J|A|C|S
ARTICLES
Published on Web 10/14/2006

A Minimalist Approach toward Protein Recognition by Epitope Transfer from Functionally Evolved β-Sheet Surfaces

Srivats Rajagopal, Scott C. Meyer, Aaron Goldman, Min Zhou, and Indraneel Ghosh*

Contribution from the Department of Chemistry, University of Arizona, P. O. Box 210041, 1306 East University Boulevard, Tucson, Arizona 85721-0041

Received July 10, 2006; E-mail: ghosh@email.arizona.edu

Abstract: New approaches for identifying small molecules that specifically target protein surfaces as opposed to active site clefts are of much current interest. Toward this goal, we describe a three-step methodology: in step one, we target a protein of interest by directed evolution of a small β-sheet scaffold; in step two, we identify residues on the scaffold that are implicated in binding; and in step three, we transfer the chemical information from the β-sheet to a small molecule mimic. As a case study, we targeted the proteolytic enzyme thrombin, involved in blood coagulation, utilizing a library of β-sheet epitopes displayed on phage that were previously selected for conservation of structure. We found that the thrombin-binding, β-sheet displaying mini-proteins retained their structure and stability while inhibiting thrombin at low micromolar inhibition constants. A conserved dityrosine recognition motif separated by 9.2 Å was found to be common among the mini-protein inhibitors and was further verified by alanine scanning. A molecule containing two tyrosine residues separated by a linker that matched the spacing on the β-sheet scaffold inhibited thrombin, whereas a similar dityrosine molecule separated by a shorter 6 Å linker could not. Moreover, kinetic analysis revealed that both the mini-protein as well as its minimalist mimic with only two functional residues exhibited noncompetitive inhibition of thrombin. Thus, this reductionist approach affords a simple methodology for transferring information from structured protein scaffolds to yield small molecule leads for targeting protein surfaces with novel mechanisms of action.

Introduction

Molecules that can selectively target protein surfaces have the potential for providing novel drug leads with new mechanisms of action and in providing new reagents in deconvoluting biological processes. Thus, the development of new methods for the identification of small proteins as well as small molecules for targeting shallow protein surfaces is of high interest.[1−6] For example, in an elegant structure-guided approach, termed "protein grafting" by Schepartz and co-workers,[7,8] the chemical information from the interaction surfaces of large natural proteins have been effectively reproduced on small helical scaffolds that have in turn been further improved by directed evolution. This approach has been utilized to target a wide array of proteins, such as Bcl-2,[8] Bcl-X_L,[9] EVH1,[10] and hDM2.[11] Kim and Sia have also used a similar grafting approach with the helical surface of GCN4 leucine zipper to target HIV-1.[12] In a different vein, Hamilton and co-workers have taken a minimalist approach[13] and successfully transferred epitope information from helical proteins to simple small molecule helix mimetics to target gp41[14] and Bcl-X_L.[15]

The above-described protein grafting and small molecule mimetic approaches require detailed knowledge of the targeted protein−protein interface and thus cannot be directly implemented for complexes that lack high-resolution structures. On the other hand, powerful selection approaches utilizing antibodies or peptides have made it possible to target most proteins without structural knowledge. Interestingly, it has been shown that de novo evolved epitopes most often target existing "hot spots"[16,17] on proteins that mediate protein−protein recognition. However, transferring the chemical information from an antibody or similarly complex protein epitopes[18] to small proteins or small molecule mimetics is not easily achieved. Thus, we

(1) Cochran, A. G. *Curr. Opin. Chem. Biol.* **2001**, *5*, 654.
(2) Boger, D. L.; Desharnais, J.; Capps, K. *Angew. Chem., Int. Ed.* **2003**, *42*, 4138.
(3) Arkin, M. R.; Wells, J. A. *Nat. Rev. Drug Discovery* **2004**, *3*, 301.
(4) Yin, H.; Hamilton, A. D. *Angew. Chem., Int. Ed.* **2005**, *44*, 4130.
(5) Zhao, L.; Chmielewski, J. *Curr. Opin. Struct. Biol.* **2005**, *15*, 31.
(6) Erlanson, D. A.; Wells, J. A.; Braisted, A. C. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 199.
(7) Zondlo, N. J.; Schepartz, A. *J. Am. Chem. Soc.* **1999**, *121*, 6938.
(8) Chin, J. W.; Schepartz, A. *Angew. Chem., Int. Ed.* **2001**, *40*, 3806.
(9) Gemperli, A. C.; Rutledge, S. E.; Maranda, A.; Schepartz, A. *J. Am. Chem. Soc.* **2005**, *127*, 1596.
(10) Golemi-Kotra, D.; Mahaffy, R.; Footer, M. J.; Holtzman, J. H.; Pollard, T. D.; Theriot, J. A.; Schepartz, A. *J. Am. Chem. Soc.* **2004**, *126*, 4.
(11) Kritzer, J. A.; Zutshi, R.; Cheah, M.; Ran, F. A.; Webman, R.; Wongjirad, T. M.; Schepartz, A. *ChemBioChem* **2006**, *7*, 29.
(12) Sia, S. K.; Kim, P. S. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 9756.
(13) Orner, B. P.; Ernst, J. T.; Hamilton, A. D. *J. Am. Chem. Soc.* **2001**, *123*, 5382.
(14) Ernst, J. T.; Kutzki, O.; Debnath, A. K.; Jiang, S.; Lu, H.; Hamilton, A. D. *Angew. Chem., Int. Ed.* **2001**, *41*, 278.
(15) Kutzki, O.; Park, H. S.; Ernst, J. T.; Orner, B. P.; Yin, H.; Hamilton, A. D. *J. Am. Chem. Soc.* **2002**, *124*, 11838.
(16) DeLano, W. L. *Curr. Opin. Struct. Biol.* **2002**, *12*, 14.
(17) DeLano, W. L.; Ultsch, M. H.; de Vos, A. M.; Wells, J. A. *Science* **2000**, *287*, 1279.
(18) Binz, H. K.; Amstutz, P.; Pluckthun, A. *Nat. Biotechnol.* **2005**, *23*, 1257.

