



# Full wwPDB X-ray Structure Validation Report ⓘ

May 14, 2020 – 12:36 pm BST

PDB ID : 2VS8  
Title : The crystal structure of I-DmoI in complex with DNA and Mn  
Authors : Marcaida, M.J.; Prieto, J.; Redondo, P.; Nadra, A.D.; Alibes, A.; Serrano, L.; Grizot, S.; Duchateau, P.; Paques, F.; Blanco, F.J.; Montoya, G.  
Deposited on : 2008-04-21  
Resolution : 2.10 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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

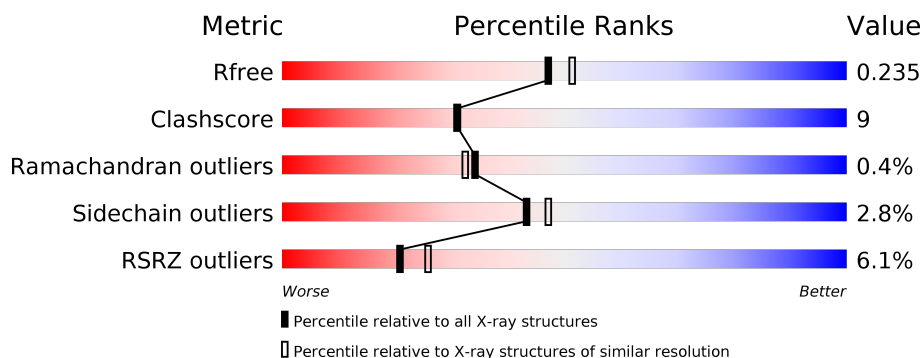
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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.





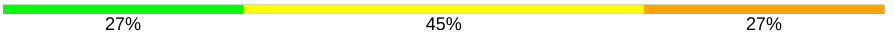


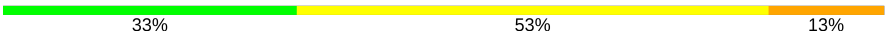



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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 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  $\leq 5\%$ . 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	A	200	<div> <div>7%</div> <div> <div></div> <div>71%</div> <div>19%</div> <div>•</div> <div>9%</div> </div> </div>
1	F	200	<div> <div>5%</div> <div> <div></div> <div>73%</div> <div>21%</div> <div>•</div> <div>5%</div> </div> </div>
1	K	200	<div> <div>9%</div> <div> <div></div> <div>71%</div> <div>17%</div> <div>•</div> <div>12%</div> </div> </div>
2	B	14	<div> <div></div> <div> <div>64%</div> <div>29%</div> <div>7%</div> </div> </div>
2	G	14	<div> <div></div> <div> <div>36%</div> <div>50%</div> <div>14%</div> </div> </div>
2	L	14	<div> <div></div> <div> <div>64%</div> <div>29%</div> <div>7%</div> </div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	C	11	
3	H	11	
3	M	11	
4	D	15	
4	I	15	
4	N	15	
5	E	10	
5	J	10	
5	O	10	

## 2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 8046 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 HOMING ENDONUCLEASE I-DMOI.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	183	Total	C	N	O	S	0	2	0
			1509	975	274	257	3			
1	F	191	Total	C	N	O	S	7	3	0
			1590	1027	289	271	3			
1	K	177	Total	C	N	O	S	0	2	0
			1474	953	268	250	3			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	14	Total	C	N	O	P	0	0	0
			285	136	53	83	13			
2	G	14	Total	C	N	O	P	0	0	0
			285	136	53	83	13			
2	L	14	Total	C	N	O	P	0	0	0
			285	136	53	83	13			

- Molecule 3 is a DNA chain called 5'-D(\*GP\*TP\*TP\*CP\*CP\*GP\*GP\*CP\* DGP\*DCP\*DGP)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	11	Total	C	N	O	P	0	0	0
			227	106	41	69	11			
3	H	11	Total	C	N	O	P	0	0	0
			227	106	41	69	11			
3	M	11	Total	C	N	O	P	0	0	0
			227	106	41	69	11			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	15	Total	C	N	O	P	0	0	0
			302	144	57	87	14			
4	I	15	Total	C	N	O	P	0	0	0
			302	144	57	87	14			
4	N	15	Total	C	N	O	P	0	0	0
			302	144	57	87	14			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	10	Total	C	N	O	P	0	0	0
			207	96	42	59	10			
5	J	10	Total	C	N	O	P	0	0	0
			207	96	42	59	10			
5	O	10	Total	C	N	O	P	0	0	0
			207	96	42	59	10			

- Molecule 6 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	2	Total	Mn	0	0
			2	2		
6	K	2	Total	Mn	0	0
			2	2		
6	F	2	Total	Mn	0	0
			2	2		

- Molecule 7 is ACETATE ION (three-letter code: ACT) (formula: C<sub>2</sub>H<sub>3</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	C	O	0	0
			4	2	2		
7	A	1	Total	C	O	0	0
			4	2	2		
7	F	1	Total	C	O	0	0
			4	2	2		

