



# wwPDB X-ray Structure Validation Summary Report

Feb 27, 2014 – 06:02 AM GMT

PDB ID : 3ERC  
Title : Crystal structure of the heterodimeric vaccinia virus mRNA polyadenylate polymerase with three fragments of RNA and 3'-deoxy ATP  
Authors : Li, C.; Li, H; Zhou, S.; Poulos, T.L.; Gershon, P.D.  
Deposited on : 2008-10-01  
Resolution : 3.21 Å(reported)

This is a wwPDB validation summary 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/ValidationPDFNotes.html>

---

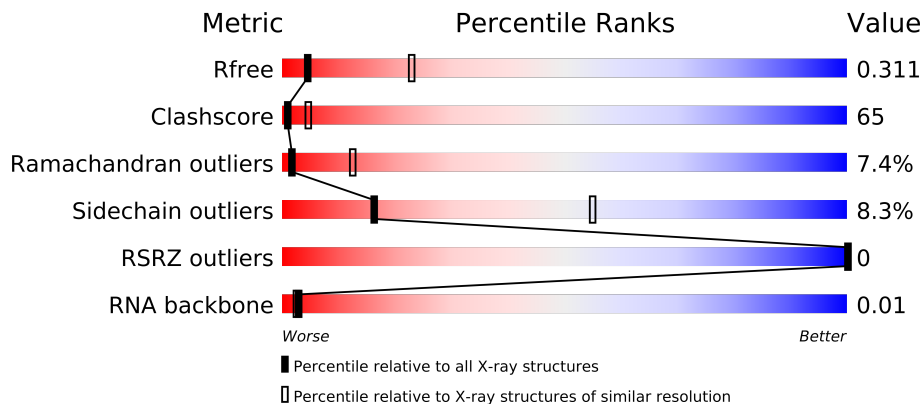
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 3.21 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))
$R_{free}$	66092	1205 (3.30-3.14)
Clashscore	79885	1072 (3.28-3.16)
Ramachandran outliers	78287	1052 (3.28-3.16)
Sidechain outliers	78261	1051 (3.28-3.16)
RSRZ outliers	66119	1206 (3.30-3.14)
RNA backbone	1838	1004 (3.74-2.70)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	297	
1	B	297	
2	C	479	
2	D	479	
3	E	5	
3	F	5	
4	G	4	
5	I	5	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
6	U	D	602	-	X

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 12480 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Cap-specific mRNA (nucleoside-2'-O-)-methyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	288	Total	C	N	O	S	0	0	0
			2382	1552	394	424	12			
1	B	288	Total	C	N	O	S	0	0	0
			2382	1552	394	424	12			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	140	ALA	ARG	ENGINEERED	UNP P07617
A	142	ALA	LYS	ENGINEERED	UNP P07617
A	143	ALA	ARG	ENGINEERED	UNP P07617
B	140	ALA	ARG	ENGINEERED	UNP P07617
B	142	ALA	LYS	ENGINEERED	UNP P07617
B	143	ALA	ARG	ENGINEERED	UNP P07617

- Molecule 2 is a protein called Poly(A) polymerase catalytic subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	445	Total	C	N	O	S	0	0	0
			3627	2318	605	680	24			
2	D	445	Total	C	N	O	S	0	0	0
			3627	2318	605	680	24			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	36	SER	LEU	ENGINEERED	UNP P23371
D	36	SER	LEU	ENGINEERED	UNP P23371

- Molecule 3 is a RNA chain called RNA/DNA chimera (5'-D(CP\*)R(UP\*UP\*)D(CP\*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	5	Total	C	N	O	P	0	0	0
			94	45	13	32	4			
3	F	5	Total	C	N	O	P	0	0	0
			94	45	13	32	4			

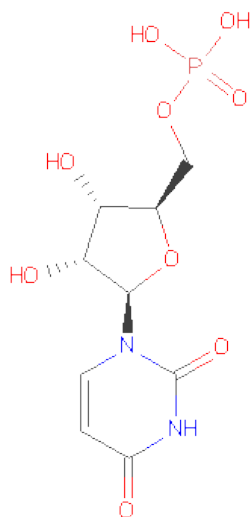
- Molecule 4 is a RNA chain called RNA/DNA chimera (5'-D(CP\*)R(UP\*UP\*)-D(C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	G	4	Total	C	N	O	P	0	0	0
			75	36	10	26	3			

- Molecule 5 is a RNA chain called RNA/DNA chimera (5'-D(CP\*CP\*)R(UP\*UP\*)D(C)-3').

