



Full wwPDB X-ray Structure Validation Report ⓘ

Dec 3, 2024 – 10:02 PM EST

PDB ID : 9EGV
Title : HOIL-1 RING2 domain bound to ubiquitin
Authors : Wang, X.S.; Lechtenberg, B.C.
Deposited on : 2024-11-21
Resolution : 2.00 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

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 : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 1.21
EDS : 3.0
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.004 (Gargrove)
Density-Fitness : 1.0.11
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.40

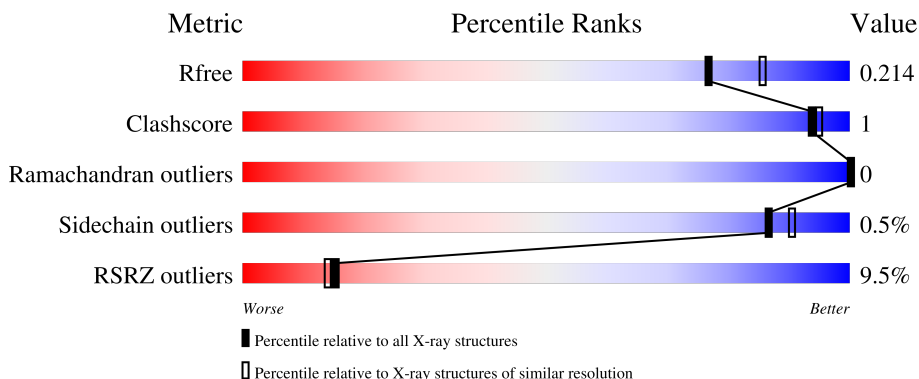
1 Overall quality at a glance i

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}	164625	9409 (2.00-2.00)
Clashscore	180529	10737 (2.00-2.00)
Ramachandran outliers	177936	10628 (2.00-2.00)
Sidechain outliers	177891	10627 (2.00-2.00)
RSRZ outliers	164620	9409 (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 ≥ 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	101	
1	B	101	
2	C	88	

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 3823 atoms, of which 1833 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 RanBP-type and C3HC4-type zinc finger-containing protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
1	A	87	Total	C	H	N	O	S	0	0	0
			1286	397	628	128	120	13			
1	B	80	Total	C	H	N	O	S	0	0	0
			1183	367	578	116	109	13			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	410	GLY	-	expression tag	UNP Q9BYM8
A	411	PRO	-	expression tag	UNP Q9BYM8
A	460	ALA	CYS	engineered mutation	UNP Q9BYM8
B	410	GLY	-	expression tag	UNP Q9BYM8
B	411	PRO	-	expression tag	UNP Q9BYM8
B	460	ALA	CYS	engineered mutation	UNP Q9BYM8

- Molecule 2 is a protein called Polyubiquitin-C.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
2	C	76	Total	C	H	N	O	S	0	0	0
			1229	378	627	105	118	1			

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-11	MET	-	initiating methionine	UNP P0CG48
C	-10	ALA	-	expression tag	UNP P0CG48
C	-9	HIS	-	expression tag	UNP P0CG48
C	-8	HIS	-	expression tag	UNP P0CG48
C	-7	HIS	-	expression tag	UNP P0CG48
C	-6	HIS	-	expression tag	UNP P0CG48
C	-5	HIS	-	expression tag	UNP P0CG48
C	-4	HIS	-	expression tag	UNP P0CG48

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-3	SER	-	expression tag	UNP P0CG48
C	-2	ALA	-	expression tag	UNP P0CG48
C	-1	ALA	-	expression tag	UNP P0CG48

