



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2018 – 02:47 pm GMT

PDB ID : 3F9K  
Title : Two domain fragment of HIV-2 integrase in complex with LEDGF IBD  
Authors : Hare, S.; Cherepanov, P.  
Deposited on : 2008-11-14  
Resolution : 3.20 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](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
Xtriage (Phenix)	:	1.13
EDS	:	trunk30967
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac	:	5.8.0158
CCP4	:	7.0 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk30967

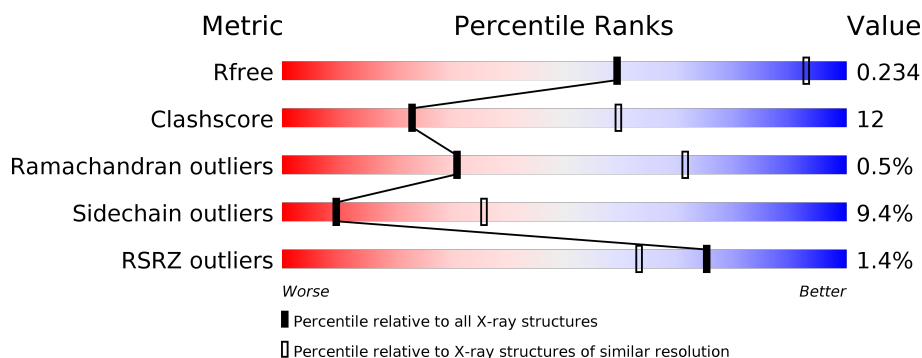
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.20 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))
$R_{free}$	111664	1121 (3.22-3.18)
Clashscore	122126	1091 (3.20-3.20)
Ramachandran outliers	120053	1074 (3.20-3.20)
Sidechain outliers	120020	1073 (3.20-3.20)
RSRZ outliers	108989	1083 (3.22-3.18)

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.

Mol	Chain	Length	Quality of chain
1	A	210	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 22%, green 71%, grey 4%);"></div> <div style="display: flex; justify-content: space-between; width: 90%; margin: 0 auto;"> <span>%</span> <span>71%</span> <span>22%</span> <span>• •</span> </div> </div>
1	B	210	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, green 70%, yellow 20%, orange 7%, grey 3%);"></div> <div style="display: flex; justify-content: space-between; width: 90%; margin: 0 auto;"> <span>70%</span> <span>20%</span> <span>• 7%</span> </div> </div>
1	E	210	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, green 72%, yellow 20%, orange 7%, grey 1%);"></div> <div style="display: flex; justify-content: space-between; width: 90%; margin: 0 auto;"> <span>72%</span> <span>20%</span> <span>• •</span> </div> </div>
1	F	210	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, green 70%, yellow 20%, orange 7%, grey 3%);"></div> <div style="display: flex; justify-content: space-between; width: 90%; margin: 0 auto;"> <span>70%</span> <span>20%</span> <span>• 7%</span> </div> </div>
1	I	210	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, green 72%, yellow 20%, orange 7%, grey 1%);"></div> <div style="display: flex; justify-content: space-between; width: 90%; margin: 0 auto;"> <span>72%</span> <span>20%</span> <span>• •</span> </div> </div>
1	J	210	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 20%, green 70%, grey 2%);"></div> <div style="display: flex; justify-content: space-between; width: 90%; margin: 0 auto;"> <span>%</span> <span>70%</span> <span>20%</span> <span>• 7%</span> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	M	210	
1	N	210	
1	Q	210	
1	R	210	
1	U	210	
1	V	210	
1	Y	210	
1	Z	210	
1	c	210	
1	d	210	
1	g	210	
1	h	210	
1	k	210	
1	l	210	
1	o	210	
1	p	210	
1	s	210	
1	t	210	
2	C	95	
2	G	95	
2	K	95	
2	O	95	
2	S	95	
2	W	95	
2	a	95	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	e	95	<div><div>%</div><div><div></div><div>85%</div><div>13%</div><div>..</div></div></div>
2	i	95	<div><div>3%</div><div><div></div><div>85%</div><div>12%</div><div>..</div></div></div>
2	m	95	<div><div>6%</div><div><div></div><div>87%</div><div>11%</div><div>..</div></div></div>
2	q	95	<div><div>6%</div><div><div></div><div>86%</div><div>13%</div><div>.</div></div></div>
2	u	95	<div><div>5%</div><div><div></div><div>86%</div><div>13%</div><div>.</div></div></div>

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 46740 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 Integrase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	204	Total	C	N	O	S	0	0	0
			1588	999	288	290	11			
1	B	196	Total	C	N	O	S	0	0	0
			1542	972	277	282	11			
1	E	204	Total	C	N	O	S	0	0	0
			1588	999	288	290	11			
1	F	196	Total	C	N	O	S	0	0	0
			1542	972	277	282	11			
1	I	204	Total	C	N	O	S	0	0	0
			1588	999	288	290	11			
1	J	196	Total	C	N	O	S	0	0	0
			1542	972	277	282	11			
1	M	204	Total	C	N	O	S	0	0	0
			1588	999	288	290	11			
1	N	196	Total	C	N	O	S	0	0	0
			1542	972	277	282	11			
1	Q	204	Total	C	N	O	S	0	0	0
			1588	999	288	290	11			
1	R	196	Total	C	N	O	S	0	0	0
			1542	972	277	282	11			
1	U	204	Total	C	N	O	S	0	0	0
			1588	999	288	290	11			
1	V	196	Total	C	N	O	S	0	0	0
			1542	972	277	282	11			
1	Y	204	Total	C	N	O	S	0	0	0
			1588	999	288	290	11			
1	Z	196	Total	C	N	O	S	0	0	0
			1542	972	277	282	11			
1	c	204	Total	C	N	O	S	0	0	0
			1588	999	288	290	11			
1	d	196	Total	C	N	O	S	0	0	0
			1542	972	277	282	11			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	g	204	Total	C	N	O	S	0	0	0
			1588	999	288	290	11			
1	h	196	Total	C	N	O	S	0	0	0
			1542	972	277	282	11			
1	k	204	Total	C	N	O	S	0	0	0
			1588	999	288	290	11			
1	l	196	Total	C	N	O	S	0	0	0
			1542	972	277	282	11			
1	o	204	Total	C	N	O	S	0	0	0
			1588	999	288	290	11			
1	p	196	Total	C	N	O	S	0	0	0
			1542	972	277	282	11			
1	s	204	Total	C	N	O	S	0	0	0
			1588	999	288	290	11			
1	t	196	Total	C	N	O	S	0	0	0
			1542	972	277	282	11			

