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 (i)

Feb 22, 2024 – 03:16 AM EST

PDB ID	:	4RIA
Title	:	FAN1 Nuclease bound to 5' phosphorylated nicked DNA
Authors	:	Pavletich, N.P.; Wang, R.
Deposited on	:	2014-10-05
Resolution	:	3.00 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

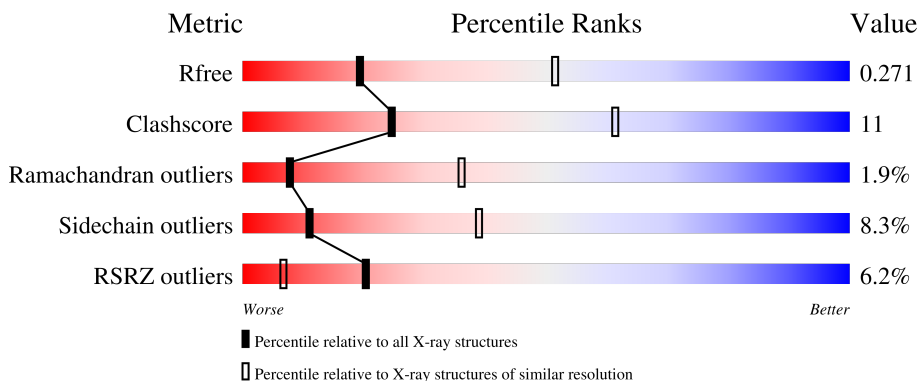|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 3.00 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	Percentile Ranks	Value
Rfree		0.271
Clashscore		11
Ramachandran outliers		1.9%
Sidechain outliers		8.3%
RSRZ outliers		6.2%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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 <=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	651	6% / 69% 21% · 6%
1	B	651	6% / 68% 23% · 6%
2	F	10	50% 50%
2	L	10	50% 50%
3	I	9	67% 33%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	K	9	67% 33%
4	G	14	29% 21% 50%
4	J	14	29% 21% 50%
5	E	12	58% 42%
5	H	12	58% 42%

2 Entry composition ⓘ

There are 6 unique types of molecules in this entry. The entry contains 11410 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 Fanconi-associated nuclease 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	613	Total 4927	C 3136	N 877	O 887	S 27	0	0	0
1	B	613	Total 4927	C 3136	N 877	O 887	S 27	0	0	0

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	358	GLY	-	expression tag	UNP Q9Y2M0
A	359	ALA	-	expression tag	UNP Q9Y2M0
A	360	HIS	-	expression tag	UNP Q9Y2M0
A	361	MET	-	expression tag	UNP Q9Y2M0
A	362	THR	-	expression tag	UNP Q9Y2M0
A	363	ARG	-	expression tag	UNP Q9Y2M0
A	364	ASN	-	expression tag	UNP Q9Y2M0
A	365	GLY	-	expression tag	UNP Q9Y2M0
A	366	PRO	-	expression tag	UNP Q9Y2M0
A	367	GLY	-	expression tag	UNP Q9Y2M0
A	368	GLN	-	expression tag	UNP Q9Y2M0
A	369	THR	-	expression tag	UNP Q9Y2M0
A	487	ALA	VAL	engineered mutation	UNP Q9Y2M0
A	?	-	CYS	deletion	UNP Q9Y2M0
A	?	-	THR	deletion	UNP Q9Y2M0
A	?	-	TRP	deletion	UNP Q9Y2M0
A	?	-	GLY	deletion	UNP Q9Y2M0
A	?	-	LYS	deletion	UNP Q9Y2M0
A	?	-	ASN	deletion	UNP Q9Y2M0
A	?	-	LYS	deletion	UNP Q9Y2M0
A	?	-	PRO	deletion	UNP Q9Y2M0
A	?	-	GLY	deletion	UNP Q9Y2M0
B	358	GLY	-	expression tag	UNP Q9Y2M0
B	359	ALA	-	expression tag	UNP Q9Y2M0
B	360	HIS	-	expression tag	UNP Q9Y2M0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	361	MET	-	expression tag	UNP Q9Y2M0
B	362	THR	-	expression tag	UNP Q9Y2M0
B	363	ARG	-	expression tag	UNP Q9Y2M0
B	364	ASN	-	expression tag	UNP Q9Y2M0
B	365	GLY	-	expression tag	UNP Q9Y2M0
B	366	PRO	-	expression tag	UNP Q9Y2M0
B	367	GLY	-	expression tag	UNP Q9Y2M0
B	368	GLN	-	expression tag	UNP Q9Y2M0
B	369	THR	-	expression tag	UNP Q9Y2M0
B	487	ALA	VAL	engineered mutation	UNP Q9Y2M0
B	?	-	CYS	deletion	UNP Q9Y2M0
B	?	-	THR	deletion	UNP Q9Y2M0
B	?	-	TRP	deletion	UNP Q9Y2M0
B	?	-	GLY	deletion	UNP Q9Y2M0
B	?	-	LYS	deletion	UNP Q9Y2M0
B	?	-	ASN	deletion	UNP Q9Y2M0
B	?	-	LYS	deletion	UNP Q9Y2M0
B	?	-	PRO	deletion	UNP Q9Y2M0
B	?	-	GLY	deletion	UNP Q9Y2M0

