



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 16, 2022 – 04:36 PM EDT

PDB ID : 4TVX
Title : Crystal structure of the E. coli CRISPR RNA-guided surveillance complex, Cascade
Authors : Jackson, R.N.; Golden, S.M.; Carter, J.; Wiedenheft, B.
Deposited on : 2014-06-28
Resolution : 3.24 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 ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) 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 : 2.27
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) : 2.27

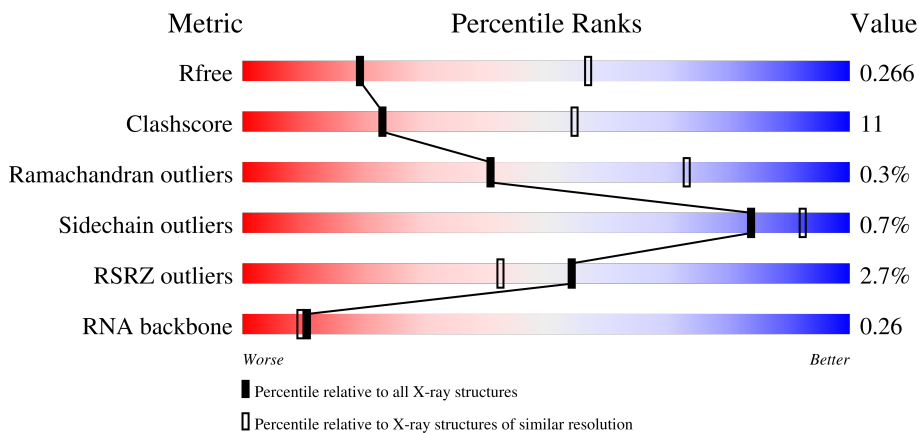
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.24 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	1619 (3.28-3.20)
Clashscore	141614	1755 (3.28-3.20)
Ramachandran outliers	138981	1728 (3.28-3.20)
Sidechain outliers	138945	1727 (3.28-3.20)
RSRZ outliers	127900	1567 (3.28-3.20)
RNA backbone	3102	1034 (3.58-2.90)

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 $\leq 5\%$. 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	J	165	
1	K	165	
1	V	165	
1	W	165	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	B	363	7% 70% 24% 5%
2	C	363	3% 75% 20%
2	D	363	3% 81% 17%
2	E	363	% 82% 15%
2	F	363	% 80% 18%
2	G	363	% 78% 19%
2	N	363	3% 76% 21%
2	O	363	4% 75% 22%
2	P	363	8% 81% 16%
2	Q	363	4% 74% 23%
2	R	363	% 80% 19%
2	S	363	% 79% 20%
3	A	199	5% 73% 23%
3	M	199	3% 82% 16%
4	H	224	% 73% 24%
4	T	224	% 72% 25%
5	L	61	2% 31% 31% 33% 5%
5	X	61	36% 33% 28%
6	I	502	3% 62% 33%
6	U	502	2% 66% 30%

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 103806 atoms, of which 50537 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 CasB.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
1	W	157	2522	795	1255	245	220	7	0	0	0
1	V	155	2322	755	1132	218	210	7	0	0	0
1	K	157	2524	795	1254	245	223	7	0	0	0
1	J	157	2551	800	1280	248	216	7	0	0	0

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
W	-4	GLY	-	expression tag	UNP P76632
W	-3	PRO	-	expression tag	UNP P76632
W	-2	GLY	-	expression tag	UNP P76632
W	-1	TYR	-	expression tag	UNP P76632
W	0	GLN	-	expression tag	UNP P76632
V	-4	GLY	-	expression tag	UNP P76632
V	-3	PRO	-	expression tag	UNP P76632
V	-2	GLY	-	expression tag	UNP P76632
V	-1	TYR	-	expression tag	UNP P76632
V	0	GLN	-	expression tag	UNP P76632
K	-4	GLY	-	expression tag	UNP P76632
K	-3	PRO	-	expression tag	UNP P76632
K	-2	GLY	-	expression tag	UNP P76632
K	-1	TYR	-	expression tag	UNP P76632
K	0	GLN	-	expression tag	UNP P76632
J	-4	GLY	-	expression tag	UNP P76632
J	-3	PRO	-	expression tag	UNP P76632
J	-2	GLY	-	expression tag	UNP P76632
J	-1	TYR	-	expression tag	UNP P76632
J	0	GLN	-	expression tag	UNP P76632

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

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	F	356	Total	C	H	N	O	S	0	0	0
			5396	1702	2681	487	511	15			
2	D	356	Total	C	H	N	O	S	0	0	0
			5165	1660	2518	476	497	14			
2	E	351	Total	C	H	N	O	S	0	0	0
			5237	1664	2582	476	501	14			
2	C	347	Total	C	H	N	O	S	0	0	0
			5023	1610	2455	469	475	14			
2	B	345	Total	C	H	N	O	S	0	0	0
			4922	1590	2389	458	471	14			
2	G	353	Total	C	H	N	O	S	0	0	0
			5306	1681	2621	482	507	15			
2	R	359	Total	C	H	N	O	S	0	0	0
			5416	1714	2686	485	516	15			
2	P	354	Total	C	H	N	O	S	0	0	0
			5117	1644	2505	471	482	15			
2	Q	354	Total	C	H	N	O	S	0	0	0
			5148	1651	2519	469	494	15			
2	O	355	Total	C	H	N	O	S	0	0	0
			5123	1645	2500	471	493	14			
2	N	358	Total	C	H	N	O	S	0	0	0
			5229	1673	2561	479	502	14			
2	S	361	Total	C	H	N	O	S	0	0	0
			5412	1715	2677	490	515	15			

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

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	A	192	Total	C	H	N	O	S	0	0	0
			2991	954	1506	262	262	7			
3	M	195	Total	C	H	N	O	S	0	1	0
			3057	971	1542	274	263	7			

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

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
4	H	219	Total	C	H	N	O	S	0	0	0
			3434	1093	1708	307	317	9			
4	T	219	Total	C	H	N	O	S	0	0	0
			3436	1093	1710	307	317	9			

- Molecule 5 is a RNA chain called Escherichia coli strain ECOR44 cluster 1 CRISPR region.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	P			
5	L	61	1957	580	657	233	426	61	0	0	0
5	X	60	1898	562	634	227	415	60	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
L	42	A	C	conflict	GB 50811866
L	53	C	U	conflict	GB 50811866
X	42	A	C	conflict	GB 50811866
X	53	C	U	conflict	GB 50811866

- Molecule 6 is a protein called CRISPR system Cascade subunit CasA.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
6	U	493	7393	2398	3633	668	675	19	0	0	0
6	I	493	7225	2356	3532	658	660	19	0	0	0


- Molecule 7 is ZINC ION (three-letter code: ZN) (formula: Zn).

