Thermodynamic Analysis of the Folding of the Streptococcal Protein G IgG-Binding Domains B1 and B2: Why Small Proteins Tend To Have High Denaturation Temperatures†

Patrick Alexander, Stephen Fahnestock,‡ Timothy Lee,§ John Orban, and Philip Bryan*
Center for Advanced Research in Biotechnology, University of Maryland, 9600 Gudelsky Drive, Rockville, Maryland 20850
Received October 22, 1991; Revised Manuscript Received January 14, 1992

ABSTRACT: We have cloned, expressed, and characterized two naturally occurring variations of the IgG-binding domain of streptococcal protein G. The domain is a stable cooperative folding unit of 56 amino acids, which maintains a unique folded structure without disulfide cross-links or tight ligand binding. We have studied the thermodynamics of the unfolding reaction for the two versions of this domain, designated B1 and B2, which differ by six amino acids. They have denaturation temperatures of 87.5 °C and 79.4 °C, respectively at pH 5.4, as determined by differential scanning calorimetry. Thermodynamic state functions for the unfolding reaction (ΔG, ΔH, ΔS, and ΔC_p) have been determined and reveal several interesting insights into the behavior of very small proteins. First, though the B1 domain has a heat denaturation point close to 90 °C, it is not unusually stable at physiologically relevant temperatures (ΔG = 25 kJ/mol at 37 °C). This behavior occurs because the stability profile (ΔG vs temperature) is flat and shallow due to the small ΔS and ΔC_p for unfolding. Related to this point is the second observation that small changes in the free energy of unfolding of the B-domain due to mutation or change in solvent conditions lead to large shifts in the heat denaturation temperature. Third, the magnitude and relative contributions of hydrophobic vs nonhydrophobic forces (per amino acid residue) to the total free energy of folding of the B-domain are remarkably typical of other globular proteins of much larger size.

Protein G is a multidomain component of the cell wall of streptococcal species of the Lancefield group G. It binds to a variety of mammalian IgG[1] by the constant Fc portion of the immunoglobulin in a manner independent of the antigen combining sites. (Myhre & Kronvall, 1977, 1981; Reis et al., 1984). Several protein G genes have been cloned and sequenced (Fahnestock et al., 1986; Filpula et al., 1987; Guss et al., 1986; Olsson et al., 1987). The IgG-binding domains have been shown to consist of 2–3 repeats of 55 amino acids separated by about 15 unique amino acids (Fahnestock et al., 1986; Guss et al., 1986; Åkerstrom et al., 1987; Sjöbring et al., 1991). Protein G is organized into functional domains similar to staphylococcal protein A (Åkerstrom & Bjorck, 1986), and the IgG-binding domains are similar in size. For a review of the cloning and characterization of IgG receptors, see Fahnestock et al. (1990), Fahnestock and Alexander (1990), Bjorck and Åkerstrom (1990), and Reis and Boyle (1990). The basic folding pattern of the 55 amino acid domain of protein G consists of a four-stranded β-sheet spanned by an α-helix (Lian et al., 1991; Gronenborn et al., 1991). For the NMR assignments and secondary structure of the B2 domain, see Orban et al. (1992). No homology exists between the IgG-binding domains of protein A and protein G on the levels of either primary or tertiary structure.

We have cloned and characterized the protein product of two 56 amino acid (55 aa plus N-terminal methionine) IgG-binding domains denoted B1 and B2. These proteins differ by six amino acid substitutions, contain no disulfide bonds, and have reversible two-state unfolding transitions at pH 5.4 of 87.5 °C and 79.4 °C, respectively [see Figure 8 in Orban et al. (1992) for the positions of the amino acid substitutions in B2 relative to B1]. Here we report some physical properties of B1 and B2 including analysis by circular dichroism spectroscopy (CD) and differential scanning calorimetry (DSC), and we examine the energetics of stabilizing such a small cooperative folding unit.

Materials and Methods

Cloning and Expression

Escherichia coli TG1 [K12, Δ(*lac–pro*), *supE*, *thi*, *hsdD5*/F'traD36, *proA*$^+$*B*$^+$, *lacI*^{q}, *lacZ*ΔM15] was obtained from the oligonucleotide-directed in vitro mutagenesis system version 2, (Amersham International plc). *E. coli* BL21(DE3) [F$^-$, *hsdS*, *gal*] has been described previously (Studier & Moffatt, 1986). All restriction enzymes and T4 DNA ligase were purchased from New England Biolabs, Inc., and alkaline phosphatase was from Boehringer Mannheim Biochemicals. Agarose for DNA gel electrophoresis was obtained from FMC Bioproducts. Deoxyribonuclease I (DNase I) and phenylmethanesulfonyl fluoride (PMSF) were obtained from Sigma Chemical Co.; isopropyl β-D-thiogalactopyranoside (IPTG) was from Gold Biotechnology.

DNA Amplification. DNA segments were amplified using the polymerase chain reaction in an Eppendorf MicroCycler

† This work was supported by the NIH (GM 42560) and the Maryland Biotechnology Institute.
* Corresponding author.
‡ Present address: Central Research and Development, Du Pont Corp., Wilmington, DE.
§ Present address: Genex Corp., Gaithersburg, MD.

[1] Abbreviations: CD, circular dichroism; C_{ex}, excess specific heat as measured by calorimetry; ΔH_{cal}, calorimetric enthalpy for unfolding; ΔH_{vH}, van't Hoff enthalpy for unfolding; DSC, differential scanning calorimetry: EDTA, disodium salt of ethylenediaminetetraacetic acid; Hepes, *N*-(2-hydroxyethyl)piperazine-*N*'-2-ethanesulfonic acid; IEF, isoelectric focusing; IgG, immunoglobulin G; IPTG, isopropyl β-D-thiogalactopyranoside; NMR, nuclear magnetic resonance; [P], protein concentration; PAGE, polyacrylamide gel electrophoresis; PCR; polymerase chain reaction; PMSF, phenylmethanesulfonyl fluoride; *T*, temperature; T_m, temperature of thermal melting.

