



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 15, 2024 – 04:57 PM EDT

PDB ID : 4U7U  
Title : Crystal structure of RNA-guided immune Cascade complex from E.coli  
Authors : Zhao, H.; Sheng, G.; Wang, J.; Wang, M.; Bunkoczi, G.; Gong, W.; Wei, Z.; Wang, Y.  
Deposited on : 2014-07-31  
Resolution : 3.00 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	:	2.37.1
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)	:	2.37.1

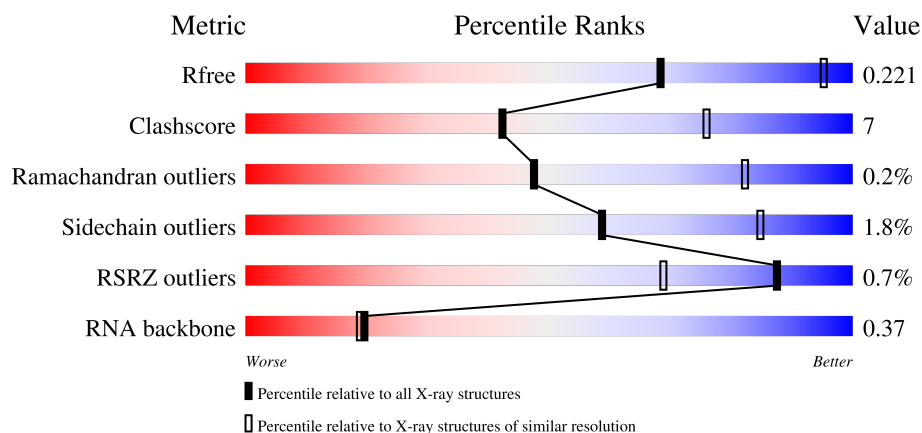
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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.








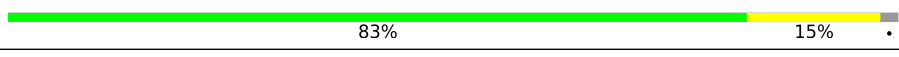

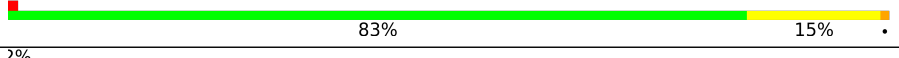
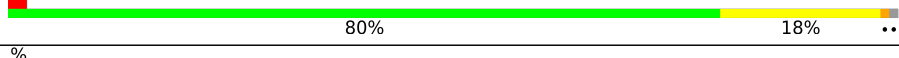





Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)
RNA backbone	3102	1173 (3.30-2.70)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments 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	A	502	
1	M	502	
2	B	160	
2	C	160	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	N	160	
2	O	160	
3	D	201	
3	P	201	
4	E	363	
4	F	363	
4	G	363	
4	H	363	
4	I	363	
4	J	363	
4	Q	363	
4	R	363	
4	S	363	
4	T	363	
4	U	363	
4	V	363	
5	K	224	
5	W	224	
6	L	61	
6	X	61	

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 54555 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 CRISPR system Cascade subunit CasA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	480	Total	C	N	O	S	0	0	0
			3739	2386	659	674	20			
1	M	489	Total	C	N	O	S	0	0	0
			3812	2428	673	691	20			

- Molecule 2 is a protein called CRISPR system Cascade subunit CasB.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	157	Total	C	N	O	S	0	0	0
			1291	808	249	227	7			
2	C	157	Total	C	N	O	S	0	0	0
			1292	809	249	227	7			
2	N	152	Total	C	N	O	S	0	0	0
			1246	784	240	215	7			
2	O	156	Total	C	N	O	S	0	0	0
			1247	785	234	221	7			

- Molecule 3 is a protein called CRISPR system Cascade subunit CasE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	191	Total	C	N	O	S	0	0	0
			1451	931	255	258	7			
3	P	191	Total	C	N	O	S	0	0	0
			1459	936	258	258	7			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	-1	ALA	-	expression tag	UNP Q46897
D	0	TRP	-	expression tag	UNP Q46897
P	-1	ALA	-	expression tag	UNP Q46897
P	0	TRP	-	expression tag	UNP Q46897

- Molecule 4 is a protein called CRISPR system Cascade subunit CasC.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	357	Total	C	N	O	S	0	0	0
			2752	1721	490	525	16			
4	F	359	Total	C	N	O	S	0	0	0
			2760	1727	489	528	16			
4	G	356	Total	C	N	O	S	0	0	0
			2753	1720	490	527	16			
4	H	363	Total	C	N	O	S	0	0	0
			2774	1736	493	529	16			
4	I	355	Total	C	N	O	S	0	0	0
			2731	1711	485	519	16			
4	J	351	Total	C	N	O	S	0	0	0
			2680	1678	473	513	16			
4	Q	362	Total	C	N	O	S	0	0	0
			2790	1744	495	535	16			
4	R	359	Total	C	N	O	S	0	0	0
			2736	1712	484	524	16			
4	S	361	Total	C	N	O	S	0	0	0
			2773	1734	494	529	16			
4	T	360	Total	C	N	O	S	0	0	0
			2777	1738	494	529	16			
4	U	354	Total	C	N	O	S	0	0	0
			2723	1704	482	521	16			
4	V	355	Total	C	N	O	S	0	0	0
			2725	1704	482	523	16			

