



# wwPDB X-ray Structure Validation Summary Report

Mar 7, 2024 – 04:54 PM EST

PDB ID : 8T48  
Title : The N4BP1 CUE-like domain in complex with linear di-Ubiquitin  
Authors : Schubert, A.F.; Harris, S.F.  
Deposited on : 2023-06-08  
Resolution : 2.00 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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
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

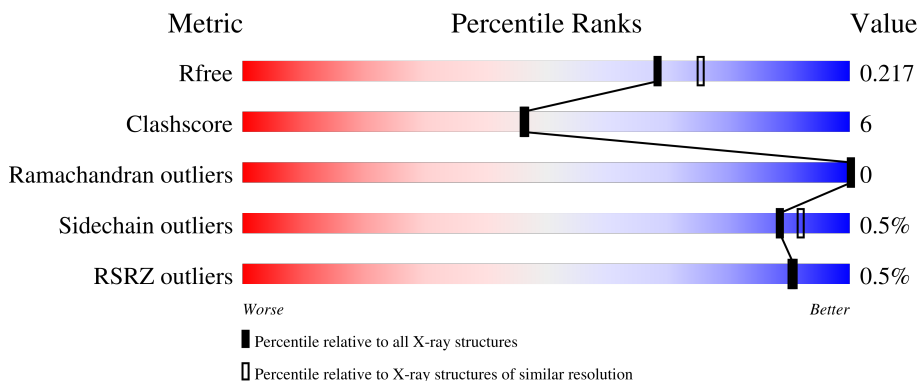
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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	172	78% 12% 10%
1	B	172	80% 10% 10%
2	C	68	72% 7% 21%
2	D	68	72% 7% 21%

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 4022 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 Di-Ubiquitin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	155	1316	818	233	263	2	0	10	0
1	B	155	1346	842	244	258	2	0	15	0

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	initiating methionine	UNP P0CG48
A	-18	GLY	-	expression tag	UNP P0CG48
A	-17	SER	-	expression tag	UNP P0CG48
A	-16	SER	-	expression tag	UNP P0CG48
A	-15	HIS	-	expression tag	UNP P0CG48
A	-14	HIS	-	expression tag	UNP P0CG48
A	-13	HIS	-	expression tag	UNP P0CG48
A	-12	HIS	-	expression tag	UNP P0CG48
A	-11	HIS	-	expression tag	UNP P0CG48
A	-10	HIS	-	expression tag	UNP P0CG48
A	-9	SER	-	expression tag	UNP P0CG48
A	-8	SER	-	expression tag	UNP P0CG48
A	-7	GLY	-	expression tag	UNP P0CG48
A	-6	LEU	-	expression tag	UNP P0CG48
A	-5	VAL	-	expression tag	UNP P0CG48
A	-4	PRO	-	expression tag	UNP P0CG48
A	-3	ARG	-	expression tag	UNP P0CG48
A	-2	GLY	-	expression tag	UNP P0CG48
A	-1	SER	-	expression tag	UNP P0CG48
A	0	HIS	-	expression tag	UNP P0CG48
B	-19	MET	-	initiating methionine	UNP P0CG48
B	-18	GLY	-	expression tag	UNP P0CG48
B	-17	SER	-	expression tag	UNP P0CG48
B	-16	SER	-	expression tag	UNP P0CG48
B	-15	HIS	-	expression tag	UNP P0CG48

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	-14	HIS	-	expression tag	UNP P0CG48
B	-13	HIS	-	expression tag	UNP P0CG48
B	-12	HIS	-	expression tag	UNP P0CG48
B	-11	HIS	-	expression tag	UNP P0CG48
B	-10	HIS	-	expression tag	UNP P0CG48
B	-9	SER	-	expression tag	UNP P0CG48
B	-8	SER	-	expression tag	UNP P0CG48
B	-7	GLY	-	expression tag	UNP P0CG48
B	-6	LEU	-	expression tag	UNP P0CG48
B	-5	VAL	-	expression tag	UNP P0CG48
B	-4	PRO	-	expression tag	UNP P0CG48
B	-3	ARG	-	expression tag	UNP P0CG48
B	-2	GLY	-	expression tag	UNP P0CG48
B	-1	SER	-	expression tag	UNP P0CG48
B	0	HIS	-	expression tag	UNP P0CG48

