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

wwPDB X-ray Structure Validation Summary Report (i)

Feb 21, 2024 – 07:06 AM EST

|		
PDB ID	:	4NL8
Title	:	PriA Helicase Bound to SSB C-terminal Tail Peptide
Authors	:	Bhattacharyya, B.; George, N.P.; Thurmes, T.M.; Keck, J.L.
Deposited on	:	2013-11-13
Resolution	:	4.08 Å(reported)

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 4.08 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.
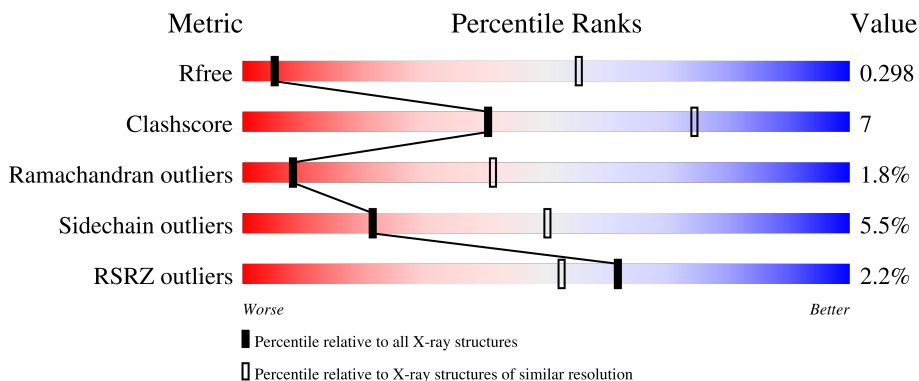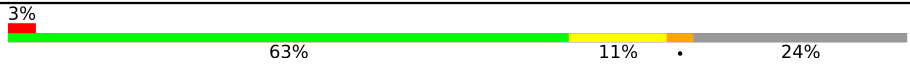

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1133 (4.46-3.70)
Clashscore	141614	1013 (4.42-3.74)
Ramachandran outliers	138981	1151 (4.46-3.70)
Sidechain outliers	138945	1139 (4.46-3.70)
RSRZ outliers	127900	1012 (4.48-3.68)

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 <=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	C	10	20% 10% 70%
1	D	10	10% 40% 50%
1	F	10	10% 20% 70%
2	A	747	2% 62% 10% · 27%
2	B	747	63% 10% · 24%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	E	747	3% 63% 11% • 24%

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 13380 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 Single-stranded DNA-binding protein.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	C	3	Total 27	C 20	N 3	O 4	0	0	0
1	D	5	Total 43	C 28	N 5	O 10	0	0	0
1	F	3	Total 27	C 20	N 3	O 4	0	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	168	TRP	-	expression tag	UNP G8W6B0
D	168	TRP	-	expression tag	UNP G8W6B0
F	168	TRP	-	expression tag	UNP G8W6B0

- Molecule 2 is a protein called Primosome assembly protein PriA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	549	Total 4335	C 2760	N 788	O 772	S 15	0	1	0
2	B	566	Total 4449	C 2842	N 805	O 787	S 15	0	1	0
2	E	571	Total 4493	C 2870	N 814	O 794	S 15	0	1	0

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-15	HIS	-	expression tag	UNP A6TGC5
A	-14	HIS	-	expression tag	UNP A6TGC5
A	-13	HIS	-	expression tag	UNP A6TGC5
A	-12	HIS	-	expression tag	UNP A6TGC5
A	-11	HIS	-	expression tag	UNP A6TGC5

