



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 10, 2023 – 07:56 AM EDT

PDB ID : 7MHP  
Title : Ensemble refinement structure of SARS-CoV-2 main protease (Mpro) at 298 K at high humidity  
Authors : Ebrahim, A.; Riley, B.T.; Kumaran, D.; Andi, B.; Fuchs, M.R.; McSweeney, S.; Keedy, D.A.  
Deposited on : 2021-04-15  
Resolution : 2.00 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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 : 2.35.1  
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

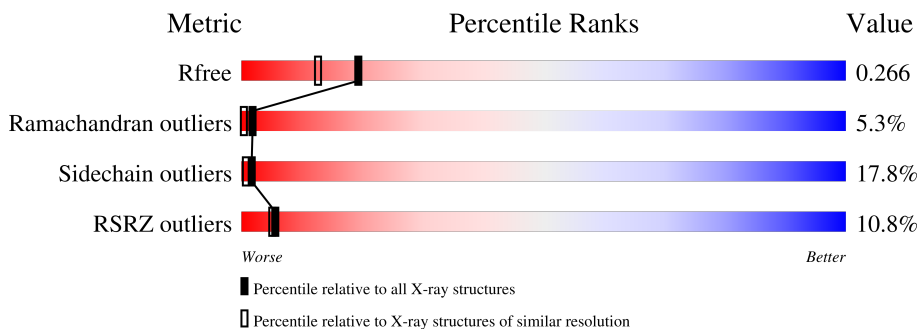
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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	8085 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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  $\geq 3$ , 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	1-A	306	11% (poor fit), 78% (0-1 outliers), 11% (2 outliers), 18% (3+ outliers)
1	10-A	306	11% (poor fit), 78% (0-1 outliers), 11% (2 outliers), 19% (3+ outliers)
1	11-A	306	11% (poor fit), 78% (0-1 outliers), 11% (2 outliers), 16% (3+ outliers), 6% (grey)
1	12-A	306	11% (poor fit), 76% (0-1 outliers), 11% (2 outliers), 21% (3+ outliers)
1	13-A	306	11% (poor fit), 76% (0-1 outliers), 11% (2 outliers), 20% (3+ outliers)
1	14-A	306	11% (poor fit), 79% (0-1 outliers), 11% (2 outliers), 18% (3+ outliers)
1	15-A	306	11% (poor fit), 79% (0-1 outliers), 11% (2 outliers), 17% (3+ outliers)

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	16-A	306	11% 79% 20% .
1	17-A	306	11% 78% 19% .
1	18-A	306	11% 81% 15% .
1	19-A	306	11% 83% 14% .
1	2-A	306	11% 77% 19% .
1	20-A	306	11% 78% 20% .
1	21-A	306	11% 82% 15% ..
1	22-A	306	11% 79% 19% ..
1	23-A	306	11% 76% 20% .
1	24-A	306	11% 78% 19% .
1	25-A	306	11% 80% 15% ..
1	26-A	306	11% 81% 17% .
1	27-A	306	11% 80% 17% ..
1	28-A	306	11% 77% 20% .
1	3-A	306	11% 80% 16% ..
1	4-A	306	11% 80% 18% ..
1	5-A	306	11% 80% 16% .
1	6-A	306	11% 83% 15% .
1	7-A	306	11% 78% 18% .
1	8-A	306	11% 82% 14% .
1	9-A	306	11% 80% 16% ..

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 132023 atoms, of which 64764 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.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	1-A	306	4680	1499	2313	402	444	22	0	0	0
1	2-A	306	4680	1499	2313	402	444	22	0	0	0
1	3-A	306	4680	1499	2313	402	444	22	0	0	0
1	4-A	306	4680	1499	2313	402	444	22	0	0	0
1	5-A	306	4680	1499	2313	402	444	22	0	0	0
1	6-A	306	4680	1499	2313	402	444	22	0	0	0
1	7-A	306	4680	1499	2313	402	444	22	0	0	0
1	8-A	306	4680	1499	2313	402	444	22	0	0	0
1	9-A	306	4680	1499	2313	402	444	22	0	0	0
1	10-A	306	4680	1499	2313	402	444	22	0	0	0
1	11-A	306	4680	1499	2313	402	444	22	0	0	0
1	12-A	306	4680	1499	2313	402	444	22	0	0	0
1	13-A	306	4680	1499	2313	402	444	22	0	0	0
1	14-A	306	4680	1499	2313	402	444	22	0	0	0
1	15-A	306	4680	1499	2313	402	444	22	0	0	0
1	16-A	306	4680	1499	2313	402	444	22	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
1	17-A	306	Total	C	H	N	O	S	0	0	0
			4680	1499	2313	402	444	22			
1	18-A	306	Total	C	H	N	O	S	0	0	0
			4680	1499	2313	402	444	22			
1	19-A	306	Total	C	H	N	O	S	0	0	0
			4680	1499	2313	402	444	22			
1	20-A	306	Total	C	H	N	O	S	0	0	0
			4680	1499	2313	402	444	22			
1	21-A	306	Total	C	H	N	O	S	0	0	0
			4680	1499	2313	402	444	22			
1	22-A	306	Total	C	H	N	O	S	0	0	0
			4680	1499	2313	402	444	22			
1	23-A	306	Total	C	H	N	O	S	0	0	0
			4680	1499	2313	402	444	22			
1	24-A	306	Total	C	H	N	O	S	0	0	0
			4680	1499	2313	402	444	22			
1	25-A	306	Total	C	H	N	O	S	0	0	0
			4680	1499	2313	402	444	22			
1	26-A	306	Total	C	H	N	O	S	0	0	0
			4680	1499	2313	402	444	22			
1	27-A	306	Total	C	H	N	O	S	0	0	0
			4680	1499	2313	402	444	22			
1	28-A	306	Total	C	H	N	O	S	0	0	0
			4680	1499	2313	402	444	22			

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	1-A	1	Total	Zn	0	0
			1	1		
2	2-A	1	Total	Zn	0	0
			1	1		
2	3-A	1	Total	Zn	0	0
			1	1		
2	4-A	1	Total	Zn	0	0
			1	1		
2	5-A	1	Total	Zn	0	0
			1	1		
2	6-A	1	Total	Zn	0	0
			1	1		
2	7-A	1	Total	Zn	0	0
			1	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	8-A	1	Total 1	Zn 1	0	0
2	9-A	1	Total 1	Zn 1	0	0
2	10-A	1	Total 1	Zn 1	0	0
2	11-A	1	Total 1	Zn 1	0	0
2	12-A	1	Total 1	Zn 1	0	0
2	13-A	1	Total 1	Zn 1	0	0
2	14-A	1	Total 1	Zn 1	0	0
2	15-A	1	Total 1	Zn 1	0	0
2	16-A	1	Total 1	Zn 1	0	0
2	17-A	1	Total 1	Zn 1	0	0
2	18-A	1	Total 1	Zn 1	0	0
2	19-A	1	Total 1	Zn 1	0	0
2	20-A	1	Total 1	Zn 1	0	0
2	21-A	1	Total 1	Zn 1	0	0
2	22-A	1	Total 1	Zn 1	0	0
2	23-A	1	Total 1	Zn 1	0	0
2	24-A	1	Total 1	Zn 1	0	0
2	25-A	1	Total 1	Zn 1	0	0
2	26-A	1	Total 1	Zn 1	0	0
2	27-A	1	Total 1	Zn 1	0	0
2	28-A	1	Total 1	Zn 1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	1-A	34	Total O 34 34	0	0
3	2-A	29	Total O 29 29	0	0
3	3-A	24	Total O 24 24	0	0
3	4-A	37	Total O 37 37	0	0
3	5-A	46	Total O 46 46	0	0
3	6-A	30	Total O 30 30	0	0
3	7-A	45	Total O 45 45	0	0
3	8-A	30	Total O 30 30	0	0
3	9-A	35	Total O 35 35	0	0
3	10-A	38	Total O 38 38	0	0
3	11-A	41	Total O 41 41	0	0
3	12-A	34	Total O 34 34	0	0
3	13-A	36	Total O 36 36	0	0
3	14-A	29	Total O 29 29	0	0
3	15-A	29	Total O 29 29	0	0
3	16-A	36	Total O 36 36	0	0
3	17-A	33	Total O 33 33	0	0
3	18-A	29	Total O 29 29	0	0
3	19-A	31	Total O 31 31	0	0
3	20-A	29	Total O 29 29	0	0
3	21-A	34	Total O 34 34	0	0

