wwPDB X-ray Structure Validation Summary Report (i)

May 7, 2025 – 01:11 pm BST

PDB ID	:	9QZ2 / pdb_00009qz2
Title	:	MINPP1 from Bacteroides thetaiotaomicron A324D/E325N mutant
Authors	:	Li, A.W.H.; Shang, X.Y.; Hemmings, A.M.
Deposited on	:	2025-04-22
Resolution	:	1.95 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 (i)) were used in the production of this report:

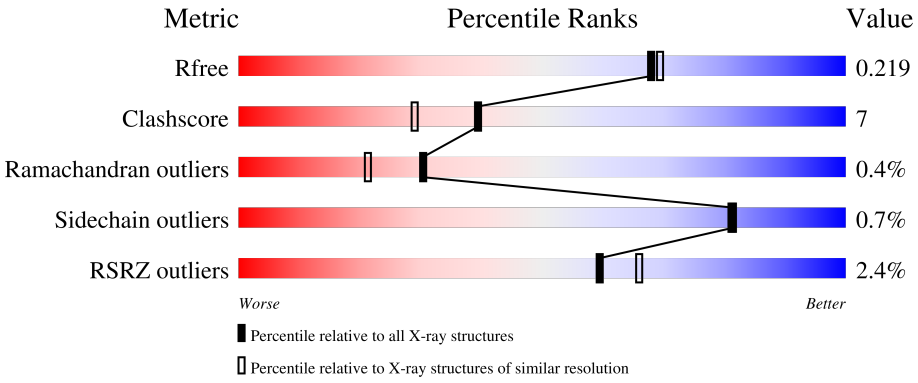
MolProbity	:	4-5-2 with Phenix2.0rc1
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.43.1

1 Overall quality at a glance (i)

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

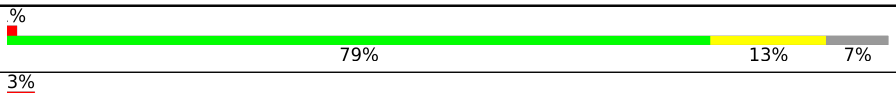
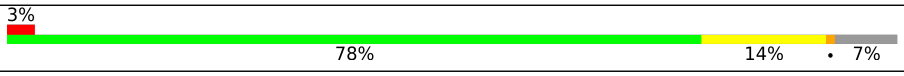
The reported resolution of this entry is 1.95 Å.

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	3187 (1.96-1.96)
Clashscore	180529	3412 (1.96-1.96)
Ramachandran outliers	177936	3390 (1.96-1.96)
Sidechain outliers	177891	3390 (1.96-1.96)
RSRZ outliers	164620	3186 (1.96-1.96)

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	426	
1	B	426	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	PO4	A	501	-	-	X	-
2	PO4	B	501	-	-	X	-

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 6795 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 Multiple inositol polyphosphate phosphatase 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	396	Total	C	N	O	P	S	0	1	0
			3247	2091	556	582	1	17			
1	B	396	Total	C	N	O	P	S	0	0	0
			3240	2086	554	582	1	17			

There are 46 discrepancies between the modelled and reference sequences:

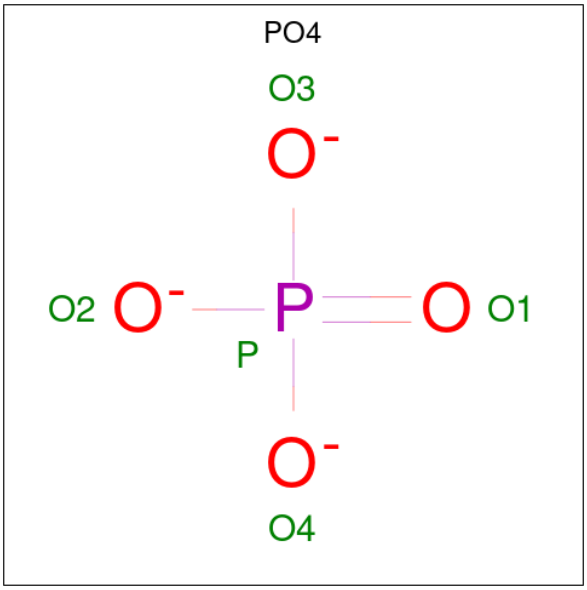
Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	initiating methionine	UNP Q89YI8
A	1	GLY	-	expression tag	UNP Q89YI8
A	2	SER	-	expression tag	UNP Q89YI8
A	3	SER	-	expression tag	UNP Q89YI8
A	4	HIS	-	expression tag	UNP Q89YI8
A	5	HIS	-	expression tag	UNP Q89YI8
A	6	HIS	-	expression tag	UNP Q89YI8
A	7	HIS	-	expression tag	UNP Q89YI8
A	8	HIS	-	expression tag	UNP Q89YI8
A	9	HIS	-	expression tag	UNP Q89YI8
A	10	SER	-	expression tag	UNP Q89YI8
A	11	SER	-	expression tag	UNP Q89YI8
A	12	GLY	-	expression tag	UNP Q89YI8
A	13	LEU	-	expression tag	UNP Q89YI8
A	14	VAL	-	expression tag	UNP Q89YI8
A	15	PRO	-	expression tag	UNP Q89YI8
A	16	ARG	-	expression tag	UNP Q89YI8
A	17	GLY	-	expression tag	UNP Q89YI8
A	18	SER	-	expression tag	UNP Q89YI8
A	19	HIS	-	expression tag	UNP Q89YI8
A	20	MET	-	expression tag	UNP Q89YI8
A	324	ASP	ALA	engineered mutation	UNP Q89YI8
A	325	ASN	GLU	engineered mutation	UNP Q89YI8
B	0	MET	-	initiating methionine	UNP Q89YI8
B	1	GLY	-	expression tag	UNP Q89YI8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	2	SER	-	expression tag	UNP Q89YI8
B	3	SER	-	expression tag	UNP Q89YI8
B	4	HIS	-	expression tag	UNP Q89YI8
B	5	HIS	-	expression tag	UNP Q89YI8
B	6	HIS	-	expression tag	UNP Q89YI8
B	7	HIS	-	expression tag	UNP Q89YI8
B	8	HIS	-	expression tag	UNP Q89YI8
B	9	HIS	-	expression tag	UNP Q89YI8
B	10	SER	-	expression tag	UNP Q89YI8
B	11	SER	-	expression tag	UNP Q89YI8
B	12	GLY	-	expression tag	UNP Q89YI8
B	13	LEU	-	expression tag	UNP Q89YI8
B	14	VAL	-	expression tag	UNP Q89YI8
B	15	PRO	-	expression tag	UNP Q89YI8
B	16	ARG	-	expression tag	UNP Q89YI8
B	17	GLY	-	expression tag	UNP Q89YI8
B	18	SER	-	expression tag	UNP Q89YI8
B	19	HIS	-	expression tag	UNP Q89YI8
B	20	MET	-	expression tag	UNP Q89YI8
B	324	ASP	ALA	engineered mutation	UNP Q89YI8
B	325	ASN	GLU	engineered mutation	UNP Q89YI8

- Molecule 2 is PHOSPHATE ION (CCD ID: PO4) (formula: O_4P).



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

Continued on next page...

Continued from previous page...

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

- Molecule 3 is water.

