Full wwPDB X-ray Structure Validation Report (i)

Feb 3, 2024 – 02:44 PM EST

PDB ID	:	1LEM
Title	:	THE MONOSACCHARIDE BINDING SITE OF LENTIL LECTIN: AN X-RAY AND MOLECULAR MODELLING STUDY
Authors	:	Loris, R.; Wyns, L.
Deposited on	:	1993-11-17
Resolution	:	3.00 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

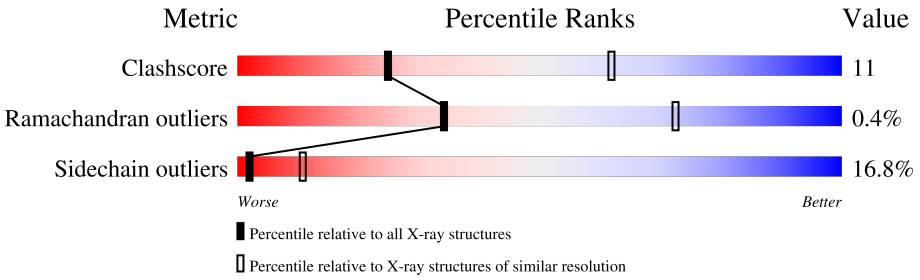
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 3.00 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	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)

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	181	64% 27% 8% ·
2	B	52	54% 33% · 10%

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 1833 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.

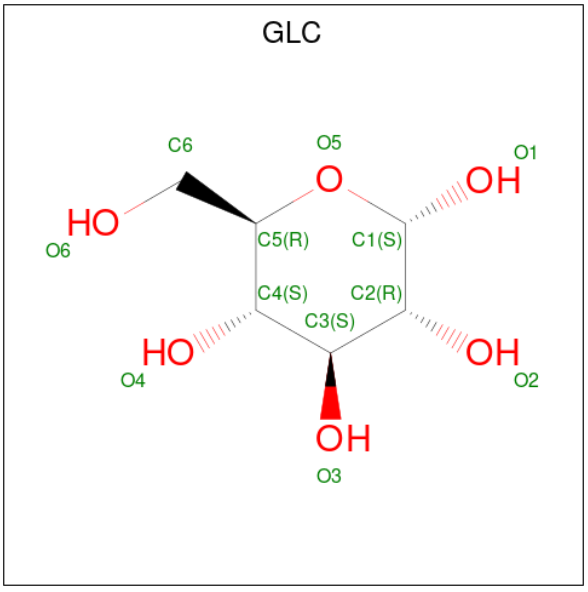
- Molecule 1 is a protein called LECTIN.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
1	A	181	1410	898	232	280	0	0	0

- Molecule 2 is a protein called LECTIN.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	B	47	366	236	59	71	0	0	0

- Molecule 3 is alpha-D-glucopyranose (three-letter code: GLC) (formula: $C_6H_{12}O_6$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
3	A	1	12	6	6	0	0

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Ca 1 1	0	0

- Molecule 5 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total Mn 1 1	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	38	Total O 38 38	0	0
6	B	5	Total O 5 5	0	0

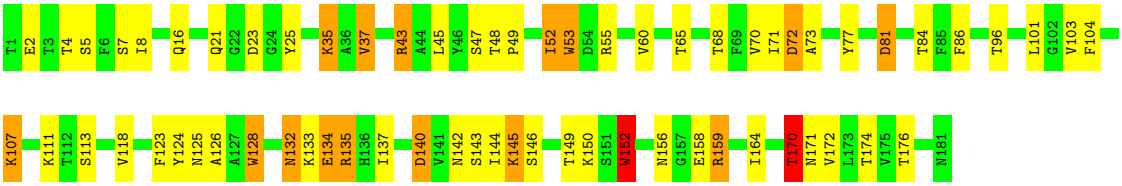
3 Residue-property plots (i)

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.

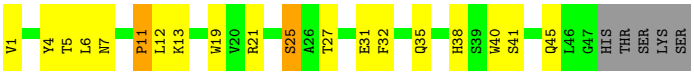
- Molecule 1: LECTIN

Chain A:



- Molecule 2: LECTIN

Chain B:



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, α, β, γ	85.71Å 85.71Å 165.38Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	(Not available) – 3.00	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-3.00)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	X-PLOR	Depositor
R, R_{free}	0.206 , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	1833	wwPDB-VP
Average B, all atoms (Å^2)	38.0	wwPDB-VP

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: MN, CA, GLC

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	1.11	2/1444 (0.1%)	1.98	46/1970 (2.3%)
2	B	1.11	0/376	1.91	9/515 (1.7%)
All	All	1.11	2/1820 (0.1%)	1.97	55/2485 (2.2%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	55	ARG	NE-CZ	7.30	1.42	1.33
1	A	53	TRP	CG-CD2	-6.36	1.32	1.43

