J. Mol. Biol. (1990) **215**, 175–182

Framework Residue 71 is a Major Determinant of the Position and Conformation of the Second Hypervariable Region in the V_H Domains of Immunoglobulins

Anna Tramontano[1], Cyrus Chothia[2,3] and Arthur M. Lesk[1,2]

[1]*European Molecular Biology Laboratory*
Meyerhofstrasse 1
Postfach 1022.09
6900 Heidelberg, F.R.G.

[2]*MRC Laboratory of Molecular Biology*
Cambridge, CB2 2QH, U.K.

[3]*Christopher Ingold Laboratory*
University College London
20 Gordon Street
London WC1H 0AJ, U.K.

(Received 2 January 1990; accepted 18 May 1990)

Analysis of the immunoglobulins of known structure reveals systematic differences in the position and main-chain conformation of the second hypervariable region of the V_H domain (H2). We show that the major determinant of the position of H2 is the size of the residue at site 71, a site that is in the conserved framework of the V_H domain. It is likely that for about two thirds of the known V_H sequences the size of the residue at this site is also a major determinant of the conformation of H2. This effect can override the predisposition of the sequence, as in the case of the H2 loop of J539, which is an exception to the rules relating sequence and conformation of short hairpin loops. Understanding the relationship between the residue at position 71 and the position and conformation of H2 has applications to the prediction and engineering of antigen-binding sites of immunoglobulins.

1. Introduction

Immunoglobulins are multi-domain proteins consisting of two chains, a light chain with one variable (V_L†) and one constant domain, and a heavy chain containing one variable domain (V_H) and three constant domains. The antigen-binding site is formed by six loops, three from the V_L and three from the V_H domains. Figure 1 shows a simplified view of the antigen-binding site, indicating the relative positions of the loops. The variability of the residues in the antigen-binding site gives rise to the high range of specificity achieved by antibodies (Wu & Kabat, 1970; Kabat *et al.*, 1987).

The atomic structures of several immunoglobulin fragments have been determined by X-ray crystallography (Davies & Metzger, 1983; Alzari *et al.*, 1988). They show that all the domains have a very similar folding pattern: two β-sheets packed face to face. A core of the double β-sheet structure, called the framework, has a very similar conformation in different variable domains because of the conservation of internal residues and the requirements of internal packing. The residues that form the interface between the V_L and V_H domains are also strongly conserved.

These results have led to the view that the framework structure plays an essentially passive role in the structural variation that occurs in the antigen-

† Abbreviations used: V_H, variable domain of immunoglobulin heavy chains; V_L, variable domain of immunoglobulin light chains; H2, second hypervariable loop of heavy chain; r.m.s., root-mean-square; α_R, right-handed α-helical conformation; α_L, left-handed α-helical

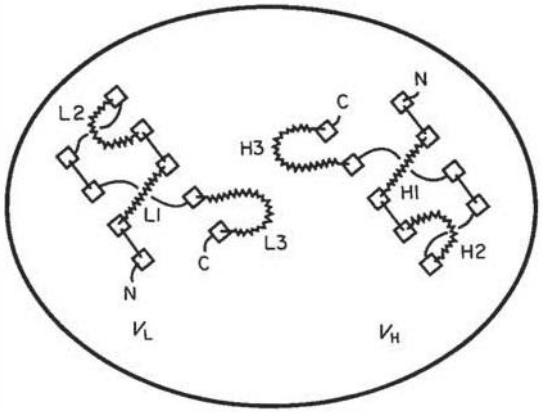

Figure 1. Outline structure of the antigen-binding site. The site is formed by 6 loops of polypeptide (∿) linked to strands in β-sheets (□).

In the immunoglobulins of known structure the conformations of the second hypervariable region in V_H (H2) differ. The position of the H2 with respect to the conserved framework is also variable. For example, in the V_H domains of immunoglobulins J539 and HyHEL-5, the H2 regions have the same number of residues. If the framework structures are superimposed, the C^β atoms in residue 53, at the tip of H2, are found to differ in position by 6·3 Å (1 Å = 0·1 nm). Here, we show that the variations in two structural features of H2, its position and its conformation, are coupled, and that they depend in large part on the nature of the amino acid residue that occupies position 71 in the heavy-chain framework.

Figure 2 shows the general structural context of H2 within the V_H domain.

