WORLDWIDE PDB

PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Feb 13, 2017 – 08:55 pm GMT

PDB ID	:	4X6Y
Title	:	Human soluble epoxide hydrolase in complex with a cyclopropyl urea derivative
Authors	:	Chiyo, N.; Takai, K.; Ishii, T.
Deposited on	:	2014-12-09
Resolution	:	2.10 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

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

|		
MolProbity	:	4.02b-467
Mogul	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

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 2.10 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.
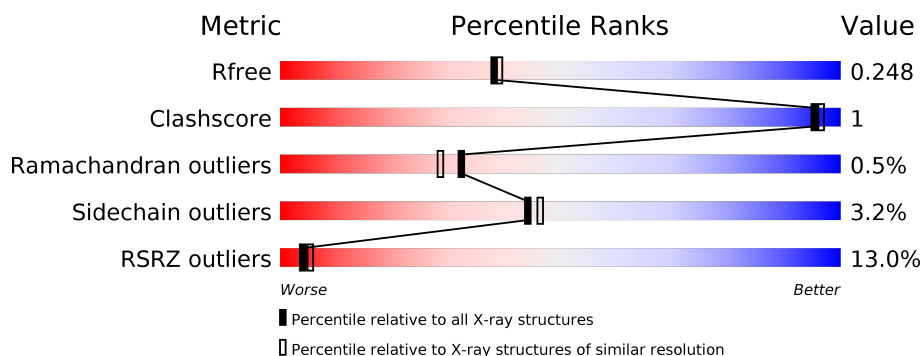
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	100719	4243 (2.10-2.10)
Clashscore	112137	4788 (2.10-2.10)
Ramachandran outliers	110173	4740 (2.10-2.10)
Sidechain outliers	110143	4741 (2.10-2.10)
RSRZ outliers	101464	4275 (2.10-2.10)

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 on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. 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	336	13% / 87% / 6% · 6%
1	B	336	12% / 89% / · · 6%

2　Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 5230 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 Bifunctional epoxide hydrolase 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	316	Total	C	N	O	S	0	0	0
			2546	1636	426	461	23			
1	B	316	Total	C	N	O	S	0	0	0
			2546	1636	426	461	23			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	220	MET	-	expression tag	UNP P34913
A	221	LYS	-	expression tag	UNP P34913
A	222	LYS	-	expression tag	UNP P34913
A	223	GLY	-	expression tag	UNP P34913
A	224	HIS	-	expression tag	UNP P34913
A	225	HIS	-	expression tag	UNP P34913
A	226	HIS	-	expression tag	UNP P34913
A	227	HIS	-	expression tag	UNP P34913
A	228	HIS	-	expression tag	UNP P34913
A	229	HIS	-	expression tag	UNP P34913
B	220	MET	-	expression tag	UNP P34913
B	221	LYS	-	expression tag	UNP P34913
B	222	LYS	-	expression tag	UNP P34913
B	223	GLY	-	expression tag	UNP P34913
B	224	HIS	-	expression tag	UNP P34913
B	225	HIS	-	expression tag	UNP P34913
B	226	HIS	-	expression tag	UNP P34913
B	227	HIS	-	expression tag	UNP P34913
B	228	HIS	-	expression tag	UNP P34913
B	229	HIS	-	expression tag	UNP P34913

- Molecule 2 is 4-phenoxy-N-[(1S,2R)-2-phenylcyclopropyl]piperidine-1-carboxamide (three-letter code: S94) (formula: $C_{21}H_{24}N_2O_2$).