- Molecule 5 is a protein called CRISPR system Cascade subunit CasD.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	K	218	Total	C	N	O	S	0	0	0
			1722	1090	305	318	9			
5	W	218	Total	C	N	O	S	0	0	0
			1726	1093	306	318	9			

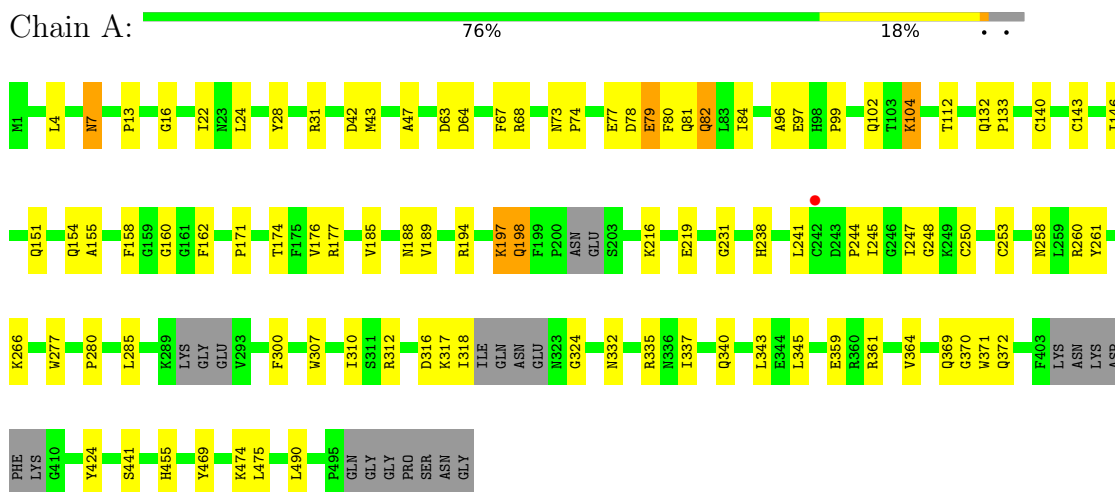
- Molecule 6 is a RNA chain called crRNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	L	61	Total	C	N	O	P	0	0	0
			1298	580	230	428	60			
6	X	61	Total	C	N	O	P	0	0	0
			1298	580	230	428	60			

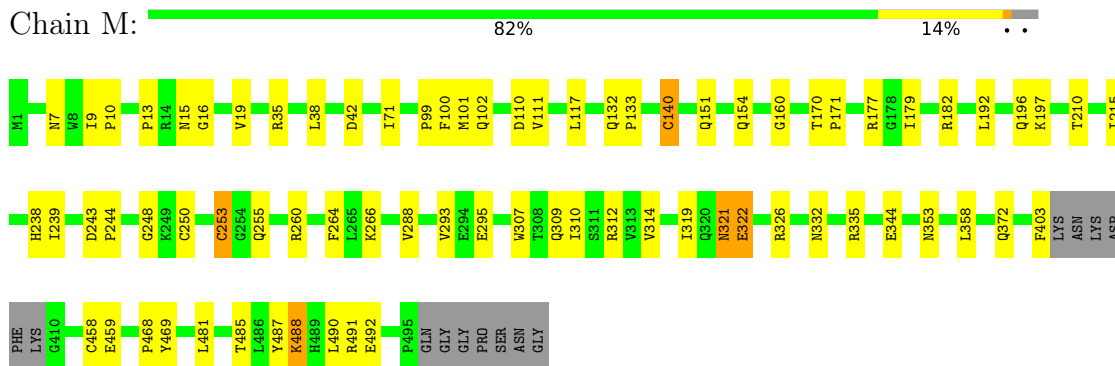
### 3 Residue-property plots

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.

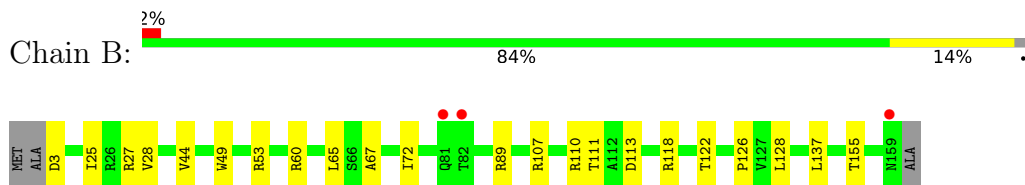
#### • Molecule 1: CRISPR system Cascade subunit CasA



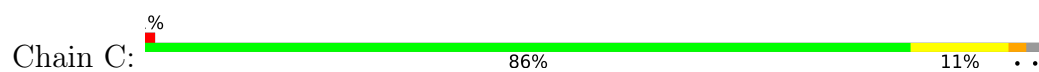
#### • Molecule 1: CRISPR system Cascade subunit CasA



