Published in final edited form as:
*Proteins*. 2011 April ; 79(4): 1048–1060. doi:10.1002/prot.22933.

# Introduction of a Disulfide Bond Leads to Stabilization and Crystallization of a Ricin Immunogen[§]

**Jaimee R. Compton**[†], **Patricia M. Legler**[†], **Benjamin V. Clingan**, **Mark A. Olson**[1], and **Charles B. Millard**[*]
Department of Structural Biology, Walter Reed Army Institute of Research, Silver Spring, MD 20910

[1]Department of Cell Biology & Biochemistry, U.S. Army Medical Research Institute of Infectious Diseases, Frederick, MD 21702.

## Abstract

RTA1-33/44-198 is a catalytically inactive, single-domain derivative of the ricin toxin A-chain (RTA) engineered to serve as a stable protein scaffold for presentation of native immunogenic epitopes (Olson, *et al*., Protein Eng Des Sel 2004 17:391-7). To improve the stability and solubility of RTA1-33/44-198 further, we have undertaken the design challenge of introducing a disulfide (SS) bond. Nine pairs of residues were selected for placement of the SS-bond based on molecular dynamics simulation studies of the modeled single-domain chain. Disulfide formation at either of two positions (R48C/T77C or V49C/E99C) involving a specific surface loop (44-55) increased the protein melting temperature by ~5°C compared with RTA1-33/44-198 and by ~13°C compared with RTA. Prolonged stability studies of the R48C/T77C variant (>60 day at 37°C, pH 7.4) confirmed a >40% reduction in self-aggregation compared with RTA1-33/44-198 lacking the SS-bond. The R48C/T77C variant retained affinity for anti-RTA antibodies capable of neutralizing ricin toxin, including a monoclonal that recognizes a human B-cell epitope. Introduction of either R48C/T77C or V49C/E99C promoted crystallization of RTA1-33/44-198, and the X-ray structures of the variants were solved to 2.3Å or 2.1Å resolution, respectively. The structures confirm formation of an intramolecular SS-bond, and reveal a single-domain fold that is significantly reduced in volume compared with RTA. Loop 44-55 is partly disordered as predicted by simulations, and is positioned to form self-self interactions between symmetry-related molecules. We discuss the importance of RTA loop 34-55 as a nucleus for unfolding and aggregation, and draw conclusions for ongoing structure-based minimalist design of RTA-based immunogens.

### Keywords

aggregation; unfolding; subunit vaccine; toxin; protein engineering; crystal packing

[§]This work was supported by the U.S. Defense Threat Reduction Agency JSTO award S.S.0003_06_WR_B (CBM) and National Institutes of Health U01 A1082120-01 (CBM). The opinions or assertions contained herein belong to the authors and are not necessarily the official views of the U.S. Army or the U.S. Department of Defense. Atomic coordinates for the crystal structures of the R48C/T77C and V49C/E99C disulfide variants of RTA1-33/44-198 have been deposited in the Protein Data Bank as RCSB codes 3LC9 and 3MK9, respectively.

[*]To whom correspondence should be addressed: U.S. Army Medical Research & Materiel Command, Frederick, MD 21702-5012. Tel: (301) 619-7400; Fax: (301) 619-2982, charles.b.millard@us.army.mil. .
[†]These authors contributed equally to this work

[1]Amino acid numbering of RTA is used throughout beginning with N-terminal sequence: IFPKQYP

## Introduction

Rational design of improved vaccines begins with the solved or modeled tertiary structure of a target macromolecule and seeks to optimize a stable sub-structure that retains sufficient mimicry of native epitopes to elicit and focus an effective immune response. 1-3 A major protein engineering challenge arises because the target immunogen of interest often has been isolated from a larger biological complex with little high-resolution information available describing either its natural quaternary structure, or the binding complexes it forms with neutralizing antibodies. Despite these difficulties, structure-based optimization has shown promise for improving recombinant immunogen homogeneity, stability and production yields, as well as rendering complex proteins suitable for a wider array of adjuvant and vaccine delivery strategies, such as virus-like particles or other multivalent display technologies.[1,4-6]

Ricin is a type II ribosome-inactivating protein (RIP) toxin that is known as a lethal poison for bioterrorism and biological warfare, as well as a potential medical treatment for cancer or other diseases of abnormal cell proliferation.[7,8] Effective human vaccines and antibodies are sought to nullify or, in the case of clinical applications, to control the cytotoxicity of ricin. Several immunogens based upon the ricin toxin A-chain (RTA) subunit have been produced and shown to protect animals against ricin exposure or to induce neutralizing antibody titers in humans.[9-11] A major shortcoming of RTA-based vaccines is the tendency of the protein immunogen to aggregate during expression, purification or storage at ambient temperatures, thereby limiting production yields and potentially requiring an impractical cold-chain.

We hypothesized that the aggregation problem was caused largely by selected portions of RTA that maintain the protein-protein interface with ricin toxin B-chain (RTB) in the natural holotoxin complex.[12] Interfacial amino acids become exposed on the surface of purified RTA where they promote self-aggregation in the absence of RTB or lipid membranes. The artificially exposed structures on RTA immunogens also may stimulate the costly production of irrelevant antibodies that are unable to bind the ricin holotoxin *in vivo* due to steric occlusion by RTB. To overcome these limitations, we have adopted a minimalist approach to shift the structure of RTA away from its natural roles as an N-glycosidase and obligate polypeptide partner of RTB in the dichain complex, and toward a smaller and more stable fold that lacks catalytic activity while retaining a high fraction of surface-based epitopes that are shared with ricin holotoxin.

RTA is classified as a complex two-domain fold with a putative N-terminal domain covering the range of residues 1-179 and a C-terminal region of 180-267 (Fig. 1).[13] Truncation of residues 199-267 removes a large segment of the RTA-RTB interface, resulting in an independently folding protein with a modest improvement in stability compared with wild-type RTA.[12,14] However, an approximately 20-residue loop segment (residues 34-55) from the RTA1-198 N-terminal domain apparently is dislodged by removal of the C-terminal region and becomes exposed to significant hydration effects (Fig. 1). To shorten this solvent-exposed loop at positional sites along the chain that permit loop closure, we also removed RTA residues 34-43. Combination of the deletions yields a protein scaffold (called RTA1-33/44-198[1]) that shows greater resistance to thermal denaturation, less aggregation under physiological conditions, and a reduction in toxic N-glycosidase activity of at least

§This work was supported by the U.S. Defense Threat Reduction Agency JSTO award S.S.0003_06_WR_B (CBM) and National Institutes of Health U01 A1082120-01 (CBM). The opinions or assertions contained herein belong to the authors and are not necessarily the official views of the U.S. Army or the U.S. Department of Defense. Atomic coordinates for the crystal structures of the R48C/T77C and V49C/E99C disulfide variants of RTA1-33/44-198 have been deposited in the Protein Data Bank as RCSB codes 3LC9 and 3MK9, respectively.

three orders of magnitude compared with RTA; the engineered immunogen combined with adjuvant protects mice or non-human primates against a supra-lethal ricin aerosol.[12]

A recent computational study of the thermal unfolding of RTA1-33/44-198 using coarse-grained lattice models with all-atom reconstruction and molecular-dynamics (MD) simulations predicted pockets of hyper-mobility.[15] Identification of specific regions of local disorder during unfolding provides a rational basis for the follow-on strategy of introducing precisely localized disulfide (SS-) bonds to improve RTA1-33/44-198. Optimally positioned SS-bonds enhance protein stability by reducing the configuration entropy of the unfolded state[16-18] and by exerting local effects on the folded state.[19-21] Additionally, SS-bonds may slow the rate of irreversible transition of unfolded states to insoluble aggregates by limiting thermal motions that expose hydrophobic surfaces.[22,23] Disulfide design is imperiled by several risks, however, including failure of the SS-bond to form during expression, the possibility of mixed disulfide formation when introducing multiple Cys residues, and by the unintentional entrapment of unproductive or misfolded states during protein folding.

We present herein the design and characterization of RTA1-33/44-198 variants containing engineered SS-bonds that confer improved resistance to thermal unfolding and aggregation. The proteins are easily expressed and purified from *E. coli* without a need to refold or otherwise induce SS-bond formation, and they retain the ability to bind toxin-neutralizing monoclonal antibodies (MAb). Variants containing an SS-bond at positions R48C/T77C or V49C/E99C readily crystallized, thereby permitting us to solve the three dimensional structure of the RTA1-33/44-198 scaffold by X-ray diffraction. The structure reveals a single domain fold with a significantly smaller volume than the RTA fold and further corroborates the hypothesis that conformational mobility within a specific loop region (residues 34-55) contributes to irreversible RTA unfolding through protein aggregation.

## Materials and Methods

### Materials

Unless otherwise stated, chemicals were purchased from Sigma-Aldrich (St. Louis, MO). Q Sepharose, SP Sepharose and PD-10 columns were purchased from GE Healthcare Life Sciences (Amersham Pharmacia, Piscataway, NJ). QuikChange™ kits were from Stratagene (Agilent Technologies, La Jolla, CA). Plasmid purification kits were from Qiagen (Hilden, Germany). BugBuster® Protein Extraction Reagent was from Novagen (Darmstadt, Germany). Syringe filters (0.1 μm) were from Millipore. *E.coli* BL-21(DE3) cells were from Invitrogen (Carlsbad, CA). Crystal Screen Cryo solution number 31 was from Hampton Research (Aliso Viejo, CA). ProtoBlot ® II AP System mouse immunoblotting kit was from Promega (Madison, WI). Purified RTA from *Ricinus communis* was purchased from Vector Laboratories (Burlingame, CA).

