



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 5, 2024 – 10:27 PM EDT

PDB ID : 2GNX  
Title : X-ray structure of a hypothetical protein from Mouse Mm.209172  
Authors : Phillips Jr., G.N.; McCoy, J.G.; Bitto, E.; Wesenberg, G.E.; Bingman, C.A.;  
Center for Eukaryotic Structural Genomics (CESG)  
Deposited on : 2006-04-11  
Resolution : 2.45 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 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	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

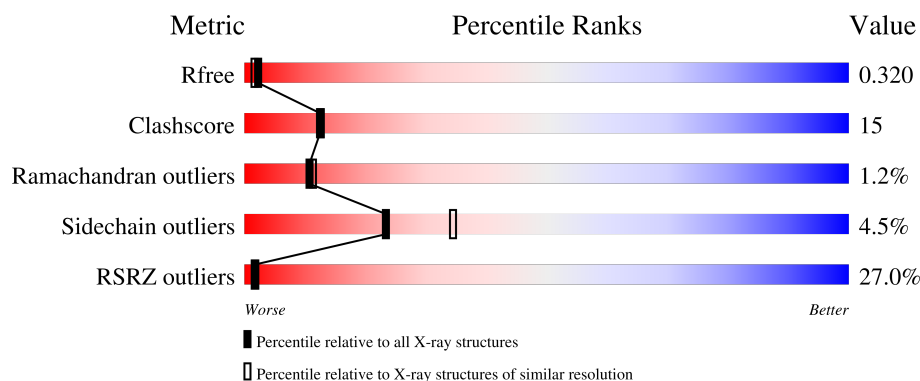
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.45 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))
$R_{free}$	164625	1096 (2.46-2.46)
Clashscore	180529	1178 (2.46-2.46)
Ramachandran outliers	177936	1170 (2.46-2.46)
Sidechain outliers	177891	1170 (2.46-2.46)
RSRZ outliers	164620	1096 (2.46-2.46)

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	344	<div> <div>20%</div> <div>56%</div> <div>26%</div> <div>•</div> <div>16%</div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2315 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 hypothetical protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	288	Total	C	N	O	S	Se	0	0	0
			2290	1489	380	412	1	8			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	137	MSE	MET	modified residue	GB 41055590
A	146	MSE	MET	modified residue	GB 41055590
A	259	MSE	MET	modified residue	GB 41055590
A	264	MSE	MET	modified residue	GB 41055590
A	286	MSE	MET	modified residue	GB 41055590
A	359	MSE	MET	modified residue	GB 41055590
A	366	MSE	MET	modified residue	GB 41055590
A	368	MSE	MET	modified residue	GB 41055590

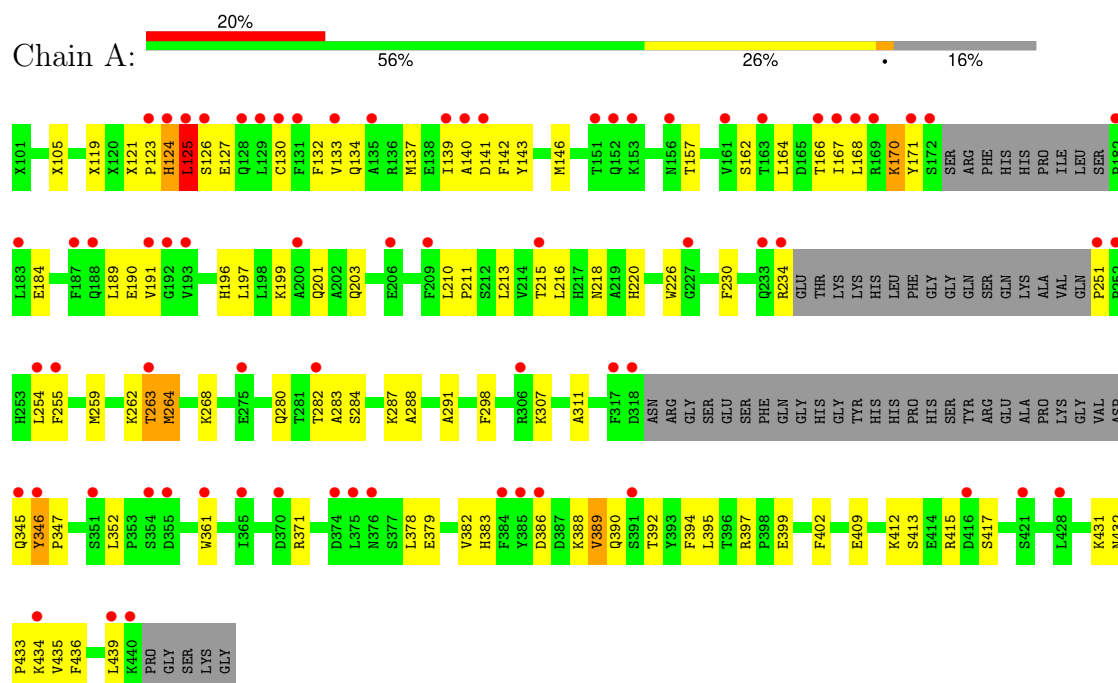
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	25	Total	O	0	0
			25	25		

