



wwPDB X-ray Structure Validation Summary Report ⓘ

May 17, 2020 – 07:43 am BST

PDB ID : 2HIK
Title : heterotrimeric PCNA sliding clamp
Authors : Pascal, J.M.; Tsodikov, O.V.; Ellenberger, T.
Deposited on : 2006-06-29
Resolution : 3.30 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<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

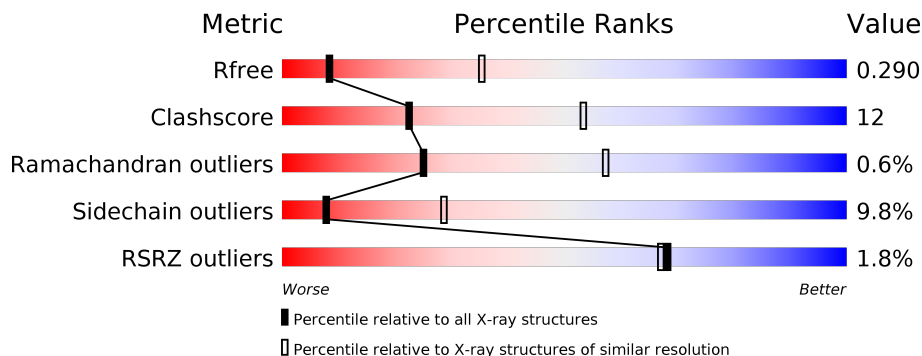
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.30 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.



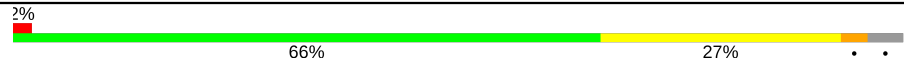

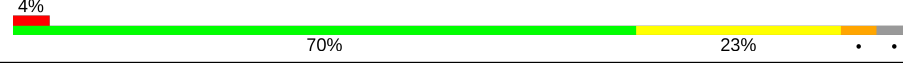
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria 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	257	
1	L	257	
1	X	257	
2	B	245	
2	M	245	
2	Y	245	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	C	252	
3	N	252	
3	Z	252	

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 17311 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 PCNA1 (SSO0397).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	Se			
1	A	249	1924	1226	309	381	8	0	0	0
1	L	248	1919	1223	308	380	8	0	0	0
1	X	248	1919	1223	308	380	8	0	0	0

There are 51 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	MODIFIED RESIDUE	UNP P57766
A	47	MSE	MET	MODIFIED RESIDUE	UNP P57766
A	50	MSE	MET	MODIFIED RESIDUE	UNP P57766
A	157	MSE	MET	MODIFIED RESIDUE	UNP P57766
A	182	MSE	MET	MODIFIED RESIDUE	UNP P57766
A	204	MSE	MET	MODIFIED RESIDUE	UNP P57766
A	220	MSE	MET	MODIFIED RESIDUE	UNP P57766
A	229	MSE	MET	MODIFIED RESIDUE	UNP P57766
A	241	MSE	MET	MODIFIED RESIDUE	UNP P57766
A	250	LEU	-	CLONING ARTIFACT	UNP P57766
A	251	GLU	-	CLONING ARTIFACT	UNP P57766
A	252	HIS	-	EXPRESSION TAG	UNP P57766
A	253	HIS	-	EXPRESSION TAG	UNP P57766
A	254	HIS	-	EXPRESSION TAG	UNP P57766
A	255	HIS	-	EXPRESSION TAG	UNP P57766
A	256	HIS	-	EXPRESSION TAG	UNP P57766
A	257	HIS	-	EXPRESSION TAG	UNP P57766
L	1	MSE	MET	MODIFIED RESIDUE	UNP P57766
L	47	MSE	MET	MODIFIED RESIDUE	UNP P57766
L	50	MSE	MET	MODIFIED RESIDUE	UNP P57766
L	157	MSE	MET	MODIFIED RESIDUE	UNP P57766
L	182	MSE	MET	MODIFIED RESIDUE	UNP P57766
L	204	MSE	MET	MODIFIED RESIDUE	UNP P57766