0006-2960/92/0431-3597$03.00/0

according to conditions outlined in the GeneAmp PCR reagent kit (Perkin Elmer Cetus). Oligonucleotides (5′-CGTGACACATATGACTTACAAACTTG-3′) (5′-GTAACCATGGATCCTTATTCAGTTACCGTAAAG-3′) and (5′-GATCGACCATATGACTTACAAATTAATCC-3′) (5′-CTGGTTTGGATCCTTATTCAGTAACTGTAAAG-3′) were synthesized on a 380B DNA synthesizer (Applied Biosystems). The underlined segments are mismatches that encode unique *Nde*I sites in the first of each pair which can be precisely fused to the ATG initiation codon of the expression plasmid and unique *Bam*HI sites downstream of a stop codon in the second of each pair. The oligonucleotides amplify a 195 base pair segment of DNA that contains the coding sequence for a single IgG-binding domain. Digestion of the amplified product with the appropriate restriction enzymes facilitates cloning and enables a precise excision of the DNA that codes for a 55 amino acid IgG-binding domain of the molecule. DNA was amplified for 24 cycles, each cycle defined as a 94 °C melt for 1 min, 42 °C annealing for 1 min, and 72 °C extension for 1.5 min. After the amplification was complete, the DNA was phenol extracted and ethanol precipitated.

Expression Plasmid. pT7-7 was a generous gift from Stanley Tabor of Harvard Medical School. pT7-7 contains bacteriophage T7 RNA polymerase promoter ϕ10 and the translation start site for gene 10. It is the basis for expression in this study with the following modification. An 880 base pair *Nru*I–*Sal*I restriction fragment from pGX 2627 (Scandella et al., 1985) containing the M13 origin of replication and λ 4S transcription terminator was excised from low-gelling agarose. pT7-7 was digested at the unique *Hin*dIII site and filled in according to the Sequenase protocol (U.S. Biochemicals). The resulting linear DNA was digested with *Sal*I, phenol extracted, and ethanol precipitated. The *Nru*I–*Sal*I M13 fragment was mixed with the digested pT7-7 and ligated according to standard procedures. TGI cells were made competent and transformed as previously described (Maniatis, 1962). The resulting plasmid was called pG5.

Production of Single-Strand DNA. TGI cells harboring pG5 were checked for their ability to produce single-strand DNA upon superinfection before cloning of gene fragments. Cells were grown and superinfected with M13K07, and single-strand plasmid DNA was isolated essentially as described (Vieira & Messing, 1987).

Subcloning of Amplified DNA. pG5 and the amplified DNA products were digested with *Nde*I and *Bam*HI. The vector's 5′ phosphates were removed using calf intestine alkaline phosphatase (Boehringer Mannheim Biochemicals). After phenol extraction and ethanol precipitation, the plasmid and fragments were mixed and ligated according to standard molecular biology techniques. TGI cells were transformed as before, and transformants were screened for the correct size *Nde*I–*Bam*HI fragments. Correct plasmids were designated pG6 and pG7.

DNA Sequencing. Single-strand plasmid DNA was isolated as described above, and sequencing of inserts was done according to the Sequenase protocol (U.S. Biochemical). The two domains have identical amino acid sequences to the B1 and B2 domains reported by Fahnestock et al.

Fermentation and Expression. The production strain, BL21(DE3), was made competent and transformed. Ten milliliters of L broth (10 g of tryptone, 5 g of yeast extract, 10 g/L NaCl) supplemented with 100 μg/mL ampicillin was inoculated with two ampicillin-resistant colonies in a 250-mL baffled flask. The culture was grown at 37 °C, 300 rpm, until mid log phase. This culture was used to inoculate 1.5 L of L broth buffered with 2.3 g of KH_2PO_4, 12.5 g of K_2HPO_4/L supplemented with 150 μg/mL ampicillin. The culture was grown at 37 °C in a BioFlo Model C30 fermenter (New Brunswick Scientific Co., Inc.) until an A_{600} 1–1.5 was attained, upon which 1 mM IPTG was added to induce the production of T7 RNA polymerase that directs synthesis of target DNA message. Four hours after induction, the cells were harvested by centrifugation for 30 min at 6000 rpm in a J2-21 centrifuge (Beckman Instruments).

Protein Purification. E. coli paste from a 1.5-L fermentation (5 g) was suspended in 50 mL of cold phosphate-buffered saline (PBS), and PMSF was added to a final concentration of 1 mM. DNase I (1 mg) in 2 mL of 40 mM Tris-HCl, 1 M $MgCl_2$, was also added. This suspension was heated to 80 °C for 5 min. After the reaction was cooled on ice, another addition of PMSF and DNase I was made. This mixture was stirred at room temperature for 30 min and centrifuged at 25000g for 30 min. The supernatant was filtered through a 0.45-μm membrane and chromatographed on a 2.5 × 20 cm column containing 70 mL of human IgG–Sepharose (Pharmacia). The column-bound protein was washed extensively with PBS followed by 5 mM ammonium acetate, pH 5.0. Protein G was eluted with 0.5 M ammonium acetate, pH 3.0. The purified protein was dialyzed extensively against 2 mM ammonium bicarbonate, pH 7.0, and lyophilized. Five grams of *E. coli* paste typically yields 150 mg of purified protein G B-domain.

Polyacrylamide Gel Electrophoresis. Protein purity was checked by discontinuous sodium dodecyl sulfate (SDS)–polyacrylamide gel electrophoresis (PAGE) on 18% running gels (Novex).

Isoelectric Focusing (IEF). The isoelectric points of the B1- and B2-domains were determined on native PAGE gels using lysine and phosphoric acid as cathode and anode buffers. The isoelectric points of B1 and B2 were estimated to be 4.0 and 4.7, respectively, by comparison with the following IEF standard proteins: acetylated horse cytochrome *c*, p*I* 3.95; phytocyanin, p*I*'s 4.75 and 4.85.

N-Terminal Analysis. The first five amino acids of the B1- and B2-domains were determined by sequential Edman degradation and HPLC analysis. This revealed that >90% of the material had the N-terminal methionine present. Otherwise, the amino acid sequences were as expected from the DNA sequence of the genes.

Determination of Extinction Coefficient. The extinction coefficient was determined by monitoring the fluorescence as described by Pajot (1976). The value obtained was A_{280} of 1.32 = 1 mg/mL.

