



# wwPDB X-ray Structure Validation Summary Report

May 14, 2014 – 06:33 PM EDT

PDB ID : 4NQK  
Title : Structure of an Ubiquitin complex  
Authors : Peisley, A.; Wu, B.; Hur, S.  
Deposited on : 2013-11-25  
Resolution : 3.70 Å(reported)

This is a wwPDB 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 <http://wwpdb.org/ValidationPDFNotes.html>

---

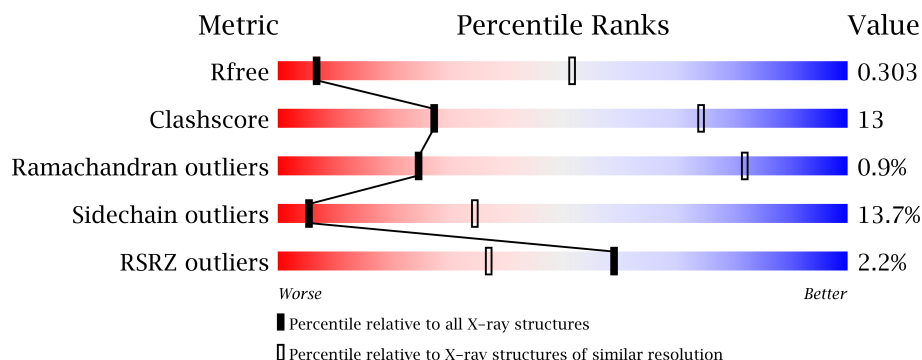
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.16 November 2013
Xtriage (Phenix)	:	dev-1439
EDS	:	stable22978
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22978

# 1 Overall quality at a glance

The reported resolution of this entry is 3.70 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))
$R_{free}$	66092	1098 (4.00-3.40)
Clashscore	79885	1009 (3.94-3.46)
Ramachandran outliers	78287	1016 (3.98-3.42)
Sidechain outliers	78261	1014 (3.98-3.42)
RSRZ outliers	66119	1099 (4.00-3.40)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	203	
1	B	203	
1	C	203	
1	D	203	
2	E	79	
2	F	79	
2	G	79	
2	H	79	
2	I	79	
2	J	79	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 19139 atoms, of which 9584 are hydrogens and 0 are deuterium.

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 Probable ATP-dependent RNA helicase DDX58.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	188	Total	C	H	N	O	S	0	0	0
			3041	990	1510	245	287	9			
1	B	188	Total	C	H	N	O	S	0	0	0
			3034	989	1505	243	288	9			
1	C	188	Total	C	H	N	O	S	0	0	0
			3000	981	1487	242	281	9			
1	D	188	Total	C	H	N	O	S	0	0	0
			3081	998	1535	248	291	9			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	EXPRESSION TAG	UNP O95786
A	-1	PRO	-	EXPRESSION TAG	UNP O95786
A	0	GLY	-	EXPRESSION TAG	UNP O95786
B	-2	GLY	-	EXPRESSION TAG	UNP O95786
B	-1	PRO	-	EXPRESSION TAG	UNP O95786
B	0	GLY	-	EXPRESSION TAG	UNP O95786
C	-2	GLY	-	EXPRESSION TAG	UNP O95786
C	-1	PRO	-	EXPRESSION TAG	UNP O95786
C	0	GLY	-	EXPRESSION TAG	UNP O95786
D	-2	GLY	-	EXPRESSION TAG	UNP O95786
D	-1	PRO	-	EXPRESSION TAG	UNP O95786
D	0	GLY	-	EXPRESSION TAG	UNP O95786

- Molecule 2 is a protein called Ubiquitin.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	E	72	Total	C	H	N	O	S	0	0	0
			1169	362	595	98	113	1			
2	F	73	Total	C	H	N	O	S	0	0	0
			1171	365	595	96	114	1			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	G	71	Total	C	H	N	O	S	0	0	0
			1128	352	570	93	112	1			
2	H	75	Total	C	H	N	O	S	0	0	0
			1202	373	611	101	116	1			
2	I	71	Total	C	H	N	O	S	0	0	0
			1125	352	570	92	110	1			
2	J	73	Total	C	H	N	O	S	0	0	0
			1188	368	606	99	114	1			

There are 12 discrepancies between the modelled and reference sequences:

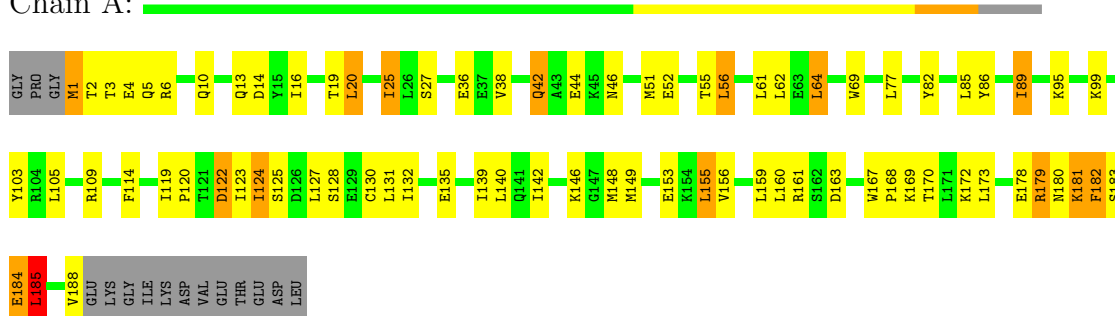
Chain	Residue	Modelled	Actual	Comment	Reference
E	-2	GLY	-	EXPRESSION TAG	UNP P0CG48
E	-1	PRO	-	EXPRESSION TAG	UNP P0CG48
F	-2	GLY	-	EXPRESSION TAG	UNP P0CG48
F	-1	PRO	-	EXPRESSION TAG	UNP P0CG48
G	-2	GLY	-	EXPRESSION TAG	UNP P0CG48
G	-1	PRO	-	EXPRESSION TAG	UNP P0CG48
H	-2	GLY	-	EXPRESSION TAG	UNP P0CG48
H	-1	PRO	-	EXPRESSION TAG	UNP P0CG48
I	-2	GLY	-	EXPRESSION TAG	UNP P0CG48
I	-1	PRO	-	EXPRESSION TAG	UNP P0CG48
J	-2	GLY	-	EXPRESSION TAG	UNP P0CG48
J	-1	PRO	-	EXPRESSION TAG	UNP P0CG48

### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: Probable ATP-dependent RNA helicase DDX58

Chain A:



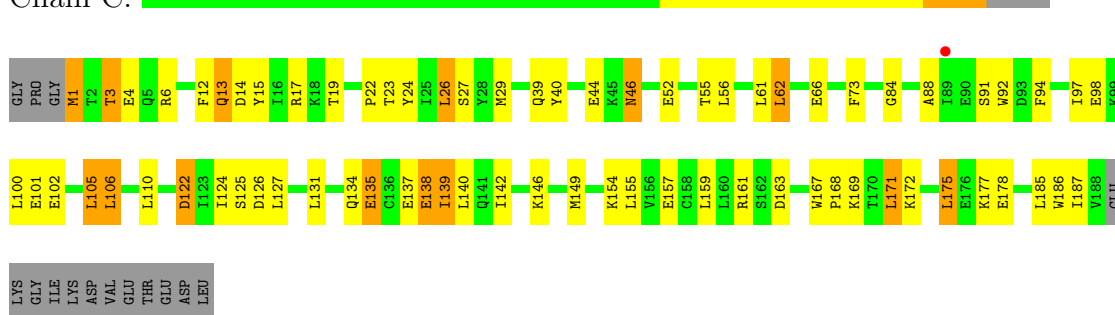
- Molecule 1: Probable ATP-dependent RNA helicase DDX58

Chain B:



- Molecule 1: Probable ATP-dependent RNA helicase DDX58

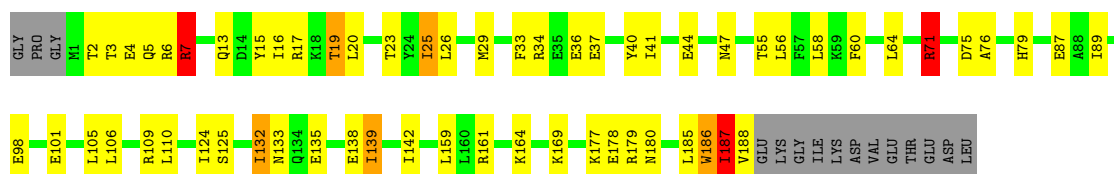
Chain C:



- Molecule 1: Probable ATP-dependent RNA helicase DDX58

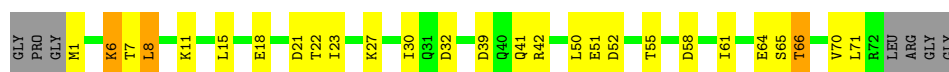
Chain D:





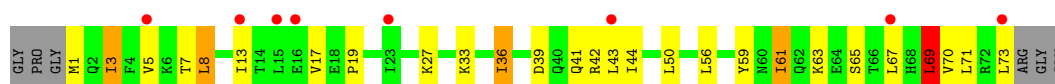
- Molecule 2: Ubiquitin

Chain E:



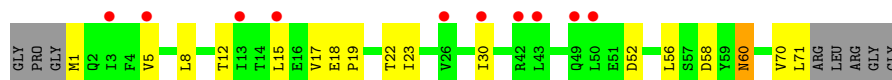
- Molecule 2: Ubiquitin

Chain F: 



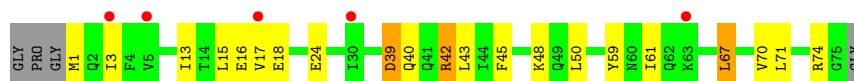
- Molecule 2: Ubiquitin

Chain G:



- Molecule 2: Ubiquitin

Chain H:



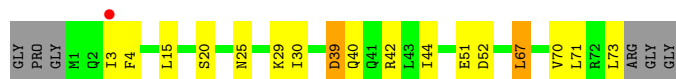
- Molecule 2: Ubiquitin

Chain I:



- Molecule 2: Ubiquitin

Chain J:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	85.54Å 101.85Å 88.15Å 90.00° 106.88° 90.00°	Depositor
Resolution (Å)	46.00 – 3.70 46.11 – 3.50	Depositor EDS
% Data completeness (in resolution range)	98.5 (46.00-3.70) 98.3 (46.11-3.50)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.16	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.17 (at 3.48Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.4_1496)	Depositor
R, $R_{free}$	0.222 , 0.285 0.242 , 0.303	Depositor DCC
$R_{free}$ test set	902 reflections (5.00%)	DCC
Wilson B-factor (Å <sup>2</sup> )	79.2	Xtriage
Anisotropy	0.189	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 46.4	EDS
Estimated twinning fraction	0.003 for l,-k,h	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtriage
Outliers	0 of 18048 reflections	Xtriage
$F_o, F_c$ correlation	0.88	EDS
Total number of atoms	19139	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	93.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.10% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 5 Model quality

### 5.1 Standard geometry

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	0.46	0/1563	0.83	1/2110 (0.0%)
1	B	0.46	0/1561	0.79	1/2108 (0.0%)
1	C	0.39	0/1545	0.75	2/2088 (0.1%)
1	D	0.42	0/1578	0.81	5/2128 (0.2%)
2	E	0.33	0/580	0.65	0/781
2	F	0.31	0/582	0.64	1/785 (0.1%)
2	G	0.35	0/564	0.66	0/761
2	H	0.32	0/597	0.65	0/804
2	I	0.65	0/561	0.99	4/757 (0.5%)
2	J	0.35	0/588	0.65	0/792
All	All	0.42	0/9719	0.77	14/13114 (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	A	0	1
1	B	0	1
1	D	0	1
2	I	0	3
All	All	0	6

There are no bond length outliers.

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	71	ARG	NE-CZ-NH1	8.41	124.51	120.30
2	I	67	LEU	N-CA-C	-7.84	89.83	111.00
1	D	71	ARG	NE-CZ-NH2	-7.15	116.73	120.30
1	A	185	LEU	CA-CB-CG	6.31	129.81	115.30
1	C	26	LEU	CA-CB-CG	6.04	129.19	115.30



There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	179	ARG	Peptide
1	B	13	GLN	Peptide
1	D	186	TRP	Peptide
2	I	3	ILE	Peptide
2	I	66	THR	Peptide

## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1531	1510	0	52	0
1	B	1529	1505	0	45	0
1	C	1513	1487	0	51	0
1	D	1546	1535	0	35	0
2	E	574	595	0	17	0
2	F	576	595	0	16	0
2	G	558	570	0	5	0
2	H	591	611	0	12	0
2	I	555	570	0	23	0
2	J	582	606	0	7	0
All	All	9555	9584	0	246	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 13.

