



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 13, 2024 – 09:32 AM EDT

PDB ID : 4FWJ  
Title : Native structure of LSD2/AOF1/KDM1b in spacegroup of I222 at 2.9A  
Authors : Zhang, Q.; Chen, Z.  
Deposited on : 2012-07-01  
Resolution : 2.90 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](mailto: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 : 4.02b-467  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 1.20.1  
EDS : 2.36.2  
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) : 2.36.2

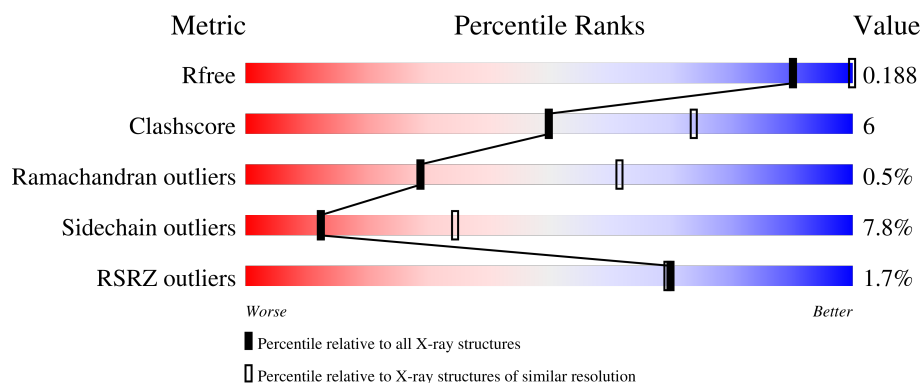
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.90 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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  $\geq 3$ , 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	796	<div> <div> <div></div> <div>78%</div> <div>14%</div> <div>6%</div> </div> </div>
1	B	796	<div> <div> <div>2%</div> <div>77%</div> <div>15%</div> <div>7%</div> </div> </div>

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 12080 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	750	Total	C	N	O	S	0	1	0
			5800	3712	992	1054	42			
1	B	742	Total	C	N	O	S	0	2	0
			5698	3641	962	1053	42			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	27	GLY	-	expression tag	UNP Q8NB78
A	28	HIS	-	expression tag	UNP Q8NB78
A	29	MET	-	expression tag	UNP Q8NB78
B	27	GLY	-	expression tag	UNP Q8NB78
B	28	HIS	-	expression tag	UNP Q8NB78
B	29	MET	-	expression tag	UNP Q8NB78

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 3 is FLAVIN-ADENINE DINUCLEOTIDE (three-letter code: FAD) (formula: C<sub>27</sub>H<sub>33</sub>N<sub>9</sub>O<sub>15</sub>P<sub>2</sub>).

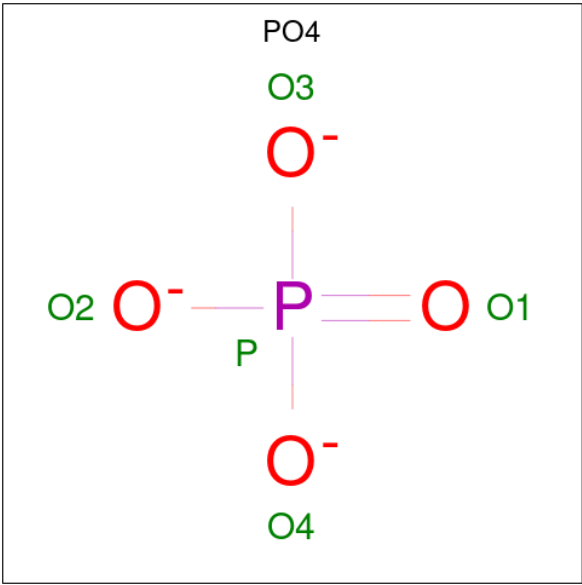


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			53	27	9	15	2		
3	B	1	Total	C	N	O	P	0	0
			53	27	9	15	2		

- Molecule 4 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	K	0	0
			1	1		

- Molecule 5 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total	O	P	0	0
			5	4	1		

