wwPDB X-ray Structure Validation Summary Report (i)

Jan 27, 2024 – 05:13 PM EST

PDB ID	:	1CU1
Title	:	CRYSTAL STRUCTURE OF AN ENZYME COMPLEX FROM HEPATITIS C VIRUS
Authors	:	Yao, N.; Weber, P.C.
Deposited on	:	1999-08-20
Resolution	:	2.50 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:

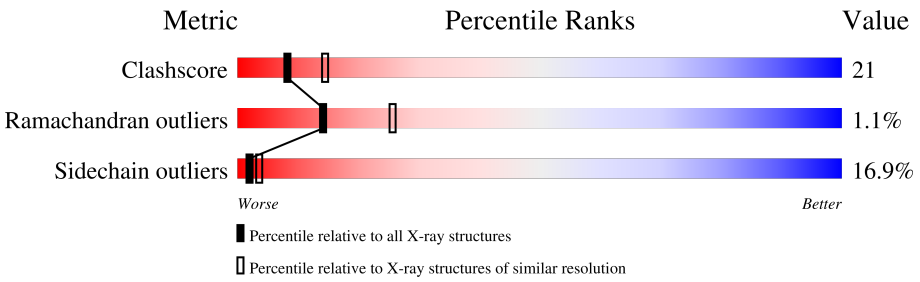
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (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 2.50 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))
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)

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	645	59% 33% 8% •
1	B	645	62% 31% 7% •

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 9894 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.

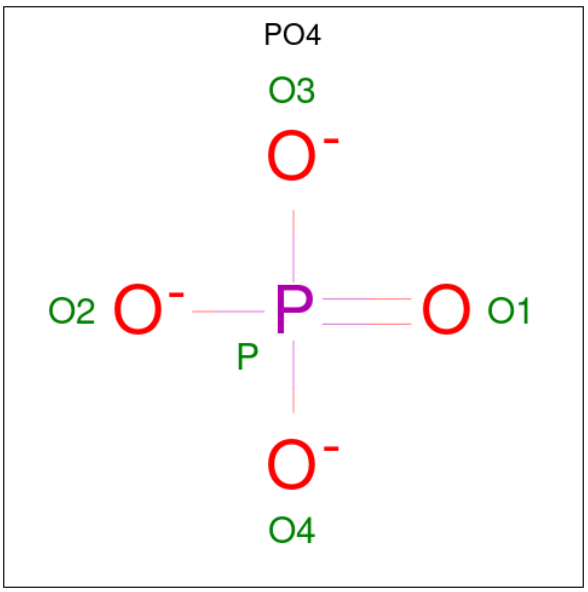
- Molecule 1 is a protein called PROTEIN (PROTEASE/HELICASE NS3).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	645	4807	3026	834	917	30	0	0	0
1	B	645	4807	3026	834	917	30	0	0	0

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
2	A	1	1	1	0	0
2	B	1	1	1	0	0

- Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	P		
3	A	1	5	4	1	0	0
			Total	O	P		
3	B	1	5	4	1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	150	150	150	0	0
			Total	O		
4	B	118	118	118	0	0

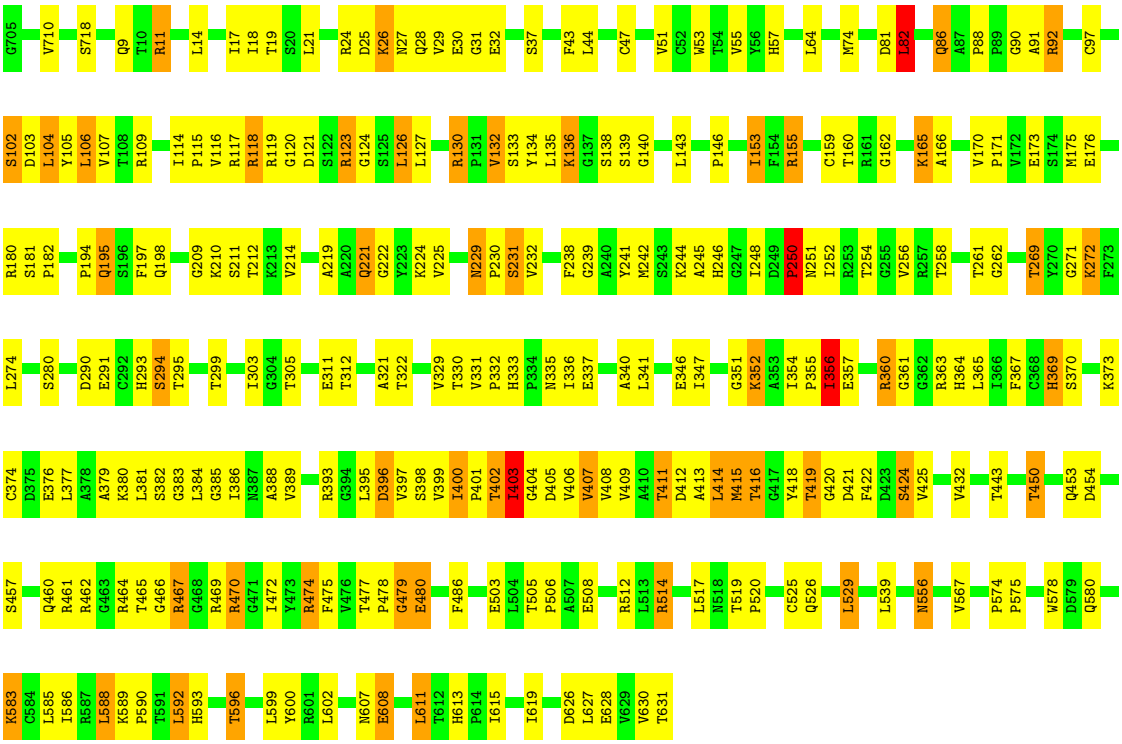
3 Residue-property plots (i)

