Three dimensional structure of an immunoglobulin light chain-binding domain of protein L. Comparison with the IgG-binding domains of protein G

ARTICLE *in* BIOCHEMISTRY · DECEMBER 1994

Impact Factor: 3.02 · DOI: 10.1021/bi00251a008 · Source: PubMed

Three-Dimensional Solution Structure of an Immunoglobulin Light Chain-Binding Domain of Protein L. Comparison with the IgG-Binding Domains of Protein G[†,‡]

Mats Wikström,[*,§] Torbjörn Drakenberg,[§] Sture Forsén,[§] Ulf Sjöbring,[∥] and Lars Björck[⊥]

Departments of Physical Chemistry 2, Medical Microbiology, and Medical and Physiological Chemistry, Lund University, Lund, Sweden

Received July 5, 1994; Revised Manuscript Received September 16, 1994[⊗]

ABSTRACT: Protein L is a multidomain protein expressed at the surface of some strains of the anaerobic bacterial species *Peptostreptococcus magnus*. It has affinity for immunoglobulin (Ig) through interaction with framework structures in the variable Ig light chain domain. The Ig-binding activity is located to five homologous repeats called B1−B5 in the N-terminal part of the protein. We have determined the three-dimensional solution structure of the 76 amino acid residue long B1 domain using NMR spectroscopy and distance geometry−restrained simulated annealing. The domain is composed of a 15 amino acid residue long disordered N-terminus followed by a folded portion comprising an α-helix packed against a four-stranded β-sheet. These secondary structural elements are well determined with a backbone atomic root mean square deviation from their mean of 0.54 Å. The B domains of protein L show very limited sequence homology to the domains of streptococcal protein G interacting with the heavy chains of IgG. However, despite this fact, and their different binding properties, the fold of the B1 domain was found to be similar to the fold of the IgG-binding protein G domains [Wikström, M., Sjöbring, U., Kastern, W., Björck, L., Drakenberg, T., & Forsén, S. (1993) *Biochemistry 32*, 3381−3386]. In the present study, the solution structure of the B1 domain enabled a more detailed comparison which can explain the different Ig-binding specificities of these two bacterial surface proteins. Among the differences observed, the α-helix orientation is the most striking. Thus, in the B1 domain of protein L the helix is almost parallel to the β-sheet, whereas in the protein G domains the helix runs diagonally across the sheet.

Over the last decade a number of immunoglobulin- (Ig-)[1] binding surface proteins have been identified and characterized in different, mostly Gram-positive, bacterial species [for references see Boyle (1990) and Kehoe (1994)]. With one exception, all of these proteins interact with heavy chains of Ig. The exception, protein L, was identified in some strains of the anaerobic bacterial species *Peptostreptococcus magnus* (Myhre & Erntell, 1985). The molecule was isolated and found to bind Ig light chains, predominantly κ light chains (Björck, 1988). Protein L is an elongated protein which, despite its high-affinity interaction with the variable domain of Ig light chains, does not interfere with the antigen-binding activity of the antibody (Åkerström Björck, 1989; Nilson et al., 1992). This is explained by the fact that protein L interacts with the framework region of the variable domain (Nilson et al., 1992). The protein L originally isolated and characterized contains five highly homologous Ig light chain-binding domains (B1−B5) of 72−76 amino acid residues each (Kastern et al., 1992) and has no albumin-binding activity (de Chateau et al., 1993). However, a recent report describes a protein L variant which apart from four B-type repeats also contains albumin-binding repeats (Murphy et al., 1994) that are closely related to the albumin-binding domains first identified in protein G (Åkerström et al., 1987), a surface protein of group C and G streptococci with affinity also for Fc and Fab fragments of IgG (Reis et al., 1984; Björck & Kronvall, 1984).

In this paper we report the three-dimensional structure of the B1 domain of protein L determined using NMR methods in combination with distance geometry−restrained simulated annealing. A comparison with the structure of the IgG-binding domains of protein G (Gronenborn et al., 1991; Lian et al., 1992; Achari et al., 1992), reveals similarities and differences with implications for the evolution and biology of Ig-binding bacterial proteins.

MATERIALS AND METHODS

Protein Purification and Sample Preparation. The expression of the protein L B1 domain and its purification have been described previously (Wikström et al., 1993). The peptide, produced in *Escherichia coli*, includes the 76 amino acids of the B1 domain (K80−G155), preceded by the two C-terminal residues (E78−N79) from the N-terminal domain of protein L.

[†] This work was supported by grants from the Swedish Medical Research Council (Projects 7480, 9926, and 10434), the Medical Faculty, Lund University, and the Swedish Research Council for Engineering Sciences (Project 123). The 500-MHz NMR spectrometer was purchased with grants from the Knut and Alice Wallenberg Foundation and the Swedish Council for Planning and Coordination of Research.

[‡] Atomic coordinates for the protein L B1 domain have been deposited in the Brookhaven Protein Data Bank with accession code 2PTL.

* To whom correspondence should be addressed.

[§] Physical Chemistry 2.

[∥] Medical Microbiology.

[⊥] Medical and Physiological Chemistry.

[⊗] Abstract published in *Advance ACS Abstracts,* November 1, 1994.

[1] Abbreviations: Ig, immunoglobulin; Fab, antigen binding fragment of Ig; Fc, constant region of the heavy Ig chain; NMR, nuclear magnetic resonance; 2D, two dimensional; NOE, nuclear Overhauser enhancement; NOESY, nuclear Overhauser enhancement spectroscopy; PE-COSY, primitive exclusive correlated spectroscopy; SA, simulated annealing; rms, root mean square.