*Continued on next page...*

*Continued from previous page...*

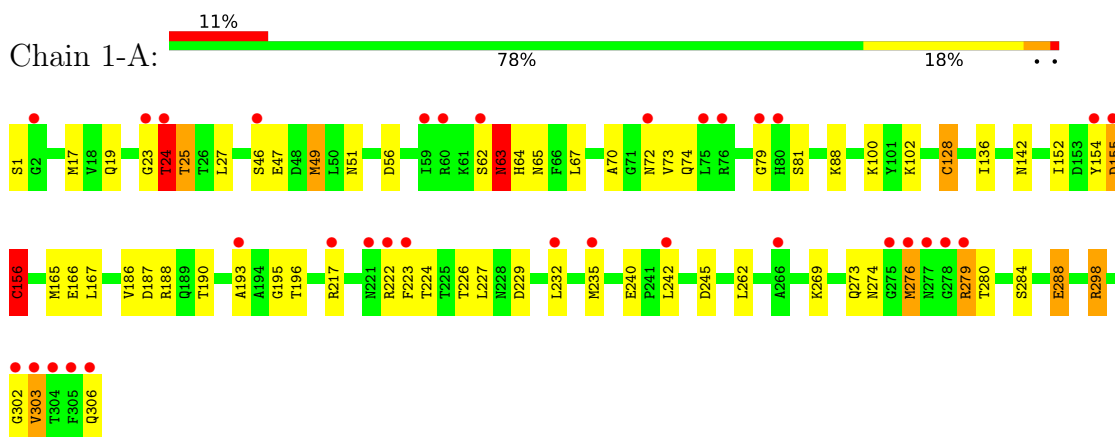
<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>	<b>ZeroOcc</b>	<b>AltConf</b>
3	22-A	37	Total O 37 37	0	0
3	23-A	37	Total O 37 37	0	0
3	24-A	38	Total O 38 38	0	0
3	25-A	40	Total O 40 40	0	0
3	26-A	33	Total O 33 33	0	0
3	27-A	30	Total O 30 30	0	0
3	28-A	31	Total O 31 31	0	0



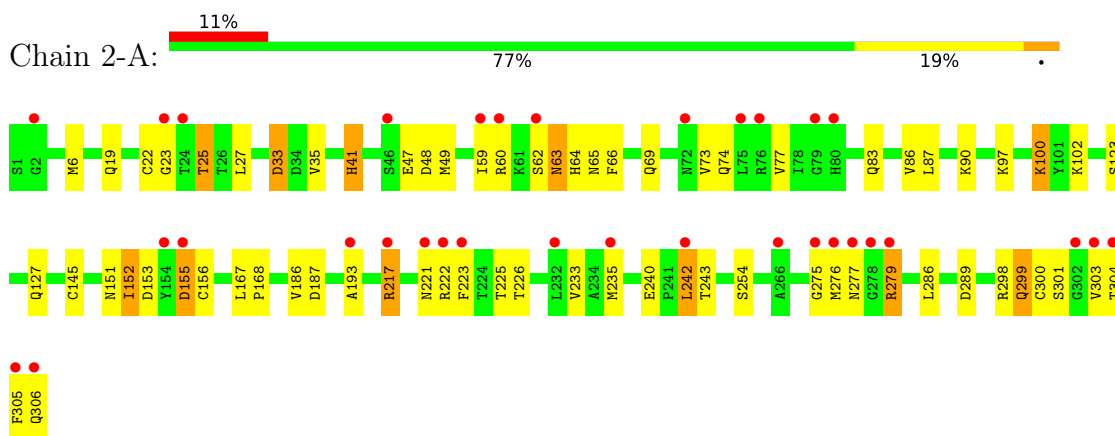
### 3 Residue-property plots [i](#)

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.

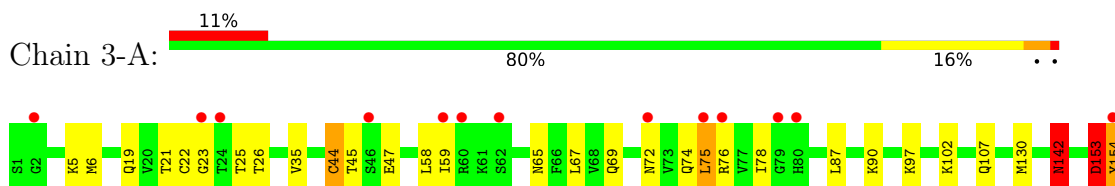
- Molecule 1: 3C-like proteinase

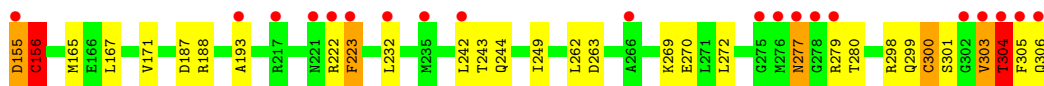


- Molecule 1: 3C-like proteinase

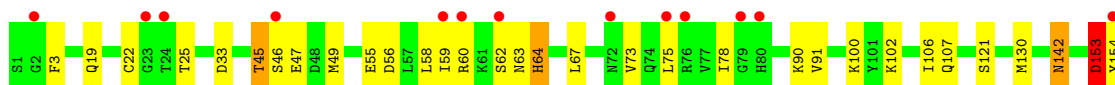
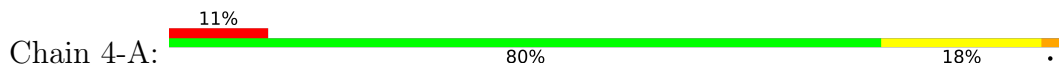


- Molecule 1: 3C-like proteinase

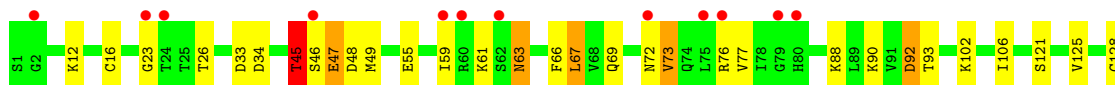
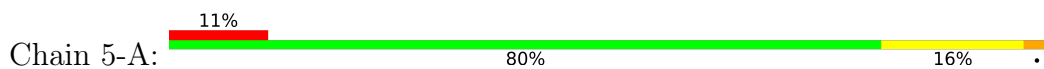




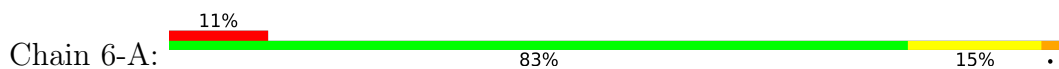
• Molecule 1: 3C-like proteinase



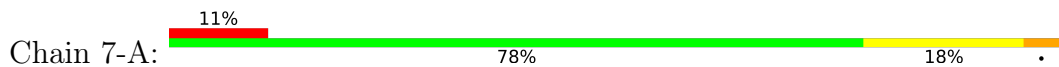
• Molecule 1: 3C-like proteinase

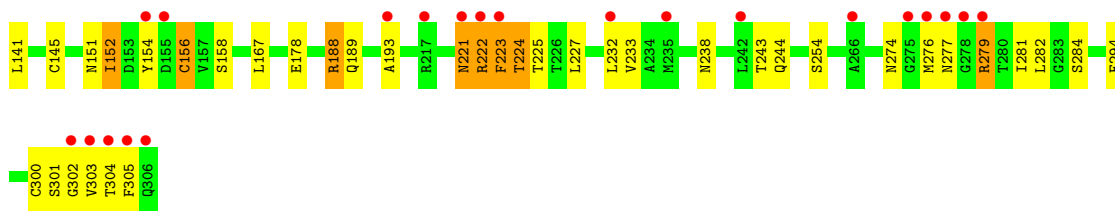


• Molecule 1: 3C-like proteinase

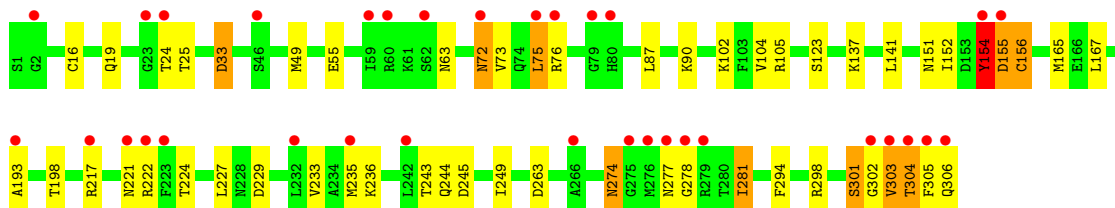
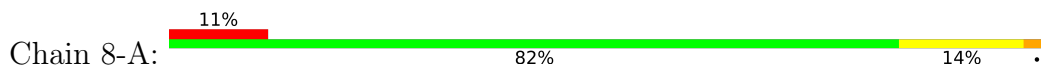


