Full wwPDB X-ray Structure Validation Report (i)

Jun 22, 2024 – 06:38 PM EDT

PDB ID	:	5NDD
Title	:	Crystal structure of a thermostabilised human protease-activated receptor-2 (PAR2) in complex with AZ8838 at 2.8 angstrom resolution
Authors	:	Cheng, R.K.Y.; Fiez-Vandal, C.; Schlenker, O.; Edman, K.; Aggeler, B.; Brown, D.G.; Brown, G.; Cooke, R.M.; Dumelin, C.E.; Dore, A.S.; Geschwindner, S.; Grebner, C.; Hermansson, N.-O.; Jazayeri, A.; Johansson, P.; Leong, L.; Prihandoko, R.; Rappas, M.; Soutter, H.; Snijder, A.; Sundstrom, L.; Tehan, B.; Thornton, P.; Troast, D.; Wiggin, G.; Zhukov, A.; Marshall, F.H.; Dekker, N.
Deposited on	:	2017-03-08
Resolution	:	2.80 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 (i) 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 (1)) were used in the production of this report:

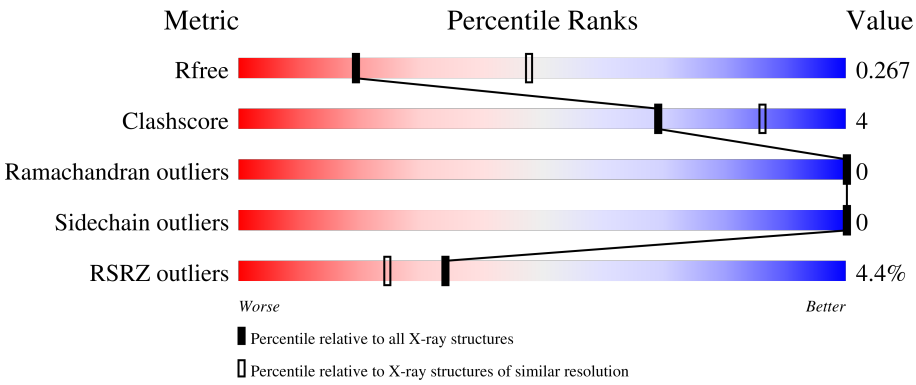
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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	619	4% 81% 11% 8%

Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.37.1

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 4531 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 Lysozyme,Proteinase-activated receptor 2,Soluble cytochrome b562,Proteinase-activated receptor 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	567	4498	2927	751	800	20	0	0	0

There are 55 discrepancies between the modelled and reference sequences:

