



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 18, 2018 – 01:01 am GMT

PDB ID : 1JD2  
Title : Crystal Structure of the yeast 20S Proteasome:TMC-95A complex: A non-covalent Proteasome Inhibitor  
Authors : Groll, M.; Koguchi, Y.; Huber, R.; Kohno, J.  
Deposited on : 2001-06-12  
Resolution : 3.00 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  
Mogul : 1.7.3 (157068), CSD as539be (2018)  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30686

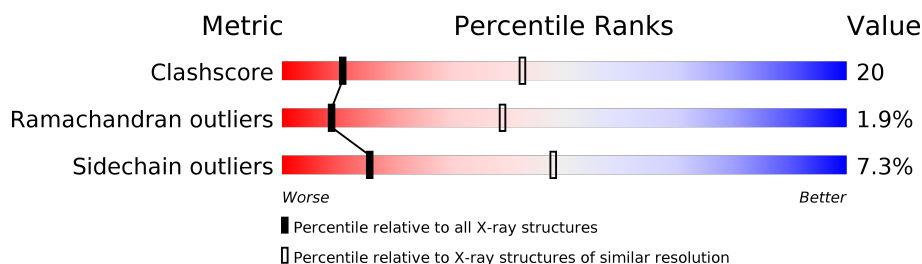
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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))
Clashscore	122078	2167 (3.00-3.00)
Ramachandran outliers	120005	2101 (3.00-3.00)
Sidechain outliers	119972	2104 (3.00-3.00)


















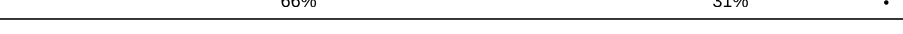

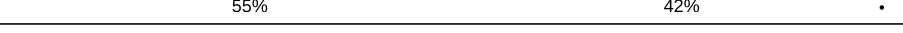



The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	250	
1	V	250	
2	B	244	
2	W	244	
3	C	241	
3	X	241	
4	D	242	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	Y	242	
5	E	233	
5	Z	233	
6	1	244	
6	F	244	
7	2	243	
7	G	243	
8	H	222	
8	O	222	
9	I	204	
9	P	204	
10	J	198	
10	Q	198	
11	K	212	
11	R	212	
12	L	222	
12	S	222	
13	M	233	
13	T	233	
14	N	196	
14	U	196	
15	8	5	
15	9	5	

## 2 Entry composition

There are 17 unique types of molecules in this entry. The entry contains 52549 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 PROTEASOME COMPONENT Y7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	250	Total	C	N	O	S	0	0	0
			1915	1219	315	377	4			
1	V	250	Total	C	N	O	S	0	0	0
			1915	1219	315	377	4			

- Molecule 2 is a protein called PROTEASOME COMPONENT Y13.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	244	Total	C	N	O	S	0	0	0
			1905	1201	321	380	3			
2	W	244	Total	C	N	O	S	0	0	0
			1905	1201	321	380	3			

- Molecule 3 is a protein called PROTEASOME COMPONENT PRE6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	241	Total	C	N	O	S	0	0	0
			1891	1181	331	375	4			
3	X	241	Total	C	N	O	S	0	0	0
			1891	1181	331	375	4			

- Molecule 4 is a protein called PROTEASOME COMPONENT PUP2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	242	Total	C	N	O	S	0	0	0
			1862	1162	314	379	7			
4	Y	242	Total	C	N	O	S	0	0	0
			1862	1162	314	379	7			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	121	LEU	-	INSERTION	UNP P32379
D	122	ARG	-	INSERTION	UNP P32379
D	123	PHE	-	INSERTION	UNP P32379
D	123A	GLY	-	INSERTION	UNP P32379
D	123B	GLU	-	INSERTION	UNP P32379
D	123C	GLY	-	INSERTION	UNP P32379
D	123E	SER	-	INSERTION	UNP P32379
D	123F	GLY	-	INSERTION	UNP P32379
D	123G	GLU	-	INSERTION	UNP P32379
D	125	GLU	ALA	CONFLICT	UNP P32379
D	126	ARG	ALA	CONFLICT	UNP P32379
D	127	LEU	ALA	CONFLICT	UNP P32379
Y	121	LEU	-	INSERTION	UNP P32379
Y	122	ARG	-	INSERTION	UNP P32379
Y	123	PHE	-	INSERTION	UNP P32379
Y	123A	GLY	-	INSERTION	UNP P32379
Y	123B	GLU	-	INSERTION	UNP P32379
Y	123C	GLY	-	INSERTION	UNP P32379
Y	123E	SER	-	INSERTION	UNP P32379
Y	123F	GLY	-	INSERTION	UNP P32379
Y	123G	GLU	-	INSERTION	UNP P32379
Y	125	GLU	ALA	CONFLICT	UNP P32379
Y	126	ARG	ALA	CONFLICT	UNP P32379
Y	127	LEU	ALA	CONFLICT	UNP P32379

- Molecule 5 is a protein called PROTEASOME COMPONENT PRE5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	233	Total	C	N	O	S	0	0	0
			1795	1129	312	350	4			
5	Z	233	Total	C	N	O	S	0	0	0
			1795	1129	312	350	4			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	127	TYR	ALA	CONFLICT	UNP P40302
Z	127	TYR	ALA	CONFLICT	UNP P40302

- Molecule 6 is a protein called PROTEASOME COMPONENT C1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	244	Total	C	N	O	S	0	0	0
			1897	1205	330	358	4			
6	1	244	Total	C	N	O	S	0	0	0
			1897	1205	330	358	4			

- Molecule 7 is a protein called PROTEASOME COMPONENT C7-ALPHA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	243	Total	C	N	O	S	0	0	0
			1921	1221	322	370	8			
7	2	243	Total	C	N	O	S	0	0	0
			1921	1221	322	370	8			

- Molecule 8 is a protein called PROTEASOME COMPONENT PUP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	H	222	Total	C	N	O	S	0	0	0
			1685	1061	293	324	7			
8	O	222	Total	C	N	O	S	0	0	0
			1685	1061	293	324	7			

- Molecule 9 is a protein called PROTEASOME COMPONENT PUP3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	I	204	Total	C	N	O	S	0	0	0
			1581	1010	258	305	8			
9	P	204	Total	C	N	O	S	0	0	0
			1581	1010	258	305	8			

- Molecule 10 is a protein called PROTEASOME COMPONENT C11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	J	198	Total	C	N	O	S	0	0	0
			1585	1005	269	305	6			
10	Q	198	Total	C	N	O	S	0	0	0
			1585	1005	269	305	6			

- Molecule 11 is a protein called PROTEASOME COMPONENT PRE2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	212	Total	C	N	O	S	0	0	0
			1644	1045	280	312	7			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	R	212	Total	C	N	O	S	0	0	0
			1644	1045	280	312	7			

- Molecule 12 is a protein called PROTEASOME COMPONENT C5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	L	222	Total	C	N	O	S	0	0	0
			1757	1115	303	335	4			
12	S	222	Total	C	N	O	S	0	0	0
			1757	1115	303	335	4			

