Protein Science (1996), 5:167–169. Cambridge University Press. Printed in the USA.

FOR THE RECORD

Conformation of the hypervariable region L3 without the key proline residue

ALBA GUARNÉ,[1] JERÓNIMO BRAVO,[1] JAVIER CALVO,[2] FRANCISCO LOZANO,[2] JORDI VIVES,[2] AND IGNASI FITA[1]

[1] Departament d'Enginyeria Química, ETSEIB, UPC, Av. Diagonal, 647. 08028 Barcelona, Spain
[2] Servei d'Inmunologia, Hospital Clínic i Provincial de Barcelona, C/ Villarroel, 170. 08036 Barcelona, Spain

(RECEIVED August 22, 1995; ACCEPTED October 17, 1995)

Abstract: The refined structure of the Fab fragment of the monoclonal antibody CRIS-1 (IgG2aκ) against the leukocyte differentiation antigen CD5, determined at 1.9 Å resolution with an agreement *R*-factor of 18.3%, reveals a variant of the canonical conformations proposed for the light chain complementarity determining region L3 (CDR-L3). This is the first Fab structure available with a κ light chain in which the CDR-L3 lacks the key proline residue in either position 94 or 95. The conformation found could be significant for about 10% of the murine IgG molecules with κ light chains without proline in their CDR-L3 sequences.

Keywords: canonical forms; complementarity determining regions (CDRs); Fab fragment

The specificity of immunoglobulins is determined by the sequence and size of the complementary determining regions (CDRs) in the variable domains. Five of the six CDR loops display only a small number of conformations (Chothia & Lesk, 1987; Chothia et al., 1989), referred to as canonical forms, and defined, in every case, by a few highly conserved key residues (Kabat et al., 1977; Kabat, 1978), which can be in or in the vicinity of the hypervariable regions.

For CDR-L3, five canonical forms have been identified, two for λ light chains and three for κ light chains. All the conformations described for κ light chains (with adjacent framework residues 90 and 97 according to the Kabat nomenclature) contain a proline residue either in position 95 or, much less frequently, position 94. Pro 94 allows a two-residue hairpin turn (canonical form 2), whereas Pro 95 constrains residues 91–96 in an extended conformation, giving the six residues canonical form 1 or the five-residue loop found in the structure of HyHEL-5 and proposed as canonical form 3 (Chothia et al., 1989). For Fabs with a λ light chain, two types of L3 loops have been described, in both cases without proline in the sequence. One type has six amino acids and again forms a hairpin turn. The second type has eight amino acids, and residues 93, 93a, 93b, and 94 define a four-residue turn (Chothia & Lesk, 1987).

About 10% of the known murine Fab sequences having a κ light chain, do not have a proline residue in the CDR-L3 region (Kabat et al., 1991). This group includes chains with six-residue loops and chains with a different number of residues in the loop. To our knowledge, no Fab structures from this group has previously been determined, although a turn conformation was proposed for the six-residue loop without proline (Chothia & Lesk, 1987). We report here the three-dimensional structure, at 1.9 Å resolution, of a murine antibody Fab fragment with a κ light chain whose CDR-L3 does not contain proline and has only five residues. The observed CDR-L3 structure extends the repertoire of possible conformations for this loop.

Results and discussion: Owing to the quality of the electron-density maps, at 1.9 Å resolution most residues of the CRIS-1 Fab fragment have been accurately built (see for example Fig. 1A). The final model, which includes 240 well-defined solvent molecules, has a crystallographic agreement *R*-factor of 18.3% for 28,776 reflections in the resolution shell 8.0–1.9 Å.

In the CRIS-1 Fab structure, only four of the six CDRs (L1, L2, H1, and H2) have conformations that correspond to known canonical forms: L1 to canonical form 2, H2 to canonical form 3, and L2 and H1 to the unique canonical forms proposed for them. Instead, CDR-L3 does not fully correspond to any of the canonical forms described for this loop, because it has only five residues (Tyr 91-Tyr 92-Asn 93-Leu 94-Tyr 96; see Fig. 1A, Kinemage 1) and lacks the key proline. In spite of these differences, the extended conformation found for L3 in CRIS-1 can still be related to canonical forms 1 and 3 with root-mean-square (RMS) deviations (for C^{α} atoms) of 1.0 Å and 1.6 Å, respectively (Fig. 1B). The amide group of the adjacent framework residue Gln 90, the only L3 key residue present in CRIS-1, forms four hydrogen bonds: three with main-chain atoms from residues 92, 93, and 94, and one with O^{γ} from side chain of Thr 97

Reprint requests to: Ignasi Fita, Departament d'Enginyeria Química, E.T.S.E.I.B., U.P.C., Av. Diagonal, 647., 08028 Barcelona, Spain; e-mail: fita@eq.upc.es.

