



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 14, 2017 – 02:17 am GMT

PDB ID : 2G8Y  
Title : The structure of a putative malate/lactate dehydrogenase from E. coli.  
Authors : Cuff, M.E.; Skarina, T.; Edwards, A.; Savchenko, A.; Cymborowski, M.; Minor, W.; Joachimiak, A.; Midwest Center for Structural Genomics (MCSG)  
Deposited on : 2006-03-03  
Resolution : 2.15 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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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) : recalc28949

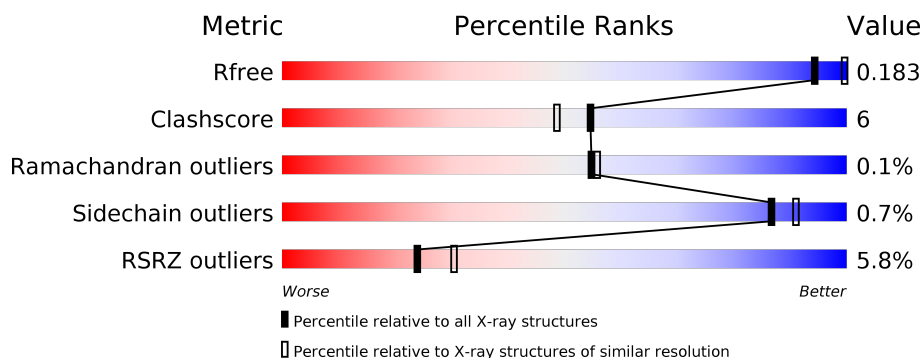
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.15 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))
$R_{free}$	100719	1170 (2.16-2.16)
Clashscore	112137	1278 (2.16-2.16)
Ramachandran outliers	110173	1256 (2.16-2.16)
Sidechain outliers	110143	1255 (2.16-2.16)
RSRZ outliers	101464	1175 (2.16-2.16)

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  $\geq 3$ , 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	385	<div> <div>5%</div> <div> <div></div> <div>82%</div> <div>11%</div> <div>7%</div> </div> </div>
1	B	385	<div> <div>5%</div> <div> <div></div> <div>85%</div> <div>7%</div> <div>7%</div> </div> </div>

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	A	1004	-	-	-	X
2	SO4	A	1007	-	-	-	X
2	SO4	B	1005	-	-	-	X
4	EDO	A	1011	-	-	-	X
4	EDO	B	1012	-	-	-	X
4	EDO	B	1014	-	-	X	-

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 6576 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 Malate/L-lactate dehydrogenases.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	359	Total	C	N	O	S	Se	0	4	0
			2761	1740	497	508	8	8			
1	B	358	Total	C	N	O	S	Se	0	4	0
			2754	1738	496	504	8	8			

There are 66 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-21	MSE	-	SEE REMARK 999	UNP P30178
A	-20	GLY	-	CLONING ARTIFACT	UNP P30178
A	-19	SER	-	CLONING ARTIFACT	UNP P30178
A	-18	SER	-	CLONING ARTIFACT	UNP P30178
A	-17	HIS	-	EXPRESSION TAG	UNP P30178
A	-16	HIS	-	EXPRESSION TAG	UNP P30178
A	-15	HIS	-	EXPRESSION TAG	UNP P30178
A	-14	HIS	-	EXPRESSION TAG	UNP P30178
A	-13	HIS	-	EXPRESSION TAG	UNP P30178
A	-12	HIS	-	EXPRESSION TAG	UNP P30178
A	-11	SER	-	CLONING ARTIFACT	UNP P30178
A	-10	SER	-	CLONING ARTIFACT	UNP P30178
A	-9	GLY	-	CLONING ARTIFACT	UNP P30178
A	-8	ARG	-	CLONING ARTIFACT	UNP P30178
A	-7	GLU	-	CLONING ARTIFACT	UNP P30178
A	-6	ASN	-	CLONING ARTIFACT	UNP P30178
A	-5	LEU	-	CLONING ARTIFACT	UNP P30178
A	-4	TYR	-	CLONING ARTIFACT	UNP P30178
A	-3	PHE	-	CLONING ARTIFACT	UNP P30178
A	-2	GLN	-	CLONING ARTIFACT	UNP P30178
A	-1	GLY	-	CLONING ARTIFACT	UNP P30178
A	0	HIS	-	CLONING ARTIFACT	UNP P30178
A	1	MSE	MET	MODIFIED RESIDUE	UNP P30178
A	23	MSE	MET	MODIFIED RESIDUE	UNP P30178
A	52	MSE	MET	MODIFIED RESIDUE	UNP P30178

