The Structural Basis of Germline-encoded V₃ Immunoglobulin Binding to Staphylococcal Protein A

By Jan L. Hillson, Nancy S. Karr, Ina R. Oppliger, Mart Mannik, and Eric H. Sasso

From the Division of Rheumatology, University of Washington, Seattle, Washington 98195

Summary

The ability of human V₃ immunoglobulins (Ig) to bind to staphylococcal protein A (SPA) via their Fab region is analogous to the binding of bacterial superantigens to T cell receptors. The present report establishes the structural basis for the interaction of SPA and V₃ Ig. We have studied a panel of 27 human monoclonal IgM that were derived from fetal B lymphocytes. As such, these IgM were expected to be encoded by unmutated germline genes. Binding to SPA in ELISA occurred with 15 of 15 V₃ IgM, but none of 12 IgM from the V₁, V₄, V₅, or V₆ families. The V₃ sequences of the 27 IgM were derived from 20 distinct V₃ elements, including 11 from the V₃ family. Use of D, J₃, and C genes was similar among V₃ and non-V₃ IgM. A comparison of the corresponding V₃ protein sequences, and those of previously studied IgM, identified a probable site for SPA binding that includes V₃ residues in framework region 3 (FR3), and perhaps FR1 and 3' complementary determining region 2. The results thus demonstrate that among human IgM, specificity for SPA is encoded by at least 11 different V₃ germline genes. Furthermore, like the T cell superantigens, SPA likely binds to residues in the V₃ framework region, outside the classical antigen-binding site of the hypervariable loops.

Materials and Methods

Monoclonal B Cell Lines. Mononuclear cells were isolated by density gradient centrifugation from second trimester fetal liver and spleen provided by the University of Washington Central Laboratory for Human Embryology (11). Fetal mononuclear cells were transformed by EBV and cloned by limiting dilution, as previously described (12). Studies described below were performed with 26 lines, randomly selected from among 62 IgM-producing cell lines (12). Beg-2, a human heterohybridoma generated from fetal spleen cells, was kindly provided by Dr. Richard Watts (University College, London, England).

Determination of IgM Concentration and L Chain Isotype. Supernatants of the 27 monoclonal IgM cell lines were assayed by ELISA in 96-well trays coated with goat F(ab')₂ anti-human F(ab')₂ (Jackson ImmunoResearch Laboratories, Inc., West Grove, PA). L chain isotype was determined by detecting bound IgM with goat F(ab')₂ anti-human κ and goat F(ab')₂ anti-human λ, both conjugated...
with horseradish peroxidase (Cooper Biomedical, Malvern, PA). IgM concentration was determined by detecting bound Ig with horseradish peroxidase-conjugated goat F(ab')2 anti-human IgM, Fc-specific (Cooper Biomedical), and comparing results with a standard curve prepared with purified polyclonal human IgM (Calbiochem Novabiochem, La Jolla, CA). For subsequent binding studies, every supernatant was adjusted to >0.1 μg IgM/ml in tissue culture media (IMDM; Sigma Chemical Co., St. Louis, MO) that was supplemented with 10% FCS, 2 mM l-glutamine, 50 U/ml penicillin, and HAT (Boehringer Mannheim, Indianapolis, IN). All supernatants were shown to be devoid of human IgG by detection with horseradish peroxidase-conjugated Fc-specific goat F(ab')2 anti-human IgG.

**IgM Binding to SPA.** IgM binding to SPA was determined by ELISA. Wells were coated with 1.5 μg SPA (Sigma Chemical Co.) in 150 μl bicarbonate buffered saline, pH 8.0, washed, incubated with 100 μl supernatant, and washed with borate buffered saline, pH 8.0, containing 0.5% Tween 20. Horseradish peroxidase-conjugated F(ab')2 goat anti-human IgM, Fc-specific, was then incubated in the wells, followed by development with 2,2'-azino-di-[3-ethylbenzthiazoline sulphonate] (ABTS) substrate (Kirkegaard-Perry, Gaithersburg, MD). By the same protocol, each supernatant was also tested for binding to wells coated with tissue culture medium supplemented with 10% FCS (negative control), and to wells coated with goat F(ab')2 anti-human F(ab')2 (positive control). All assays were performed in duplicate.

