



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 06:57 PM GMT

PDB ID : 4NCK  
Title : Crystal Structure of Pyrococcus furiosis Rad50 R797G mutation  
Authors : Classen, S.; Williams, G.J.; Arvai, A.S.; Williams, R.S.  
Deposited on : 2013-10-24  
Resolution : 1.99 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 (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

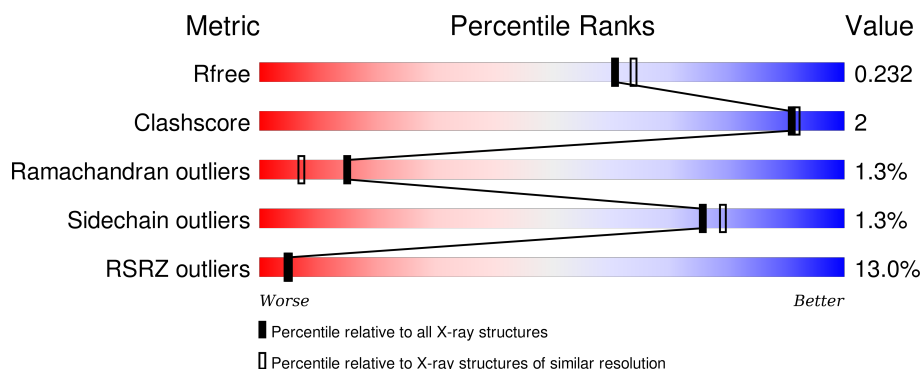
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.99 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	6249 (2.00-2.00)
Clashscore	102246	7340 (2.00-2.00)
Ramachandran outliers	100387	7248 (2.00-2.00)
Sidechain outliers	100360	7247 (2.00-2.00)
RSRZ outliers	91569	6262 (2.00-2.00)

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	339	<div> <div>13%</div> <div>84%</div> <div>8%</div> <div>8%</div> </div>
1	B	339	<div> <div>11%</div> <div>86%</div> <div>•</div> <div>9%</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
4	MG	B	903	-	-	-	X

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 9552 atoms, of which 4562 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 DNA double-strand break repair Rad50 ATPase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	313	Total	C	H	N	O	S	0	5	0
			4625	1512	2284	389	435	5			
1	B	309	Total	C	H	N	O	S	0	2	0
			4596	1491	2278	390	432	5			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	721	GLY	-	LINKER	UNP P58301
A	722	GLY	-	LINKER	UNP P58301
A	723	SER	-	LINKER	UNP P58301
A	724	GLY	-	LINKER	UNP P58301
A	725	GLY	-	LINKER	UNP P58301
A	797	GLY	ARG	ENGINEERED MUTATION	UNP P58301
B	721	GLY	-	LINKER	UNP P58301
B	722	GLY	-	LINKER	UNP P58301
B	723	SER	-	LINKER	UNP P58301
B	724	GLY	-	LINKER	UNP P58301
B	725	GLY	-	LINKER	UNP P58301
B	797	GLY	ARG	ENGINEERED MUTATION	UNP P58301

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	B	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Cl	0	0
			1	1		

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Mg	0	0
			1	1		

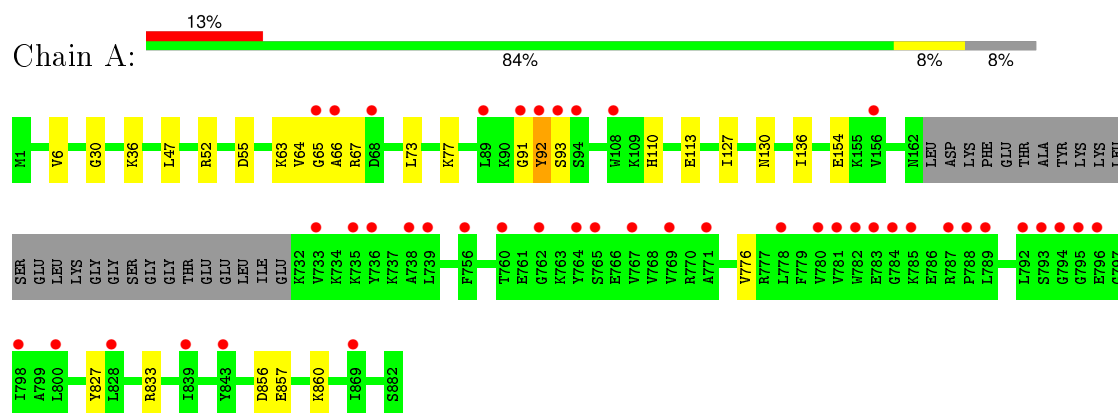
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	200	Total	O	0	0
			200	200		
5	B	124	Total	O	0	0
			124	124		