• Molecule 1: 3C-like proteinase

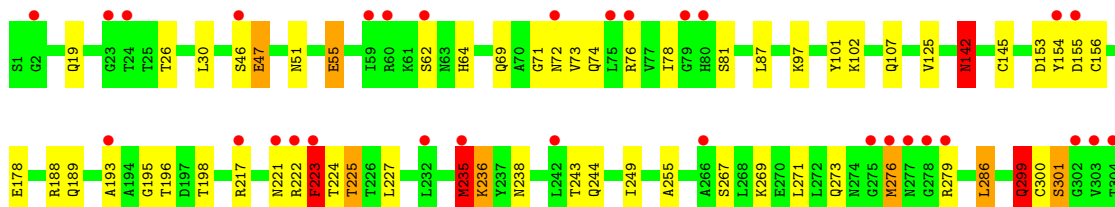
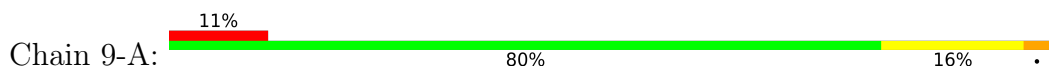




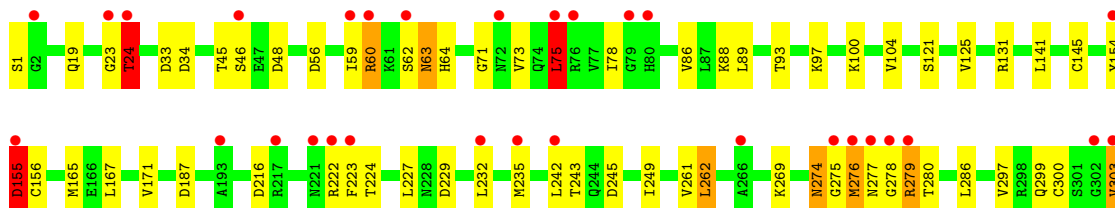
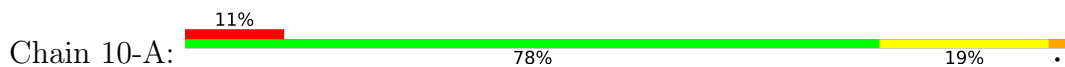
- Molecule 1: 3C-like proteinase



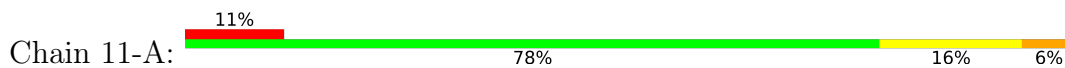
- Molecule 1: 3C-like proteinase

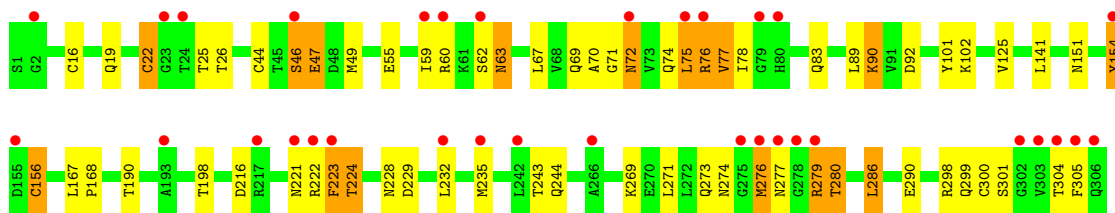


- Molecule 1: 3C-like proteinase

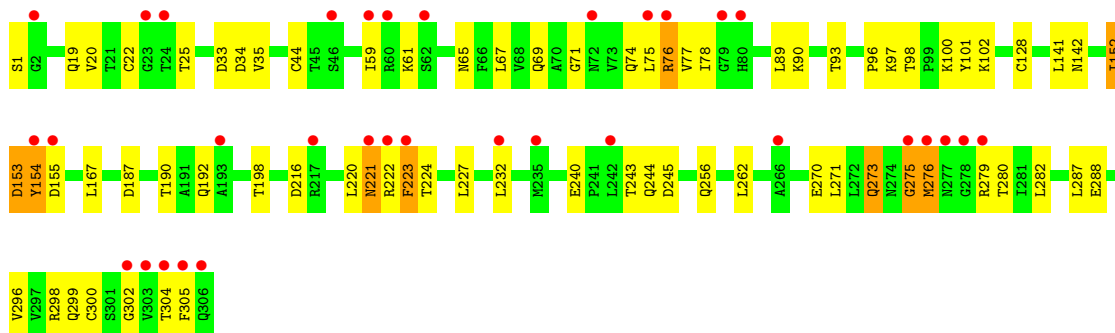
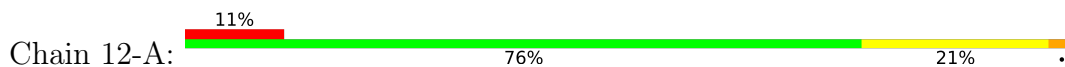


- Molecule 1: 3C-like proteinase

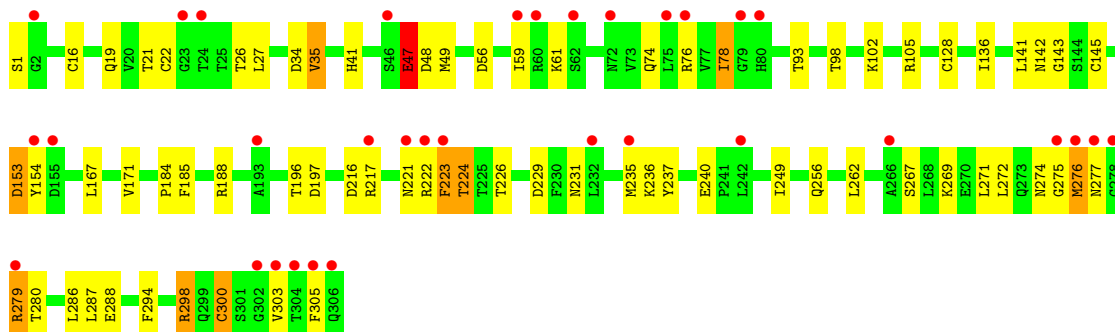
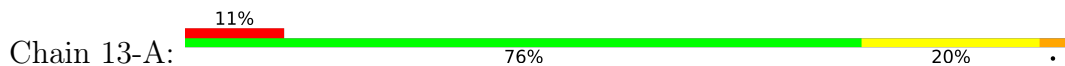




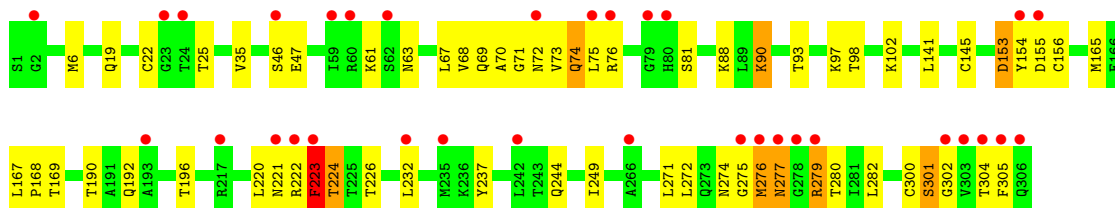
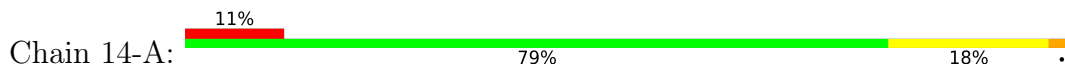
● Molecule 1: 3C-like proteinase