**IgM Vn Sequence Analysis.** As previously described, Ig H chain cDNA was prepared from 1 μl culture suspension, containing 1-25 multinucleous cells, with a synthetic deoxyoligonucleotide primer corresponding to the 5' region of Cμ (5'-GACGGAATTCTCACAG-3') (12). The cDNA was then amplified by the PCR with codons 1 to 8 (5'-CAGGTGCAGCTGGTGAATTCTGG-3') (12). The PCR product was ligated into Puc18 and sequenced by conventional methods (13, 14). Each reported sequence represents a consensus of sequences from at least two independent cDNA, each of which was sequenced the entire reported length. Nucleotide mismatches occurred at a rate of one per 10,000 bp.

**Results**

**Measurement of Binding to SPA by 27 Monoclonal IgM.** A panel of 27 monoclonal IgM, each of which was expressed by a B cell line derived from fetal liver or spleen, was tested for binding to SPA in a solid phase ELISA. All 15 IgM encoded by elements from the Vn3 gene family bound to SPA (Fig. 1). In contrast, binding to SPA was not seen with any of the nine IgM from the Vn1 family, or the IgM from the Vn4, Vn5, or Vn6 families (one each) (Fig. 1). All IgM bound well to the positive control, and poorly or not at all to the negative control (data not shown). Only the Vn4 IgM (OD = 0.05) and the Vn6 IgM (OD = 0.03) bound the negative control greater than 0.01 OD units.

**Vn Sequence Analysis of 27 Monoclonal IgM.** The Vn nucleotide sequence of each IgM was determined. The IgM were encoded by 11 different Vn3 elements, six different Vn1 elements, and one each from the Vn4, Vn5, and Vn6 families (Fig. 1). All but five of the Vn3 and two of the Vn1 sequences were identical to known germline genes (Fig. 1). A variety of Dn and Jn segments was found, with a similar distribution among the Vn3 and non-Vn3 IgM (Fig. 1). The 27 Jn sequences were all ascribable to known germline genes, and none contained somatic mutations. IgM with the same Vn sequences were all clonally distinct because they differed in their D sequences, Jn sequences, or both. k and λ L chains were equally represented among the Vn3 and the non-Vn3 IgM (Fig. 1).
Each of the 20 different \( V_n \) nucleotide sequences encoded a different translated protein sequence, i.e., six different proteins from the \( V_3 \) family, 11 from \( V_1 \), and one each from \( V_4 \), \( V_5 \), and \( V_6 \) (Fig. 2). Comparison of these protein sequences identified 23 positions at which the amino acid residue is invariant among the \( V_3 \) sequences and different from the residue(s) present in the \( V_1 \) sequences. At position 82a also, the \( V_3 \) residues are identical, and different from the \( V_1 \) residues, except for a conservative substitution in one \( V_3 \) sequence, 3G11 (Fig. 2).

Figure 2. Protein sequences of the H chain V regions of monoclonal IgM. Amino acids were translated from the obtained nucleotide sequences of 27 IgM described in Fig. 1, and are shown from codon 9 through codon 114. Sequences are grouped by \( V_n \) gene family and ability to bind to protein A. For comparison, the \( V_3 \) sequences of six previously reported \( V_3 \) IgM are shown at the bottom of the \( V_3 \) group, with their names italicized and marked with an asterisk (4, 25–28). (Top) Sequence from clone 2A10. Amino acids of other sequences are specified only where different from 2A10. (Bottom) \( X \) Position at which a single amino acid residue is identical among and unique to all our \( V_3 \) sequences. (x) All \( V_3 \) residues at that position are identical or conserved, and also present in the \( V_4 \), \( V_5 \), or \( V_6 \) sequences, but not the \( V_1 \) sequences. The sequences of the previously reported sequences (Pom, Lay, Riv, KL1, SJ1, and TS2) are excluded from the analysis leading to the assignment of \( X \) and \( x \) designators. Sequence organization and CDR are according to Kabat et al. (8). The nucleotide sequences from which these amino acid sequences were generated are available from EMBL/GenBank/DDJB among the sequences having the accession numbers L04323–L04346 and L03815-L03830.

