wwPDB X-ray Structure Validation Summary Report (i)

May 14, 2020 – 01:18 pm BST

PDB ID	:	1DLP
Title	:	STRUCTURAL CHARACTERIZATION OF THE NATIVE FETUIN-BINDING PROTEIN SCILLA CAMPANULATA AGGLUTININ (SCAFET): A NOVEL TWO-DOMAIN LECTIN
Authors	:	Wright, L.M.; Reynolds, C.D.; Rizkallah, P.J.; Allen, A.K.; VanDamme, E.J.M.; Donovan, M.J.; Peumans, W.J.
Deposited on	:	1999-12-11
Resolution	:	3.30 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 following versions of software and data (see references (1)) were used in the production of this report:

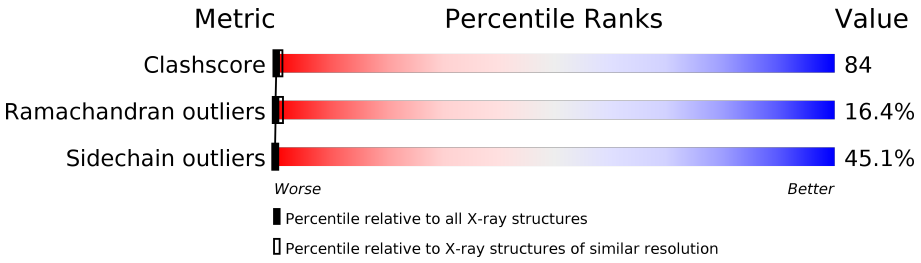
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 3.30 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))
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)

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 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	236	13% 35% 51% .
1	B	236	. 11% 34% 48% 6%
1	C	236	. 17% 32% 48% .
1	D	236	11% 35% 49% 6%
1	E	236	. 14% 39% 43% .
1	F	236	. 14% 37% 40% 8%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 10431 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 SCAFET PRECURSOR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	233	Total	C	N	O	S	18	0	0
			1778	1102	322	347	7			
1	B	221	Total	C	N	O	S	21	0	0
			1694	1053	305	329	7			
1	C	234	Total	C	N	O	S	25	0	0
			1772	1097	323	345	7			
1	D	223	Total	C	N	O	S	28	0	0
			1702	1057	307	331	7			
1	E	231	Total	C	N	O	S	17	0	0
			1759	1093	318	341	7			
1	F	216	Total	C	N	O	S	40	0	0
			1659	1034	298	320	7			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	19	Total	O	0	0
			19	19		
2	B	15	Total	O	0	0
			15	15		
2	C	6	Total	O	0	0
			6	6		
2	D	8	Total	O	0	0
			8	8		
2	E	8	Total	O	0	0
			8	8		
2	F	11	Total	O	0	0
			11	11		

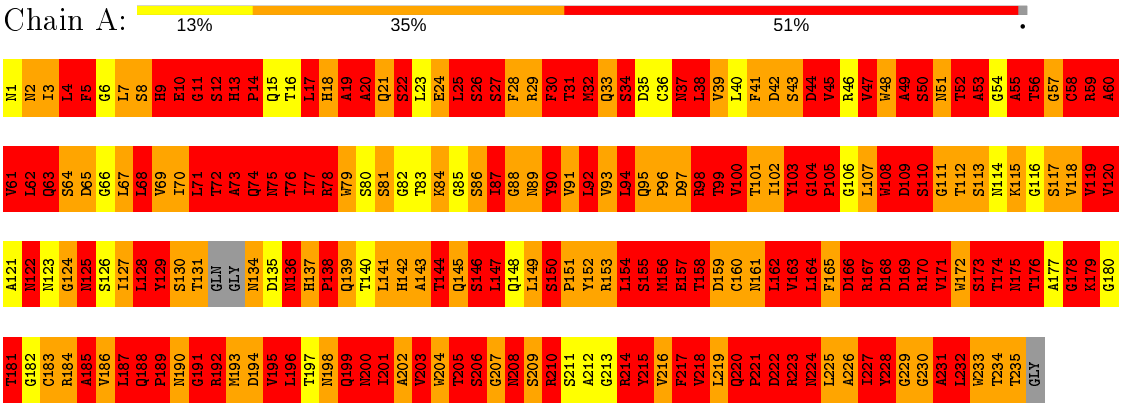
3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA 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.

