



# Full wwPDB X-ray Structure Validation Report ⓘ

Jul 21, 2016 – 08:47 PM EDT

PDB ID : 5JPI  
Title : 2.15 Angstrom Crystal Structure of S-adenosylhomocysteinase from *Cryptosporidium parvum* in Complex with D-Eritadenine and NAD  
Authors : Minasov, G.; Shuvalova, L.; Kiryukhina, O.; Dubrovskaya, I.; Bishop, B.; Kwon, K.; Anderson, W.F.; Center for Structural Genomics of Infectious Diseases (CSGID)  
Deposited on : 2016-05-03  
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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027790  
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-20027790

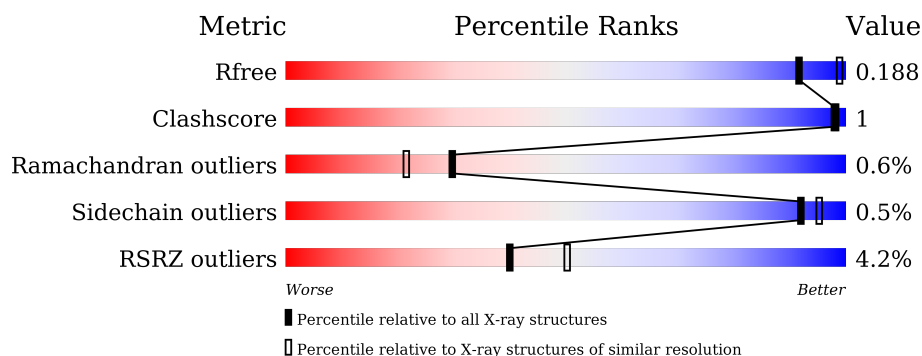
# 1 Overall quality at a glance

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))
$R_{free}$	91344	1045 (2.16-2.16)
Clashscore	102246	1152 (2.16-2.16)
Ramachandran outliers	100387	1131 (2.16-2.16)
Sidechain outliers	100360	1131 (2.16-2.16)
RSRZ outliers	91569	1050 (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	498	<div> <div>5%</div> <div>96%</div> <div>.</div> </div>
1	B	498	<div> <div>4%</div> <div>95%</div> <div>..</div> </div>
1	C	498	<div> <div>3%</div> <div>94%</div> <div>..</div> </div>
1	D	498	<div> <div>5%</div> <div>96%</div> <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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	GOL	A	503	-	-	-	X
4	GOL	A	504	-	-	-	X
4	GOL	B	503	-	-	-	X
4	GOL	D	503	-	-	-	X
5	PEG	C	504	-	-	-	X
5	PEG	C	508	-	-	-	X
8	EDO	C	509	-	-	-	X
8	EDO	D	505	-	-	-	X
9	TRS	D	504	-	-	-	X

## 2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 17474 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.

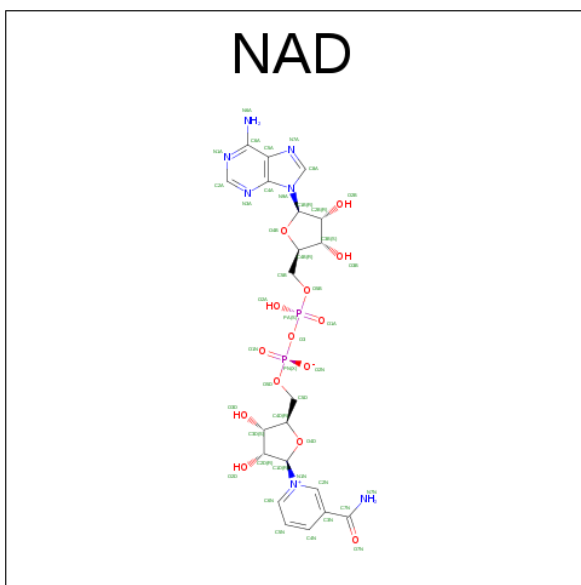
- Molecule 1 is a protein called Adenosylhomocysteinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	496	Total	C	N	O	S	0	13	0
			4024	2554	674	767	29			
1	B	491	Total	C	N	O	S	0	8	0
			3937	2497	660	752	28			
1	C	492	Total	C	N	O	S	0	8	0
			3954	2512	661	752	29			
1	D	491	Total	C	N	O	S	0	10	0
			3961	2515	662	755	29			

There are 12 discrepancies between the modelled and reference sequences:

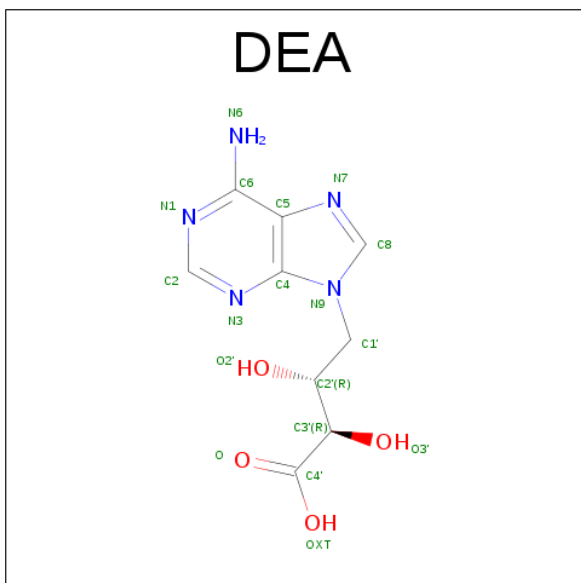
Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	expression tag	UNP Q5CPH1
A	-1	ASN	-	expression tag	UNP Q5CPH1
A	0	ALA	-	expression tag	UNP Q5CPH1
B	-2	SER	-	expression tag	UNP Q5CPH1
B	-1	ASN	-	expression tag	UNP Q5CPH1
B	0	ALA	-	expression tag	UNP Q5CPH1
C	-2	SER	-	expression tag	UNP Q5CPH1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

- 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>).



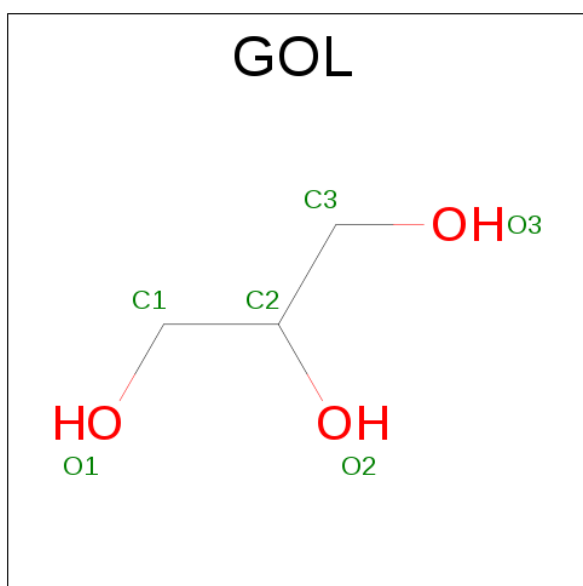
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total 44	C 21	N 7	O 14	P 2	0	0
2	B	1	Total 44	C 21	N 7	O 14	P 2	0	0
2	C	1	Total 44	C 21	N 7	O 14	P 2	0	0
2	D	1	Total 44	C 21	N 7	O 14	P 2	0	0

- Molecule 3 is D-ERITADENINE (three-letter code: DEA) (formula:  $\text{C}_9\text{H}_{11}\text{N}_5\text{O}_4$ ).



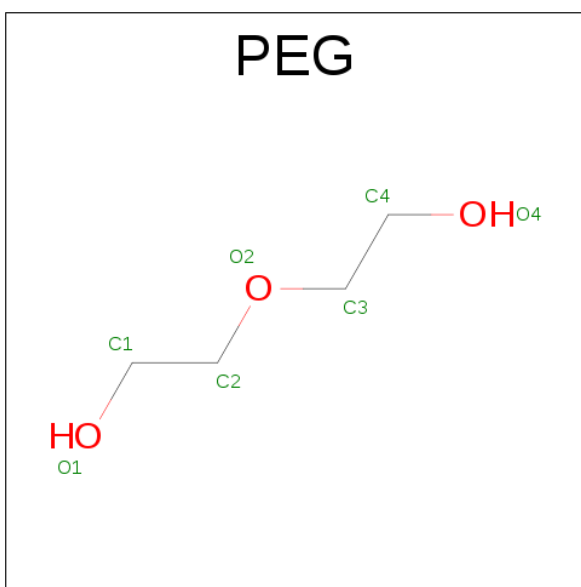
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			18	9	5	4		
3	B	1	Total	C	N	O	0	0
			18	9	5	4		
3	C	1	Total	C	N	O	0	0
			18	9	5	4		
3	D	1	Total	C	N	O	0	0
			18	9	5	4		

- 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	A	1	Total	C	O	0	0
			6	3	3		
4	A	1	Total	C	O	0	0
			6	3	3		
4	B	1	Total	C	O	0	0
			6	3	3		
4	B	1	Total	C	O	0	0
			6	3	3		
4	C	1	Total	C	O	0	0
			6	3	3		
4	D	1	Total	C	O	0	0
			6	3	3		

- Molecule 5 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>3</sub>).

