

Full wwPDB X-ray Structure Validation Report (i)

Sep 9, 2023 – 04:58 PM EDT

PDB ID : 4HOO

Title: Crystal structure of human JMJD2D/KDM4D apoenzyme

Authors : Krishnan, S.; Trievel, R.C.

Deposited on : 2012-10-22

Resolution : 2.50 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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) : 1.13 EDS : 2.35.1

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)

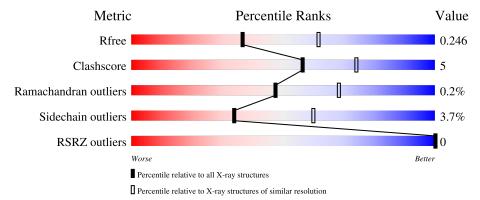
Validation Pipeline (wwPDB-VP) : 2.35.1

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 2.50 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	Similar resolution
Metric	$(\# \mathrm{Entries})$	$(\# ext{Entries}, ext{ resolution range}(ext{Å}))$
R_{free}	130704	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	330	89%	9%	
1	В	330	86%	12%	•

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	ACT	A	403	_	-	X	-



2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 5347 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 demethylase 4D.

Mol	Chain	Residues	${f Atoms}$			ZeroOcc	AltConf	Trace		
1	A	327	Total 2595	C 1673	N 441	O 468	S 13	0	1	0
1	В	330	Total 2587	C 1666	N 442	O 466	S 13	0	1	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	93	ALA	LYS	engineered mutation	UNP Q6B0I6
A	94	ALA	LYS	engineered mutation	UNP Q6B0I6
В	93	ALA	LYS	engineered mutation	UNP Q6B0I6
В	94	ALA	LYS	engineered mutation	UNP Q6B0I6

• Molecule 2 is NICKEL (II) ION (three-letter code: NI) (formula: Ni).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Ni 1 1	0	0
2	В	1	Total Ni 1 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	В	1	Total Zn 1 1	0	0

• Molecule 4 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).





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

• Molecule 5 is water.

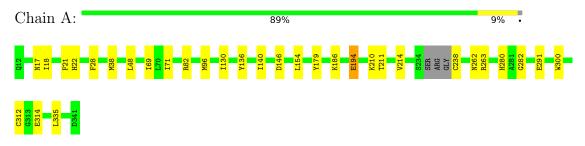
N	Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
	5	A	73	Total O 73 73	0	0
	5	В	80	Total O 80 80	0	0



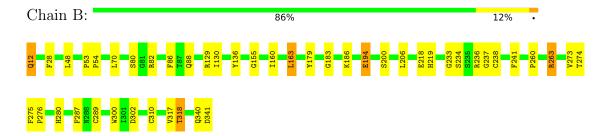
3 Residue-property plots (i)

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 demethylase 4D



• Molecule 1: Lysine-specific demethylase 4D





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32	Depositor
Cell constants	73.12Å 73.12Å 136.03Å	D
a, b, c, α , β , γ	90.00° 90.00° 120.00°	Depositor
Resolution (Å)	17.93 - 2.50	Depositor
Resolution (A)	17.93 - 2.50	EDS
% Data completeness	91.5 (17.93-2.50)	Depositor
(in resolution range)	91.5 (17.93-2.50)	EDS
R_{merge}	0.08	Depositor
R_{sum}	(Not available)	Depositor
$< I/\sigma(I) > 1$	2.47 (at 2.49Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
P.P.	0.207 , 0.246	Depositor
R, R_{free}	0.203 , 0.246	DCC
R_{free} test set	1311 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å ²)	31.2	Xtriage
Anisotropy	1.470	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$, $B_{sol}(Å^2)$	$0.31 \; , 6.9$	EDS
L-test for twinning ²	$< L > = 0.49, < L^2> = 0.33$	Xtriage
	0.027 for -h,-k,l	
Estimated twinning fraction	0.488 for h,-h-k,-l	Xtriage
	0.028 for -k,-h,-l	
F_o, F_c correlation	0.95	EDS
Total number of atoms	5347	wwPDB-VP
Average B, all atoms (Å ²)	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 4.98% of the height of the origin peak. No significant pseudotranslation is detected.

²Theoretical 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.



¹Intensities estimated from amplitudes.

