Full wwPDB X-ray Structure Validation Report (i)

Sep 28, 2024 – 08:23 AM EDT

PDB ID	:	3EOQ
Title	:	The crystal structure of putative zinc protease beta-subunit from Thermus thermophilus HB8
Authors	:	Ohtsuka, J.; Ichihara, Y.; Ebihara, A.; Yokoyama, S.; Kuramitsu, S.; Nagata, K.; Tanokura, M.
Deposited on	:	2008-09-29
Resolution	:	2.29 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 (i)) were used in the production of this report:

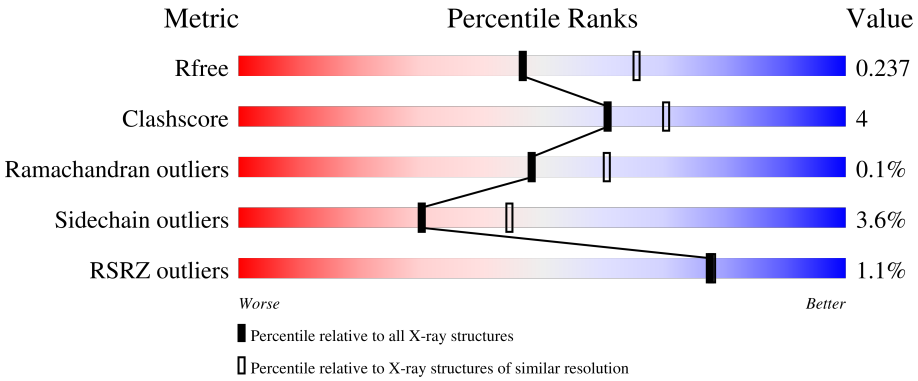
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 2.29 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}	164625	5963 (2.30-2.30)
Clashscore	180529	6698 (2.30-2.30)
Ramachandran outliers	177936	6640 (2.30-2.30)
Sidechain outliers	177891	6640 (2.30-2.30)
RSRZ outliers	164620	5963 (2.30-2.30)

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	406	89% 10%
1	B	406	2% 91% 8%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 6739 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 Putative zinc protease.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	405	Total	C	N	O	Se	0	0	0
			3215	2049	564	593	9			
1	B	406	Total	C	N	O	Se	0	0	0
			3220	2052	565	594	9			

- Molecule 2 is water.

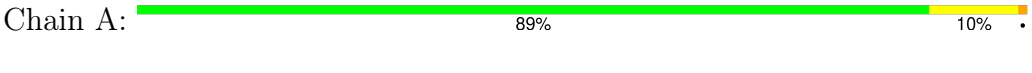
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	139	Total	O	0	0
			139	139		
2	B	165	Total	O	0	0
			165	165		

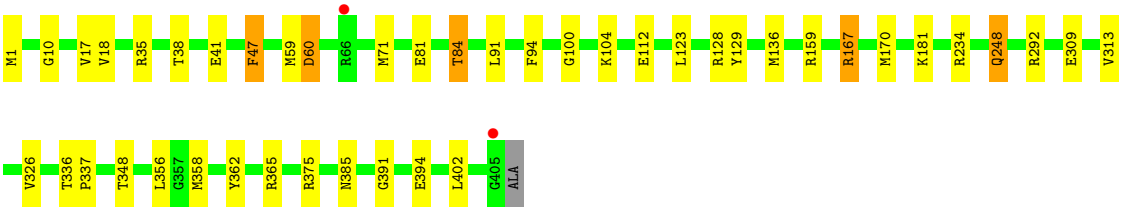
3 Residue-property plots (i)

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: Putative zinc protease

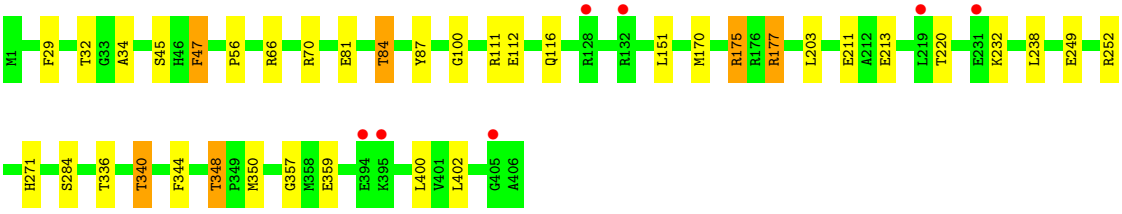
Chain A:





• Molecule 1: Putative zinc protease

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	70.34Å 94.20Å 134.45Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.62 – 2.29 48.62 – 2.29	Depositor EDS
% Data completeness (in resolution range)	99.6 (48.62-2.29) 99.6 (48.62-2.29)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	5.93 (at 2.29Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.187 , 0.237 0.187 , 0.237	Depositor DCC
R_{free} test set	2014 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å^2)	24.1	Xtriage
Anisotropy	0.111	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 31.7	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	0.94	EDS
Total number of atoms	6739	wwPDB-VP
Average B, all atoms (Å^2)	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.40% 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	0.56	0/3278	0.64	0/4417
1	B	0.58	0/3283	0.64	2/4424 (0.0%)
All	All	0.57	0/6561	0.64	2/8841 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	175	ARG	NE-CZ-NH2	-6.12	117.24	120.30
1	B	177	ARG	NE-CZ-NH2	-5.21	117.70	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-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 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	3215	0	3166	29	0
1	B	3220	0	3171	24	0
2	A	139	0	0	2	0
2	B	165	0	0	3	0
All	All	6739	0	6337	52	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 4.

All (52) 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:167:ARG:HG3	1:A:167:ARG:HH21	1.41	0.83
1:A:365:ARG:HH22	1:A:375:ARG:HH12	1.28	0.82
1:A:326:VAL:H	1:A:385:ASN:HD21	1.26	0.82
1:A:336:THR:HB	1:A:337:PRO:HD3	1.68	0.75
1:A:123:LEU:HD22	1:A:159:ARG:HG3	1.74	0.70
1:A:84:THR:HG21	2:A:407:HOH:O	1.91	0.69
1:A:248:GLN:HE22	1:A:292:ARG:H	1.40	0.68
1:B:350:MSE:HE3	1:B:350:MSE:HA	1.75	0.67
1:B:84:THR:HG21	2:B:410:HOH:O	1.95	0.67
1:B:336:THR:O	1:B:340:THR:HG23	1.96	0.65
1:A:167:ARG:HH21	1:A:167:ARG:CG	2.10	0.65
1:B:87:TYR:HD1	1:B:350:MSE:HE2	1.62	0.64
1:B:336:THR:O	1:B:340:THR:CG2	2.50	0.60
1:B:177:ARG:NH2	1:B:213:GLU:O	2.37	0.57
1:A:10:GLY:HA3	1:A:181:LYS:HG2	1.88	0.55
1:A:129:TYR:CD1	1:A:136:MSE:HE1	2.42	0.55
1:A:91:LEU:HG	1:B:340:THR:HB	1.88	0.54
1:A:129:TYR:CD1	1:A:136:MSE:CE	2.91	0.54
1:A:365:ARG:NH2	1:A:375:ARG:HH12	2.01	0.53
1:A:71:MSE:HG2	1:A:94:PHE:HB3	1.91	0.53
1:A:309:GLU:O	1:A:313:VAL:HG23	2.09	0.53
1:B:32:THR:OG1	1:B:177:ARG:CD	2.57	0.52
1:B:249:GLU:O	1:B:252:ARG:HG2	2.09	0.52
1:B:32:THR:OG1	1:B:177:ARG:HD3	2.11	0.51
1:B:87:TYR:CD1	1:B:350:MSE:HE2	2.46	0.51
1:A:100:GLY:O	1:A:104:LYS:HD3	2.10	0.51
1:A:248:GLN:NE2	1:A:292:ARG:H	2.06	0.51
1:B:238:LEU:HD13	1:B:400:LEU:HG	1.92	0.50
1:B:112:GLU:OE1	1:B:116:GLN:NE2	2.43	0.50
1:A:358:MSE:HE1	2:A:436:HOH:O	2.13	0.49
1:A:391:GLY:O	1:A:394:GLU:HG2	2.12	0.48
1:A:129:TYR:CE1	1:A:136:MSE:CE	2.97	0.48
1:A:35:ARG:HG2	1:A:81:GLU:O	2.15	0.47
1:A:248:GLN:HE21	1:A:248:GLN:H	1.62	0.46
1:B:271:HIS:HD2	2:B:430:HOH:O	1.98	0.46
1:A:112:GLU:HG2	1:A:167:ARG:CZ	2.47	0.45
1:B:344:PHE:O	1:B:348:THR:HG23	2.16	0.45
1:B:220:THR:HG23	1:B:220:THR:O	2.18	0.44
1:B:56:PRO:O	1:B:111:ARG:NH1	2.51	0.44
1:A:47:PHE:CG	1:A:170:MSE:HE1	2.52	0.44
1:A:365:ARG:HE	1:A:365:ARG:HB3	1.56	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:47:PHE:CG	1:B:170:MSE:HE1	2.53	0.43
1:A:129:TYR:CE1	1:A:136:MSE:HE1	2.52	0.43
1:B:29:PHE:CZ	1:B:357:GLY:HA3	2.54	0.42
1:A:1:MSE:N	1:A:17:VAL:O	2.52	0.42
1:B:66:ARG:O	1:B:70:ARG:HG3	2.19	0.42
1:A:358:MSE:O	1:A:362:TYR:HD1	2.03	0.41
1:B:271:HIS:CD2	2:B:430:HOH:O	2.73	0.41
1:A:38:THR:OG1	1:A:41:GLU:HG2	2.21	0.41
1:B:34:ALA:HB3	1:B:81:GLU:HA	2.03	0.41
1:B:100:GLY:HA2	1:B:203:LEU:HD13	2.03	0.40
1:B:252:ARG:HH22	1:B:359:GLU:CD	2.24	0.40

