J. Mol. Biol. (1993) **233**, 331–335

COMMUNICATION

Identification of the Contact Surface of a Streptococcal Protein G Domain Complexed with a Human Fc Fragment

Angela M. Gronenborn and G. Marius Clore

Laboratory of Chemical Physics
Building 5
National Institute of Diabetes and Digestive and Kidney Diseases
National Institutes of Health
Bethesda, MD 20814, U.S.A.

(Received 27 April 1993; accepted 28 May 1993)

The B1 domain of streptococcal protein G interacts with the C-terminal fragment of the heavy chain of immunoglobulin G (IgGFc). The binding site for the protein G domain on the antibody fragment is in close proximity or overlapping with that determined for staphylococcal protein A. The interaction of the B1 domain with IgGFc was investigated by ^{1}H-^{15}N correlation spectroscopy. The major interaction site on the B1 domain comprises parts of β-strand 3 as well as the α-helix. Comparison with the crystal structure of the protein A/IgGFc complex suggests that the mode of interaction in the two complexes is analogous, despite the lack of sequence or structural similarity between two antibody binding proteins.

Keywords: protein G IgG binding domain; human IgGFc complex; solution structure; heteronuclear NMR

A large number of staphylococcal and streptococcal strains express, on their cell surface, proteins that bind to mammalian immunoglobulins (for a review, see Boyle, 1990). The best characterized member in this group of molecules is protein A from *Staphylococcus aureus*, which is widely used as a reagent in a large number of immunological and immunochemical applications (Langone, 1982). Protein G is a large multi-domain cell surface protein of group G and C *Streptococci*, which exhibits a broader spectrum of binding to IgG† subclasses than protein A (Björck & Kronvall, 1984). It binds to all four human subclasses, as well as to a variety of mouse and rat monoclonal antibodies (Åkerström *et al.*, 1985). Protein G contains repeats of two or more IgG binding domains, each comprising 55 residues. Like protein A, protein G binds primarily and tightly to the Fc region of IgG (Fahnestock *et al.*, 1990), apparently interacting with the same site on the antibody. In addition, weak binding to Fab has been observed.

The crystal structure of a complex between one of the five homologous Fc-binding domains of protein A and human Fc has been solved (Deisenhofer *et al.*, 1978) and the binding site for the protein A fragment is located at the junction between the CH2 and CH3 domains of Fc (Deisenhofer, 1981). The protein A IgG binding domain is an all-helical structure with two helices involved in the interaction. Interestingly, protein A and protein G exhibit neither sequence nor structural homology in their IgG binding domains, although they appear to bind to the same site on IgGFc. A high-resolution solution structure has been determined for the B1 IgG binding domain of protein G (Gronenborn *et al.*, 1991). The structure comprises a four stranded β-sheet made up of two antiparallel β hairpins connected by an α-helix. The two central strands of the sheet are parallel and comprise the N and C-terminal residues. Comparison of the protein A and protein G IgG binding domain architectures reveals no immediately obvious region that could take the place of the two interacting helices of protein A in the protein G complex. It is, therefore, of considerable interest to investigate the manner in which these structurally very different molecules interact with the IgGFc. Here, we probe the interaction between the B1 domain of protein G and human Fc by means of ^{1}H-^{15}N correlation spectroscopy.

† Abbreviations used: IgG, immunoglobulin G; IgGFc, C-terminal fragment of the heavy-chain of the immunoglobulin G; Fab, antigen binding fragment of the immunoglobulin G; HPLC, high-pressure liquid chromatography; p.p.m., parts per million.

