wwPDB X-ray Structure Validation Summary Report (i)

Nov 1, 2021 – 01:31 PM EDT

PDB ID	:	1Z52
Title	:	Proaerolysin Mutant W373L
Authors	:	Parker, M.W.; Feil, S.C.; Tang, J.W.
Deposited on	:	2005-03-16
Resolution	:	2.38 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:

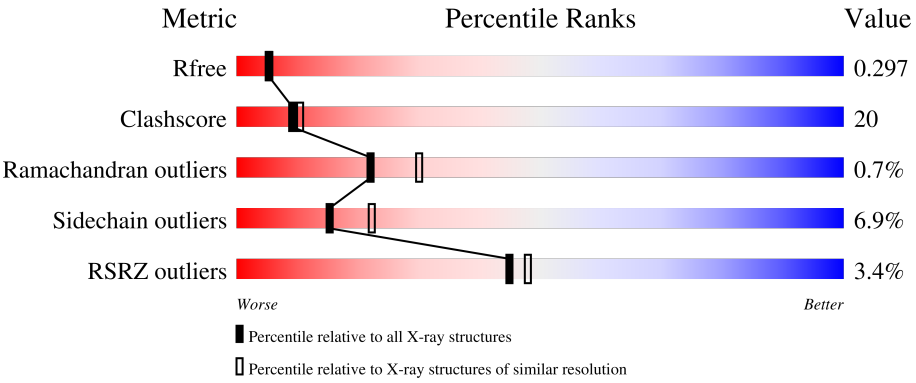
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 2.38 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	5509 (2.40-2.36)
Clashscore	141614	6082 (2.40-2.36)
Ramachandran outliers	138981	5973 (2.40-2.36)
Sidechain outliers	138945	5975 (2.40-2.36)
RSRZ outliers	127900	5397 (2.40-2.36)

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	470	4% 65% 27% . .
1	B	470	3% 61% 29% 6% .

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 7179 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 Aerolysin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	451	Total	C	N	O	S	0	0	0
			3525	2227	607	682	9			
1	B	450	Total	C	N	O	S	0	0	0
			3522	2226	605	682	9			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	373	LEU	TRP	engineered mutation	UNP P09167
B	373	LEU	TRP	engineered mutation	UNP P09167

- Molecule 2 is water.

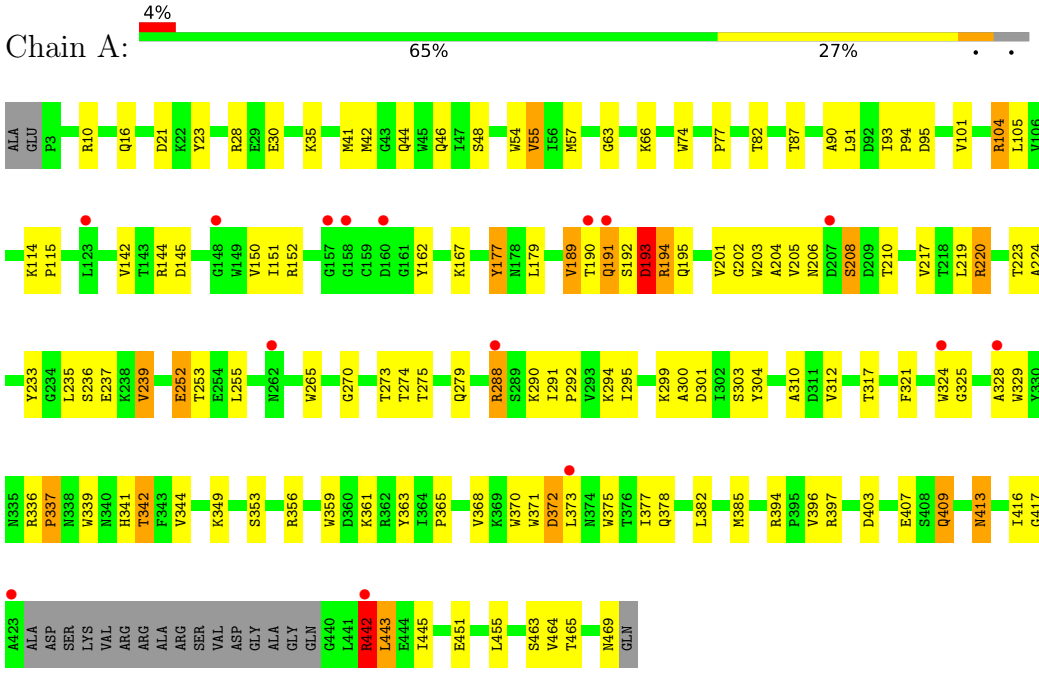
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	50	Total	O	0	0
			50	50		
2	B	82	Total	O	0	0
			82	82		

