



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 08:49 am BST

PDB ID : 4WGL  
Title : Crystal structure of a GroEL D83A/R197A double mutant  
Authors : Yang, D.; Fei, X.; LaRonde, N.A.; Beckett, D.; Lund, P.A.; Lorimer, G.H.  
Deposited on : 2014-09-19  
Resolution : 3.13 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

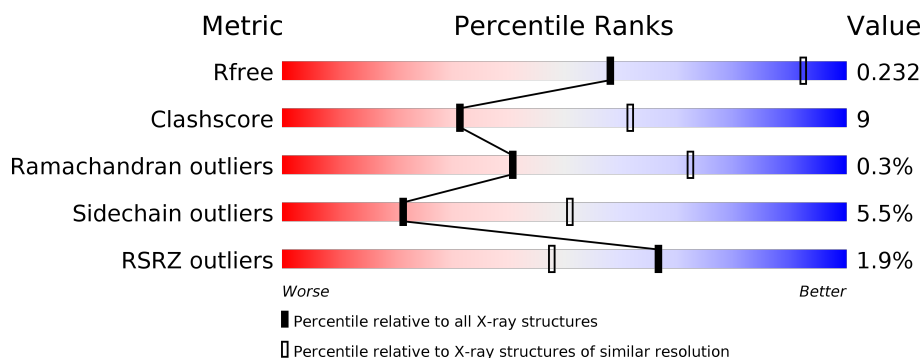
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.13 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	1626 (3.18-3.10)
Clashscore	141614	1735 (3.18-3.10)
Ramachandran outliers	138981	1677 (3.18-3.10)
Sidechain outliers	138945	1677 (3.18-3.10)
RSRZ outliers	127900	1588 (3.18-3.10)

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	548	<div> <div></div> <div>76%18% . .</div> </div>
1	B	548	<div> <div>6%</div> <div>70%24% . .</div> </div>
1	C	548	<div> <div>%</div> <div>74%19% . .</div> </div>
1	D	548	<div> <div>2%</div> <div>72%22% . .</div> </div>
1	E	548	<div> <div></div> <div>75%18% . .</div> </div>
1	F	548	<div> <div>%</div> <div>67%26% . .</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	548	<div><div>4%</div><div><div></div><div>74%</div><div>21%</div><div></div></div><div><div></div><div></div></div></div>
1	H	548	<div><div></div><div><div></div><div>74%</div><div>20%</div><div></div></div><div><div></div><div></div></div></div>
1	I	548	<div><div></div><div><div></div><div>70%</div><div>24%</div><div></div></div><div><div></div><div></div></div></div>
1	J	548	<div><div></div><div><div></div><div>70%</div><div>24%</div><div></div></div><div><div></div><div></div></div></div>
1	K	548	<div><div>9%</div><div><div></div><div>71%</div><div>23%</div><div></div></div><div><div></div><div></div></div></div>
1	L	548	<div><div></div><div><div></div><div>72%</div><div>22%</div><div></div></div><div><div></div><div></div></div></div>
1	M	548	<div><div>%</div><div><div></div><div>72%</div><div>22%</div><div></div></div><div><div></div><div></div></div></div>
1	N	548	<div><div>%</div><div><div></div><div>77%</div><div>16%</div><div></div></div><div><div></div><div></div></div></div>

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 53844 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
			3846	2393	662	771	20			
1	B	524	Total	C	N	O	S	0	0	0
			3846	2393	662	771	20			
1	C	524	Total	C	N	O	S	0	0	0
			3846	2393	662	771	20			
1	D	524	Total	C	N	O	S	0	0	0
			3846	2393	662	771	20			
1	E	524	Total	C	N	O	S	0	0	0
			3846	2393	662	771	20			
1	F	524	Total	C	N	O	S	0	0	0
			3846	2393	662	771	20			
1	G	524	Total	C	N	O	S	0	0	0
			3846	2393	662	771	20			
1	H	524	Total	C	N	O	S	0	0	0
			3846	2393	662	771	20			
1	I	524	Total	C	N	O	S	0	0	0
			3846	2393	662	771	20			
1	J	524	Total	C	N	O	S	0	0	0
			3846	2393	662	771	20			
1	K	524	Total	C	N	O	S	0	0	0
			3846	2393	662	771	20			
1	L	524	Total	C	N	O	S	0	0	0
			3846	2393	662	771	20			
1	M	524	Total	C	N	O	S	0	0	0
			3846	2393	662	771	20			
1	N	524	Total	C	N	O	S	0	0	0
			3846	2393	662	771	20			

