wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 10:37 PM GMT

PDB ID	:	1UFO
Title	:	Crystal Structure of TT1662 from Thermus thermophilus
Authors	:	Murayama, K.; Shirouzu, M.; Terada, T.; Kuramitsu, S.; Yokoyama, S.; RIKEN Structural Genomics/Proteomics Initiative (RSGI)
Deposited on	:	2003-06-03
Resolution	:	1.60 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

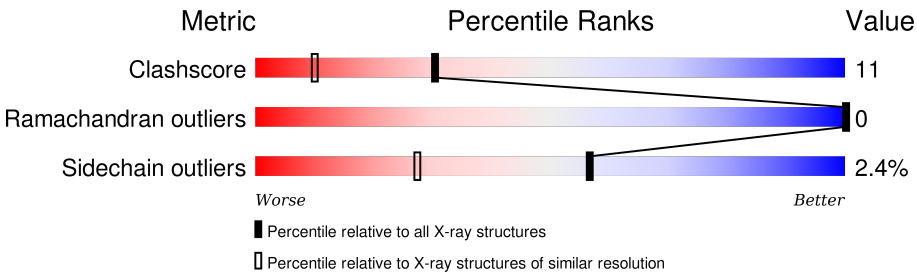
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.60 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))
Clashscore	102246	2732 (1.60-1.60)
Ramachandran outliers	100387	2654 (1.60-1.60)
Sidechain outliers	100360	2653 (1.60-1.60)

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 <=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	238	89% 10% •
1	B	238	75% 23% •
1	C	238	83% 15% •
1	D	238	87% 11% •
1	E	238	89% 9% •
1	F	238	84% 13% •

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 12988 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 hypothetical protein TT1662.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	Se			
1	A	238	1861	1197	343	317	4	0	2	0
1	B	238	1850	1191	339	316	4	0	1	0
1	C	238	1866	1203	341	318	4	0	3	0
1	D	238	1857	1195	340	318	4	0	2	0
1	E	238	1861	1197	343	317	4	0	2	0
1	F	238	1858	1197	340	317	4	0	2	0

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	319	319	319	0	0
2	B	258	258	258	0	0
2	C	327	327	327	0	0
2	D	331	331	331	0	0
2	E	299	299	299	0	0
2	F	301	301	301	0	0

3 Residue-property plots (i)

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.

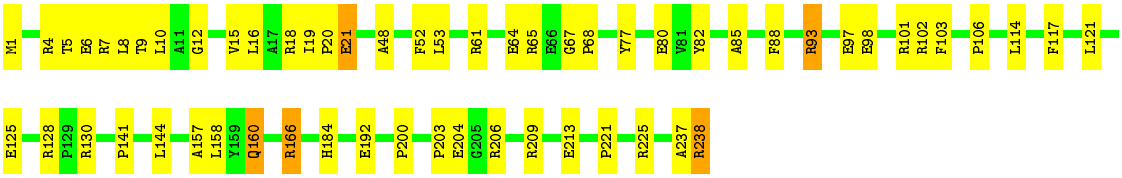
Note EDS was not executed.

- Molecule 1: hypothetical protein TT1662

Chain A:

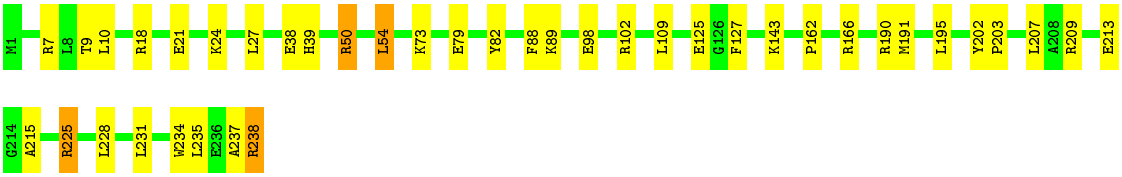
- Molecule 1: hypothetical protein TT1662

Chain B:



- Molecule 1: hypothetical protein TT1662

Chain C:



- Molecule 1: hypothetical protein TT1662

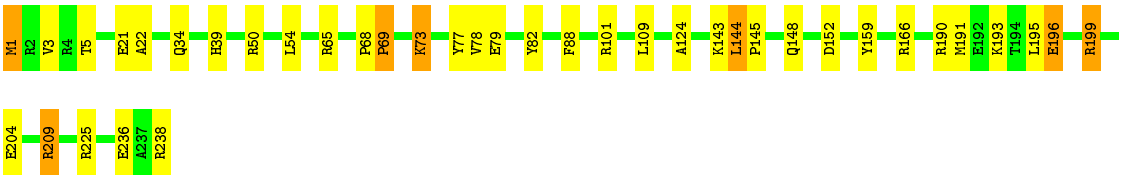
Chain D:

- Molecule 1: hypothetical protein TT1662

Chain E:

- Molecule 1: hypothetical protein TT1662

Chain F:



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, α, β, γ	129.86Å 129.86Å 70.12Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	21.99 – 1.60	Depositor
% Data completeness (in resolution range)	99.7 (21.99-1.60)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.170 , 0.207	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	12988	wwPDB-VP
Average B, all atoms (Å^2)	18.0	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	1.08	1/1906 (0.1%)	1.10	5/2575 (0.2%)
1	B	1.11	1/1895 (0.1%)	1.11	5/2561 (0.2%)
1	C	1.16	2/1911 (0.1%)	1.12	5/2583 (0.2%)
1	D	1.12	6/1902 (0.3%)	1.07	5/2571 (0.2%)
1	E	1.07	2/1906 (0.1%)	1.06	4/2575 (0.2%)
1	F	1.10	2/1903 (0.1%)	1.09	10/2573 (0.4%)
All	All	1.11	14/11423 (0.1%)	1.09	34/15438 (0.2%)

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
1	C	0	1
1	D	0	1
1	F	0	2
All	All	0	7

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	192	GLU	CD-OE1	7.70	1.34	1.25
1	F	196	GLU	CD-OE2	6.70	1.33	1.25
1	D	204	GLU	CB-CG	6.52	1.64	1.52
1	B	160	GLN	CG-CD	5.92	1.64	1.51
1	D	204	GLU	CG-CD	5.83	1.60	1.51

The worst 5 of 34 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	225	ARG	NE-CZ-NH2	9.13	124.87	120.30
1	D	199	ARG	NE-CZ-NH2	-8.95	115.82	120.30
1	E	206	ARG	NE-CZ-NH1	-8.03	116.28	120.30
1	B	128	ARG	NE-CZ-NH1	-7.50	116.55	120.30
1	E	101	ARG	NE-CZ-NH1	-7.25	116.68	120.30

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	82	TYR	Sidechain
1	B	77	TYR	Sidechain
1	B	82	TYR	Sidechain
1	C	82	TYR	Sidechain
1	D	82	TYR	Sidechain

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	1861	0	1905	45	0
1	B	1850	0	1893	72	0
1	C	1866	0	1913	47	0
1	D	1857	0	1899	24	0
1	E	1861	0	1905	32	0
1	F	1858	0	1904	36	0
2	A	319	0	0	24	0
2	B	258	0	0	16	0
2	C	327	0	0	12	0
2	D	331	0	0	11	0
2	E	299	0	0	12	0
2	F	301	0	0	15	0
All	All	12988	0	11419	249	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 11.

The worst 5 of 249 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:238:ARG:H	1:C:238:ARG:CD	1.49	1.26
1:C:238:ARG:N	1:C:238:ARG:HD2	1.38	1.19
1:C:10[A]:LEU:HD12	1:C:98:GLU:HG3	1.32	1.09
1:A:73:LYS:HE2	2:A:497:HOH:O	1.53	1.08
1:F:166:ARG:HD2	2:F:425:HOH:O	1.54	1.07

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	238/238 (100%)	234 (98%)	4 (2%)	0	100	100
1	B	237/238 (100%)	232 (98%)	5 (2%)	0	100	100
1	C	239/238 (100%)	237 (99%)	2 (1%)	0	100	100
1	D	238/238 (100%)	236 (99%)	2 (1%)	0	100	100
1	E	238/238 (100%)	236 (99%)	2 (1%)	0	100	100
1	F	238/238 (100%)	234 (98%)	4 (2%)	0	100	100
All	All	1428/1428 (100%)	1409 (99%)	19 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	185/179 (103%)	183 (99%)	2 (1%)	80	63
1	B	184/179 (103%)	177 (96%)	7 (4%)	40	13
1	C	186/179 (104%)	181 (97%)	5 (3%)	52	23
1	D	185/179 (103%)	182 (98%)	3 (2%)	70	47
1	E	185/179 (103%)	182 (98%)	3 (2%)	70	47
1	F	185/179 (103%)	178 (96%)	7 (4%)	40	13
All	All	1110/1074 (103%)	1083 (98%)	27 (2%)	57	27

5 of 27 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	162	PRO
1	D	160	GLN
1	F	73	LYS
1	C	238	ARG
1	B	88	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	34	GLN
1	A	201	HIS
1	E	160	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 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 (i)

6.1 Protein, DNA and RNA chains (i)

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

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

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

6.3 Carbohydrates (i)

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

6.4 Ligands (i)

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

6.5 Other polymers (i)

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