



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 06:06 AM GMT

PDB ID : 2VYF  
Title : CRYSTAL STRUCTURE OF THE DNAC  
Authors : Lo, Y.H.; Tsai, K.L.; Sun, Y.J.; Hsiao, C.D.  
Deposited on : 2008-07-23  
Resolution : 3.60 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](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>  
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:

MolProbity : 4.02b-467  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

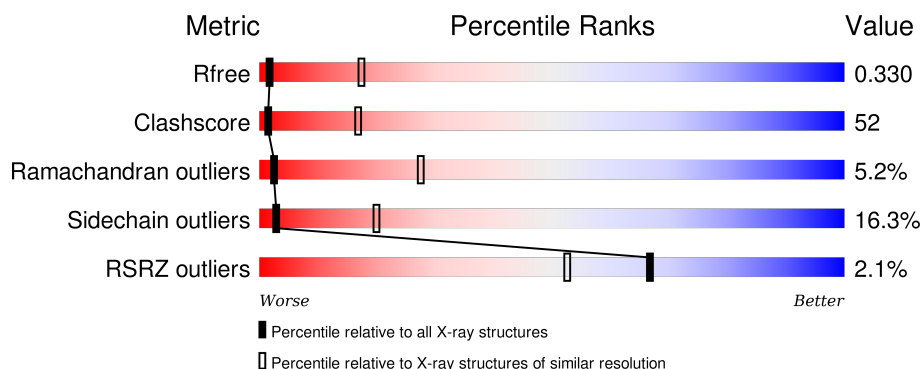
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.60 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))
$R_{free}$	91344	1408 (3.80-3.40)
Clashscore	102246	1010 (3.74-3.46)
Ramachandran outliers	100387	1007 (3.76-3.44)
Sidechain outliers	100360	1007 (3.76-3.44)
RSRZ outliers	91569	1003 (3.78-3.42)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	A	454	<div> <div>33%</div> <div>42%</div> <div>11%</div> <div>•</div> <div>12%</div> </div>
1	B	454	<div> <div>3%</div> <div>34%</div> <div>37%</div> <div>10%</div> <div>•</div> <div>17%</div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5946 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 REPLICATIVE DNA HELICASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	398	Total	C	N	O	S	0	0	4
			3062	1912	539	599	12			
1	B	377	Total	C	N	O	S	0	0	6
			2880	1798	506	564	12			

