J. Mol. Biol. (1993) 229, 597–601

Conformation of Complementarity Determining Region L1 Loop in Murine IgG λ Light Chain Extends the Repertoire of Canonical Forms

Shan Wu and Miroslaw Cygler†

Biotechnology Research Institute
National Research Council of Canada
Montreal, Quebec H4P 2R2, Canada

(Received 1 July 1992; accepted 13 October 1992)

The refined structure of Se155-4 Fab fragment, the first murine antibody with the λ light chain, reveals a novel conformation of the light chain complementarity determining region L1. This conformation extends the repertoire of canonical structures. The main determinant of this conformation is the packing of the Val27c side-chain into a hydrophobic pocket formed by the side-chains of Ala33, Leu66, Ala71 and Leu90. The framework L-FR3 loop, encompassing residues 66 to 72, which packs next to the L1 loop, bends significantly more toward the exterior of the molecule than in other Fab fragments. Sequence analysis suggests that the conformations of the L1 and L-FR3 loops observed in Se155-4 are adopted by a majority of murine λ light chains.

Keywords: Fab fragment; canonical forms; CDRs; crystal structure; *Salmonella*

Advances in genetic manipulation of antibodies achieved in recent years have made it possible and relatively easy to modify their antigenic binding sites and create smaller fragments, containing only the variable part of the antibody (Fv‡), which retains its full specificity. These advances opened the road to design and engineer antibodies with desired specificities. In order to do this in a rational way, there is a need to predict the structure of the antigen binding site from the amino acid sequence. This site is created mainly by residues from the complementarity determining regions (CDRs), L1, L2 and L3 from the light chain, and H1, H2 and H3 from the heavy chain (Wu & Kabat, 1970), forming loops extending from one end of the β-sheets of the conserved framework (FR) of the molecule (Davies & Metzger, 1983; Alzari *et al.*, 1988). The most important task in building a model is the determination of the conformation and position of CDRs (de la Paz *et al.*, 1986; Bruccoleri *et al.*, 1988).

Based on the structures of eight murine Fabs with κ light chains, human V_κ dimer, two human Fabs with λ light chains and human V_λ dimer, Chothia and colleagues (Chothia & Lesk, 1987; Chothia *et al.*, 1989) found that five out of six CDR loops display only a small number of conformations. They subsequently identified the key residues for each of the canonical forms. Five canonical forms were identified for L1 loop (1 for the λ chain and 4 for the κ chain, Table 1).

A significant part of the current understanding of the three-dimensional structure of an antigen binding site comes from the studies of Fab fragments of mouse monoclonal antibodies. Of the two types of light chains, λ chains are much less frequently found in mouse (less than 5% of the total light chain pool: McIntire & Rouse, 1970; Cotner & Eisen, 1978). It reflects the fact that there are only three genes for the V domain of λ chains ($V_{\lambda 1}$, $V_{\lambda 2}$ and $V_{\lambda x}$ (Eisen & Reilly, 1985; Sanchez *et al.*, 1990)). While the structure of both human and mouse κ chains is known, current knowledge of the λ chain conformation comes from human antibodies.

We have recently determined the first three-dimensional structure of a murine antibody Fab fragment containing a λ light chain (Cygler *et al.*, 1991). The heavy chain CDR loops, H1 and H2, and two of the light chain CDRs (L2 and L3) have been found in conformations corresponding to the known canonical forms. The L1 loop adopts, however, a

† Author to whom all correspondence should be addressed.

‡ Abbreviations used: Fv, variable fragment of an antibody; CDR, complementarity determining region; FR, conserved framework; Fab, antibody constant fragment.