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	82	Total	O	0	0
			82	82		
8	B	19	Total	O	0	0
			19	19		
8	C	17	Total	O	0	0
			17	17		
8	D	19	Total	O	0	0
			19	19		
8	E	6	Total	O	0	0
			6	6		
8	F	81	Total	O	0	0
			81	81		
8	G	5	Total	O	0	0
			5	5		
8	H	16	Total	O	0	0
			16	16		
8	I	14	Total	O	0	0
			14	14		

*Continued on next page...*

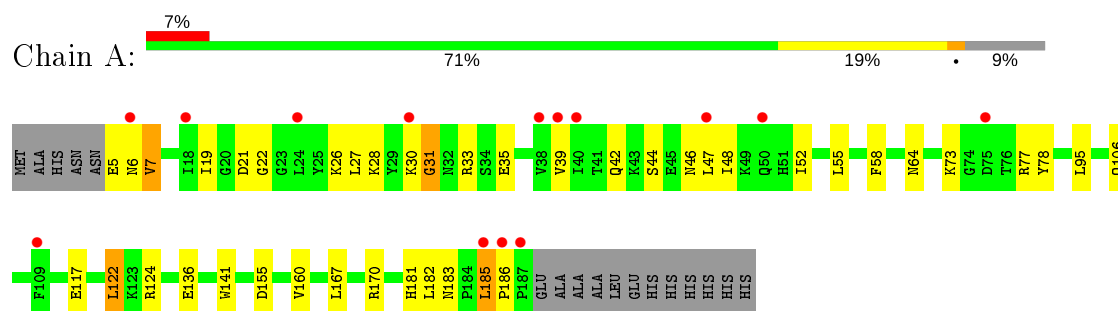
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	J	6	Total 6	O 6	0	0
8	K	77	Total 77	O 77	0	0
8	L	14	Total 14	O 14	0	0
8	M	17	Total 17	O 17	0	0
8	N	14	Total 14	O 14	0	0
8	O	5	Total 5	O 5	0	0

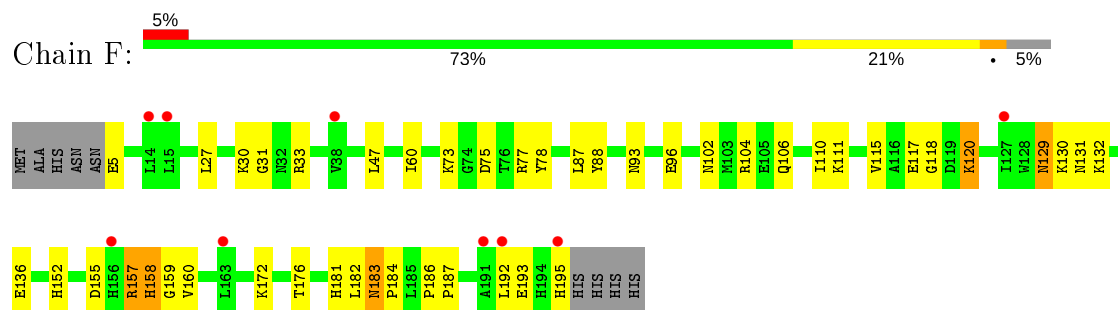
### 3 Residue-property plots [i](#)

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.

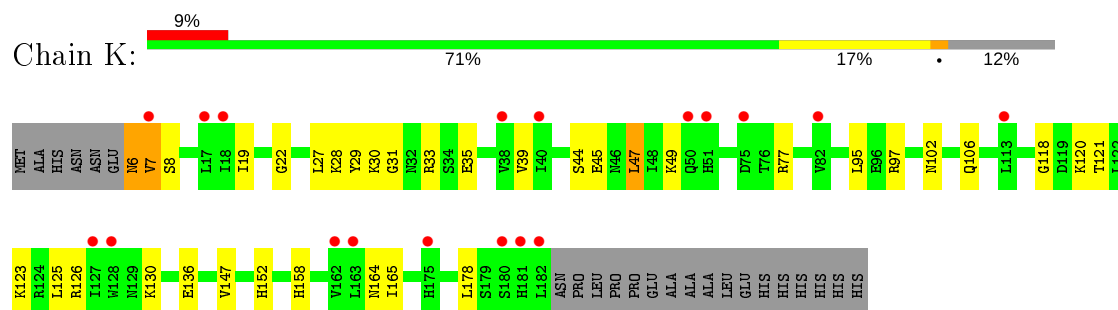
#### • Molecule 1: HOMING ENDONUCLEASE I-DMOI



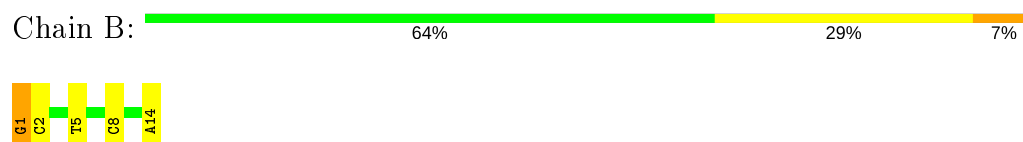
#### • Molecule 1: HOMING ENDONUCLEASE I-DMOI