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
L	220	MSE	MET	MODIFIED RESIDUE	UNP P57766
L	229	MSE	MET	MODIFIED RESIDUE	UNP P57766
L	241	MSE	MET	MODIFIED RESIDUE	UNP P57766
L	250	LEU	-	CLONING ARTIFACT	UNP P57766
L	251	GLU	-	CLONING ARTIFACT	UNP P57766
L	252	HIS	-	EXPRESSION TAG	UNP P57766
L	253	HIS	-	EXPRESSION TAG	UNP P57766
L	254	HIS	-	EXPRESSION TAG	UNP P57766
L	255	HIS	-	EXPRESSION TAG	UNP P57766
L	256	HIS	-	EXPRESSION TAG	UNP P57766
L	257	HIS	-	EXPRESSION TAG	UNP P57766
X	1	MSE	MET	MODIFIED RESIDUE	UNP P57766
X	47	MSE	MET	MODIFIED RESIDUE	UNP P57766
X	50	MSE	MET	MODIFIED RESIDUE	UNP P57766
X	157	MSE	MET	MODIFIED RESIDUE	UNP P57766
X	182	MSE	MET	MODIFIED RESIDUE	UNP P57766
X	204	MSE	MET	MODIFIED RESIDUE	UNP P57766
X	220	MSE	MET	MODIFIED RESIDUE	UNP P57766
X	229	MSE	MET	MODIFIED RESIDUE	UNP P57766
X	241	MSE	MET	MODIFIED RESIDUE	UNP P57766
X	250	LEU	-	CLONING ARTIFACT	UNP P57766
X	251	GLU	-	CLONING ARTIFACT	UNP P57766
X	252	HIS	-	EXPRESSION TAG	UNP P57766
X	253	HIS	-	EXPRESSION TAG	UNP P57766
X	254	HIS	-	EXPRESSION TAG	UNP P57766
X	255	HIS	-	EXPRESSION TAG	UNP P57766
X	256	HIS	-	EXPRESSION TAG	UNP P57766
X	257	HIS	-	EXPRESSION TAG	UNP P57766

- Molecule 2 is a protein called PCNA2 (SSO1047).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	244	Total	C	N	O	Se	0	0	0
			1927	1240	302	381	4			
2	M	243	Total	C	N	O	Se	0	0	0
			1922	1237	301	380	4			
2	Y	243	Total	C	N	O	Se	0	0	0
			1922	1237	301	380	4			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	2	MSE	MET	MODIFIED RESIDUE	UNP Q97Z84
B	199	MSE	MET	MODIFIED RESIDUE	UNP Q97Z84
B	208	MSE	MET	MODIFIED RESIDUE	UNP Q97Z84
B	215	MSE	MET	MODIFIED RESIDUE	UNP Q97Z84
M	2	MSE	MET	MODIFIED RESIDUE	UNP Q97Z84
M	199	MSE	MET	MODIFIED RESIDUE	UNP Q97Z84
M	208	MSE	MET	MODIFIED RESIDUE	UNP Q97Z84
M	215	MSE	MET	MODIFIED RESIDUE	UNP Q97Z84
Y	2	MSE	MET	MODIFIED RESIDUE	UNP Q97Z84
Y	199	MSE	MET	MODIFIED RESIDUE	UNP Q97Z84
Y	208	MSE	MET	MODIFIED RESIDUE	UNP Q97Z84
Y	215	MSE	MET	MODIFIED RESIDUE	UNP Q97Z84

- Molecule 3 is a protein called PCNA3 (SSO0405).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	243	Total	C	N	O	Se	0	0	0
			1926	1224	311	387	4			
3	N	243	Total	C	N	O	Se	0	0	0
			1926	1224	311	387	4			
3	Z	243	Total	C	N	O	Se	0	0	0
			1926	1224	311	387	4			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	1	MSE	MET	MODIFIED RESIDUE	UNP P57765
C	55	MSE	MET	MODIFIED RESIDUE	UNP P57765
C	75	MSE	MET	MODIFIED RESIDUE	UNP P57765
C	230	MSE	MET	MODIFIED RESIDUE	UNP P57765
C	245	LEU	-	CLONING ARTIFACT	UNP P57765
C	246	GLU	-	CLONING ARTIFACT	UNP P57765
C	247	HIS	-	EXPRESSION TAG	UNP P57765
C	248	HIS	-	EXPRESSION TAG	UNP P57765
C	249	HIS	-	EXPRESSION TAG	UNP P57765
C	250	HIS	-	EXPRESSION TAG	UNP P57765
C	251	HIS	-	EXPRESSION TAG	UNP P57765
C	252	HIS	-	EXPRESSION TAG	UNP P57765
N	1	MSE	MET	MODIFIED RESIDUE	UNP P57765
N	55	MSE	MET	MODIFIED RESIDUE	UNP P57765
N	75	MSE	MET	MODIFIED RESIDUE	UNP P57765
N	230	MSE	MET	MODIFIED RESIDUE	UNP P57765
N	245	LEU	-	CLONING ARTIFACT	UNP P57765

