



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 3, 2017 – 07:51 AM EDT

PDB ID : 4XTC  
Title : Crystal structure of bacterial alginate ABC transporter in complex with alginate pentasaccharide-bound periplasmic protein  
Authors : Kaneko, A.; Maruyama, Y.; Mizuno, N.; Baba, S.; Kumasaka, T.; Mikami, B.; Murata, K.; Hashimoto, W.  
Deposited on : unknown  
Resolution : 3.60 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20029824  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20029824

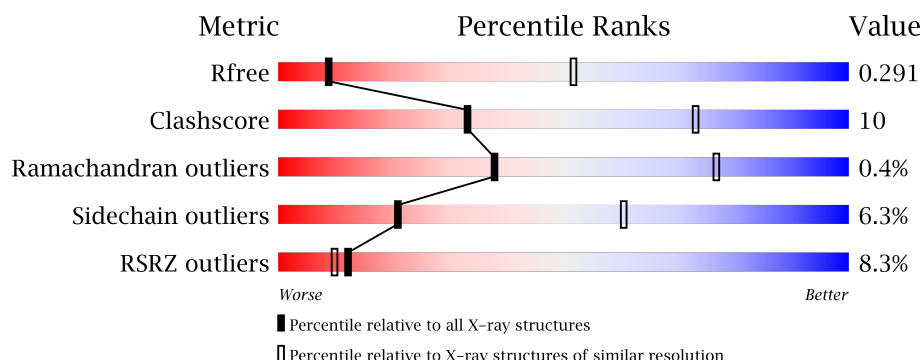
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.60 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))
$R_{free}$	100719	1026 (3.74-3.46)
Clashscore	112137	1036 (3.70-3.50)
Ramachandran outliers	110173	1030 (3.72-3.48)
Sidechain outliers	110143	1030 (3.72-3.48)
RSRZ outliers	101464	1051 (3.74-3.46)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	M	301	<div> <div>6%</div> <div>70%</div> <div>24%</div> <div>• •</div> </div>
2	N	305	<div> <div>6%</div> <div>72%</div> <div>19%</div> <div>• 6%</div> </div>
3	S	363	<div> <div>7%</div> <div>68%</div> <div>28%</div> <div>•</div> </div>
3	T	363	<div> <div>8%</div> <div>69%</div> <div>28%</div> <div>•</div> </div>
4	Q	516	<div> <div>11%</div> <div>71%</div> <div>23%</div> <div>• 5%</div> </div>

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
5	BEM	Q	503	-	-	-	X
5	BEM	Q	504	-	-	-	X
5	BEM	Q	505	-	-	-	X

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 14264 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 AlgM1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	M	288	Total	C	N	O	S	0	0	0
			2323	1554	369	390	10			

There are 23 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
M	?	-	THR	deletion	UNP Q9KWT8
M	?	-	SER	deletion	UNP Q9KWT8
M	?	-	ALA	deletion	UNP Q9KWT8
M	?	-	THR	deletion	UNP Q9KWT8
M	?	-	LYS	deletion	UNP Q9KWT8
M	?	-	ALA	deletion	UNP Q9KWT8
M	?	-	GLN	deletion	UNP Q9KWT8
M	?	-	SER	deletion	UNP Q9KWT8
M	?	-	ILE	deletion	UNP Q9KWT8
M	?	-	PRO	deletion	UNP Q9KWT8
M	?	-	LEU	deletion	UNP Q9KWT8
M	?	-	PRO	deletion	UNP Q9KWT8
M	?	-	ALA	deletion	UNP Q9KWT8
M	?	-	ALA	deletion	UNP Q9KWT8
M	?	-	THR	deletion	UNP Q9KWT8
M	?	-	LEU	deletion	UNP Q9KWT8
M	?	-	ASP	deletion	UNP Q9KWT8
M	?	-	VAL	deletion	UNP Q9KWT8
M	?	-	ARG	deletion	UNP Q9KWT8
M	?	-	SER	deletion	UNP Q9KWT8
M	?	-	LYS	deletion	UNP Q9KWT8
M	?	-	PRO	deletion	UNP Q9KWT8
M	?	-	LEU	deletion	UNP Q9KWT8

