



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 07:34 pm BST

PDB ID : 2YFY  
Title : Crystal structure of the allosteric-defective chaperonin GroEL E434K mutant  
Authors : Cabo-Bilbao, A.; Mechaly, A.E.; Agirre, J.; Spinelli, S.; Sot, B.; Muga, A.;  
Guerin, D.M.A.  
Deposited on : 2011-03-31  
Resolution : 4.50 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
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

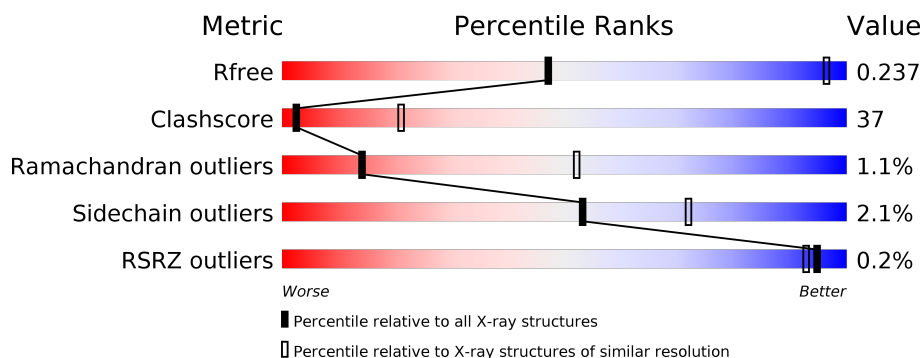
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 4.50 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	1055 (5.20-3.80)
Clashscore	141614	1123 (5.20-3.80)
Ramachandran outliers	138981	1069 (5.20-3.80)
Sidechain outliers	138945	1050 (5.20-3.80)
RSRZ outliers	127900	1101 (5.30-3.70)

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 respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	524	<div> <div></div> <div>44%54%</div> <div></div> </div>
1	B	524	<div> <div></div> <div>53%44%</div> <div></div> </div>
1	C	524	<div> <div></div> <div>47%51%</div> <div></div> </div>
1	D	524	<div> <div></div> <div>47%50%</div> <div></div> </div>
1	E	524	<div> <div></div> <div>48%48%</div> <div></div> </div>
1	F	524	<div> <div></div> <div>47%50%</div> <div></div> </div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	G	524	
1	H	524	
1	I	524	
1	J	524	
1	K	524	
1	L	524	
1	M	524	
1	N	524	

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 53984 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 60 KDA CHAPERONIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	524	Total	C	N	O	S	0	0	0
			3856	2398	666	772	20			
1	B	524	Total	C	N	O	S	0	0	0
			3856	2398	666	772	20			
1	C	524	Total	C	N	O	S	0	0	0
			3856	2398	666	772	20			
1	D	524	Total	C	N	O	S	0	0	0
			3856	2398	666	772	20			
1	E	524	Total	C	N	O	S	0	0	0
			3856	2398	666	772	20			
1	F	524	Total	C	N	O	S	0	0	0
			3856	2398	666	772	20			
1	G	524	Total	C	N	O	S	0	0	0
			3856	2398	666	772	20			
1	H	524	Total	C	N	O	S	0	0	0
			3856	2398	666	772	20			
1	I	524	Total	C	N	O	S	0	0	0
			3856	2398	666	772	20			
1	J	524	Total	C	N	O	S	0	0	0
			3856	2398	666	772	20			
1	K	524	Total	C	N	O	S	0	0	0
			3856	2398	666	772	20			
1	L	524	Total	C	N	O	S	0	0	0
			3856	2398	666	772	20			
1	M	524	Total	C	N	O	S	0	0	0
			3856	2398	666	772	20			
1	N	524	Total	C	N	O	S	0	0	0
			3856	2398	666	772	20			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	434	LYS	GLU	engineered mutation	UNP P0A6F5

