wwPDB X-ray Structure Validation Summary Report (i)

Aug 16, 2023 – 08:47 AM EDT

PDB ID	:	1YAU
Title	:	Structure of Archeabacterial 20S proteasome- PA26 complex
Authors	:	Forster, A.; Masters, E.I.; Whitby, F.G.; Robinson, H.; Hill, C.P.
Deposited on	:	2004-12-17
Resolution	:	2.40 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

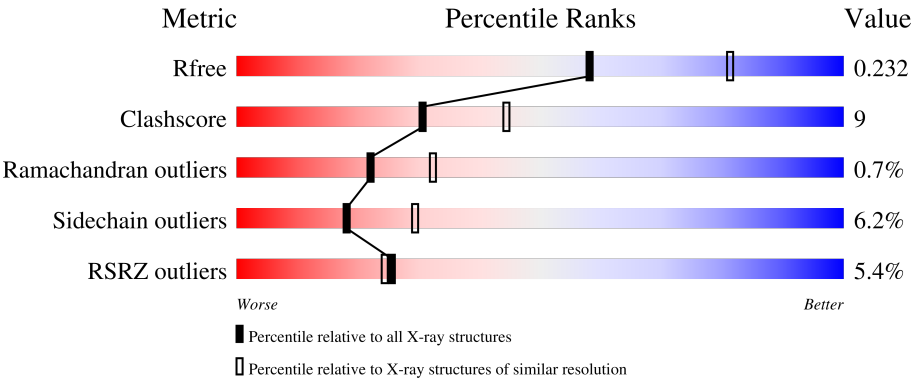
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 2.40 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}	130704	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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 of 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 <=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	233	8% 68% 21% 6% 5%
1	B	233	7% 71% 20% • 5%
1	C	233	9% 76% 14% • 5%
1	D	233	8% 70% 21% • 5%
1	E	233	5% 69% 22% 5% 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	233	8% 76% 16% • 5%
1	G	233	9% 75% 17% • 5%
2	H	217	% 73% 17% • 6%
2	I	217	% 73% 18% • 6%
2	J	217	% 76% 14% • 6%
2	K	217	2% 73% 18% • 6%
2	L	217	2% 70% 18% 5% • 6%
2	M	217	% 74% 18% • 6%
2	N	217	% 76% 14% • 6%
3	O	237	6% 75% 16% • 8%
3	P	237	5% 76% 15% • 8%
3	Q	237	6% 72% 19% 8%
3	R	237	7% 74% 17% 8%
3	S	237	5% 73% 18% 8%
3	T	237	5% 73% 18% • 8%
3	U	237	5% 77% 14% • 8%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	SO4	A	3007	-	-	X	-

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 36610 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 alpha subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	222	Total	C	N	O	S	0	0	0
			1728	1098	291	336	3			
1	B	222	Total	C	N	O	S	0	0	0
			1728	1098	291	336	3			
1	C	222	Total	C	N	O	S	0	0	0
			1728	1098	291	336	3			
1	D	222	Total	C	N	O	S	0	0	0
			1728	1098	291	336	3			
1	E	222	Total	C	N	O	S	0	0	0
			1728	1098	291	336	3			
1	F	222	Total	C	N	O	S	0	0	0
			1728	1098	291	336	3			
1	G	222	Total	C	N	O	S	0	0	0
			1728	1098	291	336	3			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	8	GLY	TYR	engineered mutation	UNP P25156
A	9	GLY	ASP	engineered mutation	UNP P25156
B	8	GLY	TYR	engineered mutation	UNP P25156
B	9	GLY	ASP	engineered mutation	UNP P25156
C	8	GLY	TYR	engineered mutation	UNP P25156
C	9	GLY	ASP	engineered mutation	UNP P25156
D	8	GLY	TYR	engineered mutation	UNP P25156
D	9	GLY	ASP	engineered mutation	UNP P25156
E	8	GLY	TYR	engineered mutation	UNP P25156
E	9	GLY	ASP	engineered mutation	UNP P25156
F	8	GLY	TYR	engineered mutation	UNP P25156
F	9	GLY	ASP	engineered mutation	UNP P25156
G	8	GLY	TYR	engineered mutation	UNP P25156
G	9	GLY	ASP	engineered mutation	UNP P25156

- Molecule 2 is a protein called Proteasome beta subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	203	Total	C	N	O	S	0	2	0
			1577	999	266	300	12			
2	I	203	Total	C	N	O	S	0	2	0
			1577	999	266	300	12			
2	J	203	Total	C	N	O	S	0	2	0
			1577	999	266	300	12			
2	K	203	Total	C	N	O	S	0	1	0
			1569	994	265	299	11			
2	L	203	Total	C	N	O	S	0	1	0
			1569	994	265	299	11			
2	M	203	Total	C	N	O	S	0	2	0
			1577	999	266	300	12			
2	N	203	Total	C	N	O	S	0	2	0
			1577	999	266	300	12			

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
H	204	HIS	-	expression tag	UNP P28061
H	205	HIS	-	expression tag	UNP P28061
H	206	HIS	-	expression tag	UNP P28061
H	207	HIS	-	expression tag	UNP P28061
H	208	HIS	-	expression tag	UNP P28061
H	209	HIS	-	expression tag	UNP P28061
I	204	HIS	-	expression tag	UNP P28061
I	205	HIS	-	expression tag	UNP P28061
I	206	HIS	-	expression tag	UNP P28061
I	207	HIS	-	expression tag	UNP P28061
I	208	HIS	-	expression tag	UNP P28061
I	209	HIS	-	expression tag	UNP P28061
J	204	HIS	-	expression tag	UNP P28061
J	205	HIS	-	expression tag	UNP P28061
J	206	HIS	-	expression tag	UNP P28061
J	207	HIS	-	expression tag	UNP P28061
J	208	HIS	-	expression tag	UNP P28061
J	209	HIS	-	expression tag	UNP P28061
K	204	HIS	-	expression tag	UNP P28061
K	205	HIS	-	expression tag	UNP P28061
K	206	HIS	-	expression tag	UNP P28061
K	207	HIS	-	expression tag	UNP P28061
K	208	HIS	-	expression tag	UNP P28061
K	209	HIS	-	expression tag	UNP P28061

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
L	204	HIS	-	expression tag	UNP P28061
L	205	HIS	-	expression tag	UNP P28061
L	206	HIS	-	expression tag	UNP P28061
L	207	HIS	-	expression tag	UNP P28061
L	208	HIS	-	expression tag	UNP P28061
L	209	HIS	-	expression tag	UNP P28061
M	204	HIS	-	expression tag	UNP P28061
M	205	HIS	-	expression tag	UNP P28061
M	206	HIS	-	expression tag	UNP P28061
M	207	HIS	-	expression tag	UNP P28061
M	208	HIS	-	expression tag	UNP P28061
M	209	HIS	-	expression tag	UNP P28061
N	204	HIS	-	expression tag	UNP P28061
N	205	HIS	-	expression tag	UNP P28061
N	206	HIS	-	expression tag	UNP P28061
N	207	HIS	-	expression tag	UNP P28061
N	208	HIS	-	expression tag	UNP P28061
N	209	HIS	-	expression tag	UNP P28061

- Molecule 3 is a protein called proteasome activator protein PA26.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	O	218	Total	C	N	O	S	0	0	0
			1679	1051	296	326	6			
3	P	218	Total	C	N	O	S	0	0	0
			1679	1051	296	326	6			
3	Q	218	Total	C	N	O	S	0	0	0
			1679	1051	296	326	6			
3	R	218	Total	C	N	O	S	0	0	0
			1679	1051	296	326	6			
3	S	218	Total	C	N	O	S	0	0	0
			1679	1051	296	326	6			
3	T	218	Total	C	N	O	S	0	0	0
			1679	1051	296	326	6			
3	U	218	Total	C	N	O	S	0	0	0
			1679	1051	296	326	6			

There are 56 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
O	-5	MET	-	initiating methionine	GB 5757773
O	-4	HIS	-	expression tag	GB 5757773

Continued on next page...