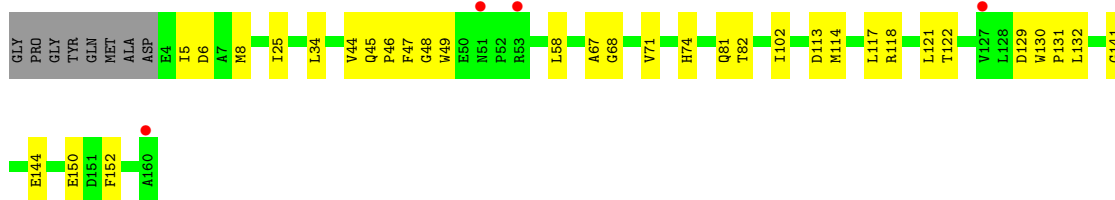
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	U	1	Total	Zn	0	0
			1	1		
7	I	1	Total	Zn	0	0
			1	1		

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.

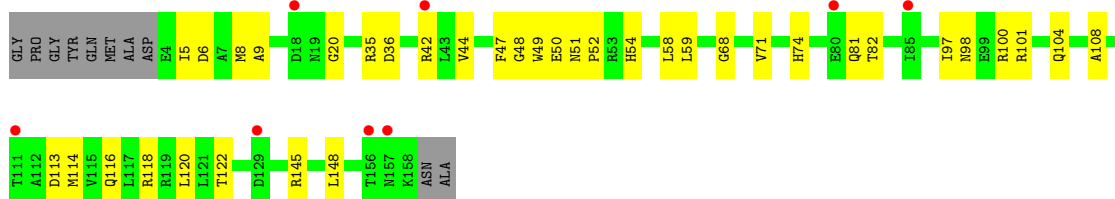
- Molecule 1: CRISPR system Cascade subunit CasB

Chain W: 



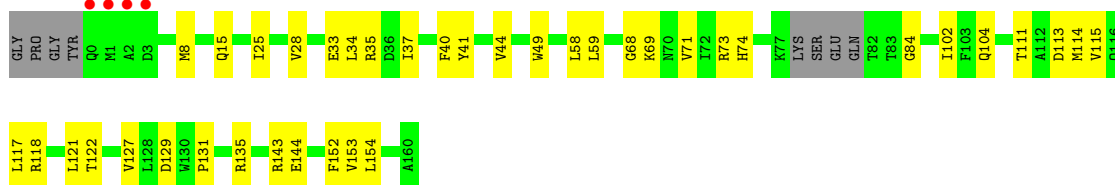
- Molecule 1: CRISPR system Cascade subunit CasB

Chain V: 



- Molecule 1: CRISPR system Cascade subunit CasB

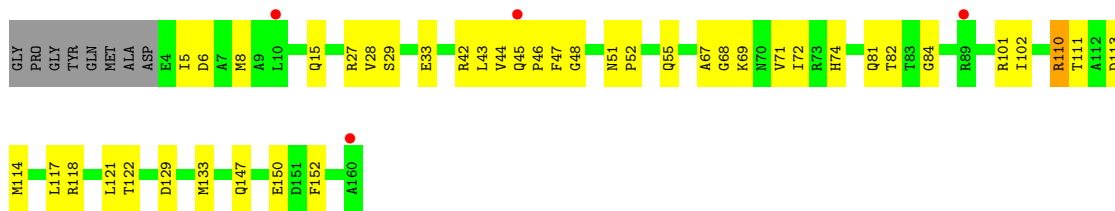
Chain K: 



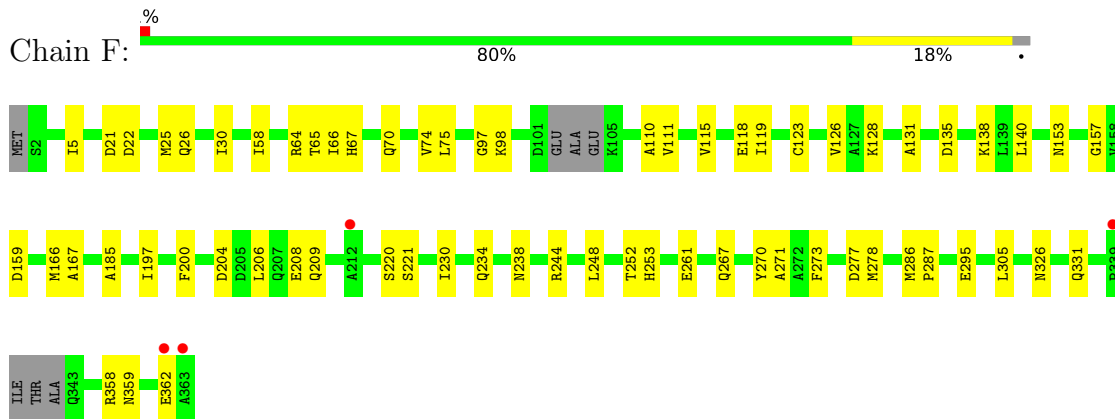
- Molecule 1: CRISPR system Cascade subunit CasB

Chain J: 

