Pergamon

Bioorganic & Medicinal Chemistry Letters 11 (2001) 2515–2519

BIOORGANIC & MEDICINAL CHEMISTRY LETTERS

# Synthesis and Evaluation of the Sunflower Derived Trypsin Inhibitor as a Potent Inhibitor of the Type II Transmembrane Serine Protease, Matriptase

Ya-Qiu Long,[a] Sheau-Ling Lee,[b] Chen-Yong Lin,[b] Istvan J. Enyedy,[b] Shaomeng Wang,[b] Peng Li,[a] Robert B. Dickson[b] and Peter P. Roller[a,*]

[a]*Laboratory of Medicinal Chemistry, National Cancer Institute, NIH, FCRDC, 376/208, PO Box B, Frederick, MD 21702, USA*
[b]*Lombardi Cancer Center, Georgetown University Medical Center, Washington, DC 20007, USA*

Received 1 February 2001; accepted 9 July 2001

**Abstract**—We report here the synthesis of a 14-amino acid long bicyclic peptide, previously isolated from sunflower seeds. This peptide, termed sunflower trypsin inhibitor (SFTI-1), is one of the most potent naturally occurring small-molecule trypsin inhibitors. In addition to inhibiting trypsin, the synthetic SFTI-1 is also a very potent inhibitor, with a $K_i$ of 0.92 nM, of the recently identified epithelial serine protease, termed 'matriptase'. Published by Elsevier Science Ltd.

Plant derived protease inhibitors serve in the defense mechanisms of plants against pests and plant pathogens.[1] These inhibitors can be classified into a number of families based on their active-site structures and their specificities to inhibit the cleavage of specific peptide sequences within proteins. The majority of these inhibitors are classified as serine protease inhibitors. One of the well known serine protease inhibitory agents, the Bowman–Birk inhibitor, found in seeds of legumes and other plants, belongs to the Bowman–Birk inhibitor family of small proteins with a $M_r$ range of 6000–8000.[2] These proteins inhibit trypsin, chymotrypsin, or elastase, depending on the configuration of the reactive site loop within the inhibitor.

Recently, a 14-amino acid peptide, termed sunflower trypsin inhibitor (SFTI-1) was isolated from sunflower seeds.[3] SFTI-1 inhibited β-trypsin with an impressive subnanomolar $K_i$ of 0.1 nM, and it inhibited cathepsin G with a comparable $K_i$.[3] SFTI-1 has considerable selectivity; for example, it proved to be 74-fold less inhibitory for chymotrypsin, and was found to be 3 orders of magnitude less inhibitory for elastase and thrombin. In contrast, it had no effect on Factor Xa.[3]

The natural product, SFTI-1 was partially characterized by classical techniques, and its structure confirmed, based on the electron density map of the inhibitor co-crystallized with bovine-β-trypsin.[3] SFTI-1 is one of the smallest, naturally occurring plant protein inhibitors reported to date, with a molecular mass of 1513. It has considerably enhanced potency relative to other peptides of similar length. Its backbone-cyclized peptide structure is additionally stabilized by a cystine disulfide bond. We report here the facile synthesis of SFTI-1. This methodology will also provide for the synthesis of various analogues, with altered inhibitory profiles, relative to various serine proteases of interest.

