



wwPDB X-ray Structure Validation Summary Report i

Feb 28, 2014 – 01:17 AM GMT

PDB ID : 3S4W
Title : Structure of the FANCI-FANCD2 complex
Authors : Pavletich, N.P.
Deposited on : 2011-05-20
Resolution : 3.41 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

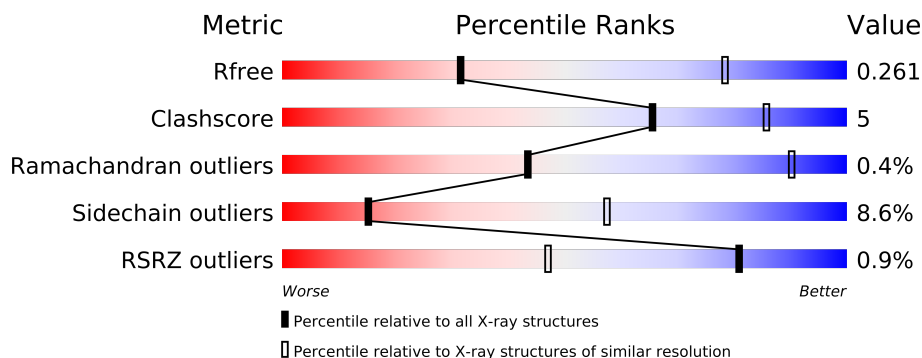
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
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)	:	stable22683

1 Overall quality at a glance

The reported resolution of this entry is 3.41 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	1017 (3.52-3.28)
Clashscore	79885	1214 (3.50-3.30)
Ramachandran outliers	78287	1177 (3.50-3.30)
Sidechain outliers	78261	1177 (3.50-3.30)
RSRZ outliers	66119	1017 (3.52-3.28)

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 ≥ 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	1308	
2	B	1323	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 18671 atoms, of which 0 are hydrogen 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 Fanconi anemia group I protein homolog.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1206	Total	C	N	O	S	0	0	0
			9506	6093	1579	1778	56			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1303	HIS	-	EXPRESSION TAG	UNP Q8K368
A	1304	HIS	-	EXPRESSION TAG	UNP Q8K368
A	1305	HIS	-	EXPRESSION TAG	UNP Q8K368
A	1306	HIS	-	EXPRESSION TAG	UNP Q8K368
A	1307	HIS	-	EXPRESSION TAG	UNP Q8K368
A	1308	HIS	-	EXPRESSION TAG	UNP Q8K368

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1146	Total	C	N	O	S	0	0	0
			9165	5905	1543	1664	53			

There are 60 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	?	-	SER	DELETION	UNP Q80V62
B	?	-	SER	DELETION	UNP Q80V62
B	?	-	SER	DELETION	UNP Q80V62
B	?	-	ALA	DELETION	UNP Q80V62
B	?	-	VAL	DELETION	UNP Q80V62
B	?	-	ALA	DELETION	UNP Q80V62
B	?	-	ALA	DELETION	UNP Q80V62
B	?	-	LYS	DELETION	UNP Q80V62
B	?	-	ASN	DELETION	UNP Q80V62
B	?	-	ARG	DELETION	UNP Q80V62

Continued on next page...

