wwPDB X-ray Structure Validation Summary Report (i)

Oct 24, 2023 – 08:51 PM EDT

PDB ID	:	3APT
Title	:	properties and crystal structure of methylenetetrahydrofolate reductase from Thermus thermophilus HB8
Authors	:	Yamada, K.
Deposited on	:	2010-10-20
Resolution	:	1.85 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 (i)) were used in the production of this report:

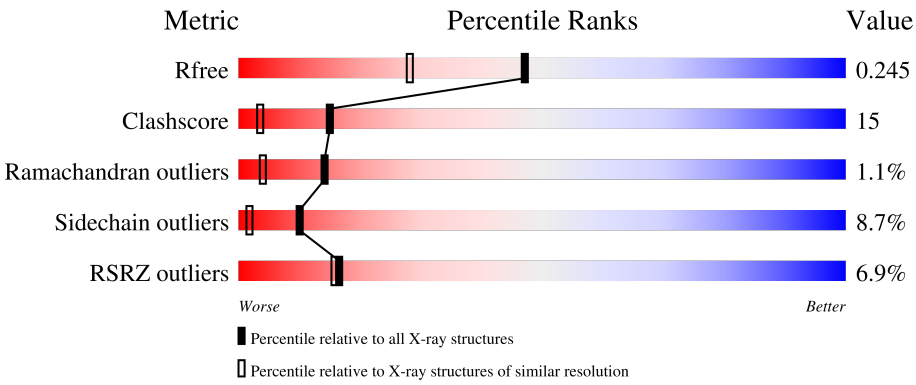
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 1.85 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}	130704	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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% 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	310	6% 72% 18% • • 6%
1	B	310	7% 66% 17% 5% • 11%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 4962 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 Methylenetetrahydrofolate reductase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	292	2307	1472	423	407	5	0	0	0
			Total	C	N	O	S			
1	B	275	2175	1391	395	385	4	0	0	0

There are 28 discrepancies between the modelled and reference sequences:

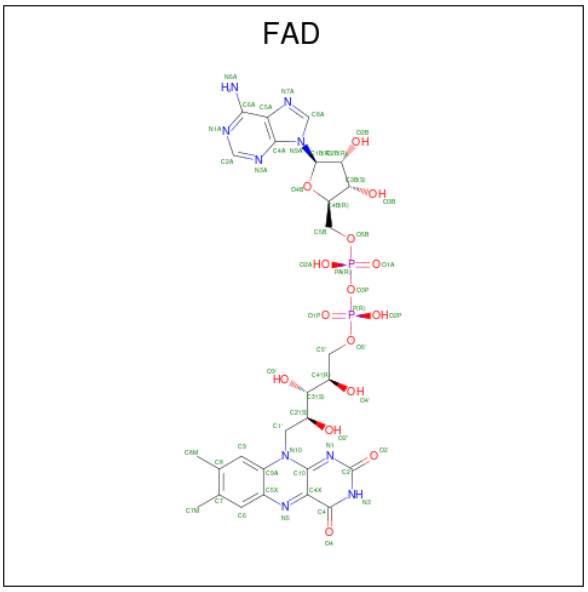
Chain	Residue	Modelled	Actual	Comment	Reference
A	297	ALA	-	expression tag	UNP Q5SLG6
A	298	LYS	-	expression tag	UNP Q5SLG6
A	299	LEU	-	expression tag	UNP Q5SLG6
A	300	ALA	-	expression tag	UNP Q5SLG6
A	301	ALA	-	expression tag	UNP Q5SLG6
A	302	ALA	-	expression tag	UNP Q5SLG6
A	303	LEU	-	expression tag	UNP Q5SLG6
A	304	GLU	-	expression tag	UNP Q5SLG6
A	305	HIS	-	expression tag	UNP Q5SLG6
A	306	HIS	-	expression tag	UNP Q5SLG6
A	307	HIS	-	expression tag	UNP Q5SLG6
A	308	HIS	-	expression tag	UNP Q5SLG6
A	309	HIS	-	expression tag	UNP Q5SLG6
A	310	HIS	-	expression tag	UNP Q5SLG6
B	297	ALA	-	expression tag	UNP Q5SLG6
B	298	LYS	-	expression tag	UNP Q5SLG6
B	299	LEU	-	expression tag	UNP Q5SLG6
B	300	ALA	-	expression tag	UNP Q5SLG6
B	301	ALA	-	expression tag	UNP Q5SLG6
B	302	ALA	-	expression tag	UNP Q5SLG6
B	303	LEU	-	expression tag	UNP Q5SLG6
B	304	GLU	-	expression tag	UNP Q5SLG6
B	305	HIS	-	expression tag	UNP Q5SLG6
B	306	HIS	-	expression tag	UNP Q5SLG6
B	307	HIS	-	expression tag	UNP Q5SLG6