Differential Scanning Calorimetry (DSC). DSC measurements were performed with a Hart 7707 DSC heat conduction scanning microcalorimeter interfaced with an IBM personal computer (Bryan et al., 1986; Pantoliano et al., 1988). The temperature was increased from 20 to 110 °C at a scan rate of 30 °C/h for the majority of experiments, but also at 60 °C/h for a few selected studies. The solution mass of all protein and control solutions was near 0.70 g/ampoule. Each experiment was comprised of four segments: (1) The first upward scan from 20 to 110 °C; (2) the first downward scan from 110 to 20 °C; (3) a second upward scan from 20 to 110 °C; and (4) a second downward scan returning to 20 °C. The power input from an upward scan of buffer vs buffer was subtracted from the first upward scan to obtain the excess power input for the unfolding transitions. The excess power thermal scans were converted to excess heat capacity vs T scans by dividing by the scan rate (Schwarz & Kirchhoff, 1988).

Table I: Thermodynamic Data for the B1- and B2-Domains of Protein G[a]

	pH	T_m (°C)	ΔH_{cal} (kJ/mol)	ΔH_{vH} (kJ/mol)	$\Delta H_{vH}/\Delta H_{cal}$	ΔS [J/(deg·mol)]	ΔC_p [kJ/(deg·mol)]	$\Delta\Delta G$ (kJ/mol)
B1	2.30	55.7 ± 0.3	178 ± 6	176 ± 6	0.99 ± 0.01	540 ± 19		
	3.51	79.8 ± 0.4	237 ± 4	229 ± 7	0.97 ± 0.04	672 ± 12		
	5.40	87.5 ± 0.1	258 ± 3	262 ± 6	1.02 ± 0.01	717 ± 12		
							2.6 ± 0.3	
B2	2.69	58.4 ± 0.3	169 ± 5	175 ± 5	1.04 ± 0.03	509 ± 15		
	2.88	64.0 ± 0.1	195 ± 1	190 ± 1	0.97 ± 0.01	579 ± 1		
	3.10	65.8 ± 0.7	189 ± 8	198 ± 4	1.05 ± 0.02	557 ± 21		
	4.00	76.1 ± 0.4	205 ± 11	214 ± 12	1.04 ± 0.01	587 ± 31		
	5.40	79.4 ± 0.2	238 ± 5	240 ± 5	1.01 ± 0.01	672 ± 10		−5.5
							2.9 ± 0.2	

[a] Thermodynamic parameters are determined as described in the text at T_m for the two domains. $\Delta\Delta G$ is the ΔG of unfolding for B2 minus the ΔG of unfolding for B1 determined at 80 °C. ΔC_p was determined from the slope of the ΔH_{cal} vs T plot as shown in Figure 3. The ΔH_{vH} values were determined by curve fitting using the exam program (Schwarz & Kirchof, 1988). The ΔC_p value of 2.9 kJ/(deg·mol) was included in the fitting procedure.

The scans were done in 50 mM NaOAc or 50 mM glycine buffer. Samples to be scanned were prepared by rehydrating lyophilized protein in the appropriate buffer and dialyzing against the same buffer. The concentration of the dialyzed protein was determined by UV absorbance using the 1 mg/mL = A_{280nm} of 1.32. The number of nanomoles of protein ranged from 290 to 490, corresponding to 3–5 mg/mL.

A sigmoidal baseline, C_{ps}, was extrapolated under the thermal transitions using a least-squares computer fit of the pre- and posttransitional baselines to linear equations in T and a, the fractional area under the transition curve at any particular T, as described by Schwarz (1988). The calorimetric enthalpy of unfolding, ΔH_{cal}, was determined from the area under the transition curve and the [P] in the ampoule. The unfolding transition temperature, T_m, was measured as the temperature at $a = 0.5$. The change in heat capacity upon unfolding, ΔC_p, was determined by measuring ΔH as a function of temperature (varying the pH between 11.3 and 2.3) using the equation $\Delta H = \Delta H_0 + \Delta C_p(T - T_0)$.

Results

Calorimetry. The thermal unfolding transitions of protein G, B1, and B2 were studied using DSC. A typical DSC scan of the unfolding transition for B2 is shown in Figure 1. When subtracted from the buffer scan (Figure 1B), a single symmetrical transition is observed with a midpoint of 64.0 °C at pH 2.88. A computer program developed by Schwarz and Kirchhoff (1988) was used to fit a theoretical curve to the experimental data by varying the adjustable parameters, T_m and ΔH_{vH} (2) to minimize the standard deviation of the observed values of excess heat capacity, C_{ex}, from the calculated curve. The fitted curve in Figure 1 is nearly superimposable on the experimental DSC scan and was obtained assuming a simple two-state model for unfolding, N ⇔ U.

The amount of excess heat absorbed by a protein sample as T is increased through a transition from the folded to unfolded state at constant pressure provides a direct measurement of the ΔH of unfolding (Privalov & Potekhin, 1986). The unfolding transitions for both B1- and B2-domains fit a two-state model and follow the dictates of equilibrium thermodynamics as expressed in the van't Hoff equation, $d \ln K/dT = \Delta H_{vH}/(RT^2)$, with ΔH_{vH}, the van't Hoff enthalpy, or apparent enthalpy, equal to the calorimetric, ΔH_{cal}, or true enthalpy. Thermodynamic data for the two domains are summarized in Table I. The maximum T_m for both domains is observed around pH 5.4. At this pH, the melting temperature of B1 is 87.5 ± 0.1 °C and of B2 is 79.4 ± 0.2 °C. The difference in free energy of unfolding between the two domains can be approximated with the expression $\Delta\Delta G = \Delta T_m(\Delta S)$ (Becktel & Schellman, 1987) and is equal to 5.5 kJ/mol around the melting temperatures.

