



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 05:37 AM GMT

PDB ID : 2R5I
Title : Pentamer Structure of Major Capsid Protein L1 of Human Papilloma Virus type 18
Authors : Bishop, B.; Dasgupta, J.; Chen, X.S.
Deposited on : 2007-09-03
Resolution : 3.40 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
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

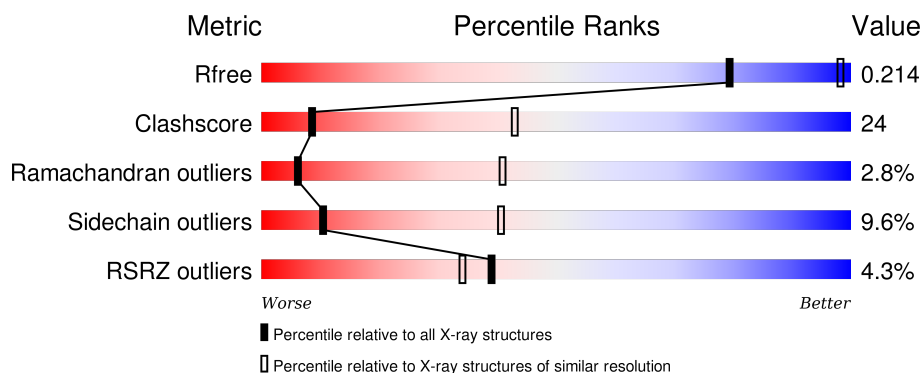
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.40 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}	91344	1476 (3.50-3.30)
Clashscore	102246	1611 (3.50-3.30)
Ramachandran outliers	100387	1571 (3.50-3.30)
Sidechain outliers	100360	1571 (3.50-3.30)
RSRZ outliers	91569	1485 (3.50-3.30)

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. 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	428	<div> <div>6%</div> <div> <div>46%</div> <div>38%</div> <div>14%</div> <div>..</div> </div> </div>
1	B	428	<div> <div>4%</div> <div> <div>46%</div> <div>39%</div> <div>13%</div> <div>..</div> </div> </div>
1	C	428	<div> <div>5%</div> <div> <div>48%</div> <div>38%</div> <div>12%</div> <div>..</div> </div> </div>
1	D	428	<div> <div>2%</div> <div> <div>46%</div> <div>40%</div> <div>11%</div> <div>..</div> </div> </div>
1	E	428	<div> <div>3%</div> <div> <div>46%</div> <div>39%</div> <div>13%</div> <div>..</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	428	
1	G	428	
1	H	428	
1	I	428	
1	J	428	
1	K	428	
1	L	428	
1	M	428	
1	N	428	
1	O	428	

2 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 49845 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 L1 protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	423	Total	C	N	O	S	0	0	0
			3323	2098	562	643	20			
1	B	423	Total	C	N	O	S	0	0	0
			3323	2098	562	643	20			
1	C	423	Total	C	N	O	S	0	0	0
			3323	2098	562	643	20			
1	D	423	Total	C	N	O	S	0	0	0
			3323	2098	562	643	20			
1	E	423	Total	C	N	O	S	0	0	0
			3323	2098	562	643	20			
1	F	423	Total	C	N	O	S	0	0	0
			3323	2098	562	643	20			
1	G	423	Total	C	N	O	S	0	0	0
			3323	2098	562	643	20			
1	H	423	Total	C	N	O	S	0	0	0
			3323	2098	562	643	20			
1	I	423	Total	C	N	O	S	0	0	0
			3323	2098	562	643	20			
1	J	423	Total	C	N	O	S	0	0	0
			3323	2098	562	643	20			
1	K	423	Total	C	N	O	S	0	0	0
			3323	2098	562	643	20			
1	L	423	Total	C	N	O	S	0	0	0
			3323	2098	562	643	20			
1	M	423	Total	C	N	O	S	0	0	0
			3323	2098	562	643	20			
1	N	423	Total	C	N	O	S	0	0	0
			3323	2098	562	643	20			
1	O	423	Total	C	N	O	S	0	0	0
			3323	2098	562	643	20			

There are 135 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	20	ALA	-	EXPRESSION TAG	UNP Q80B70
A	47	ASP	ASN	ENGINEERED	UNP Q80B70
A	175	SER	CYS	ENGINEERED	UNP Q80B70
A	393	GLN	HIS	ENGINEERED	UNP Q80B70
A	405	GLY	-	LINKER	UNP Q80B70
A	406	GLY	-	LINKER	UNP Q80B70
A	407	SER	-	LINKER	UNP Q80B70
A	408	GLY	-	LINKER	UNP Q80B70
A	409	GLY	-	LINKER	UNP Q80B70
B	20	ALA	-	EXPRESSION TAG	UNP Q80B70
B	47	ASP	ASN	ENGINEERED	UNP Q80B70
B	175	SER	CYS	ENGINEERED	UNP Q80B70
B	393	GLN	HIS	ENGINEERED	UNP Q80B70
B	405	GLY	-	LINKER	UNP Q80B70
B	406	GLY	-	LINKER	UNP Q80B70
B	407	SER	-	LINKER	UNP Q80B70
B	408	GLY	-	LINKER	UNP Q80B70
B	409	GLY	-	LINKER	UNP Q80B70
C	20	ALA	-	EXPRESSION TAG	UNP Q80B70
C	47	ASP	ASN	ENGINEERED	UNP Q80B70
C	175	SER	CYS	ENGINEERED	UNP Q80B70
C	393	GLN	HIS	ENGINEERED	UNP Q80B70
C	405	GLY	-	LINKER	UNP Q80B70
C	406	GLY	-	LINKER	UNP Q80B70
C	407	SER	-	LINKER	UNP Q80B70
C	408	GLY	-	LINKER	UNP Q80B70
C	409	GLY	-	LINKER	UNP Q80B70
D	20	ALA	-	EXPRESSION TAG	UNP Q80B70
D	47	ASP	ASN	ENGINEERED	UNP Q80B70
D	175	SER	CYS	ENGINEERED	UNP Q80B70
D	393	GLN	HIS	ENGINEERED	UNP Q80B70
D	405	GLY	-	LINKER	UNP Q80B70
D	406	GLY	-	LINKER	UNP Q80B70
D	407	SER	-	LINKER	UNP Q80B70
D	408	GLY	-	LINKER	UNP Q80B70
D	409	GLY	-	LINKER	UNP Q80B70
E	20	ALA	-	EXPRESSION TAG	UNP Q80B70
E	47	ASP	ASN	ENGINEERED	UNP Q80B70
E	175	SER	CYS	ENGINEERED	UNP Q80B70
E	393	GLN	HIS	ENGINEERED	UNP Q80B70
E	405	GLY	-	LINKER	UNP Q80B70
E	406	GLY	-	LINKER	UNP Q80B70
E	407	SER	-	LINKER	UNP Q80B70