### Selection of sites for introducing SS-bonds

All-atom MD simulations were used to identify structural regions of the RTA immunogen that contribute to early-stage protein unfolding. Within these regions, pairwise residue sites (Table I) were selected for placement of the SS-bonds to hinder thermal unraveling of the molecule. The selection of sites was also guided by a previous simulation study of RTA and the truncated chain using a coarse-grained lattice model.[15] The starting modeled structure of the RTA immunogen was built from the PDB structure 1RTC[13] and crystallographic waters were deleted. Residues in the loop region 34–43 were removed and the C-terminal region of residues 199– 267 was truncated. The model structure was subjected to energy minimization using steepest descent and adopted-basis Newton-Raphson minimization for a total of 100 steps. The forcefield was set with CHARMM22 and solvent effects during minimization

were treated by a protein distance-dependent dielectric screening model of $\varepsilon = r$. Non-bonded interaction cutoff parameters for electrostatics and van der Waals terms were set at a radius of 22 Å with a 2-Å potential switching function.

Exploration of the conformational landscape for detecting unfolding events was carried out using replica-exchange simulations methods implemented in the Multiscale Modeling Tools for Structural Biology (MMTSB).[24] A total of 32 replicas were used and exponentially spaced from a temperature *T* of 298 to 550 K. The potential energy function consisted of CHARMM22 and the generalized Born molecular volume solvent model. A simulation of 1 ns was applied and the simulation protocol was similar to that used to study thermodynamic folding-unfolding of SH3.[25] Structural regions that drive the onset of unfolding were identified by monitoring positional fluctuations and the disruption of native contacts. These regions were further analyzed by computing a Cα-Cα coordinate distance matrix and looking for residue side-chains and their torsional angles that allow for the formation of a SS-bond. A list of candidate substitutions was constructed and compared using the mutation tool of MMTSB. Sites that led to significant structural distortion or high strain energy were eliminated. In addition to the final list of proposed placements from modeling, we included substitution sites in the RTA immunogen that are structurally homologous to the SS-bond observed in the N-terminal domain of the RIP pokeweed antiviral protein (PAP).

## Site-directed mutagenesis

We started with a synthetic gene for RTA1-33/44-198 optimized for *E. coli* expression using proprietary GeneForge™ software (Aptagen Inc., Herndon, VA). Pairs of Cys residues were introduced sequentially using primer-directed site-specific mutagenesis with a QuikChange™ kit. Plasmid DNA was purified and sequenced to verify that only the intended mutations had been introduced into the gene sequence. The DNA was then transformed into *E.coli* BL21(DE3) for protein expression.

## Protein expression and purification

*E.coli* cells were grown in LB media containing kanamycin (50 μg/ml) at 37°C to a cell density of 0.8-1.0 $OD_{600}$. Cultures were induced with 0.2 mM isopropyl-b-thiogalactoside for 18-20 hours at 17°C. Cell pellets were suspended in 50 ml of lysis buffer (50 mM sodium phosphate buffer, pH 7.3, 2 mM EDTA, 30% BugBuster®) and sonicated, 30 seconds on/30 seconds off, for a total of 2 minutes. The lysate was clarified by centrifugation at 9,000 x g for 1 h at 4°C. The supernatant was loaded onto a Q Sepharose column equilibrated with 50 mM sodium phosphate buffer, 2 mM EDTA, pH 7.3, and the flow through was collected. Fractions containing the protein (~22 kDa) were combined and dialyzed overnight at 4°C against 1L of 50 mM MES, 2 mM EDTA, pH 6.4. Dialyzed protein was loaded onto an SP Sepharose column equilibrated with 50 mM MES, 2 mM EDTA, pH 6.4. Protein was eluted from the column using a 0 to 250 mM sodium chloride gradient, and was analyzed using SDS-PAGE. The fractions containing a single band on SDS-PAGE gels were combined and dialyzed into phosphate buffered saline (PBS; 137 mM NaCl, 2.7 mM KCl, 8mM sodium phosphate dibasic, 1.5 mM potassium phosphate monobasic, pH 7.4). The identity and composition of each purified protein was confirmed by trypsin digestion, separation and analysis of polypeptides by matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry (details in Supplemental Table 1).

## Thiol titrations

Free thiols were quantified by the method of Ellman.[26] Briefly, stock solutions of purified protein (0.85-1.1 mg/ml) were divided into volumes of 50, 100, 150, or 200 μl and incubated in a water bath at 100°C for 2.5 min. Tubes were cooled on ice (approximately 5 min),

centrifuged for 5 sec to remove condensation from lid, and then mixed with 25 μl of a 20 mM stock of DTNB. The appropriate amount of PBS, pH 7.4, was added to a final volume of 1 ml. Tubes were then incubated at room temperature without light for 20 min. After centrifugation at 14,000 x g to pellet denatured protein, each reaction was transferred into a cuvette and the absorbance at 412 nm was measured. The molar concentrations of $TNB^-$ versus protein were plotted and the resulting slope yielded the number of free Cys in each monomer.

### Circular dichroism (CD) and thermal denaturation studies

Protein unfolding as a function of temperature was assessed by monitoring loss of ellipticity at 222 nm using a Jasco-810 spectropolarimeter with Peltier thermal controller. Experiments were routinely conducted between 10 to 80°C at a temperature scan rate of 2 degrees per min and a protein concentration of ~0.2 mg/ml in PBS, pH 7.4. The protein melting temperature was determined as the mid-point from a four-parameter fit of the averaged CD signal from two scans versus temperature.

### Protein aggregation studies

Purified protein was sterile filtered (0.1 μm syringe filter), divided into 0.5 ml aliquots, and incubated at 37 °C. At various time intervals over the course of up to 61 days, aliquots were removed from the incubator and centrifuged at maximum speed to pellet protein aggregates. The amount of soluble protein remaining in each tube was determined from absorbance at 280 nm using a calculated extinction coefficient $\varepsilon = 18{,}005\ M^{-1}cm^{-1}$.

### Immunoblotting

Western blots containing the variants described in Table I were probed with monoclonal antibody (MAb) UNIVAX 70/138 or GD12. UNIVAX 70/138 is a murine IgG with an epitope that has been mapped to RTA N97-F108; it is one of the highest affinity ricin neutralizing MAbs studied to date.[27,28] GD12 is a murine IgG that binds RTA T163-M174[29]; this sequence corresponds with a linear, human B-cell epitope at RTA L161-I175.[30] Primary antibody binding was detected with an alkaline phosphatase-conjugated secondary antibody.

### Crystallization of SS-bond variants

Protein was purified further using gel filtration on a G-200 Superdex column (1×39 cm) equilibrated with 50 mM MES pH 6.4 with 200 mM NaCl for crystallization trials. In some trials, 2 mM BME was included in the gel filtration buffer. Protein at approximately 2.75 mg/mL was diluted 1:1 in hanging drops with each of the 48 different precipitant solutions of Hampton Crystal Screen Cryo. Suitable sheet-like protein crystals of the RTA 1-33/44-198 R48C/T77C variant were obtained from condition #31 (0.17 M ammonium sulfate, 25.5% PEG 4000, 15% glycerol) and improved in thickness over a period of 2-3 months. Once conditions were identified, crystals of the V49C/E99C variant also were grown.

### X-ray data collection and refinement

Crystals were transferred to paratone-N, flash-frozen in liquid nitrogen and mounted in a cryostream at 100 K for diffraction data collection using an in-house Bruker FR591 high flux, rotating anode X-ray diffractometer (PROTEUM) and a SMART 6000 2K CCD detector at WRAIR. Structures were solved by molecular replacement using Amore[31] with initial phases taken from a search model of RTA1-33/44-198 that was based closely on the solved structure of RTA; the overall root-mean-square deviation (rmsd) between search model and starting RCSB structure 1IFS was 0.20 Å over 174 C-α atoms. Simulated

annealing was done with CNS 1.1. Refinement was done by iterative analysis of 2Fo-Fc and Fo-Fc maps with CCP4[32]; the B-factors were refined for each atom individually.

## Results

### Site selection for SS-bonds based on MD simulations

Figure 1 shows the structure of RTA and our computer-based homology model of RTA1-33/44-198. Sites for intramolecular SS-bonds were selected based on MD simulations of the modeled RTA1-33/44-198 chain (Fig. 1). Three specific regions comprising RTA residues 44-56, 90-110 or 140-170, respectively, showed early unfolding progression as a function of the MD simulation time (Fig. 1C). SS-bonds were proposed within each hypermobile region, and several sites were linked to neighboring areas for bond formation (Fig. 1D).

### Site selection based on homology

PAP is a member of the RIP family with a homologous protein folding pattern with RTA, but PAP contains a SS-bond in the N-terminal domain at residues 86 and 105 (numbers from PDB 1GIK). Consequently, starting from structure-structure alignments with PAP, we selected RTA sites A90 and F108 or V111 for substitution with Cys (Fig. 1E). Although these residues do not align exactly with the Cys sites in the PAP structure, the coordinate distance between the residues in RTA provided the optimal location and geometry to mimic the PAP SS-bond. The secondary-structure composition for A90 and F108 is a β-strand and a coil region, respectively, where the latter follows a short α-helix segment which is variable among the RIPs (Fig. 1F).

### Effects of mutations on protein expression and folding

Of nine double mutants constructed, seven were readily expressed in the soluble protein fraction of *E. coli* during an overnight induction period at 17°C (Table I). Two of the double mutants attempted, A90C/F108C and L129C/A165C, were localized to insoluble inclusion bodies; these variants were not studied further. Mutations which altered charged residues did not change the calculated pI by more than 3 pH units, and all of the soluble proteins could be purified by the same protocol.

