Full wwPDB X-ray Structure Validation Report ⓘ

Dec 7, 2023 – 07:11 am GMT

PDB ID	:	2JEB
Title	:	Structure of a 9-subunit archaeal exosome bound to Mn ions
Authors	:	Lorentzen, E.; Conti, E.
Deposited on	:	2007-01-16
Resolution	:	2.40 Å(reported)

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

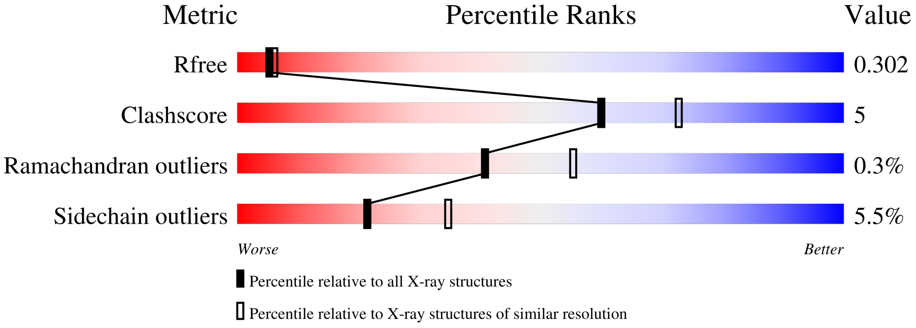
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)

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%

Mol	Chain	Length	Quality of chain
1	A	277	
2	B	250	
3	I	251	

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 5434 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 EXOSOME COMPLEX EXONUCLEASE 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	273	Total 2058	C 1313	N 335	O 405	S 5	0	1	0

- Molecule 2 is a protein called EXOSOME COMPLEX EXONUCLEASE 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	233	Total 1773	C 1123	N 307	O 333	S 10	0	1	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	182	ALA	ASP	engineered mutation	UNP Q9UXC2

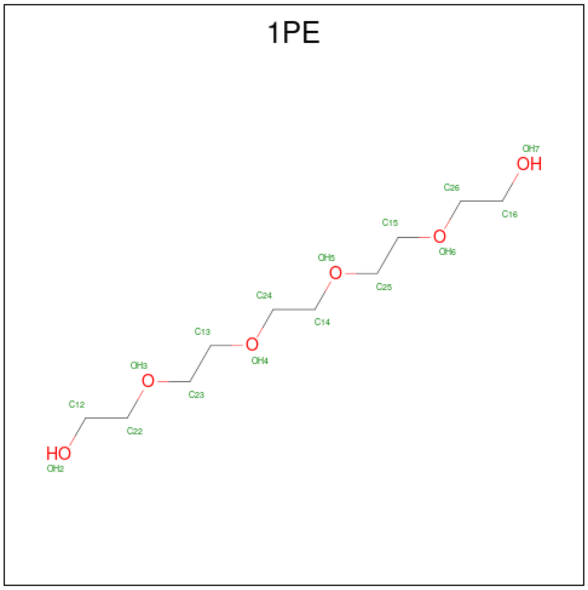
- Molecule 3 is a protein called EXOSOME COMPLEX RNA-BINDING PROTEIN 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	I	197	Total 1416	C 924	N 229	O 261	S 2	0	0	0

- Molecule 4 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total 1	Mn 1	0	0
4	I	1	Total 1	Mn 1	0	0

- Molecule 5 is PENTAETHYLENE GLYCOL (three-letter code: 1PE) (formula: $C_{10}H_{22}O_6$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 16 10 6	0	0

- Molecule 6 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	B	1	Total Cl 1 1	0	0

- Molecule 7 is water.

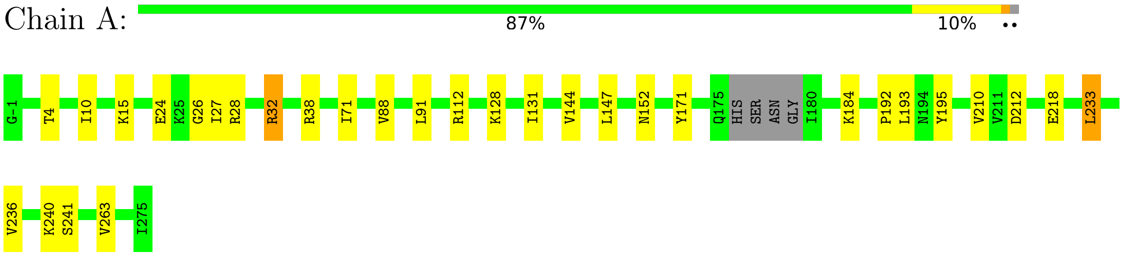
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	97	Total O 97 97	0	0
7	B	55	Total O 55 55	0	0
7	I	16	Total O 16 16	0	0

3 Residue-property plots ⓘ

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. 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. 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.

