



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 12:29 am GMT

PDB ID : 1N3F
Title : Crystal structure of I-CreI bound to a palindromic DNA sequence II (palindrome of right side of wildtype DNA target sequence)
Authors : Chevalier, B.; Turmel, M.; Lemieux, C.; Monnat, R.J.; Stoddard, B.L.
Deposited on : 2002-10-28
Resolution : 2.00 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

<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
Xtriage (Phenix) : 1.9-1692
EDS : trunk28620
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
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) : recalc28949

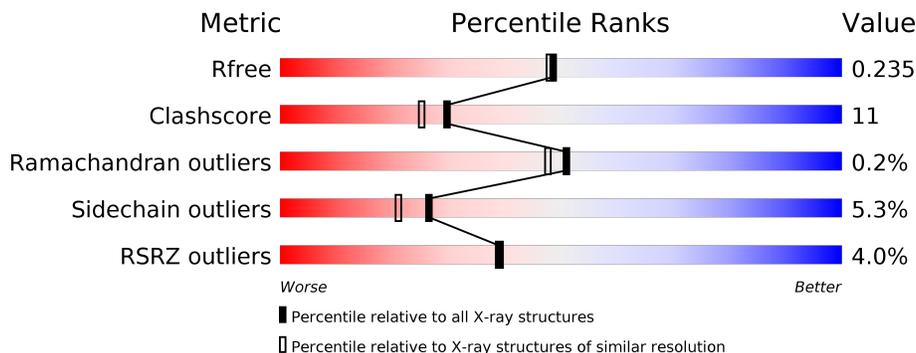
1 Overall quality at a glance

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))
R_{free}	100719	6609 (2.00-2.00)
Clashscore	112137	7775 (2.00-2.00)
Ramachandran outliers	110173	7679 (2.00-2.00)
Sidechain outliers	110143	7678 (2.00-2.00)
RSRZ outliers	101464	6696 (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 ≥ 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	C	14	
1	E	14	
1	I	14	
1	K	14	
2	D	10	
2	F	10	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	J	10	 70% 30%
2	L	10	 70% 30%
3	A	163	 4% 75% 16% • 9%
3	B	163	 4% 74% 15% • 9%
3	G	163	 5% 64% 26% • 9%
3	H	163	 4% 66% 21% • • 9%

2 Entry composition i

There are 5 unique types of molecules in this entry. The entry contains 7327 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(*CP*GP*AP*AP*AP*CP*TP*GP*TP*CP*TP*CP*GP*A)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	C	14	Total 283	C 136	N 53	O 81	P 13	0	0	0
1	E	14	Total 283	C 136	N 53	O 81	P 13	0	0	0
1	I	14	Total 283	C 136	N 53	O 81	P 13	0	0	0
1	K	14	Total 283	C 136	N 53	O 81	P 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	Total 207	C 98	N 37	O 62	P 10	0	0	0
2	F	10	Total 207	C 98	N 37	O 62	P 10	0	0	0
2	J	10	Total 207	C 98	N 37	O 62	P 10	0	0	0
2	L	10	Total 207	C 98	N 37	O 62	P 10	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	149	Total 1212	C 782	N 206	O 223	S 1	0	0	0
3	B	149	Total 1212	C 782	N 206	O 223	S 1	0	0	0
3	G	149	Total 1212	C 782	N 206	O 223	S 1	0	0	0
3	H	149	Total 1212	C 782	N 206	O 223	S 1	0	0	0

There are 12 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
G	542	THR	ALA	SEE REMARK 999	UNP P05725
G	610	GLU	TRP	SEE REMARK 999	UNP P05725
G	611	GLN	ARG	SEE REMARK 999	UNP P05725
H	742	THR	ALA	SEE REMARK 999	UNP P05725
H	810	GLU	TRP	SEE REMARK 999	UNP P05725
H	811	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	J	1	Total Ca 1 1	0	0
4	D	1	Total Ca 1 1	0	0
4	I	1	Total Ca 1 1	0	0
4	C	1	Total Ca 1 1	0	0
4	L	1	Total Ca 1 1	0	0
4	F	1	Total Ca 1 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	110	Total O 110 110	0	0
5	B	78	Total O 78 78	0	0
5	C	42	Total O 42 42	0	0
5	D	9	Total O 9 9	0	0
5	E	22	Total O 22 22	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	F	28	Total 28	O 28	0	0
5	G	81	Total 81	O 81	0	0
5	H	92	Total 92	O 92	0	0
5	I	15	Total 15	O 15	0	0
5	J	13	Total 13	O 13	0	0
5	K	15	Total 15	O 15	0	0
5	L	8	Total 8	O 8	0	0