*Continued on next page...*

*Continued from previous page...*

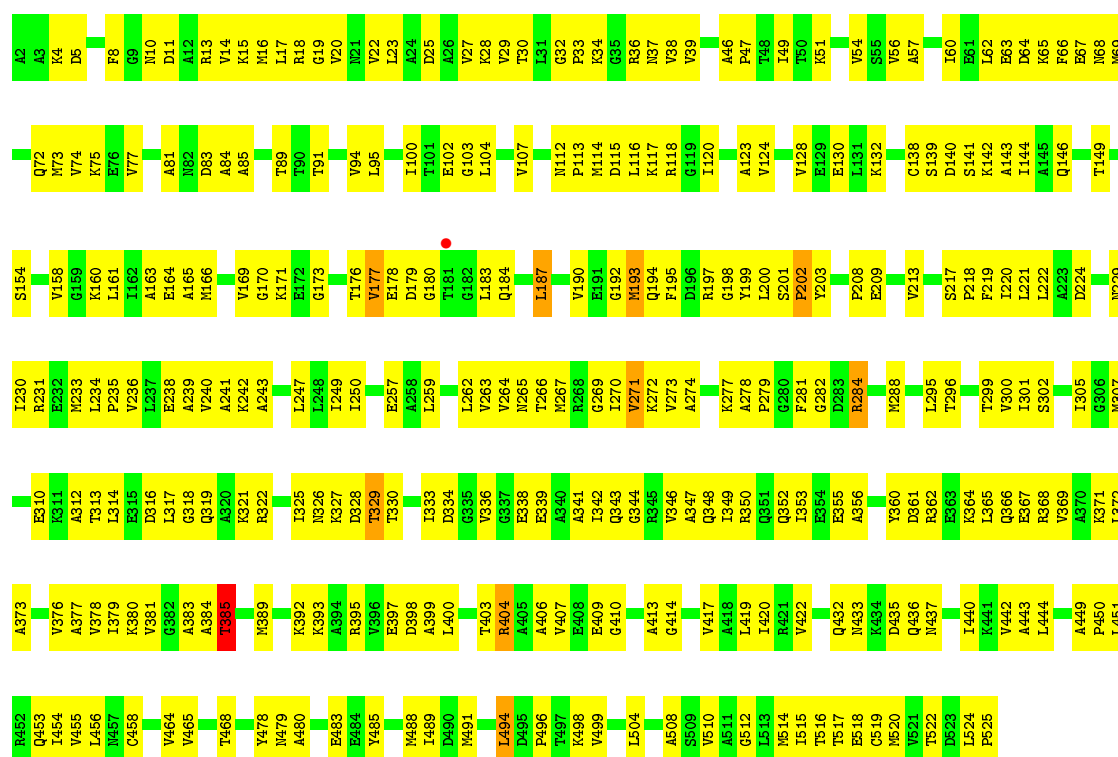
Chain	Residue	Modelled	Actual	Comment	Reference
B	434	LYS	GLU	engineered mutation	UNP P0A6F5
C	434	LYS	GLU	engineered mutation	UNP P0A6F5
D	434	LYS	GLU	engineered mutation	UNP P0A6F5
E	434	LYS	GLU	engineered mutation	UNP P0A6F5
F	434	LYS	GLU	engineered mutation	UNP P0A6F5
G	434	LYS	GLU	engineered mutation	UNP P0A6F5
H	434	LYS	GLU	engineered mutation	UNP P0A6F5
I	434	LYS	GLU	engineered mutation	UNP P0A6F5
J	434	LYS	GLU	engineered mutation	UNP P0A6F5
K	434	LYS	GLU	engineered mutation	UNP P0A6F5
L	434	LYS	GLU	engineered mutation	UNP P0A6F5
M	434	LYS	GLU	engineered mutation	UNP P0A6F5
N	434	LYS	GLU	engineered mutation	UNP P0A6F5

### 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.

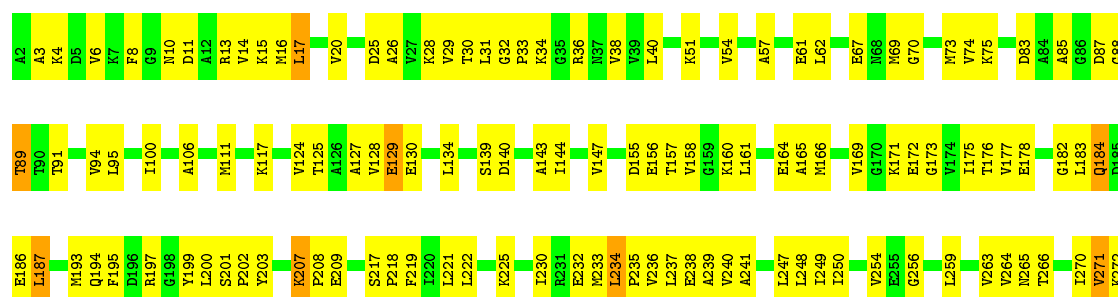
#### • Molecule 1: 60 KDA CHAPERONIN

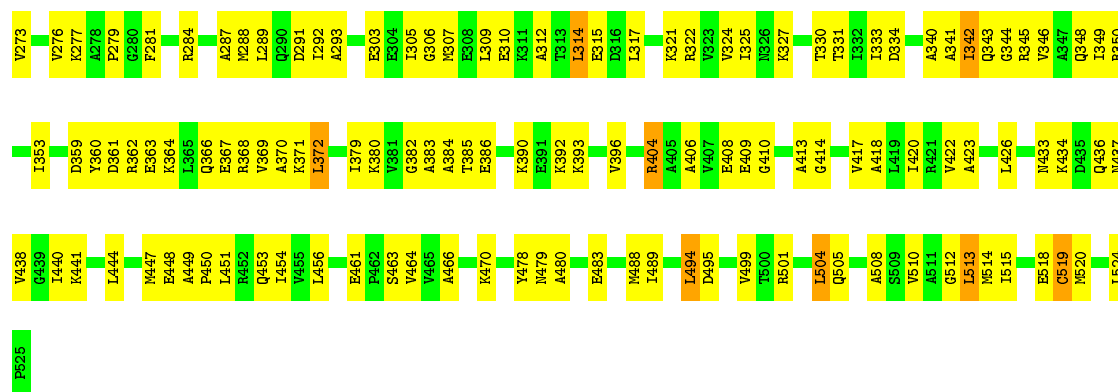
Chain A: 