2. Co-ordinates and Calculations

Protein structures used in this work are listed in Table 1. The atomic co-ordinates of these structures are distributed by the Protein Data Bank (Bernstein *et al.*, 1977), except for the refined co-ordinates of J539 which are a private communication from Drs E. A. Padlan and D. R. Davies. The structures were displayed using Insight (Dayringer *et al.*, 1986) on an Evans & Sutherland PS390. Programs written by A.M.L. (Lesk, 1986) were used for analysis of the structures and database searching.

Throughout the paper residue numbers refer to the heavy-chain numbering scheme of Kabat *et al.* (1987). In V_H domains, the conserved β-sheet framework consists of residues 3 to 12, 17 to 25, 33 to 52, 56 to 60, 68 to 82, 88 to 95 and 102 to 112 (Chothia & Lesk, 1987). These residues were used in the superpositions of V_H domains.

3. The Conformations of H2 Loops

In V_H sequences the second hypervariable region consists of a β-hairpin, comprising residues 50 through 65 (Wu & Kabat, 1970; Kabat *et al.*, 1987). In the known V_H structures the main-chain conformations of residues 50 to 52 and 56 to 65 are the same: for high-resolution well-refined structures the backbone atoms of residues 50 to 52 and 56 to 64 fit with a root-mean-square (r.m.s.) deviation between 0·4 and 0·7 Å. This region is illustrated in Figure 3:

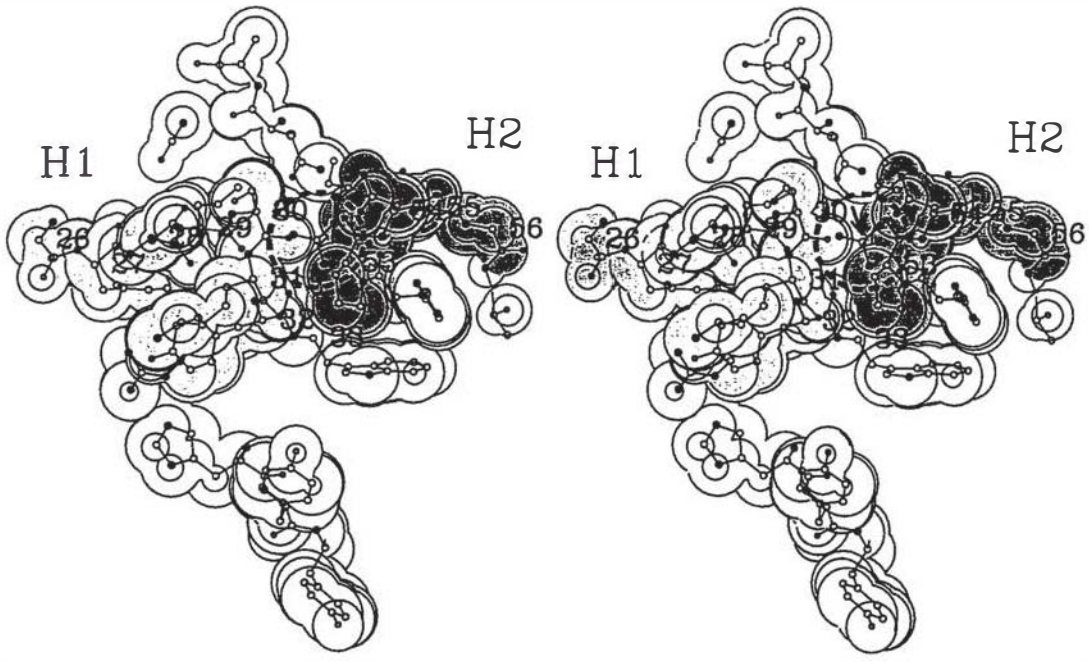

Figure 2. The structural context of H2 within the V_H domain of Fab J539. H2 is shaded relatively darkly, H1 is

Table 1

Immunoglobulin heavy chain variable domains of known atomic structure

Molecule	H2 sequence	Residue 71	Reference
NEWM	Y H G	V	Saul *et al.* (1978)
HyHEL-10	Y S G	R	Padlan *et al.* (1989)
HyHEL-5	P G S G	A	Sheriff *et al.* (1987)
KOL	D D G S	R	Marquart *et al.* (1980)
J539	P D S G	R	Suh *et al.* (1986)
McPC603	N K G N K Y	R	Satow *et al.* (1986)
4-4-20	N K P Y N Y	R	Herron *et al.* (1989)

The H2 residues are those between positions 52 and 56 (see text).

the main-chain atoms of residues 56 to 60 form hydrogen bonds to those of residues 48 to 52 to form a β-hairpin. Sequence variations in these residues have little or no effect on the main-chain conformation, because the side-chains are on the surface. The turn that links these two strands, comprising residues 52a to 55 or 53 to 55, we refer to as the H2 region. In the known structures it differs in length and conformation.

