wwPDB X-ray Structure Validation Summary Report (i)

Dec 17, 2023 – 07:01 am GMT

PDB ID	:	4CZZ
Title	:	Histone demethylase LSD1(KDM1A)-CoREST3 Complex
Authors	:	Barrios, A.P.; Gomez, A.V.; Saez, J.E.; Ciossani, G.; Toffolo, E.; Battaglioli, E.; Mattevi, A.; Andres, M.E.
Deposited on	:	2014-04-23
Resolution	:	3.00 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:

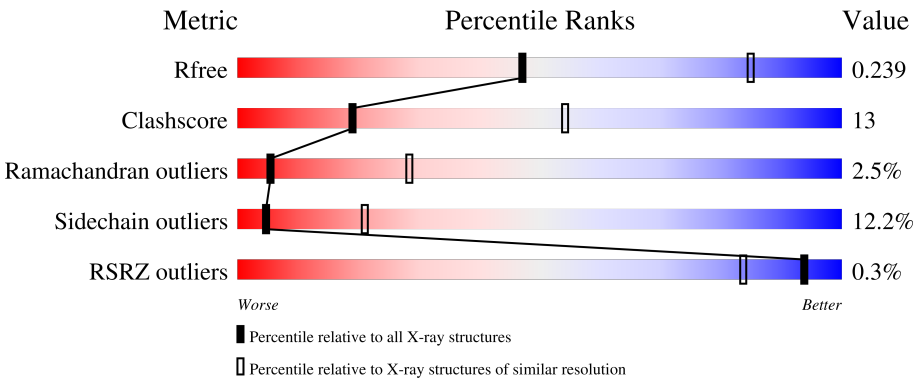
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
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 3.00 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	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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	872	50% 22% • 24%
2	B	553	14% 9% • 76%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 6339 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 LYSINE-SPECIFIC HISTONE DEMETHYLASE 1A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	666	5217	3324	906	967	20	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

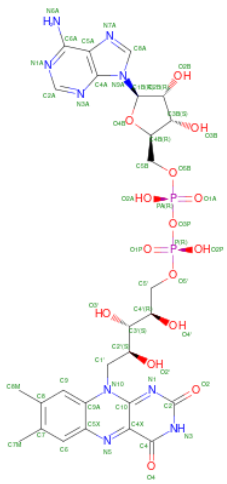
Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	ASP	deletion	UNP O60341
A	?	-	THR	deletion	UNP O60341
A	?	-	VAL	deletion	UNP O60341
A	?	-	LYS	deletion	UNP O60341

- Molecule 2 is a protein called REST COREPRESSOR 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	133	1069	670	192	202	5	0	0	0

- Molecule 3 is FLAVIN-ADENINE DINUCLEOTIDE (three-letter code: FAD) (formula: $C_{27}H_{33}N_9O_{15}P_2$).