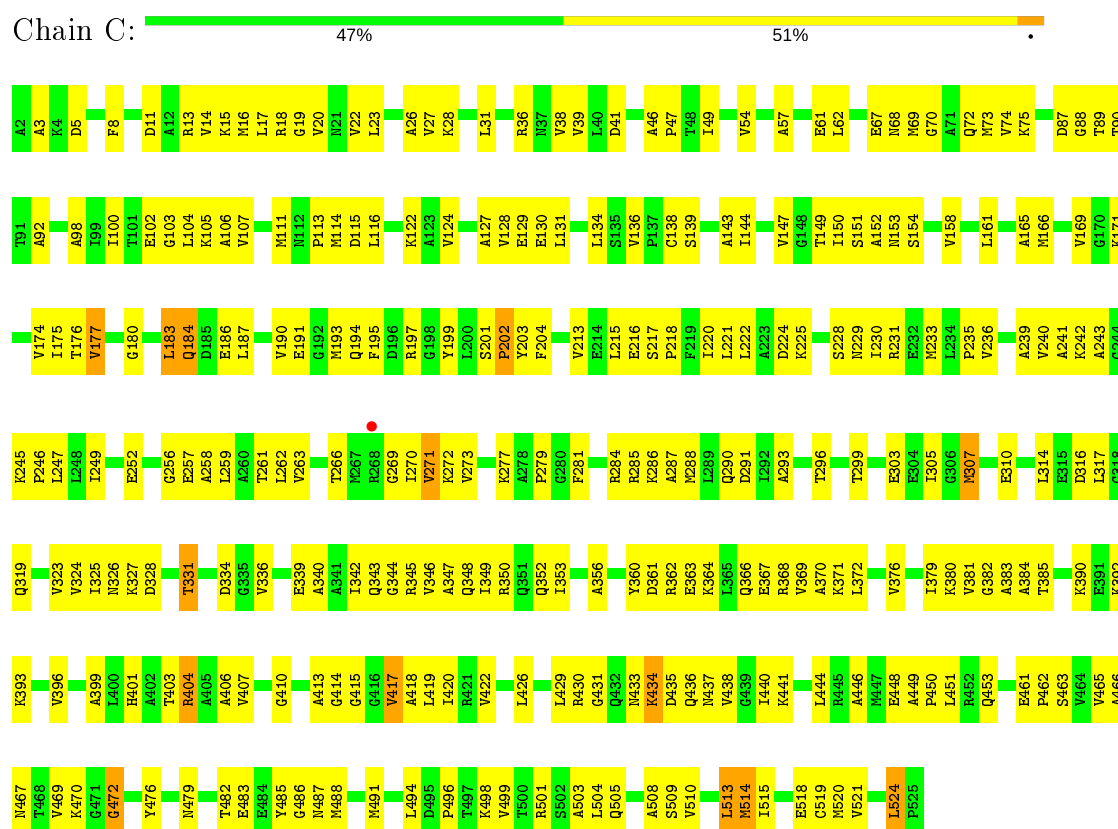
#### • Molecule 1: 60 KDA CHAPERONIN

Chain B: 

