



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 10, 2025 – 01:38 pm BST

PDB ID : 9Q9E / pdb\_00009q9e  
Title : Crystal structure of a TELSAM-SUMO1 fusion protein  
Authors : Coste, F.; Mishra, A.; Suskiewicz, M.J.  
Deposited on : 2025-02-26  
Resolution : 2.05 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4-5-2 with Phenix2.0rc1  
Mogul : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 2.0rc1  
EDS : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.003 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.43.1

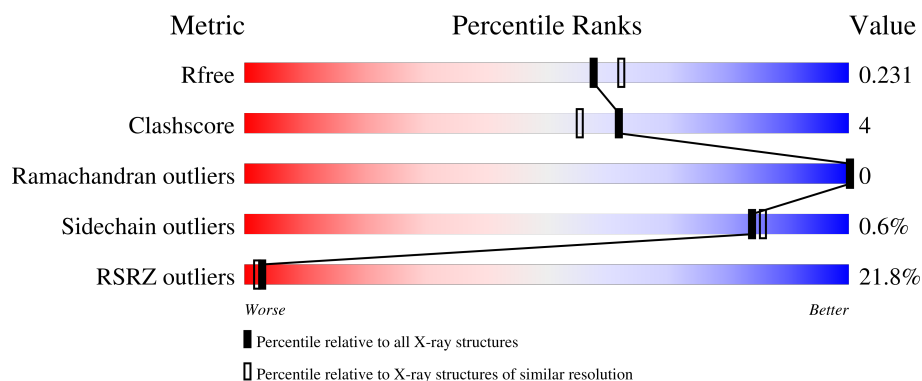
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.05 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))
$R_{free}$	164625	2096 (2.04-2.04)
Clashscore	180529	2229 (2.04-2.04)
Ramachandran outliers	177936	2217 (2.04-2.04)
Sidechain outliers	177891	2217 (2.04-2.04)
RSRZ outliers	164620	2096 (2.04-2.04)

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	167	<div> <div>7%</div> <div>82% 10% 8%</div> </div>
1	B	167	<div> <div>14%</div> <div>83% 10% 7%</div> </div>
1	C	167	<div> <div>39%</div> <div>83% 6% 11%</div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 3669 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 Transcription factor ETV6,Small ubiquitin-related modifier 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	154	Total	C	N	O	S	0	1	0
			1210	771	203	232	4			
1	B	156	Total	C	N	O	S	0	0	0
			1207	774	201	227	5			
1	C	149	Total	C	N	O	S	0	0	0
			1016	642	180	191	3			

There are 39 discrepancies between the modelled and reference sequences:

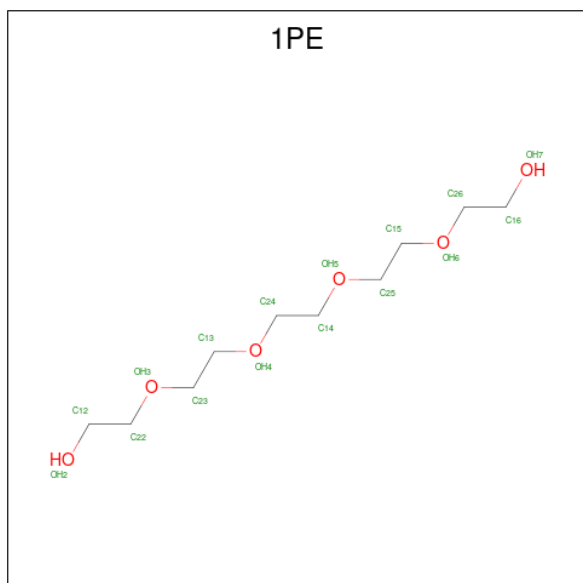
Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP P41212
A	2	HIS	-	expression tag	UNP P41212
A	3	HIS	-	expression tag	UNP P41212
A	4	HIS	-	expression tag	UNP P41212
A	5	HIS	-	expression tag	UNP P41212
A	6	HIS	-	expression tag	UNP P41212
A	7	HIS	-	expression tag	UNP P41212
A	10	ALA	ARG	engineered mutation	UNP P41212
A	73	GLU	VAL	engineered mutation	UNP P41212
A	83	ALA	LYS	engineered mutation	UNP P41212
A	85	PRO	-	linker	UNP P41212
A	86	ALA	-	linker	UNP P41212
A	87	ALA	-	linker	UNP P41212
B	1	MET	-	initiating methionine	UNP P41212
B	2	HIS	-	expression tag	UNP P41212
B	3	HIS	-	expression tag	UNP P41212
B	4	HIS	-	expression tag	UNP P41212
B	5	HIS	-	expression tag	UNP P41212
B	6	HIS	-	expression tag	UNP P41212
B	7	HIS	-	expression tag	UNP P41212
B	10	ALA	ARG	engineered mutation	UNP P41212
B	73	GLU	VAL	engineered mutation	UNP P41212
B	83	ALA	LYS	engineered mutation	UNP P41212