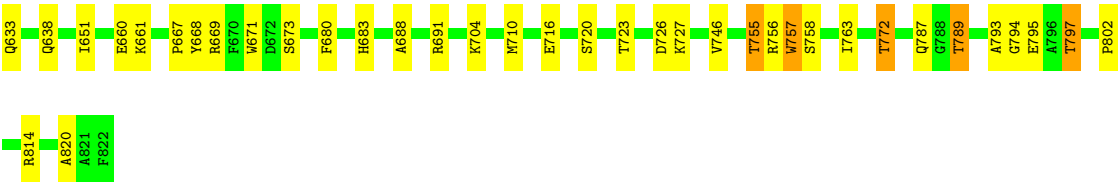
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	284	Total	O	0	0
			284	284		
6	B	180	Total	O	0	0
			180	180		



- Molecule 1: Lysine-specific histone demethylase 1B





## 4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	144.32Å 170.93Å 202.57Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.90 49.74 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.6 (50.00-2.90) 99.6 (49.74-2.90)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.39 (at 2.91Å)	Xtriage
Refinement program	REFMAC 5.6.0116	Depositor
R, $R_{free}$	0.186 , 0.222 0.189 , 0.188	Depositor DCC
$R_{free}$ test set	2818 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	45.2	Xtriage
Anisotropy	0.622	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 41.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	12080	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.87% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: PO4, K, ZN, 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	0.54	7/5947 (0.1%)	0.62	1/8067 (0.0%)
1	B	0.54	8/5846 (0.1%)	0.61	2/7941 (0.0%)
All	All	0.54	15/11793 (0.1%)	0.61	3/16008 (0.0%)

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

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	422	TRP	CD2-CE2	6.24	1.48	1.41
1	A	559	TRP	CD2-CE2	5.92	1.48	1.41
1	A	757	TRP	CD2-CE2	5.46	1.48	1.41
1	A	139	TRP	CD2-CE2	5.29	1.47	1.41
1	B	318	TRP	CD2-CE2	5.29	1.47	1.41

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	340	LEU	CA-CB-CG	6.75	130.83	115.30
1	B	631	LEU	CA-CB-CG	-5.64	102.32	115.30
1	A	664	LEU	CA-CB-CG	5.52	128.00	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	95	CYS	Peptide

## 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	A	5800	0	5607	61	0
1	B	5698	0	5408	70	0
2	A	3	0	0	0	0
2	B	3	0	0	0	0
3	A	53	0	31	2	0
3	B	53	0	31	3	0
4	A	1	0	0	0	0
5	B	5	0	0	0	0
6	A	284	0	0	4	0
6	B	180	0	0	0	0
All	All	12080	0	11077	131	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 6.

The worst 5 of 131 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:B:95:CYS:O	1:B:96[B]:PHE:HD1	1.26	1.19
1:B:332:ILE:HD12	1:B:346:VAL:HG13	1.30	1.12
1:B:95:CYS:O	1:B:96[B]:PHE:CD1	2.15	1.00
1:A:797:THR:CG2	1:A:814:ARG:HH22	1.80	0.95
1:B:797:THR:CG2	1:B:814:ARG:HH22	1.83	0.91

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	A	745/796 (94%)	699 (94%)	40 (5%)	6 (1%)	19	51
1	B	736/796 (92%)	699 (95%)	35 (5%)	2 (0%)	41	71
All	All	1481/1592 (93%)	1398 (94%)	75 (5%)	8 (0%)	29	61

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	104	LYS
1	A	105	ASP
1	A	172	LYS
1	A	371	ALA
1	B	78	TYR

### 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	602/678 (89%)	555 (92%)	47 (8%)	12	34
1	B	589/678 (87%)	543 (92%)	46 (8%)	12	34
All	All	1191/1356 (88%)	1098 (92%)	93 (8%)	12	34

5 of 93 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	235	THR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	536	LEU
1	B	266	ASN
1	B	361	LEU
1	B	572	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	266	ASN
1	B	550	ASN
1	B	505	GLN
1	B	561	HIS
1	A	505	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](#)

Of 10 ligands modelled in this entry, 7 are monoatomic - leaving 3 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
3	FAD	A	904	-	54,58,58	1.24	3 (5%)	71,89,89	1.61	9 (12%)
5	PO4	B	905	-	4,4,4	0.91	0	6,6,6	0.48	0
3	FAD	B	904	-	54,58,58	1.21	4 (7%)	71,89,89	1.88	15 (21%)

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	904	-	-	9/30/50/50	0/6/6/6
3	FAD	B	904	-	-	15/30/50/50	0/6/6/6