0022–2836/93/030597–05 $08.00/0

Table 1

Sequences of the L1 loop typical for each of the canonical forms

Kabat No.	24	25	26	27	a	b	c	d	e	f	28	29	30	31	32	33	34	2	66	71	90
Chothia No.	24	25	26	27	28	29	30	a	b	c	d	e	f	31	32	33	34	2	66	71	90
k chains																					
		•				•										•		•		•	
J539(m)	S	A	S	S	S	V	S	-	-	-	-	-	-	-	S	L	H	I		Y	
		•				•										•		•		•	
D1.3(m)	R	A	S	G	N	I	H	N	-	-	-	-	-	-	Y	L	A	I		Y	
REI(h)	Q	A	S	Q	D	I	I	K	-	-	-	-	-	-	Y	L	N	I		Y	
		•				•										•		•		•	
McPC603(m)	K	S	S	Q	S	L	L	N	S	G	N	Q	K	N	F	L	A	I		F	
		•				•										•		•		•	
4-4-20(m)	R	S	S	Q	S	L	V	H	S	-	N	G	N	T	Y	L	R	V		F	
λ chains																					
		•					•									•				•	
NEWM(g)	T	G	S	S	S	N	I	-	-	-	G	A	G	N	H	V	K			A	
KOL(h)	S	G	T	S	S	N	I	-	-	-	-	L	S	S	T	V	N			A	
RHE(h)	T	G	S	A	T	D	I	-	-	-	-	G	S	N	S	V	I			A	
		•			•		•									•			•	•	•
SALC(m)	R	S	S	T	G	T	V	-	-	-	T	S	G	N	H	A	N		L	A	L

All *λ* chains with known 3-dimensional structures are listed. The key positions for maintaining the canonical structures are marked with * (after Chothia *et al.*, 1989). (h), human; (m), mouse. The sequences, and their alignment, were taken from Chothia *et al.* (1989) and the references are cited therein. The numbering scheme used by Kabat *et al.* (1991) differs from that used by Chothia *et al.* (1989). Both numbering schemes are given.

conformation quite different from those previously observed (Fig. 1). The structure of this anti-carbohydrate Fab fragment has been determined in two different crystal environments (Rose *et al.*, 1990), in a free form, and with three different haptens, to a resolution ranging from 2·05 to 2·6 Å (1 Å = 0·1 nm: Cygler *et al.*, 1991; and our unpublished results). It was found that the conformations of the CDR loops are only marginally affected by the intermolecular interactions and it was concluded that the observed conformation of the L1 loop is therefore inherent to this antibody and determined solely by the amino acid sequence. The numbering scheme defined by Kabat *et al.* (1991) is used in this work.

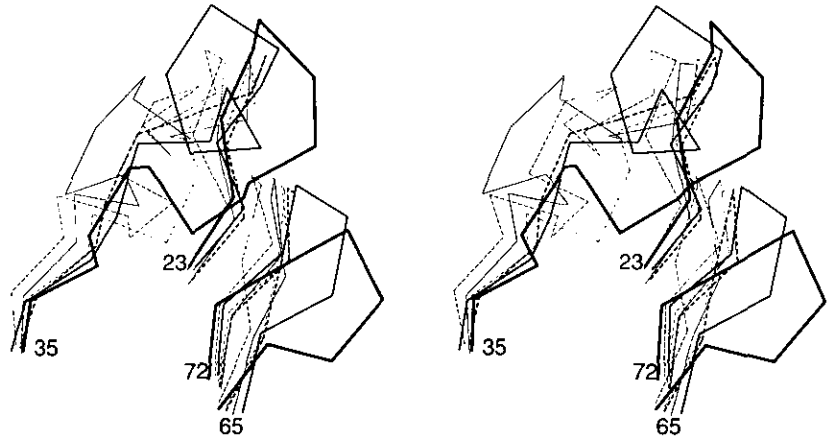

Figure 1. C^{α} tracing of L1 and L-FR3 loops representative of known canonical conformations. The superposition of the structures is based on the 2 *β*-strands connected by the L1 loop. The Se155-4 is shown as the thickest continuous line; NEWM (Saul *et al.*, 1978) as the thick continuous line; McPC603 (Satow *et al.*, 1986) as the thin continuous line; J539 (Suh *et al.*, 1986) as the heavy dotted line; D1.3 (Amit *et al.*, 1986) as the light dotted line; and 4-4-20 (Herron *et al.*, 1989) as the thin broken line.

