



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 22, 2017 – 05:26 AM EST

PDB ID : 5UTU  
Title : 2.65 Angstrom Resolution Crystal Structure of S-adenosylhomocysteinase from *Cryptosporidium parvum* in Complex with SAH and NAD  
Authors : Minasov, G.; Shuvalova, L.; Dubrovskaya, I.; Winsor, J.; Stam, J.; Kwon, K.; Anderson, W.F.; Center for Structural Genomics of Infectious Diseases (CS-GID)  
Deposited on : 2017-02-15  
Resolution : 2.65 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

<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:

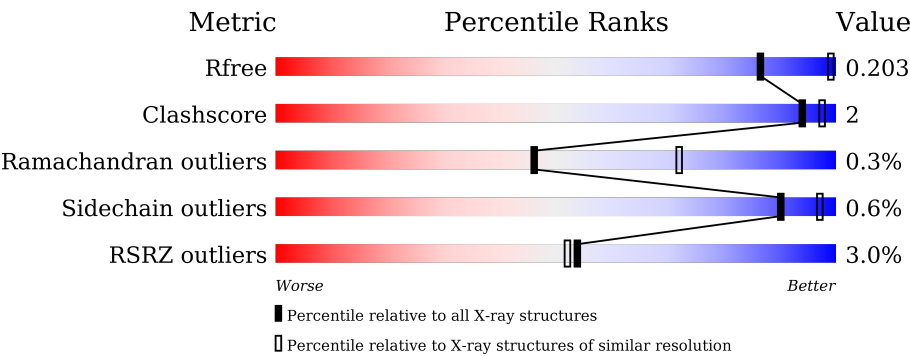
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20028442
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20028442

# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:  
*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.65 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.



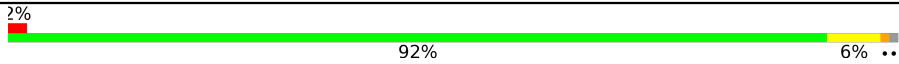
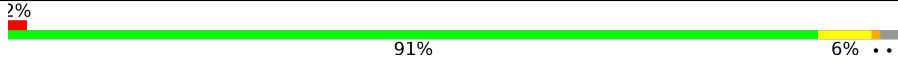
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	91344	3152 (2.70-2.62)
Clashscore	102246	3524 (2.70-2.62)
Ramachandran outliers	100387	3469 (2.70-2.62)
Sidechain outliers	100360	3469 (2.70-2.62)
RSRZ outliers	91569	3161 (2.70-2.62)

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	498	<div><div>3%</div><div><div></div><div>93%</div><div>6%</div><div>.</div></div></div>
1	B	498	<div><div>5%</div><div><div></div><div>95%</div><div>.</div><div>.</div></div></div>
1	C	498	<div><div>2%</div><div><div></div><div>93%</div><div>6%</div><div>.</div></div></div>
1	D	498	<div><div>3%</div><div><div></div><div>93%</div><div>5%</div><div>.</div></div></div>
1	E	498	<div><div>2%</div><div><div></div><div>93%</div><div>6%</div><div>.</div></div></div>
1	F	498	<div><div>3%</div><div><div></div><div>92%</div><div>6%</div><div>..</div></div></div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	G	498	
1	H	498	

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
5	SO4	D	506	-	-	-	X
6	P33	A	508	-	-	-	X

## 2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 33201 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 Adenosylhomocysteinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	494	Total	C	N	O	S	0	8	0
			3979	2528	666	756	29			
1	B	494	Total	C	N	O	S	0	0	0
			3902	2478	654	742	28			
1	C	491	Total	C	N	O	S	0	0	0
			3877	2464	650	735	28			
1	D	493	Total	C	N	O	S	0	1	0
			3901	2477	653	742	29			
1	E	494	Total	C	N	O	S	0	5	0
			3950	2507	665	749	29			
1	F	493	Total	C	N	O	S	0	6	0
			3943	2503	660	751	29			
1	G	495	Total	C	N	O	S	0	4	0
			3949	2511	660	749	29			
1	H	487	Total	C	N	O	S	0	9	0
			3922	2495	654	744	29			

There are 24 discrepancies between the modelled and reference sequences:

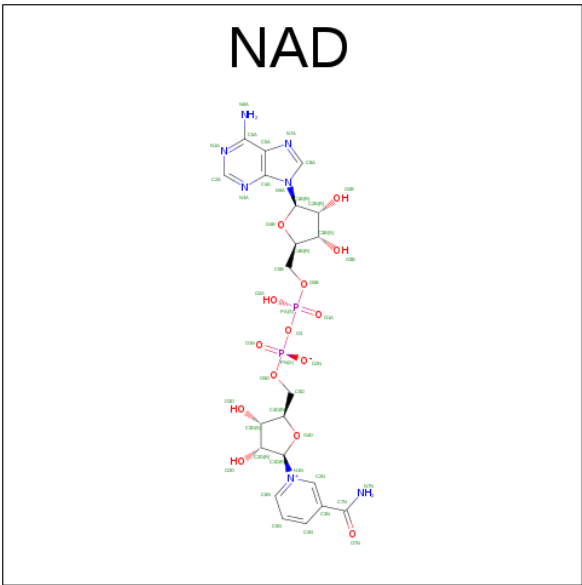
Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	expression tag	UNP Q5CPH1
A	-1	ASN	-	expression tag	UNP Q5CPH1
A	0	ALA	-	expression tag	UNP Q5CPH1
B	-2	SER	-	expression tag	UNP Q5CPH1
B	-1	ASN	-	expression tag	UNP Q5CPH1
B	0	ALA	-	expression tag	UNP Q5CPH1
C	-2	SER	-	expression tag	UNP Q5CPH1
C	-1	ASN	-	expression tag	UNP Q5CPH1
C	0	ALA	-	expression tag	UNP Q5CPH1
D	-2	SER	-	expression tag	UNP Q5CPH1
D	-1	ASN	-	expression tag	UNP Q5CPH1
D	0	ALA	-	expression tag	UNP Q5CPH1
E	-2	SER	-	expression tag	UNP Q5CPH1

