



wwPDB X-ray Structure Validation Summary Report ⓘ

May 9, 2016 – 12:00 AM EDT

PDB ID : 4YPL
Title : Crystal structure of a hexameric LonA protease bound to three ADPs
Authors : Lin, C.-C.; Chang, C.-I.
Deposited on : 2015-03-13
Resolution : 3.45 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
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20027457
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20027457

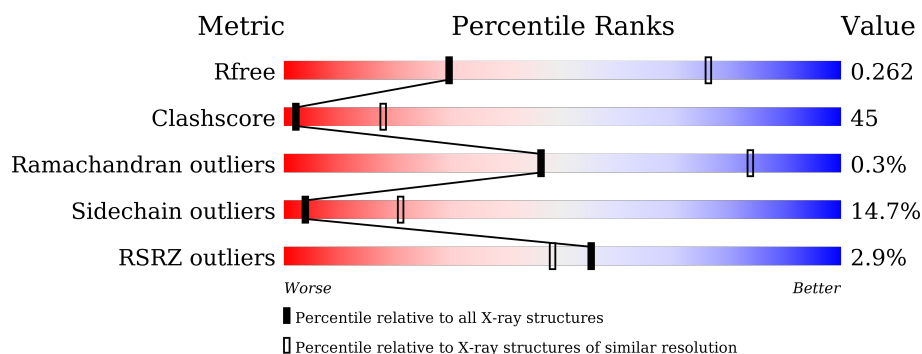
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.45 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}	91344	1000 (3.56-3.36)
Clashscore	102246	1090 (3.56-3.36)
Ramachandran outliers	100387	1057 (3.56-3.36)
Sidechain outliers	100360	1058 (3.56-3.36)
RSRZ outliers	91569	1005 (3.56-3.36)

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 ≥ 3 , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	555	<div> <div>0%</div> <div> <div>43%</div> <div>46%</div> <div>7%</div> </div> </div>
1	B	555	<div> <div>0%</div> <div> <div>48%</div> <div>42%</div> <div>7%</div> </div> </div>
1	C	555	<div> <div>5%</div> <div> <div>39%</div> <div>46%</div> <div>12%</div> </div> </div>
1	D	555	<div> <div>2%</div> <div> <div>43%</div> <div>47%</div> <div>7%</div> </div> </div>
1	E	555	<div> <div>5%</div> <div> <div>36%</div> <div>49%</div> <div>12%</div> </div> </div>
1	F	555	<div> <div>2%</div> <div> <div>46%</div> <div>43%</div> <div>7%</div> </div> </div>

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
2	4KZ	A	801	-	-	-	X
2	4KZ	C	801	-	-	-	X
2	4KZ	E	801	-	-	-	X

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 25358 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 Lon protease.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	537	Total	C	N	O	S	0	0	0
			4179	2643	732	788	16			
1	B	537	Total	C	N	O	S	0	0	0
			4179	2643	732	788	16			
1	C	538	Total	C	N	O	S	0	0	0
			4187	2649	733	789	16			
1	F	538	Total	C	N	O	S	0	0	0
			4188	2648	734	790	16			
1	D	537	Total	C	N	O	S	0	0	0
			4179	2643	732	788	16			
1	E	537	Total	C	N	O	S	0	0	0
			4179	2643	732	788	16			

There are 24 discrepancies between the modelled and reference sequences:

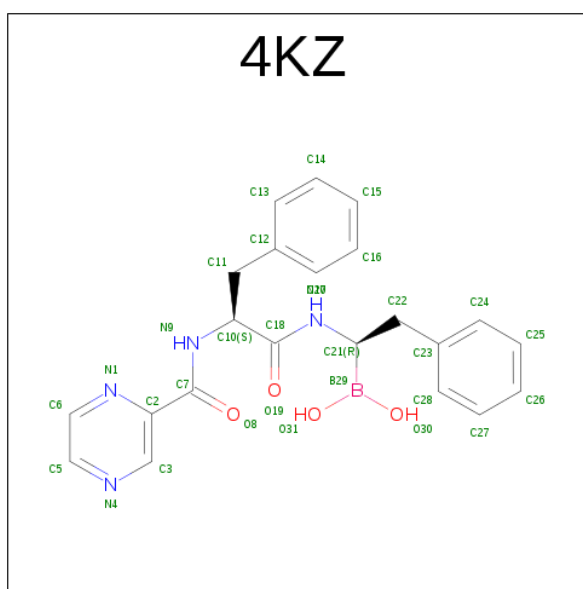
Chain	Residue	Modelled	Actual	Comment	Reference
A	239	GLY	-	expression tag	UNP A0A059VAZ3
A	240	HIS	-	expression tag	UNP A0A059VAZ3
A	241	MET	-	expression tag	UNP A0A059VAZ3
A	423	GLN	GLU	engineered mutation	UNP A0A059VAZ3
B	239	GLY	-	expression tag	UNP A0A059VAZ3
B	240	HIS	-	expression tag	UNP A0A059VAZ3
B	241	MET	-	expression tag	UNP A0A059VAZ3
B	423	GLN	GLU	engineered mutation	UNP A0A059VAZ3
C	239	GLY	-	expression tag	UNP A0A059VAZ3
C	240	HIS	-	expression tag	UNP A0A059VAZ3
C	241	MET	-	expression tag	UNP A0A059VAZ3
C	423	GLN	GLU	engineered mutation	UNP A0A059VAZ3
F	239	GLY	-	expression tag	UNP A0A059VAZ3
F	240	HIS	-	expression tag	UNP A0A059VAZ3
F	241	MET	-	expression tag	UNP A0A059VAZ3
F	423	GLN	GLU	engineered mutation	UNP A0A059VAZ3
D	239	GLY	-	expression tag	UNP A0A059VAZ3

Continued on next page...

Continued from previous page...

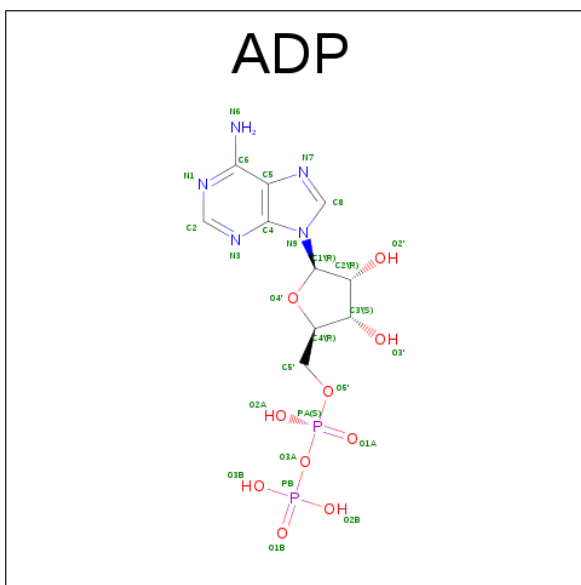
Chain	Residue	Modelled	Actual	Comment	Reference
D	240	HIS	-	expression tag	UNP A0A059VAZ3
D	241	MET	-	expression tag	UNP A0A059VAZ3
D	423	GLN	GLU	engineered mutation	UNP A0A059VAZ3
E	239	GLY	-	expression tag	UNP A0A059VAZ3
E	240	HIS	-	expression tag	UNP A0A059VAZ3
E	241	MET	-	expression tag	UNP A0A059VAZ3
E	423	GLN	GLU	engineered mutation	UNP A0A059VAZ3

- Molecule 2 is N-[(1R)-1-(dihydroxyboranyl)-2-phenylethyl]-Nalpha-(pyrazin-2-ylcarbonyl)-L-phenylalaninamide (three-letter code: 4KZ) (formula: C₂₂H₂₃BN₄O₄).