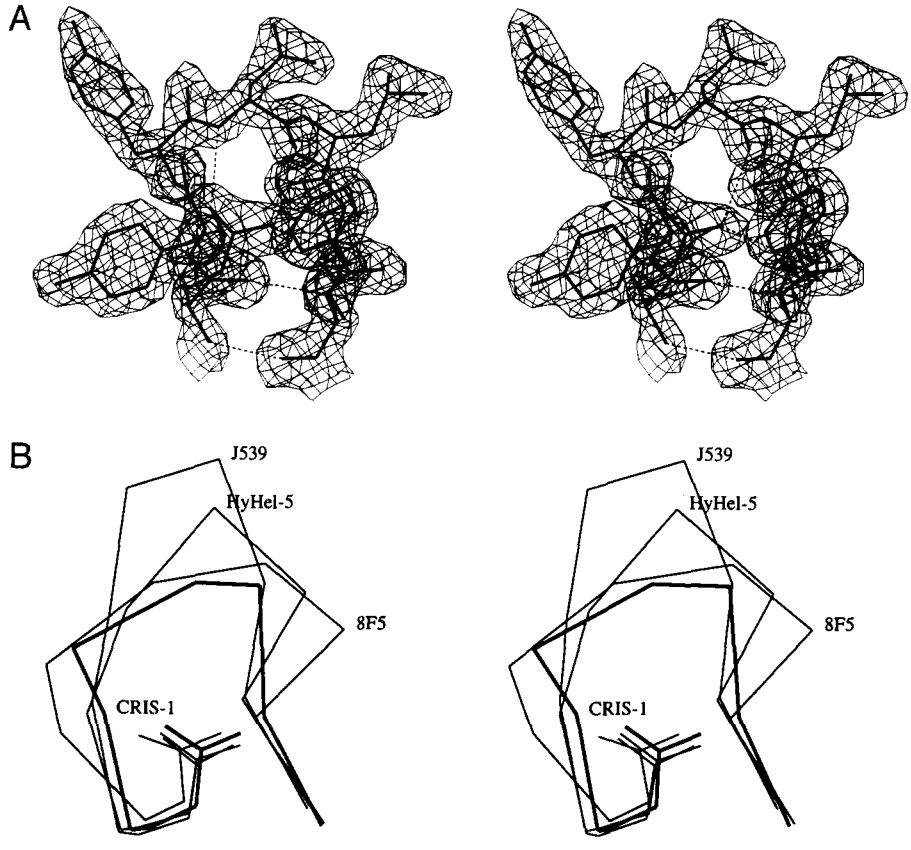

Fig. 1. **(A)** Stereo view of CDR-L3 ($2F_o - F_c$) electron density map. The corresponding molecular model, residues Gln 90-Tyr 91-Tyr 92-Asn 93-Leu 94-Tyr 96-Thr 97, is also shown. Intraloop hydrogen bonds are represented by dotted lines. **(B)** Stereo view showing the superimposition of the CDR-L3 alpha carbon tracing of CRIS-1 (thicker line) with the loops from 8F5 (Tormo et al., 1994), J539 (Suh et al., 1986), and HyHel-5 (Sheriff et al., 1987), which represent, respectively, canonical forms 1, 2, and 3 for this hypervariable region. As a reference, the side chain of residue 90 is also shown in some detail.

(Fig. 1A, Kinemage 1). The hydrogen bonds between the main chain of Gln 90 and Thr 97 initiate the β-sheet defined by strands $\beta 6$ and $\beta 7$ of the light chain variable domain. The L3 loop side chains of Tyr 91, Tyr 92, Asn 93, and Tyr 96 are mostly oriented toward the Fab binding pocket and have dispositions similar to the ones found in the canonical form 1. In the crystal, the polar atoms of these residues form a number of hydrogen bonds with solvent molecules and with residues from a symmetrically related molecule. In the structure of CRIS-1, Leu 94 has a left-handed main-chain conformation ($\Phi = 52$ and $\Psi = -136$), and its side chain fills a hydrophobic pocket defined by the heavy-chain framework residues Trp 47 and Pro 62 and the aliphatic part of the side chain of Lys 59 from CDR-H2. In about 80% of the κ light chain sequences, when the L3 hypervariable region has only five residues, a leucine is located in position 94 (Kabat, 1991).

The determinants of structure that stabilize the L3 loop in CRIS-1 differ from the ones described for the canonical forms of this hypervariable region. However, the conformation of L3 in CRIS-1 retains the structural features of the canonical form 1 that facilitate the proper orientation of the side chains and the packing of the CDR loops in the paratope. Thus, in CRIS-1, the presence of only five residues in the L3 loop prevents the formation of a hairpin turn, and the conformation and interactions of Leu 94 appear to compensate for the shorter sequence and the absence of the key residue proline.