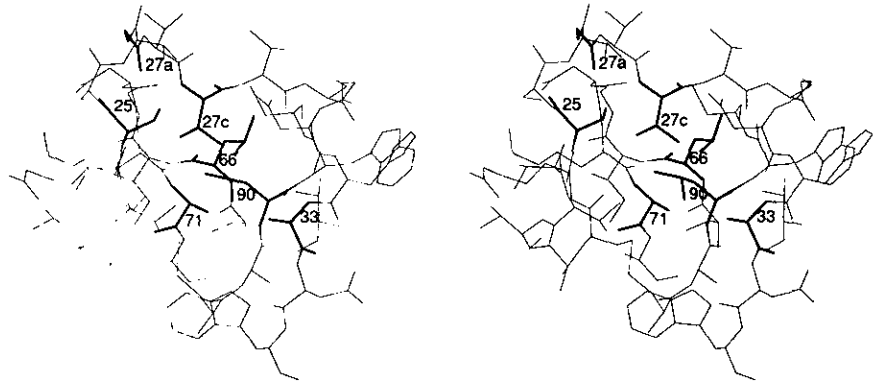

Figure 2. Environment of the L1 loop of the Se155-4 antibody. Essential residues are shown as thick lines.

The canonical forms that have been identified for the L1 region, residues 24 to 34, are different for κ and for (human) λ chains (Fig. 1). In mouse and human κ chains, the L1 loop has an extended form, while in the human λ chains, the L1 loop forms a type I β-turn from residue 26 to 27b, followed by an irregular helix. The hydrophobic pocket formed by the main chain of residue 25 and the side-chain of residue 33 and of the framework residue 71 (FR3) is filled by the side-chain of residue 27c. Burying the side-chain of Ile27c in a hydrophobic environment was suggested to be the driving force toward assuming this conformation. The key determinants of this conformation are residues at positions 25, 27c, 33 and 71 (Chothia & Lesk, 1987; see also Table 1) and in the three known human λ chain structures, they are Gly, Ile, Val and Ala, respectively.

The L1 loop of Se155-4 Fab has 14 residues. Residues 25 to 27a form a type I β-turn, stabilized by hydrogen bonds Ser25 $O^{\gamma}\ldots O^{\gamma 1}$ Thr27 and Ser25 NH ... Wat ... OC Gly27a. The three residues Thr27b, Val27c and Thr28 are in an extended conformation. The side-chains of Thr27b and Thr28 are exposed to the solvent, while the side-chain of Val27c is buried in a hydrophobic pocket formed by the side-chains of Ala33 and Leu90 (L3), and Leu66 and Ala71 (L-FR3, Fig. 2). Residues Thr28 to His32 forms a 3_{10} helical turn. The stability of this turn is enhanced by hydrogen bonds from a bridging water molecule to Ala33 NH and to OC Thr28.

The side-chain of Asn31 further stabilizes the L1 loop conformation by hydrogen bonding with Thr27 through a water molecule, with Thr28 and with Trp91 (L3). This asparagine residue is conserved in mouse λ chain sequences.

The L1 loop in Se155-4 has the same length as the corresponding loop of NEWM (Saul *et al.*, 1978). Comparison of their sequences (Table 1) shows that the character of the key position identified for NEWM is maintained in the Se155-4 (Ile27c → Val, Val33 → Ala, Ala71 → Ala), with the exception of Gly25, which is replaced by Ser in Se155-4. Moreover, the Gly30-Asn31-His32 sequence is common to both. In their analysis, Chothia & Lesk (1987) concluded that the murine V_{λ} L1 region is likely to be a distorted version of that found in the human structures. Why then, is the conformation of this loop very different? The answer to this question is provided by a close examination of the role of the side-chain in position 25. In NEWM, the residue is glycine. Its C^{α} atom is facing the interior of the protein and contributes to a hydrophobic pocket capped tightly by side-chain of Ile27c (Fig. 3). In Se155-4, the hydrophobic pocket is maintained in the same place and is lined up with some of the same residues as in NEWM; Ala33, Leu66, Ala71 and Leu90. However, residue 25 is serine instead of

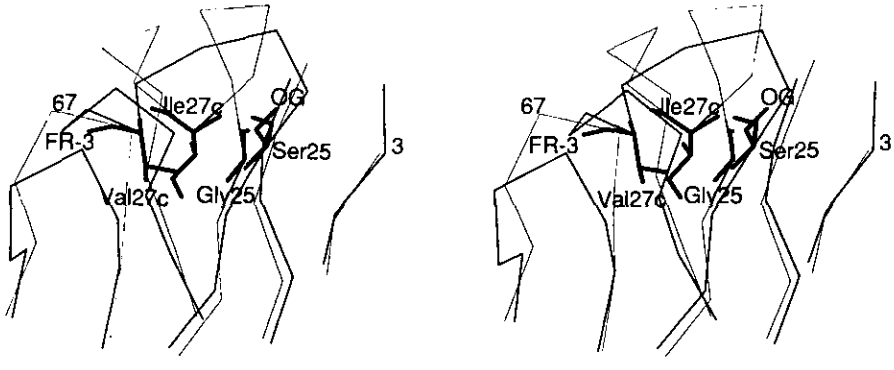

Figure 3. C^{α} tracing of the 2 λ chain L1 loops representing 2 canonical forms. Residues in positions 25 and 27c are shown in full. Note a similar position taken by Val27c and Ile27c in the 2 structures.