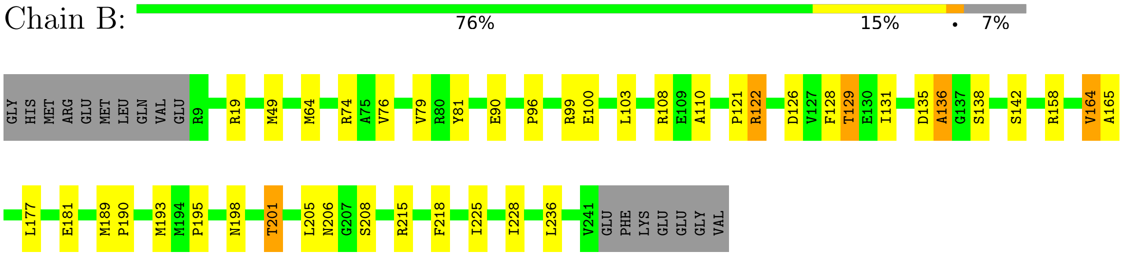
- Molecule 1: EXOSOME COMPLEX EXONUCLEASE 2

Chain A:



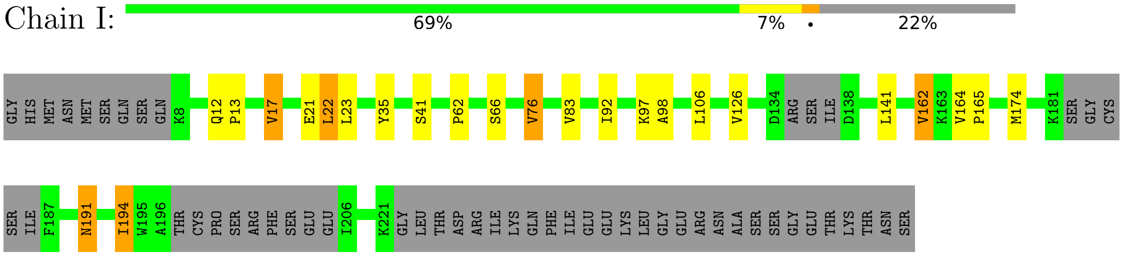
- Molecule 2: EXOSOME COMPLEX EXONUCLEASE 1

Chain B:



- Molecule 3: EXOSOME COMPLEX RNA-BINDING PROTEIN 1

Chain I:



4 Data and refinement statistics ⓘ

Property	**Value**	**Source**
Space group	P 21 3	Depositor
Cell constants a, b, c, α, β, γ	133.64Å 133.64Å 133.64Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.54 – 2.40 42.26 – 2.40	Depositor EDS
% Data completeness (in resolution range)	100.0 (44.54-2.40) 100.0 (42.26-2.40)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.56 (at 2.39Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.198 , 0.258 0.250 , 0.302	Depositor DCC
R_{free} test set	1567 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	44.4	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 48.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.034 for l,-k,h	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	5434	wwPDB-VP
Average B, all atoms (Å^2)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.56% 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 ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: 1PE, CL, MN

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	0.46	0/2093	0.64	2/2851 (0.1%)
2	B	0.50	0/1803	0.70	2/2440 (0.1%)
3	I	0.39	0/1441	0.56	0/1972
All	All	0.46	0/5337	0.64	4/7263 (0.1%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	215	ARG	NE-CZ-NH1	6.68	123.64	120.30
2	B	215	ARG	NE-CZ-NH2	-6.60	117.00	120.30
1	A	112	ARG	NE-CZ-NH1	5.47	123.03	120.30
1	A	112	ARG	NE-CZ-NH2	-5.07	117.77	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	2058	0	2074	15	0
2	B	1773	0	1805	24	0
3	I	1416	0	1340	10	0
4	A	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	I	1	0	0	0	0
5	A	16	0	22	0	0
6	B	1	0	0	0	0
7	A	97	0	0	4	0
7	B	55	0	0	2	0
7	I	16	0	0	0	0
All	All	5434	0	5241	48	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 5.

