



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 6, 2020 – 10:32 AM BST

PDB ID : 6QZA  
Title : HLA-DR1 with GMF Influenza PB1 Peptide  
Authors : Greenshields-Watson, A.; Rizkallah, P.J.  
Deposited on : 2019-03-11  
Resolution : 3.09 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](mailto: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 ⓘ symbol.

---

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

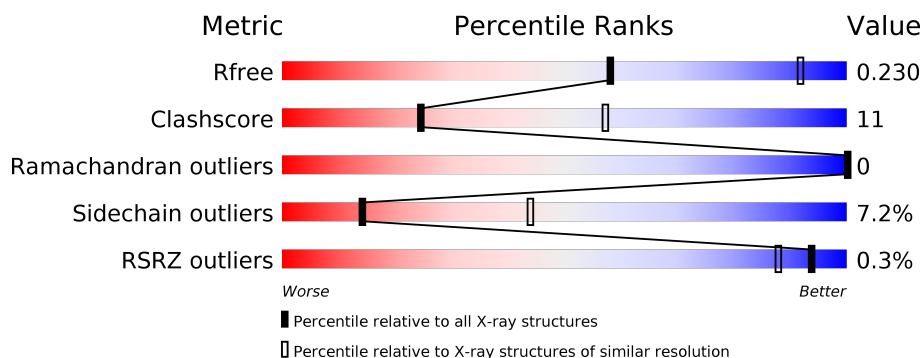
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.09 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))
$R_{free}$	130704	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

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  $\geq 3$ , 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  $\leq 5\%$ . 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.

Mol	Chain	Length	Quality of chain
1	AAA	183	<div> <div>69%</div> <div>27%</div> <div>• •</div> </div>
1	DDD	183	<div> <div>5%</div> <div>69%</div> <div>27%</div> <div>• •</div> </div>
2	BBB	191	<div> <div>71%</div> <div>24%</div> <div>• •</div> </div>
2	EEE	191	<div> <div>68%</div> <div>25%</div> <div>• 5%</div> </div>
3	CCC	20	<div> <div>65%</div> <div>10%</div> <div>25%</div> </div>
3	FFF	20	<div> <div>5%</div> <div>40%</div> <div>25%</div> <div>5%</div> <div>30%</div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 6148 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 HLA class II histocompatibility antigen, DR alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	179	Total	C	N	O	S	0	0	0
			1469	951	238	275	5			
1	DDD	179	Total	C	N	O	S	0	0	0
			1465	948	237	275	5			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	0	MET	-	initiating methionine	UNP P01903
DDD	0	MET	-	initiating methionine	UNP P01903

- Molecule 2 is a protein called HLA class II histocompatibility antigen, DRB1-1 beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	BBB	187	Total	C	N	O	S	0	0	0
			1525	957	273	289	6			
2	EEE	181	Total	C	N	O	S	0	0	0
			1481	933	263	279	6			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
BBB	0	MET	-	initiating methionine	UNP P04229
EEE	0	MET	-	initiating methionine	UNP P04229

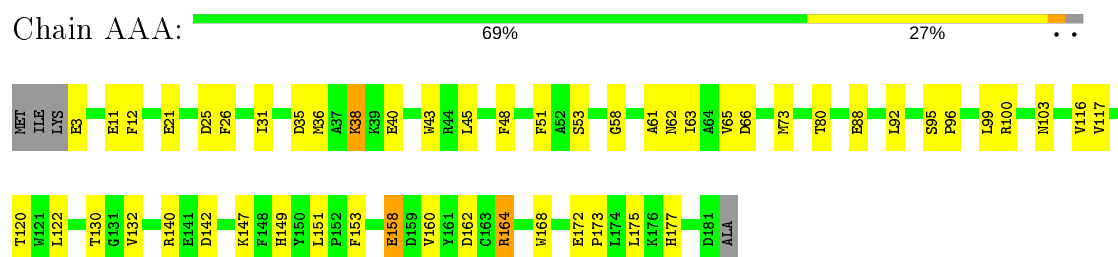
- Molecule 3 is a protein called PB1-410-422-GMF-Peptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	CCC	15	Total	C	N	O	S	0	0	0
			108	70	16	19	3			
3	FFF	14	Total	C	N	O	S	0	0	0
			100	65	15	18	2			

