



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 3, 2024 – 10:13 AM EDT

PDB ID : 8UKH
Title : Crystal structure of Plasmodium falciparum CeTOS in complex with antibody 4h12
Authors : Tang, W.K.; Tolia, N.H.; Urusova, D.
Deposited on : 2023-10-13
Resolution : 3.52 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) : 1.20.1
EDS : 3.0
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.002 (Gargrove)
Density-Fitness : 1.0.11
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.38.3

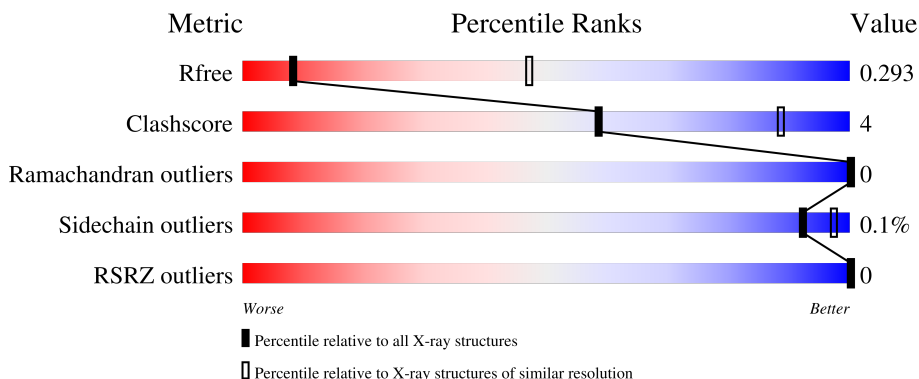
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.52 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.



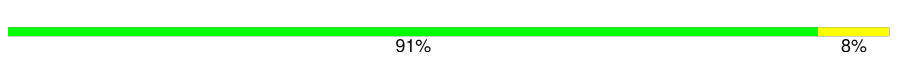
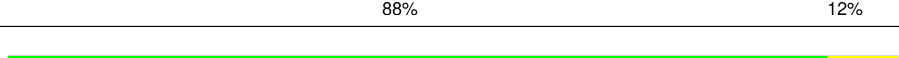
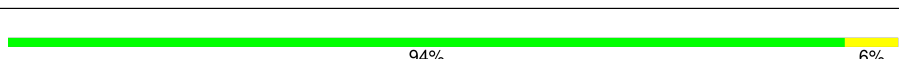
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	1089 (3.58-3.46)
Clashscore	180529	1165 (3.58-3.46)
Ramachandran outliers	177936	1150 (3.58-3.46)
Sidechain outliers	177891	1151 (3.58-3.46)
RSRZ outliers	164620	1088 (3.58-3.46)

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 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	168	
1	B	168	
1	C	168	
1	D	168	
2	H	218	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	I	218	 91% 8%
2	J	218	 94% 6%
2	K	218	 88% 12%
3	L	210	 92% 8%
3	M	210	 94% 6%
3	N	210	 89% 10%
3	O	210	 87% 13%

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 29232 atoms, of which 14469 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 Cell traversal protein for ookinetes and sporozoites.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
1	A	57	899	277	461	71	88	2	0	0	0
1	B	56	888	274	456	70	86	2	0	0	0
1	C	58	915	282	468	73	90	2	0	0	0
1	D	58	915	282	468	73	90	2	0	0	0