0022-2836/93/190331-05 $08.00/0

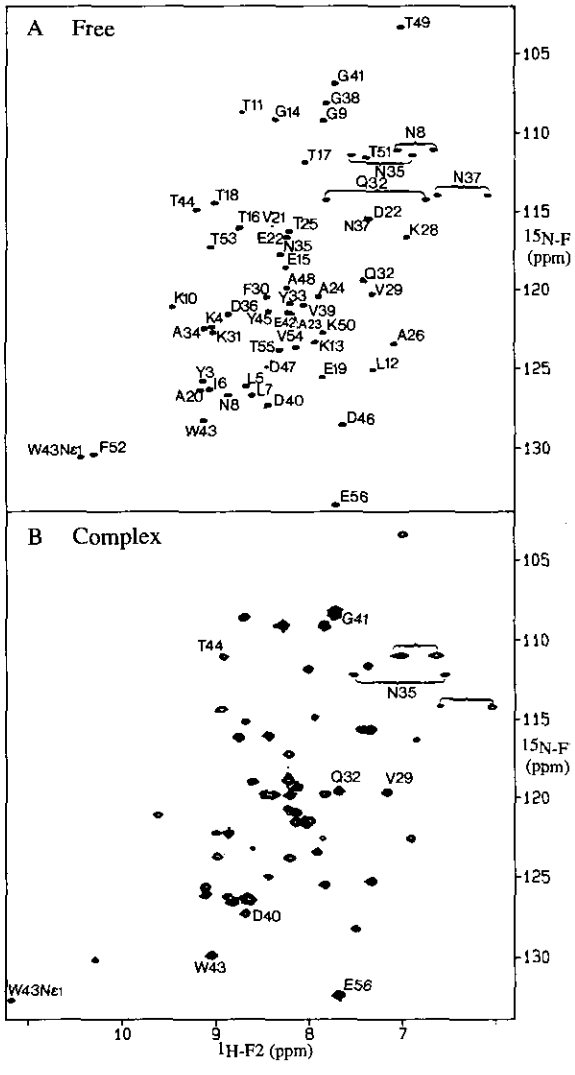

Figure 1. ^{15}N (F1 axis)- NH (F2 axis) region of the ^{15}N-^{1}H Overbodenhausen correlation spectrum of 1·1 mM IgG binding domain at 50 °C in 50 mM sodium phosphate buffer (pH 5·8), in the absence (A) and presence (B) of 1 equivalent Fc fragment. These spectra, as well as another set of spectra at 35 °C (not shown), were recorded on a Bruker AMX600 spectrometer as described previously (Bax *et al.*, 1990). ^{15}N-^{1}H correlation peaks are labelled for all resonances in A and for selected ones in B. The IgG binding domain was expressed in *Escherichia coli* grown on minimal medium with $^{15}NH_4Cl$ as the sole nitrogen source to obtain uniform (>95%) ^{15}N labelling, and partially purified as described previously (Gronenborn *et al.*, 1991). This purification procedure yields 2 species of protein, with and without the N-terminal methionine residue. The major species with the N-terminal methionine residue was then purified to homogeneity by reverse phase HPLC and was greater than 99% pure as judged by N-terminal sequencing and NMR spectroscopy. Assignments were made in a straightforward manner by reference to the previously published assignments at different pH and temperature values (Gronenborn *et al.*, 1991; PDB accession number R1GB1MR). Residues with backbone amide chemical shifts that are identical in both spectra are: T2, Y3, K4, L5, I6, L7, N8 (also N8 side-chain amino), G9, T11, L12, K13, G14, E15, T16, T17, T18, E19, A20, V21, D22, N37 (also N37 side-chain amino), G38, D47, A48, T49, K50, T51, F52, V54, T55.

NMR spectroscopy is a very powerful method for determining three-dimensional structures of proteins in solution *via* nuclear Overhauser enhancement measurements. However, even in cases where a complete structure determination is impossible, NMR can provide useful information about the system on a more qualitative basis. This is due to the fact that individual nuclei or atoms can be monitored, yielding valuable information about the molecule under study. In particular, the chemical shift of a nucleus can be employed as a probe for changes in the atomic environment, since it is extremely sensitive to any conformational or electronic influences. Thus, one can regard the pattern of resonance frequencies for a particular protein as a fingerprint for its structure under defined conditions. In this regard, the perturbation of ^{15}N and NH chemical shifts upon complex formation with either a ligand or another protein provides a highly sensitive tool for the mapping of binding sites on a protein (Chen *et al.*, 1993; van Nuland *et al.*, 1993). We have exploited this property to identify the contact region of a protein G IgG binding domain when complexed with human Fc.