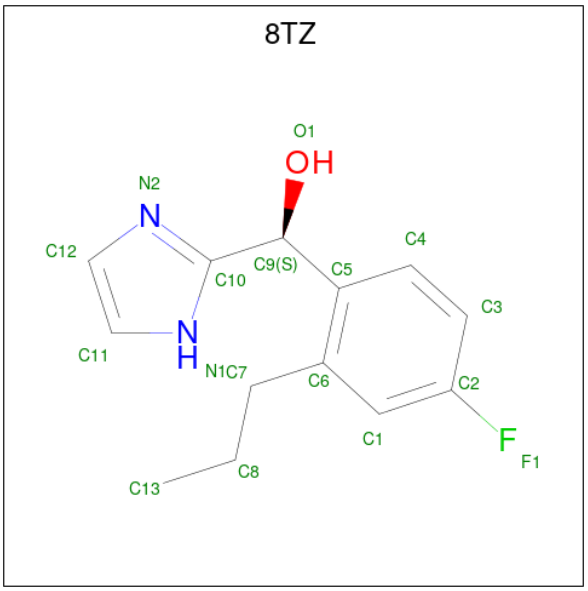
Chain	Residue	Modelled	Actual	Comment	Reference
A	979	MET	-	initiating methionine	UNP D9IEF7
A	980	VAL	-	expression tag	UNP D9IEF7
A	981	SER	-	expression tag	UNP D9IEF7
A	982	ALA	-	expression tag	UNP D9IEF7
A	983	ILE	-	expression tag	UNP D9IEF7
A	984	VAL	-	expression tag	UNP D9IEF7
A	985	LEU	-	expression tag	UNP D9IEF7
A	986	TYR	-	expression tag	UNP D9IEF7
A	987	VAL	-	expression tag	UNP D9IEF7
A	988	LEU	-	expression tag	UNP D9IEF7
A	989	LEU	-	expression tag	UNP D9IEF7
A	990	ALA	-	expression tag	UNP D9IEF7
A	991	ALA	-	expression tag	UNP D9IEF7
A	992	ALA	-	expression tag	UNP D9IEF7
A	993	ALA	-	expression tag	UNP D9IEF7
A	994	HIS	-	expression tag	UNP D9IEF7
A	995	SER	-	expression tag	UNP D9IEF7
A	996	ALA	-	expression tag	UNP D9IEF7
A	997	PHE	-	expression tag	UNP D9IEF7
A	998	ALA	-	expression tag	UNP D9IEF7
A	999	ALA	-	expression tag	UNP D9IEF7
A	1000	ALA	-	expression tag	UNP D9IEF7
A	1001	SER	-	expression tag	UNP D9IEF7
A	1054	THR	CYS	conflict	UNP D9IEF7
A	1097	ALA	CYS	conflict	UNP D9IEF7
A	1162	ILE	-	linker	UNP D9IEF7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	1163	TYR	-	linker	UNP D9IEF7
A	1164	GLU	-	linker	UNP D9IEF7
A	1165	PHE	-	linker	UNP D9IEF7
A	89	ALA	GLY	engineered mutation	UNP P55085
A	108	ALA	HIS	engineered mutation	UNP P55085
A	157	ALA	GLY	engineered mutation	UNP P55085
A	166	LEU	MET	engineered mutation	UNP P55085
A	174	ALA	TYR	engineered mutation	UNP P55085
A	176	GLU	VAL	engineered mutation	UNP P55085
A	222	GLN	ASN	engineered mutation	UNP P55085
A	268	ALA	MET	engineered mutation	UNP P55085
A	2006	TRP	MET	conflict	UNP P0ABE7
A	2101	ILE	HIS	conflict	UNP P0ABE7
A	2105	LEU	-	linker	UNP P0ABE7
A	289	ALA	ILE	engineered mutation	UNP P55085
A	293	ALA	LEU	engineered mutation	UNP P55085
A	378	ALA	-	expression tag	UNP P55085
A	379	ALA	-	expression tag	UNP P55085
A	380	ALA	-	expression tag	UNP P55085
A	381	HIS	-	expression tag	UNP P55085
A	382	HIS	-	expression tag	UNP P55085
A	383	HIS	-	expression tag	UNP P55085
A	384	HIS	-	expression tag	UNP P55085
A	385	HIS	-	expression tag	UNP P55085
A	386	HIS	-	expression tag	UNP P55085
A	387	HIS	-	expression tag	UNP P55085
A	388	HIS	-	expression tag	UNP P55085
A	389	HIS	-	expression tag	UNP P55085
A	390	HIS	-	expression tag	UNP P55085

- Molecule 2 is ({S})-(4-fluoranyl-2-propyl-phenyl)-(1 {H}-imidazol-2-yl)methanol (three-letter code: 8TZ) (formula: $C_{13}H_{15}FN_2O$).

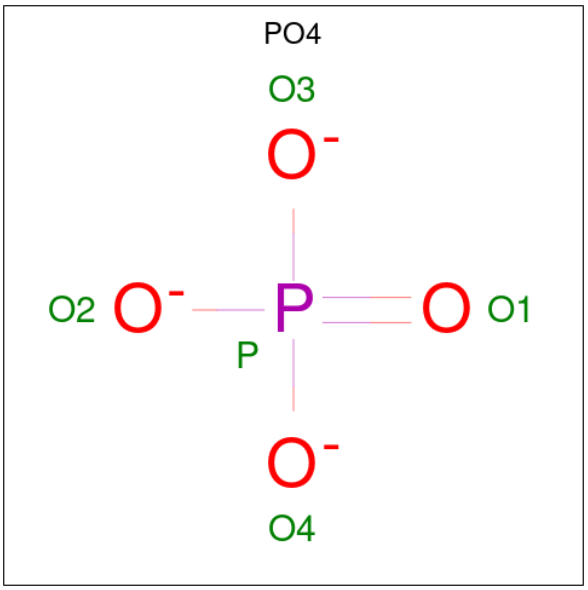


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	F	N	O		
2	A	1	17	13	1	2	1	0	0

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

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

- Molecule 4 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	P		
4	A	1	5	4	1	0	0

- Molecule 5 is water.

