Unexpected fold in the circumsporozoite protein target of malaria vaccines

Michael B. Doud, Adem C. Koksal, Li-Zhi Mi, Gaojie Song, Chafen Lu, and Timothy A. Springer

Immune Disease Institute, Children’s Hospital Boston and Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115

Contributed by Timothy A. Springer, April 4, 2012 (sent for review March 19, 2012)

Circumsporozoite (CS) protein is the major surface component of Plasmodium falciparum sporozoites and is essential for host cell invasion. A vaccine containing tandem repeats, region III, and thrombospondin type-I repeat (TSR) of CS is efficacious in phase III trials but gives only a 35% reduction in severe malaria in the first year postimmunization. We solved crystal structures showing that region III and TSR fold into a single unit, an “αTSR” domain. The αTSR domain possesses a hydrophobic pocket and core, missing in TSR domains. CS binds heparin, but αTSR does not. Interestingly, polymorphic T-cell epitopes map to specialized αTSR regions. The N and C termini are unexpectedly close, providing clues for sporozoite sheath organization. Elucidation of a unique structure of a domain within CS enables rational design of next-generation subunit vaccines and functional and medicinal chemical investigation of the conserved hydrophobic pocket.

During a blood meal, Plasmodium sporozoites are transmitted from an infected mosquito to a vertebrate host. Sporozoites are uniformly covered with circumsporozoite protein (CS), a highly expressed and immunodominant surface protein antigen necessary for sporozoite development and host cell targeting (1, 2). Gliding motility enables sporozoites to move directly through host cells to reach the vasculature and later to penetrate the liver. During gliding, CS moves from the anterior to the posterior end of the sporozoite, where it is shed and deposited in a trail (3). CS is composed of an N-terminal domain, a conserved pentapeptide protease cleavage site termed region I, a repeat region, a short sequence containing the linker to the repeat region by their structural basis for vaccine improvement.

The authors declare no conflict of interest.

Data deposition: The atomic coordinates reported in this paper have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 3VDJ, 3VDK, and 3VDL).


1To whom correspondence should be addressed. E-mail: springer@idi.harvard.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1205737109/-/DCSupplemental.
flap at the apex of the domain where the aliphatic portion of the Lys sidechain caps the hydrophobic core (Fig. 2A).

CS binds heparan sulfate proteoglycans (HSPGs) on hepatocytes (11–13). The highly sulfated HSPGs of hepatocytes are proposed to mediate targeting of sporozoites to the liver and trigger sporozoite cell invasion (14, 15). Heparin and HSPG binding sites have been identified in the CS ectodomain (11, 16) and in conserved N-terminal region I and αTSR region II+ peptides (12, 13, 17, 18). The two basic residues in region II+, Arg-345, and Lys-347, are neither highly exposed nor present in a basic patch, and Arg-345 in the β2-strand interacts with Trp-331 in the π-cation bond described above and is also neutralized by a salt bridge to Glu-365 at the adjacent position in the β3-strand (Fig. 2A and B). The complete Plasmodium falciparum CS ectodomain bound to heparin-sepharose under stringent conditions, eluting in 0.35 M NaCl; however, the αTSR domain did not bind, even in absence of salt (Fig. 2H). These results are consistent with mapping of HSPG binding to the conserved basic residues clustered around region I (16, 17) (Fig. L4). In vertebrate hosts, cleavage in region I unmasks the αTSR domain and is thought to permit it to play a role in liver cell invasion (19).

As sometimes seen with short tags that are appended to proteins used in crystallization, the N-terminal YVEF peptide encoded by a restriction enzyme site in the Pichia construct played a role in forming the crystallization lattice. Essentially, this tetrapeptide acted as a hydrophobic probe for the hydrophobic pocket of the αTSR domain. Similar trimers were formed in each of the two independent crystal forms, in which the YVEF peptide from one monomer bound to the hydrophobic pocket of a neighboring monomer around a threefold symmetry or pseudosymmetry axis (Fig. S1). The resulting trimeric, triangular prisms are provocatively similar in shape to that of Plasmodium vivax p25, which is proposed to tile the surface of the ookinete mosquito stage (20); however, we have no evidence that the trimeric form seen in crystals (Fig. S2) is physiologically relevant to packing on the sporozoite surface. Thus, most of the trimeric interface residues are shown below the alignment. (D) Diversity of TSR domain architectures, modified from ref. 7.