FAD



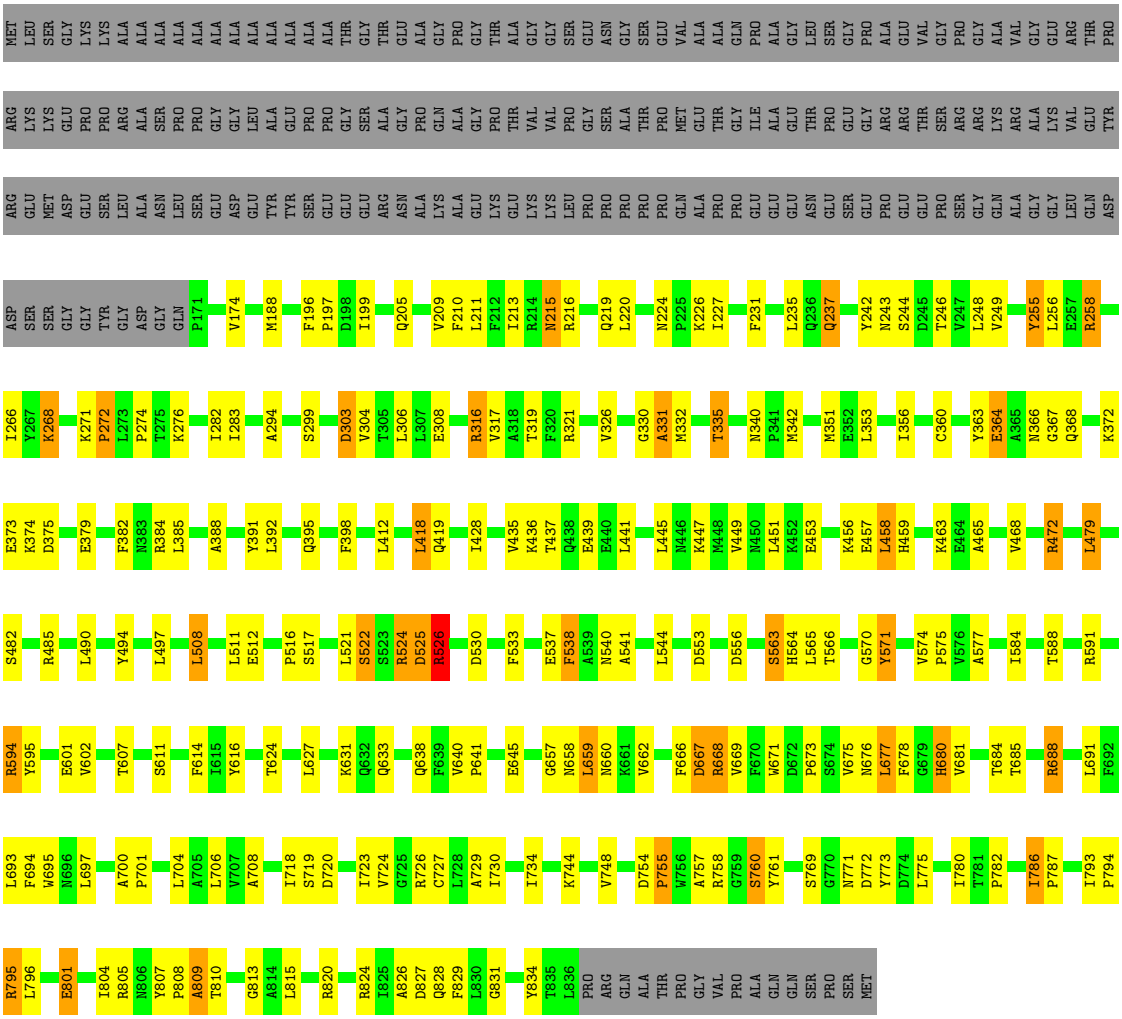
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
3	A	1	53	27	9	15	2	0	0

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.

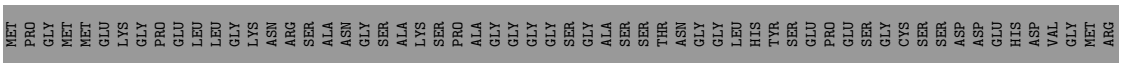
- Molecule 1: LYSINE-SPECIFIC HISTONE DEMETHYLASE 1A

