



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 19, 2024 – 11:01 PM EST

PDB ID : 4MHC
Title : Crystal Structure of a Nucleoporin
Authors : Seo, H.S.; Blus, B.J.; Blobel, G.
Deposited on : 2013-08-29
Resolution : 2.40 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 ⓘ 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
Xtriage (Phenix) : 1.13
EDS : 2.36
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) : 2.36

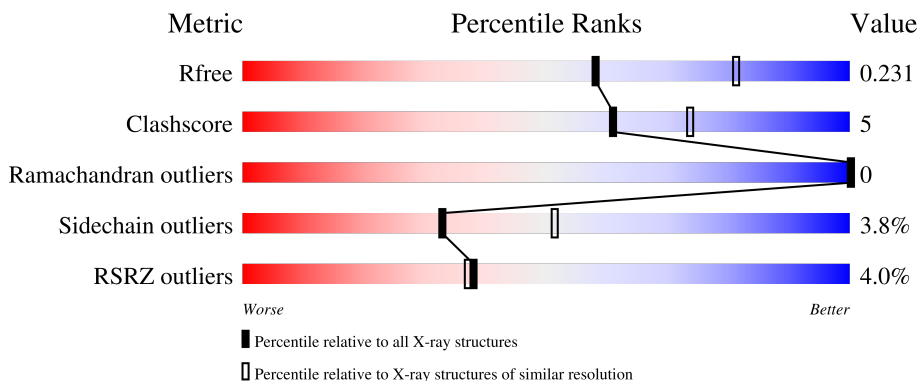
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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 $\leq 5\%$. 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	826	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5447 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 Nucleoporin NUP157.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	671	5299	3407	863	1015	14	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	68	SER	-	expression tag	UNP P40064
A	69	MET	-	expression tag	UNP P40064

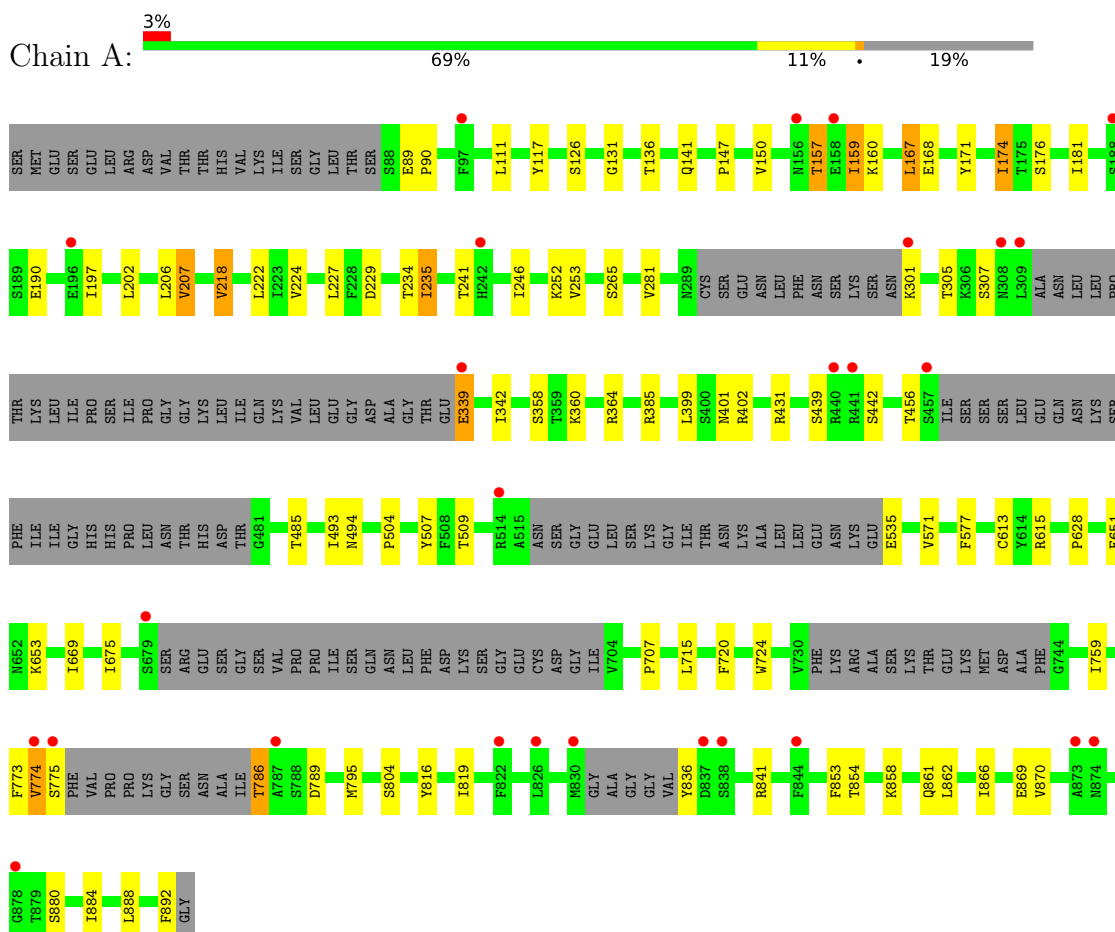
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	148	Total	O	0	0
			148	148		