### 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: hypothetical protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	82.69Å 198.96Å 67.08Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.97 – 2.45 40.97 – 2.45	Depositor EDS
% Data completeness (in resolution range)	90.2 (40.97-2.45) 90.4 (40.97-2.45)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.68 (at 2.45Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.281 , 0.322 0.280 , 0.320	Depositor DCC
$R_{free}$ test set	939 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	69.9	Xtriage
Anisotropy	0.235	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 61.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	2315	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	76.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.76% 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	0.43	0/2233	0.58	0/3002

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 [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	2290	0	2225	68	0
2	A	25	0	0	2	0
All	All	2315	0	2225	68	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 15.

All (68) 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:432:ASN:O	1:A:435:VAL:HG12	1.81	0.80
1:A:123:PRO:O	1:A:124:HIS:HB2	1.81	0.80
1:A:105:UNK:CB	1:A:254:LEU:H	1.96	0.78
1:A:220:HIS:HA	1:A:262:LYS:HD2	1.75	0.69

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:311:ALA:HB2	1:A:409:GLU:HG3	1.76	0.67
1:A:189:LEU:HD23	1:A:255:PHE:HE1	1.60	0.67
1:A:133:VAL:O	1:A:137:MSE:HG3	1.95	0.66
1:A:168:LEU:HD13	1:A:191:VAL:HG11	1.77	0.66
1:A:392:THR:HG21	1:A:412:LYS:HB2	1.80	0.64
1:A:436:PHE:O	1:A:439:LEU:HB2	1.98	0.63
1:A:345:GLN:O	1:A:346:TYR:HB2	1.99	0.63
1:A:199:LYS:O	1:A:203:GLN:HG3	2.00	0.61
1:A:282:THR:HG22	1:A:284:SER:H	1.65	0.60
1:A:413:SER:C	1:A:415:ARG:H	2.06	0.58
1:A:371:ARG:NH1	1:A:382:VAL:HG13	2.20	0.57
1:A:162:SER:O	1:A:166:THR:HG23	2.05	0.56
1:A:394:PHE:C	1:A:395:LEU:HD12	2.26	0.55
1:A:230:PHE:CZ	1:A:234:ARG:HD3	2.41	0.55
1:A:298:PHE:CZ	1:A:352:LEU:HD21	2.42	0.54
1:A:264:MSE:HE1	1:A:432:ASN:ND2	2.22	0.54
1:A:346:TYR:N	1:A:347:PRO:CD	2.71	0.53
1:A:298:PHE:HZ	1:A:352:LEU:HD21	1.74	0.53
1:A:264:MSE:HE1	1:A:432:ASN:HD22	1.74	0.52
1:A:139:ILE:O	1:A:142:PHE:HB3	2.10	0.52
1:A:170:LYS:HA	1:A:170:LYS:HE2	1.91	0.52
1:A:190:GLU:OE2	1:A:254:LEU:HD21	2.09	0.52
1:A:397:ARG:HG3	1:A:402:PHE:O	2.11	0.51
1:A:389:VAL:HG13	1:A:389:VAL:O	2.11	0.51
1:A:134:GLN:HA	1:A:137:MSE:HE3	1.92	0.51
1:A:383:HIS:HD2	1:A:394:PHE:CE1	2.30	0.50
1:A:263:THR:HB	2:A:8:HOH:O	2.12	0.49
1:A:288:ALA:O	1:A:291:ALA:HB3	2.13	0.49
1:A:199:LYS:HB3	1:A:215:THR:HG21	1.95	0.49
1:A:134:GLN:HG2	1:A:171:TYR:OH	2.13	0.48
1:A:386:ASP:OD2	1:A:389:VAL:HB	2.14	0.48
1:A:164:LEU:HD12	1:A:167:ILE:HD11	1.96	0.48
1:A:311:ALA:CB	1:A:409:GLU:HG3	2.42	0.48
1:A:143:TYR:HD2	1:A:146:MSE:HE2	1.78	0.47
1:A:130:CYS:SG	1:A:132:PHE:HB2	2.54	0.47
1:A:268:LYS:NZ	1:A:399:GLU:OE2	2.33	0.46
1:A:157:THR:HG21	1:A:201:GLN:HB3	1.97	0.46
1:A:378:LEU:HD22	1:A:397:ARG:HG2	1.98	0.46
1:A:196:HIS:CE1	1:A:218:ASN:HB3	2.51	0.45
1:A:210:LEU:HB2	1:A:211:PRO:HD3	1.98	0.45
1:A:413:SER:C	1:A:415:ARG:N	2.70	0.45