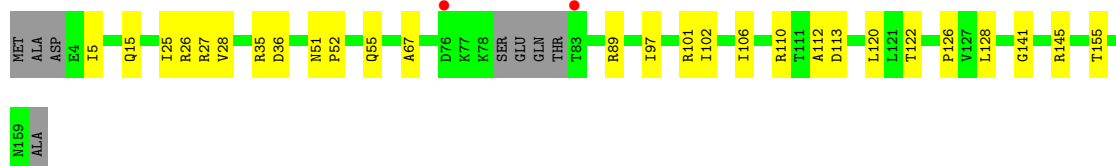
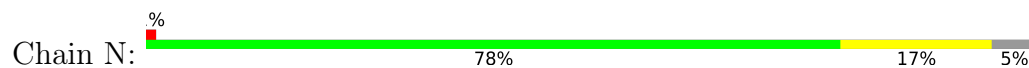
#### • Molecule 2: CRISPR system Cascade subunit CasB



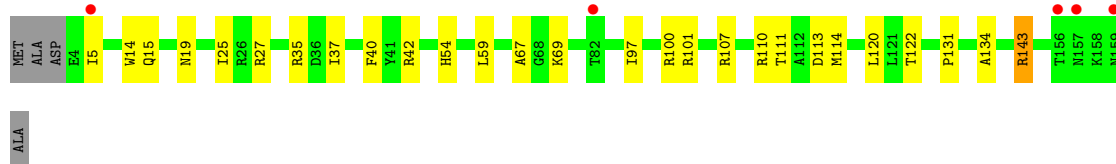
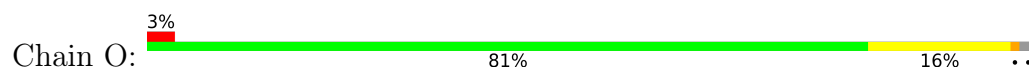
#### • Molecule 2: CRISPR system Cascade subunit CasB



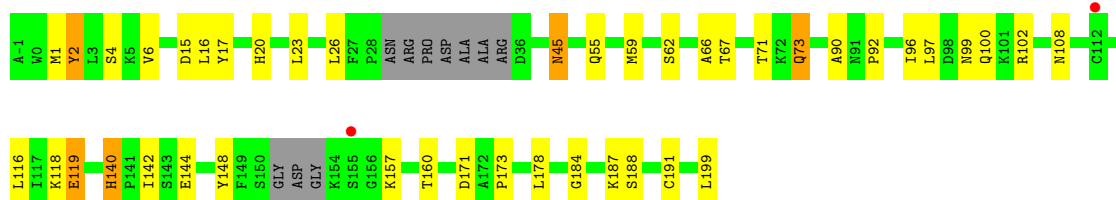
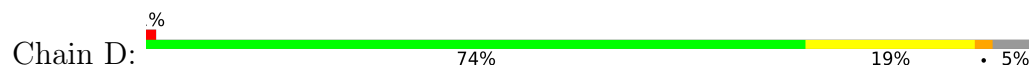
• Molecule 2: CRISPR system Cascade subunit CasB



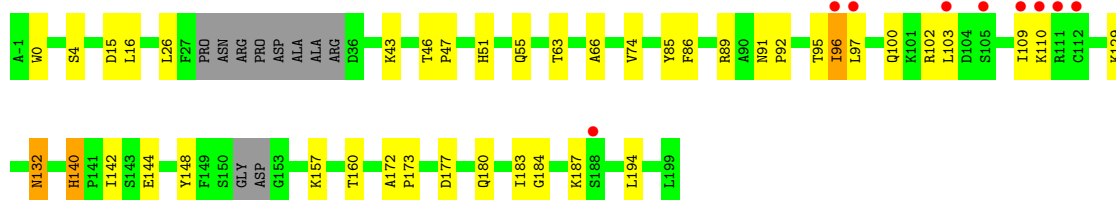
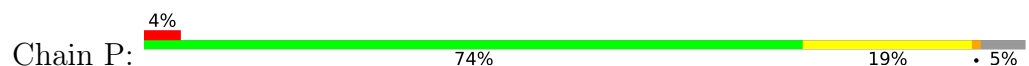
• Molecule 2: CRISPR system Cascade subunit CasB



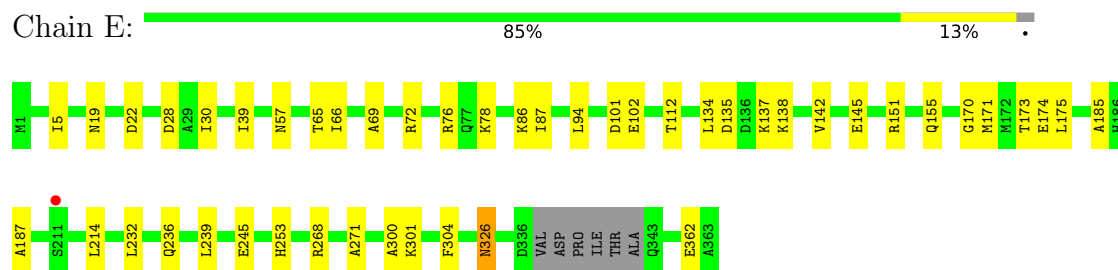
• Molecule 3: CRISPR system Cascade subunit CasE