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	85	PRO	-	linker	UNP P41212
B	86	ALA	-	linker	UNP P41212
B	87	ALA	-	linker	UNP P41212
C	1	MET	-	initiating methionine	UNP P41212
C	2	HIS	-	expression tag	UNP P41212
C	3	HIS	-	expression tag	UNP P41212
C	4	HIS	-	expression tag	UNP P41212
C	5	HIS	-	expression tag	UNP P41212
C	6	HIS	-	expression tag	UNP P41212
C	7	HIS	-	expression tag	UNP P41212
C	10	ALA	ARG	engineered mutation	UNP P41212
C	73	GLU	VAL	engineered mutation	UNP P41212
C	83	ALA	LYS	engineered mutation	UNP P41212
C	85	PRO	-	linker	UNP P41212
C	86	ALA	-	linker	UNP P41212
C	87	ALA	-	linker	UNP P41212

- Molecule 2 is PENTAETHYLENE GLYCOL (CCD ID: 1PE) (formula: C<sub>10</sub>H<sub>22</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			10	6	4		
2	A	1	Total	C	O	0	0
			13	8	5		

- Molecule 3 is PHOSPHATE ION (CCD ID: PO4) (formula: O<sub>4</sub>P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	B	1	Total	O	P	0	0
			5	4	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	86	Total	O	0	0
			86	86		
4	B	71	Total	O	0	0
			71	71		
4	C	51	Total	O	0	0
			51	51		



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	54.73Å 99.69Å 113.58Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.97 – 2.05 47.97 – 2.05	Depositor EDS
% Data completeness (in resolution range)	99.9 (47.97-2.05) 96.5 (47.97-2.05)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.15 (at 2.05Å)	Xtriage
Refinement program	PHENIX 1.20.1-4487	Depositor
R, $R_{free}$	0.209 , 0.231 0.209 , 0.231	Depositor DCC
$R_{free}$ test set	1917 reflections (4.82%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	24.2	Xtriage
Anisotropy	0.354	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 85.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	3669	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.59% 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, 1PE

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	0.09	0/1243	0.28	0/1694
1	B	0.12	0/1236	0.28	0/1681
1	C	0.09	0/1038	0.26	0/1424
All	All	0.10	0/3517	0.27	0/4799

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

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	1210	0	1103	11	0
1	B	1207	0	1129	10	0
1	C	1016	0	810	6	0
2	A	23	0	30	3	0
3	B	5	0	0	0	0
4	A	86	0	0	1	0
4	B	71	0	0	0	0
4	C	51	0	0	1	0
All	All	3669	0	3072	26	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 4.

