9256 *Biochemistry* **2010,** *49,* 9256–9268
DOI: 10.1021/bi1008968

BIOCHEMISTRY
including biophysical chemistry & molecular biology
*Article*

# Assessing Energetic Contributions to Binding from a Disordered Region in a Protein−Protein Interaction[†,‡]

Sangwoo Cho,[§] Chittoor P. Swaminathan,[§] Daniel A. Bonsor,[∥] Melissa C. Kerzic,[§] Rongjin Guan,[§,△] Jianying Yang,[§,◇] Michele C. Kieke,[⊥,▽] Peter S. Andersen,[#] David M. Kranz,[⊥] Roy A. Mariuzza,[§] and Eric J. Sundberg*[,∥]

[§]*W. M. Keck Laboratory for Structural Biology, University of Maryland Institute for Bioscience and Biotechnology Research, Rockville, Maryland 20850,* [∥]*Boston Biomedical Research Institute, Watertown, Massachusetts 02472,* [⊥]*Department of Biochemistry, University of Illinois, Urbana, Illinois 61801, and* [#]*Symphogen A/S, DK-2800 Lyngby, Denmark.* [△]*Present address: Center for Advanced Biotechnology and Medicine, Rutgers University, Piscataway, NJ 08854.* [◇]*Present address: Max-Planck-Institut für Immunobiologie, D79108 Freiburg, Germany.* [▽]*Present address: Department of Biology, Concordia University, St. Paul, MN 55104.*

*Received June 3, 2010; Revised Manuscript Received September 13, 2010*

ABSTRACT: Many functional proteins are at least partially disordered prior to binding. Although the structural transitions upon binding of disordered protein regions can influence the affinity and specificity of protein complexes, their precise energetic contributions to binding are unknown. Here, we use a model protein−protein interaction system in which a locally disordered region has been modified by directed evolution to quantitatively assess the thermodynamic and structural contributions to binding of disorder-to-order transitions. Through X-ray structure determination of the protein binding partners before and after complex formation and isothermal titration calorimetry of the interactions, we observe a correlation between protein ordering and binding affinity for complexes along this affinity maturation pathway. Additionally, we show that discrepancies between observed and calculated heat capacities based on buried surface area changes in the protein complexes can be explained largely by heat capacity changes that would result solely from folding the locally disordered region. Previously developed algorithms for predicting binding energies of protein−protein interactions, however, are unable to correctly model the energetic contributions of the structural transitions in our model system. While this highlights the shortcomings of current computational methods in modeling conformational flexibility, it suggests that the experimental methods used here could provide training sets of molecular interactions for improving these algorithms and further rationalizing molecular recognition in protein−protein interactions.



Although proteins have numerous capabilities, one of their most important functions is to bind other proteins. Interactions between proteins are essential for nearly all cellular processes (*1*−*3*), and aberrant protein−protein interactions contribute to the pathogenesis of numerous human diseases (*4*). Because of the importance of protein−protein interactions in nearly all aspects of biology, efforts to decipher the rules that govern these associations have been underway for many decades. Genome-wide mapping of protein−protein interactions has identified many of the molecular components of physiological and pathological processes (*5*−*9*), and current structural genomics efforts are aimed at expanding the structural database of the constituent protein domains involved in these interactions (*10*). Thus, the ability to predict the binding specificities and energies of protein complexes from protein structures alone has reached paramount importance as it represents a way in which to translate the vast and growing interactome and protein structure databases into novel insights to biological function.

In order to better quantify the various effects that contribute to protein molecular recognition, it is often necessary to create model protein−protein interaction systems that can be perturbed in a controllable manner to alter one factor that affects binding in isolation and subsequently assess the model for structural and energetic changes resulting from that perturbation. Perhaps the most common method of perturbation and assessment is to perform alanine-scanning mutagenesis in order to measure the energetic contribution of individual amino acid residues within a protein−protein interface (*11*, *12*). This technique has been used to map the functional epitopes of numerous protein−protein interfaces. Following a similar mutagenesis approach, quantitative estimations of biophysical parameters affecting protein−protein interactions such as the hydrophobic effect can be made by mutating a single large hydrophobic residue within an interface to various residues with smaller and less hydrophobic side chains. The thermodynamic and structural changes associated with these mutations can then be measured by isothermal titration calorimetry and X-ray crystallography, respectively, to yield a measure of the binding free energy change per unit buried apolar surface area (*13*, *14*).

Although such studies provide powerful means by which to improve predictive algorithms for protein−protein interactions, many properties of proteins that affect binding are not restricted to the effect of a single amino acid residue but instead are dependent on the coordinated behavior of numerous residues within an interface. One such complex property of protein−protein

[†]This work was supported in part by National Institutes of Health Grants AI55882 and AI65690 (to E.J.S.), AI036900 (to R.A.M.), and AI06461 (to D.M.K.).

[‡]The coordinate and structure factor files for the structures have been deposited in the Protein Data Bank with the following accession codes: SEC3-1D8, 3BVG; SEC3-1D3, 3BVM; SEC3-1A4, 3BVZ; SEC3-1D8/mV$\beta$8.2, 3BYT; SEC3-1D3/mV$\beta$8.2, 3BYY; SEC3-1A4/mV$\beta$8.2, 3BZD.

*To whom correspondence should be addressed: 617-658-7882 (phone); 617-972-1761 (fax); sundberg@bbri.org (e-mail).

 Published on Web 09/13/2010 

interactions is energetic cooperativity between amino acid residues by which the summation of binding free energies of several interface residues individually is not equivalent to the binding free energy resulting from the entire set of those residues together within a single protein−protein interface. This type of behavior in protein complexes implies that there exists some form of networked communication between interface residues and has led to the proposal that protein binding sites possess a modular architecture that is a significant energetic driver for interaction (*15*, *16*).

Combinatorial effects in protein−protein interactions, such as energetic cooperativity, cannot be assessed by mutating a single amino acid residue and measuring how binding of the resulting variant differs from that of the wild-type protein. Instead, alternative methods that can address the coordinated nature of protein−protein interactions have been used, including computational biology approaches (*17*−*19*) and multiple mutant analysis (*15*). An additional approach is to create unique model systems of protein−protein interactions by subjecting one of the proteins in a complex to directed evolution (*20*), such as by phage or yeast display (*21*, *22*). This iterative process of mutation and selection for tighter binding to the unmodified target protein describes an affinity maturation pathway of protein variants that, in total, can span many orders of magnitude in affinity. Because numerous mutations are made that together increase the affinity, the dissection of these affinity maturation pathways by measuring the structural and energetic changes associated with different combinations of mutations makes this is an especially powerful method for investigating biophysical parameters that are combinatorial by definition, including, but not limited to, energetic cooperativity (*23*−*25*).

Despite these and other efforts, an understanding of protein−protein interactions at the level of accurately predicting their specificities and affinities remains elusive. One notable behavior of proteins that can have an effect on molecular recognition, and is especially difficult to assess quantitatively, is their inherent flexibility (*26*), although some recent progress has been made in incorporating conformational change into protein−protein docking algorithms (*27*−*31*). There exists a broad continuum of protein flexibility that can influence interaction energetics, from side chain sampling of rotamer conformations to structural changes within protein regions that are disordered prior to binding. Previous studies using single-site alanine-scanning mutagenesis have shown that disordered residues can be critical for protein binding (*32*, *33*). However, the contributions of disordered regions to protein−protein interactions cannot be dependent upon single amino acid residues acting in isolation but undoubtedly always involve the concerted structural variations, and consequent energetic effects, of combinations of residues. Thus, a comprehensive understanding of the role that disordered regions play in protein−protein associations requires experimental strategies that can address the combinatorial nature of protein molecular recognition, such as those that have been applied to the study of energetic cooperativity.

To assess the role of a disordered protein region in complex formation, we have utilized a model protein−protein interaction system involving the complex between the *Staphylococcus aureus* enterotoxin C3 (SEC3)[1] and its cellular receptor, a variable region of the murine T cell receptor (TCR) $\beta$ chain 8.2 (mV$\beta$8.2). A region of local disorder within SEC3 was previously altered by directed evolution (*34*), the resulting affinity-matured variants differing only in a linear sequence of five residues within a disulfide loop that is disordered in the wild-type protein in the unbound state. This modified set of protein−protein interactions provides the molecular tools by which structural changes can be correlated with thermodynamic changes in order to assess the energetic contribution to binding by a disordered region in a protein−protein interaction.

## EXPERIMENTAL PROCEDURES

*Protein Expression and Purification*. Murine TCR V$\beta$ 8.2 (mV$\beta$8.2) was produced by *in vitro* refolding from *Escherichia coli* inclusion bodies and purified following previously established protocols (*25*). Wild-type SEC3 (SEC3-WT) and three variants thereof (SEC3-1D8, SEC3-1D3, and SEC3-1A4) were prepared from bacterial periplasm and purified as described previously (*23*). SEC3/mV$\beta$8.2 complexes were formed by mixing SEC3 and mV$\beta$8.2 in a 1:1.2 molar ratio and incubated for 24 h at 4 °C. Each of these protein mixtures was further purified by size exclusion chromatography using a Superdex 75 HR 10/30 column (Amersham Biosciences) equilibrated with 50 mM Tris-HCl, pH 8.5, and the purified complex was concentrated to 8 mg/mL by centrifugation using a Centricon device (Amicon).

*Crystallization and Data Collection*. For the crystallization of uncomplexed SEC3 variants, purified protein was concentrated to 5−8 mg/mL, and crystals were grown at room temperature by hanging drop vapor diffusion using a mixture of 1 $\mu$L of concentrated protein solution and an equal volume of reservoir buffer containing 0.1 M Tris-HCl, pH 8.5, 2.4 M ammonium sulfate, 2.25% polyethylene glycol (PEG) 400, and 2.25% Tween 20. Cocrystals of SEC3-1D8/mV$\beta$8.2 were grown at room temperature by hanging drop vapor diffusion using a mixture of 1 $\mu$L of concentrated protein solution (8 mg/mL) and an equal volume of reservoir buffer containing 20% PEG 3350, 0.2 M triammonium citrate, pH 7.0, and 0.3% dioxane. Showers of thin, needle-like crystals were formed within 1 week. These crystals gradually dissolved, and prismatic crystals as large as $0.05 \times 0.1 \times 0.3$ mm were obtained after approximately 1 month. Cocrystals of SEC3-1D3/mV$\beta$8.2 and SEC3-1A4/mV$\beta$8.2 were grown at room temperature by hanging drop vapor diffusion by mixing 1 $\mu$L of concentrated protein solution and an equal volume of reservoir buffer containing 2.0 M ammonium sulfate, 0.1 M Tris-HCl, pH 7.0, and 0.3% 1,6-diaminohexane. Hexagonal crystals as large as $0.05 \times 0.05 \times 0.1$ mm were obtained after 2 months. For the uncomplexed SEC3 variants, diffraction data were collected from frozen crystals at 100 K using an R-axis IV++ image-plate system equipped with Osmic mirror and a Rigaku rotating anode Cu K$\alpha$ X-ray generator. The same reservoir buffer including 20% (v/v) of glycerol was used as cryoprotectant. The collected data were processed d*TREK incorporated in CrystalClear v1.35 (Molecular Structure Corp.). Diffraction data of the SEC3/mV$\beta$8.2 complexes were collected from frozen crystals in liquid nitrogen on beamline X12B of Brookhaven National Synchrotron Laboratory with a Quantum-4 CCD detector. The data were processed and scaled with DENZO/SCALEPACK (*35*).

