



wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 26, 2024 – 08:20 AM EDT

PDB ID : 6S1C
Title : P3221 crystal form of the Ctf18-1-8/Pol2(1-528) complex
Authors : Grabarczyk, D.B.
Deposited on : 2019-06-18
Resolution : 6.10 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) : 1.13
EDS : 2.37.1
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) : 2.37.1

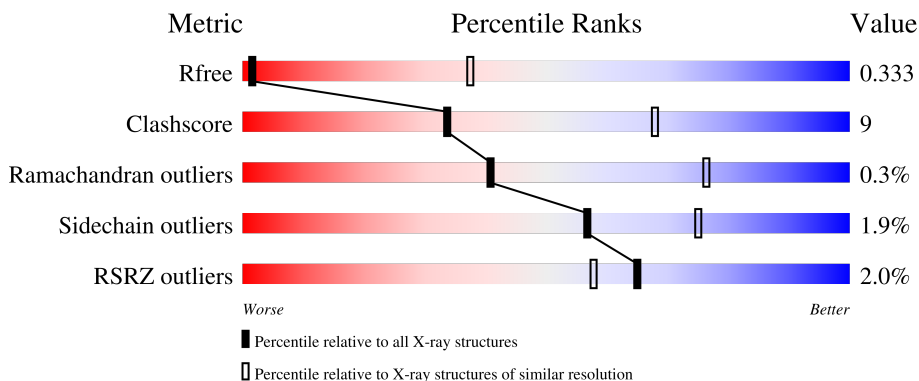
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 6.10 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.



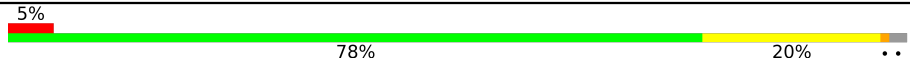


Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1002 (8.30-3.88)
Clashscore	141614	1051 (8.30-3.90)
Ramachandran outliers	138981	1018 (8.30-3.86)
Sidechain outliers	138945	1019 (8.30-3.82)
RSRZ outliers	127900	1015 (8.20-3.78)

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	549	 2% 65% 12% 21%
1	E	549	 2% 62% 11% 27%
2	B	380	 2% 66% 23% 10%
2	F	380	 2% 57% 20% 21%
3	C	133	 4% 68% 28% 2%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	G	133	
4	D	33	
4	H	33	

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 14378 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 DNA polymerase epsilon catalytic subunit A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	426	Total 3465	C 2232	N 555	O 659	S 19	0	0	0
1	E	400	Total 3244	C 2094	N 516	O 616	S 18	0	0	0

There are 54 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-24	MET	-	initiating methionine	UNP P21951
A	-23	LYS	-	expression tag	UNP P21951
A	-22	HIS	-	expression tag	UNP P21951
A	-21	HIS	-	expression tag	UNP P21951
A	-20	HIS	-	expression tag	UNP P21951
A	-19	HIS	-	expression tag	UNP P21951
A	-18	HIS	-	expression tag	UNP P21951
A	-17	HIS	-	expression tag	UNP P21951
A	-16	SER	-	expression tag	UNP P21951
A	-15	ALA	-	expression tag	UNP P21951
A	-14	GLY	-	expression tag	UNP P21951
A	-13	LEU	-	expression tag	UNP P21951
A	-12	GLU	-	expression tag	UNP P21951
A	-11	VAL	-	expression tag	UNP P21951
A	-10	LEU	-	expression tag	UNP P21951
A	-9	PHE	-	expression tag	UNP P21951
A	-8	GLN	-	expression tag	UNP P21951
A	-7	GLY	-	expression tag	UNP P21951
A	-6	PRO	-	expression tag	UNP P21951
A	-5	GLY	-	expression tag	UNP P21951
A	-4	THR	-	expression tag	UNP P21951
A	-3	GLY	-	expression tag	UNP P21951
A	-2	SER	-	expression tag	UNP P21951
A	-1	GLU	-	expression tag	UNP P21951
A	0	PHE	-	expression tag	UNP P21951

