



Full wwPDB X-ray Structure Validation Report ⓘ

Nov 4, 2024 – 12:30 am GMT

PDB ID : 8Q00
Title : TssM-Ub-PA complex - A USP-like DUB from *B. pseudomallei* (193-430) re-acted with Ub-PA
Authors : Uthoff, M.; Hermanns, T.; Hofmann, K.; Baumann, U.
Deposited on : 2023-07-27
Resolution : 1.62 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

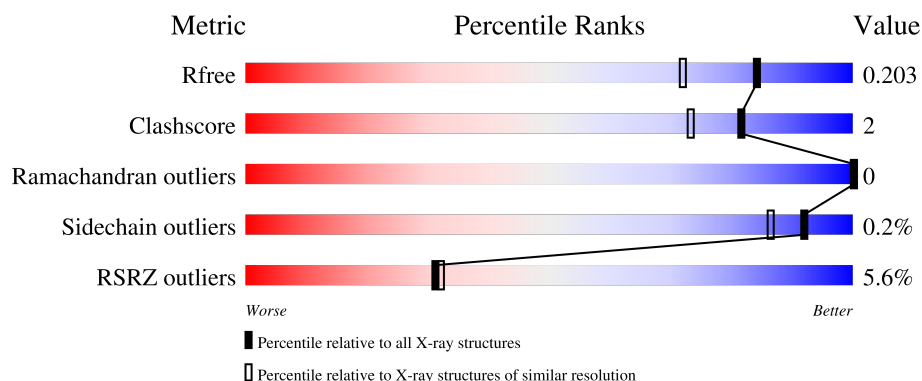
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.62 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}	164625	6077 (1.64-1.60)
Clashscore	180529	6617 (1.64-1.60)
Ramachandran outliers	177936	6498 (1.64-1.60)
Sidechain outliers	177891	6497 (1.64-1.60)
RSRZ outliers	164620	6075 (1.64-1.60)

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	300	<div> <div>5%</div> <div>89%</div> <div>9%</div> <div>.</div> </div>
1	C	300	<div> <div>5%</div> <div>97%</div> <div>.</div> </div>
2	B	76	<div> <div>11%</div> <div>95%</div> <div>5%</div> <div>.</div> </div>
2	D	76	<div> <div>4%</div> <div>97%</div> <div>.</div> </div>

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 6582 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 TssM.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	295	Total	C	N	O	S	0	14	0
			2298	1451	399	442	6			
1	C	300	Total	C	N	O	S	0	11	0
			2316	1457	406	449	4			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	191	GLY	-	expression tag	UNP D7SFB8
C	191	GLY	-	expression tag	UNP D7SFB8

- Molecule 2 is a protein called Polyubiquitin-B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	76	Total	C	N	O	S	0	0	0
			601	379	105	116	1			
2	D	76	Total	C	N	O	S	0	3	0
			620	391	109	119	1			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	76	AYE	GLY	engineered mutation	UNP P0CG47
D	76	AYE	GLY	engineered mutation	UNP P0CG47

- Molecule 3 is CITRATE ANION (three-letter code: FLC) (formula: C₆H₅O₇).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			13	6	7		

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



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

- Molecule 5 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total 1	Na 1	0	0
5	C	1	Total 1	Na 1	0	0

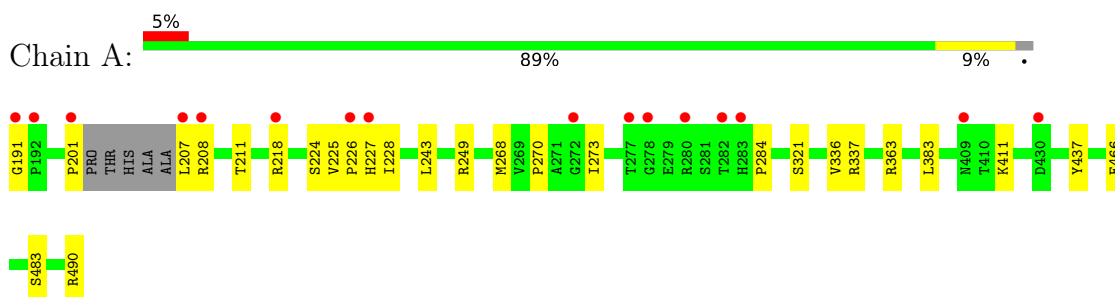
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	324	Total 324	O 324	0	0
6	B	50	Total 50	O 50	0	0
6	C	277	Total 277	O 277	0	0
6	D	73	Total 73	O 73	0	0

