



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2018 – 04:00 pm GMT

PDB ID : 2B4R
Title : Crystal structure of glyceraldehyde-3-phosphate dehydrogenase from Plasmodium falciparum at 2.25 Angstrom Resolution reveals intriguing extra electron density in the active site
Authors : Robien, M.A.; Bosch, J.; Hol, W.G.J.; Structural Genomics of Pathogenic Protozoa Consortium (SGPP)
Deposited on : 2005-09-26
Resolution : 2.25 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

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

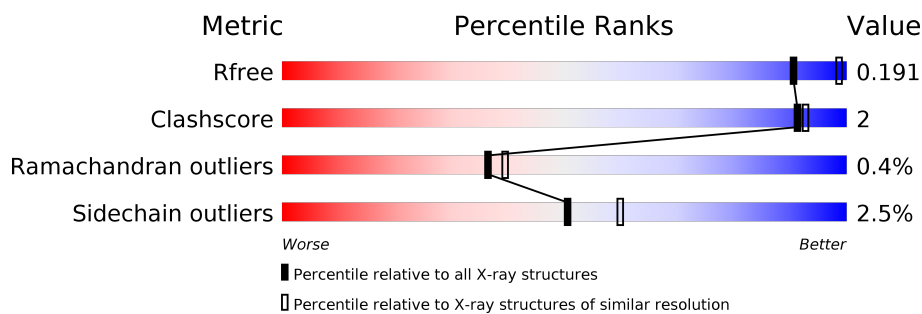
The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.7.3 (157068), CSD as539be (2018)
Xtriage (Phenix) : 1.13
EDS : trunk30967
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk30967

i

X-RAY DIFFRACTION

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}	111664	1178 (2.26-2.26)
Clashscore	122126	1286 (2.26-2.26)
Ramachandran outliers	120053	1253 (2.26-2.26)
Sidechain outliers	120020	1254 (2.26-2.26)

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 ≥ 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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	O	345	<div><div></div><div>88%</div><div>8%</div><div>••</div></div>
1	P	345	<div><div></div><div>87%</div><div>9%</div><div>••</div></div>
1	Q	345	<div><div></div><div>90%</div><div>6%</div><div>••</div></div>
1	R	345	<div><div></div><div>87%</div><div>8%</div><div>••</div></div>

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 10700 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 glyceraldehyde-3-phosphate dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	O	334	Total	C	N	O	S	58	0	0
			2551	1627	438	473	13			
1	P	334	Total	C	N	O	S	73	0	0
			2551	1627	438	473	13			
1	Q	334	Total	C	N	O	S	57	0	0
			2551	1627	438	473	13			
1	R	334	Total	C	N	O	S	47	0	0
			2551	1627	438	473	13			

There are 44 discrepancies between the modelled and reference sequences:

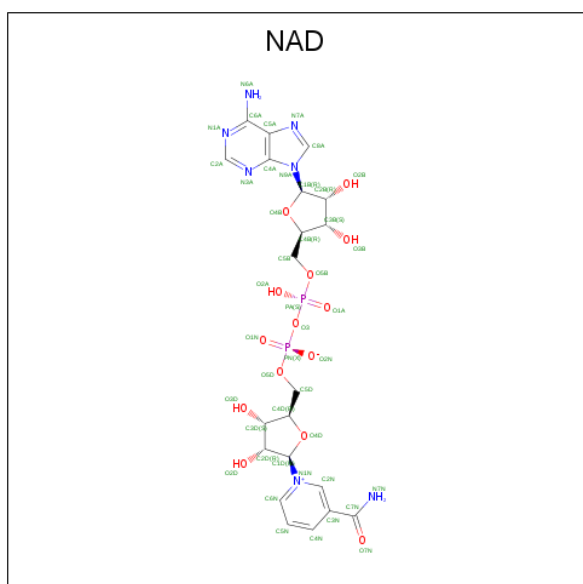
Chain	Residue	Modelled	Actual	Comment	Reference
O	-7	MET	-	CLONING ARTIFACT	UNP Q8T6B1
O	-6	ALA	-	CLONING ARTIFACT	UNP Q8T6B1
O	-5	HIS	-	EXPRESSION TAG	UNP Q8T6B1
O	-4	HIS	-	EXPRESSION TAG	UNP Q8T6B1
O	-3	HIS	-	EXPRESSION TAG	UNP Q8T6B1
O	-2	HIS	-	EXPRESSION TAG	UNP Q8T6B1
O	-1	HIS	-	EXPRESSION TAG	UNP Q8T6B1
O	0	HIS	-	EXPRESSION TAG	UNP Q8T6B1
O	3	ALA	VAL	ENGINEERED	UNP Q8T6B1
O	336	THR	ASN	ENGINEERED	UNP Q8T6B1
O	337	SER	ASN	ENGINEERED	UNP Q8T6B1
P	-7	MET	-	CLONING ARTIFACT	UNP Q8T6B1
P	-6	ALA	-	CLONING ARTIFACT	UNP Q8T6B1
P	-5	HIS	-	EXPRESSION TAG	UNP Q8T6B1
P	-4	HIS	-	EXPRESSION TAG	UNP Q8T6B1
P	-3	HIS	-	EXPRESSION TAG	UNP Q8T6B1
P	-2	HIS	-	EXPRESSION TAG	UNP Q8T6B1
P	-1	HIS	-	EXPRESSION TAG	UNP Q8T6B1
P	0	HIS	-	EXPRESSION TAG	UNP Q8T6B1
P	3	ALA	VAL	ENGINEERED	UNP Q8T6B1
P	336	THR	ASN	ENGINEERED	UNP Q8T6B1

Continued on next page...

Continued from previous page...

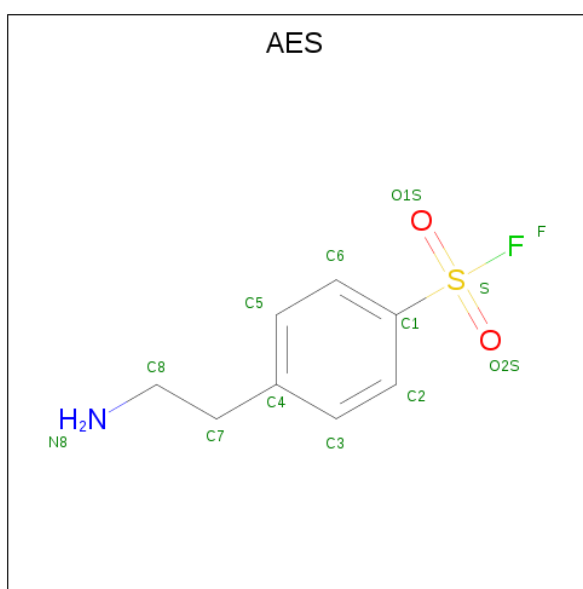
Chain	Residue	Modelled	Actual	Comment	Reference
P	337	SER	ASN	ENGINEERED	UNP Q8T6B1
Q	-7	MET	-	CLONING ARTIFACT	UNP Q8T6B1
Q	-6	ALA	-	CLONING ARTIFACT	UNP Q8T6B1
Q	-5	HIS	-	EXPRESSION TAG	UNP Q8T6B1
Q	-4	HIS	-	EXPRESSION TAG	UNP Q8T6B1
Q	-3	HIS	-	EXPRESSION TAG	UNP Q8T6B1
Q	-2	HIS	-	EXPRESSION TAG	UNP Q8T6B1
Q	-1	HIS	-	EXPRESSION TAG	UNP Q8T6B1
Q	0	HIS	-	EXPRESSION TAG	UNP Q8T6B1
Q	3	ALA	VAL	ENGINEERED	UNP Q8T6B1
Q	336	THR	ASN	ENGINEERED	UNP Q8T6B1
Q	337	SER	ASN	ENGINEERED	UNP Q8T6B1
R	-7	MET	-	CLONING ARTIFACT	UNP Q8T6B1
R	-6	ALA	-	CLONING ARTIFACT	UNP Q8T6B1
R	-5	HIS	-	EXPRESSION TAG	UNP Q8T6B1
R	-4	HIS	-	EXPRESSION TAG	UNP Q8T6B1
R	-3	HIS	-	EXPRESSION TAG	UNP Q8T6B1
R	-2	HIS	-	EXPRESSION TAG	UNP Q8T6B1
R	-1	HIS	-	EXPRESSION TAG	UNP Q8T6B1
R	0	HIS	-	EXPRESSION TAG	UNP Q8T6B1
R	3	ALA	VAL	ENGINEERED	UNP Q8T6B1
R	336	THR	ASN	ENGINEERED	UNP Q8T6B1
R	337	SER	ASN	ENGINEERED	UNP Q8T6B1