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	290	ALA	ASP	engineered mutation	UNP P21951
A	292	ALA	GLU	engineered mutation	UNP P21951
E	-24	MET	-	initiating methionine	UNP P21951
E	-23	LYS	-	expression tag	UNP P21951
E	-22	HIS	-	expression tag	UNP P21951
E	-21	HIS	-	expression tag	UNP P21951
E	-20	HIS	-	expression tag	UNP P21951
E	-19	HIS	-	expression tag	UNP P21951
E	-18	HIS	-	expression tag	UNP P21951
E	-17	HIS	-	expression tag	UNP P21951
E	-16	SER	-	expression tag	UNP P21951
E	-15	ALA	-	expression tag	UNP P21951
E	-14	GLY	-	expression tag	UNP P21951
E	-13	LEU	-	expression tag	UNP P21951
E	-12	GLU	-	expression tag	UNP P21951
E	-11	VAL	-	expression tag	UNP P21951
E	-10	LEU	-	expression tag	UNP P21951
E	-9	PHE	-	expression tag	UNP P21951
E	-8	GLN	-	expression tag	UNP P21951
E	-7	GLY	-	expression tag	UNP P21951
E	-6	PRO	-	expression tag	UNP P21951
E	-5	GLY	-	expression tag	UNP P21951
E	-4	THR	-	expression tag	UNP P21951
E	-3	GLY	-	expression tag	UNP P21951
E	-2	SER	-	expression tag	UNP P21951
E	-1	GLU	-	expression tag	UNP P21951
E	0	PHE	-	expression tag	UNP P21951
E	290	ALA	ASP	engineered mutation	UNP P21951
E	292	ALA	GLU	engineered mutation	UNP P21951

- Molecule 2 is a protein called Sister chromatid cohesion protein DCC1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	342	2782	1798	464	507	13	0	0	0
2	F	299	2423	1571	401	438	13	0	0	0

- Molecule 3 is a protein called Chromosome transmission fidelity protein 8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	131	1020	646	177	191	6	0	0	0
3	G	131	1020	646	177	191	6	0	0	0

- Molecule 4 is a protein called Chromosome transmission fidelity protein 18.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
4	D	24	212	138	37	37	0	0	0
4	H	24	212	138	37	37	0	0	0

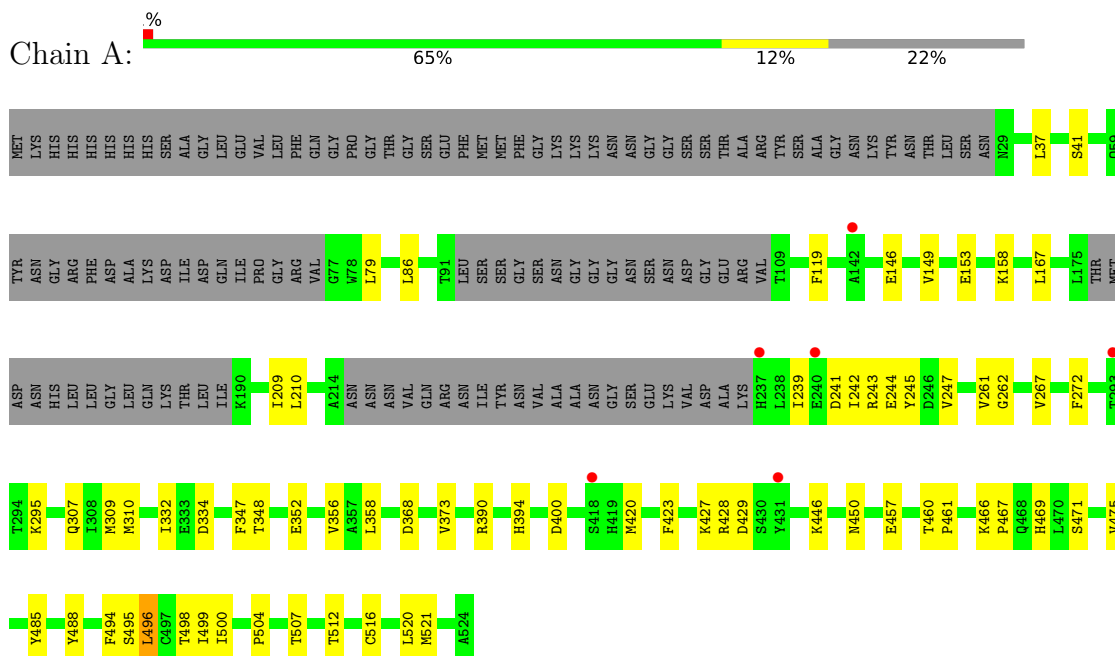
There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	709	GLY	-	expression tag	UNP P49956
D	710	ALA	-	expression tag	UNP P49956
D	711	MET	-	expression tag	UNP P49956
D	712	GLY	-	expression tag	UNP P49956
H	709	GLY	-	expression tag	UNP P49956
H	710	ALA	-	expression tag	UNP P49956
H	711	MET	-	expression tag	UNP P49956
H	712	GLY	-	expression tag	UNP P49956

