



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 5, 2023 – 02:37 AM EST

PDB ID : 5UC7  
Title : Crystal structure of BioA / 7,8-diaminopelargonic acid aminotransferase / DAPA synthase from *Citrobacter rodentium*, PLP complex  
Authors : Stogios, P.J.; Evdokimova, E.; Di Leo, R.; Savchenko, A.; Anderson, W.F.; Center for Structural Genomics of Infectious Diseases (CSGID)  
Deposited on : 2016-12-21  
Resolution : 1.83 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](mailto: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 : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

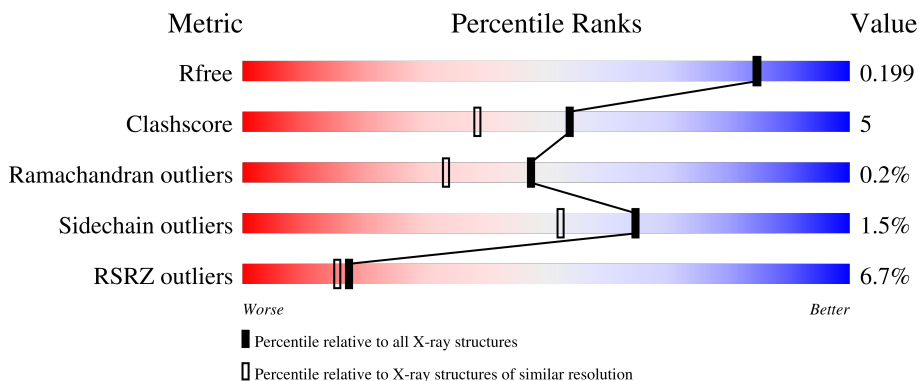
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.83 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))
$R_{free}$	130704	4003 (1.86-1.82)
Clashscore	141614	4233 (1.86-1.82)
Ramachandran outliers	138981	4185 (1.86-1.82)
Sidechain outliers	138945	4186 (1.86-1.82)
RSRZ outliers	127900	3957 (1.86-1.82)

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  $\geq 3$ , 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	431	

## 2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 3873 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.

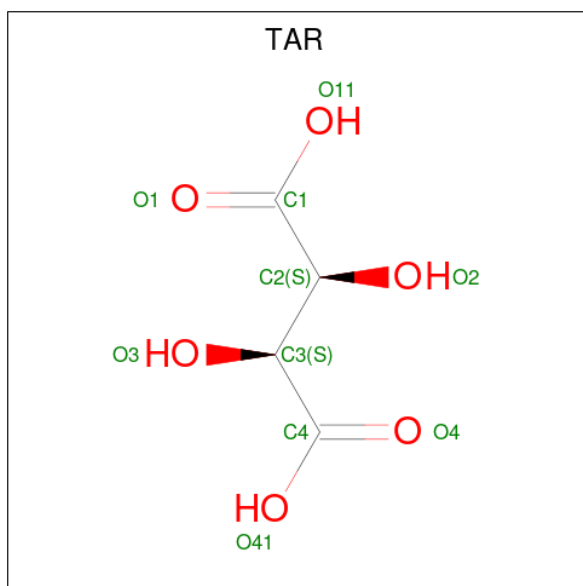
- Molecule 1 is a protein called Adenosylmethionine-8-amino-7-oxononanoate aminotransferase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	P	S			
1	A	431	3396	2149	607	609	1	30	0	9	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLN	-	expression tag	UNP D2TPI3
A	0	GLY	-	expression tag	UNP D2TPI3

- Molecule 2 is D(-)-TARTARIC ACID (three-letter code: TAR) (formula: C<sub>4</sub>H<sub>6</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
2	A	1	10	4	6	0	0
2	A	1	10	4	6	0	0

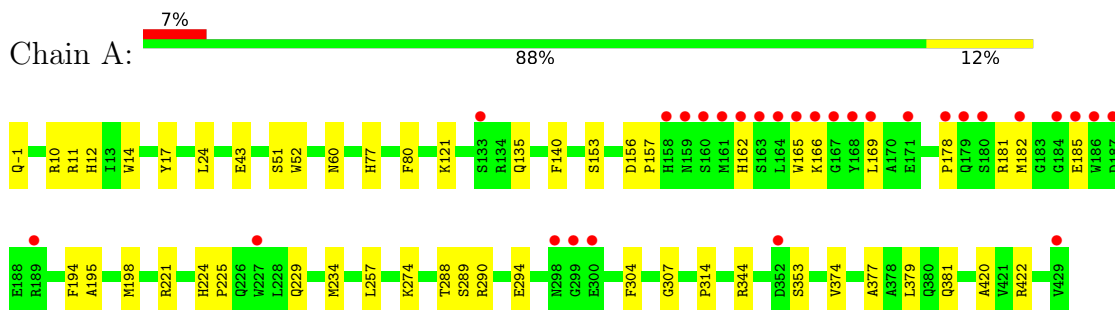
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	450	Total	O	0	7
			457	457		

