W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 13, 2017 – 09:18 pm GMT

PDB ID	:	3PMQ
Title	:	Crystal structure of the outer membrane decaheme cytochrome MtrF
Authors	:	Clarke, T.A.; Edwards, M.J.; Richardson, D.J.
Deposited on	:	2010-11-17
Resolution	:	3.20 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
http://wwpdb.org/validation/2016/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
Mogul	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

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 3.20 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.
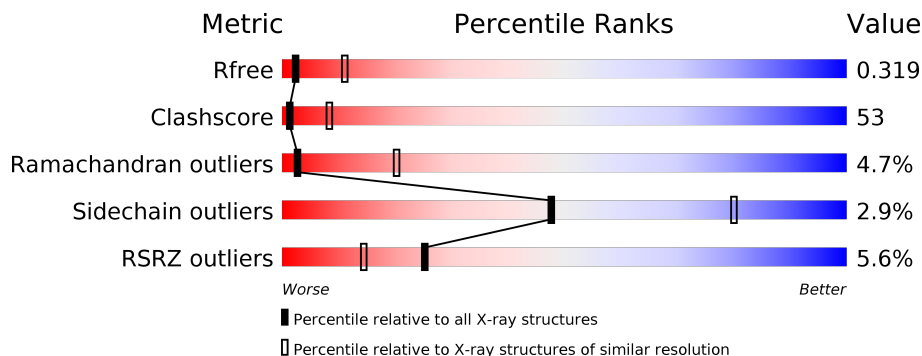
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	100719	1015 (3.22-3.18)
Clashscore	112137	1009 (3.20-3.20)
Ramachandran outliers	110173	1118 (3.22-3.18)
Sidechain outliers	110143	1117 (3.22-3.18)
RSRZ outliers	101464	1020 (3.22-3.18)

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 >=3, 2, 1 and 0 types of geometric quality criteria. 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 <=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	669	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	HEC	A	678	-	-	-	X
2	HEC	A	679	-	-	-	X

2   Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 4826 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 Decaheme cytochrome c MtrF.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	593	4395	2688	791	883	33	0	0	0

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	640	LYS	-	EXPRESSION TAG	UNP Q8EG32
A	641	GLY	-	EXPRESSION TAG	UNP Q8EG32
A	642	GLU	-	EXPRESSION TAG	UNP Q8EG32
A	643	LEU	-	EXPRESSION TAG	UNP Q8EG32
A	644	LYS	-	EXPRESSION TAG	UNP Q8EG32
A	645	LEU	-	EXPRESSION TAG	UNP Q8EG32
A	646	GLU	-	EXPRESSION TAG	UNP Q8EG32
A	647	GLY	-	EXPRESSION TAG	UNP Q8EG32
A	648	LYS	-	EXPRESSION TAG	UNP Q8EG32
A	649	PRO	-	EXPRESSION TAG	UNP Q8EG32
A	650	ILE	-	EXPRESSION TAG	UNP Q8EG32
A	651	PRO	-	EXPRESSION TAG	UNP Q8EG32
A	652	ASN	-	EXPRESSION TAG	UNP Q8EG32
A	653	PRO	-	EXPRESSION TAG	UNP Q8EG32
A	654	LEU	-	EXPRESSION TAG	UNP Q8EG32
A	655	LEU	-	EXPRESSION TAG	UNP Q8EG32
A	656	GLY	-	EXPRESSION TAG	UNP Q8EG32
A	657	LEU	-	EXPRESSION TAG	UNP Q8EG32
A	658	ASP	-	EXPRESSION TAG	UNP Q8EG32
A	659	SER	-	EXPRESSION TAG	UNP Q8EG32
A	660	THR	-	EXPRESSION TAG	UNP Q8EG32
A	661	ARG	-	EXPRESSION TAG	UNP Q8EG32
A	662	THR	-	EXPRESSION TAG	UNP Q8EG32
A	663	GLY	-	EXPRESSION TAG	UNP Q8EG32
A	664	HIS	-	EXPRESSION TAG	UNP Q8EG32
A	665	HIS	-	EXPRESSION TAG	UNP Q8EG32
A	666	HIS	-	EXPRESSION TAG	UNP Q8EG32