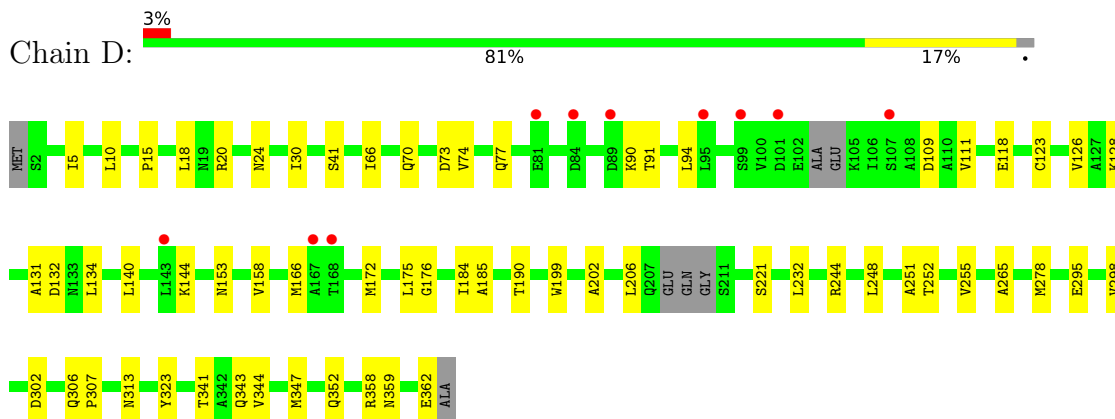




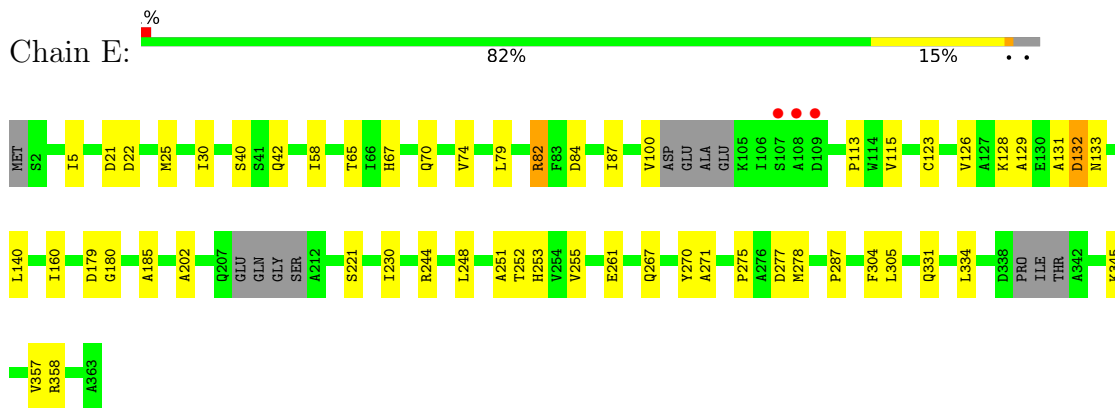
● Molecule 2: CRISPR system Cascade subunit CasC



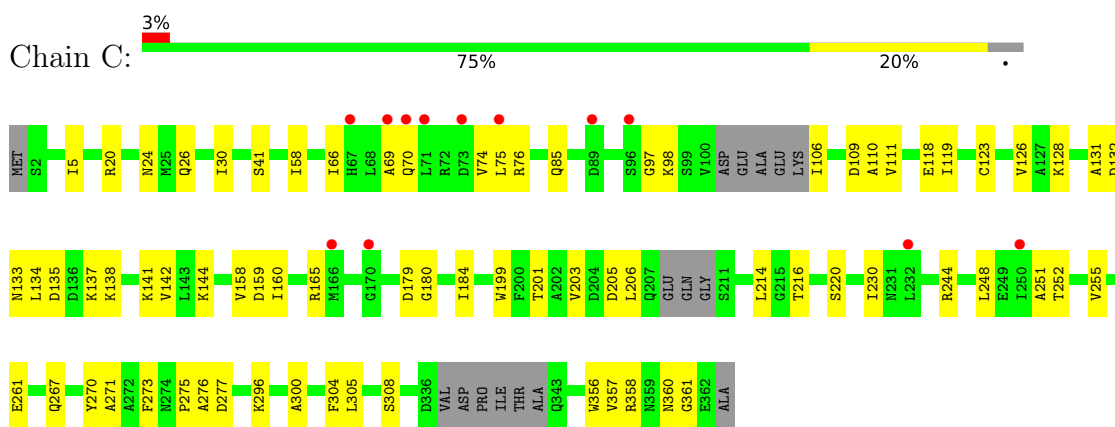
● Molecule 2: CRISPR system Cascade subunit CasC



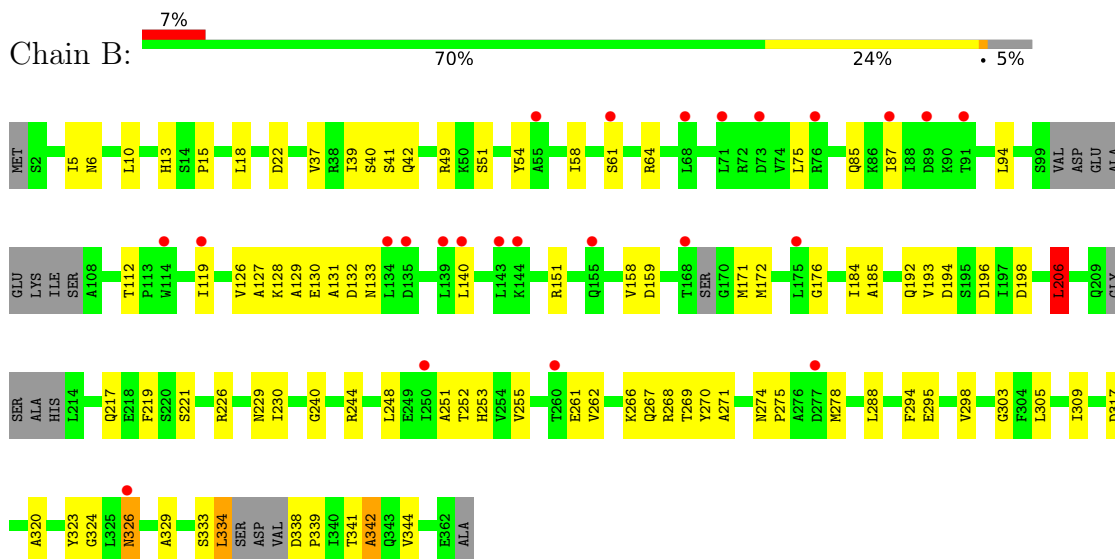
● Molecule 2: CRISPR system Cascade subunit CasC



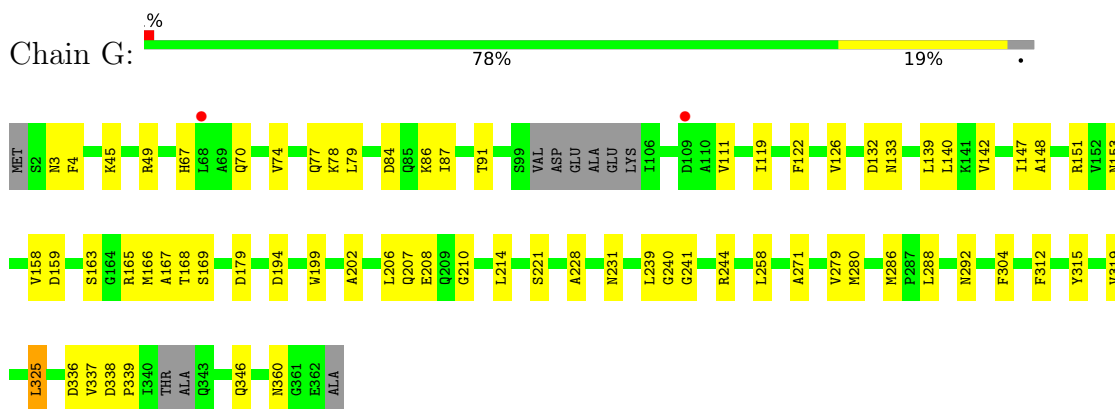
● Molecule 2: CRISPR system Cascade subunit CasC



