WORLDWIDE PDB

PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Oct 27, 2022 – 10:26 am BST

PDB ID	:	7OB1
Title	:	OLIGOPEPTIDASE B FROM S. PROTEOMACULANS WITH MODIFIED HINGE
Authors	:	Petrenko, D.E.; Nikolaeva, A.Y.; Lazarenko, V.A.; Dorovatovskiy, P.V.; Vlaskina, A.V.; Korzhenevskiy, D.A.; Mikhailova, A.G.; Rakitina, T.V.; Timofeev, V.I.
Deposited on	:	2021-04-20
Resolution	:	2.00 Å(reported)

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

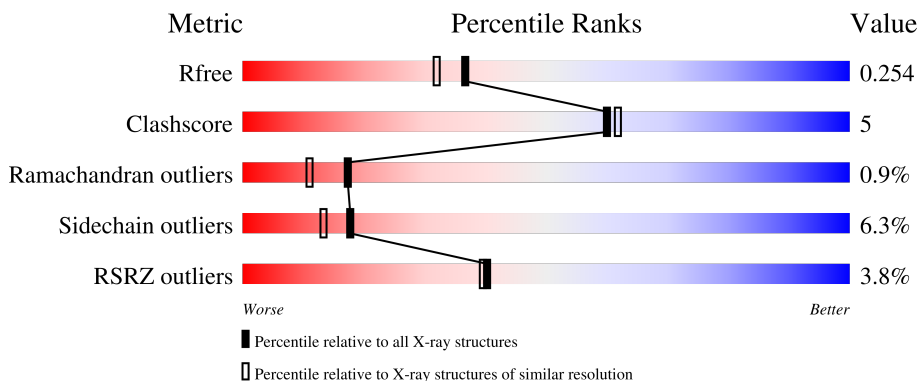|		
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.31.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

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 2.00 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		0.254
Clashscore		5
Ramachandran outliers		0.9%
Sidechain outliers		6.3%
RSRZ outliers		3.8%

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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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	677	4% 85% 13% •

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 5840 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.
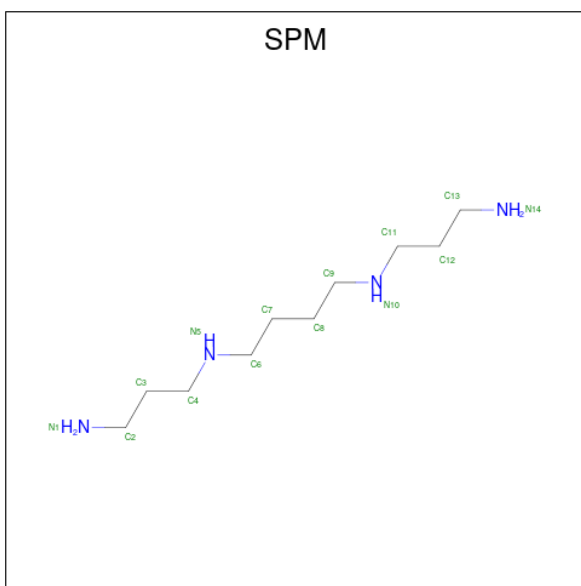
- Molecule 1 is a protein called Oligopeptidase B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	677	Total 5554	C 3527	N 933	O 1081	S 13	0	1	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	HIS	-	expression tag	UNP B3VI58
A	71	GLU	ILE	conflict	UNP B3VI58
A	72	ASN	PRO	conflict	UNP B3VI58
A	73	LEU	GLN	conflict	UNP B3VI58
A	74	TYR	GLN	conflict	UNP B3VI58
A	75	PHE	GLU	conflict	UNP B3VI58
A	76	GLN	HIS	conflict	UNP B3VI58

- Molecule 2 is SPERMINE (three-letter code: SPM) (formula: $C_{10}H_{26}N_4$).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total 14	C 10	N 4	0	0
2	A	1	Total 14	C 10	N 4	0	0
2	A	1	Total 14	C 10	N 4	0	0
2	A	1	Total 14	C 10	N 4	0	0
2	A	1	Total 14	C 10	N 4	0	0

• Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	216	Total 216	O 216	0	0

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: Oligopeptidase B
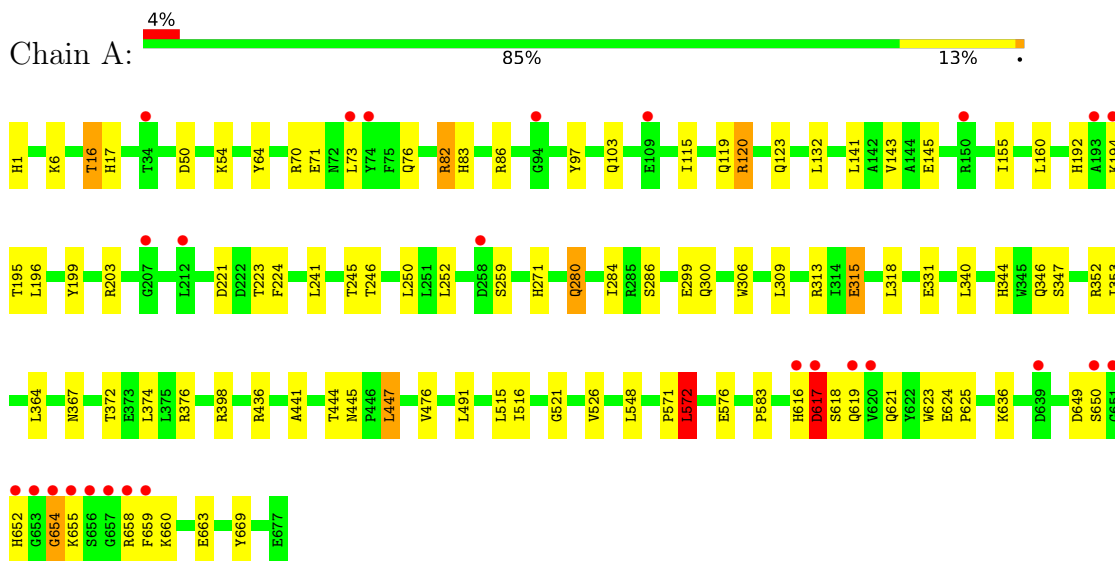
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	73.21Å 101.02Å 108.89Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.00 20.00 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.8 (20.00-2.00) 99.9 (20.00-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	6.74 (at 2.01Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.206 , 0.250 0.214 , 0.254	Depositor DCC
R_{free} test set	2760 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	25.1	Xtriage
Anisotropy	0.024	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< |L| > = 0.51$, $< L^2 > = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	5840	wwPDB-VP
Average B, all atoms (Å^2)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.56% 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: SPM

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	0.75	0/5700	0.89	7/7748 (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	1

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	436	ARG	NE-CZ-NH1	7.83	124.22	120.30
1	A	436	ARG	NE-CZ-NH2	-6.79	116.91	120.30
1	A	70	ARG	NE-CZ-NH1	-6.43	117.09	120.30
1	A	669	TYR	CB-CG-CD1	5.50	124.30	121.00
1	A	70	ARG	CG-CD-NE	-5.21	100.85	111.80
1	A	669	TYR	CB-CG-CD2	-5.09	117.95	121.00
1	A	398	ARG	NE-CZ-NH2	-5.01	117.80	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	618	SER	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	5554	0	5262	51	0
2	A	70	0	130	0	0
3	A	216	0	0	5	0
All	All	5840	0	5392	51	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 5.

All (51) 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:221:ASP:OD1	1:A:223:THR:HG22	1.37	1.21
1:A:367:ASN:HD21	1:A:376:ARG:H	1.23	0.85
1:A:441:ALA:O	1:A:444:THR:HB	1.75	0.85
1:A:372:THR:HG22	1:A:374:LEU:H	1.45	0.81
1:A:221:ASP:OD1	1:A:223:THR:CG2	2.27	0.80
1:A:441:ALA:H	1:A:445:ASN:HD21	1.37	0.73
1:A:571:PRO:O	1:A:572:LEU:HB2	1.88	0.72
1:A:120:ARG:HD3	1:A:145:GLU:OE1	1.93	0.69
1:A:132:LEU:CD2	1:A:141:LEU:HD11	2.23	0.69
1:A:245:THR:HG23	1:A:246:THR:OG1	1.97	0.65
1:A:616:HIS:O	1:A:617:ASP:HB3	1.98	0.62
1:A:516:ILE:HD11	1:A:526:VAL:HG21	1.81	0.61
1:A:83:HIS:HD2	3:A:892:HOH:O	1.85	0.60
1:A:115:ILE:HD11	1:A:120:ARG:HD2	1.84	0.59
1:A:16:THR:HG23	3:A:961:HOH:O	2.03	0.58
1:A:132:LEU:HD23	1:A:141:LEU:HD11	1.84	0.58
1:A:344:HIS:HE1	3:A:912:HOH:O	1.86	0.57
1:A:224:PHE:O	3:A:801:HOH:O	2.17	0.57
1:A:271:HIS:HE1	1:A:286:SER:OG	1.88	0.56
1:A:572:LEU:O	1:A:576:GLU:HG3	2.07	0.55
1:A:624:GLU:HB2	1:A:625:PRO:HD3	1.89	0.54
1:A:64:TYR:OH	1:A:659:PHE:CD1	2.56	0.54
1:A:115:ILE:HD12	1:A:143:VAL:HG11	1.90	0.53
1:A:659:PHE:CE2	1:A:663:GLU:OE1	2.61	0.53