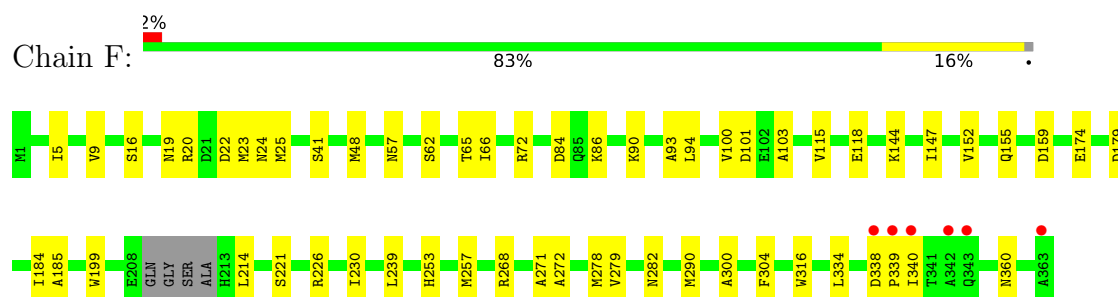
• Molecule 3: CRISPR system Cascade subunit CasE



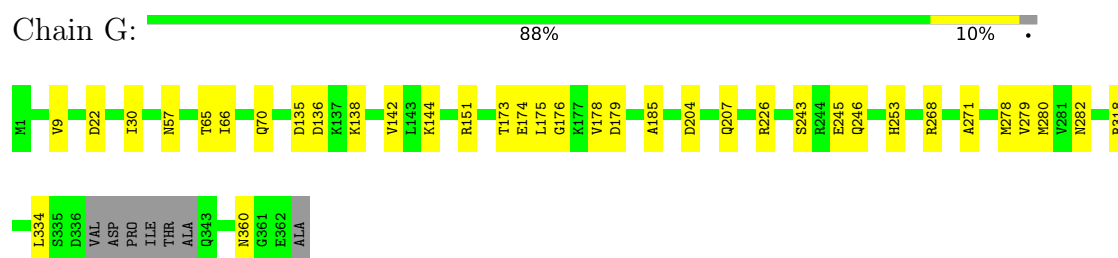
• Molecule 4: CRISPR system Cascade subunit CasC



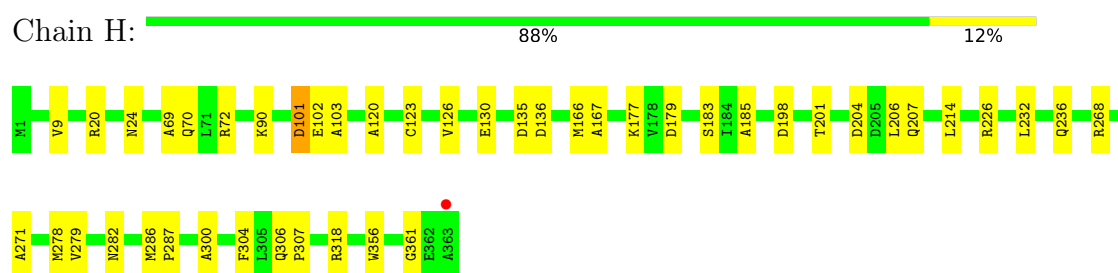
• Molecule 4: CRISPR system Cascade subunit CasC



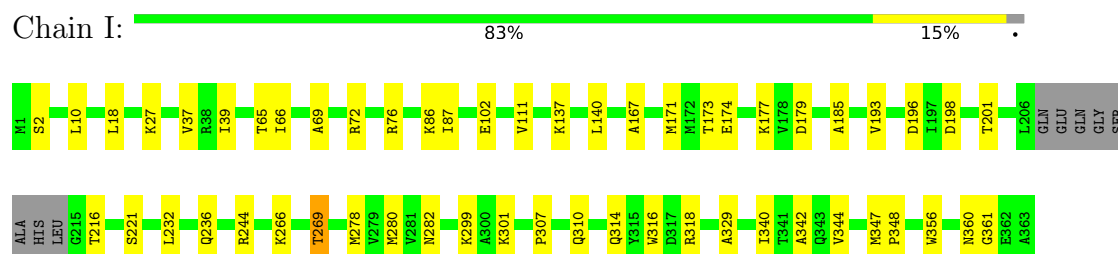
• Molecule 4: CRISPR system Cascade subunit CasC