0006-2960/94/0433-14011$04.50/0

Derivation of Structural Constraints. NMR spectra were recorded on a GE Ω 500 spectrometer operating at 500.13 MHz for protons (Wikström et al., 1993). NOE distance constraints were derived from NOESY spectra (Macura & Ernst, 1980) that have been processed on SUN Sparc Workstations using the Felix program (BIOSYM Technologies, San Diego). NOESY spectra with five different mixing times (40, 60, 80, 120, and 200 ms) were obtained in both H_2O and D_2O at pH 6.0 and 300 K. Template cross-peak files were generated from the 200-ms NOESY experiments, and the intensities were measured by integrating all cross peaks at the different mixing times using the Felix program. Intensities corresponding to equivalent and degenerate proton resonances were scaled according to Yip (1990). The NOE intensity as a function of mixing time was fit to a second degree polynomial for each cross peak to obtain the initial slope. The initial slope is proportional to the distance between the protons giving rise to the specific cross peak. The distances were thereafter calibrated using the $H\delta-H\epsilon$ distance of 2.48 Å in the tyrosine side chain or using standard distances from the earlier identified secondary structural elements (Wüthrich, 1986). The upper distance limits were set to 10% over the calculated distance. The lower distance limits were conservatively set to 1.8 Å throughout.

Backbone ϕ dihedral angle constraints were obtained from $^3J(H^\alpha,H^N)$ coupling constants using a curve-fitting procedure as described previously (Wikström et al., 1993). A $^3J(H^\alpha,H^N)$ coupling constant less than 5 Hz indicates an α-helical conformation, and the dihedral angle ϕ is constrained as $-60° \pm 20°$. On the other hand, a $^3J(H^\alpha,H^N)$ coupling constant larger than 8 Hz is indicative of an extended β-conformation, and the dihedral constraint is then $-120° \pm 40°$. χ^1 dihedral angles and stereospecific assignment of eight β-methylenes were obtained using the $^3J(H^\alpha,H^\beta)$ coupling constants from a PE-COSY (Müller, 1987) together with related distances determined using the initial buildup approach (Wagner et al., 1987). The χ^1 dihedral angles were set to one of the three staggered rotamer conformations $(-60°, 180°, +60°) \pm 40°$.

No hydrogen bond constraints were used in any stage of the structure calculations. The evaluation of hydrogen bonds was performed with the program MUMLOOK (M. Ullner and O. Teleman). The acceptance criteria for a hydrogen bond were that the distance between the acceptor oxygen and the donor hydrogen was ≤2.5 Å and the angle between the acceptor oxygen, the donor hydrogen, and the donor nitrogen was ≥120°.

Structure Calculations. Structures were generated using the program X-PLOR (Brünger, 1992). Initial structures were obtained by distance geometry embedding using all atoms without metrization. These initial structures were subjected to restrained simulated annealing followed by a SA refinement. In the last two steps the dynamical annealing starts at 2000 K, is cooled down to 100 K, and ends with a Powell minimization at 0 K. Instead of using pseudoatoms for equivalent and nonstereospecifically assigned protons, we incorporated these distance constraints as $(\Sigma r^{-6})^{-1/6}$ effective distances (Constantine et al., 1992). This approach gives tighter and physically more realistic constraints. An additional correction of 0.3 Å was added to the upper distance limits for methyl groups to take into account the fast reorientation of these groups. None of the final 21 structures had a distance constraint violation >0.3 Å and no dihedral angle constraint violation >4°.

The structures were visually inspected and molecular graphics images produced using the MidasPlus software system (Ferrin et al., 1988) from the Computer Graphics Laboratory, University of California, San Francisco, running on Silicon Graphic workstations.

The coordinates for the final set of structures of the protein L B1 domain have been deposited with the Protein Data Bank, Chemistry Department, Brookhaven National Laboratory, Upton, NY 11973.

RESULTS AND DISCUSSION

We have previously reported on the ^1H resonance assignment and the secondary structure of the B1 domain of protein L (Wikström et al., 1993). The presently calculated structure supports our earlier finding that the 17 amino acid residues in the N-terminus are highly flexible and without a regular secondary structure (15 amino acids from the B1 domain and the two N-terminal amino acids E78 and N79). This part of the peptide chain is highly charged, containing ten charged residues (one lysine, one aspartate, and eight glutamates). Similar charged regions are present also in the other four B domains and have been proposed to function as linking segments between the Ig-binding domains (Wikström et al., 1993). The high flexibility of this part of the B1 domain has also been confirmed by ^{15}N relaxation NMR experiments (Wikström et al., unpublished experiments). Our discussion on the structure of the protein L B1 domain is therefore focused on the folded portion of the peptide, i.e., residues Val95−Gly155.

Structure of the B1 Domain of Protein L. The structure was calculated with a total of 925 experimental constraints. Of these, 881 were NOE constraints comprising 312 intraresidue, 187 sequential, 95 medium-range, and 287 long-range constraints. In addition, 44 dihedral angle constraints for 36 ϕ and 8 χ^1 angles were used. No hydrogen bond constraints were used in any stage of the structure calculations. In the final round, 100 structures were calculated, and of these, 59 had no NOE distance constraints violation >0.5 Å and no dihedral constraints violation >5°. A subset of 21 structures with no distance constraint violation >0.3 Å and no dihedral constraint violation >4° was chosen as the representative set of structures for the protein L B1 domain. A Ramachandran plot (Ramachandran et al., 1963) of residues Val95−Ala154 shows that the majority of non-glycyl residues lie within energetically allowed ϕ,ψ regions (Figure 1). The structures have very small deviations from ideal geometry as can be seen in Table 1. A mean structure representing the geometric average of the 21 refined structures was calculated by superimposing the backbone (N, C^α, C′) atoms of the secondary structural elements using the X-PLOR program (Brünger, 1992).

