



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 17, 2020 – 10:57 pm BST

PDB ID : 2PK2  
Title : Cyclin box structure of the P-TEFb subunit Cyclin T1 derived from a fusion complex with EIAV Tat  
Authors : Anand, K.; Schulte, A.; Fujinaga, K.; Scheffzek, K.; Geyer, M.  
Deposited on : 2007-04-17  
Resolution : 2.67 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 : 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

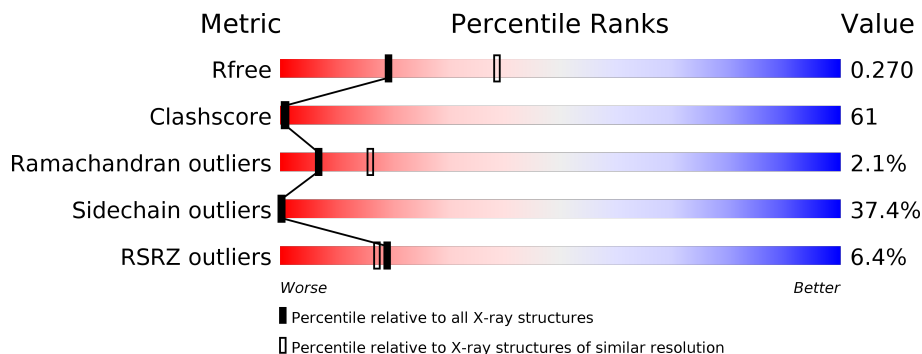
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.67 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	3863 (2.70-2.66)
Clashscore	141614	4210 (2.70-2.66)
Ramachandran outliers	138981	4141 (2.70-2.66)
Sidechain outliers	138945	4141 (2.70-2.66)
RSRZ outliers	127900	3780 (2.70-2.66)

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	358	
1	B	358	
1	C	358	
1	D	358	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 8396 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 Cyclin-T1, Protein Tat.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	256	2076	1331	360	375	10	0	0	0
1	B	256	2080	1333	360	377	10	0	0	0
1	C	256	2082	1334	361	377	10	0	0	0
1	D	256	2079	1332	360	377	10	0	0	0

There are 84 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	77	ARG	GLN	SEE REMARK 999	UNP O60563
A	282	GLY	-	linker	UNP O60563
A	283	GLY	-	linker	UNP O60563
A	284	THR	-	linker	UNP O60563
A	285	GLY	-	linker	UNP O60563
A	286	GLY	-	linker	UNP O60563
A	287	GLY	-	linker	UNP O60563
A	288	SER	-	linker	UNP O60563
A	289	GLY	-	linker	UNP O60563
A	290	GLY	-	linker	UNP O60563
A	291	GLY	-	linker	UNP O60563
A	292	SER	-	linker	UNP O60563
A	293	GLY	-	linker	UNP O60563
A	294	GLY	-	linker	UNP O60563
A	295	GLY	-	linker	UNP O60563
A	296	SER	-	linker	UNP O60563
A	297	GLY	-	linker	UNP O60563
A	298	GLY	-	linker	UNP O60563
A	299	GLY	-	linker	UNP O60563
A	300	THR	-	linker	UNP O60563
A	301	SER	-	linker	UNP O60563

