



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 06:38 AM GMT

PDB ID : 2XRZ  
Title : X-RAY STRUCTURE OF ARCHAEAL CLASS II CPD PHOTOLYASE  
FROM METHANOSARCINA MAZEI IN COMPLEX WITH INTACT CPD-  
LESION  
Authors : Kiontke, S.; Geisselbrecht, Y.; Pokorny, R.; Carell, T.; Batschauer, A.; Essen,  
L.O.  
Deposited on : 2010-09-24  
Resolution : 2.20 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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
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) : trunk26865

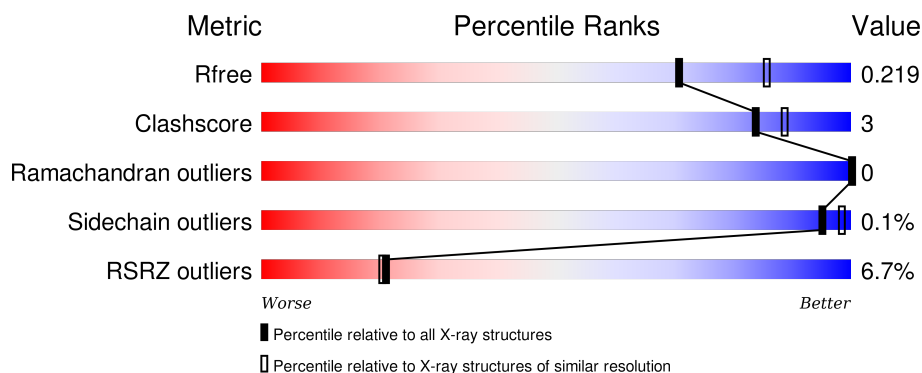
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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	3774 (2.20-2.20)
Clashscore	102246	4477 (2.20-2.20)
Ramachandran outliers	100387	4404 (2.20-2.20)
Sidechain outliers	100360	4405 (2.20-2.20)
RSRZ outliers	91569	3781 (2.20-2.20)

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	482	<div> <div>2%</div> <div>89%</div> <div>7%</div> </div>
1	B	482	<div> <div>6%</div> <div>85%</div> <div>5%</div> <div>10%</div> </div>
2	C	14	<div> <div>36%</div> <div>57%</div> <div>21%</div> <div>21%</div> </div>
2	E	14	<div> <div>29%</div> <div>36%</div> <div>29%</div> <div>7%</div> <div>29%</div> </div>
3	D	14	<div> <div>43%</div> <div>71%</div> <div>29%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	F	14	

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
6	GOL	A	1465	-	-	-	X
6	GOL	A	1466	-	-	-	X
7	ACT	A	1467	-	-	-	X
7	ACT	A	1470	-	-	-	X
7	ACT	B	1463	-	-	X	X

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 8978 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 DEOXYRIBODIPYRIMIDINE PHOTOLYASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	450	Total	C	N	O	S	0	4	0
			3651	2350	611	676	14			
1	B	434	Total	C	N	O	S	0	6	0
			3522	2266	592	649	15			

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-17	MET	-	EXPRESSION TAG	UNP Q8PYK9
A	-16	GLY	-	EXPRESSION TAG	UNP Q8PYK9
A	-15	SER	-	EXPRESSION TAG	UNP Q8PYK9
A	-14	SER	-	EXPRESSION TAG	UNP Q8PYK9
A	-13	HIS	-	EXPRESSION TAG	UNP Q8PYK9
A	-12	HIS	-	EXPRESSION TAG	UNP Q8PYK9
A	-11	HIS	-	EXPRESSION TAG	UNP Q8PYK9
A	-10	HIS	-	EXPRESSION TAG	UNP Q8PYK9
A	-9	HIS	-	EXPRESSION TAG	UNP Q8PYK9
A	-8	HIS	-	EXPRESSION TAG	UNP Q8PYK9
A	-7	SER	-	EXPRESSION TAG	UNP Q8PYK9
A	-6	SER	-	EXPRESSION TAG	UNP Q8PYK9
A	-5	GLY	-	EXPRESSION TAG	UNP Q8PYK9
A	-4	LEU	-	EXPRESSION TAG	UNP Q8PYK9
A	-3	VAL	-	EXPRESSION TAG	UNP Q8PYK9
A	-2	PRO	-	EXPRESSION TAG	UNP Q8PYK9
A	-1	ARG	-	EXPRESSION TAG	UNP Q8PYK9
A	0	GLY	-	EXPRESSION TAG	UNP Q8PYK9
A	1	SER	-	EXPRESSION TAG	UNP Q8PYK9
A	2	HIS	-	EXPRESSION TAG	UNP Q8PYK9
A	377	THR	MET	ENGINEERED MUTATION	UNP Q8PYK9
B	-17	MET	-	EXPRESSION TAG	UNP Q8PYK9
B	-16	GLY	-	EXPRESSION TAG	UNP Q8PYK9
B	-15	SER	-	EXPRESSION TAG	UNP Q8PYK9
B	-14	SER	-	EXPRESSION TAG	UNP Q8PYK9

