



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 08:03 PM GMT

PDB ID : 4QSK  
Title : Crystal Structure of *L. monocytogenes* Pyruvate Carboxylase in complex with Cyclic-di-AMP  
Authors : Choi, P.H.; Tong, L.  
Deposited on : 2014-07-04  
Resolution : 2.70 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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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) : trunk26865

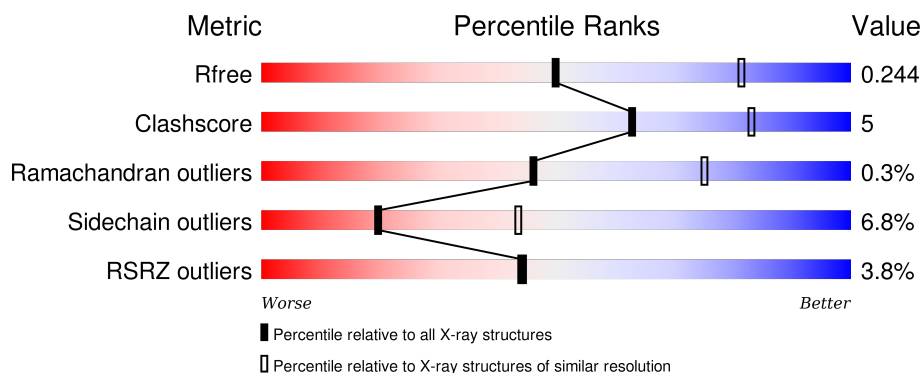
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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))
$R_{free}$	91344	2103 (2.70-2.70)
Clashscore	102246	2422 (2.70-2.70)
Ramachandran outliers	100387	2382 (2.70-2.70)
Sidechain outliers	100360	2382 (2.70-2.70)
RSRZ outliers	91569	2107 (2.70-2.70)

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	A	1147	<div> <div>4%</div> <div>79%</div> <div>13%</div> <div>6%</div> </div>
1	B	1147	<div> <div>3%</div> <div>78%</div> <div>14%</div> <div>6%</div> </div>

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 16959 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.

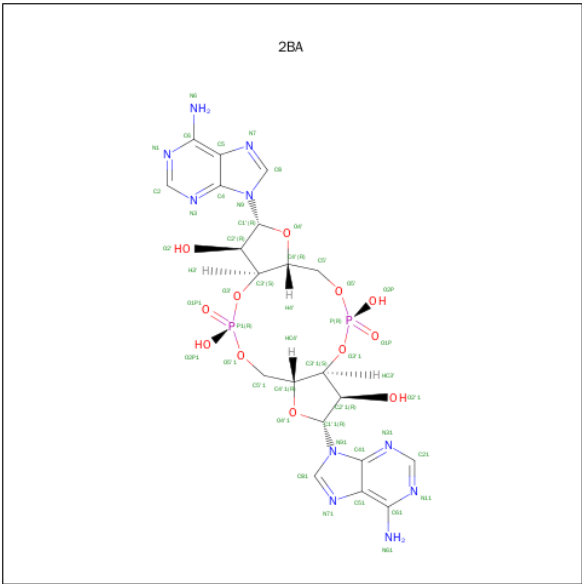
- Molecule 1 is a protein called Pyruvate carboxylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1073	Total	C	N	O	S	0	0	0
			8438	5355	1433	1609	41			
1	B	1073	Total	C	N	O	S	0	0	0
			8449	5362	1433	1613	41			

There are 2 discrepancies between the modelled and reference sequences:

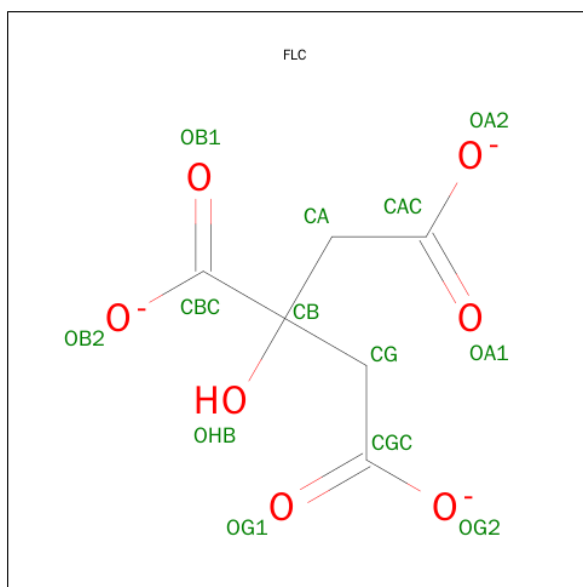
Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	EXPRESSION TAG	UNP W6G6F5
B	0	MET	-	EXPRESSION TAG	UNP W6G6F5