- Molecule 2 is a DNA chain called DNA (5'-D(P*AP*GP*CP*CP*AP*CP*GP*CP*CP*T)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	10	Total 202	C 95	N 37	O 60	P 10	0	0	0
2	L	10	Total 202	C 95	N 37	O 60	P 10	0	0	0

- Molecule 3 is a DNA chain called DNA (5'-D(P*AP*GP*AP*CP*TP*CP*CP*TP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	K	9	Total 180	C 86	N 31	O 54	P 9	0	0	0
3	I	9	Total 180	C 86	N 31	O 54	P 9	0	0	0

- Molecule 4 is a DNA chain called DNA (5'-D(*TP*TP*TP*TP*TP*TP*G*AP*GP*GP*CP*GP*TP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	G	7	Total	C	N	O	P	0	0	0
			145	69	30	40	6			
4	J	7	Total	C	N	O	P	0	0	0
			145	69	30	40	6			

- Molecule 5 is a DNA chain called DNA (5'-D(P*GP*CP*TP*GP*AP*GP*GP*AP*GP*T P*CP*T)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	12	Total	C	N	O	P	0	0	0
			250	118	47	73	12			
5	H	12	Total	C	N	O	P	0	0	0
			250	118	47	73	12			

- Molecule 6 is BARIUM ION (three-letter code: BA) (formula: Ba).

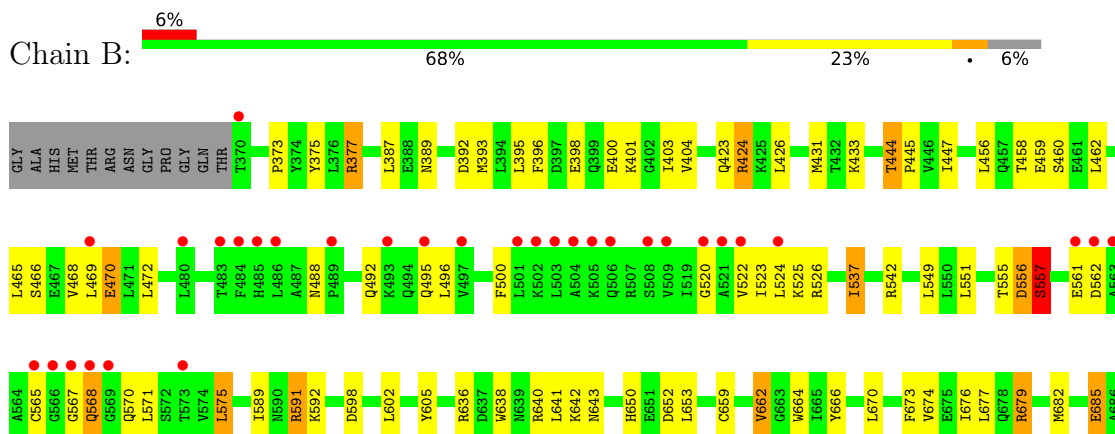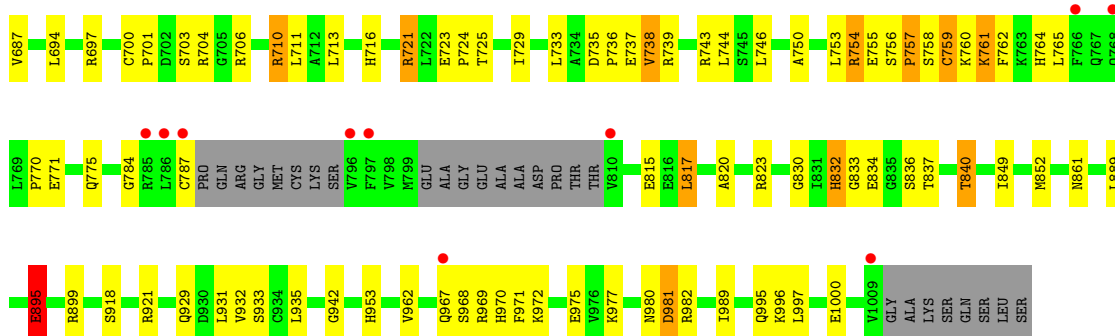| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total	Ba	0	0
			1	1		
6	B	1	Total	Ba	0	0
			1	1		