### 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: Adenosylmethionine-8-amino-7-oxononanoate aminotransferase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	75.37Å 88.31Å 63.27Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.15 – 1.83 51.43 – 1.84	Depositor EDS
% Data completeness (in resolution range)	97.6 (44.15-1.83) 97.6 (51.43-1.84)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.99 (at 1.83Å)	Xtrriage
Refinement program	PHENIX dev_2481	Depositor
R, $R_{free}$	0.152 , 0.198 0.153 , 0.199	Depositor DCC
$R_{free}$ test set	2000 reflections (5.44%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	19.3	Xtrriage
Anisotropy	0.690	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 66.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	3873	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	34.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.95% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: TAR, LLP

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	0.82	0/3478	0.80	0/4716

There are no bond length outliers.

There are no bond angle outliers.

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	3396	0	3376	36	0
2	A	20	0	8	0	0
3	A	457	0	0	11	0
All	All	3873	0	3384	36	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 (36) 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:-1:GLN:HG2	3:A:940:HOH:O	2.02	0.59

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:12:HIS:HE1	3:A:933:HOH:O	1.90	0.54
1:A:43[A]:GLU:HG3	3:A:699:HOH:O	2.08	0.52
1:A:381:GLN:HG3	3:A:920:HOH:O	2.10	0.51
1:A:165:TRP:O	1:A:169:LEU:N	2.42	0.50
1:A:17:TYR:HD1	3:A:816:HOH:O	1.94	0.50
1:A:10[B]:ARG:NH2	3:A:618:HOH:O	2.46	0.49
1:A:162:HIS:O	1:A:166:LYS:N	2.46	0.47
1:A:24:LEU:HD22	1:A:377:ALA:HB2	1.97	0.47
1:A:224:HIS:ND1	1:A:225:PRO:HD2	2.29	0.47
1:A:121[A]:LYS:HG2	1:A:304:PHE:HE1	1.80	0.46
1:A:221:ARG:NH1	3:A:624:HOH:O	2.49	0.46
1:A:181:ARG:C	1:A:182:MET:HG2	2.36	0.46
1:A:11:ARG:O	1:A:11:ARG:HG3	2.15	0.45
1:A:374:VAL:HB	1:A:379:LEU:HD12	1.99	0.45
1:A:353:SER:HB2	1:A:422[B]:ARG:NH1	2.32	0.44
1:A:353:SER:HB2	1:A:422[B]:ARG:HH12	1.82	0.44
1:A:288[B]:THR:HG23	1:A:289:SER:O	2.17	0.43
1:A:344:ARG:NH1	3:A:632:HOH:O	2.52	0.43
1:A:10[A]:ARG:NH1	3:A:629:HOH:O	2.51	0.43
1:A:77:HIS:HA	1:A:314:PRO:HD2	2.01	0.43
1:A:198:MET:HE1	1:A:234:MET:C	2.38	0.43
1:A:153:SER:HB3	1:A:169:LEU:HG	2.00	0.43
1:A:185:GLU:HG2	1:A:224:HIS:HD2	1.83	0.43
1:A:51:SER:O	1:A:52:TRP:HB2	2.19	0.43
1:A:194:PHE:HD1	1:A:234:MET:HE1	1.84	0.43
1:A:156:ASP:HA	1:A:157:PRO:HD3	1.88	0.42
1:A:185:GLU:HG2	1:A:224:HIS:CD2	2.54	0.42
1:A:195:ALA:HA	1:A:234:MET:HE1	2.01	0.42
1:A:229:GLN:HB3	3:A:899:HOH:O	2.18	0.42
1:A:224:HIS:CE1	1:A:225:PRO:HD2	2.55	0.42
1:A:140:PHE:CD2	1:A:178:PRO:HG3	2.55	0.42
1:A:165:TRP:O	1:A:169:LEU:HB2	2.20	0.41
1:A:379:LEU:HD21	1:A:420:ALA:HB1	2.02	0.41
1:A:290:ARG:O	1:A:294:GLU:HG3	2.22	0.40
1:A:422[B]:ARG:NE	3:A:638:HOH:O	2.54	0.40

