Conformational Stability, Folding, and Ligand-Binding Affinity of Single-Chain Fv Immunoglobulin Fragments Expressed in *Escherichia coli*†

Michael W. Pantoliano,*,‡ Robert E. Bird,§ Syd Johnson,∥ Eric D. Asel,‡ Steven W. Dodd,⊥ Jay F. Wood, and Karl D. Hardman‡

Genex Corporation, Gaithersburg, Maryland 20872

Received March 27, 1991; Revised Manuscript Received July 12, 1991

ABSTRACT: A fluorescein-binding single-chain Fv (scFv) was chosen as a model for the study of the physicochemical parameters associated with synthetic IgG fragments. Three such scFv proteins were designed from the primary sequences of one anti-fluorescyl monoclonal antibody (Mab 4.4.20). These were constructed with varying-length interdomain peptide linkers of between 12 and 25 residues, expressed in *Escherichia coli*, and the protein folding, stability, and antigen-binding characteristics were assessed. Efficient renaturation could be accomplished in vitro to yield ~26 mg of active scFv/L of fermentation. Scatchard analysis for fluorescein ligand binding revealed that the scFv designs come within 2-fold of the K_a = 1.99 (±0.18) × 10^9 observed for the parental 4.4.20 Fab and have identical stoichiometries (n ~0.99). Reversible solvent denaturation studies demonstrated that the unfolding/refolding equilibria for the scFv proteins can be fit to a simple two-state model and that two of the scFv designs were found to be slightly more stable than single IgG domains (V_L and C_L) when assessed in terms of the free energy of unfolding, $\Delta G°_{n-u}$, or nearly identical to other multiple domain immunoglobulin proteins such as light chains and Fab's when relative transition midpoints, C_m, are compared. Linkers which conferred conformational flexibility beyond the minimally required length of 12 residues were found to have a stabilizing effect. By these criteria of ligand-binding function and protein stability, the scFv proteins were found to be bona fide minimal replicas of their parental IgG molecules.

Several recent reports have described the genetic construction of antibody fragments such as Fab's[1] (Better et al., 1988; Horwitz et al., 1988), Fv's (Skerra & Plückthun, 1988; Reichmann et al., 1988), single-chain Fv's (Bird et al., 1988; Huston et al., 1988), and even single V_H domains (Ward et al., 1989) with their subsequent expression in *Escherichia coli*, yeast (Horwitz et al., 1988), myeloma cells (Reichmann et al., 1988). These proteins have been designed to mimic the binding functionality of the parental 12-domain IgG molecules (150 kDa) but attempt to retain the minimal amount of essential polypeptide chain by embracing one, two, or four domains of ~12.5 kDa each, depending on inclination.

The single-chain Fv (scFv) antibody fragment is a two-domain minimal strategy (see Figure 1) and is comprised of the V_L and V_H domains with a short peptide linker that connects the carboxy-terminal residue, L107 [numbering of Kabat et al. (1987)], with that of the amino terminus, H1, of the V_H domain (Bird et al., 1988). The thermodynamic stability of scFv proteins has not been addressed, to date, but will be of fundamental importance in assessing their utility as framework strategies for the delivery of drugs or diagnostic imaging applications (Chaudhary et al., 1989; Colcher et al., 1990; Condra et al., 1990). Moreover, nothing is known about the effect of different peptide linkers on thermodynamic stability/folding and ligand-binding affinity of scFv's or whether there is a preferred way of connecting these two domains.

A fluorescein- (FL-) binding single-chain Fv (scFv) was, therefore, chosen as a model for the exploration of the physicochemical parameters of scFv proteins and for the eventual incremental optimization (Matsumura et al., 1989a,b; Pantoliano et al., 1989) to expand the utility of synthetic IgG fragments in the medical and industrial arenas (Haber et al., 1989; Schultz et al., 1990; Winter & Milstein, 1991). Three scFv proteins were designed from the primary sequence of the anti-fluorescyl Mab 4.4.20 (Voss, 1984; Bedzyk et al., 1989) to contain varying-length interdomain peptide linkers of between 12 and 25 residues. The protein folding, conformational stability, and antigen-binding affinity of these new proteins were then investigated.[2]

Materials and Methods

Molecular Biology. The anti-fluorescyl scFv genes were expressed in *E. coli* from plasmids using the hybrid λ phage O_L/P_R promoter (Scandella et al., 1985) to transcribe the scFv gene as a fusion protein with the *E. coli* ompA signal sequence

† Presented in part at the Sixth Du Pont Immunology Symposium on Antibody Engineering, Wilmington, DE, Oct 1990.

* Corresponding author.

‡ Present address: Du Pont-Merck Pharmaceuticals Co., E328/B48B, Wilmington, DE 19880.

§ Present address: Molecular Oncology Inc., Gaithersburg, MD 20878.

∥ Present address: MedImmune Inc., Gaithersburg, MD 20878.

⊥ Present address: Eli Lilly, Indianapolis, IN 46285.

[1] Abbreviations: D, denaturant; EDTA, disodium salt of ethylenediaminetetraacetic acid; Fab, antibody fragment obtained after papain digestion and comprised of V_L, V_H, C_L, and C_H1 IgG domains; F, fluorescence intensity with F_x representing F at x nm; FL, fluorescein; f(u), fraction unfolded; Fv, antibody fragment that contains only the variable domains V_H and V_L; Gdn-HCl, guanidine hydrochloride; Hepes, *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid; kDa, kilodaltons; Mops, 3-(*N*-morpholino)propanesulfonic acid; N, native folded state of protein; Q, fluorescence quenching, where $Q = (F_0 - F)/(F_0 - F_{bkg})$ and F_0, F, and F_{bkg} are the fluorescence intensities of the reference, sample, and background, respectively; Q_{max}, maximum fluorescence quenching; scFv, single-chain Fv, the subclass of Fv's which contain the V_H and V_L domains connected by a peptide linker; SD, standard deviation; Tris, tris(hydroxymethyl)aminomethane; $t_{1/2}$, half-life; U, unfolded state of protein; V_H, variable heavy domain; V_L, variable light domain.

[2] For a preliminary account of the stability/folding for one scFv protein, see Bedzyk et al. (1990).

0006-2960/91/0430-10117$02.50/0

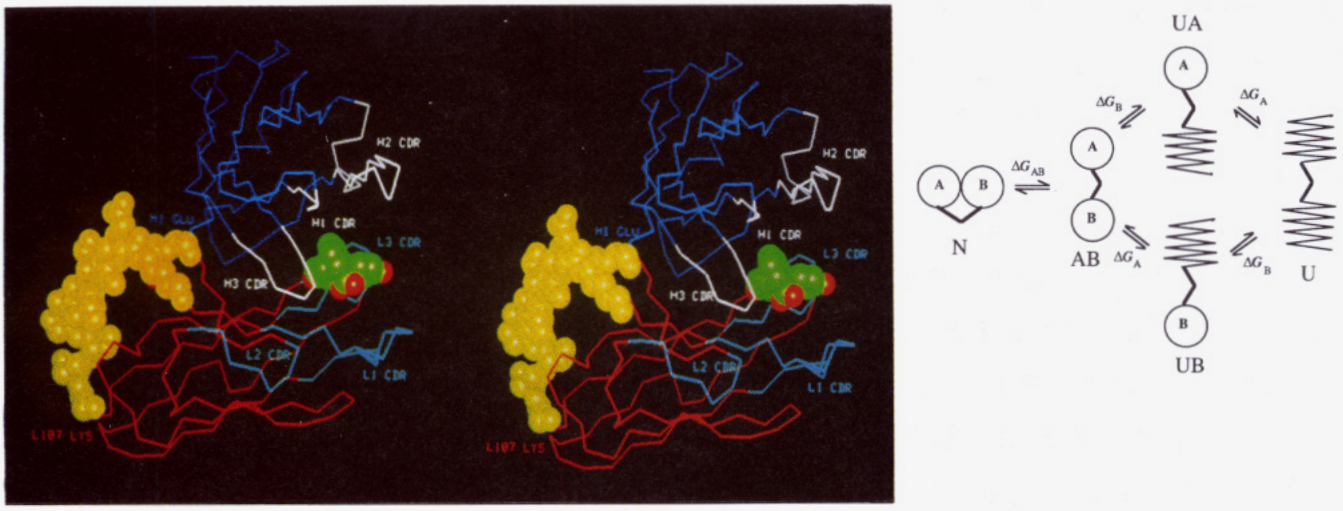

FIGURE 1: (A, left) Structural model of a single-chain Fv antibody fragment. The stereo model is based on the available X-ray structure of an antigen-bound Fab fragment of the antifluorescyl Mab 4.4.20 (see text) and is comprised of four main components: (i) the α-carbon backbone coordinates of the V_L domain (red); (ii) the α-carbon coordinates of the V_H domain (blue); (iii) a peptide linker (all atoms with van der Waals' surfaces shown in yellow) which connects the C-terminus residue, L107, of V_L to the N-terminus, H1, of V_H (the peptide linker shown here is the 14-residue 212 linker in a hypothetical conformation); (iv) the antigen, fluorescein, represented by van der Waals' surfaces (green and red for carbon and oxygen atoms, respectively) and bound at the complementary determining regions, CDR's (cyan for V_L and white for V_H). (B, right) Simple domain model for the unfolding/refolding reactions of scFv proteins. Circles represent native domains, while zigzag patterns represent unfolded domains. The interconnecting polypeptide linker is represented by a thick line. Possible partially unfolded intermediates between the native state, N, and the completely unfolded state, U, are represented by AB, UA, and UB. Their thermodynamic relevance is determined by the relative magnitudes of the respective equilibria as represented by ΔG_{AB}, the free energy of interdomain association, and the free energies of unfolding for the individual domains, ΔG_A and ΔG_B, as described in the text.