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 the various outlier classes 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: 5'-D(*CP*GP*AP*AP*AP*CP*TP*GP*TP*CP*TP*CP*GP*A)-3'

Chain C: 



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

Chain E: 



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

Chain I: 



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

Chain K: 



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

Chain D: 



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

Chain F: 



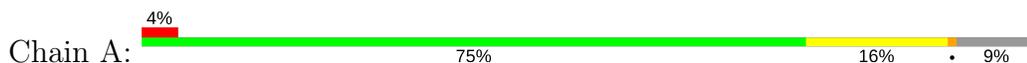
- Molecule 2: 5'-D(P*GP*AP*CP*AP*GP*TP*TP*TP*CP*G-3')



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



- Molecule 3: DNA endonuclease I-CreI



- Molecule 3: DNA endonuclease I-CreI

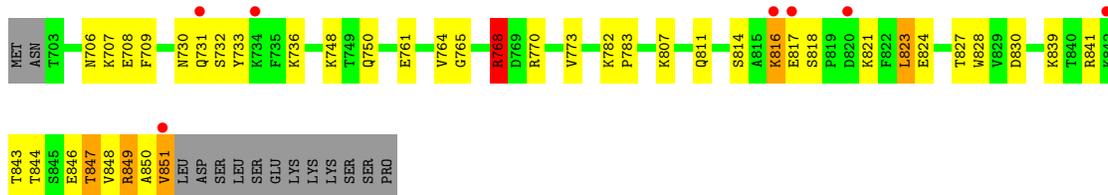


- Molecule 3: DNA endonuclease I-CreI



- Molecule 3: DNA endonuclease I-CreI





4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	78.32Å 76.39Å 81.06Å 90.00° 108.75° 90.00°	Depositor
Resolution (Å)	19.85 – 2.00 19.84 – 2.00	Depositor EDS
% Data completeness (in resolution range)	98.2 (19.85-2.00) 98.3 (19.84-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.32 (at 2.01Å)	Xtrriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.195 , 0.228 0.204 , 0.235	Depositor DCC
R_{free} test set	2993 reflections (4.99%)	DCC
Wilson B-factor (Å ²)	31.4	Xtrriage
Anisotropy	0.041	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 55.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	0.005 for l,-k,h	Xtrriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	7327	wwPDB-VP
Average B, all atoms (Å ²)	38.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 20.52 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to $8.5453e-03$. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

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

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	0.37	0/317	0.82	0/487
1	E	0.31	0/317	0.78	0/487
1	I	0.29	0/317	0.76	0/487
1	K	0.27	0/317	0.76	0/487
2	D	0.63	1/231 (0.4%)	0.82	0/353
2	F	0.73	1/231 (0.4%)	0.87	1/353 (0.3%)
2	J	0.69	1/231 (0.4%)	0.85	0/353
2	L	0.65	1/231 (0.4%)	0.83	0/353
3	A	0.39	0/1235	0.65	0/1667
3	B	0.38	0/1235	0.67	2/1667 (0.1%)
3	G	0.39	0/1235	0.62	0/1667
3	H	0.36	0/1235	0.63	0/1667
All	All	0.42	4/7132 (0.1%)	0.70	3/10028 (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
1	C	0	1
1	E	0	1
1	I	0	1
1	K	0	1
3	H	0	1
All	All	0	5

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	J	915	DG	OP3-P	-7.95	1.51	1.61

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	465	DG	OP3-P	-7.64	1.51	1.61
2	L	965	DG	OP3-P	-7.33	1.52	1.61
2	D	415	DG	OP3-P	-6.88	1.52	1.61

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	260	ASP	CB-CG-OD2	-8.48	110.67	118.30
3	B	260	ASP	CB-CG-OD1	5.52	123.27	118.30
2	F	465	DG	OP1-P-OP2	-5.19	111.82	119.60

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	414	DA	Sidechain
1	E	464	DA	Sidechain
3	H	768	ARG	Sidechain
1	I	914	DA	Sidechain
1	K	964	DA	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	283	0	158	6	0
1	E	283	0	158	5	0
1	I	283	0	158	6	0
1	K	283	0	158	9	0
2	D	207	0	114	3	0
2	F	207	0	114	1	0
2	J	207	0	114	1	0
2	L	207	0	114	1	0
3	A	1212	0	1249	20	0
3	B	1212	0	1249	22	1
3	G	1212	0	1249	38	1
3	H	1212	0	1249	46	1

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	C	1	0	0	0	0
4	D	1	0	0	0	0
4	F	1	0	0	0	0
4	I	1	0	0	0	0
4	J	1	0	0	0	0
4	L	1	0	0	0	0
5	A	110	0	0	5	3
5	B	78	0	0	2	2
5	C	42	0	0	2	0
5	D	9	0	0	0	0
5	E	22	0	0	0	0
5	F	28	0	0	0	0
5	G	81	0	0	2	1
5	H	92	0	0	2	3
5	I	15	0	0	2	0
5	J	13	0	0	0	0
5	K	15	0	0	0	0
5	L	8	0	0	0	0
All	All	7327	0	6084	138	6

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 11.