S94
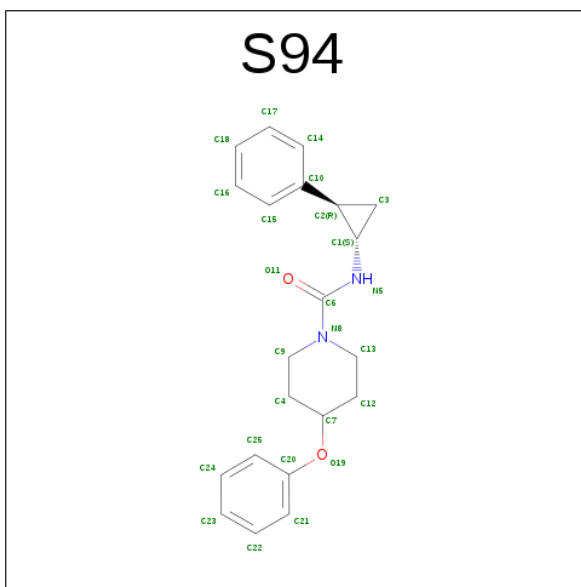
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			25	21	2	2		
2	B	1	Total	C	N	O	0	0
			25	21	2	2		

• Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	49	Total	O	0	0
			49	49		
3	B	39	Total	O	0	0
			39	39		

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA 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: Bifunctional epoxide hydrolase 2

Chain A:

• Molecule 1: Bifunctional epoxide hydrolase 2

Chain B:
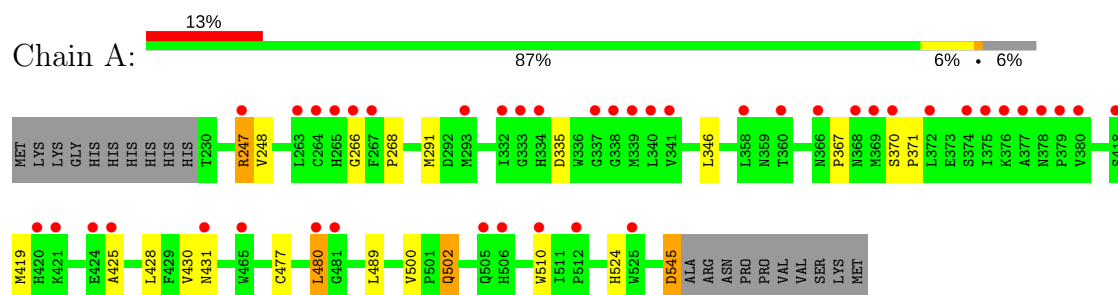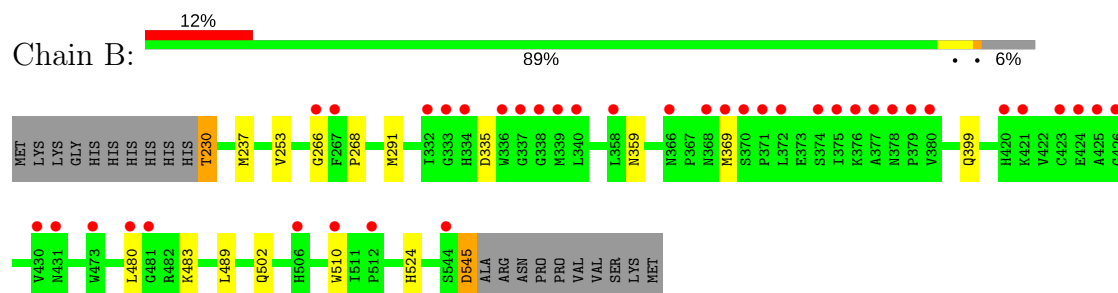
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	129.23Å 80.72Å 87.33Å 90.00° 126.22° 90.00°	Depositor
Resolution (Å)	50.00 – 2.10 33.50 – 2.07	Depositor EDS
% Data completeness (in resolution range)	79.9 (50.00-2.10) 79.9 (33.50-2.07)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.78 (at 2.08Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.198 , 0.244 0.203 , 0.248	Depositor DCC
R_{free} test set	1748 reflections (5.31%)	DCC
Wilson B-factor (Å^2)	35.0	Xtriage
Anisotropy	0.053	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.40 , 45.2	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	0.94	EDS
Total number of atoms	5230	wwPDB-VP
Average B, all atoms (Å^2)	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.14% 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: S94