Continued from previous page...

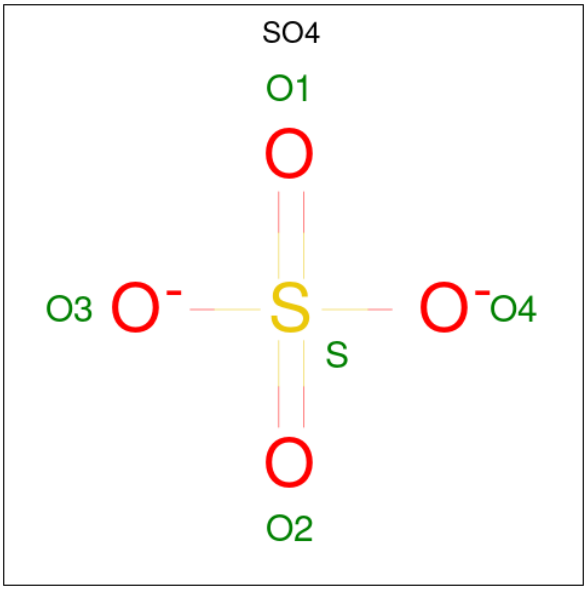
Chain	Residue	Modelled	Actual	Comment	Reference
O	-3	HIS	-	expression tag	GB 5757773
O	-2	HIS	-	expression tag	GB 5757773
O	-1	HIS	-	expression tag	GB 5757773
O	0	HIS	-	expression tag	GB 5757773
O	1	HIS	-	expression tag	GB 5757773
O	49	VAL	THR	variant	GB 5757773
P	-5	MET	-	initiating methionine	GB 5757773
P	-4	HIS	-	expression tag	GB 5757773
P	-3	HIS	-	expression tag	GB 5757773
P	-2	HIS	-	expression tag	GB 5757773
P	-1	HIS	-	expression tag	GB 5757773
P	0	HIS	-	expression tag	GB 5757773
P	1	HIS	-	expression tag	GB 5757773
P	49	VAL	THR	variant	GB 5757773
Q	-5	MET	-	initiating methionine	GB 5757773
Q	-4	HIS	-	expression tag	GB 5757773
Q	-3	HIS	-	expression tag	GB 5757773
Q	-2	HIS	-	expression tag	GB 5757773
Q	-1	HIS	-	expression tag	GB 5757773
Q	0	HIS	-	expression tag	GB 5757773
Q	1	HIS	-	expression tag	GB 5757773
Q	49	VAL	THR	variant	GB 5757773
R	-5	MET	-	initiating methionine	GB 5757773
R	-4	HIS	-	expression tag	GB 5757773
R	-3	HIS	-	expression tag	GB 5757773
R	-2	HIS	-	expression tag	GB 5757773
R	-1	HIS	-	expression tag	GB 5757773
R	0	HIS	-	expression tag	GB 5757773
R	1	HIS	-	expression tag	GB 5757773
R	49	VAL	THR	variant	GB 5757773
S	-5	MET	-	initiating methionine	GB 5757773
S	-4	HIS	-	expression tag	GB 5757773
S	-3	HIS	-	expression tag	GB 5757773
S	-2	HIS	-	expression tag	GB 5757773
S	-1	HIS	-	expression tag	GB 5757773
S	0	HIS	-	expression tag	GB 5757773
S	1	HIS	-	expression tag	GB 5757773
S	49	VAL	THR	variant	GB 5757773
T	-5	MET	-	initiating methionine	GB 5757773
T	-4	HIS	-	expression tag	GB 5757773
T	-3	HIS	-	expression tag	GB 5757773
T	-2	HIS	-	expression tag	GB 5757773

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
T	-1	HIS	-	expression tag	GB 5757773
T	0	HIS	-	expression tag	GB 5757773
T	1	HIS	-	expression tag	GB 5757773
T	49	VAL	THR	variant	GB 5757773
U	-5	MET	-	initiating methionine	GB 5757773
U	-4	HIS	-	expression tag	GB 5757773
U	-3	HIS	-	expression tag	GB 5757773
U	-2	HIS	-	expression tag	GB 5757773
U	-1	HIS	-	expression tag	GB 5757773
U	0	HIS	-	expression tag	GB 5757773
U	1	HIS	-	expression tag	GB 5757773
U	49	VAL	THR	variant	GB 5757773

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



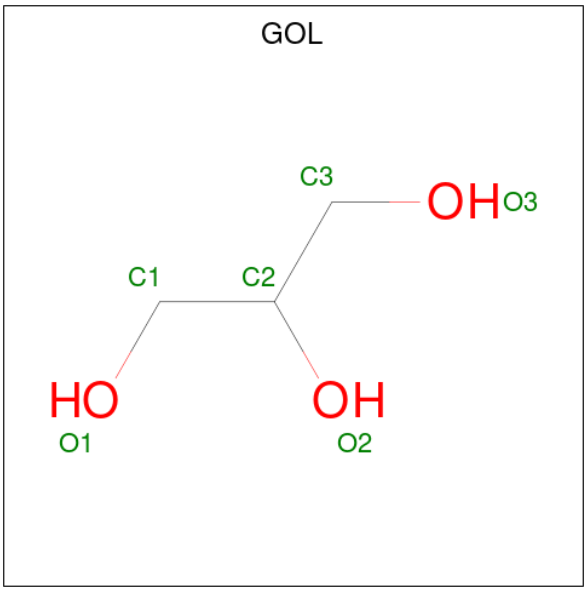
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total O S 5 4 1	0	0
4	B	1	Total O S 5 4 1	0	0
4	C	1	Total O S 5 4 1	0	0
4	D	1	Total O S 5 4 1	0	0
4	E	1	Total O S 5 4 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	F	1	Total O S 5 4 1	0	0
4	G	1	Total O S 5 4 1	0	0
4	H	1	Total O S 5 4 1	0	0
4	H	1	Total O S 5 4 1	0	0
4	I	1	Total O S 5 4 1	0	0
4	I	1	Total O S 5 4 1	0	0
4	I	1	Total O S 5 4 1	0	0
4	J	1	Total O S 5 4 1	0	0
4	J	1	Total O S 5 4 1	0	0
4	K	1	Total O S 5 4 1	0	0
4	K	1	Total O S 5 4 1	0	0
4	L	1	Total O S 5 4 1	0	0
4	L	1	Total O S 5 4 1	0	0
4	M	1	Total O S 5 4 1	0	0
4	M	1	Total O S 5 4 1	0	0
4	M	1	Total O S 5 4 1	0	0
4	M	1	Total O S 5 4 1	0	0
4	N	1	Total O S 5 4 1	0	0
4	N	1	Total O S 5 4 1	0	0
4	N	1	Total O S 5 4 1	0	0

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	H	1	Total C O 6 3 3	0	0
5	I	1	Total C O 6 3 3	0	0
5	J	1	Total C O 6 3 3	0	0
5	K	1	Total C O 6 3 3	0	0
5	L	1	Total C O 6 3 3	0	0
5	M	1	Total C O 6 3 3	0	0
5	N	1	Total C O 6 3 3	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	67	Total O 67 67	0	0
6	B	58	Total O 58 58	0	0
6	C	69	Total O 69 69	0	0
6	D	49	Total O 49 49	0	0
6	E	67	Total O 67 67	0	0

Continued on next page...

