wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 07:07 PM GMT

PDB ID	:	1E5F
Title	:	METHIONINE GAMMA-LYASE (MGL) FROM TRICHOMONAS VAGINALIS
Authors	:	Goodall, G.; Mottram, J.C.; Coombs, G.H.; Lapthorn, A.J.
Deposited on	:	2000-07-25
Resolution	:	2.18 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

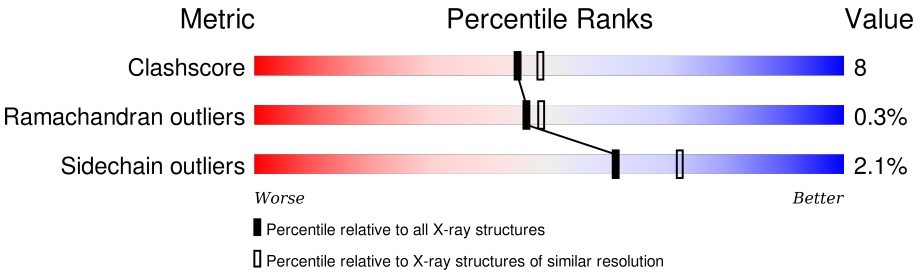
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 2.18 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))
Clashscore	102246	5965 (2.20-2.16)
Ramachandran outliers	100387	5863 (2.20-2.16)
Sidechain outliers	100360	5864 (2.20-2.16)

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 >=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 <=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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	404	66% 25% 6%
1	B	404	64% 26% 6%

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
3	SO4	A	407	-	-	X	-
3	SO4	B	411	-	-	X	-

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 6649 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.

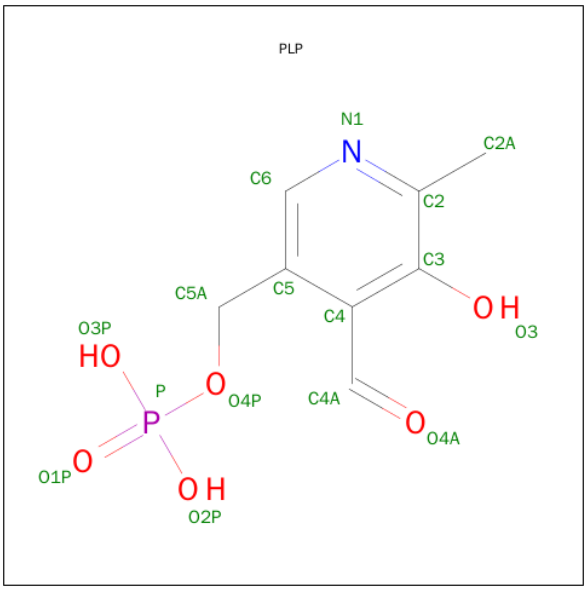
- Molecule 1 is a protein called METHIONINE GAMMA-LYASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	393	Total	C	N	O	S	0	5	0
			2969	1882	501	562	24			
1	B	393	Total	C	N	O	S	0	4	0
			2969	1877	505	563	24			

There are 4 discrepancies between the modelled and reference sequences:

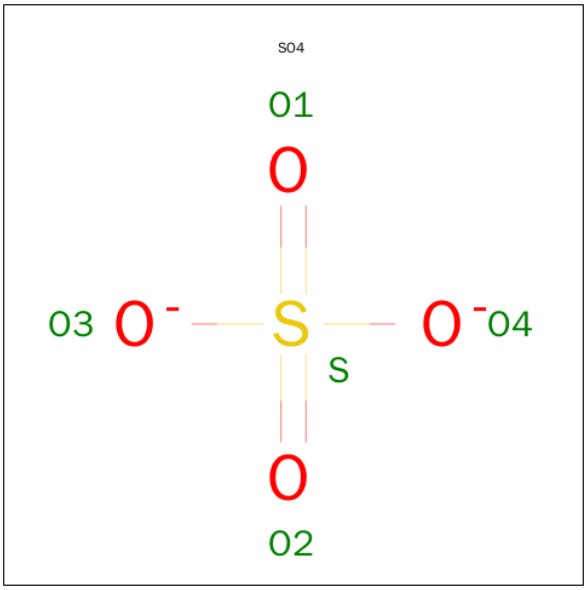
Chain	Residue	Modelled	Actual	Comment	Reference
A	2	ALA	SER	CLONING ARTIFACT	UNP O15564
A	308	TYR	SER	CLONING ARTIFACT	UNP O15564
B	2	ALA	SER	CLONING ARTIFACT	UNP O15564
B	308	TYR	SER	CLONING ARTIFACT	UNP O15564