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	-1	ASN	-	expression tag	UNP Q5CPH1
E	0	ALA	-	expression tag	UNP Q5CPH1
F	-2	SER	-	expression tag	UNP Q5CPH1
F	-1	ASN	-	expression tag	UNP Q5CPH1
F	0	ALA	-	expression tag	UNP Q5CPH1
G	-2	SER	-	expression tag	UNP Q5CPH1
G	-1	ASN	-	expression tag	UNP Q5CPH1
G	0	ALA	-	expression tag	UNP Q5CPH1
H	-2	SER	-	expression tag	UNP Q5CPH1
H	-1	ASN	-	expression tag	UNP Q5CPH1
H	0	ALA	-	expression tag	UNP Q5CPH1

- Molecule 2 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>).



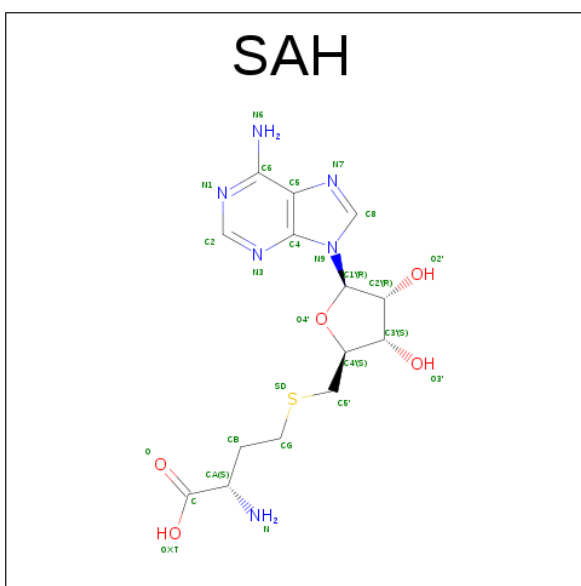
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
2	B	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
2	C	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
2	D	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
2	E	1	Total	C	N	O	P	0	0
			44	21	7	14	2		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	F	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
2	G	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
2	H	1	Total	C	N	O	P	0	0
			44	21	7	14	2		

- Molecule 3 is S-ADENOSYL-L-HOMOCYSTEINE (three-letter code: SAH) (formula:  $C_{14}H_{20}N_6O_5S$ ).

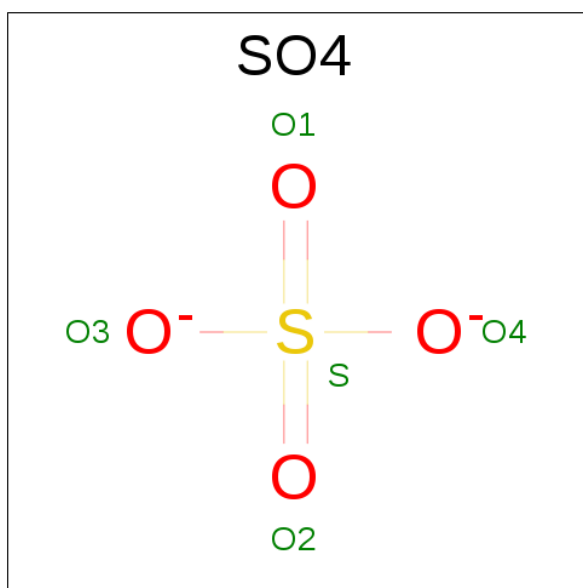


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	S	0	0
			26	14	6	5	1		
3	B	1	Total	C	N	O	S	0	0
			26	14	6	5	1		
3	C	1	Total	C	N	O	S	0	0
			26	14	6	5	1		
3	D	1	Total	C	N	O	S	0	0
			26	14	6	5	1		
3	E	1	Total	C	N	O	S	0	0
			26	14	6	5	1		
3	F	1	Total	C	N	O	S	0	1
			26	14	6	5	1		
3	G	1	Total	C	N	O	S	0	0
			26	14	6	5	1		
3	H	1	Total	C	N	O	S	0	1
			26	14	6	5	1		

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	G	2	Total	Cl	0	0
			2	2		
4	D	1	Total	Cl	0	0
			1	1		
4	E	2	Total	Cl	0	0
			2	2		
4	H	1	Total	Cl	0	0
			1	1		
4	B	2	Total	Cl	0	0
			2	2		
4	C	1	Total	Cl	0	0
			1	1		
4	A	1	Total	Cl	0	0
			1	1		
4	F	1	Total	Cl	0	0
			1	1		

- 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		
5	A	1	Total	O	S	0	0
			5	4	1		
5	A	1	Total	O	S	0	0
			5	4	1		

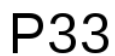
*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	O	S	0	0
			5	4	1		
5	B	1	Total	O	S	0	0
			5	4	1		
5	B	1	Total	O	S	0	0
			5	4	1		
5	C	1	Total	O	S	0	0
			5	4	1		
5	C	1	Total	O	S	0	0
			5	4	1		
5	C	1	Total	O	S	0	0
			5	4	1		
5	C	1	Total	O	S	0	0
			5	4	1		
5	D	1	Total	O	S	0	0
			5	4	1		
5	D	1	Total	O	S	0	0
			5	4	1		
5	D	1	Total	O	S	0	0
			5	4	1		
5	E	1	Total	O	S	0	0
			5	4	1		
5	E	1	Total	O	S	0	0
			5	4	1		
5	F	1	Total	O	S	0	0
			5	4	1		
5	F	1	Total	O	S	0	0
			5	4	1		
5	G	1	Total	O	S	0	0
			5	4	1		
5	G	1	Total	O	S	0	0
			5	4	1		
5	G	1	Total	O	S	0	0
			5	4	1		
5	H	1	Total	O	S	0	0
			5	4	1		
5	H	1	Total	O	S	0	0
			5	4	1		

- Molecule 6 is 3,6,9,12,15,18-HEXAOXAICOSANE-1,20-DIOL (three-letter code: P33) (formula: C<sub>14</sub>H<sub>30</sub>O<sub>8</sub>).





