Sequence-Specific ^{1}H NMR Assignments and Secondary Structure of the Streptococcal Protein G B2-Domain

John Orban,* Patrick Alexander, and Philip Bryan

Center for Advanced Research in Biotechnology, University of Maryland, 9600 Gudelsky Drive, Rockville, Maryland 20850

Received October 22, 1991

ABSTRACT: Two-dimensional NMR spectroscopy has been used to obtain sequence-specific ^{1}H NMR assignments for the IgG-binding B2-domain of streptococcal protein G. Secondary structure elements were identified from analysis of characteristic backbone–backbone NOE patterns and amide proton exchange data. The B2-domain contains a four-stranded β-sheet region in which the two inner strands form a parallel β-sheet with each other and antiparallel β-sheets with the outer strands. The outer strands are connected via a 16-residue α-helix and short loops on both ends of the helix. The α-helix and β-sheet structures contain well-defined polar and apolar sides, and numerous long-range NOEs from the apolar helix to apolar sheet regions were used to derive a model for the global fold of the B2-domain. While the overall fold is similar to that obtained for B1-type domains, differences in amide proton exchange rates and hydrophobic packing are observed.

Protein G is a multidomain protein consisting of 2–3 immunoglobulin G (IgG)[1]-binding domains, B1, B2, and B3 (Fahnestock et al., 1986; Fahnestock & Alexander, 1990; Fahnestock et al., 1990). Each domain is 55 amino acids long (55 residues plus N-terminal methionine) with B1 and B2 differing by six amino acid residues in their primary structure. Differences in domain stability and IgG binding have been observed, and a detailed discussion of the comparative thermodynamics of B1 and B2 is presented in the preceding paper in this issue (Alexander et al., 1992). At pH 5.4, B1 and B2 have T_m's of 87.5 °C and 79.4 °C, respectively. Each of these domains binds to the Fc portion of IgG with binding constants of 3×10^7 (B1) and 2.1×10^8 (B2) (T. Lee, unpublished results). Protein A from *Staphylococcus aureus* also binds to IgG but shows no sequence homology to the domains from protein G. The X-ray structure of protein A complexed with the Fc fragment of IgG shows that protein A is α-helical (Deisenhofer, 1981), and solution NMR studies of protein A confirm this although there is an extra helix in the unbound form (Torigoe et al., 1990). In contrast, recent NMR work (Lian et al., 1991) on a B1-like domain of protein G (see Figure 8 for a definition of the B1-like domain) showed that the secondary structure consists of four strands of β-sheet with a 16-residue α-helix joining the two outer strands. However, the orientation of the helix with respect to the sheets was not determined. The structure of the B1-domain in solution, obtained using distance geometry and simulated annealing methods, was recently reported (Gronenborn et al., 1991). In this paper, we report the sequential ^{1}H NMR assignments and secondary structure of the B2-domain and provide a summary of the long-range helix–sheet NOE contacts which determine the tertiary organization of the molecule.

[1] Abbreviations: IgG, immunoglobulin G; NMR, nuclear magnetic resonance; DQF-COSY, two-dimensional double-quantum filtered correlated spectroscopy; RELAY, two-dimensional relayed coherence transfer spectroscopy; NOE, nuclear Overhauser effect; NOESY, two-dimensional NOE spectroscopy; ppm, parts per million; $d_{AB}(i,j)$, the NOE connectivity between protons A and B on residues i and j, respectively. Protons A and B are designated N for amide protons, α for αCH, β for βCH, and γ for γCH.

Materials and Methods

The B2-domain of protein G was prepared and purified using the procedures outlined in the previous paper in this journal (Alexander et al., 1992). N-Terminal sequence analysis showed that the sample also contained a methionine-processed 55-mer which was at too low a level (10%) to interfere in this NMR study. NMR samples were prepared by dissolving lyophilized protein in 0.4 mL of 90% H_2O/10% D_2O containing 100 mM NaOAc-d_3 at pH 5.4, and also in 99.99% D_2O at the same pH. Sample concentrations were 9 mM, and chemical shifts were referenced to external sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS). All NMR spectra were recorded at 29 °C on a Bruker AMX-500 spectrometer and processed on a Silicon Graphics 4D/35 workstation using FTNMR (Hare Research Inc., Woodinville, WA).

DQF-COSY (Piantini et al., 1982), RELAY (Eich et al., 1982; Bax & Drobny, 1985), and NOESY (Jeener et al., 1979; Kumar et al., 1980) experiments were acquired in hypercomplex (States et al., 1982) or time-proportional phase incrementation (Drobny et al., 1979; Bodenhausen et al., 1980) mode with standard phase-cycling schemes. All two-dimensional experiments were recorded with 1024 complex points in t_2 and 400 increments in t_1 with a spectral width of 6024 Hz in both dimensions. For each t_1 value, 32 transients were collected with a relaxation delay of 2.0 s between scans. Solvent suppression was achieved by presaturation during the relaxation delay and during the mixing times of the RELAY and NOESY experiments. No presaturation was employed in D_2O spectra where the residual HOD signal was minimal.

RELAY spectra were acquired with mixing times of 30 ms (D_2O spectra) and 36 ms (H_2O spectra). NOESY spectra were recorded for several mixing times ranging from 50 to 180 ms with random variation of the mixing time by up to 10% (Macura et al., 1981).

COSY and RELAY spectra were apodized with skewed sine-bell functions (phase shift 0°, skew factor 0.7) in both dimensions. NOESY spectra were multiplied in t_2 by a window function with a value of 1.0 for the first 256 points which dropped smoothly to 0 at 1024 points as a sine-bell-squared