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	437/431 (101%)	420 (96%)	16 (4%)	1 (0%)	47 33

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	307	GLY

### 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	351/342 (103%)	346 (99%)	5 (1%)	67 55

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

Mol	Chain	Res	Type
1	A	14	TRP
1	A	60	ASN
1	A	80	PHE
1	A	135	GLN
1	A	257	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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	274	1	23,24,25	2.76	7 (30%)	25,32,34	1.70	5 (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	LLP	A	274	1	-	4/16/17/19	0/1/1/1

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	274	LLP	C4-C4'	8.65	1.63	1.46
1	A	274	LLP	C4'-NZ	5.48	1.45	1.27
1	A	274	LLP	C2'-C2	3.96	1.57	1.50
1	A	274	LLP	C6-N1	3.60	1.42	1.34
1	A	274	LLP	C4-C5	-2.88	1.38	1.42
1	A	274	LLP	CB-CA	2.82	1.57	1.53
1	A	274	LLP	C3-C2	2.74	1.43	1.40

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	274	LLP	C4-C4'-NZ	-4.63	103.05	124.31
1	A	274	LLP	C3-C4-C5	3.19	120.71	118.26

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	274	LLP	C5-C6-N1	-2.57	119.54	123.82
1	A	274	LLP	OP4-C5'-C5	2.48	114.08	109.35
1	A	274	LLP	C3-C2-N1	-2.10	118.05	120.77

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	274	LLP	CG-CD-CE-NZ
1	A	274	LLP	C3-C4-C4'-NZ
1	A	274	LLP	CD-CE-NZ-C4'
1	A	274	LLP	C4-C4'-NZ-CE

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no monosaccharides 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	TAR	A	502	-	9,9,9	1.14	0	12,12,12	1.16	1 (8%)
2	TAR	A	501	-	9,9,9	1.67	2 (22%)	12,12,12	1.29	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	TAR	A	502	-	-	6/12/12/12	-
2	TAR	A	501	-	-	6/12/12/12	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	501	TAR	C2-C1	2.89	1.56	1.52
2	A	501	TAR	C3-C4	2.10	1.55	1.52

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	502	TAR	O1-C1-C2	-2.04	116.26	121.63

There are no chirality outliers.

All (12) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	501	TAR	O1-C1-C2-O2
2	A	501	TAR	O11-C1-C2-O2
2	A	501	TAR	O3-C3-C4-O4
2	A	501	TAR	O3-C3-C4-O41
2	A	502	TAR	O1-C1-C2-O2
2	A	501	TAR	O1-C1-C2-C3
2	A	501	TAR	O11-C1-C2-C3
2	A	502	TAR	O1-C1-C2-C3
2	A	502	TAR	O11-C1-C2-C3
2	A	502	TAR	O11-C1-C2-O2
2	A	502	TAR	C2-C3-C4-O4
2	A	502	TAR	C2-C3-C4-O41

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers [i](#)

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 [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	430/431 (99%)	0.19	29 (6%) <b>17</b> <b>15</b>	11, 27, 73, 127	0

All (29) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	164	LEU	5.7
1	A	186	TRP	4.6
1	A	298	ASN	4.6
1	A	161	MET	4.3
1	A	162	HIS	3.8
1	A	168	TYR	3.7
1	A	299	GLY	3.7
1	A	166	LYS	3.6
1	A	187	ASP	3.6
1	A	182	MET	3.4
1	A	178	PRO	3.3
1	A	159	ASN	3.2
1	A	429	VAL	3.1
1	A	180	SER	3.0
1	A	185	GLU	2.9
1	A	165	TRP	2.9
1	A	169	LEU	2.9
1	A	300	GLU	2.8
1	A	167	GLY	2.8
1	A	352	ASP	2.8
1	A	189	ARG	2.7
1	A	179	GLN	2.7
1	A	163	SER	2.5
1	A	133[A]	SER	2.5
1	A	184	GLY	2.3
1	A	227	TRP	2.3
1	A	171	GLU	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	160	SER	2.1
1	A	158	HIS	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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	LLP	A	274	24/25	0.92	0.14	16,32,36,36	0

## 6.3 Carbohydrates [i](#)

There are no monosaccharides 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
2	TAR	A	501	10/10	0.78	0.18	45,57,66,68	0
2	TAR	A	502	10/10	0.81	0.21	90,92,93,93	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.