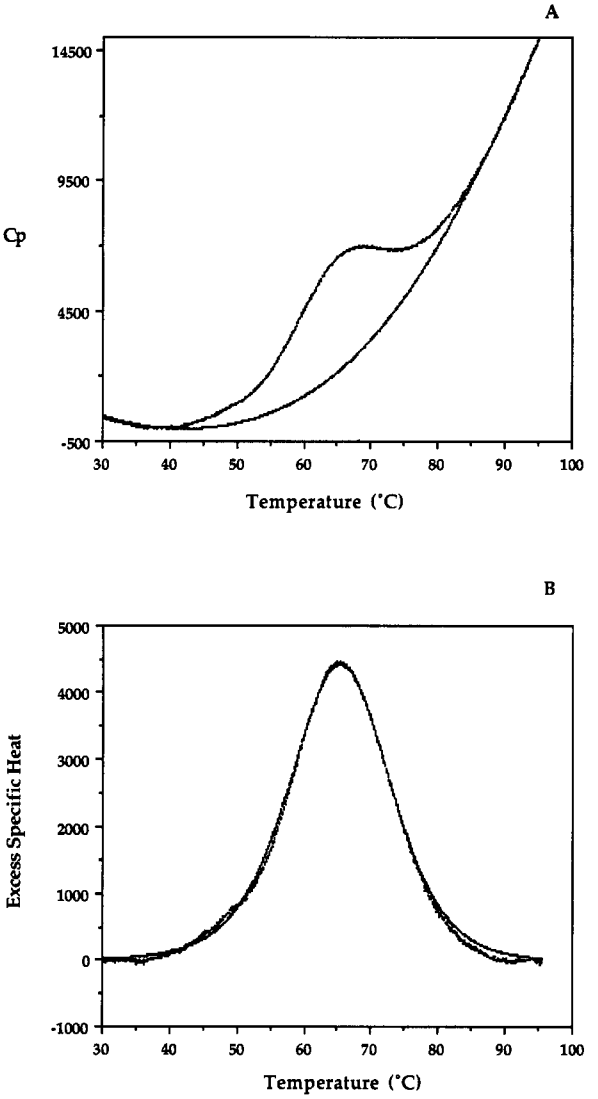

FIGURE 1: Differential scanning calorimetry. (A) Raw calorimetric data of B2 in 50 mM glycine, pH 2.88, and the corresponding buffer baseline. (B) The subtracted calorimetric data from panel A and a theoretical two-state fit are shown (T_m = 64.0 °C). The experimentally measured excess heat capacity was fit to a computer-derived theoretical curve that simulates a two-state unfolding process not involving association or dissociation (Schwarz & Kirchhoff, 1988). Excess heat capacity is in units of microjoules per degree. The calorimeter ampoule contained 2.1 mg of protein.

The unfolding reaction is quantitatively reversible throughout the transition. This has been determined in the DSC by reversing the temperature scan and also by monitoring

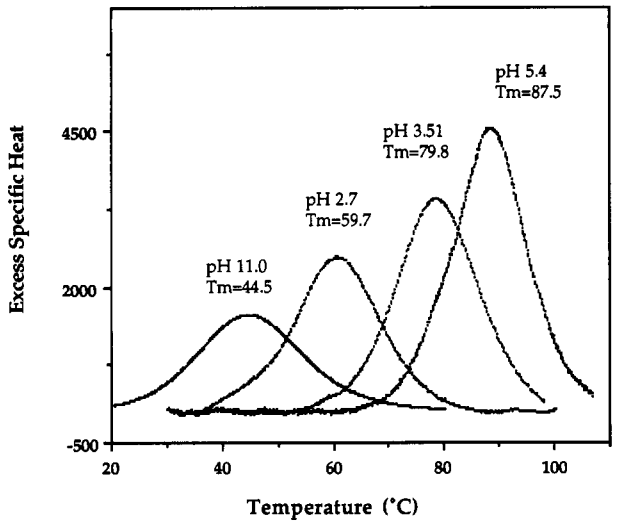

FIGURE 2: Calorimetric scans of B1 taken in 50 mM sodium acetate, pH 5.40 and 3.51, and 50 mM glycine, pH 2.70, and 50 mM potassium phosphate, pH 11.0. Excess heat capacity is in units of microjoules per degree. The calorimeter ampoules contained 2.1 mg of protein.

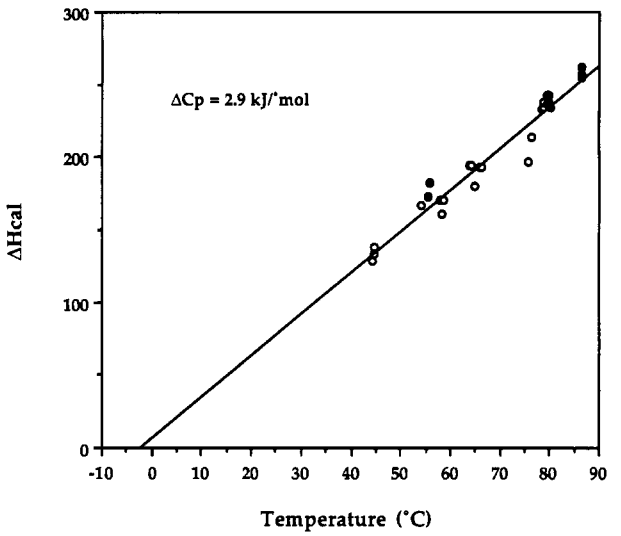

FIGURE 3: Plot of ΔH_{cal} for B1-domain (●) and B2-domain (○) vs temperature. Calorimetric experiments were carried out at pH's ranging from 2.3 to 11.0, such that the temperature of denaturation was shifted between 44 and 87 °C. Each point represents one calorimetric experiment. The line was fitted by linear regression using only data points for B2 (○).

unfolding and refolding in the CD. After scanning to 110 °C to obtain good posttransitional baselines, however, less than 25% of the protein melts at the same temperature in a second heating scan.

Temperature Dependence of the Free Energy of Unfolding. For globular proteins, the heat capacity of the unfolded state is significantly larger than the heat capacity of the folded state (Privalov, 1979; Becktel & Schellman, 1987). This heat capacity difference (ΔC_p) has been shown to be precisely correlated with the increased exposure of hydrophobic surface area upon unfolding (Privalov & Gill, 1988; Livingstone et al., 1991). If ΔC_p is known, the free energy of unfolding can be calculated at temperatures beyond the experimentally measurable range using the equation

$$\Delta G = \Delta H_0 - T\Delta S_0 + \Delta C_p[T - T_0 - T \ln (T/T_0)] \quad (1)$$

where ΔH_0 and ΔS_0 are the enthalpy and entropy of unfolding evaluated at a reference temperature T_0, and ΔG is the free energy of unfolding at a temperature T (Brandts, 1964; Pace & Tanford, 1968; Privalov & Khechinashvili, 1974; Privalov, 1979; Becktel & Schellman, 1987).