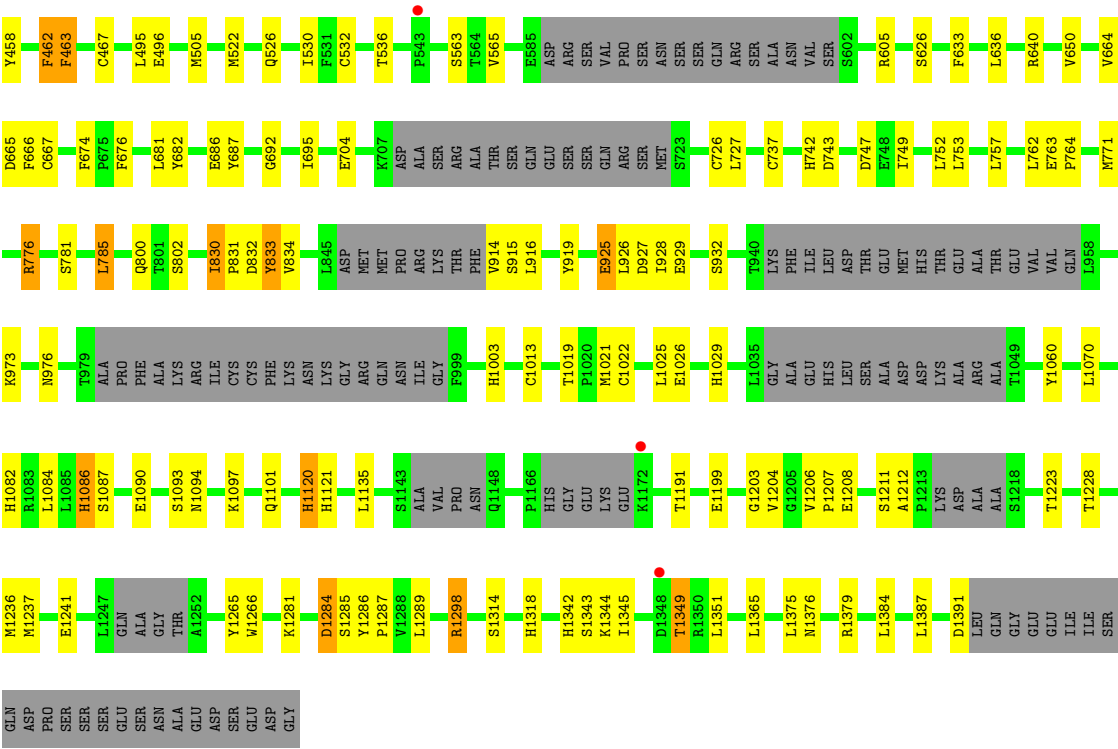
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	?	-	ASN	DELETION	UNP Q80V62
B	?	-	LYS	DELETION	UNP Q80V62
B	?	-	GLY	DELETION	UNP Q80V62
B	?	-	LYS	DELETION	UNP Q80V62
B	?	-	THR	DELETION	UNP Q80V62
B	?	-	GLY	DELETION	UNP Q80V62
B	?	-	GLY	DELETION	UNP Q80V62
B	?	-	LYS	DELETION	UNP Q80V62
B	?	-	LYS	DELETION	UNP Q80V62
B	?	-	GLN	DELETION	UNP Q80V62
B	?	-	LYS	DELETION	UNP Q80V62
B	?	-	ALA	DELETION	UNP Q80V62
B	?	-	ASP	DELETION	UNP Q80V62
B	?	-	SER	DELETION	UNP Q80V62
B	?	-	ASN	DELETION	UNP Q80V62
B	?	-	LYS	DELETION	UNP Q80V62
B	?	-	ALA	DELETION	UNP Q80V62
B	?	-	SER	DELETION	UNP Q80V62
B	?	-	CYS	DELETION	UNP Q80V62
B	?	-	SER	DELETION	UNP Q80V62
B	?	-	ASP	DELETION	UNP Q80V62
B	?	-	THR	DELETION	UNP Q80V62
B	?	-	LEU	DELETION	UNP Q80V62
B	?	-	LEU	DELETION	UNP Q80V62
B	?	-	THR	DELETION	UNP Q80V62
B	?	-	GLU	DELETION	UNP Q80V62
B	?	-	ASP	DELETION	UNP Q80V62
B	?	-	THR	DELETION	UNP Q80V62
B	?	-	SER	DELETION	UNP Q80V62
B	?	-	GLU	DELETION	UNP Q80V62
B	?	-	CYS	DELETION	UNP Q80V62
B	?	-	ASP	DELETION	UNP Q80V62
B	?	-	MET	DELETION	UNP Q80V62
B	?	-	ALA	DELETION	UNP Q80V62
B	?	-	PRO	DELETION	UNP Q80V62
B	?	-	SER	DELETION	UNP Q80V62
B	?	-	GLY	DELETION	UNP Q80V62
B	?	-	ARG	DELETION	UNP Q80V62
B	?	-	SER	DELETION	UNP Q80V62
B	?	-	HIS	DELETION	UNP Q80V62
B	?	-	VAL	DELETION	UNP Q80V62
B	?	-	ASP	DELETION	UNP Q80V62

