



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 06:55 PM GMT

PDB ID : 1D5Y  
Title : CRYSTAL STRUCTURE OF THE E. COLI ROB TRANSCRIPTION FACTOR IN COMPLEX WITH DNA  
Authors : Kwon, H.J.; Bennik, M.H.J.; Demple, B.; Ellenberger, T.  
Deposited on : 1999-10-12  
Resolution : 2.70 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) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

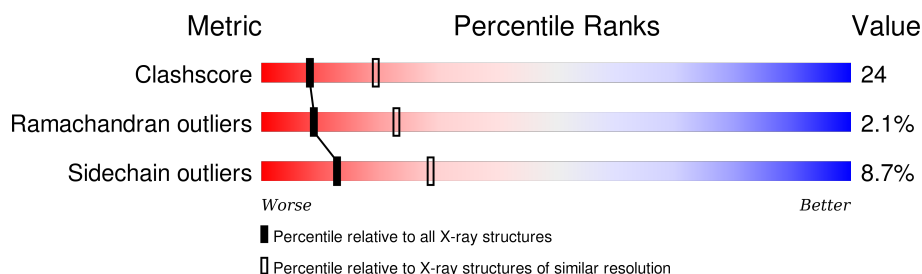
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.70 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	102246	2422 (2.70-2.70)
Ramachandran outliers	100387	2382 (2.70-2.70)
Sidechain outliers	100360	2382 (2.70-2.70)


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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	M	21	<div> <div>86%</div> <div>14%</div> </div>
1	O	21	<div> <div>86%</div> <div>14%</div> </div>
2	N	21	<div> <div>57%</div> <div>43%</div> </div>
2	P	21	<div> <div>38%</div> <div>62%</div> </div>
3	A	292	<div> <div>58%</div> <div>35%</div> <div>5%</div> </div>
3	B	292	<div> <div>57%</div> <div>37%</div> <div>5%</div> </div>
3	C	292	<div> <div>57%</div> <div>36%</div> <div>6%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	D	292	 A horizontal bar chart showing the quality of chain D. The bar is divided into three segments: green (54%), yellow (37%), and orange (7%). The segments are labeled with their respective percentages: 54%, 37%, and 7%. The bar ends with a small black dot.

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 10986 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\*GP\*AP\*CP\*AP\*GP\*CP\*AP\*CP\*TP\*GP\*AP\*AP\*TP\*GP\*TP\*CP\*AP\*AP\*AP\*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	M	21	Total	C	N	O	P	0	0	0
			431	206	85	120	20			
1	O	21	Total	C	N	O	P	0	0	0
			431	206	85	120	20			

- Molecule 2 is a DNA chain called DNA (5'-D(\*AP\*CP\*TP\*TP\*TP\*GP\*AP\*CP\*AP\*TP\*TP\*CP\*AP\*GP\*TP\*GP\*CP\*TP\*GP\*TP\*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	N	21	Total	C	N	O	P	0	0	0
			424	205	71	128	20			
2	P	21	Total	C	N	O	P	0	0	0
			424	205	71	128	20			

- Molecule 3 is a protein called ROB TRANSCRIPTION FACTOR.


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	288	Total	C	N	O	S	0	0	0
			2331	1490	408	423	10			
3	D	288	Total	C	N	O	S	0	0	0
			2307	1471	405	421	10			
3	C	288	Total	C	N	O	S	0	0	0
			2331	1490	408	423	10			
3	B	288	Total	C	N	O	S	0	0	0
			2307	1471	405	421	10			

### 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.

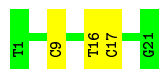
- Molecule 1: DNA (5'-D(\*TP\*GP\*AP\*CP\*AP\*GP\*CP\*AP\*CP\*TP\*GP\*AP\*AP\*TP\*GP\*TP\*CP\*AP\*AP\*AP\*G)-3')

Chain M: 



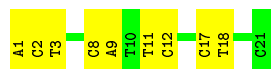
- Molecule 1: DNA (5'-D(\*TP\*GP\*AP\*CP\*AP\*GP\*CP\*AP\*CP\*TP\*GP\*AP\*AP\*TP\*GP\*TP\*CP\*AP\*AP\*AP\*G)-3')

Chain O: 

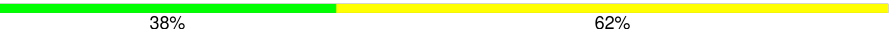


- Molecule 2: DNA (5'-D(\*AP\*CP\*TP\*TP\*TP\*GP\*AP\*CP\*AP\*TP\*TP\*CP\*AP\*GP\*TP\*GP\*CP\*TP\*GP\*TP\*C)-3')

