



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 22, 2023 – 06:05 PM EDT

PDB ID : 8G0Q  
Title : Crystal structure of the yeast Ndc80:Nuf2 head region with a bound Dam1 segment  
Authors : Zahm, J.A.; Jenni, S.; Harrison, S.C.  
Deposited on : 2023-02-01  
Resolution : 3.22 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

<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 : 2.32.1  
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) : 2.32.1

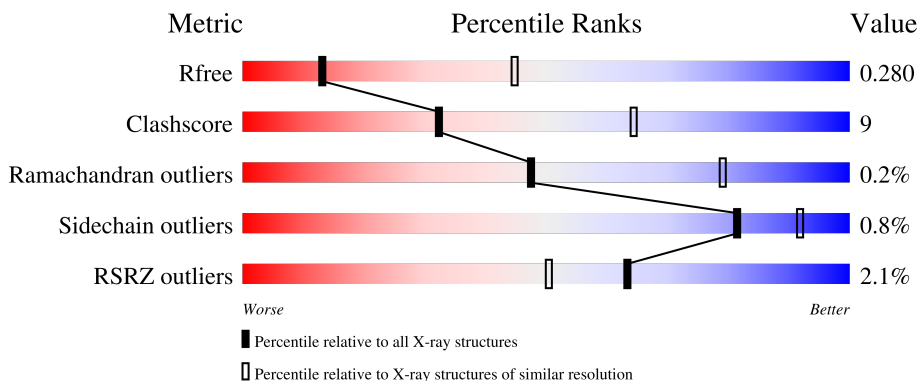
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.22 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	1335 (3.24-3.20)
Clashscore	141614	1460 (3.24-3.20)
Ramachandran outliers	138981	1437 (3.24-3.20)
Sidechain outliers	138945	1436 (3.24-3.20)
RSRZ outliers	127900	1291 (3.24-3.20)

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	287	
1	C	287	
2	B	233	
2	D	233	

## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 15883 atoms, of which 7943 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 Kinetochores protein NDC80.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	269	4560	1457	2292	381	424	6	0	0	0
1	C	268	4536	1451	2279	377	423	6	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	103	SER	-	expression tag	UNP P40460
A	104	GLY	-	expression tag	UNP P40460
A	105	SER	-	expression tag	UNP P40460
C	103	SER	-	expression tag	UNP P40460
C	104	GLY	-	expression tag	UNP P40460
C	105	SER	-	expression tag	UNP P40460

- Molecule 2 is a protein called DASH complex subunit DAM1, Kinetochores protein NUF2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
2	B	230	3700	1166	1843	316	363	12	0	0	0
2	D	190	3087	985	1529	251	310	12	0	0	0

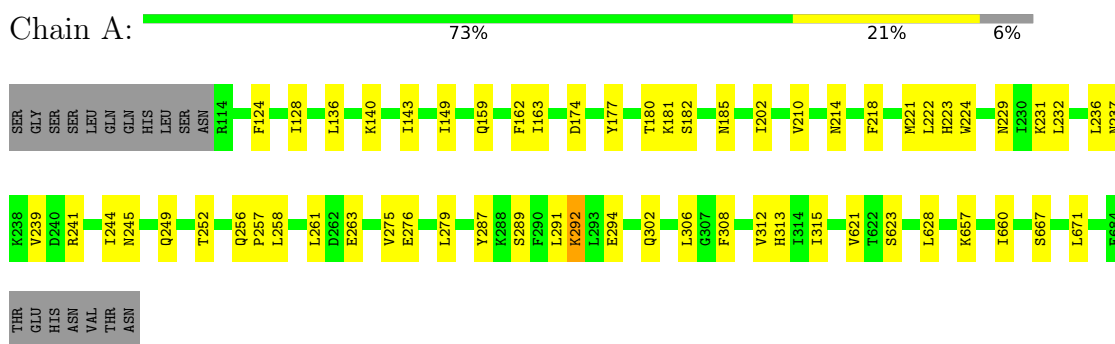
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	251	MET	-	initiating methionine	UNP P53267
D	966	MET	-	initiating methionine	UNP P53267

