



Università degli Studi di Cagliari

DOTTORATO DI RICERCA

In Ingegneria Biomedica

Ciclo XXVI

**Modellistica computazionale molecolare di complessi proteici con
applicazioni a Sclerosi Multipla, Sindrome Fibromialgica e
Funzionamento Muscolare**

Settore/i scientifico disciplinari di afferenza

09/G2

Presentata dal Dott. Enrico Pieroni

Coordinatore Dottorato Prof. Giacomo Cao

Tutor/Relatrice Prof.ssa Maria Giovanna Marrosu

Esame finale anno accademico 2012 – 2013

INDICE

Abstract	2
Premessa	3
1. Introduzione	5
1.1. <i>La modellistica computazionale</i>	6
1.2. <i>La modellistica dei sistemi biologici</i>	6
2. Modellistica molecolare: strumenti, applicazioni e protocolli sviluppati	9
2.1. <i>Modellistica per Omologia di sequenza, Dinamica Molecolare e Docking</i>	10
2.2. <i>La proteina di membrana MHC</i>	15
2.2.1. <i>Struttura</i>	
2.2.2. <i>L'interazione peptide:MHC e TCR è correlata alle caratteristiche dinamiche del complesso p:MHC. Una analisi della letteratura.</i>	
2.2.3. <i>Protocollo SM-P: preparazione del sistema peptide:MHC</i> ✨	
2.2.4. <i>Protocollo SM-D: DM dei complessi peptide:MHC</i> ✨	
2.2.5. <i>Protocollo SM-A:</i>	
<i>analisi delle simulazioni dell'interazione peptide:MHC</i> ✨	
2.3. <i>Il sistema PAF:PAFR</i>	26
2.3.1. <i>Struttura</i>	
2.3.2. <i>Protocollo FMS-P: preparazione del sistema legando:PAFR</i> ✨	
2.3.3. <i>Protocollo FMS-D: DM dei complessi legando:PAFR</i> ✨	
2.3.4. <i>Protocollo FMS-A: Analisi della DM dei complessi legando:PAFR</i> ✨	

2.4. La proteina CASQ1	30
2.4.1. <i>Struttura</i>	
2.4.2. <i>Protocollo CASQ1-P: Preparazione della proteina CASQ1</i> ✱	
2.4.3. <i>Protocollo CASQ1-D: DM del sistema CASQ1 – Ca⁺⁺</i> ✱	
2.4.4. <i>Protocollo CASQ1-A: Analisi del sistema CASQ1 – Ca⁺⁺</i> ✱	

3. Modellistica della presentazione antigenica, applicata alla Sclerosi

Multippla	34
3.1. <i>La sclerosi multipla</i>	35
3.2. <i>Accenni al funzionamento del sistema immunitario: la presentazione degli antigeni da parte della glicoproteina di membrana MHC</i>	35
3.3. <i>Il mimetismo molecolare e l'ipotesi di una sua connessione con le malattie autoimmuni</i>	39
3.4 <i>L'architettura della regione genetica che codifica per la proteina MHC-II</i>	41
3.5. <i>Analisi di sequenza degli aplotipi del serotipo DR2</i> ✱	43
3.6. <i>Rassegna di letteratura sul ruolo del sito DRB1 86 e DQB1 38</i>	44
3.7. <i>Antigeni rilevanti per la SM</i>	46
3.8. <i>Analisi molecolare degli alleli DRB1 del gruppo serologico DR2</i> ✱	46
3.9. <i>Uno studio più ampio, a livello di sequenza e funzione, di nove aplotipi predisponenti e protettivi</i> ✱	53
3.10. <i>Analisi molecolare degli aplotipi DRB1-DQB1 del gruppo DR2</i> ✱	58
3.11. <i>Analisi molecolare dei mutanti DRB1-DQB1 del gruppo DR2</i> ✱	63
3.12. <i>Limiti e prospettive dell'analisi svolta</i> ✱	69

3.13 Conclusioni ✳	70
4. Modellistica dei meccanismi biologici del dolore, applicata alla Sindrome Fibromialgica	73
4.1. Introduzione: il sistema PAF / PAFR	74
4.2. La Sindrome Fibromialgica e l'approccio metabolomico	74
4.3. Modellistica molecolare ✳	75
4.4. Conclusioni ✳	80
5. Modellistica della proteina CASQ1, applicata al funzionamento muscolare	81
5.1 Stabilità della proteina al variare della $[Ca^{2+}]$ ✳	82
5.2. Organizzazione e struttura dei siti di legame ionico al variare della $[Ca^{2+}]$ ✳	84
5.3. Interazioni proteina-proteina che stabilizzano il dimero ✳	88
5.4. Alterazioni del dominio CAS ✳	89
5.5. Aspetti biologici ✳	90
5.6. Modello di polimerizzazione ed interazione CASQ1:Ca ²⁺ ✳	91
6. Prospettive	92
7. Conclusioni	93
8. Bibliografia	95
Ringraziamenti	111
A1. Lavori Scientifici Pubblicati, Presentazioni a Conferenze e Partecipazioni a Scuole	112

Biomedical engineering ...

... is a discipline that advances knowledge in engineering, biology and medicine, and improves human health through cross-disciplinary activities that integrate the engineering sciences with the biomedical sciences and clinical practice. It includes:

- The acquisition of new knowledge and understanding of living systems through [...] experimental and analytical techniques [...].*

- The development of new devices, algorithms, processes and systems [...].*

(London Imperial College, Department of Bioengineering Website)

Abstract

The main objective of this work is to show that clinical individual information and experimental findings on biological systems can be effectively integrated by computational modelling at an atomistic scale, to provide elucidation of, support to or proposal of functional mechanisms.

The complexity of the living systems requires specific approaches for specific diseases. In this work, I focused on Multiple Sclerosis and the antigen presentation step in the adaptive immune system response as a possible contribution to disease onset and progression, pinpointing a functional mimicry in the *self* and *nonself* peptide recognition operated by the predisposing HLA alleles. The Fibromyalgic Syndrome is also discussed and a potential link between diet-derived PAF-like lipids and the activation of a pain molecular pathway mediated by PAFR receptor is proposed. Finally the interaction of Calcium ions and the CASQ protein – relevant for muscle contractions – is analysed, and a model of the ion binding mechanisms and protein polymerization/ de-polymerization is discussed.

The investigation instrument adopted is the Molecular Dynamics Modeling, and the developed protocols for the preparation, simulation and analyses of each biological system are provided.

The therapeutical perspectives of the findings here presented are related to the design of i) peptides able to interfere with the antigen presentation mechanism and modulate autoimmune response; ii) an individual risk model for multiple sclerosis patients; iii) molecules able to interfere with the PAFR and modulate pain molecular signalling; iv) biomarkers for diagnosis support and clinical tests for the Fibromyalgic Syndrome; v) molecules able to interfere with the Calcium and CASQ interaction, able to modulate heart beat; vi) personalised diet modulating the introduction of fat possibly contributing to Fibromyalgy.

Premessa

L'obiettivo generale di questo lavoro é mostrare come le informazioni cliniche individuali e lo studio sperimentale dei sistemi biologici di rilevanza clinica possano essere integrati ed affiancati efficacemente dalla modellistica computazionale, al fine di sostenere, chiarire o proporre meccanismi funzionali. Data la complessità dei sistemi viventi, in generale l'approccio adottato é quello di analizzare malattie e problemi specifici.

Come di norma succede nell'ambito della Biomedicina e della Bioingegneria, anche questo lavoro si é svolto all'interno di una collaborazione multi-disciplinare.

Per la parte relativa alla Sclerosi Multipla, il Centro Sclerosi Multipla dell'Ospedale Binaghi (Cagliari) ed il suo Laboratorio di Biochimica, hanno fornito gli aplotipi rilevanti emersi dai test genetici sui loro pazienti, contribuendo pertanto alla scelta iniziale dei modelli di complesso proteina:peptide oltre che alla supervisione medico-clinica.

Per quanto concerne la Sindrome Fibromialgica, si é collaborato col Dipartimento di Scienze della Vita e dell'Ambiente (Unica), che ha eseguito l'analisi metabolomica dei campioni dei pazienti, fornendo i profili metabolici ematici, e coi Dipartimenti di Scienze Mediche e Biomediche (Unica), per la parte di biochimica clinica e la necessaria supervisione medico-clinica.

Lo studio dell'imagazzinamento di ioni Calcio e della polimerizzazione della proteina CASQ1 é stato svolto in collaborazione col Dipartimento di Chimica Applicata del Birla Institute of Technology (India), che ha realizzato il primo modello statico del dimero, definito il design biochimico degli esperimenti ed effettuato alcuni saggi biochimici, e col Dipartimento di Fisiologia e Biologia della Cellula, Ohio University (USA), che ha eseguito le mutazioni in vitro ed i saggi funzionali sulla proteina.

La parte di modellistica, ed in particolare la Dinamica Molecolare si é svolta integralmente al CRS4, dove assieme al Dott. A. Kumar é stato progettato il design degli esperimenti virtuali, sono stati definiti i protocolli computazionali di simulazione ed é stata effettuata l'analisi dei risultati. L'interpretazione biologica e biomedica nell'ambito della Sclerosi Multipla e della Sindrome Fibromialgica é stata principalmente curata dall'Autore.

Questo lavoro di tesi é organizzato in maniera da presentare in capitoli distinti gli aspetti tecnici dei protocolli di simulazione e le conclusioni biologiche e biomediche. Questo sia per semplificare la lettura che per la presenza di significative sovrapposizioni nei protocolli adottati, che richiedevano una sintesi unitaria. Si é comunque cercato di rendere i capitoli sulle conclusioni biologiche e biomediche abbastanza indipendenti, riportando laddove necessario una sintesi semplificata degli aspetti tecnici rilevanti.

Il materiale presentato risulta dunque organizzato nel seguente modo:

- i protocolli originali sviluppati per la preparazione dei sistemi, per l'esecuzione delle dinamiche molecolari e per l'analisi strutturale dell'output delle dinamiche sono presentati nel Capitolo 2, Sezioni 2.2 (2.2.2-5), 2.3 (2.3.2-4) e 2.4 (2.4.2-4);
- gli aspetti biologici e medici vengono presentati nei Capitoli 3, 4, e 5.

Al fine di evidenziare ulteriormente quanto originalmente sviluppato in questo lavoro di tesi, distinguendolo dalla discussione generale e della letteratura, si é evidenziata ogni sezione che presenta risultati originali con il simbolo ✳, la ruota di un timone.

Infine, si é cercato di presentare gli strumenti della modellistica in maniera semplice ma completa e, per una migliore comprensione dei problemi medici, si sono introdotti, dal punto di vista di chi non é biologo o immunologo, dei concetti chiave relativi al funzionamento del sistema immunitario ed alla segnalazione cellulare.

1. Introduzione

Biomedical informatics ...

... can be defined as the intersection of computer and information sciences with an application domain such as health care, public health, basic biomedical research, or clinical translational research.

... is the science of optimal organization, management, presentation and utilization of information relevant to human health and biology.

... research produces concepts, tools and approaches that advance what is known in the field and have the capacity to improve human health.

NIH, NLM, Department of Health and Human Services

1.1. La modellistica computazionale

Negli ultimi anni, il progresso tecnologico ha consentito l'accesso a risorse di calcolo ad alte prestazioni (HPC) a costi estremamente ridotti. Questo ha rappresentato un forte stimolo all'utilizzo ed al perfezionamento del software, delle tecniche e degli algoritmi di modellistica dei sistemi complessi, composti da un numero enorme di elementi interagenti a differenti scale di risoluzione temporale e spaziale. D'altra parte, il costo elevato, e spesso i tempi lunghi, delle ricerche in ambito clinico, farmacologico e medico, quali ad esempio lo sviluppo di nuovi farmaci, l'utilizzo di modelli animali e cellulari, gli esperimenti chimici e biochimici, ha facilitato la penetrazione della modellistica computazionale nell'ambito delle scienze della vita. Quest'ultimo rappresenta un settore che racchiude una sfida notevole per la ricerca scientifica, a causa della complessità intrinseca dei sistemi biologici, che veniva ritenuta precedentemente inattaccabile, perlomeno in maniera efficace, tramite la simulazione (Quarteroni 2006, Schlick 2002). Lo stesso progresso tecnologico ha altresì prodotto dispositivi automatici (*high throughput*) in grado di generare una enorme quantità di dati (genetici, strutturali, metabolici, ...) in tempi brevi, rendendo in generale difficili le operazioni di analisi ed elaborazione, e richiedendo dunque l'utilizzo di modelli computazionali per il data mining e lo screening non supervisionato.

