



wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 18, 2024 – 04:10 PM EST

PDB ID : 8SJ4
Title : 8F3-1H9-Ara h 6
Authors : Spiller, B.W.; Shrem, R.A.
Deposited on : 2023-04-17
Resolution : 2.67 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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.21
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

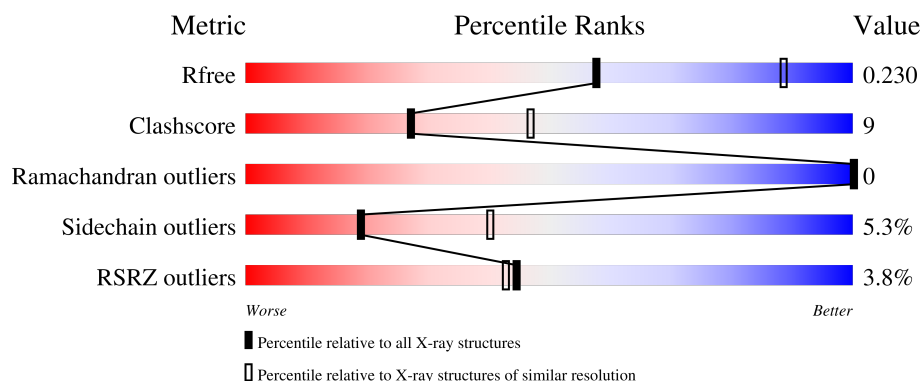
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION






The reported resolution of this entry is 2.67 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))
R_{free}	164625	4708 (2.70-2.66)
Clashscore	180529	5138 (2.70-2.66)
Ramachandran outliers	177936	5071 (2.70-2.66)
Sidechain outliers	177891	5071 (2.70-2.66)
RSRZ outliers	164620	4708 (2.70-2.66)

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 $\leq 5\%$. 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	H	218	
2	L	214	
3	D	214	
4	F	225	
5	A	122	

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 7429 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 1H9 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	H	215	Total	C	N	O	S	0	0	0
			1596	1008	263	319	6			

- Molecule 2 is a protein called 1H9 light chain kappa.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	L	212	Total	C	N	O	S	0	0	0
			1621	1010	277	329	5			

- Molecule 3 is a protein called 8F3 light chain kappa.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	203	Total	C	N	O	S	0	0	0
			1571	987	271	309	4			

- Molecule 4 is a protein called 8F3 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	F	222	Total	C	N	O	S	0	0	0
			1687	1074	279	329	5			

- Molecule 5 is a protein called Conglutin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	A	98	Total	C	N	O	S	0	0	0
			791	464	154	155	18			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	12	MET	-	initiating methionine	UNP Q647G9
A	70	GLY	ARG	conflict	UNP Q647G9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	102	SER	ASN	conflict	UNP Q647G9
A	125	SER	-	expression tag	UNP Q647G9
A	126	GLY	-	expression tag	UNP Q647G9
A	127	SER	-	expression tag	UNP Q647G9
A	128	HIS	-	expression tag	UNP Q647G9
A	129	HIS	-	expression tag	UNP Q647G9
A	130	HIS	-	expression tag	UNP Q647G9
A	131	HIS	-	expression tag	UNP Q647G9
A	132	HIS	-	expression tag	UNP Q647G9
A	133	HIS	-	expression tag	UNP Q647G9

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	H	70	Total O 70 70	0	0
6	L	56	Total O 56 56	0	0
6	D	9	Total O 9 9	0	0
6	F	19	Total O 19 19	0	0
6	A	9	Total O 9 9	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.

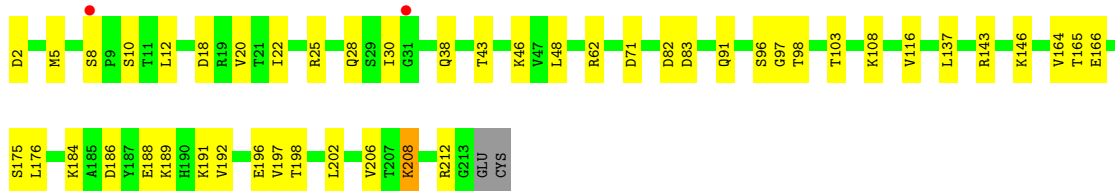
- Molecule 1: 1H9 heavy chain

Chain H: 



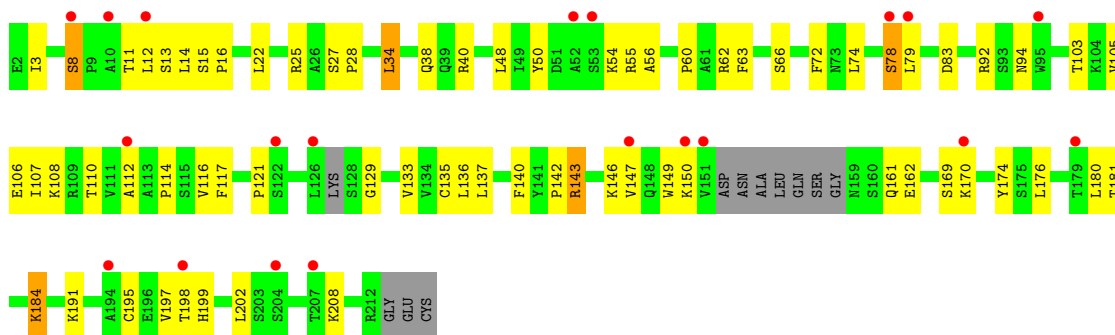
- Molecule 2: 1H9 light chain kappa

Chain L: 




