wwPDB X-ray Structure Validation Summary Report (i)

Apr 27, 2024 – 04:16 pm BST

PDB ID	:	2XJO
Title	:	Crystal structure of Streptococcus suis Dpr with nickel
Authors	:	Haikarainen, T.; Thanassoulas, A.; Stavros, P.; Nounesis, G.; Haataja, S.; Papageorgiou, A.C.
Deposited on	:	2010-07-06
Resolution	:	2.10 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

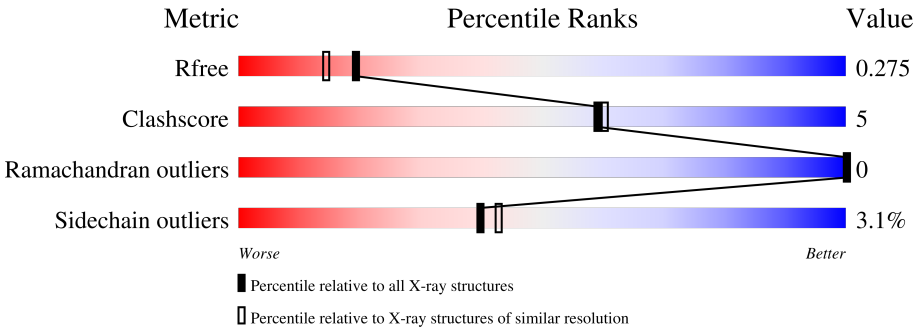
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.10 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)

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 >=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 <=5%

Mol	Chain	Length	Quality of chain
1	A	165	80% 11% 9%
1	B	165	79% 12% • 8%
1	C	165	77% 14% 9%
1	D	165	82% 10% 8%
1	E	165	81% 10% 8%
1	F	165	85% 7% 8%
1	G	165	79% 11% 10%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	H	165	81% 10% 8%
1	I	165	73% 16% • 8%
1	J	165	81% 10% • 8%
1	K	165	81% 10% • 8%
1	L	165	80% 11% • 8%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 15560 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 DNA PROTECTION DURING STARVATION PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	150	Total	C	N	O	S	0	2	0
			1218	775	200	236	7			
1	B	151	Total	C	N	O	S	0	3	0
			1223	779	201	236	7			
1	C	150	Total	C	N	O	S	0	0	0
			1197	762	197	232	6			
1	D	152	Total	C	N	O	S	0	5	0
			1259	798	208	246	7			
1	E	151	Total	C	N	O	S	0	3	0
			1221	777	200	238	6			
1	F	151	Total	C	N	O	S	0	3	0
			1235	785	202	242	6			
1	G	149	Total	C	N	O	S	0	4	0
			1227	778	201	242	6			
1	H	151	Total	C	N	O	S	0	2	0
			1216	775	200	235	6			
1	I	151	Total	C	N	O	S	0	3	0
			1212	772	202	232	6			
1	J	151	Total	C	N	O	S	0	3	0
			1223	778	200	239	6			
1	K	152	Total	C	N	O	S	0	6	0
			1252	793	206	247	6			
1	L	152	Total	C	N	O	S	0	6	0
			1243	789	203	245	6			

- Molecule 2 is NICKEL (II) ION (three-letter code: NI) (formula: Ni).

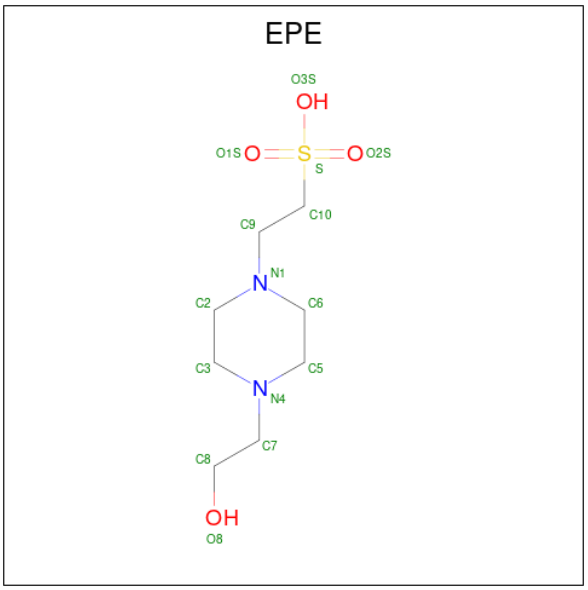
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Ni	0	0
			1	1		
2	B	1	Total	Ni	0	0
			1	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	C	1	Total Ni 1 1	0	0
2	D	1	Total Ni 1 1	0	0
2	E	1	Total Ni 1 1	0	0
2	F	1	Total Ni 1 1	0	0
2	G	1	Total Ni 1 1	0	0
2	H	1	Total Ni 1 1	0	0
2	I	1	Total Ni 1 1	0	0
2	J	1	Total Ni 1 1	0	0
2	K	1	Total Ni 1 1	0	0
2	L	1	Total Ni 1 1	0	0

- Molecule 3 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: $C_8H_{18}N_2O_4S$).