All (26) 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:12:PRO:HD2	1:A:15:LEU:HD12	1.74	0.70
1:A:106:PHE:HA	2:A:202:1PE:H142	1.74	0.68
1:C:146:THR:HG22	1:C:149:GLU:HB3	1.78	0.64
1:B:9:ILE:H	1:B:84:GLN:HE21	1.46	0.64
1:A:97:ILE:HD12	1:A:103:GLU:HG2	1.81	0.61
1:A:38:PHE:HB3	1:B:54:ALA:HB2	1.85	0.59
1:A:41:ARG:NH1	2:A:201:1PE:OH3	2.36	0.59
1:B:38:PHE:HB3	1:C:54:ALA:HB2	1.85	0.59
1:A:17:LEU:HD13	1:A:21:TYR:HB2	1.86	0.57
1:B:98:GLY:HA2	1:B:160:VAL:HB	1.89	0.55
1:C:93:LYS:O	1:C:155:GLU:N	2.37	0.50
1:C:45:SER:O	4:C:201:HOH:O	2.20	0.47
1:B:72:ASP:OD1	1:B:73:GLU:N	2.48	0.46
1:B:14:HIS:HD2	1:B:29:GLN:OE1	1.99	0.45
1:A:64:ARG:NH1	4:A:308:HOH:O	2.47	0.45
1:A:19:PRO:O	1:A:52:GLY:HA3	2.17	0.44
1:B:131:SER:O	1:B:163:GLU:N	2.44	0.44
1:B:12:PRO:HG2	1:B:15:LEU:HD12	2.00	0.44
1:A:33:TRP:HZ3	1:A:77:LEU:HD21	1.83	0.43
1:B:103:GLU:O	1:B:104:ILE:HD13	2.18	0.43
1:A:60:LYS:HD3	1:A:75:TYR:CD2	2.52	0.43
2:A:202:1PE:H141	2:A:202:1PE:H132	1.78	0.43
1:B:12:PRO:HB2	1:B:14:HIS:CE1	2.55	0.41
1:C:114:LEU:HG	1:C:147:PRO:HD3	2.02	0.41
1:A:41:ARG:HE	1:A:41:ARG:HB3	1.72	0.41
1:C:22:TRP:HB2	1:C:52:GLY:HA3	2.04	0.40

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	153/167 (92%)	151 (99%)	2 (1%)	0	100	100
1	B	154/167 (92%)	153 (99%)	1 (1%)	0	100	100
1	C	147/167 (88%)	143 (97%)	4 (3%)	0	100	100
All	All	454/501 (91%)	447 (98%)	7 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	122/149 (82%)	121 (99%)	1 (1%)	79	80
1	B	122/149 (82%)	121 (99%)	1 (1%)	79	80
1	C	75/149 (50%)	75 (100%)	0	100	100
All	All	319/447 (71%)	317 (99%)	2 (1%)	84	86

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	14	HIS
1	B	20	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	18	GLN
1	A	84	GLN
1	B	14	HIS
1	B	36	ASN
1	B	84	GLN
1	B	105	HIS
1	C	18	GLN

### 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 oligosaccharides in this entry.

## 5.6 Ligand geometry ⓘ

3 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	PO4	B	201	-	4,4,4	0.92	0	6,6,6	0.43	0
2	1PE	A	201	-	9,9,15	0.09	0	8,8,14	0.16	0
2	1PE	A	202	-	12,12,15	0.13	0	11,11,14	0.09	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	1PE	A	201	-	-	1/7/7/13	-
2	1PE	A	202	-	-	3/10/10/13	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	202	1PE	C14-C24-OH4-C13
2	A	202	1PE	C25-C15-OH6-C26
2	A	202	1PE	OH5-C14-C24-OH4
2	A	201	1PE	OH2-C12-C22-OH3

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	201	1PE	1	0
2	A	202	1PE	2	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	154/167 (92%)	0.24	12 (7%)	20 22	16, 34, 63, 101	1 (0%)
1	B	156/167 (93%)	0.59	23 (14%)	7 6	17, 40, 84, 103	0
1	C	149/167 (89%)	1.53	65 (43%)	1 0	18, 75, 157, 189	0
All	All	459/501 (91%)	0.78	100 (21%)	3 2	16, 40, 141, 189	1 (0%)

