



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 16, 2021 – 09:10 PM EDT

PDB ID : 1S7V  
Title : Crystal structures of the murine class I major histocompatibility complex H-2Db in complex with LCMV-derived gp33 index peptide and three of its escape variants  
Authors : Velloso, L.M.; Michaelsson, J.; Ljunggren, H.G.; Schneider, G.; Achour, A.  
Deposited on : 2004-01-30  
Resolution : 2.20 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	:	2.23.2
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)	:	2.23.2

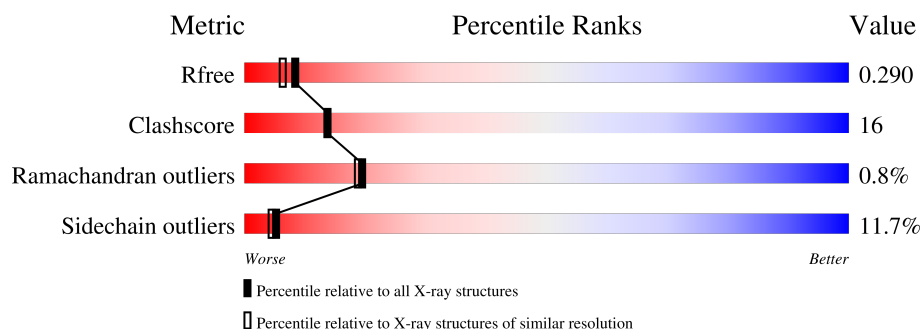
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)

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

Mol	Chain	Length	Quality of chain
1	A	338	
1	D	338	
2	B	99	
2	E	99	
3	C	9	
3	F	9	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 6651 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 H-2 class I histocompatibility antigen, D-B alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	276	Total	C	N	O	S	0	0	0
			2215	1398	393	415	9			
1	D	276	Total	C	N	O	S	0	0	0
			2210	1395	393	413	9			

- Molecule 2 is a protein called Beta-2-microglobulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	99	Total	C	N	O	S	0	0	0
			821	524	138	152	7			
2	E	99	Total	C	N	O	S	0	0	0
			821	524	138	152	7			

- Molecule 3 is a protein called Glycoprotein 9-residue peptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	9	Total	C	N	O	S	0	0	0
			70	45	11	13	1			
3	F	9	Total	C	N	O	S	0	0	0
			70	45	11	13	1			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	6	LEU	PHE	engineered mutation	UNP P07399
C	9	MET	CYS	SEE REMARK 999	UNP P07399
F	6	LEU	PHE	engineered mutation	UNP P07399
F	9	MET	CYS	SEE REMARK 999	UNP P07399

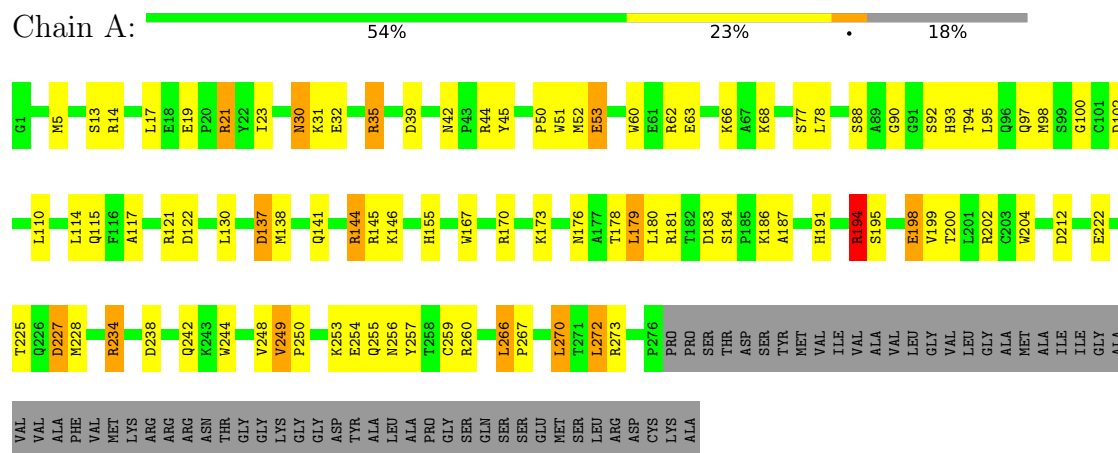
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	144	Total 144	O 144	0	0
4	B	75	Total 75	O 75	0	0
4	C	6	Total 6	O 6	0	0
4	D	154	Total 154	O 154	0	0
4	E	61	Total 61	O 61	0	0
4	F	4	Total 4	O 4	0	0