- Molecule 2 is PYRIDOXAL-5'-PHOSPHATE (three-letter code: PLP) (formula: $C_8H_{10}NO_6P$).



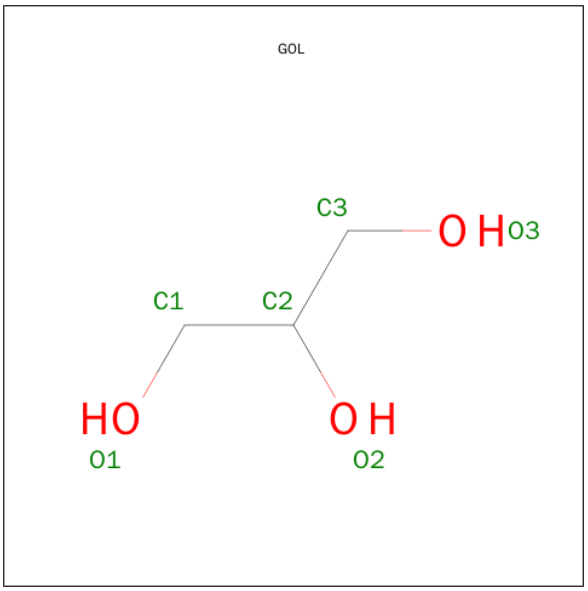
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C N O P 15 8 1 5 1	0	0
2	B	1	Total C N O P 15 8 1 5 1	0	0

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



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

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	329	Total O 329 329	0	0
5	B	316	Total O 316 316	0	0

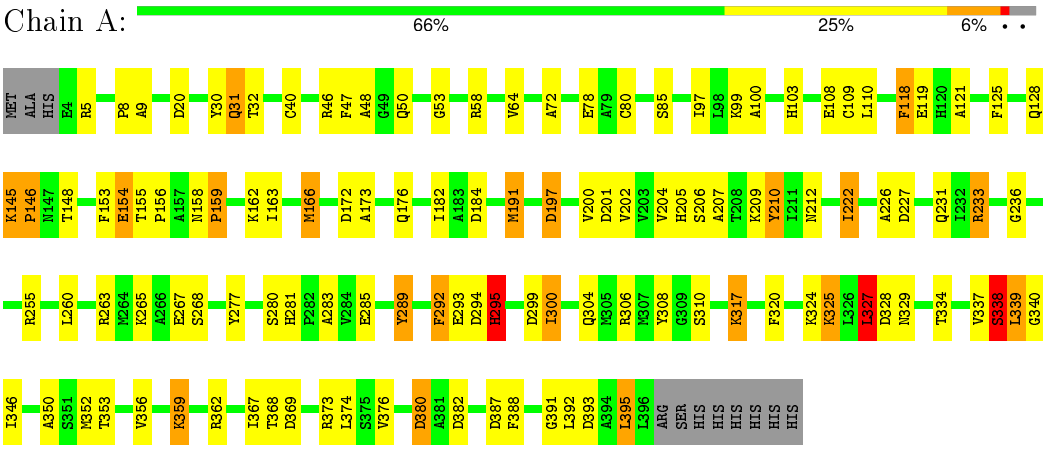
3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

Note EDS was not executed.