*Continued on next page...*

*Continued from previous page...*

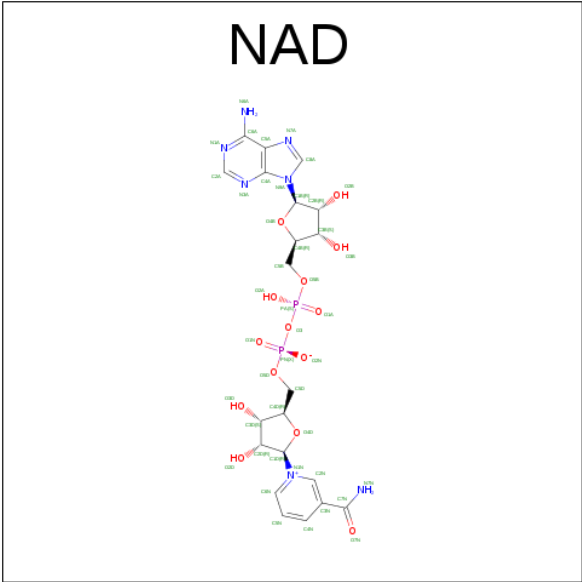
Chain	Residue	Modelled	Actual	Comment	Reference
A	98	MSE	MET	MODIFIED RESIDUE	UNP P30178
A	149	MSE	MET	MODIFIED RESIDUE	UNP P30178
A	220	MSE	MET	MODIFIED RESIDUE	UNP P30178
A	242	MSE	MET	MODIFIED RESIDUE	UNP P30178
A	272	MSE	MET	MODIFIED RESIDUE	UNP P30178
A	347	MSE	MET	MODIFIED RESIDUE	UNP P30178
A	362	GLY	-	CLONING ARTIFACT	UNP P30178
A	363	SER	-	CLONING ARTIFACT	UNP P30178
B	-21	MSE	-	SEE REMARK 999	UNP P30178
B	-20	GLY	-	CLONING ARTIFACT	UNP P30178
B	-19	SER	-	CLONING ARTIFACT	UNP P30178
B	-18	SER	-	CLONING ARTIFACT	UNP P30178
B	-17	HIS	-	EXPRESSION TAG	UNP P30178
B	-16	HIS	-	EXPRESSION TAG	UNP P30178
B	-15	HIS	-	EXPRESSION TAG	UNP P30178
B	-14	HIS	-	EXPRESSION TAG	UNP P30178
B	-13	HIS	-	EXPRESSION TAG	UNP P30178
B	-12	HIS	-	EXPRESSION TAG	UNP P30178
B	-11	SER	-	CLONING ARTIFACT	UNP P30178
B	-10	SER	-	CLONING ARTIFACT	UNP P30178
B	-9	GLY	-	CLONING ARTIFACT	UNP P30178
B	-8	ARG	-	CLONING ARTIFACT	UNP P30178
B	-7	GLU	-	CLONING ARTIFACT	UNP P30178
B	-6	ASN	-	CLONING ARTIFACT	UNP P30178
B	-5	LEU	-	CLONING ARTIFACT	UNP P30178
B	-4	TYR	-	CLONING ARTIFACT	UNP P30178
B	-3	PHE	-	CLONING ARTIFACT	UNP P30178
B	-2	GLN	-	CLONING ARTIFACT	UNP P30178
B	-1	GLY	-	CLONING ARTIFACT	UNP P30178
B	0	HIS	-	CLONING ARTIFACT	UNP P30178
B	1	MSE	MET	MODIFIED RESIDUE	UNP P30178
B	23	MSE	MET	MODIFIED RESIDUE	UNP P30178
B	52	MSE	MET	MODIFIED RESIDUE	UNP P30178
B	98	MSE	MET	MODIFIED RESIDUE	UNP P30178
B	149	MSE	MET	MODIFIED RESIDUE	UNP P30178
B	220	MSE	MET	MODIFIED RESIDUE	UNP P30178
B	242	MSE	MET	MODIFIED RESIDUE	UNP P30178
B	272	MSE	MET	MODIFIED RESIDUE	UNP P30178
B	347	MSE	MET	MODIFIED RESIDUE	UNP P30178
B	362	GLY	-	CLONING ARTIFACT	UNP P30178
B	363	SER	-	CLONING ARTIFACT	UNP P30178