To calculate ΔC_p according to the equation $\Delta H = \Delta H_0 + \Delta C_p(T - T_0)$, ΔH was measured at a number of reference temperatures by varying the pH in calorimetric experiments (Figure 2). Glycine or acetate buffers were used to cancel out the heat of ionization of carboxylate groups upon unfolding, so that ΔH will be strictly a function of the temperature at which it is measured. The B-domains contain no histidines, which could become protonated during unfolding at acid pH, producing heat effects. The plots of ΔH vs temperature for B1 and B2 are shown in Figure 3. To improve the accuracy of the ΔC_p determination, ΔH was measured over a range of more than 40 °C. Within the accuracy of the measurements of ΔH, the plots are linear, indicating that ΔC_p is constant over this temperature range. ΔC_p from these plots is 2.9 kJ/(deg·mol) and, within experimental error, is the same for B1 and B2.

Using the reference values of ΔH_{cal} for B1 and B2 from Table I and the value of ΔC_p of 2.9 kJ/(deg·mol), ΔG and ΔH were calculated as a function of temperature (Figure 4). The stability profiles (ΔG vs T) of B1 and B2 are similar, reaching a maximum ΔG at about 5 °C. The maximum for B1 is 30 kJ/mol and for B2 is 25 kJ/mol. At this maximum $\Delta S = 0$

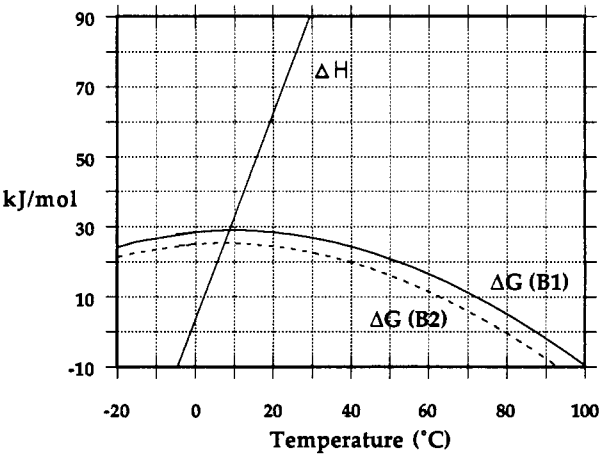

FIGURE 4: Free energy profile of B1 and B2. ΔG for unfolding is plotted vs temperature according to the equation $\Delta G = \Delta H_0 - T\Delta S_0 + \Delta C_p[T - T_0 - T \ln (T/T_0)]$, where ΔH_0 and ΔS_0 are the enthalpy and entropy of unfolding evaluated at a reference temperature T_0 (see text). Also shown is ΔH for unfolding vs temperature. The ΔH line intersects the ΔG curves ($\Delta G = \Delta H$) where ΔS is equal to 0. This is the temperature where ΔG is maximal.

and stabilization is entirely enthalpic (Becktel & Schellman, 1987). When evaluated at a common reference temperature, the difference in ΔH for B1 and B2 is not larger than the uncertainty in our measurements (±5 kJ/mol). Therefore, the difference of about 5 kJ/mol in ΔG of B1 and B2 could be due to small differences in ΔH or ΔS or both.

Paradoxically, in spite of their high T_m, the maximum stability of these domains is well below the average of 50 kJ/mol for single-domain proteins for which data are available (Privalov, 1988). The stability of these domains at 37 °C, the physiologically relevant temperature for *Streptococcus*, is only 20–25 kJ/mol. This is about 10 times the thermal energy (RT) and would therefore seem to be an adequate but not excessive free energy of stabilization.

Dependence of T_m on pH. The melting transition temperature for the B-domains decreases rapidly below pH 4.0 due to the increasing positive charge of the protein. The change in the number of protons bound to the unfolded state, n_u, and the folded state, n_f, can be determined over the pH range 2–4

Table II: Comparison of Thermodynamic Parameters of Globular Proteins[a]

	molecular weight	ΔH(25 °C) (kJ/mol per residue)	ΔS(110 °C) [J/(deg·mol) per residue]	ΔC_p [J/(deg·mol) per residue]
protein G B1	7179	1.4	16.1	53
parvalbumin	11 500	1.4	16.8	46
cytochrome *c*	12 400	0.65	17.8	67
ribonuclease A	13 600	2.4	17.8	44
hen lysozyme	14 300	2.0	17.6	52
staphylococcal nuclease	16 800	0.85	17.5	61
myoglobin	17 900	0.04	17.9	75
papain	23 400	0.93	17.0	60
β-trypsin	23 800	1.3	17.9	58
α-chymotrypsin	25 200	1.1	18.0	58
average		1.2 ± 0.7	17.4 ± 0.6	57 ± 9

[a] Data for the other proteins were taken from Privalov (1988).

from the dependence of T_m on pH, according to the following equation (Privalov, 1979):

$$n_u - n_f = (\Delta H_s/2.303RT_m^2)(\partial(T_m)/\partial(\mathrm{pH})) \quad (2)$$

Sufficient data points were obtained over this pH range for B2 to fit the T_m dependence on pH to the equation

$$T_m(\mathrm{Kelvin}) = 298.3 + 12.93\mathrm{pH} \quad (3)$$

Using ΔH = 195 kJ/mol and T_m = 338 °K at pH 3.0 and the slope from eq 3, we calculate that 1.1 proton/molecule is absorbed upon unfolding at pH 3.0.

