



Full wwPDB X-ray Structure Validation Report ⓘ

Dec 12, 2023 – 08:10 PM JST

PDB ID : 8HSW
Title : CRYSTAL STRUCTURE OF BAT MHC CLASS I MYLU-B-67 FOR 2.1
ANGSTROM
Authors : Wang, S.Q.; Zhang, N.Z.
Deposited on : 2022-12-20
Resolution : 2.03 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

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
Xtrriage (Phenix) : 1.13
EDS : 2.36
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

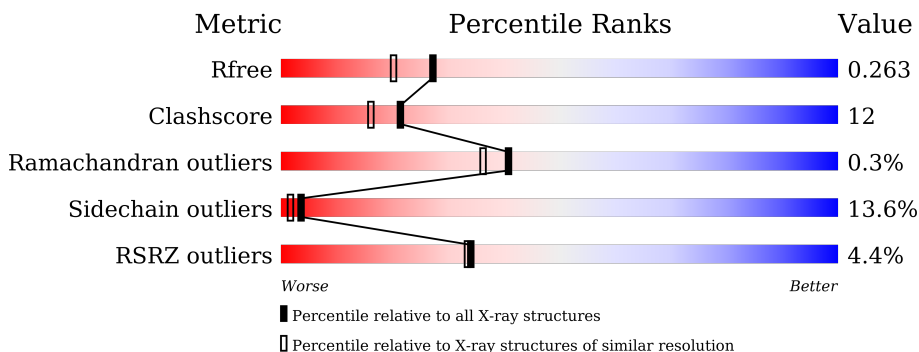
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION



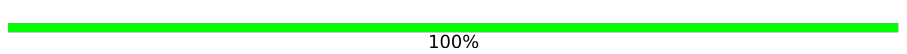
The reported resolution of this entry is 2.03 Å.

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	10434 (2.04-2.00)
Clashscore	141614	11643 (2.04-2.00)
Ramachandran outliers	138981	11493 (2.04-2.00)
Sidechain outliers	138945	11492 (2.04-2.00)
RSRZ outliers	127900	10220 (2.04-2.00)

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	280	 6% 69% 26% 6%
2	B	97	 74% 22% .
3	C	9	 100%

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 3163 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 Ig-like domain-containing protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	280	2267	1428	403	428	8	0	0	0

- Molecule 2 is a protein called Beta-2-microglobulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	97	800	509	138	149	4	0	0	0

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1	GLU	-	expression tag	UNP L5K3Y9
B	2	PRO	-	expression tag	UNP L5K3Y9
B	3	ARG	-	expression tag	UNP L5K3Y9
B	96	ARG	-	expression tag	UNP L5K3Y9
B	97	ASN	-	expression tag	UNP L5K3Y9

- Molecule 3 is a protein called THR-PRO-GLN-SER-ALA-PRO-HIS-GLY-VAL.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	C	9	62	38	12	12	0	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	18	Total	O	0	0
			18	18		
4	B	15	Total	O	0	0
			15	15		

Continued on next page...

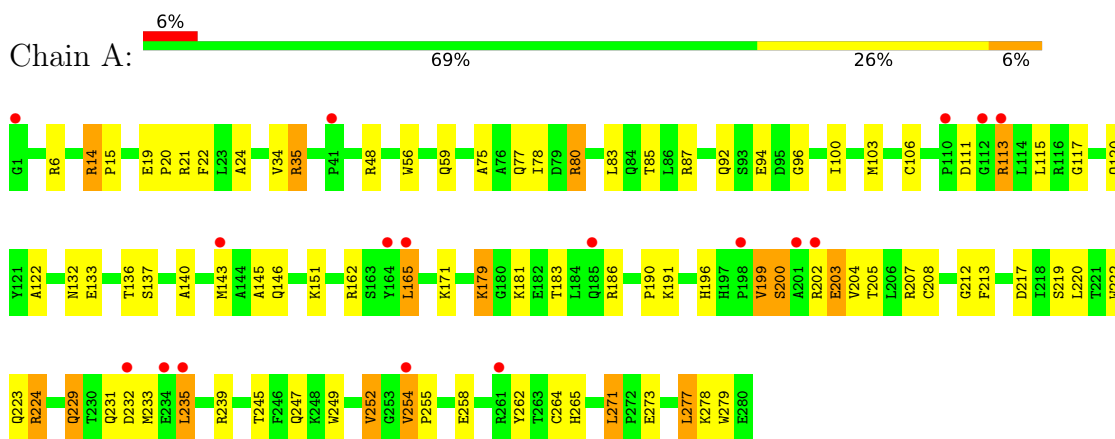
Continued from previous page...