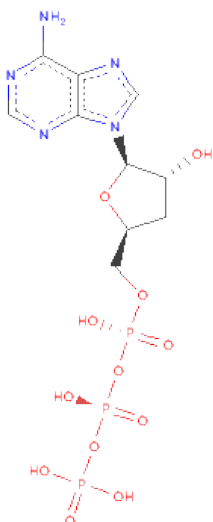
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	I	5	Total	C	N	O	P	0	0	0
			94	45	13	32	4			

- Molecule 6 is URIDINE-5'-MONOPHOSPHATE (three-letter code: U) (formula: C<sub>9</sub>H<sub>13</sub>N<sub>2</sub>O<sub>9</sub>P).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	C	1	Total	C	N	O	P	0	0
			20	9	2	8	1		
6	D	1	Total	C	N	O	P	0	0
			20	9	2	8	1		

- Molecule 7 is 3'-DEOXYADENOSINE-5'-TRIPHOSPHATE (three-letter code: 3AT) (formula: C<sub>10</sub>H<sub>16</sub>N<sub>5</sub>O<sub>12</sub>P<sub>3</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
7	D	1	Total	C	N	O	P	0	0
			30	10	5	12	3		
7	C	1	Total	C	N	O	P	0	0
			30	10	5	12	3		

- Molecule 8 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	D	2	Total	Ca	0	0
			2	2		
8	C	2	Total	Ca	0	0
			2	2		

- Molecule 9 is water.

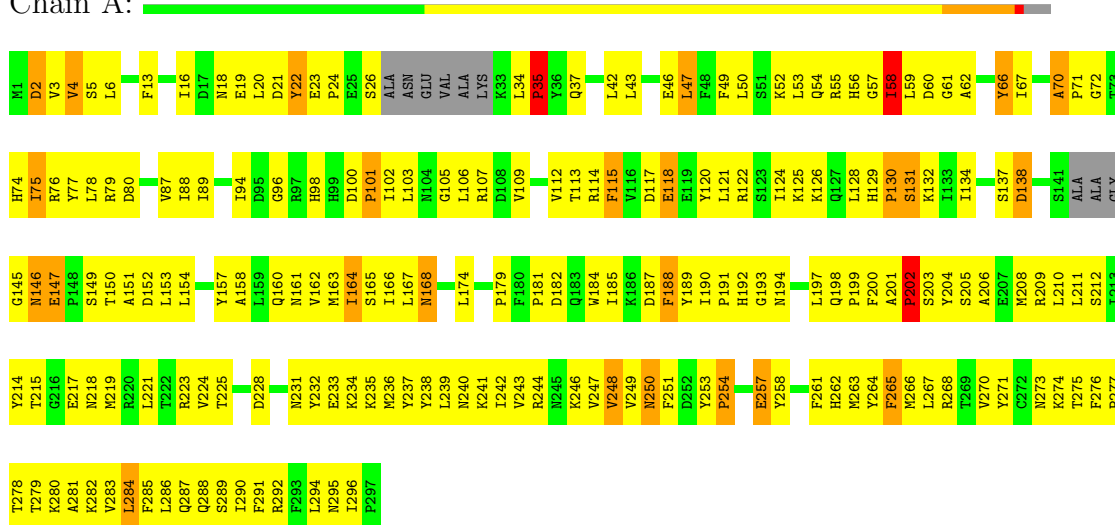
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	C	1	Total	O	0	0
			1	1		

### 3 Residue-property plots

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.

