



## wwPDB EM Validation Summary Report ⓘ

Nov 21, 2022 – 12:20 pm GMT

PDB ID : 7ZF1  
EMDB ID : EMD-14694  
Title : Structure of ubiquitinated FANCI in complex with FANCD2 and double-stranded DNA  
Authors : Lemonidis, K.; Rennie, M.L.; Arkinson, C.; Streetley, J.; Clarke, M.; Chaugule, V.K.; Walden, H.  
Deposited on : 2022-03-31  
Resolution : 4.14 Å (reported)  
Based on initial model : 6VAE

This is a wwPDB EM 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/EMValidationReportHelp>

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:

EMDB validation analysis : 0.0.0.dev97  
MolProbity : 4.02b-467  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.26

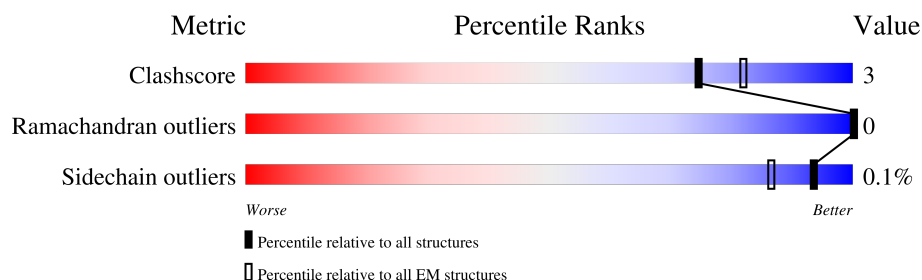
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 4.14 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)	EM structures (#Entries)
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the 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  $\leq 5\%$ .

Mol	Chain	Length	Quality of chain
1	A	1355	71% 6% 23%
2	B	1468	73% 7% 20%
3	C	80	91% . 5%
4	S	29	52% 41% 7%
5	T	29	48% 45% 7%

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 38889 atoms, of which 19431 are hydrogens and 0 are deuteriums.

In the tables below, 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 Fanconi anemia group I protein.

Mol	Chain	Residues	Atoms						AltConf	Trace
1	A	1042	Total	C	H	N	O	S	0	0
			16846	5311	8555	1387	1544	49		

There are 27 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-26	MET	-	initiating methionine	UNP Q9NVI1
A	-25	HIS	-	expression tag	UNP Q9NVI1
A	-24	HIS	-	expression tag	UNP Q9NVI1
A	-23	HIS	-	expression tag	UNP Q9NVI1
A	-22	HIS	-	expression tag	UNP Q9NVI1
A	-21	HIS	-	expression tag	UNP Q9NVI1
A	-20	HIS	-	expression tag	UNP Q9NVI1
A	-19	GLU	-	expression tag	UNP Q9NVI1
A	-18	ASN	-	expression tag	UNP Q9NVI1
A	-17	LEU	-	expression tag	UNP Q9NVI1
A	-16	TYR	-	expression tag	UNP Q9NVI1
A	-15	PHE	-	expression tag	UNP Q9NVI1
A	-14	GLN	-	expression tag	UNP Q9NVI1
A	-13	GLY	-	expression tag	UNP Q9NVI1
A	-12	LYS	-	expression tag	UNP Q9NVI1
A	-11	PRO	-	expression tag	UNP Q9NVI1
A	-10	ILE	-	expression tag	UNP Q9NVI1
A	-9	PRO	-	expression tag	UNP Q9NVI1
A	-8	ASN	-	expression tag	UNP Q9NVI1
A	-7	PRO	-	expression tag	UNP Q9NVI1
A	-6	LEU	-	expression tag	UNP Q9NVI1
A	-5	LEU	-	expression tag	UNP Q9NVI1
A	-4	GLY	-	expression tag	UNP Q9NVI1
A	-3	LEU	-	expression tag	UNP Q9NVI1
A	-2	ASP	-	expression tag	UNP Q9NVI1
A	-1	SER	-	expression tag	UNP Q9NVI1
A	0	THR	-	expression tag	UNP Q9NVI1

- Molecule 2 is a protein called Fanconi anemia group D2 protein.

Mol	Chain	Residues	Atoms						AltConf	Trace
2	B	1179	Total	C	H	N	O	S	0	0
			19103	6099	9644	1566	1741	53		

There are 17 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-16	MET	-	initiating methionine	UNP Q9BXW9
B	-15	HIS	-	expression tag	UNP Q9BXW9
B	-14	HIS	-	expression tag	UNP Q9BXW9
B	-13	HIS	-	expression tag	UNP Q9BXW9
B	-12	HIS	-	expression tag	UNP Q9BXW9
B	-11	HIS	-	expression tag	UNP Q9BXW9
B	-10	HIS	-	expression tag	UNP Q9BXW9
B	-9	LEU	-	expression tag	UNP Q9BXW9
B	-8	GLU	-	expression tag	UNP Q9BXW9
B	-7	VAL	-	expression tag	UNP Q9BXW9
B	-6	LEU	-	expression tag	UNP Q9BXW9
B	-5	PHE	-	expression tag	UNP Q9BXW9
B	-4	GLN	-	expression tag	UNP Q9BXW9
B	-3	GLY	-	expression tag	UNP Q9BXW9
B	-2	PRO	-	expression tag	UNP Q9BXW9
B	-1	GLY	-	expression tag	UNP Q9BXW9
B	0	SER	-	expression tag	UNP Q9BXW9

- Molecule 3 is a protein called Ubiquitin-60S ribosomal protein L40.