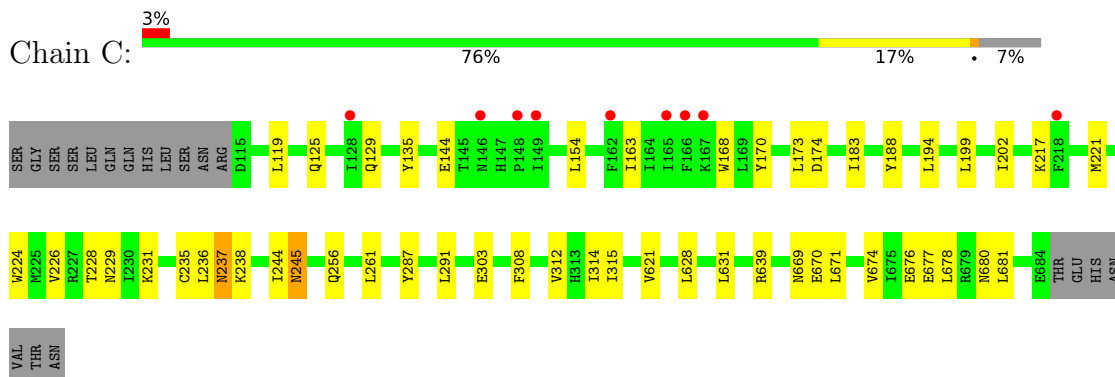
### 3 Residue-property plots [i](#)

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.

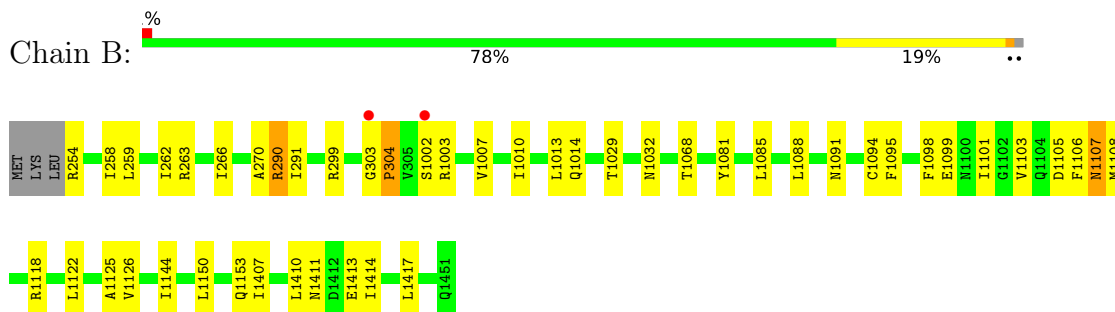
- Molecule 1: Kinetochores protein NDC80



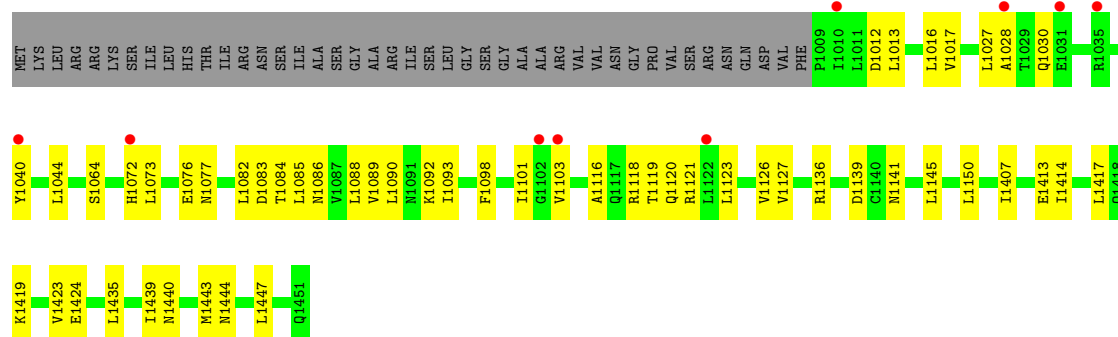
- Molecule 1: Kinetochores protein NDC80