Relative Contribution of the Hydrophobic Effect to the Stability of the B-Domains. One key question in studying these domains is to determine how such a small protein with limited potential for burying hydrophobic surface can fold at all without the benefit of disulfide cross-links or ligand binding. We have approached this question by using the liquid hydrocarbon model of the hydrophobic effect to calculate the contribution of the hydrophobic (hφ) and residual (res) forces to ΔH, ΔS, and ΔG of unfolding of the B-domains (Kauzman, 1959; Tanford, 1962, 1970; Chothia, 1974; Baldwin, 1986; Dill et al., 1989). Baldwin has suggested that $\Delta H_{h\phi}$ and $\Delta S_{h\phi}$ for a protein can be approximated using the expression

$$\Delta G_{h\phi} = \Delta C_p(T - T_h) + \Delta C_p T \ln (T/T_s) \quad (4)$$

where T_s is the temperature at which $\Delta S_{h\phi} = 0$ and T_h is the temperature where $\Delta H_{h\phi} = 0$. On the basis of the experimentally measured values for the transfer of various liquid hydrocarbons to water, T_s = 113 °C (386 K) and T_h = 22 °C (295 K) (Gill et al., 1976; Sturtevant, 1977). The calculation of the hydrophobic interactions in a protein assumes that ΔC_p is completely due to the hydrophobic effect and is constant with temperature (Baldwin, 1986; Privalov & Gill, 1988). Residual forces are calculated by subtracting hydrophobic forces from observed forces (obs).

We have compared our data to the values calculated by Baldwin for hen lysozyme (Sturtevant, 1977), a protein of 123 amino acids. As for lysozyme the ΔH_{res} favors folding at all temperatures but is much smaller for B1 (78 kJ/mol) than the ΔH_{res} of lysozyme (216 kJ/mol). If this nonhydrophobic enthalpy contribution is normalized per residue, then ΔH_{res}/residue is 1.4 kJ/mol for B1 and 1.8 kJ/mol for lysozyme. Privalov has compared the ΔH per residue for a number of single domain proteins at 25 °C. At this temperature, the hydrophobic contribution to the total enthalpy of unfolding is small according to the liquid hydrocarbon model. Table II compares the ΔH per residue at 25 °C for several proteins. The average value obtained is 1.2 kJ/mol per residue. B1 is roughly average in comparison. Clearly, there is not a large nonhydrophobic contribution to enthalpy to compensate for their small size. By comparing ΔC_p per residue (Table II), one observes that although ΔC_p for the B-domains is small compared to larger globular proteins when evaluated on a per molecule basis, it is remarkably large when compared on a per residue basis. Its ΔC_p of 53 J/mol per residue is comparable to ΔC_p for lysozyme, staphylococcal nuclease, trypsin, chymotrypsin, papain, subtilisin, and carbonic anhydrase, which have ΔC_p's in the range of 50–60 J/mol per residue but are 2–5 times the molecular weight of the B-domains. The magnitude of ΔC_p per residue indicates that the B-domain manages to overcome the unfavorably large surface to volume ratio expected with a small structure and bury a significant hydrophobic core in spite of its size (Richards, 1977).

The ΔS_{res} is equal to about 900 J/(deg·mol). Thus, the nonhydrophobic (conformational entropy) contribution per residue is 16.1 J/(deg·mol) which is probably within experimental error the same as values obtained by Baldwin for hen lysozyme and by Privalov for numerous proteins. This can also be observed from plots of ΔS_{obs} vs ln (T_s/T). When $T = T_s$, the entire ΔS_{obs} is due to nonhydrophobic contributions and extrapolates to 940 J/(deg·mol). We conclude that, in spite of its small size, the contribution of the hydrophobic interactions to the energetics of unfolding is substantial and typical of other globular proteins.

Circular Dichroism. The circular dichroism spectrum of the B2 domain reflects the distribution of 50% β-sheet and 30% α-helix observed from the analysis of secondary structure of this domain by NMR (Orban et al., 1992). The melting of B2 as followed by the loss of secondary structure in the CD closely parallels the results observed by DSC. Figure 5 shows the cooperative changes in CD spectra as a function of temperature. The thermally denatured state of B2 resembles a random-coil structure with a minimum at 200 nm. The melting midpoint at pH 4.0 for B2 determined after correcting for the pretransitional slope was about 75 °C. The ΔH_{vH} determined for these data was 202 kJ/mol at the melting point. Both these values are in reasonable agreement with DSC data (Table I).

Discussion

The 56 amino acid IgG-binding domain of protein G is a stable cooperative folding unit which maintains a stable, unique folded state without the benefit of disulfide cross-links or tight ligand binding. The temperature-induced unfolded state appears by CD analysis to be much like a random coil, and DSC has shown the unfolding reaction to precisely follow a two-state model. The thermodynamics of the unfolding reactions for the B1- and B2-domains of protein G have been analyzed. B1 and B2 differ from one another by six amino acid substitutions, and B1 has a T_m 8.1 °C higher and a ΔG of unfolding about 5 kJ/mol larger than that of B2. Since $\Delta\Delta G \sim \Delta T_m(\Delta S)$, note that for a small protein with a small ΔS value at T_m, that changes in the free energy of unfolding resulting from mutations or changes in solvent conditions such as pH will cause relatively large changes in T_m. For example, mutations in subtilisin BPN′ (MW = 27 700) which add 5 kJ/mol to the free energy of unfolding result in a less than a 4 °C increase in T_m (Pantoliano et al., 1988), whereas a 5 kJ/mol change in the free energy of folding between the B1- and B2-domains (MW = 7180) results in an 8 °C difference in the T_m's.

The B2-domain binds human IgG with about 7-fold higher affinity than B1, which corresponds to an increase of ~5 kJ/mol in the free energy of binding at 25 °C. B2 therefore may have traded about 5 kJ/mol in its free energy of unfolding for about 5 kJ/mol in binding energy to IgG. Comparison

