



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 1, 2026 – 08:50 AM UTC

PDB ID : 9CF4 / pdb_00009cf4
Title : Crystal structure of the dimeric transaminase DoeD from *C. salexigens* DSM 3043.
Authors : Skogvold, A.C.; Brakestad, H.T.; Erlandsen, H.; Leiros, I.
Deposited on : 2024-06-27
Resolution : 1.50 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

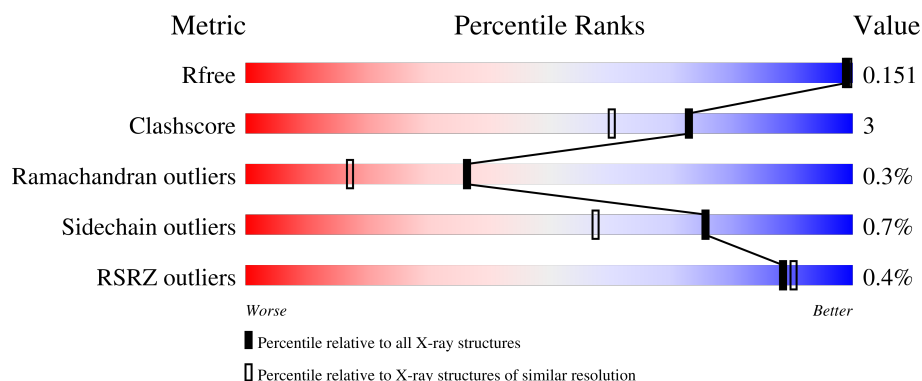
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION



The reported resolution of this entry is 1.50 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	4037 (1.50-1.50)
Clashscore	190562	4235 (1.50-1.50)
Ramachandran outliers	187476	4153 (1.50-1.50)
Sidechain outliers	187428	4150 (1.50-1.50)
RSRZ outliers	180081	4039 (1.50-1.50)

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	474	 92% 6% .
2	B	474	 93% . .

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	GOL	A	502	-	-	X	-

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 15609 atoms, of which 7145 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 Aminotransferase.

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
1	A	465	Total	C	H	N	O	P	S	0	13	0
			7145	2277	3537	637	676	1	17			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	469	HIS	-	expression tag	UNP Q1QTY7
A	470	HIS	-	expression tag	UNP Q1QTY7
A	471	HIS	-	expression tag	UNP Q1QTY7
A	472	HIS	-	expression tag	UNP Q1QTY7
A	473	HIS	-	expression tag	UNP Q1QTY7
A	474	HIS	-	expression tag	UNP Q1QTY7

- Molecule 2 is a protein called Aminotransferase.

Mol	Chain	Residues	Atoms								ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S					
2	B	465	7174	2285	3553	640	678	18			0	17	0

There are 6 discrepancies between the modelled and reference sequences:

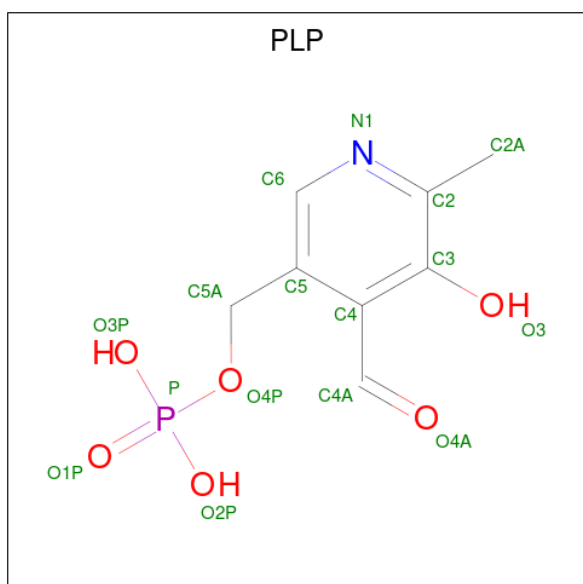
Chain	Residue	Modelled	Actual	Comment	Reference
B	469	HIS	-	expression tag	UNP Q1QTY7
B	470	HIS	-	expression tag	UNP Q1QTY7
B	471	HIS	-	expression tag	UNP Q1QTY7
B	472	HIS	-	expression tag	UNP Q1QTY7
B	473	HIS	-	expression tag	UNP Q1QTY7
B	474	HIS	-	expression tag	UNP Q1QTY7

- Molecule 3 is GLYCEROL (CCD ID: GOL) (formula: C₃H₈O₃).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	H	O	0	0
			14	3	8	3		
3	A	1	Total	C	H	O	0	0
			13	3	7	3		
3	A	1	Total	C	H	O	0	0
			14	3	8	3		
3	A	1	Total	C	H	O	0	0
			14	3	8	3		
3	B	1	Total	C	H	O	0	1
			28	6	16	6		

- Molecule 4 is PYRIDOXAL-5'-PHOSPHATE (CCD ID: PLP) (formula: $C_8H_{10}NO_6P$).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
4	B	1	Total	C	H	N	O	P	0	0
			23	8	8	1	5	1		

- Molecule 5 is water.

