Full wwPDB X-ray Structure Validation Report (i)

Jan 29, 2024 – 05:09 PM EST

PDB ID	:	1F08
Title	:	CRYSTAL STRUCTURE OF THE DNA-BINDING DOMAIN OF THE REPLICATION INITIATION PROTEIN E1 FROM PAPILLOMAVIRUS
Authors	:	Enemark, E.J.; Chen, G.; Vaughn, D.E.; Stenlund, A.; Joshua-Tor, L.
Deposited on	:	2000-05-15
Resolution	:	1.90 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 (i)) were used in the production of this report:

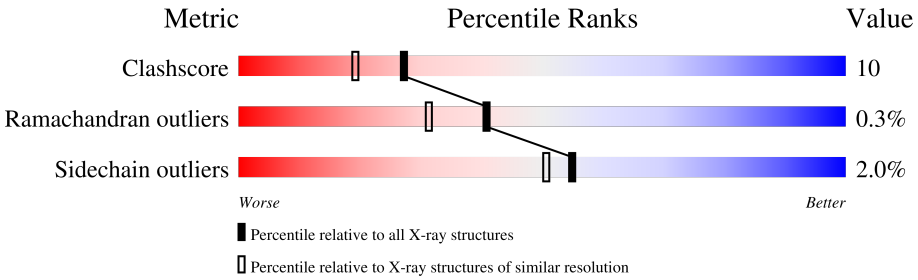
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 1.90 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	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)

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	148	85% 12% ·
1	B	148	89% 8% ··

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 2463 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 REPLICATION PROTEIN E1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	148	Total	C	N	O	S	7	0	0
			1185	765	206	205	9			
1	B	145	Total	C	N	O	S	7	0	0
			1164	754	200	201	9			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	156	GLY	-	cloning artifact	UNP P03116
A	157	SER	-	cloning artifact	UNP P03116
A	158	ARG	-	cloning artifact	UNP P03116
A	253	MET	THR	SEE REMARK 999	UNP P03116
A	266	PRO	ALA	SEE REMARK 999	UNP P03116
B	156	GLY	-	cloning artifact	UNP P03116
B	157	SER	-	cloning artifact	UNP P03116
B	158	ARG	-	cloning artifact	UNP P03116
B	253	MET	THR	SEE REMARK 999	UNP P03116
B	266	PRO	ALA	SEE REMARK 999	UNP P03116

- Molecule 2 is BROMIDE ION (three-letter code: BR) (formula: Br).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	13	Total	Br	0	0
			13	13		
2	B	1	Total	Br	0	0
			1	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	45	Total	O	0	0
			45	45		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	55	Total O 55 55	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.

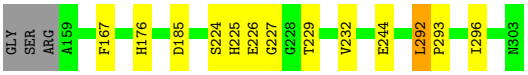
Note EDS was not executed.

• Molecule 1: REPLICATION PROTEIN E1

Chain A:

• Molecule 1: REPLICATION PROTEIN E1

Chain B:



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	42.17Å 84.90Å 124.28Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	25.00 – 1.90	Depositor
% Data completeness (in resolution range)	93.3 (25.00-1.90)	Depositor
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS	Depositor
R, R_{free}	0.241 , 0.271	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2463	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: BR

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	0.62	0/1211	0.73	0/1629
1	B	0.64	0/1190	0.74	0/1602
All	All	0.63	0/2401	0.74	0/3231

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

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	1185	0	1210	26	0
1	B	1164	0	1189	22	0
2	A	13	0	0	0	0
2	B	1	0	0	0	0
3	A	45	0	0	0	0
3	B	55	0	0	1	0
All	All	2463	0	2399	48	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 10.

