



# wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 06:13 AM GMT

PDB ID : 1Q5R  
Title : The Rhodococcus 20S proteasome with unprocessed pro-peptides  
Authors : Kwon, Y.D.; Nagy, I.; Adams, P.D.; Baumeister, W.; Jap, B.K.  
Deposited on : 2003-08-08  
Resolution : 3.10 Å(reported)

This is a wwPDB 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 <http://wwpdb.org/ValidationPDFNotes.html>

---

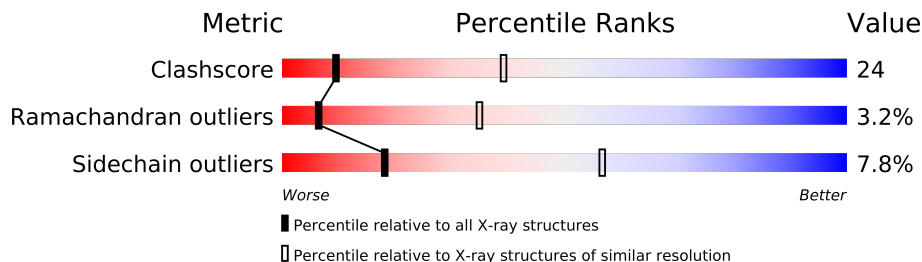
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	21963
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 3.10 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	79885	1078 (3.16-3.04)
Ramachandran outliers	78287	1044 (3.16-3.04)
Sidechain outliers	78261	1044 (3.16-3.04)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	259	
1	B	259	
1	C	259	
1	D	259	
1	E	259	
1	F	259	
1	G	259	
2	H	294	
2	I	294	
2	J	294	
2	K	294	
2	L	294	
2	M	294	
2	N	294	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 25013 atoms, of which 0 are hydrogen and 0 are deuterium.

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 alpha-type subunit 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	219	Total	C	N	O	S	0	0	0
			1695	1057	312	321	5			
1	B	219	Total	C	N	O	S	0	0	0
			1695	1057	312	321	5			
1	C	219	Total	C	N	O	S	0	0	0
			1695	1057	312	321	5			
1	D	219	Total	C	N	O	S	0	0	0
			1695	1057	312	321	5			
1	E	219	Total	C	N	O	S	0	0	0
			1695	1057	312	321	5			
1	F	219	Total	C	N	O	S	0	0	0
			1695	1057	312	321	5			
1	G	219	Total	C	N	O	S	0	0	0
			1695	1057	312	321	5			

There are 49 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	INITIATING METHIONINE	UNP Q53080
A	2	HIS	-	EXPRESSION TAG	UNP Q53080
A	3	HIS	-	EXPRESSION TAG	UNP Q53080
A	4	HIS	-	EXPRESSION TAG	UNP Q53080
A	5	HIS	-	EXPRESSION TAG	UNP Q53080
A	6	HIS	-	EXPRESSION TAG	UNP Q53080
A	7	HIS	-	EXPRESSION TAG	UNP Q53080
B	1	MET	-	INITIATING METHIONINE	UNP Q53080
B	2	HIS	-	EXPRESSION TAG	UNP Q53080
B	3	HIS	-	EXPRESSION TAG	UNP Q53080
B	4	HIS	-	EXPRESSION TAG	UNP Q53080
B	5	HIS	-	EXPRESSION TAG	UNP Q53080
B	6	HIS	-	EXPRESSION TAG	UNP Q53080
B	7	HIS	-	EXPRESSION TAG	UNP Q53080
C	1	MET	-	INITIATING METHIONINE	UNP Q53080

