Small Molecule Peptidomimetic Trypsin Inhibitors: Validation Of An EKO Binding Mode, But With A Twist

Journal:	*Organic & Biomolecular Chemistry*
Manuscript ID	OB-COM-10-2021-002127.R1
Article Type:	Communication
Date Submitted by the Author:	24-Nov-2021
Complete List of Authors:	Lyu, Rui-Liang; Texas A&M University, Chemistry Joy, Shaon; Texas A&M University, Chemistry Packianathan, Charles; Texas A&M University Laganowsky, Arthur; Texas A&M University, Department of Chemistry Burgess, Kevin; Texas A&M University College Station, Chemistry

SCHOLARONE™
Manuscripts

Small Molecule Peptidomimetic Trypsin Inhibitors: Validation Of An EKO Binding Mode, But With A Twist

Rui-Liang Lyu, Shaon Joy, Charles Packianathan, Arthur Laganowsky, and Kevin Burgess*

Department of Chemistry, Texas A & M University, Box 30012, College Station, TX 77842-3012, USA.

E-mail: burgess@tamu.edu

TITLE RUNNING HEAD: small molecule peptidomimetic inhibitor

ABSTRACT

Examination of a series of naturally-occurring trypsin inhibitor proteins, led to identification of a set of three residues (which we call the "interface triplet") to be determinant of trypsin binding affinity, hence excellent templates for small molecule mimicry. Consequently, we attempted to use the *Exploring Key Orientation* (EKO) strategy developed in our lab to evaluate small molecules that mimic the interface triplet regions of natural trypsin inhibitors, hence potentially might bind and inhibit the catalytic activity of trypsin. A bis-triazole scaffold ("TT-mer") was the most promising of the molecules evaluated *in silico*. Twelve such compounds were synthesized and assayed against trypsin, among which the best showed a K_d of 2.1 μM. X-ray crystallography revealed a high degree of matching between an illustrative TT-mer's actual

binding mode and that of the mimics that overlaid the interface triplet in the crystal structure. Deviation of the third side chain from the PPI structure seems to be due to alleviation of an unfavorable a *dipole-dipole* interaction in the small molecule's actual bound confmation.