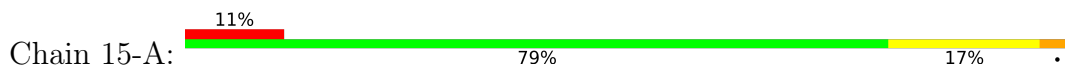
● Molecule 1: 3C-like proteinase

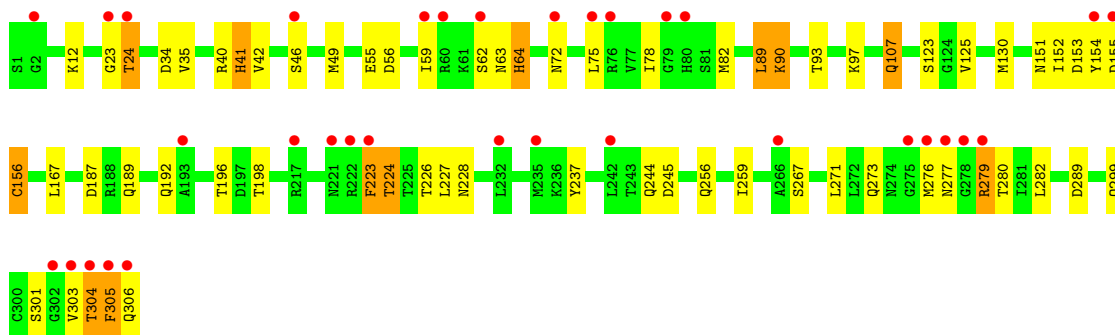


● Molecule 1: 3C-like proteinase

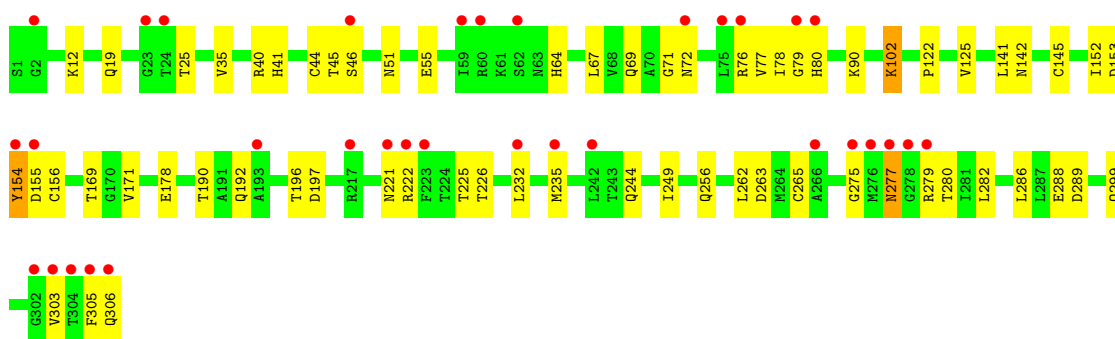
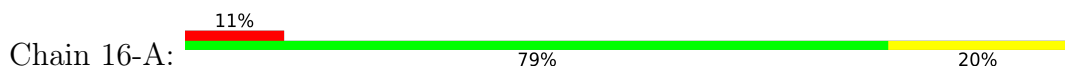


● Molecule 1: 3C-like proteinase

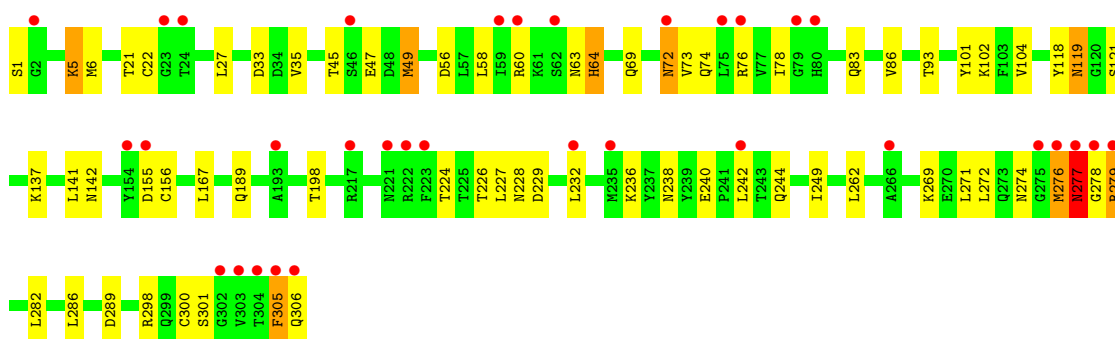
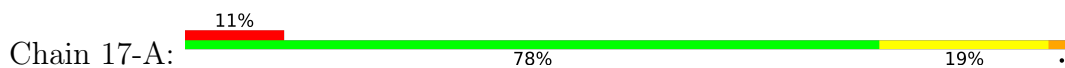




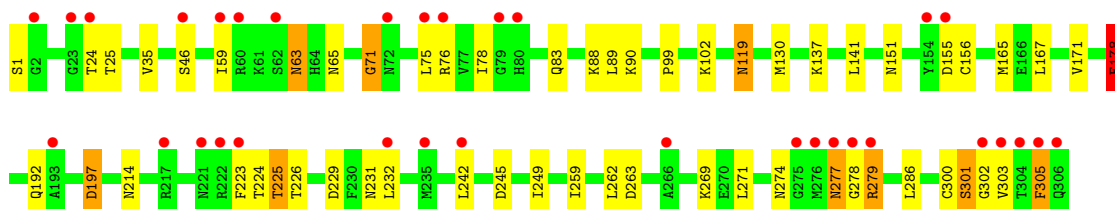
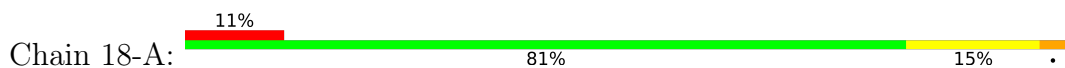
- Molecule 1: 3C-like proteinase