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:447:LEU:HD12	1:A:476:VAL:HB	1.92	0.51
1:A:621:GLN:HB3	1:A:623:TRP:CZ3	2.45	0.51
1:A:120:ARG:CD	1:A:145:GLU:OE1	2.58	0.50
1:A:86:ARG:HH11	1:A:103:GLN:HE21	1.59	0.50
1:A:331:GLU:HG2	1:A:364:LEU:HG	1.95	0.48
1:A:192:HIS:O	1:A:196:LEU:N	2.44	0.48
1:A:17:HIS:HE1	1:A:583:PRO:O	1.98	0.47
1:A:132:LEU:HD21	1:A:141:LEU:HD11	1.97	0.47
1:A:280:GLN:HE21	1:A:280:GLN:HA	1.81	0.46
1:A:284:ILE:HG21	1:A:306:TRP:CH2	2.51	0.46
1:A:313:ARG:HB3	1:A:315:GLU:OE1	2.17	0.45
1:A:340:LEU:HB2	1:A:353:ILE:CG1	2.47	0.44
1:A:97:TYR:CE1	1:A:654:GLY:HA2	2.52	0.44
1:A:344:HIS:HD2	1:A:346:GLN:H	1.66	0.44
1:A:340:LEU:HB2	1:A:353:ILE:HG13	2.00	0.44
1:A:86:ARG:HD2	1:A:103:GLN:HE21	1.82	0.43
1:A:344:HIS:CD2	1:A:346:GLN:H	2.35	0.43
1:A:82:ARG:NH1	3:A:815:HOH:O	2.51	0.43
1:A:132:LEU:CD2	1:A:141:LEU:CD1	2.94	0.43
1:A:447:LEU:HD13	1:A:521:GLY:HA3	1.99	0.43
1:A:132:LEU:HD23	1:A:141:LEU:CD1	2.47	0.42
1:A:616:HIS:O	1:A:617:ASP:CB	2.67	0.42
1:A:143:VAL:O	1:A:155:ILE:HA	2.19	0.42
1:A:344:HIS:CD2	1:A:347:SER:H	2.39	0.41
1:A:203:ARG:NH2	1:A:259:SER:O	2.49	0.40
1:A:659:PHE:CD2	1:A:663:GLU:OE1	2.74	0.40
1:A:621:GLN:HG3	1:A:623:TRP:CH2	2.56	0.40

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	676/677 (100%)	647 (96%)	23 (3%)	6 (1%)	17	11

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	194	LYS
1	A	617	ASP
1	A	572	LEU
1	A	649	ASP
1	A	655	LYS
1	A	654	GLY

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	592/591 (100%)	554 (94%)	38 (6%)	17	13

All (38) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1	HIS
1	A	6	LYS
1	A	16	THR
1	A	50	ASP
1	A	54	LYS
1	A	71[A]	GLU
1	A	71[B]	GLU
1	A	73	LEU
1	A	76	GLN
1	A	82	ARG
1	A	119	GLN
1	A	120	ARG
1	A	123	GLN
1	A	160	LEU
1	A	195	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	199	TYR
1	A	241	LEU
1	A	250	LEU
1	A	252	LEU
1	A	280	GLN
1	A	299	GLU
1	A	300	GLN
1	A	309	LEU
1	A	315	GLU
1	A	318	LEU
1	A	352	ARG
1	A	447	LEU
1	A	491	LEU
1	A	515	LEU
1	A	548	LEU
1	A	572	LEU
1	A	617	ASP
1	A	619	GLN
1	A	636	LYS
1	A	650	SER
1	A	652	HIS
1	A	658	ARG
1	A	660	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (17) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1	HIS
1	A	17	HIS
1	A	83	HIS
1	A	103	GLN
1	A	172	ASN
1	A	181	ASN
1	A	271	HIS
1	A	280	GLN
1	A	300	GLN
1	A	339	GLN
1	A	344	HIS
1	A	367	ASN
1	A	445	ASN
1	A	518	GLN
1	A	585	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	619	GLN
1	A	621	GLN