- Molecule 13 is a protein called PROTEASOME COMPONENT PRE4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	M	233	Total	C	N	O	S	0	0	0
			1824	1154	312	351	7			
13	T	233	Total	C	N	O	S	0	0	0
			1824	1154	312	351	7			

- Molecule 14 is a protein called PROTEASOME COMPONENT PRE3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	N	196	Total	C	N	O	S	0	0	0
			1512	955	250	300	7			
14	U	196	Total	C	N	O	S	0	0	0
			1512	955	250	300	7			

- Molecule 15 is a protein called TMC-95A inhibitor.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
15	8	5	Total	C	N	O	0	0	0
			49	33	6	10			
15	9	5	Total	C	N	O	0	0	0
			49	33	6	10			

- Molecule 16 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
16	P	1	Total	Mg	0	0
			1	1		
16	G	1	Total	Mg	0	0
			1	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
16	D	1	Total 1	Mg 1	0	0
16	H	1	Total 1	Mg 1	0	0
16	I	2	Total 2	Mg 2	0	0
16	N	1	Total 1	Mg 1	0	0
16	L	1	Total 1	Mg 1	0	0
16	F	2	Total 2	Mg 2	0	0

- Molecule 17 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
17	A	105	Total 105	O 105	0	0
17	B	80	Total 80	O 80	0	0
17	C	69	Total 69	O 69	0	0
17	D	90	Total 90	O 90	0	0
17	E	55	Total 55	O 55	0	0
17	F	99	Total 99	O 99	0	0
17	G	106	Total 106	O 106	0	0
17	H	114	Total 114	O 114	0	0
17	I	107	Total 107	O 107	0	0
17	J	116	Total 116	O 116	0	0
17	K	103	Total 103	O 103	0	0
17	L	140	Total 140	O 140	0	0
17	M	150	Total 150	O 150	0	0

*Continued on next page...*



*Continued from previous page...*

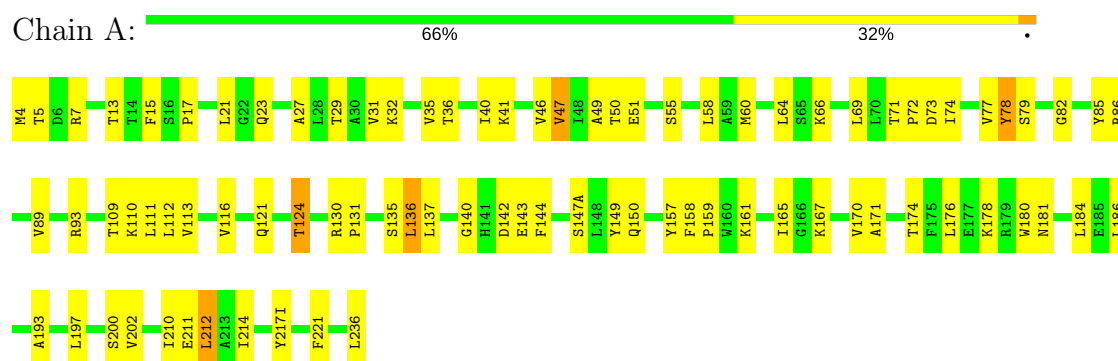
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
17	N	116	Total 116	O 116	0	0
17	O	121	Total 121	O 121	0	0
17	P	101	Total 101	O 101	0	0
17	Q	121	Total 121	O 121	0	0
17	R	107	Total 107	O 107	0	0
17	S	144	Total 144	O 144	0	0
17	T	142	Total 142	O 142	0	0
17	U	118	Total 118	O 118	0	0
17	V	104	Total 104	O 104	0	0
17	W	74	Total 74	O 74	0	0
17	X	70	Total 70	O 70	0	0
17	Y	89	Total 89	O 89	0	0
17	Z	53	Total 53	O 53	0	0
17	1	93	Total 93	O 93	0	0
17	2	103	Total 103	O 103	0	0
17	8	2	Total 2	O 2	0	0
17	9	1	Total 1	O 1	0	0

### 3 Residue-property plots

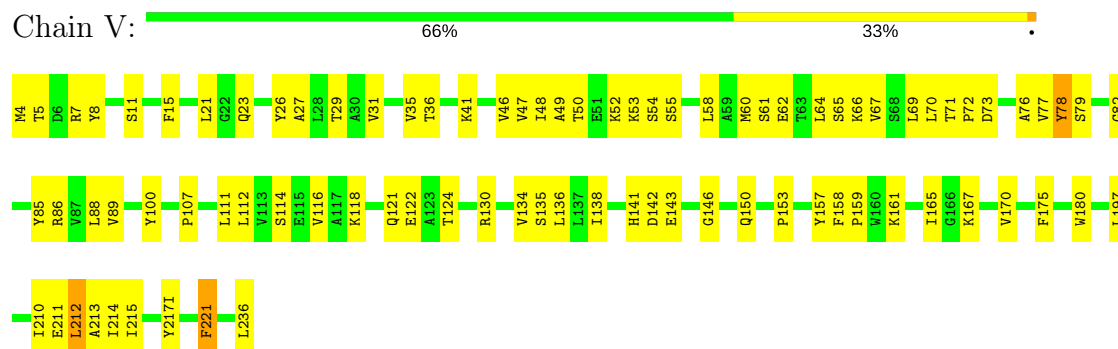
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are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

Note EDS was not executed.

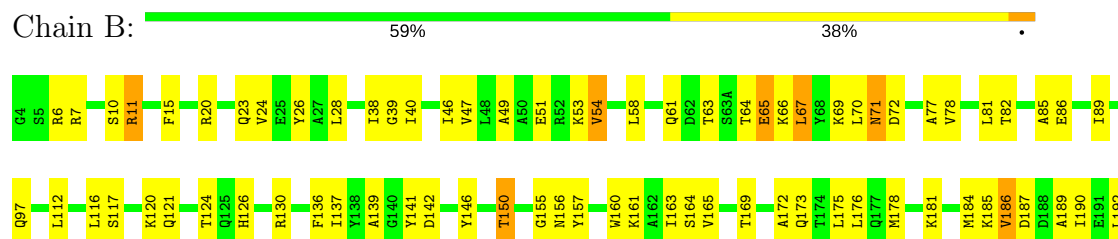
#### • Molecule 1: PROTEASOME COMPONENT Y7

