



# Full wwPDB X-ray Structure Validation Report

Feb 26, 2014 – 08:09 PM GMT

PDB ID : 1BC8  
Title : STRUCTURES OF SAP-1 BOUND TO DNA SEQUENCES FROM THE  
E74 AND C-FOS PROMOTERS PROVIDE INSIGHTS INTO HOW ETS  
PROTEINS DISCRIMINATE BETWEEN RELATED DNA TARGETS  
Authors : Mo, Y.; Vaessen, B.; Johnston, K.; Marmorstein, R.  
Deposited on : 1998-05-05  
Resolution : 1.93 Å(reported)

This is a full wwPDB validation 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>

---

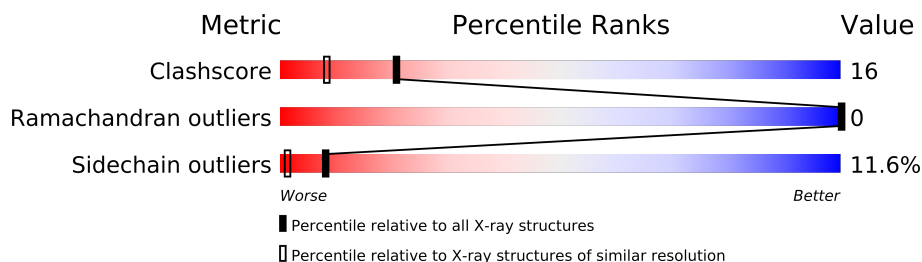
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 21963  
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 1.93 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))
Clashscore	79885	2281 (1.96-1.92)
Ramachandran outliers	78287	2255 (1.96-1.92)
Sidechain outliers	78261	2255 (1.96-1.92)

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 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	10	
2	B	10	
3	C	93	

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 1356 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 DNA chain called DNA (5'-D(\*TP\*AP\*CP\*CP\*GP\*GP\*AP\*AP\*GP\*T)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	10	Total	C	N	O	P	0	0	0
			204	98	40	57	9			

- Molecule 2 is a DNA chain called DNA (5'-D(\*AP\*AP\*CP\*TP\*TP\*CP\*CP\*GP\*GP\*T)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	10	Total	C	N	O	P	0	0	0
			200	97	35	59	9			

- Molecule 3 is a protein called PROTEIN (SAP-1 ETS DOMAIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	93	Total	C	N	O	S	0	0	0
			790	516	136	133	5			

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Zn	0	0
			1	1		
4	C	1	Total	Zn	0	0
			1	1		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	40	Total	O	0	0
			40	40		
5	B	38	Total	O	0	0
			38	38		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	C	82	Total	O	0	0
			82	82		

### 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.

Note EDS was not executed.

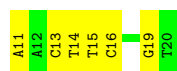
- Molecule 1: DNA (5'-D(\*TP\*AP\*CP\*CP\*GP\*GP\*AP\*AP\*GP\*T)-3')

Chain A: 



- Molecule 2: DNA (5'-D(\*AP\*AP\*CP\*TP\*TP\*CP\*CP\*GP\*GP\*T)-3')

Chain B: 



- Molecule 3: PROTEIN (SAP-1 ETS DOMAIN)

Chain C: 



## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	31.59Å 55.67Å 44.23Å 90.00° 107.16° 90.00°	Depositor
Resolution (Å)	28.90 – 1.93	Depositor
% Data completeness (in resolution range)	96.4 (28.90-1.93)	Depositor
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	X-PLOR 3.851	Depositor
R, $R_{free}$	0.220 , 0.277	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	1356	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	16.0	wwPDB-VP

## 5 Model quality

### 5.1 Standard geometry

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

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	0.46	0/229	0.73	0/352
2	B	0.55	0/223	0.86	0/342
3	C	0.42	0/809	0.59	0/1087
All	All	0.45	0/1261	0.68	0/1781

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	B	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	B	19	DG	Sidechain

### 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	204	0	114	3	0
2	B	200	0	115	8	0
3	C	790	0	815	26	0
4	B	1	0	0	0	0
4	C	1	0	0	0	0
5	A	40	0	0	1	0
5	B	38	0	0	3	0
5	C	82	0	0	1	0
All	All	1356	0	1044	35	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 16.