Chain N: 



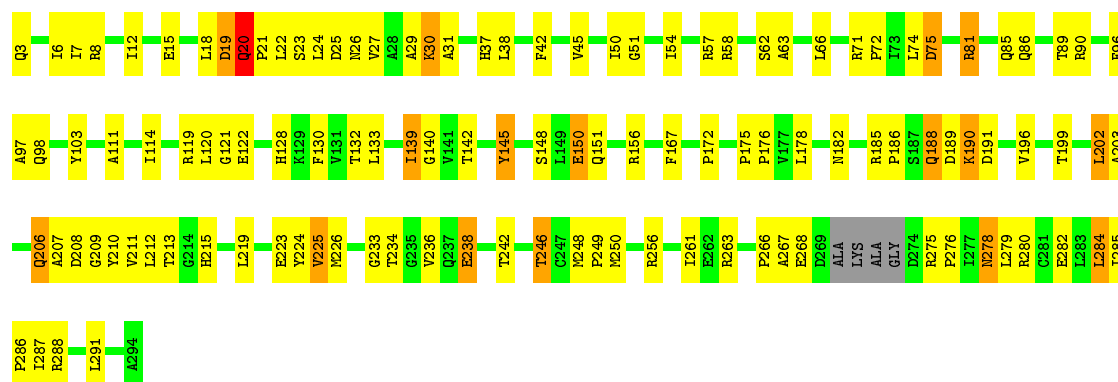
- Molecule 2: DNA (5'-D(\*AP\*CP\*TP\*TP\*TP\*GP\*AP\*CP\*AP\*TP\*TP\*CP\*AP\*GP\*TP\*GP\*CP\*TP\*GP\*TP\*C)-3')

Chain P: 



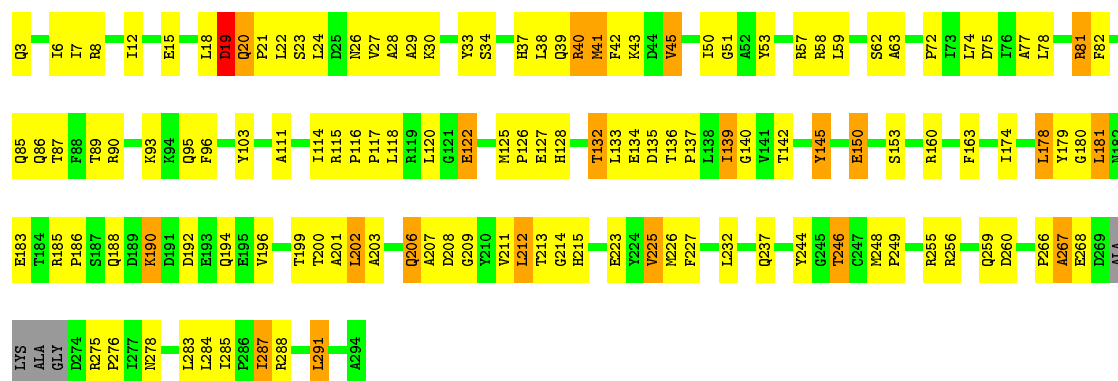
- Molecule 3: ROB TRANSCRIPTION FACTOR

Chain A: 



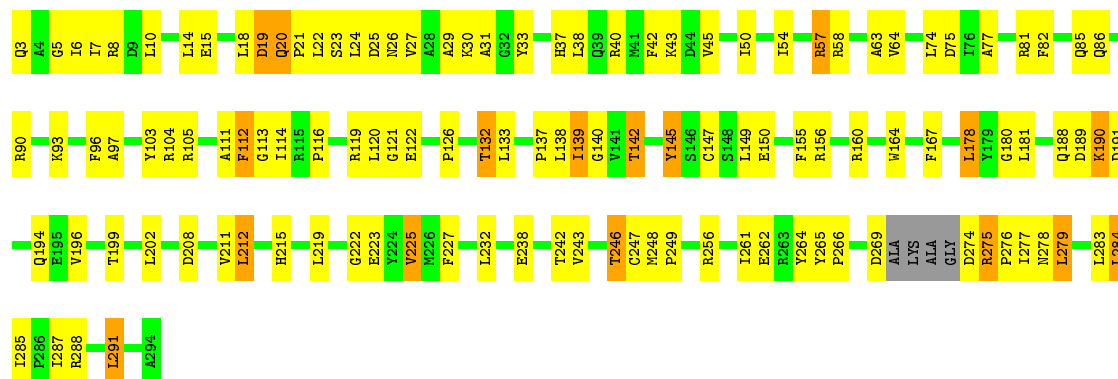
• Molecule 3: ROB TRANSCRIPTION FACTOR

Chain D: 54% 37% 7% .