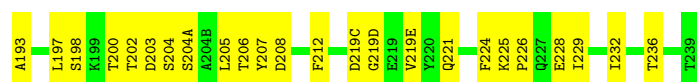


#### • Molecule 1: PROTEASOME COMPONENT Y7



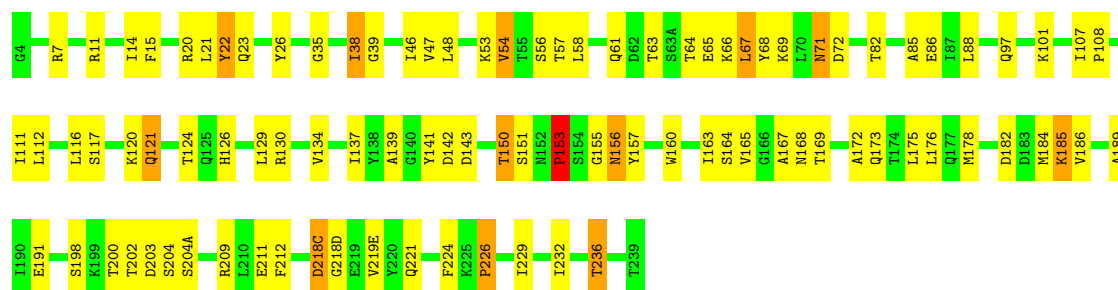
#### • Molecule 2: PROTEASOME COMPONENT Y13