10.1021/ja064885b CCC: $33.50

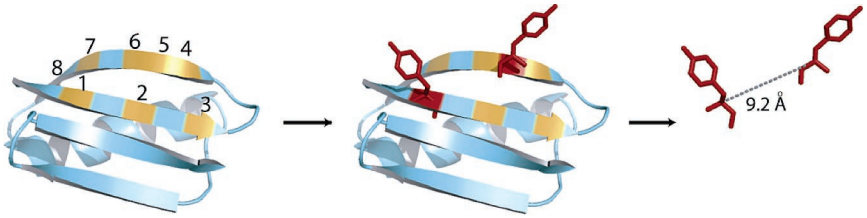

Figure 1. Strategy for small molecule protein recognition by epitope transfer from evolved mini-proteins.

posed the question, could we (a) use a scaffold that displayed a structurally well-defined α-helix or β-sheet for targeting a protein surface by directed evolution; (b) elucidate the minimal structural and chemical information necessary for binding; and (c) subsequently transfer the well-defined epitope information to a small molecule?

To implement our strategy, we first chose a small β-sheet scaffold (HTB1) designed for thermostability by Mayo and Malakauskas[19] (Figure 1). Minimal β-sheet surfaces or β-sheet mimetic designs have been challenging and thus are underutilized in targeting protein surfaces. However, recent advances in designing small β-sheets by Cochran et al.,[20] Gellman et al.,[21,22] and Waters and Tatko[23,24] as well as advances in β-sheet mimetics by Nowick and co-workers[25−27] provide new opportunities in this area. We envisioned that randomization of eight positions in the first two strands of HTB1 (Figure 1) with an amino acid set containing residues known to be over-represented at protein interfaces would provide a minimum of 200 Å^2 of a potential interaction surface. Since the helix lying across the β-sheet of HTB1 confers IgG binding capability,[28,29] we thought that if the mini-protein library were selected to bind IgG, then the resultant sub-library would retain nearly native tertiary structure. This would produce a new library with a well-defined β-sheet scaffold displaying the randomized residues. As a target protein, we chose thrombin, a serine protease that is representative of generic globular proteins.[30−32] Thrombin is involved in the blood-clotting cascade and possesses several protein−protein recognition sites, such as Exosite I, Exosite II, and the platelet-binding site.[30−32] Any of these existing sites could potentially be targeted by our pre-organized β-sheet. Herein we detail the results of selections with our mini-protein library, the validation of structural conservation, the identification of key residues implicated in thrombin binding, and finally the successful transfer of the chemical information from the structured β-sheet to a small molecule mimetic.

Materials and Methods

All amino acid derivatives, HOBT, BOP, and resins used for peptide synthesis were purchased from Novabiochem. Enzymes and M13KO7 helper phage were purchased from New England Biolabs, XL1-Blue *Escherichia coli* was obtained from Stratagene, custom oligonucleotides were obtained from Integrated DNA Technologies, and all other materials were purchased from Sigma, unless otherwise noted.

Library Construction and Phage Display Selection. The phage display vector pHTB1 was obtained by cloning a synthetic construct encoding for the HTB1 gene between the *SfiI* and *NotI* restriction sites into the phagemid vector pCANTAB 5E (Pharmacia).[32,33] There is a unique *PstI* site incorporated in the HTB1 construct sequence that enables the creation of two cassettes for mutagenesis (*SfiI*/*PstI* or *PstI*/*NotI*). The rTB1 library-containing cassette was created between the two restriction sites *SfiI*/*PstI* and incorporated the codon NRK at each position marked for randomization. N corresponds to either A, T, G, or C. R represents A or G, and K represents T or G. (See Supporting Information for the library primer sequences, Figure S1.) This library codes for residues biased toward those known to be over-represented in protein−protein interfaces.[34,35] After transformation of the library into XL-1 Blue *E. coli* cells, the total library size was 2.2×10^9.

The phage display selection was carried out as described previously.[32] Briefly, the HTB1 library was displayed as a gene III fusion on M13 phage particles using the phagemid vector library described above. The library-containing phage particles were isolated after superinfection with M13KO7 helper phage from the supernatant of an XL-1 Blue *E. coli* culture via poly(ethylene glycol) (PEG) precipitation. The isolated phage particles were then used to pan against targets immobilized on polystyrene plates via standard methods. Bound phage particles were eluted under acidic high-salt conditions (1 M NaCl, pH = 2). The output phage particles were then used to infect fresh XL-1 Blue *E. coli*, which became, after amplification, the input for the next round of selection. Phage-infected *E. coli* were plated on agar plates with glucose and ampicillin selection to achieve single colonies for isolation and DNA sequencing.

Protein Cloning, Expression, and Purification. Selected HTB1 mutants were cloned into appropriate restriction sites of the pQE30 His-tag expression vector (Qiagen) using appropriate primers. The sequences of the cloned mutants were confirmed by DNA sequencing and expressed as His-tag fusion proteins to facilitate purification. (See Supporting Information for primer sequences and amino acid sequences of the complete, expressed proteins, Figures S2 and S3.) All HTB1 mutants were expressed in XL-1 Blue *E. coli* and induced with 1 mM IPTG at an $OD_{600} = 1$. Cells were harvested after 10−16 h and lysed. The soluble fraction of the cell lysate was bound to nickel agarose resin (Qiagen) in a buffer containing 50 mM Na_2HPO_4 and 0.3 M NaCl at pH 8.0. The proteins were eluted with increasing concentrations of imidazole, with all mutants eluting at 0.5 M imidazole. These proteins were subsequently purified on a size-exclusion HiLoad 16/60 Superdex 75 prep grade FPLC column (Amersham Pharmacia). Protein molecular weights were determined via MALDI mass spectrometry (see Supporting Information, Table S1).

HTB1(Q33C) Labeling and IgG Binding Assay. The extent to which the mutants could bind HTB1's native target, IgG, was measured via an assay first developed by Hellinga and Sloan.[36,37] The Q33C mutant of HTB1 was labeled with the environmentally sensitive dye,

