



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 14, 2018 – 11:24 am GMT

PDB ID : 2V0C  
Title : LEUCYL-TRNA SYNTHETASE FROM THERMUS THERMOPHILUS COMPLEXED WITH A SULPHAMOYL ANALOGUE OF LEUCYL-ADENYLATE In the synthetic site and an adduct of AMP with 5-Fluoro-1,3-dihydro-1-hydroxy-2,1-benzoxaborole (AN2690) in the editing site  
Authors : Rock, F.; Mao, W.; Yaremchuk, A.; Tukalo, M.; Crepin, T.; Zhou, H.; Zhang, Y.; Hernandez, V.; Akama, T.; Baker, S.; Plattner, J.; Shapiro, L.; Martinis, S.A.; Benkovic, S.J.; Cusack, S.; Alley, M.R.K.  
Deposited on : 2007-05-14  
Resolution : 1.85 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	:	1.7.3 (157068), CSD as539be (2018)
Xtriage (Phenix)	:	1.13
EDS	:	trunk31020
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac	:	5.8.0158
CCP4	:	7.0 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk31020

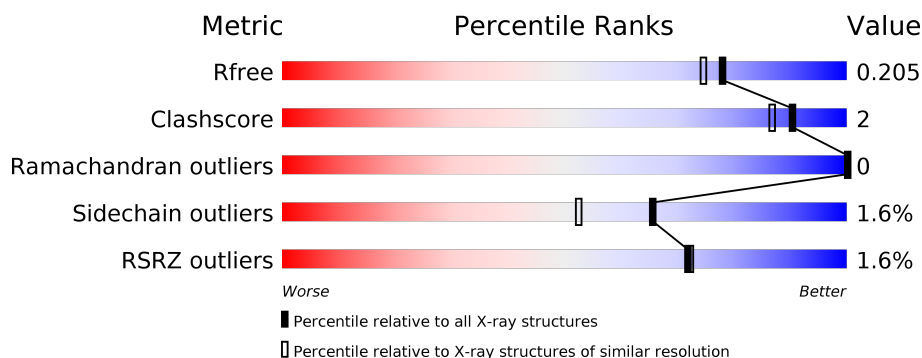
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.85 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))
$R_{free}$	111664	2111 (1.86-1.86)
Clashscore	122126	2258 (1.86-1.86)
Ramachandran outliers	120053	2234 (1.86-1.86)
Sidechain outliers	120020	2234 (1.86-1.86)
RSRZ outliers	108989	2075 (1.86-1.86)

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	878	<div> <div></div> <div>87%</div> <div>5%</div> <div>8%</div> </div>

## 2 Entry composition [i](#)

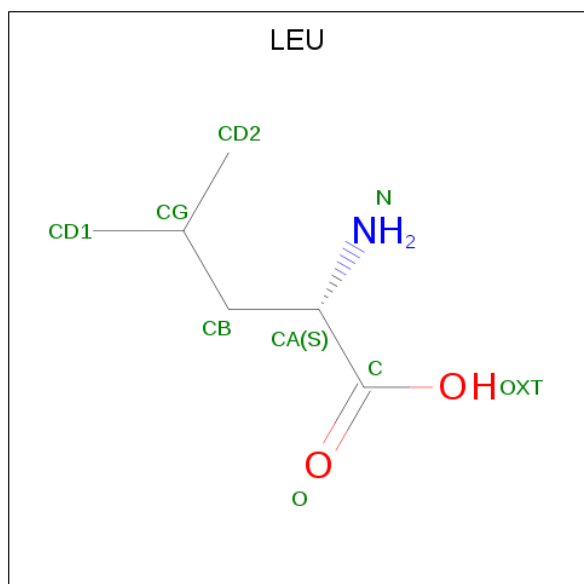
There are 7 unique types of molecules in this entry. The entry contains 7383 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 AMINOACYL-TRNA SYNTHETASE.

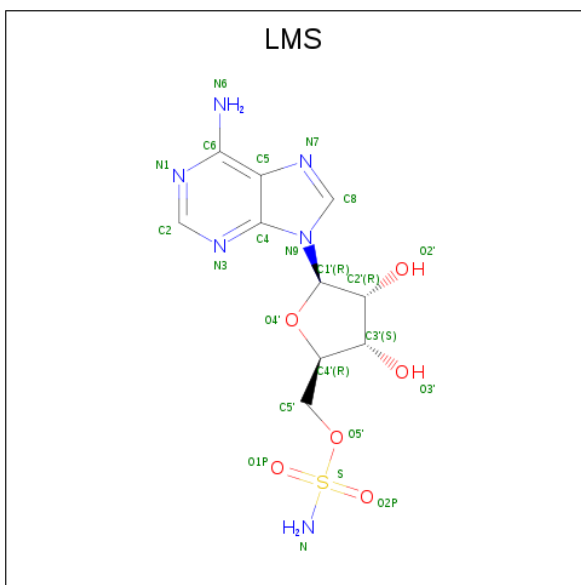
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	810	6608	4266	1125	1186	31	0	0	0

- Molecule 2 is LEUCINE (three-letter code: LEU) (formula:  $C_6H_{13}NO_2$ ).



