



# wwPDB X-ray Structure Validation Summary Report i

Feb 28, 2014 – 10:09 AM GMT

PDB ID : 3U44  
Title : Crystal structure of AddAB-DNA complex  
Authors : Saikrishnan, K.; Krajewski, W.; Wigley, D.  
Deposited on : 2011-10-07  
Resolution : 3.20 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

---

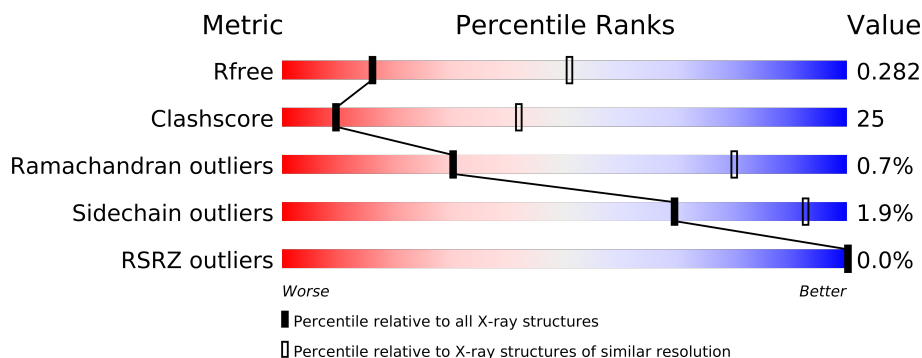
The following versions of software and data (see [references](#)) were used in the production of this report:

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 3.20 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	1824 (3.30-3.10)
Clashscore	79885	1078 (3.26-3.14)
Ramachandran outliers	78287	1059 (3.26-3.14)
Sidechain outliers	78261	1058 (3.26-3.14)
RSRZ outliers	66119	1825 (3.30-3.10)

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. 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	1232	
2	B	1166	
3	X	48	

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 19156 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 ATP-dependent helicase/nuclease subunit A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1122	Total	C	N	O	S	0	0	0
			9117	5821	1548	1721	27			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	780	GLY	ALA	CONFLICT	UNP P23478
A	1172	ALA	ASP	ENGINEERED MUTATION	UNP P23478

- Molecule 2 is a protein called ATP-dependent helicase/deoxyribonucleasesubunit B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1141	Total	C	N	O	S	0	0	0
			9261	5885	1583	1749	44			

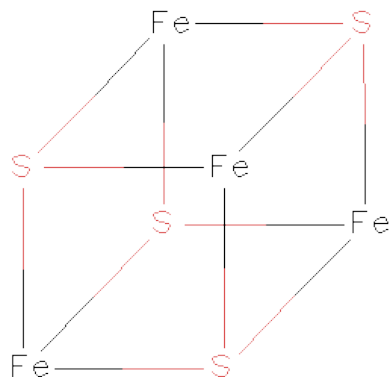
There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	843	ASP	GLU	CONFLICT	UNP P23477
B	844	GLU	GLN	CONFLICT	UNP P23477
B	961	ALA	ASP	ENGINEERED MUTATION	UNP P23477

- Molecule 3 is a DNA chain called DNA (36-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	X	36	Total	C	N	O	P	0	0	0
			730	352	128	216	34			

- Molecule 4 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe<sub>4</sub>S<sub>4</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total	Fe	S	0	0
			8	4	4		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	18	Total	O	0	0
			18	18		
5	B	22	Total	O	0	0
			22	22		

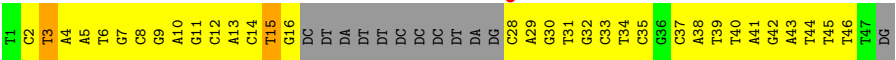






• Molecule 3: DNA (36-MER)

Chain X: A horizontal bar chart representing the DNA sequence. The bar is divided into segments of different colors: red, green, yellow, orange, and grey. The segments are labeled with nucleotide codes: T, C, A, G, and D.