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	0.39	0/2622	0.61	0/3557
1	B	0.39	0/2622	0.58	0/3557
All	All	0.39	0/5244	0.60	0/7114

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
1	B	0	1
All	All	0	3

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	247	ARG	Peptide
1	A	291	MET	Peptide
1	B	291	MET	Peptide

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	2546	0	2473	9	0
1	B	2546	0	2473	5	0
2	A	25	0	23	1	0
2	B	25	0	23	1	0
3	A	49	0	0	0	0
3	B	39	0	0	1	0
All	All	5230	0	4992	14	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 1.

All (14) 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:545:ASP:OD2	1:A:545:ASP:N	2.25	0.69
1:A:500:VAL:HG12	1:A:502:GLN:HG3	1.85	0.57
1:B:266:GLY:HA3	1:B:335:ASP:HB3	1.89	0.55
1:A:247:ARG:HB2	1:A:248:VAL:HG23	1.89	0.55
1:A:524:HIS:NE2	2:A:601:S94:H16	2.22	0.55
1:B:545:ASP:N	1:B:545:ASP:OD2	2.43	0.51
1:A:266:GLY:HA3	1:A:335:ASP:HB3	1.94	0.49
1:A:477:CYS:O	1:A:480:LEU:HD22	2.17	0.45
1:B:524:HIS:NE2	2:B:601:S94:H16	2.33	0.44
1:B:230:THR:HG23	3:B:718:HOH:O	2.17	0.44
1:A:425:ALA:O	1:A:430:VAL:HG21	2.20	0.42
1:B:237:MET:HB3	1:B:253:VAL:HG13	2.02	0.41
1:A:367:PRO:HG3	1:A:480:LEU:HD21	2.03	0.41
1:A:370:SER:CB	1:A:371:PRO:CD	2.99	0.41

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	314/336 (94%)	301 (96%)	11 (4%)	2 (1%)	28	24
1	B	314/336 (94%)	301 (96%)	12 (4%)	1 (0%)	44	44
All	All	628/672 (94%)	602 (96%)	23 (4%)	3 (0%)	32	28

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	419	MET
1	A	268	PRO
1	B	268	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	278/296 (94%)	270 (97%)	8 (3%)	48	51
1	B	278/296 (94%)	268 (96%)	10 (4%)	40	41
All	All	556/592 (94%)	538 (97%)	18 (3%)	44	46

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	346	LEU
1	A	428	LEU
1	A	431	ASN
1	A	480	LEU
1	A	489	LEU
1	A	502	GLN
1	A	510	TRP
1	A	545	ASP
1	B	230	THR
1	B	359	ASN
1	B	369	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	399	GLN
1	B	480	LEU
1	B	483	LYS
1	B	489	LEU
1	B	502	GLN
1	B	510	TRP
1	B	545	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	378	ASN
1	A	431	ASN
1	A	535	ASN
1	B	384	GLN
1	B	452	GLN
1	B	535	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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

2 ligands are modelled in this entry.

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	#$|Z| > 2$
2	S94	A	601	-	26,28,28	0.94	1 (3%)	32,38,38	1.32	5 (15%)
2	S94	B	601	-	26,28,28	0.89	1 (3%)	32,38,38	1.29	4 (12%)

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	S94	A	601	-	-	0/16/31/31	0/3/4/4
2	S94	B	601	-	-	0/16/31/31	0/3/4/4

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	601	S94	C10-C2	-3.09	1.47	1.51
2	B	601	S94	C10-C2	-2.96	1.47	1.51