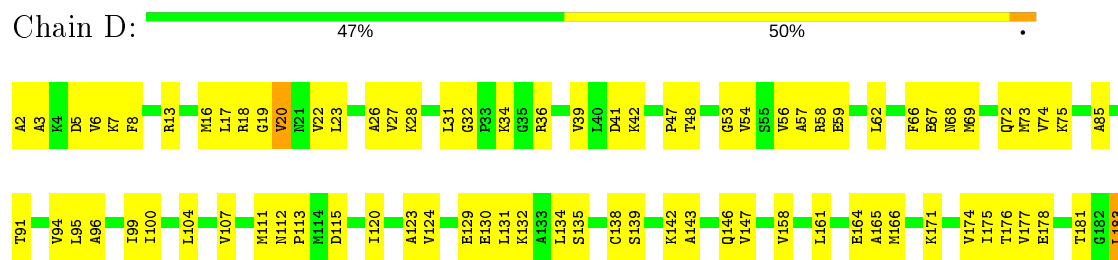


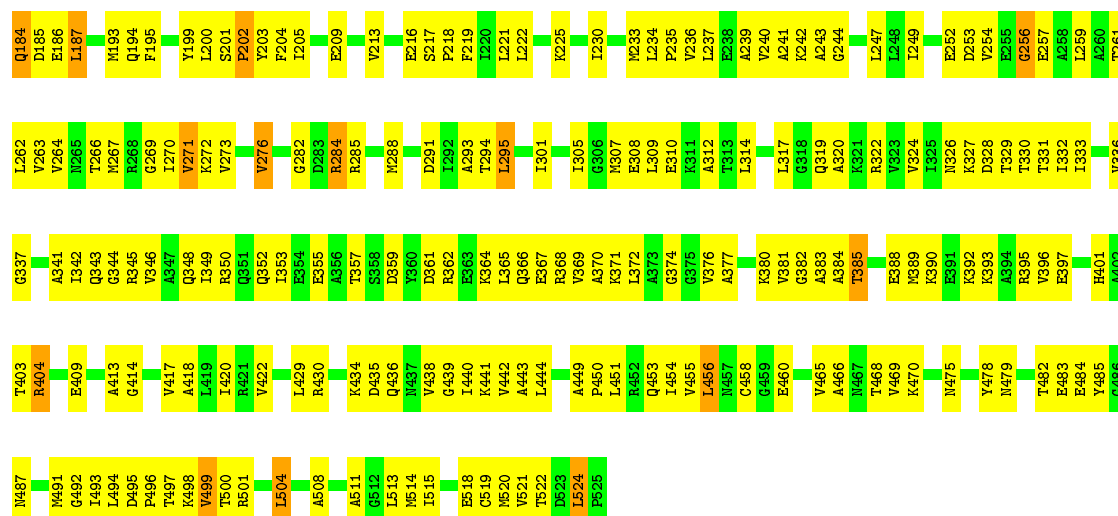


### • Molecule 1: 60 KDA CHAPERONIN

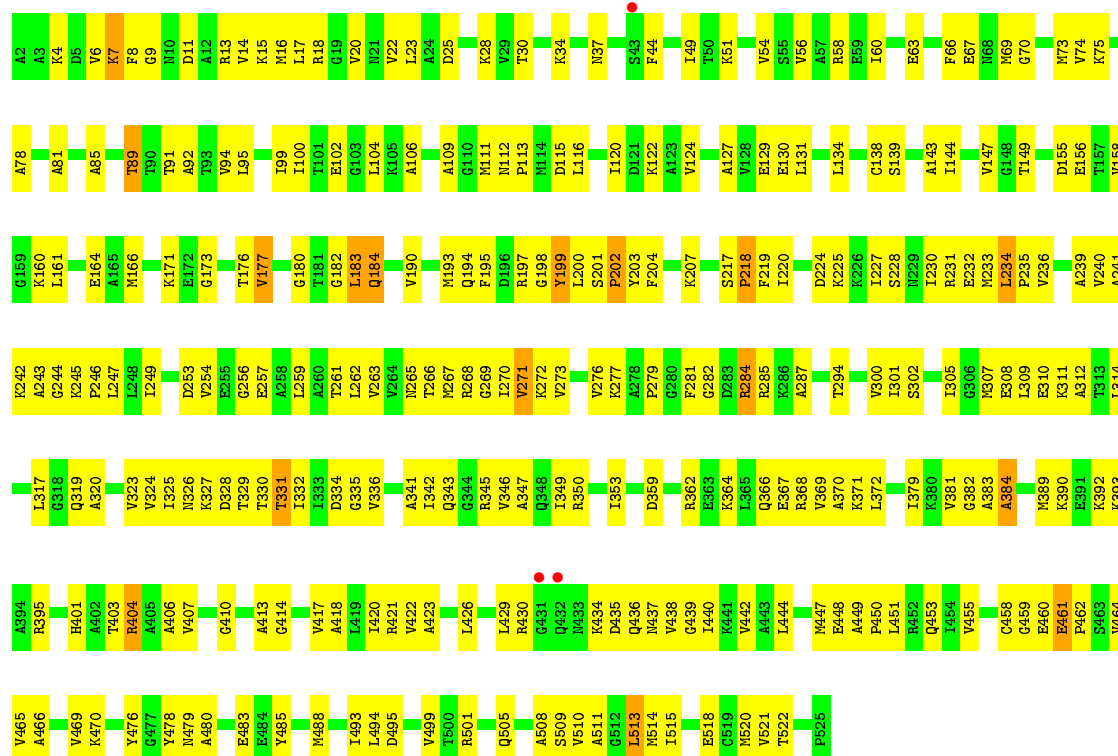


### • Molecule 1: 60 KDA CHAPERONIN





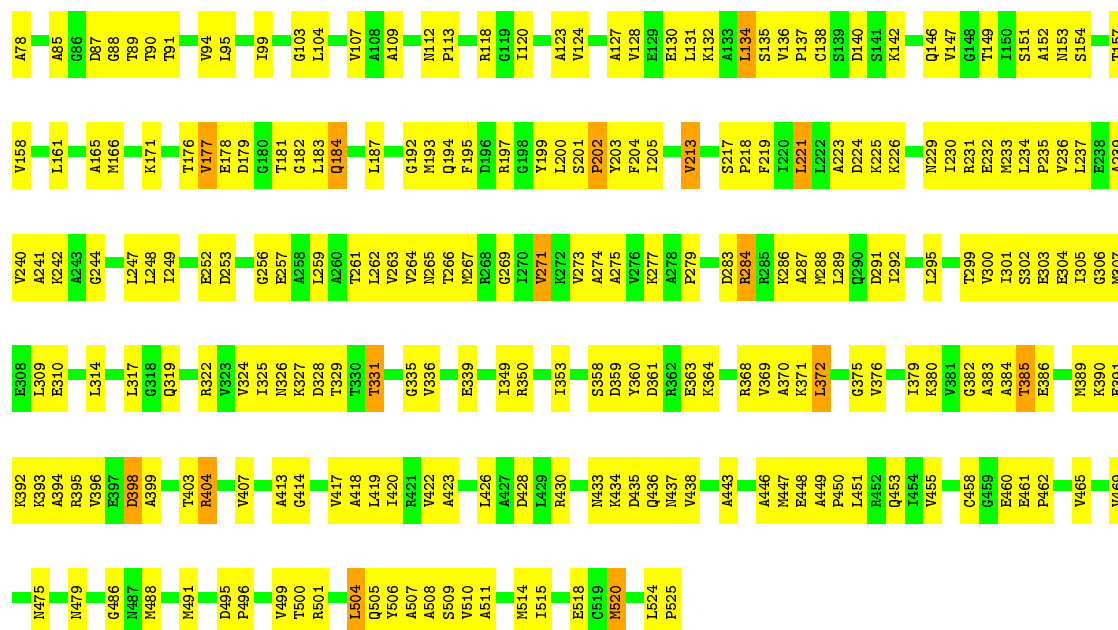
### • Molecule 1: 60 KDA CHAPERONIN



### • Molecule 1: 60 KDA CHAPERONIN

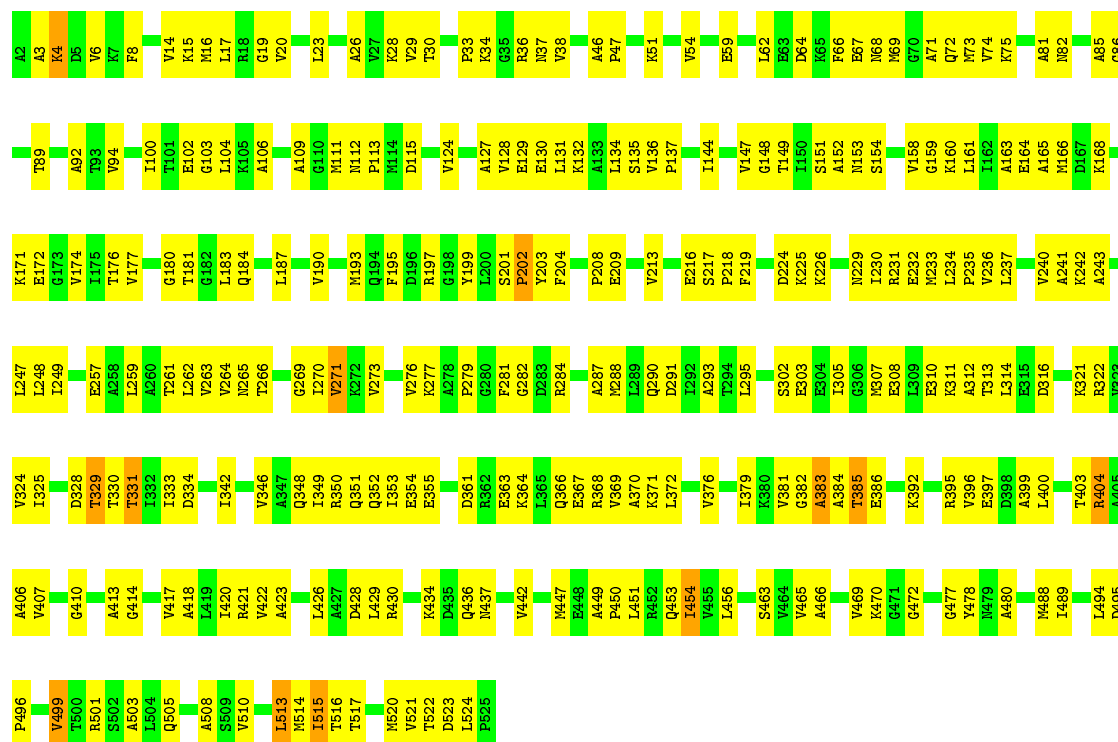






• Molecule 1: 60 KDA CHAPERONIN

Chain G: 50% 48%

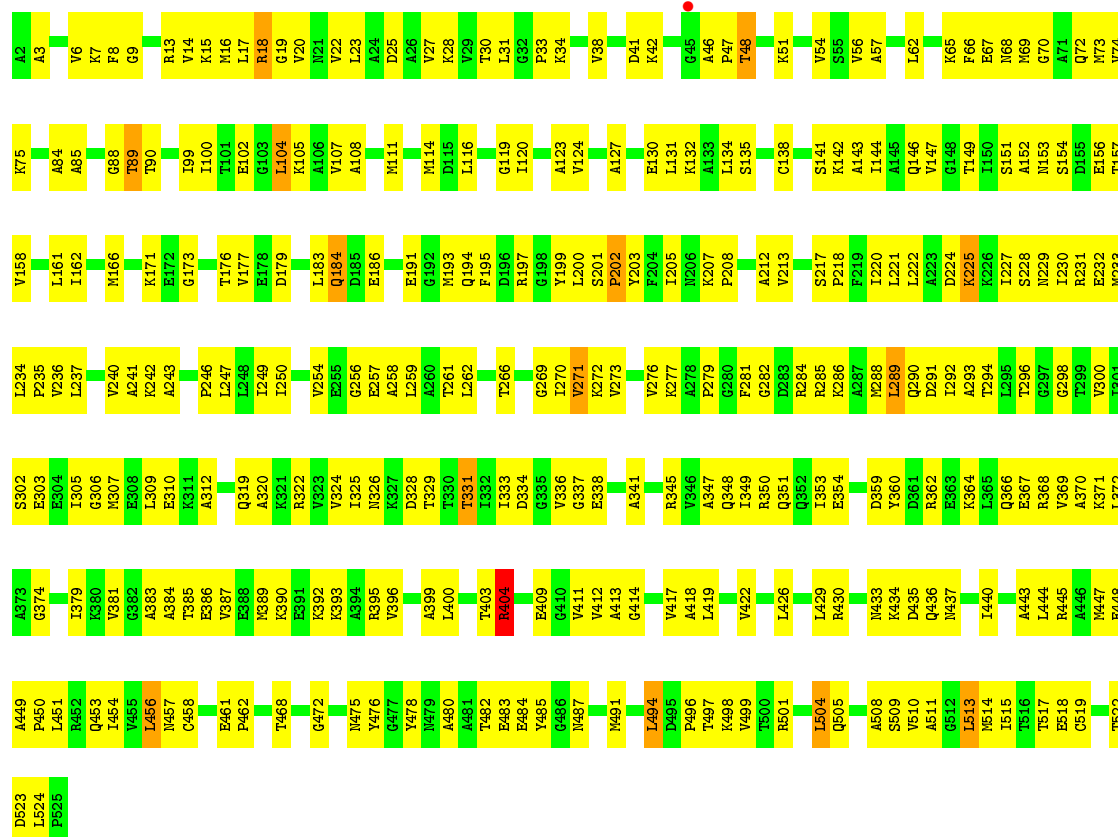


• Molecule 1: 60 KDA CHAPERONIN

Chain H: 46% 51%

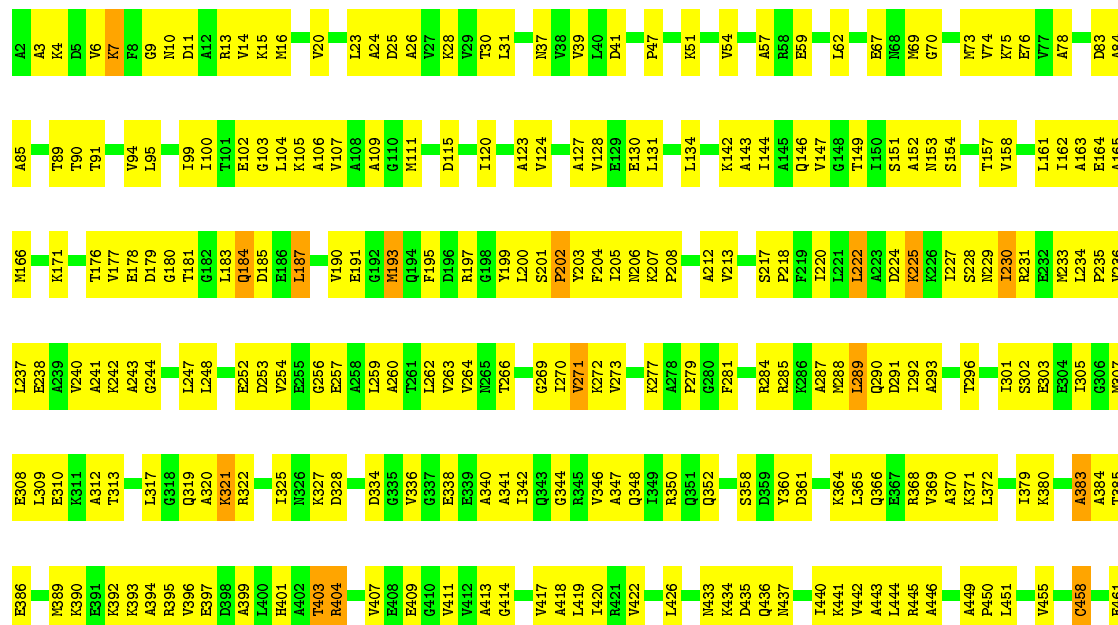


Chain J:  44% 53%



• Molecule 1: 60 KDA CHAPERONIN

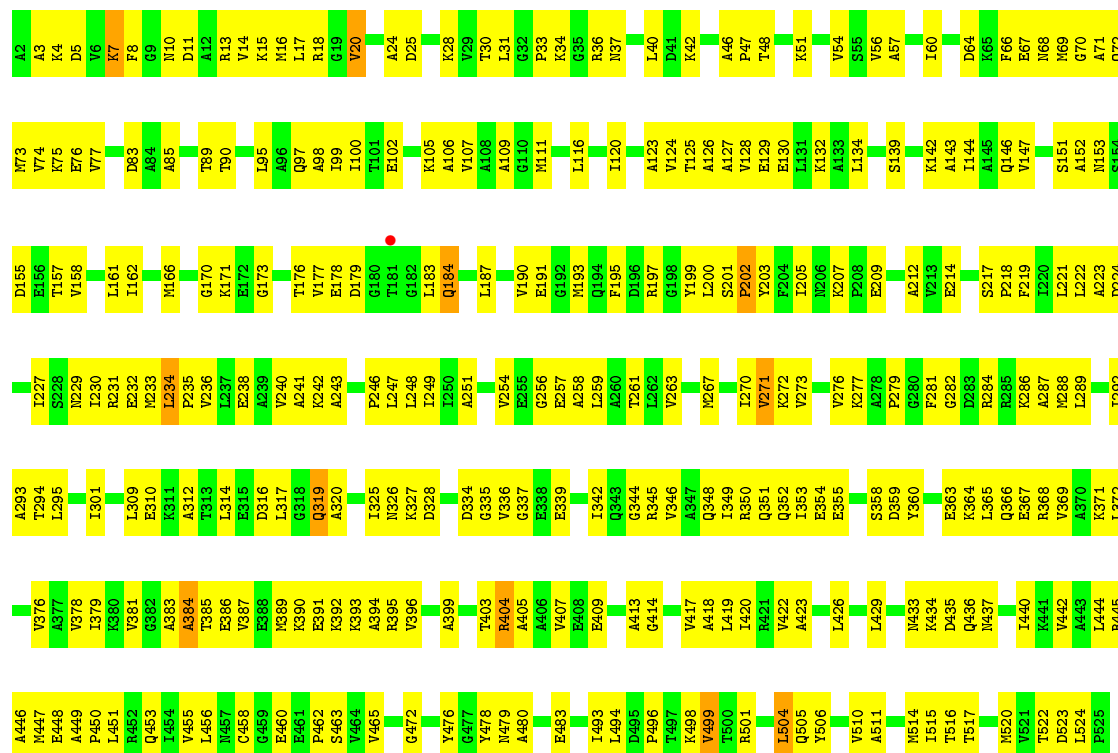
Chain K:  47% 50%





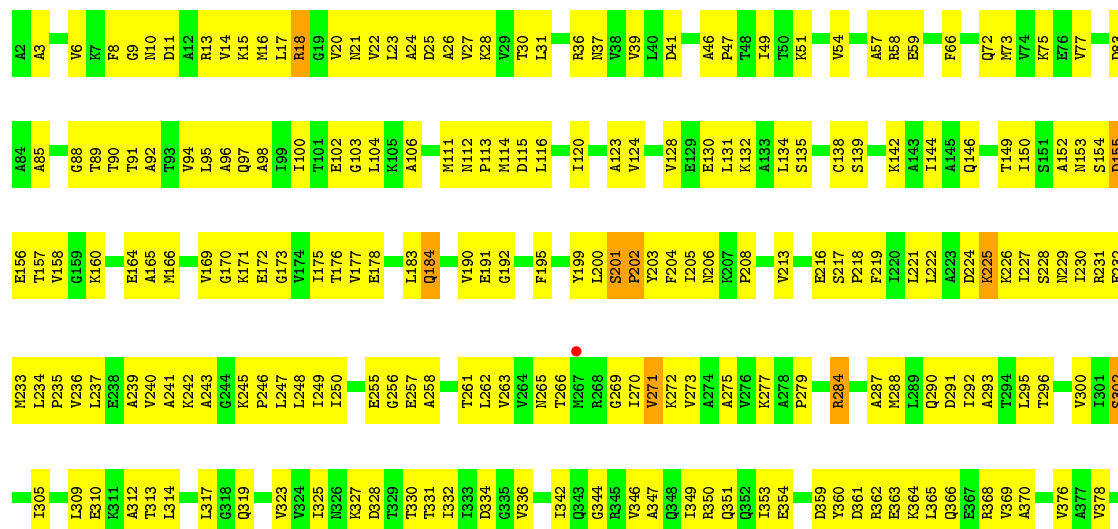
• Molecule 1: 60 KDA CHAPERONIN

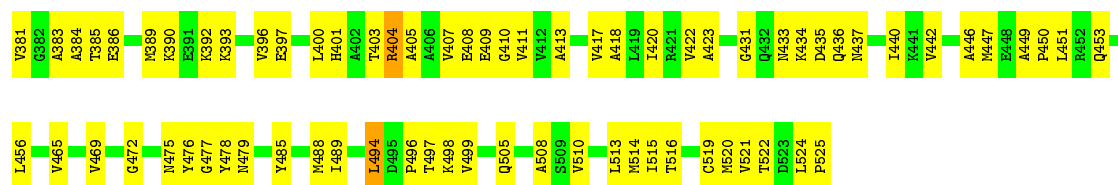
Chain L: 45% 53%



• Molecule 1: 60 KDA CHAPERONIN

Chain M: 46% 52%





• Molecule 1: 60 KDA CHAPERONIN



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	171.91Å 171.91Å 454.59Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.99 – 4.50 19.99 – 4.50	Depositor EDS