Fig. 1. Sequence alignments and fold schematics. (A) Gene architecture of P. falciparum circumsporozoite protein. The portion of CS fused to hepatitis B surface antigen in the RTS,S vaccine is indicated by "vaccine." (B) C-terminal portion of CS from different species. Conserved residues are colored according to chemical type. Disulfide-linked cysteines are colored according to connectivity. Dots below green (hydrophobic) columns mark residues in the hydrophobic core (blue) or exposed hydrophobic pocket (orange). Secondary structure is designated below the alignment by thick lines (helices and β-strands) colored by TSR strand direction (cyan and pink are antiparallel strand directions). Shannon sequence entropy was calculated from 42 distinct P. falciparum isolates from Freetown, Sierra Leone (27) using the Protein Variability Server (46). (C) Alignment with TSR domains from F-spondin (FSP) (9), thrombospondin-1 (TSP1) (7), and ADAMTS13 (10). Unique portions of PFin αTSR are overlined in orange. Only sequences overlined in identical colors are structurally equivalent. Segments forming sheets or helices are shown with thickened lines. Disulfide connectivity is color coded. P. falciparum CS (BD7 strain) mature numbering is shown above the sequences. TSR consensus residues are shown below the alignment. (D) Diversity of TSR domain architectures, modified from ref. 7.
involves burial of the tetrapeptide in the hydrophobic pocket, with only minor contributions from other portions of αTSR monomers. We did not obtain diffraction-quality crystals of the αTSR domain expressed in mammalian cells or expressed in Pichia after replacing the YVEF peptide with the native CS tetra- or nonapeptide N-terminal sequences. αTSR is highly soluble and analytical ultracentrifugation and small angle X-ray scattering showed it to be monomeric, even at concentrations as high as 130 mg/mL (Fig. 2F and Fig. S3). The molecular envelope revealed by combined small and wide-angle X-ray scattering (SWAXS) is in excellent agreement with the crystal structure (Fig. 2F and G), with a $\chi^2$ of 0.93. An indent seen in the SWAXS structure (Fig. 2F, arrows) corresponds to the hydrophobic pocket; however, the YVEF peptide appears to be disordered in the monomer structure in solution because it did not contribute to the SWAXS envelope. Thus, the αTSR monomer in solution adopts essentially the same structure as in trimers in crystals. Furthermore, SWAXS from the mammalian material lacking the N-terminal YVEF peptide shows a scattering profile similar to that of the Pichia material with YVEF and

Fig. 2. Structure of αTSR. (A) CS αTSR showing layer residues and conserved flap fastening residues. (B) Layers of the TSR fold. Dashed boxes indicate layers lost in the αTSR. Dashed blue lines indicate $\pi$-cation and salt bridge interactions retained in αTSR, adapted from ref. 7. (C) Thrombospondin-1 TSR domain 1 (7) superimposed on the αTSR domain in A, Right and separated horizontally. Trp layers are colored orange and stacking layers are labeled by residue type. (D) Hydrophobic core of the αTSR domain. Conserved hydrophobic sidechains (with the exception of I346, in which many species have arginine or lysine sidechains) and their interactions are shown. A portion of the structure is omitted for clarity. (E) Transparent surface representation of αTSR with conserved hydrophobic pocket residues and their surface representation shaded orange. (F) Fitting of the crystal structure of αTSR to the final SWAXS envelope, shown as a transparent surface. The αTSR crystal structure is shown in cartoon in green, except the N-terminal YVEF is in orange. The YVEF sequence may be flexible in solution, because it is not included in the SWAXS envelope. Arrows indicate an observed indent on the SWAXS model corresponding to the approximate location of the hydrophobic pocket in the crystal structure. (G) Scattering curve calculated from the crystal structure and excluding the YVEF sequence (black) is overlaid to the raw experimental SWAXS scattering data for Pichia αTSR (red) and HEK293T αTSR (cyan). (H) Heparin-sepharose affinity assay of Pichia αTSR (red) and CS (black). Each construct was passed over a HiTrap heparin-sepharose column at a flow rate of 0.2 mL/min in 10 mM Tris pH 7.4 buffer and washed with a sodium chloride gradient (brown) over 40 column volumes.
to the profile predicted by the crystal structure (Fig. 2G). Thus, the hydrophobic pocket is unoccupied in solution. Furthermore, the unique αTSR fold is independent of the eukaryotic expression system used, and the N-terminal tetrapeptide extension (like the C-terminal His tag, which is also present in the crystallized construct) has little or no impact on the folded structure.