There are 28 discrepancies between the modelled and reference sequences:

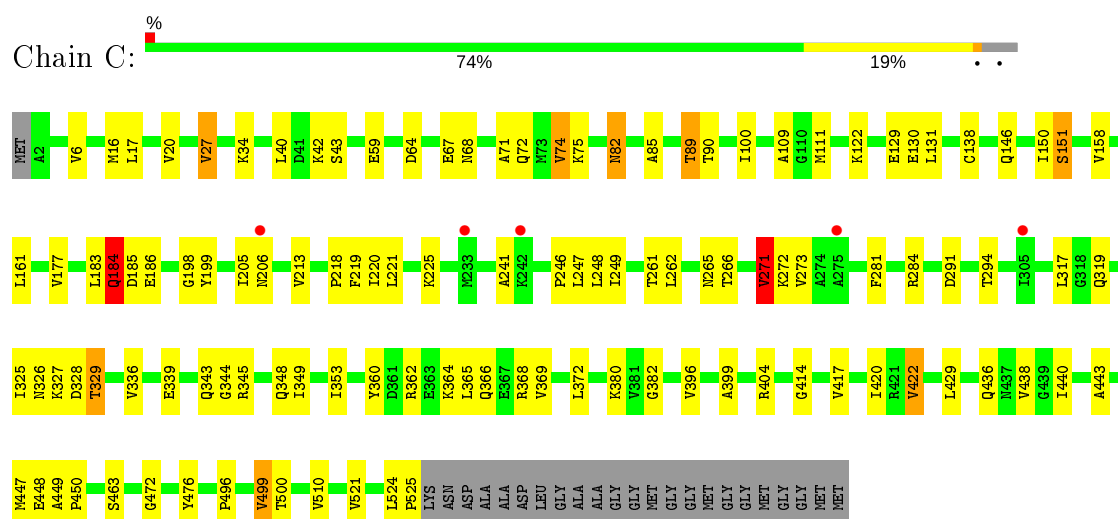
Chain	Residue	Modelled	Actual	Comment	Reference
A	83	ALA	ASP	engineered mutation	UNP P0A6F5

