



wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 10:52 PM GMT

PDB ID : 3TBY
Title : CRYSTAL STRUCTURE OF THE MURINE CLASS I MAJOR HISTO-COMPATIBILITY COMPLEX H-2DB IN COMPLEX WITH THE LCMV-DERIVED GP33 ALTERED PEPTIDE ligand (V3P, Y4F)
Authors : Duru, A.D.; Allerbring, E.B.; Uchtenhagen, H.; Mazumdar, P.A.; Badia-Martinez, D.; Madhurantakam, C.; Sandalova, T.; Nygren, P.; Achour, A.
Deposited on : 2011-08-08
Resolution : 2.50 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

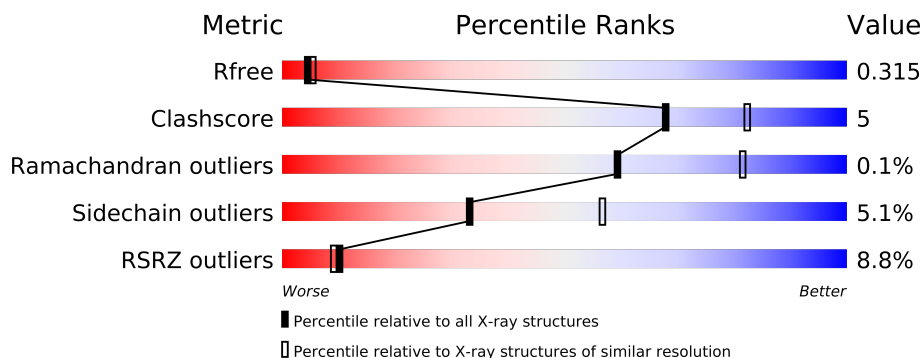
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance

The reported resolution of this entry is 2.50 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}	66092	2784 (2.50-2.50)
Clashscore	79885	3562 (2.50-2.50)
Ramachandran outliers	78287	3480 (2.50-2.50)
Sidechain outliers	78261	3482 (2.50-2.50)
RSRZ outliers	66119	2785 (2.50-2.50)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	338	
1	D	338	
1	G	338	
1	J	338	
2	B	99	
2	E	99	
2	H	99	
2	K	99	
3	C	9	
3	F	9	
3	I	9	
3	L	9	

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 12496 atoms, of which 0 are hydrogen and 0 are deuterium.

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	266	Total	C	N	O	S	0	0	0
			2187	1383	386	409	9			
1	D	266	Total	C	N	O	S	0	0	0
			2187	1383	386	409	9			
1	G	264	Total	C	N	O	S	0	0	0
			2175	1376	384	406	9			
1	J	262	Total	C	N	O	S	0	0	0
			2159	1366	381	403	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
			817	521	137	152	7			
2	E	99	Total	C	N	O	S	0	0	0
			817	521	137	152	7			
2	H	99	Total	C	N	O	S	0	0	0
			817	521	137	152	7			
2	K	99	Total	C	N	O	S	0	0	0
			817	521	137	152	7			

- Molecule 3 is a protein called GLYCOPROTEIN GPC.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	9	Total	C	N	O	S	0	0	0
			72	48	11	12	1			
3	F	9	Total	C	N	O	S	0	0	0
			72	48	11	12	1			
3	I	9	Total	C	N	O	S	0	0	0
			72	48	11	12	1			
3	L	9	Total	C	N	O	S	0	0	0
			72	48	11	12	1			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	3	PRO	VAL	ENGINEERED MUTATION	UNP P07399
C	4	PHE	TYR	ENGINEERED MUTATION	UNP P07399
C	9	MET	CYS	ENGINEERED MUTATION	UNP P07399
F	3	PRO	VAL	ENGINEERED MUTATION	UNP P07399
F	4	PHE	TYR	ENGINEERED MUTATION	UNP P07399
F	9	MET	CYS	ENGINEERED MUTATION	UNP P07399
I	3	PRO	VAL	ENGINEERED MUTATION	UNP P07399
I	4	PHE	TYR	ENGINEERED MUTATION	UNP P07399
I	9	MET	CYS	ENGINEERED MUTATION	UNP P07399
L	3	PRO	VAL	ENGINEERED MUTATION	UNP P07399
L	4	PHE	TYR	ENGINEERED MUTATION	UNP P07399
L	9	MET	CYS	ENGINEERED MUTATION	UNP P07399

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	38	Total O 38 38	0	0
4	B	26	Total O 26 26	0	0
4	C	2	Total O 2 2	0	0
4	D	35	Total O 35 35	0	0
4	E	24	Total O 24 24	0	0
4	F	1	Total O 1 1	0	0
4	G	49	Total O 49 49	0	0
4	H	19	Total O 19 19	0	0
4	I	2	Total O 2 2	0	0
4	J	22	Total O 22 22	0	0
4	K	13	Total O 13 13	0	0
4	L	1	Total O 1 1	0	0