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

- Molecule 3 is [(2R,3S,4R,5R)-5-(6-AMINO-9H-PURIN-9-YL)-3,4-DIHYDROXYTE TRAHYDRO-2-FURANYL]METHYL SULFAMATE (three-letter code: LMS) (formula:  $C_{10}H_{14}N_6O_6S$ ).

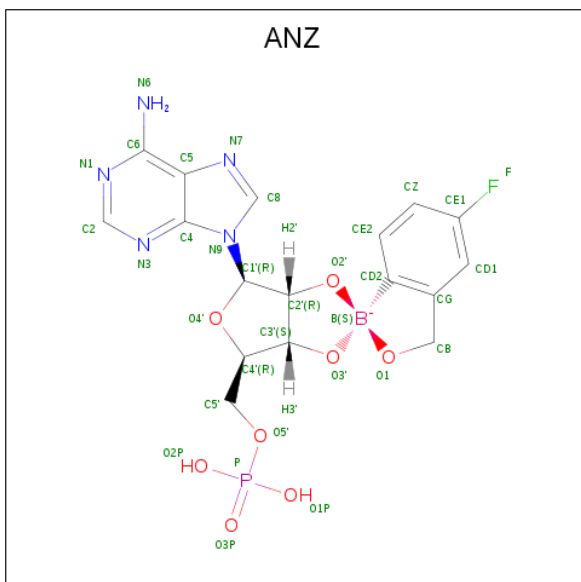


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	S	0	0
			23	10	6	6	1		

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	2	Total	Zn	0	0
			2	2		

- Molecule 5 is [(6-AMINO-9H-PURIN-9-YL)-[5-FLUORO-1,3-DIHYDRO-1-HYDROXY-2,1-BENZOXABOROLE]-4'YL]METHYL DIHYDROGEN PHOSPHATE (three-letter code: ANZ) (formula: C<sub>17</sub>H<sub>17</sub>BFN<sub>5</sub>O<sub>8</sub>P).



Mol	Chain	Residues	Atoms							ZeroOcc	AltConf
5	A	1	Total	B	C	F	N	O	P	0	0
			33	1	17	1	5	8	1		

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

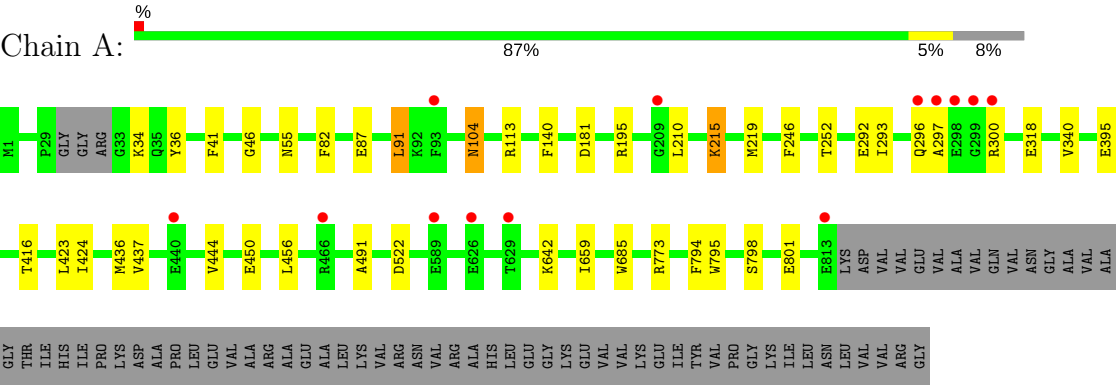
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	684	Total	O	0	0
			684	684		

### 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 the various outlier classes 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.

- Molecule 1: AMINOACYL-TRNA SYNTHETASE



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	101.83Å 154.23Å 174.95Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	87.37 – 1.85 48.08 – 1.85	Depositor EDS
% Data completeness (in resolution range)	99.5 (87.37-1.85) 99.5 (48.08-1.85)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.15 (at 1.86Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.180 , 0.202 0.185 , 0.205	Depositor DCC
$R_{free}$ test set	3477 reflections (2.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	24.7	Xtriage
Anisotropy	0.334	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 48.2	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	0.96	EDS
Total number of atoms	7383	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.85% 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: ZN, LMS, SO4, ANZ

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	0.69	1/6805 (0.0%)	0.73	4/9237 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	798	SER	CB-OG	7.11	1.51	1.42

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	773	ARG	NE-CZ-NH2	-5.53	117.54	120.30
1	A	773	ARG	NE-CZ-NH1	5.49	123.04	120.30
1	A	195	ARG	NE-CZ-NH2	-5.25	117.68	120.30
1	A	522	ASP	CB-CG-OD2	5.12	122.91	118.30

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	6608	0	6474	25	0
2	A	8	0	10	0	0
3	A	23	0	14	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	2	0	0	0	0
5	A	33	0	15	0	0
6	A	25	0	0	0	0
7	A	684	0	0	7	0
All	All	7383	0	6513	25	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 2.

