wwPDB X-ray Structure Validation Summary Report (i)

Aug 6, 2023 – 05:46 AM EDT

PDB ID	:	1KAE
Title	:	L-HISTIDINOL DEHYDROGENASE (HISD) STRUCTURE COMPLEXED WITH L-HISTIDINOL (SUBSTRATE), ZINC AND NAD (COFACTOR)
Authors	:	Barbosa, J.A.R.G.; Sivaraman, J.; Li, Y.; Larocque, R.; Matte, A.; Schrag, J.D.; Cygler, M.
Deposited on	:	2001-11-01
Resolution	:	1.70 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 (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

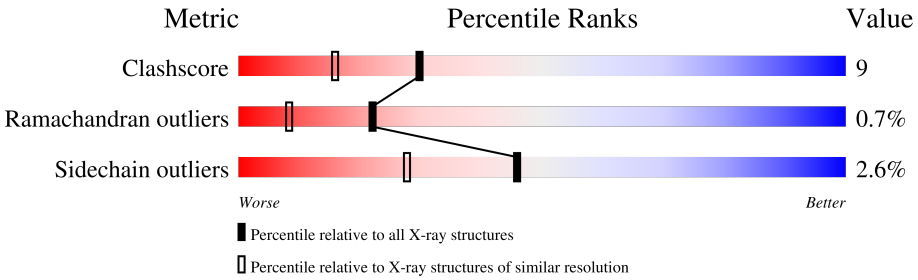
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 1.70 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))
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)

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 of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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 <=5%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	434	86% 13% .
1	B	434	82% 17% .

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
4	DTT	A	901	X	X	-	-

2 Entry composition (i)

There are 9 unique types of molecules in this entry. The entry contains 7249 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.

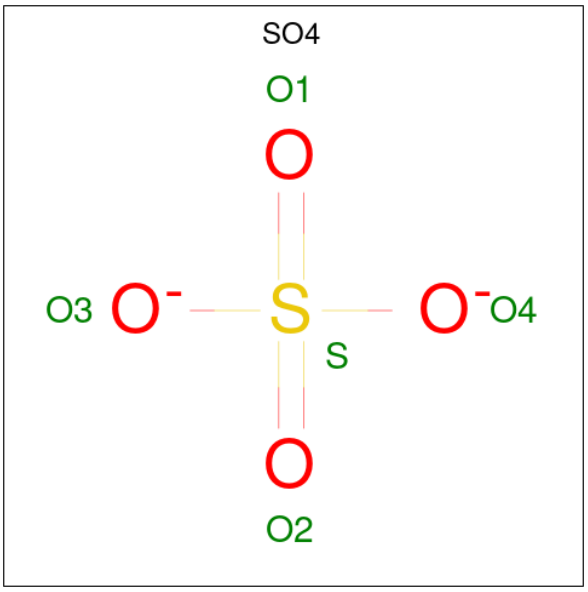
- Molecule 1 is a protein called Histidinol dehydrogenase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	434	Total	C	N	O	S	Se	31	5	0
			3247	2021	565	646	8	7			
1	B	434	Total	C	N	O	S	Se	66	2	0
			3237	2017	562	644	7	7			

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	modified residue	UNP P06988
A	15	GLU	VAL	SEE REMARK 999	UNP P06988
A	22	MSE	MET	modified residue	UNP P06988
A	88	MSE	MET	modified residue	UNP P06988
A	144	MSE	MET	modified residue	UNP P06988
A	150	SER	ARG	SEE REMARK 999	UNP P06988
A	232	MSE	MET	modified residue	UNP P06988
A	277	MSE	MET	modified residue	UNP P06988
A	313	LEU	SER	SEE REMARK 999	UNP P06988
A	390	MSE	MET	modified residue	UNP P06988
A	403	LEU	VAL	SEE REMARK 999	UNP P06988
B	1	MSE	MET	modified residue	UNP P06988
B	15	GLU	VAL	SEE REMARK 999	UNP P06988
B	22	MSE	MET	modified residue	UNP P06988
B	88	MSE	MET	modified residue	UNP P06988
B	144	MSE	MET	modified residue	UNP P06988
B	150	SER	ARG	SEE REMARK 999	UNP P06988
B	232	MSE	MET	modified residue	UNP P06988
B	277	MSE	MET	modified residue	UNP P06988
B	313	LEU	SER	SEE REMARK 999	UNP P06988
B	390	MSE	MET	modified residue	UNP P06988
B	403	LEU	VAL	SEE REMARK 999	UNP P06988