(19) Malakauskas, S. M.; Mayo, S. L. *Nat. Struct. Biol.* **1998**, *5*, 47.
(20) Cochran, A. G.; Skelton, N. J.; Starovasnik, M. A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 5578.
(21) Syud, F. A.; Stanger, H. E.; Gellman, S. H. *J. Am. Chem. Soc.* **2001**, *123*, 8667.
(22) Espinosa, J. F.; Syud, F. A.; Gellman, S. H. *Protein Sci.* **2002**, *11*, 1492.
(23) Tatko, C. D.; Waters, M. L. *J. Am. Chem. Soc.* **2002**, *124*, 9372.
(24) Tatko, C. D.; Waters, M. L. *Protein Sci.* **2003**, *12*, 2443.
(25) Nowick, J. S.; Cary, J. M.; Tsai, J. H. *J. Am. Chem. Soc.* **2001**, *123*, 5176.
(26) Nowick, J. S. *Acc. Chem. Res.* **1999**, *32*, 287.
(27) Nowick, J. S.; Brower, J. O. *J. Am. Chem. Soc.* **2003**, *125*, 876.
(28) Sauer-Eriksson, A. E.; Kleywegt, G. J.; Uhlen, M.; Jones, T. A. *Structure* **1995**, *3*, 265.
(29) Gronenborn, A. M.; Clore, G. M. *J. Mol. Biol.* **1993**, *233*, 331.
(30) Stubbs, M. T.; Bode, W. *Curr. Opin. Struct. Biol.* **1994**, *4*, 823.
(31) Lane, D. A.; Philippou, H.; Huntington, J. A. *Blood* **2005**, *106*, 2605.
(32) Rajagopal, S.; Meza-Romero, R.; Ghosh, I. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 1389.
(33) Meyer, S. C.; Huerta, C.; Ghosh, I. *Biochemistry* **2005**, *44*, 2360.
(34) Janin, J.; Chothia, C. *J. Biol. Chem.* **1990**, *265*, 16027.
(35) Bogan, A. A.; Thorn, K. S. *J. Mol. Biol.* **1998**, *280*, 1.

acrylodan: 4.7 mg (21 μmol) of acrylodan (Molecular Probes) was dissolved in 1 mL of dimethyl sulfoxide, which was then added to 1.05 μmol of the Q33C mutant in 15 mL of PBS buffer (70 μM final concentration). The reaction vessel was shaken for 2 h, at which point the reaction was quenched with 20 μL of β-mercaptoethanol. The labeled protein, FHTB1, was first purified on a nickel-agarose affinity column (Qiagen), and then on a NAP-10 Sephadex G-25 column (Amersham Pharmacia). Labeling of the protein was confirmed by visualization of an SDS-PAGE gel by UV fluorescence and coomassie blue staining (data not shown). The molecular weight of FHTB1 was confirmed via MALDI mass spectrometry (expected mass: 8571 g/mol; found: 8569 *m*/*z*).

IgG binding assays were run on a Photon Technology International (PTI) spectrofluorometer in 20 mM Tris-HCl at pH = 7.4 with an excitation wavelength of 392 nm. An internal standard of a water-soluble 20 nM quantum dot solution (λ_{max} = 595 nM, Evident Technologies) was used. Fluorescence emission of the labeled FHTB1 was monitored at 480 nm and expressed as a fraction of the quantum dot emission intensity.

To determine the binding affinity of FHTB1 to the Fc portion of IgG, whole human IgG was added in increasing concentrations to FHTB1 (with a final FHTB1 concentration of 300 nM). The results with whole human IgG correspond well with those obtained using the Fc portion of human IgG, assuming two Fc regions per IgG molecule (data not shown). To determine the IgG binding affinity of the thrombin-binding HTB1 mutants, a competition assay was run with FHTB1. Increasing amounts of the selected thrombin binders were premixed with FHTB1 (300 nM), and human IgG was added (500 nM putative Fc concentration). Fluorescence was measured after a 30-min incubation and expressed as a fraction of the quantum dot internal standard (20 nM). The resulting binding curves were fit to the following equation:

$$F = \frac{F_0}{1 + \left(\frac{K_{\mathrm{DFHTB1}}}{[\mathrm{FHTB1}]}\right)\left(1 + \frac{[B1_{\mathrm{mut}}]}{K_{\mathrm{Dmut}}}\right)} + F_{\mathrm{f}} \qquad (1)$$

where F is the measured fluorescence, F_0 is the initial fluorescence reading, F_f is the final fluorescence reading, [FHTB1] is the concentration of FHTB1, K_{DFHTB1} is the dissociation constant of FHTB1 for IgG, [$B1_{mut}$] is the concentration of the selected thrombin-binding mutant, and K_{Dmut} is the dissociation constant of the thrombin binders for their interaction with IgG.[36]

Circular Dichroism Spectroscopy. Circular dichroism (CD) data were recorded on an Aviv 62A-DS spectropolarimeter using a 0.1-cm path length cell. All spectra were taken at 25 °C in 20 mM sodium phosphate at pH 8.0. Protein concentrations were determined by UV spectrophotometry at 280 nm under denaturing conditions. The mean residue ellipticity was calculated as: $(\theta \times 100)/(0.1\ \mathrm{cm} \times [\mathrm{P}] \times n)$, where n is the number of residues, [P] is the protein concentration (μM), θ is in units of (degree × cm^2/dmol), and the spectra were acquired in a 0.1-cm path length cell. Thermal melts were monitored at 218 nm and were run in 3 M guanidine HCl. Data points were collected every 2 °C with an averaging time of 10 s and an equilibration time of 2 min.

HTB1 Mutant Labeling and Analytical Ultracentrifugation. Single Cys residues were incorporated into the N-termini of the proteins 9sr15C and 9sr4C (Supporting Information, Figure S3), which were then labeled with the thiol-specific probe, Alexa Fluor 488 C_5 maleimide (Molecular Probes). The Cys-containing proteins were separately dialyzed into 20 mM Tris buffer containing 20 mM NaCl and 1 mM DTT at pH 7.5 and 4 °C. The proteins (2 mL) were next dialyzed into 800 mL of buffer without DTT for a period of 2 h at 4 °C. The buffer was then replaced with fresh buffer, and the proteins were allowed to dialyze for 2 h more. To 700 μL of 50 μM protein 9sr15C or 9sr4C was added a 10-fold mol excess of Alexa Fluor 488 C_5 maleimide dye, and the reaction was allowed to proceed for 3 h at room temperature on the benchtop shaker. Excess dye was removed with a NAP-10 Sephadex G-25 column (Amersham Pharmacia). A 15% SDS-PAGE gel was run to confirm the labeling of the protein (not shown). The dye-labeled proteins were then further purified via FPLC with a size-exclusion HiLoad 16/60 Superdex 75 prep grade column (Amersham Pharmacia).