INTRODUCTION

High affinity inhibitors to cationic trypsin-1 are desirable because overactivation of trypsin, and inactivation of endogenous trypsin inhibitors, causes inflammation, fibrosis, and pancreatic dysfunction.[1] Furthermore, trypsin inhibitors are emblematic of drugs used to treat maladies involving other members of the trypsin family. For example, *Momordica cochinchinensis* trypsin inhibitor II (MCoTI-II), a natural trypsin-inhibitory miniprotein, has been adapted to inhibit pathogenic proteases including β-tryptase (implicated in asthma),[2] neutrophil elastase (chronic obstructive pulmonary disorder),[3] coagulation factor XIIa (thrombosis),[4] kallikrein-related peptidase 4 (KLK4, prostate cancer),[5] and matriptase (cancers).[6] Novartis adapted a small molecule trypsin ligand to inhibit complement factor D, a structurally similar protease and a major component in immune activation;[7] dysfunction of factor D leads to disorders including age-related macular degeneration. Other trypsin-family proteases of therapeutic interest include urokinase plasminogen activator (uPA) and plasmin (important in cancer metastasis), thrombin, factor VII, and factor X (regulate blood coagulation). In view of this therapeutic impact, we set out to develop peptidomimetic inhibitors to the flagship, trypsin. Trypsin was chosen because its active site structure is well understood, and there are numerous high affinity natural ligands to serve as design starting points. However, our prime motivation was as a test case for our *Exploring Key Orientations* (EKO) strategy.[8]
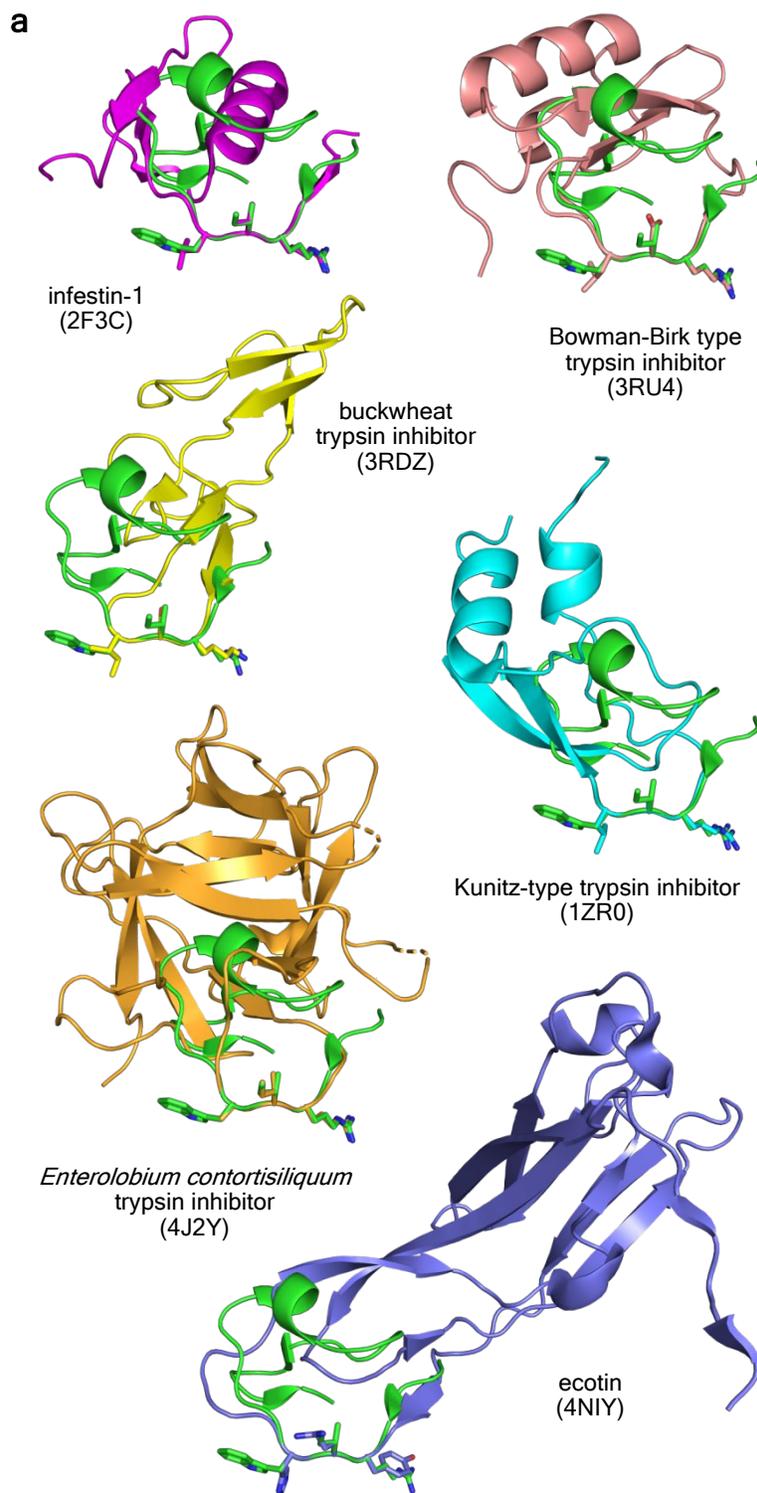
EKO compares PPI interface regions with favored small molecule conformations that present three amino acid side chains, based on degree of fit of side chain Cα and Cβ coordinates. Validation for EKO has been reported for the HIV-1 protease dimer,[8] antithrombin dimer,[9] PCSK9•LDLR,[10] and uPA•uPAR.[11,12] However, none of those cases are supported by

2

crystallographic data to confirm the actual binding mode of the small molecules. Trypsin is relatively easy to crystallize, so we anticipated it could be co-crystalized to obtain the first structure of a small molecule evaluated by EKO bound to a protein receptor.

RESULTS AND DISCUSSION

Structural Commonalities Between Naturally-Occurring Trypsin Inhibitors

We examined interface regions of natural trypsin inhibitors (Fig 1) and observed a conformational similarity in a trypsin-interface binding segment common to all. This is somewhat surprising in view of the diverse global structures and evolutionary origins of the parent protein. This common segment, which we refer to as the "interface triplet", comprises residues occupying the S1, S1', and S2' pockets in trypsin's active site, which appear to dominate the interaction energy in the protein-protein interaction (PPI).

a

infestin-1
(2F3C)

Bowman-Birk type
trypsin inhibitor
(3RU4)

buckwheat
trypsin inhibitor
(3RDZ)

Kunitz-type trypsin inhibitor
(1ZR0)

Enterolobium contortisiliquum
trypsin inhibitor
(4J2Y)

ecotin
(4NIY)