- Molecule 2 is NICOTINAMIDE-ADENINE-DINUCLEOTIDE (three-letter code: NAD) (formula: $C_{21}H_{27}N_7O_{14}P_2$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	O	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
2	P	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
2	Q	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
2	R	1	Total	C	N	O	P	0	0
			44	21	7	14	2		

- Molecule 3 is 4-(2-AMINOETHYL)BENZENESULFONYL FLUORIDE (three-letter code: AES) (formula: C₈H₁₀FNO₂S).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
3	O	1	Total 13	C 8	F 1	N 1	O 2	S 1	0	0
3	P	1	Total 13	C 8	F 1	N 1	O 2	S 1	0	0
3	R	1	Total 13	C 8	F 1	N 1	O 2	S 1	0	0

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	O	1	Total	C	O	0	0
			6	3	3		
4	P	1	Total	C	O	0	0
			6	3	3		
4	R	1	Total	C	O	0	0
			6	3	3		
4	R	1	Total	C	O	0	0
			6	3	3		

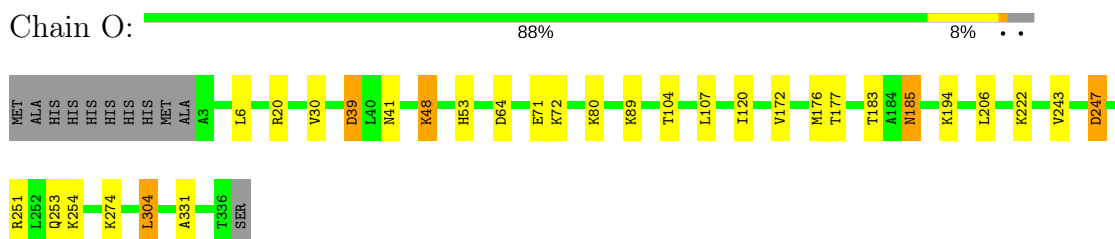
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	O	63	Total	O	0	0
			63	63		
5	P	67	Total	O	0	0
			67	67		
5	Q	53	Total	O	0	0
			53	53		
5	R	74	Total	O	0	0
			74	74		

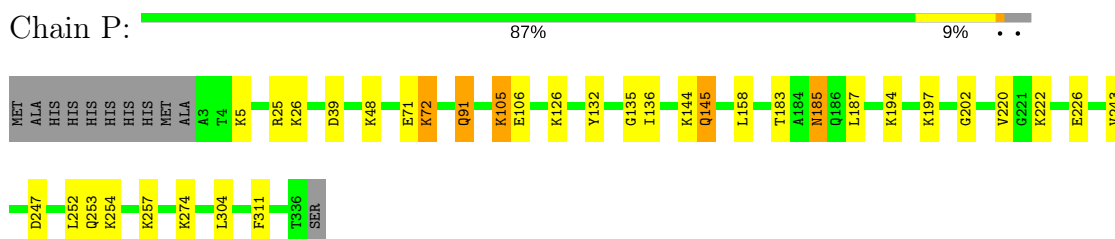
3 Residue-property plots [i](#)

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. 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. 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.

