



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 19, 2016 – 07:58 PM GMT

PDB ID : 4XES  
Title : Structure of active-like neurotensin receptor  
Authors : Krumm, B.E.; White, J.F.; Shah, P.; Grisshammer, R.  
Deposited on : 2014-12-24  
Resolution : 2.60 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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026982  
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) : rb-20026982

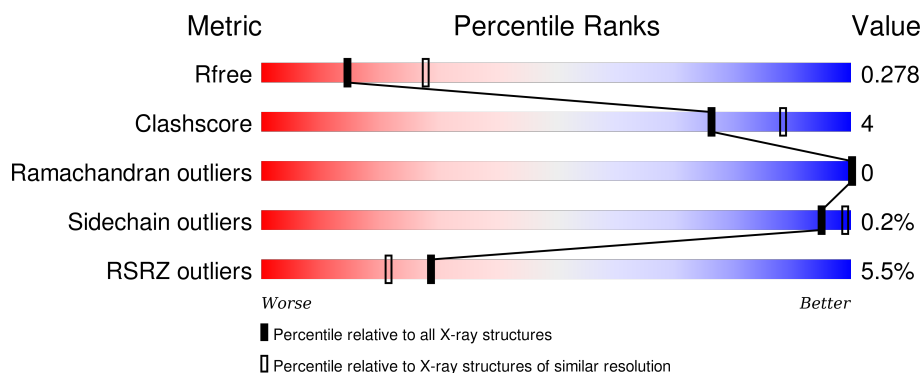
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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	2328 (2.60-2.60)
Clashscore	102246	2679 (2.60-2.60)
Ramachandran outliers	100387	2635 (2.60-2.60)
Sidechain outliers	100360	2635 (2.60-2.60)
RSRZ outliers	91569	2334 (2.60-2.60)

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	541	<div> <div>5%</div> <div>80%</div> <div>7%</div> <div>13%</div> </div>
2	B	6	<div> <div>17%</div> <div>100%</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	PEG	A	1204	-	-	-	X
4	PEG	A	1205	-	-	-	X
4	PEG	A	1206	-	-	-	X

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 3923 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 Neurotensin receptor type 1, Endolysin chimera.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	471	3714	2409	634	652	19	164	0	0

There are 37 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	33	ASP	-	expression tag	UNP P20789
A	34	TYR	-	expression tag	UNP P20789
A	35	LYS	-	expression tag	UNP P20789
A	36	ASP	-	expression tag	UNP P20789
A	37	ASP	-	expression tag	UNP P20789
A	38	ASP	-	expression tag	UNP P20789
A	39	ASP	-	expression tag	UNP P20789
A	40	ALA	-	expression tag	UNP P20789
A	41	THR	-	expression tag	UNP P20789
A	42	SER	-	expression tag	UNP P20789
A	86	LEU	ALA	engineered mutation	UNP P20789
A	166	ALA	GLU	engineered mutation	UNP P20789
A	215	ALA	GLY	engineered mutation	UNP P20789
A	360	ALA	VAL	engineered mutation	UNP P20789
A	989	ALA	-	linker	UNP P20789
A	990	HIS	-	linker	UNP P20789
A	991	HIS	-	linker	UNP P20789
A	992	HIS	-	linker	UNP P20789
A	993	HIS	-	linker	UNP P20789
A	994	HIS	-	linker	UNP P20789
A	995	HIS	-	linker	UNP P20789
A	996	HIS	-	linker	UNP P20789
A	997	HIS	-	linker	UNP P20789
A	998	HIS	-	linker	UNP P20789
A	999	HIS	-	linker	UNP P20789
A	1000	GLY	-	linker	UNP P20789
A	1001	GLY	-	linker	UNP P20789

*Continued on next page...*

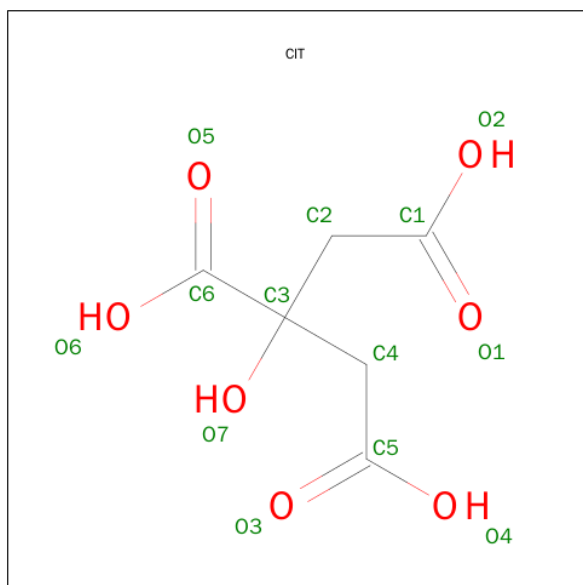
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	1012	GLY	ARG	conflict	UNP P00720
A	1054	THR	CYS	conflict	UNP P00720
A	1097	ALA	CYS	conflict	UNP P00720
A	1122	ASN	GLN	conflict	UNP P00720
A	1123	ASN	GLN	conflict	UNP P00720
A	1137	ARG	ILE	conflict	UNP P00720
A	1162	GLY	-	expression tag	UNP P00720
A	1163	SER	-	expression tag	UNP P00720
A	1164	GLY	-	expression tag	UNP P00720
A	1165	SER	-	expression tag	UNP P00720