There are 72 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	EXPRESSION TAG	UNP P04584
A	1	VAL	-	EXPRESSION TAG	UNP P04584
A	180	VAL	ILE	VARIANT	UNP P04584
B	0	MET	-	EXPRESSION TAG	UNP P04584
B	1	VAL	-	EXPRESSION TAG	UNP P04584
B	180	VAL	ILE	VARIANT	UNP P04584
E	0	MET	-	EXPRESSION TAG	UNP P04584
E	1	VAL	-	EXPRESSION TAG	UNP P04584
E	180	VAL	ILE	VARIANT	UNP P04584
F	0	MET	-	EXPRESSION TAG	UNP P04584
F	1	VAL	-	EXPRESSION TAG	UNP P04584
F	180	VAL	ILE	VARIANT	UNP P04584
I	0	MET	-	EXPRESSION TAG	UNP P04584
I	1	VAL	-	EXPRESSION TAG	UNP P04584
I	180	VAL	ILE	VARIANT	UNP P04584
J	0	MET	-	EXPRESSION TAG	UNP P04584
J	1	VAL	-	EXPRESSION TAG	UNP P04584
J	180	VAL	ILE	VARIANT	UNP P04584
M	0	MET	-	EXPRESSION TAG	UNP P04584
M	1	VAL	-	EXPRESSION TAG	UNP P04584
M	180	VAL	ILE	VARIANT	UNP P04584
N	0	MET	-	EXPRESSION TAG	UNP P04584
N	1	VAL	-	EXPRESSION TAG	UNP P04584

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
N	180	VAL	ILE	VARIANT	UNP P04584
Q	0	MET	-	EXPRESSION TAG	UNP P04584
Q	1	VAL	-	EXPRESSION TAG	UNP P04584
Q	180	VAL	ILE	VARIANT	UNP P04584
R	0	MET	-	EXPRESSION TAG	UNP P04584
R	1	VAL	-	EXPRESSION TAG	UNP P04584
R	180	VAL	ILE	VARIANT	UNP P04584
U	0	MET	-	EXPRESSION TAG	UNP P04584
U	1	VAL	-	EXPRESSION TAG	UNP P04584
U	180	VAL	ILE	VARIANT	UNP P04584
V	0	MET	-	EXPRESSION TAG	UNP P04584
V	1	VAL	-	EXPRESSION TAG	UNP P04584
V	180	VAL	ILE	VARIANT	UNP P04584
Y	0	MET	-	EXPRESSION TAG	UNP P04584
Y	1	VAL	-	EXPRESSION TAG	UNP P04584
Y	180	VAL	ILE	VARIANT	UNP P04584
Z	0	MET	-	EXPRESSION TAG	UNP P04584
Z	1	VAL	-	EXPRESSION TAG	UNP P04584
Z	180	VAL	ILE	VARIANT	UNP P04584
c	0	MET	-	EXPRESSION TAG	UNP P04584
c	1	VAL	-	EXPRESSION TAG	UNP P04584
c	180	VAL	ILE	VARIANT	UNP P04584
d	0	MET	-	EXPRESSION TAG	UNP P04584
d	1	VAL	-	EXPRESSION TAG	UNP P04584
d	180	VAL	ILE	VARIANT	UNP P04584
g	0	MET	-	EXPRESSION TAG	UNP P04584
g	1	VAL	-	EXPRESSION TAG	UNP P04584
g	180	VAL	ILE	VARIANT	UNP P04584
h	0	MET	-	EXPRESSION TAG	UNP P04584
h	1	VAL	-	EXPRESSION TAG	UNP P04584
h	180	VAL	ILE	VARIANT	UNP P04584
k	0	MET	-	EXPRESSION TAG	UNP P04584
k	1	VAL	-	EXPRESSION TAG	UNP P04584
k	180	VAL	ILE	VARIANT	UNP P04584
l	0	MET	-	EXPRESSION TAG	UNP P04584
l	1	VAL	-	EXPRESSION TAG	UNP P04584
l	180	VAL	ILE	VARIANT	UNP P04584
o	0	MET	-	EXPRESSION TAG	UNP P04584
o	1	VAL	-	EXPRESSION TAG	UNP P04584
o	180	VAL	ILE	VARIANT	UNP P04584
p	0	MET	-	EXPRESSION TAG	UNP P04584
p	1	VAL	-	EXPRESSION TAG	UNP P04584

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
p	180	VAL	ILE	VARIANT	UNP P04584
s	0	MET	-	EXPRESSION TAG	UNP P04584
s	1	VAL	-	EXPRESSION TAG	UNP P04584
s	180	VAL	ILE	VARIANT	UNP P04584
t	0	MET	-	EXPRESSION TAG	UNP P04584
t	1	VAL	-	EXPRESSION TAG	UNP P04584
t	180	VAL	ILE	VARIANT	UNP P04584

- Molecule 2 is a protein called PC4 and SFRS1-interacting protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	94	Total 761	C 479	N 133	O 142	S 7	3	0	0
2	G	94	Total 761	C 479	N 133	O 142	S 7	3	0	0
2	K	94	Total 761	C 479	N 133	O 142	S 7	3	0	0
2	O	94	Total 761	C 479	N 133	O 142	S 7	3	0	0
2	S	94	Total 761	C 479	N 133	O 142	S 7	3	0	0
2	W	94	Total 761	C 479	N 133	O 142	S 7	3	0	0
2	a	94	Total 761	C 479	N 133	O 142	S 7	3	0	0
2	e	94	Total 761	C 479	N 133	O 142	S 7	3	0	0
2	i	94	Total 761	C 479	N 133	O 142	S 7	3	0	0
2	m	94	Total 761	C 479	N 133	O 142	S 7	3	0	0
2	q	94	Total 761	C 479	N 133	O 142	S 7	3	0	0
2	u	94	Total 761	C 479	N 133	O 142	S 7	3	0	0

There are 72 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	436	LEU	-	EXPRESSION TAG	UNP O75475
C	437	GLU	-	EXPRESSION TAG	UNP O75475
C	438	VAL	-	EXPRESSION TAG	UNP O75475

*Continued on next page...*



*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	439	LEU	-	EXPRESSION TAG	UNP O75475
C	440	PHE	-	EXPRESSION TAG	UNP O75475
C	441	GLN	-	EXPRESSION TAG	UNP O75475
G	436	LEU	-	EXPRESSION TAG	UNP O75475
G	437	GLU	-	EXPRESSION TAG	UNP O75475
G	438	VAL	-	EXPRESSION TAG	UNP O75475
G	439	LEU	-	EXPRESSION TAG	UNP O75475
G	440	PHE	-	EXPRESSION TAG	UNP O75475
G	441	GLN	-	EXPRESSION TAG	UNP O75475
K	436	LEU	-	EXPRESSION TAG	UNP O75475
K	437	GLU	-	EXPRESSION TAG	UNP O75475
K	438	VAL	-	EXPRESSION TAG	UNP O75475
K	439	LEU	-	EXPRESSION TAG	UNP O75475
K	440	PHE	-	EXPRESSION TAG	UNP O75475
K	441	GLN	-	EXPRESSION TAG	UNP O75475
O	436	LEU	-	EXPRESSION TAG	UNP O75475
O	437	GLU	-	EXPRESSION TAG	UNP O75475
O	438	VAL	-	EXPRESSION TAG	UNP O75475
O	439	LEU	-	EXPRESSION TAG	UNP O75475
O	440	PHE	-	EXPRESSION TAG	UNP O75475
O	441	GLN	-	EXPRESSION TAG	UNP O75475
S	436	LEU	-	EXPRESSION TAG	UNP O75475
S	437	GLU	-	EXPRESSION TAG	UNP O75475
S	438	VAL	-	EXPRESSION TAG	UNP O75475
S	439	LEU	-	EXPRESSION TAG	UNP O75475
S	440	PHE	-	EXPRESSION TAG	UNP O75475
S	441	GLN	-	EXPRESSION TAG	UNP O75475
W	436	LEU	-	EXPRESSION TAG	UNP O75475
W	437	GLU	-	EXPRESSION TAG	UNP O75475
W	438	VAL	-	EXPRESSION TAG	UNP O75475
W	439	LEU	-	EXPRESSION TAG	UNP O75475
W	440	PHE	-	EXPRESSION TAG	UNP O75475
W	441	GLN	-	EXPRESSION TAG	UNP O75475
a	436	LEU	-	EXPRESSION TAG	UNP O75475
a	437	GLU	-	EXPRESSION TAG	UNP O75475
a	438	VAL	-	EXPRESSION TAG	UNP O75475
a	439	LEU	-	EXPRESSION TAG	UNP O75475
a	440	PHE	-	EXPRESSION TAG	UNP O75475
a	441	GLN	-	EXPRESSION TAG	UNP O75475
e	436	LEU	-	EXPRESSION TAG	UNP O75475
e	437	GLU	-	EXPRESSION TAG	UNP O75475
e	438	VAL	-	EXPRESSION TAG	UNP O75475