1.2. La modellistica dei sistemi biologici

Data la complessità e l'estrema variabilità dei sistemi biologici, ad oggi non esistono modelli concettuali unificanti per la loro modellistica (Lesk 2005). In tal modo, ogni problema va affrontato singolarmente, sia esso l'interazione farmaco-recettore (ad esempio in *drug design*), proteina-proteina (contrazione muscolare, segnalazione e metabolismo cellulare), lipide-proteina (sindrome metabolica, sistema immunitario,

segnalazione cellulare), legando-proteina (*drug design*, modulazione del sistema immunitario), peptide-proteina (segnalazione e regolazione cellulare, modulazione del sistema immunitario), carboidrato-proteina (adesione e segnalazione cellulare, riconoscimento patogeni ed infiammazione), polimero-proteina (*drug delivery*), gene-proteina (*pathways* e *networks*).

In questo lavoro si è affrontato il tema generale della modellistica dei sistemi biologici, con una attenzione particolare ai sistemi che sono di interesse medico e clinico, affrontando in dettaglio: i) la sclerosi multipla e la presentazione di peptidi in grado di avviare una risposta immune ed autoimmune, ii) la sindrome fibromialgica e l'interazione fra lipidi e uno specifico recettore collegato ai meccanismi biologici del dolore; iii) il legame ed il rilascio degli ioni Calcio da parte di una specifica proteina, CASQ1, per il funzionamento delle fibre muscolari. La motivazione della ricerca è quella di fornire supporto a dei risultati sperimentali o clinici, fornire una spiegazione funzionale ad evidenze di natura statistica e proporre meccanismi funzionali da sottoporre a verifica sperimentale. L'obiettivo finale è la comprensione dettagliata di parte dei meccanismi biologici legati alla malattia, nelle sue varie fasi, così da poter progettare la strategia e la modalità operativa ottimale per poter interferire con essi, e bloccarne o rallentarne il decorso.

Per quanto riguarda i tre casi applicativi affrontati durante la tesi, le potenzialità di intervento sono ad esempio legate al design di: a) peptidi terapeutici, in grado di interferire nella presentazione antigenica e modulare la reazione autoimmune nella sclerosi multipla; b) un modello di rischio e progressione di malattia individualizzato, per lo specifico paziente di sclerosi multipla; c) molecole in grado di interferire con il recettore responsabile del pathway del dolore nella sindrome fibromialgica; d) test di laboratorio per la sindrome fibromialgica, basati su marcatori nel profilo metabolico individuale; e) molecole in grado di interferire con

l'immagazzinamento ed il rilascio del Calcio da parte della proteina CASQ2 per la modulazione del battito cardiaco; f) diete personalizzate, che bilancino i lipidi di natura esogena che possono contribuire ai meccanismi biologici del dolore nella sindrome fibromialgica.

2. Modellistica molecolare: strumenti, protocolli sviluppati

Molecular modeling [...] covers computational chemistry, life science modeling and materials modeling, and includes:

- *computer-aided molecular design;*
- *rational drug design, de novo ligand design, receptor modeling and docking;*
- *cheminformatics, data analysis, visualization and mining;*
- *computational medicinal chemistry;*
- *homology modeling;*
- *simulation of peptides, DNA and other biopolymers;*
- *quantitative structure-activity relationships (QSAR) and ADME-modeling;*
- *modeling of biological reaction mechanisms;*
- *and combined experimental and computational studies in which calculations play a major role.*

(Springer – Verlag rd.springer.com/journal/894)

Molecular Dynamics is a method to calculate the structure and energy of molecules based on nuclear motions. [Unlike QM approaches,] Electrons are not considered explicitly, but rather it is assumed that they will find their optimum distribution once the positions of the nuclei are known. This assumption is based on the Born-Oppenheimer approximation of the Schrödinger equation. The Born-Oppenheimer approximation states that nuclei are much heavier and move much more slowly than electrons. Thus, nuclear motions, vibrations and rotations can be studied separately from electrons; the electrons are assumed to move fast enough to adjust to any movement of the nuclei. In a very crude sense molecular dynamics treats a molecule as a collection of weights [(atomistic level)] connected with springs, where the weights represent the nuclei and the springs represent the bonds.

(Glactone Project, chemistry.gsu.edu/Glactone/)

2.1. Modellistica per Omologia di sequenza, Dinamica Molecolare e Docking

Homology Modeling. Una proteina è composta da una catena lineare di amminoacidi legati dal legame peptidico (Figura 2.1). Per molte proteine, di molte specie, sono note le sequenze dei nucleotidi e dunque degli amminoacidi, ma solo per alcune si hanno le strutture tridimensionali, ottenute tramite spettroscopia NMR, tecniche cristallografiche di diffrazione ai raggi X o, più recentemente, microscopia a scansione elettronica (SEM). Per una proteina obiettivo (*target*) di cui non sia nota la struttura, è possibile scegliere una proteina di struttura nota, che funga da riferimento (*template*) per costruire un modello tridimensionale della proteina obiettivo (Tramontano 2006). Per fare questo, uno dei metodi più semplici, ma anche più efficaci, è quello di effettuare una ricerca sulle banche dati genetiche (ad esempio nel portale EMBL-EBI, www.ebi.ac.uk/), e trovare le proteine che hanno la maggiore identità di sequenza (omologia) di amminoacidi con la nostra proteina obiettivo. Qualora l'identità di sequenza sia bassa, è necessario procedere ad un accurato allineamento multiplo (di più proteine, di più specie) al fine di verificare i domini e i motivi funzionali evolutivamente conservati e selezionare così il miglior template. Una volta ottenuto il template, esistono software appositi (modeler, SwissPDBViewer) (Eswar 2006, Guex 1997) che consentono di realizzare una struttura tridimensionale della proteina obiettivo (ossia assegnare ad ogni atomo una tripletta di coordinate spaziali), utilizzando l'ossatura del legame peptidico (*backbone*) della proteina template come una specie di calco. Naturalmente, essendo in taluni casi gli amminoacidi diversi e avendo ciascun amminoacido caratteristiche di ingombro sterico e gradi di libertà rotazionali distinti, è necessario procedere successivamente all'ottimizzazione delle catene laterali dei residui. Questo può essere fatto tramite software opportuni (ad esempio Scwrl, reperibile all'URL www1.jcsg.org/prod/scripts/scwrl/) che provano le varie combinazioni di angoli di rotazione delle catene laterali, sino a trovare la configurazione ottimale che non

presenta sovrapposizioni (*clash*) steriche. La struttura così ottenuta viene infine ottimizzata variando leggermente, e con algoritmi stocastici appositi (ad esempio tecniche Montecarlo affiancate da gradiente coniugato, semplice, algoritmi genetici, simulated annealing) (Hammond), tutte le posizioni atomiche sino a trovare un minimo della sua energia.

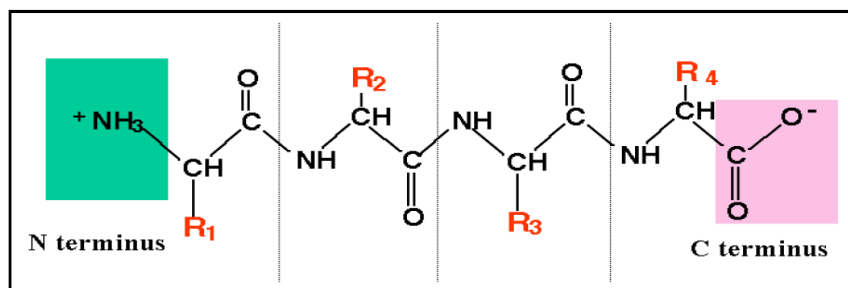


Figura 2.1. Struttura di un peptide o di una proteina: catena lineare di amminoacidi, ciascuno con distinte catene laterali, legati da legame peptidico e con una parte N-terminale e C-terminale (www.ucl.ac.uk/~sjjgsc/ProteinStructure.html).

Molecular Dynamics. Il problema del calcolo dell'energia di una proteina, ed in generale di una qualsiasi struttura molecolare, ci porta direttamente all'introduzione della Dinamica Molecolare (DM). Una molecola, per quanto complessa, può essere rappresentata come un insieme di mattoni fondamentali, detti tipi atomici, che in sintesi si comportano come delle masse cariche puntiformi soggette alle leggi classiche di Newton (Rapaport). Per visualizzare il concetto dei tipi atomici, in buona sostanza si può pensare che essi siano gli atomi canonici della tavola chimica di Mendeleev, distinti per le loro diverse geometrie di ibridazione, ad esempio l'atomo di Carbonio avrà un diverso tipo atomico per ogni stato di ibridazione che chimicamente può assumere (sp , sp^2 , sp^3). La distinzione delle diverse geometrie di coordinazione che un atomo può avere è importante, in quanto ogni geometria corrisponde a dei legami atomici ben precisi, con caratteristiche specifiche di energia di legame, distanza interatomica e vincoli rotazionali. I tipi atomici vengono fatti interagire tramite potenziali empirici, fra cui: forze elastiche di legame, angolari e

torsionali, e forze elettrostatiche di Coulomb (Figura 2.2). In particolare si introduce il potenziale semiclassico di Lennard-Jones per le forze di Van der Waals del legame di una coppia di atomi neutri, che è repulsivo a bassa distanza, così da simulare la presenza di una sfera impenetrabile, e dotato di un minimo alla distanza di legame tipica. Particolare attenzione deve essere posta nella valutazione delle forze a lungo raggio (elettrostatica e Van der Waals), che richiedono l'adozione di una qualche forma di troncamento per essere computazionalmente calcolabili in tempi ragionevoli. L'insieme dei parametri di tutte queste forze definisce uno specifico *Force Field*, che di norma deve essere specializzato a seconda delle categorie di molecole o delle molecole stesse che si considerano (Guvench 2008, Ponder 2003). Ad esempio, le proteine sono composte da pochi tipi di atomi (C, N, O, H, P, S) e quindi hanno dei force fields sviluppati appositamente per esse. Un farmaco invece ha una varietà maggiore di tipi di atomi, e di norma richiede un force field ottimizzato. L'acqua, spesso presente nella modellistica dei sistemi biologici per simulare la matrice extracellulare e come ambiente ove la proteina assume la sua conformazione strutturale, richiede un force field apposito, e ne sono stati sviluppati svariati (ad esempio TIP, SPC, BF, BNF, ST) (Spoel 1998), la cui differenza principale consiste nel numero di siti adottati (da 2 a 6) per descrivere una molecola d'acqua, e nell'ammettere flessibilità e polarizzazione. Una volta definiti gli oggetti fondamentali che interagiscono (i tipi atomici) ed i potenziali di interazione (i force fields), la DM non farà altro che risolvere numericamente le equazioni di Newton associate a questo sistema di corpi (esempi di software: NAMD, gromacs) (Kalé 1995, Berendsen 1995), e definirà una evoluzione temporale dell'intera struttura a partire da una configurazione iniziale. I problemi principali sono: i) la dimensione del sistema, ossia il numero totale di gradi di libertà; ii) le frequenze massime dei moti locali che si ritengono importanti per una corretta dinamica; iii) la scala temporale del cambiamento di configurazione globale che si intende studiare. Questi tre aspetti, congiuntamente alla valutazione delle risorse di calcolo ed al tempo massimo di

simulazione a disposizione, consentono di definire il passo temporale ottimale per la discretizzazione delle equazioni ed il tempo totale di simulazione. In maniera semplificata, ma piuttosto realistica, l'output di una simulazione di DM è costituito da un film che rappresenta l'evoluzione temporale della molecola, con ciascun frame preso ad uno specifico tempo di campionamento dell'evoluzione.

