



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 09:07 PM GMT

PDB ID : 1NJT
Title : COMPLEX STRUCTURE OF HCMV PROTEASE AND A PEP-
TIDOMIMETIC INHIBITOR
Authors : Khayat, R.; Batra, R.; Qian, C.; Halmos, T.; Bailey, M.; Tong, L.
Deposited on : 2003-01-02
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
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

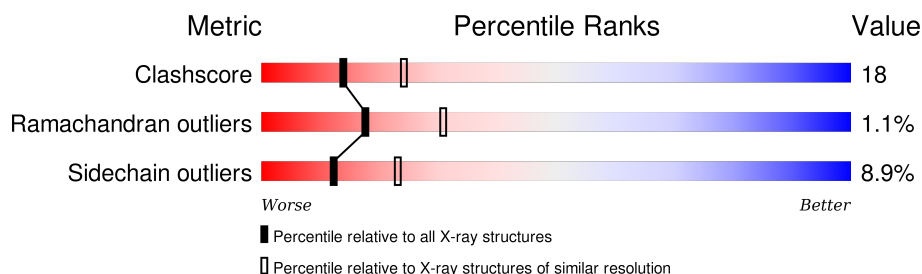
1 Overall quality at a glance

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	102246	4242 (2.50-2.50)
Ramachandran outliers	100387	4156 (2.50-2.50)
Sidechain outliers	100360	4158 (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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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 $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	256	
1	B	256	
1	C	256	
1	D	256	
2	E	6	
2	F	6	
2	G	6	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	H	6	 83% 17%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 7444 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 Capsid protein P40.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	226	Total	C	N	O	S	0	0	0
			1773	1109	325	332	7			
1	B	227	Total	C	N	O	S	0	0	0
			1782	1113	326	336	7			
1	C	222	Total	C	N	O	S	0	0	0
			1741	1090	317	329	5			
1	D	223	Total	C	N	O	S	0	0	0
			1745	1094	318	328	5			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	143	GLN	ALA	ENGINEERED	UNP P16753
B	443	GLN	ALA	ENGINEERED	UNP P16753
C	1143	GLN	ALA	ENGINEERED	UNP P16753
D	1443	GLN	ALA	ENGINEERED	UNP P16753

- Molecule 2 is a protein called Peptidomimetic Inhibitor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	6	Total	C	F	N	O	0	0	0
			39	23	3	5	8			
2	F	6	Total	C	F	N	O	0	0	0
			39	23	3	5	8			
2	G	6	Total	C	F	N	O	0	0	0
			39	23	3	5	8			
2	H	6	Total	C	F	N	O	0	0	0
			39	23	3	5	8			

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total Cl 1 1	0	0
3	A	1	Total Cl 1 1	0	0
3	D	1	Total Cl 1 1	0	0
3	C	1	Total Cl 1 1	0	0

- Molecule 4 is water.

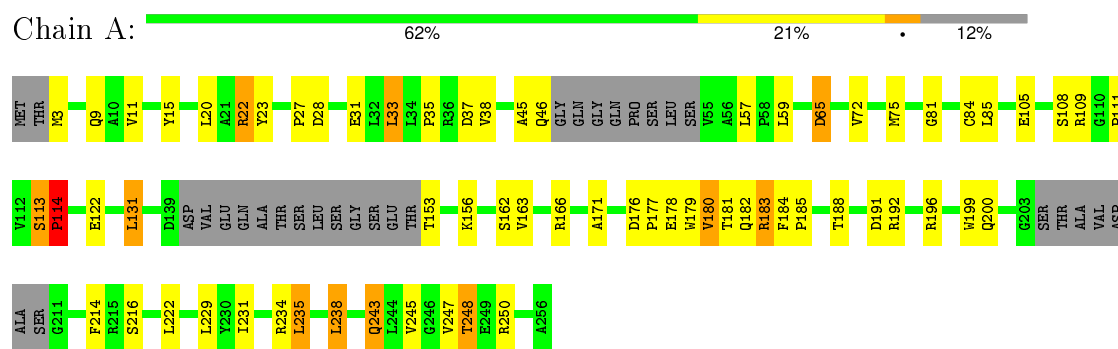
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	53	Total O 53 53	0	0
4	B	83	Total O 83 83	0	0
4	C	58	Total O 58 58	0	0
4	D	45	Total O 45 45	0	0
4	E	1	Total O 1 1	0	0
4	F	1	Total O 1 1	0	0
4	G	1	Total O 1 1	0	0
4	H	1	Total O 1 1	0	0

