



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 5, 2021 – 04:02 PM EDT

PDB ID : 7LKM  
Title : The PilB(N-terminal\_P70S mutant)-PilZ complex of the Type IV pilus from *Xanthomonas citri*  
Authors : Llontop, E.E.; Guzzo, C.R.; Farah, C.S.  
Deposited on : 2021-02-02  
Resolution : 2.00 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

<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 : 2.23.1  
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) : 2.23.1

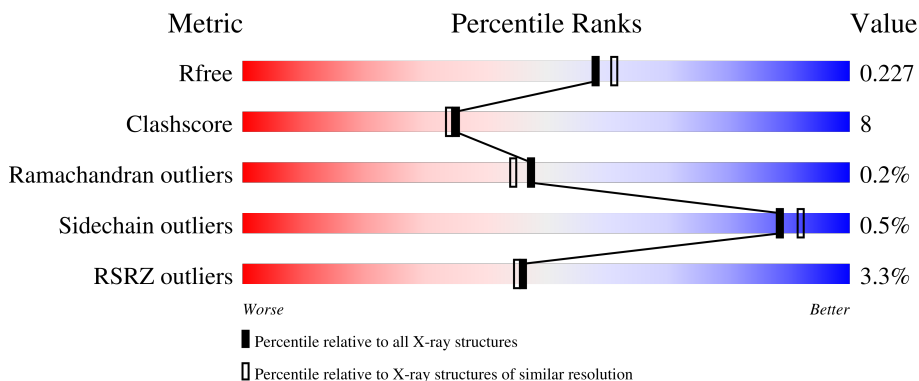
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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 <sub>free</sub>	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (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 of 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	173	 3% 69% 14% •• 14%
1	B	173	 2% 75% 9% • 14%
2	C	117	 3% 79% 14% • 6%
2	D	117	 3% 81% 11% • 7%

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 4417 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 Pilus biogenesis protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	149	1170	737	210	219	4	0	4	0
1	B	148	1167	736	208	219	4	0	4	0

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-9	MET	-	initiating methionine	UNP Q8PHL2
A	-8	GLY	-	expression tag	UNP Q8PHL2
A	-7	SER	-	expression tag	UNP Q8PHL2
A	-6	SER	-	expression tag	UNP Q8PHL2
A	-5	HIS	-	expression tag	UNP Q8PHL2
A	-4	HIS	-	expression tag	UNP Q8PHL2
A	-3	HIS	-	expression tag	UNP Q8PHL2
A	-2	HIS	-	expression tag	UNP Q8PHL2
A	-1	HIS	-	expression tag	UNP Q8PHL2
A	0	HIS	-	expression tag	UNP Q8PHL2
A	1	SER	-	expression tag	UNP Q8PHL2
A	2	SER	-	expression tag	UNP Q8PHL2
A	3	GLY	-	expression tag	UNP Q8PHL2
A	4	LEU	-	expression tag	UNP Q8PHL2
A	5	VAL	-	expression tag	UNP Q8PHL2
A	6	PRO	-	expression tag	UNP Q8PHL2
A	7	ARG	-	expression tag	UNP Q8PHL2
A	8	GLY	-	expression tag	UNP Q8PHL2
A	9	SER	-	expression tag	UNP Q8PHL2
A	10	HIS	-	expression tag	UNP Q8PHL2
A	11	MET	-	expression tag	UNP Q8PHL2
A	70	SER	PRO	engineered mutation	UNP Q8PHL2
B	-9	MET	-	initiating methionine	UNP Q8PHL2
B	-8	GLY	-	expression tag	UNP Q8PHL2
B	-7	SER	-	expression tag	UNP Q8PHL2

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	-6	SER	-	expression tag	UNP Q8PHL2
B	-5	HIS	-	expression tag	UNP Q8PHL2
B	-4	HIS	-	expression tag	UNP Q8PHL2
B	-3	HIS	-	expression tag	UNP Q8PHL2
B	-2	HIS	-	expression tag	UNP Q8PHL2
B	-1	HIS	-	expression tag	UNP Q8PHL2
B	0	HIS	-	expression tag	UNP Q8PHL2
B	1	SER	-	expression tag	UNP Q8PHL2
B	2	SER	-	expression tag	UNP Q8PHL2
B	3	GLY	-	expression tag	UNP Q8PHL2
B	4	LEU	-	expression tag	UNP Q8PHL2
B	5	VAL	-	expression tag	UNP Q8PHL2
B	6	PRO	-	expression tag	UNP Q8PHL2
B	7	ARG	-	expression tag	UNP Q8PHL2
B	8	GLY	-	expression tag	UNP Q8PHL2
B	9	SER	-	expression tag	UNP Q8PHL2
B	10	HIS	-	expression tag	UNP Q8PHL2
B	11	MET	-	expression tag	UNP Q8PHL2
B	70	SER	PRO	engineered mutation	UNP Q8PHL2

