wwPDB X-ray Structure Validation Summary Report (i)

Nov 10, 2024 – 07:42 PM EST

PDB ID	:	1L8W
Title	:	Crystal Structure of Lyme Disease Variable Surface Antigen VlsE of Borrelia burgdorferi
Authors	:	Eicken, C.; Sharma, V.; Klabunde, T.; Lawrenz, M.B.; Hardham, J.M.; Norris, S.J.; Sacchettini, J.C.
Deposited on	:	2002-03-21
Resolution	:	2.30 Å(reported)

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

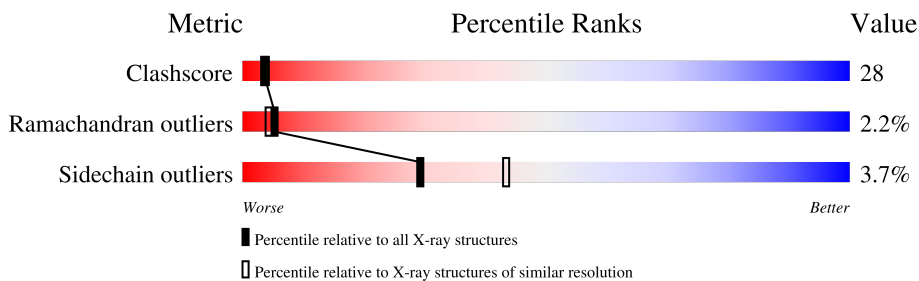|		
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

1   Overall quality at a glance ⓘ

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

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	Percentile Ranks	Value
Clashscore		28
Ramachandran outliers		2.2%
Sidechain outliers		3.7%

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))
Clashscore	180529	6698 (2.30-2.30)
Ramachandran outliers	177936	6640 (2.30-2.30)
Sidechain outliers	177891	6640 (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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	348	49% / 28% / • / 22%
1	B	348	49% / 32% / 5% / 15%
1	C	348	52% / 25% / • / 21%
1	D	348	45% / 37% / • / 14%

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 8994 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 VlsE1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	271	Total	C	N	O	Se	0	0	0
			1885	1167	327	389	2			
1	B	296	Total	C	N	O	Se	0	0	0
			2087	1302	358	425	2			
1	C	274	Total	C	N	O	Se	0	0	0
			1902	1176	331	393	2			
1	D	300	Total	C	N	O	Se	0	0	0
			2110	1315	362	431	2			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	9	MET	-	expression tag	UNP O06878
A	10	ARG	-	expression tag	UNP O06878
A	11	GLY	-	expression tag	UNP O06878
A	12	SER	-	expression tag	UNP O06878
A	13	HIS	-	expression tag	UNP O06878
A	14	HIS	-	expression tag	UNP O06878
A	15	HIS	-	expression tag	UNP O06878
A	16	HIS	-	expression tag	UNP O06878
A	17	HIS	-	expression tag	UNP O06878
A	18	HIS	-	expression tag	UNP O06878
A	19	GLY	-	expression tag	UNP O06878
A	20	SER	-	expression tag	UNP O06878
B	9	MET	-	expression tag	UNP O06878
B	10	ARG	-	expression tag	UNP O06878
B	11	GLY	-	expression tag	UNP O06878
B	12	SER	-	expression tag	UNP O06878
B	13	HIS	-	expression tag	UNP O06878
B	14	HIS	-	expression tag	UNP O06878
B	15	HIS	-	expression tag	UNP O06878
B	16	HIS	-	expression tag	UNP O06878
B	17	HIS	-	expression tag	UNP O06878

