



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 16, 2019 – 11:26 AM EDT

PDB ID : 6R4M  
Title : Crystal structure of S. cerevisia Niemann-Pick type C protein NPC2  
Authors : Winkler, M.B.L.; Kidmose, R.T.; Pedersen, B.P.  
Deposited on : 2019-03-22  
Resolution : 2.80 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.0 (224370), CSD as540be (2019)  
Xtriage (Phenix) : 1.13  
EDS : 2.4  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.4

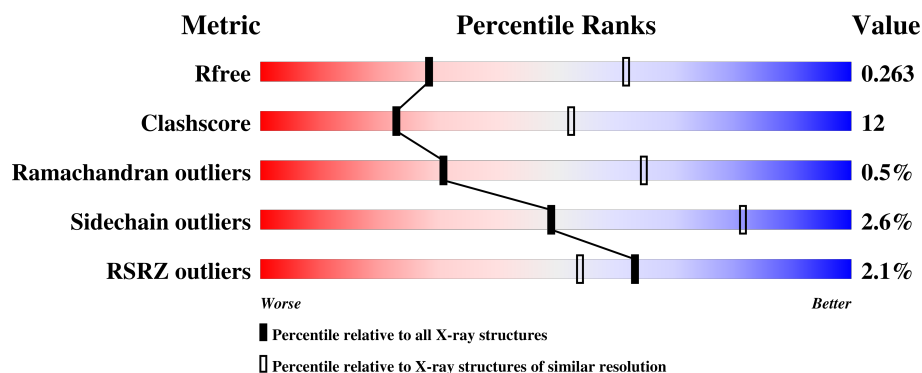
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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))
$R_{free}$	111664	2792 (2.80-2.80)
Clashscore	122126	3209 (2.80-2.80)
Ramachandran outliers	120053	3158 (2.80-2.80)
Sidechain outliers	120020	3160 (2.80-2.80)
RSRZ outliers	108989	2726 (2.80-2.80)

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	201	<div> <div>%</div> <div> <div></div> <div>45%</div> <div>23%</div> <div>•</div> <div>29%</div> </div> </div>
1	B	201	<div> <div>2%</div> <div> <div></div> <div>55%</div> <div>15%</div> <div>•</div> <div>29%</div> </div> </div>
1	C	201	<div> <div>%</div> <div> <div></div> <div>58%</div> <div>14%</div> <div>•</div> <div>25%</div> </div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3431 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 Phosphatidylglycerol/phosphatidylinositol transfer protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	142	Total	C	N	O	S	0	0	0
			1106	705	172	225	4			
1	B	143	Total	C	N	O	S	0	0	0
			1114	711	173	226	4			
1	C	150	Total	C	N	O	S	0	0	0
			1169	747	184	234	4			

There are 84 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	174	LEU	-	expression tag	UNP Q12408
A	175	VAL	-	expression tag	UNP Q12408
A	176	PRO	-	expression tag	UNP Q12408
A	177	ARG	-	expression tag	UNP Q12408
A	178	GLY	-	expression tag	UNP Q12408
A	179	SER	-	expression tag	UNP Q12408
A	180	GLY	-	expression tag	UNP Q12408
A	181	GLY	-	expression tag	UNP Q12408
A	182	GLY	-	expression tag	UNP Q12408
A	183	GLY	-	expression tag	UNP Q12408
A	184	SER	-	expression tag	UNP Q12408
A	185	GLY	-	expression tag	UNP Q12408
A	186	GLY	-	expression tag	UNP Q12408
A	187	GLY	-	expression tag	UNP Q12408
A	188	GLY	-	expression tag	UNP Q12408
A	189	SER	-	expression tag	UNP Q12408
A	190	GLY	-	expression tag	UNP Q12408
A	191	GLY	-	expression tag	UNP Q12408
A	192	HIS	-	expression tag	UNP Q12408
A	193	HIS	-	expression tag	UNP Q12408
A	194	HIS	-	expression tag	UNP Q12408
A	195	HIS	-	expression tag	UNP Q12408
A	196	HIS	-	expression tag	UNP Q12408