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.

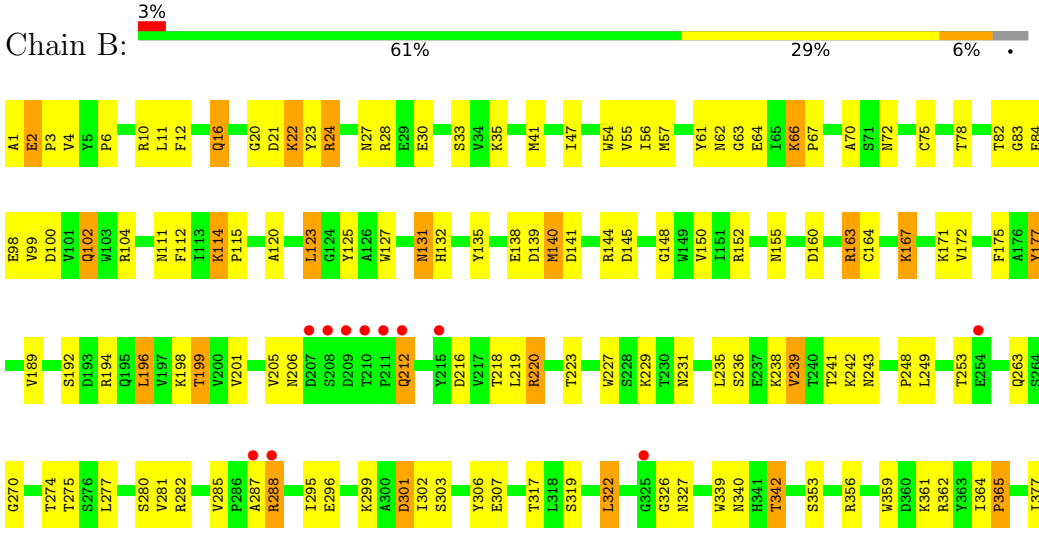
• Molecule 1: Aerolysin

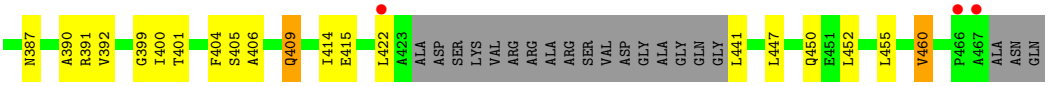
Chain A:



• Molecule 1: Aerolysin

Chain B:




N387
A390
R391
V392
G399
I400
T401
F404
S405
A406
Q409
I414
E415
L422
A423
ALA
ASP
SER
LYS
VAL
ARG
ARG
ALA
ARG
SER
VAL
ASP
GLY
ALA
GLY
GLN
GLY
L441
L447
Q450
E451
L452
L455
V460
P466
A467
ALA
ASN
GLN

