Contribution of Domain Interface Residues to the Stability of Antibody C_H3 Domain Homodimers

William Dall'Acqua,[‡] Alexander L. Simon,[‡,§] Michael G. Mulkerrin,[||] and Paul Carter*[,‡]

Departments of Molecular Oncology and Medicinal & Analytical Chemistry, Genentech Inc., 1 DNA Way, South San Francisco, California 94080

Received February 3, 1998; Revised Manuscript Received April 22, 1998

ABSTRACT: Dimers of C_H3 domains from human IgG_1 were used to study the effect of mutations constructed at a domain−domain interface upon domain dissociation and unfolding, "complex stability". Alanine replacement mutants were constructed on one side of the interface for each of the sixteen interdomain contact residues by using a single-chain C_H3 dimer in which the carboxyl terminus of one domain was joined to the amino terminus of the second domain via a $(G_4S)_4$ linker. Single-chain variants were expressed in *Escherichia coli* grown in a fermentor and recovered in yields of 6−90 mg L^{-1} by immobilized metal affinity chromatography. Guanidine hydrochloride-induced denaturation was used to follow domain dissociation and unfolding. Surprisingly, the linker did not perturb the complex stability for either the wild type or two destabilizing mutants. The C_H3 domain dissociation and unfolding energetics are dominated by six contact residues where corresponding alanine mutations each destabilize the complex by >2.0 kcal mol^{-1}. Five of these residues (T366, L368, F405, Y407, and K409) form a patch at the center of the interface and are located on the two internal antiparallel β-strands. These energetically key residues are surrounded by 10 residues on the two external β-strands whose contribution to complex stability is small (three have a $\Delta\Delta G$ of 1.1−1.3 kcal mol^{-1}) or very small (seven have a $\Delta\Delta G$ of ≤0.7 kcal mol^{-1}). Thus, at the center of the C_H3 structural interface there is a small "functional interface" of residues that make significant contributions to complex stability.

Understanding protein−protein recognition is of fundamental significance as many biological processes depend on the ability of proteins to form specific and stable complexes. Protein−protein contact surfaces have been defined at atomic resolution by structure determination for a rapidly expanding number of protein complexes (reviewed in refs *1*−*4*). Extensive mutational analyses of such protein complexes have identified the residues that make significant energetic contributions to binding. In some cases, the energetically important residues represent only a small subset of the contact residues (*5*, *6*), whereas for others, a large proportion of the contact residues are functionally important (*7*).

The most detailed mutational analyses of protein−protein interactions to date have involved pairs of different proteins that can exist both free in solution and also in a bound state, such as ligand−receptor complexes (*5*, *6*) and antibody−antigen complexes (*7*), "noncompulsory complexes". In contrast, there is a paucity of data for the exceedingly common biological motif of "compulsory complexes" (*8*) in which free monomers are not readily detected. This likely reflects problems of interpreting the effects of mutations made simultaneously on both sides of an interface and the technical difficulty in making replacements on only one surface. For example, coexpression of mutant and wild-type chains typically leads to the formation of parental homodimers as well as the desired heterodimer, thereby confounding subsequent analysis of the mutational effects. Appending one of the chains with a charge tag (*9*, *10*) or each chain with either an epitope or polyhistidine tag (*11*) permits purification of the heterodimer away from the two parental homodimers. Nevertheless, subunit dissociation and exchange may lead to formation of parental homodimers that interfere with interpretation of mutational analysis.

An additional difficulty in functional analysis of domain interfaces of compulsory complexes is that mutations may affect unfolding of domains as well as their dissociation. Close coupling of these processes may confound partitioning of mutational effects between them. "Complex stability" here encompasses effects on both domain dissociation and unfolding of a compulsory complex.

We have extended the study of protein−protein recognition to investigate the effect of mutations constructed on one contact surface of a compulsory complex, the homodimer of the C_H3 domain (hdC_H3) from human IgG_1. The C_H3 homodimer was converted to a single-chain format (scC_H3) analogous to that originally described for single-chain Fv (scFv) fragments (*12*, *13*) by installing a peptide linker for connecting the carboxyl terminus of one domain to the amino terminus of the second domain. The linker was found not to significantly modify the stability of wild-type C_H3 or for two of the most destabilizing C_H3 mutants. This allowed us to adopt the scC_H3 format in analyzing the effects of

* Corresponding author. Telephone: (650) 225-1932. Fax: (650) 225-1716. E-mail: pjc@gene.com.

[‡] Department of Molecular Oncology.

[§] Present address: Cancer Biology, MSOB 309, Mail Code 5404, Stanford University, Stanford, CA 94305.

[||] Department of Medicinal & Analytical Chemistry.

S0006-2960(98)00270-0 CCC: $15.00
Published on Web 06/06/1998

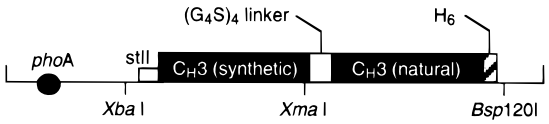

FIGURE 1: Schematic representation of an operon for the expression of scC_H3 in *E. coli* using the phagemid vector pRA4. A synthetic C_H3 gene (*19*) is connected to a natural C_H3 gene (*75*) via a linker encoding the sequence $(G_4S)_4$. The second C_H3 gene is followed by a sequence encoding six histidine residues (H_6) that allows purification by IMAC. The operon is under the transcriptional control of the *E. coli* alkaline phosphatase (*pho*A) promoter (*76*) which can be induced by phosphate starvation. Secretion is directed by the heat stable enterotoxin signal sequence [stII, (*77*)].

mutations installed on only one side of the domain interface upon domain dissociation and unfolding.

The C_H3 interface involves 16 residues located on four antiparallel β-strands that make intermolecular contacts and are buried 1090 Å^2 from each surface (*14*, *15*). These 16 residues were individually replaced with alanine on one side of the interface. Alanine replacements were chosen to minimize unfavorable steric contacts and to avoid imposing new charge interactions or hydrogen bonds from the substituted side chains while minimizing the risk of altering the main chain conformation (*16*−*18*). The contribution of individual residues to complex stability was assessed by comparing guanidine hydrochloride (GuHCl[1])-induced unfolding of wild-type and alanine mutant scC_H3 by far-ultraviolet circular dichroism (far-UV CD) spectroscopy.

MATERIALS AND METHODS

Reagents. Restriction enzymes, Vent DNA polymerase, and other DNA-modifying enzymes were purchased from New England Biolabs, Inc. (Beverly, MA), whereas Sequenase version 2.0 and Muta-Gene kits were obtained from United States Biochemicals (Cleveland, OH) and Bio-Rad (Hercules, CA), respectively. [^{35}S]dATPαS was obtained from Amersham Life Science Inc. (Arlington Heights, IL), and ultrapure guanidine hydrochloride (GuHCl) was purchased from Gibco-BRL. HiTrap nickel nitrilotriacetic acid (Ni^{2+}-NTA) and Superdex 75 columns were obtained from Pharmacia (Uppsala, Sweden). Oligonucleotides were synthesized on an Applied Biosystems (Foster City, CA) model 394 DNA synthesizer.