- Molecule 7 is HEXAETHYLENE GLYCOL (three-letter code: P6G) (formula:  $\text{C}_{12}\text{H}_{26}\text{O}_7$ ).

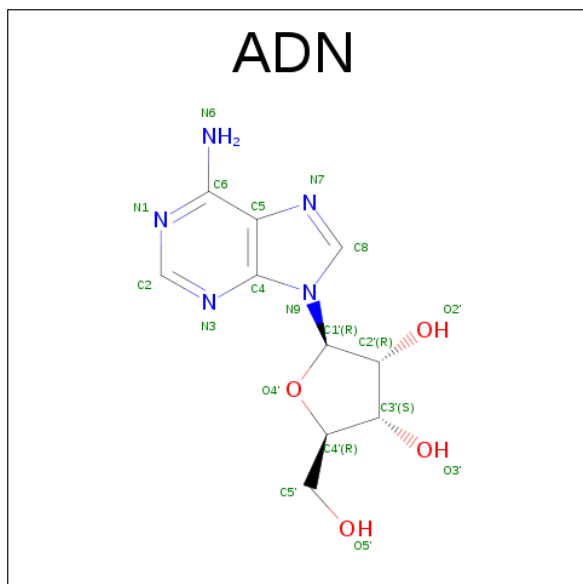


*Continued on next page...*

*Continued from previous page...*

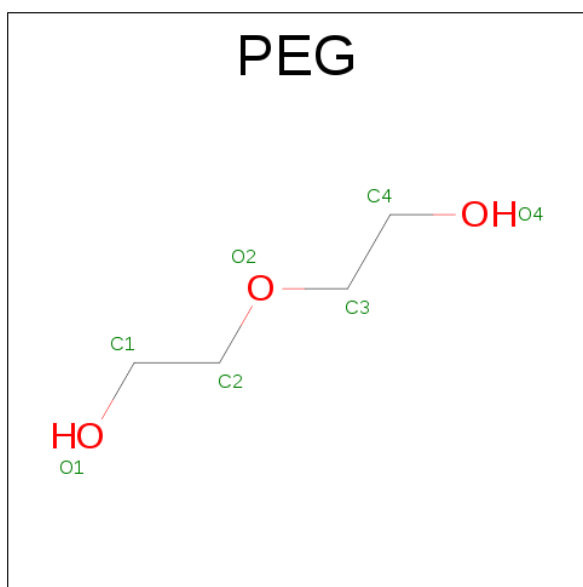
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	F	1	Total	C	O	0	0
			19	12	7		

- Molecule 8 is ADENOSINE (three-letter code: ADN) (formula:  $C_{10}H_{13}N_5O_4$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	F	1	Total	C	N	O	0	1
			19	10	5	4		
8	H	1	Total	C	N	O	0	1
			19	10	5	4		

- Molecule 9 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula:  $C_4H_{10}O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	G	1	Total	C	O	0	0
			7	4	3		
9	G	1	Total	C	O	0	0
			7	4	3		

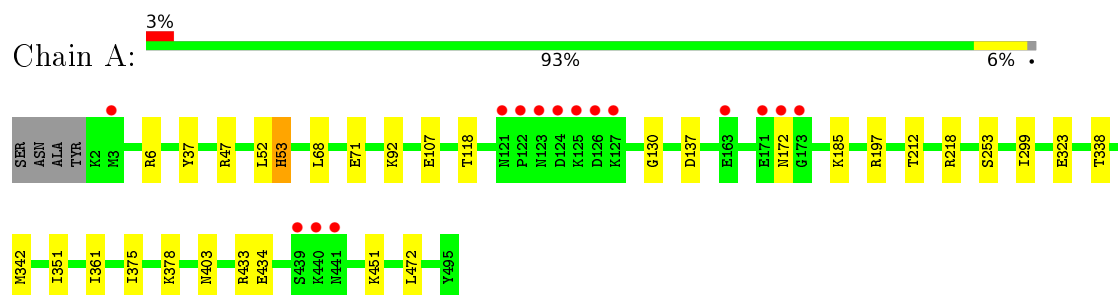
- Molecule 10 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	157	Total	O	0	2
			159	159		
10	B	84	Total	O	0	1
			85	85		
10	C	128	Total	O	0	2
			130	130		
10	D	103	Total	O	0	1
			104	104		
10	E	146	Total	O	0	2
			148	148		
10	F	90	Total	O	0	1
			91	91		
10	G	123	Total	O	0	0
			123	123		
10	H	100	Total	O	0	2
			101	101		

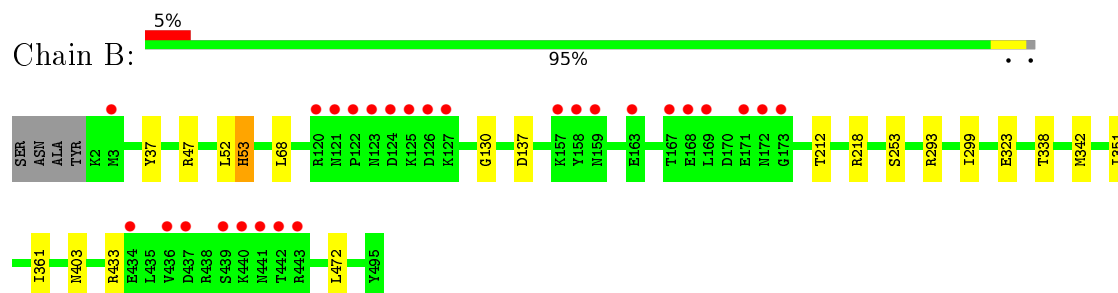
### 3 Residue-property plots [i](#)

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 ( $\text{RSRZ} > 2$ ). 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.