THR	L180	G1
ASP	R181	R6
TYR	K186	G16
MET	V189	L17
VAL	P193	E18
VAL	R194	R35
ALA	SER	E41
VAL	LYS	W42
LEU	GLY	Y45
GLY	E188	R48
VAL	V199	E58
LEU	T200	R62
GLY	L206	Q65
VAL	T209	K66
ILE	W217	W73
GLY	Q218	L78
ALA	LEU	R89
VAL	ASN	G90
ALA	G221	T94
PHE	E222	S105
VAL	E223	Y113
MET	LEU	L114
LYS	THR	Q115
ARG	GLN	D122
ARG	ASP	T134
ARG	M228	T143
ASN	E229	T147
THR	R234	F148
GLY		Q149
GLY	W244	H155
GLY	A245	E166
ASP	V248	R170
TYR	V249	G175
ALA	P250	ASN
LEU	L251	ALA
ALA	G252	PRO
PRO		PRO
GLY	V257	GLY
SER		GLU
GLN	R260	THR
SER	L266	THR
GLU	L270	THR
SER		THR
LEU	R273	THR
ARG	W274	THR
ASP	E275	THR
CYS	PRO	THR
LYS	PRO	THR
ALA	PRO	THR

- Chain B:

Country	Publications
I1	10
Q2	10
C25	10
Q29	10
Q38	10
N42	10
K45	10
I46	10
P47	10
K48	10
M54	10
K58	10
D59	10
W60	10
S61	10
A66	10
F70	10
D76	10
D85	10
V93	10
M99	10

- Chain E:

Category	Count
I1	10
Q2	10
N21	10
I22	10
N42	10
P47	10
M54	10
S57	10
K58	10
S61	10
H67	10
F70	10
D76	10
K91	10
T92	20
V93	10
M99	10

- Chain H: 

Category	Count
I1	100
Q2	100
E16	100
P47	100
M54	100
K58	100
I64	100
F70	100
T75	100
V93	100
M99	100

- Chain K:

The bar chart displays the number of nodes in each layer of the network. The x-axis lists nodes from I1 to M99. The y-axis represents the number of nodes, with a scale from 0 to 100. The chart shows a repeating pattern of 10 nodes per layer, with some nodes highlighted in red.

Node	Count
I1	10
Q2	10
G18	10
N21	10
L22	10
P33	10
H34	10
L35	10
P47	10
D53	10
M54	10
K58	10
D59	10
W60	10
T73	10
D76	10
K83	10
A88	10
K91	10
T92	10
V93	10
M99	10

- Chain C: 

- Chain E:

A diagram showing a sequence of nodes: K1, F4, N5, and M9. The nodes are represented by yellow rectangular boxes. K1 is a single box. F4 and N5 are stacked vertically within a single box. M9 is a single box. Green horizontal lines connect K1 to the F4/N5 box, and the F4/N5 box to M9.

- 

Chain I: 



- Molecule 3: GLYCOPROTEIN GPC

Chain L: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	92.82Å 124.27Å 99.65Å 90.00° 103.29° 90.00°	Depositor
Resolution (Å)	20.00 – 2.50 19.97 – 2.50	Depositor EDS
% Data completeness (in resolution range)	97.9 (20.00-2.50) 97.9 (19.97-2.50)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.91 (at 2.50Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.284 , 0.315 0.285 , 0.315	Depositor DCC
R_{free} test set	3747 reflections (5.31%)	DCC
Wilson B-factor (Å ²)	40.2	Xtriage
Anisotropy	0.044	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 10.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.26$	Xtriage
Outliers	1 of 74349 reflections (0.001%)	Xtriage
F_o, F_c correlation	0.86	EDS
Total number of atoms	12496	wwPDB-VP
Average B, all atoms (Å ²)	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 11.22% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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	0.50	0/2250	0.59	0/3052
1	D	0.48	0/2250	0.60	0/3052
1	G	0.50	0/2237	0.58	0/3032
1	J	0.49	0/2221	0.59	0/3010
2	B	0.54	0/843	0.68	0/1144
2	E	0.58	0/843	0.73	2/1144 (0.2%)
2	H	0.60	0/843	0.72	0/1144
2	K	0.61	0/843	0.72	0/1144
3	C	0.77	0/74	0.85	0/97
3	F	0.67	0/74	0.84	0/97
3	I	0.68	0/74	0.87	0/97
3	L	0.82	0/74	0.69	0/97
All	All	0.53	0/12626	0.63	2/17110 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	47	PRO	N-CA-C	7.36	131.25	112.10
2	E	47	PRO	CB-CA-C	-5.54	98.16	112.00