Continued on next page...

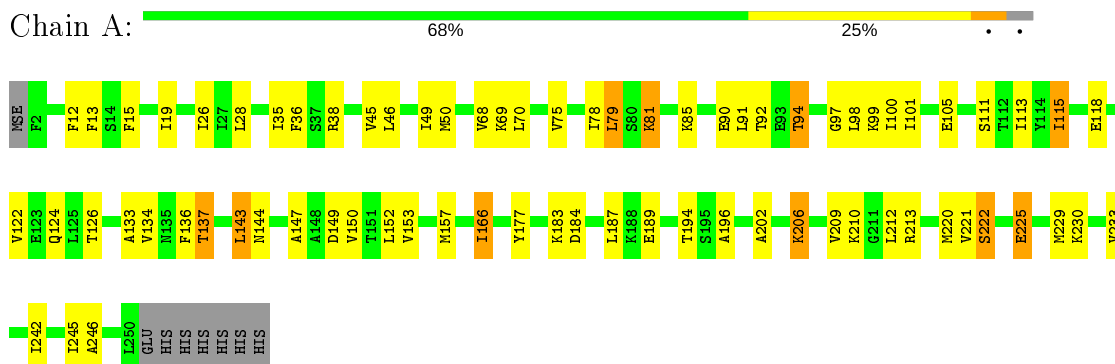
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
N	246	GLU	-	CLONING ARTIFACT	UNP P57765
N	247	HIS	-	EXPRESSION TAG	UNP P57765
N	248	HIS	-	EXPRESSION TAG	UNP P57765
N	249	HIS	-	EXPRESSION TAG	UNP P57765
N	250	HIS	-	EXPRESSION TAG	UNP P57765
N	251	HIS	-	EXPRESSION TAG	UNP P57765
N	252	HIS	-	EXPRESSION TAG	UNP P57765
Z	1	MSE	MET	MODIFIED RESIDUE	UNP P57765
Z	55	MSE	MET	MODIFIED RESIDUE	UNP P57765
Z	75	MSE	MET	MODIFIED RESIDUE	UNP P57765
Z	230	MSE	MET	MODIFIED RESIDUE	UNP P57765
Z	245	LEU	-	CLONING ARTIFACT	UNP P57765
Z	246	GLU	-	CLONING ARTIFACT	UNP P57765
Z	247	HIS	-	EXPRESSION TAG	UNP P57765
Z	248	HIS	-	EXPRESSION TAG	UNP P57765
Z	249	HIS	-	EXPRESSION TAG	UNP P57765
Z	250	HIS	-	EXPRESSION TAG	UNP P57765
Z	251	HIS	-	EXPRESSION TAG	UNP P57765
Z	252	HIS	-	EXPRESSION TAG	UNP P57765

