



wwPDB EM Validation Summary Report ⓘ

Jul 15, 2025 – 08:05 PM JST

PDB ID : 8K22 / pdb_00008k22
EMDB ID : EMD-36825
Title : ICP1 Csy-dsDNA-Cas1-Cas2/3 complex (half form)
Authors : Zhang, L.X.; Feng, Y.
Deposited on : 2023-07-12
Resolution : 2.92 Å(reported)

This is a wwPDB EM 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/EMValidationReportHelp>

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:

EMDB validation analysis : **FAILED**
MolProbity : 4-5-2 with Phenix2.0rc1
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ : **FAILED**
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.44

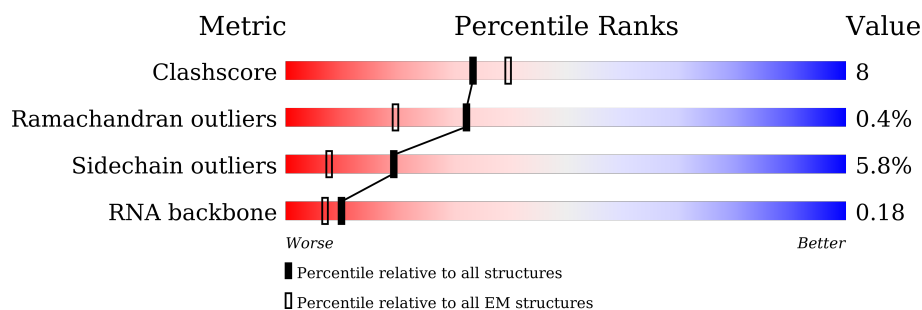
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 2.92 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.



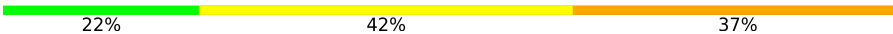
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415
RNA backbone	6643	2191

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the 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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	B	178	80% 19% ..
2	C	248	79% 18% .
3	D	295	77% 18% .
3	E	295	78% 22%
3	d	295	73% 26% .
3	e	295	73% 23% .
4	G	306	76% 22% ..
4	I	306	77% 20% ..

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	J	306	 75% 22% ..
4	K	306	 74% 25% .
4	L	306	 71% 26% ..
4	M	306	 72% 26% ..
5	H	168	 77% 19% ..
6	P	60	 22% 42% 37%
7	Q	43	 74% 23% .
7	U	43	 37% 12% 51%
8	R	31	 81% 19%
9	A	930	 73% 26% .
9	a	930	 6% . 92%
10	V	21	 81% 19%

2 Entry composition

There are 11 unique types of molecules in this entry. The entry contains 39285 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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 Csy1.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	B	177	Total	C	N	O	S	0	0
			1415	896	235	274	10		

- Molecule 2 is a protein called Csy2.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	C	247	Total	C	N	O	S	0	0
			1903	1214	320	355	14		

- Molecule 3 is a protein called Cas1.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	D	295	Total	C	N	O	S	2	0
			2323	1484	398	429	12		
3	E	295	Total	C	N	O	S	0	0
			2329	1491	394	431	13		
3	d	295	Total	C	N	O	S	2	0
			2337	1492	401	431	13		
3	e	295	Total	C	N	O	S	0	0
			2324	1489	393	429	13		

- Molecule 4 is a protein called Csy3.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	G	302	Total	C	N	O	S	0	0
			2314	1456	392	459	7		
4	I	302	Total	C	N	O	S	0	0
			2298	1447	389	455	7		
4	J	302	Total	C	N	O	S	0	0
			2318	1459	393	459	7		
4	K	302	Total	C	N	O	S	0	0
			2318	1459	393	459	7		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf	Trace
4	L	302	Total	C	N	O	S	0	0
			2307	1450	391	459	7		
4	M	302	Total	C	N	O	S	0	0
			2314	1456	392	459	7		

- Molecule 5 is a protein called Csy4.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	H	165	Total	C	N	O	S	0	0
			1290	818	221	245	6		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
H	51	ILE	VAL	conflict	UNP F1D5V5