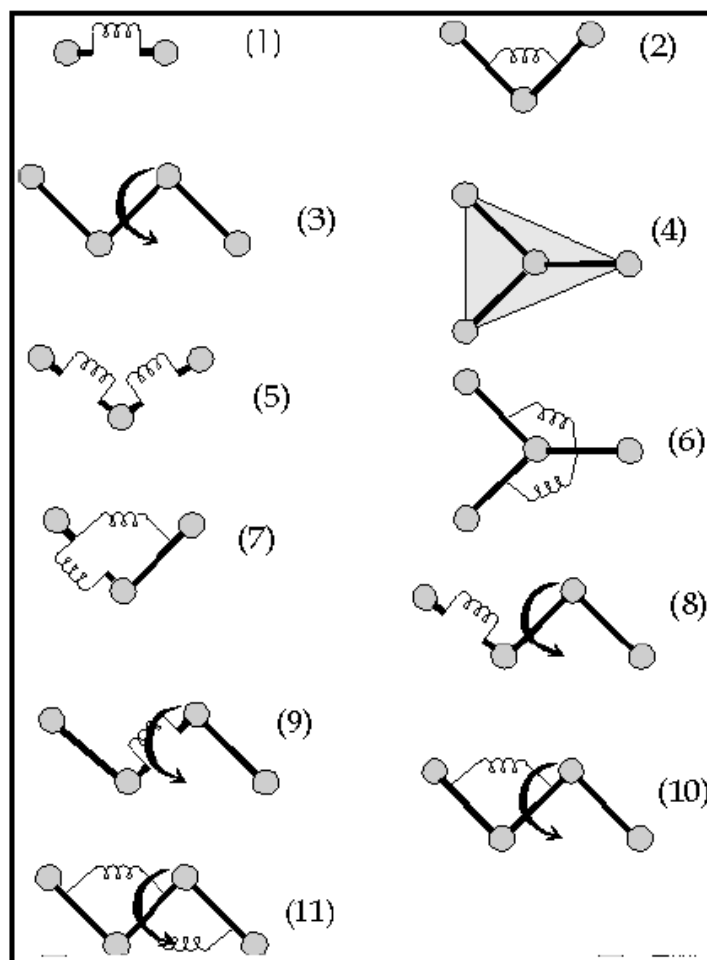


Figura 2.2. Diagramma con un esempio dei termini di interazione che sono tipicamente considerati per valutare analiticamente un force field, e che descrivono le varie tipologie di interazione dei tipi atomici. Ogni termine contiene dei parametri che vengono ottimizzati a seconda della classe di molecole o delle molecole stesse (immagine dal sito

edoc.hu-berlin.de/dissertationen/physik/severin-nikolai/HTML/severin-ch2.html)

L'analisi delle caratteristiche geometriche, fisiche e chimiche del sistema, mediate nel tempo o tracciate nel tempo, consente di stimare parametri biologici, fisici e chimici

importanti quali: energie di legame (nel caso di un complesso) e suoi contributi entropici ed entalpici, stabilità della struttura, interazione con le molecole d'acqua, flessibilità di specifici domini, cambiamenti di conformazione chimica, siti di legame, reti di interazione (legami idrogeno, interazione di stacking, interazioni idrofobiche), effetti strutturali di mutazioni di amminoacidi in una proteina. La dinamica molecolare è una tecnica molto efficace, in grado di fornire un quadro dinamico ad alta risoluzione delle molecole e delle loro interazioni (Andricioaei 2001, Glykos 2006, Naim 2007), contrariamente alla rappresentazione statica fornita dalla cristallografia. La DM è stata utilizzata per modellare molti sistemi biologici (Kumar 2010a, Kumar 2010b, Dror 2011, Balarju 2012), includendo cambiamenti di conformazione del sito di legame delle proteine MHC (Ghosh 1995, Cui 2008, Painter 2011), effetti di mutazione virtuale del peptide o della proteina MHC (Knapp 2010) e l'interazione fra legandi endogeni ed agonisti del recettore PAFR (Gui 2007).

Docking. Mediante la DM si considerano tutti gli atomi del sistema liberi di muoversi, e questo richiede una buona preparazione del sistema iniziale e lunghi tempi di simulazione. Esistono tecniche semplificate che consentono invece di valutare l'interazione fra una proteina ed una piccola molecola (legando) in tempi molto rapidi, anche in assenza di conoscenza del sito di legame (*blind mode*). Queste sono le tecniche di Docking (ad esempio Autodock, Zdock) (Morris 2009, Pierce 2011), nelle quali la proteina viene tenuta rigorosamente fissa ed il legando viene generato in tutti i suoi vari stati di conformazione chimica (*conformers*) e viene posizionato in tutti i modi possibili rispetto alla struttura proteica. Per ogni posa viene valutata l'energia del legame e le varie pose vengono così classificate. Oltre all'energia di legame, ottenuta tramite gli stessi force field della DM, sono state sviluppate moltissime altre tecniche di *scoring* delle pose, basate su metodologie di intelligenza artificiale e data mining, di norma ritenute più specializzate ed efficaci (Leach).

2.2. La proteina di membrana MHC

2.2.1. Struttura

Molte strutture tridimensionali di complessi peptide:MHC di classe II sono state risolte ad alta risoluzione (Stern 1994, Murthy 1997, Brown 1993, Smith 1998, Ghosh 1995, Gunther 2010, Madden 1995, Hammer 1993), fornendo dunque una solida base per la comprensione dell'interazione peptide:proteina e nel contempo indizi utili sul meccanismo di riconoscimento molecolare. Come si evince dalle strutture, la proteina MHC di classe II é un etero-dimero costituito da due catene A e B, ciascuna formata da due domini, uno trans-membrana ed uno extra-cellulare (Figura 2.3). La proteina presenta un sito di legame per peptidi, ossia corte catene di amminoacidi. Questo sito é costituito dal dominio extra-cellulare composto dai domini $\alpha 1$ e $\beta 1$ della catena A e B rispettivamente, che hanno la struttura secondaria di α -eliche. La caratterizzazione strutturale ha ulteriormente portato all'identificazione di specifici residui di ancoraggio del peptide sulla proteina, detti tasche (pockets, si veda la Figura 2.4 (A)), ed in corrispondenza di questi a specifici pattern preferenziali di amminoacidi di legame (binding motifs) sul peptide. In generale, gli amminoacidi di legame sul peptide sono localizzati nei residui 1, 4, 6 e 9, ed il tipo di amminoacido varia notevolmente a seconda dell'allele HLA (Hammer 1993, Rammensee 1995, Southwood 1998) Studi recenti di modellistica hanno utilizzato la variazione della mappa elettrostatica esterna al sito di legame, per distinguere le tasche di legame con proprietà di ancoraggio specifico (pocket P1 e P4) o di riconoscimento (pocket P4 e P7) (Agudelo 2009). Il sito di legame é una specie di solco con le estremità aperte (contrariamente alla classe I), lungo circa 40 Å, e con larghezza variabile (Figura 2.4 (A)).

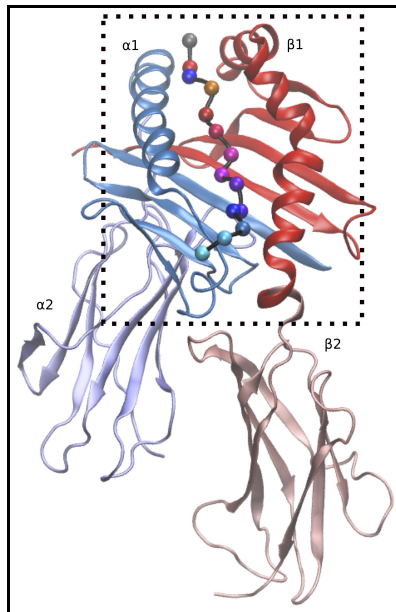


Figura 2.3. Proteina di membrana MHC di classe II. Sono evidenti i due domini intramembrana $\alpha 2$ e $\beta 2$, ed i due domini che costituiscono il sito di legame $\alpha 1$ e $\beta 1$. L'immagine presenta la configurazione tipica di presentazione di un peptide all'interno del sito di legame.

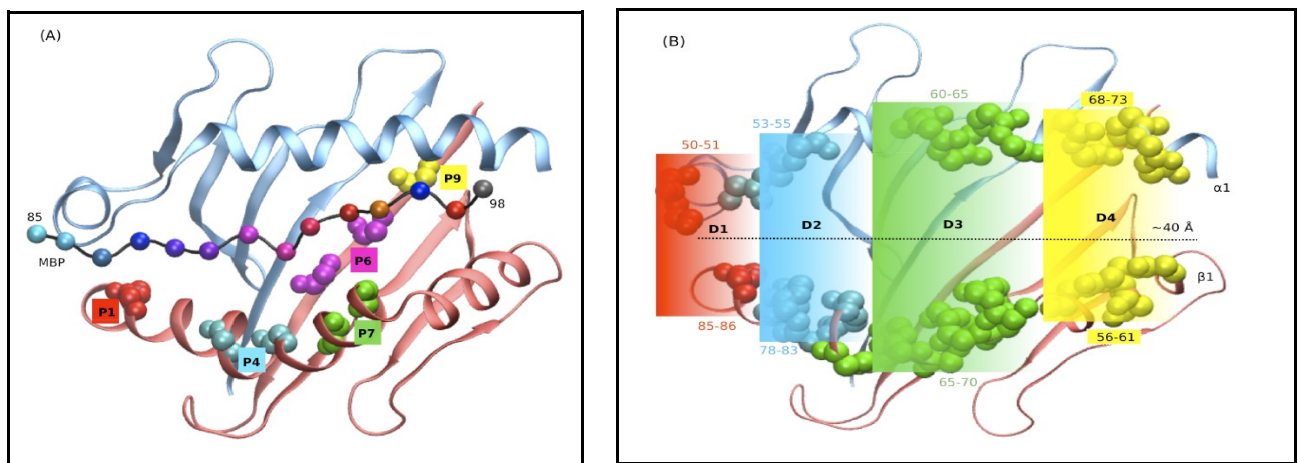


Figura 2.4. Le tasche di legame e le zone di flessibilità della proteina MHC. (A) P1, P4, P6, P7 e P9 sono le aree di superficie della proteina MHC che formano l'interazione di legame col peptide. Il numero rappresenta la posizione dell'amminoacido corrispondente sul peptide. (B) Suddivisione della proteina MHC in quattro aree, D1, D2, D3, D4, all'interno delle quali viene analizzata la variazione della larghezza del sito di legame, per valutare le caratteristiche di flessibilità ed ancoraggio col peptide.

Con lo scopo di analizzare le caratteristiche locali del sito di legame durante l'interazione col peptide, esso é stato diviso in quattro aree (D1, D2, D3, e D4), si veda la (Figura 2.4 (B)), corrispondenti alle aree di influenza delle quattro tasche P1, P4, P7 e P9, in maniera simile a quanto fatto precedentemente da altri autori (Yaneva 2009). In dettaglio, per gli alleli DQB1 le regioni sono così definite: D1(α 53-54, β 84-85); D2 (α 57-59, β 77-82); D3 (α 64-69, β 67-72) and D4 (α 72-77, β 56-61), e la larghezza di ciascuna regione viene valutata come la distanza dei centri di massa dei due fianchi che la definiscono.

2.2.2. L'interazione peptide:MHC e TCR è correlata alle caratteristiche dinamiche del complesso peptide:MHC. Una analisi della letteratura.