Continued from previous page...

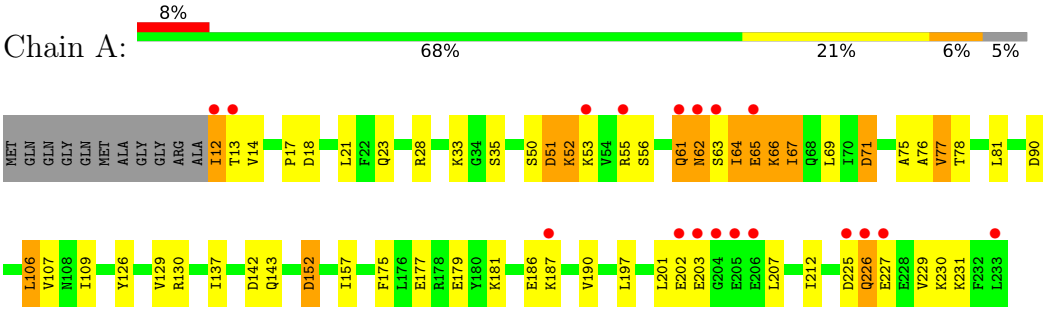
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	F	61	Total O 61 61	0	0
6	G	51	Total O 51 51	0	0
6	H	93	Total O 93 93	0	0
6	I	100	Total O 100 100	0	0
6	J	94	Total O 94 94	0	0
6	K	85	Total O 85 85	0	0
6	L	93	Total O 93 93	0	0
6	M	100	Total O 100 100	0	0
6	N	107	Total O 107 107	0	0
6	O	52	Total O 52 52	0	0
6	P	47	Total O 47 47	0	0
6	Q	47	Total O 47 47	0	0
6	R	69	Total O 69 69	0	0
6	S	90	Total O 90 90	0	0
6	T	97	Total O 97 97	0	0
6	U	75	Total O 75 75	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

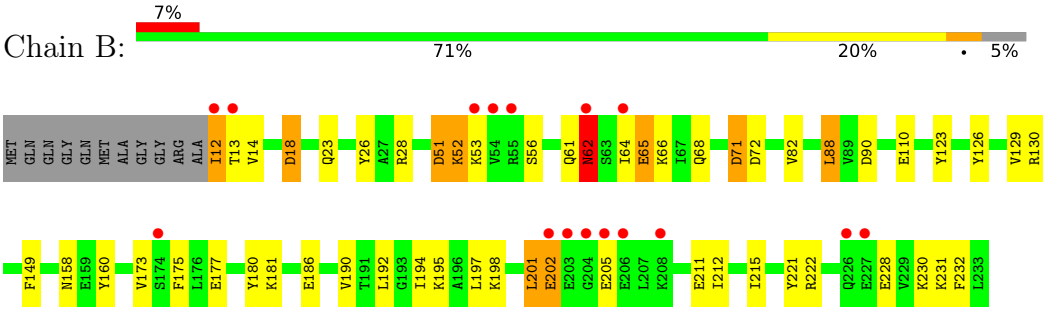
• Molecule 1: Proteasome alpha subunit

Chain A:



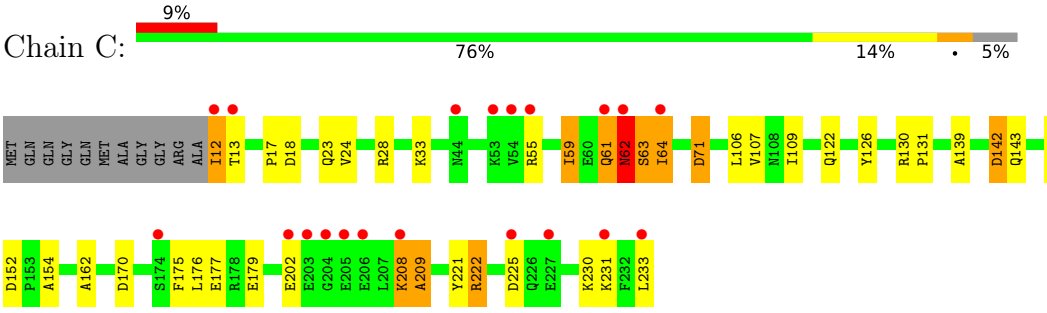
• Molecule 1: Proteasome alpha subunit

Chain B:



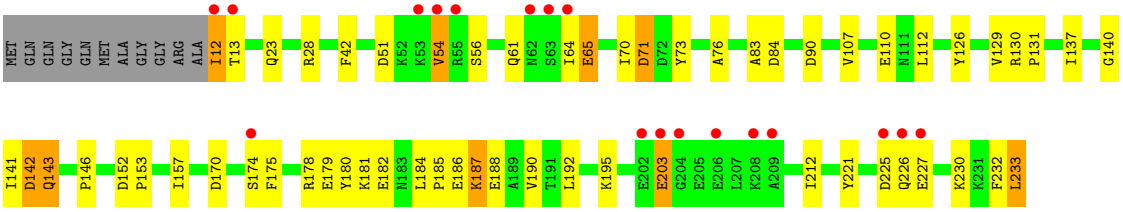
• Molecule 1: Proteasome alpha subunit

Chain C:

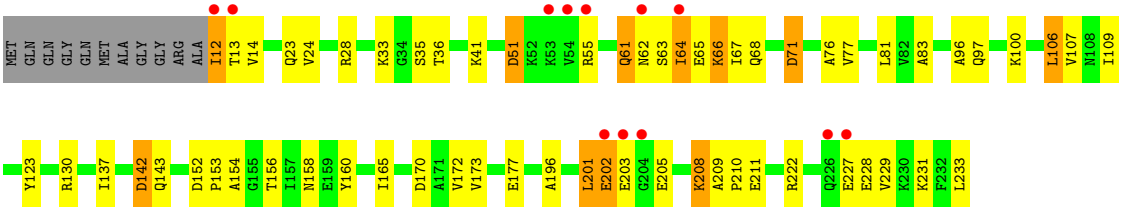


• Molecule 1: Proteasome alpha subunit

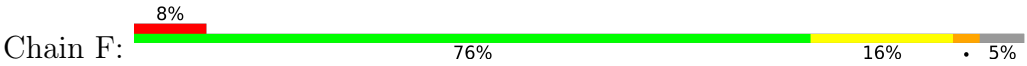
Chain D:

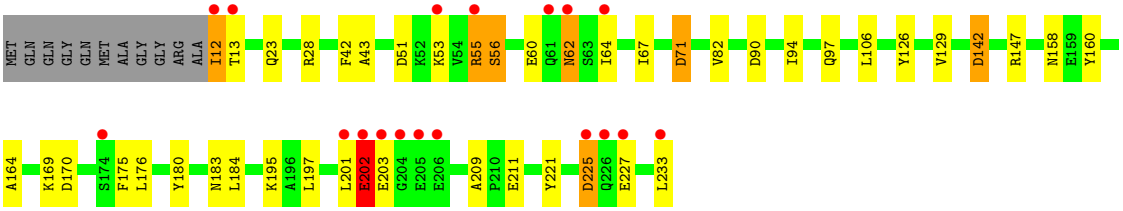


• Molecule 1: Proteasome alpha subunit

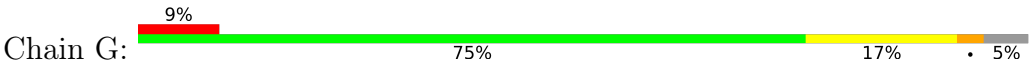


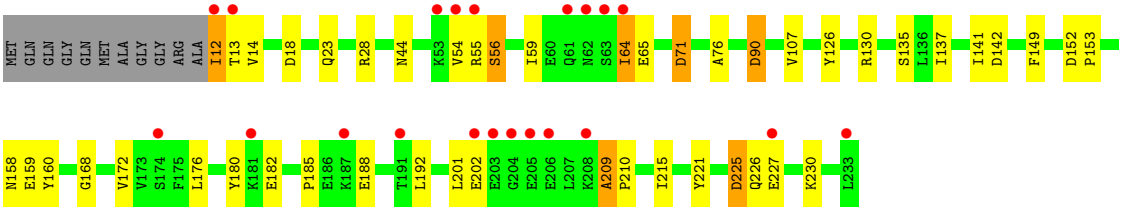
• Molecule 1: Proteasome alpha subunit



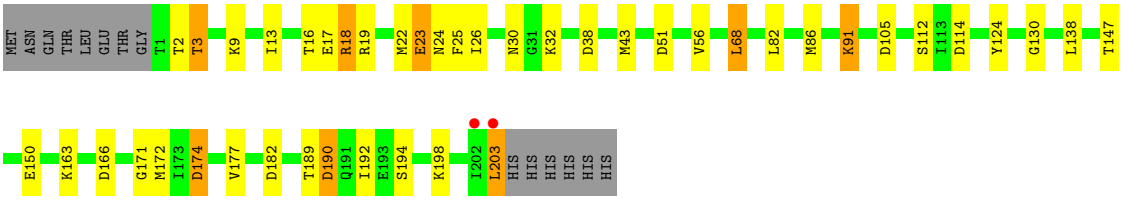


• Molecule 1: Proteasome alpha subunit



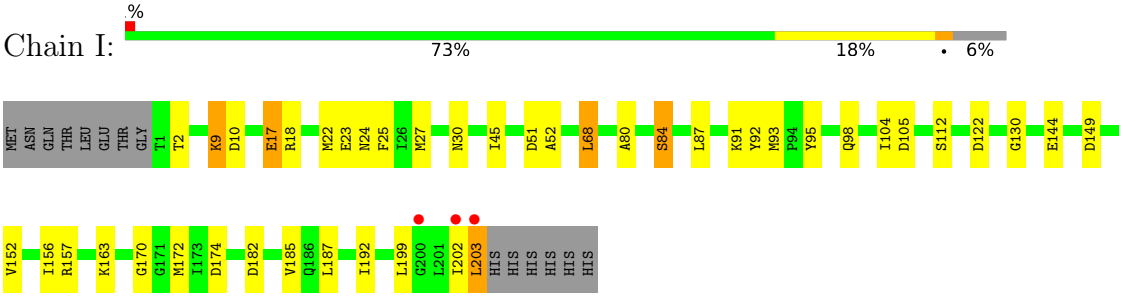


• Molecule 2: Proteasome beta subunit



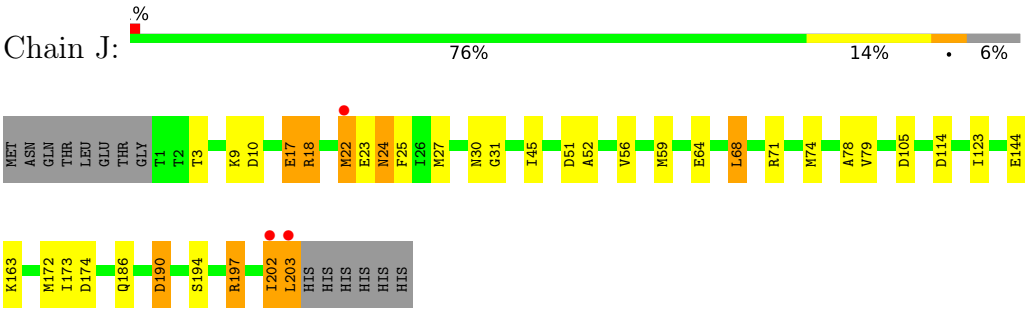
• Molecule 2: Proteasome beta subunit

Chain I:



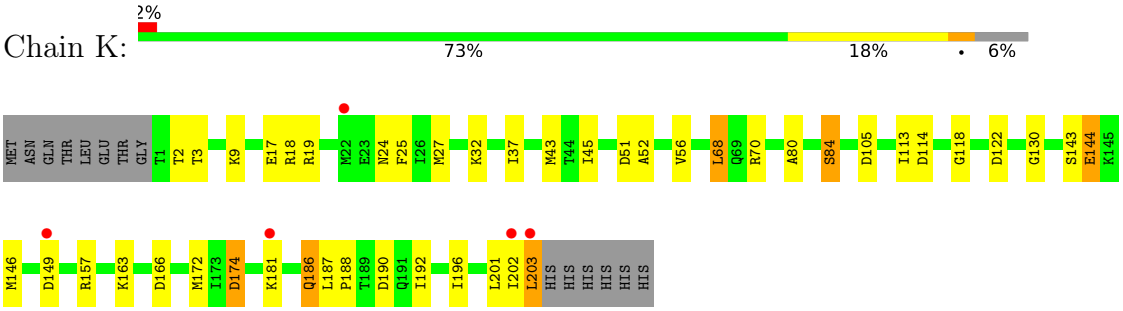
• Molecule 2: Proteasome beta subunit

Chain J:



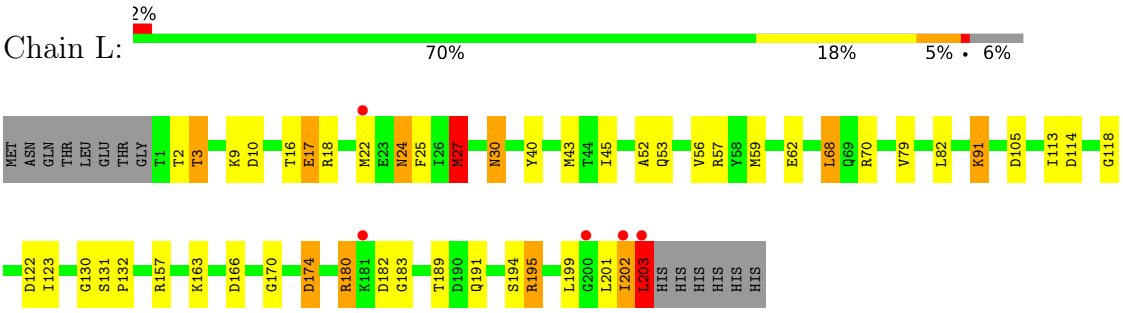
• Molecule 2: Proteasome beta subunit

Chain K:



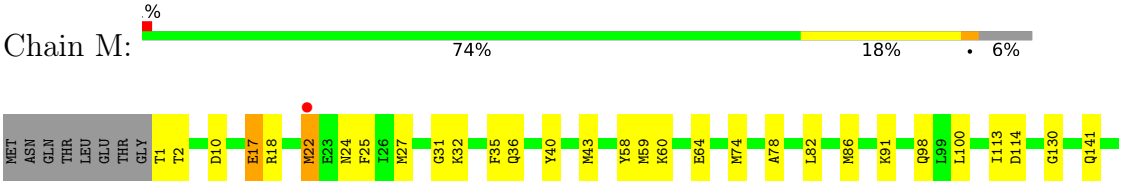
• Molecule 2: Proteasome beta subunit

Chain L:

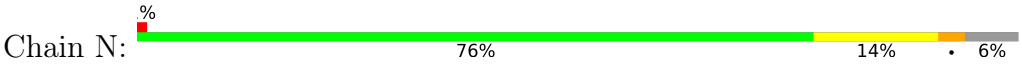


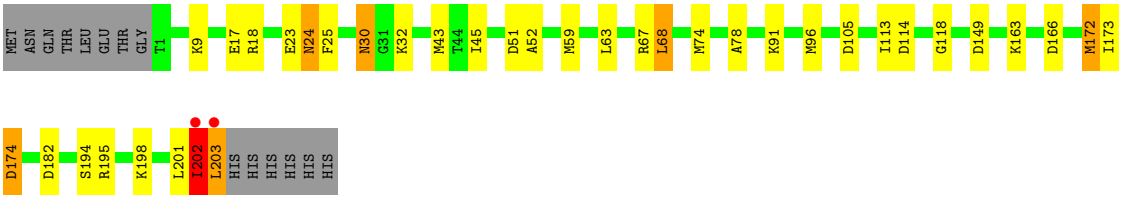
• Molecule 2: Proteasome beta subunit

Chain M:

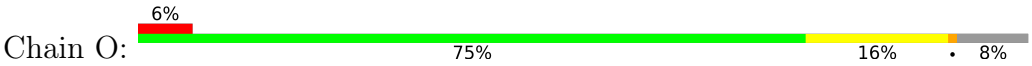


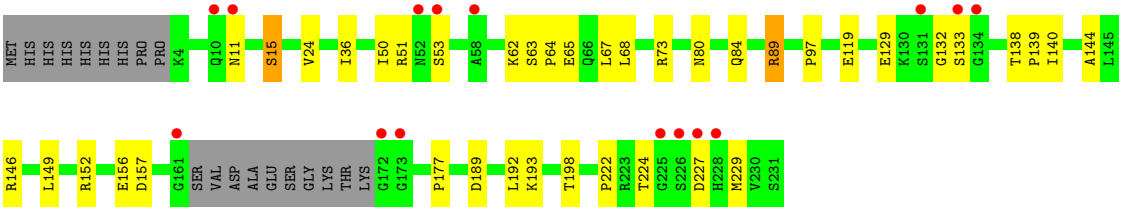
• Molecule 2: Proteasome beta subunit



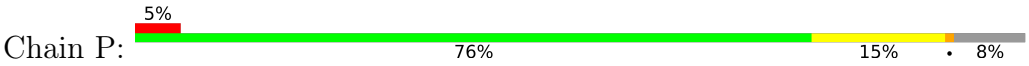


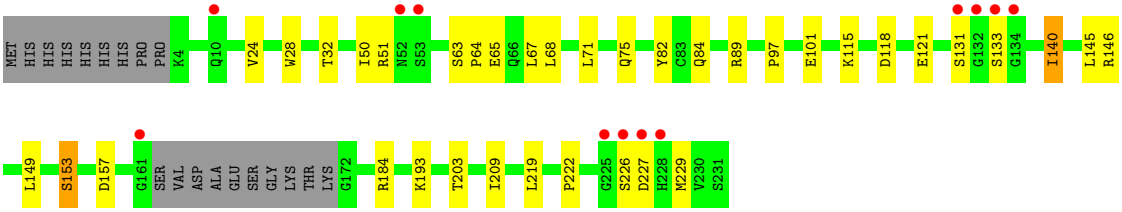
• Molecule 3: proteasome activator protein PA26