0006-2960/92/0431-3604$03.00/0

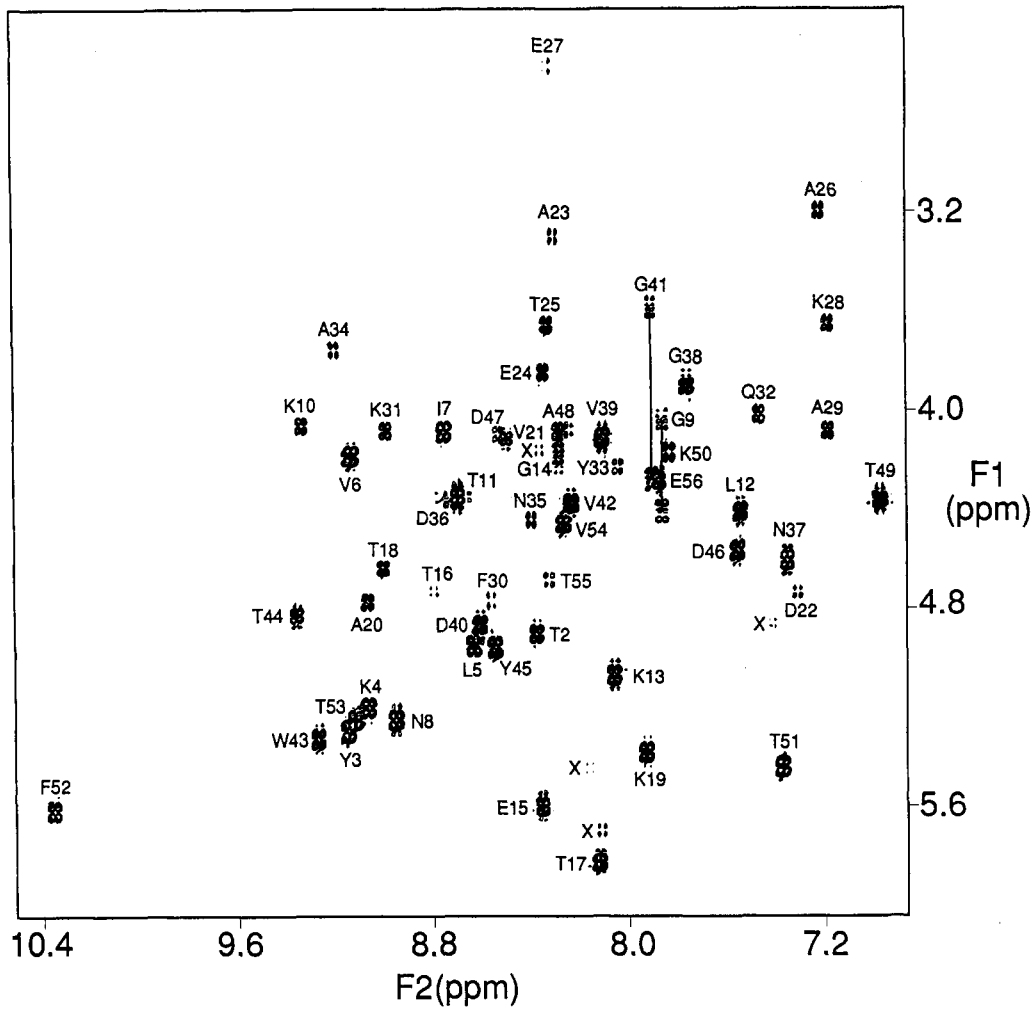

FIGURE 1: DQF-COSY fingerprint region of the protein G B2-domain in H_2O, at 29 °C and pH 5.4. Peaks due to minor contaminants are labeled with an X.

function. In the t_1 dimension, NOESY data were multiplied with a 90°-shifted sine-bell-squared function of 400 points. For all two-dimensional data, the first t_1 experiment was multiplied by 0.5 to suppress t_1 ridge noise (Otting et al., 1985).

Amide proton exchange experiments were carried out by collecting NOESY spectra of the B2 domain freshly dissolved in D_2O, at 29 °C and pH 5.4. NOESY data were recorded in the manner described above except that eight transients were obtained per t_1 experiment. This decreased the data collection time to 4.5 h for a NOESY spectrum while a sufficiently high signal to noise ratio was maintained. Three NOESY spectra were recorded over the 15-h period immediately after the B2-domain was dissolved in D_2O.

Results and Discussion

Sequence-specific assignments for the B2-domain of protein G were made using standard procedures (Wuthrich, 1986). First, amino acid spin systems were identified using DQF-COSY and RELAY spectra, recorded in D_2O, to obtain scalar through-bond connectivities. The αCH and side-chain assignments were then connected to their labile amide protons using H_2O DQF-COSY and RELAY spectra. Sequential assignments were obtained using $d_{\alpha N}(i,i+1)$, $d_{\beta N}(i,i+1)$, and $d_{NN}(i,i+1)$ H_2O NOESY connectivities for unique di- and tripeptides which could then be matched with the known amino acid sequence of the B2-domain.

Amino Acid Spin System Identification. The B2-domain contains the following amino acid composition: 11 Thr, 7 Lys, 6 Ala, 5 Val, 5 Asp, 4 Gly, 4 Glu, 3 Asn, 3 Tyr, 2 Phe, 2 Leu, 1 Trp, 1 Ile, 1 Gln, and 1 Met. The N-terminal Met was not detected in any of the NMR experiments as is often the case in protein studies. The DQF-COSY spectrum in H_2O shows a total of 57 out of the 59 expected NH–αCH cross-peaks in the fingerprint region. Three out of the four glycines were readily identified by the presence of NH–αCH cross-peak pairs (Figure 1) as well as strong αCH–αCH crosspeaks near the diagonal. The remaining Gly residue was characterized during the sequential assignment procedure. All 11 Thr residues were identified by their αCH–βCH and βCH–γCH_3 DQF-COSY connectivities (Figure 2) in conjunction with αCH–βCH–γCH_3 RELAY peaks (D_2O spectra). The five Val residues were identified in a similar manner. D_2O spectra were particularly useful for identification of three of the Thr spin systems which were either very close to or on top of the HOD resonance. The six Ala residues were characterized by their αCH–βCH_3 COSY peaks in D_2O spectra (Figure 2). The two leucines and one isoleucine were assigned by tracing cross-peaks back to the remaining methyl resonances in the aliphatic region using COSY and RELAY spectra.

The αCH–βCH_2 spin system fragments were broadly classified as belonging to aromatic or nonaromatic side chains. Through-space connectivities from aromatic δCH protons to αCH and βCH_2 protons were established from NOESY spectra in D_2O for the six aromatic amino acids. The Phe and Tyr resonances were differentiated by comparison of DQF-COSY and RELAY spectra in D_2O, where the δCH–ϵCH–ζCH relayed coherence path was diagnostic of a Phe aromatic spin system. The Trp43 aromatic protons were readily

