Full wwPDB X-ray Structure Validation Report (i)

Dec 10, 2022 – 12:54 PM EST

PDB ID	:	2OBQ
Title	:	Discovery of the HCV NS3/4A Protease Inhibitor SCH503034. Key Steps in Structure-Based Optimization
Authors	:	Prongay, A.J.; Guo, Z.; Yao, N.; Fischmann, T.; Strickland, C.; Myers Jr., J.; Weber, P.C.; Malcolm, B.; Beyer, B.M.; Ingram, R.; Pichardo, J.; Hong, Z.; Prosise, W.W.; Ramanathan, L.; Taremi, S.S.; Yarosh-Tomaine, T.; Zhang, R.; Senior, M.; Yang, R.; Arasappan, A.; Bennett, F.; Bogen, S.F.; Chen, K.; Jao, E.; Liu, Y.; Love, R.G.; Saksena, A.K.; Venkatraman, S.; Girijavallabhan, V.; Njoroge, F.G.; Madison, V.
Deposited on	:	2006-12-19
Resolution	:	2.50 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 (1)) were used in the production of this report:

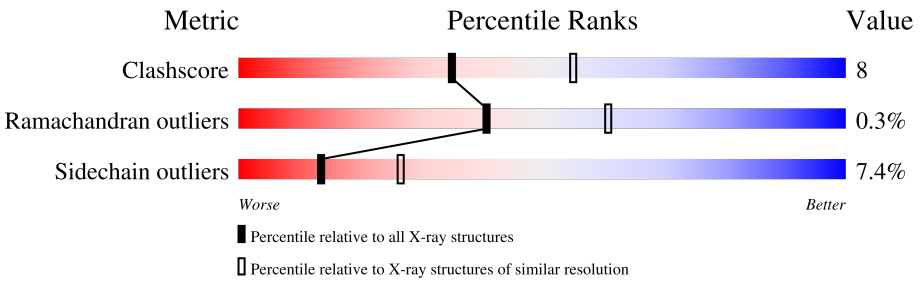
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

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	200	70% 18% • 10%
1	C	200	59% 14% • 24%
2	B	23	87% 9% •
2	D	23	43% 26% 30%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 2881 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 Hepatitis C virus.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	180	1335	833	243	250	9	0	0	0
			Total	C	N	O	S			
1	C	151	1116	699	200	209	8	0	0	0

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-10	MET	-	cloning artifact	UNP Q9ELS8
A	-9	ALA	-	cloning artifact	UNP Q9ELS8
A	-8	SER	-	cloning artifact	UNP Q9ELS8
A	-7	MET	-	cloning artifact	UNP Q9ELS8
A	-6	THR	-	cloning artifact	UNP Q9ELS8
A	-5	GLY	-	cloning artifact	UNP Q9ELS8
A	-4	GLY	-	cloning artifact	UNP Q9ELS8
A	-3	GLN	-	cloning artifact	UNP Q9ELS8
A	-2	GLN	-	cloning artifact	UNP Q9ELS8
A	-1	MET	-	cloning artifact	UNP Q9ELS8
A	0	GLY	-	cloning artifact	UNP Q9ELS8
A	119	ARG	GLN	conflict	UNP Q9ELS8
A	182	GLY	-	cloning artifact	UNP Q9ELS8
A	183	SER	-	cloning artifact	UNP Q9ELS8
A	184	HIS	-	expression tag	UNP Q9ELS8
A	185	HIS	-	expression tag	UNP Q9ELS8
A	186	HIS	-	expression tag	UNP Q9ELS8
A	187	HIS	-	expression tag	UNP Q9ELS8
A	188	HIS	-	expression tag	UNP Q9ELS8
A	189	HIS	-	expression tag	UNP Q9ELS8
C	-10	MET	-	cloning artifact	UNP Q9ELS8
C	-9	ALA	-	cloning artifact	UNP Q9ELS8
C	-8	SER	-	cloning artifact	UNP Q9ELS8
C	-7	MET	-	cloning artifact	UNP Q9ELS8
C	-6	THR	-	cloning artifact	UNP Q9ELS8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-5	GLY	-	cloning artifact	UNP Q9ELS8
C	-4	GLY	-	cloning artifact	UNP Q9ELS8
C	-3	GLN	-	cloning artifact	UNP Q9ELS8
C	-2	GLN	-	cloning artifact	UNP Q9ELS8
C	-1	MET	-	cloning artifact	UNP Q9ELS8
C	0	GLY	-	cloning artifact	UNP Q9ELS8
C	119	ARG	GLN	conflict	UNP Q9ELS8
C	182	GLY	-	cloning artifact	UNP Q9ELS8
C	183	SER	-	cloning artifact	UNP Q9ELS8
C	184	HIS	-	expression tag	UNP Q9ELS8
C	185	HIS	-	expression tag	UNP Q9ELS8
C	186	HIS	-	expression tag	UNP Q9ELS8
C	187	HIS	-	expression tag	UNP Q9ELS8
C	188	HIS	-	expression tag	UNP Q9ELS8
C	189	HIS	-	expression tag	UNP Q9ELS8