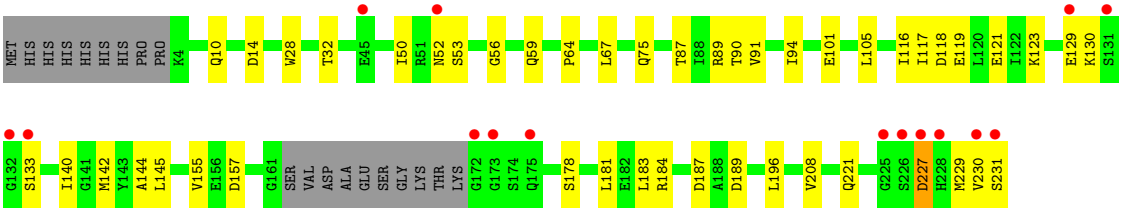


• Molecule 3: proteasome activator protein PA26

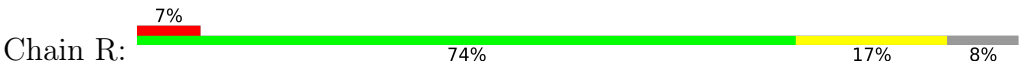


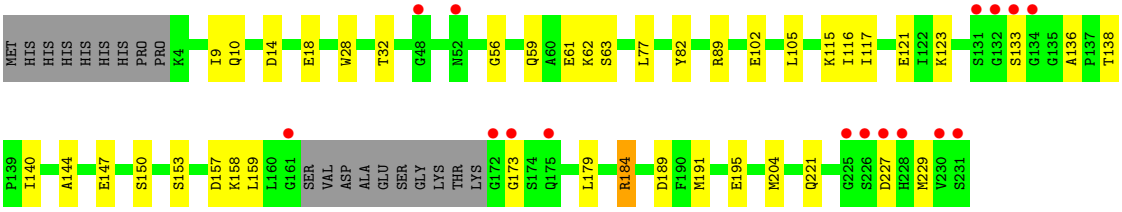


• Molecule 3: proteasome activator protein PA26

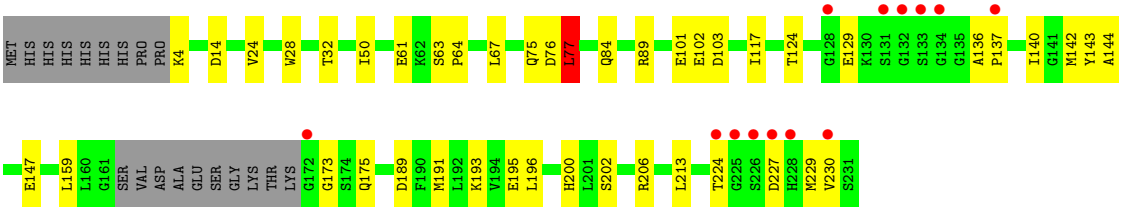


• Molecule 3: proteasome activator protein PA26

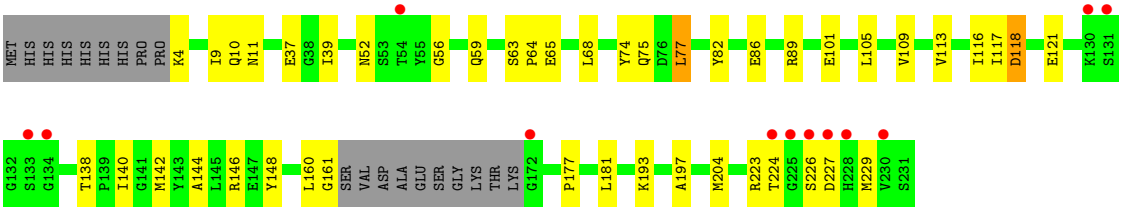




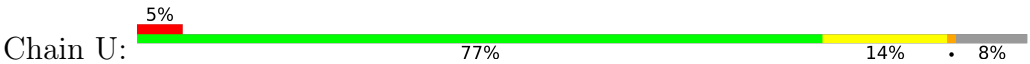
• Molecule 3: proteasome activator protein PA26

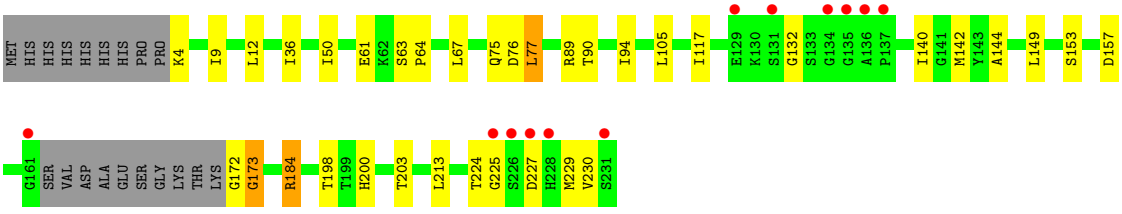


• Molecule 3: proteasome activator protein PA26



• Molecule 3: proteasome activator protein PA26





4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	255.22Å 126.91Å 181.03Å 90.00° 92.42° 90.00°	Depositor
Resolution (Å)	8.00 – 2.40 8.00 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.8 (8.00-2.40) 99.8 (8.00-2.40)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	3.29 (at 2.40Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.180 , 0.233 0.180 , 0.232	Depositor DCC
R_{free} test set	1069 reflections (0.49%)	wwPDB-VP
Wilson B-factor (Å^2)	35.2	Xtriage
Anisotropy	0.028	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.51 , 74.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.000 for -h,-k,l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	36610	wwPDB-VP
Average B, all atoms (Å^2)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.75% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: SO4, GOL

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	0.80	0/1751	0.88	3/2359 (0.1%)
1	B	0.83	0/1751	0.91	5/2359 (0.2%)
1	C	0.80	0/1751	0.87	6/2359 (0.3%)
1	D	0.77	0/1751	0.91	6/2359 (0.3%)
1	E	0.81	0/1751	0.88	5/2359 (0.2%)
1	F	0.74	0/1751	0.85	4/2359 (0.2%)
1	G	0.78	0/1751	0.89	3/2359 (0.1%)
2	H	0.85	0/1597	0.94	6/2157 (0.3%)
2	I	0.84	0/1597	0.95	5/2157 (0.2%)
2	J	0.88	0/1597	0.93	5/2157 (0.2%)
2	K	0.82	0/1589	1.00	10/2147 (0.5%)
2	L	0.84	0/1589	0.97	8/2147 (0.4%)
2	M	0.87	0/1597	0.96	4/2157 (0.2%)
2	N	0.82	0/1597	0.95	6/2157 (0.3%)
3	O	0.66	0/1702	0.79	2/2299 (0.1%)
3	P	0.64	0/1702	0.78	2/2299 (0.1%)
3	Q	0.66	0/1702	0.79	3/2299 (0.1%)
3	R	0.71	0/1702	0.81	3/2299 (0.1%)
3	S	0.75	0/1702	0.84	5/2299 (0.2%)
3	T	0.77	0/1702	0.83	3/2299 (0.1%)
3	U	0.76	0/1702	0.82	2/2299 (0.1%)
All	All	0.78	0/35334	0.88	96/47685 (0.2%)

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	C	0	1

There are no bond length outliers.

The worst 5 of 96 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	K	174	ASP	CB-CG-OD2	10.79	128.01	118.30
2	N	203	LEU	CA-CB-CG	9.32	136.73	115.30
3	S	227	ASP	CB-CG-OD2	8.09	125.58	118.30
1	A	142	ASP	CB-CG-OD2	7.99	125.49	118.30
1	F	142	ASP	CB-CG-OD2	7.72	125.25	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	62	ASN	Peptide

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	1728	0	1764	54	0
1	B	1728	0	1764	44	0
1	C	1728	0	1764	41	0
1	D	1728	0	1764	45	0
1	E	1728	0	1764	54	0
1	F	1728	0	1764	38	0
1	G	1728	0	1764	32	0
2	H	1577	0	1625	27	0
2	I	1577	0	1625	29	0
2	J	1577	0	1625	24	0
2	K	1569	0	1617	30	0
2	L	1569	0	1617	29	0
2	M	1577	0	1625	33	0
2	N	1577	0	1625	28	0
3	O	1679	0	1700	32	0
3	P	1679	0	1700	28	0
3	Q	1679	0	1700	32	0
3	R	1679	0	1700	34	0
3	S	1679	0	1700	32	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	T	1679	0	1700	38	0
3	U	1679	0	1700	36	0
4	A	5	0	0	2	0
4	B	5	0	0	0	0
4	C	5	0	0	0	0
4	D	5	0	0	1	0
4	E	5	0	0	0	0
4	F	5	0	0	0	0
4	G	5	0	0	0	0
4	H	10	0	0	0	0
4	I	15	0	0	1	0
4	J	10	0	0	1	0
4	K	10	0	0	0	0
4	L	10	0	0	1	0
4	M	20	0	0	1	0
4	N	15	0	0	1	0
5	H	6	0	8	0	0
5	I	6	0	8	0	0
5	J	6	0	8	0	0
5	K	6	0	8	0	0
5	L	6	0	8	2	0
5	M	6	0	8	0	0
5	N	6	0	8	0	0
6	A	67	0	0	7	0
6	B	58	0	0	2	0
6	C	69	0	0	2	0
6	D	49	0	0	3	0
6	E	67	0	0	10	0
6	F	61	0	0	5	0
6	G	51	0	0	3	0
6	H	93	0	0	3	0
6	I	100	0	0	9	0
6	J	94	0	0	14	0
6	K	85	0	0	12	0
6	L	93	0	0	1	0
6	M	100	0	0	6	0
6	N	107	0	0	8	0
6	O	52	0	0	3	0
6	P	47	0	0	2	0
6	Q	47	0	0	1	0
6	R	69	0	0	4	0
6	S	90	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	T	97	0	0	1	0
6	U	75	0	0	5	0
All	All	36610	0	35663	652	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 9.