All (55) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	55	ARG	NE-CZ-NH1	17.80	129.20	120.30
1	A	55	ARG	NH1-CZ-NH2	-12.16	106.02	119.40
2	B	21	ARG	NE-CZ-NH1	10.50	125.55	120.30
1	A	152	TRP	CD1-CG-CD2	8.89	113.41	106.30
1	A	152	TRP	CG-CD2-CE3	8.77	141.79	133.90
1	A	152	TRP	CE2-CD2-CG	-7.97	100.92	107.30
2	B	19	TRP	CE2-CD2-CG	-7.75	101.10	107.30
2	B	19	TRP	CG-CD2-CE3	7.59	140.74	133.90
1	A	170	THR	N-CA-CB	-7.03	96.95	110.30
1	A	159	ARG	NE-CZ-NH2	-7.02	116.79	120.30
2	B	21	ARG	NE-CZ-NH2	-7.00	116.80	120.30
1	A	135	ARG	NE-CZ-NH2	-6.97	116.81	120.30
1	A	53	TRP	CE2-CD2-CE3	6.96	127.05	118.70
1	A	55	ARG	NE-CZ-NH2	6.95	123.78	120.30
1	A	77	TYR	CB-CG-CD2	-6.94	116.83	121.00
1	A	43	ARG	NE-CZ-NH1	6.91	123.75	120.30
1	A	152	TRP	CB-CG-CD1	-6.89	118.04	127.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	53	TRP	CD2-CE2-CZ2	-6.58	114.40	122.30
2	B	19	TRP	CD1-CG-CD2	6.44	111.45	106.30
1	A	53	TRP	CE3-CZ3-CH2	-6.44	114.12	121.20
1	A	53	TRP	CD1-CG-CD2	6.28	111.33	106.30
1	A	72	ASP	CA-C-N	-6.23	103.50	117.20
1	A	16	GLN	CA-CB-CG	-6.20	99.75	113.40
2	B	19	TRP	CB-CG-CD1	-6.18	118.97	127.00
1	A	55	ARG	CD-NE-CZ	6.12	132.17	123.60
2	B	19	TRP	NE1-CE2-CZ2	-6.10	123.69	130.40
2	B	40	TRP	CG-CD2-CE3	6.05	139.34	133.90
1	A	128	TRP	CD1-CG-CD2	6.02	111.11	106.30
1	A	128	TRP	CE2-CD2-CG	-6.01	102.49	107.30
1	A	170	THR	CA-CB-CG2	5.99	120.79	112.40
1	A	2	GLU	CA-CB-CG	5.98	126.55	113.40
1	A	125	ASN	N-CA-C	-5.77	95.41	111.00
1	A	8	ILE	CA-C-N	5.77	129.90	117.20
1	A	133	LYS	CA-CB-CG	-5.77	100.70	113.40
1	A	4	THR	CA-CB-CG2	5.75	120.45	112.40
1	A	37	VAL	CB-CA-C	-5.73	100.51	111.40
2	B	40	TRP	CE2-CD2-CG	-5.72	102.72	107.30
1	A	60	VAL	CG1-CB-CG2	-5.69	101.80	110.90
1	A	8	ILE	O-C-N	-5.67	113.62	122.70
1	A	149	THR	CA-CB-CG2	5.64	120.30	112.40
1	A	152	TRP	CG-CD1-NE1	-5.60	104.50	110.10
1	A	140	ASP	CB-CA-C	-5.57	99.26	110.40
1	A	55	ARG	CB-CG-CD	5.53	125.98	111.60
1	A	96	THR	CA-C-N	-5.42	105.36	116.20
1	A	25	TYR	CB-CG-CD2	-5.42	117.75	121.00
1	A	53	TRP	CZ3-CH2-CZ2	5.40	128.08	121.60
1	A	53	TRP	CE2-CD2-CG	-5.38	103.00	107.30
1	A	159	ARG	NE-CZ-NH1	5.30	122.95	120.30
1	A	135	ARG	CA-CB-CG	-5.30	101.74	113.40
1	A	4	THR	CA-CB-OG1	-5.22	98.04	109.00
1	A	43	ARG	NE-CZ-NH2	-5.19	117.70	120.30
1	A	47	SER	CA-C-N	5.18	128.60	117.20
1	A	123	PHE	CB-CG-CD1	-5.17	117.18	120.80
1	A	35	LYS	CA-C-N	5.07	128.35	117.20
1	A	128	TRP	CG-CD2-CE3	5.01	138.41	133.90

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	1410	0	1363	38	0
2	B	366	0	350	9	0
3	A	12	0	12	0	0
4	A	1	0	0	0	0
5	A	1	0	0	0	0
6	A	38	0	0	3	0
6	B	5	0	0	0	0
All	All	1833	0	1725	38	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.

