



# wwPDB X-ray Structure Validation Summary Report

Apr 10, 2014 – 09:40 PM EDT

PDB ID : 4MX2  
Title : Crystal Structure of adenylosuccinate lyase from Leishmania donovani  
Authors : Wernimont, A.K.; Loppnau, P.; Dong, A.; Krojer, T.; Bradley, A.; Bushell, S.; von Delft, F.; Robinson, D.; Gilbert, I.; Bountra, C.; Arrowsmith, C.H.; Edwards, A.M.; Hui, R.; Mottaghi, K.; Structural Genomics Consortium (SGC)  
Deposited on : 2013-09-25  
Resolution : 1.90 Å(reported)

This is a wwPDB 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 <http://wwpdb.org/ValidationPDFNotes.html>

---

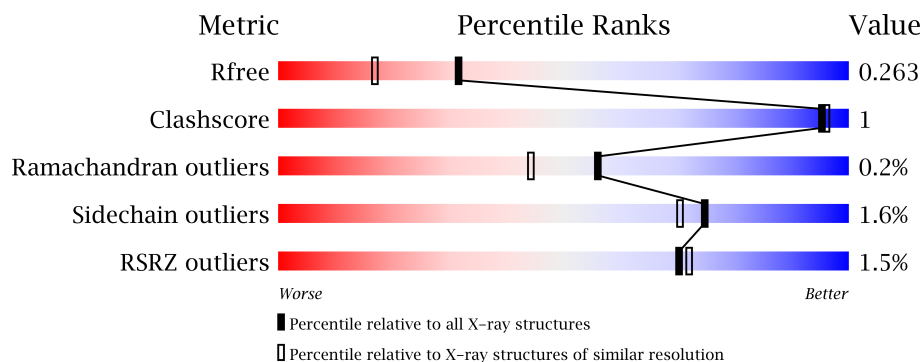
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : **FAILED**  
Xtriage (Phenix) : dev-1439  
EDS : stable22978  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22978

# 1 Overall quality at a glance

The reported resolution of this entry is 1.90 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))
$R_{free}$	66092	3684 (1.90-1.90)
Clashscore	79885	4465 (1.90-1.90)
Ramachandran outliers	78287	4413 (1.90-1.90)
Sidechain outliers	78261	4414 (1.90-1.90)
RSRZ outliers	66119	3686 (1.90-1.90)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	480	
1	B	480	
1	C	480	
1	D	480	
1	E	480	
1	F	480	
1	G	480	
1	H	480	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
4	GOL	E	501	-	X

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 29460 atoms, of which 0 are hydrogen and 0 are deuterium.

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.

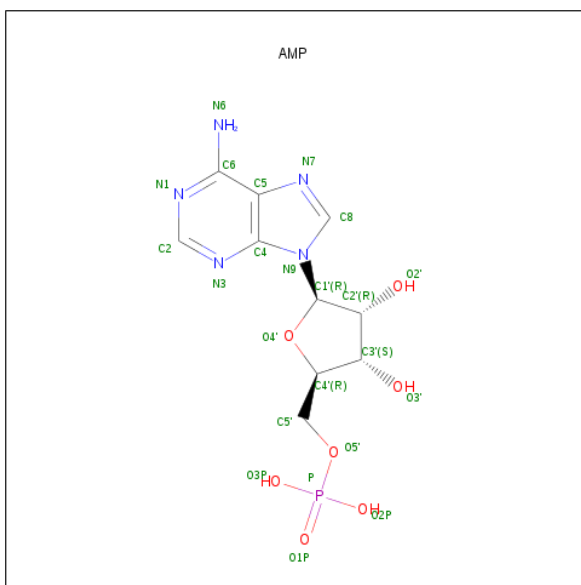
- Molecule 1 is a protein called Adenylosuccinate lyase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	424	Total	C	N	O	S	0	8	0
			3333	2145	557	617	14			
1	B	439	Total	C	N	O	S	0	15	0
			3522	2261	595	652	14			
1	C	426	Total	C	N	O	S	0	8	0
			3367	2167	568	618	14			
1	D	439	Total	C	N	O	S	0	9	0
			3487	2239	588	647	13			
1	E	425	Total	C	N	O	S	0	9	0
			3347	2154	560	619	14			
1	F	431	Total	C	N	O	S	0	12	0
			3450	2222	577	637	14			
1	G	427	Total	C	N	O	S	0	9	0
			3375	2174	567	620	14			
1	H	438	Total	C	N	O	S	0	14	0
			3490	2247	584	644	15			