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	-13	HIS	-	EXPRESSION TAG	UNP Q8PYK9
B	-12	HIS	-	EXPRESSION TAG	UNP Q8PYK9
B	-11	HIS	-	EXPRESSION TAG	UNP Q8PYK9
B	-10	HIS	-	EXPRESSION TAG	UNP Q8PYK9
B	-9	HIS	-	EXPRESSION TAG	UNP Q8PYK9
B	-8	HIS	-	EXPRESSION TAG	UNP Q8PYK9
B	-7	SER	-	EXPRESSION TAG	UNP Q8PYK9
B	-6	SER	-	EXPRESSION TAG	UNP Q8PYK9
B	-5	GLY	-	EXPRESSION TAG	UNP Q8PYK9
B	-4	LEU	-	EXPRESSION TAG	UNP Q8PYK9
B	-3	VAL	-	EXPRESSION TAG	UNP Q8PYK9
B	-2	PRO	-	EXPRESSION TAG	UNP Q8PYK9
B	-1	ARG	-	EXPRESSION TAG	UNP Q8PYK9
B	0	GLY	-	EXPRESSION TAG	UNP Q8PYK9
B	1	SER	-	EXPRESSION TAG	UNP Q8PYK9
B	2	HIS	-	EXPRESSION TAG	UNP Q8PYK9
B	377	THR	MET	ENGINEERED MUTATION	UNP Q8PYK9

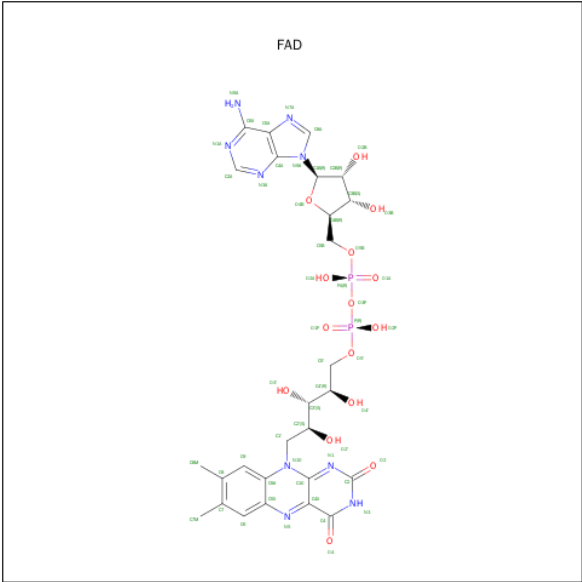
- Molecule 2 is a DNA chain called CPD-COMPRISING OLIGONUCLEOTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	11	Total	C	N	O	P	0	0	0
			242	116	44	71	11			
2	E	10	Total	C	N	O	P	0	0	0
			223	107	41	65	10			

- Molecule 3 is a DNA chain called COUNTERSTRAND-OLIGONUCLEOTIDE.

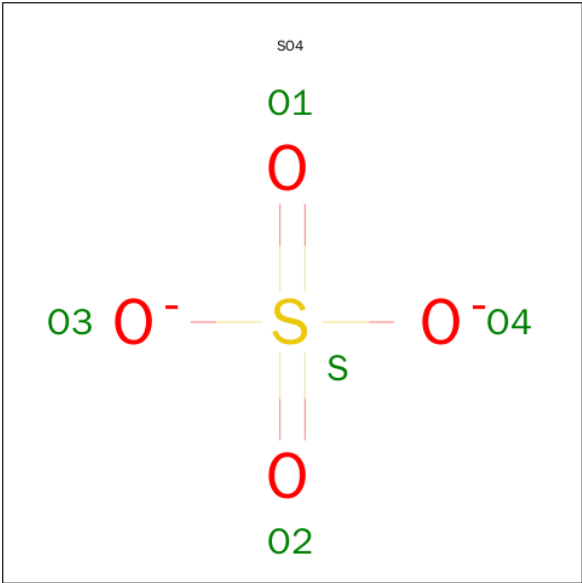
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	14	Total	C	N	O	P	0	0	0
			286	136	56	81	13			
3	F	13	Total	C	N	O	P	0	0	0
			266	126	54	74	12			

- Molecule 4 is FLAVIN-ADENINE DINUCLEOTIDE (three-letter code: FAD) (formula:  $C_{27}H_{33}N_9O_{15}P_2$ ).



