WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2024 – 12:56 PM EST

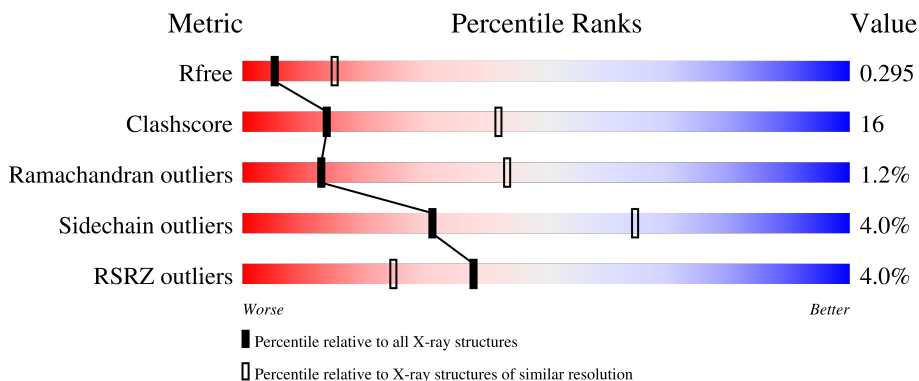|		
PDB ID	:	3K75
Title	:	X-ray crystal structure of reduced XRCC1 bound to DNA pol beta catalytic domain
Authors	:	Cuneo, M.J.; London, R.E.
Deposited on	:	2009-10-12
Resolution	:	2.95 Å(reported)

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 2.95 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	Percentile Ranks	Value
Rfree		0.295
Clashscore		16
Ramachandran outliers		1.2%
Sidechain outliers		4.0%
RSRZ outliers		4.0%

Worse .... *Better*

▐ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3104 (3.00-2.92)
Clashscore	141614	3462 (3.00-2.92)
Ramachandran outliers	138981	3340 (3.00-2.92)
Sidechain outliers	138945	3343 (3.00-2.92)
RSRZ outliers	127900	2986 (3.00-2.92)

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 <=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	B	189	% / 54% 25% • 20%
1	C	189	7% / 50% 24% 5% • 21%
2	D	252	3% / 66% 28% • •
2	E	252	4% / 71% 22% • •

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 6363 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 DNA repair protein XRCC1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	151	Total	C	N	O	S	7	1	0
			1161	722	209	225	5			
1	C	149	Total	C	N	O	S	7	2	0
			1159	718	211	225	5			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	184	HIS	-	expression tag	UNP P18887
B	185	HIS	-	expression tag	UNP P18887
B	186	HIS	-	expression tag	UNP P18887
B	187	HIS	-	expression tag	UNP P18887
B	188	HIS	-	expression tag	UNP P18887
B	189	HIS	-	expression tag	UNP P18887
C	184	HIS	-	expression tag	UNP P18887
C	185	HIS	-	expression tag	UNP P18887
C	186	HIS	-	expression tag	UNP P18887
C	187	HIS	-	expression tag	UNP P18887
C	188	HIS	-	expression tag	UNP P18887
C	189	HIS	-	expression tag	UNP P18887

- Molecule 2 is a protein called DNA polymerase beta.

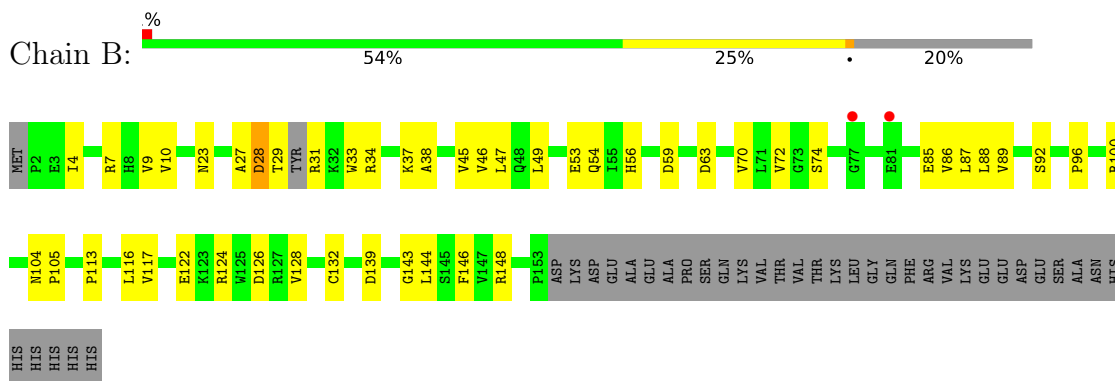| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	248	Total	C	N	O	S	0	4	0
			2044	1282	366	388	8			
2	E	243	Total	C	N	O	S	0	3	0
			1999	1256	354	381	8			