Discussion

The TSR superfamily is widespread among animals and protozoans, with 187 TSR domains in 41 human proteins (21). However, the αTSR domain in CS is strikingly divergent in sequence and structure from other TSR domains, including those in other Plasmodium surface proteins (8). Residues that form the unique αTSR hydrophobic core and α-helix, and CS flap are highly conserved across 10 Plasmodium species infecting a range of vertebrate hosts (Fig. 1B). The structural adaptations that evolved specifically in the CS αTSR domain change its overall shape. The thin and elongated prototypical TSR domain is shortened on its long axis by 15 Å and widened 10 Å on the other axis by the α-helix and flap (Figs. 1D and 2A–C). Moreover, in TSR domains the N- and C termini are located on opposite ends of the long axis, enabling their tandem use in multidomain proteins (7) (Fig. 1A). In contrast, the N terminus of the CS αTSR domain is located on the same end of the domain as the C terminus (Figs. 1D and 2A and E). This has important implications for how the αTSR is oriented on the P. falciparum sporozoite surface, between the tetrapeptide repeats and C-terminal GPI anchor.

CS is a major sporozoite protein, accounting for 5–15% of total sporozoite [35S]-methionine incorporation (22), and uniformly coats the sporozoite surface (23). Before proteolytic cleavage by a parasite protease in region I, CS is sufficiently densely packed on sporozoites to protect αTSR from antibody recognition (19). How CS forms a protective sheath over the sporozoite plasma membrane and whether dense packing is facilitated by specific packing interactions between neighboring αTSR domains are important questions for further research. The atomic structure of αTSR defined here enables lower resolution investigation of whether the sporozoite sheath contains ordered assemblies of smaller units.

As the immunodominant sporozoite antigen, for many years CS has been the main target of exoerythrocytic vaccine research. During natural infections, young children in endemic regions and adults in nonendemic regions initially develop antibody responses to all three major regions of CS: the N terminus, the central repeat region, and the C terminus containing the αTSR, as shown by reaction with synthetic peptides (24, 25). Repeated infections in endemic regions skew the antibody response to the central repeat region, and reinfections continue to occur.

An important caveat to all previous studies on antibody responses to CS subregions is that the antibodies were interrogated using short synthetic peptides. However, we found here that the αTSR domain adopts a highly specific 3D shape, in which amino residues that are distant in sequence are close in structure. Hence the important epitopes on the αTSR domain are conformational and are unlikely to be well represented by synthetic peptide fragments. Our study now makes available to research a well-folded, highly characterized domain from CS. It will be important not only to test this domain as a vaccine, but also to characterize the antibody responses to the αTSR domain of individuals from endemic regions, and of individuals vaccinated with sporozoites and RTS,S.

The αTSR domain also contains several T-cell epitopes, one of which, CS.T3, is responsible for a CD4+ T-cell response that correlates with protection (26) (Fig. 1B). The other T-cell epitopes, Th2R and Th3R, are polymorphic regions of the αTSR (27). These polymorphic T-cell epitopes and the highest sequence variation between CS isolates locate to the structurally unique elements of the αTSR: the hydrophilic faces of the α-helix and the CS flap (Fig. 1B). Thus, the unique features of the αTSR domain, which endow it with a hydrophobic core, a hydrophobic pocket, and close proximity between its

Table 1. X-ray diffraction data collection, phasing, and refinement statistics for αTSR structure