Description of the Structure. Figure 2 depicts the structure of the B1 domain. The folded portion contains a four-stranded β-sheet with a central α-helix on top. These parts of the protein molecule are well-defined with a backbone root mean square deviation from their mean structure of 0.54 Å. The α-helix involves residues from Phe117 to Leu131, as previously stated (Wikström et al., 1993). The β-sheet is of an unusual mixed type with two outer antiparallel pairs of strands in combination with a central pair of parallel

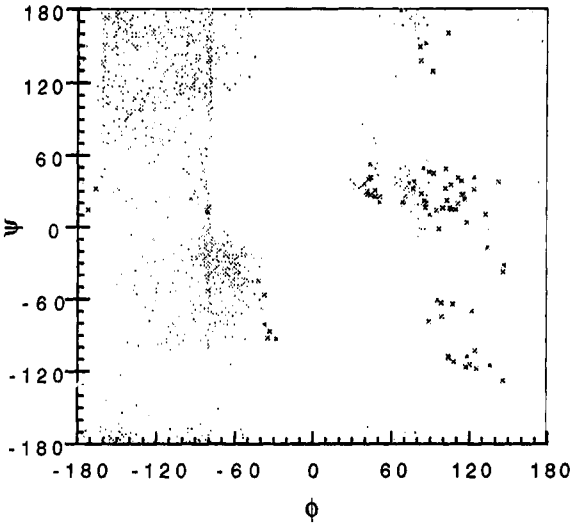

FIGURE 1: Ramachandran plot of residues Val95−Ala154 for the 21 refined structures. Glycine residues are represented by (×) and the other residues by (●).

strands. The β-sheet shows the characteristic right-handed twist. A five-residue turn from Phe103 to Ser107 connects the first two β-strands, β1 (Val95−Phe103) and β2 (Ser107−Gly115). In the turn connecting strands β3 (Glu137−Asp144) and β4 (Tyr147−Ala154), both Lys145 and Gly146 have φ,ψ angles (in the positive φ region of Figure 1) and a NOE pattern suggesting a type I′ β-turn. An evaluation of the calculated structures shows, however, that the expected hydrogen bond between Asp144(CO) and Tyr147(NH) has a rather low occurrence (present in 4 of the 21 refined structures). This agrees with the observation that the amide proton of Tyr147 has a relatively fast exchange rate (Wikström et al., 1993). We have therefore assigned this turn as a distorted type I′ β-turn. The angle between the helix and the β-sheet is $\Omega \sim -10°$ in the notation of Janin and Chothia (1980). Both the α-helix and the β-sheet show a pronounced amphipathic character, combining their hydrophobic surfaces into an inner hydrophobic core. The segment following the helix, starting at Lys132, is a rather compact loop connecting the α-helix to the third β-strand at Glu137. Figure 3 shows the root mean square deviation from the mean structure for the backbone and for all heavy atoms and the number of NOEs, all plotted against residue number. The most well-defined regions of the structure are not surprisingly the four-stranded sheet and the helix. On the other hand, the first turn and the loop between the helix and the third strand and the C-terminus are the most ill-defined regions. Except for the previously discussed distorted type I′ turn, the main, non-glycine contributors to the positive φ region in Figure 1 are also found in these regions. The bad precision of the turn, Phe103 to Ser107, and the C-terminus can be explained by the low NOE density (with no long-range contacts), whereas the loop has relatively many both medium- and long-range NOEs. We are currently investigating if this loop could be more flexible than other parts of the protein by the application of ^{15}N relaxation NMR experiments on a ^{15}N-labeled B1 domain.

As previously mentioned, the calculation of the structure involves no hydrogen bond constraints. An evaluation of the calculated structures allowed the identification of a number of hydrogen bonds present in at least 10 of the 21 refined structures using the criteria outlined in the Materials and Methods section. The identified hydrogen bonds are summarized in Table 2. The first two helical residues, Phe117 and Glu118, show only CO(i)−NH(i+3) hydrogen bonds, which agrees with the slow amide proton exchange observed for both Ala120 and Thr121. The C-terminal end of the helix has a more irregular hydrogen-bonding pattern with Ala126 CO showing hydrogen bonds to both Asp129 NH (11 of 21 structures) and Thr130 NH (18 of 21 structures), whereas the last helical residue, L131 NH, has only weakly populated hydrogen bonds to A128 CO and Y127 CO (5 of 21 structures). Two inner residues in the α-helix, Ser122 and Asp129, were reported to have relatively fast exchange with solvent (Wikström et al., 1993). It is interesting to note that both these residues have more weakly populated CO(i)−NH(i+4) hydrogen bonds (Ser122, 9 of 21 structures; Asp129, 7 of 21 structures).

