.1
6a	CKRKS<u>Cpa</u>-NH_2	**6b**		92	8.5 ± 0.1
7a	CKRKW<u>Cpa</u>-NH_2	**7b**		97	7.3 ± 0.1
8a	CGKRK<u>Cpa</u>-NH_2	**8b**		97	8.4 ± 0.1
9a	CRKS<u>Cpa</u>-NH_2	**9b**		91	2.7 ± 0.1
10a	CKRK<u>Cpa</u>-NH_2	**10b**		94	2.8 ± 0.1

[a] Peptides synthesized by Fmoc solid-phase peptide synthesis using Rink amide resin. [b] Cpa, L-3-(2-cyano-4-pyridyl)alanine. [c] Cyclization was monitored by LC-MS in 10 mM Tris-HCl pH 7.5, 1 mM TCEP buffer at a peptide concentration of 75 µM. [d] Yields were calculated from LC-MS data. [e] Half-times were calculated from LC-MS data assuming first-order reaction kinetics.

Compound **12** can readily be incorporated into peptides by standard Fmoc solid-phase peptide synthesis without any observed side reactions under standard coupling and cleavage conditions (Scheme 1c). Following cleavage from the solid support, the linear peptides (**1a-10a**) are isolated under acidic conditions. Subsequent exposure to buffered aqueous solution at pH 7.5 results in spontaneous cyclization (**1b-10b**) via dihydrothiazole formation (Scheme 1c).

We analyzed the macrocyclization process for ten peptides of different length and sequence (Table 1). With 5-7 amino acid residues they are just in the intriguing chemical space between small molecules and much larger, less-strained macrocycles. Figure 1 shows detailed results for the model peptide pair **1** (**a**, linear; **b**, cyclic) that comprises seven amino acids, including Cpa and the N-terminal cysteine. As expected, the cyclization reaction follows first-order kinetics. The observed half-time at pH 7.5 was below 10 minutes (Figure 1b), indicating full completion after 1 hour. No significant by-products were detected during cyclization (Figure 1c), and the absence of linear or cyclic oligomers was confirmed by mass spectrometry (Figure 1d). The high specificity of the macrocyclization is reflected by excellent cyclization yields greater than or equal to 90% for all peptides analyzed (Table 1). We observed full conversion of **1a** to **1b** after 3 hours at different pH values ranging from 6 to 8 in commonly used buffer systems (MES, Tris, HEPES, MOPS). In addition, we confirmed the tolerance for organic solvents and found that the cyclization completes within 2 hours in excess of DMF (Table S5).

We further analyzed the orthogonality of the cyclization reaction with respect to other potentially interfering amino acid side chains. The presence of a second cysteine residue next to the N-terminal cysteine (**2**) only marginally reduces the reaction yield from 98% to 90% compared to a glycine residue in the same position (**1**). A bulkier phenylalanine side chain (**3**) or a negatively charged glutamate (**4**) next to the N-terminal cysteine as well as C-terminal extensions beyond the Cpa residue (**5**) are equally well tolerated. The cyclization half-times of all macrocycles comprised of seven amino acids (**1b**-**5b**) were between 2.5 and 8.3 minutes. We investigated three peptides that contain only six amino acids (**6**-**8**) and found that, on average, they cyclize only insignificantly slower (half-times between 7.3 and 8.5 min). Selective cyclization of these constrained peptides was also not impacted by a bulky and hydrophobic tryptophan side chain directly next to the Cpa residue (**7**). We further generated constrained macrocycles that contain only five amino acids (including the cysteine and Cpa residues), which cyclized even faster than those comprised of six residues with reaction half-times below 3 minutes (**9**, **10**).