Single-Chain C_H3 Construction. Phagemid pRA2 contains a dicistronic operon in which C_H3 expressed from a synthetic gene is cosecreted with a second copy of C_H3 expressed from the natural gene as a fusion with M13 gene III (*19*). Phagemid pRA2 was modified to create pRA4 for scC_H3 expression by connecting two C_H3 domains via a $(G_4S)_4$ linker (Figure 1). In addition, the second C_H3 gene was mutagenized to replace M13 gene III with a sequence encoding six histidines to facilitate purification by immobilized metal affinity chromatography (IMAC; *20*). Briefly, the natural C_H3 gene (100 ng) was subjected to a two-step PCR mutagenesis using 50 pmol of the primer 5′-CAAACGAAGGGCCCTTATTAGTGATGGTGATGGTGATGTTTACCCGGAGACAGGGAGAGG-3′ (*Bsp*120I site underlined) with successively 50 pmol of the primer 5′-GGCGGAGGTGGTTCTGGCGGTGGCGGATCGGGGCAGCCCCGAGAACCACAGGTGTACACCCTGCCCCCATCTCGGGAAG-3′ and 50 pmol of the primer 5′-AAAAAGGGTATCTAGAGGTTGAGGTGATTTTTCCCTGTCTCCCGGGAAAGGTGGAGGCGGTTCAGGTGGAGGCGGTTCAGGCGGAGGTGGTT-3′ (underlined sites for *Xba*I and *Xma*I, respectively), each dNTP at 200 μM, 100 mM Tris-HCl (pH 8.3), 500 mM KCl, 1.5 mM $MgCl_2$, 0.01% (w/v) gelatin, and 3 units of Vent DNA polymerase. The sample was incubated at 94 °C for 5 min followed by 30 rounds of thermocycling (94 °C for 1 min, 50 °C for 1 min, and 72 °C for 1 min) using a GeneAmp PCR System 9600 (Perkin-Elmer, Norwalk, CT). The amplified DNA was digested with *Xba*I and *Bsp*120I and cloned into pRA2 (*19*) restricted with the same enzymes to create phagemid pRA3. The scC_H3 expression phagemid, pRA4, was constructed by subcloning the synthetic C_H3 gene as an *Xba*I−*Xma*I fragment into similarly cleaved pRA3. The sequence was verified by dideoxynucleotide sequencing (*21*) using Sequenase version 2.0.

Site-Directed Mutagenesis. Mutants were prepared by site-directed mutagenesis using the method of Kunkel (*22*) in conjunction with a Muta-Gene kit. The oligonucleotides for mutagenesis were designed to replace the chosen wild-type codons in the synthetic C_H3 gene with the alanine codon, GCT. Mutations were constructed using phagemid pRA2 (*19*) modified to contain either synthetic or natural C_H3 and verified by dideoxynucleotide sequencing.

Expression and Purification of C_H3 Variants. Phagemids for the expression of hdC_H3 and scC_H3 variants were transformed into *Escherichia coli* 33B6 (*23*). The C_H3 variants were secreted from corresponding *E. coli* cells cultured for approximately 48 h at 30 °C in an aerated 10 L fermentor (*24*). The cell density at harvest was in the range of 100−150 OD_{550}. The hdC_H3 and scC_H3 variants were purified as described previously (*19*) or below, respectively.

A periplasmic fraction was prepared (*19*) from 100 g of fermentation paste for each scC_H3 variant and dialyzed extensively against 50 mM Tris-HCl (pH 8.0). The dialysate was adjusted to 25 mM imidazole and applied to a 5 mL Ni^{2+}-NTA affinity column. The resin was washed with 20 column volumes of 50 mM Tris-HCl (pH 8.0), and the scC_H3 variants were eluted with a linear gradient of 0 to 500 mM imidazole in 50 mM Tris-HCl (pH 7.5). The purified scC_H3 variants were dialyzed against 50 mM sodium phosphate buffer (pH 7.4) or 10 mM Tris-HCl (pH 8.0) and 10 mM NaCl. The scC_H3 variants were analyzed by size exclusion chromatography on a Superdex 75 column in phosphate-buffered saline.

The purified hdC_H3 and scC_H3 variants were concentrated using Centriprep-10 concentrators (Amicon, Beverley, MA), flash-frozen in liquid nitrogen, and stored at −70 °C. The protein yields were estimated from the absorbance at 280 nm using the extinction coefficients (ϵ_{280}) for the wild-type hdC_H3 and scC_H3 determined by amino acid hydrolysis to be 32 204 and 38 107 cm^{-1} M^{-1}, respectively.

Far-UV CD Spectroscopy. Data were obtained as described by Zhukovsky et al. (*25*) on an AVIV (Lakewood, NJ) 60DS spectropolarimeter. Briefly, spectra for scC_H3 and hdC_H3 variants (25−80 μM) were recorded at 20.0 ± 0.1

[1] Abbreviations: Flk-1, fetal liver kinase 1; GuHCl, guanidine hydrochloride; HEL, hen egg white lysozyme; hdC_H3, homodimer C_H3; IMAC, immobilized metal affinity chromatography; KDR, kinase insert domain-containing receptor; NTA, nitrilotriacetic acid; scC_H3, single-chain C_H3; UV CD, ultraviolet circular dichroism; VEGF, vascular endothelial growth factor; θ_{obs}, ellipticity.

°C in 10 mM Tris-HCl (pH 8.0) and 10 mM NaCl. Data were collected over the wavelength range of 250−200 nm in 0.25 nm intervals in a thermostated circular cuvette with a 0.1 cm path length. The final spectra represent the mean of five far-UV CD scans with an integration time of 2 s per data point. Results are reported as mean residue weight ellipticity (θ_{MRW}, in deg cm^2 dmol^{-1}) calculated using the appropriate mean residue weights for each hdC_H3 and scC_H3 variant.

Equilibrium Denaturation Assays. Guanidine hydrochloride-induced unfolding of hdC_H3 and scC_H3 variants was followed by far-UV CD spectroscopy at a wavelength of 225 nm. The C_H3 variants (25−30 samples at ∼1 mg mL^{-1}) were prepared with varying concentrations of GuHCl (0−8 M) in 50 mM sodium phosphate (pH 7.4) and equilibrated for 24 h at 20.0 ± 0.1 °C prior to data collection. Ellipticity θ_{obs} was measured at a wavelength of 225 nm in circular cuvettes with a 0.1 cm path length. A total of 120 data points were taken in a 2 min trace, and the mean was calculated. The final concentrations of GuHCl in each sample were determined by measurement of the refractive index (*26*).