Una delle applicazioni discusse nel presente lavoro di tesi relativamente alla SM, è quella di correlare le caratteristiche locali di flessibilità del sistema peptide-proteina MHC alla capacità di legare uno specifico recettore (TCR) sulla membrana di un linfocita T ed alla eventuale conseguente attivazione del linfocita T. Dato che questo è un argomento piuttosto complesso e molto dibattuto (Knapp 2010, Adams 2011), si ritiene utile riportare una breve analisi della letteratura rilevante.

Recentemente, alcuni autori (Yaneva 2010) hanno evidenziato tramite simulazione un legame diretto fra stati flessibile della proteina MHC e capacità aumentate di legame del peptide. Essi hanno inoltre identificato un interruttore sensibile al pH nel sito α 1 His 33, un residuo vicino alla P1 che interagisce con α 2 Val 136: a pH neutro o elevato, l'interazione di questi due residui vincola i movimenti relativi delle due α eliche, bloccando i cambiamenti di conformazione della proteina MHC. Gli autori riportano inoltre come le regioni più flessibili in assenza del peptide siano le α 51-54 e α 58-70. In un altro studio sperimentale (Sadegh-Nasseri 2011), gli autori propongono in maniera simile che uno stato flessibile del complesso peptide:MHC

(denotato per brevità pMHC o p:MHC) possa essere convertito in una forma attiva dall'azione della proteina catalizzante DM, mentre tale conversione risulta inibita per complessi pMHC maggiormente rigidi e compatti. Altri lavori (Rupp 2011) evidenziano la chiusura del sito di legame in una proteina MHC, dovuta al legame idrogeno stabilizzante fra i residui Asn 82 e Gln 9 per ridurre l'esposizione al solvente nella proteina MHC non caricata con un peptide. Uno studio precedente (Painter 2008) aveva già descritto il ruolo funzionale della flessibilità differenziale lungo il sito di legame, particolarmente della regione α 50-59, che nella proteina MHC libera si muove per occupare i siti riservati al legame col peptide (P1 e P4), saturando quasi tutti i legami idrogeno. Sono infine state effettuate anche delle analisi entropiche-entalpiche del processo di legame peptide e proteina (Ferrante 2012), sia per spiegare le proprietà di flessibilità richieste prima del legame che i cambi di configurazione della proteina MHC che consentono di arrivare a complessi finali p:MHC stabili.

Per quanto riguarda l'interazione col linfocita T, i modelli più accreditati di legame pMHC e TCR (Mazza 2007, Gagnon 2006, Armstrong 2008, Harkioliaki 2009, Scott 2011, Adams 2011, Holland 2012), sono basati su una plasticità conformazionale sia del TCR che del complesso pMHC, e richiedono dunque un certo livello flessibilità del complesso pMHC per giungere ad un riconoscimento da parte del TCR e consentire i necessari cambiamenti conformazionali in seguito al legame col TCR (Csermely 2010, Fenwick 2011). Recentemente alcuni autori (Call 2011) hanno osservato che le caratteristiche dell'interazione pMHC:TCR portano la memoria dell'ambiente dove il peptide è stato reclutato, includendo sia le modalità di scansione extra-cellulare del recettore TCR sul linfocita T, che l'eventuale competizione con altri peptidi. In particolare, gli autori hanno evidenziato l'importanza di una interazione p:MHC sufficientemente prolungata, e dunque stabile, così da consentire l'emergere di superfici di interazione uniche e durature.

Gli stessi autori hanno rivisto i meccanismi diversi attraverso i quali piccole molecole possono influenzare la risposta immunitaria, ad esempio modulando la stabilità del sistema p:MHC, e suggerendo un meccanismo d'azione, indipendente dalla proteina DM, ma che invece perturbi la rete dei legami idrogeno nativa del complesso p:MHC. Recentemente, sono state anche analizzate le caratteristiche di flessibilità del peptide per alcune molecole MHC di classe I, ed il loro impatto sulle modalità di scansione del complesso pMHC da parte del TCR e sul seguente meccanismo di attivazione (Reboul 2012). Sebbene questo lavoro sia relativo alla sola classe I, gli autori enfatizzano il ruolo, di carattere generale, dei pattern di interazione dinamici e complementari alla superficie fra TCR e complesso pMHC, per poter giungere con successo ad un complesso ternario finale.

Complessi p:MHC con emivita lunga consentono di esporre il peptide antigenico per un tempo prolungato, fattore che potrebbe essere critico per il lento processo di reclutamento di un linfocita T reattivo per quel pMHC specifico (Berg 2007). Ciò nondimeno, sono stati proposti dei meccanismi di ingaggio sequenziali multipli (Jansson 2011), che richiedono complessi ad emivita più breve, per consentire allo stesso peptide di ingaggiare molti linfociti T sequenzialmente. In entrambi i casi, viene comunque richiesta una conformazione flessibile ottimale del complesso pMHC al fine di un riconoscimento efficiente da parte del TCR (Armstrong 2008, Scott 2011, Borbulevych 2009, Reboul 2012).

***2.2.3. Protocollo SM-P: preparazione del sistema peptide:MHC* ✨**

Viene qui presentato il protocollo sviluppato per la preparazione dei complessi peptide e proteina MHC utilizzati nello studio del meccanismo di presentazione del peptide antigenico nel contesto della Sclerosi Multipla. I peptidi utilizzati sono due

(si veda la Figura 2.5): uno virale proveniente dalla proteina EBNA-1 del virus EBV (residui 400-413), ed uno mielinico proveniente dalla auto-proteina MBP (residui 85-98).

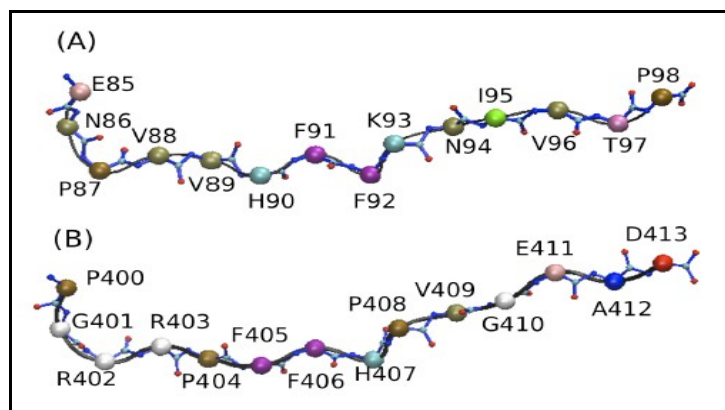


Figura 2.5. Sequenza e conformazione tipica dei due peptidi studiati all'interno del sito di legame della proteina MHC. (A) MBP 85-98, sequenza ENPVVNFFKNIVTP e (B) EBNA-1 400-413, sequenza PGRRPFFHPVGEAD.

DQB1*05:02 (protettivo). Le strutture di partenza per la proteina MHC ed il complesso MHC:MBP sono state estratte dal database PDB(www.rcsb.org/), numero di accesso 3pl6, relativo al complesso ternario MHC:MBP:TCR (Figura 2.6).

DQB1*06:02 (predisponente). La struttura di partenza per la sola proteina MHC é stata estratta dal database PDB, numero di accesso 1uvq, relativo al complesso MHC con un peptide della ipocreatina (Figura 2.7).

DRB1*15:01 (predisponente). Le strutture di partenza per la proteina MHC ed il complesso MHC:MBP sono state estratte dal database PDB, numero di accesso 1bx2, relativo proprio al complesso con MBP (Figura 2.8).

DRB1*16:01. L'allele protettivo DRB1*16:01 non é stato strutturalmente caratterizzato, pertanto si é utilizzato come template per la modellistica per omologia l'allele DRB1*01:01 che condivide col target una similarità di sequenza del 94%. Si é utilizzata la struttura presente nel database PDB, numero di accesso 3pdo, relativa al complesso del DRB1*01:01 col peptide CLIP (Figura 2.9). La qualità del modello, ottenuto tramite il software modeler (Eswar 2006), é stata convalidata

tramite grafico di Ramachandran, mostrando come oltre il 92% dei residui fosse nella regione consentita dello spazio delle fasi dei gradi di libertà angolari del backbone peptidico ed oltre il 7% nella regione di tolleranza (PROCHECK e SwissPDBViewer, www.ebi.ac.uk/thornton-srv/software/PROCHECK/).



Figura 2.6. MHC:MBP:TCR (3pl6) utilizzato per la struttura del DQB1*05:02.

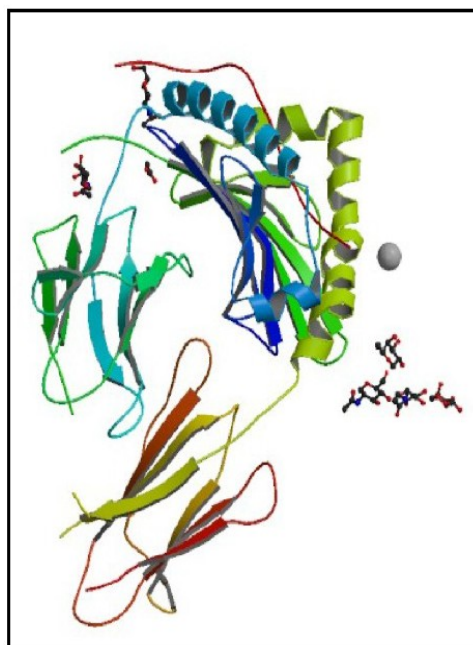


Figura 2.7. MHC:ipocreatina (1uvq) utilizzato per la struttura del DQB1*06:02.



Figura 2.8. MHC:MBP (1bx2) utilizzato per la struttura del DRB1*15:01.

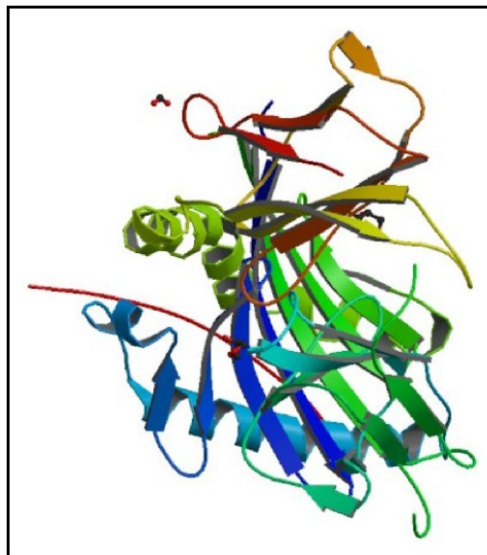


Figura 2.9. DRB1*01:01 in complesso con il CLIP (3pdo), utilizzata per il modello per omologia del DRB1*16:01.

I modelli per i complessi DQB1*05:02 : EBNA-1, DQB1*06:02 : MBP, DQB1*06:02 : EBNA-1, DRB1*15:01 : EBNA-1 sono stati ottenuti dai template disponibili per la proteina MHC o per il complesso MHC : MBP, utilizzando il server web MODPROPEP (Kumar 2007). Le configurazioni dei peptidi all'interno del sito di

legame sono state ulteriormente verificati tramite il software di docking AUTO-DOCK (Morris 2009).

Il modello per il complesso DRB1*1601 : EBNA-1 é stato direttamente ottenuto tramite AUTO-DOCK.

Le mutazioni di singolo amminoacido alla posizione DQB1 38 e DRB1 86 degli alleli in complesso coi due peptidi, sono state ottenute tramite il modulo Mutator disponibile nel software VMD. Gli atomi di idrogeno mancanti nei modelli finali sono stati aggiunti attraverso il software VMD (Humphrey 1996). I sistemi sono stati successivamente posti in una scatola riempita con molecole d'acqua e gli opportuni contro-ioni sono stati aggiunti al fine di neutralizzare il sistema (Figura 2.10).

2.2.4. Protocollo SM-D: DM dei complessi peptide:MHC ✱

Viene qui presentato il protocollo sviluppato per la DM dei complessi peptide e proteina MHC utilizzati nello studio del meccanismo di presentazione del peptide antigenico nel contesto della Sclerosi Multipla. L'elaborazione é stata effettuata tramite software NAMD su cluster di calcolo multi-processore del CRS4.