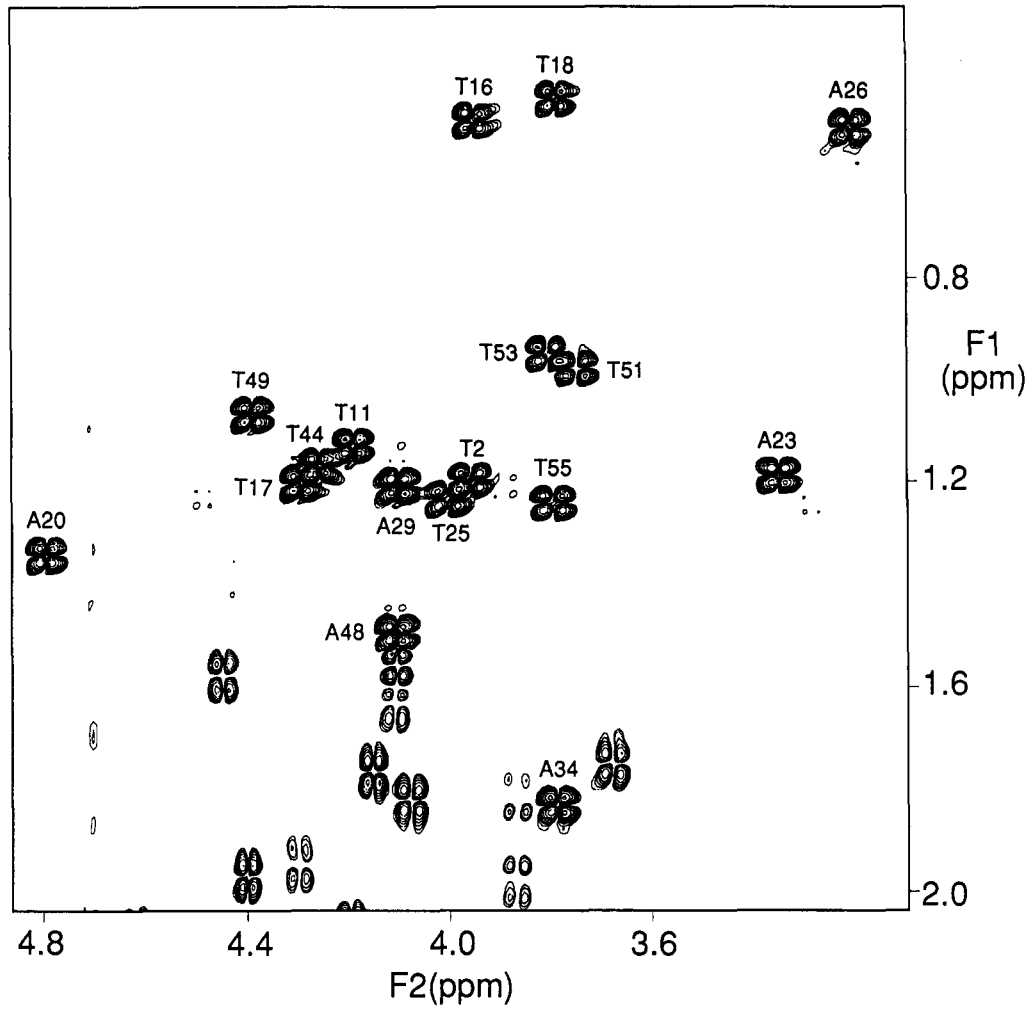

FIGURE 2: Region of the DQF-COSY spectrum in D_2O illustrating the $\beta CH-\gamma CH_3$ connectivities for the 11 threonine residues and the $\alpha CH-\beta CH_3$ cross-peaks for the six alanine residues.

characterized by an extended $\epsilon^3CH-\zeta^3CH-\eta CH-\zeta^2CH$ connectivity path in D_2O COSY and RELAY spectra and an $\epsilon^1NH-\delta^1CH$ COSY peak in the H_2O COSY spectrum. The δ^1CH and ϵ^3CH aromatic Trp protons showed NOEs to their $\alpha CH-\beta CH_2$ spin system.

The remaining 20 spin systems belonged to 3 Asn, 5 Asp, 1 Gln, 4 Glu, and 7 Lys. The Asp/Asn residues were identified by their AMX $\alpha CH-\beta CH_2$ COSY coupling patterns. These residues were distinguished from eath other during the sequential assignment procedure, and further confirmation was obtained by the observation of $\beta CH_2-\delta NH_2$ NOEs for the three Asn residues. The Glu/Gln residues were assigned by the observation of relayed $\alpha CH-\beta CH_2-\gamma CH_2$ cross-peaks which differentiated these spin systems from Asp/Asn. As for the Asp/Asn residues, no distinction was made between Glu and Gln at this stage but this was determined during sequence-specific assignment and confirmed by the observation of $\gamma CH_2-\epsilon NH_2$ NOEs for Gln32. By elimination, the remaining seven spin systems belonged to Lys residues.

For unique αCH chemical shifts, amide protons were connected directly to their spin systems using the H_2O DQF-COSY spectrum. Where ambiguities existed, relayed NH-βCH peaks were used to complete the identity of amino acids.

Sequence-Specific Assignments. Sequential assignments were made from the 180-ms NOESY spectrum in H_2O using the standard $d_{\alpha N}(i,i+1)$, $d_{\beta N}(i,i+1)$, and $d_{NN}(i,i+1)$ connectivities. A logical starting point for sequence-specific assignment was Trp43. $d_{\alpha N}(i,i+1)$ NOESY connectivities were made in both directions to Val42 and Thr44. The other aromatic residues also served as useful starting points. Phe52 was uniquely defined since it gave $d_{\alpha N}(i,i+1)$ connectivities in both directions to two threonines, Thr51 and Thr53. The remaining Phe was therefore Phe30 by elimination. Gly9 gave $d_{\alpha N}(i,i+1)$ connectivities to Lys10 and Asn8. In this manner, oligopeptide segments were built up and matched with the known sequence of the B2-domain. The $d_{\alpha N}(i,i+1)$ connectivities for the peptide fragments Gly38–Asp47 and Thr49–Val54 are shown in Figure 3. In addition to $d_{\alpha N}(i,i+1)$ interresidue connectivities, $d_{NN}(i,i+1)$, $d_{\beta N}(i,i+1)$, and $d_{\gamma N}(i,i+1)$ NOE contacts were also used to obtain sequence-specific assignments. Therefore, ambiguities or interruptions in the NH–αCH connectivity pattern could be bypassed. The additional interresidue connectivities also provided internal cross-checks for the sequential assignment procedure. Figure 4 shows the $d_{NN}(i,i+1)$ NOE contacts for the peptide fragment Ala29–Val39. Medium-range i to $i+2$, $i+3$, and $i+4$ NOE connectivities, which appear in turns and α-helical regions, were also used for sequence-specific assignments. Figure 5 summarizes the sequential and medium-range NOE contacts observed for the B2-domain, and sequence-specific chemical shifts are given in Table I.