*Structure Determination and Refinement*. The apo SEC3 and SEC3/mV$\beta$8.2 complex structures were solved by molecular replacement with the program Molrep in CCP4 (*36*) using components or the entirety of the SEC3-WT/mV$\beta$8.2 complex crystal structure (PDB accession code 2AQ3) (*23*) as search probes.

[1]Abbreviations: ASA, accessible surface area; CDR, complementarity determining region; $C_P$, heat capacity; HV4, hypervariable region 4; ITC, isothermal titration calorimetry; MHC, major histocompatibility complex; mV$\beta$8.2, murine T cell receptor $\beta$ chain 8.2; SAG, superantigen; SEC3, staphylococcal enterotoxin C3.

Refinement was initially performed with CNS (*37*) with rigid-body refinement followed by positional, simulated annealing and individual *B* factor refinement. Manual model rebuilding was carried out between each set of refinement steps using $\sigma_A$-weighted $2F_o - F_c$ maps in the program xfit of the XtalView package (*38*). After CNS refinement converged, further refinement was carried out with the program Refmac5 in CCP4. During final refinement, solvent water molecules based on higher than $2\sigma$ peaks in the $\sigma_A$-weighted $F_o - F_c$ maps were added gradually and conservatively with regard for their environment including potential interactions with hydrogen bond partners. The solvent model was refined further by omitting all water molecules that exhibited high *B* values (>60 Å$^2$) or poor hydrogen-bonding distances or geometries.

*Isothermal Titration Calorimetry*. ITC measurements were performed by titrating mV$\beta$8.2 into solutions of SEC3 using a MicroCal VP-ITC titration microcalorimeter (MicroCal). Protein solutions were extensively dialyzed in phosphate-buffered saline, pH 7.2, filtered, and degassed prior to measurement. In a typical experiment, 2−5 $\mu$L aliquots of 0.8−2.4 mM mV$\beta$8.2 were injected from a 250 $\mu$L rotating syringe at 310 rpm into the sample cell containing 1.37 mL of 0.011−0.14 mM SEC3. For each titration experiment, an identical buffer dilution correction was conducted, and these heats of dilution were subtracted from the corresponding binding experiment. Titration data were analyzed using a single-site fitting model, and a nonlinear least-squares fitting method was used to determine the change in enthalpy, $\Delta H^\circ_b$, the equilibrium association constant, $K_b$, and the molar stoichiometry, *n*. Data acquisition and analysis were performed using the software package Origin provided by the manufacturer.

## RESULTS

*Constructing a Model Protein−Protein Interaction System To Evaluate a Disordered Region*. The binding of the bacterial superantigen (SAG) staphylococcal enterotoxin C3 (SEC3) to the variable domain of the murine T cell receptor (TCR) $\beta$ chain 8.2 (mV$\beta$8.2) has been characterized extensively (*23*, *25*, *39*−*41*). SAGs constitute a large family of protein toxins secreted from *S. aureus* and *Streptococcus pyogenes* bacteria that bind simultaneously to TCR and major histocompatibility complex (MHC) class II molecules that can activate up to 20% of all T cells (*42*−*44*). This initiates a systemic release of inflammatory cytokines that can lead to a condition known as toxic shock syndrome, which is characterized by high fever, erythematous rash, and hypotension, and can result in multiorgan failure and death.

X-ray crystal structures of wild-type SEC3 (SEC3-WT) alone (*45*) and in complex with mV$\beta$8.2 (*23*, *40*) or with the MHC class II molecule HLA-DR1 (*46*) have been determined previously. Like many other bacterial SAGs, SEC3 contains a single disulfide loop of 16 residues, which is disordered in each of these crystal structures, except for the higher resolution of the two available SEC3-WT/mV$\beta$8.2 structures that we had most recently determined (*23*). In this structure, the disulfide loop adopts diverse, but predominantly ordered, conformations within the four complexes that comprise the asymmetric unit of the crystal, one of which makes an intermolecular contact with mV$\beta$8.2, as will be discussed below.

The disulfide loop of SEC3-WT had been subjected previously to phage-display directed evolution in order to produce variants that bound with higher affinity to mV$\beta$8.2 (*34*). After randomization of five residues within the disulfide loop and selection for improved mV$\beta$8.2 binders, SEC3 variants that exhibited significantly higher affinity than the wild-type interaction were produced.

Table 1: Data Collection and Refinement Statistics for Unbound SEC3 Variants[a]

| | SEC3-1D8 | SEC3-1D3 | SEC3-1A4 |
|---|---|---|---|
| data collection | | | |
| space group | $P4_32_12$ | $P4_32_12$ | $P4_32_12$ |
| unit cell dimensions | | | |
| $a = b$ (Å) | 42.70 | 42.62 | 42.60 |
| $c$ (Å) | 287.14 | 286.52 | 287.42 |
| $\alpha = \beta = \gamma$ (deg) | 90.00 | 90.00 | 90.00 |
| molecules per asymmetric unit | 1 | 1 | 1 |
| resolution limit (Å) | 2.0 | 2.0 | 2.3 |
| observations | 87452 | 94326 | 75647 |
| unique reflections | 16498 | 17009 | 11546 |
| completeness (%) | 85.8 (82.16) | 89.0 (81.24) | 90.2 (70.5) |
| $R_{merge}$ (%)[b] | 6.3 (35.7) | 7.3 (37.4) | 5.6 (35.6) |
| refinement | | | |
| resolution (Å) | 30−2.0 | 30.0−2.0 | 30.0−2.3 |
| $R_{cryst}$ (%)[c] | 21.2 (30.9) | 22.5 (27.6) | 19.7 (23.1) |
| $R_{free}$ (%)[d] | 27.6 (41.5) | 27.4 (36.4) | 25.8 (34.6) |
| protein residues | 228 | 226 | 228 |
| Zn ions | 1 | 1 | 1 |
| water molecules | 176 | 227 | 140 |
| average *B* factors (Å$^2$) | | | |
| SEC3 | 16.7 | 22.8 | 21.8 |
| water | 25.5 | 33.3 | 27.7 |
| rms deviations | | | |
| bonds (Å) | 0.020 | 0.022 | 0.023 |
| angles (deg) | 1.755 | 1.912 | 1.963 |
| Ramachandran plot statistics | | | |
| core (%) | 88.0 | 87.8 | 87.5 |
| allowed (%) | 11.5 | 11.2 | 10.6 |
| generous (%) | 0.5 | 1.0 | 1.9 |
| disallowed (%) | 0 | 0 | 0 |

[a]Values in parentheses correspond to the highest resolution shell: SEC3-1D8 (2.05−2.00 Å); SEC3-1D3 (2.05−2.00 Å); SEC3-1A4 (2.46−2.40 Å). [b]$R_{merge}(I) = \sum|I(i) - \langle I(h)\rangle|/\sum I(i)$, where $I(i)$ is the *i*th observation of the intensity of the *hkl* reflection and $\langle I\rangle$ is the mean intensity from multiple measurements of the *h*,*k*,*l* reflection. [c]$R_{cryst}(F) = \sum_h||F_o(h)| - |F_c(h)||/\sum_h|F_o(h)|$, where $|F_o(h)|$ and $|F_c(h)|$ are the observed and calculated structure factor amplitudes for the *h*,*k*,*l* reflection. [d]$R_{free}$ is calculated over reflections in a test set not included in atomic refinement: SEC3-1D8, 844 reflections, 5.1%; SEC3-1D3, 857 reflections, 5.0%; SEC3-1A4, 505 reflections, 4.7%.

Hypothesizing that this sequence variation resulting from the directed evolution process affected the dynamic properties of the SEC3 disulfide loop with a concomitant effect on mV$\beta$8.2 binding, we selected three such variants for comprehensive structural and thermodynamic analysis alone and in complex with mV$\beta$8.2 to compare to SEC3-WT.

These variants, named SEC3-1D8 ($K_D$ = 4 $\mu$M), SEC3-1D3 ($K_D$ = 570 nM), and SEC3-1A4 ($K_D$ = 100 nM), represent roughly equally spaced points along the affinity maturation pathway at which the strength of binding is increased stepwise 3−6-fold beyond that of SEC3-WT ($K_D$ = 12 $\mu$M) or the next highest affinity SEC3 variant, for a total affinity increase of 120-fold. Notably, the two highest affinity variants, SEC3-1D3 and SEC3-1A4, each have two fewer residues in their disulfide loops than SEC3-WT and SEC3-1D8, as shorter loops were preferentially selected in latter rounds of the directed evolution process. Our goal in this study was not to describe the mechanism of binding in the SEC3/mV$\beta$8.2 system *per se*, as this has been addressed adequately in previous studies, but instead to utilize a discrete affinity-matured set of variant SEC3/mV$\beta$8.2 complexes as a model for better assessing the energetic contributions of disordered regions in protein−protein interactions.