The worst 5 of 652 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)
3:U:184:ARG:HG2	3:U:184:ARG:HH11	0.95	1.11
1:E:143:GLN:HG3	6:E:3051:HOH:O	1.49	1.10
1:D:28:ARG:HH22	3:U:229:MET:HG2	1.02	1.08
2:L:27:MET:HG2	6:M:3108:HOH:O	1.53	1.08
1:G:28:ARG:HH22	3:Q:229:MET:HG2	1.20	1.05

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	220/233 (94%)	208 (94%)	6 (3%)	6 (3%)	5 5
1	B	220/233 (94%)	204 (93%)	13 (6%)	3 (1%)	11 15
1	C	220/233 (94%)	211 (96%)	5 (2%)	4 (2%)	8 10
1	D	220/233 (94%)	202 (92%)	16 (7%)	2 (1%)	17 25
1	E	220/233 (94%)	210 (96%)	8 (4%)	2 (1%)	17 25
1	F	220/233 (94%)	206 (94%)	11 (5%)	3 (1%)	11 15
1	G	220/233 (94%)	205 (93%)	10 (4%)	5 (2%)	6 7
2	H	203/217 (94%)	195 (96%)	8 (4%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	I	203/217 (94%)	196 (97%)	7 (3%)	0	100	100
2	J	203/217 (94%)	194 (96%)	9 (4%)	0	100	100
2	K	202/217 (93%)	196 (97%)	6 (3%)	0	100	100
2	L	202/217 (93%)	197 (98%)	4 (2%)	1 (0%)	29	41
2	M	203/217 (94%)	196 (97%)	7 (3%)	0	100	100
2	N	203/217 (94%)	195 (96%)	7 (3%)	1 (0%)	29	41
3	O	214/237 (90%)	208 (97%)	6 (3%)	0	100	100
3	P	214/237 (90%)	209 (98%)	4 (2%)	1 (0%)	29	41
3	Q	214/237 (90%)	206 (96%)	8 (4%)	0	100	100
3	R	214/237 (90%)	206 (96%)	8 (4%)	0	100	100
3	S	214/237 (90%)	206 (96%)	8 (4%)	0	100	100
3	T	214/237 (90%)	208 (97%)	6 (3%)	0	100	100
3	U	214/237 (90%)	208 (97%)	5 (2%)	1 (0%)	29	41
All	All	4457/4809 (93%)	4266 (96%)	162 (4%)	29 (1%)	22	32

5 of 29 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	226	GLN
1	C	62	ASN
1	C	209	ALA
1	D	54	VAL
1	D	56	SER

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	185/191 (97%)	165 (89%)	20 (11%)	6	9
1	B	185/191 (97%)	172 (93%)	13 (7%)	15	24
1	C	185/191 (97%)	174 (94%)	11 (6%)	19	32

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	D	185/191 (97%)	171 (92%)	14 (8%)	13 20
1	E	185/191 (97%)	173 (94%)	12 (6%)	17 27
1	F	185/191 (97%)	176 (95%)	9 (5%)	25 40
1	G	185/191 (97%)	174 (94%)	11 (6%)	19 32
2	H	172/183 (94%)	160 (93%)	12 (7%)	15 24
2	I	172/183 (94%)	158 (92%)	14 (8%)	11 18
2	J	172/183 (94%)	152 (88%)	20 (12%)	5 7
2	K	171/183 (93%)	160 (94%)	11 (6%)	17 28
2	L	171/183 (93%)	154 (90%)	17 (10%)	8 11
2	M	172/183 (94%)	163 (95%)	9 (5%)	23 38
2	N	172/183 (94%)	159 (92%)	13 (8%)	13 20
3	O	179/196 (91%)	171 (96%)	8 (4%)	27 44
3	P	179/196 (91%)	174 (97%)	5 (3%)	43 63
3	Q	179/196 (91%)	171 (96%)	8 (4%)	27 44
3	R	179/196 (91%)	170 (95%)	9 (5%)	24 40
3	S	179/196 (91%)	176 (98%)	3 (2%)	60 78
3	T	179/196 (91%)	171 (96%)	8 (4%)	27 44
3	U	179/196 (91%)	175 (98%)	4 (2%)	52 71
All	All	3750/3990 (94%)	3519 (94%)	231 (6%)	18 29

5 of 231 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	I	203	LEU
3	T	37	GLU
2	K	144	GLU
3	T	4	LYS
3	Q	10	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 45 such sidechains are listed below:

Mol	Chain	Res	Type
2	N	191	GLN
3	R	79	HIS
3	O	52	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	Q	59	GLN
3	S	47	HIS

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

32 ligands 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	$\#\|Z\| > 2$
4	SO4	M	3024	-	4,4,4	0.22	0	6,6,6	0.20	0
4	SO4	K	3013	-	4,4,4	0.16	0	6,6,6	0.52	0
4	SO4	M	3025	-	4,4,4	0.17	0	6,6,6	0.41	0
4	SO4	F	3005	-	4,4,4	0.25	0	6,6,6	0.13	0
4	SO4	L	3014	-	4,4,4	0.16	0	6,6,6	0.51	0
5	GOL	N	2003	-	5,5,5	0.34	0	5,5,5	0.35	0
4	SO4	I	3011	-	4,4,4	0.19	0	6,6,6	0.66	0
4	SO4	L	3019	-	4,4,4	0.07	0	6,6,6	0.56	0
4	SO4	M	3020	-	4,4,4	0.15	0	6,6,6	0.58	0
4	SO4	E	3004	-	4,4,4	0.17	0	6,6,6	0.17	0
4	SO4	J	3012	-	4,4,4	0.15	0	6,6,6	0.41	0
5	GOL	J	2006	-	5,5,5	0.47	0	5,5,5	0.35	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
5	GOL	L	2001	-	5,5,5	0.33	0	5,5,5	0.72	0
4	SO4	H	3008	-	4,4,4	0.29	0	6,6,6	0.58	0
5	GOL	M	2002	-	5,5,5	0.54	0	5,5,5	0.87	0
5	GOL	K	2007	-	5,5,5	0.75	0	5,5,5	1.19	1 (20%)
4	SO4	J	3017	-	4,4,4	0.20	0	6,6,6	0.41	0
4	SO4	I	3016	-	4,4,4	0.18	0	6,6,6	0.79	0
4	SO4	N	3009	-	4,4,4	0.12	0	6,6,6	0.65	0
4	SO4	K	3018	-	4,4,4	0.17	0	6,6,6	0.57	0
4	SO4	M	3010	-	4,4,4	0.22	0	6,6,6	0.67	0
4	SO4	A	3007	-	4,4,4	0.14	0	6,6,6	0.17	0
4	SO4	N	3022	-	4,4,4	0.18	0	6,6,6	0.19	0
4	SO4	N	3021	-	4,4,4	0.13	0	6,6,6	0.33	0
4	SO4	D	3003	-	4,4,4	0.12	0	6,6,6	0.21	0
4	SO4	H	3015	-	4,4,4	0.12	0	6,6,6	0.99	0
4	SO4	G	3006	-	4,4,4	0.15	0	6,6,6	0.16	0
4	SO4	B	3001	-	4,4,4	0.15	0	6,6,6	0.12	0
4	SO4	C	3002	-	4,4,4	0.14	0	6,6,6	0.09	0
4	SO4	I	3023	-	4,4,4	0.28	0	6,6,6	0.21	0
5	GOL	I	2005	-	5,5,5	0.31	0	5,5,5	0.81	0
5	GOL	H	2004	-	5,5,5	0.38	0	5,5,5	0.98	0