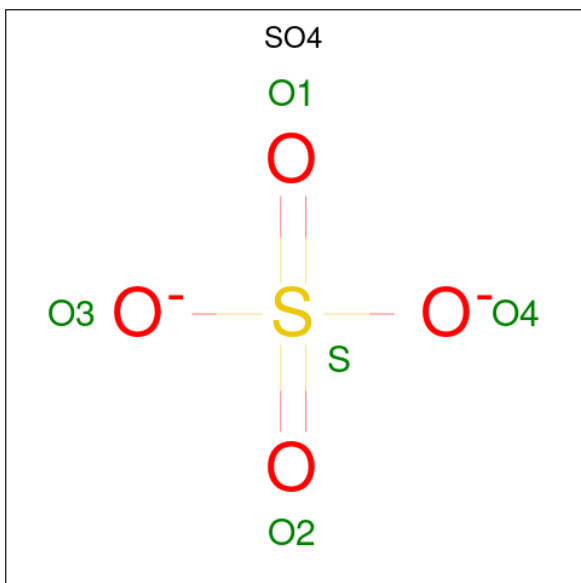
- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	3	Total Zn 3 3	0	0
3	B	3	Total Zn 3 3	0	0

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Cl 1 1	0	0

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total O S 5 4 1	0	0
5	B	1	Total O S 5 4 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	C	1	Total	O	S	0	0
			5	4	1		

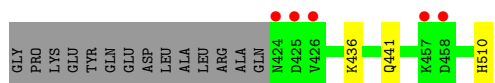
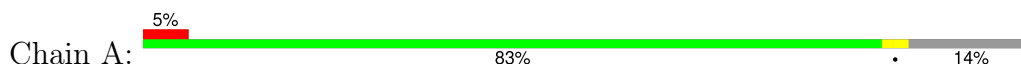
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	55	Total	O	0	0
			55	55		
6	B	44	Total	O	0	0
			44	44		
6	C	4	Total	O	0	0
			4	4		

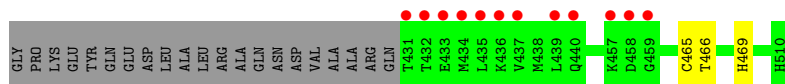
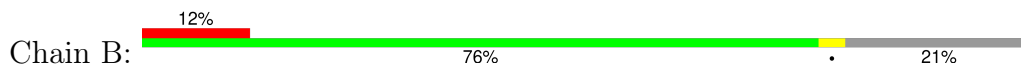
3 Residue-property plots [i](#)

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.

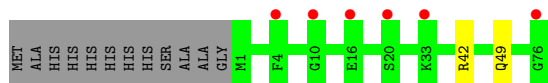
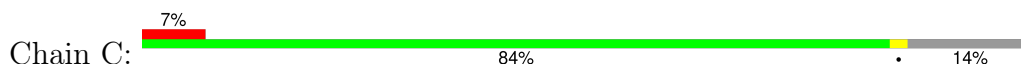
- Molecule 1: RanBP-type and C3HC4-type zinc finger-containing protein 1



- Molecule 1: RanBP-type and C3HC4-type zinc finger-containing protein 1



- Molecule 2: Polyubiquitin-C



4 Data and refinement statistics

Property	Value	Source
Space group	I 41 2 2	Depositor
Cell constants a, b, c, α , β , γ	95.92Å 95.92Å 187.43Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.96 – 2.00 47.96 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.9 (47.96-2.00) 89.0 (47.96-2.00)	Depositor EDS
R_{merge}	0.23	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.33 (at 2.00Å)	Xtrriage
Refinement program	PHENIX 1.21.1_5286	Depositor
R, R_{free}	0.180 , 0.208 0.187 , 0.214	Depositor DCC
R_{free} test set	1576 reflections (5.21%)	wwPDB-VP
Wilson B-factor (Å ²)	26.5	Xtrriage
Anisotropy	0.279	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 41.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	3823	wwPDB-VP
Average B, all atoms (Å ²)	61.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.83% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: ZN, SO4, CL

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	0.56	0/672	0.73	0/908
1	B	0.54	0/619	0.69	0/836
2	C	0.36	0/608	0.56	0/816
All	All	0.50	0/1899	0.67	0/2560

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	658	628	628	2	0
1	B	605	578	578	2	0
2	C	602	627	629	1	0
3	A	3	0	0	0	0
3	B	3	0	0	0	0
4	A	1	0	0	0	0
5	B	10	0	0	0	0
5	C	5	0	0	0	0
6	A	55	0	0	2	0
6	B	44	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	C	4	0	0	0	0
All	All	1990	1833	1835	5	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 1.