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	408	GLY	-	LINKER	UNP Q80B70
E	409	GLY	-	LINKER	UNP Q80B70
F	20	ALA	-	EXPRESSION TAG	UNP Q80B70
F	47	ASP	ASN	ENGINEERED	UNP Q80B70
F	175	SER	CYS	ENGINEERED	UNP Q80B70
F	393	GLN	HIS	ENGINEERED	UNP Q80B70
F	405	GLY	-	LINKER	UNP Q80B70
F	406	GLY	-	LINKER	UNP Q80B70
F	407	SER	-	LINKER	UNP Q80B70
F	408	GLY	-	LINKER	UNP Q80B70
F	409	GLY	-	LINKER	UNP Q80B70
G	20	ALA	-	EXPRESSION TAG	UNP Q80B70
G	47	ASP	ASN	ENGINEERED	UNP Q80B70
G	175	SER	CYS	ENGINEERED	UNP Q80B70
G	393	GLN	HIS	ENGINEERED	UNP Q80B70
G	405	GLY	-	LINKER	UNP Q80B70
G	406	GLY	-	LINKER	UNP Q80B70
G	407	SER	-	LINKER	UNP Q80B70
G	408	GLY	-	LINKER	UNP Q80B70
G	409	GLY	-	LINKER	UNP Q80B70
H	20	ALA	-	EXPRESSION TAG	UNP Q80B70
H	47	ASP	ASN	ENGINEERED	UNP Q80B70
H	175	SER	CYS	ENGINEERED	UNP Q80B70
H	393	GLN	HIS	ENGINEERED	UNP Q80B70
H	405	GLY	-	LINKER	UNP Q80B70
H	406	GLY	-	LINKER	UNP Q80B70
H	407	SER	-	LINKER	UNP Q80B70
H	408	GLY	-	LINKER	UNP Q80B70
H	409	GLY	-	LINKER	UNP Q80B70
I	20	ALA	-	EXPRESSION TAG	UNP Q80B70
I	47	ASP	ASN	ENGINEERED	UNP Q80B70
I	175	SER	CYS	ENGINEERED	UNP Q80B70
I	393	GLN	HIS	ENGINEERED	UNP Q80B70
I	405	GLY	-	LINKER	UNP Q80B70
I	406	GLY	-	LINKER	UNP Q80B70
I	407	SER	-	LINKER	UNP Q80B70
I	408	GLY	-	LINKER	UNP Q80B70
I	409	GLY	-	LINKER	UNP Q80B70
J	20	ALA	-	EXPRESSION TAG	UNP Q80B70
J	47	ASP	ASN	ENGINEERED	UNP Q80B70
J	175	SER	CYS	ENGINEERED	UNP Q80B70
J	393	GLN	HIS	ENGINEERED	UNP Q80B70

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
J	405	GLY	-	LINKER	UNP Q80B70
J	406	GLY	-	LINKER	UNP Q80B70
J	407	SER	-	LINKER	UNP Q80B70
J	408	GLY	-	LINKER	UNP Q80B70
J	409	GLY	-	LINKER	UNP Q80B70
K	20	ALA	-	EXPRESSION TAG	UNP Q80B70
K	47	ASP	ASN	ENGINEERED	UNP Q80B70
K	175	SER	CYS	ENGINEERED	UNP Q80B70
K	393	GLN	HIS	ENGINEERED	UNP Q80B70
K	405	GLY	-	LINKER	UNP Q80B70
K	406	GLY	-	LINKER	UNP Q80B70
K	407	SER	-	LINKER	UNP Q80B70
K	408	GLY	-	LINKER	UNP Q80B70
K	409	GLY	-	LINKER	UNP Q80B70
L	20	ALA	-	EXPRESSION TAG	UNP Q80B70
L	47	ASP	ASN	ENGINEERED	UNP Q80B70
L	175	SER	CYS	ENGINEERED	UNP Q80B70
L	393	GLN	HIS	ENGINEERED	UNP Q80B70
L	405	GLY	-	LINKER	UNP Q80B70
L	406	GLY	-	LINKER	UNP Q80B70
L	407	SER	-	LINKER	UNP Q80B70
L	408	GLY	-	LINKER	UNP Q80B70
L	409	GLY	-	LINKER	UNP Q80B70
M	20	ALA	-	EXPRESSION TAG	UNP Q80B70
M	47	ASP	ASN	ENGINEERED	UNP Q80B70
M	175	SER	CYS	ENGINEERED	UNP Q80B70
M	393	GLN	HIS	ENGINEERED	UNP Q80B70
M	405	GLY	-	LINKER	UNP Q80B70
M	406	GLY	-	LINKER	UNP Q80B70
M	407	SER	-	LINKER	UNP Q80B70
M	408	GLY	-	LINKER	UNP Q80B70
M	409	GLY	-	LINKER	UNP Q80B70
N	20	ALA	-	EXPRESSION TAG	UNP Q80B70
N	47	ASP	ASN	ENGINEERED	UNP Q80B70
N	175	SER	CYS	ENGINEERED	UNP Q80B70
N	393	GLN	HIS	ENGINEERED	UNP Q80B70
N	405	GLY	-	LINKER	UNP Q80B70
N	406	GLY	-	LINKER	UNP Q80B70
N	407	SER	-	LINKER	UNP Q80B70
N	408	GLY	-	LINKER	UNP Q80B70
N	409	GLY	-	LINKER	UNP Q80B70
O	20	ALA	-	EXPRESSION TAG	UNP Q80B70

Continued on next page...