- Molecule 2 is a protein called AlgM2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	N	287	Total	C	N	O	S	0	0	0
			2278	1521	362	382	13			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
N	294	LEU	-	expression tag	UNP Q9KWT7
N	295	GLU	-	expression tag	UNP Q9KWT7
N	296	HIS	-	expression tag	UNP Q9KWT7
N	297	HIS	-	expression tag	UNP Q9KWT7
N	298	HIS	-	expression tag	UNP Q9KWT7
N	299	HIS	-	expression tag	UNP Q9KWT7
N	300	HIS	-	expression tag	UNP Q9KWT7
N	301	HIS	-	expression tag	UNP Q9KWT7
N	302	HIS	-	expression tag	UNP Q9KWT7
N	303	HIS	-	expression tag	UNP Q9KWT7
N	304	HIS	-	expression tag	UNP Q9KWT7
N	305	HIS	-	expression tag	UNP Q9KWT7

- Molecule 3 is a protein called AlgS.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	S	363	Total	C	N	O	S	0	0	0
			2777	1745	503	518	11			
3	T	363	Total	C	N	O	S	0	0	0
			2777	1745	503	518	11			

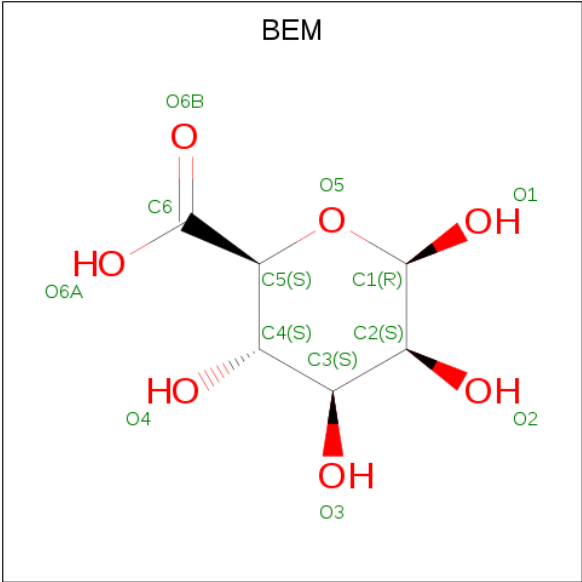
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
S	160	GLN	GLU	engineered mutation	UNP Q9KWT9
T	160	GLN	GLU	engineered mutation	UNP Q9KWT9

- Molecule 4 is a protein called AlgQ2.

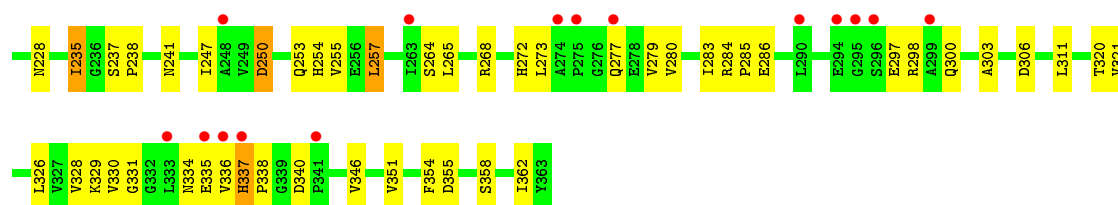
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	Q	492	Total	C	N	O	S	0	0	0
			4048	2604	693	735	16			

- Molecule 5 is beta-D-mannuronic acid (three-letter code: BEM) (formula: C<sub>6</sub>H<sub>10</sub>O<sub>7</sub>).