- Molecule 6 is a RNA chain called RNA (60-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
6	P	60	Total	C	N	O	P	0	0
			1260	565	210	425	60		

- Molecule 7 is a DNA chain called DNA (43-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
7	Q	43	Total	C	N	O	P	0	0
			898	425	184	246	43		
7	U	21	Total	C	N	O	P	0	0
			439	209	88	121	21		

- Molecule 8 is a DNA chain called DNA (31-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
8	R	31	Total	C	N	O	P	0	0
			620	299	94	196	31		

- Molecule 9 is a protein called HD Cas3-type domain-containing protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	a	75	Total	C	N	O	S	0	0
			579	370	101	104	4		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf	Trace
9	A	929	Total	C	N	O	S	0	0
			7278	4632	1245	1367	34		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
a	834	LYS	ARG	conflict	UNP F1D5V9
A	834	LYS	ARG	conflict	UNP F1D5V9

- Molecule 10 is a DNA chain called DNA (5'-D(*AP*TP*CP*TP*TP*CP*CP*CP*TP*AP*TP*TP*TP*AP*AP*AP*TP*TP*GP*CP*T)-3').

Mol	Chain	Residues	Atoms					AltConf	Trace
10	V	21	Total	C	N	O	P	0	0
			419	205	65	129	20		

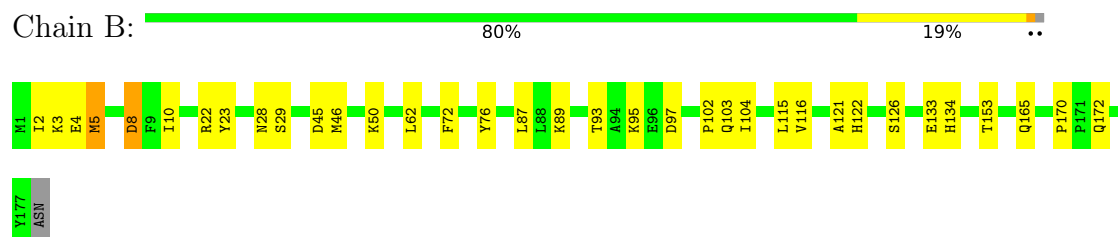
- Molecule 11 is MANGANESE (II) ION (CCD ID: MN) (formula: Mn) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		AltConf
11	A	2	Total	Mn	0
			2	2	

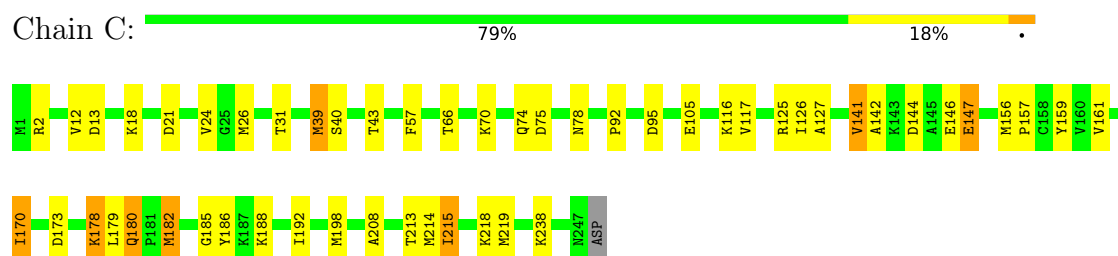
3 Residue-property plots [i](#)

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. 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. 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.

