



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 06:45 PM GMT

PDB ID : 1C7Y  
Title : E.COLI RUVA-HOLLIDAY JUNCTION COMPLEX  
Authors : Ariyoshi, M.; Nishino, T.; Iwasaki, H.; Shinagawa, H.; Morikawa, K.  
Deposited on : 2000-04-03  
Resolution : 3.10 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

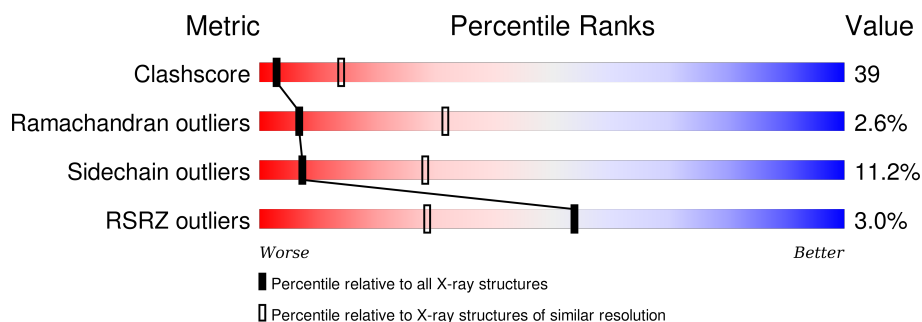
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.10 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	1222 (3.14-3.06)
Ramachandran outliers	100387	1174 (3.14-3.06)
Sidechain outliers	100360	1174 (3.14-3.06)
RSRZ outliers	91569	1119 (3.14-3.06)

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	B	13	
2	C	12	
3	D	13	
4	E	12	
5	F	13	
6	G	12	
7	H	13	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	I	12	<div><div></div><div>8%92%</div></div>
9	A	203	<div><div></div><div>46%45%6% ..</div></div>

## 2 Entry composition [i](#)

There are 9 unique types of molecules in this entry. The entry contains 3565 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(P\*DAP\*DAP\*DGP\*DTP\*DTP\*DGP\*DGP\*DGP\*DAP\*DTP\*DTP\*DGP\*DT)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	13	Total	C	N	O	P	0	13	0
			270	130	50	78	12			

- Molecule 2 is a DNA chain called DNA (5'-D(P\*DCP\*DTP\*DGP\*DTP\*DGP\*DTP\*DGP\*DTP\*DAP\*DAP\*DGP\*DC)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	12	Total	C	N	O	P	0	12	0
			248	118	44	74	12			

- Molecule 3 is a DNA chain called DNA (5'-D(P\*DGP\*DCP\*DTP\*DTP\*DAP\*DCP\*DAP\*DCP\*DAP\*DCP\*DAP\*DGP\*DA)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	13	Total	C	N	O	P	0	13	0
			262	126	51	73	12			

- Molecule 4 is a DNA chain called DNA (5'-D(P\*DGP\*DGP\*DTP\*DTP\*DAP\*DGP\*DGP\*DGP\*DTP\*DGP\*DAP\*DA)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	12	Total	C	N	O	P	0	12	0
			255	120	51	72	12			

- Molecule 5 is a DNA chain called DNA (5'-D(P\*DTP\*DTP\*DCP\*DAP\*DCP\*DCP\*DCP\*DTP\*DAP\*DAP\*DCP\*DCP\*DA)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	13	Total	C	N	O	P	0	13	0
			255	124	44	75	12			

- Molecule 6 is a DNA chain called DNA (5'-D(P\*DGP\*DAP\*DCP\*DAP\*DCP\*DAP\*DCP\*DAP\*DTP\*DTP\*DCP\*DG)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	G	12	Total	C	N	O	P	0	12	0
			244	116	46	70	12			

- Molecule 7 is a DNA chain called DNA (5'-D(P\*DCP\*DGP\*DAP\*DAP\*DTP\*DGP\*DTP\*DGP\*DTP\*DGP\*DTP\*DCP\*DT)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	H	13	Total	C	N	O	P	0	13	0
			265	128	46	79	12			

- Molecule 8 is a DNA chain called DNA (5'-D(P\*DCP\*DAP\*DAP\*DTP\*DCP\*DCP\*DCP\*DAP\*DAP\*DCP\*DTP\*DT)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	I	12	Total	C	N	O	P	0	12	0
			239	115	41	71	12			

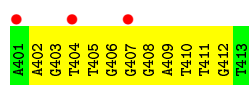
- Molecule 9 is a protein called HOLLIDAY JUNCTION DNA HELICASE RUVA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	A	199	Total	C	N	O	S	0	0	0
			1527	974	262	284	7			

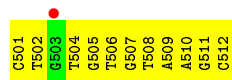
### 3 Residue-property plots

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.