#### • Molecule 1: HOMING ENDONUCLEASE I-DMOI



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





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

Chain G:  36% 50% 14%



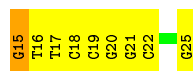
- Molecule 2: 5'-D(\*GP\*CP\*CP\*TP\*TP\*GP\*CP\*CP\*GP\*GP \*GP\*TP\*AP\*A)-3'

Chain L:  64% 29% 7%

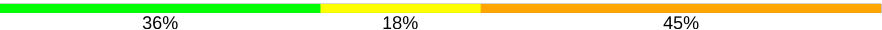


- Molecule 3: 5'-D(\*GP\*TP\*TP\*CP\*CP\*GP\*GP\*CP\* DGP\*DCP\*DGP)-3'

Chain C:  18% 73% 9%



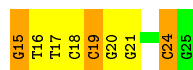
- Molecule 3: 5'-D(\*GP\*TP\*TP\*CP\*CP\*GP\*GP\*CP\* DGP\*DCP\*DGP)-3'

Chain H:  36% 18% 45%



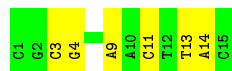
- Molecule 3: 5'-D(\*GP\*TP\*TP\*CP\*CP\*GP\*GP\*CP\* DGP\*DCP\*DGP)-3'

Chain M:  27% 45% 27%



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

Chain D:  60% 40%



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

Chain I:  53% 33% 13%



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

Chain N:  33% 53% 13%



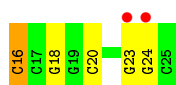
● Molecule 5: 5'-D(\*CP\*CP\*GP\*GP\*CP\*AP\*AP\* GP\*GP\*CP)-3'

Chain E:  70% 20% 10%



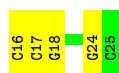
● Molecule 5: 5'-D(\*CP\*CP\*GP\*GP\*CP\*AP\*AP\* GP\*GP\*CP)-3'

Chain J:  20% 50% 40% 10%



● Molecule 5: 5'-D(\*CP\*CP\*GP\*GP\*CP\*AP\*AP\* GP\*GP\*CP)-3'

Chain O:  60% 40%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	107.13 Å   70.63 Å   106.95 Å 90.00°   119.85°   90.00°	Depositor
Resolution (Å)	25.76 – 2.10 25.76 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.7 (25.76-2.10) 99.7 (25.76-2.10)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.10 (at 2.10 Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.201 , 0.248 0.192 , 0.235	Depositor DCC
$R_{free}$ test set	4048 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	38.0	Xtriage
Anisotropy	0.823	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 47.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.020 for l,k,-h-l 0.020 for -h-l,k,h 0.017 for -h-l,-k,l 0.016 for h,-k,-h-l 0.016 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	8046	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.00% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

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

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	0.64	0/1541	0.73	0/2076
1	F	0.60	1/1631 (0.1%)	0.73	1/2197 (0.0%)
1	K	0.67	0/1506	0.75	1/2023 (0.0%)
2	B	1.00	0/319	1.77	6/491 (1.2%)
2	G	0.93	0/319	1.55	4/491 (0.8%)
2	L	0.94	1/319 (0.3%)	1.58	4/491 (0.8%)
3	C	1.31	1/253 (0.4%)	1.84	9/387 (2.3%)
3	H	1.40	1/253 (0.4%)	1.99	9/387 (2.3%)
3	M	1.19	1/253 (0.4%)	1.80	8/387 (2.1%)
4	D	1.11	1/338 (0.3%)	1.85	7/519 (1.3%)
4	I	1.22	0/338	2.21	17/519 (3.3%)
4	N	1.01	0/338	1.90	10/519 (1.9%)
5	E	1.12	1/232 (0.4%)	1.39	1/354 (0.3%)
5	J	1.09	1/232 (0.4%)	1.66	9/354 (2.5%)
5	O	1.21	1/232 (0.4%)	1.58	4/354 (1.1%)
All	All	0.88	9/8104 (0.1%)	1.32	90/11549 (0.8%)

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	A	0	3
1	F	0	4
1	K	0	3
All	All	0	10

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	15	DG	OP3-P	-12.90	1.45	1.61
3	H	15	DG	OP3-P	-11.88	1.46	1.61
3	M	15	DG	OP3-P	-11.51	1.47	1.61
5	O	16	DC	OP3-P	-10.34	1.48	1.61
5	E	16	DC	OP3-P	-10.15	1.49	1.61
5	J	16	DC	OP3-P	-8.50	1.50	1.61
4	D	4	DG	C3'-O3'	6.19	1.51	1.44
2	L	2	DC	C3'-O3'	-5.18	1.37	1.44
1	F	88	TYR	CD2-CE2	5.16	1.47	1.39