• Molecule 3: ROB TRANSCRIPTION FACTOR

Chain C: 57% 36% 6% .



• Molecule 3: ROB TRANSCRIPTION FACTOR

Chain B: 57% 37% 5% .



E282	A201	R90
L283	L202	A91
L284	A203	F92
	Q204	K93
L287		K94
E288	D208	
		R104
L291	L212	R105
A294	L219	S106
		P107
	E223	E108
	T224	A111
	V225	
	M226	T114
	L232	R115
		P116
	V236	R119
	Q237	L120
	E238	G121
	F239	E122
	L240	F123
	L241	
	T242	P126
	T246	E127
	C247	H128
	M248	
	P249	T132
	P250	P137
	L251	L138
	N252	T139
	L253	G140
	T254	V141
	R255	T142
	R256	
	D260	Y145
	L261	E150
	E262	
	R263	S153
	L264	
	Y265	R156
	P266	
	A267	F163
	E268	
	D269	L178
	ALA	
	LVS	Q188
	ALA	D189
	GLY	K190
	D274	D191
	R275	
	P276	Q194
	L277	E195
	N278	V196
	L279	
	R280	T199
	C281	T200

## 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, $\alpha$ , $\beta$ , $\gamma$	75.38 Å   208.01 Å   67.73 Å 90.00°   90.00°   90.00°	Depositor
Resolution (Å)	30.00 – 2.70	Depositor
% Data completeness (in resolution range)	95.1 (30.00-2.70)	Depositor
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	CNS	Depositor
R, $R_{free}$	0.254 , 0.302	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	10986	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	58.0	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	M	0.38	0/485	0.67	0/747
1	O	0.38	0/485	0.67	0/747
2	N	0.38	0/473	0.71	0/728
2	P	0.38	0/473	0.69	0/728
3	A	0.42	0/2389	0.70	0/3237
3	B	0.44	0/2362	0.69	1/3201 (0.0%)
3	C	0.43	0/2389	0.69	0/3237
3	D	0.43	0/2362	0.69	0/3201
All	All	0.42	0/11418	0.69	1/15826 (0.0%)

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
3	B	0	1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
3	B	284	LEU	CA-CB-CG	5.98	129.05	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	B	145	TYR	Sidechain

## 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	M	431	0	237	2	0
1	O	431	0	237	2	0
2	N	424	0	241	5	0
2	P	424	0	241	12	0
3	A	2331	0	2304	124	0
3	B	2307	0	2272	106	0
3	C	2331	0	2304	132	0
3	D	2307	0	2272	137	0
All	All	10986	0	10108	515	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 24.

The worst 5 of 515 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)
3:C:291:LEU:H	3:C:291:LEU:HD12	1.33	0.94
3:B:269:ASP:HB3	3:B:277:ILE:HD12	1.50	0.93
3:A:20:GLN:HB2	3:A:21:PRO:HD3	1.47	0.93
3:A:190:LYS:H	3:A:190:LYS:HZ2	1.16	0.93
3:A:7:ILE:HD12	3:A:45:VAL:HB	1.53	0.90

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	284/292 (97%)	256 (90%)	20 (7%)	8 (3%)	6	15
3	B	284/292 (97%)	239 (84%)	41 (14%)	4 (1%)	14	35
3	C	284/292 (97%)	252 (89%)	27 (10%)	5 (2%)	11	27
3	D	284/292 (97%)	252 (89%)	25 (9%)	7 (2%)	7	18
All	All	1136/1168 (97%)	999 (88%)	113 (10%)	24 (2%)	9	23

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	19	ASP
3	A	267	ALA
3	D	19	ASP
3	D	33	TYR
3	D	267	ALA

### 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	A	246/247 (100%)	225 (92%)	21 (8%)	13	30
3	B	242/247 (98%)	221 (91%)	21 (9%)	13	29
3	C	246/247 (100%)	223 (91%)	23 (9%)	11	25
3	D	242/247 (98%)	222 (92%)	20 (8%)	14	31
All	All	976/988 (99%)	891 (91%)	85 (9%)	13	29

5 of 85 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	278	ASN
3	C	122	GLU
3	B	242	THR
3	D	287	ILE
3	C	38	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 32 such sidechains are listed below:

Mol	Chain	Res	Type
3	D	252	ASN
3	C	3	GLN
3	B	98	GLN
3	D	278	ASN
3	C	17	HIS

### 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 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.