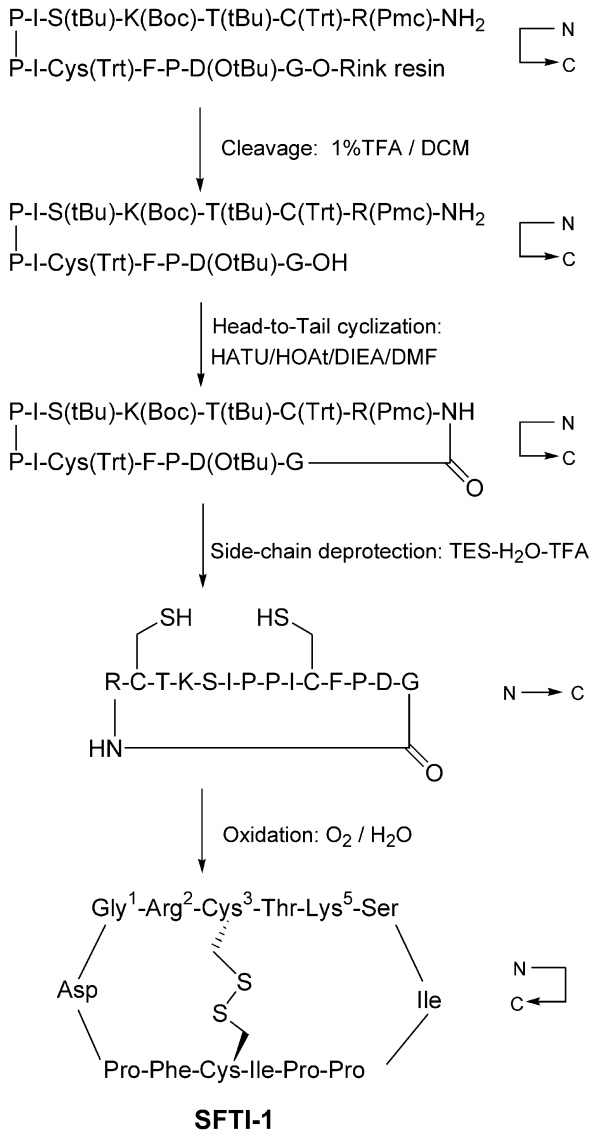
One such serine protease of interest is known as matriptase (or MT-SP1), a member of the emerging class of type II transmembrane serine proteases.[4–6] The mouse homologue of matriptase has also been described and is termed epithin.[7] Matriptase/epithin is of considerable interest for the development and pathogenesis of epithelial tissues. Although matriptase is initially synthesized by multiple types of epithelial cells as a transmembrane serine protease, it was isolated originally from human milk in its activated form complexed with its cognate Kunitz type of serine inhibitor (KSPI), the hepatocyte growth factor activator inhibitor (HAI-1). Additionally, matriptase was isolated from human breast cancer cells in culture.[4,5,8] Matriptase may function to degrade the extracellular matrix, as well as several cellular regulatory proteins; specifically, it may activate hepatocyte growth factor (HGF) by cleaving its inactive proform, it may activate urokinase by cleaving its zymogen, and it may cleave and activate the protease

*Corresponding author. Tel.: +1-301-846-5904; fax: +1-301-846-6033; e-mail: proll@helix.nih.gov

0960-894X/01/$ - see front matter. Published by Elsevier Science Ltd.
PII: S0960-894X(01)00493-0

activated receptor-2 (PAR-2).[9–11] Thus, matriptase blockade could potentially modulate cell proliferation, motility, invasion, and differentiation of cells.[6,10] We report here the synthesis of SFTI-1 and its potent enzyme inhibitory effectiveness with matriptase, in comparison to its potency for related serine proteases. We carried out molecular modeling studies to provide a structural basis for the interactions between SFTI-1 and matriptase. SFTI-1 represents a useful pharmacological tool to study the biological function of matriptase, as well as a potentially useful therapeutic agent.

Fmoc chemistry-based solid-phase peptide synthesis methodology was used for the synthesis of the linear peptide RCTKSIPPICFPDG-Rink resin. The total synthesis of bicyclic peptide SFTI-1 is described in Scheme 1. The acid sensitive 4-(2′,4′-dimethoxyphenyl-hydroxyphenyl)-phenoxy('Rink resin') was purchased from Bachem California Inc. (Torrance, CA, USA). Fmoc derivatives of standard amino acids were obtained from Perkin–Elmer/Applied Biosystems Division. Side-chain protections are as follows: Arg(Pmc), Cys(Trt), Thr(*t*-Bu), Lys(Boc), Ser(*t*-Bu), Asp(O*t*-Bu). HBTU/HOBt activation of $N^{\alpha}$-protected amino acids was employed for coupling, and 20% piperidine/DMF was used for Fmoc deprotection. HATU/HOAt/DIEA in DMF was used for backbone cyclization. The crude peptide was purified by reverse-phase high-performance liquid chromatography (RP-HPLC). HPLC conditions: Vydac C4 column (20×250 mm); solvent gradient, A, 0.05% TFA in water; B, 0.05% TFA in 90% acetonitrile in water with gradient indicated below; flow rate, 10 mL/min; UV detector, 225 nm. FABMS (unit resolution, glycerol matrix) was performed on a VG Analytical 7070E-HF mass spectrometer. Amino acid analysis (6 N HCl, 110 °C, 24 h) was carried out at the Protein and Carbohydrate Structure Facility (University of Michigan, Ann Arbor, MI, USA).




