



# Full wwPDB X-ray Structure Validation Report ⓘ

Oct 16, 2021 – 07:52 PM EDT

PDB ID : 1T9I  
Title : I-CreI(D20N)/DNA complex  
Authors : Chevalier, B.; Sussman, D.; Otis, C.; Boudreau, D.; Turmel, M.; Lemieux, C.;  
Stephens, K.; Monnat Jr., R.J.; Stoddard, B.L.  
Deposited on : 2004-05-17  
Resolution : 1.60 Å(reported)

This is a Full wwPDB X-ray Structure 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

<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) : 2.23.2

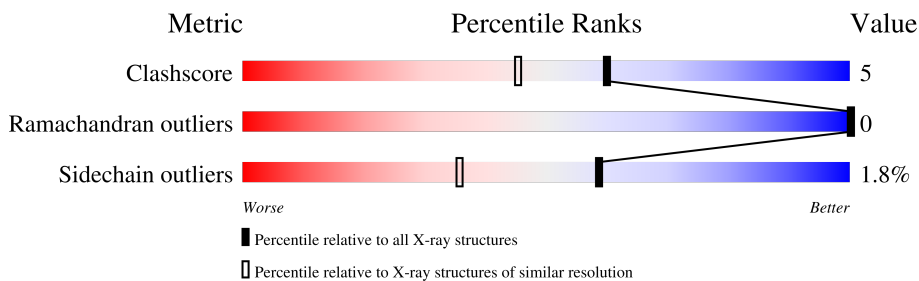
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.60 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	3665 (1.60-1.60)
Ramachandran outliers	138981	3564 (1.60-1.60)
Sidechain outliers	138945	3563 (1.60-1.60)

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 of 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	C	24	75% (green), 21% (yellow), 4% (orange), 0% (red), 0% (grey)
2	D	24	63% (green), 38% (yellow), 0% (orange), 0% (red), 0% (grey)
3	A	163	81% (green), 12% (yellow), 6% (orange), 0% (red), 0% (grey)
3	B	163	87% (green), 8% (yellow), 5% (orange), 0% (red), 0% (grey)

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 4132 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 5'-D(\*GP\*CP\*AP\*AP\*AP\*AP\*CP\*GP\*TP\*CP\*GP\*TP\*GP\*AP\*GP\*AP\*CP\*AP\*GP\*TP\*TP\*TP\*CP\*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	C	24	493	235	95	140	23	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	D	24	485	233	85	144	23	0	0	0

- Molecule 3 is a protein called DNA endonuclease I-CreI.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	A	153	1241	799	212	229	1	0	0	0
3	B	155	1251	805	214	231	1	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	20	ASN	ASP	engineered mutation	UNP P05725
A	42	THR	ALA	SEE REMARK 999	UNP P05725
A	110	GLU	TRP	SEE REMARK 999	UNP P05725
A	111	GLN	ARG	SEE REMARK 999	UNP P05725
B	320	ASN	ASP	engineered mutation	UNP P05725
B	342	THR	ALA	SEE REMARK 999	UNP P05725
B	410	GLU	TRP	SEE REMARK 999	UNP P05725
B	411	GLN	ARG	SEE REMARK 999	UNP P05725

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	C	2	Total Ca 2 2	0	0

- Molecule 5 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	C	1	Total Na 1 1	0	0
5	D	1	Total Na 1 1	0	0
5	A	1	Total Na 1 1	0	0
5	B	1	Total Na 1 1	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	C	110	Total O 110 110	0	0
6	D	94	Total O 94 94	0	0
6	A	226	Total O 226 226	0	0
6	B	226	Total O 226 226	0	0

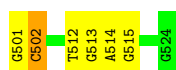
### 3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: 5'-D(\*GP\*CP\*AP\*AP\*AP\*AP\*CP\*GP\*TP\*CP\*GP\*TP\*GP\*AP\*GP\*AP\*CP\*AP\*GP\*TP\*TP\*TP\*CP\*G)-3'

Chain C: 




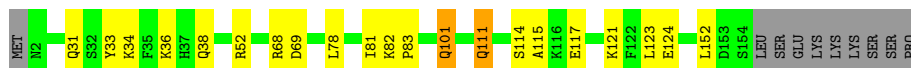
- Molecule 2: 5'-D(\*CP\*GP\*AP\*AP\*AP\*CP\*TP\*GP\*TP\*CP\*TP\*CP\*AP\*CP\*GP\*AP\*CP\*GP\*TP\*TP\*TP\*TP\*GP\*C)-3'

Chain D: 




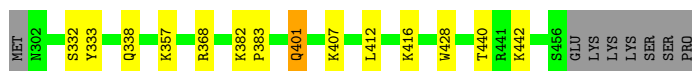
- Molecule 3: DNA endonuclease I-CreI

Chain A: 



- Molecule 3: DNA endonuclease I-CreI

Chain B: 



## 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, $\alpha$ , $\beta$ , $\gamma$	43.80Å 67.31Å 89.35Å 90.00° 92.40° 90.00°	Depositor
Resolution (Å)	19.97 – 1.60	Depositor
% Data completeness (in resolution range)	99.0 (19.97-1.60)	Depositor
$R_{merge}$	0.04	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.194 , 0.211	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	4132	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	20.0	wwPDB-VP