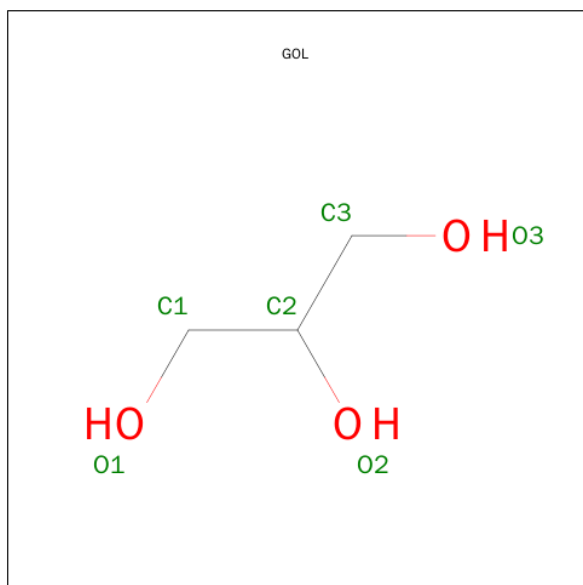
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total 53	C 27	N 9	O 15	P 2	0	0
4	B	1	Total 53	C 27	N 9	O 15	P 2	0	0

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



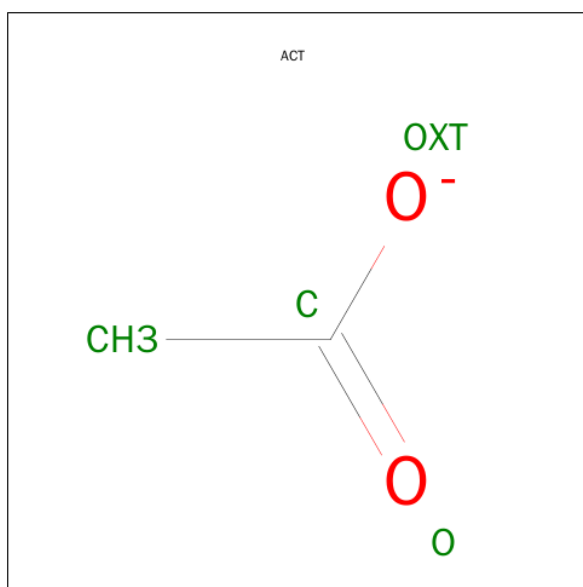
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	O	S	0	0
			5	4	1		

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



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

- Molecule 7 is ACETATE ION (three-letter code: ACT) (formula:  $C_2H_3O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	C	O	0	0
			4	2	2		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total 4	C 2	O 2	0	0
7	A	1	Total 4	C 2	O 2	0	0
7	A	1	Total 4	C 2	O 2	0	0
7	B	1	Total 4	C 2	O 2	0	0
7	B	1	Total 4	C 2	O 2	0	0

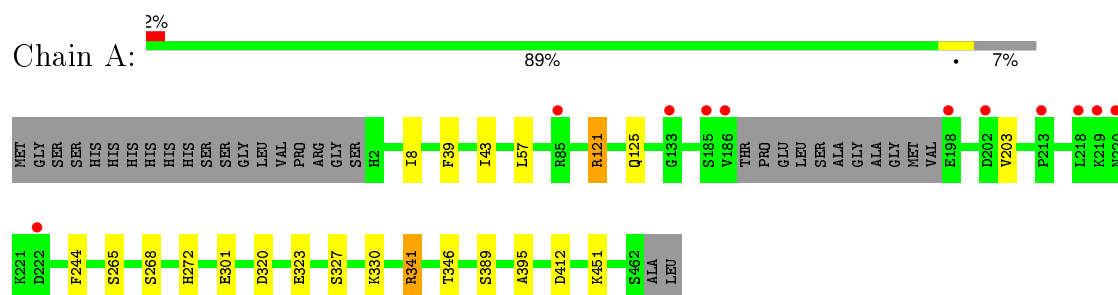
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	333	Total 333	O 333	0	0
8	B	265	Total 265	O 265	0	0
8	C	15	Total 15	O 15	0	0
8	D	9	Total 9	O 9	0	0
8	E	8	Total 8	O 8	0	0
8	F	5	Total 5	O 5	0	0