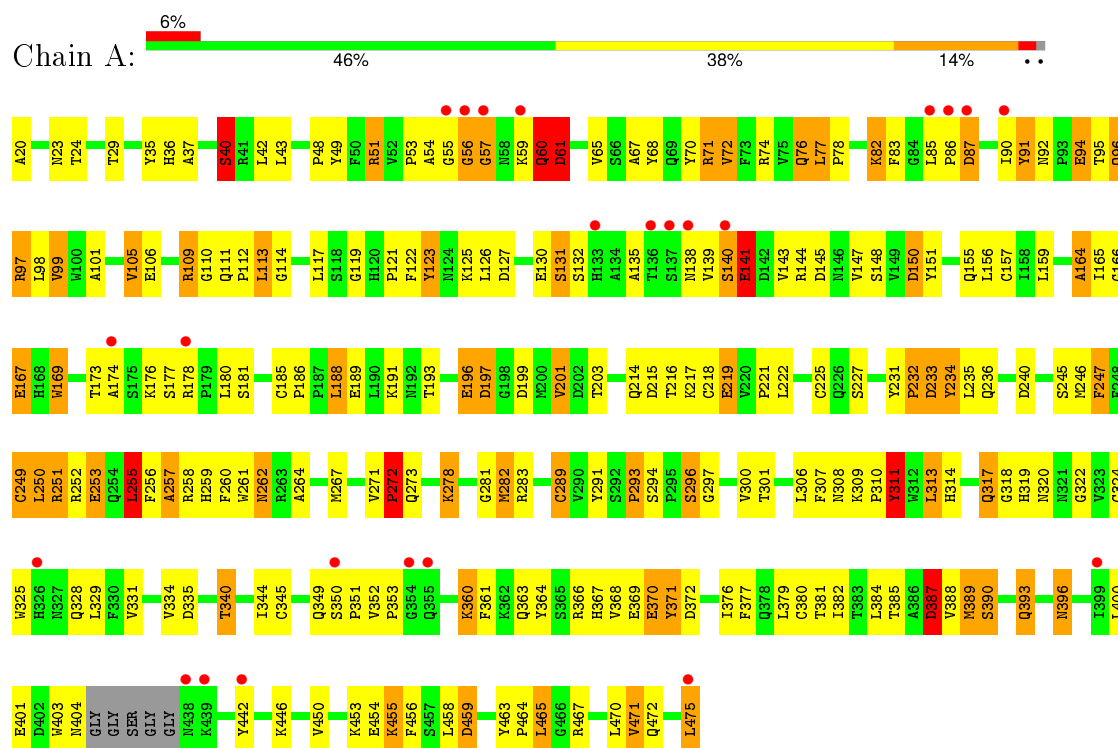
Continued from previous page...

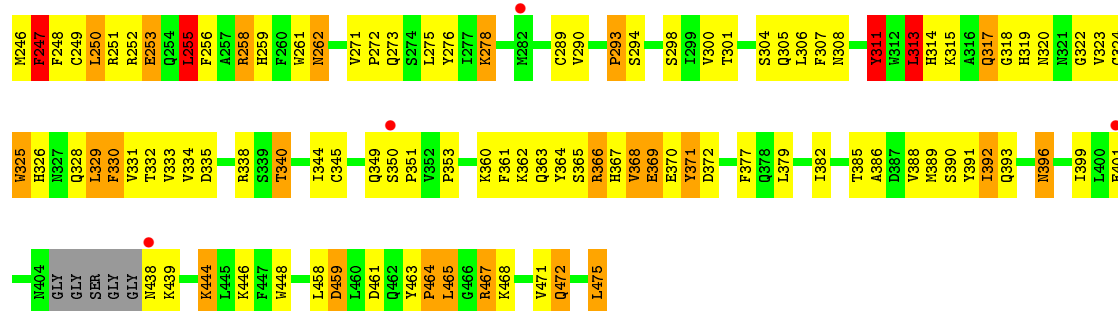
Chain	Residue	Modelled	Actual	Comment	Reference
O	47	ASP	ASN	ENGINEERED	UNP Q80B70
O	175	SER	CYS	ENGINEERED	UNP Q80B70
O	393	GLN	HIS	ENGINEERED	UNP Q80B70
O	405	GLY	-	LINKER	UNP Q80B70
O	406	GLY	-	LINKER	UNP Q80B70
O	407	SER	-	LINKER	UNP Q80B70
O	408	GLY	-	LINKER	UNP Q80B70
O	409	GLY	-	LINKER	UNP Q80B70