- Molecule 2: DASH complex subunit DAM1, Kinetochores protein NUF2



- Molecule 2: DASH complex subunit DAM1, Kinetochores protein NUF2



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 32 1 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	130.00Å 130.00Å 216.31Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	99.86 – 3.22 99.86 – 3.22	Depositor EDS
% Data completeness (in resolution range)	72.8 (99.86-3.22) 72.8 (99.86-3.22)	Depositor EDS
$R_{merge}$	0.17	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.22 (at 3.19Å)	Xtrriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, $R_{free}$	0.263 , 0.285 0.266 , 0.280	Depositor DCC
$R_{free}$ test set	1236 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	80.0	Xtrriage
Anisotropy	0.000	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 25.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	0.042 for -h,-k,l	Xtrriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	15883	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	100.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.24% 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	0.26	0/2312	0.47	0/3112
1	C	0.26	0/2301	0.45	0/3098
2	B	0.26	0/1882	0.49	0/2534
2	D	0.26	0/1580	0.45	0/2127
All	All	0.26	0/8075	0.47	0/10871

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
2	B	0	1
All	All	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	292	LYS	Peptide
2	B	304	PRO	Peptide

### 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	2268	2292	2292	47	0
1	C	2257	2279	2279	34	0
2	B	1857	1843	1843	41	0
2	D	1558	1529	1529	41	0
All	All	7940	7943	7943	139	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 9.

The worst 5 of 139 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)
2:D:1013:LEU:HD21	2:D:1030:GLN:NE2	2.00	0.77
2:B:1029:THR:OG1	2:B:1032:ASN:OD1	2.03	0.75
2:B:1105:ASP:O	2:B:1118:ARG:NH2	2.22	0.73
1:C:669:ASN:OD1	1:C:670:GLU:N	2.22	0.72
1:A:202:ILE:HD11	1:A:218:PHE:HE1	1.57	0.70

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	267/287 (93%)	260 (97%)	7 (3%)	0	100	100
1	C	266/287 (93%)	263 (99%)	3 (1%)	0	100	100
2	B	228/233 (98%)	209 (92%)	17 (8%)	2 (1%)	17	55
2	D	188/233 (81%)	183 (97%)	5 (3%)	0	100	100
All	All	949/1040 (91%)	915 (96%)	32 (3%)	2 (0%)	47	79

All (2) Ramachandran outliers are listed below:



Mol	Chain	Res	Type
2	B	290	ARG
2	B	1107	ASN

### 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	259/276 (94%)	256 (99%)	3 (1%)	71	87
1	C	258/276 (94%)	255 (99%)	3 (1%)	71	87
2	B	211/214 (99%)	210 (100%)	1 (0%)	88	94
2	D	179/214 (84%)	179 (100%)	0	100	100
All	All	907/980 (93%)	900 (99%)	7 (1%)	81	92

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	254	ARG
1	C	229	ASN
1	C	245	ASN
1	C	237	ASN
1	A	237	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	256	GLN
1	C	245	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 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 [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	269/287 (93%)	-0.00	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	40, 72, 121, 156	0
1	C	268/287 (93%)	0.15	9 (3%) <span style="border: 1px solid red; padding: 2px;">45</span> <span style="border: 1px solid red; padding: 2px;">31</span>	51, 103, 149, 166	0
2	B	230/233 (98%)	-0.09	2 (0%) <span style="border: 1px solid blue; padding: 2px;">84</span> <span style="border: 1px solid blue; padding: 2px;">76</span>	43, 85, 143, 162	0
2	D	190/233 (81%)	0.13	9 (4%) <span style="border: 1px solid red; padding: 2px;">31</span> <span style="border: 1px solid red; padding: 2px;">20</span>	43, 108, 160, 180	0
All	All	957/1040 (92%)	0.05	20 (2%) <span style="border: 1px solid blue; padding: 2px;">63</span> <span style="border: 1px solid blue; padding: 2px;">50</span>	40, 88, 150, 180	0

The worst 5 of 20 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	149	ILE	6.0
1	C	148	PRO	4.4
2	D	1035	ARG	4.1
2	B	1002	SER	3.9
1	C	162	PHE	2.8

### 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.