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	904	FAD	C4X-N5	4.12	1.39	1.30
3	A	904	FAD	C4X-N5	4.07	1.39	1.30
3	A	904	FAD	C10-N1	3.68	1.40	1.33
3	A	904	FAD	P-O3P	3.47	1.63	1.59
3	B	904	FAD	C1'-C2'	3.22	1.57	1.52

The worst 5 of 24 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	904	FAD	C4B-O4B-C1B	-7.19	103.34	109.92
3	B	904	FAD	C4'-C3'-C2'	-6.57	102.64	113.57
3	B	904	FAD	C1'-C2'-C3'	5.43	124.39	109.66
3	B	904	FAD	N3A-C2A-N1A	-5.31	121.46	128.67
3	A	904	FAD	N3A-C2A-N1A	-5.03	121.84	128.67

There are no chirality outliers.

5 of 24 torsion outliers are listed below:

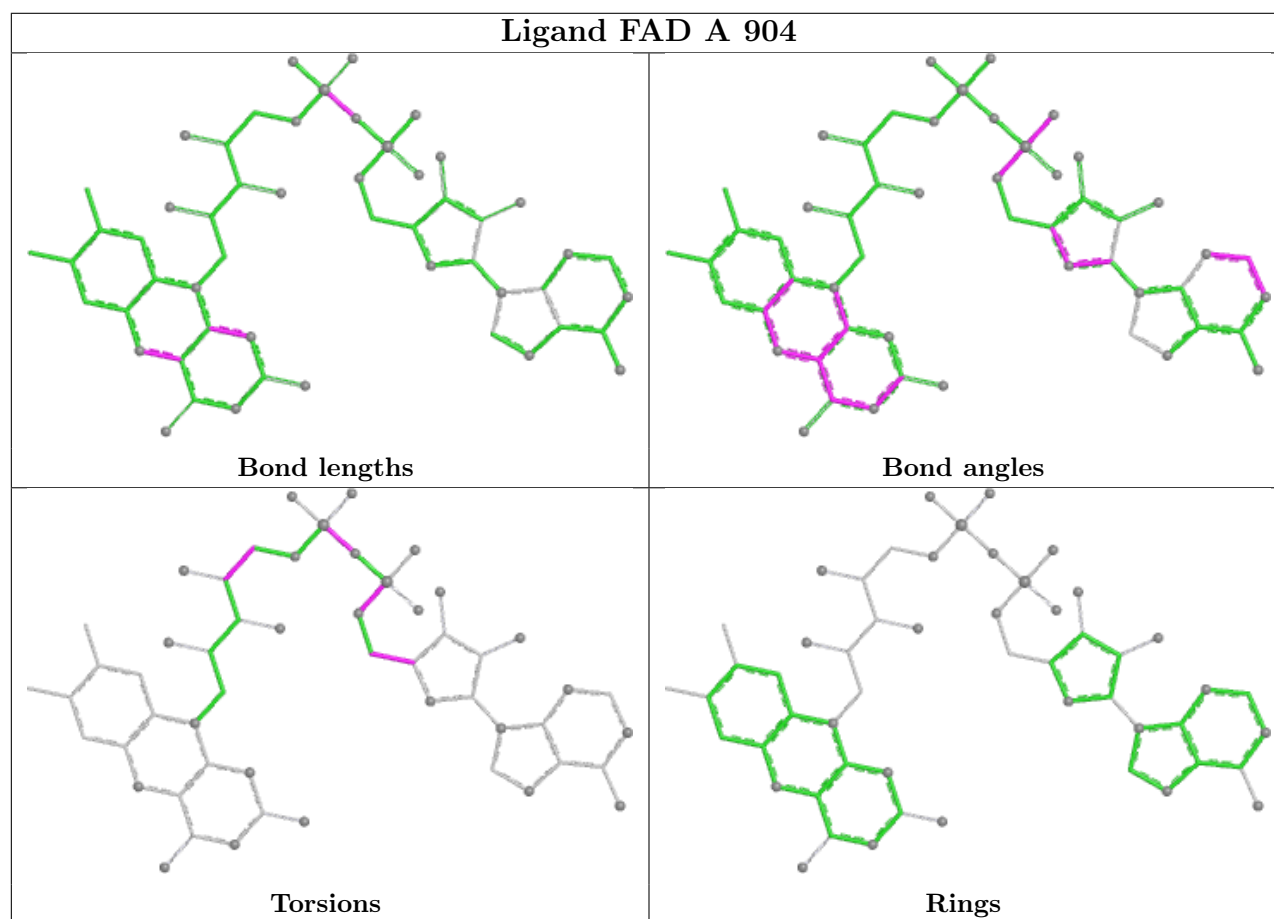
Mol	Chain	Res	Type	Atoms
3	A	904	FAD	C5B-O5B-PA-O1A
3	A	904	FAD	C5B-O5B-PA-O2A
3	A	904	FAD	C5B-O5B-PA-O3P
3	A	904	FAD	C3'-C4'-C5'-O5'
3	A	904	FAD	O4'-C4'-C5'-O5'

