



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 11:44 pm BST

PDB ID : 6AZ2  
Title : Crystal structure of Asf1-Fab 12E complex  
Authors : Bailey, L.J.; Kossiakoff, A.A.  
Deposited on : 2017-09-09  
Resolution : 2.48 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](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 : 2.11  
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) : 2.11

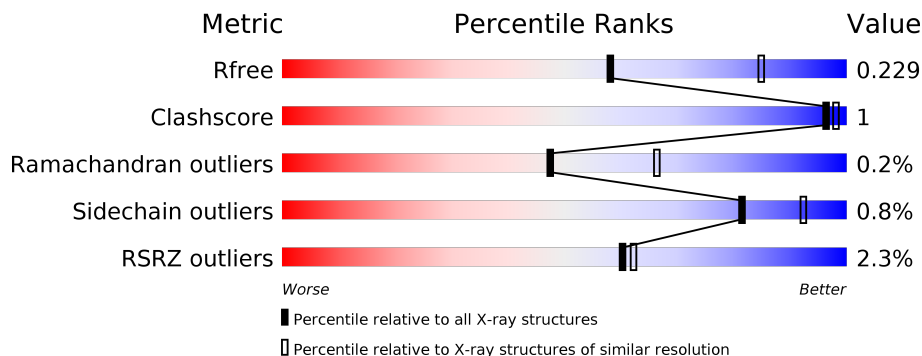
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.48 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	5857 (2.50-2.46)
Clashscore	141614	6594 (2.50-2.46)
Ramachandran outliers	138981	6469 (2.50-2.46)
Sidechain outliers	138945	6471 (2.50-2.46)
RSRZ outliers	127900	5738 (2.50-2.46)

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 on 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	A	229	 89% 7%
1	C	229	 90% 7%
2	B	154	 86% 12%
2	D	154	 91% 5% 5%
3	E	215	 95%
3	F	215	 96%

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 17431 atoms, of which 8481 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 Fab Heavy Chain.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	213	3148	1011	1555	264	312	6	0	0	0
1	C	214	3157	1016	1558	265	312	6	0	0	0

- Molecule 2 is a protein called Histone chaperone ASF1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
2	B	136	2162	706	1072	176	206	2	0	0	0
2	D	147	2329	758	1156	188	225	2	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1	GLY	-	expression tag	UNP P32447
D	1	GLY	-	expression tag	UNP P32447

- Molecule 3 is a protein called Fab Light Chain.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
3	E	210	3181	1011	1569	269	327	5	0	0	0
3	F	211	3190	1014	1571	272	328	5	0	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	47	Total	O	0	0
			47	47		

*Continued on next page...*

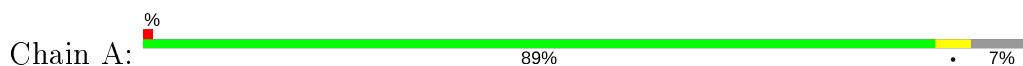
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>	<b>ZeroOcc</b>	<b>AltConf</b>
4	C	49	Total O 49 49	0	0
4	B	52	Total O 52 52	0	0
4	D	61	Total O 61 61	0	0
4	E	24	Total O 24 24	0	0
4	F	31	Total O 31 31	0	0

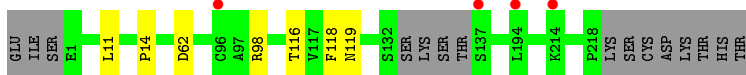
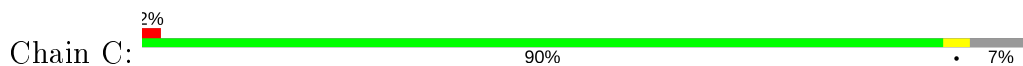
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.

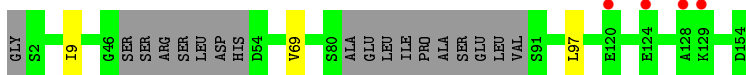
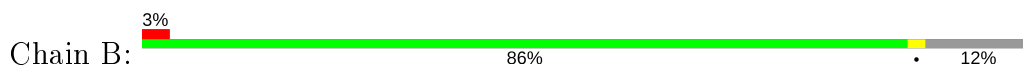
- Molecule 1: Fab Heavy Chain



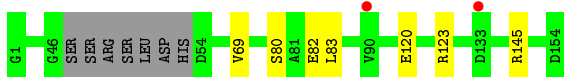
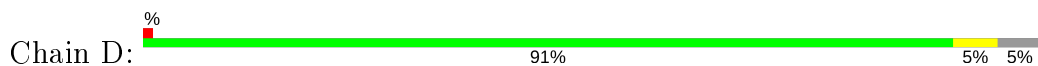
- Molecule 1: Fab Heavy Chain