Figure 1 shows the ^{1}H-^{15}N Overbodenhausen correlation spectrum (Bodenhausen & Ruben, 1980; Bax *et al.*, 1990; Norwood *et al.*, 1990) of the uniformly ^{15}N-labelled B1 IgG binding domain of protein G (Fig.1A), as well as that obtained for the G domain/IgGFc complex (Fig. 1B). The most prominent difference between the two spectra consists of the markedly increased linewidths for the crosspeaks arising from the complex. This is not surprising since the total molecular mass of the complex is approximately 62 kDa. Superposition of the two spectra reveals that, in addition to the line broadening, approximately half of the resonances are shifted in the complex, while the other half remain at exactly the same positions found in the uncomplexed protein G domain. Since both spectra were recorded under identical buffer and temperature conditions, the observed shifts in the protein G resonances have to arise either directly or indirectly from contacts with the Fc fragment. More strikingly, the resonances that do not show any shift at all must arise from those regions of the protein G domain that are not in contact with Fc and, hence, unaffected by binding. The unaffected ^{1}H-^{15}N correlation peaks ($\Delta\delta$ ^{15}N <0·2 p.p.m.; $\Delta\delta$ ^{1}H <0·02 p.p.m.) comprise those from residues 2 to 22, that is amino acid residues in β-strands 1 and 2, as well as residues 37, 38 and 47 to 52. The latter are located in β-strand 4. Shifted resonances arise from residues within the stretches 23 to 36 and 40 to 46, located in the α-helix and β-strand 3. Thus shifted and non-shifted resonances fall into two clearly defined regions of the protein structure. Figure 2 presents two views of the protein G domain in a ribbon representation, highlighting the two regions. The beginning of β-strand 4 and β-strands 1 and 2, located at the back of the structure in the view shown in Figure 2B, are clearly not involved in IgGFc binding, whereas the α-helix and β-strand 3

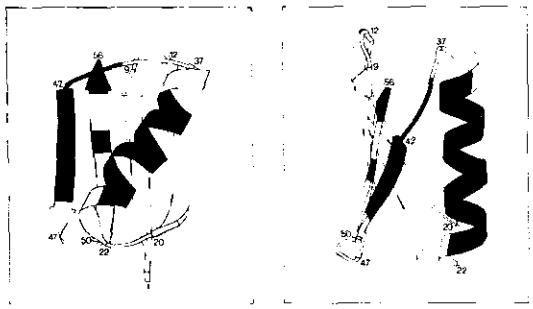

Figure 2. A ribbon drawing of 2 views of the solution structure of the IgG binding domain of protein G illustrating the location of the IgGFc binding induced chemical shift changes. Those residues whose NH proton and/or ^{15}N chemical shifts are altered in the complex as compared to the free protein G domain are shown in black. The schematic ribbon diagram was produced with the program MOLSCRIPT (Kraulis, 1991) and the structure is from Gronenborn *et al.* (1991: PDB Accession number 1GB1).

most likely form the contact site. Since chemical shift changes can arise from direct protein-protein contacts or conformational changes induced by complex formation, one cannot state with certainty that all of the α-helix and β-strand 3 contacts the Fc fragment. Nevertheless, the interaction site has to lie within these boundaries. In this regard, it is interesting to note that a small peptide fragment derived from the protein G domain consisting of the 11 amino acid residue stretch from residue 35 to 45, that is the end of the α-helix and β-strand 3, blocks the binding of the protein G domain to IgGFc (Frick *et al.*, 1992). This 11 residue peptide segment is located within the region of shifted resonances in the protein G domain/IgGFc complex, supporting the notion that indeed the α-helix and β-strand 3 constitute the binding site.