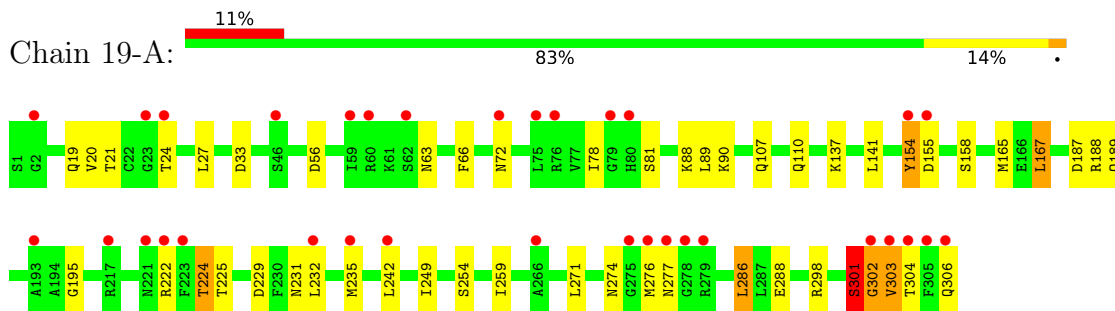
- Molecule 1: 3C-like proteinase



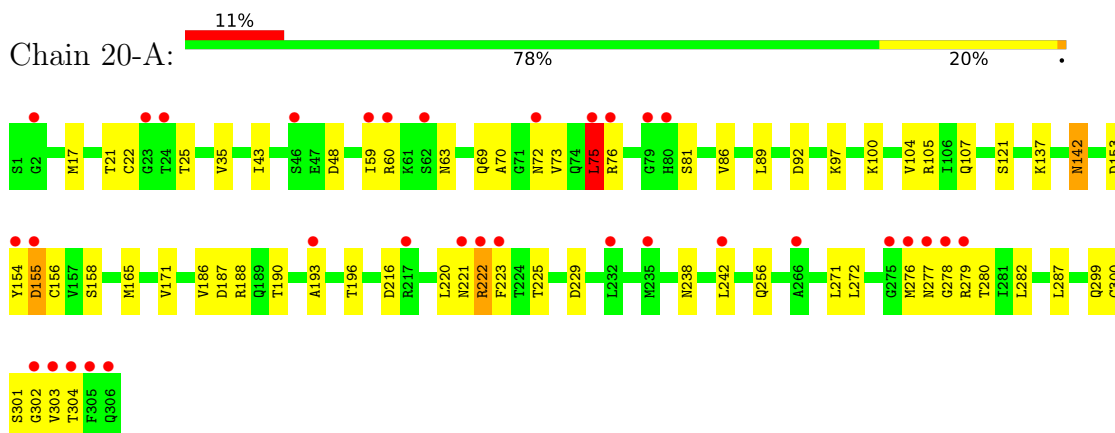
- Molecule 1: 3C-like proteinase



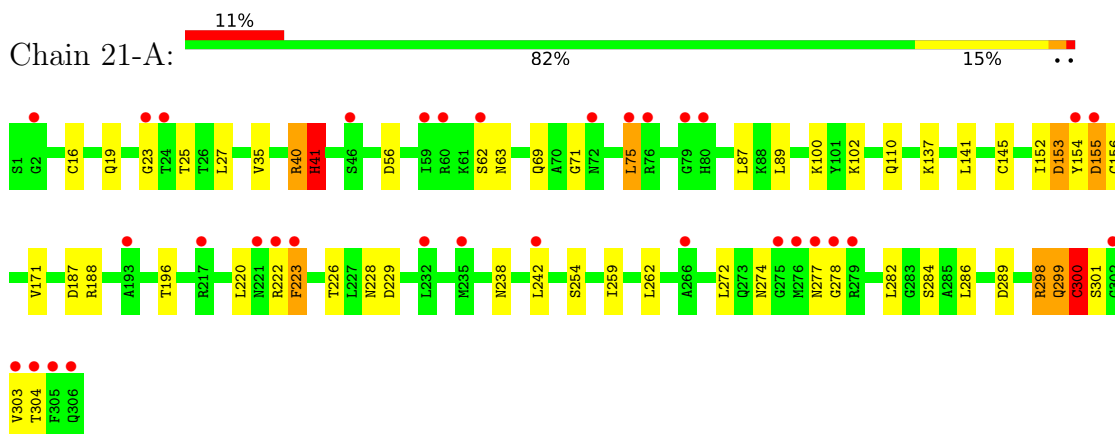
- Molecule 1: 3C-like proteinase



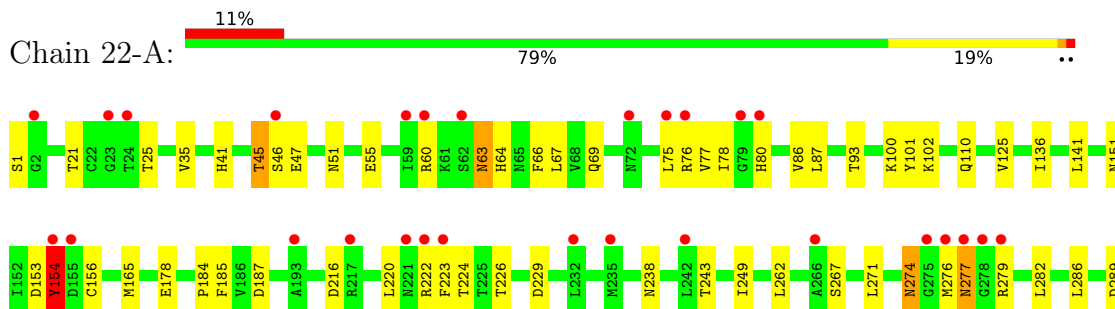
- Molecule 1: 3C-like proteinase

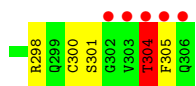


- Molecule 1: 3C-like proteinase

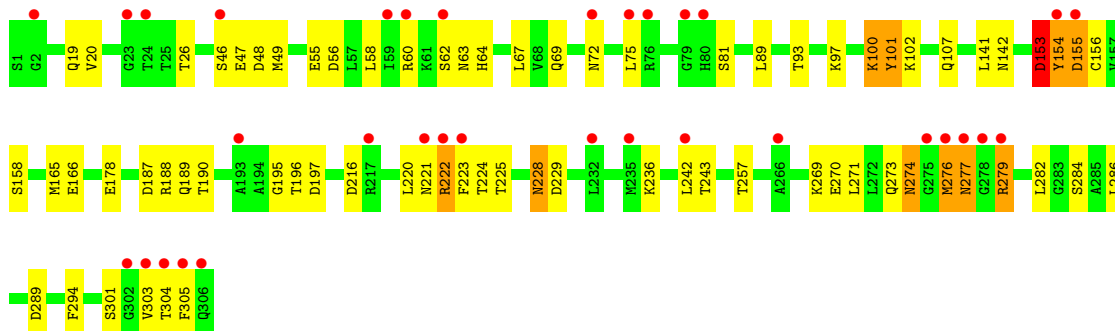
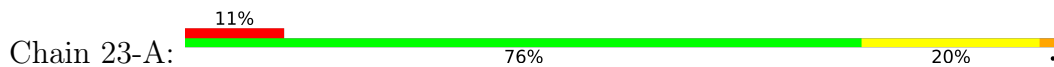


- Molecule 1: 3C-like proteinase

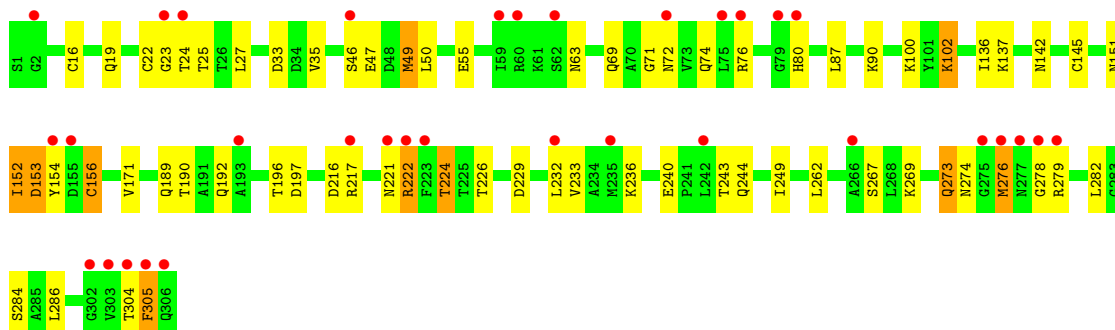
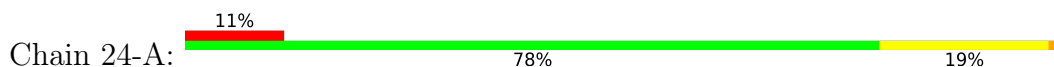




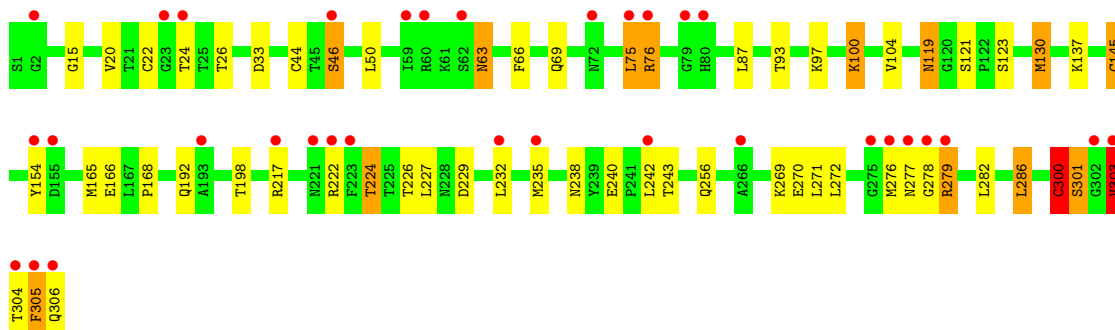
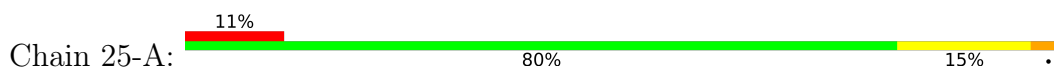
• Molecule 1: 3C-like proteinase