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	2	HIS	-	EXPRESSION TAG	UNP Q53080
C	3	HIS	-	EXPRESSION TAG	UNP Q53080
C	4	HIS	-	EXPRESSION TAG	UNP Q53080
C	5	HIS	-	EXPRESSION TAG	UNP Q53080
C	6	HIS	-	EXPRESSION TAG	UNP Q53080
C	7	HIS	-	EXPRESSION TAG	UNP Q53080
D	1	MET	-	INITIATING METHIONINE	UNP Q53080
D	2	HIS	-	EXPRESSION TAG	UNP Q53080
D	3	HIS	-	EXPRESSION TAG	UNP Q53080
D	4	HIS	-	EXPRESSION TAG	UNP Q53080
D	5	HIS	-	EXPRESSION TAG	UNP Q53080
D	6	HIS	-	EXPRESSION TAG	UNP Q53080
D	7	HIS	-	EXPRESSION TAG	UNP Q53080
E	1	MET	-	INITIATING METHIONINE	UNP Q53080
E	2	HIS	-	EXPRESSION TAG	UNP Q53080
E	3	HIS	-	EXPRESSION TAG	UNP Q53080
E	4	HIS	-	EXPRESSION TAG	UNP Q53080
E	5	HIS	-	EXPRESSION TAG	UNP Q53080
E	6	HIS	-	EXPRESSION TAG	UNP Q53080
E	7	HIS	-	EXPRESSION TAG	UNP Q53080
F	1	MET	-	INITIATING METHIONINE	UNP Q53080
F	2	HIS	-	EXPRESSION TAG	UNP Q53080
F	3	HIS	-	EXPRESSION TAG	UNP Q53080
F	4	HIS	-	EXPRESSION TAG	UNP Q53080
F	5	HIS	-	EXPRESSION TAG	UNP Q53080
F	6	HIS	-	EXPRESSION TAG	UNP Q53080
F	7	HIS	-	EXPRESSION TAG	UNP Q53080
G	1	MET	-	INITIATING METHIONINE	UNP Q53080
G	2	HIS	-	EXPRESSION TAG	UNP Q53080
G	3	HIS	-	EXPRESSION TAG	UNP Q53080
G	4	HIS	-	EXPRESSION TAG	UNP Q53080
G	5	HIS	-	EXPRESSION TAG	UNP Q53080
G	6	HIS	-	EXPRESSION TAG	UNP Q53080
G	7	HIS	-	EXPRESSION TAG	UNP Q53080

- Molecule 2 is a protein called proteasome beta-type subunit 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	251	Total	C	N	O	S	0	0	0
			1864	1164	323	375	2			
2	I	251	Total	C	N	O	S	0	0	0
			1864	1164	323	375	2			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	J	251	Total	C	N	O	S	0	0	0
			1864	1164	323	375	2			
2	K	251	Total	C	N	O	S	0	0	0
			1864	1164	323	375	2			
2	L	251	Total	C	N	O	S	0	0	0
			1864	1164	323	375	2			
2	M	251	Total	C	N	O	S	0	0	0
			1864	1164	323	375	2			
2	N	251	Total	C	N	O	S	0	0	0
			1864	1164	323	375	2			

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
H	-63	ALA	LYS	ENGINEERED	UNP Q53079
I	-63	ALA	LYS	ENGINEERED	UNP Q53079
J	-63	ALA	LYS	ENGINEERED	UNP Q53079
K	-63	ALA	LYS	ENGINEERED	UNP Q53079
L	-63	ALA	LYS	ENGINEERED	UNP Q53079
M	-63	ALA	LYS	ENGINEERED	UNP Q53079
N	-63	ALA	LYS	ENGINEERED	UNP Q53079

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	4	Total	O	0	0
			4	4		
3	B	5	Total	O	0	0
			5	5		
3	C	3	Total	O	0	0
			3	3		
3	D	6	Total	O	0	0
			6	6		
3	E	5	Total	O	0	0
			5	5		
3	F	2	Total	O	0	0
			2	2		
3	G	1	Total	O	0	0
			1	1		
3	H	13	Total	O	0	0
			13	13		
3	I	14	Total	O	0	0
			14	14		

*Continued on next page...*

*Continued from previous page...*

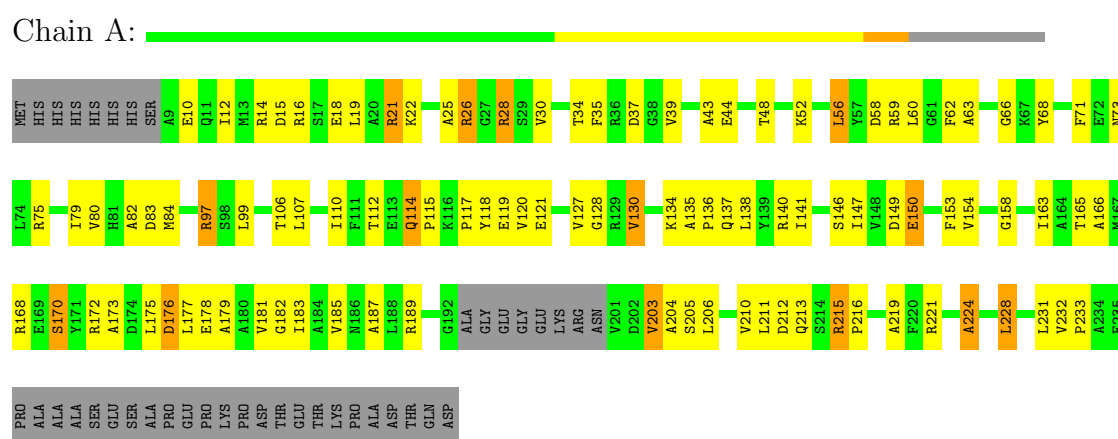
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	J	13	Total 13	O 13	0	0
3	K	8	Total 8	O 8	0	0
3	L	12	Total 12	O 12	0	0
3	M	6	Total 6	O 6	0	0
3	N	8	Total 8	O 8	0	0

