



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 10:02 PM UTC

PDB ID : 5AOU / pdb_00005aou
Title : Structure of the engineered retro-aldolase RA95.5-8F apo
Authors : Obexer, R.; Mittl, P.; Hilvert, D.
Deposited on : 2015-09-11
Resolution : 1.10 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 ⓘ 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	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

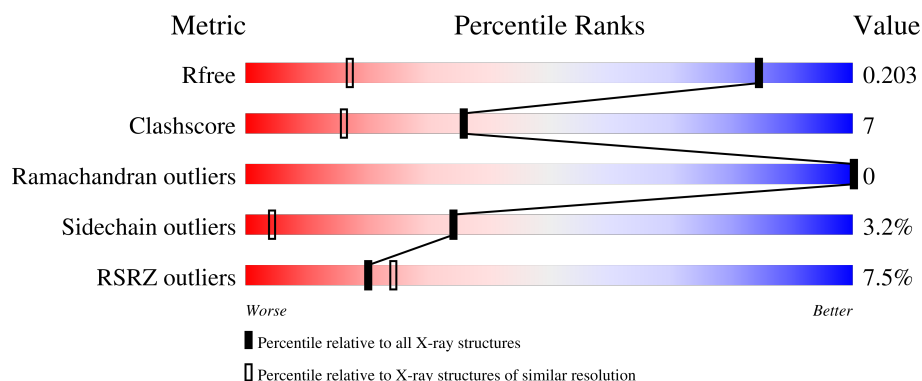
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.10 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}	180053	1625 (1.12-1.08)
Clashscore	190562	1656 (1.12-1.08)
Ramachandran outliers	187476	1616 (1.12-1.08)
Sidechain outliers	187428	1612 (1.12-1.08)
RSRZ outliers	180081	1625 (1.12-1.08)

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	258	

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
3	PO4	A	5006	-	X	-	-

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 2304 atoms, of which 9 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 INDOLE-3-GLYCEROL PHOSPHATE SYNTHASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	242	Total	C	H	N	O	S	0	13	1
			2079	1324	9	348	390	8			

There are 43 discrepancies between the modelled and reference sequences:

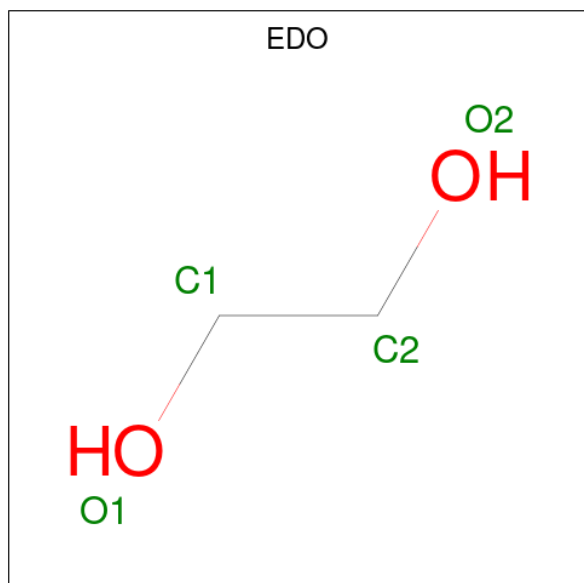
Chain	Residue	Modelled	Actual	Comment	Reference
A	2246	LEU	-	expression tag	UNP Q06121
A	2247	ILE	-	expression tag	UNP Q06121
A	2248	GLU	-	expression tag	UNP Q06121
A	2249	GLY	-	expression tag	UNP Q06121
A	2250	SER	-	expression tag	UNP Q06121
A	2251	LEU	-	expression tag	UNP Q06121
A	2252	GLU	-	expression tag	UNP Q06121
A	2253	HIS	-	expression tag	UNP Q06121
A	2254	HIS	-	expression tag	UNP Q06121
A	2255	HIS	-	expression tag	UNP Q06121
A	2256	HIS	-	expression tag	UNP Q06121
A	2257	HIS	-	expression tag	UNP Q06121
A	2258	HIS	-	expression tag	UNP Q06121
A	1010	GLU	LYS	engineered mutation	UNP Q06121
A	1022	VAL	PHE	engineered mutation	UNP Q06121
A	1023	HIS	ARG	engineered mutation	UNP Q06121
A	1051	TYR	GLU	engineered mutation	UNP Q06121
A	1053	LEU	LYS	engineered mutation	UNP Q06121
A	2070	ALA	SER	engineered mutation	UNP Q06121
A	2072	TYR	PHE	engineered mutation	UNP Q06121
A	2075	PRO	ARG	engineered mutation	UNP Q06121
A	2083	LYS	LEU	engineered mutation	UNP Q06121
A	2090	ASP	ASN	engineered mutation	UNP Q06121
A	2095	MET	THR	engineered mutation	UNP Q06121
A	2110	ASN	LYS	engineered mutation	UNP Q06121
A	2135	GLU	LYS	engineered mutation	UNP Q06121
A	2151	GLY	SER	engineered mutation	UNP Q06121

