wwPDB X-ray Structure Validation Summary Report (i)

Aug 2, 2023 – 11:17 PM EDT

PDB ID	:	1G9Z
Title	:	LAGLIDADG HOMING ENDONUCLEASE I-CREI / DNA PRODUCT COMPLEX WITH MAGNESIUM
Authors	:	Chevalier, B.; Monnat, R.J.; Stoddard, B.L.
Deposited on	:	2000-11-28
Resolution	:	1.80 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

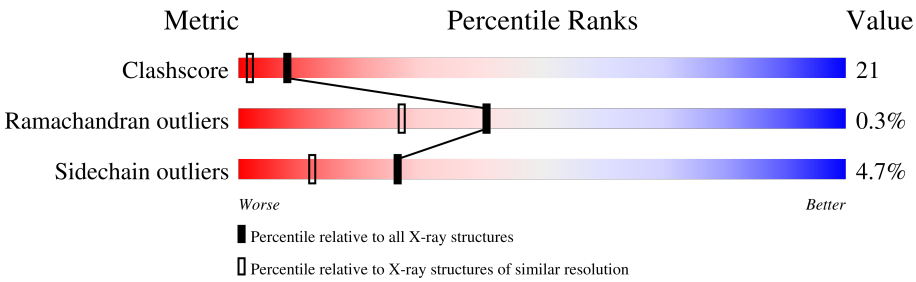
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.34

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 1.80 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	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)

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 >=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 <=5%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	C	14	50% 50%
2	D	10	40% 50% 10%
3	E	14	36% 57% 7%
4	F	10	10% 80% 10%
5	A	152	66% 33% •
5	B	152	68% 30% •

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 4314 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*A)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	C	14	287	137	58	79	13	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	D	10	207	98	37	62	10	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	E	14	280	135	51	81	13	0	0	0

- Molecule 4 is a DNA chain called 5'-D(P*GP*AP*CP*GP*TP*TP*TP*TP*GP*C)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
4	F	10	206	98	34	64	10	0	0	0

- Molecule 5 is a protein called DNA ENDONUCLEASE I-CREI.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	A	152	1237	796	210	230	1	0	0	0
5	B	152	1237	796	210	230	1	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
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	242	THR	ALA	SEE REMARK 999	UNP P05725
B	310	GLU	TRP	SEE REMARK 999	UNP P05725
B	311	GLN	ARG	SEE REMARK 999	UNP P05725

- Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	C	1	Total Mg 1 1	0	0
6	D	1	Total Mg 1 1	0	0
6	F	1	Total Mg 1 1	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	C	88	Total O 88 88	0	0
7	D	59	Total O 59 59	0	0
7	E	74	Total O 74 74	0	0
7	F	62	Total O 62 62	0	0
7	A	279	Total O 279 279	0	0
7	B	295	Total O 295 295	0	0

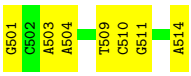
3 Residue-property plots (i)

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.

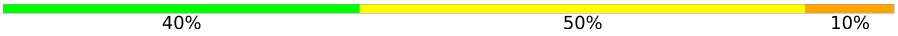
• Molecule 1: 5'-D(*GP*CP*AP*AP*AP*AP*CP*GP*TP*CP*GP*TP*GP*A)-3'

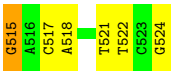
Chain C:



• Molecule 2: 5'-D(P*GP*AP*CP*AP*GP*TP*TP*TP*CP*G)-3'

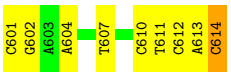
Chain D:





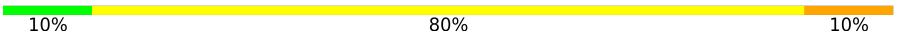
• Molecule 3: 5'-D(*CP*GP*AP*AP*AP*CP*TP*GP*TP*CP*TP*CP*AP*C)-3'