The ellipticity data were fitted by a nonlinear least-squares method to both two-state (*27*) and three-state (*28*) models of unfolding using KaleidaGraph, version 3.0.8 (Synergy Software, PCS Inc., Reading, PA), taking into account the slopes of the pre- and post-transitional baselines. Virtually identical estimates were obtained using the linear extrapolation method (*29*), as already noted (*25*) (data not shown). Denaturation parameters that were estimated include ΔG_0, the free energy of unfolding in the absence of denaturant, and $D_{50\%}$, the denaturant concentration at which the apparent free energy change of unfolding, ΔG_U, is zero. In addition, estimates were made for *m*, the slope of ΔG_U versus the denaturant concentration *D*. The errors associated with $D_{50\%}$ values are commonly less significant than those associated with ΔG_0 values which are extrapolated to 0 M GuHCl (*30*, *31*). Consequently, the difference in the free energy of domain dissociation and unfolding between the wild type and the mutant, $\Delta\Delta G$, is more accurately determined in the transition region, using the following equation:

$$\Delta\Delta G = \langle m \rangle (D_{50\%\ wt} - D_{50\%\ mut}) \quad (1)$$

where $\langle m \rangle$ is the mean value of *m* for all of the variants (*25*, *32*). A one-way analysis of variance (ANOVA) was consistent with *m* being independent of variant except for the F405A/F405′A[2] mutant in hdC_H3 and scC_H3 formats. The F405A/F405′A mutant was therefore excluded in the calculation of $\langle m \rangle$ (*25*).

Mass Spectrometry. C_H3 variants in 50 mM sodium phosphate buffer (pH 7.4) were mass analyzed with a packed capillary liquid chromatography system coupled directly to a PE-Sciex API 3 triple-quadrupole mass spectrometer as described (*33*).

[2] Mutations in C_H3 domains are denoted by the amino acid residue and number [Eu numbering scheme of Kabat et al. (*78*)], followed by the replacement amino acid. Mutations constructed in a pair of C_H3 domains are separated by a slash with a prime sign used to indicate the residue on the second domain. For example, F405A and F405A/F405′A represent replacements of the phenylalanine at position 405 with alanine on one or both domains, respectively.

Table 1: Expression, Purification, and Mass Analysis of C_H3 Variants

C_H3 variant	yield (mg L^{-1})	observed mass (Da)	expected mass (Da)[a]
homodimers			
wild-type	310	12 106 ± 2	12 107
F405A/F405′A	20	12 032 ± 1	12 031
Y407A/Y407′A	35	12 015 ± 1	12 015
single chains			
wild-type	18	26 281 ± 1	26 281
Q347A	61	26 222 ± 2	26 224
Y349A	36	26 190 ± 1	26 189
T350A	81	26 252 ± 1	26 251
L351A	41	26 239 ± 1	26 239
T366A	75	26 249 ± 2	26 251
L368A	90	26 239 ± 1	26 239
K370A	72	26 224 ± 1	26 224
K392A	9	26 222 ± 2	26 224
T394A	41	26 249 ± 2	26 251
P395A	52	26 252 ± 3	26 255
V397A	65	26 252 ± 1	26 253
L398A	6	26 240 ± 2	26 239
D399A	10	26 237 ± 1	26 237
F405A	14	26 204 ± 1	26 205
F405A/F405′A	8	26 129 ± 1	26 129
Y407A	37	26 189 ± 2	26 189
Y407A/Y407′A	57	26 097 ± 3	26 097
K409A	8	26 223 ± 2	26 224

[a] By high-resolution mass spectrometry.

RESULTS

Design of scC_H3 and Mutational Analysis Strategy. A single-chain form of C_H3 was designed by utilizing a flexible linker to span from K453 at the carboxyl terminus of C_H3 to G341 at the amino terminus of the partner C_H3 domain. Unfortunately, this distance could not be accurately estimated since the 10 carboxyl-terminal residues, S444−K453, are not defined in the original 2.8 Å X-ray crystallographic structure of human IgG_1 Fc (*14*) or in a more recent higher-resolution (2.0 Å) structure (M. Ultsch and A. M. de Vos, personal communication). L443, the carboxyl-terminal structurally defined residue, is approximately 37 Å from residue G341 of the partner C_H3 domain. scC_H3 domains were constructed with a long linker, $(G_4S)_4$, in an attempt to accommodate the uncertainty in the distance to be spanned. The second copy of C_H3 was followed by a sequence encoding six histidines to permit purification by IMAC (Figure 1).

The 16 domain interface residues were individually replaced with alanine on one side of the interface, the amino-terminal C_H3 domain of scC_H3. Mutants with replacements on both sides of the interface, namely, F405A/F405′A and Y407A/Y407′A, together with the wild type were constructed in both hdC_H3 and scC_H3 formats to assess the effect of the linker upon complex stability.

Expression and Purification of hdC_H3 and scC_H3 Variants. All C_H3 variants were secreted from *E. coli* grown to high cell density in a fermentor. The scC_H3 variants were purified by IMAC and eluted as a single peak at ∼150 mM imidazole (not shown), whereas the hdC_H3 variants were recovered by chromatography using ion exchange and ABx resins as described previously (*19*). Sufficient protein was recovered for all C_H3 variants to support further analysis, although the recovered yield varied from 35−310 to 6−90 mg L^{-1} for hdC_H3 and scC_H3 variants, respectively (Table 1). The molecular masses of all C_H3 variants determined by high-

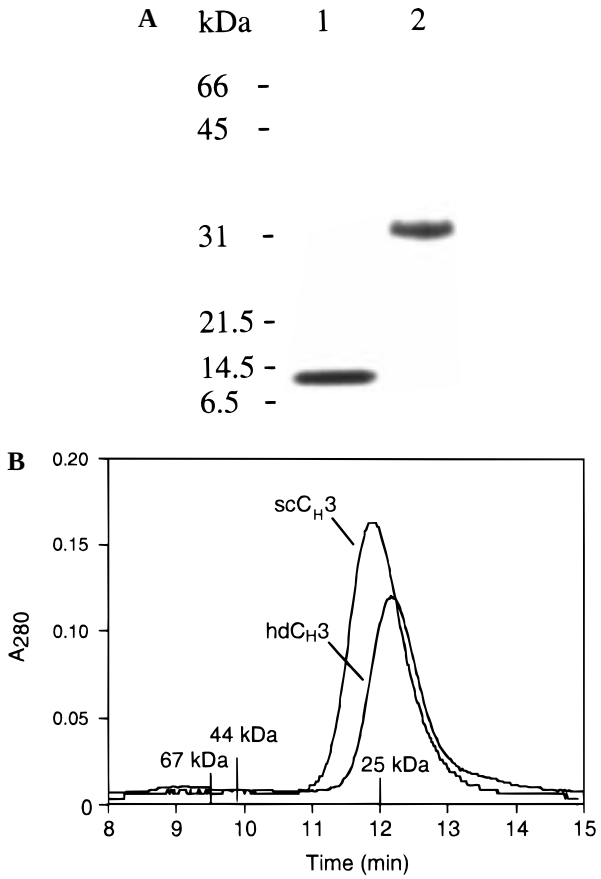

FIGURE 2: Purification of wild-type hdC_H3 and scC_H3 variants. (A) SDS−PAGE analysis using 4 to 20% acrylamide (Novex, San Diego, CA) after staining with Coomassie Blue: lane 1, purified wild-type scC_H3; and lane 2, purified wild-type hdC_H3. (B) Size exclusion chromatography of wild-type hdC_H3 and scC_H3 variants. The molecular mass standards elute as follows: bovine chymotrypsinogen A, 25 kDa, 12.0 min; chicken ovalbumin, 44 kDa, 9.9 min; and bovine serum albumin, 67 kDa, 9.4 min.

resolution electrospray mass spectrometry are consistent with those expected (Table 1).

