Supplementary materials for the manuscript:

Dynamical Insights into *Mycobacterium avium* subsp. *paratuberculosis* peptide binding differential characteristics to HLA-DRB1 proteins associated to Multiple Sclerosis

Amit Kumar[a,b,*], Leonardo A Sechi[c], Pierluigi Caboni[d], Maria Giovanna Marrosu[e], Luigi Atzori[b], Enrico Pieroni [a,*]

[a] CRS4 Science and Technology Park Polaris, Biomedicine Dept., Piscina Manna, 09010, Pula (CA) Italy.
[b] Department of Biomedical Sciences, Oncology and Molecular Pathology Unit, University of Cagliari, Cagliari, 09100, Italy.
[c] Department of Biomedical Sciences, Microbiology and Virology Unit, University of Sassari, Sassari, Italy.
[d] Department of Life and Environmental Sciences, University of Cagliari, Cagliari, Italy.
[e] Multiple Sclerosis Center, Department of Public Health and Clinical and Molecular Medicine, University of Cagliari, Cagliari, 09100, Italy.

*Corresponding authors

Dr. Amit Kumar
Visiting Scientist, Biomedicine Dept.
CRS4 Science and Technology Park Polaris,
Piscina Manna, Pula (CA). 09010.
Italy.
Tel. +39-070-9250355
Email. amit369@gmail.com, akumar@crs4.it

Section 1 Supplementary Tables

Table S1. Details of Simulation system

System	Box Size	No of water	Total of atoms
*03:01	[86 86 74]	14975	50939
*15:01	[76 90 81]	15025	51217
*16:01	[77 96 75]	15078	51429
*15:02	[78 96 76]	15083	51434

Table S2. Cluster analysis on HLA peptide binding groove residues. Grouping of binding groove conformational states and population probability (in %) in different clusters from dPCA is reported. We also report the fluctuations (in Å) between binding groove structures in the same cluster. Absence of a cluster is denoted by symbol x.

Clusters	**C-1**		**C-2**		**C-3**		**C-4**		**C-5**	
	%	Å	%	Å	%	Å	%	Å	%	Å
*03:01	18	1.2	8	1.1	2	1.0	0.8	1.2	x	x
*15:01	18	1.0	6	0.9	4	1.0	1	1.1	0.7	0.9
*16:01	27	1.2	12	1.1	9	1.0	4	1.0	0.9	0.8
*15:02	27	1.1	10	1.1	2	1.0	x	x	x	x

Table S3. H-bond between MAP and DRB1 residues. In the third column "REF" refers to H-bonds present (represented with symbol ✓) in our starting MD structure, and if absent with a X symbol. The fourth column "MD" refers the percentage of trajectory frames that satisfy the H-bond criteria during 100 ns of MD simulations. In (A) for *03:01 predisposing protein in (B) for *15:01 predisposing protein, in (C) for *16:01 protective protein and in (D) for *15:02 protective protein.

(A)

MAP	DRB1*03:01	REF	MD
D301	H81	✓	57.5
T303	N82	✓	61.2
N305	Y26, R74, Y78	✓, ✓, ✓	91.5, 8.3, 52.9
V306	K71, R74	✓, X	23.8, 11.2
K307	S11, T12, R29, Y30	✓, X, X, ✓	8.9, 50.5, 16.4, 3.5
G308	Y30	✓	4.9
D309	Y60, Q64	X, X	36.3, 20.2

(B)

MAP	DRB1*15:01	REF	MD
D301	H81	✓	78.2
T303	N82, H81	✓, X	90.4, 3.3
N305	R13, Q70,	X, ✓	17.9, 12.3
V306	R13, Q70,	✓, X	97.9, 25.8
G308	W61	X	2.3
D309	Y60, Q64	✓, X	79.8, 69.2

(C)

MAP	DRB1*16:01	REF	MD
D301	H81	✓	72.4
T303	N82, T77	✓, X	83.2, 21.9
N305	R13, D28, R71	✓, ✓, X	84.9, 48.2, 16.2
V306	R13, R71	✓, ✓	2.3, 98.9
D309	W61, Y60	✓, X	5.8, 35.5