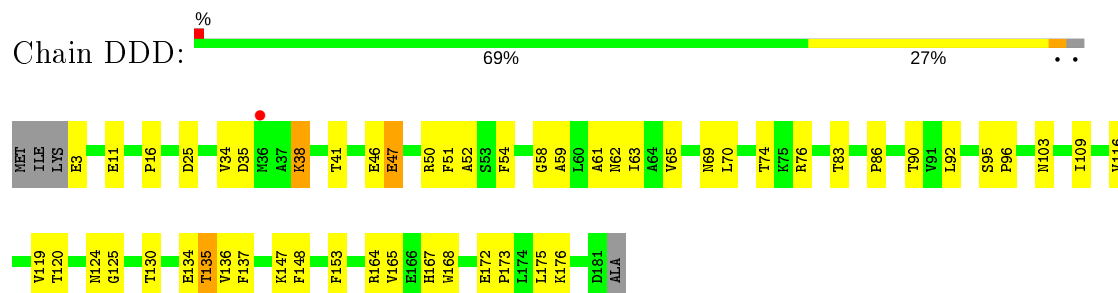
### 3 Residue-property plots [i](#)

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 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.

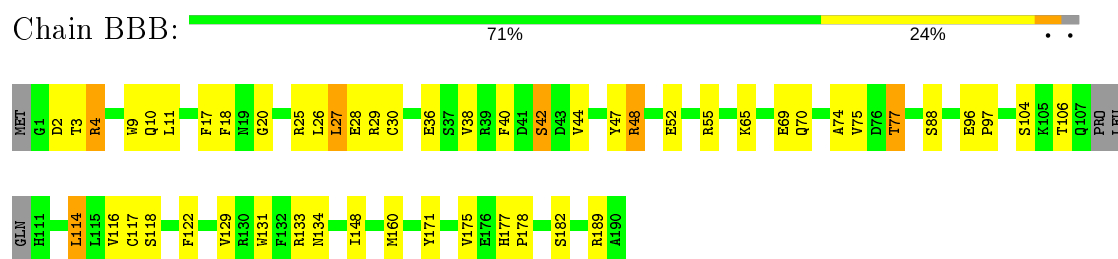
- Molecule 1: HLA class II histocompatibility antigen, DR alpha chain



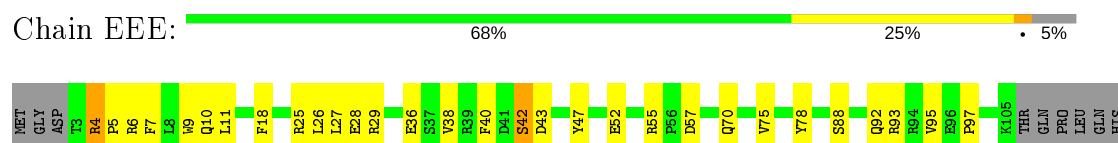
- Molecule 1: HLA class II histocompatibility antigen, DR alpha chain

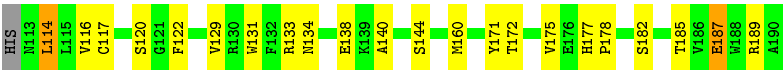


- Molecule 2: HLA class II histocompatibility antigen, DRB1-1 beta chain



- Molecule 2: HLA class II histocompatibility antigen, DRB1-1 beta chain

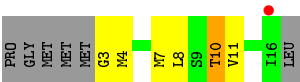




● Molecule 3: PB1-410-422-GMF-Peptide



● Molecule 3: PB1-410-422-GMF-Peptide



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	184.02 Å   184.02 Å   87.28 Å 90.00°   90.00°   120.00°	Depositor
Resolution (Å)	92.01 – 3.09 79.68 – 3.09	Depositor EDS
% Data completeness (in resolution range)	99.6 (92.01-3.09) 99.3 (79.68-3.09)	Depositor EDS
$R_{merge}$	0.19	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.38 (at 3.07 Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, $R_{free}$	0.183 , 0.226 0.185 , 0.230	Depositor DCC
$R_{free}$ test set	1531 reflections (4.88%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	69.1	Xtriage
Anisotropy	0.089	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 47.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.019 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	6148	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	81.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.53% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 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	AAA	0.46	0/1514	0.88	3/2066 (0.1%)
1	DDD	0.39	0/1510	0.81	1/2062 (0.0%)
2	BBB	0.53	0/1563	0.89	2/2122 (0.1%)
2	EEE	0.44	0/1517	0.81	0/2058
3	CCC	0.54	0/108	0.98	1/143 (0.7%)
3	FFF	0.49	0/100	0.81	0/133
All	All	0.46	0/6312	0.85	7/8584 (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	DDD	0	1

There are no bond length outliers.