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	77	ARG	GLN	SEE REMARK 999	UNP O60563
B	282	GLY	-	linker	UNP O60563
B	283	GLY	-	linker	UNP O60563
B	284	THR	-	linker	UNP O60563
B	285	GLY	-	linker	UNP O60563
B	286	GLY	-	linker	UNP O60563
B	287	GLY	-	linker	UNP O60563
B	288	SER	-	linker	UNP O60563
B	289	GLY	-	linker	UNP O60563
B	290	GLY	-	linker	UNP O60563
B	291	GLY	-	linker	UNP O60563
B	292	SER	-	linker	UNP O60563
B	293	GLY	-	linker	UNP O60563
B	294	GLY	-	linker	UNP O60563
B	295	GLY	-	linker	UNP O60563
B	296	SER	-	linker	UNP O60563
B	297	GLY	-	linker	UNP O60563
B	298	GLY	-	linker	UNP O60563
B	299	GLY	-	linker	UNP O60563
B	300	THR	-	linker	UNP O60563
B	301	SER	-	linker	UNP O60563
C	77	ARG	GLN	SEE REMARK 999	UNP O60563
C	282	GLY	-	linker	UNP O60563
C	283	GLY	-	linker	UNP O60563
C	284	THR	-	linker	UNP O60563
C	285	GLY	-	linker	UNP O60563
C	286	GLY	-	linker	UNP O60563
C	287	GLY	-	linker	UNP O60563
C	288	SER	-	linker	UNP O60563
C	289	GLY	-	linker	UNP O60563
C	290	GLY	-	linker	UNP O60563
C	291	GLY	-	linker	UNP O60563
C	292	SER	-	linker	UNP O60563
C	293	GLY	-	linker	UNP O60563
C	294	GLY	-	linker	UNP O60563
C	295	GLY	-	linker	UNP O60563
C	296	SER	-	linker	UNP O60563
C	297	GLY	-	linker	UNP O60563
C	298	GLY	-	linker	UNP O60563
C	299	GLY	-	linker	UNP O60563
C	300	THR	-	linker	UNP O60563
C	301	SER	-	linker	UNP O60563

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	77	ARG	GLN	SEE REMARK 999	UNP O60563
D	282	GLY	-	linker	UNP O60563
D	283	GLY	-	linker	UNP O60563
D	284	THR	-	linker	UNP O60563
D	285	GLY	-	linker	UNP O60563
D	286	GLY	-	linker	UNP O60563
D	287	GLY	-	linker	UNP O60563
D	288	SER	-	linker	UNP O60563
D	289	GLY	-	linker	UNP O60563
D	290	GLY	-	linker	UNP O60563
D	291	GLY	-	linker	UNP O60563
D	292	SER	-	linker	UNP O60563
D	293	GLY	-	linker	UNP O60563
D	294	GLY	-	linker	UNP O60563
D	295	GLY	-	linker	UNP O60563
D	296	SER	-	linker	UNP O60563
D	297	GLY	-	linker	UNP O60563
D	298	GLY	-	linker	UNP O60563
D	299	GLY	-	linker	UNP O60563
D	300	THR	-	linker	UNP O60563
D	301	SER	-	linker	UNP O60563