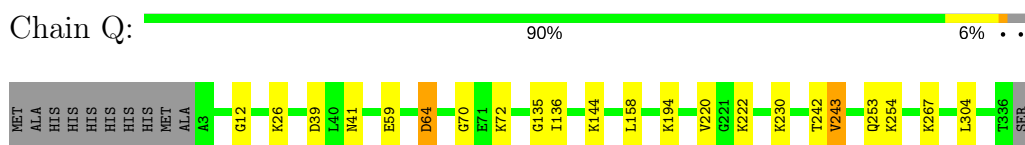
- Molecule 1: glyceraldehyde-3-phosphate dehydrogenase



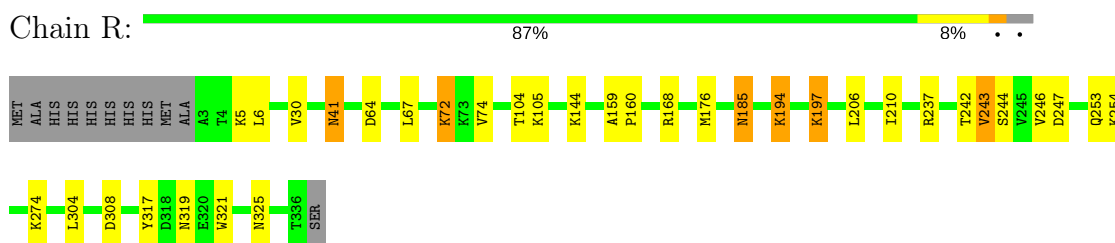
- Molecule 1: glyceraldehyde-3-phosphate dehydrogenase



- Molecule 1: glyceraldehyde-3-phosphate dehydrogenase



- Molecule 1: glyceraldehyde-3-phosphate dehydrogenase



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	69.33Å 104.58Å 90.84Å 90.00° 107.35° 90.00°	Depositor
Resolution (Å)	46.37 – 2.25 46.36 – 2.25	Depositor EDS
% Data completeness (in resolution range)	96.9 (46.37-2.25) 96.9 (46.36-2.25)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.21 (at 2.24Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.183 , 0.243 0.193 , 0.191	Depositor DCC
R_{free} test set	2861 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å ²)	46.7	Xtriage
Anisotropy	0.149	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.27 , 47.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	10700	wwPDB-VP
Average B, all atoms (Å ²)	71.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.98% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: GOL, AES, NAD

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	O	0.99	13/2603 (0.5%)	0.81	11/3528 (0.3%)
1	P	1.25	18/2603 (0.7%)	0.86	16/3528 (0.5%)
1	Q	0.92	11/2603 (0.4%)	0.92	14/3528 (0.4%)
1	R	0.77	9/2603 (0.3%)	0.74	11/3528 (0.3%)
All	All	1.00	51/10412 (0.5%)	0.84	52/14112 (0.4%)

The worst 5 of 51 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	P	71	GLU	CG-CD	-30.16	1.06	1.51
1	O	48	LYS	CD-CE	-24.67	0.89	1.51
1	P	253	GLN	CG-CD	-24.05	0.95	1.51
1	P	106	GLU	CG-CD	-22.93	1.17	1.51
1	O	247	ASP	CB-CG	-21.29	1.07	1.51

The worst 5 of 52 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	Q	41	ASN	CA-CB-CG	20.48	158.45	113.40
1	Q	72	LYS	CB-CG-CD	18.86	160.64	111.60
1	P	72	LYS	CG-CD-CE	18.08	166.15	111.90
1	O	247	ASP	CB-CG-OD1	-15.35	104.48	118.30
1	Q	41	ASN	CB-CG-OD1	-14.96	91.67	121.60

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	O	2551	0	2578	13	0
1	P	2551	0	2578	9	0
1	Q	2551	0	2578	4	0
1	R	2551	0	2578	13	0
2	O	44	0	26	0	0
2	P	44	0	26	0	0
2	Q	44	0	26	1	0
2	R	44	0	26	1	0
3	O	13	0	10	0	0
3	P	13	0	10	0	0
3	R	13	0	10	0	0
4	O	6	0	8	0	0
4	P	6	0	8	0	0
4	R	12	0	16	0	0
5	O	63	0	0	0	0
5	P	67	0	0	0	0
5	Q	53	0	0	0	0
5	R	74	0	0	0	0
All	All	10700	0	10478	37	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 37 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:P:185:ASN:H	1:P:185:ASN:HD22	1.40	0.69
1:R:185:ASN:H	1:R:185:ASN:HD22	1.38	0.69
1:P:183:THR:OG1	1:P:185:ASN:ND2	2.30	0.65
1:O:185:ASN:H	1:O:185:ASN:HD22	1.50	0.58
1:Q:158:LEU:HD23	1:Q:220:VAL:HG21	1.86	0.57

