wwPDB X-ray Structure Validation Summary Report (i)

May 29, 2024 – 05:24 PM EDT

PDB ID	:	1FRG
Title	:	CRYSTAL STRUCTURE, SEQUENCE, AND EPITOPE MAPPING OF A PEPTIDE COMPLEX OF AN ANTI-INFLUENZA HA PEPTIDE ANTIBODY FAB 26(SLASH)9: FINE-TUNING ANTIBODY SPECIFICITY
Authors	:	Churchill, M.E.A.; Wilson, I.A.
Deposited on	:	1994-01-17
Resolution	:	2.80 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

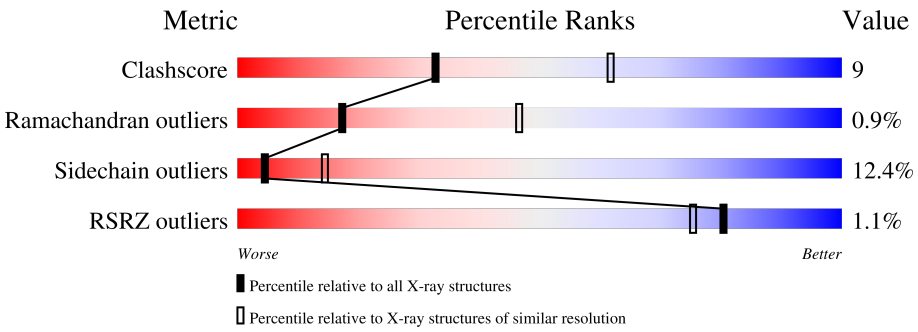
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.80 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	L	217	65% 28% 6% •
2	H	220	2% 63% 32% 5%
3	P	10	90% 10%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 3427 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 IGG2A 26/9 FAB (LIGHT CHAIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	L	217	1689	1056	284	342	7	0	0	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
L	24	LYS	THR	conflict	PIR A31790
L	35	ARG	GLN	conflict	PIR A31790
L	38	PHE	TYR	conflict	PIR A31790
L	43	HIS	GLN	conflict	PIR A31790
L	52	LEU	VAL	conflict	PIR A31790
L	69	SER	THR	conflict	PIR A31790
L	82	THR	SER	conflict	PIR A31790
L	91	ILE	VAL	conflict	PIR A31790
L	100	HIS	ASN	conflict	PIR A31790
L	106	ALA	GLY	conflict	PIR A31790

- Molecule 2 is a protein called IGG2A 26/9 FAB (HEAVY CHAIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	H	220	1672	1060	278	327	7	0	0	0

There are 15 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
H	220	LEU	GLN	conflict	GB 533229
H	234	PHE	SER	conflict	GB 533229
H	245	THR	SER	conflict	GB 533229
H	249	PHE	TYR	conflict	GB 533229
H	256	HIS	GLN	conflict	GB 533229
H	278	GLN	PRO	conflict	GB 533229

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
H	297	PHE	TYR	conflict	GB 533229
H	299	GLU	GLN	conflict	GB 533229
H	301	THR	SER	conflict	GB 533229
H	308	ALA	SER	conflict	GB 533229
H	309	GLY	ALA	conflict	GB 533229
H	310	LEU	MET	conflict	GB 533229
H	322	LYS	ASN	conflict	GB 533229
H	329	ARG	GLN	conflict	GB 533229
H	436	PRO	-	insertion	GB 533229

- Molecule 3 is a protein called INFLUENZA HEMAGGLUTININ HA1 (STRAIN X47) (RESIDUES 101 - 108).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	P	10	65	41	9	15	0	0	1

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	L	1	1	1	0	0

3 Residue-property plots (i)

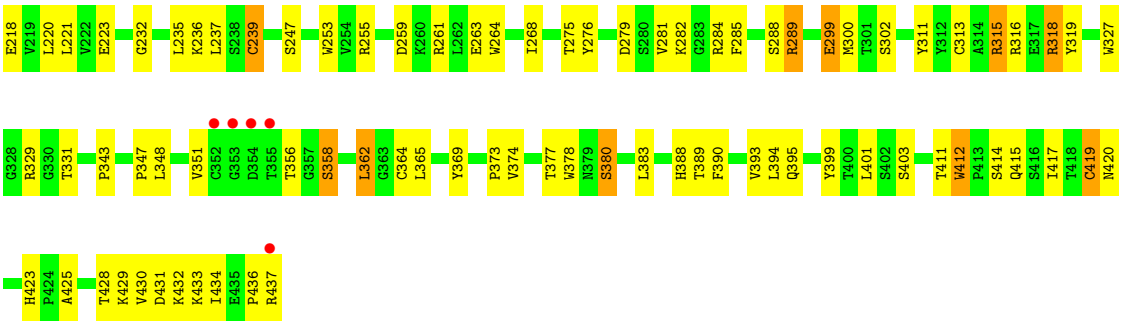
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). 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: IGG2A 26/9 FAB (LIGHT CHAIN)