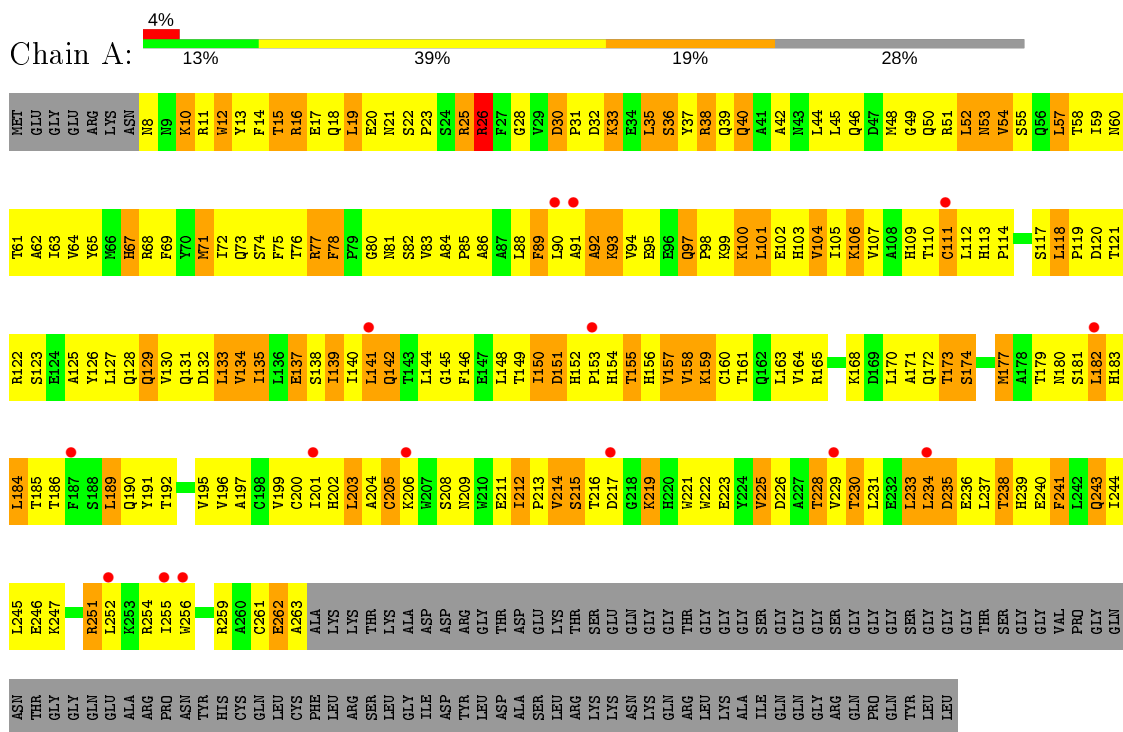
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	25	Total O 25 25	0	0
2	B	16	Total O 16 16	0	0
2	C	21	Total O 21 21	0	0
2	D	17	Total O 17 17	0	0

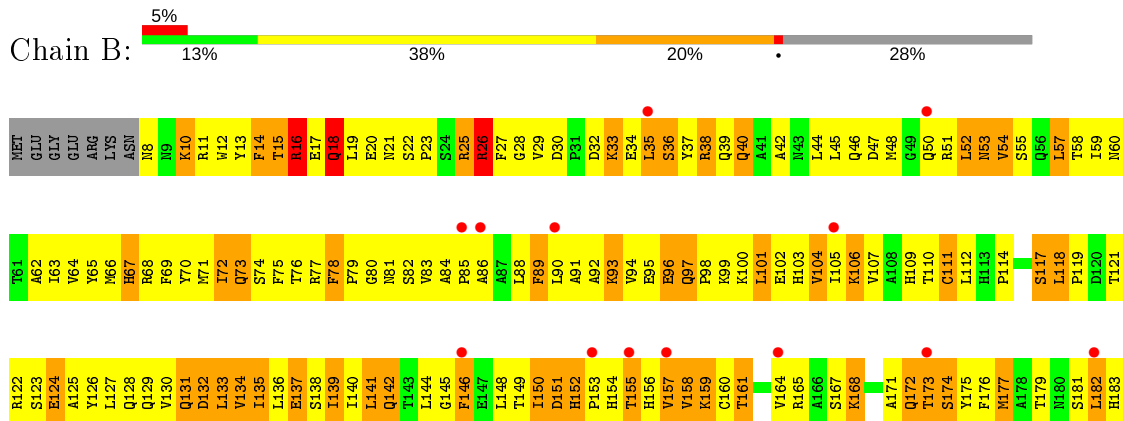
### 3 Residue-property plots

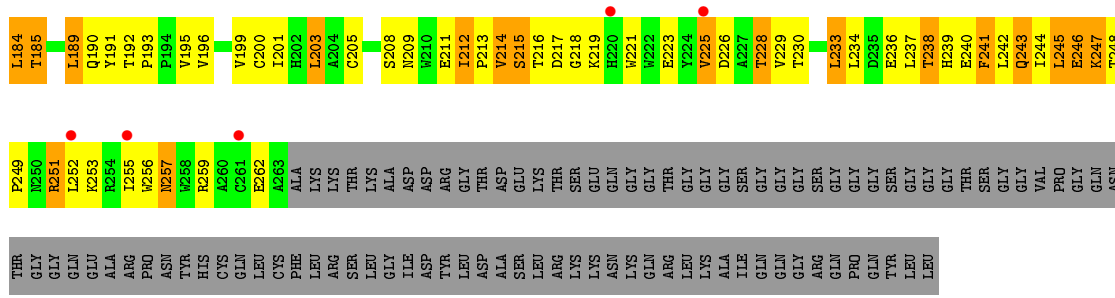
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: Cyclin-T1, Protein Tat