All (90) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	H	16	DT	O4'-C1'-N1	-12.88	98.99	108.00
4	N	9	DA	O4'-C1'-N9	12.51	116.75	108.00
4	D	9	DA	O4'-C1'-N9	12.06	116.44	108.00
3	C	16	DT	O4'-C1'-N1	-11.06	100.26	108.00
4	D	14	DA	O4'-C1'-N9	-10.72	100.49	108.00
3	C	21	DG	O4'-C1'-N9	9.72	114.80	108.00
3	H	17	DT	C6-C5-C7	-9.67	117.10	122.90
2	G	12	DT	O4'-C1'-N1	-9.52	101.33	108.00
2	B	5	DT	C5-C4-O4	-9.01	118.60	124.90
3	H	17	DT	C4-C5-C7	8.99	124.39	119.00
4	I	14	DA	C5-N7-C8	-8.97	99.41	103.90
2	B	1	DG	O4'-C1'-N9	-8.77	101.86	108.00
2	B	5	DT	N3-C4-O4	8.75	125.15	119.90
4	I	9	DA	O4'-C1'-N9	8.26	113.78	108.00
5	E	18	DG	O4'-C1'-N9	8.09	113.67	108.00
2	G	2	DC	O4'-C1'-N1	8.09	113.66	108.00
4	I	14	DA	N9-C1'-C2'	7.72	127.26	112.60
3	C	25	DG	C5-C6-O6	-7.64	124.01	128.60
4	I	11	DC	O4'-C1'-N1	7.59	113.31	108.00
3	H	20	DG	O4'-C1'-N9	7.58	113.30	108.00
2	L	4	DT	O4'-C1'-N1	-7.57	102.70	108.00
4	I	14	DA	N7-C8-N9	7.55	117.58	113.80
4	D	4	DG	O4'-C1'-N9	-7.55	102.72	108.00
5	O	18	DG	O4'-C1'-N9	7.54	113.28	108.00
5	J	16	DC	N3-C2-O2	-7.38	116.73	121.90
3	M	24	DC	O4'-C1'-N1	7.35	113.14	108.00
3	H	19	DC	O4'-C1'-N1	7.19	113.03	108.00
5	J	16	DC	C4-C5-C6	7.13	120.96	117.40
4	I	14	DA	C2-N3-C4	-7.12	107.04	110.60
4	I	14	DA	C4-C5-N7	7.03	114.21	110.70

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	J	20	DC	O4'-C1'-N1	7.01	112.91	108.00
4	I	2	DG	O5'-P-OP2	-6.93	99.46	105.70
3	M	21	DG	O4'-C1'-N9	6.75	112.73	108.00
3	H	22	DC	O4'-C1'-N1	6.75	112.72	108.00
4	D	13	DT	P-O3'-C3'	6.68	127.71	119.70
3	C	17	DT	C4-C5-C7	6.64	122.99	119.00
4	I	7	DG	P-O3'-C3'	6.64	127.67	119.70
4	I	14	DA	C3'-C2'-C1'	-6.60	94.58	102.50
2	B	8	DC	N1-C2-O2	-6.57	114.96	118.90
4	I	9	DA	P-O3'-C3'	6.57	127.58	119.70
4	D	3	DC	O4'-C1'-N1	6.53	112.57	108.00
1	F	27	LEU	CA-CB-CG	6.45	130.12	115.30
3	M	19	DC	O4'-C1'-N1	6.42	112.50	108.00
4	D	3	DC	C2-N3-C4	6.39	123.09	119.90
3	M	18	DC	O4'-C1'-N1	6.36	112.45	108.00
4	I	3	DC	O5'-P-OP1	-6.35	99.98	105.70
4	I	2	DG	P-O3'-C3'	6.25	127.20	119.70
3	M	16	DT	O4'-C1'-N1	-6.19	103.67	108.00
5	J	16	DC	C5-C6-N1	-6.14	117.93	121.00
5	J	24	DG	O4'-C1'-N9	6.14	112.30	108.00
5	O	18	DG	C1'-O4'-C4'	-5.95	104.15	110.10
4	I	2	DG	O4'-C1'-N9	-5.92	103.86	108.00
3	H	21	DG	C1'-O4'-C4'	-5.90	104.20	110.10
2	L	12	DT	O4'-C1'-N1	-5.88	103.88	108.00
3	M	17	DT	C4-C5-C7	5.84	122.50	119.00
4	I	7	DG	C8-N9-C4	-5.77	104.09	106.40
4	N	11	DC	O4'-C1'-N1	5.75	112.03	108.00
5	O	24	DG	O4'-C1'-N9	5.74	112.02	108.00
3	C	20	DG	O4'-C1'-N9	5.73	112.01	108.00
3	H	21	DG	C4'-C3'-C2'	-5.72	97.95	103.10
3	C	19	DC	O4'-C1'-N1	5.69	111.98	108.00
4	N	9	DA	C1'-O4'-C4'	-5.65	104.45	110.10
2	G	9	DG	P-O3'-C3'	5.63	126.46	119.70
5	J	16	DC	N1-C2-N3	5.62	123.14	119.20
2	G	13	DA	P-O3'-C3'	5.60	126.42	119.70
5	J	23	DG	O4'-C1'-N9	5.57	111.90	108.00
4	I	7	DG	C5-C6-O6	-5.55	125.27	128.60
4	N	5	DC	O4'-C1'-N1	5.54	111.88	108.00
4	N	6	DC	O4'-C1'-N1	-5.52	104.14	108.00
4	N	4	DG	C5-C6-N1	5.50	114.25	111.50
3	C	17	DT	C6-C5-C7	-5.49	119.61	122.90
4	N	3	DC	C2-N3-C4	5.46	122.63	119.90