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.

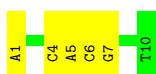• Molecule 1: Fanconi-associated nuclease 1
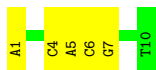
Chain A:

• Molecule 1: Fanconi-associated nuclease 1

Chain B:

V687 L694 R697 C700 P701 D702 S703 R704 R706 G705 R710 L711 A712 L713 H716 R721 L722 E723 P724 T725 I729 L733 A734 D735 P736 E737 V738 R739 R743 L744 S745 L746 A750 L753 R754 E755 S756 P757 S758 C759 K760 K761 F762 K763 H764 L765 F766 Q767 Q768

L769 P770 E771 Q775 G784 R785 L786 C787 PRO GLN ARG GLY MET CYS LYS SER V796 F797 V798 M799 GLU ALA GLY GLU ALA ALA ASP PRO THR THR V810 E815 E816 L817 A820 R823 G830 I831 H832 G833 E834 G835 S836 T837 T840 I849 M852 N861 L889

E895 R899 S918 R921 Q929 D930 L931 V932 S933 C934 L935 G942 H953 V962 Q967 S968 R969 H970 F971 K972 E975 V976 K977 N980 D981 R982 I989 Q995 K996 L997 E1000 V1009 GLY ALA LYS SER GLN SER LEU SER

• Molecule 2: DNA (5'-D(P*AP*GP*CP*CP*AP*CP*GP*CP*CP*T)-3')

Chain F: 50%   50%

A1 C4 A5 C6 G7 T10

• Molecule 2: DNA (5'-D(P*AP*GP*CP*CP*AP*CP*GP*CP*CP*T)-3')

Chain L: 50%   50%

A1 C4 A5 C6 G7 T10

• Molecule 3: DNA (5'-D(P*AP*GP*AP*CP*TP*CP*CP*TP*C)-3')

Chain K: 67%   33%

A11 C14 T15 C16 C17 T18 C19

• Molecule 3: DNA (5'-D(P*AP*GP*AP*CP*TP*CP*CP*TP*C)-3')

Chain I: 67%   33%

A11 C14 T15 C16 C17 T18 C19

• Molecule 4: DNA (5'-D(*TP*TP*TP*TP*TP*TP*G*AP*GP*GP*CP*GP*TP*G)-3')

Chain G: 29%   21%   50%

DT DT DT DT DT DG A8 G12 T13 G14
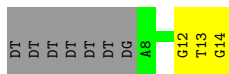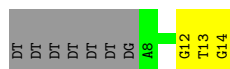
• Molecule 4: DNA (5'-D(*TP*TP*TP*TP*TP*TP*G*AP*GP*GP*CP*GP*TP*G)-3')

Chain J: 29%   21%   50%

• Molecule 5: DNA (5'-D(P*GP*CP*TP*GP*AP*GP*GP*AP*GP*TP*CP*T)-3')

Chain E:

• Molecule 5: DNA (5'-D(P*GP*CP*TP*GP*AP*GP*GP*AP*GP*TP*CP*T)-3')

Chain H:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 2 21 21	Depositor
Cell constants a, b, c, α, β, γ	87.81Å 100.32Å 212.47Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 3.00 48.82 – 2.99	Depositor EDS
% Data completeness (in resolution range)	93.6 (40.00-3.00) 93.3 (48.82-2.99)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.87 (at 3.01Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.224 , 0.267 0.224 , 0.271	Depositor DCC
R_{free} test set	1512 reflections (4.02%)	wwPDB-VP
Wilson B-factor (Å^2)	55.4	Xtriage
Anisotropy	0.514	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 46.7	EDS
L-test for twinning[2]	$< |L| > = 0.46, < L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	11410	wwPDB-VP
Average B, all atoms (Å^2)	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 51.72 % 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 5.3913e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[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 (i)

