



# Full wwPDB X-ray Structure Validation Report ⓘ

May 4, 2022 – 06:02 PM EDT

PDB ID : 7ST8  
Title : Crystal structure of 7H2.2 Fab in complex with SAS1B C-terminal region  
Authors : Legg, M.S.G.; Evans, S.V.  
Deposited on : 2021-11-12  
Resolution : 2.75 Å(reported)

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

We welcome your comments at [validation@mail.wwpdb.org](mailto: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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.13  
EDS : 2.28.1  
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) : 2.28.1

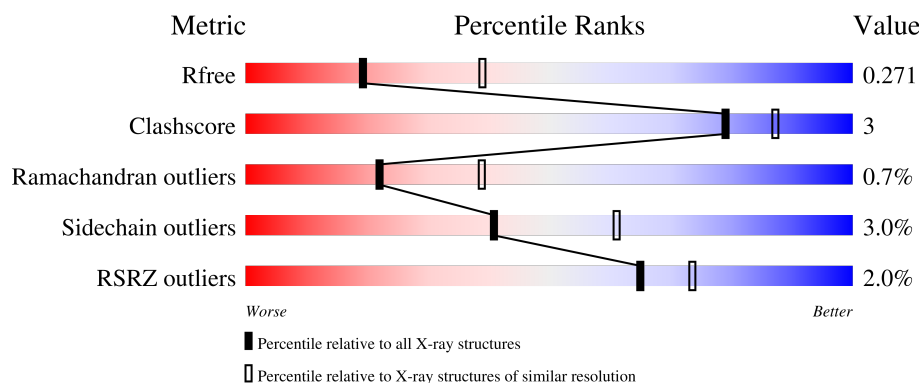
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.75 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1235 (2.78-2.74)
Clashscore	141614	1277 (2.78-2.74)
Ramachandran outliers	138981	1257 (2.78-2.74)
Sidechain outliers	138945	1257 (2.78-2.74)
RSRZ outliers	127900	1207 (2.78-2.74)

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  $\geq 3$ , 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	223	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 1%, green 98%);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> <span>1%</span> <span>92%</span> <span>5%</span> </div> </div>
2	L	217	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 2%, orange 1%, yellow 1%, green 96%);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> <span>2%</span> <span>90%</span> <span>10%</span> </div> </div>
3	S	170	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 1%, green 8%, grey 91%);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> <span>1%</span> <span>8%</span> <span>91%</span> </div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 3381 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 7H2.2 Fab Heavy Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	H	223	Total	C	N	O	S	0	0	0
			1629	1038	267	317	7			

- Molecule 2 is a protein called 7H2.2 Fab Light Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	L	217	Total	C	N	O	S	0	0	0
			1620	1015	271	328	6			

- Molecule 3 is a protein called Astacin-like metalloendopeptidase.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	S	15	Total	C	N	O	0	0	0
			106	67	19	20			

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
S	262	MET	-	initiating methionine	UNP Q6HA08
S	263	GLY	-	expression tag	UNP Q6HA08
S	264	SER	-	expression tag	UNP Q6HA08
S	265	SER	-	expression tag	UNP Q6HA08
S	266	HIS	-	expression tag	UNP Q6HA08
S	267	HIS	-	expression tag	UNP Q6HA08
S	268	HIS	-	expression tag	UNP Q6HA08
S	269	HIS	-	expression tag	UNP Q6HA08
S	270	HIS	-	expression tag	UNP Q6HA08
S	271	HIS	-	expression tag	UNP Q6HA08
S	272	SER	-	expression tag	UNP Q6HA08
S	273	SER	-	expression tag	UNP Q6HA08
S	274	GLY	-	expression tag	UNP Q6HA08
S	275	LEU	-	expression tag	UNP Q6HA08

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
S	276	VAL	-	expression tag	UNP Q6HA08
S	277	PRO	-	expression tag	UNP Q6HA08
S	278	ARG	-	expression tag	UNP Q6HA08
S	279	GLY	-	expression tag	UNP Q6HA08
S	280	SER	-	expression tag	UNP Q6HA08
S	281	HIS	-	expression tag	UNP Q6HA08
S	282	MET	-	expression tag	UNP Q6HA08
S	283	ALA	-	expression tag	UNP Q6HA08

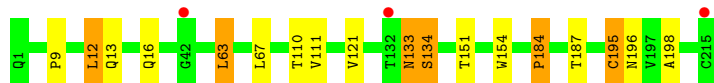
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	H	11	Total O 11 11	0	0
4	L	14	Total O 14 14	0	0
4	S	1	Total O 1 1	0	0

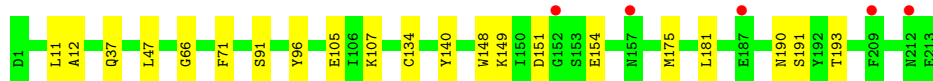
### 3 Residue-property plots [i](#)

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.