Secondary Structure. The B2-domain of protein G contains both antiparallel and parallel β-sheets which are characterized by strong $d_{\alpha N}(i,i+1)$ connectivities (Figure 3) as well as long-range $d_{\alpha\alpha}(i,j)$, $d_{\alpha N}(i,j)$, and $d_{NN}(i,j)$ NOE contacts, summarized in Figure 6. In addition to backbone–backbone NOEs, numerous side chain–backbone and side chain–side chain NOEs (Figure 7) were helpful in defining both the

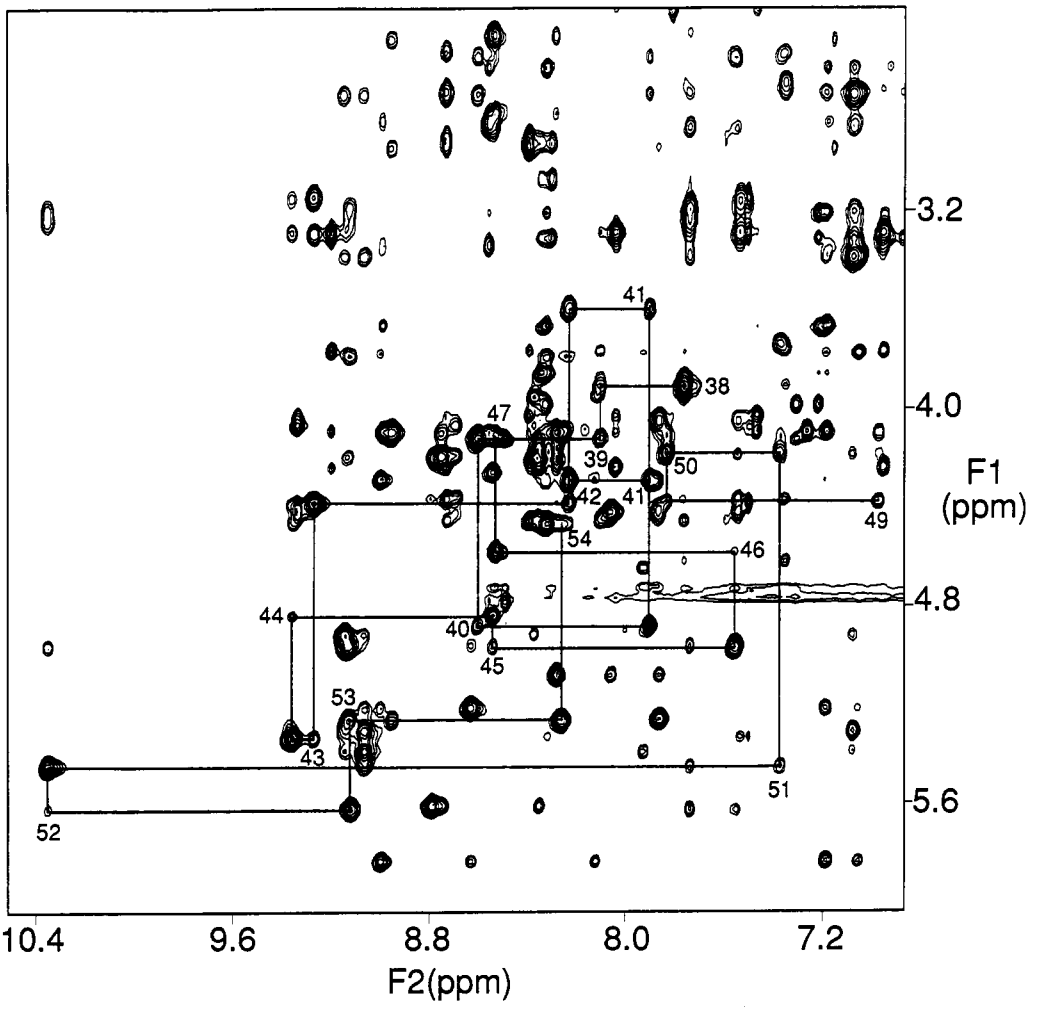

FIGURE 3: NOESY fingerprint region of the protein G B2-domain in H_2O, at 29 °C and pH 5.4. Assignment of residues 38–47 and 49–54 using sequential $d_{\alpha N}(i,i+1)$ connectivities is shown; the intraresidue $d_{\alpha N}$ connectivities are numbered.

parallel and antiparallel β-sheets. The slowly exchanging amide protons observed in D_2O NOESY spectra (Figure 5) were also consistent with the secondary structure proposed in Figure 6. However, it is interesting to note that the amide protons for residues 9, 12, and 14, all of which are involved in an antiparallel β-sheet, were undetectable after only 4.5 h in D_2O. This is probably due to some fraying or distortion of the β-sheet structure near the turn at residues 10 and 11, and it is consistent with the medium interstrand NOE intensity for $d_{\alpha\alpha}(8,13)$ compared with strong interstrand NOE connectivities for $d_{\alpha\alpha}(2,19)$, $d_{\alpha\alpha}(4,17)$, and $d_{\alpha\alpha}(6,15)$ (Figure 6).

Residues 10 and 11 are involved in a type I tight turn which is uniquely defined by strong $d_{NN}(10,11)$ and $d_{NN}(11,12)$ NOE connectivities as well as a medium $d_{\alpha N}(10,11)$ NOE contact (Wuthrich, 1986). Residues 47–50 are involved in a larger turn with strong $d_{NN}(i,i+1)$ NOEs from 47–48, 48–49, 49–50, and 50–51. These observations contrast the type II (residue 9–12) and type I (residue 48–50) turns proposed for a B1-like domain (Lian et al., 1991) but are consistent with the turn structures obtained in solution for the B1-domain using distance geometry calculations (Gronenborn et al., 1991).

The α-helical section of the B2-domain is characterized by strong $d_{NN}(i,i+1)$ sequential NOE connectivities (Figure 4) as well as medium-range i to $i+2$, $i+3$, and $i+4$ NOEs. These diagnostic NOE connectivities are summarized in Figures 5 and 7 and indicate that an α-helical region exists from Ala23 to Gly38. Slow amide proton exchange in this region is also consistent with an α-helix. The Ala23, Glu24, and Gly38 amide protons at each end of the helix exchange rapidly and cannot be detected after 4.5 h in D_2O, while residues 25, 36, and 37 have intermediate amide proton exchange rates. The remaining helical amide proton resonances are still detectable after 15 h in D_2O, with the exception of Gln32 and Asn35 which are detectable after 9 h but are completely exchanged after 15 h. Gln32 and Asn35 are separated by approximately one turn of the helix, and their similar amide proton exchange rates may be due to the comparable degree of solvent exposure (see below and Figure 8A). Short loop regions at Val21–Asp22 and Val39–Asp40 connect the α-helix to the β-sheets.

