



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2026 – 03:29 PM UTC

PDB ID : 4GU1 / pdb_00004gu1
Title : Crystal structure of LSD2
Authors : Chen, F.; Dong, Z.; Fang, J.; Yang, Y.; Li, Z.; Xu, Y.; Yang, H.; Wang, P.;
Fang, R.; Shi, Y.; Xu, Y.
Deposited on : 2012-08-29
Resolution : 2.94 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity : 4-5-2 with Phenix2.0
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0
EDS : 3.0
Buster-report : wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

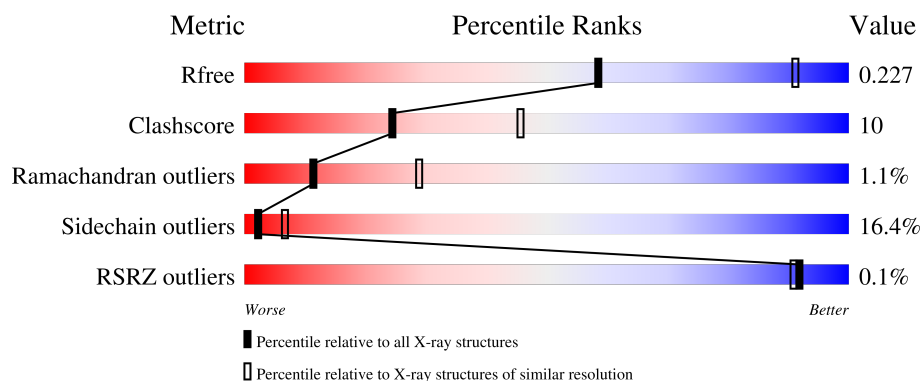
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.94 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}	180053	1159 (2.96-2.92)
Clashscore	190562	1184 (2.96-2.92)
Ramachandran outliers	187476	1131 (2.96-2.92)
Sidechain outliers	187428	1131 (2.96-2.92)
RSRZ outliers	180081	1159 (2.96-2.92)

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 $\leq 5\%$. 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	784	 65% 25% 5% 5%
1	B	784	 63% 27% 5% 5%

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 11971 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.

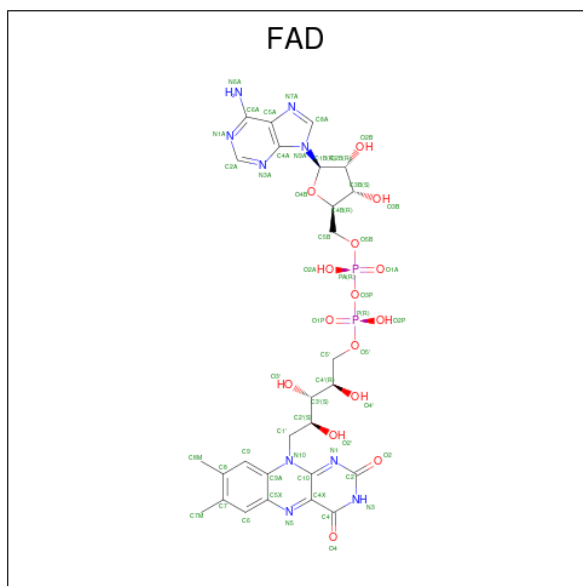
- Molecule 1 is a protein called Lysine-specific histone demethylase 1B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	746	Total	C	N	O	S	0	0	0
			5915	3775	1013	1086	41			
1	B	746	Total	C	N	O	S	0	0	0
			5915	3775	1013	1086	41			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	39	PRO	-	expression tag	UNP Q8NB78
A	40	LEU	-	expression tag	UNP Q8NB78
A	41	GLY	-	expression tag	UNP Q8NB78
A	42	SER	-	expression tag	UNP Q8NB78
A	43	GLU	-	expression tag	UNP Q8NB78
A	44	PHE	-	expression tag	UNP Q8NB78
A	45	LYS	-	expression tag	UNP Q8NB78
A	46	GLY	-	expression tag	UNP Q8NB78
A	47	LEU	-	expression tag	UNP Q8NB78
A	48	ARG	-	expression tag	UNP Q8NB78
A	49	ARG	-	expression tag	UNP Q8NB78
A	50	ARG	-	expression tag	UNP Q8NB78
B	39	PRO	-	expression tag	UNP Q8NB78
B	40	LEU	-	expression tag	UNP Q8NB78
B	41	GLY	-	expression tag	UNP Q8NB78
B	42	SER	-	expression tag	UNP Q8NB78
B	43	GLU	-	expression tag	UNP Q8NB78
B	44	PHE	-	expression tag	UNP Q8NB78
B	45	LYS	-	expression tag	UNP Q8NB78
B	46	GLY	-	expression tag	UNP Q8NB78
B	47	LEU	-	expression tag	UNP Q8NB78
B	48	ARG	-	expression tag	UNP Q8NB78
B	49	ARG	-	expression tag	UNP Q8NB78
B	50	ARG	-	expression tag	UNP Q8NB78