3 Residue-property plots

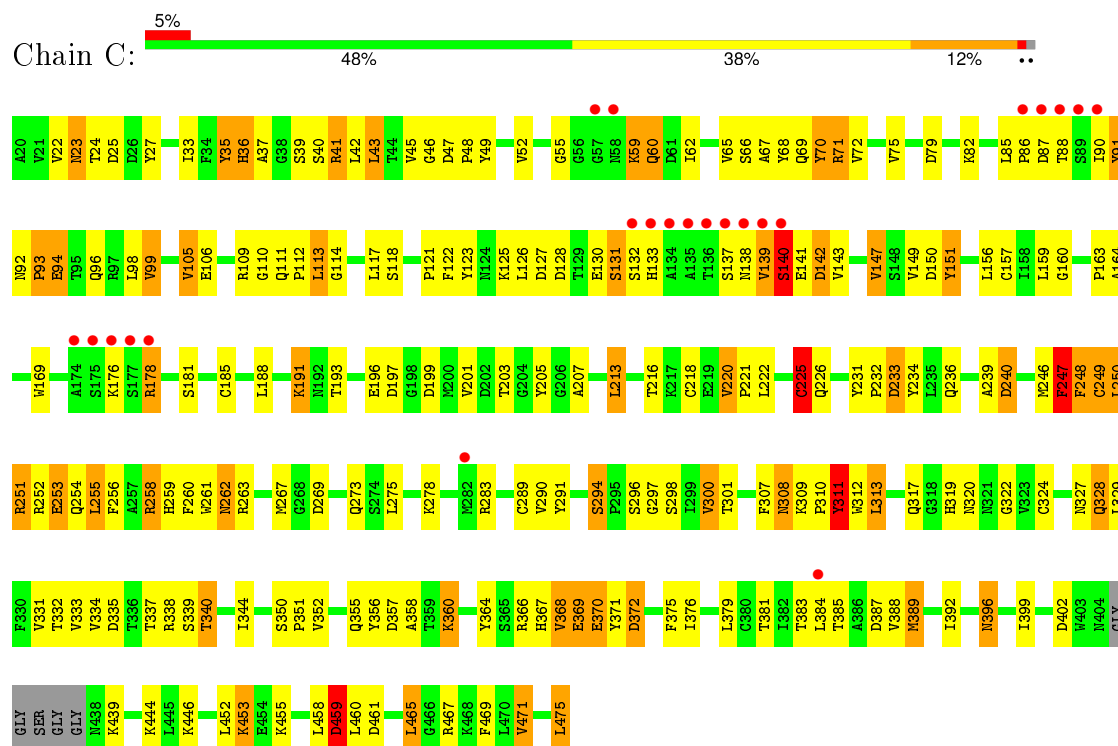
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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: L1 protein

