WORLDWIDE PDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Feb 28, 2024 – 12:46 AM EST

PDB ID : 5QJ9
Title : PanDDA analysis group deposition of models with modelled events (e.g. bound ligands) – Crystal Structure of NUDT5 in complex with Z768399682
Authors : Dubianok, Y.; Collins, P.; Krojer, T.; Wright, N.; Strain-Damerell, C.; Burgess-Brown, N.; Bountra, C.; Arrowsmith, C.H.; Edwards, A.; Huber, K.; von Delft, F.
Deposited on : 2018-11-12
Resolution : 1.85 Å(reported)

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

|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
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 1.85 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.
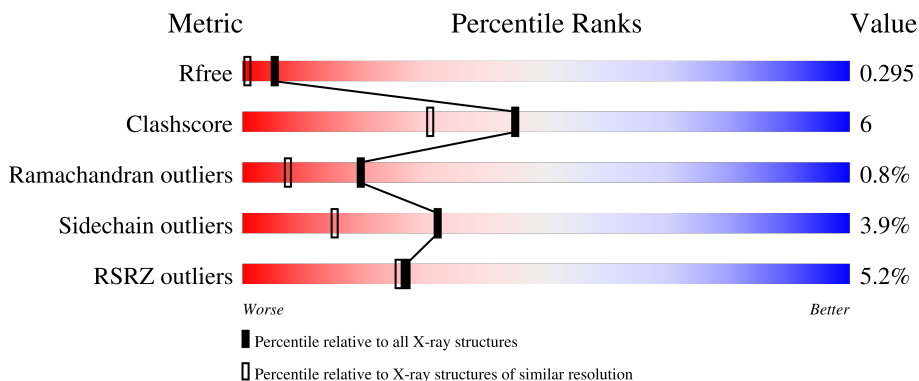
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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	209	2% 84% 8% 7%
1	B	209	4% 82% 10% · 7%
1	C	209	4% 78% 13% 8%
1	D	209	8% 70% 19% · 9%

2 Entry composition ⓘ

There are 6 unique types of molecules in this entry. The entry contains 6060 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 ADP-sugar pyrophosphatase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	194	Total	C	N	O	S	0	0	0
			1450	917	242	283	8			
1	B	194	Total	C	N	O	S	0	0	0
			1464	924	244	288	8			
1	C	192	Total	C	N	O	S	0	0	0
			1412	890	238	277	7			
1	D	190	Total	C	N	O	S	0	1	0
			1406	888	238	272	8			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	SER	-	expression tag	UNP Q9UKK9
B	0	SER	-	expression tag	UNP Q9UKK9
C	0	SER	-	expression tag	UNP Q9UKK9
D	0	SER	-	expression tag	UNP Q9UKK9

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	2	Total	Mg	0	0
			2	2		
2	B	2	Total	Mg	0	0
			2	2		
2	C	2	Total	Mg	0	0
			2	2		
2	D	2	Total	Mg	0	0
			2	2		

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

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

- Molecule 4 is [(2R,5S)-2,5-dimethylmorpholin-4-yl](1,2,5-thiadiazol-3-yl)methanone (three-letter code: K0D) (formula: $C_9H_{13}N_3O_2S$).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total 15	C 9	N 3	O 2	S 1	0	0
4	B	1	Total 15	C 9	N 3	O 2	S 1	0	0
4	C	1	Total 15	C 9	N 3	O 2	S 1	0	0
4	C	1	Total 15	C 9	N 3	O 2	S 1	0	0

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).