Hairpin structures have been classified according to their length and conformation (Venkatachalam, 1968; Efimov, 1986; Sibanda & Thornton, 1985; Sibanda *et al.*, 1989). Particular conformations are usually associated with characteristic sequence patterns. The positions of Gly, Asn, Asp and Pro residues are important because these residues allow main-chain conformations that in other residues cause strain.

(a) *Three-residue H2 regions*

In NEWM and HyHEL-10, the H2 loop is a three-residue hairpin, residues 53 to 55. The NEWM H2 loop is shown as conformation 1 in Figure 3. The usual sequence requirement for this conformation is a Gly (or Asn or Asp) at the third position (residue 55), which can take up a + + conformation (that is, $\phi>0, \psi>0$) (Sibanda *et al.*, 1989). Both NEWM and HyHEL-10 have a glycine at this position:

	53	54	55	71
NEWM	Tyr	His	Gly	Val
HyHEL-10	Tyr	Ser	Gly	Arg

and in both cases the Gly is in a + + conformation.

(b) *Four-residue H2 regions*

The H2 loop of HyHEL-5 is a four-residue hairpin, residues 52a to 55. This is shown as conformation 2 in Figure 3. The conformation is close to the one most commonly observed in four-residue turns, in which the first three residues are in an α_R conformation and the fourth in an α_L conformation. These turns normally require Gly in the fourth position (Efimov, 1986; Sibanda & Thornton, 1985; Sibanda *et al.*, 1989), as observed in HyHEL-5.

The H2 regions in KOL and J539 form four-residue turns with a conformation different from HyHEL-5. They both have the third residue (54) in the α_L conformation and the first, second and fourth in the α_R conformation. This is shown as conformation 3 in Figure 3.

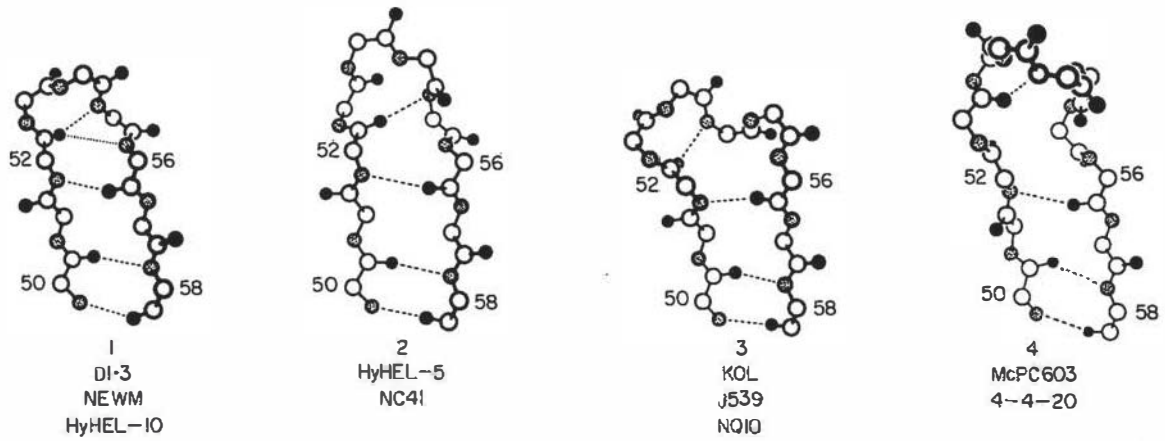

Figure 3. The main-chain conformations of the 2nd hypervariable region in V_H domains in the immunoglobulins of

Table 2

Results of a database search for main-chain conformations the same as that of the H2 loop of KOL

