



# Full wwPDB X-ray Structure Validation Report ⓘ

May 25, 2020 – 07:02 am BST

PDB ID : 1HF8  
Title : CALM-N N-terminal domain of clathrin assembly lymphoid myeloid leukaemia protein  
Authors : Ford, M.G.J.; Evans, P.R.; McMahon, H.T.  
Deposited on : 2000-11-30  
Resolution : 2.00 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](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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
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) : 2.11

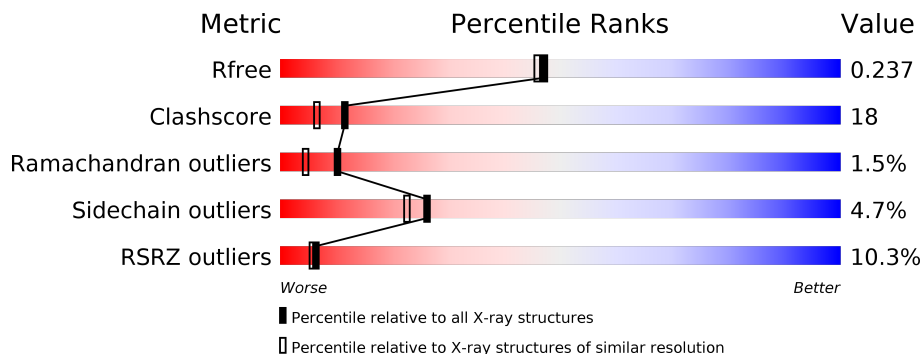
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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 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	289	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2244 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 CLATHRIN ASSEMBLY PROTEIN SHORT FORM.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	263	2114	1345	363	393	13	0	0	0

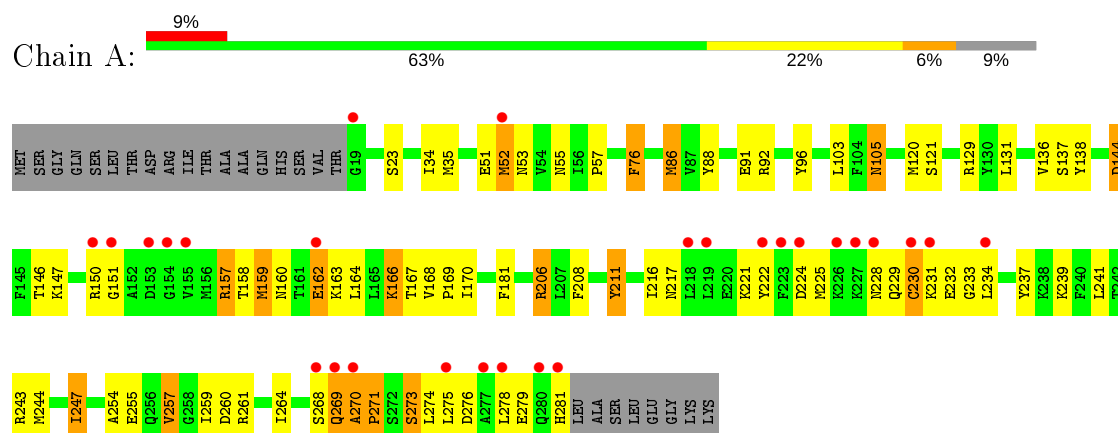
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	130	Total	O	0	0
			130	130		

### 3 Residue-property plots [i](#)

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.

- Molecule 1: CLATHRIN ASSEMBLY PROTEIN SHORT FORM



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	77.93Å 77.93Å 121.85Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	65.94 – 2.00 36.02 – 2.00	Depositor EDS
% Data completeness (in resolution range)	98.7 (65.94-2.00) 100.0 (36.02-2.00)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.18 (at 2.00Å)	Xtrriage
Refinement program	REFMAC 5.0	Depositor
R, $R_{free}$	0.190 , 0.223 0.200 , 0.237	Depositor DCC
$R_{free}$ test set	1324 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	35.2	Xtrriage
Anisotropy	0.530	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 54.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	2244	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	48.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.49% 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

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	1.46	15/2150 (0.7%)	1.31	21/2896 (0.7%)

All (15) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	86	MET	SD-CE	-13.32	1.03	1.77
1	A	211	TYR	CD1-CE1	9.01	1.52	1.39
1	A	206	ARG	CG-CD	-7.99	1.31	1.51
1	A	257	VAL	CB-CG2	-7.46	1.37	1.52
1	A	96	TYR	CD2-CE2	-6.80	1.29	1.39
1	A	96	TYR	CD1-CE1	-6.27	1.29	1.39
1	A	255	GLU	CD-OE2	-5.79	1.19	1.25
1	A	92	ARG	CG-CD	5.78	1.66	1.51
1	A	121	SER	CA-CB	5.77	1.61	1.52
1	A	264	ILE	CA-CB	-5.48	1.42	1.54
1	A	35	MET	SD-CE	-5.37	1.47	1.77
1	A	52	MET	SD-CE	5.29	2.07	1.77
1	A	136	VAL	CB-CG2	-5.29	1.41	1.52
1	A	257	VAL	C-O	-5.10	1.13	1.23
1	A	88	TYR	CB-CG	-5.04	1.44	1.51