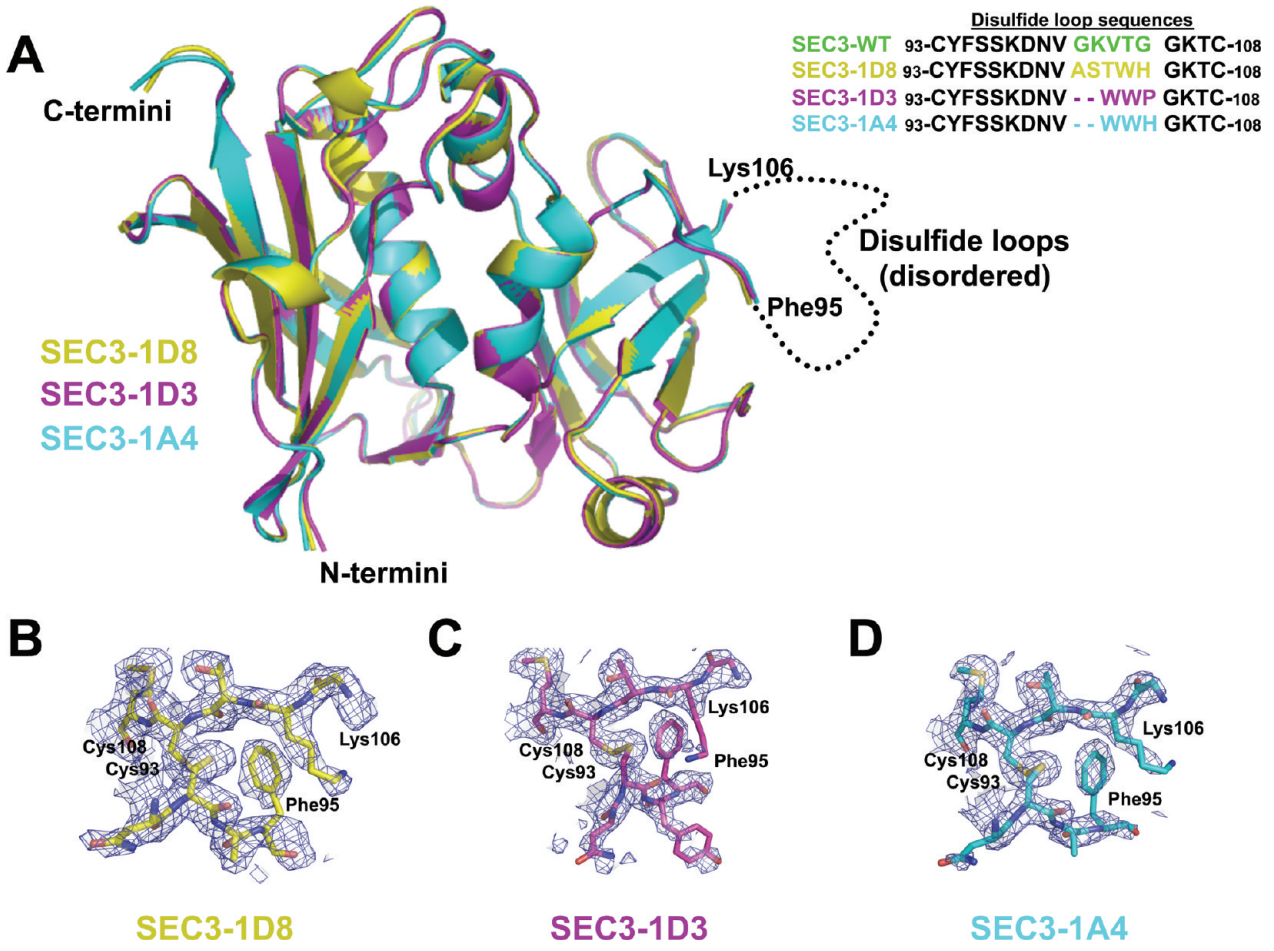


FIGURE 1: Crystallographic analysis of variant SEC3 molecules in the unbound state. (A) Superposition of the X-ray crystal structures of SEC3-1D8 (yellow), SEC3-1D3 (magenta), and SEC3-1A4 (cyan). The disulfide loop sequences altered by phage display are also shown. $\sigma_A$-weighted $2F_o - F_c$ electron density maps contoured at $1.2\sigma$ of the disulfide loop regions of (B) SEC3-1D8, (C) SEC3-1D3, and (D) SEC3-1A4. Residues labeled include Cys93 and Cys108, which form the disulfide bridge, as well as Phe95 and Lys106, which represent the furthest residues on the N- and C-termini of the disulfide loop, respectively, that can be built into the final electron maps.

*Increased Binding in the Model System Is Not Due to Increasingly Preorganized Unbound States*. In order to rule out the possibility that the increased affinity of the SEC3 variants was due to structural changes made prior to binding mV$\beta$8.2, we determined X-ray crystal structures of SEC3-1D8, SEC3-1D3, and SEC3-1A4 each in the unbound state. Each of the SEC3 variants crystallized isomorphously to one another in a tetragonal space group with a single molecule per asymmetric unit and diffracted to resolutions between 2.0 and 2.4 Å (Table 1). All of these structures are nearly perfectly superimposable and are well ordered at every position except for most residues within the disulfide loop (Figure 1A).

Electron density in each of the unbound variant SEC3 structures is interpretable only to the extent of two residues beyond the disulfide bond, such that only the loop residues Cys93-Tyr94-Phe95 and Lys106-Thr107-Cys108 have definable positions (Figure 1B–D). The remainder of the residues within the disulfide loop (i.e., residues 96–105) cannot be modeled into electron density, including those spanning positions 100–104 that were altered by phage display. Within the limits of our crystallographic assessment of protein conformation, directed evolution of this region of SEC3 does not change the degree of local disorder in the protein in the unbound state, and therefore, we have no evidence by which to ascribe increases in affinity to a preorganization of the binding site. In the absence of nuclear magnetic resonance or similar analysis, though, the possibility that the disulfide loops of the higher affinity variants are rigidified, but still not rigid enough to form a well-defined conformation in the crystal, cannot be ruled out entirely.

*Affinity Maturation in the Model System Imparts a Distinct Thermodynamic Signature*. We measured the thermodynamic parameters of the wild-type and variant SEC3 interactions with mV$\beta$8.2 by ITC at 25 °C (Figure 2A–D, Table 2). Relative to the wild-type interaction, each of the variant complexes exhibits both enthalpic gains and entropic penalties. In each case, the former is greater than the latter, resulting in a reduction of the free energy term and higher affinity, a thermodynamic indication of the ordering of a previously less ordered region during a binding reaction. The enthalpic gains and entropic penalties correlate with the affinities of the complexes. Additionally, the affinities determined by ITC in this study are nearly identical to those determined previously by surface plasmon resonance analysis, despite the different analytical methods employed and the use of TCR V$\beta$ single domain proteins in this study versus the use of full-length $\alpha\beta$TCR in the previous study (*34*).

We also measured enthalpic changes at numerous temperatures from 2 to 42 °C and plotted these values versus temperature to determine the observed heat capacity changes ($\Delta C_{P\,OBS}$) associated with each complex (Figure 2E). A relative increase in $\Delta C_{P\,OBS}$ for the two highest affinity SEC3 variant complexes

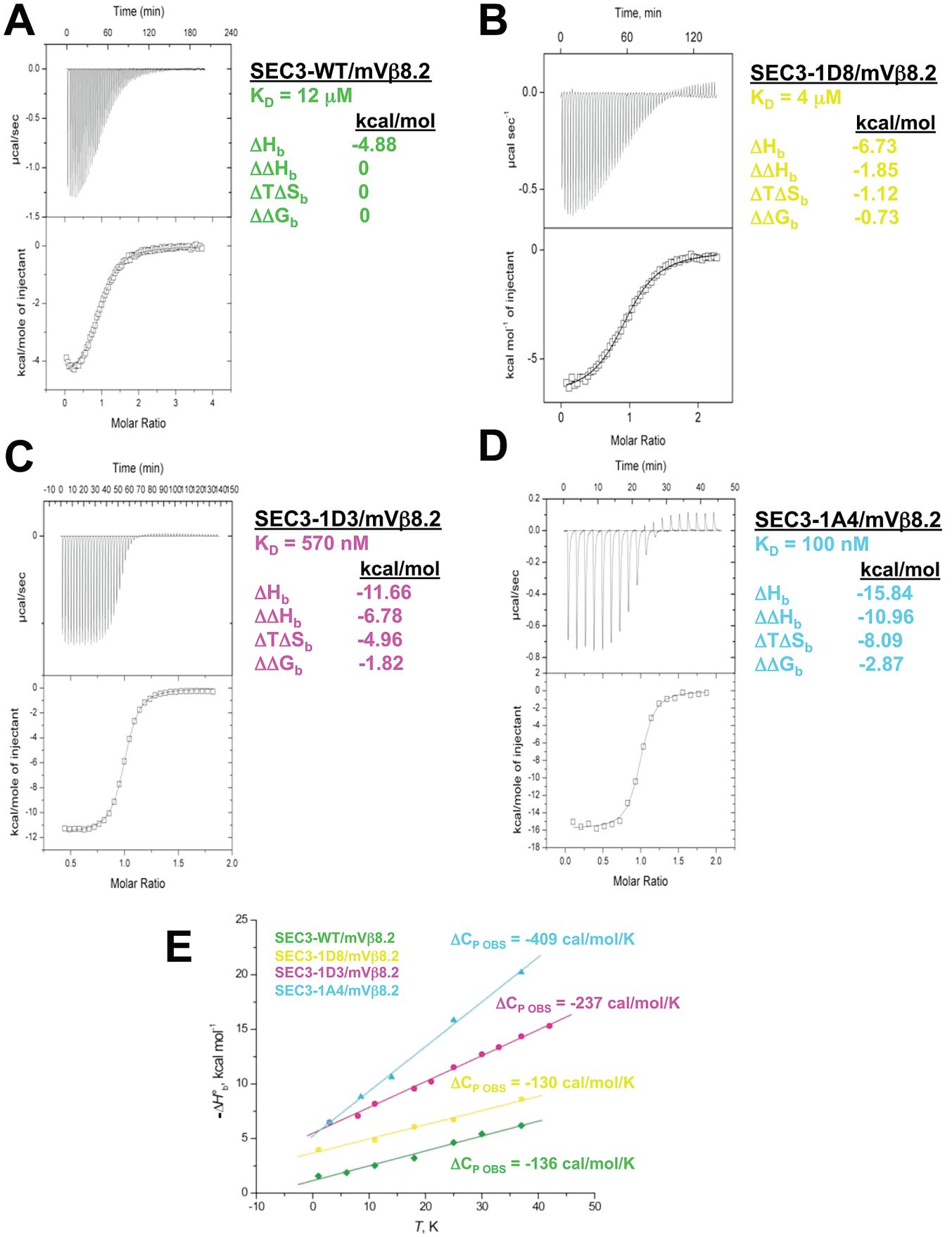


FIGURE 2: Thermodynamic analysis of SEC3/mV$\beta$8.2 complexes. ITC titrations for the (A) SEC3-WT/mV$\beta$8.2, (B) SEC3-1D8/mV$\beta$8.2, (C) SEC3-1D3/mV$\beta$8.2, and (D) SEC3-1A4/mV$\beta$8.2 interactions. The thermodynamic parameters for each interaction are listed. (E) Heat capacity plots for the four SEC3/mV$\beta$8.2 interactions in which enthalpy measurements from at least five temperatures for each interaction are plotted, from which the heat capacity changes for the wild-type and variant SEC3/mV$\beta$8.2 complexes are determined.

