wwPDB X-ray Structure Validation Summary Report (i)

Sep 27, 2023 – 05:54 PM EDT

PDB ID	:	1M1C
Title	:	Structure of the L-A virus
Authors	:	Naitow, H.; Tang, J.; Canady, M.; Wickner, R.B.; Johnson, J.E.
Deposited on	:	2002-06-18
Resolution	:	3.50 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 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:

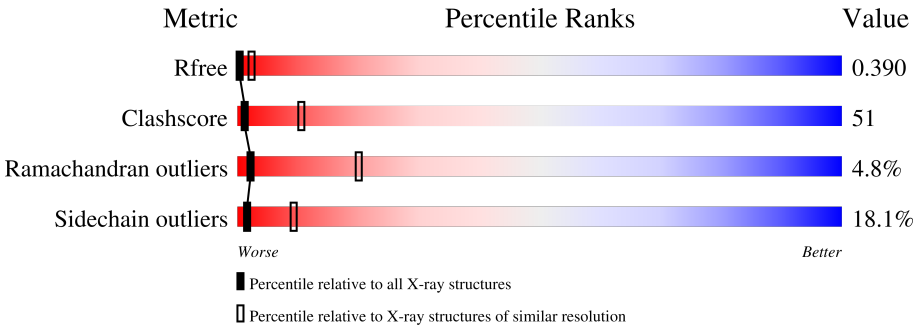
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 3.50 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}	130704	1659 (3.60-3.40)
Clashscore	141614	1036 (3.58-3.42)
Ramachandran outliers	138981	1005 (3.58-3.42)
Sidechain outliers	138945	1006 (3.58-3.42)

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%

Mol	Chain	Length	Quality of chain
1	A	680	34% 52% 10% .
1	B	680	26% 53% 15% . .

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 10302 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 Major coat protein.

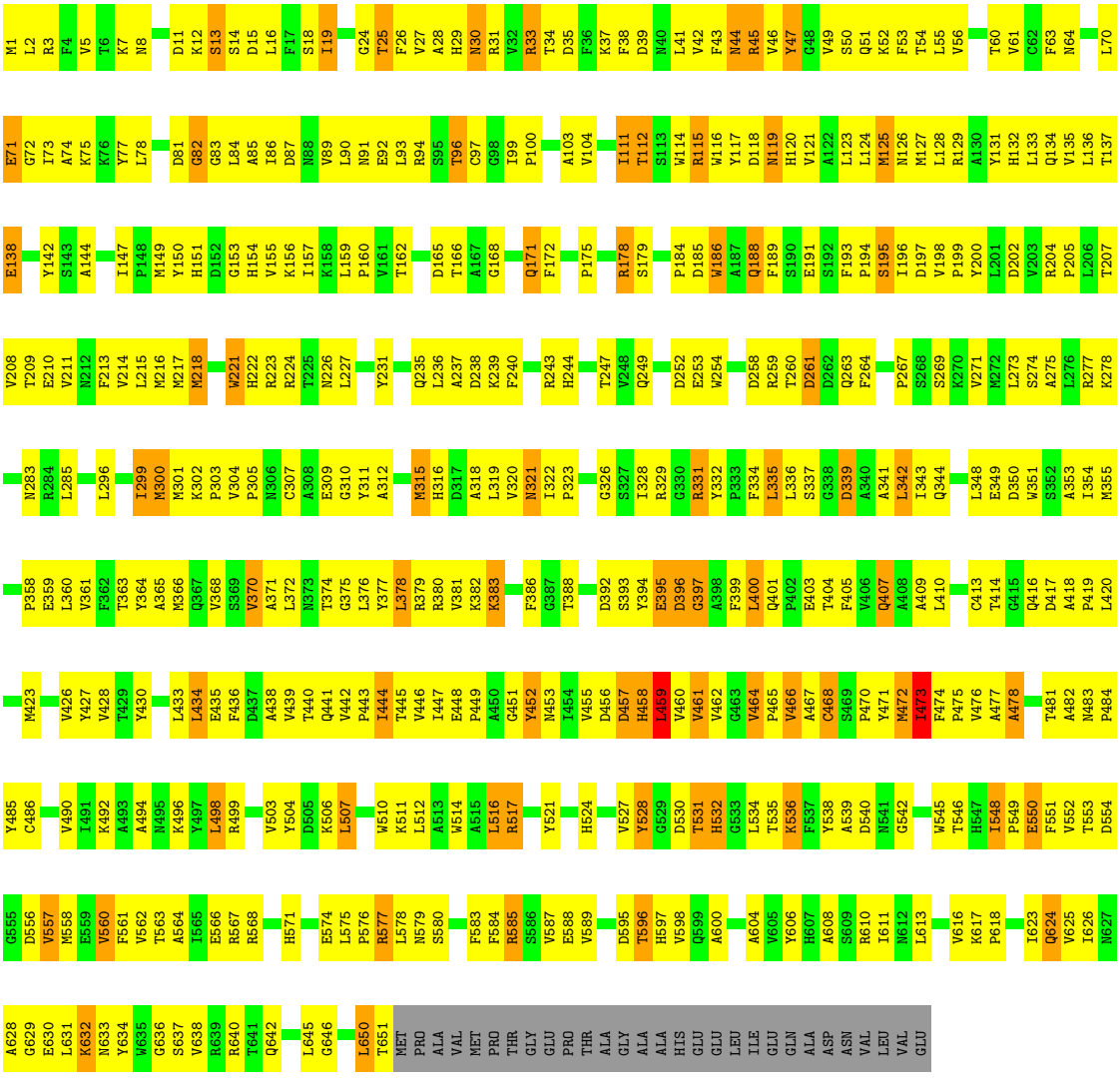
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	651	Total	C	N	O	S	0	0	0
			5151	3302	871	955	23			
1	B	651	Total	C	N	O	S	0	0	0
			5151	3302	871	955	23			

