



# wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 7, 2022 – 06:03 PM EDT

PDB ID : 2AV9  
Title : Crystal Structure of the PA5185 protein from Pseudomonas Aeruginosa Strain PAO1.  
Authors : Chruszcz, M.; Wang, S.; Cymborowski, M.; Kudritska, M.; Evdokimova, E.; Edwards, A.; Savchenko, A.; Joachimiak, A.; Minor, W.; Midwest Center for Structural Genomics (MCSG)  
Deposited on : 2005-08-29  
Resolution : 2.40 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 : 2.27  
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) : 2.27

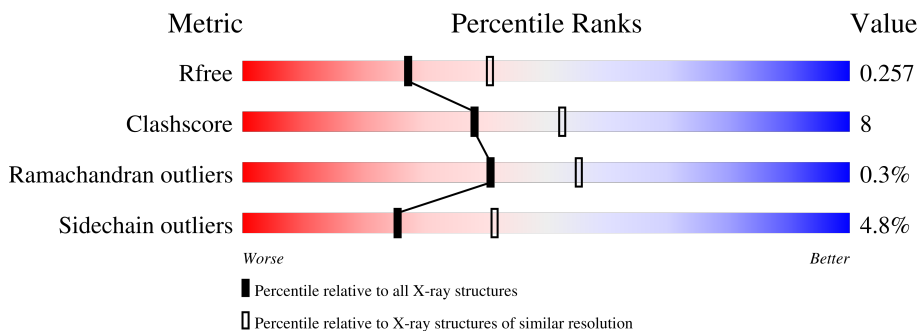
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)

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  $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	147	
1	B	147	
1	C	147	
1	D	147	
1	E	147	
1	F	147	
1	G	147	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	H	147	 80% 13% • 5%
1	I	147	 76% 16% • 5%
1	J	147	 73% 18% • 6%
1	K	147	 71% 20% • 7%
1	L	147	 81% 12% 7%

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
2	SO4	F	150	-	-	X	-

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 12653 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 Thioesterase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	142	1101	707	195	196	2	1	0	1	0
1	B	139	1070	687	183	196	3	1	0	1	0
1	C	139	1050	675	180	192	2	1	0	0	0
1	D	142	1047	672	178	194	2	1	0	0	0
1	E	137	1033	668	172	190	2	1	0	0	0
1	F	139	1072	687	184	198	2	1	0	0	0
1	G	136	989	643	163	180	2	1	0	0	0
1	H	139	1041	668	185	185	2	1	0	0	0
1	I	139	1058	680	181	194	2	1	0	0	0
1	J	138	1057	679	185	189	3	1	0	1	0
1	K	137	997	646	167	181	2	1	0	0	0
1	L	137	1004	650	165	186	2	1	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

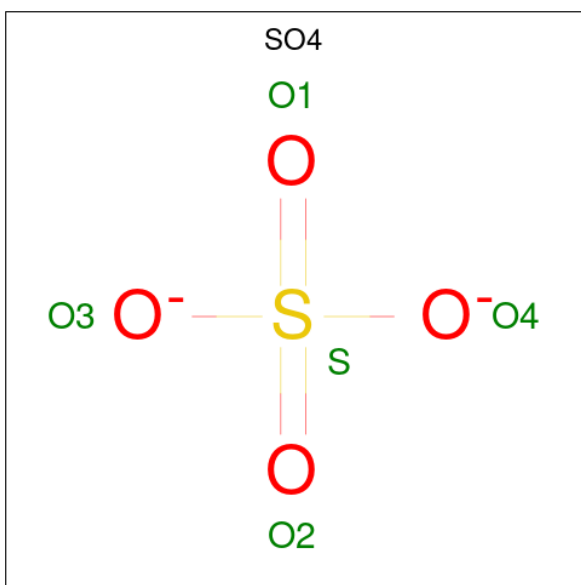
Chain	Residue	Modelled	Actual	Comment	Reference
A	84	MSE	MET	modified residue	UNP Q9HU04
B	84	MSE	MET	modified residue	UNP Q9HU04
C	84	MSE	MET	modified residue	UNP Q9HU04
D	84	MSE	MET	modified residue	UNP Q9HU04
E	84	MSE	MET	modified residue	UNP Q9HU04