### Verification of SS-bond formation

Using titration of free thiols with DTNB, the number of reduced Cys in each purified protein was determined to verify the formation of SS-bonds (Table I). RTA 1-33/44-198 yielded a molar reaction stoichiometry with $TNB^-$ of 0.97 ± 0.03, consistent with a single free thiol in the protein at Cys171. Likewise, the variants R48C/T77C and V49C/E99C yielded a stoichiometry of 1.00 ± 0.08 and 0.96 ± 0.05, respectively, consistent with one free thiol at Cys171 and formation of a single intramolecular SS-bond between the introduced Cys pairs. Non-reducing SDS-PAGE gels and Western blots confirmed that each purified protein was primarily a monomer (Supplemental Fig. 1). Variants that failed to increase $T_m$ values also consistently yielded a $TNB^-$ reaction stoichiometry > 1, demonstrating that the designed SS-bonds did not form fully (Table I).

### Effects of SS-bonds on protein melting temperature ($T_m$)

Thermal denaturation of RTA1-33/44-198 and the SS-bonded variants was irreversible and linked to aggregation in 1x PBS pH 7.4 as judged by an increase in solution turbidity and failure to reverse the heat-induced loss of ellipticity upon cooling. This behavior precluded calculation of ΔG of folding. Nevertheless, a single sigmoidal transition was observed for

each variant in the absence of reducing agents (Supplemental Fig. 2), thus permitting consistent measurement and comparison of $T_m$.

Of the seven different double mutants purified and studied, three SS-bonded variants showed higher $T_m$ compared with RTA1-33/44-198 lacking the SS-bond; two variants had no effect, and two resulted in a lower $T_m$ (Table I). Two of the sites that led to a significant increase in $T_m$ inserted an SS-bond near the truncated RTA loop (34-43). The measured $T_m$ values of the R48C/T77C and V49C/E99C variants were 5°C higher than RTA 1-33/44-198. When compared to RTA alone, the $T_m$ values of the R48C/T77C and V49C/E99C variants were approximately 13°C higher.[12] Additionally, one of the sites selected based on homology with PAP, A90C/V111C, formed a disulfide bond that contributed to an increase in $T_m$ of 2.7 degrees (Table I, Figure 1). A fourth variant, N141C/I192C, also showed a small but significant increase in $T_m$ of 1.5 degrees.

### Effects of SS-bonds on protein aggregation

Introduction of the R48C/T77C SS-bond reduced the rate and total amount of aggregation of RTA1-33/44-198 in solution (Fig. 2; Table II). The effect on aggregation was dependent upon the total protein concentration (Table II).

### Addition of an SS-bond promotes crystallization

Extensive attempts to crystallize RTA1-33/44-198 using conventional hanging drop, sitting drop or free-interface diffusion under thousands of different combinations of buffer, additives and precipitants all failed. In contrast, two of the SS-bond variants (R48C/T77C or V49C/E99C) produced near diffraction-quality crystals in 2 of the first 48 screening conditions attempted at 17°C. Protein drops of RTA1-33/44-198 lacking the SS-bond failed to crystallize under identical conditions. Even when micro-seeded with crystals of the SS-bond variants, diffraction quality crystals of RTA1-33/44-198 were not obtained.

### X-ray structure of the RTA1-33/44-198 R48C/T77C variant

Orthorhombic crystals of the R48C/T77C variant diffracted to 2.1 Å, and all data from 57-2.3 Å were used for structure refinement.. We compared two possible space groups for refinement, I222 or pseudo-symmetry with P21212. The specific parameters considered included the figure-of-merit for data indexing, Patterson maps, comparison of R- and $R_{free}$-values during refinement in each space group, as well as inspection of the quality of the electron density maps. Based on our assessment of these data, the space group I222 was selected for refinement with one molecule in the asymmetric unit (Table III). The refined crystal structure is a monomer with a single free Cys171 and an SS-bond between Cys48 and Cys77. Density for the C48-C77 SS-bond could be seen in the 2Fo-Fc map (Fig. 3) but uniform density in the loop containing R48C (residues 44-55) was not observed, consistent with disorder.

An X-ray diffraction dataset also was collected from isomorphous crystals of the V49C/E99C variant of RTA1-33/44-198 and refined from 58-2.1Å resolution (Table III). The refined structure produced results that are similar with those of the R48C/T77C variant, confirming the essential features described herein for the RTA1-33/44-198 protein fold.

Superposition of the refined structure of RTA1-33/44-198 R48C/T77C with the corresponding sequence of RTA (PDB 1IFS) resulted in an overall rmsd of 0.59 Å over 156 Calpha. The close similarity of the RTA1-33/44-198 crystal structures described herein with the first 198 residues of RTA (Fig. 4) is consistent with earlier predictions of secondary structure content based upon CD and FTIR.[14] Subtle structural differences between the two structures include the C-terminal residues 181-188, and a loop between residues 132-137 at

a crystal packing interface. In RTA, the residues 181-183 form part of a reverse turn between two helical segments, whereas in the SS-bonded variant these residues form a helix without a turn.

The structures demonstrate that the N-terminal region of RTA can fold independently in the absence of the C-terminal domain or RTB, supporting theoretical prediction of this region as a distinct domain. The N-terminal domain is significantly more compact than RTA because it lacks 79 amino acids; the volume of RTA1-33/44-198 V49C/E99C was 27,177Å$^3$ (1,444 atoms) compared with 36,707Å$^3$ (2,038 atoms) for RTA (1IFS) when calculated from the crystal structures using the solvent accessible surface with a 1.4Å radius rolling probe.[33]

From the refined coordinates of the two RTA1-33/44-198 crystal structures, we compared the theoretical dihedral strain energies for the engineered SS-bonds by the method of Hogg and colleagues[34], as well as the programs *SS-bond*[35] and *Disulfide by Design*.[36] All three approaches calculated significantly higher strain energy for the C48-C77 SS-bond compared with the C49-C99 SS-bond. Although these approximate methods have been correlated with empirical data for other proteins,[34,37,38] our results must be considered preliminary until crystal structures of sufficient resolution are available to permit more rigorous energy calculations.

## Comparison of refined B-factors

Comparison of the B-factors among the three structures shows that loop 44-55 is the most disordered region found within each of the SS-bond stabilized variants (cf., Fig. 4A-C). The electron density was improved for the loop 44-55 in the V49C/E99C structure compared with that of the R48C/T77C structure, but still disordered relative to the rest of the protein. Figure 4D shows the relative average B-factors by residue for the refined X-ray structures of RTA (PDB 1IFS) and SS-bonded variant RTA1-33/44-198 R48C/T77C. To normalize differences introduced by refinement approaches for the two structures, we divided the average B-factor for the main chain atoms of each residue by the average B-factor calculated for all atoms of the structure. The most striking finding of the plot is the increased B-factors for loop 44-55 in the structure of the SS-bond variants. Additionally, average B-factors are lower for the RTA1-33/44-198 across residues 150-165 (Fig. 4D); these residues form part of a helix-turn-helix motif and participate in intermolecular crystal packing for both proteins.

## Intermolecular crystal packing

An examination of intermolecular packing within the crystals revealed that loop 44-55 in RTA1-33/44-198 is positioned to form self-self interactions between symmetry-related molecules (Fig. 5). Hyper-mobility of this loop in the absence of the SS-bond could interfere with intermolecular packing of the crystal. As noted, this loop also was among the most mobile regions in MD simulations. The closest distance of the Cys pairs between two symmetry molecules is approximately 14Å, making formation of an intermolecular SS-bond unlikely. Because of the disorder observed for this loop, however, we cannot rigorously exclude the possibility that a small fraction of intermolecular SS-bonds could form in the crystalline state.[39]

## Neutralizing antibodies bind purified SS-bond variants

The locations of several epitopes that consistently elicit toxin-neutralizing antibodies in mice or humans have been identified on RTA.[30,40] Of particular interest, a human B-cell epitope is formed by RTA residues L161-I175.[30] It was suggested previously that the structural alterations in RTA required to arrive at the smaller RTA1-33/44-198 fold might compromise the binding of antibodies to important toxin neutralizing epitopes.[9] Comparison of the X-ray

structure of RTA1-33/44-198 R48C/T77C with that of RTA (PDB 1IFS) shows that residues L161-I175 are superposable for the two structures (Supplemental Fig. 3). Moreover, RTA1-33/44-198 and each of the SS-bond variants tested bound the UNIVAX 70/138 and GD12 MAbs on Western blots (Supplemental Fig. 1). One of the variants, V49C/E99C, showed weaker or variable binding with GD12 using this method. Although the Western blots provide suggestive evidence of retention of selected neutralizing epitopes, this method is not a substitute for follow-on immunological studies of the capacity of the RTA variants to induce protective antibodies *in vivo*.