% Data completeness (in resolution range)	96.4 (19.99-4.50) 99.4 (19.99-4.50)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.23 (at 4.54Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, $R_{free}$	0.167 , 0.240 0.181 , 0.237	Depositor DCC
$R_{free}$ test set	2006 reflections (2.29%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	157.4	Xtriage
Anisotropy	0.151	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.25 , 96.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.36$ , $\langle L^2 \rangle = 0.18$	Xtriage
Estimated twinning fraction	0.118 for -h,-k,l 0.369 for h,-h-k,-l 0.118 for -k,-h,-l	Xtriage
Reported twinning fraction	0.500 for h,-h-k,-l	Depositor
Outliers	0 of 87570 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	53984	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	174.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.19% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup> Intensities estimated from amplitudes.

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

## 5 Model quality

### 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	0.79	0/3884	1.11	6/5242 (0.1%)
1	B	0.84	2/3884 (0.1%)	1.09	9/5242 (0.2%)
1	C	0.77	0/3884	1.08	6/5242 (0.1%)
1	D	0.76	1/3884 (0.0%)	1.05	7/5242 (0.1%)
1	E	0.77	1/3884 (0.0%)	1.04	6/5242 (0.1%)
1	F	0.93	4/3884 (0.1%)	1.18	11/5242 (0.2%)
1	G	0.81	1/3884 (0.0%)	1.08	4/5242 (0.1%)
1	H	0.87	0/3884	1.13	7/5242 (0.1%)
1	I	0.85	2/3884 (0.1%)	1.13	10/5242 (0.2%)
1	J	0.84	0/3884	1.10	8/5242 (0.2%)
1	K	0.82	3/3884 (0.1%)	1.10	10/5242 (0.2%)
1	L	0.81	0/3884	1.07	5/5242 (0.1%)
1	M	0.78	0/3884	1.05	2/5242 (0.0%)
1	N	0.76	0/3884	1.04	5/5242 (0.1%)
All	All	0.82	14/54376 (0.0%)	1.09	96/73388 (0.1%)

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	A	0	1
1	D	0	1
1	H	0	2
1	I	0	1
1	J	0	1
1	M	0	1
All	All	0	7

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	I	90	THR	CA-CB	6.74	1.70	1.53
1	F	458	CYS	CB-SG	-6.39	1.71	1.82