### 3 Residue-property plots

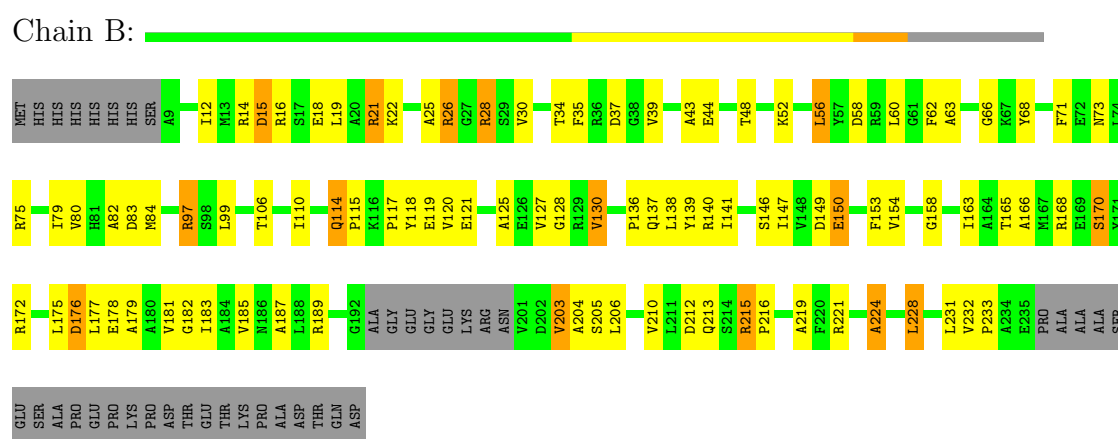
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according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

Note EDS was not executed.

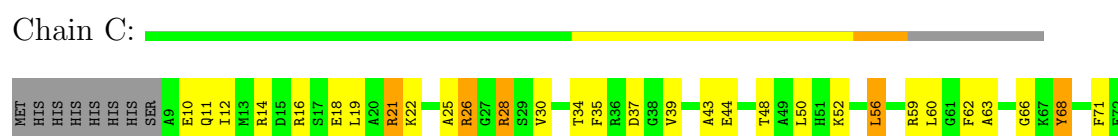
- Molecule 1: proteasome alpha-type subunit 1