The fetal origin of the studied IgM makes it likely that the \( V_3 \) sequences they contained were encoded by unmutated genes, and were not subject to selection by exogenous Ag. In fact, 12 of the 12 distinct \( V_n \) sequences, including six from the \( V_3 \) family, were identical to those of known \( V_3 \) germline genes. The finding that all IgM had unmutated \( J_n \) sequences provides additional evidence that the novel \( V_n \) sequences, four \( V_3 \) and one \( V_1 \), are probably unmutated, and therefore identical to \( V_n \) germline sequences that have not yet been reported. Thus, the data clearly identify 11 distinct \( V_n \) genes that encode SPA-binding IgM. Furthermore, specificity for SPA was encoded by the unmutated form of at least six, and probably all 11, of these \( V_3 \) germline genes.

The haploid genome is estimated to contain between 25 and 50 \( V_3 \) germline gene loci, up to two thirds of which are functional (16, 29). Thus, about 20–40% of \( V_3 \) germline loci, and a larger portion of functional \( V_3 \) germline loci, have now been directly demonstrated to encode SPA-binding Ig. It seems likely that many of the remaining \( V_n \) loci will also encode SPA-binding proteins. In previous studies of polyclonal Ig purified from blood, a small minority of total \( V_3 \) IgM, and a larger subset of total \( V_3 \) IgA and \( V_3 \) IgG...
did not demonstrate Fab-mediated binding to SPA (4, 5). The elements encoding these SPA-nonbinding V\textsubscript{\textalpha}3 Ig could have lost specificity for SPA through somatic modification, or they might belong to a subset of V\textsubscript{\textalpha}3 germline genes that encode SPA-nonbinding proteins. It is also possible that certain combinations of D, J\textsubscript{\textalpha}, V\textsubscript{\textalpha}, and J\textsubscript{L} genes abrogate V\textsubscript{\textalpha}3 binding to SPA. However, the broad variety of these genes (D, J\textsubscript{\textalpha}, V\textsubscript{\textalpha}, J\textsubscript{L}) that have been found to encode SPA-binding IgM argues that this effect occurs infrequently, if at all (4, 28, 30).

The site on Ig Fab that binds to SPA has been localized to the variable region of the H chain (31), and shown to be functionally distinct from a conventional hapten-binding site (32). Our data now provide structural evidence that SPA binds outside the classical antibody binding site. The V\textsubscript{\alpha} sequences from our IgM identify 24 amino acid positions at which all V\textsubscript{\alpha}3 sequences have a conserved residue, and all V\textsubscript{\alpha}1 sequences differ by a nonconservative change. Two of these positions are in FR2, which is inaccessible to solvent (8). The remaining 22 positions localize a candidate binding site for SPA to two peptides, one in FR1 (residues 9-27), the other in 3' CDR2/FR3 (residues 62-84). Seven of these positions, indicated at the bottom of Fig. 2 (X), are strongly associated with IgM ability to bind to SPA, because their V\textsubscript{\alpha}3 residues are unique. The other 15 positions, (x), could also have a role in SPA binding, even though their V\textsubscript{\alpha}3 amino acids appear in the V\textsubscript{\alpha}6 sequence (13 of 15 positions), the V\textsubscript{\alpha}4 sequence (12 positions), or the V\textsubscript{\alpha}5 sequence (5 positions). In an intact, folded Ig molecule, the two peptides reside in closely adjacent, solvent-exposed, \beta-pleated sheets that define a region on the lateral aspect of the Fab molecule that is removed from the V\textsubscript{\alpha} hypervariable loops (8) (Fig. 3). This V\textsubscript{\alpha} structure is structurally analogous to the region bound by T cell superantigens on the \beta chain of TCRs (33, 34).