(Movva et al., 1980). The expression plasmids were transformed into a strain containing the C_{I857} repressor gene, and the synthesis of scFv protein was induced by raising the culture temperature to 42 °C after growth at 30 °C. The ompA signal was expected to cause the scFv protein to be secreted into the periplasmic space and/or into the culture media. The ompA signal was found to be correctly processed, but the majority of the protein was produced in an insoluble inclusion body like form (Mitraki & King, 1989). Thus, under these conditions, the ompA signal does not provide significant secretion; however, the ompA sequence does facilitate expression and allows production of the scFv protein at >10% of the total cell protein, as estimated by densitometer scans of stained SDS–PAGE gels.

Recovery of Insoluble scFv from E. coli. A three-step process was used. (1) Cell paste recovery: *E. coli* cell paste was recovered by centrifugation at 11000*g* for 30 min and stored at −20 °C. (2) Cell homogenization: Frozen cell paste was suspended in 10 volumes of buffer (50 mM Tris-HCl, pH 8.0, 1.0 mM EDTA, 0.1 mM PMSF) at 4 °C and passed through a pressurized cell homogenizer (9000 psi).[3] The homogenate was then centrifuged at 23000*g* for 30 min, and the insoluble pellet was saved and then resuspended with the same buffer (4 °C) and passed through the homogenizer once again. (3) Recovery of insoluble fraction: After centrifugation the pellet was washed four times with the same buffer (4 °C). SDS–PAGE analysis of the isolated protein followed by densitometer scanning indicated that the scFv was ~40% pure at the end of this step. The final pellet was stored at −20 °C.

Renaturation of Active scFv Proteins. A three-step process was used. (1) Dissolution in 6 M guanidine hydrochloride (Gdn-HCl): The insoluble fraction isolated above was dissolved in 6 M Gdn-HCl (50 mM Tris-HCl, pH 8.0, 50 mM KCl) so that [protein] ≈5–10 mg/mL and then centrifuged to remove particulates. (2) Renaturation: The dissolved scFv protein was then diluted 200-fold with buffer (50 mM Tris-HCl, pH 8.0, 10 mM $CaCl_2$, 50 mM KCl, 0.1 mM PMSF) thermostated at 7.0 °C.[4] The solution was allowed to sit for 24 h without stirring (7.0 °C). The progress of folding was followed as shown in Figure 2. (3) Ultrafiltration: The protein solution was concentrated on a Millipore Pellicon ultrafiltration apparatus (10 kDa cutoff).

Protein Purification. Purification of the renatured scFv's was accomplished by cation-exchange HPLC using a Waters Delta Prep 3000 preparative system. Typically, ~1 L of the concentrated scFv (0.14 mg/mL by activity) was equilibrated with the HPLC starting buffer, 40 mM Mops, pH 6.4, and 1.0 mM $Ca(OAc)_2$ and then pumped onto a 21.5 × 250 mm semipreparative polycation A column (Nest Group) also equilibrated with the same buffer. A pH elution gradient was then employed by linearly increasing the ratio of 40 mM Mops, pH 7.4, and 10 mM $Ca(OAc)_2$ (buffer B) to the starting buffer. Purified scFv's were verified for homogeneity through SDS–PAGE analysis (Figure 3A) or isoelectric focusing.[5] The scFv concentration was estimated spectrophotometrically using ϵ_{280} = 59055 M^{-1} cm^{-1} calculated from the Trp and Tyr content of the translated amino acid sequences of each scFv gene. This calculation was made using ϵ = 1413 M^{-1} cm^{-1} for Tyr (×15) and ϵ = 6310 M^{-1} cm^{-1} for Trp (×6) near 280 nm (*Handbook of Chemistry and Physics*, 1974–1975). This was converted to $E_{280}^{0.1\%}$ values of 2.21, 2.20, and 2.09 for 4.4.20/202′, 4.4.20/212, and 4.4.20/205, respectively, with the corresponding MW's of 26 700, 26 800, and 28 200 (MW's calculated from sequences minus ompA signal). The Fab was calculated to have ϵ_{280} = 96 186 M^{-1} cm^{-1} ($E_{280}^{0.1\%}$ = 1.96) on the basis of 10 Trp and 22 Tyr residues (MW = 48 100).

Preparation of Fab. The anti-fluorescyl 4.4.20 Mab was subjected to limited papain hydrolysis as described by Gibson

[3] The extent of lysis was >95%, as judged by phase-contrast microscopy.

[4] A preliminary screen of refolding buffer components and *T* revealed these conditions to be best.

[5] Observed p*I*'s are 8.6 for 4.4.20/202′, 9.4 for 4.4.20/212, and 8.6 for 4.4.20/205.