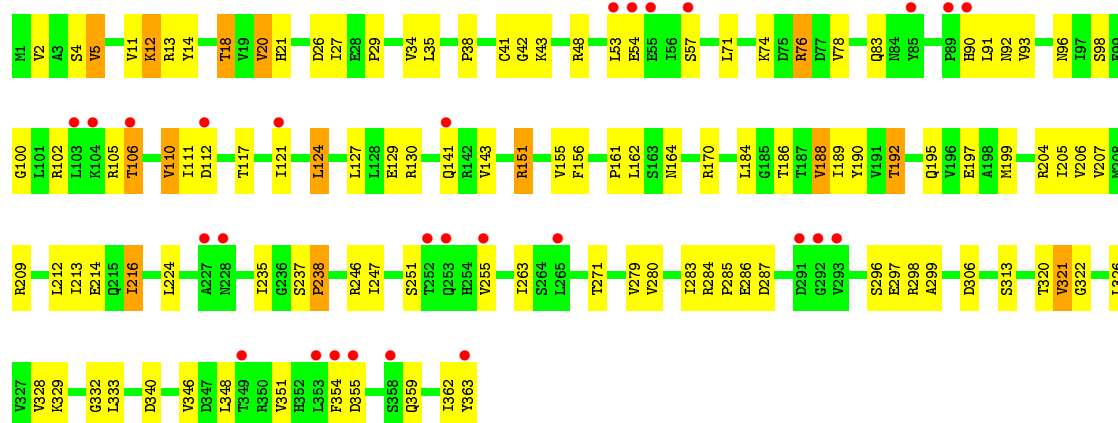


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	Q	1	Total	C	O	0	0
			12	6	6		
5	Q	1	Total	C	O	0	0
			12	6	6		
5	Q	1	Total	C	O	0	0
			12	6	6		
5	Q	1	Total	C	O	0	0
			12	6	6		
5	Q	1	Total	C	O	0	0
			13	6	7		

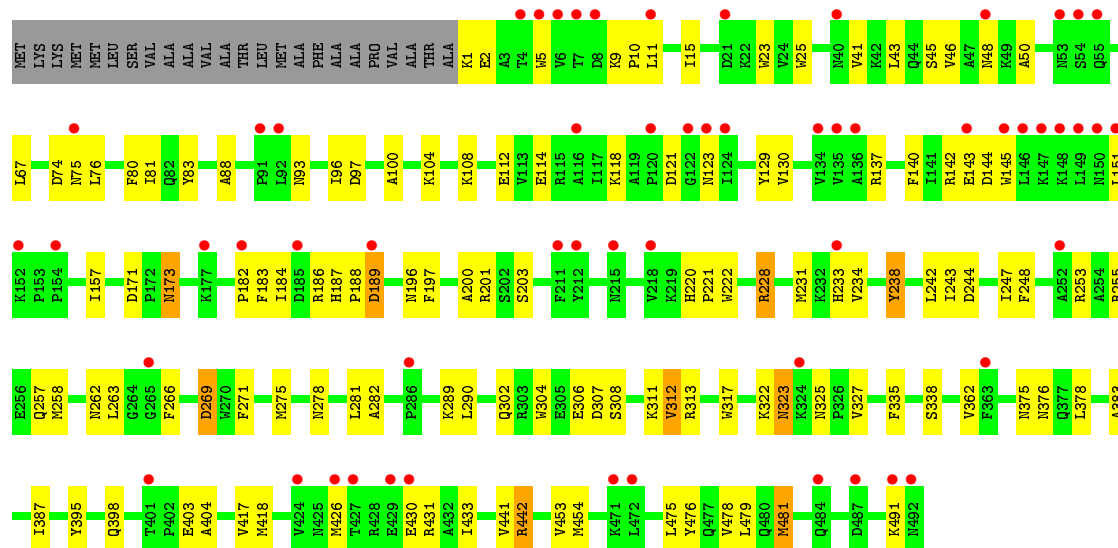




### • Molecule 3: AlgS



### • Molecule 4: AlgQ2



GLOBAL-STATISTICS INFOmissingINFO



## 4 Model quality [i](#)

### 4.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: BEM

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	M	0.23	0/2384	0.42	0/3250
2	N	0.24	0/2337	0.42	0/3178
3	S	0.22	0/2822	0.48	1/3826 (0.0%)
3	T	0.21	0/2822	0.46	0/3826
4	Q	0.21	0/4168	0.37	0/5640
All	All	0.22	0/14533	0.43	1/19720 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	S	257	LEU	CA-CB-CG	5.34	127.58	115.30

There are no chirality outliers.

There are no planarity outliers.

### 4.2 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	M	2323	0	2412	46	0
2	N	2278	0	2347	51	0
3	S	2777	0	2854	60	0
3	T	2777	0	2854	67	0
4	Q	4048	0	3920	78	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	Q	61	0	33	2	0
All	All	14264	0	14420	280	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 10.

The worst 5 of 280 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)
3:S:34:VAL:HG22	3:S:190:TYR:HB3	1.63	0.79
3:T:34:VAL:HG22	3:T:190:TYR:HB3	1.63	0.79
3:S:235:ILE:HD11	3:S:284:ARG:HH12	1.50	0.76
3:S:192:THR:HG22	3:S:194:ASP:H	1.55	0.72
4:Q:201:ARG:NH1	4:Q:418:MET:SD	2.64	0.71

There are no symmetry-related clashes.