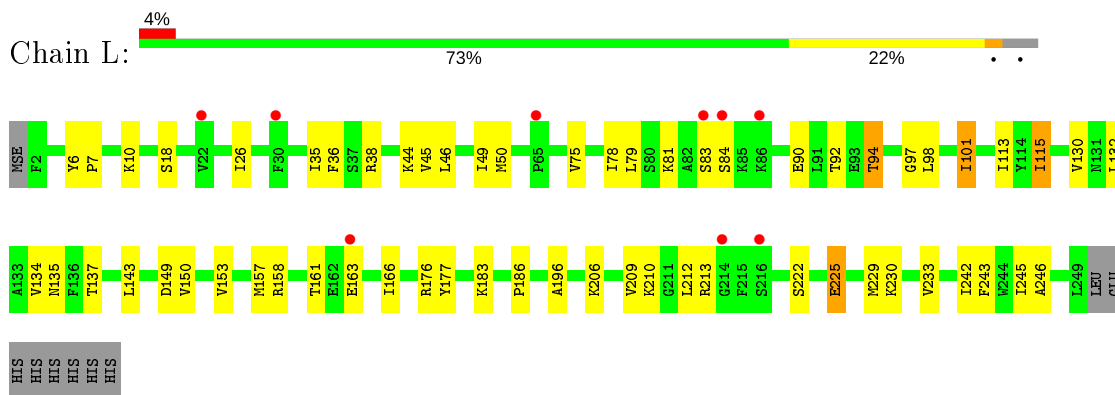
3 Residue-property plots [i](#)

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.

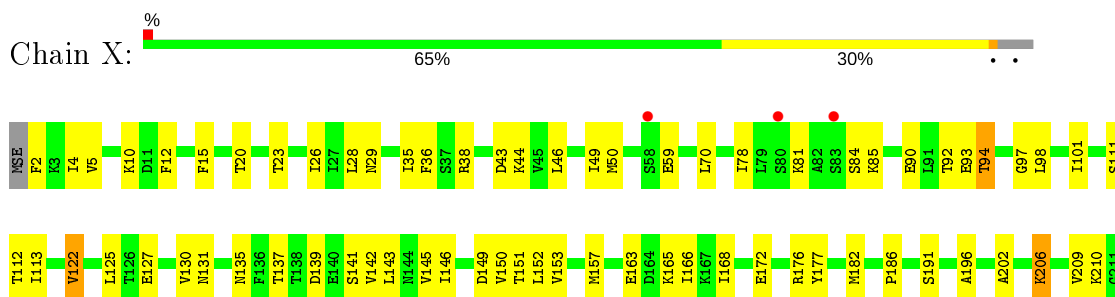
- Molecule 1: PCNA1 (SSO0397)



- Molecule 1: PCNA1 (SSO0397)

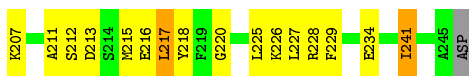
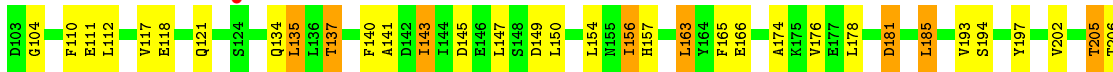


- Molecule 1: PCNA1 (SSO0397)

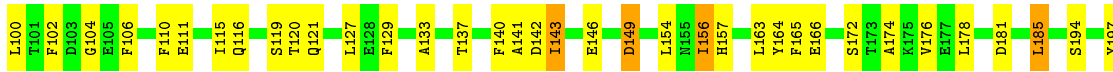




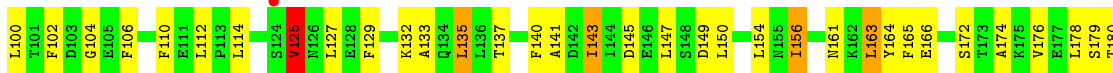
- Molecule 2: PCNA2 (SSO1047)



- Molecule 2: PCNA2 (SSO1047)

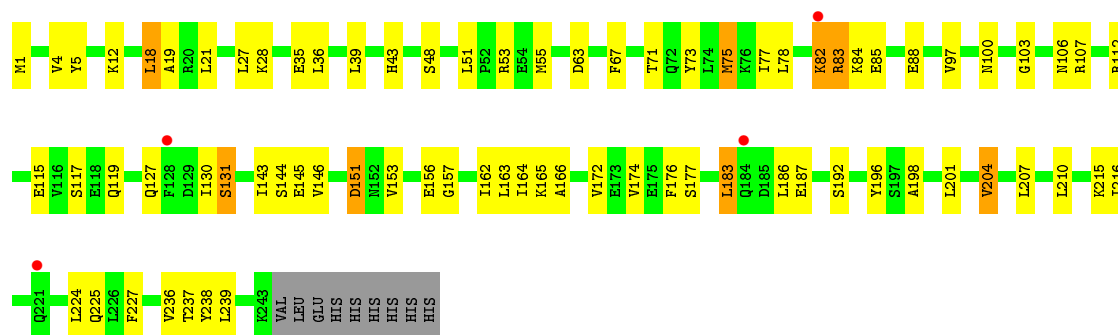


- Molecule 2: PCNA2 (SSO1047)