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.

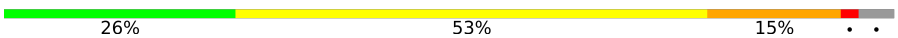
- Molecule 1: Major coat protein

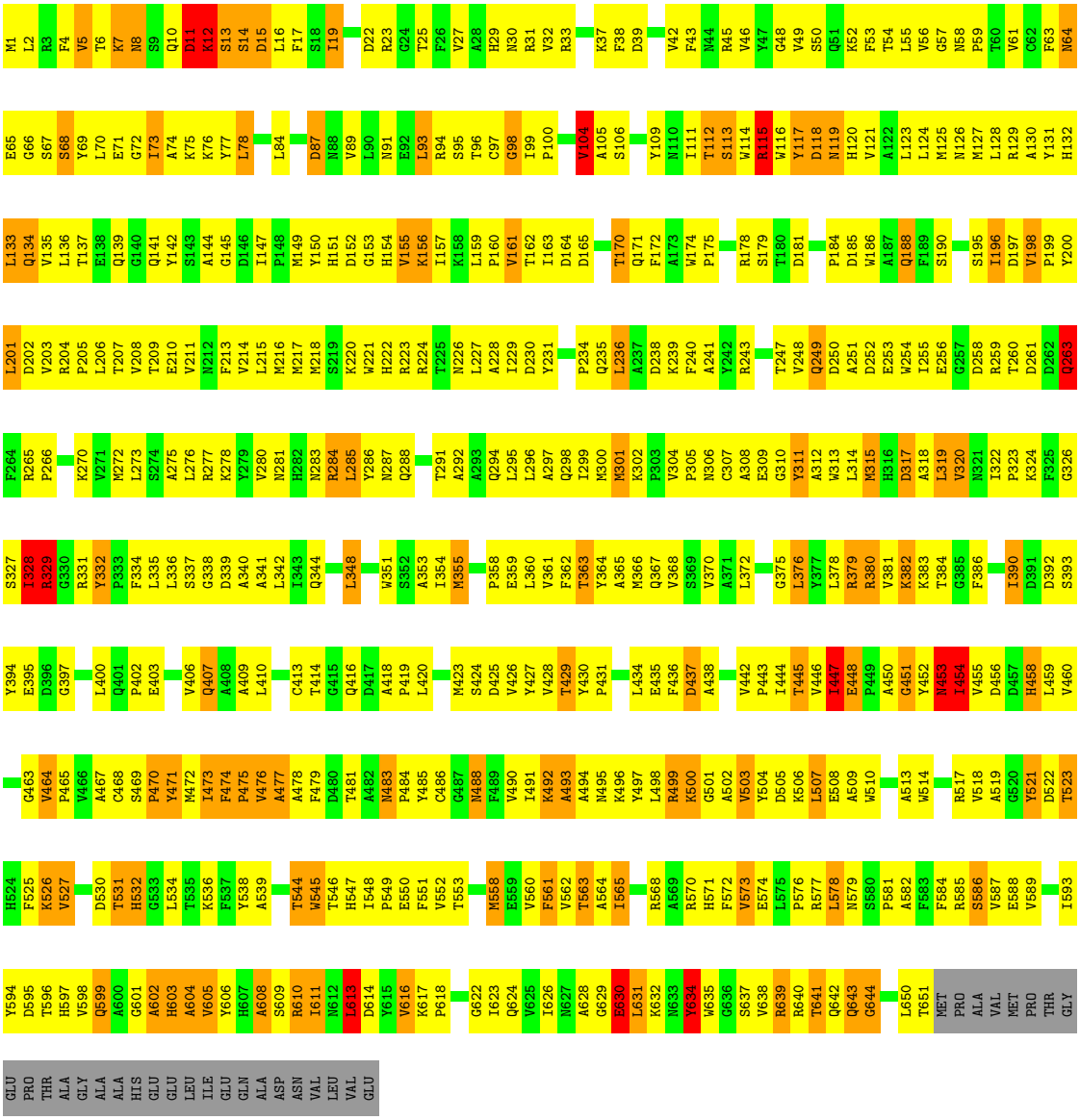
Chain A:



- Molecule 1: Major coat protein

Chain B:




M1 L2 R3 F4 V5 T6 K7 M8 S9 Q10 D11 K12 S13 S14 D15 L16 F17 S18 I19 D22 R23 G24 T25 F26 V27 A28 H29 N30 R31 V32 R33 K37 F38 D39 V42 F43 N44 R45 V46 Y47 G48 V49 S50 Q51 K52 F53 T54 L55 V56 G57 N58 P59 T60 V61 C62 F63 N64
E65 G66 S67 S68 Y69 L70 E71 G72 I73 A74 K75 K76 Y77 L78 L84 D87 N88 V89 L90 N91 E92 L93 R94 S95 T96 C97 G98 I99 P100 V104 A105 S106 Y109 N110 I111 T112 S113 W114 R115 W116 Y117 D118 N119 H120 V121 A122 L123 L124 M125 N126 M127 L128 R129 A130 Y131 H132
L133 Q134 V135 L136 T137 E138 Q139 G140 Q141 Y142 S143 A144 G145 D146 I147 P148 M149 Y150 H151 D152 G153 H154 V155 K156 I157 K158 L159 P160 V161 T162 I163 D164 D165 T170 Q171 F172 A173 W174 P175 R178 S179 T180 D181 P184 D185 W186 A187 Q188 F189 S190 S195 I196 D197 V198 P199 Y200
L201 D202 V203 R204 P205 L206 T207 V208 T209 E210 V211 N212 F213 V214 L215 M216 M217 M218 S219 K220 W221 H222 R223 R224 T225 N226 L227 A228 I229 D230 Y231 P234 Q235 L236 A237 D238 K239 F240 A241 Y242 R243 T247 V248 Q249 D250 A251 D252 E253 W254 I255 E256 G257 D258 R259 T260 D261 D262 Q263
F264 R265 P266 K270 V271 M272 L273 S274 A275 L276 R277 K278 Y279 V280 N281 H282 N283 R284 L285 Y286 N287 Q288 T291 A292 A293 Q294 L295 L296 A297 Q298 I299 M300 M301 K302 P303 V304 P305 N306 C307 A308 E309 G310 Y311 A312 W313 L314 M315 H316 D317 A318 L319 V320 N321 I322 P323 K324 F325 G326
S327 I328 R329 G330 R331 Y332 P333 F334 L335 L336 S337 G338 D339 A340 A341 L342 I343 Q344 L348 W351 S352 A353 I354 M355 P358 E359 L360 V361 F362 T363 Y364 A365 M366 Q367 V368 S369 V370 A371 L372 G375 L376 Y377 L378 R379 R380 V381 K382 K383 T384 G385 F386 I390 D391 D392 S393
Y394 E395 D396 G397 L400 Q401 P402 E403 V406 Q407 A408 A409 L410 C413 T414 G415 Q416 D417 A418 P419 L420 M423 S424 D425 V426 Y427 V428 T429 Y430 P431 L434 E435 F436 D437 A438 V442 P443 I444 T445 V446 I447 E448 P449 A450 G451 Y452 N453 I454 V455 D456 D457 H458 L459 V460
G463 V464 P465 V466 A467 C468 S469 P470 Y471 M472 I473 P474 P475 V476 A477 A478 F479 D480 T481 A482 N483 P484 Y485 C486 G487 W488 F489 V490 I491 K492 A493 A494 N495 K496 Y497 L498 R499 K500 G501 A502 V503 Y504 D505 K506 L507 E508 A509 W510 A513 W514 R517 V518 A519 G520 Y521 D522 T523
H524 F525 K526 V527 D530 T531 H532 G533 L534 T535 K536 F537 Y538 A539 T544 W545 T546 H547 I548 P549 E550 F551 V552 T553 M558 E559 V560 F561 V562 T563 A564 I565 R568 A569 R570 H571 F572 V573 E574 L575 P576 R577 L578 N579 S580 P581 A582 F583 F584 R585 S586 V587 E588 V589 I593
Y594 D595 T596 H597 V598 Q599 A600 G601 A602 H603 A604 V605 Y606 H607 A608 S609 R610 I611 M612 L613 D614 Y615 V616 K617 P618 G622 I623 Q624 V625 I626 N627 A628 G629 E630 L631 K632 N633 Y634 W635 G636 S637 V638 R639 R640 T641 Q642 Q643 G644 L650 T651 MET PRO ALA VAL MET PRO THR GLY
GLU PRO THR ALA GLY ALA ALA HIS GLU GLU LEU ILE GLU GLN ALA ASP ASN VAL LEU VAL GLU