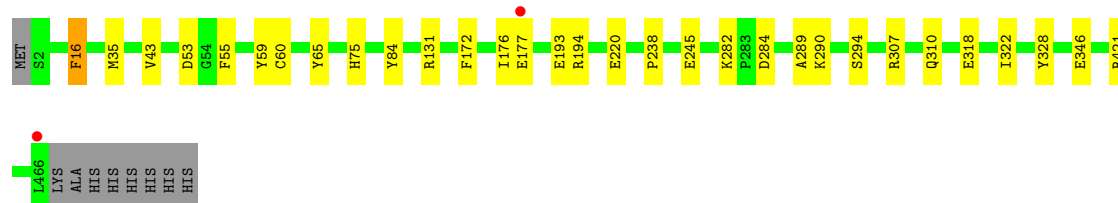
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	627	Total	O	0	0
			627	627		
5	B	557	Total	O	0	0
			557	557		

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: Aminotransferase

Chain A:  92% 6% .



• Molecule 2: Aminotransferase

Chain B:  93% . .



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	82.38Å 94.10Å 219.98Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	35.91 – 1.50 35.91 – 1.50	Depositor EDS
% Data completeness (in resolution range)	99.7 (35.91-1.50) 99.7 (35.91-1.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.16 (at 1.50Å)	Xtriage
Refinement program	PHENIX (1.13_2998: ???)	Depositor
R, R_{free}	0.112 , 0.151 0.114 , 0.151	Depositor DCC
R_{free} test set	2291 reflections (1.68%)	wwPDB-VP
Wilson B-factor (Å ²)	16.0	Xtriage
Anisotropy	0.029	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.42 , 53.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.98	EDS
Total number of atoms	15609	wwPDB-VP
Average B, all atoms (Å ²)	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.78% 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: LLP, GOL, PLP

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	0.50	4/3706 (0.1%)	0.64	5/5019 (0.1%)
2	B	0.35	0/3754	0.59	0/5083
All	All	0.43	4/7460 (0.1%)	0.62	5/10102 (0.0%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	177	GLU	C-N	14.29	1.55	1.33
1	A	176	ILE	C-N	-11.37	1.18	1.33
1	A	131	ARG	C-O	-5.33	1.18	1.24
1	A	310	GLN	C-O	-5.31	1.18	1.24

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	177	GLU	O-C-N	12.96	138.82	122.87
1	A	176	ILE	CA-C-N	5.78	128.92	120.71
1	A	176	ILE	C-N-CA	5.78	128.92	120.71
1	A	177	GLU	CA-C-N	-5.12	113.85	122.25
1	A	177	GLU	C-N-CA	-5.12	113.85	122.25

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	3608	3537	3481	23	0
2	B	3621	3553	3483	22	0
3	A	24	31	32	9	0
3	B	12	16	16	0	0
4	B	15	8	6	3	0
5	A	627	0	0	12	7
5	B	557	0	0	7	6
All	All	8464	7145	7018	46	10

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 3.

The worst 5 of 46 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)
2:B:290:LYS:NZ	4:B:502:PLP:C4A	1.91	1.33
2:B:290:LYS:HZ1	4:B:502:PLP:C4A	1.58	1.04
2:B:290:LYS:HZ2	4:B:502:PLP:C4A	1.61	1.03
2:B:95[B]:GLN:NE2	5:B:602:HOH:O	2.04	0.90
1:A:220:GLU:OE2	3:A:504:GOL:H11	1.80	0.82

The worst 5 of 10 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)
5:A:1053:HOH:O	5:B:1057:HOH:O[5_445]	1.88	0.32
5:A:1094:HOH:O	5:A:1143:HOH:O[8_445]	1.91	0.29
5:A:609:HOH:O	5:A:665:HOH:O[5_455]	2.01	0.19
5:B:954:HOH:O	5:B:1125:HOH:O[5_555]	2.09	0.11
5:B:660:HOH:O	5:B:896:HOH:O[3_554]	2.12	0.08

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	475/474 (100%)	457 (96%)	17 (4%)	1 (0%)	43	22
2	B	480/474 (101%)	459 (96%)	19 (4%)	2 (0%)	30	12
All	All	955/948 (101%)	916 (96%)	36 (4%)	3 (0%)	36	17

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	290	LYS
1	A	289	ALA
2	B	289	ALA

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	366/362 (101%)	363 (99%)	3 (1%)	73	54
2	B	371/363 (102%)	369 (100%)	2 (0%)	81	66
All	All	737/725 (102%)	732 (99%)	5 (1%)	76	57