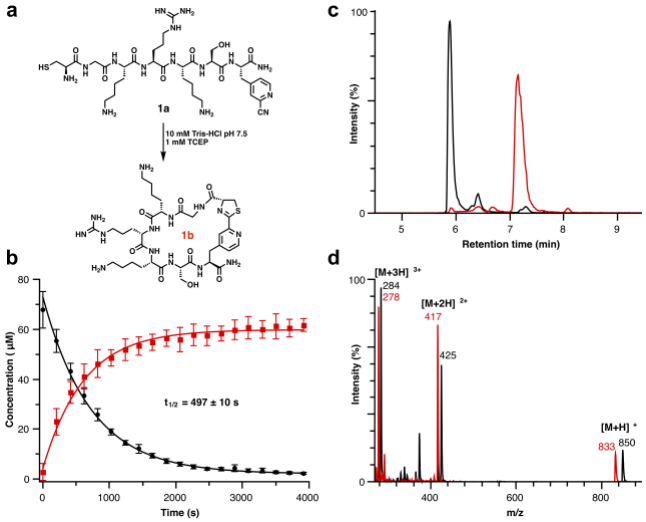


Figure 1. Selective macrocyclization of peptide **1a** to **1b** in aqueous solution (pH 7.5) at a concentration of 75 µM. (a) Scheme of spontaneous intramolecular condensation of compound **1a** to form macrocycle **1b**. (b) First-order reaction kinetics of the transformation of **1a** (black) into **1b** (red) with a reactant half-life time of 8.3 minutes. Measurements were performed in triplicate with error bars representing one standard deviation. The continuous lines correspond to mono-exponential fits. (c) Superimposed LC-MS chromatograms of **1a** (black) and its subsequent transformation into **1b** (red) after 60 min of incubation demonstrating quantitative conversion. (d) Superimposed mass spectra of **1a** (black) and **1b** (red) corresponding to the LC-MS chromatograms presented in (c).

The presented peptides are part of a series of protease inhibitors targeting Zika and related flaviviruses. The Zika virus protease NS2B-NS3 (ZiPro) is responsible for the procession of a viral precursor-polyprotein that is essential for viral replication in host cells.[7] Following the nomenclature for protease subsites introduced by Schechter and Berger,[8] the proteases from Zika and related flaviviruses recognize basic side chain residues (arginine or lysine) in the P_1–P_3 and smaller uncharged residues in the P_4 and P_1' positions.[9] Macrocyclic

peptides that act as allosteric inhibitors of ZiPro have recently been reported,[10] but no cyclic substrate-analogs that bind to the active site are known. Using the method described here, we readily cyclized the ZiPro substrate spanning residues P_4 to P_1' (GKRKS) by adding N-terminal cysteine and C-terminal Cpa residues to the sequence (**1a**). This simple procedure generated a competitive active-site inhibitor of ZiPro of remarkable affinity (K_i = 0.14 μM, Figure 2d) and stability. Although cyclic peptide **1b** reflects the substrate sequence, the proteolytic cleavage is very slow, with only about half of cyclic **1b** being processed by ZiPro to the linear peptide **1c** after 20 h of incubation (Figure 2a-c, Figure S1a). In striking contrast, the half-life of the linear ZiPro substrate (Bz-nKKR-AMC; n, norleucine; AMC, 7-amino-4-methylcoumarin) under identical conditions is less than one minute (Figure S1b). To assess the conformational constraint induced by macrocyclization, we assigned the ^{1}H and ^{13}C NMR resonances of **1b** (Table S3), measured $^3J(H_N,H_\alpha)$ couplings of the peptide backbone (Table S4), and analyzed the data with the dihedral angle prediction program TALOS+,[11] which indicated a disordered structure. Despite the apparent flexibility of the macrocycle, the cyclization of **1a** to **1b** may pre-organize the substrate residues for recognition yet compromise the optimal positioning of the cleavage site in the proteolytic center.