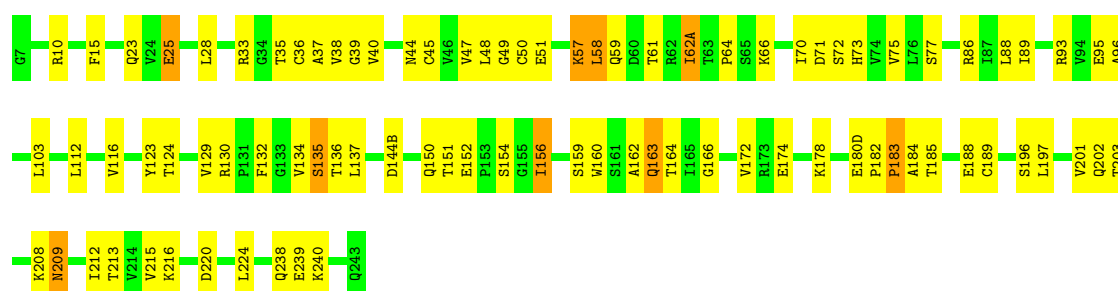
• Molecule 2: PROTEASOME COMPONENT Y13

Chain W: 61% 34% 5%



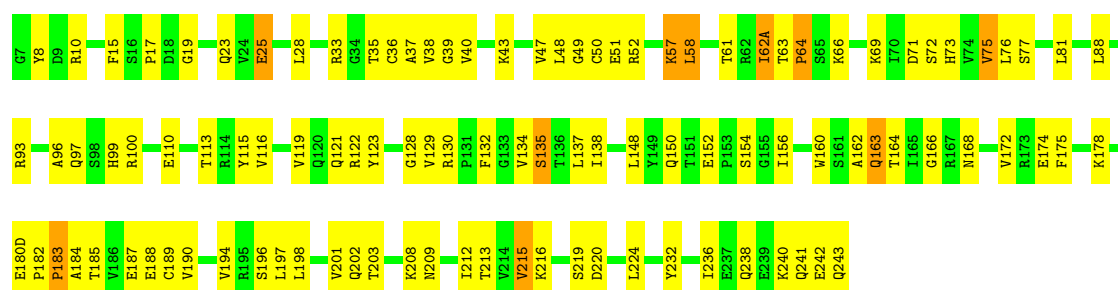
• Molecule 3: PROTEASOME COMPONENT PRE6

Chain C: 63% 33% 4%



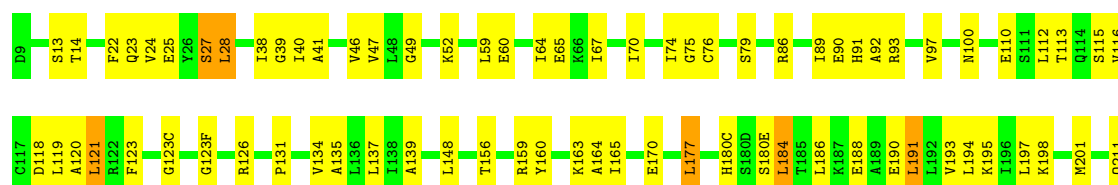
• Molecule 3: PROTEASOME COMPONENT PRE6

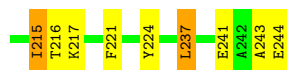
Chain X: 56% 40% 4%



• Molecule 4: PROTEASOME COMPONENT PUP2

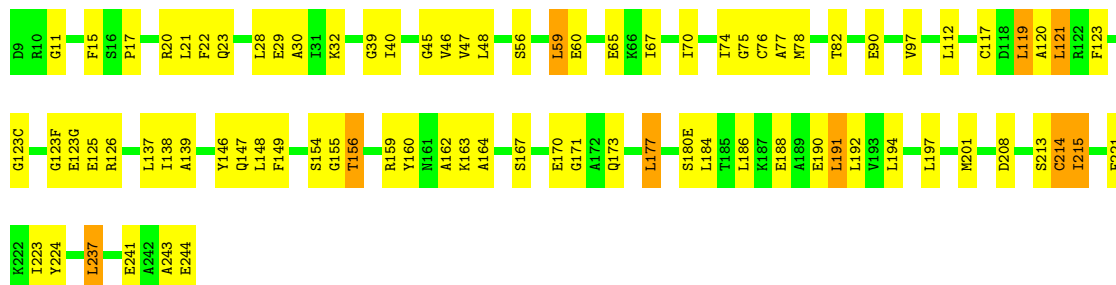
Chain D: 65% 31% 4%





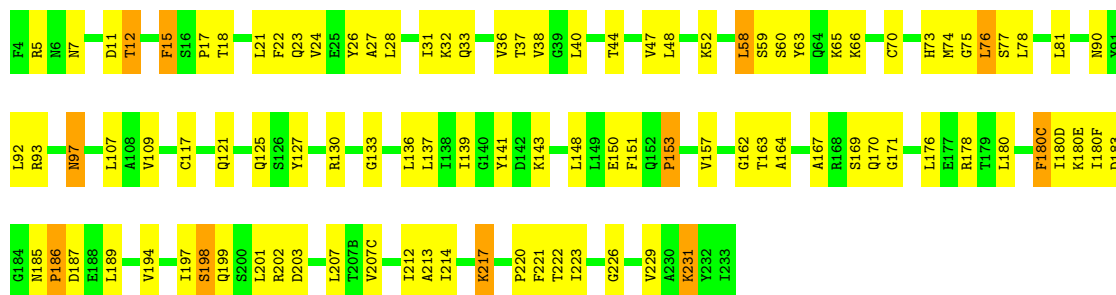
• Molecule 4: PROTEASOME COMPONENT PUP2

Chain Y: 66% 31%



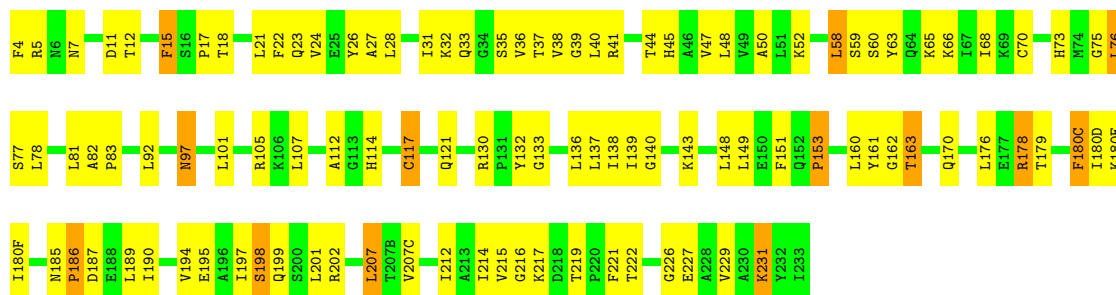
• Molecule 5: PROTEASOME COMPONENT PRE5

Chain E: 57% 38% 5%



• Molecule 5: PROTEASOME COMPONENT PRE5

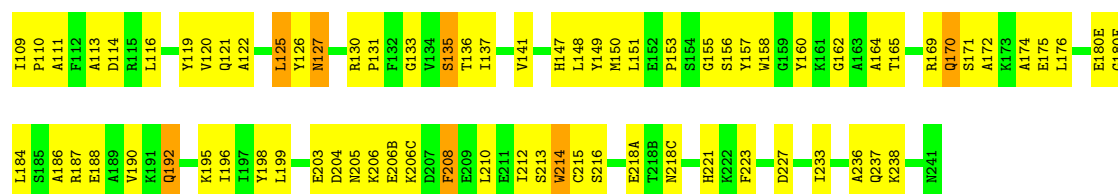
Chain Z: 54% 40% 6%



• Molecule 6: PROTEASOME COMPONENT C1

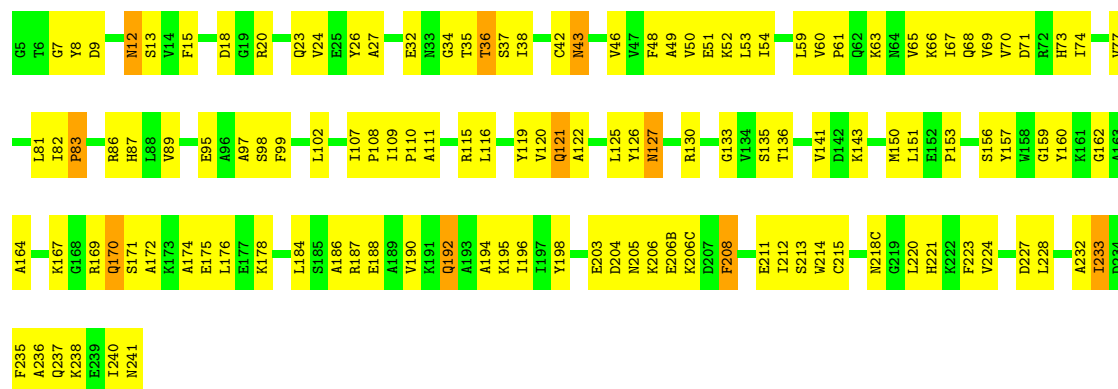
Chain F: 53% 44%





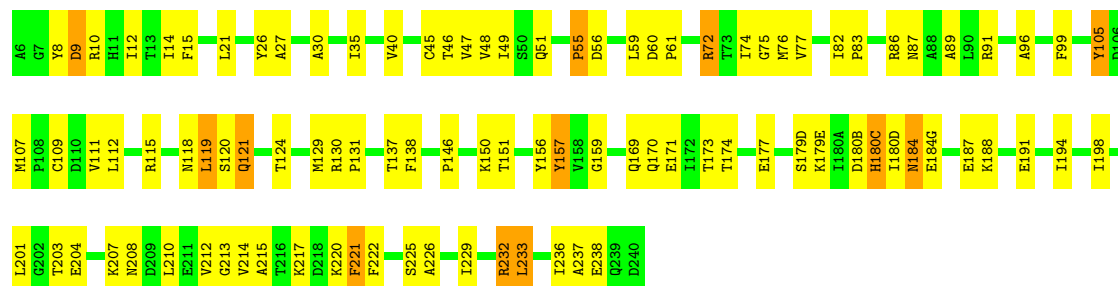
### • Molecule 6: PROTEASOME COMPONENT C1

Chain 1: 48% 48% .



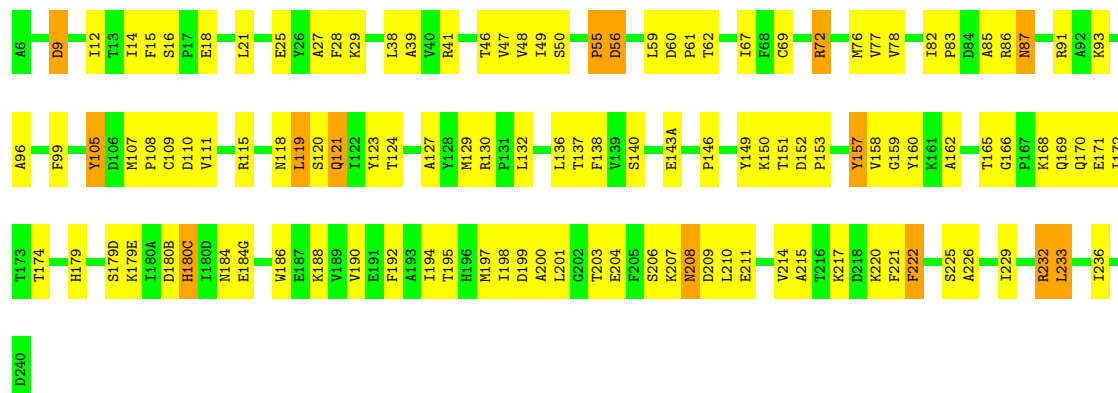
### • Molecule 7: PROTEASOME COMPONENT C7-ALPHA