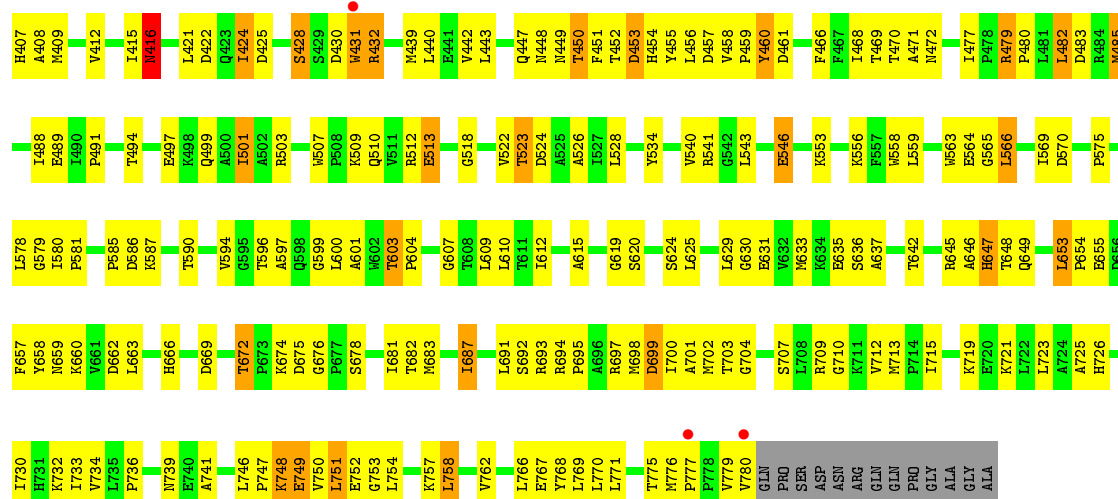


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	B	C	N	O	0	0
			31	1	22	4	4		
2	B	1	Total	B	C	N	O	0	0
			31	1	22	4	4		
2	C	1	Total	B	C	N	O	0	0
			31	1	22	4	4		
2	F	1	Total	B	C	N	O	0	0
			31	1	22	4	4		
2	D	1	Total	B	C	N	O	0	0
			31	1	22	4	4		
2	E	1	Total	B	C	N	O	0	0
			31	1	22	4	4		

