wwPDB X-ray Structure Validation Summary Report (i)

Feb 27, 2014 – 11:52 PM GMT

PDB ID	:	3D7L
Title	:	The crystal structure of the protein lin1944 from Listeria innocua .
Authors	:	Tan, K.; Bigelow, L.; Gu, M.; Joachimiak, A.; Midwest Center for Structural Genomics (MCSG)
Deposited on	:	2008-05-21
Resolution	:	2.06 Å(reported)

1 Overall quality at a glance ⓘ

The reported resolution of this entry is 2.06 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	Percentile Ranks	Value
Rfree		0.209
Clashscore		4
Ramachandran outliers		0
Sidechain outliers		3.2%
RSRZ outliers		0.4%

Worse   *Better*

▮ Percentile relative to all X-ray structures

⧠ Percentile relative to X-ray structures of similar resolution
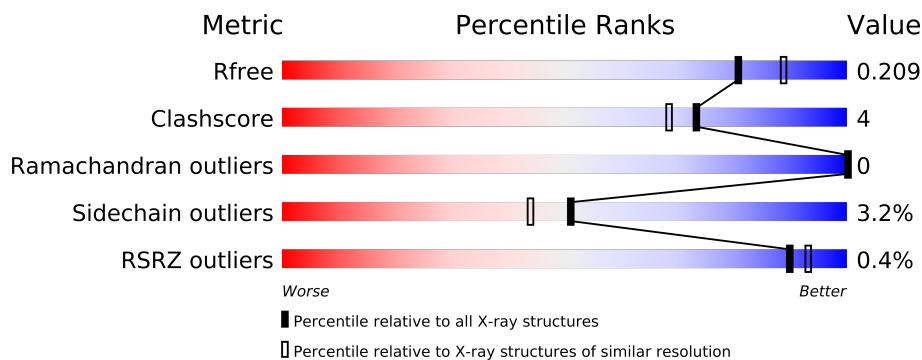
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	66092	1224 (2.08-2.04)
Clashscore	79885	1390 (2.08-2.04)
Ramachandran outliers	78287	1381 (2.08-2.04)
Sidechain outliers	78261	1381 (2.08-2.04)
RSRZ outliers	66119	1225 (2.08-2.04)

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	202	
1	B	202	
1	C	202	
1	D	202	
1	E	202	
1	F	202	
1	G	202	
1	H	202	
1	I	202	

2   Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 14457 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 Lin1944 protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	202	Total	C	N	O	Se	0	0	0
			1492	943	247	296	6			
1	B	200	Total	C	N	O	Se	0	0	0
			1478	936	244	292	6			
1	C	200	Total	C	N	O	Se	0	0	0
			1478	936	244	292	6			
1	D	201	Total	C	N	O	Se	0	0	0
			1486	940	246	294	6			
1	E	200	Total	C	N	O	Se	0	0	0
			1478	936	244	292	6			
1	F	199	Total	C	N	O	Se	0	0	0
			1473	933	243	291	6			
1	G	200	Total	C	N	O	Se	0	0	0
			1478	936	244	292	6			
1	H	202	Total	C	N	O	Se	0	0	0
			1492	943	247	296	6			
1	I	200	Total	C	N	O	Se	0	0	0
			1478	936	244	292	6			

There are 27 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	expression tag	UNP Q92AH8
A	-1	ASN	-	expression tag	UNP Q92AH8
A	0	ALA	-	expression tag	UNP Q92AH8
B	-2	SER	-	expression tag	UNP Q92AH8
B	-1	ASN	-	expression tag	UNP Q92AH8
B	0	ALA	-	expression tag	UNP Q92AH8
C	-2	SER	-	expression tag	UNP Q92AH8
C	-1	ASN	-	expression tag	UNP Q92AH8
C	0	ALA	-	expression tag	UNP Q92AH8
D	-2	SER	-	expression tag	UNP Q92AH8
D	-1	ASN	-	expression tag	UNP Q92AH8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	0	ALA	-	expression tag	UNP Q92AH8
E	-2	SER	-	expression tag	UNP Q92AH8
E	-1	ASN	-	expression tag	UNP Q92AH8
E	0	ALA	-	expression tag	UNP Q92AH8
F	-2	SER	-	expression tag	UNP Q92AH8
F	-1	ASN	-	expression tag	UNP Q92AH8
F	0	ALA	-	expression tag	UNP Q92AH8
G	-2	SER	-	expression tag	UNP Q92AH8
G	-1	ASN	-	expression tag	UNP Q92AH8
G	0	ALA	-	expression tag	UNP Q92AH8
H	-2	SER	-	expression tag	UNP Q92AH8
H	-1	ASN	-	expression tag	UNP Q92AH8
H	0	ALA	-	expression tag	UNP Q92AH8
I	-2	SER	-	expression tag	UNP Q92AH8
I	-1	ASN	-	expression tag	UNP Q92AH8
I	0	ALA	-	expression tag	UNP Q92AH8