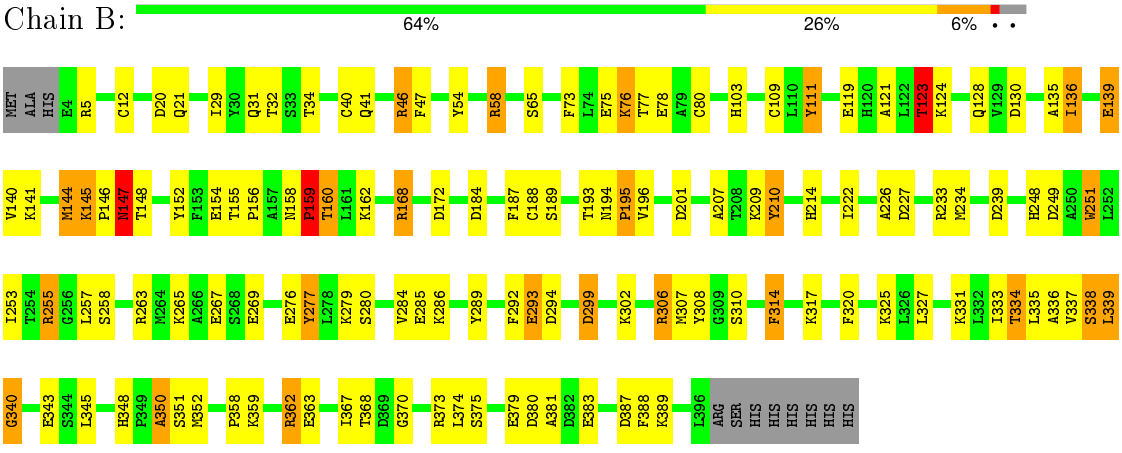
- Molecule 1: METHIONINE GAMMA-LYASE

Chain A:



- Molecule 1: METHIONINE GAMMA-LYASE

Chain B:



4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 31 1 2	Depositor
Cell constants a, b, c, α, β, γ	88.26Å 88.26Å 217.85Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	25.00 – 2.18	Depositor
% Data completeness (in resolution range)	98.6 (25.00-2.18)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
Refinement program	REFMAC	Depositor
R, R_{free}	0.162 , 0.212	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	6649	wwPDB-VP
Average B, all atoms (Å^2)	31.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, SO4, PLP

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	1.29	10/3050 (0.3%)	2.20	100/4133 (2.4%)
1	B	1.31	10/3046 (0.3%)	2.22	104/4131 (2.5%)
All	All	1.30	20/6096 (0.3%)	2.21	204/8264 (2.5%)

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	24
1	B	0	13
All	All	0	37

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	280	SER	CB-OG	7.79	1.52	1.42
1	A	343	GLU	CB-CG	-6.88	1.39	1.52
1	B	12[A]	CYS	CB-SG	-6.80	1.70	1.82
1	B	12[B]	CYS	CB-SG	-6.80	1.70	1.82
1	B	338	SER	CB-OG	6.64	1.50	1.42

The worst 5 of 204 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	46	ARG	CD-NE-CZ	40.20	179.87	123.60
1	A	46	ARG	NE-CZ-NH2	-20.89	109.86	120.30
1	A	46	ARG	NE-CZ-NH1	20.67	130.64	120.30
1	B	306[A]	ARG	NE-CZ-NH1	20.02	130.31	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	306[B]	ARG	NE-CZ-NH1	20.02	130.31	120.30

There are no chirality outliers.

5 of 37 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	31	GLN	Mainchain
1	A	47	PHE	Mainchain
1	A	50	GLN	Mainchain
1	A	53	GLY	Mainchain
1	A	9	ALA	Mainchain

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	2969	0	2961	45	0
1	B	2969	0	2942	49	0
2	A	15	0	6	5	0
2	B	15	0	7	5	0
3	A	15	0	0	2	0
3	B	15	0	0	3	0
4	A	6	0	8	3	0
5	A	329	0	0	5	0
5	B	316	0	0	9	0
All	All	6649	0	5924	94	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 94 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:209:LYS:NZ	2:A:405:PLP:C4A	2.45	0.79
1:A:209:LYS:HZ1	2:A:405:PLP:C4A	1.97	0.77