- Molecule 2 is a protein called Type IV fimbriae assembly protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	110	Total	C	N	O	S	0	1	0
			814	526	134	151	3			
2	D	109	Total	C	N	O	S	0	0	0
			798	516	134	145	3			

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	C	1	Total	O	S	0	0
			5	4	1		
3	C	1	Total	O	S	0	0
			5	4	1		
3	C	1	Total	O	S	0	0
			5	4	1		
3	D	1	Total	O	S	0	0
			5	4	1		
3	D	1	Total	O	S	0	0
			5	4	1		

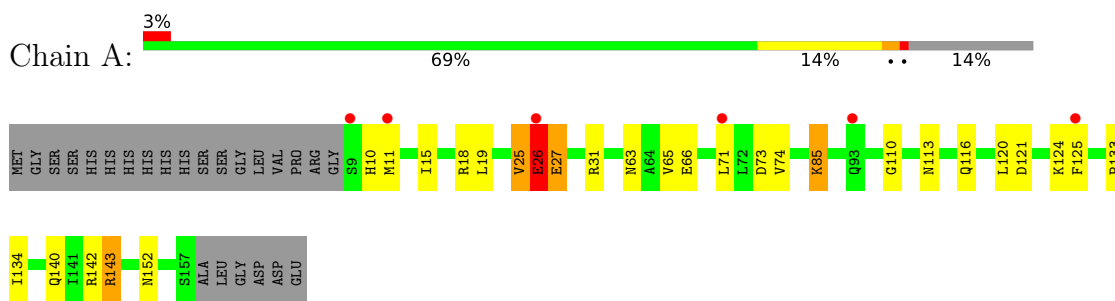
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	135	Total	O	0	0
			135	135		
4	B	123	Total	O	0	0
			123	123		
4	C	101	Total	O	0	0
			101	101		
4	D	84	Total	O	0	0
			84	84		

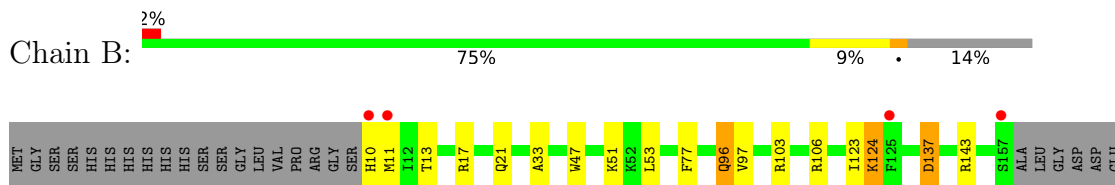
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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 ( $RSRZ > 2$ ). 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.

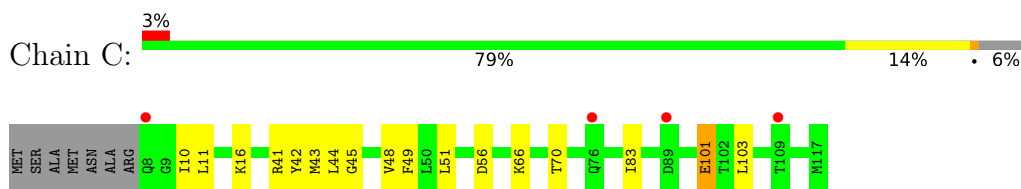
- Molecule 1: Pilus biogenesis protein



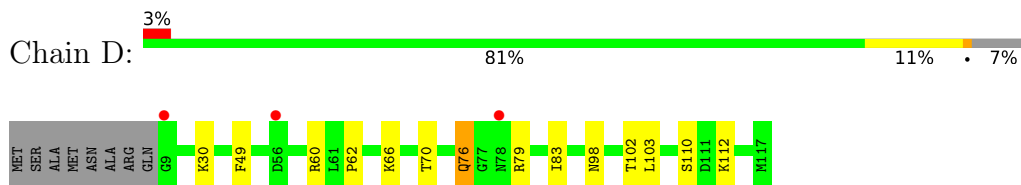
- Molecule 1: Pilus biogenesis protein



- Molecule 2: Type IV fimbriae assembly protein



- Molecule 2: Type IV fimbriae assembly protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	39.06Å 122.41Å 62.73Å 90.00° 98.98° 90.00°	Depositor
Resolution (Å)	28.00 – 2.00 28.03 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.7 (28.00-2.00) 99.7 (28.03-2.00)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.45 (at 1.99Å)	Xtrriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, $R_{free}$	0.179 , 0.227 0.179 , 0.227	Depositor DCC
$R_{free}$ test set	1936 reflections (4.94%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	23.4	Xtrriage
Anisotropy	0.050	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 57.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	4417	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.69% 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: SO4

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	0.51	0/1187	0.88	7/1608 (0.4%)
1	B	0.47	0/1186	1.02	4/1610 (0.2%)
2	C	0.47	0/833	0.75	1/1134 (0.1%)
2	D	0.51	0/817	0.80	2/1112 (0.2%)
All	All	0.49	0/4023	0.88	14/5464 (0.3%)

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	2
1	B	0	1
2	C	0	1
All	All	0	4

There are no bond length outliers.