Continued on next page...

Continued from previous page...

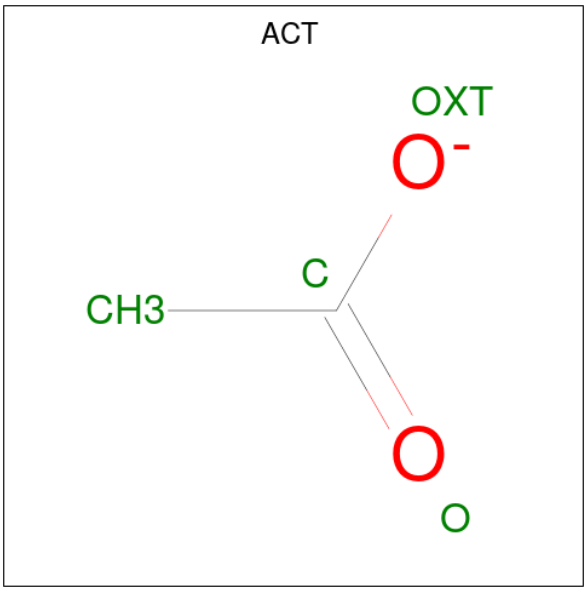
Chain	Residue	Modelled	Actual	Comment	Reference
B	308	HIS	-	expression tag	UNP Q5SLG6
B	309	HIS	-	expression tag	UNP Q5SLG6
B	310	HIS	-	expression tag	UNP Q5SLG6

- Molecule 2 is FLAVIN-ADENINE DINUCLEOTIDE (three-letter code: FAD) (formula: $C_{27}H_{33}N_9O_{15}P_2$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
2	A	1	53	27	9	15	2	0	0

- Molecule 3 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
3	B	1	4	2	2	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	244	244	244	0	0
			Total	O		
4	B	179	179	179	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 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: Methylenetetrahydrofolate reductase

Chain A: 6% 72% 18% • • 6%

M1 R4 D5 L6 L7 K8 A9 R10 R11 E18 F19 F20 P21 P22 K23 R33 T34 L35 E36 K39 V46 S47 T48 T49 Y50 G51 A52 M53 G54 S55 T56 R57 E58 R59 S84 R85 K86 E87 R94 F95 V96 E97 R107 P111 R112 G113 F117 H120 P121

R125 Y126 A127 V131 Y149 H153 E157 D162 F189 R194 A195 R196 R197 P203 P206 M209 R215 Q216 L217 R218 I228 L232 Q240 D241 D242 P243 K244 A245 I249 V258 E267 G268 Y272 T273 L274 N275 K276 A292 SER GLY

GLN PRO ALA LYS LEU ALA ALA ALA LEU GLU HIS HIS HIS HIS HIS HIS

• Molecule 1: Methylenetetrahydrofolate reductase

Chain B: 7% 66% 17% 5% • 11%

M1 L7 K8 A9 R10 R11 L14 F20 P21 P22 K23 R33 E37 L38 K39 A40 I48 T49 Y50 GLY ALA MET GLY SER T56 R57 E58 L70 H77 L78 T79 G82 Q83 S84 R85 K86 E87 V91 R94 E97 E101 R107 G108 D109 R112

