Conformational Similarity and Systematic Displacement of Complementarity Determining Region Loops in High Resolution Antibody X-ray Structures*

(Received for publication, April 26, 1995, and in revised form, July 26, 1995)

Jürgen Bajorath‡§, Linda Harris‡, and Jiri Novotny¶

From the Bristol-Myers Squibb Research Institute,‡ Seattle, Washington 98121 and ¶Princeton, New Jersey 08543

Comparison of seven high resolution x-ray structures shows that the conformations of canonical complementarity determining region (CDR) loops, which are shared by these antibodies, are very similar. However, large spatial displacements (up to 2.7 Å) of the essentially identical CDR loops become evident when the antibody β-sheet frameworks, to which the loops are attached, are least-squares superposed. The loop displacements follow, and amplify, small positional differences in framework/loop splice points. Intradomain structural variability and, to a lesser extent, domain-domain orientation appear to cause the observed loop divergences. The results suggest that the selection of framework regions for loop grafting procedures is more critical than previously thought.

Immunoglobulin variable domains, VL and VH, associate noncovalently to form the Fv, a dimer of antiparallel, eight-stranded β-sandwiches (1, 2). The VL and VH β-sheets from different antibodies are nearly identical in three dimensions. However, the six complementarity determining region (CDR) loops (L1–L3, H1–H3), which connect the β-strands of the conserved framework and encode antigenic specificities, are much more variable in both sequence and conformation (2, 3). Chothia, Lesk, and colleagues (4, 5) identified sets of similar “canonical” conformations for all CDR loops except H3. Between 50 and 95% of antibody sequences are consistent with the classified canonical conformations (4), which are determined by conserved interactions of only a few key residues (“structural determinants”) within the loop and/or the framework regions. Some differences in the position of canonical CDR loops relative to superposed framework regions by comparing x-ray structures (4) or x-ray and modeled structures (5) were previously observed. However, these effects were observed in structures determined at medium resolution, generally considered minor (4, 5), and not systematically explored.

We have compared, via least-squares superpositions, the Fvs of seven x-ray structures refined to high resolution (better than 2 Å) available in the Brookhaven Protein Data Bank (6). It was anticipated that a systematic comparison of structures determined to such high precision would (i) shed light on the general relation between CDR loop conformation and position and (ii) help to assess the limitations of a procedure widely used in comparative model building, i.e. splicing of loops from a known

---

*The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§To whom correspondence should be addressed: Bristol-Myers Squibb Research Inst., 3005 First Ave., Seattle, WA 98121. Tel.: 206-727-3612; Fax: 206-727-3602; E-mail: bajorath@protocol.bms.com.

1 The abbreviations used are: VL, variable light chain; VH, variable heavy chain; CDR, complementarity determining region; Fv, variable fragment; rms, root mean square.
x-ray structure onto the conserved structural scaffold of an antibody model (7, 8).

Fig. 1 shows a comparison (9) of the amino acid sequences of the seven antibodies, which include various heavy and light chains (both \( \kappa \) and \( \lambda \)) from free as well as antigen-complexed antibodies. The 4-4-20 Fv, the highest resolution structure with \( \kappa \) light chain, was used as the template on which backbone segments of the other Fvs were superposed. Cumulative backbone root mean square (rms) deviations of the \( \beta \)-strands were determined after superposing each of these antibodies on the 4-4-20 Fv. Two alternative least-squares superpositions were used, employing different pairs of equivalence residues. The S1 set of residues consisted of only the most conserved regions of the Fvs, i.e. the four short 4-residue segments (10) of the central \( \beta \)-sheets. The S2 set consisted of a more extended set of residues and included the majority of the \( \beta \)-sheet framework (Fig. 1) akin to Stanfield et al. (11). As can be seen in Table I, the cumulative backbone rms deviations were small and the results obtained with the two superposition sets were similar.

Two different measures of differences in loop conformations were calculated. The “direct rms” value (d_rms) was calculated for the direct, pairwise, least-squares superposition of the CDR backbones only. This value can be regarded as a similarity measure of the loop conformations per se, regardless of any structural context. The “spatial rms” (s_rms) value, on the other hand, was calculated for the loop backbones overlaid as a result of the superposition of the conserved framework regions, i.e. the S1 and S2 sets. Thus, the s_rms value includes not only the differences in CDR loop conformations but also those in CDR loop positions. This value reflects the take-off angles and the ensuing rigid-body shifts in the spatial orientations of CDR loops. A comparison of two CDR loops, which have a similar conformation but are spatially displaced, will yield a low d_rms and a high s_rms value.