All (38) 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:A:73:ALA:H	1:A:156:ASN:HD21	1.14	0.95
1:A:73:ALA:H	1:A:156:ASN:ND2	1.84	0.75
1:A:52:ILE:HG23	1:A:53:TRP:HD1	1.55	0.72
1:A:81:ASP:HB2	2:B:27:THR:O	1.92	0.70
1:A:128:TRP:HB2	1:A:145:LYS:HG2	1.74	0.69
1:A:68:THR:HG22	2:B:38:HIS:HB2	1.77	0.67
1:A:72:ASP:HA	1:A:156:ASN:HD21	1.61	0.66
1:A:21:GLN:HB2	1:A:43:ARG:HB2	1.78	0.64
1:A:152:TRP:HA	2:B:4:TYR:CE2	2.35	0.62
1:A:176:THR:HG23	6:A:292:HOH:O	2.02	0.59
1:A:113:SER:HB3	1:A:142:ASN:ND2	2.18	0.58
1:A:170:THR:HG23	1:A:172:VAL:HG23	1.86	0.57
1:A:137:ILE:HG22	2:B:6:LEU:HD23	1.86	0.57
1:A:21:GLN:HE22	1:A:43:ARG:HH21	1.53	0.57
1:A:72:ASP:HA	1:A:156:ASN:ND2	2.21	0.56
1:A:103:VAL:HG23	1:A:104:PHE:CD2	2.41	0.56
1:A:70:VAL:CG1	1:A:159:ARG:HG2	2.36	0.55
1:A:21:GLN:NE2	1:A:43:ARG:HE	2.08	0.52
1:A:65:THR:HG22	1:A:164:ILE:HB	1.91	0.52
1:A:170:THR:HG23	1:A:172:VAL:CG2	2.40	0.51

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:124:TYR:CE2	1:A:126:ALA:HA	2.47	0.50
1:A:73:ALA:N	1:A:156:ASN:HD21	1.96	0.50
1:A:84:THR:OG1	2:B:25:SER:HB3	2.13	0.48
1:A:124:TYR:CZ	1:A:126:ALA:HA	2.50	0.47
1:A:135:ARG:HB2	6:A:287:HOH:O	2.14	0.47
1:A:101:LEU:HA	2:B:27:THR:OG1	2.14	0.46
1:A:174:THR:HG23	6:A:276:HOH:O	2.15	0.46
1:A:52:ILE:HG23	1:A:53:TRP:CD1	2.43	0.46
1:A:132:ASN:CG	1:A:134:GLU:HG3	2.36	0.45
1:A:71:ILE:HG12	2:B:35:GLN:HG3	1.97	0.45
1:A:70:VAL:HG12	1:A:159:ARG:HA	2.01	0.43
1:A:107:LYS:HA	1:A:144:ILE:HG22	2.01	0.43
1:A:144:ILE:O	1:A:144:ILE:HG13	2.19	0.42
1:A:68:THR:CG2	2:B:38:HIS:HB2	2.48	0.42
1:A:176:THR:HG22	2:B:5:THR:HG23	2.00	0.42
1:A:84:THR:HG22	1:A:118:VAL:O	2.20	0.41
1:A:86:PHE:CD1	1:A:86:PHE:N	2.89	0.41
1:A:70:VAL:HG12	1:A:159:ARG:HG2	2.02	0.40

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	179/181 (99%)	168 (94%)	11 (6%)	0	100 100
2	B	45/52 (86%)	43 (96%)	1 (2%)	1 (2%)	6 31
All	All	224/233 (96%)	211 (94%)	12 (5%)	1 (0%)	34 72

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	11	PRO

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	156/156 (100%)	133 (85%)	23 (15%)	3 15
2	B	40/45 (89%)	30 (75%)	10 (25%)	0 3
All	All	196/201 (98%)	163 (83%)	33 (17%)	2 11

All (33) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	5	SER
1	A	7	SER
1	A	23	ASP
1	A	35	LYS
1	A	37	VAL
1	A	45	LEU
1	A	48	THR
1	A	49	PRO
1	A	52	ILE
1	A	81	ASP
1	A	107	LYS
1	A	111	LYS
1	A	132	ASN
1	A	134	GLU
1	A	140	ASP
1	A	143	SER
1	A	145	LYS
1	A	146	SER
1	A	150	LYS
1	A	152	TRP
1	A	158	GLU
1	A	170	THR
1	A	171	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	1	VAL
2	B	7	ASN
2	B	11	PRO
2	B	12	LEU
2	B	13	LYS
2	B	25	SER
2	B	31	GLU
2	B	32	PHE
2	B	41	SER
2	B	45	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	21	GLN
1	A	59	ASN
1	A	105	ASN
1	A	142	ASN
1	A	155	GLN
1	A	156	ASN

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 3 ligands modelled in this entry, 2 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	$\#\|Z\| > 2$
3	GLC	A	248	-	12,12,12	1.49	3 (25%)	17,17,17	1.45	1 (5%)

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
3	GLC	A	248	-	-	2/2/22/22	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	248	GLC	C6-C5	2.83	1.61	1.51
3	A	248	GLC	C3-C2	2.46	1.58	1.52
3	A	248	GLC	O1-C1	2.21	1.46	1.39

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	248	GLC	C1-O5-C5	4.67	122.48	113.66

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	248	GLC	O5-C5-C6-O6
3	A	248	GLC	C4-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

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

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.