- Molecule 2 is a protein called Neurotensin/neuromedin N.

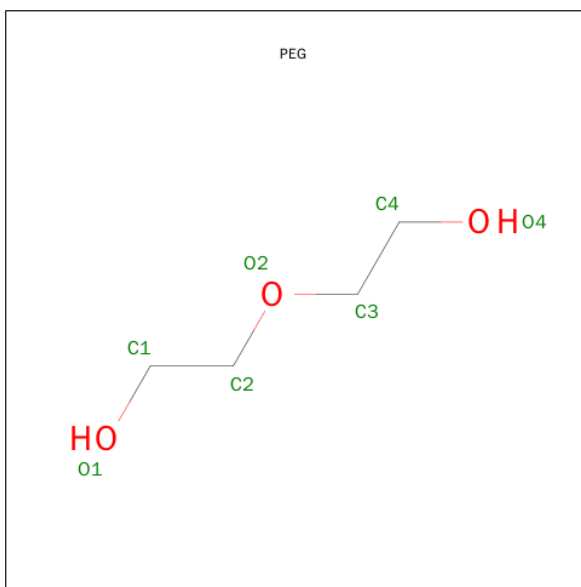
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	6	Total	C	N	O	0	0	0
			58	38	12	8			

- Molecule 3 is CITRIC ACID (three-letter code: CIT) (formula: C<sub>6</sub>H<sub>8</sub>O<sub>7</sub>).



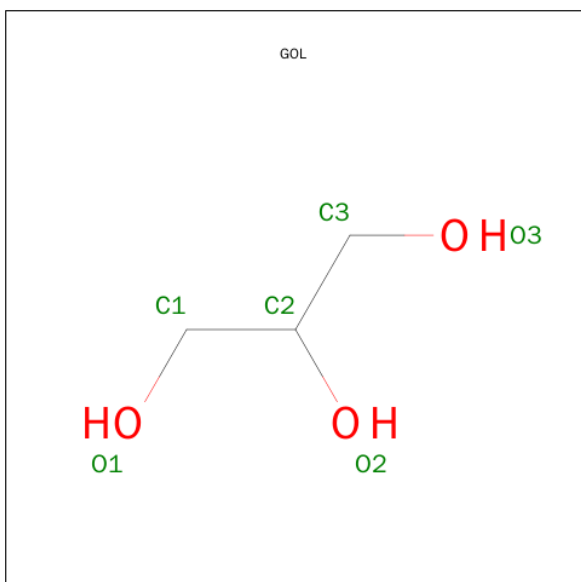
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			13	6	7		

- Molecule 4 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>3</sub>).



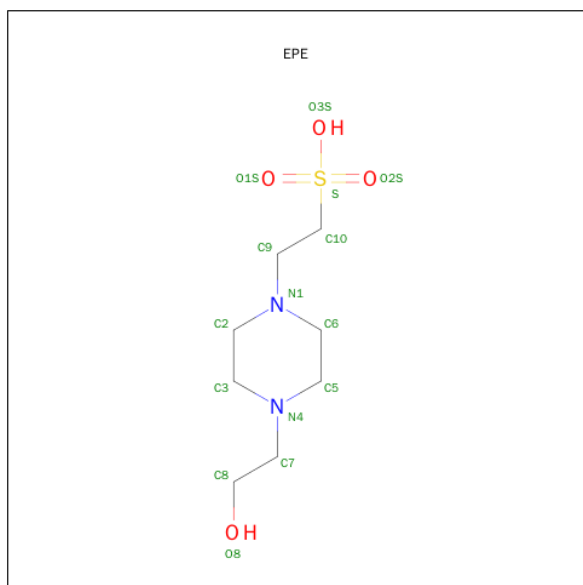
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			7	4	3		
4	A	1	Total	C	O	0	0
			7	4	3		
4	A	1	Total	C	O	0	0
			7	4	3		
4	A	1	Total	C	O	0	0
			7	4	3		
4	A	1	Total	C	O	0	0
			7	4	3		