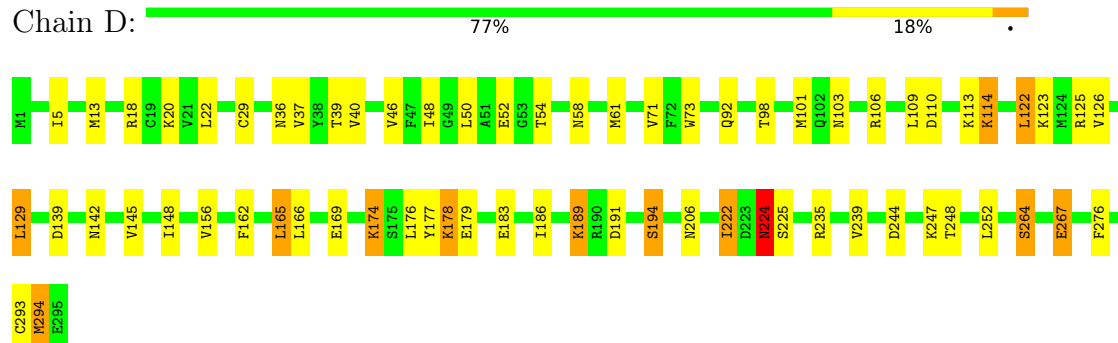
• Molecule 1: Csy1



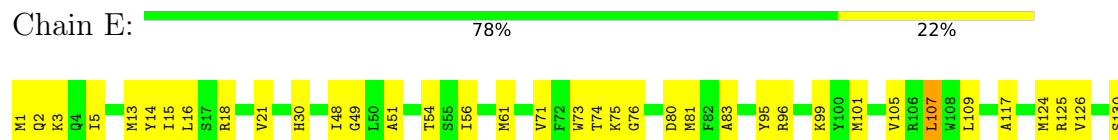
• Molecule 2: Csy2



• Molecule 3: Cas1



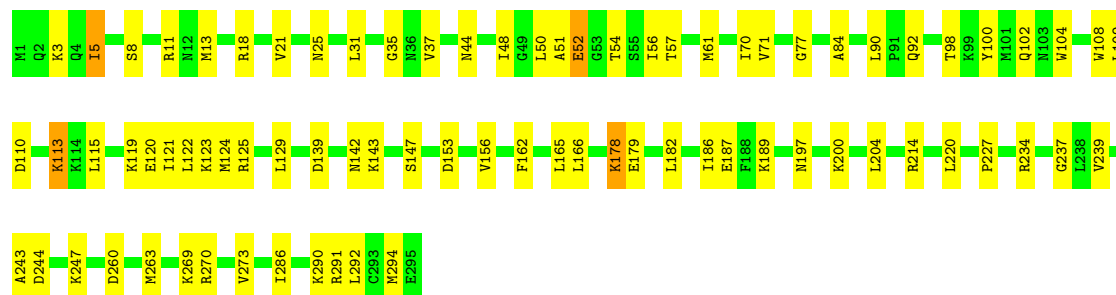
• Molecule 3: Cas1





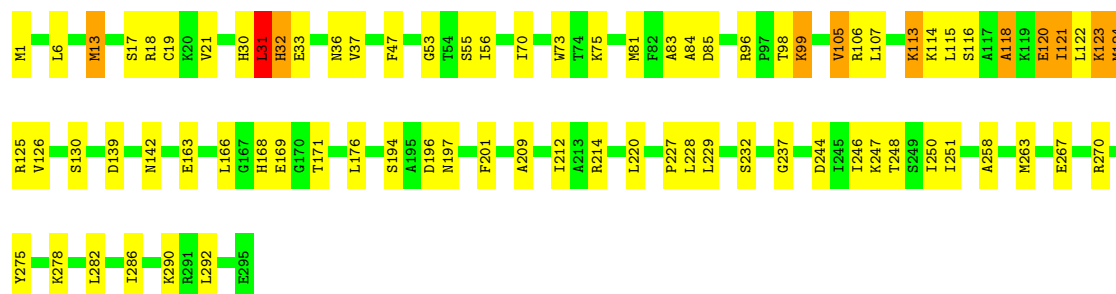
• Molecule 3: Cas1

Chain d: 73% 26%



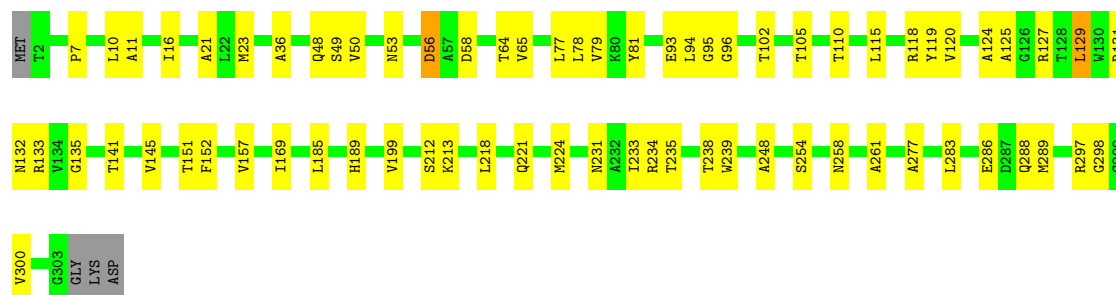