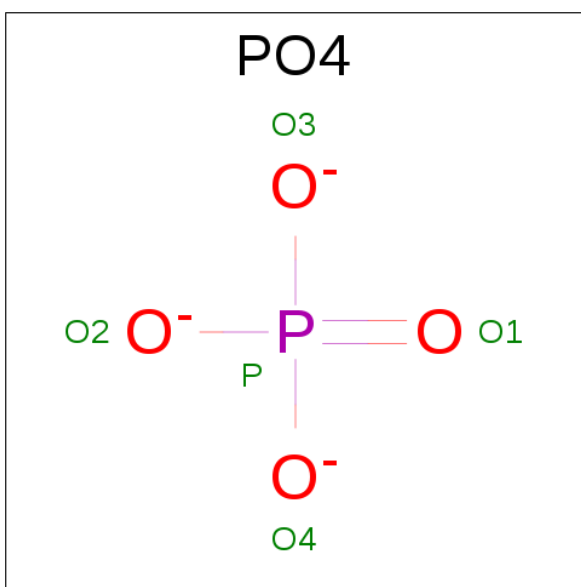


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total	C	O	0	0
			7	4	3		
5	C	1	Total	C	O	0	0
			7	4	3		
5	C	1	Total	C	O	0	0
			7	4	3		

- Molecule 6 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

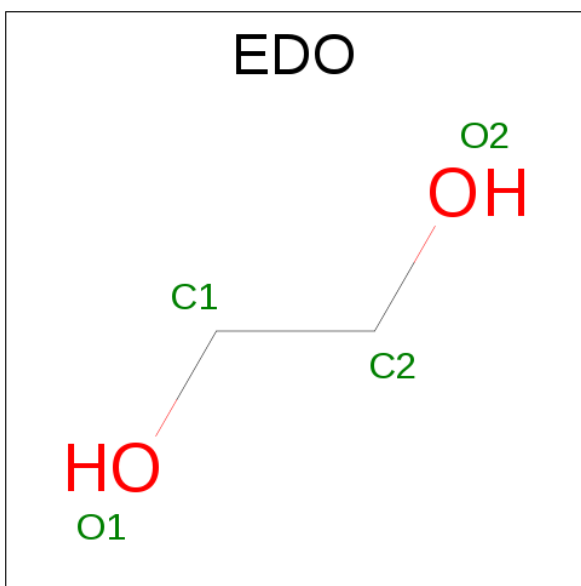
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	2	Total	Cl	0	0
			2	2		
6	C	2	Total	Cl	0	0
			2	2		

- Molecule 7 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	C	1	Total	O	P	0	0
			5	4	1		

- Molecule 8 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).

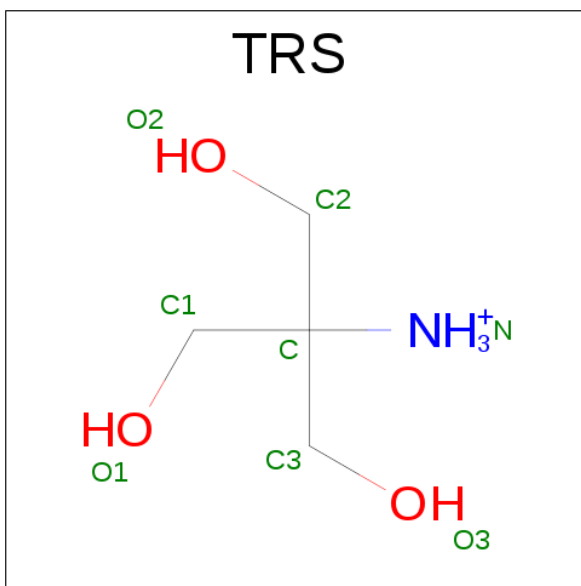


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	C	1	Total	C	O	0	0
			4	2	2		
8	D	1	Total	C	O	0	0
			4	2	2		

- Molecule 9 is 2-AMINO-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code:



TRS) (formula: C<sub>4</sub>H<sub>12</sub>NO<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
9	D	1	Total	C	N	O	0	0
			8	4	1	3		