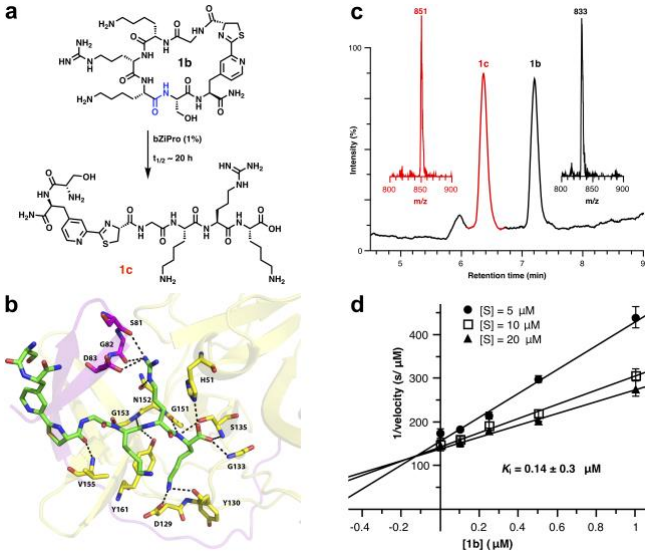


Figure 2. Macrocyclic peptide **1b** interacting with the Zika virus protease NS2B-NS3 (bZiPro). (a) Scheme of proteolytic digestion of **1b** to **1c** by bZiPro with a half-life of approximately 20 h. The proteolytic cleavage site between the P_1 lysine and P_1' serine residues is highlighted in blue. (b) Crystal structure of bZiPro in complex with **1c** at a resolution of 2.0 Å (PDB code: 6JPW). NS2B (magenta) and NS3 (yellow) are shown as ribbons and **1c** (green) is represented as sticks. The bZiPro residues involved in interactions with **1c** are labeled and shown as sticks and described in detail in the Supporting Information. (c) LC-MS chromatogram and corresponding mass spectra after 20 h of incubation of **1b** (5 μM) in the presence of 1% bZiPro (50 nM) in 10 mM Tris-HCl pH 8.5, 20% (v/v) glycerol, 1 mM CHAPS. About half of **1b** was proteolytically digested to **1c** after 20 h of incubation. (d) Dixon plot for macrocyclic peptide **1b** and bZiPro, demonstrating a competitive inhibition mechanism with a K_i value of 0.14 μM. Measurements were performed in triplicate in 10 mM Tris-HCl pH 8.5, 20% (v/v) glycerol, 1 mM CHAPS using Bz-nKKR-AMC (n, norleucine; AMC, 7-amino-4-methylcoumarin) as fluorescent substrate. Error bars indicated one standard deviation.

We set out to determine the crystal structure of ZiPro (prepared from a construct in which NS2B and NS3 are unlinked, referred to as bZiPro)[12] in complex with the cyclic peptide **1b**, but observed the cleavage product **1c** (Figure 2b, Figure S2). The asymmetric unit contained four bZiPro molecules, three of which were occupied by peptide **1c** with nearly identical binding modes (Figure S3). Comparison of the complex with previously reported bZiPro–ligand co-crystal structures containing substrate-derived ligands reveal similar P_1 and P_2 interactions and P_1–P_3 backbone conformations (Figure S4).[13] Instead of adopting a horse-shoe like conformation as observed previously, the P_3 lysine side chain of **1c** is flipped (Figure S4b) and the main chain residues exit bZiPro at a hydrophobic grove (Figure S2d).

The phenomenon that the proteolytic cleavage product remains bound to bZiPro has been described recently for a linear substrate-derived peptide.[13b] In contrast to linear peptides, where the prime-site residues dissociate after cleavage, however, all parts of the bZiPro inhibitor **1c** continue to engage in binding interactions with the protease after proteolysis of the cyclic inhibitor **1b**. As **1b** gets proteolytically cleaved only very slowly ($t_{1/2}$ ~ 20 h) and remains bound to the catalytically active center after cleavage, it has promising characteristics for lead compounds.

In conclusion, the condensation reaction of cysteine with 3-(2-cyano-4-pyridyl)alanine represents an extraordinarily straightforward and selective strategy for peptide cyclization. The reaction is orthogonal to all proteinogenic amino acids, proceeds very quickly at neutral pH and can be employed to generate constrained macrocycles in high yield. Synthesis of the unnatural amino acid Cpa is straightforward and inexpensive, and Fmoc-Cpa (**12**) is easily accessible and suited for automatic solid-phase peptide synthesis, making peptide cyclization amenable to researchers with limited access to synthetic chemistry. With growing demand for constrained peptide macrocycles in the biosciences,[1] the method presented here can be expected to find many applications, including drug design campaigns for targets with shallow ligand binding sites such as the Zika virus NS2B-NS3 protease.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

Experimental details and supporting data for chemical synthesis, kinetic and inhibition assays, protein expression and crystallization, crystallographic and NMR data, Tables S1-S4, Figures S1-S4.