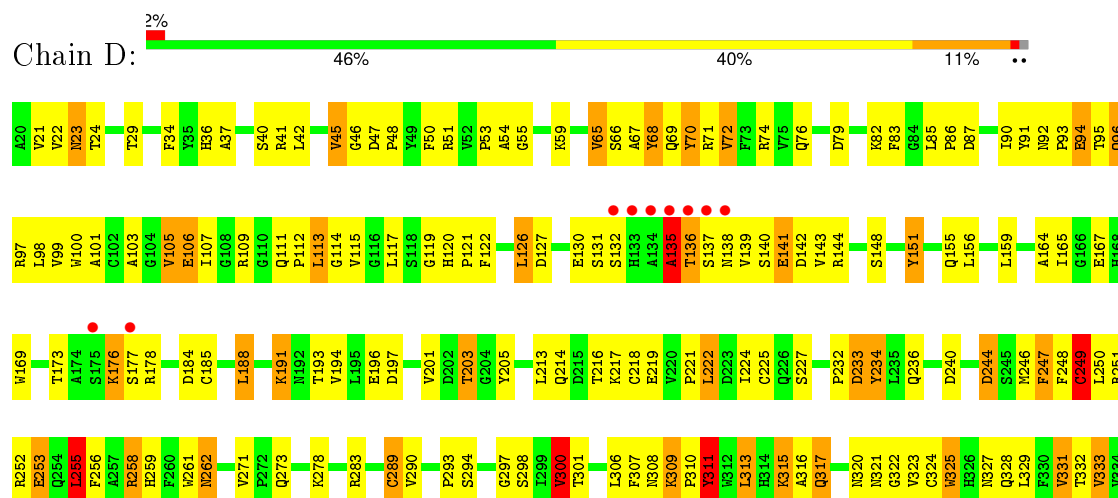


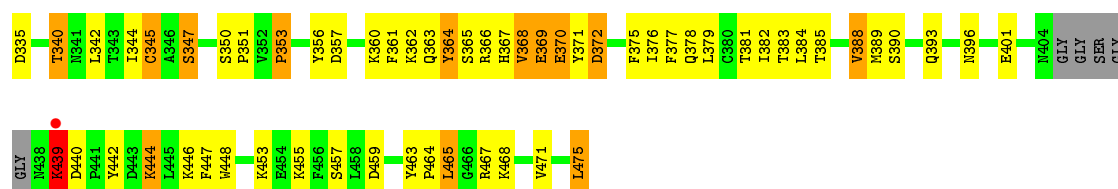


• Molecule 1: L1 protein



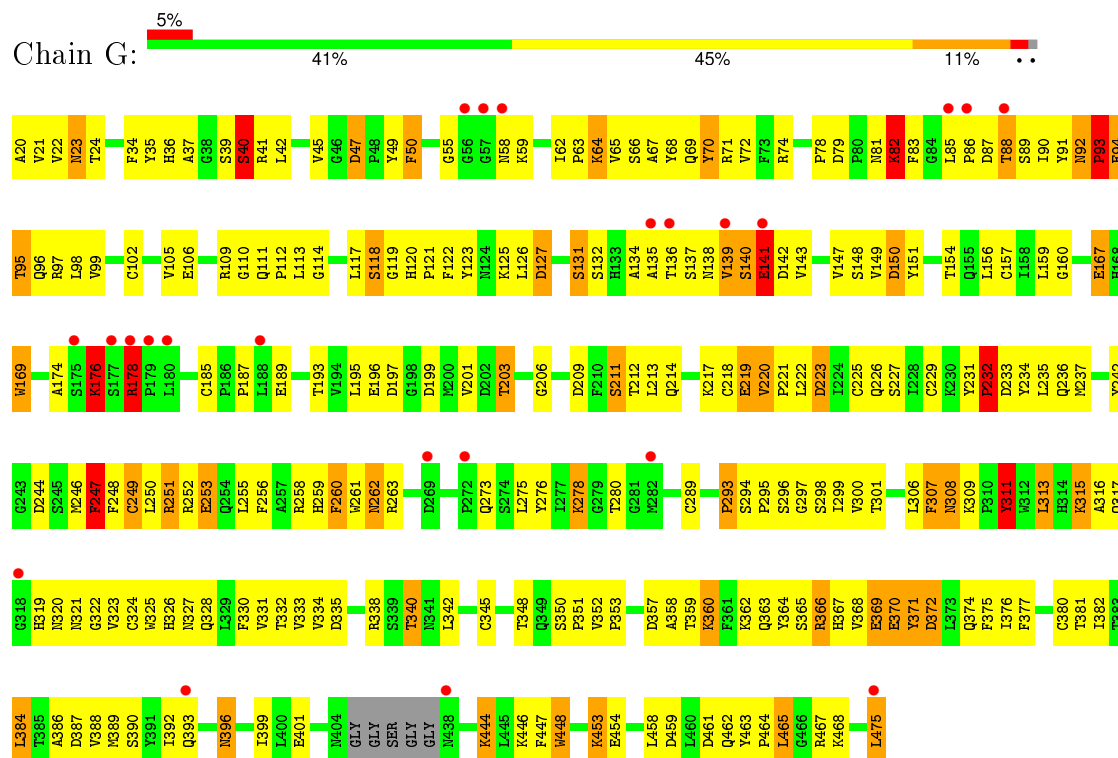
• Molecule 1: L1 protein



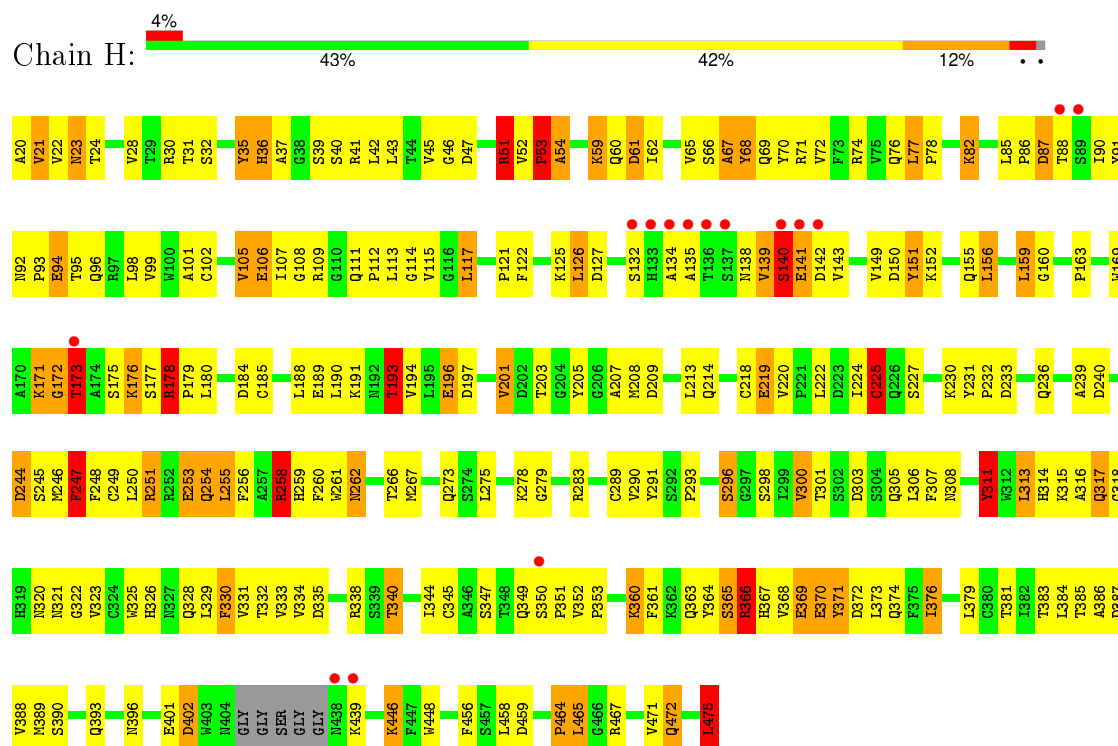




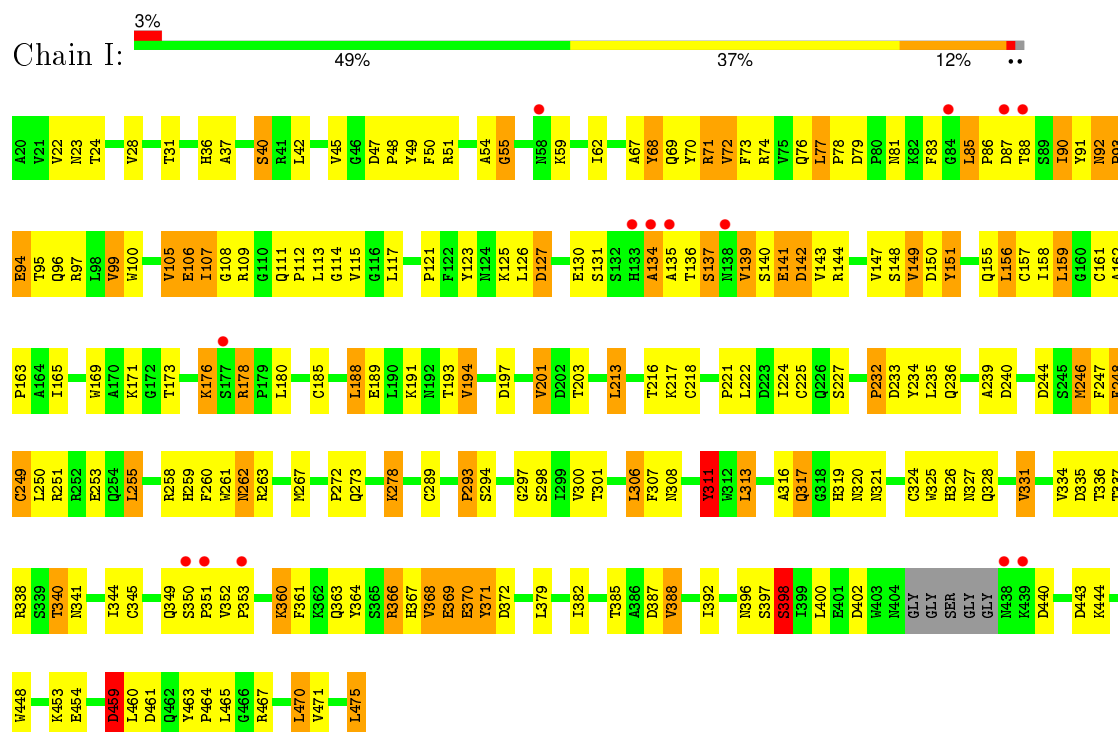
• Molecule 1: L1 protein



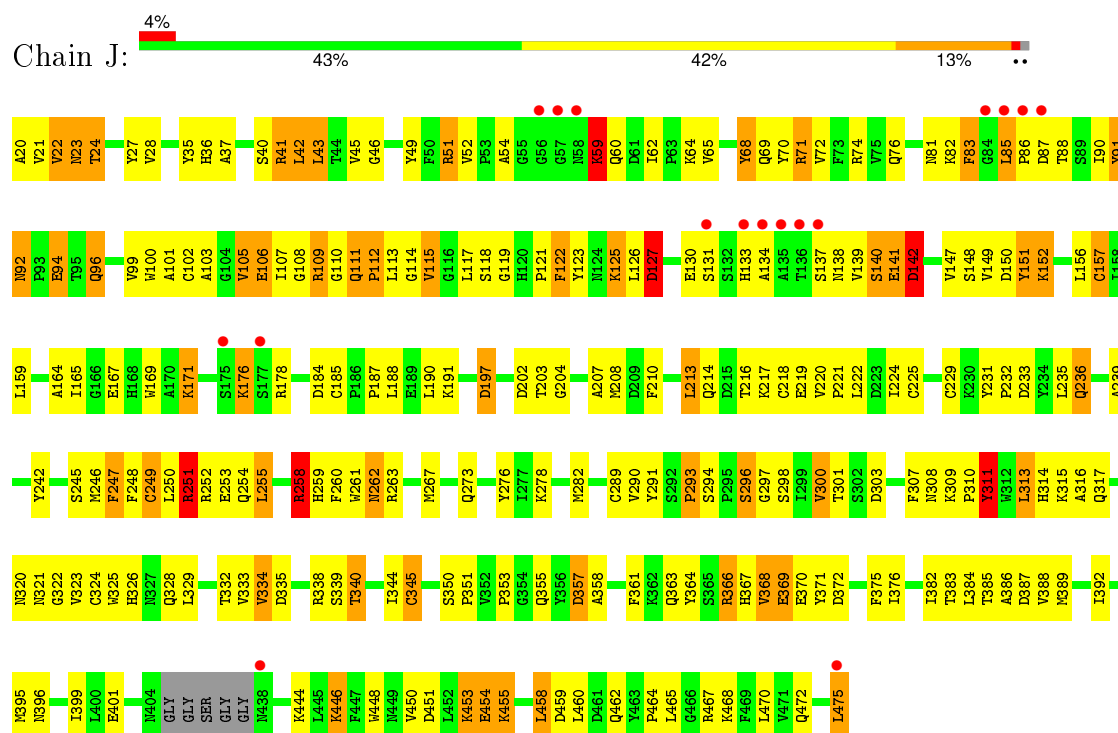
• Molecule 1: L1 protein



- Molecule 1: L1 protein

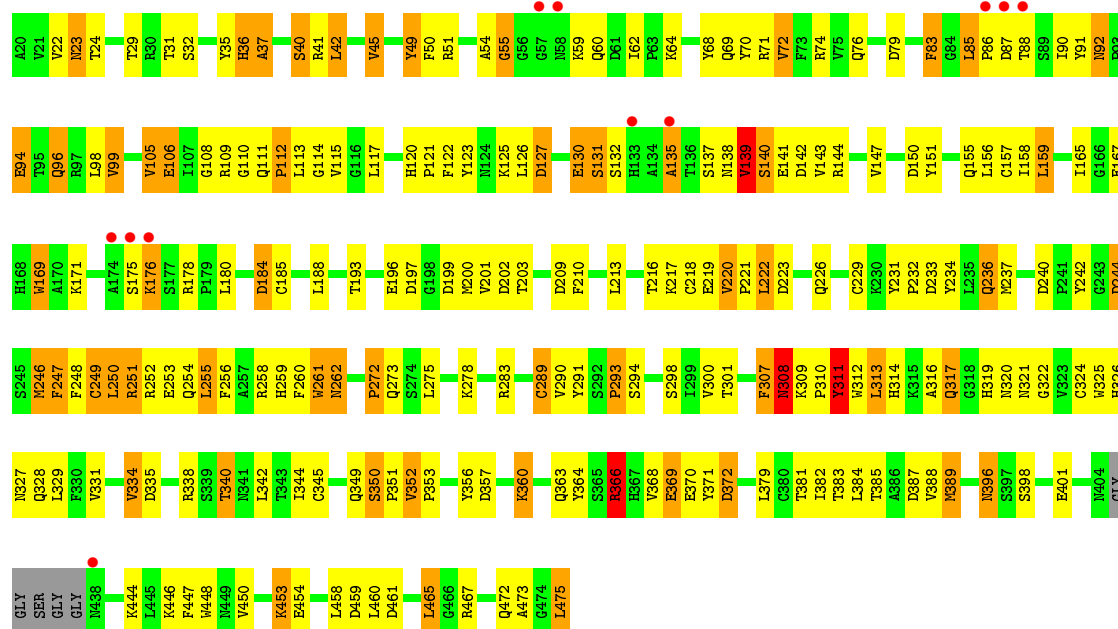


- Molecule 1: L1 protein

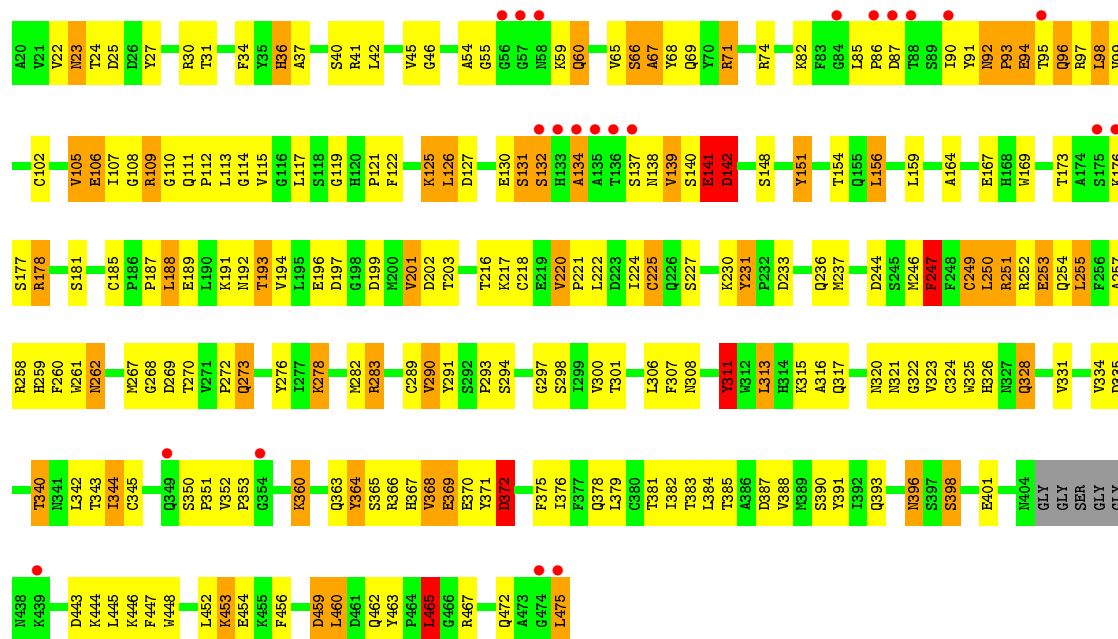


- Molecule 1: L1 protein





• Molecule 1: L1 protein



• Molecule 1: L1 protein



SER	GLY	GLY	N438	K446	F447	W448	K453	E454	K455	D459	L460	D461	L465	K468	G474	L475	T332	T333	Y334	T340	N341	L342	T343	I344	C345	A346	S347	S350	P351	Y352	P353	K360	F361	K362	Q363	Y364	S365	R366	H367	Y368	E369	E370	Y371	D372	F375	I376	F377	Q378	T381	I382	T383	L384	D387	Y388	H389	S390	T391	I392	Q393	N396	S397	S398	I399	L400	E401	M404	GLY	GLY	P179	L180	G183	D184	C185	F186	P187	L188	E189	L190	K191	N192	T193	E196	D197	V201	D202	T203	G204	Y205	M208	L213	Q214	D215	T216	K217	C218	E219	V220	P221	L222	C226	C229	K230	Y231	P232	D233	Y234	L235	Q236	M237	Y242	G243	D244	S245	M246	F247	F248	C249	L250	R251	R252	E253	Q254	L255	F256	A257	R258	H259	F260	W261	N262	R263	L264	Q273	S274	K278	G279	M282	R283	A284	S285	C289	Y290	Y291	S292	P293	S294	P295	S296	G297	S298	I299	V300	T301	F307	N308	Y311	W312	I313	H314	K315	A316	Q317	N320	N321	C324	W325	H326	M327	Q328	L329	F330	V331