All (100) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	160	VAL	6.7
1	C	132	LEU	5.9
1	C	161	TYR	5.7
1	C	158	ILE	5.5
1	C	121	TYR	5.4
1	C	91	TYR	4.9
1	C	135	LEU	4.9
1	C	120	SER	4.6
1	C	133	ARG	4.5
1	C	122	CYS	4.5
1	B	10	ALA	4.4
1	C	102	SER	4.3
1	C	98	GLY	4.2
1	A	11	LEU	4.1
1	C	94	LEU	4.1
1	B	87	ALA	4.1
1	C	157	VAL	4.1
1	C	134	PHE	4.0
1	C	97	ILE	4.0
1	C	156	ASP	3.8
1	C	130	ASN	3.8
1	C	96	VAL	3.8
1	A	12	PRO	3.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	131	SER	3.8
1	C	90	GLU	3.7
1	C	103	GLU	3.7
1	C	104	ILE	3.7
1	B	14	HIS	3.7
1	C	128	PRO	3.6
1	C	152	MET	3.5
1	B	101	SER	3.5
1	C	108	VAL	3.5
1	C	13	ALA	3.4
1	C	101	SER	3.4
1	B	88	GLU	3.4
1	B	102	SER	3.4
1	B	129	MET	3.3
1	C	95	LYS	3.3
1	C	92	ILE	3.2
1	B	130	ASN	3.2
1	C	155	GLU	3.2
1	C	111	THR	3.1
1	C	129	MET	3.1
1	C	14	HIS	3.1
1	C	118	LYS	3.1
1	C	15	LEU	3.1
1	C	140	ARG	3.1
1	C	153	GLU	3.1
1	B	11	LEU	3.0
1	C	106	PHE	3.0
1	C	117	LEU	3.0
1	C	159	GLU	3.0
1	C	110	MET	2.9
1	C	105	HIS	2.9
1	C	119	GLU	2.9
1	A	129	MET	2.9
1	B	13	ALA	2.8
1	C	136	PHE	2.8
1	B	99	GLN	2.8
1	C	87	ALA	2.8
1	C	126	GLY	2.8
1	C	142	ALA	2.7
1	B	103	GLU	2.7
1	A	161	TYR	2.6
1	B	97	ILE	2.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	101	SER	2.6
1	C	127	VAL	2.6
1	B	85	PRO	2.6
1	C	125	GLN	2.5
1	C	154	GLU	2.5
1	C	144	ASN	2.5
1	B	15	LEU	2.4
1	C	150	LEU	2.4
1	A	99	GLN	2.4
1	B	126	GLY	2.4
1	A	13	ALA	2.4
1	A	14	HIS	2.4
1	C	141	ILE	2.4
1	C	100	ASP	2.4
1	B	16	ARG	2.4
1	B	8	SER	2.3
1	B	84	GLN	2.3
1	B	125	GLN	2.3
1	B	9	ILE	2.3
1	C	113	HIS	2.3
1	B	127	VAL	2.2
1	A	164	GLN	2.2
1	C	99	GLN	2.2
1	B	163	GLU	2.2
1	C	138	GLY	2.2
1	C	148	LYS	2.2
1	A	17	LEU	2.2
1	A	84	GLN	2.1
1	C	89	GLY	2.1
1	C	124	ARG	2.1
1	C	86	ALA	2.1
1	C	107	LYS	2.1
1	C	115	LYS	2.1
1	C	114	LEU	2.0
1	A	110	MET	2.0

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

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
2	1PE	A	201	10/16	0.73	0.18	43,57,71,74	0
2	1PE	A	202	13/16	0.74	0.16	53,62,67,68	0
3	PO4	B	201	5/5	0.90	0.13	15,19,29,37	5

### 6.5 Other polymers [i](#)

There are no such residues in this entry.