



wwPDB X-ray Structure Validation Summary Report ⓘ

May 26, 2020 – 08:30 pm BST

PDB ID : 3O44
Title : Crystal Structure of the Vibrio cholerae Cytolysin (HlyA) Heptameric Pore
Authors : De, S.; Olson, R.
Deposited on : 2010-07-26
Resolution : 2.88 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
Xtriage (Phenix) : 1.13
EDS : 2.11
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) : 2.11

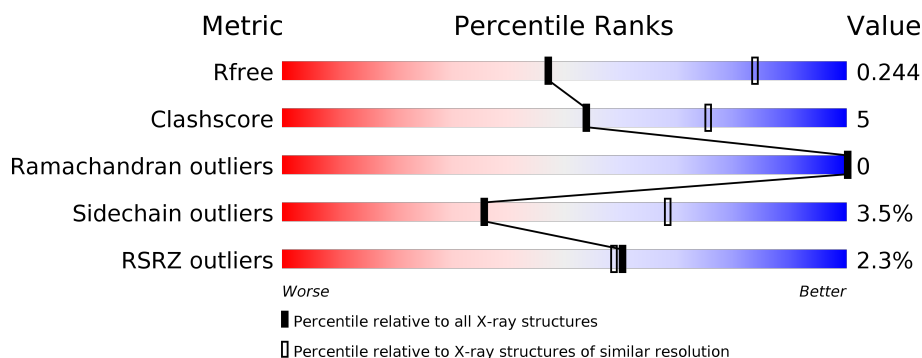
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.88 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	2691 (2.90-2.86)
Clashscore	141614	2947 (2.90-2.86)
Ramachandran outliers	138981	2868 (2.90-2.86)
Sidechain outliers	138945	2871 (2.90-2.86)
RSRZ outliers	127900	2629 (2.90-2.86)

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 $\leq 5\%$. 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	593	83% 14% ..
1	B	593	85% 12% ..
1	C	593	86% 11% ..
1	D	593	84% 13% ..
1	E	593	85% 12% ..
1	F	593	86% 11% ..

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	593	<p>5% 84% 13% ..</p>
1	H	593	<p>2% 85% 13% .</p>
1	I	593	<p>% 85% 13% ..</p>
1	J	593	<p>2% 85% 12% ..</p>
1	K	593	<p>5% 83% 14% ..</p>
1	L	593	<p>81% 16% ..</p>
1	M	593	<p>2% 83% 14% ..</p>
1	N	593	<p>4% 83% 14% ..</p>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 62948 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 Hemolysin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	581	4464	2784	779	892	9	0	0	0
1	B	581	4457	2783	778	887	9	0	0	0
1	C	581	4431	2759	776	887	9	0	0	0
1	D	581	4471	2794	778	890	9	0	0	0
1	E	581	4460	2784	777	890	9	0	0	0
1	F	581	4412	2746	768	889	9	0	0	0
1	G	581	4436	2764	777	886	9	0	0	0
1	H	581	4468	2791	774	894	9	0	0	0
1	I	581	4479	2797	780	893	9	0	0	0
1	J	581	4487	2804	782	892	9	0	0	0
1	K	581	4417	2755	772	881	9	0	0	0
1	L	581	4449	2773	779	888	9	0	0	0
1	M	581	4447	2773	777	888	9	0	0	0
1	N	581	4420	2759	769	883	9	0	0	0

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	606	ASN	HIS	CONFLICT	UNP C2C744

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	606	ASN	HIS	CONFLICT	UNP C2C744
C	606	ASN	HIS	CONFLICT	UNP C2C744
D	606	ASN	HIS	CONFLICT	UNP C2C744
E	606	ASN	HIS	CONFLICT	UNP C2C744
F	606	ASN	HIS	CONFLICT	UNP C2C744
G	606	ASN	HIS	CONFLICT	UNP C2C744
H	606	ASN	HIS	CONFLICT	UNP C2C744
I	606	ASN	HIS	CONFLICT	UNP C2C744
J	606	ASN	HIS	CONFLICT	UNP C2C744
K	606	ASN	HIS	CONFLICT	UNP C2C744
L	606	ASN	HIS	CONFLICT	UNP C2C744
M	606	ASN	HIS	CONFLICT	UNP C2C744
N	606	ASN	HIS	CONFLICT	UNP C2C744

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	58	Total O 58 58	0	0
2	B	62	Total O 62 62	0	0
2	C	57	Total O 57 57	0	0
2	D	65	Total O 65 65	0	0
2	E	44	Total O 44 44	0	0
2	F	26	Total O 26 26	0	0
2	G	50	Total O 50 50	0	0
2	H	48	Total O 48 48	0	0
2	I	61	Total O 61 61	0	0
2	J	32	Total O 32 32	0	0
2	K	25	Total O 25 25	0	0
2	L	43	Total O 43 43	0	0
2	M	43	Total O 43 43	0	0

Continued on next page...