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: ZN, ACT, NI

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		nd angles
MIOI	Chain	RMSZ	# Z > 5	RMSZ	# Z > 5
1	A	0.69	$1/2678 \; (0.0\%)$	0.74	0/3642
1	В	0.70	$1/2673 \ (0.0\%)$	0.74	2/3638 (0.1%)
All	All	0.69	$2/5351 \ (0.0\%)$	0.74	2/7280 (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	В	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$\operatorname{Observed}(\text{\AA})$	Ideal(A)
1	A	194	GLU	CD-OE2	5.99	1.32	1.25
1	В	194	GLU	CD-OE2	5.54	1.31	1.25

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	\mathbf{Z}	$Observed(^o)$	$\operatorname{Ideal}({}^o)$
1	В	263	ARG	NE-CZ-NH2	-5.39	117.61	120.30
1	В	263	ARG	NE-CZ-NH1	5.07	122.83	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

\mathbf{Mol}	Chain	Res	Type	Group
1	В	318	THR	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	2595	0	2421	18	2
1	В	2587	0	2404	33	2
2	A	1	0	0	0	0
2	В	1	0	0	0	0
3	A	1	0	0	0	0
3	В	1	0	0	0	0
4	A	8	0	6	3	0
5	A	73	0	0	4	0
5	В	80	0	0	11	0
All	All	5347	0	4831	50	2

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 5.

All (50) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	$\begin{array}{c} \text{Interatomic} \\ \text{distance (Å)} \end{array}$	Clash overlap (Å)
1:B:238:CYS:SG	1:B:310:CYS:HA	1.94	1.08
1:B:289:CYS:HB3	5:B:580:HOH:O	1.76	0.84
1:B:238:CYS:SG	1:B:310:CYS:CA	2.65	0.81
1:B:218:GLU:HB3	5:B:577:HOH:O	1.81	0.79
1:B:238:CYS:SG	1:B:310:CYS:CB	2.74	0.75
1:B:289:CYS:O	5:B:580:HOH:O	2.11	0.69
1:A:262:ASN:OD1	4:A:403:ACT:C	2.41	0.68
1:B:218:GLU:CB	5:B:577:HOH:O	2.39	0.68
1:B:341:ASP:O	5:B:524:HOH:O	2.15	0.65
1:A:194:GLU:OE2	1:A:280:HIS:CE1	2.50	0.64
1:B:206:LEU:N	5:B:580:HOH:O	2.31	0.64
1:B:206:LEU:HG	5:B:580:HOH:O	2.01	0.60
1:B:48:LEU:HD13	1:B:274:THR:HG22	1.85	0.58
1:B:48:LEU:HD12	1:B:273:VAL:O	2.06	0.56
1:B:194:GLU:OE2	1:B:280:HIS:CE1	2.59	0.54
1:B:219:HIS:CE1	1:B:260:PRO:HG2	2.43	0.54
1:A:17:ASN:OD1	4:A:403:ACT:H1	2.09	0.52
1:B:88:GLN:OE1	1:B:129:ARG:NH2	2.42	0.52

Continued on next page...



 $Continued\ from\ previous\ page...$

A	_	Interatomic	Clash
Atom-1	Atom-2	distance (Å)	overlap (Å)
1:A:210:LYS:HE3	5:A:541:HOH:O	2.09	0.51
1:A:186:LYS:HE2	5:A:548:HOH:O	2.11	0.49
1:A:21:PHE:CE1	1:A:38:MET:HG2	2.48	0.49
1:B:317:VAL:CB	5:B:522:HOH:O	2.62	0.48
1:A:300:TRP:O	1:A:300:TRP:CG	2.66	0.47
1:A:238:CYS:HA	1:A:312:CYS:HB3	1.97	0.47
1:B:194:GLU:HG3	1:B:200:SER:HB3	1.97	0.46
1:B:289:CYS:CB	5:B:580:HOH:O	2.46	0.46
1:B:12:GLN:HE21	1:B:12:GLN:HA	1.81	0.46
1:A:154:LEU:HD12	1:A:291:GLU:HG2	1.98	0.46
1:A:18:ILE:HD11	1:A:214:VAL:CG1	2.45	0.46
1:A:130:ILE:HG12	5:A:556:HOH:O	2.15	0.46
1:A:210:LYS:CE	5:A:541:HOH:O	2.63	0.45
1:A:18:ILE:HD11	1:A:214:VAL:HG11	1.98	0.45
1:A:335:LEU:HD21	1:B:302:ASP:HB3	1.99	0.45
1:B:233:GLY:O	1:B:236:ARG:CB	2.64	0.45
1:B:130:ILE:HG12	5:B:559:HOH:O	2.17	0.44
1:B:219:HIS:ND1	1:B:260:PRO:HG2	2.32	0.44
1:B:155:GLY:HA3	5:B:518:HOH:O	2.16	0.44
1:A:18:ILE:HD13	1:A:48:LEU:HD23	1.98	0.44
1:B:183:GLY:O	1:B:287:PHE:HA	2.17	0.44
1:B:300:TRP:O	1:B:300:TRP:CG	2.72	0.43
1:A:71:ILE:HG23	1:A:140:ILE:HD11	2.01	0.42
1:B:160:ILE:O	1:B:163:LEU:HB2	2.20	0.42
1:A:262:ASN:OD1	4:A:403:ACT:O	2.37	0.42
1:B:234:SER:HB3	1:B:241:PHE:CD1	2.54	0.42
1:B:53:PRO:HA	1:B:54:PRO:HD3	1.92	0.42
1:B:238:CYS:SG	1:B:310:CYS:HB2	2.55	0.41
1:B:275:PHE:O	1:B:276:PRO:C	2.59	0.41
1:B:317:VAL:O	1:B:318:THR:HB	2.19	0.41
1:A:211:THR:O	1:A:282:GLY:HA3	2.21	0.41
1:B:80:SER:O	1:B:86:PHE:HA	2.21	0.40