AUTHOR INFORMATION

Corresponding Author

* christoph.nitsche@anu.edu.au
* t.huber@anu.edu.au

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

ORCID

Christoph Nitsche: 0000-0002-3704-2699

Hideki Onagi: 0000-0003-0085-7349
Jun-Ping Quek: 0000-0002-3827-8580
Gottfried Otting: 0000-0002-0563-0146
Dahai Luo: 0000-0001-7637-7275
Thomas Huber: 0000-0002-3680-8699

ACKNOWLEDGMENT

C.N. thanks the Alexander von Humboldt Foundation for a Feodor Lynen Fellowship and the Australian Research Council for a Discovery Early Career Research Award (DE190100015). D.L. acknowledges the support from a start-up grant from Lee Kong Chian School of Medicine, Nanyang Technological University and Singapore National Research Foundation grant NRF2016NRF-CRP001-063. We thank scientists from Australian Light Source MX beam-lines for their help on diffraction data collection.

REFERENCES

1. Morrison, C., Constrained peptides' time to shine? *Nat. Rev. Drug Discov.* **2018,** *17*, 531–533.
2. Bhat, A.; Roberts, L. R.; Dwyer, J. J., Lead discovery and optimization strategies for peptide macrocycles. *Eur. J. Med. Chem.* **2015,** *94*, 471–479.
3. White, C. J.; Yudin, A. K., Contemporary strategies for peptide macrocyclization. *Nat. Chem.* **2011,** *3*, 509–524.
4. (a) Berteotti, A.; Vacondio, F.; Lodola, A.; Bassi, M.; Silva, C.; Mor, M.; Cavalli, A., Predicting the reactivity of nitrile-carrying compounds with cysteine: a combined computational and experimental study. *ACS Med. Chem. Lett.* **2014,** *5*, 501–505; (b) Maltsev, O. V.; Walter, V.; Brandl, M. J.; Hintermann, L., Medium buffer effects on the condensation of L-cysteine and aryl nitriles to (*R*)-2-aryl-4,5-dihydrothiazole-4-carboxylic acids. *Synthesis* **2013,** *45*, 2763–2767; (c) Toyama, T.; Saitoh, T.; Takahashi, Y.; Oka, K.; Citterio, D.; Suzuki, K.; Nishiyama, S., Click reaction based on the biosynthesis of firefly luciferin. *Chem. Lett.* **2017,** *46*, 753–755; (d) Yuan, Y.; Liang, G., A biocompatible, highly efficient click reaction and its applications. *Org. Biomol. Chem.* **2014,** *12*, 865–871.
5. (a) Cheng, Y.; Peng, H.; Chen, W.; Ni, N.; Ke, B.; Dai, C.; Wang, B., Rapid and specific post-synthesis modification of DNA through a biocompatible condensation of 1,2-aminothiols with 2-cyanobenzothiazole. *Chem. Eur. J.* **2013,** *19*, 4036–4042; (b) Deng, Y.; Liu, S.; Mei, K.; Tang, A. M.; Cao, C. Y.; Liang, G. L., Multifunctional small molecule for controlled assembly of oligomeric nanoparticles and crosslinked polymers. *Org. Biomol. Chem.* **2011,** *9*, 6917–6919; (c) Inoue, M.; Tamura, T.; Yoshimitsu, Y.; Hohsaka, T.; Watanabe, T. Peptide compound and method for producing same, composition for screening use, and method for selecting peptide compound. WO2018174078, 2018; (d) Liang, G.; Ren, H.; Rao, J., A biocompatible condensation reaction for controlled assembly of nanostructures in living cells. *Nat. Chem.* **2010,** *2*, 54–60; (e) Ren, H.; Xiao, F.; Zhan, K.; Kim, Y. P.; Xie, H.; Xia, Z.; Rao, J., A biocompatible condensation reaction for the labeling of terminal cysteine residues on proteins. *Angew. Chem. Int. Ed.* **2009,** *48*, 9658–9662; (f) Sinha, S.; Ahire, D.; Wagh, S.; Mullick, D.; Sistla, R.; Selvakumar, K.; Cortes, J. C.; Putlur, S. P.; Mandlekar, S.; Johnson, B. M., Electrophilicity of pyridazine-3-carbonitrile, pyrimidine-2-carbonitrile, and pyridine-carbonitrile derivatives: a chemical model to describe the formation of thiazoline derivatives in human liver microsomes. *Chem. Res. Toxicol.