*Continued on next page...*

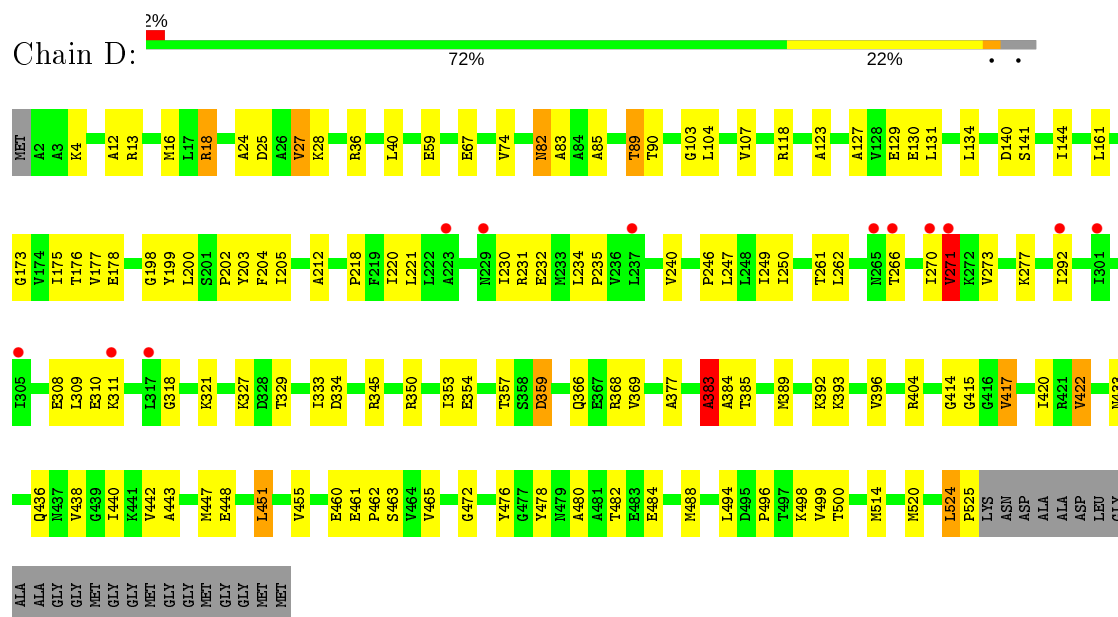
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	197	ALA	ARG	engineered mutation	UNP P0A6F5
B	83	ALA	ASP	engineered mutation	UNP P0A6F5
B	197	ALA	ARG	engineered mutation	UNP P0A6F5
C	83	ALA	ASP	engineered mutation	UNP P0A6F5
C	197	ALA	ARG	engineered mutation	UNP P0A6F5
D	83	ALA	ASP	engineered mutation	UNP P0A6F5
D	197	ALA	ARG	engineered mutation	UNP P0A6F5
E	83	ALA	ASP	engineered mutation	UNP P0A6F5
E	197	ALA	ARG	engineered mutation	UNP P0A6F5
F	83	ALA	ASP	engineered mutation	UNP P0A6F5
F	197	ALA	ARG	engineered mutation	UNP P0A6F5
G	83	ALA	ASP	engineered mutation	UNP P0A6F5
G	197	ALA	ARG	engineered mutation	UNP P0A6F5
H	83	ALA	ASP	engineered mutation	UNP P0A6F5
H	197	ALA	ARG	engineered mutation	UNP P0A6F5
I	83	ALA	ASP	engineered mutation	UNP P0A6F5
I	197	ALA	ARG	engineered mutation	UNP P0A6F5
J	83	ALA	ASP	engineered mutation	UNP P0A6F5
J	197	ALA	ARG	engineered mutation	UNP P0A6F5
K	83	ALA	ASP	engineered mutation	UNP P0A6F5
K	197	ALA	ARG	engineered mutation	UNP P0A6F5
L	83	ALA	ASP	engineered mutation	UNP P0A6F5
L	197	ALA	ARG	engineered mutation	UNP P0A6F5
M	83	ALA	ASP	engineered mutation	UNP P0A6F5
M	197	ALA	ARG	engineered mutation	UNP P0A6F5
N	83	ALA	ASP	engineered mutation	UNP P0A6F5
N	197	ALA	ARG	engineered mutation	UNP P0A6F5