*Continued on next page...*

*Continued from previous page...*

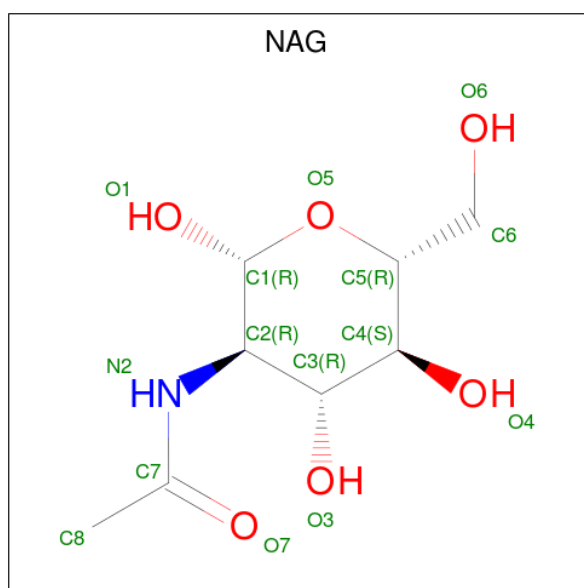
Chain	Residue	Modelled	Actual	Comment	Reference
A	197	HIS	-	expression tag	UNP Q12408
A	198	HIS	-	expression tag	UNP Q12408
A	199	HIS	-	expression tag	UNP Q12408
A	200	HIS	-	expression tag	UNP Q12408
A	201	HIS	-	expression tag	UNP Q12408
B	174	LEU	-	expression tag	UNP Q12408
B	175	VAL	-	expression tag	UNP Q12408
B	176	PRO	-	expression tag	UNP Q12408
B	177	ARG	-	expression tag	UNP Q12408
B	178	GLY	-	expression tag	UNP Q12408
B	179	SER	-	expression tag	UNP Q12408
B	180	GLY	-	expression tag	UNP Q12408
B	181	GLY	-	expression tag	UNP Q12408
B	182	GLY	-	expression tag	UNP Q12408
B	183	GLY	-	expression tag	UNP Q12408
B	184	SER	-	expression tag	UNP Q12408
B	185	GLY	-	expression tag	UNP Q12408
B	186	GLY	-	expression tag	UNP Q12408
B	187	GLY	-	expression tag	UNP Q12408
B	188	GLY	-	expression tag	UNP Q12408
B	189	SER	-	expression tag	UNP Q12408
B	190	GLY	-	expression tag	UNP Q12408
B	191	GLY	-	expression tag	UNP Q12408
B	192	HIS	-	expression tag	UNP Q12408
B	193	HIS	-	expression tag	UNP Q12408
B	194	HIS	-	expression tag	UNP Q12408
B	195	HIS	-	expression tag	UNP Q12408
B	196	HIS	-	expression tag	UNP Q12408
B	197	HIS	-	expression tag	UNP Q12408
B	198	HIS	-	expression tag	UNP Q12408
B	199	HIS	-	expression tag	UNP Q12408
B	200	HIS	-	expression tag	UNP Q12408
B	201	HIS	-	expression tag	UNP Q12408
C	174	LEU	-	expression tag	UNP Q12408
C	175	VAL	-	expression tag	UNP Q12408
C	176	PRO	-	expression tag	UNP Q12408
C	177	ARG	-	expression tag	UNP Q12408
C	178	GLY	-	expression tag	UNP Q12408
C	179	SER	-	expression tag	UNP Q12408
C	180	GLY	-	expression tag	UNP Q12408
C	181	GLY	-	expression tag	UNP Q12408
C	182	GLY	-	expression tag	UNP Q12408

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	183	GLY	-	expression tag	UNP Q12408
C	184	SER	-	expression tag	UNP Q12408
C	185	GLY	-	expression tag	UNP Q12408
C	186	GLY	-	expression tag	UNP Q12408
C	187	GLY	-	expression tag	UNP Q12408
C	188	GLY	-	expression tag	UNP Q12408
C	189	SER	-	expression tag	UNP Q12408
C	190	GLY	-	expression tag	UNP Q12408
C	191	GLY	-	expression tag	UNP Q12408
C	192	HIS	-	expression tag	UNP Q12408
C	193	HIS	-	expression tag	UNP Q12408
C	194	HIS	-	expression tag	UNP Q12408
C	195	HIS	-	expression tag	UNP Q12408
C	196	HIS	-	expression tag	UNP Q12408
C	197	HIS	-	expression tag	UNP Q12408
C	198	HIS	-	expression tag	UNP Q12408
C	199	HIS	-	expression tag	UNP Q12408
C	200	HIS	-	expression tag	UNP Q12408
C	201	HIS	-	expression tag	UNP Q12408

- Molecule 2 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	C	1	Total	C	N	O	0	0
			14	8	1	5		