In the β-sheet the residues identified to be involved in hydrogen bonds in most cases also show slow NH exchange with solvent. Val95 and Ala154 are, however, exceptions showing fast NH exchange but still a high occurrence in the respective interstrand hydrogen bond (Table 2). These two residues appear at the N- and C-termini of strands β1 and β4, respectively. The faster NH exchange of Val95 and

Table 1

(A) Summary of Experimental Constraints Used in the Structure Calculation and Structural Statistics of the Calculated Structures of the B1 Domain of Protein L

interproton NOE distances[a] (Å)	
total	**881**
intraresidue	312
sequential ($i-j=0$)	187
medium range ($2 \le \lvert i-j \rvert \le 4$)	95
long range ($\lvert i-j \rvert \ge 5$)	287
dihedral angle (deg)	
total	**44**
ϕ	36
χ^1	8
total energy (kcal mol^{-1})	194 ± 19
bonds	13 ± 1
angles	84 ± 8
impropers	14 ± 3
NOE restraints[b]	62 ± 7
dihedral restraints[b]	1.9 ± 0.5
rms deviation from experimental restraints	
NOE distance restraints (Å)	0.036 ± 0.002
dihedral restraints (deg)	0.83 ± 0.12
rms deviation from ideal covalent geometry	
bonds (Å)	0.0033 ± 0.0001
angles (deg)	0.51 ± 0.02
impropers (deg)	0.39 ± 0.04

(B) Atomic rms Deviations from the Respective Mean[c] Structure for the 21 Refined Structures

	backbone (N, Cα, C′)	all heavy atoms
α-helix and β-sheet	0.54 ± 0.11	0.98 ± 0.09
α-helix	0.34 ± 0.08	0.93 ± 0.09
β-sheet	0.54 ± 0.12	1.09 ± 0.14
Val95−Gly155	0.75 ± 0.15	1.16 ± 0.12

[a] The number of constraints are listed only for the folded portion of the molecule, Val95−Gly155. [b] The values of the square-well NOE and dihedral angle potentials are calculated with force constants of 50 kcal mol^{-1} Å^{-2} and 200 kcal mol^{-1} rad^{-2}, respectively. [c] The mean structures is the coordinates of the geometric average for the respective regions (α-helix, β-sheet, etc.) calculated from the 21 refined structures. The mathematical average is obtained by superimposing only N, Cα, and C′ atoms using the program X-PLOR (Brünger, 1992). The α-helix involves residues Phe117−Leu131, while the β-sheet is composed of residues Val95−Phe103, Ser107−Gly115, Glu137−Asp144, and Tyr147−Ala154.

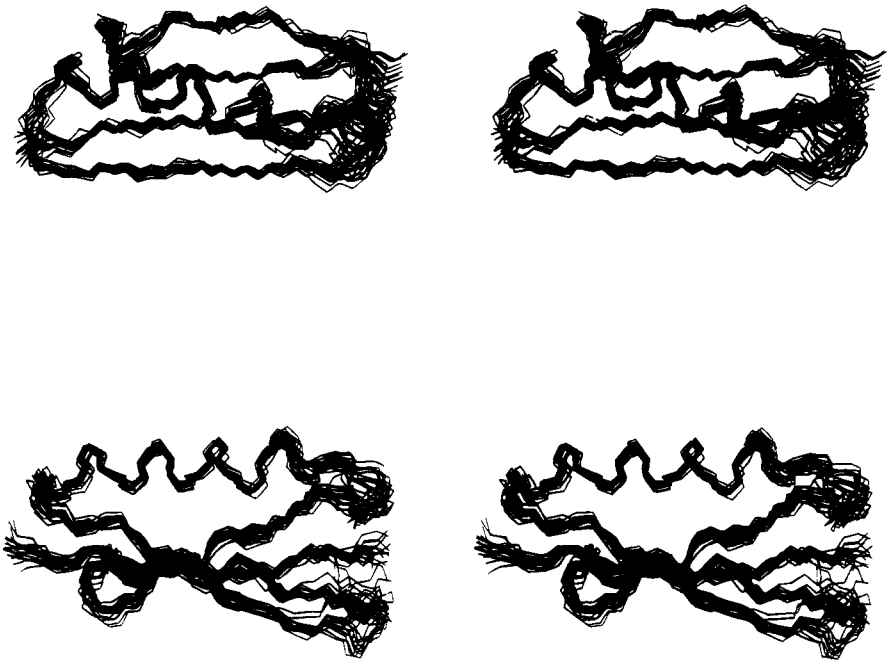

FIGURE 2: Structure of the protein L B1 domain. Stereo diagrams of the selected 21 refined structures after they have been superimposed on their mean structure using the backbone (N, C^{α}, C′) atoms of the secondary structural elements.

Ala154 is therefore probably related to the solvent exposure of these two residues.

Due to the high sequence homology within the Ig-binding domains of protein L (89% sequence identity B1:B2 down to 61% sequence identity B1:B5), we expect that all five domains will have a similar three-dimensional structure. The five homologous domains are aligned in Figure 4. Strands β1, β2, and β4 show a high degree of conservation. β-Strands β1 and β4 form the inner parallel pair of strands in the B1 domain, and the residues in these two strands are therefore involved in a number of interactions with amino acid residues in both the other strands as well with residues within the α-helix. The conservation of these regions points to the importance of keeping the residues in the interior of the protein invariant to allow the formation of an inner hydrophobic protein core. Four of the residues in the helix are different in domains B2−B5 as compared to the B1 sequence. Analysis of the structure of the B1 domain reveals that these four residues are exposed on the surface of the α-helix and that the amino acid residues making contacts with the β-sheet are all conserved. Again, this is probably due to interactions in the interior of the peptide.

There are five conserved glycine residues in the protein L B1 domain. They are located in the two turns connecting the β-strands, at the end of the second β-strand, in the loop connecting the helix with the third strand, and as the C-terminal residue. Thus it seems that glycine with its conformational flexibility is required in these positions. This notion is supported by the fact that the glycine residues (Gly106, Gly115, Gly136, and Gly146) occupy the positive ϕ region in a large majority of the 21 refined structures (Figure 1). The contribution to the lower, right quadrant of Figure 1 ($\phi \sim +120°$, $\psi \sim -90°$) is solely from the last residue in strand β2, Gly115, indicating the importance of a glycine residue at this position.