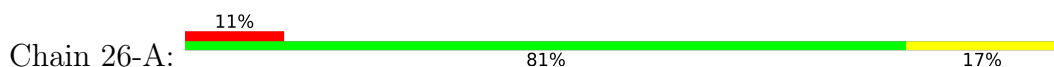
• Molecule 1: 3C-like proteinase

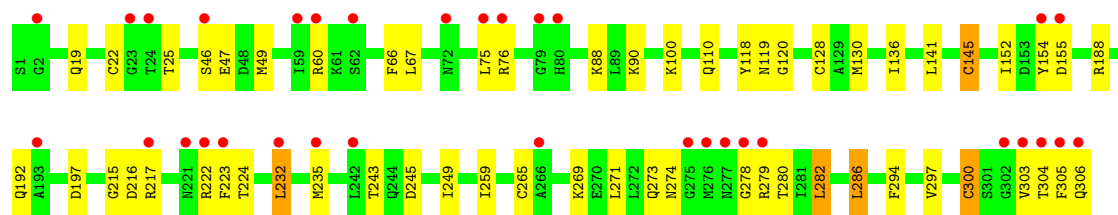


• Molecule 1: 3C-like proteinase

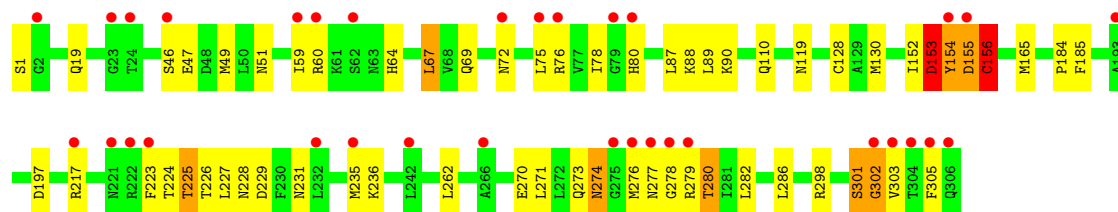
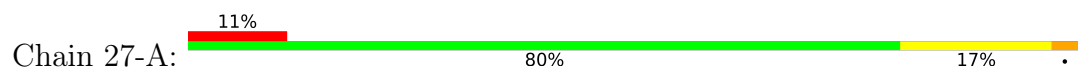


• Molecule 1: 3C-like proteinase

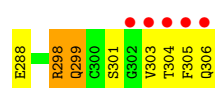
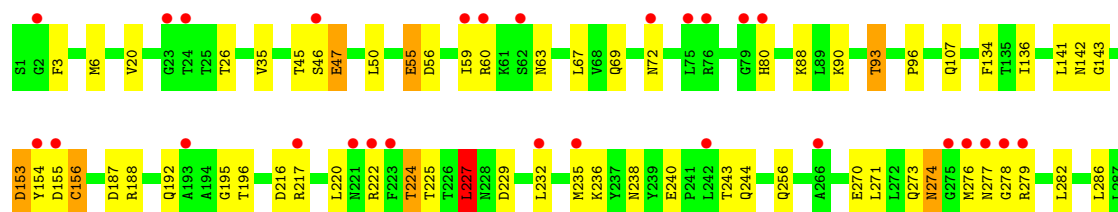
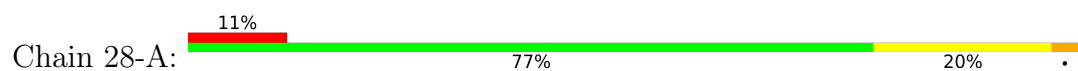




● Molecule 1: 3C-like proteinase



● Molecule 1: 3C-like proteinase





## 4 Data and refinement statistics i

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	114.88Å 54.74Å 45.24Å 90.00° 101.42° 90.00°	Depositor
Resolution (Å)	56.30 – 2.00 56.30 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.0 (56.30-2.00) 88.6 (56.30-2.00)	Depositor EDS
$R_{merge}$	0.18	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.59 (at 2.00Å)	Xtriage
Refinement program	PHENIX (phenix.ensemble_refinement:1.19.2_4158)	Depositor
R, $R_{free}$	0.159 , 0.221 0.188 , 0.266	Depositor DCC
$R_{free}$ test set	924 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	34.3	Xtriage
Anisotropy	0.135	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.55 , 712.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	132023	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.00% 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 i

### 5.1 Standard geometry i

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

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	1-A	0.73	3/2420 (0.1%)	0.83	1/3289 (0.0%)
1	2-A	0.72	1/2420 (0.0%)	0.86	4/3289 (0.1%)
1	3-A	0.69	3/2420 (0.1%)	0.81	3/3289 (0.1%)
1	4-A	0.72	3/2420 (0.1%)	0.86	0/3289
1	5-A	0.74	5/2420 (0.2%)	0.84	1/3289 (0.0%)
1	6-A	0.66	1/2420 (0.0%)	0.80	1/3289 (0.0%)
1	7-A	0.72	2/2420 (0.1%)	0.86	4/3289 (0.1%)
1	8-A	0.69	1/2420 (0.0%)	0.85	1/3289 (0.0%)
1	9-A	0.69	2/2420 (0.1%)	0.84	3/3289 (0.1%)
1	10-A	0.79	4/2420 (0.2%)	0.84	4/3289 (0.1%)
1	11-A	0.76	4/2420 (0.2%)	0.88	3/3289 (0.1%)
1	12-A	0.78	4/2420 (0.2%)	0.83	4/3289 (0.1%)
1	13-A	0.69	2/2420 (0.1%)	0.87	2/3289 (0.1%)
1	14-A	0.70	1/2420 (0.0%)	0.83	0/3289
1	15-A	0.69	2/2420 (0.1%)	0.86	2/3289 (0.1%)
1	16-A	0.72	4/2420 (0.2%)	0.82	2/3289 (0.1%)
1	17-A	0.72	2/2420 (0.1%)	0.84	3/3289 (0.1%)
1	18-A	0.75	2/2420 (0.1%)	0.85	3/3289 (0.1%)
1	19-A	0.68	0/2420	0.85	4/3289 (0.1%)
1	20-A	0.67	1/2420 (0.0%)	0.86	2/3289 (0.1%)
1	21-A	0.75	3/2420 (0.1%)	0.86	9/3289 (0.3%)
1	22-A	0.67	0/2420	0.86	5/3289 (0.2%)
1	23-A	0.71	2/2420 (0.1%)	0.84	3/3289 (0.1%)
1	24-A	0.71	3/2420 (0.1%)	0.89	3/3289 (0.1%)
1	25-A	0.73	3/2420 (0.1%)	0.87	4/3289 (0.1%)
1	26-A	0.75	3/2420 (0.1%)	0.85	4/3289 (0.1%)
1	27-A	0.72	2/2420 (0.1%)	0.86	2/3289 (0.1%)
1	28-A	0.75	4/2420 (0.2%)	0.88	2/3289 (0.1%)
All	All	0.72	67/67760 (0.1%)	0.85	79/92092 (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	1-A	0	11
1	2-A	0	8
1	3-A	0	7
1	4-A	0	4
1	5-A	0	7
1	6-A	0	4
1	7-A	0	5
1	8-A	0	8
1	9-A	0	9
1	10-A	0	7
1	11-A	0	11
1	12-A	0	9
1	13-A	0	14
1	14-A	0	5
1	15-A	0	6
1	16-A	0	4
1	17-A	0	5
1	18-A	0	7
1	19-A	0	5
1	20-A	0	7
1	21-A	0	8
1	22-A	0	5
1	23-A	0	6
1	24-A	0	4
1	25-A	0	8
1	26-A	0	4
1	27-A	0	8
1	28-A	0	7
All	All	0	193

The worst 5 of 67 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	10-A	300	CYS	CB-SG	13.92	2.06	1.82
1	21-A	145	CYS	CB-SG	12.98	2.04	1.82
1	4-A	156	CYS	CB-SG	-12.70	1.60	1.82
1	26-A	145	CYS	CB-SG	-10.89	1.63	1.82
1	17-A	156	CYS	CB-SG	10.27	1.99	1.82

The worst 5 of 79 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	24-A	153	ASP	CB-CG-OD1	8.90	126.31	118.30
1	21-A	145	CYS	CA-CB-SG	8.29	128.93	114.00
1	25-A	303	VAL	CA-CB-CG1	7.65	122.37	110.90
1	25-A	286	LEU	CA-CB-CG	7.29	132.06	115.30
1	19-A	167	LEU	CA-CB-CG	7.15	131.74	115.30

There are no chirality outliers.