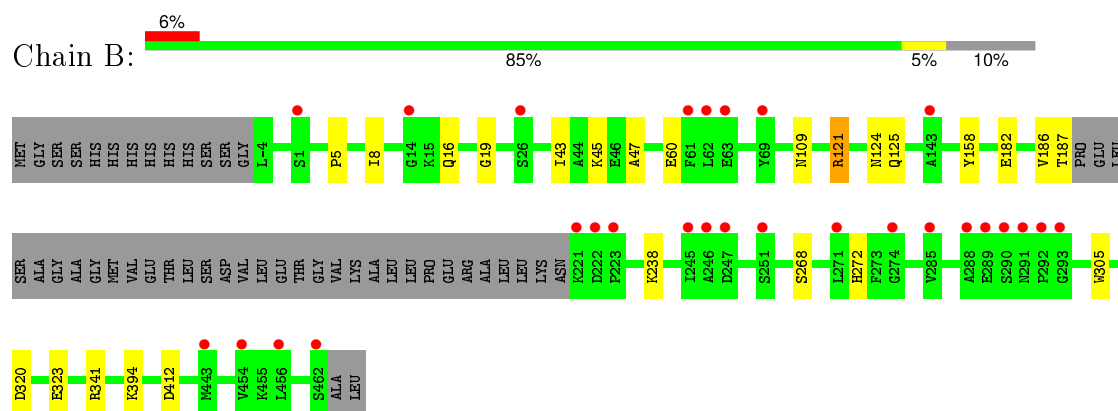
### 3 Residue-property plots [i](#)

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.

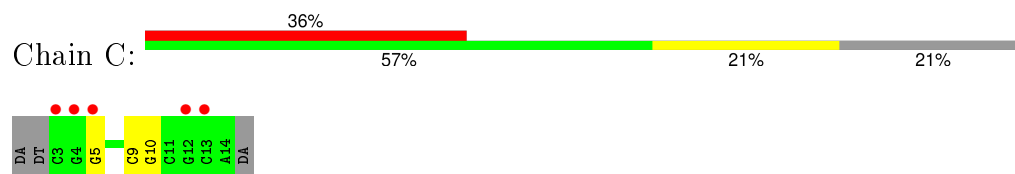
- Molecule 1: DEOXYRIBODIPYRIMIDINE PHOTOLYASE



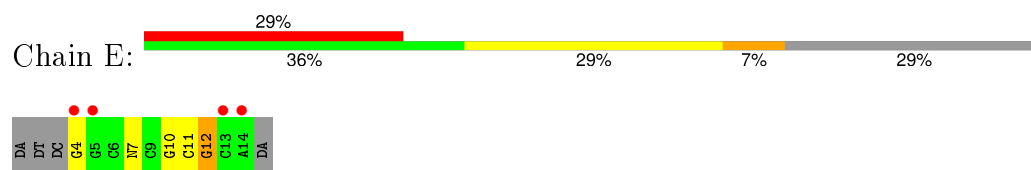
- Molecule 1: DEOXYRIBODIPYRIMIDINE PHOTOLYASE



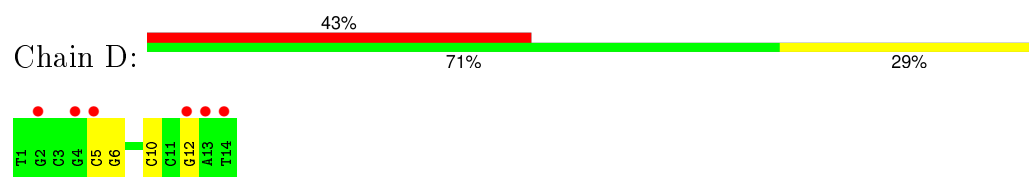
- Molecule 2: CPD-COMPRISING OLIGONUCLEOTIDE



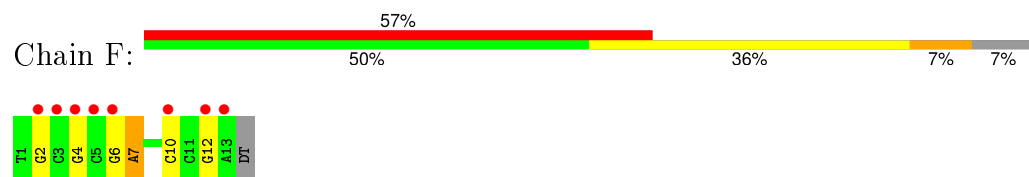
- Molecule 2: CPD-COMPRISING OLIGONUCLEOTIDE