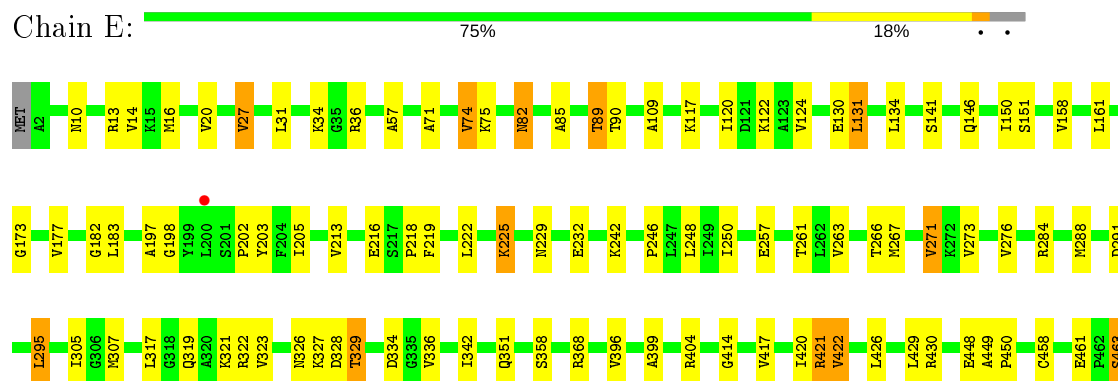




- Molecule 1: 60 kDa chaperonin

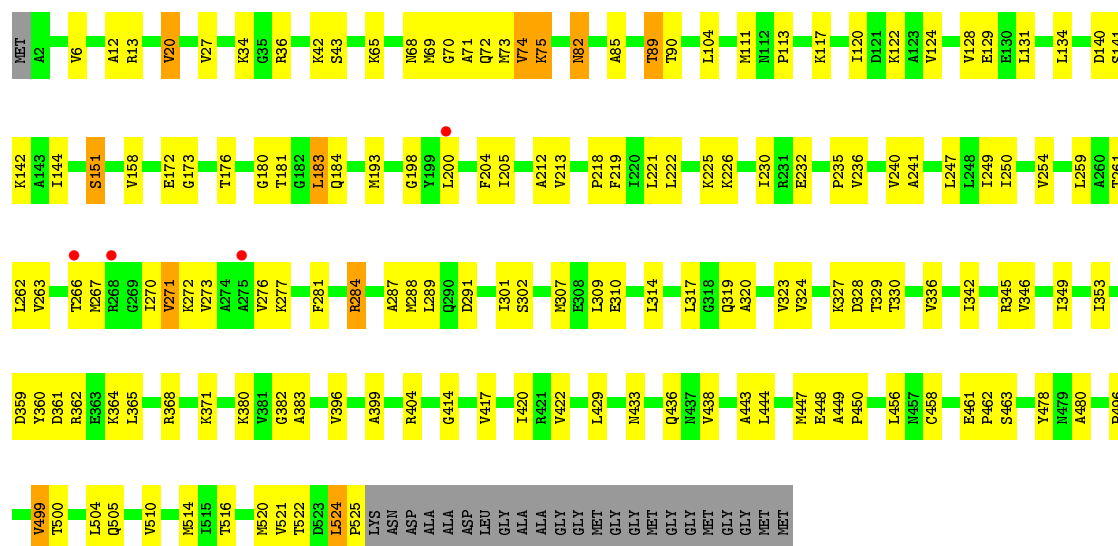


- Molecule 1: 60 kDa chaperonin

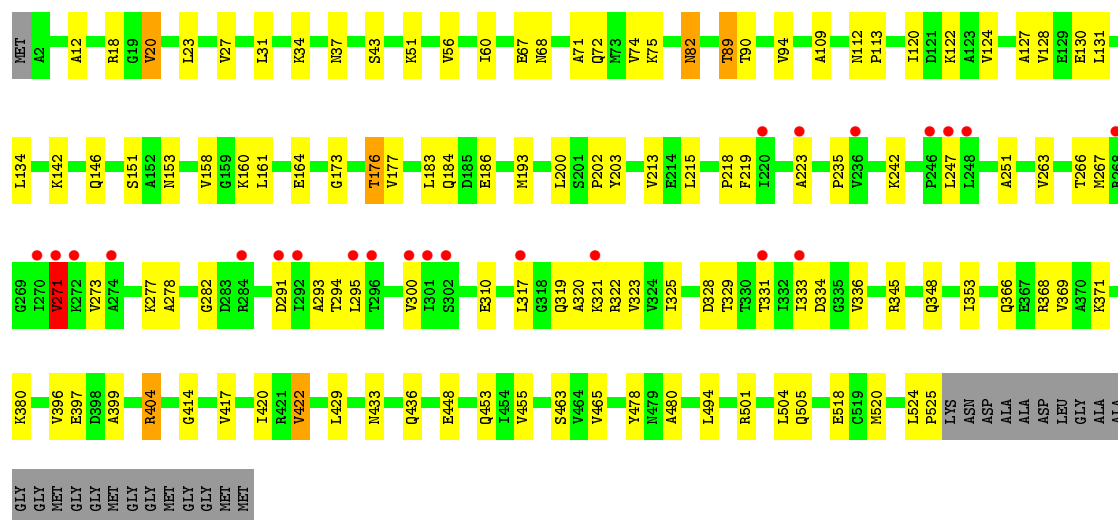
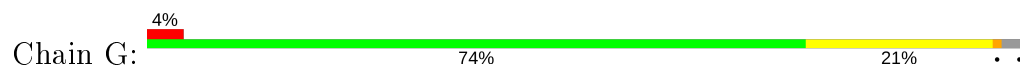