• Molecule 4: CRISPR system Cascade subunit CasC

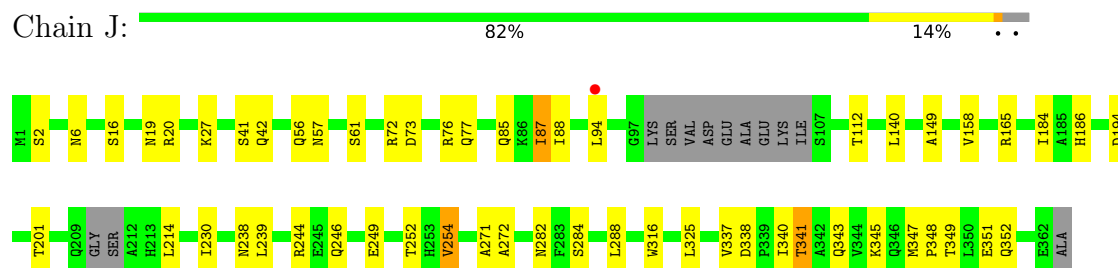


• Molecule 4: CRISPR system Cascade subunit CasC

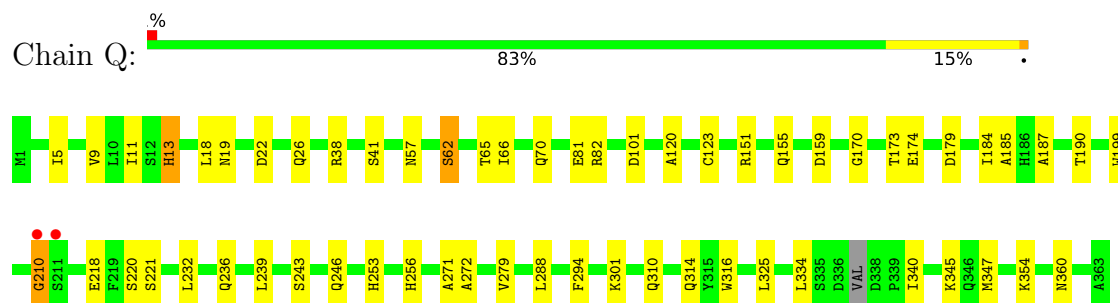




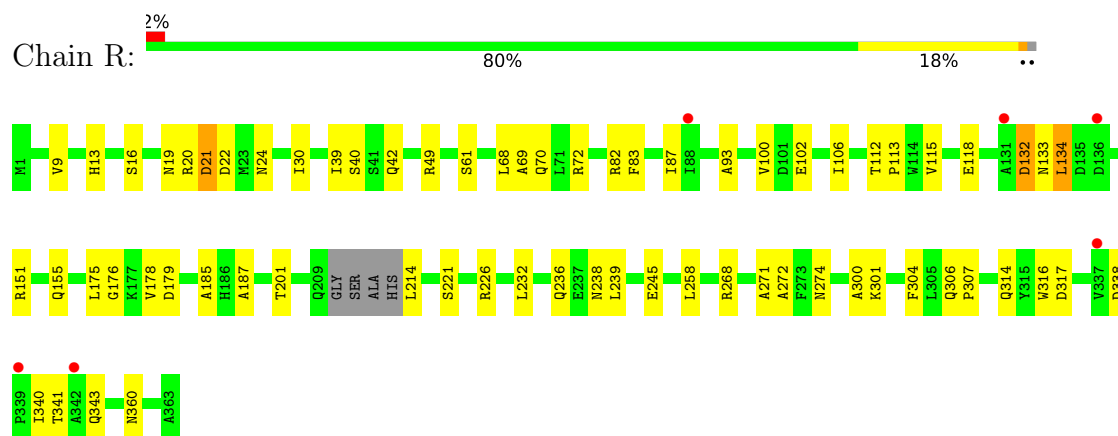
- Molecule 4: CRISPR system Cascade subunit CasC



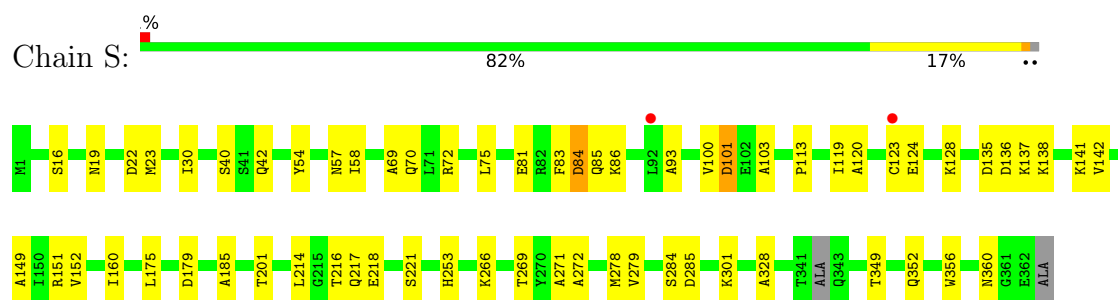
- Molecule 4: CRISPR system Cascade subunit CasC



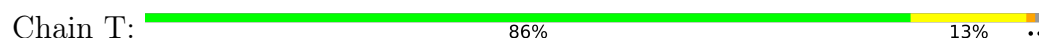
- Molecule 4: CRISPR system Cascade subunit CasC

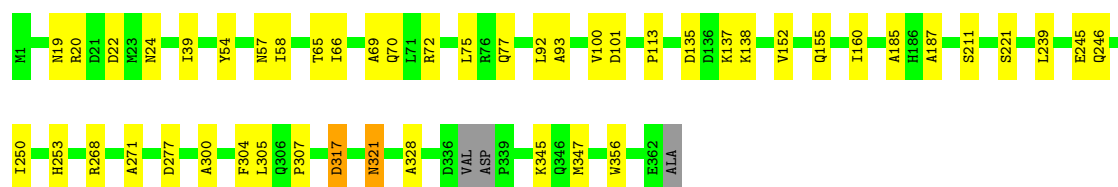


- Molecule 4: CRISPR system Cascade subunit CasC



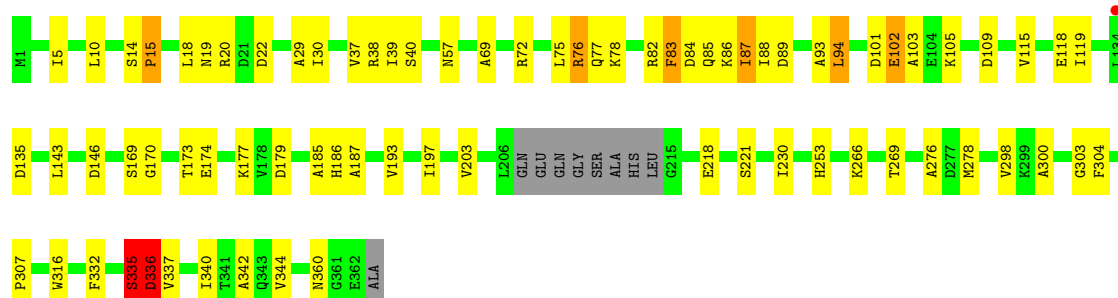
- Molecule 4: CRISPR system Cascade subunit CasC