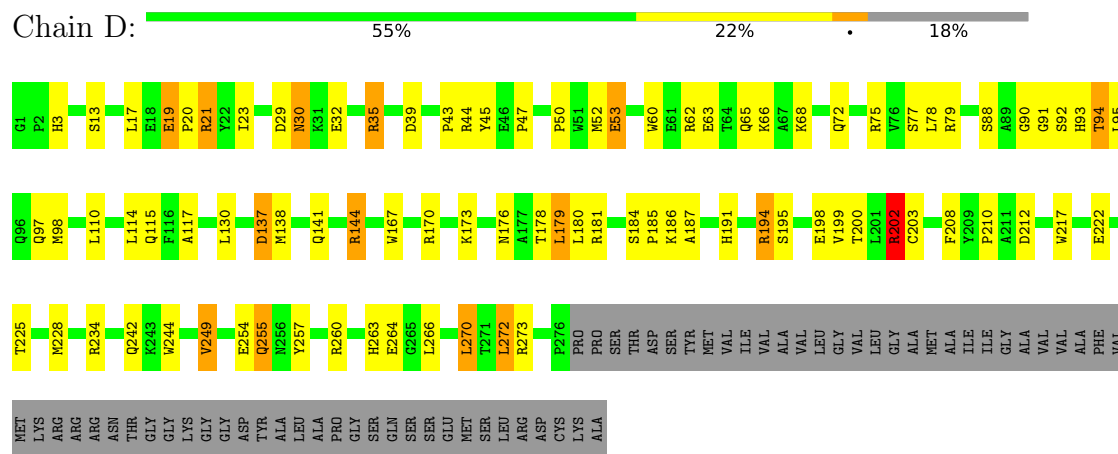
### 3 Residue-property plots

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.

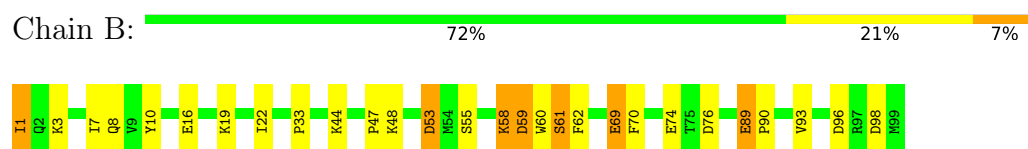
- Molecule 1: H-2 class I histocompatibility antigen, D-B alpha chain




- Molecule 1: H-2 class I histocompatibility antigen, D-B alpha chain

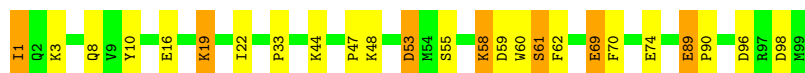


- Molecule 2: Beta-2-microglobulin



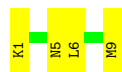
- Molecule 2: Beta-2-microglobulin

Chain E:  75% 18% 7%



- Molecule 3: Glycoprotein 9-residue peptide

Chain C:  56% 44%



- Molecule 3: Glycoprotein 9-residue peptide

Chain F:  67% 33%



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	160.47Å 91.18Å 92.15Å 90.00° 125.00° 90.00°	Depositor
Resolution (Å)	32.44 – 2.20 32.41 – 2.20	Depositor EDS
% Data completeness (in resolution range)	94.4 (32.44-2.20) 94.2 (32.41-2.20)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.63 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, $R_{free}$	0.226 , 0.267 0.277 , 0.290	Depositor DCC
$R_{free}$ test set	2639 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	47.4	Xtriage
Anisotropy	0.051	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 15.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.001 for $1/2^*h+1/2^*k+2^*l, 1/2^*h+1/2^*k, -1/2^*h+1/2^*k-l$ 0.001 for $-1/2^*h-3/2^*k-l, -1/2^*h+1/2^*k-l, 1/2^*h+1/2^*k$ 0.009 for $-1/2^*h+3/2^*k-l, 1/2^*h+1/2^*k+l, 1/2^*h-1/2^*k$ 0.009 for $1/2^*h-1/2^*k+2^*l, -1/2^*h+1/2^*k, -1/2^*h-1/2^*k-l$ 0.016 for $-h+k-l, -l, -k$ 0.016 for $-h-k-l, l, k$ 0.013 for $-1/2^*h-1/2^*k+l, -1/2^*h-1/2^*k-l, 1/2^*h-1/2^*k$ 0.003 for $-1/2^*h+1/2^*k+l, 1/2^*h-1/2^*k+l, 1/2^*h+1/2^*k$ 0.014 for $1/2^*h+3/2^*k, 1/2^*h-1/2^*k, -1/2^*h-1/2^*k-l$ 0.004 for $1/2^*h-3/2^*k, -1/2^*h-1/2^*k, -1/2^*h+1/2^*k-l$ 0.467 for $-h-2^*l, -k, l$	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	6651	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	23.0	wwPDB-VP

<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.