• Molecule 3: Cas1

Chain e: 73% 23%



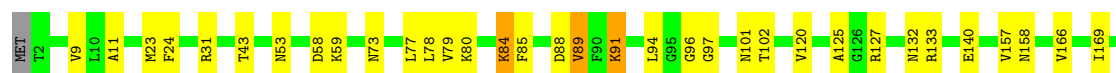
• Molecule 4: Csy3

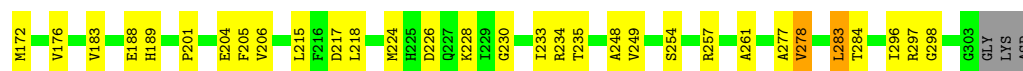
Chain G: 76% 22%



• Molecule 4: Csy3

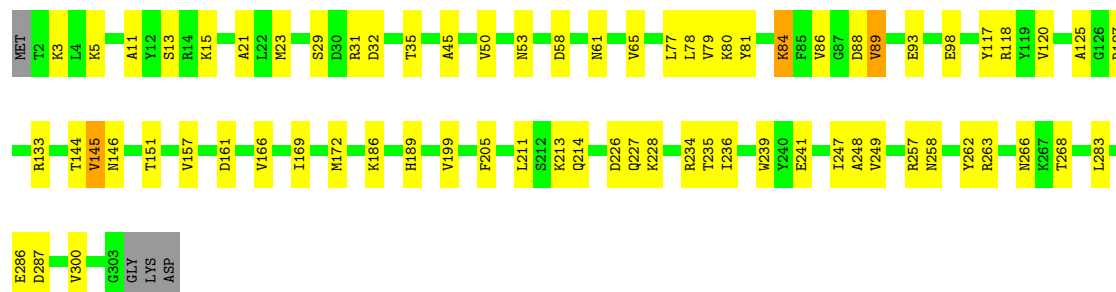
Chain I: 77% 20%





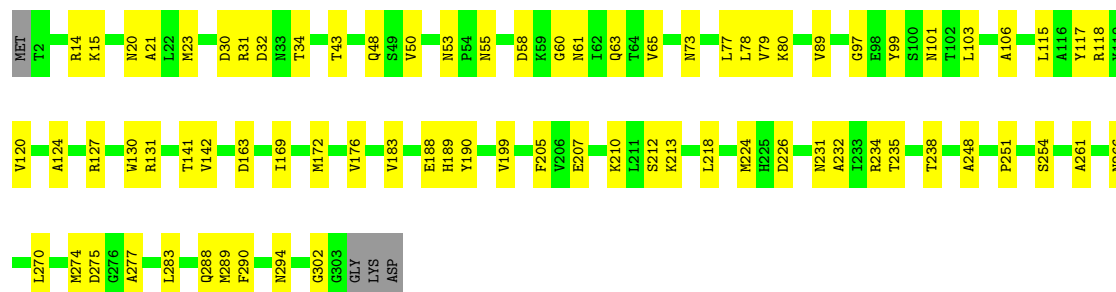
• Molecule 4: Csy3

Chain J: 75% 22% ..