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	2159	LEU	GLU	engineered mutation	UNP Q06121
A	2178	THR	GLY	engineered mutation	UNP Q06121
A	2180	TYR	ASN	engineered mutation	UNP Q06121
A	2182	MET	ARG	engineered mutation	UNP Q06121
A	2183	ASN	ASP	engineered mutation	UNP Q06121
A	2184	PHE	LEU	engineered mutation	UNP Q06121
A	2187	GLY	LEU	engineered mutation	UNP Q06121
A	2209	PRO	ALA	engineered mutation	UNP Q06121
A	2210	LEU	GLU	engineered mutation	UNP Q06121
A	2211	LEU	SER	engineered mutation	UNP Q06121
A	2212	ASP	GLY	engineered mutation	UNP Q06121
A	2213	PHE	ILE	engineered mutation	UNP Q06121
A	2214	PHE	SER	engineered mutation	UNP Q06121
A	2216	PRO	ARG	engineered mutation	UNP Q06121
A	2231	MET	LEU	engineered mutation	UNP Q06121
A	2233	SER	GLY	engineered mutation	UNP Q06121

- Molecule 2 is 1,2-ETHANEDIOL (CCD ID: EDO) (formula: C₂H₆O₂).



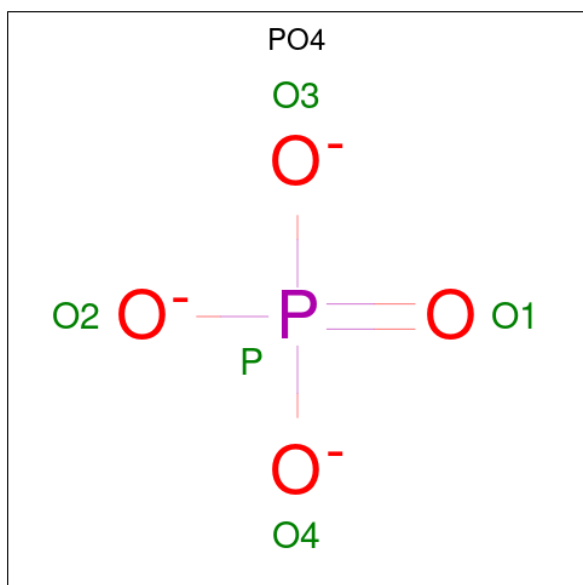
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0

Continued on next page...

Continued from previous page...

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

- Molecule 3 is PHOSPHATE ION (CCD ID: PO4) (formula: O₄P).



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

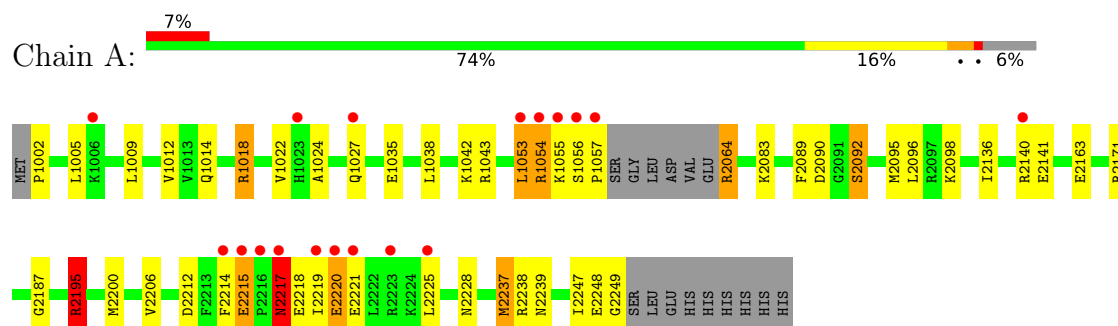
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	190	Total	O	0	0
			190	190		

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: INDOLE-3-GLYCEROL PHOSPHATE SYNTHASE



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	75.09Å 84.76Å 37.66Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	56.20 – 1.10 56.20 – 1.10	Depositor EDS
% Data completeness (in resolution range)	88.8 (56.20-1.10) 90.3 (56.20-1.10)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.09 (at 1.10Å)	Xtriage
Refinement program	SHELXL	Depositor
R, R_{free}	0.169 , 0.219 0.164 , 0.203	Depositor DCC
R_{free} test set	4272 reflections (4.81%)	wwPDB-VP
Wilson B-factor (Å ²)	11.3	Xtriage
Anisotropy	0.225	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 60.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.98	EDS
Total number of atoms	2304	wwPDB-VP
Average B, all atoms (Å ²)	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.08% 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: PO4, MHO, 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	0.90	1/2093 (0.0%)	1.64	40/2822 (1.4%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	2248	GLU	C-N	-5.30	1.25	1.33

The worst 5 of 40 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1027	GLN	OE1-CD-NE2	-13.19	109.41	122.60
1	A	1035	GLU	CB-CG-CD	11.18	131.61	112.60
1	A	1054	ARG	NE-CZ-NH2	9.71	127.94	119.20
1	A	1054	ARG	NE-CZ-NH1	-9.16	112.34	121.50
1	A	1056	SER	O-C-N	9.00	127.47	121.71

There are no chirality outliers.