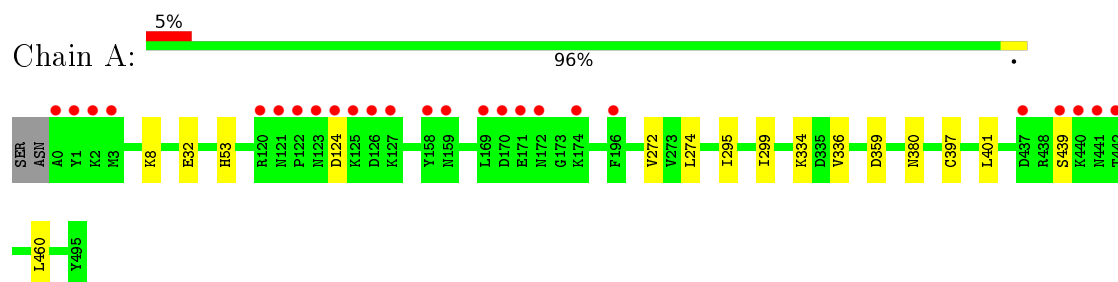
- Molecule 10 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	295	Total	O	0	13
			308	308		
10	B	299	Total	O	0	14
			313	313		
10	C	344	Total	O	0	11
			354	354		
10	D	285	Total	O	0	9
			293	293		

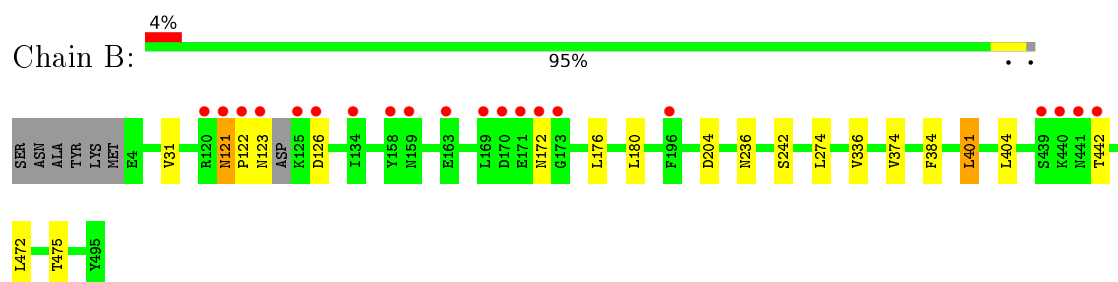
### 3 Residue-property plots

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.