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



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

- Molecule 3 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	2	Total	Cl	0	0
			2	2		
3	B	2	Total	Cl	0	0
			2	2		

- Molecule 4 is SODIUM ION (CCD ID: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Na	0	0
			1	1		
4	B	1	Total	Na	0	0
			1	1		

- Molecule 5 is ZINC ION (CCD ID: ZN) (formula: Zn).

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

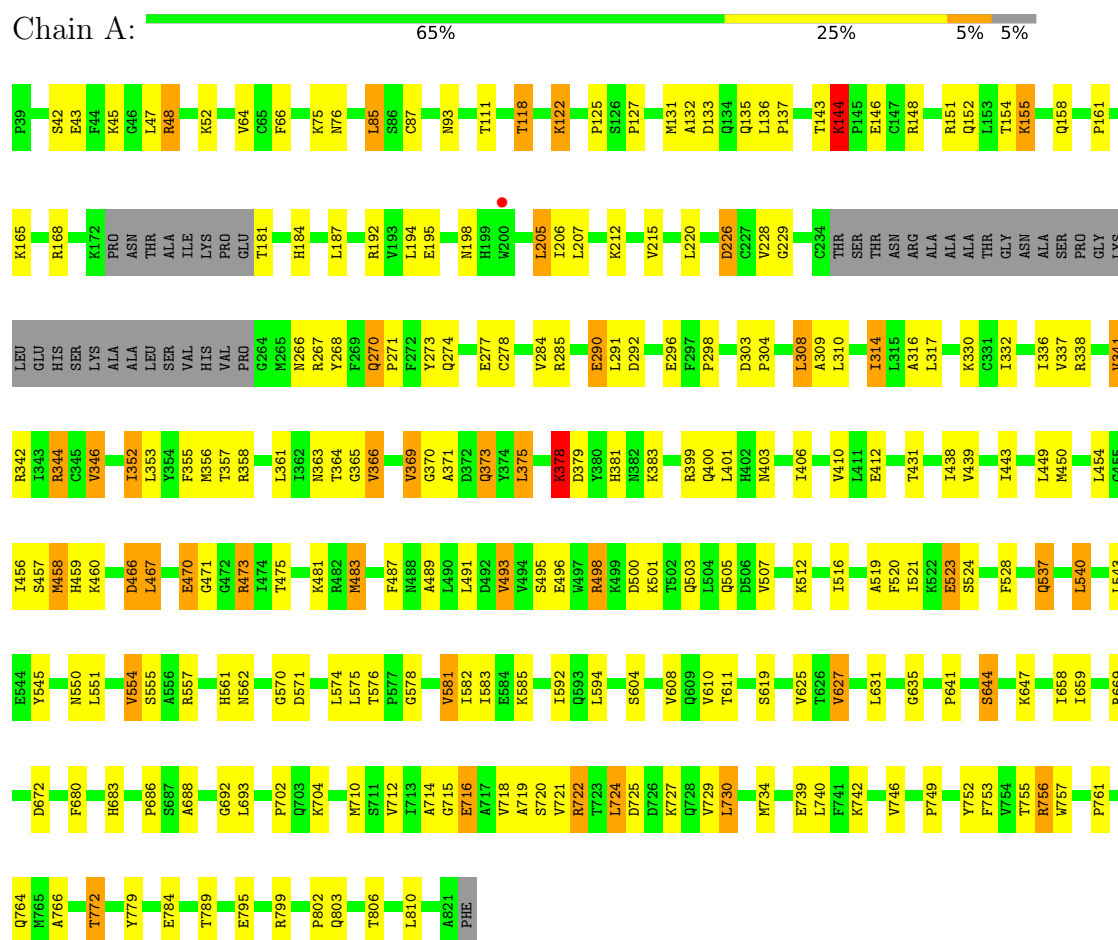
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	11	Total 11	O 11	0	0
6	B	12	Total 12	O 12	0	0