Δ (Å)	Molecule (Protein Data Bank code)	Starting residue	Sequence			
0·18	Rhizopuspepsin (3APR)	145	S	Q	G	L
0·19	Subtilisin Carlsberg (2SEC)	10	L	I	K	A
0·22	Ribonuclease A (7RSA)	32	S	R	N	L
0·22	Pepsinogen (1PSG)	142	D	Q	G	L
0·22	434 repressor protein (1R69)	35	E	N	G	K
0·23	Calmodulin (3CLN)	57	A	D	G	N
0·24	Calmodulin (3CLN)	21	K	D	G	D
0·28	Adenylate kinase (3ADK)	166	K	R	G	I
0·29	Fab J539	353	P	D	S	G
0·29	Cytochrome c551 (451C)	9	N	K	G	C

Δ, root-mean-square deviation of N, C^{α}, C and ● atoms of residues 53 to 56 of the V_H domain of KOL and well-fitting regions from other known structures.

This type of turn has not been described previously, but we find that it occurs fairly often in proteins. We searched the database of solved structures for regions similar in main-chain conformation to the H2 loop of KOL. Table 2 lists the closest matches: ten loops, including J539 H2, for which the r.m.s. difference in position of main-chain atoms is less than 0·3 Å. There are 61 such loops with r.m.s. deviation less than 0·5 Å. For KOL and J539 H2 and the nine best-fitting non-homologous loops, the average values of the conformational angles and their standard deviations are:

Angle	ϕ_1	ψ_1	ϕ_2	ψ_2	ϕ_3	ψ_3	ϕ_4	ψ_4
Mean (deg.)	−61	−35	−95	77	65	22	−78	−18
Standard deviation (deg.)	12	8	12	14	11	11	13	12

Of the nine loops in Table 2, excluding J539, seven have a Gly in the third position, like KOL, one has Asn and one has Lys. Of all the loops with r.m.s. deviation less than 0·5 Å, none is like J539 in having Gly at only the fourth position.

These results show that H2 in J539 is an exception to the rules relating sequence and structure in short hairpins. Both HyHEL-5 and J539 have Gly in the fourth position of the loop:

	52a	53	54	55
KOL	Asp	Asp	Gly	Ser
J539	Pro	Asp	Ser	Gly
HyHEL-5	Pro	Gly	Ser	Gly

The position of Gly in J539 should imply a conformation of H2 similar to that of HyHEL-5. Instead the conformation observed in J539 is the same as in KOL (see Fig. 4(b) and Fig. 5(a)). The r.m.s. deviation in the position of the H2 main-chain atoms in J539 and HyHEL-5 is 1·9 Å; for J539 and KOL it is 0·3 Å. The residues of H2 in J539 make no non-bonded contacts to residues other than those in H1 and Arg71 and Asn73 (see Fig. 2).

(c) *Six-residue H2 regions*

In McPC603 and 4-4-20, the H2 loops are six-residue (see Fig. 5(b)). The r.m.s. deviation of all N, C^{α}, C and O atoms is 0·96 Å. The McPC603 H2 loop is shown as conformation 4 in Figure 3. The sequences in these regions are:

	52a	52b	52c	53	54	55	71
McPC603	Asn	Lys	Gly	Asn	Lys	Tyr	Arg
4-4-20	Asn	Lys	Pro	Tyr	Asn	Tyr	Arg

In both structures residue 54 is in the α_L conformation. In the other V_H sequences with six-residue H2 loops, the residues found at this position are Gly, Asn or Asp (Kabat *et al.*, 1987). It is interesting to note in this context that the Lys at position 54 in McPC603 is the result of a somatic mutation from a germ-line gene that contains a Gly.

The Interactions of H2 with the Framework

Examination of the interactions of the H2 loops with the rest of the V_H domain shows that the determinants of the conformations of four-residue H2 loops are not entirely within the sequence of the loop itself, but involve the packing of the loop against the rest of the structure.

In Figures 4 and 5 we show, for pairs of antibodies of known structure, the relative positions of the H1 and H2 loops and the contacts made by certain side-chains. The relative positions of these loops in these Figures are those induced by the superposition of the framework structures. The Figures show that the H1 loops occupy rather similar positions with respect to the framework in all the known structures. But the positions of the H2 loops are in some cases very different. These differences are related to the size of the residue at position 71.