Sedimentation equilibrium analysis was performed on a Beckman Optima XL-1 analytical ultracentrifuge equipped with an An60Ti rotor at 25 °C with rotor speeds of 10 000, 15 000, and 20 000 rpm. A control experiment was run simultaneously in another rotor cell containing only dye-labeled protein. All protein samples were prepared and dialyzed overnight in 20 mM Tris buffer containing 20 mM NaCl at pH 8.0. The dialysates were then used as blanks during the sedimentation analysis. Scans were performed by measuring the absorbance at 494 or 280 nm with a step size of 0.001 cm. Samples equilibrated for 24 h and duplicate scans 2 h apart were overlaid to determine when equilibrium had been reached. Data points were collected for (a) a mixture of thrombin (Sigma) at a final concentration of 4 μM with dye-labeled 9sr15C (11 μM final) and (b) thrombin (2.2 μM final) with labeled 9sr4C (9 μM final). By fitting the data to a theoretical curve, a fraction for the bound complex was obtained that enabled an estimation of the equilibrium dissociation constant (K_d) for the interaction between thrombin and the HTB1 variant. In all cases, the data sets were analyzed by SEDEQ 4.1 (Allen Minton, Laboratory of Biochemical Pharmacology, NIH).

Thrombin Inhibition Assay. The inhibitory activities of the HTB1 mini-proteins for thrombin were measured in a kinetic assay using a fluorogenic substrate, benzoyl-Phe-Val-Arg-AMC (Calbiochem). These assays were performed at room temperature in 20 mM Tris buffer at pH 8.0, containing 20 mM NaCl and 0.05% Triton X-100 with 1 nM thrombin. After 30 min of incubation of thrombin with increasing concentrations of the HTB1 mutant proteins, the assays were started by the addition of the fluorogenic substrate. The proteolytic release of coumarin was monitored for 60 s by fluorescence emission at 450 nm using a PTI spectrofluorimeter, with an excitation wavelength of 370 nm. The raw data points were fit linearly, and the resulting slopes were used to calculate IC_{50} values, which were determined by plotting the data as percent inhibition versus the log of the concentration of the HTB1 mutant proteins. The data sets were fit to the variable slope sigmoidal eq 2 using Kaleidagraph (version 3.52, Synergy Software).

$$y = a + \frac{(b - a)}{1 + \left(\frac{x}{c}\right)^d} \qquad (2)$$

where a is the value of dependent variable y for the minimal curve asymptote, b denotes the value of y for the maximal curve asymptote, x is the logarithm of the inhibitor concentration, c is the IC_{50} value, and d is the Hill slope, which describes the steepness of the inhibitor-response curve. The data set was further analyzed by the Cheng−Prusoff [38] eq 3:

$$K_{\mathrm{i}} = \frac{IC_{50}}{1 + \left(\frac{[\mathrm{S}]}{K_{\mathrm{m}}}\right)} \qquad (3)$$

where IC_{50} is the concentration of the inhibitor producing a 50% inhibition, [S] is the substrate concentration, and K_m (21 ± 3 μM) is the Michaelis constant of the substrate for the enzyme obtained from a Michaelis Menton plot (data not shown). A thrombin inhibition assay was also performed as described above using small molecule dityrosine

(36) Sloan, D. J.; Hellinga, H. W. *Protein Sci.* **1999**, *8*, 1643.
(37) Sloan, D. J.; Hellinga, H. W. *Protein Eng.* **1998**, *11*, 819.
(38) Cheng, Y.; Prusoff, W. H. *Biochem. Pharmacol.* **1973**, *22*, 3099.

mimics: (1) Tyr-βAla-βAla-Tyr and (2) Tyr-Gly-Tyr (see below for details of the peptide synthesis).

Kinetic assays were performed on the protein 9sr4 using different substrate concentrations that were varied from 25 to 150 μM. Inhibitor concentrations were likewise varied, but with a concentration range of 5−50 μM. The same experiment was repeated with the tetrapeptide (1) (Tyr-βAla-βAla-Tyr), which was varied from 500 to 1000 μM. A plot of $1/v$ versus 1/[S] gave the best fit to a noncompetitive model (SigmaPlot) expressed in eq 4 in terms of v:

$$v = \frac{V_{max}}{\left(1 + \frac{I}{K_i}\right)\left(1 + \frac{K_m}{[S]}\right)} \quad (4)$$

where v is the velocity of the reaction in arbitrary fluorescence units per second and V_{max} is the maximum velocity of the reaction.

Alanine-Scanning Mutagenesis. Single alanine mutants for both 9sr4 and 9sr15 at positions 1 and 5 were constructed by site-directed mutagenesis using the QuikChange site-directed mutagenesis protocol (Stratagene). Utilizing the appropriate primers (see Supporting Information for sequences, Figure S2), we constructed the Y1A and Y5A mutants of 9sr4 and 9sr15 separately via polymerase chain reaction with the parent pQE30 constructs as templates. The four Ala mutant plasmids were isolated, and the proteins were expressed in the same manner as the parent mutants, as described above.

Peptide Synthesis and Purification. Peptides were synthesized by solid-phase peptide synthesis using standard fluorenylmethoxycarbonyl (Fmoc) strategy on Rink Amide AM resin. The final peptides were cleaved from the resin, and global deprotection was achieved by TFA cleavage with the appropriate scavengers (95% TFA, 2.5% water, 2.5% triisopropylsilane) for 120 min at room temperature. The solution was filtered and precipitated by transferring to a centrifuge tube that contained pre-cooled ether. After centrifugation and three washes with fresh ether, the white crude product was collected. The crude peptide was purified by reverse-phase HPLC (Varian) using a C18 column (Vydac). Peptides were purified with a linear gradient of 20−80% acetonitrile in water containing 0.1% TFA at a flow rate of 8 mL/min, for 50 min. Pure peptide (9.5 and 27.8 mg) was obtained for peptides (1) and (2), respectively, after purification.

Peptide sequences: (**1**) NH_2-Tyr-βAla-βAla-Tyr-$CONH_2$ and (**2**) NH_2-Tyr-Gly-Tyr-$CONH_2$.

MALDI mass spectra confirmed the peptide molecular weights:

Peptide (**1**) expected: 485.5; found: 486.1 (*m*/*z*).

Peptide (**2**) expected: 400.4; found: 401.1 (*m*/*z*).

Results and Discussion

Our overall goal was to systematically apply a three-step approach to generate minimalist mimics of protein epitopes capable of binding to a selected target protein. In step 1, we describe the directed evolution of a small β-sheet epitope (HTB1) to target the serine protease, thrombin. In this step, we utilize a series of biophysical and biochemical methods to establish the retention of structure in our scaffold and also establish thrombin binding and inhibition. In step 2, we detail the identification of key residues implicated in thrombin binding. Finally, in step 3 we transfer the chemical information from the structured β-sheet to minimalist small molecules and interrogate them for inhibition of thrombin's proteolytic activity.