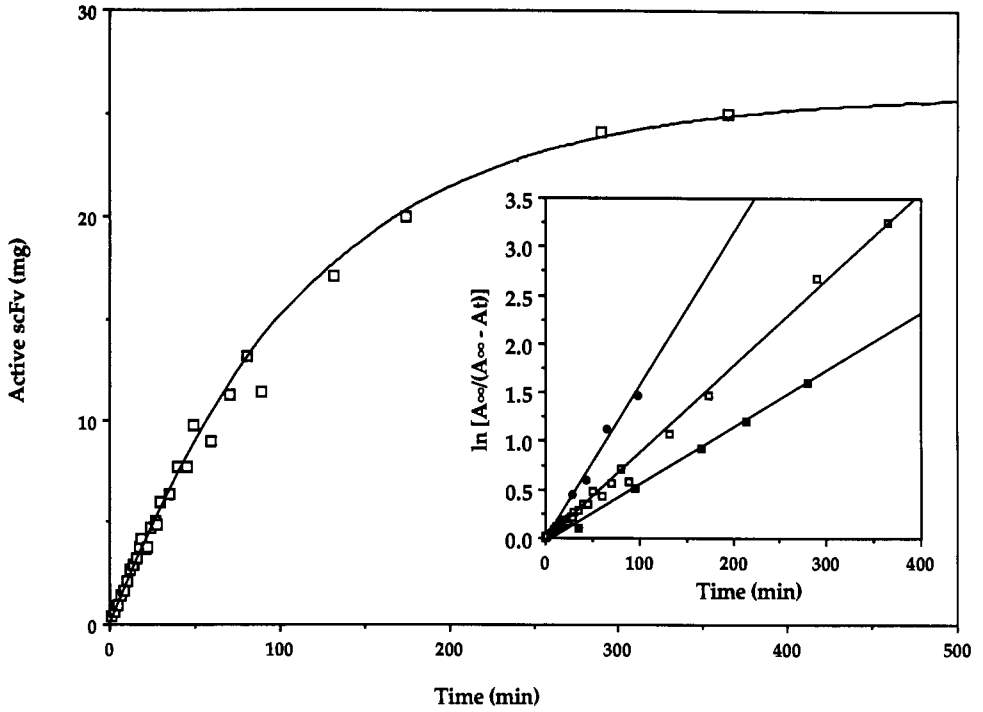

FIGURE 2: Refolding kinetics of scFv proteins at 7.0 °C and pH 8.0. A solubilized sample of 4.4.20/212 in 6 M Gdn-HCl (25 mL of [protein] = 10 mg/mL) was diluted 200-fold with 50 mM Tris-HCl, pH 8.0, and 10 mM $CaCl_2$ at $t = 0$, and aliquots were removed at the indicated times (□) and assayed for FL binding/quenching by titrating a 1.0 nM solution of FL with small aliquots of scFv solution until 40% quench (Q) of the initial F_{510} (excitation at 495 nm). A 40% Q is equivalent to an active [scFv] = ~1.0 nM (26.8 ng/mL) when log K_a = 9.0 for the binding of FL to scFv (see text) as judged from simple titration curves where half of the FL is bound (see footnote 14). The Q_{max} was found to be ~96% for concentrated solutions of purified scFv proteins. The curve drawn through the data is for a first-order process of $t_{1/2}$ and 80 min, and A_∞ = 26 mg. The results were normalized for 1 L of *E. coli* broth. The final diluted [protein] = 50 μg/mL (1.9 μM) thermostated at T = 7.0 °C.[4] (Inset) First-order analysis for refolding of the scFv proteins: 4.4.20/202′ (■), 4.4.20/212 (□), and 4.4.20/205 (●). Straight lines represent linear least-squares fits to the data. Activity was measured for time t (A_t) and time ∞ (A_∞), respectively. The A_∞ values for the three proteins are 3, 26, and 12 mg, respectively. The k_{obs}'s for the rates of folding for the individual scFv proteins were obtained from the slopes and found to be 0.97×10^{-4} s^{-1}, 1.50×10^{-4} s^{-1}, and 2.61×10^{-4} s^{-1} for 4.4.20/202′, 4.4.20/212, and 4.4.20/205, respectively. All conditions are the same as described for the 4.4.20/212 protein.

et al. (1988). The resulting Fab's were purified by HPLC as described above and affinity chromatography as described by Voss (1984). The purified Fab yielded diffraction-quality crystals when crystallized using the methods of Gibson et al. (1988).

Equilibrium Fluorescein Ligand Binding. The affinity, K_a, and stoichiometry, n, for the equilibrium association reaction of the FL ligand were determined by Scatchard analysis as previously described (Scatchard, 1949; Voss, 1984; Bird et al., 1988).

Gdn-HCl- and Urea-Induced Unfolding/Refolding Equilibria. Solvent denaturation studies were conducted using the intrinsic Trp fluorescence emission spectra of the scFv proteins to monitor the extent of the unfolding/refolding transition (Pace, 1986, 1990). The fluorescence emission maximum at 350 nm (excitation at 290 nm) was found to decrease by ~40% upon unfolding in 4.0 M Gdn-HCl with no noticeable change in λ_{max}. Fluorescence intensity (F) was measured with a Perkin-Elmer LS-5 spectrometer. Stock solutions of 6 M Gdn-HCl or 8.0 M urea[6] were gravimetrically prepared by dissolution in 150 mM NaCl and 50 mM Tris-HCl, pH 8.0. These solutions were then diluted with the same buffer to the required [D] before the addition of protein. The [protein] was 10 μg/mL (0.38 μM) but varied between 1.0 and 25 μg/mL in specified cases. The unfolding reactions were initiated by diluting solutions of purified scFv's into the individual solutions of [D] and then incubating at 25 °C in a circulating water bath for 5 h to reach equilibrium.[7] F_{350} was then measured for the various [D]'s. Refolding reactions were initiated by first diluting purified scFv's into 4 M Gdn-HCl or 7 M urea to obtain the unfolded form of these proteins and then diluting to the desired [D].

A nonlinear least-squares fit of the fluorescence data to a two-state model for protein unfolding was performed using the BCONG subroutine of the IMSL Math/Library v. 1.0 (IMSL, Houston, TX). BCONG uses a quasi-Newton method with a user-supplied analytical gradient function for parameter optimization and was run on a Cray X-MP computer. Error determinations are based upon standard error matrix (inverse of the Hessian) analysis.

Molecular Modeling. Molecular modeling simulations were conducted with an Evans and Sutherland computer graphic terminal, driven with FRODO software (Jones, 1978), or a Silicon Graphics workstation, driven with Biosym INSIGHT II software. Three linkers were designed by first choosing the minimal-length peptide necessary to span the ~40-Å distance separating the C-terminus, L107, of the V_L domain to the N-terminus, H1, of the V_H domain (Figure 1A), as judged from high-resolution X-ray structures of Fab fragments [for example, Satow et al. (1986)] available in the Brookhaven Protein Data Bank (Bernstein et al., 1977). It was also possible to use a partially refined 1.8-Å resolution X-ray structure obtained through the crystallization of the anti-FL 4.4.20 Fab

[6] Urea solutions were frozen in 10-mL aliquots immediately after mixing and stored at −20 °C to minimize ammonium cyanate formation (Kellis et al., 1989). Aliquots were then used as needed.

[7] The time course for this reaction was found to be nearly complete after ~0.5 h.

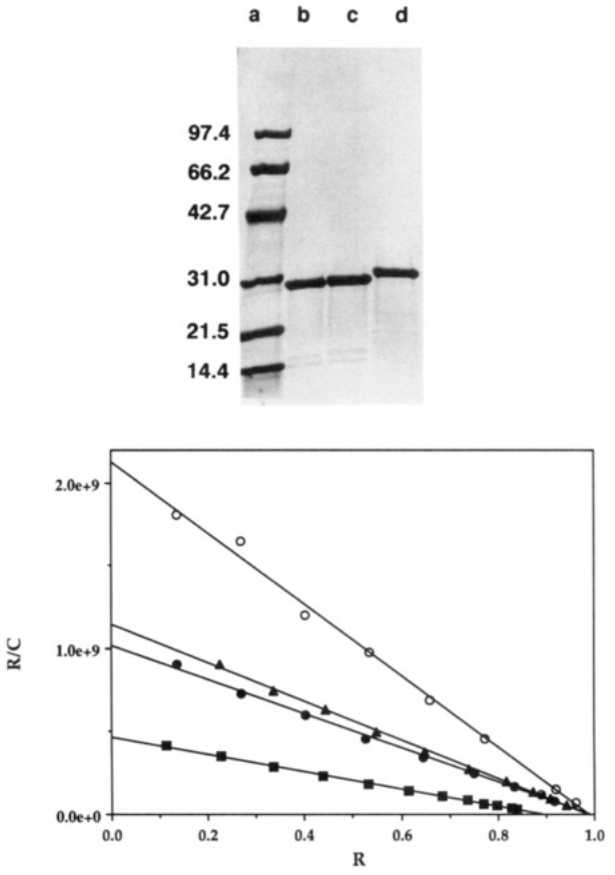

FIGURE 3: (A, top) SDS–PAGE gel analysis (Laemmli, 1970): lane a, MW standards ($\times 10^3$); lane b, 4.4.20/202′; lane c, 4.4.20/212; lane d, 4.4.20/205. Lanes were loaded with 20 μg of protein. (B, bottom) Scatchard analysis of FL ligand binding to scFv proteins and parental Fab. FL binding was assayed by titration of a 70 nM solution of purified scFv or Fab with small aliquots of a 400 nM solution of FL while monitoring F_{510} (excitation at 495 nm). The R values, $[\text{scFv·FL}]/[\text{scFv}]_t$, were calculated from Q/Q_{max} as described by Voss (1984). The observed K_a values (M^{-1}) are 0.49 ($\pm$0.03) $\times 10^9$, n = 0.90, for 4.4.20/202′ (■); 1.07 ($\pm$0.04) $\times 10^9$, n = 0.99, for 4.4.20/212 (●); 1.15 ($\pm$0.02) $\times 10^9$, n = 0.99, for 4.4.20/205 (▲); and 1.99 ($\pm$0.18) $\times 10^9$, n = 0.99, for 4.4.20 Fab (○); and are the average for three separate experiments with the errors $\pm$1 SD (T = 25 °C, in 20 mM Hepes–NaOH, pH 8.0, with I = 0.10 adjusted with NaCl).

derived from proteolysis of the parental Mab (manuscript in preparation). A similar X-ray structure for a 4.4.20-derived Fab has since been reported by Herron et al. (1989) but was not yet available at the time of the design of these peptide linkers.

The minimum-length peptide was judged to be 12 residues and -G-K-S-S-G-S-G-S-E-S-K-S- was chosen as the shortest possible linker and called 202′. The second peptide linker, 212, was two residues longer, with the sequence -G-S-T-S-G-S-G-K-S-S-E-G-K-G-, and was chosen to test the effect of linker length on scFv functionality. Finally, a much longer 25-residue linker, 205, was chosen for its potential to pack up against the V domains and possibly introduce specific stabilizing interactions. The sequence is -S-S-A-D-D-A-K-K-D-A-A-K-K-D-D-A-K-K-D-D-A-K-K-D-G- and was chosen for its overabundance of relatively hydrophilic residues.