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	L	2	DC	O4'-C1'-N1	5.42	111.79	108.00
3	C	21	DG	O4'-C1'-C2'	-5.42	101.57	105.90
4	N	9	DA	C3'-C2'-C1'	-5.41	96.01	102.50
4	I	14	DA	C6-C5-N7	-5.41	128.52	132.30
3	M	24	DC	N3-C4-N4	5.34	121.74	118.00
2	B	2	DC	P-O3'-C3'	5.32	126.08	119.70
3	M	17	DT	C6-C5-C7	-5.29	119.73	122.90
4	D	11	DC	O4'-C1'-N1	5.27	111.69	108.00
5	J	23	DG	O4'-C1'-C2'	-5.26	101.69	105.90
4	N	11	DC	C1'-O4'-C4'	-5.22	104.88	110.10
4	N	4	DG	P-O3'-C3'	5.18	125.92	119.70
3	C	18	DC	O4'-C1'-N1	5.18	111.62	108.00
5	J	18	DG	C3'-C2'-C1'	-5.12	96.35	102.50
2	B	8	DC	P-O3'-C3'	5.09	125.81	119.70
1	K	27	LEU	CA-CB-CG	5.09	127.00	115.30
3	H	16	DT	O4'-C1'-C2'	-5.03	101.88	105.90
2	L	4	DT	C5-C4-O4	-5.02	121.38	124.90
5	O	17	DC	P-O3'-C3'	5.01	125.71	119.70

There are no chirality outliers.

All (10) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	185	LEU	Peptide
1	A	30	LYS	Peptide
1	A	31	GLY	Peptide
1	F	157	ARG	Peptide
1	F	158	HIS	Peptide
1	F	30	LYS	Peptide
1	F	31	GLY	Peptide
1	K	30	LYS	Peptide
1	K	31	GLY	Peptide
1	K	6	ASN	Peptide

## 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	A	1509	0	1578	39	0
1	F	1590	0	1663	36	0
1	K	1474	0	1553	28	0
2	B	285	0	159	3	0
2	G	285	0	159	7	0
2	L	285	0	159	4	0
3	C	227	0	124	2	0
3	H	227	0	124	6	0
3	M	227	0	124	5	0
4	D	302	0	169	0	0
4	I	302	0	169	6	0
4	N	302	0	169	5	0
5	E	207	0	111	3	0
5	J	207	0	111	2	0
5	O	207	0	111	0	0
6	A	2	0	0	0	0
6	F	2	0	0	0	0
6	K	2	0	0	0	0
7	A	8	0	6	0	0
7	F	4	0	3	1	0
8	A	82	0	0	6	0
8	B	19	0	0	1	0
8	C	17	0	0	0	0
8	D	19	0	0	0	0
8	E	6	0	0	1	0
8	F	81	0	0	1	0
8	G	5	0	0	0	0
8	H	16	0	0	1	0
8	I	14	0	0	1	0
8	J	6	0	0	2	0
8	K	77	0	0	3	0
8	L	14	0	0	0	0
8	M	17	0	0	1	0
8	N	14	0	0	1	0
8	O	5	0	0	0	0
All	All	8046	0	6492	128	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 9.