Comparison of the Ig-Binding Domains of Protein L and Protein G. As mentioned above, streptococcal protein G contains separate binding domains for albumin and IgG. The IgG-binding domains have affinity for two regions in the heavy (γ) chain: the Cγ2−Cγ3 interface region (Stone et al., 1989) and the Cγ1 domain (Derrick & Wigley, 1992). The structure of single IgG-binding domains from protein G has been solved both by NMR methods (Gronenborn et al., 1991; Lian et al., 1992) and by X-ray crystallography (Achari et al., 1992). The sequences for these domains are highly homologous (only 2−6 of 55 amino acids differ), and the three-dimensional structures are very similar. The determination of the secondary structure of the B1 domain of protein L (Wikström et al., 1993) revealed that the fold is similar to the one described for the protein G domains. Both have a central α-helix on top of a mixed four-stranded β-sheet in a $(-1, +3x, -1)$ topology. The amino acid sequences for protein L and G are aligned in Figure 5, where the alignment is adopted to secondary structural elements showing an overall low homology (16% sequence identity). A schematic comparison of the two structures (mean structures) is shown in Figure 6. In Figure 7, the aligned β-strands (from Figure 5) of the two domains are superimposed, using backbone (N, C^{α}, C′) atoms, giving an rms difference of 2.42 Å. In spite of the overall resemblance of the Ig-binding domains of protein L and protein G, several differences are found when the three-dimensional structures of the two proteins are compared. The most striking difference is the orientation of the α-helix with respect to the β-sheet. The helix in protein G runs diagonally across the β-sheet with an angle of $\Omega \sim -40°$, whereas in protein L the helix is more parallel to the sheet ($\Omega \sim -10°$). A difference related to this is the conformation of the loop connecting the helix to the third β-strand. In protein G, the C-terminus of the α-helix is more distant from the third β-strand. The connecting loop between the helix and the third strand is consequently extended. In protein L, on the other hand, the C-terminus of the helix is closer to the third strand, which makes the loop more compact.

Sequence comparisons demonstrate that the Ig-binding B1 domains of proteins L and G have the sequence DNG in

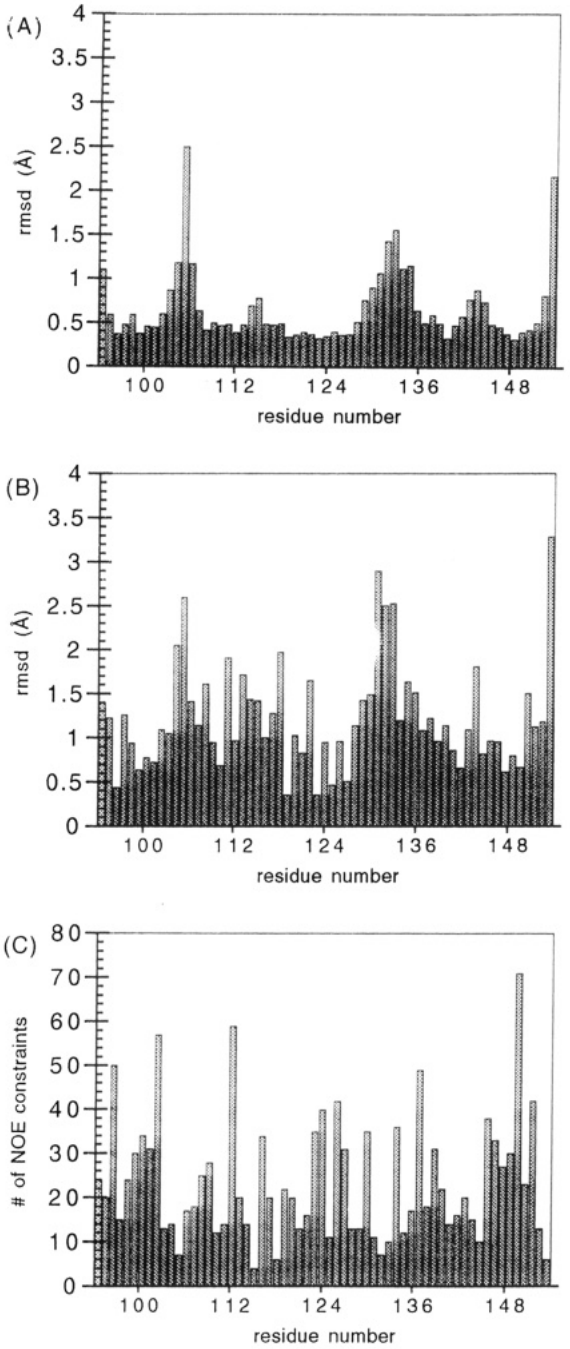

FIGURE 3: Graphs of the deviation from the mean structure and the NOE density. Backbone (A) and all heavy atom (B) root mean square deviation from their mean is shown as a function of residue number. As a comparison, the number of experimental NOE constraints as a function of residue number is depicted in (C).

structurally related parts of the molecule. In protein G, DNG is located at the end of the α-helix, whereas in protein L, the sequence is found in the loop between the α-helix and the third strand. In protein G, the helix is terminated by the glycine residue in the DNG sequence. Aurora et al. (1994) have shown that one-third of all α-helices in proteins, with high-resolution structures deposited in the Brookhaven Protein Data Bank, terminate with a glycine residue, and they have classified these helices into two primary motifs on the basis of their hydrogen-bonding pattern. Inspection of the protein G structure (Gronenborn et al., 1991; Achari et al., 1992) shows that it belongs to the larger class, called the Schellman motif, due to hydrogen bonds between G38(NH) and N35(CO) and between V39(NH) and A34(CO), respectively. In the B1 domain of protein L, the helix, instead of being terminated by a glycine, ends with two lysine residues, and evaluation of the protein L structures shows that the hydrogen-bonding pattern observed in the protein G domain indicative of a Schellman motif is absent.