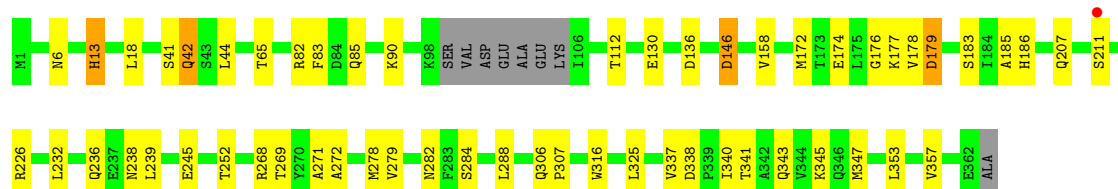
• Molecule 4: CRISPR system Cascade subunit CasC

Chain U: 77% 19% ...



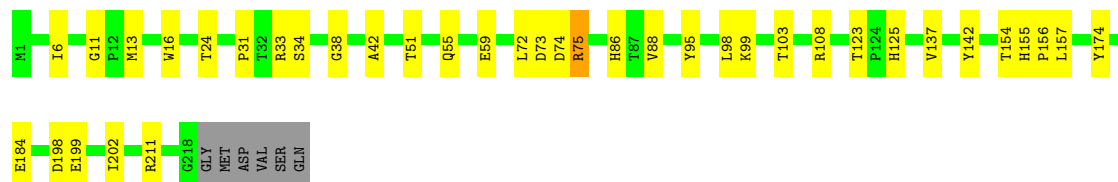
• Molecule 4: CRISPR system Cascade subunit CasC

Chain V: 82% 14% ..



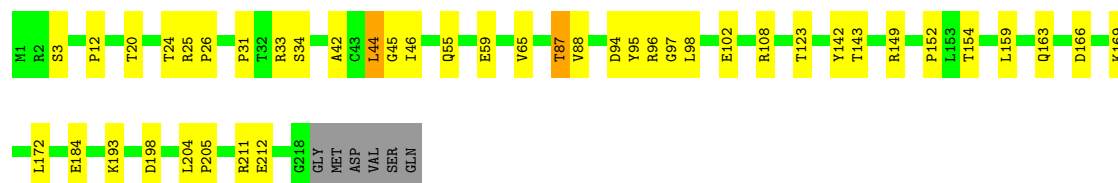
• Molecule 5: CRISPR system Cascade subunit CasD

Chain K: 80% 17% .

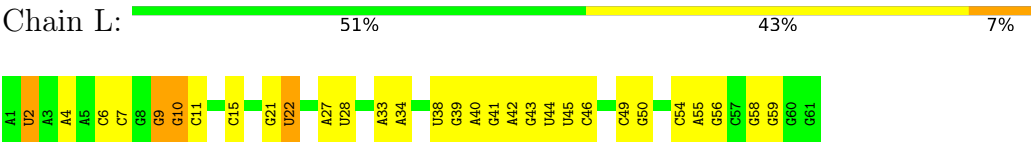


• Molecule 5: CRISPR system Cascade subunit CasD

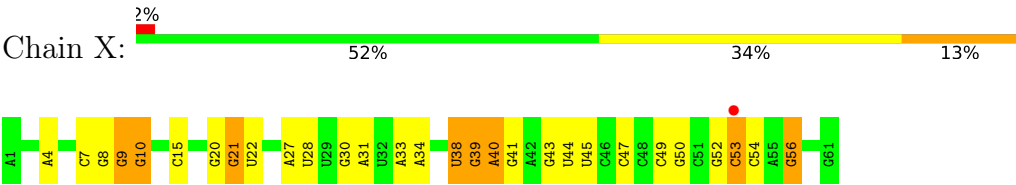
Chain W: 78% 18% ..



