wwPDB X-ray Structure Validation Summary Report (i)

Oct 3, 2021 – 04:43 PM EDT

PDB ID	:	3I41
Title	:	Crystal structure of beta toxin from Staphylococcus aureus F277A, P278A mutant
Authors	:	Huseby, M.; Shi, K.; Kruse, A.C.; Ohlendorf, D.H.
Deposited on	:	2009-07-01
Resolution	:	1.75 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:

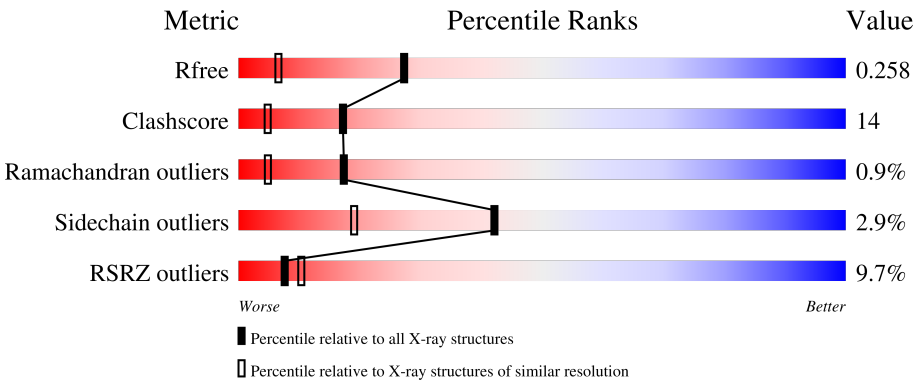
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 1.75 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	2340 (1.76-1.76)
Clashscore	141614	2466 (1.76-1.76)
Ramachandran outliers	138981	2437 (1.76-1.76)
Sidechain outliers	138945	2437 (1.76-1.76)
RSRZ outliers	127900	2298 (1.76-1.76)

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 of 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	317	9% 70% 19% • 9%
1	B	317	9% 76% 14% • 9%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 5192 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 Beta-hemolysin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	290	2329	1477	393	454	5	0	0	0
1	B	290	2327	1475	393	454	5	0	0	0

There are 46 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	expression tag	UNP A7LAI8
A	-18	ARG	-	expression tag	UNP A7LAI8
A	-17	SER	-	expression tag	UNP A7LAI8
A	-16	SER	-	expression tag	UNP A7LAI8
A	-15	HIS	-	expression tag	UNP A7LAI8
A	-14	HIS	-	expression tag	UNP A7LAI8
A	-13	HIS	-	expression tag	UNP A7LAI8
A	-12	HIS	-	expression tag	UNP A7LAI8
A	-11	HIS	-	expression tag	UNP A7LAI8
A	-10	HIS	-	expression tag	UNP A7LAI8
A	-9	SER	-	expression tag	UNP A7LAI8
A	-8	SER	-	expression tag	UNP A7LAI8
A	-7	GLY	-	expression tag	UNP A7LAI8
A	-6	LEU	-	expression tag	UNP A7LAI8
A	-5	VAL	-	expression tag	UNP A7LAI8
A	-4	PRO	-	expression tag	UNP A7LAI8
A	-3	ARG	-	expression tag	UNP A7LAI8
A	-2	GLY	-	expression tag	UNP A7LAI8
A	-1	SER	-	expression tag	UNP A7LAI8
A	0	HIS	-	expression tag	UNP A7LAI8
A	1	MET	-	expression tag	UNP A7LAI8
A	277	ALA	PHE	engineered mutation	UNP A7LAI8
A	278	ALA	PRO	engineered mutation	UNP A7LAI8
B	-19	MET	-	expression tag	UNP A7LAI8
B	-18	ARG	-	expression tag	UNP A7LAI8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-17	SER	-	expression tag	UNP A7LAI8
B	-16	SER	-	expression tag	UNP A7LAI8
B	-15	HIS	-	expression tag	UNP A7LAI8
B	-14	HIS	-	expression tag	UNP A7LAI8
B	-13	HIS	-	expression tag	UNP A7LAI8
B	-12	HIS	-	expression tag	UNP A7LAI8
B	-11	HIS	-	expression tag	UNP A7LAI8
B	-10	HIS	-	expression tag	UNP A7LAI8
B	-9	SER	-	expression tag	UNP A7LAI8
B	-8	SER	-	expression tag	UNP A7LAI8
B	-7	GLY	-	expression tag	UNP A7LAI8
B	-6	LEU	-	expression tag	UNP A7LAI8
B	-5	VAL	-	expression tag	UNP A7LAI8
B	-4	PRO	-	expression tag	UNP A7LAI8
B	-3	ARG	-	expression tag	UNP A7LAI8
B	-2	GLY	-	expression tag	UNP A7LAI8
B	-1	SER	-	expression tag	UNP A7LAI8
B	0	HIS	-	expression tag	UNP A7LAI8
B	1	MET	-	expression tag	UNP A7LAI8
B	277	ALA	PHE	engineered mutation	UNP A7LAI8
B	278	ALA	PRO	engineered mutation	UNP A7LAI8

- Molecule 2 is water.