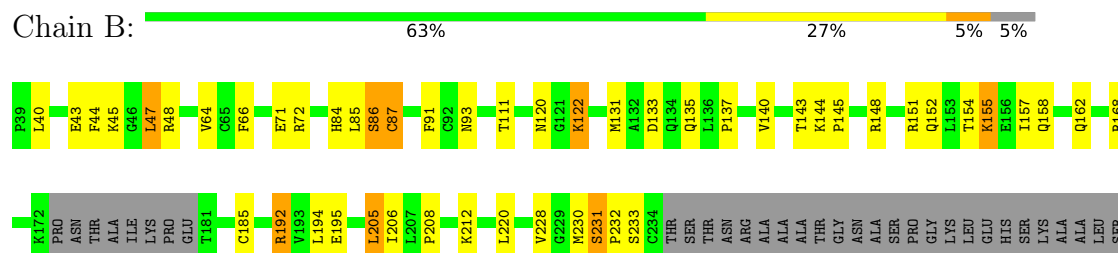
3 Residue-property plots

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 1B



• Molecule 1: Lysine-specific histone demethylase 1B



T750	L664	F568	L467	VAL
A751	Q665	A569	E470	HIS
Y752	F666	G570	G471	VAL
		D571		PRO
T755	V675		T475	G264
R756	Q676			Q270
W757		L574	L367	
	H683	L575	S368	
P761	V684	L576	V369	Y273
W762	P685	P577		Q274
I763	P686	G578	L375	P275
Q764	S687		L376	N276
W765	A688	V581	P377	E277
A766		I582	K378	
	R691	I583	D379	K280
T772	G692	E584	Y380	
	L693	K585		C283
Y779	V696	I592	K383	E290
E784		P597	S384	L291
	P702	I601	V385	
T789	Q703		G394	L294
	K704	V608	R399	Y295
R799		T611	Q400	E296
	M710	T612	L401	Y300
	S711	T613	N403	S301
T806	V712			R302
E815	I713	T619	V408	D303
	A714	A620	T409	P304
K818	G715	Q621	V410	T305
	E716	V625	L411	M306
A821	I717	T626	E412	Y307
PHE	V718	V627	R416	L308
	A719		I417	A309
	S720		G418	L310
	V721	L631		I314
		L632	T431	
	L724		V439	Q329
	D725	I637	M440	K330
	D726	Q638	G442	C331
	K727	F639	I443	I332
	Q728			
	V729	L643	L543	I336
	L730	S644	G548	V337
	Q731	K647	S549	R338
	Q732		N550	V341
	C733	A650	L551	R342
	M734	I651	V554	I343
	A735	S653	A556	R344
	T736	K742	R557	C345
	T737		D560	V346
	L737		H561	Q347
	R738		N562	E348
	E739			
	E739			R351
	L740			I352
	F741			L353
	K742			
	V746			
	P749			

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	89.16Å 89.22Å 342.54Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.56 – 2.94 39.56 – 2.94	Depositor EDS
% Data completeness (in resolution range)	97.5 (39.56-2.94) 97.5 (39.56-2.94)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.32 (at 2.96Å)	Xtriage
Refinement program	PHENIX (phenix.refine)	Depositor
R, R_{free}	0.185 , 0.228 0.187 , 0.227	Depositor DCC
R_{free} test set	2926 reflections (4.94%)	wwPDB-VP
Wilson B-factor (Å ²)	83.1	Xtriage
Anisotropy	0.401	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 77.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.467 for k,h,-l	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	11971	wwPDB-VP
Average B, all atoms (Å ²)	83.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.09% 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: FAD, CL, NA, ZN

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	0.57	0/6060	0.93	9/8203 (0.1%)
1	B	0.55	0/6060	0.93	7/8203 (0.1%)
All	All	0.56	0/12120	0.93	16/16406 (0.1%)

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	303	ASP	CA-C-N	8.03	128.22	119.87
1	B	303	ASP	C-N-CA	8.03	128.22	119.87
1	A	520	PHE	N-CA-C	-6.87	105.43	113.88
1	B	675	VAL	CB-CA-C	-5.87	103.89	111.53
1	B	520	PHE	N-CA-C	-5.61	106.98	113.88

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	5915	0	5842	114	0
1	B	5915	0	5842	122	0
2	A	53	0	31	5	0
2	B	53	0	31	6	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	2	0	0	1	0
3	B	2	0	0	1	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
5	A	3	0	0	0	0
5	B	3	0	0	0	0
6	A	11	0	0	0	0
6	B	12	0	0	0	0
All	All	11971	0	11746	234	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 10.