## 4.3 Torsion angles [i](#)

### 4.3.1 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	M	284/301 (94%)	264 (93%)	19 (7%)	1 (0%)	38	77
2	N	285/305 (93%)	263 (92%)	20 (7%)	2 (1%)	25	68
3	S	361/363 (99%)	336 (93%)	23 (6%)	2 (1%)	28	70
3	T	361/363 (99%)	335 (93%)	24 (7%)	2 (1%)	28	70
4	Q	490/516 (95%)	464 (95%)	25 (5%)	1 (0%)	51	85
All	All	1781/1848 (96%)	1662 (93%)	111 (6%)	8 (0%)	38	77

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	M	253	ILE
3	S	362	ILE
3	T	111	ILE
4	Q	312	VAL
3	T	110	VAL

#### 4.3.2 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	M	251/260 (96%)	238 (95%)	13 (5%)	27	65
2	N	242/258 (94%)	227 (94%)	15 (6%)	21	61
3	S	307/307 (100%)	282 (92%)	25 (8%)	14	50
3	T	307/307 (100%)	284 (92%)	23 (8%)	16	53
4	Q	424/440 (96%)	403 (95%)	21 (5%)	28	66
All	All	1531/1572 (97%)	1434 (94%)	97 (6%)	21	60

5 of 97 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	S	257	LEU
4	Q	46	VAL
3	T	279	VAL
3	S	279	VAL
3	S	328	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
4	Q	55	GLN
4	Q	323	ASN
4	Q	375	ASN
3	S	193	HIS
4	Q	372	GLN

### 4.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 4.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 4.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 4.6 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$
5	BEM	Q	501	5	9,12,13	0.25	0	13,17,19	0.98	1 (7%)
5	BEM	Q	502	5	9,12,13	0.21	0	13,17,19	0.76	0
5	BEM	Q	503	5	9,12,13	0.21	0	13,17,19	0.83	0
5	BEM	Q	504	5	9,12,13	0.23	0	13,17,19	0.75	0
5	BEM	Q	505	5	10,13,13	0.44	0	15,19,19	0.73	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
5	BEM	Q	501	5	-	0/0/21/24	0/1/1/1
5	BEM	Q	502	5	-	0/0/21/24	0/1/1/1
5	BEM	Q	503	5	-	0/0/21/24	0/1/1/1
5	BEM	Q	504	5	-	0/0/21/24	0/1/1/1
5	BEM	Q	505	5	-	0/0/24/24	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
5	Q	501	BEM	C1-O5-C5	2.06	115.88	112.17

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	Q	503	BEM	1	0
5	Q	505	BEM	1	0

## 4.7 Other polymers [i](#)

There are no such residues in this entry.

## 4.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 5 Fit of model and data [i](#)

### 5.1 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	M	288/301 (95%)	0.29	19 (6%)	19 13	58, 103, 179, 224	0
2	N	287/305 (94%)	0.17	17 (5%)	23 16	60, 99, 148, 213	0
3	S	363/363 (100%)	0.28	25 (6%)	18 13	59, 103, 178, 259	0
3	T	363/363 (100%)	0.44	28 (7%)	14 11	75, 126, 201, 281	0
4	Q	492/516 (95%)	0.64	59 (11%)	5 5	89, 137, 196, 252	0
All	All	1793/1848 (97%)	0.40	148 (8%)	12 10	58, 118, 186, 281	0

The worst 5 of 148 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	M	317	ARG	6.9
4	Q	150	ASN	6.7
3	S	1	MET	6.4
3	S	295	GLY	6.0
1	M	318	ILE	5.4

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

There are no non-standard protein/DNA/RNA residues in this entry.

### 5.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 5.4 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. LLDF column lists the quality of electron

density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
5	BEM	Q	504	12/13	0.84	0.80	3.20	87,122,132,136	0
5	BEM	Q	505	13/13	0.89	0.83	2.85	119,137,147,156	0
5	BEM	Q	503	12/13	0.86	0.46	0.52	93,114,122,125	0
5	BEM	Q	501	12/13	0.89	0.29	-0.90	95,121,135,144	0
5	BEM	Q	502	12/13	0.82	0.34	-1.11	100,116,124,125	0

## 5.5 Other polymers [i](#)

There are no such residues in this entry.