



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 5, 2025 – 03:25 PM JST

PDB ID : 8YOR
Title : Complex of HARSWHEP and a-HARSWHEP Fab
Authors : Zhai, L.
Deposited on : 2024-03-13
Resolution : 2.30 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

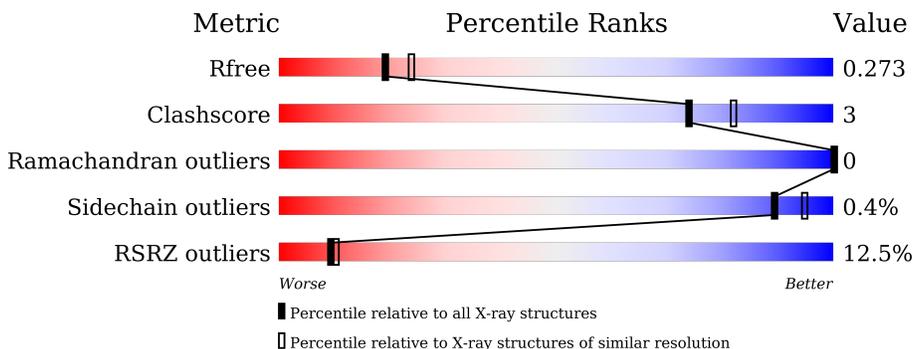
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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	5963 (2.30-2.30)
Clashscore	180529	6698 (2.30-2.30)
Ramachandran outliers	177936	6640 (2.30-2.30)
Sidechain outliers	177891	6640 (2.30-2.30)
RSRZ outliers	164620	5963 (2.30-2.30)

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	A	217	 6% 89% 11%
1	C	217	 14% 94% 6%
2	B	234	 6% 89% 7% .
2	D	234	 6% 91% 5% .
3	E	66	 32% 58% 6% 36%
3	F	66	 42% 44% 17% 39%

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 7924 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 a-HARSWHEP Fab-light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	217	Total 1681	C 1054	N 281	O 340	S 6	0	0	0
1	C	217	Total 1681	C 1054	N 281	O 340	S 6	0	0	0

- Molecule 2 is a protein called a-HARSWHEP Fab-heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	224	Total 1670	C 1042	N 279	O 340	S 9	0	0	0
2	D	225	Total 1676	C 1045	N 280	O 341	S 10	0	0	0

- Molecule 3 is a protein called Histidine-tRNA ligase, cytoplasmic.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	E	42	Total 331	C 209	N 61	O 61	0	0	0
3	F	40	Total 312	C 197	N 56	O 59	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	68	HIS	-	expression tag	UNP P12081
E	69	HIS	-	expression tag	UNP P12081
E	70	HIS	-	expression tag	UNP P12081
E	71	HIS	-	expression tag	UNP P12081
E	72	HIS	-	expression tag	UNP P12081
E	73	HIS	-	expression tag	UNP P12081
F	68	HIS	-	expression tag	UNP P12081
F	69	HIS	-	expression tag	UNP P12081

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	70	HIS	-	expression tag	UNP P12081
F	71	HIS	-	expression tag	UNP P12081
F	72	HIS	-	expression tag	UNP P12081
F	73	HIS	-	expression tag	UNP P12081

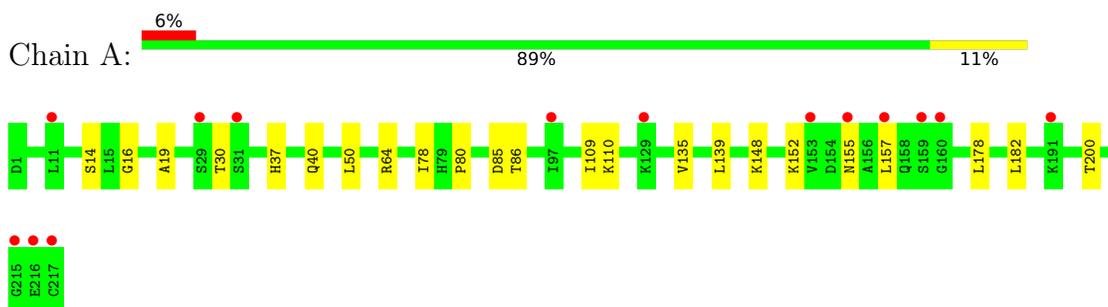
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	129	Total 129	O 129	0	0
4	B	179	Total 179	O 179	0	0
4	E	19	Total 19	O 19	0	0
4	C	82	Total 82	O 82	0	0
4	F	18	Total 18	O 18	0	0
4	D	146	Total 146	O 146	0	0

