wwPDB X-ray Structure Validation Summary Report (i)

Aug 29, 2020 – 06:21 PM BST

PDB ID	:	6QD6
Title	:	Molecular scaffolds expand the nanobody toolkit for cryo-EM applications: crystal structure of Mb-cHopQ-Nb207
Authors	:	Uchanski, T.; Masiulis, S.; Fischer, B.; Kalichuk, V.; Wohlkonig, A.; Zogg, T.; Remaut, H.; Vranken, W.; Aricescu, A.R.; Pardon, E.; Steyaert, J.
Deposited on	:	2018-12-31
Resolution	:	2.84 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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

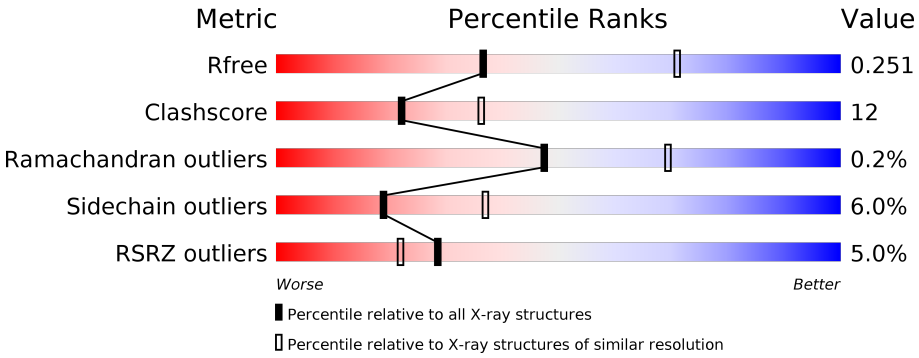
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.13
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.84 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.



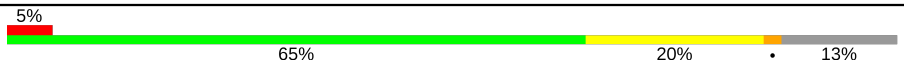
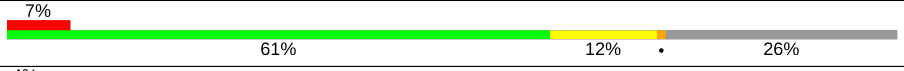
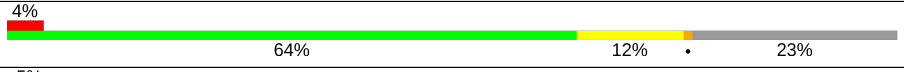
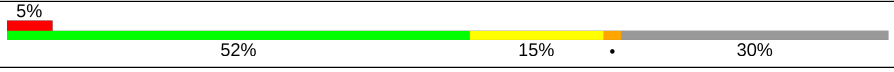
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1031 (2.86-2.82)
Clashscore	141614	1078 (2.86-2.82)
Ramachandran outliers	138981	1050 (2.86-2.82)
Sidechain outliers	138945	1051 (2.86-2.82)
RSRZ outliers	127900	1019 (2.86-2.82)

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 on 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 <=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	542	2% 69% 13% • 17%
1	B	542	5% 68% 13% • 18%
1	C	542	3% 59% 19% • 20%
1	D	542	2% 68% 13% • 18%
1	E	542	4% 64% 17% • 17%
1	F	542	3% 65% 18% • 16%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	542	
1	H	542	
1	I	542	
1	J	542	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 33077 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 Mb-cHopQ-Nb207,Outer membrane protein,Mb-cHopQ-Nb207,Outer membrane protein,Mb-cHopQ-Nb207.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	448	Total	C	N	O	S	0	0	0
			3409	2107	603	686	13			
1	B	446	Total	C	N	O	S	0	0	0
			3372	2081	594	682	15			
1	C	431	Total	C	N	O	S	0	0	0
			3263	2015	573	660	15			
1	D	446	Total	C	N	O	S	0	0	0
			3384	2090	598	683	13			
1	E	451	Total	C	N	O	S	0	0	0
			3406	2103	599	689	15			
1	F	457	Total	C	N	O	S	0	0	0
			3465	2144	606	702	13			
1	G	469	Total	C	N	O	S	0	0	0
			3561	2197	626	723	15			
1	H	399	Total	C	N	O	S	0	0	0
			3015	1868	525	609	13			
1	I	419	Total	C	N	O	S	0	0	0
			3187	1976	558	640	13			
1	J	380	Total	C	N	O	S	0	0	0
			2900	1795	508	584	13			

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Cl	0	0
			1	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	18	Total	O	0	0
			18	18		

Continued on next page...

Continued from previous page...