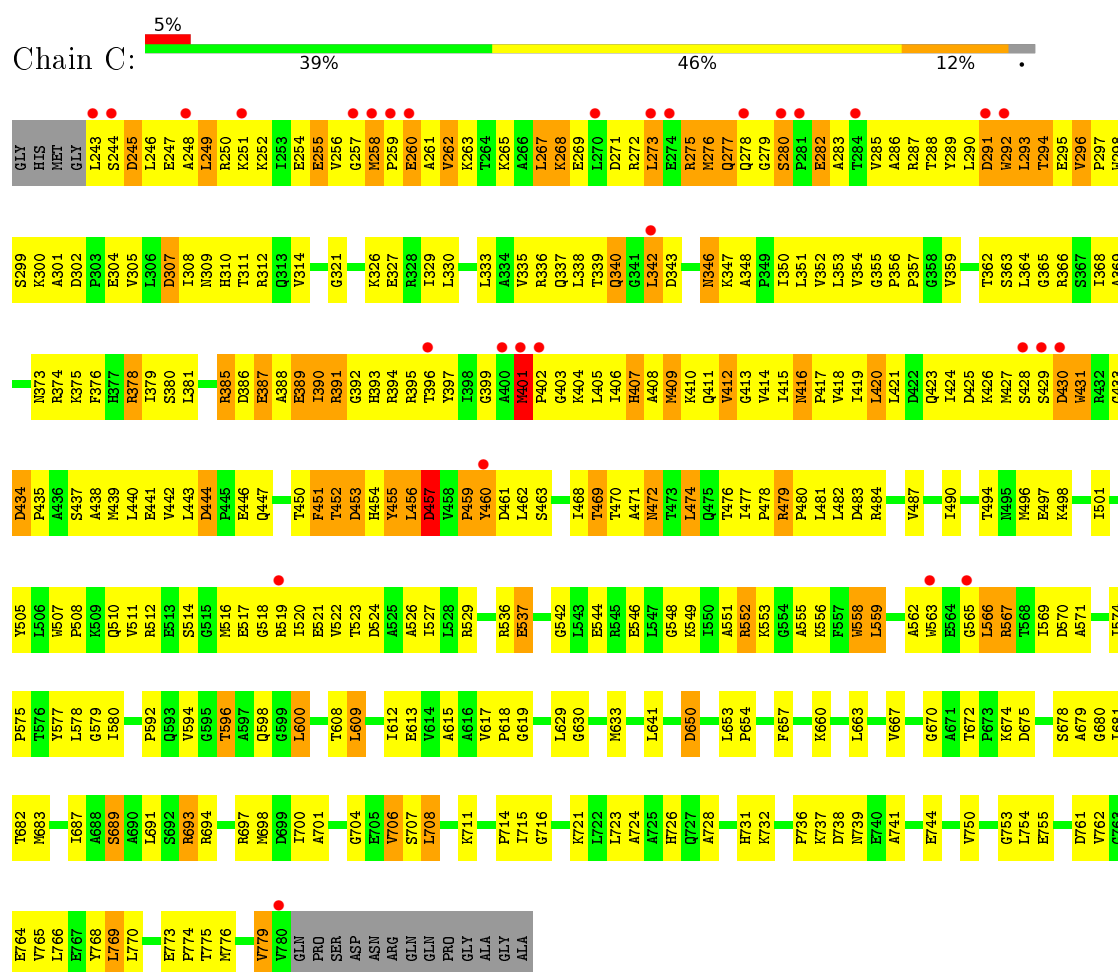
- Molecule 3 is ADENOSINE-5'-DIPHOSPHATE (three-letter code: ADP) (formula: C₁₀H₁₅N₅O₁₀P₂).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	B	1	Total 27	C 10	N 5	O 10	P 2	0	0
3	F	1	Total 27	C 10	N 5	O 10	P 2	0	0
3	D	1	Total 27	C 10	N 5	O 10	P 2	0	0