Further insight into the protein A-binding site of V\textsubscript{\alpha}3 Ig can be gained by examining six previously reported V\textsubscript{\alpha} sequences, from IgM Pom, Lay, Riv, KL1, SJ1, and TS2 (Fig. 2) (25-27). IgM Pom, Lay, and KL1 bind to SPA (4, 28). In contrast, IgM SJ1 and TS2 have been reported to not bind to SPA (28), even though their protein sequences are 94 and 96% identical, respectively, to those encoded by 1.9III and 56p1, which we found to bind to SPA (Fig. 1). A sequence comparison that includes SJ1 and TS2 identifies five FR3 positions, 75, 76, 80, 82a, and 84, at which nonconservative substitutions in a V\textsubscript{\alpha}3 sequence are associated with inability to bind protein A (Fig. 2). Each of these positions was also identified by the above analysis of our V\textsubscript{\alpha}3 sequences (Fig. 2). Thus, some of these substitutions could have abrogated SPA binding. The other substitutions in SJ1 and TS2 are unlikely to have abrogated SPA binding because, either (a) they resulted in a conservative amino acid change (residues 28 in FR1, and 59 in CDR2); (b) they occurred at a site that is inaccessible to SPA (residue 40 in FR2); or (c) in other V\textsubscript{\alpha}3 IgM, nonconservative substitutions at that position were associated with retained ability to bind SPA (residues 52, 52a, and 57 in CDR2) (Fig. 2).

These findings suggest that the FR3 75–84 peptide contains a determinant that is critical for binding to SPA. Residues 75–84 begin in a FR3 loop, and extend to include a portion of FR3 that is nearer to the C region than to the CDR (Fig. 3). Among residues 75–84, nonconservative substitutions that do not abrogate SPA binding are found in IgM KL1 (T to P, residue 77); Lay (R to Q, residue 83); and Pom (Q to L and R to Q, residues 81 and 83) (Fig. 2). It seems likely, therefore, that only a limited number of residues in the 75 to 84 peptide is directly involved in the SPA-V\textsubscript{\alpha}3 interaction. This prediction can be directly tested by site-directed mutagenesis.

In conclusion, we have studied the SPA-binding properties of a panel of monoclonal human IgM produced by B lymphocytes derived from fetal liver or spleen. Binding occurred with all V\textsubscript{\alpha}3 molecules, but no others, indicating that specificity for SPA is encoded by at least 11 different V\textsubscript{\alpha}3 germline genes. Analysis of IgM V\textsubscript{\alpha} sequences indicated that conserved residues in V\textsubscript{\alpha}3 FR1 and 3' CDR2/FR3 likely play a role in SPA recognition by V\textsubscript{\alpha}3 proteins. Furthermore, the site to which SPA binds might directly involve residues in the FR3 75–84 peptide. These findings elucidate the structural basis of the nonclassical binding specificity that V\textsubscript{\alpha}3 Ig have for SPA, and demonstrate a structural analogy between the SPA-Fab interaction, and the binding of bacterial superantigens to TCRs.
The excellent technical assistance of Ms. Cynthia E. Merrill is greatly appreciated.

This work was supported in part by National Institutes of Health grants AR-40237, AR-12839, and AR-40561.

Address correspondence to Dr. Eric H. Sasso, Division of Rheumatology RG-28, University of Washington, Seattle, WA 98195.

Received for publication 3 December 1992.

References

1. Langone, J.J. 1982. Staphylococcus aureus and related immunoglobulin receptors produced by streptococci and pneumococci. Adv. Immunol. 32:157.