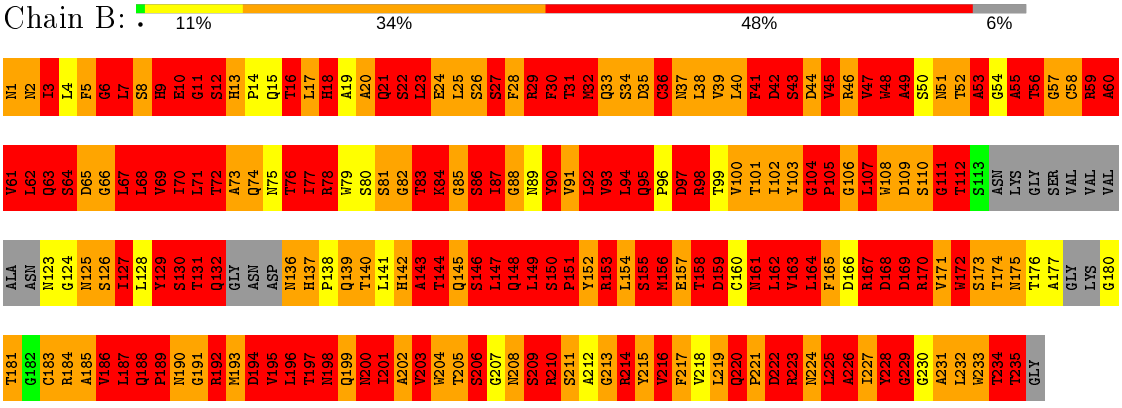
- Molecule 1: LECTIN SCAFET PRECURSOR

Chain A:



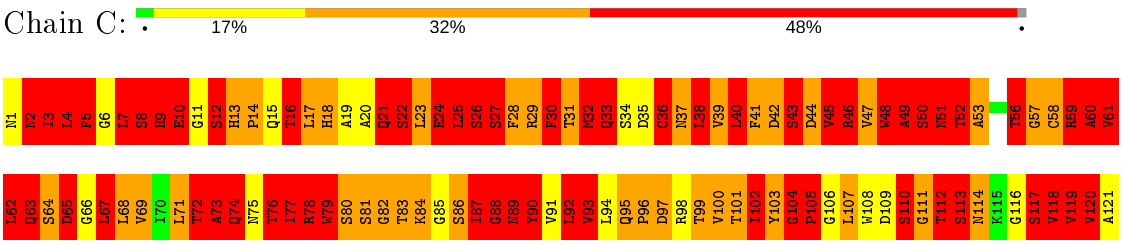
- Molecule 1: LECTIN SCAFET PRECURSOR

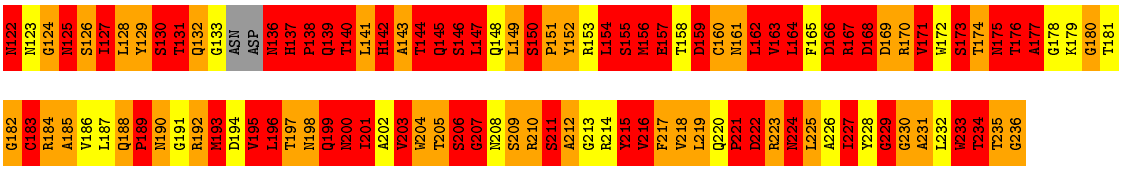
Chain B:



- Molecule 1: LECTIN SCAFET PRECURSOR

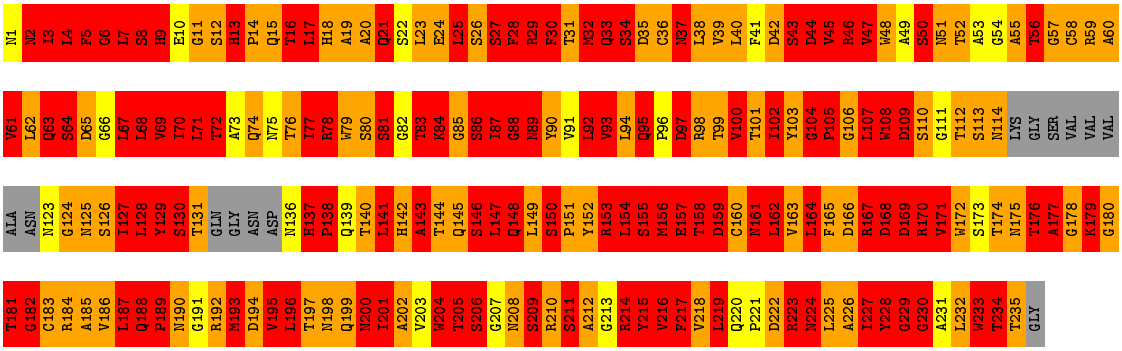
Chain C:





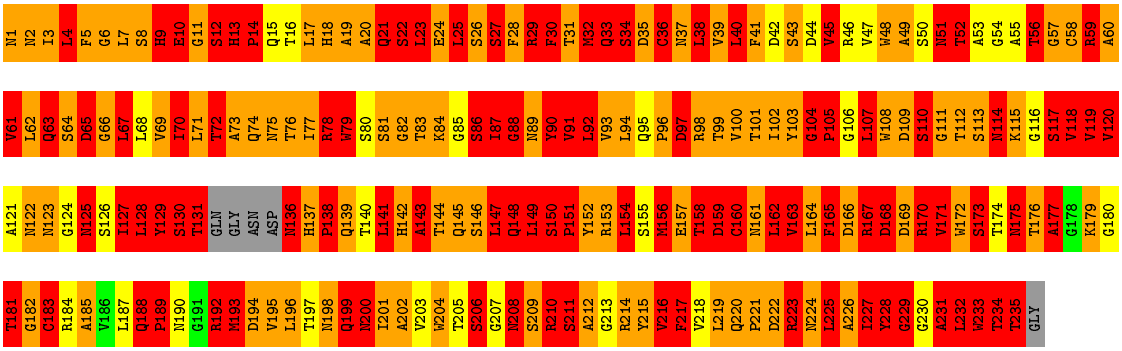
• Molecule 1: LECTIN SCAFET PRECURSOR

Chain D:



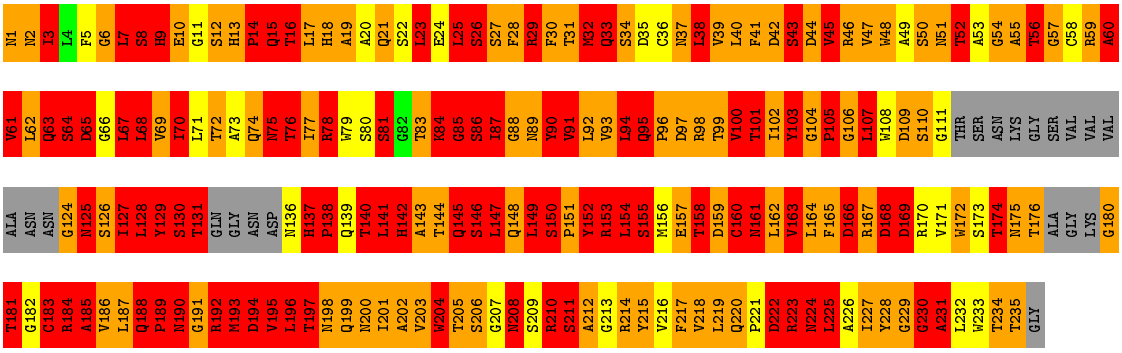
• Molecule 1: LECTIN SCAFET PRECURSOR

Chain E:



• Molecule 1: LECTIN SCAFET PRECURSOR

Chain F:



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	277.94Å 164.10Å 53.60Å 90.00° 95.38° 90.00°	Depositor
Resolution (Å)	20.00 – 3.30	Depositor
% Data completeness (in resolution range)	95.3 (20.00-3.30)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	REFMAC	Depositor
R, R_{free}	0.189 , 0.293	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	10431	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

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	2.62	91/1812 (5.0%)	6.89	912/2470 (36.9%)
1	B	2.44	80/1726 (4.6%)	6.84	897/2351 (38.2%)
1	C	2.34	58/1805 (3.2%)	6.48	838/2459 (34.1%)
1	D	3.14	88/1735 (5.1%)	7.13	928/2365 (39.2%)
1	E	2.06	35/1793 (2.0%)	5.99	765/2444 (31.3%)
1	F	2.14	38/1691 (2.2%)	6.60	810/2303 (35.2%)
All	All	2.48	390/10562 (3.7%)	6.66	5150/14392 (35.8%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	2	60
1	B	2	73
1	C	0	66
1	D	0	51
1	E	0	39
1	F	2	53
All	All	6	342