All (128) 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:K:6:ASN:HA	1:K:7:VAL:HB	1.09	1.06
1:K:6:ASN:CA	1:K:7:VAL:HB	1.86	1.04
2:G:10:DG:H1	5:J:16:DC:H5	1.03	0.98
1:K:6:ASN:HA	1:K:7:VAL:CB	1.93	0.96
2:B:1:DG:H1	5:E:25:DC:H5	1.04	0.95
1:F:129:ASN:HD22	1:F:131:ASN:H	1.15	0.91
8:H:2003:HOH:O	4:I:14:DA:H8	1.52	0.90
2:L:14:DA:O3'	3:M:15:DG:OP3	1.92	0.88
1:A:136[A]:GLU:HG3	8:A:2050:HOH:O	1.82	0.80
2:G:14:DA:O3'	3:H:15:DG:OP3	2.01	0.79
2:G:12:DT:H3	4:I:14:DA:H62	1.31	0.79
1:A:26:LYS:HE2	1:A:28:LYS:NZ	1.97	0.78
1:F:5:GLU:N	1:F:104:ARG:NH1	2.35	0.75
5:J:16:DC:OP2	8:J:2002:HOH:O	2.04	0.74
4:I:2:DG:OP2	8:I:2003:HOH:O	2.04	0.74
1:F:130:LYS:HD2	1:F:158:HIS:O	1.89	0.72
1:K:35:GLU:HG3	8:K:2031:HOH:O	1.88	0.71
1:K:102:ASN:O	1:K:106:GLN:HG3	1.91	0.70
1:A:47:LEU:O	1:A:52:ILE:HD12	1.91	0.70
1:A:46:ASN:O	1:A:47:LEU:HB3	1.91	0.69
2:L:14:DA:HO3'	3:M:15:DG:P	2.15	0.69
1:A:48:ILE:CD1	1:A:78:TYR:HB2	2.25	0.67
2:B:1:DG:N1	5:E:25:DC:H5	1.85	0.67
1:A:26:LYS:HE2	1:A:28:LYS:HZ3	1.60	0.66
1:F:181:HIS:HE1	1:F:193:GLU:OE1	1.78	0.66
3:M:24:DC:OP2	8:M:2010:HOH:O	2.13	0.66
1:K:130:LYS:HG3	4:N:15:DC:OP2	1.97	0.64
1:F:186:PRO:HB2	7:F:1198:ACT:H2	1.80	0.64
1:A:21:ASP:OD1	8:A:2002:HOH:O	2.15	0.64
3:C:22:DC:P	1:K:97:ARG:HH22	2.21	0.63
1:F:5:GLU:N	1:F:104:ARG:HH12	1.95	0.62
4:N:7:DG:H2''	4:N:8:DG:H5''	1.82	0.62
1:A:42:GLN:HB3	1:A:48:ILE:HD11	1.80	0.62
1:F:155:ASP:O	1:F:159:GLY:HA2	2.00	0.62
1:K:6:ASN:CA	1:K:7:VAL:CB	2.64	0.61
2:G:14:DA:HO3'	3:H:15:DG:P	2.23	0.61
5:E:16:DC:OP1	8:E:2002:HOH:O	2.16	0.60
1:F:158:HIS:CE1	4:I:14:DA:H2	2.21	0.58
1:F:73:LYS:HD2	1:F:78:TYR:OH	2.02	0.58
1:K:118:GLY:O	1:K:120:LYS:HE3	2.04	0.57
1:F:93:ASN:O	1:F:96:GLU:HB2	2.04	0.57
4:N:5:DC:OP2	8:N:2007:HOH:O	2.18	0.56

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:155:ASP:HB3	1:A:160:VAL:HB	1.88	0.56
1:A:26:LYS:HE2	1:A:28:LYS:HZ1	1.71	0.56
1:F:73:LYS:HD2	1:F:78:TYR:CZ	2.41	0.55
1:K:147:VAL:CG1	1:K:165:ILE:HG23	2.37	0.55
1:F:117:GLU:OE1	8:J:2002:HOH:O	2.18	0.55
1:F:106:GLN:O	1:F:110:ILE:HG13	2.06	0.55
1:F:183:ASN:HB2	1:F:184:PRO:CD	2.37	0.54
1:F:157:ARG:HB2	1:F:158:HIS:CD2	2.42	0.54
1:A:48:ILE:HD13	1:A:78:TYR:HB2	1.90	0.54
1:A:22:GLY:HA3	1:A:39:VAL:O	2.08	0.54
1:A:44:SER:O	1:A:48:ILE:HD12	2.08	0.53
1:F:183:ASN:HB2	1:F:184:PRO:HD2	1.89	0.53
1:A:47:LEU:O	1:A:52:ILE:CD1	2.56	0.53
1:A:46:ASN:O	1:A:47:LEU:CB	2.57	0.52
1:A:7:VAL:HG23	1:A:58:PHE:CE2	2.43	0.52
1:F:33:ARG:NH1	3:H:21:DG:O6	2.42	0.52
1:F:172:LYS:O	1:F:176:THR:HB	2.09	0.51
1:K:6:ASN:N	1:K:7:VAL:HB	2.26	0.51
1:F:60[B]:ILE:HD11	1:F:87:LEU:HD22	1.93	0.51
1:A:55:LEU:HD23	1:A:185:LEU:HD12	1.94	0.50
1:A:27:LEU:O	1:A:28:LYS:HD3	2.11	0.50
1:A:167:LEU:HD12	1:A:170:ARG:CZ	2.41	0.50
2:G:12:DT:H3	4:I:14:DA:N6	2.03	0.50
1:F:93:ASN:HD21	4:N:1:DC:N4	2.10	0.49
1:K:7:VAL:O	1:K:7:VAL:HG12	2.10	0.49
1:K:6:ASN:C	1:K:8:SER:H	2.15	0.49
1:A:167:LEU:HD12	1:A:170:ARG:NH2	2.27	0.49
1:F:102:ASN:O	1:F:106:GLN:HG3	2.12	0.49
1:K:152[B]:HIS:HD2	8:K:2061:HOH:O	1.95	0.49
1:A:185:LEU:HG	1:A:186:PRO:CD	2.43	0.49
3:H:19:DC:H2''	3:H:20:DG:H5'	1.94	0.49
1:K:118:GLY:HA2	1:K:126:ARG:O	2.12	0.49
1:F:155:ASP:OD1	1:F:157:ARG:HB2	2.14	0.48
1:A:27:LEU:HD22	1:A:27:LEU:N	2.28	0.48
1:A:167:LEU:CD1	1:A:170:ARG:CZ	2.92	0.48
1:F:158:HIS:HB3	1:F:160:VAL:HG23	1.95	0.47
1:A:167:LEU:HD23	8:B:2005:HOH:O	2.13	0.47
1:A:182:LEU:O	1:A:183:ASN:OD1	2.33	0.47
3:H:21:DG:H2''	3:H:22:DC:H5'	1.97	0.47
1:F:181:HIS:CE1	1:F:193:GLU:OE1	2.64	0.47
1:K:44:SER:OG	1:K:47:LEU:HB2	2.15	0.47