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	AAA	73	MET	CG-SD-CE	-7.68	87.92	100.20
3	CCC	6	ASN	CB-CA-C	-6.40	97.60	110.40
1	DDD	124	ASN	CB-CA-C	5.96	122.31	110.40
1	AAA	164	ARG	CG-CD-NE	-5.34	100.58	111.80
2	BBB	4	ARG	CG-CD-NE	-5.27	100.73	111.80

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	DDD	125	GLY	Peptide

## 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	AAA	1469	0	1396	32	0
1	DDD	1465	0	1385	41	0
2	BBB	1525	0	1439	34	0
2	EEE	1481	0	1411	37	0
3	CCC	108	0	115	1	0
3	FFF	100	0	106	6	0
All	All	6148	0	5852	131	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 11.

The worst 5 of 131 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)
2:BBB:48:ARG:HH11	2:BBB:48:ARG:HG2	1.03	1.09
1:DDD:135:THR:HG22	1:DDD:148:PHE:H	1.37	0.89
2:EEE:114:LEU:HD12	2:EEE:160:MET:HB3	1.57	0.86
2:EEE:133:ARG:HD2	2:EEE:171:TYR:CE1	2.11	0.85
2:BBB:74:ALA:HA	2:BBB:77:THR:HG23	1.64	0.80

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	AAA	177/183 (97%)	169 (96%)	8 (4%)	0	100 100

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	DDD	177/183 (97%)	171 (97%)	6 (3%)	0	100	100
2	BBB	183/191 (96%)	171 (93%)	12 (7%)	0	100	100
2	EEE	177/191 (93%)	170 (96%)	7 (4%)	0	100	100
3	CCC	13/20 (65%)	13 (100%)	0	0	100	100
3	FFF	12/20 (60%)	12 (100%)	0	0	100	100
All	All	739/788 (94%)	706 (96%)	33 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 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	AAA	163/167 (98%)	155 (95%)	8 (5%)	25	57
1	DDD	162/167 (97%)	152 (94%)	10 (6%)	18	49
2	BBB	166/172 (96%)	152 (92%)	14 (8%)	11	38
2	EEE	162/172 (94%)	148 (91%)	14 (9%)	10	37
3	CCC	13/17 (76%)	13 (100%)	0	100	100
3	FFF	12/17 (71%)	9 (75%)	3 (25%)	0	2
All	All	678/712 (95%)	629 (93%)	49 (7%)	14	44

5 of 49 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	DDD	25	ASP
1	DDD	90	THR
2	EEE	189	ARG
1	DDD	38	LYS
1	DDD	92	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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](#)

There are no ligands in this entry.

### 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](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	AAA	179/183 (97%)	-0.18	0 <span>100</span> <span>100</span>	32, 61, 106, 118	0
1	DDD	179/183 (97%)	0.13	1 (0%) <span>89</span> <span>78</span>	59, 96, 141, 164	0
2	BBB	187/191 (97%)	-0.14	0 <span>100</span> <span>100</span>	39, 62, 106, 155	0
2	EEE	181/191 (94%)	0.10	0 <span>100</span> <span>100</span>	54, 88, 136, 165	0
3	CCC	15/20 (75%)	0.05	0 <span>100</span> <span>100</span>	64, 70, 119, 141	0
3	FFF	14/20 (70%)	0.53	1 (7%) <span>16</span> <span>6</span>	94, 117, 161, 169	0
All	All	755/788 (95%)	-0.01	2 (0%) <span>94</span> <span>88</span>	32, 77, 133, 169	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	DDD	36	MET	2.9
3	FFF	16	ILE	2.4

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

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

### 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

### 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.