



# Full wwPDB X-ray Structure Validation Report ⓘ

May 24, 2020 – 04:32 am BST

PDB ID : 6QB7  
Title : Structure of the H1 domain of human KCTD16  
Authors : Pinkas, D.M.; Bufton, J.C.; Williams, E.P.; Strain-Damerell, C.; Kupinska, K.; Burgess-Brown, N.A.; von Delft, F.; Arrowsmith, C.H.; Edwards, A.M.; Bountra, C.; Bullock, A.N.; Structural Genomics Consortium (SGC)  
Deposited on : 2018-12-20  
Resolution : 2.23 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 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 : 2.11  
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) : 2.11

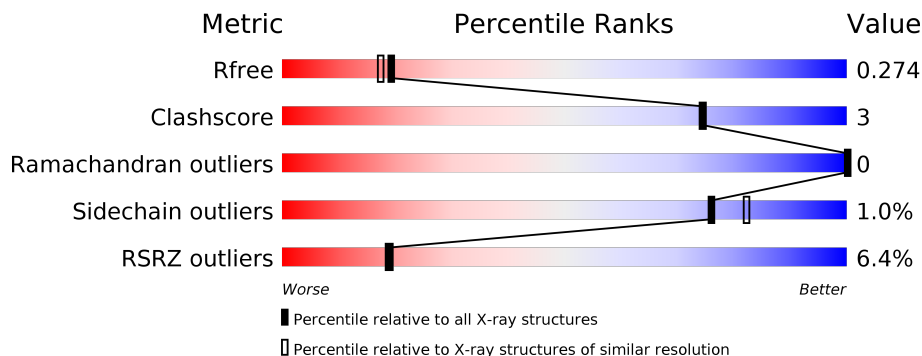
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.23 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	2391 (2.26-2.22)
Clashscore	141614	2539 (2.26-2.22)
Ramachandran outliers	138981	2489 (2.26-2.22)
Sidechain outliers	138945	2490 (2.26-2.22)
RSRZ outliers	127900	2353 (2.26-2.22)

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 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	163	<p>4% 64% 6% 30%</p>
1	B	163	<p>5% 64% 6% 30%</p>
1	C	163	<p>6% 61% 8% 31%</p>
1	D	163	<p>4% 62% 7% 31%</p>
1	E	163	<p>4% 65% 7% 31%</p>

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:

<b>Mol</b>	<b>Type</b>	<b>Chain</b>	<b>Res</b>	<b>Chirality</b>	<b>Geometry</b>	<b>Clashes</b>	<b>Electron density</b>
2	PO4	E	305	-	-	-	X

## 2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 4838 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.

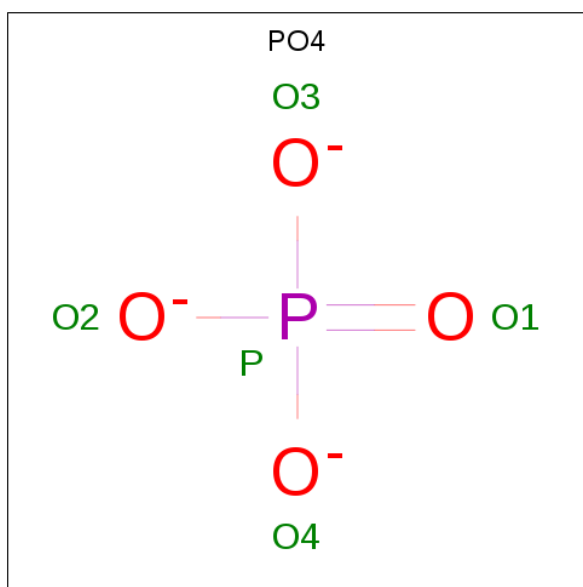
- Molecule 1 is a protein called BTB/POZ domain-containing protein KCTD16.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	114	926	593	160	168	5	0	0	0
1	B	114	931	595	163	168	5	0	0	0
1	C	112	929	594	163	167	5	0	1	0
1	D	112	908	581	156	166	5	0	0	0
1	E	113	933	596	165	167	5	0	0	0