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



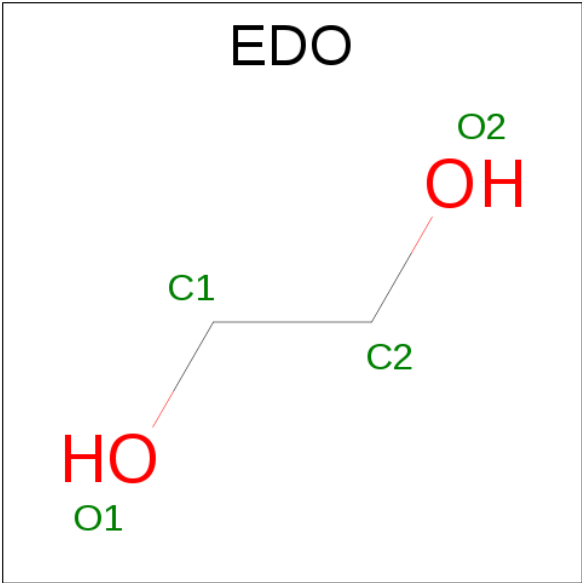
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 is NICOTINAMIDE-ADENINE-DINUCLEOTIDE (three-letter code: NAD) (formula: C<sub>21</sub>H<sub>27</sub>N<sub>7</sub>O<sub>14</sub>P<sub>2</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total 44	C 21	N 7	O 14	P 2	0	0
3	B	1	Total 44	C 21	N 7	O 14	P 2	0	0

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



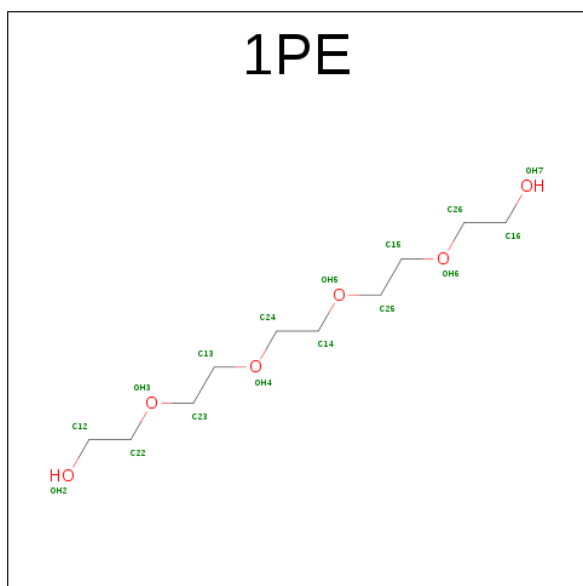
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		

Continued on next page...

Continued from previous page...

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

- Molecule 5 is PENTAETHYLENE GLYCOL (three-letter code: 1PE) (formula:  $C_{10}H_{22}O_6$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			16	10	6		

- Molecule 6 is water.