The worst 5 of 138 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:H:844:THR:HG22	3:H:847:THR:CG2	1.94	0.97
3:G:602:ALA:O	3:G:606:LEU:HD23	1.66	0.95
3:G:648:VAL:O	3:G:651:VAL:HG23	1.65	0.94
3:H:823:LEU:HD11	3:H:849:ARG:HD2	1.52	0.89
3:H:844:THR:HG22	3:H:847:THR:HG23	1.54	0.89

The worst 5 of 6 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)
3:B:260:ASP:OD2	3:H:849:ARG:NH2[2_655]	1.60	0.60
3:G:561:GLU:OE1	5:B:1482:HOH:O[1_655]	1.70	0.50
5:A:1443:HOH:O	5:H:1453:HOH:O[1_455]	2.06	0.14
5:B:1482:HOH:O	5:G:1478:HOH:O[1_455]	2.08	0.12

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:A:1443:HOH:O	5:H:1472:HOH:O[1_455]	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	
3	A	147/163 (90%)	139 (95%)	8 (5%)	0	100	100
3	B	147/163 (90%)	139 (95%)	8 (5%)	0	100	100
3	G	147/163 (90%)	137 (93%)	10 (7%)	0	100	100
3	H	147/163 (90%)	139 (95%)	7 (5%)	1 (1%)	25	18
All	All	588/652 (90%)	554 (94%)	33 (6%)	1 (0%)	51	48

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	H	850	ALA

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	136/150 (91%)	132 (97%)	4 (3%)	48	47
3	B	136/150 (91%)	129 (95%)	7 (5%)	28	22
3	G	136/150 (91%)	126 (93%)	10 (7%)	16	10

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	H	136/150 (91%)	128 (94%)	8 (6%)	23	17
All	All	544/600 (91%)	515 (95%)	29 (5%)	26	21

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	G	550	GLN
3	G	577	ILE
3	H	847	THR
3	G	552	ARG
3	G	606	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
3	G	550	GLN
3	H	750	GLN
3	G	603	ASN
3	B	250	GLN
3	H	706	ASN

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](#)

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

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, 95th 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 ²)	Q<0.9
1	C	14/14 (100%)	-0.39	0 100 100	25, 30, 45, 54	0
1	E	14/14 (100%)	-0.35	0 100 100	22, 42, 51, 54	0
1	I	14/14 (100%)	-0.28	0 100 100	31, 41, 52, 56	0
1	K	14/14 (100%)	-0.27	0 100 100	27, 42, 51, 53	0
2	D	10/10 (100%)	0.19	1 (10%) 8 8	24, 50, 72, 76	0
2	F	10/10 (100%)	-0.46	0 100 100	22, 29, 43, 50	0
2	J	10/10 (100%)	0.08	0 100 100	28, 55, 60, 61	0
2	L	10/10 (100%)	0.08	0 100 100	29, 50, 68, 69	0
3	A	149/163 (91%)	-0.03	6 (4%) 39 39	18, 28, 65, 83	0
3	B	149/163 (91%)	0.08	6 (4%) 39 39	20, 33, 70, 87	0
3	G	149/163 (91%)	0.16	8 (5%) 26 26	20, 36, 71, 83	0
3	H	149/163 (91%)	0.07	7 (4%) 32 32	18, 33, 73, 89	0
All	All	692/748 (92%)	0.03	28 (4%) 39 39	18, 34, 71, 89	0

The worst 5 of 28 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	B	317	GLU	6.2
3	G	650	ALA	4.5
3	B	351	VAL	4.3
3	G	617	GLU	4.1
3	H	817	GLU	3.9

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 [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å ²)	Q<0.9
4	CA	C	498	1/1	0.98	0.07	-1.82	34,34,34,34	0
4	CA	I	998	1/1	0.99	0.07	-2.68	41,41,41,41	0
4	CA	D	499	1/1	0.99	0.03	-2.98	36,36,36,36	0
4	CA	J	997	1/1	0.99	0.02	-4.13	38,38,38,38	0
4	CA	F	497	1/1	0.99	0.03	-4.38	37,37,37,37	0
4	CA	L	999	1/1	0.96	0.03	-4.87	40,40,40,40	0

6.5 Other polymers [i](#)

There are no such residues in this entry.