Analysis of loop conformations shows that all seven antibodies in our comparison set share canonical structure types (5) for CDR loops H1 and L2 and that three antibodies share equivalent canonical structures for CDR loops H2 and L3. Table II lists the results of the pairwise rms comparisons for the H1, H2, L2, and L3 loops with equivalent canonical structure. As can be seen, the backbone conformations (d_rms) of loops from the same canonical group are remarkably similar (≈0.5 Å for H1 and ≈0.2 Å for L2) in the high resolution structures. In contrast, the s_rms values show much greater differences. The superposition sets S1 and S2 gave similar cumulative backbone rms deviations for the framework segments (d_rms in Table I) and similar s_rms values (s_rms1 and s_rms2, respectively) for the compared CDR loops. Fig. 2 provides graphical demonstrations of the CDR loop d_rms and s_rms superpositions. The results indicate that positional differences of CDR loops are a rule in highly refined antibody x-ray structures. Thus, whereas the conformations of equivalent canonical CDR loops are very similar in the set of highly refined antibody x-ray structures, their spatial positions are different.

Relative orientations of the VH-VL domains in the Fvs included in our comparison showed variations typically seen before (12), i.e. up to ≈10°. It should be noted that our way of superposing and comparing the Fv structures differs fundamentally from conventional comparisons, which rely on optimal
least-squares superposition of one domain followed by measurement of angular deviation (from the pseudo-dyad axis of the Fv) of the other domain. Since different variable domains have different primary structures, numbers of residues, and molecular masses, the results of superpositions of selected domain pairs are expected to vary somewhat from case to case, which is critical for Fv comparison. Thus, we have tried to employ a more uniform procedure to compare Fvs by superposing the few invariant VL and VH residues that participate in the formation of the VH-VL domain interface (10).

In addition to the s_rms deviations reported in Table II, we have also measured s_rms values for CDRs after independent superposition of the isolated VH and VL domains, respectively. These calculations were carried out for VH and VL domains that share equivalent canonical structures for H1 and H2 and for L2 and L3, respectively. The results are shown in Table III. The comparison shows that the VH-VL domain orientation contributes to the positional differences of the CDRs. However, large s_rms values are also observed after superposition of single domains. In this comparison, intradomain framework/loop variability appears to be the primary cause of loop positional differences in both VH and VL domains, contributing, on the average, about 75% to the structural variability (as expressed in terms of s_rms values). Therefore, effects such as differences in side chain packing of the β-sheet framework, CDR-CDR interactions, or intermolecular interactions in the crystal environment must contribute to positional CDR loop differences. These spatial deviations cannot simply be a consequence of antigen binding as they are observed even when comparing the five structures crystallized in the absence of antigen.

Fig. 3 shows the superposition of the Fv framework regions and illustrates the presence of systematic structural deviations. The central strands of the β-sheets gradually diverge toward their termini, and edge strands often display even wider overall differences. The positions of the framework termini differ with average pairwise differences of the terminal α-carbon positions of typically ~1 Å or more. Relative displacements of CDR loops roughly correlate with positional differences at the CDR-framework junctions. By comparing some Fvs, we noted concerted displacements of light and heavy chain CDR loops, suggesting a role of CDR-CDR interactions.

The analysis has implications for comparative modeling of an unknown antibody binding site (i.e., the Fv), which usually proceeds in two steps: (i) a framework is chosen from among the available x-ray structures via amino acid sequence alignments and (ii) desired CDR loops are identified and implanted onto the selected framework (7, 8). Given the many x-ray crystallographic structures of antibody fragments now available, the framework selection has generally been considered as a straightforward and almost trivial exercise. However, the data presented here suggest that the subtle positional differences in framework end points define an inherent limit to the accuracy of CDR loop modeling even if high-resolution structures are used as templates. Although the canonical loop motifs in general agree to within 0.5 Å, framework-dependent differences in loop take-off angles set the actual accuracy limit, i.e., the s_rms value, to at least 1.5 Å per loop. For example, the error would be 1.9–2.4 Å if the J539 H2 loop were implanted.