● Molecule 6: crRNA



● Molecule 6: crRNA



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	111.41Å 118.14Å 225.87Å 92.23° 93.55° 106.06°	Depositor
Resolution (Å)	43.80 – 3.00 50.05 – 3.00	Depositor EDS
% Data completeness (in resolution range)	88.8 (43.80-3.00) 88.8 (50.05-3.00)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.22 (at 3.01Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.4_1496)	Depositor
R, $R_{free}$	0.179 , 0.218 0.184 , 0.221	Depositor DCC
$R_{free}$ test set	9794 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	55.6	Xtriage
Anisotropy	0.128	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 28.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	54555	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.83% 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	0.44	0/3825	0.55	0/5197
1	M	0.65	1/3901 (0.0%)	0.57	1/5303 (0.0%)
2	B	0.58	0/1316	0.53	0/1777
2	C	0.55	0/1317	0.60	2/1778 (0.1%)
2	N	0.50	0/1270	0.52	0/1714
2	O	0.42	0/1272	0.57	0/1724
3	D	0.50	0/1480	0.65	0/2011
3	P	0.40	0/1488	0.60	0/2019
4	E	0.60	0/2795	0.53	0/3768
4	F	0.60	0/2803	0.52	0/3785
4	G	0.64	0/2796	0.52	0/3771
4	H	0.63	0/2819	0.54	0/3809
4	I	0.61	0/2774	0.55	0/3744
4	J	0.53	0/2722	0.55	0/3681
4	Q	0.70	0/2834	0.56	0/3823
4	R	0.59	0/2779	0.57	2/3758 (0.1%)
4	S	0.59	0/2817	0.56	0/3802
4	T	0.60	0/2821	0.55	0/3805
4	U	0.51	0/2766	0.59	0/3737
4	V	0.72	0/2768	0.62	1/3741 (0.0%)
5	K	0.47	0/1764	0.54	0/2398
5	W	0.69	0/1768	0.61	1/2402 (0.0%)
6	L	1.04	0/1450	1.06	1/2259 (0.0%)
6	X	1.04	0/1450	0.99	0/2259
All	All	0.62	1/55795 (0.0%)	0.60	8/76065 (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
4	R	0	2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	#Chirality outliers	#Planarity outliers
4	U	0	2
All	All	0	4

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	M	458	CYS	CB-SG	-5.56	1.72	1.81

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	112	ALA	N-CA-C	-6.88	92.42	111.00
5	W	44	LEU	CA-CB-CG	5.83	128.71	115.30
1	M	250	CYS	CA-CB-SG	5.62	124.11	114.00
6	L	22	U	O4'-C1'-N1	5.32	112.46	108.20
4	R	21	ASP	CB-CG-OD1	-5.32	113.51	118.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	R	132	ASP	Peptide
4	R	338	ASP	Peptide
4	U	335	SER	Peptide
4	U	94	LEU	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	A	3739	0	3705	64	0
1	M	3812	0	3775	47	0
2	B	1291	0	1295	15	0
2	C	1292	0	1300	25	0
2	N	1246	0	1253	15	0
2	O	1247	0	1226	17	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	D	1451	0	1447	37	0
3	P	1459	0	1466	31	0
4	E	2752	0	2729	31	0
4	F	2760	0	2730	36	0
4	G	2753	0	2727	27	0
4	H	2774	0	2731	30	0
4	I	2731	0	2714	38	0
4	J	2680	0	2624	36	0
4	Q	2790	0	2764	39	0
4	R	2736	0	2672	41	0
4	S	2773	0	2734	42	0
4	T	2777	0	2761	32	0
4	U	2723	0	2690	84	0
4	V	2725	0	2682	37	0
5	K	1722	0	1700	29	0
5	W	1726	0	1711	31	0
6	L	1298	0	658	13	0
6	X	1298	0	658	16	0
All	All	54555	0	52752	738	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 738 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)
4:U:82:ARG:CG	4:U:83:PHE:CE2	1.79	1.60
4:U:82:ARG:CG	4:U:83:PHE:HE2	1.03	1.58
4:U:82:ARG:CB	4:U:83:PHE:CE2	1.92	1.50
4:U:82:ARG:C	4:U:83:PHE:CD2	1.88	1.47
4:U:82:ARG:HB3	4:U:83:PHE:CE2	1.50	1.40

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	470/502 (94%)	432 (92%)	36 (8%)	2 (0%)	34	72
1	M	485/502 (97%)	457 (94%)	24 (5%)	4 (1%)	19	57
2	B	155/160 (97%)	149 (96%)	6 (4%)	0	100	100
2	C	155/160 (97%)	147 (95%)	8 (5%)	0	100	100
2	N	148/160 (92%)	140 (95%)	7 (5%)	1 (1%)	22	60
2	O	154/160 (96%)	147 (96%)	7 (4%)	0	100	100
3	D	185/201 (92%)	178 (96%)	7 (4%)	0	100	100
3	P	185/201 (92%)	176 (95%)	9 (5%)	0	100	100
4	E	353/363 (97%)	333 (94%)	20 (6%)	0	100	100
4	F	355/363 (98%)	335 (94%)	20 (6%)	0	100	100
4	G	352/363 (97%)	335 (95%)	17 (5%)	0	100	100
4	H	361/363 (99%)	348 (96%)	12 (3%)	1 (0%)	41	76
4	I	351/363 (97%)	335 (95%)	16 (5%)	0	100	100
4	J	345/363 (95%)	324 (94%)	21 (6%)	0	100	100
4	Q	358/363 (99%)	347 (97%)	10 (3%)	1 (0%)	41	76
4	R	355/363 (98%)	335 (94%)	20 (6%)	0	100	100
4	S	357/363 (98%)	343 (96%)	13 (4%)	1 (0%)	41	76
4	T	356/363 (98%)	343 (96%)	13 (4%)	0	100	100
4	U	350/363 (96%)	325 (93%)	23 (7%)	2 (1%)	25	64
4	V	351/363 (97%)	331 (94%)	20 (6%)	0	100	100
5	K	216/224 (96%)	211 (98%)	5 (2%)	0	100	100
5	W	216/224 (96%)	210 (97%)	5 (2%)	1 (0%)	29	68
All	All	6613/6850 (96%)	6281 (95%)	319 (5%)	13 (0%)	47	82

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	U	336	ASP
5	W	45	GLY
1	M	160	GLY
1	M	321	ASN
1	M	488	LYS