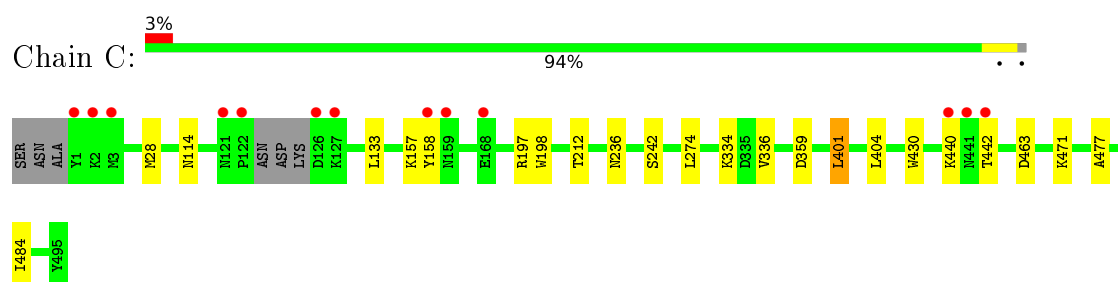
- Molecule 1: Adenosylhomocysteinase



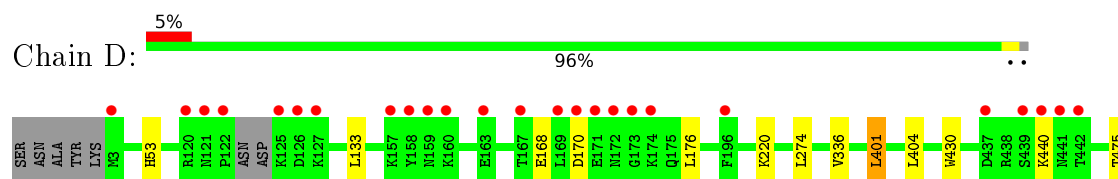
- Molecule 1: Adenosylhomocysteinase



- Molecule 1: Adenosylhomocysteinase



- Molecule 1: Adenosylhomocysteinase



1495

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	115.85Å 121.22Å 178.77Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.97 – 2.15 29.97 – 2.15	Depositor EDS
% Data completeness (in resolution range)	99.9 (29.97-2.15) 99.9 (29.97-2.15)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.30 (at 2.16Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, $R_{free}$	0.142 , 0.181 0.155 , 0.188	Depositor DCC
$R_{free}$ test set	6830 reflections (5.26%)	DCC
Wilson B-factor (Å <sup>2</sup> )	32.2	Xtriage
Anisotropy	0.221	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 49.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.012 for k,h,-l	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	17474	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.50% 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: GOL, NAD, CL, PO4, EDO, DEA, TRS, 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	0.48	0/4089	0.70	0/5510
1	B	0.49	0/4000	0.70	0/5392
1	C	0.49	0/4018	0.72	0/5414
1	D	0.48	0/4024	0.69	0/5424
All	All	0.49	0/16131	0.70	0/21740

There are no bond length outliers.

There are no bond angle outliers.

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	4024	0	4078	8	0
1	B	3937	0	3984	11	0
1	C	3954	0	4011	11	0
1	D	3961	0	4010	7	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
3	A	18	0	10	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	18	0	10	0	0
3	C	18	0	10	0	0
3	D	18	0	10	0	0
4	A	12	0	16	0	0
4	B	12	0	16	0	0
4	C	6	0	8	0	0
4	D	6	0	8	0	0
5	B	7	0	10	1	0
5	C	14	0	20	0	0
6	B	2	0	0	0	0
6	C	2	0	0	0	0
7	C	5	0	0	0	0
8	C	4	0	6	0	0
8	D	4	0	6	0	0
9	D	8	0	12	0	0
10	A	308	0	0	0	0
10	B	313	0	0	0	0
10	C	354	0	0	0	0
10	D	293	0	0	2	0
All	All	17474	0	16329	33	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 1.

All (33) 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:D:220[B]:LYS:NZ	10:D:601:HOH:O	2.30	0.65
1:B:121:ASN:O	1:B:123:ASN:N	2.38	0.55
1:D:220[B]:LYS:CE	10:D:601:HOH:O	2.56	0.53
1:B:401:LEU:HD13	1:B:404:LEU:HB2	1.90	0.53
1:D:274:LEU:HD22	1:D:336[B]:VAL:HG12	1.90	0.53
1:A:334:LYS:CE	1:B:475:THR:HG22	2.43	0.48
1:B:442:THR:HG22	1:B:442:THR:O	2.13	0.48
1:C:197:ARG:NH1	1:C:198:TRP:CZ2	2.81	0.47
1:C:28[B]:MET:CE	1:C:463[B]:ASP:HB3	2.44	0.47
1:B:176:LEU:HD22	1:B:180:LEU:HD23	1.97	0.47
1:A:274:LEU:HD22	1:A:336:VAL:HG12	1.97	0.46
1:D:133:LEU:HD11	1:D:430:TRP:HB2	1.98	0.46
1:C:133:LEU:HD11	1:C:430:TRP:HB2	1.97	0.46
1:C:401:LEU:HD13	1:C:404:LEU:HB2	1.97	0.45

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:374:VAL:HG22	1:B:384:PHE:CD2	2.52	0.44
1:C:442:THR:O	1:C:442:THR:HG22	2.17	0.44
1:D:401:LEU:HD22	1:D:404:LEU:HD12	1.99	0.43
1:B:31:VAL:HG12	5:B:505:PEG:H11	2.00	0.43
1:C:236:ASN:O	1:C:242:SER:HB3	2.19	0.43
1:C:274:LEU:HD22	1:C:336:VAL:HG12	2.00	0.43
1:C:334:LYS:CE	1:D:475:THR:HG22	2.49	0.43
1:A:32:GLU:HG2	1:A:460:LEU:HD22	2.01	0.42
1:D:170:ASP:HB3	1:D:176:LEU:HD23	2.01	0.42
1:A:334:LYS:HE3	1:B:475:THR:HG22	2.02	0.42
1:A:299:ILE:HD13	1:B:472:LEU:HD21	2.03	0.41
1:C:477:ALA:HB2	1:C:484:ILE:HD13	2.03	0.41
1:A:380:ASN:HB3	1:A:397:CYS:HB3	2.03	0.41
1:B:274:LEU:HD22	1:B:336:VAL:HG12	2.01	0.41
1:A:8[B]:LYS:O	1:A:8[B]:LYS:HG3	2.20	0.41
1:C:157[A]:LYS:HG3	1:C:158:TYR:CD2	2.55	0.41
1:C:114:ASN:OD1	1:C:197:ARG:NH1	2.51	0.41
1:B:236:ASN:O	1:B:242:SER:HB3	2.20	0.41
1:A:272:VAL:HA	1:A:295:ILE:O	2.22	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	507/498 (102%)	487 (96%)	17 (3%)	3 (1%)	30	21
1	B	495/498 (99%)	473 (96%)	18 (4%)	4 (1%)	24	15
1	C	496/498 (100%)	480 (97%)	15 (3%)	1 (0%)	52	51
1	D	497/498 (100%)	477 (96%)	17 (3%)	3 (1%)	30	21
All	All	1995/1992 (100%)	1917 (96%)	67 (3%)	11 (1%)	30	21