Results

Kinetics of Refolding of scFv Proteins. The time course for folding of the scFv proteins was followed by monitoring the increase in FL binding activity as a function of time after dilution from 6 M Gdn-HCl. The refolding reactions appeared to follow first-order kinetics under the conditions employed as demonstrated in Figure 2 for 4.4.20/212. The $t_{1/2}$ of this reaction was found to be 80 min at 7 °C and yielded 26 mg of active scFv/L of fermentation after 500 min.[8] The nature of the first-order folding reaction is shown in Figure 2 (inset). The data for 4.4.20/212, which was monitored for nearly five $t_{1/2}$'s, were found to fit a straight line with the y-intercept extrapolation very close to 0 at zero time, suggestive of a first-order process with no evidence of stable intermediates during the course of this reaction. Similarly, the data for the 4.4.20/202′ and 4.4.20/205 were also found to be linear under these conditions, implying that all three scFv proteins follow first-order folding reactions. Variation of [protein] was not found to appreciably affect the rates of folding, only the yield due to aggregation of the unfolded, U, form (see Figure 1B) of these proteins at >25 μg/mL. The presence of reducing agents (10 mM DTT) in the initial 6 M Gdn-HCl buffer was found to decrease the rates of folding by 2- or 3-fold, with greatly reduced yields. Thus, the folding reactions reported here are for the oxidized forms of these proteins where the disulfide bonds are intact (vide infra).[9]

Equilibrium Fluorescein Ligand Binding. Scatchard analyses for the binding of the FL ligand to 4.4.20/202′, 4.4.20/212, and 4.4.20/205 scFv proteins are shown in Figure 3B together with the parental Fab. These plots are linear over the R ($[\text{scFv·FL}]/[\text{scFv}]_t$) interval of 0.10–0.95 as expected for simple single-valency binding equilibra with no cooperativity. Moreover, the data for the scFv proteins revealed that K_a, as determined from the slope, and n, as determined from the x-axis intercept, were in reasonable agreement with those of the 4.4.20-derived Fab. This was especially true for the 4.4.20/212 and 4.4.20/205 scFv's where the K_a's were only 2-fold smaller than that for the Fab.

Assay of Disulfide Linkages and Free Thiols. The number of disulfides in a purified sample of the 4.4.20/212 scFv was determined by the method of Thannhauser et al. (1984) in 3.0 M Gdn-SCN. A 4.36 μM solution of 4.4.20/212 yielded 8.83 μM 2-nitro-5-thiobenzoic acid, thereby identifying the presence of 2.03 disulfide linkages/mol of scFv. In addition, free thiols were assayed by the method of Riddles et al. (1983) in 6.4 M Gdn-HCl. No free thiol groups were detected in a 4.36 μM solution of 4.4.20/212. Together, these results suggest that the disulfide bonds L23C/L88C and H22C/H92C are intact in the refolded scFv's.

N-Terminal Amino Acid Sequence. A purified sample of 4.4.20/202′ was sequenced via Edman degradation (Edman & Begg, 1967). The first 22 cycles yielded NH_2-D-V-V-M-T^5-Q-T-P-L-S^{10}-L-P-V-S-L^{15}-G-D-Q-A-S^{20}-I-S-, as expected for the N-terminus of the V_L domain (Bedzyk et al., 1989). Thus, the ompA signal sequence is correctly removed from the scFv proteins in vivo.

Unfolding/Refolding Transitions. (A) Data Analysis. The transition curves for the solvent-induced denaturation of the scFv proteins were analyzed by assuming a two-state approximation, N ↔ U, where only the N and U states are present in signficant concentration during the reaction. The equilibrium constant for unfolding, K_{obs}, is defined by K_{obs} =

[8] The rates were measured at 7 °C because of the failure of 4.4.20/202′ to fold at 25 °C.

[9] The oxidation of thiols in inclusion bodies to form disulfide bonds presumably occurs during lysis of the cells (Mitraki & King, 1989), and for a protein containing four thiols a random oxidation would be expected to produce ten different forms of a scFv protein that contain zero, one, *or* two disulfide bonds, and only one of these contains the correct two disulfide linkages. This may also explain why only 26 mg/L fermentation was obtained and not the 250 mg expected from estimates of protein expression.

Table I: Changes in Free Energy of Unfolding/Refolding for scFv Proteins at pH 8.0 and 25 °C

denaturant	protein/linker	$\Delta G°_{n-u}$[a] (kcal/mol)	$\Delta G°_{u-n}$[a] (kcal/mol)	m_G[b] (kcal/mol·M)	C_m[c] (M)	$\Delta(\Delta G°_{n-u})$[d] (kcal/mol)
Gdn-HCl	4.4.20/202′					
	4.4.20/212	4.33 ± 0.30	−4.12 ± 0.64[e]	4.18 ± 0.26	1.04	0.00
	4.4.20/205	4.86 ± 0.36		3.98 ± 0.28	1.22	0.53
urea	4.4.20/202′	3.87 ± 0.26	−3.41 ± 0.42	1.47 ± 0.09	2.63	−1.25
	4.420/212	5.12 ● 0.30	−5.15 ± 0.61	1.39 ± 0.08	3.68	0.00
	4.4.20/205	5.43 ± 0.29	−5.45 ± 1.36	1.44 ± 0.08	3.77	0.31
lit. Gdn-HCl	V_L (κ, Oku)[f]	4.7		3.7	1.3	
	C_L (κ, Oku)[f]	3.8		4.1	0.9	
	V_L (κ, Wes)[g]	5.5			1.15	
	C_L (κ, Wes)[g]	5.5			1.15	
	Fab (κ, Wes)[h]				1.4	
	C_L (λ, Nag)[i]	5.7			1.2	
	C_L (λ, Nag)[i]	1.6 (-SH HS-)			0.4	
	C_L (λ, Nag)[j]	1.4 (-S-Hg-S-)			0.4	

[a] Nonlinear least-squares fit of the F_{350} data as shown in Figures 4–6 and described in the text. Errors represent ±1 SD (67% confidence limits) for at least three separate experiments. Unless noted otherwise, the $\Delta G°_{n-u}$ values are for the form of the protein that has the disulfide bonds intact. The $\Delta G°_{u-n}$ values are for the refolding reaction. [b] m_G is the slope of the linear denaturation plot, $-d(\Delta G_{obs})/d[D]$. [c] C_m = [D] at the midpoint of unfolding transitions. [d] $\Delta(\Delta G°_{n-u})$ was calculated by subtraction of the $\Delta G°_{n-u}$ values relative to that of the 4.4.20/212 protein. [e] The m_G best-fit values for the refolding reactions in urea are 1.43 ± 0.16, 1.56 ± 0.18, and 1.35 ± 0.33 kcal/mol·M for 4.4.20/202′, 4.4.20/212, and 4.4.20/205, respectively. The best-fit m_G value for the refolding of 4.4.20/212 in Gdn-HCl was 3.74 ± 0.49. [f] Goto et al. (1988), pH 7.5, 25 °C. $\Delta G°_{n-u}$ was measured by the linear extrapolation method (Pace, 1986). [g] Rowe and Tanford (1973), pH 7.0, 25 °C. $\Delta G°_{n-u}$ was measured by the free energy of transfer model (Tanford, 1970). [h] Rowe (1976), pH 7.0, 25 °C. [i] Goto and Hamaguchi (1979), pH 7.5, 25 °C. $\Delta G°_{n-u}$ was measured by the D binding model (Pace, 1986). [j] Goto and Hamaguchi (1986), pH 7.5, 25 °C. $\Delta G°_{n-u}$ was measured by the D binding model.

$f(u)/[1 - f(u)]$ at each [D], where $f(u) = [U]/[N]$. The equations that describe the linear effect of solvent on the N and U forms of the proteins are

$$y_n = y_n° + m_n[D] \quad \text{and} \quad y_u = y_u° + m_u[D] \qquad (1)$$

where y_n and y_u are F values for the N and U forms of these proteins as a function of [D] and where $y_n°$ and m_n, and $y_u°$ and m_u, are the y-intercept and slope of the pre- and post-transitional baselines, respectively. The fraction unfolded, $f(u)$, can then be evaluated from the observed F_{350}, y_{obs}, by

$$f(u) = (y_n - y_{obs})/(y_n - y_u) \qquad (2)$$

The observed differences in free energy between the N and U forms, ΔG_{obs}, at each [D] can then be assessed by

$$\Delta G_{obs} = -RT \ln \{f(u)/[1 - f(u)]\} = -RT \ln [(y_n - y_{obs})/(y_{obs} - y_u)] \qquad (3)$$

For the two-state model, ΔG_{obs} will, in general, be a linear function of [D] and can be expressed as

$$\Delta G_{obs} = \Delta G°_{n-u} - m_G[D] \qquad (4)$$

where m_G is the slope of the linear denaturation plot, $-d(\Delta G_{obs})/d[D]$, and $\Delta G°_{n-u}$ is the free energy change for unfolding at [D] = 0, or the y-intercept.