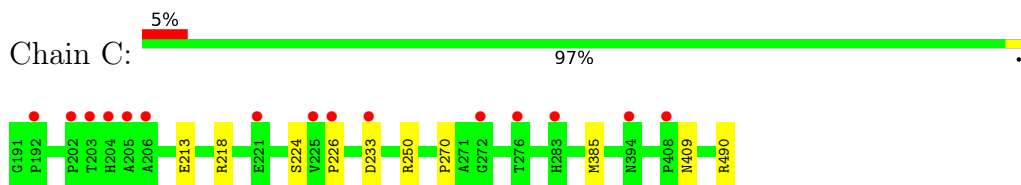
3 Residue-property plots [i](#)

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 in grey.

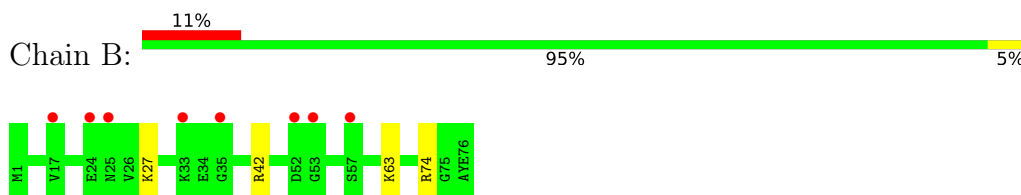
- Molecule 1: TssM



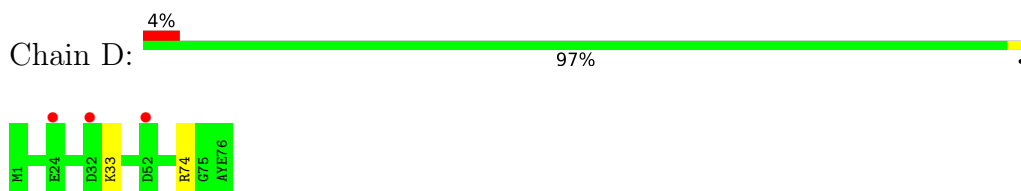
- Molecule 1: TssM



- Molecule 2: Polyubiquitin-B



- Molecule 2: Polyubiquitin-B



4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α , β , γ	103.75Å 103.75Å 191.52Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	91.22 – 1.62 91.22 – 1.62	Depositor EDS
% Data completeness (in resolution range)	100.0 (91.22-1.62) 93.2 (91.22-1.62)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.03 (at 1.62Å)	Xtriage
Refinement program	REFMAC 5.8.0419, PHENIX 1.20.1-4487	Depositor
R, R_{free}	0.174 , 0.198 0.185 , 0.203	Depositor DCC
R_{free} test set	130701 reflections (1.51%)	wwPDB-VP
Wilson B-factor (Å ²)	24.9	Xtriage
Anisotropy	0.188	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 38.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	6582	wwPDB-VP
Average B, all atoms (Å ²)	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.08% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: AYE, NA, FLC, EDO

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	0.59	1/2389 (0.0%)	0.93	2/3274 (0.1%)
1	C	0.53	0/2407	0.88	2/3299 (0.1%)
2	B	0.58	0/603	0.91	2/811 (0.2%)
2	D	0.61	0/631	0.89	0/848
All	All	0.57	1/6030 (0.0%)	0.91	6/8232 (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	A	0	2
2	B	0	1
2	D	0	2
All	All	0	5

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	466	GLU	CD-OE2	-5.38	1.19	1.25