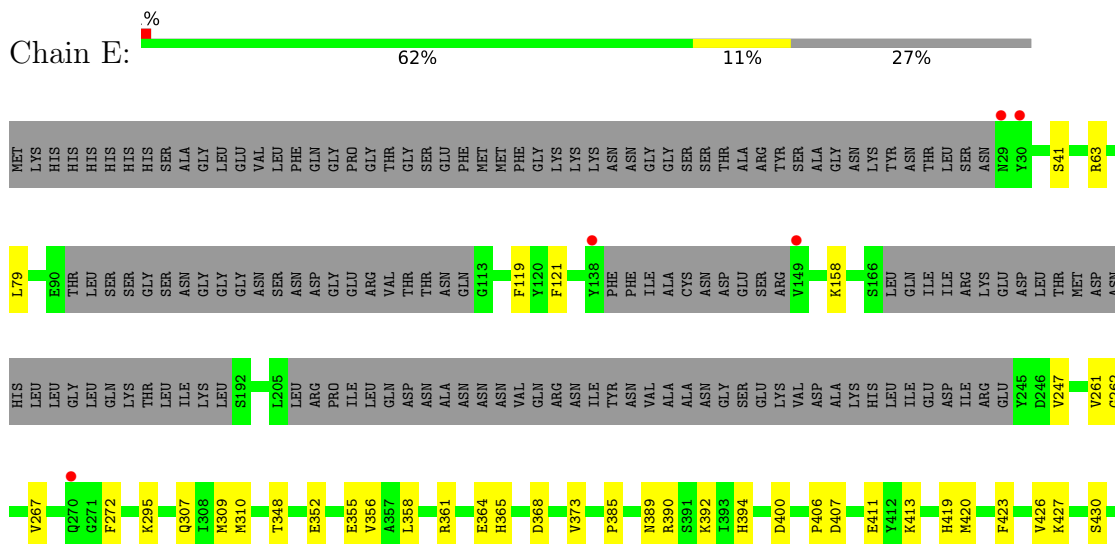
3 Residue-property plots

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: DNA polymerase epsilon catalytic subunit A

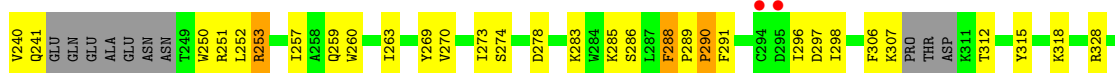


- Molecule 1: DNA polymerase epsilon catalytic subunit A

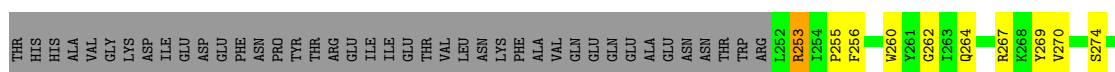
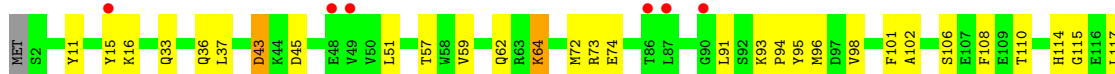




- Molecule 2: Sister chromatid cohesion protein DCC1




- Molecule 2: Sister chromatid cohesion protein DCC1

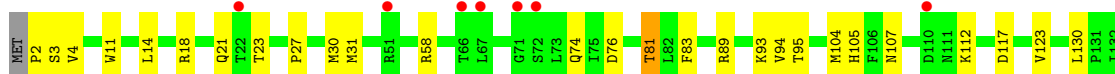


- Molecule 3: Chromosome transmission fidelity protein 8



- Molecule 3: Chromosome transmission fidelity protein 8

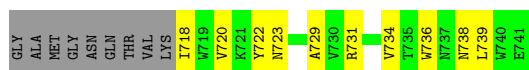
Chain G:  5% 78% 20% ..



MET

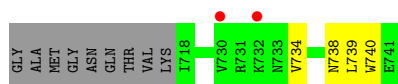
- Molecule 4: Chromosome transmission fidelity protein 18