There are 8 discrepancies between the modelled and reference sequences:

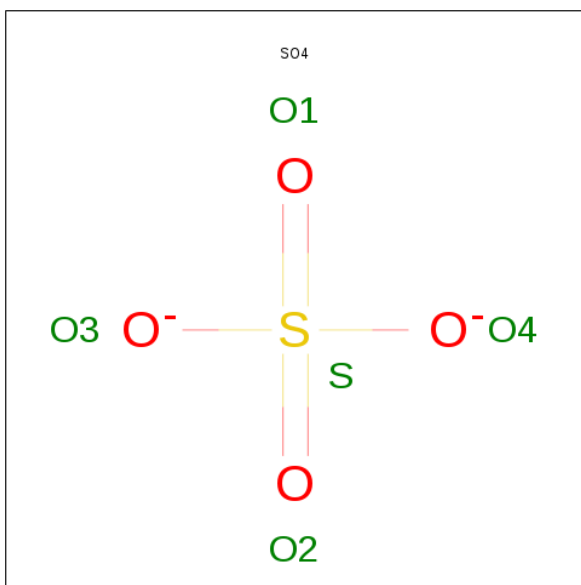
Chain	Residue	Modelled	Actual	Comment	Reference
A	0	GLY	-	EXPRESSION TAG	UNP A7LBL3
B	0	GLY	-	EXPRESSION TAG	UNP A7LBL3
C	0	GLY	-	EXPRESSION TAG	UNP A7LBL3
D	0	GLY	-	EXPRESSION TAG	UNP A7LBL3
E	0	GLY	-	EXPRESSION TAG	UNP A7LBL3
F	0	GLY	-	EXPRESSION TAG	UNP A7LBL3
G	0	GLY	-	EXPRESSION TAG	UNP A7LBL3
H	0	GLY	-	EXPRESSION TAG	UNP A7LBL3

- Molecule 2 is ADENOSINE MONOPHOSPHATE (three-letter code: AMP) (formula: C<sub>10</sub>H<sub>14</sub>N<sub>5</sub>O<sub>7</sub>P).



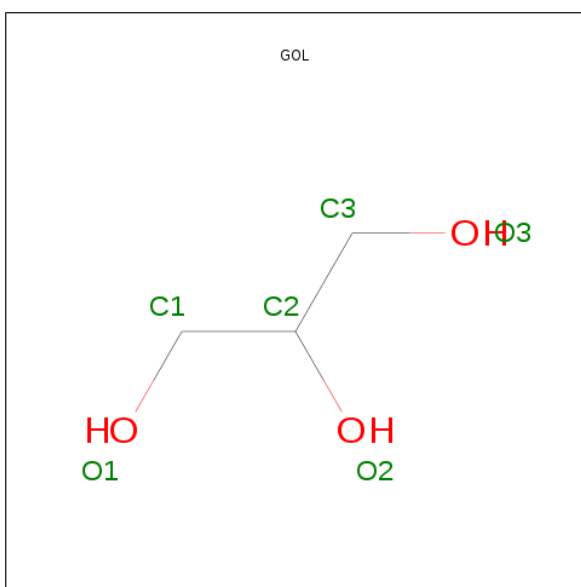
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			23	10	5	7	1		
2	B	1	Total	C	N	O	P	0	0
			23	10	5	7	1		
2	C	1	Total	C	N	O	P	0	0
			23	10	5	7	1		
2	D	1	Total	C	N	O	P	0	0
			23	10	5	7	1		
2	E	1	Total	C	N	O	P	0	0
			23	10	5	7	1		
2	F	1	Total	C	N	O	P	0	0
			23	10	5	7	1		
2	G	1	Total	C	N	O	P	0	0
			23	10	5	7	1		
2	H	1	Total	C	N	O	P	0	0
			23	10	5	7	1		