The worst 5 of 390 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	1	ASN	CG-ND2	50.13	2.58	1.32
1	D	78	ARG	CG-CD	46.82	2.69	1.51
1	A	44	ASP	CB-CG	-34.10	0.80	1.51
1	C	168	ASP	CG-OD1	30.06	1.94	1.25
1	D	167	ARG	CD-NE	-29.28	0.96	1.46

The worst 5 of 5150 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	214	ARG	NE-CZ-NH2	59.84	150.22	120.30
1	E	59	ARG	NE-CZ-NH2	-55.14	92.73	120.30
1	B	184	ARG	NE-CZ-NH2	-53.22	93.69	120.30
1	F	223	ARG	NE-CZ-NH2	-49.78	95.41	120.30
1	A	184	ARG	NE-CZ-NH1	48.82	144.71	120.30

5 of 6 chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	9	HIS	CA
1	A	87	ILE	CA
1	B	60	ALA	CA
1	B	168	ASP	CA
1	F	60	ALA	CA

5 of 342 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	13	HIS	Mainchain
1	A	19	ALA	Mainchain
1	A	2	ASN	Peptide
1	A	25	LEU	Mainchain
1	A	5	PHE	Mainchain

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	1778	0	1700	282	0
1	B	1694	0	1600	265	0
1	C	1772	0	1695	270	0
1	D	1702	0	1613	321	0
1	E	1759	0	1684	310	0
1	F	1659	0	1579	305	0
2	A	19	0	0	1	0
2	B	15	0	0	0	0
2	C	6	0	0	0	0
2	D	8	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	E	8	0	0	0	0
2	F	11	0	0	0	0
All	All	10431	0	9871	1677	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 84.

The worst 5 of 1677 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:D:77:ILE:CD1	1:D:77:ILE:CG1	1.80	1.57
1:E:131:THR:CA	1:E:131:THR:C	1.75	1.50
1:E:136:ASN:CA	1:E:136:ASN:CB	1.88	1.49
1:A:131:THR:CB	1:A:131:THR:OG1	1.64	1.46
1:D:64:SER:OG	1:D:64:SER:CB	1.67	1.43

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	229/236 (97%)	158 (69%)	41 (18%)	30 (13%)	0	1
1	B	213/236 (90%)	132 (62%)	42 (20%)	39 (18%)	0	1
1	C	230/236 (98%)	150 (65%)	42 (18%)	38 (16%)	0	1
1	D	217/236 (92%)	148 (68%)	30 (14%)	39 (18%)	0	1
1	E	227/236 (96%)	150 (66%)	39 (17%)	38 (17%)	0	1
1	F	208/236 (88%)	138 (66%)	37 (18%)	33 (16%)	0	1
All	All	1324/1416 (94%)	876 (66%)	231 (17%)	217 (16%)	0	1

5 of 217 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	10	GLU
1	A	11	GLY
1	A	14	PRO
1	A	19	ALA
1	A	20	ALA

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	197/198 (100%)	99 (50%)	98 (50%)	0	0
1	B	187/198 (94%)	96 (51%)	91 (49%)	0	0
1	C	195/198 (98%)	111 (57%)	84 (43%)	0	0
1	D	187/198 (94%)	101 (54%)	86 (46%)	0	0
1	E	194/198 (98%)	112 (58%)	82 (42%)	0	0
1	F	183/198 (92%)	108 (59%)	75 (41%)	0	0
All	All	1143/1188 (96%)	627 (55%)	516 (45%)	0	0

5 of 516 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	137	HIS
1	D	48	TRP
1	F	127	ILE
1	C	150	SER
1	C	203	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 65 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	125	ASN
1	D	18	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	F	161	ASN
1	C	136	ASN
1	C	161	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 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

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	B	3
1	A	1
1	C	1
1	E	1

The worst 5 of 6 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	25:LEU	C	26:SER	N	1.19

Continued on next page...

Continued from previous page...

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	8:SER	C	9:HIS	N	1.19
1	C	229:GLY	C	230:GLY	N	1.19
1	E	85:GLY	C	86:SER	N	1.19
1	B	30:PHE	C	31:THR	N	1.16

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.