There are no planarity outliers.

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	2070	9	2114	31	0
2	A	20	0	30	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	15	0	0	0	0
4	A	190	0	0	7	0
All	All	2295	9	2144	31	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 7.

The worst 5 of 31 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:2171:ARG:HH12	2:A:3009:EDO:H22	1.30	0.95
1:A:1054:ARG:HD2	4:A:4079:HOH:O	1.81	0.80
1:A:2215[B]:GLU:HB2	1:A:2217:ASN:OD1	1.83	0.77
1:A:2215[A]:GLU:HA	1:A:2215[A]:GLU:OE1	1.88	0.72
1:A:2092[B]:SER:HB3	4:A:4103:HOH:O	1.98	0.64

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	250/258 (97%)	248 (99%)	2 (1%)	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	231/234 (99%)	223 (96%)	8 (4%)	32 3

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2247	ILE
1	A	2220	GLU
1	A	2195	ARG
1	A	2092[B]	SER
1	A	2217	ASN

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	2194	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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	MHO	A	2237	1	6,8,9	1.60	1 (16%)	3,9,11	1.90	1 (33%)

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	MHO	A	2237	1	-	2/6/7/9	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	2237	MHO	OD1-SD	3.88	1.61	1.50

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	2237	MHO	CE-SD-CG	3.11	105.22	97.45

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	2237	MHO	CB-CG-SD-CE
1	A	2237	MHO	CB-CG-SD-OD1

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	2237	MHO	1	0

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

8 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	$\# Z > 2$
3	PO4	A	5005	-	4,4,4	1.63	1 (25%)	6,6,6	0.92	0
2	EDO	A	3009	-	3,3,3	0.50	0	2,2,2	0.17	0
3	PO4	A	5001	-	4,4,4	0.30	0	6,6,6	0.68	0
2	EDO	A	3003	-	3,3,3	0.94	0	2,2,2	0.22	0
3	PO4	A	5006	-	4,4,4	1.72	1 (25%)	6,6,6	2.42	3 (50%)
2	EDO	A	3004	-	3,3,3	0.39	0	2,2,2	1.36	0
2	EDO	A	3007	-	3,3,3	0.55	0	2,2,2	0.42	0
2	EDO	A	3005	-	3,3,3	0.50	0	2,2,2	0.62	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	EDO	A	3009	-	-	1/1/1/1	-
2	EDO	A	3003	-	-	0/1/1/1	-
2	EDO	A	3004	-	-	1/1/1/1	-
2	EDO	A	3007	-	-	0/1/1/1	-
2	EDO	A	3005	-	-	0/1/1/1	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	5006	PO4	P-O3	2.82	1.62	1.54
3	A	5005	PO4	P-O3	2.42	1.61	1.54

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	5006	PO4	O4-P-O2	4.05	120.51	107.91
3	A	5006	PO4	O4-P-O3	-3.27	97.73	107.91
3	A	5006	PO4	O3-P-O2	-2.74	99.39	107.91

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	3004	EDO	O1-C1-C2-O2
2	A	3009	EDO	O1-C1-C2-O2

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	3009	EDO	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 [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	241/258 (93%)	0.22	18 (7%) 20 24	5, 14, 41, 68	13 (5%)

The worst 5 of 18 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1053	LEU	5.1
1	A	2216	PRO	4.2
1	A	1023	HIS	3.4
1	A	1027	GLN	2.8
1	A	1056	SER	2.7

6.2 Non-standard residues in protein, DNA, RNA chains [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
1	MHO	A	2237	9/10	0.94	0.12	14,35,44,47	1

6.3 Carbohydrates [i](#)

There are no oligosaccharides 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($\text{\AA}^2$)	Q<0.9
2	EDO	A	3009	4/4	0.74	0.21	72,74,77,78	0
2	EDO	A	3004	4/4	0.86	0.18	22,32,34,45	4
2	EDO	A	3003	4/4	0.93	0.10	14,20,20,25	0
3	PO4	A	5006	5/5	0.96	0.09	20,23,45,48	0
2	EDO	A	3005	4/4	0.97	0.07	14,16,22,24	0
3	PO4	A	5005	5/5	0.97	0.07	15,21,33,38	5
2	EDO	A	3007	4/4	0.97	0.08	14,23,28,33	4
3	PO4	A	5001	5/5	0.99	0.06	15,19,28,29	5

6.5 Other polymers ⓘ

There are no such residues in this entry.