The worst 5 of 234 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:498:ARG:HD3	1:A:557:ARG:HA	1.57	0.87
1:B:192:ARG:HG3	1:B:192:ARG:HH11	1.46	0.80
1:B:439:VAL:HG21	1:B:458:MET:HE2	1.72	0.71
1:A:458:MET:HE3	1:A:575:LEU:HD13	1.73	0.71
1:A:332:ILE:HG12	1:A:346:VAL:HB	1.75	0.68

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	740/784 (94%)	681 (92%)	52 (7%)	7 (1%)	14	34
1	B	740/784 (94%)	674 (91%)	57 (8%)	9 (1%)	10	27
All	All	1480/1568 (94%)	1355 (92%)	109 (7%)	16 (1%)	11	29

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	120	ASN
1	A	369	VAL
1	A	635	GLY
1	B	155	LYS
1	B	525	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	A	640/669 (96%)	536 (84%)	104 (16%)	2	7
1	B	640/669 (96%)	534 (83%)	106 (17%)	2	6
All	All	1280/1338 (96%)	1070 (84%)	210 (16%)	2	7

5 of 210 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	111	THR
1	B	341	VAL
1	B	729	VAL
1	B	151	ARG
1	B	270	GLN

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	A	537	GLN
1	A	764	GLN
1	B	676	GLN
1	B	665	GLN
1	A	440	ASN

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 oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 14 ligands modelled in this entry, 12 are monoatomic - leaving 2 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	FAD	B	901	-	58,58,58	2.18	19 (32%)	85,89,89	1.84	25 (29%)
2	FAD	A	901	-	58,58,58	1.29	6 (10%)	85,89,89	1.62	16 (18%)

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	B	901	-	-	9/34/50/50	0/6/6/6
2	FAD	A	901	-	-	3/34/50/50	0/6/6/6

The worst 5 of 25 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	901	FAD	PA-O3P	-8.07	1.50	1.59

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	901	FAD	C5A-N7A	-4.57	1.30	1.39
2	B	901	FAD	P-O3P	-4.52	1.54	1.59
2	A	901	FAD	C4X-N5	4.47	1.40	1.30
2	B	901	FAD	C8A-N9A	-3.66	1.31	1.37

The worst 5 of 41 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	901	FAD	C5A-C4A-N3A	-5.76	118.79	126.72
2	A	901	FAD	C5A-C4A-N3A	-4.98	119.85	126.72
2	A	901	FAD	N3A-C2A-N1A	-4.51	121.75	128.58
2	B	901	FAD	N3A-C2A-N1A	-4.36	121.98	128.58
2	B	901	FAD	C4-N3-C2	-4.28	118.03	125.64

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	901	FAD	C5B-O5B-PA-O3P
2	B	901	FAD	C3'-C4'-C5'-O5'
2	B	901	FAD	O4'-C4'-C5'-O5'
2	B	901	FAD	C5'-O5'-P-O1P
2	B	901	FAD	C5'-O5'-P-O2P