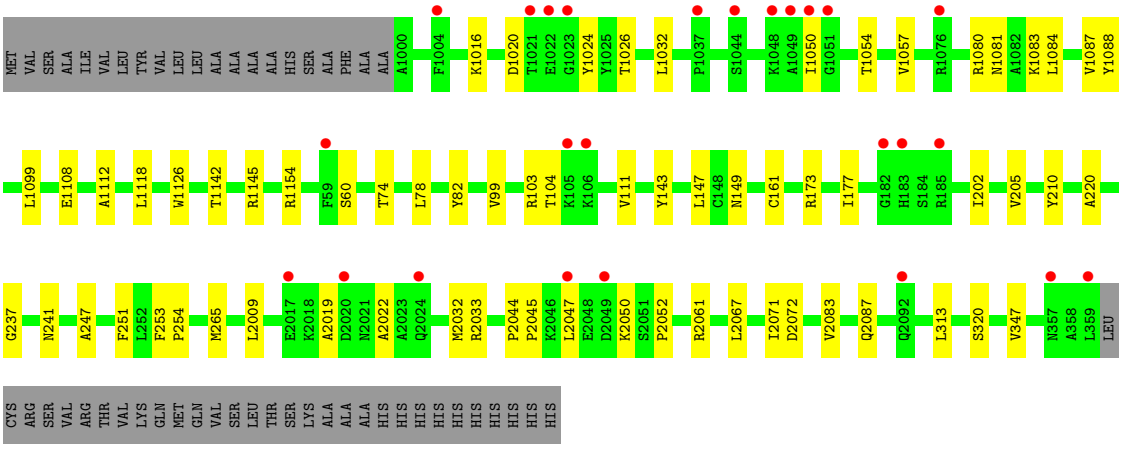
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	10	10	10	0	0

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: Lysozyme,Proteinase-activated receptor 2,Soluble cytochrome b562,Proteinase-activated receptor 2

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	37.11Å 62.26Å 87.34Å 104.88° 90.88° 96.89°	Depositor
Resolution (Å)	34.34 – 2.80 34.34 – 2.80	Depositor EDS
% Data completeness (in resolution range)	96.7 (34.34-2.80) 96.7 (34.34-2.80)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.20 (at 2.81Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.222 , 0.264 0.231 , 0.267	Depositor DCC
R_{free} test set	858 reflections (4.81%)	wwPDB-VP
Wilson B-factor (Å^2)	51.4	Xtriage
Anisotropy	1.020	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 42.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	4531	wwPDB-VP
Average B, all atoms (Å^2)	62.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.47% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: NA, 8TZ, PO4

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	0.23	0/4596	0.38	0/6239

There are no bond length outliers.

There are no bond angle outliers.

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	4498	0	4605	34	0
2	A	17	0	0	1	0
3	A	1	0	0	0	0
4	A	5	0	0	0	0
5	A	10	0	0	0	0
All	All	4531	0	4605	34	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 4.

