wwPDB EM Validation Summary Report (i)

Nov 20, 2022 – 01:22 AM EST

PDB ID	:	3IXV
EMDB ID	:	EMD-5100
Title	:	Scorpion Hemocyanin resting state pseudo atomic model built based on cryo-EM density map
Authors	:	Cong, Y.; Zhang, Q.; Woolford, D.; Schweikardt, T.; Khant, H.; Ludtke, S.; Chiu, W.; Decker, H.
Deposited on	:	2009-02-13
Resolution	:	6.80 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 (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

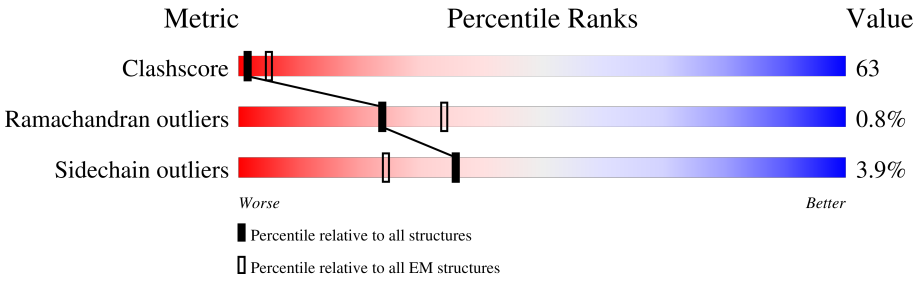
EMDB validation analysis	:	0.0.1.dev43
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	FAILED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.3

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 6.80 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	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

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 <=5%

Mol	Chain	Length	Quality of chain
1	A	626	37% 58% •
1	C	626	38% 55% 6% •
1	D	626	38% 57% 5%
1	E	626	37% 57% 5% •
1	F	626	38% 58% •
1	G	626	40% 55% 5%
1	H	626	40% 54% 5% •
1	I	626	38% 57% 5%
1	J	626	42% 52% 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	K	626	37% 56% 6%
1	L	626	38% 57% 5%
1	M	626	40% 54%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 74208 atoms, of which 13476 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 Hemocyanin AA6 chain.

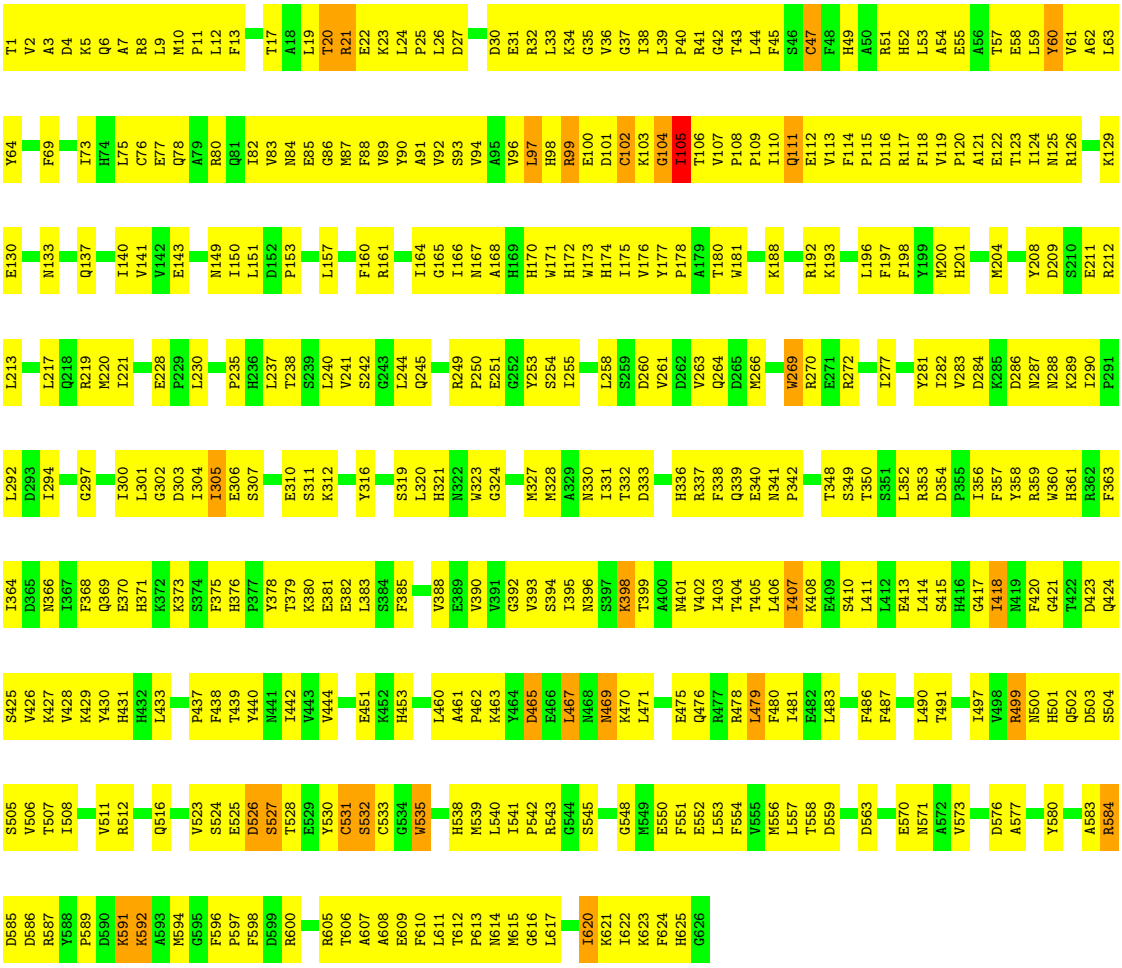
Mol	Chain	Residues	Atoms						AltConf	Trace
1	A	626	Total	C	H	N	O	S	0	0
			6184	3199	1123	877	961	24		
1	C	626	Total	C	H	N	O	S	0	0
			6184	3199	1123	877	961	24		
1	D	626	Total	C	H	N	O	S	0	0
			6184	3199	1123	877	961	24		
1	E	626	Total	C	H	N	O	S	0	0
			6184	3199	1123	877	961	24		
1	F	626	Total	C	H	N	O	S	0	0
			6184	3199	1123	877	961	24		
1	G	626	Total	C	H	N	O	S	0	0
			6184	3199	1123	877	961	24		
1	H	626	Total	C	H	N	O	S	0	0
			6184	3199	1123	877	961	24		
1	I	626	Total	C	H	N	O	S	0	0
			6184	3199	1123	877	961	24		
1	J	626	Total	C	H	N	O	S	0	0
			6184	3199	1123	877	961	24		
1	K	626	Total	C	H	N	O	S	0	0
			6184	3199	1123	877	961	24		
1	L	626	Total	C	H	N	O	S	0	0
			6184	3199	1123	877	961	24		
1	M	626	Total	C	H	N	O	S	0	0
			6184	3199	1123	877	961	24		

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.