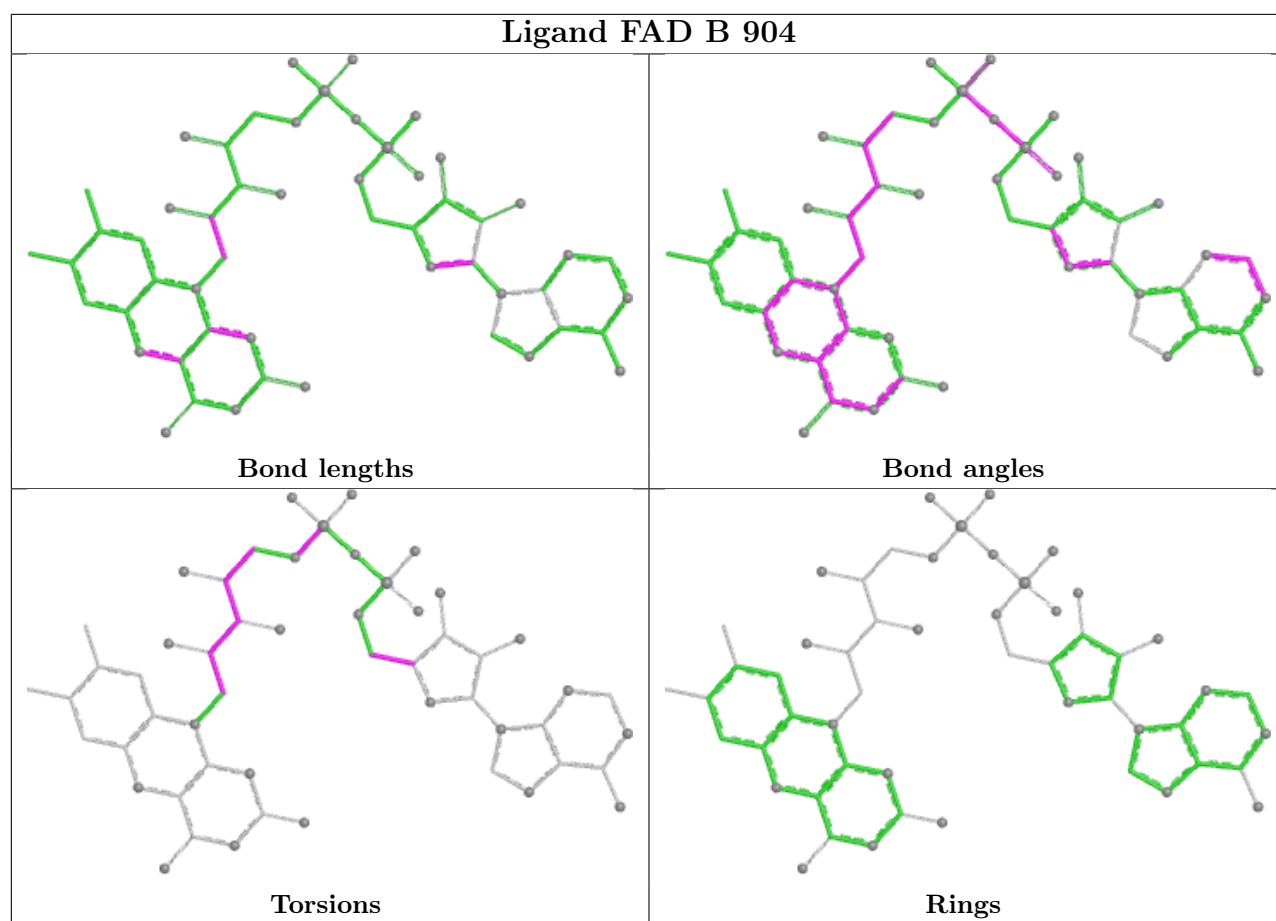
There are no ring outliers.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	904	FAD	2	0
3	B	904	FAD	3	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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	750/796 (94%)	-0.39	11 (1%) 73 73	21, 36, 77, 179	1 (0%)
1	B	742/796 (93%)	-0.26	14 (1%) 66 65	26, 46, 106, 269	1 (0%)
All	All	1492/1592 (93%)	-0.33	25 (1%) 70 69	21, 41, 99, 269	2 (0%)

