



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 12:07 pm GMT

PDB ID : 1M7W  
Title : HNF4a ligand binding domain with bound fatty acid  
Authors : Dhe-Paganon, S.; Duda, K.; Iwamoto, M.; Chi, Y.I.; Shoelson, S.E.  
Deposited on : 2002-07-22  
Resolution : 2.80 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

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.7.2 (RC1), CSD as538be (2017)  
Xtrriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : recalc28949

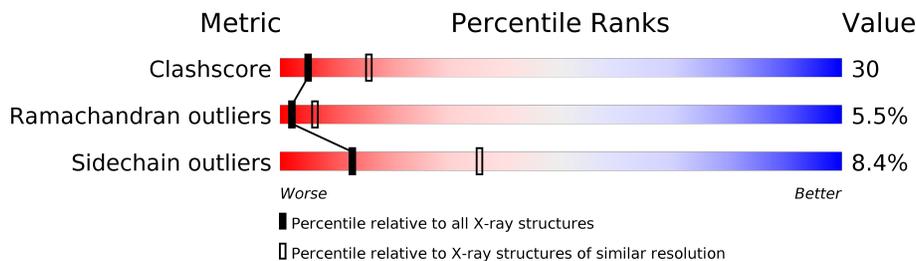
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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))
Clashscore	112137	3033 (2.80-2.80)
Ramachandran outliers	110173	2983 (2.80-2.80)
Sidechain outliers	110143	2985 (2.80-2.80)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	250	
1	B	250	
1	C	250	
1	D	250	

## 2 Entry composition [i](#)

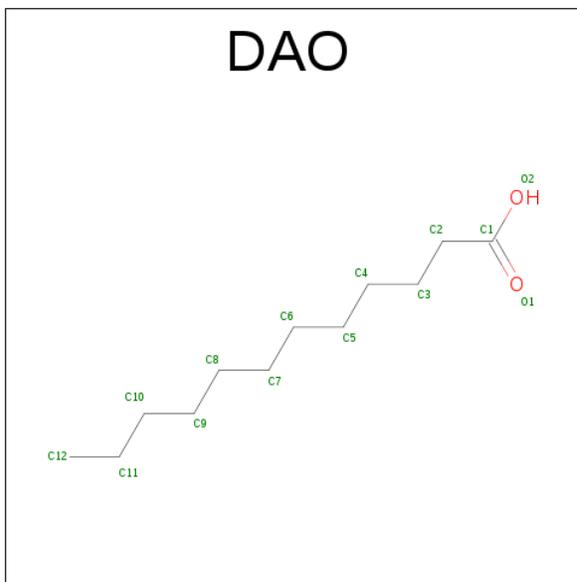
There are 3 unique types of molecules in this entry. The entry contains 7176 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Hepatocyte nuclear factor 4-alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	223	Total 1777	C 1140	N 298	O 329	S 10	0	0	0
1	B	218	Total 1745	C 1122	N 293	O 320	S 10	0	0	0
1	C	220	Total 1759	C 1131	N 295	O 323	S 10	0	0	0
1	D	221	Total 1765	C 1134	N 296	O 325	S 10	0	0	0

- Molecule 2 is LAURIC ACID (three-letter code: DAO) (formula: C<sub>12</sub>H<sub>24</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
2	A	1	Total 14	C 12	O 2	0	0
2	B	1	Total 14	C 12	O 2	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	C	1	Total	C	O	0	0
			14	12	2		
2	D	1	Total	C	O	0	0
			14	12	2		

- Molecule 3 is water.