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	?	-	LYS	DELETION	UNP Q80V62
B	?	-	GLU	DELETION	UNP Q80V62
B	?	-	SER	DELETION	UNP Q80V62
B	?	-	THR	DELETION	UNP Q80V62
B	?	-	GLY	DELETION	UNP Q80V62
B	?	-	LYS	DELETION	UNP Q80V62
B	?	-	GLU	DELETION	UNP Q80V62
B	?	-	GLY	DELETION	UNP Q80V62



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	83.80Å 110.40Å 350.50Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.86 – 3.41 20.02 – 3.41	Depositor EDS
% Data completeness (in resolution range)	90.0 (19.86-3.41) 88.9 (20.02-3.41)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.72 (at 3.44Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.1_743)	Depositor
R, R_{free}	0.240 , 0.272 0.228 , 0.261	Depositor DCC
R_{free} test set	1622 reflections (4.04%)	DCC
Wilson B-factor (Å ²)	90.3	Xtriage
Anisotropy	0.261	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 53.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 42934 reflections	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	18671	wwPDB-VP
Average B, all atoms (Å ²)	122.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.68% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities 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	0.28	0/9655	0.47	0/13038
2	B	0.26	0/9336	0.46	0/12621
All	All	0.27	0/18991	0.46	0/25659

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
2	B	0	3
All	All	0	4

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1062	ILE	Peptide
2	B	1094	ASN	Peptide
2	B	1284	ASP	Peptide
2	B	290	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	9506	0	0	44	0
2	B	9165	0	0	53	0
All	All	18671	0	0	97	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 5.

The worst 5 of 97 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:B:771:MET:O	2:B:776:ARG:NH1	2.16	0.78
1:A:34:GLN:O	1:A:37:ALA:N	2.18	0.76
1:A:373:HIS:O	1:A:374:SER:OG	2.06	0.73
1:A:1285:ASP:OD1	1:A:1286:PHE:N	2.21	0.73
1:A:818:ARG:NH1	1:A:825:GLU:OE1	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	1190/1308 (91%)	1108 (93%)	75 (6%)	7 (1%)	33	85
2	B	1120/1323 (85%)	1043 (93%)	75 (7%)	2 (0%)	56	95
All	All	2310/2631 (88%)	2151 (93%)	150 (6%)	9 (0%)	43	90

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	122	VAL
1	A	568	VAL
2	B	1285	SER
1	A	25	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	717	LEU

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	1093/1188 (92%)	992 (91%)	101 (9%)	13	52
2	B	1030/1193 (86%)	949 (92%)	81 (8%)	18	62
All	All	2123/2381 (89%)	1941 (91%)	182 (9%)	15	57

5 of 182 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1011	LYS
2	B	56	THR
2	B	1266	TRP
1	A	1021	CYS
1	A	1206	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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, 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	1206/1308 (92%)	-0.12	17 (1%)	72 35	48, 100, 204, 311	0
2	B	1146/1323 (86%)	-0.06	4 (0%)	91 73	74, 127, 192, 236	0
All	All	2352/2631 (89%)	-0.09	21 (0%)	81 47	48, 116, 198, 311	0

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	397	LYS	4.3
1	A	399	LEU	4.1
1	A	14	THR	3.8
2	B	543	PRO	3.5
1	A	1189	ARG	3.4

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.