1	B	519	CYS	CB-SG	-6.07	1.72	1.82
1	F	339	GLU	CG-CD	-5.82	1.43	1.51
1	F	77	VAL	CA-CB	-5.65	1.42	1.54

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	I	504	LEU	CB-CG-CD1	-13.84	87.47	111.00
1	I	221	LEU	CA-CB-CG	-9.57	93.29	115.30
1	C	513	LEU	CA-CB-CG	-9.52	93.39	115.30
1	K	371	LYS	CD-CE-NZ	8.82	131.99	111.70
1	F	504	LEU	CB-CG-CD1	-8.67	96.26	111.00

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	432	GLN	Mainchain
1	D	256	GLY	Mainchain
1	H	163	ALA	Mainchain
1	H	385	THR	Peptide
1	I	359	ASP	Mainchain

## 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	3856	0	3983	311	0
1	B	3856	0	3983	272	0
1	C	3856	0	3983	305	0
1	D	3856	0	3983	310	0
1	E	3856	0	3983	313	0
1	F	3856	0	3983	307	1
1	G	3856	0	3983	297	0
1	H	3856	0	3983	321	1
1	I	3856	0	3983	317	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	J	3856	0	3983	336	0
1	K	3856	0	3983	330	0
1	L	3856	0	3983	327	0
1	M	3856	0	3983	306	0
1	N	3856	0	3983	297	0
All	All	53984	0	55762	4071	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 37.