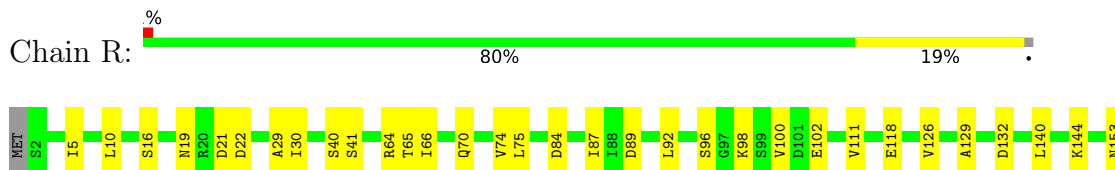
- Molecule 2: CRISPR system Cascade subunit CasC

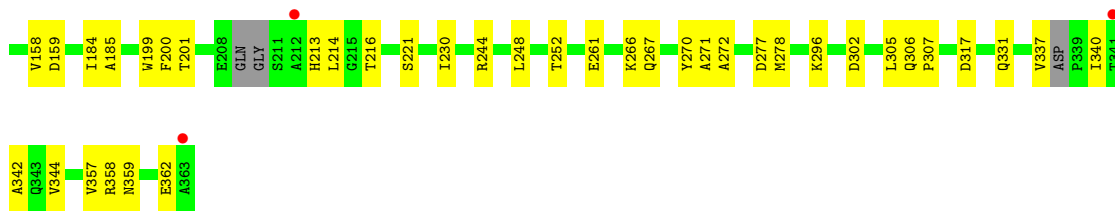


- Molecule 2: CRISPR system Cascade subunit CasC

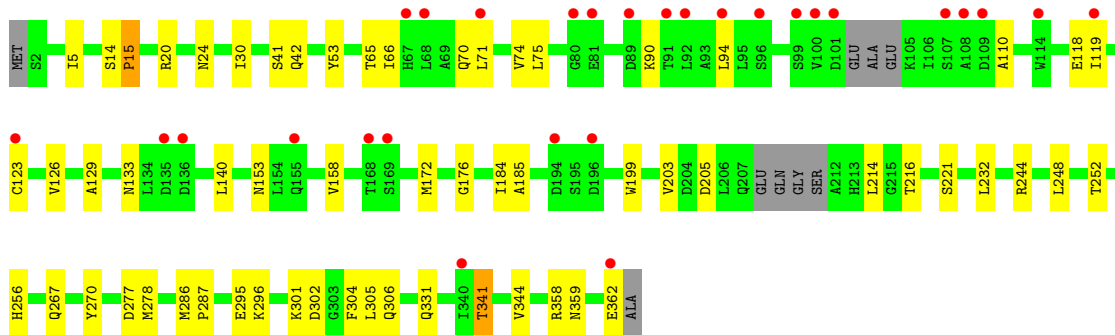
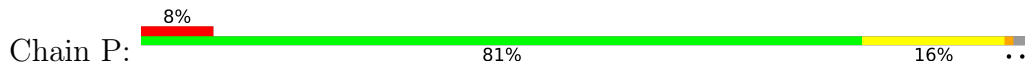


- Molecule 2: CRISPR system Cascade subunit CasC

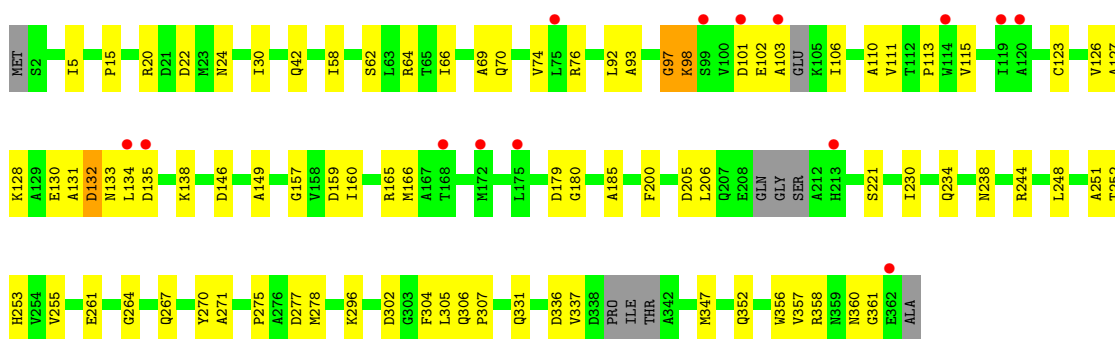
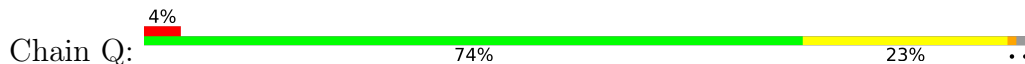




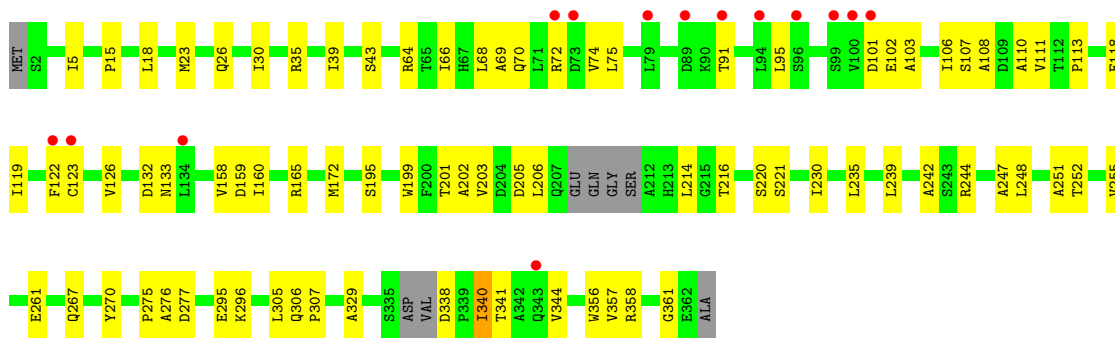
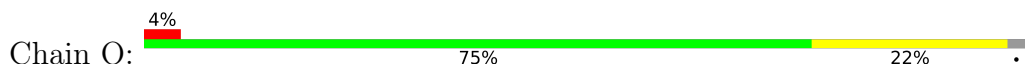
● Molecule 2: CRISPR system Cascade subunit CasC