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	23	MET	-	initiating methionine	UNP Q8I5P1
A	24	GLY	-	expression tag	UNP Q8I5P1
A	183	LEU	-	expression tag	UNP Q8I5P1
A	184	GLU	-	expression tag	UNP Q8I5P1
A	185	HIS	-	expression tag	UNP Q8I5P1
A	186	HIS	-	expression tag	UNP Q8I5P1
A	187	HIS	-	expression tag	UNP Q8I5P1
A	188	HIS	-	expression tag	UNP Q8I5P1
A	189	HIS	-	expression tag	UNP Q8I5P1
A	190	HIS	-	expression tag	UNP Q8I5P1
B	23	MET	-	initiating methionine	UNP Q8I5P1
B	24	GLY	-	expression tag	UNP Q8I5P1
B	183	LEU	-	expression tag	UNP Q8I5P1
B	184	GLU	-	expression tag	UNP Q8I5P1
B	185	HIS	-	expression tag	UNP Q8I5P1
B	186	HIS	-	expression tag	UNP Q8I5P1
B	187	HIS	-	expression tag	UNP Q8I5P1
B	188	HIS	-	expression tag	UNP Q8I5P1
B	189	HIS	-	expression tag	UNP Q8I5P1
B	190	HIS	-	expression tag	UNP Q8I5P1
C	23	MET	-	initiating methionine	UNP Q8I5P1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	24	GLY	-	expression tag	UNP Q8I5P1
C	183	LEU	-	expression tag	UNP Q8I5P1
C	184	GLU	-	expression tag	UNP Q8I5P1
C	185	HIS	-	expression tag	UNP Q8I5P1
C	186	HIS	-	expression tag	UNP Q8I5P1
C	187	HIS	-	expression tag	UNP Q8I5P1
C	188	HIS	-	expression tag	UNP Q8I5P1
C	189	HIS	-	expression tag	UNP Q8I5P1
C	190	HIS	-	expression tag	UNP Q8I5P1
D	23	MET	-	initiating methionine	UNP Q8I5P1
D	24	GLY	-	expression tag	UNP Q8I5P1
D	183	LEU	-	expression tag	UNP Q8I5P1
D	184	GLU	-	expression tag	UNP Q8I5P1
D	185	HIS	-	expression tag	UNP Q8I5P1
D	186	HIS	-	expression tag	UNP Q8I5P1
D	187	HIS	-	expression tag	UNP Q8I5P1
D	188	HIS	-	expression tag	UNP Q8I5P1
D	189	HIS	-	expression tag	UNP Q8I5P1
D	190	HIS	-	expression tag	UNP Q8I5P1

- Molecule 2 is a protein called 4h12 heavy chain.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
2	H	217	3209	1023	1596	268	317	5	0	0	0
2	I	217	3209	1023	1596	268	317	5	0	0	0
2	J	217	3209	1023	1596	268	317	5	0	0	0
2	K	218	3221	1027	1600	269	320	5	0	0	0

- Molecule 3 is a protein called 4h12 light chain.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
3	L	210	3189	1014	1557	279	332	7	0	0	0
3	M	210	3189	1014	1557	279	332	7	0	0	0
3	N	210	3189	1014	1557	279	332	7	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
3	O	210	3189	1014	1557	279	332	7	0	0	0

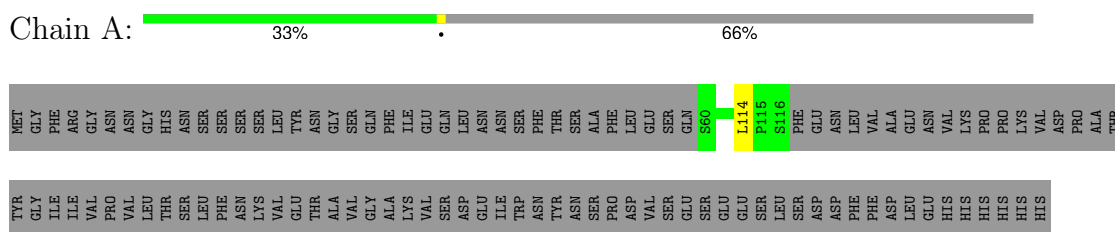
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	C	1	Total 1	O 1	0	0
4	H	1	Total 1	O 1	0	0
4	J	3	Total 3	O 3	0	0
4	K	1	Total 1	O 1	0	0
4	N	3	Total 3	O 3	0	0
4	O	2	Total 2	O 2	0	0

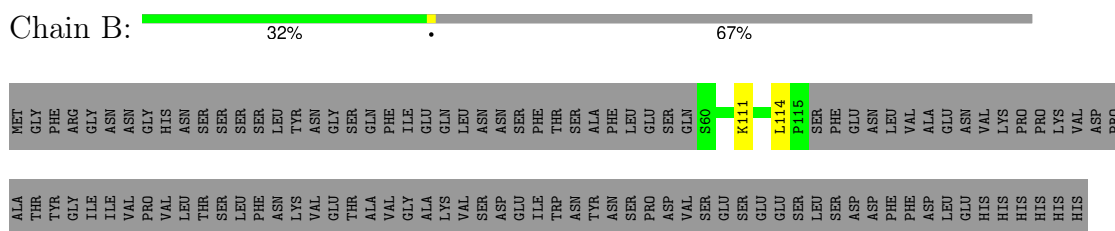
3 Residue-property plots [i](#)

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.