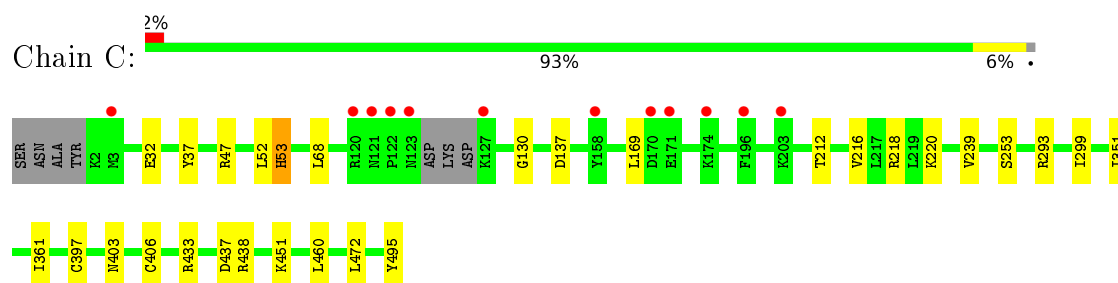
- Molecule 1: Adenosylhomocysteinase



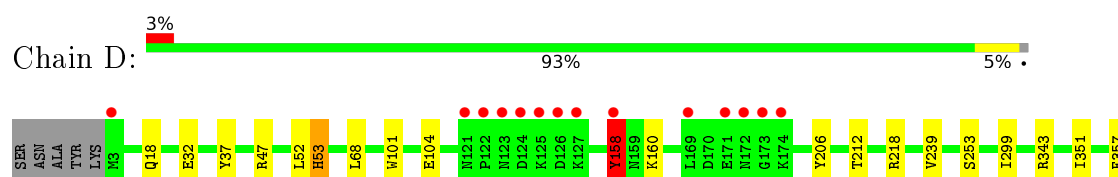
- Molecule 1: Adenosylhomocysteinase

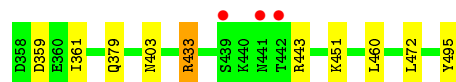


- Molecule 1: Adenosylhomocysteinase

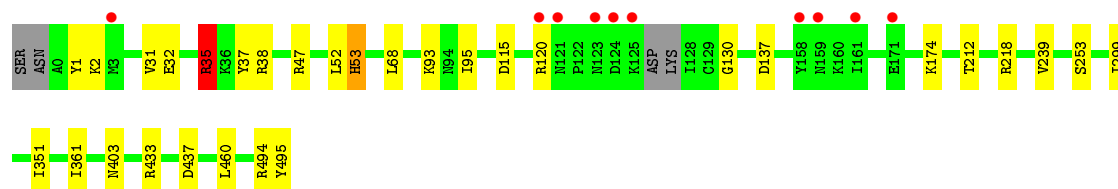


- Molecule 1: Adenosylhomocysteinase

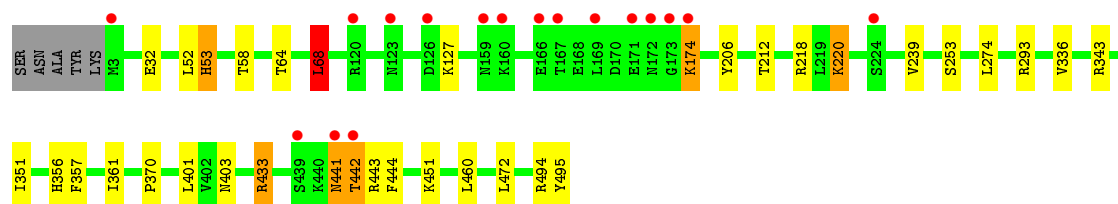
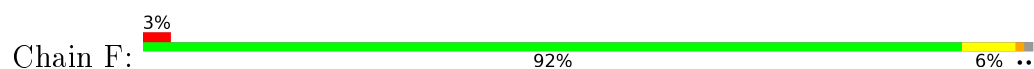




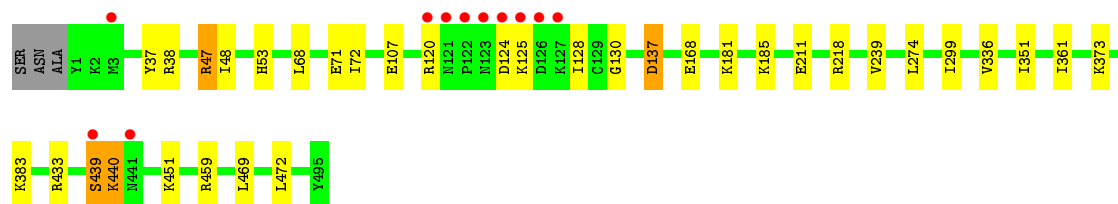
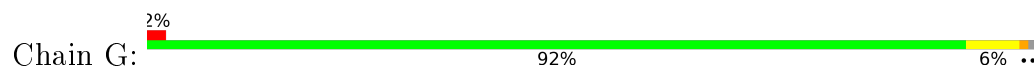
• Molecule 1: Adenosylhomocysteinase