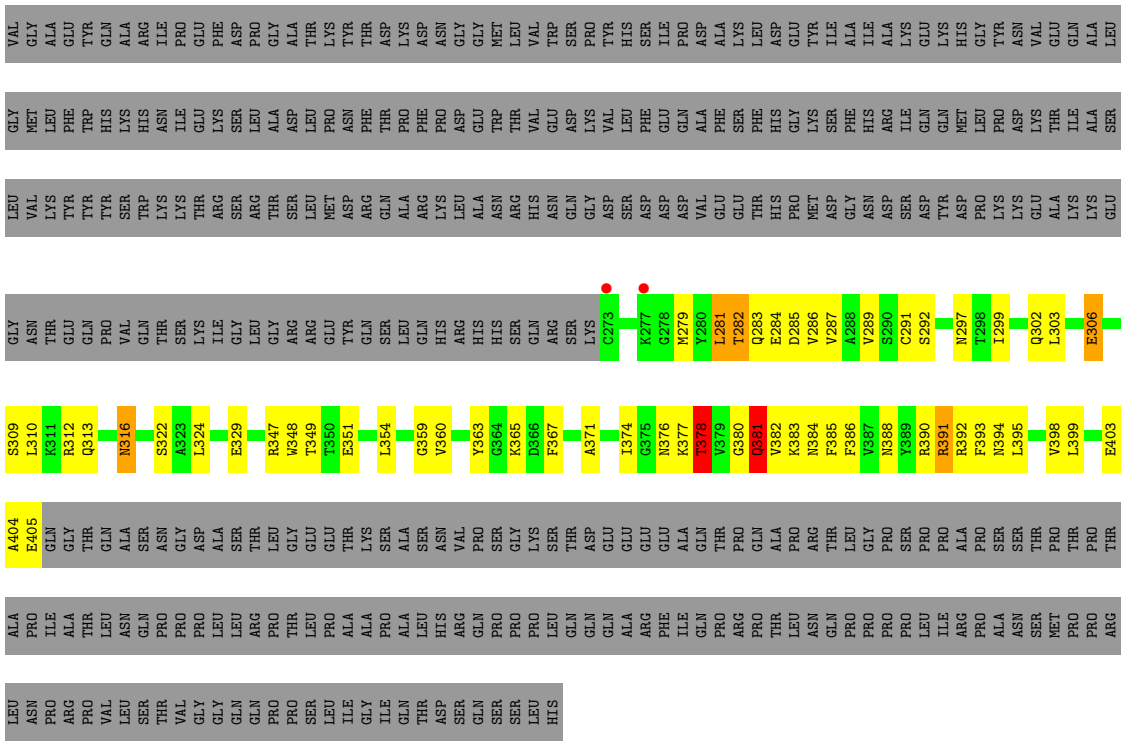
Chain A: 50% 22% • 24%



- Molecule 2: REST COREPRESSOR 3

Chain B: 14% 9% • 76%




VAL GLY ALA GLU TYR GLN ALA ARG ILE PRO GLU PHE ASP PRO GLY ALA THR LYS TYR THR ASP LYS ASP ASN GLY GLY MET LEU VAL TRP SER PRO TYR HIS SER ILE PRO ASP ALA LYS LEU ASP GLU TYR ILE ALA ILE ALA LYS GLU LYS HIS GLY TYR ASN VAL GLU GLN ALA LEU
GLY MET LEU PHE TRP HIS LYS HIS ASN ILE GLU LYS SER LEU ALA ASP LEU PRO ASN PHE THR PRO PHE PRO ASP GLU TRP THR VAL GLU ASP LYS VAL LEU PHE GLU GLN ALA PHE SER PHE HIS GLY LYS SER PHE HIS ARG ILE GLN GLN MET LEU PRO ASP LYS THR ILE ALA SER
LEU VAL LYS TYR TYR TYR SER TRP LYS LYS THR ARG SER ARG THR SER LEU MET ASP ARG GLN ALA ARG LYS LEU ALA ASN ARG HIS ASN GLN GLY ASP SER ASP ASP ASP VAL GLU GLU THR HIS PRO MET ASP GLY ASN ASP SER ASP TYR ASP PRO LYS LYS GLU ALA LYS LYS GLU
GLY ASN THR GLU GLN PRO VAL GLN THR SER LYS ILE GLY LEU GLY ARG ARG GLU TYR GLN SER LEU GLN HIS ARG HIS HIS SER GLN ARG SER LYS C273 K277 G278 N279 Y280 L281 T282 Q283 E284 D285 V286 V287 A288 V289 S290 C291 S292 N297 T298 I299 Q302 L303 E306
S309 L310 K311 R312 Q313 N316 S322 A323 L324 E329 R347 W348 T349 T350 E351 L354 G359 V360 Y363 G364 K365 D366 F367 A371 I374 G375 N376 K377 T378 V379 G380 Q381 V382 K383 N384 F385 F386 V387 N388 Y389 R390 R391 R392 F393 N394 L395 V398 L399 E403
A404 E405 GLN GLY THR GLN ALA SER ASN GLY ASP ALA SER THR LEU GLY GLU GLU THR LYS SER ALA SER ASN VAL PRO SER GLY LYS SER THR ASP GLU GLU GLU GLU ALA GLN THR PRO GLN ALA PRO ARG THR LEU LEU GLY PRO SER PRO PRO PRO LEU ILE ARG PRO ALA PRO SER SER THR PRO THR PRO THR
ALA PRO ILE ALA THR LEU ASN GLN PRO PRO PRO LEU LEU ARG PRO THR LEU PRO ALA ALA PRO ALA LEU HIS ARG GLN PRO PRO PRO LEU GLN GLN GLN ALA ARG PHE ILE GLN PRO ARG PRO THR LEU ASN GLN PRO PRO PRO PRO LEU ILE ARG PRO ALA ASN SER MET PRO THR PRO ARG