- Molecule 2 is a protein called NEDD4-binding protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	54	454	281	77	95	1	0	5	0
2	D	54	469	291	81	96	1	0	7	0

There are 46 discrepancies between the modelled and reference sequences:

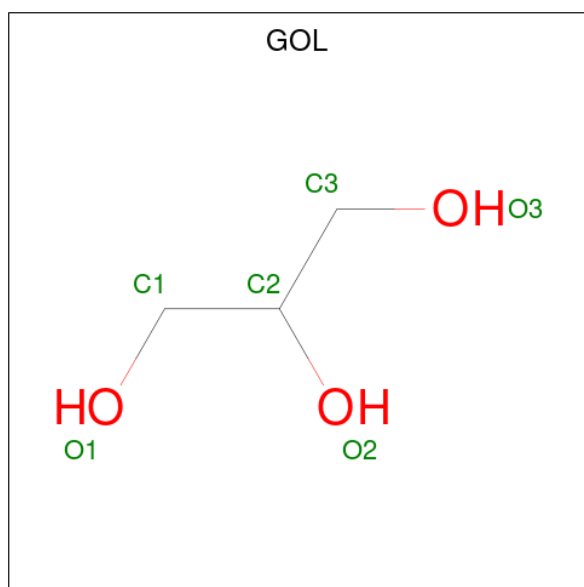
Chain	Residue	Modelled	Actual	Comment	Reference
C	829	MET	-	initiating methionine	UNP O75113
C	830	HIS	-	expression tag	UNP O75113
C	831	HIS	-	expression tag	UNP O75113
C	832	HIS	-	expression tag	UNP O75113
C	833	HIS	-	expression tag	UNP O75113
C	834	HIS	-	expression tag	UNP O75113
C	835	HIS	-	expression tag	UNP O75113
C	836	GLY	-	expression tag	UNP O75113
C	837	GLU	-	expression tag	UNP O75113
C	838	ASN	-	expression tag	UNP O75113
C	839	LEU	-	expression tag	UNP O75113
C	840	TYR	-	expression tag	UNP O75113
C	841	PHE	-	expression tag	UNP O75113
C	842	GLN	-	expression tag	UNP O75113
C	843	GLY	-	expression tag	UNP O75113

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	844	SER	-	expression tag	UNP O75113
C	852	SER	ASN	variant	UNP O75113
C	868	LYS	ARG	conflict	UNP O75113
C	876	ALA	VAL	conflict	UNP O75113
C	890	LEU	MET	conflict	UNP O75113
C	894	GLY	-	expression tag	UNP O75113
C	895	ASN	-	expression tag	UNP O75113
C	896	SER	-	expression tag	UNP O75113
D	829	MET	-	initiating methionine	UNP O75113
D	830	HIS	-	expression tag	UNP O75113
D	831	HIS	-	expression tag	UNP O75113
D	832	HIS	-	expression tag	UNP O75113
D	833	HIS	-	expression tag	UNP O75113
D	834	HIS	-	expression tag	UNP O75113
D	835	HIS	-	expression tag	UNP O75113
D	836	GLY	-	expression tag	UNP O75113
D	837	GLU	-	expression tag	UNP O75113
D	838	ASN	-	expression tag	UNP O75113
D	839	LEU	-	expression tag	UNP O75113
D	840	TYR	-	expression tag	UNP O75113
D	841	PHE	-	expression tag	UNP O75113
D	842	GLN	-	expression tag	UNP O75113
D	843	GLY	-	expression tag	UNP O75113
D	844	SER	-	expression tag	UNP O75113
D	852	SER	ASN	variant	UNP O75113
D	868	LYS	ARG	conflict	UNP O75113
D	876	ALA	VAL	conflict	UNP O75113
D	890	LEU	MET	conflict	UNP O75113
D	894	GLY	-	expression tag	UNP O75113
D	895	ASN	-	expression tag	UNP O75113
D	896	SER	-	expression tag	UNP O75113

