



wwPDB/EMDatabank EM Map/Model Validation Summary Report

Oct 10, 2017 – 05:59 PM EDT

PDB ID : 5H1B
EMDB ID: : EMD-9566
Title : Human RAD51 presynaptic complex
Authors : Xu, J.; Zhao, L.; Xu, Y.; Zhao, W.; Sung, P.; Wang, H.W.
Deposited on : unknown
Resolution : 4.40 Å(reported)
Based on PDB ID : 1SZP

This is a wwPDB/EMDatabank EM Map/Model Validation Summary Report
for a publicly released PDB/EMDB entry.

We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
<http://wwpdb.org/validation/2016/EMValidationReportHelp>
with specific help available everywhere you see the  symbol.

MolProbity : 4.02b-467
Mogul : 1.7.2 (RC1), CSD as538be (2017)
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20030345

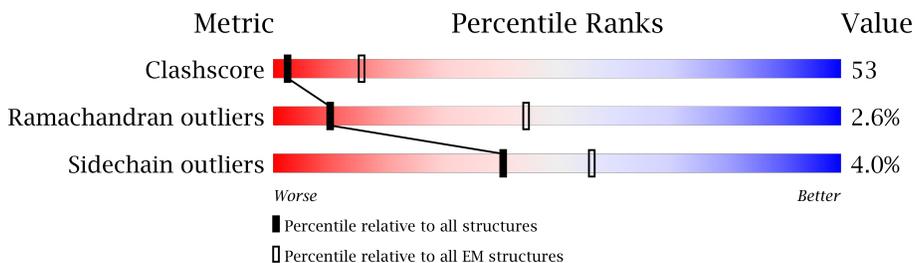
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 4.40 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	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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	339	
1	B	339	
1	C	339	
2	D	9	

2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 7359 atoms, of which 0 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 DNA repair protein RAD51 homolog 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	308	2361	1479	417	452	13	0	0
1	B	308	2361	1479	417	452	13	0	0
1	C	308	2361	1479	417	452	13	0	0

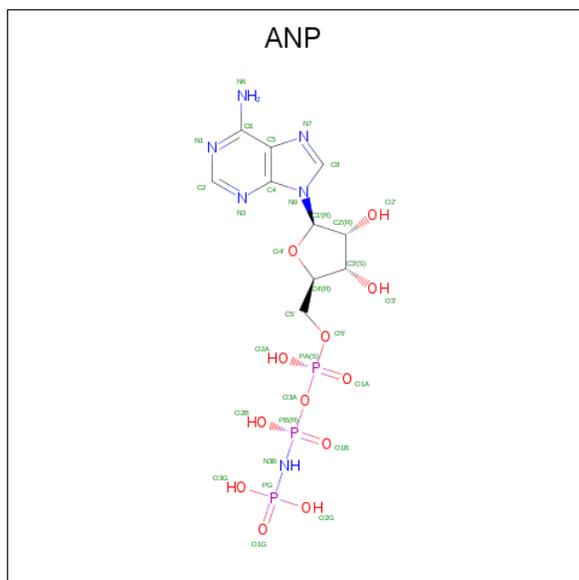
There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	313	GLN	LYS	engineered mutation	UNP Q06609
B	313	GLN	LYS	engineered mutation	UNP Q06609
C	313	GLN	LYS	engineered mutation	UNP Q06609

- Molecule 2 is a DNA chain called DNA (5'-D(P*TP*TP*TP*TP*TP*TP*TP*TP*T)-3').

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	P		
2	D	9	180	90	18	63	9	0	0

- Molecule 3 is PHOSPHOAMINOPHOSPHONIC ACID-ADENYLATE ESTER (three-letter code: ANP) (formula: C₁₀H₁₇N₆O₁₂P₃).



Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	P	
3	A	1	Total	C	N	O	P	0
			31	10	6	12	3	
3	B	1	Total	C	N	O	P	0
			31	10	6	12	3	
3	C	1	Total	C	N	O	P	0
			31	10	6	12	3	