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	O	332/345 (96%)	314 (95%)	17 (5%)	1 (0%)	43	48
1	P	332/345 (96%)	314 (95%)	17 (5%)	1 (0%)	43	48
1	Q	332/345 (96%)	313 (94%)	17 (5%)	2 (1%)	27	27
1	R	332/345 (96%)	310 (93%)	21 (6%)	1 (0%)	43	48
All	All	1328/1380 (96%)	1251 (94%)	72 (5%)	5 (0%)	36	39

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	O	243	VAL
1	P	243	VAL
1	Q	243	VAL
1	R	243	VAL
1	Q	70	GLY

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	O	277/286 (97%)	270 (98%)	7 (2%)	50	60
1	P	277/286 (97%)	269 (97%)	8 (3%)	45	54
1	Q	277/286 (97%)	274 (99%)	3 (1%)	76	84
1	R	277/286 (97%)	267 (96%)	10 (4%)	38	45
All	All	1108/1144 (97%)	1080 (98%)	28 (2%)	50	60

5 of 28 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	P	197	LYS
1	Q	59	GLU
1	R	247	ASP
1	P	222	LYS
1	P	304	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	P	185	ASN
1	R	185	ASN
1	Q	139	HIS
1	P	62	HIS
1	R	138	HIS

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](#)

11 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	AES	O	601	-	11,13,13	0.53	0	16,18,18	1.88	2 (12%)
2	NAD	O	801	-	40,48,48	1.64	3 (7%)	44,73,73	1.78	3 (6%)
4	GOL	O	804	-	5,5,5	0.35	0	5,5,5	0.23	0
3	AES	P	602	-	11,13,13	0.52	0	16,18,18	1.96	2 (12%)
2	NAD	P	701	-	40,48,48	1.65	3 (7%)	44,73,73	1.74	2 (4%)
4	GOL	P	703	-	5,5,5	0.38	0	5,5,5	0.26	0
2	NAD	Q	501	-	40,48,48	1.75	3 (7%)	44,73,73	1.87	2 (4%)
4	GOL	R	503	-	5,5,5	0.33	0	5,5,5	0.31	0
3	AES	R	603	-	11,13,13	0.52	0	16,18,18	1.91	2 (12%)
2	NAD	R	901	-	40,48,48	1.64	3 (7%)	44,73,73	1.86	3 (6%)
4	GOL	R	903	-	5,5,5	0.32	0	5,5,5	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
3	AES	O	601	-	-	0/9/9/9	0/1/1/1
2	NAD	O	801	-	-	0/22/62/62	0/5/5/5
4	GOL	O	804	-	-	0/4/4/4	0/0/0/0
3	AES	P	602	-	-	0/9/9/9	0/1/1/1
2	NAD	P	701	-	-	0/22/62/62	0/5/5/5
4	GOL	P	703	-	-	0/4/4/4	0/0/0/0
2	NAD	Q	501	-	-	0/22/62/62	0/5/5/5
4	GOL	R	503	-	-	0/4/4/4	0/0/0/0
3	AES	R	603	-	-	0/9/9/9	0/1/1/1
2	NAD	R	901	-	-	0/22/62/62	0/5/5/5
4	GOL	R	903	-	-	0/4/4/4	0/0/0/0

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	R	901	NAD	C2A-N1A	2.57	1.38	1.33
2	P	701	NAD	C2A-N1A	2.76	1.39	1.33
2	O	801	NAD	C2A-N1A	2.80	1.39	1.33
2	Q	501	NAD	C2A-N1A	2.83	1.39	1.33
2	P	701	NAD	C2A-N3A	3.95	1.38	1.32

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	Q	501	NAD	N3A-C2A-N1A	-10.79	119.63	128.86
2	R	901	NAD	N3A-C2A-N1A	-10.62	119.77	128.86
2	O	801	NAD	N3A-C2A-N1A	-10.07	120.24	128.86
2	P	701	NAD	N3A-C2A-N1A	-10.04	120.27	128.86
3	P	602	AES	O2S-S-C1	-6.72	103.88	110.74

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	Q	501	NAD	1	0
2	R	901	NAD	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 ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.