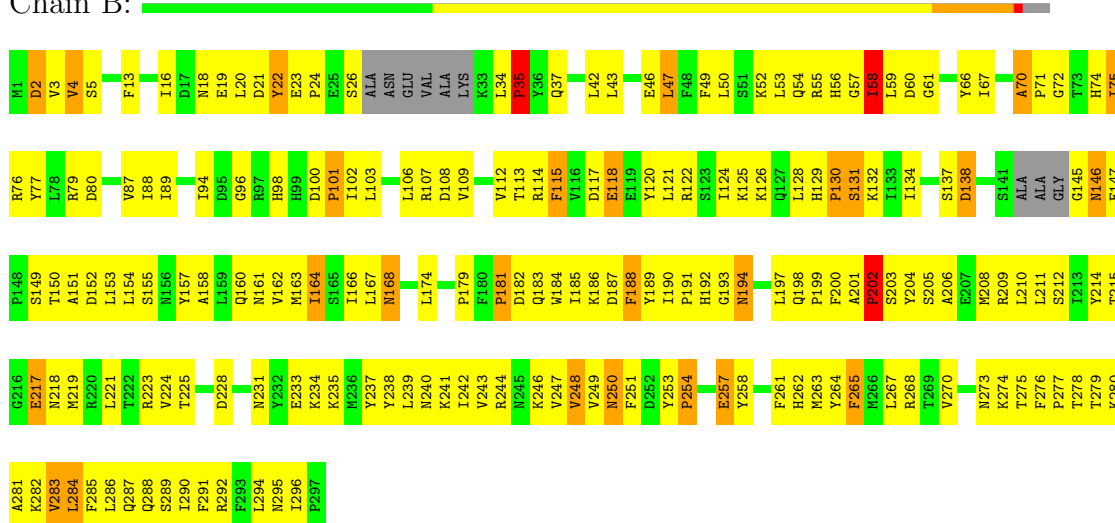
- Molecule 1: Cap-specific mRNA (nucleoside-2'-O-)-methyltransferase

Chain A:



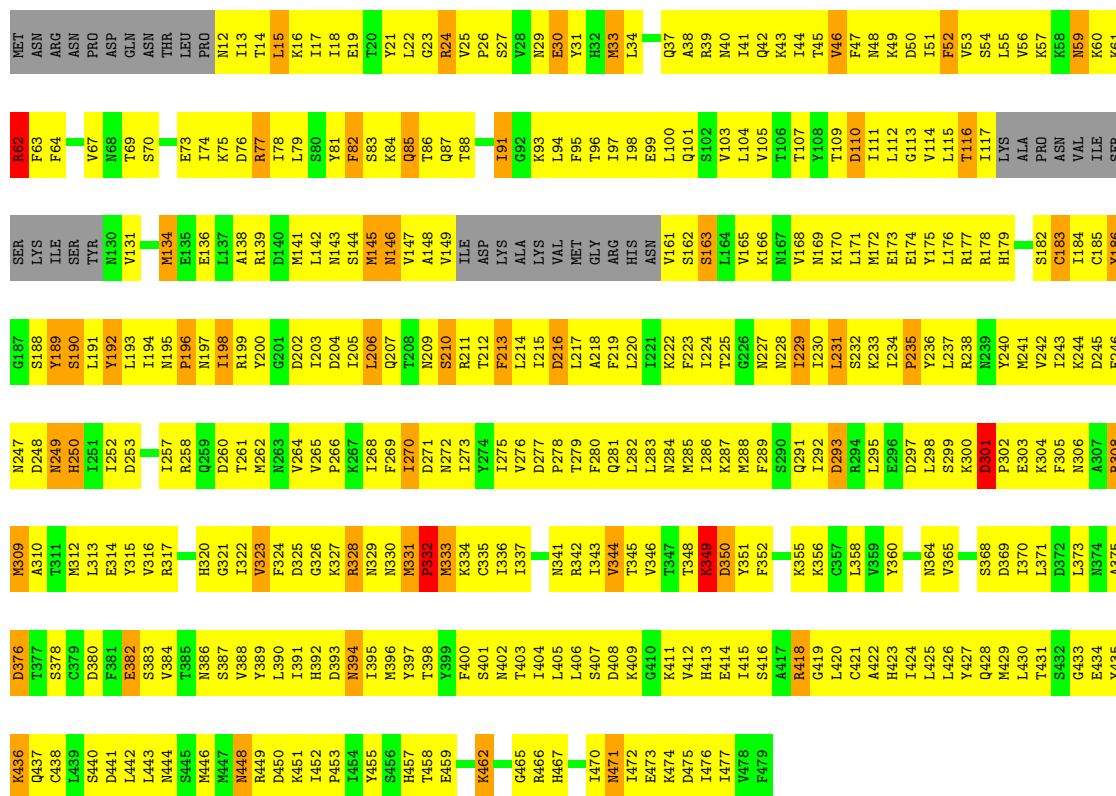
- Molecule 1: Cap-specific mRNA (nucleoside-2'-O-)-methyltransferase