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:298:PHE:CZ	1:A:352:LEU:HD11	2.52	0.45
1:A:395:LEU:HD12	1:A:395:LEU:N	2.32	0.44
1:A:386:ASP:O	1:A:390:GLN:N	2.51	0.43
1:A:415:ARG:O	1:A:417:SER:N	2.51	0.43
1:A:197:LEU:HD23	1:A:197:LEU:HA	1.86	0.43
1:A:432:ASN:N	1:A:433:PRO:CD	2.81	0.43
1:A:130:CYS:C	1:A:132:PHE:H	2.20	0.43
1:A:226:TRP:HZ3	1:A:259:MSE:CE	2.31	0.43
1:A:283:ALA:O	1:A:287:LYS:HG2	2.19	0.43
1:A:379:GLU:HG3	2:A:19:HOH:O	2.18	0.43
1:A:216:LEU:HD23	1:A:216:LEU:HA	1.76	0.42
1:A:431:LYS:HD3	1:A:434:LYS:HE2	2.01	0.42
1:A:119:UNK:C	1:A:121:UNK:N	2.82	0.42
1:A:139:ILE:HG13	1:A:140:ALA:N	2.35	0.41
1:A:213:LEU:HD12	1:A:213:LEU:O	2.20	0.41
1:A:383:HIS:HD2	1:A:394:PHE:CZ	2.38	0.41
1:A:435:VAL:HG13	1:A:436:PHE:HD1	1.85	0.41
1:A:230:PHE:CE2	1:A:251:PRO:HG3	2.56	0.41
1:A:298:PHE:CE1	1:A:352:LEU:HD11	2.56	0.41
1:A:125:LEU:CD1	1:A:126:SER:N	2.84	0.41
1:A:282:THR:O	1:A:283:ALA:C	2.58	0.41
1:A:184:GLU:OE2	1:A:184:GLU:C	2.59	0.40
1:A:371:ARG:HG3	1:A:382:VAL:HG11	2.04	0.40

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	259/344 (75%)	237 (92%)	19 (7%)	3 (1%)	<b>11</b> <b>11</b>

All (3) Ramachandran outliers are listed below:



Mol	Chain	Res	Type
1	A	124	HIS
1	A	125	LEU
1	A	346	TYR

### 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	246/284 (87%)	235 (96%)	11 (4%)	23	34

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

Mol	Chain	Res	Type
1	A	125	LEU
1	A	127	GLU
1	A	141	ASP
1	A	170	LYS
1	A	263	THR
1	A	264	MSE
1	A	280	GLN
1	A	307	LYS
1	A	361	TRP
1	A	388	LYS
1	A	389	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	152	GLN
1	A	188	GLN
1	A	196	HIS
1	A	201	GLN
1	A	217	HIS
1	A	233	GLN
1	A	345	GLN
1	A	401	HIS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	432	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 oligosaccharides 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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	121:UNK	C	123:PRO	N	6.26

## 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	259/344 (75%)	1.39	70 (27%) 2 2	25, 70, 118, 131	1 (0%)

All (70) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	130	CYS	7.2
1	A	361	TRP	6.2
1	A	227	GLY	5.2
1	A	167	ILE	5.1
1	A	129	LEU	4.9
1	A	131	PHE	4.6
1	A	428	LEU	4.4
1	A	128	GLN	4.3
1	A	133	VAL	4.3
1	A	140	ALA	4.3
1	A	166	THR	4.2
1	A	183	LEU	3.9
1	A	345	GLN	3.8
1	A	234	ARG	3.7
1	A	123	PRO	3.7
1	A	156	ASN	3.5
1	A	346	TYR	3.5
1	A	187	PHE	3.3
1	A	191	VAL	3.3
1	A	152	GLN	3.3
1	A	125	LEU	3.2
1	A	182	PRO	3.2
1	A	255	PHE	3.2
1	A	370	ASP	3.1
1	A	318	ASP	3.1
1	A	386	ASP	3.0
1	A	385	TYR	3.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	440	LYS	3.0
1	A	171	TYR	2.9
1	A	188	GLN	2.9
1	A	124	HIS	2.9
1	A	233	GLN	2.9
1	A	376	ASN	2.8
1	A	126	SER	2.7
1	A	251	PRO	2.7
1	A	421	SER	2.7
1	A	416	ASP	2.6
1	A	306	ARG	2.6
1	A	439	LEU	2.6
1	A	151	THR	2.6
1	A	275	GLU	2.6
1	A	317	PHE	2.5
1	A	200	ALA	2.5
1	A	282	THR	2.5
1	A	193	VAL	2.5
1	A	351	SER	2.5
1	A	391	SER	2.5
1	A	168	LEU	2.5
1	A	172	SER	2.5
1	A	384	PHE	2.5
1	A	169	ARG	2.4
1	A	254	LEU	2.4
1	A	135	ALA	2.4
1	A	434	LYS	2.3
1	A	192	GLY	2.3
1	A	374	ASP	2.3
1	A	206	GLU	2.2
1	A	139	ILE	2.2
1	A	354	SER	2.2
1	A	141	ASP	2.2
1	A	375	LEU	2.1
1	A	161	VAL	2.1
1	A	365	ILE	2.1
1	A	263	THR	2.1
1	A	153	LYS	2.1
1	A	163	THR	2.0
1	A	215	THR	2.0
1	A	355	ASP	2.0
1	A	252	PRO	2.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	209	PHE	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 monosaccharides 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.