## Discussion

General rules for introducing stabilizing SS-bonds have proven difficult to derive, even for single domain globular proteins that undergo reversible folding-unfolding transitions under laboratory conditions.[18,20,21,41] Many proteins, like RTA, undergo essentially irreversible thermal unfolding linked partly to aggregation. Aggregation limits strict interpretation of SS-bond effects because there is no general relationship among thermodynamic stability, the thermal dependence of aggregation kinetics, or the final percentage of protein that remains soluble.[42,43] Nevertheless, improved resistance to thermal denaturation has been achieved empirically for several proteins of pharmaceutical or industrial importance by introduction of a single SS-bond, including carbonic anhydrase II (A23C/L203C)[44], *Drosophila* cholinesterase (I327C/D375C)[22], alkaline protease AprP (G199C/F236C)[45], single-chain immunoglobulins (A49C/I70C)[46], manganese peroxidase (A48C/A63C)[47], alpha dextranase (D279C/S289C)[48], and xylanase (V98C/A152C).[49]

Factors typically considered in selecting sites for introduction of stabilizing intramolecular SS-bonds include geometric constraints extracted from databases of solved structures, MD simulations, local entropic and enthalpic effects of specific residue substitutions with Cys and, if available, comparison with structurally similar or homologous proteins.[41,50] Estimates of geometric limits are useful for excluding Cys pairs that are unlikely to form SS-bonds, but this approach is insufficient to predict the effect of the proposed SS-bonds on protein stability. Experimental studies of several different proteins, including RTA1-33/44-198 (Table I), consistently show that formation of an SS-bond can have positive, negative, or no effect on protein stability.[41,51]

In addition to multi-resolution MD simulations, we considered homology for selection of two of the RTA1-33/44-198 variants because several X-ray structures of RTA and a close structural neighbor, PAP, are available. One of the homology variants (A90C/V111C) did result in a new disulfide and an increase in the $T_m$ by 2.7 degrees (Table I). Of the presently known RIP family, only PAP contains this SS-bond in the N-terminal domain. Consequently, the observation that C90-C111 forms and enhances the $T_m$ for RTA1-33/44-198 raises the possibility of exploring this position for stabilization of other RIPs.

### Effects of SS-bond on RTA1-33/44-198

Disulfide bonds promote protein folding primarily by decreasing the configuration entropy of the unfolded state.[16,52] This effect on $\Delta S$ has been shown to be proportional to the separation between bonded Cys residues for several globular proteins.[18,21] Of nine Cys pairs studied herein, the two that resulted in the largest increases in $T_m$ for RTA1-33/44-198 were both to the same loop but separated by 29 (C48-C77) or 50 (C49-C99) residues (Table I). Despite a large difference in the length of intervening polypeptide sequences for these two proteins, an SS-bond to the loop 44-55 resulted in comparable effects on $T_m$. The lack of correlation between the length of intervening polypeptide sequence and observed $T_m$

suggests that configuration entropy alone does not fully account for the stabilization observed.

Local entropy and enthalpy effects on the folded state, especially strain energy introduced by the engineered SS-bond, also can make important contributions to the net protein stability.[19-21,53] Significant deviation in the SS-bond dihedral χ-angle values in the refined crystal structure of the R48C/T77C variant compared with dihedral values commonly observed implies a greater degree of local strain. The C48-C77 dihedral angles are -120, 53, 81, 121, -86 degrees for $\chi_1$, $\chi_2$, $\chi_3$, $\chi_2$', and $\chi_1$' respectively; whereas, expected angles of a comparable class 4 right-handed disulfide are -47, -116, 94, -63, and 172, respectively.[53] Estimation of strain energy for C48-C77 bond using the program *SS-bond*[35] yielded a much less favorable value of 6.3 kcal/mol compared with either the C49-C99 bond (2.3 kcal/mol) or the average strain energy (3.1 kcal/mol) calculated for over 42,000 different SS-bonds from X-ray structures in the PDB.[34] The large difference in theoretical strain energy between the two bonds may be related to the fact that C48-C77 links a mobile 44-55 loop with a turn, whereas C49-C99 links the loop to an alpha helix (Fig. 3).

Reduction in local strain energy may be possible by optimizing the length or composition of the 44-55 loop, for example by replacement of one or both proline residues bounding the loop (Pro46 and Pro52). Disulfide stabilization energy has been shown to depend upon the perturbation energy for the peptide bond isomerization of nearby Pro residues.[54] Failure to clearly resolve loop electron density in the R48C/T77C variant is consistent with multiple local configuration states and heterogeneity in the cis-/trans-isomers at Pro46 and Pro52.

## Importance of kinetic processes

Configuration entropy and disulfide strain both would predict that the V49C/E99C variant should be more stable than the R48C/T77C variant, yet these proteins exhibit essentially identical $T_m$ values (Table 1). This underscores the difficulty of predicting protein stability or designing SS-bonds *de novo* from geometry or thermodynamic principles alone. In addition to thermodynamic effects, the kinetics of irreversible aggregation likely contributes to the observed increase in $T_m$ for the RTA1-33/44-198 variants (Table II). It has been pointed out that partial (local) unfolding may initiate irreversible processes such as aggregation or proteolysis, underscoring the importance of local stabilization in conjunction with global thermodynamic stabilization of the protein fold.[55] Recognizing that RTA unfolding was linked with irreversible aggregation, we focused intentionally on disulfide stabilization of three specific structural regions that contribute to early-stage, partial protein unfolding as judged by high thermal motion observed in RTA1-33/44-198 during MD simulations. It was our contention that anchoring these regions should help reduce conformational disorder and thus provide favorable resistance to thermal denaturation of the RTA immunogen.

## Hypothesis that RTA34-55 loop is a nucleus for unfolding

Stabilizing mutations tend to cluster in specific regions of protein folds, especially regions involved in the early stages of unfolding.[43] One of the regions we selected for analysis in RTA is the residual portion (residues 44-55) of a longer surface loop in RTA (residues 34-55). Its high thermal motion in MD simulations, as well as the disorder observed in the crystal structures (Fig. 4), are consistent with the hypothesis that loop 34-55 is a potential nucleus for RTA unfolding. A protein construct comprising RTA_1-198 that retains loop 34-43 was intermediate between full length RTA and RTA-133/44-198 with respect to its $T_m$ and aggregation propensity[14]; thus, retention of the first 10 residues of loop 34-55 destabilizes RTA. Moreover, introduction of an SS-bond to the residual loop (residues 44-55) significantly reduces both the rate and total amount of self-aggregation observed for

RTA1-33/44-198 (Table II: Fig. 2). By restricting local disorder and partial unfolding, the SS-bonds may reduce the prevalence of conformational states with exposed hydrophobic surfaces that promote self-aggregation.

The potentially overlapping mechanism(s) by which the Cys mutations or SS-bonds led to crystallization cannot be determined precisely but could involve slowed protein aggregation kinetics, local structural effects of the SS-bond on the conformation of the loop 44-55, or effects of the residue substitutions themselves. It is intriguing to note that the mobile loop 44-55 is positioned to form a self-self intermolecular interaction within the higher order crystal structure (Fig. 5). While residues participating in such interactions are typically among the most ordered in a protein crystal structure, we observe significant disorder in loop 44-55. This may be caused by its inherent flexibility, as well as partial SS-bond reduction under conditions of crystallization and X-ray analysis.

A hypothetical role for loop 34-55 as a nucleus for partial unfolding and aggregation raises the question of why natural RTA retains such intrinsically destabilizing features. It has been proposed that quasi-native “molten globule” states of RTA have a biological function in efficient toxin translocation across amphipathic biological membranes.[56] Insertion of an SS-bond within the RTA C-terminal region at position Cys215-Cys255, for example, impaired a conformational change important for toxin function.[57] Consistent with this proposal, monitoring secondary structure during chemical denaturation showed that RTA unfolding fits a multi-state model that likely includes at least one non-native state with exposed hydrophobic surfaces, whereas the unfolding of RTA1-33/44-198 more closely approximates a two-state model.[14] Complete removal of the C-terminal residues 199-267 likely precludes formation of such non-native states and, thereby, slows the thermodynamically favorable self-aggregation of RTA. Interestingly, even in the apparent absence of RIP activity, at least one protein with the RTA protein fold has been linked to a natural self-aggregation or agglutination function *in vivo*.[58] Thus, RTA function may depend upon formation of reversible, partly-unfolded states prone to self-aggregation, in addition to the well-characterized N-glycosidase activity and binding of RTB. In this sense, our immunogen design efforts may be viewed as “de-differentiating” RTA away from biological functions requiring a complex, two-domain fold, and toward a more stable but non-functional single-domain fold.

## Conclusion

We increased the $T_m$ and reduced self-aggregation of RTA 1-33/44-198 by engineering an SS-bond into a specific loop (residues 44-55) predicted to be involved in the early stages of thermal unfolding. Stabilization of this loop by either R48C/T77C or V49C/E99C also facilitated protein crystallization. The importance of the mobile loop 44-55 conformation for crystallization was corroborated by the observation that it is positioned to form self-self interactions between molecules in the crystal. The refined X-ray structure confirmed SS-bond formation and demonstrated that the N-terminal region of RTA can function as an independent protein folding unit (domain) with an approximately 25% reduction in volume compared with RTA. This is consistent with the original hypothesis and earlier experimental data supporting global stabilization of RTA by elimination of unnecessary interfacial regions that interact with RTB in the holotoxin. Future improvement of the RTA1-33/44-198 immunogen may be gained by further optimization of the sequence and/or length of loop 44-55 containing the engineered SS-bond.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

We are grateful to Dr. Nick Mantis for providing MAb GD12. Dr. Matthew Benning (Bruker AXS Inc.) provided extensive assistance with the *Proteum* software package. Drs. Michael Lee and I.C. Yeh contributed valuable discussions of methods for modeling protein folding-unfolding dynamics.