Chain G: 60% 35% 5%

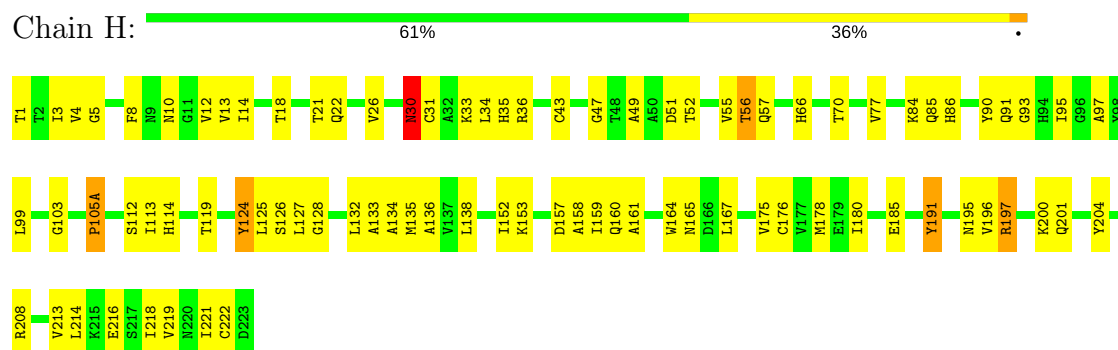


### • Molecule 7: PROTEASOME COMPONENT C7-ALPHA

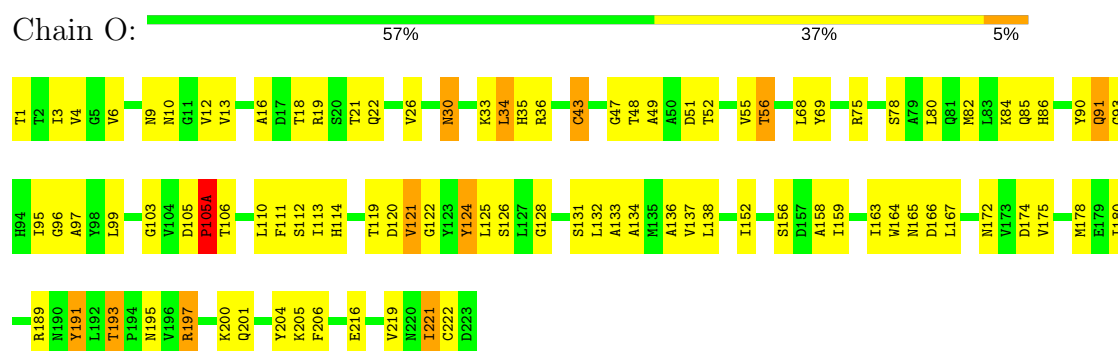
Chain 2: 51% 43% 6%



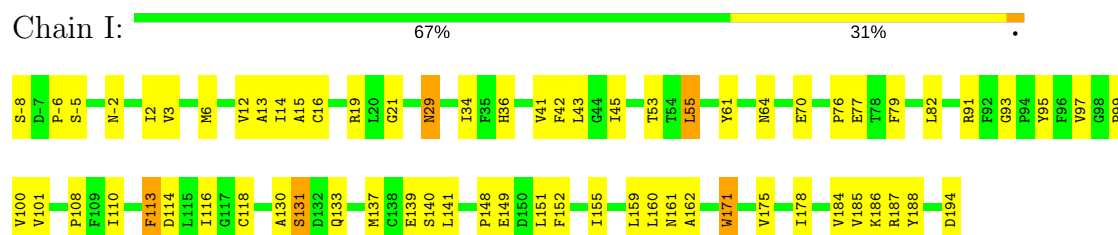
- Molecule 8: PROTEASOME COMPONENT PUP1



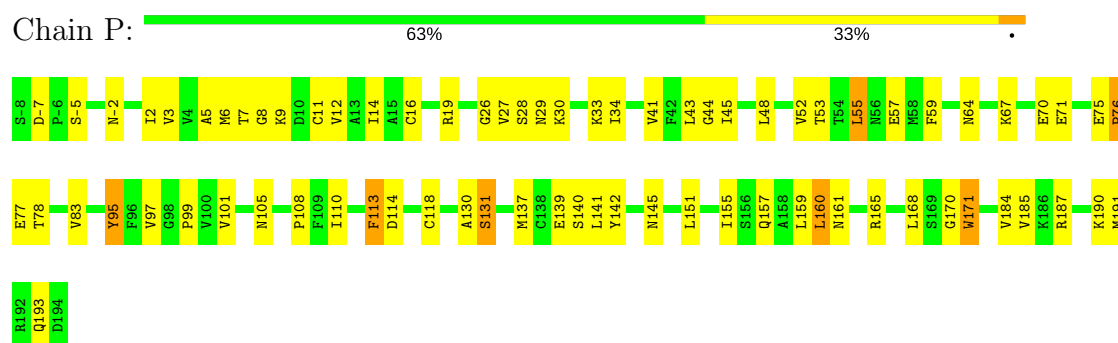
- Molecule 8: PROTEASOME COMPONENT PUP1



- Molecule 9: PROTEASOME COMPONENT PUP3

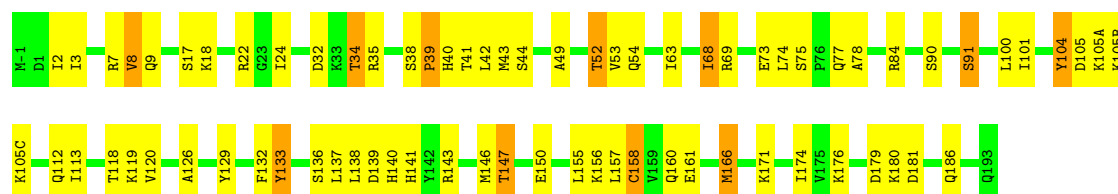


- Molecule 9: PROTEASOME COMPONENT PUP3



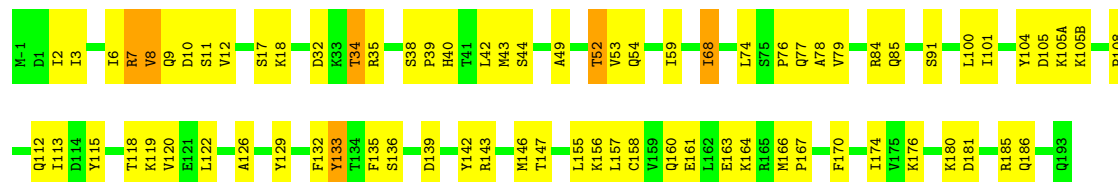
- Molecule 10: PROTEASOME COMPONENT C11





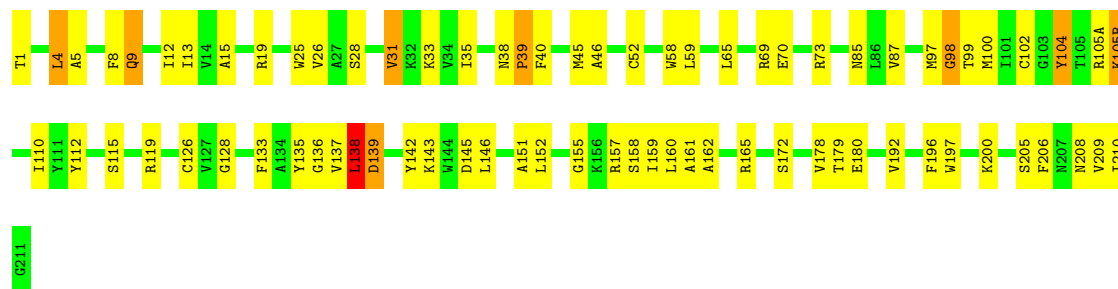
• Molecule 10: PROTEASOME COMPONENT C11

Chain Q: 62% 35%



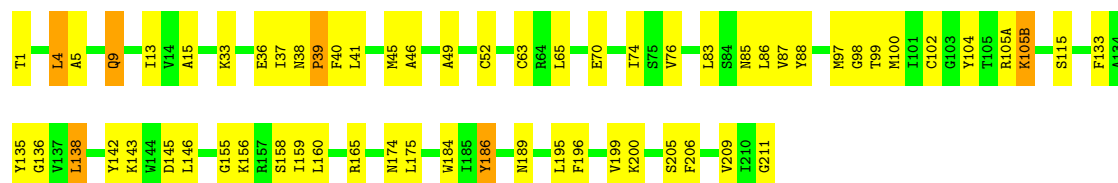
• Molecule 11: PROTEASOME COMPONENT PRE2