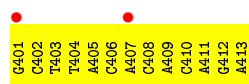
- Molecule 1: DNA (5'-D(P\*DAP\*DAP\*DGP\*DTP\*DTP\*DGP\*DGP\*DGP\*DAP\*DTP\*DTP\*DGP\*DT)-3')



- Molecule 2: DNA (5'-D(P\*DCP\*DTP\*DGP\*DTP\*DGP\*DTP\*DGP\*DTP\*DAP\*DAP\*DGP\*DC)-3')



- Molecule 3: DNA (5'-D(P\*DGP\*DCP\*DTP\*DTP\*DAP\*DCP\*DAP\*DCP\*DAP\*DCP\*DAP\*DGP\*DA)-3')

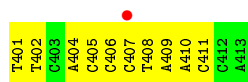


- Molecule 4: DNA (5'-D(P\*DGP\*DGP\*DTP\*DTP\*DAP\*DGP\*DGP\*DGP\*DTP\*DGP\*DAP\*DA)-3')

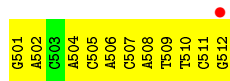


- Molecule 5: DNA (5'-D(P\*DTP\*DTP\*DCP\*DAP\*DCP\*DCP\*DCP\*DTP\*DAP\*DAP\*DCP\*DCP\*DA)-3')

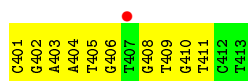




- Molecule 6: DNA (5'-D(P\*DGP\*DAP\*DCP\*DAP\*DCP\*DAP\*DCP\*DAP\*DTP\*DTP\*DCP\*DG)-3')



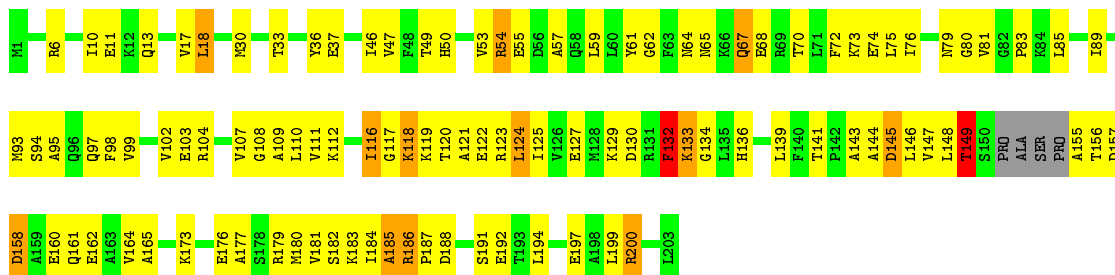
- Molecule 7: DNA (5'-D(P\*DCP\*DGP\*DAP\*DAP\*DTP\*DGP\*DTP\*DGP\*DTP\*DGP\*DTP\*DCP\*DT)-3')



- Molecule 8: DNA (5'-D(P\*DCP\*DAP\*DAP\*DTP\*DCP\*DCP\*DCP\*DAP\*DAP\*DCP\*DTP\*DT)-3')



- Molecule 9: HOLLIDAY JUNCTION DNA HELICASE RUVA



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 4 3 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	158.65Å 158.65Å 158.65Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.80 – 3.10 50.17 – 3.00	Depositor EDS
% Data completeness (in resolution range)	99.9 (45.80-3.10) 99.9 (50.17-3.00)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	9.44 (at 3.01Å)	Xtriage
Refinement program	CNS 0.9	Depositor
R, $R_{free}$	0.249 , 0.279 0.238 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	DCC
Wilson B-factor (Å <sup>2</sup> )	64.6	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 82.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 7133 reflections	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	3565	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	74.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.63% 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](#)