LEU ASN PRO ARG PRO VAL LEU SER THR VAL GLY GLY GLN GLN PRO PRO SER LEU ILE GLY ILE GLN THR ASP SER GLN SER SER LEU HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	118.04Å 177.44Å 235.79Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	141.78 – 3.00 70.89 – 3.00	Depositor EDS
% Data completeness (in resolution range)	99.3 (141.78-3.00) 99.4 (70.89-3.00)	Depositor EDS
R_{merge}	0.19	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.58 (at 3.01Å)	Xtriage
Refinement program	REFMAC 5.8.0069	Depositor
R, R_{free}	0.197 , 0.236 0.206 , 0.239	Depositor DCC
R_{free} test set	971 reflections (1.96%)	wwPDB-VP
Wilson B-factor (Å^2)	75.2	Xtriage
Anisotropy	0.091	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 66.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6339	wwPDB-VP
Average B, all atoms (Å^2)	80.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.98% 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

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	0.73	0/5331	0.89	6/7232 (0.1%)
2	B	0.57	0/1085	0.80	1/1461 (0.1%)
All	All	0.71	0/6416	0.87	7/8693 (0.1%)

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	820	ARG	NE-CZ-NH2	-6.95	116.83	120.30
1	A	820	ARG	NE-CZ-NH1	6.63	123.61	120.30
1	A	815	LEU	CA-CB-CG	6.42	130.08	115.30
1	A	795	ARG	NE-CZ-NH2	-5.43	117.58	120.30
1	A	795	ARG	NE-CZ-NH1	5.15	122.87	120.30

There are no chirality outliers.

There are no planarity outliers.

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	5217	0	5252	135	0
2	B	1069	0	1075	34	0
3	A	53	0	31	4	0
All	All	6339	0	6358	165	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 13.

The worst 5 of 165 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:526:ARG:NH2	1:A:530:ASP:OD1	1.86	1.08
2:B:306:GLU:O	2:B:310:LEU:HD12	1.68	0.93
1:A:726:ARG:O	1:A:730:ILE:HG12	1.84	0.77
1:A:760:SER:HB2	3:A:900:FAD:HM83	1.69	0.75
2:B:378:THR:O	2:B:381:GLN:NE2	2.18	0.75

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	664/872 (76%)	586 (88%)	63 (10%)	15 (2%)	6 30
2	B	131/553 (24%)	113 (86%)	13 (10%)	5 (4%)	3 18
All	All	795/1425 (56%)	699 (88%)	76 (10%)	20 (2%)	5 28