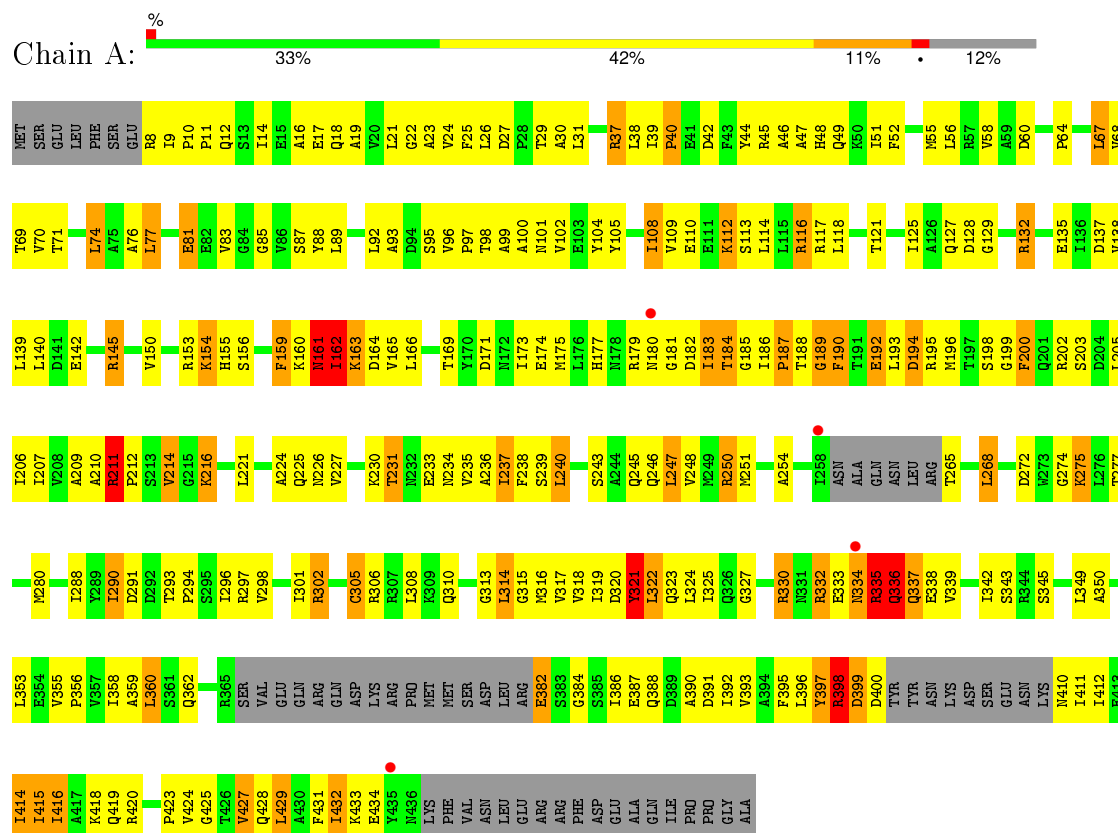
- Molecule 2 is GOLD ION (three-letter code: AU) (formula: Au).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	2	Total	Au	0	0
			2	2		
2	A	2	Total	Au	0	0
			2	2		

### 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 errors 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: REPLICATIVE DNA HELICASE



VAL	GLN	LEU	ALA	PHE	ILE	LYS	GLU	TYR	ASN	LYS	PHE	VAL	ASN	GLU	LEU	ARG	ARG	PHE	ASP	GLU	ALA	GLN	PRO	GLY	ALA																																	
R1365	SER	VAL	GLU	GLN	ARG	GLN	ASP	LYS	ARG	PRO	MET	MET	SER	ASP	LEU	ARG	E1382	I1386	E1387	Q1388	D1389	I1390	D1391	I1392	V1393	A1394	F1395	L1396	V1397	ARG	ASP	ASP	TYR	TYR	ASN	LYS	ASP	SER	GLY	ASN	LYS	ASN	I1411	I1412	E1413	I1414	I1415	I1416	A1417	K1418	Q1419	R1420	N1421	G1422	P1423	V1424	G1425	T1426
S1295	I1296	R1297	V1298	S1299	D1300	I1301	A1302	A1303	K1304	C1305	R1306	R1307	L1308	K1309	G1313	L1314	G1315	R1316	V1317	V1318	I1319	D1320	Y1321	L1322	GLN	LEU	ILE	GLN	GLY	SER	GLY	ARG	ASN	R1332	E1333	M1334	E1341	I1342	S1343	R1344	K1347	R1351	E1352	L1353	E1354	V1355	P1356	V1357	I1358	A1359	L1360	S1361	Q1362	L1363	S1364			
T1217	L1221	N1222	I1223	A1224	Q1225	N1226	V1227	A1228	T1229	K1230	T1231	M1232	E1233	M1234	V1235	A1236	I1237	F1238	S1239	L1240	E1241	M1242	S1243	A1244	Q1245	L1247	V1248	M1251	L1252	C1253	N1257	I1258	M1259	A1260	Q1261	M1262	L1263	R1264	T1265	G1266	R1267	L1268	T1269	P1270	E1271	A1286	G1287	I1288	Y1289	I1290	D1291	D1292	T1293	P1294				