Mol	Chain	Residues	Atoms						AltConf	Trace
3	C	76	Total	C	H	N	O	S	0	0
			1228	378	627	105	117	1		

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-3	GLY	-	expression tag	UNP P62987
C	-2	PRO	-	expression tag	UNP P62987
C	-1	GLY	-	expression tag	UNP P62987
C	0	SER	-	expression tag	UNP P62987

- Molecule 4 is a DNA chain called DNA (61-MER).

Mol	Chain	Residues	Atoms						AltConf	Trace
4	S	27	Total	C	H	N	O	P	0	0
			860	263	302	106	162	27		

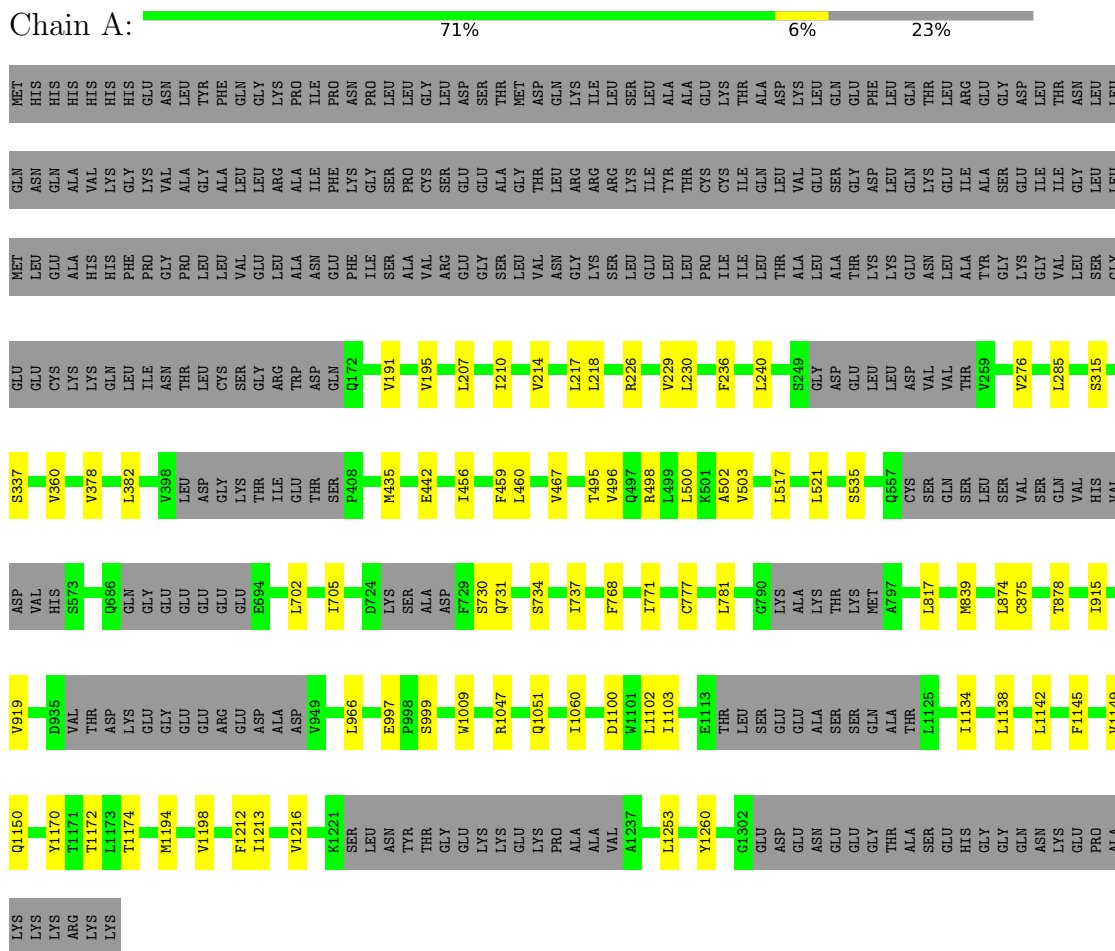
- Molecule 5 is a DNA chain called DNA (61-MER).