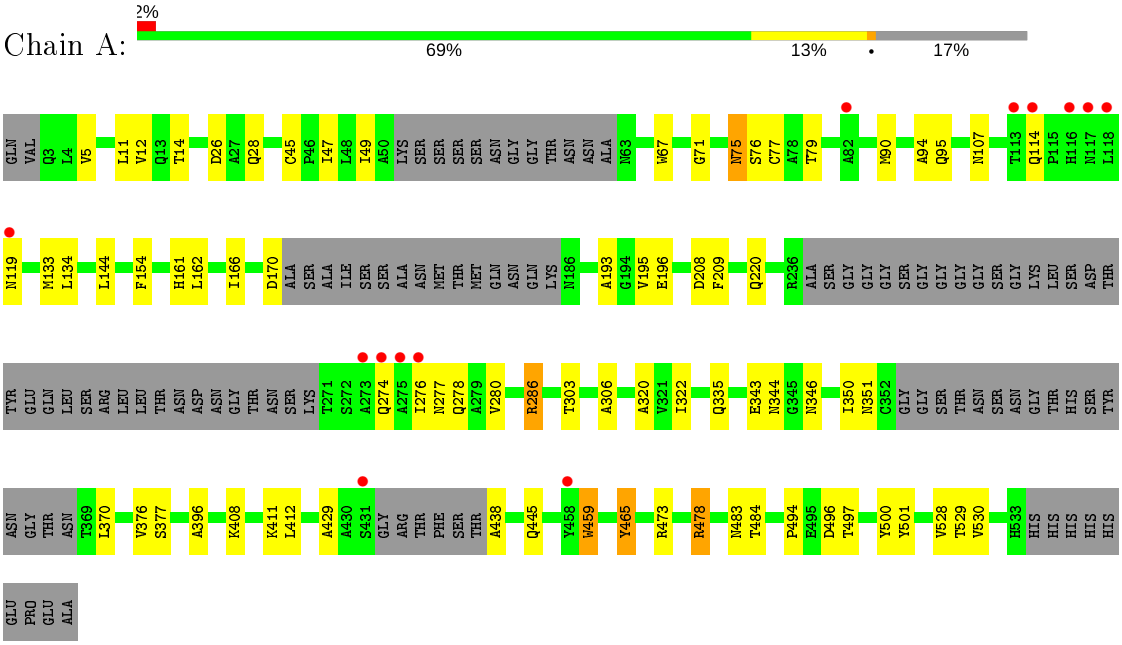
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	15	Total O 15 15	0	0
3	C	11	Total O 11 11	0	0
3	D	12	Total O 12 12	0	0
3	E	8	Total O 8 8	0	0
3	F	14	Total O 14 14	0	0
3	G	14	Total O 14 14	0	0
3	H	5	Total O 5 5	0	0
3	I	12	Total O 12 12	0	0
3	J	5	Total O 5 5	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

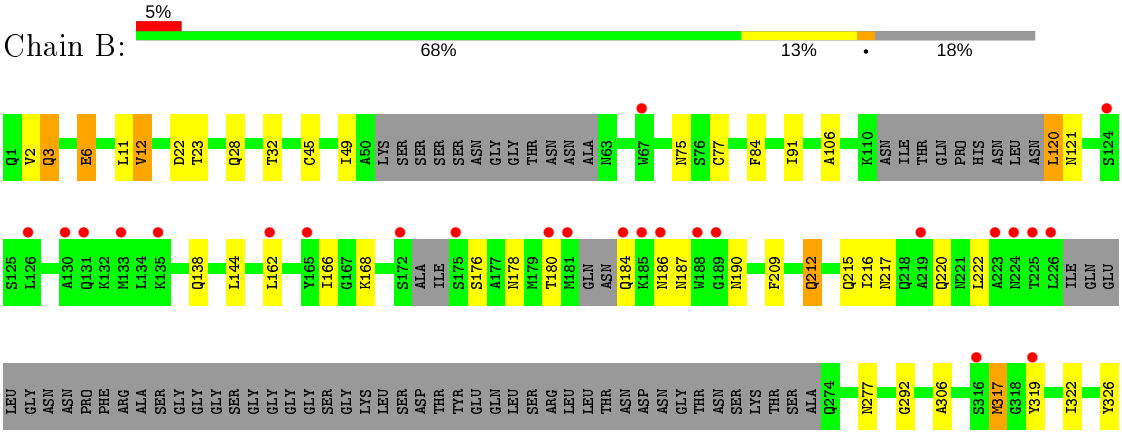
• Molecule 1: Mb-cHopQ-Nb207,Outer membrane protein,Mb-cHopQ-Nb207,Outer membrane protein,Mb-cHopQ-Nb207