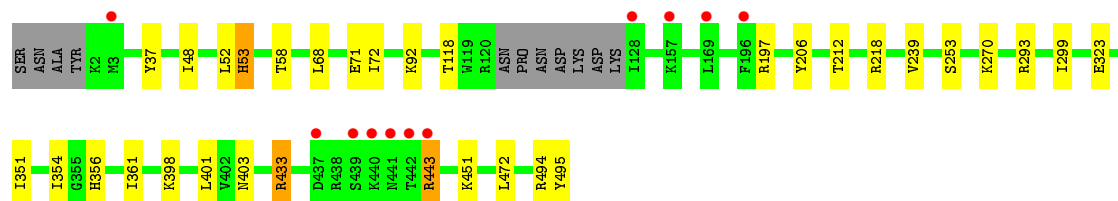
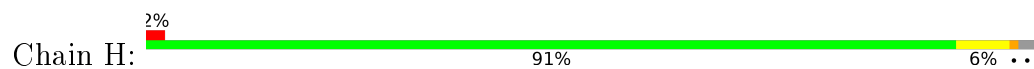
• Molecule 1: Adenosylhomocysteinase



• Molecule 1: Adenosylhomocysteinase



• Molecule 1: Adenosylhomocysteinase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	107.91Å 185.83Å 121.62Å 90.00° 97.98° 90.00°	Depositor
Resolution (Å)	30.00 – 2.65 29.76 – 2.65	Depositor EDS
% Data completeness (in resolution range)	99.9 (30.00-2.65) 100.0 (29.76-2.65)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.13 (at 2.64Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, $R_{free}$	0.170 , 0.202 0.172 , 0.203	Depositor DCC
$R_{free}$ test set	6930 reflections (5.32%)	DCC
Wilson B-factor (Å <sup>2</sup> )	45.4	Xtriage
Anisotropy	0.032	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 51.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	33201	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.66% 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: SAH, NAD, CL, SO4, P6G, ADN, P33, PEG

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	0.47	0/4045	0.77	6/5451 (0.1%)
1	B	0.44	0/3966	0.74	5/5348 (0.1%)
1	C	0.48	0/3940	0.74	6/5312 (0.1%)
1	D	0.47	1/3965 (0.0%)	0.79	10/5347 (0.2%)
1	E	0.48	0/4014	0.77	7/5410 (0.1%)
1	F	0.48	0/4009	0.76	8/5406 (0.1%)
1	G	0.50	0/4015	0.84	13/5414 (0.2%)
1	H	0.47	0/3987	0.73	4/5374 (0.1%)
All	All	0.48	1/31941 (0.0%)	0.77	59/43062 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	359	ASP	CB-CG	-6.50	1.38	1.51

The worst 5 of 59 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	47	ARG	NE-CZ-NH1	-16.30	112.15	120.30
1	G	459	ARG	NE-CZ-NH2	15.01	127.80	120.30
1	D	359	ASP	CB-CA-C	-14.08	82.23	110.40
1	E	35[A]	ARG	CG-CD-NE	-10.33	90.11	111.80
1	E	35[B]	ARG	CG-CD-NE	-10.33	90.11	111.80

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	3979	0	4024	17	0
1	B	3902	0	3961	11	0
1	C	3877	0	3939	13	0
1	D	3901	0	3956	15	0
1	E	3950	0	4004	18	0
1	F	3943	0	3983	22	0
1	G	3949	0	4004	18	0
1	H	3922	0	3965	17	0
2	A	44	0	26	0	0
2	B	44	0	26	0	0
2	C	44	0	26	0	0
2	D	44	0	26	0	0
2	E	44	0	26	0	0
2	F	44	0	26	1	0
2	G	44	0	26	0	0
2	H	44	0	26	0	0
3	A	26	0	19	0	0
3	B	26	0	19	2	0
3	C	26	0	19	0	0
3	D	26	0	19	1	0
3	E	26	0	19	0	0
3	F	26	0	19	3	0
3	G	26	0	19	0	0
3	H	26	0	19	1	0
4	A	1	0	0	0	0
4	B	2	0	0	0	0
4	C	1	0	0	0	0
4	D	1	0	0	0	0
4	E	2	0	0	0	0
4	F	1	0	0	0	0
4	G	2	0	0	0	0
4	H	1	0	0	0	0
5	A	20	0	0	0	0
5	B	10	0	0	0	0
5	C	20	0	0	0	0
5	D	15	0	0	0	0
5	E	10	0	0	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	F	10	0	0	0	0
5	G	15	0	0	0	0
5	H	10	0	0	0	0
6	A	22	0	30	0	0
6	B	22	0	30	0	0
6	H	22	0	30	0	0
7	D	19	0	26	0	0
7	F	19	0	26	0	0
8	F	19	0	13	1	0
8	H	19	0	13	0	0
9	G	14	0	20	0	0
10	A	159	0	0	0	0
10	B	85	0	0	0	0
10	C	130	0	0	0	0
10	D	104	0	0	0	0
10	E	148	0	0	0	0
10	F	91	0	0	0	0
10	G	123	0	0	0	0
10	H	101	0	0	0	0
All	All	33201	0	32384	119	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 119 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:H:118:THR:OG1	1:H:197:ARG:NH2	2.09	0.85
1:A:118:THR:OG1	1:A:197:ARG:NH1	2.09	0.84
1:E:2:LYS:HE2	1:E:115:ASP:OD2	1.81	0.79
1:E:35[B]:ARG:HG2	1:E:38[B]:ARG:HH21	1.52	0.74
1:E:47:ARG:NH1	1:E:130:GLY:O	2.21	0.73