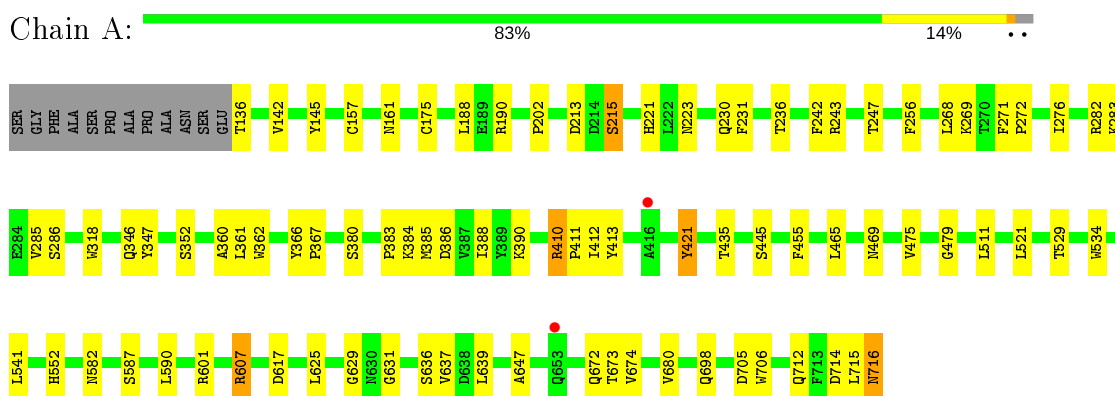
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	N	36	Total	O	0	0
			36	36		

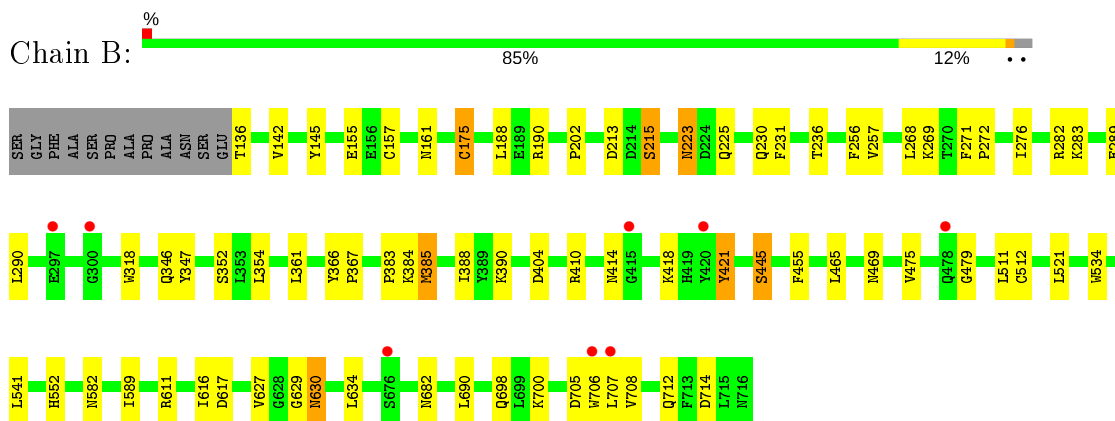
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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.