Mol	Chain	Residues	Atoms						AltConf	Trace
5	T	27	Total	C	H	N	O	P	0	0
			852	260	303	100	162	27		

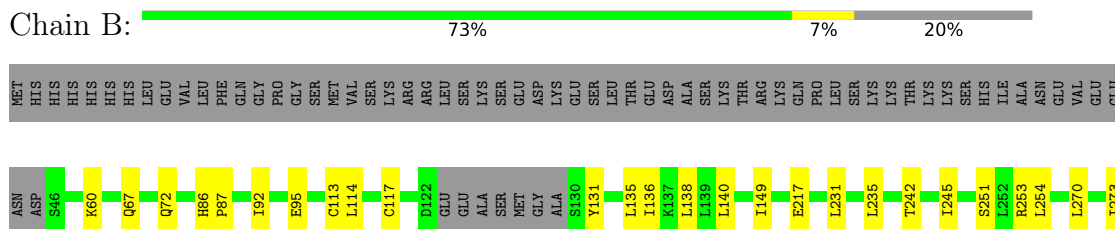
### 3 Residue-property plots

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. 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. 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.

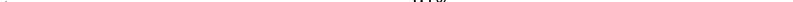
- Molecule 1: Fanconi anemia group I protein



- Molecule 2: Fanconi anemia group D2 protein



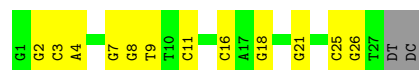
- Molecule 3: Ubiquitin-60S ribosomal protein L40

Chain C:  91% 5%



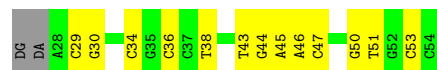
- Molecule 4: DNA (61-MER)

Chain S:  52% 41% 7%



- Molecule 5: DNA (61-MER)

Chain T:  48% 45% 7%



## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	139601	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	JEOL CRYO ARM 300	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	46	Depositor
Minimum defocus (nm)	500	Depositor
Maximum defocus (nm)	3800	Depositor
Magnification	Not provided	
Image detector	DIRECT ELECTRON DE-64 (8k x 8k)	Depositor



## 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	0.29	0/8422	0.55	0/11359
2	B	0.29	0/9637	0.55	1/13036 (0.0%)
3	C	0.30	0/607	0.57	0/816
4	S	0.63	0/626	0.97	0/965
5	T	0.65	0/614	0.97	0/944
All	All	0.32	0/19906	0.59	1/27120 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
2	B	346	LEU	CA-CB-CG	5.08	126.98	115.30

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	8291	8555	8547	49	0
2	B	9459	9644	9638	59	0
3	C	601	627	629	2	0
4	S	558	302	303	14	0
5	T	549	303	303	14	0
All	All	19458	19431	19420	123	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 3.

The worst 5 of 123 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:B:231:LEU:HD13	2:B:245:ILE:HG23	1.75	0.68
1:A:467:VAL:HG21	1:A:502:ALA:HB1	1.78	0.66
4:S:18:DG:O6	5:T:36:DC:N4	2.32	0.63
1:A:460:LEU:HD23	1:A:498:ARG:HD2	1.80	0.62
2:B:768:GLY:O	2:B:772:GLU:N	2.32	0.62

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 PDB entries followed by that with respect to all EM entries.

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	1022/1355 (75%)	989 (97%)	33 (3%)	0	100	100
2	B	1157/1468 (79%)	1131 (98%)	26 (2%)	0	100	100
3	C	74/80 (92%)	73 (99%)	1 (1%)	0	100	100
All	All	2253/2903 (78%)	2193 (97%)	60 (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 PDB entries followed by that with respect to all EM entries.

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	962/1227 (78%)	962 (100%)	0	100	100
2	B	1084/1339 (81%)	1082 (100%)	2 (0%)	93	96
3	C	68/70 (97%)	68 (100%)	0	100	100
All	All	2114/2636 (80%)	2112 (100%)	2 (0%)	93	97

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

Mol	Chain	Res	Type
2	B	292	MET
2	B	431	MET

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 [i](#)

There are no chain breaks in this entry.

## 6 Map visualisation

This section contains visualisations of the EMDB entry EMD-14694. These allow visual inspection of the internal detail of the map and identification of artifacts.

Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

### 6.1 Orthogonal projections

This section was not generated.

### 6.2 Central slices

This section was not generated.

### 6.3 Largest variance slices

This section was not generated.

### 6.4 Orthogonal surface views

This section was not generated.

### 6.5 Mask visualisation

This section was not generated. No masks/segmentation were deposited.

## 7 Map analysis ⓘ

This section contains the results of statistical analysis of the map.

### 7.1 Map-value distribution ⓘ

This section was not generated.

### 7.2 Volume estimate versus contour level ⓘ

This section was not generated.

### 7.3 Rotationally averaged power spectrum ⓘ

This section was not generated. The rotationally averaged power spectrum had issues being displayed.

## 8 Fourier-Shell correlation

This section was not generated. No FSC curve or half-maps provided.

## 9 Map-model fit

This section was not generated.