wwPDB X-ray Structure Validation Summary Report (i)

Oct 9, 2023 – 04:57 PM EDT

PDB ID	:	8DS1
Title	:	Structure of SARS-CoV-2 Mpro in complex with nsp12-nsp13 (C12) cut site sequence
Authors	:	Lee, J.; Kenward, C.; Worrall, L.J.; Vuckovic, M.; Paetzel, M.; Strynadka, N.C.J.
Deposited on	:	2022-07-21
Resolution	:	2.19 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:

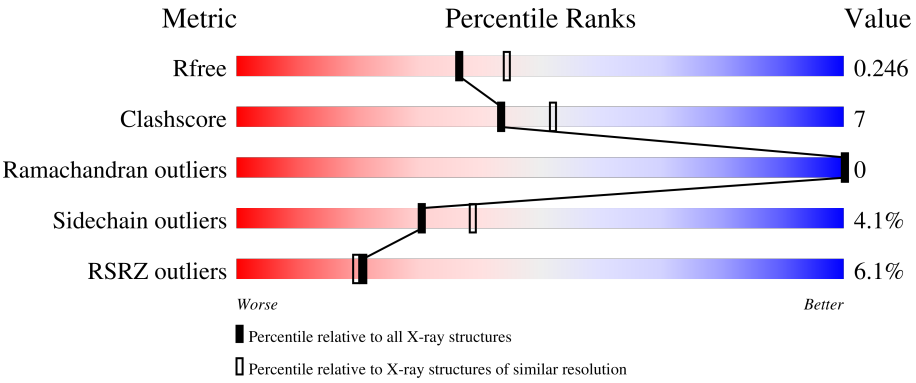
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 2.19 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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	306	2% 79% 19%
1	B	306	3% 77% 20%
1	C	306	19% 73% 24%
1	D	306	21% 78% 20%
1	E	306	2% 82% 15%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	306	3% 86% 12%
1	G	306	% 84% 15%
1	H	306	82% 17%
1	I	306	6% 83% 16%
1	J	306	10% 80% 19%
1	K	306	% 84% 15%
1	L	306	4% 83% 15%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 28279 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 3C-like proteinase nsp5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	302	Total	C	N	O	S	0	0	0
			2335	1479	397	437	22			
1	B	302	Total	C	N	O	S	0	1	0
			2343	1484	400	437	22			
1	C	302	Total	C	N	O	S	0	0	0
			2285	1451	385	428	21			
1	D	306	Total	C	N	O	S	0	0	0
			2288	1446	388	434	20			
1	E	302	Total	C	N	O	S	0	0	0
			2326	1473	394	437	22			
1	F	302	Total	C	N	O	S	0	0	0
			2332	1476	397	437	22			
1	G	302	Total	C	N	O	S	0	0	0
			2320	1471	395	432	22			
1	H	302	Total	C	N	O	S	0	1	0
			2343	1484	400	437	22			
1	I	306	Total	C	N	O	S	0	0	0
			2367	1500	404	441	22			
1	J	306	Total	C	N	O	S	0	0	0
			2347	1484	400	441	22			
1	K	302	Total	C	N	O	S	0	1	0
			2337	1481	396	438	22			
1	L	302	Total	C	N	O	S	0	0	0
			2330	1475	397	436	22			