All (11) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	121	ASN
1	B	122	PRO
1	A	439	SER
1	B	172	ASN
1	C	440	LYS
1	D	440	LYS
1	A	124	ASP
1	B	126	ASP
1	D	53	HIS
1	A	53	HIS
1	D	168	GLU

### 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	446/435 (102%)	444 (100%)	2 (0%)	93	96
1	B	437/435 (100%)	435 (100%)	2 (0%)	92	95
1	C	438/435 (101%)	434 (99%)	4 (1%)	84	89
1	D	439/435 (101%)	438 (100%)	1 (0%)	95	97
All	All	1760/1740 (101%)	1751 (100%)	9 (0%)	92	95

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	359	ASP
1	A	401	LEU
1	B	204	ASP
1	B	401	LEU
1	C	212	THR
1	C	359	ASP
1	C	401	LEU
1	C	471	LYS
1	D	401	LEU



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

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 25 ligands modelled in this entry, 4 are monoatomic - leaving 21 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	0.84	1 (2%)	46,73,73	1.70	6 (13%)
3	DEA	A	502	-	13,19,19	1.07	1 (7%)	9,27,27	2.76	1 (11%)
4	GOL	A	503	-	5,5,5	0.19	0	5,5,5	0.22	0
4	GOL	A	504	-	5,5,5	0.22	0	5,5,5	0.28	0
2	NAD	B	501	-	42,48,48	0.81	1 (2%)	46,73,73	1.59	4 (8%)
3	DEA	B	502	-	13,19,19	1.05	1 (7%)	9,27,27	2.57	2 (22%)
4	GOL	B	503	-	5,5,5	0.32	0	5,5,5	0.19	0
4	GOL	B	504	-	5,5,5	0.23	0	5,5,5	0.19	0
5	PEG	B	505	-	6,6,6	0.67	0	5,5,5	0.55	0
2	NAD	C	501	-	42,48,48	0.89	2 (4%)	46,73,73	1.60	4 (8%)
3	DEA	C	502	-	13,19,19	1.05	1 (7%)	9,27,27	2.73	1 (11%)
4	GOL	C	503	-	5,5,5	0.29	0	5,5,5	0.22	0
5	PEG	C	504	-	6,6,6	0.48	0	5,5,5	0.28	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
7	PO4	C	505	-	4,4,4	0.59	0	6,6,6	0.23	0
5	PEG	C	508	-	6,6,6	0.46	0	5,5,5	0.34	0
8	EDO	C	509	-	3,3,3	0.48	0	2,2,2	0.32	0
2	NAD	D	501	-	42,48,48	0.86	1 (2%)	46,73,73	1.59	4 (8%)
3	DEA	D	502	-	13,19,19	0.96	1 (7%)	9,27,27	2.68	1 (11%)
4	GOL	D	503	-	5,5,5	0.29	0	5,5,5	0.15	0
9	TRS	D	504	-	7,7,7	0.48	0	9,9,9	0.40	0
8	EDO	D	505	-	3,3,3	0.48	0	2,2,2	0.22	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	DEA	A	502	-	-	0/6/12/12	0/2/2/2
4	GOL	A	503	-	-	0/4/4/4	0/0/0/0
4	GOL	A	504	-	-	0/4/4/4	0/0/0/0
2	NAD	B	501	-	-	0/22/62/62	0/5/5/5
3	DEA	B	502	-	-	0/6/12/12	0/2/2/2
4	GOL	B	503	-	-	0/4/4/4	0/0/0/0
4	GOL	B	504	-	-	0/4/4/4	0/0/0/0
5	PEG	B	505	-	-	0/4/4/4	0/0/0/0
2	NAD	C	501	-	-	0/22/62/62	0/5/5/5
3	DEA	C	502	-	-	0/6/12/12	0/2/2/2
4	GOL	C	503	-	-	0/4/4/4	0/0/0/0
5	PEG	C	504	-	-	0/4/4/4	0/0/0/0
7	PO4	C	505	-	-	0/0/0/0	0/0/0/0
5	PEG	C	508	-	-	0/4/4/4	0/0/0/0
8	EDO	C	509	-	-	0/1/1/1	0/0/0/0
2	NAD	D	501	-	-	0/22/62/62	0/5/5/5
3	DEA	D	502	-	-	0/6/12/12	0/2/2/2
4	GOL	D	503	-	-	0/4/4/4	0/0/0/0
9	TRS	D	504	-	-	0/9/9/9	0/0/0/0
8	EDO	D	505	-	-	0/1/1/1	0/0/0/0