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: Nucleoporin NUP157



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	62.11Å 76.94Å 82.63Å 90.00° 106.05° 90.00°	Depositor
Resolution (Å)	45.10 – 2.40 45.11 – 2.40	Depositor EDS
% Data completeness (in resolution range)	96.4 (45.10-2.40) 96.4 (45.11-2.40)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.88 (at 2.39Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, R_{free}	0.189 , 0.231 0.189 , 0.231	Depositor DCC
R_{free} test set	1432 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å ²)	44.3	Xtrriage
Anisotropy	0.201	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 49.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	5447	wwPDB-VP
Average B, all atoms (Å ²)	37.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.55% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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	0.28	0/5402	0.46	0/7330

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	5299	0	5310	48	0
2	A	148	0	0	1	0
All	All	5447	0	5310	48	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 5.

All (48) 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:207:VAL:HG22	1:A:265:SER:HB3	1.79	0.64
1:A:535:GLU:N	2:A:1033:HOH:O	2.31	0.64
1:A:431:ARG:HD2	1:A:507:TYR:CE2	2.34	0.62
1:A:339:GLU:HB3	1:A:364:ARG:HH22	1.65	0.62

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:202:LEU:HD11	1:A:227:LEU:HD23	1.85	0.58
1:A:786:THR:HG23	1:A:789:ASP:HB2	1.85	0.57
1:A:439:SER:OG	1:A:442:SER:OG	2.22	0.54
1:A:141:GLN:HG3	1:A:613:CYS:SG	2.47	0.54
1:A:675:ILE:HG23	1:A:707:PRO:HG2	1.88	0.54
1:A:197:ILE:HD13	1:A:224:VAL:HG11	1.90	0.54
1:A:167:LEU:HD12	1:A:571:VAL:HG13	1.89	0.53
1:A:339:GLU:N	1:A:358:SER:HG	2.06	0.53
1:A:111:LEU:HD22	1:A:504:PRO:HB2	1.90	0.53
1:A:888:LEU:HA	1:A:892:PHE:HB2	1.91	0.53
1:A:222:LEU:HB2	1:A:235:ILE:HD11	1.92	0.52
1:A:147:PRO:HG2	1:A:150:VAL:HB	1.92	0.50
1:A:774:VAL:HG12	1:A:775:SER:H	1.75	0.50
1:A:171:TYR:OH	1:A:190:GLU:OE1	2.31	0.49
1:A:159:ILE:HG12	1:A:160:LYS:HB2	1.95	0.48
1:A:339:GLU:HG2	1:A:360:LYS:HE2	1.95	0.48
1:A:168:GLU:OE1	1:A:168:GLU:N	2.43	0.47
1:A:150:VAL:HG13	1:A:174:ILE:HD11	1.96	0.47
1:A:804:SER:OG	1:A:853:PHE:O	2.26	0.47
1:A:628:PRO:HG2	1:A:669:ILE:HD11	1.97	0.47
1:A:720:PHE:HE2	1:A:759:ILE:HD13	1.79	0.46
1:A:385:ARG:NH1	1:A:401:ASN:OD1	2.47	0.46