Table 2: Hydrogen Bonds Identified in at Least 10 of the 21 Refined Structures of the B1 Domain of Protein L[a]

donor[b]	acceptor	occurrence
G115 NH	V96 O	14
F113 NH	I97 O	11
A111 NH	A99 O	12
Q109 NH	L101O	18
F153 NH	I102 O	20
L101 NH	Q109 O	17
A99 NH	A111 O	10
A120 NH	F117 O	11
T121 NH	E118 O	12
E123 NH	K119 O	10
A124 NH	A120 O	19
Y125 NH	T121 O	21
A126 NH	S122 O	21
Y127 NH	E123 O	17
A128 NH	A124 O	21
D129 NH	A126 O	11
T130 NH	A126 O	18
A154 NH	E137 O	19
Y147 NH	A143 O	11
T148 NH	A143 O	15
A143 NH	T148 O	11
N100 NH	L149 O	11
D141 NH	N150 O	16
I102 NH	I151 O	17
T139 NH	K152 O	13

[a] Hydrogen bonds are listed if they are observed in at least 10 of the 21 refined structures using the criteria described in Materials and Methods. The listed hydrogen bonds involve in all cases atoms in the peptide backbone. [b] Residues showing slow amide proton exchange (Wikström et al., 1993) are underlined.

```
B1: VTIKANLIFANGSTQTAEFKGTFEKATSEAYAYADTLKKDNGEYTVDVADKGYTLNIKFAG
B2: --------V-D-K----------E--A---R---A-----------------------
B3: --------V-D-K----------E--A---R---L-A-E--K------------------
B4: --------V-D-K----------AE--A---R---L-A-E--K--A-LE-G---I--R---
B5: ----E-IYYED-TV---T-----AE--A---R---L-S-EH-K--A-LE-G---I--R---
```

FIGURE 4: Alignment of the five homologous B domains in protein L. B2−B5 are compared to the B1 domain, and identical residues are indicated (−).

In addition to the characteristic CO(i)−NH(i+4) hydrogen bonds the helices in proteins L and G also show several CO(i)−NH(i+3) hydrogen bonds. The first two residues in protein L show only CO(i)−NH(i+3) hydrogen bonds. This is in contrast to the helix of protein G where the first residues have either both CO(i)−NH(i+3) and CO(i)−NH(i+4) or only the characteristic latter one. In the protein G structure by Gronenborn et al. (1991) the last five residues in the helix are tightened into a 3_{10} helix showing only CO(i)−NH(i+3) hydrogen bonds, whereas in the structures by Lian et al. (1992) the helix seems to be α-helical over its entire length. In the B1 domain of protein L, there is no indication of a tightening toward the C-terminus. On the contrary, analysis of the hydrogen-bonding pattern in protein L suggests some fraying at the end of the α-helix.

The first two β-strands are of the same length in proteins L and G, both containing nine amino acid residues. The turns connecting these two pairs of β-strands are, however, different. In protein G, strands β1 and β2 are connected by a type I β-turn while in protein L the linking segment is a five-residue turn. In protein L the length of the other two

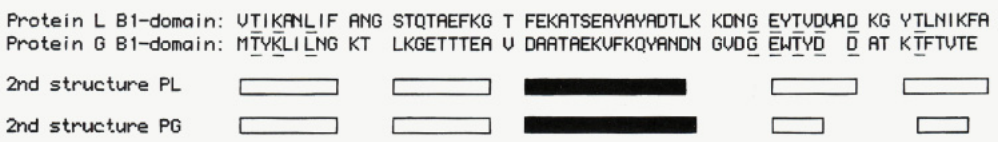

FIGURE 5: Alignment of the B1 domains of protein L and protein G (Fahnestock et al., 1986). The alignment is based on secondary structural elements. Identical residues are underlined. The secondary structural elements for protein L (PL) (Wikström et al., 1993) and protein G (PG) (Gronenborn et al., 1991) are shown below the sequences. β-strands are represented by open boxes and the two α-helices by filled boxes.

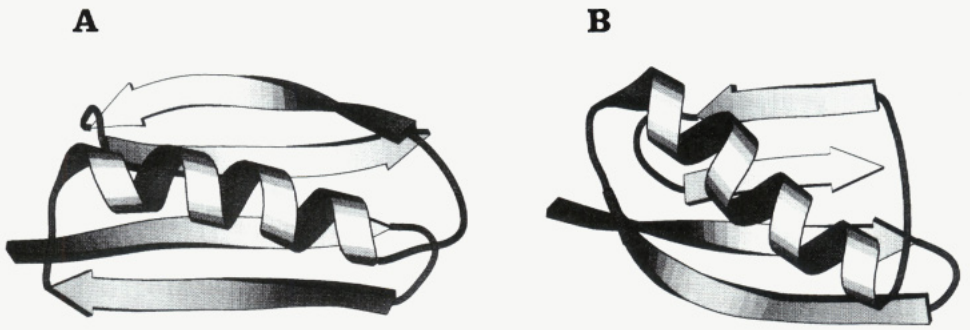

FIGURE 6: Schematic ribbon representation of the B1 domains (mean structures) of protein L (A) and protein G (B) (Gronenborn et al., 1991) using the program MOLSCRIPT (Kraulis, 1991).