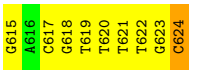
Chain E:



• Molecule 4: 5'-D(P*GP*AP*CP*GP*TP*TP*TP*TP*GP*C)-3'

Chain F:



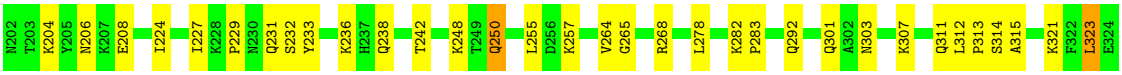


• Molecule 5: DNA ENDONUCLEASE I-CREI

Chain A:

• Molecule 5: DNA ENDONUCLEASE I-CREI

Chain B:



4 Data and refinement statistics (i)

Xtriage (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, α, β, γ	43.00Å 67.90Å 88.30Å 90.00° 91.60° 90.00°	Depositor
Resolution (Å)	19.92 – 1.80	Depositor
% Data completeness (in resolution range)	95.2 (19.92-1.80)	Depositor
R_{merge}	0.04	Depositor
R_{sym}	3.80	Depositor
Refinement program		Depositor
R, R_{free}	0.204 , 0.249	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4314	wwPDB-VP
Average B, all atoms (Å^2)	32.0	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: MG

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	0.26	0/323	0.69	0/497
2	D	0.61	1/231 (0.4%)	0.71	0/353
3	E	0.29	0/313	0.72	0/480
4	F	0.76	1/229 (0.4%)	0.88	1/350 (0.3%)
5	A	0.34	0/1260	0.60	0/1700
5	B	0.34	0/1260	0.60	0/1700
All	All	0.39	2/3616 (0.1%)	0.65	1/5080 (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	E	0	1
4	F	0	1
All	All	0	2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	F	615	DG	OP3-P	-7.47	1.52	1.61
2	D	515	DG	OP3-P	-7.09	1.52	1.61

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	F	615	DG	OP1-P-OP2	-5.05	112.02	119.60

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	E	614	DC	Sidechain
4	F	624	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	287	0	158	13	0
2	D	207	0	114	9	0
3	E	280	0	159	15	0
4	F	206	0	115	11	0
5	A	1237	0	1270	57	0
5	B	1237	0	1270	46	0
6	C	1	0	0	0	0
6	D	1	0	0	0	0
6	F	1	0	0	0	0
7	A	279	0	0	15	0
7	B	295	0	0	14	0
7	C	88	0	0	1	0
7	D	59	0	0	1	0
7	E	74	0	0	1	0
7	F	62	0	0	3	0
All	All	4314	0	3086	140	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 21.

The worst 5 of 140 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)
5:A:101:GLN:H	5:A:101:GLN:HE21	1.03	0.98
2:D:524:DG:H1	3:E:601:DC:H42	1.06	0.94
5:A:82:LYS:HE3	5:A:82:LYS:HA	1.49	0.93
5:B:326:CYS:O	5:B:329:VAL:HG22	1.75	0.87
3:E:613:DA:H1'	3:E:614:DC:O4'	1.78	0.83

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	
5	A	150/152 (99%)	144 (96%)	6 (4%)	0	100	100
5	B	150/152 (99%)	142 (95%)	7 (5%)	1 (1%)	22	10
All	All	300/304 (99%)	286 (95%)	13 (4%)	1 (0%)	41	27

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	B	349	ARG

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	
5	A	139/139 (100%)	132 (95%)	7 (5%)	24	10
5	B	139/139 (100%)	133 (96%)	6 (4%)	29	14
All	All	278/278 (100%)	265 (95%)	13 (5%)	26	12

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	B	250	GLN
5	B	255	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
5	B	323	LEU
5	B	278	LEU
5	B	301	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
5	B	231	GLN
5	B	250	GLN
5	B	292	GLN
5	A	92	GLN
5	A	101	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 3 ligands modelled in this entry, 3 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.