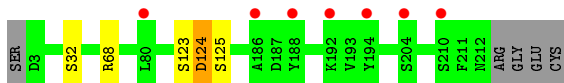
- Molecule 2: Histone chaperone ASF1



- Molecule 2: Histone chaperone ASF1

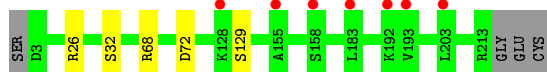


- Molecule 3: Fab Light Chain



- Molecule 3: Fab Light Chain

Chain F:  3% 96%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	99.67Å 109.74Å 122.17Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.05 – 2.48 50.05 – 2.48	Depositor EDS
% Data completeness (in resolution range)	98.8 (50.05-2.48) 98.8 (50.05-2.48)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.92 (at 2.48Å)	Xtrriage
Refinement program	PHENIX (1.10_2152: ???)	Depositor
R, $R_{free}$	0.193 , 0.228 0.195 , 0.229	Depositor DCC
$R_{free}$ test set	2416 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	40.1	Xtrriage
Anisotropy	0.142	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 38.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	17431	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	54.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.59% 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 [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	0.27	0/1632	0.48	0/2225
1	C	0.28	0/1639	0.48	0/2236
2	B	0.26	0/1115	0.47	0/1517
2	D	0.27	0/1200	0.50	0/1635
3	E	0.26	0/1648	0.47	0/2240
3	F	0.27	0/1655	0.46	0/2250
All	All	0.27	0/8889	0.48	0/12103

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 [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	A	1593	1555	1554	3	0
1	C	1599	1558	1561	2	1
2	B	1090	1072	1069	2	0
2	D	1173	1156	1160	5	1
3	E	1612	1569	1569	3	0
3	F	1619	1571	1571	3	0
4	A	47	0	0	0	0
4	B	52	0	0	0	0
4	C	49	0	0	0	0
4	D	61	0	0	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	E	24	0	0	0	0
4	F	31	0	0	1	0
All	All	8950	8481	8484	17	1

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 1.

The worst 5 of 17 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)
3:F:32:SER:O	3:F:68:ARG:NH1	2.13	0.82
3:F:129:SER:OG	4:F:301:HOH:O	2.08	0.72
3:E:32:SER:O	3:E:68:ARG:NH1	2.23	0.71
2:D:82:GLU:O	2:D:83:LEU:HB2	2.03	0.59
2:D:120:GLU:OE1	2:D:123:ARG:NH2	2.39	0.55

All (1) 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)
1:C:11:LEU:O	2:D:145:ARG:HH22[4_455]	1.58	0.02

## 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	A	209/229 (91%)	206 (99%)	3 (1%)	0	100 100
1	C	210/229 (92%)	207 (99%)	2 (1%)	1 (0%)	29 46
2	B	130/154 (84%)	125 (96%)	5 (4%)	0	100 100
2	D	143/154 (93%)	138 (96%)	5 (4%)	0	100 100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	E	208/215 (97%)	198 (95%)	9 (4%)	1 (0%)	29	46
3	F	209/215 (97%)	201 (96%)	8 (4%)	0	100	100
All	All	1109/1196 (93%)	1075 (97%)	32 (3%)	2 (0%)	47	66

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	E	124	ASP
1	C	119	ASN

### 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	A	177/193 (92%)	173 (98%)	4 (2%)	50	74
1	C	177/193 (92%)	175 (99%)	2 (1%)	73	88
2	B	123/141 (87%)	123 (100%)	0	100	100
2	D	133/141 (94%)	132 (99%)	1 (1%)	81	92
3	E	186/190 (98%)	185 (100%)	1 (0%)	88	95
3	F	186/190 (98%)	186 (100%)	0	100	100
All	All	982/1048 (94%)	974 (99%)	8 (1%)	81	92

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	209	ASN
3	E	125	SER
1	C	98	ARG
1	A	98	ARG
1	C	62	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 carbohydrates 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	A	213/229 (93%)	0.16	2 (0%) 84 86	25, 48, 71, 95	0
1	C	214/229 (93%)	0.30	4 (1%) 66 68	23, 45, 83, 105	0
2	B	136/154 (88%)	0.31	4 (2%) 51 54	23, 41, 92, 117	0
2	D	147/154 (95%)	0.18	2 (1%) 75 77	22, 34, 70, 89	0
3	E	210/215 (97%)	0.50	7 (3%) 46 49	26, 56, 87, 104	0
3	F	211/215 (98%)	0.40	7 (3%) 46 49	22, 53, 87, 102	0
All	All	1131/1196 (94%)	0.32	26 (2%) 60 62	22, 48, 85, 117	0

The worst 5 of 26 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	F	128	LYS	4.6
3	E	80	LEU	4.5
1	C	137	SER	4.1
2	B	124	GLU	3.9
3	E	186	ALA	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 carbohydrates 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.