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 ⓘ

5 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	SPM	A	701	-	13,13,13	0.20	0	12,12,12	0.44	0
2	SPM	A	705	-	13,13,13	0.20	0	12,12,12	0.46	0
2	SPM	A	704	-	13,13,13	0.20	0	12,12,12	0.35	0
2	SPM	A	703	-	13,13,13	0.28	0	12,12,12	0.51	0
2	SPM	A	702	-	13,13,13	0.19	0	12,12,12	0.30	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
2	SPM	A	701	-	-	4/11/11/11	-
2	SPM	A	705	-	-	7/11/11/11	-
2	SPM	A	704	-	-	5/11/11/11	-
2	SPM	A	703	-	-	8/11/11/11	-
2	SPM	A	702	-	-	5/11/11/11	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (29) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	704	SPM	C8-C9-N10-C11
2	A	704	SPM	N10-C11-C12-C13
2	A	701	SPM	C7-C8-C9-N10
2	A	702	SPM	C7-C8-C9-N10
2	A	703	SPM	C2-C3-C4-N5
2	A	701	SPM	C2-C3-C4-N5
2	A	705	SPM	C2-C3-C4-N5
2	A	703	SPM	C12-C11-N10-C9
2	A	702	SPM	N1-C2-C3-C4
2	A	705	SPM	C11-C12-C13-N14
2	A	705	SPM	N5-C6-C7-C8
2	A	703	SPM	C7-C8-C9-N10
2	A	705	SPM	C6-C7-C8-C9
2	A	702	SPM	N5-C6-C7-C8
2	A	703	SPM	C3-C4-N5-C6
2	A	705	SPM	C7-C6-N5-C4
2	A	702	SPM	C11-C12-C13-N14
2	A	703	SPM	C11-C12-C13-N14
2	A	704	SPM	C11-C12-C13-N14
2	A	702	SPM	C7-C6-N5-C4
2	A	704	SPM	N1-C2-C3-C4
2	A	705	SPM	N1-C2-C3-C4
2	A	701	SPM	N10-C11-C12-C13
2	A	703	SPM	C7-C6-N5-C4
2	A	703	SPM	C8-C9-N10-C11
2	A	701	SPM	C8-C9-N10-C11
2	A	704	SPM	C7-C6-N5-C4
2	A	705	SPM	C3-C4-N5-C6
2	A	703	SPM	N1-C2-C3-C4

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(Å^2)	Q<0.9
1	A	677/677 (100%)	0.03	26 (3%)	40	39	14, 26, 52, 142	0

All (26) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	655	LYS	12.6
1	A	652	HIS	11.0
1	A	656	SER	11.0
1	A	654	GLY	8.1
1	A	74	TYR	5.9
1	A	653	GLY	4.8
1	A	94	GLY	4.2
1	A	616	HIS	4.1
1	A	193	ALA	4.1
1	A	650	SER	4.0
1	A	658	ARG	3.7
1	A	659	PHE	3.7
1	A	619	GLN	3.6
1	A	73	LEU	3.6
1	A	657	GLY	3.5
1	A	150	ARG	3.2
1	A	639	ASP	3.1
1	A	207	GLY	2.8
1	A	194	LYS	2.8
1	A	651	GLY	2.6
1	A	620	VAL	2.6
1	A	34	THR	2.5
1	A	109	GLU	2.4
1	A	617	ASP	2.3
1	A	212	LEU	2.3
1	A	258	ASP	2.1

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
2	SPM	A	702	14/14	0.43	0.38	55,65,121,124	0
2	SPM	A	704	14/14	0.53	0.31	60,63,73,74	0
2	SPM	A	705	14/14	0.62	0.30	61,69,81,82	0
2	SPM	A	703	14/14	0.70	0.24	47,51,55,57	0
2	SPM	A	701	14/14	0.79	0.18	42,48,54,54	0

6.5 Other polymers ⓘ

There are no such residues in this entry.