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

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	G	1	Total C N O S 15 8 2 4 1	0	0
3	K	1	Total C N O S 15 8 2 4 1	0	0
3	L	1	Total C N O S 15 8 2 4 1	0	0

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total Ca 1 1	0	0
4	D	1	Total Ca 1 1	0	0
4	F	2	Total Ca 2 2	0	0
4	I	1	Total Ca 1 1	0	0
4	J	1	Total Ca 1 1	0	0
4	K	1	Total Ca 1 1	0	0

- Molecule 5 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total Cl 1 1	0	0
5	D	1	Total Cl 1 1	0	0
5	E	1	Total Cl 1 1	0	0
5	G	1	Total Cl 1 1	0	0
5	I	1	Total Cl 1 1	0	0
5	J	1	Total Cl 1 1	0	0

- Molecule 6 is water.

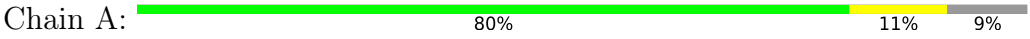
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	58	Total O 58 58	0	0
6	B	92	Total O 92 92	0	0
6	C	42	Total O 42 42	0	0
6	D	96	Total O 96 96	0	0
6	E	38	Total O 38 38	0	0
6	F	79	Total O 79 79	0	0
6	G	46	Total O 46 46	0	0
6	H	48	Total O 48 48	0	0
6	I	54	Total O 54 54	0	0
6	J	61	Total O 61 61	0	0
6	K	75	Total O 75 75	0	0
6	L	60	Total O 60 60	0	0

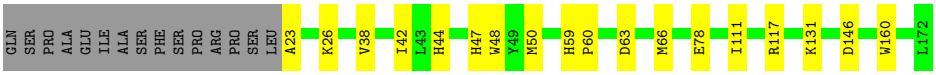
3 Residue-property plots (i)

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. 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. 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 PROTECTION DURING STARVATION PROTEIN

Chain A:





• Molecule 1: DNA PROTECTION DURING STARVATION PROTEIN

Chain B:

• Molecule 1: DNA PROTECTION DURING STARVATION PROTEIN

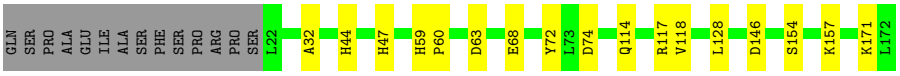
Chain C:

• Molecule 1: DNA PROTECTION DURING STARVATION PROTEIN

Chain D:

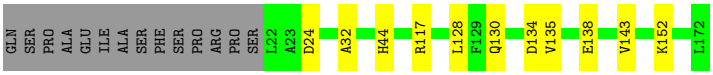
• Molecule 1: DNA PROTECTION DURING STARVATION PROTEIN

Chain E:



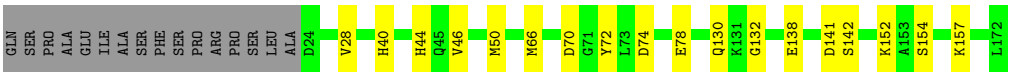
• Molecule 1: DNA PROTECTION DURING STARVATION PROTEIN

Chain F:



• Molecule 1: DNA PROTECTION DURING STARVATION PROTEIN

Chain G:

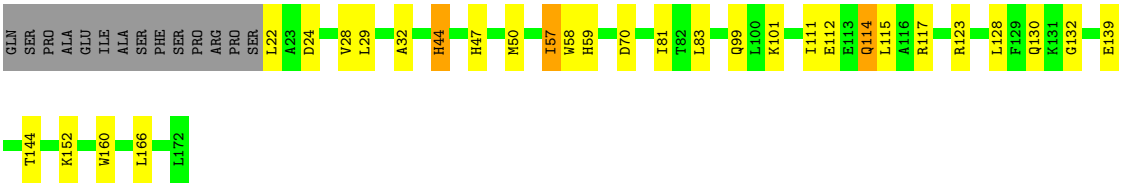


• Molecule 1: DNA PROTECTION DURING STARVATION PROTEIN

Chain H:

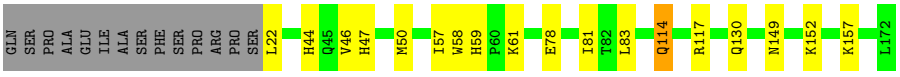
• Molecule 1: DNA PROTECTION DURING STARVATION PROTEIN

Chain I:



• Molecule 1: DNA PROTECTION DURING STARVATION PROTEIN

Chain J:



• Molecule 1: DNA PROTECTION DURING STARVATION PROTEIN

Chain K:

• Molecule 1: DNA PROTECTION DURING STARVATION PROTEIN

Chain L:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	104.86Å 137.46Å 141.73Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.10 20.00 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.1 (20.00-2.10) 99.1 (20.00-2.10)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.05 (at 2.09Å)	Xtriage
Refinement program	REFMAC 5.4.0078	Depositor
R, R_{free}	0.176 , 0.222 0.242 , 0.275	Depositor DCC
R_{free} test set	5940 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	41.8	Xtriage
Anisotropy	0.151	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 30.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.018 for -h,l,k	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	15560	wwPDB-VP
Average B, all atoms (Å^2)	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 11.31% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: CA, EPE, CL, NI

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	0.72	0/1243	0.66	0/1677
1	B	0.78	0/1252	0.72	2/1690 (0.1%)
1	C	0.69	0/1222	0.68	1/1651 (0.1%)
1	D	0.81	0/1284	0.68	0/1732
1	E	0.64	0/1250	0.67	0/1689
1	F	0.83	0/1260	0.71	0/1701
1	G	0.68	1/1252 (0.1%)	0.70	1/1690 (0.1%)
1	H	0.69	0/1245	0.69	1/1681 (0.1%)
1	I	0.70	0/1245	0.69	0/1682
1	J	0.75	0/1252	0.64	0/1692
1	K	0.76	0/1285	0.69	0/1735
1	L	0.71	0/1280	0.72	0/1728
All	All	0.73	1/15070 (0.0%)	0.69	5/20348 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	G	157	LYS	CE-NZ	5.52	1.62	1.49

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	141	ASP	CB-CG-OD1	5.74	123.47	118.30
1	B	73	LEU	CA-CB-CG	5.59	128.17	115.30
1	H	141	ASP	CB-CG-OD1	5.32	123.09	118.30
1	B	141	ASP	CB-CG-OD1	5.24	123.01	118.30
1	G	141	ASP	CB-CG-OD1	5.18	122.97	118.30

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	1218	0	1176	8	0
1	B	1223	0	1182	14	0
1	C	1197	0	1152	11	0
1	D	1259	0	1210	12	0
1	E	1221	0	1170	10	0
1	F	1235	0	1187	9	0
1	G	1227	0	1173	8	0
1	H	1216	0	1176	9	0
1	I	1212	0	1169	18	0
1	J	1223	0	1169	12	0
1	K	1252	0	1195	20	0
1	L	1243	0	1192	14	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
2	G	1	0	0	0	0
2	H	1	0	0	0	0
2	I	1	0	0	0	0
2	J	1	0	0	0	0
2	K	1	0	0	0	0
2	L	1	0	0	0	0
3	B	15	0	17	1	0
3	G	15	0	17	1	0
3	K	15	0	17	1	0
3	L	15	0	17	0	0
4	B	1	0	0	0	0
4	D	1	0	0	0	0
4	F	2	0	0	0	0
4	I	1	0	0	0	0
4	J	1	0	0	0	0
4	K	1	0	0	0	0
5	B	1	0	0	0	0
5	D	1	0	0	0	0
5	E	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	G	1	0	0	0	0
5	I	1	0	0	0	0
5	J	1	0	0	0	0
6	A	58	0	0	0	0
6	B	92	0	0	1	0
6	C	42	0	0	0	0
6	D	96	0	0	4	0
6	E	38	0	0	3	0
6	F	79	0	0	1	0
6	G	46	0	0	0	0
6	H	48	0	0	0	0
6	I	54	0	0	4	0
6	J	61	0	0	0	0
6	K	75	0	0	0	0
6	L	60	0	0	0	0
All	All	15560	0	14219	136	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 5.