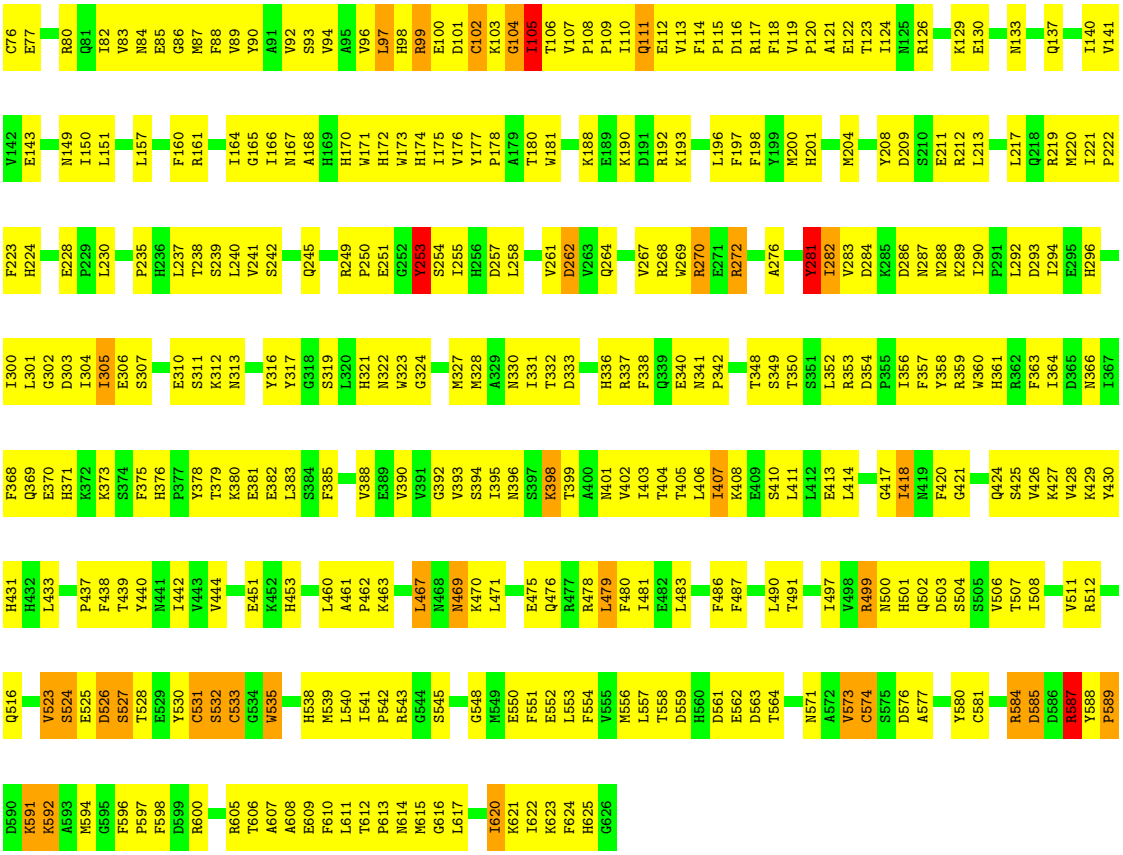
- Molecule 1: Hemocyanin AA6 chain

Chain A:



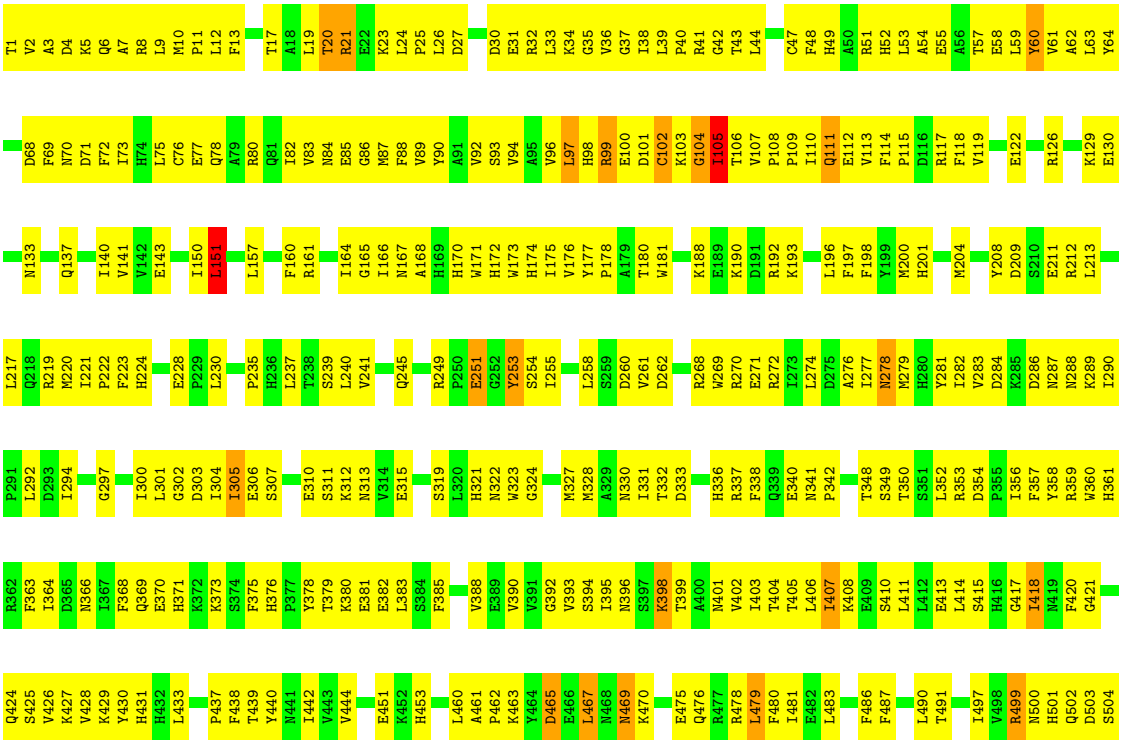
- Molecule 1: Hemocyanin AA6 chain

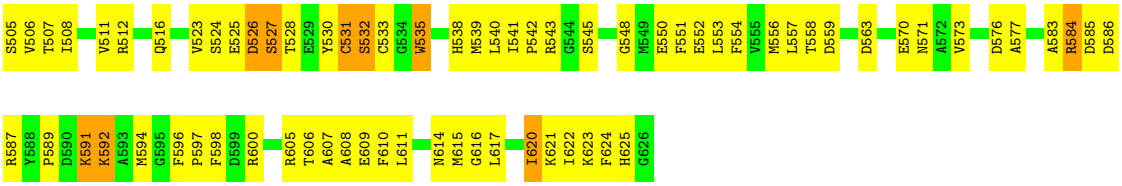
Chain C:



• Molecule 1: Hemocyanin AA6 chain

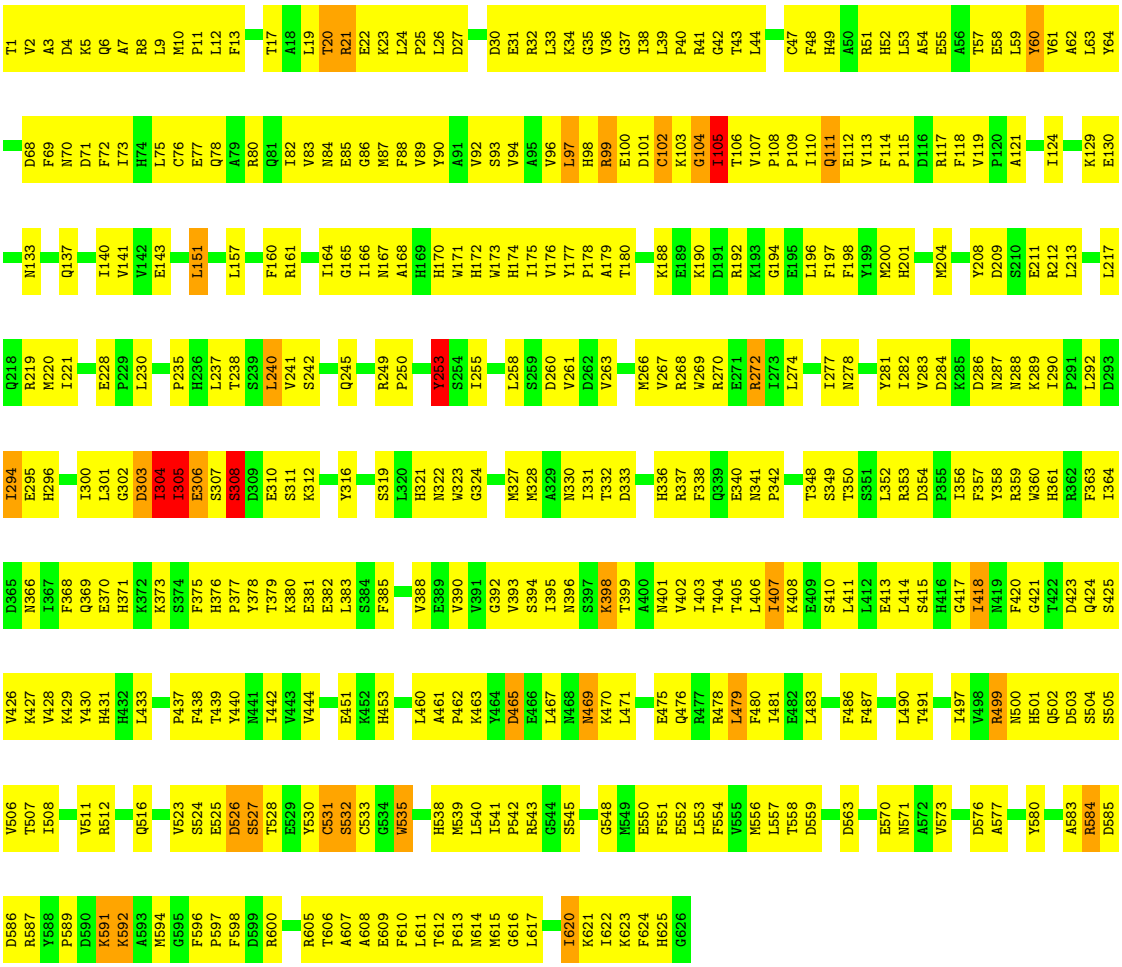
Chain D:





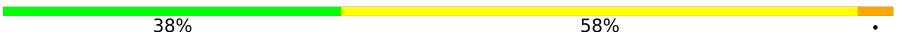
• Molecule 1: Hemocyanin AA6 chain

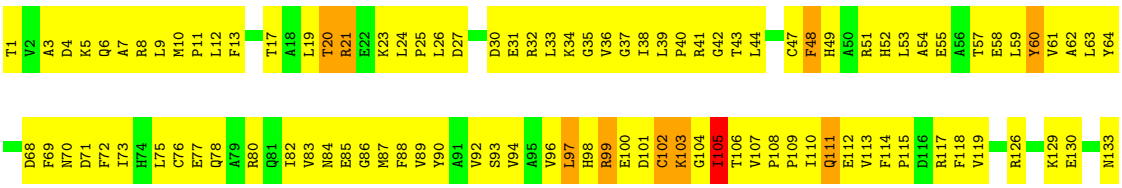
Chain E:

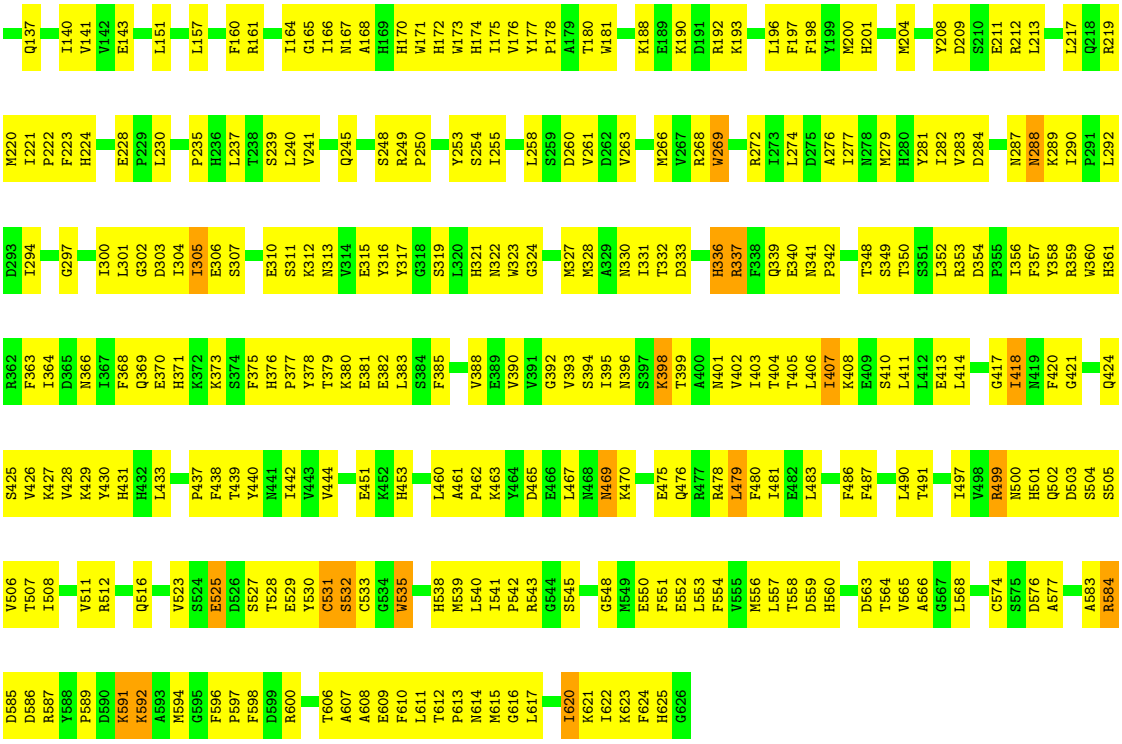


• Molecule 1: Hemocyanin AA6 chain

Chain F:

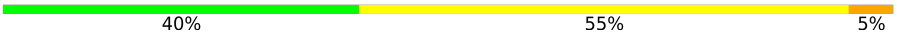


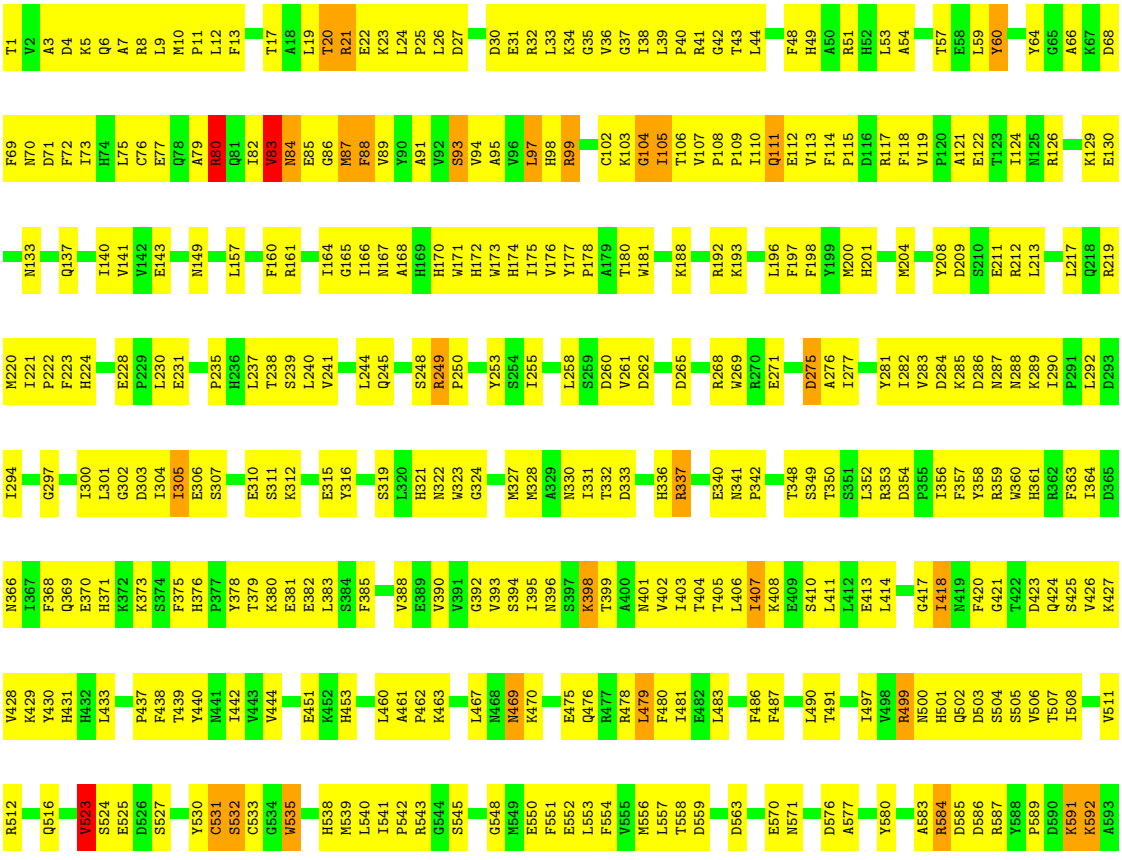


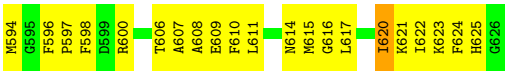


• Molecule 1: Hemocyanin AA6 chain

Chain G:

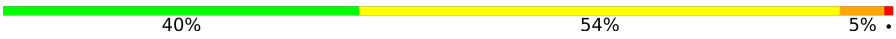


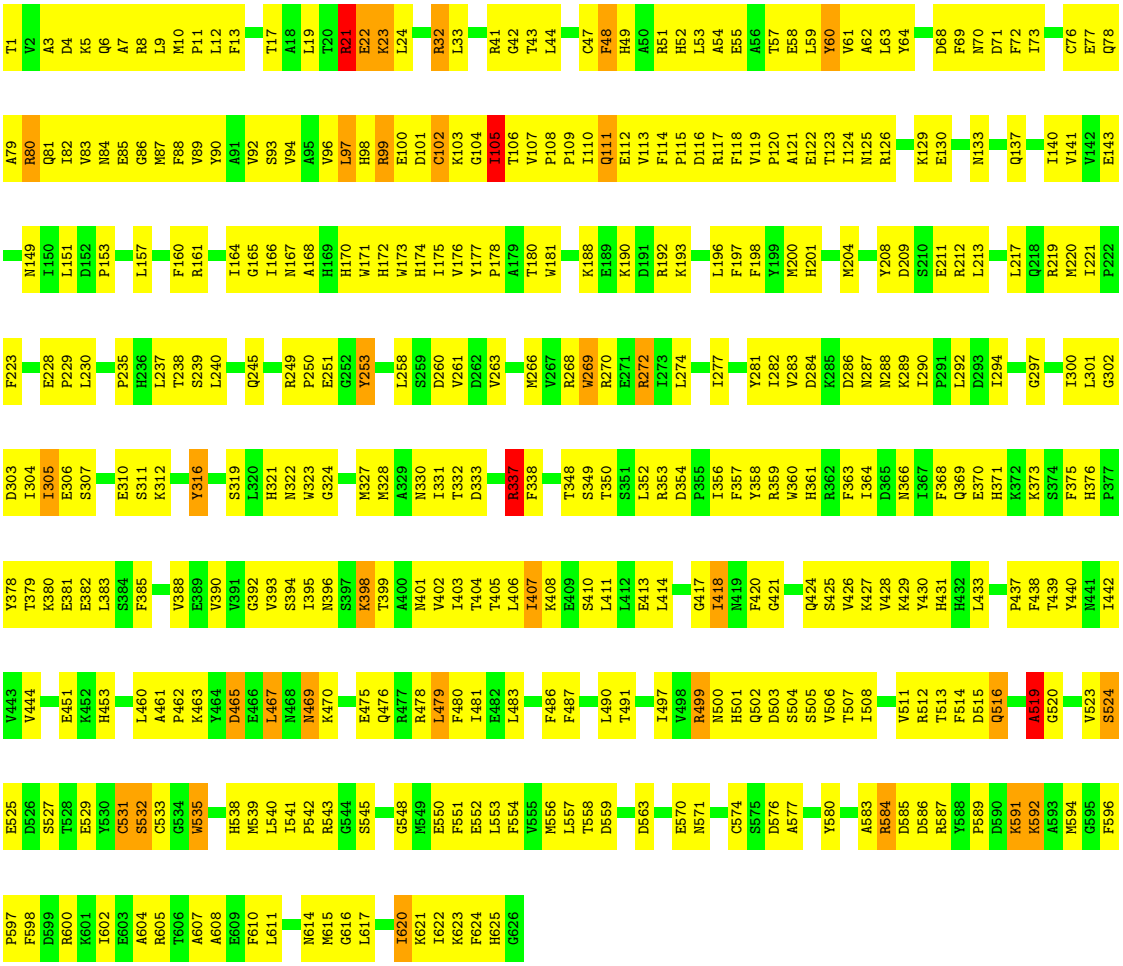




• Molecule 1: Hemocyanin AA6 chain

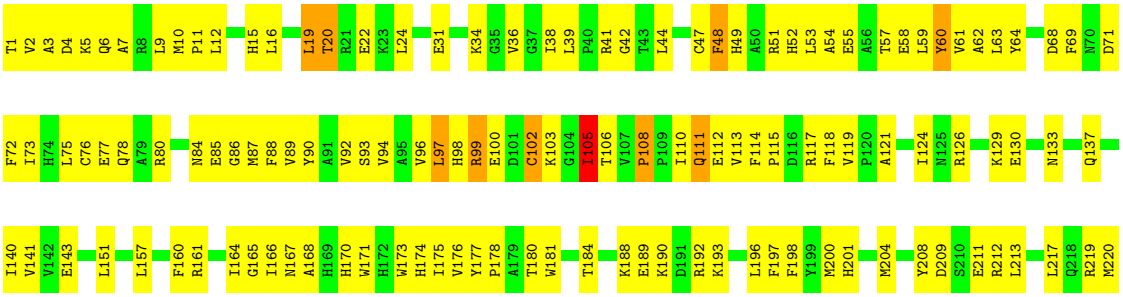
Chain H:

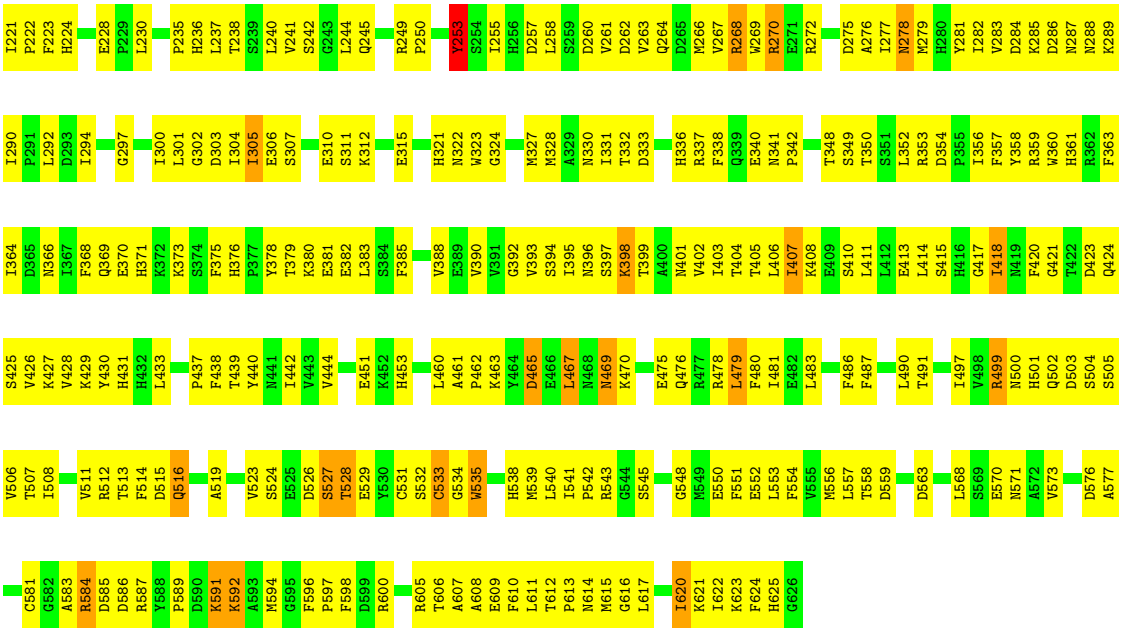




• Molecule 1: Hemocyanin AA6 chain

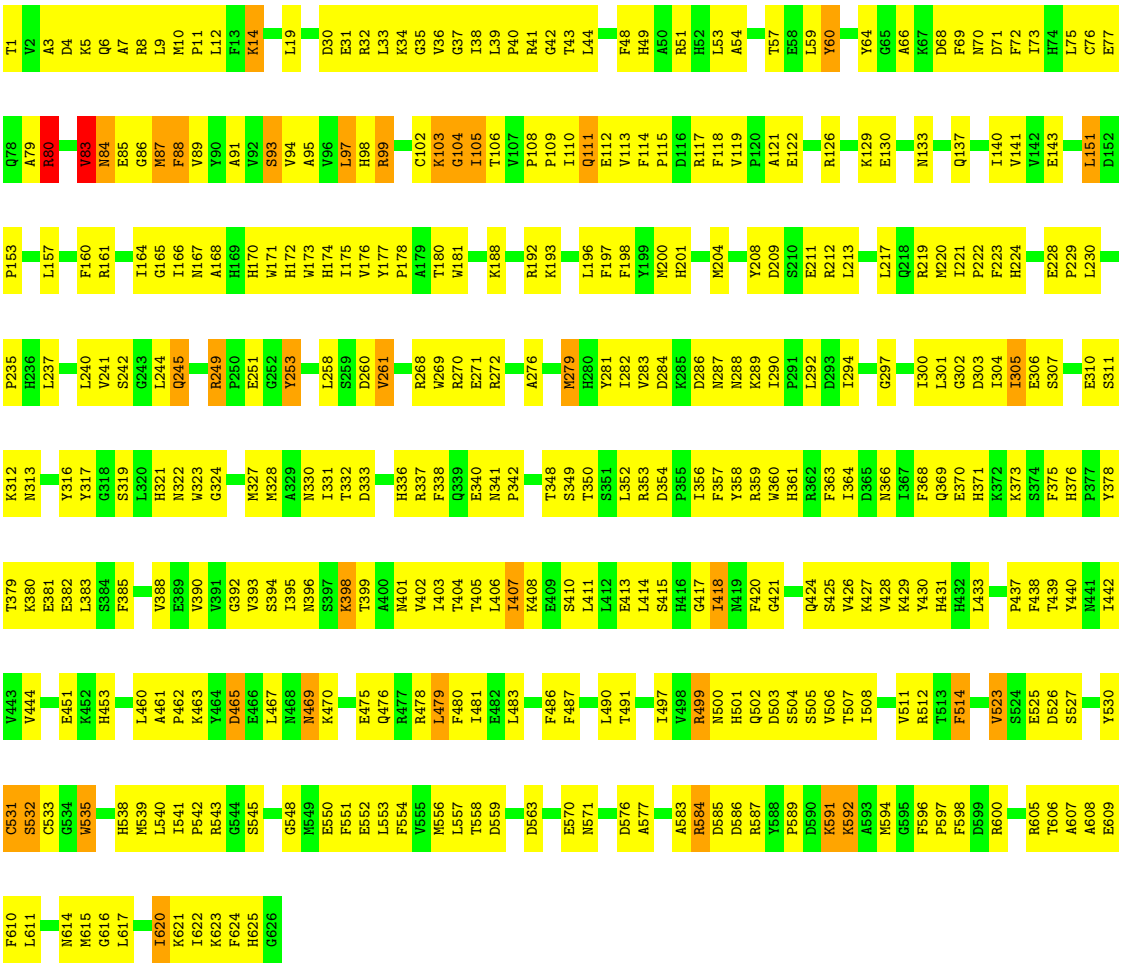
Chain I:





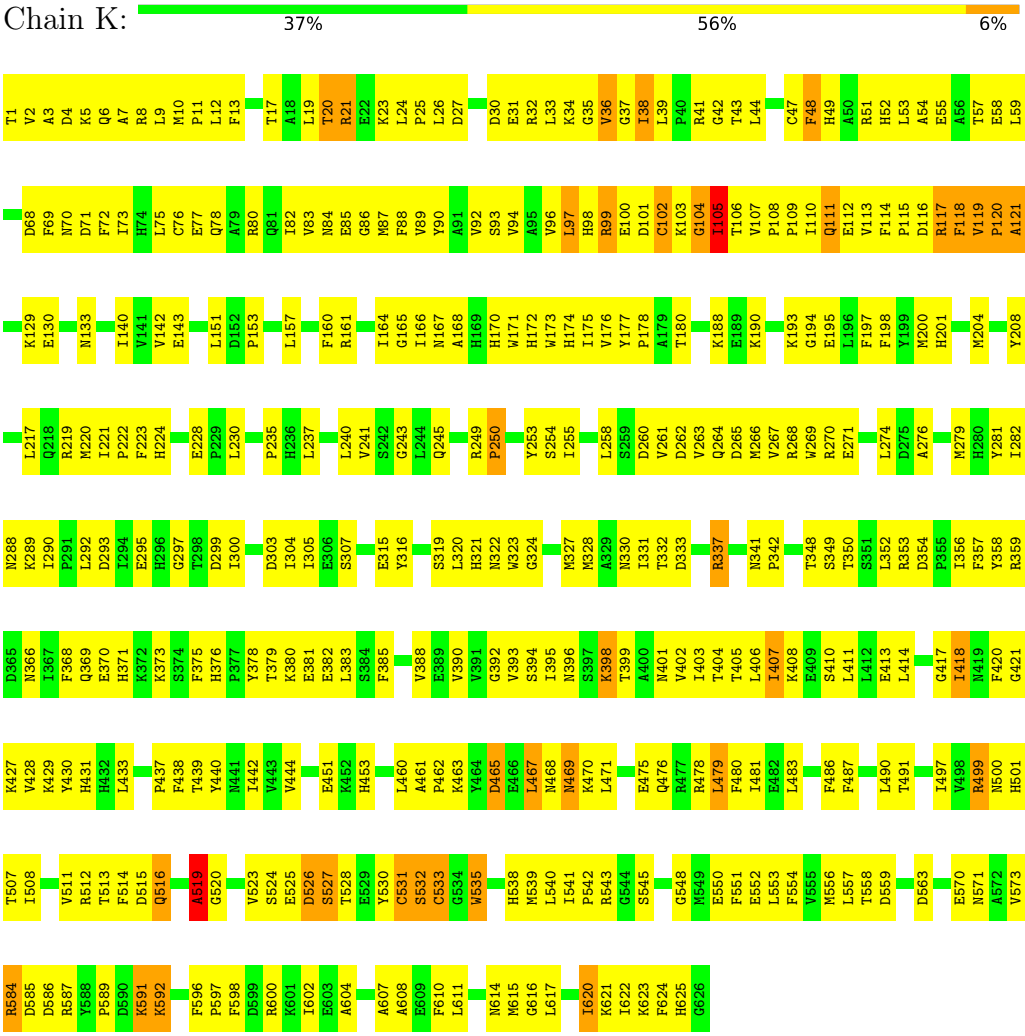
• Molecule 1: Hemocyanin AA6 chain

Chain J:



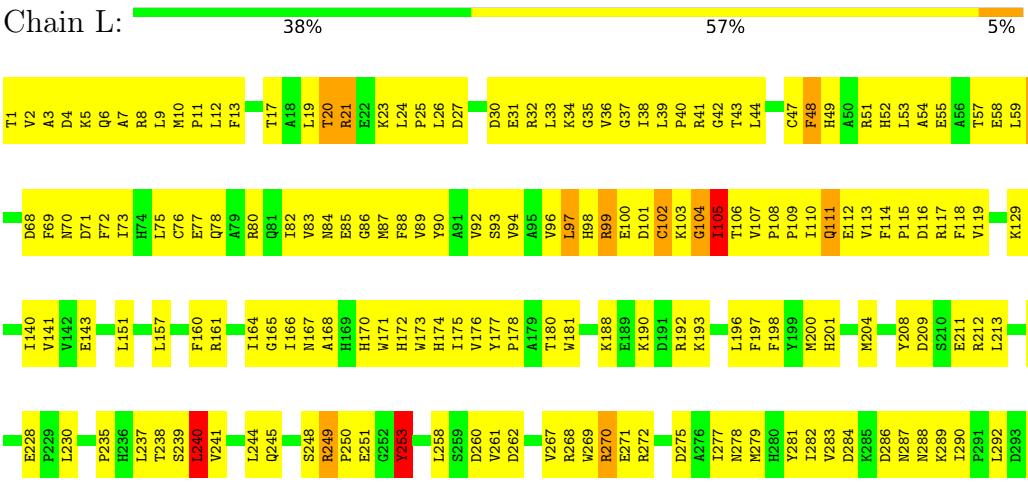
• Molecule 1: Hemocyanin AA6 chain

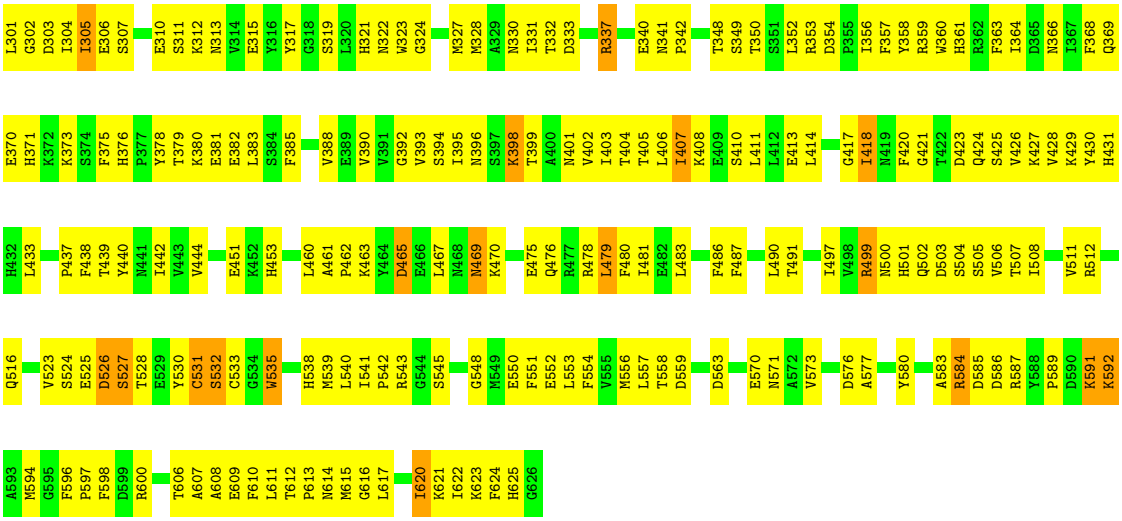
Chain K:



• Molecule 1: Hemocyanin AA6 chain

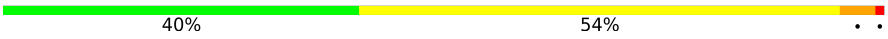
Chain L:

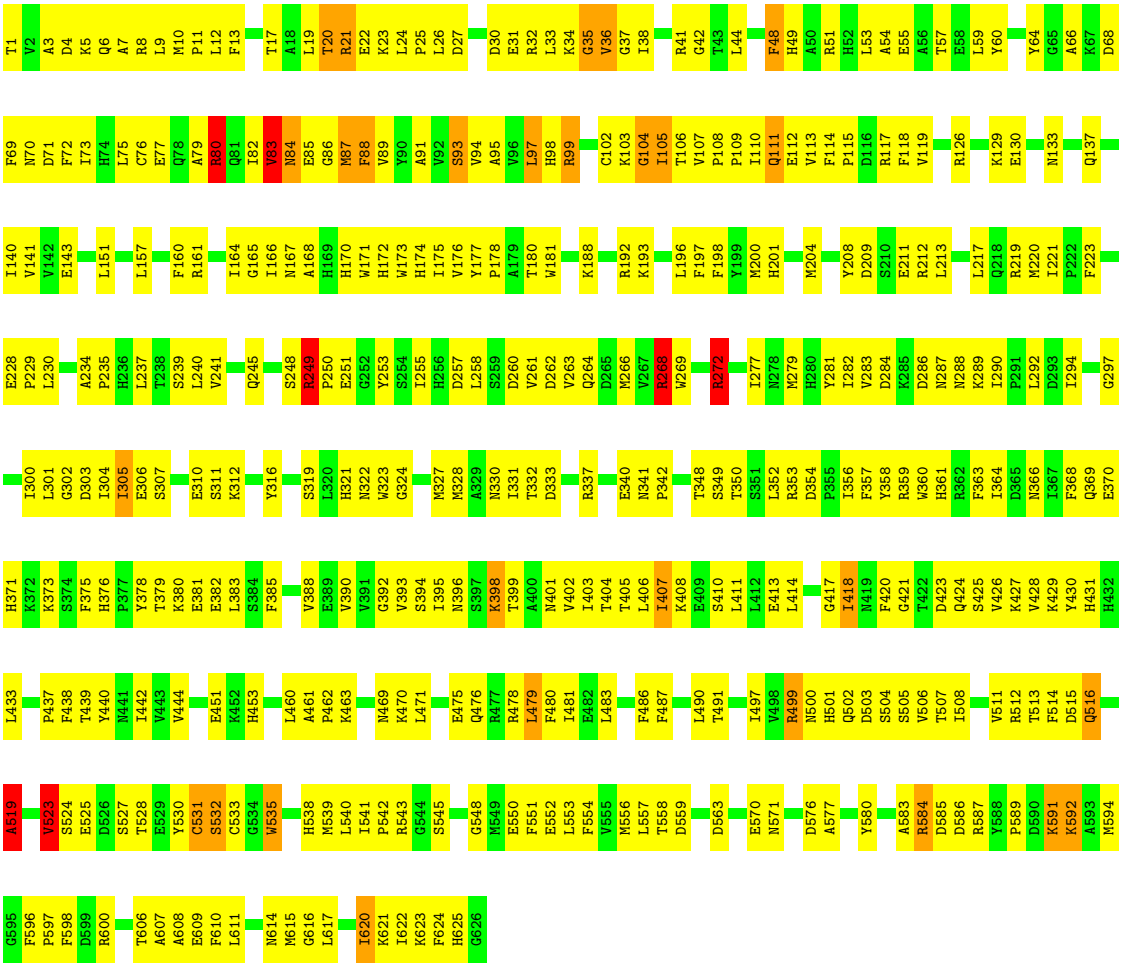




• Molecule 1: Hemocyanin AA6 chain

Chain M:





4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C2	Depositor
Number of particles used	17500	Depositor
Resolution determination method	FSC 0.5 CUT-OFF	Depositor
CTF correction method	each micrograph	Depositor
Microscope	JEOL 3200FSC	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	18	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	3500	Depositor
Magnification	50000	Depositor
Image detector	KODAK SO-163 FILM	Depositor