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:281:HIS:HD2	1:A:283:ALA:H	1.32	0.73
1:B:146:PRO:HA	5:B:2184:HOH:O	1.89	0.73
1:B:103:HIS:HB3	1:B:148:THR:HA	1.74	0.70

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	396/404 (98%)	382 (96%)	13 (3%)	1 (0%)	46	48
1	B	395/404 (98%)	386 (98%)	8 (2%)	1 (0%)	46	48
All	All	791/808 (98%)	768 (97%)	21 (3%)	2 (0%)	46	48

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	338	SER
1	B	338	SER

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	320/334 (96%)	314 (98%)	6 (2%)	65	75

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	B	319/334 (96%)	311 (98%)	8 (2%)	55 65
All	All	639/668 (96%)	625 (98%)	14 (2%)	61 70

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	123	THR
1	B	124	LYS
1	B	314	PHE
1	A	359[B]	LYS
1	B	159	PRO

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	281	HIS
1	B	147	ASN
1	B	41	GLN
1	A	128	GLN
1	A	295	HIS

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 carbohydrates in this entry.

5.6 Ligand geometry (i)

9 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	PLP	A	405	-	15,15,16	1.83	4 (26%)	21,22,23	3.19	8 (38%)
3	SO4	A	407	-	4,4,4	1.42	1 (25%)	6,6,6	1.22	1 (16%)
3	SO4	A	408	-	4,4,4	0.98	0	6,6,6	0.59	0
3	SO4	A	409	-	4,4,4	1.09	0	6,6,6	0.74	0
4	GOL	A	412	-	5,5,5	1.00	0	5,5,5	1.49	1 (20%)
2	PLP	B	405	-	15,15,16	1.56	2 (13%)	21,22,23	2.13	8 (38%)
3	SO4	B	406	-	4,4,4	1.44	1 (25%)	6,6,6	0.90	0
3	SO4	B	410	-	4,4,4	1.30	0	6,6,6	0.84	0
3	SO4	B	411	-	4,4,4	1.26	0	6,6,6	1.49	1 (16%)

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	PLP	A	405	-	-	0/6/6/8	0/1/1/1
3	SO4	A	407	-	-	0/0/0/0	0/0/0/0
3	SO4	A	408	-	-	0/0/0/0	0/0/0/0
3	SO4	A	409	-	-	0/0/0/0	0/0/0/0
4	GOL	A	412	-	-	0/4/4/4	0/0/0/0
2	PLP	B	405	-	-	0/6/6/8	0/1/1/1
3	SO4	B	406	-	-	0/0/0/0	0/0/0/0
3	SO4	B	410	-	-	0/0/0/0	0/0/0/0
3	SO4	B	411	-	-	0/0/0/0	0/0/0/0

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	405	PLP	C5-C4	-4.51	1.35	1.40
2	B	405	PLP	C5-C4	-4.25	1.35	1.40
2	A	405	PLP	P-O2P	-2.12	1.47	1.54
3	A	407	SO4	O2-S	2.12	1.54	1.47
3	B	406	SO4	O3-S	2.47	1.56	1.47

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	405	PLP	C3-C2-N1	-7.90	109.70	120.61
2	A	405	PLP	C4A-C4-C3	-5.02	111.27	120.36
2	B	405	PLP	C3-C2-N1	-4.16	114.87	120.61
2	A	405	PLP	C2A-C2-C3	-3.83	116.42	121.04
2	B	405	PLP	O4P-C5A-C5	-3.25	103.62	108.99

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 16 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	405	PLP	5	0
3	A	407	SO4	2	0
4	A	412	GOL	3	0
2	B	405	PLP	5	0
3	B	406	SO4	1	0
3	B	411	SO4	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

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates

EDS was not executed - this section will therefore be empty.

6.4 Ligands

EDS was not executed - this section will therefore be empty.

6.5 Other polymers

EDS was not executed - this section will therefore be empty.