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-10	HIS	-	expression tag	UNP A6TGC5
A	-9	SER	-	expression tag	UNP A6TGC5
A	-8	SER	-	expression tag	UNP A6TGC5
A	-7	GLY	-	expression tag	UNP A6TGC5
A	-6	LEU	-	expression tag	UNP A6TGC5
A	-5	VAL	-	expression tag	UNP A6TGC5
A	-4	PRO	-	expression tag	UNP A6TGC5
A	-3	ARG	-	expression tag	UNP A6TGC5
A	-2	GLY	-	expression tag	UNP A6TGC5
A	-1	SER	-	expression tag	UNP A6TGC5
A	0	HIS	-	expression tag	UNP A6TGC5
B	-15	HIS	-	expression tag	UNP A6TGC5
B	-14	HIS	-	expression tag	UNP A6TGC5
B	-13	HIS	-	expression tag	UNP A6TGC5
B	-12	HIS	-	expression tag	UNP A6TGC5
B	-11	HIS	-	expression tag	UNP A6TGC5
B	-10	HIS	-	expression tag	UNP A6TGC5
B	-9	SER	-	expression tag	UNP A6TGC5
B	-8	SER	-	expression tag	UNP A6TGC5
B	-7	GLY	-	expression tag	UNP A6TGC5
B	-6	LEU	-	expression tag	UNP A6TGC5
B	-5	VAL	-	expression tag	UNP A6TGC5
B	-4	PRO	-	expression tag	UNP A6TGC5
B	-3	ARG	-	expression tag	UNP A6TGC5
B	-2	GLY	-	expression tag	UNP A6TGC5
B	-1	SER	-	expression tag	UNP A6TGC5
B	0	HIS	-	expression tag	UNP A6TGC5
E	-15	HIS	-	expression tag	UNP A6TGC5
E	-14	HIS	-	expression tag	UNP A6TGC5
E	-13	HIS	-	expression tag	UNP A6TGC5
E	-12	HIS	-	expression tag	UNP A6TGC5
E	-11	HIS	-	expression tag	UNP A6TGC5
E	-10	HIS	-	expression tag	UNP A6TGC5
E	-9	SER	-	expression tag	UNP A6TGC5
E	-8	SER	-	expression tag	UNP A6TGC5
E	-7	GLY	-	expression tag	UNP A6TGC5
E	-6	LEU	-	expression tag	UNP A6TGC5
E	-5	VAL	-	expression tag	UNP A6TGC5
E	-4	PRO	-	expression tag	UNP A6TGC5
E	-3	ARG	-	expression tag	UNP A6TGC5
E	-2	GLY	-	expression tag	UNP A6TGC5
E	-1	SER	-	expression tag	UNP A6TGC5

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	0	HIS	-	expression tag	UNP A6TGC5

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	2	Total 2	Zn 2	0	0
3	B	2	Total 2	Zn 2	0	0
3	E	2	Total 2	Zn 2	0	0

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.
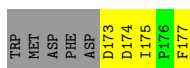
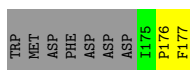- Molecule 1: Single-stranded DNA-binding protein