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	124	SER	-	expression tag	UNP Q68DU8
A	125	MET	-	expression tag	UNP Q68DU8
B	124	SER	-	expression tag	UNP Q68DU8
B	125	MET	-	expression tag	UNP Q68DU8
C	124	SER	-	expression tag	UNP Q68DU8
C	125	MET	-	expression tag	UNP Q68DU8
D	124	SER	-	expression tag	UNP Q68DU8
D	125	MET	-	expression tag	UNP Q68DU8
E	124	SER	-	expression tag	UNP Q68DU8
E	125	MET	-	expression tag	UNP Q68DU8

- 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 5 4 1	0	0
2	A	1	Total O P 5 4 1	0	0
2	B	1	Total O P 5 4 1	0	0
2	B	1	Total O P 5 4 1	0	0
2	C	1	Total O P 5 4 1	0	0
2	C	1	Total O P 5 4 1	0	0
2	C	1	Total O P 5 4 1	0	0
2	D	1	Total O P 5 4 1	0	0
2	D	1	Total O P 5 4 1	0	0
2	D	1	Total O P 5 4 1	0	0
2	E	1	Total O P 5 4 1	0	0
2	E	1	Total O P 5 4 1	0	0
2	E	1	Total O P 5 4 1	0	0
2	E	1	Total O P 5 4 1	0	0

*Continued on next page...*

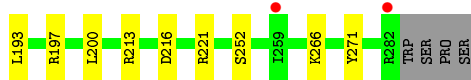
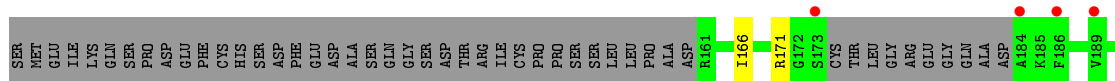
*Continued from previous page...*

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

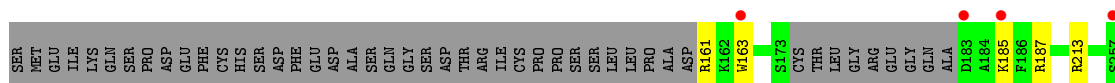
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	16	Total	O	0	0
			16	16		
3	B	27	Total	O	0	0
			27	27		
3	C	27	Total	O	0	0
			27	27		
3	D	36	Total	O	0	0
			36	36		
3	E	30	Total	O	0	0
			30	30		





● Molecule 1: BTB/POZ domain-containing protein KCTD16





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	74.43Å 105.62Å 107.55Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	60.00 – 2.23 75.36 – 2.23	Depositor EDS
% Data completeness (in resolution range)	97.4 (60.00-2.23) 98.2 (75.36-2.23)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.64 (at 2.22Å)	Xtrriage
Refinement program	PHENIX (1.13_2998: ???)	Depositor
R, $R_{free}$	0.230 , 0.275 0.231 , 0.274	Depositor DCC
$R_{free}$ test set	2071 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	38.9	Xtrriage
Anisotropy	0.557	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 55.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	0.002 for -h,l,k	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	4838	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	56.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 15.94% 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: 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	0.24	0/950	0.42	0/1280
1	B	0.24	0/955	0.43	0/1286
1	C	0.25	0/953	0.42	0/1283
1	D	0.25	0/932	0.43	0/1258
1	E	0.24	0/957	0.43	0/1287
All	All	0.25	0/4747	0.42	0/6394

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	926	0	878	5	0
1	B	931	0	884	7	0
1	C	929	0	887	10	0
1	D	908	0	852	10	0
1	E	933	0	898	7	0
2	A	10	0	0	0	0
2	B	10	0	0	0	0
2	C	15	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	15	0	0	1	0
2	E	25	0	0	0	0
3	A	16	0	0	0	0
3	B	27	0	0	0	0
3	C	27	0	0	1	0
3	D	36	0	0	2	0
3	E	30	0	0	1	0
All	All	4838	0	4399	29	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 3.