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	262	VAL	6.5
1	B	237	THR	6.1
1	A	50	TYR	4.6
1	A	237	THR	3.7
1	A	48	LYS	3.5

### 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.

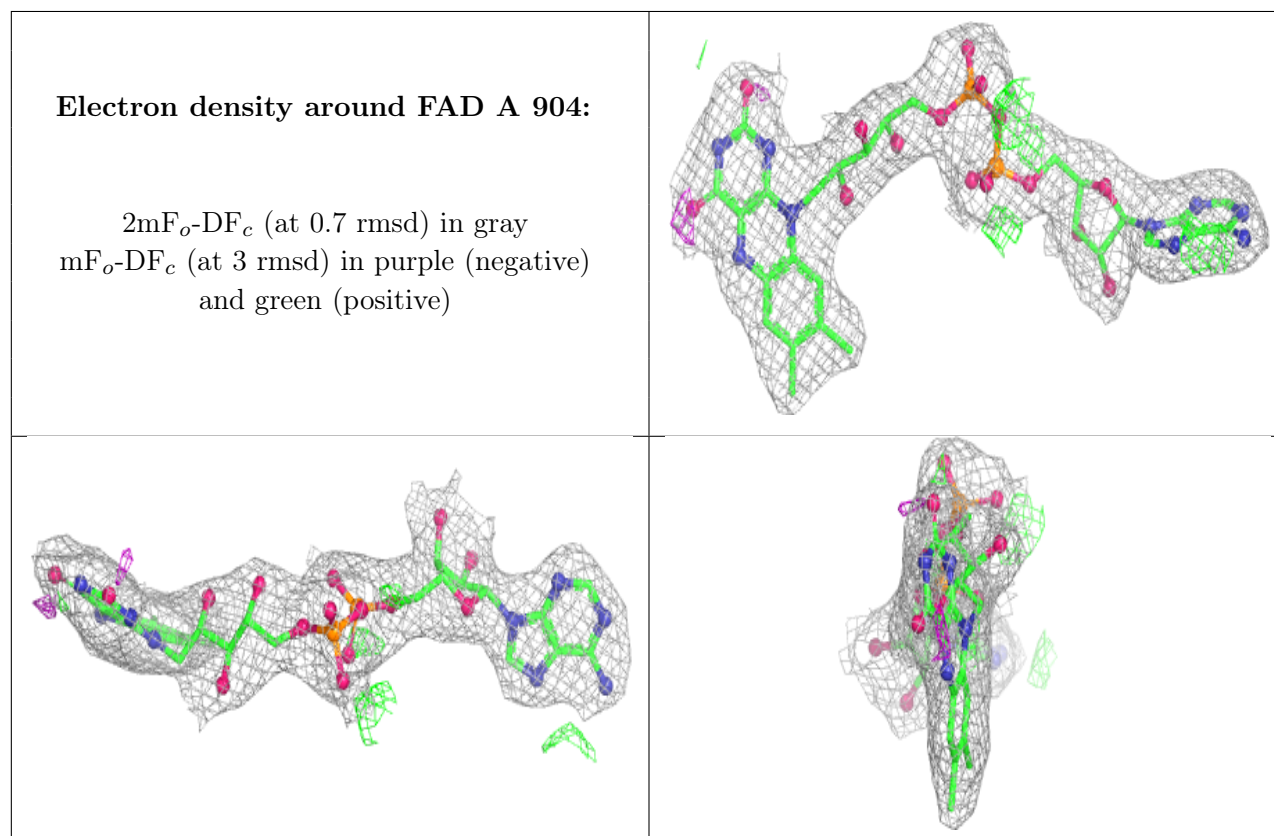
### 6.4 Ligands [i](#)