- Molecule 3: COUNTERSTRAND-OLIGONUCLEOTIDE



● Molecule 3: COUNTERSTRAND-OLIGONUCLEOTIDE



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	71.70Å 116.20Å 168.30Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.48 – 2.20 29.47 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.9 (29.48-2.20) 99.9 (29.47-2.20)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.35 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.170 , 0.208 0.182 , 0.219	Depositor DCC
$R_{free}$ test set	1094 reflections (1.54%)	DCC
Wilson B-factor (Å <sup>2</sup> )	32.5	Xtriage
Anisotropy	0.296	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 59.4	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	0 of 72057 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	8978	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.42% 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: GOL, TT, SO4, FAD, ACT

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	0.64	0/3763	0.66	3/5094 (0.1%)
1	B	0.57	0/3640	0.63	2/4930 (0.0%)
2	C	0.85	0/227	1.38	1/345 (0.3%)
2	E	0.89	0/206	1.50	3/313 (1.0%)
3	D	0.90	0/321	1.40	1/494 (0.2%)
3	F	0.83	0/299	1.36	4/460 (0.9%)
All	All	0.64	0/8456	0.79	14/11636 (0.1%)

There are no bond length outliers.

All (14) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	121	ARG	NE-CZ-NH2	-8.47	116.06	120.30
1	B	121	ARG	NE-CZ-NH2	-7.88	116.36	120.30
1	A	121	ARG	NE-CZ-NH1	7.71	124.16	120.30
2	E	12	DG	O4'-C1'-N9	6.20	112.34	108.00
1	B	121	ARG	NE-CZ-NH1	6.12	123.36	120.30
2	E	10	DG	O4'-C1'-N9	5.90	112.13	108.00
2	E	4	DG	P-O3'-C3'	5.83	126.69	119.70
3	F	7	DA	O4'-C1'-N9	-5.65	104.05	108.00
3	D	10	DC	P-O3'-C3'	5.48	126.27	119.70
2	C	5	DG	O4'-C1'-N9	5.32	111.72	108.00
3	F	12	DG	O4'-C1'-N9	5.29	111.70	108.00
1	A	341	ARG	NE-CZ-NH1	5.21	122.91	120.30
3	F	2	DG	P-O3'-C3'	5.13	125.85	119.70
3	F	4	DG	P-O3'-C3'	5.05	125.76	119.70

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	3651	0	3540	23	0
1	B	3522	0	3396	21	0
2	C	242	0	136	2	0
2	E	223	0	125	3	0
3	D	286	0	158	2	0
3	F	266	0	146	2	0
4	A	53	0	31	2	0
4	B	53	0	31	1	0
5	A	5	0	0	0	0
6	A	18	0	24	1	0
7	A	16	0	12	3	0
7	B	8	0	6	3	0
8	A	333	0	0	3	0
8	B	265	0	0	3	0
8	C	15	0	0	0	0
8	D	9	0	0	1	0
8	E	8	0	0	0	0
8	F	5	0	0	0	0
All	All	8978	0	7605	47	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 3.

All (47) 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:8:ILE:HD13	1:A:39[B]:PHE:CZ	1.99	0.98
1:A:8:ILE:CD1	1:A:39[B]:PHE:CE2	2.49	0.95
1:A:39[B]:PHE:CE2	1:A:43:ILE:HD11	2.05	0.91
1:A:8:ILE:HD13	1:A:39[B]:PHE:CE2	2.07	0.89
1:A:125:GLN:HE22	1:B:341:ARG:H	1.36	0.74
1:B:43:ILE:CG2	7:B:1463:ACT:H3	2.24	0.67
1:A:341:ARG:H	1:B:125:GLN:HE22	1.43	0.66
4:A:998:FAD:N1A	8:A:2296:HOH:O	2.30	0.63
1:B:124[B]:ASN:ND2	8:B:2087:HOH:O	2.30	0.63