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
e	439	LEU	-	EXPRESSION TAG	UNP O75475
e	440	PHE	-	EXPRESSION TAG	UNP O75475
e	441	GLN	-	EXPRESSION TAG	UNP O75475
i	436	LEU	-	EXPRESSION TAG	UNP O75475
i	437	GLU	-	EXPRESSION TAG	UNP O75475
i	438	VAL	-	EXPRESSION TAG	UNP O75475
i	439	LEU	-	EXPRESSION TAG	UNP O75475
i	440	PHE	-	EXPRESSION TAG	UNP O75475
i	441	GLN	-	EXPRESSION TAG	UNP O75475
m	436	LEU	-	EXPRESSION TAG	UNP O75475
m	437	GLU	-	EXPRESSION TAG	UNP O75475
m	438	VAL	-	EXPRESSION TAG	UNP O75475
m	439	LEU	-	EXPRESSION TAG	UNP O75475
m	440	PHE	-	EXPRESSION TAG	UNP O75475
m	441	GLN	-	EXPRESSION TAG	UNP O75475
q	436	LEU	-	EXPRESSION TAG	UNP O75475
q	437	GLU	-	EXPRESSION TAG	UNP O75475
q	438	VAL	-	EXPRESSION TAG	UNP O75475
q	439	LEU	-	EXPRESSION TAG	UNP O75475
q	440	PHE	-	EXPRESSION TAG	UNP O75475
q	441	GLN	-	EXPRESSION TAG	UNP O75475
u	436	LEU	-	EXPRESSION TAG	UNP O75475
u	437	GLU	-	EXPRESSION TAG	UNP O75475
u	438	VAL	-	EXPRESSION TAG	UNP O75475
u	439	LEU	-	EXPRESSION TAG	UNP O75475
u	440	PHE	-	EXPRESSION TAG	UNP O75475
u	441	GLN	-	EXPRESSION TAG	UNP O75475

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	g	1	Total Zn 1 1	0	0
3	h	1	Total Zn 1 1	0	0
3	B	1	Total Zn 1 1	0	0
3	c	1	Total Zn 1 1	0	0
3	t	1	Total Zn 1 1	0	0
3	N	1	Total Zn 1 1	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	o	1	Total 1	Zn 1	0	0
3	p	1	Total 1	Zn 1	0	0
3	J	1	Total 1	Zn 1	0	0
3	k	1	Total 1	Zn 1	0	0
3	E	1	Total 1	Zn 1	0	0
3	V	1	Total 1	Zn 1	0	0
3	A	1	Total 1	Zn 1	0	0
3	R	1	Total 1	Zn 1	0	0
3	s	1	Total 1	Zn 1	0	0
3	M	1	Total 1	Zn 1	0	0
3	I	1	Total 1	Zn 1	0	0
3	Z	1	Total 1	Zn 1	0	0
3	U	1	Total 1	Zn 1	0	0
3	Q	1	Total 1	Zn 1	0	0
3	d	1	Total 1	Zn 1	0	0
3	Y	1	Total 1	Zn 1	0	0
3	l	1	Total 1	Zn 1	0	0
3	F	1	Total 1	Zn 1	0	0

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	g	1	Total 1	Mg 1	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	h	1	Total 1	Mg 1	0	0
4	B	1	Total 1	Mg 1	0	0
4	c	1	Total 1	Mg 1	0	0
4	t	1	Total 1	Mg 1	0	0
4	N	1	Total 1	Mg 1	0	0
4	o	1	Total 1	Mg 1	0	0
4	p	1	Total 1	Mg 1	0	0
4	J	1	Total 1	Mg 1	0	0
4	k	1	Total 1	Mg 1	0	0
4	E	1	Total 1	Mg 1	0	0
4	V	1	Total 1	Mg 1	0	0
4	A	1	Total 1	Mg 1	0	0
4	R	1	Total 1	Mg 1	0	0
4	s	1	Total 1	Mg 1	0	0
4	M	1	Total 1	Mg 1	0	0
4	I	1	Total 1	Mg 1	0	0
4	Z	1	Total 1	Mg 1	0	0
4	U	1	Total 1	Mg 1	0	0
4	Q	1	Total 1	Mg 1	0	0
4	d	1	Total 1	Mg 1	0	0
4	Y	1	Total 1	Mg 1	0	0

*Continued on next page...*

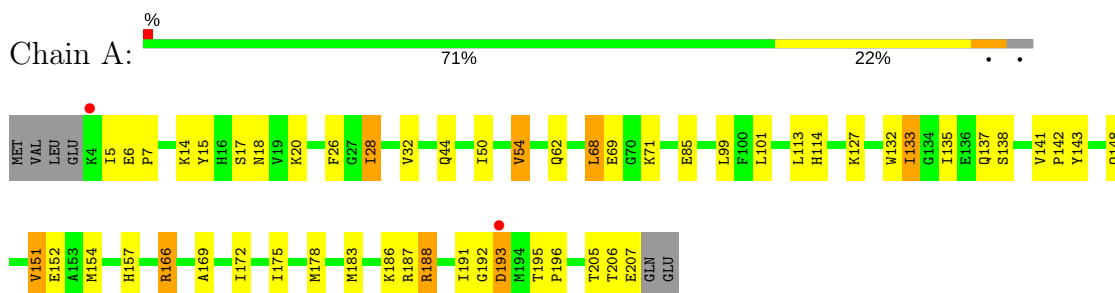
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	1	1	Total 1	Mg 1	0	0
4	F	1	Total 1	Mg 1	0	0

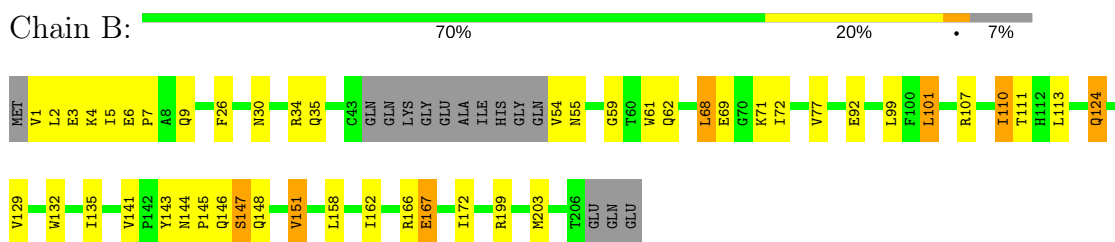
### 3 Residue-property plots [i](#)

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.

