



wwPDB X-ray Structure Validation Summary Report

Nov 2, 2023 – 07:03 PM EDT

PDB ID : 3W03
Title : XLF-XRCC4 complex
Authors : Wu, Q.; Ochi, T.; Matak-Vinkovic, D.; Robinson, C.V.; Chirgadze, D.Y.;
Blundell, T.L.
Deposited on : 2012-10-17
Resolution : 8.49 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

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
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

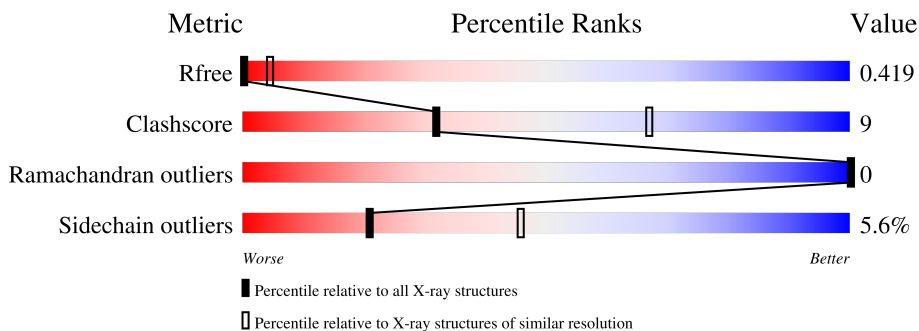
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 8.49 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	1005 (11.50-3.90)
Clashscore	141614	1070 (11.50-3.90)
Ramachandran outliers	138981	1003 (11.50-3.90)
Sidechain outliers	138945	1003 (11.50-3.86)

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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	235	75% (green), 21% (yellow), 4% (orange), 1% (red), 1% (grey)
1	B	235	76% (green), 18% (yellow), 4% (orange), 1% (red), 1% (grey)
2	C	184	57% (green), 30% (yellow), 11% (orange), 2% (red), 1% (grey)
2	D	184	58% (green), 29% (yellow), 11% (orange), 2% (red), 1% (grey)

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 6229 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 Non-homologous end-joining factor 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	231	Total 1818	C 1159	N 306	O 338	S 15	0	0	0
1	B	225	Total 1757	C 1126	N 292	O 324	S 15	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	SER	-	expression tag	UNP Q9H9Q4
A	0	GLY	-	expression tag	UNP Q9H9Q4
B	-1	SER	-	expression tag	UNP Q9H9Q4
B	0	GLY	-	expression tag	UNP Q9H9Q4

- Molecule 2 is a protein called DNA repair protein XRCC4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	164	Total 1327	C 838	N 224	O 259	S 6	0	0	0
2	D	164	Total 1327	C 838	N 224	O 259	S 6	0	0	0

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-19	MET	-	expression tag	UNP Q13426
C	-18	GLY	-	expression tag	UNP Q13426
C	-17	SER	-	expression tag	UNP Q13426
C	-16	SER	-	expression tag	UNP Q13426
C	-15	HIS	-	expression tag	UNP Q13426
C	-14	HIS	-	expression tag	UNP Q13426
C	-13	HIS	-	expression tag	UNP Q13426
C	-12	HIS	-	expression tag	UNP Q13426

Continued on next page...