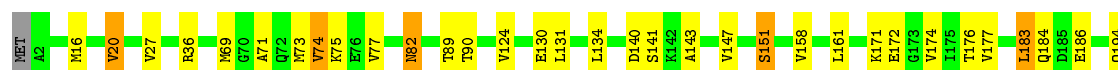
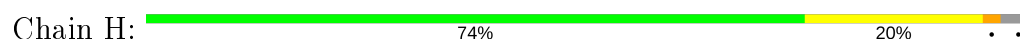
- Molecule 1: 60 kDa chaperonin



- Molecule 1: 60 kDa chaperonin



- Molecule 1: 60 kDa chaperonin





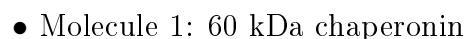
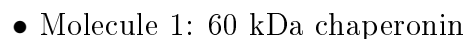
- Molecule 1: 60 kDa chaperonin

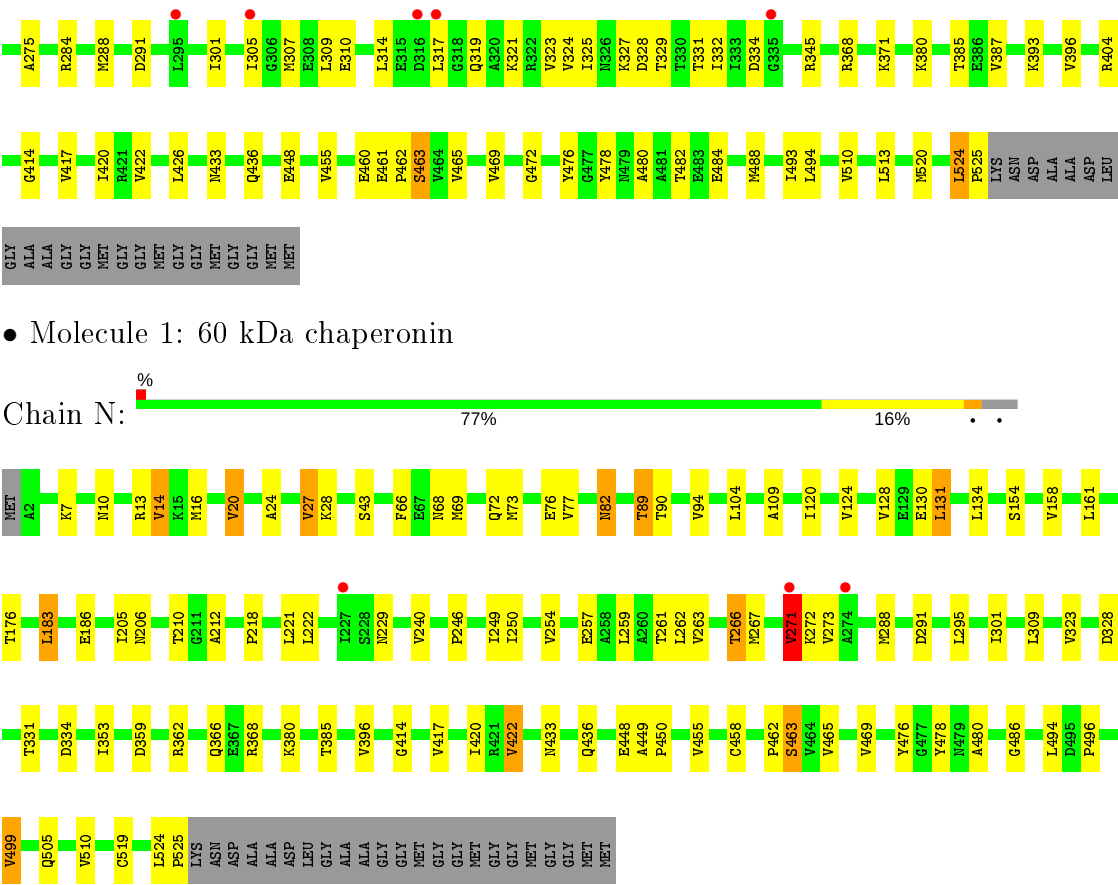
Chain I:  70% 24% . .