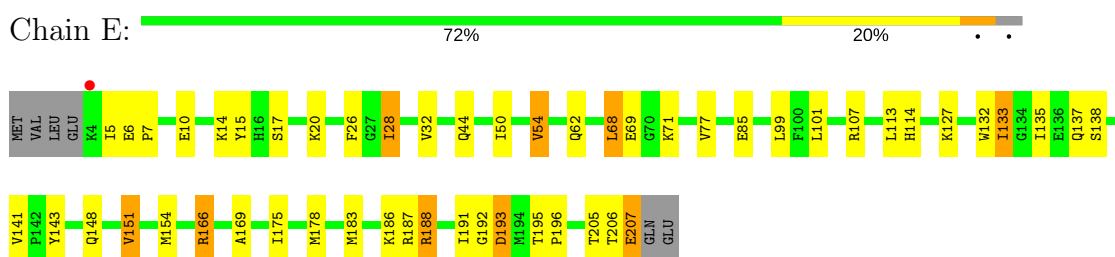
- Molecule 1: Integrase



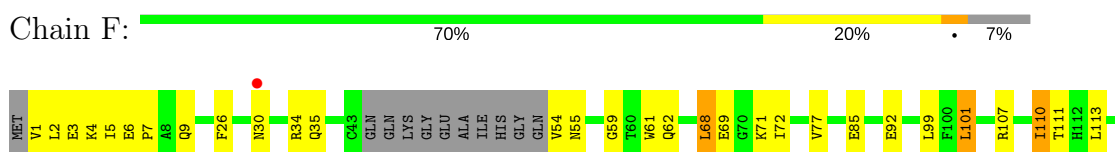
- Molecule 1: Integrase



- Molecule 1: Integrase

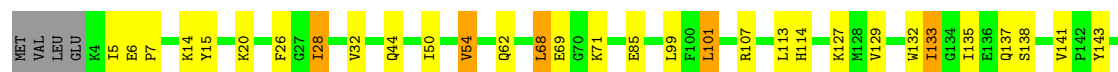


- Molecule 1: Integrase

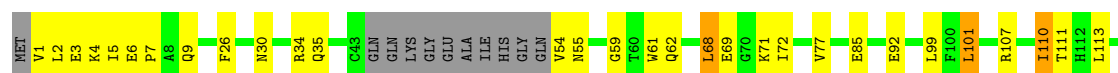




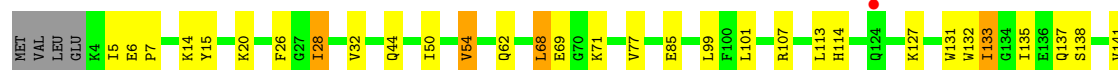
• Molecule 1: Integrase



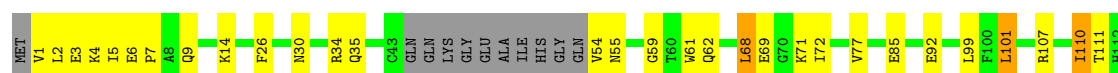
• Molecule 1: Integrase



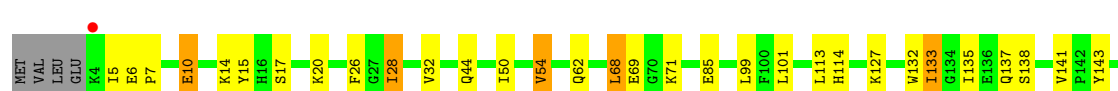
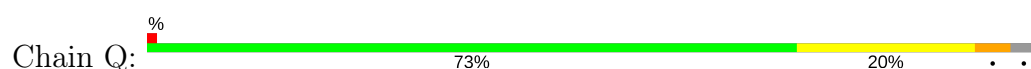
• Molecule 1: Integrase



• Molecule 1: Integrase

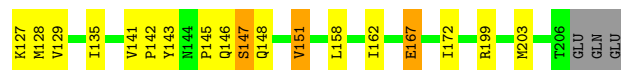
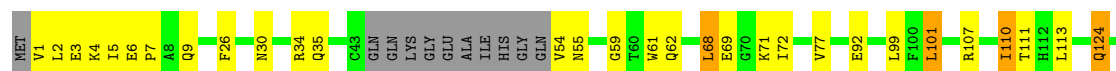


• Molecule 1: Integrase

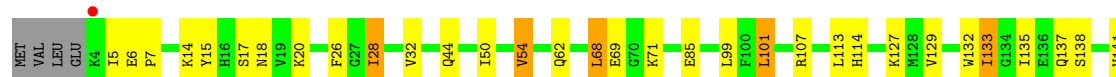
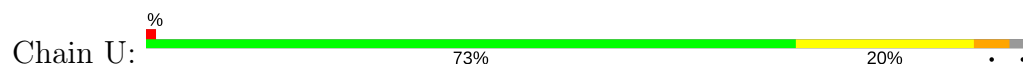




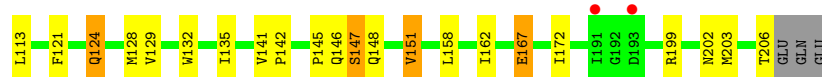
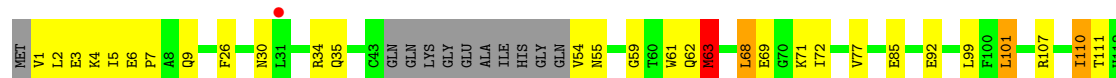
• Molecule 1: Integrase