-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	81.16Å 105.04Å 235.75Å 87.56° 85.66° 68.16°	Depositor
Resolution (Å)	40.00 – 3.40 51.09 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.7 (40.00-3.40) 98.8 (51.09-2.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	19.60	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.34 (at 2.91Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.214 , 0.260 0.203 , 0.214	Depositor DCC
R_{free} test set	7362 reflections (8.06%)	DCC
Wilson B-factor (Å ²)	63.4	Xtriage
Anisotropy	0.374	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 60.1	EDS
Estimated twinning fraction	0.000 for -h,-h+k,-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.52$, $\langle L^2 \rangle = 0.35$	Xtriage
Outliers	0 of 158695 reflections	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	49845	wwPDB-VP
Average B, all atoms (Å ²)	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.88% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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	1.92	72/3408 (2.1%)	1.47	35/4633 (0.8%)
1	B	1.93	78/3408 (2.3%)	1.44	36/4633 (0.8%)
1	C	1.89	64/3408 (1.9%)	1.47	38/4633 (0.8%)
1	D	1.96	79/3408 (2.3%)	1.44	29/4633 (0.6%)
1	E	1.99	80/3408 (2.3%)	1.47	41/4633 (0.9%)
1	F	2.02	90/3408 (2.6%)	1.46	34/4633 (0.7%)
1	G	2.03	90/3408 (2.6%)	1.48	33/4633 (0.7%)
1	H	1.93	71/3408 (2.1%)	1.52	46/4633 (1.0%)
1	I	1.84	46/3408 (1.3%)	1.42	34/4633 (0.7%)