2. Inganas, M., G.O. Johansson, and J. Sjoquist. 1981. Further characterization of the alternative protein-A interaction of immunoglobulins: demonstration of a Fc-binding fragment of protein A expressing the alternative reactivity. Scand. J. Immunol. 14:379.

3. Deisenhofer, J. 1981. Crystallographic refinement and atomic models of human Fc fragment and its complex with fragment B of protein A from Staphylococcus aureus at 2.9- and 2.8-A resolution. Biochemistry. 20:2361.

4. Sasso, E.H., G.J. Silverman, and M. Mannik. 1989. Human IgM molecules that bind Staphylococcal protein A contain V\(\mu\)III H chains. J. Immunol. 142:2778.

5. Sasso, E.H., G.J. Silverman, and M. Mannik. 1991. Human IgA and IgG Fab' fragments that bind to Staphylococcal protein A belong to the V\(\mu\)III subgroup. J. Immunol. 147:1877.

6. Ibrahim, S., M. Kaartinen, I. Seppala, A. Matoso-Ferreira, and O. Makela. 1993. The alternative binding site for Protein A in the Fab fragment of immunoglobulins. J. Immunol. 150:257.

7. Capra, J.D., and J.M. Kehoe. 1974. Variable region sequences of five human immunoglobulin heavy chains in the V\(\mu\)III subgroup: definitive identification of four heavy chain hypervariable regions. Proc. Natl. Acad. Sci. USA. 71:845.

8. Kabat, E.A., T.T. Wu, M. Reid-Miller, H.M. Perry, K.S. Gottesman, and C. Foeller. 1987. Sequences of Proteins of Immunological Interest, U.S. Department of Health and Human Services, United States Government Printing Office, Bethesda, Md. 2003-220.

9. Capra, J.D., and D.G. Klapper. 1976. Complete acid sequence of the Vd FLP which confers susceptibility to autoimmunity diseases. J. Clin. Invest. 88:193.

10. Chen, P.P., P.-F. Liu, S. Sinha, and D.A. Carson. 1988. A 16/6 idiotype-positive anti-DNA antibody is encoded by a conserved VH gene with no somatic mutation. Arthritis Rheum. 31:1429.

11. Ravetch, J.V., U. Siebenlist, S. Korsmeyer, T. Waldmann, and P. Leder. 1981. Structure of the human immunoglobulin \(\mu\) locus: characterization of embryonic and rearranged J and D genes. Cell. 27:583.

12. Ichihara, Y., H. Matsuoka, and Y. Kurosawa. 1988. Organization of human immunoglobulin heavy chain diversity gene loci. EMBO (Fur. Mol. Biol. Organ.) J. 7:4141.

13. Matsuda, F., E.K. Shin, Y. Hirabayashi, H. Nagaoaka, M.C. Yoshida, S.Q. Zong, and T. Honjo. 1990. Organization of variable region segments of the human immunoglobulin heavy chain: duplication of the D\(\delta\) cluster within the locus and interchromosomal translocation of variable region segments. EMBO (Fur. Mol. Biol. Organ.) J. 9:2501.

14. Buluwela, L., D.G. Albertson, P. Sherrington, P.H. Rabbitts, N. Spurr, and T.H. Rabbitts. 1988. The use of chromosomal translocations to study human immunoglobulin gene organization: mapping D\(\delta\) segments within 35 kb of the C\(\mu\) gene and identification of a new D\(\delta\) locus. EMBO (Fur. Mol. Biol. Organ.) J. 7:2003.

15. Schroeder, H.W., Jr., J.L. Hillson, and R.M. Perlmutter. 1987. Early restriction of the human antibody repertoire. Science (Wash. DC). 238:791.

