



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 12, 2017 – 07:46 pm GMT

PDB ID : 2YH5  
Title : Structure of the C-terminal domain of BamC  
Authors : Zeth, K.; Albrecht, R.  
Deposited on : 2011-04-27  
Resolution : 1.25 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

<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	:	trunk28620
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)	:	recalc28949

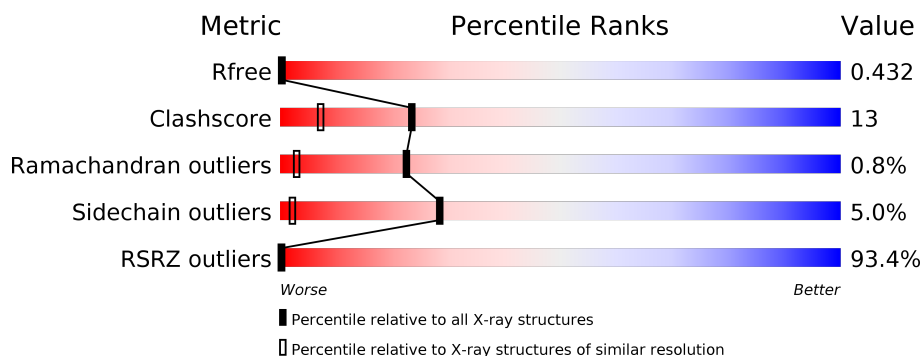
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.25 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	1690 (1.30-1.22)
Clashscore	112137	1781 (1.30-1.22)
Ramachandran outliers	110173	1716 (1.30-1.22)
Sidechain outliers	110143	1714 (1.30-1.22)
RSRZ outliers	101464	1693 (1.30-1.22)

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	127	<div> <div>89%</div> <div>78%14%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
2	PO4	A	1323	-	X	-	-
2	PO4	A	1324	-	-	-	X

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 1018 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.

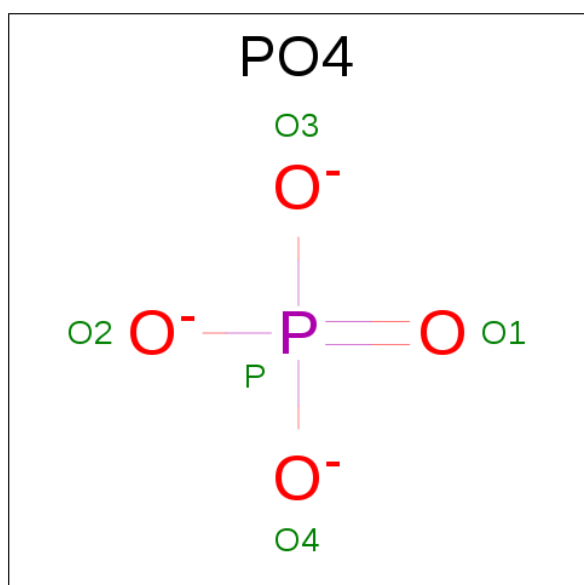
- Molecule 1 is a protein called DAPX PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	121	948	596	157	190	5	5	7	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	321	LEU	-	EXPRESSION TAG	UNP Q47548
A	322	GLU	-	EXPRESSION TAG	UNP Q47548
A	323	HIS	-	EXPRESSION TAG	UNP Q47548
A	324	HIS	-	EXPRESSION TAG	UNP Q47548
A	325	HIS	-	EXPRESSION TAG	UNP Q47548
A	326	HIS	-	EXPRESSION TAG	UNP Q47548
A	327	HIS	-	EXPRESSION TAG	UNP Q47548
A	328	HIS	-	EXPRESSION TAG	UNP Q47548

- Molecule 2 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



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

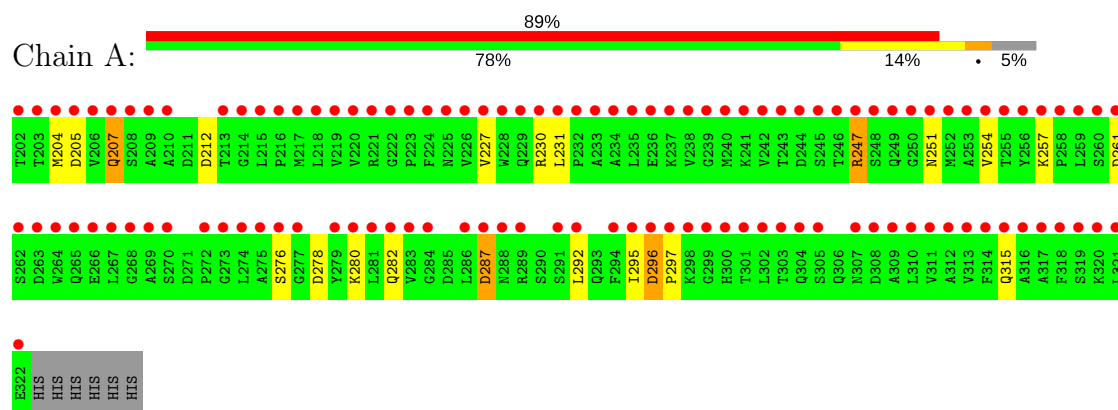
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	60	Total	O	0	0
			60	60		

### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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 ( $\text{RSRZ} > 2$ ). 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: DAPX PROTEIN



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	29.74Å 59.12Å 31.06Å 90.00° 116.37° 90.00°	Depositor
Resolution (Å)	29.56 – 1.25 26.65 – 1.24	Depositor EDS
% Data completeness (in resolution range)	100.0 (29.56-1.25) 86.8 (26.65-1.24)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.29 (at 1.24Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, $R_{free}$	0.146 , 0.184 0.439 , 0.432	Depositor DCC
$R_{free}$ test set	1169 reflections (5.00%)	DCC
Wilson B-factor (Å <sup>2</sup> )	13.3	Xtriage
Anisotropy	0.113	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 55.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.046 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.75	EDS
Total number of atoms	1018	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.95% 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 ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: PO4

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	1.31	3/985 (0.3%)	1.23	5/1338 (0.4%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	276	SER	CB-OG	-10.51	1.28	1.42
1	A	254	VAL	CB-CG2	-7.95	1.36	1.52
1	A	207	GLN	CB-CG	-5.84	1.36	1.52

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	247	ARG	NE-CZ-NH1	16.96	128.78	120.30
1	A	247	ARG	NE-CZ-NH2	-12.85	113.87	120.30
1	A	212	ASP	CB-CG-OD2	-6.85	112.14	118.30
1	A	287	ASP	CB-CG-OD2	-6.58	112.38	118.30
1	A	278	ASP	CB-CG-OD1	5.27	123.04	118.30

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	948	0	949	24	0
2	A	10	0	0	0	0
3	A	60	0	0	8	0
All	All	1018	0	949	24	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 13.

All (24) 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:207:GLN:HB2	3:A:2005:HOH:O	1.42	1.17
1:A:292:LEU:HG	3:A:2049:HOH:O	1.60	0.99
1:A:227:VAL:O	1:A:231[A]:LEU:HD23	1.68	0.94
1:A:230:ARG:NH2	3:A:2019:HOH:O	2.05	0.89
1:A:207:GLN:CG	3:A:2005:HOH:O	2.30	0.78
1:A:205:ASP:OD2	3:A:2004:HOH:O	2.09	0.70
1:A:251:ASN:OD1	1:A:280:LYS:HG3	1.92	0.69
1:A:251:ASN:HB2	1:A:282:GLN:HE22	1.57	0.69
1:A:280:LYS:HE3	1:A:282:GLN:HG2	1.76	0.68
1:A:280:LYS:CE	1:A:295:ILE:HD12	2.27	0.64
1:A:292:LEU:CD2	3:A:2049:HOH:O	2.46	0.62
1:A:280:LYS:HE2	1:A:295:ILE:HD12	1.84	0.59
1:A:280:LYS:HE3	1:A:282:GLN:CG	2.32	0.59
1:A:205:ASP:HA	1:A:315[A]:GLN:HE22	1.68	0.58
1:A:227:VAL:O	1:A:231[A]:LEU:CD2	2.49	0.58
1:A:227:VAL:CG1	1:A:231[A]:LEU:HD21	2.37	0.55
1:A:207:GLN:HG3	3:A:2005:HOH:O	2.01	0.53
1:A:296:ASP:CB	1:A:297:PRO:CD	2.87	0.52
1:A:207:GLN:CB	3:A:2005:HOH:O	2.15	0.52
1:A:296:ASP:HB2	1:A:297:PRO:CD	2.43	0.48
1:A:227:VAL:HG12	1:A:231[A]:LEU:CD2	2.44	0.47
1:A:227:VAL:CG1	1:A:231[A]:LEU:CD2	2.93	0.46
1:A:205:ASP:HA	1:A:315[A]:GLN:NE2	2.31	0.43
1:A:227:VAL:HG13	1:A:231[A]:LEU:HD21	2.00	0.42

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	126/127 (99%)	125 (99%)	0	1 (1%)	22	3

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	287	ASP

### 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	108/107 (101%)	102 (94%)	6 (6%)	25	2

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	204[A]	MET
1	A	204[B]	MET
1	A	247	ARG
1	A	257	LYS
1	A	261	ASP
1	A	296	ASP

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

Mol	Chain	Res	Type
1	A	229	GLN
1	A	282	GLN
1	A	293	GLN
1	A	307	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

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	PO4	A	1323	-	4,4,4	2.40	2 (50%)	6,6,6	1.66	2 (33%)
2	PO4	A	1324	-	4,4,4	0.86	0	6,6,6	1.50	2 (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
2	PO4	A	1323	-	-	0/0/0/0	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	PO4	A	1324	-	-	0/0/0/0	0/0/0/0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1323	PO4	P-O2	-4.02	1.40	1.54
2	A	1323	PO4	P-O1	2.63	1.56	1.50

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1323	PO4	O3-P-O1	-2.49	100.37	110.97
2	A	1324	PO4	O4-P-O1	-2.05	102.23	110.97
2	A	1323	PO4	O3-P-O2	2.36	116.59	107.90
2	A	1324	PO4	O3-P-O2	2.85	118.37	107.90

There are no chirality outliers.

