



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 04:47 AM GMT

PDB ID : 2O7M  
Title : The C-terminal loop of the homing endonuclease I-CreI is essential for DNA binding and cleavage. Identification of a novel site for specificity engineering in the I-CreI scaffold  
Authors : Prieto, J.; Redondo, P.; Padro, D.; Blanco, F.J.; Paques, F.; Montoya, G.  
Deposited on : 2006-12-11  
Resolution : 2.00 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  
<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

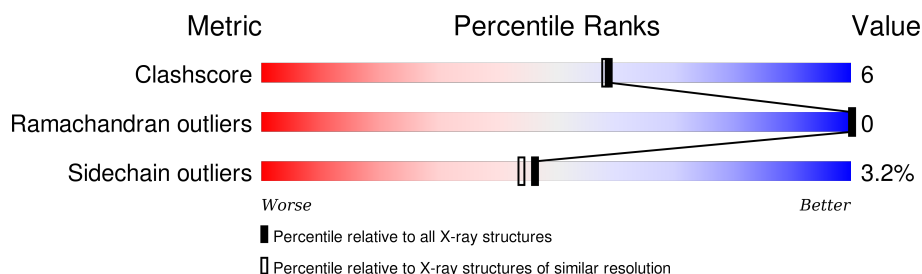
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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	7340 (2.00-2.00)
Ramachandran outliers	100387	7248 (2.00-2.00)
Sidechain outliers	100360	7247 (2.00-2.00)

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	A	156	 85% 11% . .
1	B	156	 81% 15% . . .

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2746 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 protein called DNA endonuclease I-CreI.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	153	Total	C	N	O	S	0	0	0
			1242	799	211	231	1			
1	B	153	Total	C	N	O	S	0	0	0
			1242	799	211	231	1			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	42	THR	ALA	CONFLICT	UNP P05725
A	110	GLU	TRP	CONFLICT	UNP P05725
A	111	GLN	ARG	CONFLICT	UNP P05725
B	42	THR	ALA	CONFLICT	UNP P05725
B	110	GLU	TRP	CONFLICT	UNP P05725
B	111	GLN	ARG	CONFLICT	UNP P05725

- Molecule 2 is water.


Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	129	Total	O	0	0
			129	129		
2	B	133	Total	O	0	0
			133	133		

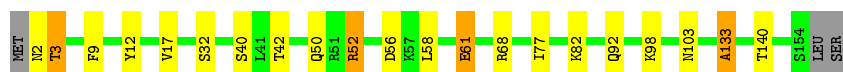
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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 endonuclease I-CreI

Chain A:  85% 11% . .



- Molecule 1: DNA endonuclease I-CreI

Chain B:  81% 15% . . .



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	P 43	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	69.09 Å 69.09 Å 93.04 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	34.54 – 2.00	Depositor
% Data completeness (in resolution range)	93.2 (34.54-2.00)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.184 , 0.235	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2746	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	34.0	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	A	1.34	2/1265 (0.2%)	1.05	2/1708 (0.1%)
1	B	1.35	6/1265 (0.5%)	1.00	3/1708 (0.2%)
All	All	1.35	8/2530 (0.3%)	1.03	5/3416 (0.1%)

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	61	GLU	CD-OE2	5.61	1.31	1.25
1	A	61	GLU	CD-OE2	5.44	1.31	1.25
1	B	125	VAL	CB-CG2	5.38	1.64	1.52
1	B	8	GLU	CG-CD	5.34	1.59	1.51
1	B	87	PHE	CE1-CZ	5.20	1.47	1.37
1	B	120	ASP	CB-CG	5.12	1.62	1.51
1	B	102	ALA	CA-CB	5.08	1.63	1.52
1	A	133	ALA	CA-CB	5.07	1.63	1.52

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	52	ARG	NE-CZ-NH1	-6.85	116.87	120.30
1	A	52	ARG	NE-CZ-NH1	-6.64	116.98	120.30
1	A	40	SER	CB-CA-C	-5.99	98.72	110.10
1	B	52	ARG	NE-CZ-NH2	5.99	123.29	120.30
1	B	107	LYS	CD-CE-NZ	-5.37	99.35	111.70

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	A	1242	0	1275	13	0
1	B	1242	0	1275	17	0
2	A	129	0	0	2	0
2	B	133	0	0	3	0
All	All	2746	0	2550	29	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 6.