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	146	Total 146	O 146	0	0
2	B	122	Total 122	O 122	0	0
2	C	122	Total 122	O 122	0	0
2	D	122	Total 122	O 122	0	0
2	E	117	Total 117	O 117	0	0
2	F	91	Total 91	O 91	0	0
2	G	135	Total 135	O 135	0	0
2	H	140	Total 140	O 140	0	0
2	I	129	Total 129	O 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.
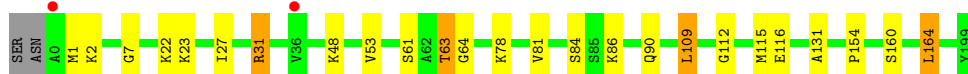
- Molecule 1: Lin1944 protein

Chain A:
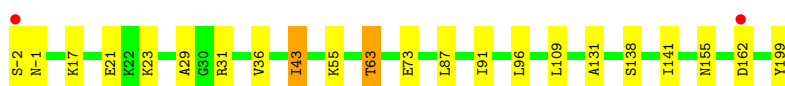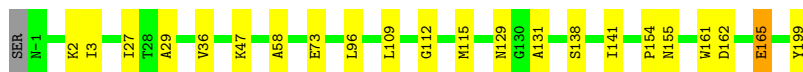

- Molecule 1: Lin1944 protein

Chain B:

- Molecule 1: Lin1944 protein

Chain C:

- Molecule 1: Lin1944 protein

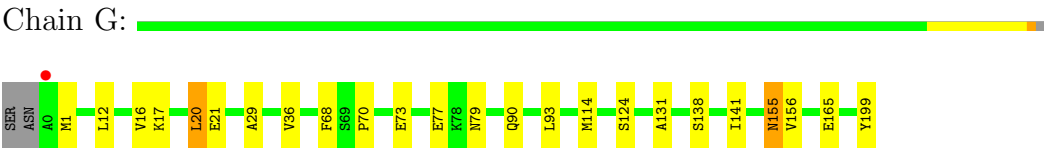Chain D:

- Molecule 1: Lin1944 protein
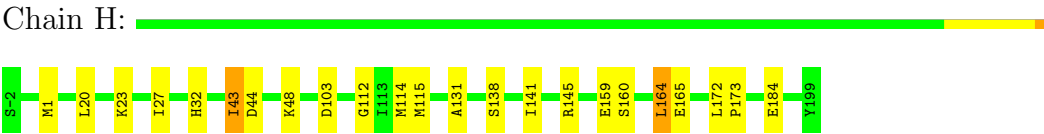
Chain E:

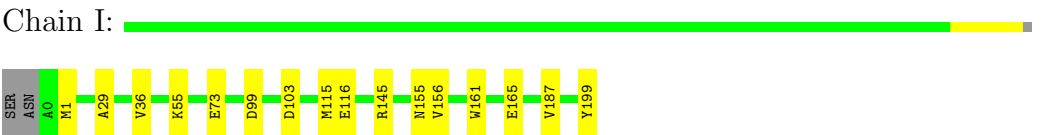- Molecule 1: Lin1944 protein

Chain F:

• Molecule 1: Lin1944 protein

Chain G:

• Molecule 1: Lin1944 protein

Chain H:

• Molecule 1: Lin1944 protein

Chain I:

4   Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	69.81Å 128.20Å 404.74Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.47 – 2.06 46.47 – 2.06	Depositor EDS
% Data completeness (in resolution range)	97.9 (46.47-2.06) 97.9 (46.47-2.06)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.73 (at 2.07Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.163 , 0.206 0.167 , 0.209	Depositor DCC
R_{free} test set	5541 reflections (5.27%)	DCC
Wilson B-factor (Å^2)	26.7	Xtriage
Anisotropy	0.229	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 46.1	EDS
Estimated twinning fraction	0.035 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.043 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
L-test for twinning	$< |L| > = 0.46, < L^2 > = 0.29$	Xtriage
Outliers	0 of 110663 reflections	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	14457	wwPDB-VP
Average B, all atoms (Å^2)	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.60% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

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	0.68	0/1507	0.71	0/2028
1	B	0.68	0/1493	0.70	0/2009
1	C	0.63	0/1493	0.66	0/2009
1	D	0.64	0/1501	0.67	0/2020
1	E	0.62	0/1493	0.66	0/2009
1	F	0.61	0/1489	0.63	0/2005
1	G	0.70	1/1493 (0.1%)	0.66	0/2009
1	H	0.69	0/1507	0.67	0/2028
1	I	0.67	0/1493	0.68	0/2009
All	All	0.66	1/13469 (0.0%)	0.67	0/18126

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	G	77	GLU	CG-CD	5.09	1.59	1.51

There are no bond angle outliers.