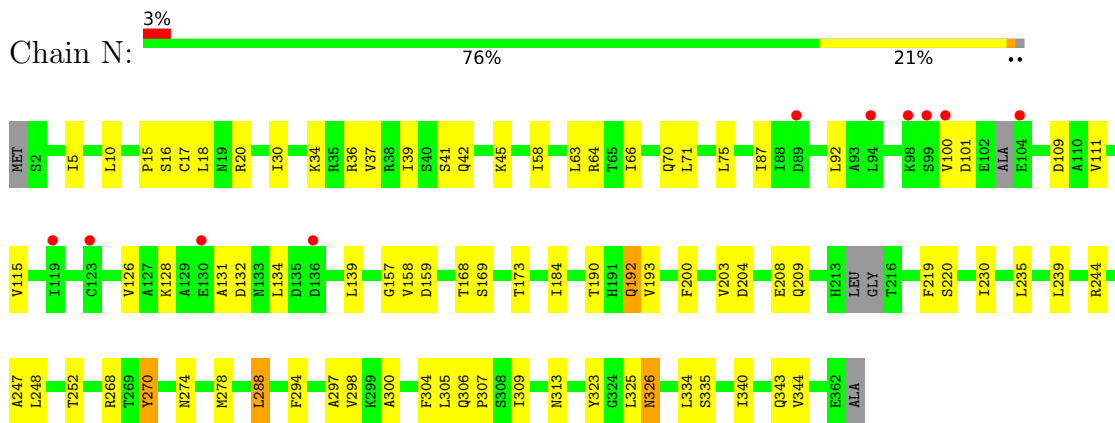
● Molecule 2: CRISPR system Cascade subunit CasC



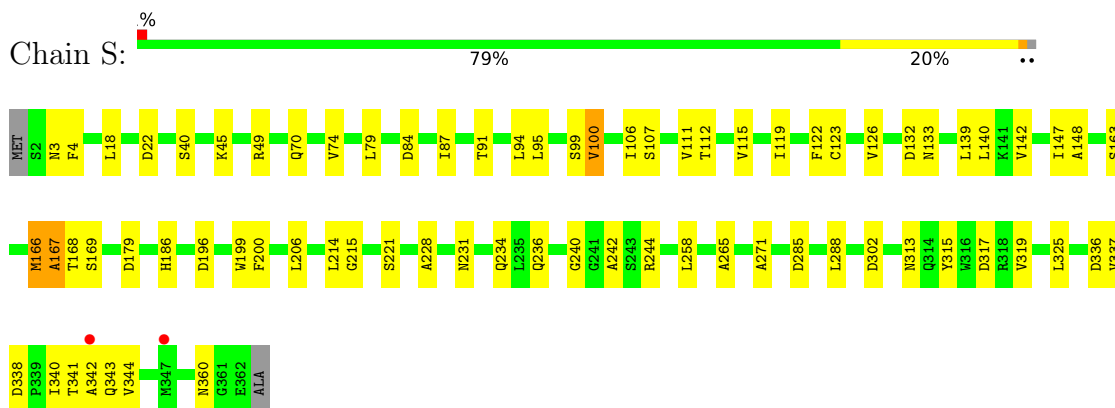
● Molecule 2: CRISPR system Cascade subunit CasC



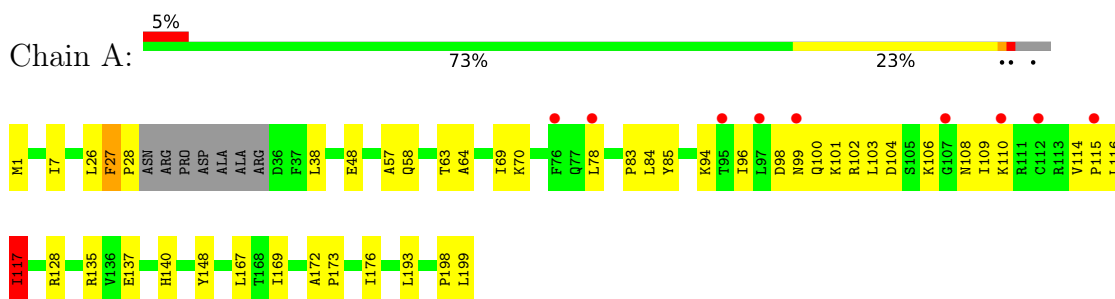
- Molecule 2: CRISPR system Cascade subunit CasC



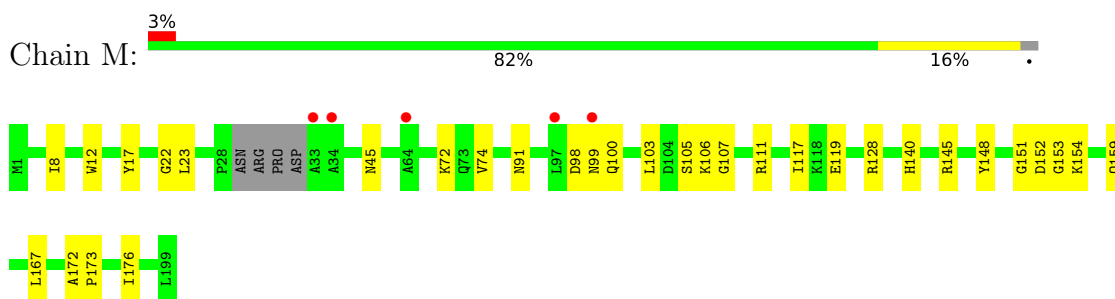
- Molecule 2: CRISPR system Cascade subunit CasC



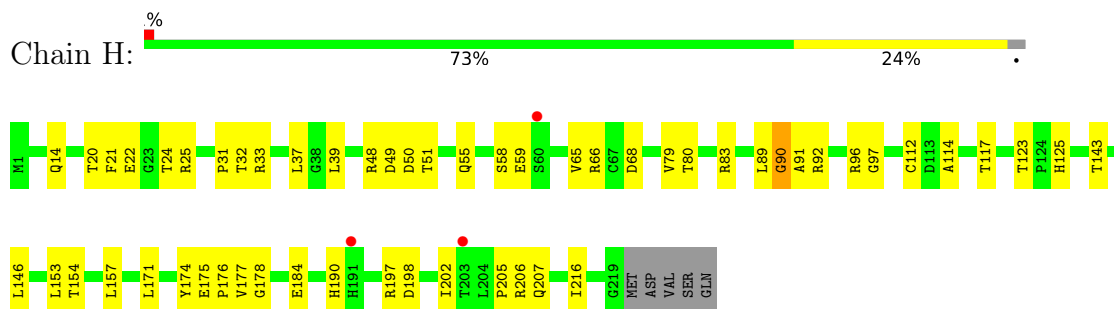
- Molecule 3: CRISPR system Cascade subunit CasE



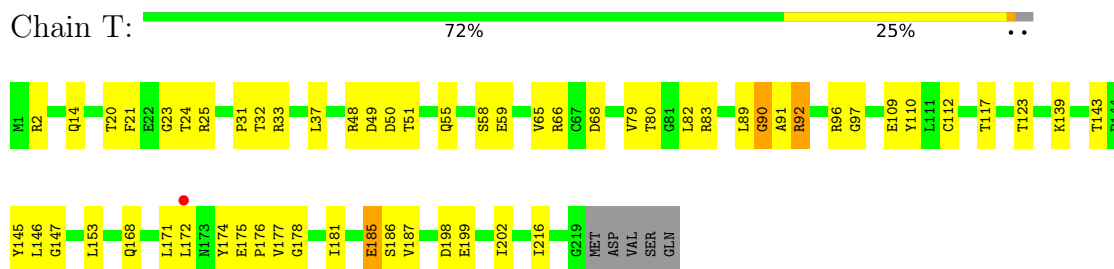
- Molecule 3: CRISPR system Cascade subunit CasE