As suggested by Santoro and Bolen (1988) eqs 1, 3, and 4 can be combined to obtain an expression for the observed fluorescence change, y_{obs}, as a function of [D]:

$$y_{obs} = \frac{\left[(y_u° + m_u[D])\left\{\exp\left[-\left(\frac{\Delta G°_{n-u} - m_G[D]}{RT}\right)\right]\right\} + (y_n° + m_n[D])\right]}{\left[\exp\left[-\left(\frac{\Delta G°_{n-u} - m_G[D]}{RT}\right)\right] + 1\right]} \qquad (5)$$

The six fitting parameters then become $y_n°$, m_n, $y_u°$, m_u, m_G, and $\Delta G°_{n-u}$. The last parameter is a measure of the stability of the proteins in the reference state, [D] = 0, pH = 8.0, T = 25 °C, and is equivalent to that $\Delta G°_{n-u}$ commonly measured by the linear extrapolation method for assessing protein stability (Pace, 1986), except that eq 5 allows a nonlinear least-squares computer fit to all the data in Figure 4A and not just the data in the transition region. An advantage of the nonlinear least-squares fit is that the errors are more realistic (and larger) due to the inclusion of the uncertainties for the pre- and posttransitional solvent effects (Santoro & Bolen, 1988).

(B) Gdn-HCl-Induced Unfolding. The scFv's were equilibrated with increasing [Gdn-HCl], and the observed change in Trp fluorescence was plotted in Figure 4A. The data for 4.4.20/202′ failed to converge in the fitting program due to the lack of a pretransition baseline. From the computer-fit parameters it was possible to generate $f(u)$ vs [D] plots for the two scFv proteins that converged (Figure 4B). A summary for the fitted m_G and $\Delta G°_{n-u}$ are shown in Table I.

(C) Urea-Induced Unfolding. The urea-induced transitions for all three scFv proteins were found to be sufficient for convergence, due to the milder denaturing power of urea compared to Gdn-HCl, and the results are plotted in Figure 5 in terms of $f(u)$ vs [urea] after calculation from the optimized best-fit parameters using eqs 2–5. A summary for the fitted m_G and $\Delta G°_{n-u}$ is shown in Table I. The $\Delta G°_{n-u}$ values obtained in urea are close to those in Gdn-HCl but are consistently ~10–15% higher.[10] This was not found to be statistically significant because the $\Delta G°_{n-u}$ values overlap in the 2σ range (95% confidence limits) and the variance ratio F tests indicate that there is a 95% chance that they are the same.

(D) Reversibility of Solvent-Induced Equilibria. The unfolding processes for all three scFv proteins were found to be nearly completely reversible in both Gdn-HCl and urea, as demonstrated in Figure 6 by the close agreement between the transitions for the unfolding and refolding reactions. The four refolding reactions, for which there were sufficient data for convergence of the fitting program, gave $\Delta G°_{u-n}$ values (opposite sign) that were in good agreement with the corresponding $\Delta G°_{n-u}$ (see Table I). The errors associated with the refolding reactions were found to be larger than those for

[10] This result is similar to that found by Santoro and Bolen (1988) where the urea data for chymotrypsin was found to be 7% higher than that in Gdn-HCl.

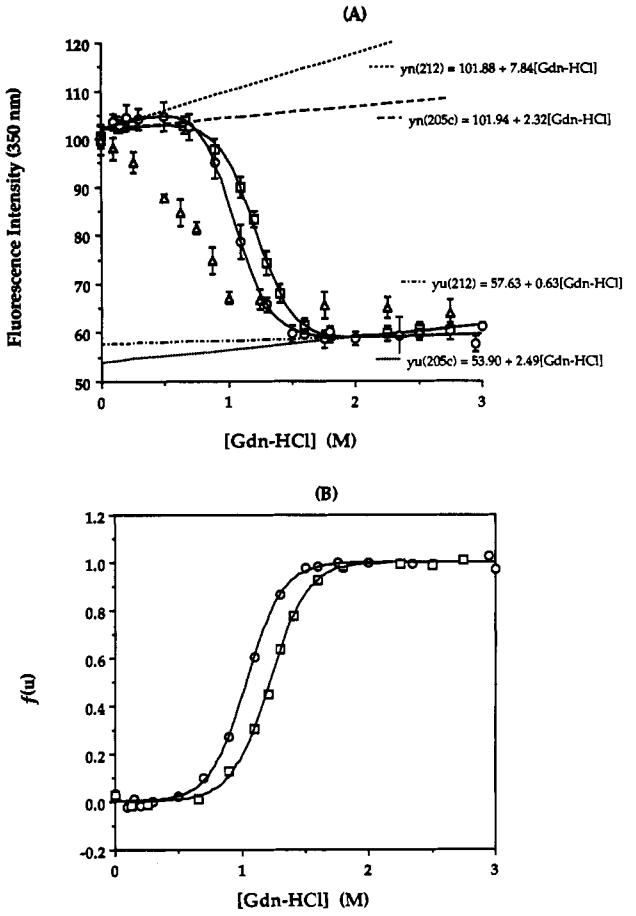

FIGURE 4: Unfolding reactions for scFv proteins in Gdn-HCl. (A) Intrinsic Trp F_{350} vs [Gdn-HCl] for scFv proteins: 4.4.20/202′ (△), 4.4.20/212 (○), and 4.4.20/205 (□). The solid lines represent the nonlinear least-squares fit of the experimental points for the scFv proteins 4.4.20/212 and 4.4.20/205 to eq 5 in the text, constructed using the best-fit estimations of the six fitting parameters, $y_n°$, m_n, $y_u°$, m_u, m_G, and $\Delta G°_{n-u}$, expected for a simple two-state N ↔ U assumption. The dotted and dashed lines represent the pre- and posttransitional lines for 4.4.20/212 and 4.4.20/205 and were drawn using the parameters derived from the optimization analysis. The 4.4.20/202′ data could not be fit due to lack of unique convergence. The [scFv] was 10 μg/mL (0.38 μM) and T = 25 °C for all experiments in Figures 4–6. Also, each data point in all three figures represents the average of at least three individual reactions with the error bars representing ±1 SD. (B) Fraction unfolded vs [Gdn-HCl]. The fraction unfolded, f(u), for each original F_{350} data point at a given [D] in (A) was calculated using eq 2 and the best-fit parameters found above for the pre- and posttransitional lines. The continuous solid lines represent the nonlinear least-squares best fit through all the data points, as described above, but expressed in terms of f(u) instead of y_{obs}.

unfolding and are attributed to our tendency to neglect data collection at low [D], thereby increasing the error contribution from the post-folding transition baseline.

A further test of the reversibility of these transitions is through the assay of FL ligand binding to scFv proteins after one denaturation–renaturation cycle (second time if you include original refolding from inclusion bodies). The results indicated that these proteins retained ~95% of their original FL binding function as measured in side by side fluorescence quench assays with scFv or Fab proteins already assayed by Scatchard analysis in Figure 3B.

(*E*) *Molecularity of Solvent-Induced Equilibria.* The [protein] dependence for the unfolding transitions was investigated in order to establish whether significant association occurs during denaturation. A 20-fold range of [protein] between 1.25 and 25 μg/mL (44–885 nM) was found to have

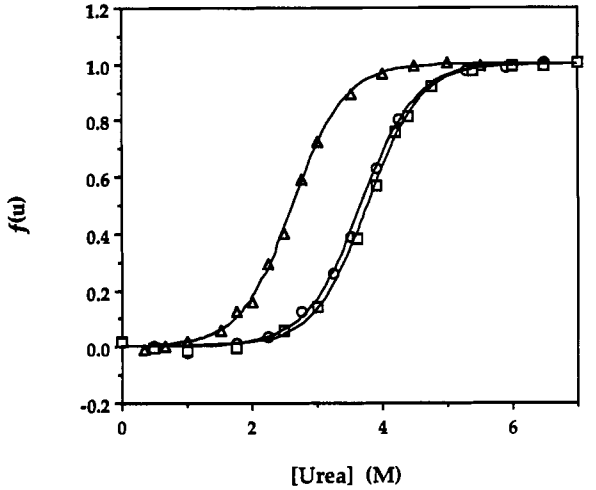

FIGURE 5: Unfolding reactions for scFv proteins in urea. The original F_{350} data for the scFV proteins were studied as a function of [urea] and analyzed by the nonlinear least-squares fitting program. For simplicity sake, only the fraction unfolded, f(u) is plotted vs [urea]. Legend: 4.4.20/202′ (△), 4.4.20/212 (○), and 4.4.20/205 (□). The solid lines represent the best computer fit through all the observed data points to the two-state model.

little effect on the conformational state of the scFv, 4.4.20/205, when analyzed in an f(u) vs [D] plot, implying that the unfolding of scFv's is unimolecular under these conditions.