(D)

MAP	DRB1*15:02	REF	MD
D301	H81	✓	50.9
T303	N82, H81	✓, X	90.8, 16.2
N305	R13, F26	✓, ✓	3.6, 2.3
V306	R13, Q70	✓, ✓	88.9, 21.4
K307	D28, Y30	X, X	2.2, 18.8
G308	D66, Q70	X, X	7.3, 7.4
D309	W61, Y60	✓, X,	57.2, 35.5

Section 2 Supplementary Figures

Fig. S1. RMSD plot vs simulation time. RMSD of C-alpha atoms of binding groove residues as function of simulation time is shown here.

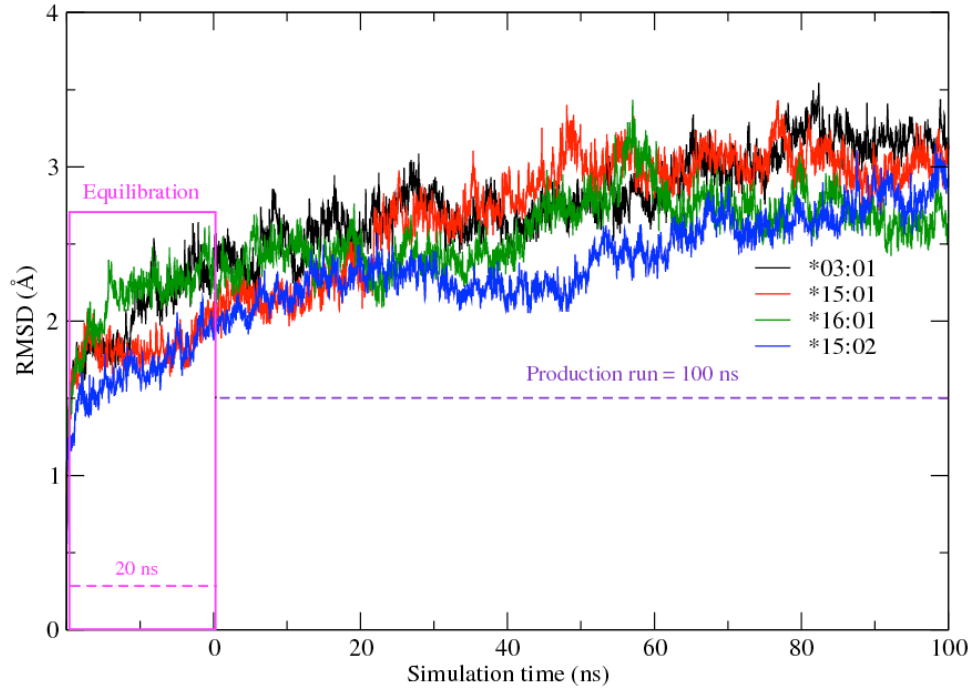


Fig. S2 RMSD probability distribution for *03:01 protein complex.