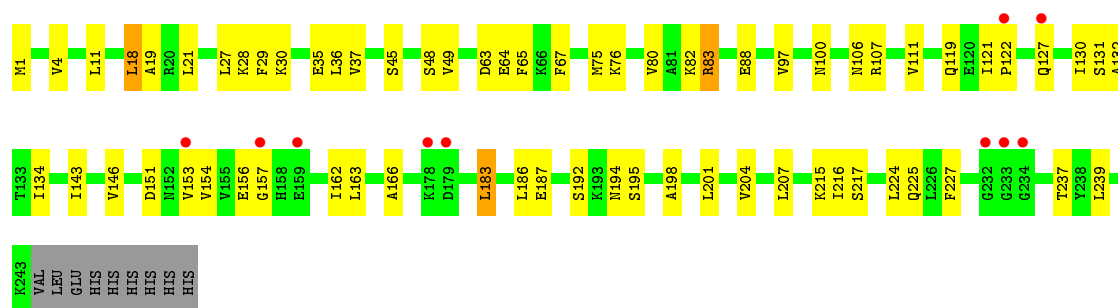


- Molecule 3: PCNA3 (SSO0405)

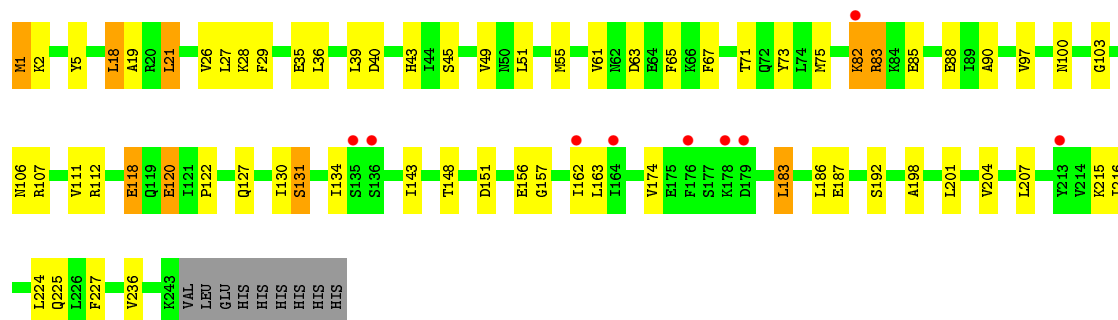




- Molecule 3: PCNA3 (SSO0405)



- Molecule 3: PCNA3 (SSO0405)



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	142.28Å 223.09Å 79.14Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 3.30 47.80 – 3.22	Depositor EDS
% Data completeness (in resolution range)	88.8 (20.00-3.30) 84.9 (47.80-3.22)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.21 (at 3.25Å)	Xtrriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.236 , 0.299 0.231 , 0.290	Depositor DCC
R_{free} test set	1805 reflections (5.12%)	wwPDB-VP
Wilson B-factor (Å ²)	67.3	Xtrriage
Anisotropy	0.445	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 51.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.27$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	17311	wwPDB-VP
Average B, all atoms (Å ²)	71.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.66% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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	0/1944	0.53	0/2607
1	L	0.34	0/1939	0.51	0/2600
1	X	0.34	0/1939	0.52	0/2600
2	B	0.36	0/1959	0.58	0/2643
2	M	0.34	0/1954	0.56	0/2636
2	Y	0.36	0/1954	0.58	1/2636 (0.0%)
3	C	0.36	0/1949	0.53	0/2622
3	N	0.33	0/1949	0.51	0/2622
3	Z	0.33	0/1949	0.51	0/2622
All	All	0.35	0/17536	0.54	1/23588 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
2	Y	217	LEU	CA-CB-CG	5.33	127.56	115.30

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	1924	0	1966	49	0
1	L	1919	0	1964	28	0
1	X	1919	0	1964	49	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	1927	0	1925	56	0
2	M	1922	0	1920	66	0
2	Y	1922	0	1920	68	0
3	C	1926	0	1936	42	0
3	N	1926	0	1936	36	0
3	Z	1926	0	1936	40	0
All	All	17311	0	17467	402	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 12.