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	500/498 (100%)	486 (97%)	13 (3%)	1 (0%)	52	77
1	B	492/498 (99%)	477 (97%)	14 (3%)	1 (0%)	52	77
1	C	487/498 (98%)	473 (97%)	13 (3%)	1 (0%)	52	77
1	D	492/498 (99%)	478 (97%)	13 (3%)	1 (0%)	52	77
1	E	495/498 (99%)	480 (97%)	14 (3%)	1 (0%)	52	77
1	F	497/498 (100%)	484 (97%)	10 (2%)	3 (1%)	30	54
1	G	497/498 (100%)	482 (97%)	13 (3%)	2 (0%)	39	65
1	H	492/498 (99%)	479 (97%)	12 (2%)	1 (0%)	52	77
All	All	3952/3984 (99%)	3839 (97%)	102 (3%)	11 (0%)	46	72

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	G	440	LYS
1	F	442	THR
1	A	53	HIS
1	B	53	HIS
1	C	53	HIS

### 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	440/435 (101%)	437 (99%)	3 (1%)	88	96

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	432/435 (99%)	431 (100%)	1 (0%)	95	99
1	C	429/435 (99%)	428 (100%)	1 (0%)	95	99
1	D	432/435 (99%)	429 (99%)	3 (1%)	88	96
1	E	436/435 (100%)	433 (99%)	3 (1%)	88	96
1	F	436/435 (100%)	431 (99%)	5 (1%)	80	93
1	G	437/435 (100%)	435 (100%)	2 (0%)	92	98
1	H	433/435 (100%)	429 (99%)	4 (1%)	84	95
All	All	3475/3480 (100%)	3453 (99%)	22 (1%)	90	97

5 of 22 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	35[B]	ARG
1	F	127	LYS
1	H	398	LYS
1	E	174	LYS
1	F	68	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 58 ligands modelled in this entry, 11 are monoatomic - leaving 47 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	NAD	A	501	-	42,48,48	0.82	1 (2%)	46,73,73	1.55	3 (6%)
3	SAH	A	502	-	22,28,28	1.15	2 (9%)	18,40,40	2.16	2 (11%)
5	SO4	A	504	-	4,4,4	0.34	0	6,6,6	0.16	0
5	SO4	A	505	-	4,4,4	0.43	0	6,6,6	0.14	0
5	SO4	A	506	-	4,4,4	0.40	0	6,6,6	0.18	0
5	SO4	A	507	-	4,4,4	0.35	0	6,6,6	0.25	0
6	P33	A	508	-	21,21,21	0.60	0	20,20,20	0.25	0
2	NAD	B	501	-	42,48,48	0.82	1 (2%)	46,73,73	1.51	5 (10%)
3	SAH	B	502	-	22,28,28	1.23	2 (9%)	18,40,40	2.62	5 (27%)
5	SO4	B	505	-	4,4,4	0.36	0	6,6,6	0.16	0
5	SO4	B	506	-	4,4,4	0.39	0	6,6,6	0.12	0
6	P33	B	507	-	21,21,21	0.57	0	20,20,20	0.27	0
2	NAD	C	501	-	42,48,48	0.90	3 (7%)	46,73,73	1.53	6 (13%)
3	SAH	C	502	-	22,28,28	1.04	2 (9%)	18,40,40	2.15	3 (16%)
5	SO4	C	504	-	4,4,4	0.38	0	6,6,6	0.12	0
5	SO4	C	505	-	4,4,4	0.40	0	6,6,6	0.15	0
5	SO4	C	506	-	4,4,4	0.37	0	6,6,6	0.12	0
5	SO4	C	507	-	4,4,4	0.41	0	6,6,6	0.14	0
2	NAD	D	501	-	42,48,48	0.80	2 (4%)	46,73,73	1.58	4 (8%)
3	SAH	D	502	-	22,28,28	1.30	3 (13%)	18,40,40	2.57	5 (27%)
5	SO4	D	504	-	4,4,4	0.43	0	6,6,6	0.08	0
5	SO4	D	505	-	4,4,4	0.39	0	6,6,6	0.27	0
5	SO4	D	506	-	4,4,4	0.39	0	6,6,6	0.08	0
7	P6G	D	507	-	18,18,18	0.54	0	17,17,17	0.29	0
2	NAD	E	501	-	42,48,48	0.87	1 (2%)	46,73,73	1.48	3 (6%)
3	SAH	E	502	-	22,28,28	1.09	2 (9%)	18,40,40	2.13	3 (16%)
5	SO4	E	505	-	4,4,4	0.39	0	6,6,6	0.11	0
5	SO4	E	506	-	4,4,4	0.38	0	6,6,6	0.11	0
2	NAD	F	501	-	42,48,48	0.86	2 (4%)	46,73,73	1.74	4 (8%)
3	SAH	F	502[A]	-	22,28,28	1.21	2 (9%)	18,40,40	2.32	4 (22%)
8	ADN	F	503[B]	-	18,21,21	1.07	1 (5%)	16,31,31	2.60	4 (25%)
5	SO4	F	505	-	4,4,4	0.43	0	6,6,6	0.08	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	SO4	F	506	-	4,4,4	0.37	0	6,6,6	0.19	0
7	P6G	F	507	-	18,18,18	0.54	0	17,17,17	0.25	0
2	NAD	G	501	-	42,48,48	0.80	2 (4%)	46,73,73	1.58	4 (8%)
3	SAH	G	502	-	22,28,28	1.15	2 (9%)	18,40,40	2.15	3 (16%)
5	SO4	G	505	-	4,4,4	0.40	0	6,6,6	0.15	0
5	SO4	G	506	-	4,4,4	0.38	0	6,6,6	0.30	0
5	SO4	G	507	-	4,4,4	0.43	0	6,6,6	0.14	0
9	PEG	G	508	-	6,6,6	0.55	0	5,5,5	0.25	0
9	PEG	G	509	-	6,6,6	0.54	0	5,5,5	0.27	0
2	NAD	H	501	-	42,48,48	0.81	1 (2%)	46,73,73	1.53	4 (8%)
3	SAH	H	502[A]	-	22,28,28	1.13	2 (9%)	18,40,40	1.86	2 (11%)
8	ADN	H	503[B]	-	18,21,21	0.99	1 (5%)	16,31,31	2.41	4 (25%)
5	SO4	H	505	-	4,4,4	0.40	0	6,6,6	0.10	0
5	SO4	H	506	-	4,4,4	0.36	0	6,6,6	0.11	0
6	P33	H	507	-	21,21,21	0.54	0	20,20,20	0.32	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	NAD	A	501	-	-	0/22/62/62	0/5/5/5
3	SAH	A	502	-	-	0/7/31/31	0/3/3/3
5	SO4	A	504	-	-	0/0/0/0	0/0/0/0
5	SO4	A	505	-	-	0/0/0/0	0/0/0/0
5	SO4	A	506	-	-	0/0/0/0	0/0/0/0
5	SO4	A	507	-	-	0/0/0/0	0/0/0/0
6	P33	A	508	-	-	0/19/19/19	0/0/0/0
2	NAD	B	501	-	-	0/22/62/62	0/5/5/5
3	SAH	B	502	-	-	0/7/31/31	0/3/3/3
5	SO4	B	505	-	-	0/0/0/0	0/0/0/0
5	SO4	B	506	-	-	0/0/0/0	0/0/0/0
6	P33	B	507	-	-	0/19/19/19	0/0/0/0
2	NAD	C	501	-	-	0/22/62/62	0/5/5/5
3	SAH	C	502	-	-	0/7/31/31	0/3/3/3
5	SO4	C	504	-	-	0/0/0/0	0/0/0/0
5	SO4	C	505	-	-	0/0/0/0	0/0/0/0
5	SO4	C	506	-	-	0/0/0/0	0/0/0/0
5	SO4	C	507	-	-	0/0/0/0	0/0/0/0
2	NAD	D	501	-	-	0/22/62/62	0/5/5/5