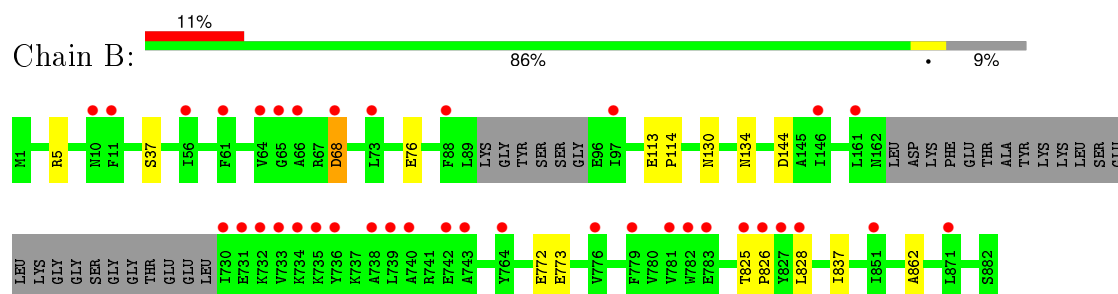
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: DNA double-strand break repair Rad50 ATPase



- Molecule 1: DNA double-strand break repair Rad50 ATPase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 2 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	49.67Å 68.06Å 202.90Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.24 – 1.99 48.24 – 1.99	Depositor EDS
% Data completeness (in resolution range)	98.6 (48.24-1.99) 99.3 (48.24-1.99)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.63 (at 2.00Å)	Xtriage
Refinement program	PHENIX (phenix.refine: dev_1233)	Depositor
R, $R_{free}$	0.181 , 0.230 0.194 , 0.232	Depositor DCC
$R_{free}$ test set	2000 reflections (4.39%)	DCC
Wilson B-factor (Å <sup>2</sup> )	43.0	Xtriage
Anisotropy	0.281	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 66.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 47574 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	9552	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	63.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.83% 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 [i](#)

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

Bond lengths and bond angles in the following residue types are not validated in this section: MG, SO4, CL

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	0/2390	0.69	0/3233
1	B	0.53	0/2358	0.63	2/3194 (0.1%)
All	All	0.57	0/4748	0.66	2/6427 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	5	ARG	NE-CZ-NH2	-6.79	116.91	120.30
1	B	5	ARG	NE-CZ-NH1	5.08	122.84	120.30

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	2341	2284	2290	15	0
1	B	2318	2278	2268	7	0
2	B	5	0	0	0	0
3	B	1	0	0	0	0
4	B	1	0	0	0	0
5	A	200	0	0	1	2
5	B	124	0	0	3	2

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	4990	4562	4558	20	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 2.

All (20) 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:144:ASP:OD2	5:B:1045:HOH:O	2.12	0.67
1:A:6:VAL:HG23	1:A:73[A]:LEU:HD23	1.83	0.58
1:B:76:GLU:OE2	5:B:1032:HOH:O	2.18	0.57
1:A:77:LYS:NZ	5:A:924:HOH:O	2.35	0.53
1:A:6:VAL:HG23	1:A:73[A]:LEU:CD2	2.40	0.50
1:A:30:GLY:O	1:A:36:LYS:HE3	2.12	0.50
1:B:837:ILE:HG23	1:B:862:ALA:HB2	1.96	0.47
1:A:93:SER:OG	1:B:68:ASP:OD1	2.25	0.47
1:A:91:GLY:O	1:A:92:TYR:CB	2.62	0.47
1:A:63:LYS:O	1:A:65:GLY:N	2.48	0.47
1:A:52:ARG:NH2	1:A:55:ASP:OD2	2.49	0.45
1:A:857:GLU:OE1	1:A:860:LYS:NZ	2.31	0.45
1:A:833:ARG:NE	1:A:856:ASP:OD2	2.49	0.45
1:B:113:GLU:OE1	5:B:1030:HOH:O	2.21	0.43
1:A:30:GLY:O	1:A:36:LYS:CE	2.67	0.42
1:A:110:HIS:CD2	1:B:114:PRO:HG2	2.55	0.41
1:B:130:ASN:ND2	1:B:134:ASN:HD22	2.18	0.41
1:A:154:GLU:HA	1:A:776:VAL:HG23	2.01	0.41
1:A:127:ILE:HD11	1:A:136:ILE:CD1	2.50	0.41
1:A:47:LEU:HD21	1:A:73[B]:LEU:HD21	2.03	0.40

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)
5:A:946:HOH:O	5:B:1085:HOH:O[1_655]	2.16	0.04
5:A:930:HOH:O	5:B:1087:HOH:O[1_655]	2.18	0.02