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	$egin{aligned} ext{Interatomic} \ ext{distance} \ (ext{Å}) \end{aligned}$	Clash overlap (Å)
1:A:22:HIS:CD2	1:B:318:THR:CG2[2_554]	1.75	0.45
1:A:22:HIS:NE2	1:B:318:THR:CG2[2_554]	2.11	0.09



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	Perce	entiles
1	A	324/330~(98%)	313 (97%)	11 (3%)	0	100	100
1	В	330/330 (100%)	312 (94%)	17 (5%)	1 (0%)	41	61
All	All	654/660 (99%)	625 (96%)	28 (4%)	1 (0%)	47	68

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	В	237	GLY

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	$258/280 \ (92\%)$	249 (96%)	9 (4%)	36 62
1	В	255/280~(91%)	245 (96%)	10 (4%)	32 57
All	All	513/560 (92%)	494 (96%)	19 (4%)	34 60

All (19) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	28	PHE
1	A	69	ILE
1	A	82	ARG
1	A	96	MET
1	A	136	TYR

Continued on next page...



Continued from previous page...

Mol	Chain	Res	Type
1	A	146	ASP
1	A	179	TYR
1	A	263	ARG
1	A	314	GLU
1	В	12	GLN
1	В	28	PHE
1	В	70	LEU
1	В	82	ARG
1	В	136	TYR
1	В	163	LEU
1	В	179	TYR
1	В	186	LYS
1	В	263	ARG
1	В	340	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	340	GLN
1	В	12	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 6 ligands modelled in this entry, 4 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	Trino	Chain	Res	Link	Bond lengths			Bond angles		
	туре				Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
4	ACT	A	403	-	3,3,3	0.84	0	3,3,3	1.67	1 (33%)
4	ACT	A	404	-	3,3,3	0.67	0	3,3,3	1.32	0

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^o)$	$\operatorname{Ideal}({}^{o})$
4	A	403	ACT	OXT-C-CH3	2.19	124.25	115.18

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	403	ACT	3	0

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^{th} 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 $>$	$\#RSRZ{>}2$		Z>2	$OWAB(A^2)$	Q<0.9
1	A	327/330~(99%)	-0.60	0	100	100	23, 38, 58, 68	1 (0%)
1	В	330/330 (100%)	-0.60	0	100	100	23, 38, 59, 74	1 (0%)
All	All	657/660 (99%)	-0.60	0	100	100	23, 38, 59, 74	2 (0%)

There are no RSRZ outliers to report.

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^{th} 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	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q<0.9
3	ZN	В	402	1/1	0.63	0.18	57,57,57,57	0
4	ACT	A	404	4/4	0.94	0.09	53,55,55,56	0
4	ACT	A	403	4/4	0.95	0.13	39,40,41,44	0
3	ZN	A	402	1/1	0.96	0.17	43,43,43,43	0
2	NI	A	401	1/1	0.97	0.15	36,36,36,36	0
2	NI	В	401	1/1	0.97	0.15	38,38,38,38	0



6.5 Other polymers (i)

There are no such residues in this entry.

