



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 09:35 PM GMT

PDB ID : 1PRE
Title : PROAEROLYSIN
Authors : Parker, M.W.; Buckley, J.T.; Postma, J.P.M.; Tucker, A.D.; Tsernoglou, D.
Deposited on : 1995-09-15
Resolution : 2.80 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
<http://wwpdb.org/validation/2016/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
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

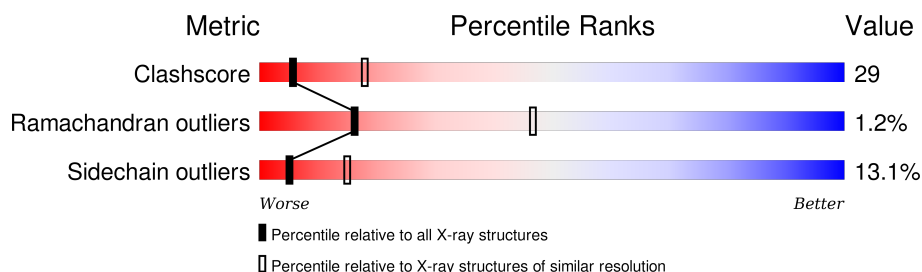
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 Å.

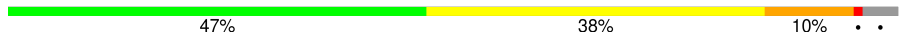
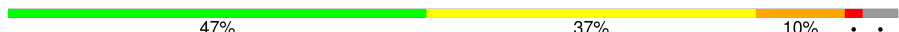
Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)

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. 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	470	 47% 38% 10% . .
1	B	470	 47% 37% 10% . .

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 7089 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 PROAEROLYSIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	449	Total	C	N	O	S	0	0	0
			3523	2228	607	679	9			
1	B	451	Total	C	N	O	S	0	0	0
			3533	2234	607	683	9			

- Molecule 2 is water.

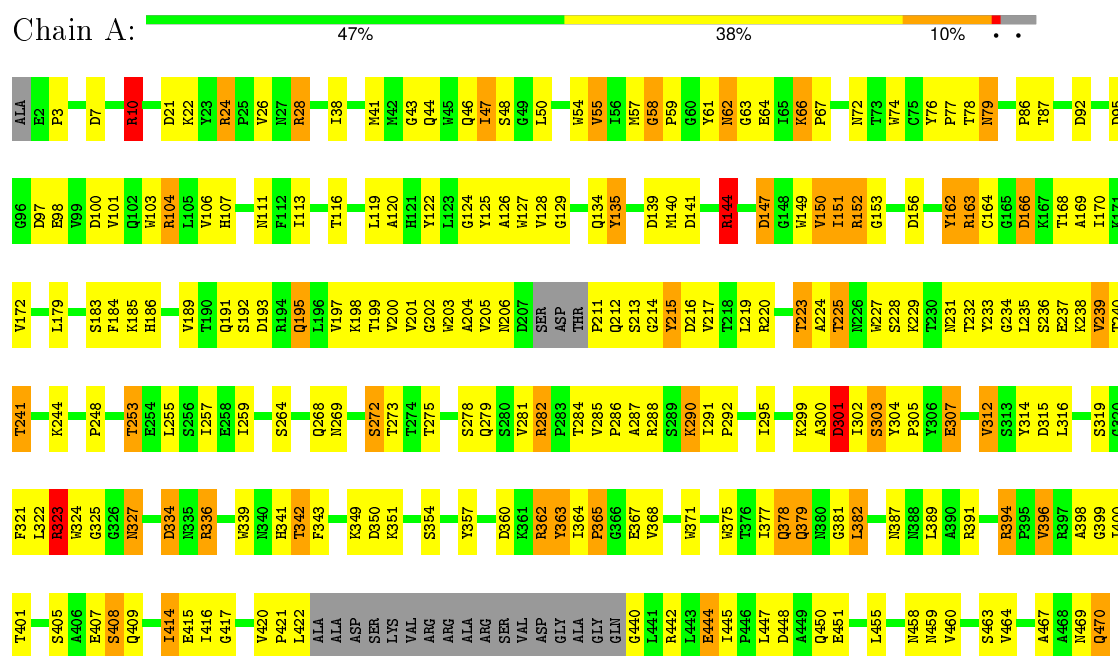
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	18	Total	O	0	0
			18	18		
2	B	15	Total	O	0	0
			15	15		