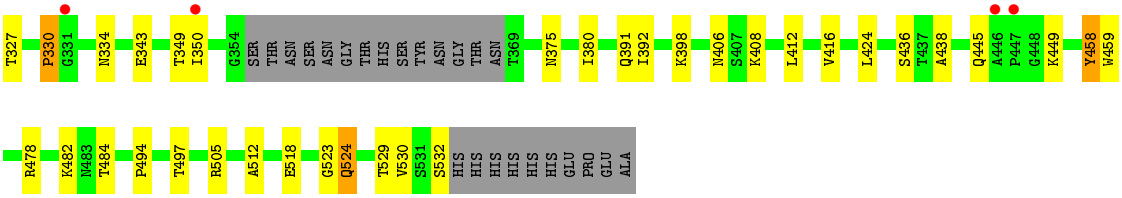
Chain A:



• Molecule 1: Mb-cHopQ-Nb207,Outer membrane protein,Mb-cHopQ-Nb207,Outer membrane protein,Mb-cHopQ-Nb207

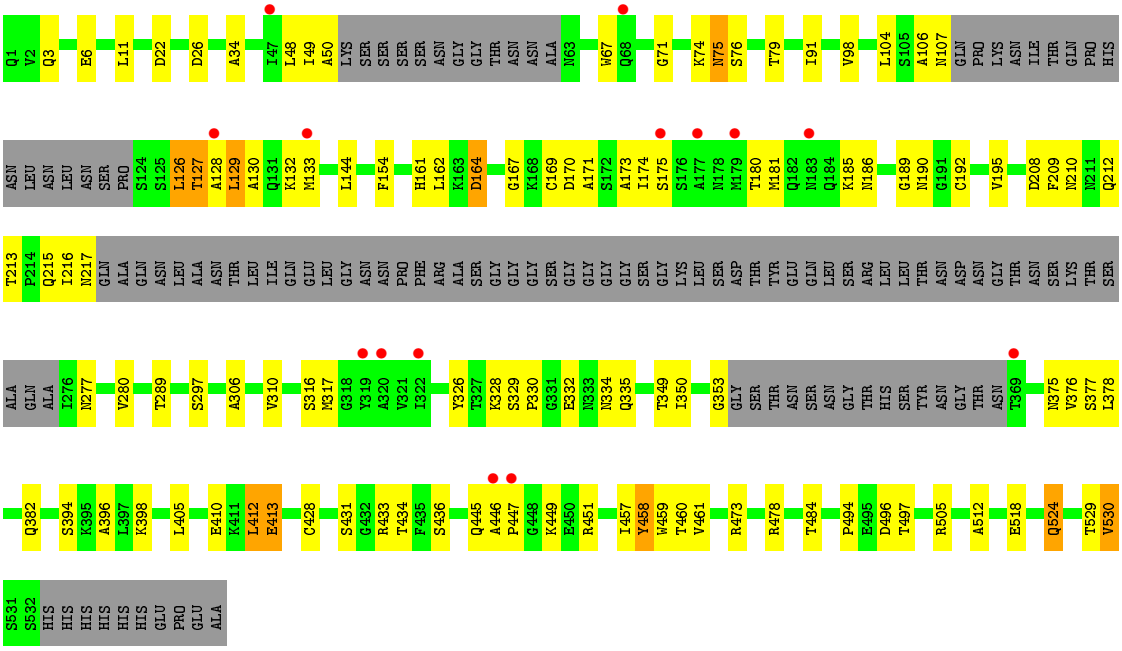
Chain B:





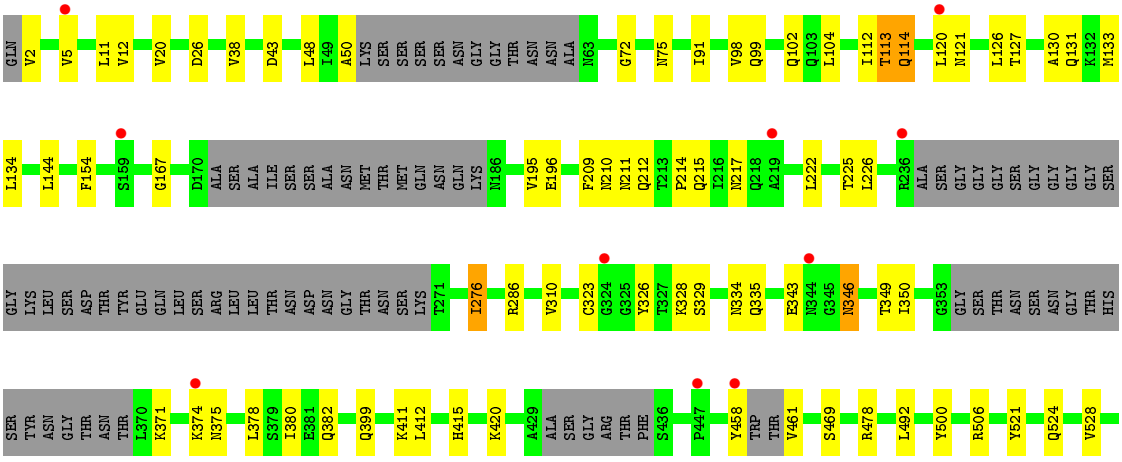
• Molecule 1: Mb-cHopQ-Nb207,Outer membrane protein,Mb-cHopQ-Nb207,Outer membrane pr otein,Mb-cHopQ-Nb207

