WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jun 18, 2024 – 01:15 PM EDT

PDB ID	:	4FQV
Title	:	Crystal structure of broadly neutralizing antibody CR9114 bound to H7 influenza hemagglutinin
Authors	:	Ekiert, D.C.; Dreyfus, C.; Wilson, I.A.
Deposited on	:	2012-06-25
Resolution	:	5.75 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

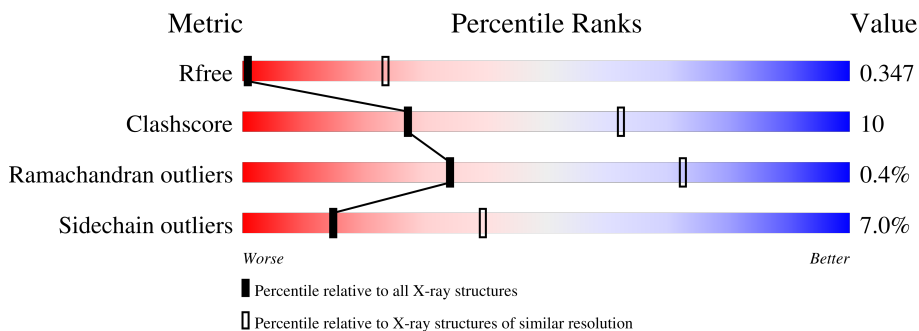|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 5.75 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.

Metric	Percentile Ranks	Value
Rfree		0.347
Clashscore		10
Ramachandran outliers		0.4%
Sidechain outliers		7.0%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1007 (7.66-3.86)
Clashscore	141614	1033 (7.60-3.90)
Ramachandran outliers	138981	1002 (7.60-3.86)
Sidechain outliers	138945	1005 (7.70-3.82)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5%

Mol	Chain	Length	Quality of chain
1	A	327	75% 18% • • •
1	C	327	73% 20% • • •
1	E	327	79% 15% • •
2	B	176	68% 23% 6% •
2	D	176	66% 27% 5% •
2	F	176	69% 23% 5% •
3	H	224	80% 15% • •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	I	224	86% 10% •
3	J	224	87% 8% •
4	L	216	77% 19% ••
4	M	216	85% 11% ••
4	N	216	81% 16% ••

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 20938 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Hemagglutinin HA1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	316	Total 2423	C 1503	N 435	O 471	S 14	0	0	0
1	C	316	Total 2423	C 1503	N 435	O 471	S 14	0	0	0
1	E	317	Total 2427	C 1505	N 436	O 472	S 14	0	0	0

There are 15 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	7	ALA	-	EXPRESSION TAG	UNP Q6VMK1
A	8	ASP	-	EXPRESSION TAG	UNP Q6VMK1
A	9	PRO	-	EXPRESSION TAG	UNP Q6VMK1
A	10	GLY	-	EXPRESSION TAG	UNP Q6VMK1
A	254	LEU	PRO	CONFLICT	UNP Q6VMK1
C	7	ALA	-	EXPRESSION TAG	UNP Q6VMK1
C	8	ASP	-	EXPRESSION TAG	UNP Q6VMK1
C	9	PRO	-	EXPRESSION TAG	UNP Q6VMK1
C	10	GLY	-	EXPRESSION TAG	UNP Q6VMK1
C	254	LEU	PRO	CONFLICT	UNP Q6VMK1
E	7	ALA	-	EXPRESSION TAG	UNP Q6VMK1
E	8	ASP	-	EXPRESSION TAG	UNP Q6VMK1
E	9	PRO	-	EXPRESSION TAG	UNP Q6VMK1
E	10	GLY	-	EXPRESSION TAG	UNP Q6VMK1
E	254	LEU	PRO	CONFLICT	UNP Q6VMK1

- Molecule 2 is a protein called Hemagglutinin HA2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	170	Total 1380	C 851	N 243	O 279	S 7	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	171	Total 1388	C 857	N 244	O 280	S 7	0	0	0
2	F	169	Total 1369	C 845	N 239	O 278	S 7	0	0	0