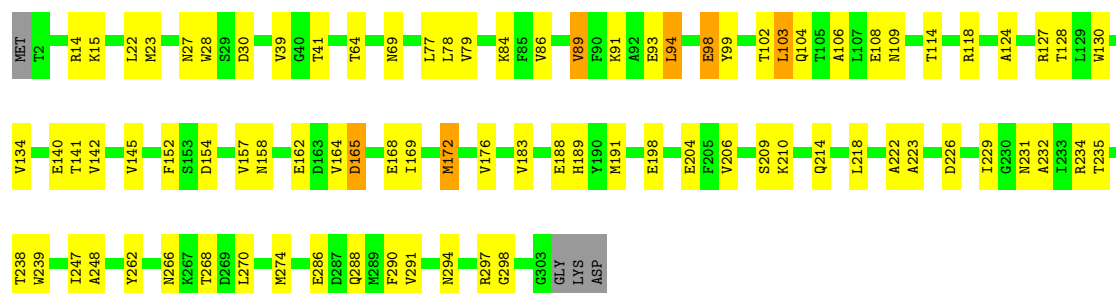
• Molecule 4: Csy3

Chain K: 74% 25% .



• Molecule 4: Csy3

Chain L: 71% 26% ..

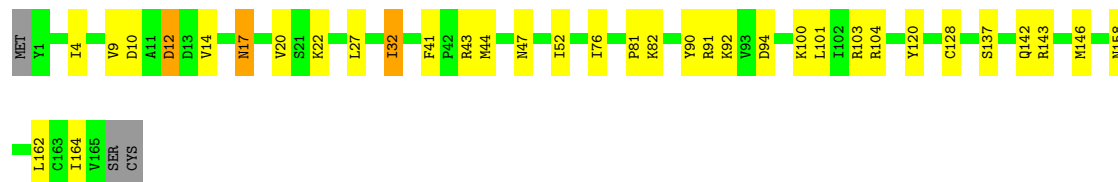
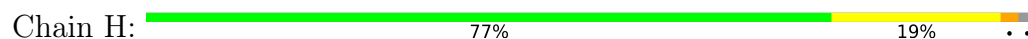


• Molecule 4: Csy3

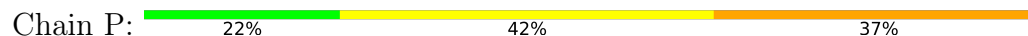
Chain M: 72% 26% ..



- Molecule 5: Csy4



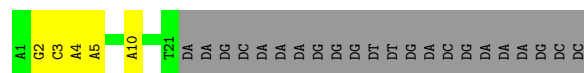
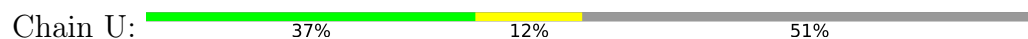
- Molecule 6: RNA (60-MER)



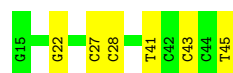
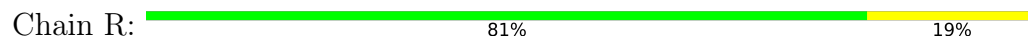
- Molecule 7: DNA (43-MER)



- Molecule 7: DNA (43-MER)