Step 1: (a) Identification of Structured Mini-Proteins that Bind Thrombin. As our scaffold, we chose HTB1, a thermostable mini-protein redesigned by Mayo and Malakauskas[19] that binds IgG through residues on a short α-helix buttressed by a four-stranded β-sheet (Figure 1). We synthesized a 10^8 member

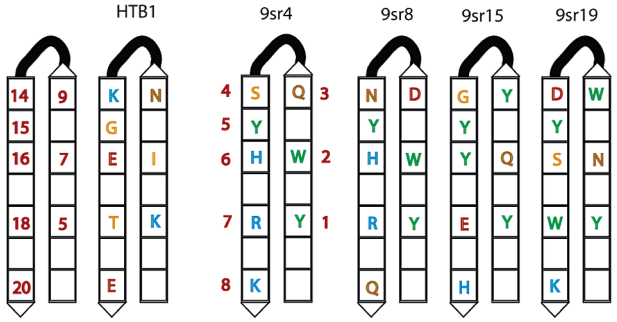

Figure 2. Thrombin-selected HTB1 mini-protein sequences represented as an antiparallel β-sheet motif. The residue numbering, as well as the parent HTB1 sequence, is depicted on the left, with the four thrombin-binding mini-proteins with Tyr in randomized positions 1 and 5 on the right.

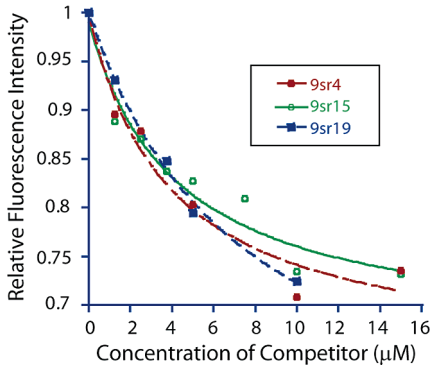

Figure 3. IgG binding assay. Addition of IgG to an acrylodan-labeled HTB1 (FHTB1) causes an increase in fluorescence emission at 480 nm. Unlabeled thrombin-binding HTB1 mutant proteins were added to the IgG/FHTB1 complex, and the decrease in fluorescence emission was measured as a function of competitor concentration.

β-sheet library displayed on phage within the context of the first two β-strands of HTB1. To obtain structurally well-defined epitopes, we carried out in vitro selections against IgG (structural selection) followed by thrombin (functional selection). Nine rounds of selection against thrombin resulted in the identification of a set of mini-proteins where the first and fifth positions were conserved (Figure 2), which likely represented a thrombin-binding epitope. The selected mini-proteins 9sr4, 9sr8, 9sr15, 9sr19, and the parent HTB1 were cloned into an expression vector, expressed in *E. coli*, purified by metal affinity chromatography followed by gel-filtration chromatography, and subsequently characterized by mass spectrometry. Typical yields of isolated proteins were ∼40 mg/L.

(b) IgG Binding of Thrombin-Binding Mini-Proteins. We wanted to verify that our so-called "structural selection" step against IgG allowed our mini-proteins to retain their ability to bind IgG. We modified an IgG binding assay developed by Hellinga and Sloan in the context of the native GB1 protein.[37] An environmentally sensitive fluorophore, acrylodan, was site specifically attached to a cysteine mutation at position 33 of the HTB1 protein to form FHTB1. Direct binding of IgG to FHTB1 resulted in an increase in fluorescence at 480 nM (data not shown). The binding affinity of individual mini-proteins was determined by competition against the IgG/FHTB1 complex. On the basis of the K_d of the FHTB1−IgG interaction (320 ± 5 nM), we found that the binding constants for our mini-proteins were 2.1 μM for 9sr4, 2.2 μM for 9sr15, and 3.9 μM for 9sr19. Thus, the ability of the thrombin-binding mini-proteins to bind IgG strongly suggests retention of tertiary structure (Figure 3).

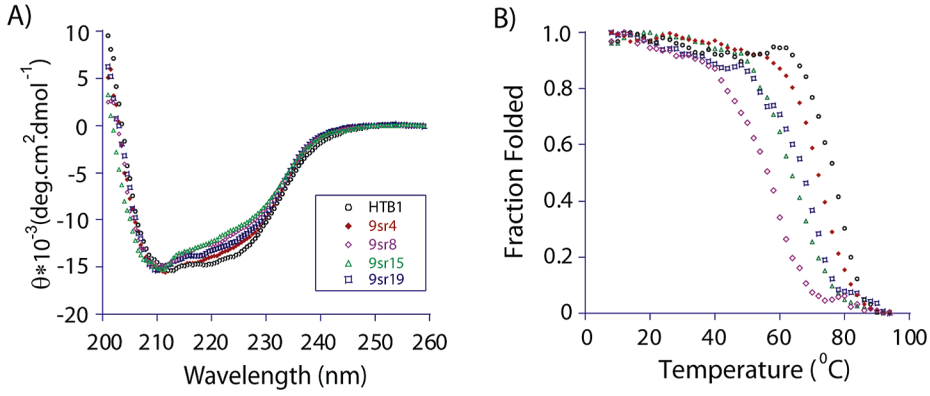

Figure 4. CD spectroscopy of the thrombin-binding miniproteins. (A) CD spectra of the mini-proteins were recorded at 25 μM protein in a 20 mM sodium phosphate buffer at pH 8.0. (B) Thermal melts of the selected thrombin-binding mini-proteins were monitored at 218 nm in 3 M guanidine HCl with 20 mM sodium phosphate buffer at pH 8.0.