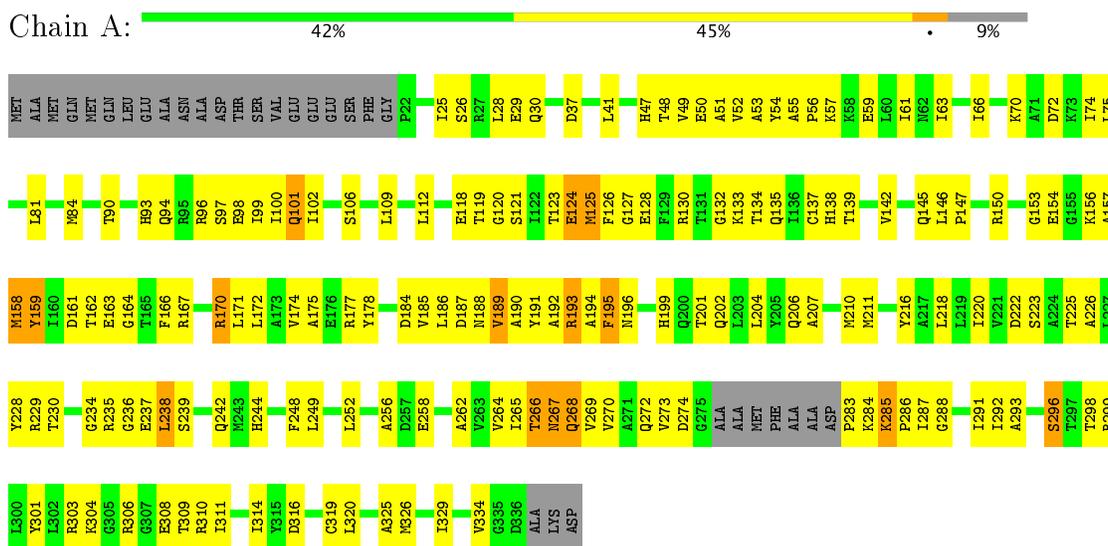
- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
			Total	Mg	
4	B	1	Total	Mg	0
			1	1	
4	A	1	Total	Mg	0
			1	1	
4	C	1	Total	Mg	0
			1	1	

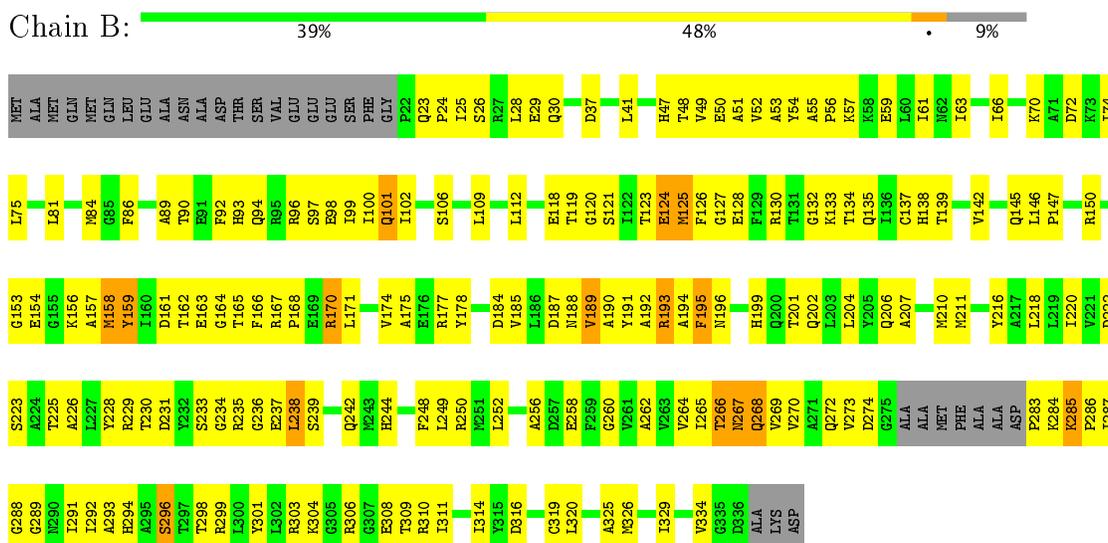
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

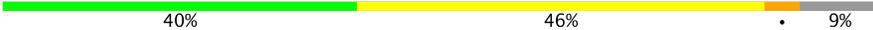
- Molecule 1: DNA repair protein RAD51 homolog 1