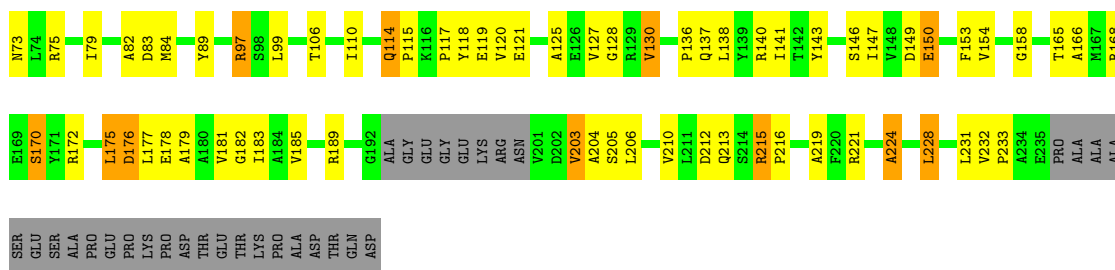


- Molecule 1: proteasome alpha-type subunit 1



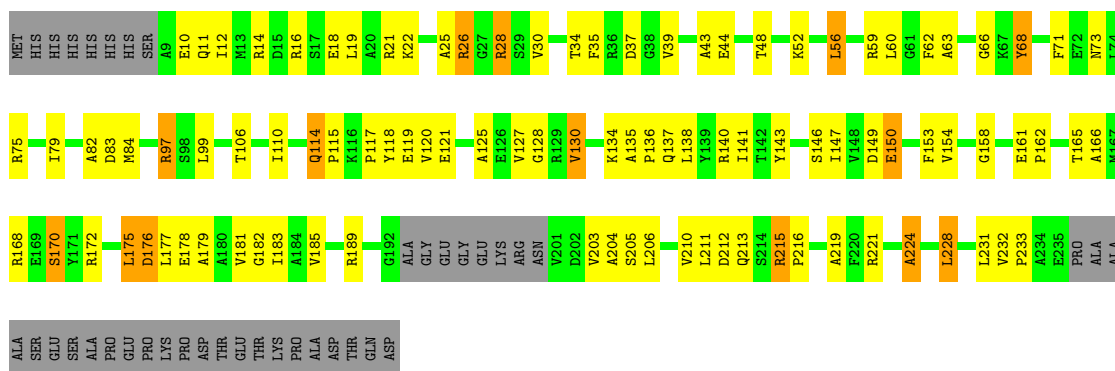
- Molecule 1: proteasome alpha-type subunit 1





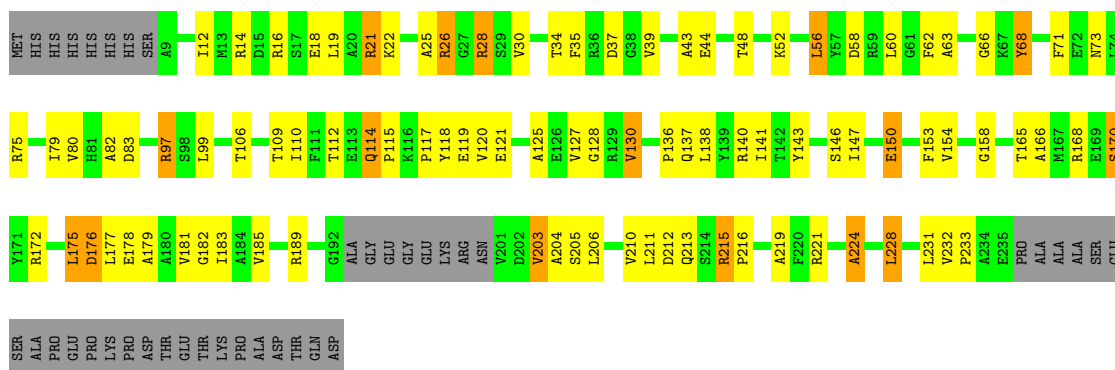
- Molecule 1: proteasome alpha-type subunit 1

Chain D:



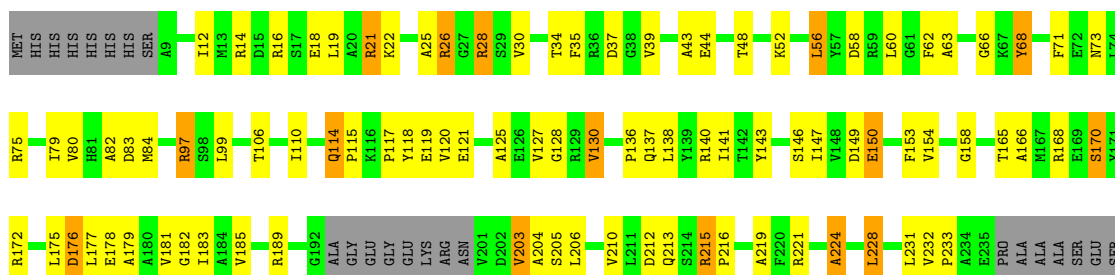
- Molecule 1: proteasome alpha-type subunit 1

Chain E:



- Molecule 1: proteasome alpha-type subunit 1

Chain F:

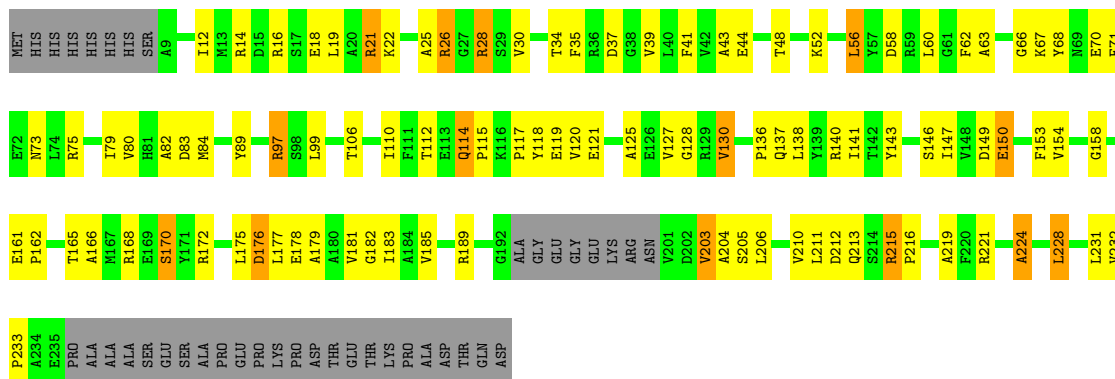




ALA  
PRO  
GLU  
PRO  
LYS  
PRO  
ASP  
THR  
GLU  
THR  
LYS  
PRO  
ALA  
ASP  
THR  
GLN  
ASP

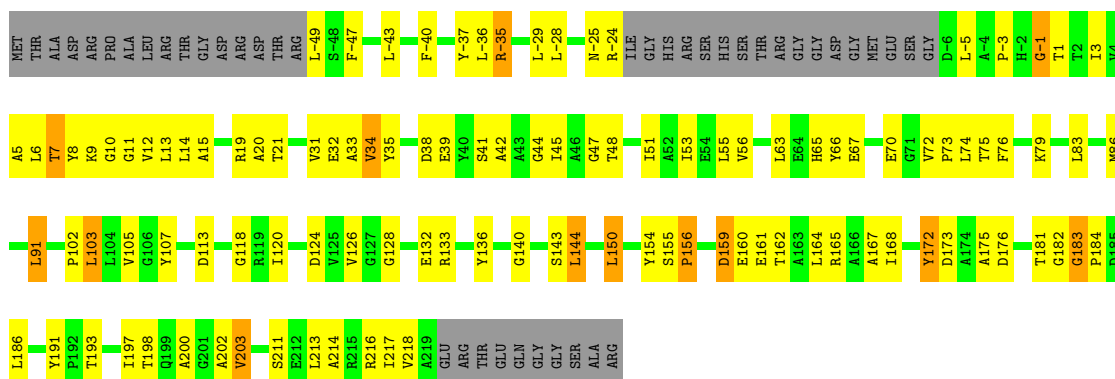
• Molecule 1: proteasome alpha-type subunit 1

Chain G:



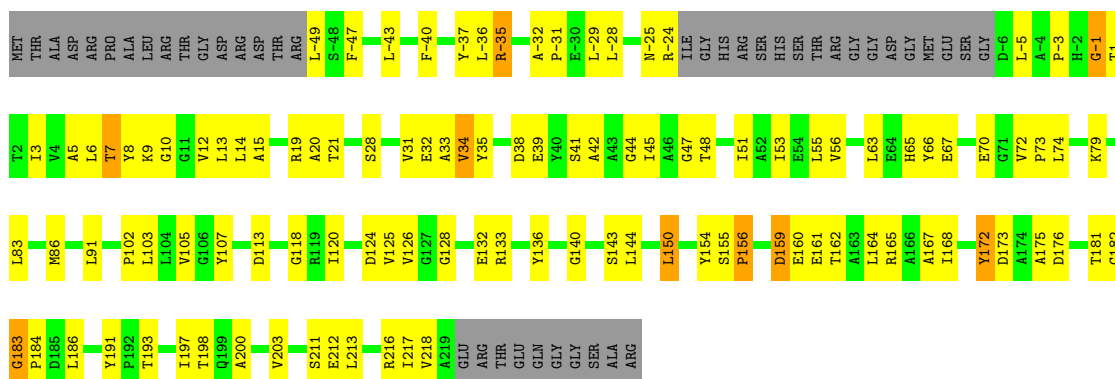
• Molecule 2: proteasome beta-type subunit 1

Chain H:



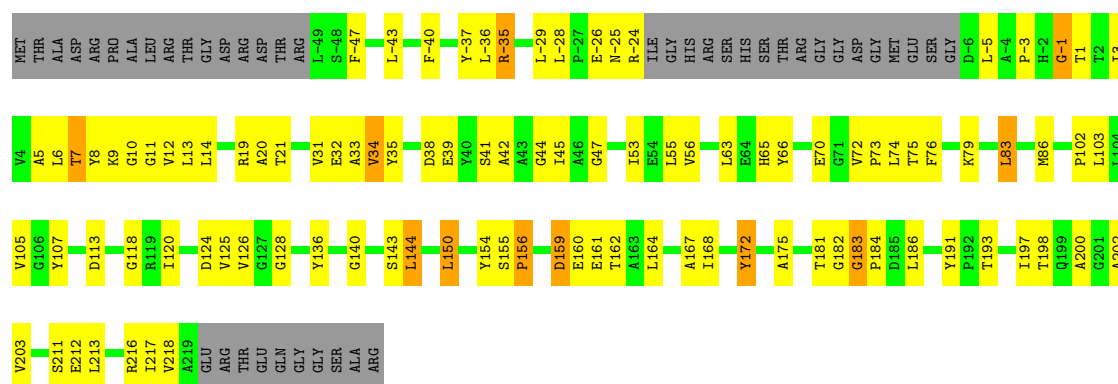
• Molecule 2: proteasome beta-type subunit 1

Chain I:



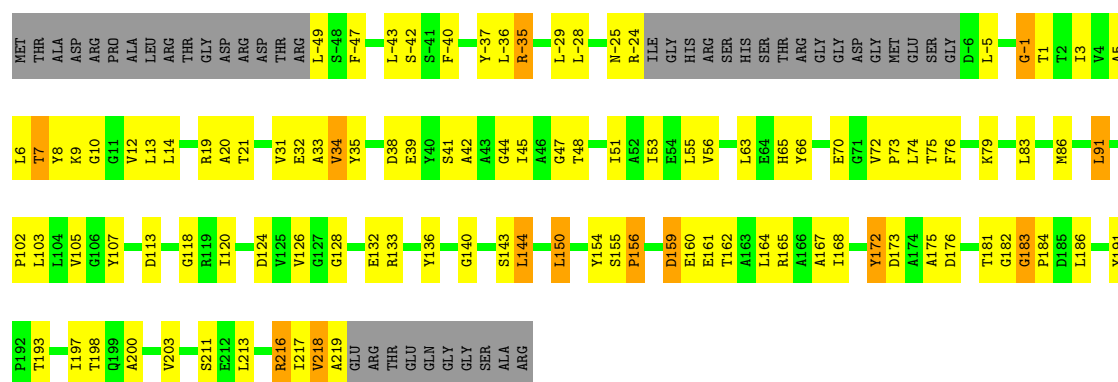
• Molecule 2: proteasome beta-type subunit 1

Chain J:



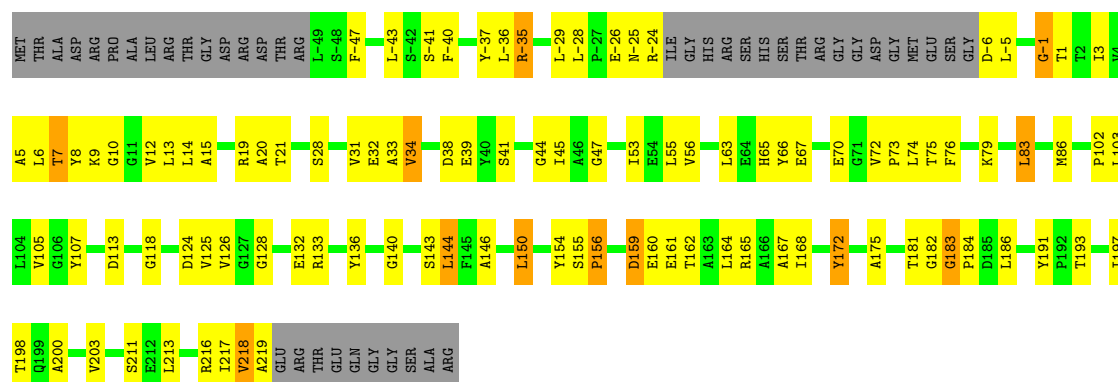
• Molecule 2: proteasome beta-type subunit 1

Chain K:



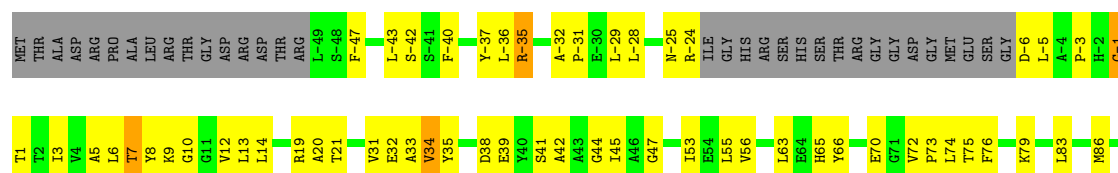
• Molecule 2: proteasome beta-type subunit 1