Wild-type scC_H3 and hdC_H3 each migrate as a single major band when analyzed by SDS−PAGE with apparent molecular masses of 31 and 12 kDa, respectively (Figure 2A). Wild-type scC_H3 has an electrophoretic mobility lower than expected from its molecular mass (27 kDa), which may reflect the presence of the hexahistidine tag. Similar results were obtained with all other scC_H3 and hdC_H3 variants (not shown). Size exclusion chromatographs of wild-type (Figure 2B) and mutant scC_H3 (not shown) are consistent with the presence of the anticipated ~25 kDa monomer with no evidence for dimers or higher-order polymers.

Far-UV CD Spectroscopy and Equilibrium Denaturation Assays for C_H3 Variants. The wild-type scC_H3 fragment gave rise to a far-UV CD spectrum that is indistinguishable from that of the corresponding hdC_H3 (Figure 3), including a minimum at a wavelength of 225 nm characteristic of β-sheet structures (*34*). This wavelength was chosen for following the GuHCl-induced unfolding and dissociation of C_H3 variants. Unfolding and dissociation is fully reversible as judged by the complete recovery of original ellipticity upon dilution of the denatured samples as described (*25*, *35*). In contrast, thermal unfolding of hdC_H3 is not fully reversible and aggregation occurs at elevated temperatures (*19*).

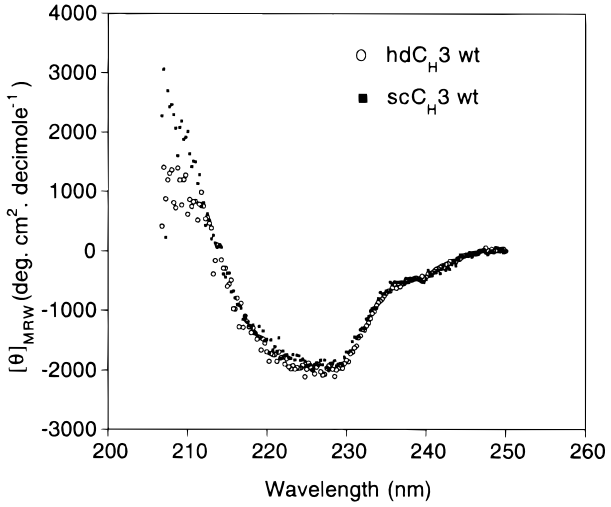

FIGURE 3: Far-UV CD spectra of wild-type scC_H3 and hdC_H3 variants recorded at 20.0 ± 0.1 °C in 10 mM Tris-HCl (pH 8.0) and 10 mM NaCl.

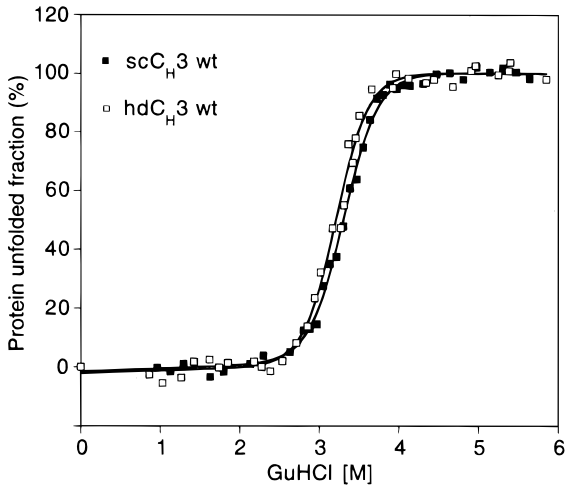

FIGURE 4: Equilibration denaturation profiles of wild-type scC_H3 and hdC_H3 recorded in 50 mM sodium phosphate (pH 7.4) at 20.0 ± 0.1 °C in the presence of varying amounts of GuHCl. Data were fitted using a nonlinear least-squares method to a two-state model (*27*).

GuHCl-induced denaturation data were analyzed by the nonlinear least-squares method of Santoro and Bolen (*27*) to estimate ΔG_0. All data were found to fit better to two-state than to three-state models as judged by lower χ^2 values (not shown). In addition, the transitions observed were continuous, an indication of rapid equilibration between two states (*27*) consistent with close coupling between dissociation of C_H3 domains and their unfolding.

Validation of the Use of the Linker. Wild-type C_H3 in single-chain and homodimer formats show identical stability profiles (Figure 4) and fitted parameters (Table 2) following GuHCl-induced denaturation, indicating that the linker does not significantly modify complex stability. Variants F405A/F405′A and Y407A/Y407′A containing destabilizing mutations on both sides of the C_H3 interface also gave very similar denaturation profiles and parameters in hdC_H3 and scC_H3 formats (Table 2). Taken together, these data support the use of the scC_H3 format in analyzing the effect of alanine scanning mutations constructed on one side of the C_H3 domain interface.

Table 2: Equilibrium Denaturation Parameters for hdC_H3 and scC_H3 Variants[a]

C_H3 variant	format	ΔG_0[b] (kcal mol^{-1})	$D_{50\%}$[c] (M)	m[d] (kcal mol^{-2} L)	$\Delta\Delta G$[e] (kcal mol^{-1})
wild-type	hd	9.1 ± 0.4	3.19 ± 0.14	2.92 ± 0.24	(0)
wild-type	sc	9.1 ± 0.3	3.25 ± 0.08	2.80 ± 0.06	(0)
F405A/F405′A	hd	1.1 ± 0.2	1.20 ± 0.02	0.92 ± 0.22	6.4
F405A/F405′A	sc	1.2 ± 0.3	1.05 ± 0.17	1.15 ± 0.07	6.1
Y407A/Y407′A	hd	4.9 ± 0.4	1.66 ± 0.08	2.81 ± 0.06	4.0
Y407A/Y407′A	sc	4.3 ± 0.4	1.74 ± 0.15	2.52 ± 0.11	4.0

[a] Estimates of the denaturation parameters were performed in duplicate or triplicate. Errors shown are twice the standard deviation of the measurements. [b] ΔG_0, free energy of unfolding in the absence of the denaturant estimated by nonlinear least-squares fitting of the raw data to a two-state model for unfolding. [c] $D_{50\%}$, denaturant concentration at which the apparent free energy of unfolding $\Delta G_u = 0$ kcal mol^{-1}. [d] m, slope of ΔG_u vs D. [e] $\Delta\Delta G$, difference in the free energy of unfolding between wild-type scC_H3 and hdC_H3 and variants in the same format. $\Delta\Delta G$ values were calculated from eq 1 using 2.65 kcal mol^{-1} (SD of 0.18 kcal mol^{-1}), the mean value of m for all proteins, excluding hdC_H3 F405A/F405′A and scC_H3 F405A/F405′A. The standard deviation for $\Delta\Delta G$ values is 0.1−0.15 kcal mol^{-1} except for that for hdC_H3 F405A/F405′A (0.25 kcal mol^{-1}).