Table 2: Thermodynamic Parameters of the SEC3/mV$\beta$8.2 Interactions

| | SEC3-WT/mV$\beta$8.2 | SEC3-1D8/mV$\beta$8.2 | SEC3-1D3/mV$\beta$8.2 | SEC3-1A4/mV$\beta$8.2 |
|---|---|---|---|---|
| | | Isothermal Titration Calorimetry | | |
| $K_b$ ($M^{-1}$) | $(8.10 \pm 0.30) \times 10^4$ | $(2.78 \pm 0.12) \times 10^5$ | $(1.76 \pm 0.07) \times 10^6$ | $(1.04 \pm 0.12) \times 10^7$ |
| $K_D$ (M) | $1.2 \times 10^{-5}$ | $3.6 \times 10^{-6}$ | $5.7 \times 10^{-7}$ | $9.6 \times 10^{-8}$ |
| $\Delta G°_b$ (kcal/mol) | −6.70 | −7.43 | −8.52 | −9.57 |
| $\Delta\Delta G°_b$ (kcal/mol) | 0 | −0.73 | −1.82 | −2.78 |
| $\Delta H°_b$ (kcal/mol) | −4.88 ± 0.04 | −6.73 ± 0.06 | 11.66 ± 0.04 | −15.84 ± 0.14 |
| $T\Delta S°_b$ (kcal/mol) | 1.82 | 0.70 | −3.14 | −6.27 |
| $\Delta C_{P OBS}$ (cal/mol/K) | −136 ± 8 | −130 ± 8 | −237 ± 5 | −409 ± 11 |
| | | Heat Capacity Changes: Calculated, Conformational, and Folding | | |
| | | $\Delta C_{P CALC}$ (cal/mol/K) | | |
| ref *48* | −106 | −79 | −134 | −137 |
| ref *51* | −112 | −70 | −148 | −153 |
| ref *49* | −111 | −90 | −137 | −140 |
| ref *50* | −123 | −75 | −163 | −170 |
| | | $\Delta C_{P CONF}$ (cal/mol/K) | | |
| ref *48* | −30 | −51 | −103 | −272 |
| ref *51* | −24 | −60 | −89 | −256 |
| ref *49* | −25 | −40 | −100 | −269 |
| ref *50* | −13 | −55 | −74 | −239 |
| | | $\Delta C_{P FOLD}$ (cal/mol/K) | | |
| ref *48* | −27 | −49 | −121 | −188 |
| ref *51* | −26 | −65 | −157 | −154 |
| ref *49* | −30 | −45 | −112 | −108 |
| ref *50* | −28 | −73 | −178 | −175 |
| | | Protein−Protein Interaction Prediction Algorithms | | |
| | | experimentally determined values | | |
| $\Delta\Delta G_{MEASURED}$ | | −0.73 | −1.82 | −2.87 |
| | | calculated values: mutated *in silico* from wild-type complex | | |
| $\Delta\Delta G_{CALC\text{-}FoldX}$ | | −2.86 | −2.59 | −2.84 |
| $\Delta\Delta G_{CALC\text{-}Concoord/PBSA}$ | | −1.59 | −1.69 | −1.79 |
| | | calculated values: using X-ray crystal structure atomic coordinates | | |
| $\Delta G_{CALC\text{-}FoldX}$ | −5.32 | −5.67 | −10.94 | −7.32 |
| $\Delta\Delta G_{CALC\text{-}FoldX}$ | 0 | −0.35 | −5.62 | −2.00 |

was observed. Substantial changes in heat capacities, such as observed in our model system, have been taken to suggest a role for conformational flexibility in binding, as will be discussed below.

*Disorder-to-Order Structural Transitions Are Observed by Crystallographic Analysis of the Protein Complexes*. To determine whether disorder-to-order structural transitions had occurred upon binding, we solved the crystal structures of each of the SEC3 variants in complex with their common molecular target, mV$\beta$8.2. All of these structures were determined to a resolution of 2.2−2.3 Å (Table 3). These complexes are essentially structurally indistinguishable from one another, save for the disulfide loops, and the remainder of the interface is essentially the same for all of the complexes, which has been extensively detailed elsewhere (*23*). Also unchanging between the complex structures is the conformation of the TCR V$\beta$ domain, including its CDR1, HV4, and CDR2 loops with which the wild-type and variant SEC3 proteins interact (Figure 3A). The most obvious structural changes observed in mV$\beta$8.2 are restricted to the CDR3 and framework region apical loops, neither of which contacts SEC3. While our calorimetric experiments cannot eliminate the possibility that minor, or even imperceptible, differences in packing throughout the protein and outside of the interface itself contribute to some of the differences in affinity and thermodynamics, no clear structural evidence exists to implicate that the TCR V$\beta$ domain contributes differentially to complex formation throughout the affinity maturation process.

The disulfide loops of the wild-type and variant SEC3 proteins each adopted a different conformation when bound to the common protein ligand, mV$\beta$8.2 (Figure 3B). The disulfide loops in SEC3-WT (*23*) and the lowest affinity variant, SEC3-1D8, when in complex with mV$\beta$8.2, exhibit electron density in which the majority of the main chain and most side chains can be readily modeled (Figure 3C,D). Electron density maps are markedly improved in complexes formed by the next highest affinity variant, SEC3-1D3, and the highest affinity variant, SEC3-1A4, into which essentially all atoms can be unambiguously built (Figure 3E,F).

Inspection of the structures of the wild-type and variant SEC3 proteins alone and in complex with mV$\beta$8.2 demonstrates that

Table 3: Data Collection and Refinement Statistics for SEC3 Variants in Complex with mV$\beta$8.2[a]

| | SEC3-WT (*23*) | SEC3-1D8 | SEC3-1D3 | SEC3-1A4 |
|---|---|---|---|---|
| data collection | | | | |
| space group | *P*1 | *P*1 | $P6_5$ | $P6_5$ |
| unit cell dimensions | | | | |
| *a* (Å) | 64.16 | 63.05 | 96.69 | 96.97 |
| *b* (Å) | 70.46 | 70.23 | 96.69 | 96.97 |
| *c* (Å) | 98.37 | 98.51 | 91.84 | 92.51 |
| α (deg) | 74.18 | 74.85 | 90.00 | 90.00 |
| $\beta$ (deg) | 75.76 | 75.10 | 90.00 | 90.00 |
| $\gamma$ (deg) | 88.40 | 88.41 | 120.00 | 120.00 |
| complexes per asymmetric unit | 4 | 4 | 1 | 1 |
| resolution limit (Å) | 2.3 | 2.3 | 2.2 | 2.3 |
| observations | 272478 | 354953 | 219846 | 193253 |
| unique reflections | 63758 | 67140 | 24784 | 21938 |
| completeness (%) | 94.21 (73.7) | 95.9 (85.2) | 99.9 (94.9) | 99.7 (93.3) |
| $R_{merge}$ (%)[b] | 6.8 (27.3) | 4.4 (31.3) | 6.1 (33.3) | 5.8 (35.6) |
| refinement | | | | |
| $R_{cryst}$ (%)[c] | 21.2 (28.6) | 22.3 (36.9) | 20.5 | 22.9 |
| $R_{free}$ (%)[d] | 27.1 (34.4) | 27.9 (43.6) | 23.1 | 25.3 |
| protein residues | 1380 | 1365 | 343 | 343 |
| water molecules | 198 | 146 | 109 | 85 |
| average *B* factors (Å$^2$) | | | | |
| mV$\beta$8.2 | 40.8 | 46.0 | 43.8 | 26.9 |
| SEC3 | 47.5 | 53.4 | 35.3 | 18.5 |
| disulfide loop | 67.5 | 78.8 | 39.5 | 23.9 |
| rms deviations | | | | |
| bonds (Å) | 0.032 | 0.038 | 0.021 | 0.019 |
| angles (deg) | 2.531 | 2.870 | 1.721 | 1.843 |
| Ramachandran plot statistics | | | | |
| core (%) | 80.8 | 79.5 | 87.8 | 85.2 |
| allowed (%) | 17.5 | 18.3 | 11.2 | 13.3 |
| generous (%) | 1.4 | 1.8 | 1.0 | 1.5 |
| disallowed (%) | 0.2 | 0.3 | 0 | 0 |

[a]Values in parentheses correspond to the highest resolution shell: SEC3-WT (2.36−2.30 Å); SEC3-1D8 (2.36−2.30 Å); SEC3-1D3 (2.05−2.00 Å); SEC3-1A4 (1.95−1.90 Å). [b]$R_{merge}(I) = \sum|I(i) - \langle I(h)\rangle|/\sum I(i)$, where $I(i)$ is the $i$th observation of the intensity of the $hkl$ reflection and $\langle I\rangle$ is the mean intensity from multiple measurements of the $h,k,l$ reflection. [c]$R_{cryst}(F) = \sum_h||F_o(h)| - |F_c(h)||/\sum_h|F_o(h)|$, where $|F_o(h)|$ and $|F_c(h)|$ are the observed and calculated structure factor amplitudes for the $h,k,l$ reflection. [d]$R_{free}$ is calculated over reflections in a test set not included in atomic refinement: SEC3-WT, 3394 reflections, 5.1%; SEC3-1D8, 3424 reflections, 5.1%; SEC3-1D3, 960 reflections, 5.2%; SEC3-1A4, 630 reflections, 5.2%.

they all exhibit distinct disorder-to-order transitions upon binding. These findings are consistent with the premise that, in this model protein−protein interaction system, directed evolution does not select for increased affinity through preorganization of the unbound state but through some combination of a greater propensity to order upon binding and a reduction of the interaction off-rate (*34*). Possessing such a series of affinity-matured variant protein complexes in which the most notable difference is the dynamic behavior of a disordered region provides a unique opportunity to assess the energetic contributions to binding from disordered elements of a protein.