- Molecule 2 is a protein called Hepatitis C virus.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	22	159	105	29	24	1	0	0	0
			Total	C	N	O	S			
2	D	16	108	70	20	17	1	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	19	LYS	-	cloning artifact	UNP P27958
B	20	LYS	-	cloning artifact	UNP P27958
B	40	LYS	-	cloning artifact	UNP P27958
B	41	LYS	-	cloning artifact	UNP P27958
D	19	LYS	-	cloning artifact	UNP P27958
D	20	LYS	-	cloning artifact	UNP P27958
D	40	LYS	-	cloning artifact	UNP P27958
D	41	LYS	-	cloning artifact	UNP P27958

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
3	A	1	1	1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
3	C	1	1	1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	92	92	92	0	0
			Total	O		
4	B	12	12	12	0	0
			Total	O		
4	C	49	49	49	0	0
			Total	O		
4	D	8	8	8	0	0

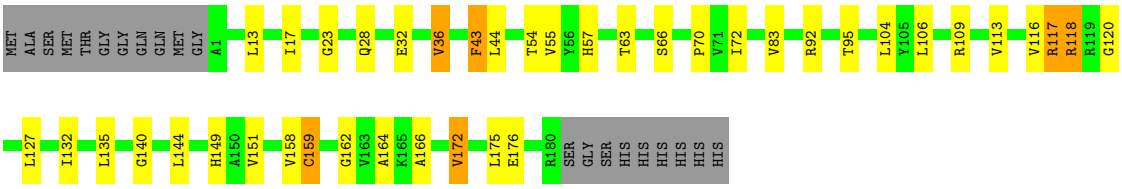
3 Residue-property plots (i)

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.

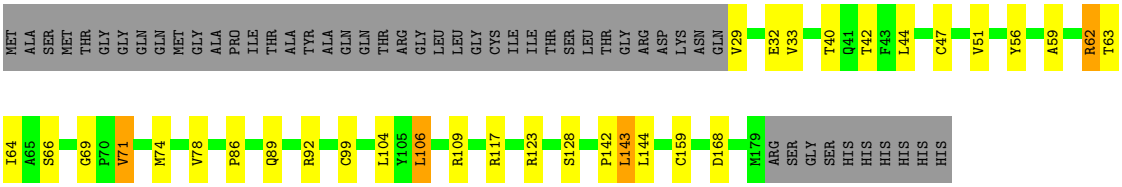
• Molecule 1: Hepatitis C virus

Chain A: 70% 18% • 10%



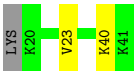
• Molecule 1: Hepatitis C virus

Chain C: 59% 14% • 24%



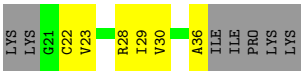
• Molecule 2: Hepatitis C virus

Chain B: 87% 9% •



• Molecule 2: Hepatitis C virus

Chain D: 43% 26% 30%



4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, α, β, γ	225.57Å 225.57Å 75.72Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	8.00 – 2.50	Depositor
% Data completeness (in resolution range)	89.4 (8.00-2.50)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
Refinement program	X-PLOR	Depositor
R, R_{free}	0.203 , 0.264	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2881	wwPDB-VP
Average B, all atoms ($Å^2$)	33.0	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: 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	0.42	0/1361	0.71	0/1856
1	C	0.39	0/1140	0.74	0/1557
2	B	0.46	0/160	0.77	0/212
2	D	0.48	0/108	0.70	0/145
All	All	0.42	0/2769	0.73	0/3770

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	1335	0	1352	24	0
1	C	1116	0	1119	19	0
2	B	159	0	193	2	0
2	D	108	0	125	8	0
3	A	1	0	0	0	0
3	C	1	0	0	0	0
4	A	92	0	0	1	0
4	B	12	0	0	0	0
4	C	49	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	D	8	0	0	1	0
All	All	2881	0	2789	45	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 8.