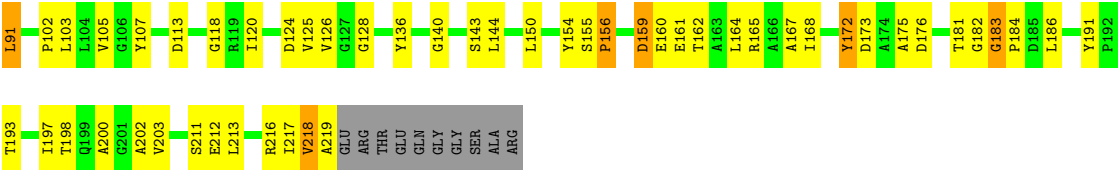
Chain L:



• Molecule 2: proteasome beta-type subunit 1

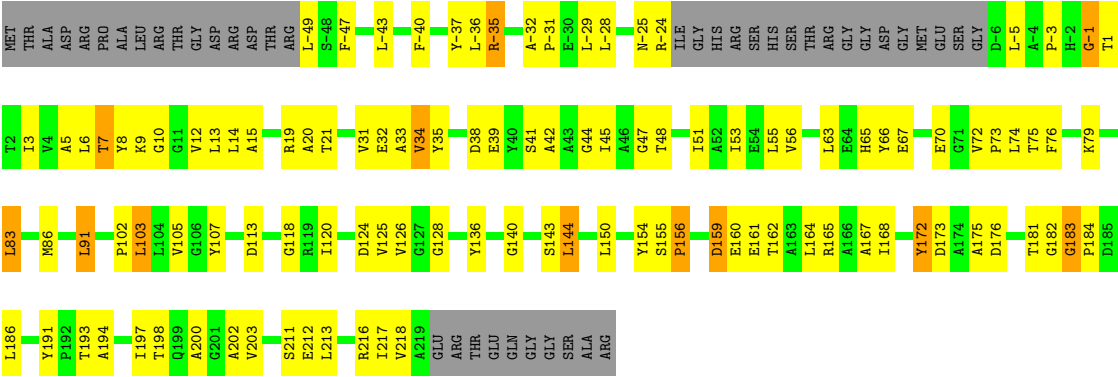
Chain M:





• Molecule 2: proteasome beta-type subunit 1

Chain N:



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	152.09Å 212.37Å 250.21Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.96 – 3.10	Depositor
% Data completeness (in resolution range)	88.7 (19.96-3.10)	Depositor
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.229 , 0.246	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	25013	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	54.0	wwPDB-VP

## 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	0.45	0/1720	0.64	0/2325
1	B	0.44	0/1720	0.64	0/2325
1	C	0.45	0/1720	0.64	0/2325
1	D	0.44	0/1720	0.63	0/2325
1	E	0.44	0/1720	0.63	0/2325
1	F	0.45	0/1720	0.64	0/2325
1	G	0.46	0/1720	0.64	0/2325
2	H	0.53	0/1892	0.72	0/2569
2	I	0.52	0/1892	0.72	0/2569
2	J	0.52	0/1892	0.72	0/2569
2	K	0.52	0/1892	0.72	0/2569
2	L	0.51	0/1892	0.72	0/2569
2	M	0.52	0/1892	0.71	0/2569
2	N	0.52	0/1892	0.72	0/2569
All	All	0.49	0/25284	0.68	0/34258

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1695	0	1693	110	0
1	B	1695	0	1693	105	0
1	C	1695	0	1693	104	1
1	D	1695	0	1693	103	0
1	E	1695	0	1693	93	0
1	F	1695	0	1693	99	0
1	G	1695	0	1693	103	0
2	H	1864	0	1837	88	0
2	I	1864	0	1837	79	0
2	J	1864	0	1837	75	0
2	K	1864	0	1837	78	0
2	L	1864	0	1837	78	0
2	M	1864	0	1837	76	0
2	N	1864	0	1837	84	0
3	A	4	0	0	0	0
3	B	5	0	0	1	0
3	C	3	0	0	3	0
3	D	6	0	0	4	0
3	E	5	0	0	1	0
3	F	2	0	0	1	0
3	G	1	0	0	0	0
3	H	13	0	0	3	0
3	I	14	0	0	3	0
3	J	13	0	0	0	0
3	K	8	0	0	2	0
3	L	12	0	0	3	0
3	M	6	0	0	2	0
3	N	8	0	0	1	0
All	All	25013	0	24710	1216	1

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 24.