5 of 20 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	331	ALA
1	A	364	GLU
1	A	516	PRO
1	A	522	SER
1	A	834	TYR

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	566/711 (80%)	500 (88%)	66 (12%)	5 22
2	B	116/473 (24%)	99 (85%)	17 (15%)	3 15
All	All	682/1184 (58%)	599 (88%)	83 (12%)	5 21

5 of 83 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	744	LYS
2	B	306	GLU
1	A	760	SER
1	A	824	ARG
2	B	324	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	283	GLN
2	B	302	GLN
2	B	388	ASN
2	B	313	GLN
1	A	410	GLN

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)

1 ligand is 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	FAD	A	900	-	53,58,58	1.34	5 (9%)	68,89,89	1.72	16 (23%)

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	FAD	A	900	-	-	3/30/50/50	0/6/6/6

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	900	FAD	C9A-C5X	4.48	1.48	1.41
3	A	900	FAD	C5X-N5	-3.56	1.32	1.39
3	A	900	FAD	C4-N3	-3.01	1.33	1.38
3	A	900	FAD	C8-C7	2.86	1.48	1.40
3	A	900	FAD	C2B-C1B	-2.02	1.50	1.53

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	900	FAD	N3A-C2A-N1A	-4.55	121.57	128.68
3	A	900	FAD	P-O3P-PA	-3.57	120.58	132.83
3	A	900	FAD	C4X-C10-N1	-3.28	117.12	124.73
3	A	900	FAD	O4-C4-C4X	-3.27	117.92	126.60

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	900	FAD	C6-C5X-C9A	3.27	123.56	118.94

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	900	FAD	O4'-C4'-C5'-O5'
3	A	900	FAD	PA-O3P-P-O5'
3	A	900	FAD	O4B-C4B-C5B-O5B

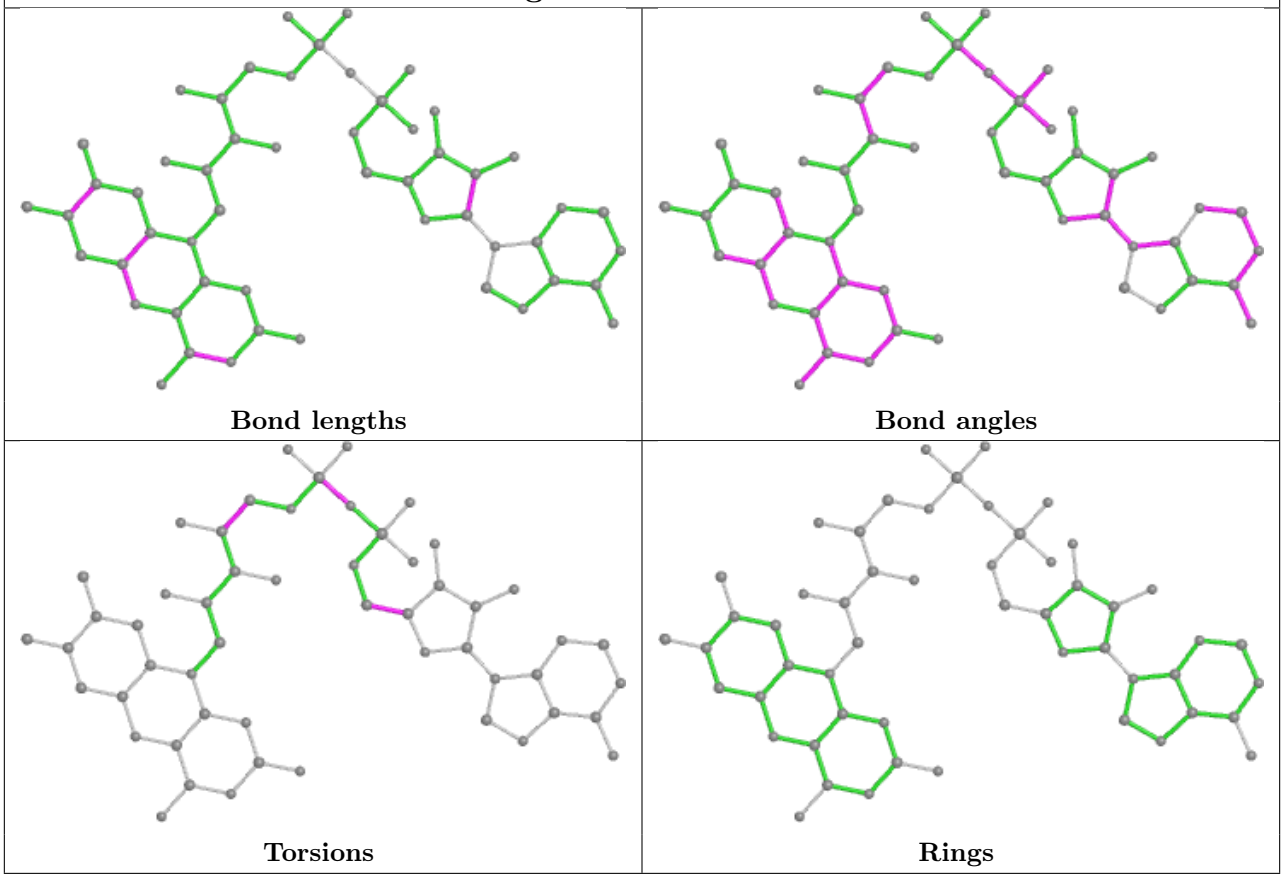
There are no ring outliers.