EDO

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total 4	C 2	O 2	0	0
5	B	1	Total 4	C 2	O 2	0	0
5	C	1	Total 4	C 2	O 2	0	0

• Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	74	Total 74	O 74	0	0
6	B	73	Total 73	O 73	0	0
6	C	70	Total 70	O 70	0	0
6	D	30	Total 30	O 30	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 as grey.
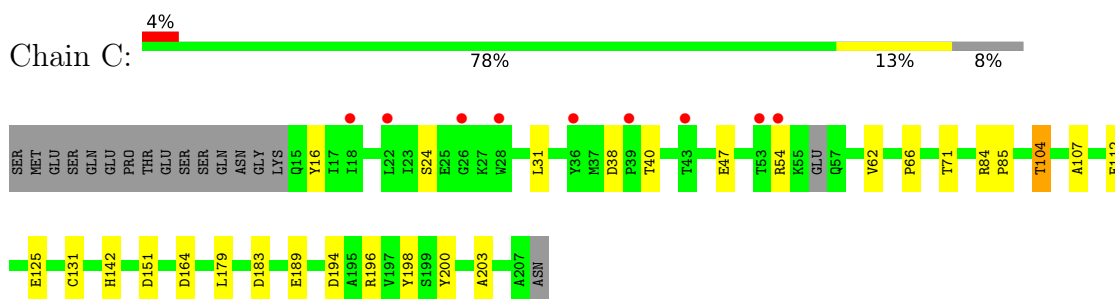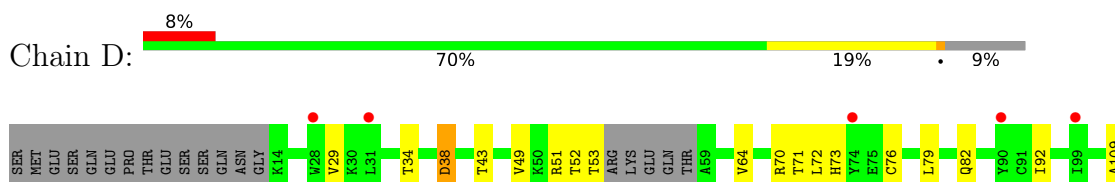
• Molecule 1: ADP-sugar pyrophosphatase
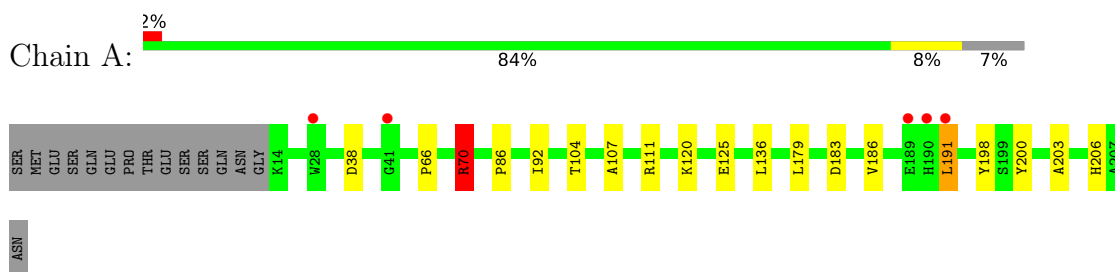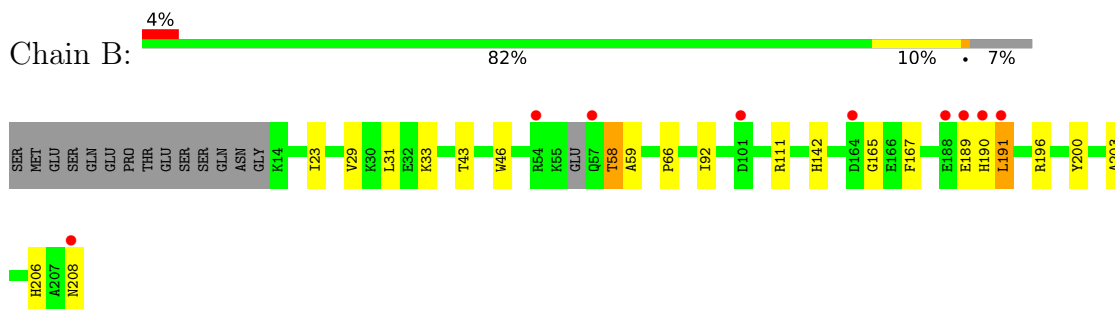
Chain A:

• Molecule 1: ADP-sugar pyrophosphatase

Chain B:

• Molecule 1: ADP-sugar pyrophosphatase

Chain C:

• Molecule 1: ADP-sugar pyrophosphatase

Chain D:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	49.59Å 59.88Å 79.67Å 79.45° 81.18° 75.37°	Depositor
Resolution (Å)	77.83 – 1.85 77.83 – 1.85	Depositor EDS
% Data completeness (in resolution range)	96.3 (77.83-1.85) 96.3 (77.83-1.85)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.43 (at 1.84Å)	Xtriage
Refinement program	REFMAC 5.8.0189	Depositor
R, R_{free}	0.242 , 0.288 0.256 , 0.295	Depositor DCC
R_{free} test set	3566 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	28.1	Xtriage
Anisotropy	0.160	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 41.5	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	6060	wwPDB-VP
Average B, all atoms (Å^2)	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.74% 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: MG, CL, EDO, K0D

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	0.89	0/1478	0.95	4/2019 (0.2%)
1	B	0.79	0/1492	0.88	2/2035 (0.1%)
1	C	0.83	0/1439	0.93	2/1969 (0.1%)
1	D	0.78	0/1432	0.90	1/1956 (0.1%)
All	All	0.83	0/5841	0.91	9/7979 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	1

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	70	ARG	NE-CZ-NH1	7.44	124.02	120.30
1	A	183	ASP	CB-CG-OD1	6.41	124.07	118.30
1	C	151	ASP	CB-CG-OD1	6.12	123.81	118.30
1	A	70	ARG	NE-CZ-NH2	-6.03	117.29	120.30
1	D	38	ASP	CB-CG-OD1	5.61	123.34	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	58	THR	Peptide

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	1450	0	1405	18	0
1	B	1464	0	1414	25	0
1	C	1412	0	1320	14	0
1	D	1406	0	1345	30	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
2	C	2	0	0	0	0
2	D	2	0	0	0	0
3	A	1	0	0	0	0
4	A	15	0	0	0	0
4	B	15	0	0	0	0
4	C	30	0	0	1	0
5	A	4	0	6	0	0
5	B	4	0	6	0	0
5	C	4	0	6	0	0
6	A	74	0	0	2	0
6	B	73	0	0	0	0
6	C	70	0	0	2	0
6	D	30	0	0	1	0
All	All	6060	0	5502	73	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 6.

The worst 5 of 73 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:191:LEU:O	1:A:191:LEU:HD23	1.30	1.25
1:D:207:ALA:O	1:D:208:ASN:ND2	1.76	1.17
1:A:191:LEU:HD23	1:A:191:LEU:C	1.66	1.12
1:B:189:GLU:HB3	1:B:191:LEU:HD12	1.48	0.95
1:A:191:LEU:O	1:A:191:LEU:CD2	2.13	0.95

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	192/209 (92%)	188 (98%)	3 (2%)	1 (0%)	29	15
1	B	190/209 (91%)	184 (97%)	4 (2%)	2 (1%)	14	4
1	C	188/209 (90%)	181 (96%)	5 (3%)	2 (1%)	14	4
1	D	187/209 (90%)	169 (90%)	17 (9%)	1 (0%)	29	15
All	All	757/836 (91%)	722 (95%)	29 (4%)	6 (1%)	19	7

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	207	ALA
1	B	59	ALA
1	C	24	SER
1	B	66	PRO
1	C	66	PRO

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	151/179 (84%)	148 (98%)	3 (2%)	55	40
1	B	154/179 (86%)	151 (98%)	3 (2%)	57	43
1	C	141/179 (79%)	133 (94%)	8 (6%)	20	6
1	D	143/179 (80%)	134 (94%)	9 (6%)	18	4
All	All	589/716 (82%)	566 (96%)	23 (4%)	32	15

