



Full wwPDB X-ray Structure Validation Report

Feb 27, 2014 – 07:57 PM GMT

PDB ID : 1YDX
Title : Crystal structure of Type-I restriction-modificationsystem S subunit from M. genitalium
Authors : Machado, B.; Quijada, O.; Pinol, J.; Fita, I.; Querol, E.; Carpena, X.
Deposited on : 2004-12-27
Resolution : 2.30 Å(reported)

This is a full wwPDB validation 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/ValidationPDFNotes.html>

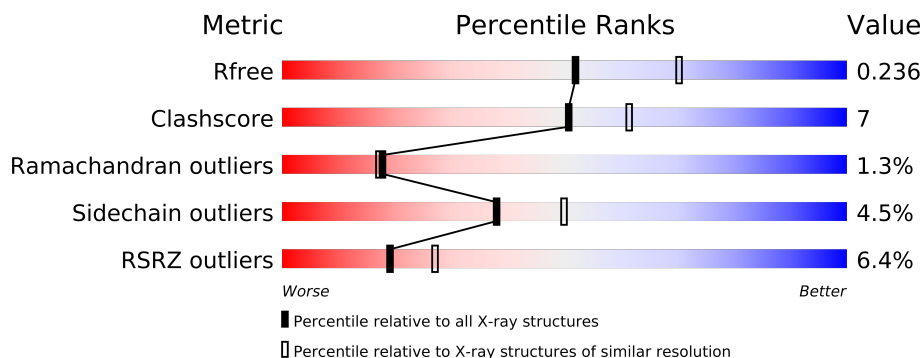
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable22639
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable22683

1 Overall quality at a glance

The reported resolution of this entry is 2.30 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}	66092	2929 (2.30-2.30)
Clashscore	79885	3679 (2.30-2.30)
Ramachandran outliers	78287	3642 (2.30-2.30)
Sidechain outliers	78261	3641 (2.30-2.30)
RSRZ outliers	66119	2930 (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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	406	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3163 atoms, of which 0 are hydrogen and 0 are deuterium.

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 type I restriction enzyme specificity protein MG438.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	374	3032	1950	500	572	4	6	0	0	0

There are 29 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-22	MSE	-	CLONING ARTIFACT	UNP Q49434
A	-21	GLY	-	CLONING ARTIFACT	UNP Q49434
A	-20	HIS	-	CLONING ARTIFACT	UNP Q49434
A	-19	HIS	-	CLONING ARTIFACT	UNP Q49434
A	-18	HIS	-	CLONING ARTIFACT	UNP Q49434
A	-17	HIS	-	CLONING ARTIFACT	UNP Q49434
A	-16	HIS	-	CLONING ARTIFACT	UNP Q49434
A	-15	HIS	-	CLONING ARTIFACT	UNP Q49434
A	-14	HIS	-	CLONING ARTIFACT	UNP Q49434
A	-13	HIS	-	CLONING ARTIFACT	UNP Q49434
A	-12	HIS	-	CLONING ARTIFACT	UNP Q49434
A	-11	HIS	-	CLONING ARTIFACT	UNP Q49434
A	-10	SER	-	CLONING ARTIFACT	UNP Q49434
A	-9	SER	-	CLONING ARTIFACT	UNP Q49434
A	-8	GLY	-	CLONING ARTIFACT	UNP Q49434
A	-7	HIS	-	CLONING ARTIFACT	UNP Q49434
A	-6	ILE	-	CLONING ARTIFACT	UNP Q49434
A	-5	ASP	-	CLONING ARTIFACT	UNP Q49434
A	-4	ASP	-	CLONING ARTIFACT	UNP Q49434
A	-3	ASP	-	CLONING ARTIFACT	UNP Q49434
A	-2	ASP	-	CLONING ARTIFACT	UNP Q49434
A	-1	LYS	-	CLONING ARTIFACT	UNP Q49434
A	0	HIS	-	CLONING ARTIFACT	UNP Q49434
A	1	MSE	MET	MODIFIED RESIDUE	UNP Q49434
A	29	MSE	MET	MODIFIED RESIDUE	UNP Q49434
A	201	MSE	MET	MODIFIED RESIDUE	UNP Q49434
A	283	MSE	MET	MODIFIED RESIDUE	UNP Q49434

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	307	MSE	MET	MODIFIED RESIDUE	UNP Q49434
A	373	MSE	MET	MODIFIED RESIDUE	UNP Q49434

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	2	Total Cl 2 2	0	0

- Molecule 3 is water.