16. Berman, J.E., S.J. Mellis, R. Pollock, C.L. Smith, H. Suh, B. Heinke, C. Kowal, U. Surti, L. Chess, C.R. Cantor, and F.W. Alt. 1988. Content and organization of the human Ig V\(\kappa\) locus: definition of three new V\(\kappa\) families and linkage to the Ig C\(\kappa\) locus. EMBO (Fur. Mol. Biol. Organ.) J. 7:727.

17. O. Makela. 1993. The alternative binding site for Protein A expressed in the alternative reactivity. J. Immunol. 147:1877.

18. Hillson, J.W., I.R. Oppliger, E.H. Sasso, E.C.B. Milner, and T.H. Rabbitts. 1988. The use of chromosomal translocations to study human immunoglobulin gene organization: mapping D\(\delta\) segments within 35 kb of the C\(\mu\) gene and identification of a new D\(\delta\) locus. EMBO (Fur. Mol. Biol. Organ.) J. 7:2003.

19. Sonntag, E., B. Weingartner, and R. Grutzmann. 1989. Structure and evolution of mammalian V\(\kappa\) families. Int. Immunol. 2:41.

20. Boyum, A. 1967. Isolation of leukocytes from human blood. Scand. J. Clin. Lab. Invest. 9:745.

21. Boyum, A., I.R. Oppliger, E.H. Sasso, E.C.B. Milner, and M.H. Wener. 1981. Human heavy-chain variable-region genes in human autoantibodies. In Immunoglobulin Genes. Academic Press Limited, London. 203-220.
27. Pascual, V., I. Randen, K. Thompson, M. Sioud, O. Forre, J. Natvig, and J.D. Capra. 1990. The complete nucleotide sequences of the heavy chain variable regions of six monospecific rheumatoid factors derived from Epstein-Barr virus-transformed B cells isolated from the synovial tissue of patients with rheumatoid arthritis. Further evidence that some autoantibodies are unmutated copies of germ line genes. J. Clin. Invest. 86:1320.

28. Thompson, K.M., I. Randen, J.B. Natvig, R.A. Mageed, R. Jefferis, D.A. Carson, H. Tighe, and O. Forre. 1990. Human monoclonal rheumatoid factors derived from the polyclonal repertoire of rheumatoid synovial tissue: incidence of cross-reactive idiotopes and expression of \( V_\alpha \) and \( V_\kappa \) subgroups. Eur. J. Immunol. 20:863.

29. Kodaira, M., T. Kinashi, I. Umemura, M. Fumihiko, N. Takafumi, Y. Ono, and T. Honjo. 1986. Organization and evolution of variable region genes of the human immunoglobulin heavy chain. J. Mol. Biol. 190:529.

30. Pascual, V., K. Victor, I. Randen, K. Thompson, J.B. Natvig, and J.D. Capra. 1992. IgM rheumatoid factors in patients with rheumatoid arthritis derive from a diverse array of germline immunoglobulin genes and display little evidence of somatic variation. J. Rheumatol. (Suppl. 32) 19:50.

31. Vidal, M.A., and F.P. Conde. 1985. Alternative mechanism reactivity of Protein A-immunoglobulin interaction: the \( V_\alpha \)-associated reactivity of a monoclonal human IgM. J. Immunol. 135:1232.

32. Young, W.W., Y. Tamura, D.M. Wolock, and J.W. Fox. 1984. Staphylococcal protein A binding to the Fab fragments of mouse monoclonal antibodies. J. Immunol. 133:3163.

33. Pullen, A., T. Wade, P. Marrack, and J. Kappler. 1992. Identification of the region of the T cell receptor \( \beta \) chain that interacts with the self-superantigen Mls-1\(^+\). Cell. 61:1365.

34. Choi, Y., A. Herman, D. DiGiusto, T. Wade, P. Marrack, and J. Kappler. 1990. Residues of the variable region of the T-cell receptor \( \beta \) chain that interact with S. aureus toxin superantigens. Nature (Lond.). 346:471.

336 Germline V\( \alpha \)3 Ig Binding to Protein A