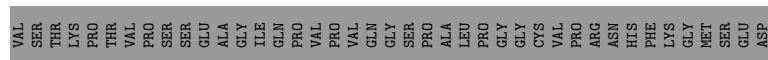
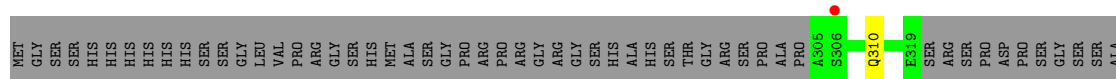
- Molecule 1: 7H2.2 Fab Heavy Chain



- Molecule 2: 7H2.2 Fab Light Chain



- Molecule 3: Astacin-like metalloendopeptidase



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	132.35Å 43.03Å 106.23Å 90.00° 123.89° 90.00°	Depositor
Resolution (Å)	26.41 – 2.75 26.40 – 2.68	Depositor EDS
% Data completeness (in resolution range)	93.3 (26.41-2.75) 90.1 (26.40-2.68)	Depositor EDS
$R_{merge}$	0.16	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.06 (at 2.68Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, $R_{free}$	0.231 , 0.270 0.232 , 0.271	Depositor DCC
$R_{free}$ test set	662 reflections (5.16%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	43.6	Xtriage
Anisotropy	0.592	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 44.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	3381	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.21% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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

### 5.1 Standard geometry

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	0.66	0/1675	0.77	0/2303
2	L	0.65	0/1659	0.75	0/2260
3	S	0.73	0/105	0.72	0/141
All	All	0.66	0/3439	0.76	0/4704

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	H	1629	0	1550	8	0
2	L	1620	0	1516	10	0
3	S	106	0	110	1	0
4	H	11	0	0	0	0
4	L	14	0	0	0	0
4	S	1	0	0	0	0
All	All	3381	0	3176	18	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 3.

All (18) 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:L:151:ASP:N	2:L:191:SER:O	2.38	0.57
1:H:154:TRP:CZ3	1:H:195:CYS:HB3	2.41	0.55
2:L:91:SER:OG	3:S:310:GLN:NE2	2.44	0.51
1:H:184:PRO:HD2	1:H:187:THR:HG21	1.92	0.50
2:L:12:ALA:HA	2:L:105:GLU:O	2.13	0.49
2:L:91:SER:HA	2:L:96:TYR:CD1	2.47	0.49
2:L:149:LYS:HB2	2:L:193:THR:CG2	2.46	0.45
2:L:66:GLY:HA3	2:L:71:PHE:HA	1.99	0.45
2:L:37:GLN:HB2	2:L:47:LEU:HD11	2.00	0.44
1:H:12:LEU:CD1	1:H:16:GLN:HB3	2.49	0.43
2:L:107:LYS:HA	2:L:140:TYR:OH	2.19	0.42
1:H:13:GLN:HB3	1:H:16:GLN:HE21	1.84	0.41
1:H:133:ASN:O	1:H:134:SER:CB	2.69	0.41
2:L:148:TRP:O	2:L:154:GLU:HA	2.20	0.41
1:H:63:LEU:CD1	1:H:67:LEU:HD11	2.51	0.41
1:H:151:THR:CG2	1:H:198:ALA:HB3	2.51	0.41
1:H:9:PRO:HG2	1:H:12:LEU:HD23	2.03	0.40
2:L:181:LEU:HD23	2:L:181:LEU:N	2.36	0.40

There are no symmetry-related clashes.

## 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	H	221/223 (99%)	201 (91%)	17 (8%)	3 (1%)	11	19
2	L	215/217 (99%)	204 (95%)	11 (5%)	0	100	100
3	S	13/170 (8%)	12 (92%)	1 (8%)	0	100	100
All	All	449/610 (74%)	417 (93%)	29 (6%)	3 (1%)	22	39

All (3) Ramachandran outliers are listed below:



Mol	Chain	Res	Type
1	H	134	SER
1	H	133	ASN
1	H	184	PRO

### 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	176/193 (91%)	169 (96%)	7 (4%)	31	51
2	L	176/190 (93%)	172 (98%)	4 (2%)	50	69
3	S	10/129 (8%)	10 (100%)	0	100	100
All	All	362/512 (71%)	351 (97%)	11 (3%)	41	61

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	12	LEU
1	H	63	LEU
1	H	110	THR
1	H	111	VAL
1	H	121	VAL
1	H	195	CYS
1	H	196	ASN
2	L	11	LEU
2	L	134	CYS
2	L	175	MET
2	L	190	ASN

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

Mol	Chain	Res	Type
1	H	16	GLN
1	H	102	HIS
1	H	105	GLN
3	S	310	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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	H	223/223 (100%)	-0.06	3 (1%) 77 84	26, 45, 70, 94	0
2	L	217/217 (100%)	-0.15	5 (2%) 60 69	26, 39, 81, 99	0
3	S	15/170 (8%)	-0.18	1 (6%) 17 21	32, 36, 61, 62	0
All	All	455/610 (74%)	-0.11	9 (1%) 65 73	26, 43, 79, 99	0

All (9) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	132	THR	3.6
2	L	212	ASN	3.2
3	S	306	SER	2.6
2	L	209	PHE	2.4
1	H	215	CYS	2.3
1	H	42	GLY	2.3
2	L	152	GLY	2.3
2	L	187	GLU	2.3
2	L	157	ASN	2.1

### 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 ⓘ

There are no such residues in this entry.