- Molecule 1: 60 kDa chaperonin

Chain J:  70% 24% . .

- Molecule 1: 60 kDa chaperonin





• Molecule 1: 60 kDa chaperonin

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	135.62Å 259.71Å 280.85Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	123.52 – 3.13 123.52 – 3.13	Depositor EDS
% Data completeness (in resolution range)	99.6 (123.52-3.13) 99.6 (123.52-3.13)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	0.15	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.65 (at 3.13Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.4_1496)	Depositor
R, $R_{free}$	0.166 , 0.233 0.168 , 0.232	Depositor DCC
$R_{free}$ test set	2003 reflections (1.15%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	69.5	Xtriage
Anisotropy	0.402	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 70.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	53844	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	80.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.62% 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	0.43	0/3874	0.61	0/5232
1	B	0.43	0/3874	0.62	0/5232
1	C	0.42	0/3874	0.60	0/5232
1	D	0.42	0/3874	0.59	1/5232 (0.0%)
1	E	0.42	0/3874	0.61	0/5232
1	F	0.41	0/3874	0.59	0/5232
1	G	0.41	0/3874	0.58	0/5232
1	H	0.45	0/3874	0.62	0/5232
1	I	0.46	0/3874	0.66	1/5232 (0.0%)
1	J	0.46	0/3874	0.63	1/5232 (0.0%)
1	K	0.43	0/3874	0.61	1/5232 (0.0%)
1	L	0.47	0/3874	0.64	0/5232
1	M	0.41	0/3874	0.59	1/5232 (0.0%)
1	N	0.43	0/3874	0.61	0/5232
All	All	0.43	0/54236	0.61	5/73248 (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
1	D	0	1
1	J	0	1
All	All	0	2

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	I	295	LEU	CA-CB-CG	6.30	129.78	115.30
1	M	183	LEU	CA-CB-CG	5.83	128.70	115.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	J	131	LEU	CA-CB-CG	5.14	127.11	115.30
1	K	182	GLY	N-CA-C	5.13	125.94	113.10
1	D	383	ALA	N-CA-C	5.11	124.78	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	383	ALA	Peptide
1	J	31	LEU	Peptide

## 5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3846	0	3969	59	0
1	B	3846	0	3969	83	0
1	C	3846	0	3969	73	0
1	D	3846	0	3969	78	0
1	E	3846	0	3969	64	0
1	F	3846	0	3969	89	0
1	G	3846	0	3969	68	0
1	H	3846	0	3969	71	0
1	I	3846	0	3969	88	0
1	J	3846	0	3969	79	0
1	K	3846	0	3969	87	0
1	L	3846	0	3969	72	0
1	M	3846	0	3969	70	0
1	N	3846	0	3969	54	0
All	All	53844	0	55566	986	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 986 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:K:232:GLU:HB3	1:K:309:LEU:HD11	1.46	0.95
1:F:241:ALA:HA	1:F:271:VAL:HG11	1.51	0.92
1:L:130:GLU:HB3	1:L:422:VAL:HG22	1.51	0.91
1:E:173:GLY:O	1:E:404:ARG:NH2	2.08	0.86
1:E:524:LEU:HD12	1:E:525:PRO:HD2	1.56	0.85