All (21) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	206	ARG	CA-CB-CG	-7.25	97.44	113.40
1	A	261	ARG	NE-CZ-NH1	7.25	123.92	120.30
1	A	211	TYR	CB-CG-CD2	-6.81	116.92	121.00
1	A	268	SER	N-CA-C	-6.59	93.21	111.00
1	A	181	PHE	CB-CG-CD2	-6.42	116.30	120.80
1	A	260	ASP	CB-CG-OD1	6.38	124.04	118.30
1	A	129	ARG	NE-CZ-NH2	-6.37	117.11	120.30
1	A	144	ASP	CB-CG-OD2	6.33	124.00	118.30
1	A	76	PHE	CB-CG-CD1	6.30	125.21	120.80

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	211	TYR	CB-CG-CD1	6.21	124.73	121.00
1	A	76	PHE	CB-CG-CD2	-6.05	116.56	120.80
1	A	224	ASP	CB-CG-OD2	5.71	123.44	118.30
1	A	268	SER	CB-CA-C	5.54	120.62	110.10
1	A	34	ILE	CG1-CB-CG2	-5.47	99.37	111.40
1	A	129	ARG	NE-CZ-NH1	5.46	123.03	120.30
1	A	120	MET	CG-SD-CE	5.40	108.84	100.20
1	A	103	LEU	CB-CG-CD2	-5.39	101.84	111.00
1	A	35	MET	CG-SD-CE	5.39	108.82	100.20
1	A	86	MET	CG-SD-CE	-5.30	91.72	100.20
1	A	247	ILE	CG1-CB-CG2	-5.17	100.03	111.40
1	A	131	LEU	CB-CG-CD2	5.14	119.74	111.00

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	2114	0	2137	75	0
2	A	130	0	0	6	1
All	All	2244	0	2137	75	1

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 18.

All (75) 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:52:MET:SD	1:A:52:MET:CE	2.07	1.42
1:A:86:MET:CG	1:A:86:MET:CE	2.08	1.30
1:A:86:MET:SD	1:A:86:MET:CE	1.03	1.13
1:A:86:MET:SD	1:A:86:MET:HE2	1.63	1.05
1:A:86:MET:SD	1:A:86:MET:HE3	1.63	1.02
1:A:86:MET:SD	1:A:86:MET:HE1	1.63	1.01

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:86:MET:HE2	1:A:86:MET:HG3	1.56	0.85
1:A:244:MET:HE3	1:A:244:MET:HA	1.61	0.82
1:A:208:PHE:HE1	1:A:244:MET:CE	1.94	0.81
1:A:86:MET:CG	1:A:86:MET:HE2	1.85	0.81
1:A:208:PHE:CE1	1:A:244:MET:HE2	2.16	0.80
1:A:225:MET:HE2	1:A:230:CYS:HA	1.62	0.80
1:A:279:GLU:O	1:A:279:GLU:HG2	1.82	0.79
1:A:159:MET:HE2	1:A:163:LYS:HG3	1.66	0.78
1:A:208:PHE:CE1	1:A:244:MET:CE	2.70	0.74
1:A:159:MET:HE3	1:A:163:LYS:HG2	1.72	0.72
1:A:208:PHE:HE1	1:A:244:MET:HE1	1.55	0.72
1:A:243:ARG:O	1:A:247:ILE:HG13	1.91	0.71
1:A:228:ASN:O	1:A:231:LYS:HB3	1.92	0.70
1:A:274:LEU:O	1:A:274:LEU:HG	1.92	0.70
1:A:164:LEU:HD22	1:A:217:ASN:ND2	2.09	0.67
1:A:105:ASN:HD22	1:A:105:ASN:C	1.98	0.66
1:A:151:GLY:O	1:A:157:ARG:HD3	1.96	0.65
1:A:270:ALA:N	1:A:271:PRO:HD3	2.13	0.63
1:A:158:THR:O	1:A:158:THR:HG22	2.00	0.61
1:A:147:LYS:HE3	2:A:2083:HOH:O	2.02	0.59
1:A:225:MET:CE	1:A:230:CYS:HA	2.33	0.59
1:A:144:ASP:OD1	1:A:146:THR:OG1	2.21	0.58
1:A:159:MET:CE	1:A:163:LYS:CG	2.81	0.58
1:A:237:TYR:CE1	1:A:241:LEU:HD11	2.38	0.58
1:A:105:ASN:C	1:A:105:ASN:ND2	2.57	0.57
1:A:159:MET:CE	1:A:163:LYS:HG3	2.35	0.56
1:A:168:VAL:N	1:A:169:PRO:CD	2.70	0.55
1:A:229:GLN:O	1:A:232:GLU:N	2.40	0.54
1:A:159:MET:CE	1:A:163:LYS:HG2	2.39	0.52
1:A:163:LYS:O	1:A:166:LYS:HB2	2.11	0.51
1:A:159:MET:HE3	1:A:163:LYS:CG	2.39	0.50
1:A:160:ASN:OD1	1:A:162:GLU:N	2.43	0.50
1:A:229:GLN:O	1:A:230:CYS:C	2.51	0.49
1:A:222:TYR:HE2	1:A:276:ASP:HB2	1.78	0.48
1:A:91:GLU:OE2	1:A:138:TYR:OH	2.27	0.48
1:A:233:GLY:O	1:A:234:LEU:C	2.50	0.48
1:A:222:TYR:CE2	1:A:276:ASP:HB2	2.49	0.48
1:A:55:ASN:OD1	1:A:57:PRO:HD2	2.14	0.48
1:A:147:LYS:CE	2:A:2083:HOH:O	2.59	0.48
1:A:276:ASP:C	1:A:278:LEU:H	2.16	0.47
1:A:167:THR:O	1:A:170:ILE:N	2.44	0.46