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	667	HIS	-	EXPRESSION TAG	UNP Q8EG32
A	668	HIS	-	EXPRESSION TAG	UNP Q8EG32
A	669	HIS	-	EXPRESSION TAG	UNP Q8EG32

• Molecule 2 is HEME C (three-letter code: HEC) (formula: $C_{34}H_{34}FeN_4O_4$).
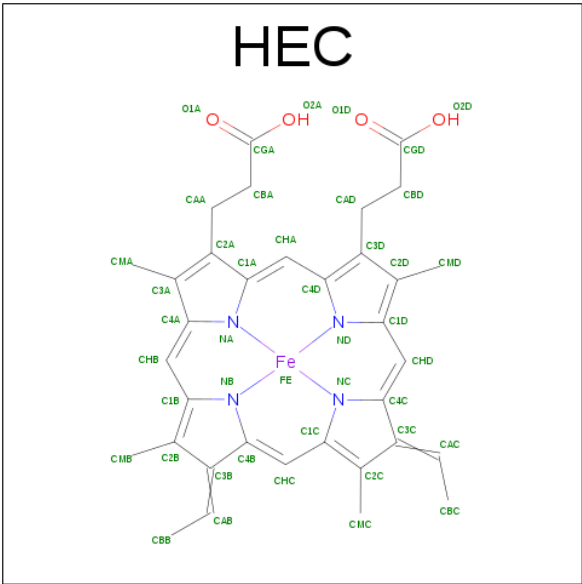
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
2	A	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
2	A	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
2	A	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
2	A	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
2	A	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
2	A	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
2	A	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
2	A	1	Total 43	C 34	Fe 1	N 4	O 4	0	0
2	A	1	Total 43	C 34	Fe 1	N 4	O 4	0	0

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total 1	Ca 1	0	0

3 Residue-property plots ⓘ

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: Decaheme cytochrome c MtrF
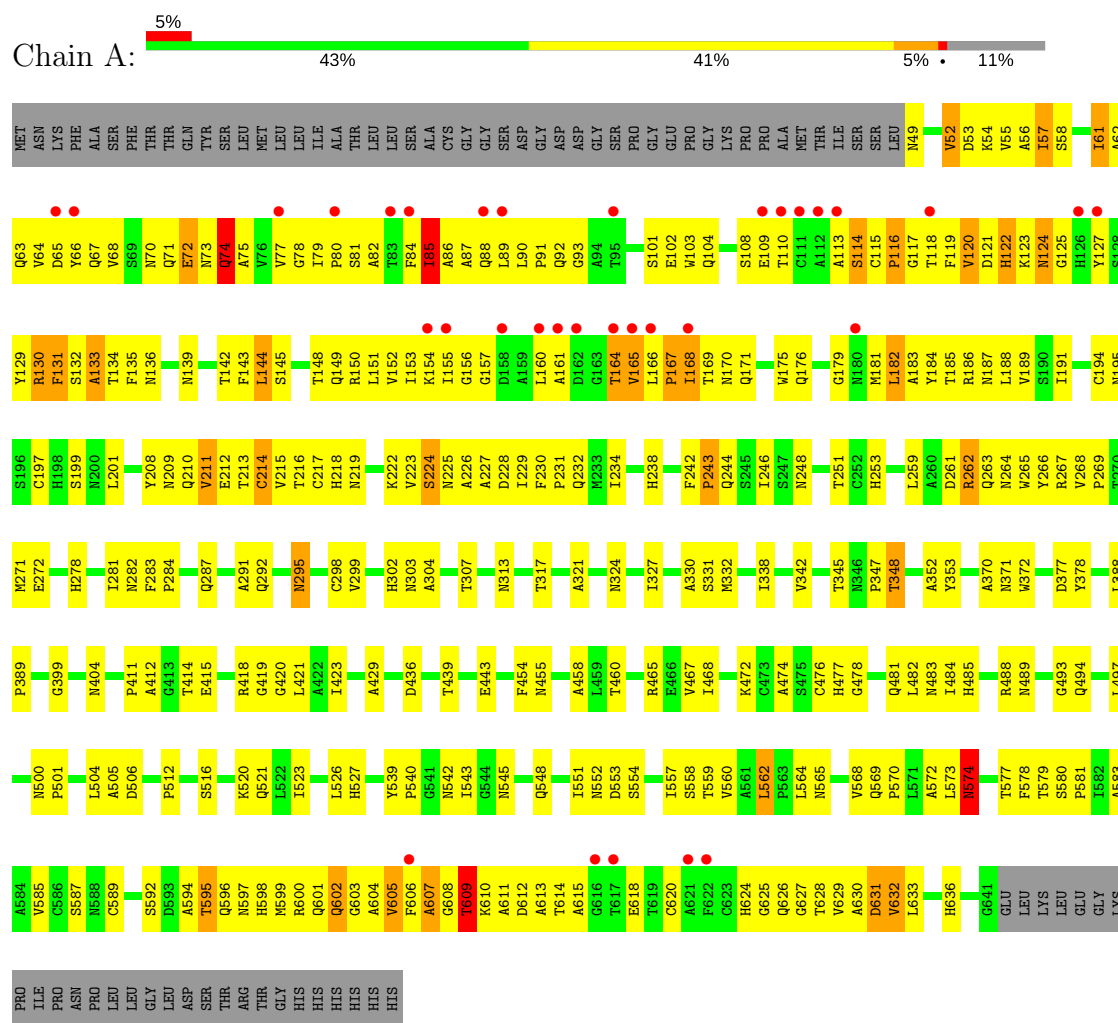
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	I 2 3	Depositor
Cell constants a, b, c, α, β, γ	256.64Å 256.64Å 256.64Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 3.20 57.39 – 3.20	Depositor EDS
% Data completeness (in resolution range)	98.0 (40.00-3.20) 97.7 (57.39-3.20)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.35 (at 3.19Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.6.4_486)	Depositor
R, R_{free}	0.305 , 0.321 0.301 , 0.319	Depositor DCC
R_{free} test set	2295 reflections (5.06%)	DCC
Wilson B-factor (Å^2)	89.6	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 86.9	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.024 for -l,-k,-h	Xtriage
F_o,F_c correlation	0.83	EDS
Total number of atoms	4826	wwPDB-VP
Average B, all atoms (Å^2)	110.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.76% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ 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 ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: CA, HEC