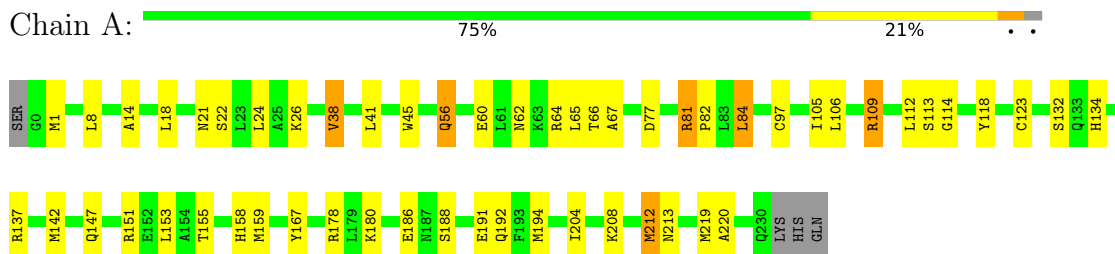
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-11	HIS	-	expression tag	UNP Q13426
C	-10	HIS	-	expression tag	UNP Q13426
C	-9	SER	-	expression tag	UNP Q13426
C	-8	SER	-	expression tag	UNP Q13426
C	-7	GLY	-	expression tag	UNP Q13426
C	-6	LEU	-	expression tag	UNP Q13426
C	-5	VAL	-	expression tag	UNP Q13426
C	-4	PRO	-	expression tag	UNP Q13426
C	-3	ARG	-	expression tag	UNP Q13426
C	-2	GLY	-	expression tag	UNP Q13426
C	-1	SER	-	expression tag	UNP Q13426
C	0	HIS	-	expression tag	UNP Q13426
D	-19	MET	-	expression tag	UNP Q13426
D	-18	GLY	-	expression tag	UNP Q13426
D	-17	SER	-	expression tag	UNP Q13426
D	-16	SER	-	expression tag	UNP Q13426
D	-15	HIS	-	expression tag	UNP Q13426
D	-14	HIS	-	expression tag	UNP Q13426
D	-13	HIS	-	expression tag	UNP Q13426
D	-12	HIS	-	expression tag	UNP Q13426
D	-11	HIS	-	expression tag	UNP Q13426
D	-10	HIS	-	expression tag	UNP Q13426
D	-9	SER	-	expression tag	UNP Q13426
D	-8	SER	-	expression tag	UNP Q13426
D	-7	GLY	-	expression tag	UNP Q13426
D	-6	LEU	-	expression tag	UNP Q13426
D	-5	VAL	-	expression tag	UNP Q13426
D	-4	PRO	-	expression tag	UNP Q13426
D	-3	ARG	-	expression tag	UNP Q13426
D	-2	GLY	-	expression tag	UNP Q13426
D	-1	SER	-	expression tag	UNP Q13426
D	0	HIS	-	expression tag	UNP Q13426

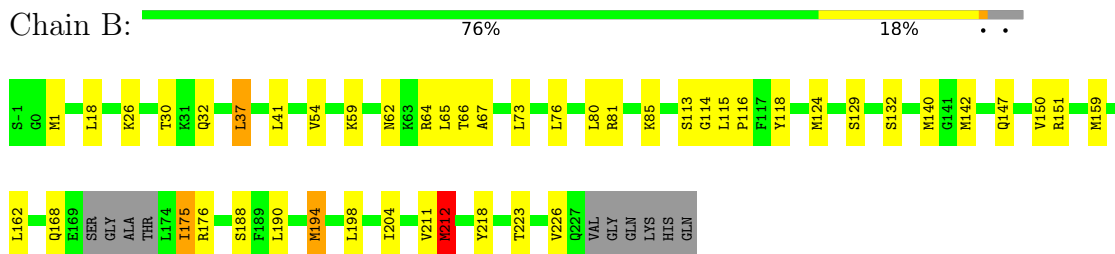
3 Residue-property plots [i](#)

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. 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.

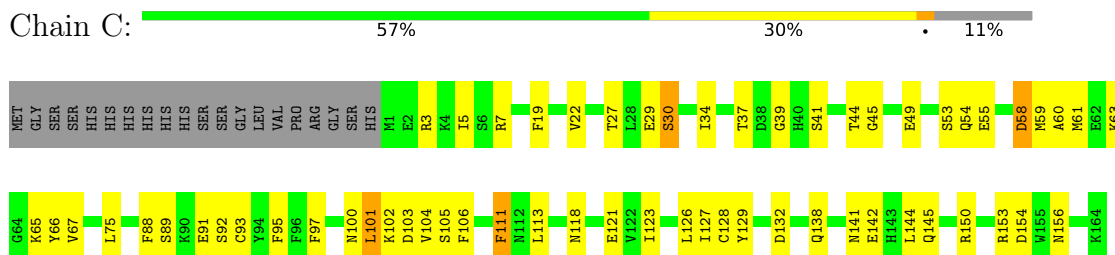
- Molecule 1: Non-homologous end-joining factor 1



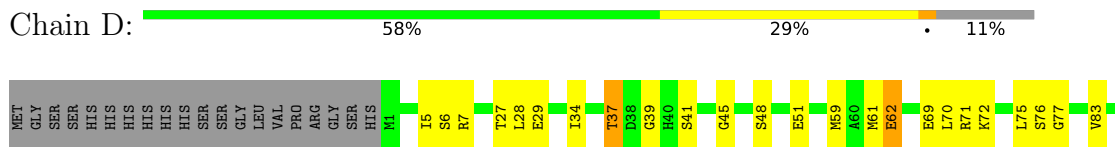
- Molecule 1: Non-homologous end-joining factor 1