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	42	ARG	NE-CZ-NH2	-8.56	116.02	120.30
1	C	490	ARG	NE-CZ-NH1	-6.70	116.95	120.30
2	B	42	ARG	NE-CZ-NH1	6.44	123.52	120.30
1	C	218	ARG	NE-CZ-NH1	5.66	123.13	120.30
1	A	490	ARG	NE-CZ-NH2	-5.34	117.63	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	249	ARG	NE-CZ-NH1	5.10	122.85	120.30

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	337	ARG	Sidechain
1	A	363	ARG	Sidechain
2	B	74	ARG	Sidechain
2	D	74[A]	ARG	Sidechain
2	D	74[B]	ARG	Sidechain

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	2298	0	2286	21	0
1	C	2316	0	2294	4	0
2	B	601	0	630	2	0
2	D	620	0	655	1	0
3	A	13	0	5	0	0
4	A	4	0	6	0	0
4	C	4	0	6	0	0
5	A	1	0	0	0	0
5	C	1	0	0	0	0
6	A	324	0	0	0	3
6	B	50	0	0	2	1
6	C	277	0	0	0	0
6	D	73	0	0	0	1
All	All	6582	0	5882	28	3

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 2.

All (28) 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:226[B]:PRO:O	1:A:270:PRO:HG2	1.82	0.78
1:C:233:ASP:HB2	1:C:409:ASN:OD1	1.95	0.67
1:A:226[B]:PRO:HG3	1:A:270:PRO:O	1.94	0.66
1:C:213[B]:GLU:OE2	1:C:250:ARG:NE	2.31	0.62
1:A:207:LEU:HG	1:A:208:ARG:H	1.68	0.58
1:A:224:SER:O	1:A:226[B]:PRO:HD3	2.05	0.57
1:A:207:LEU:HG	1:A:208:ARG:N	2.22	0.54
1:A:191:GLY:O	1:A:218:ARG:NH1	2.41	0.53
1:A:226[B]:PRO:O	1:A:227[B]:HIS:HB2	2.09	0.51
1:C:224:SER:O	1:C:226:PRO:HD3	2.11	0.51
1:A:268[B]:MET:SD	1:A:284:PRO:HB3	2.51	0.51
1:A:225[B]:VAL:CG1	1:A:228:ILE:HD11	2.40	0.50
1:A:225[B]:VAL:HG11	1:A:243:LEU:HD11	1.94	0.49
1:A:224:SER:O	1:A:226[B]:PRO:N	2.45	0.49
1:A:224:SER:O	1:A:226[B]:PRO:CD	2.61	0.48
1:A:226[B]:PRO:HB2	1:A:270:PRO:HB2	1.94	0.48
1:A:336:VAL:HG11	1:A:383:LEU:CD2	2.45	0.47
1:A:208:ARG:HG3	1:A:208:ARG:HH11	1.82	0.45
1:C:226:PRO:HB2	1:C:270:PRO:HD2	2.00	0.43
2:D:33:LYS:O	2:D:33:LYS:HG3	2.18	0.43
1:A:225[B]:VAL:HG11	1:A:243:LEU:CD1	2.49	0.43
1:A:321[B]:SER:OG	1:A:411:LYS:HD3	2.18	0.42
1:A:226[B]:PRO:HG2	1:A:270:PRO:HD2	2.03	0.41
1:A:437:TYR:CE1	1:A:483[B]:SER:HB2	2.56	0.41
1:A:201:PRO:HD2	1:A:211:THR:O	2.21	0.40
2:B:63:LYS:HB2	6:B:108:HOH:O	2.21	0.40
1:A:225[A]:VAL:CG1	1:A:273:ILE:HD12	2.50	0.40
2:B:27:LYS:NZ	6:B:101:HOH:O	2.32	0.40

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
6:A:848:HOH:O	6:B:144:HOH:O[4_454]	1.94	0.26
6:A:867:HOH:O	6:A:867:HOH:O[7_555]	1.98	0.22
6:A:660:HOH:O	6:D:130:HOH:O[6_454]	2.12	0.08

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	305/300 (102%)	296 (97%)	9 (3%)	0	100	100
1	C	309/300 (103%)	302 (98%)	7 (2%)	0	100	100
2	B	73/76 (96%)	72 (99%)	1 (1%)	0	100	100
2	D	76/76 (100%)	75 (99%)	1 (1%)	0	100	100
All	All	763/752 (102%)	745 (98%)	18 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	253/242 (104%)	253 (100%)	0	100	100
1	C	253/242 (104%)	252 (100%)	1 (0%)	89	83
2	B	68/68 (100%)	68 (100%)	0	100	100
2	D	71/68 (104%)	71 (100%)	0	100	100
All	All	645/620 (104%)	644 (100%)	1 (0%)	92	86

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	385	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are

no such sidechains identified.