<table>
<thead>
<tr>
<th>Protein Data Bank deposition code*</th>
<th>3VDJ</th>
<th>3VDK</th>
<th>3VDL</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data collection</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>H32</td>
<td>H32</td>
<td>P41,2,2</td>
</tr>
<tr>
<td>Molecules/asymmetric unit</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td><strong>Cell dimensions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>66.52, 66.52, 85.51</td>
<td>66.53, 66.53, 86.49</td>
<td>63.37, 63.37, 118.16</td>
</tr>
<tr>
<td>α, β, γ (°)</td>
<td>90, 90, 120</td>
<td>90, 90, 120</td>
<td>90, 90, 90</td>
</tr>
<tr>
<td>Wavelength, Å</td>
<td>0.97949</td>
<td>1.07195</td>
<td>0.97949</td>
</tr>
<tr>
<td>Resolution, Å</td>
<td>50–1.70 (1.73–1.70)†</td>
<td>50–1.85 (1.92–1.85)†</td>
<td>50–2.04 (2.11–2.04)†</td>
</tr>
<tr>
<td>Observations (unique)</td>
<td>40,160 (8,160)</td>
<td>41,336 (6,375)</td>
<td>176,960 (15,611)</td>
</tr>
<tr>
<td>l/h overall</td>
<td>22.7 (3.6)</td>
<td>42.6 (2.9)</td>
<td>21.6 (1.44)</td>
</tr>
<tr>
<td>Completeness</td>
<td>99.1 (91.1)</td>
<td>98.1 (86.9)</td>
<td>97.1 (74.7)</td>
</tr>
<tr>
<td>Rmerge</td>
<td>5.7 (35.1)</td>
<td>9.1 (64.3)</td>
<td>9.6 (69.4)</td>
</tr>
<tr>
<td>SAD phasing</td>
<td>0.34/0.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Structure refinement</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution</td>
<td>50–1.70</td>
<td>50–1.85</td>
<td>50–2.04</td>
</tr>
<tr>
<td>Rmerge/Rfree %§</td>
<td>16.9/19.8</td>
<td>17.9/21.5</td>
<td>19.3/24.3</td>
</tr>
<tr>
<td>Protein/solvent atoms</td>
<td>568/103</td>
<td>572/53</td>
<td>1701/90</td>
</tr>
<tr>
<td>Bond length deviation from ideal geometry (Å)</td>
<td>0.006</td>
<td>0.006</td>
<td>0.007</td>
</tr>
<tr>
<td>Bond angle deviation from ideal geometry (°)</td>
<td>0.986</td>
<td>1.0646</td>
<td>1.040</td>
</tr>
<tr>
<td>Ramachandran favored/allowed/outliers (%)</td>
<td>98.6/1.4/0</td>
<td>100/0/0</td>
<td>98.6/1.4/0</td>
</tr>
<tr>
<td>Average B factor for protein/solvent</td>
<td>22.0/33.8</td>
<td>34.9/39.3</td>
<td>32.9/33.8</td>
</tr>
</tbody>
</table>

*Experimental data has been deposited with the indicated deposition ID codes.
†Values in parentheses are for highest-resolution shell.
‡Rmerge = ∑ h|I(h)−〈I(h)〉|/∑ h|I(h)|, where I(h) and 〈I(h)〉 are the nth and mean measurement of intensity of reflection h.
§The Rfree value was calculated using 5% of the data.
N- and C-terminal ends with implications for orientation in the parasite sheath, evolved despite considerable immunological cost, and must be adaptive for the parasite.

The highly conserved hydrophobic pocket in the αTSR domain is particularly fascinating and completely unexpected. Most previous work on the role of the TSR domain of CS in host cell recognition and invasion has focused on putative heparin-binding peptides; however, whereas peptides derived from the TSR region have been shown to bind to heparin, we found that the intact, folded αTSR domain does not. The pocket may have a hydrophobic ligand that is important in parasite-host interactions. The most definitive way to test this hypothesis would be to introduce mutations in this pocket into the parasite genome and test their function in vivo.

A CS fragment containing 19 tetrapeptide repeats and the sequence of the αTSR region is in large phase III trials; however, to date infants and 1-y olds receiving three vaccine doses show only a 35% lower rate of infection in the first year after immunization (6). The current vaccine is produced in the yeast Saccharomyces cerevisiae (28, 29). The CS fragment is fused N terminal to the small hepatitis B virus surface antigen (s-HBsAg) to make the component termed RTS (repeat/TSR/s-HBsAG).

There is no secretion signal sequence N terminal to the CS in this fusion protein. The same yeast cells coexpress intact s-HBsAg, which is termed S, and thus the vaccine product is termed RTS:S. It is unclear whether the αTSR sequence present in RTS:S has the same disulfide-bonded structure or fold as in our crystal structure. Although both proteins are made in yeast, the s-HBsAg displays unusual characteristics. s-HBsAg has an N-terminal transmembrane domain 1 (TM-I) that is not removed, a cytoplasmic domain, a TM-II domain followed by an extracellular loop with a single N-linked glycosylation site, and two further TM domains near the C terminus. Expression of s-HBsAg in mammalian cells results in formation of virus-like particles, lipid-associated particles that are N-glycosylated and secreted (30). However, in yeast, s-HBsAg accumulates in the endoplasmic reticulum (ER), and virus-like particles form after the cell wall is broken and cells are treated with detergents (31). In further contrast to s-HBsAg made in mammalian cells, and to proteins secreted by yeast cells, s-HBsAg made in yeast cells is not N-glycosylated (28). Therefore, it is not known whether the CS fragment fused to s-HBsAg and expressed in yeast has been exposed to the lumen of the ER where disulfide iso- mersases, chaperonins, and other machinery regulate disulfide bond formation, folding, and quality control. Our αTSR construct has a signal sequence, which is removed as shown by mass spectrometry, and is secreted into the medium, and thus has passed ER quality control and is expected to be properly folded.