Chain D:  42% 30% 27%



- Molecule 4: Chromosome transmission fidelity protein 18

Chain H:  6% 61% 12% 27%



4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α , β , γ	126.33Å 126.33Å 378.43Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.97 – 6.10 19.97 – 6.10	Depositor EDS
% Data completeness (in resolution range)	99.9 (19.97-6.10) 100.0 (19.97-6.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.24 (at 5.92Å)	Xtrriage
Refinement program	PHENIX 1.14 3260	Depositor
R, R_{free}	0.275 , 0.333 0.276 , 0.333	Depositor DCC
R_{free} test set	841 reflections (9.78%)	wwPDB-VP
Wilson B-factor (Å ²)	419.0	Xtrriage
Anisotropy	0.128	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.17 , 326.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.42$, $\langle L^2 \rangle = 0.25$	Xtrriage
Estimated twinning fraction	0.065 for -h,-k,l	Xtrriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	14378	wwPDB-VP
Average B, all atoms (Å ²)	511.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.18% 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	0.27	0/3551	0.49	0/4809
1	E	0.26	0/3329	0.48	0/4509
2	B	0.25	0/2845	0.43	0/3843
2	F	0.27	0/2477	0.45	0/3341
3	C	0.28	0/1036	0.56	1/1395 (0.1%)
3	G	0.26	0/1036	0.50	0/1395
4	D	0.35	0/219	0.56	0/297
4	H	0.29	0/219	0.48	0/297
All	All	0.27	0/14712	0.48	1/19886 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	97	LEU	CA-CB-CG	-6.12	101.23	115.30

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	3465	0	3353	43	0
1	E	3244	0	3120	40	0
2	B	2782	0	2784	73	0
2	F	2423	0	2446	66	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	1020	0	1031	36	0
3	G	1020	0	1031	17	0
4	D	212	0	197	13	0
4	H	212	0	197	3	0
All	All	14378	0	14159	243	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 9.

The worst 5 of 243 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:B:11:TYR:HB3	2:B:102:ALA:HA	1.65	0.79
1:A:309:MET:HG3	1:A:310:MET:HG3	1.63	0.78
3:C:93:LYS:HE2	3:C:95:THR:HB	1.65	0.76
3:C:97:LEU:HD11	3:C:122:LYS:HD3	1.67	0.76
1:E:368:ASP:OD2	2:F:367:ARG:NH2	2.18	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	416/549 (76%)	401 (96%)	15 (4%)	0	100	100
1	E	390/549 (71%)	380 (97%)	10 (3%)	0	100	100
2	B	332/380 (87%)	303 (91%)	26 (8%)	3 (1%)	17	56
2	F	289/380 (76%)	262 (91%)	24 (8%)	3 (1%)	15	54
3	C	129/133 (97%)	125 (97%)	4 (3%)	0	100	100
3	G	129/133 (97%)	128 (99%)	1 (1%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	D	22/33 (67%)	22 (100%)	0	0	100	100
4	H	22/33 (67%)	22 (100%)	0	0	100	100
All	All	1729/2190 (79%)	1643 (95%)	80 (5%)	6 (0%)	41	76

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	290	PRO
2	F	290	PRO
2	B	364	LYS
2	F	364	LYS
2	B	45	ASP

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	385/484 (80%)	380 (99%)	5 (1%)	69	82
1	E	358/484 (74%)	352 (98%)	6 (2%)	60	78
2	B	311/352 (88%)	308 (99%)	3 (1%)	76	86
2	F	272/352 (77%)	265 (97%)	7 (3%)	46	66
3	C	112/120 (93%)	106 (95%)	6 (5%)	22	47
3	G	112/120 (93%)	108 (96%)	4 (4%)	35	59
4	D	22/28 (79%)	22 (100%)	0	100	100
4	H	22/28 (79%)	22 (100%)	0	100	100
All	All	1594/1968 (81%)	1563 (98%)	31 (2%)	57	75

5 of 31 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	23	THR
3	G	23	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	C	76	ASP
3	G	76	ASP
2	F	253	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	83	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 [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	426/549 (77%)	-0.25	6 (1%) 75 66	329, 448, 605, 684	0
1	E	400/549 (72%)	-0.17	5 (1%) 77 68	388, 490, 597, 736	0
2	B	342/380 (90%)	-0.18	2 (0%) 89 84	386, 511, 600, 748	0
2	F	299/380 (78%)	0.10	8 (2%) 54 47	421, 572, 675, 745	0
3	C	131/133 (98%)	0.15	5 (3%) 40 36	381, 482, 578, 684	0
3	G	131/133 (98%)	0.36	7 (5%) 26 26	480, 577, 637, 700	0
4	D	24/33 (72%)	-0.50	0 100 100	391, 483, 534, 590	0
4	H	24/33 (72%)	-0.09	2 (8%) 11 13	502, 541, 583, 630	0
All	All	1777/2190 (81%)	-0.09	35 (1%) 65 58	329, 506, 630, 748	0

The worst 5 of 35 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	G	66	THR	4.7
2	B	295	ASP	4.5
1	E	29	ASN	4.5
3	G	71	GLY	3.6
3	C	98	ASP	3.4

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

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.