All (48) 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:158:ARG:CZ	1:A:159:ALA:HB3	1.75	1.15
1:B:292:LEU:HD13	1:B:293:PRO:HD2	1.52	0.90
1:B:225:HIS:HD2	1:B:227:GLY:H	1.22	0.88
1:A:292:LEU:HD11	1:A:296:ILE:HB	1.60	0.82
1:A:158:ARG:HD3	1:A:159:ALA:N	1.95	0.81
1:A:226:GLU:H	1:A:226:GLU:CD	1.83	0.79
1:B:224:SER:CB	1:B:229:THR:HG22	2.14	0.78
1:A:292:LEU:CD1	1:A:296:ILE:HB	2.17	0.75
1:B:225:HIS:CD2	1:B:227:GLY:H	2.04	0.74
1:B:224:SER:HB3	1:B:229:THR:HG22	1.72	0.71
1:B:292:LEU:HD13	1:B:293:PRO:CD	2.20	0.71
1:A:158:ARG:NE	1:A:159:ALA:HB3	2.08	0.66
1:A:224:SER:HB2	1:A:229:THR:HG22	1.80	0.64
1:A:158:ARG:NH1	1:A:159:ALA:HB3	2.13	0.64
1:A:158:ARG:HD3	1:A:159:ALA:H	1.62	0.63
1:B:224:SER:HB2	1:B:229:THR:HG22	1.82	0.61
1:B:292:LEU:CD1	1:B:296:ILE:HB	2.29	0.61
1:A:292:LEU:HD11	1:A:296:ILE:CB	2.29	0.61
1:A:225:HIS:CD2	1:A:227:GLY:H	2.20	0.60
1:A:292:LEU:HD13	1:A:293:PRO:HD2	1.83	0.60
1:B:225:HIS:HD2	1:B:227:GLY:N	1.98	0.59
1:B:226:GLU:H	1:B:226:GLU:CD	2.07	0.57
1:A:157:SER:OG	1:A:259:GLU:HG3	2.07	0.55
1:A:301:THR:O	1:A:302:LEU:CB	2.56	0.54
1:B:292:LEU:HD12	1:B:296:ILE:HB	1.89	0.54
1:A:292:LEU:HD13	1:A:293:PRO:CD	2.38	0.53
1:A:159:ALA:CB	1:A:262:MET:SD	2.97	0.53
1:A:159:ALA:HB2	1:A:262:MET:SD	2.50	0.52
1:A:301:THR:O	1:A:302:LEU:HB3	2.11	0.51
1:A:301:THR:HG22	1:A:302:LEU:N	2.25	0.51
1:A:156:GLY:N	1:A:258:GLU:OE1	2.44	0.51
1:A:292:LEU:HD13	1:A:293:PRO:N	2.27	0.50
1:A:292:LEU:HD11	1:A:296:ILE:CG2	2.43	0.48
1:B:185:ASP:O	1:B:244:GLU:HG3	2.12	0.48
1:B:292:LEU:HD11	1:B:296:ILE:HB	1.96	0.48
1:B:167:PHE:HE2	1:B:232:VAL:HG12	1.80	0.46
1:A:224:SER:CB	1:A:229:THR:HG22	2.45	0.46
1:B:176:HIS:HD2	3:B:71:HOH:O	1.99	0.45
1:B:224:SER:HB3	1:B:229:THR:CG2	2.45	0.45
1:B:167:PHE:CE2	1:B:232:VAL:CG1	3.01	0.44
1:A:210:LEU:O	1:A:213:LYS:HG2	2.18	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:156:GLY:O	1:A:157:SER:HB2	2.18	0.43
1:A:185:ASP:O	1:A:244:GLU:HG3	2.18	0.43
1:B:292:LEU:HD11	1:B:296:ILE:HG21	1.99	0.43
1:B:167:PHE:CE2	1:B:232:VAL:HG11	2.54	0.43
1:B:167:PHE:HE2	1:B:232:VAL:CG1	2.31	0.43
1:B:292:LEU:HD11	1:B:296:ILE:CG2	2.49	0.42
1:B:292:LEU:HD11	1:B:296:ILE:CB	2.50	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	146/148 (99%)	142 (97%)	3 (2%)	1 (1%)	22 12
1	B	143/148 (97%)	140 (98%)	3 (2%)	0	100 100
All	All	289/296 (98%)	282 (98%)	6 (2%)	1 (0%)	41 31

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	302	LEU

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	129/129 (100%)	125 (97%)	4 (3%)	40 32
1	B	127/129 (98%)	126 (99%)	1 (1%)	81 82
All	All	256/258 (99%)	251 (98%)	5 (2%)	55 51

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	158	ARG
1	A	225	HIS
1	A	233	TYR
1	A	292	LEU
1	B	292	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	225	HIS
1	B	176	HIS
1	B	225	HIS

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 14 ligands modelled in this entry, 14 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 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.