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0
5	A	1	Total C O 6 3 3	0	0

- Molecule 6 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: C<sub>8</sub>H<sub>18</sub>N<sub>2</sub>O<sub>4</sub>S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total C N O S 15 8 2 4 1	0	0
6	A	1	Total C N O S 15 8 2 4 1	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	53	Total O 53 53	0	0
7	B	2	Total O 2 2	0	0





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	49.84Å 88.44Å 161.33Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	34.00 – 2.60 34.16 – 2.60	Depositor EDS
% Data completeness (in resolution range)	99.3 (34.00-2.60) 99.4 (34.16-2.60)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.68 (at 2.61Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, $R_{free}$	0.230 , 0.280 0.235 , 0.278	Depositor DCC
$R_{free}$ test set	1209 reflections (5.66%)	DCC
Wilson B-factor (Å <sup>2</sup> )	56.3	Xtriage
Anisotropy	0.551	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 54.7	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	3 of 22562 reflections (0.013%)	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	3923	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.67% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: EPE, GOL, PEG, CIT

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	0.49	0/3797	0.63	0/5159
2	B	0.41	0/59	0.56	0/77
All	All	0.48	0/3856	0.63	0/5236

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	3714	0	3787	28	0
2	B	58	0	63	0	0
3	A	13	0	5	0	0
4	A	35	0	50	1	0
5	A	18	0	24	2	0
6	A	30	0	36	0	0
7	A	53	0	0	1	0
7	B	2	0	0	0	0
All	All	3923	0	3965	28	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 (28) 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:234:VAL:O	1:A:238:ILE:HG13	1.83	0.78
1:A:1052:ARG:HH11	1:A:1052:ARG:HG2	1.56	0.68
1:A:371:LEU:HA	7:A:1335:HOH:O	1.94	0.66
1:A:1009:ILE:HD11	1:A:1164:GLY:HA2	1.77	0.66
1:A:1052:ARG:NH1	1:A:1052:ARG:HG2	2.08	0.64
1:A:1126:TRP:HB3	1:A:1154:ARG:HA	1.87	0.56
1:A:1050:ILE:HD11	1:A:1054:THR:CG2	2.36	0.56
1:A:1052:ARG:HH12	1:A:1054:THR:HA	1.71	0.55
1:A:1052:ARG:HH11	1:A:1052:ARG:CG	2.20	0.54
1:A:207:THR:HG21	1:A:237:VAL:HG21	1.90	0.54
1:A:160:VAL:HG12	1:A:253:ILE:HD11	1.91	0.52
1:A:127:ASN:HD21	1:A:136:ALA:H	1.60	0.50
1:A:263:LYS:NZ	4:A:1203:PEG:O2	2.46	0.47
1:A:1052:ARG:NH1	1:A:1054:THR:HA	2.30	0.47
1:A:1144:ASN:HB2	5:A:1209:GOL:H11	1.99	0.45
1:A:1081:ASN:HD22	1:A:1108:GLU:CG	2.31	0.44
1:A:262:ASN:OD1	1:A:311:ARG:NH2	2.51	0.44
1:A:137:PHE:O	1:A:141:GLY:HA3	2.17	0.44
1:A:203:PRO:O	1:A:207:THR:HG23	2.18	0.44
1:A:375:ASN:O	1:A:378:GLN:N	2.51	0.43
1:A:1054:THR:O	1:A:1055:ASN:CB	2.67	0.43
1:A:322:LEU:N	1:A:323:PRO:HD2	2.34	0.43
1:A:145:TYR:OH	1:A:149:ARG:NH1	2.52	0.42
1:A:253:ILE:HG22	1:A:314:VAL:HG22	2.02	0.42
1:A:172:CYS:SG	1:A:263:LYS:HD3	2.59	0.42
1:A:143:ARG:NH1	1:A:205:LEU:O	2.51	0.42
1:A:264:LEU:O	1:A:268:VAL:HG23	2.21	0.41
1:A:1145:ARG:H	5:A:1209:GOL:H11	1.86	0.41

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	465/541 (86%)	443 (95%)	22 (5%)	0	100	100
2	B	4/6 (67%)	4 (100%)	0	0	100	100
All	All	469/547 (86%)	447 (95%)	22 (5%)	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	400/459 (87%)	399 (100%)	1 (0%)	94	99
2	B	6/6 (100%)	6 (100%)	0	100	100
All	All	406/465 (87%)	405 (100%)	1 (0%)	95	99

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	376	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

