wwPDB X-ray Structure Validation Summary Report (i)

Nov 14, 2016 – 04:35 PM EST

PDB ID	:	1KG0
Title	:	Structure of the Epstein-Barr Virus gp42 Protein Bound to the MHC class II Receptor HLA-DR1
Authors	:	Mullen, M.M.; Haan, K.M.; Longnecker, R.; Jardetzky, T.S.
Deposited on	:	2001-11-25
Resolution	:	2.65 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 (i) symbol.

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

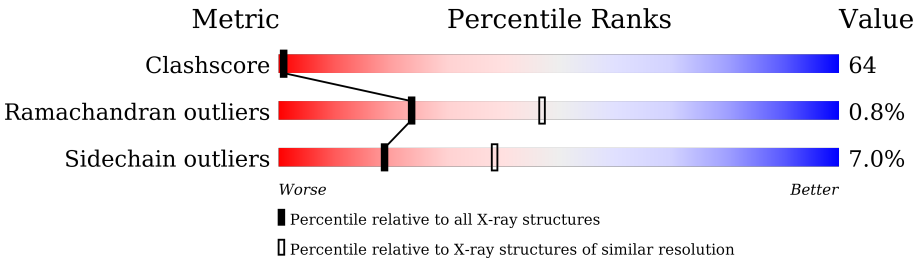
MolProbity	:	4.02b-467
Mogul	:	unknown
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20028320

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 2.65 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	3524 (2.70-2.62)
Ramachandran outliers	100387	3469 (2.70-2.62)
Sidechain outliers	100360	3469 (2.70-2.62)

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 <=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	180	47% 49% .
2	B	188	45% 50% . .
3	D	13	23% 77%
4	C	136	36% 54% 10% .

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 4844 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 MHC class II Receptor HLA-DR1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	180	Total	C	N	O	S	0	0	0
			1479	957	240	277	5			

- Molecule 2 is a protein called MHC class II Receptor HLA-DR1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	188	Total	C	N	O	S	0	0	0
			1545	973	277	289	6			

- Molecule 3 is a protein called Hemagglutinin HA Peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	D	13	Total	C	N	O	0	0	0
			106	69	18	19			

- Molecule 4 is a protein called gp42 Protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	C	136	Total	C	N	O	S	0	0	0
			1098	708	181	198	11			

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	232	Total	O	0	0
			232	232		
5	B	190	Total	O	0	0
			190	190		
5	C	176	Total	O	0	0
			176	176		
5	D	18	Total	O	0	0
			18	18		

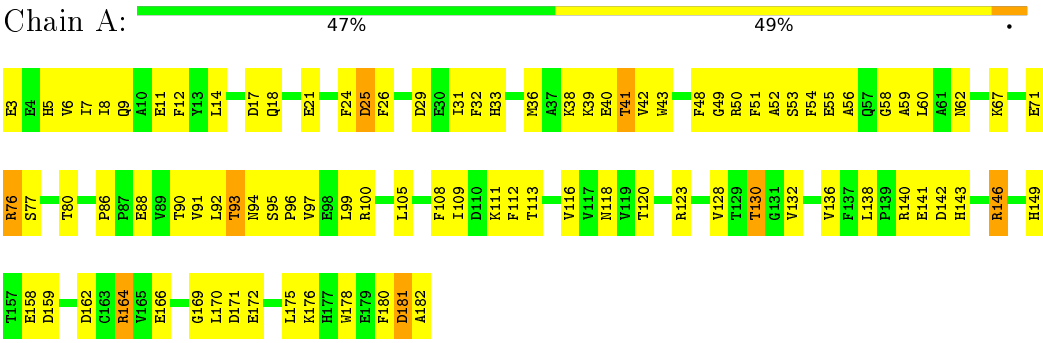
3 Residue-property plots (i)

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.

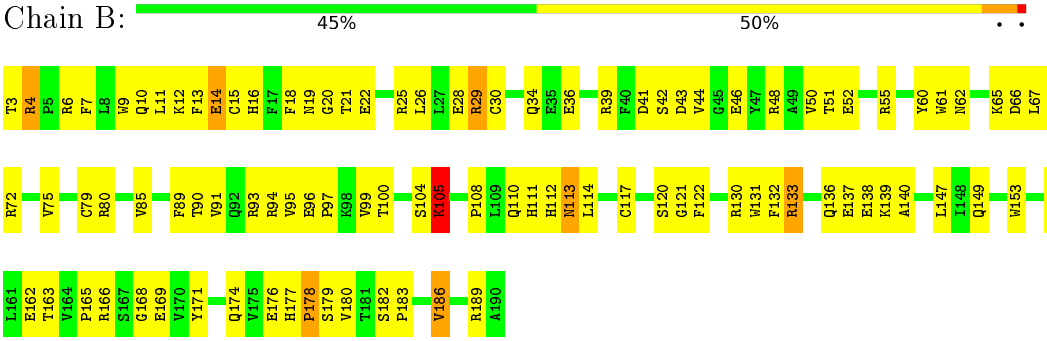
- Molecule 1: MHC class II Receptor HLA-DR1

Chain A:



- Molecule 2: MHC class II Receptor HLA-DR1

Chain B:

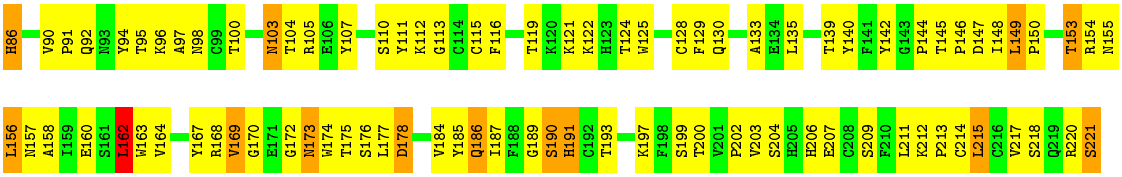


- Molecule 3: Hemagglutinin HA Peptide

Chain D:

- Molecule 4: gp42 Protein

Chain C:


H86
V90
P91
Q92
N93
Y94
T95
K96
A97
N98
C99
T100
N103
T104
R105
E106
Y107
S110
Y111
K112
G113
C114
C115
F116
T119
K120
K121
K122
H123
T124
W125
C128
F129
Q130
A133
E134
L135
T139
Y140
F141
Y142
G143
P144
T145
P146
D147
I148
L149
P150
T153
R154
N155
L156
N157
A158
I159
E160
S161
L162
W163
V164
Y167
R168
V169
G170
E171
G172
N173
W174
T175
S176
L177
D178
V184
Y185
Q186
I187
F188
G189
S190
H191
C192
T193
K197
F198
S199
T200
V201
P202
V203
S204
H205
H206
E207
C208
S209
F210
L211
K212
P213
C214
L215
C216
V217
S218
Q219
R220
S221

4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	P 63 2 2	Depositor
Cell constants a, b, c, α, β, γ	170.40Å 170.40Å 101.00Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	30.00 – 2.65	Depositor
% Data completeness (in resolution range)	(Not available) (30.00-2.65)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.221 , 0.247	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4844	wwPDB-VP
Average B, all atoms (Å^2)	62.0	wwPDB-VP

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	0.45	0/1524	0.66	0/2077
2	B	0.44	0/1585	0.71	2/2152 (0.1%)
3	D	0.43	0/107	0.71	0/141
4	C	0.50	0/1136	0.88	3/1545 (0.2%)
All	All	0.46	0/4352	0.74	5/5915 (0.1%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	105	LYS	CA-CB-CG	11.49	138.68	113.40
4	C	113	GLY	N-CA-C	8.96	135.49	113.10
2	B	105	LYS	CB-CG-CD	8.64	134.08	111.60
4	C	191	HIS	N-CA-C	6.60	128.81	111.00
4	C	162	LEU	CA-CB-CG	5.38	127.67	115.30

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	1479	0	1412	146	0
2	B	1545	0	1478	199	2
3	D	106	0	119	25	0
4	C	1098	0	1023	186	1

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	232	0	0	111	1
5	B	190	0	0	109	4
5	C	176	0	0	125	4
5	D	18	0	0	16	0
All	All	4844	0	4032	526	7

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 64.

The worst 5 of 526 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)
5:A:248:HOH:O	3:D:307:LYS:HG2	1.36	1.25
4:C:100:THR:HB	5:C:383:HOH:O	1.32	1.24
1:A:172:GLU:HA	5:A:271:HOH:O	1.35	1.24
5:B:229:HOH:O	4:C:104:THR:HB	1.35	1.24
4:C:119:THR:HB	5:C:316:HOH:O	1.35	1.24

The worst 5 of 7 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)
5:B:347:HOH:O	5:B:347:HOH:O[9_765]	2.06	0.14
2:B:174:GLN:NE2	5:C:382:HOH:O[12_555]	2.10	0.10
5:B:287:HOH:O	5:C:342:HOH:O[12_555]	2.11	0.09
5:B:344:HOH:O	5:C:263:HOH:O[3_665]	2.12	0.08
4:C:92:GLN:NE2	5:C:272:HOH:O[11_655]	2.12	0.08

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	178/180 (99%)	172 (97%)	5 (3%)	1 (1%)	30 54
2	B	186/188 (99%)	176 (95%)	9 (5%)	1 (0%)	34 59
3	D	11/13 (85%)	11 (100%)	0	0	100 100
4	C	134/136 (98%)	123 (92%)	9 (7%)	2 (2%)	13 28
All	All	509/517 (98%)	482 (95%)	23 (4%)	4 (1%)	24 47

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	C	169	VAL
4	C	190	SER
1	A	130	THR
2	B	178	PRO

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	164/164 (100%)	154 (94%)	10 (6%)	23 46
2	B	170/170 (100%)	161 (95%)	9 (5%)	28 54
3	D	12/12 (100%)	12 (100%)	0	100 100
4	C	123/123 (100%)	109 (89%)	14 (11%)	7 14
All	All	469/469 (100%)	436 (93%)	33 (7%)	19 38

5 of 33 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	69	GLU
2	B	186	VAL
4	C	209	SER
2	B	105	LYS
2	B	113	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such

sidechains are listed below:

Mol	Chain	Res	Type
2	B	156	GLN
4	C	173	ASN
4	C	103	ASN
2	B	19	ASN
3	D	311	GLN

5.3.3 RNA

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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.