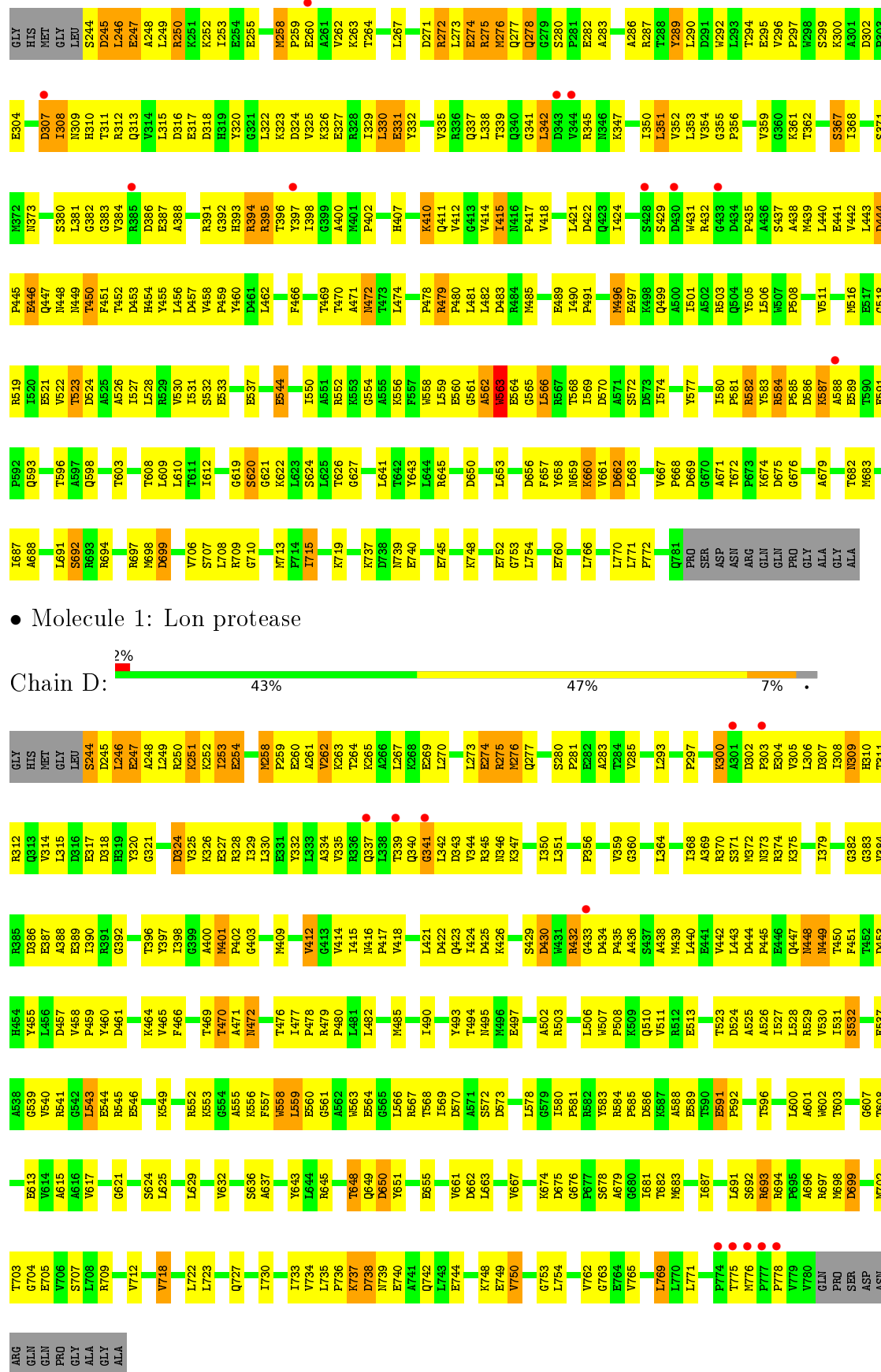


• Molecule 1: Lon protease

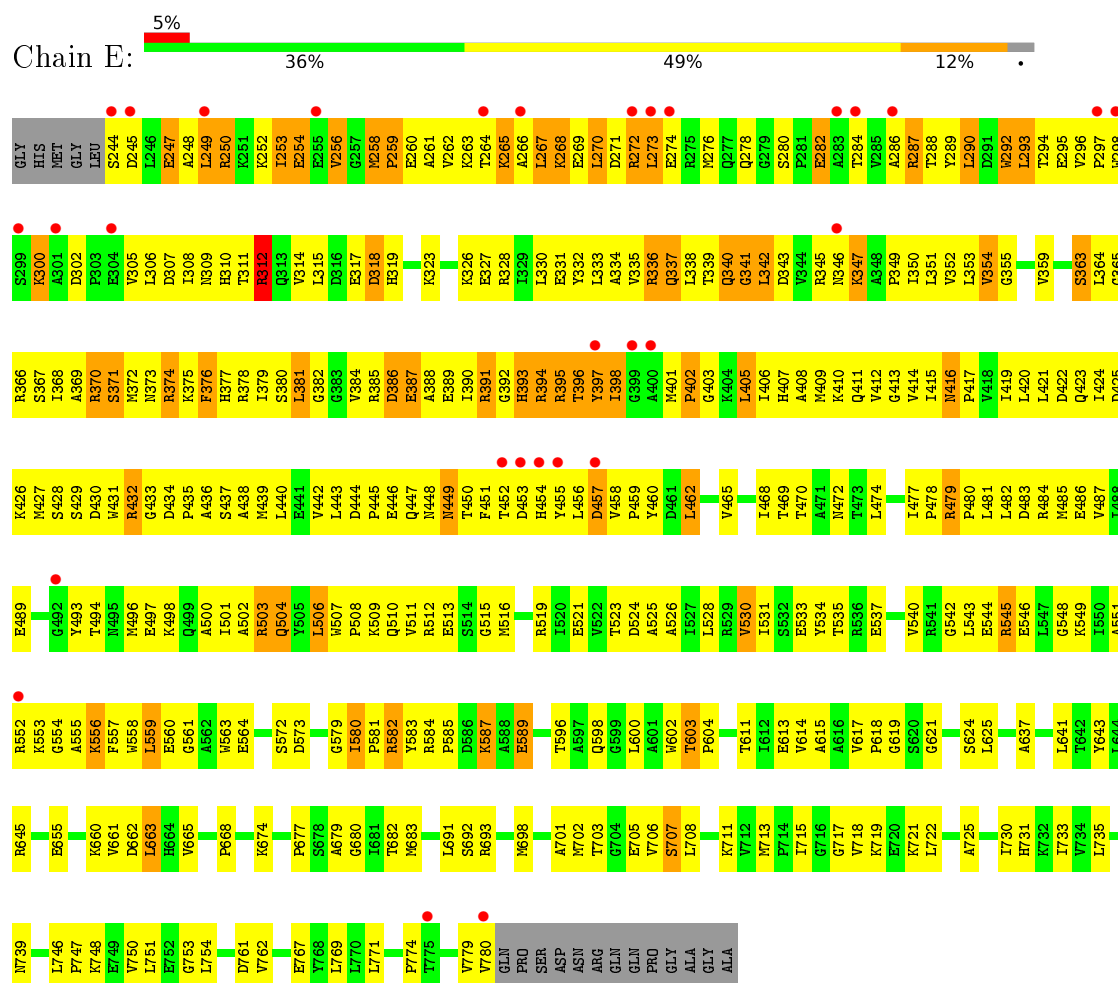


• Molecule 1: Lon protease





● Molecule 1: Lon protease



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	169.04Å 169.13Å 135.36Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 3.45 46.90 – 3.44	Depositor EDS
% Data completeness (in resolution range)	93.5 (30.00-3.45) 93.0 (46.90-3.44)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.38 (at 3.40Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, R_{free}	0.237 , 0.265 0.237 , 0.262	Depositor DCC
R_{free} test set	2459 reflections (5.32%)	DCC
Wilson B-factor (Å ²)	46.1	Xtriage
Anisotropy	0.204	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 29.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.000 for k,h,-l	Xtriage
F_o, F_c correlation	0.88	EDS
Total number of atoms	25358	wwPDB-VP
Average B, all atoms (Å ²)	67.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.13% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: 4KZ, ADP

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	0.42	0/4258	0.72	4/5774 (0.1%)
1	B	0.41	0/4258	0.71	3/5774 (0.1%)
1	C	0.42	0/4266	0.75	1/5785 (0.0%)
1	D	0.41	0/4258	0.68	0/5774
1	E	0.42	0/4258	0.74	4/5774 (0.1%)
1	F	0.42	0/4267	0.72	2/5786 (0.0%)
All	All	0.42	0/25565	0.72	14/34667 (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	A	0	1
1	B	0	2
1	C	0	4
1	D	0	2
1	E	0	5
1	F	0	1
All	All	0	15

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	391	ARG	N-CA-C	-8.40	88.33	111.00
1	E	312	ARG	NE-CZ-NH2	8.09	124.34	120.30
1	F	563	TRP	N-CA-C	-7.10	91.84	111.00
1	E	312	ARG	NE-CZ-NH1	-5.93	117.33	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	312	ARG	CG-CD-NE	5.78	123.94	111.80

There are no chirality outliers.