5.1 Standard geometry (i)

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

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	0.44	0/5029	0.65	0/6799
1	B	0.46	0/5029	0.65	0/6799
2	F	0.81	1/225 (0.4%)	0.89	0/342
2	L	0.78	1/225 (0.4%)	0.87	0/342
3	I	0.37	0/200	0.82	0/305
3	K	0.31	0/200	0.82	0/305
4	G	0.42	0/163	0.64	0/251
4	J	0.41	0/163	0.70	0/251
5	E	0.30	0/280	0.79	0/431
5	H	0.32	0/280	0.77	1/431 (0.2%)
All	All	0.46	2/11794 (0.0%)	0.68	1/16256 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	1	DA	OP3-P	-10.59	1.48	1.61
2	L	1	DA	OP3-P	-10.04	1.49	1.61

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	H	9	DG	C1'-O4'-C4'	-5.13	104.97	110.10

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	4927	0	4960	108	0
1	B	4927	0	4960	114	0
2	F	202	0	112	2	0
2	L	202	0	112	4	0
3	I	180	0	102	2	0
3	K	180	0	102	2	0
4	G	145	0	80	7	0
4	J	145	0	80	6	0
5	E	250	0	136	8	0
5	H	250	0	136	5	0
6	A	1	0	0	0	0
6	B	1	0	0	0	0
All	All	11410	0	10780	241	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 11.

The worst 5 of 241 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:389:ASN:HB3	1:A:591:ARG:HH21	1.07	1.09
1:A:682:MET:HB3	1:A:685:GLU:HG3	1.34	1.09
1:B:682:MET:HB3	1:B:685:GLU:HG3	1.39	1.05
1:B:836:SER:O	1:B:840:THR:HG23	1.66	0.95
1:A:836:SER:O	1:A:840:THR:HG23	1.67	0.93

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	607/651 (93%)	549 (90%)	44 (7%)	14 (2%)	6	30
1	B	607/651 (93%)	546 (90%)	52 (9%)	9 (2%)	10	42
All	All	1214/1302 (93%)	1095 (90%)	96 (8%)	23 (2%)	8	36

5 of 23 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	557	SER
1	A	558	MET
1	A	564	ALA
1	B	557	SER
1	A	760	LYS

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	A	536/563 (95%)	491 (92%)	45 (8%)	11	38
1	B	536/563 (95%)	492 (92%)	44 (8%)	11	39
All	All	1072/1126 (95%)	983 (92%)	89 (8%)	11	39

5 of 89 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	556	ASP
1	B	679	ARG
1	B	561	GLU
1	B	636	ARG
1	B	738	VAL

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

Mol	Chain	Res	Type
1	B	953	HIS
1	B	926	GLN
1	A	995	GLN
1	B	775	GLN
1	A	953	HIS

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 2 ligands modelled in this entry, 2 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	A	613/651 (94%)	0.03	39 (6%)	19	6	31, 54, 145, 173	0
1	B	613/651 (94%)	0.08	42 (6%)	16	5	33, 54, 144, 176	0
2	F	10/10 (100%)	-0.49	0	100	100	55, 67, 79, 86	0
2	L	10/10 (100%)	-0.49	0	100	100	59, 65, 73, 78	0
3	I	9/9 (100%)	-0.54	0	100	100	57, 63, 72, 73	0
3	K	9/9 (100%)	-0.58	0	100	100	53, 62, 70, 70	0
4	G	7/14 (50%)	-0.44	0	100	100	46, 51, 64, 68	0
4	J	7/14 (50%)	-0.59	0	100	100	45, 53, 72, 77	0
5	E	12/12 (100%)	-0.49	0	100	100	62, 79, 83, 84	0
5	H	12/12 (100%)	-0.54	0	100	100	55, 78, 84, 85	0
All	All	1302/1392 (93%)	0.02	81 (6%)	20	7	31, 55, 143, 176	0

The worst 5 of 81 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	568	GLN	9.1
1	A	568	GLN	7.1
1	B	484	PHE	6.4
1	B	796	VAL	6.2
1	B	480	LEU	5.7

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
6	BA	B	1101	1/1	0.99	0.20	61,61,61,61	0
6	BA	A	1101	1/1	1.00	0.19	54,54,54,54	0

6.5 Other polymers ⓘ

There are no such residues in this entry.