## 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	314/339 (93%)	303 (96%)	7 (2%)	4 (1%)	15	7
1	B	305/339 (90%)	293 (96%)	8 (3%)	4 (1%)	15	7
All	All	619/678 (91%)	596 (96%)	15 (2%)	8 (1%)	15	7

All (8) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	64	VAL
1	A	92	TYR
1	B	773	GLU
1	B	826	PRO
1	A	66	ALA
1	A	67	ARG
1	B	772	GLU
1	B	828	LEU

### 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	225/293 (77%)	222 (99%)	3 (1%)	76	79
1	B	229/293 (78%)	226 (99%)	3 (1%)	76	79
All	All	454/586 (78%)	448 (99%)	6 (1%)	76	79

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

Mol	Chain	Res	Type
1	A	113	GLU
1	A	130	ASN
1	A	827	TYR
1	B	37	SER
1	B	68	ASP
1	B	825	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	855	HIS
1	B	130	ASN

### 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 3 ligands modelled in this entry, 2 are monoatomic - leaving 1 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	SO4	B	901	-	4,4,4	0.19	0	6,6,6	0.12	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	SO4	B	901	-	-	0/0/0/0	0/0/0/0

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.

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	313/339 (92%)	0.98	44 (14%) 4 4	25, 49, 110, 140	0
1	B	309/339 (91%)	0.84	37 (11%) 6 6	27, 59, 107, 137	0
All	All	622/678 (91%)	0.91	81 (13%) 5 5	25, 55, 109, 140	0

All (81) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	736	TYR	8.1
1	B	66	ALA	7.8
1	B	733	VAL	7.2
1	A	764	TYR	6.9
1	A	767	VAL	6.8
1	B	65	GLY	6.8
1	A	733	VAL	6.7
1	A	92	TYR	6.4
1	A	781	VAL	6.0
1	B	739	LEU	6.0
1	A	782	TRP	5.9
1	B	61	PHE	5.6
1	A	792	LEU	5.6
1	B	827	TYR	5.5
1	B	731	GLU	5.0
1	A	789	LEU	5.0
1	B	730	ILE	4.8
1	A	93	SER	4.8
1	A	787	ARG	4.8
1	B	64	VAL	4.7
1	A	784	GLY	4.4
1	A	843	TYR	4.3
1	B	782	TRP	4.3
1	A	735	LYS	4.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	94	SER	4.3
1	A	739	LEU	4.2
1	B	88	PHE	4.2
1	A	762	GLY	4.1
1	A	795	GLY	4.0
1	A	771	ALA	4.0
1	A	756	PHE	3.8
1	B	734	LYS	3.7
1	A	765	SER	3.6
1	B	779	PHE	3.5
1	A	760	THR	3.5
1	A	788	PRO	3.5
1	A	66	ALA	3.4
1	B	828	LEU	3.4
1	A	738	ALA	3.3
1	A	794	GLY	3.3
1	A	785	LYS	3.3
1	B	740	ALA	3.2
1	B	825	THR	3.2
1	A	91	GLY	3.1
1	A	798	ILE	3.0
1	B	732	LYS	3.0
1	B	735	LYS	3.0
1	A	800	LEU	3.0
1	A	796	GLU	2.9
1	B	68	ASP	2.9
1	B	826	PRO	2.9
1	B	781	VAL	2.9
1	A	778	LEU	2.9
1	A	68	ASP	2.8
1	A	780	VAL	2.8
1	B	10	ASN	2.7
1	B	743	ALA	2.6
1	B	736	TYR	2.5
1	A	793	SER	2.4
1	A	828	LEU	2.4
1	B	764	TYR	2.4
1	B	146	ILE	2.3
1	B	783	GLU	2.3
1	B	871	LEU	2.3
1	A	156	VAL	2.3
1	A	839	ILE	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	869[A]	ILE	2.3
1	B	161	LEU	2.3
1	A	89	LEU	2.2
1	B	776	VAL	2.2
1	B	97	ILE	2.2
1	B	73	LEU	2.2
1	A	65	GLY	2.2
1	A	783	GLU	2.2
1	A	108	TRP	2.2
1	B	742	GLU	2.1
1	B	11	PHE	2.1
1	B	56	ILE	2.1
1	B	851	ILE	2.0
1	A	769	VAL	2.0
1	B	738	ALA	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. 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
4	MG	B	903	1/1	0.75	0.67	38.10	80,80,80,80	0
2	SO4	B	901	5/5	0.81	0.16	-0.07	118,119,120,121	0
3	CL	B	902	1/1	0.96	0.07	-2.64	50,50,50,50	0

## 6.5 Other polymers

There are no such residues in this entry.