### 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	401/426 (94%)	391 (98%)	10 (2%)	47	79
1	M	409/426 (96%)	405 (99%)	4 (1%)	76	91
2	B	136/138 (99%)	136 (100%)	0	100	100
2	C	136/138 (99%)	135 (99%)	1 (1%)	84	94
2	N	130/138 (94%)	128 (98%)	2 (2%)	65	87
2	O	128/138 (93%)	125 (98%)	3 (2%)	50	80
3	D	151/171 (88%)	144 (95%)	7 (5%)	27	64
3	P	152/171 (89%)	146 (96%)	6 (4%)	32	69
4	E	289/298 (97%)	284 (98%)	5 (2%)	60	85
4	F	290/298 (97%)	286 (99%)	4 (1%)	67	88
4	G	291/298 (98%)	291 (100%)	0	100	100
4	H	288/298 (97%)	286 (99%)	2 (1%)	84	94
4	I	287/298 (96%)	285 (99%)	2 (1%)	84	94
4	J	279/298 (94%)	274 (98%)	5 (2%)	59	85
4	Q	294/298 (99%)	289 (98%)	5 (2%)	60	85
4	R	282/298 (95%)	279 (99%)	3 (1%)	73	90
4	S	290/298 (97%)	285 (98%)	5 (2%)	60	85
4	T	293/298 (98%)	290 (99%)	3 (1%)	76	91
4	U	287/298 (96%)	276 (96%)	11 (4%)	33	69
4	V	286/298 (96%)	276 (96%)	10 (4%)	36	71
5	K	185/192 (96%)	179 (97%)	6 (3%)	39	74
5	W	186/192 (97%)	180 (97%)	6 (3%)	39	74
All	All	5470/5706 (96%)	5370 (98%)	100 (2%)	59	85

5 of 100 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	Q	13	HIS
4	T	317	ASP
5	W	143	THR
4	Q	62	SER
4	S	19	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
4	Q	13	HIS
4	R	314	GLN
4	R	19	ASN
4	S	85	GLN
4	F	282	ASN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
6	L	60/61 (98%)	19 (31%)	2 (3%)
6	X	60/61 (98%)	19 (31%)	2 (3%)
All	All	120/122 (98%)	38 (31%)	4 (3%)

5 of 38 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
6	L	2	U
6	L	4	A
6	L	9	G
6	L	10	G
6	L	15	C

All (4) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
6	L	27	A
6	L	45	U
6	X	15	C
6	X	45	U

## 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](#)

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	480/502 (95%)	-0.33	1 (0%) 95 87	37, 80, 111, 135	0
1	M	489/502 (97%)	-0.51	0 100 100	26, 42, 68, 122	0
2	B	157/160 (98%)	-0.23	3 (1%) 66 37	30, 44, 84, 120	0
2	C	157/160 (98%)	-0.20	1 (0%) 89 72	32, 52, 85, 115	0
2	N	152/160 (95%)	-0.06	2 (1%) 77 51	40, 56, 85, 114	0
2	O	156/160 (97%)	0.27	5 (3%) 47 20	53, 84, 124, 146	0
3	D	191/201 (95%)	-0.20	2 (1%) 82 59	41, 70, 112, 122	0
3	P	191/201 (95%)	0.17	9 (4%) 31 11	52, 78, 111, 127	0
4	E	357/363 (98%)	-0.36	1 (0%) 94 84	30, 48, 78, 104	0
4	F	359/363 (98%)	-0.29	6 (1%) 70 41	28, 48, 86, 121	0
4	G	356/363 (98%)	-0.43	0 100 100	24, 44, 76, 101	0
4	H	363/363 (100%)	-0.32	1 (0%) 94 84	25, 47, 93, 116	0
4	I	355/363 (97%)	-0.49	0 100 100	26, 47, 79, 104	0
4	J	351/363 (96%)	-0.32	1 (0%) 94 84	35, 62, 110, 129	0
4	Q	362/363 (99%)	-0.40	2 (0%) 89 72	23, 39, 72, 114	0
4	R	359/363 (98%)	-0.17	6 (1%) 70 41	27, 50, 119, 155	0
4	S	361/363 (99%)	-0.16	2 (0%) 89 72	30, 52, 99, 118	0
4	T	360/363 (99%)	-0.34	0 100 100	32, 50, 75, 121	0
4	U	354/363 (97%)	-0.19	1 (0%) 94 84	41, 68, 129, 146	0
4	V	355/363 (97%)	-0.43	1 (0%) 94 84	21, 38, 85, 120	0
5	K	218/224 (97%)	-0.51	0 100 100	44, 67, 90, 106	0
5	W	218/224 (97%)	-0.58	0 100 100	22, 35, 53, 81	0
6	L	61/61 (100%)	-0.24	0 100 100	31, 47, 101, 118	0
6	X	61/61 (100%)	0.06	1 (1%) 72 44	25, 50, 120, 150	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
All	All	6823/6972 (97%)	-0.31	45 (0%) 87 69	21, 52, 103, 155	0

The worst 5 of 45 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	P	112	CYS	4.6
2	B	82	THR	4.4
3	P	97	LEU	4.1
3	P	103	LEU	4.0
2	O	82	THR	3.9

## 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 monosaccharides in this entry.

## 6.4 Ligands ⓘ

There are no ligands in this entry.

## 6.5 Other polymers ⓘ

There are no such residues in this entry.