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	84	MSE	MET	modified residue	UNP Q9HU04
G	84	MSE	MET	modified residue	UNP Q9HU04
H	84	MSE	MET	modified residue	UNP Q9HU04
I	84	MSE	MET	modified residue	UNP Q9HU04
J	84	MSE	MET	modified residue	UNP Q9HU04
K	84	MSE	MET	modified residue	UNP Q9HU04
L	84	MSE	MET	modified residue	UNP Q9HU04

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



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	C	1	Total O S 5 4 1	0	0
2	D	1	Total O S 5 4 1	0	0
2	E	1	Total O S 5 4 1	0	0
2	F	1	Total O S 5 4 1	0	0
2	F	1	Total O S 5 4 1	0	0

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	F	1	Total 5	O 4	S 1	0	0
2	G	1	Total 5	O 4	S 1	0	0
2	H	1	Total 5	O 4	S 1	0	0
2	I	1	Total 5	O 4	S 1	0	0
2	J	1	Total 5	O 4	S 1	0	0
2	K	1	Total 5	O 4	S 1	0	0
2	L	1	Total 5	O 4	S 1	0	0


- Molecule 3 is water.

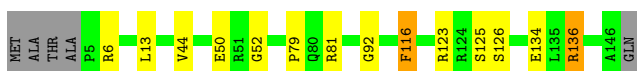
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	11	Total 11	O 11	0	0
3	B	10	Total 10	O 10	0	0
3	C	3	Total 3	O 3	0	0
3	D	8	Total 8	O 8	0	0
3	E	6	Total 6	O 6	0	0
3	F	4	Total 4	O 4	0	0
3	G	2	Total 2	O 2	0	0
3	I	1	Total 1	O 1	0	0
3	J	12	Total 12	O 12	0	0
3	K	1	Total 1	O 1	0	0
3	L	1	Total 1	O 1	0	0

### 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.

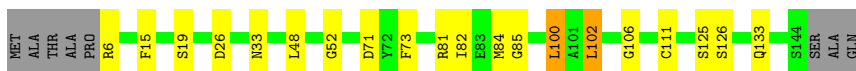
- Molecule 1: Thioesterase

Chain A:  87% 8% ..



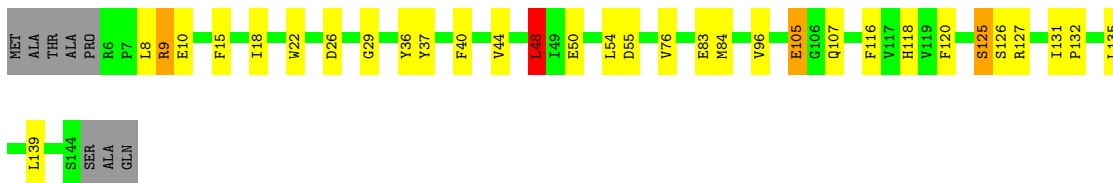
- Molecule 1: Thioesterase

Chain B:  81% 12% • 5%




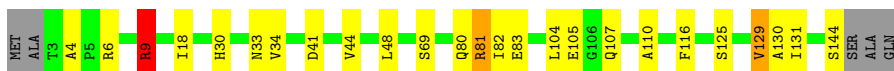
- Molecule 1: Thioesterase

Chain C:  73% 19% .. 5%




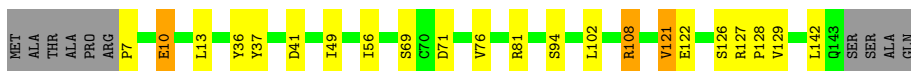
- Molecule 1: Thioesterase

Chain D:  80% 15% ...