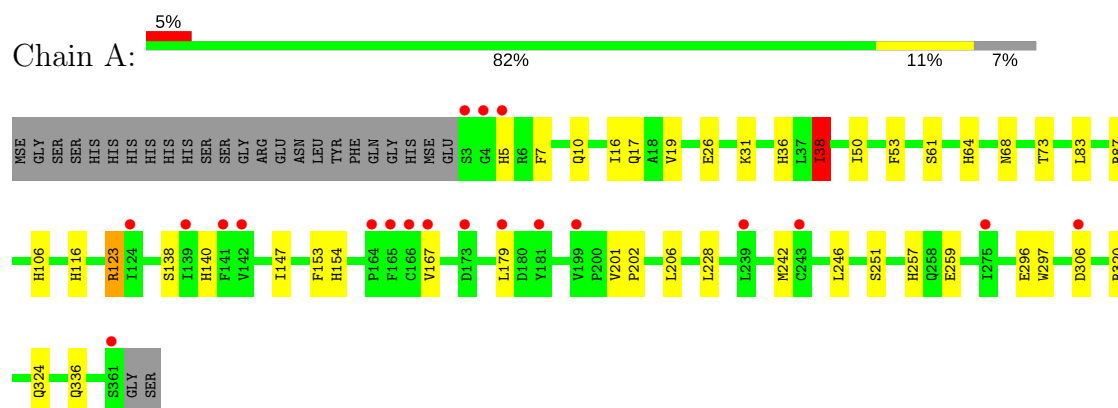


Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	448	Total 448	O 448	0	0
6	B	429	Total 429	O 429	0	0

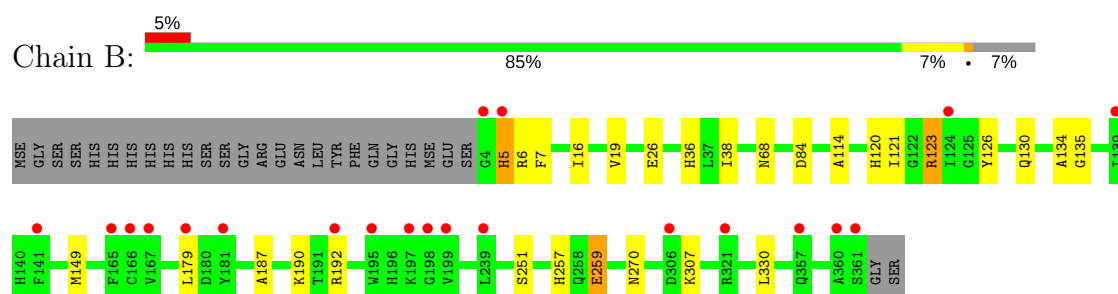
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: Malate/L-lactate dehydrogenases



- Molecule 1: Malate/L-lactate dehydrogenases



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	139.09Å 139.09Å 151.65Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	35.50 – 2.15 35.49 – 2.15	Depositor EDS
% Data completeness (in resolution range)	96.1 (35.50-2.15) 96.1 (35.49-2.15)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.51 (at 2.16Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.158 , 0.186 0.155 , 0.183	Depositor DCC
$R_{free}$ test set	4456 reflections (5.29%)	DCC
Wilson B-factor (Å <sup>2</sup> )	29.1	Xtriage
Anisotropy	0.621	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 36.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.018 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	6576	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.25% 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: 1PE, EDO, NAD, 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	0.68	0/2826	0.70	3/3829 (0.1%)
1	B	0.67	1/2819 (0.0%)	0.66	2/3820 (0.1%)
All	All	0.67	1/5645 (0.0%)	0.68	5/7649 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	259	GLU	CG-CD	6.28	1.61	1.51

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	38	ILE	CG1-CB-CG2	-9.50	90.51	111.40
1	B	123	ARG	NE-CZ-NH1	-7.08	116.76	120.30
1	A	123	ARG	NE-CZ-NH2	-5.88	117.36	120.30
1	A	167	VAL	CB-CA-C	-5.33	101.27	111.40
1	B	84	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	2761	0	2666	38	0
1	B	2754	0	2666	27	0
2	A	15	0	0	0	0
2	B	25	0	0	0	0
3	A	44	0	26	1	0
3	B	44	0	26	0	0
4	A	20	0	30	2	0
4	B	20	0	30	9	0
5	A	16	0	22	1	0
6	A	448	0	0	11	0
6	B	429	0	0	8	0
All	All	6576	0	5466	67	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 6.