4
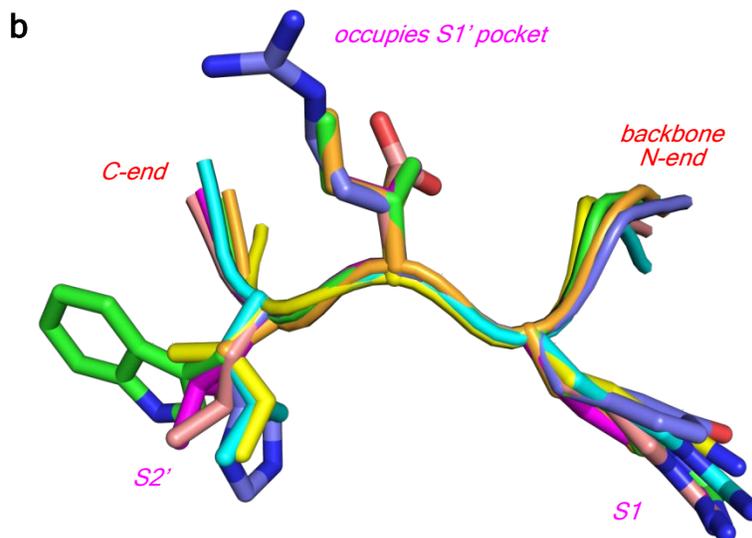
Figure 1. Overlay of the interface regions of natural trypsin inhibitors showing similarities in the trypsin-binding segment, *ie* the highly conserved PPI interface. **a** *Momordica charantia* trypsin inhibitor A (MCTI-A, green, PDB ID 1F2S) is overlaid on six other trypsin inhibitors. **b** Enlarged view of the overlay with side chains of the P1, P1', P2' residues are highlighted as sticks.

Peptidomimetic Design

Amides between amino acids in dipeptides can often be effectively substituted by five-membered ring motifs that replace four consecutive peptide backbone atoms, either C-Cα-N-C or N-C-Cα-N, and "stitching" them together by an additional fifth atom (Fig 2a). These five-membered rings may be saturated (illustrative[13-24]) or unsaturated (*eg*[25-27]) where the optimal for is determined by the degree of curvature in the bound state of the parent peptide. Tripeptides can be similarly mimicked by joining two such five-membered rings.

The strategy explored in this work featured two chemically synthesizable five-membered rings, such as triazole, hydantoin, oxazole, or oxazolidinone to mimic the interface triplet referred to above. EKO was deployed to evaluate how well these scaffolds can adopt conformations corresponding to inhibitor•trypsin PPI interface triplet region. Thus, each tripeptide mimic was installed with three methyl side chains (as the R groups in Fig 2a) before a molecular dynamics routine (quenched molecular dynamics, QMD)[28,29] was employed to sample its conformers. These conformers were clustered based on methyl side chains orientations, and the lowest energy conformer of each cluster was selected as a representative. The potential

energy of each representative conformer was calculated, and those more than 3 kcal/mol higher than the lowest-energy one were discarded because they are unlikely to be populated. The remaining stable conformers were then overlaid on the peptide template using the Kabsch algorithm. Overlay "goodnesses of fit" were quantified in terms of the root mean square deviation (RMSD) of the Cα-Cβ vectors of the three methyl side chains to that of the peptide-segment template. In our experience, an RMSD less than 0.5 Å may be considered a good overlay.