Since protein A and protein G apparently compete for the same binding site on IgGFc, we examined the X-ray structure of the B fragment of *S. aureus* protein A bound to Fc (Deisenhofer *et al.*, 1978; Deisenhofer, 1981) with the aim of elucidating how such seemingly different topological units as the all-helical B fragment and the protein G domain are able to undergo similar interactions. It is possible to superimpose one of the two interacting helices of protein A on the helix of the protein G domain, which in turn locates β-strand 3 in a similar position with respect to the Fc as the second interacting helix of protein A. Two stereoviews of this alignment between the protein G and A domains are shown in Figure 3. Thus, it is quite possible that protein A and protein G cover an almost identical region of the Fc with amino acid side-chains from β-strand 3 occupying sites used by residues from the second α-helix in protein A.

The information derived from the ^{1}H-^{15}N Overbodenhausen correlation spectrum of the complex pertains mainly to the polypeptide backbone, since most of the observed crosspeaks arise from the backbone amide proton and nitrogen atoms. There are, however, several crosspeaks that yield information about selected side-chains, such as those arising from the amino groups of glutamine and asparagine and the N^{ε} amide of the tryptophan ring. Whereas the crosspeaks assigned to Asn8 and Asn37 show no chemical shift changes in the complex, those of Asn35 are clearly shifted and those of Gln32 disappear all together. One of the most substantial shifts is observed for the Trp43 N^{ε} amide proton upon complexation, clearly indicating the involvement of this side-chain in complex formation. All of the above side-chains fall into the appropriate regions of the structure as classified by the backbone amide shifts.

Having identified the contact region on the B1 IgG binding domain of protein G in the Fc complex, we attempted to find an explanation for the difference observed in the binding of the B1 and B2 domains of protein G. The association constants for the binding of IgG to the B1 and B2 domains of protein G are $0{\cdot}3 \times 10^{8}$ and $2{\cdot}1 \times 10^{8}\ M^{-1}$, respectively (S. R. Fahnestock, personal communication). The sequence comparison shows six amino acid substitutions between the two domains: Ile6→Val, Leu7→Ile, Glu19→Lys, Ala24→Glu, Val29→Ala and Glu42→Val. Leu7 is completely buried and cannot, therefore, be involved in binding. The other five residues, on the other hand, are solvent accessible and hence potential contact residues. Of those, only three fall into the identified contact region; namely, Ala24, Val29 and Glu42. In particular, Glu42 lies on the solvent-exposed surface of the β-sheet in β-strand 3 and exhibits markedly perturbed resonances, like its immediate neighbors. Since the contact between the protein A fragment and IgGFc is predominantly formed by hydrophobic interactions (Deisenhofer, 1981) and Glu42 is part of the major contact site in the B1 complex, it seems likely that substituting this charged side-chain for a valine residue would result in a more favorable contact and thus tighter binding.

Recently, the X-ray structure of a protein G domain bound to an Fab fragment was reported (Derrick & Wigley, 1992). In this complex, β-strand 2 of protein G forms an antiparallel interaction with the last β-strand of the CH1 domain. This interaction is predominantly a backbone-backbone interaction, connecting the β-sheet of the immunoglobulin with that of the protein G domain in a contiguous fashion, reminiscent of the interaction observed in the crystal lattice of a protein G domain X-ray structure (Achari *et al.*, 1992). This type of interaction contrasts markedly with the arrangement observed in the protein A fragment/Fc complex (Deisenhofer *et al.*, 1978; Deisenhofer, 1981). In addition, the parts of the protein G domain that are found to be in contact with the CH1 domain cannot be those involved in the Fc interactions, since no chemical shift change is observed for residues located in β-strand 2 in the Fc complex. Thus, it may well be that the association observed between Fab and the protein G domain is a non-specific interaction brought about by crystal