**Scheme 1.** Synthetic route for sunflower trypsin inhibitor SFTI-1.

The first amino acid, glycine, was attached to the acid labile Rink resin. Positioning Gly at the C-terminal end minimizes the possibility of racemization on subsequent backbone cyclization. Fmoc-Gly-OH (297 mg, 1.0 mmol) was double coupled with the Rink resin (357 mg, 0.1 mmol) by using 1,3-diisopropylcarbodiimide (1 mmol), 4-(dimethylamino)pyridine (0.1 mmol) and *N*-methylmorpholine (0.1 mmol) in 3 mL DMF (2 h). The Fmoc-Gly-Rink resin was dried under a vacuum and treated as a preloaded resin for automated peptide synthesis with an ABI 433A peptide synthesizer using the FastMoc protocol. On completion of the sequence RCTKSIPPICFPDG-Rink resin, the side-chain protected peptide was cleaved from the resin with 20 mL of 1% TFA/DCM (1 min). The solution was collected by filtration and neutralized with NMM in an ice bath. The resin was treated similarly two more times with 1% TFA/DCM (5 min) and the combined neutralized solution was evaporated to dryness in vacuo. For backbone cyclization, the crude side-chain protected peptide was dissolved in anhydrous DMF (80 mL), and added DIEA (0.6 mmol) and HATU/HOAt (0.30 mmol each) at ice bath temperature. The reaction mixture was stirred for 24 h, neutralized with 30% AcOH and evaporated in vacuo. The residue was treated with 95% TFA containing 2.5% each of triethylsilane (TES) and $H_2O$ to deprotect the side chains. Subsequently, for disulfide oxidization, the head-to-tail cyclic peptide was dissolved in 20 mL of water and added dropwise to 250 mL of water solution, which was previously adjusted to pH 8.5 with ammonium acetate and ammonium hydroxide. The disulfide cyclization occurred spontaneously under the basic conditions while bubbling with oxygen for 6 h, and then stirring overnight rt. The reaction was quenched by adding AcOH and the solution was lyophilized. The synthetic SFTI-1 was purified by RP-HPLC, rt = 17.6 min (gradient 20–60% B over 40 min); FABMS $(M+H)^+$ 1513.0 (calcd 1513.7). Amino acid analysis: Asp 1.02(1), Ser 0.91(1), Thr 1.15(1), Pro 2.79(3), Lys 0.97(1), Gly 1.15(1), Arg 1.16(1), Ile 1.83(2), Phe 1.02(1).

The 70-kDa activated matriptase was isolated as described previously.[5,7,8] Urokinase-type plasminogen activator (uPA) was purified by aminobenzamidine-Sepharose 6B

(Amersham Pharmacia, Piscataway, NJ, USA) from a partially purified uPA from human urine. Bovine β-trypsin, bovine thrombin, Bowman–Birk inhibitor (BBI), and the fluorescent substrates were purchased from Sigma (Sigma Chemical Co., St. Louis, MO, USA). Inhibitory activity of SFTI-1 to proteases was measured at room temperature in a reaction buffer of 100 mM Tris–HCl (pH 8.5) containing 100 mg/mL of bovine serum albumin, using the fluorescent substrate peptides. To a cuvette containing 170 μL of reaction buffer was added 10 μL of enzyme solution and 10 μL of inhibitor solution. After preincubation, 10 μL of substrate solution was added and the cuvette content mixed thoroughly. The residual enzyme activity was determined by following the change of fluorescence released by hydrolysis of the substrates, using a fluorescent spectrophotometer (Hitachi F4500) with excitation wavelength of 360 nm and emission at 480 nm. Fluorescent peptide *N*-*t*-Boc Gln-Ala-Arg-AMC was used as substrate for matriptase and trypsin, peptide *N*-*t*-Boc-Leu-Gly-Arg-AMC was used as substrate for uPA, and peptide *N*-*t*-Boc-Leu-Arg-Arg-AMC was used as substrate for thrombin. Hydrolysis rates were recorded in presence of 6–7 different concentrations of SFTI-1. The $K_i$ values were determined by Dixon plots from two sets of data with different concentrations of substrate.