Among all the tripeptide mimics evaluated (summary in SI), one that consists of two triazoles, which we colloquially refer to as the "TT-mer", emerged as a fine target for experimental evaluation since it appeared easy to synthesize (Fig 2b) and overlaid on the interface triplet with a RMSD of 0.37 Å (Fig 2c).
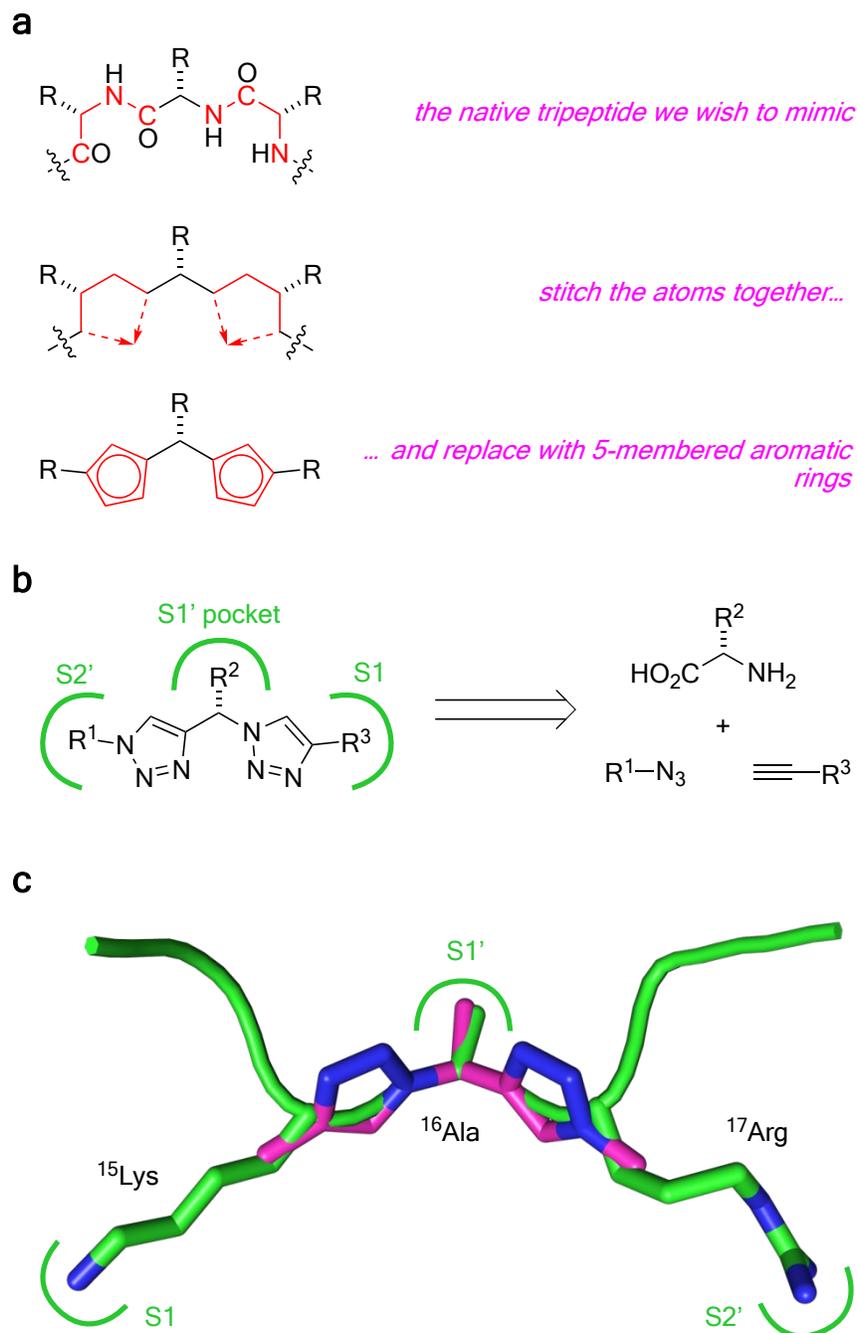
Figure 2. Design and EKO evaluation of target compounds. **a** The peptide backbone is replaced with two 5-membered aromatic rings. **b** The structure and retrosynthesis of TT-mer. **c** A simulated conformer of TT-mer overlays on the interface triplet found on bovine pancreatic trypsin inhibitor (BPTI) with an RMSD of 0.37 Å.
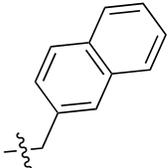
Binding Affinities of TT-mers to Trypsin

Twelve TT-mers with various side chain combinations were synthesized and their binding

affinities to *bovine* cationic trypsin measured by a kinetic assay (Zimmerman *et al*,[30] and SI).

Bovine cationic trypsin is typically used as a substitute for the human variant in early stage

inhibitor discovery (example[1]) as they have high sequence and structural similarity, and

segments around the active site (residues 189 – 215) are almost completely the same. The

compounds that gave greatest inhibition against bovine trypsin (**1k** and **1l**) were assayed

against *human* trypsin using a spectrophotometric probe, H-Glu-Gly-Arg-pNA (Table 1, and SI);

data from the two assays (using bovine and human trypsin) correlate within reasonable ranges.

These data were referenced to benzamidine positive control, known to bind trypsin with a K_d

18.4 μM, which compares well to the value of 22.2 μM in our assay H-Glu-Gly-Arg-pNA.[31]

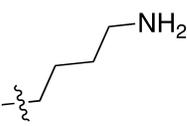**Table 1.** Binding affinities of TT-mers to bovine cationic trypsin.