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	SAH	D	502	-	-	0/7/31/31	0/3/3/3
5	SO4	D	504	-	-	0/0/0/0	0/0/0/0
5	SO4	D	505	-	-	0/0/0/0	0/0/0/0
5	SO4	D	506	-	-	0/0/0/0	0/0/0/0
7	P6G	D	507	-	-	0/16/16/16	0/0/0/0
2	NAD	E	501	-	-	0/22/62/62	0/5/5/5
3	SAH	E	502	-	-	0/7/31/31	0/3/3/3
5	SO4	E	505	-	-	0/0/0/0	0/0/0/0
5	SO4	E	506	-	-	0/0/0/0	0/0/0/0
2	NAD	F	501	-	-	0/22/62/62	0/5/5/5
3	SAH	F	502[A]	-	-	0/7/31/31	0/3/3/3
8	ADN	F	503[B]	-	-	0/2/22/22	0/3/3/3
5	SO4	F	505	-	-	0/0/0/0	0/0/0/0
5	SO4	F	506	-	-	0/0/0/0	0/0/0/0
7	P6G	F	507	-	-	0/16/16/16	0/0/0/0
2	NAD	G	501	-	-	0/22/62/62	0/5/5/5
3	SAH	G	502	-	-	0/7/31/31	0/3/3/3
5	SO4	G	505	-	-	0/0/0/0	0/0/0/0
5	SO4	G	506	-	-	0/0/0/0	0/0/0/0
5	SO4	G	507	-	-	0/0/0/0	0/0/0/0
9	PEG	G	508	-	-	0/4/4/4	0/0/0/0
9	PEG	G	509	-	-	0/4/4/4	0/0/0/0
2	NAD	H	501	-	-	0/22/62/62	0/5/5/5
3	SAH	H	502[A]	-	-	0/7/31/31	0/3/3/3
8	ADN	H	503[B]	-	-	0/2/22/22	0/3/3/3
5	SO4	H	505	-	-	0/0/0/0	0/0/0/0
5	SO4	H	506	-	-	0/0/0/0	0/0/0/0
6	P33	H	507	-	-	0/19/19/19	0/0/0/0

The worst 5 of 32 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	502	SAH	C5'-SD	-3.75	1.73	1.81
3	F	502[A]	SAH	C5'-SD	-3.33	1.74	1.81
3	A	502	SAH	C5'-SD	-3.30	1.74	1.81
3	B	502	SAH	C5'-SD	-3.29	1.74	1.81
3	G	502	SAH	C5'-SD	-2.95	1.75	1.81

The worst 5 of 68 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	501	NAD	N3A-C2A-N1A	-8.43	122.25	128.87

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	501	NAD	N3A-C2A-N1A	-8.06	122.54	128.87
2	A	501	NAD	N3A-C2A-N1A	-7.95	122.62	128.87
3	E	502	SAH	N3-C2-N1	-7.60	122.90	128.87
3	C	502	SAH	N3-C2-N1	-7.50	122.98	128.87

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	502	SAH	2	0
3	D	502	SAH	1	0
2	F	501	NAD	1	0
3	F	502[A]	SAH	3	0
8	F	503[B]	ADN	1	0
3	H	502[A]	SAH	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	494/498 (99%)	-0.41	15 (3%) 54 52	30, 44, 84, 143	0
1	B	494/498 (99%)	-0.21	27 (5%) 29 26	30, 54, 117, 170	0
1	C	491/498 (98%)	-0.43	12 (2%) 62 60	31, 45, 86, 151	0
1	D	493/498 (98%)	-0.31	17 (3%) 49 47	29, 50, 105, 157	0
1	E	494/498 (99%)	-0.45	10 (2%) 68 67	30, 44, 85, 143	0
1	F	493/498 (98%)	-0.27	17 (3%) 49 47	27, 49, 105, 128	0
1	G	495/498 (99%)	-0.43	11 (2%) 65 64	28, 42, 85, 140	0
1	H	487/498 (97%)	-0.39	11 (2%) 64 62	29, 52, 95, 136	0
All	All	3941/3984 (98%)	-0.36	120 (3%) 54 52	27, 47, 97, 170	0