All (29) 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:A:50:GLN:HB2	2:A:221:HOH:O	1.62	0.95
1:B:9:PHE:CE1	1:B:61:GLU:HG2	2.09	0.87
1:B:50:GLN:HB2	2:B:180:HOH:O	1.77	0.84
1:A:9:PHE:CE1	1:A:61:GLU:HG2	2.13	0.83
1:A:50:GLN:HG2	2:A:200:HOH:O	1.79	0.81
1:B:52:ARG:NH1	1:B:56:ASP:OD2	2.21	0.73
1:A:52:ARG:NH1	1:A:56:ASP:OD2	2.25	0.70
1:B:82:LYS:HB3	1:B:83:PRO:HD3	1.81	0.62
1:A:68:ARG:HG2	1:A:77:ILE:HD12	1.84	0.58
1:B:6:ASN:HD22	1:B:9:PHE:H	1.52	0.57
1:A:2:ASN:CG	1:A:3:THR:H	2.09	0.55
1:A:42:THR:HG23	1:A:77:ILE:HG23	1.90	0.53
1:A:92:GLN:HG3	1:A:103:ASN:OD1	2.12	0.50
1:A:17:VAL:O	1:A:98:LYS:HE2	2.12	0.49
1:B:107:LYS:HD2	1:B:128:TRP:CZ2	2.49	0.48
1:A:68:ARG:HG2	1:A:77:ILE:CD1	2.43	0.47
1:A:133:ALA:HB3	1:A:140:THR:HG22	1.99	0.45
1:B:6:ASN:ND2	1:B:9:PHE:H	2.15	0.45
1:B:134:ALA:HB2	1:B:140:THR:HG21	2.00	0.43
1:B:50:GLN:HG3	2:B:180:HOH:O	2.18	0.42
1:B:50:GLN:H	1:B:50:GLN:HE21	1.65	0.42
1:B:52:ARG:CG	1:B:52:ARG:HH11	2.32	0.42
1:A:2:ASN:CG	1:A:3:THR:N	2.72	0.42
1:B:142:LYS:HE3	2:B:255:HOH:O	2.19	0.42
1:B:39:LEU:HD22	1:B:112:LEU:HD21	2.01	0.41
1:A:12:TYR:CE1	1:B:15:GLY:HA2	2.55	0.41
1:B:81:ILE:HG23	1:B:109:ILE:HG23	2.01	0.41

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:9:PHE:CE1	1:B:61:GLU:CG	2.94	0.41
1:B:42:THR:HG23	1:B:77:ILE:HG23	2.03	0.41

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	
1	A	151/156 (97%)	148 (98%)	3 (2%)	0	100	100
1	B	151/156 (97%)	148 (98%)	3 (2%)	0	100	100
All	All	302/312 (97%)	296 (98%)	6 (2%)	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	
1	A	140/143 (98%)	136 (97%)	4 (3%)	50	49
1	B	140/143 (98%)	135 (96%)	5 (4%)	42	39
All	All	280/286 (98%)	271 (97%)	9 (3%)	46	44

All (9) residues with a non-rotameric sidechain are listed below:



Mol	Chain	Res	Type
1	A	3	THR
1	A	32	SER
1	A	58	LEU
1	A	82	LYS
1	B	32	SER
1	B	50	GLN
1	B	52	ARG
1	B	107	LYS
1	B	139	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	2	ASN
1	A	31	GLN
1	A	47	GLN
1	A	92	GLN
1	B	2	ASN
1	B	6	ASN
1	B	50	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

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.