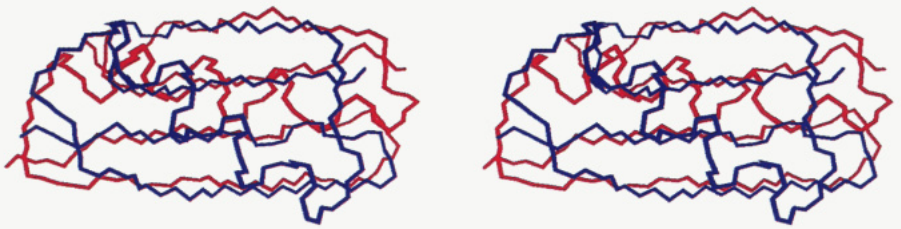

FIGURE 7: Superposition of the aligned β-strands in the two B1 domains (from Figure 5) of protein L (red) and protein G (Gronenborn et al., 1991) (blue). Only backbone (N, C^{α}, C′) atoms are superimposed and shown (rms difference: 2.42 Å).

strands, β3 and β4, is eight amino acid residues whereas in protein G the corresponding strands are only five residues long. In protein G the loop connecting these strands has been assigned a six-residue turn (Gronenborn et al., 1991), which also has the effect of making the strands one amino acid residue shorter. In protein L this connecting loop has been assigned a distorted type I′ β-turn.

To summarize, the conformations of the Ig-binding domains of proteins L and G are quite different. These differences pertain not only to the different helix orientations but also to differences in the turns between the β-strands, as well as the lengths of strands β3 and β4.

Immunoglobulin-Binding Surfaces. The IgG-binding domains of protein G seem to have physically separated binding sites for Fab and Fc, since a large molar excess of Fc does not inhibit the binding of Fab to protein G and *vice versa* (Erntell et al., 1988). The Fc binding has been shown to involve some residues from the α-helix, the extended loop, and the third β-strand (Gronenborn & Clore, 1993). A crystallographic study of the complex formed by a protein G domain and a Cγ1 domain has shown that this interaction involves a connection between the second β-strand from protein G and the C-terminal strand from the antibody molecule in a contiguous way (Derrick & Wigley, 1992; Lian et al., 1994). Some of the side chains in the C-terminus of the α-helix in protein G also provide ligands for the interaction. A single IgG-binding protein G domain therefore contains two binding regions with affinity for two different surfaces in the γ chain. When the primary and tertiary structures of the Ig-binding domains of proteins L and G are compared, the third β-strands are most homologous. This makes the involvement of this strand in the interaction between protein L and the Ig light chain highly unlikely since protein L cannot inhibit the binding of protein G to IgGFc (Wikström et al., unpublished results). The least homologous and surface-exposed regions in the comparison are the second β-strands, parts of the α-helix, and the loop following the helix. In this context it should be noted that the second β-strand shows a high degree of conservation when the five B domains of protein L are compared.

Evolutionary and Biological Considerations. Bacteria expressing surface molecules like proteins L and G will *in vivo* be coated with host Ig (in the case of protein G also with albumin), which changes their physicochemical surface properties (Miörner et al., 1980). It appears likely that the host−parasite relationship is also influenced by these protein−protein interactions. The biological consequences of the binding of Ig to bacterial surfaces are still poorly understood, although indirect evidence suggests that microorganisms gain selective advantages from the interaction with Ig. Thus, it has been demonstrated that not only protein G but also staphylococcal protein A (Forsgren & Sjöquist, 1966) and protein H of group A streptococci (Åkesson et al., 1990) all bind to the Cγ2−Cγ3 interface region, despite the fact that their IgGFc-binding regions are nonhomologous (Frick et al., 1992, 1994). The convergence in the evolution of these regions indicates that they are connected with essential microbial functions. From an evolutionary−functional point of view, the structure of the Ig-binding domains of protein L is of interest. Compared to protein G, the overall organization is similar but there are also distinct structural and binding differences which is consistent with a divergent

evolution from a common origin. Recent work has shown that a functional domain, a protein G-related albumin-binding molecule, can be transferred from streptococci to *P. magnus* and be introduced in protein L-related genes (de Chateau & Björck, 1994; Murphy et al., 1994). It remains to be demonstrated that similar events can explain the emergence of the Ig-binding protein L domains. Nevertheless, in order to understand how the Ig-binding region of these domains has evolved, especially in relation to corresponding regions in protein G and other Ig-binding bacterial proteins, it is necessary to identify and define the region interacting with Ig light chains. Using heteronuclear NMR techniques, we are currently addressing this question.

The affinity for Ig light chains enables protein L to bind antibodies of all classes, whereas, for instance, protein G binds exclusively IgG. This could have implications for pathogenicity and virulence. Streptococci and staphylococci are virulent bacteria causing an array of clinical infections in humans, but it has been difficult to ascribe a role for Ig-binding proteins in the virulence of these bacteria. *P. magnus*, on the other hand, is a member of the normal flora on most body surfaces, and compared to staphylococci and streptococci these bacteria are much less virulent. Only about 10% of *P. magnus* strains carry the protein L gene, but these strains are more frequencly associated with clinical infections, suggesting that protein L could be a virulence determinant (Kastern et al., 1990). The reason why protein L-expressing bacteria appear more virulent is not clear, but such bacteria and protein L itself can release histamine from human mast cells and basophils through interaction with cytophilic IgE (Patella et al., 1990), which could be one possible virulence mechanism. The connection between protein L and virulence adds interest to structural–evolutionary studies of this unique Ig light chain-binding bacterial protein.