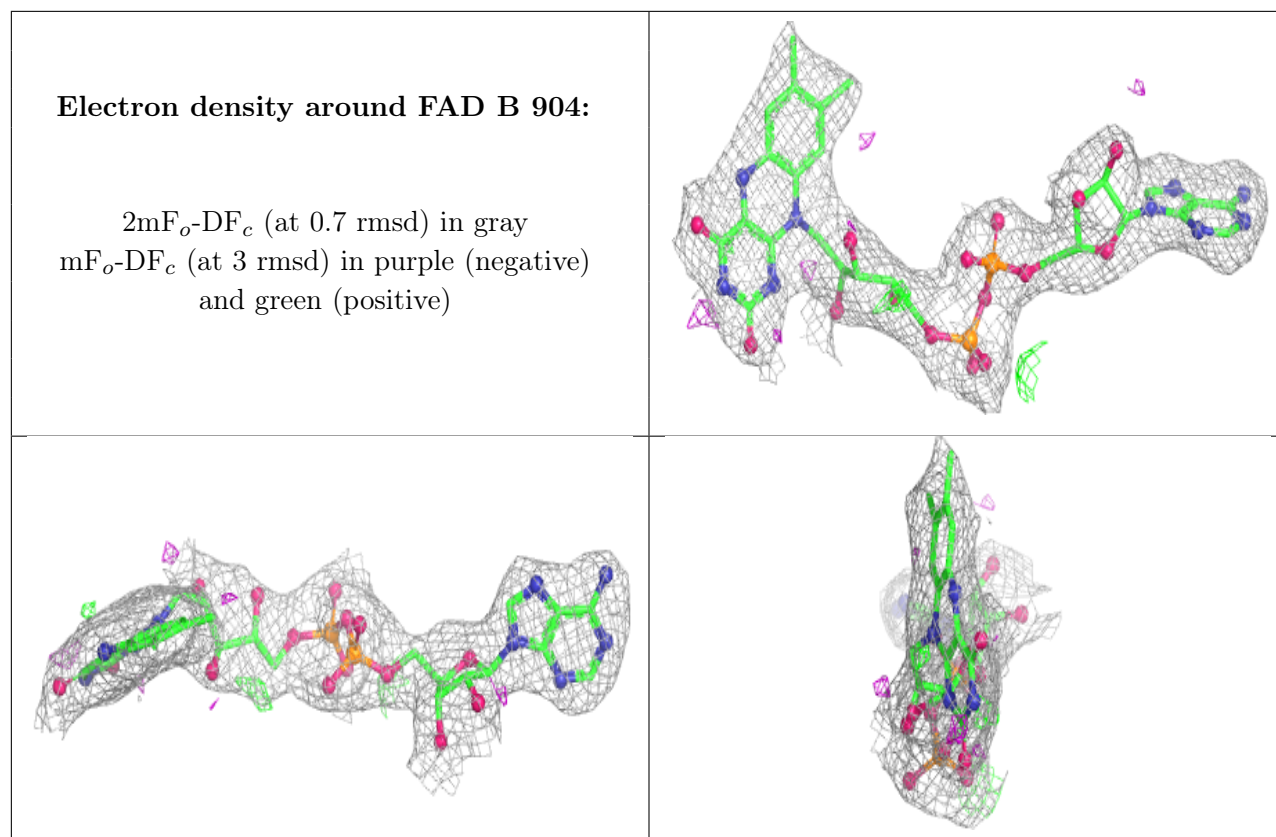
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q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



Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	ZN	B	903	1/1	0.92	0.05	102,102,102,102	0
4	K	A	905	1/1	0.95	0.04	57,57,57,57	0
2	ZN	A	902	1/1	0.98	0.06	52,52,52,52	0
3	FAD	A	904	53/53	0.98	0.15	20,22,25,26	0
3	FAD	B	904	53/53	0.98	0.14	27,31,34,36	0
2	ZN	B	902	1/1	0.98	0.03	88,88,88,88	0
5	PO4	B	905	5/5	0.98	0.12	42,45,48,48	0
2	ZN	A	901	1/1	0.99	0.05	59,59,59,59	0
2	ZN	A	903	1/1	0.99	0.03	62,62,62,62	0
2	ZN	B	901	1/1	0.99	0.06	51,51,51,51	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.





## 6.5 Other polymers [i](#)

There are no such residues in this entry.