All (35) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:B:11:DA:HO5'	2:B:11:DA:H8	0.90	0.88
2:B:11:DA:O5'	2:B:11:DA:H8	1.63	0.81
3:C:6:THR:HG22	3:C:9:GLN:H	1.47	0.80
2:B:15:DT:H2''	2:B:16:DC:H5''	1.67	0.76
3:C:15:LEU:HD13	3:C:24:ILE:HG12	1.78	0.65
3:C:5:ILE:HD13	3:C:13:GLN:HE22	1.63	0.64
3:C:12:LEU:HD13	3:C:84:PHE:CZ	2.36	0.60
2:B:11:DA:N6	5:B:154:HOH:O	2.35	0.58
5:B:63:HOH:O	3:C:49:LYS:HE2	2.02	0.58
2:B:15:DT:H2'	3:C:58:LYS:HD3	1.86	0.56
3:C:17:LYS:HB3	3:C:19:GLN:OE1	2.07	0.55
2:B:13:DC:H2'	2:B:14:DT:H72	1.92	0.52
3:C:15:LEU:CD1	3:C:24:ILE:HG12	2.40	0.51
3:C:34:LYS:HZ3	3:C:80:PHE:HB2	1.76	0.50
3:C:34:LYS:HE2	5:C:97:HOH:O	2.12	0.49
3:C:6:THR:HG22	3:C:9:GLN:N	2.22	0.48
3:C:17:LYS:HB2	3:C:20:ASN:ND2	2.29	0.48
1:A:8:DA:N7	5:A:19:HOH:O	2.35	0.48
1:A:4:DC:H2''	1:A:5:DG:C8	2.48	0.48
2:B:11:DA:O5'	2:B:11:DA:C8	2.46	0.48
3:C:15:LEU:HG	3:C:26:TRP:CE2	2.49	0.48
3:C:45:TRP:CZ2	3:C:49:LYS:HE3	2.49	0.48
2:B:13:DC:H2'	2:B:14:DT:C7	2.44	0.47
3:C:23:MET:HB2	3:C:37:GLN:HB2	1.97	0.46
3:C:6:THR:HG21	3:C:8:TRP:NE1	2.30	0.46
3:C:6:THR:HG21	3:C:8:TRP:CE2	2.51	0.45
5:B:68:HOH:O	3:C:49:LYS:HB2	2.14	0.45
3:C:6:THR:CG2	3:C:9:GLN:H	2.24	0.45

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
3:C:5:ILE:HG13	3:C:6:THR:N	2.32	0.44
1:A:2:DA:H4'	3:C:79:LYS:HG2	1.99	0.44
3:C:6:THR:HG21	3:C:8:TRP:CD1	2.55	0.42
3:C:28:SER:HB3	3:C:32:GLN:HB3	2.01	0.41
3:C:92:ASN:ND2	3:C:92:ASN:N	2.69	0.41
3:C:17:LYS:HB3	3:C:19:GLN:CD	2.41	0.41
3:C:6:THR:CG2	3:C:9:GLN:HG3	2.51	0.40

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	
3	C	91/93 (98%)	87 (96%)	4 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 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	
3	C	86/86 (100%)	76 (88%)	10 (12%)	8	1

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	6	THR
3	C	12	LEU
3	C	13	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
3	C	15	LEU
3	C	19	GLN
3	C	21	LYS
3	C	50	ASN
3	C	73	LYS
3	C	86	SER
3	C	92	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
3	C	13	GLN
3	C	37	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 ⓘ

Of 2 ligands modelled in this entry, 2 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.

## 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 ⓘ

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

### 6.3 Carbohydrates ⓘ

EDS was not executed - this section will therefore be empty.

### 6.4 Ligands ⓘ

EDS was not executed - this section will therefore be empty.

### 6.5 Other polymers ⓘ

EDS was not executed - this section will therefore be empty.