



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 12:42 am BST

PDB ID : 1BDT  
Title : WILD TYPE GENE-REGULATING PROTEIN ARC/DNA COMPLEX  
Authors : Schilbach, J.F.; Karzai, A.W.; Raumann, B.E.; Sauer, R.T.  
Deposited on : 1998-05-11  
Resolution : 2.50 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

<https://www.wwpdb.org/validation/2017/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  
Xtrriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

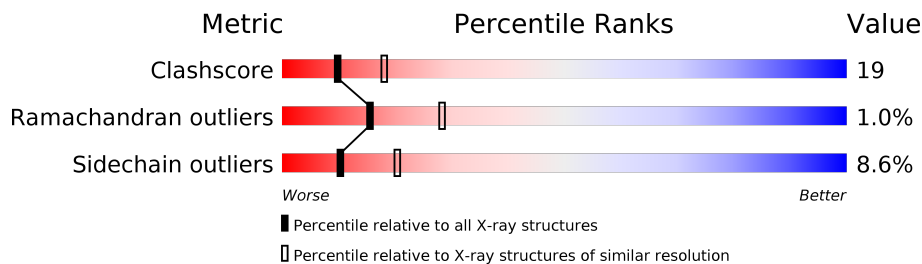
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

## X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)

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 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  $\leq 5\%$

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	E	22	41% 59%
2	F	22	27% 73%
3	A	53	57% 36% 6% .
3	B	53	70% 25% 6%
3	C	53	58% 30% 6% 6%
3	D	53	68% 23% . 6%

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 2676 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 DNA chain called DNA (5'-D(\*TP\*AP\*TP\*AP\*GP\*TP\*AP\*GP\*AP\*GP\*TP\*GP\*CP\*TP\*TP\*CP\*TP\*AP\*TP\*CP\*AP\*T)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	E	22	448	217	77	133	21	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	F	22	448	216	84	127	21	0	0	0

- Molecule 3 is a protein called PROTEIN (GENE-REGULATING PROTEIN ARC).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	A	52	429	268	81	76	4	0	0	0
3	B	53	434	271	82	77	4	0	0	0
3	C	50	417	260	79	74	4	0	0	0
3	D	50	417	260	79	74	4	0	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	E	24	Total 24 O 24	0	0
4	F	18	Total 18 O 18	0	0
4	A	12	Total 12 O 12	0	0

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
4	B	11	Total 11	O 11	0	0
4	C	7	Total 7	O 7	0	0
4	D	11	Total 11	O 11	0	0

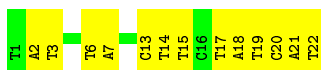
### 3 Residue-property plots [i](#)

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.

Note EDS was not executed.

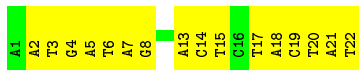
- Molecule 1: DNA (5'-D(\*TP\*AP\*TP\*AP\*GP\*TP\*AP\*GP\*AP\*GP\*TP\*GP\*CP\*TP\*TP\*CP\*TP\*AP\*TP\*CP\*AP\*T)-3')

Chain E: 



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

Chain F: 



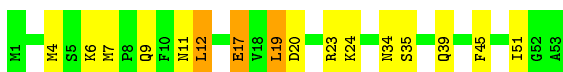
- Molecule 3: PROTEIN (GENE-REGULATING PROTEIN ARC)

Chain A: 



- Molecule 3: PROTEIN (GENE-REGULATING PROTEIN ARC)

Chain B: 



- Molecule 3: PROTEIN (GENE-REGULATING PROTEIN ARC)

Chain C: 



- Molecule 3: PROTEIN (GENE-REGULATING PROTEIN ARC)

Chain D:  68% 23% • 6%



## 4 Data and refinement statistics

Xtrriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	62.13Å 56.41Å 52.33Å 90.00° 104.14° 90.00°	Depositor
Resolution (Å)	20.00 – 2.50	Depositor
% Data completeness (in resolution range)	98.3 (20.00-2.50)	Depositor
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	X-PLOR 3.1	Depositor
R, $R_{free}$	0.236 , 0.282	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	2676	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	51.0	wwPDB-VP

## 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	E	0.52	0/501	0.76	0/772
2	F	0.45	0/503	0.75	0/774
3	A	0.39	0/435	0.60	0/577
3	B	0.47	0/440	0.66	0/584
3	C	0.39	0/423	0.61	0/561
3	D	0.43	0/423	0.60	0/561
All	All	0.44	0/2725	0.68	0/3829

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	E	448	0	253	21	0
2	F	448	0	250	26	0
3	A	429	0	443	17	0
3	B	434	0	448	13	0
3	C	417	0	429	27	0
3	D	417	0	429	20	0
4	A	12	0	0	2	0
4	B	11	0	0	0	0
4	C	7	0	0	1	0
4	D	11	0	0	0	0

*Continued on next page...*



Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	E	24	0	0	1	0
4	F	18	0	0	0	0
All	All	2676	0	2252	93	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 19.

The worst 5 of 93 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)
2:F:5:DA:H2''	2:F:6:DT:C5'	1.81	1.10
2:F:5:DA:H2''	2:F:6:DT:H5'	1.11	1.09
1:E:13:DC:H2''	1:E:14:DT:H5'	1.28	1.07
2:F:13:DA:H2''	2:F:14:DC:H5'	1.42	1.02
1:E:3:DT:H2'	3:A:4:MET:HE1	1.50	0.93

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 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	A	50/53 (94%)	48 (96%)	1 (2%)	1 (2%)	7	12
3	B	51/53 (96%)	51 (100%)	0	0	100	100
3	C	48/53 (91%)	44 (92%)	3 (6%)	1 (2%)	7	11
3	D	48/53 (91%)	47 (98%)	1 (2%)	0	100	100
All	All	197/212 (93%)	190 (96%)	5 (2%)	2 (1%)	15	28

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	5	SER
3	C	5	SER

### 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
3	A	47/47 (100%)	41 (87%)	6 (13%)	4 8
3	B	47/47 (100%)	43 (92%)	4 (8%)	10 21
3	C	46/47 (98%)	43 (94%)	3 (6%)	17 33
3	D	46/47 (98%)	43 (94%)	3 (6%)	17 33
All	All	186/188 (99%)	170 (91%)	16 (9%)	10 20

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	B	12	LEU
3	B	17	GLU
3	C	16	ARG
3	B	4	MET
3	D	2	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
3	B	11	ASN
3	C	29	ASN
3	D	29	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](#)

There are no ligands in this entry.

#### 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

### 6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

### 6.4 Ligands

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers

EDS was not executed - this section is therefore empty.