Global Fold of the B2-Domain of Protein G. Both the α-helical and β-sheet regions of the B2-domain have well-defined polar and apolar sides illustrated in Figure 8A,B. The polar face of the helix would be expected to be solvent-exposed while the apolar side may form a hydrophobic core with the apolar sides of the parallel and antiparallel β-sheets. Indeed, a number of long-range helix–sheet and loop–sheet NOEs were observed which defined the orientation of the helix with respect to the sheets, and these are summarized in Figure 7. In particular, weak NOE connectivities were observed for $d_{\beta N}$-(47,21) and $d_{\alpha N}(22,47)$ and a number of side chain–side chain contacts were detected between Asp22 and Lys50, suggesting a possible stabilizing electrostatic interaction between these two residues. Tyr3 showed numerous backbone and peripheral NOE contacts to Asp22, Ala23, Glu24, and Ala26 as well as a weak NOE between Tyr3β,β′ and Phe30ζ. Likewise, Thr18 showed NOEs to Ala26 and Phe30. It was possible to continue along the first antiparallel β-sheet and observe NOE connectivities from apolar sheet residues to apolar helix residues.

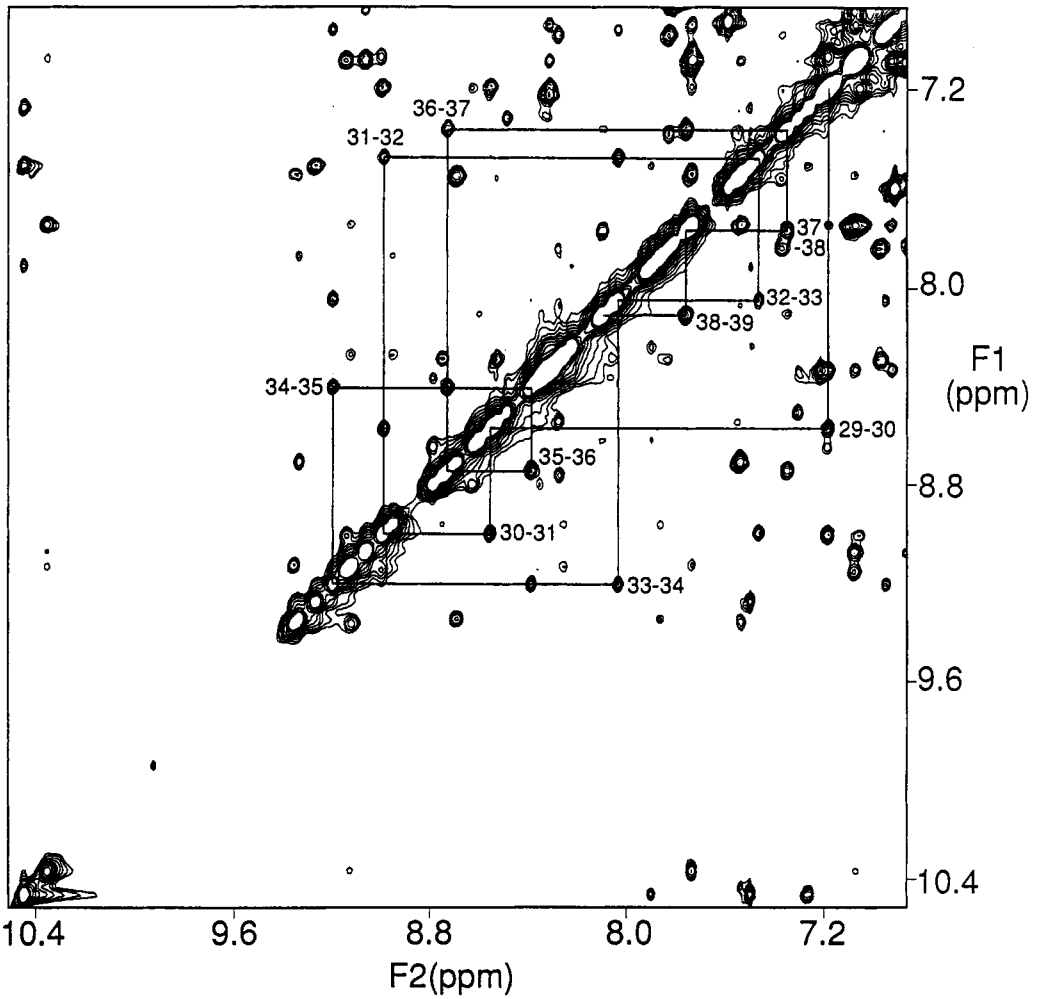

FIGURE 4: NH–NH region of the protein G B2-domain NOESY spectrum in H_2O, at 29 °C and pH 5.4. Assignment of residues 29–39 using sequential $d_{NN}(i,i+1)$ connectivities is illustrated (residue number in F2–residue number in F1).