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.

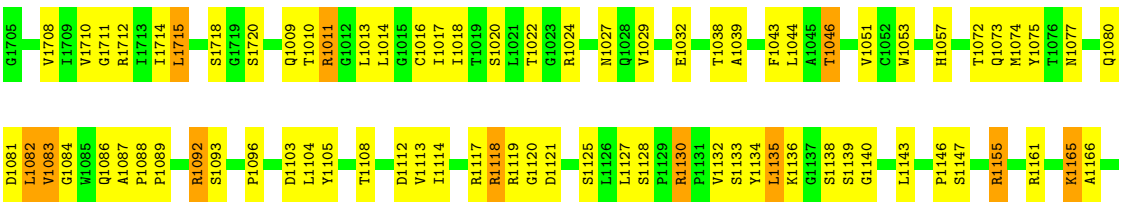
• Molecule 1: PROTEIN (PROTEASE/HELICASE NS3)

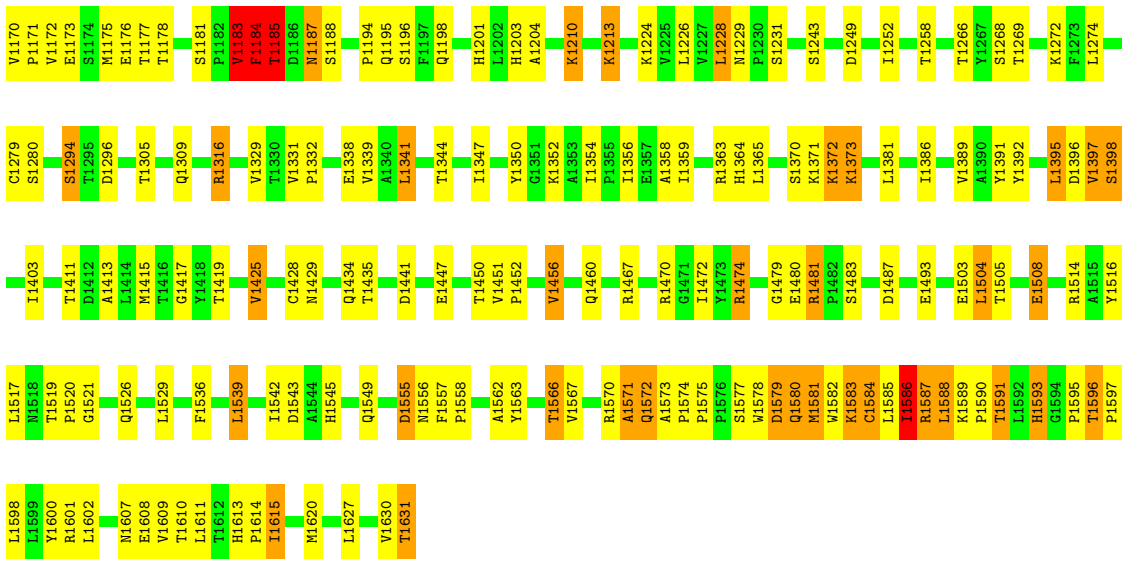
Chain A: 59% 33% 8% •



• Molecule 1: PROTEIN (PROTEASE/HELICASE NS3)