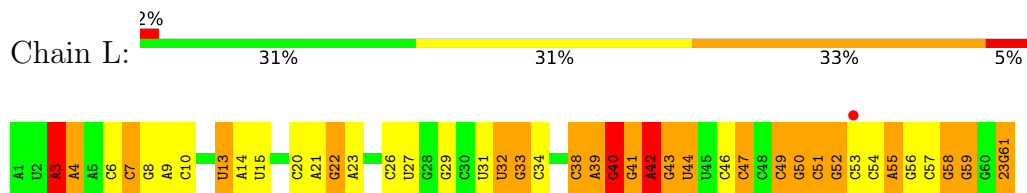
- Molecule 4: CRISPR system Cascade subunit CasD



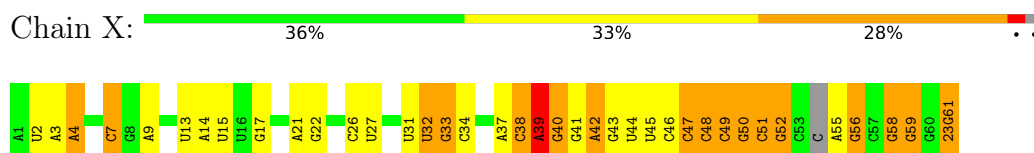
- Molecule 4: CRISPR system Cascade subunit CasD



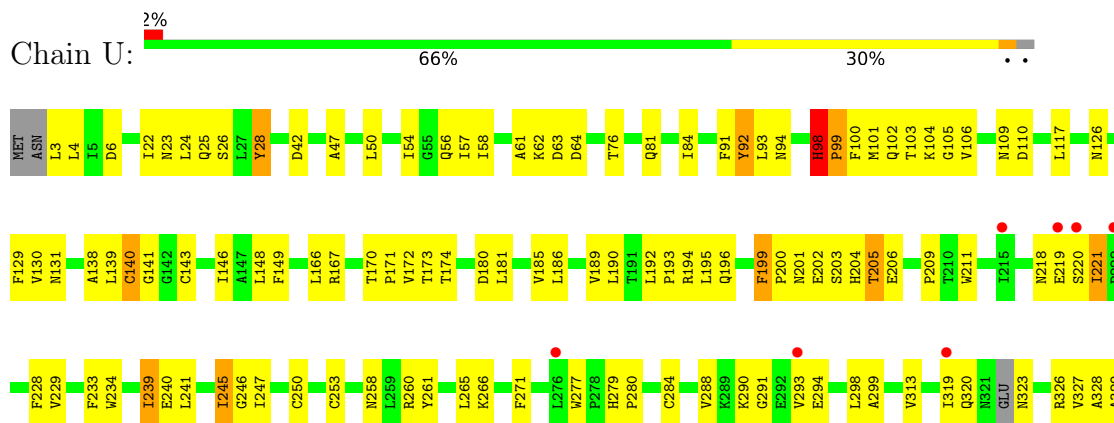
- Molecule 5: Escherichia coli strain ECOR44 cluster 1 CRISPR region

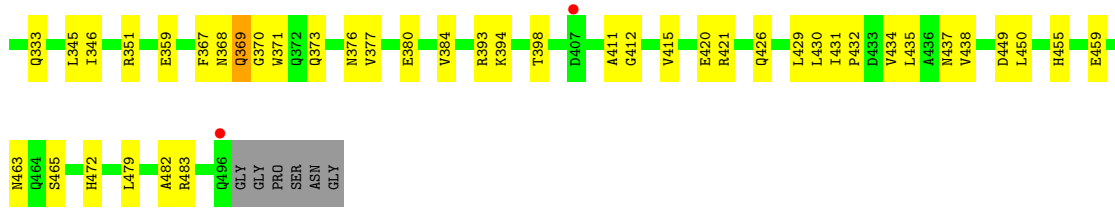


- Molecule 5: Escherichia coli strain ECOR44 cluster 1 CRISPR region

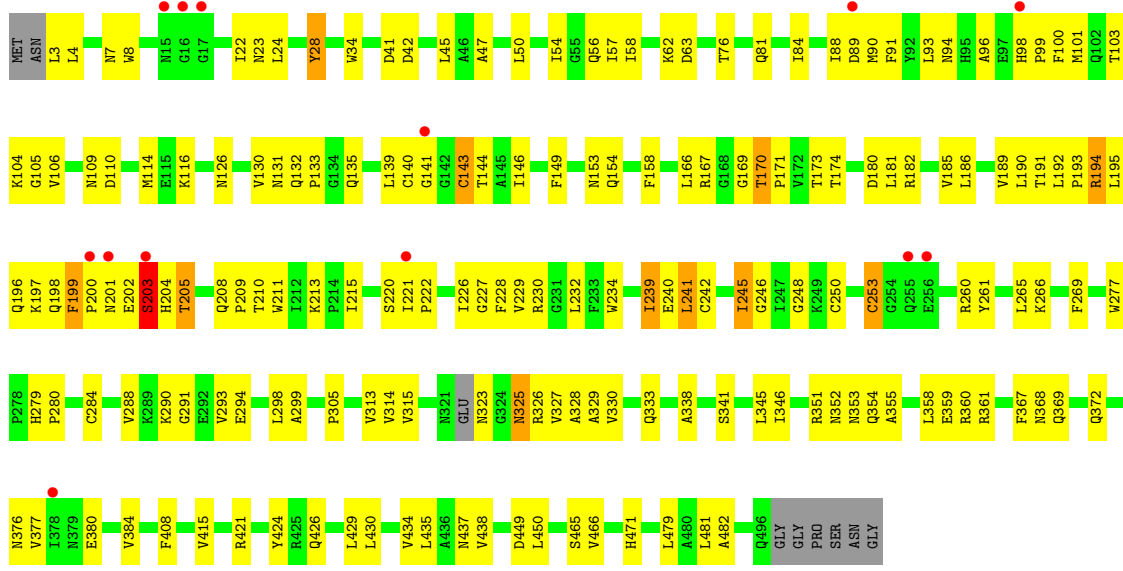


- Molecule 6: CRISPR system Cascade subunit CasA





● Molecule 6: CRISPR system Cascade subunit CasA



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	200.85Å 214.54Å 217.38Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.61 – 3.24 39.61 – 3.24	Depositor EDS
% Data completeness (in resolution range)	99.8 (39.61-3.24) 96.1 (39.61-3.24)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.13 (at 3.25Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: dev_1760)	Depositor
R, R_{free}	0.234 , 0.268 0.237 , 0.266	Depositor DCC
R_{free} test set	2034 reflections (1.37%)	wwPDB-VP
Wilson B-factor (Å ²)	74.0	Xtrriage
Anisotropy	0.395	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 45.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	0.016 for -h,l,k	Xtrriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	103806	wwPDB-VP
Average B, all atoms (Å ²)	94.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.64% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 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: ZN, 23G