## Abbreviations and symbols

The abbreviations used are:

| | |
|---|---|
| **BME** | beta-mercaptoethanol |
| **CD** | circular dichroism |
| **DTNB** | 5,5′-dithiobis(2-nitrobenzoic acid) |
| **EDTA** | ethylenediaminetetraacetic acid |
| **MALDI-TOF** | matrix-assisted laser desorption/ionization time-of-flight |
| **MES** | 2-(N-morpholino)ethanesulfonic acid |
| **MAb** | monoclonal antibody |
| **MD** | molecular dynamics |
| **PAP** | pokeweed antiviral protein |
| **PBS** | phosphate buffered saline |
| **RIP** | ribosome inactivating protein |
| **rmsd** | root-mean-square deviation |
| **RTA** | ricin toxin A-chain |
| **RTA1-33/44-198** | ricin toxin A-chain containing residues 1-198 and a deletion of residues 34-43 |
| **RTB** | ricin toxin B-chain |
| **SDS-PAGE** | sodium dodecyl sulfate polyacrylamide gel electrophoresis |
| **SS-bond** | disulfide bond |

## References

1. Dormitzer PR, Ulmer JB, Rappuoli R. Structure-based antigen design: a strategy for next generation vaccines. Trends Biotechnol. 2008; 26(12):659–667. [PubMed: 18977045]
2. Serruto D, Rappuoli R. Post-genomic vaccine development. FEBS Lett. 2006; 580(12):2985–2992. [PubMed: 16716781]
3. Rinaudo CD, Telford JL, Rappuoli R, Seib KL. Vaccinology in the genome era. J Clin Invest. 2009; 119(9):2515–2525. [PubMed: 19729849]
4. Daugherty PS. Protein engineering with bacterial display. Curr Opin Struct Biol. 2007; 17(4):474–480. [PubMed: 17728126]
5. Georgiou G, Stathopoulos C, Daugherty PS, Nayak AR, Iverson BL, Curtiss R 3rd. Display of heterologous proteins on the surface of microorganisms: from the screening of combinatorial libraries to live recombinant vaccines. Nat Biotechnol. 1997; 15(1):29–34. [PubMed: 9035102]

6. Kumar S, Ochoa W, Singh P, Hsu C, Schneemann A, Manchester M, Olson M, Reddy V. Tomato bushy stunt virus (TBSV), a versatile platform for polyvalent display of antigenic epitopes and vaccine design. Virology. 2009; 388(1):185–190. [PubMed: 19344926]
7. Millard, CB.; LeClaire, RD. Ricin and related toxins. In: Romano, JR.; Salem, H., editors. Chemical Warfare Agents: Chemistry, Pharmacology, Toxicology and Therapeutics. CRC Press; Boca Raton, FL: 2007. p. 466
8. Schnell R, Borchmann P, Staak JO, Schindler J, Ghetie V, Vitetta ES, Engert A. Clinical evaluation of ricin A-chain immunotoxins in patients with Hodgkin's lymphoma. Ann Oncol. 2003; 14(5):729–736. [PubMed: 12702527]
9. Vitetta ES, Smallshaw JE, Coleman E, Jafri H, Foster C, Munford R, Schindler J. A pilot clinical trial of a recombinant ricin vaccine in normal humans. Proc Natl Acad Sci U S A. 2006; 103(7): 2268–2273. [PubMed: 16461456]
10. Kende M, Del Giudice G, Rivera N, Hewetson J. Enhancement of intranasal vaccination in mice with deglycosylated chain A ricin by LTR72, a novel mucosal adjuvant. Vaccine. 2006; 24(12): 2213–2221. [PubMed: 16325310]
11. Griffiths GD, Bailey SC, Hambrook JL, Keyte MP. Local and systemic responses against ricin toxin promoted by toxoid or peptide vaccines alone or in liposomal formulations. Vaccine. 1998; 16(5):530–535. [PubMed: 9491508]
12. Olson MA, Carra JH, Roxas-Duncan V, Wannemacher RW, Smith LA, Millard CB. Finding a new vaccine in the ricin protein fold. Protein Eng Des Sel. 2004; 17(4):391–397. [PubMed: 15187223]
13. Misna D, Monzingo AF, Katzin BJ, Ernst S, Robertus JD. Structure of recombinant ricin A chain at 2.3 A. Protein Sci. 1993; 2(3):429–435. [PubMed: 8453380]
14. McHugh CA, Tammariello RF, Millard CB, Carra JH. Improved stability of a protein vaccine through elimination of a partially unfolded state. Protein Sci. 2004; 13(10):2736–2743. [PubMed: 15340172]
15. Olson MA, Yeh IC, Lee MS. Coarse-grained lattice model simulations of sequence-structure fitness of a ribosome-inactivating protein. Biopolymers. 2008; 89(2):153–159. [PubMed: 17985366]
16. Flory PJ. Theory of elastic mechanisms of fibrous proteins. J Am Chem Soc. 1956; 78:5222–5235.
17. Matsumura M, Becktel WJ, Levitt M, Matthews BW. Stabilization of phage T4 lysozyme by engineered disulfide bonds. Proc Natl Acad Sci U S A. 1989; 86(17):6562–6566. [PubMed: 2671995]
18. Pace CN, Grimsley GR, Thomson JA, Barnett BJ. Conformational stability and activity of ribonuclease T1 with zero, one, and two intact disulfide bonds. J Biol Chem. 1988; 263(24): 11820–11825. [PubMed: 2457027]
19. Doig AJ, Williams DH. Is the hydrophobic effect stabilizing or destabilizing in proteins? The contribution of disulphide bonds to protein stability. J Mol Biol. 1991; 217(2):389–398. [PubMed: 1992169]
20. Wetzel R. Harnessing disulfide bonds using protein engineering. Trends Biochem Sci. 1987; 12:478–482.
21. Betz SF. Disulfide bonds and the stability of globular proteins. Protein Sci. 1993; 2(10):1551–1558. [PubMed: 8251931]
22. Siadat OR, Lougarre A, Lamouroux L, Ladurantie C, Fournier D. The effect of engineered disulfide bonds on the stability of Drosophila melanogaster acetylcholinesterase. BMC Biochem. 2006; 7:12. [PubMed: 16686937]
23. Clarke J, Fersht AR. Engineered disulfide bonds as probes of the folding pathway of barnase: increasing the stability of proteins against the rate of denaturation. Biochemistry. 1993; 32(16): 4322–4329. [PubMed: 8476861]
24. Feig M, Karanicolas J, Brooks CL 3rd. MMTSB Tool Set: enhanced sampling and multiscale modeling methods for applications in structural biology. J Mol Graph Model. 2004; 22(5):377–395. [PubMed: 15099834]
25. Yeh IC, Lee MS, Olson MA. Calculation of protein heat capacity from replica-exchange molecular dynamics simulations with different implicit solvent models. J Phys Chem B. 2008; 112(47): 15064–15073. [PubMed: 18959439]

26. Ellman GL. Tissue sulfhydryl groups. Arch Biochem Biophys. 1959; 82(1):70–77. [PubMed: 13650640]
27. Lemley PV, Amanatides P, Wright DC. Identification and characterization of a monoclonal antibody that neutralizes ricin toxicity in vitro and in vivo. Hybridoma. 1994; 13(5):417–421. [PubMed: 7860097]
28. Mantis NJ, McGuinness CR, Sonuyi O, Edwards G, Farrant SA. Immunoglobulin A antibodies against ricin A and B subunits protect epithelial cells from ricin intoxication. Infect Immun. 2006; 74(6):3455–3462. [PubMed: 16714576]
29. Neal LM, O'Hara J, Brey RN 3rd, Mantis NJ. A monoclonal immunoglobulin G antibody directed against an immunodominant linear epitope on the ricin A chain confers systemic and mucosal immunity to ricin. Infect Immun. 2010; 78(1):552–561. [PubMed: 19858297]
30. Castelletti D, Fracasso G, Righetti S, Tridente G, Schnell R, Engert A, Colombatti M. A dominant linear B-cell epitope of ricin A-chain is the target of a neutralizing antibody response in Hodgkin's lymphoma patients treated with an anti-CD25 immunotoxin. Clin Exp Immunol. 2004; 136(2): 365–372. [PubMed: 15086403]
31. Navaza J. Implementation of molecular replacement in AMoRe. Acta Crystallogr D Biol Crystallogr. 2001; 57(Pt 10):1367–1372. [PubMed: 11567147]
32. The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr. 1994; 50(Pt 5):760–763. [PubMed: 15299374]
33. Voss NR, Gerstein M, Steitz TA, Moore PB. The geometry of the ribosomal polypeptide exit tunnel. J Mol Biol. 2006; 360(4):893–906. [PubMed: 16784753]
34. Schmidt B, Hogg PJ. Search for allosteric disulfide bonds in NMR structures. BMC Struct Biol. 2007; 7:49. [PubMed: 17640393]
35. Hazes B, Dijkstra BW. Model building of disulfide bonds in proteins with known three-dimensional structure. Protein Eng. 1988; 2(2):119–125. [PubMed: 3244694]
36. Dombkowski AA. Disulfide by Design: a computational method for the rational design of disulfide bonds in proteins. Bioinformatics. 2003; 19(14):1852–1853. [PubMed: 14512360]
37. Wetzel R, Perry LJ, Baase WA, Becktel WJ. Disulfide bonds and thermal stability in T4 lysozyme. Proc Natl Acad Sci U S A. 1988; 85(2):401–405. [PubMed: 3277175]
38. Wells JA, Powers DB. In vivo formation and stability of engineered disulfide bonds in subtilisin. J Biol Chem. 1986; 261(14):6564–6570. [PubMed: 3516996]
39. Heinz DW, Matthews BW. Rapid crystallization of T4 lysozyme by intermolecular disulfide cross-linking. Protein Eng. 1994; 7(3):301–307. [PubMed: 8177878]
40. Castelletti D, Colombatti M. Peptide analogues of a T-cell epitope of ricin toxin A-chain prevent agonist-mediated human T-cell response. Int Immunol. 2005; 17(4):365–372. [PubMed: 15724064]
41. Matsumura M, Matthews BW. Stabilization of functional proteins by introduction of multiple disulfide bonds. Methods Enzymol. 1991; 202:336–356. [PubMed: 1784181]
42. Chrunyk BA, Evans J, Lillquist J, Young P, Wetzel R. Inclusion body formation and protein stability in sequence variants of interleukin-1 beta. J Biol Chem. 1993; 268(24):18053–18061. [PubMed: 8394358]
43. Eijsink VG, Bjork A, Gaseidnes S, Sirevag R, Synstad B, van den Burg B, Vriend G. Rational engineering of enzyme stability. J Biotechnol. 2004; 113(1-3):105–120. [PubMed: 15380651]
44. Martensson LG, Karlsson M, Carlsson U. Dramatic stabilization of the native state of human carbonic anhydrase II by an engineered disulfide bond. Biochemistry. 2002; 41(52):15867–15875. [PubMed: 12501217]
45. Ko JH, Jang WH, Kim EK, Lee HB, Park KD, Chung JH, Yoo OJ. Enhancement of thermostability and catalytic efficiency of AprP, an alkaline protease from Pseudomonas sp., by the introduction of a disulfide bond. Biochem Biophys Res Commun. 1996; 221(3):631–635. [PubMed: 8630012]
46. Hagihara Y, Mine S, Uegaki K. Stabilization of an immunoglobulin fold domain by an engineered disulfide bond at the buried hydrophobic region. J Biol Chem. 2007; 282(50):36489–36495. [PubMed: 17932041]
47. Reading NS, Aust SD. Engineering a disulfide bond in recombinant manganese peroxidase results in increased thermostability. Biotechnol Prog. 2000; 16(3):326–333. [PubMed: 10835231]