Chain B: 62% 31% 7% •




V1170 P1171 V1172 E1173 S1174 M1175 E1176 T1177 T1178 S1181 P1182 V1183 F1184 T1185 D1186 N1187 S1188 P1194 Q1195 S1196 F1197 Q1198 H1201 L1202 H1203 A1204 K1210 K1213 K1224 V1225 L1226 V1227 L1228 N1229 P1230 S1231 S1243 D1249 I1252 T1258 T1266 Y1267 S1268 T1269 K1272 F1273 L1274
C1279 S1280 S1294 T1295 D1296 T1305 Q1309 R1316 V1329 T1330 V1331 P1332 E1338 V1339 A1340 L1341 T1344 I1347 Y1350 G1351 K1352 A1353 I1354 P1355 I1356 E1357 A1358 I1359 R1363 H1364 L1365 S1370 K1371 K1372 K1373 L1381 I1386 V1389 A1390 Y1391 Y1392 L1395 D1396 V1397 S1398
I1403 T1411 D1412 A1413 L1414 M1415 T1416 G1417 Y1418 T1419 V1425 C1428 N1429 Q1434 T1435 D1441 E1447 T1450 V1451 P1452 V1456 Q1460 R1467 R1470 G1471 I1472 Y1473 R1474 G1479 E1480 R1481 P1482 S1483 D1487 E1493 E1503 L1504 T1505 E1508 R1514 A1515 Y1516
L1517 M1518 T1519 P1520 G1521 Q1526 L1529 F1536 L1539 I1542 D1543 A1544 H1545 Q1549 D1555 N1556 F1557 P1558 A1562 Y1563 T1566 V1567 R1570 A1571 Q1572 A1573 P1574 P1575 P1576 S1577 W1578 D1579 Q1580 M1581 W1582 K1583 C1584 L1585 I1586 R1587 L1588 K1589 P1590 T1591 L1592 H1593 G1594 P1595 T1596 P1597
L1598 L1599 Y1600 R1601 L1602 N1607 E1608 V1609 T1610 L1611 T1612 H1613 P1614 I1615 M1620 L1627 V1630 T1631

4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	91.36Å 110.51Å 141.20Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 2.50	Depositor
% Data completeness (in resolution range)	(Not available) (40.00-2.50)	Depositor
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	X-PLOR 98.1	Depositor
R, R_{free}	0.200 , 0.260	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	9894	wwPDB-VP
Average B, all atoms (Å^2)	35.0	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: PO4, ZN

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	0.67	1/4916 (0.0%)	0.83	7/6714 (0.1%)
1	B	0.67	2/4916 (0.0%)	0.83	7/6714 (0.1%)
All	All	0.67	3/9832 (0.0%)	0.83	14/13428 (0.1%)

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	0	2
1	B	0	2
All	All	0	4

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	245	ALA	C-N	9.91	1.56	1.34
1	B	1183	VAL	C-N	-8.09	1.15	1.34
1	B	1081	ASP	C-N	-5.32	1.21	1.34

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	250	PRO	O-C-N	-11.66	104.04	122.70
1	B	1183	VAL	C-N-CA	8.79	143.66	121.70
1	A	245	ALA	C-N-CA	-7.39	103.23	121.70
1	B	1092	ARG	NE-CZ-NH2	6.53	123.57	120.30
1	A	556	ASN	N-CA-C	6.31	128.05	111.00

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	105	TYR	Sidechain
1	A	250	PRO	Mainchain
1	B	1084	GLY	Mainchain
1	B	1183	VAL	Peptide

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	4807	0	4781	222	0
1	B	4807	0	4779	185	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	5	0	0	1	0
3	B	5	0	0	0	0
4	A	150	0	0	9	1
4	B	118	0	0	7	1
All	All	9894	0	9560	404	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 21.

The worst 5 of 404 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:400:ILE:HD12	1:A:400:ILE:H	1.18	1.07
1:A:596:THR:HG22	1:A:607:ASN:HD22	1.19	1.06
1:A:269:THR:HG22	1:A:272:LYS:H	1.21	1.02
1:A:360:ARG:HD2	1:A:361:GLY:H	1.27	0.99
1:B:1411:THR:HG23	1:B:1413:ALA:H	1.26	0.99

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)
4:A:783:HOH:O	4:B:1838:HOH:O[4_567]	2.11	0.09

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	643/645 (100%)	609 (95%)	28 (4%)	6 (1%)	17 31
1	B	643/645 (100%)	609 (95%)	26 (4%)	8 (1%)	13 24
All	All	1286/1290 (100%)	1218 (95%)	54 (4%)	14 (1%)	14 26

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	402	THR
1	B	1571	ALA
1	A	356	ILE
1	A	397	VAL
1	A	479	GLY

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	526/526 (100%)	440 (84%)	86 (16%)	2 4
1	B	526/526 (100%)	434 (82%)	92 (18%)	2 3
All	All	1052/1052 (100%)	874 (83%)	178 (17%)	2 3

5 of 178 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	1196	SER
1	B	1425	VAL
1	B	1224	LYS
1	B	1344	THR
1	B	1481	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	1086	GLN
1	B	1613	HIS
1	B	1198	GLN
1	B	1556	ASN
1	B	1187	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 4 ligands modelled in this entry, 2 are monoatomic - leaving 2 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	PO4	B	1800	-	4,4,4	1.63	0	6,6,6	0.76	0
3	PO4	A	800	-	4,4,4	1.56	0	6,6,6	0.51	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	800	PO4	1	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	1183:VAL	C	1184:PHE	N	1.15

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.