There are 72 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	301	PRO	-	expression tag	UNP P0DTD1
A	302	HIS	-	expression tag	UNP P0DTD1
A	303	THR	-	expression tag	UNP P0DTD1
A	304	VAL	-	expression tag	UNP P0DTD1
A	305	LEU	-	expression tag	UNP P0DTD1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	306	GLN	-	expression tag	UNP P0DTD1
B	301	PRO	-	expression tag	UNP P0DTD1
B	302	HIS	-	expression tag	UNP P0DTD1
B	303	THR	-	expression tag	UNP P0DTD1
B	304	VAL	-	expression tag	UNP P0DTD1
B	305	LEU	-	expression tag	UNP P0DTD1
B	306	GLN	-	expression tag	UNP P0DTD1
C	301	PRO	-	expression tag	UNP P0DTD1
C	302	HIS	-	expression tag	UNP P0DTD1
C	303	THR	-	expression tag	UNP P0DTD1
C	304	VAL	-	expression tag	UNP P0DTD1
C	305	LEU	-	expression tag	UNP P0DTD1
C	306	GLN	-	expression tag	UNP P0DTD1
D	301	PRO	-	expression tag	UNP P0DTD1
D	302	HIS	-	expression tag	UNP P0DTD1
D	303	THR	-	expression tag	UNP P0DTD1
D	304	VAL	-	expression tag	UNP P0DTD1
D	305	LEU	-	expression tag	UNP P0DTD1
D	306	GLN	-	expression tag	UNP P0DTD1
E	301	PRO	-	expression tag	UNP P0DTD1
E	302	HIS	-	expression tag	UNP P0DTD1
E	303	THR	-	expression tag	UNP P0DTD1
E	304	VAL	-	expression tag	UNP P0DTD1
E	305	LEU	-	expression tag	UNP P0DTD1
E	306	GLN	-	expression tag	UNP P0DTD1
F	301	PRO	-	expression tag	UNP P0DTD1
F	302	HIS	-	expression tag	UNP P0DTD1
F	303	THR	-	expression tag	UNP P0DTD1
F	304	VAL	-	expression tag	UNP P0DTD1
F	305	LEU	-	expression tag	UNP P0DTD1
F	306	GLN	-	expression tag	UNP P0DTD1
G	301	PRO	-	expression tag	UNP P0DTD1
G	302	HIS	-	expression tag	UNP P0DTD1
G	303	THR	-	expression tag	UNP P0DTD1
G	304	VAL	-	expression tag	UNP P0DTD1
G	305	LEU	-	expression tag	UNP P0DTD1
G	306	GLN	-	expression tag	UNP P0DTD1
H	301	PRO	-	expression tag	UNP P0DTD1
H	302	HIS	-	expression tag	UNP P0DTD1
H	303	THR	-	expression tag	UNP P0DTD1
H	304	VAL	-	expression tag	UNP P0DTD1
H	305	LEU	-	expression tag	UNP P0DTD1

Continued on next page...

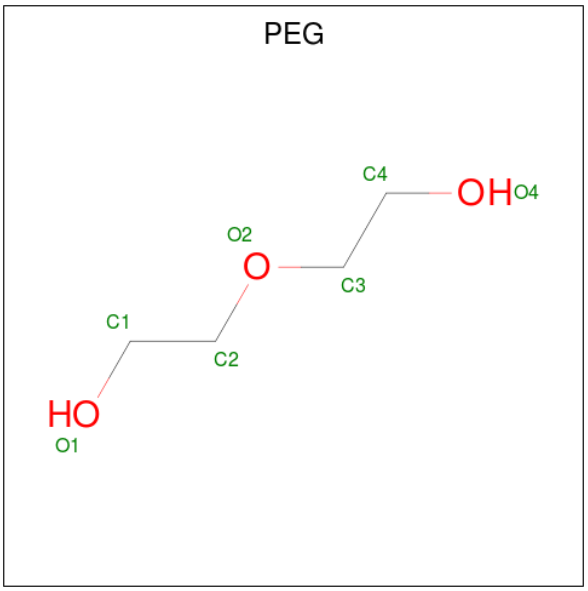
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
H	306	GLN	-	expression tag	UNP P0DTD1
I	301	PRO	-	expression tag	UNP P0DTD1
I	302	HIS	-	expression tag	UNP P0DTD1
I	303	THR	-	expression tag	UNP P0DTD1
I	304	VAL	-	expression tag	UNP P0DTD1
I	305	LEU	-	expression tag	UNP P0DTD1
I	306	GLN	-	expression tag	UNP P0DTD1
J	301	PRO	-	expression tag	UNP P0DTD1
J	302	HIS	-	expression tag	UNP P0DTD1
J	303	THR	-	expression tag	UNP P0DTD1
J	304	VAL	-	expression tag	UNP P0DTD1
J	305	LEU	-	expression tag	UNP P0DTD1
J	306	GLN	-	expression tag	UNP P0DTD1
K	301	PRO	-	expression tag	UNP P0DTD1
K	302	HIS	-	expression tag	UNP P0DTD1
K	303	THR	-	expression tag	UNP P0DTD1
K	304	VAL	-	expression tag	UNP P0DTD1
K	305	LEU	-	expression tag	UNP P0DTD1
K	306	GLN	-	expression tag	UNP P0DTD1
L	301	PRO	-	expression tag	UNP P0DTD1
L	302	HIS	-	expression tag	UNP P0DTD1
L	303	THR	-	expression tag	UNP P0DTD1
L	304	VAL	-	expression tag	UNP P0DTD1
L	305	LEU	-	expression tag	UNP P0DTD1
L	306	GLN	-	expression tag	UNP P0DTD1

- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total Na 1 1	0	0
2	F	1	Total Na 1 1	0	0
2	G	1	Total Na 1 1	0	0
2	K	1	Total Na 1 1	0	0
2	L	1	Total Na 1 1	0	0

- Molecule 3 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 7 4 3	0	0
3	F	1	Total C O 7 4 3	0	0
3	G	1	Total C O 7 4 3	0	0
3	I	1	Total C O 7 4 3	0	0
3	L	1	Total C O 7 4 3	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	29	Total O 29 29	0	0
4	B	22	Total O 22 22	0	0
4	C	8	Total O 8 8	0	0
4	D	1	Total O 1 1	0	0
4	E	22	Total O 22 22	0	0
4	F	12	Total O 12 12	0	0
4	G	58	Total O 58 58	0	0

Continued on next page...

Continued from previous page...

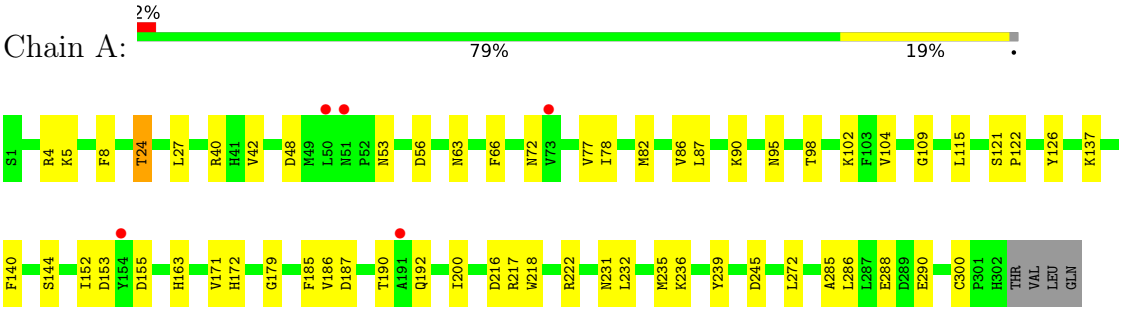
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	H	50	Total O 50 50	0	0
4	I	26	Total O 26 26	0	0
4	J	13	Total O 13 13	0	0
4	K	22	Total O 22 22	0	0
4	L	23	Total O 23 23	0	0

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.

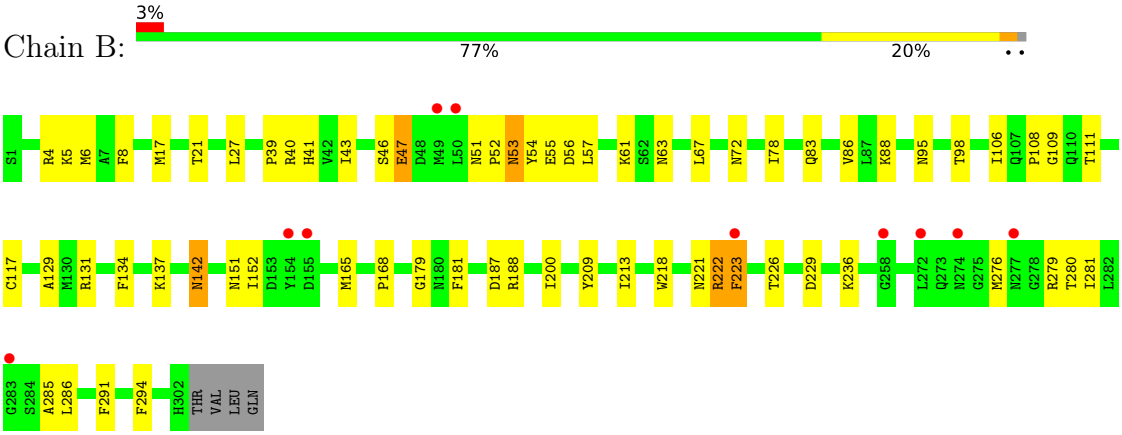
• Molecule 1: 3C-like proteinase nsp5

Chain A:



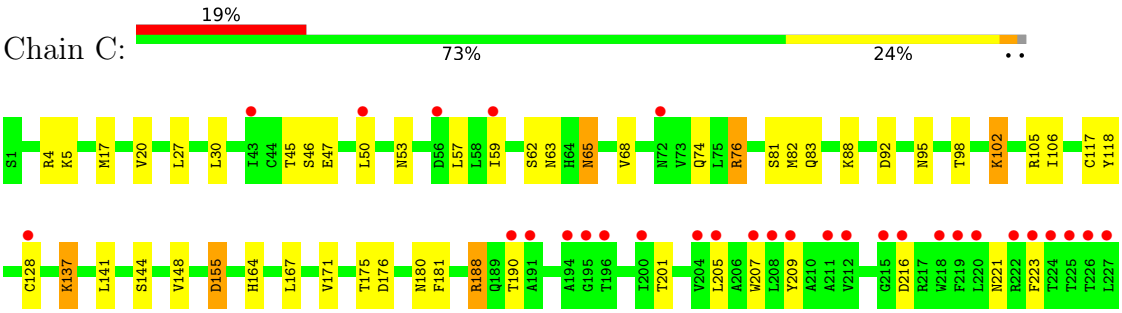
• Molecule 1: 3C-like proteinase nsp5

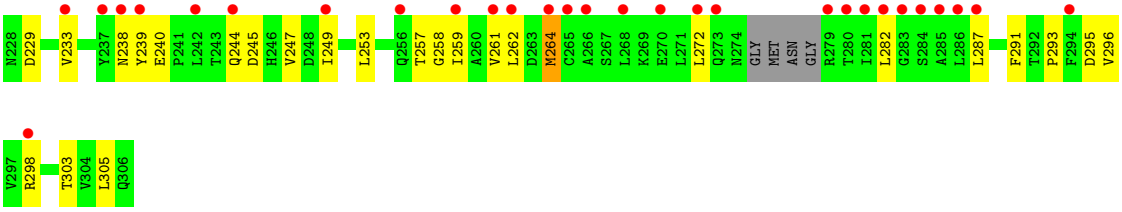
Chain B:



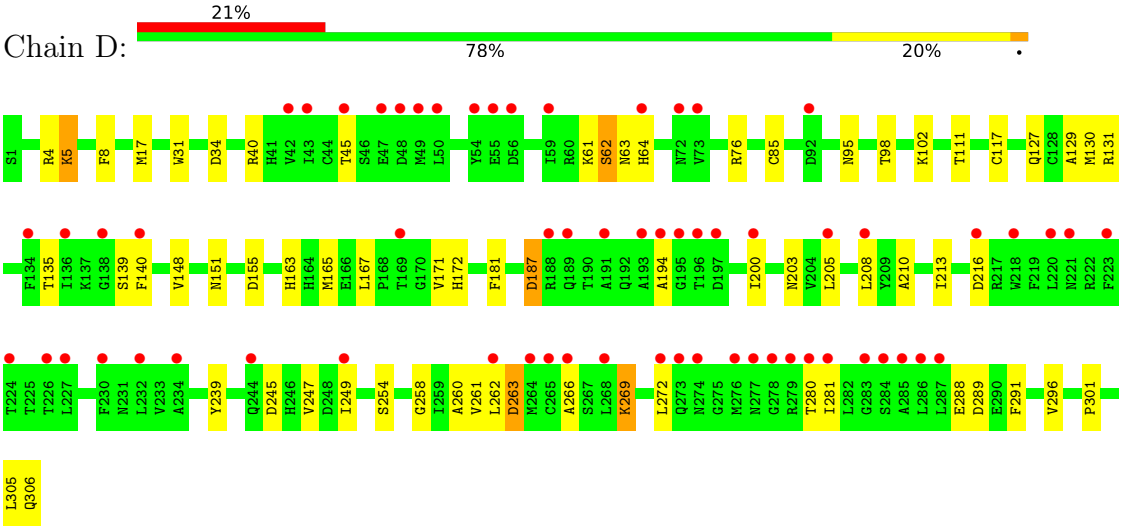
• Molecule 1: 3C-like proteinase nsp5

Chain C:

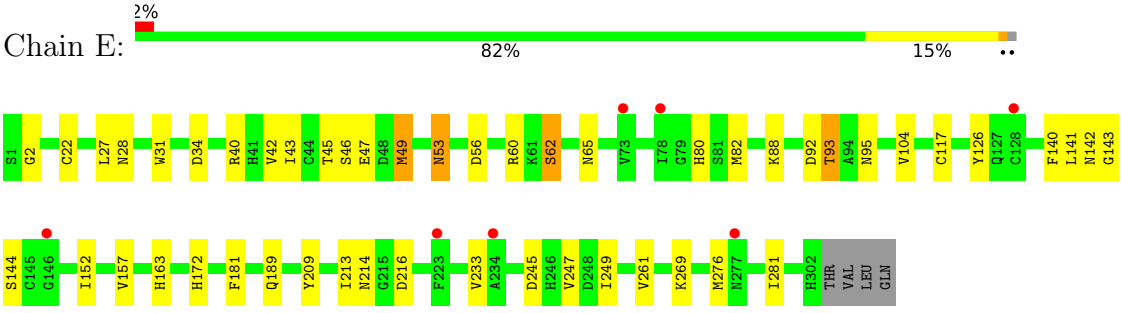




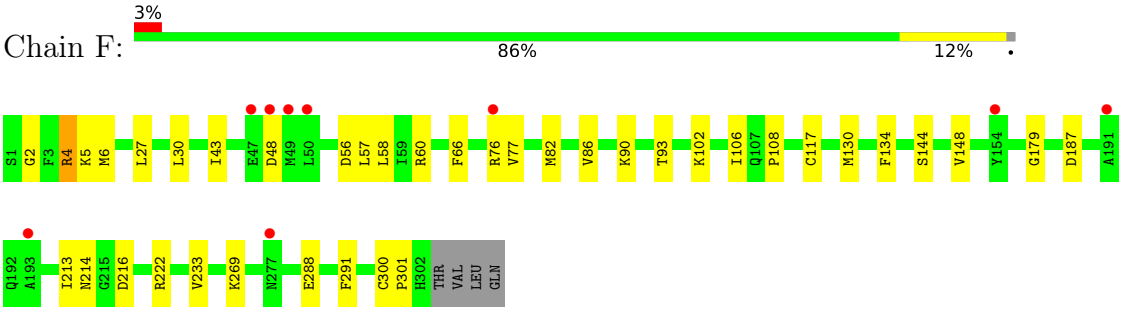
• Molecule 1: 3C-like proteinase nsp5



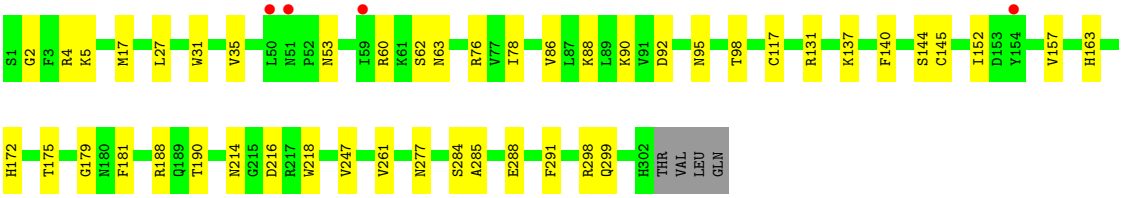
• Molecule 1: 3C-like proteinase nsp5



• Molecule 1: 3C-like proteinase nsp5

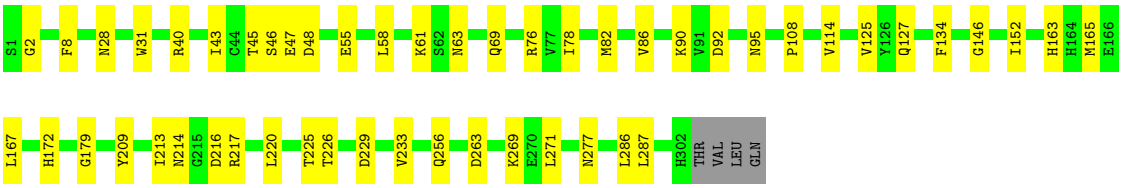


• Molecule 1: 3C-like proteinase nsp5



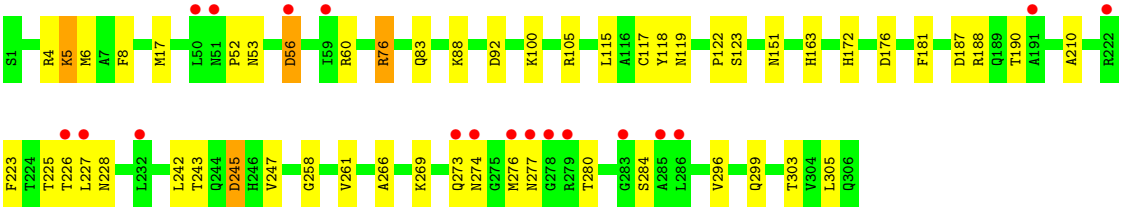
• Molecule 1: 3C-like proteinase nsp5

Chain H:



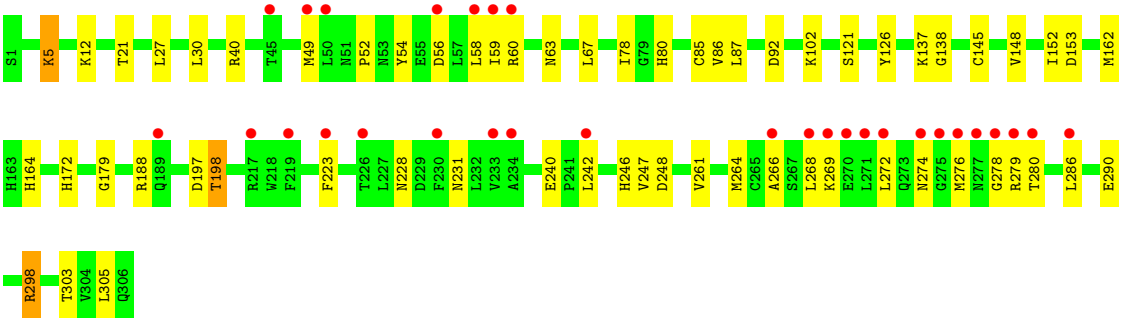
• Molecule 1: 3C-like proteinase nsp5

Chain I:



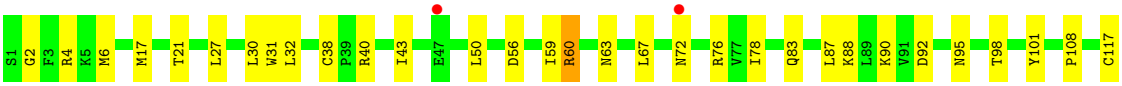
• Molecule 1: 3C-like proteinase nsp5

Chain J:

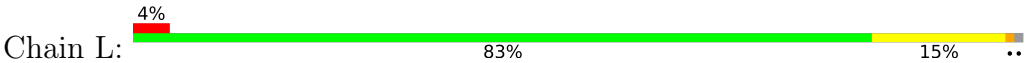


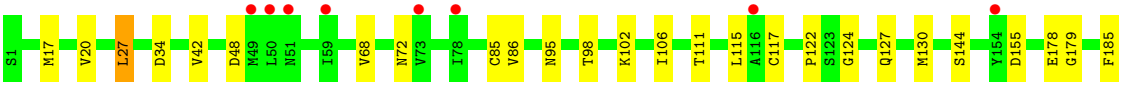
• Molecule 1: 3C-like proteinase nsp5

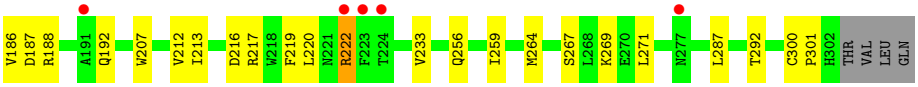
Chain K:



• Molecule 1: 3C-like proteinase nsp5







4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	67.00Å 105.78Å 276.34Å 90.00° 90.93° 90.00°	Depositor
Resolution (Å)	48.52 – 2.19 48.52 – 2.19	Depositor EDS
% Data completeness (in resolution range)	67.3 (48.52-2.19) 67.3 (48.52-2.19)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.48 (at 2.20Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.198 , 0.248 0.198 , 0.246	Depositor DCC
R_{free} test set	6703 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å^2)	38.5	Xtriage
Anisotropy	0.087	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 42.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.039 for h,-k,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	28279	wwPDB-VP
Average B, all atoms (Å^2)	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 23.86 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 4.2929e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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: NA, PEG

