wwPDB X-ray Structure Validation Summary Report (i)

Sep 28, 2023 – 10:14 AM EDT

PDB ID	:	2E0Z
Title	:	Crystal structure of virus-like particle from Pyrococcus furiosus
Authors	:	Akita, F.; Chong, K.T.; Tanaka, H.; Yamashita, E.; Miyazaki, N.; Nakaishi, Y.; Namba, K.; Ono, Y.; Suzuki, M.; Tsukihara, T.; Nakagawa, A.
Deposited on	:	2006-10-16
Resolution	:	3.60 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

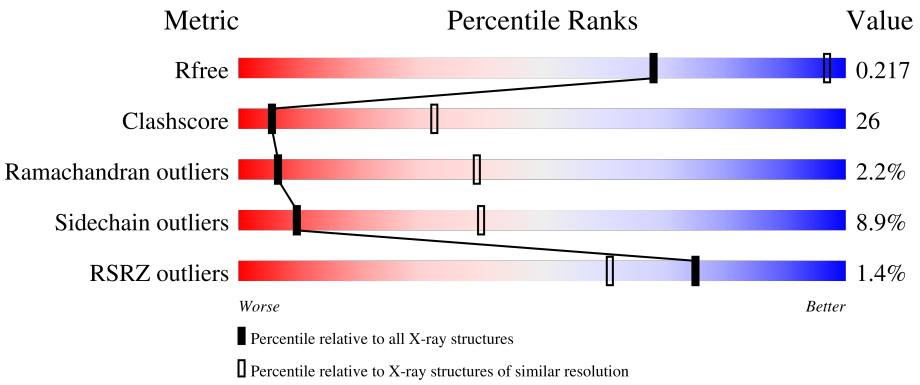
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 3.60 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	1257 (3.70-3.50)
Clashscore	141614	1353 (3.70-3.50)
Ramachandran outliers	138981	1307 (3.70-3.50)
Sidechain outliers	138945	1307 (3.70-3.50)
RSRZ outliers	127900	1161 (3.70-3.50)

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	345	37% 28% • 32%
1	B	345	% 39% 25% • • 32%
1	C	345	% 37% 24% • 34%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 5488 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 Virus-like particle.

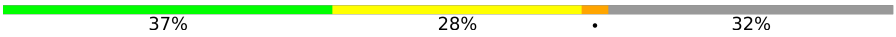
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	236	1851	1191	307	350	3	0	0	0
			Total	C	N	O	S			
1	B	236	1851	1191	307	350	3	0	0	0
			Total	C	N	O	S			
1	C	228	1786	1149	297	338	2	0	0	0

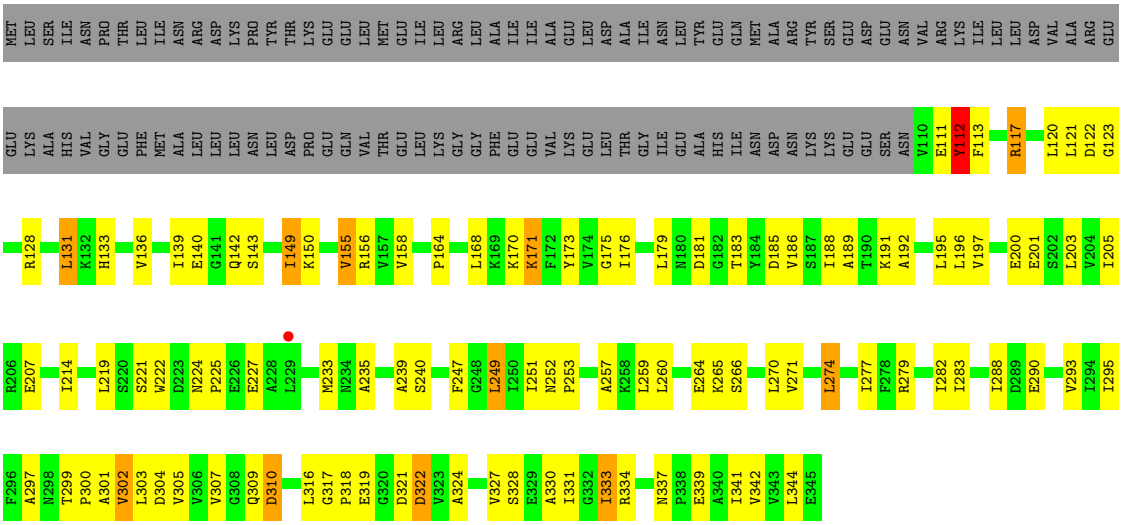
3 Residue-property plots (i)