Discussion

The kinetics and equilibria for the folding/unfolding reactions of scFv proteins were investigated primarily to obtain sufficient quantities of purified samples that are necessary to explore the in vivo antigen targeting (Colcher et al., 1990), immunological (Bedzyk et al., 1990), and physicochemical properties of this new class of designed proteins. The criteria used to evaluate the function of scFv proteins are binding affinity, K_a, stoichiometry, n, and conformational stability in terms of free energy, $\Delta G°_{n-u}$, or relative transition midpoints, C_m. By these criteria, the scFv designs represented by 4.4.20/212 and 4.4.20/205 come within 2-fold of the K_a for the naturally derived design, i.e., Fab, and are indistinguishable with regard to n, i.e., 0.99 (Figure 3A). This is of particular importance because previous reports for scFv proteins (Bird et al., 1988; Huston et al., 1988) contained low n values (~0.25) which could have been interpreted in terms of partly unfolded and unstable protein designs. In this regard, the scFv designs for 4.4.20/212 and 4.4.20/205 were found to be slightly more stable than single IgG domains for which there are sufficient data in the literature, i.e., κ V_L and C_L, when assessed in terms of $\Delta G°_{n-u}$, or nearly identical to κ light chains (Rowe & Tanford, 1973) and Fab's (Rowe, 1976) when comparing C_m (see Table I). Thus, by these criteria of ligand-binding function and protein stability, the 4.4.20/212 and 4.4.20/205 scFv proteins were found to be bona fide minimal monovalent replicas of their parental IgG molecules.

The second consideration in pursuing the kinetics and equilibra for the folding/unfolding reactions of scFv proteins was to explore this new class of proteins as potential models for the investigation of protein folding and stability. In this regard, the first-order kinetics for scFv folding reactions suggest that the folding occurs with no evidence for intermediate species along the reaction pathway between U and N states. Similar results have been observed for isolated type κ V_L domains obtained by limited proteolytic cleavage. The refolding and unfolding kinetics of the Oku V_L could be described by a single exponential term, with the refolding reaction

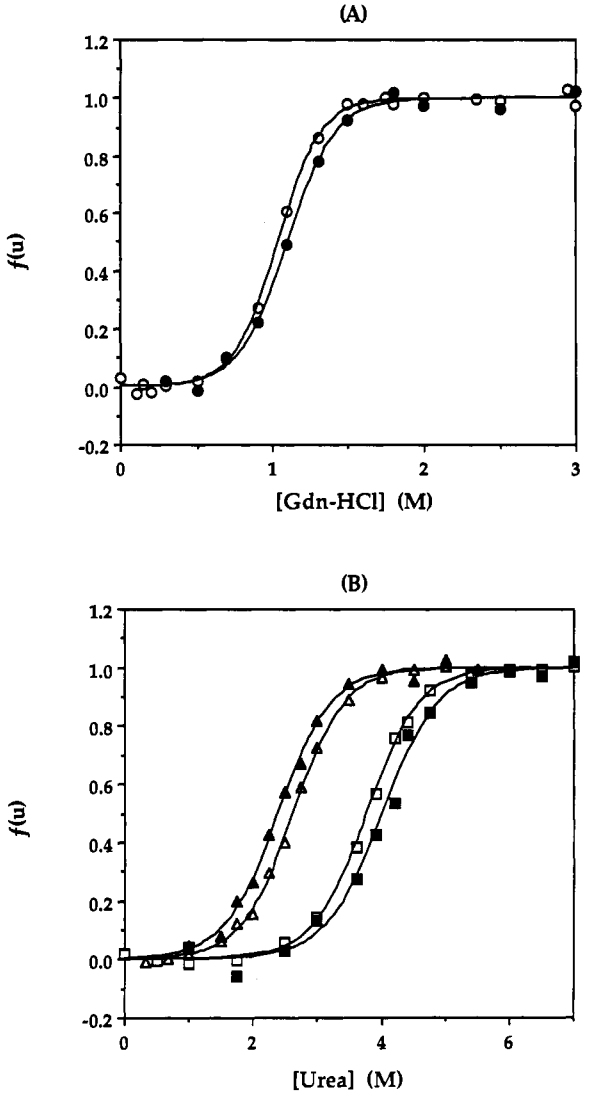

FIGURE 6: Reversibility of the unfolding/refolding reactions for scFv proteins. (A) Unfolding/refolding reactions of 4.4.20/212 in Gdn-HCl. The fraction unfolded, f(u), for the unfolding reaction of 4.4.20/212 (○) is compared with the data for the refolding reaction (●) where the protein begins at equilibrium in 4.0 M Gdn-HCl and is diluted to lower [D]. The solid lines represent the nonlinear least squares computer best fit of all the data points to the two state assumption after optimization of the six fitting parameters as described in the text ($\Delta G°_{n-u}$ is replaced by $\Delta G°_{u-n}$ and then plotted in terms of f(u). (B) Unfolding/refolding reactions in urea. The fraction unfolded, f(u) for the unfolding reactions of 4.4.20/202′ (△) and 4.4.20/205 (□) is shown alongside the data for the refolding reactions (▲ and ■, respectively). Fitting parameters are summarized in Table I.

characterized by a $t_{1/2}$ of 2.2 min at 25 °C, pH 7.5, and 0.3 M Gdn-HCl (Goto et al., 1988; Tsunenaga et al., 1987). Only double-jump experiments suggested that a more complicated mechanism may be necessary to explain the kinetics of folding of V_L domains. The V_L fragments are known to be in equilibrium with the homodimer, a persistent problem in the study of single V domains. Additional scrutiny with respect to the effect of [D] and/or T on the folding kinetics of scFv's may be useful for detecting intermediates. Continuous monitoring of the intrinsic Trp F_{350} may also facilitate the detection of early fast-folding phases that are over within the dead time of the activity assays.

Further support for a simple two-state model, N ↔ U, for the folding of scFv proteins was obtained through the investigation of unfolding/refolding equilibria as a function of [Gdn-HCl] and [urea]. The [D]-dependent reversible transitions were found to closely approximate a simple two-state equilibrium when analyzed via a nonlinear least-squares fit of the experimental data to eq 5 through the optimization of the six fitting parameters, $y_n°$, m_n, $y_u°$, m_u, m_G, and $\Delta G°_{n-u}$ (see Figures 4–6). Both the unfolding and refolding transitions could be fit in this manner (Figure 6), demonstrating the reversibility of this reaction and implying that the same simple reaction mechanism applies in either direction. The apparent close fit of the experimental unfolding/refolding transitions to two-state behavior implies that the concentration of intermediates along the reaction coordinate is quite low and only N and U are present in significant amounts.[11]

In order for a two-domain protein such as scFv to follow a two-state behavior, coupling of the individual domains into a cooperative folding unit must occur. This coupling will depend on the relative differences in intrinsic stability for each domain, as well as on the interaction between domains. Figure 1B is a model that describes the possible steps and intermediate states that may be expected for the unfolding of scFv proteins. Similar models have been used for the interpretation of the thermodynamics of two-domain proteins (Rowe & Tanford, 1973; Rowe, 1976; Brandts et al., 1989; Ramsay & Freire, 1990). Using such a multistate thermodynamic model, Brandts et al. (1989) have shown that if $\Delta G_A = \Delta G_B$, then ΔG_{AB} interactions of ~2.5 kcal/mol are sufficient to couple unfolding transitions, under the assumptions used in their analysis. Moreover, the ΔG_{AB} interaction necessary to couple domains A and B as a single thermal transition when domain A alone has a T_m^A = 50 °C while domain B alone has a T_m^B = 60 °C may be estimated (Brandts et al., 1989):

$$\Delta G_{AB} = 2.5 + \Delta H_A(T_m^B/T_m^A - 1) \quad \text{(kcal/mol)} \quad (6)$$

where 2.5 kcal/mol is the interaction energy necessary to couple two domains of equal T_m and ΔH of unfolding so that the coupling index $\Delta H_{cal}/\Delta H_{vH}$, the ratio of calorimetric and van't Hoff enthalpies (Privalov & Potekhin, 1986), is 1.1. This simulation assumes a ΔH_A of 60.0 kcal/mol at T_m^A [see Goto & Hamaguchi (1987)] and $\Delta C_p = 0$. Thus, the observed two-state behavior for transitions of scFv proteins is consistent with these simulations, whether $\Delta G_{V_L} = \Delta G_{V_H}$ or $\Delta G_{V_L} \neq \Delta G_{V_H}$, as long as ΔG_{AB} is 2.5–4.4 kcal/mol, depending on the extent of inequality between ΔG_{V_L} and ΔG_{V_H}. In either case, the transitions for a coupled cooperative folding unit would be expected to occur at a slightly higher C_m or T_m (ΔT_m ~ 5 °C) than the most stable single domain.[12]