3 Residue-property plots

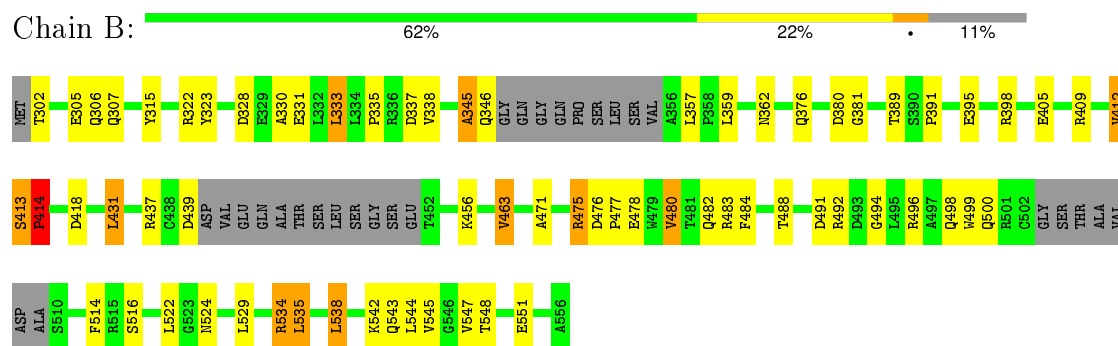
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). 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.

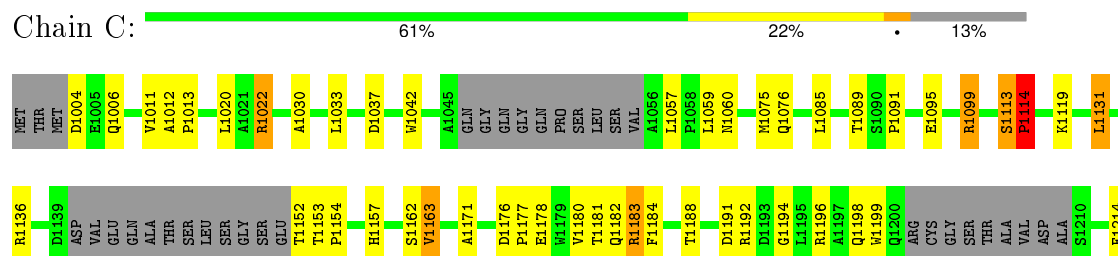
• Molecule 1: Capsid protein P40



• Molecule 1: Capsid protein P40



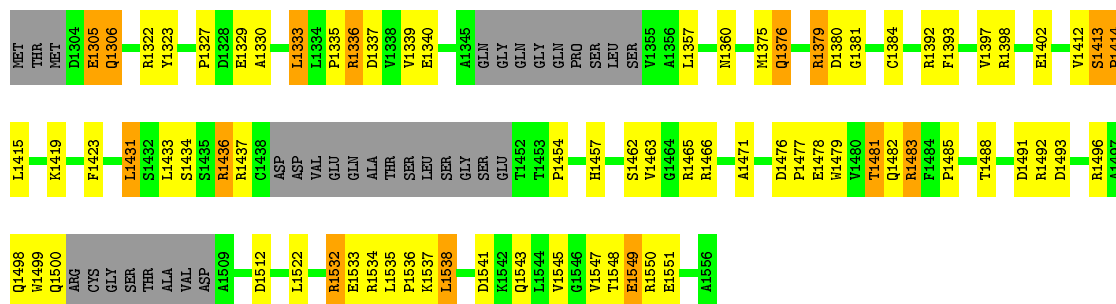
• Molecule 1: Capsid protein P40





• Molecule 1: Capsid protein P40

Chain D: 57% 24% 6% 13%



• Molecule 2: Peptidomimetic Inhibitor

Chain E: 67% 33%



• Molecule 2: Peptidomimetic Inhibitor

Chain F: 83% 17%



• Molecule 2: Peptidomimetic Inhibitor

Chain G: 83% 17%



• Molecule 2: Peptidomimetic Inhibitor

Chain H: 83% 17%



4 Data and refinement statistics

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

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	107.07Å 213.31Å 52.80Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.95 – 2.50	Depositor
% Data completeness (in resolution range)	87.6 (19.95-2.50)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.228 , 0.271	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	7444	wwPDB-VP
Average B, all atoms (Å ²)	28.0	wwPDB-VP

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: DMH, CFT, DMK, ACE, CL

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	0.38	0/1808	0.61	0/2447
1	B	0.38	0/1817	0.62	0/2460
1	C	0.38	0/1776	0.61	0/2406
1	D	0.37	0/1780	0.60	0/2412
2	E	2.57	1/11 (9.1%)	1.63	1/13 (7.7%)
2	F	2.59	1/11 (9.1%)	1.68	1/13 (7.7%)
2	G	2.48	1/11 (9.1%)	1.75	1/13 (7.7%)
2	H	2.57	1/11 (9.1%)	1.69	1/13 (7.7%)
All	All	0.43	4/7225 (0.1%)	0.62	4/9777 (0.0%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	564	ALA	C-O	8.48	1.39	1.23
2	H	1564	ALA	C-O	8.43	1.39	1.23
2	E	264	ALA	C-O	8.43	1.39	1.23
2	G	1264	ALA	C-O	8.12	1.38	1.23

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	G	1264	ALA	CA-C-O	-5.96	107.58	120.10
2	H	1564	ALA	CA-C-O	-5.80	107.92	120.10
2	F	564	ALA	CA-C-O	-5.72	108.08	120.10
2	E	264	ALA	CA-C-O	-5.57	108.41	120.10

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	1773	0	1743	65	0
1	B	1782	0	1750	62	0
1	C	1741	0	1708	67	0
1	D	1745	0	1718	71	0
2	E	39	0	33	1	0
2	F	39	0	33	0	0
2	G	39	0	33	0	0
2	H	39	0	33	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	53	0	0	3	0
4	B	83	0	0	5	0
4	C	58	0	0	5	0
4	D	45	0	0	6	0
4	E	1	0	0	0	0
4	F	1	0	0	0	0
4	G	1	0	0	0	0
4	H	1	0	0	0	0
All	All	7444	0	7051	262	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 18.