5 of 193 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	1-A	23	GLY	Peptide
1	1-A	24	THR	Peptide
1	1-A	62	SER	Peptide
1	1-A	63	ASN	Peptide
1	1-A	79	GLY	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	1-A	2367	2313	2313	0	0
1	2-A	2367	2313	2313	0	0
1	3-A	2367	2313	2313	0	0
1	4-A	2367	2313	2313	0	0
1	5-A	2367	2313	2313	0	0
1	6-A	2367	2313	2313	0	0
1	7-A	2367	2313	2313	0	0
1	8-A	2367	2313	2313	0	0
1	9-A	2367	2313	2313	0	0
1	10-A	2367	2313	2313	0	0
1	11-A	2367	2313	2313	0	0
1	12-A	2367	2313	2313	0	0
1	13-A	2367	2313	2313	0	0
1	14-A	2367	2313	2313	0	0
1	15-A	2367	2313	2313	0	0
1	16-A	2367	2313	2313	0	0
1	17-A	2367	2313	2313	0	0
1	18-A	2367	2313	2313	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	19-A	2367	2313	2313	0	0
1	20-A	2367	2313	2313	0	0
1	21-A	2367	2313	2313	0	0
1	22-A	2367	2313	2313	0	0
1	23-A	2367	2313	2313	0	0
1	24-A	2367	2313	2313	0	0
1	25-A	2367	2313	2313	0	0
1	26-A	2367	2313	2313	0	0
1	27-A	2367	2313	2313	0	0
1	28-A	2367	2313	2313	0	0
2	1-A	1	0	0	0	0
2	2-A	1	0	0	0	0
2	3-A	1	0	0	0	0
2	4-A	1	0	0	0	0
2	5-A	1	0	0	0	0
2	6-A	1	0	0	0	0
2	7-A	1	0	0	0	0
2	8-A	1	0	0	0	0
2	9-A	1	0	0	0	0
2	10-A	1	0	0	0	0
2	11-A	1	0	0	0	0
2	12-A	1	0	0	0	0
2	13-A	1	0	0	0	0
2	14-A	1	0	0	0	0
2	15-A	1	0	0	0	0
2	16-A	1	0	0	0	0
2	17-A	1	0	0	0	0
2	18-A	1	0	0	0	0
2	19-A	1	0	0	0	0
2	20-A	1	0	0	0	0
2	21-A	1	0	0	0	0
2	22-A	1	0	0	0	0
2	23-A	1	0	0	0	0
2	24-A	1	0	0	0	0
2	25-A	1	0	0	0	0
2	26-A	1	0	0	0	0
2	27-A	1	0	0	0	0
2	28-A	1	0	0	0	0
3	1-A	34	0	0	0	0
3	2-A	29	0	0	0	0
3	3-A	24	0	0	0	0
3	4-A	37	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	5-A	46	0	0	0	0
3	6-A	30	0	0	0	0
3	7-A	45	0	0	0	0
3	8-A	30	0	0	0	0
3	9-A	35	0	0	0	0
3	10-A	38	0	0	0	0
3	11-A	41	0	0	0	0
3	12-A	34	0	0	0	0
3	13-A	36	0	0	0	0
3	14-A	29	0	0	0	0
3	15-A	29	0	0	0	0
3	16-A	36	0	0	0	0
3	17-A	33	0	0	0	0
3	18-A	29	0	0	0	0
3	19-A	31	0	0	0	0
3	20-A	29	0	0	0	0
3	21-A	34	0	0	0	0
3	22-A	37	0	0	0	0
3	23-A	37	0	0	0	0
3	24-A	38	0	0	0	0
3	25-A	40	0	0	0	0
3	26-A	33	0	0	0	0
3	27-A	30	0	0	0	0
3	28-A	31	0	0	0	0
All	All	67259	64764	64764	0	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). Clashscore could not be calculated for this entry.

There are no clashes within the asymmetric unit.

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	1-A	304/306 (99%)	253 (83%)	35 (12%)	16 (5%)	2	0
1	2-A	304/306 (99%)	246 (81%)	36 (12%)	22 (7%)	1	0
1	3-A	304/306 (99%)	257 (84%)	29 (10%)	18 (6%)	1	0
1	4-A	304/306 (99%)	262 (86%)	25 (8%)	17 (6%)	2	0
1	5-A	304/306 (99%)	256 (84%)	33 (11%)	15 (5%)	2	0
1	6-A	304/306 (99%)	261 (86%)	28 (9%)	15 (5%)	2	0
1	7-A	304/306 (99%)	255 (84%)	28 (9%)	21 (7%)	1	0
1	8-A	304/306 (99%)	253 (83%)	36 (12%)	15 (5%)	2	0
1	9-A	304/306 (99%)	243 (80%)	42 (14%)	19 (6%)	1	0
1	10-A	304/306 (99%)	256 (84%)	26 (9%)	22 (7%)	1	0
1	11-A	304/306 (99%)	251 (83%)	34 (11%)	19 (6%)	1	0
1	12-A	304/306 (99%)	258 (85%)	26 (9%)	20 (7%)	1	0
1	13-A	304/306 (99%)	259 (85%)	32 (10%)	13 (4%)	2	0
1	14-A	304/306 (99%)	260 (86%)	23 (8%)	21 (7%)	1	0
1	15-A	304/306 (99%)	258 (85%)	26 (9%)	20 (7%)	1	0
1	16-A	304/306 (99%)	262 (86%)	30 (10%)	12 (4%)	3	1
1	17-A	304/306 (99%)	253 (83%)	37 (12%)	14 (5%)	2	0
1	18-A	304/306 (99%)	266 (88%)	25 (8%)	13 (4%)	2	0
1	19-A	304/306 (99%)	268 (88%)	30 (10%)	6 (2%)	7	3
1	20-A	304/306 (99%)	267 (88%)	24 (8%)	13 (4%)	2	0
1	21-A	304/306 (99%)	265 (87%)	28 (9%)	11 (4%)	3	1
1	22-A	304/306 (99%)	258 (85%)	33 (11%)	13 (4%)	2	0
1	23-A	304/306 (99%)	263 (86%)	25 (8%)	16 (5%)	2	0
1	24-A	304/306 (99%)	250 (82%)	39 (13%)	15 (5%)	2	0
1	25-A	304/306 (99%)	261 (86%)	29 (10%)	14 (5%)	2	0
1	26-A	304/306 (99%)	263 (86%)	27 (9%)	14 (5%)	2	0
1	27-A	304/306 (99%)	262 (86%)	26 (9%)	16 (5%)	2	0
1	28-A	304/306 (99%)	248 (82%)	39 (13%)	17 (6%)	2	0
All	All	8512/8568 (99%)	7214 (85%)	851 (10%)	447 (5%)	2	0

5 of 447 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	1-A	25	THR
1	1-A	63	ASN
1	1-A	193	ALA
1	1-A	274	ASN
1	1-A	276	MET