The individual V_L and V_H domains of the 4.4.20 scFv have not yet been examined as a function of [D]. The relative conformational stabilities of the V_L and the V_H domains for this protein, or any IgG molecule for that matter, are therefore unknown. In a recent report, Glockshuber et al. (1990) mention that the rate of thermal inactivation of McPC603 V_H domains is "rapid" relative to that of the V_L domain at 37 °C without elaboration as to magnitude or molecularity (extent to which homodimers exist). The conformational stabilities of V_L and C_L domains, however, have been explored in greater depth and are summarized in Table I. Stability measurements through Gdn-HCl-induced unfolding (Goto et al., 1988) reveal that the V_L (Oku, κ) domain obtained through clostripain digestion of light chains undergoes an unfolding transition at

[11] Computational experiments were conducted using an equation for a three-state model similar to eq 5. It was found that unless the intermediate state was quite distinct, strongly disturbing the apparent linearity of the transition region in ΔG_{obs} vs [D] plots, it was difficult to obtain converged solutions with reasonable error limits on the fitting parameters.

[12] Preliminary DSC experiments for scFv proteins were plagued by the low solubility of the unfolded form, resulting in aggregation near the T_m (~49 °C for 4.4.20/202′).

C_m = 1.3 M Gdn-HCl (pH 7.5, 25 °C) and was found to have a $\Delta G°_{n-u}$ = 4.7 kcal/mol when analyzed using the two-state assumption and the linear extrapolation method. Similarly, the C_L fragment from this same light chain was found to have C_m = 0.9 M Gdn-HCl and $\Delta G°_{n-u}$ = 3.8 kcal/mol (pH 7.5, 25 °C). The molecularity of the unfolding reaction for the C_L fragment was clearly unimolecular, while the V_L fragments were found to be in a monomer/dimer equilibrium with the V_L dimer contributing to the stability of the V_L fragment (Goto et al., 1988; Tsunenaga et al., 1987). The 4.7 kcal/mol is therefore considered an upper limit for $\Delta G°_{n-u}$ of the isolated Oku V_L. These are the only $\Delta G°_{n-u}$ data for single IgG domains evaluated through the linear extrapolation method, which has only recently become widely accepted (Pace, 1986, 1990). The older data in Table I were obtained either through the free energy of transfer model (Tanford, 1970) or the denaturant binding method (Aune & Tanford, 1969a,b) which invariably yields higher values (35–75% higher) for $\Delta G°_{n-u}$.[13] Therefore, the Oku data for V_L and C_L are the only literature $\Delta G°_{n-u}$ data directly comparable with those reported here for scFv's.

The bimolecular association constants, K_{AB}, for Fv antibody fragments (no linker) have been observed to range between 10^5 and 10^8 M^{-1} (Hochmann et al., 1976; Klein et al., 1979; Horne et al., 1982; Glockshuber et al., 1990). Since they are bimolecular constants, the association of V_L and V_H domains will be [protein] dependent near the value of $1/K_{AB}$ and the corresponding ΔG_{AB} should be expressed as $\Delta G_{AB} = -RT \ln (K_{AB}[B])$. On the other hand, the equilibrium constant, K_{intra}, for an intramolecular interaction between domains A and B in a scFv can be expressed in terms of K_{AB} by

$$K_{intra} = K_{AB}[A/B]_{sc} \quad (7)$$

where $[A/B]_{sc}$ is the effective molar concentration of domain A relative to domain B in the scFv molecule (Creighton, 1984, and references cited therein). The magnitude of $[A/B]_{sc}$ expected for a scFv is unknown and will be dependent on the degrees of translation and rotational freedom in the AB form of the molecule, but estimates of 10^{-1}–10^{-4} M have been suggested as theoretical values for the $[A/B]_{sc}$ of the ends of random polypeptides of 10–20 residues (Mutter, 1977; Creighton, 1984) relative to each other, i.e., [N-terminus/C-terminus]$_{sc}$. Therefore, when working at [protein] = ~1 μM, one may estimate that the unimolecular K_{intra} is ≥100 times higher than the corresponding bimolecular $K_{AB}[B]$, resulting in a ΔG_{AB} for scFv's that is at least 2.7 kcal/mol greater than that for the comparable Fv, primarily due to the entropic effect of covalent linkage of the two domains in scFv designs. This was the original reason why the scFv strategy was preferred by Bird et al. (1988) and is borne out by the observation of Glockshuber et al. (1990) that conversion of McPC603 Fv's to scFv's through introduction of an interdomain peptide, disulfide, or chemical linkage causes a 10–60-fold increase in kinetic thermal stability (37 °C). Moreover, the extra 2.7 kcal/mol expected for ΔG_{AB} of scFv vs Fv proteins near [protein] = 1 μM may alone be sufficient to couple the V_L and V_H domains of scFv proteins into a single cooperative folding unit, as suggested by eq 6, possibly explaining the apparent two-state behavior observed in Figures 2 and 4–6.

In assessing the effects of the three peptide linkers, it was observed that the relative order of conformational stability of the scFv's, 4.4.20/205 > 4.4.20/212 > 4.4.20/202′ (see Table I), also parallels the relative rates of folding: 1.8, 1.0, and 0.67, respectively (212 reference). This suggests that a part of the stabilizing effect of the 205 linker on the scFv protein results from destabilization of the U form of this protein. Indeed, the $\Delta(\Delta G^{\ddagger}_{u-n})$ calculated for 205 [$\Delta G^{\ddagger}_{u-n} = -RT \ln (kh/k_BT)$] was found to be 0.31 kcal/mol (280 K), which agrees with the $\Delta(\Delta G°_{n-u})$ or $\Delta(\Delta G°_{u-n})$ values that appear in Table I. The destabilizing effect of the 202′ linker relative to that of 212 can also be partially explained in terms of the U form of these scFv proteins, but the magnitude of the stability loss, −1.25 kcal/mol, is much larger than could be explained on the basis of the rates. Therefore, the larger part of $\Delta(\Delta G°_{n-u})$ observed for the 4.4.20/202′ scFv is probably due to an effect on the N state, perhaps due to steric strain introduced by a minimal-length linker.

The apparent close fit of the experimental unfolding/refolding transitions to two-state behavior is a necessary precondition for the support of a simple two-state model but is not sufficient for proof. With sufficient quantities of scFv proteins now available, it may be possible to obtain further evidence either for or against the two-state model through differential scanning calorimetry (DSC) or other physical techniques which are more sensitive and better suited for the thermodynamic analysis of multidomain proteins (Brandts et al., 1989; Pantoliano et al., 1989).

We conclude by observing that the scFv two-domain minimal strategy for synthetic antibodies appears to have advantages over the Fab, Fv, and single V domain designs, especially in the area of protein folding and stability. One advantage is that the unfolding transitions are unimolecular, independent of [protein], while the noncovalent heterodimer association of Fv's and the homodimer association of single V domains are expected to yield unfolding transitions that are [protein] dependent (Rowe, 1976). Moreover, scFv proteins appear to follow reversible two-state behavior during solvent-induced unfolding/refolding equilibria, also suggesting markedly simplified thermodynamic models, especially when compared with Fab's (Rowe, 1976). These advantages afforded the scFv design, presumably through the entropic effect of the peptide linker on ΔG_{AB}, will make them desirable targets for protein folding and stability experiments. The thermodynamic stability of scFv proteins will be of fundamental importance for their utility as frame-work strategies for the delivery of drugs (Chaudhary et al., 1989), as tools in diagnostic imaging applications (Colcher et al., 1990), and as catalysts in organic synthesis (Wong et al., 1990).

Acknowledgments

We gratefully acknowledge R. Ladner for his early scFv protein designs which were prototypes of those described in this study. We thank P. Bryan and M. Whitlow for their help in establishing binding assays and D. Filpula for DNA sequencing. We thank F. Van-Catledege of Du Pont CR & D, Scientific Computing Division, for setting up the nonlinear least-squares fit of the unfolding/refolding reactions on the Du Pont Cray X-MP computer. We also thank J. Krywko for his help in generating the computer-graphic reproduction shown in Figure 1A.