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	J	0.28	0/1296	0.53	0/1751
1	K	0.29	0/1294	0.55	0/1751
1	V	0.28	0/1215	0.55	0/1656
1	W	0.27	0/1292	0.55	0/1749
2	B	0.31	0/2573	0.65	2/3491 (0.1%)
2	C	0.30	0/2608	0.59	0/3530
2	D	0.29	0/2690	0.58	0/3649
2	E	0.31	0/2696	0.58	2/3648 (0.1%)
2	F	0.31	0/2757	0.59	0/3723
2	G	0.32	0/2728	0.65	2/3687 (0.1%)
2	N	0.30	0/2711	0.63	0/3676
2	O	0.34	1/2665 (0.0%)	0.60	0/3617
2	P	0.29	0/2654	0.58	0/3601
2	Q	0.30	0/2670	0.59	2/3618 (0.1%)
2	R	0.31	0/2773	0.60	0/3747
2	S	0.32	0/2780	0.64	0/3763
3	A	0.28	0/1516	0.55	0/2056
3	M	0.28	0/1549	0.56	0/2100
4	H	0.28	0/1768	0.58	1/2402 (0.0%)
4	T	0.27	0/1768	0.57	1/2402 (0.0%)
5	L	0.38	0/1423	1.02	4/2216 (0.2%)
5	X	0.39	0/1382	1.04	2/2151 (0.1%)
6	I	0.33	0/3782	0.65	2/5159 (0.0%)
6	U	0.37	0/3851	0.66	0/5244
All	All	0.31	1/54441 (0.0%)	0.64	18/74387 (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
2	B	0	3
2	E	0	1
2	F	0	1
2	G	0	1
2	N	0	1
2	P	0	1
2	Q	0	2
2	R	0	2
2	S	0	3
3	A	0	1
4	H	0	2
4	T	0	3
6	I	0	3
6	U	0	7
All	All	0	31

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	O	102	GLU	CB-CG	6.10	1.63	1.52

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	I	194	ARG	NE-CZ-NH1	8.31	124.46	120.30
2	G	86	LYS	CD-CE-NZ	-7.60	94.22	111.70
2	E	82	ARG	NE-CZ-NH1	7.44	124.02	120.30
6	I	194	ARG	NE-CZ-NH2	-6.94	116.83	120.30
2	Q	97	GLY	N-CA-C	-5.96	98.21	113.10

There are no chirality outliers.

5 of 31 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	206	LEU	Peptide
2	B	221	SER	Peptide
2	B	342	ALA	Peptide
2	E	132	ASP	Peptide
2	F	326	ASN	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	J	1271	1280	1280	30	0
1	K	1270	1254	1257	25	0
1	V	1190	1132	1131	29	0
1	W	1267	1255	1255	22	0
2	B	2533	2389	2389	74	0
2	C	2568	2455	2455	59	0
2	D	2647	2518	2518	47	0
2	E	2655	2582	2582	48	0
2	F	2715	2681	2681	56	0
2	G	2685	2621	2621	52	0
2	N	2668	2561	2561	72	0
2	O	2623	2500	2500	65	0
2	P	2612	2505	2505	47	0
2	Q	2629	2519	2519	63	0
2	R	2730	2686	2685	57	0
2	S	2735	2677	2677	69	0
3	A	1485	1506	1505	36	0
3	M	1515	1542	1541	30	0
4	H	1726	1708	1708	40	0
4	T	1726	1710	1710	43	0
5	L	1300	657	656	39	0
5	X	1264	634	633	35	0
6	I	3693	3532	3533	155	0
6	U	3760	3633	3634	136	0
7	I	1	0	0	0	0
7	U	1	0	0	0	0
All	All	53269	50537	50536	1183	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 11.

The worst 5 of 1183 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)
2:D:73:ASP:OD1	2:D:77:GLN:NE2	1.90	1.03

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:L:52:G:N7	5:L:55:A:N6	2.06	1.03
6:U:202:GLU:O	6:U:205:THR:OG1	1.84	0.94
2:B:58:ILE:HG22	2:B:158:VAL:HG11	1.51	0.92
2:S:244:ARG:NH1	2:S:360:ASN:OD1	2.01	0.92

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	J	155/165 (94%)	154 (99%)	1 (1%)	0	100	100
1	K	153/165 (93%)	151 (99%)	2 (1%)	0	100	100
1	V	153/165 (93%)	152 (99%)	1 (1%)	0	100	100
1	W	155/165 (94%)	153 (99%)	2 (1%)	0	100	100
2	B	335/363 (92%)	315 (94%)	18 (5%)	2 (1%)	25	61
2	C	339/363 (93%)	323 (95%)	16 (5%)	0	100	100
2	D	350/363 (96%)	327 (93%)	20 (6%)	3 (1%)	17	52
2	E	343/363 (94%)	328 (96%)	15 (4%)	0	100	100
2	F	350/363 (96%)	330 (94%)	20 (6%)	0	100	100
2	G	347/363 (96%)	322 (93%)	25 (7%)	0	100	100
2	N	352/363 (97%)	333 (95%)	18 (5%)	1 (0%)	41	73
2	O	349/363 (96%)	325 (93%)	19 (5%)	5 (1%)	11	43
2	P	348/363 (96%)	332 (95%)	14 (4%)	2 (1%)	25	61
2	Q	346/363 (95%)	329 (95%)	15 (4%)	2 (1%)	25	61
2	R	353/363 (97%)	336 (95%)	16 (4%)	1 (0%)	41	73
2	S	359/363 (99%)	336 (94%)	22 (6%)	1 (0%)	41	73

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	A	188/199 (94%)	182 (97%)	4 (2%)	2 (1%)	14	48
3	M	192/199 (96%)	186 (97%)	6 (3%)	0	100	100
4	H	217/224 (97%)	208 (96%)	9 (4%)	0	100	100
4	T	217/224 (97%)	209 (96%)	8 (4%)	0	100	100
6	I	489/502 (97%)	460 (94%)	28 (6%)	1 (0%)	47	78
6	U	489/502 (97%)	458 (94%)	30 (6%)	1 (0%)	47	78
All	All	6579/6866 (96%)	6249 (95%)	309 (5%)	21 (0%)	41	73

5 of 21 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	117	ILE
2	S	100	VAL
6	I	170	THR
2	D	343	GLN
2	N	343	GLN