The worst 5 of 262 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:1413:SER:HB2	1:D:1414:PRO:HD3	1.45	0.97
1:D:1436:ARG:HH11	1:D:1454:PRO:HA	1.35	0.91
1:A:113:SER:HB3	1:A:114:PRO:CD	2.03	0.88
1:D:1413:SER:CB	1:D:1414:PRO:HD3	2.03	0.88
1:B:418:ASP:OD1	1:B:475:ARG:HD2	1.72	0.88

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	218/256 (85%)	206 (94%)	9 (4%)	3 (1%)	14	24
1	B	219/256 (86%)	206 (94%)	10 (5%)	3 (1%)	14	24
1	C	214/256 (84%)	203 (95%)	9 (4%)	2 (1%)	21	37
1	D	215/256 (84%)	204 (95%)	9 (4%)	2 (1%)	21	37
2	E	1/6 (17%)	1 (100%)	0	0	100	100
2	F	1/6 (17%)	1 (100%)	0	0	100	100
2	G	1/6 (17%)	1 (100%)	0	0	100	100
2	H	1/6 (17%)	1 (100%)	0	0	100	100
All	All	870/1048 (83%)	823 (95%)	37 (4%)	10 (1%)	17	31

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	45	ALA
1	A	113	SER
1	B	345	ALA
1	B	413	SER
1	C	1113	SER

5.3.2 Protein sidechains

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	191/215 (89%)	173 (91%)	18 (9%)	11	20

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	193/215 (90%)	174 (90%)	19 (10%)	10	19
1	C	188/215 (87%)	175 (93%)	13 (7%)	19	35
1	D	188/215 (87%)	170 (90%)	18 (10%)	10	19
2	E	1/1 (100%)	1 (100%)	0	100	100
2	F	1/1 (100%)	1 (100%)	0	100	100
2	G	1/1 (100%)	1 (100%)	0	100	100
2	H	1/1 (100%)	1 (100%)	0	100	100
All	All	764/864 (88%)	696 (91%)	68 (9%)	12	23

5 of 68 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	480	VAL
1	C	1037	ASP
1	D	1483	ARG
1	B	483	ARG
1	B	535	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	362	ASN
1	B	376	GLN
1	C	1243	GLN
1	B	309	GLN
1	B	325	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

8 non-standard protein/DNA/RNA residues are modelled in this entry.

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
2	DMK	E	262	2	5,9,10	1.43	1 (20%)	5,13,15	0.84	0
2	DMH	E	263	2	8,9,10	0.65	0	7,11,13	0.84	0
2	DMK	F	562	2	5,9,10	1.24	1 (20%)	5,13,15	1.19	1 (20%)
2	DMH	F	563	2	8,9,10	0.70	0	7,11,13	1.02	0
2	DMK	G	1262	2	5,9,10	1.31	1 (20%)	5,13,15	0.86	0
2	DMH	G	1263	2	8,9,10	0.61	0	7,11,13	1.13	0
2	DMK	H	1562	2	5,9,10	1.54	1 (20%)	5,13,15	0.96	0
2	DMH	H	1563	2	8,9,10	0.55	0	7,11,13	1.02	0

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
2	DMK	E	262	2	-	0/4/14/16	0/0/0/0
2	DMH	E	263	2	-	0/8/10/12	0/0/0/0
2	DMK	F	562	2	-	0/4/14/16	0/0/0/0
2	DMH	F	563	2	-	0/8/10/12	0/0/0/0
2	DMK	G	1262	2	-	0/4/14/16	0/0/0/0
2	DMH	G	1263	2	-	0/8/10/12	0/0/0/0
2	DMK	H	1562	2	-	0/4/14/16	0/0/0/0
2	DMH	H	1563	2	-	0/8/10/12	0/0/0/0

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	562	DMK	CB-CG1	2.37	1.55	1.50
2	G	1262	DMK	CB-CG1	2.49	1.55	1.50
2	E	262	DMK	CB-CG1	2.80	1.55	1.50
2	H	1562	DMK	CB-CG1	2.91	1.56	1.50

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	562	DMK	CG3-CB-CG2	2.51	110.42	107.33

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
2	E	262	DMK	1	0

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 4 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 ⓘ

6.1 Protein, DNA and RNA chains ⓘ

EDS was not executed - this section will therefore be empty.

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

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates ⓘ

EDS was not executed - this section will therefore be empty.

6.4 Ligands ⓘ

EDS was not executed - this section will therefore be empty.

6.5 Other polymers ⓘ

EDS was not executed - this section will therefore be empty.