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 63	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	176.97Å 176.97Å 108.82Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	24.87 – 3.60 24.87 – 3.59	Depositor EDS
% Data completeness (in resolution range)	(Not available) (24.87-3.60) 91.9 (24.87-3.59)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.54 (at 3.64Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, $R_{free}$	0.293 , 0.324 0.305 , 0.330	Depositor DCC
$R_{free}$ test set	1770 reflections (9.25%)	DCC
Wilson B-factor (Å <sup>2</sup> )	142.8	Xtriage
Anisotropy	0.540	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 93.8	EDS
Estimated twinning fraction	0.038 for h,-h-k,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	1 of 22531 reflections (0.004%)	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	5946	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	167.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.85% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 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: AU

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	0.40	0/3096	0.77	3/4184 (0.1%)
1	B	0.43	0/2908	0.95	18/3930 (0.5%)
All	All	0.42	0/6004	0.86	21/8114 (0.3%)

There are no bond length outliers.

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	1297	ARG	N-CA-C	9.29	136.08	111.00
1	B	1260	ALA	N-CA-C	9.18	135.79	111.00
1	A	162	ILE	CG1-CB-CG2	-8.29	93.16	111.40
1	B	1411	ILE	N-CA-C	8.28	133.36	111.00
1	B	1419	GLN	N-CA-C	7.66	131.68	111.00

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	3062	0	3098	343	2
1	B	2880	0	2922	299	2
2	A	2	0	0	0	0
2	B	2	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	5946	0	6020	621	2

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 52.

The worst 5 of 621 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:211:ARG:HD3	1:A:212:PRO:CD	1.49	1.41
1:A:297:ARG:HH22	1:B:1154:LYS:N	1.15	1.39
1:A:145:ARG:NH2	1:A:145:ARG:HB3	1.35	1.37
1:B:1145:ARG:NH2	1:B:1145:ARG:HB2	1.41	1.34
1:A:297:ARG:NH2	1:B:1154:LYS:H	1.24	1.33

All (2) 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:154:LYS:O	1:B:1307:ARG:NH1[2_655]	1.77	0.43
1:A:87:SER:OG	1:B:1163:LYS:NZ[4_665]	1.89	0.31

## 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	390/454 (86%)	319 (82%)	54 (14%)	17 (4%)	3	32
1	B	365/454 (80%)	305 (84%)	38 (10%)	22 (6%)	2	24
All	All	755/908 (83%)	624 (83%)	92 (12%)	39 (5%)	2	27

5 of 39 Ramachandran outliers are listed below:



Mol	Chain	Res	Type
1	A	40	PRO
1	A	161	ASN
1	A	190	PHE
1	A	192	GLU
1	A	211	ARG

### 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	332/388 (86%)	274 (82%)	58 (18%)	2	16
1	B	313/388 (81%)	266 (85%)	47 (15%)	3	25
All	All	645/776 (83%)	540 (84%)	105 (16%)	3	20

5 of 105 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	360	LEU
1	B	1031	LEU
1	B	1341	GLU
1	A	362	GLN
1	A	414	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	1101	ASN
1	B	1201	GLN
1	B	1419	GLN
1	B	1127	GLN
1	B	1152	GLN

### 5.3.3 RNA ⓘ

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 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	398/454 (87%)	-0.15	4 (1%) 84 73	112, 165, 187, 187	0
1	B	377/454 (83%)	-0.09	12 (3%) 51 37	96, 175, 187, 187	0
All	All	775/908 (85%)	-0.12	16 (2%) 67 52	96, 171, 187, 187	0

The worst 5 of 16 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	1265	THR	5.3
1	B	1266	GLY	3.0
1	B	1259	ASN	2.9
1	B	1322	LEU	2.9
1	A	435	TYR	2.8

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	AU	B	2426	1/1	0.97	0.11	-2.32	186,186,186,186	0
2	AU	A	1436	1/1	0.99	0.04	-3.75	186,186,186,186	0
2	AU	A	1437	1/1	0.97	0.09	-	186,186,186,186	0
2	AU	B	2427	1/1	0.98	0.05	-	186,186,186,186	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.