ACKNOWLEDGMENT

We thank Maarten de Chateau and Bryan Finn for stimulating discussions and Ing-Britt Gustafsson and Ulrika Ringdahl for excellent technical assistance.

REFERENCES

Achari, A., Hale, S. P., Howard, A. J., Clore, A. M., Gronenborn, A. M., Hardman, K. D., & Whitlow, M. (1992) *Biochemistry 31*, 10449–10457.

Åkerström, B., & Björck, L. (1989) *J. Biol. Chem. 264*, 19740–19746.

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

Åkesson, P., Cooney, J., Kishimodo, F., & Björck, L. (1990) *Mol. Immunol. 27*, 523–531.

Aurora, R., Srinivasan, S., & Rose, G. D. (1994) *Science 264*, 1126–1130.

Björck, L. (1988) *J. Immunol. 140*, 1194–1197.

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

Boyle, M. D. P., Ed. (1990) *Bacterial Immunoglobulin Binding Proteins*, Vol. I, Academic Press, San Diego.

Brünger, A. T. (1992) in *X-PLOR. Version 3.1. A system for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT.

Constantine, K. L., Madrid, M., Banyai, L., Trexler, M., Patthy, L., & Llinas, M. (1992) *J. Mol. Biol. 223*, 281–298.

de Chateau, M., & Björck, L. (1994) *J. Biol. Chem. 269*, 12147–12151.

de Chateau, M., Nilson, B. H. K., Erntell, M., Myhre, E., Magnusson, C. G. M., Åkerström, B., & Björck, L. (1993) *Scand. J. Immunol. 37*, 399–405.

Derrick, J. P., & Wigley, D. B. (1992) *Nature 359*, 752–754.

Erntell, M., Myhre, E., Sjöbring, U., & Björck, L. (1988) *Mol. Immunol. 25*, 121–126.

Fahnestock, S. R., Alexander, P., Nagle, J., & Filpula, D. (1992) *J. Bacteriol. 167*, 870–880.

Ferrin, T. E., et al. (1988) The MIDAS display system, *J. Mol. Graphics 6*, 13–27.

Forsgren, A., & Sjöquist, J. (1966) *J. Immunol. 97*, 822–827.

Frick, I.-M., Wikström, M., Forsén, S., Drakenberg, T., Gomi, H., Sjöbring, U., & Björck, L. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 8532–8536.

Frick, I.-M., Åkesson, P., Cooney, J., Sjöbring, U., Schmidt, K.-H., Gomi, H., Hattori, S., Tagawa, C., Kishimoto, F., & Björck, L. (1994) *Mol. Microbiol. 12*, 143–151.

Gronenborn, A. M., & Clore, G. M. (1993) *J. Mol. Biol. 233*, 331–335.

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

Janin, J., & Chothia, C. (1980) *J. Mol. Biol. 143*, 95–128.

Kastern, W., Holst, E., Nielsen, E., Sjöbring, U., & Björck, L. (1990) *Infect. Immunol. 58*, 1217–1222.

Kastern, W., Sjöbring, U., & Björck, L. (1992) *J. Biol. Chem. 267*, 12820–12825.

Kehoe, M. A. (1994) *New Compr. Biochem. 27*, 217–261.

Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946–950.

Lian, L.-Y., Derrick, J. P., Sutcliffe, M. J., Yang, J. C., & Roberts, G. C. K. (1992) *J. Mol. Biol. 228*, 1219–1234.

Lian, L.-Y., Barsukov, I. L., Derrick, J. P., & Roberts, G. C. K. (1994) *Nature Struct. Biol. 6*, 355–357.

Macura, A., & Ernst, R. R. (1980) *Mol. Phys. 41*, 95–117.

Miörner, H., Myhre, E., Björck, L., & Kronvall, G. (1980) *Infect. Immun. 29*, 879–885.

Müller, L. (1987) *J. Magn. Reson. 72*, 191–196.

Murphy, P. P., Duggleby, C. J., Atkinson, M. A., Trowern, A. R., Atkinson, T., & Goward, C. R. (1994) *Mol. Microbiol. 12*, 911–920.

Myhre, E. B., & Erntell, M. (1985) *Mol. Immunol. 22*, 879–885.

Nilson, B. H. K., Solomon, A., Björck, L., & Åkerström, B. (1992) *J. Biol. Chem. 267*, 2234–2239.

Patella, V., Casolaro, V., Björck, L., & Marone, G. (1990) *J. Immunol. 145*, 3054–3061.

Ramachandran, G. N., Ramakrishnan, C., & Sasisekharan, V. (1963) *J. Mol. Biol. 23*, 283–437.

Reis, K. J., Ayoub, E. M., & Boyle, M. D. P. (1984) *J. Immunol. 132*, 3091–3097.

Stone, C. G., Sjöbring, U., Björck, L., Sjöquist, J., Barber, C. V., & Nardella, F. A. (1989) *J. Immunol. 143*, 565–570.

Wagner, G., Braun, W., Havel, T. F., Schaumann, T., Go, N., & Wüthrich, K. (1987) *J. Mol. Biol. 196*, 611–639.

Wikström, M., Sjöbring, U., Kastern, W., Björck, L., Drakenberg, T., & Forsén, S. (1993) *Biochemistry 32*, 3381–3386.

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

Yip, P. F. (1990) *J. Magn. Reson. 90*, 382–383.