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	100.23Å 138.96Å 102.97Å 90.00° 105.33° 90.00°	Depositor
Resolution (Å)	29.88 – 3.20 56.93 – 3.20	Depositor EDS
% Data completeness (in resolution range)	99.0 (29.88-3.20) 99.1 (56.93-3.20)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.99 (at 3.19Å)	Xtriage
Refinement program	PHENIX (phenix.refine: dev_750)	Depositor
R, $R_{free}$	0.233 , 0.284 0.228 , 0.282	Depositor DCC
$R_{free}$ test set	2203 reflections (4.95%)	DCC
Wilson B-factor (Å <sup>2</sup> )	72.0	Xtriage
Anisotropy	0.354	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 35.0	EDS
Estimated twinning fraction	0.028 for l,-k,h	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 44538 reflections	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	19156	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	76.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.39% of the height of the origin peak. No significant pseudotranslation is detected.*

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



## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: SF4

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	0.22	0/9294	0.39	0/12535
2	B	0.23	0/9448	0.40	0/12732
3	X	0.47	0/816	1.17	4/1256 (0.3%)
All	All	0.24	0/19558	0.46	4/26523 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	X	15	DT	O4'-C1'-N1	7.54	113.28	108.00
3	X	3	DT	C5-C4-O4	-6.52	120.34	124.90
3	X	15	DT	C3'-C2'-C1'	-6.10	95.18	102.50
3	X	15	DT	O4'-C1'-C2'	-5.18	101.76	105.90

There are no chirality outliers.

There are no planarity outliers.

### 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	9117	0	9063	448	0
2	B	9261	0	9222	474	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	X	730	0	412	63	0
4	B	8	0	0	1	0
5	A	18	0	0	0	0
5	B	22	0	0	0	0
All	All	19156	0	18697	928	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 25.

The worst 5 of 928 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:845:THR:HG22	1:A:847:PRO:HD2	1.30	1.12
1:A:172:ARG:H	2:B:888:LYS:HD3	1.20	1.04
2:B:864:VAL:HG11	2:B:890:GLN:HG2	1.38	1.01
2:B:972:ALA:HA	2:B:1152:ILE:HD12	1.43	1.01
1:A:243:LEU:HB3	1:A:301:LEU:HD21	1.45	0.99

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	1096/1232 (89%)	1024 (93%)	63 (6%)	9 (1%)	27 77
2	B	1137/1166 (98%)	1053 (93%)	77 (7%)	7 (1%)	33 83
All	All	2233/2398 (93%)	2077 (93%)	140 (6%)	16 (1%)	30 80

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	52	GLU
1	A	284	VAL
1	A	172	ARG
1	A	567	PHE

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	887	HIS

### 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	977/1062 (92%)	963 (99%)	14 (1%)	78	95
2	B	1007/1029 (98%)	983 (98%)	24 (2%)	61	91
All	All	1984/2091 (95%)	1946 (98%)	38 (2%)	69	93

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

Mol	Chain	Res	Type
2	B	78	HIS
2	B	353	GLU
2	B	1012	HIS
2	B	269	LEU
2	B	532	ILE

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

Mol	Chain	Res	Type
2	B	230	HIS
2	B	1045	GLN
2	B	809	HIS
1	A	836	HIS
2	B	587	GLN

### 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 ⓘ

1 ligand is modelled in this entry.

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	$\# Z  > 2$
4	SF4	B	1167	2	12,12,12	5.94	12 (100%)	0,24,24	0.00	-

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
4	SF4	B	1167	2	-	0/0/48/48	0/0/5/5

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	1167	SF4	S4-FE3	-6.47	2.28	2.33
4	B	1167	SF4	S3-FE1	-6.28	2.29	2.33
4	B	1167	SF4	S2-FE3	-6.22	2.29	2.33
4	B	1167	SF4	S4-FE2	-6.16	2.29	2.33
4	B	1167	SF4	S1-FE4	-6.16	2.29	2.33

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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, 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	1122/1232 (91%)	-0.03	0	100	100	40, 70, 114, 133	0
2	B	1141/1166 (97%)	-0.06	0	100	100	40, 72, 108, 148	0
3	X	36/48 (75%)	0.05	1 (2%)	50	11	80, 119, 147, 158	0
All	All	2299/2446 (93%)	-0.04	1 (0%)	100	100	40, 71, 113, 158	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	X	28	DC	2.9

### 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 ⓘ

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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
4	SF4	B	1167	8/8	0.13	-1.40	69,77,84,112	8

## 6.5 Other polymers ⓘ

There are no such residues in this entry.