All (67) 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:16:ILE:HD12	1:A:38:ILE:HD13	1.34	1.09
1:B:149:MSE:HE3	6:B:1374:HOH:O	1.57	1.04
1:A:61:SER:HA	5:A:1021:1PE:H232	1.43	1.00
1:A:16:ILE:HD12	1:A:38:ILE:CD1	1.98	0.94
1:A:26:GLU:H	1:A:68:ASN:HD21	1.18	0.88
1:B:26:GLU:H	1:B:68:ASN:HD21	1.18	0.86
1:B:121:ILE:H	1:B:270:ASN:HD22	1.21	0.86
1:A:5[A]:HIS:CE1	6:A:1434:HOH:O	2.29	0.84
1:A:296:GLU:HB2	6:A:1438:HOH:O	1.79	0.81
1:A:242:MSE:HE3	6:A:1456:HOH:O	1.82	0.78
1:A:16:ILE:CD1	1:A:38:ILE:CD1	2.66	0.73
1:A:116:HIS:HD2	1:A:251:SER:O	1.73	0.71
1:A:246:LEU:HG	6:A:1456:HOH:O	1.90	0.70
1:B:257:HIS:HD2	1:B:259:GLU:OE1	1.78	0.66
1:B:187:ALA:HA	6:B:1422:HOH:O	1.96	0.66
1:B:5[A]:HIS:HB2	1:B:7:PHE:CE2	2.33	0.63
1:B:192:ARG:HD2	4:B:1020:EDO:H22	1.81	0.62
1:A:17:GLN:HE21	1:A:31:LYS:HA	1.63	0.61
4:B:1014:EDO:H21	6:B:1331:HOH:O	2.00	0.61
1:B:121:ILE:H	1:B:270:ASN:ND2	1.97	0.61
1:B:149:MSE:CE	6:B:1374:HOH:O	2.28	0.61
1:B:6:ARG:HH12	4:B:1014:EDO:C1	2.14	0.60
1:B:130:GLN:NE2	6:B:1192:HOH:O	2.35	0.59

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:306:ASP:HB3	6:A:1344:HOH:O	2.03	0.59
1:B:5[B]:HIS:HA	6:B:1330:HOH:O	2.01	0.58
1:A:257:HIS:HD2	1:A:259[B]:GLU:OE1	1.91	0.54
1:A:16:ILE:CD1	1:A:38:ILE:HD12	2.37	0.54
1:A:116:HIS:CD2	1:A:251:SER:O	2.59	0.53
1:B:135:GLY:HA3	4:B:1012:EDO:H11	1.90	0.52
1:B:6:ARG:HH12	4:B:1014:EDO:H12	1.75	0.51
1:A:106:HIS:HB3	6:A:1036:HOH:O	2.11	0.51
1:A:116:HIS:HE1	6:A:1251:HOH:O	1.95	0.49
1:B:16:ILE:HD12	1:B:38:ILE:HG12	1.95	0.48
1:A:73:THR:HA	1:A:83:LEU:HD23	1.93	0.48
1:A:5[B]:HIS:HB2	1:A:7:PHE:CE2	2.49	0.47
1:B:192:ARG:CD	4:B:1020:EDO:H22	2.44	0.47
1:A:36:HIS:HE1	1:A:123:ARG:H	1.61	0.46
1:A:10[B]:GLN:NE2	6:A:1062:HOH:O	2.48	0.46
1:A:5[A]:HIS:HE1	6:A:1434:HOH:O	1.85	0.45
1:A:138:SER:OG	1:A:140:HIS:HE1	1.98	0.45
4:B:1014:EDO:C2	6:B:1331:HOH:O	2.61	0.45
1:A:206:LEU:CD2	1:A:228:LEU:HD23	2.46	0.45
1:B:179:LEU:C	1:B:179:LEU:HD23	2.37	0.45
1:A:10[B]:GLN:NE2	6:A:1396:HOH:O	2.47	0.45
4:A:1017:EDO:H12	1:B:307:LYS:NZ	2.33	0.44
1:A:19:VAL:HG13	1:A:53:PHE:CZ	2.52	0.44
1:B:6:ARG:HH12	4:B:1014:EDO:H11	1.83	0.44
1:A:123:ARG:HG3	3:A:1001:NAD:O3D	2.18	0.44
1:B:134:ALA:O	4:B:1012:EDO:H22	2.18	0.43
1:A:179:LEU:HD23	1:A:179:LEU:C	2.40	0.43
1:A:153:PHE:O	1:A:154:HIS:HB2	2.20	0.42
1:B:36:HIS:HE1	1:B:123:ARG:H	1.67	0.42
1:A:147:ILE:O	1:A:147:ILE:HG22	2.20	0.42
1:A:320:ARG:O	1:A:324:GLN:HG2	2.20	0.41
1:B:187:ALA:HB3	1:B:190:LYS:HG2	2.02	0.41
1:A:50:ILE:HG21	1:A:50:ILE:HD13	1.76	0.41
1:B:36:HIS:HD2	1:B:126:TYR:CD2	2.38	0.41
1:A:16:ILE:HD11	1:A:50:ILE:HB	2.03	0.41
1:B:114:ALA:HB1	1:B:251:SER:HA	2.02	0.41
1:A:297:TRP:CD2	1:B:257:HIS:HA	2.55	0.41
4:A:1016:EDO:H22	6:A:1433:HOH:O	2.21	0.41
1:A:36:HIS:CE1	1:A:123:ARG:H	2.37	0.41
1:B:7:PHE:HE2	1:B:330:LEU:HB2	1.86	0.41
1:A:201:VAL:HB	1:A:202:PRO:HD2	2.03	0.41