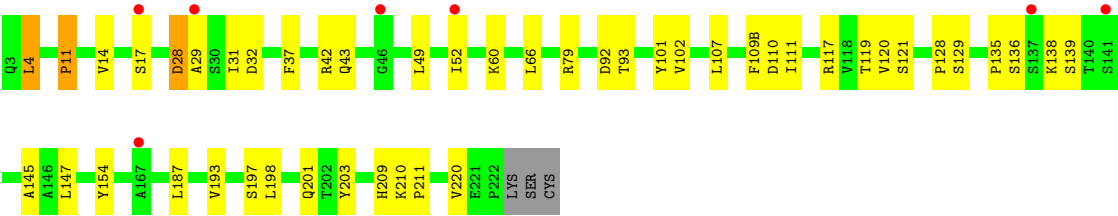
- Molecule 3: 8F3 light chain kappa

Chain D: 

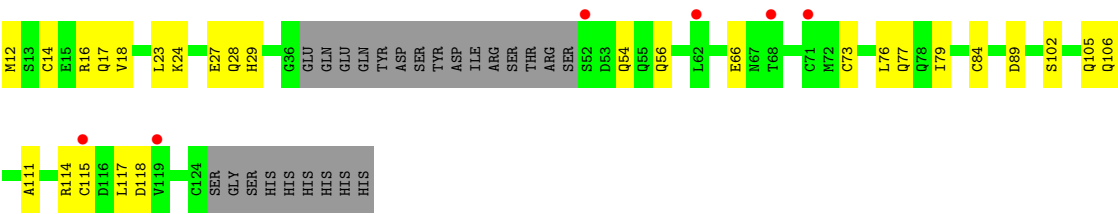


- Molecule 4: 8F3 heavy chain

Chain F: 



● Molecule 5: Conglutin



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	39.38Å 85.45Å 165.83Å 90.00° 96.17° 90.00°	Depositor
Resolution (Å)	54.96 – 2.67 54.96 – 2.67	Depositor EDS
% Data completeness (in resolution range)	99.3 (54.96-2.67) 99.3 (54.96-2.67)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.14 (at 2.69Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.190 , 0.232 0.190 , 0.230	Depositor DCC
R_{free} test set	1529 reflections (4.91%)	wwPDB-VP
Wilson B-factor (Å ²)	44.2	Xtriage
Anisotropy	0.322	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 52.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.031 for h,-k,-h-l	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	7429	wwPDB-VP
Average B, all atoms (Å ²)	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.38% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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	H	0.26	0/1635	0.48	0/2231
2	L	0.28	0/1655	0.51	0/2249
3	D	0.28	0/1607	0.54	0/2183
4	F	0.69	1/1733 (0.1%)	0.59	1/2371 (0.0%)
5	A	0.24	0/795	0.49	0/1057
All	All	0.41	1/7425 (0.0%)	0.53	1/10091 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	F	11	PRO	N-CD	26.30	1.84	1.47

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	F	11	PRO	N-CD-CG	-12.23	84.85	103.20

There are no chirality outliers.

There are no planarity outliers.

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	H	1596	0	1562	10	0
2	L	1621	0	1577	27	0
3	D	1571	0	1528	52	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	F	1687	0	1668	26	0
5	A	791	0	751	14	0
6	A	9	0	0	1	0
6	D	9	0	0	1	0
6	F	19	0	0	2	0
6	H	70	0	0	2	0
6	L	56	0	0	6	0
All	All	7429	0	7086	129	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 129 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)
4:F:11:PRO:N	4:F:11:PRO:CD	1.84	1.23
3:D:13:SER:O	3:D:108:LYS:HE2	1.51	1.10
3:D:146:LYS:HD3	3:D:198:THR:OG1	1.53	1.08
5:A:56:GLN:OE1	6:A:201:HOH:O	1.94	0.85
3:D:11:THR:HG22	3:D:105:VAL:HA	1.64	0.79

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	H	213/218 (98%)	212 (100%)	1 (0%)	0	100	100
2	L	210/214 (98%)	205 (98%)	5 (2%)	0	100	100
3	D	197/214 (92%)	182 (92%)	15 (8%)	0	100	100
4	F	220/225 (98%)	215 (98%)	5 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
5	A	94/122 (77%)	92 (98%)	2 (2%)	0	100	100
All	All	934/993 (94%)	906 (97%)	28 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	H	179/182 (98%)	174 (97%)	5 (3%)	38	65
2	L	185/187 (99%)	180 (97%)	5 (3%)	40	67
3	D	175/183 (96%)	158 (90%)	17 (10%)	6	15
4	F	194/197 (98%)	184 (95%)	10 (5%)	19	41
5	A	92/115 (80%)	85 (92%)	7 (8%)	11	24
All	All	825/864 (96%)	781 (95%)	44 (5%)	19	40

5 of 44 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	F	17	SER
4	F	198	LEU
4	F	28	ASP
4	F	129	SER
5	A	14	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
3	D	125	GLN
4	F	180	GLN
5	A	54	GLN
5	A	82	ASN

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 oligosaccharides 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	H	215/218 (98%)	-0.51	1 (0%) 87 86	21, 36, 64, 105	0
2	L	212/214 (99%)	-0.34	2 (0%) 81 80	19, 39, 61, 83	0
3	D	203/214 (94%)	0.97	20 (9%) 14 13	52, 89, 119, 131	0
4	F	222/225 (98%)	0.17	7 (3%) 50 48	32, 59, 111, 138	0
5	A	98/122 (80%)	0.34	6 (6%) 28 26	30, 72, 118, 138	0
All	All	950/993 (95%)	0.09	36 (3%) 44 43	19, 53, 111, 138	0

The worst 5 of 36 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	D	95	TRP	6.7
4	F	46	GLY	4.4
2	L	8	SER	4.2
3	D	198	THR	4.0
3	D	12	LEU	3.7

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.