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:33:ARG:HD2	4:I:4:DG:O6	2.15	0.46
1:K:126:ARG:HG2	1:K:164:ASN:HD22	1.80	0.46
1:K:121:THR:HB	1:K:123:LYS:HE3	1.97	0.46
1:F:111:LYS:HE2	1:F:186:PRO:HD3	1.97	0.46
1:K:130:LYS:HB3	1:K:130:LYS:HE2	1.72	0.45
2:B:14:DA:O3'	3:C:15:DG:OP3	2.32	0.45
1:K:136[A]:GLU:OE2	8:K:2056:HOH:O	2.20	0.45
1:A:167:LEU:CD1	1:A:170:ARG:NH2	2.79	0.45
1:A:181:HIS:CE1	8:A:2080:HOH:O	2.69	0.45
1:A:19:ILE:HD11	1:A:95:LEU:HD22	1.99	0.45
1:F:152[B]:HIS:CD2	8:F:2072:HOH:O	2.68	0.45
1:A:44:SER:O	1:A:48:ILE:CD1	2.65	0.45
1:K:130:LYS:HE3	1:K:158:HIS:O	2.17	0.45
2:L:14:DA:O3'	3:M:15:DG:P	2.71	0.45
1:A:35:GLU:HG3	8:A:2025:HOH:O	2.16	0.44
1:A:106:GLN:HG2	1:A:141:TRP:CD2	2.53	0.44
2:L:12:DT:H2''	2:L:13:DA:C8	2.52	0.44
1:A:5:GLU:HA	1:A:6:ASN:HA	1.51	0.44
1:K:29:TYR:HB2	1:K:33:ARG:HB2	2.00	0.44
1:K:7:VAL:O	1:K:7:VAL:CG1	2.66	0.44
1:A:48:ILE:HD13	1:A:78:TYR:CB	2.48	0.43
3:H:19:DC:H2'	3:H:20:DG:C8	2.53	0.43
1:A:42:GLN:HB3	1:A:48:ILE:CD1	2.48	0.43
1:A:106:GLN:NE2	8:A:2042:HOH:O	2.52	0.43
1:F:47:LEU:HD11	1:F:187:PRO:CG	2.48	0.43
1:K:19:ILE:HD11	1:K:95:LEU:HD22	2.01	0.43
1:A:124:ARG:NH1	8:A:2045:HOH:O	2.26	0.43
3:M:19:DC:H2'	3:M:20:DG:C8	2.54	0.43
1:F:192:LEU:O	1:F:195:HIS:HD2	2.02	0.42
1:K:45:GLU:HG2	1:K:49:LYS:HD2	2.01	0.42
1:A:73:LYS:HG3	1:A:78:TYR:CE1	2.54	0.42
1:A:31:GLY:HA2	1:A:33:ARG:H	1.85	0.42
1:A:106:GLN:HG2	1:A:141:TRP:CE2	2.55	0.41
1:F:158:HIS:H	1:F:159:GLY:HA2	1.84	0.41
1:F:104:ARG:HG2	1:F:182:LEU:HD13	2.02	0.41
1:F:118:GLY:O	1:F:120:LYS:HE3	2.21	0.41
1:K:33:ARG:HD2	4:N:4:DG:O6	2.20	0.41
1:F:47:LEU:HD11	1:F:187:PRO:HG3	2.02	0.41
1:K:125:LEU:HB2	1:K:165:ILE:HB	2.03	0.41
1:K:22:GLY:HA3	1:K:39:VAL:O	2.21	0.41
1:F:115:VAL:CG2	1:F:186:PRO:HB3	2.51	0.40

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:G:3:DC:H2''	2:G:4:DT:O5'	2.21	0.40
1:F:129:ASN:ND2	1:F:131:ASN:H	1.98	0.40
1:F:132:LYS:O	1:F:136:GLU:HG2	2.21	0.40
2:G:8:DC:H2''	2:G:9:DG:C8	2.56	0.40

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	182/200 (91%)	173 (95%)	8 (4%)	1 (0%)	29	26
1	F	192/200 (96%)	182 (95%)	10 (5%)	0	100	100
1	K	177/200 (88%)	170 (96%)	6 (3%)	1 (1%)	25	21
All	All	551/600 (92%)	525 (95%)	24 (4%)	2 (0%)	34	32

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	K	7	VAL
1	A	122	LEU

### 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	165/179 (92%)	160 (97%)	5 (3%)	41	44
1	F	174/179 (97%)	169 (97%)	5 (3%)	42	46
1	K	161/179 (90%)	157 (98%)	4 (2%)	47	52
All	All	500/537 (93%)	486 (97%)	14 (3%)	43	47