glycine, and its side-chain prefers a hydrophilic environment. Through some adjustment of the main-chain torsion angles in this region, the O^{γ} of Ser25 is repositioned to point outside. The direction of the main chain in Se155-4 departs here from that of NEWM. Yet the Val in position 27c still finds its way into hydrophobic pocket! Its final position is very close to that taken by Ile27c in NEWM (Fig. 3). This is in part possible due to a flexible Gly27a, which assumes a conformation unfavorable for other amino acids ($\phi, \psi = 130°, -150°$). In this sense it is the requirement of the hydrophobic side-chain at position 27c to fit into a hydrophobic environment that takes precedence over other factors in both molecules. Although replacement of Gly25 by Ser prevents the same folding as in NEWM, a different fold of this loop is found that, nevertheless, brings Val27c to the hydrophobic pocket.

From this analysis, the essential residues for maintaining the conformation of L1 loop in Se155-4 are those at positions 25 (Ser), 27a (Gly), 27c (Val), 33 (Ala), 66 (Leu), 71 (Ala) and 90 (Leu).

Two of the residues contributing to the hydrophobic cavity filled by Val27c come from the framework loop L-FR3 (Leu66 and Ala71). This loop always interacts with the L1 loop, covering it from the outside. The position of the L-FR3 loop relative to the rest of the protein (β-sheet scaffold) depends on the shape of the L1 loop. In the known structures of Fab fragments with κ light chains, the L1 loop extends along the direction of the pseudo 2-fold axis and the L-FR3 loop is approximately in the plane of the β-strands it links. In human Fab fragments with λ light chains, this loop is bent outside and has a more "open" position to give space for a more voluminous helical conformation of the L1 loop. In Se155-4, the L1 loop extends further outside and, as a result, the L-FR3 loop moves out even more (Fig. 1). This rather large global movement is achieved by small adjustments of the main chain torsion angles ($\sim$20 to 40°) of residues in this loop, with the largest changes occurring at positions 66 and 69.

There are only three genes, $V_{\lambda 1}$, $V_{\lambda 2}$ and $V_{\lambda x}$, encoding the variable region of λ chains (Eisen & Reilly, 1985; Sanchez *et al.*, 1987). The germline sequences for all of them are known (Kabat *et al.*, 1991). The amino acid sequences encoded by $V_{\lambda 1}$ and $V_{\lambda 2}$ genes are very similar (only 7 differences), while the amino acid sequence encoded by the $V_{\lambda x}$ gene is only 30 to 33% identical with those encoded by the $V_{\lambda 1}$, $V_{\lambda 2}$, and V_{κ} genes. The key positions responsible for the conformation of the L1 loop observed in Se155-4 are the same in all sequences encoded by both $V_{\lambda 1}$ and $V_{\lambda 2}$ genes (except for a Val27c → Ala substitution in one sequence). The L1 loop encoded by the $V_{\lambda x}$ gene is two residues shorter and has only one residue in common with L1 of Se155-4. We predict that murine λ chains encoded by $V_{\lambda 1}$ and $V_{\lambda 2}$ genes, will have L1 and L-FR3 loops in a conformation similar to those observed in Se155-4, while those encoded by the $V_{\lambda x}$ gene will have a different conformation. The conformation of the L1 loop observed in Se155-4 extends the repertoir of canonical forms.

This is NRCC publication number 33704.

References

Alzari, P. M., Lascombe, M.-B. & Poljak, R. J. (1988). Three-dimensional structure of antibodies. *Annu. Rev. Immunol.* **6**, 555–580.

Amit, A. G., Mariuzza, R. A., Philips, S. E. V. & Poljak, R. J. (1986). Three dimensional structure of an antigen-antibody complex at 2·8 Å resolution. *Science*, **233**, 747–753.