G113 GLU ARG VAL PHE ARG PRO HIS PRO GLU G123 F124 R125 P154 E155 S156 E157 S158 L159 E160 A161 D162 L163 R164 Q180 Y188 F191 L192 E193 R194 I204 V211 T212 R215 Q216 L217 R218 R219 F220 T221 E222 V223 C224 L236 E237 R238 H239 Q240 D241 D242 P243 K244

A245 V246 G250 V251 V255 R256 E267 L274 N275 K276 R281 E285 L289 ARG PRO ALA SER GLY GLN PRO ALA LYS LEU ALA ALA ALA LEU GLU HIS HIS HIS HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	43.92Å 89.67Å 160.72Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.95 – 1.85 19.95 – 1.85	Depositor EDS
% Data completeness (in resolution range)	99.6 (19.95-1.85) 99.6 (19.95-1.85)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	5.02 (at 1.85Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, R_{free}	0.203 , 0.252 0.198 , 0.245	Depositor DCC
R_{free} test set	2796 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å^2)	32.4	Xtriage
Anisotropy	0.020	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 50.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	4962	wwPDB-VP
Average B, all atoms (Å^2)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.64% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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: FAD, ACT

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	1.16	7/2362 (0.3%)	1.18	12/3192 (0.4%)
1	B	1.04	0/2223	1.05	8/3001 (0.3%)
All	All	1.10	7/4585 (0.2%)	1.12	20/6193 (0.3%)

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	2	0
All	All	2	1

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	194	ARG	CD-NE	-6.32	1.35	1.46
1	A	267	GLU	CG-CD	6.17	1.61	1.51
1	A	46	VAL	CB-CG2	5.84	1.65	1.52
1	A	189	PHE	CD2-CE2	5.59	1.50	1.39
1	A	131	VAL	CB-CG1	5.45	1.64	1.52

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	194	ARG	NE-CZ-NH2	-20.24	110.18	120.30
1	A	107	ARG	NE-CZ-NH1	12.04	126.32	120.30
1	B	33	ARG	NE-CZ-NH1	10.67	125.63	120.30
1	A	194	ARG	NE-CZ-NH1	9.80	125.20	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	194	ARG	CG-CD-NE	-7.97	95.07	111.80

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	B	48	ILE	CB
1	B	204	ILE	CB

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	51	GLY	Peptide

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	2307	0	2324	67	0
1	B	2175	0	2195	70	0
2	A	53	0	31	1	0
3	B	4	0	3	0	0
4	A	244	0	0	8	0
4	B	179	0	0	11	0
All	All	4962	0	4553	137	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 15.

The worst 5 of 137 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:11:ARG:HH21	1:A:11:ARG:CG	1.69	1.04
1:A:33:ARG:HH11	1:A:33:ARG:HG3	1.26	1.01
1:A:11:ARG:HG2	1:A:11:ARG:NH2	1.57	0.99
1:B:86:LYS:O	1:B:86:LYS:HD3	1.62	0.97
1:A:55:SER:HB2	1:A:56:THR:C	1.92	0.89

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	290/310 (94%)	280 (97%)	8 (3%)	2 (1%)	22 9
1	B	269/310 (87%)	258 (96%)	7 (3%)	4 (2%)	10 2
All	All	559/620 (90%)	538 (96%)	15 (3%)	6 (1%)	14 4

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	158	SER
1	B	242	ASP
1	B	276	LYS
1	A	55	SER
1	B	240	GLN

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	236/249 (95%)	223 (94%)	13 (6%)	21 7
1	B	223/249 (90%)	196 (88%)	27 (12%)	5 0
All	All	459/498 (92%)	419 (91%)	40 (9%)	10 2

5 of 40 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	155	GLU
1	B	241	ASP
1	B	204	ILE
1	B	219	ARG
1	B	246	VAL