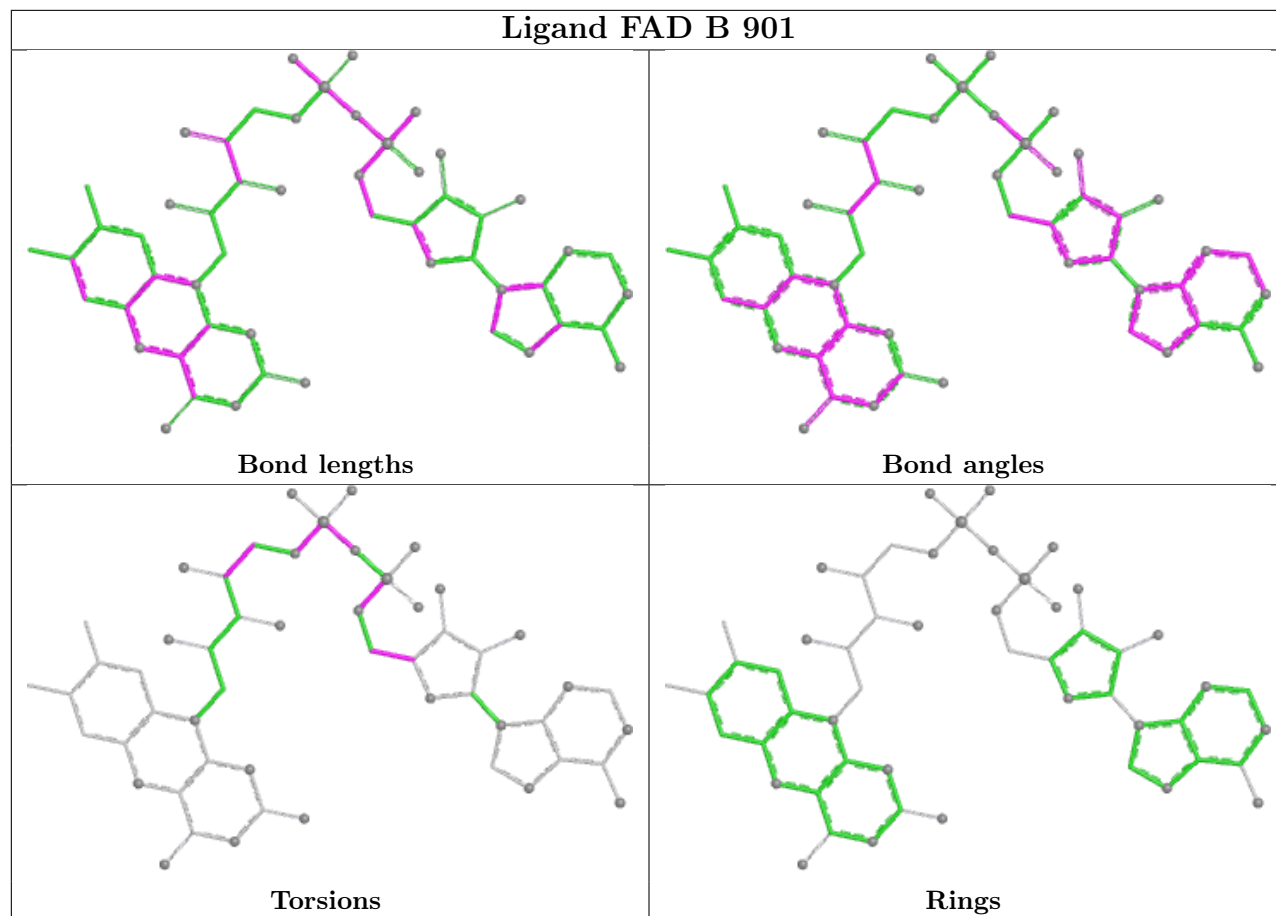
There are no ring outliers.