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	407.00Å 403.20Å 572.00Å 90.00° 90.46° 90.00°	Depositor
Resolution (Å)	30.00 – 3.50 47.85 – 3.28	Depositor EDS
% Data completeness (in resolution range)	(Not available) (30.00-3.50) 22.6 (47.85-3.28)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.31 (at 3.25Å)	Xtriage
Refinement program	X-PLOR 3.851	Depositor
R, R_{free}	0.266 , 0.268 0.390 , 0.390	Depositor DCC
R_{free} test set	96853 reflections (9.99%)	wwPDB-VP
Wilson B-factor (Å^2)	57.0	Xtriage
Anisotropy	0.158	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , -17.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.41$, $< L^2 > = 0.23$	Xtriage
Estimated twinning fraction	0.064 for k,h,-l 0.065 for -k,-h,-l 0.067 for h,-k,-l	Xtriage
F_o,F_c correlation	0.59	EDS
Total number of atoms	10302	wwPDB-VP
Average B, all atoms (Å^2)	62.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.47% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

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	0.29	0/5289	0.57	4/7212 (0.1%)
1	B	0.29	0/5289	0.60	4/7212 (0.1%)
All	All	0.29	0/10578	0.59	8/14424 (0.1%)

There are no bond length outliers.

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	532	HIS	N-CA-C	-8.36	88.42	111.00
1	A	452	TYR	N-CA-C	-7.01	92.06	111.00
1	B	492	LYS	N-CA-C	-5.83	95.27	111.00
1	B	474	PHE	N-CA-C	5.65	126.25	111.00
1	B	613	LEU	N-CA-C	5.56	126.02	111.00

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-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 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	5151	0	5012	470	0
1	B	5151	0	5012	579	0
All	All	10302	0	10024	1036	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 51.

The worst 5 of 1036 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:16:LEU:HB2	1:B:363:THR:HG21	1.23	1.14
1:B:445:THR:HG23	1:B:446:VAL:H	1.05	1.14
1:A:15:ASP:HB2	1:A:608:ALA:HB1	1.25	1.13
1:A:144:ALA:HB3	1:A:165:ASP:HA	1.26	1.10
1:B:376:LEU:HB3	1:B:464:VAL:HG21	1.31	1.10

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	649/680 (95%)	597 (92%)	36 (6%)	16 (2%)	5 34
1	B	649/680 (95%)	540 (83%)	63 (10%)	46 (7%)	1 12
All	All	1298/1360 (95%)	1137 (88%)	99 (8%)	62 (5%)	2 20

5 of 62 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	13	SER
1	A	397	GLY
1	A	473	ILE
1	A	531	THR
1	A	536	LYS

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	550/573 (96%)	462 (84%)	88 (16%)	2 14
1	B	550/573 (96%)	439 (80%)	111 (20%)	1 6
All	All	1100/1146 (96%)	901 (82%)	199 (18%)	1 9

5 of 199 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	178	ARG
1	B	363	THR
1	B	196	ILE
1	B	301	MET
1	B	407	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 39 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	154	HIS
1	B	597	HIS
1	B	235	GLN
1	B	416	GLN
1	B	642	GLN

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)

There are no ligands in this entry.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.