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

Mol	Chain	Res	Type
1	B	65	GLN
1	B	68	GLN
1	B	240	GLN
1	B	77	HIS
1	A	275	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)

2 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	ACT	B	311	-	3,3,3	1.74	1 (33%)	3,3,3	1.51	1 (33%)
2	FAD	A	311	-	53,58,58	1.51	4 (7%)	68,89,89	1.82	18 (26%)

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	FAD	A	311	-	-	0/30/50/50	0/6/6/6

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	311	FAD	C2A-N3A	5.71	1.41	1.32
2	A	311	FAD	C4X-N5	5.41	1.41	1.30
2	A	311	FAD	C2A-N1A	3.27	1.40	1.33
2	A	311	FAD	C10-N1	2.70	1.38	1.33
3	B	311	ACT	O-C	2.30	1.32	1.22

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	311	FAD	N3A-C2A-N1A	-5.96	119.36	128.68
2	A	311	FAD	C8M-C8-C9	-3.85	112.37	119.49
2	A	311	FAD	C7M-C7-C6	-3.53	112.97	119.49
2	A	311	FAD	C1B-N9A-C4A	-2.89	121.56	126.64
2	A	311	FAD	C4X-C10-N10	2.80	120.57	116.48

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

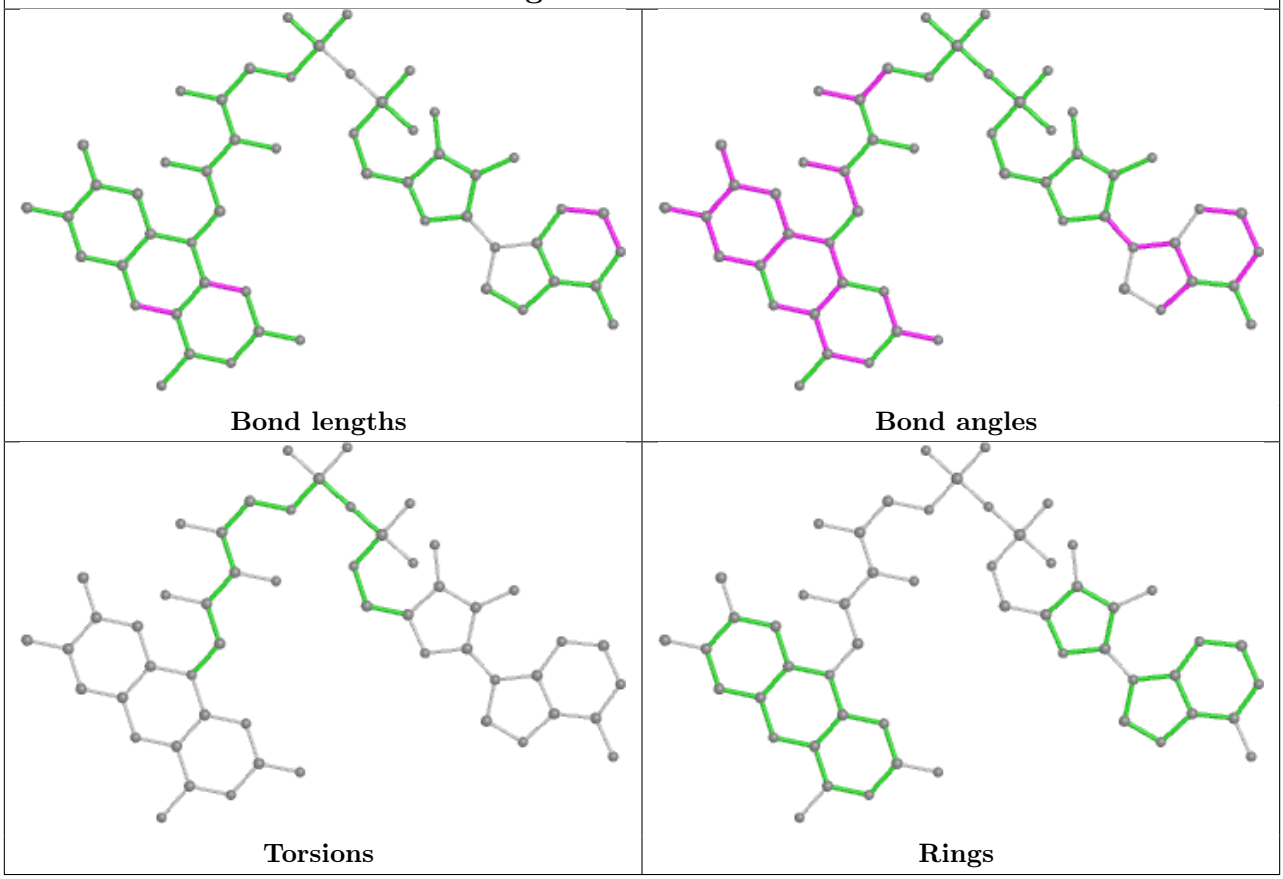
1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	311	FAD	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 FAD A 311