The homology modeling approach was used to build the 3-D structure of matriptase. A search of the protein databank,[12] using the program BLAST,[13] has identified thrombin (PDB entry: 1hxe) as a good template for homology modeling, with 34% identities, 53% positives, and 6% gaps. The structure of matriptase was built using the program MODELLER.[14] The structure derived in this manner was refined by its solvation with TIP3P water molecules, minimization to eliminate bad contacts, and equilibration at 300 K using molecular dynamics simulation with the program CHARMM.[15] For docking purposes, only the active site of matriptase was solvated with a 20 Å radius water sphere of TIP3P water, with the center defined by the average of SFTI-1 Cys3 $S^{\gamma}$, and matriptase Ser805 $O^{\gamma}$, Ser805 $C^{\beta}$, and Ser805 $C^{\alpha}$. Only the surface residues of matriptase were considered flexible, and the rest were fixed during the simulation. The structure of SFTI-1 was obtained from the protein databank (PDB entry: 1sfi). The starting orientation of the inhibitor is the same as the one found in the X-ray structure with trypsin, 1SFI.[3] The docking of matriptase with the inhibitor was performed using molecular dynamics with Tsallis effective potential,[16,17] as implemented in CHARMM.[15] The simulation was done using the all atom parameter set CHARMM22, the temperature was set to 300 K, the time step for integration was 1 fs, the potential shift parameter was 11,000 kcal/mol, and q, the final value of the Tsallis coefficient, was 1.001. The q-annealing schedule was set to 20 ps, for increasing q from 1.0 to 1.001, 20 ps for decreasing q from 1.001 to 1.0, and 10 ps normal molecular dynamics. The total length of simulation was 4 ns. Long-range electrostatic forces were treated with the force switch method in a switching range of 8–12 Å. Van der Waals forces were calculated with the shift method and a cutoff of 12 Å. The nonbond list was kept to 14 Å, and updated heuristically. Solvent waters were kept from evaporating by using a spherical miscellaneous mean field potential as implemented in CHARMM. An NOE restraint was used on the distance between SFTI-1 Lys5 $N^{\zeta}$ and Asp799 $C^{\gamma}$ of the protein. This restraint kept Lys5 of SFTI-1 bound to the S1 pocket of matriptase; Asp799 is at the bottom of the binding pocket. The maximum value of the distance was set to 4 Å, similar to the distance in the X-ray structure of trypsin complexed with SFTI-1.[3]

## Results and Discussion

Endogenous proteases play a pivotal role in the normal cellular physiology of the cell, such as the proteolytic activation of peptide hormones and the activation of message-transmitting peptides and proteins.[18] Our efforts are focused on regulation of a recently identified protease, termed matriptase. Matriptase was initially purified from human milk, but it is also produced by normal and cancerous epithelial cells in culture. Recent studies have suggested that inappropriate expression of its active form has the potential of producing deleterious effect in tissues, contributing to pathogenic states, such as cancer.[5,8] This particular serine protease can degrade extracellular matrix proteins, and activate specific proteins, such as HGF, uPA, and PAR-2, by cleaving its inactive pro-form.[9,10] These effects are likely to contribute to abnormal cell proliferation, motility, and states of differentiation. In efforts to identify inhibitors of matriptase for experimental purposes, we evaluated the recently identified sunflower derived trypsin inhibitor (SFTI-1), and found it to be a highly effective inhibitor of the enzyme.