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	18	HIS	-	expression tag	UNP O06878
B	19	GLY	-	expression tag	UNP O06878
B	20	SER	-	expression tag	UNP O06878
C	9	MET	-	expression tag	UNP O06878
C	10	ARG	-	expression tag	UNP O06878
C	11	GLY	-	expression tag	UNP O06878
C	12	SER	-	expression tag	UNP O06878
C	13	HIS	-	expression tag	UNP O06878
C	14	HIS	-	expression tag	UNP O06878
C	15	HIS	-	expression tag	UNP O06878
C	16	HIS	-	expression tag	UNP O06878
C	17	HIS	-	expression tag	UNP O06878
C	18	HIS	-	expression tag	UNP O06878
C	19	GLY	-	expression tag	UNP O06878
C	20	SER	-	expression tag	UNP O06878
D	9	MET	-	expression tag	UNP O06878
D	10	ARG	-	expression tag	UNP O06878
D	11	GLY	-	expression tag	UNP O06878
D	12	SER	-	expression tag	UNP O06878
D	13	HIS	-	expression tag	UNP O06878
D	14	HIS	-	expression tag	UNP O06878
D	15	HIS	-	expression tag	UNP O06878
D	16	HIS	-	expression tag	UNP O06878
D	17	HIS	-	expression tag	UNP O06878
D	18	HIS	-	expression tag	UNP O06878
D	19	GLY	-	expression tag	UNP O06878
D	20	SER	-	expression tag	UNP O06878

- Molecule 2 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	275	Total 275	O 275	0	0
2	B	250	Total 250	O 250	0	0
2	C	256	Total 256	O 256	0	0
2	D	229	Total 229	O 229	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.

Note EDS was not executed.
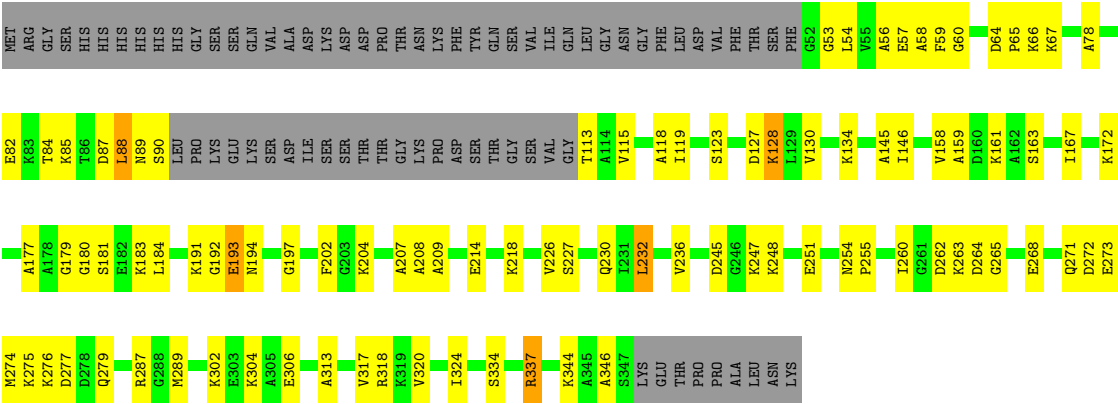
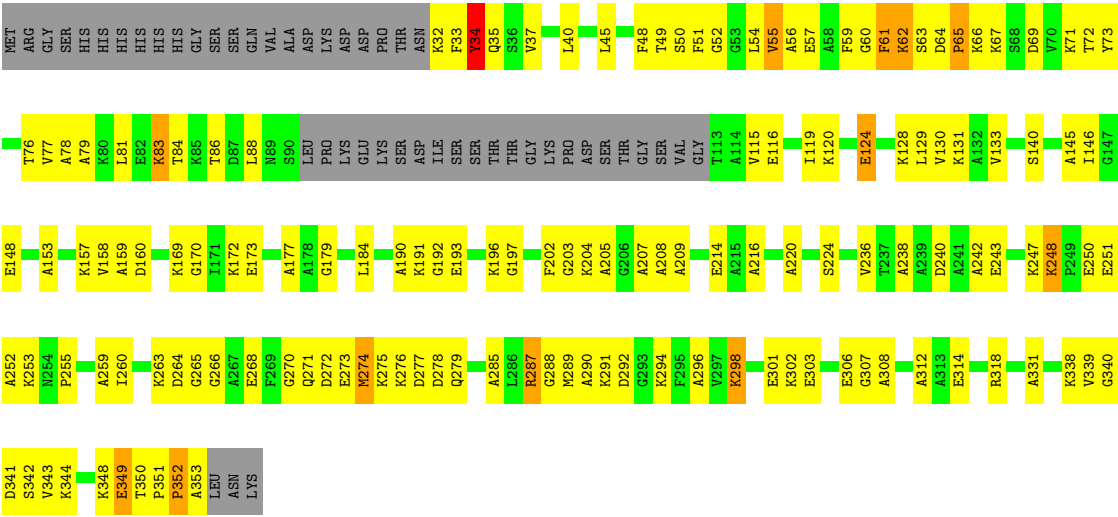- Molecule 1: VlsE1