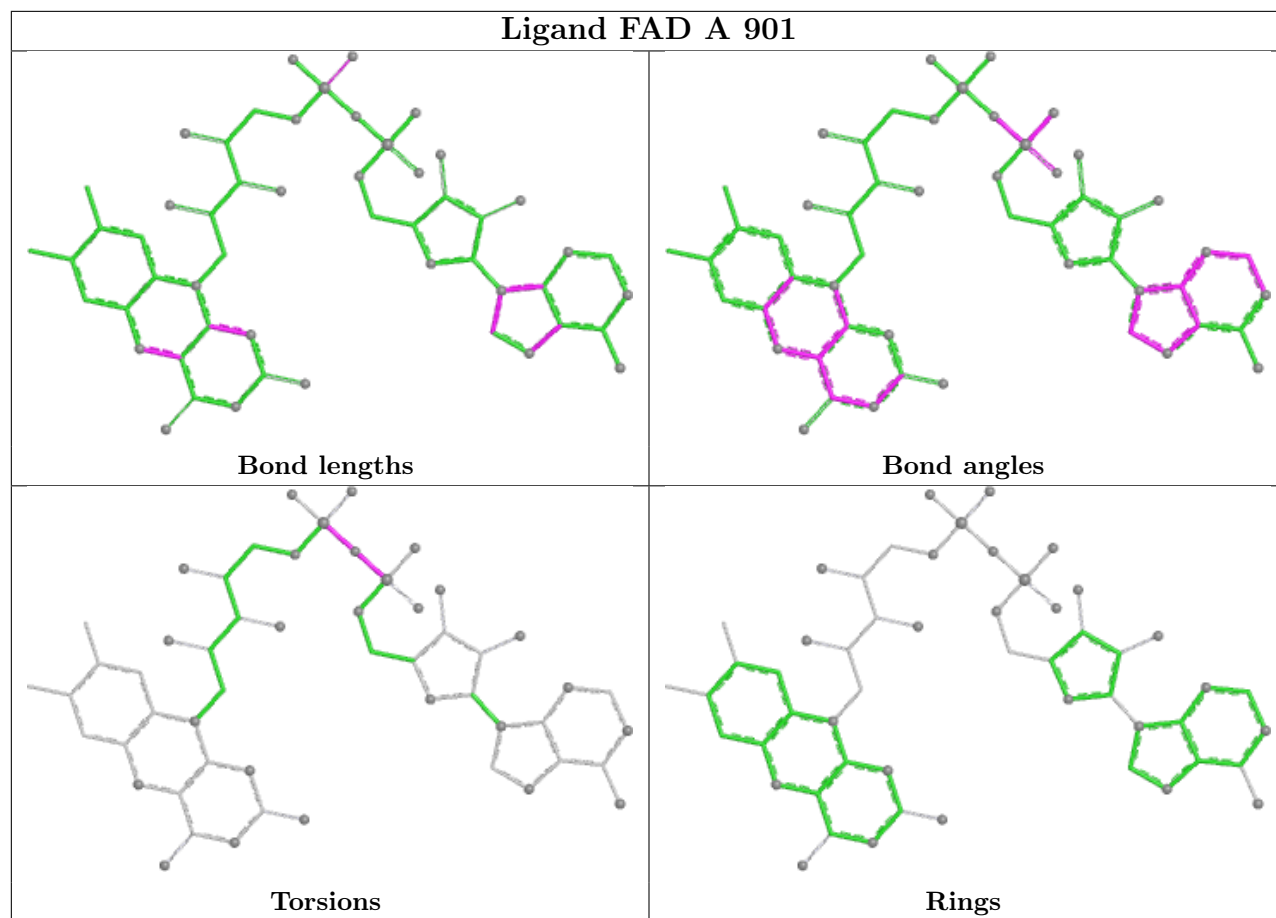
2 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	901	FAD	6	0
2	A	901	FAD	5	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 than 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.





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	746/784 (95%)	-1.31	1 (0%) 92 91	54, 79, 116, 161	0
1	B	746/784 (95%)	-1.27	0 100 100	56, 80, 117, 153	0
All	All	1492/1568 (95%)	-1.29	1 (0%) 92 91	54, 80, 117, 161	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	200	TRP	2.4

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 oligosaccharides 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.