ID	R^1	R^2	R^3	K_d (μM)
1a				954
1b				355
1c				>2,000

1d   1,730

1e   622

1f   410

1g   39
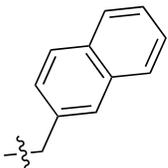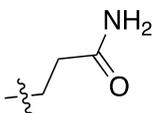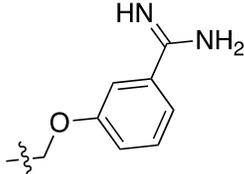
1h   28
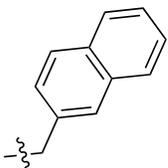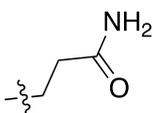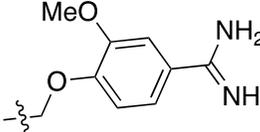
1i   23
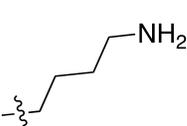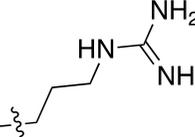
1j   342
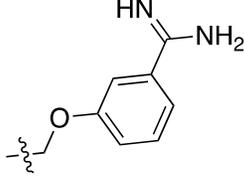
1k   26 (11)
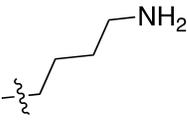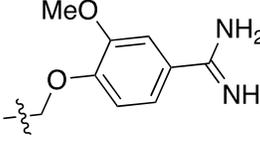
1l   2.1 (0.43)

Trends in the binding constants of the test molecules are largely uninteresting since better trypsin inhibitors exist. More importantly, the data pointed to superior binders might therefore be good candidates for co-crystallization studies. After some experimentation, a co-crystal of **1l** with trypsin obtained and the crystal structure was solved.

Binding Mode and of 1l to Trypsin

EKO evaluations will only reveal "hits" corresponding to compounds that can attain low energy conformations that overlay on the appropriate protein ligand segment. For example, in the case featured here, a conformation of **1l** overlaid R^1, R^2, R^3 side chains on the P2', P1', and P1 residues of the interface triplet, hence that compound was selected for experimental evaluation.

Compound **1l**•trypsin bound in the targeted trypsin pocket, but in a slightly different conformation (Fig 3). Side chains R^2 and R^3 occupied the S1' and S1 pockets that correspond to the interface triplet in the parent PPI used for the EKO analysis. However, the benzamidine group on R^3 formed a salt bridge with Asp189, as well as hydrogen bonds with Ser190 and the backbone carbonyl of Gly219, similar to arginine in naturally-occurring trypsin inhibitors and other benzamidine derivatives (examples[32,33]). EKO would not have evaluated this pose since it only considers small molecule conformers that overlay on interface segments of the protein ligand.
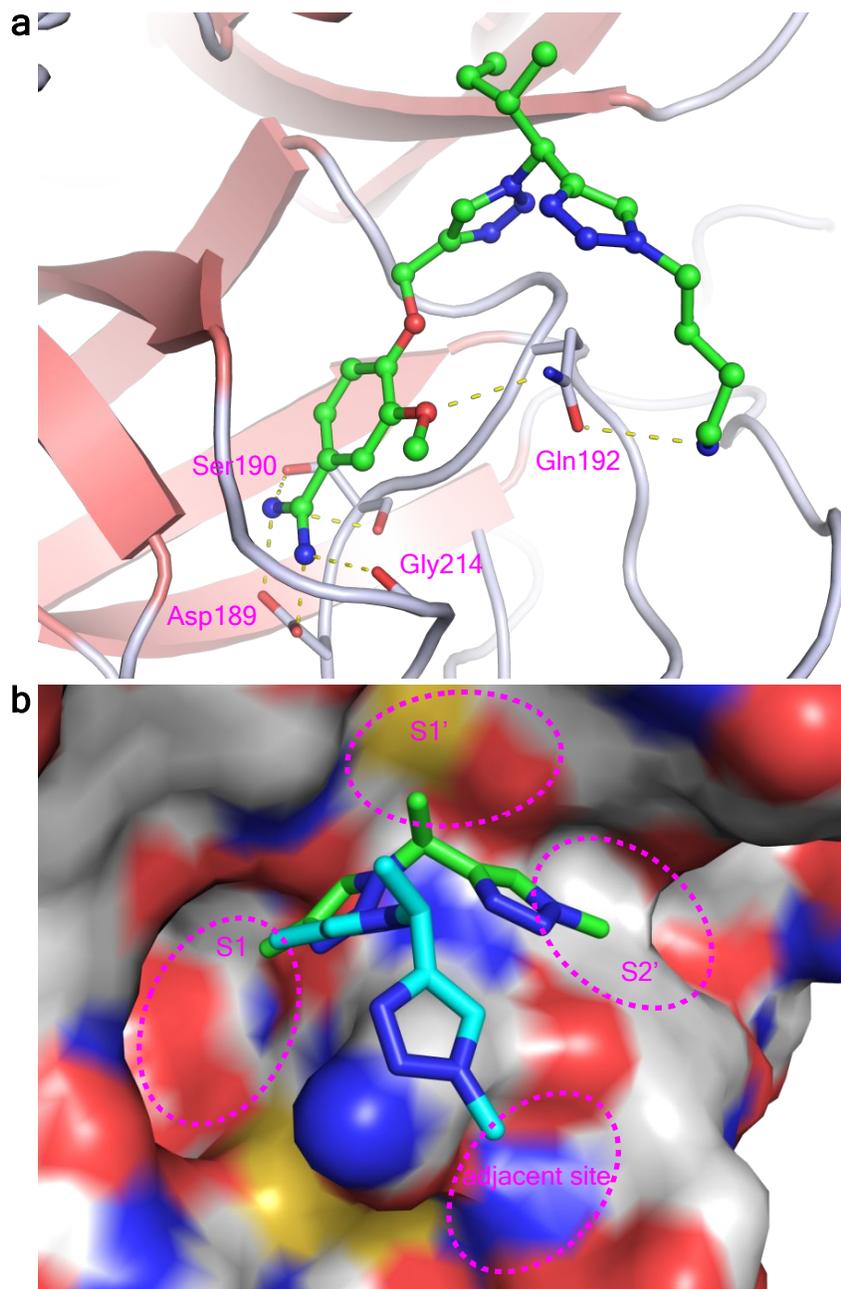
Figure 3. a X-ray structure of **1l** bound to bovine trypsin (PDB ID 7jwx). b Actual solid-state conformation (green) compared with interface triplet one as illustrated by the overlay of **1l** on this generated by EKO (cyan), (structure of the trypsin inhibitor removed for clarity).