1:A:493:ILE:HA	1:A:494:ASN:HA	1.61	0.46
1:A:858:LYS:O	1:A:861:GLN:HG2	2.16	0.45
1:A:720:PHE:O	1:A:724:TRP:HB2	2.18	0.44
1:A:126:SER:OG	1:A:131:GLY:HA3	2.17	0.44
1:A:456:THR:HG21	1:A:485:THR:OG1	2.18	0.44
1:A:253:VAL:HA	1:A:301:LYS:HD2	1.99	0.44
1:A:773:PHE:HB3	1:A:795:MET:HG2	1.98	0.44
1:A:836:TYR:HD1	1:A:841:ARG:HH21	1.64	0.44
1:A:229:ASP:CG	1:A:252:LYS:HE3	2.38	0.43
1:A:206:LEU:HD11	1:A:235:ILE:HD13	2.00	0.43
1:A:89:GLU:HA	1:A:90:PRO:HD3	1.87	0.43
1:A:305:THR:HG22	1:A:307:SER:H	1.83	0.43
1:A:136:THR:O	1:A:615:ARG:HD3	2.18	0.42
1:A:653:LYS:HD3	1:A:653:LYS:HA	1.75	0.42
1:A:218:VAL:HG21	1:A:234:THR:HB	2.01	0.42
1:A:234:THR:O	1:A:246:ILE:HA	2.19	0.42
1:A:281:VAL:HA	1:A:342:ILE:HG12	2.02	0.42
1:A:174:ILE:CD1	1:A:181:ILE:HB	2.50	0.42
1:A:281:VAL:HG11	1:A:339:GLU:HA	2.02	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:866:ILE:O	1:A:870:VAL:HG23	2.20	0.41
1:A:157:THR:HB	1:A:176:SER:HB2	2.02	0.41
1:A:880:SER:O	1:A:884:ILE:HG13	2.20	0.41

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	653/826 (79%)	630 (96%)	23 (4%)	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	604/736 (82%)	581 (96%)	23 (4%)	33 51

All (23) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	117	TYR
1	A	157	THR
1	A	159	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	167	LEU
1	A	174	ILE
1	A	207	VAL
1	A	218	VAL
1	A	235	ILE
1	A	241	THR
1	A	339	GLU
1	A	399	LEU
1	A	402	ARG
1	A	509	THR
1	A	577	PHE
1	A	651	PHE
1	A	715	LEU
1	A	774	VAL
1	A	786	THR
1	A	816	TYR
1	A	819	ILE
1	A	854	THR
1	A	862	LEU
1	A	869	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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](#)

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, 95th 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 ²)	Q<0.9
1	A	671/826 (81%)	0.11	27 (4%) 38 37	14, 34, 71, 93	0

All (27) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	309	LEU	4.6
1	A	440	ARG	4.2
1	A	844	PHE	3.9
1	A	874	ASN	3.8
1	A	158	GLU	3.5
1	A	873	ALA	3.2
1	A	774	VAL	3.1
1	A	787	ALA	3.0
1	A	301	LYS	2.9
1	A	826	LEU	2.9
1	A	156	ASN	2.9
1	A	514	ARG	2.9
1	A	830	MET	2.8
1	A	308	ASN	2.7
1	A	822	PHE	2.6
1	A	188	SER	2.5
1	A	878	GLY	2.5
1	A	837	ASP	2.5
1	A	775	SER	2.4
1	A	339	GLU	2.3
1	A	441	ARG	2.3
1	A	457	SER	2.2
1	A	838	SER	2.2
1	A	196	GLU	2.1
1	A	242	HIS	2.1
1	A	97	PHE	2.0
1	A	679	SER	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.