The worst 5 of 246 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:135:GLU:OE1	1:A:161:ARG:NH1	2.17	0.77
1:C:154:LYS:NZ	1:C:157:GLU:OE2	2.19	0.76
1:D:179:ARG:CB	1:D:180:ASN:HA	2.18	0.73
1:D:15:TYR:O	1:D:19:THR:OG1	2.05	0.73
1:C:6:ARG:NH1	1:C:62:LEU:O	2.22	0.72

There are no symmetry-related clashes.

## 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	186/203 (92%)	172 (92%)	11 (6%)	3 (2%)	14	72
1	B	186/203 (92%)	171 (92%)	14 (8%)	1 (0%)	38	88
1	C	186/203 (92%)	170 (91%)	13 (7%)	3 (2%)	14	72
1	D	186/203 (92%)	176 (95%)	8 (4%)	2 (1%)	21	79
2	E	70/79 (89%)	69 (99%)	1 (1%)	0	100	100
2	F	71/79 (90%)	69 (97%)	2 (3%)	0	100	100
2	G	69/79 (87%)	68 (99%)	1 (1%)	0	100	100
2	H	73/79 (92%)	71 (97%)	2 (3%)	0	100	100
2	I	69/79 (87%)	65 (94%)	3 (4%)	1 (1%)	16	74
2	J	71/79 (90%)	69 (97%)	2 (3%)	0	100	100
All	All	1167/1286 (91%)	1100 (94%)	57 (5%)	10 (1%)	25	82

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	187	ILE
1	C	66	GLU
1	C	139	ILE
1	D	139	ILE
1	A	180	ASN

### 5.3.2 Protein sidechains ⓘ

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	162/182 (89%)	137 (85%)	25 (15%)	4	28
1	B	162/182 (89%)	142 (88%)	20 (12%)	7	40
1	C	158/182 (87%)	135 (85%)	23 (15%)	5	31
1	D	166/182 (91%)	142 (86%)	24 (14%)	5	31
2	E	66/69 (96%)	58 (88%)	8 (12%)	7	41
2	F	66/69 (96%)	57 (86%)	9 (14%)	5	35
2	G	64/69 (93%)	55 (86%)	9 (14%)	5	33
2	H	67/69 (97%)	58 (87%)	9 (13%)	6	36
2	I	63/69 (91%)	54 (86%)	9 (14%)	5	32
2	J	67/69 (97%)	60 (90%)	7 (10%)	10	50
All	All	1041/1142 (91%)	898 (86%)	143 (14%)	5	34

5 of 143 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	171	LEU
1	D	71	ARG
2	I	52	ASP
1	C	185	LEU
1	D	17	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
2	E	41	GLN
2	F	41	GLN
2	H	25	ASN
1	D	47	ASN
2	H	60	ASN

### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

There are no ligands in this entry.

## 5.7 Other polymers ⓘ

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 ⓘ

### 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, 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	188/203 (92%)	0.10	0 100 100	29, 64, 122, 168	0
1	B	188/203 (92%)	0.09	0 100 100	37, 79, 125, 170	0
1	C	188/203 (92%)	0.16	1 (0%) 88 73	43, 87, 135, 163	0
1	D	188/203 (92%)	0.08	0 100 100	36, 68, 106, 128	0
2	E	72/79 (91%)	0.23	0 100 100	45, 78, 121, 145	0
2	F	73/79 (92%)	0.74	8 (10%) 6 6	86, 147, 188, 214	1 (1%)
2	G	71/79 (89%)	0.99	10 (14%) 3 4	109, 167, 209, 230	0
2	H	75/79 (94%)	0.71	5 (6%) 17 13	79, 130, 172, 200	1 (1%)
2	I	71/79 (89%)	0.22	1 (1%) 72 48	65, 98, 147, 170	0
2	J	73/79 (92%)	0.31	1 (1%) 72 48	46, 73, 111, 140	1 (1%)
All	All	1187/1286 (92%)	0.26	26 (2%) 59 37	29, 83, 167, 230	3 (0%)

The worst 5 of 26 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	G	15	LEU	4.0
2	H	3	ILE	3.8
2	G	26	VAL	3.3
2	G	50	LEU	3.1
2	G	30	ILE	3.0

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

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

### 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

### 6.4 Ligands ⓘ

There are no ligands in this entry.

### 6.5 Other polymers ⓘ

There are no such residues in this entry.