Chain L:

• Molecule 2: IGG2A 26/9 FAB (HEAVY CHAIN)

Chain H:



• Molecule 3: INFLUENZA HEMAGGLUTININ HA1 (STRAIN X47) (RESIDUES 101 - 108)

Chain P:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	66.60Å 115.10Å 73.30Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	8.00 – 2.80 37.44 – 2.63	Depositor EDS
% Data completeness (in resolution range)	(Not available) (8.00-2.80) 82.2 (37.44-2.63)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.20 (at 2.65Å)	Xtriage
Refinement program	X-PLOR	Depositor
R, R_{free}	0.190 , (Not available) 0.180 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	37.1	Xtriage
Anisotropy	0.219	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 99.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	3427	wwPDB-VP
Average B, all atoms (Å^2)	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.20% 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)

Bond lengths and bond angles in the following residue types are not validated in this section: NH2, ACE

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	L	0.95	0/1729	1.69	29/2346 (1.2%)
2	H	0.92	0/1714	1.75	36/2332 (1.5%)
3	P	0.90	0/63	1.60	1/87 (1.1%)
All	All	0.94	0/3506	1.72	66/4765 (1.4%)

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	L	0	1

There are no bond length outliers.

The worst 5 of 66 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	L	41	TRP	CD1-CG-CD2	9.25	113.70	106.30
1	L	161	ARG	NE-CZ-NH1	8.99	124.80	120.30
2	H	378	TRP	CD1-CG-CD2	8.89	113.41	106.30
2	H	264	TRP	CD1-CG-CD2	8.88	113.41	106.30
1	L	154	TRP	CD1-CG-CD2	8.77	113.32	106.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	L	92	TYR	Sidechain

5.2 Too-close contacts (i)

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	L	1689	0	1631	30	0
2	H	1672	0	1638	27	0
3	P	65	0	57	0	0
4	L	1	0	0	0	0
All	All	3427	0	3326	55	0

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 9.

The worst 5 of 55 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:H:362:LEU:HD22	2:H:434:ILE:HG21	1.69	0.74
1:L:139:VAL:HG22	1:L:184:THR:HG23	1.76	0.68
1:L:142:LEU:HD12	1:L:150:ILE:HD13	1.76	0.66
1:L:52:LEU:HD22	1:L:61:GLU:HG3	1.82	0.61
2:H:235:LEU:HB3	2:H:300:MET:HE3	1.82	0.61

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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	L	215/217 (99%)	201 (94%)	13 (6%)	1 (0%)	29 61
2	H	218/220 (99%)	203 (93%)	12 (6%)	3 (1%)	11 34

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	P	8/10 (80%)	8 (100%)	0	0	100	100
All	All	441/447 (99%)	412 (93%)	25 (6%)	4 (1%)	17	46

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	L	197	SER
2	H	425	ALA
2	H	358	SER
2	H	380	SER

5.3.2 Protein sidechains (i)

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	L	194/194 (100%)	170 (88%)	24 (12%)	4	14
2	H	186/186 (100%)	163 (88%)	23 (12%)	4	14
3	P	7/7 (100%)	6 (86%)	1 (14%)	3	10
All	All	387/387 (100%)	339 (88%)	48 (12%)	4	14

5 of 48 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	H	288	SER
2	H	362	LEU
2	H	299	GLU
2	H	351	VAL
2	H	373	PRO

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	L	143	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	L	204	HIS
2	H	415	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	L	217/217 (100%)	-0.63	0 100 100	3, 18, 49, 63	0
2	H	220/220 (100%)	-0.47	5 (2%) 60 51	4, 22, 56, 117	0
3	P	8/10 (80%)	-0.50	0 100 100	7, 20, 27, 31	0
All	All	445/447 (99%)	-0.55	5 (1%) 80 75	3, 19, 54, 117	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	353	GLY	13.7
2	H	354	ASP	7.8
2	H	355	THR	4.1
2	H	437	ARG	2.6
2	H	352	CYS	2.5

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

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.