5 of 23 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	38	ASP
1	D	72	LEU
1	D	70	ARG
1	D	76	CYS
1	C	16	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	149	ASN
1	D	155	ASN
1	D	208	ASN
1	D	206	HIS
1	B	206	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 16 ligands modelled in this entry, 9 are monoatomic - leaving 7 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
5	EDO	A	305	-	3,3,3	0.41	0	2,2,2	0.22	0
5	EDO	B	301	-	3,3,3	0.61	0	2,2,2	0.35	0
4	K0D	C	305	-	16,16,16	2.62	6 (37%)	17,22,22	1.88	4 (23%)
4	K0D	B	304	-	16,16,16	2.35	4 (25%)	17,22,22	1.20	2 (11%)
5	EDO	C	301	-	3,3,3	0.58	0	2,2,2	0.12	0
4	K0D	C	304	-	16,16,16	2.54	5 (31%)	17,22,22	3.12	10 (58%)
4	K0D	A	304	-	16,16,16	2.35	6 (37%)	17,22,22	1.39	3 (17%)

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
5	EDO	A	305	-	-	0/1/1/1	-
5	EDO	B	301	-	-	0/1/1/1	-
4	K0D	C	305	-	-	0/4/21/21	0/2/2/2
4	K0D	B	304	-	-	0/4/21/21	0/2/2/2
5	EDO	C	301	-	-	0/1/1/1	-
4	K0D	C	304	-	-	1/4/21/21	0/2/2/2
4	K0D	A	304	-	-	0/4/21/21	0/2/2/2

The worst 5 of 21 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	304	K0D	C8-C7	7.02	1.49	1.37
4	B	304	K0D	C8-C7	6.71	1.49	1.37
4	A	304	K0D	C8-C7	6.13	1.48	1.37
4	C	305	K0D	C3-N	5.41	1.54	1.47
4	C	304	K0D	C3-N	4.89	1.53	1.47

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	C	304	K0D	C2-N-C3	-8.42	107.56	115.70
4	C	305	K0D	N2-S-N1	-4.91	95.03	100.42
4	C	304	K0D	C7-C6-N	4.52	124.97	119.76
4	C	304	K0D	C4-C3-N	4.43	118.94	111.64
4	C	305	K0D	C7-C8-N1	3.01	114.18	108.95

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	C	304	K0D	O1-C6-N-C3

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	C	305	K0D	1	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 ⓘ

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	194/209 (92%)	-0.13	5 (2%)	56	54	18, 28, 50, 62	6 (3%)
1	B	194/209 (92%)	-0.06	9 (4%)	32	31	19, 33, 56, 74	6 (3%)
1	C	192/209 (91%)	0.09	9 (4%)	31	30	19, 36, 64, 78	10 (5%)
1	D	190/209 (90%)	0.63	17 (8%)	9	9	24, 45, 68, 91	19 (10%)
All	All	770/836 (92%)	0.13	40 (5%)	27	26	18, 36, 63, 91	41 (5%)

The worst 5 of 40 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	207	ALA	8.0
1	B	191	LEU	6.9
1	A	191	LEU	6.7
1	D	206	HIS	5.8
1	D	208	ASN	4.5

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
4	K0D	C	305	15/15	0.83	0.18	45,52,61,63	15
2	MG	A	302	1/1	0.84	0.04	52,52,52,52	0
3	CL	A	303	1/1	0.85	0.07	66,66,66,66	0
4	K0D	B	304	15/15	0.87	0.19	31,39,43,45	15
4	K0D	A	304	15/15	0.88	0.15	54,58,64,64	0
4	K0D	C	304	15/15	0.90	0.19	45,53,62,64	15
2	MG	D	302	1/1	0.92	0.14	50,50,50,50	0
5	EDO	B	301	4/4	0.92	0.11	36,36,38,39	0
5	EDO	A	305	4/4	0.93	0.13	37,38,39,41	0
2	MG	D	301	1/1	0.93	0.05	37,37,37,37	0
5	EDO	C	301	4/4	0.95	0.07	33,42,44,48	0
2	MG	B	302	1/1	0.96	0.04	30,30,30,30	0
2	MG	C	302	1/1	0.96	0.05	50,50,50,50	0
2	MG	B	303	1/1	0.97	0.10	61,61,61,61	0
2	MG	C	303	1/1	0.98	0.05	27,27,27,27	0
2	MG	A	301	1/1	0.99	0.07	29,29,29,29	0

6.5 Other polymers (i)

There are no such residues in this entry.