1	J	1.94	71/3408 (2.1%)	1.50	41/4633 (0.9%)
1	K	1.89	66/3408 (1.9%)	1.45	41/4633 (0.9%)
1	L	1.99	81/3408 (2.4%)	1.51	42/4633 (0.9%)
1	M	1.91	70/3408 (2.1%)	1.44	33/4633 (0.7%)
1	N	1.90	72/3408 (2.1%)	1.43	31/4633 (0.7%)
1	O	1.95	80/3408 (2.3%)	1.51	37/4633 (0.8%)
All	All	1.94	1110/51120 (2.2%)	1.47	551/69495 (0.8%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	D	0	1

The worst 5 of 1110 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	439	LYS	CE-NZ	16.04	1.89	1.49
1	L	311	TYR	CE2-CZ	15.05	1.58	1.38
1	O	311	TYR	CE2-CZ	15.03	1.58	1.38
1	E	311	TYR	CE2-CZ	14.64	1.57	1.38
1	I	311	TYR	CE2-CZ	14.61	1.57	1.38

The worst 5 of 551 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	L	109	ARG	NE-CZ-NH1	-17.51	111.55	120.30
1	F	263	ARG	NE-CZ-NH2	-16.83	111.89	120.30
1	L	109	ARG	NE-CZ-NH2	16.15	128.38	120.30
1	L	366	ARG	NE-CZ-NH2	-13.69	113.45	120.30
1	O	252	ARG	NE-CZ-NH2	-13.15	113.72	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	53	PRO	Peptide

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3323	0	3192	188	1
1	B	3323	0	3192	170	0
1	C	3323	0	3192	154	1
1	D	3323	0	3192	151	0
1	E	3323	0	3192	161	0
1	F	3323	0	3192	199	0
1	G	3323	0	3192	197	0
1	H	3323	0	3192	202	0
1	I	3323	0	3192	165	0
1	J	3323	0	3192	166	0
1	K	3323	0	3192	152	0
1	L	3323	0	3192	175	0
1	M	3323	0	3192	166	0
1	N	3323	0	3192	174	0
1	O	3323	0	3192	157	0
All	All	49845	0	47880	2325	1

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 24.