- Molecule 2 is (2R,3R,3AS,5R,7AR,9R,10R,10AS,12R,14AR)-2,9-BIS(6-AMINO-9H-PURIN-9-YL)OCTAHYDRO-2H,7H-DIFURO[3,2-D:3',2'-J][1,3,7,9,2,8]TETRAOXADIPHOSPHACYCLODODECINE-3,5,10,12-TETROL 5,12-DIOXIDE (three-letter code: 2BA) (formula: C<sub>20</sub>H<sub>24</sub>N<sub>10</sub>O<sub>12</sub>P<sub>2</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			22	10	5	6	1		
2	B	1	Total	C	N	O	P	0	0
			22	10	5	6	1		

- Molecule 3 is CITRATE ANION (three-letter code: FLC) (formula:  $C_6H_5O_7$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			13	6	7		
3	B	1	Total	C	O	0	0
			13	6	7		

- Molecule 4 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Mn	0	0
			1	1		
4	A	1	Total	Mn	0	0
			1	1		

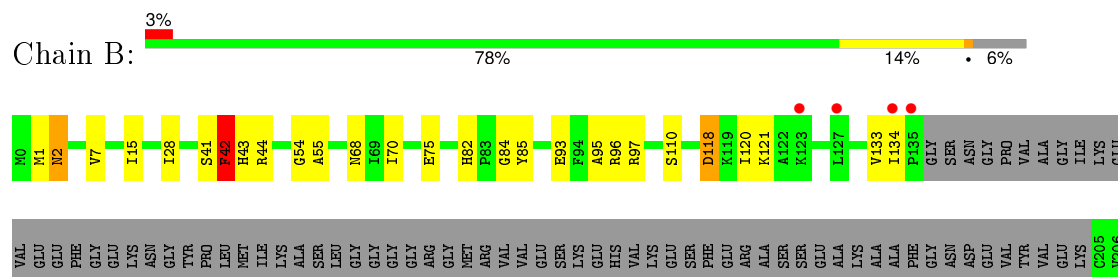
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: Pyruvate carboxylase



#### • Molecule 1: Pyruvate carboxylase



ILE	L1025	Q791	R586	Q333	N207
	T1039	E807	F587	Q344	N208
	Q1066	A819	D591	S345	P209
	M1058	I820	P592	R346	K210
	M1059	N821	R600	M357	D219
	V1060	V827	M607	Y366	I224
	Q1061	H830	R613	R367	V225
	SER	E831	M616	S368	H226
	THR	M832	Y620	T369	K240
	VAL	G835	I376	R375	V241
	ILE	Q836	M622	D377	V242
	A1066	N839	G639	F382	E243
	R1067	Q842	V640	Q383	N248
	E1076	V847	D641	V387	S252
	Q1088	D851	R644	S393	E253
	V1089	M866	R663	C399	L254
	V1090	D869	I668	R414	N255
	V1091	T874	E670	N415	R256
	S1096	P875	C674	E418	R257
	V1097	K878	D682	F419	I258
	G1100	D882	T688	R420	C259
	D1101	E894	Y691	D451	K267
	P1102	E917	G703	I462	N273
	I1103	E927	I709	R483	L280
	I1104	R837	M712	K468	V281
	I1105	L940	L728	Y502	E282
	T1106	T941	K729	T510	G283
	E1107	D942	I736	Q544	Y287
	A1108	E967	T740	T549	F288
	M1109	K981	Y751	R552	I289
	K1110	V982	D760	F571	E290
	M1111	K992	I761	M576	R294
	E1112	T997	S766	D582	V295
	T1113	E1015	V786		E298
	I1115				H299
	I1116				T300
	Q1117				I301
	A1117				K304
	P1118				Q314
	F1119				D319
	D1120				A322
	V1123				L323
	S1124				H324
	S1129				D325
	D1130				Q326
	I1141				
	E1142				
	V1143				
	M1144				
	R1145				

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	227.00 Å   146.86 Å   96.06 Å 90.00°   103.54°   90.00°	Depositor
Resolution (Å)	45.68 – 2.70 45.64 – 2.70	Depositor EDS
% Data completeness (in resolution range)	96.9 (45.68-2.70) 96.9 (45.64-2.70)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.83 (at 2.69 Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, $R_{free}$	0.198   ,   0.242 0.198   ,   0.244	Depositor DCC
$R_{free}$ test set	4071 reflections (5.26%)	DCC
Wilson B-factor (Å <sup>2</sup> )	65.2	Xtriage
Anisotropy	0.113	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 40.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtriage
Outliers	0 of 81819 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	16959	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	72.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.00% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: FLC, MN, 2BA

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	0.61	0/8605	0.80	4/11659 (0.0%)
1	B	0.68	0/8616	0.82	3/11673 (0.0%)
All	All	0.64	0/17221	0.81	7/23332 (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	A	0	1

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	B	42	PHE	CB-CA-C	6.27	122.93	110.40
1	A	420	ARG	NE-CZ-NH2	-5.54	117.53	120.30
1	A	42	PHE	CB-CA-C	5.36	121.11	110.40
1	A	420	ARG	NE-CZ-NH1	5.12	122.86	120.30
1	B	304	MET	CG-SD-CE	-5.06	92.10	100.20

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	621	LYS	Peptide



## 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	8438	0	8361	92	0
1	B	8449	0	8378	88	0
2	A	22	0	11	0	1
2	B	22	0	11	0	1
3	A	13	0	5	0	0
3	B	13	0	5	1	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
All	All	16959	0	16771	177	2

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.