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	C	1	Total	O	S	0	0
			5	4	1		
3	G	1	Total	O	S	0	0
			5	4	1		
3	H	1	Total	O	S	0	0
			5	4	1		

- Molecule 4 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
4	E	1	Total	C	O	0	0
			6	3	3		

- Molecule 5 is water.

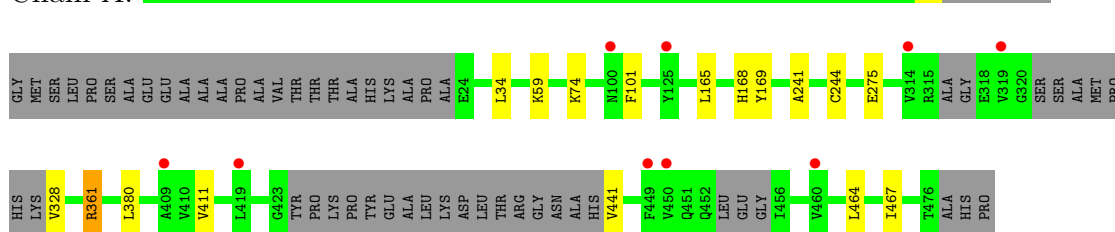
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	232	Total 233	O 233	0	1
5	B	244	Total 244	O 244	0	0
5	C	212	Total 212	O 212	0	0
5	D	264	Total 265	O 265	0	1
5	E	215	Total 215	O 215	0	0
5	F	239	Total 239	O 239	0	0
5	G	217	Total 217	O 217	0	0
5	H	259	Total 259	O 259	0	0

### 3 Residue-property plots

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.

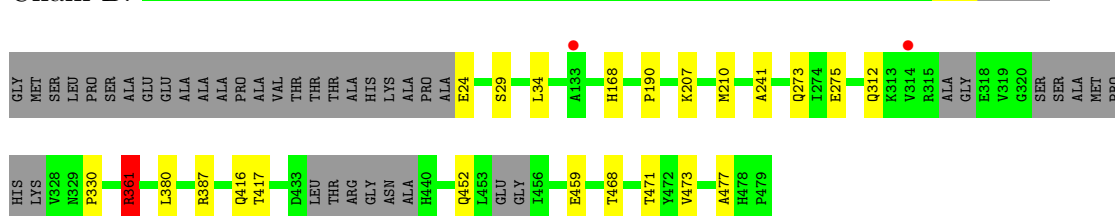
- Molecule 1: Adenylosuccinate lyase

Chain A:



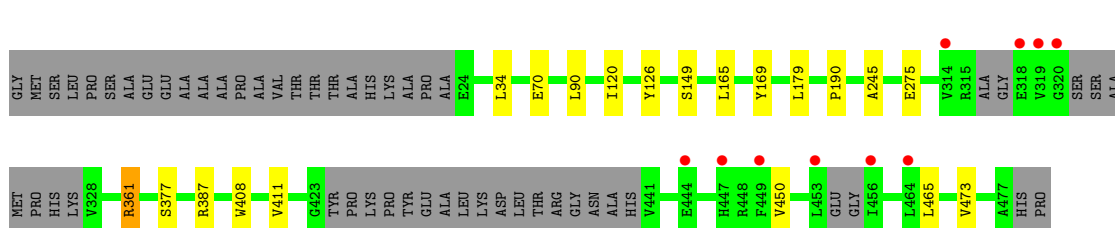
- Molecule 1: Adenylosuccinate lyase

Chain B:



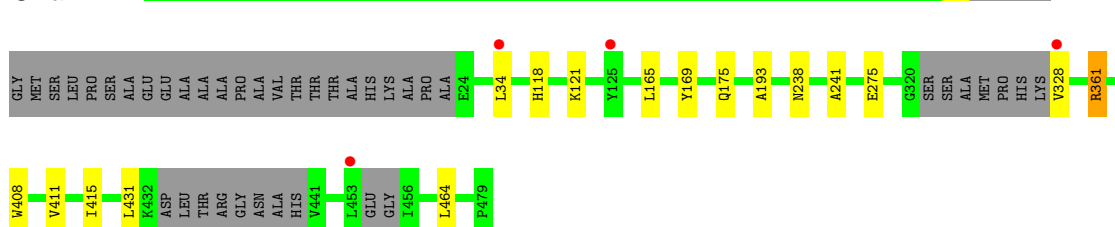
- Molecule 1: Adenylosuccinate lyase

Chain C:



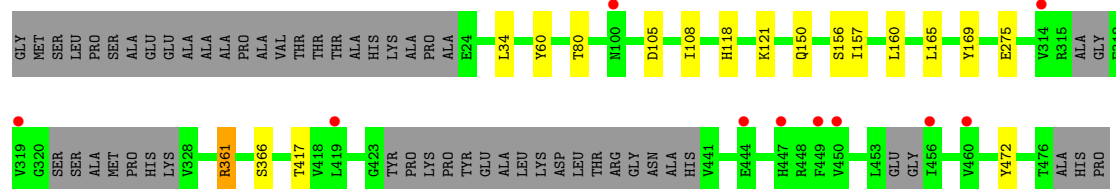
- Molecule 1: Adenylosuccinate lyase

Chain D:

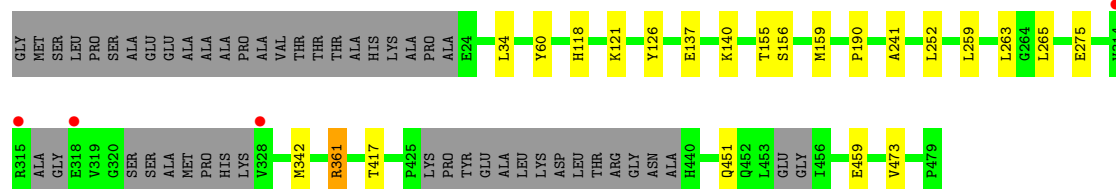




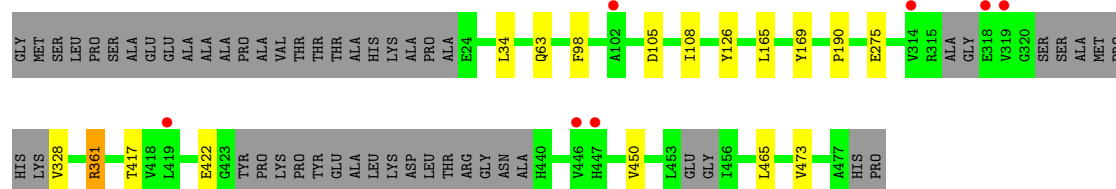
- Molecule 1: Adenylosuccinate lyase

Chain E: 

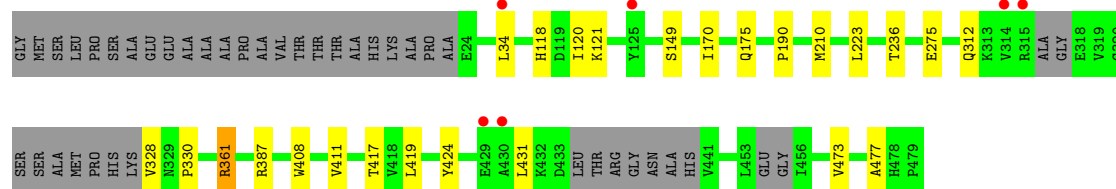
- Molecule 1: Adenylosuccinate lyase

Chain F: 