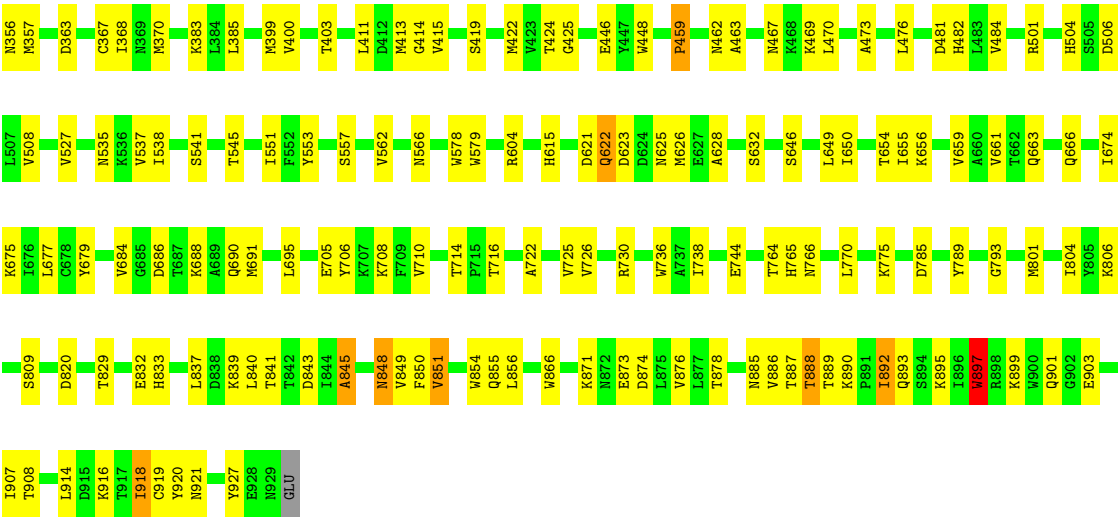


- Molecule 8: DNA (31-MER)

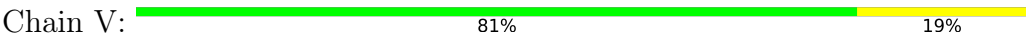


- Molecule 9: HD Cas3-type domain-containing protein





● Molecule 10: DNA (5'-D(*AP*TP*CP*TP*TP*CP*CP*CP*TP*AP*TP*TP*TP*AP*AP*AP*TP*TP*GP*CP*T)-3')



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	235177	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	50	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	1500	Depositor
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor

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: MN

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	B	0.32	0/1448	0.61	2/1956 (0.1%)
2	C	0.32	0/1940	0.67	4/2625 (0.2%)
3	D	0.48	0/2377	0.76	2/3215 (0.1%)
3	E	0.26	0/2377	0.62	3/3210 (0.1%)
3	d	0.34	0/2391	0.74	0/3231
3	e	0.51	2/2373 (0.1%)	0.75	6/3208 (0.2%)
4	G	0.34	0/2354	0.60	1/3198 (0.0%)
4	I	0.36	0/2338	0.65	2/3180 (0.1%)
4	J	0.35	0/2358	0.67	2/3202 (0.1%)
4	K	0.36	0/2358	0.64	0/3202
4	L	0.42	0/2347	0.78	5/3190 (0.2%)
4	M	0.28	0/2354	0.63	2/3198 (0.1%)
5	H	0.37	0/1308	0.77	1/1757 (0.1%)
6	P	0.34	0/1404	0.62	2/2181 (0.1%)
7	Q	0.31	0/1014	0.53	1/1565 (0.1%)
7	U	0.31	0/495	0.48	0/763
8	R	0.28	0/688	0.53	0/1057
9	A	0.37	1/7432 (0.0%)	0.72	14/10072 (0.1%)
9	a	0.34	0/588	0.79	0/791
10	V	0.54	0/466	0.80	1/716 (0.1%)
All	All	0.37	3/40410 (0.0%)	0.68	48/55517 (0.1%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	e	31	LEU	CA-C	-6.96	1.44	1.53
9	A	83	LYS	CA-C	-6.25	1.44	1.52
3	e	96	ARG	C-N	5.02	1.39	1.33

The worst 5 of 48 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	L	98	GLU	N-CA-C	-10.38	100.69	113.97
4	J	53	ASN	CB-CA-C	10.11	118.38	111.20
9	A	13	ILE	N-CA-C	-9.93	103.36	112.90
9	A	77	ASP	N-CA-C	-9.79	101.21	113.55
5	H	12	ASP	N-CA-C	-8.65	102.53	113.43

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	1415	0	1353	17	0
2	C	1903	0	1916	34	0
3	D	2323	0	2294	40	0
3	E	2329	0	2300	39	0
3	d	2337	0	2318	46	0
3	e	2324	0	2296	40	0
4	G	2314	0	2249	38	0
4	I	2298	0	2216	37	0
4	J	2318	0	2260	37	0
4	K	2318	0	2260	46	0
4	L	2307	0	2229	46	0
4	M	2314	0	2249	47	0
5	H	1290	0	1336	26	0
6	P	1260	0	641	36	0
7	Q	898	0	481	9	0
7	U	439	0	237	5	0
8	R	620	0	355	5	0
9	A	7278	0	7186	129	0
9	a	579	0	595	14	0
10	V	419	0	243	4	0
11	A	2	0	0	0	0
All	All	39285	0	37014	609	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 8.