There are 14 discrepancies between the modelled and reference sequences:

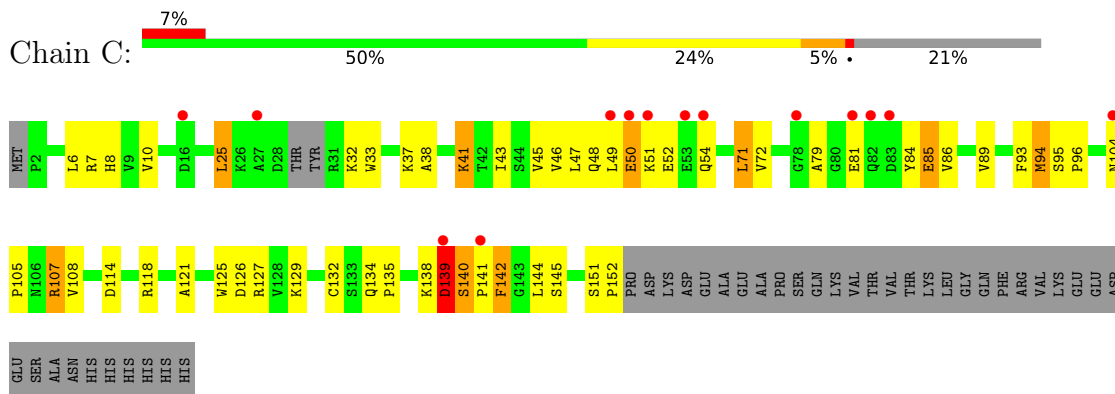| Chain	Residue	Modelled	Actual	Comment	Reference
D	90	MET	-	expression tag	UNP P06766
D	336	HIS	-	expression tag	UNP P06766
D	337	HIS	-	expression tag	UNP P06766
D	338	HIS	-	expression tag	UNP P06766
D	339	HIS	-	expression tag	UNP P06766
D	340	HIS	-	expression tag	UNP P06766
D	341	HIS	-	expression tag	UNP P06766
E	90	MET	-	expression tag	UNP P06766
E	336	HIS	-	expression tag	UNP P06766
E	337	HIS	-	expression tag	UNP P06766
E	338	HIS	-	expression tag	UNP P06766
E	339	HIS	-	expression tag	UNP P06766
E	340	HIS	-	expression tag	UNP P06766
E	341	HIS	-	expression tag	UNP P06766

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 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.
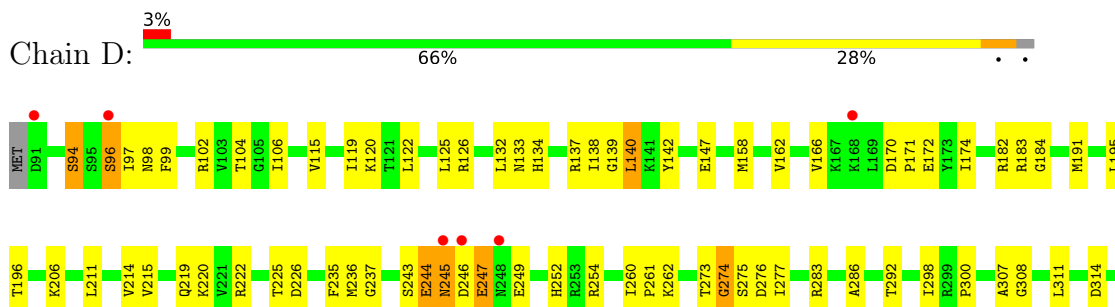
• Molecule 1: DNA repair protein XRCC1

• Molecule 1: DNA repair protein XRCC1

• Molecule 2: DNA polymerase beta

Q317 Y322 W325 R326 Y327 R328 E329 P330 K331 E335 H338 HIS HIS HIS

• Molecule 2: DNA polymerase beta

Chain E:

4%

71%   22%

MET ASP ASP T93 S96 I97 L100 T101 R102 V103 T104 G105 I106 R112 K113 L114 V115 D116 I119 K120 T121 L122 L125 K131 H134 H135 Q136 R137 I138 G139 L140 K141 E147 V162 L163 N164 E165 V166 K167 K168 L169 D170 P171 E172 Y173 R183 V215 Q219