Conformational Preference of 1l

In retrospect, the experimentally observed orientation of the R^1 side chain in the solid state structure can be attributed to a positive factor in an alternative binding mode, and a negative one which would have existed if the mimic had bound in the conformation that overlaid the triplet interface region well. Binding of the lysine ammonium cation in the R^1 residue is

positively stabilized by hydrogen bonding with the Gln192 carbonyl (Fig 3a). Conversely, the

two triazoles of **1l** may adopt "*cis*" and "*trans*" orientations (Fig 4). In the *cis*-conformation that

overlays with the interface triplet, the triazole dipoles are unfavorably aligned, while in the

trans they are favorably opposed. In other words, lone pairs on the closest nitrogen atoms in

each ring repel each other in the *cis*-state, destabilizing that conformation, while this effect is

alleviated in the *trans*-form.

EKO samples populated solution state conformations of the small molecule core backbone

(represented by the case where $R^1 – R^3$ are all methyl) to see if they will overlay with interface

protein ligand orientations in the PPI solid state structure. The key hypothesis underlying EKO

is that conformations that do overlay well will be reinforced in the protein receptor when $R^1 –$

R^3 correspond to appropriate regions of the protein ligand interface. Justification for the

assertions they are reinforced is that the protein receptor interacts with those particular side

chains in the bound conformation. Consequently, calculations of relative energies of

alternative conformations of the small molecule core backbone *in the absence of the protein*

receptor are *not* directly relevant to EKO. Nevertheless, those types of calculations were

performed here to assess stability differences for the *cis*- and *trans*-orientations of the TT-mers

in the gas phase.

Calculations using the Merck molecular force field (MMFF94) estimated the *trans*-conformer is

~5.86 kcal/mol more stable than the *cis*; QM calculations (DFT at B3LYP/3-21g, Gaussian 16)

of the same compound were also performed and these experiments gave a similar difference

to the molecular mechanics experiments: 4.845 kcal/mol (gas phase, throughout). The force

field used for molecular mechanics in EKO (AMBER10) also indicated *trans* was more stable

than *cis* but indicated an energy gap between the two forms that is below the cutoff threshold

that was applied in the featured EKO analysis (3 kcal/mol); consequently, both forms were

classified as "populated". The EKO analyses are performed using a continuous dielectric