All (45) 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:C:63:THR:HG23	2:D:22:CYS:SG	2.10	0.91
1:A:176:GLU:HG2	4:A:987:HOH:O	1.91	0.70
1:C:66:SER:HB3	1:C:69:GLY:O	1.92	0.70
1:C:59:ALA:O	1:C:62:ARG:HB2	1.93	0.68
1:C:106:LEU:HD12	1:C:143:LEU:HG	1.82	0.61
1:A:32:GLU:HA	1:A:92:ARG:O	2.06	0.53
1:A:158:VAL:HB	1:A:166:ALA:HB3	1.90	0.53
1:C:56:TYR:CE2	1:C:78:VAL:HG22	2.43	0.53
1:C:33:VAL:HB	2:D:29:ILE:HB	1.91	0.53
1:A:36:VAL:HG22	2:B:23:VAL:HG12	1.91	0.52
1:C:64:ILE:HG23	1:C:71:VAL:HG13	1.92	0.52
1:C:64:ILE:CG2	1:C:71:VAL:HG13	2.41	0.51
1:C:42:THR:HG22	1:C:109:ARG:HH12	1.75	0.51
1:A:172:VAL:HG23	1:A:175:LEU:HD22	1.93	0.51
1:A:13:LEU:O	1:A:17:ILE:HG13	2.11	0.50
1:C:32:GLU:HG3	1:C:92:ARG:HG3	1.95	0.49
1:A:127:LEU:HD12	1:C:99:CYS:HA	1.95	0.49
1:C:117:ARG:HH11	1:C:117:ARG:HG3	1.77	0.49
1:A:104:LEU:HD22	1:A:151:VAL:HG21	1.95	0.48
1:C:29:VAL:HG21	2:D:36:ALA:HA	1.95	0.48
1:A:104:LEU:HB2	1:A:116:VAL:CG2	2.44	0.48
1:C:44:LEU:HD21	1:C:109:ARG:HA	1.96	0.47
1:A:132:ILE:HG22	1:A:162:GLY:O	2.15	0.47
1:A:55:VAL:HG21	1:A:57:HIS:CE1	2.50	0.47
2:D:30:VAL:HG13	2:D:30:VAL:O	2.15	0.46
1:A:44:LEU:O	1:A:140:GLY:HA3	2.16	0.45
2:D:28:ARG:HD3	4:D:42:HOH:O	2.15	0.45
1:A:23:GLY:HA3	1:A:70:PRO:HG3	1.98	0.45
1:A:117:ARG:HB2	1:A:127:LEU:HD21	2.00	0.44
1:A:92:ARG:HB3	1:A:92:ARG:NH1	2.33	0.44
1:A:104:LEU:HB2	1:A:116:VAL:HG22	1.99	0.43
1:C:47:CYS:HA	1:C:51:VAL:O	2.18	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:40:LYS:NZ	2:B:40:LYS:HB2	2.33	0.43
1:C:74:MET:SD	1:C:86:PRO:HG3	2.57	0.43
1:C:29:VAL:CG2	2:D:36:ALA:HA	2.49	0.42
1:C:63:THR:CG2	2:D:22:CYS:SG	2.97	0.42
1:A:63:THR:HG22	1:A:72:ILE:HA	2.00	0.42
1:A:118:ARG:HD2	1:A:120:GLY:O	2.18	0.42
1:A:106:LEU:O	1:A:113:VAL:HA	2.19	0.42
1:C:62:ARG:HB3	2:D:23:VAL:HG23	2.01	0.42
1:A:95:THR:OG1	1:A:149:HIS:CD2	2.73	0.42
1:A:44:LEU:HD21	1:A:109:ARG:HG3	2.01	0.41
1:A:159:CYS:HB3	1:A:164:ALA:HA	2.01	0.41
1:A:54:THR:CG2	1:A:83:VAL:HG22	2.51	0.41
1:A:43:PHE:CD1	1:A:43:PHE:N	2.89	0.41

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	178/200 (89%)	169 (95%)	8 (4%)	1 (1%)	25 43
1	C	149/200 (74%)	143 (96%)	6 (4%)	0	100 100
2	B	20/23 (87%)	20 (100%)	0	0	100 100
2	D	14/23 (61%)	14 (100%)	0	0	100 100
All	All	361/446 (81%)	346 (96%)	14 (4%)	1 (0%)	41 61

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	172	VAL

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	146/161 (91%)	137 (94%)	9 (6%)	18 35
1	C	123/161 (76%)	110 (89%)	13 (11%)	6 13
2	B	18/19 (95%)	18 (100%)	0	100 100
2	D	12/19 (63%)	12 (100%)	0	100 100
All	All	299/360 (83%)	277 (93%)	22 (7%)	13 27

All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	28	GLN
1	A	36	VAL
1	A	43	PHE
1	A	66	SER
1	A	117	ARG
1	A	118	ARG
1	A	135	LEU
1	A	144	LEU
1	A	159	CYS
1	C	40	THR
1	C	62	ARG
1	C	71	VAL
1	C	89	GLN
1	C	104	LEU
1	C	106	LEU
1	C	123	ARG
1	C	128	SER
1	C	142	PRO
1	C	143	LEU
1	C	144	LEU
1	C	159	CYS
1	C	168	ASP

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	27	ASN
1	A	34	GLN
1	A	80	GLN
1	A	149	HIS
1	C	49	ASN
1	C	89	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)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

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.

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 (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.