- Molecule 3 is a protein called Antibody CR9114 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	H	215	Total 1608	C 1013	N 268	O 320	S 7	0	2	0
3	I	215	Total 1608	C 1013	N 268	O 320	S 7	0	2	0
3	J	215	Total 1608	C 1013	N 268	O 320	S 7	0	2	0

- Molecule 4 is a protein called Antibody CR9114 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	L	211	Total 1568	C 978	N 266	O 320	S 4	0	0	0
4	M	211	Total 1568	C 978	N 266	O 320	S 4	0	0	0
4	N	211	Total 1568	C 978	N 266	O 320	S 4	0	0	0

wwPDB X-ray Structure Validation Summary Report
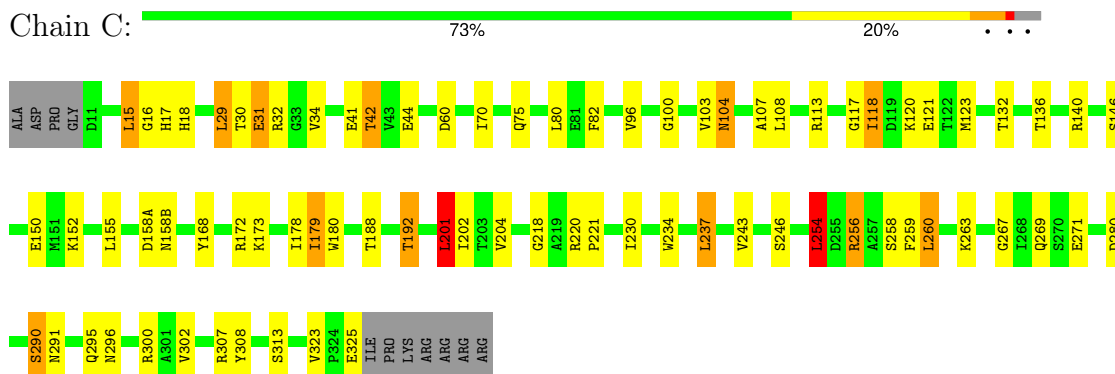
3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.
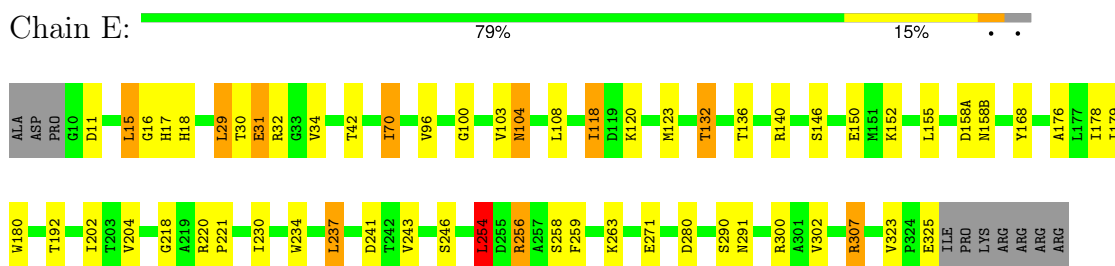
• Molecule 1: Hemagglutinin HA1

Chain A:

• Molecule 1: Hemagglutinin HA1

Chain C:

• Molecule 1: Hemagglutinin HA1

Chain E:

• Molecule 2: Hemagglutinin HA2

Chain B:

• Molecule 2: Hemagglutinin HA2

Chain D:

• Molecule 2: Hemagglutinin HA2

Chain F:

• Molecule 3: Antibody CR9114 heavy chain

Chain H:

• Molecule 3: Antibody CR9114 heavy chain

Chain I:

• Molecule 3: Antibody CR9114 heavy chain

Chain J:

Q1 W36 T68 F74 E81 C92 N97 Y98 Y99 Y100 Y100A P126 SER SER LYS SER THR SER G133 L138 G157 L178 V184 P185 S186 L189 N197 K206 V207 D208 P213 LYS SER CYS

● Molecule 4: Antibody CR9114 light chain

Chain L:

77%   19%

GLN S2 V11 S12 G13 T14 P15 I21 W35 A43 Y49 S50 N51 D52 Q53 R54 V58 P59 D60 S67 S80 E83 C88 A96 V106 L106A G107 Q108 P109 K110 F118 P119 P120 L132 F139 Y140 P141 V144 T145 K149 S152 S153

K156 K166 Q167 A173 W185 R189 S192 T196 E203 K204 P208 THR GLU CYS SER

● Molecule 4: Antibody CR9114 light chain

Chain M:

85%   11%

GLN S2 R30 S50 N51 R54 V58 R61 S67 S76 D93 Q108 L132 T145 K149 S152 K156 L178 T181 P182 E183 Q184 W185 K186 S187 H188 R189 S192 T196 E203 K204 P208 THR GLU CYS SER

● Molecule 4: Antibody CR9114 light chain

Chain N:

81%   16%

GLN S2 S12 G13 D27 R30 S50 N51 R54 V58 R61 F62 S67 I75 S76 S80 E83 L104 T105 V106 L106A G107 Q108 E123 E124 K129 L132 V144 T145 W148 K149 S152 K156 L178 S187 H188 R189 S192

T196 H197 E203 K204 P208 THR GLU CYS SER

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 42 21 2	Depositor
Cell constants a, b, c, α, β, γ	197.57Å 197.57Å 223.62Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.35 – 5.75 49.35 – 5.75	Depositor EDS
% Data completeness (in resolution range)	99.4 (49.35-5.75) 99.0 (49.35-5.75)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.14	Depositor
$< I/\sigma(I) >$ [1]	1.35 (at 5.73Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8_1069)	Depositor
R, R_{free}	0.314 , 0.339 0.323 , 0.347	Depositor DCC
R_{free} test set	642 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å^2)	319.3	Xtriage
Anisotropy	0.293	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 229.2	EDS
L-test for twinning[2]	$< |L| > = 0.44, < L^2 > = 0.26$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.81	EDS
Total number of atoms	20938	wwPDB-VP
Average B, all atoms (Å^2)	244.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.52% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#|Z| >5	RMSZ	#|Z| >5
1	A	0.32	0/2468	0.61	1/3336 (0.0%)
1	C	0.32	0/2468	0.63	2/3336 (0.1%)
1	E	0.32	0/2472	0.61	1/3341 (0.0%)
2	B	0.32	0/1404	0.53	0/1892
2	D	0.36	0/1412	0.56	0/1903
2	F	0.35	0/1393	0.54	0/1878
3	H	0.36	0/1647	0.75	1/2245 (0.0%)
3	I	0.33	0/1647	0.57	0/2245
3	J	0.37	0/1647	0.61	0/2245
4	L	0.33	0/1606	0.59	0/2193
4	M	0.38	0/1606	0.56	0/2193
4	N	0.31	0/1606	0.55	0/2193
All	All	0.34	0/21376	0.60	5/29000 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	C	0	1
1	E	0	1
2	B	0	1
2	D	0	1
2	F	0	1
4	M	0	1
All	All	0	7

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	254	LEU	CB-CG-CD1	-8.52	96.51	111.00
1	A	254	LEU	CB-CG-CD1	-7.97	97.45	111.00
1	C	254	LEU	CB-CG-CD1	-7.07	98.98	111.00
3	H	92	CYS	N-CA-C	-6.61	93.14	111.00
1	C	201	LEU	CA-CB-CG	5.40	127.72	115.30

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	29	LEU	Peptide
2	B	56	ILE	Peptide
1	C	29	LEU	Peptide
2	D	56	ILE	Peptide
1	E	29	LEU	Peptide

5.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2423	0	2373	61	0
1	C	2423	0	2373	57	0
1	E	2427	0	2376	50	0
2	B	1380	0	1284	48	0
2	D	1388	0	1295	54	0
2	F	1369	0	1271	44	0
3	H	1608	0	1556	59	0
3	I	1608	0	1556	34	0
3	J	1608	0	1556	36	2
4	L	1568	0	1517	37	2
4	M	1568	0	1517	14	1
4	N	1568	0	1517	29	0
All	All	20938	0	20191	417	3

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 10.