Dapprima, ciascun complesso peptide:MHC ha subito una minimizzazione energetica, seguita da un graduale riscaldamento del sistema a passi da 30 K, sino alla temperatura finale di 310 K, mantenendo nel contempo vincolate le posizioni degli atomi di Carbonio del backbone peptidico (50kcal/(mol A)) per 0.2 ns. Successivamente tali vincoli sono stati rilasciati a passi di 10kcal/(mol A) durante una simulazione per un tempo complessivo di 0.3 ns. Ciascun sistema é stato equilibrato tramite una simulazione di 3 ns, per poi dare luogo ai run di produzione nell'ensemble statistico NPT (ossia mantenendo costante numero di atomi, pressione e temperatura) a 310 K ed 1 atm, con condizioni al contorno periodiche, di lunghezza temporale adeguata a seconda dell'allele. Per gli alleli DRB1 sono state effettuati

dapprima runs di produzione di 120 ns. In seguito, avendo verificato nel caso del DRB1 che il tempo era più che sufficiente per campionare le caratteristiche di interesse del sistema, per gli alleli DQB1 sono stati effettuati runs di soli 100 ns.

Per la proteina ed i contro-ioni si è utilizzato il force field Charmm27 (Foloppe 2000) ed AMBER-99 (Salomon 2013), mentre per le molecole d'acqua è stato scelto il force field TIP3P.

Gli stati di carica degli amminoacidi sono stati assegnati in modalità standard, ad eccezione del residuo Asp 66, presente nella catena A di alcune proteine MHC, dato che si tratta di un residuo chiave in grado di condizionare pesantemente l'intera dinamica, e che è stato pertanto trattato come recentemente proposto (Yaneva 2009).

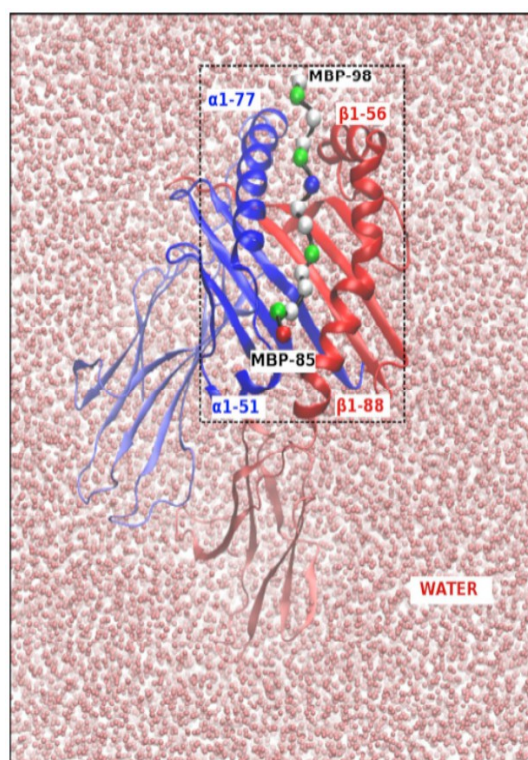


Figura 2.10. Tipico sistema peptide:MHC immerso in acqua.

La dimensione iniziale della box di simulazione per i sistemi DQB1 e DRB1 è pari, rispettivamente, a [80 Å, 96 Å, 74 Å] e [77 Å, 75 Å, 96 Å], con un numero totale di atomi in ciascun sistema pari a ~50 000 (Figura 2.10). Tutti i legami con gli atomi di idrogeno sono stati vincolati utilizzando SHAKE (Ryckaert 1977), il che ha consentito un tempo di integrazione pari a 2 fs. Le interazioni a lungo raggio sono state valutate

tramite metodo PME (Ewald 1921, Darden 1999), con una griglia di dimensioni [96 Å, 96 Å, 96 Å]. Le interazioni di Van der Waals ed elettrostatiche sono state calcolate sino ad una distanza di 10 Å.

2.2.5. Protocollo SM-A: analisi delle simulazioni dell'interazione peptide:MHC ✨

- 1) La stabilità del sistema p:MHC è stata verificata tramite il software VMD, valutando l'andamento temporale della root mean square deviation (RMSD) degli atomi di Carbonio del backbone peptidico per un insieme di residui selezionati;
- 2) la flessibilità del sito di legame della proteina MHC è stata analizzata dividendo il sito di legame peptidico in quattro regioni, come descritto nella Sezione 2.2.1 (Yaneva 2009) e valutando in ciascuna regione l'istogramma delle fluttuazioni della larghezza;
- 3) Le interazioni idrogeno fra i residui del peptide e del sito di legame sono state valutate utilizzando una distanza Donatore-Acettore inferiore a 3.1 Å e un angolo di cutoff di 130°;
- 4) Le interazioni di stacking aromatiche (McGaughey 1998) sono state calcolate utilizzando il software EUCEB (Tsoulos 2011) con un angolo di cutoff diedrico di 30° fra le catene laterali planari e quelle con anelli, cutoff di distanza centroide fra catene laterali di 5.0 Å, ed una durata minima di 20% del tempo totale di simulazione;
- 5) L'energia di legame per il complesso p:MHC è stata valutata utilizzando il metodo SIE (Solvated Interaction Energy) (Sulea 2012). Il requisito per poter utilizzare questo approccio è l'utilizzo del force field Amber 99, pertanto è stato necessario integrare le simulazioni già svolte con ulteriori 30 ns con questo specifico force field; l'energia di legame è stata valutata con passi temporali di 20 ps. Questa duplicazione delle simulazioni è stata utilizzata anche per verificare che i risultati fossero indipendenti da piccole variazioni dei parametri del force field.

2.3. Il sistema PAF:PAFR

2.3.1. Struttura

Il fattore di attivazione delle piastrine (Platelet-Activating Factor, PAF, si veda la Figura 2.11) é un fosfolipide in grado di attivare e mediare svariate funzioni dei leucociti, quali la degranulazione e l'aggregazione delle piastrine, i processi di infiammazione, le reazioni allergiche gravi, il rigetto dei tessuti, il danno ischemico, la permeabilità vascolare, lo stress ossidativo cellulare, la chemiotassi dei leucociti e il metabolismo dell'acido arachidonico (Montrucchio 2000). Il PAF viene espresso da svariati tipi di cellule, particolarmente quelle specializzate nelle difese immunitarie, quali piastrine, cellule endoteliali, neutrofilo, monociti e macrofagi. Data la sua versatilità e potenza, il livello ematico del PAF viene finemente regolato a livelli molto bassi, ma si é osservato un suo incremento significativo in risposta a specifici stimoli pro-infiammatori. Alcuni studi farmacologici hanno dimostrato come alcuni dei suoi antagonisti possano essere utilizzati per la cura delle malattie mediate dal PAF. Il PAF si lega ad un recettore di membrana, il PAFR, che appartiene alla vasta famiglia dei recettori accoppiati alla proteina G (GPCR) (Venkatakrishnan 2013), caratterizzati dalla presenza di sette eliche trans-membrana che subiscono cambiamenti conformazionali in seguito al legame di una molecola esterna, e avviano così la catena della segnalazione cellulare all'interno della cellula. Molti degli effetti del PAF vengono in tal modo mediati dal legame col PAFR, e dunque il PAFR rappresenta un bersaglio farmacologico interessante (Iovino 2013, Grigg 2012, Edwards 2009, Kasperska-Zajac 2008, Chen 2008). Sfortunatamente la struttura del PAFR non é stata risolta, pertanto, di recente, alcuni autori (Gui 2007) hanno proposto un primo modello del PAFR, realizzato per omologia col template strutturale della rodopsina bovina (PDB access id 1L9H), che appartiene anch'essa

alla stessa sottofamiglia dei GPCR del PAFR. Come si vede dalla Figura 2.12, il recettore PAFR é costituito da sette domini transmembrana aventi struttura secondaria di α -eliche, ed identificati dal numero romano da I a VII, un dominio N-terminale extracellulare, un dominio C-terminale intracellulare, tre loops extracellulari (E1, E2, E3) agenti come sensori ambientali e tre loops intracellulari (I1, I2, I3) agenti come attuatori della segnalazione.

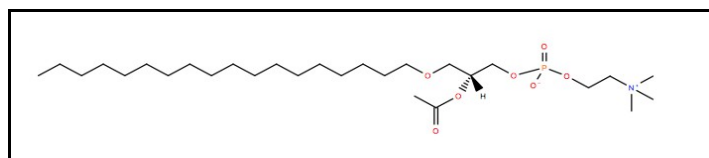


Figura 2.11. Struttura chimica del PAF, o fosfocolina PC(O-18:0/2:0), ottenuta dalla banca dati LIPIDOMICS (www.lipidmaps.org/), id LMGP01020094.

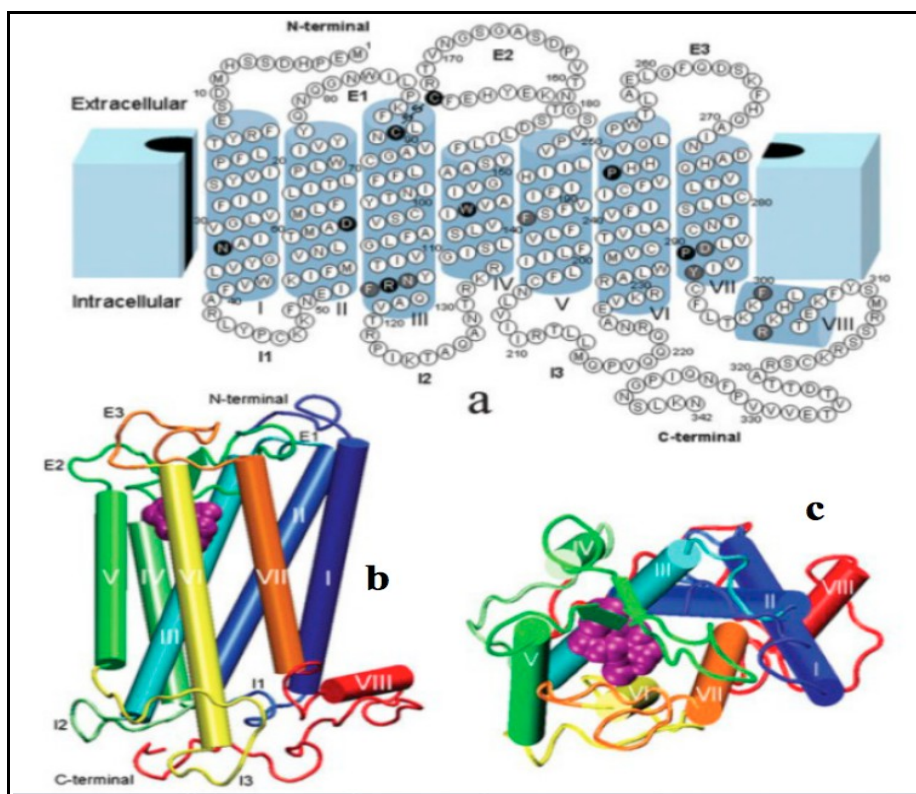


Figura 2.12. Il PAFR. (a) Struttura secondaria, che evidenzia i domini intracellulari (I1, I2, I3), quelli intramembrana (I, II, ..., VII), quelli extracellulari (E1, E2, E3) e le due parti(N- e C) terminali. (b,c) modello della struttura 3D (Gui 2007), e posizione tipica di legame del PAF.

2.3.2. Protocollo FMS-P: preparazione del sistema legando:PAFR ✱

La struttura iniziale per il recettore PAFR (Platelet Activating Factor Receptor) è stata ottenuta da un modello per omologia realizzato da altri autori (Gui 2007).