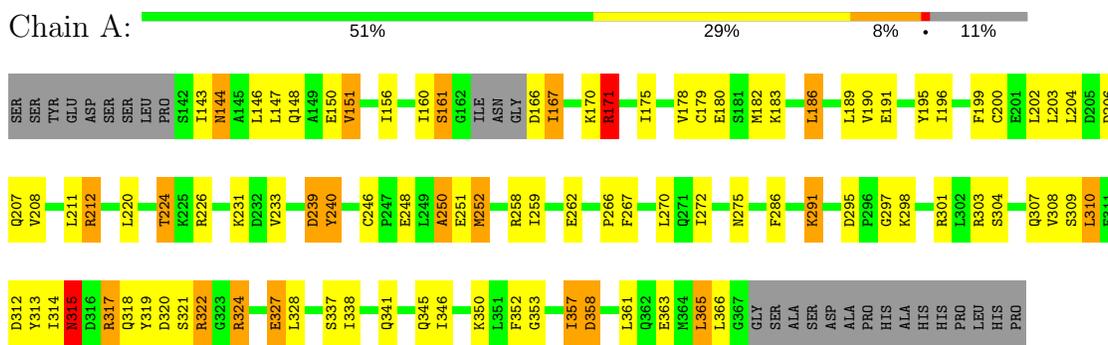
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	17	Total	O	0	0
			17	17		
3	B	22	Total	O	0	0
			22	22		
3	C	19	Total	O	0	0
			19	19		
3	D	16	Total	O	0	0
			16	16		

### 3 Residue-property plots

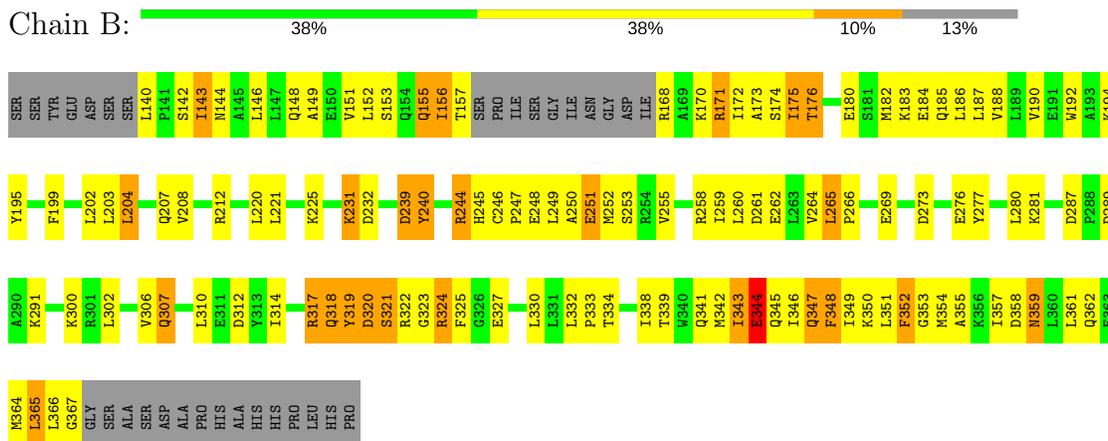
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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.

Note EDS was not executed.

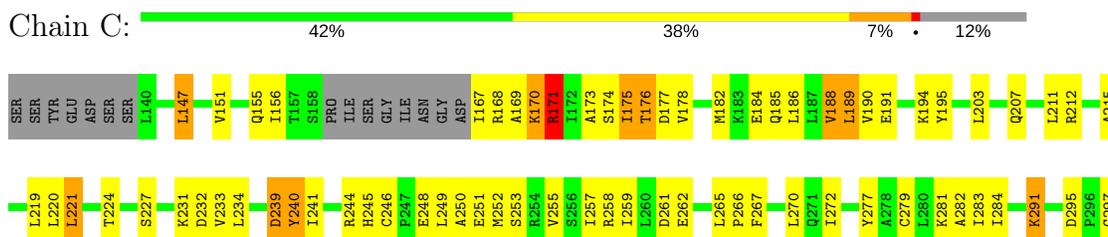
- Molecule 1: Hepatocyte nuclear factor 4-alpha

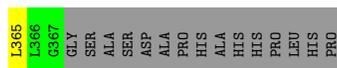


- Molecule 1: Hepatocyte nuclear factor 4-alpha



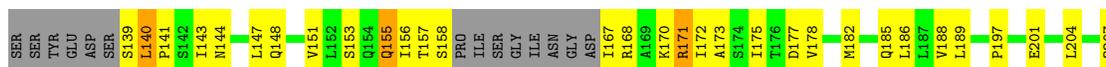
- Molecule 1: Hepatocyte nuclear factor 4-alpha





