



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 24, 2018 – 02:52 AM EST

PDB ID : 1DIB  
Title : HUMAN METHYLENETETRAHYDROFOLATE DEHYDROGENASE  
/ CYCLOHYDROLASE COMPLEXED WITH NADP AND INHIBITOR  
LY345899  
Authors : Schmidt, A.; Wu, H.; MacKenzie, R.E.; Chen, V.J.; Bewly, J.R.; Ray, J.E.;  
Toth, J.E.; Cygler, M.  
Deposited on : 1999-11-29  
Resolution : 2.70 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)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20030736  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030736

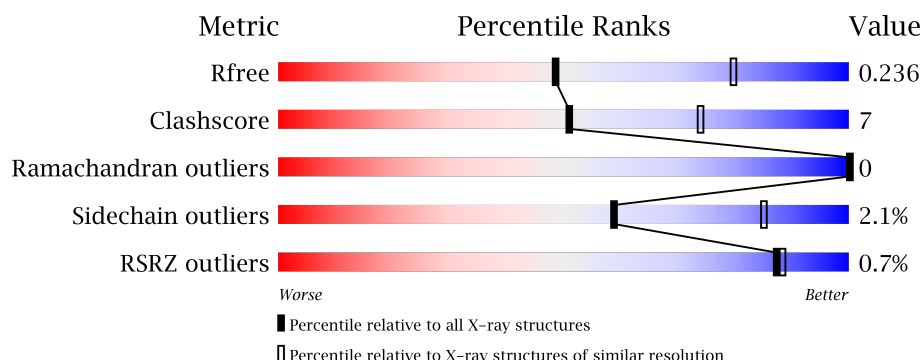
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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))
$R_{free}$	100719	2259 (2.70-2.70)
Clashscore	112137	2590 (2.70-2.70)
Ramachandran outliers	110173	2550 (2.70-2.70)
Sidechain outliers	110143	2550 (2.70-2.70)
RSRZ outliers	101464	2275 (2.70-2.70)

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.

Mol	Chain	Length	Quality of chain
1	A	306	
1	B	306	

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
2	NAP	A	2001	-	-	-	X
2	NAP	B	2002	-	-	-	X
3	L34	A	3001	-	-	-	X

2 Entry composition ⓘ

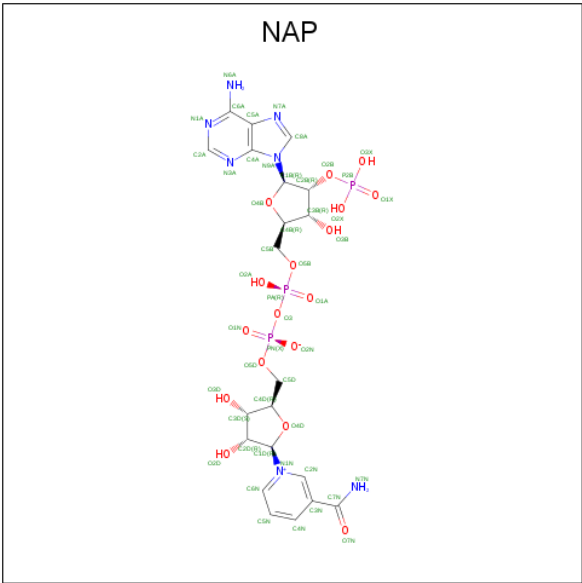
There are 4 unique types of molecules in this entry. The entry contains 4629 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 METHYLENETETRAHYDROFOLATE DEHYDROGENASE/CYCLOHYDROLASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	285	Total	C	N	O	S	4	0	0
			2152	1351	374	416	11			
1	B	295	Total	C	N	O	S	4	0	0
			2237	1404	390	432	11			

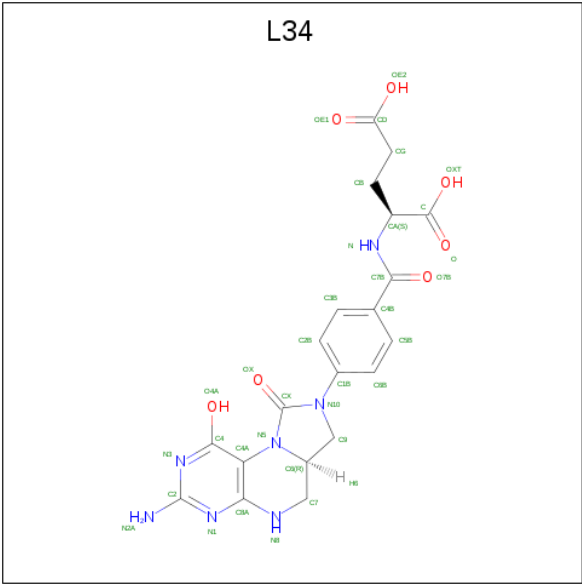
- Molecule 2 is NADP NICOTINAMIDE-ADENINE-DINUCLEOTIDE PHOSPHATE (three-letter code: NAP) (formula: C<sub>21</sub>H<sub>28</sub>N<sub>7</sub>O<sub>17</sub>P<sub>3</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			48	21	7	17	3		
2	B	1	Total	C	N	O	P	0	0
			48	21	7	17	3		