The worst 5 of 1216 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:M:219:ALA:HA	3:M:231:HOH:O	1.59	1.02
2:I:-35:ARG:HH11	2:I:-35:ARG:HB2	1.26	0.97
2:L:-35:ARG:HH11	2:L:-35:ARG:HB2	1.29	0.97
2:N:-35:ARG:HB2	2:N:-35:ARG:HH11	1.28	0.97
2:M:-35:ARG:HH11	2:M:-35:ARG:HB2	1.30	0.96

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	Distance(Å)	Clash(Å)
1:C:172:ARG:NH1	1:C:172:ARG:NH1[4.576]	1.55	0.65

## 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	215/259 (83%)	178 (83%)	30 (14%)	7 (3%)	6	33
1	B	215/259 (83%)	179 (83%)	29 (14%)	7 (3%)	6	33
1	C	215/259 (83%)	181 (84%)	25 (12%)	9 (4%)	4	27
1	D	215/259 (83%)	179 (83%)	28 (13%)	8 (4%)	5	31
1	E	215/259 (83%)	180 (84%)	26 (12%)	9 (4%)	4	27
1	F	215/259 (83%)	178 (83%)	28 (13%)	9 (4%)	4	27
1	G	215/259 (83%)	178 (83%)	29 (14%)	8 (4%)	5	31
2	H	247/294 (84%)	224 (91%)	17 (7%)	6 (2%)	9	43
2	I	247/294 (84%)	221 (90%)	20 (8%)	6 (2%)	9	43
2	J	247/294 (84%)	222 (90%)	19 (8%)	6 (2%)	9	43
2	K	247/294 (84%)	222 (90%)	18 (7%)	7 (3%)	8	39
2	L	247/294 (84%)	223 (90%)	17 (7%)	7 (3%)	8	39
2	M	247/294 (84%)	224 (91%)	16 (6%)	7 (3%)	8	39
2	N	247/294 (84%)	224 (91%)	17 (7%)	6 (2%)	9	43
All	All	3234/3871 (84%)	2813 (87%)	319 (10%)	102 (3%)	6	35

5 of 102 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	28	ARG
1	A	224	ALA
1	B	28	ARG
1	B	224	ALA
1	C	28	ARG

### 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	174/206 (84%)	162 (93%)	12 (7%)	22	62
1	B	174/206 (84%)	162 (93%)	12 (7%)	22	62
1	C	174/206 (84%)	162 (93%)	12 (7%)	22	62
1	D	174/206 (84%)	163 (94%)	11 (6%)	25	66
1	E	174/206 (84%)	162 (93%)	12 (7%)	22	62
1	F	174/206 (84%)	163 (94%)	11 (6%)	25	66
1	G	174/206 (84%)	163 (94%)	11 (6%)	25	66
2	H	189/222 (85%)	173 (92%)	16 (8%)	15	51
2	I	189/222 (85%)	172 (91%)	17 (9%)	14	47
2	J	189/222 (85%)	172 (91%)	17 (9%)	14	47
2	K	189/222 (85%)	172 (91%)	17 (9%)	14	47
2	L	189/222 (85%)	173 (92%)	16 (8%)	15	51
2	M	189/222 (85%)	172 (91%)	17 (9%)	14	47
2	N	189/222 (85%)	171 (90%)	18 (10%)	12	42
All	All	2541/2996 (85%)	2342 (92%)	199 (8%)	18	57

5 of 199 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	H	113	ASP
2	I	172	TYR
2	N	21	THR
2	H	150	LEU
2	I	34	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 38 such sidechains are listed below:

Mol	Chain	Res	Type
2	H	65	HIS
2	I	81	ASN
2	N	-44	ASN

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
2	I	-44	ASN
2	J	-44	ASN

### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

There are no ligands in this entry.

### 5.7 Other polymers ⓘ

There are no such residues in this entry.

### 5.8 Polymer linkage issues

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

### 6.3 Carbohydrates ⓘ

EDS was not executed - this section will therefore be empty.

### 6.4 Ligands ⓘ

EDS was not executed - this section will therefore be empty.

### 6.5 Other polymers ⓘ

EDS was not executed - this section will therefore be empty.