W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Feb 14, 2017 – 07:12 am GMT

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)

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

|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

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	Percentile Ranks	Value
Rfree		0.306
Clashscore		16
Ramachandran outliers		0.8%
Sidechain outliers		11.7%
RSRZ outliers		10.7%

Worse   *Better*

▌Percentile relative to all X-ray structures

▯Percentile relative to X-ray structures of similar resolution
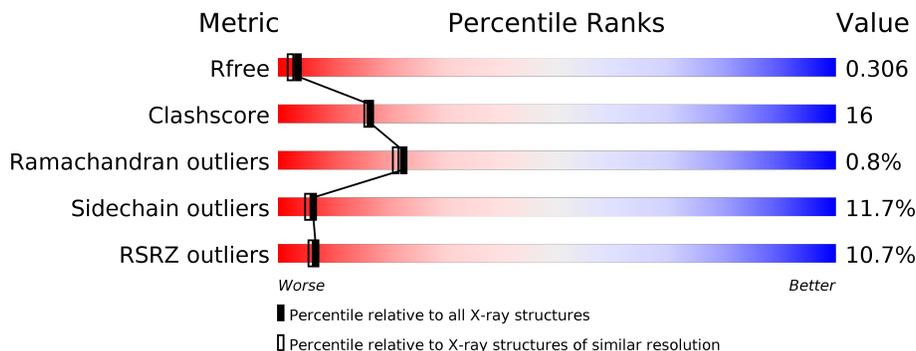
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	100719	4002 (2.20-2.20)
Clashscore	112137	4730 (2.20-2.20)
Ramachandran outliers	110173	4656 (2.20-2.20)
Sidechain outliers	110143	4657 (2.20-2.20)
RSRZ outliers	101464	4033 (2.20-2.20)

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 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.

Mol	Chain	Length	Quality of chain
1	A	338	11% / 54% 23% • 18%
1	D	338	12% / 55% 22% • 18%
2	B	99	2% / 72% 21% 7%
2	E	99	2% / 75% 18% 7%
3	C	9	56% 44%
3	F	9	67% 33%

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	UNP P07399
C	9	MET	CYS	SEE REMARK 999	UNP P07399
F	6	LEU	PHE	ENGINEERED	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 and DNA 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: H-2 class I histocompatibility antigen, D-B alpha chain
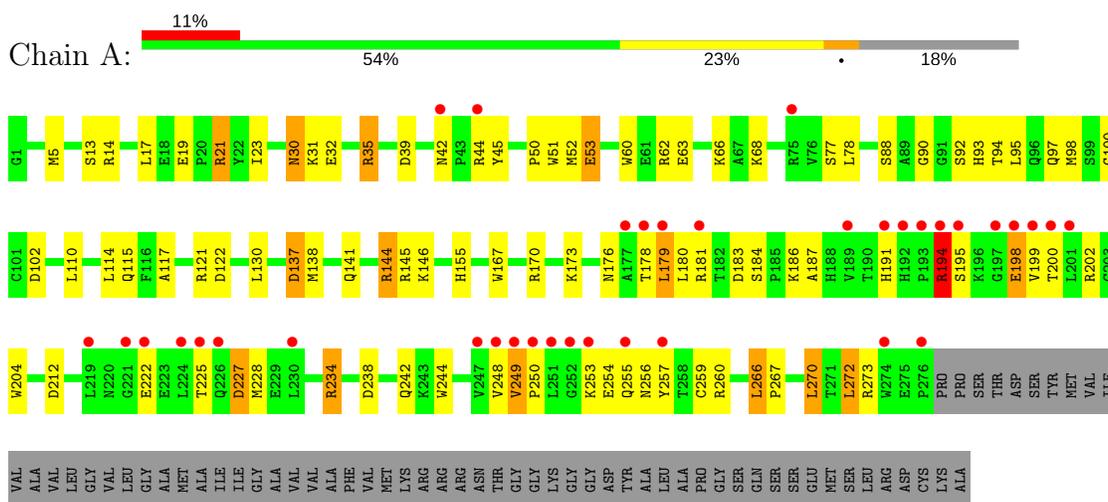
Chain A:

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

Chain D:

• Molecule 2: Beta-2-microglobulin

Chain B:

I1 Q2 K3 I7 Q8 V9 Y10 E16 K19 I22 P33 K44 P47 K48 D53 M54 S55 K58 D59 W60 S61 F62 E69 F70 E74 T75 D76 E89 P90 V93 D96 R97 D98 M99

● Molecule 2: Beta-2-microglobulin

Chain E:

2%
75%　　　　　　　　18%　　　　7%

I1 Q2 K3 Q8 V9 Y10 E16 N17 G18 K19 I22 P33 K44 P47 K48 D53 M54 S55 K58 D59 W60 S61 F62 E69 F70 E74 E89 P90 D96 R97 D98 M99

● Molecule 3: Glycoprotein 9-residue peptide

Chain C:

56%　　　　　　　　　44%

K1 N5 L6 M9

● Molecule 3: Glycoprotein 9-residue peptide

Chain F:

67%　　　　　　　33%

K1 N5 M9
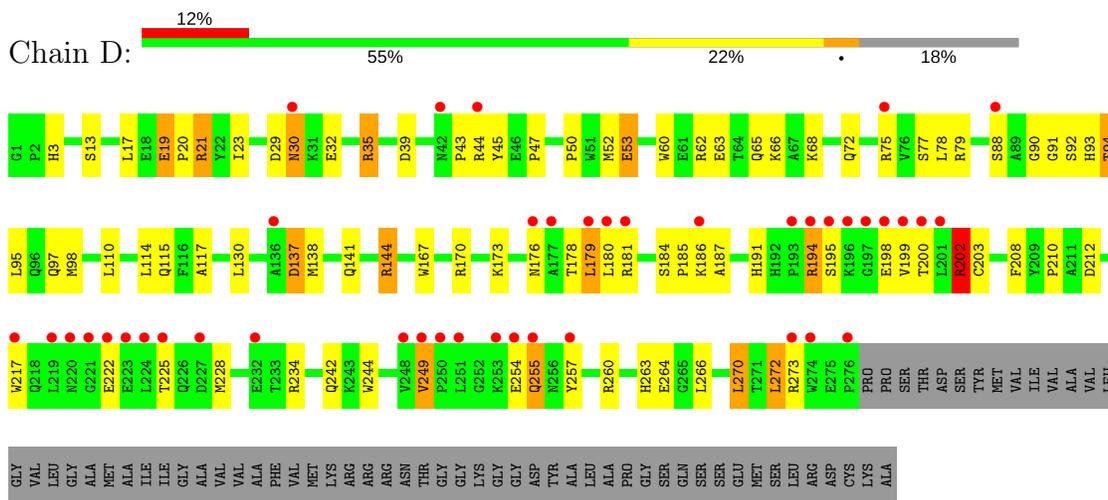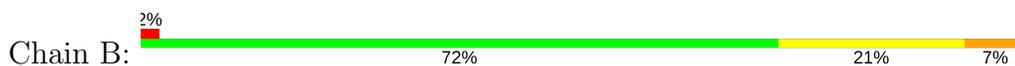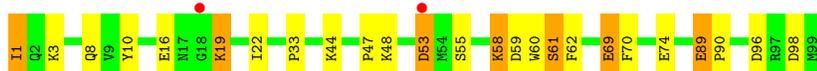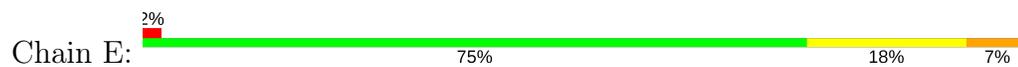
4   Data and refinement statistics   ⓘ

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	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
$<I/\sigma(I)>$ [1]	1.63 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.226 , 0.267 0.277 , 0.306	Depositor DCC
R_{free} test set	2638 reflections (5.06%)	DCC
Wilson B-factor (Å^2)	47.4	Xtriage
Anisotropy	0.051	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 15.8	EDS
L-test for twinning[2]	$<|L|> = 0.49, <L^2> = 0.33$	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^2)	23.0	wwPDB-VP

[1] Intensities estimated from amplitudes.
[2] Theoretical values of $<|L|>$, $<L^2>$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

Xtriage'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 (i)

5.1 Standard geometry (i)

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(^{o})	Ideal(^{o})
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:443:HOH:O	4:B:158:HOH:O[4_545]	2.10	0.10
4:A:411:HOH:O	4:B:151: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 (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	274/338 (81%)	261 (95%)	11 (4%)	2 (1%)	25	24
1	D	274/338 (81%)	258 (94%)	14 (5%)	2 (1%)	25	24
2	B	97/99 (98%)	92 (95%)	4 (4%)	1 (1%)	18	16
2	E	97/99 (98%)	92 (95%)	4 (4%)	1 (1%)	18	16
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%)	22	21

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 (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	224/280 (80%)	194 (87%)	30 (13%)	4	3
1	D	223/280 (80%)	196 (88%)	27 (12%)	6	5
2	B	94/94 (100%)	85 (90%)	9 (10%)	10	9

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	E	94/94 (100%)	84 (89%)	10 (11%)	8	7
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%)	6	5

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

Mol	Chain	Res	Type
2	B	58	LYS
1	D	35	ARG
2	E	55	SER
2	B	61	SER
2	B	89	GLU

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

Mol	Chain	Res	Type
2	B	2	GLN
3	C	5	ASN
2	E	8	GLN
2	B	8	GLN
1	D	3	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 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 ⓘ

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, 95th 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(Å²)	Q<0.9
1	A	276/338 (81%)	0.96	36 (13%)	4	3	11, 20, 34, 64	0
1	D	276/338 (81%)	0.88	42 (15%)	2	2	11, 20, 34, 64	0
2	B	99/99 (100%)	0.49	2 (2%)	65	63	12, 20, 33, 35	0
2	E	99/99 (100%)	0.44	2 (2%)	65	63	12, 20, 33, 35	0
3	C	9/9 (100%)	0.02	0	100	100	6, 7, 11, 13	0
3	F	9/9 (100%)	0.14	0	100	100	6, 8, 11, 13	0
All	All	768/892 (86%)	0.78	82 (10%)	7	6	6, 20, 34, 64	0

The worst 5 of 82 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	248	VAL	9.3
1	D	248	VAL	8.8
1	A	250	PRO	6.7
1	D	201	LEU	6.3
1	D	220	ASN	6.2

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

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

6.3   Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4   Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.