- Molecule 1: Adenylosuccinate lyase

Chain G: 

- Molecule 1: Adenylosuccinate lyase

Chain H: 

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	124.02Å 194.70Å 147.51Å 90.00° 91.31° 90.00°	Depositor
Resolution (Å)	34.98 – 1.90 34.86 – 1.90	Depositor EDS
% Data completeness (in resolution range)	95.7 (34.98-1.90) 89.4 (34.86-1.90)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.94 (at 1.89Å)	Xtriage
Refinement program	BUSTER 2.10.0	Depositor
R, $R_{free}$	0.218 , 0.253 0.226 , 0.263	Depositor DCC
$R_{free}$ test set	13161 reflections (5.29%)	DCC
Wilson B-factor (Å <sup>2</sup> )	16.5	Xtriage
Anisotropy	0.309	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 32.1	EDS
Estimated twinning fraction	0.199 for -h,-k,l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Outliers	6 of 261938 reflections (0.002%)	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	29460	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 44.58 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 1.4942e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup>Intensities estimated from amplitudes.

## 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, AMP

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	0.50	0/3424	0.59	0/4649
1	B	0.51	0/3626	0.59	0/4921
1	C	0.50	0/3460	0.59	0/4692
1	D	0.50	0/3585	0.59	0/4870
1	E	0.48	0/3442	0.59	0/4673
1	F	0.50	0/3556	0.59	0/4827
1	G	0.49	0/3471	0.59	0/4710
1	H	0.50	0/3607	0.59	0/4900
All	All	0.50	0/28171	0.59	0/38242

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3333	0	3288	7	0
1	B	3522	0	3471	9	0
1	C	3367	0	3322	9	0
1	D	3487	0	3454	10	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	3347	0	3299	12	0
1	F	3450	0	3429	13	0
1	G	3375	0	3335	6	0
1	H	3490	0	3460	13	0
2	A	23	0	12	0	0
2	B	23	0	12	0	0
2	C	23	0	12	0	0
2	D	23	0	12	0	0
2	E	23	0	12	0	0
2	F	23	0	12	0	0
2	G	23	0	12	0	0
2	H	23	0	12	0	0
3	C	5	0	0	0	0
3	G	5	0	0	0	0
3	H	5	0	0	0	0
4	E	6	0	8	4	0
5	A	233	0	0	0	0
5	B	244	0	0	0	0
5	C	212	0	0	0	0
5	D	265	0	0	1	0
5	E	215	0	0	0	0
5	F	239	0	0	2	0
5	G	217	0	0	0	0
5	H	259	0	0	0	0
All	All	29460	0	27162	70	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 1.

The worst 5 of 70 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:E:157:ILE:HD11	4:E:501:GOL:H32	1.39	1.05
1:A:244[B]:CYS:HG	1:E:472:TYR:HD2	1.05	0.98
1:H:170:ILE:HD11	1:H:223:LEU:HD23	1.67	0.76
1:E:366:SER:HA	4:E:501:GOL:H12	1.67	0.75
1:E:157:ILE:CD1	4:E:501:GOL:H32	2.18	0.72

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	422/480 (88%)	409 (97%)	12 (3%)	1 (0%)	56	44
1	B	444/480 (92%)	422 (95%)	21 (5%)	1 (0%)	56	44
1	C	424/480 (88%)	408 (96%)	15 (4%)	1 (0%)	56	44
1	D	441/480 (92%)	421 (96%)	19 (4%)	1 (0%)	56	44
1	E	424/480 (88%)	408 (96%)	15 (4%)	1 (0%)	56	44
1	F	433/480 (90%)	415 (96%)	17 (4%)	1 (0%)	56	44
1	G	427/480 (89%)	411 (96%)	15 (4%)	1 (0%)	56	44
1	H	442/480 (92%)	422 (96%)	19 (4%)	1 (0%)	56	44
All	All	3457/3840 (90%)	3316 (96%)	133 (4%)	8 (0%)	56	44