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 oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 5 ligands modelled in this entry, 2 are monoatomic - leaving 3 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	$\# Z > 2$
3	FLC	A	501	-	12,12,12	0.98	1 (8%)	17,17,17	1.33	1 (5%)
4	EDO	C	501	-	3,3,3	0.19	0	2,2,2	0.17	0
4	EDO	A	502	-	3,3,3	0.34	0	2,2,2	0.43	0

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
3	FLC	A	501	-	-	2/16/16/16	-
4	EDO	C	501	-	-	1/1/1/1	-
4	EDO	A	502	-	-	0/1/1/1	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	501	FLC	CB-CBC	2.09	1.55	1.53

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	501	FLC	OB1-CBC-CB	-3.14	117.81	122.25

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	C	501	EDO	O1-C1-C2-O2
3	A	501	FLC	CB-CG-CGC-OG1
3	A	501	FLC	CB-CG-CGC-OG2

There are no ring outliers.

No monomer is involved in short contacts.

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 ⓘ

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(Å ²)	Q<0.9
1	A	295/300 (98%)	0.16	16 (5%)	32 33	14, 25, 50, 71	14 (4%)
1	C	300/300 (100%)	0.35	15 (5%)	35 36	17, 30, 52, 79	11 (3%)
2	B	75/76 (98%)	0.69	8 (10%)	12 12	21, 34, 53, 60	0
2	D	75/76 (98%)	0.47	3 (4%)	43 44	18, 30, 44, 51	3 (4%)
All	All	745/752 (99%)	0.32	42 (5%)	31 32	14, 28, 51, 79	28 (3%)

All (42) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	207	LEU	7.7
1	A	226[A]	PRO	4.5
1	C	226	PRO	3.9
1	A	278	GLY	3.8
1	C	205	ALA	3.5
1	C	202	PRO	3.4
1	C	283	HIS	3.4
1	A	409	ASN	3.4
1	A	277	THR	3.2
1	A	282	THR	3.1
1	C	203	THR	3.0
2	B	24	GLU	3.0
1	C	394[A]	ASN	2.7
1	A	208	ARG	2.7
2	B	53	GLY	2.7
1	A	201	PRO	2.6
2	B	33	LYS	2.5
1	C	221	GLU	2.5
1	C	276	THR	2.5
1	A	218	ARG	2.4
1	C	204	HIS	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	233	ASP	2.4
2	B	35	GLY	2.4
1	A	430	ASP	2.4
1	C	225	VAL	2.3
1	A	272	GLY	2.3
1	A	283	HIS	2.3
2	B	57	SER	2.3
2	D	52	ASP	2.3
1	C	272	GLY	2.3
1	A	280	ARG	2.3
2	D	32	ASP	2.3
1	C	192	PRO	2.2
2	B	25	ASN	2.2
1	A	227[A]	HIS	2.2
1	A	191	GLY	2.2
1	A	192	PRO	2.1
2	B	17	VAL	2.1
2	B	52	ASP	2.1
2	D	24	GLU	2.0
1	C	408	PRO	2.0
1	C	206	ALA	2.0

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 [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. 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(\AA^2)	Q<0.9
4	EDO	C	501	4/4	0.82	0.16	48,52,58,60	0
3	FLC	A	501	13/13	0.89	0.11	28,34,52,52	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
4	EDO	A	502	4/4	0.90	0.12	35,39,42,49	0
5	NA	C	502	1/1	0.97	0.07	31,31,31,31	0
5	NA	A	503	1/1	1.00	0.08	28,28,28,28	0

6.5 Other polymers [i](#)

There are no such residues in this entry.