La struttura dei due legandi lipidici esogeni, PC(14:0) (Figura 2.13) e PC(16:0) (Figura 2.14), e del legando endogeno fisiologico del PAFR, il PAF, sono state scaricate dal database LIPIDOMICS (www.lipidmaps.org/) e modellate tramite il web server CORINA (Corina). Come si vede dalle Figure 2.11, 2.13, 2.14, tutti e tre i legandi hanno lunghe catene lipidiche ed una testa polare. Le strutture dei legandi sono state successivamente sottoposte ad ottimizzazione geometrica utilizzando il metodo di Hartree-Fock HF-6-31G* disponibile all'interno del pacchetto Gaussian (Frisch 2004). Le cariche ed i parametri dei force fields per i legandi sono state poi ottenute tramite il protocollo standard integrato nel pacchetto AMBER (Salomon 2013).

Le strutture dei legandi così ottimizzate sono state infine soggette a docking col recettore PAFR tramite il software Autodock (Morris 2009), al fine di ottenere i siti di legame senza alcun *a priori* sulla loro localizzazione e identificare le migliori configurazioni di legame (pose) da utilizzare nelle dinamiche molecolari.

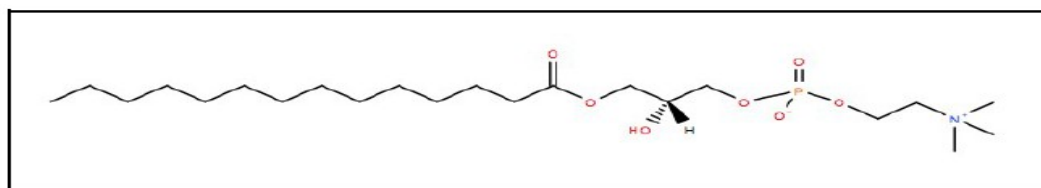


Figura 2.13. Struttura della lisofosfocolina PC(14:0/0:0), estratta dal database LIPIDOMICS con id LMGP01050012.

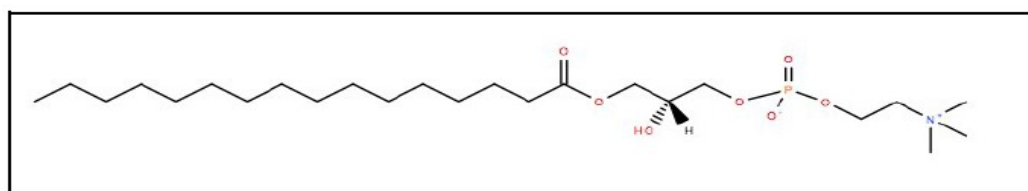


Figura 2.14. Struttura della lisofosfocolina PC(16:0/0:0), estratta dal database LIPIDOMICS con id LMGP01050018.

2.3.3. Protocollo FMS-D: DM dei complessi legando:PAFR ✳

Prima della simulazione di DM, ciascun complesso identificato dalla miglior posa offerta da Autodock, è stato integrato in un bistrato lipidico POPC (Kumar 2010) così da simulare la membrana cellulare, e successivamente inserito in una scatola d'acqua (force field TIP3P). Dopo una simulazione di 3 ns per consentire al sistema di assumere una configurazione energetica bassa ed abbastanza stabile (rilassamento), sono iniziati i runs di produzione di 50 ns per ciascun sistema. Le simulazioni hanno come al solito utilizzato il software NAMD con force field AMBER 99 (Phillips 2005). I parametri generali della simulazione di DM per i complessi ed il modello d'acqua sono gli stessi utilizzati per il caso della Sclerosi Multipla (Sezione 2.2.4).

2.3.4. Protocollo FMS-A: Analisi della DM dei complessi legando:PAFR ✳

L'analisi delle dinamiche dei complessi legando:PAFR si è in particolare focalizzata sulla valutazione di:

- 1) RMSD degli atomi pesanti dei legandi, così da valutare la stabilità del legame ed i cambiamenti di conformazione reciproca recettore-legando durante la simulazione;
- 2) l'esistenza di reti stabili (almeno il 10% del tempo di simulazione) di interazione tramite legame idrogeno, che forniscono un contributo significativo all'energia di legame complessiva;
- 3) le energie di legame per ciascun complesso, valutata tramite il software SIETRAJ (Naim 2007).

2.4. La proteina CASQ1

2.4.1. Struttura

La Calsequestrina è una proteina in grado di immagazzinare ioni di Calcio all'interno delle cellule dei tessuti muscolari, rilasciandoli per attivare la contrazione muscolare. Ciascuna proteina può legarsi ad un enorme numero di ioni Calcio, con un meccanismo ancora poco conosciuto, che è stato proposto sia basato sulla neutralizzazione della superficie polare della proteina. Sono state identificate due isoforme della proteina: la CASQ1, coinvolta nel funzionamento dei muscoli rapidi, e la CASQ2, coinvolta nel funzionamento del cuore. La proteina presenta per circa metà della sua struttura una configurazione casuale, ed il legame degli ioni Calcio induce una forma più compatta e maggiormente strutturata. La proteina viene attivata tramite processo di fosforilazione operato dalla caseina-chinasi 2 (Wang 1998). La struttura cristallografica della proteina CASQ1 (Sanchez 2012)(Figura 2.15) mostra come il monomero sia composto da tre domini globulari α - β , formati da cinque foglietti β disposti planarmente ed alternati a quattro α eliche. Il dominio C-terminale è formato da ripetizioni di residui di acido aspartico (CAS, consecutive aspartate stretch), la cui lunghezza varia fra le diverse specie, ed pari a circa 9-14 amminoacidi per i mammiferi. Alcuni studi biofisici mostrano come il dominio CAS, pur mancando di domini strutturali di legame noti, sia in grado di legare un enorme numero di ioni con bassa affinità (Maurer 1985, Cozens 1984, Aaron 1984, Ikemoto 1989, Park 2004). La natura del legame non è stata ancora sufficientemente chiarita, ed è stato proposto che il meccanismo di legame sia dovuto a puri fenomeni superficiali, governati dalla elevata carica negativa della superficie molecolare (Park 2003, 2004). La forma monomerica della proteina è in grado di formare reti di interazione proteina-proteina dando luogo alla realizzazione di omo-polimeri, solo alcuni dei quali con struttura nota, al fine di ottimizzare l'immagazzinamento ed il

rilascio degli ioni Calcio. In dettaglio, è stato proposto che il processo di polimerizzazione sia modulato dalla concentrazione degli ioni Calcio (Wang 1998, Park 2003, Park 2004, Aaron 1984, Kim 2007), tramite un meccanismo progressivo da monomeri a dimeri, e poi da dimeri a tetrameri. Il dominio CAS è localizzato all'interfaccia fra i monomeri nel dimerico, e dunque può influenzare la polimerizzazione, ma il suo ruolo è poco compreso poiché la sua struttura non è mai stata risolta. In generale l'assenza di informazione strutturale delle interazioni in gioco nella formazione dei polimeri CASQ1, consente solo speculazioni sui meccanismi coinvolti e sulla loro dipendenza dalla concentrazione di ioni Calcio.

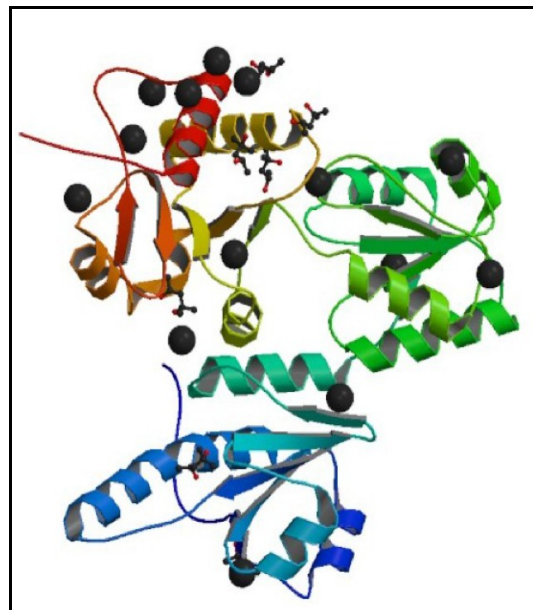


Figura 2.15. Struttura del complesso degli ioni Calcio e proteina calsequestrina umana di tipo 1, ritrovata nei muscoli scheletrici. La struttura monomerica è stata estrapolata dall'esamero risolto (PDB id 3uom, Sanchez 2012). .

2.4.2. Protocollo CASQ1-P: Preparazione della proteina CASQ1 ✳

Il modello per la proteina è stato realizzato a partire dalla struttura esamerica disponibile sul database PDB (id 3uom) (Sanchez 2012), composta dalla ripetizione

di tre dimeri (Figura 2.15), ma mancante delle catene C-terminali (residui 350-364 di ciascun monomero). Tramite il software di visualizzazione e manipolazione VMD è stato isolato un singolo dimero. Successivamente, il dimero così ottenuto è stato sottoposto ad una analisi di coerenza chimica strutturale ed integrato di tutti gli aspetti assenti, incompleti o richiesti dalla DM, utilizzando sia VMD che la suite Schrodinger (www.schrodinger.com). In dettaglio: sono state aggiunte le due catene C-terminali mancanti e gli altri elementi chimici assenti a causa della metodologia sperimentale (in particolare: gli atomi idrogeno ed alcune catene laterali); sono stati definiti gli ordini di legame e le cariche degli atomi, ed ottimizzati gli orientamenti dei gruppi simmetrici. Lo stato di carica dell'amminoacido His è stato trattato ottimizzato in maniera non standard (Sezione 2.2.4). Alla fine è stato aggiunto un numero variabile di ioni Calcio, Ca^{++} , corrispondente alla concentrazione di interesse.

2.4.3. Protocollo CASQ1-D: DM del sistema CASQ1 - Ca^{++} ✱

Per la proteina e gli ioni si è utilizzato il force field CHARMM27 (MacKerell 2000). Il sistema dimero e ioni Calcio è stato quindi inserito in una scatola d'acqua, aggiungendo contro-ioni addizionali per neutralizzare il sistema (Humphrey 1996). Le molecole d'acqua sono descritte, come per le proteine MHC e PAFR, dal force field TIP3P (Jorgensen 1983). Gli stati di carica dei residui sono stati assegnati utilizzando il software propKa (Li 2005). Ciascun sistema è stato quindi sottoposto al solito protocollo presentato per MHC e PAFR (Sezioni 2.2.4 e 2.3.3): i) minimizzazione dell'energia, ii) riscaldamento a 300 K con vincoli posizionali di 50 kcal/(mol Å²) per 0.2 ns, iii) successivo rilascio dei vincoli per 0.3 ns ed iv) equilibratura per 3 ns. Sono stati poi fatte le simulazioni per 30 ns a 300 K ed 1 atm, con condizioni periodiche. La dimensione iniziale della scatola di simulazione è di [138 Å, 110 Å, 95 Å], per un numero totale di ~130.000 atomi. Le interazioni

elettrostatiche a lungo raggio e le forze di Van der Waals sono state trattate come precedentemente descritto, Sezione 2.2.4, con una griglia di [148 Å 128 Å 96Å] e cut-off di 12 Å.

2.4.4. Protocollo CASQ1-A: Analisi del sistema CASQ1 - Ca⁺⁺ ✨

Come al solito, la stabilità della proteina durante la DM è stata verificata tramite valutazione della RMSD degli atomi pesanti del backbone peptidico. Tramite il software CARMA è stata anche valutata l'entropia configurazionale del dominio CAS, a partire dalla matrice di covarianza delle fluttuazioni degli atomi pesanti con passi temporali di 200 ps (Andricioaei 2001, Glykos 2006, Balaraju 2012). I residui identificativi di ciascun sito di legame sono stati identificati dalla permanenza degli ioni di Calcio durante la simulazione entro una distanza di 5 Å dai residui stessi. Per ciascun sito, è stata valutata l'energia di interazione elettrostatica e di Van der Waals fra gli ioni Calcio ed i residui del sito (con una distanza di cutoff di 12 Å). In maniera simile si è analizzato il ruolo dell'acqua nelle interazioni dirette e nella mediazioni di interazioni, per la proteina intera, con una distanza di cut-off di 3.0 Å. Le interazioni idrogeno ed idrofobiche sono state valutate tramite VMD sulle strutture medie della proteina calcolate ogni 2 ns. Infine le reti di interazione sono state visualizzate tramite il software LIGOPLOT (Wallace 1995).