SFTI-1 was recently isolated from sunflower seeds as a complex with trypsin, and its structure was determined by NMR spectroscopy and by X-ray crystallography.[3] Its promising protease inhibitory profile, reported in the original work,[3] prompted us to develop the current synthetic methodology for preparation of this 14-amino acid bicyclic peptide, as outlined in Scheme 1. The amino acid backbone was assembled on Rink resin. After resin cleavage, the fully side chain-protected peptide was backbone cyclized, followed by deprotection of all side chains. Air oxidation, in weakly basic medium, smoothly provided in good yield the intramolecularly bridged cyclic peptide, SFTI-1. NMR and X-ray studies demonstrated that SFTI-1 has considerable structural rigidity,[3] imparted by the intramolecular disulfide bond. The overall conformation and the amino acid sequence was very similar to the reported structures for the equivalent trypsin-inhibitory loop of the reactive site segments within the BBI. That segment (–CTKSIPP–) is also very similar to the trypsin inhibitory segment of the mung bean inhibitory protein.[19]

Table 1 summarizes the results for a relevant set of serine proteases. The synthesized SFTI-1, just as its natural form, is a potent inhibitor of trypsin with a $K_i$ value of 1.06 nM. In addition, we found that SFTI-1 inhibited matriptase with comparable effectiveness to trypsin.

**Table 1.** Protease inhibitory properties of SFTI-1

| Protease | $K_i$ (nM)[a] |
|---|---|
| Matriptase | 0.92 |
| Trypsin | 1.06[b] |
| Thrombin | 5050 |
| uPA | 500,000 |

[a]Assay conditions as described in the text.
[b]$K_i$ of 0.1 nM was reported by Luckett et al.[3] using competitive assay conditions.

Although the modeled 3-D structure of matriptase was built using thrombin as template, SFTI-1 is a poor inhibitor of thrombin ($K_i = 5{,}050$ nM). SFTI-1 is also non-inhibitory for uPA, an important serine protease in the extracellular matrix degradative network. This selectivity of SFTI-1, thus, makes it a valuable tool to study the function of matriptase in biological matrices.

On account of the overall similarity of the secondary structure and 3-D structure between SFTI-1 and the trypsin inhibitory loop of BBI, we tested the inhibitory activity of BBI to matriptase. As shown in Figure 1, at concentrations of 1.25, 2.5, and 5 μM, BBI significantly reduced the matriptase hydrolytic activity of standard substrates, compared with the SFTI-1 at the same concentration; however, the inhibitory potency of SFTI-1 is 2- to 3-fold higher than that of BBI.

A molecular modeling study was initiated in order to better understand the structural features that contribute to the high inhibitory activity of SFTI-1. The predicted structure of matriptase complexed with SFTI-1 was obtained after docking using 4 ns molecular dynamics simulations with the Tsallis effective potential, as shown in Figure 2. This structure was compared to the crystal structure of trypsin, in complex with SFTI-1,[3] to investigate the similarities and differences between the interactions of SFTI-1 with these two proteins. In both structures, Lys5 of the SFTI-1 binds to the S1 site of protease; interacting with Asp189 in trypsin or Asp799 in matriptase. Arg2 of the SFTI-1 mainly interacts with the backbone carbonyl group of Asn97 of trypsin and solvent molecules in the X-ray structure. Arg2 forms an H-bond with the Phe706 main chain carbonyl and interacts with Phe708 and Phe706 side chains of matriptase through π-cation interactions. Asp14 in SFTI-1 is largely exposed to solvent in both structures; it forms a hydrogen bond with Asn72 in the trypsin X-ray complex structure, but mainly interacts with Arg2 of SFTI-1 in the docked matriptase structure. Phe12 of SFTI-1 is in close proximity with only one hydrophobic residue, Leu99, in the trypsin complex structure, while this residue interacts with Phe708 in the docked complex structure with matriptase. Our docking studies suggest that SFTI-1 has similar interactions with matriptase and trypsin, and these results provide a rational explanation for the comparative inhibitory potency of SFTI-1 to these two proteases.