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:272:HIS:HD2	1:A:412:ASP:OD1	1.82	0.63
1:A:301:GLU:OE2	7:A:1470:ACT:H1	1.99	0.62
1:B:5:PRO:HA	1:B:8:ILE:HD12	1.84	0.60
1:A:320:ASP:OD2	7:A:1467:ACT:H1	2.03	0.58
1:A:39[B]:PHE:HE2	1:A:43:ILE:HD11	1.66	0.57
1:B:186:VAL:HG12	1:B:187:THR:HG23	1.88	0.56
1:A:323[B]:GLU:OE1	1:A:330:LYS:NZ	2.38	0.54
1:B:45:LYS:NZ	1:B:182:GLU:OE1	2.31	0.53
1:A:323[B]:GLU:OE2	1:B:323:GLU:OE1	2.26	0.53
1:A:323[B]:GLU:HG2	8:A:2230:HOH:O	2.09	0.52
1:B:394:LYS:NZ	8:B:2220:HOH:O	2.43	0.51
1:B:305:TRP:CE2	2:E:7:TT:H5A2	2.46	0.51
1:B:272:HIS:HD2	1:B:412:ASP:OD1	1.93	0.51
3:D:12:DG:OP2	8:D:2009:HOH:O	2.19	0.51
1:A:121:ARG:NH2	1:A:320:ASP:OD1	2.34	0.50
1:A:346:THR:HG21	1:B:60:GLU:HG2	1.95	0.49
7:A:1469:ACT:H1	8:A:2055:HOH:O	2.10	0.49
1:B:19:GLY:HA3	1:B:109:ASN:O	2.14	0.48
1:B:121:ARG:NH2	1:B:320:ASP:OD1	2.43	0.47
2:C:9:DC:H2'	2:C:10:DG:C8	2.49	0.47
1:B:16:GLN:NE2	7:B:1463:ACT:H2	2.31	0.46
1:A:451:LYS:NZ	2:C:10:DG:OP2	2.46	0.46
3:F:6:DG:H2''	3:F:7:DA:C8	2.51	0.46
2:E:11:DC:H2''	2:E:12:DG:C8	2.51	0.45
1:B:124[B]:ASN:ND2	8:B:2086:HOH:O	2.48	0.45
1:A:327:SER:HB2	6:A:1465:GOL:H31	1.98	0.45
1:B:158:TYR:CZ	3:F:10:DC:H4'	2.52	0.45
1:B:43:ILE:HD13	1:B:43:ILE:HG21	1.70	0.44
1:A:8:ILE:HD12	1:A:39[B]:PHE:CE2	2.48	0.44
1:B:305:TRP:CD2	2:E:7:TT:H5A2	2.53	0.44
1:A:389[A]:SER:OG	1:A:395:ALA:HB2	2.18	0.44
1:A:268:SER:HB3	4:A:998:FAD:H5'2	2.00	0.43
1:B:47:ALA:HB2	7:B:1463:ACT:H1	2.00	0.42
1:B:268:SER:HB3	4:B:998:FAD:H5'2	2.01	0.42
1:A:8:ILE:HD12	1:A:39[B]:PHE:CD2	2.54	0.42
1:A:57:LEU:HD13	1:A:203:VAL:HG11	2.03	0.41
3:D:5:DC:H2''	3:D:6:DG:C8	2.56	0.40
1:A:244:PHE:CD1	1:A:265:SER:HA	2.57	0.40

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	450/482 (93%)	442 (98%)	8 (2%)	0	100	100
1	B	436/482 (90%)	428 (98%)	8 (2%)	0	100	100
All	All	886/964 (92%)	870 (98%)	16 (2%)	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	388/416 (93%)	388 (100%)	0	100	100
1	B	373/416 (90%)	372 (100%)	1 (0%)	94	98
All	All	761/832 (92%)	760 (100%)	1 (0%)	95	98

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

Mol	Chain	Res	Type
1	B	238	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (10) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	16	GLN
1	A	48	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	125	GLN
1	A	161	HIS
1	A	272	HIS
1	B	16	GLN
1	B	125	GLN
1	B	272	HIS
1	B	279	GLN
1	B	343	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

2 non-standard protein/DNA/RNA residues 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	TT	C	7	2	38,43,44	0.94	1 (2%)	54,69,72	1.20	8 (14%)
2	TT	E	7	2	38,43,44	0.88	2 (5%)	54,69,72	1.58	5 (9%)

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	TT	C	7	2	-	0/18/105/106	0/3/6/6
2	TT	E	7	2	-	0/18/105/106	0/3/6/6

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	7	TT	C1R-N1T	2.97	1.49	1.45
2	E	7	TT	C1'-N1	3.12	1.50	1.45
2	C	7	TT	C1'-N1	3.74	1.50	1.45

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	7	TT	O4R-C1R-C2R	-2.89	100.51	106.28
2	C	7	TT	O2-C2-N1	-2.72	119.26	123.36
2	C	7	TT	C5M-C5T-C5	-2.37	108.98	116.37
2	C	7	TT	O2T-C2T-N3T	-2.33	116.92	121.48
2	E	7	TT	O4R-C1R-C2R	-2.18	101.92	106.28
2	C	7	TT	C5A-C5-C5T	2.03	122.69	116.37
2	C	7	TT	C5T-C6T-N1T	2.14	118.73	115.70
2	C	7	TT	C5-C5T-C4T	2.27	120.82	113.09
2	C	7	TT	C2R-C1R-N1T	2.37	118.88	115.64
2	E	7	TT	C5T-C6T-N1T	2.47	119.20	115.70
2	E	7	TT	C2R-C1R-N1T	3.13	119.93	115.64
2	E	7	TT	C2'-C1'-N1	4.16	121.35	115.64
2	E	7	TT	C7-O3'-C3'	7.56	130.26	115.39