All (29) 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:E:187:ARG:NH1	3:E:401:HOH:O	2.30	0.63
1:C:218:ALA:HB3	1:C:221:ARG:HG3	1.80	0.62
1:C:202:LYS:NZ	1:C:211:GLU:OE2	2.34	0.59
1:D:171:ARG:HG3	1:E:185:LYS:HG3	1.85	0.58
1:D:221:ARG:NH1	2:D:303:PO4:O3	2.38	0.57
1:A:266:LYS:HZ3	1:B:268:TRP:HD1	1.53	0.56
1:A:268:TRP:HD1	1:E:266:LYS:HZ3	1.55	0.55
1:C:217:ARG:NH2	3:C:404:HOH:O	2.39	0.54
1:D:216:ASP:OD1	3:D:401:HOH:O	2.18	0.54
1:D:166:ILE:HD11	1:D:200:LEU:HB3	1.92	0.51
1:A:172:GLY:HA3	1:A:270:SER:HB3	1.94	0.49
1:A:218:ALA:HB3	1:A:221:ARG:HG2	1.94	0.49
1:D:266:LYS:HZ3	1:E:268:TRP:HD1	1.61	0.49
1:E:161:ARG:HD3	1:E:163:TRP:CZ2	2.49	0.48
1:D:266:LYS:NZ	1:E:266:LYS:O	2.37	0.47
1:C:238:ASP:O	1:C:242:GLU:HG3	2.15	0.46
1:B:171:ARG:NH2	1:C:186:PHE:O	2.43	0.46
1:B:273:GLU:OE1	1:C:274:TYR:OH	2.27	0.46
1:C:251[B]:ASN:OD1	1:D:252:SER:HB2	2.17	0.45
1:D:197:ARG:HD2	1:D:200:LEU:HD13	1.99	0.45
1:D:193:LEU:HD12	1:D:271:TYR:HE1	1.83	0.44
1:D:221:ARG:NH1	3:D:405:HOH:O	2.44	0.44
1:C:271:TYR:CE2	1:C:273:GLU:HG3	2.53	0.44
1:B:163:TRP:CH2	1:C:242:GLU:HG2	2.52	0.43
1:A:242:GLU:HG2	1:E:163:TRP:CH2	2.53	0.43
1:B:161:ARG:HD3	1:B:163:TRP:CZ2	2.55	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:198:ILE:HG22	1:C:202:LYS:HD2	2.02	0.42
1:B:218:ALA:HB3	1:B:221:ARG:HG2	2.02	0.42
1:B:238:ASP:O	1:B:242:GLU:HG3	2.21	0.41

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	110/163 (68%)	107 (97%)	3 (3%)	0	100	100
1	B	110/163 (68%)	108 (98%)	2 (2%)	0	100	100
1	C	109/163 (67%)	108 (99%)	1 (1%)	0	100	100
1	D	108/163 (66%)	107 (99%)	1 (1%)	0	100	100
1	E	109/163 (67%)	107 (98%)	2 (2%)	0	100	100
All	All	546/815 (67%)	537 (98%)	9 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	97/145 (67%)	95 (98%)	2 (2%)	53	60

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	97/145 (67%)	96 (99%)	1 (1%)	76	82
1	C	98/145 (68%)	98 (100%)	0	100	100
1	D	95/145 (66%)	94 (99%)	1 (1%)	73	80
1	E	99/145 (68%)	98 (99%)	1 (1%)	76	82
All	All	486/725 (67%)	481 (99%)	5 (1%)	76	82

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

Mol	Chain	Res	Type
1	A	173	SER
1	A	191	ARG
1	B	267	ILE
1	D	213	ARG
1	E	213	ARG

Some 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](#)

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](#)