3 Residue-property plots

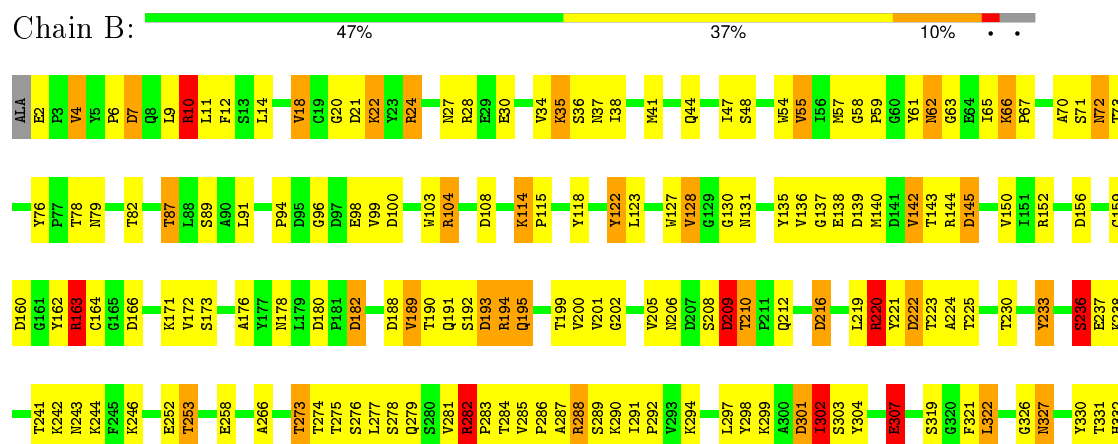
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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.

Note EDS was not executed.

• Molecule 1: PROAEROLYSIN



• Molecule 1: PROAEROLYSIN



Q409	R336	P337	I338	I339	I340	H341	T342	G346	P347	Y348	I349	D350	R351	A352	S353	S354	T355	R356	Y357	Q358	K361	R362	T363	I364	P365	I366	E367	Y368	T369	W370	W371	D372	W373	I374	H375	T376	I377	Q378	Q379	S383	T384	I388	I389	A390	R391	Y392	I393	R394	P395	Y396	R397	G402	D403
	N413	I416	G417	V420	P421	L422	A423		A424	ASP	SER	LYS	VAL	ARG	ARG	ALA	ARG	SER	VAL	ASP		GLY	ALA	GLY	GLN	GLY	L441	R442	I443	E444	I445	D448	A449	Q450	E451	L452	L455	G456	F457		V460		S461	A468	ASN	GLN							

4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α , β , γ	104.00Å 104.00Å 222.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	6.00 – 2.80	Depositor
% Data completeness (in resolution range)	98.0 (6.00-2.80)	Depositor
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	PROLSQ	Depositor
R, R_{free}	(Not available) , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	7089	wwPDB-VP
Average B, all atoms (Å ²)	33.0	wwPDB-VP

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	0.69	0/3620	1.86	72/4938 (1.5%)
1	B	0.68	0/3631	1.82	67/4957 (1.4%)
All	All	0.69	0/7251	1.84	139/9895 (1.4%)

There are no bond length outliers.

The worst 5 of 139 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	B	24	ARG	NE-CZ-NH2	18.22	129.41	120.30
1	A	394	ARG	CD-NE-CZ	16.01	146.01	123.60
1	B	152	ARG	NE-CZ-NH1	14.62	127.61	120.30
1	A	152	ARG	NE-CZ-NH2	-14.48	113.06	120.30
1	B	391	ARG	NE-CZ-NH1	14.31	127.45	120.30

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	3523	0	3357	193	0
1	B	3533	0	3366	218	0
2	A	18	0	0	5	0
2	B	15	0	0	2	0
All	All	7089	0	6723	401	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 29.

The worst 5 of 401 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:B:351:LYS:HD3	1:B:357:TYR:CE2	1.88	1.09
1:B:362:ARG:HG3	1:B:374:ASN:ND2	1.69	1.06
1:B:100:ASP:O	1:B:104:ARG:HG3	1.59	1.03
1:B:292:PRO:HG2	1:B:417:GLY:HA3	1.45	0.97
1:B:14:LEU:HB2	1:B:18:VAL:CG1	1.97	0.95

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	443/470 (94%)	401 (90%)	39 (9%)	3 (1%)	26	62
1	B	447/470 (95%)	404 (90%)	35 (8%)	8 (2%)	11	34
All	All	890/940 (95%)	805 (90%)	74 (8%)	11 (1%)	16	47

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	86	PRO
1	B	190	THR
1	B	362	ARG
1	A	236	SER
1	B	58	GLY

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	378/392 (96%)	333 (88%)	45 (12%)	6	19
1	B	379/392 (97%)	325 (86%)	54 (14%)	4	12
All	All	757/784 (97%)	658 (87%)	99 (13%)	5	15

5 of 99 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	470	GLN
1	B	66	LYS
1	B	368	VAL
1	B	10	ARG
1	B	24	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 29 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	62	ASN
1	B	195	GLN
1	B	387	ASN
1	B	72	ASN
1	B	206	ASN

5.3.3 RNA ⓘ

There are no RNA molecules 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 [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 ⓘ

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates ⓘ

EDS was not executed - this section will therefore be empty.

6.4 Ligands ⓘ

EDS was not executed - this section will therefore be empty.

6.5 Other polymers ⓘ

EDS was not executed - this section will therefore be empty.