*The Evolved Protein Variants Exhibit Increased Structural Ordering with Increased Affinity*. Differences in the degree of ordering in the wild-type and variant SEC3/mV$\beta$8.2 complexes become more evident when all molecules within the asymmetric unit are examined. Complexes of SEC3-WT and SEC3-1D8 with mV$\beta$8.2 crystallize isomorphously in the monoclinic space group with four complexes in the asymmetric unit. When the disulfide loops from each of the four complexes are superimposed, the conformational variability between like loops is greater in SEC3-WT as compared to that of the affinity-matured variant SEC3-1D8 (Figure 4A,B). The conformational variability in these two lower affinity complexes is also reflected in the average temperature factors of residues in these crystals, in which those for the disulfide loop residues are higher than those for non-disulfide loop residues in the SEC3 molecules or for residues in mV$\beta$8.2 (Table 3). The higher affinity variants SEC3-1D3 and SEC3-1A4 in complex with mV$\beta$8.2 crystallize isomorphously with one another in a hexagonal space group with only a single complex per asymmetric unit, and thus, each disulfide loop from SEC3-1D3 and SEC3-1A4 was observed in only a single defined conformation in its bound state (Figure 4C,D). It is likely that the structural homogeneity of the disulfide loops in these two variant complex structures is maintained by the intermolecular contacts made with mV$\beta$8.2 (see below) and is what allows for crystal packing in a higher symmetry space group. In analogy to the lower affinity complexes, this suggested restriction of conformational space of the disulfide loops in these higher affinity complexes is reflected in the average temperature factors of these crystals, as those for the disulfide loop residues are intermediate to those for non-disulfide loop residues in the SEC3 molecules and for residues in mV$\beta$8.2 (Table 3).

*Protein Ordering Results in Increased Intermolecular Contacts*. The disorder-to-order transitions of the disulfide loops of the SEC3 proteins result in intermolecular contacts being made between the previously disordered region and the molecular target, mV$\beta$8.2. For SEC3-WT and the lowest affinity variant SEC3-1D8, only residues from the five amino acid stretch that had been altered during the directed evolution procedure contact mV$\beta$8.2, while the remaining nine genetically constant disulfide loop residues

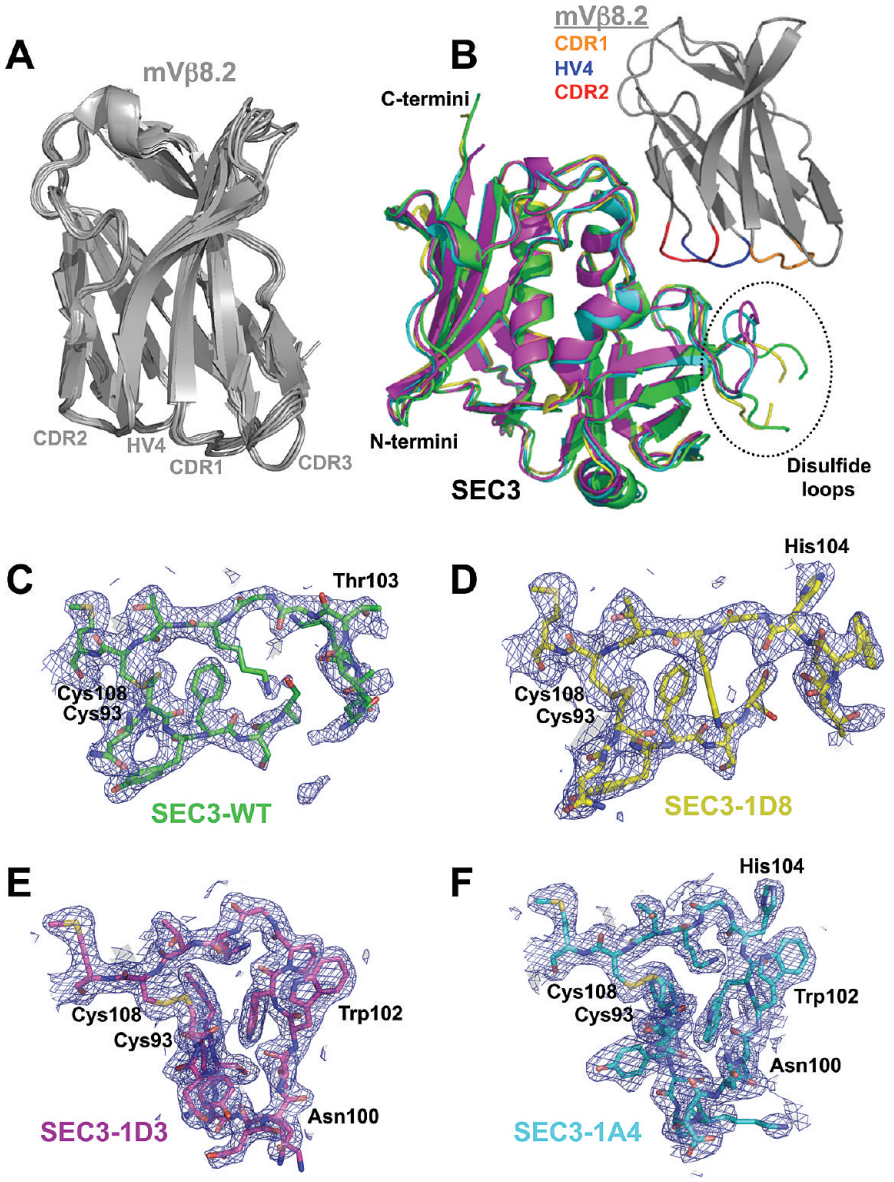


FIGURE 3: Crystallographic analysis of wild-type and variant SEC3 molecules in complex with their common molecular target, mV$\beta$8.2. (A) Superposition of all crystallographically observed mV$\beta$8.2 molecules. (B) Superposition of the X-ray crystal structures of the SEC3-WT/mV$\beta$8.2 (green), SEC3-1D8/mV$\beta$8.2 (yellow), SEC3-1D3/mV$\beta$8.2 (magenta), and SEC3-1A4/mV$\beta$8.2 (cyan) complexes. The hypervariable loops of mV$\beta$8.2 contacted by the wild-type and variant SEC3 molecules are highlighted: CDR1 (orange), HV4 (blue), and CDR2 (red). The SEC3 disulfide loops are highlighted by the oval. Shown are $\sigma_A$-weighted $2F_o - F_c$ electron density maps contoured at 1.2$\sigma$ of the disulfide loop regions of (C) SEC3-WT, (D) SEC3-1D8, (E) SEC3-1D3, and (F) SEC3-1A4 in the complex structures. Residues labeled include Cys93 and Cys108, which form the disulfide bridge, as well as those that make intermolecular contacts with mV$\beta$8.2.

do not. Although the majority of disulfide loop contacts between the two highest affinity variants, SEC3-1D3 and SEC3-1A4, and mV$\beta$8.2 are made by the altered residues, some contacts are made by one residue preceding the five residues targeted by phage display. For all of the complexes in this model system, the intermolecular contacts of the disulfide loop residues in the wild-type and evolved complexes increase with increasing affinity.

In one of the four SEC3-WT/mV$\beta$8.2 complexes in the asymmetric unit, residues Thr103$^{\text{SEC3-WT}}$ and Asn28$^{\text{mV}\beta\text{8.2}}$ interact (Figure 5A), while there are no contacts in this region for protein complexes in any of the three asymmetric units. This is similar to the SEC3-1D8/mV$\beta$8.2 complex, in which His104$^{\text{SEC-1D8}}$ contacts the same mV$\beta$8.2 residue, Asn28$^{\text{mV}\beta\text{8.2}}$ (Figure 5B). A single residue, Trp102$^{\text{SEC3-1D3}}$, in the next highest affinity variant, SEC3-1D3, makes numerous van der Waals contacts with residues Asn28$^{\text{mV}\beta\text{8.2}}$, Asn30$^{\text{mV}\beta\text{8.2}}$, and Gln72$^{\text{mV}\beta\text{8.2}}$ (Figure 5C). A single hydrogen bond is made between the main chain carboxyl group of Asn100$^{\text{SEC3-1D3}}$ and Asn30$^{\text{mV}\beta\text{8.2}}$. The highest affinity variant, SEC3-1A4, contains the only disulfide loop that utilizes more than one of the altered residue side chains to form the interface with mV$\beta$8.2. In the SEC3-1A4/mV$\beta$8.2 complex, Trp102$^{\text{SEC3-1A4}}$ adopts a conformation and makes contacts similar to the same residue in SEC3-1D3 (Figure 5D). The side chains of two other residues, though, Asn100$^{\text{SEC3-1A4}}$ and His104$^{\text{SEC3-1A4}}$, make water-mediated hydrogen bonds to residues Asn30$^{\text{mV}\beta\text{8.2}}$ and Glu73$^{\text{mV}\beta\text{8.2}}$, respectively.

*Comprehensive Analysis of Structural and Thermodynamic Changes Suggests That Conformational Changes Are Responsible for Affinity Increases*. In macromolecular interactions, the observed heat capacity ($\Delta C_{P\,OBS}$), as measured by ITC analysis at multiple temperatures, is composed of contributions from both the nonpolar and polar buried surface areas

($\Delta ASA_N$ and $\Delta ASA_P$, respectively). These buried surface areas are readily obtained from high-resolution crystal structures. For rigid-body interactions, or those protein complexes involving little conformational flexibility, heat capacity changes can be calculated ($\Delta C_{P\,CALC}$) that closely approximate $\Delta C_{P\,OBS}$, according to the equation:

$$\Delta C_{P\,CALC} = C_N \Delta ASA_N + C_P \Delta ASA_P \quad (1)$$

where $C_N$ and $C_P$ are coefficients that relate the heat capacity changes associated with the hydration of nonpolar and polar groups, respectively. These coefficients have been experimentally determined using a variety of data sets. While the values of $C_N$ and $C_P$ vary depending on the data set, they generally agree that the hydration of nonpolar groups results in heat capacity increases while the hydration of polar groups results in heat capacity decreases (*47*).