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2187	0	0	8	0
1	D	2187	0	0	14	0
1	G	2175	0	0	10	0
1	J	2159	0	0	12	0
2	B	817	0	0	6	0
2	E	817	0	0	6	0
2	H	817	0	0	4	0
2	K	817	0	0	7	0
3	C	72	0	72	4	0
3	F	72	0	72	3	0
3	I	72	0	72	4	0
3	L	72	0	72	6	0
4	A	38	0	0	1	0
4	B	26	0	0	0	0
4	C	2	0	0	0	0
4	D	35	0	0	7	0
4	E	24	0	0	1	0
4	F	1	0	0	0	0
4	G	49	0	0	2	0
4	H	19	0	0	1	0
4	I	2	0	0	0	0
4	J	22	0	0	3	0
4	K	13	0	0	0	0
4	L	1	0	0	0	0
All	All	12496	0	288	62	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 5.

The worst 5 of 62 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:4:SER:OG	4:A:360:HOH:O	2.04	0.76
1:G:35:ARG:NH2	2:H:54:MET:O	2.25	0.70
3:L:2:ALA:HB1	3:L:3:PRO:HD2	1.74	0.68
2:B:42:ASN:ND2	2:B:76:ASP:OD1	2.27	0.67
2:H:58:LYS:NZ	4:H:104:HOH:O	2.30	0.64

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	A	258/338 (76%)	242 (94%)	16 (6%)	0	100	100
1	D	258/338 (76%)	242 (94%)	16 (6%)	0	100	100
1	G	254/338 (75%)	237 (93%)	17 (7%)	0	100	100
1	J	252/338 (75%)	236 (94%)	16 (6%)	0	100	100
2	B	97/99 (98%)	94 (97%)	3 (3%)	0	100	100
2	E	97/99 (98%)	93 (96%)	4 (4%)	0	100	100
2	H	97/99 (98%)	93 (96%)	3 (3%)	1 (1%)	22	38
2	K	97/99 (98%)	94 (97%)	3 (3%)	0	100	100
3	C	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
3	F	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
3	I	7/9 (78%)	6 (86%)	1 (14%)	0	100	100
3	L	7/9 (78%)	5 (71%)	2 (29%)	0	100	100
All	All	1438/1784 (81%)	1354 (94%)	83 (6%)	1 (0%)	59	81

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	H	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%)	214 (96%)	10 (4%)	38	63
1	D	224/280 (80%)	216 (96%)	8 (4%)	47	73

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	G	223/280 (80%)	212 (95%)	11 (5%)	35	59
1	J	221/280 (79%)	210 (95%)	11 (5%)	34	58
2	B	93/94 (99%)	88 (95%)	5 (5%)	31	53
2	E	93/94 (99%)	87 (94%)	6 (6%)	24	42
2	H	93/94 (99%)	86 (92%)	7 (8%)	19	34
2	K	93/94 (99%)	89 (96%)	4 (4%)	40	65
3	C	7/7 (100%)	7 (100%)	0	100	100
3	F	7/7 (100%)	6 (86%)	1 (14%)	5	8
3	I	7/7 (100%)	6 (86%)	1 (14%)	5	8
3	L	7/7 (100%)	5 (71%)	2 (29%)	0	0
All	All	1292/1524 (85%)	1226 (95%)	66 (5%)	33	57

5 of 66 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	F	4	PHE
1	G	166	GLU
2	K	2	GLN
1	G	41	GLU
1	G	115	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
3	L	5	ASN

5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

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(Å ²)	Q<0.9
1	A	266/338 (78%)	0.82	28 (10%) 7 6	18, 38, 73, 109	0
1	D	266/338 (78%)	0.74	25 (9%) 9 8	17, 38, 68, 124	0
1	G	264/338 (78%)	0.85	31 (11%) 5 5	17, 37, 76, 101	0
1	J	262/338 (77%)	0.88	38 (14%) 3 3	16, 41, 77, 110	0
2	B	99/99 (100%)	0.36	3 (3%) 48 50	19, 33, 44, 53	0
2	E	99/99 (100%)	0.42	1 (1%) 79 81	21, 33, 46, 48	0
2	H	99/99 (100%)	0.25	0 100 100	17, 32, 43, 51	0
2	K	99/99 (100%)	0.59	5 (5%) 27 27	22, 37, 46, 51	0
3	C	9/9 (100%)	0.21	0 100 100	18, 25, 30, 33	0
3	F	9/9 (100%)	0.05	0 100 100	18, 21, 26, 30	0
3	I	9/9 (100%)	0.41	0 100 100	21, 23, 30, 34	0
3	L	9/9 (100%)	0.34	0 100 100	22, 24, 26, 26	0
All	All	1490/1784 (83%)	0.70	131 (8%) 10 9	16, 36, 70, 124	0

The worst 5 of 131 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	177	ALA	8.7
1	D	178	THR	7.1
1	A	251	LEU	6.0
1	G	180	LEU	5.4
1	J	251	LEU	4.9

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.