Chain C:



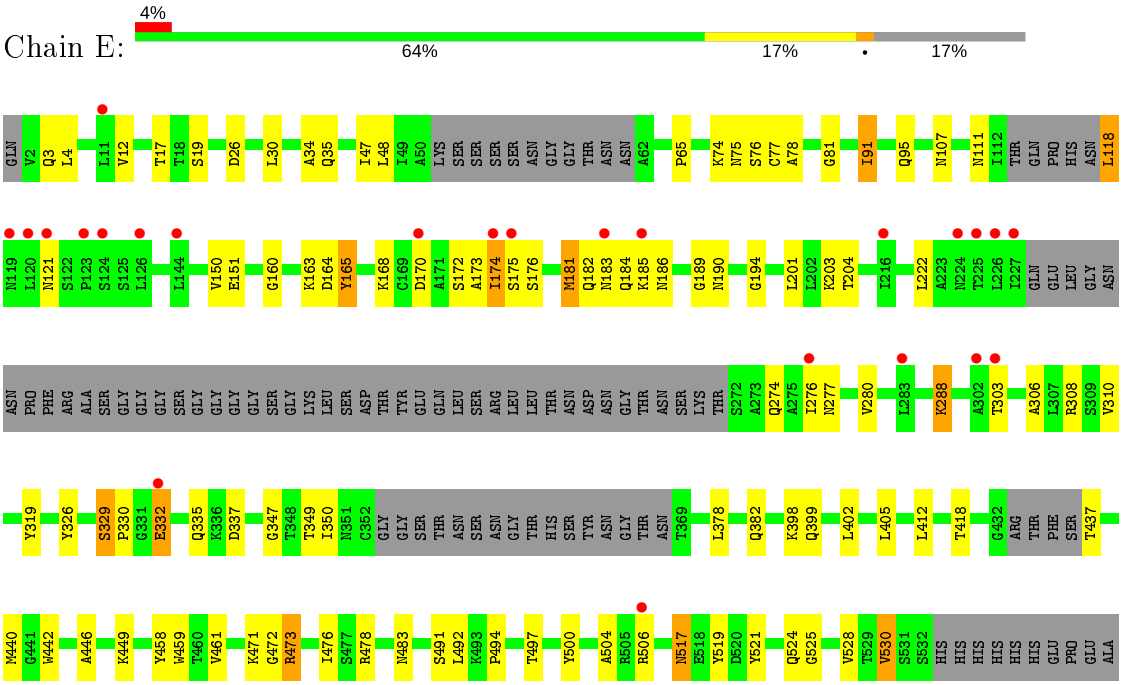
• Molecule 1: Mb-cHopQ-Nb207,Outer membrane protein,Mb-cHopQ-Nb207,Outer membrane pr otein,Mb-cHopQ-Nb207

Chain D:



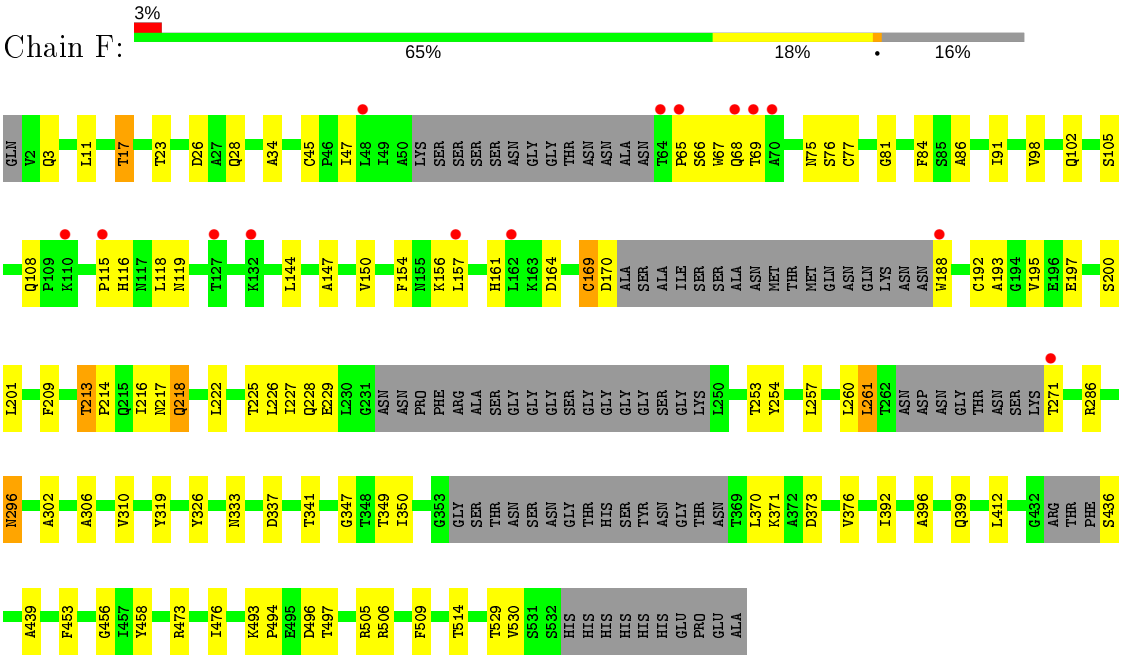
• Molecule 1: Mb-cHopQ-Nb207,Outer membrane protein,Mb-cHopQ-Nb207,Outer membrane pr otein,Mb-cHopQ-Nb207