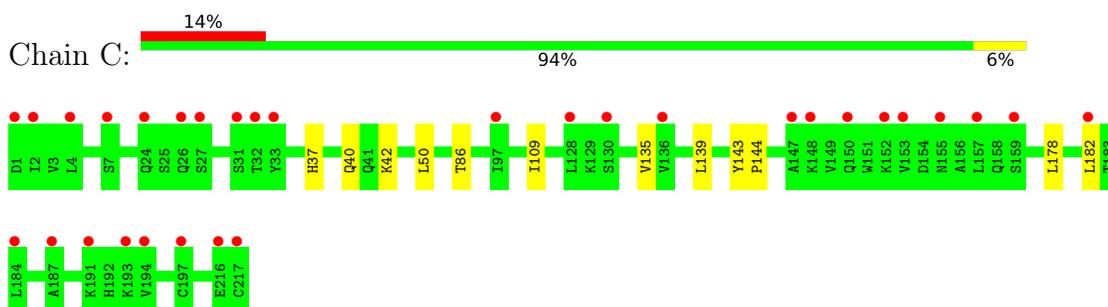
3 Residue-property plots [i](#)

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.

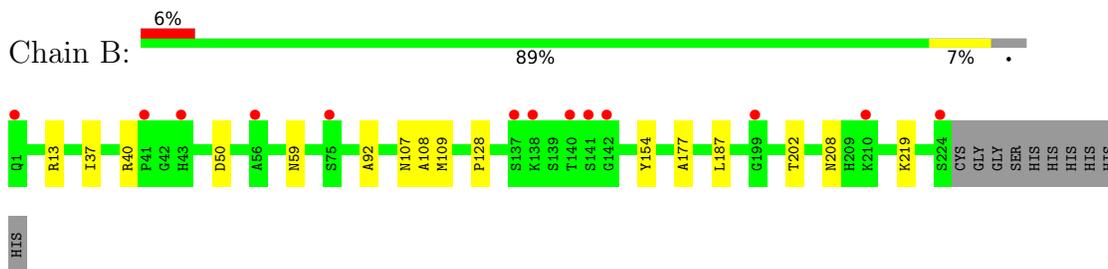
- Molecule 1: α -HARSWHEP Fab-light chain



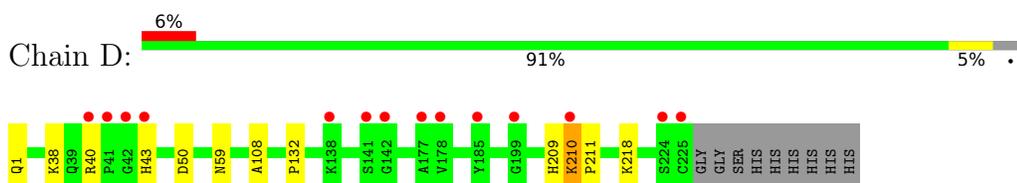
- Molecule 1: α -HARSWHEP Fab-light chain



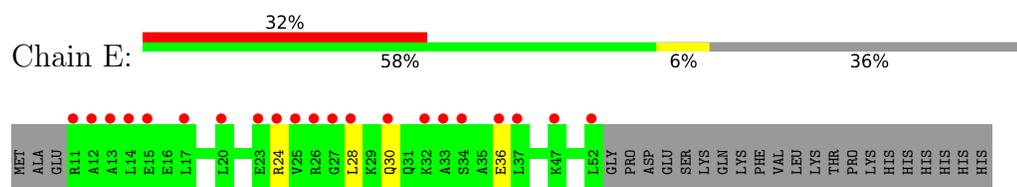
- Molecule 2: α -HARSWHEP Fab-heavy chain



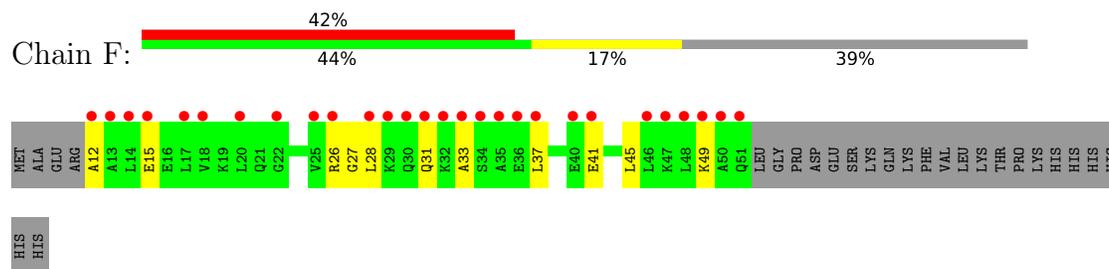
- Molecule 2: α -HARSWHEP Fab-heavy chain