There are no torsion outliers.

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 [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

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	121/127 (95%)	4.25	113 (93%) 0 0	10, 15, 29, 51	0

All (113) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	202	THR	10.8
1	A	321	LEU	10.6
1	A	267	LEU	10.1
1	A	286	LEU	8.8
1	A	287	ASP	7.5
1	A	242	VAL	7.5
1	A	250	GLY	7.4
1	A	269	ALA	7.0
1	A	220[A]	VAL	6.8
1	A	249	GLN	6.7
1	A	281	LEU	6.6
1	A	283	VAL	6.6
1	A	297	PRO	6.5
1	A	244	ASP	6.4
1	A	243	THR	6.3
1	A	221	ARG	6.3
1	A	224	PHE	6.3
1	A	317	ALA	6.1
1	A	226	VAL	6.0
1	A	247	ARG	6.0
1	A	310	LEU	6.0
1	A	318	PHE	5.9
1	A	228	TRP	5.9
1	A	279	TYR	5.7
1	A	259	LEU	5.5
1	A	215	LEU	5.4
1	A	203	THR	5.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	311	VAL	5.2
1	A	315[A]	GLN	5.2
1	A	261	ASP	5.0
1	A	295	ILE	5.0
1	A	205	ASP	5.0
1	A	299	GLY	5.0
1	A	231[A]	LEU	5.0
1	A	255	THR	4.9
1	A	302	LEU	4.9
1	A	314	PHE	4.9
1	A	292	LEU	4.9
1	A	264	TRP	4.8
1	A	301	THR	4.8
1	A	316	ALA	4.8
1	A	320	LYS	4.7
1	A	246	THR	4.7
1	A	204[A]	MET	4.6
1	A	277	GLY	4.6
1	A	256	TYR	4.5
1	A	305[A]	SER	4.5
1	A	233	ALA	4.5
1	A	206	VAL	4.4
1	A	238	VAL	4.3
1	A	245	SER	4.3
1	A	210	ALA	4.3
1	A	230	ARG	4.3
1	A	248	SER	4.3
1	A	251	ASN	4.3
1	A	319	SER	4.3
1	A	235	LEU	4.2
1	A	227	VAL	4.2
1	A	254	VAL	4.1
1	A	236[A]	GLU	4.1
1	A	217	MET	4.1
1	A	209	ALA	4.0
1	A	312	ALA	3.9
1	A	294	PHE	3.8
1	A	252	MET	3.8
1	A	270	SER	3.8
1	A	273	GLY	3.8
1	A	241	LYS	3.7
1	A	253	ALA	3.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	284	GLY	3.7
1	A	282	GLN	3.7
1	A	322	GLU	3.6
1	A	234	ALA	3.6
1	A	308	ASP	3.6
1	A	207	GLN	3.6
1	A	263	ASP	3.6
1	A	219	VAL	3.5
1	A	288	ASN	3.5
1	A	222	GLY	3.5
1	A	237	LYS	3.4
1	A	229	GLN	3.4
1	A	274	LEU	3.4
1	A	223	PRO	3.3
1	A	313	VAL	3.3
1	A	266	GLU	3.1
1	A	258	PRO	3.1
1	A	232	PRO	3.1
1	A	272	PRO	3.1
1	A	268	GLY	3.0
1	A	225	ASN	3.0
1	A	218	LEU	3.0
1	A	275	ALA	2.9
1	A	216	PRO	2.9
1	A	276	SER	2.9
1	A	240	MET	2.8
1	A	304	GLN	2.8
1	A	260	SER	2.7
1	A	289	ARG	2.7
1	A	296	ASP	2.6
1	A	298	LYS	2.5
1	A	239	GLY	2.5
1	A	309	ALA	2.5
1	A	291	SER	2.4
1	A	300	HIS	2.3
1	A	213	THR	2.3
1	A	214	GLY	2.3
1	A	303	THR	2.3
1	A	208[A]	SER	2.2
1	A	265	GLN	2.2
1	A	262	SER	2.2
1	A	257	LYS	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	307	ASN	2.1
1	A	280	LYS	2.1

## 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(Å <sup>2</sup> )	Q<0.9
2	PO4	A	1324	5/5	0.67	0.46	9.27	9,10,12,20	0
2	PO4	A	1323	5/5	0.87	0.38	1.87	5,9,10,13	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.