- Molecule 2: DNA repair protein XRCC4



- Molecule 2: DNA repair protein XRCC4



F88	S89	K90	E91	S92	C93	Y94	F95	E98	K99	M100	L101	V104	S105	F106	R107	L108	F111	M112	L113	M118	R124	E125	L126	Q138	M141	L144	Q145	K146	E147	M148	E149	M156	D157	W158	Q159	G160	R161	F162	E163	K164
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

4 Data and refinement statistics

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, α , β , γ	236.76Å 236.76Å 103.23Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.81 – 8.49 49.81 – 8.49	Depositor EDS
% Data completeness (in resolution range)	99.3 (49.81-8.49) 99.9 (49.81-8.49)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	0.05	Depositor
$\langle I/\sigma(I) \rangle$ ¹	10.54 (at 8.33Å)	Xtrriage
Refinement program	PHENIX (1.7.2_869)	Depositor
R, R_{free}	0.363 , 0.360 0.376 , 0.419	Depositor DCC
R_{free} test set	149 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å ²)	931.0	Xtrriage
Anisotropy	0.402	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 194.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	0.055 for h,-h-k,-l	Xtrriage
F_o, F_c correlation	0.65	EDS
Total number of atoms	6229	wwPDB-VP
Average B, all atoms (Å ²)	73.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 11.34% 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](#)

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	0.75	6/1854 (0.3%)	0.69	0/2513
1	B	0.76	7/1793 (0.4%)	0.66	1/2434 (0.0%)
2	C	0.39	0/1353	0.55	0/1821
2	D	0.38	0/1353	0.57	0/1821
All	All	0.63	13/6353 (0.2%)	0.63	1/8589 (0.0%)

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	1	MET	CG-SD	6.90	1.99	1.81
1	B	194	MET	CG-SD	6.81	1.98	1.81
1	A	1	MET	CG-SD	6.76	1.98	1.81
1	B	159	MET	CG-SD	6.75	1.98	1.81
1	A	194	MET	CG-SD	6.40	1.97	1.81

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	B	37	LEU	CA-CB-CG	5.54	128.03	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	1818	0	1818	29	69
1	B	1757	0	1741	20	25
2	C	1327	0	1290	39	36
2	D	1327	0	1290	33	58
All	All	6229	0	6139	110	94

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 9.

The worst 5 of 110 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:109:ARG:HH11	1:A:109:ARG:HG3	1.24	1.02
1:A:62:ASN:HD21	1:A:118:TYR:H	1.20	0.87
2:D:27:THR:HG22	2:D:29:GLU:H	1.39	0.87
2:C:37:THR:HG22	2:C:39:GLY:H	1.39	0.86
2:C:89:SER:HB3	2:C:92:SER:HB3	1.56	0.85

The worst 5 of 94 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)
1:A:64:ARG:NH1	2:C:59:MET:C[6_554]	0.71	1.49
1:B:66:THR:CG2	2:D:105:SER:O[1_556]	0.71	1.49
1:B:65:LEU:CD1	2:D:106:PHE:CE2[1_556]	0.79	1.41
1:A:64:ARG:NE	2:C:59:MET:CB[6_554]	0.88	1.32
1:A:186:GLU:CG	2:D:164:LYS:CD[4_655]	0.93	1.27

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	229/235 (97%)	220 (96%)	9 (4%)	0	100	100
1	B	221/235 (94%)	210 (95%)	11 (5%)	0	100	100
2	C	162/184 (88%)	154 (95%)	8 (5%)	0	100	100
2	D	162/184 (88%)	155 (96%)	7 (4%)	0	100	100
All	All	774/838 (92%)	739 (96%)	35 (4%)	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	200/207 (97%)	191 (96%)	9 (4%)	27	52
1	B	191/207 (92%)	179 (94%)	12 (6%)	18	43
2	C	146/163 (90%)	137 (94%)	9 (6%)	18	43
2	D	146/163 (90%)	138 (94%)	8 (6%)	21	47
All	All	683/740 (92%)	645 (94%)	38 (6%)	21	46

5 of 38 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	153	ARG
2	D	141	ASN
2	C	156	ASN
2	D	70	LEU
2	D	156	ASN

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

Mol	Chain	Res	Type
2	D	141	ASN
2	D	137	ASN
2	C	137	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	D	118	ASN
2	C	118	ASN

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 Fit of model and data

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.