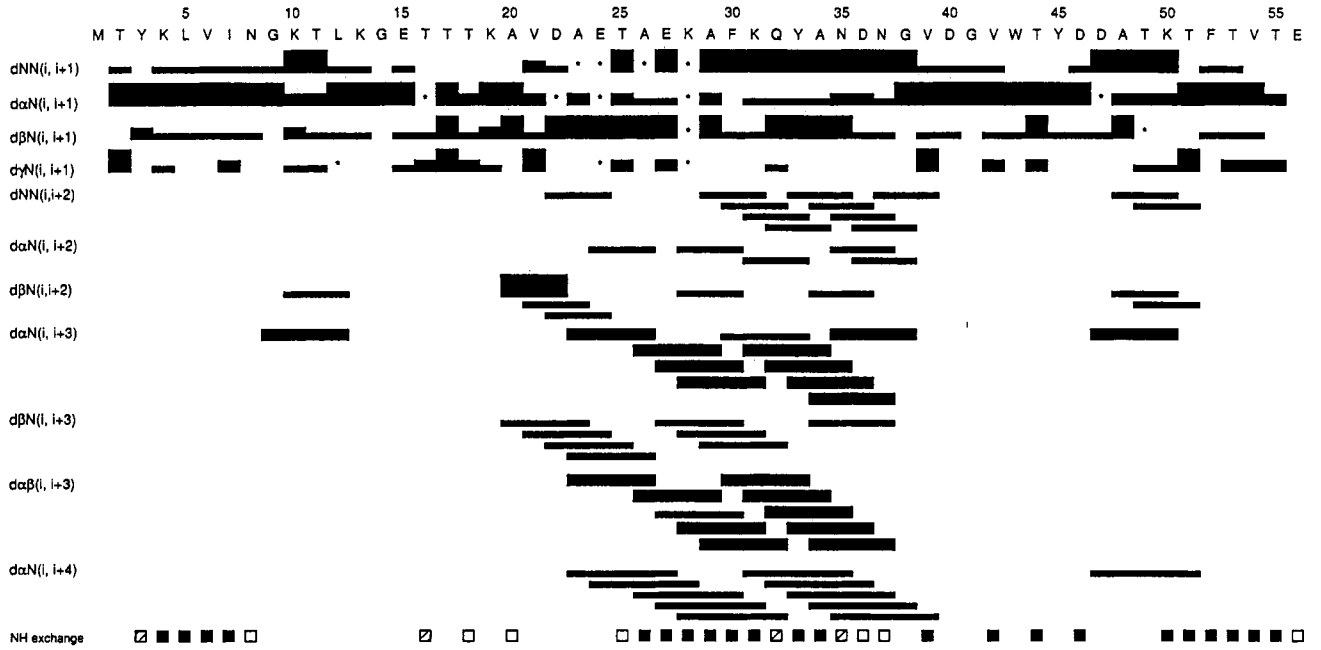

FIGURE 5: Amino acid sequence of the protein G B2-domain with a summary of NOE contacts used in the sequential assignment procedure and amide proton exchange rates. The NOE intensities are proportional to the height of the bars. Asterisks indicate positions where sequential connectivities could not be established due to peak overlap. Medium-range NOE connectivities such as $d_{NN}(i,i+2)$, $d_{\alpha N}(i,i+3)$, and $d_{\alpha N}(i,i+4)$ are represented by lines between the two interacting residues. Amide proton exchange rates are summarized in the following way: open squares indicate that the amide signals were detected in a NOESY spectrum of the B2-domain freshly dissolved in D_2O, recorded over a 4.5-h period at 29 °C and pH 5.4; hatched squares indicate that the amide signals were still observable after 9 h in D_2O at 29 °C and pH 5.4; filled squares identify amide signals which were still observable after 15 h in D_2O using the above conditions.

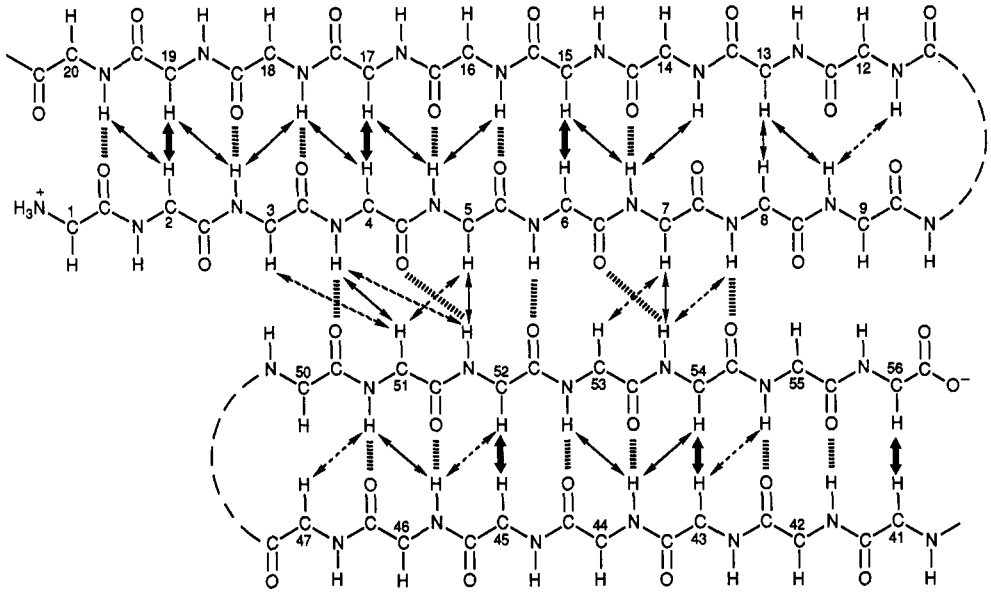

FIGURE 6: Schematic illustration of the four-stranded β-sheet region of the protein G B2-domain. The two central strands form a parallel β-sheet structure while the outer strands are involved in antiparallel β-sheets. The experimental interstrand NOE connectivities between backbone protons are represented by arrows with the NOE intensity being proportional to the thickness of the lines. Hydrogen bonds implied by the observation of slowly exchanging amide protons and at least one interstrand long-range NOE are indicated by hashed lines.

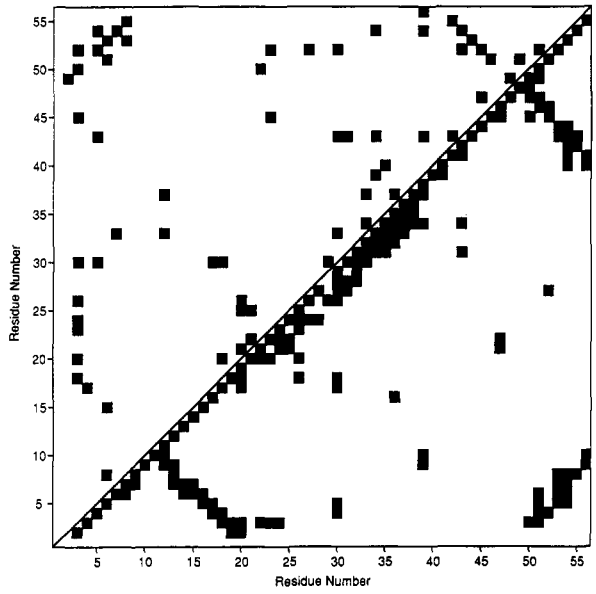

FIGURE 7: Interresidue NOE map for the protein G B2-domain. Filled squares identify residues between which NOE connectivities were observed. Interresidue NOE contacts involving at least one backbone proton are shown below the diagonal; connectivities involving side chain–side chain NOEs are represented above the diagonal.

Thus, the Leu5 δCH_3 protons gave strong NOEs to Phe30 δCH aromatic protons, an Ile7 γCH_2 proton gave a strong NOE contact to Tyr33 δCH and ϵCH aromatic protons, and a Leu12 δCH_3 group gave a medium NOE to the Asn37 βCH_2 protons. Residues 9 and 10 showed weak NOEs to Val39, and Glu56 showed weak sheet–loop connectivities to Val39 and Asp40.