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	7	VAL
1	A	64	ASN
1	A	77	ARG
1	A	117	GLU
1	A	122	LEU
1	F	75	ASP
1	F	77	ARG
1	F	120	LYS
1	F	129	ASN
1	F	183	ASN
1	K	28	LYS
1	K	47	LEU
1	K	77	ARG
1	K	178	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	93	ASN
1	A	106	GLN
1	A	158	HIS
1	F	50	GLN
1	F	93	ASN
1	F	106	GLN
1	F	129	ASN
1	F	158	HIS
1	F	181	HIS
1	F	195	HIS
1	K	106	GLN
1	K	164	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 9 ligands modelled in this entry, 6 are monoatomic - leaving 3 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
7	ACT	A	1190	-	1,3,3	1.04	0	0,3,3	0.00	-
7	ACT	A	1191	-	1,3,3	0.88	0	0,3,3	0.00	-
7	ACT	F	1198	-	1,3,3	2.52	1 (100%)	0,3,3	0.00	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	F	1198	ACT	CH3-C	2.52	1.51	1.48

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	F	1198	ACT	1	0

## 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

### 6.1 Protein, DNA and RNA chains

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	A	183/200 (91%)	0.49	14 (7%) 13 17	28, 40, 65, 96	1 (0%)
1	F	191/200 (95%)	0.34	9 (4%) 31 37	26, 41, 59, 66	0
1	K	177/200 (88%)	0.39	18 (10%) 6 8	27, 37, 59, 81	0
2	B	14/14 (100%)	-0.60	0 100 100	39, 46, 58, 59	0
2	G	14/14 (100%)	-0.28	0 100 100	34, 63, 85, 86	0
2	L	14/14 (100%)	-0.48	0 100 100	41, 49, 59, 60	0
3	C	11/11 (100%)	-0.55	0 100 100	38, 47, 69, 71	0
3	H	11/11 (100%)	-0.66	0 100 100	35, 45, 55, 63	0
3	M	11/11 (100%)	-0.52	0 100 100	35, 47, 63, 64	0
4	D	15/15 (100%)	-0.39	0 100 100	34, 47, 74, 76	0
4	I	15/15 (100%)	-0.60	0 100 100	35, 43, 58, 63	0
4	N	15/15 (100%)	-0.31	0 100 100	34, 46, 67, 70	0
5	E	10/10 (100%)	-0.16	0 100 100	42, 62, 68, 73	0
5	J	10/10 (100%)	1.02	2 (20%) 1 1	44, 87, 109, 119	0
5	O	10/10 (100%)	0.23	0 100 100	39, 64, 77, 77	0
All	All	701/750 (93%)	0.25	43 (6%) 21 26	26, 42, 68, 119	1 (0%)

All (43) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	187	PRO	9.9
1	K	182	LEU	7.4
1	K	181	HIS	4.8
1	F	156	HIS	4.2
1	K	180	SER	4.2
1	A	185	LEU	3.8
1	A	47	LEU	3.8

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	F	127	ILE	3.6
1	K	51	HIS	3.5
1	F	191	ALA	3.5
1	A	6	ASN	3.5
1	A	75	ASP	3.4
1	A	30	LYS	3.3
1	K	127	ILE	3.2
1	K	50	GLN	3.2
1	A	39	VAL	3.2
1	F	192	LEU	3.0
1	A	18	ILE	2.9
1	A	186	PRO	2.9
1	K	75	ASP	2.9
1	K	18	ILE	2.7
1	A	50	GLN	2.7
1	A	38	VAL	2.7
1	F	14	LEU	2.6
5	J	24	DG	2.6
1	K	113	LEU	2.5
1	K	175	HIS	2.5
1	K	38	VAL	2.5
1	F	163	LEU	2.4
1	K	17	LEU	2.4
1	F	15	LEU	2.4
1	K	7	VAL	2.3
1	F	195	HIS	2.3
1	F	38	VAL	2.2
1	K	40	ILE	2.2
1	K	128	TRP	2.2
1	A	40	ILE	2.2
5	J	23	DG	2.2
1	A	109	PHE	2.1
1	K	163	LEU	2.1
1	K	162	VAL	2.1
1	K	82	VAL	2.1
1	A	24	LEU	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	B-factors( $\text{\AA}^2$ )	Q<0.9
7	ACT	F	1198	4/4	0.85	0.22	51,51,51,51	0
7	ACT	A	1190	4/4	0.93	0.13	49,49,49,50	0
7	ACT	A	1191	4/4	0.94	0.10	59,59,59,59	0
6	MN	K	1184	1/1	0.98	0.11	39,39,39,39	0
6	MN	A	1189	1/1	0.98	0.10	40,40,40,40	0
6	MN	F	1196	1/1	0.99	0.11	38,38,38,38	0
6	MN	K	1183	1/1	0.99	0.14	40,40,40,40	0
6	MN	A	1188	1/1	1.00	0.11	40,40,40,40	0
6	MN	F	1197	1/1	1.00	0.12	38,38,38,38	0

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

There are no such residues in this entry.