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	0.26	0/4485	0.53	1/6111 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	224	SER	N-CA-C	5.07	124.69	111.00

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	A	4395	0	4133	466	0
2	A	430	0	300	68	0
3	A	1	0	0	0	0
All	All	4826	0	4433	487	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 53.

The worst 5 of 487 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:155:ILE:HG12	1:A:171:GLN:HB3	1.26	1.17
1:A:116:PRO:HB3	1:A:124:ASN:HA	1.29	1.15
1:A:133:ALA:HB1	1:A:134:THR:HB	1.15	1.15
1:A:55:VAL:HG21	1:A:61:ILE:H	1.11	1.15
1:A:527:HIS:HD1	2:A:677:HEC:HBC1	0.98	1.14

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	A	591/669 (88%)	454 (77%)	109 (18%)	28 (5%)	3	20

5 of 28 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	57	ILE
1	A	72	GLU
1	A	114	SER
1	A	120	VAL
1	A	262	ARG

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.

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	479/542 (88%)	465 (97%)	14 (3%)	48	80

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	214	CYS
1	A	295	ASN
1	A	574	ASN
1	A	144	LEU
1	A	562	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	149	GLN
1	A	176	GLN
1	A	280	GLN
1	A	122	HIS
1	A	295	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 11 ligands modelled in this entry, 1 is monoatomic - leaving 10 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$|Z| > 2$	Counts	RMSZ	#$|Z| > 2$
2	HEC	A	670	1	28,50,50	2.29	3 (10%)	16,82,82	1.77	4 (25%)
2	HEC	A	671	1	28,50,50	2.42	4 (14%)	16,82,82	1.65	4 (25%)
2	HEC	A	672	1	28,50,50	2.35	3 (10%)	16,82,82	1.58	3 (18%)
2	HEC	A	673	1	28,50,50	2.37	3 (10%)	16,82,82	1.46	2 (12%)
2	HEC	A	674	1	28,50,50	2.24	3 (10%)	16,82,82	1.51	3 (18%)
2	HEC	A	675	1	28,50,50	2.33	3 (10%)	16,82,82	1.56	2 (12%)
2	HEC	A	676	1	28,50,50	2.29	3 (10%)	16,82,82	1.57	4 (25%)
2	HEC	A	677	1	28,50,50	2.35	3 (10%)	16,82,82	1.60	3 (18%)
2	HEC	A	678	1	28,50,50	2.31	3 (10%)	16,82,82	1.53	3 (18%)
2	HEC	A	679	1	28,50,50	2.28	3 (10%)	16,82,82	1.63	3 (18%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	HEC	A	670	1	-	0/6/54/54	0/0/8/8
2	HEC	A	671	1	-	0/6/54/54	0/0/8/8
2	HEC	A	672	1	-	0/6/54/54	0/0/8/8
2	HEC	A	673	1	-	0/6/54/54	0/0/8/8
2	HEC	A	674	1	-	0/6/54/54	0/0/8/8
2	HEC	A	675	1	-	0/6/54/54	0/0/8/8
2	HEC	A	676	1	-	0/6/54/54	0/0/8/8
2	HEC	A	677	1	-	0/6/54/54	0/0/8/8
2	HEC	A	678	1	-	0/6/54/54	0/0/8/8
2	HEC	A	679	1	-	0/6/54/54	0/0/8/8

The worst 5 of 31 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	671	HEC	C3C-C2C	-6.66	1.33	1.40

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	671	HEC	C3B-C2B	-6.51	1.33	1.40
2	A	677	HEC	C3C-C2C	-6.50	1.33	1.40
2	A	673	HEC	C3B-C2B	-6.45	1.33	1.40
2	A	677	HEC	C3B-C2B	-6.41	1.34	1.40

The worst 5 of 31 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	670	HEC	CMC-C2C-C1C	-4.01	122.31	128.46
2	A	675	HEC	CAD-CBD-CGD	-2.88	107.73	112.66
2	A	677	HEC	CMC-C2C-C1C	-2.83	124.12	128.46
2	A	672	HEC	CBA-CAA-C2A	-2.81	107.10	112.47
2	A	677	HEC	CAD-CBD-CGD	-2.72	108.02	112.66

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

10 monomers are involved in 68 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	670	HEC	4	0
2	A	671	HEC	13	0
2	A	672	HEC	7	0
2	A	673	HEC	2	0
2	A	674	HEC	5	0
2	A	675	HEC	8	0
2	A	676	HEC	6	0
2	A	677	HEC	11	0
2	A	678	HEC	8	0
2	A	679	HEC	5	0

5.7 Other polymers ⓘ

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	593/669 (88%)	0.20	33 (5%)	25	14	46, 99, 206, 272	0

The worst 5 of 33 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	126	HIS	6.8
1	A	162	ASP	4.7
1	A	84	PHE	4.6
1	A	161	ALA	4.1
1	A	111	CYS	4.1

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 carbohydrates in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
2	HEC	A	678	43/43	0.95	0.41	3.11	114,162,187,202	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
2	HEC	A	679	43/43	0.89	0.38	<mark>3.09</mark>	106,176,195,203	0
2	HEC	A	670	43/43	0.97	0.30	1.55	55,90,118,129	0
2	HEC	A	677	43/43	0.97	0.28	1.24	73,103,115,122	0
2	HEC	A	672	43/43	0.98	0.29	1.18	40,69,89,104	0
2	HEC	A	673	43/43	0.97	0.25	0.89	29,61,98,106	0
2	HEC	A	676	43/43	0.96	0.26	0.34	39,73,127,151	0
2	HEC	A	674	43/43	0.98	0.22	0.32	31,52,83,105	0
2	HEC	A	675	43/43	0.98	0.24	0.03	53,83,105,143	0
2	HEC	A	671	43/43	0.96	0.25	-0.19	63,86,140,148	0
3	CA	A	680	1/1	0.90	0.14	-1.14	101,101,101,101	0

6.5 Other polymers ⓘ

There are no such residues in this entry.