The worst 5 of 120 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	3	MET	7.1
1	D	126	ASP	6.8
1	B	3	MET	6.2
1	F	442	THR	6.0
1	C	123	ASN	5.9

### 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	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
6	P33	A	508	22/22	0.86	0.25	4.88	62,77,84,90	0
5	SO4	D	506	5/5	0.86	0.28	4.13	47,50,57,58	5
6	P33	H	507	22/22	0.90	0.22	1.89	53,66,69,71	0
7	P6G	F	507	19/19	0.91	0.18	1.87	48,59,70,71	0
5	SO4	C	507	5/5	0.88	0.28	1.82	94,99,104,106	0
5	SO4	C	505	5/5	0.96	0.21	1.76	72,72,82,83	0
6	P33	B	507	22/22	0.91	0.23	1.68	51,66,71,75	0
5	SO4	A	507	5/5	0.97	0.24	1.55	64,65,66,77	0
3	SAH	F	502[A]	26/26	0.88	0.27	1.42	37,65,72,73	26
9	PEG	G	508	7/7	0.83	0.20	1.38	62,67,73,76	0
3	SAH	H	502[A]	26/26	0.93	0.19	1.32	47,93,98,100	26
5	SO4	D	505	5/5	0.95	0.16	1.20	73,75,76,78	0
8	ADN	F	503[B]	19/19	0.86	0.26	1.18	40,52,58,60	19
8	ADN	H	503[B]	19/19	0.92	0.18	0.97	29,35,37,37	19
3	SAH	B	502	26/26	0.93	0.18	0.75	42,52,80,96	0
7	P6G	D	507	19/19	0.94	0.16	0.68	46,62,69,70	0
5	SO4	H	506	5/5	0.94	0.16	0.58	81,94,99,101	0
3	SAH	E	502	26/26	0.96	0.16	0.55	38,47,59,63	0
3	SAH	D	502	26/26	0.94	0.19	0.45	48,57,64,72	0
5	SO4	G	506	5/5	0.97	0.15	0.23	58,61,71,74	0
3	SAH	C	502	26/26	0.97	0.15	0.17	29,41,63,66	0
2	NAD	E	501	44/44	0.98	0.13	0.11	29,35,39,42	0
5	SO4	F	506	5/5	0.96	0.15	0.10	74,74,77,87	0
3	SAH	G	502	26/26	0.97	0.15	0.07	29,41,56,60	0
3	SAH	A	502	26/26	0.96	0.16	0.07	32,38,45,54	0
2	NAD	C	501	44/44	0.97	0.13	-0.04	31,36,40,43	0
5	SO4	H	505	5/5	0.74	0.20	-0.26	93,101,109,119	0
2	NAD	D	501	44/44	0.98	0.12	-0.26	27,31,35,37	0
2	NAD	A	501	44/44	0.98	0.12	-0.32	29,32,36,37	0
2	NAD	F	501	44/44	0.98	0.11	-0.59	26,30,34,36	0
5	SO4	E	506	5/5	0.98	0.10	-0.68	62,70,71,75	0
2	NAD	G	501	44/44	0.99	0.11	-0.70	27,31,34,39	0
2	NAD	B	501	44/44	0.98	0.10	-0.91	30,35,40,43	0
2	NAD	H	501	44/44	0.99	0.09	-0.97	33,36,38,40	0
4	CL	H	504	1/1	0.92	0.09	-1.72	67,67,67,67	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
4	CL	G	503	1/1	0.69	0.18	-	71,71,71,71	0
5	SO4	C	504	5/5	0.93	0.41	-	95,96,97,116	0
5	SO4	F	505	5/5	0.86	0.30	-	85,91,97,106	0
5	SO4	C	506	5/5	0.83	0.30	-	102,108,115,119	0
4	CL	G	504	1/1	0.85	0.09	-	65,65,65,65	0
5	SO4	E	505	5/5	0.95	0.34	-	86,91,96,101	0
5	SO4	A	506	5/5	0.84	0.32	-	91,98,105,107	0
5	SO4	G	505	5/5	0.92	0.35	-	77,80,87,94	0
9	PEG	G	509	7/7	0.73	0.35	-	82,94,103,104	0
4	CL	B	504	1/1	0.63	0.15	-	84,84,84,84	0
5	SO4	B	505	5/5	0.84	0.39	-	99,107,116,123	0
5	SO4	B	506	5/5	0.78	0.21	-	101,102,116,118	0
4	CL	E	503	1/1	0.91	0.11	-	66,66,66,66	0
5	SO4	A	504	5/5	0.96	0.36	-	80,83,94,98	0
5	SO4	D	504	5/5	0.95	0.13	-	85,86,90,100	0
4	CL	C	503	1/1	0.90	0.23	-	74,74,74,74	0
4	CL	B	503	1/1	0.97	0.09	-	52,52,52,52	0
5	SO4	G	507	5/5	0.88	0.15	-	74,81,87,95	0
4	CL	A	503	1/1	0.82	0.19	-	86,86,86,86	0
4	CL	F	504	1/1	0.82	0.27	-	80,80,80,80	0
4	CL	D	503	1/1	0.77	0.34	-	80,80,80,80	0
5	SO4	A	505	5/5	0.86	0.19	-	80,82,86,98	0
4	CL	E	504	1/1	0.99	0.04	-	65,65,65,65	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.