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	501	NAD	O4D-C1D	2.17	1.44	1.41
3	B	502	DEA	C5-C4	2.39	1.45	1.40

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	502	DEA	C5-C4	2.53	1.46	1.40
3	A	502	DEA	C5-C4	2.62	1.46	1.40
2	B	501	NAD	C5A-C4A	2.66	1.46	1.40
3	C	502	DEA	C5-C4	2.73	1.46	1.40
2	C	501	NAD	C5A-C4A	2.78	1.46	1.40
2	A	501	NAD	C5A-C4A	2.91	1.47	1.40
2	D	501	NAD	C5A-C4A	3.02	1.47	1.40

All (23) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	501	NAD	N3A-C2A-N1A	-8.34	122.32	128.87
2	B	501	NAD	N3A-C2A-N1A	-7.66	122.85	128.87
2	C	501	NAD	N3A-C2A-N1A	-7.64	122.87	128.87
2	D	501	NAD	N3A-C2A-N1A	-7.57	122.93	128.87
3	A	502	DEA	N3-C2-N1	-7.50	122.98	128.87
3	C	502	DEA	N3-C2-N1	-7.42	123.05	128.87
3	D	502	DEA	N3-C2-N1	-7.38	123.07	128.87
3	B	502	DEA	N3-C2-N1	-6.90	123.45	128.87
2	A	501	NAD	C1B-N9A-C4A	-4.34	121.96	126.81
2	B	501	NAD	C1B-N9A-C4A	-3.53	122.86	126.81
2	D	501	NAD	C1B-N9A-C4A	-3.44	122.97	126.81
2	C	501	NAD	C1B-N9A-C4A	-2.81	123.67	126.81
2	A	501	NAD	O7N-C7N-N7N	-2.03	119.69	122.58
3	B	502	DEA	N6-C6-N1	2.12	122.07	118.52
2	A	501	NAD	C2A-N1A-C6A	2.15	122.60	118.77
2	A	501	NAD	C3N-C7N-N7N	2.57	120.73	117.82
2	A	501	NAD	C2D-C1D-N1N	2.74	118.90	113.53
2	B	501	NAD	C3N-C7N-N7N	2.92	121.13	117.82
2	D	501	NAD	C3N-C7N-N7N	2.95	121.15	117.82
2	B	501	NAD	C2D-C1D-N1N	3.02	119.46	113.53
2	C	501	NAD	C2D-C1D-N1N	3.17	119.75	113.53
2	C	501	NAD	C3N-C7N-N7N	3.26	121.51	117.82
2	D	501	NAD	C2D-C1D-N1N	3.40	120.20	113.53

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	505	PEG	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	496/498 (99%)	-0.26	25 (5%)	32 43	23, 34, 81, 131	0
1	B	491/498 (98%)	-0.34	20 (4%)	41 51	24, 34, 70, 144	0
1	C	492/498 (98%)	-0.36	13 (2%)	59 68	23, 32, 58, 119	0
1	D	491/498 (98%)	-0.25	25 (5%)	32 42	25, 38, 84, 123	0
All	All	1970/1992 (98%)	-0.30	83 (4%)	40 50	23, 35, 76, 144	0