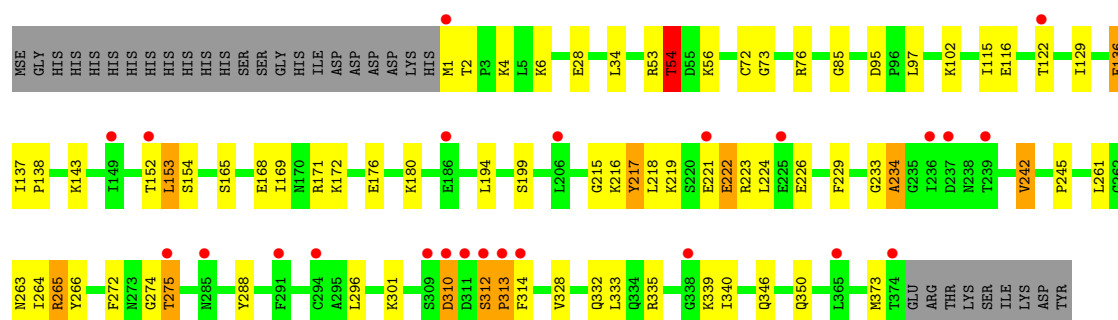
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	129	Total O 129 129	0	0

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.

- Molecule 1: type I restriction enzyme specificity protein MG438

Chain A: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α , β , γ	76.60Å 76.60Å 174.85Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 2.30 24.82 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.8 (20.00-2.30) 99.6 (24.82-2.20)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.96 (at 2.19Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.197 , 0.231 0.207 , 0.236	Depositor DCC
R_{free} test set	1558 reflections (5.32%)	DCC
Wilson B-factor (Å ²)	43.6	Xtriage
Anisotropy	0.015	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 60.6	EDS
Estimated twinning fraction	0.033 for -h,-k,l	Xtriage
L-test for twinning	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	1 of 30971 reflections (0.003%)	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	3163	wwPDB-VP
Average B, all atoms (Å ²)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.11% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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

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	1.07	10/3088 (0.3%)	0.92	11/4149 (0.3%)

All (10) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	153	LEU	C-N	16.68	1.72	1.34
1	A	223	ARG	CZ-NH1	13.81	1.51	1.33
1	A	223	ARG	NE-CZ	13.40	1.50	1.33
1	A	223	ARG	CZ-NH2	12.73	1.49	1.33
1	A	223	ARG	CD-NE	9.90	1.63	1.46
1	A	221	GLU	CD-OE1	8.03	1.34	1.25
1	A	219	LYS	C-N	7.75	1.51	1.34
1	A	152	THR	C-N	6.45	1.48	1.34
1	A	221	GLU	CD-OE2	5.30	1.31	1.25
1	A	136	GLU	CG-CD	-5.27	1.44	1.51

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	223	ARG	NE-CZ-NH2	-23.69	108.45	120.30
1	A	223	ARG	NE-CZ-NH1	12.61	126.61	120.30
1	A	153	LEU	O-C-N	-11.22	104.75	122.70
1	A	76	ARG	NE-CZ-NH2	-9.85	115.38	120.30
1	A	76	ARG	NE-CZ-NH1	9.77	125.18	120.30
1	A	223	ARG	CD-NE-CZ	-8.02	112.37	123.60
1	A	53	ARG	NE-CZ-NH2	-6.08	117.26	120.30
1	A	152	THR	C-N-CA	5.64	135.80	121.70
1	A	53	ARG	NE-CZ-NH1	5.20	122.90	120.30
1	A	95	ASP	CB-CG-OD1	5.14	122.92	118.30
1	A	54	THR	CB-CA-C	-5.01	98.07	111.60

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3032	0	3051	45	0
2	A	2	0	0	0	0
3	A	129	0	0	2	0
All	All	3163	0	3051	45	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 7.

All (45) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:153:LEU:C	1:A:154:SER:N	1.72	1.41
1:A:217:TYR:HB2	1:A:275:THR:HG23	1.68	0.75
1:A:34:LEU:O	1:A:54:THR:HG22	1.88	0.73
1:A:1:MSE:HE1	1:A:373:MSE:HG2	1.74	0.69
1:A:265:ARG:HD2	1:A:266:TYR:H	1.57	0.67
1:A:215:GLY:HA3	1:A:274:GLY:O	1.94	0.67
1:A:34:LEU:O	1:A:54:THR:CG2	2.45	0.65
1:A:328:VAL:HG21	1:A:333:LEU:HD23	1.79	0.65
1:A:265:ARG:CD	1:A:266:TYR:H	2.12	0.63
1:A:218:LEU:HD22	1:A:222:GLU:HG2	1.81	0.62
1:A:312:SER:N	1:A:313:PRO:HD3	2.15	0.62
1:A:335:ARG:O	1:A:339:LYS:HB2	1.99	0.62
1:A:346:GLN:O	1:A:350:GLN:HG3	2.01	0.61
1:A:153:LEU:O	1:A:154:SER:N	2.35	0.58
1:A:72:CYS:O	1:A:116:GLU:HG3	2.03	0.58
1:A:54:THR:HB	1:A:56:LYS:H	1.67	0.57
1:A:233:GLY:O	1:A:234:ALA:HB2	2.05	0.56
1:A:217:TYR:HB2	1:A:275:THR:CG2	2.35	0.54
1:A:216:LYS:O	1:A:217:TYR:HB3	2.08	0.54
1:A:97:LEU:O	1:A:143:LYS:HE2	2.12	0.50