All (5) 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:C:42:ARG:NE	2:C:49:GLN:OE1	2.21	0.72
1:B:466:THR:HG22	6:B:728:HOH:O	2.04	0.58
1:A:436:LYS:NZ	6:A:701:HOH:O	2.25	0.53
1:A:441:GLN:NE2	6:A:702:HOH:O	2.43	0.51
1:B:465:CYS:O	1:B:469:HIS:HA	2.21	0.40

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	85/101 (84%)	81 (95%)	4 (5%)	0	100	100
1	B	78/101 (77%)	77 (99%)	1 (1%)	0	100	100
2	C	74/88 (84%)	73 (99%)	1 (1%)	0	100	100
All	All	237/290 (82%)	231 (98%)	6 (2%)	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	73/84 (87%)	72 (99%)	1 (1%)	62	68
1	B	68/84 (81%)	68 (100%)	0	100	100
2	C	68/76 (90%)	68 (100%)	0	100	100
All	All	209/244 (86%)	208 (100%)	1 (0%)	86	90

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	510	HIS

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 10 ligands modelled in this entry, 7 are monoatomic - leaving 3 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	$\# Z > 2$
5	SO4	C	101	-	4,4,4	0.82	0	6,6,6	0.31	0
5	SO4	B	604	-	4,4,4	0.80	0	6,6,6	0.38	0
5	SO4	B	605	-	4,4,4	0.69	0	6,6,6	0.10	0

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.

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

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, 95th 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 ²)	Q<0.9
1	A	87/101 (86%)	-0.02	5 (5%) 30 28	27, 37, 84, 126	0
1	B	80/101 (79%)	0.40	12 (15%) 6 5	30, 40, 121, 129	0
2	C	76/88 (86%)	1.18	6 (7%) 20 19	46, 85, 118, 126	0
All	All	243/290 (83%)	0.49	23 (9%) 15 14	27, 48, 116, 129	0

All (23) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	432	THR	5.6
1	B	431	THR	4.7
1	B	436	LYS	4.4
1	B	434	MET	4.4
1	B	437	VAL	4.3
1	B	440	GLN	3.9
1	B	458	ASP	3.4
1	B	435	LEU	3.2
2	C	10	GLY	3.1
1	A	424	ASN	3.0
1	A	425	ASP	3.0
1	B	433	GLU	2.9
1	A	457	LYS	2.9
1	A	426	VAL	2.7
1	B	457	LYS	2.6
1	A	458	ASP	2.5
1	B	439	LEU	2.4
2	C	33	LYS	2.4
1	B	459	GLY	2.3
2	C	76	GLY	2.2
2	C	16	GLU	2.1
2	C	4	PHE	2.1
2	C	20	SER	2.1

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, 95th 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(Å ²)	Q<0.9
5	SO4	B	605	5/5	0.54	0.13	88,103,124,153	0
5	SO4	B	604	5/5	0.83	0.10	63,68,88,105	0
5	SO4	C	101	5/5	0.90	0.10	65,67,75,86	0
4	CL	A	604	1/1	0.99	0.06	33,33,33,33	0
3	ZN	A	603	1/1	0.99	0.02	27,27,27,27	0
3	ZN	B	601	1/1	0.99	0.02	29,29,29,29	0
3	ZN	B	603	1/1	0.99	0.02	31,31,31,31	0
3	ZN	B	602	1/1	1.00	0.04	29,29,29,29	0
3	ZN	A	601	1/1	1.00	0.02	29,29,29,29	0
3	ZN	A	602	1/1	1.00	0.02	25,25,25,25	0

6.5 Other polymers [i](#)

There are no such residues in this entry.