KOL and J539 have four-residue H2 loops in very similar positions and conformations (Fig. 4(b)). The

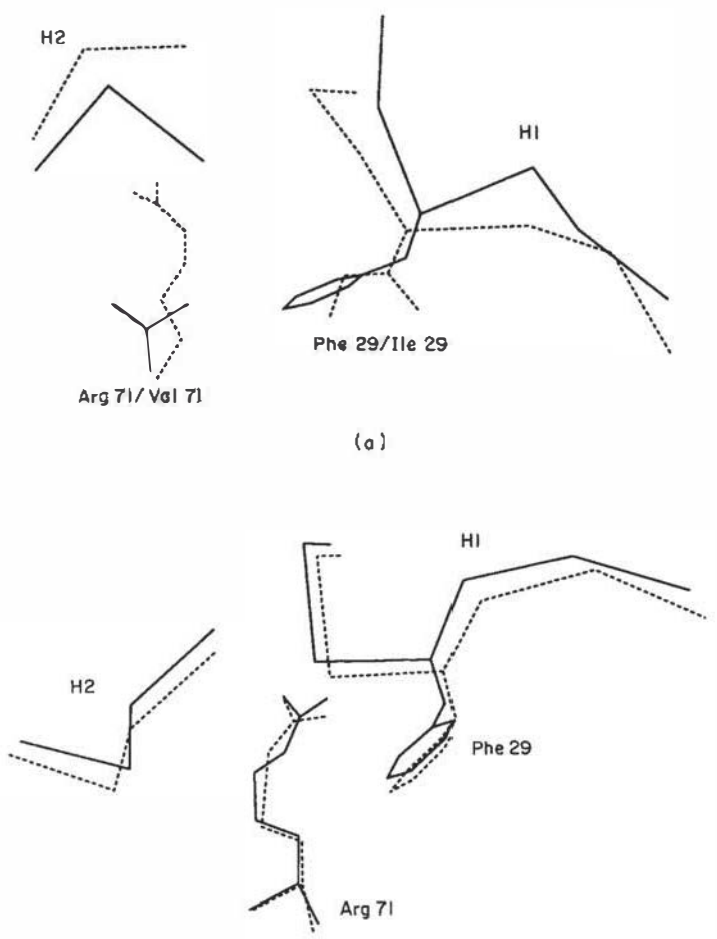

Figure 4. The relative positions of the H1 and H2 hypervariable regions and of framework residue 71, in different pairs of immunoglobulins. The H1 and H2 regions are represented by their C^{α} atoms. The positions shown here are those found after the superposition of the V_H framework residues (see text). (a) NEWM (continuous lines) and HyHEL-10 (broken lines). (b) KOL (continuous lines) and J539 (broken lines).

side-chains of these arginine residues are buried. They form hydrogen bonds to main-chain atoms of residues in the H1 and H2 loops and pack against the Phe at position 29.

The superposition of J539 and HyHEL-5 shown in Figure 5(a) illustrates the case of two immunoglobulin structures with H2 loops of the same length but different conformation and position. In J539, in which residue 71 is an Arg, residue Pro52a in the H2 loop is on the surface. In HyHEL-5, in which residue 71 is an Ala, Pro52a is buried, filling the cavity that would be created by the absence of a long side-chain at position 71. The manner in which these H2 loops pack against the rest of the V_H domain explains why the H2 region of J539 does not have the conformation that we would expect from Gly at position 56. If it did have the expected conformation, like that in HyHEL-5, the Pro52a side-chain would occupy the same space as the side-chains that move the side-chain of Pro52a away from Arg71 require an H2 conformation different from that in HyHEL-5.

In both McPC603 and 4-4-20, H2 is a six-residue turn, and residue 71 is an Arg. In McPC603 Arg71 has its side-chain buried, and is hydrogen bonded to the main-chain of H1 and H2, as in KOL and J539 (Fig. 5(b)). The Tyr at the sixth position (55) packs against Arg71.

5. The Role of Residue 71

These observations can be summarized as follows.

(1) Position 71 contains a small or medium-sized residue. For three and four-residue H2 loops the residue at position 53/52a packs against residues at positions 71 and 29. This brings the H1 and H2 loops close together and puts four-residue H2 loops in conformation 2 (Fig. 3).

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.