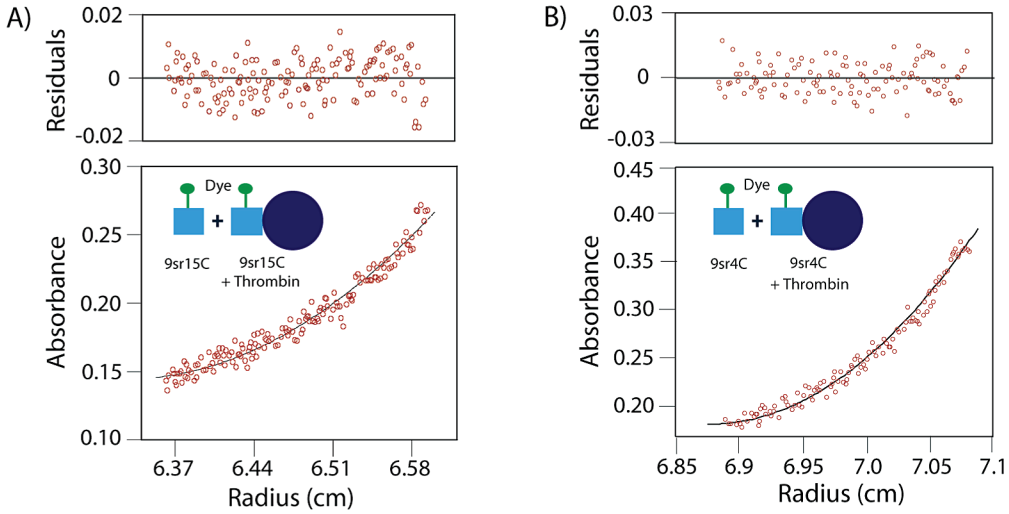

Figure 5. Sedimentation equilibrium analysis of 9sr15C−Alexa Fluor 488 and 9sr4C−Alexa Fluor 488 with thrombin. The sedimentation equilibrium absorbance profile of the labeled thrombin-binding mini-proteins (A) 9sr15C and (B) 9sr4C in the presence of thrombin recorded at 494 nm. Data shown for 15 000 rpm.

(c) Circular Dichroism Spectroscopy of Thrombin-Binding Mini-Proteins. Having verified IgG binding, we wanted to clearly establish that the mini-proteins were well-folded despite incorporating eight mutant surface residues on adjacent β-strands. In our second set of experiments to test for structural conservation, we used CD spectroscopy. CD spectroscopy demonstrated that the selected mini-proteins were well-structured in solution and possessed spectra that were very similar to the parent HTB1 parent scaffold (Figure 4A). To investigate the relative stability of these proteins as compared to the thermostable parent HTB1, we carried out CD thermal denaturation experiments under extreme conditions (3 M guanidine hydrochloride), where most proteins of 56 residues would be unfolded. All the mini-proteins showed significant stability (Figure 4B) displaying reversible folding and unfolding transitions, with melting temperatures between 55 and 75 °C in 3 M guanidine hydrochloride. Thus, we had realized our first goal of obtaining stable, structured domains.

(d) Direct Thrombin Binding by Sedimentation Equilibrium Experiments. The next step in the characterization of the HTB1 variants was to establish that they could indeed bind thrombin. For measuring direct binding, we chose to use sedimentation equilibrium experiments by analytical ultracentrifugation where complexation can be unambiguously established. However, it is challenging to directly determine the complexation of two macromolecules of significantly different molecular weights using intrinsic absorbance. Typically, one of the macromolecules has to be selectively labeled to follow complexation.[39] Thus, a cysteine residue was introduced in the mini-proteins 9sr15 and 9sr4 (to form 9sr15C and 9sr4C, respectively) by site-directed mutagenesis at the N-terminal end of the mini-proteins. This Cys could then be selectively labeled with the Alexa Fluor 488 C_5 maleimide dye. The labeled protein and thrombin were mixed at molar ratios of 2.8:1 for 9sr15 C and 4:1 for 9sr4C and allowed to reach equilibrium at different rotor speeds (10 000, 15 000, and 20 000 rpm) (Supporting Information, Figure S5 and Table S2). Analysis of the sedimentation equilibrium data (Figure 5) clearly showed that both 9SR15 and 9SR4 formed 1:1 complexes with thrombin with estimated dissociation constants (K_d) of 9.6 and 13.6 μM, respectively (see Supporting Information for details). This experiment validated our two-step selection process and established that the mini-proteins could evolve a new thrombin-binding function while retaining their β-sheet structure and IgG binding ability. This is significantly different from selections with antibodies, antibody mimetics, and unstructured peptides,

(39) Weinberg, R. L.; Veprintsev, D. B.; Fersht, A. R. *J. Mol. Biol.* **2004**, *341*, 1145.

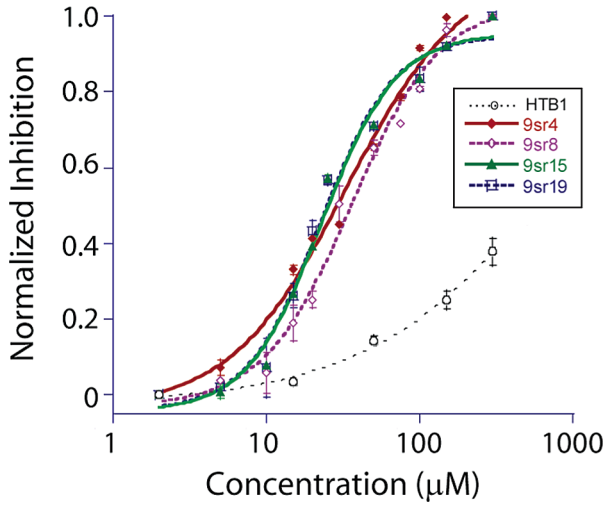

Figure 6. Thrombin inhibition by selected mini-proteins. Kinetic inhibition assays of thrombin were run using the fluorogenic substrate benzoyl-Phe-Val-Arg-AMC. The normalized inhibition of the HTB1 mini-proteins was plotted against the inhibitor concentration to determine IC_{50} values, which were then used to calculate the inhibition constant, K_i, based on the Cheng−Prusoff equation.

where structure is not predetermined. In these cases, the structures of the selected ligands remain unknown, unlike that for our HTB1 scaffold, without high-resolution structural determination of the complexes.

(e) Thrombin Inhibition Assay. Once the thrombin-binding ability of the selected HTB1 mutants was established, we investigated whether binding could perhaps lead to the inhibition of thrombin's proteolytic activity. The mini-proteins were tested for thrombin inhibition utilizing a commercially available substrate benzoyl-Phe-Val-Arg-AMC ($K_m = 21 \pm 3\ \mu M$), which releases a fluorescent coumarin derivative when cleaved by thrombin. Inhibitory constants, K_i, were determined from the Cheng−Prusoff equation.[38] All the mini-proteins were found to inhibit thrombin in the low micromolar regime (Figure 6), even though we had not explicitly selected for thrombin inhibitors but only thrombin binders. The specific inhibition constants, K_i, for the mini-proteins were found to be: 9sr4, 27 $\pm$ 5 μM; 9sr8, 28 $\pm$ 4 μM; 9sr15, 18 $\pm$ 3 μM; and 9sr19, 17 $\pm$ 3 μM. Thus, low micromolar inhibitors could be evolved in the context of our pre-organized β-sheet.