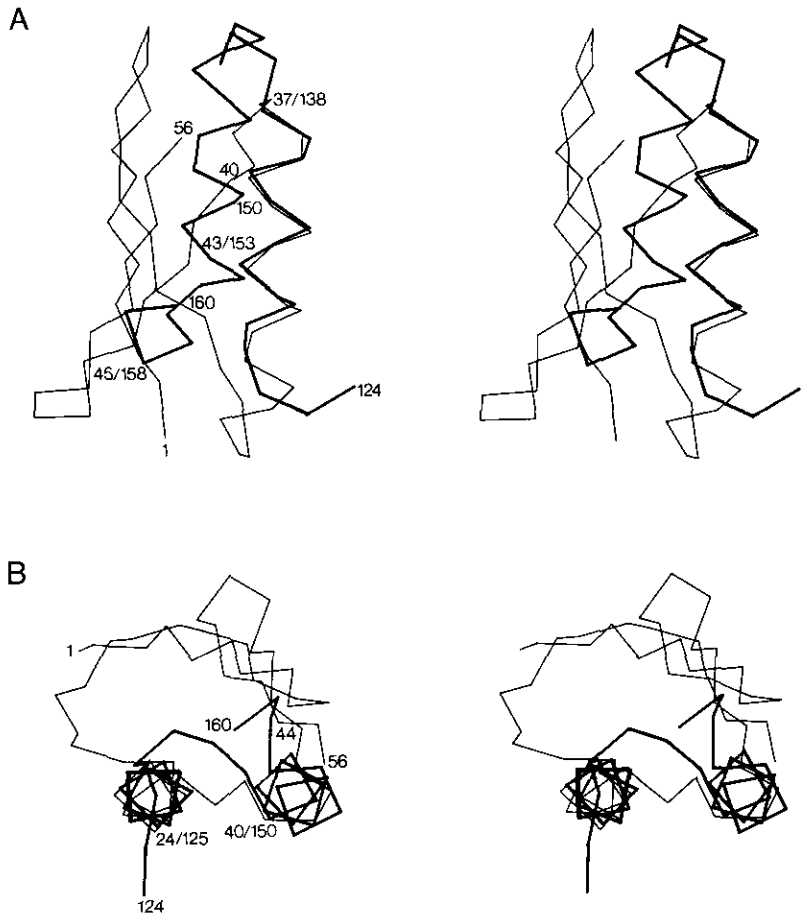

Figure 3. Two stereoviews of superpositions of the IgGFc bound protein A fragment (thick lines) and the protein G IgG binding domain (thin lines). Only residues 124 to 160 of the protein A fragment, which comprise the 2 helices (residues 127 to 138 and 143 to 156) that contact the IgGFc, are shown. The first helix of protein A is aligned with the *helix of the protein G domain, while the second helix of protein A is aligned with β-strand 3 of the protein G domain.* The C^{α} atoms of 18 residues can be superimposed with root-mean-square difference of 1·3 Å (namely, residues 24 to 37 and 40 to 43 of the protein G domain on residues 125 to 138 and 150 to 153, respectively, of the protein A fragment). The view in A is the same as that shown in the right-hand panel of Fig. 2.

packing forces, similar to that observed between neighboring molecules in the crystal lattice of the protein G IgG binding domain alone. Thus, we believe that the primary binding site on the protein G domain for IgG is made up of β-strand 3 and the α-helix, resulting in similar interactions to those observed in the protein A/Fc complex.

We thank Drs Marc Whitlow, Nina Essig and David Filpula for providing us with the IgG binding domain clone of streptococcal protein G. This work was supported by the AIDS Targeted Anti-Viral Program of the Office of the Director of the National Institutes of Health.

References

Achari, A., Hale, S. P., Howard, A. J., Clore, G. M., Gronenborn, A. M., Hardman, K. D. & Whitlow, M. (1992). 1·67-Å X-ray structure of the B2 immunoglobulin-binding domain of streptococcal protein G and comparison to the NMR structure of the B1 domain. *Biochemistry*, **31**, 10449–10457.