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	E	7	TT	2	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

12 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
5	SO4	A	1463	-	4,4,4	0.30	0	6,6,6	0.30	0
6	GOL	A	1464	-	5,5,5	0.50	0	5,5,5	0.30	0
6	GOL	A	1465	-	5,5,5	0.38	0	5,5,5	0.58	0
6	GOL	A	1466	-	5,5,5	0.35	0	5,5,5	0.37	0
7	ACT	A	1467	-	1,3,3	2.03	1 (100%)	0,3,3	0.00	-
7	ACT	A	1468	-	1,3,3	1.99	0	0,3,3	0.00	-
7	ACT	A	1469	-	1,3,3	1.03	0	0,3,3	0.00	-
7	ACT	A	1470	-	1,3,3	1.20	0	0,3,3	0.00	-
4	FAD	A	998	-	48,58,58	1.31	5 (10%)	54,89,89	2.07	9 (16%)
7	ACT	B	1463	-	1,3,3	1.31	0	0,3,3	0.00	-
7	ACT	B	1464	-	1,3,3	2.15	1 (100%)	0,3,3	0.00	-
4	FAD	B	998	-	48,58,58	1.34	7 (14%)	54,89,89	2.08	9 (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
5	SO4	A	1463	-	-	0/0/0/0	0/0/0/0
6	GOL	A	1464	-	-	0/4/4/4	0/0/0/0
6	GOL	A	1465	-	-	0/4/4/4	0/0/0/0
6	GOL	A	1466	-	-	0/4/4/4	0/0/0/0
7	ACT	A	1467	-	-	0/0/0/0	0/0/0/0
7	ACT	A	1468	-	-	0/0/0/0	0/0/0/0
7	ACT	A	1469	-	-	0/0/0/0	0/0/0/0
7	ACT	A	1470	-	-	0/0/0/0	0/0/0/0
4	FAD	A	998	-	-	0/30/50/50	0/6/6/6
7	ACT	B	1463	-	-	0/0/0/0	0/0/0/0
7	ACT	B	1464	-	-	0/0/0/0	0/0/0/0
4	FAD	B	998	-	-	0/30/50/50	0/6/6/6

All (14) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	998	FAD	C10-N1	2.02	1.39	1.35
7	A	1467	ACT	CH3-C	2.03	1.51	1.48
7	B	1464	ACT	CH3-C	2.15	1.51	1.48

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	998	FAD	C2A-N1A	2.15	1.38	1.33
4	B	998	FAD	C2A-N1A	2.44	1.38	1.33
4	B	998	FAD	C5X-N5	2.46	1.39	1.35
4	B	998	FAD	C4-N3	2.52	1.37	1.33
4	A	998	FAD	C1'-N10	2.75	1.51	1.48
4	A	998	FAD	C4-N3	3.24	1.39	1.33
4	B	998	FAD	C1'-N10	3.53	1.52	1.48
4	A	998	FAD	C2A-N3A	3.62	1.38	1.32
4	B	998	FAD	C2A-N3A	3.68	1.38	1.32
4	A	998	FAD	C4X-N5	3.88	1.39	1.33
4	B	998	FAD	C4X-N5	4.13	1.39	1.33

All (18) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	998	FAD	N3A-C2A-N1A	-10.36	120.96	128.89
4	B	998	FAD	N3A-C2A-N1A	-10.34	120.97	128.89
4	A	998	FAD	C4X-C4-N3	-3.26	119.14	123.59
4	A	998	FAD	C2B-C1B-N9A	-3.04	109.64	114.29
4	B	998	FAD	C4X-C4-N3	-2.78	119.79	123.59
4	B	998	FAD	C4A-C5A-N7A	-2.34	107.32	109.48
4	B	998	FAD	C1B-N9A-C4A	-2.05	123.85	126.94
4	A	998	FAD	O2P-P-O3P	2.26	115.35	105.09
4	A	998	FAD	C1'-N10-C9A	2.44	121.59	118.86
4	A	998	FAD	O4B-C1B-N9A	2.54	113.41	108.10
4	A	998	FAD	C4X-N5-C5X	2.71	119.88	116.76
4	B	998	FAD	C1'-N10-C9A	2.73	121.93	118.86
4	B	998	FAD	O2P-P-O3P	2.78	117.69	105.09
4	B	998	FAD	C5X-C9A-N10	3.36	120.17	117.62
4	B	998	FAD	C4X-N5-C5X	3.57	120.87	116.76
4	A	998	FAD	C5X-C9A-N10	3.65	120.39	117.62
4	A	998	FAD	C4-N3-C2	5.65	120.13	115.25
4	B	998	FAD	C4-N3-C2	6.45	120.82	115.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