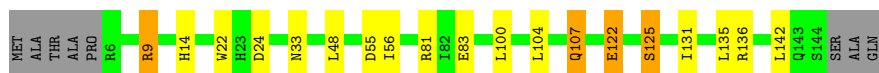
- Molecule 1: Thioesterase

Chain E:  78% 13% • 7%



- Molecule 1: Thioesterase

Chain F:  82% 10% 5%




- Molecule 1: Thioesterase

Chain G:  73% 18% 7%



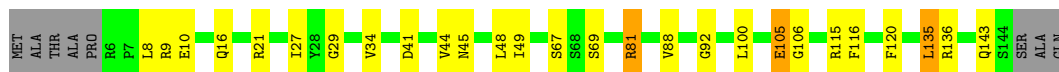
- Molecule 1: Thioesterase

Chain H:  80% 13% 5%



- Molecule 1: Thioesterase

Chain I:  76% 16% 5%



- Molecule 1: Thioesterase

Chain J:  73% 18% 6%




- Molecule 1: Thioesterase

Chain K:  71% 20% 7%



- Molecule 1: Thioesterase

Chain L:  81% 12% 7%





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	241.66Å 64.38Å 117.47Å 90.00° 105.40° 90.00°	Depositor
Resolution (Å)	36.94 – 2.40 36.92 – 2.40	Depositor EDS
% Data completeness (in resolution range)	100.0 (36.94-2.40) 100.0 (36.92-2.40)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.64 (at 2.39Å)	Xtrriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.204 , 0.253 0.260 , 0.257	Depositor DCC
$R_{free}$ test set	3448 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	41.6	Xtrriage
Anisotropy	0.161	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 46.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	12653	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	45.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.28% 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	0.43	0/1131	1.06	4/1538 (0.3%)
1	B	0.44	0/1095	1.11	5/1489 (0.3%)
1	C	0.33	0/1075	0.94	2/1464 (0.1%)
1	D	0.38	0/1073	1.02	3/1465 (0.2%)
1	E	0.37	0/1058	0.97	4/1443 (0.3%)
1	F	0.34	0/1097	1.08	5/1493 (0.3%)
1	G	0.31	0/1013	0.84	4/1385 (0.3%)
1	H	0.31	0/1067	0.82	1/1455 (0.1%)
1	I	0.33	0/1083	0.89	1/1475 (0.1%)
1	J	0.38	0/1082	1.06	4/1472 (0.3%)
1	K	0.70	2/1022 (0.2%)	0.89	1/1396 (0.1%)
1	L	0.32	0/1029	0.84	0/1405
All	All	0.40	2/12825 (0.0%)	0.97	34/17480 (0.2%)

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	I	0	1
1	J	0	1
All	All	0	2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	K	108	ARG	CG-CD	15.33	1.90	1.51
1	K	108	ARG	CB-CG	12.09	1.85	1.52

The worst 5 of 34 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	9	ARG	NE-CZ-NH2	-12.13	114.23	120.30
1	J	81	ARG	NE-CZ-NH1	-10.60	115.00	120.30
1	F	9	ARG	NE-CZ-NH1	9.79	125.19	120.30
1	B	102	LEU	CB-CG-CD1	-7.92	97.53	111.00
1	K	108	ARG	CB-CG-CD	-7.60	91.85	111.60

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	I	115	ARG	Peptide
1	J	106	GLY	Peptide

## 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	1101	0	1056	6	0
1	B	1070	0	1005	15	0
1	C	1050	0	976	24	0
1	D	1047	0	941	16	0
1	E	1033	0	961	15	0
1	F	1072	0	1014	13	0
1	G	989	0	900	26	0
1	H	1041	0	963	17	0
1	I	1058	0	988	20	0
1	J	1057	0	992	23	0
1	K	997	0	903	30	0
1	L	1004	0	908	10	0
2	A	5	0	0	0	0
2	B	10	0	0	1	0
2	C	5	0	0	0	0
2	D	5	0	0	0	0
2	E	5	0	0	0	0
2	F	15	0	0	4	0
2	G	5	0	0	0	0
2	H	5	0	0	0	0
2	I	5	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	J	5	0	0	1	0
2	K	5	0	0	1	0
2	L	5	0	0	0	0
3	A	11	0	0	0	0
3	B	10	0	0	0	0
3	C	3	0	0	0	0
3	D	8	0	0	1	0
3	E	6	0	0	0	0
3	F	4	0	0	0	0
3	G	2	0	0	0	0
3	I	1	0	0	0	0
3	J	12	0	0	0	0
3	K	1	0	0	0	0
3	L	1	0	0	0	0
All	All	12653	0	11607	184	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 8.