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	J	130/141 (92%)	128 (98%)	2 (2%)	65	83
1	K	130/141 (92%)	130 (100%)	0	100	100
1	V	115/141 (82%)	115 (100%)	0	100	100
1	W	129/141 (92%)	129 (100%)	0	100	100
2	B	243/298 (82%)	240 (99%)	3 (1%)	71	86
2	C	249/298 (84%)	248 (100%)	1 (0%)	91	95
2	D	259/298 (87%)	259 (100%)	0	100	100
2	E	271/298 (91%)	271 (100%)	0	100	100
2	F	282/298 (95%)	282 (100%)	0	100	100
2	G	277/298 (93%)	274 (99%)	3 (1%)	73	87

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	N	266/298 (89%)	263 (99%)	3 (1%)	73	87
2	O	257/298 (86%)	257 (100%)	0	100	100
2	P	254/298 (85%)	254 (100%)	0	100	100
2	Q	260/298 (87%)	260 (100%)	0	100	100
2	R	282/298 (95%)	282 (100%)	0	100	100
2	S	280/298 (94%)	280 (100%)	0	100	100
3	A	158/170 (93%)	157 (99%)	1 (1%)	86	93
3	M	159/170 (94%)	159 (100%)	0	100	100
4	H	185/192 (96%)	185 (100%)	0	100	100
4	T	185/192 (96%)	184 (100%)	1 (0%)	88	94
6	I	371/426 (87%)	360 (97%)	11 (3%)	41	70
6	U	386/426 (91%)	376 (97%)	10 (3%)	46	74
All	All	5128/5716 (90%)	5093 (99%)	35 (1%)	84	92

5 of 35 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
6	I	199	PHE
6	I	203	SER
6	I	241	LEU
2	N	326	ASN
2	N	288	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
2	D	292	ASN
2	B	209	GLN
2	G	153	ASN
2	N	331	GLN

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
5	L	59/61 (96%)	33 (55%)	3 (5%)
5	X	56/61 (91%)	28 (50%)	1 (1%)

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
All	All	115/122 (94%)	61 (53%)	4 (3%)

5 of 61 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
5	L	3	A
5	L	4	A
5	L	7	C
5	L	9	A
5	L	13	U

All (4) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
5	L	9	A
5	L	40	G
5	L	42	A
5	X	9	A

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

2 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
5	23G	X	61	5	19,29,30	5.55	12 (63%)	21,45,48	3.27	12 (57%)
5	23G	L	61	5	19,29,30	5.44	12 (63%)	21,45,48	3.26	9 (42%)

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
5	23G	X	61	5	-	0/3/35/36	0/4/4/4
5	23G	L	61	5	-	1/3/35/36	0/4/4/4

The worst 5 of 24 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	X	61	23G	C4-N3	11.81	1.54	1.35
5	L	61	23G	C4-N3	11.54	1.53	1.35
5	X	61	23G	C2-N2	8.84	1.51	1.33
5	L	61	23G	C2-N2	8.76	1.51	1.33
5	X	61	23G	C3'-C4'	-8.17	1.30	1.52

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	X	61	23G	N3-C2-N1	-8.69	115.63	127.22
5	L	61	23G	N3-C2-N1	-8.48	115.91	127.22
5	X	61	23G	C2-N3-C4	6.49	122.77	115.36
5	L	61	23G	C2-N3-C4	6.38	122.65	115.36
5	L	61	23G	C1'-N9-C4	-5.15	117.59	126.64

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	L	61	23G	C4'-C5'-O5'-P

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	X	61	23G	2	0
5	L	61	23G	1	0

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 2 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, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	J	157/165 (95%)	0.21	4 (2%) 57 46	74, 95, 126, 153	0
1	K	157/165 (95%)	0.17	4 (2%) 57 46	51, 75, 108, 152	0
1	V	155/165 (93%)	0.38	8 (5%) 27 17	78, 105, 137, 153	0
1	W	157/165 (95%)	0.04	4 (2%) 57 46	57, 82, 108, 122	0
2	B	345/363 (95%)	0.42	24 (6%) 16 11	74, 115, 160, 168	0
2	C	347/363 (95%)	0.30	12 (3%) 44 32	56, 99, 149, 164	0
2	D	356/363 (98%)	0.31	10 (2%) 53 41	50, 89, 147, 160	0
2	E	351/363 (96%)	0.04	3 (0%) 84 78	36, 67, 98, 121	0
2	F	356/363 (98%)	0.10	4 (1%) 80 73	31, 66, 118, 151	0
2	G	353/363 (97%)	-0.01	2 (0%) 89 85	30, 60, 96, 109	0
2	N	358/363 (98%)	0.26	10 (2%) 53 41	49, 87, 140, 160	0
2	O	355/363 (97%)	0.25	14 (3%) 39 28	50, 89, 145, 164	0
2	P	354/363 (97%)	0.39	28 (7%) 12 8	54, 93, 152, 172	0
2	Q	354/363 (97%)	0.22	14 (3%) 38 28	45, 73, 132, 148	0
2	R	359/363 (98%)	0.09	3 (0%) 86 80	31, 61, 99, 122	0
2	S	361/363 (99%)	-0.01	2 (0%) 89 85	33, 62, 113, 153	0
3	A	192/199 (96%)	0.44	9 (4%) 31 21	75, 96, 122, 208	0
3	M	195/199 (97%)	0.23	5 (2%) 56 44	67, 93, 122, 139	0
4	H	219/224 (97%)	0.12	3 (1%) 75 66	42, 74, 105, 124	0
4	T	219/224 (97%)	0.12	1 (0%) 91 87	36, 66, 94, 110	0
5	L	60/61 (98%)	0.27	1 (1%) 70 60	43, 83, 167, 186	0
5	X	59/61 (96%)	0.18	0 100 100	40, 79, 141, 174	0
6	I	493/502 (98%)	0.34	13 (2%) 56 44	46, 113, 176, 267	0
6	U	493/502 (98%)	0.16	9 (1%) 68 58	50, 84, 120, 227	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
All	All	6805/6988 (97%)	0.21	187 (2%) 54 42	30, 84, 142, 267	0

The worst 5 of 187 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	363	ALA	6.7
2	R	363	ALA	6.3
2	B	175	LEU	5.2
2	Q	103	ALA	4.4
1	V	156	THR	4.3

6.2 Non-standard residues in protein, DNA, RNA chains [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, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
5	23G	L	61	26/27	0.86	0.23	62,102,137,142	0
5	23G	X	61	26/27	0.86	0.22	74,105,132,145	0

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, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
7	ZN	I	601	1/1	0.62	0.14	181,181,181,181	0
7	ZN	U	601	1/1	0.88	0.21	181,181,181,181	0

6.5 Other polymers [i](#)

There are no such residues in this entry.