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

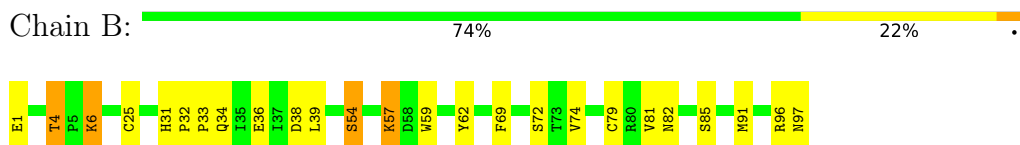
3 Residue-property plots [i](#)

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: Ig-like domain-containing protein



- Molecule 2: Beta-2-microglobulin



- Molecule 3: THR-PRO-GLN-SER-ALA-PRO-HIS-GLY-VAL



There are no outlier residues recorded for this chain.

4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	75.74Å 80.15Å 144.24Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.80 – 2.03 30.78 – 2.03	Depositor EDS
% Data completeness (in resolution range)	67.8 (30.80-2.03) 67.9 (30.78-2.03)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.25 (at 2.03Å)	Xtrriage
Refinement program	REFMAC 5.8.0123	Depositor
R, R_{free}	0.234 , 0.262 0.241 , 0.263	Depositor DCC
R_{free} test set	1024 reflections (5.25%)	wwPDB-VP
Wilson B-factor (Å ²)	43.6	Xtrriage
Anisotropy	0.030	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 23.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	3163	wwPDB-VP
Average B, all atoms (Å ²)	49.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.10% 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	A	0.71	2/2336 (0.1%)	0.90	0/3178
2	B	0.63	0/826	0.88	0/1123
3	C	0.70	0/64	0.86	0/88
All	All	0.69	2/3226 (0.1%)	0.89	0/4389

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	271	LEU	C-N	8.56	1.50	1.34
1	A	273	GLU	C-N	8.29	1.50	1.34

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	2267	0	2125	59	0
2	B	800	0	762	16	0
3	C	62	0	60	0	0
4	A	18	0	0	1	0
4	B	15	0	0	2	0
4	C	1	0	0	0	0
All	All	3163	0	2947	73	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 12.