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:16:ILE:HD11	1:A:38:ILE:HD12	2.03	0.40
1:A:64:HIS:O	1:A:87:ARG:HD2	2.20	0.40
1:B:5[B]:HIS:HD2	6:B:1447:HOH:O	2.04	0.40

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	361/385 (94%)	352 (98%)	9 (2%)	0	100	100
1	B	360/385 (94%)	352 (98%)	7 (2%)	1 (0%)	44	41
All	All	721/770 (94%)	704 (98%)	16 (2%)	1 (0%)	55	56

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	120	HIS

### 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	284/292 (97%)	282 (99%)	2 (1%)	87	91
1	B	283/292 (97%)	280 (99%)	3 (1%)	78	82

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
All	All	567/584 (97%)	562 (99%)	5 (1%)	87	87

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	38	ILE
1	A	336	GLN
1	B	5[A]	HIS
1	B	5[B]	HIS
1	B	19	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (19) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	17	GLN
1	A	36	HIS
1	A	62	GLN
1	A	68	ASN
1	A	106	HIS
1	A	116	HIS
1	A	140	HIS
1	A	257	HIS
1	A	258	GLN
1	A	353	GLN
1	B	17	GLN
1	B	36	HIS
1	B	66	GLN
1	B	68	ASN
1	B	106	HIS
1	B	130	GLN
1	B	257	HIS
1	B	270	ASN
1	B	358	GLN

### 5.3.3 RNA ⓘ

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](#)