The worst 5 of 184 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:108:ARG:CB	1:K:108:ARG:CG	1.85	1.54
1:K:108:ARG:CG	1:K:108:ARG:CD	1.90	1.49
1:G:80:GLN:NE2	1:K:106:GLY:HA3	1.71	1.05
1:B:84:MSE:HE1	1:B:100:LEU:HB3	1.40	0.99
1:D:81:ARG:HD2	1:I:81:ARG:HD2	1.45	0.97

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	141/147 (96%)	139 (99%)	2 (1%)	0	100	100
1	B	138/147 (94%)	134 (97%)	4 (3%)	0	100	100
1	C	137/147 (93%)	134 (98%)	3 (2%)	0	100	100
1	D	140/147 (95%)	135 (96%)	3 (2%)	2 (1%)	11	15
1	E	135/147 (92%)	130 (96%)	5 (4%)	0	100	100
1	F	137/147 (93%)	132 (96%)	5 (4%)	0	100	100
1	G	134/147 (91%)	132 (98%)	2 (2%)	0	100	100
1	H	137/147 (93%)	134 (98%)	3 (2%)	0	100	100
1	I	137/147 (93%)	131 (96%)	5 (4%)	1 (1%)	22	32
1	J	137/147 (93%)	135 (98%)	2 (2%)	0	100	100
1	K	135/147 (92%)	129 (96%)	4 (3%)	2 (2%)	10	14
1	L	135/147 (92%)	131 (97%)	4 (3%)	0	100	100
All	All	1643/1764 (93%)	1596 (97%)	42 (3%)	5 (0%)	41	55

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	125	SER
1	K	107	GLN
1	K	124	ARG
1	D	4	ALA
1	I	105	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	110/120 (92%)	108 (98%)	2 (2%)	59	76
1	B	105/120 (88%)	102 (97%)	3 (3%)	42	62
1	C	101/120 (84%)	95 (94%)	6 (6%)	19	32
1	D	96/120 (80%)	88 (92%)	8 (8%)	11	17

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	E	100/120 (83%)	95 (95%)	5 (5%)	24	40
1	F	108/120 (90%)	103 (95%)	5 (5%)	27	43
1	G	91/120 (76%)	88 (97%)	3 (3%)	38	57
1	H	98/120 (82%)	93 (95%)	5 (5%)	24	39
1	I	103/120 (86%)	95 (92%)	8 (8%)	12	19
1	J	102/120 (85%)	99 (97%)	3 (3%)	42	62
1	K	91/120 (76%)	87 (96%)	4 (4%)	28	45
1	L	93/120 (78%)	88 (95%)	5 (5%)	22	36
All	All	1198/1440 (83%)	1141 (95%)	57 (5%)	25	41

5 of 57 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	125	SER
1	L	126	SER
1	H	100	LEU
1	L	118	HIS
1	K	16	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	G	57	GLN
1	L	97	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

15 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
2	SO4	B	149	-	4,4,4	0.15	0	6,6,6	0.87	0
2	SO4	L	148	-	4,4,4	0.13	0	6,6,6	0.49	0
2	SO4	K	148	-	4,4,4	0.14	0	6,6,6	0.66	0
2	SO4	E	148	-	4,4,4	0.13	0	6,6,6	0.86	0
2	SO4	F	149	-	4,4,4	0.14	0	6,6,6	0.51	0
2	SO4	H	148	-	4,4,4	0.16	0	6,6,6	0.58	0
2	SO4	I	148	-	4,4,4	0.18	0	6,6,6	0.78	0
2	SO4	B	148	-	4,4,4	0.14	0	6,6,6	0.80	0
2	SO4	A	148	-	4,4,4	0.17	0	6,6,6	0.55	0
2	SO4	G	148	-	4,4,4	0.14	0	6,6,6	0.53	0
2	SO4	F	148	-	4,4,4	0.15	0	6,6,6	0.84	0
2	SO4	J	148	-	4,4,4	0.16	0	6,6,6	1.08	0
2	SO4	F	150	-	4,4,4	0.12	0	6,6,6	0.78	0
2	SO4	C	148	-	4,4,4	0.13	0	6,6,6	0.79	0
2	SO4	D	148	-	4,4,4	0.16	0	6,6,6	0.64	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.

5 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	K	148	SO4	1	0
2	F	149	SO4	1	0
2	B	148	SO4	1	0
2	J	148	SO4	1	0
2	F	150	SO4	3	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.