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	403/406 (99%)	397 (98%)	5 (1%)	1 (0%)	44 55
1	B	404/406 (100%)	401 (99%)	3 (1%)	0	100 100
All	All	807/812 (99%)	798 (99%)	8 (1%)	1 (0%)	48 60

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	60	ASP

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	319/310 (103%)	307 (96%)	12 (4%)	28	42
1	B	319/310 (103%)	308 (97%)	11 (3%)	32	47
All	All	638/620 (103%)	615 (96%)	23 (4%)	30	44

All (23) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	18	VAL
1	A	47	PHE
1	A	59	MSE
1	A	60	ASP
1	A	84	THR
1	A	128	ARG
1	A	167	ARG
1	A	234	ARG
1	A	248	GLN
1	A	348	THR
1	A	356	LEU
1	A	402	LEU
1	B	45	SER
1	B	47	PHE
1	B	84	THR
1	B	151	LEU
1	B	175	ARG
1	B	211	GLU
1	B	232	LYS
1	B	284	SER
1	B	340	THR
1	B	348	THR
1	B	402	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (10) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	76	ASN
1	A	147	GLN
1	A	248	GLN
1	A	256	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	271	HIS
1	A	315	GLN
1	A	385	ASN
1	B	153	ASN
1	B	260	HIS
1	B	271	HIS

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

There are no oligosaccharides 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, 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	396/406 (97%)	-0.27	2 (0%)	87	88	12, 22, 36, 46	15 (3%)
1	B	397/406 (97%)	-0.32	7 (1%)	67	68	11, 21, 33, 49	13 (3%)
All	All	793/812 (97%)	-0.29	9 (1%)	77	78	11, 21, 35, 49	28 (3%)

All (9) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	66	ARG	4.4
1	B	132	ARG	4.1
1	A	405	GLY	3.4
1	B	231	GLU	3.4
1	B	219	LEU	3.2
1	B	128	ARG	2.8
1	B	395	LYS	2.6
1	B	394	GLU	2.5
1	B	405	GLY	2.2

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 monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.