Chain A:

- Molecule 1: VlsE1

Chain B:

- Molecule 1: VlsE1

Chain C:

• Molecule 1: VlsE1

Chain D:

45%   37%   •   14%
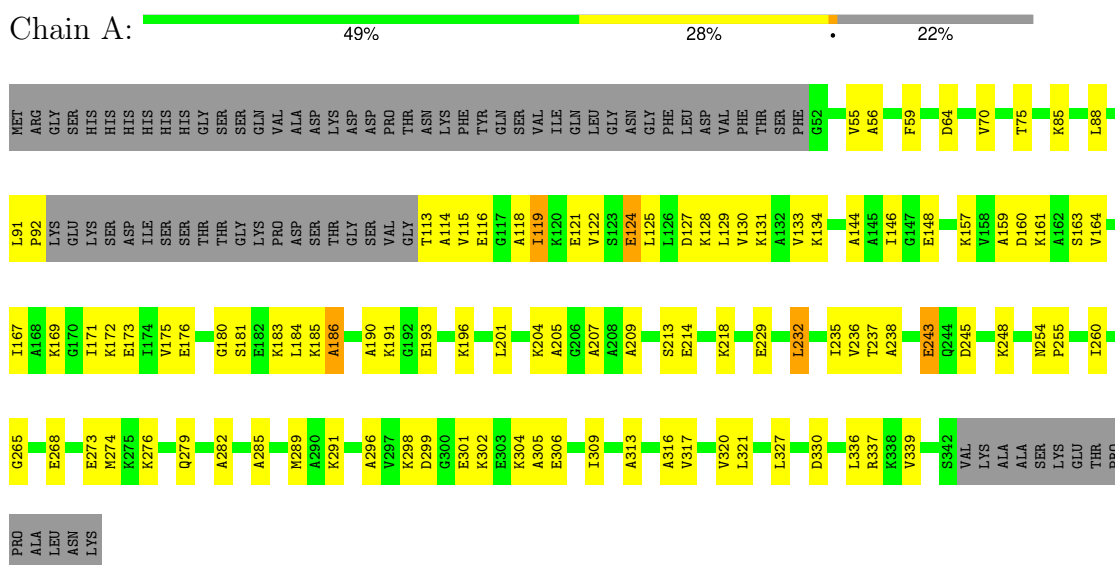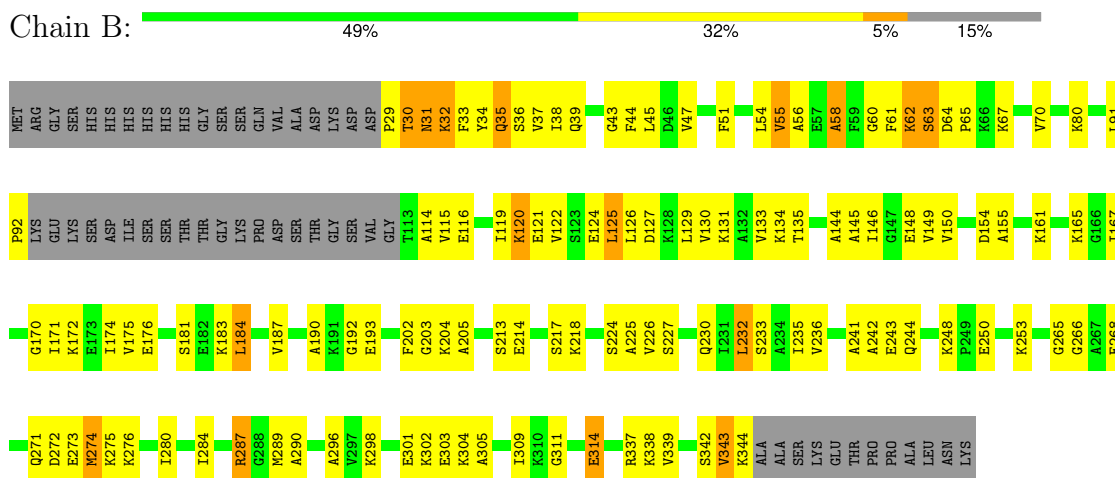
4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	85.17Å 59.18Å 116.15Å 90.00° 104.58° 90.00°	Depositor
Resolution (Å)	82.43 – 2.30	Depositor
% Data completeness (in resolution range)	96.7 (82.43-2.30)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.216 , 0.291	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	8994	wwPDB-VP
Average B, all atoms (Å^2)	41.0	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	0.34	0/1893	0.55	0/2529
1	B	0.33	0/2101	0.54	0/2810
1	C	0.32	0/1909	0.55	0/2549
1	D	0.30	0/2124	0.52	0/2841
All	All	0.32	0/8027	0.54	0/10729