5 of 15 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	457	ASP	Peptide
1	B	431	TRP	Peptide
1	B	564	GLU	Peptide
1	C	451	PHE	Peptide
1	C	453	ASP	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	4179	0	4264	298	0
1	B	4179	0	4264	297	0
1	C	4187	0	4275	583	0
1	D	4179	0	4264	306	0
1	E	4179	0	4262	524	0
1	F	4188	0	4272	334	0
2	A	31	0	0	4	0
2	B	31	0	0	2	0
2	C	31	0	0	0	0
2	D	31	0	0	4	0
2	E	31	0	0	0	0
2	F	31	0	0	1	0
3	B	27	0	12	0	0
3	D	27	0	12	0	0
3	F	27	0	12	3	0
All	All	25358	0	25637	2297	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 45.

The worst 5 of 2297 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:C:393:HIS:CB	1:C:402:PRO:HD3	1.35	1.54
1:C:393:HIS:HB2	1:C:402:PRO:CD	1.41	1.45
1:E:402:PRO:HD3	1:E:455:TYR:CD2	1.59	1.36
1:C:393:HIS:N	1:C:454:HIS:HB2	1.40	1.32
1:E:381:LEU:HD23	1:E:424:ILE:CG2	1.60	1.31

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	535/555 (96%)	498 (93%)	36 (7%)	1 (0%)	52	87
1	B	535/555 (96%)	501 (94%)	33 (6%)	1 (0%)	52	87
1	C	536/555 (97%)	492 (92%)	42 (8%)	2 (0%)	39	80
1	D	535/555 (96%)	499 (93%)	36 (7%)	0	100	100
1	E	535/555 (96%)	505 (94%)	27 (5%)	3 (1%)	30	74
1	F	536/555 (97%)	499 (93%)	35 (6%)	2 (0%)	39	80
All	All	3212/3330 (96%)	2994 (93%)	209 (6%)	9 (0%)	46	84

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	348	ALA
1	C	401	MET
1	F	308	ILE
1	E	416	ASN
1	B	416	ASN

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	448/460 (97%)	388 (87%)	60 (13%)	5	25
1	B	448/460 (97%)	398 (89%)	50 (11%)	7	33
1	C	449/460 (98%)	370 (82%)	79 (18%)	2	12
1	D	448/460 (97%)	387 (86%)	61 (14%)	5	24
1	E	448/460 (97%)	363 (81%)	85 (19%)	2	9
1	F	449/460 (98%)	389 (87%)	60 (13%)	5	25
All	All	2690/2760 (98%)	2295 (85%)	395 (15%)	4	21

5 of 395 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	693	ARG
1	F	444	ASP
1	E	503	ARG
1	C	737	LYS
1	F	299	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 44 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	340	GLN
1	C	454	HIS
1	E	377	HIS
1	C	393	HIS
1	C	447	GLN

5.3.3 RNA ⓘ

There are no RNA molecules 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 ⓘ

9 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$
2	4KZ	A	801	1	32,33,33	0.60	0	35,43,43	1.22	5 (14%)
3	ADP	B	801	-	24,29,29	1.02	1 (4%)	23,45,45	1.80	2 (8%)
2	4KZ	B	802	1	32,33,33	0.76	0	35,43,43	1.19	5 (14%)
2	4KZ	C	801	1	32,33,33	0.48	0	35,43,43	1.22	5 (14%)
3	ADP	D	801	-	24,29,29	0.92	1 (4%)	23,45,45	2.05	4 (17%)
2	4KZ	D	802	1	32,33,33	0.73	1 (3%)	35,43,43	1.18	2 (5%)
2	4KZ	E	801	1	32,33,33	0.83	2 (6%)	35,43,43	1.21	5 (14%)
3	ADP	F	801	-	24,29,29	1.01	1 (4%)	23,45,45	2.00	1 (4%)
2	4KZ	F	802	1	32,33,33	0.62	0	35,43,43	1.07	4 (11%)