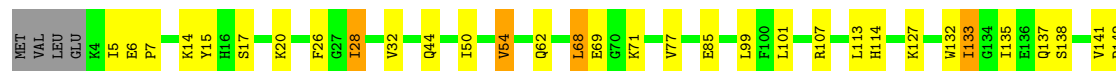
• Molecule 1: Integrase



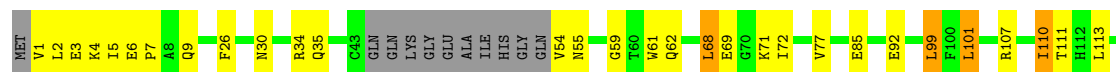
• Molecule 1: Integrase



• Molecule 1: Integrase



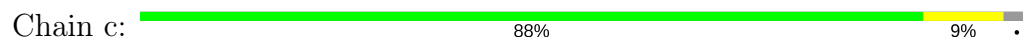
• Molecule 1: Integrase



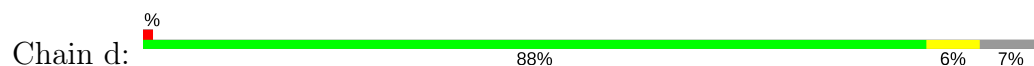




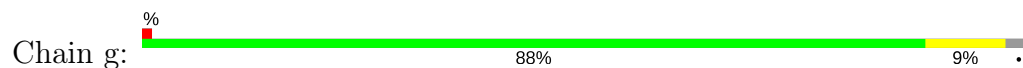
- Molecule 1: Integrase



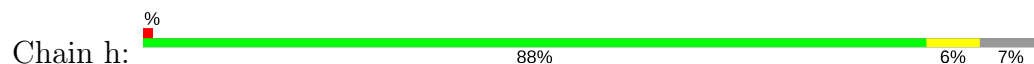
- Molecule 1: Integrase



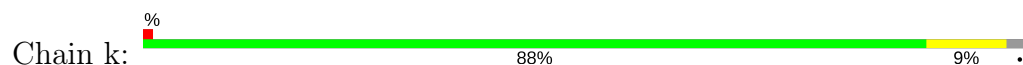
- Molecule 1: Integrase



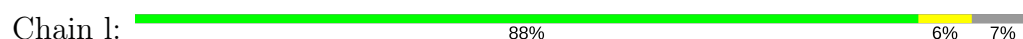
- Molecule 1: Integrase



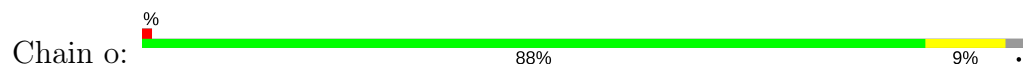
- Molecule 1: Integrase



- Molecule 1: Integrase

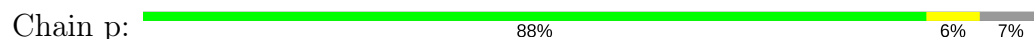


- Molecule 1: Integrase

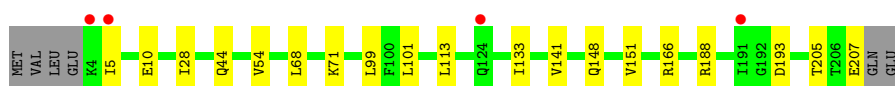
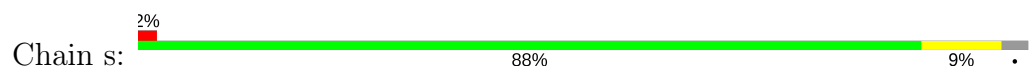




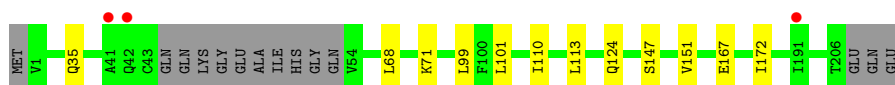
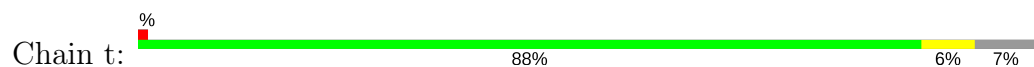
- Molecule 1: Integrase



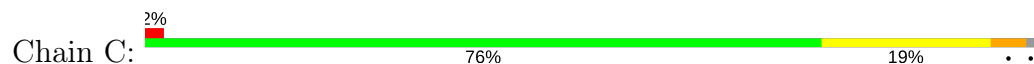
- Molecule 1: Integrase



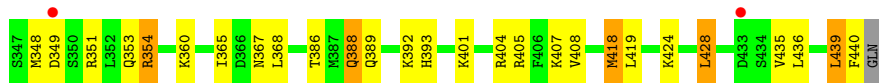
- Molecule 1: Integrase



- Molecule 2: PC4 and SFRS1-interacting protein



- Molecule 2: PC4 and SFRS1-interacting protein



- Molecule 2: PC4 and SFRS1-interacting protein



- Molecule 2: PC4 and SFRS1-interacting protein





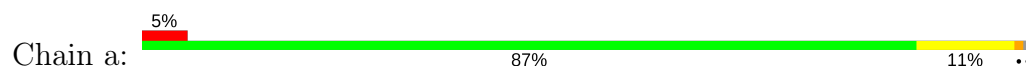
- Molecule 2: PC4 and SFRS1-interacting protein



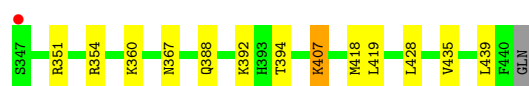
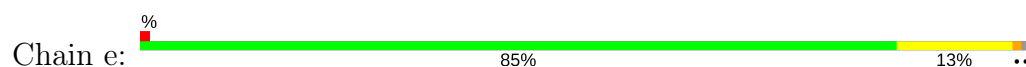
- Molecule 2: PC4 and SFRS1-interacting protein



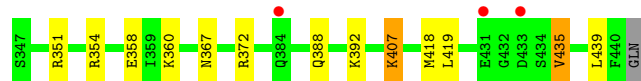
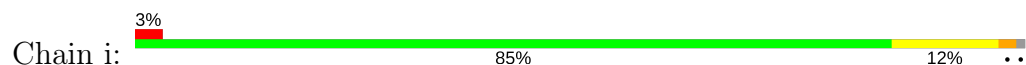
- Molecule 2: PC4 and SFRS1-interacting protein



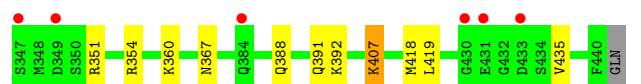
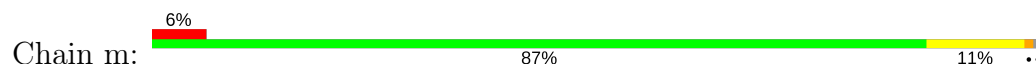
- Molecule 2: PC4 and SFRS1-interacting protein



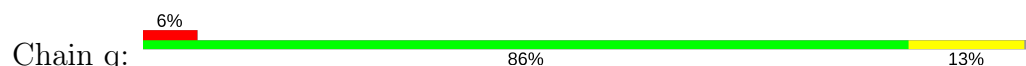
- Molecule 2: PC4 and SFRS1-interacting protein

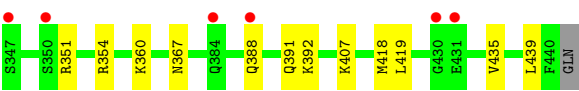


- Molecule 2: PC4 and SFRS1-interacting protein

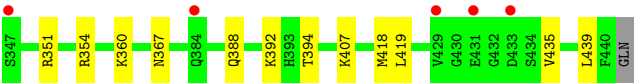
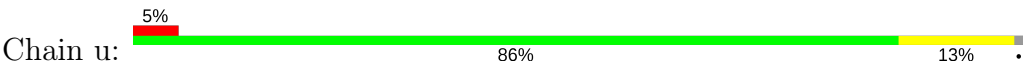