3. Modellistica della presentazione antigenica, applicata alla Sclerosi

Multipla

La sclerosi multipla (SM) è una malattia autoimmune che colpisce il sistema nervoso centrale, con una incidenza in Sardegna maggiore di quanto ci si aspetterebbe dagli studi di distribuzione geografica su scala mondiale. Nella SM il sistema immunitario viene in qualche modo fuorviato, e le sue cellule specializzate attivano un meccanismo di reazione contro le proteine che compongono la mielina, che costituisce il rivestimento dei nervi, compromettendo la trasmissione degli impulsi nervosi. Ad oggi, si ritiene che la malattia sia dovuta ad una predisposizione genetica sulla quale incide un fattore scatenante, attualmente sconosciuto, di origine ambientale. Gli studi di associazione genetica, sia con approccio a gene candidato, che sull'intero genoma (GWAS), hanno confermato la regione genetica nota come HLA come il principale determinante genetico. Fra i vari fattori ambientali studiati, hanno un posto di rilievo alcuni virus e batteri, in particolare il virus di Epstein-Barr (EBV). L'approccio modellistico adottato per la sclerosi multipla consiste nel valutare le caratteristiche dinamiche, a livello molecolare, della presentazione di peptidi effettuata dai vari alplotipi HLA predisponenti e protettivi presenti nella popolazione sarda. Sono stati in particolare considerati peptidi provenienti da proteine del proprio corpo, specificatamente da alcune proteine mieliniche, e provenienti da agenti patogeni ambientali, in particolare da proteine del virus EBV.

3.1. La Sclerosi Multipla

La sclerosi multipla è una malattia infiammatoria cronica, caratterizzata dalla perdita di tolleranza del sistema immunitario verso le proteine del proprio corpo (*self*), con l'effetto di una demielinizzazione degli assoni del sistema nervoso centrale (Oksenberg 2008). La ricerca ha prodotto evidenze crescenti che la malattia sia causata dall'interazione di un profilo genetico predisponente ed una esposizione a fattori ambientali, che intervengono con una precisa cadenza temporale ma sono largamente sconosciuti (Baranzini 2011). Questi aspetti formano un quadro complesso e multi-fattoriale, formato dalla convergenza di molti elementi, ciascuno con un peso relativamente basso. Sin dai primi studi genetici, la regione genetica nota come HLA (Human Leukocyte Antigen), particolarmente la classe II, è stata identificata come la principale determinante genetica della SM (Ramagopalan 2009). Successivi studi a livello di genoma hanno confermato questi risultati con una accuratezza maggiore, nel contempo rivelando anche nuovi geni collegati alla MS (IMSG 2011). tutti gli studi di natura genetica hanno un valore principalmente statistico, con delle limitazioni dovute alla dimensione del genoma campionato, al numero dei marcatori utilizzato ed agli effetti di disequilibrio di linkage del DNA, ed infine non tengono in considerazione effetti epigenetici o post-translazionali. Pertanto, è importante integrare questi risultati con metodologie che consentano di studiare dinamicamente nel tempo i meccanismi funzionali della malattia, considerando sia il piano genetico che quello proteico (Fugger 2009).

3.2. Accenni al funzionamento del sistema immunitario: la presentazione degli antigeni da parte della glicoproteina di membrana MHC

Tutti gli agenti patogeni stranieri (*non-self*) presenti nel nostro corpo e le proteine stesse dei nostri tessuti (*self*), vengono continuamente elaborate e demolite da

specifici enzimi della cellula, ottenendo in tal modo piccoli frammenti lineari composti da pochi amminoacidi, detti peptidi. Questi peptidi si legano successivamente a specifiche glicoproteine di membrana, appartenenti al sistema maggiore di istocompatibilità (MHC), codificato dalla regione genetica HLA, per venire esposti alla matrice extracellulare da cellule specializzate nella presentazione antigenica (APC, particolarmente macrofagi, cellule dendritiche e linfociti B) (Kuby Immunology). Il sistema HLA si divide in classe I e classe II. In linea di massima, le proteine di classe I presentano antigeni derivanti da agenti patogeni già all'interno della cellula (pathway endogeno, attivo ad esempio per i virus, per i quali la replicazione avviene nel nucleo cellulare), mentre le proteine di classe II sono specializzate per agenti patogeni o proteine *self* che penetrano dall'esterno della cellula (pathway esogeno, ad esempio per i batteri) (Figura 3.1) (Kuby Immunology).

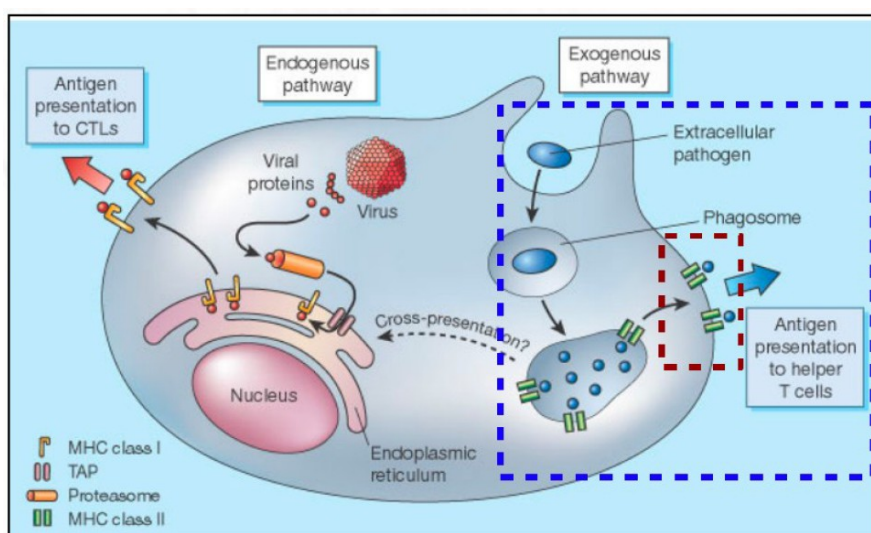


Figura 3.1. Pathway esogeno (tramite MHC-II) ed esogeno (MHC-I) per la presentazione di peptidi antigenici da parte delle cellule APC (Craig 2003).

A questo punto, in seguito alla presentazione del peptide, intervengono altre cellule specializzate del sistema immunitario, i linfociti T, che sono dotati di specifici recettori di membrana (TCR) in grado di legare il complesso peptide-MHC (p:MHC) (Figura 3.2). A seconda che si tratti di proteina MHC della classe I o II, vengono

reclutate due tipologie distinte di linfociti T, note rispettivamente come linfociti T citotossici (CTL) o aiutanti (Th) (Kuby Immunology).

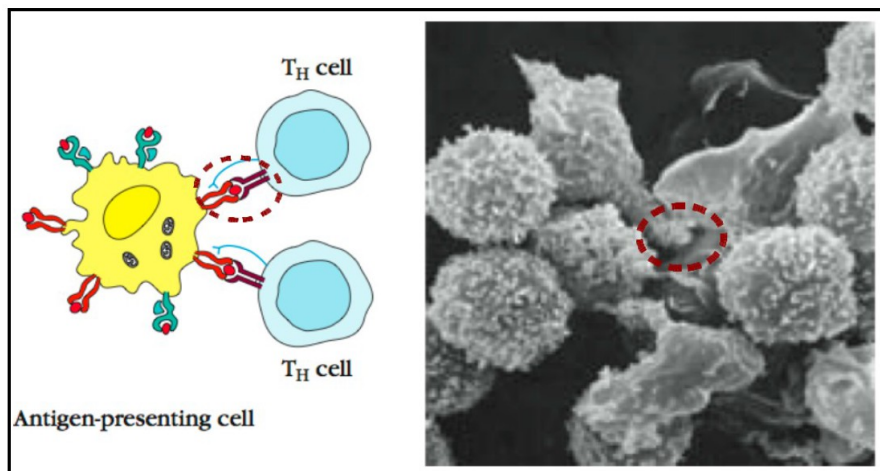


Figura 3.2. Presentazione di peptidi antigenici da parte delle cellule APC verso i linfociti T.

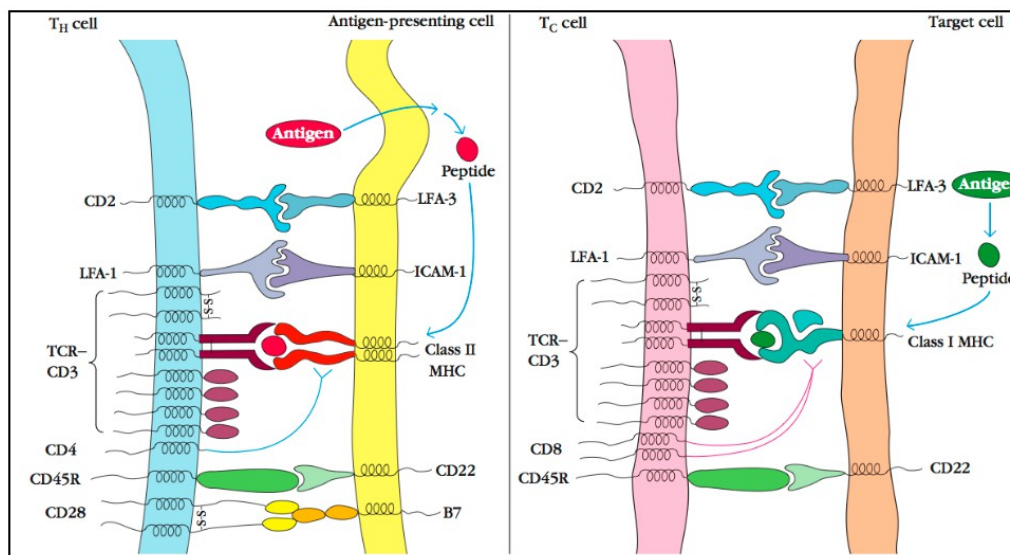


Figura 3.3. A seconda che si tratti della presentazione dell'antigene da parte della classe I (destra) o II (sinistra), si ha una interazione APC e TCR all'interno di uno specifico contesto di recettori ausiliari, col risultato di una attivazione del linfocita T in direzioni diverse (CTL o Th).

In sintesi: nella tipica risposta del sistema immunitario adattativo, gli antigeni derivanti da agenti patogeni verranno quindi riconosciuti dai linfociti T tramite il legame fra TCR e p:MHC, nel contesto di alcune molecole ausiliarie sulla membrana

sia del linfocita T che della APC, aventi funzione di co-stimolo, col risultato finale che il linfocita T verrà dunque attivato in maniera specifica (Figura 3.3).

In questa fase, all'interno della cellula del linfocita T attivato, inizia un complesso processo di arruolamento di chinasi citoplasmatiche, con il conseguente rilascio a cascata di una varietà di segnali chimici (citochine), producendo un pathway di segnalazione mediato da diverse proteine, sino agli ultimi fattori di trascrizione che penetrano nel nucleo cellulare e definiscono lo sviluppo ulteriore di questo specifico linfocita T (Figura 3.4) (Heldin).