The worst 5 of 25 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:801:GLU:OE1	7:A:2664:HOH:O	2.03	0.77
1:A:395:GLU:HG3	7:A:2142:HOH:O	1.85	0.76
1:A:181:ASP:OD1	7:A:2175:HOH:O	2.12	0.68
1:A:297:ALA:HB3	1:A:300:ARG:HD2	1.76	0.67
1:A:318:GLU:OE1	7:A:2324:HOH:O	2.15	0.64

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	806/878 (92%)	796 (99%)	10 (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	686/740 (93%)	675 (98%)	11 (2%)	65	52

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	82	PHE
1	A	91	LEU
1	A	215	LYS
1	A	55	ASN
1	A	210	LEU

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

### 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](#)

Of 10 ligands modelled in this entry, 2 are monoatomic - leaving 8 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
2	LEU	A	1501	3	6,7,8	1.84	1 (16%)	6,8,10	0.50	0
3	LMS	A	1814	2	22,25,25	4.66	5 (22%)	23,38,38	4.52	8 (34%)
5	ANZ	A	1817	-	29,38,38	0.99	1 (3%)	31,60,60	2.13	8 (25%)
6	SO4	A	1818	-	4,4,4	0.08	0	6,6,6	0.53	0
6	SO4	A	1819	-	4,4,4	0.11	0	6,6,6	0.38	0
6	SO4	A	1820	-	4,4,4	0.14	0	6,6,6	0.14	0
6	SO4	A	1821	-	4,4,4	0.13	0	6,6,6	0.21	0
6	SO4	A	1822	-	4,4,4	0.19	0	6,6,6	0.18	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	LEU	A	1501	3	-	0/4/6/8	0/0/0/0
3	LMS	A	1814	2	-	0/6/26/26	0/3/3/3
5	ANZ	A	1817	-	-	0/6/48/48	0/6/6/6
6	SO4	A	1818	-	-	0/0/0/0	0/0/0/0
6	SO4	A	1819	-	-	0/0/0/0	0/0/0/0
6	SO4	A	1820	-	-	0/0/0/0	0/0/0/0
6	SO4	A	1821	-	-	0/0/0/0	0/0/0/0
6	SO4	A	1822	-	-	0/0/0/0	0/0/0/0

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1814	LMS	S-N	-11.61	1.46	1.58
3	A	1814	LMS	C8-N9	-2.10	1.34	1.36
5	A	1817	ANZ	CD1-CE1	2.30	1.41	1.37
3	A	1814	LMS	C5-C4	2.94	1.47	1.40
2	A	1501	LEU	CA-C	4.30	1.55	1.50

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1814	LMS	O2P-S-O1P	-18.43	102.37	119.97
5	A	1817	ANZ	N3-C2-N1	-9.21	120.98	128.86
3	A	1814	LMS	N3-C2-N1	-6.34	123.43	128.86
5	A	1817	ANZ	CZ-CE1-CD1	-2.66	119.78	123.29
5	A	1817	ANZ	CB-CG-CD1	-2.15	125.34	129.36

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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	810/878 (92%)	-0.19	13 (1%) 72 72	14, 24, 40, 60	0

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	296	GLN	5.0
1	A	297	ALA	4.7
1	A	299	GLY	3.4
1	A	298	GLU	3.2
1	A	209	GLY	3.1

### 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. 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(Å <sup>2</sup> )	Q<0.9
5	ANZ	A	1817	33/33	0.91	0.09	24,26,48,54	0
6	SO4	A	1821	5/5	0.94	0.12	83,83,83,84	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
6	SO4	A	1820	5/5	0.94	0.23	85,85,86,86	0
6	SO4	A	1818	5/5	0.96	0.14	46,47,50,51	0
4	ZN	A	1815	1/1	0.96	0.05	35,35,35,35	0
6	SO4	A	1822	5/5	0.98	0.18	53,54,55,55	0
3	LMS	A	1814	23/23	0.98	0.12	14,17,18,19	0
6	SO4	A	1819	5/5	0.99	0.09	34,34,37,38	0
4	ZN	A	1816	1/1	0.99	0.06	27,27,27,27	0
2	LEU	A	1501	8/9	0.99	0.15	13,14,15,15	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.