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	522/548 (95%)	512 (98%)	8 (2%)	2 (0%)	34	67
1	B	522/548 (95%)	510 (98%)	10 (2%)	2 (0%)	34	67
1	C	522/548 (95%)	510 (98%)	10 (2%)	2 (0%)	34	67
1	D	522/548 (95%)	511 (98%)	10 (2%)	1 (0%)	47	78
1	E	522/548 (95%)	510 (98%)	10 (2%)	2 (0%)	34	67
1	F	522/548 (95%)	511 (98%)	10 (2%)	1 (0%)	47	78
1	G	522/548 (95%)	512 (98%)	9 (2%)	1 (0%)	47	78
1	H	522/548 (95%)	514 (98%)	8 (2%)	0	100	100
1	I	522/548 (95%)	507 (97%)	13 (2%)	2 (0%)	34	67
1	J	522/548 (95%)	514 (98%)	7 (1%)	1 (0%)	47	78
1	K	522/548 (95%)	506 (97%)	14 (3%)	2 (0%)	34	67
1	L	522/548 (95%)	511 (98%)	10 (2%)	1 (0%)	47	78
1	M	522/548 (95%)	507 (97%)	14 (3%)	1 (0%)	47	78
1	N	522/548 (95%)	513 (98%)	8 (2%)	1 (0%)	47	78
All	All	7308/7672 (95%)	7148 (98%)	141 (2%)	19 (0%)	41	72

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	271	VAL
1	J	271	VAL
1	M	271	VAL
1	N	271	VAL
1	A	183	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	402/413 (97%)	379 (94%)	23 (6%)	20	49
1	B	402/413 (97%)	382 (95%)	20 (5%)	24	55
1	C	402/413 (97%)	382 (95%)	20 (5%)	24	55
1	D	402/413 (97%)	380 (94%)	22 (6%)	21	51
1	E	402/413 (97%)	381 (95%)	21 (5%)	23	53
1	F	402/413 (97%)	380 (94%)	22 (6%)	21	51
1	G	402/413 (97%)	385 (96%)	17 (4%)	30	60
1	H	402/413 (97%)	383 (95%)	19 (5%)	26	57
1	I	402/413 (97%)	376 (94%)	26 (6%)	17	45
1	J	402/413 (97%)	377 (94%)	25 (6%)	18	47
1	K	402/413 (97%)	383 (95%)	19 (5%)	26	57
1	L	402/413 (97%)	376 (94%)	26 (6%)	17	45
1	M	402/413 (97%)	379 (94%)	23 (6%)	20	49
1	N	402/413 (97%)	375 (93%)	27 (7%)	16	44
All	All	5628/5782 (97%)	5318 (94%)	310 (6%)	21	51

5 of 310 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	331	THR
1	I	138	CYS
1	N	43	SER

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	G	494	LEU
1	H	328	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	H	326	ASN
1	M	433	ASN
1	H	479	ASN
1	B	206	ASN
1	J	146	GLN

### 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 carbohydrates in this entry.

### 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

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

There are no such residues in this entry.

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

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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	524/548 (95%)	-0.18	2 (0%) 92 86	22, 60, 120, 174	0
1	B	524/548 (95%)	0.17	31 (5%) 22 10	24, 65, 207, 301	0
1	C	524/548 (95%)	-0.12	5 (0%) 82 70	25, 70, 166, 204	0
1	D	524/548 (95%)	-0.02	12 (2%) 60 40	28, 73, 161, 232	0
1	E	524/548 (95%)	-0.15	1 (0%) 95 91	28, 69, 136, 190	0
1	F	524/548 (95%)	-0.07	4 (0%) 86 74	31, 77, 135, 190	0
1	G	524/548 (95%)	0.09	23 (4%) 34 17	27, 80, 167, 208	0
1	H	524/548 (95%)	-0.25	0 100 100	25, 63, 118, 168	0
1	I	524/548 (95%)	-0.13	2 (0%) 92 86	21, 54, 146, 227	0
1	J	524/548 (95%)	-0.16	1 (0%) 95 91	18, 63, 134, 189	0
1	K	524/548 (95%)	0.29	48 (9%) 9 3	22, 66, 239, 292	0
1	L	524/548 (95%)	-0.23	1 (0%) 95 91	25, 58, 125, 177	0
1	M	524/548 (95%)	-0.04	7 (1%) 77 61	28, 77, 172, 233	0
1	N	524/548 (95%)	-0.12	3 (0%) 89 80	22, 64, 152, 207	0
All	All	7336/7672 (95%)	-0.07	140 (1%) 66 48	18, 66, 167, 301	0

The worst 5 of 140 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	K	223	ALA	12.0
1	B	251	ALA	8.8
1	K	301	ILE	7.6
1	K	203	TYR	7.2
1	K	251	ALA	7.2

## 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.