All (73) 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)
1:A:264:CYS:HB3	1:A:277:LEU:CD1	2.00	0.91
1:A:117:GLY:HA3	1:A:165:LEU:HD11	1.51	0.88
2:B:1:GLU:HA	4:B:115:HOH:O	1.84	0.77
1:A:111:ASP:OD2	1:A:113:ARG:HB2	1.87	0.73
1:A:235:LEU:HD23	1:A:235:LEU:O	1.89	0.73
1:A:122:ALA:HB2	2:B:59:TRP:CE2	2.25	0.72
1:A:35:ARG:HG2	1:A:48:ARG:HD3	1.74	0.68
1:A:75:ALA:O	1:A:78:ILE:HG22	1.93	0.67
2:B:6:LYS:HA	2:B:6:LYS:HE2	1.77	0.67
1:A:77:GLN:HE22	1:A:80:ARG:NH1	1.94	0.65
2:B:31:HIS:CD2	2:B:32:PRO:HA	2.32	0.65
1:A:224:ARG:HB2	1:A:229:GLN:HE22	1.61	0.64
1:A:117:GLY:CA	1:A:165:LEU:HD11	2.27	0.63
1:A:199:VAL:O	1:A:199:VAL:HG13	2.01	0.60
2:B:57:LYS:HE3	2:B:57:LYS:H	1.66	0.60
1:A:35:ARG:HG2	1:A:48:ARG:CD	2.31	0.59
2:B:1:GLU:HG2	2:B:1:GLU:O	2.05	0.57
1:A:264:CYS:HB3	1:A:277:LEU:HD12	1.87	0.56
1:A:106:CYS:HB3	1:A:165:LEU:HD21	1.88	0.56
1:A:196:HIS:CE1	1:A:204:VAL:HG21	2.41	0.56
1:A:132:ASN:HD22	1:A:137:SER:HB3	1.71	0.55
1:A:223:GLN:HG3	1:A:265:HIS:CD2	2.41	0.55
1:A:133:GLU:O	1:A:133:GLU:HG2	2.07	0.55
1:A:146:GLN:NE2	1:A:146:GLN:HA	2.23	0.53
1:A:117:GLY:HA3	1:A:165:LEU:CD1	2.34	0.52
1:A:224:ARG:HB2	1:A:229:GLN:NE2	2.24	0.52
2:B:36:GLU:O	2:B:81:VAL:HA	2.09	0.52
1:A:207:ARG:HD3	1:A:249:TRP:CE3	2.45	0.51
1:A:277:LEU:HD12	1:A:277:LEU:H	1.75	0.51
1:A:277:LEU:HD12	1:A:277:LEU:N	2.23	0.51
2:B:32:PRO:HB2	2:B:33:PRO:HD2	1.91	0.51
1:A:19:GLU:OE1	1:A:80:ARG:NH2	2.45	0.50
1:A:77:GLN:HE22	1:A:80:ARG:HH11	1.59	0.50
1:A:103:MET:CE	1:A:120:GLN:HE21	2.24	0.50
1:A:136:THR:HG23	1:A:162:ARG:NH1	2.27	0.50
1:A:202:ARG:O	1:A:255:PRO:HA	2.12	0.49
1:A:15:PRO:HG2	1:A:96:GLY:O	2.13	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:6:ARG:HD2	1:A:103:MET:SD	2.53	0.49
1:A:106:CYS:HB3	1:A:165:LEU:CD2	2.42	0.49
1:A:264:CYS:HB3	1:A:277:LEU:HD11	1.93	0.48
1:A:136:THR:CG2	1:A:162:ARG:HH11	2.27	0.48
1:A:14:ARG:NH1	1:A:21:ARG:HB2	2.29	0.47
2:B:4:THR:HA	2:B:85:SER:OG	2.14	0.47
1:A:196:HIS:HB2	1:A:279:TRP:CH2	2.51	0.46
1:A:132:ASN:ND2	1:A:137:SER:HB3	2.31	0.46
1:A:233:MET:HB2	1:A:252:VAL:HG13	1.98	0.45
1:A:122:ALA:HB2	2:B:59:TRP:CD2	2.50	0.45
1:A:83:LEU:HD21	1:A:100:ILE:HG22	1.99	0.45
1:A:190:PRO:HD2	1:A:271:LEU:HG	1.99	0.44
2:B:25:CYS:HB2	2:B:39:LEU:HD21	1.99	0.44
1:A:140:ALA:HB1	1:A:145:ALA:CB	2.48	0.44
1:A:146:GLN:HA	1:A:146:GLN:HE21	1.81	0.44
1:A:203:GLU:HA	1:A:254:VAL:O	2.17	0.44
1:A:207:ARG:HD2	1:A:249:TRP:CG	2.52	0.44
1:A:224:ARG:HG3	1:A:262:TYR:CZ	2.53	0.44
1:A:103:MET:HE3	1:A:120:GLN:HE21	1.83	0.44
2:B:79:CYS:O	2:B:91:MET:HA	2.18	0.43
2:B:57:LYS:H	2:B:57:LYS:CE	2.30	0.43
1:A:19:GLU:HB3	1:A:20:PRO:HD2	1.99	0.43
1:A:224:ARG:HB2	1:A:229:GLN:OE1	2.19	0.42
2:B:38:ASP:HA	4:B:109:HOH:O	2.19	0.42
1:A:35:ARG:HG2	1:A:48:ARG:CG	2.50	0.42
1:A:239:ARG:HG2	1:A:247:GLN:HB2	2.02	0.41
2:B:54:SER:HB3	2:B:62:TYR:CZ	2.55	0.41
1:A:208:CYS:HB2	1:A:222:TRP:CZ2	2.55	0.41
2:B:4:THR:HG22	2:B:85:SER:HB2	2.03	0.41
1:A:22:PHE:CE1	1:A:24:ALA:HB2	2.56	0.41
1:A:56:TRP:CG	1:A:179:LYS:HG3	2.56	0.41
1:A:212:GLY:HA2	1:A:245:THR:OG1	2.21	0.41
1:A:87:ARG:NH1	1:A:92:GLN:O	2.55	0.40
1:A:190:PRO:CA	1:A:213:PHE:HB3	2.51	0.40
1:A:196:HIS:HB2	1:A:279:TRP:CZ3	2.57	0.40
1:A:219:SER:HA	4:A:309:HOH:O	2.21	0.40