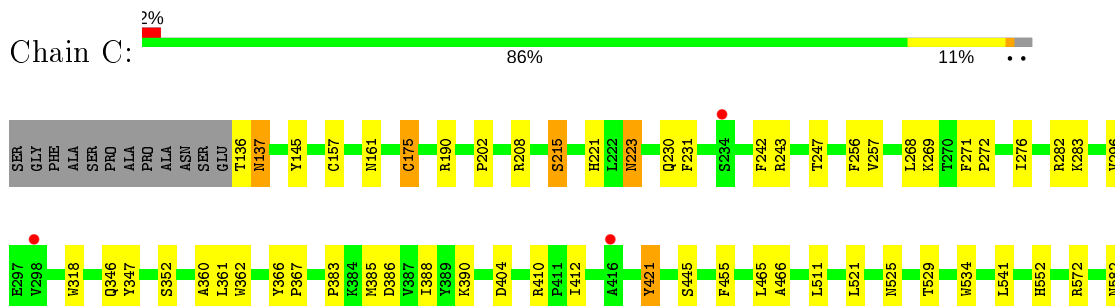
- Molecule 1: Hemolysin

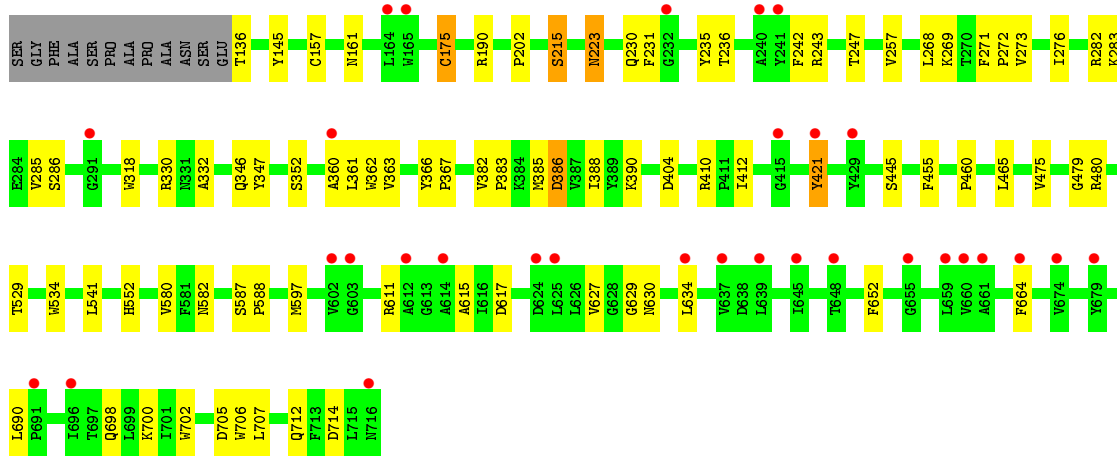


- Molecule 1: Hemolysin

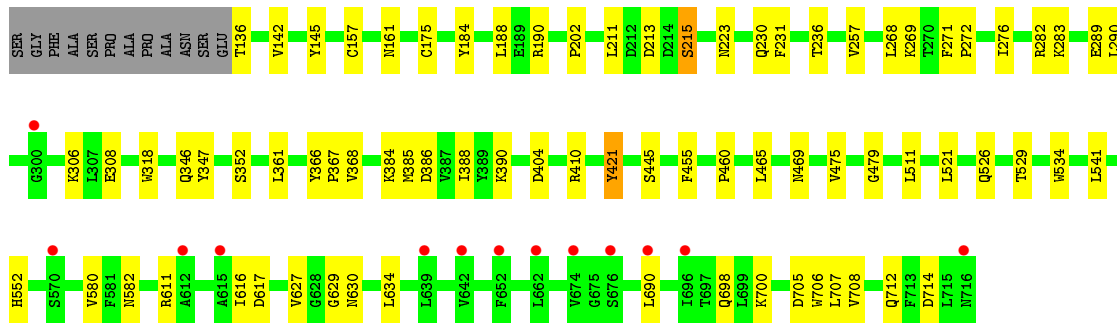
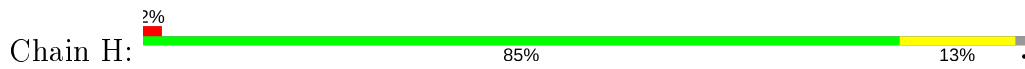