15 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	E	302	-	4,4,4	0.93	0	6,6,6	0.43	0
2	PO4	C	302	-	4,4,4	0.92	0	6,6,6	0.41	0
2	PO4	C	303	-	4,4,4	0.91	0	6,6,6	0.45	0
2	PO4	D	303	-	4,4,4	0.92	0	6,6,6	0.42	0
2	PO4	E	305	-	4,4,4	0.93	0	6,6,6	0.40	0
2	PO4	E	303	-	4,4,4	0.92	0	6,6,6	0.47	0
2	PO4	B	302	-	4,4,4	0.92	0	6,6,6	0.41	0
2	PO4	D	301	-	4,4,4	0.93	0	6,6,6	0.42	0
2	PO4	E	304	-	4,4,4	0.92	0	6,6,6	0.43	0
2	PO4	D	302	-	4,4,4	0.92	0	6,6,6	0.43	0
2	PO4	B	301	-	4,4,4	0.92	0	6,6,6	0.44	0
2	PO4	A	301	-	4,4,4	0.92	0	6,6,6	0.43	0
2	PO4	C	301	-	4,4,4	0.93	0	6,6,6	0.45	0
2	PO4	E	301	-	4,4,4	0.92	0	6,6,6	0.44	0
2	PO4	A	302	-	4,4,4	0.93	0	6,6,6	0.41	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	D	303	PO4	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

### 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	114/163 (69%)	0.50	6 (5%) 26 26	33, 54, 91, 107	0
1	B	114/163 (69%)	0.49	8 (7%) 16 15	29, 48, 90, 115	0
1	C	112/163 (68%)	0.64	9 (8%) 12 12	27, 49, 89, 107	0
1	D	112/163 (68%)	0.56	6 (5%) 25 25	28, 45, 87, 97	0
1	E	113/163 (69%)	0.53	7 (6%) 20 20	28, 49, 90, 111	0
All	All	565/815 (69%)	0.54	36 (6%) 19 19	27, 49, 91, 115	0

All (36) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	259	ILE	4.6
1	B	186	PHE	4.3
1	E	259	ILE	4.1
1	D	186	PHE	4.1
1	A	184	ALA	4.0
1	B	184	ALA	3.9
1	C	259	ILE	3.8
1	E	258	PHE	3.7
1	C	216	ASP	3.5
1	C	262	TYR	3.4
1	E	183	ASP	3.2
1	D	259	ILE	3.2
1	D	173	SER	3.0
1	C	281	SER	3.0
1	C	184	ALA	3.0
1	C	186	PHE	2.9
1	B	189	VAL	2.9
1	B	257	SER	2.8
1	A	173	SER	2.7
1	D	189	VAL	2.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	184	ALA	2.7
1	E	257	SER	2.5
1	D	282	ARG	2.5
1	C	258	PHE	2.4
1	B	185	LYS	2.2
1	B	258	PHE	2.2
1	E	185	LYS	2.2
1	B	256	ALA	2.1
1	A	189	VAL	2.1
1	A	186	PHE	2.1
1	E	163	TRP	2.1
1	C	185	LYS	2.1
1	C	261	GLN	2.1
1	A	262	TYR	2.0
1	E	262	TYR	2.0
1	A	160	ASP	2.0

## 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. 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	PO4	E	305	5/5	0.75	0.48	80,82,97,115	0
2	PO4	B	302	5/5	0.85	0.15	53,69,88,92	0
2	PO4	B	301	5/5	0.86	0.11	93,94,118,123	0
2	PO4	E	304	5/5	0.87	0.28	67,82,96,106	5
2	PO4	E	303	5/5	0.90	0.26	47,50,72,78	5
2	PO4	E	301	5/5	0.90	0.18	93,111,116,125	0
2	PO4	A	302	5/5	0.91	0.20	62,65,70,73	5

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	PO4	C	303	5/5	0.92	0.17	63,67,89,90	0
2	PO4	D	301	5/5	0.92	0.11	99,100,111,116	0
2	PO4	C	301	5/5	0.93	0.16	92,100,103,107	0
2	PO4	D	302	5/5	0.94	0.15	62,65,76,80	0
2	PO4	E	302	5/5	0.94	0.12	52,53,65,72	0
2	PO4	A	301	5/5	0.94	0.15	98,99,106,119	0
2	PO4	D	303	5/5	0.95	0.20	36,46,53,64	5
2	PO4	C	302	5/5	0.95	0.14	46,60,73,74	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.