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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

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	C	0.63	0/554	0.89	0/854
2	D	0.59	0/542	0.91	0/834
3	A	0.59	0/1264	0.74	0/1707
3	B	0.58	0/1274	0.77	0/1721
All	All	0.59	0/3634	0.81	0/5116

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
1	C	0	5
2	D	0	5
All	All	0	10

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (10) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	502	DC	Sidechain
1	C	512	DT	Sidechain
1	C	513	DG	Sidechain
1	C	514	DA	Sidechain
1	C	515	DG	Sidechain
2	D	562	DC	Sidechain
2	D	563	DA	Sidechain

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Group
2	D	564	DC	Sidechain
2	D	565	DG	Sidechain
2	D	567	DC	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	C	493	0	271	2	0
2	D	485	0	273	5	0
3	A	1241	0	1274	16	0
3	B	1251	0	1278	11	0
4	C	2	0	0	0	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
5	C	1	0	0	0	0
5	D	1	0	0	0	0
6	A	226	0	0	4	1
6	B	226	0	0	1	1
6	C	110	0	0	1	0
6	D	94	0	0	1	0
All	All	4132	0	3096	30	1

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 5.

All (30) 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:A:78:LEU:HD12	6:A:1192:HOH:O	1.72	0.87
3:A:101:GLN:H	3:A:101:GLN:HE21	1.25	0.84
3:B:401:GLN:H	3:B:401:GLN:HE21	1.24	0.84
6:C:1056:HOH:O	3:B:440:THR:HG23	1.87	0.74
2:D:552:DG:H5'	6:D:1588:HOH:O	1.89	0.72
3:A:124:GLU:HB2	6:A:1202:HOH:O	1.89	0.71
3:A:82:LYS:HB3	3:A:83:PRO:HD3	1.76	0.67

*Continued on next page...*



Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:B:412:LEU:O	3:B:416:LYS:HD3	1.95	0.66
3:A:123:LEU:HD21	3:A:152:LEU:HD12	1.77	0.64
3:A:114:SER:O	3:A:117:GLU:HG2	2.04	0.57
3:B:401:GLN:HE21	3:B:401:GLN:N	1.99	0.56
1:C:502:DC:H2'	3:B:333:TYR:CE2	2.40	0.56
2:D:568:DG:H2'	2:D:569:DT:H72	1.90	0.54
3:A:52:ARG:NH1	3:A:69:ASP:OD2	2.41	0.54
3:A:31:GLN:NE2	3:A:36:LYS:HD2	2.26	0.50
3:A:101:GLN:HE21	3:A:101:GLN:N	2.02	0.49
3:A:115:ALA:HA	3:A:121:LYS:HG3	1.94	0.49
3:A:36:LYS:NZ	6:A:1654:HOH:O	2.34	0.48
3:A:33:TYR:CZ	3:A:38:GLN:HB2	2.49	0.47
3:A:31:GLN:HE22	3:A:36:LYS:HD2	1.80	0.46
3:A:34:LYS:HD3	6:A:1640:HOH:O	2.15	0.46
3:B:357:LYS:HD3	6:B:1163:HOH:O	2.15	0.46
3:A:111:GLN:CG	3:A:121:LYS:NZ	2.79	0.46
1:C:501:DG:H2'	3:B:332:SER:O	2.17	0.44
2:D:554:DA:OP2	3:A:81:ILE:HG13	2.17	0.43
2:D:568:DG:H2'	2:D:569:DT:C7	2.49	0.42
3:B:407:LYS:HE2	3:B:428:TRP:CZ2	2.55	0.42
2:D:568:DG:H5''	3:B:442:LYS:HG3	2.01	0.41
3:B:382:LYS:HB2	3:B:383:PRO:HD3	2.02	0.41
3:B:333:TYR:CZ	3:B:338:GLN:HB2	2.56	0.40

All (1) 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)
6:A:1290:HOH:O	6:B:1641:HOH:O[1_545]	2.02	0.18

## 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	151/163 (93%)	145 (96%)	6 (4%)	0	100	100
3	B	153/163 (94%)	146 (95%)	7 (5%)	0	100	100
All	All	304/326 (93%)	291 (96%)	13 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 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	139/150 (93%)	136 (98%)	3 (2%)	52	27
3	B	139/150 (93%)	137 (99%)	2 (1%)	67	47
All	All	278/300 (93%)	273 (98%)	5 (2%)	59	36

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	68	ARG
3	A	101	GLN
3	A	111	GLN
3	B	368	ARG
3	B	401	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
3	A	2	ASN
3	A	26	GLN
3	A	31	GLN
3	A	44	GLN
3	A	101	GLN
3	A	103	ASN
3	B	350	GLN
3	B	401	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 monosaccharides in this entry.

### 5.6 Ligand geometry [i](#)

Of 6 ligands modelled in this entry, 6 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.

No monomer is involved in short contacts.

### 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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

EDS was not executed - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates [i](#)

EDS was not executed - this section is therefore empty.

### 6.4 Ligands [i](#)

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers [i](#)

EDS was not executed - this section is therefore empty.