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
5	GOL	J	2006	-	-	2/4/4/4	-
5	GOL	L	2001	-	-	1/4/4/4	-
5	GOL	M	2002	-	-	4/4/4/4	-
5	GOL	K	2007	-	-	2/4/4/4	-
5	GOL	N	2003	-	-	2/4/4/4	-
5	GOL	I	2005	-	-	4/4/4/4	-
5	GOL	H	2004	-	-	4/4/4/4	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	K	2007	GOL	O3-C3-C2	-2.11	100.10	110.20

There are no chirality outliers.

5 of 19 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	H	2004	GOL	C1-C2-C3-O3
5	H	2004	GOL	O2-C2-C3-O3
5	I	2005	GOL	C1-C2-C3-O3
5	I	2005	GOL	O2-C2-C3-O3
5	J	2006	GOL	O1-C1-C2-C3

There are no ring outliers.

8 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	L	3019	SO4	1	0
5	L	2001	GOL	2	0
4	J	3017	SO4	1	0
4	I	3016	SO4	1	0
4	M	3010	SO4	1	0
4	A	3007	SO4	2	0
4	N	3021	SO4	1	0
4	D	3003	SO4	1	0

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, 95^{th} 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	222/233 (95%)	-0.26	18 (8%)	12	11	21, 36, 77, 85	0
1	B	222/233 (95%)	-0.25	16 (7%)	15	14	21, 36, 72, 86	0
1	C	222/233 (95%)	-0.25	20 (9%)	9	8	21, 36, 73, 86	0
1	D	222/233 (95%)	-0.16	18 (8%)	12	11	23, 39, 76, 87	0
1	E	222/233 (95%)	-0.34	12 (5%)	25	24	23, 35, 66, 78	0
1	F	222/233 (95%)	-0.16	18 (8%)	12	11	24, 39, 77, 88	0
1	G	222/233 (95%)	-0.10	21 (9%)	8	7	24, 38, 75, 88	0
2	H	203/217 (93%)	-0.72	2 (0%)	82	80	20, 28, 48, 78	0
2	I	203/217 (93%)	-0.68	3 (1%)	73	72	21, 29, 50, 71	0
2	J	203/217 (93%)	-0.66	3 (1%)	73	72	22, 30, 51, 76	0
2	K	203/217 (93%)	-0.67	5 (2%)	57	55	22, 30, 51, 74	0
2	L	203/217 (93%)	-0.63	5 (2%)	57	55	22, 30, 50, 73	0
2	M	203/217 (93%)	-0.73	3 (1%)	73	72	21, 29, 46, 69	0
2	N	203/217 (93%)	-0.72	2 (0%)	82	80	20, 29, 49, 71	0
3	O	218/237 (91%)	-0.00	15 (6%)	16	15	30, 46, 72, 80	0
3	P	218/237 (91%)	0.08	12 (5%)	25	24	32, 49, 77, 86	0
3	Q	218/237 (91%)	0.13	15 (6%)	16	15	30, 50, 76, 93	0
3	R	218/237 (91%)	-0.03	16 (7%)	15	13	25, 44, 76, 85	0
3	S	218/237 (91%)	-0.28	13 (5%)	21	20	23, 37, 72, 85	0
3	T	218/237 (91%)	-0.23	12 (5%)	25	24	24, 38, 69, 84	0
3	U	218/237 (91%)	-0.20	12 (5%)	25	24	25, 38, 73, 87	0
All	All	4501/4809 (93%)	-0.32	241 (5%)	25	24	20, 36, 72, 93	0

The worst 5 of 241 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	O	225	GLY	10.0
3	U	225	GLY	8.8
3	Q	227	ASP	8.6
2	L	203	LEU	8.5
2	I	203	LEU	8.5

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 monosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	SO4	E	3004	5/5	0.78	0.37	120,121,121,121	0
4	SO4	A	3007	5/5	0.82	0.30	116,117,117,117	0
4	SO4	F	3005	5/5	0.83	0.30	112,112,113,113	0
4	SO4	B	3001	5/5	0.85	0.25	128,128,128,128	0
4	SO4	D	3003	5/5	0.90	0.21	105,105,105,106	0
4	SO4	G	3006	5/5	0.90	0.20	114,114,114,114	0
4	SO4	I	3023	5/5	0.92	0.15	83,83,84,84	0
4	SO4	C	3002	5/5	0.93	0.19	110,110,110,110	0
4	SO4	M	3024	5/5	0.93	0.14	82,84,85,86	0
4	SO4	M	3025	5/5	0.94	0.27	77,77,79,79	0
4	SO4	N	3022	5/5	0.94	0.15	89,90,91,91	0
5	GOL	J	2006	6/6	0.94	0.10	37,41,46,50	0
5	GOL	K	2007	6/6	0.94	0.11	38,39,41,42	0
5	GOL	H	2004	6/6	0.95	0.10	42,42,44,44	0
5	GOL	L	2001	6/6	0.95	0.11	41,41,42,44	0
4	SO4	H	3015	5/5	0.96	0.17	47,49,51,51	0
4	SO4	L	3019	5/5	0.96	0.14	56,58,58,59	0
4	SO4	I	3016	5/5	0.96	0.14	51,52,56,56	0
5	GOL	M	2002	6/6	0.96	0.09	49,51,52,52	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	SO4	H	3008	5/5	0.97	0.17	47,47,50,52	0
4	SO4	N	3021	5/5	0.97	0.12	54,55,58,59	0
4	SO4	M	3020	5/5	0.97	0.15	63,64,64,65	0
4	SO4	K	3018	5/5	0.97	0.16	56,57,57,57	0
5	GOL	N	2003	6/6	0.97	0.08	44,45,46,46	0
4	SO4	J	3012	5/5	0.98	0.13	44,44,47,48	0
4	SO4	M	3010	5/5	0.98	0.18	50,53,55,55	0
5	GOL	I	2005	6/6	0.98	0.09	36,38,41,44	0
4	SO4	J	3017	5/5	0.98	0.12	49,50,50,52	0
4	SO4	K	3013	5/5	0.98	0.15	42,43,47,49	0
4	SO4	I	3011	5/5	0.98	0.10	38,40,43,46	0
4	SO4	N	3009	5/5	0.98	0.16	46,47,50,50	0
4	SO4	L	3014	5/5	0.98	0.18	46,47,50,51	0

6.5 Other polymers (i)

There are no such residues in this entry.