7 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	1465	GOL	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	A	1467	ACT	1	0
7	A	1469	ACT	1	0
7	A	1470	ACT	1	0
4	A	998	FAD	2	0
7	B	1463	ACT	3	0
4	B	998	FAD	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 ⓘ

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	450/482 (93%)	-0.12	11 (2%) 62 61	33, 46, 74, 85	0
1	B	434/482 (90%)	0.15	28 (6%) 22 22	38, 55, 83, 98	0
2	C	10/14 (71%)	1.81	5 (50%) 0 0	46, 93, 133, 137	0
2	E	9/14 (64%)	1.80	4 (44%) 0 0	62, 90, 120, 131	0
3	D	14/14 (100%)	1.79	6 (42%) 0 0	65, 97, 126, 148	0
3	F	13/14 (92%)	3.25	8 (61%) 0 0	79, 111, 140, 167	0
All	All	930/1020 (91%)	0.12	62 (6%) 21 20	33, 51, 88, 167	0

All (62) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	F	4	DG	6.7
3	F	13	DA	6.7
2	C	3	DC	5.6
2	E	4	DG	5.5
3	F	5	DC	5.2
1	B	223	PRO	5.1
1	A	186	VAL	4.9
3	F	3	DC	4.9
1	A	198	GLU	4.5
3	D	14	DT	4.4
3	F	12	DG	4.0
1	B	62	LEU	3.8
3	D	5	DC	3.6
1	A	219	LYS	3.5
1	B	245	ILE	3.5
2	E	5	DG	3.5
2	C	5	DG	3.4
1	A	220	ASN	3.4
3	D	13	DA	3.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	462	SER	3.2
1	B	291	ASN	3.2
1	B	247	ASP	3.1
1	B	69	TYR	3.0
3	F	2	DG	3.0
1	B	246	ALA	2.9
1	B	222	ASP	2.9
1	B	271	LEU	2.9
1	B	63	GLU	2.8
1	B	274	GLY	2.8
2	C	4	DG	2.7
1	A	185	SER	2.7
2	E	13	DC	2.7
1	A	213	PRO	2.7
1	B	26	SER	2.7
1	B	456	LEU	2.6
1	B	292	PRO	2.6
1	B	290	SER	2.6
1	B	289	GLU	2.6
3	F	10	DC	2.6
1	B	293	GLY	2.6
1	B	288	ALA	2.6
2	C	12	DG	2.5
3	D	4	DG	2.5
1	B	251	SER	2.4
1	A	85	ARG	2.4
1	B	454	VAL	2.4
2	C	13	DC	2.3
3	D	2	DG	2.3
2	E	14	DA	2.3
1	A	222	ASP	2.2
1	B	443	MET	2.2
1	B	221	LYS	2.2
1	B	1	SER	2.2
1	B	14	GLY	2.2
1	A	218	LEU	2.1
1	B	285	VAL	2.1
1	A	133	GLY	2.1
1	B	61	PHE	2.0
1	B	143	ALA	2.0
1	A	202	ASP	2.0
3	D	12	DG	2.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
3	F	6	DG	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [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
2	TT	E	7	38/39	0.96	0.13	-	55,62,79,84	0
2	TT	C	7	38/39	0.96	0.12	-	40,44,50,53	0

## 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
6	GOL	A	1466	6/6	0.84	0.38	5.40	85,87,87,87	0
7	ACT	A	1470	4/4	0.85	0.21	5.15	86,86,86,87	0
7	ACT	B	1463	4/4	0.94	0.25	4.18	73,73,73,73	0
6	GOL	A	1465	6/6	0.78	0.27	3.15	59,61,62,63	0
7	ACT	A	1467	4/4	0.84	0.20	2.85	67,67,67,67	0
6	GOL	A	1464	6/6	0.87	0.15	1.15	54,55,56,56	0
7	ACT	A	1468	4/4	0.95	0.17	0.34	50,50,51,51	0
7	ACT	B	1464	4/4	0.91	0.18	0.20	59,59,59,60	0
4	FAD	A	998	53/53	0.97	0.14	-0.20	31,36,38,40	0
4	FAD	B	998	53/53	0.96	0.15	-0.39	42,50,57,59	0
5	SO4	A	1463	5/5	0.99	0.09	-0.77	53,53,57,57	0
7	ACT	A	1469	4/4	0.92	0.26	-	76,76,76,76	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.