TSR domains can be fucosylated by the protein O-fucosyl transferase 2 (POFUT2). In TSR domain 1 of thrombospandin, the CSVTCDGG sequence is fucosylated on Thr (7, 32). The fucosylation motif is highly conserved in the CSVTCDGG sequence in CS αTSR. Fucose can be further modified by addition of β-1,3 glucosyl (33, 34). PSI-BLAST searches suggest POFUT2 (and not POFUT1, a homolog involved in Notch fucosylation) is conserved in P. falciparum (gene ID PF04454c) and other apicomplexans. Mass spectrometry of our purified αTSR preparations showed absence of glycosylation of the Pichia protein (calculated 8880.0 Da with four half-cystines, found 8879.9 Da) and presence of both fucose and hexose (e.g., glucose) on the HEK293T protein (calculated 8649.4 Da with four half-cystines, found 8649.1 Da). Lack of fucosylation in Pichia is in agreement with lack of density for fucose in our crystal structure and lack of a POFUT2 homolog in yeast. The possibility of fucosylation of CS has not previously been raised in the malaria literature. Whether CS is fucosylated on the sporozoite surface and how fucosylation affects vaccine efficacy are important questions for the future.

The elucidation of a unique structured domain in CS as the αTSR region, comprising not just the TSR homology region but also region III, is a significant step forward for preerythrocytic malaria vaccines. This work enables subunit vaccine design to transition from sequence based to structure based. Furthermore, it will now be possible to correlate protective and broadly cross-reactive immune responses in vaccinated individuals with specific structural features of αTSR, much as is now being done in the analysis of protective antibodies to influenza and HIV (35, 36).

Materials and Methods
Plasmid Construction, Protein Expression, and Purification. A codon-optimized gene encoding E310-C374 of P. falciparum 37D CS followed by sequence PHHHHHHHH (DNA2.0) was cloned in frame to the S. cerevisiae α-factor signal sequence of pPIC9K (Invitrogen) using EcoRI and NotI sites. The EcoRI site introduced a YVEF sequence before E310. P. pastoris strain GS115 was electroporated with Sall linearized vector. Standard protocols were used for culture growth in buffered complex glycerol media followed by protein expression for 72 h in buffered complex methanol media (Invitrogen; Pichia Expression kit, manual 25-0043).

Culture supernatant supplemented with 0.5 M NaCl, 10 mM imidazole, and 0.5 mM nickel chloride and sodium hydroxide to pH 7.8 was allowed to precipitate overnight at 4 °C. Supernatant (2 L) was applied to a 10-M nickel-NTA agarose (Qiagen) at 5 mL/min. After wash with 0.3 M NaCl, 10 mM Tris pH 8.0, 0.5 M imidazole, and then elution with 0.5 M Tris pH 8.0, 0.5 M imidazole. Fractions containing CS protein were concentrated and further purified by gel filtration on Superdex 75 (GE Healthcare).

The same E310-C374 CS sequence followed by PHHHHHHHH was expressed in HEK293T cell transfectants using the pET8b/PLEXm vectors. Material was purified by nickel-NTA agarose and gel filtration with the same protocol as the Pichia protein.

Recombinant CS was from Science Applications International Corporation (SAIC), the primary contractor for the Malaria Vaccine Production and Support Services (MVPPS) contract (AI-N01-054210) administered through the Division of Microbiology and Infectious Diseases/National Institute of Allergy and Infectious Diseases/National Institutes of Health. (Resides 21–382 of P. falciparum 37D strain (GenBank accession no. CAB383998) was expressed in Pseudomonas fluorescens at Pfenex. CS was 96% pure and 95% monomeric. SDS/PAGE showed a single band at ~55 kDa.