48. Chen L, Yu C, Zhou X, Zhang Y. Rational introduction of disulfide bond to enhance optimal temperature of Lipomyces starkeyi alpha-dextranase expressed in Pichia pastoris. J Microbiol Biotechnol. 2009; 19(12):1506–1513. [PubMed: 20075611]

49. Wakarchuk WW, Sung WL, Campbell RL, Cunningham A, Watson DC, Yaguchi M. Thermostabilization of the Bacillus circulans xylanase by the introduction of disulfide bonds. Protein Eng. 1994; 7(11):1379–1386. [PubMed: 7700870]

50. Thornton JM. Disulphide bridges in globular proteins. J Mol Biol. 1981; 151(2):261–287. [PubMed: 7338898]

51. Mitchinson C, Wells JA. Protein engineering of disulfide bonds in subtilisin BPN'. Biochemistry. 1989; 28(11):4807–4815. [PubMed: 2504281]

52. Schellman JA. The stability of hydrogen-bonded peptide structures in aqueous solution. C R Trav Lab Carlsberg Chim. 1955; 29(14-15):230–259. [PubMed: 13305131]

53. Katz BA, Kossiakoff A. The crystallographically determined structures of atypical strained disulfides engineered into subtilisin. J Biol Chem. 1986; 261(33):15480–15485. [PubMed: 3096989]

54. Hinck AP, Truckses DM, Markley JL. Engineered disulfide bonds in staphylococcal nuclease: effects on the stability and conformation of the folded protein. Biochemistry. 1996; 35(32):10328–10338. [PubMed: 8756688]

55. van den Burg B, Eijsink VG. Selection of mutations for increased protein stability. Curr Opin Biotechnol. 2002; 13(4):333–337. [PubMed: 12323355]

56. Argent RH, Parrott AM, Day PJ, Roberts LM, Stockley PG, Lord JM, Radford SE. Ribosome-mediated folding of partially unfolded ricin A-chain. J Biol Chem. 2000; 275(13):9263–9269. [PubMed: 10734065]

57. Argent RH, Roberts LM, Wales R, Robertus JD, Lord JM. Introduction of a disulfide bond into ricin A chain decreases the cytotoxicity of the ricin holotoxin. J Biol Chem. 1994; 269(43):26705–26710. [PubMed: 7929403]

58. Arreguin-Espinosa R, Fenton B, Vazquez-Contreras E, Arreguin B, Garcia-Hernandez E. PFA, a novel mollusk agglutinin, is structurally related to the ribosome-inactivating protein superfamily. Arch Biochem Biophys. 2001; 394(2):151–155. [PubMed: 11594727]

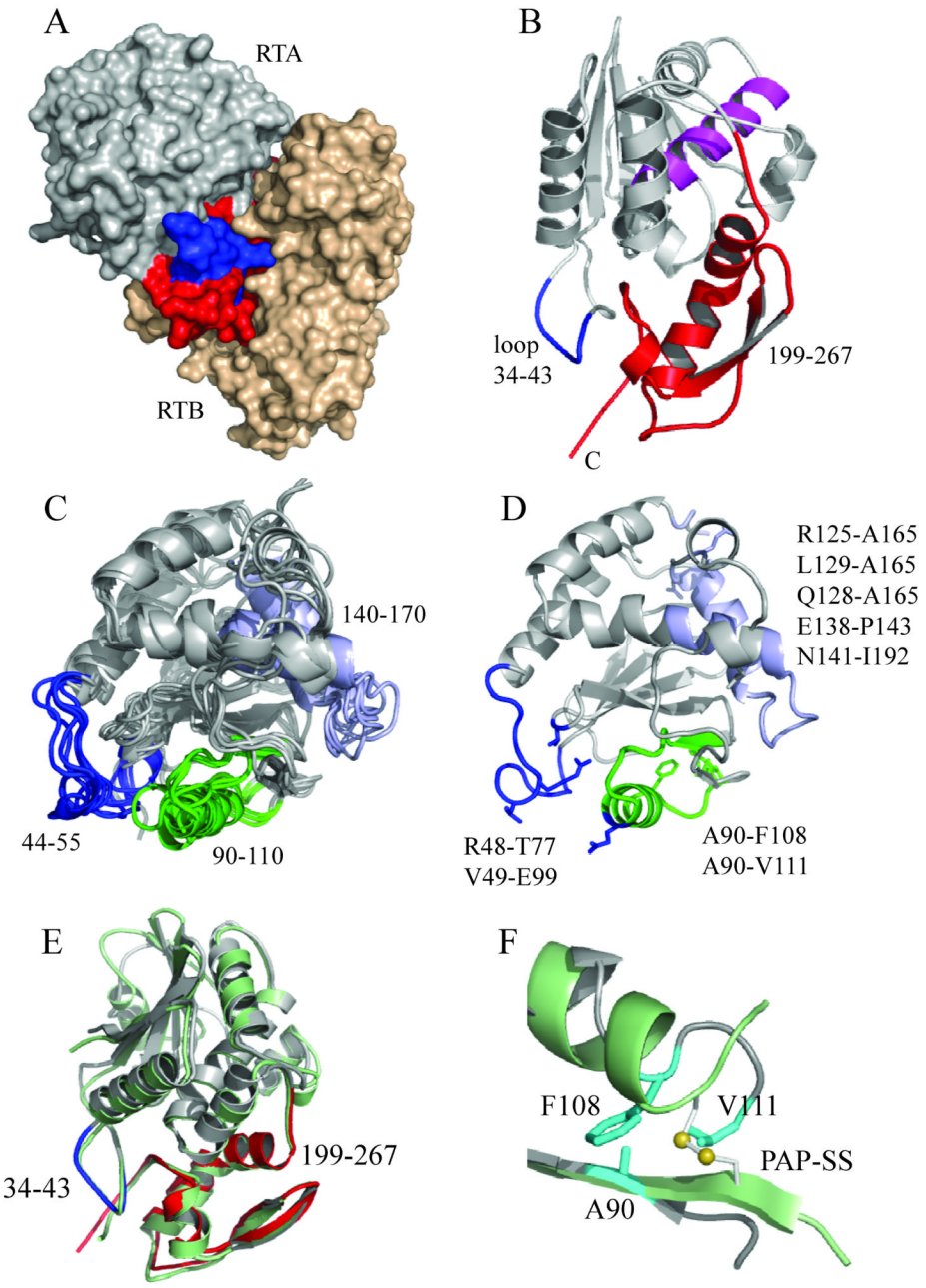


**Figure 1.**

Illustration of the protein design for the RTA immunogen and the selection of substitution sites for placement of SS-bonds. (A) RTA-RTB crystallographic structure (PDB 2AAI) highlighting removal of loop residues 34-43 (colored blue) and truncation of the C-terminal residues, 199-267 (colored red). Non-truncated RTA regions are colored grey. (B) Residues L161-I175 of a known human B-cell epitope are colored magenta on the structure of RTA. (C) Conformations of a modeled RTA1-33/44-198 protein from replica-exchange molecular-dynamics simulations culled at 298K. Structural regions that show early-stage unfolding are highlighted. (D) The nine pairwise residue sites proposed for placement of the SS-bonds. (E)

Overlay of RTA and PAP (lime green). (F) Enlargement of the loop region from the superposition showing the location of the A90-F108 and A90-V111 side chains.