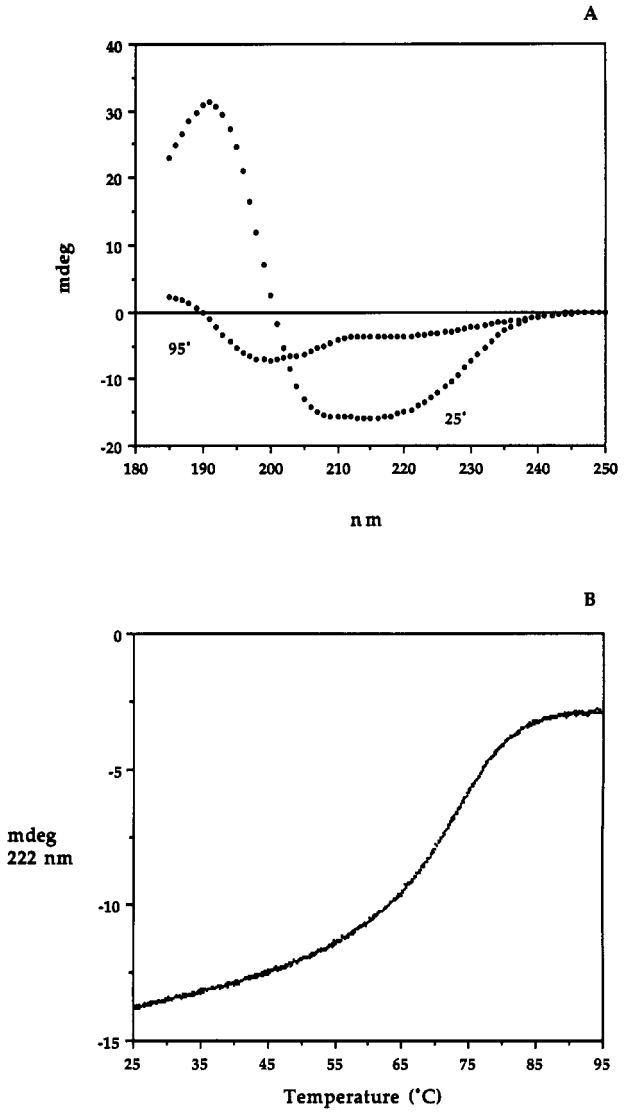

FIGURE 5: Thermal denaturation of B2 measured by CD. (A) CD spectra of 1 mg/mL B2 in 50 mM sodium acetate, pH 4.0, at 25 and 95 °C. (B) CD data points collected at 222 nm every 5 s, while the sample was heated at a rate of 1 °C/min from 25 to 95 °C. [P] was 1 mg/mL in 50 mM sodium acetate, pH 4.0. Spectra were obtained in a Jasco 720 spectropolarimeter using a water-jacketed cell of 0.1 mm path length. Midpoint of denaturation is ~75 °C.

of the solution structures of the B1- and B2-domains should illuminate the structural bases for the differences in stability and binding (Gronenborn et al., 1991; Orban et al., 1992).

The lower size limit for a stable folding domain has been postulated to be about 50 amino acids (Privalov, 1988). This estimate is based on the observation that the free energy of folding for single-domain globular proteins rarely exceeds about 500 J/mol per residue. Therefore, in order to obtain an overall ΔG of unfolding which is 10 times greater than the energy due to thermal motion at 25 °C, about 50 residues would be required. This prediction holds remarkably well for B1 and B2. Their ΔG's for unfolding at 25 °C are about 25–30 kJ/mol which is about 500 J/mol per residue. In comparison with other single-domain proteins, they are therefore typical in terms of ΔG per residue and somewhat marginally stable in terms of ΔG per molecule. Since melting temperature is often used synonymously with stability, it may thus seem ironic that B1, in particular, has a such high denaturation temperature (87.5 °C). The reason for the high melting temperature, however, is that the slope of the stability curve (ΔG vs T) is equal to $-\Delta S$ and the curvature is determined by $\Delta C_p/T$ (Schellman & Becktel, 1987). For a very small protein, where ΔS and ΔC_p per mole of protein are small, the stability curve will be broader and more shallow than the curve for a larger protein. Thus, if a small globular protein has the same stability at 25 °C as a large single-domain protein, the small protein will usually have a higher thermal denaturation point. We note that several very small proteins have high heat denaturation temperatures: BPTI (58 aa, T_m = 100 °C); neurotoxin II (62 aa, T_m = 96 °C; Creighton, 1986); parvalbumin (100 aa, T_m = 90 °C; Filimonov et al., 1978). Therefore, the high unfolding temperature of B1 and B2 does *not* reflect a high stability at physiologically relevant temperature (e.g., 37 °C) but, instead, is a consequence of the small size of the folding unit. On the contrary, the B-domains seem to have acquired only a minimum stability required for function at 37 °C (20–25 kJ/mol).

The increase in heat capacity of the unfolded state is directly related to the amount of hydrophobic surface which is buried in the folded state relative to the unfolded state (Livingstone et al., 1991). Although ΔC_p for the B-domains is small compared to larger globular proteins when evaluated on a per molecule basis, it is remarkably large when compared on a per residue basis. Apparently, the B-domain manages to bury a significant hydrophobic core in spite of its small size. We have calculated the hydrophobic and nonhydrophobic contributions to ΔH, ΔS, and ΔG for unfolding according to the treatment of Baldwin. Near 25 °C, the contribution of the hydrophobic effect to ΔH is close to zero, on the basis of hydrocarbon transfer data (Gill et al., 1976). Thus, the ΔH for unfolding at this temperature should have little contribution from the hydrophobic effect. The ΔH per residue at 25 °C for the B-domains is average compared to other single-domain proteins for which data are available. Contrary to our initial expectations, the B-domains, therefore, are not stable because of some unusual, nonhydrophobic contribution to ΔH, such as an unusually complete network of protein–protein hydrogen bonds. Rather, the B-domain appears to be stable because it efficiently buries hydrophobic surface in spite of its small size and concomitant unfavorable surface to volume ratio. One clear conclusion of this work is that the B-domain is remarkably similar to larger proteins in terms of ΔH, ΔS, and ΔG per residue and the relative contributions of hydrophobic and nonhydrophobic forces.

Our analysis of the thermodynamics of unfolding of the B-domain leads to three conclusions concerning the behavior of very small folding units relative to larger cooperative folding units. First, the high T_m of some very small proteins does not mean they are particularly stable at lower temperatures. The B1-domain has a heat denaturation point close to 90 °C, yet it has ΔG = 25 kJ/mol at 37 °C. Related to the first point is that changes in the free energy of unfolding of a small protein will lead to large shifts in the heat denaturation temperature. Finally, the forces stabilizing the folded state of the B-domain (hydrophobic vs nonhydrophobic contributions) are remarkably typical of other globular proteins of much larger size.

Streptococcus seems to have been able to achieve a remarkable economy in the design of the B-domains [Figure 9 in Orban et al. (1992)]. Likely, protein binding function is more easily designed into a compact globular structure than an enzymatic activity. The IgG-binding function of the protein G B-domain appears to easily be accommodated by convex surfaces on the exterior of the protein. For this reason, structural integrity and efficient hydrophobic packaging need not be compromised, allowing the IgG-binding function of

protein G to be designed into an very small domain.