Chain C:
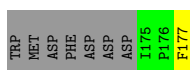
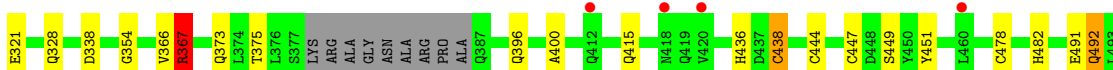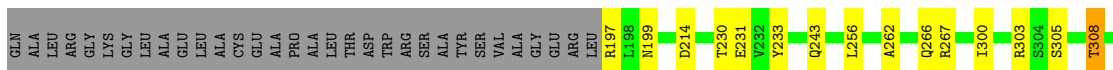- Molecule 1: Single-stranded DNA-binding protein
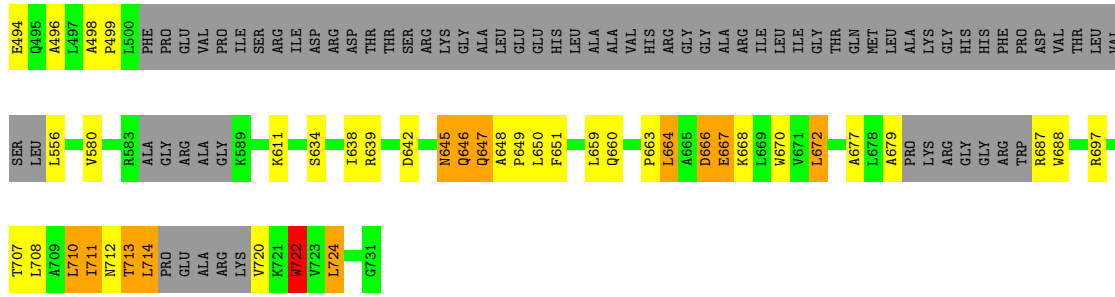
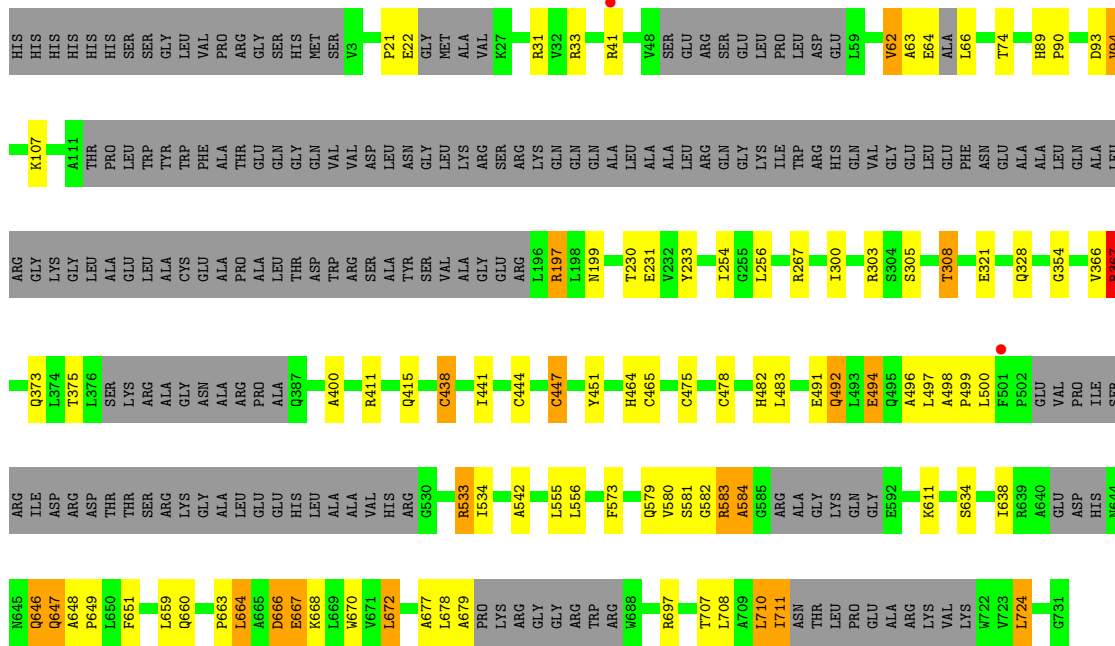Chain D:

- Molecule 1: Single-stranded DNA-binding protein

Chain F:

- Molecule 2: Primosome assembly protein PriA

Chain A:

• Molecule 2: Primosome assembly protein PriA
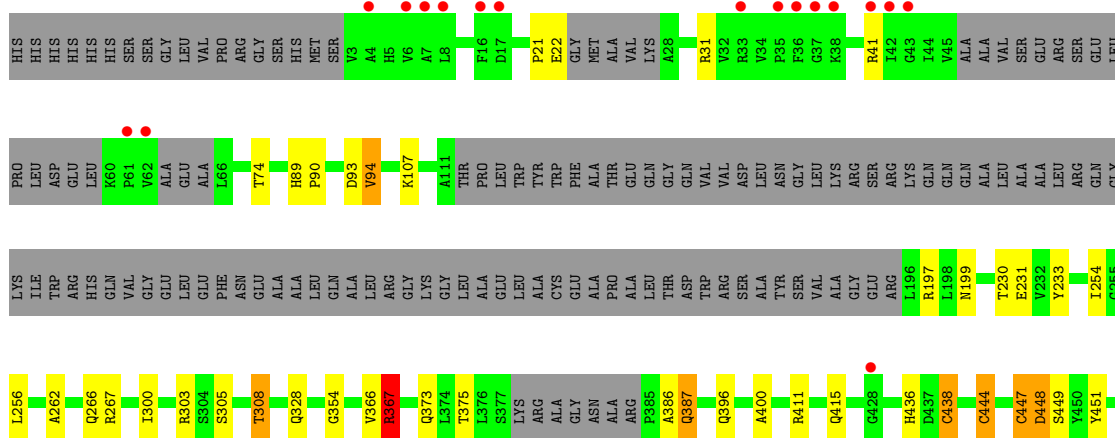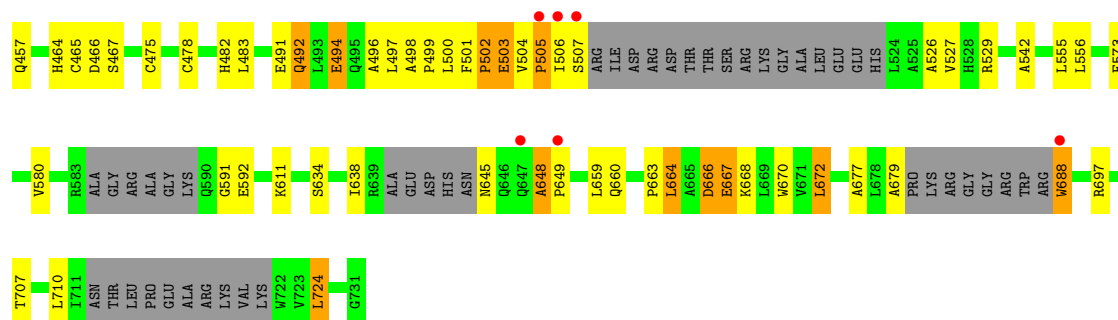
Chain B:

• Molecule 2: Primosome assembly protein PriA

Chain E:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	106.61Å 153.25Å 192.37Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 4.08 48.71 – 4.08	Depositor EDS
% Data completeness (in resolution range)	84.2 (50.00-4.08) 84.3 (48.71-4.08)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.11	Depositor
$< I/\sigma(I) >$ [1]	2.82 (at 4.14Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.240 , 0.298 0.249 , 0.298	Depositor DCC
R_{free} test set	1103 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å^2)	124.8	Xtriage
Anisotropy	0.301	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.26 , 102.8	EDS
L-test for twinning[2]	$< |L| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.87	EDS
Total number of atoms	13380	wwPDB-VP
Average B, all atoms (Å^2)	143.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.09% 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:
ZN

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	C	0.78	0/28	0.67	0/36
1	D	0.77	0/44	0.80	0/58
1	F	0.64	0/28	0.50	0/36
2	A	0.48	1/4439 (0.0%)	0.69	2/6035 (0.0%)
2	B	0.46	0/4557	0.68	1/6197 (0.0%)
2	E	0.45	0/4604	0.66	1/6262 (0.0%)
All	All	0.46	1/13700 (0.0%)	0.68	4/18624 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	338	ASP	CB-CG	6.75	1.66	1.51

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	338	ASP	CB-CG-OD2	6.20	123.88	118.30
2	E	367	ARG	NE-CZ-NH1	5.73	123.17	120.30
2	A	367	ARG	NE-CZ-NH1	5.58	123.09	120.30
2	B	367	ARG	NE-CZ-NH1	5.45	123.02	120.30

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	C	27	0	26	2	0
1	D	43	0	34	0	0
1	F	27	0	26	1	0
2	A	4335	0	4334	60	0
2	B	4449	0	4452	62	0
2	E	4493	0	4496	71	0
3	A	2	0	0	0	0
3	B	2	0	0	0	0
3	E	2	0	0	0	0
All	All	13380	0	13368	192	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 7.