From these long-range NOEs, a model for the global fold of the B2-domain was constructed and is shown in Figure 9. The helix is positioned diagonally across the four strands involved in β-sheets with Asp22 located near Asp47 and Lys50, and Asn37 near Leu12. The abundance of long-range NOEs between apolar residues suggests an efficient packing of the hydrophobic core. This is also evident from the upfield-shifted methylene and methyl resonances of a number of hydrophobic residues in close contact with aromatic amino acids. For example, the high-field chemical shift of the βCH_2 resonances

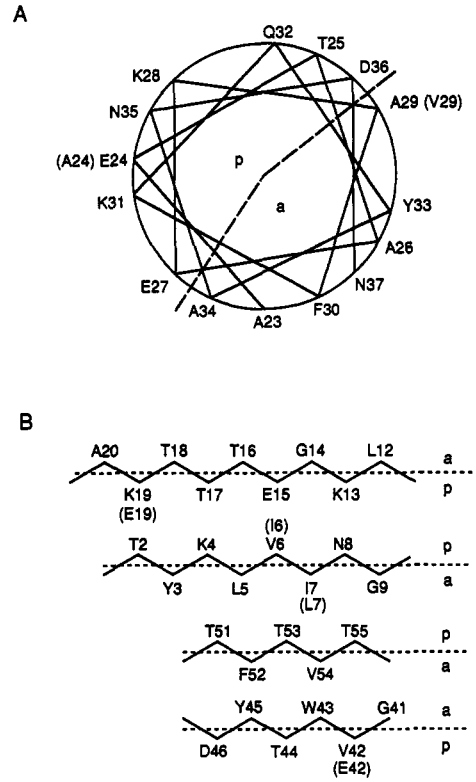

FIGURE 8: Secondary structure elements in the protein G B2-domain illustrating the well-defined polar (p) and apolar (a) regions in the α-helix (A) and the β-sheets (B). The six amino acid substitutions observed for the B1-domain are shown in brackets. In the B1-like domain, V6 and I7 [numbered V11 and I12 in Lian et al. (1991)] are preserved and the other four amino acid substitutions are as for B1.

of Leu5 at −1.16 and 0.78 ppm can be attributed to their proximity to the aromatic rings of Phe30, Phe52, and Trp43. Likewise, the Ile7 γCH_2 resonances occur upfield shifted at 0.58 and 0.75 ppm due to the nearby Tyr33 side chain. Interestingly, the corresponding chemical shifts in a B1-like domain (Lian et al., 1991) are Leu 0.95 ppm and Ile 1.41 ppm, where the data were obtained at 25 °C and pH 4.2. The relatively large chemical shift differences, particularly the

Table I: ^{1}H Chemical Shift Assignments for the B2-Domain of Protein G at pH 5.4 and 29 °C

residue	NH	αCH	βCH	others
Met1				
Thr2	8.39	4.92	3.96	γCH_3 1.19
Tyr3	9.15	5.31	2.74, 3.39	δCH 7.08; εCH 6.88
Lys4	9.07	5.22	1.87, 2.01	γCH_2 1.33, 1.45
Leu5	8.64	4.96	−1.16, 0.78	γCH 0.84; δCH_3 0.46, 0.50
Val6	9.14	4.20	2.06	γCH_3 0.83
Ile7	8.76	4.10	1.55	γCH_2 0.58, 0.75; γCH_3 0.70; δCH_3 1.31
Asn8	8.96	5.26	2.52, 2.95	δNH_2 6.76, 7.16
Gly9	7.87	4.05, 4.42		
Lys10	9.34	4.07	1.81	γCH_2 1.46
Thr11	8.71	4.37	4.19	γCH_3 1.12
Leu12	7.55	4.42	1.37	γCH 1.34; δCH_3 0.77, 0.87
Lys13	8.07	5.08	1.73, 1.87	γCH_2 1.44
Gly14	8.29	4.18, 4.21		
Glu15	8.36	5.61	1.93, 2.03	γCH_2 2.18
Thr16	8.80	4.73	3.95	γCH_3 0.50
Thr17	8.13	5.83	4.29	γCH_3 1.20
Thr18	9.01	4.65	3.78	γCH_3 0.43
Lys19	7.94	5.39	1.70	γCH_2 1.45
Ala20	9.07	4.79	1.34	
Val21	8.50	4.13	2.19	γCH_3 0.98
Asp22	7.32	4.73	2.93, 3.08	
Ala23	8.32	3.32	1.17	
Glu24	8.35	3.86	1.82, 1.98	γCH_2 2.21
Thr25	8.33	3.67	3.98	γCH_3 1.22
Ala26	7.23	3.21	0.52	
Glu27	8.33	2.64	1.84, 1.89	γCH_2 1.59
Lys28	7.19	3.67	1.74	γCH_2 1.53
Ala29	7.19	4.10	1.21	
Phe30	8.57	4.78	2.84, 3.35	δCH 6.57; εCH 7.07; ζCH 7.19
Lys31	9.00	4.10	1.54, 1.62	γCH_2 0.43, 0.77; δCH_2 1.08, 1.61
Gln32	7.48	4.03	2.19	γCH_2 2.40; ϵNH_2 6.84, 7.82
Tyr33	8.05	4.24	3.29	δCH 6.96; εCH 6.69
Ala34	9.20	3.77	1.81	
Asn35	8.40	4.46	2.93	δNH_2 6.92, 7.60
Asp36	8.74	4.37	2.56, 2.73	
Asn37	7.36	4.62	2.07, 2.69	δNH_2 6.34, 6.62
Gly38	7.77	3.91		
Val39	8.11	4.13	1.76	γCH_3 0.65, 0.77
Asp40	8.61	4.88	2.59, 2.74	
Gly41	7.91	3.61, 4.29		
Val42	8.24	4.39	1.96	γCH_3 0.98, 1.03
Trp43	9.27	5.34	3.15, 3.30	δ^1CH 7.51; ϵ^1NH 10.45; ζ^2CH 7.27; ηCH 6.71; ζ^3CH 6.60; ϵ^3CH 7.55
Thr44	9.36	4.84	4.26	γCH_3 1.16
Tyr45	8.55	4.97	2.49, 2.87	δCH 5.91; εCH 6.33
Asp46	7.57	4.58	2.26, 2.59	
Asp47	8.54	4.11	2.51, 2.81	
Ala48	8.29	4.10	1.48	
Thr49	6.98	4.38	4.38	γCH_3 1.06
Lys50	7.84	4.18	2.07, 2.22	γCH_2 1.73
Thr51	7.39	5.45	3.74	γCH_3 0.97
Phe52	10.36	5.63	3.20, 3.25	δCH 7.75; εCH 7.08; ζCH 6.94
Thr53	9.13	5.27	3.79	γCH_3 0.95
Val54	8.27	4.47	−0.29	γCH_3 −0.37, 0.34
Thr55	8.33	4.73	3.80	γCH_3 1.23
Glu56	7.88	4.30	1.95, 2.16	γCH_2 2.31