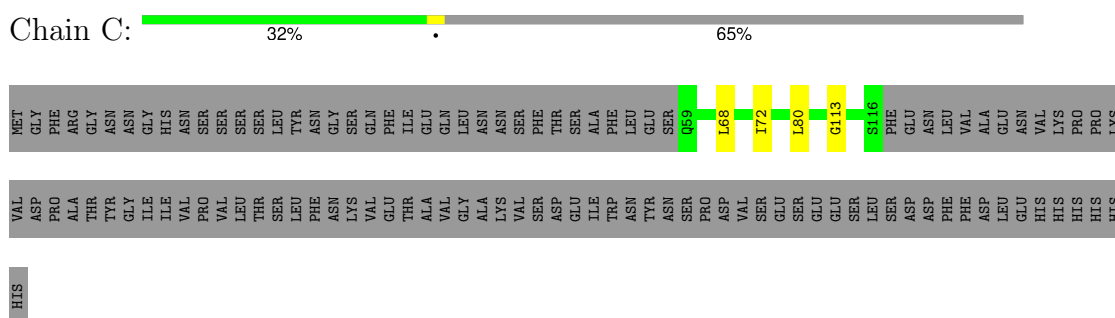
- Molecule 1: Cell traversal protein for ookinetes and sporozoites



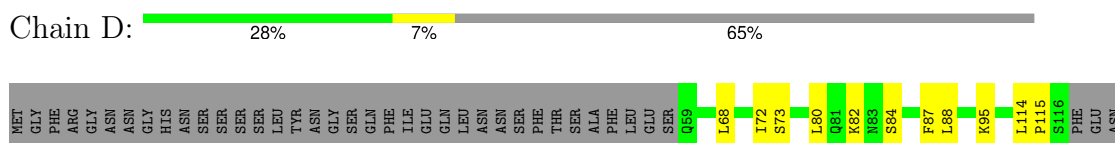
- Molecule 1: Cell traversal protein for ookinetes and sporozoites



- Molecule 1: Cell traversal protein for ookinetes and sporozoites



- Molecule 1: Cell traversal protein for ookinetes and sporozoites



LEU VAL ALA GLU ASN VAL LYS PRO HIS VAL ASP PRO ALA THR GLY ILE ILE VAL PRO VAL THR LEU SER LEU PHE ASN LYS VAL GLU THR ALA VAL LYS VAL SER ASP GLU ILE TRP ASN TYR SER PRO ASP VAL SER GLU GLU SER LEU SER ASP

PHE ASP LEU HIS HIS HIS HIS

- Molecule 2: 4h12 heavy chain

Chain H: 88% 12%



- Molecule 2: 4h12 heavy chain

Chain I: 91% 8%



- Molecule 2: 4h12 heavy chain

Chain J: 94% 6%



- Molecule 2: 4h12 heavy chain

Chain K: 88% 12%



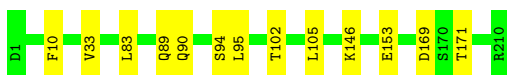
- Molecule 3: 4h12 light chain

Chain L: 92% 8%



- Molecule 3: 4h12 light chain

Chain M: 94% 6%




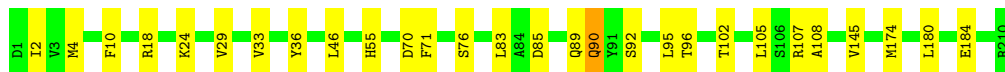
- Molecule 3: 4h12 light chain

Chain N: 89% 10%



- Molecule 3: 4h12 light chain

Chain O:  87% 13%



4 Data and refinement statistics i

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	59.69Å 59.73Å 153.44Å 82.07° 82.20° 72.27°	Depositor
Resolution (Å)	19.86 – 3.52 19.86 – 3.52	Depositor EDS
% Data completeness (in resolution range)	97.1 (19.86-3.52) 97.7 (19.86-3.52)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	5.68 (at 3.52Å)	Xtrriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.268 , 0.293 0.266 , 0.293	Depositor DCC
R_{free} test set	1215 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å ²)	37.0	Xtrriage
Anisotropy	1.391	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.24 , 46.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	0.398 for -k,-h,-l	Xtrriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	29232	wwPDB-VP
Average B, all atoms (Å ²)	54.0	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 44.32 % 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 1.5516e-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	0.25	0/442	0.45	0/591
1	B	0.25	0/436	0.42	0/583
1	C	0.24	0/451	0.46	0/603
1	D	0.24	0/451	0.41	0/603
2	H	0.26	0/1655	0.55	0/2264
2	I	0.26	0/1655	0.53	0/2264
2	J	0.27	0/1655	0.54	0/2264
2	K	0.28	0/1663	0.54	0/2275
3	L	0.25	0/1668	0.52	0/2264
3	M	0.26	0/1668	0.53	0/2264
3	N	0.26	0/1668	0.54	0/2264
3	O	0.26	0/1668	0.53	0/2264
All	All	0.26	0/15080	0.52	0/20503