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

Mol	Chain	Res	Type
1	A	16	PHE
1	A	59	TYR
1	A	322	ILE
2	B	16	PHE
2	B	59	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	5	GLN
1	A	97	GLN
2	B	77	GLN
2	B	97	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	LLP	A	290	1	23,24,25	2.30	6 (26%)	25,32,34	3.65	8 (32%)

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	LLP	A	290	1	-	4/16/17/19	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	290	LLP	C4'-NZ	7.90	1.53	1.27
1	A	290	LLP	P-OP4	3.28	1.70	1.60
1	A	290	LLP	CE-NZ	2.95	1.53	1.46
1	A	290	LLP	OP4-C5'	-2.76	1.34	1.44
1	A	290	LLP	C4-C5	-2.74	1.38	1.42

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	290	LLP	CE-NZ-C4'	11.90	156.84	118.72
1	A	290	LLP	OP3-P-OP4	11.08	135.56	106.67
1	A	290	LLP	OP4-P-OP1	-3.34	97.41	106.44
1	A	290	LLP	OP3-P-OP2	-3.25	95.60	107.80
1	A	290	LLP	O3-C3-C2	2.97	123.74	117.58

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	290	LLP	C4-C4'-NZ-CE
1	A	290	LLP	CG-CD-CE-NZ
1	A	290	LLP	C3-C4-C4'-NZ
1	A	290	LLP	C5-C4-C4'-NZ

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

7 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	GOL	B	501[A]	-	5,5,5	0.70	0	5,5,5	0.63	0
3	GOL	A	502	-	5,5,5	1.10	1 (20%)	5,5,5	1.07	0
4	PLP	B	502	-	15,15,16	2.00	7 (46%)	21,22,23	4.11	6 (28%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	GOL	A	504	-	5,5,5	0.85	0	5,5,5	1.08	0
3	GOL	A	501	-	5,5,5	0.69	0	5,5,5	0.64	0
3	GOL	B	501[B]	-	5,5,5	0.20	0	5,5,5	0.83	0
3	GOL	A	503	-	5,5,5	0.64	0	5,5,5	1.39	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
3	GOL	B	501[A]	-	-	0/4/4/4	-
3	GOL	A	502	-	-	1/4/4/4	-
4	PLP	B	502	-	-	2/6/6/8	0/1/1/1
3	GOL	A	504	-	-	3/4/4/4	-
3	GOL	A	501	-	-	2/4/4/4	-
3	GOL	B	501[B]	-	-	0/4/4/4	-
3	GOL	A	503	-	-	2/4/4/4	-

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	502	PLP	P-O2P	3.46	1.67	1.54
4	B	502	PLP	P-O4P	3.06	1.70	1.60
4	B	502	PLP	C5-C4	-2.57	1.37	1.40
4	B	502	PLP	O4P-C5A	-2.45	1.35	1.44
4	B	502	PLP	O3-C3	-2.38	1.31	1.36

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	502	PLP	O3P-P-O4P	16.90	150.74	106.67
4	B	502	PLP	O3P-P-O1P	-4.99	91.41	110.83
4	B	502	PLP	O2P-P-O4P	-2.85	99.25	106.67
4	B	502	PLP	O4P-P-O1P	-2.69	99.17	106.44
3	A	503	GOL	C3-C2-C1	-2.57	102.36	111.80

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	501	GOL	C1-C2-C3-O3
4	B	502	PLP	C5A-O4P-P-O1P
4	B	502	PLP	C5A-O4P-P-O3P
3	A	502	GOL	O1-C1-C2-C3
3	A	503	GOL	O1-C1-C2-C3

There are no ring outliers.

4 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	502	GOL	6	0
4	B	502	PLP	3	0
3	A	504	GOL	2	0
3	A	501	GOL	1	0

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	176:ILE	C	177:GLU	N	1.18

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	464/474 (97%)	-0.99	2 (0%) 88 90	8, 17, 31, 61	7 (1%)
2	B	465/474 (98%)	-0.94	2 (0%) 88 90	8, 19, 38, 69	10 (2%)
All	All	929/948 (97%)	-0.96	4 (0%) 88 90	8, 18, 36, 69	17 (1%)

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	466	LEU	2.7
1	A	177	GLU	2.4
1	A	466	LEU	2.1
2	B	421[A]	ARG	2.1

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	LLP	A	290	24/25	0.94	0.11	13,21,48,68	0

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(Å ²)	Q<0.9
3	GOL	A	504	6/6	0.78	0.19	38,60,117,133	0
3	GOL	A	502	6/6	0.84	0.12	50,59,67,70	0
3	GOL	A	501	6/6	0.88	0.11	39,47,50,50	0
3	GOL	A	503	6/6	0.89	0.11	38,48,56,57	0
3	GOL	B	501[A]	6/6	0.89	0.15	37,50,65,78	14
3	GOL	B	501[B]	6/6	0.89	0.15	32,49,65,78	14
4	PLP	B	502	15/16	0.94	0.08	16,22,29,60	0

6.5 Other polymers [i](#)

There are no such residues in this entry.