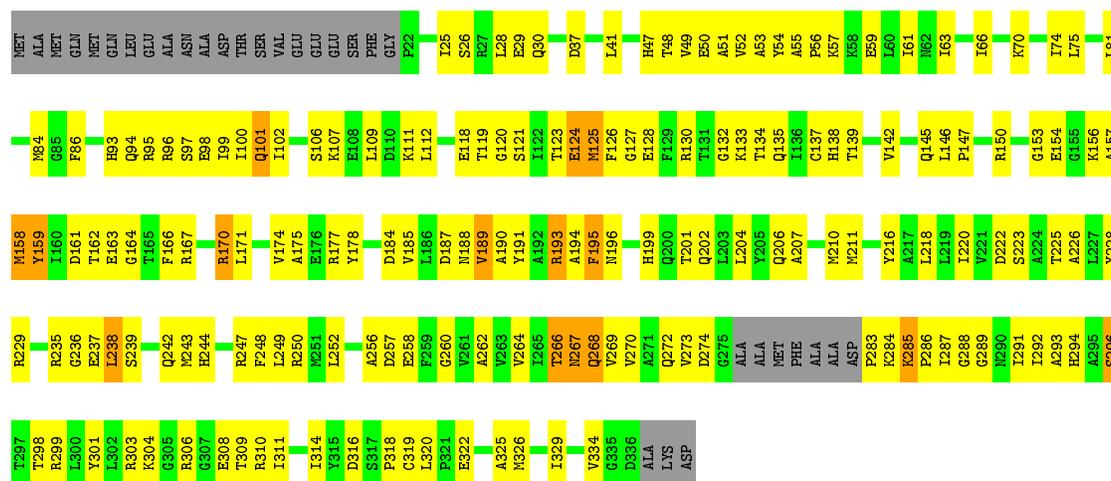


- Molecule 1: DNA repair protein RAD51 homolog 1



- Molecule 1: DNA repair protein RAD51 homolog 1

Chain C:  40% 46% 9%



• Molecule 2: DNA (5'-D(P*TP*TP*TP*TP*TP*TP*TP*TP*T)-3')

Chain D:  33% 67%



4 Experimental information

Property	Value	Source
Reconstruction method	HELICAL	Depositor
Imposed symmetry	HELICAL, twist=56.77°, rise=15.88 Å, axial sym=C1	Depositor
Number of segments used	33838	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING ONLY	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{Å}^2$)	50	Depositor
Minimum defocus (nm)	1500	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	22500	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

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: MG, ANP

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 >2	RMSZ	# Z >2
1	A	0.42	0/2395	0.57	0/3229
1	B	0.42	0/2395	0.57	0/3229
1	C	0.42	0/2395	0.57	0/3229
2	D	0.97	0/197	1.26	0/302
All	All	0.45	0/7382	0.60	0/9989

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	2361	0	2376	260	0
1	B	2361	0	2377	302	0
1	C	2361	0	2377	272	0
2	D	180	0	109	19	0
3	A	31	0	13	7	0
3	B	31	0	13	7	0
3	C	31	0	13	6	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	C	1	0	0	0	0
All	All	7359	0	7278	770	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 53.

The worst 5 of 770 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:B:158:MET:HE2	1:B:216:TYR:CG	1.63	1.33
1:C:158:MET:HE3	1:C:216:TYR:CG	1.62	1.33
1:A:158:MET:HE2	1:A:216:TYR:CG	1.63	1.33
1:C:222:ASP:O	1:C:266:THR:HB	1.28	1.33
1:C:161:ASP:OD1	1:C:164:GLY:N	1.63	1.30

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	304/339 (90%)	277 (91%)	19 (6%)	8 (3%)	6	42
1	B	304/339 (90%)	277 (91%)	19 (6%)	8 (3%)	6	42
1	C	304/339 (90%)	277 (91%)	19 (6%)	8 (3%)	6	42
All	All	912/1017 (90%)	831 (91%)	57 (6%)	24 (3%)	10	42

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	101	GLN
1	A	189	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	238	LEU
1	B	189	VAL
1	B	238	LEU

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	247/269 (92%)	237 (96%)	10 (4%)	36 67
1	B	247/269 (92%)	237 (96%)	10 (4%)	36 67
1	C	247/269 (92%)	237 (96%)	10 (4%)	36 67
All	All	741/807 (92%)	711 (96%)	30 (4%)	40 67

5 of 30 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	158	MET
1	B	193	ARG
1	C	266	THR
1	B	170	ARG
1	B	266	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	199	HIS
1	B	268	GLN
1	C	199	HIS
1	B	188	ASN
1	C	188	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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 6 ligands modelled in this entry, 3 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
3	ANP	A	600	1,4	29,33,33	2.70	6 (20%)	28,52,52	1.11	1 (3%)
3	ANP	B	600	1,4	29,33,33	2.70	6 (20%)	28,52,52	1.11	1 (3%)
3	ANP	C	600	1,4	29,33,33	2.71	6 (20%)	28,52,52	1.11	1 (3%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	ANP	A	600	1,4	-	0/13/38/38	0/3/3/3
3	ANP	B	600	1,4	-	0/13/38/38	0/3/3/3
3	ANP	C	600	1,4	-	0/13/38/38	0/3/3/3

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	600	ANP	PB-O3A	-3.00	1.55	1.59
3	B	600	ANP	PB-O3A	-2.97	1.55	1.59
3	C	600	ANP	PB-O3A	-2.94	1.55	1.59
3	C	600	ANP	PB-O2B	-2.71	1.49	1.56

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	600	ANP	PB-O2B	-2.69	1.49	1.56

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	600	ANP	PA-O3A-PB	-4.01	118.22	132.38
3	B	600	ANP	PA-O3A-PB	-4.01	118.22	132.38
3	C	600	ANP	PA-O3A-PB	-4.01	118.23	132.38

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 20 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	600	ANP	7	0
3	B	600	ANP	7	0
3	C	600	ANP	6	0

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.