Chain K: 64% 32%



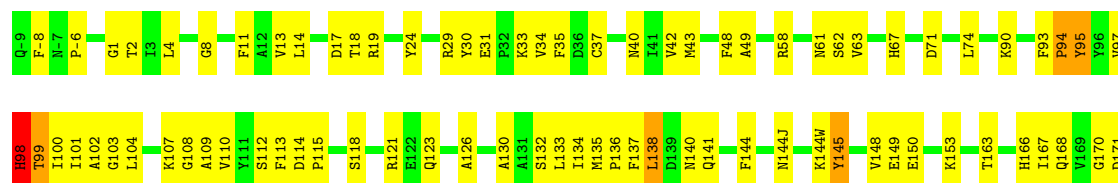
• Molecule 11: PROTEASOME COMPONENT PRE2

Chain R: 70% 27%



• Molecule 12: PROTEASOME COMPONENT C5

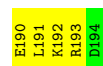
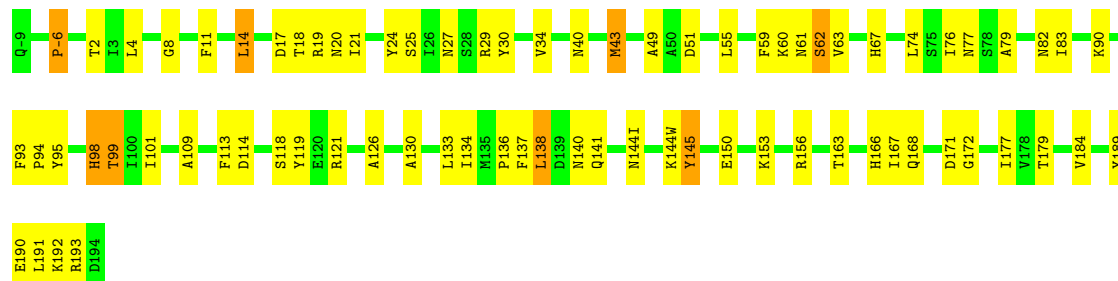
Chain L: 60% 37%





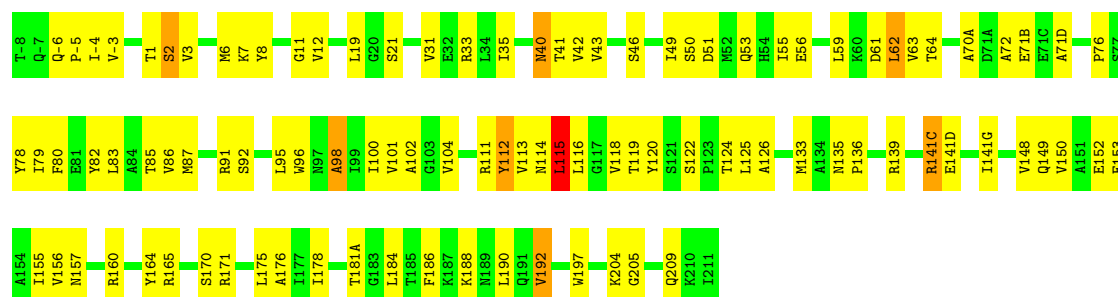
• Molecule 12: PROTEASOME COMPONENT C5

Chain S: 66% 31%



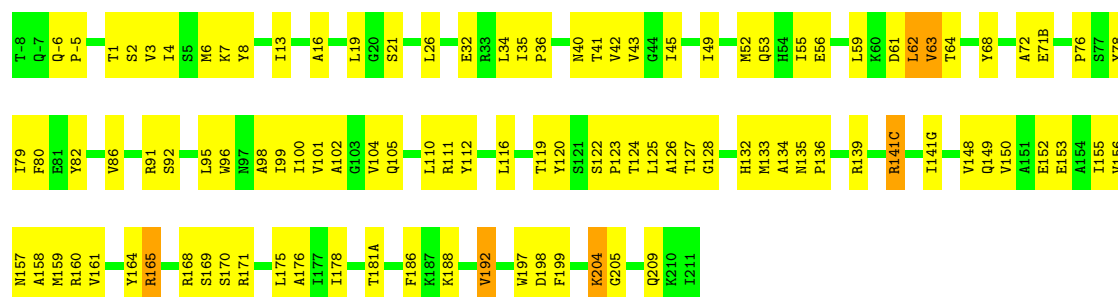
• Molecule 13: PROTEASOME COMPONENT PRE4

Chain M: 57% 40%



• Molecule 13: PROTEASOME COMPONENT PRE4

Chain T: 55% 42%

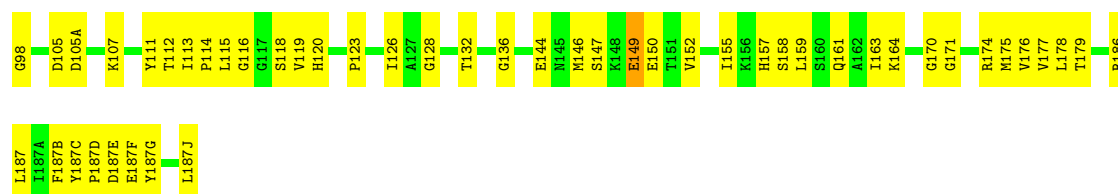


• Molecule 14: PROTEASOME COMPONENT PRE3

Chain N: 56% 42%

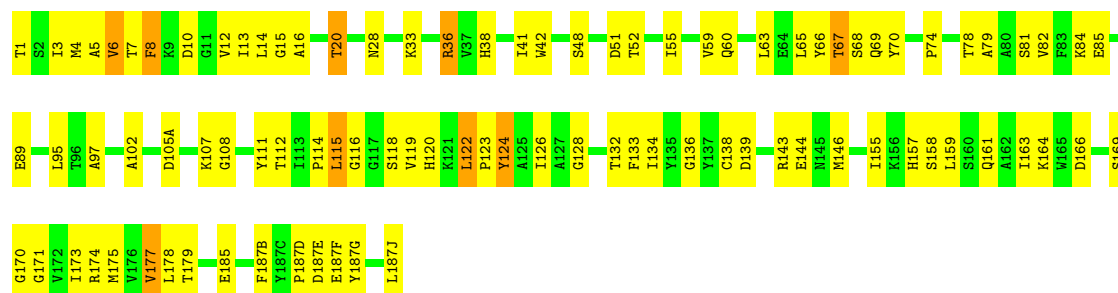






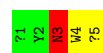
• Molecule 14: PROTEASOME COMPONENT PRE3