The worst 5 of 2325 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:J:446:LYS:CD	1:J:446:LYS:CG	1.75	1.63
1:B:141:GLU:CG	1:B:141:GLU:CB	1.75	1.63
1:C:191:LYS:CE	1:C:191:LYS:CD	1.76	1.62
1:L:360:LYS:CD	1:L:360:LYS:CE	1.78	1.61
1:F:465:LEU:CD1	1:F:465:LEU:CG	1.74	1.61

All (1) 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:A:387:ASP:OD2	1:C:360:LYS:NZ[1_565]	2.12	0.08

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	419/428 (98%)	372 (89%)	34 (8%)	13 (3%)	5	39
1	B	419/428 (98%)	370 (88%)	34 (8%)	15 (4%)	4	34
1	C	419/428 (98%)	372 (89%)	36 (9%)	11 (3%)	7	42
1	D	419/428 (98%)	373 (89%)	38 (9%)	8 (2%)	10	49
1	E	419/428 (98%)	373 (89%)	34 (8%)	12 (3%)	6	40
1	F	419/428 (98%)	367 (88%)	43 (10%)	9 (2%)	9	47
1	G	419/428 (98%)	366 (87%)	42 (10%)	11 (3%)	7	42
1	H	419/428 (98%)	362 (86%)	46 (11%)	11 (3%)	7	42
1	I	419/428 (98%)	373 (89%)	32 (8%)	14 (3%)	5	37
1	J	419/428 (98%)	369 (88%)	39 (9%)	11 (3%)	7	42
1	K	419/428 (98%)	370 (88%)	39 (9%)	10 (2%)	7	44

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	L	419/428 (98%)	369 (88%)	36 (9%)	14 (3%)	5	37
1	M	419/428 (98%)	367 (88%)	37 (9%)	15 (4%)	4	34
1	N	419/428 (98%)	376 (90%)	30 (7%)	13 (3%)	5	39
1	O	419/428 (98%)	372 (89%)	38 (9%)	9 (2%)	9	47
All	All	6285/6420 (98%)	5551 (88%)	558 (9%)	176 (3%)	6	41

5 of 176 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	61	ASP
1	A	131	SER
1	A	135	ALA
1	B	59	LYS
1	B	131	SER

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	370/371 (100%)	331 (90%)	39 (10%)	8	36
1	B	370/371 (100%)	339 (92%)	31 (8%)	14	49
1	C	370/371 (100%)	334 (90%)	36 (10%)	10	40
1	D	370/371 (100%)	337 (91%)	33 (9%)	12	45
1	E	370/371 (100%)	337 (91%)	33 (9%)	12	45
1	F	370/371 (100%)	332 (90%)	38 (10%)	9	37
1	G	370/371 (100%)	334 (90%)	36 (10%)	10	40
1	H	370/371 (100%)	335 (90%)	35 (10%)	11	41
1	I	370/371 (100%)	337 (91%)	33 (9%)	12	45
1	J	370/371 (100%)	335 (90%)	35 (10%)	11	41
1	K	370/371 (100%)	331 (90%)	39 (10%)	8	36
1	L	370/371 (100%)	330 (89%)	40 (11%)	8	35

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	M	370/371 (100%)	335 (90%)	35 (10%)	11	41
1	N	370/371 (100%)	333 (90%)	37 (10%)	9	38
1	O	370/371 (100%)	338 (91%)	32 (9%)	13	48
All	All	5550/5565 (100%)	5018 (90%)	532 (10%)	10	40

5 of 532 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	372	ASP
1	I	251	ARG
1	N	355	GLN
1	H	36	HIS
1	H	317	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 221 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	462	GLN
1	I	308	ASN
1	N	396	ASN
1	H	138	ASN
1	H	363	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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 [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

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, 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	423/428 (98%)	0.08	24 (5%)	27	25	16, 46, 103, 121	0
1	B	423/428 (98%)	-0.05	19 (4%)	37	33	19, 45, 101, 127	0
1	C	423/428 (98%)	0.10	23 (5%)	29	27	15, 45, 103, 123	0
1	D	423/428 (98%)	-0.13	10 (2%)	62	57	14, 44, 102, 126	0
1	E	423/428 (98%)	-0.17	11 (2%)	59	54	15, 45, 104, 123	0
1	F	423/428 (98%)	0.22	29 (6%)	20	19	19, 50, 105, 123	0
1	G	423/428 (98%)	0.18	23 (5%)	29	27	22, 50, 103, 127	0
1	H	423/428 (98%)	-0.07	15 (3%)	48	42	19, 46, 103, 124	0
1	I	423/428 (98%)	-0.11	14 (3%)	50	45	17, 45, 101, 123	0
1	J	423/428 (98%)	-0.05	17 (4%)	42	37	18, 46, 103, 123	0
1	K	423/428 (98%)	-0.11	11 (2%)	59	54	18, 44, 103, 122	0
1	L	423/428 (98%)	0.01	22 (5%)	31	28	16, 46, 103, 122	0
1	M	423/428 (98%)	-0.01	20 (4%)	35	32	18, 46, 102, 126	0
1	N	423/428 (98%)	-0.09	19 (4%)	37	33	19, 45, 103, 125	0
1	O	423/428 (98%)	-0.01	17 (4%)	42	37	17, 46, 103, 126	0
All	All	6345/6420 (98%)	-0.01	274 (4%)	39	34	14, 46, 104, 127	0

The worst 5 of 274 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	135	ALA	8.1
1	D	133	HIS	7.8
1	M	57	GLY	7.7
1	O	134	ALA	7.4
1	N	86	PRO	7.4

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 [i](#)

There are no such residues in this entry.