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:312:SER:H	1:A:313:PRO:HD3	1.78	0.49
1:A:137:ILE:HB	1:A:138:PRO:HD2	1.94	0.49
1:A:301:LYS:HD2	3:A:508:HOH:O	2.13	0.48
1:A:332:GLN:OE1	1:A:335:ARG:NH2	2.46	0.48
1:A:73:GLY:HA2	1:A:115:ILE:HG22	1.95	0.47
1:A:242:VAL:HG11	1:A:272:PHE:HB2	1.97	0.47
1:A:168:GLU:O	1:A:172:LYS:HG3	2.16	0.46
1:A:288:TYR:CD2	1:A:335:ARG:HA	2.50	0.45
1:A:242:VAL:HG21	1:A:272:PHE:CD1	2.51	0.45
1:A:229:PHE:CD2	1:A:245:PRO:HD3	2.51	0.45
1:A:264:ILE:HD13	1:A:296:LEU:HB3	1.99	0.44
1:A:2:THR:CG2	1:A:6:LYS:HE3	2.48	0.43
1:A:194:LEU:HD11	1:A:340:ILE:HD12	2.01	0.43
1:A:153:LEU:CA	1:A:154:SER:N	2.73	0.43
1:A:72:CYS:O	1:A:116:GLU:CG	2.67	0.42
1:A:73:GLY:HA3	1:A:116:GLU:HG3	2.01	0.42
1:A:233:GLY:O	1:A:234:ALA:CB	2.69	0.41
1:A:28:GLU:O	1:A:85:GLY:HA2	2.20	0.41
1:A:229:PHE:HB2	1:A:242:VAL:HG12	2.02	0.41
1:A:328:VAL:CG2	1:A:333:LEU:HD23	2.48	0.41
1:A:176:GLU:HG3	1:A:180:LYS:HE2	2.02	0.41
1:A:165:SER:O	1:A:169:ILE:HG13	2.21	0.41
1:A:313:PRO:HB2	1:A:314:PHE:H	1.73	0.40
1:A:129:ILE:HA	1:A:129:ILE:HD13	1.91	0.40
1:A:4:LYS:HD2	3:A:440:HOH:O	2.22	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	372/406 (92%)	363 (98%)	4 (1%)	5 (1%)	18 17

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	234	ALA
1	A	310	ASP
1	A	313	PRO
1	A	217	TYR
1	A	312	SER

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	335/358 (94%)	320 (96%)	15 (4%)	38 50

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	54	THR
1	A	102	LYS
1	A	122	THR
1	A	136	GLU
1	A	171	ARG
1	A	199	SER
1	A	222	GLU
1	A	224	LEU
1	A	226	GLU
1	A	242	VAL
1	A	261	LEU
1	A	263	ASN
1	A	265	ARG
1	A	275	THR
1	A	310	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	238	ASN
1	A	273	ASN
1	A	285	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	350	GLN

5.3.3 RNA ⓘ

There are no RNA chains 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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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, 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	374/406 (92%)	0.42	24 (6%)	19 27	31, 40, 49, 61	0

All (24) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	312	SER	10.8
1	A	310	ASP	4.8
1	A	311	ASP	4.2
1	A	236	ILE	3.4
1	A	1	MSE	3.4
1	A	374	THR	3.2
1	A	314	PHE	3.1
1	A	313	PRO	2.9
1	A	309	SER	2.9
1	A	239	THR	2.8
1	A	365	LEU	2.7
1	A	221	GLU	2.7
1	A	294	CYS	2.6
1	A	149	ILE	2.6
1	A	152	THR	2.5
1	A	285	ASN	2.5
1	A	186	GLU	2.5
1	A	122	THR	2.5
1	A	338	GLY	2.4
1	A	206	LEU	2.1
1	A	237	ASP	2.1
1	A	275	THR	2.1
1	A	225	GLU	2.0
1	A	291	PHE	2.0

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 carbohydrates in this entry.

6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(Å ²)	Q<0.9
2	CL	A	385	1/1	0.15	-0.14	50,50,50,50	0
2	CL	A	384	1/1	0.06	-1.84	43,43,43,43	0

6.5 Other polymers ⓘ

There are no such residues in this entry.