- Molecule 1: Hemolysin

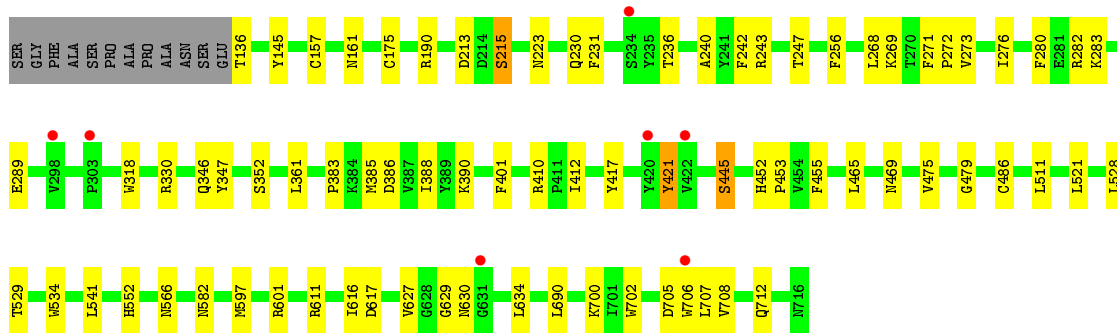
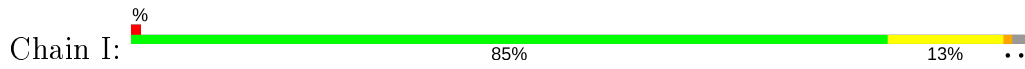




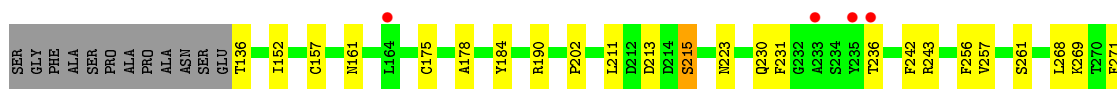
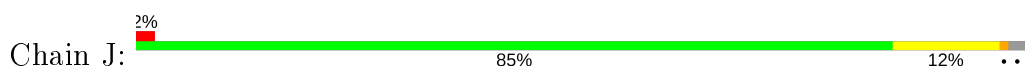
• Molecule 1: Hemolysin

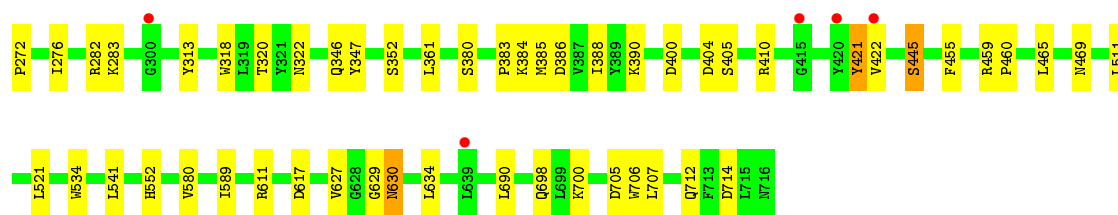


• Molecule 1: Hemolysin

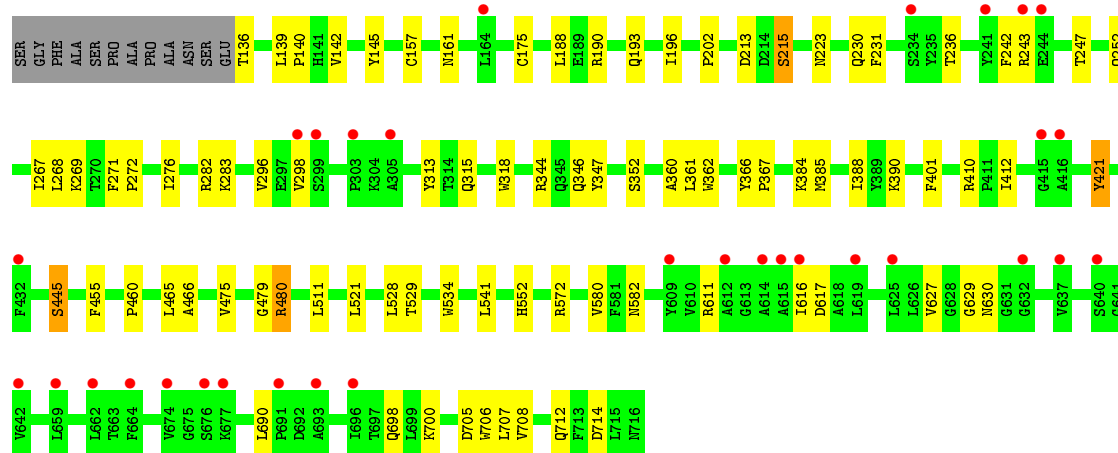
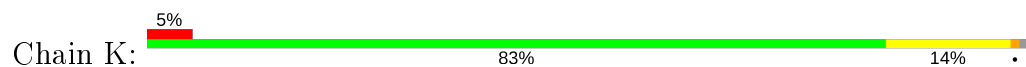