Materials and methods: *Cloning and sequencing of heavy and light chain variable regions:* Cytoplasmic RNA from the CRIS-1 hybridoma was isolated as described previously (Lozano et al., 1994). The new procedure followed for cloning and sequencing the light chain (see supplementary material in the Electronic Appendix) has recently been published (Heinrichs et al., 1995). The product of the secondary PCR was cloned into the pUC119 *Sfi* I/*Not* I/His 6 vector and then subjected to direct double-stranded DNA sequencing. Sequencing reactions were performed using the Taq DyeDeoxi Terminator Cycle Sequencing Kit (Applied Biosystems, Inc., Foster City, California) with 500 ng of each dsDNA and 5 pmol of sequencing primers.

Preparation and crystallization of the Fab fragment: Fab fragments were obtained and purified by a standard protocol, and crystallization was performed by hanging-drop vapor diffusion (Guarné et al., 1995). The crystals were orthorhombic with space group $P2_12_12_1$ and unit cell parameters $a = 37.16$ Å, $b = 86.31$ Å, and $c = 134.01$ Å. One Fab fragment per asymmetric unit was present, and the estimated solvent content was 44% (Matthews, 1968). Diffraction data were collected using the synchrotron radiation facilities at the European Molecular Biology Laboratory outstation at Deutsches Elektronen-Synchronotron in Hamburg, to 1.9 Å resolution. Integration and reduction of intensities were carried out using the program DENZO, giving

a completeness of 94.1% and an overall internal R-factor of 8.9%.

Structure determination and refinement: The structure was determined by molecular replacement with the program AMoRe. The Fab fragment 8F5 (Tormo et al., 1994) was used as a starting model, and the Fv and constant module (C_L:C_H1) were positioned separately. Refinement and map calculation were done with the program XPLOR. Maps were examined using the graphic program TURBO. Omitted residues and side chains from the variable domain were introduced in the model when visible in ($2F_o - F_c$) or ($F_o - F_c$) maps during the refinement process.

Supplementary material in Electronic Appendix: Amino acid and nucleotide sequence of CRIS-1 light chain variable region, numbered according to Kabat et al. (1991) are available. The CDRs are underlined.

Acknowledgments: We thank Dr.Z.Otwinowski for his help with the program DENZO and A. Barceló for technical assistance with automatic DNA sequencing. This work has been supported by grants PB92-0707 and PM92-0190 (DGICYT) 95/1102(FLS), and GRQ93-9118 (CIRIT).

References

Chothia C, Lesk AM. 1987. Canonical structures for the hypervariable regions of immunoglobulins. *J Mol Biol 196*:901–917.

Chothia C, Lesk AM, Tramontano A, Levitt M, Smith-Gill SJ, Air G, Sheriff S, Padlan EA, Davies D, Tulip WR. 1989. Conformations of immunoglobulin hypervariable regions. *Nature 342*:877–883.

Guarné A, Bravo J, Calvo J, Lozano F, Vives J, Fita I. 1995. Three-dimensional structure of a monoclonal antibody Fab fragment against CD5. *An Quím 91*:296–299.

Heinrichs A, Milstein C, Gherardi E. 1995. Universal cloning and direct sequencing of rearranged antibody V genes using C region primers, biotin-captured cDNA and one-side PCR. *J Immunol 178*:241–251.

Kabat EA. 1978. The structural basis of antibody complementarity. *Adv Protein Chem 32*:1–75.

Kabat EA, Wu TT, Bilofsky H. 1977. Unusual distributions of amino acids in complementarity-determining (hypervariable) segments of heavy and light chains of immunoglobulins and their possible roles in specificity of antibody-combining sites. *J Biol Chem 252*:6609–6616.

Kabat EA, Wu TT, Perry HM, Gottesman KS, Foeller C. 1991. *Sequences of proteins of immunological interest, 5th ed.* Bethesda, Maryland: National Institutes of Health.

Lozano F, Maertzdorf B, Pannell R, Milstein C. 1994. Low cytoplasmic mRNA levels of immunoglobulin κ light chain genes containing nonsense codons correlate with inefficient splicing. *EMBO J 13*:4617–4622.

Matthews BW. 1968. Solvent content of protein crystals. *J Mol Biol 33*: 491–497.

Sheriff S, Silverton EW, Padlan EA, Cohen GH, Smith-Gill SJ, Funzel BC, Davies DR. 1987. Three-dimensional structure of an antibody–antigen complex. *Proc Natl Acad Sci USA 84*:8075–8079.

Suh SW, Bhat TN, Navia MA, Cohen GH, Rao DN, Rudikoff S, Davies DR. 1986. The galactan-binding immunoglobulin Fab J539: An x-ray diffraction study at 2.6 Å resolution. *Proteins Struct Funct Genet 1*:74–80.

Tormo J, Blaas D, Parry NR, Rowlands D, Stuart D, Fita I. 1994. Crystal structure of a human rhinovirus neutralizing antibody complexed with a peptide derived from viral capsid protein VP2. *EMBO J 13*:2247–2256.