When protein plasticity plays a significant role in complex formation, however, it is often the case that $\Delta C_{P\,CALC}$ is much less than $\Delta C_{P\,OBS}$. The heat capacity changes potentially resulting from conformational changes induced upon complex formation ($\Delta C_{P\,CONF}$) can be estimated by the difference between the observed and calculated heat capacity changes as follows:

$$\Delta C_{P\,CONF} = \Delta C_{P\,OBS} - \Delta C_{P\,CALC} \quad (2)$$

In the SEC3/mV$\beta$8.2 protein−protein interaction system, changes in $\Delta C_{P\,CONF}$, relative to that of the wild-type complex, as a measure of conformational change upon binding, are indicative of the degree of ordering induced during the binding process. We used four distinct sets (*48*−*51*) of previously published $C_N$ and $C_P$ coefficients, along with $\Delta ASA_N$ and $\Delta ASA_P$ values from our

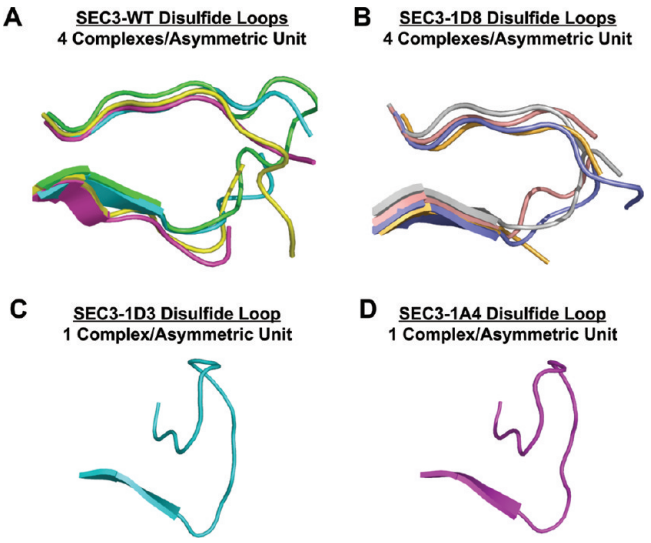


FIGURE 4: In-crystal conformational variability of, and intermolecular contacts made by, the SEC3 disulfide loops. Superposition of the disulfide loops from each of four complexes per asymmetric unit in the (A) SEC3-WT/mV$\beta$8.2 and (B) SEC3-1D8/mV$\beta$8.2 complexes. The disulfide loops from the single complex per asymmetric unit in each of the (C) SEC3-1D3/mV$\beta$8.2 and (D) SEC3-1A4/mV$\beta$8.2 complexes are shown.

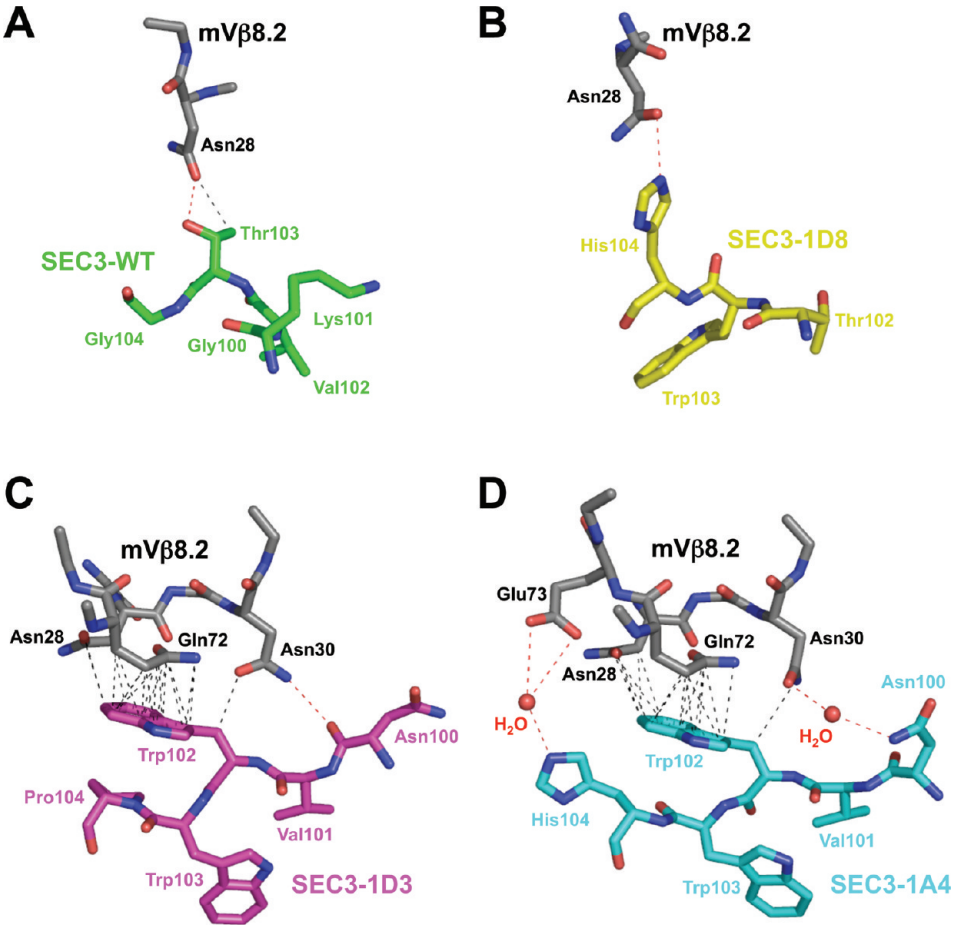


FIGURE 5: Intermolecular contacts made by SEC3 disulfide residues to mV$\beta$8.2 in the (A) SEC3-WT/mV$\beta$8.2, (B) SEC3-1D8/mV$\beta$8.2, (C) SEC3-1D3/mV$\beta$8.2, and (D) SEC3-1A4/mV$\beta$8.2 complexes. Hydrogen bonds are represented by red dashed lines and van der Waals interactions by black dashed lines.

SEC3/mV$\beta$8.2 complex crystal structures (in this case from the one complex in each asymmetric unit in which the SEC3 disulfide loop residues make intermolecular contacts with mV$\beta$8.2) to determine $\Delta C_{P\,CALC}$ and $\Delta C_{P\,CONF}$ values for each of the these complexes (Table 2). When the logarithms of these $\Delta C_{P\,CONF}$ values are plotted versus their affinities, they exhibit linear correlations (Supporting Information Figure 1), with $R^2$ values ranging from 0.90 to 0.99. Since the enthalpic component of binding is largely captured in the $\Delta C_{P\,CALC}$ term, this result suggests that the degree of ordering in the SEC3 disulfide loops imposed by directed evolution is likely responsible for much of the affinity increases in these protein−protein complexes.

*Modeled Heat Capacity Changes Resulting from the Folding of the Disulfide Loop Largely Account for the Discrepancies between Observed and Calculated Heat Capacities*. Discrepancies between observed and calculated heat capacities in macromolecular interactions can be derived from factors outside of conformational changes, including, but not limited to, the hydration of hydrophobic groups, hydrogen bonding, electrostatics, protein conformational entropy, vibrational terms, and linked equilibria (*52*). Thus, in order to implicate more directly the ordering of the wild-type and variant SEC3 disulfide loops upon binding to mV$\beta$8.2 as a source of the $\Delta C_{P\,CONF}$ values exhibited by these protein complexes, we attempted to quantify the likely heat capacity changes that would derive specifically from these structural transitions. We calculated the changes in nonpolar and polar buried surface areas between fully extended conformations of each of the SEC3 disulfide loops versus the conformations of each disulfide loop in the SEC3/mV$\beta$8.2 complex structures (Supporting Information Figure 2A). The extended conformations were modeled by mutating the residues at positions 9 through 24 of an extended region of the OmpF protein (PDB ID 2FZG), as a surrogate protein backbone, to those of the SEC3 disulfide loops. Notably, these analyses were done in the absence of the remainder of the SEC3/mV$\beta$8.2 complex. Then, using the four distinct sets (*48*−*51*) of previously published $C_N$ and $C_P$ coefficients as above, we calculated four corresponding heat capacity changes that could be solely ascribed to the folding of each of the wild-type and variant SEC3 disulfide loops. We designated these values as $\Delta C_{P\,FOLD}$ and list them in Table 2.

If the folding of the disulfide loop is responsible for the discrepancies between observed and calculated heat capacities in our model system, then for each SEC3/mV$\beta$8.2 complex, $\Delta C_{P\,OBS}$ should be equal to the sum of the $\Delta C_{P\,CALC}$ and $\Delta C_{P\,FOLD}$ values (or, analogously, $\Delta C_{P\,FOLD}$ should equal $\Delta C_{P\,CONF}$). To evaluate this, we plotted $\Delta C_{P\,OBS}$ and $[\Delta C_{P\,CALC} + \Delta C_{P\,FOLD}]$ for each of the SEC3/mV$\beta$8.2 complexes using each set of $C_N$ and $C_P$ coefficients (Supporting Information Figure 2B).

We observed that for SEC3-WT, SEC3-1D8, and SEC3-1D3 the sum of the $\Delta C_{P\,CALC}$ and $\Delta C_{P\,FOLD}$ values closely approximates or overestimates $\Delta C_{P\,OBS}$, suggesting that the folding of the SEC3-1D3 disulfide loop makes significant contributions to the observed heat capacity changes in these complexes.

For SEC3-1A4, each of the four sets of $C_N$ and $C_P$ coefficients resulted in sums of $\Delta C_{P\,CALC}$ and $\Delta C_{P\,FOLD}$ values that underestimated that of $\Delta C_{P\,OBS}$. Although the folding of the SEC3-1A4 loop likely contributes substantially to the $\Delta C_{P\,OBS}$ of the SEC3-1A4/mV$\beta$8.2 complex, some other factor, such as the inclusion of two ordered water molecules in the variant interface (Figure 5D), appears to also contribute to the measured change in heat capacity. Bridging water molecules have been shown to increase the magnitude of the negative heat capacity term for complex formation in other macromolecular interactions (*53*) and, upon being incorporated into a macromolecular interface, may suffer a reduction of their "soft" vibrational modes with similar effects (*52*, *54*, *55*).

*Energetic Contributions of the Structural Transitions Are Poorly Estimated by Currently Available Protein−Protein Interaction Prediction Algorithms*. The large conformational heat capacity differences correlating with increased affinity suggests that a large portion of the measured energetic differences between the wild-type and variant SEC3 molecules binding to mV$\beta$8.2 are due to a disorder−order transition of the disulfide loop. We estimated binding energies resulting from the affinity maturation of the SEC3/mV$\beta$8.2 model system using state-of-the-art modeling algorithms.

Numerous docking algorithms, including RosettaDock (*56*, *57*) and FoldX(*58*), as well as a more recent algorithm, Concoord/PBSA (*59*) that is purported to better take into account conformational changes, are publically available to perform such calculations. Not only are there multiple algorithms but there exist a number of ways in which one can conceivably perform these calculations, depending on the starting structures and mutation scenarios. We used a combination of both predictive algorithms and mutation scenarios in an attempt to computationally model the structural transitions in the SEC3/mV$\beta$8.2 complexes that we observed experimentally, as described below.

Starting from the structure of the wild-type SEC3/mV$\beta$8.2 structure, the sequences of the disulfide loop were mutated *in silico* according to the sequences of each of the SEC3 variants, and changes in binding free energies as a result of the mutations were calculated using the programs FoldX and Concoord/PBSA. The relative free energy changes for the three mutant complexes are similar to one another regardless of the mutations made, such that there is no clear pattern in the calculated affinity changes along the affinity maturation pathway, in contrast to what is observed experimentally (Table 2). Concoord/PBSA generates an ensemble of structures based on the information provided, while FoldX generates a single atomic coordinate file. The mutated structures generated by both of these programs (with no information from the actual variant crystal structures) are not noticeably different from the original wild-type structure, with only minor changes in main chain conformation (data not shown), and are thus entirely different from the experimentally determined crystal structures. Neither the energetic changes nor the structural changes generated by FoldX and Concoord/PBSA realistically model the free energy changes in this system.