Chain U: 53% 43% 5%



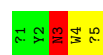
• Molecule 15: TMC-95A inhibitor

Chain 8: 40% 40% 20%



• Molecule 15: TMC-95A inhibitor

Chain 9: 40% 40% 20%



## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	135.20Å 301.30Å 144.80Å 90.00° 112.60° 90.00°	Depositor
Resolution (Å)	50.00 – 3.00	Depositor
% Data completeness (in resolution range)	(Not available) (50.00-3.00)	Depositor
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	X-PLOR	Depositor
R, $R_{free}$	0.251 , 0.336	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	52549	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	36.0	wwPDB-VP

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: AKK, MG, 1QQ, R4K

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	0.65	0/1952	0.81	0/2642
1	V	0.62	0/1952	0.79	0/2642
2	B	0.65	0/1935	0.79	0/2618
2	W	0.62	0/1935	0.79	0/2618
3	C	0.65	0/1920	0.80	1/2598 (0.0%)
3	X	0.59	0/1920	0.78	0/2598
4	D	0.63	0/1887	0.78	1/2541 (0.0%)
4	Y	0.63	0/1887	0.78	2/2541 (0.1%)
5	E	0.57	0/1823	0.77	0/2463
5	Z	0.57	0/1823	0.78	0/2463
6	1	0.64	0/1937	0.79	0/2614
6	F	0.61	0/1937	0.78	1/2614 (0.0%)
7	2	0.71	0/1959	0.80	0/2652
7	G	0.65	0/1959	0.80	0/2652
8	H	0.72	0/1716	0.85	0/2326
8	O	0.69	1/1716 (0.1%)	0.81	0/2326
9	I	0.72	0/1611	0.84	0/2174
9	P	0.76	0/1611	0.85	2/2174 (0.1%)
10	J	0.70	1/1613 (0.1%)	0.82	0/2173
10	Q	0.69	0/1613	0.80	0/2173
11	K	0.71	0/1681	0.81	1/2274 (0.0%)
11	R	0.69	0/1681	0.80	0/2274
12	L	0.75	1/1795 (0.1%)	0.82	2/2420 (0.1%)
12	S	0.68	0/1795	0.83	1/2420 (0.0%)
13	M	0.68	0/1855	0.80	1/2514 (0.0%)
13	T	0.70	0/1855	0.84	3/2514 (0.1%)
14	N	0.75	0/1541	0.81	1/2087 (0.0%)
14	U	0.73	0/1541	0.78	1/2087 (0.0%)
15	8	1.96	0/20	2.20	2/26 (7.7%)
15	9	2.53	0/20	2.48	2/26 (7.7%)
All	All	0.67	3/50490 (0.0%)	0.81	21/68244 (0.0%)

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
11	R	0	1
15	8	0	1
15	9	0	1
All	All	0	3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
12	L	37	CYS	CB-SG	-5.27	1.73	1.81
8	O	43	CYS	CB-SG	-5.14	1.73	1.81
10	J	158	CYS	CB-SG	-5.06	1.73	1.81

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
15	9	3	ASN	CB-CG-OD1	-7.24	107.12	121.60
15	8	3	ASN	CB-CG-OD1	-7.21	107.18	121.60
13	T	98	ALA	N-CA-C	-6.22	94.20	111.00
15	9	3	ASN	OD1-CG-ND2	-6.11	107.85	121.90
9	P	95	TYR	N-CA-C	-6.02	94.74	111.00

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
15	8	3	ASN	Sidechain
15	9	3	ASN	Sidechain
11	R	186	TYR	Sidechain

## 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	1915	0	1926	59	0
1	V	1915	0	1926	71	0
2	B	1905	0	1901	75	0
2	W	1905	0	1901	75	0
3	C	1891	0	1900	68	0
3	X	1891	0	1900	92	0
4	D	1862	0	1836	75	0
4	Y	1862	0	1836	67	0
5	E	1795	0	1797	78	0
5	Z	1795	0	1797	111	0
6	1	1897	0	1886	111	0
6	F	1897	0	1886	101	0
7	2	1921	0	1910	95	0
7	G	1921	0	1910	83	0
8	H	1685	0	1688	75	0
8	O	1685	0	1688	83	0
9	I	1581	0	1574	56	0
9	P	1581	0	1574	67	0
10	J	1585	0	1590	55	0
10	Q	1585	0	1590	62	0
11	K	1644	0	1595	72	0
11	R	1644	0	1595	54	0
12	L	1757	0	1711	79	0
12	S	1757	0	1711	80	0
13	M	1824	0	1832	74	0
13	T	1824	0	1832	90	0
14	N	1512	0	1481	75	0
14	U	1512	0	1481	71	0
15	8	49	0	36	8	0
15	9	49	0	36	7	0
16	D	1	0	0	0	0
16	F	2	0	0	0	0
16	G	1	0	0	0	0
16	H	1	0	0	0	0
16	I	2	0	0	0	0
16	L	1	0	0	0	0
16	N	1	0	0	0	0
16	P	1	0	0	0	0
17	1	93	0	0	3	0
17	2	103	0	0	6	0
17	8	2	0	0	0	0
17	9	1	0	0	0	0
17	A	105	0	0	1	0

Continued on next page...

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
17	B	80	0	0	1	0
17	C	69	0	0	3	0
17	D	90	0	0	6	0
17	E	55	0	0	3	0
17	F	99	0	0	6	0
17	G	106	0	0	3	0
17	H	114	0	0	2	0
17	I	107	0	0	3	0
17	J	116	0	0	6	0
17	K	103	0	0	4	0
17	L	140	0	0	8	0
17	M	150	0	0	3	0
17	N	116	0	0	5	0
17	O	121	0	0	3	0
17	P	101	0	0	4	0
17	Q	121	0	0	5	0
17	R	107	0	0	4	0
17	S	144	0	0	6	0
17	T	142	0	0	8	0
17	U	118	0	0	1	0
17	V	104	0	0	2	0
17	W	74	0	0	2	0
17	X	70	0	0	4	0
17	Y	89	0	0	5	0
17	Z	53	0	0	2	0
All	All	52549	0	49326	1936	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 20.