Chain E:



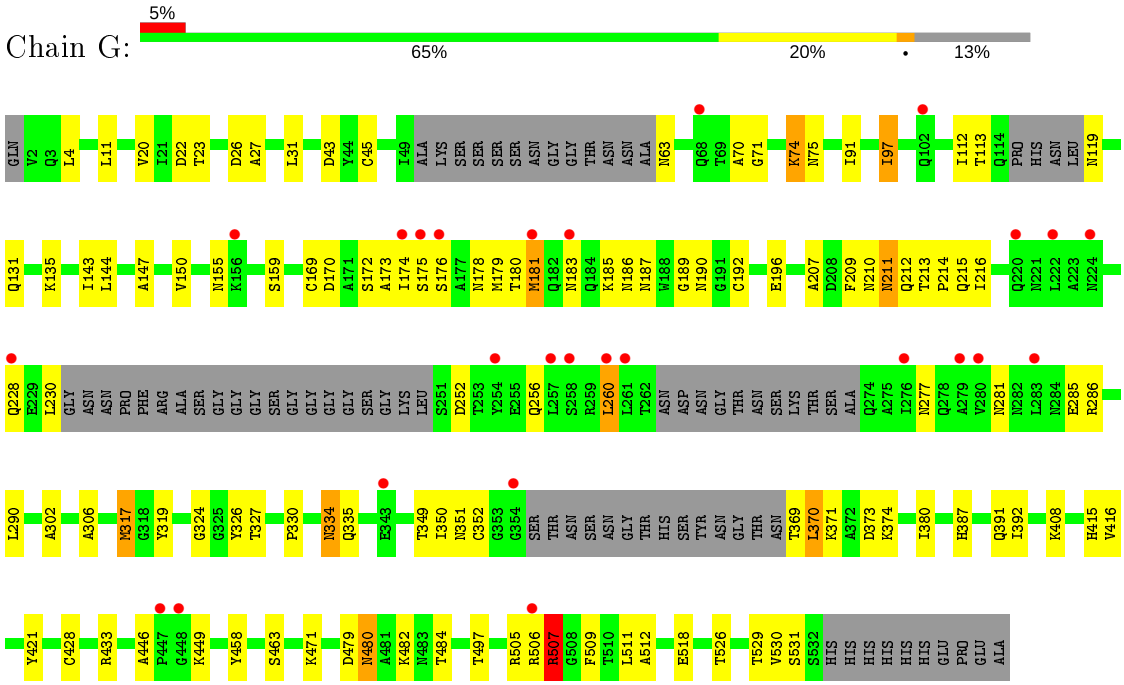
• Molecule 1: Mb-cHopQ-Nb207,Outer membrane protein,Mb-cHopQ-Nb207,Outer membrane pr otein,Mb-cHopQ-Nb207

Chain F:



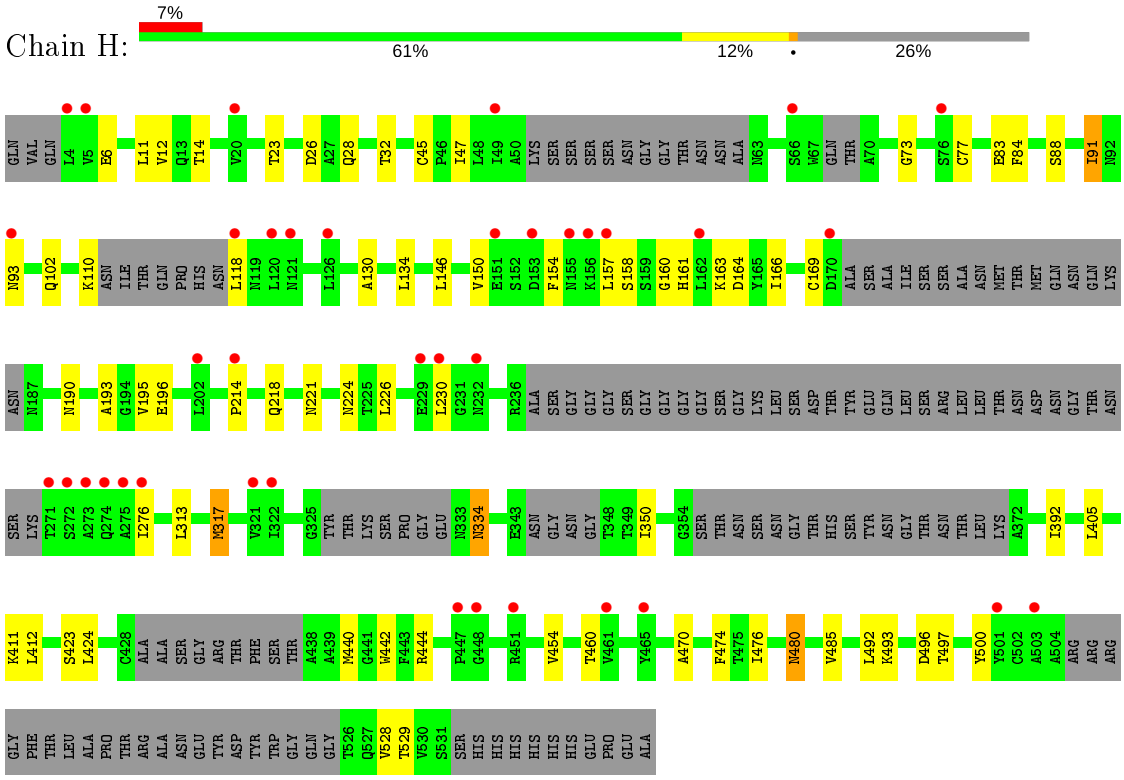
• Molecule 1: Mb-cHopQ-Nb207,Outer membrane protein,Mb-cHopQ-Nb207,Outer membrane pr otein,Mb-cHopQ-Nb207

Chain G:



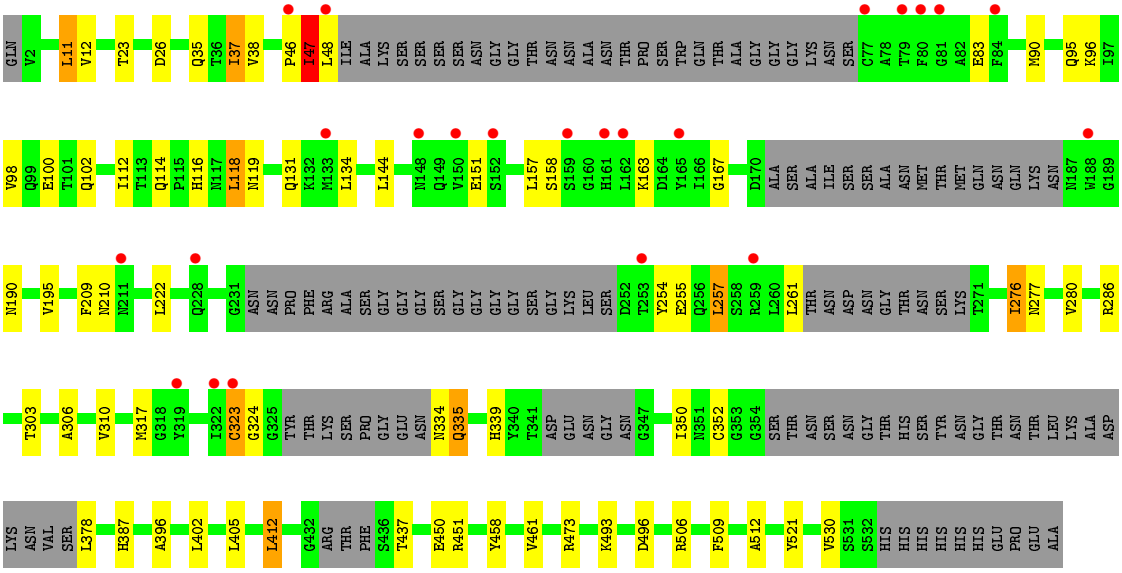
• Molecule 1: Mb-cHopQ-Nb207,Outer membrane protein,Mb-cHopQ-Nb207,Outer membrane pr otein,Mb-cHopQ-Nb207

Chain H:



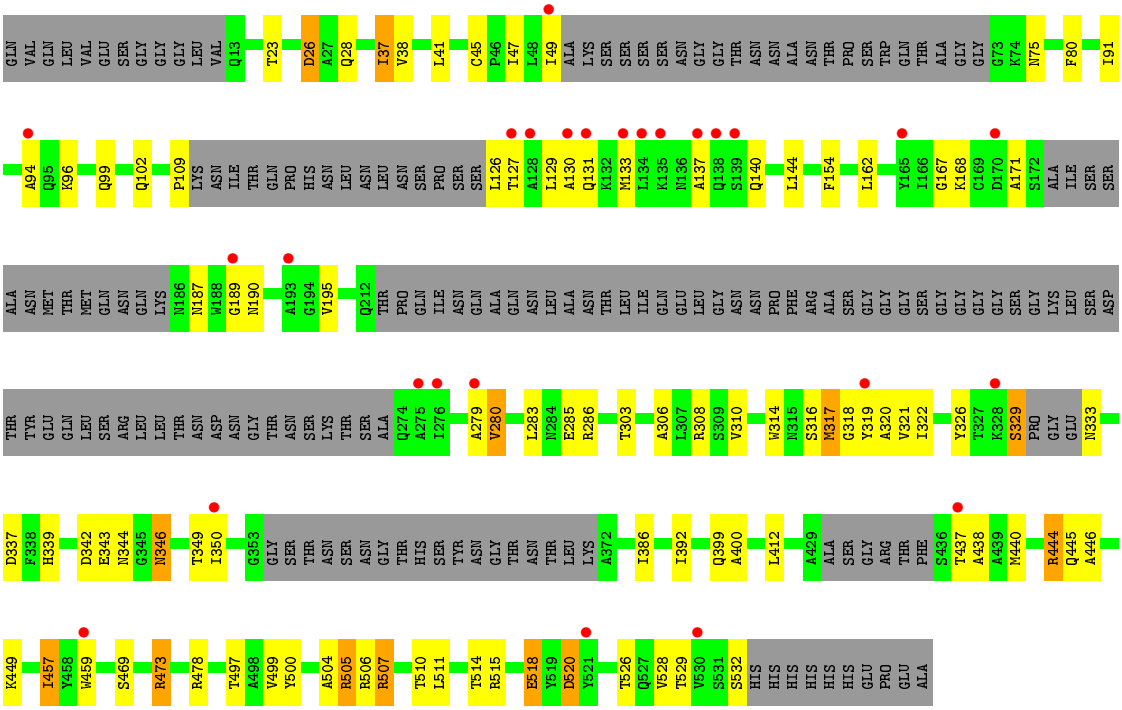
• Molecule 1: Mb-cHopQ-Nb207,Outer membrane protein,Mb-cHopQ-Nb207,Outer membrane pr otein,Mb-cHopQ-Nb207

Chain I:



• Molecule 1: Mb-cHopQ-Nb207,Outer membrane protein,Mb-cHopQ-Nb207,Outer membrane pr otein,Mb-cHopQ-Nb207

Chain J:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	71.17Å 92.92Å 244.22Å 92.05° 96.93° 112.15°	Depositor
Resolution (Å)	41.45 – 2.84 44.77 – 2.84	Depositor EDS
% Data completeness (in resolution range)	95.5 (41.45-2.84) 95.4 (44.77-2.84)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.86 (at 2.86Å)	Xtriage
Refinement program	BUSTER 2.10.3	Depositor
R, R_{free}	0.204 , 0.239 0.225 , 0.251	Depositor DCC
R_{free} test set	6538 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	82.3	Xtriage
Anisotropy	0.101	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 72.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.039 for h,-h-k,-h-l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	33077	wwPDB-VP
Average B, all atoms (Å^2)	99.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.40% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: CL

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	0.50	0/3461	0.73	0/4690
1	B	0.56	0/3419	0.74	0/4625
1	C	0.55	0/3310	0.77	0/4479
1	D	0.51	0/3432	0.74	0/4647
1	E	0.50	0/3453	0.73	0/4676
1	F	0.55	0/3514	0.74	0/4760
1	G	0.52	0/3610	0.74	0/4887
1	H	0.48	0/3051	0.72	0/4124
1	I	0.48	0/3229	0.72	0/4369
1	J	0.53	0/2939	0.73	0/3971
All	All	0.52	0/33418	0.74	0/45228

There are no bond length outliers.

There are no bond angle outliers.