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	601	S94	C20-O19-C7	-3.48	111.97	119.06
2	B	601	S94	C4-C7-C12	-3.16	105.99	111.58
2	A	601	S94	C1-N5-C6	-3.07	115.16	122.45
2	B	601	S94	C1-N5-C6	-2.77	115.86	122.45
2	A	601	S94	C3-C2-C10	-2.43	117.85	122.31
2	A	601	S94	C4-C7-C12	-2.37	107.39	111.58
2	B	601	S94	O11-C6-N5	-2.10	117.35	123.15
2	A	601	S94	C12-C13-N8	2.12	113.94	110.82
2	A	601	S94	C13-C12-C7	2.90	113.56	110.30

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	601	S94	1	0
2	B	601	S94	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(Å^2)	Q<0.9
1	A	316/336 (94%)	0.48	43 (13%)	3	5	26, 40, 82, 111	0
1	B	316/336 (94%)	0.49	39 (12%)	5	6	27, 44, 91, 116	0
All	All	632/672 (94%)	0.49	82 (12%)	4	5	26, 42, 85, 116	0

All (82) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	420	HIS	7.9
1	A	374	SER	6.3
1	A	369	MET	6.1
1	A	375	ILE	6.1
1	B	510	TRP	6.0
1	B	379	PRO	5.3
1	A	420	HIS	5.1
1	A	380	VAL	4.9
1	A	510	TRP	4.9
1	B	369	MET	4.7
1	B	376	LYS	4.5
1	A	377	ALA	4.5
1	B	426	GLY	4.5
1	A	372	LEU	4.5
1	A	480	LEU	4.4
1	B	372	LEU	4.4
1	B	377	ALA	4.2
1	B	421	LYS	4.0
1	B	424	GLU	3.9
1	B	480	LEU	3.8
1	A	512	PRO	3.7
1	A	264	CYS	3.7
1	B	430	VAL	3.6
1	B	473	TRP	3.6

Continued on next page...

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

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	380	VAL	3.5
1	A	332	ILE	3.5
1	B	378	ASN	3.5
1	A	506	HIS	3.4
1	B	267	PHE	3.4
1	A	263	LEU	3.4
1	A	424	GLU	3.3
1	B	339	MET	3.3
1	A	341	VAL	3.3
1	B	370	SER	3.3
1	A	370	SER	3.2
1	A	337	GLY	3.2
1	B	333	GLY	3.2
1	B	337	GLY	3.2
1	A	368	ASN	3.2
1	A	378	ASN	3.2
1	A	340	LEU	3.2
1	B	371	PRO	3.2
1	B	366	ASN	3.1
1	A	266	GLY	3.1
1	A	339	MET	2.9
1	B	544	SER	2.9
1	A	481	GLY	2.8
1	B	425	ALA	2.7
1	B	506	HIS	2.7
1	A	425	ALA	2.7
1	B	266	GLY	2.7
1	A	421	LYS	2.7
1	B	358	LEU	2.6
1	B	336	TRP	2.6
1	A	376	LYS	2.5
1	B	332	ILE	2.5
1	B	334	HIS	2.5
1	A	267	PHE	2.4
1	B	423	CYS	2.4
1	A	366	ASN	2.4
1	A	465	TRP	2.4
1	A	360	THR	2.3
1	B	338	GLY	2.3
1	A	418	SER	2.3
1	B	368	ASN	2.3
1	A	505	GLN	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	358	LEU	2.3
1	A	334	HIS	2.3
1	B	375	ILE	2.3
1	B	340	LEU	2.2
1	A	333	GLY	2.2
1	B	512	PRO	2.2
1	B	374	SER	2.2
1	A	338	GLY	2.1
1	A	431	ASN	2.1
1	A	265	HIS	2.1
1	B	431	ASN	2.1
1	B	481	GLY	2.1
1	A	247	ARG	2.0
1	A	379	PRO	2.0
1	A	293	MET	2.0
1	A	525	TRP	2.0

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 carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å^2)	Q<0.9
2	S94	B	601	25/25	0.86	0.26	1.51	47,53,69,70	0
2	S94	A	601	25/25	0.91	0.24	1.46	44,48,62,64	0

6.5 Other polymers (i)

There are no such residues in this entry.