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	137	ASP	CB-CG-OD2	-21.08	99.33	118.30
1	B	137	ASP	CB-CG-OD1	17.47	134.03	118.30
1	A	143	ARG	NE-CZ-NH1	-14.06	113.27	120.30
1	A	26	GLU	CA-CB-CG	7.99	130.98	113.40
1	B	137	ASP	OD1-CG-OD2	-7.09	109.83	123.30

There are no chirality outliers.

All (4) planarity outliers are listed below:



Mol	Chain	Res	Type	Group
1	A	143	ARG	Sidechain
1	A	26	GLU	Sidechain
1	B	137	ASP	Sidechain
2	C	101	GLU	Sidechain

## 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	1170	0	1179	26	0
1	B	1167	0	1167	17	0
2	C	814	0	826	17	0
2	D	798	0	814	13	0
3	C	15	0	0	0	0
3	D	10	0	0	0	0
4	A	135	0	0	3	1
4	B	123	0	0	3	2
4	C	101	0	0	3	1
4	D	84	0	0	1	0
All	All	4417	0	3986	65	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 8.

The worst 5 of 65 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:65:VAL:HG21	2:C:41:ARG:HE	1.41	0.85
1:A:85:LYS:HE3	1:A:152:ASN:HB3	1.65	0.77
1:A:142:ARG:NH2	4:A:202:HOH:O	2.18	0.76
1:A:65:VAL:HG21	2:C:41:ARG:NE	2.01	0.75
2:C:11:LEU:HD12	2:C:48:VAL:HG21	1.76	0.67

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)
4:A:327:HOH:O	4:B:317:HOH:O[2_447]	1.93	0.27
4:B:279:HOH:O	4:C:307:HOH:O[1_656]	2.16	0.04

### 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	151/173 (87%)	147 (97%)	3 (2%)	1 (1%)	22	16
1	B	150/173 (87%)	149 (99%)	1 (1%)	0	100	100
2	C	109/117 (93%)	108 (99%)	1 (1%)	0	100	100
2	D	107/117 (92%)	106 (99%)	1 (1%)	0	100	100
All	All	517/580 (89%)	510 (99%)	6 (1%)	1 (0%)	47	44

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	27	GLU

#### 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	123/141 (87%)	121 (98%)	2 (2%)	62	67
1	B	123/141 (87%)	121 (98%)	2 (2%)	62	67
2	C	84/92 (91%)	84 (100%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	D	82/92 (89%)	82 (100%)	0	100	100
All	All	412/466 (88%)	408 (99%)	4 (1%)	88	81

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	11[A]	MET
1	A	11[B]	MET
1	B	96[A]	GLN
1	B	96[B]	GLN

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

Mol	Chain	Res	Type
1	B	93	GLN
2	D	76	GLN

### 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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

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	$\# Z  > 2$
3	SO4	C	201	-	4,4,4	0.16	0	6,6,6	0.19	0
3	SO4	D	202	-	4,4,4	0.15	0	6,6,6	0.15	0
3	SO4	C	203	-	4,4,4	0.19	0	6,6,6	0.28	0
3	SO4	D	201	-	4,4,4	0.15	0	6,6,6	0.12	0
3	SO4	C	202	-	4,4,4	0.17	0	6,6,6	0.16	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 [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	149/173 (86%)	0.03	6 (4%) 38 37	16, 24, 47, 69	0
1	B	148/173 (85%)	-0.03	4 (2%) 54 53	11, 25, 42, 65	0
2	C	110/117 (94%)	0.05	4 (3%) 42 42	13, 23, 50, 68	0
2	D	109/117 (93%)	-0.05	3 (2%) 53 51	13, 24, 52, 67	0
All	All	516/580 (88%)	-0.00	17 (3%) 46 45	11, 24, 48, 69	0

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	8	GLN	5.5
1	A	9	SER	4.2
1	A	11[A]	MET	4.2
1	A	26	GLU	3.7
2	C	109	THR	3.4

### 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 monosaccharides 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( $\text{\AA}^2$ )	Q<0.9
3	SO4	C	203	5/5	0.88	0.33	49,57,66,79	0
3	SO4	C	201	5/5	0.90	0.20	37,40,60,61	0
3	SO4	C	202	5/5	0.94	0.29	50,59,61,87	0
3	SO4	D	202	5/5	0.94	0.46	60,71,86,90	0
3	SO4	D	201	5/5	0.97	0.20	48,48,69,73	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.