The FoldX program is capable of calculating an absolute binding free energy for a single structure, in addition to relative changes in binding free energy between two structures as described above. Thus, we used FoldX to calculate absolute binding free energies of the wild-type and three variant SEC3/mV$\beta$8.2 complexes for which we had determined the X-ray crystal structures, using the experimentally determined atomic coordinates. We then compared the resulting changes in calculated binding free energies to those measured experimentally by ITC. While FoldX is able to predict that SEC3-1D3 and SEC3-1A4 bind mV$\beta$8.2 significantly tighter than do SEC3-1D8 and SEC3-WT, the calculated binding free energy changes reflect neither the magnitude of the measured values nor the rank order of affinities for the evolved variant complexes (Table 2). These results indicate that current computational approaches are unable to account for the disordered-to-ordered structural transitions exhibited by this set of affinity-matured protein−protein complexes but instead derive

their estimations largely from the intermolecular contacts contained within the experimentally determined atomic coordinates.

## DISCUSSION

The set of evolved protein−protein interactions presented here provides a unique opportunity to assess the energetic contributions to binding from a disordered protein region. Although such contributions can be significant, quantitatively assessing them separately from those made by ordered regions of the protein, which often constitute the majority of the binding interface, is not a straightforward task. By combining directed evolution with comprehensive crystallographic and thermodynamic analyses of variants along the resulting affinity maturation pathway, we have quantified these effects in one particular protein−protein complex. To our knowledge, no other affinity-matured protein−protein interaction has been evaluated at this level of detail for this purpose. Short of identifying a series of naturally arising mutations in a protein complex that serendipitously alters protein conformational flexibility in a systematic way, concomitant with a change in protein complex affinity, it is difficult to develop appropriate and well-controlled model systems for quantifying such energetic contributions in the absence of directed evolution techniques.

We show that affinity changes of several orders of magnitude can result predominantly from changes in the dynamic properties of a disordered protein region. Our results suggest that the ability to modify the affinity of a given protein−protein complex by applying evolutionary pressure to a disordered protein region could emerge as a viable route for protein engineering and design in which proteins that exhibit disordered regions near the protein−protein interface could be altered to produce molecules with improved binding capacities or even unique functions.

In our model system, we find that a disordered region of a protein can contribute in an energetically important way to protein binding and that this energetic significance can be altered without preorganization of the disordered region. In this affinity-matured SEC3/mV$\beta$8.2 complex, the structural changes (assessed by crystallographic studies) and the thermodynamic changes (determined by calorimetric methods) associated with complex formation are in excellent agreement. As affinity increases throughout the directed evolution process, an increasingly large entropic penalty is paid upon binding. However, as is expected from the ordering of previously disordered regions that influence binding affinity, each entropic cost along the affinity maturation pathway is more than compensated by an enthalpic gain.

The relative thermodynamic changes between the wild-type and variant SEC3/mV$\beta$8.2 complexes can be attributed largely to structural transitions that the disulfide loops undergo for several reasons. First, portions of the protein−protein interface, beyond those involving the evolved SEC3 disulfide loops, have not been altered genetically, and we have shown by X-ray crystallography that they are structurally indistinguishable among all of the affinity-matured complexes. Thus, their effects on binding should remain constant among all of the protein complexes that we have evaluated. Second, by calculating heat capacity changes derived solely from the folding of the wild-type and variant SEC3 disulfide loops upon complex formation, in the absence of the remainder of the protein complex, we show that these loop-derived heat capacity changes either fully approximate (in the cases of the SEC3-WT, SEC3-1D8, and SEC3-1D3 complexes with mV$\beta$8.2) or make substantial contributions (in the case of the SEC3-1A4/mV$\beta$8.2 complex) to the discrepancies between the overall observed and calculated heat capacity changes in each of the four SEC3/mV$\beta$8.2 complexes. The most notable difference between the two highest affinity variant complexes, those involving SEC3-1D3 and SEC3-1A4, is two water molecules that are present, according to discrete electron density on the crystal structure.

The inability of several state-of-the-art algorithms to model the energetic contributions of the structural transitions induced upon binding in the affinity-matured SEC3/mV$\beta$8.2 protein−protein interaction system indicates that additional improvements to these algorithms are necessary. Such modifications are often made using a controlled training set of interactions that exhibit a characteristic biophysical trait that is to be modeled. One major impediment to proper prediction of the energetic contributions to protein−protein interactions due to conformational changes may be the lack of appropriate training sets. The methods employed here, including directed evolution, X-ray crystallography, and ITC, could provide a precise and targeted training set of interactions to aid in overcoming this current shortcoming of predictive algorithms for protein−protein interactions. Indeed, another set of SEC3/mV$\beta$8.2 affinity-matured variant complexes that we had previously generated and analyzed (*23*, *25*, *41*) has been used recently to model intermolecular cooperativity and to establish a general computational approach to handle explicit interfacial water molecules (*60*).

## ACKNOWLEDGMENT

We thank the staff at beamline X12B at the National Synchrotron Light Source, Brookhaven National Laboratory.

## SUPPORTING INFORMATION AVAILABLE

Figures showing correlations between affinities and conformational heat capacities, as well as the modeling of conformational heat capacity changes due to loop disorder-to-order transitions. This material is available free of charge via the Internet at http://pubs.acs.org.

## REFERENCES

1. Gascoigne, N. R., and Zal, T. (2004) Molecular interactions at the T cell-antigen-presenting cell interface. *Curr. Opin. Immunol. 16*, 114–119.
2. Pawson, T., and Nash, P. (2000) Protein-protein interactions define specificity in signal transduction. *Genes Dev. 14*, 1027–1047.
3. Warren, A. J. (2002) Eukaryotic transcription factors. *Curr. Opin. Struct. Biol. 12*, 107–114.
4. Rual, J. F., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T., Dricot, A., Li, N., Berriz, G. F., Gibbons, F. D., Dreze, M., Ayivi-Guedehoussou, N., Klitgord, N., Simon, C., Boxem, M., Milstein, S., Rosenberg, J., Goldberg, D. S., Zhang, L. V., Wong, S. L., Franklin, G., Li, S., Albala, J. S., Lim, J., Fraughton, C., Llamosas, E., Cevik, S., Bex, C., Lamesch, P., Sikorski, R. S., Vandenhaute, J., Zoghbi, H. Y., Smolyar, A., Bosak, S., Sequerra, R., Doucette-Stamm, L., Cusick, M. E., Hill, D. E., Roth, F. P., and Vidal, M. (2005) Towards a proteome-scale map of the human protein-protein interaction network. *Nature 437*, 1173–1178.
5. Li, S., Armstrong, C. M., Bertin, N., Ge, H., Milstein, S., Boxem, M., Vidalain, P. O., Han, J. D., Chesneau, A., Hao, T., Goldberg, D. S., Li, N., Martinez, M., Rual, J. F., Lamesch, P., Xu, L., Tewari, M., Wong, S. L., Zhang, L. V., Berriz, G. F., Jacotot, L., Vaglio, P., Reboul, J., Hirozane-Kishikawa, T., Li, Q., Gabel, H. W., Elewa, A., Baumgartner, B., Rose, D. J., Yu, H., Bosak, S., Sequerra, R., Fraser, A., Mango, S. E., Saxton, W. M., Strome, S., Van Den Heuvel, S., Piano, F., Vandenhaute, J., Sardet, C., Gerstein, M., Doucette-Stamm, L., Gunsalus, K. C., Harper, J. W., Cusick, M. E., Roth, F. P., Hill, D. E., and Vidal, M. (2004) A map of the interactome network of the metazoan *C. elegans*. *Science 303*, 540–543.
6. Bouwmeester, T., Bauch, A., Ruffner, H., Angrand, P. O., Bergamini, G., Croughton, K., Cruciat, C., Eberhard, D., Gagneur, J., Ghidelli, S., Hopf, C., Huhse, B., Mangano, R., Michon, A. M., Schirle, M.,