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.49% of the height of the origin peak. No significant pseudotranslation is detected.*



## 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	0.72	2/2281 (0.1%)	0.87	12/3098 (0.4%)
1	D	0.88	3/2277 (0.1%)	0.88	7/3092 (0.2%)
2	B	0.70	0/847	0.84	5/1148 (0.4%)
2	E	0.69	0/847	0.84	3/1148 (0.3%)
3	C	0.83	0/70	0.78	0/92
3	F	0.89	0/70	0.81	0/92
All	All	0.78	5/6392 (0.1%)	0.86	27/8670 (0.3%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	255	GLN	C-O	22.19	1.65	1.23
1	D	200	THR	C-O	9.07	1.40	1.23
1	A	255	GLN	C-O	7.10	1.36	1.23
1	A	250	PRO	C-O	5.92	1.35	1.23
1	D	255	GLN	C-N	5.49	1.46	1.34

The worst 5 of 27 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	202	ARG	NE-CZ-NH1	7.50	124.05	120.30
2	E	96	ASP	CB-CG-OD2	6.51	124.16	118.30
1	D	255	GLN	CA-C-N	-6.07	103.84	117.20
2	B	96	ASP	CB-CG-OD2	6.04	123.74	118.30
1	A	238	ASP	CB-CG-OD2	5.84	123.56	118.30

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	2215	0	2058	70	0
1	D	2210	0	2055	83	0
2	B	821	0	796	31	0
2	E	821	0	796	33	0
3	C	70	0	76	6	0
3	F	70	0	76	6	0
4	A	144	0	0	14	3
4	B	75	0	0	8	3
4	C	6	0	0	0	0
4	D	154	0	0	25	1
4	E	61	0	0	9	1
4	F	4	0	0	0	0
All	All	6651	0	5857	192	4

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 16.

The worst 5 of 192 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:255:GLN:C	1:D:255:GLN:O	1.65	1.34
1:D:52:MET:HB3	4:D:489:HOH:O	1.41	1.19
2:B:69:GLU:HG3	4:B:148:HOH:O	1.53	1.05
2:E:61:SER:HB3	4:E:118:HOH:O	1.66	0.96
1:A:98:MET:CE	1:A:115:GLN:HE21	1.87	0.88

All (4) 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)
4:D:466:HOH:O	4:E:155:HOH:O[4_556]	2.08	0.12
4:A:411:HOH:O	4:B:151:HOH:O[4_545]	2.10	0.10
4:A:443:HOH:O	4:B:158:HOH:O[4_545]	2.10	0.10
4:A:381:HOH:O	4:B:170:HOH:O[4_545]	2.16	0.04

## 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	274/338 (81%)	261 (95%)	11 (4%)	2 (1%)	22	22
1	D	274/338 (81%)	258 (94%)	14 (5%)	2 (1%)	22	22
2	B	97/99 (98%)	92 (95%)	4 (4%)	1 (1%)	15	14
2	E	97/99 (98%)	92 (95%)	4 (4%)	1 (1%)	15	14
3	C	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
3	F	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
All	All	756/892 (85%)	715 (95%)	35 (5%)	6 (1%)	19	19

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	195	SER
1	D	195	SER
1	A	88	SER
1	D	88	SER
2	B	47	PRO

### 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	224/280 (80%)	194 (87%)	30 (13%)	4	3
1	D	223/280 (80%)	196 (88%)	27 (12%)	5	4
2	B	94/94 (100%)	85 (90%)	9 (10%)	8	8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	E	94/94 (100%)	84 (89%)	10 (11%)	6	6
3	C	7/7 (100%)	7 (100%)	0	100	100
3	F	7/7 (100%)	7 (100%)	0	100	100
All	All	649/762 (85%)	573 (88%)	76 (12%)	5	4

5 of 76 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	202	ARG
2	E	58	LYS
1	D	225	THR
1	D	273	ARG
2	E	89	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	3	HIS
1	D	115	GLN
3	F	5	ASN
1	D	87	GLN
1	D	127	ASN

### 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 monosaccharides 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 ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.