All (34) 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:2052:PRO:O	1:A:2061:ARG:NH2	2.27	0.67
1:A:1016:LYS:HG2	1:A:1057:VAL:HG22	1.80	0.63
1:A:1026:THR:HG22	1:A:1032:LEU:HA	1.85	0.59
1:A:173:ARG:HD2	1:A:347:VAL:HG13	1.86	0.58
1:A:1126:TRP:HB3	1:A:1154:ARG:HA	1.87	0.57
1:A:1081:ASN:HB3	1:A:1084:LEU:HB2	1.87	0.57
1:A:2009:LEU:HD12	1:A:2032:MET:HG3	1.88	0.55
1:A:60:SER:HB3	1:A:220:ALA:HB3	1.89	0.55
1:A:99:VAL:HA	1:A:103:ARG:HB2	1.90	0.53
1:A:1020:ASP:OD1	1:A:1024:TYR:N	2.42	0.52
1:A:237:GLY:O	1:A:241:ASN:ND2	2.43	0.52
1:A:143:TYR:HB3	1:A:147:LEU:HD23	1.93	0.50
1:A:1142:THR:HB	1:A:1145:ARG:HB3	1.95	0.49
1:A:1080:ARG:NH2	1:A:1108:GLU:OE1	2.43	0.47
1:A:82:TYR:OH	2:A:2201:8TZ:N1	2.47	0.47
1:A:161:CYS:SG	1:A:202:ILE:HD12	2.55	0.47
1:A:2019:ALA:O	1:A:2087:GLN:NE2	2.48	0.47
1:A:2067:LEU:O	1:A:2071:ILE:HG13	2.15	0.46
1:A:202:ILE:HA	1:A:205:VAL:HG12	1.98	0.46
1:A:247:ALA:O	1:A:251:PHE:HB3	2.16	0.46
1:A:1088:TYR:HD1	1:A:1099:LEU:HD23	1.82	0.44
1:A:253:PHE:HB3	1:A:254:PRO:HD3	1.99	0.44
1:A:2033:ARG:NH1	1:A:2072:ASP:OD1	2.51	0.44
1:A:2047:LEU:HB3	1:A:2050:LYS:HB2	2.00	0.44
1:A:2022:ALA:HB2	1:A:2083:VAL:HG22	2.00	0.44
1:A:1087:VAL:HG21	1:A:1118:LEU:HB3	1.99	0.43
1:A:1083:LYS:HE3	1:A:1112:ALA:HB1	2.00	0.43
1:A:104:THR:HG21	1:A:111:VAL:HG21	2.00	0.42
1:A:177:ILE:HD12	1:A:265:MET:HG3	2.02	0.42
1:A:313:LEU:HB2	1:A:320:SER:HB3	2.01	0.42
1:A:2044:PRO:HA	1:A:2045:PRO:HD3	1.90	0.42
1:A:1050:ILE:HG13	1:A:1054:THR:HG21	2.02	0.41
1:A:149:ASN:ND2	1:A:210:TYR:O	2.45	0.41
1:A:74:THR:HA	1:A:78:LEU:HG	2.03	0.40

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	565/619 (91%)	545 (96%)	20 (4%)	0	100 100

There are no Ramachandran outliers to report.

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	480/521 (92%)	480 (100%)	0	100 100

There are no protein residues with a non-rotameric sidechain to report.

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	158	ASN

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)

Of 3 ligands modelled in this entry, 1 is monoatomic - leaving 2 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	$\#\|Z\| > 2$
4	PO4	A	2203	-	4,4,4	0.90	0	6,6,6	0.45	0
2	8TZ	A	2201	-	15,18,18	1.24	2 (13%)	15,24,24	1.37	2 (13%)

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
2	8TZ	A	2201	-	-	1/7/11/11	0/2/2/2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	2201	8TZ	C1-C2	3.12	1.42	1.37
2	A	2201	8TZ	C3-C2	2.07	1.41	1.37

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	2201	8TZ	C7-C6-C5	-3.33	119.97	123.73
2	A	2201	8TZ	C3-C2-C1	-2.44	120.12	123.29

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	2201	8TZ	C5-C6-C7-C8

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	2201	8TZ	1	0

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	567/619 (91%)	0.12	25 (4%) 34 24	38, 59, 98, 116	0

All (25) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1023	GLY	5.3
1	A	1021	THR	5.1
1	A	182	GLY	4.7
1	A	183	HIS	3.7
1	A	2020	ASP	3.5
1	A	1048	LYS	3.4
1	A	1037	PRO	3.3
1	A	1051	GLY	3.0
1	A	185	ARG	3.0
1	A	1049	ALA	2.9
1	A	105	LYS	2.7
1	A	1022	GLU	2.7
1	A	1004	PHE	2.6
1	A	1050	ILE	2.5
1	A	2024	GLN	2.5
1	A	2047	LEU	2.5
1	A	2092	GLN	2.4
1	A	106	LYS	2.4
1	A	1044	SER	2.3
1	A	2017	GLU	2.3
1	A	357	ASN	2.2
1	A	59	PHE	2.2
1	A	359	LEU	2.0
1	A	2049	ASP	2.0
1	A	1076	ARG	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(Å^2)	Q<0.9
3	NA	A	2202	1/1	0.79	0.18	49,49,49,49	0
4	PO4	A	2203	5/5	0.81	0.30	68,75,90,91	0
2	8TZ	A	2201	17/17	0.95	0.20	38,41,45,48	0

6.5 Other polymers (i)

There are no such residues in this entry.