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	205.32Å 205.32Å 39.08Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	44.45 – 2.80 44.45 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.5 (44.45-2.80) 99.6 (44.45-2.80)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.09 (at 2.81Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.236 , 0.262 0.236 , 0.263	Depositor DCC
$R_{free}$ test set	1197 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	90.4	Xtriage
Anisotropy	0.333	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 64.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.034 for h,-h-k,-l	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	3431	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	107.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.99% 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:  
NAG

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	0.41	0/1129	0.70	1/1543 (0.1%)
1	B	0.39	0/1137	0.71	2/1555 (0.1%)
1	C	0.36	0/1193	0.72	1/1630 (0.1%)
All	All	0.39	0/3459	0.71	4/4728 (0.1%)

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	C	0	1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	123	CYS	CA-CB-SG	-5.55	104.01	114.00
1	C	123	CYS	CA-CB-SG	-5.52	104.06	114.00
1	A	157	LYS	CD-CE-NZ	5.38	124.07	111.70
1	B	46	GLN	CA-CB-CG	5.33	125.13	113.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	25	GLY	Peptide

## 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	1106	0	1091	42	1
1	B	1114	0	1101	24	0
1	C	1169	0	1159	26	1
2	A	14	0	13	0	0
2	B	14	0	13	0	0
2	C	14	0	13	0	0
All	All	3431	0	3390	84	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 12.

The worst 5 of 84 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:C:119:GLU:OE2	1:C:133:LYS:N	2.07	0.88
1:B:63:ASP:HB2	1:B:74:THR:HB	1.58	0.84
1:C:25:GLY:O	1:C:27:PHE:HD1	1.62	0.81
1:A:91:ILE:HD11	1:A:109:LEU:HD13	1.63	0.79
1:A:72:ASN:OD1	1:A:137:GLU:HG2	1.84	0.77

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)
1:A:157:LYS:NZ	1:C:41:GLU:OE2[1_556]	2.15	0.05

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	140/201 (70%)	134 (96%)	6 (4%)	0	100	100
1	B	141/201 (70%)	137 (97%)	4 (3%)	0	100	100
1	C	148/201 (74%)	141 (95%)	5 (3%)	2 (1%)	12	38
All	All	429/603 (71%)	412 (96%)	15 (4%)	2 (0%)	31	65

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	26	ILE
1	C	27	PHE

### 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	128/168 (76%)	124 (97%)	4 (3%)	43	77
1	B	129/168 (77%)	128 (99%)	1 (1%)	83	95
1	C	134/168 (80%)	129 (96%)	5 (4%)	37	71
All	All	391/504 (78%)	381 (97%)	10 (3%)	49	81

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

Mol	Chain	Res	Type
1	B	147	TYR
1	C	27	PHE
1	C	101	ARG
1	A	154	TYR
1	C	52	LYS

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

Mol	Chain	Res	Type
1	B	105	GLN
1	C	105	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

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	NAG	A	301	1	14,14,15	0.98	2 (14%)	17,19,21	0.67	1 (5%)
2	NAG	B	301	1	14,14,15	0.60	0	17,19,21	0.40	0
2	NAG	C	301	1	14,14,15	0.58	0	17,19,21	0.70	1 (5%)

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	NAG	A	301	1	-	2/6/23/26	0/1/1/1
2	NAG	B	301	1	-	0/6/23/26	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAG	C	301	1	-	2/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	301	NAG	O5-C1	2.82	1.48	1.43
2	A	301	NAG	C1-C2	2.21	1.55	1.52

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	301	NAG	C1-O5-C5	2.45	115.53	112.20
2	A	301	NAG	C1-O5-C5	2.30	115.33	112.20

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	301	NAG	C4-C5-C6-O6
2	A	301	NAG	O5-C5-C6-O6
2	C	301	NAG	C4-C5-C6-O6
2	C	301	NAG	O5-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

## 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	142/201 (70%)	-0.10	2 (1%) 75 69	89, 117, 159, 190	0
1	B	143/201 (71%)	-0.12	5 (3%) 44 33	67, 104, 153, 187	0
1	C	150/201 (74%)	-0.27	2 (1%) 77 71	61, 84, 135, 180	0
All	All	435/603 (72%)	-0.16	9 (2%) 63 54	61, 101, 156, 190	0

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	30	LEU	5.6
1	B	172	PRO	4.1
1	C	24	ILE	3.6
1	C	26	ILE	3.4
1	B	102	LEU	3.3

### 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. 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	B-factors( $\text{\AA}^2$ )	Q<0.9
2	NAG	B	301	14/15	0.73	0.22	161,177,192,195	0
2	NAG	A	301	14/15	0.81	0.24	166,175,182,185	0
2	NAG	C	301	14/15	0.88	0.22	147,160,172,182	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.