Bruccoleri, R. E., Haber, E. & Novotny, J. (1988). Structure of antibody hypervariable loops reproduced by a conformational search algorithm. *Nature (London)*, **335**, 564–568.

Chothia, C. & Lesk, A. M. (1987). Canonical structures for the hypervariable region of immunoglobulins. *J. Mol. Biol.* **196**, 901–917.

Chothia, C., Lesk, A. M., Tramontano, A., Levitt, M., Smith-Gill, S. J., Air, G., Sheriff, S., Padlan, E. A., Davies, D. R., Tulip, W. R., Colman, P. M., Spinelli, S., Alzari, P. M. & Poljak, R. L. (1989). Conformations of immunoglobulin hypervariable regions. *Nature (London)*, **342**, 877–883.

Cotner, T. & Eisen, H. N. (1978). The nature abundance of λ_2-light chains in inbred mice. *J. Expt. Med.* **148**, 1388–1399.

Cygler, M., Rose, D. R. & Bundle, D. R. (1991). Recognition of a cell-surface oligosaccharide of pathogenic *Salmonella* by an antibody Fab fragment. *Science*, **253**, 442–445.

Davies, D. R. & Metzger, H. A. (1983). Structural basis of antibody function. *Annu. Rev. Immunol.* **1**, 87–117.

de la Paz, P., Sutton, B. J., Darsley, M. J. & Rees, A. R. (1986). Modelling of the combining sites of the anti-lysozyme monoclonal antibodies and of the complex between one of the antibodies and its epitope *EMBO J.* **5**, 415–425.

Eisen, H. N. & Reilly, E. B. (1985). Lambda chains and genes in inbred mice. *Annu. Rev. Immunol.* **3**, 337–365.

Herron, J. N., He, X. M., Masson, M. L., Voss, E. W., Jr & Edmundson, A. B. (1989). Three dimensional structure of a fluorescein-Fab complex crystallized in MPD. *Protein Eng.* **5**, 271–280.

Kabat, E. A., Wu, T. T., Perry, H. M., Gottesman, K. S. & Foeller, C. (1991). *Sequences of Proteins of Immunological Interest*, 5th edit., NIH, Bethesda, MD.

McIntire, K. R. & Rouse, A. M. (1970). Immunoglobulin light chains: alteration of $\kappa:\lambda$ ratio. *Fed. Proc. Fed. Amer. Soc. Expt. Biol.* **29**, 704.

Rose, D. R., Cygler, M., To, R. J., Przybylska, M., Sinnott, B. & Bundle, D. R. (1990). Preliminary crystal structure analysis of an Fab specific for a *Salmonella* O-polysaccharide antigen. *J. Mol. Biol.* **215**, 489–492.

Sanchez, P., Marche, P. N., Le Guern, C. & Cazenave, P.-A. (1987). Structure of a third murine immunoglobulin λ light chain variable region that is expressed in laboratory mice. *Proc. Nat. Acad. Sci., U.S.A.* **84**, 9185–9188.

Sanchez, P., Marche, P. N., Ruffe-Juy, D. & Cazenave, P.-A. (1990). Mouse $V_{\lambda x}$ gene sequence generates no

junctional diversity and is conserved in mammalian species. *J. Immunol.* **144**, 2816–2820.

Satow, Y., Cohen, G. H., Padlan, E. A. & Davies, D. R. (1986). The phosphocholine binding immunoglobulin Fab McPC603: an X-ray diffraction study at 2·7 Å. *J. Mol. Biol.* **190**, 593–604.

Saul, F. A., Amzel, L. M. & Poljak, R. J. (1978). Preliminary refinement and structural analysis of the Fab fragment from human immunoglobulin NEW at 2·0 Å. *J. Mol. Biol.* **253**, 585–597.

Suh, S. W., Bhat, T. N., Navia, M. A., Cohen, G. H., Rao, D. N., Rudikoff, S. & Davies, D. R. (1986). The galactan-binding immunoglobulin Fab J539: an X-ray diffraction study at 2·6 Å resolution. *Protein Eng.* **1**, 74–80.

Wu, T. T. & Kabat, E. A. (1970). An analysis of the sequences of the variable regions of Bence–Jones proteins and myeloma light chains and their implications for antibody complementarity. *J. Expt. Med.* **132**, 211–249.

Edited by A. R. Fersht