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

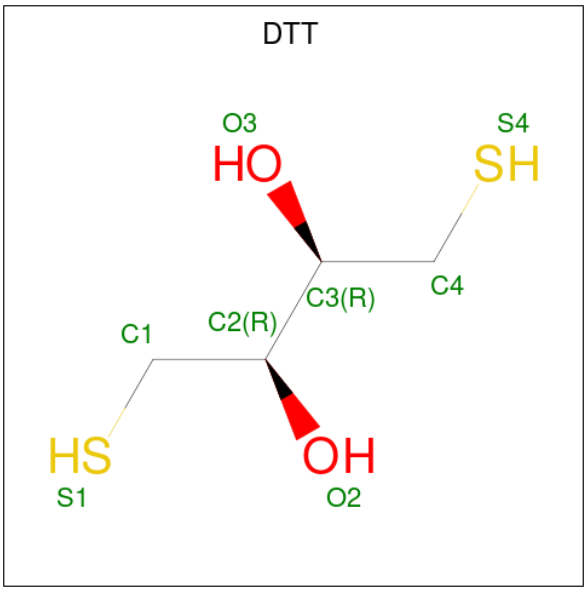


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

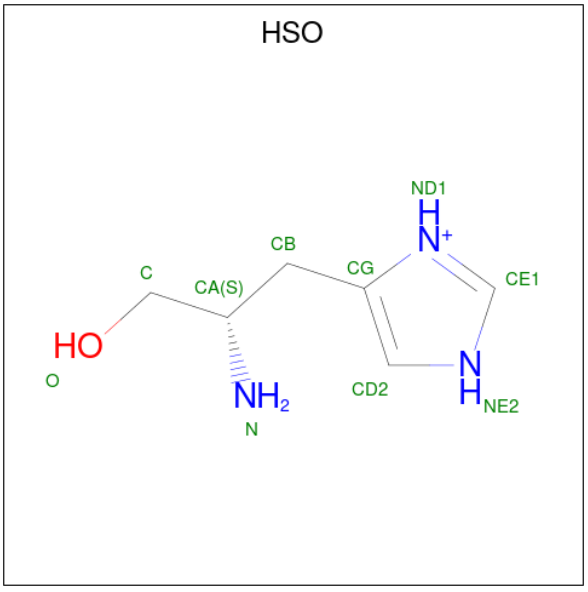
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Zn 1 1	0	0
3	B	1	Total Zn 1 1	0	0

- Molecule 4 is 2,3-DIHYDROXY-1,4-DITHIOBUTANE (three-letter code: DTT) (formula: $C_4H_{10}O_2S_2$).



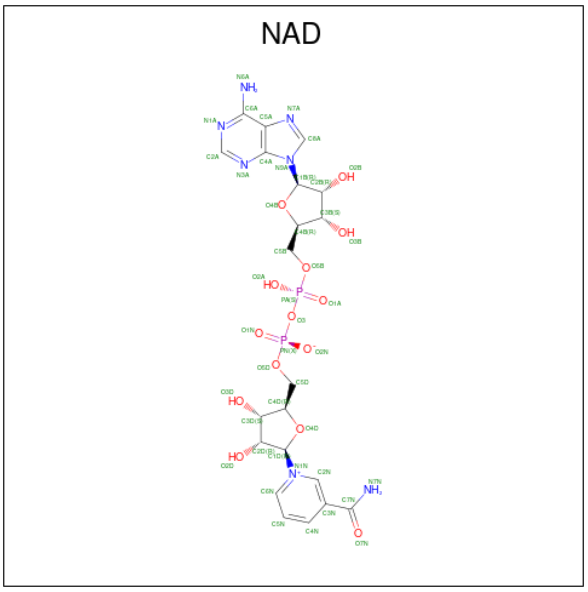
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O	S		
4	A	1	8	4	2	2	5	0