The worst 5 of 192 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:62:VAL:HG12	2:B:63:ALA:H	1.32	0.93
2:E:494:GLU:HG3	2:E:505:PRO:HG3	1.50	0.92
2:B:63:ALA:O	2:B:64:GLU:HG3	1.69	0.92
2:B:708:LEU:O	2:B:711:ILE:HD11	1.71	0.88
2:E:502:PRO:O	2:E:503:GLU:HG3	1.74	0.88

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	C	1/10 (10%)	1 (100%)	0	0	100	100
1	D	3/10 (30%)	0	1 (33%)	2 (67%)	0	0

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	F	1/10 (10%)	0	1 (100%)	0	100	100
2	A	530/747 (71%)	493 (93%)	29 (6%)	8 (2%)	10	45
2	B	545/747 (73%)	504 (92%)	32 (6%)	9 (2%)	9	43
2	E	550/747 (74%)	512 (93%)	27 (5%)	11 (2%)	7	40
All	All	1630/2271 (72%)	1510 (93%)	90 (6%)	30 (2%)	8	41

5 of 30 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	664	LEU
2	A	667	GLU
2	A	713	THR
1	D	174	ASP
2	B	62	VAL

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.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	3/10 (30%)	3 (100%)	0	100	100
1	D	5/10 (50%)	3 (60%)	2 (40%)	0	0
1	F	3/10 (30%)	3 (100%)	0	100	100
2	A	464/614 (76%)	437 (94%)	27 (6%)	20	48
2	B	473/614 (77%)	448 (95%)	25 (5%)	22	51
2	E	480/614 (78%)	456 (95%)	24 (5%)	24	52
All	All	1428/1872 (76%)	1350 (94%)	78 (6%)	21	50

5 of 78 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	E	199	ASN
2	E	645	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	E	308	THR
2	E	447	CYS
2	E	688	TRP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	373	GLN
2	B	647	GLN
2	E	495	GLN
2	B	495	GLN
2	B	657	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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 6 ligands modelled in this entry, 6 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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 ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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(Å²)	Q<0.9
1	C	3/10 (30%)	-0.32	0	100	100	119, 119, 160, 167	0
1	D	5/10 (50%)	0.41	0	100	100	169, 192, 198, 199	0
1	F	3/10 (30%)	-0.50	0	100	100	148, 148, 187, 208	0
2	A	549/747 (73%)	-0.05	13 (2%)	59	49	80, 129, 197, 257	0
2	B	566/747 (75%)	-0.13	2 (0%)	92	87	81, 134, 206, 266	0
2	E	571/747 (76%)	0.09	23 (4%)	38	30	92, 146, 224, 285	0
All	All	1697/2271 (74%)	-0.03	38 (2%)	62	52	80, 137, 210, 285	0

The worst 5 of 38 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	A	59	LEU	4.9
2	E	38	LYS	4.8
2	E	505	PRO	4.5
2	E	506	ILE	4.2
2	E	8	LEU	4.1

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

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

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

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. 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	B-factors(Å^2)	Q<0.9
3	ZN	A	801	1/1	0.99	0.13	103,103,103,103	0
3	ZN	A	802	1/1	0.99	0.06	129,129,129,129	0
3	ZN	B	802	1/1	0.99	0.11	118,118,118,118	0
3	ZN	E	801	1/1	0.99	0.16	117,117,117,117	0
3	ZN	E	802	1/1	0.99	0.15	102,102,102,102	0
3	ZN	B	801	1/1	1.00	0.12	100,100,100,100	0

6.5 Other polymers ⓘ

There are no such residues in this entry.