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:229:GLN:O	1:A:231:LYS:N	2.48	0.46
1:A:86:MET:CB	1:A:86:MET:CE	2.88	0.46
1:A:254:ALA:HB1	1:A:259:ILE:HD12	1.97	0.46
1:A:51:GLU:HB3	2:A:2018:HOH:O	2.15	0.46
1:A:274:LEU:O	1:A:274:LEU:CG	2.63	0.45
1:A:232:GLU:O	1:A:233:GLY:C	2.55	0.45
1:A:160:ASN:ND2	2:A:2088:HOH:O	2.44	0.45
1:A:162:GLU:HG2	1:A:163:LYS:N	2.32	0.44
1:A:53:ASN:HB2	2:A:2018:HOH:O	2.17	0.44
1:A:257:VAL:O	1:A:257:VAL:CG1	2.65	0.43
1:A:244:MET:CE	1:A:244:MET:HA	2.35	0.43
1:A:160:ASN:OD1	1:A:160:ASN:C	2.57	0.43
1:A:150:ARG:NH2	2:A:2086:HOH:O	2.50	0.43
1:A:229:GLN:C	1:A:231:LYS:N	2.71	0.43
1:A:276:ASP:C	1:A:278:LEU:N	2.72	0.42
1:A:206:ARG:HD3	1:A:206:ARG:HH11	1.68	0.42
1:A:217:ASN:OD1	1:A:221:LYS:HD2	2.20	0.42
1:A:239:LYS:O	1:A:243:ARG:HD2	2.19	0.42
1:A:234:LEU:HD12	1:A:234:LEU:HA	1.76	0.42
1:A:157:ARG:HH21	1:A:216:ILE:HG21	1.85	0.41
1:A:273:SER:C	1:A:275:LEU:H	2.24	0.41
1:A:273:SER:O	1:A:275:LEU:N	2.53	0.41
1:A:105:ASN:ND2	1:A:105:ASN:O	2.45	0.41
1:A:208:PHE:CZ	1:A:244:MET:HE2	2.54	0.41
1:A:150:ARG:NH2	1:A:206:ARG:HD2	2.36	0.41
1:A:269:GLN:C	1:A:271:PRO:HD3	2.41	0.41
1:A:208:PHE:CE1	1:A:244:MET:HE1	2.42	0.40
1:A:157:ARG:HH11	1:A:157:ARG:HD3	1.71	0.40

All (1) 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	Interatomic distance (Å)	Clash overlap (Å)
2:A:2044:HOH:O	2:A:2044:HOH:O[7_556]	2.02	0.18

## 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	261/289 (90%)	242 (93%)	15 (6%)	4 (2%)	10 4

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	270	ALA
1	A	230	CYS
1	A	273	SER
1	A	271	PRO

### 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	235/256 (92%)	224 (95%)	11 (5%)	26 22

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	23	SER
1	A	76	PHE
1	A	105	ASN
1	A	137	SER
1	A	157	ARG
1	A	159	MET
1	A	162	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	166	LYS
1	A	211	TYR
1	A	269	GLN
1	A	281	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	95	GLN
1	A	105	ASN
1	A	140	GLN
1	A	212	ASN
1	A	269	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 carbohydrates in this entry.

### 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

### 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

### 6.1 Protein, DNA and RNA chains

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	263/289 (91%)	0.27	27 (10%) <b>6</b> <b>6</b>	22, 41, 90, 125	0

All (27) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	281	HIS	6.3
1	A	224	ASP	4.4
1	A	230	CYS	4.2
1	A	19	GLY	4.1
1	A	234	LEU	4.1
1	A	223	PHE	3.5
1	A	280	GLN	3.4
1	A	278	LEU	3.4
1	A	227	LYS	3.1
1	A	155	VAL	3.0
1	A	153	ASP	2.9
1	A	222	TYR	2.9
1	A	269	GLN	2.9
1	A	219	LEU	2.8
1	A	231	LYS	2.5
1	A	275	LEU	2.5
1	A	228	ASN	2.5
1	A	277	ALA	2.5
1	A	150	ARG	2.4
1	A	52	MET	2.4
1	A	268	SER	2.3
1	A	154	GLY	2.3
1	A	162	GLU	2.3
1	A	270	ALA	2.2
1	A	226	LYS	2.2
1	A	151	GLY	2.2
1	A	218	LEU	2.0

## 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](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.