ACKNOWLEDGMENTS

We thank Fredrick Schwarz for valuable assistance in the calorimetric experiments, Joel Hoskins for synthesizing the oligonucleotides used in the PCR reactions and DNA sequencing, and Kathryn Fisher and Mathew Mauro for advice on gene expression and protein production. We also acknowledge Michael Pantoliano for his preliminary calorimetric experiments on protein G.

REFERENCES

Åkerstrom, B., & Björck, L. (1986) *J. Biol. Chem. 261*, 10240–10247.

Åkerstrom, B., Nielsen, E., & Björck, L. (1987) *J. Biol. Chem. 262*, 13388–13391.

Baldwin, R. L. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 8069–8072.

Becktel, W. J., & Schellman, J. A. (1987) *Biopolymers 26*, 1859–1877.

Björck, L., & Kronvall, G. (1984) *J. Immunol. 133*, 969–974.

Björck, L., & Åkerstrom, B. (1990) in *Bacterial Immunoglobulin Binding Proteins* (Boyle, M. P. D., Ed.) Vol. 1, pp 113–126, Academic Press, San Diego.

Brandts, J. F. (1964) *J. Am. Chem. Soc. 86*, 4291–4301.

Bryan, P. N., Rollence, M. L., Pantoliano, M. W., Wood, J., Finzel, B. C., Gilliland, G. L., Howard, A. J., & Poulos, T. L. (1986a) *Proteins: Struct., Funct., Genet. 1*, 326–334.

Chothia, C. (1974) *Nature 248*, 338–339.

Creighton, T. E. (1986) *Methods Enzymol. 131*, 157–.

Dill, K. A., Alonso, D. O. V., & Hutchinson, K. (1989) *Biochemistry 28*, 5439–5449.

Fahnestock, S. R., & Alexander, P. (1990) in *Bacterial Immunoglobulin Binding Proteins* (Boyle, M. P. D., Ed.) Vol. 2, pp 417–424, Academic Press, San Diego.

Fahnestock, S. R., Alexander, P., Nagle, J., & Filpula, D. (1986) *J. Bacteriol. 167*, 870–880.

Fahnestock, S. R., Alexander, P., Filpula, D., & Nagle, J. (1990) in *Bacterial Immunoglobulin Binding Proteins* (Boyle, M. P. D., Ed.) Vol. 1, pp 133–148, Academic Press, San Diego.

Filimonov, F. F., Pfeil, W., Tsalkova, T. N., & Privalov, P. (1978) *Biophys. Chem. 8*, 117–122.

Filpula, D., Alexander, P., & Fahnestock, S. R. (1987) *Nucleic Acids Res. 15*, 7210.

Gill, S. J., Nichols, N. F., & Wadsö, I. (1976) *J. Chem. Thermodyn. 8*, 445–452.

Gronenborn, A. M., Filpula, D. R., Essig, N. Z., Achari, A., Whitlow, M., Wingfield, P. T., & Clore, M. G. (1991) *Science 253*, 657–661.

Guss, B., Eliasson, M., Olsson, A., Uhlen, M., Frej, A. K., Jornvall, H., Flock, J. I., & Lindberg, M. (1986) *EMBO J. 5*, 1567–1575.

Kauzmann, W. (1959) *Adv. Protein Chem. 14*, 1–63.

Lian, L. Y., Yang, J. C., Derrick, J. P., Sutcliffe, M. J., Roberts, G. C. K., Murphy, J. P., Goward, C. R., & Atkinson, T. (1991) *Biochemistry 30*, 5335–5340.

Livingstone, J. R., Spolar, R. S., & Record, M. T. (1991) *Biochemistry 30*, 4237–4244.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1962) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Myhre, E. B., & Kronvall, G. (1977) *Infect. Immun. 17*, 475–482.

Myhre, E. B., & Kronvall, G. (1981) *Basic concepts of streptococci and streptococcal diseases* (Holm, E., & Christensen, P., Eds.) pp 209–210, Redbook Ltd., Chertsey Surrey, U.K.

Olsson, A., Eliasson, M., Guss, B., Nilsson, B., Hellman, U., Lindberg, M., & Uhlen, M. (1987) *Eur. J. Biochem. 186*, 319–324.

Orban, J., Alexander, P., & Bryan, P. (1992) *Biochemistry* (following paper in this issue).

Pace, C. N., & Tanford, C. (1968) *Biochemistry 7*, 198–208.

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.

Pajot, P. (1976) *Eur. J. Biochem. 63*, 263–269.

Privalov, P. L. (1979) *Adv. Protein Chem. 33*, 167–241.

Privalov, P. L., & Khechinashvili, N. N. (1974) *J. Mol. Biol. 86*, 665–684.

Privalov, P. L., & Potekhin, S. A. (1986) *Methods Enzymol. 131*, 4–51.

Privalov, P. L., & Gill, S. J. (1988) *Adv. Protein Chem. 39*, 191–234.

Reis, K J., & Boyle, M. D. P. (1990) in *Bacterial Immunoglobulin Binding Proteins* (Boyle, M. P. D., Ed.) Vol. 1, pp 101–112, Academic Press, San Diego.

Reis, K J., Ayoub, E. M., & Boyle, M. D. P. (1984) *J. Immunol. 132*, 3098–3102.

Richards, F. M. (1977) *Annu. Rev. Biophys. Bioeng. 6*, 151–176.

Scandella, D., Arthur, P., Mattingly, M., & Neuhold, L. (1985) *J. Cell Biochem. Suppl. 9, Part B*, 203.

Schwarz, F. P., & Kirchhoff, W. H. (1988) *Thermochim. Acta 128*, 267–295.

Sjöbring, U., Björck, L., & Kastern, W. (1991) *J. Biol. Chem. 266*, 399–405.

Studier, F. W., & Moffatt, B. A. (1986) *J. Mol. Biol. 189*, 113–130.

Sturtevant, J. M. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 2236–2240.

Tanford, C. (1962) *J. Am. Chem. Soc. 84*, 4240–4247.

Tanford, C. (1970) *Adv. Protein Chem. 24*, 1–95.

Vierra, J., & Messing, J. (1987) *Methods Enzymol. 153*, 3–11.