The worst 5 of 177 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:B:582:ASP:OD2	1:B:586:ARG:NH1	2.01	0.92
1:A:582:ASP:OD2	1:A:586:ARG:NH1	2.03	0.91
1:A:663:ARG:NH1	1:A:703:GLY:O	2.13	0.81
1:B:663:ARG:NH1	1:B:703:GLY:O	2.14	0.80
1:B:510:THR:HG21	1:B:571:PHE:O	1.82	0.80

All (2) 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)
2:A:2000:2BA:P	2:A:2000:2BA:O3'[2_757]	1.79	0.41
2:B:1202:2BA:P	2:B:1202:2BA:O3'[2_757]	1.82	0.38

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	1067/1147 (93%)	1004 (94%)	60 (6%)	3 (0%)	46	75
1	B	1067/1147 (93%)	1007 (94%)	57 (5%)	3 (0%)	46	75
All	All	2134/2294 (93%)	2011 (94%)	117 (6%)	6 (0%)	46	75

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	283	GLY
1	B	283	GLY
1	A	273	ASN
1	B	273	ASN
1	A	869	ASP

### 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	917/982 (93%)	855 (93%)	62 (7%)	20	43
1	B	920/982 (94%)	857 (93%)	63 (7%)	20	43
All	All	1837/1964 (94%)	1712 (93%)	125 (7%)	20	43

5 of 125 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1111	MET

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	120	ILE
1	B	1067	ARG
1	A	1119	PHE
1	B	2	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 31 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	830	HIS
1	B	82	HIS
1	B	797	GLN
1	B	43	HIS
1	B	112	HIS

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 6 ligands modelled in this entry, 2 are monoatomic - leaving 4 for Mogul analysis.

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	2BA	A	2000	-	15,24,50	1.25	1 (6%)	16,35,78	2.21	3 (18%)
3	FLC	A	2001	-	3,12,12	1.51	1 (33%)	3,17,17	2.73	2 (66%)
2	2BA	B	1202	-	15,24,50	1.28	1 (6%)	16,35,78	2.37	3 (18%)
3	FLC	B	1203	-	3,12,12	1.25	0	3,17,17	1.22	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	2BA	A	2000	-	-	0/3/25/62	0/3/3/7
3	FLC	A	2001	-	-	0/6/16/16	0/0/0/0
2	2BA	B	1202	-	-	0/3/25/62	0/3/3/7
3	FLC	B	1203	-	-	0/6/16/16	0/0/0/0

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	2001	FLC	CG-CB	2.31	1.58	1.54
2	B	1202	2BA	C5-C4	2.62	1.46	1.40
2	A	2000	2BA	C5-C4	2.76	1.46	1.40

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	1202	2BA	N3-C2-N1	-7.40	123.23	128.89
2	A	2000	2BA	N3-C2-N1	-6.85	123.65	128.89
2	B	1202	2BA	C1'-N9-C4	-4.57	120.05	126.94
3	A	2001	FLC	CB-CA-CAC	-3.37	109.56	114.96
2	A	2000	2BA	C4-C5-N7	-2.89	106.82	109.48

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	2000	2BA	0	1
2	B	1202	2BA	0	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	1203	FLC	1	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, 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	1073/1147 (93%)	0.04	42 (3%) 43 43	47, 73, 115, 152	0
1	B	1073/1147 (93%)	-0.10	39 (3%) 46 46	34, 64, 104, 140	0
All	All	2146/2294 (93%)	-0.03	81 (3%) 44 44	34, 68, 109, 152	0

The worst 5 of 81 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	906	PHE	5.6
1	B	1120	ASP	4.9
1	A	904	ILE	4.6
1	B	135	PRO	4.5
1	A	903	THR	4.4

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

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

### 6.3 Carbohydrates [i](#)

There are no carbohydrates 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. 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( $\text{\AA}^2$ )	Q<0.9
4	MN	A	2002	1/1	0.99	0.28	1.03	91,91,91,91	0
4	MN	B	1201	1/1	0.97	0.26	0.95	65,65,65,65	0
2	2BA	A	2000	22/44	0.97	0.12	-0.64	52,62,76,79	0
3	FLC	A	2001	13/13	0.92	0.13	-1.19	73,82,88,91	0
3	FLC	B	1203	13/13	0.93	0.13	-1.47	64,78,86,87	0
2	2BA	B	1202	22/44	0.98	0.10	-1.78	41,49,53,54	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.