The worst 5 of 4071 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:H:511:ALA:O	1:H:515:ILE:HD12	1.38	1.21
1:B:85:ALA:HB1	1:B:499:VAL:HG12	1.22	1.19
1:E:319:GLN:O	1:E:336:VAL:HG23	1.43	1.19
1:F:131:LEU:CD1	1:F:422:VAL:HG21	1.73	1.18
1:F:57:ALA:O	1:F:75:LYS:HE3	1.43	1.15

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:358:SER:CB	1:H:167:ASP:OD1[2_554]	2.19	0.01

## 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	522/524 (100%)	497 (95%)	21 (4%)	4 (1%)	19	60
1	B	522/524 (100%)	495 (95%)	23 (4%)	4 (1%)	19	60

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	522/524 (100%)	500 (96%)	18 (3%)	4 (1%)	19	60
1	D	522/524 (100%)	498 (95%)	19 (4%)	5 (1%)	15	54
1	E	522/524 (100%)	496 (95%)	21 (4%)	5 (1%)	15	54
1	F	522/524 (100%)	496 (95%)	20 (4%)	6 (1%)	14	52
1	G	522/524 (100%)	499 (96%)	18 (3%)	5 (1%)	15	54
1	H	522/524 (100%)	489 (94%)	26 (5%)	7 (1%)	12	48
1	I	522/524 (100%)	496 (95%)	18 (3%)	8 (2%)	10	46
1	J	522/524 (100%)	496 (95%)	21 (4%)	5 (1%)	15	54
1	K	522/524 (100%)	494 (95%)	20 (4%)	8 (2%)	10	46
1	L	522/524 (100%)	495 (95%)	22 (4%)	5 (1%)	15	54
1	M	522/524 (100%)	490 (94%)	25 (5%)	7 (1%)	12	48
1	N	522/524 (100%)	495 (95%)	21 (4%)	6 (1%)	14	52
All	All	7308/7336 (100%)	6936 (95%)	293 (4%)	79 (1%)	14	52

5 of 79 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	184	GLN
1	B	271	VAL
1	C	184	GLN
1	C	271	VAL
1	D	184	GLN

### 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	404/404 (100%)	399 (99%)	5 (1%)	71	84
1	B	404/404 (100%)	396 (98%)	8 (2%)	55	73
1	C	404/404 (100%)	391 (97%)	13 (3%)	39	62
1	D	404/404 (100%)	392 (97%)	12 (3%)	41	63

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	E	404/404 (100%)	392 (97%)	12 (3%)	41	63
1	F	404/404 (100%)	394 (98%)	10 (2%)	47	68
1	G	404/404 (100%)	392 (97%)	12 (3%)	41	63
1	H	404/404 (100%)	398 (98%)	6 (2%)	65	80
1	I	404/404 (100%)	398 (98%)	6 (2%)	65	80
1	J	404/404 (100%)	398 (98%)	6 (2%)	65	80
1	K	404/404 (100%)	397 (98%)	7 (2%)	60	78
1	L	404/404 (100%)	397 (98%)	7 (2%)	60	78
1	M	404/404 (100%)	398 (98%)	6 (2%)	65	80
1	N	404/404 (100%)	394 (98%)	10 (2%)	47	68
All	All	5656/5656 (100%)	5536 (98%)	120 (2%)	53	72

5 of 120 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	331	THR
1	G	361	ASP
1	N	20	VAL
1	F	361	ASP
1	G	94	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 37 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	475	ASN
1	H	21	ASN
1	M	401	HIS
1	G	229	ASN
1	G	326	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

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

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	524/524 (100%)	-0.47	1 (0%) 95 93	88, 174, 251, 296	0
1	B	524/524 (100%)	-0.52	0 100 100	96, 172, 231, 280	0
1	C	524/524 (100%)	-0.50	1 (0%) 95 93	84, 177, 238, 293	0
1	D	524/524 (100%)	-0.48	0 100 100	89, 180, 263, 409	0
1	E	524/524 (100%)	-0.47	3 (0%) 89 84	78, 191, 247, 350	0
1	F	524/524 (100%)	-0.63	0 100 100	54, 140, 193, 240	0
1	G	524/524 (100%)	-0.54	0 100 100	85, 165, 231, 320	0
1	H	524/524 (100%)	-0.59	0 100 100	64, 154, 237, 308	0
1	I	524/524 (100%)	-0.52	1 (0%) 95 93	68, 174, 241, 308	0
1	J	524/524 (100%)	-0.56	1 (0%) 95 93	82, 155, 227, 300	0
1	K	524/524 (100%)	-0.57	0 100 100	75, 162, 228, 263	0
1	L	524/524 (100%)	-0.51	1 (0%) 95 93	70, 184, 260, 317	0
1	M	524/524 (100%)	-0.49	1 (0%) 95 93	92, 181, 259, 500	0
1	N	524/524 (100%)	-0.38	7 (1%) 77 68	92, 190, 273, 428	0
All	All	7336/7336 (100%)	-0.52	16 (0%) 95 93	54, 171, 246, 500	0

The worst 5 of 16 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	L	181	THR	4.4
1	N	44	PHE	3.9
1	A	181	THR	3.7
1	C	268	ARG	3.4
1	E	431	GLY	3.3

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.