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	1885	0	1942	107	0
1	B	2087	0	2143	146	0
1	C	1902	0	1961	82	0
1	D	2110	0	2164	140	0
2	A	275	0	0	17	0
2	B	250	0	0	21	0
2	C	256	0	0	16	0
2	D	229	0	0	20	1
All	All	8994	0	8210	451	1

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 28.

The worst 5 of 451 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:B:35:GLN:HA	1:B:38:ILE:HD13	1.35	1.09
1:A:113:THR:HG23	1:A:114:ALA:H	1.35	0.91
1:A:175:VAL:HG12	1:A:180:GLY:HA2	1.57	0.87
1:A:285:ALA:HB3	2:A:1001:HOH:O	1.74	0.86
1:D:302:LYS:O	1:D:306:GLU:HG3	1.76	0.85

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:D:2516:HOH:O	2:D:2516:HOH:O[2_656]	2.09	0.11

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	267/348 (77%)	244 (91%)	21 (8%)	2 (1%)	19	23
1	B	292/348 (84%)	263 (90%)	19 (6%)	10 (3%)	3	2
1	C	270/348 (78%)	252 (93%)	15 (6%)	3 (1%)	12	13
1	D	296/348 (85%)	254 (86%)	32 (11%)	10 (3%)	3	2
All	All	1125/1392 (81%)	1013 (90%)	87 (8%)	25 (2%)	5	4

5 of 25 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	186	ALA
1	A	207	ALA
1	B	30	THR
1	D	34	TYR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	62	LYS

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	179/243 (74%)	172 (96%)	7 (4%)	27	41
1	B	202/243 (83%)	194 (96%)	8 (4%)	27	40
1	C	180/243 (74%)	174 (97%)	6 (3%)	33	48
1	D	203/243 (84%)	196 (97%)	7 (3%)	32	47
All	All	764/972 (79%)	736 (96%)	28 (4%)	29	43

5 of 28 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	314	GLU
1	D	298	LYS
1	C	193	GLU
1	D	240	ASP
1	C	128	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	210	HIS
1	D	271	GLN
1	B	210	HIS
1	B	244	GLN
1	C	210	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 ⓘ

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3   Carbohydrates ⓘ

EDS was not executed - this section is therefore empty.

6.4   Ligands ⓘ

EDS was not executed - this section is therefore empty.

6.5   Other polymers ⓘ

EDS was not executed - this section is therefore empty.