The worst 5 of 136 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:K:117[A]:ARG:HH11	1:K:117[A]:ARG:CG	1.55	1.16
1:K:117[A]:ARG:HG3	1:K:117[A]:ARG:NH1	1.52	1.04
1:C:130:GLN:NE2	1:C:152:LYS:HE3	1.82	0.95
1:C:130:GLN:HE21	1:C:152:LYS:HE3	1.32	0.94
1:L:21:SER:N	1:L:22:LEU:HA	1.86	0.88

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	150/165 (91%)	149 (99%)	1 (1%)	0	100	100
1	B	152/165 (92%)	151 (99%)	1 (1%)	0	100	100
1	C	148/165 (90%)	147 (99%)	1 (1%)	0	100	100
1	D	155/165 (94%)	154 (99%)	1 (1%)	0	100	100
1	E	152/165 (92%)	151 (99%)	1 (1%)	0	100	100
1	F	152/165 (92%)	151 (99%)	1 (1%)	0	100	100
1	G	151/165 (92%)	149 (99%)	2 (1%)	0	100	100
1	H	151/165 (92%)	149 (99%)	2 (1%)	0	100	100
1	I	152/165 (92%)	150 (99%)	2 (1%)	0	100	100
1	J	152/165 (92%)	151 (99%)	1 (1%)	0	100	100
1	K	156/165 (94%)	155 (99%)	1 (1%)	0	100	100
1	L	156/165 (94%)	154 (99%)	2 (1%)	0	100	100
All	All	1827/1980 (92%)	1811 (99%)	16 (1%)	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	130/141 (92%)	125 (96%)	5 (4%)	33	34
1	B	131/141 (93%)	127 (97%)	4 (3%)	40	43
1	C	127/141 (90%)	121 (95%)	6 (5%)	26	25
1	D	135/141 (96%)	133 (98%)	2 (2%)	65	71
1	E	130/141 (92%)	127 (98%)	3 (2%)	50	55
1	F	132/141 (94%)	131 (99%)	1 (1%)	81	86
1	G	132/141 (94%)	130 (98%)	2 (2%)	65	71
1	H	130/141 (92%)	126 (97%)	4 (3%)	40	43
1	I	129/141 (92%)	121 (94%)	8 (6%)	18	15

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	J	130/141 (92%)	127 (98%)	3 (2%)	50 55
1	K	135/141 (96%)	131 (97%)	4 (3%)	41 44
1	L	135/141 (96%)	127 (94%)	8 (6%)	19 17
All	All	1576/1692 (93%)	1526 (97%)	50 (3%)	40 41

5 of 50 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	I	57	ILE
1	J	44	HIS
1	L	160	TRP
1	I	70	ASP
1	I	114	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 58 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	163	GLN
1	L	130	GLN
1	H	130	GLN
1	L	114	GLN
1	K	114	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)

Of 29 ligands modelled in this entry, 25 are monoatomic - leaving 4 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	$\#\|Z\| > 2$
3	EPE	G	1173	-	15,15,15	0.73	0	18,20,20	1.89	4 (22%)
3	EPE	K	1173	-	15,15,15	1.03	1 (6%)	18,20,20	1.89	5 (27%)
3	EPE	L	1173	-	15,15,15	0.83	1 (6%)	18,20,20	2.29	6 (33%)
3	EPE	B	1173	-	15,15,15	0.92	1 (6%)	18,20,20	2.20	5 (27%)

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	EPE	G	1173	-	-	6/9/19/19	0/1/1/1
3	EPE	K	1173	-	-	3/9/19/19	0/1/1/1
3	EPE	L	1173	-	-	4/9/19/19	0/1/1/1
3	EPE	B	1173	-	-	6/9/19/19	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	K	1173	EPE	C10-S	3.39	1.82	1.77
3	L	1173	EPE	C10-S	2.73	1.81	1.77
3	B	1173	EPE	C10-S	2.60	1.81	1.77

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	1173	EPE	O1S-S-C10	6.26	114.45	106.92
3	L	1173	EPE	O1S-S-C10	5.17	113.14	106.92
3	G	1173	EPE	O2S-S-C10	4.62	112.48	106.92

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	L	1173	EPE	C5-N4-C3	4.25	118.40	108.83
3	K	1173	EPE	C5-N4-C3	4.02	117.88	108.83

There are no chirality outliers.

5 of 19 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	1173	EPE	C10-C9-N1-C2
3	G	1173	EPE	C10-C9-N1-C2
3	K	1173	EPE	C10-C9-N1-C2
3	L	1173	EPE	N4-C7-C8-O8
3	L	1173	EPE	C9-C10-S-O2S

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	G	1173	EPE	1	0
3	K	1173	EPE	1	0
3	B	1173	EPE	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

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.