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	438	461	460	1	0
1	B	432	456	455	1	0
1	C	447	468	468	7	0
1	D	447	468	468	12	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	H	1613	1596	1596	25	0
2	I	1613	1596	1596	13	0
2	J	1613	1596	1596	9	0
2	K	1621	1600	1600	20	0
3	L	1632	1557	1557	13	0
3	M	1632	1557	1557	10	0
3	N	1632	1557	1557	14	0
3	O	1632	1557	1557	18	0
4	C	1	0	0	0	0
4	H	1	0	0	0	0
4	J	3	0	0	0	0
4	K	1	0	0	0	0
4	N	3	0	0	0	0
4	O	2	0	0	0	0
All	All	14763	14469	14467	126	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 4.

The worst 5 of 126 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:H:12:VAL:HG11	2:H:86:LEU:HD12	1.43	1.00
2:H:12:VAL:HG21	2:H:18:VAL:HG13	1.61	0.82
3:N:107:ARG:NH1	3:N:108:ALA:O	2.15	0.80
2:H:12:VAL:CG1	2:H:16:ALA:HB3	2.12	0.79
2:H:198:THR:HG21	2:H:211:ASP:OD1	1.85	0.75

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	55/168 (33%)	52 (94%)	3 (6%)	0	100	100
1	B	54/168 (32%)	53 (98%)	1 (2%)	0	100	100
1	C	56/168 (33%)	53 (95%)	3 (5%)	0	100	100
1	D	56/168 (33%)	53 (95%)	3 (5%)	0	100	100
2	H	215/218 (99%)	206 (96%)	9 (4%)	0	100	100
2	I	215/218 (99%)	213 (99%)	2 (1%)	0	100	100
2	J	215/218 (99%)	207 (96%)	8 (4%)	0	100	100
2	K	216/218 (99%)	204 (94%)	12 (6%)	0	100	100
3	L	208/210 (99%)	200 (96%)	8 (4%)	0	100	100
3	M	208/210 (99%)	202 (97%)	6 (3%)	0	100	100
3	N	208/210 (99%)	200 (96%)	8 (4%)	0	100	100
3	O	208/210 (99%)	196 (94%)	12 (6%)	0	100	100
All	All	1914/2384 (80%)	1839 (96%)	75 (4%)	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	53/153 (35%)	53 (100%)	0	100	100
1	B	52/153 (34%)	52 (100%)	0	100	100
1	C	54/153 (35%)	54 (100%)	0	100	100
1	D	54/153 (35%)	54 (100%)	0	100	100
2	H	183/184 (100%)	183 (100%)	0	100	100
2	I	183/184 (100%)	183 (100%)	0	100	100
2	J	183/184 (100%)	183 (100%)	0	100	100
2	K	184/184 (100%)	184 (100%)	0	100	100
3	L	185/185 (100%)	185 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	M	185/185 (100%)	185 (100%)	0	100	100
3	N	185/185 (100%)	184 (100%)	1 (0%)	86	92
3	O	185/185 (100%)	184 (100%)	1 (0%)	86	92
All	All	1686/2088 (81%)	1684 (100%)	2 (0%)	92	97

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	N	90	GLN
3	O	90	GLN

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	57/168 (33%)	-0.49	0 100 100	61, 69, 84, 86	0
1	B	56/168 (33%)	-0.53	0 100 100	64, 71, 85, 94	0
1	C	58/168 (34%)	-0.51	0 100 100	28, 41, 71, 80	0
1	D	58/168 (34%)	-0.59	0 100 100	26, 40, 68, 74	0
2	H	217/218 (99%)	-0.42	0 100 100	63, 66, 78, 85	0
2	I	217/218 (99%)	-0.37	0 100 100	56, 63, 72, 84	0
2	J	217/218 (99%)	-0.63	0 100 100	29, 34, 53, 72	0
2	K	218/218 (100%)	-0.56	0 100 100	29, 37, 53, 76	0
3	L	210/210 (100%)	-0.42	0 100 100	62, 72, 81, 86	0
3	M	210/210 (100%)	-0.40	0 100 100	58, 67, 76, 82	0
3	N	210/210 (100%)	-0.55	0 100 100	27, 38, 56, 65	0
3	O	210/210 (100%)	-0.57	0 100 100	30, 42, 61, 67	0
All	All	1938/2384 (81%)	-0.50	0 100 100	26, 60, 77, 94	0

There are no RSRZ outliers to report.

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

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.