- Molecule 2: PC4 and SFRS1-interacting protein





● Molecule 2: PC4 and SFRS1-interacting protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	201.36Å 202.50Å 280.48Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	34.99 – 3.20 34.98 – 3.20	Depositor EDS
% Data completeness (in resolution range)	100.0 (34.99-3.20) 100.0 (34.98-3.20)	Depositor EDS
$R_{merge}$	0.16	Depositor
$R_{sym}$	0.16	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.44 (at 3.18Å)	Xtriage
Refinement program	REFMAC 5.4.0067	Depositor
R, $R_{free}$	0.225 , 0.234 0.225 , 0.234	Depositor DCC
$R_{free}$ test set	9721 reflections (5.16%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	51.7	Xtriage
Anisotropy	0.339	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 42.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.013 for k,h,-l	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	46740	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.48% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, 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	A	0.35	0/1621	0.50	0/2195
1	B	0.33	0/1573	0.49	0/2130
1	E	0.35	0/1621	0.50	0/2195
1	F	0.33	0/1573	0.50	0/2130
1	I	0.34	0/1621	0.49	0/2195
1	J	0.33	0/1573	0.50	0/2130
1	M	0.34	0/1621	0.49	0/2195
1	N	0.33	0/1573	0.50	0/2130
1	Q	0.34	0/1621	0.49	0/2195
1	R	0.34	0/1573	0.50	0/2130
1	U	0.35	0/1621	0.50	0/2195
1	V	0.34	0/1573	0.56	2/2130 (0.1%)
1	Y	0.35	0/1621	0.50	0/2195
1	Z	0.33	0/1573	0.50	0/2130
1	c	0.35	0/1621	0.50	0/2195
1	d	0.33	0/1573	0.50	0/2130
1	g	0.35	0/1621	0.52	1/2195 (0.0%)
1	h	0.33	0/1573	0.49	0/2130
1	k	0.35	0/1621	0.50	0/2195
1	l	0.33	0/1573	0.50	0/2130
1	o	0.36	0/1621	0.50	0/2195
1	p	0.35	0/1573	0.50	0/2130
1	s	0.36	0/1621	0.50	0/2195
1	t	0.35	0/1573	0.50	0/2130
2	C	0.33	0/767	0.52	0/1024
2	G	0.38	0/767	0.51	0/1024
2	K	0.44	1/767 (0.1%)	0.51	0/1024
2	O	0.46	1/767 (0.1%)	0.52	0/1024
2	S	0.35	0/767	0.52	0/1024
2	W	0.36	0/767	0.54	0/1024
2	a	0.36	0/767	0.52	0/1024
2	e	0.56	1/767 (0.1%)	0.51	0/1024

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
2	i	0.42	1/767 (0.1%)	0.50	0/1024
2	m	0.43	1/767 (0.1%)	0.50	0/1024
2	q	0.38	0/767	0.51	0/1024
2	u	0.35	0/767	0.50	0/1024
All	All	0.36	5/47532 (0.0%)	0.50	3/64188 (0.0%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	e	407	LYS	CG-CD	-12.20	1.10	1.52
2	O	407	LYS	CG-CD	-8.77	1.22	1.52
2	K	407	LYS	CG-CD	-8.04	1.25	1.52
2	m	407	LYS	CG-CD	-6.27	1.31	1.52
2	i	407	LYS	CG-CD	-5.61	1.33	1.52

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	V	63	MET	CG-SD-CE	-8.59	86.46	100.20
1	g	10	GLU	CB-CA-C	-7.20	96.00	110.40
1	V	63	MET	CA-CB-CG	5.75	123.07	113.30

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	1588	0	1565	49	0
1	B	1542	0	1535	42	0
1	E	1588	0	1565	46	0
1	F	1542	0	1535	42	0
1	I	1588	0	1565	47	0
1	J	1542	0	1535	38	0
1	M	1588	0	1565	47	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	N	1542	0	1535	40	1
1	Q	1588	0	1565	44	0
1	R	1542	0	1535	39	0
1	U	1588	0	1565	46	0
1	V	1542	0	1535	39	0
1	Y	1588	0	1565	44	0
1	Z	1542	0	1535	37	2
1	c	1588	0	1565	0	0
1	d	1542	0	1535	0	0
1	g	1588	0	1565	0	0
1	h	1542	0	1535	0	0
1	k	1588	0	1565	0	0
1	l	1542	0	1535	0	0
1	o	1588	0	1565	0	0
1	p	1542	0	1535	0	0
1	s	1588	0	1565	0	0
1	t	1542	0	1535	0	0
2	C	761	0	796	17	0
2	G	761	0	796	26	1
2	K	761	0	796	17	0
2	O	761	0	796	15	0
2	S	761	0	796	26	0
2	W	761	0	796	23	0
2	a	761	0	796	0	0
2	e	761	0	796	0	0
2	i	761	0	796	0	2
2	m	761	0	796	0	0
2	q	761	0	796	0	0
2	u	761	0	796	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	E	1	0	0	0	0
3	F	1	0	0	0	0
3	I	1	0	0	0	0
3	J	1	0	0	0	0
3	M	1	0	0	0	0
3	N	1	0	0	0	0
3	Q	1	0	0	0	0
3	R	1	0	0	0	0
3	U	1	0	0	0	0
3	V	1	0	0	0	0
3	Y	1	0	0	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	Z	1	0	0	0	0
3	c	1	0	0	0	0
3	d	1	0	0	0	0
3	g	1	0	0	0	0
3	h	1	0	0	0	0
3	k	1	0	0	0	0
3	l	1	0	0	0	0
3	o	1	0	0	0	0
3	p	1	0	0	0	0
3	s	1	0	0	0	0
3	t	1	0	0	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	E	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	M	1	0	0	0	0
4	N	1	0	0	0	0
4	Q	1	0	0	0	0
4	R	1	0	0	0	0
4	U	1	0	0	0	0
4	V	1	0	0	0	0
4	Y	1	0	0	0	0
4	Z	1	0	0	0	0
4	c	1	0	0	0	0
4	d	1	0	0	0	0
4	g	1	0	0	0	0
4	h	1	0	0	0	0
4	k	1	0	0	0	0
4	l	1	0	0	0	0
4	o	1	0	0	0	0
4	p	1	0	0	0	0
4	s	1	0	0	0	0
4	t	1	0	0	0	0
All	All	46740	0	46752	651	3

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 12.

The worst 5 of 651 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:F:54:VAL:HG12	1:F:55:ASN:H	1.17	1.10
1:R:54:VAL:HG12	1:R:55:ASN:H	1.16	1.09
1:V:54:VAL:HG12	1:V:55:ASN:H	1.17	1.07
1:B:54:VAL:HG12	1:B:55:ASN:H	1.18	1.06
1:J:54:VAL:HG12	1:J:55:ASN:H	1.17	1.06

All (3) 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)
1:Z:170:ASN:OD1	2:i:372:ARG:NH2[2_555]	2.13	0.07
2:G:354:ARG:NH1	1:N:156:HIS:NE2[3_544]	2.17	0.03
1:Z:170:ASN:ND2	2:i:358:GLU:OE2[2_555]	2.17	0.03