21 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	$\# Z  > 2$
3	NAD	A	1001	-	41,48,48	1.58	3 (7%)	43,73,73	1.86	4 (9%)
2	SO4	A	1003	-	4,4,4	0.26	0	6,6,6	0.77	0
2	SO4	A	1004	-	4,4,4	0.23	0	6,6,6	0.40	0
2	SO4	A	1007	-	4,4,4	0.54	0	6,6,6	0.55	0
4	EDO	A	1011	-	3,3,3	0.50	0	2,2,2	0.26	0
4	EDO	A	1016	-	3,3,3	0.36	0	2,2,2	0.40	0
4	EDO	A	1017	-	3,3,3	0.79	0	2,2,2	0.76	0
4	EDO	A	1018	-	3,3,3	0.59	0	2,2,2	0.30	0
4	EDO	A	1019	-	3,3,3	0.52	0	2,2,2	0.49	0
5	1PE	A	1021	-	15,15,15	0.43	0	14,14,14	0.46	0
3	NAD	B	1002	-	41,48,48	1.63	3 (7%)	43,73,73	1.69	4 (9%)
2	SO4	B	1005	-	4,4,4	0.35	0	6,6,6	0.41	0
2	SO4	B	1006	-	4,4,4	0.29	0	6,6,6	0.44	0
2	SO4	B	1008	-	4,4,4	0.18	0	6,6,6	0.43	0
2	SO4	B	1009	-	4,4,4	0.23	0	6,6,6	0.09	0
2	SO4	B	1010	-	4,4,4	0.62	0	6,6,6	0.13	0
4	EDO	B	1012	-	3,3,3	0.33	0	2,2,2	0.76	0
4	EDO	B	1013	-	3,3,3	0.51	0	2,2,2	0.27	0
4	EDO	B	1014	-	3,3,3	0.13	0	2,2,2	1.15	0
4	EDO	B	1015	-	3,3,3	0.36	0	2,2,2	0.78	0
4	EDO	B	1020	-	3,3,3	0.50	0	2,2,2	0.15	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
3	NAD	A	1001	-	-	0/22/62/62	0/5/5/5
2	SO4	A	1003	-	-	0/0/0/0	0/0/0/0
2	SO4	A	1004	-	-	0/0/0/0	0/0/0/0
2	SO4	A	1007	-	-	0/0/0/0	0/0/0/0
4	EDO	A	1011	-	-	0/1/1/1	0/0/0/0
4	EDO	A	1016	-	-	0/1/1/1	0/0/0/0
4	EDO	A	1017	-	-	0/1/1/1	0/0/0/0
4	EDO	A	1018	-	-	0/1/1/1	0/0/0/0
4	EDO	A	1019	-	-	0/1/1/1	0/0/0/0
5	1PE	A	1021	-	-	0/13/13/13	0/0/0/0
3	NAD	B	1002	-	-	0/22/62/62	0/5/5/5
2	SO4	B	1005	-	-	0/0/0/0	0/0/0/0
2	SO4	B	1006	-	-	0/0/0/0	0/0/0/0
2	SO4	B	1008	-	-	0/0/0/0	0/0/0/0
2	SO4	B	1009	-	-	0/0/0/0	0/0/0/0
2	SO4	B	1010	-	-	0/0/0/0	0/0/0/0
4	EDO	B	1012	-	-	0/1/1/1	0/0/0/0
4	EDO	B	1013	-	-	0/1/1/1	0/0/0/0
4	EDO	B	1014	-	-	0/1/1/1	0/0/0/0
4	EDO	B	1015	-	-	0/1/1/1	0/0/0/0
4	EDO	B	1020	-	-	0/1/1/1	0/0/0/0

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1001	NAD	C2A-N1A	3.04	1.39	1.33
3	B	1002	NAD	C2A-N1A	3.05	1.39	1.33
3	A	1001	NAD	C2A-N3A	3.58	1.38	1.32
3	B	1002	NAD	C2A-N3A	3.71	1.38	1.32
3	A	1001	NAD	O7N-C7N	7.45	1.39	1.24
3	B	1002	NAD	O7N-C7N	7.76	1.40	1.24

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1001	NAD	N3A-C2A-N1A	-9.97	120.17	128.86
3	B	1002	NAD	N3A-C2A-N1A	-8.80	121.20	128.86
3	A	1001	NAD	C1B-N9A-C4A	-3.17	121.15	126.64

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	1002	NAD	C1B-N9A-C4A	-2.89	121.64	126.64
3	A	1001	NAD	O7N-C7N-C3N	-2.32	116.91	119.62
3	B	1002	NAD	O2D-C2D-C1D	-2.07	105.15	111.61
3	B	1002	NAD	C3N-C2N-N1N	-2.01	118.40	120.43
3	A	1001	NAD	C3N-C7N-N7N	2.59	120.74	117.77