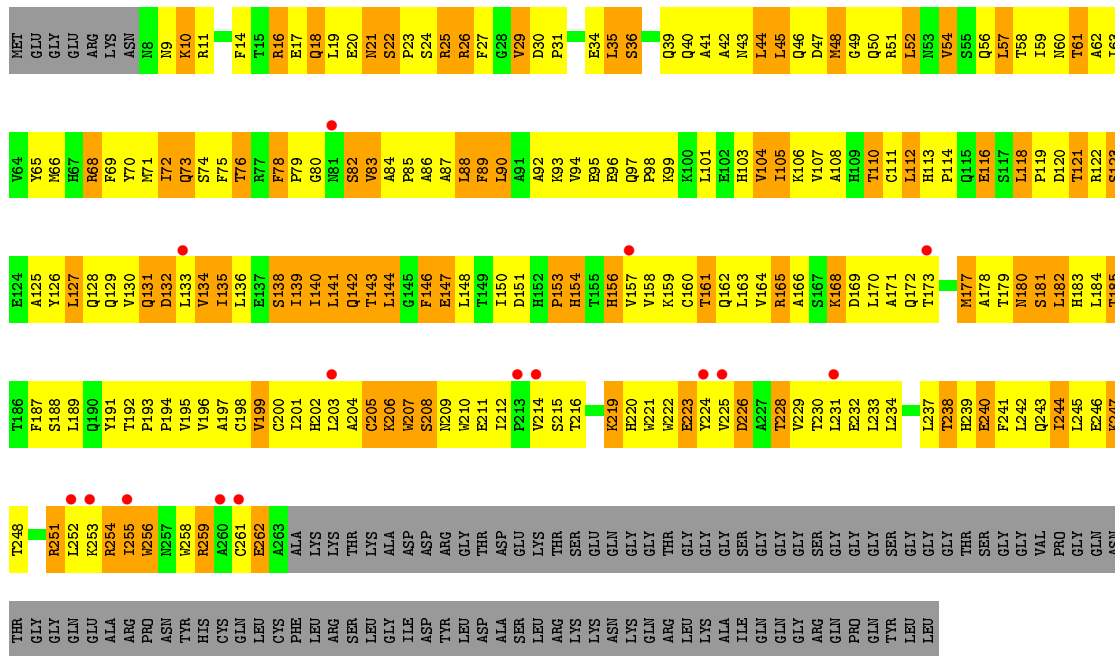


#### • Molecule 1: Cyclin-T1, Protein Tat

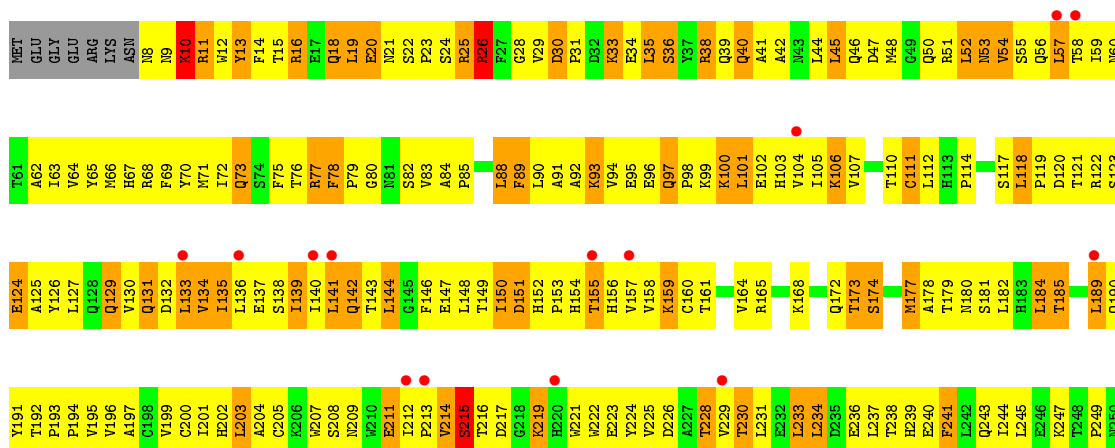




• Molecule 1: Cyclin-T1, Protein Tat



• Molecule 1: Cyclin-T1, Protein Tat



R251	GLY
I252	GLN
R253	GLU
R254	ALA
I255	ARG
R256	PRO
N257	ASN
N258	TRP
R259	HIS
A260	CYS
C261	GLN
E262	LEU
A263	CYS
ALA	PHE
LYS	LEU
LYS	ARG
THR	SER
LYS	LEU
LYS	GLY
ALA	ILE
ASP	ASP
ARG	TYR
GLY	LEU
THR	ASP
ASP	ALA
GLU	SER
LYS	LEU
THR	ARG
THR	LYS
SER	LYS
GLU	LYS
GLN	ASN
GLY	LYS
GLY	GLN
THR	ARG
GLY	LEU
GLY	LYS
GLY	ALA
SER	ILE
GLY	GLN
GLY	GLY
GLY	ARG
GLY	GLN
GLY	PRO
GLY	GLN
GLY	TYR
GLY	LEU
GLY	LEU
THR	THR
SER	SER
GLY	GLY
GLY	VAL
PRO	PRO
GLY	GLY
GLN	GLN
ASN	ASN
THR	THR
GLY	GLY

GLY
GLN
GLU
ALA
ARG
PRO
ASN
TRP
HIS
CYS
GLN
LEU
CYS
PHE
LEU
ARG
SER
LEU
GLY
ILE
ASP
TYR
LEU
ASP
ALA
SER
LEU
ARG
LYS
LYS
ASN
LYS
GLN
ARG
LEU
LYS
ALA
ILE
GLN
GLY
ARG
GLN
PRO
GLN
TYR
LEU
LEU



## 4 Data and refinement statistics i

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	203.83Å 203.83Å 124.79Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	50.00 – 2.67 19.99 – 2.66	Depositor EDS