● Molecule 1: Hepatocyte nuclear factor 4-alpha

Chain D: 43% 39% 6% 12%



## 4 Data and refinement statistics

Xtrriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	102.29Å 102.29Å 227.69Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.80	Depositor
% Data completeness (in resolution range)	(Not available) (20.00-2.80)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	CNS	Depositor
R, $R_{free}$	0.237 , 0.282	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	7176	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	55.0	wwPDB-VP

## 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: DAO

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	0.40	0/1803	0.63	0/2435
1	B	0.40	0/1771	0.64	0/2392
1	C	0.41	0/1785	0.68	1/2411 (0.0%)
1	D	0.39	0/1791	0.63	0/2419
All	All	0.40	0/7150	0.65	1/9657 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	347	GLN	N-CA-C	-5.43	96.33	111.00

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	1777	0	1834	89	0
1	B	1745	0	1806	119	0
1	C	1759	0	1822	119	0
1	D	1765	0	1827	117	0
2	A	14	0	23	1	0
2	B	14	0	23	3	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	14	0	23	1	0
2	D	14	0	23	1	0
3	A	17	0	0	3	0
3	B	22	0	0	3	0
3	C	19	0	0	3	0
3	D	16	0	0	3	0
All	All	7176	0	7381	435	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 30.

The worst 5 of 435 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:D:340:TRP:NE1	3:D:48:HOH:O	1.86	1.07
1:C:291:LYS:HD2	1:C:291:LYS:H	1.24	1.01
1:C:324:ARG:HG2	1:C:324:ARG:HH11	1.25	1.01
1:D:356:LYS:HA	1:D:356:LYS:HE3	1.41	1.00
1:C:170:LYS:O	1:C:171:ARG:HB2	1.64	0.98

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	219/250 (88%)	189 (86%)	19 (9%)	11 (5%)	<b>2</b> <b>7</b>
1	B	214/250 (86%)	180 (84%)	19 (9%)	15 (7%)	<b>1</b> <b>3</b>
1	C	216/250 (86%)	186 (86%)	20 (9%)	10 (5%)	<b>3</b> <b>9</b>
1	D	217/250 (87%)	183 (84%)	22 (10%)	12 (6%)	<b>2</b> <b>6</b>
All	All	866/1000 (87%)	738 (85%)	80 (9%)	48 (6%)	<b>2</b> <b>6</b>

5 of 48 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	161	SER
1	A	250	ALA
1	A	358	ASP
1	B	143	ILE
1	B	175	ILE

### 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	198/220 (90%)	177 (89%)	21 (11%)	8 23
1	B	194/220 (88%)	178 (92%)	16 (8%)	13 37
1	C	196/220 (89%)	179 (91%)	17 (9%)	12 33
1	D	197/220 (90%)	185 (94%)	12 (6%)	22 53
All	All	785/880 (89%)	719 (92%)	66 (8%)	13 35

5 of 66 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	312	ASP
1	B	365	LEU
1	D	305	GLN
1	B	317	ARG
1	B	348	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	154	GLN
1	C	341	GLN
1	D	345	GLN
1	C	268	GLN
1	C	315	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

4 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	$\# Z  > 2$
2	DAO	A	700	-	10,13,13	0.38	0	9,13,13	0.30	0
2	DAO	B	701	-	10,13,13	0.29	0	9,13,13	0.42	0
2	DAO	C	702	-	10,13,13	0.24	0	9,13,13	0.34	0
2	DAO	D	703	-	10,13,13	0.24	0	9,13,13	0.29	0

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	DAO	A	700	-	-	0/9/11/11	0/0/0/0
2	DAO	B	701	-	-	0/9/11/11	0/0/0/0
2	DAO	C	702	-	-	0/9/11/11	0/0/0/0
2	DAO	D	703	-	-	0/9/11/11	0/0/0/0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	700	DAO	1	0
2	B	701	DAO	3	0
2	C	702	DAO	1	0
2	D	703	DAO	1	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](#)

EDS was not executed - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates [i](#)

EDS was not executed - this section is therefore empty.

### 6.4 Ligands [i](#)

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers [i](#)

EDS was not executed - this section is therefore empty.