



wwPDB X-ray Structure Validation Summary Report

Feb 26, 2014 – 04:36 PM GMT

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 validation summary report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

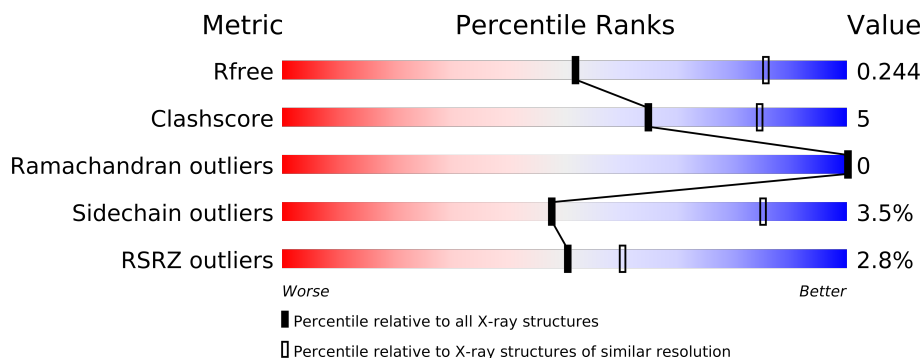
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable22639
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable22683

1 Overall quality at a glance

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}	66092	1360 (2.90-2.86)
Clashscore	79885	1696 (2.90-2.86)
Ramachandran outliers	78287	1647 (2.90-2.86)
Sidechain outliers	78261	1650 (2.90-2.86)
RSRZ outliers	66119	1362 (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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	593	
1	B	593	
1	C	593	
1	D	593	
1	E	593	
1	F	593	
1	G	593	
1	H	593	
1	I	593	
1	J	593	
1	K	593	
1	L	593	
1	M	593	
1	N	593	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 62948 atoms, of which 0 are hydrogen and 0 are deuterium.

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

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...

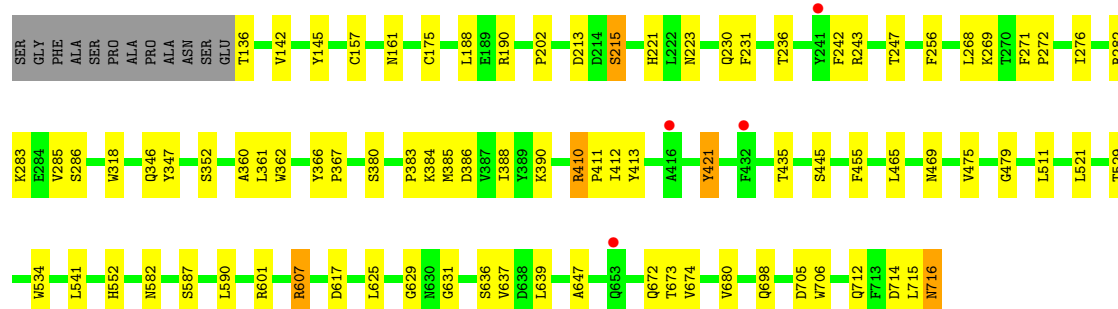
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	N	36	Total	O	0	0
			36	36		

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 errors 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.

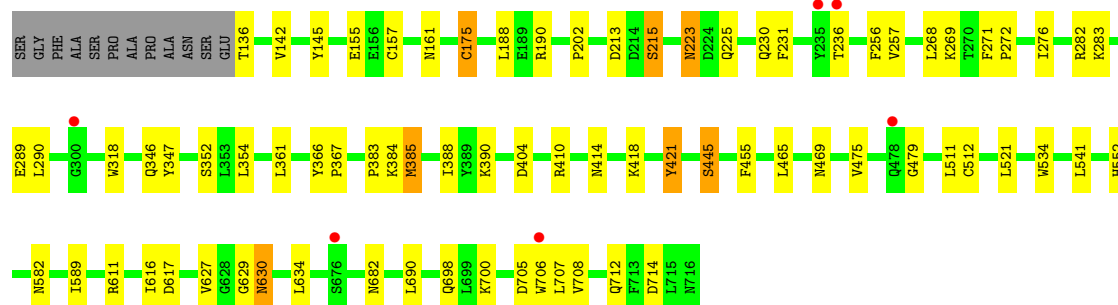
• Molecule 1: Hemolysin

Chain A: 