**Table III**

| CDR Loop Comparison | 22083 |
|---------------------|-------|
| a CDR H1 | b CDR H2 |
| KOL | d_rms | s_rms1(VL) | s_rms1(Fv) | KOL | d_rms | s_rms1(VL) | s_rms1(Fv) |
| HIL | d_rms | 0.23 | 0.67 | 1.15 | d_rms | 0.23 | 1.23 | 1.44 |
| s_rms1(VL) | 0.07 | 0.59 | 1.56 | s_rms1(VL) | 0.19 | 0.82 | 2.41 |
| s_rms1(Fv) | 1.15 | 1.43 | d_rms | 0.27 | 0.17 | 0.77 | 1.77 |
| s_rms1(VL) | 1.11 | 0.57 | s_rms1(VL) | 1.11 | 0.57 | 1.77 | 0.82 |
| s_rms1(Fv) | 1.56 | 1.43 | s_rms1(VL) | 1.56 | 1.43 | 2.41 | 1.53 |

**Fig. 2. Comparison of canonical CDR loops.** Backbone (N, Cα, C, O) comparisons of equivalent CDR loop structures are shown: a, H1; b, L2; c, H2. The top of each figure shows the direct least-squares fit of the loops (corresponding to d_rms values); the bottom shows comparison after S1 set framework superposition (corresponding to s_rms1 values). Color code: 4-4-20, magenta; D1.3, green; H52, yellow; H1l, red; J539, silver; KOL, blue S.E.155–4, gold.
onto the HIL framework after Fv superposition. Even if it were possible to accurately model VH-VL domain orientation angles, relative spatial displacement of CDR loops would not be controlled. The choice of framework regions and the definition of loop/framework junctions are important components of model-building protocols employing loop transfers and loop simulations. The exact position of the splice point or framework terminus will strongly influence the take-off angle of the loop and, in consequence, its spatial position.

The humanization of antibodies presents a similar problem. In this case, the rodent CDRs are experimentally grafted onto human framework regions (13). Despite conservation of murine structural determinants, moderate loss of antibody affinity is often observed (14). This may arise from changes in CDR loop position relative to the framework and to each other and may explain why the selection of a highly homologous human framework template is often critical (15).

In summary, our comparison of high precision x-ray structures confirms the similarity of equivalent canonical conformations and shows the presence of significant systematic differences in their spatial positions. The analysis implies that procedures for the selection of the best possible framework template in antibody modeling and humanization should be given much more attention than in the past. Future assessments of prediction accuracy should be expressed in, and discussed with respect to, the s_rms similarity measures in addition to the more commonly used, and better looking, d_rms values.

Acknowledgment—We thank Dr. Steven Sheriff for encouragement and helpful suggestions.

REFERENCES
1. Alzari, P. M., Lascombe, M.-B., and Poljak, R. J. (1988) Annu. Rev. Immunol. 6, 555–580
2. Padlan, E. A. (1994) Mol. Immunol. 31, 169–218
3. Kabat, E. A., Wu, T. T., Perry, H. M., Gottesman, K. S., and Foeller, C. (1991) Sequences of Proteins of Immunological Interest, 5th Ed., National Institutes of Health, Bethesda, MD
4. Chothia, C., and Lesk, A. M. (1987) J. Mol. Biol. 196, 901–917
5. Chothia, C., Lesk, A. M., Tramontano, A., Levitt, M., Smith-Gill, S., Air, G., Sheriff, S., Padlan, E. A., Davies, D., Tulip, W. R., Colman, P. M., Spinelli, S., Alzari, P. M., and Poljak, R. J. (1989) Nature 342, 877–883
6. Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. E., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., and Tasumi, M. (1977) J. Mol. Biol. 112, 535–542
7. Lesk, A. M., and Tramontano, A. (1991) in Antibody Engineering: A Practical Guide (Borrebaeck, W. H., ed) pp. 1–38, W. H. Freeman & Co, New York
8. Jones, T. A., and Thirup, S. (1986) EMBO J. 5, 819–822
9. Harris, L., and Bajorath, J. (1995) Protein Sci. 4, 306–310
10. Novotny, J., and Sharp, K. A. (1992) Prog. Biophys. Mol. Biol. 58, 203–224
11. Stanfield, R. L., Takimoto-Kamimura, M., Rini, J. M., Profy, A. T., and Wilson, F. A. (1993) Structure 1, 83–93
12. Wilson, I. A., and Stanfield, R. L. (1994) Curr. Opin. Struct. Biol. 4, 857–868
13. Jones, P. T., Dear, P. H., Foote, J., Neuberger, M. S., and Winter, G. (1986) Nature 321, 522–525
14. Carter, P., Presta, L., Gorman, C. M., Ridgeway, J. B. B., Henner, D., Wong, W. L. T., Rawlind, A. M., Kotses, C., Carver, M. E., and Shepard, H. M. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 4285–4289
15. Studnicka, G. M., Soares, S., Better, M., Williams, R. E., Nadell, R., and Horwitz, A. H. (1994) Protein Eng. 7, 805–814