## 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	202/210 (96%)	192 (95%)	9 (4%)	1 (0%)	31	71
1	B	192/210 (91%)	185 (96%)	7 (4%)	0	100	100
1	E	202/210 (96%)	192 (95%)	9 (4%)	1 (0%)	31	71
1	F	192/210 (91%)	185 (96%)	7 (4%)	0	100	100
1	I	202/210 (96%)	192 (95%)	9 (4%)	1 (0%)	31	71
1	J	192/210 (91%)	186 (97%)	6 (3%)	0	100	100
1	M	202/210 (96%)	192 (95%)	9 (4%)	1 (0%)	31	71
1	N	192/210 (91%)	186 (97%)	6 (3%)	0	100	100
1	Q	202/210 (96%)	192 (95%)	9 (4%)	1 (0%)	31	71
1	R	192/210 (91%)	185 (96%)	7 (4%)	0	100	100
1	U	202/210 (96%)	192 (95%)	9 (4%)	1 (0%)	31	71

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	V	192/210 (91%)	185 (96%)	7 (4%)	0	100	100
1	Y	202/210 (96%)	191 (95%)	10 (5%)	1 (0%)	31	71
1	Z	192/210 (91%)	185 (96%)	7 (4%)	0	100	100
1	c	202/210 (96%)	192 (95%)	9 (4%)	1 (0%)	31	71
1	d	192/210 (91%)	185 (96%)	7 (4%)	0	100	100
1	g	202/210 (96%)	192 (95%)	9 (4%)	1 (0%)	31	71
1	h	192/210 (91%)	185 (96%)	7 (4%)	0	100	100
1	k	202/210 (96%)	191 (95%)	10 (5%)	1 (0%)	31	71
1	l	192/210 (91%)	185 (96%)	7 (4%)	0	100	100
1	o	202/210 (96%)	192 (95%)	9 (4%)	1 (0%)	31	71
1	p	192/210 (91%)	185 (96%)	7 (4%)	0	100	100
1	s	202/210 (96%)	192 (95%)	9 (4%)	1 (0%)	31	71
1	t	192/210 (91%)	185 (96%)	7 (4%)	0	100	100
2	C	92/95 (97%)	88 (96%)	3 (3%)	1 (1%)	16	56
2	G	92/95 (97%)	88 (96%)	3 (3%)	1 (1%)	16	56
2	K	92/95 (97%)	88 (96%)	2 (2%)	2 (2%)	7	39
2	O	92/95 (97%)	87 (95%)	4 (4%)	1 (1%)	16	56
2	S	92/95 (97%)	88 (96%)	3 (3%)	1 (1%)	16	56
2	W	92/95 (97%)	86 (94%)	5 (5%)	1 (1%)	16	56
2	a	92/95 (97%)	85 (92%)	5 (5%)	2 (2%)	7	39
2	e	92/95 (97%)	88 (96%)	3 (3%)	1 (1%)	16	56
2	i	92/95 (97%)	86 (94%)	4 (4%)	2 (2%)	7	39
2	m	92/95 (97%)	87 (95%)	4 (4%)	1 (1%)	16	56
2	q	92/95 (97%)	87 (95%)	4 (4%)	1 (1%)	16	56
2	u	92/95 (97%)	86 (94%)	5 (5%)	1 (1%)	16	56
All	All	5832/6180 (94%)	5568 (96%)	237 (4%)	27 (0%)	31	71

5 of 27 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	K	435	VAL
2	a	367	ASN
2	C	367	ASN

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
2	G	367	ASN
2	K	367	ASN

### 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	170/181 (94%)	153 (90%)	17 (10%)	8	33
1	B	169/181 (93%)	157 (93%)	12 (7%)	16	51
1	E	170/181 (94%)	152 (89%)	18 (11%)	7	30
1	F	169/181 (93%)	157 (93%)	12 (7%)	16	51
1	I	170/181 (94%)	153 (90%)	17 (10%)	8	33
1	J	169/181 (93%)	157 (93%)	12 (7%)	16	51
1	M	170/181 (94%)	153 (90%)	17 (10%)	8	33
1	N	169/181 (93%)	157 (93%)	12 (7%)	16	51
1	Q	170/181 (94%)	152 (89%)	18 (11%)	7	30
1	R	169/181 (93%)	157 (93%)	12 (7%)	16	51
1	U	170/181 (94%)	153 (90%)	17 (10%)	8	33
1	V	169/181 (93%)	156 (92%)	13 (8%)	14	48
1	Y	170/181 (94%)	153 (90%)	17 (10%)	8	33
1	Z	169/181 (93%)	157 (93%)	12 (7%)	16	51
1	c	170/181 (94%)	152 (89%)	18 (11%)	7	30
1	d	169/181 (93%)	157 (93%)	12 (7%)	16	51
1	g	170/181 (94%)	153 (90%)	17 (10%)	8	33
1	h	169/181 (93%)	157 (93%)	12 (7%)	16	51
1	k	170/181 (94%)	152 (89%)	18 (11%)	7	30
1	l	169/181 (93%)	157 (93%)	12 (7%)	16	51
1	o	170/181 (94%)	152 (89%)	18 (11%)	7	30
1	p	169/181 (93%)	157 (93%)	12 (7%)	16	51

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	s	170/181 (94%)	152 (89%)	18 (11%)	7	30
1	t	169/181 (93%)	157 (93%)	12 (7%)	16	51
2	C	88/89 (99%)	78 (89%)	10 (11%)	6	27
2	G	88/89 (99%)	77 (88%)	11 (12%)	5	23
2	K	88/89 (99%)	78 (89%)	10 (11%)	6	27
2	O	88/89 (99%)	78 (89%)	10 (11%)	6	27
2	S	88/89 (99%)	77 (88%)	11 (12%)	5	23
2	W	88/89 (99%)	77 (88%)	11 (12%)	5	23
2	a	88/89 (99%)	78 (89%)	10 (11%)	6	27
2	e	88/89 (99%)	76 (86%)	12 (14%)	4	19
2	i	88/89 (99%)	78 (89%)	10 (11%)	6	27
2	m	88/89 (99%)	78 (89%)	10 (11%)	6	27
2	q	88/89 (99%)	77 (88%)	11 (12%)	5	23
2	u	88/89 (99%)	77 (88%)	11 (12%)	5	23
All	All	5124/5412 (95%)	4642 (91%)	482 (9%)	9	36

5 of 482 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	V	68	LEU
2	a	351	ARG
1	s	68	LEU
1	V	124	GLN
1	Y	71	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 147 such sidechains are listed below:

Mol	Chain	Res	Type
1	V	112	HIS
2	a	388	GLN
1	s	51	HIS
1	V	137	GLN
1	Y	124	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](#)

Of 48 ligands modelled in this entry, 48 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 ⓘ