11 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	CIT	A	1201	-	3,12,12	0.95	0	3,17,17	1.73	1 (33%)
4	PEG	A	1202	-	6,6,6	0.59	0	5,5,5	0.16	0
4	PEG	A	1203	-	6,6,6	0.70	0	5,5,5	0.55	0
4	PEG	A	1204	-	6,6,6	0.31	0	5,5,5	0.38	0
4	PEG	A	1205	-	6,6,6	0.57	0	5,5,5	0.13	0
4	PEG	A	1206	-	6,6,6	0.67	0	5,5,5	0.57	0
5	GOL	A	1207	-	5,5,5	0.32	0	5,5,5	0.35	0
5	GOL	A	1208	-	5,5,5	0.32	0	5,5,5	0.21	0
5	GOL	A	1209	-	5,5,5	0.16	0	5,5,5	0.58	0
6	EPE	A	1210	-	15,15,15	1.99	1 (6%)	19,20,20	1.71	3 (15%)
6	EPE	A	1211	-	15,15,15	1.81	1 (6%)	19,20,20	1.74	3 (15%)

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
3	CIT	A	1201	-	-	0/6/16/16	0/0/0/0
4	PEG	A	1202	-	-	0/4/4/4	0/0/0/0
4	PEG	A	1203	-	-	0/4/4/4	0/0/0/0
4	PEG	A	1204	-	-	0/4/4/4	0/0/0/0
4	PEG	A	1205	-	-	0/4/4/4	0/0/0/0
4	PEG	A	1206	-	-	0/4/4/4	0/0/0/0
5	GOL	A	1207	-	-	0/4/4/4	0/0/0/0
5	GOL	A	1208	-	-	0/4/4/4	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	GOL	A	1209	-	-	0/4/4/4	0/0/0/0
6	EPE	A	1210	-	-	0/9/19/19	0/1/1/1
6	EPE	A	1211	-	-	0/9/19/19	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	1210	EPE	C10-S	-7.25	1.66	1.77
6	A	1211	EPE	C10-S	-6.54	1.67	1.77

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1201	CIT	C3-C4-C5	-2.73	110.69	114.95
6	A	1210	EPE	O2S-S-O1S	-2.09	108.07	113.96
6	A	1211	EPE	O2S-S-O1S	-2.06	108.13	113.96
6	A	1210	EPE	O3S-S-C10	2.41	110.00	104.99
6	A	1211	EPE	O3S-S-C10	2.81	110.83	104.99
6	A	1210	EPE	O2S-S-C10	5.81	110.98	106.87
6	A	1211	EPE	O1S-S-C10	6.21	111.26	106.87

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1203	PEG	1	0
5	A	1209	GOL	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	471/541 (87%)	0.18	25 (5%) 30 23	29, 63, 88, 108	53 (11%)
2	B	6/6 (100%)	0.45	1 (16%) 2 1	49, 52, 61, 69	0
All	All	477/547 (87%)	0.18	26 (5%) 29 21	29, 63, 88, 108	53 (11%)

All (26) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	381	LEU	4.8
1	A	175	PHE	4.6
1	A	95	LEU	4.3
1	A	94	SER	4.1
1	A	382	SER	3.6
1	A	216	ASP	3.6
1	A	378	GLN	3.4
1	A	180	LEU	3.2
1	A	218	THR	3.1
1	A	50	GLY	3.1
1	A	177	ALA	3.0
1	A	49	ALA	2.9
1	A	217	GLY	2.9
1	A	220	PRO	2.7
1	A	1047	ASP	2.6
1	A	179	THR	2.5
1	A	182	SER	2.5
1	A	215	ALA	2.4
2	B	13	LEU	2.3
1	A	61	ILE	2.2
1	A	130	TRP	2.2
1	A	92	LYS	2.1
1	A	376	PHE	2.1
1	A	219	HIS	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	1052	ARG	2.0
1	A	380	PHE	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	PEG	A	1204	7/7	0.83	0.29	4.02	86,88,94,95	0
4	PEG	A	1206	7/7	0.73	0.30	3.33	63,64,72,74	0
4	PEG	A	1205	7/7	0.85	0.23	2.49	73,76,78,78	0
5	GOL	A	1207	6/6	0.79	0.22	1.55	69,73,74,77	0
4	PEG	A	1203	7/7	0.88	0.22	1.00	48,56,60,64	0
6	EPE	A	1211	15/15	0.80	0.26	0.75	82,92,113,121	0
3	CIT	A	1201	13/13	0.92	0.20	0.70	51,56,62,62	0
5	GOL	A	1209	6/6	0.92	0.20	0.26	56,57,64,68	0
6	EPE	A	1210	15/15	0.94	0.16	-0.19	47,49,62,62	0
4	PEG	A	1202	7/7	0.81	0.17	-0.46	58,60,62,63	0
5	GOL	A	1208	6/6	0.90	0.33	-	64,70,73,74	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.