medium of 80 wherein the negative effects of aligned dipoles would be dampened, so it is unsurprising that the predicted energy difference between the *cis*- and *trans*-TTmer states was less for the AMBER10 calculations than in the gas phase ones.
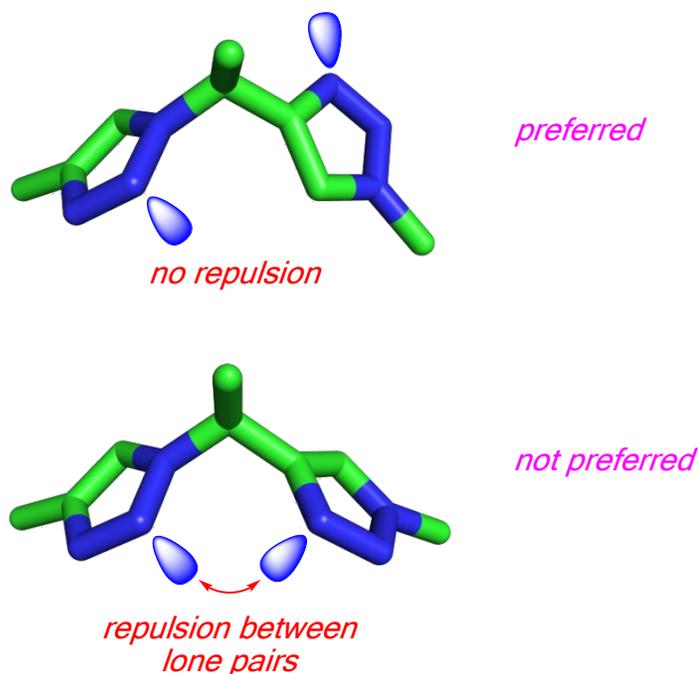

Figure 4. *Trans* (top) and *cis* (bottom) conformers of TT-mers. The repulsion between the lone pairs of triazole nitrogens makes *cis*-conformer less favorable compared to *trans*-conformers.

CONCLUSIONS

In summary, we identified a highly conserved tripeptide conformation in a number of naturally-occurring trypsin inhibitors bound to the trypsin active site. A bis-triazole scaffold ("TT-mer") was designed to mimic this triplet and was verified by EKO. A library of only 18 TT-mers were synthesized and their binding affinities against trypsin were measured. Co-crystallization of **1l** and trypsin revealed that in the solid state the bound *trans*-TT-mer conformation had two of the three side chains mimicking those of the interface tripeptide. In that *trans*-TT-mer conformation, R^3 could not overlay with the interface triplet. This alternative conformation was

favored because it leads to a favorable hydrogen-bonding interaction, and alleviates unfavorable dipole alignments in the *cis*- form.

This work successfully highlights how EKO can be used in design of small molecule drug candidates that have three side chains corresponding to protein ligand interface segments. The structural commonality between interface segments in the natural trypsin-family inhibitor proteins (Fig 1) was leveraged using EKO to validate small molecule mimicry featuring one conceptual type of core structure but different side chains to impart selectivity.

ASSOCIATED CONTENT

Electronic Supplementary Information (ESI) available: protocols for EKO modeling, synthesis, assays, and X-ray crystallography; characterization data of compounds synthesized.

AUTHOR INFORMATION

Corresponding Authors

burgess@tamu.edu

Notes

The authors declare no competing financial interests. The study was designed by RL and KB, syntheses of the compounds was performed by SJ and RL, assays and modeling was performed by RL, and CP with AL formed the co-crystals and solved the structure.

ACKNOWLEDGEMENTS

REFERENCE

(1) T. Brandl, O. Simic, P. R. Skaanderup, K. Namoto, F. Berst, C. Ehrhardt, N. Schiering, I. Mueller and J. Woelcke, *Bioorg. Med. Chem. Lett.*, 2016, **26**, 4340-4344.

(2) C. P. Sommerhoff, O. Avrutina, H.-U. Schmoldt, D. Gabrijelcic-Geiger, U. Diederichsen and H. Kolmar, *J. Mol. Biol.*, 2010, **395**, 167-175.

(3) P. Thongyoo, C. Bonomelli, R. J. Leatherbarrow and E. W. Tate, *J. Med. Chem.*, 2009, **52**, 6197-6200.

(4) J. E. Swedberg, T. Mahatmanto, H. Abdul Ghani, S. J. de Veer, C. I. Schroeder, J. M. Harris and D. J. Craik, *J. Med. Chem.*, 2016, 7287-7292.

(5) J. E. Swedberg, H. A. Ghani, J. M. Harris, S. J. de Veer and D. J. Craik, *ACS Med. Chem. Lett.*, 2018, **9**, 1258-1262.

(6) P. Quimbar, U. Malik, C. P. Sommerhoff, Q. Kaas, L. Y. Chan, Y.-H. Huang, M. Grundhuber, K. Dunse, D. J. Craik, M. A. Anderson and N. L. Daly, *J. Biol. Chem.*, 2013, **288**, 13885-13896.

(7) J. Maibaum, S.-M. Liao, A. Vulpetti, N. Ostermann, S. Randl, S. Rudisser, E. Lorthiois, P. Erbel, B. Kinzel, F. A. Kolb, S. Barbieri, J. Wagner, C. Durand, K. Fettis, S. Dussauge, N. Hughes, O. Delgado, U. Hommel, T. Gould, A. Mac Sweeney, B. Gerhartz, F. Cumin, S. Flohr, A. Schubart, B. Jaffee, R. Harrison, A. M. Risitano, J. Eder and K. Anderson, *Nat. Chem. Biol.*, 2016, **12**, 1105-1110.