The worst 5 of 417 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:19:ASP:OD2	3:H:100(A):TYR:OH	1.65	1.11
1:C:123:MET:SD	1:C:254:LEU:HD13	2.00	1.01
3:H:11:VAL:CG2	3:H:147:PRO:HG3	1.94	0.97
1:E:123:MET:SD	1:E:254:LEU:CD1	2.55	0.95
2:F:38:TYR:HE1	3:I:97:ASN:HB2	1.30	0.94

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:L:189:ARG:NH2	3:J:157:GLY:O[6_554]	1.99	0.21
4:L:189:ARG:NE	3:J:157:GLY:O[6_554]	2.10	0.10
4:M:187:SER:OG	4:M:187:SER:OG[2_465]	2.14	0.06

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	314/327 (96%)	295 (94%)	19 (6%)	0	100	100
1	C	314/327 (96%)	297 (95%)	17 (5%)	0	100	100
1	E	315/327 (96%)	296 (94%)	19 (6%)	0	100	100
2	B	168/176 (96%)	151 (90%)	14 (8%)	3 (2%)	8	40
2	D	169/176 (96%)	152 (90%)	14 (8%)	3 (2%)	8	40
2	F	167/176 (95%)	151 (90%)	14 (8%)	2 (1%)	13	50
3	H	213/224 (95%)	210 (99%)	3 (1%)	0	100	100
3	I	213/224 (95%)	212 (100%)	1 (0%)	0	100	100
3	J	213/224 (95%)	212 (100%)	1 (0%)	0	100	100
4	L	209/216 (97%)	204 (98%)	4 (2%)	1 (0%)	29	69
4	M	209/216 (97%)	206 (99%)	2 (1%)	1 (0%)	29	69

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	N	209/216 (97%)	205 (98%)	3 (1%)	1 (0%)	29	69
All	All	2713/2829 (96%)	2591 (96%)	111 (4%)	11 (0%)	34	72

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	57	GLU
2	D	57	GLU
2	F	57	GLU
2	B	119	TYR
2	D	119	TYR

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	268/277 (97%)	236 (88%)	32 (12%)	5	20
1	C	268/277 (97%)	236 (88%)	32 (12%)	5	20
1	E	268/277 (97%)	239 (89%)	29 (11%)	6	23
2	B	145/151 (96%)	129 (89%)	16 (11%)	6	22
2	D	146/151 (97%)	131 (90%)	15 (10%)	7	25
2	F	144/151 (95%)	128 (89%)	16 (11%)	6	22
3	H	180/187 (96%)	177 (98%)	3 (2%)	60	78
3	I	180/187 (96%)	178 (99%)	2 (1%)	73	84
3	J	180/187 (96%)	178 (99%)	2 (1%)	73	84
4	L	175/180 (97%)	170 (97%)	5 (3%)	42	64
4	M	175/180 (97%)	171 (98%)	4 (2%)	50	70
4	N	175/180 (97%)	171 (98%)	4 (2%)	50	70
All	All	2304/2385 (97%)	2144 (93%)	160 (7%)	15	40

5 of 160 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	246	SER
4	L	67	SER
1	E	271	GLU
2	F	65	LEU
4	M	67	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
3	H	164	HIS
4	N	108	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6   Fit of model and data ⓘ

6.1   Protein, DNA and RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2   Non-standard residues in protein, DNA, RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3   Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4   Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5   Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.