All (83) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	125	LYS	8.2
1	A	122	PRO	7.5
1	A	124	ASP	6.7
1	A	125	LYS	6.5
1	B	122	PRO	6.2
1	A	441	ASN	6.2
1	B	121	ASN	5.9
1	D	441	ASN	5.5
1	D	171	GLU	5.5
1	D	125	LYS	5.4
1	D	172	ASN	5.2
1	A	440	LYS	5.1
1	D	174	LYS	4.7
1	B	172	ASN	4.7
1	C	441	ASN	4.6
1	D	122	PRO	4.5
1	D	158	TYR	4.5
1	A	123	ASN	4.5
1	B	158	TYR	4.4
1	A	126	ASP	4.3
1	C	122	PRO	4.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	440	LYS	4.1
1	A	0	ALA	4.1
1	D	126	ASP	4.0
1	A	1	TYR	3.9
1	A	171	GLU	3.7
1	C	440	LYS	3.7
1	A	121	ASN	3.7
1	C	1	TYR	3.6
1	B	441	ASN	3.6
1	D	127	LYS	3.5
1	D	169	LEU	3.5
1	A	158	TYR	3.5
1	C	127	LYS	3.5
1	A	437	ASP	3.4
1	A	127	LYS	3.4
1	D	196	PHE	3.4
1	D	159	ASN	3.4
1	D	121	ASN	3.4
1	B	120	ARG	3.4
1	D	439	SER	3.3
1	C	121	ASN	3.3
1	D	3	MET	3.2
1	A	172	ASN	3.1
1	C	3	MET	3.1
1	B	126	ASP	3.1
1	B	171	GLU	3.1
1	D	442	THR	3.1
1	A	170	ASP	3.0
1	C	442	THR	3.0
1	A	159	ASN	3.0
1	A	3	MET	2.9
1	D	120	ARG	2.9
1	D	167	THR	2.8
1	B	123	ASN	2.8
1	B	159	ASN	2.8
1	A	120	ARG	2.8
1	C	2	LYS	2.8
1	A	196	PHE	2.6
1	A	174	LYS	2.5
1	B	440	LYS	2.5
1	A	2	LYS	2.5
1	D	157	LYS	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	442	THR	2.5
1	D	170	ASP	2.4
1	D	160	LYS	2.4
1	C	158	TYR	2.4
1	A	169	LEU	2.4
1	C	168	GLU	2.4
1	B	196	PHE	2.3
1	B	439	SER	2.3
1	B	170	ASP	2.3
1	C	126	ASP	2.3
1	B	173	GLY	2.3
1	B	442	THR	2.2
1	D	163	GLU	2.2
1	A	439	SER	2.2
1	D	437	ASP	2.2
1	D	173	GLY	2.2
1	B	163	GLU	2.1
1	C	159	ASN	2.1
1	B	134	ILE	2.1
1	B	169	LEU	2.1

## 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

## 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. 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
9	TRS	D	504	8/8	0.86	0.17	12.73	54,63,66,72	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
8	EDO	C	509	4/4	0.41	0.18	7.25	73,78,78,80	0
4	GOL	B	503	6/6	0.92	0.18	4.29	39,57,60,64	0
5	PEG	C	504	7/7	0.73	0.24	3.74	67,73,78,83	0
4	GOL	D	503	6/6	0.95	0.16	3.60	44,61,64,68	0
4	GOL	A	504	6/6	0.91	0.18	3.20	61,69,71,74	0
5	PEG	C	508	7/7	0.86	0.23	3.03	71,72,82,88	0
8	EDO	D	505	4/4	0.90	0.24	2.84	75,75,76,77	0
4	GOL	A	503	6/6	0.94	0.16	2.23	58,60,67,74	0
3	DEA	B	502	18/18	0.97	0.12	0.48	25,27,35,36	0
2	NAD	B	501	44/44	0.99	0.10	0.15	22,27,29,31	0
2	NAD	D	501	44/44	0.99	0.10	0.13	25,29,34,38	0
2	NAD	C	501	44/44	0.99	0.09	-0.13	22,24,26,28	0
2	NAD	A	501	44/44	0.99	0.09	-0.37	23,26,28,29	0
3	DEA	D	502	18/18	0.98	0.09	-0.74	27,30,38,41	0
3	DEA	C	502	18/18	0.98	0.09	-0.89	23,26,32,36	0
3	DEA	A	502	18/18	0.98	0.07	-1.26	23,25,34,38	0
6	CL	B	506	1/1	0.99	0.03	-2.75	35,35,35,35	0
4	GOL	C	503	6/6	0.94	0.13	-	59,62,63,64	0
7	PO4	C	505	5/5	0.84	0.21	-	77,78,88,92	0
5	PEG	B	505	7/7	0.77	0.34	-	42,49,52,56	0
4	GOL	B	504	6/6	0.86	0.13	-	63,64,67,67	0
6	CL	C	507	1/1	0.93	0.11	-	95,95,95,95	0
6	CL	C	506	1/1	0.82	0.08	-	89,89,89,89	0
6	CL	B	507	1/1	0.83	0.08	-	80,80,80,80	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.