We next investigated whether the selected β-sheet epitopes were specific for thrombin or if they inhibited other proteases as well. The proteolytic assay was carried out with trypsin, a promiscuous protease,[40] which also cleaves the thrombin substrate, benzoyl-Phe-Val-Arg-AMC. 9SR4 was assayed against trypsin and showed very weak inhibitory activity, which strongly suggests that the mini-proteins are selective for thrombin and are not generic protease inhibitors (Supporting Information, Table S3).

Step 2: Epitope Validation by Alanine-Scanning Mutagenesis. Having established that we had identified structured β-sheet epitopes capable of binding and inhibiting thrombin, we wished to identify the residues on our scaffold that were implicated in thrombin binding. Since each of the evolved β-sheet epitopes possessed tyrosine residues at the first and fifth positions, we chose to sequentially replace them by alanine.

9sr Proteins	% Thrombin Inhibition	
	15 μM	50 μM
9sr4	28.6	65.3
9sr4(Y1A)	2.5	20.2
9sr4(Y5A)	0	6.9
9sr15	43.5	89.6
9sr15(Y1A)	8.6	13.1
9sr15(Y5A)	8.3	22.7

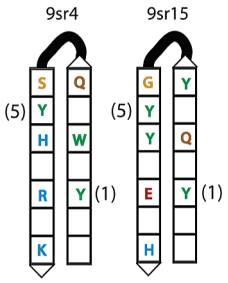

Figure 7. Thrombin inhibition by single alanine mutants of the 9sr4 and 9sr15 mini-proteins. The table on the left shows the % inhibition of thrombin by the alanine mutants in comparison to that of the parent thrombin-binding mini-proteins at the same concentrations.

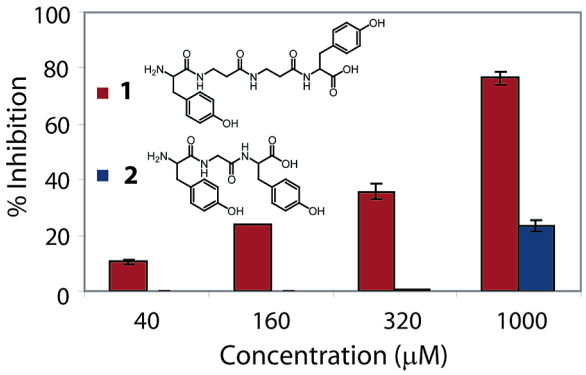

Figure 8. Inhibition of thrombin with minimal mimics of the dityrosine epitope. The % inhibition of thrombin with tetrapeptide **1** (Tyr-β-Ala-β-Ala-Tyr) and tripeptide **2** (Tyr-Gly-Tyr). The calculated distance between the two tyrosines in **1** mimics that of the selected thrombin-binding peptides (11 Å) based on the NMR structure of the parent HTB1, whereas the same distance in **2** is shorter (6 Å).

Thus, individual tyrosine to alanine mutants of 9sr4 and 9sr15 were cloned, expressed, and purified. We first established through CD spectroscopy that the secondary structure of the mutants was not grossly perturbed (Supporting Information, Figures S6A,B). Following secondary structural validation, each of the mini-protein mutants was tested for their ability to inhibit thrombin. The inhibitory activities of the single alanine mutants were significantly lower, and thus K_i values could not be accurately determined, as 100% inhibition was not achieved. Instead, we evaluated the relative activity of these alanine mutants at two specific concentrations (15 and 50 μM) (Figure 7). A significant loss in thrombin inhibitory activity was observed for each of the alanine mutants of 9sr4 and 9sr15 compared to the parent mini-proteins. The conservation of the dityrosine together with the loss of activity upon mutation of either tyrosine to alanine strongly suggests that the two tyrosine residues may comprise a minimal thrombin-binding epitope. It is very interesting to note that, when Sidhu and co-workers utilized a simple four amino acid code for recognition of the vascular endothelial growth factor (VEGF) with an antibody, tyrosines provided 71% of the surface area required for specific VEGF recognition.[41]

(40) Parry, M. A.; Stone, S. R.; Hofsteenge, J.; Jackman, M. P. *Biochem. J.* **1993**, *290*, 665.

(41) Fellouse, F.; Weismann, C.; Sidhu, S. S. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 12467.

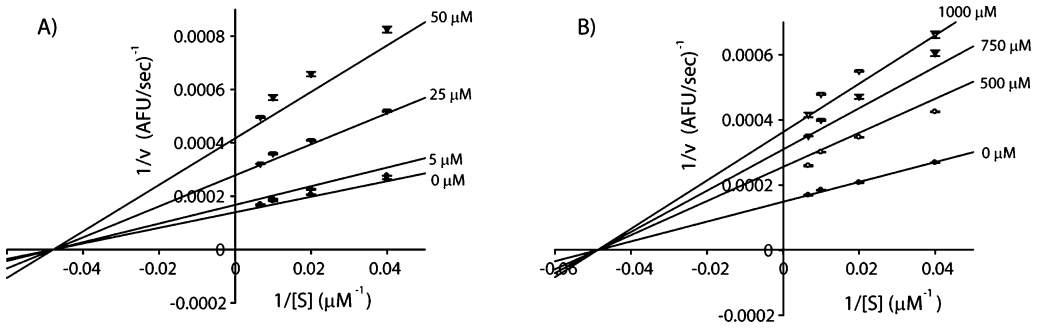

Figure 9. Kinetic analysis of the inhibition of thrombin by 9sr4 and the tetrapeptide **1** (Tyr-β-Ala-β-Ala-Tyr). Thrombin was incubated with varying concentrations of (A) the thrombin-binding mini-protein 9sr4 (5, 25, and 50 μM) and (B) the tetrapeptide **1** (500, 750, and 1000 μM) with the substrate concentration increasing from 25 to 150 μM. The data sets were globally fit to a noncompetitive binding model. Fits to competitive and uncompetitive binding models are available in the Supporting Information.