The worst 5 of 402 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)
2:M:137:THR:HA	2:M:215:MSE:HE3	1.37	1.06
2:Y:137:THR:HA	2:Y:215:MSE:HE3	1.43	0.97
1:A:149:ASP:HB3	2:B:82:ARG:HE	1.29	0.97
2:B:137:THR:HA	2:B:215:MSE:HE3	1.45	0.96
2:B:218:TYR:HB2	2:B:226:LYS:HB3	1.54	0.90

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	247/257 (96%)	230 (93%)	15 (6%)	2 (1%)	19 51
1	L	246/257 (96%)	233 (95%)	10 (4%)	3 (1%)	13 42
1	X	246/257 (96%)	229 (93%)	15 (6%)	2 (1%)	19 51
2	B	242/245 (99%)	225 (93%)	16 (7%)	1 (0%)	34 66
2	M	241/245 (98%)	222 (92%)	18 (8%)	1 (0%)	34 66

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	Y	241/245 (98%)	225 (93%)	14 (6%)	2 (1%)	19	51
3	C	241/252 (96%)	226 (94%)	14 (6%)	1 (0%)	34	66
3	N	241/252 (96%)	232 (96%)	8 (3%)	1 (0%)	34	66
3	Z	241/252 (96%)	229 (95%)	12 (5%)	0	100	100
All	All	2186/2262 (97%)	2051 (94%)	122 (6%)	13 (1%)	25	57

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	Y	125	VAL
1	A	85	LYS
1	X	163	GLU
3	C	117	SER
1	X	85	LYS

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	218/219 (100%)	200 (92%)	18 (8%)	11	36
1	L	218/219 (100%)	200 (92%)	18 (8%)	11	36
1	X	218/219 (100%)	202 (93%)	16 (7%)	14	41
2	B	218/215 (101%)	185 (85%)	33 (15%)	3	13
2	M	218/215 (101%)	192 (88%)	26 (12%)	5	21
2	Y	218/215 (101%)	187 (86%)	31 (14%)	3	15
3	C	219/224 (98%)	197 (90%)	22 (10%)	7	27
3	N	219/224 (98%)	205 (94%)	14 (6%)	17	46
3	Z	219/224 (98%)	204 (93%)	15 (7%)	16	44
All	All	1965/1974 (100%)	1772 (90%)	193 (10%)	8	29

5 of 193 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	L	166	ILE
2	M	185	LEU
3	Z	1	MSE
1	L	206	LYS
2	M	79	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
2	M	116	GLN
3	N	100	ASN
3	Z	43	HIS
3	C	110	ASN
2	Y	121	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](#)

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, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	241/257 (93%)	0.05	0 100 100	65, 71, 74, 76	0
1	L	240/257 (93%)	0.30	9 (3%) 40 37	65, 71, 74, 76	0
1	X	240/257 (93%)	0.13	3 (1%) 77 77	63, 71, 74, 76	0
2	B	240/245 (97%)	-0.01	1 (0%) 92 93	69, 71, 75, 81	0
2	M	239/245 (97%)	0.09	1 (0%) 92 93	69, 71, 76, 82	0
2	Y	239/245 (97%)	0.00	1 (0%) 92 93	69, 71, 75, 83	0
3	C	239/252 (94%)	0.06	4 (1%) 70 68	69, 71, 79, 85	0
3	N	239/252 (94%)	0.34	10 (4%) 36 34	69, 71, 80, 85	0
3	Z	239/252 (94%)	0.22	9 (3%) 40 37	69, 71, 81, 86	0
All	All	2156/2262 (95%)	0.13	38 (1%) 68 67	63, 71, 75, 86	0

The worst 5 of 38 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	Z	162	ILE	3.6
2	B	124	SER	3.5
3	C	128	PHE	3.3
3	Z	82	LYS	3.2
3	N	159	GLU	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

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.