- Molecule 3: Histidine-tRNA ligase, cytoplasmic



- Molecule 3: Histidine-tRNA ligase, cytoplasmic



4 Data and refinement statistics

Property	Value	Source
Space group	I 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	86.75Å 113.56Å 149.50Å 90.00° 106.60° 90.00°	Depositor
Resolution (Å)	89.00 – 2.30 89.00 – 2.30	Depositor EDS
% Data completeness (in resolution range)	96.8 (89.00-2.30) 96.9 (89.00-2.30)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.24 (at 2.29Å)	Xtrriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.227 , 0.273 0.228 , 0.273	Depositor DCC
R_{free} test set	3068 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å ²)	32.9	Xtrriage
Anisotropy	0.033	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 46.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	7924	wwPDB-VP
Average B, all atoms (Å ²)	45.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 35.95 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 5.3859e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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	A	0.21	0/1722	0.38	0/2342
1	C	0.21	0/1722	0.38	0/2342
2	B	0.22	0/1707	0.41	0/2324
2	D	0.22	0/1713	0.41	0/2332
3	E	0.19	0/330	0.32	0/437
3	F	0.19	0/311	0.34	0/412
All	All	0.21	0/7505	0.39	0/10189

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	1681	0	1616	13	0
1	C	1681	0	1615	7	0
2	B	1670	0	1621	9	0
2	D	1676	0	1625	9	0
3	E	331	0	368	2	0
3	F	312	0	344	6	0
4	A	129	0	0	1	0
4	B	179	0	0	1	0
4	C	82	0	0	1	0
4	D	146	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	19	0	0	0	0
4	F	18	0	0	1	0
All	All	7924	0	7189	43	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.

The worst 5 of 43 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:D:210:LYS:HG3	2:D:211:PRO:HD3	1.80	0.64
3:F:12:ALA:HB1	3:F:15:GLU:HB3	1.82	0.62
1:C:86:THR:HB	1:C:109:ILE:HG12	1.87	0.57
1:C:135:VAL:HB	1:C:182:LEU:HB3	1.86	0.57
1:A:155:ASN:ND2	4:A:304:HOH:O	2.37	0.54

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	A	215/217 (99%)	208 (97%)	7 (3%)	0	100	100
1	C	215/217 (99%)	205 (95%)	10 (5%)	0	100	100
2	B	222/234 (95%)	219 (99%)	3 (1%)	0	100	100
2	D	223/234 (95%)	217 (97%)	6 (3%)	0	100	100
3	E	40/66 (61%)	40 (100%)	0	0	100	100
3	F	38/66 (58%)	38 (100%)	0	0	100	100
All	All	953/1034 (92%)	927 (97%)	26 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	191/191 (100%)	191 (100%)	0	100	100
1	C	191/191 (100%)	191 (100%)	0	100	100
2	B	189/197 (96%)	188 (100%)	1 (0%)	86	93
2	D	190/197 (96%)	189 (100%)	1 (0%)	86	93
3	E	34/56 (61%)	33 (97%)	1 (3%)	37	54
3	F	32/56 (57%)	32 (100%)	0	100	100
All	All	827/888 (93%)	824 (100%)	3 (0%)	89	95

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	208	ASN
3	E	30	GLN
2	D	210	LYS

Sometimes 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 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	A	217/217 (100%)	0.43	14 (6%) 26 28	24, 39, 65, 121	0
1	C	217/217 (100%)	0.96	31 (14%) 7 8	30, 52, 90, 132	0
2	B	224/234 (95%)	0.30	13 (5%) 30 32	23, 32, 55, 96	0
2	D	225/234 (96%)	0.51	14 (6%) 28 29	23, 37, 58, 99	0
3	E	42/66 (63%)	1.90	21 (50%) 0 0	38, 60, 89, 97	0
3	F	40/66 (60%)	2.73	28 (70%) 0 0	57, 78, 106, 109	0
All	All	965/1034 (93%)	0.70	121 (12%) 9 10	23, 40, 80, 132	0

The worst 5 of 121 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	F	13	ALA	5.6
3	F	12	ALA	5.5
3	F	14	LEU	4.8
3	F	20	LEU	4.7
2	D	225	CYS	4.4

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.