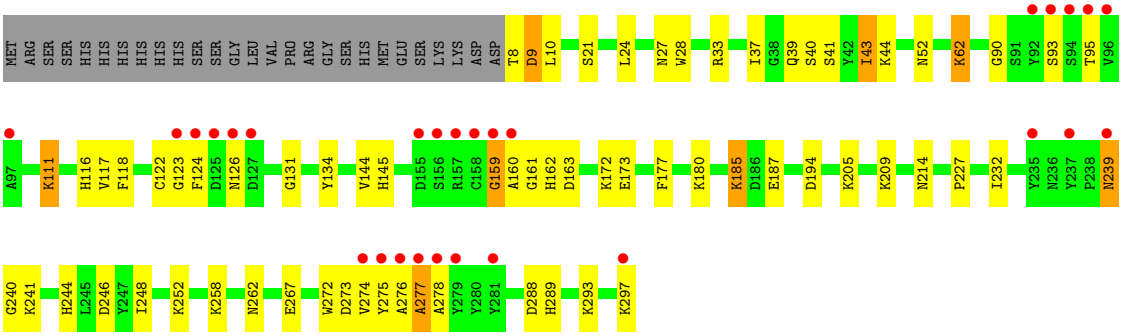
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	289	289	289	0	0
			Total	O		
2	B	247	247	247	0	0

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.

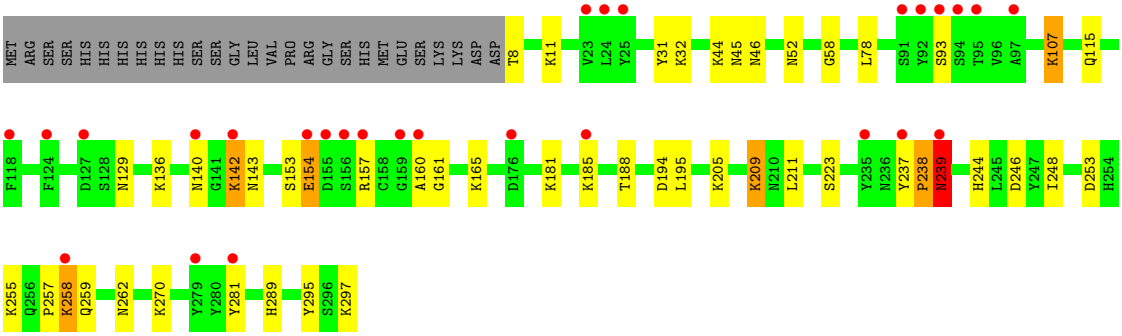
• Molecule 1: Beta-hemolysin

Chain A:



• Molecule 1: Beta-hemolysin

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	63.18Å 68.67Å 127.84Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	26.18 – 1.75 25.38 – 1.75	Depositor EDS
% Data completeness (in resolution range)	95.3 (26.18-1.75) 95.3 (25.38-1.75)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.36 (at 1.75Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.201 , 0.260 0.200 , 0.258	Depositor DCC
R_{free} test set	2741 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	21.6	Xtriage
Anisotropy	0.112	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 61.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	5192	wwPDB-VP
Average B, all atoms (Å^2)	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.52% 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 (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	0.79	0/2388	0.81	2/3231 (0.1%)
1	B	0.67	0/2385	0.71	1/3226 (0.0%)
All	All	0.73	0/4773	0.76	3/6457 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	43	ILE	CG1-CB-CG2	5.82	124.19	111.40
1	B	238	PRO	N-CA-CB	5.43	109.81	103.30
1	A	9	ASP	N-CA-CB	5.43	120.37	110.60

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	2329	0	2247	87	2
1	B	2327	0	2241	42	0
2	A	289	0	0	34	0
2	B	247	0	0	14	0
All	All	5192	0	4488	127	2

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 14.

The worst 5 of 127 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:A:185:LYS:HB3	2:A:574:HOH:O	1.24	1.37
1:A:62:LYS:HD2	2:A:568:HOH:O	1.27	1.28
1:A:62:LYS:HB3	2:A:449:HOH:O	1.38	1.19
1:A:62:LYS:CD	2:A:568:HOH:O	1.78	1.18
1:A:185:LYS:CD	2:A:574:HOH:O	1.83	1.17

All (2) 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	Interatomic distance (Å)	Clash overlap (Å)
1:A:111:LYS:NZ	1:A:267:GLU:OE1[4_445]	1.78	0.42
1:A:41:SER:CB	1:A:111:LYS:CE[4_545]	2.04	0.16

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	288/317 (91%)	273 (95%)	13 (4%)	2 (1%)	22 8
1	B	288/317 (91%)	277 (96%)	8 (3%)	3 (1%)	15 4
All	All	576/634 (91%)	550 (96%)	21 (4%)	5 (1%)	17 5

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	160	ALA
1	B	239	ASN
1	A	159	GLY
1	B	258	LYS
1	A	277	ALA

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	257/282 (91%)	249 (97%)	8 (3%)	40	17
1	B	256/282 (91%)	249 (97%)	7 (3%)	44	22
All	All	513/564 (91%)	498 (97%)	15 (3%)	42	19

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	241	LYS
1	B	239	ASN
1	B	107	LYS
1	B	281	TYR
1	B	154	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	239	ASN
1	B	262	ASN
1	B	289	HIS
1	B	259	GLN
1	A	289	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 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)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

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	290/317 (91%)	0.64	28 (9%)	7	10	10, 20, 46, 57	0
1	B	290/317 (91%)	0.67	28 (9%)	7	10	15, 26, 40, 49	0
All	All	580/634 (91%)	0.66	56 (9%)	7	10	10, 23, 43, 57	0

The worst 5 of 56 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	124	PHE	10.4
1	A	279	TYR	8.7
1	A	156	SER	8.1
1	B	124	PHE	8.1
1	A	160	ALA	7.9

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 monosaccharides 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.