- Molecule 5 is L-histidinol (three-letter code: HSO) (formula: $C_6H_{12}N_3O$).



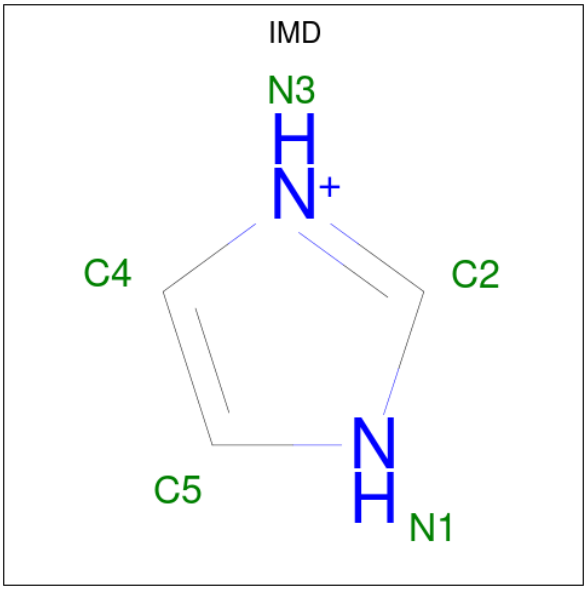
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
5	A	1	10	6	3	1	0	0

- Molecule 6 is NICOTINAMIDE-ADENINE-DINUCLEOTIDE (three-letter code: NAD) (formula: $C_{21}H_{27}N_7O_{14}P_2$).



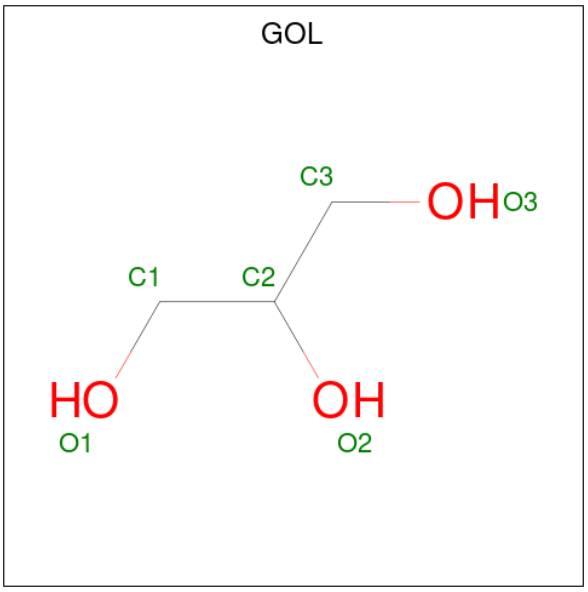
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	A	1	Total	C	N	O	P	17	0
			44	21	7	14	2		
6	B	1	Total	C	N	O	P	35	0
			44	21	7	14	2		

- Molecule 7 is IMIDAZOLE (three-letter code: IMD) (formula: $C_3H_5N_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	B	1	Total	C	N		0	0
			5	3	2			

- Molecule 8 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	B	1	Total C O 6 3 3	0	0

- Molecule 9 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	327	Total O 327 327	0	0
9	B	299	Total O 299 299	0	0