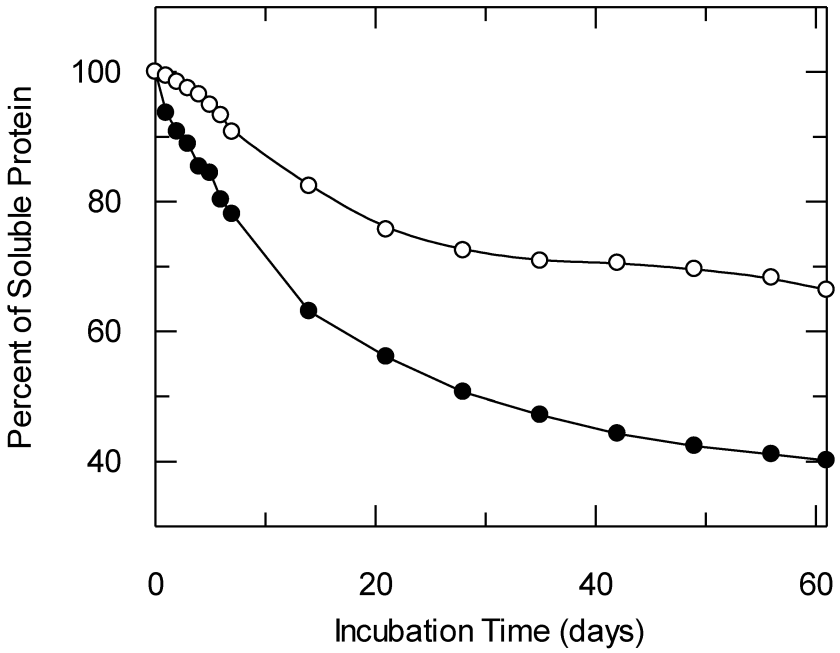


**Figure 2.**

Introduction of a disulfide bond in the R48C/T77C variant reduces the aggregation rate for RTA1-33/44-198. Solutions of each highly purified (>90% purity by PAGE) recombinant immunogen were prepared in PBS pH 7.4 and stored at a fixed starting concentration (0.2-, 1.0-, or 2.0-mg/mL) in a sealed tube at 37°C. At various time points, tubes were removed from incubation and centrifuged (14,000 rpm for 5 min) to pellet insoluble protein aggregates. The amount of protein remaining in solution was measured by absorbance at 280 nm, and expressed as a percentage of the starting protein concentration. For each time point, three separate tubes were sampled and the average percentage and standard error was tabulated. Forty percent of the parent RTA 1-33/44-198 molecule (●) remained soluble after 61 days at 37°C (1 mg/mL protein in PBS), whereas 66% of the R48C/T77C variant (○) remained soluble after 61 days under identical conditions (Table II).

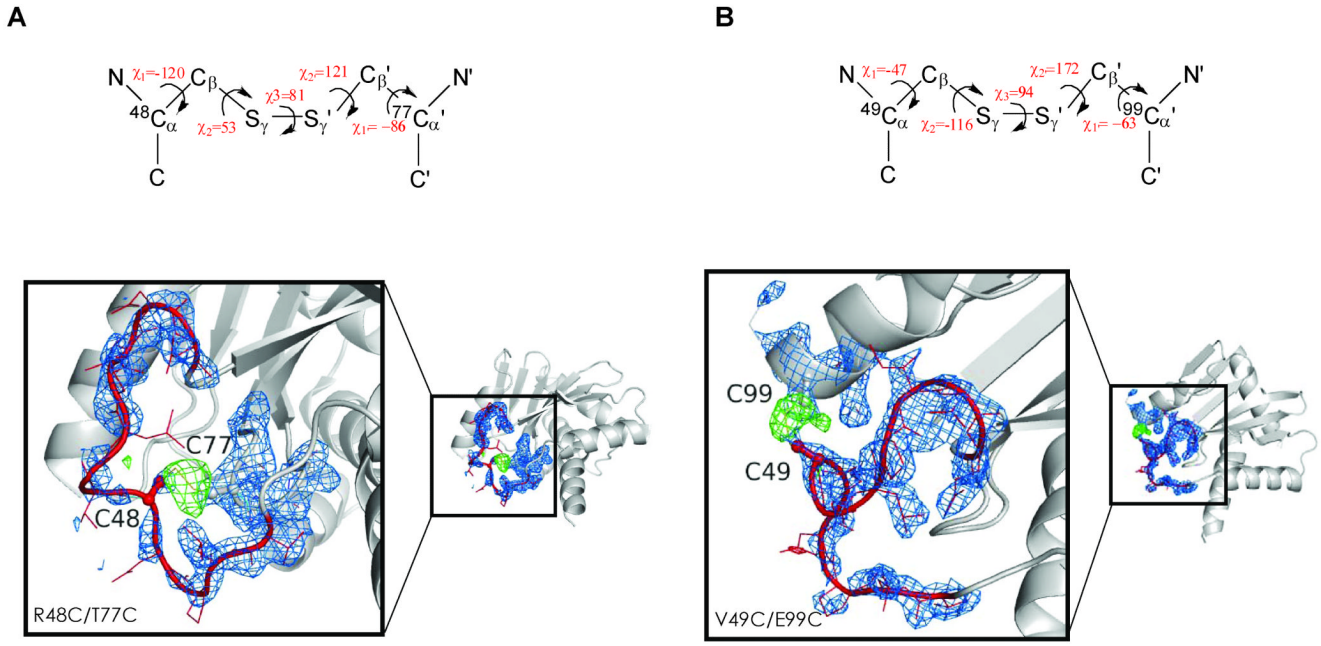


**Figure 3.**

Comparison of engineered SS-bonds for R48C/T77C (panel A) and V49C/E99C (panel B) variants of RTA1-33/44-198. Top of each panel depicts the final refined values of the χ-angles around each SS-bond (top). Bottom of each panel shows the electron density around the engineered SS-bonds with the loop 44-55 colored red in a ribbon diagram of the overall structure. Electron density difference maps ($F_o$-$F_c$) were calculated by truncating each Cys residue at Cβ (converted to Ala) to visualize residual positive difference density for modeling and refinement of the intramolecular SS-bonds. Positive difference density map is displayed as green mesh at the 3.5 sigma contour level, and negative difference density map is red at -3.5 sigma contour level; no negative density is visible at this contour level. The $2F_o$-$F_c$ map is shown in blue at the 1.0 sigma level.

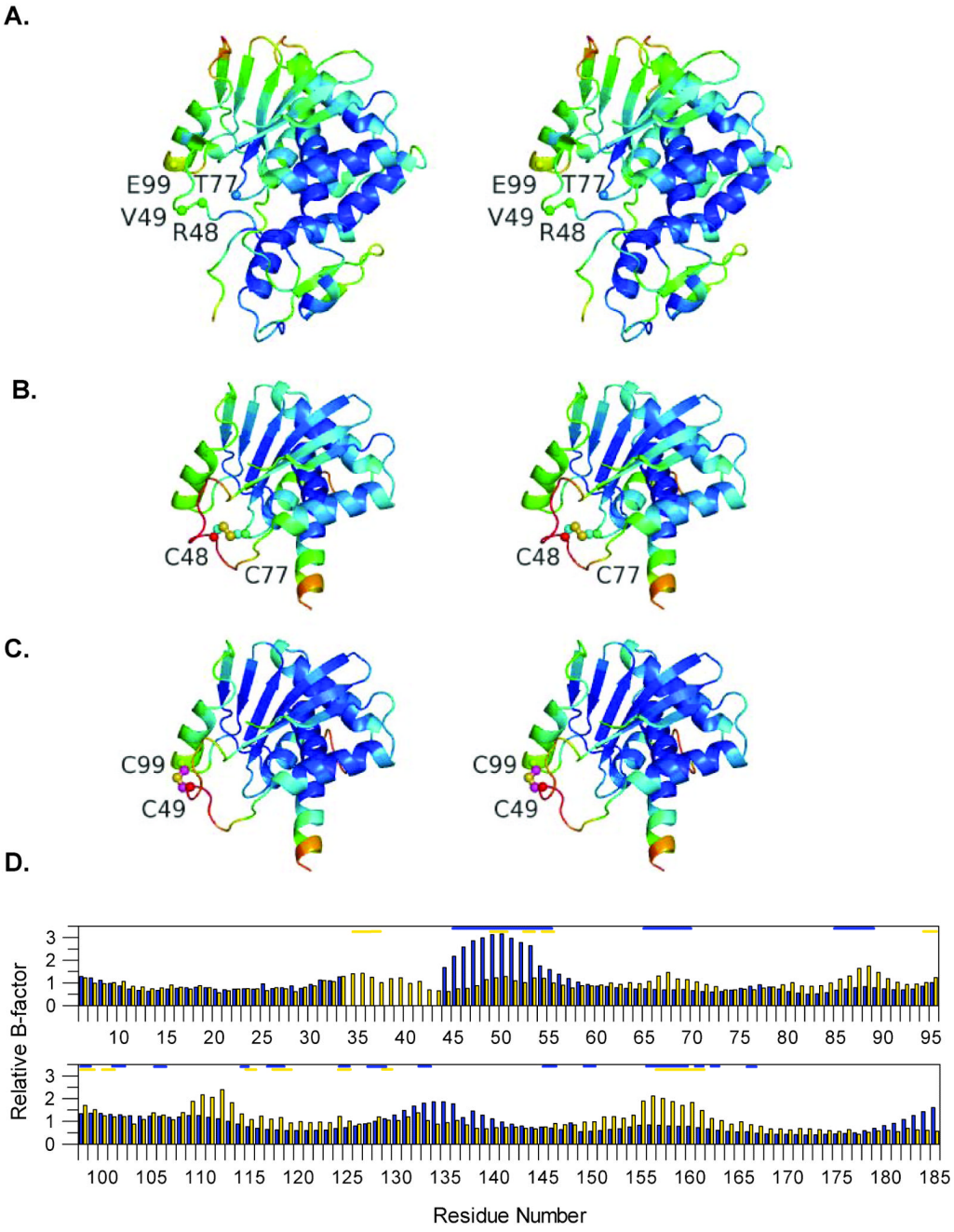


**Figure 4.**

Comparison of relative B-factors for RTA with those of the R48C/T77C and V49C/E99C variants of RTA1-33/44-198. Structures colored by B-factor for (A) RTA (PDB 1IFS); (B) R48C/T77C variant; and (C) V49C/E99C variant. Panel (D) shows a plot comparing the relative B-factors versus residue number for RTA (yellow) and the R48C/T77C variant (blue). For each protein the relative B-factors were calculated by dividing the average B-factor for the main chain of each residue by the average B-factor calculated for the protein. Crystal contacts of 5.0Å are indicated by the horizontal yellow lines for RTA and blue lines for the R48C/T77C variant.