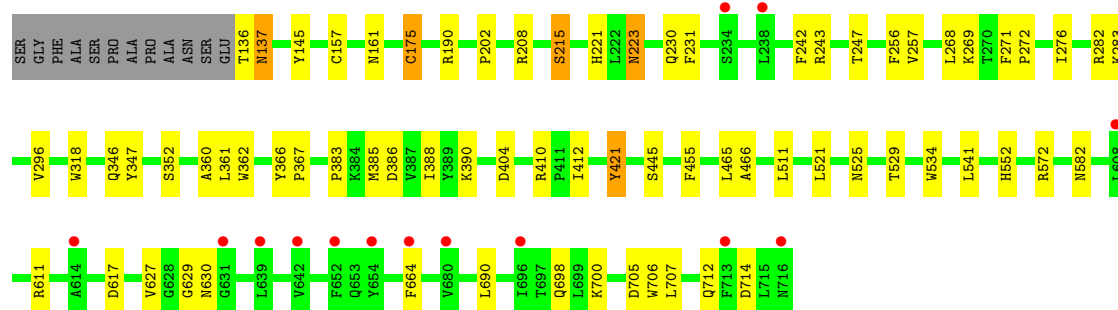
• Molecule 1: Hemolysin

Chain B: 



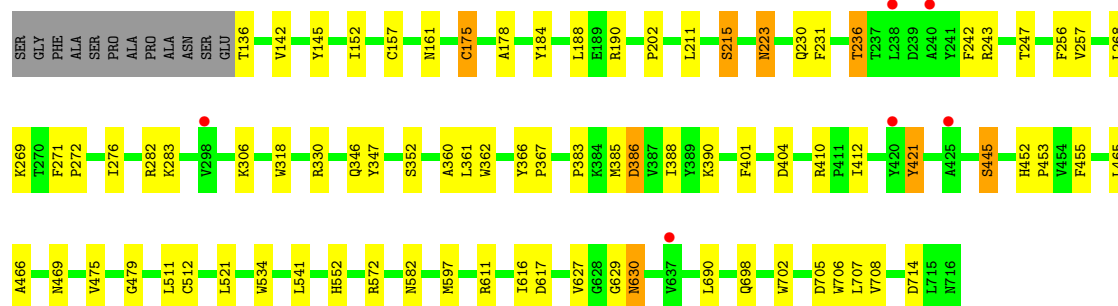
• Molecule 1: Hemolysin

Chain C: 



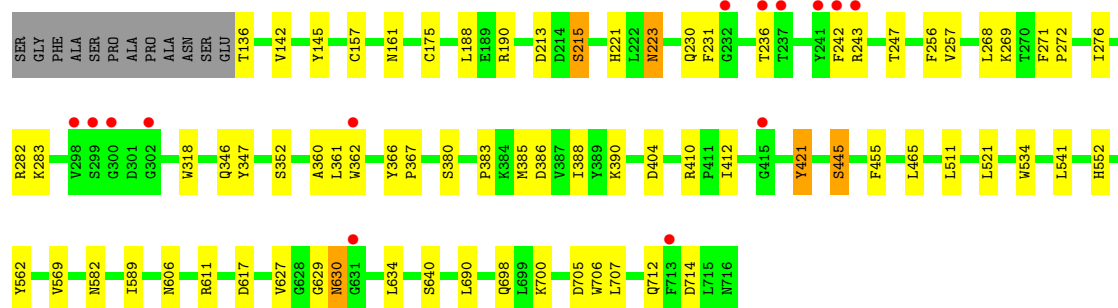
- Molecule 1: Hemolysin

Chain D:



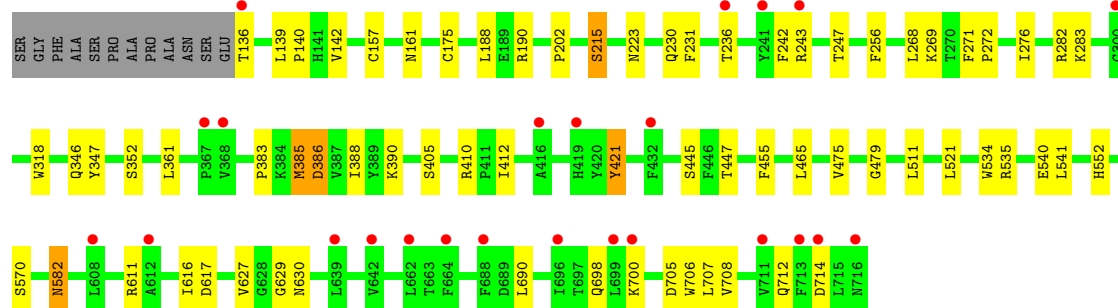
- Molecule 1: Hemolysin

Chain E:



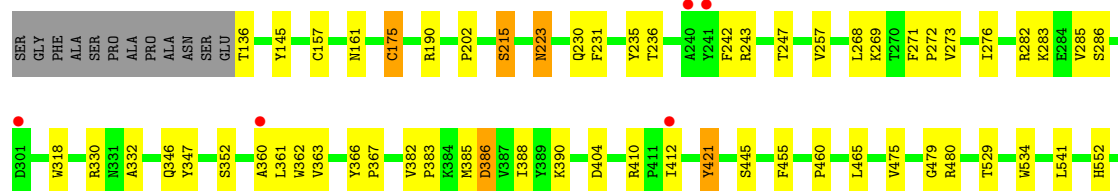
- Molecule 1: Hemolysin

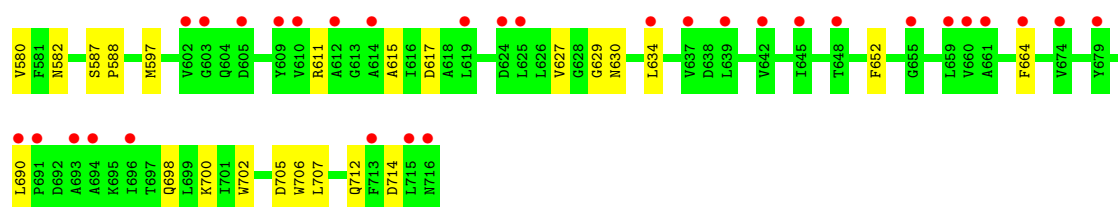
Chain F:



- Molecule 1: Hemolysin

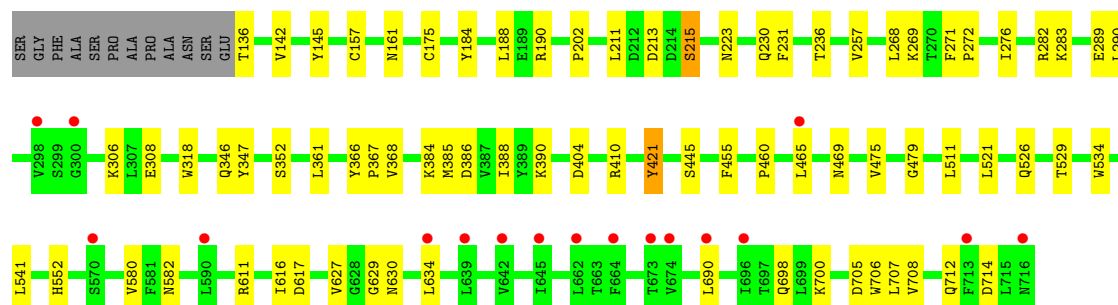
Chain G:





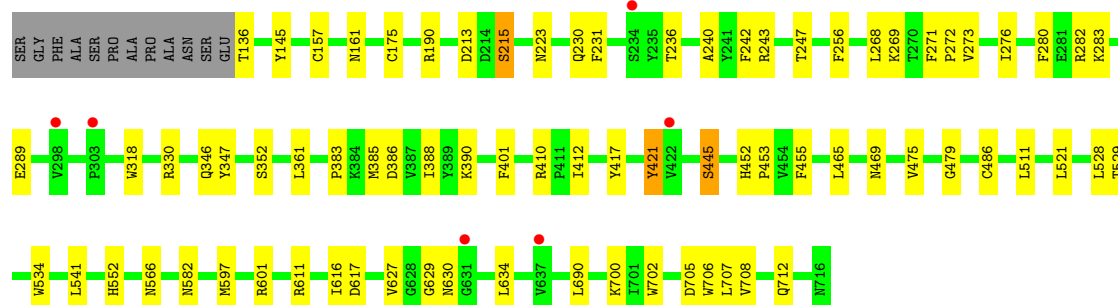
• Molecule 1: Hemolysin

Chain H:



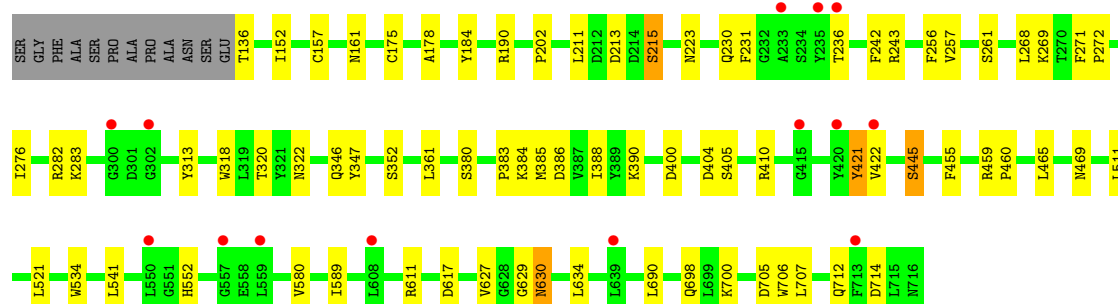
• Molecule 1: Hemolysin

Chain I:



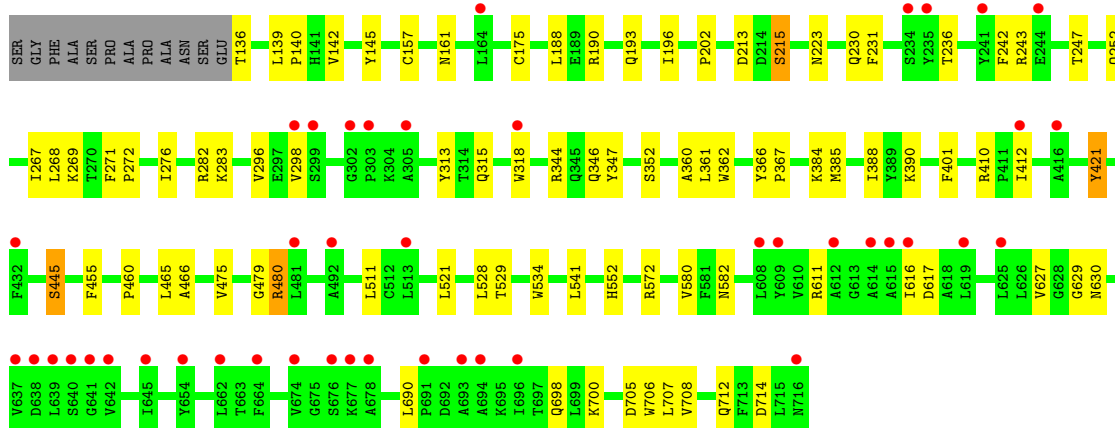
• Molecule 1: Hemolysin

Chain J:



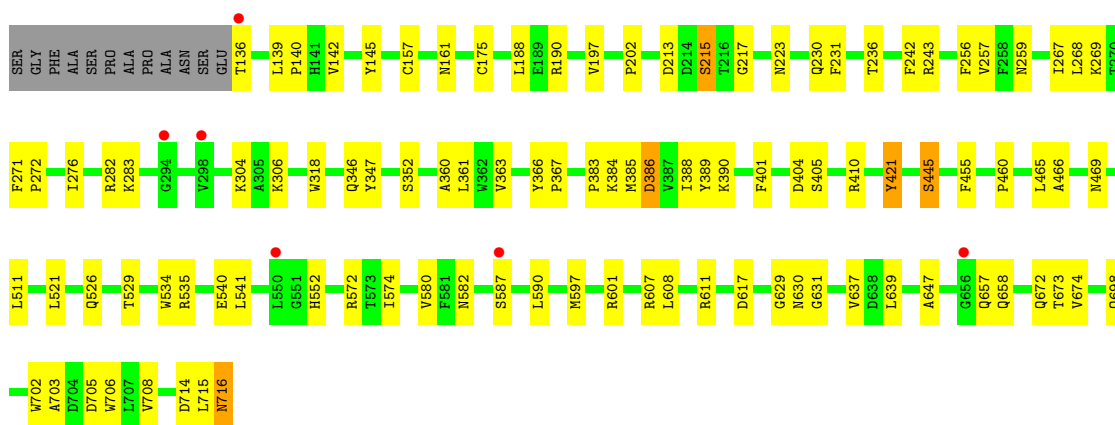
• Molecule 1: Hemolysin

Chain K:



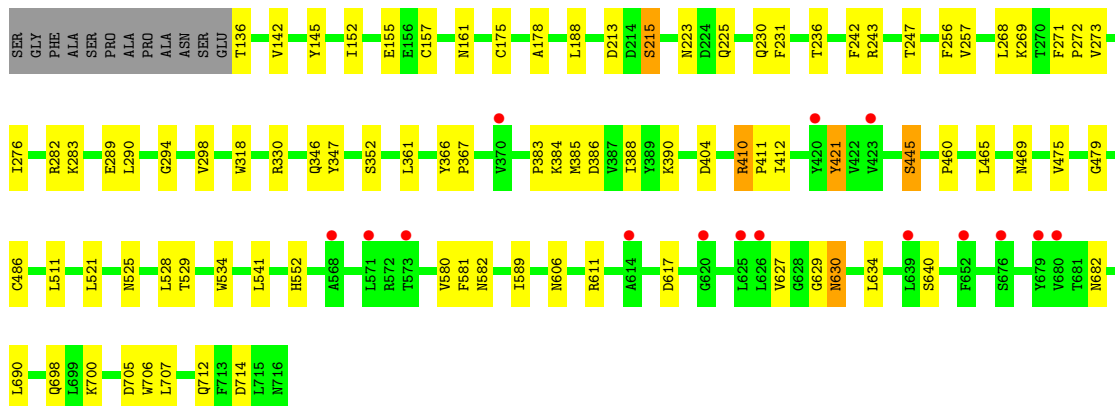
• Molecule 1: Hemolysin

Chain L:



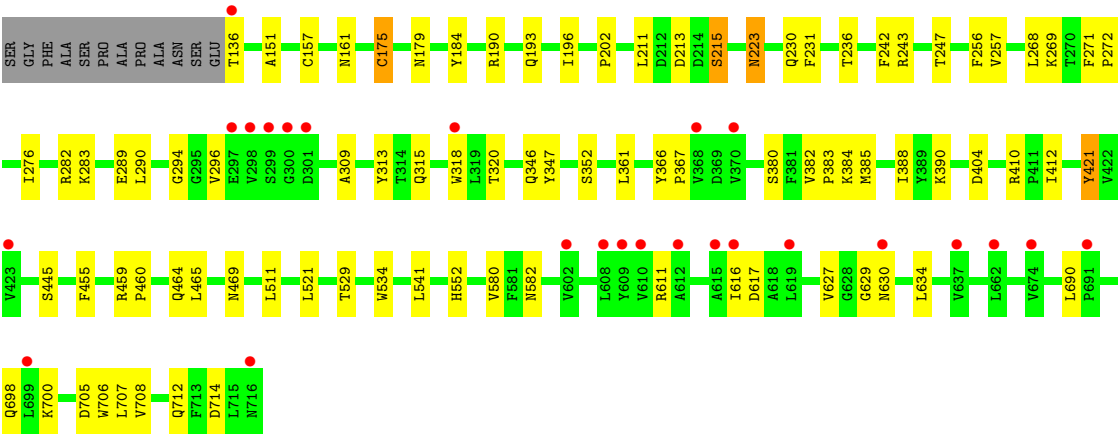
• Molecule 1: Hemolysin

Chain M:



• Molecule 1: Hemolysin

Chain N:



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)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.5_2)	Depositor
R, R_{free}	0.218 , 0.249 0.215 , 0.244	Depositor DCC
R_{free} test set	15094 reflections (5.05%)	DCC
Wilson B-factor (Å ²)	52.7	Xtriage
Anisotropy	0.687	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 20.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	2 of 299067 reflections (0.001%)	Xtriage
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

Xtriage'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.

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	0.44	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 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

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
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

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 5.

The worst 5 of 639 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(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 symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:D:236:THR:OG1	1:G:235:TYR:CE2[4.456]	1.90	0.30

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

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 ⓘ

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%)	43	82
1	B	462/495 (93%)	446 (96%)	16 (4%)	48	85
1	C	456/495 (92%)	440 (96%)	16 (4%)	48	85
1	D	464/495 (94%)	448 (97%)	16 (3%)	49	86
1	E	461/495 (93%)	445 (96%)	16 (4%)	48	85
1	F	453/495 (92%)	436 (96%)	17 (4%)	44	83
1	G	457/495 (92%)	440 (96%)	17 (4%)	45	83
1	H	464/495 (94%)	450 (97%)	14 (3%)	53	88
1	I	466/495 (94%)	451 (97%)	15 (3%)	51	87
1	J	468/495 (94%)	451 (96%)	17 (4%)	47	84
1	K	453/495 (92%)	437 (96%)	16 (4%)	48	85
1	L	460/495 (93%)	442 (96%)	18 (4%)	43	82
1	M	461/495 (93%)	444 (96%)	17 (4%)	45	83
1	N	454/495 (92%)	440 (97%)	14 (3%)	52	88
All	All	6443/6930 (93%)	6216 (96%)	227 (4%)	48	85

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
1	H	346	GLN
1	M	427	GLN
1	G	315	GLN
1	G	426	HIS

5.3.3 RNA ⓘ

There are no RNA chains in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

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, 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%)	-0.05	4 (0%) 84 90	29, 48, 92, 128	0
1	B	581/593 (97%)	-0.00	6 (1%) 79 86	29, 50, 93, 132	0
1	C	581/593 (97%)	0.12	14 (2%) 56 66	28, 51, 94, 129	0
1	D	581/593 (97%)	0.17	6 (1%) 79 86	29, 49, 93, 128	0
1	E	581/593 (97%)	0.07	14 (2%) 56 66	31, 50, 94, 131	0
1	F	581/593 (97%)	0.14	24 (4%) 35 43	31, 53, 96, 130	0
1	G	581/593 (97%)	0.32	36 (6%) 20 24	31, 53, 96, 127	0
1	H	581/593 (97%)	0.13	17 (2%) 49 59	29, 51, 94, 132	0
1	I	581/593 (97%)	0.13	6 (1%) 79 86	29, 49, 89, 129	0
1	J	581/593 (97%)	0.06	14 (2%) 56 66	31, 51, 94, 128	0
1	K	581/593 (97%)	0.22	44 (7%) 14 17	33, 53, 96, 130	0
1	L	581/593 (97%)	-0.15	6 (1%) 79 86	31, 51, 93, 128	0
1	M	581/593 (97%)	0.05	15 (2%) 53 63	31, 52, 92, 127	0
1	N	581/593 (97%)	0.13	25 (4%) 34 41	31, 52, 96, 132	0
All	All	8134/8302 (97%)	0.10	231 (2%) 50 60	28, 51, 94, 132	0

The worst 5 of 231 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	716	ASN	7.2
1	I	298	VAL	5.2
1	H	716	ASN	5.2
1	G	664	PHE	5.2
1	N	300	GLY	5.2

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 ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.