The worst 5 of 1936 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)
7:2:123:TYR:O	17:2:2009:HOH:O	1.54	1.25
14:N:136:GLY:HA2	14:U:161:GLN:HE21	1.13	1.10
8:H:21:THR:O	15:8:4:R4K:O33	1.75	1.05
14:N:161:GLN:HE21	14:U:136:GLY:HA2	1.21	1.05
8:O:21:THR:O	15:9:4:R4K:O33	1.75	1.04

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	248/250 (99%)	207 (84%)	38 (15%)	3 (1%)	14	51
1	V	248/250 (99%)	212 (86%)	32 (13%)	4 (2%)	11	43
2	B	242/244 (99%)	210 (87%)	26 (11%)	6 (2%)	6	31
2	W	242/244 (99%)	209 (86%)	24 (10%)	9 (4%)	4	21
3	C	239/241 (99%)	207 (87%)	25 (10%)	7 (3%)	5	26
3	X	239/241 (99%)	203 (85%)	29 (12%)	7 (3%)	5	26
4	D	240/242 (99%)	205 (85%)	30 (12%)	5 (2%)	8	36
4	Y	240/242 (99%)	210 (88%)	24 (10%)	6 (2%)	6	31
5	E	231/233 (99%)	196 (85%)	30 (13%)	5 (2%)	7	34
5	Z	231/233 (99%)	196 (85%)	28 (12%)	7 (3%)	5	26
6	1	242/244 (99%)	203 (84%)	33 (14%)	6 (2%)	6	31
6	F	242/244 (99%)	206 (85%)	34 (14%)	2 (1%)	21	61
7	2	241/243 (99%)	211 (88%)	24 (10%)	6 (2%)	6	31
7	G	241/243 (99%)	209 (87%)	27 (11%)	5 (2%)	8	36
8	H	220/222 (99%)	192 (87%)	25 (11%)	3 (1%)	12	47
8	O	220/222 (99%)	198 (90%)	17 (8%)	5 (2%)	7	33
9	I	202/204 (99%)	181 (90%)	20 (10%)	1 (0%)	31	71
9	P	202/204 (99%)	185 (92%)	16 (8%)	1 (0%)	31	71
10	J	196/198 (99%)	175 (89%)	17 (9%)	4 (2%)	8	37
10	Q	196/198 (99%)	175 (89%)	16 (8%)	5 (3%)	6	30
11	K	210/212 (99%)	190 (90%)	17 (8%)	3 (1%)	12	47
11	R	210/212 (99%)	187 (89%)	21 (10%)	2 (1%)	17	56
12	L	220/222 (99%)	191 (87%)	27 (12%)	2 (1%)	19	59
12	S	220/222 (99%)	196 (89%)	23 (10%)	1 (0%)	31	71
13	M	231/233 (99%)	202 (87%)	25 (11%)	4 (2%)	10	42

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
13	T	231/233 (99%)	200 (87%)	29 (13%)	2 (1%)	19	59
14	N	194/196 (99%)	175 (90%)	16 (8%)	3 (2%)	11	45
14	U	194/196 (99%)	171 (88%)	20 (10%)	3 (2%)	11	45
15	8	1/5 (20%)	1 (100%)	0	0	100	100
15	9	1/5 (20%)	1 (100%)	0	0	100	100
All	All	6314/6378 (99%)	5504 (87%)	693 (11%)	117 (2%)	9	39

5 of 117 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	143	GLU
1	A	167	LYS
2	B	54	VAL
2	B	204(A)	SER
3	C	58	LEU

### 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	209/209 (100%)	198 (95%)	11 (5%)	25	62
1	V	209/209 (100%)	200 (96%)	9 (4%)	32	70
2	B	203/203 (100%)	188 (93%)	15 (7%)	15	47
2	W	203/203 (100%)	186 (92%)	17 (8%)	12	41
3	C	213/213 (100%)	194 (91%)	19 (9%)	11	38
3	X	213/213 (100%)	196 (92%)	17 (8%)	13	43
4	D	198/198 (100%)	185 (93%)	13 (7%)	18	53
4	Y	198/198 (100%)	185 (93%)	13 (7%)	18	53
5	E	192/192 (100%)	171 (89%)	21 (11%)	7	27
5	Z	192/192 (100%)	173 (90%)	19 (10%)	8	32
6	1	201/201 (100%)	181 (90%)	20 (10%)	8	31

Continued on next page...



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
6	F	201/201 (100%)	183 (91%)	18 (9%)	10	38
7	2	207/207 (100%)	188 (91%)	19 (9%)	10	37
7	G	207/207 (100%)	191 (92%)	16 (8%)	14	46
8	H	181/181 (100%)	169 (93%)	12 (7%)	18	53
8	O	181/181 (100%)	167 (92%)	14 (8%)	14	46
9	I	172/172 (100%)	160 (93%)	12 (7%)	16	50
9	P	172/172 (100%)	163 (95%)	9 (5%)	25	63
10	J	175/175 (100%)	164 (94%)	11 (6%)	20	55
10	Q	175/175 (100%)	165 (94%)	10 (6%)	23	60
11	K	169/169 (100%)	156 (92%)	13 (8%)	14	46
11	R	169/169 (100%)	158 (94%)	11 (6%)	19	54
12	L	185/185 (100%)	173 (94%)	12 (6%)	19	54
12	S	185/185 (100%)	173 (94%)	12 (6%)	19	54
13	M	199/199 (100%)	186 (94%)	13 (6%)	19	54
13	T	199/199 (100%)	188 (94%)	11 (6%)	24	61
14	N	162/162 (100%)	152 (94%)	10 (6%)	20	56
14	U	162/162 (100%)	147 (91%)	15 (9%)	10	36
15	8	2/2 (100%)	2 (100%)	0	100	100
15	9	2/2 (100%)	2 (100%)	0	100	100
All	All	5336/5336 (100%)	4944 (93%)	392 (7%)	15	48

5 of 392 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
13	M	112	TYR
10	Q	7	ARG
6	1	170	GLN
13	M	184	LEU
8	O	56	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 187 such sidechains are listed below:

Mol	Chain	Res	Type
13	M	135	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
10	Q	54	GLN
6	1	123	HIS
13	M	172	ASN
8	O	30	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# $ Z  > 2$	Counts	RMSZ	# $ Z  > 2$
15	R4K	8	4	15	17,18,19	2.98	6 (35%)	19,27,29	2.52	8 (42%)
15	R4K	9	4	15	17,18,19	3.75	9 (52%)	19,27,29	3.61	10 (52%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
15	R4K	8	4	15	-	0/9/28/30	0/2/2/2
15	R4K	9	4	15	-	0/9/28/30	0/2/2/2

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
15	9	4	R4K	C32-N31	-8.82	1.24	1.35
15	9	4	R4K	C55-C50	-8.14	1.30	1.39
15	8	4	R4K	C29-C50	-6.97	1.41	1.51

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
15	8	4	R4K	C55-C50	-6.22	1.32	1.39
15	8	4	R4K	C32-N31	-5.28	1.28	1.35

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
15	9	4	R4K	C54-C55-C50	-6.76	115.56	121.93
15	8	4	R4K	O-C-CA	-4.51	114.58	125.09
15	8	4	R4K	C54-C55-C50	-4.10	118.06	121.93
15	9	4	R4K	C51-C50-C29	-3.69	124.91	130.68
15	8	4	R4K	CB-CA-C	-3.48	107.29	112.15

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
15	8	4	R4K	2	0
15	9	4	R4K	2	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 10 ligands modelled in this entry, 10 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 [i](#)

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

EDS was not executed - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates [i](#)

EDS was not executed - this section is therefore empty.

### 6.4 Ligands [i](#)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.