Alanine Scanning Mutagenesis of the C_H3 Domain Interface. The largest reductions in complex stability ($\Delta\Delta G >$ 2.0 kcal mol^{-1}) occurred on alanine substitution of residues T366, L368, P395, F405, Y407, and K409. Significant destabilizing effects ($\Delta\Delta G = 1.0-1.3$ kcal mol^{-1}) were also seen on substitution of residues Q347, L351, and K370. Alanine replacements of the seven other interface residues (Y349, T350, K392, T394, V397, L398, and D399) had small or insignificant effects upon complex stability ($\Delta\Delta G \leq 0.7$ kcal mol^{-1}). These data from alanine scanning mutagenesis indicate that residues in the two internal antiparallel β-strands contribute significantly more to complex stability than residues in the two external antiparallel β-strands (Figure 5).

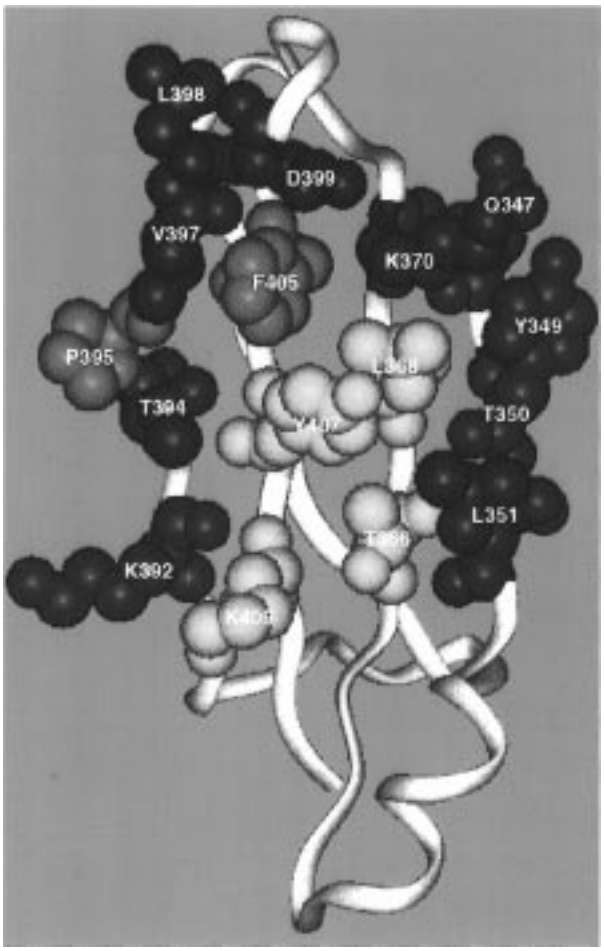

FIGURE 5: Space-filling model of the surface of C_H3 in contact in the C_H3−C_H3 complex based upon the human IgG_1 structure of Deisenhofer (*14*). Residues are color-coded according to the loss of free energy of stabilization upon alanine substitution: red, ≥2.5 kcal mol^{-1}; yellow, 2.0−2.4 kcal mol^{-1}; and blue, <1.5 kcal mol^{-1}.

DISCUSSION

Single-Chain Strategy. A mutational analysis of the human IgG_1 C_H3 domain was undertaken to quantify the contribution of interdomain contact residues to domain dissociation and unfolding of this compulsory complex. A scC_H3 dimer was constructed by joining the carboxyl terminus of one domain to the amino terminus of the second domain via a 20-residue linker. This permitted mutations to be constructed on one side of the interface, thereby avoiding the complication of attempting to interpret mutations constructed simultaneously on both sides of the interface.

The linker was anticipated to stabilize the C_H3 dimer by virtue of increasing the local concentration of monomers. Such stabilization has been previously demonstrated for analogous linkers in a scFv fragment (*36*) and a single-chain version of the homodimeric bacteriophage P22 Arc repressor (*37*). Surprisingly, the linker did not significantly affect the free energy change upon domain dissociation and unfolding for wild-type C_H3 and two destabilized mutants, F405A/F405′A and Y407A/Y407′A (Table 2). The $(G_4S)_4$ linker may be sufficiently long and flexible that the increased stability anticipated from a high local concentration of monomers (*37*) is lost. Alternatively, the concentration of C_H3 variants (1 mg mL^{-1}) used for denaturation experiments may be too high to observe the effects due to the covalent linkage. Wild-type C_H3 in homodimeric and single-chain formats exhibits indistinguishable CD spectra (Figure 3), indicating no gross perturbation of the secondary structure (*38*) or of the β-sheets packing (*39*, *40*). Taken together, these data suggest that scC_H3 is a good substitute for heterodimeric C_H3 for analyzing the effect of mutations on only one side of the domain interface.

No free monomer has been detected in C_H3 (pFc′) preparations at concentrations as low as 10^{-10} M (*41*). This is consistent with C_H3 dimers having a dissociation constant (K_D) of $<10^{-10}$ M (*42*) corresponding to a free energy change upon dissociation of >13 kcal mol^{-1} calculated using the well-known relationship

$$\Delta G = RT \ln K_D \qquad (2)$$

where R is the gas constant and T the absolute temperature (293 K). In contrast, the free energy change upon hdC_H3 dissociation and unfolding was estimated to be 9.1 kcal mol^{-1} (Table 2). These data are consistent with the C_H3 domains being stabilized by their association.

A concern with the interdomain linker is that it might promote C_H3 dimerization or multimerization as shown with single-chain Fv fragments, particularly with linkers of 10 or fewer residues (*43*−*45*). The scC_H3 constructed with a 20-residue linker is monomeric as judged by size exclusion chromatography, with no evidence for multimerization (Figure 2B). In contrast, size exclusion chromatography of scC_H3 constructed with a shorter linker, $(G_4S)_3$, is consistent with a monomer−dimer equilibrium (data not shown).