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
2	4KZ	A	801	1	-	0/22/28/28	0/3/3/3
3	ADP	B	801	-	-	0/12/32/32	0/3/3/3
2	4KZ	B	802	1	-	0/22/28/28	0/3/3/3
2	4KZ	C	801	1	-	0/22/28/28	0/3/3/3

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	ADP	D	801	-	-	0/12/32/32	0/3/3/3
2	4KZ	D	802	1	-	0/22/28/28	0/3/3/3
2	4KZ	E	801	1	-	0/22/28/28	0/3/3/3
3	ADP	F	801	-	-	0/12/32/32	0/3/3/3
2	4KZ	F	802	1	-	0/22/28/28	0/3/3/3

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	802	4KZ	C21-N20	-2.23	1.41	1.47
2	E	801	4KZ	C10-C18	-2.13	1.46	1.52
2	E	801	4KZ	C21-N20	-2.04	1.41	1.47
3	D	801	ADP	C5-C4	2.55	1.46	1.40
3	F	801	ADP	C5-C4	2.57	1.46	1.40

The worst 5 of 33 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	F	801	ADP	N3-C2-N1	-8.08	122.53	128.87
3	D	801	ADP	N3-C2-N1	-8.02	122.57	128.87
3	B	801	ADP	N3-C2-N1	-6.97	123.39	128.87
3	D	801	ADP	C1'-N9-C4	-3.16	123.28	126.81
2	E	801	4KZ	C2-C3-N4	-2.84	118.59	122.05

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	801	4KZ	4	0
2	B	802	4KZ	2	0
2	D	802	4KZ	4	0
3	F	801	ADP	3	0
2	F	802	4KZ	1	0

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 ⓘ

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, 95th 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	537/555 (96%)	-0.06	4 (0%) 89 84	31, 57, 94, 130	0
1	B	537/555 (96%)	0.01	7 (1%) 79 72	32, 62, 112, 141	0
1	C	538/555 (96%)	0.23	30 (5%) 28 25	27, 63, 121, 148	0
1	D	537/555 (96%)	0.01	11 (2%) 68 62	27, 59, 109, 127	0
1	E	537/555 (96%)	0.28	30 (5%) 28 25	31, 75, 132, 155	0
1	F	538/555 (96%)	-0.11	10 (1%) 70 63	25, 55, 104, 129	0
All	All	3224/3330 (96%)	0.06	92 (2%) 55 48	25, 61, 115, 155	0

The worst 5 of 92 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	255	GLU	4.4
1	C	281	PRO	4.1
1	C	429	SER	3.9
1	C	396	THR	3.6
1	E	273	LEU	3.5

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

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

6.3 Carbohydrates ⓘ

There are no carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th 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	LLDF	B-factors(Å ²)	Q<0.9
2	4KZ	E	801	31/31	0.94	0.26	2.41	37,42,47,48	0
2	4KZ	A	801	31/31	0.94	0.30	2.19	46,52,58,59	0
2	4KZ	C	801	31/31	0.94	0.28	2.04	31,32,37,39	0
2	4KZ	D	802	31/31	0.92	0.24	1.54	31,38,50,55	0
2	4KZ	B	802	31/31	0.92	0.24	1.52	42,47,64,70	0
2	4KZ	F	802	31/31	0.94	0.21	1.49	34,41,44,44	0
3	ADP	B	801	27/27	0.95	0.19	-0.46	38,45,52,55	0
3	ADP	D	801	27/27	0.96	0.17	-0.48	41,45,49,51	0
3	ADP	F	801	27/27	0.96	0.18	-0.54	38,46,50,52	0

6.5 Other polymers [i](#)

There are no such residues in this entry.