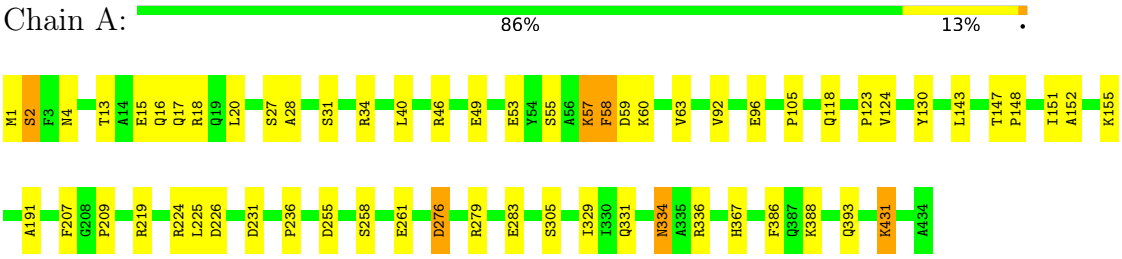
3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

Note EDS was not executed.

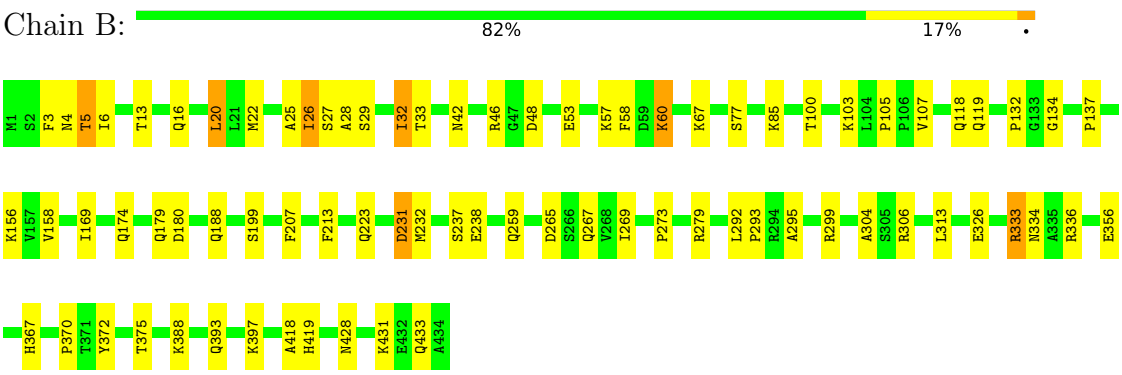
• Molecule 1: Histidinol dehydrogenase

Chain A:



• Molecule 1: Histidinol dehydrogenase

Chain B:



4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	54.93Å 107.94Å 156.71Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.50 – 1.70	Depositor
% Data completeness (in resolution range)	87.4 (38.50-1.70)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.05	Depositor
Refinement program	CNS	Depositor
R, R_{free}	0.213 , 0.241	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	7249	wwPDB-VP
Average B, all atoms (Å^2)	26.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, ZN, DTT, SO4, IMD, HSO, 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	A	0.69	0/3314	0.80	1/4496 (0.0%)
1	B	0.72	1/3290 (0.0%)	0.83	5/4466 (0.1%)
All	All	0.70	1/6604 (0.0%)	0.81	6/8962 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	B	0	1
All	All	0	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	27	SER	CA-C	6.16	1.69	1.52

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	26	ILE	O-C-N	-8.24	109.51	122.70
1	B	60	LYS	C-N-CA	-7.82	102.16	121.70
1	B	231	ASP	CB-CG-OD1	5.79	123.51	118.30
1	A	124	VAL	N-CA-C	-5.62	95.82	111.00
1	B	60	LYS	N-CA-C	-5.20	96.96	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	130	TYR	Sidechain
1	B	26	ILE	Mainchain

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3247	0	3265	53	0
1	B	3237	0	3256	81	0
2	A	10	0	0	0	0
2	B	10	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	8	0	8	0	0
5	A	10	0	11	0	0
6	A	44	0	26	0	0
6	B	44	0	26	1	0
7	B	5	0	4	0	0
8	B	6	0	8	2	0
9	A	327	0	0	4	0
9	B	299	0	0	9	0
All	All	7249	0	6604	121	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 9.