• Molecule 1: Hemolysin

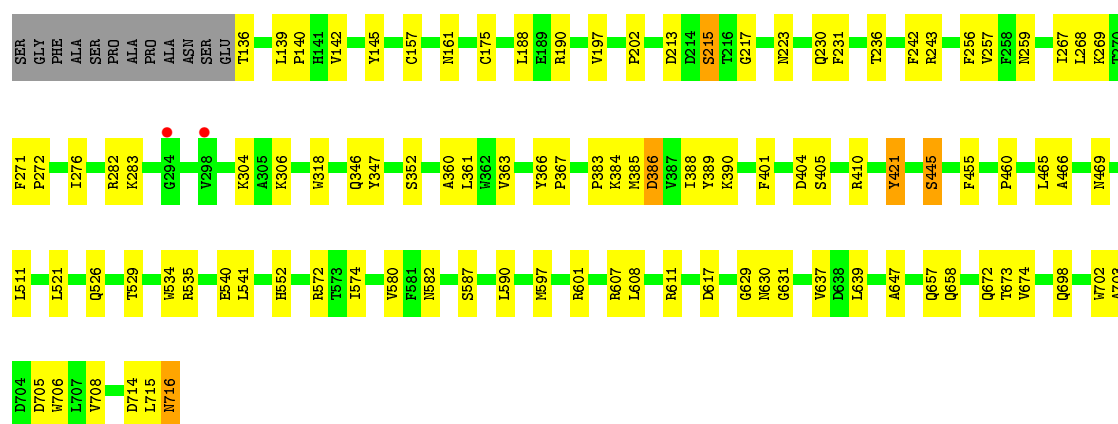
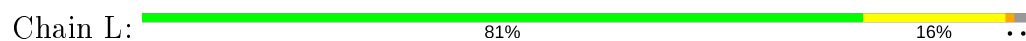