% Data completeness (in resolution range)	(Not available) (50.00-2.67) 99.2 (19.99-2.66)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.75 (at 2.67Å)	Xtriage
Refinement program	SHELXL-97	Depositor
R, $R_{free}$	0.274 , 0.306 0.249 , 0.270	Depositor DCC
$R_{free}$ test set	1547 reflections (2.80%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	48.0	Xtriage
Anisotropy	0.046	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 45.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.30$ , $\langle L^2 \rangle = 0.13$	Xtriage
Estimated twinning fraction	0.188 for $-2/3^*h-1/3^*k-4/3^*l,-1/3^*h-2/3^*k+4/3^*l,-1/3^*h+1/3^*k+1/3^*l$ 0.186 for $-h,1/3^*h-1/3^*k-4/3^*l,-1/3^*h-2/3^*k+1/3^*l$ 0.186 for $-1/3^*h+1/3^*k+4/3^*l,-k,2/3^*h+1/3^*k+1/3^*l$ 0.437 for $-h,2/3^*h+1/3^*k+4/3^*l,1/3^*h+2/3^*k-1/3^*l$ 0.438 for $-1/3^*h-2/3^*k+4/3^*l,-2/3^*h-1/3^*k-4/3^*l,1/3^*h-1/3^*k-1/3^*l$ 0.430 for $1/3^*h+2/3^*k-4/3^*l,-k,-2/3^*h-1/3^*k-1/3^*l$ 0.187 for $h,-h-k,-l$	Xtriage
$F_o, F_c$ correlation	0.83	EDS
Total number of atoms	8396	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.25% 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](#)