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	361	ARG
1	B	361	ARG
1	C	361	ARG
1	D	361	ARG
1	E	361	ARG

### 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	350/403 (87%)	344 (98%)	6 (2%)	73	68
1	B	371/403 (92%)	361 (97%)	10 (3%)	57	47
1	C	350/403 (87%)	346 (99%)	4 (1%)	84	82

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	370/403 (92%)	365 (99%)	5 (1%)	78	75
1	E	351/403 (87%)	346 (99%)	5 (1%)	78	75
1	F	368/403 (91%)	362 (98%)	6 (2%)	75	70
1	G	353/403 (88%)	347 (98%)	6 (2%)	73	68
1	H	372/403 (92%)	367 (99%)	5 (1%)	80	77
All	All	2885/3224 (90%)	2838 (98%)	47 (2%)	75	70

5 of 47 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	275	GLU
1	E	80	THR
1	H	275	GLU
1	D	361	ARG
1	E	150	GLN

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	D	166	HIS
1	H	312	GLN
1	D	312	GLN
1	B	452	GLN
1	D	167	HIS

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

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

Mogul failed to run properly - this section will therefore be empty.

## 5.5 Carbohydrates ⓘ

Mogul failed to run properly - this section will therefore be empty.

## 5.6 Ligand geometry ⓘ

Mogul failed to run properly - this section will therefore be empty.

## 5.7 Other polymers ⓘ

Mogul failed to run properly - this section will therefore be empty.

## 5.8 Polymer linkage issues

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	424/480 (88%)	0.08	9 (2%) 60 62	9, 17, 43, 83	0
1	B	439/480 (91%)	0.02	2 (0%) 88 90	8, 16, 40, 75	0
1	C	426/480 (88%)	0.16	10 (2%) 57 59	9, 18, 47, 75	0
1	D	439/480 (91%)	0.02	4 (0%) 81 83	8, 16, 38, 64	0
1	E	425/480 (88%)	0.16	10 (2%) 56 57	10, 19, 46, 74	0
1	F	431/480 (89%)	0.08	4 (0%) 81 83	9, 17, 38, 81	0
1	G	427/480 (88%)	0.15	7 (1%) 68 70	9, 19, 43, 77	0
1	H	438/480 (91%)	0.04	6 (1%) 72 74	8, 16, 39, 73	0
All	All	3449/3840 (89%)	0.09	52 (1%) 70 72	8, 17, 43, 83	0

The worst 5 of 52 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	314	VAL	6.1
1	A	314	VAL	6.0
1	B	314	VAL	5.7
1	A	319	VAL	5.0
1	G	314	VAL	4.4

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

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.



## 6.4 Ligands ⓘ

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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
4	GOL	E	501	6/6	0.17	3.63	26,27,30,31	0
2	AMP	E	502	23/23	0.12	0.12	15,22,23,24	0
2	AMP	A	501	23/23	0.09	-0.43	8,14,16,17	0
2	AMP	B	501	23/23	0.08	-0.57	8,16,19,19	0
2	AMP	D	501	23/23	0.09	-0.58	10,17,19,20	0
2	AMP	C	501	23/23	0.09	-0.70	8,14,18,20	0
2	AMP	F	501	23/23	0.09	-0.77	14,19,22,23	0
3	SO4	H	502	5/5	0.10	-0.85	44,47,48,48	0
3	SO4	C	502	5/5	0.10	-1.32	29,30,32,32	0
2	AMP	H	501	23/23	0.08	-1.43	11,18,20,21	0
2	AMP	G	501	23/23	0.09	-1.50	11,15,19,20	0
3	SO4	G	502	5/5	0.07	-1.60	30,34,36,38	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.