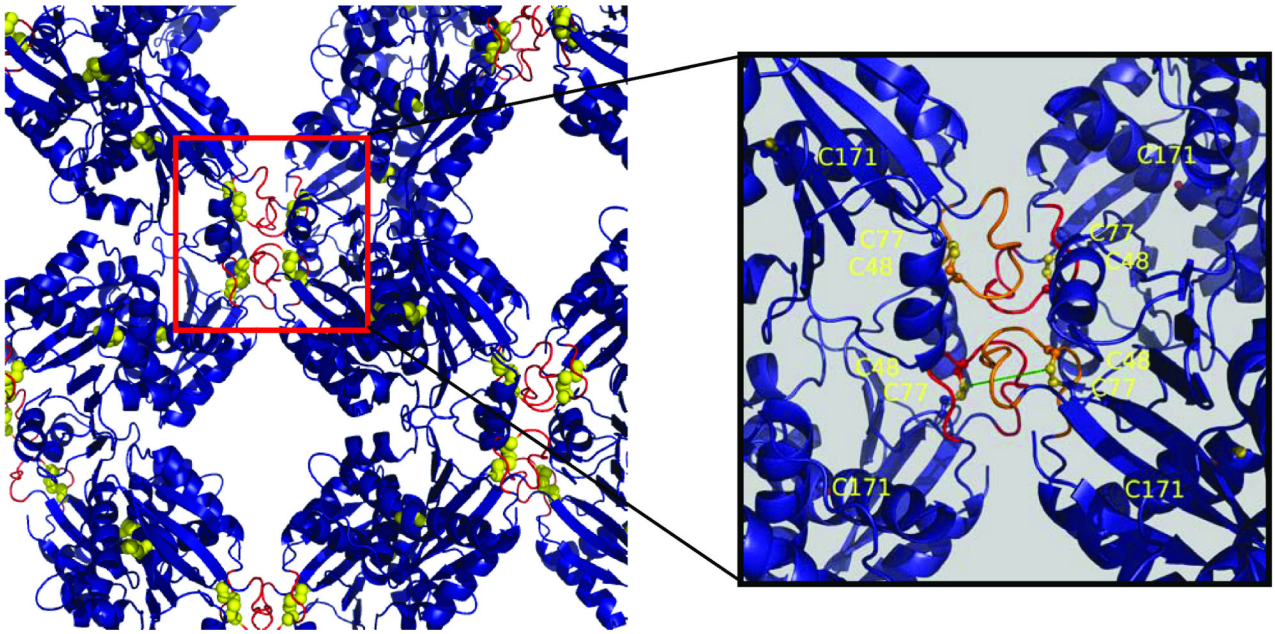


**Figure 5.**

Overview of intermolecular interactions of symmetry-related molecules within the RTA 1-33/44-198 R48C/T77C crystals. All cysteines are shown in yellow. The inset is a close-up view of loop 44-55 (shown in orange and red) which is positioned to participate in self-self interactions across an intermolecular monomer-monomer interfaces. The distance between the two closest cysteines, shown in green, is 13.7Å.

**Table I**

Attempts to introduce a stabilizing SS-bond at different locations in RTA1-33/44-198

| Ricin Vaccine Immunogen[1,2] | | Number of Free Cys[3] | Measured $T_m$[1] |
|---|---|---|---|
| RTA1-33/44-198 | | 0.97 ± 0.03 | 57.90 ± 0.03 |
| *Substitutions that Increase $T_m$* | | | |
| RTA1-33/44-198 | R48C/T77C | 1.00 ± 0.08 | 62.9 ± 0.2 |
| RTA1-33/44-198 | V49C/E99C | 0.96 ± 0.05 | 62.9 ± 0.2 |
| RTA1-33/44-198 | A90C/V111C | 1.08 ± 0.03 | 60.6 ± 0.2 |
| RTA1-33/44-198 | N141C/I192C | 1.09 ± 0.04 | 59.4 ± 0.5 |
| *Substitutions with No Significant Effect on $T_m$* | | | |
| RTA1-33/44-198 | R125C/A165C | 2.47 ± 0.01 | 58.20 ± 0.01 |
| *Substitutions that Decrease $T_m$* | | | |
| RTA1-33/44-198 | E138C/P143C | 1.45 ± 0.07 | 55.9 ± 0.4 |
| RTA1-33/44-198 | Q128C/A165C | 1.88 ± 0.04 | 51.5 ± 0.8 |

[1] Presence of desired substitution and no other was confirmed for each protein by re-sequencing of each entire DNA construct.

[2] Purity of each final product was estimated from Coomassie blue-stained SDS-PAGE to be >95%. The identity of purified protein was confirmed by reaction with MAb on Western blots.

[3] The number of free Cys residues was quantified for purified proteins by titrating free –SH using DTNB. Titrations were conducted in triplicate at five different protein concentrations between 40-50 μM; the mean + SE was determined from linear regression of the titration plots. Note that the parent protein molecule (RTA1-33/44-198) contains one Cys, whereas each mutant contains a total of three Cys residues. The DTNB assay was optimized to probe the reduced Cys residues such that values significantly greater than one reflect incomplete formation of a stable SS-bond between the engineered Cys residues.

## Table II

Disulfide bond R48C-T77C reduces aggregation kinetic of RTA1-33/44-198

| Initial Protein Concentration | Days at 37°C[1] | % Protein Remaining in Solution (mean ± standard error) | |
|---|---|---|---|
| | | RTA1-33/44-198 | *RTA1-33/44-198 R48C/T77C* |
| 0.2 mg/ml | 10 | 51 ± 3 | 87 ± 5 |
| | 28 | | 73 ± 5 |
| | 49 | | 59 ± 5 |
| 1.0 mg/ml | 7 | 78 ± 1 | 91 ± 2 |
| | 28 | 51 ± 1 | 72 ± 1 |
| | 61 | 40 ± 1 | 66 ± 1 |
| 2.0 mg/ml | 7 | 71.6 ± 0.3 | 94.5 ± 0.3 |
| | 28 | 45.1 ± 0.6 | 65.7 ± 1.0 |

[1] Experiments were conducted in PBS, pH 7.4

## Table III

X-ray crystallography data collection and refinement summary statistics

| | RTA1-33/44-198 R48C/T77C | RTA1-33/44-198 V49C/E99C |
|---|---|---|
| Space group | I222 | I222 |
| Unit Cell Dimensions (Å) | a=51.5, b=72.3, c=94.2 | a= 51.7, b=72.8, c=96.4 |
| Wavelength (Å) | 1.54 | 1.54 |
| Resolution Range (Å) [a] | 57.4-2.3 (2.38-2.28) | 58.1-2.1 (2.17-2.08) |
| Unique Reflections | 8,333 (979) | 10,903 (1,009) |
| $R_{sym}$ [b] | 0.098 (0.385) | 0.054 (0.326) |
| I/σI | 19.05 (5.52) | 20.18 (3.88) |
| Completeness | 100.0 (100.0) | 96.6 (77.1) |
| Redundancy | 16.0 (11.9) | 9.6 (3.0) |
| Refinement Statistics: | | |
| Resolution (Å) | 2.3 | 2.1 |
| No. of reflections | 7,938 | 10,377 |
| $R_{factor}$ [c] | 0.214 | 0.211 |
| $R_{free}$ [d] | 0.245(5%) [d] | 0.236(5%) [d] |
| Number of Atoms: | | |
| Protein | 1362 | 1362 |
| Solvent | 85 | 76 |
| Other (sulfate ion) | 5 | 5 |
| Average B-factors (Å$^2$) | | |
| Protein | 18.4 | 20.8 |
| Solvent | 25.7 | 28.8 |
| R.m.s.d. from ideal geometry: | | |
| Bond lengths (Å) | 0.010 | 0.006 |
| Bond angles (degrees) | 1.12 | 1.11 |
| Ramachandran plot | | |
| Most favored regions (%) | 90.1% | 90.1% |
| Additional allowed regions (%) | 9.9% | 9.2% |
| Generously allowed regions (%) | 0.0% | 0.7% |
| Disallowed regions (%) | 0.0% | 0.0 |

[a] Values in parentheses are for the highest resolution shell

[b] $R_{sym}$ was calculated from $R_{sym} = \sum |I_i - \langle I \rangle| / \sum I_i$

[c] $R_{factor}$ for working set of reflections was calculated using: $R_{factor} = \sum ||F_o| - |F_c|| / \sum |F_o|$

[d] $R_{free}$ for test set and size of test set as % total reflections in parentheses