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	0.33	0/2130	0.62	1/2904 (0.0%)
1	B	0.28	0/2134	0.66	3/2909 (0.1%)
1	C	0.25	0/2136	0.59	0/2911
1	D	0.28	0/2133	0.66	5/2908 (0.2%)
All	All	0.29	0/8533	0.63	9/11632 (0.1%)

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	259	ARG	CA-CB-CG	7.41	129.70	113.40
1	A	26	ARG	NE-CZ-NH1	7.30	123.95	120.30
1	B	18	GLN	C-N-CA	6.66	138.35	121.70
1	B	26	ARG	CD-NE-CZ	6.53	132.75	123.60
1	B	16	ARG	C-N-CA	5.96	136.60	121.70

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	2076	0	2043	259	0
1	B	2080	0	2045	241	0
1	C	2082	0	2052	323	0
1	D	2079	0	2045	224	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	25	0	0	6	0
2	B	16	0	0	4	0
2	C	21	0	0	3	0
2	D	17	0	0	0	0
All	All	8396	0	8185	1008	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 61.

The worst 5 of 1008 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:A:15:THR:HG23	1:A:18:GLN:HB2	1.38	1.03
1:D:215:SER:HG	1:D:224:TYR:HE1	1.04	1.02
1:A:177:MET:HB3	1:A:200:CYS:HB3	1.42	0.98
1:C:63:ILE:HG21	1:C:182:LEU:HB3	1.47	0.97
1:C:239:HIS:HA	1:C:242:LEU:HD12	1.47	0.97

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	254/358 (71%)	223 (88%)	26 (10%)	5 (2%)	<a href="#">7</a> <a href="#">17</a>
1	B	254/358 (71%)	219 (86%)	30 (12%)	5 (2%)	<a href="#">7</a> <a href="#">17</a>
1	C	254/358 (71%)	186 (73%)	63 (25%)	5 (2%)	<a href="#">7</a> <a href="#">17</a>
1	D	254/358 (71%)	222 (87%)	26 (10%)	6 (2%)	<a href="#">6</a> <a href="#">13</a>
All	All	1016/1432 (71%)	850 (84%)	145 (14%)	21 (2%)	<a href="#">7</a> <a href="#">16</a>

5 of 21 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	15	THR
1	C	214	VAL
1	D	10	LYS
1	D	215	SER
1	A	38	ARG

### 5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	227/307 (74%)	145 (64%)	82 (36%)	0	0
1	B	228/307 (74%)	145 (64%)	83 (36%)	0	0
1	C	229/307 (75%)	144 (63%)	85 (37%)	0	0
1	D	228/307 (74%)	137 (60%)	91 (40%)	0	0
All	All	912/1228 (74%)	571 (63%)	341 (37%)	0	0

5 of 341 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	230	THR
1	C	88	LEU
1	D	182	LEU
1	B	241	PHE
1	C	29	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 46 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	202	HIS
1	C	129	GLN
1	D	190	GLN
1	C	39	GLN
1	C	60	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](#)

There are no ligands in this entry.

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

There are no such residues in this entry.

### 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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	256/358 (71%)	0.54	15 (5%) 22 20	23, 60, 69, 76	0
1	B	256/358 (71%)	0.58	18 (7%) 16 14	23, 60, 69, 76	0
1	C	256/358 (71%)	0.42	15 (5%) 22 20	22, 53, 67, 80	0
1	D	256/358 (71%)	0.59	18 (7%) 16 14	24, 61, 69, 76	0
All	All	1024/1432 (71%)	0.53	66 (6%) 19 17	22, 59, 69, 80	0

The worst 5 of 66 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	157	VAL	5.8
1	A	255	ILE	5.7
1	D	257	ASN	5.3
1	D	57	LEU	4.3
1	B	255	ILE	4.1

### 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](#)

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.