The worst 5 of 121 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:431:LYS:HE3	1:A:431:LYS:O	1.56	1.04
1:B:4:ASN:ND2	1:B:5:THR:H	1.65	0.94
1:A:118:GLN:HE22	1:B:336:ARG:HH22	1.18	0.92
1:A:336:ARG:HH12	1:B:118:GLN:HE22	1.18	0.87
1:A:336:ARG:HH12	1:B:118:GLN:NE2	1.74	0.85

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	437/434 (101%)	419 (96%)	14 (3%)	4 (1%)	17 5
1	B	434/434 (100%)	413 (95%)	19 (4%)	2 (0%)	29 13
All	All	871/868 (100%)	832 (96%)	33 (4%)	6 (1%)	22 8

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	2	SER
1	A	28	ALA
1	A	57	LYS
1	B	5	THR
1	B	60	LYS

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	352/340 (104%)	341 (97%)	11 (3%)	40 21
1	B	349/340 (103%)	341 (98%)	8 (2%)	50 33
All	All	701/680 (103%)	682 (97%)	19 (3%)	46 26

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	32	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	388	LYS
1	B	393	GLN
1	B	367	HIS
1	A	388	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	223	GLN
1	B	259	GLN
1	B	433	GLN
1	B	387	GLN
1	B	42	ASN

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 12 ligands modelled in this entry, 2 are monoatomic - leaving 10 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	SO4	A	702	-	4,4,4	0.29	0	6,6,6	0.10	0
5	HSO	A	1001	3	6,10,10	0.77	0	4,12,12	1.24	0
8	GOL	B	801	-	5,5,5	0.33	0	5,5,5	0.24	0
2	SO4	B	701	-	4,4,4	0.23	0	6,6,6	0.24	0
4	DTT	A	901	1	7,7,7	6.46	5 (71%)	4,8,8	5.52	4 (100%)
2	SO4	A	704	-	4,4,4	0.18	0	6,6,6	0.22	0
2	SO4	B	703	-	4,4,4	0.37	0	6,6,6	0.12	0
6	NAD	A	1201	-	42,48,48	2.32	9 (21%)	50,73,73	1.86	9 (18%)
6	NAD	B	1202	-	42,48,48	2.42	9 (21%)	50,73,73	2.10	9 (18%)
7	IMD	B	902	3	3,5,5	0.19	0	4,5,5	0.70	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
5	HSO	A	1001	3	-	0/6/6/6	0/1/1/1
8	GOL	B	801	-	-	0/4/4/4	-
4	DTT	A	901	1	2/2/2/2	2/8/8/8	-
6	NAD	A	1201	-	-	5/26/62/62	0/5/5/5
6	NAD	B	1202	-	-	10/26/62/62	0/5/5/5
7	IMD	B	902	3	-	-	0/1/1/1

The worst 5 of 23 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	901	DTT	O2-C2	-15.52	1.10	1.43
6	B	1202	NAD	C2N-N1N	7.88	1.44	1.35
6	B	1202	NAD	C2N-C3N	7.39	1.50	1.39
6	A	1201	NAD	C2N-C3N	7.33	1.50	1.39
6	A	1201	NAD	C2N-N1N	7.25	1.43	1.35

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	901	DTT	C3-C4-S4	-8.41	90.02	114.47
6	B	1202	NAD	C3N-C7N-N7N	6.44	125.47	117.75
6	B	1202	NAD	O4B-C1B-C2B	-6.32	97.69	106.93
6	A	1201	NAD	C3N-C7N-N7N	6.31	125.32	117.75
6	A	1201	NAD	O7N-C7N-C3N	-5.83	112.66	119.63

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
4	A	901	DTT	C2
4	A	901	DTT	C3