1 monomer is involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	900	FAD	4	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 900



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	666/872 (76%)	-0.22	0	100	100	41, 73, 110, 151	0
2	B	133/553 (24%)	-0.01	2 (1%)	73	46	68, 101, 130, 150	0
All	All	799/1425 (56%)	-0.18	2 (0%)	94	84	41, 78, 118, 151	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	277	LYS	2.5
2	B	273	CYS	2.3

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.

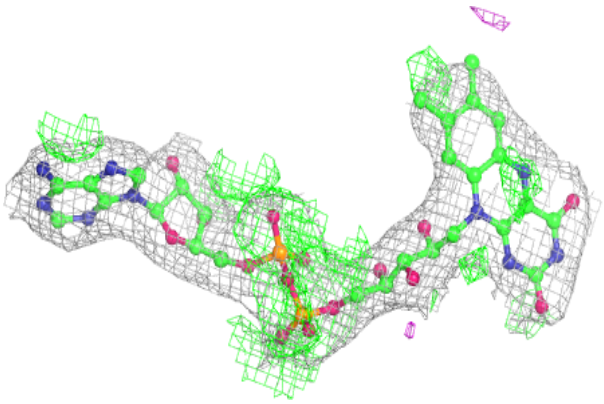
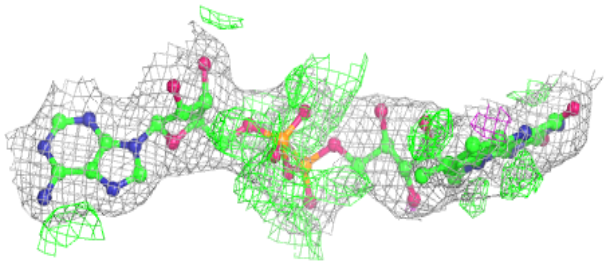
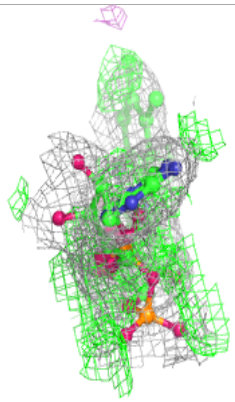
6.4 Ligands (i)

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	FAD	A	900	53/53	0.99	0.21	38,51,72,79	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 900: $2mF_o$-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.