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	B	0.28	0/303	0.62	0/468
2	C	0.33	0/277	0.66	0/426
3	D	0.29	0/294	0.64	0/451
4	E	0.31	0/287	0.63	0/443
5	F	0.32	0/284	0.66	0/434
6	G	0.38	0/273	0.76	0/418
7	H	0.35	0/296	0.75	0/456
8	I	0.33	0/266	0.66	0/406
9	A	0.85	5/1549 (0.3%)	1.02	7/2088 (0.3%)
All	All	0.60	5/3829 (0.1%)	0.82	7/5590 (0.1%)

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
9	A	0	1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	A	116	ILE	C-N	15.58	1.61	1.33
9	A	117	GLY	CA-C	7.73	1.64	1.51
9	A	53	VAL	C-N	-5.83	1.20	1.34
9	A	118	LYS	N-CA	5.34	1.57	1.46
9	A	155	ALA	CA-CB	5.17	1.63	1.52

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
9	A	116	ILE	CB-CA-C	-14.01	83.58	111.60
9	A	134	GLY	N-CA-C	-11.92	83.31	113.10

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
9	A	53	VAL	C-N-CA	7.16	139.59	121.70
9	A	116	ILE	C-N-CA	-7.14	107.30	122.30
9	A	54	ARG	CB-CA-C	6.26	122.92	110.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
9	A	132	PHE	Mainchain

## 5.2 Too-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 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	B	270	0	124	25	0
2	C	248	0	118	28	0
3	D	262	0	125	34	0
4	E	255	0	114	32	0
5	F	255	0	127	17	0
6	G	244	0	116	21	0
7	H	265	0	126	24	0
8	I	239	0	120	28	0
9	A	1527	0	1579	101	3
All	All	3565	0	2549	240	3

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 39.

The worst 5 of 240 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:B:403[A]:DG:N1	2:C:510[A]:DA:C2	1.81	1.48
1:B:403[A]:DG:N1	2:C:510[A]:DA:H2	0.93	1.40
7:H:410[D]:DG:N1	8:I:503[D]:DA:N1	1.72	1.36
7:H:410[D]:DG:N2	8:I:503[D]:DA:H2	1.30	1.26
1:B:403[A]:DG:O6	2:C:510[A]:DA:N1	1.66	1.23

All (3) 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)
9:A:104:ARG:NE	9:A:104:ARG:NH1[13_555]	1.89	0.31
9:A:186:ARG:NH2	9:A:188:ASP:OD2[5_555]	1.91	0.29
9:A:104:ARG:NH1	9:A:104:ARG:NH1[13_555]	2.09	0.11

## 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	
9	A	195/203 (96%)	174 (89%)	16 (8%)	5 (3%)	7	32

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
9	A	132	PHE
9	A	133	LYS
9	A	124	LEU
9	A	185	ALA
9	A	129	LYS

### 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	
9	A	160/163 (98%)	142 (89%)	18 (11%)	7	28

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
9	A	145	ASP
9	A	149	THR
9	A	194	LEU
9	A	130	ASP
9	A	141	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
9	A	13	GLN
9	A	67	GLN
9	A	96	GLN
9	A	100	ASN
9	A	175	GLN

### 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 ⓘ

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	B	13/13 (100%)	1.13	3 (23%) 1 0	64, 86, 118, 120	13 (100%)
2	C	12/12 (100%)	0.67	1 (8%) 14 5	69, 86, 102, 112	12 (100%)
3	D	13/13 (100%)	1.14	2 (15%) 3 1	64, 97, 117, 117	13 (100%)
4	E	12/12 (100%)	0.70	0 100 100	64, 97, 105, 110	12 (100%)
5	F	13/13 (100%)	1.10	1 (7%) 16 5	66, 101, 125, 126	13 (100%)
6	G	12/12 (100%)	0.66	1 (8%) 14 5	33, 94, 112, 127	12 (100%)
7	H	13/13 (100%)	1.07	1 (7%) 16 5	69, 91, 122, 126	13 (100%)
8	I	12/12 (100%)	0.61	0 100 100	68, 99, 110, 112	12 (100%)
9	A	199/203 (98%)	-0.24	0 100 100	13, 40, 85, 104	0
All	All	299/303 (98%)	0.14	9 (3%) 54 29	13, 60, 115, 127	100 (33%)

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	407[A]	DG	2.8
3	D	407[B]	DA	2.6
5	F	407[C]	DC	2.5
3	D	401[B]	DG	2.4
7	H	407[D]	DT	2.2

### 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

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.