5 of 17 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	901	DTT	C2-C3-C4-S4
4	A	901	DTT	O3-C3-C4-S4
6	A	1201	NAD	O4D-C1D-N1N-C2N
6	A	1201	NAD	C2D-C1D-N1N-C2N
6	B	1202	NAD	C5D-O5D-PN-O1N

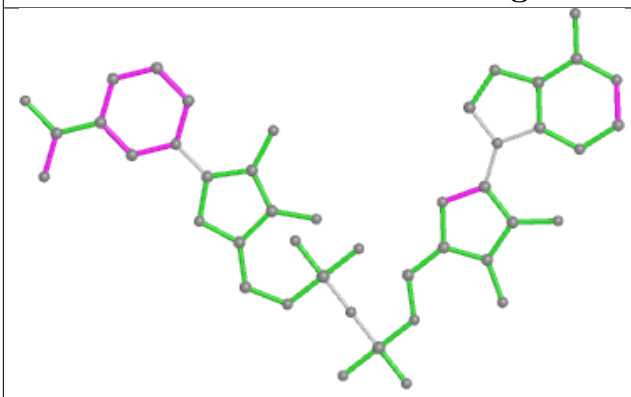
There are no ring outliers.

2 monomers are involved in 3 short contacts:

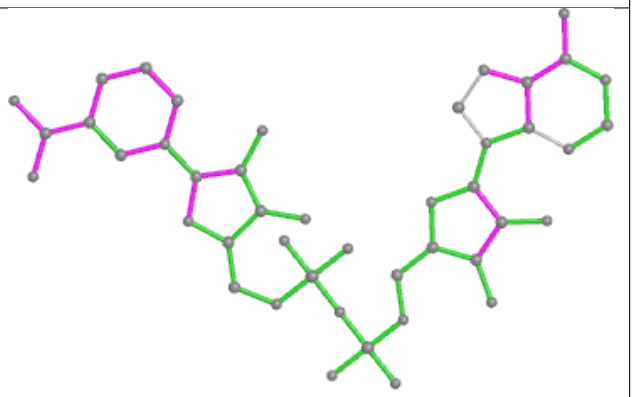
Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	B	801	GOL	2	0
6	B	1202	NAD	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