### 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	1-A	263/263 (100%)	211 (80%)	52 (20%)	1	0
1	2-A	263/263 (100%)	218 (83%)	45 (17%)	2	1
1	3-A	263/263 (100%)	216 (82%)	47 (18%)	2	1
1	4-A	263/263 (100%)	216 (82%)	47 (18%)	2	1
1	5-A	263/263 (100%)	219 (83%)	44 (17%)	2	1
1	6-A	263/263 (100%)	225 (86%)	38 (14%)	3	1
1	7-A	263/263 (100%)	217 (82%)	46 (18%)	2	1
1	8-A	263/263 (100%)	221 (84%)	42 (16%)	2	1
1	9-A	263/263 (100%)	218 (83%)	45 (17%)	2	1
1	10-A	263/263 (100%)	219 (83%)	44 (17%)	2	1
1	11-A	263/263 (100%)	217 (82%)	46 (18%)	2	1
1	12-A	263/263 (100%)	218 (83%)	45 (17%)	2	1
1	13-A	263/263 (100%)	210 (80%)	53 (20%)	1	0
1	14-A	263/263 (100%)	217 (82%)	46 (18%)	2	1
1	15-A	263/263 (100%)	216 (82%)	47 (18%)	2	1
1	16-A	263/263 (100%)	218 (83%)	45 (17%)	2	1
1	17-A	263/263 (100%)	208 (79%)	55 (21%)	1	0
1	18-A	263/263 (100%)	219 (83%)	44 (17%)	2	1
1	19-A	263/263 (100%)	219 (83%)	44 (17%)	2	1
1	20-A	263/263 (100%)	215 (82%)	48 (18%)	1	1

*Continued on next page...*



Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	21-A	263/263 (100%)	225 (86%)	38 (14%)	3	1
1	22-A	263/263 (100%)	211 (80%)	52 (20%)	1	0
1	23-A	263/263 (100%)	204 (78%)	59 (22%)	1	0
1	24-A	263/263 (100%)	211 (80%)	52 (20%)	1	0
1	25-A	263/263 (100%)	213 (81%)	50 (19%)	1	0
1	26-A	263/263 (100%)	223 (85%)	40 (15%)	3	1
1	27-A	263/263 (100%)	217 (82%)	46 (18%)	2	1
1	28-A	263/263 (100%)	210 (80%)	53 (20%)	1	0
All	All	7364/7364 (100%)	6051 (82%)	1313 (18%)	2	1

5 of 1313 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	21-A	35	VAL
1	25-A	192	GLN
1	21-A	272	LEU
1	21-A	27	LEU
1	23-A	158	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 230 such sidechains are listed below:

Mol	Chain	Res	Type
1	14-A	69	GLN
1	28-A	69	GLN
1	17-A	273	GLN
1	28-A	63	ASN
1	25-A	256	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 28 ligands modelled in this entry, 28 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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	1-A	306/306 (100%)	0.72	33 (10%) 5 5	39, 44, 52, 56	306 (100%)
1	2-A	306/306 (100%)	0.72	33 (10%) 5 5	39, 44, 52, 56	306 (100%)
1	3-A	306/306 (100%)	0.72	33 (10%) 5 5	39, 44, 52, 56	306 (100%)
1	4-A	306/306 (100%)	0.72	33 (10%) 5 5	39, 44, 52, 56	306 (100%)
1	5-A	306/306 (100%)	0.72	33 (10%) 5 5	39, 44, 52, 56	306 (100%)
1	6-A	306/306 (100%)	0.72	33 (10%) 5 5	39, 44, 52, 56	306 (100%)
1	7-A	306/306 (100%)	0.72	33 (10%) 5 5	39, 44, 52, 56	306 (100%)
1	8-A	306/306 (100%)	0.72	33 (10%) 5 5	39, 44, 52, 56	306 (100%)
1	9-A	306/306 (100%)	0.72	33 (10%) 5 5	39, 44, 52, 56	306 (100%)
1	10-A	306/306 (100%)	0.72	33 (10%) 5 5	39, 44, 52, 56	306 (100%)
1	11-A	306/306 (100%)	0.72	33 (10%) 5 5	39, 44, 52, 56	306 (100%)
1	12-A	306/306 (100%)	0.72	33 (10%) 5 5	39, 44, 52, 56	306 (100%)
1	13-A	306/306 (100%)	0.72	33 (10%) 5 5	39, 44, 52, 56	306 (100%)
1	14-A	306/306 (100%)	0.72	33 (10%) 5 5	39, 44, 52, 56	306 (100%)
1	15-A	306/306 (100%)	0.72	33 (10%) 5 5	39, 44, 52, 56	306 (100%)
1	16-A	306/306 (100%)	0.72	33 (10%) 5 5	39, 44, 52, 56	306 (100%)
1	17-A	306/306 (100%)	0.72	33 (10%) 5 5	39, 44, 52, 56	306 (100%)
1	18-A	306/306 (100%)	0.72	33 (10%) 5 5	39, 44, 52, 56	306 (100%)
1	19-A	306/306 (100%)	0.72	33 (10%) 5 5	39, 44, 52, 56	306 (100%)
1	20-A	306/306 (100%)	0.72	33 (10%) 5 5	39, 44, 52, 56	306 (100%)
1	21-A	306/306 (100%)	0.72	33 (10%) 5 5	39, 44, 52, 56	306 (100%)
1	22-A	306/306 (100%)	0.72	33 (10%) 5 5	39, 44, 52, 56	306 (100%)
1	23-A	306/306 (100%)	0.72	33 (10%) 5 5	39, 44, 52, 56	306 (100%)
1	24-A	306/306 (100%)	0.72	33 (10%) 5 5	39, 44, 52, 56	306 (100%)

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	25-A	306/306 (100%)	0.72	33 (10%) 5 5	39, 44, 52, 56	306 (100%)
1	26-A	306/306 (100%)	0.72	33 (10%) 5 5	39, 44, 52, 56	306 (100%)
1	27-A	306/306 (100%)	0.72	33 (10%) 5 5	39, 44, 52, 56	306 (100%)
1	28-A	306/306 (100%)	0.72	33 (10%) 5 5	39, 44, 52, 56	306 (100%)
All	All	8568/8568 (100%)	0.72	924 (10%) 5 5	39, 44, 52, 56	8568 (100%)

The worst 5 of 924 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	1-A	303	VAL	35.0
1	2-A	303	VAL	35.0
1	3-A	303	VAL	35.0
1	4-A	303	VAL	35.0
1	5-A	303	VAL	35.0

## 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
2	ZN	1-A	401	1/1	0.91	0.35	42,42,42,42	1
2	ZN	2-A	401	1/1	0.91	0.35	42,42,42,42	1
2	ZN	3-A	401	1/1	0.91	0.35	42,42,42,42	1
2	ZN	4-A	401	1/1	0.91	0.35	42,42,42,42	1
2	ZN	5-A	401	1/1	0.91	0.35	42,42,42,42	1
2	ZN	6-A	401	1/1	0.91	0.35	42,42,42,42	1
2	ZN	7-A	401	1/1	0.91	0.35	42,42,42,42	1

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
2	ZN	8-A	401	1/1	0.91	0.35	42,42,42,42	1
2	ZN	9-A	401	1/1	0.91	0.35	42,42,42,42	1
2	ZN	10-A	401	1/1	0.91	0.35	42,42,42,42	1
2	ZN	11-A	401	1/1	0.91	0.35	42,42,42,42	1
2	ZN	12-A	401	1/1	0.91	0.35	42,42,42,42	1
2	ZN	13-A	401	1/1	0.91	0.35	42,42,42,42	1
2	ZN	14-A	401	1/1	0.91	0.35	42,42,42,42	1
2	ZN	15-A	401	1/1	0.91	0.35	42,42,42,42	1
2	ZN	16-A	401	1/1	0.91	0.35	42,42,42,42	1
2	ZN	17-A	401	1/1	0.91	0.35	42,42,42,42	1
2	ZN	18-A	401	1/1	0.91	0.35	42,42,42,42	1
2	ZN	19-A	401	1/1	0.91	0.35	42,42,42,42	1
2	ZN	20-A	401	1/1	0.91	0.35	42,42,42,42	1
2	ZN	21-A	401	1/1	0.91	0.35	42,42,42,42	1
2	ZN	22-A	401	1/1	0.91	0.35	42,42,42,42	1
2	ZN	23-A	401	1/1	0.91	0.35	42,42,42,42	1
2	ZN	24-A	401	1/1	0.91	0.35	42,42,42,42	1
2	ZN	25-A	401	1/1	0.91	0.35	42,42,42,42	1
2	ZN	26-A	401	1/1	0.91	0.35	42,42,42,42	1
2	ZN	27-A	401	1/1	0.91	0.35	42,42,42,42	1
2	ZN	28-A	401	1/1	0.91	0.35	42,42,42,42	1

## 6.5 Other polymers [i](#)

There are no such residues in this entry.