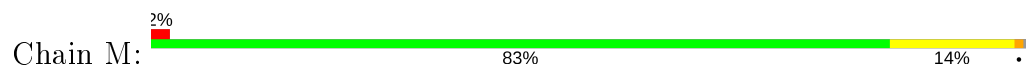
- Molecule 1: Hemolysin

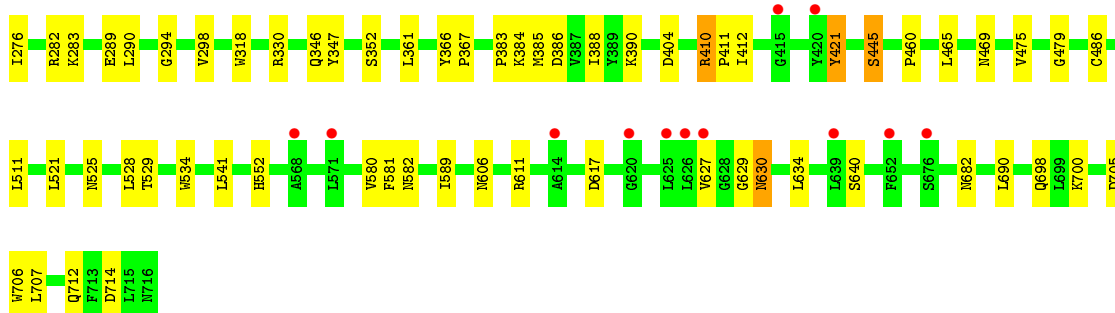


- Molecule 1: Hemolysin

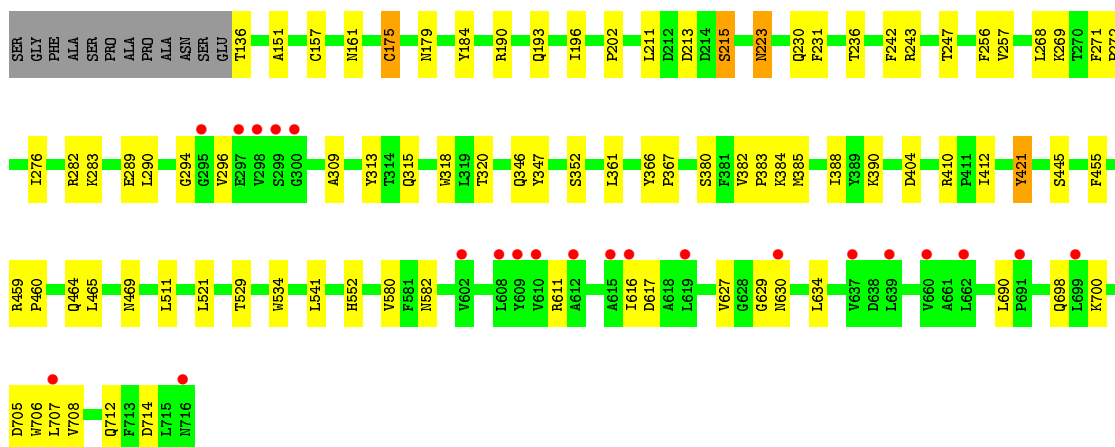
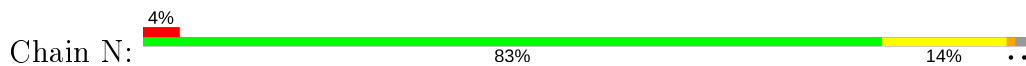


- Molecule 1: Hemolysin





● Molecule 1: Hemolysin



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	172.35Å 182.86Å 430.39Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	43.96 – 2.88 86.18 – 2.89	Depositor EDS
% Data completeness (in resolution range)	90.8 (43.96-2.88) 98.4 (86.18-2.89)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.86 (at 2.91Å)	Xtrriage
Refinement program	PHENIX 1.5_2	Depositor
R, R_{free}	0.218 , 0.249 0.215 , 0.244	Depositor DCC
R_{free} test set	15098 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å ²)	52.7	Xtrriage
Anisotropy	0.687	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 45.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	62948	wwPDB-VP
Average B, all atoms (Å ²)	55.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.27% 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 [i](#)

5.1 Standard geometry [i](#)

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	0/4557	0.57	0/6204
1	B	0.43	0/4551	0.56	0/6198
1	C	0.44	0/4523	0.55	0/6164
1	D	0.45	0/4567	0.56	0/6221
1	E	0.43	0/4555	0.55	0/6205
1	F	0.42	0/4503	0.56	0/6139
1	G	0.45	0/4528	0.56	0/6168
1	H	0.40	0/4564	0.55	0/6219
1	I	0.44	0/4575	0.56	0/6233
1	J	0.42	0/4583	0.55	0/6240
1	K	0.43	0/4509	0.56	0/6146
1	L	0.43	0/4541	0.58	0/6183
1	M	0.43	0/4539	0.57	0/6183
1	N	0.42	0/4513	0.56	0/6153
All	All	0.43	0/63608	0.56	0/86656

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 [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	4464	0	4169	57	0
1	B	4457	0	4163	51	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	4431	0	4100	50	0
1	D	4471	0	4172	57	1
1	E	4460	0	4146	46	0
1	F	4412	0	4065	47	0
1	G	4436	0	4120	55	1
1	H	4468	0	4156	51	0
1	I	4479	0	4178	55	0
1	J	4487	0	4205	50	0
1	K	4417	0	4093	58	0
1	L	4449	0	4145	67	0
1	M	4447	0	4147	61	0
1	N	4420	0	4094	61	0
2	A	58	0	0	1	0
2	B	62	0	0	1	0
2	C	57	0	0	1	0
2	D	65	0	0	2	0
2	E	44	0	0	0	0
2	F	26	0	0	2	0
2	G	50	0	0	0	0
2	H	48	0	0	0	0
2	I	61	0	0	1	0
2	J	32	0	0	0	0
2	K	25	0	0	0	0
2	L	43	0	0	1	0
2	M	43	0	0	2	0
2	N	36	0	0	1	0
All	All	62948	0	57953	639	1

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 5.