F223 I224 T225 D226 V238 C239 S243 E244 N245 D246 E247 N248 T273 G274 S275 D276 L287 E288 N294 E295 Y296 T297 I298 R299 P300 L301 G308 L311 Q317 Q324 W325 R326 Y327 R328 E329 P330 S334 E335 HIS HIS HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	140.77Å 44.05Å 152.77Å 90.00° 107.21° 90.00°	Depositor
Resolution (Å)	24.47 – 2.95 24.47 – 2.95	Depositor EDS
% Data completeness (in resolution range)	100.0 (24.47-2.95) 99.0 (24.47-2.95)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	4.35 (at 2.94Å)	Xtriage
Refinement program	REFMAC 5.5.0066	Depositor
R, R_{free}	0.242 , 0.295 0.241 , 0.295	Depositor DCC
R_{free} test set	957 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	43.8	Xtriage
Anisotropy	0.311	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 29.6	EDS
L-test for twinning[2]	$<|L|> = 0.43, <L^2> = 0.26$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	6363	wwPDB-VP
Average B, all atoms (Å^2)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.44% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $<|L|>$, $<L^2>$ 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	B	0.32	0/1183	0.52	0/1596
1	C	0.27	0/1180	0.46	0/1590
2	D	0.37	0/2086	0.50	0/2812
2	E	0.34	0/2038	0.45	0/2745
All	All	0.34	0/6487	0.48	0/8743

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
2	D	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	D	244	GLU	Peptide

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	B	1161	0	1139	34	1

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	1159	0	1135	48	0
2	D	2044	0	2001	69	0
2	E	1999	0	1980	58	0
All	All	6363	0	6255	199	1

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 16.

The worst 5 of 199 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:C:89:VAL:CG2	2:E:308:GLY:HA2	1.74	1.18
2:D:245:ASN:O	2:D:246:ASP:CG	1.87	1.11
2:D:245:ASN:O	2:D:246:ASP:OD1	1.67	1.10
2:E:247:GLU:OE1	2:E:248:ASN:N	1.90	1.04
2:E:170:ASP:OD1	2:E:171:PRO:HD2	1.58	1.02

All (1) 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:B:7:ARG:NH2	1:B:139:ASP:OD1[4_445]	2.07	0.13

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	B	148/189 (78%)	130 (88%)	17 (12%)	1 (1%)	22	56
1	C	147/189 (78%)	127 (86%)	15 (10%)	5 (3%)	3	17
2	D	250/252 (99%)	237 (95%)	12 (5%)	1 (0%)	34	69

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	E	244/252 (97%)	223 (91%)	19 (8%)	2 (1%)	19	53
All	All	789/882 (90%)	717 (91%)	63 (8%)	9 (1%)	13	46

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	28	ASP
1	C	139	ASP
1	C	142	PHE
2	D	274	GLY
2	E	274	GLY

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	B	129/164 (79%)	126 (98%)	3 (2%)	50	78
1	C	129/164 (79%)	122 (95%)	7 (5%)	22	54
2	D	226/230 (98%)	214 (95%)	12 (5%)	22	55
2	E	223/230 (97%)	216 (97%)	7 (3%)	40	71
All	All	707/788 (90%)	678 (96%)	29 (4%)	31	64

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	245	ASN
2	E	328	ARG
2	D	317[A]	GLN
2	E	317	GLN
2	D	314	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
2	E	217	GLN
2	E	317	GLN
2	D	252	HIS
2	E	128	ASN
2	E	157	GLN

5.3.3 RNA ⓘ

There are no RNA molecules 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 monosaccharides 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	B	151/189 (79%)	0.00	2 (1%) 77	61	17, 45, 87, 100	0
1	C	149/189 (78%)	0.66	14 (9%) 8	5	36, 58, 96, 110	0
2	D	248/252 (98%)	-0.02	7 (2%) 53	36	20, 41, 92, 128	0
2	E	243/252 (96%)	0.01	9 (3%) 41	27	24, 47, 95, 134	0
All	All	791/882 (89%)	0.12	32 (4%) 38	25	17, 48, 95, 134	0

The worst 5 of 32 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	50	GLU	6.0
2	D	246	ASP	5.3
2	E	93	THR	5.0
1	C	82	GLN	4.0
1	C	81	GLU	3.7

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 monosaccharides 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.