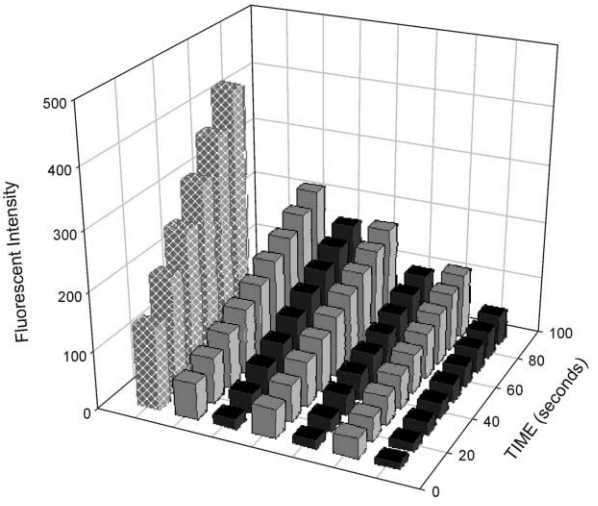



**Figure 1.** SFTI-1 is several fold more potent inhibitor of matriptase than BBI at low micromolar concentrations. The released fluorescence of the peptide substrate resulting from the proteolytic cleavage by matriptase is shown. The assays were carried out either in the absence of inhibitors (hatched bars), or in the presence of various concentrations of SFTI-1 (black bars) or of BBI (gray bars), as described in the text.

In complexes of SFTI-1 with trypsin and with matriptase the Ser6 OH and the Thr4 OH side chains of the inhibitor interact through hydrogen bonds with the catalytic triad amino acids Ser195/Ser805 and the His57/His656, of both trypsin and matriptase, respectively. These interactions may thwart the activation of the catalytic triad Ser by His57/His656 for nucleophilic attack at the scissile bond. Although the scissile bond in both cases is oriented properly for nucleophilic attack, the distance between Ser195/Ser805 $O^{\gamma}$ and SFTI-1 Lys5 $C^{\zeta}$ is 2.8 and 3.6 Å in the complex with trypsin and matriptase, respectively, longer than the requirement of a covalent bond. This suggests that the inhibitor is not hydrolyzed by the protease, which is in general agreement with the observation that SFTI-1 is resistant to proteolysis by trypsin.

In order to investigate the observed selectivity of SFTI-1 between matriptase and thrombin, we built the complex of SFTI-1 and thrombin. This was facilitated by our

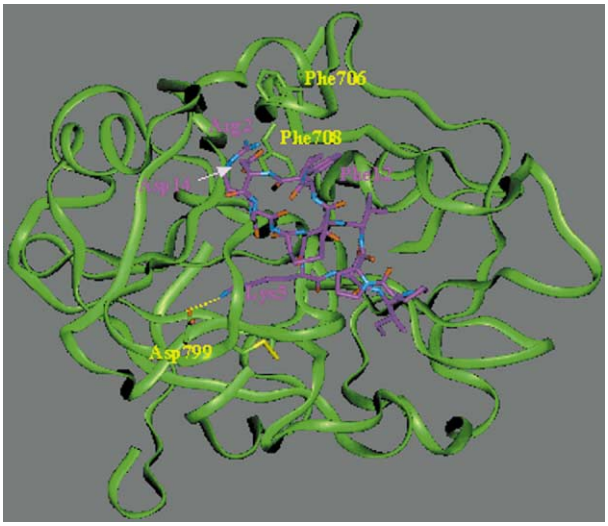



**Figure 2.** Modeling-based structure of matriptase complexed with SFTI-1.

earlier results of homology modeling of both the thrombin and the matriptase structure.[20] The inhibitor was placed into the active site of thrombin using the same relative position and orientation of the inhibitor as found in the X-ray structure of SFTI-1 complexed with trypsin. These structures show that, in order to accommodate the inhibitor in the active site of thrombin (PDB entry: 1hxe),[21] the loop between Tyr60.A and Asn60.G has to undergo significant conformational change, because Trp60.D is too bulky within the protein for effective binding of SFTI-1.