The worst 5 of 639 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:C:283:LYS:HB2	1:C:318:TRP:HB2	1.36	1.04
1:K:480:ARG:HH11	1:K:480:ARG:HB3	1.22	1.02
1:N:283:LYS:HB2	1:N:318:TRP:HB2	1.41	1.02
1:F:283:LYS:HB2	1:F:318:TRP:HB2	1.43	1.01
1:K:283:LYS:HB2	1:K:318:TRP:HB2	1.42	1.01

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the sym-

metry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:236:THR:OG1	1:G:235:TYR:CE2[4_456]	1.90	0.30

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	579/593 (98%)	557 (96%)	22 (4%)	0	100	100
1	B	579/593 (98%)	556 (96%)	23 (4%)	0	100	100
1	C	579/593 (98%)	554 (96%)	25 (4%)	0	100	100
1	D	579/593 (98%)	556 (96%)	23 (4%)	0	100	100
1	E	579/593 (98%)	553 (96%)	26 (4%)	0	100	100
1	F	579/593 (98%)	555 (96%)	24 (4%)	0	100	100
1	G	579/593 (98%)	556 (96%)	23 (4%)	0	100	100
1	H	579/593 (98%)	555 (96%)	24 (4%)	0	100	100
1	I	579/593 (98%)	552 (95%)	27 (5%)	0	100	100
1	J	579/593 (98%)	554 (96%)	25 (4%)	0	100	100
1	K	579/593 (98%)	553 (96%)	26 (4%)	0	100	100
1	L	579/593 (98%)	555 (96%)	24 (4%)	0	100	100
1	M	579/593 (98%)	554 (96%)	25 (4%)	0	100	100
1	N	579/593 (98%)	559 (96%)	20 (4%)	0	100	100
All	All	8106/8302 (98%)	7769 (96%)	337 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	464/495 (94%)	446 (96%)	18 (4%)	32	64
1	B	462/495 (93%)	446 (96%)	16 (4%)	36	68
1	C	456/495 (92%)	440 (96%)	16 (4%)	36	68
1	D	464/495 (94%)	448 (97%)	16 (3%)	37	69
1	E	461/495 (93%)	445 (96%)	16 (4%)	36	68
1	F	453/495 (92%)	436 (96%)	17 (4%)	33	65
1	G	457/495 (92%)	440 (96%)	17 (4%)	34	66
1	H	464/495 (94%)	450 (97%)	14 (3%)	41	73
1	I	466/495 (94%)	451 (97%)	15 (3%)	39	71
1	J	468/495 (94%)	451 (96%)	17 (4%)	35	67
1	K	453/495 (92%)	437 (96%)	16 (4%)	36	68
1	L	460/495 (93%)	442 (96%)	18 (4%)	32	64
1	M	461/495 (93%)	444 (96%)	17 (4%)	34	66
1	N	454/495 (92%)	440 (97%)	14 (3%)	40	72
All	All	6443/6930 (93%)	6216 (96%)	227 (4%)	36	68

5 of 227 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	215	SER
1	H	445	SER
1	M	445	SER
1	G	276	ILE
1	G	630	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 63 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	179	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	H	346	GLN
1	M	427	GLN
1	G	315	GLN
1	G	426	HIS

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 carbohydrates 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, 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	A	581/593 (97%)	0.01	2 (0%) 94 94	29, 48, 92, 128	0
1	B	581/593 (97%)	0.09	8 (1%) 75 75	29, 50, 93, 132	0
1	C	581/593 (97%)	0.15	12 (2%) 63 62	28, 51, 94, 129	0
1	D	581/593 (97%)	0.21	7 (1%) 79 78	29, 49, 93, 128	0
1	E	581/593 (97%)	0.10	10 (1%) 70 70	31, 50, 94, 131	0
1	F	581/593 (97%)	0.17	20 (3%) 45 40	31, 53, 96, 130	0
1	G	581/593 (97%)	0.34	31 (5%) 26 22	31, 53, 96, 127	0
1	H	581/593 (97%)	0.16	13 (2%) 62 60	29, 51, 94, 132	0
1	I	581/593 (97%)	0.18	7 (1%) 79 78	29, 49, 89, 129	0
1	J	581/593 (97%)	0.07	9 (1%) 73 73	31, 51, 94, 128	0
1	K	581/593 (97%)	0.26	32 (5%) 25 21	33, 53, 96, 130	0
1	L	581/593 (97%)	-0.08	2 (0%) 94 94	31, 51, 93, 128	0
1	M	581/593 (97%)	0.06	12 (2%) 63 62	31, 52, 92, 127	0
1	N	581/593 (97%)	0.20	22 (3%) 40 36	31, 52, 96, 132	0
All	All	8134/8302 (97%)	0.14	187 (2%) 60 59	28, 51, 94, 132	0

The worst 5 of 187 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	N	300	GLY	5.0
1	G	716	ASN	4.8
1	G	679	TYR	4.7
1	F	416	ALA	4.6
1	I	298	VAL	4.5

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.