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: Virus-like particle

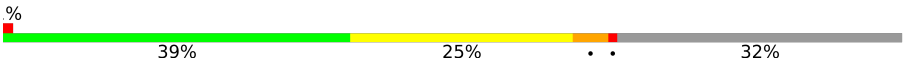
Chain A:

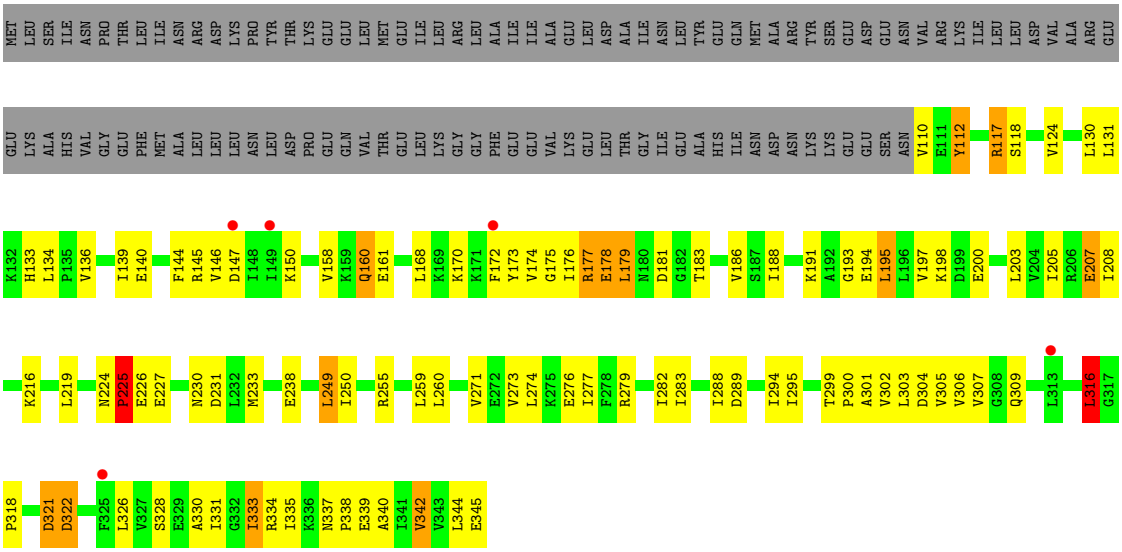




• Molecule 1: Virus-like particle

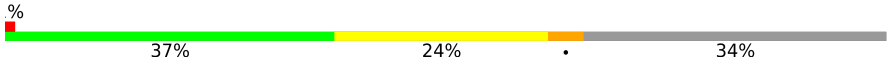
Chain B:





- Molecule 1: Virus-like particle

Chain C:



MET LEU SER ILE ASN PRO THR LEU ILE ASN ARG ASP LYS PRO TYR THR LYS GLU GLU LEU MET GLU ILE LEU ARG LEU ALA ILE ILE ALA GLU LEU ASP ALA ILE ASN LEU TYR GLU GLN MET ALA ARG TYR SER GLU ASP GLU ASN VAL ARG LYS ILE LEU LEU ASP VAL ALA ARG GLU

GLU LYS ALA HIS VAL GLY GLU PHE MET ALA LEU LEU LEU ASN LEU ASP PRO GLU GLN VAL THR GLU LEU LYS GLY GLY PHE GLU GLU VAL LYS GLU LEU THR GLY ILE GLU ALA HIS ILE ASN ASP ASN LYS LYS GLU GLU SER ASN V110 E111 Y112 L116 R117 L120 L121 D122 G123