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	1492	0	1526	12	0

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	1478	0	1515	15	0
1	C	1478	0	1515	17	0
1	D	1486	0	1521	14	0
1	E	1478	0	1515	18	0
1	F	1473	0	1510	12	0
1	G	1478	0	1515	14	0
1	H	1492	0	1526	12	0
1	I	1478	0	1515	9	0
2	A	146	0	0	3	0
2	B	122	0	0	3	0
2	C	122	0	0	0	0
2	D	122	0	0	0	0
2	E	117	0	0	2	0
2	F	91	0	0	0	0
2	G	135	0	0	1	0
2	H	140	0	0	1	0
2	I	129	0	0	1	0
All	All	14457	0	13658	114	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 4.

The worst 5 of 114 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:2:LYS:HE2	1:B:27:ILE:HD11	1.19	1.15
1:E:143:MSE:HE3	1:F:71:LEU:CD2	1.87	1.04
1:E:143:MSE:HE3	1:F:71:LEU:HD23	1.41	1.02
1:G:155:ASN:HD22	1:G:156:VAL:H	1.12	0.96
1:B:2:LYS:HE2	1:B:27:ILE:CD1	1.95	0.95

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	200/202 (99%)	194 (97%)	6 (3%)	0	100	100
1	B	198/202 (98%)	192 (97%)	6 (3%)	0	100	100
1	C	198/202 (98%)	192 (97%)	6 (3%)	0	100	100
1	D	199/202 (98%)	191 (96%)	8 (4%)	0	100	100
1	E	198/202 (98%)	191 (96%)	7 (4%)	0	100	100
1	F	197/202 (98%)	189 (96%)	8 (4%)	0	100	100
1	G	198/202 (98%)	192 (97%)	6 (3%)	0	100	100
1	H	200/202 (99%)	195 (98%)	5 (2%)	0	100	100
1	I	198/202 (98%)	192 (97%)	6 (3%)	0	100	100
All	All	1786/1818 (98%)	1728 (97%)	58 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	161/155 (104%)	155 (96%)	6 (4%)	45	37
1	B	159/155 (103%)	150 (94%)	9 (6%)	29	18
1	C	159/155 (103%)	153 (96%)	6 (4%)	44	36
1	D	160/155 (103%)	156 (98%)	4 (2%)	60	55
1	E	159/155 (103%)	155 (98%)	4 (2%)	60	55
1	F	159/155 (103%)	155 (98%)	4 (2%)	60	55
1	G	159/155 (103%)	154 (97%)	5 (3%)	52	45
1	H	161/155 (104%)	155 (96%)	6 (4%)	45	37
1	I	159/155 (103%)	157 (99%)	2 (1%)	80	79
All	All	1436/1395 (103%)	1390 (97%)	46 (3%)	51	44

5 of 46 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	164	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	E	2	LYS
1	H	164	LEU
1	D	73	GLU
1	D	162	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 22 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	197	GLN
1	G	52	GLN
1	I	79	ASN
1	F	52	GLN
1	F	79	ASN

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 ⓘ

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, 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(Å²)	Q<0.9
1	A	202/202 (100%)	-0.33	2 (0%)	79	81	25, 29, 37, 43	0
1	B	200/202 (99%)	-0.35	2 (1%)	79	81	24, 29, 36, 46	0
1	C	200/202 (99%)	-0.46	1 (0%)	88	92	24, 29, 37, 46	0
1	D	201/202 (99%)	-0.47	0	100	100	24, 29, 36, 44	0
1	E	200/202 (99%)	-0.32	1 (0%)	88	92	23, 29, 37, 48	0
1	F	199/202 (98%)	-0.20	0	100	100	23, 30, 37, 45	0
1	G	200/202 (99%)	-0.42	1 (0%)	88	92	22, 27, 35, 43	0
1	H	202/202 (100%)	-0.48	0	100	100	21, 27, 34, 39	0
1	I	200/202 (99%)	-0.49	0	100	100	22, 27, 36, 42	0
All	All	1804/1818 (99%)	-0.39	7 (0%)	90	93	21, 28, 36, 48	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	0	ALA	3.4
1	G	0	ALA	3.2
1	A	162	ASP	2.9
1	C	0	ALA	2.5
1	A	-2	SER	2.4

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 (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.