Chain B:



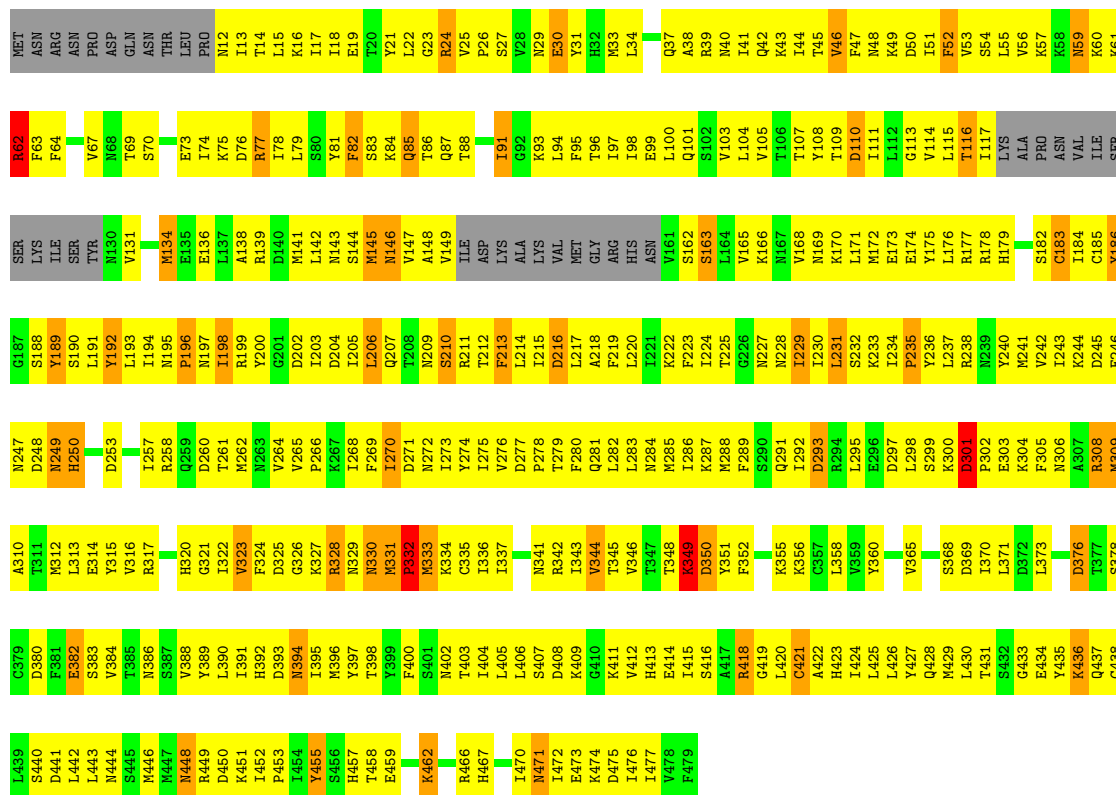
- Molecule 2: Poly(A) polymerase catalytic subunit

Chain C:



● Molecule 2: Poly(A) polymerase catalytic subunit

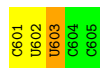
Chain D:





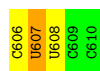
- Molecule 3: RNA/DNA chimera (5'-D(CP\*)R(UP\*UP\*)D(CP\*C)-3')

Chain E: 



- Molecule 3: RNA/DNA chimera (5'-D(CP\*)R(UP\*UP\*)D(CP\*C)-3')

Chain F: 



- Molecule 4: RNA/DNA chimera (5'-D(CP\*)R(UP\*UP\*)-D(C)-3')

Chain G: 



- Molecule 5: RNA/DNA chimera (5'-D(CP\*CP\*)R(UP\*UP\*)D(C)-3')

Chain I: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	69.92Å 77.15Å 108.03Å 89.52° 73.45° 63.76°	Depositor
Resolution (Å)	38.55 – 3.21 38.55 – 3.16	Depositor EDS
% Data completeness (in resolution range)	93.8 (38.55-3.21) 91.4 (38.55-3.16)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	0.11	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.13 (at 3.18Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, $R_{free}$	0.245 , 0.309 0.250 , 0.311	Depositor DCC
$R_{free}$ test set	1453 reflections (5.05%)	DCC
Wilson B-factor (Å <sup>2</sup> )	78.7	Xtriage
Anisotropy	0.318	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 50.2	EDS
Estimated twinning fraction	0.056 for h,h-k,h-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 30209 reflections	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	12480	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	84.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.62% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CA, 3AT

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	0.48	0/2444	0.69	0/3308
1	B	0.49	0/2444	0.69	0/3308
2	C	0.48	0/3686	0.70	0/4972
2	D	0.48	0/3686	0.70	0/4972
3	E	1.01	0/103	0.96	0/156
3	F	0.76	0/103	0.74	0/156
4	G	0.92	0/82	0.78	0/124
5	I	0.75	0/103	0.71	0/156
All	All	0.50	0/12651	0.70	0/17152

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
2	D	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	D	455	TYR	Sidechain

## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2382	0	2389	223	0
1	B	2382	0	2389	229	0
2	C	3627	0	3689	595	0
2	D	3627	0	3689	600	0
3	E	94	0	55	7	0
3	F	94	0	55	14	0
4	G	75	0	44	5	0
5	I	94	0	55	11	0
6	C	20	0	11	3	0
6	D	20	0	11	2	0
7	C	30	0	12	3	0
7	D	30	0	12	4	0
8	C	2	0	0	0	0
8	D	2	0	0	0	0
9	C	1	0	0	2	0
All	All	12480	0	12411	1620	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 65.

The worst 5 of 1620 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:C:168:VAL:HA	2:C:171:LEU:HD12	1.23	1.11
2:D:168:VAL:HA	2:D:171:LEU:HD12	1.22	1.10
1:A:219:MET:HE3	1:A:221:LEU:HD11	1.32	1.10
2:C:86:THR:HB	2:C:87:GLN:HE22	1.20	1.06
2:C:91:ILE:H	2:C:91:ILE:HD13	1.21	1.06

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone ⓘ

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	282/297 (95%)	207 (73%)	57 (20%)	18 (6%)	2	17
1	B	282/297 (95%)	203 (72%)	61 (22%)	18 (6%)	2	17
2	C	439/479 (92%)	293 (67%)	110 (25%)	36 (8%)	1	10
2	D	439/479 (92%)	295 (67%)	109 (25%)	35 (8%)	1	11
All	All	1442/1552 (93%)	998 (69%)	337 (23%)	107 (7%)	2	12

5 of 107 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	2	ASP
1	A	146	ASN
1	A	147	GLU
1	A	277	PRO
1	B	2	ASP

### 5.3.2 Protein sidechains ⓘ

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	266/270 (98%)	241 (91%)	25 (9%)	13	45
1	B	266/270 (98%)	241 (91%)	25 (9%)	13	45
2	C	422/453 (93%)	390 (92%)	32 (8%)	19	60
2	D	422/453 (93%)	390 (92%)	32 (8%)	19	60
All	All	1376/1446 (95%)	1262 (92%)	114 (8%)	16	55