The worst 5 of 609 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)
9:A:663:GLN:HG2	9:A:892:ILE:HD11	1.49	0.94
4:K:21:ALA:HA	4:K:80:LYS:O	1.83	0.79
3:D:224:ASN:HD21	3:D:239:VAL:HG12	1.52	0.73
5:H:10:ASP:CG	4:L:158:ASN:HD21	1.96	0.73
9:A:174:GLN:HB3	9:A:257:TRP:HE1	1.56	0.71

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 PDB entries followed by that with respect to all EM entries.

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	B	175/178 (98%)	165 (94%)	10 (6%)	0	100	100
2	C	245/248 (99%)	225 (92%)	20 (8%)	0	100	100
3	D	295/295 (100%)	279 (95%)	14 (5%)	2 (1%)	19	47
3	E	293/295 (99%)	286 (98%)	7 (2%)	0	100	100
3	d	295/295 (100%)	278 (94%)	16 (5%)	1 (0%)	37	65
3	e	293/295 (99%)	279 (95%)	14 (5%)	0	100	100
4	G	300/306 (98%)	290 (97%)	10 (3%)	0	100	100
4	I	300/306 (98%)	284 (95%)	16 (5%)	0	100	100
4	J	300/306 (98%)	286 (95%)	11 (4%)	3 (1%)	13	38
4	K	300/306 (98%)	286 (95%)	14 (5%)	0	100	100
4	L	300/306 (98%)	285 (95%)	15 (5%)	0	100	100
4	M	300/306 (98%)	286 (95%)	13 (4%)	1 (0%)	37	65
5	H	163/168 (97%)	148 (91%)	14 (9%)	1 (1%)	22	51
9	A	927/930 (100%)	850 (92%)	69 (7%)	8 (1%)	14	41

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
9	a	73/930 (8%)	71 (97%)	1 (1%)	1 (1%)	9	30
All	All	4559/5470 (83%)	4298 (94%)	244 (5%)	17 (0%)	32	59

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	M	62	ILE
9	A	159	ILE
9	A	851	VAL
3	D	5	ILE
4	J	50	VAL

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 PDB entries followed by that with respect to all EM entries.

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	B	155/157 (99%)	146 (94%)	9 (6%)	17	44
2	C	208/211 (99%)	195 (94%)	13 (6%)	15	40
3	D	247/254 (97%)	224 (91%)	23 (9%)	7	22
3	E	247/254 (97%)	239 (97%)	8 (3%)	34	67
3	d	250/254 (98%)	232 (93%)	18 (7%)	12	33
3	e	247/254 (97%)	227 (92%)	20 (8%)	9	28
4	G	246/251 (98%)	238 (97%)	8 (3%)	33	66
4	I	241/251 (96%)	227 (94%)	14 (6%)	17	44
4	J	247/251 (98%)	232 (94%)	15 (6%)	15	42
4	K	247/251 (98%)	235 (95%)	12 (5%)	21	51
4	L	244/251 (97%)	225 (92%)	19 (8%)	10	29
4	M	246/251 (98%)	239 (97%)	7 (3%)	38	71
5	H	148/151 (98%)	143 (97%)	5 (3%)	32	65
9	A	792/817 (97%)	743 (94%)	49 (6%)	15	41
9	a	63/817 (8%)	62 (98%)	1 (2%)	58	83

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
All	All	3828/4675 (82%)	3607 (94%)	221 (6%)	19	44

5 of 221 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	L	108	GLU
3	d	124	MET
9	A	916	LYS
9	A	837	LEU
4	L	165	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 53 such sidechains are listed below:

Mol	Chain	Res	Type
4	M	101	ASN
3	d	92	GLN
9	A	622	GLN
4	M	146	ASN
9	a	51	ASN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
6	P	59/60 (98%)	37 (62%)	3 (5%)

5 of 37 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
6	P	-6	U
6	P	-5	U
6	P	-2	A
6	P	0	A
6	P	1	G

All (3) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
6	P	11	U
6	P	22	C
6	P	41	C

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 oligosaccharides 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.