All (48) 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)
2:B:129:THR:HG21	2:B:142:SER:OG	1.75	0.87
3:I:92:ILE:O	3:I:191:ASN:ND2	2.18	0.76
2:B:110:ALA:HB1	2:B:201:THR:HG23	1.74	0.68
1:A:240:LYS:NZ	7:A:2087:HOH:O	2.19	0.60
1:A:218:GLU:OE2	7:A:2080:HOH:O	2.17	0.59
1:A:10:ILE:HG22	7:A:2007:HOH:O	2.03	0.59
2:B:19:ARG:HD2	2:B:181:GLU:OE2	2.02	0.58
2:B:164:VAL:HG22	2:B:225:ILE:HG13	1.85	0.58
2:B:164:VAL:HG13	2:B:228:ILE:CD1	2.35	0.57
2:B:81:TYR:HA	2:B:129:THR:HG22	1.87	0.56
1:A:91:LEU:HD12	1:A:147:LEU:HD23	1.87	0.56
2:B:198:ASN:O	2:B:198:ASN:ND2	2.42	0.52
3:I:194:ILE:HG13	3:I:194:ILE:O	2.08	0.52
1:A:233:LEU:HD21	1:A:263:VAL:HG22	1.92	0.50
2:B:164:VAL:HG13	2:B:228:ILE:HD11	1.95	0.49
2:B:96:PRO:HA	2:B:100:GLU:OE1	2.13	0.48
3:I:17:VAL:HG21	3:I:23:LEU:CD2	2.43	0.48
1:A:28:ARG:NH1	1:A:210:VAL:HG13	2.28	0.48
2:B:189:MET:HE2	2:B:205:LEU:HD22	1.95	0.47
1:A:171:TYR:CD2	1:A:184:LYS:HA	2.49	0.47
2:B:49:MET:CE	7:B:2032:HOH:O	2.62	0.47
3:I:76:VAL:HG22	3:I:191:ASN:CG	2.35	0.46
2:B:110:ALA:CB	2:B:201:THR:HG23	2.45	0.45
1:A:152:ASN:ND2	1:A:212:ASP:HA	2.32	0.45
1:A:71:ILE:HD12	1:A:131:ILE:HD13	1.99	0.44
2:B:121:PRO:O	2:B:122:ARG:HG2	2.16	0.44
2:B:79:VAL:HG21	2:B:108:ARG:HA	1.99	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:26:GLY:O	1:A:27:ILE:HD12	2.18	0.44
2:B:126:ASP:HB3	2:B:128:PHE:CE1	2.53	0.44
2:B:198:ASN:O	2:B:198:ASN:CG	2.56	0.44
3:I:164:VAL:N	3:I:165:PRO:HD2	2.34	0.43
3:I:12:GLN:HE22	3:I:21:GLU:HG2	1.83	0.42
1:A:32:ARG:CZ	1:A:38:ARG:HG3	2.49	0.42
2:B:99:ARG:O	2:B:103:LEU:HG	2.19	0.42
1:A:24:GLU:HG3	7:A:2010:HOH:O	2.20	0.42
3:I:98:ALA:HB1	3:I:141:LEU:HG	2.02	0.42
2:B:64:MET:HE1	2:B:76:VAL:CG2	2.49	0.42
2:B:193:MET:O	2:B:195:PRO:HD3	2.19	0.42
2:B:131:ILE:HD12	2:B:138:SER:HB2	2.01	0.41
3:I:12:GLN:NE2	3:I:22:LEU:O	2.53	0.41
3:I:62:PRO:O	3:I:162:VAL:HG13	2.21	0.41
3:I:35:TYR:CZ	3:I:62:PRO:HG3	2.56	0.41
1:A:4:THR:HG21	2:B:74:ARG:CZ	2.51	0.41
2:B:135:ASP:O	2:B:136:ALA:HB3	2.21	0.41
1:A:88:VAL:HG22	1:A:144:VAL:HB	2.03	0.40
2:B:208:SER:HB3	7:B:2045:HOH:O	2.20	0.40
2:B:165:ALA:HA	2:B:190:PRO:HA	2.03	0.40
1:A:128:LYS:HD3	1:A:192:PRO:HB2	2.03	0.40

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	270/277 (98%)	261 (97%)	9 (3%)	0	100	100
2	B	232/250 (93%)	225 (97%)	6 (3%)	1 (0%)	34	48
3	I	189/251 (75%)	178 (94%)	10 (5%)	1 (0%)	29	41
All	All	691/778 (89%)	664 (96%)	25 (4%)	2 (0%)	41	55

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	I	13	PRO
2	B	136	ALA

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	226/243 (93%)	219 (97%)	7 (3%)	40	60
2	B	186/208 (89%)	175 (94%)	11 (6%)	19	32
3	I	137/223 (61%)	124 (90%)	13 (10%)	8	12
All	All	549/674 (82%)	518 (94%)	31 (6%)	21	34

All (31) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	15	LYS
1	A	32	ARG
1	A	193	LEU
1	A	195	TYR
1	A	233	LEU
1	A	236	VAL
1	A	241	SER
2	B	90	GLU
2	B	122	ARG
2	B	129	THR
2	B	158	ARG
2	B	164	VAL
2	B	177	LEU
2	B	201	THR
2	B	206[A]	ASN
2	B	206[B]	ASN
2	B	218	PHE
2	B	236	LEU
3	I	17	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	I	22	LEU
3	I	41	SER
3	I	66	SER
3	I	76	VAL
3	I	83	VAL
3	I	97	LYS
3	I	106	LEU
3	I	126	VAL
3	I	162	VAL
3	I	174	MET
3	I	191	ASN
3	I	194	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	198	ASN
3	I	191	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 ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 4 ligands modelled in this entry, 3 are monoatomic - leaving 1 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#|Z| > 2$	Counts	RMSZ	$\#|Z| > 2$
5	1PE	A	401	-	15,15,15	0.50	0	14,14,14	0.24	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	1PE	A	401	-	-	2/13/13/13	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	401	1PE	C14-C24-OH4-C13
5	A	401	1PE	C16-C26-OH6-C15

There are no ring outliers.

No monomer is involved in short contacts.

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 ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.