Schlegl, J., Schwab, M., Stein, M. A., Bauer, A., Casari, G., Drewes, G., Gavin, A. C., Jackson, D. B., Joberty, G., Neubauer, G., Rick, J., Kuster, B., and Superti-Furga, G. (2004) A physical and functional map of the human TNF-alpha/NF-kappa B signal transduction pathway. *Nat. Cell Biol. 6*, 97–105.
7. Giot, L., Bader, J. S., Brouwer, C., Chaudhuri, A., Kuang, B., Li, Y., Hao, Y. L., Ooi, C. E., Godwin, B., Vitols, E., Vijayadamodar, G., Pochart, P., Machineni, H., Welsh, M., Kong, Y., Zerhusen, B., Malcolm, R., Varrone, Z., Collis, A., Minto, M., Burgess, S., McDaniel, L., Stimpson, E., Spriggs, F., Williams, J., Neurath, K., Ioime, N., Agee, M., Voss, E., Furtak, K., Renzulli, R., Aanensen, N., Carrolla, S., Bickelhaupt, E., Lazovatsky, Y., DaSilva, A., Zhong, J., Stanyon, C. A., Finley, R. L., Jr., White, K. P., Braverman, M., Jarvie, T., Gold, S., Leach, M., Knight, J., Shimkets, R. A., McKenna, M. P., Chant, J., and Rothberg, J. M. (2003) A protein interaction map of *Drosophila melanogaster*. *Science 302*, 1727–1736.
8. Uetz, P., Giot, L., Cagney, G., Mansfield, T. A., Judson, R. S., Knight, J. R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P., Qureshi-Emili, A., Li, Y., Godwin, B., Conover, D., Kalbfleisch, T., Vijayadamodar, G., Yang, M., Johnston, M., Fields, S., and Rothberg, J. M. (2000) A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature 403*, 623–627.
9. Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M., and Sakaki, Y. (2001) A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc. Natl. Acad. Sci. U.S.A. 98*, 4569–4574.
10. Dessailly, B. H., Nair, R., Jaroszewski, L., Fajardo, J. E., Kouranov, A., Lee, D., Fiser, A., Godzik, A., Rost, B., and Orengo, C. (2009) PSI-2: structural genomics to cover protein domain family space. *Structure 17*, 869–881.
11. Clackson, T., and Wells, J. A. (1995) A hot spot of binding energy in a hormone-receptor interface. *Science 267*, 383–386.
12. DeLano, W. L. (2002) Unraveling hot spots in binding interfaces: progress and challenges. *Curr. Opin. Struct. Biol. 12*, 14–20.
13. Sundberg, E. J., Urrutia, M., Braden, B. C., Isern, J., Tsuchiya, D., Fields, B. A., Malchiodi, E. L., Tormo, J., Schwarz, F. P., and Mariuzza, R. A. (2000) Estimation of the hydrophobic effect in an antigen-antibody protein-protein interface. *Biochemistry 39*, 15375–15387.
14. Li, Y., Huang, Y., Swaminathan, C. P., Smith-Gill, S. J., and Mariuzza, R. A. (2005) Magnitude of the hydrophobic effect at central versus peripheral sites in protein-protein interfaces. *Structure* (*Cambridge*) *13*, 297–307.
15. Reichmann, D., Rahat, O., Albeck, S., Meged, R., Dym, O., and Schreiber, G. (2005) The modular architecture of protein-protein binding interfaces. *Proc. Natl. Acad. Sci. U.S.A. 102*, 57–62.
16. Reichmann, D., Rahat, O., Cohen, M., Neuvirth, H., and Schreiber, G. (2007) The molecular architecture of protein-protein binding sites. *Curr. Opin. Struct. Biol. 17*, 67–76.
17. Keskin, O., Ma, B., and Nussinov, R. (2005) Hot regions in protein–protein interactions: the organization and contribution of structurally conserved hot spot residues. *J. Mol. Biol. 345*, 1281–1294.
18. Keskin, O., Ma, B., Rogale, K., Gunasekaran, K., and Nussinov, R. (2005) Protein-protein interactions: organization, cooperativity and mapping in a bottom-up Systems Biology approach. *Phys. Biol. 2*, S24–S35.
19. Keskin, O., Tsai, C. J., Wolfson, H., and Nussinov, R. (2004) A new, structurally nonredundant, diverse data set of protein-protein interfaces and its implications. *Protein Sci. 13*, 1043–1055.
20. Yuen, C. M., and Liu, D. R. (2007) Dissecting protein structure and function using directed evolution. *Nat. Methods 4*, 995–997.
21. Campbell, C. J., and Wittrup, K. D. (2007) Protein technologies. *Curr. Opin. Biotechnol. 18*, 293–294.
22. Gai, S. A., and Wittrup, K. D. (2007) Yeast surface display for protein engineering and characterization. *Curr. Opin. Struct. Biol. 17*, 467–473.
23. Cho, S., Swaminathan, C. P., Yang, J., Kerzic, M. C., Guan, R., Kieke, M. C., Kranz, D. M., Mariuzza, R. A., and Sundberg, E. J. (2005) Structural basis of affinity maturation and intramolecular cooperativity in a protein-protein interaction. *Structure* (*Cambridge*) *13*, 1775–1787.
24. Moza, B., Buonpane, R. A., Zhu, P., Herfst, C. A., Rahman, A. K., McCormick, J. K., Kranz, D. M., and Sundberg, E. J. (2006) Long-range cooperative binding effects in a T cell receptor variable domain. *Proc. Natl. Acad. Sci. U.S.A. 103*, 9867–9872.
25. Yang, J., Swaminathan, C. P., Huang, Y., Guan, R., Cho, S., Kieke, M. C., Kranz, D. M., Mariuzza, R. A., and Sundberg, E. J. (2003) Dissecting cooperative and additive binding energetics in the affinity maturation pathway of a protein-protein interface. *J. Biol. Chem. 278*, 50412–50421.
26. Goh, C. S., Milburn, D., and Gerstein, M. (2004) Conformational changes associated with protein-protein interactions. *Curr. Opin. Struct. Biol. 14*, 104–109.
27. Bastard, K., Prevost, C., and Zacharias, M. (2006) Accounting for loop flexibility during protein-protein docking. *Proteins 62*, 956–969.
28. Dominguez, C., Boelens, R., and Bonvin, A. M. (2003) HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc. 125*, 1731–1737.
29. Schneidman-Duhovny, D., Inbar, Y., Nussinov, R., and Wolfson, H. J. (2005) Geometry-based flexible and symmetric protein docking. *Proteins 60*, 224–231.
30. Smith, G. R., Sternberg, M. J., and Bates, P. A. (2005) The relationship between the flexibility of proteins and their conformational states on forming protein-protein complexes with an application to protein-protein docking. *J. Mol. Biol. 347*, 1077–1101.
31. Wang, C., Bradley, P., and Baker, D. (2007) Protein-protein docking with backbone flexibility. *J. Mol. Biol. 373*, 503–519.
32. Fasolini, M., Wu, X., Flocco, M., Trosset, J. Y., Oppermann, U., and Knapp, S. (2003) Hot spots in Tcf4 for the interaction with beta-catenin. *J. Biol. Chem. 278*, 21092–21098.
33. Hu, D., Crist, M., Duan, X., and Gimble, F. S. (1999) Mapping of a DNA binding region of the PI-sceI homing endonuclease by affinity cleavage and alanine-scanning mutagenesis. *Biochemistry 38*, 12621–12628.
34. Andersen, P. S., Geisler, C., Buus, S., Mariuzza, R. A., and Karjalainen, K. (2001) Role of the T cell receptor ligand affinity in T cell activation by bacterial superantigens. *J. Biol. Chem. 276*, 33452–33457.
35. Otwinowski, Z., and Minor, W. (1997) Processing X-ray diffraction data collected in oscillation mode. *Methods Enzymol. 276*, 307–326.
36. CCP4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 50*, 760–763.
37. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D54*, 905–921.
38. McRee, D. E. (1999) XtalView/Xfit—a versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol. 125*, 156–165.
39. Churchill, H. R., Andersen, P. S., Parke, E. A., Mariuzza, R. A., and Kranz, D. M. (2000) Mapping the energy of superantigen *Staphylococcus* enterotoxin C3 recognition of an alpha/beta T cell receptor using alanine scanning mutagenesis. *J. Exp. Med. 191*, 835–846.
40. Fields, B. A., Malchiodi, E. L., Li, H., Ysern, X., Stauffacher, C. V., Schlievert, P. M., Karjalainen, K., and Mariuzza, R. A. (1996) Crystal structure of a T-cell receptor beta-chain complexed with a superantigen. *Nature 384*, 188–192.
41. Kieke, M. C., Sundberg, E., Shusta, E. V., Mariuzza, R. A., Wittrup, K. D., and Kranz, D. M. (2001) High affinity T cell receptors from yeast display libraries block T cell activation by superantigens. *J. Mol. Biol. 307*, 1305–1315.
42. Marrack, P., and Kappler, J. (1990) The staphylococcal enterotoxins and their relatives. *Science 248*, 705–711.
43. McCormick, J. K., Yarwood, J. M., and Schlievert, P. M. (2001) Toxic shock syndrome and bacterial superantigens: an update. *Annu. Rev. Microbiol. 55*, 77–104.
44. Sundberg, E. J., Li, Y., and Mariuzza, R. A. (2002) So many ways of getting in the way: diversity in the molecular architecture of superantigen-dependent T-cell signaling complexes. *Curr. Opin. Immunol. 14*, 36–44.
45. Chi, Y. I., Sadler, I., Jablonski, L. M., Callantine, S. D., Deobald, C. F., Stauffacher, C. V., and Bohach, G. A. (2002) Zinc-mediated dimerization and its effect on activity and conformation of staphylococcal enterotoxin type C. *J. Biol. Chem. 277*, 22839–22846.
46. Sundberg, E. J., Andersen, P. S., Schlievert, P. M., Karjalainen, K., and Mariuzza, R. A. (2003) Structural, energetic and functional analysis of a protein-protein interface at distinct stages of affinity maturation. *Structure 11*, 1151–1161.
47. Prabhu, N. V., and Sharp, K. A. (2005) Heat capacity in proteins. *Annu. Rev. Phys. Chem. 56*, 521–548.
48. Spolar, R. S., Livingstone, J. R., and Record, M. T., Jr. (1992) Use of liquid hydrocarbon and amide transfer data to estimate contributions to thermodynamic functions of protein folding from the removal of nonpolar and polar surface from water. *Biochemistry 31*, 3947–3955.
49. Myers, J. K., Pace, C. N., and Scholtz, J. M. (1995) Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci. 4*, 2138–2148.

50. Makhatadze, G. I., and Privalov, P. L. (1995) Energetics of protein structure. *Adv. Protein Chem. 47*, 307–425.
51. Murphy, K. P., and Freire, E. (1992) Thermodynamics of structural stability and cooperative folding behavior in proteins. *Adv. Protein Chem. 43*, 313–361.
52. Sturtevant, J. M. (1977) Heat capacity and entropy changes in processes involving proteins. *Proc. Natl. Acad. Sci. U.S.A. 74*, 2236–2240.
53. Bergqvist, S., Williams, M. A., O'Brien, R., and Ladbury, J. E. (2004) Heat capacity effects of water molecules and ions at a protein-DNA interface. *J. Mol. Biol. 336*, 829–842.
54. Morton, C. J., and Ladbury, J. E. (1996) Water-mediated protein-DNA interactions: the relationship of thermodynamics to structural detail. *Protein Sci. 5*, 2115–2118.
55. Ladbury, J. E., Wright, J. G., Sturtevant, J. M., and Sigler, P. B. (1994) A thermodynamic study of the trp repressor-operator interaction. *J. Mol. Biol. 238*, 669–681.
56. Chaudhury, S., Sircar, A., Sivasubramanian, A., Berrondo, M., and Gray, J. J. (2007) Incorporating biochemical information and backbone flexibility in RosettaDock for CAPRI rounds 6−12. *Proteins 69*, 793–800.
57. Gray, J. J., Moughon, S., Wang, C., Schueler-Furman, O., Kuhlman, B., Rohl, C. A., and Baker, D. (2003) Protein-protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations. *J. Mol. Biol. 331*, 281–299.
58. Guerois, R., Nielsen, J. E., and Serrano, L. (2002) Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. *J. Mol. Biol. 320*, 369–387.
59. Benedix, A., Becker, C. M., de Groot, B. L., Caflisch, A., and Bockmann, R. A. (2009) Predicting free energy changes using structural ensembles. *Nat. Methods 6*, 3–4.
60. Wong, S., Amaro, R. E., and McCammon, J. A. (2009) MM-PBSA captures key role of intercalating water molecules at a protein-protein interface. *J. Chem. Theory Comput. 5*, 422–429.