Step 3: Small Molecule Mimic of the Dityrosine Motif. With evidence that the dityrosine motif is important for binding thrombin, we finally turned to mimicking this minimal motif in a small molecule. Our previous experiments utilizing IgG binding and CD spectroscopy strongly suggested that the thrombin-binding mutants retain the parent HTB1 structure. This allowed us to assign a distance of 9.2 Å between the Cα carbons of tyrosine residues at positions 1 and 5 (Figure 1) on the basis of the NMR structure of HTB1.[19] This structural model of the mini-proteins indicates that an absolutely minimal epitope comprising a tetrapeptide Tyr-(βAla)$_2$-Tyr would mimic the distance between the two Tyr residues in question and would span an average $\sim$11 Å. A second peptide, Tyr-Gly-Tyr, was designed to restrict the two tyrosine residues to a shorter distance of $\sim$6 Å. We hypothesized that Tyr-(βAla)$_2$-Tyr would mimic the epitope and inhibit thrombin whereas Tyr-Gly-Tyr would not span the requisite distance to mimic the dityrosine epitope and therefore be a poor inhibitor. Both peptides were synthesized, purified, and subsequently tested in the thrombin inhibition assays. The minimal tetrapeptide epitope was shown to have $IC_{50} = 681 \pm 70$ μM, whereas the shorter Tyr-Gly-Tyr exhibited negligible thrombin inhibition (Figure 8) even though they both contained the same dityrosine residues. This experiment demonstrated that it is possible to transfer the epitope information from a mini-protein to a very simple small molecule, where appropriate separation between the tyrosine residues is critical for inhibitory activity.

Mode of Inhibition. After the successful design of the minimal small molecule that could recapitulate the activity of the 56-residue mini-protein, we wanted to establish that the mode of inhibition for both Tyr-(βAla)$_2$-Tyr and the parent mini-proteins was identical. Toward this goal, kinetic analysis of the mini-protein, 9sr4, and the tetrapeptide, Tyr-(βAla)$_2$-Tyr, was carried out. The resulting kinetic data were analyzed using equations for competitive, noncompetitive, and uncompetitive inhibition models (Supporting Information, Figures S7 and S8). Neither the competitive nor the uncompetitive model agreed with the kinetic data. Interestingly, the data for both mini-protein and its mimetic, Tyr-(βAla)$_2$-Tyr, were best fit to a noncompetitive inhibition model (Figure 9) and provided K_i values of 25 $\pm$ 4 μM for 9sr4 and 705 $\pm$ 26 μM, respectively. The K_i values obtained from this analysis are in good agreement with the values previously determined at a single substrate concentration using the Cheng−Prusoff method (Figure 6). Dynamic light scattering as well as 8-anilino-1-naphthalene sulfonic acid binding experiments ruled out possible aggregation-induced inhibition of our small molecule inhibitor (Supporting Information, Figures S9 and S10). Furthermore, extended dialysis experiments demonstrated that the Tyr-(βAla)$_2$-Tyr inhibitor was a reversible noncovalent inhibitor (Supporting Information, Figure S11). Since both our mini-protein as well as the small molecule inhibitor exhibit similar modes of inhibition, it clearly suggests that they very likely target the same site on thrombin, as designed. Furthermore, the observed noncompetitive inhibition also suggests that both these molecules bind to a protein surface distal to the active site of thrombin and exert an allosteric effect that prevents substrate turnover. Preliminary protease footprinting experiments to localize the site on thrombin targeted by 9SR4 indicate that the mini-proteins likely target the heparin-binding site (Supporting Information, Figure S12). Noncompetitive inhibition is an emerging area in enzyme inhibition and has proven to be successful in inhibiting kinases such as MEK1 and MEK2, where highly selective small molecules target a site adjacent to the ATP pocket.[42,43] Furthermore, recent results from the Alzheimer's field indicate that several potent small molecules specifically inhibit γ-secretase by a noncompetitive mechanism.[44]

Conclusion

We have demonstrated that the small thermostable HTB1 scaffold can be utilized as a platform for generating a well-folded β-sheet library, which in turn can be selected for binding a globular target, thrombin. Interestingly, the selected thrombin-binding mini-proteins also inhibited thrombin's proteolytic activity with K_i values in the 17−27 μM range. Homology between the mini-proteins as well as alanine-scanning experiments identified a dityrosine motif across the two β-strands that was necessary for thrombin inhibition. Simple molecular modeling allowed transfer of the dityrosine epitope to a flexible small molecule that maintained the 9.2 Å spacing between the tyrosine residues. This small molecule was found to inhibit thrombin with a K_i of 705 μM in a noncompetitive manner as observed for the parent mini-protein ($K_i = 25$ μM). The 28-fold difference in activity by reduction of a 56-residue mini-protein to only two functional residues on a flexible scaffold is comparable to the 10-fold difference in activity observed in the reduction of a 20-residue peptide inhibitor to three functional residues on a

(42) Ohren, J. F.; et al. *Nat. Struct. Mol. Biol.* **2004**, *11*, 1192.
(43) Sebolt-Leopold, J. S.; English, J. M. *Nature (London)* **2006**, *441*, 457.
(44) Tian, G.; Sobotka-Briner, C. D.; Zysk, J.; Liu, X.; Birr, C.; Sylvester, M. A.; Edwards, P. D.; Scott, C. D.; Greenberg, B. D. *J. Biol. Chem.* **2002**, *277*, 31499.

helix mimetic reported by Hamilton and co-workers.[13] Thus, we demonstrate the feasibility of a simple three-step approach for obtaining a minimalist small molecule with only two functional residues that targets a protein surface. Future work will focus on chemical elaboration of the small dityrosine motif to increase specificity and by incorporation of structural constraints to increase affinity. At the same time, we will also elucidate the cognate binding surface on thrombin that allows for the novel mechanism of action of this new class of thrombin inhibitors. We believe that the three-step approach described herein for the de novo discovery of small molecule mimics for protein surfaces is general and can be extended to other mini-protein motifs, such as α-helix scaffolds,[7] that could be further reduced to small molecule helix mimetics.[13] Furthermore, recent advances in phage display with unnatural amino acids may eventually allow for a larger diversity in scaffold surface chemistry.[45]

Acknowledgment. We thank members of the Ghosh laboratory, Professor Scott Saavedra, and Professor Wolfram Bode for helpful comments. This work was partially supported by grants from the ABRC, R21AG025954 (NIH), and R01AI068414 (NIH). The model in Figure 1 was rendered using PyMol (www.pymol.org).

Supporting Information Available: Details of cloning, sequencing, purification, biophysical characterization (S1−S14), and complete ref 42 (S14) citation. This material is available free of charge via the Internet at http://pubs.acs.org.

JA064885B

(45) Feng, T.; Tsao, M. L.; Schultz, P. G. *J. Am. Chem. Soc.* **2004**, *126*, 15962.