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

7 monomers are involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1001	NAD	1	0
4	A	1016	EDO	1	0
4	A	1017	EDO	1	0
5	A	1021	1PE	1	0
4	B	1012	EDO	2	0
4	B	1014	EDO	5	0
4	B	1020	EDO	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

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	351/385 (91%)	-0.11	20 (5%) 24 32	29, 34, 47, 66	0
1	B	350/385 (90%)	0.09	21 (6%) 23 29	29, 35, 58, 75	0
All	All	701/770 (91%)	-0.01	41 (5%) 24 30	29, 34, 51, 75	0

All (41) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	4	GLY	7.0
1	B	4	GLY	6.8
1	A	3	SER	5.2
1	B	5[A]	HIS	4.3
1	B	199	VAL	4.2
1	B	139	ILE	3.6
1	B	195	TRP	3.2
1	A	142	VAL	3.1
1	A	361	SER	3.0
1	B	357	GLN	3.0
1	B	167	VAL	2.9
1	A	165	PHE	2.9
1	B	239	LEU	2.9
1	A	141	PHE	2.8
1	A	306	ASP	2.8
1	B	306	ASP	2.7
1	B	197	LYS	2.7
1	B	141	PHE	2.7
1	B	165	PHE	2.7
1	A	239	LEU	2.7
1	A	139	ILE	2.6
1	A	124	ILE	2.6
1	A	164	PRO	2.5
1	B	360	ALA	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	179	LEU	2.5
1	A	173	ASP	2.4
1	B	179	LEU	2.4
1	A	167	VAL	2.3
1	B	166	CYS	2.3
1	B	192	ARG	2.3
1	A	5[A]	HIS	2.2
1	A	199	VAL	2.2
1	B	124	ILE	2.2
1	B	321	ARG	2.2
1	A	181	TYR	2.1
1	B	198	GLY	2.1
1	B	181	TYR	2.1
1	A	243	CYS	2.1
1	B	361	SER	2.0
1	A	166	CYS	2.0
1	A	275	ILE	2.0

## 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. 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( $\text{\AA}^2$ )	Q<0.9
4	EDO	A	1011	4/4	0.93	0.16	8.81	56,57,58,60	0
2	SO4	A	1007	5/5	0.93	0.21	4.00	61,61,62,63	0
2	SO4	A	1004	5/5	0.91	0.25	2.67	80,82,83,84	0
4	EDO	B	1012	4/4	0.93	0.18	2.48	55,57,57,58	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	SO4	B	1005	5/5	0.96	0.15	2.09	62,63,64,65	0
2	SO4	B	1008	5/5	0.87	0.17	1.93	101,102,102,102	0
4	EDO	A	1018	4/4	0.88	0.18	1.12	56,57,59,61	0
2	SO4	B	1006	5/5	0.95	0.21	0.91	80,81,81,81	0
4	EDO	A	1019	4/4	0.89	0.17	0.45	48,49,51,54	0
2	SO4	A	1003	5/5	0.93	0.12	-0.16	68,69,69,71	0
4	EDO	B	1014	4/4	0.94	0.11	-0.39	42,42,46,47	0
4	EDO	A	1016	4/4	0.93	0.09	-1.01	47,48,50,50	0
4	EDO	B	1013	4/4	0.91	0.12	-1.19	44,46,49,52	0
3	NAD	A	1001	44/44	0.98	0.07	-1.35	26,32,34,35	0
3	NAD	B	1002	44/44	0.98	0.07	-1.63	23,29,34,34	0
4	EDO	A	1017	4/4	0.59	0.32	-	85,86,86,87	0
2	SO4	B	1009	5/5	0.94	0.27	-	99,99,100,100	0
5	1PE	A	1021	16/16	0.98	0.23	-	23,28,43,44	16
2	SO4	B	1010	5/5	0.91	0.35	-	133,133,133,133	0
4	EDO	B	1015	4/4	0.94	0.23	-	47,52,54,56	0
4	EDO	B	1020	4/4	0.83	0.25	-	84,85,85,85	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.