5 of 114 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	46	VAL
2	C	229	ILE
2	D	332	PRO
2	C	59	ASN
2	C	110	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 55 such sidechains are listed below:

Mol	Chain	Res	Type
2	C	87	GLN
2	C	247	ASN
2	D	341	ASN
2	C	101	GLN
2	C	167	ASN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	E	1/5 (20%)	1 (100%)	0
3	F	2/5 (40%)	1 (50%)	0
4	G	2/4 (50%)	1 (50%)	0
5	I	2/5 (40%)	1 (50%)	0
All	All	7/19 (36%)	4 (57%)	0

All (4) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
3	E	603	U
3	F	607	U
4	G	502	U
5	I	703	U

There are no RNA pucker outliers to report.

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

Of 8 ligands modelled in this entry, 4 are monoatomic - leaving 4 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
6	U	C	600	-	19,21,22	1.53	2 (10%)	23,30,33	0.90	1 (4%)
7	3AT	C	801	8	32,32,32	1.89	9 (28%)	50,50,50	2.67	14 (28%)
6	U	D	602	-	19,21,22	1.69	2 (10%)	23,30,33	0.81	1 (4%)
7	3AT	D	800	8	32,32,32	2.07	12 (37%)	50,50,50	2.70	12 (24%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	U	C	600	-	-	0/6/25/26	0/2/2/2
7	3AT	C	801	8	-	0/22/34/34	0/1/3/3
6	U	D	602	-	-	0/6/25/26	0/2/2/2
7	3AT	D	800	8	-	0/22/34/34	0/1/3/3

The worst 5 of 25 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	D	602	U	P-OP1	6.36	1.53	1.46
6	C	600	U	P-OP1	5.65	1.53	1.46
7	D	800	3AT	PB-O3B	4.94	1.68	1.59
7	C	801	3AT	PA-O3A	4.03	1.67	1.59
7	C	801	3AT	PB-O3A	3.94	1.67	1.59

The worst 5 of 28 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	D	800	3AT	N3-C2-N1	-11.92	118.75	128.71
7	C	801	3AT	N3-C2-N1	-11.83	118.82	128.71
7	D	800	3AT	PA-O3A-PB	-6.58	112.38	131.68
7	C	801	3AT	PA-O3A-PB	-5.67	115.05	131.68
7	C	801	3AT	PB-O3B-PG	-5.39	115.87	131.68

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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 ⓘ

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	288/297 (96%)	-0.33	0 100 100	41, 68, 94, 133	0
1	B	288/297 (96%)	-0.35	0 100 100	41, 66, 93, 124	0
2	C	445/479 (92%)	-0.21	0 100 100	47, 90, 123, 156	0
2	D	445/479 (92%)	-0.17	0 100 100	41, 94, 130, 162	0
3	E	5/5 (100%)	0.17	0 100 100	101, 102, 150, 157	0
3	F	5/5 (100%)	0.29	0 100 100	97, 100, 138, 157	0
4	G	4/4 (100%)	-0.40	0 100 100	81, 97, 119, 122	0
5	I	5/5 (100%)	-0.36	0 100 100	95, 101, 116, 116	0
All	All	1485/1571 (94%)	-0.24	0 100 100	41, 82, 122, 162	0

There are no RSRZ outliers to report.

### 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 ⓘ

There are no carbohydrates in this entry.

### 6.4 Ligands ⓘ

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, 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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
6	U	D	602	20/21	0.32	2.37	79,79,79,79	0
7	3AT	D	800	30/30	0.25	0.21	79,108,108,108	0
7	3AT	C	801	30/30	0.20	0.02	79,93,97,97	0
6	U	C	600	20/21	0.16	-0.13	79,79,79,79	0
8	CA	C	901	1/1	0.20	-0.30	79,79,79,79	0
8	CA	C	902	1/1	0.14	-0.90	79,79,79,79	0
8	CA	D	904	1/1	0.15	-1.37	79,79,79,79	0
8	CA	D	903	1/1	0.14	-1.90	79,79,79,79	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.