* **2014,** *27*, 2052–2061; (g) Wang, Y.; Hu, X.; Weng, J.; Li, J.; Fan, Q.; Zhang, Y.; Ye, D., A photoacoustic probe for the imaging of tumor apoptosis by caspase-mediated macrocyclization and self-assembly. *Angew. Chem. Int. Ed.* **2019,** *58*, 4886–4890; (h) Ye, D.; Pandit, P.; Kempen, P.; Lin, J.; Xiong, L.; Sinclair, R.; Rutt, B.; Rao, J., Redox-triggered self-assembly of gadolinium-based MRI probes for sensing reducing environment. *Bioconjug. Chem.* **2014,** *25*, 1526–1536; (i) Ye, D.; Shuhendler, A. J.; Cui, L.; Tong, L.; Tee, S. S.; Tikhomirov, G.; Felsher, D. W.; Rao, J., Bioorthogonal cyclization-mediated *in situ* self-assembly of small-molecule probes for imaging caspase activity *in vivo*. *Nat. Chem.* **2014,** *6*, 519–526; (j) Yuan, Y.; Zhang, J.; Wang, M.; Mei, B.; Guan, Y.; Liang, G., Detection of glutathione *in vitro* and in cells by the controlled self-assembly of nanorings. *Anal. Chem.* **2013,** *85*, 1280–1284; (k) Ye, D.; Liang, G.; Ma, M. L.; Rao, J., Controlling intracellular macrocyclization for the imaging of protease activity. *Angew. Chem. Int. Ed.* **2011,** *50*, 2275–2279.
6. Fife, W. K., Regioselective cyanation of pyridine 1-oxides with trimethylsilanecarbonitrile: a modified Reissert-Henze reaction. *J. Org. Chem.* **1983,** *48*, 1375–1377.
7. (a) Nitsche, C., Strategies towards protease inhibitors for emerging flaviviruses. *Adv. Exp. Med. Biol.* **2018,** *1062*, 175-186; (b) Nitsche, C., Proteases from dengue, West Nile and Zika viruses as drug targets. *Biophys. Rev.* **2019**, *11*, 157–165.
8. Schechter, I.; Berger, A., On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun.* **1967,** *27*, 157–162.
9. (a) Nitsche, C.; Holloway, S.; Schirmeister, T.; Klein, C. D., Biochemistry and medicinal chemistry of the dengue virus protease. *Chem. Rev.* **2014,** *114*, 11348–11381; (b) Gruba, N.; Martinez, J. I. R.; Grzywa, R.; Wysocka, M.; Skoreński, M.; Dabrowska, A.; Łęcka, M.; Suder, P.; Sieńczyk, M.; Pyrc, K.; Lesner, A., One step beyond: design of substrates spanning primed positions of Zika virus NS2B-NS3 protease. *ACS Med. Chem. Lett.* **2018,** *9*, 1025–1029.
10. Nitsche, C.; Passioura, T.; Varava, P.; Mahawaththa, M. C.; Leuthold, M. M.; Klein, C. D.; Suga, H.; Otting, G., *De novo* discovery of nonstandard macrocyclic peptides as noncompetitive inhibitors of the Zika virus NS2B-NS3 protease. *ACS. Med. Chem. Lett.* **2019,** *10*, 168–174.
11. Shen, Y.; Delaglio, F.; Cornilescu, G.; Bax, A., TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* **2009,** *44*, 213–223.
12. Zhang, Z.; Li, Y.; Loh, Y. R.; Phoo, W. W.; Hung, A. W.; Kang, C.; Luo, D., Crystal structure of unlinked NS2B-NS3 protease from Zika virus. *Science* **2016,** *354*, 1597–1600.
13. (a) Phoo, W. W.; Li, Y.; Zhang, Z.; Lee, M. Y.; Loh, Y. R.; Tan, Y. B.; Ng, E. Y.; Lescar, J.; Kang, C.; Luo, D., Structure of the NS2B-NS3 protease from Zika virus after self-cleavage. *Nat. Commun.* **2016,** *7*, 13410; (b) Phoo, W. W.; Zhang, Z.; Wirawan, M.; Chew, E. J. C.; Chew, A. B. L.; Kouretova, J.; Steinmetzer, T.; Luo, D., Structures of Zika virus NS2B-NS3 protease in complex with peptidomimetic inhibitors. *Antiviral Res.* **2018,** *160*, 17–24.