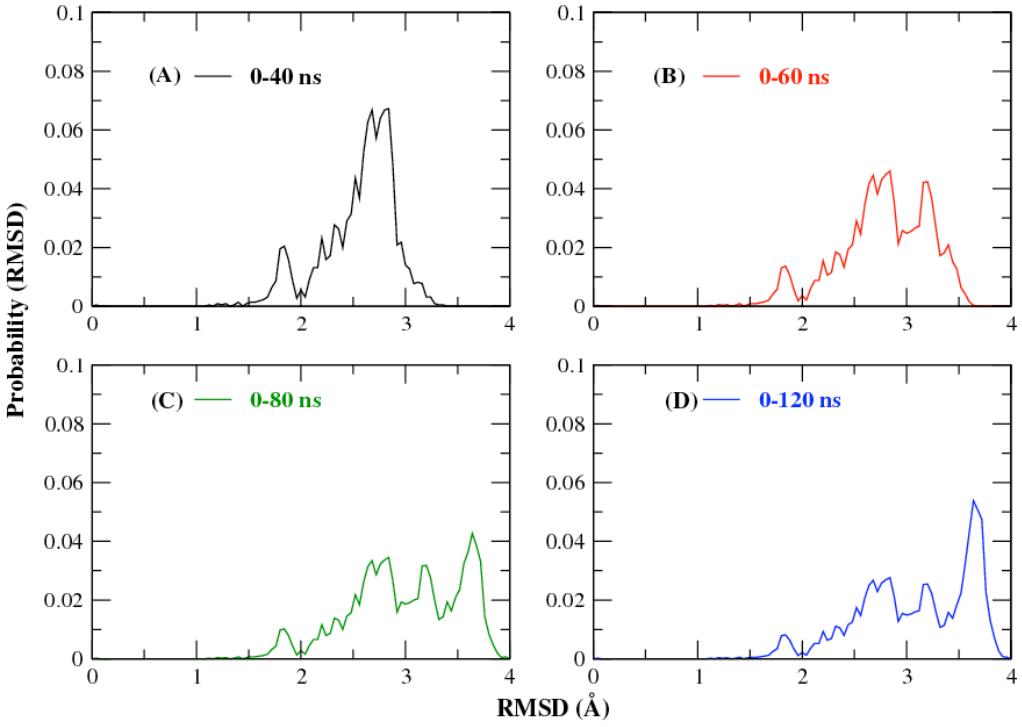


Fig. S3. Good-Turing convergence test. Probability$_{unobserved}$ (RMSD) of unobserved configurations as a function of RMSD distance.

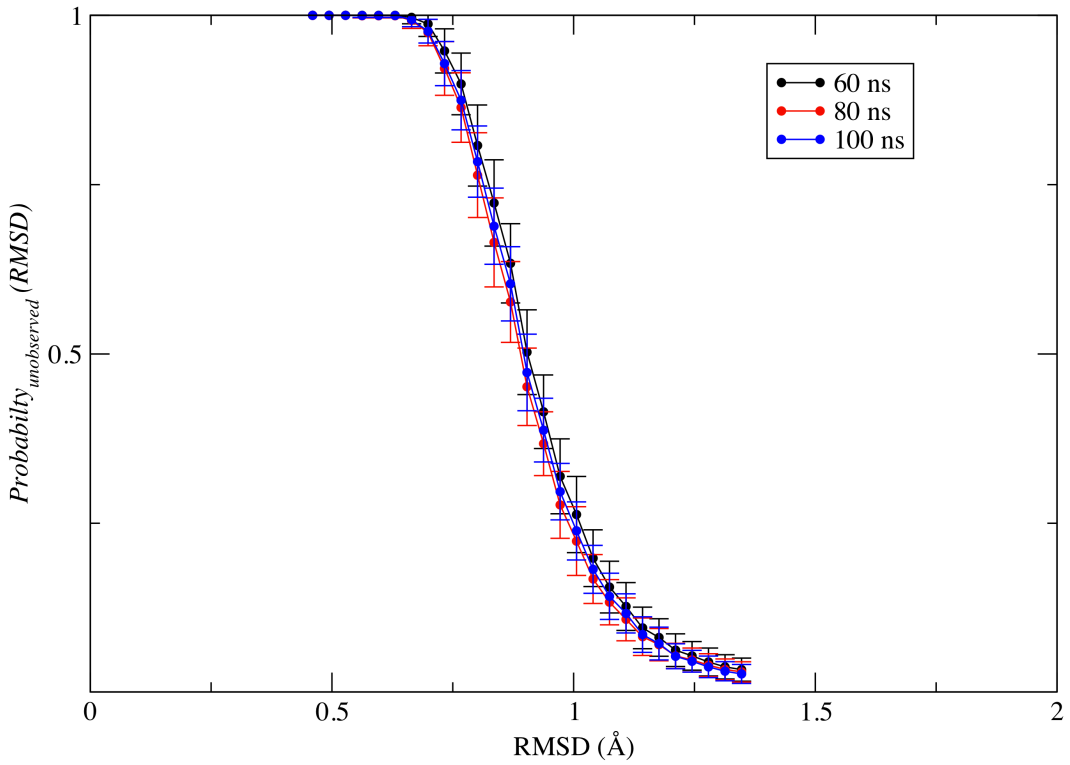


Fig. S4. Binding groove conformations during MD simulations for *03:01 and *15:01 molecular systems along the top two principal components.