Crystalization and Structure Determination. Protein in 10 mM Tris pH 7.5, 0.25 M sodium chloride (20–40 mg/mL) formed tetragonal crystals in 1.4 M sodium citrate pH 6.1 and trigonal crystals in 0.1 M citrate pH 4.0, 1 M lithium chloride, 20% (wt/vol) PEG 6000. Trigonal crystals were soaked in 0.1 M citrate pH 4.0, 1.0 M lithium chloride, 25% (wt/vol) PEG 6000 with 2.4 mM potassium tetraniplotiniate (II) for 18 h (Hampton Research). Crystals were cryopreserved by incrementing PEG by 5% (in the case of the trigonal crys-

als) and then successive 10-s soaks with 5% (wt/vol) increments of glycerol up to 20% before freezing in liquid nitrogen. Diffraction data were collected on beamline 24ID at APS and indexed and scaled with HKL 2000 (37). The structure was determined by SAD using CCP4 and PHENIX software suites (38, 39). Structure refinement was with COOT and PHENIX (39, 40).

SWAXS Data Collection, Analysis, and Model Reconstruction. Small and wide-angle X-ray scattering was at beam line X9A at the National Synchrotron Light Source (Upton, NY). αTSR samples were passed through a flow capillary, collecting data at 20-s exposures in triplicate at concentrations from 12 to 130 mg/mL (Pichia protein) or 15–45 mg/mL (HEK293T protein). Scattering from 10 mM Tris pH 7.5, 0.25 M NaCl buffer was subtracted. Guinier analysis showed no signs of radiation damage or aggregation. Data were reduced by circular averaging and analyzed using COOT and PHENIX software suites (38, 39). Structure reassembly was with COOT and PHENIX (39, 40).

Mass Spectrometry. Intact protein samples were desalted by reversed-phase microbore HPLC and subjected to electrospray ionization mass spectrometry on a Bruker discovery LC-ion trap mass spectrometer (mass accuracy within 1 Da).

ACKNOWLEDGMENTS. We thank beamlines X9 at National Synchrotron Light Source and 24ID at Advanced Photon Source and David King (Howard Hughes Medical Institute, Department of Molecular and Cell Biology, University of California at Berkeley) for mass spectrometry. This work was supported by National Institutes of Health Grant AI095686.


33. Sato T, et al. (2006) Molecular cloning and characterization of a novel human j1,3-glucosyltransferase, which is localized at the endoplasmic reticulum and glucosylates O-linked fucosylglycan on thrombospondin type 1 repeat domain. Glycobiochemistry 16: 1194–1206.


Supporting Information

Doud et al. 10.1073/pnas.1205737109

Fig. S1. Tyr-Val-Glu-Phe (YVEF) tetrapeptide binding to the hydrophobic pocket. Portion of a trimer is shown. Each monomer is shown in cartoon in a different color. Transparent surface is shown for the non-YVEF portion of the green αTSR monomer. Residues forming the pocket and its surface are colored orange as in Fig. 2E. Val and Phe are buried in the pocket. Each monomer's tetrapeptide is buried in a neighboring monomer's hydrophobic pocket.

Fig. S2. Crystallographic trimer assembly of the αTSR domain and comparison with Pvs25. (A) Crystallographic αTSR trimer viewed parallel to (Upper) and orthogonal to (Lower) the threefold axis. Two nearly identical trimer assemblies were observed in different crystal lattices and space groups. (B) Similar views of the Pvs25 monomer. The Pvs25 monomer was fit to the αTSR trimer by cross-correlation after both structures were resolution filtered to 15 Å. “Superimposed” structures were then separated horizontally on the page to create common views in A and B. Four EGF-like domains are individually colored.
Theoretical monomer Mw (Da) = 8880
AUC Calculated Mw (Da) = 8793

Fig. S3. Analytical ultracentrifugation sedimentation equilibrium. Symbols represent the recorded absorbance values at the given radius after the system reached equilibrium for three αTSR concentrations. Lines represent fit to a self-association model assuming partial specific volume 0.7309 mL/g, solvent density 1.00702 g/mL, and viscosity $1.5422 \times 10^{-2}$ cm$^2$/s and no self-association ($n = 1$). Best fit was achieved for a monomeric model (Mw = 8.79 kDa, $n = 1$). Purified Pichia αTSR domain was diluted in 10 mM Tris pH 7.5, 0.25 M NaCl to final concentrations of 24, 44, and 72 μM and used in sedimentation equilibrium analysis on a Beckman XL-A analytical ultracentrifuge (Beckman Coulter). Samples were centrifuged at 27,000, 32,000, and 45,000 rpm using an An-60 Ti rotor. Sedimentation equilibrium was achieved as verified during data analysis. Sednterp was used to calculate v-bar and buffer density and viscosity. SEDFIT was used to process the raw data and SEDPHAT was further used to perform a global fitting of the data to self-association models (1).