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 6 3 3	0	0

- Molecule 4 is LITHIUM ION (three-letter code: LI) (formula: Li).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total Li 1 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	146	Total O 146 146	0	0
5	B	154	Total O 154 154	0	0
5	C	61	Total O 61 61	0	0
5	D	69	Total O 69 69	0	0



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	35.79Å 118.72Å 50.70Å 90.00° 91.72° 90.00°	Depositor
Resolution (Å)	46.60 – 2.00 59.36 – 2.00	Depositor EDS
% Data completeness (in resolution range)	97.2 (46.60-2.00) 97.2 (59.36-2.00)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.24 (at 2.00Å)	Xtrriage
Refinement program	PHENIX 1.14_3260	Depositor
R, $R_{free}$	0.160 , 0.218 0.160 , 0.217	Depositor DCC
$R_{free}$ test set	1277 reflections (4.60%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	20.5	Xtrriage
Anisotropy	0.221	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 56.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	0.196 for h,-k,-l	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	4022	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.24% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, LI

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	0.38	0/1330	0.54	0/1788
1	B	0.35	0/1363	0.56	0/1830
2	C	0.48	0/458	0.55	0/616
2	D	0.42	0/473	0.54	0/634
All	All	0.39	0/3624	0.55	0/4868

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	1316	0	1338	16	0
1	B	1346	0	1394	19	0
2	C	454	0	449	7	0
2	D	469	0	472	8	0
3	A	6	0	8	2	0
4	B	1	0	0	0	0
5	A	146	0	0	11	1
5	B	154	0	0	8	0
5	C	61	0	0	4	1

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	D	69	0	0	2	0
All	All	4022	0	3661	46	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 6.

The worst 5 of 46 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:49:GLN:NE2	5:A:601:HOH:O	1.94	0.97
1:A:49:GLN:OE1	5:A:602:HOH:O	1.95	0.84
2:C:845:GLN:OE1	5:C:901:HOH:O	1.94	0.83
2:D:860[A]:LYS:NZ	5:D:901:HOH:O	2.14	0.80
2:C:864[B]:ASP:OD2	5:C:902:HOH:O	1.98	0.80

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)
5:A:734:HOH:O	5:C:933:HOH:O[1_655]	1.92	0.28

## 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	164/172 (95%)	164 (100%)	0	0	100	100
1	B	167/172 (97%)	166 (99%)	1 (1%)	0	100	100
2	C	57/68 (84%)	56 (98%)	1 (2%)	0	100	100
2	D	59/68 (87%)	58 (98%)	1 (2%)	0	100	100
All	All	447/480 (93%)	444 (99%)	3 (1%)	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	149/153 (97%)	148 (99%)	1 (1%)	84	88
1	B	152/153 (99%)	151 (99%)	1 (1%)	84	88
2	C	51/59 (86%)	51 (100%)	0	100	100
2	D	53/59 (90%)	53 (100%)	0	100	100
All	All	405/424 (96%)	403 (100%)	2 (0%)	88	92

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	6	LYS
1	B	39	ASP

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](#)

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 for Mogul analysis.

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
3	GOL	A	501	-	5,5,5	0.94	0	5,5,5	0.90	0

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
3	GOL	A	501	-	-	0/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	501	GOL	2	0

## 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	155/172 (90%)	-0.20	0 <b>100</b>   <b>100</b>	10, 22, 44, 58	0
1	B	155/172 (90%)	-0.16	1 (0%) <b>89</b>   <b>88</b>	12, 24, 48, 59	0
2	C	54/68 (79%)	-0.32	1 (1%) <b>66</b>   <b>65</b>	12, 17, 34, 53	0
2	D	54/68 (79%)	-0.27	0 <b>100</b>   <b>100</b>	10, 16, 36, 57	0
All	All	418/480 (87%)	-0.21	2 (0%) <b>91</b>   <b>90</b>	10, 22, 46, 59	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	10	GLY	3.1
2	C	896	SER	2.3

### 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 monosaccharides 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
4	LI	B	201	1/1	0.81	0.79	18,18,18,18	0
3	GOL	A	501	6/6	0.89	0.20	28,43,62,66	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.