



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 14, 2023 – 01:19 PM EDT

PDB ID : 8EUS  
Title : Crystal structure of NPC1 luminal domain C  
Authors : Odongo, L.; Pornillos, O.  
Deposited on : 2022-10-19  
Resolution : 2.30 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  
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

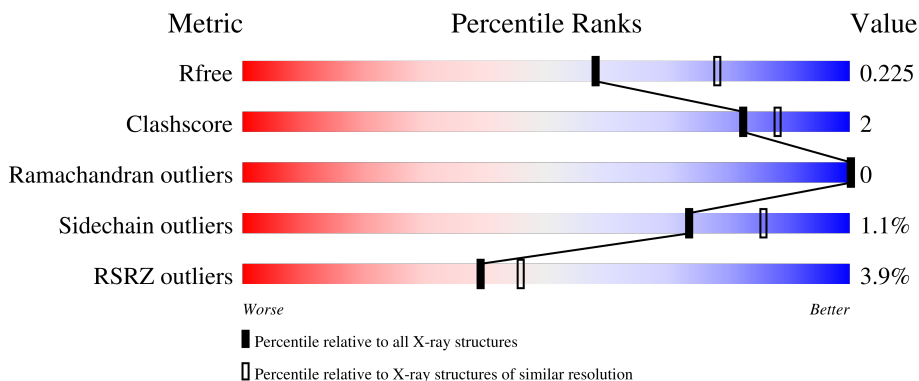
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.30 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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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	262	 3% 71% 8% 21%
1	C	262	 3% 75% . 21%

## 2 Entry composition i

There are 2 unique types of molecules in this entry. The entry contains 6688 atoms, of which 3185 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 NPC intracellular cholesterol transporter 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	C	207	3269	1083	1596	270	316	4	0	1	0
1	A	207	3259	1082	1589	270	314	4	0	1	0

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	370	MET	-	initiating methionine	UNP O15118
C	622	LEU	-	expression tag	UNP O15118
C	623	GLU	-	expression tag	UNP O15118
C	624	HIS	-	expression tag	UNP O15118
C	625	HIS	-	expression tag	UNP O15118
C	626	HIS	-	expression tag	UNP O15118
C	627	HIS	-	expression tag	UNP O15118
C	628	HIS	-	expression tag	UNP O15118
C	629	HIS	-	expression tag	UNP O15118
C	630	HIS	-	expression tag	UNP O15118
C	631	HIS	-	expression tag	UNP O15118
A	370	MET	-	initiating methionine	UNP O15118
A	622	LEU	-	expression tag	UNP O15118
A	623	GLU	-	expression tag	UNP O15118
A	624	HIS	-	expression tag	UNP O15118
A	625	HIS	-	expression tag	UNP O15118
A	626	HIS	-	expression tag	UNP O15118
A	627	HIS	-	expression tag	UNP O15118
A	628	HIS	-	expression tag	UNP O15118
A	629	HIS	-	expression tag	UNP O15118
A	630	HIS	-	expression tag	UNP O15118
A	631	HIS	-	expression tag	UNP O15118

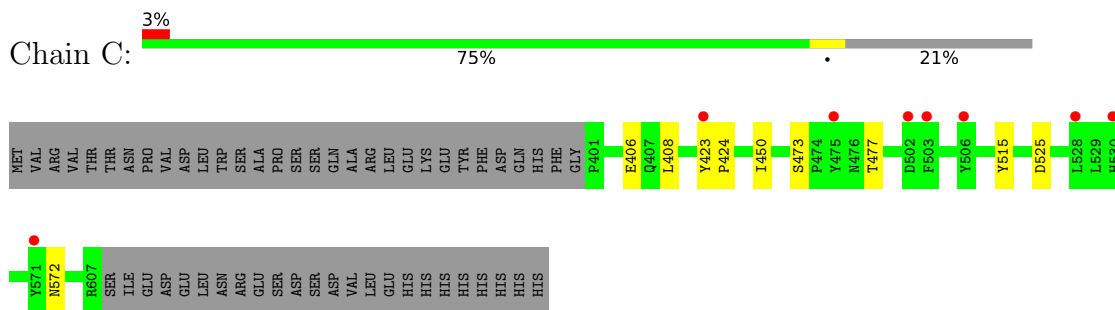
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	C	98	Total O 98 98	0	0
2	A	62	Total O 62 62	0	0