V124 L131 P135 V136 T137 R138 V146 K150 V157 V158 K159 Q160 E161 Y162 L168 K169 K170 K171 V174 G175 I176 R177 E178 L179 V186 S187 I188 A189 T190 E194 V204 L209 S210 T211 E212 G213 I214 L219 W222 D223 N224 P225 E226 E227 A228 N234

A235 L236 Q237 S243 A244 G245 L249 I250 I251 N252 P253 K254 R255 Y256 A257 K258 L259 L260 K261 I262 TYR GLU LYS SER GLY LYS MET LEU V271 E272 V273 L274 I277 F278 I282 I283 L286 N287 I288 D289 E290 N291 K292 V293 I294 I295 F296 A297 N298 T299 P300 A301 V302 L303 D304

V305 V306 V307 G308 Q309 D310 L313 L316 D321 D322 V323 L326 V327 S328 E329 A330 I331 G332 I333 R334 I335 K336 N337 P338 E339 A340 I341 V342 V343 L344 E345

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	631.50Å 631.50Å 351.30Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 3.60 79.61 – 3.61	Depositor EDS
% Data completeness (in resolution range)	(Not available) (20.00-3.60) 96.8 (79.61-3.61)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.85 (at 3.58Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.268 , 0.267 0.215 , 0.217	Depositor DCC
R_{free} test set	38570 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	138.0	Xtriage
Anisotropy	0.051	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 60.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	5488	wwPDB-VP
Average B, all atoms (Å^2)	99.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.29% 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 (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	0.54	0/1876	0.67	0/2530
1	B	0.55	0/1876	0.69	0/2530
1	C	0.49	0/1809	0.66	0/2441
All	All	0.53	0/5561	0.67	0/7501

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	1851	0	1940	100	0
1	B	1851	0	1940	99	0
1	C	1786	0	1870	99	0
All	All	5488	0	5750	292	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 26.

The worst 5 of 292 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:171:LYS:H	1:C:171:LYS:HE2	1.00	1.14

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:181:ASP:HB2	1:A:183:THR:HG22	1.27	1.06
1:A:333:ILE:HD13	1:A:333:ILE:H	1.23	1.02
1:B:117:ARG:HH11	1:B:117:ARG:HB2	1.26	0.99
1:C:333:ILE:HD12	1:C:333:ILE:H	1.25	0.98

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	234/345 (68%)	205 (88%)	24 (10%)	5 (2%)	7	40
1	B	234/345 (68%)	197 (84%)	33 (14%)	4 (2%)	9	45
1	C	224/345 (65%)	192 (86%)	26 (12%)	6 (3%)	5	35
All	All	692/1035 (67%)	594 (86%)	83 (12%)	15 (2%)	6	39

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	225	PRO
1	A	112	TYR
1	A	316	LEU
1	A	318	PRO
1	B	225	PRO

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	205/302 (68%)	186 (91%)	19 (9%)	9 38
1	B	205/302 (68%)	188 (92%)	17 (8%)	11 42
1	C	198/302 (66%)	180 (91%)	18 (9%)	9 39
All	All	608/906 (67%)	554 (91%)	54 (9%)	9 40

5 of 54 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	207	GLU
1	B	342	VAL
1	C	303	LEU
1	B	225	PRO
1	B	321	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	133	HIS
1	C	234	ASN
1	C	291	ASN
1	C	287	ASN
1	A	337	ASN

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	236/345 (68%)	0.10	1 (0%)	92	86	63, 92, 121, 160	0
1	B	236/345 (68%)	0.12	5 (2%)	63	48	64, 90, 120, 155	0
1	C	228/345 (66%)	0.11	4 (1%)	68	53	72, 102, 181, 195	0
All	All	700/1035 (67%)	0.11	10 (1%)	75	61	63, 95, 145, 195	0

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	243	SER	2.8
1	C	313	LEU	2.6
1	B	313	LEU	2.5
1	A	229	LEU	2.5
1	B	172	PHE	2.2

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.