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 (i)

6.1 Protein, DNA and RNA chains (i)

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, 95th 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^2)	Q<0.9
1	A	292/310 (94%)	0.38	18 (6%)	20	20	17, 30, 54, 76	1 (0%)
1	B	275/310 (88%)	0.53	21 (7%)	13	13	18, 35, 58, 73	0
All	All	567/620 (91%)	0.45	39 (6%)	16	16	17, 33, 57, 76	1 (0%)

The worst 5 of 39 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	220	PHE	9.0
1	A	52	ALA	8.4
1	A	55	SER	7.5
1	B	40	ALA	5.8
1	A	53	MET	5.5

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 monosaccharides in this entry.

6.4 Ligands (i)

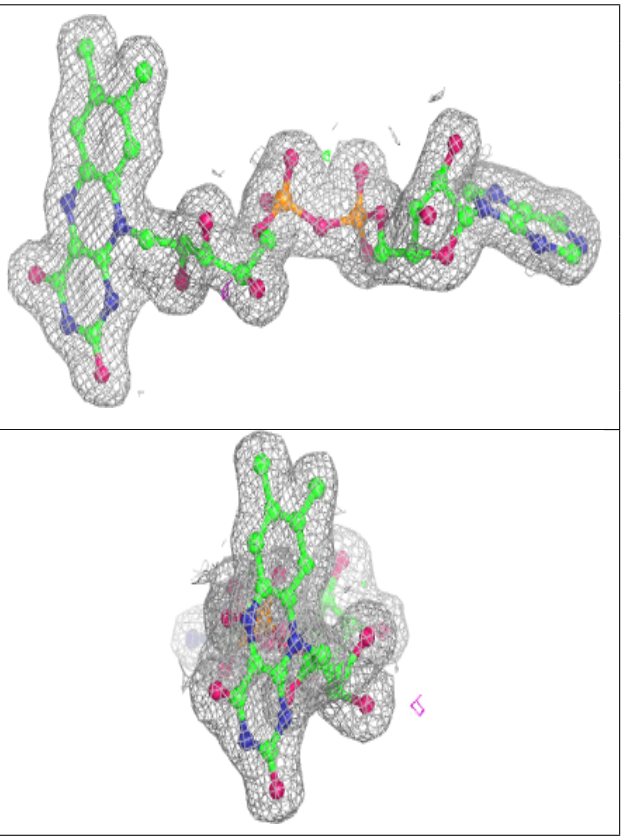
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	ACT	B	311	4/4	0.79	0.18	27,30,32,32	0
2	FAD	A	311	53/53	0.96	0.09	21,25,34,34	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around FAD A 311:

2mF$_o$-DF$_c$ (at 0.7 rmsd) in gray
mF$_o$-DF$_c$ (at 3 rmsd) in purple (negative) and green (positive)



6.5 Other polymers (i)

There are no such residues in this entry.