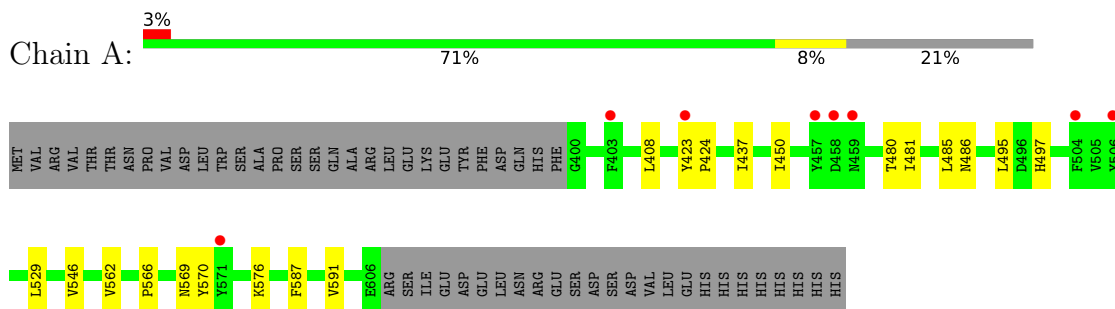
### 3 Residue-property plots [i](#)

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: NPC intracellular cholesterol transporter 1



- Molecule 1: NPC intracellular cholesterol transporter 1



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	53.15Å 65.21Å 68.53Å 90.00° 91.69° 90.00°	Depositor
Resolution (Å)	41.19 – 2.30 41.18 – 2.00	Depositor EDS
% Data completeness (in resolution range)	97.4 (41.19-2.30) 76.4 (41.18-2.00)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.12 (at 2.00Å)	Xtrriage
Refinement program	PHENIX 1.20.1_4487, PHENIX 1.20.1_4487	Depositor
R, $R_{free}$	0.191 , 0.229 0.189 , 0.225	Depositor DCC
$R_{free}$ test set	1263 reflections (5.23%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	26.7	Xtrriage
Anisotropy	0.152	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 42.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.046 for h,-k,-l	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	6688	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	35.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.84% 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	0.27	0/1725	0.49	0/2365
1	C	0.27	0/1727	0.49	0/2367
All	All	0.27	0/3452	0.49	0/4732

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	1670	1589	1588	12	0
1	C	1673	1596	1593	4	0
2	A	62	0	0	0	0
2	C	98	0	0	0	0
All	All	3503	3185	3181	16	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 2.

All (16) 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:546:VAL:HA	1:A:562:VAL:HG13	1.83	0.61

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:481:ILE:O	1:A:486:ASN:ND2	2.38	0.56
1:C:408:LEU:HD11	1:C:450:ILE:CD1	2.40	0.52
1:A:497[A]:HIS:CG	1:A:529:LEU:HD22	2.48	0.48
1:C:473:SER:HB2	1:C:477:THR:HG22	1.97	0.47
1:A:423:TYR:CG	1:A:424:PRO:HA	2.50	0.46
1:A:570:TYR:CE1	1:A:576:LYS:HG2	2.51	0.46
1:A:480:THR:O	1:A:566:PRO:HD2	2.16	0.46
1:A:437:ILE:HD11	1:A:495:LEU:HG	1.98	0.45
1:C:423:TYR:CG	1:C:424:PRO:HA	2.53	0.44
1:A:408:LEU:HD11	1:A:450:ILE:HD11	2.00	0.43
1:A:408:LEU:HD21	1:A:591:VAL:HG11	2.01	0.43
1:A:587:PHE:O	1:A:591:VAL:HG22	2.18	0.43
1:C:515:TYR:CE2	1:C:525:ASP:HA	2.53	0.42
1:A:408:LEU:HD11	1:A:450:ILE:CD1	2.49	0.42
1:A:497[A]:HIS:ND1	1:A:529:LEU:HD22	2.34	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	206/262 (79%)	200 (97%)	6 (3%)	0	100	100
1	C	206/262 (79%)	199 (97%)	7 (3%)	0	100	100
All	All	412/524 (79%)	399 (97%)	13 (3%)	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	185/238 (78%)	183 (99%)	2 (1%)	73	86
1	C	186/238 (78%)	184 (99%)	2 (1%)	73	86
All	All	371/476 (78%)	367 (99%)	4 (1%)	73	86

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	406	GLU
1	C	572	ASN
1	A	485	LEU
1	A	569	ASN

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

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	207/262 (79%)	0.21	8 (3%) 39 46	14, 33, 63, 82	0
1	C	207/262 (79%)	-0.03	8 (3%) 39 46	13, 25, 52, 72	0
All	All	414/524 (79%)	0.09	16 (3%) 39 46	13, 29, 61, 82	0

All (16) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	475	TYR	4.5
1	A	504	PHE	3.5
1	C	503	PHE	3.5
1	A	403	PHE	3.0
1	C	423	TYR	3.0
1	C	571	TYR	2.9
1	A	571	TYR	2.8
1	C	528	LEU	2.8
1	C	502	ASP	2.7
1	A	459	ASN	2.5
1	A	458	ASP	2.4
1	C	530	HIS	2.1
1	A	423	TYR	2.1
1	C	506	TYR	2.1
1	A	506	TYR	2.1
1	A	457	TYR	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.