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	278/280 (99%)	260 (94%)	17 (6%)	1 (0%)	34	28
2	B	95/97 (98%)	92 (97%)	3 (3%)	0	100	100
3	C	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
All	All	380/386 (98%)	358 (94%)	21 (6%)	1 (0%)	41	36

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	200	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	A	233/233 (100%)	199 (85%)	34 (15%)	3	1
2	B	91/91 (100%)	80 (88%)	11 (12%)	5	2
3	C	7/7 (100%)	7 (100%)	0	100	100
All	All	331/331 (100%)	286 (86%)	45 (14%)	3	2

All (45) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	14	ARG
1	A	34	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	35	ARG
1	A	59	GLN
1	A	80	ARG
1	A	85	THR
1	A	94	GLU
1	A	113	ARG
1	A	115	LEU
1	A	143	MET
1	A	151	LYS
1	A	165	LEU
1	A	171	LYS
1	A	179	LYS
1	A	181	LYS
1	A	183	THR
1	A	186	ARG
1	A	191	LYS
1	A	199	VAL
1	A	200	SER
1	A	203	GLU
1	A	205	THR
1	A	217	ASP
1	A	220	LEU
1	A	224	ARG
1	A	229	GLN
1	A	231	GLN
1	A	232	ASP
1	A	235	LEU
1	A	252	VAL
1	A	254	VAL
1	A	258	GLU
1	A	277	LEU
1	A	278	LYS
2	B	4	THR
2	B	6	LYS
2	B	34	GLN
2	B	54	SER
2	B	57	LYS
2	B	69	PHE
2	B	72	SER
2	B	74	VAL
2	B	82	ASN
2	B	96	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	97	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (11) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	32	GLN
1	A	77	GLN
1	A	120	GLN
1	A	132	ASN
1	A	146	GLN
1	A	196	HIS
1	A	231	GLN
1	A	260	GLN
2	B	31	HIS
2	B	82	ASN
3	C	7	HIS

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

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	280/280 (100%)	0.40	17 (6%) 21 20	29, 50, 78, 96	0
2	B	97/97 (100%)	-0.13	0 100 100	27, 39, 56, 66	0
3	C	9/9 (100%)	0.42	0 100 100	36, 40, 44, 45	0
All	All	386/386 (100%)	0.27	17 (4%) 34 33	27, 46, 74, 96	0

All (17) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	235	LEU	4.7
1	A	110	PRO	4.3
1	A	165	LEU	4.1
1	A	112	GLY	3.2
1	A	201	ALA	3.2
1	A	232	ASP	3.0
1	A	234	GLU	2.9
1	A	254	VAL	2.6
1	A	202	ARG	2.5
1	A	261	ARG	2.4
1	A	164	TYR	2.4
1	A	113	ARG	2.4
1	A	185	GLN	2.4
1	A	41	PRO	2.2
1	A	198	PRO	2.2
1	A	1	GLY	2.1
1	A	143	MET	2.0

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.