Åkerström, B., Brodin, T., Reis, K. & Björck, L. (1985). Protein G: a powerful tool for binding and detection of monoclonal and polyclonal antibodies. *J. Immunol.* **135**, 2589–2592.

Bax, A., Ikura, M., Kay, L. E., Torchia, D. A. & Tschudin, R. (1990). Comparison of different modes of two-dimensional reverse correlation NMR for the study of proteins. *J. Magn. Reson.* **86**, 304–318.

Björck, L. & Kornvall, G. (1984). Purification and some properties of streptococcal protein G, a novel IgG-binding reagent. *J. Immunol.* **133**, 969–974.

Bodenhausen, G. & Ruben, D. J. (1980). Natural abundance 15-nitrogen NMR by enhanced heteronuclear spectroscopy. *Chem. Phys. Letters*, **69**, 185–189.

Boyle, N. D. P. (1990). *Bacterial Immunoglobulin Binding Proteins*, Academic Press, San Diego.

Chen, Y., Reizer, J., Saier, M. H., Fairbrother, W. J. & Wright, P. E. (1993). Mapping of the binding interfaces of proteins of the bacterial phosphotransferase system HPr and IIA^{glc}. *Biochemistry*, **32**, 32–37.

Deisenhofer, J. (1981). Crystallographic refinement and atomic models of a human Fc fragment and its complex with fragment B of protein A from *Staphylococcus aureus* at 2·9- and 2·8-Å resolution. *Biochemistry*, **20**, 2361–2370.

Deisenhofer, J., Jones, T. A., Huber, R., Sjödahl, J. & Sjöquist, J. (1978). Crystallization, crystal structure

analysis and atomic model of the complex formed by a human Fc fragment and fragment B of protein A from *Staphylococcus aureus*. *Hoppe-Seyler's Z. Physiol. Chem.* **359**, 975–985.

Derrick, J. P. & Wigley, D. B. (1992). Crystal structure of a streptococcal protein G domain bound to an Fab fragment. *Nature (London)*, **359**, 752–754.

Fahnestock, S. R., Alexander, P., Filpula, D. & Nagle, J. (1990). Structure and evolution of the streptococcal genes encoding protein G. In *Bacterial Immunoglobulin Binding Proteins* (Boyle, N.D.P., ed.), vol. 1, pp. 133–148, Academic Press, San Diego.

Frick, I.-M., Wikström, M., Forsén, S., Drakenberg, T., Gomi, H., Sjöbring, U. & Björck, L. (1992). Convergent evolution among immunoglobulin G-binding bacterial proteins. *Proc. Nat. Acad. Sci., U.S.A.* **89**, 8532-8536.

Gronenborn, A. M., Filpula, D. R., Essig, N. Z., Achari, A., Whitlow, M., Wingfield, P. T. & Clore, G. M. (1991). A novel highly stable fold of the immunoglobulin binding domain of Streptococcal protein G. *Science*, **253**, 657–661.

Kraulis, P. J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950.

Langone, J. J. (1982). Protein A of *Staphylococcus aureus* and related immunoglobulin receptors produced by Streptococci and Pneumococci. *Advan. Immunol.* **32**, 157–252.

Norwood, T. J., Boyd, J., Heritage, J. E., Soffe, N. & Campbell, I. D. (1990). Comparison of techniques for ^{1}H-detected heteronuclear ^{1}H-^{15}N spectroscopy. *J. Magn. Reson.* **87**, 488–501.

van Nuland, N. A. J., Kroon, G. J. A., Dijkstra, K., Wolters, G. K., Scheek, R. M. & Robillard, G. T. (1993). The NMR determination of the IIA^{mtl} binding site on HPr of the *Escherichia coli* phosphoenol pyruvate-dependent phosphotransferase system. *FEBS Letters*, **315**, 11–15.

Edited by P. E. Wright