Interpretation of the Denaturation Parameters. Stabilization of C_H3 dimers is largely mediated by six of the 16 residues tested: T366, L368, P395, F405, Y407, and K409 (Figure 5). Residues T366 and Y407 form a hydrogen bond and represent the principal intermolecular and intramolecular contact with each other. Alanine substitution of Y407 and T366 reduces the complex stability by 2.3 and 2.1 kcal mol^{-1}, respectively. This loss in stability is consistent with the 1.0−2.5 kcal mol^{-1} reduction in free energy seen on replacing hydrogen-bonding residues in other proteins where both donor and acceptor residues are uncharged (*46*−*49*). Unfortunately, the exact strength of hydrogen bonds cannot be accurately determined from such single-mutation data since other interactions (van der Waals and hydrophobic) that may contribute to stabilization of the interface are concomitantly perturbed (*32*, *50*, *51*). Furthermore, the introduction of cavity-creating mutations might be followed by compensatory structural rearrangements (*52*−*54*) or solvent reorganization (*55*). A more accurate estimation of strength of the hydrogen bond between residues T366 and Y407′ awaits the construction of additional mutations and structural analysis of the corresponding mutant proteins.

Residue K409 forms a hydrogen bond with residue D399 on the partner C_H3 domain (*14*, *15*). Truncation of the K409 side chain to alanine significantly decreases complex stability ($\Delta\Delta G$ = 2.4 kcal mol^{-1}), whereas pruning the D399 side chain results in only a small reduction in stability ($\Delta\Delta G$ = 0.7 kcal mol^{-1}) (Table 3). There are numerous possible explanations for the asymmetry of these mutational effects at residues K409 and D399, including the possibility that these residues are not the major energetic contacts with each other. For example, the near functional silence of D399 may reflect that the fact that the cost of desolvating this residue is comparable to the energy gained through intermolecular contact with K409. The release of water molecules is not a major contributor to the stabilization of some protein−protein complexes (*56*, *57*). Alternatively, the larger effect observed for substitution of residue K409 than for D399, its hydrogen bond partner, can be explained by an effect of K409 on folding of the corresponding domain.

Residues L368 and F405 form intermolecular and intramolecular contacts through van der Waals interactions only. The large decrease in stability upon substitution of these residues with alanine likely reflects the contribution of the hydrophobic effect and of van der Waals contacts as well as probable solvent and protein structure rearrangements. The strength of the van der Waals contacts has been estimated to be approximately 0.5 kcal mol^{-1} (*49*, *55*).

Table 3: Equilibrium Denaturation Parameters for scC_H3 Alanine Variants[a]

C_H3 variant	ΔG_0 (kcal mol^{-1})	$D_{50\%}$ (M)	m (kcal mol^{-2} L)	$\Delta\Delta G^b$ (kcal mol^{-1})
Q347A	8.4 ± 0.4	2.86 ± 0.06	2.93 ± 0.20	1.1
Y349A	8.5 ± 0.2	3.00 ± 0.01	2.83 ± 0.06	0.7
T350A	8.8 ± 0.1	3.21 ± 0.01	2.76 ± 0.06	0.1
L351A	7.0 ± 0.4	2.76 ± 0.02	2.52 ± 0.18	1.3
T366A	6.3 ± 0.2	2.38 ± 0.02	2.65 ± 0.11	2.3
L368A	6.1 ± 0.3	2.40 ± 0.04	2.52 ± 0.20	2.2
K370A	6.9 ± 0.3	2.83 ± 0.08	2.45 ± 0.12	1.1
K392A	8.5 ± 0.3	3.11 ± 0.19	2.75 ± 0.04	0.4
T394A	8.1 ± 0.1	3.02 ± 0.01	2.67 ± 0.04	0.6
P395A	5.3 ± 0.4	2.08 ± 0.03	2.62 ± 0.11	3.2
V397A	7.7 ± 0.3	3.02 ± 0.01	2.57 ± 0.09	0.6
L398A	8.2 ± 0.3	3.23 ± 0.04	2.55 ± 0.08	0.1
D399A	8.2 ± 0.3	2.99 ± 0.02	2.72 ± 0.06	0.7
F405A	5.8 ± 0.1	2.32 ± 0.03	2.50 ± 0.07	2.5
Y407A	6.6 ± 0.4	2.49 ± 0.02	2.65 ± 0.16	2.1
K409A	5.2 ± 0.1	2.37 ± 0.11	2.21 ± 0.04	2.4

[a] Details as described in Table 2. [b] The standard deviation for $\Delta\Delta G$ values was 0.1−0.15 kcal mol^{-1} except for that for K392A (0.25 kcal mol^{-1}).

The most destabilizing mutation ($\Delta\Delta G$ = 3.2 kcal mol^{-1}), P395A, is located in a peripheral β-strand. Replacement of this proline residue likely perturbs the local backbone conformation, thereby indirectly affecting intermolecular interactions involving neighboring residues. Cis to trans isomerization of the peptide bond at P395 is not a plausible hypothesis for explaining the effects of the P395A replacement since this residue in wild-type C_H3 is in the trans rather than in the cis conformation (M. Ultsch and A. M. de Vos, personal communication). Substitution of proline residues in proteins with smaller amino acids is believed to be destabilizing through local conformational effects (*58*−*60*) or through stabilization of the unfolded state (*61*). A proline to serine substitution in nucleoside diphosphate kinase compromised the ability of the isolated subunits to associate into a hexamer (*60*). This effect was attributed to a change in the flexibility of the proline-containing loop constituting the domain−domain interface.

The m parameter reflects the difference between the solvent-exposed nonpolar surface area of the denatured and native states of proteins and/or the cooperativity of protein unfolding (*29*, *62*). Thus, the significant change in m for the F405A/F405′A double mutant may reflect the presence of equilibrium intermediates during unfolding or changes in the solvation properties of the denatured state.

Additivity of Mutational Effects. The reduction in stability for the double mutant, Y407A/Y407′A scC_H3 ($\Delta\Delta G$ = 4.0 kcal mol^{-1}), approximately equals the sum of stability reductions of the constituent single mutants Y407A and Y407′A ($\Delta\Delta G$ = 2.1 kcal mol^{-1}). This functional additivity (*63*) is consistent with Y407A and Y407′A having only local effects that are independent of each other. Residues Y407 and Y407′ are ∼5 Å apart at the interface and do not form any intermolecular contacts (*14*). In contrast, the double mutant F405A/F405′A scC_H3 is destabilized by 1.1 kcal mol^{-1} more than predicted by adding the $\Delta\Delta G$ terms for the component single mutants. Thus, residues F405 and F405′ are energetically coupled although separated by >12 Å. Nonadditivity of distant mutations has been observed in several other systems, including barnase−barstar (*64*), sta-

phylococcal nuclease (*65*), human hemoglobin (*66*), and an antibody−anti-idiotype (*49*). Several possible explanations for long distance nonadditivity have been proposed (*67*), including solvent rearrangements, as described at the domain interface for a complex between a mutant antibody and chicken lysozyme (*55*, *68*).

Relationship between Structural and Functional Interface. A priori, the binding interaction between two proteins might be mediated primarily by a few strong interactions involving a small subset of interface residues or through a larger number of contact residues each making small contributions. In the case of the scC_H3 compulsory complex, six of the 16 interface residues account for the majority of the energy of stabilization. The residues most critical for scC_H3 stabilization are located on the two interior β-strands, forming a patch at the center of the interface (Figure 5). P395 notwithstanding, residues at the periphery make very little net contribution to the free energy of scC_H3 stabilization. Thus, scC_H3 is stabilized by an "energetic interface" between C_H3 domains located at the heart of a larger "structural interface".