### 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	204/210 (97%)	-0.12	2 (0%) 82 73	30, 43, 66, 83	0
1	B	196/210 (93%)	-0.08	0 100 100	31, 45, 72, 100	0
1	E	204/210 (97%)	-0.19	1 (0%) 90 86	30, 42, 66, 83	0
1	F	196/210 (93%)	-0.07	1 (0%) 90 86	31, 45, 72, 100	0
1	I	204/210 (97%)	-0.11	0 100 100	31, 42, 66, 83	0
1	J	196/210 (93%)	-0.14	2 (1%) 82 73	31, 45, 72, 100	0
1	M	204/210 (97%)	-0.14	1 (0%) 90 86	31, 43, 66, 83	0
1	N	196/210 (93%)	-0.09	1 (0%) 90 86	31, 45, 72, 100	0
1	Q	204/210 (97%)	-0.15	2 (0%) 82 73	31, 42, 66, 83	0
1	R	196/210 (93%)	-0.03	0 100 100	31, 45, 72, 100	0
1	U	204/210 (97%)	-0.12	2 (0%) 82 73	31, 43, 66, 83	0
1	V	196/210 (93%)	-0.09	3 (1%) 73 61	31, 45, 72, 100	0
1	Y	204/210 (97%)	-0.12	1 (0%) 90 86	31, 42, 66, 83	0
1	Z	196/210 (93%)	-0.12	1 (0%) 90 86	31, 45, 72, 100	0
1	c	204/210 (97%)	-0.11	1 (0%) 90 86	31, 43, 66, 83	0
1	d	196/210 (93%)	-0.08	2 (1%) 82 73	31, 45, 72, 100	0
1	g	204/210 (97%)	-0.17	2 (0%) 82 73	30, 43, 66, 83	0
1	h	196/210 (93%)	-0.16	2 (1%) 82 73	31, 45, 72, 100	0
1	k	204/210 (97%)	-0.19	3 (1%) 73 61	31, 43, 66, 83	0
1	l	196/210 (93%)	-0.20	1 (0%) 90 86	31, 45, 72, 100	0
1	o	204/210 (97%)	-0.13	2 (0%) 82 73	31, 43, 66, 83	0
1	p	196/210 (93%)	-0.05	0 100 100	31, 45, 72, 100	0
1	s	204/210 (97%)	-0.14	4 (1%) 65 51	31, 43, 66, 83	0
1	t	196/210 (93%)	-0.09	3 (1%) 73 61	31, 45, 72, 100	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
2	C	94/95 (98%)	0.00	2 (2%) 63 49	32, 48, 68, 81	1 (1%)
2	G	94/95 (98%)	-0.07	2 (2%) 63 49	32, 49, 68, 81	1 (1%)
2	K	94/95 (98%)	0.03	3 (3%) 47 31	32, 53, 70, 81	1 (1%)
2	O	94/95 (98%)	0.14	6 (6%) 19 10	32, 55, 72, 81	1 (1%)
2	S	94/95 (98%)	-0.01	2 (2%) 63 49	32, 50, 68, 81	1 (1%)
2	W	94/95 (98%)	-0.06	4 (4%) 35 22	32, 49, 68, 81	1 (1%)
2	a	94/95 (98%)	0.23	5 (5%) 26 14	32, 56, 75, 81	1 (1%)
2	e	94/95 (98%)	-0.18	1 (1%) 80 69	32, 50, 68, 81	1 (1%)
2	i	94/95 (98%)	0.23	3 (3%) 47 31	32, 56, 76, 81	1 (1%)
2	m	94/95 (98%)	0.40	6 (6%) 19 10	32, 56, 79, 81	1 (1%)
2	q	94/95 (98%)	0.25	6 (6%) 19 10	32, 55, 76, 81	1 (1%)
2	u	94/95 (98%)	0.27	5 (5%) 26 14	32, 56, 76, 81	1 (1%)
All	All	5928/6180 (95%)	-0.08	82 (1%) 75 63	30, 45, 72, 100	12 (0%)

The worst 5 of 82 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	S	349	ASP	3.5
2	m	431	GLU	3.4
2	C	384	GLN	3.4
2	W	347	SER	3.4
2	W	349	ASP	3.2

## 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. 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(Å <sup>2</sup> )	Q<0.9
4	MG	t	211	1/1	0.76	0.16	36,36,36,36	0
4	MG	c	211	1/1	0.78	0.10	28,28,28,28	0
4	MG	E	211	1/1	0.81	0.19	14,14,14,14	0
4	MG	l	211	1/1	0.82	0.27	20,20,20,20	0
4	MG	h	211	1/1	0.83	0.21	17,17,17,17	0
4	MG	g	211	1/1	0.86	0.13	21,21,21,21	0
4	MG	F	211	1/1	0.88	0.17	29,29,29,29	0
4	MG	d	211	1/1	0.91	0.19	28,28,28,28	0
4	MG	R	211	1/1	0.92	0.16	14,14,14,14	0
4	MG	s	211	1/1	0.92	0.17	41,41,41,41	0
4	MG	J	211	1/1	0.93	0.23	52,52,52,52	0
4	MG	Q	211	1/1	0.93	0.17	20,20,20,20	0
4	MG	M	211	1/1	0.94	0.20	24,24,24,24	0
4	MG	o	211	1/1	0.94	0.21	32,32,32,32	0
4	MG	A	211	1/1	0.95	0.16	19,19,19,19	0
4	MG	k	211	1/1	0.95	0.24	41,41,41,41	0
4	MG	p	211	1/1	0.95	0.22	38,38,38,38	0
4	MG	U	211	1/1	0.95	0.21	29,29,29,29	0
4	MG	B	211	1/1	0.95	0.23	39,39,39,39	0
4	MG	N	211	1/1	0.95	0.16	33,33,33,33	0
4	MG	V	211	1/1	0.96	0.23	32,32,32,32	0
4	MG	I	211	1/1	0.96	0.14	20,20,20,20	0
4	MG	Y	211	1/1	0.96	0.18	27,27,27,27	0
3	ZN	F	210	1/1	0.96	0.06	53,53,53,53	0
3	ZN	J	210	1/1	0.97	0.05	49,49,49,49	0
3	ZN	N	210	1/1	0.97	0.06	52,52,52,52	0
4	MG	Z	211	1/1	0.97	0.21	31,31,31,31	0
3	ZN	h	210	1/1	0.97	0.06	49,49,49,49	0
3	ZN	l	210	1/1	0.97	0.04	65,65,65,65	0
3	ZN	p	210	1/1	0.98	0.04	67,67,67,67	0
3	ZN	d	210	1/1	0.98	0.06	51,51,51,51	0
3	ZN	B	210	1/1	0.98	0.05	46,46,46,46	0
3	ZN	E	210	1/1	0.98	0.07	40,40,40,40	0
3	ZN	Z	210	1/1	0.98	0.07	45,45,45,45	0
3	ZN	o	210	1/1	0.99	0.06	51,51,51,51	0
3	ZN	M	210	1/1	0.99	0.06	41,41,41,41	0
3	ZN	t	210	1/1	0.99	0.04	69,69,69,69	0
3	ZN	R	210	1/1	0.99	0.05	47,47,47,47	0
3	ZN	c	210	1/1	0.99	0.08	43,43,43,43	0
3	ZN	V	210	1/1	0.99	0.03	54,54,54,54	0
3	ZN	A	210	1/1	0.99	0.08	38,38,38,38	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	ZN	Q	210	1/1	0.99	0.07	39,39,39,39	0
3	ZN	g	210	1/1	0.99	0.09	49,49,49,49	0
3	ZN	U	210	1/1	0.99	0.06	41,41,41,41	0
3	ZN	Y	210	1/1	0.99	0.05	43,43,43,43	0
3	ZN	I	210	1/1	0.99	0.07	38,38,38,38	0
3	ZN	s	210	1/1	1.00	0.09	51,51,51,51	0
3	ZN	k	210	1/1	1.00	0.06	49,49,49,49	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.