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	70.71Å 91.14Å 168.22Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	25.00 – 2.38 25.00 – 2.39	Depositor EDS
% Data completeness (in resolution range)	(Not available) (25.00-2.38) 86.7 (25.00-2.39)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.79 (at 2.39Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.250 , 0.301 0.246 , 0.297	Depositor DCC
R_{free} test set	1908 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	33.2	Xtriage
Anisotropy	0.127	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 32.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	7179	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 4.29% 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

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	0.69	5/3621 (0.1%)	0.87	7/4941 (0.1%)
1	B	0.58	1/3618 (0.0%)	0.79	6/4938 (0.1%)
All	All	0.64	6/7239 (0.1%)	0.83	13/9879 (0.1%)

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	1
1	B	0	2
All	All	0	3

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	193	ASP	C-N	22.70	1.86	1.34
1	A	370	TRP	C-N	-9.24	1.12	1.34
1	A	189	VAL	C-N	9.00	1.54	1.34
1	B	102	GLN	C-N	8.59	1.53	1.34
1	A	203	TRP	C-N	-6.58	1.19	1.34

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	193	ASP	CA-C-N	-22.34	68.05	117.20
1	A	193	ASP	O-C-N	16.85	149.66	122.70
1	A	104	ARG	NE-CZ-NH2	7.74	124.17	120.30
1	A	442	ARG	NE-CZ-NH2	7.55	124.07	120.30
1	A	220	ARG	NE-CZ-NH2	7.51	124.05	120.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	443	LEU	Mainchain
1	B	172	VAL	Mainchain
1	B	196	LEU	Mainchain

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	3525	0	3362	134	1
1	B	3522	0	3365	145	1
2	A	50	0	0	15	2
2	B	82	0	0	27	2
All	All	7179	0	6727	275	3

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 20.

The worst 5 of 275 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:193:ASP:CA	1:A:194:ARG:N	1.92	1.33
1:A:193:ASP:C	1:A:194:ARG:N	1.86	1.29
1:A:193:ASP:HA	1:A:194:ARG:N	1.50	1.26
1:A:442:ARG:HH11	1:A:442:ARG:CG	1.53	1.21
1:B:353:SER:HB3	2:B:535:HOH:O	1.41	1.18

All (3) 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)
2:A:517:HOH:O	2:B:534:HOH:O[4_456]	1.94	0.26
1:A:210:THR:OG1	2:B:548:HOH:O[4_556]	2.14	0.06
1:B:1:ALA:CB	2:A:518:HOH:O[4_556]	2.16	0.04

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	447/470 (95%)	414 (93%)	30 (7%)	3 (1%)	22 30
1	B	446/470 (95%)	414 (93%)	29 (6%)	3 (1%)	22 30
All	All	893/940 (95%)	828 (93%)	59 (7%)	6 (1%)	22 30

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	191	GLN
1	A	145	ASP
1	A	193	ASP
1	B	365	PRO
1	B	2	GLU

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	379/392 (97%)	358 (94%)	21 (6%)	21 32
1	B	379/392 (97%)	348 (92%)	31 (8%)	11 15
All	All	758/784 (97%)	706 (93%)	52 (7%)	15 22

5 of 52 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	87	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	167	LYS
1	B	422	LEU
1	B	114	LYS
1	B	140	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 27 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	131	ASN
1	B	231	ASN
1	B	413	ASN
1	B	212	GLN
1	B	263	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	A	3

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	193:ASP	C	194:ARG	N	1.86
1	A	203:TRP	C	204:ALA	N	1.18
1	A	370:TRP	C	371:TRP	N	1.12

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	451/470 (95%)	0.19	17 (3%)	40	43	10, 26, 47, 66	0
1	B	450/470 (95%)	0.14	14 (3%)	49	51	9, 23, 49, 72	0
All	All	901/940 (95%)	0.17	31 (3%)	45	48	9, 25, 49, 72	0

The worst 5 of 31 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	190	THR	6.3
1	A	423	ALA	6.2
1	A	324	TRP	4.4
1	B	210	THR	4.2
1	B	211	PRO	4.1

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 monosaccharides 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.