References

Aune, K. C., & Tanford, C. (1969a) *Biochemistry 8*, 4579–4585.

Aune, K. C., & Tanford, C. (1969a) *Biochemistry 8*,

[13] Indeed, the V_L and C_L domains of Oku κ chain were found to be $\Delta G°_{n-u}$ = 6.3 ± 0.2 and 5.2 ● 0.2 kcal/mol, respectively, when evaluated by the D binding method (Tsunenaga et al., 1987).

[14] Q = 40% is used for $Q_{max}/2$ instead of 48% because of the correction for the scFv·FL complex in calculations of $[FL]_{free}$ for Q vs log $[FL]_{free}$ plots.

4586–4590.
Bedzyk, W. D., Johnson, L. S., Riordan, G. S., & Voss, E. W., Jr. (1989) *J. Biol. Chem. 264*, 1565–1569.
Bedzyk, W. D., Weidner, K. M., Denzin, L. K., Johnson, L. S., Hardman, K. D., Pantoliano, M. W., Asel, E. D., & Voss, E. W., Jr. (1990) *J. Biol. Chem. 265*, 18615–18620.
Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535–542.
Better, M., Chang, C. P., Robinson, R. R., & Horwitz, A. H. (1988) *Science 240*, 1041–1043.
Bird, R. E., Hardman, K. D., Jacobson, J. W., Johnson, S., Kaufman, B. M., Lee S.-M., Lee, T., Pope, H. S., Riordan, G. S., & Whitlow, M. (1988) *Science 242*, 423–426.
Brandts, J. F., Hu, C. Q., Lin, L.-N., & Mas, M. T. (1989) *Biochemistry 28*, 8588–8596.
Chaudhary, V. K., Queen, C., Junghans, R. P., Waldmann, T. A., FitzGerald, D. J., & Pastan, I. (1989) *Nature 339*, 394–397.
Colcher, D., Bird, R. E., Roselli, M., Hardman, K. D., Johnson, S., Pope, S., Dodd, S., Pantoliano, M. W., Milenic, D., & Schlom, J. (1990) *J. Natl. Cancer Inst. 82*, 1191–1197.
Condra, J. H., Sardana, V. V., Tomassini, J. E., Schlabach, A. J., Davies, M. E., Lineberger, D. W., Graham, D. J., Gotlib, L., & Colonno, R. J. (1990) *J. Biol. Chem. 265*, 2292–2295.
Creighton, T. E. (1984) in *Proteins: Structures and Molecular Properties*, Freeman, New York
Edman, P., & Begg, G. (1967) *Eur. J. Biochem. 1*, 80–91.
Gibson, A. L., Herron, J. N., He, X.-M., Patrick, V. A., Mason, M. L., Lin, J.-N., Kranz, D. M., Voss, E. W., Jr., & Edmundson, A. B. (1988) *Proteins: Struct., Funct. Genet. 3*, 155.
Glockshuber, R., Malia, M., Pfitzinger, I., & Plückthun, A. (1990) *Biochemistry 29*, 1362–1367.
Goto, Y., & Hamaguchi, K. (1979) *J. Biochem. (Tokyo) 86*, 1433–1441.
Goto, Y., & Hamaguchi, K. (1986) *Biochemistry 25*, 2821–2828.
Goto, Y., & Hamaguchi, K. (1987) *Biochemistry 26*, 1879–1884.
Goto, Y., Ichimura, N., & Hamaguchi, K. (1988) *Biochemistry 27*, 1670–1677.
Haber, E., Quertermous, T., Matsueda, G. R., & Runge, M. S. (1989) *Science 243*, 51–56.
Handbook of Chemistry and Physics (1974–1975) 55 ed., CRC Press, Boca Raton, FL.
Herron, J. N., He, X.-M., Mason, M. L., Voss, E. W., Jr., & Edmundson, A. B. (1989) *Proteins: Struct., Funct. Genet. 5*, 271–280.
Hochmann, J., Gavish, M., Inbar, D., & Givol, D. (1976) *Biochemistry 15*, 2706–2710.
Horne, C., Klein, M., Polidoulis, I., & Dorrington, K. J. (1982) *J. Immunol. 129*, 660–664.
Horwitz, A. H., Chang, C. P., Better, M., Hellstrom, E., & Robinson, R. R. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 8678–8682.
Huston, J. S., Levinson, D., Mudgett-Hunter, M., Tai, M.-S., Novotny, J., Margolies, M. N., Ridge, R. J., Bruccoleri, R. E., Haber, E., Crea, R., & Oppermann, H. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 5879–5883.
Jones, T. A. (1978) *J. Appl. Crystallogr. 11*, 268.
Kabat, E. A., Wu, T. T., Reid-Miller, M., Perry, H. M., & Gottesman, K. S. (1987) in *Sequences of Proteins of Immunological Interest*, 4th ed., U.S. Department of Health and Human Services, Washington, DC.
Kellis, J. T., Jr., Nyberg, K., & Fersht, A. R. (1989) *Biochemistry 28*, 4914–4922.
Klein, M., Kortan, C., Kells, D. I. C., & Dorrington, K. J. (1979) *Biochemistry 18*, 1473.
Laemmli, U. K. (1970) *Nature 227*, 680–685.
Matsumura, M., Becktel, W. J., Levitt, M., & Matthews, B. W. (1989a) *Proc. Natl. Acad. Sci. U.S.A. 86*, 6562–6566.
Matsumura, M., Signor, G., & Matthews, B. W. (1989b) *Nature 342*, 291–293.
Mitraki, A., & King, J. (1989) *Bio/Technology 7*, 690–697.
Movva, N. R., Nakamura, K., & Inouye, M. (1980) *J. Mol. Biol. 143*, 317–328.
Mutter, M. (1977) *J. Am. Chem. Soc. 99*, 8307–8314.
Pace, C. N. (1986) *Methods Enzymol. 131*, 266–280.
Pace, C. N. (1990) *Trends Biotechnol. 8*, 93–98.
Pantoliano, M. W., Whitlow, M., Wood, J. F., Dodd, S. W., Hardman, K. D., Rollence, M. L., & Bryan, P. N. (1989) *Biochemistry 28* 7205–7213.
Privalov, P. L., & Potekhin, S. A. (1986) *Methods Enzymol. 131*, 4–51.
Ramsay, G., & Freire, E. (1990) *Biochemistry 29*, 8677–8683.
Riddles, P. W., Blakeley, R. L., & Zerner, B. (1983) *Methods Enzymol. 91*, 49–60.
Riechmann, L., Foote, J., & Winter, G. (1988) *J. Mol. Biol. 203*, 825–828.
Rowe, E. S. (1976) *Biochemistry 15*, 905–916.
Rowe, E. S., & Tanford, C. (1973) *Biochemistry 12*, 4822–4827.
Santoro, M. M., & Bolen, D. W. (1988) *Biochemistry 27*, 8063–8068.
Satow, Y., Cohen, G. H., Padlan, E. A., & Davies, D. R. (1986) *J. Mol. Biol. 190*, 593–604.
Scandella, D., Arthur, P., Mattingly, M., & Neuhold, L. (1985) *J. Cell. Biochem. 9B*, 203.
Scatchard, G. (1949) *Ann. N.Y. Acad. Sci. 51*, 660–672.
Schultz, P. G., Lerner, R. A., & Benkovic, S. J. (1990) *Chem. Eng. News* (May 28 issue).
Skerra, A., & Plückthun, A. (1988) *Science 240*, 1038–1041.
Tanford, C. (1970) *Adv. Protein Chem. 24*, 1–95.
Thannhauser, T. W., Konishi, Y., & Scheraga, H. A. (1984) *Anal. Biochem. 138*, 181–188.
Tsunenaga, M., Goto, Y., Kawata, Y., & Hamaguchi, K. (1987) *Biochemistry 26*, 6044–6051.
Voss, E. W., Jr. (1984) in *Fluorescein Hapten: An Immunological Probe* (Voss, E. W., Jr., Ed.) CRC Press, Boca Raton, FL.
Ward, E. S., Güssow, D., Griffiths, A. D., Jones, P. T., & Winter, G. (1989) *Nature 341*, 544–546.
Winter, G., & Milstein, C. (1991) *Nature 349*, 293–299.
Wong, C.-H., Chen, S.-T., Hennen, W. J., Bibbs, J. A., Wang, Y.-F., Liu, J. L.-C., Pantoliano, M. W., Whitlow, M., & Bryan, P. N. (1990) *J. Am. Chem. Soc. 112*, 945–953.