Small energetic interfaces have been observed previously for the noncompulsory complexes between human growth hormone and its receptor (*5*, *69*) and for vascular endothelial growth factor (VEGF) and its KDR and Flk-1 receptors (*6*). A few strong interactions at the center of an interface also dominate the energetics of binding of the complex between hen egg white lysozyme and the monoclonal antibody D1.3 (*7*). In contrast, the majority of the contact residues play a significant role in binding (>1.5 kcal mol^{-1}) for an antibody−anti-idiotype complex (*7*).

Nature of the Functional Interface. Mutational analysis of the compulsory complex of scC_H3 indicates that residues making major contributions to domain stability and/or association have side chains that are hydrophobic (F405 and L368), polar (T366 and Y407), or charged (K409). Similarly, both polar and nonpolar residues play an important energetic role in association of the noncompulsory complexes of chicken lysozyme with a cognate monoclonal antibody (*7*) and of VEGF with its KDR receptor (*6*). In contrast, the energetics of human growth hormone associating with its receptor relies primarily upon hydrophobic residues (*5*). Thus, there is much variation in the chemical nature of the energetic interface residues among compulsory and noncompulsory complexes studied to date.

The free energy of the C_H3-C_H3 complex stabilization appears to arise from only half of the antiparallel β-strand constituting the interface. Several different models of β-sheet packing have been described, including aligned (*70*), three-layer (*71*), and intercalated (*15*). Whether the main features of the energetic map we describe here are to be generalized for those different kinds of β-sheets-involving interfaces is still to be determined. Since the interior of many monomeric proteins is formed by β-sheet packing, the quantitative information we gained from this study is also highly relevant to the issue of protein stabilization.

Summary. We have shown that approximately one-third of the contact side chains at the human IgG1 C_H3 domain interface can account for the majority of contributions to domain folding and association. These residues cluster near the center of the structural interface. This study will help us to progress from purely anatomical descriptions of protein−protein interfaces such as their sizes, the identity of contact residues, the amount of buried surface areas, the number of hydrogen bonds, and van der Waals interactions (*72*−*74*) to a detailed understanding of how these structural features contribute to the affinity and specificity of the binding reactions. To our knowledge, this is the first extensive mutational analysis of the domain interface of a compulsory protein complex.

ACKNOWLEDGMENT

We thank Bart de Vos and Mark Ultsch for sharing unpublished X-ray crystallographic coordinates of a human IgG_1 Fc, Mark Vasser, Parkash Jhurani, Peter Ng, and Kristina Azizian for synthetic oligonucleotides, Brad Snedecor and Han Chen for *E. coli* fermentations, Jim Bourell for mass spectrometry of the proteins, Allan Padua for amino acid hydrolysis analysis, and Bob Kelley for helpful discussions and review of the manuscript.

REFERENCES

1. Davies, D. R., Padlan, E., and Sheriff, S. (1990) *Annu. Rev. Biochem. 59*, 439−473.
2. Mariuzza, R. A., and Poljak, R. J. (1993) *Curr. Opin. Immunol. 5*, 50−55.
3. Webster, D. M., Henry, A. H., and Rees, A. R. (1994) *Curr. Opin. Struct. Biol. 4*, 123−129.
4. Braden, B. C., and Poljak, R. J. (1995) *FASEB J. 9*, 9−16.
5. Clackson, T., and Wells, J. A. (1995) *Science 267*, 383−386.
6. Muller, Y. A., Li, B., Christinger, H. W., Wells, J. A., Cunningham, B. C., and de Vos, A. M. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 7192−7197.
7. Dall'Acqua, W., Goldman, E. R., Eisenstein, E., and Mariuzza, R. A. (1996) *Biochemistry 35*, 9667−9676.
8. Goodsell, D. S., and Olson, A. J. (1993) *Trends Biochem. Sci. 18*, 65−68.
9. Deonarain, M. P., Scrutton, N. S., and Perham, R. N. (1992) *Biochemistry 31*, 1491−1497.
10. Deonarain, M. P., Scrutton, N. S., and Perham, R. N. (1992) *Biochemistry 31*, 1498−1504.
11. Elsevier, J. P., Wells, L., Quimby, B. B., and Fridovich-Keil, J. L. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 7166−7171.
12. Bird, R. E., Hardman, K. D., Jacobson, J. W., Johnson, S., Kaufman, B. M., Lee, S. M., Lee, T., Pope, S. H., Riordan, G. S., and Whitlow, M. (1988) *Science 242*, 423−426.
13. Huston, J. S., Levinson, D., Mudgett-Hunter, M., Tai, M. S., Novotny, J., Margolies, M. N., Ridge, R. J., Bruccoleri, R. E., Haber, E., and Crea, R. (1988*) Proc. Natl. Acad. Sci. U.S.A. 85*, 5879−5883.
14. Deisenhofer, J. (1981) *Biochemistry 20*, 2361−2370.
15. Miller, S. (1990) *J. Mol. Biol. 216*, 965−973.
16. Carter, P., and Wells, J. A. (1988) *Nature 332*, 564−568.
17. Duncan, A. R., and Winter, G. (1988) *Nature 332*, 563−564.
18. Cunningham, B. C., and Wells, J. A. (1989) *Science 244*, 1081−1085.
19. Atwell, S., Ridgway, J. B. B., Wells, J. A., and Carter, P. (1997) *J. Mol. Biol. 270*, 26−35.
20. Hochuli, E. (1988) *J. Chromatogr. 444*, 293−302.
21. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463−5467.
22. Kunkel, T. A., Roberts, J. D., and Zakour, R. A. (1987) *Methods Enzymol. 154*, 367−382.
23. Rodrigues, M. L., Presta, L. G., Kotts, C. E., Wirth, C., Mordenti, J., Osaka, G., Wong, W. L. T., Nuijens, A., Blackburn, B., and Carter, P. (1995) *Cancer Res. 55*, 63−70.
24. Carter, P., Kelley, R. F., Rodrigues, M. L., Snedecor, B., Covarrubias, M., Velligan, M. D., Wong, W. L. T., Rowland, A. M., Kotts, C. E., Carver, M. E., Yang, M., Bourell, J. H., Shephard, H. M., and Henner, D. (1992) *Bio/Technology 10*, 163−168.
25. Zhukovsky, E. A., Mulkerrin, M. G., and Presta, L. G. (1994) *Biochemistry 33*, 9856−9864.