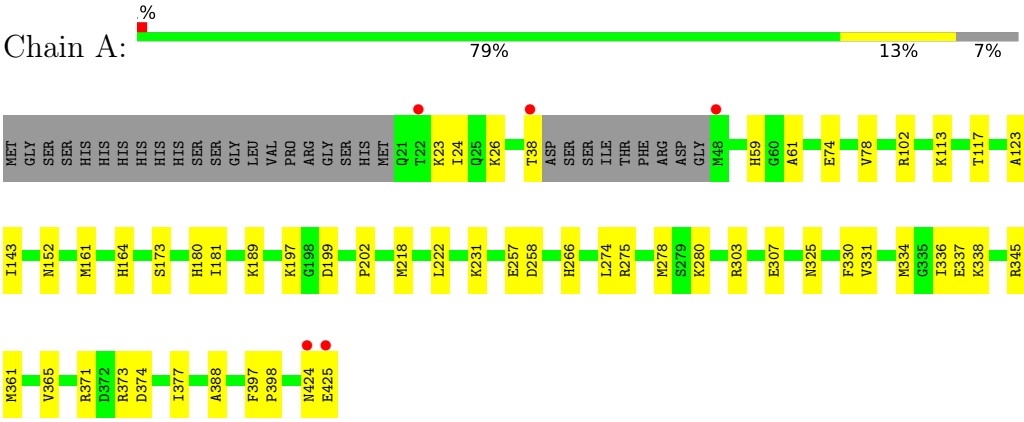
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	162	Total O 162 162	0	0
3	B	136	Total O 136 136	0	0

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.

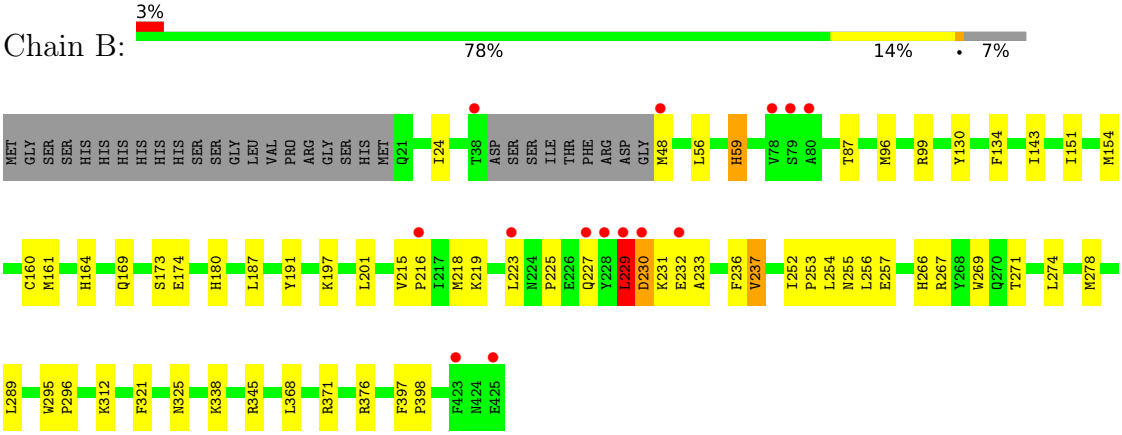
• Molecule 1: Multiple inositol polyphosphate phosphatase 1

Chain A:



• Molecule 1: Multiple inositol polyphosphate phosphatase 1

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	52.47Å 117.91Å 76.11Å 90.00° 108.13° 90.00°	Depositor
Resolution (Å)	26.61 – 1.95 26.61 – 1.95	Depositor EDS
% Data completeness (in resolution range)	98.0 (26.61-1.95) 90.4 (26.61-1.95)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.72 (at 1.95Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.187 , 0.219 0.187 , 0.219	Depositor DCC
R_{free} test set	4408 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å^2)	26.9	Xtriage
Anisotropy	0.859	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 39.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.040 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	6795	wwPDB-VP
Average B, all atoms (Å^2)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.43% 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: NEP, 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	0.35	0/3318	0.56	0/4487
1	B	0.32	0/3307	0.54	2/4472 (0.0%)
All	All	0.34	0/6625	0.55	2/8959 (0.0%)

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.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	56	LEU	CA-C-N	-5.16	118.14	121.65
1	B	56	LEU	C-N-CA	-5.16	118.14	121.65

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	345	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	3247	0	3239	44	1
1	B	3240	0	3232	42	1
2	A	5	0	0	2	0
2	B	5	0	0	2	0
3	A	162	0	0	11	0
3	B	136	0	0	6	0
All	All	6795	0	6471	85	1

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 7.