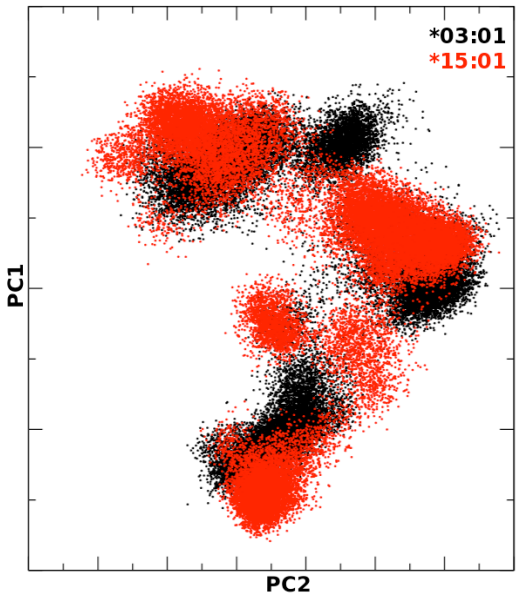


Fig. S5. **Cross correlation matrices of C-alpha fluctuations.** In (A) protective protein *16:01 and in (B) predisposing protein. The color gradient for low values of correlation (dark blue), medium correlation (yellow) and high correlation (dark red), with origin at low left corner.

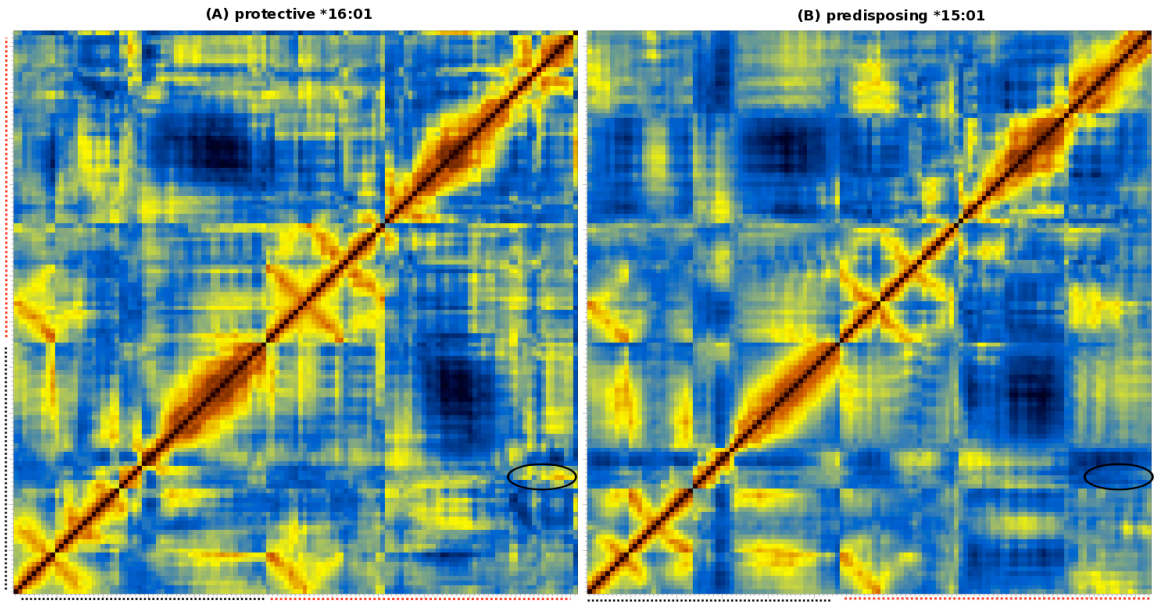
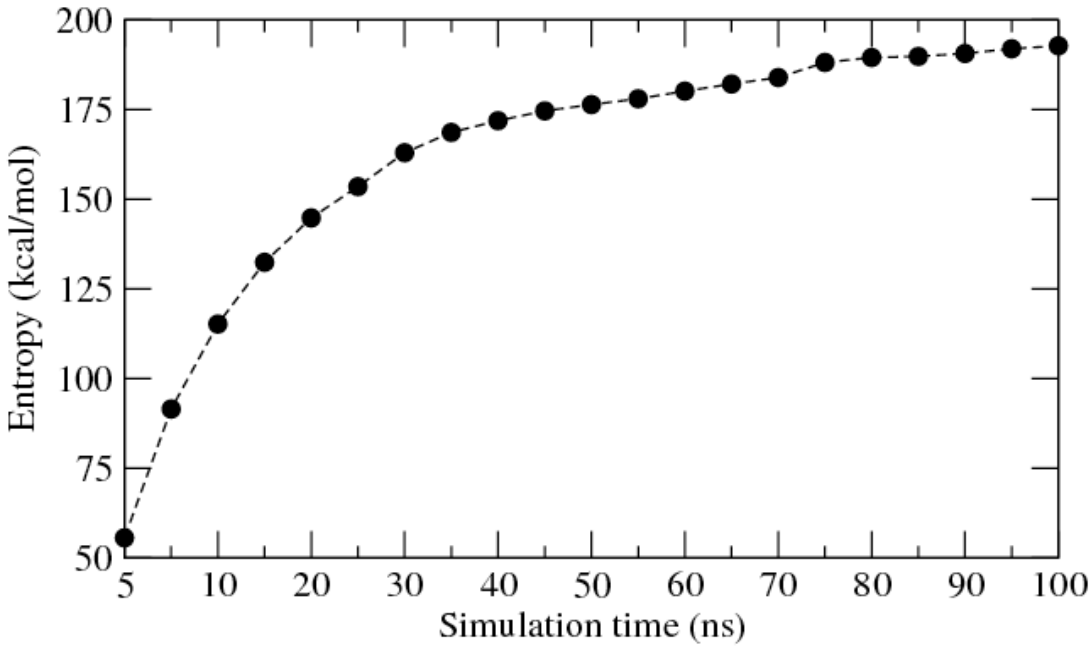