(8) E. Ko, A. Raghuraman, L. M. Perez, T. R. Ioerger and K. Burgess, *J. Am. Chem. Soc.*, 2013, **135**, 167-173.

(9) D. Xin, A. Holzenburg and K. Burgess, *Chem. Sci.*, 2014, **5**, 4914-4921.

(10) J. Taechalertpaisarn, B. Zhao, X. Liang and K. Burgess, *J. Am. Chem. Soc.*, 2018, **140**, 3242-3249.

(11) M. Arancillo, J. Taechalertpaisarn, X. Liang and K. Burgess, *Angew. Chem. Int.*, 2021, **60**, 6653-6659.

(12) M. Arancillo , C. M. Lin, K. Churion and K. Burgess, *J. Med. Chem.*, 2021.

(13) S. Borg, G. Estenne-Bouhtou, K. Luthman, I. Csoregh, W. Hesselink and U. Hacksell, *J. Org. Chem.*, 1995, **60**, 3112-3120.

(14) K. Luthman, S. Borg and U. Hacksell, *Methods in Molecular Medicine*, 1999, **23**, 1-23.

(15) D. Boeglin, S. Cantel, A. Heitz, J. Martinez and J.-A. Fehrentz, *Org. Lett.*, 2003, **5**, 4465-4468.

(16) A. Brik, J. Alexandratos, Y.-C. Lin, J. H. Elder, A. J. Olson, A. Wlodawer, D. S. Goodsell and C.-H. Wong, *CHEMBIOCHEM*, 2005, **6**, 1167-1169.

(17) Z. Zhang and E. Fan, *Tetrahedron Lett.*, 2006, **47**, 665-669.

(18) N. G. Angelo and P. S. Arora, *J. Org. Chem.*, 2007, **72**, 7963-7967.

(19) V. D. Bock, D. Speijer, H. Hiemstra and J. H. Van Maarseveen, *Org. Biomol. Chem.*, 2007, **5**, 971-975.

(20) A. L. Jochim, S. E. Miller, N. G. Angelo and P. S. Arora, *Bioorg. Med. Chem. Lett.*, 2009, **19**, 6023-6026.

(21) N. Narendra, T. M. Vishwanatha and V. V. Sureshbabu, *Int. J. Pept. Res. Ther.*, 2010, **16**, 283-290.

(22) S. Petit, C. Fruit and L. Bischoff, *Org. Lett.*, 2010, ASAP.

(23) B. C. Doak, M. J. Scanlon and J. S. Simpson, *Org. Lett.*, 2011, **13**, 537-539.

(24) Z. Ke, H.-F. Chow, M.-C. Chan, Z. Liu and K.-H. Sze, *Org. Lett.*, 2012, **14**, 394-397.

(25) W. J. Hoekstra, B. L. Hulshizer, D. F. McComsey, P. Andrade-Gordon, J. A. Kauffman, M. F. Addo, D. Oksenberg, R. M. Scarborough and B. E. Maryanoff, *Bioorg. Med. Chem. Lett.*, 1998, **8**, 1649-1654.

(26) P. Garcia-Reynaga and M. S. VanNieuwenhze, *Org. Lett.*, 2008, **10**, 4621-4623.

(27) E. Biron, J. Chatterjee and H. Kessler, *Org. Lett.*, 2006, **8**, 2417-2420.

(28) B. M. Pettitt, T. Matsunaga, F. Al-Obeidi, C. Gehrig, V. J. Hruby and M. Karplus, *Biophys. J.*, 1991, **60**, 1540-1544.

(29) S. D. O'Connor, P. E. Smith, F. Al-Obeidi and B. M. Pettitt, *J. Med. Chem.*, 1992, **35**, 2870-2881.

(30) M. Zimmerman, E. Yurewicz and G. Patel, *Anal. Biochem.*, 1976, **70**, 258-262.

(31) M. Mares-Guia and E. Shaw, *Journal of Biological Chemistry*, 1965, **240**, 1579-1585.

(32) L. Muley, B. Baum, M. Smolinski, M. Freindorf, A. Heine, G. Klebe and D. G. Hangauer, *J. Med. Chem.*, 2010, **53**, 2126-2135.

(33) A. M. Said, M. W. Parker and C. W. Vander Kooi, *Bioorg. Chem.*, 2020, **100**, 103856.