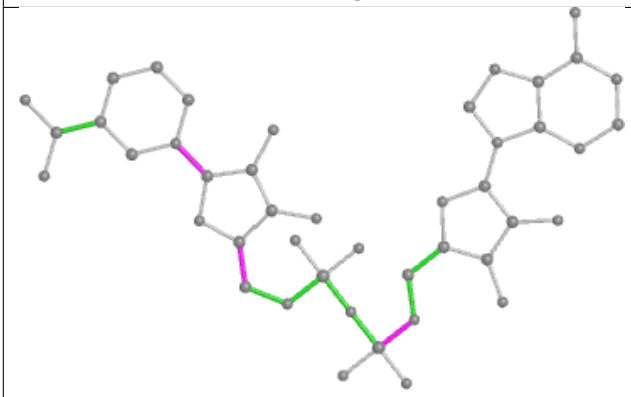
Ligand NAD A 1201



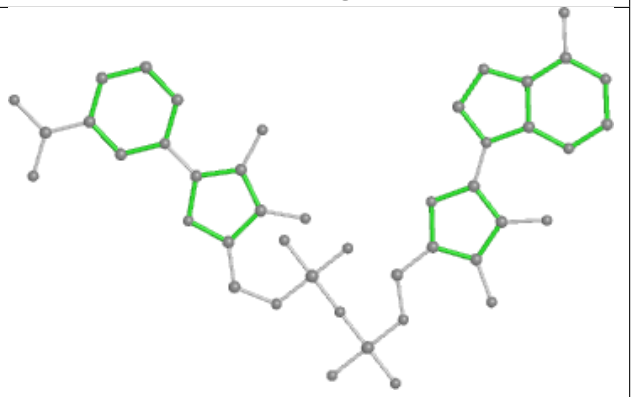
Bond lengths



Bond angles

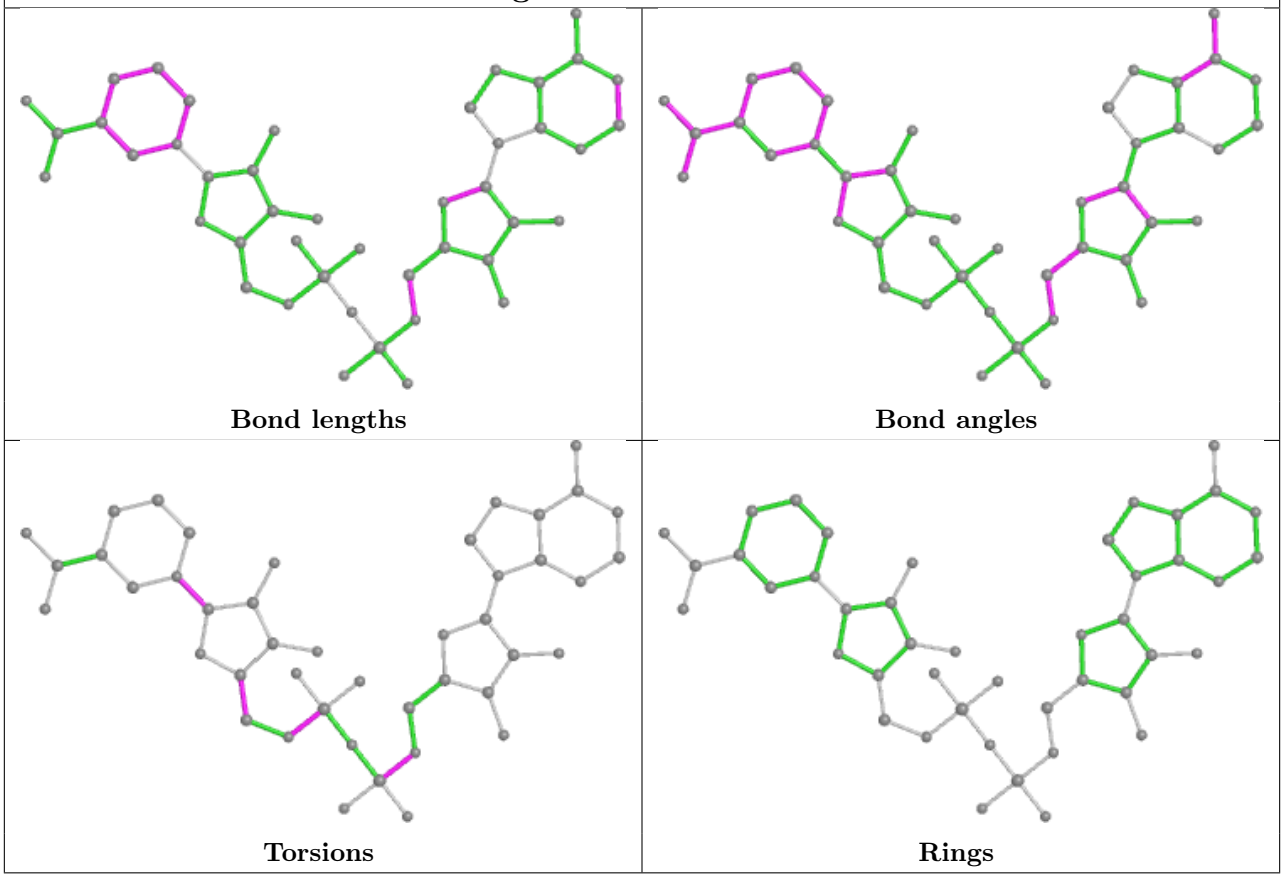


Torsions



Rings

Ligand NAD B 1202



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

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.