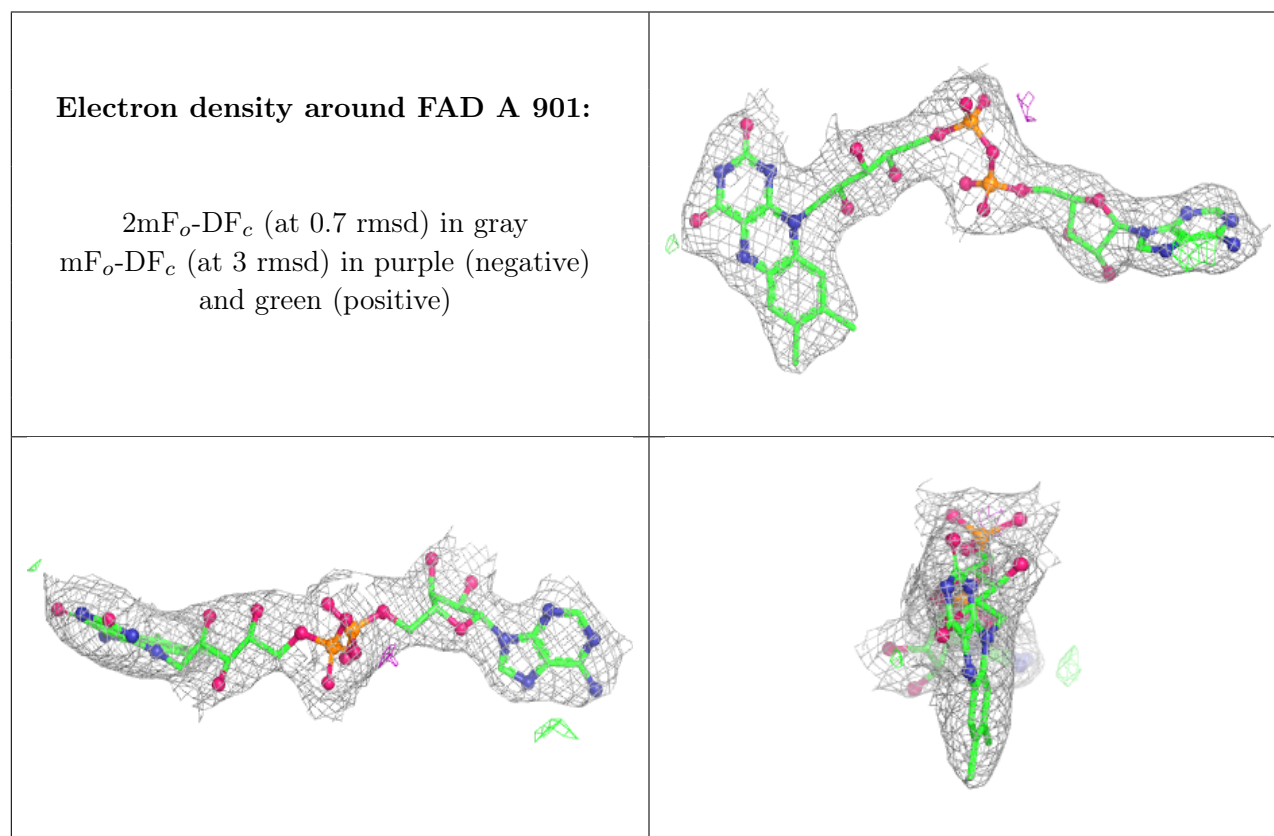
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
4	NA	B	904	1/1	0.98	0.10	86,86,86,86	0
3	CL	A	903	1/1	0.99	0.06	102,102,102,102	0
3	CL	B	902	1/1	0.99	0.04	102,102,102,102	0
3	CL	B	903	1/1	0.99	0.06	98,98,98,98	0

Continued on next page...

Continued from previous page...

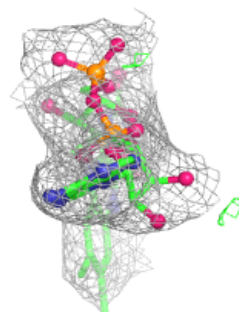
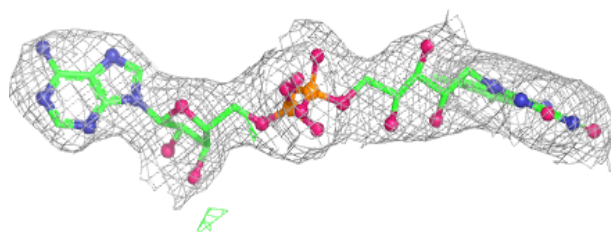
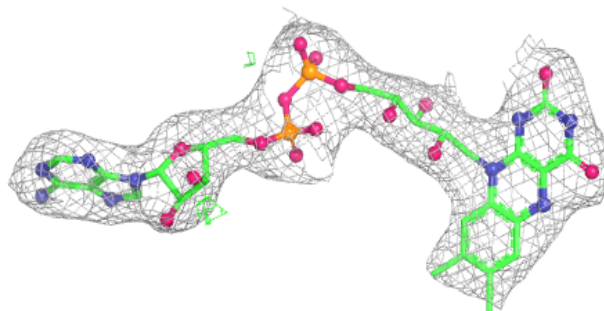
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
4	NA	A	904	1/1	0.99	0.08	88,88,88,88	0
3	CL	A	902	1/1	0.99	0.05	104,104,104,104	0
2	FAD	A	901	53/53	1.00	0.03	51,68,77,80	0
2	FAD	B	901	53/53	1.00	0.03	54,71,79,80	0
5	ZN	A	905	1/1	1.00	0.01	72,72,72,72	0
5	ZN	A	906	1/1	1.00	0.01	96,96,96,96	0
5	ZN	A	907	1/1	1.00	0.01	106,106,106,106	0
5	ZN	B	905	1/1	1.00	0.01	73,73,73,73	0
5	ZN	B	906	1/1	1.00	0.01	98,98,98,98	0
5	ZN	B	907	1/1	1.00	0.01	101,101,101,101	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 B 901:

$2mF_o - DF_c$ (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.