26. Nozaki, Y. (1972) *Methods Enzymol. 26*, 43−50.
27. Santoro, M. M., and Bolen, D. W. (1988) *Biochemistry 27*, 8063−8068.
28. Barrick, D., and Baldwin, R. L. (1993) *Biochemistry 32*, 3790−3796.
29. Greene, R. F., and Pace, C. N. (1974) *J. Biol. Chem. 249*, 5388−5393.
30. Kellis, J. T., Jr., Nyberg, K., and Fersht, A. R. (1989) *Biochemistry 28*, 4914−4922.
31. Horovitz, A., Matthews, J. M., and Fersht, A. R. (1992) *J. Mol. Biol. 227*, 560−568.
32. Serrano, L., Horovitz, A., Avron, B., Bycroft, M., and Fersht, A. R. (1990) *Biochemistry 29*, 9343−9352.
33. Bourell, J. H., Clauser, K. P., Kelley, R., Carter, P., and Stults, J. T. (1994) *Anal. Chem. 66*, 2088−2095.
34. Venyaminov, S. Y., and Yang, J. S. (1996) in *Circular Dichroism and the Conformational Analysis of Biomolecules*, pp 69−107, Plenum Press, New York.
35. Missiakas, D., Betton, J. M., Minard, P., and Yon, J. M. (1990) *Biochemistry 29*, 8683−8689.
36. Glockshuber, R., Malia, M., Pfitzinger, I., and Pluckthun, A. A. (1990) *Biochemistry 29*, 1362−1367.
37. Robinson, C. R., and Sauer, R. T. (1996) *Biochemistry 35*, 13878−13884.
38. Provencher, S. W., and Glockner, J. (1981) *Biochemistry 20*, 33−37.
39. Manning, M. C., and Woody, R. W. (1987) *Biopolymers 26*, 1731−1752.
40. Manning, M. C., Illangasekare, M., and Woody, R. W. (1988) *Biophys. Chem. 31*, 77−86.
41. Ellerson, J. R., Yasmeen, D., Painter, R. H., and Dorrington, K. J. (1976) *J. Immunol. 116*, 510−517.
42. Schiffer, M., Chang, C. H., Naik, V. M., and Stevens, F. J. (1988) *J. Mol. Biol. 203*, 799−802.
43. Holliger, P., Prospero, T., and Winter, G. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 6444−6448.
44. Whitlow, M., Filpula, D., Rollence, M. L., Feng, S. L., and Wood, J. F. (1994) *Protein Eng. 7*, 1017−1026.
45. Desplancq, D., King, D. J., Lawson, A. D. G., and Mountain, A. (1994) *Protein Eng. 7*, 1027−1033.
46. Fersht, A. R., Shi, J., Knill-Jones, J., Lowe, D. M., Wilkinson, A. J., Blow, D. M., Brick, P., Carter, P., Waye, M. M. Y., and Winter, G. (1985) *Nature 314*, 235−238.
47. Street, I. P., Armstrong, C. R., and Withers, S. G. (1986) *Biochemistry 25*, 6021−6027.
48. Shirley, B. A., Stanssens, P., Hahn, U., and Pace, C. N. (1992) *Biochemistry 31*, 725−732.
49. Goldman, E. R., Dall'Acqua, W., Braden, B. C., and Mariuzza, R. A. (1997) *Biochemistry 36*, 49−56.
50. Ackers, G. K., and Smith, F. R. (1985) *Annu. Rev. Biochem. 54*, 597−629.
51. Fersht, A. R. (1988) *Biochemistry 27*, 1577−1580.
52. Eriksson, A. E., Baase, W. A., Zhang, X. J., Heinz, D. W., Blaber, M., Baldwin, E. P., and Matthews, B. W. (1992) *Science 255*, 178−183.
53. Rashin, A. A., Rashin, B. H., Rashin, A., and Abagyan, R. (1997) *Protein Sci. 6*, 2143−2158.
54. Atwell, S., Ultsch, M., de Vos, A. M., and Wells, J. A. (1997) *Science 278*, 1125−1128.
55. Ysern, X., Fields, B. A., Bhat, T. N., Goldbaum, F. A., Dall'Acqua, W., Schwartz, F. P., Poljak, R. J., and Mariuzza, R. A. (1994) *J. Mol. Biol. 238*, 496−500.
56. Ross, P. D., and Subramanian, S. (1981) *Biochemistry 20*, 3096−3102.
57. Bhat, T. N., Bentley, G. A., Boulot, G., Greene, M. I., Tello, D., Dall'Acqua, W., Souchon, H., Schwartz, F. P., Mariuzza, R. A., and Poljak, R. J. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 1089−1093.
58. Koshy, T. I., Luntz, T. L., Schejter, A., and Margoliash, E. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 8697−8701.
59. Schejter, A., Luntz, T. L., Koshy, T. I., and Margoliash, E. (1992) *Biochemistry 31*, 8336−8343.
60. Lascu, I., Deville-Bonne, D., Glaser, P., and Veron, M. (1993) *J. Biol. Chem. 268*, 20268−20275.
61. Alber, T. (1989) *Annu. Rev. Biochem. 58*, 765−798.
62. Shortle, D., and Meeker, A. K. (1986) *Proteins: Struct., Funct., Genet. 1*, 81−89.
63. Wells, J. A. (1991) *Methods Enzymol. 202*, 390−414.
64. Schreiber, G., and Fersht, A. R. (1995) *J. Mol. Biol. 248*, 478−486.
65. Green, S. M., and Shortle, D. (1993) *Biochemistry 32*, 10131−10139.
66. Speros, P. C., LiCata, V. J., Yontani, T., and Ackers, G. K. (1991) *Biochemistry 30*, 7254−7262.
67. LiCata, V. J., and Ackers, G. K. (1995) *Biochemistry 34*, 3133−3139.
68. Fields, B. A., Goldbaum, F. A., Dall'Acqua, W., Malchiodi, E. L., Cauerhff, A., Schwartz, F. P., Ysern, X., Poljak, R. J., and Mariuzza, R. A. (1996) *Biochemistry 35*, 15494−15503.
69. Cunningham, B. C., and Wells, J. A. (1993) *J. Mol. Biol. 234*, 554−563.
70. Chothia, C., and Janin, J. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 4146−4150.
71. Chothia, C., Novotny, J., Bruccoleri, R., and Karplus, M. (1985) *J. Mol. Biol. 186*, 651−663.
72. Janin, J., and Chothia, C. (1990) *J. Biol. Chem. 265*, 16027−16030.
73. Wilson, L. A., and Stanfield, R. L. (1993) *Curr. Opin. Struct. Biol. 3*, 113−118.
74. Davies, D. R., and Cohen, G. H. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 7−12.
75. Ellison, J. W., Berson, B. J., and Hood, L. E. (1982) *Nucleic Acids Res. 10*, 4071−4079.
76. Chang, C. N., Kuang, W. J., and Chen, E. Y. (1986) *Gene 44*, 121−125.
77. Picken, R. N., Mazaitis, A. J., Maas, W. K., Rey, M., and Heyneker, H. (1983) *Infect. Immun. 42*, 269−275.
78. Kabat, E. A., Wu, T. T., Perry, H. M., Gottesman, K. S., and Foeller, C. (1991) *Sequences of Proteins of Immunological Interest*, 5th ed., Vol. 1, pp 688−696, National Institutes of Health, Bethesda, MD.

BI980270I