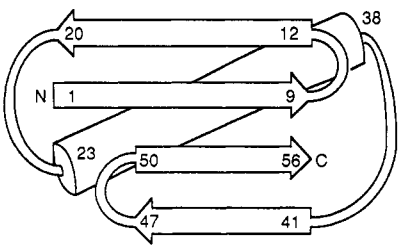

FIGURE 9: Schematic diagram of the global fold for the protein G B2-domain.

greater than 2 ppm difference for one of the Leu5 βCH_2 resonances, are not likely due to solution conditions but rather to differences in packing of the hydrophobic core between the two domains. These differences may contribute to the thermal stability difference between B1 and B2 since the efficient burial of hydrophobic surface has been shown to be the major factor contributing to protein G stability (Alexander et al., 1992).

In general, amide proton exchange rates for the B2-domain and B1-like domain (Lian et al., 1991) are similar, with the notable exception of Val39. In the B1-like domain, the Val39 amide proton [numbered Val44 in Lian et al. (1991)] exchanged with D_2O after 12 h at 25 °C, pH 4.2. By comparison, the Val39 amide proton in the B2-domain was still detectable after 15 h in D_2O at 29 °C, pH 5.4, indicating that structural differences in this loop region exist between the B1- and B2-domains. These structural differences are probably subtle since none of the six variant residues in B2 (Figure 8A,B) show direct NOE contacts to Val39.

While the overall model of the B2-domain global fold derived in the present study is consistent with a recent high-resolution NMR structure obtained for the B1-domain

(Gronenborn et al., 1991), some differences in both hydrophobic packing and amide proton exchange characteristics are apparent. To further understand the differences in thermal stability and IgG binding between B1 and B2, a more detailed structure of the B2-domain using distance geometry methods is required. This work is currently in progress.

REFERENCES

Alexander, P., Fahnestock, S., Lee, T., Orban, J., & Bryan, P. (1992) *Biochemistry* (preceding paper in this issue).

Bax, A., & Drobny, G. P. (1985) *J. Magn. Reson. 61*, 306-320.

Bodenhausen, G., Vold, R. L., & Vold, R. R. (1980) *J. Magn. Reson. 37*, 93-106.

Deisenhofer, J. (1981) *Biochemistry 20*, 2361-2370.

Drobny, G. P., Pines, A., Sinton, S., Weitekamp, D. P., & Wemmer, D. E. (1979) *Symp. Faraday Soc. 23*, 49-55.

Eich, G. Bodenhausen, G., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 3732-3733.

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 Immunoglobin Binding Proteins* (Boyle, M. P. D., Ed.) Vol. 1, pp 133-148, Academic Press, San Diego.

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

Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546-4553.

Kumar, A., Ernst, R. R., & Wuthrich, K. (1980) *Biochem. Biophys. Res. Commun. 95*, 1-6.

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.

Macura, S., Huang, Y., Sater, D., & Ernst, R. R. (1981) *J. Magn. Reson. 43*, 259-281.

Otting, G., Widmer, H., Wagner, G., & Wuthrich, K. (1985) *J. Magn. Reson. 66*, 187-193.

Piantini, U., Sorensen, O. W., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 6800-6801.

States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286-292.

Torigoe, H., Shimada, I., Saito, A., Sato, M., & Arata, Y. (1990) *Biochemistry 29*, 8787-8793.

Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.

Complete Primary Structure of Bovine β_2-Glycoprotein I: Localization of the Disulfide Bridges[†,‡]

Emøke Bendixen, Torben Halkier,* Staffan Magnusson,[§] Lars Sottrup-Jensen, and Torsten Kristensen

Department of Molecular Biology and Plant Physiology, University of Aarhus, DK-8000 Aarhus C, Denmark

Received September 25, 1991; Revised Manuscript Received January 24, 1992

ABSTRACT: The complete primary structure of bovine β_2-glycoprotein I was determined by a combination of cDNA and peptide sequencing. Bovine β_2-glycoprotein I was purified from citrated plasma, and by sequencing selected peptides, the complete disulfide bridge patterns of the 11 disulfide bridges were established as well as the positions of the five asparagine-linked carbohydrate groups. Bovine β_2-glycoprotein I comprises five mutually homologous domains or Short Consensus Repeats, each containing two disulfide bridges, except for the fifth most C-terminal domain which diverges from the Short Consensus Repeat consensus by containing an additional disulfide bridge. In the four N-terminal domains, the first and third and the second and fourth half-cystines are disulfide-linked, while in the fifth domain the first and fourth, the second and fifth, and the third and sixth half-cystines are disulfide-linked.

β_2-Glycoprotein I is a perchloric acid soluble protein in normal human plasma, first purified in 1961 (Schultze et al., 1961). Later, β_2-glycoprotein I has been purified from rat plasma (Polz et al., 1980). The complete amino acid sequence of human β_2-glycoprotein I has been determined by peptide sequencing (Lozier et al., 1984) as well as by cDNA sequencing (Steinkasserer et al., 1991), and it reveals that β_2-glycoprotein I is a single-chain polypeptide, consisting of 326 amino acid residues, with four or five asparagine-linked oligosaccharide groups. The calculated molecular mass of the polypeptide chain is 36 281 Da which is considerably less than the 54 200 Da estimated by SDS-PAGE (Lozier et al., 1984). Sedimentation equilibrium studies of human β_2-glycoprotein I indicate a molecular mass of 40 000-48 000 Da (Heimburger et al., 1964; Finlayson et al., 1967). Glycosylation could account for the discrepancy between the calculated and the estimated molecular mass values. The cDNA derived amino

[†] This work has been supported by The Danish Cancer Society (88-043; T.H.), The NOVO Foundation (1988-12-19 and 1989-12-18; T.H.), The Danish Science Research Council (11-8256; T.K.), The Carlsberg Foundation (88-0274/20; T.K.), The Danish Cancer Society (90-025; L.S.-J.), and The Center for Eucaryotic Gene Regulation, University of Aarhus.

[‡] The nucleotide sequence data reported will appear in the EMBL, GenBank, and DDBJ Nucleotide Sequence Databases under Accession Number X60065.

* Address correspondence to this author at NOVO-Nordisk, NOVO Allé, DK-2880 Bagsvaerd, Denmark.

[§] Deceased.

0006-2960/92/0431-3611$03.00/0