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	0.25	0/2388	0.47	0/3247
1	B	0.26	0/2399	0.50	0/3261
1	C	0.26	0/2337	0.48	0/3186
1	D	0.25	0/2340	0.49	0/3193
1	E	0.26	0/2379	0.48	0/3236
1	F	0.25	0/2385	0.49	0/3243
1	G	0.26	0/2373	0.48	0/3228
1	H	0.25	0/2399	0.49	0/3261
1	I	0.26	0/2421	0.48	0/3293
1	J	0.25	0/2401	0.49	0/3267
1	K	0.26	0/2393	0.49	0/3254
1	L	0.26	0/2383	0.49	0/3240
All	All	0.26	0/28598	0.49	0/38909

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	2335	0	2282	34	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	2343	0	2294	42	0
1	C	2285	0	2190	49	0
1	D	2288	0	2159	41	0
1	E	2326	0	2262	33	0
1	F	2332	0	2273	21	0
1	G	2320	0	2256	27	0
1	H	2343	0	2294	29	0
1	I	2367	0	2318	37	0
1	J	2347	0	2264	39	0
1	K	2337	0	2282	26	0
1	L	2330	0	2268	28	0
2	A	1	0	0	0	0
2	F	1	0	0	0	0
2	G	1	0	0	0	0
2	K	1	0	0	0	0
2	L	1	0	0	0	0
3	A	7	0	10	0	0
3	F	7	0	10	2	0
3	G	7	0	10	1	0
3	I	7	0	10	2	0
3	L	7	0	10	0	0
4	A	29	0	0	3	0
4	B	22	0	0	1	0
4	C	8	0	0	0	0
4	D	1	0	0	0	0
4	E	22	0	0	0	0
4	F	12	0	0	0	0
4	G	58	0	0	2	0
4	H	50	0	0	0	0
4	I	26	0	0	0	0
4	J	13	0	0	1	0
4	K	22	0	0	0	0
4	L	23	0	0	1	0
All	All	28279	0	27192	372	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 7.

The worst 5 of 372 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:J:138:GLY:O	1:J:172:HIS:HE1	1.58	0.86
1:I:4:ARG:H	1:I:299:GLN:HE22	1.22	0.85
1:C:257:THR:HG23	1:C:259:ILE:H	1.45	0.81
1:E:62:SER:H	1:E:65:ASN:HD22	1.29	0.80
1:A:186:VAL:H	1:A:192:GLN:HE22	1.29	0.79

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	300/306 (98%)	296 (99%)	4 (1%)	0	100	100
1	B	301/306 (98%)	295 (98%)	6 (2%)	0	100	100
1	C	298/306 (97%)	293 (98%)	5 (2%)	0	100	100
1	D	304/306 (99%)	293 (96%)	11 (4%)	0	100	100
1	E	300/306 (98%)	291 (97%)	9 (3%)	0	100	100
1	F	300/306 (98%)	291 (97%)	9 (3%)	0	100	100
1	G	300/306 (98%)	296 (99%)	4 (1%)	0	100	100
1	H	301/306 (98%)	293 (97%)	8 (3%)	0	100	100
1	I	304/306 (99%)	303 (100%)	1 (0%)	0	100	100
1	J	304/306 (99%)	299 (98%)	5 (2%)	0	100	100
1	K	301/306 (98%)	295 (98%)	6 (2%)	0	100	100
1	L	300/306 (98%)	294 (98%)	6 (2%)	0	100	100
All	All	3613/3672 (98%)	3539 (98%)	74 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	259/264 (98%)	247 (95%)	12 (5%)	27	34
1	B	260/264 (98%)	250 (96%)	10 (4%)	33	42
1	C	247/264 (94%)	227 (92%)	20 (8%)	11	12
1	D	242/264 (92%)	229 (95%)	13 (5%)	22	26
1	E	257/264 (97%)	248 (96%)	9 (4%)	36	46
1	F	258/264 (98%)	249 (96%)	9 (4%)	36	46
1	G	254/264 (96%)	247 (97%)	7 (3%)	43	56
1	H	260/264 (98%)	253 (97%)	7 (3%)	44	57
1	I	263/264 (100%)	251 (95%)	12 (5%)	27	34
1	J	257/264 (97%)	244 (95%)	13 (5%)	24	29
1	K	259/264 (98%)	252 (97%)	7 (3%)	44	57
1	L	257/264 (97%)	249 (97%)	8 (3%)	40	51
All	All	3073/3168 (97%)	2946 (96%)	127 (4%)	30	39