The worst 5 of 85 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:345:ARG:NH2	3:A:601:HOH:O	1.80	1.02
1:B:24:ILE:HD11	1:B:257:GLU:HB2	1.52	0.90
1:B:215:VAL:HG23	1:B:216:PRO:HD3	1.60	0.84
1:B:338:LYS:HE3	3:B:656:HOH:O	1.84	0.77
1:A:274:LEU:HG	1:A:278:MET:HE2	1.67	0.77

All (1) 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)
1:A:258:ASP:OD2	1:B:376:ARG:NH1[1_455]	2.05	0.15

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	392/426 (92%)	383 (98%)	9 (2%)	0	100 100
1	B	391/426 (92%)	375 (96%)	13 (3%)	3 (1%)	16 8
All	All	783/852 (92%)	758 (97%)	22 (3%)	3 (0%)	30 21

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	230	ASP
1	B	227	GLN
1	B	229	LEU

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	351/376 (93%)	349 (99%)	2 (1%)	84 83
1	B	350/376 (93%)	347 (99%)	3 (1%)	75 75
All	All	701/752 (93%)	696 (99%)	5 (1%)	81 81

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	38	THR
1	A	127	ILE
1	B	197	LYS
1	B	229	LEU
1	B	237	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	255	ASN
1	B	193	ASN
1	B	164	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	25	GLN
1	B	169	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

2 non-standard protein/DNA/RNA residues 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	$\#\|Z\| > 2$
1	NEP	A	59	1	10,14,15	2.16	3 (30%)	5,20,22	2.09	3 (60%)
1	NEP	B	59	1	10,14,15	1.75	3 (30%)	5,20,22	1.85	1 (20%)

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
1	NEP	A	59	1	-	0/5/12/14	0/1/1/1
1	NEP	B	59	1	-	0/5/12/14	0/1/1/1

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	59	NEP	CD2-CG	4.18	1.42	1.36
1	A	59	NEP	P-O2P	3.90	1.62	1.54
1	B	59	NEP	CD2-CG	3.62	1.41	1.36
1	B	59	NEP	P-O1P	2.57	1.60	1.54
1	A	59	NEP	CD2-NE2	2.40	1.43	1.39

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	59	NEP	O1P-P-O2P	-3.32	93.51	106.57
1	B	59	NEP	O1P-P-O2P	-3.14	94.24	106.57
1	A	59	NEP	O1P-P-O3P	-2.03	109.04	113.44
1	A	59	NEP	O2P-P-O3P	2.02	117.81	113.44

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	B	59	NEP	1	0

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

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	PO4	B	501	-	4,4,4	1.18	0	6,6,6	0.87	0
2	PO4	A	501	-	4,4,4	1.24	0	6,6,6	0.79	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	501	PO4	2	0
2	A	501	PO4	2	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

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, 95^{th} 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	395/426 (92%)	0.05	5 (1%)	74	79	21, 35, 49, 65	1 (0%)
1	B	395/426 (92%)	0.24	14 (3%)	47	54	27, 38, 62, 81	0
All	All	790/852 (92%)	0.14	19 (2%)	59	66	21, 36, 57, 81	1 (0%)

The worst 5 of 19 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	229	LEU	4.0
1	B	228	TYR	3.2
1	B	38	THR	2.9
1	B	227	GLN	2.9
1	B	232	GLU	2.8

6.2 Non-standard residues in protein, DNA, RNA chains

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, 95^{th} 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
1	NEP	B	59	14/15	0.89	0.11	29,39,47,47	0
1	NEP	A	59	14/15	0.90	0.12	31,35,45,47	0

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, 95^{th} 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	PO4	A	501	5/5	0.84	0.12	49,52,62,71	0
2	PO4	B	501	5/5	0.84	0.11	52,58,63,65	0

6.5 Other polymers

There are no such residues in this entry.