Fig. S6. Configurational Entropy estimation. Entropy calculation at different time intervals for MAP complexed to *15:01 protein.



Section 3. Molecular Dynamics simulation of *16:01 protective protein with *15:01 as template

As no X-ray structures are available for the protective proteins investigated in our study, therefore the structures were homology modeled previously using same template (*01:01) for the two protective proteins. To rule out a potential bias in choice of our template structure, we repeated MD simulations for the *16:01 protein, in this case modeled using *15:01 (1BX2) as template. We summarize the analysis of *16:01 simulations using the two template X-ray structures (i) *01:01 (pdb: 3PDO) (ii) *15:01 (pdb: 1BX2).

MAP	DRB1*16:01	*01:01 template	*15:01 template
D301	H81	✓	✓
T303	T77, N82	✓, ✓	✓, ✓
N305	R13, D28	✓, ✓	✓, ✓
V306	R71	✓	✓
G308	W61	X	✓
D309	Y60	✓	✓

Table S4. Persistent H-bond interaction between peptide and DRB1 residues

Only difference in between the two simulations is noted in additional H-bond interaction observed for *15:01 template simulation involving interacting pair G308-W61.

DRB1*16:01	*01:01 template	*15:01 template
$\Delta G_{(SIE)}$ (kcal/mol)	-10.1 ± 0.7	-10.8 ± 0.6
Configurational Entropy (kcal/mol)	168	178
Buried Surface area	545 ± 40	580 ± 28

Table S5. Physicochemical properties analysis.

We observed a slightly higher value for binding free energy, configurational entropy and buried surface area estimations for *16:01 protective protein complex simulations using *15:01 as template structure. The difference is values noted for the binding energy estimation and buried surface area calculations are within the error estimates. Moreover, the comparison with the predisposing proteins simulations results as consistent for either choice of the template in our MD simulations.