5 of 127 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	53	ASN
1	J	280	THR
1	F	144	SER
1	J	274	ASN
1	K	216	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 78 such sidechains are listed below:

Mol	Chain	Res	Type
1	I	53	ASN
1	K	214	ASN
1	I	83	GLN
1	J	164	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	L	192	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 10 ligands modelled in this entry, 5 are monoatomic - leaving 5 for Mogul analysis.

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	$\#\|Z\| > 2$
3	PEG	L	402	-	6,6,6	0.12	0	5,5,5	0.08	0
3	PEG	F	402	-	6,6,6	0.10	0	5,5,5	0.10	0
3	PEG	G	402	-	6,6,6	0.12	0	5,5,5	0.06	0
3	PEG	A	402	-	6,6,6	0.10	0	5,5,5	0.10	0
3	PEG	I	401	-	6,6,6	0.10	0	5,5,5	0.09	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
3	PEG	L	402	-	-	0/4/4/4	-
3	PEG	F	402	-	-	2/4/4/4	-
3	PEG	G	402	-	-	1/4/4/4	-
3	PEG	A	402	-	-	3/4/4/4	-
3	PEG	I	401	-	-	0/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	F	402	PEG	O1-C1-C2-O2
3	F	402	PEG	O2-C3-C4-O4
3	G	402	PEG	C1-C2-O2-C3
3	A	402	PEG	C1-C2-O2-C3
3	A	402	PEG	C4-C3-O2-C2

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	F	402	PEG	2	0
3	G	402	PEG	1	0
3	I	401	PEG	2	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	302/306 (98%)	0.00	5 (1%) 70 68	26, 42, 78, 108	0
1	B	302/306 (98%)	0.16	10 (3%) 46 44	30, 49, 85, 127	0
1	C	302/306 (98%)	0.97	59 (19%) 1 1	34, 70, 117, 152	0
1	D	306/306 (100%)	1.06	63 (20%) 1 1	39, 75, 121, 139	0
1	E	302/306 (98%)	0.13	7 (2%) 60 58	28, 50, 78, 127	0
1	F	302/306 (98%)	0.19	9 (2%) 50 48	27, 51, 90, 137	0
1	G	302/306 (98%)	-0.05	4 (1%) 77 75	19, 33, 77, 121	0
1	H	302/306 (98%)	-0.06	0 100 100	19, 35, 64, 115	0
1	I	306/306 (100%)	0.25	18 (5%) 22 21	25, 46, 86, 113	0
1	J	306/306 (100%)	0.45	30 (9%) 7 6	27, 53, 104, 178	0
1	K	302/306 (98%)	0.07	4 (1%) 77 75	27, 45, 74, 130	0
1	L	302/306 (98%)	0.25	13 (4%) 35 33	29, 50, 91, 140	0
All	All	3636/3672 (99%)	0.29	222 (6%) 21 20	19, 49, 100, 178	0

The worst 5 of 222 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	283	GLY	11.8
1	D	191	ALA	7.8
1	D	283	GLY	7.7
1	C	285	ALA	6.9
1	J	230	PHE	6.5

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 (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
2	NA	L	401	1/1	0.69	0.25	83,83,83,83	0
2	NA	F	401	1/1	0.87	0.09	54,54,54,54	0
2	NA	A	401	1/1	0.87	0.10	55,55,55,55	0
3	PEG	F	402	7/7	0.89	0.18	43,57,65,67	0
3	PEG	I	401	7/7	0.89	0.15	42,51,57,59	0
3	PEG	A	402	7/7	0.90	0.20	51,56,65,66	0
3	PEG	L	402	7/7	0.91	0.22	47,53,58,60	0
3	PEG	G	402	7/7	0.92	0.21	38,43,48,49	0
2	NA	G	401	1/1	0.97	0.12	42,42,42,42	0
2	NA	K	401	1/1	0.97	0.19	56,56,56,56	0

6.5 Other polymers (i)

There are no such residues in this entry.