- Molecule 3 is 4-(7-AMINO-9-HYDROXY-1-OXO-3,3A,4,5-TETRAHYDRO-2,5,6,8,9B-PENTAAZA-CYCLOPENTA[A]NAPHTHALEN-2-YL)-PHENYL CARBONYL-GLUTAMIC

ACID (three-letter code: L34) (formula: C<sub>20</sub>H<sub>21</sub>N<sub>7</sub>O<sub>7</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			34	20	7	7		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	52	Total	O	0	0
			52	52		
4	B	58	Total	O	0	0
			58	58		



- Molecule 1: METHYLENETETRAHYDROFOLATE DEHYDROGENASE/CYCLOHYDROLASE



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	67.49 Å   136.37 Å   61.38 Å 90.00°   90.00°   90.00°	Depositor
Resolution (Å)	15.00 – 2.70 19.60 – 2.70	Depositor EDS
% Data completeness (in resolution range)	97.1 (15.00-2.70) 94.7 (19.60-2.70)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.12 (at 2.71 Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.208   ,   0.240 0.210   ,   0.236	Depositor DCC
$R_{free}$ test set	1545 reflections (10.17%)	DCC
Wilson B-factor (Å <sup>2</sup> )	25.9	Xtriage
Anisotropy	0.139	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 28.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	4629	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.90% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NAP, L34

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	0.38	0/2183	0.65	0/2962
1	B	0.38	0/2272	0.63	0/3084
All	All	0.38	0/4455	0.64	0/6046

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	2152	0	2201	35	0
1	B	2237	0	2288	24	0
2	A	48	0	25	6	0
2	B	48	0	25	4	0
3	A	34	0	19	5	0
4	A	52	0	0	2	0
4	B	58	0	0	3	0
All	All	4629	0	4558	59	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 7.



The worst 5 of 59 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:37:ARG:H	1:A:95:HIS:HD2	1.28	0.80
1:A:44:GLY:HA3	1:A:104:ASP:HB3	1.71	0.70
1:A:37:ARG:H	1:A:95:HIS:CD2	2.10	0.69
1:A:237:GLY:HA2	2:A:2001:NAP:O3D	2.00	0.60
1:B:1134:ARG:HD3	4:B:4065:HOH:O	2.03	0.59

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	281/306 (92%)	273 (97%)	8 (3%)	0	100	100
1	B	293/306 (96%)	283 (97%)	10 (3%)	0	100	100
All	All	574/612 (94%)	556 (97%)	18 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	238/258 (92%)	231 (97%)	7 (3%)	48	77

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	248/258 (96%)	245 (99%)	3 (1%)	75	92
All	All	486/516 (94%)	476 (98%)	10 (2%)	59	85

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	142	ASP
1	A	251	LYS
1	B	1051	LEU
1	A	125	ASP
1	A	259	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	22	ASN
1	A	95	HIS
1	B	1022	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](#)

3 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	NAP	A	2001	-	44,52,52	1.98	5 (11%)	51,80,80	1.98	7 (13%)
3	L34	A	3001	-	28,37,37	1.85	7 (25%)	33,54,54	2.51	11 (33%)
2	NAP	B	2002	-	44,52,52	1.93	4 (9%)	51,80,80	1.86	4 (7%)

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	NAP	A	2001	-	-	0/27/67/67	0/5/5/5
3	L34	A	3001	-	-	0/15/46/46	0/3/4/4
2	NAP	B	2002	-	-	0/27/67/67	0/5/5/5

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	2001	NAP	C5A-C4A	-6.01	1.26	1.40
2	B	2002	NAP	C5A-C4A	-5.74	1.27	1.40
3	A	3001	L34	C4A-C8A	-4.19	1.32	1.41
3	A	3001	L34	CX-N10	-3.77	1.34	1.38
3	A	3001	L34	CX-N5	-2.96	1.33	1.38

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	2002	NAP	N3A-C2A-N1A	-11.16	119.14	128.86
2	A	2001	NAP	N3A-C2A-N1A	-11.09	119.20	128.86
3	A	3001	L34	N3-C2-N1	-4.25	118.57	125.45
3	A	3001	L34	CB-CA-C	-4.02	106.44	112.28
3	A	3001	L34	C9-N10-C1B	-4.01	114.24	121.44

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	2001	NAP	6	0
3	A	3001	L34	5	0
2	B	2002	NAP	4	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<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	285/306 (93%)	-0.34	2 (0%) 87 88	8, 20, 33, 50	1 (0%)
1	B	295/306 (96%)	-0.38	2 (0%) 87 88	9, 19, 32, 45	1 (0%)
All	All	580/612 (94%)	-0.36	4 (0%) 87 88	8, 19, 33, 50	2 (0%)

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	2	ALA	4.1
1	A	251	LYS	2.5
1	B	1030	GLN	2.3
1	B	1002	ALA	2.2

### 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 carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	NAP	B	2002	48/48	0.75	0.60	12.20	24,37,45,49	48
2	NAP	A	2001	48/48	0.75	0.64	9.40	13,28,37,42	48
3	L34	A	3001	34/34	0.73	0.48	7.43	13,15,17,17	34

## 6.5 Other polymers [i](#)

There are no such residues in this entry.