The constrained, conformationally rigid structure of the bicyclic peptide, SFTI-1, provides for a promising template towards further development of more specific inhibitors of extracellular matrix serine proteases, such as matriptase. Synthetic methodologies allow for the design and synthesis of various analogues, based on the homology-modeled matriptase catalytic site. Synthetic approaches are also available now for the design and generation of libraries of back bone cyclized peptides with bisecting cystine disulfide bridges, as demonstrated for the development of enzyme inhibitory BPTI mimetics.[22]

## Acknowledgements

S.-L. Lee was supported by Department of Defense Fellowship DAMD 17-00-1-0269.

## References and Notes

1. Shewry, P. R.; Lukas, J. A. *Advan. Botan. Res.* **1997**, *26*, 135.
2. Laskowski, M.; Kato, L. *Annu. Rev. Biochem.* **1980**, *49*, 593.
3. Luckett, S.; Santiago Garcia, R.; Barker, J. J.; Konarev, A. V.; Shewry, P. R.; Clarke, A. R.; Brady, R. L. *J. Mol. Biol.* **1999**, *290*, 525.
4. Lin, C.-Y.; Wang, J.-K.; Torri, J.; Dou, L.; Sang, Q. X. A.; Dickson, R. B. *J. Biol. Chem.* **1997**, *272*, 9147.
5. Lin, C.-Y.; Anders, J.; Johnson, M.; Sang, Q. A.; Dickson, R. B. *J. Biol. Chem.* **1999**, *274*, 18231.
6. Hooper, J. D.; Clements, J. A.; Quigley, J. P.; Antalis, T. M. *J. Biol. Chem.* **2001**, *276*, 857.
7. Kim, M. G.; Chen, C.; Lyu, M. S.; Cho, E. G.; Park, D.; Schwartz, R. H. *Immunogenetics* **1999**, *49*, 420.
8. Lin, C.-Y.; Anders, J.; Johnson, M.; Dickson, R. B. *J. Biol. Chem.* **1999**, *274*, 18237.
9. Takeuchi, T.; Harris, S.; Hwang, W.; Yan, K. W.; Coughlin, S. R.; Craik, C. S. *J. Biol. Chem.* **2000**, *275*, 26333.
10. Lee, S.-L.; Dickson, R. B.; Lin, C. Y. *J. Biol. Chem.* **2000**, *275*, 36720.
11. Kawaguchi, T.; Qin, L.; Shimomura, T; Kondo, J.; Matsumoto, K.; Denda, K.; Kitamura, N. *J. Biol. Chem.* **1997**, *272*, 27558.
12. (a) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Mayer, E. F.; Brice, J. M. D., Jr.; Rogers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. *J. Mol. Biol.* **1977**, *112*, 535. (b) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235.
13. Altschul, S. F.; Madden, T. L.; Schäffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. *Nucleic Acids Res.* **1997**, *25*, 3389.
14. Sali, A.; Potterton, L.; Yuan, F.; van Vlijmen, H.; Karplus, M. *Proteins: Struct. Funct. Genet.* **1995**, *23*, 318.
15. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
16. Andricioaei, I.; Straub, J. E. *Physical Rev. E.* **1996**, *53*, R3055.
17. Pak, Y.; Wang, S. *J. Phys. Chem. B* **2000**, *104*, 354.
18. Wenzel, H. R.; Tschesche, H. In *Peptides: Synthesis, Structure and Applications*; Gutte, B., Ed.; Academic: San Diego, 1995; pp 321–362.
19. Li, Y.; Huang, Q.; Zhang, S.; Liu, S.; Chi, C.; Tang, Y. *J. Biochem.* **1994**, *116*, 18.
20. Enyedy, I. J.; Lee, S.-L.; Kuo, A. H.; Dickson, R. B.; Lin, C.-Y.; Wang, S. *J. Med. Chem.* **2001**, *44*, 1349.
21. Rydel, T. J.; Tulinsky, A.; Bode, W.; Huber, R. *J. Mol. Biol.* **1991**, *221*, 583.
22. Kasher, R.; Oren, D. A.; Barda, Y.; Gilon, C. *J. Mol. Biol.* **1999**, *292*, 421.