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	0.94	1/5191 (0.0%)	1.19	21/7033 (0.3%)
1	C	0.90	0/5191	1.21	30/7033 (0.4%)
1	D	1.75	6/5191 (0.1%)	1.18	22/7033 (0.3%)
1	E	1.32	3/5191 (0.1%)	1.24	27/7033 (0.4%)
1	F	0.90	0/5191	1.18	23/7033 (0.3%)
1	G	0.89	0/5191	1.18	19/7033 (0.3%)
1	H	1.73	7/5191 (0.1%)	1.24	28/7033 (0.4%)
1	I	1.32	3/5191 (0.1%)	1.21	25/7033 (0.4%)
1	J	0.90	0/5191	1.19	20/7033 (0.3%)
1	K	0.91	1/5191 (0.0%)	1.26	32/7033 (0.5%)
1	L	0.90	0/5191	1.20	23/7033 (0.3%)
1	M	0.91	1/5191 (0.0%)	1.24	27/7033 (0.4%)
All	All	1.16	22/62292 (0.0%)	1.21	297/84396 (0.4%)

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	A	0	3
1	C	0	5
1	D	0	2
1	E	0	6
1	F	0	6
1	G	0	10
1	H	0	7
1	I	0	4
1	J	0	11
1	K	0	6
1	L	0	8
1	M	0	13
All	All	0	81

The worst 5 of 22 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	I	253	TYR	CZ-OH	69.59	2.56	1.37
1	E	253	TYR	CZ-OH	68.02	2.53	1.37
1	D	253	TYR	CG-CD2	46.95	2.00	1.39
1	D	253	TYR	CE1-CZ	46.45	1.99	1.38
1	D	253	TYR	CE2-CZ	46.30	1.98	1.38

The worst 5 of 297 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	K	519	ALA	O-C-N	-20.49	88.36	123.20
1	H	519	ALA	O-C-N	-20.01	89.19	123.20
1	M	519	ALA	O-C-N	-19.78	89.58	123.20
1	A	47	CYS	O-C-N	14.24	145.49	122.70
1	I	253	TYR	CE1-CZ-CE2	-12.27	100.17	119.80

There are no chirality outliers.

5 of 81 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	104	GLY	Peptide
1	A	270	ARG	Sidechain
1	A	523	VAL	Peptide
1	C	104	GLY	Peptide
1	C	253	TYR	Sidechain

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	5061	1123	4864	715	0
1	C	5061	1123	4865	656	0
1	D	5061	1123	4865	687	0
1	E	5061	1123	4865	685	0
1	F	5061	1123	4865	627	0
1	G	5061	1123	4865	611	0
1	H	5061	1123	4864	626	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	I	5061	1123	4865	645	0
1	J	5061	1123	4865	599	0
1	K	5061	1123	4864	604	0
1	L	5061	1123	4865	639	0
1	M	5061	1123	4864	603	0
All	All	60732	13476	58376	7547	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 63.

The worst 5 of 7547 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:H:253:TYR:CE1	1:H:253:TYR:CZ	1.96	1.54
1:H:253:TYR:CG	1:H:253:TYR:CD1	1.96	1.53
1:D:253:TYR:CD1	1:D:253:TYR:CG	1.97	1.53
1:D:253:TYR:CD1	1:D:253:TYR:CE1	1.97	1.52
1:H:253:TYR:CZ	1:H:253:TYR:CE2	1.98	1.52

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	A	624/626 (100%)	590 (95%)	30 (5%)	4 (1%)	25	66
1	C	624/626 (100%)	583 (93%)	32 (5%)	9 (1%)	11	46
1	D	624/626 (100%)	590 (95%)	30 (5%)	4 (1%)	25	66
1	E	624/626 (100%)	583 (93%)	34 (5%)	7 (1%)	14	52
1	F	624/626 (100%)	589 (94%)	31 (5%)	4 (1%)	25	66

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	G	624/626 (100%)	581 (93%)	41 (7%)	2 (0%)	41 77
1	H	624/626 (100%)	577 (92%)	41 (7%)	6 (1%)	15 54
1	I	624/626 (100%)	591 (95%)	29 (5%)	4 (1%)	25 66
1	J	624/626 (100%)	587 (94%)	34 (5%)	3 (0%)	29 69
1	K	624/626 (100%)	576 (92%)	38 (6%)	10 (2%)	9 44
1	L	624/626 (100%)	586 (94%)	33 (5%)	5 (1%)	19 60
1	M	624/626 (100%)	584 (94%)	37 (6%)	3 (0%)	29 69
All	All	7488/7512 (100%)	7017 (94%)	410 (6%)	61 (1%)	24 60

5 of 61 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	20	THR
1	A	524	SER
1	C	20	THR
1	C	34	LYS
1	C	524	SER

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 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	A	556/556 (100%)	537 (97%)	19 (3%)	37 60
1	C	556/556 (100%)	534 (96%)	22 (4%)	31 55
1	D	556/556 (100%)	534 (96%)	22 (4%)	31 55
1	E	556/556 (100%)	534 (96%)	22 (4%)	31 55
1	F	556/556 (100%)	536 (96%)	20 (4%)	35 59
1	G	556/556 (100%)	535 (96%)	21 (4%)	33 57
1	H	556/556 (100%)	534 (96%)	22 (4%)	31 55
1	I	556/556 (100%)	532 (96%)	24 (4%)	29 53
1	J	556/556 (100%)	533 (96%)	23 (4%)	30 55

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	K	556/556 (100%)	532 (96%)	24 (4%)	29	53
1	L	556/556 (100%)	537 (97%)	19 (3%)	37	60
1	M	556/556 (100%)	537 (97%)	19 (3%)	37	60
All	All	6672/6672 (100%)	6415 (96%)	257 (4%)	36	56

5 of 257 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	L	469	ASN
1	M	83	VAL
1	F	469	ASN
1	F	407	ILE
1	M	105	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 146 such sidechains are listed below:

Mol	Chain	Res	Type
1	K	81	GLN
1	M	371	HIS
1	K	224	HIS
1	L	172	HIS
1	E	138	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)

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)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	K	1
1	H	1
1	M	1
1	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	K	519:ALA	C	520:GLY	N	1.14
1	H	519:ALA	C	520:GLY	N	1.10
1	M	519:ALA	C	520:GLY	N	1.05
1	A	47:CYS	C	48:PHE	N	0.99

6 Map visualisation

This section contains visualisations of the EMDB entry EMD-5100. These allow visual inspection of the internal detail of the map and identification of artifacts.

No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

This section was not generated.

6.2 Central slices

This section was not generated.

6.3 Largest variance slices

This section was not generated.

6.4 Orthogonal surface views

This section was not generated.

6.5 Mask visualisation

This section was not generated. No masks/segmentation were deposited.

7 Map analysis

This section contains the results of statistical analysis of the map.

7.1 Map-value distribution

This section was not generated.

7.2 Volume estimate versus contour level

This section was not generated.

7.3 Rotationally averaged power spectrum

This section was not generated. The rotationally averaged power spectrum had issues being displayed.

8 Fourier-Shell correlation

This section was not generated. No FSC curve or half-maps provided.

9 Map-model fit

This section was not generated.