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	A	3409	0	3318	64	0
1	B	3372	0	3291	62	0
1	C	3263	0	3181	77	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	3384	0	3298	48	0
1	E	3406	0	3329	96	0
1	F	3465	0	3388	85	0
1	G	3561	0	3477	122	0
1	H	3015	0	2948	57	0
1	I	3187	0	3118	56	0
1	J	2900	0	2817	104	0
2	A	1	0	0	0	0
3	A	18	0	0	0	0
3	B	15	0	0	0	0
3	C	11	0	0	0	0
3	D	12	0	0	1	0
3	E	8	0	0	3	0
3	F	14	0	0	0	0
3	G	14	0	0	0	0
3	H	5	0	0	0	0
3	I	12	0	0	2	0
3	J	5	0	0	0	0
All	All	33077	0	32165	761	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 12.

The worst 5 of 761 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:J:326:TYR:CD1	1:J:329:SER:HB3	1.50	1.46
1:F:45:CYS:SG	1:F:77:CYS:CB	2.23	1.27
1:J:326:TYR:CG	1:J:329:SER:HB3	1.70	1.26
1:A:45:CYS:SG	1:A:77:CYS:SG	1.45	1.24
1:F:45:CYS:CB	1:F:77:CYS:SG	2.25	1.23

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	A	436/542 (80%)	417 (96%)	19 (4%)	0	100	100
1	B	432/542 (80%)	412 (95%)	19 (4%)	1 (0%)	47	69
1	C	421/542 (78%)	398 (94%)	21 (5%)	2 (0%)	29	51
1	D	432/542 (80%)	417 (96%)	15 (4%)	0	100	100
1	E	439/542 (81%)	422 (96%)	17 (4%)	0	100	100
1	F	443/542 (82%)	424 (96%)	18 (4%)	1 (0%)	47	69
1	G	457/542 (84%)	441 (96%)	15 (3%)	1 (0%)	47	69
1	H	377/542 (70%)	364 (97%)	12 (3%)	1 (0%)	41	61
1	I	401/542 (74%)	388 (97%)	12 (3%)	1 (0%)	47	69
1	J	364/542 (67%)	350 (96%)	14 (4%)	0	100	100
All	All	4202/5420 (78%)	4033 (96%)	162 (4%)	7 (0%)	47	69

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	H	161	HIS
1	I	47	ILE
1	C	75	ASN
1	F	333	ASN
1	G	507	ARG

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	A	366/438 (84%)	351 (96%)	15 (4%)	30	56
1	B	361/438 (82%)	342 (95%)	19 (5%)	22	43
1	C	349/438 (80%)	325 (93%)	24 (7%)	15	31
1	D	364/438 (83%)	342 (94%)	22 (6%)	19	37

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	E	365/438 (83%)	346 (95%)	19 (5%)	23	44
1	F	373/438 (85%)	354 (95%)	19 (5%)	24	45
1	G	384/438 (88%)	359 (94%)	25 (6%)	17	33
1	H	325/438 (74%)	305 (94%)	20 (6%)	18	35
1	I	342/438 (78%)	318 (93%)	24 (7%)	15	30
1	J	309/438 (70%)	284 (92%)	25 (8%)	11	24
All	All	3538/4380 (81%)	3326 (94%)	212 (6%)	19	37

5 of 212 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	478	ARG
1	G	11	LEU
1	J	308	ARG
1	E	530	VAL
1	F	213	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 109 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	281	ASN
1	E	277	ASN
1	J	35	GLN
1	D	301	GLN
1	E	28	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

Of 1 ligands modelled in this entry, 1 is 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

There are no such residues in this entry.

5.8 Polymer linkage issues

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^{th} 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	A	448/542 (82%)	0.27	13 (2%)	51	45	56, 89, 132, 163	0
1	B	446/542 (82%)	0.30	29 (6%)	18	13	47, 89, 149, 197	0
1	C	431/542 (79%)	0.19	14 (3%)	47	41	46, 88, 138, 187	0
1	D	446/542 (82%)	0.28	10 (2%)	62	57	54, 89, 135, 160	0
1	E	451/542 (83%)	0.26	24 (5%)	26	20	61, 103, 160, 179	0
1	F	457/542 (84%)	0.18	14 (3%)	49	42	55, 88, 130, 157	0
1	G	469/542 (86%)	0.27	26 (5%)	25	18	59, 95, 144, 178	0
1	H	399/542 (73%)	0.45	38 (9%)	8	4	76, 110, 152, 175	0
1	I	419/542 (77%)	0.35	23 (5%)	25	18	63, 107, 151, 183	0
1	J	380/542 (70%)	0.35	26 (6%)	17	11	69, 110, 152, 184	0
All	All	4346/5420 (80%)	0.29	217 (4%)	28	22	46, 97, 146, 197	0

The worst 5 of 217 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	225	THR	6.3
1	B	185	LYS	6.1
1	H	272	SER	6.1
1	G	176	SER	6.0
1	B	184	GLN	5.4

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 (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} 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
2	CL	A	601	1/1	0.84	0.25	91,91,91,91	0

6.5 Other polymers (i)

There are no such residues in this entry.