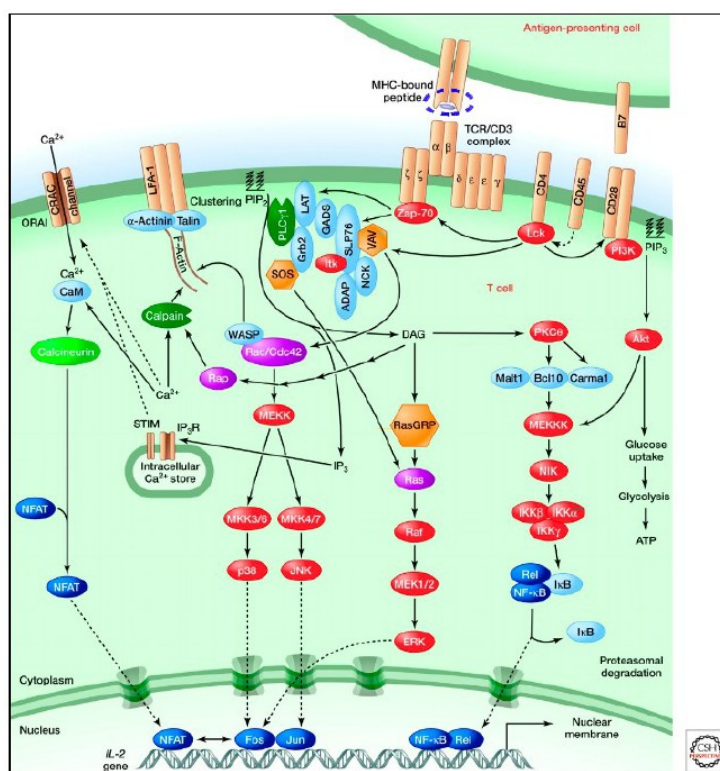


Figura 3.4. Il riconoscimento del complesso peptide:MHC da parte del TCR avvia una complessa cascata di segnalazione chimica che trasporta l'informazione sino al nucleo della cellula, avviando l'espansione clonale della cellula verso distinti percorsi di differenziazione cellulare (Lawrence 2011).

In tal modo, in dipendenza anche dal contesto specifico dell'infiammazione, come espresso dal profilo delle molecole di segnalazione presenti, verrà iniziato il processo di divisione cellulare, differenziazione e proliferazione del linfocita per dare luogo

ad un repertorio di linfociti in grado di riconoscere l'agente patogeno e eliminarlo (CTL) o coadiuvare l'attacco Th).

3.3. Il mimetismo molecolare e l'ipotesi di una sua connessione con le malattie autoimmuni

Data una specifica proteina MHC e due peptidi p_1 e p_2 , ad esempio uno *self* ed uno *non-self*, è quindi possibile che i due complessi p_1 MHC e p_2 MHC abbiano caratteristiche geometriche o fisico-chimiche simili, in termini, ad esempio, di interfaccia esterna, mappe di polarità superficiale, superficie del sito di legame, specifici pattern di amminoacidi all'interno del sito di legame. Questo può avvenire sia qualora i due peptidi p_1 e p_2 possiedano somiglianza di sequenza, sia qualora i due peptidi abbiano sequenza diverse, ma i complessi p_1 MHC e p_2 MHC risultanti abbiano una somiglianza strutturale o funzionale. In entrambi i casi si parla di mimetismo molecolare (*molecular mimicry*) fra i due peptidi (Chastain 2012) (Figura 3.5). Entrambi i peptidi sono potenzialmente riconoscibili dallo stesso linfocita T, eventualmente in momenti temporali ed in tessuti distinti. Ad esempio, come visto precedentemente, un virus che penetra nel corpo viene demolito per produrre peptidi virali che, una volta presentati dalla proteina MHC apposita, portano alla selezione di un linfocita T reattivo per quello specifico peptide, e quindi alla proliferazione di un insieme di linfociti T in grado di riconoscere ed attaccare direttamente (CTL, pathway tipico per i virus) o coadiuvare la distruzione (Th) dello stesso virus, localizzato anche in altri tessuti.

Qualora ci sia mimetismo molecolare fra un peptide virale ed un peptide *self*, è dunque possibile che tali linfociti T, attivati dal virus, attacchino le proteine che contengono il peptide *self*, arrecando in tal modo un danno ai tessuti del corpo stesso. Questa è, in forma semplificata, una delle ipotesi più accreditate per spiegare

alcuni aspetti funzionali della malattia autoimmune tramite la presenza di mimetismo molecolare fra *self* e *non-self*.

Perché questo meccanismo sia effettivo, occorre un altro elemento. Il sistema immunitario ha sviluppato, durante l'evoluzione, dei meccanismi molto raffinati e complessi per la selezione del repertorio iniziale di linfociti T in circolazione nel corpo, che devono soddisfare essenzialmente due requisiti: i) non essere auto-reattivi, ossia reattivi verso il *self*, ii) essere sufficientemente vari da poter riconoscere una varietà potenzialmente enorme di peptidi patogenici (Kuby Immunology).

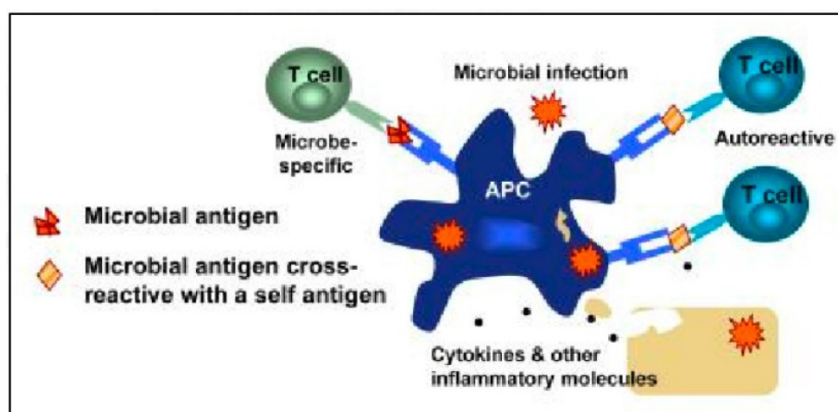


Figura 3.5. Sul lato sinistro dell'immagine si ha un linfocita T che è specifico per un solo tipo di microbo, che quindi avvia una corretta linea d'azione del sistema immunitario. Sul lato destro si ha mimetismo molecolare: l'antigene microbico presentato dalla proteina MHC attiva un linfocita T (autoreattivo e crossreattivo) che in seguito darà luogo ad una reazione autoimmune verso auto-proteine (Chastain 2012)

Questo avviene durante il processo di selezione timica, per cui un eventuale linfocita T cross-reattivo, ossia reattivo sia verso uno specifico *self* che un specifico *non-self*, deve anche essere in qualche modo sfuggito alla selezione operata a livello del timo. Il processo di mimetismo molecolare per un linfocita T auto-reattivo, può essere visto complessivamente come una perdita sia di tolleranza verso il *self* che in generale della omeostasi immunologica, ed è in grado di attivare reazioni

immunitarie contro proteine funzionalmente molto importanti del nostro corpo, potenzialmente in grado di portare a malattie autoimmuni (Chastain 2012). Nel caso specifico della SM, uno specifico linfocita T cross-reattivo potrebbe essere stato i) attivato nei tessuti periferici, in seguito al contatto con antigene virale (ad esempio derivato da EBV), e ii) dopo proliferazione e differenziazione, iii) successivamente riattivato nel sistema nervoso centrale da antigeni mielinici (presentati tipicamente da cellule dendritiche, astrociti e microglia), portando così alla distruzione di proteine mieliniche, e quindi all'inizio o al sostegno di un processo infiammatorio degenerativo autoimmune.

3.4 L'architettura della regione genetica che codifica per la proteina MHC-II

Le proteine MHC di classe II vengono codificate dalla regione genetica HLA. La struttura di questa regione é piuttosto complessa (Figura 3.6)(Klein 2000). In generale, i geni codificanti per le catene α delle proteine MHC vengono identificati dalla lettera A, mentre quelli codificanti per le catene β vengono identificati dalla lettera B. All'interno di tale regione, si trovano:

- il gene HLA-DRA;
- il gene HLA-DRB1, comune a tutti gli individui, e otto altri pseudogeni (non funzionali) e geni variabili individualmente, denotati da un diverso numero finale, DRB x ($x=2,\dots,9$);
- il gene HLA-DQA1, ed un secondo pseudogene HLA-DQ2;
- il gene HLA-DQB1, e due altri pseudogeni HLA-DQ x ($x=2,3$);

I diversi arrangiamenti dei DRB x (o dei DQB x) riscontrati nella popolazione si chiamano aplotipi. Ciascun gene può avere una specifica sequenza che varia da individuo ad individuo. Tali variazioni individuali sono rappresentazioni diverse dello stesso gene, con caratteristiche funzionali della proteina codificata diverse, e vengono chiamati alleli o polimorfismi del gene. Il gene che presenta il maggior

numero di polimorfismi nella regione HLA di classe II, oltre 200, é il DRB1. Al contrario il DRA1 ha solamente due polimorfismi. In maniera simile il DQB1 ha quasi una cinquantina di alleli, mentre il DQA1 ne ha circa la metà (Marsh). I diversi alleli di uno stesso gene sono identificati attraverso la sigla numerica che segue il nome del gene (Figura 3.7), ad esempio HLA-DRB1*01:01 (Marsh 2010).

I geni con un elevato numero di alleli sono ovviamente quelli più interessanti dal punto di vista clinico, perché gli studi statistici possono consentire di associare potenzialmente un determinato stato/allele ad una condizione di malattia specifica. Questo é quanto avviene nella Sclerosi Multipla per la quale, appunto, le sotto-regioni di interesse sono principalmente i geni HLA-DRB1 e HLA-DQB1.

In merito alla notazione adottata in questo capitolo, si utilizzeranno spesso termini quali la “proteina HLA-DRB1*15:01”, o più semplicemente “proteina DRB1*15:01”, intendendo per questo una abbreviazione per “la proteina MHC di classe II la cui catena B viene codificata dall'allele HLA-DRB1*15:01 e la cui catena A viene codificata dal gene HLA-DRA”.

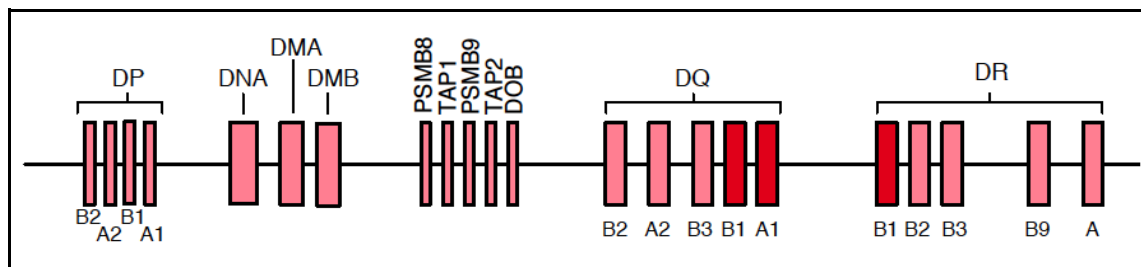


Figura 3.6. Diagramma schematico che mostra l'architettura della regione genetica HLA di classe II (disegno realizzato da O. Brand, S. Gough, J. Heward

journals.cambridge.org/fulltext_content/ERM/ERM7_23/S1462399405009981sup003.htm)

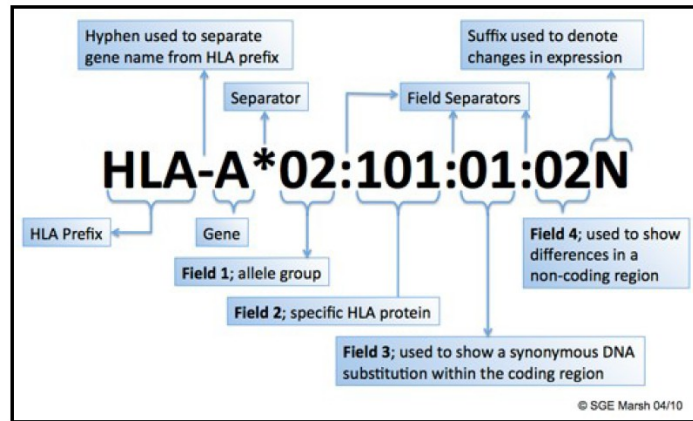


Figura 3.7. Nomenclatura dell'HLA (hla.alleles.org/nomenclature/naming.html).

3.5. Analisi di sequenza degli aplotipi del serotipo DR2 ✨

La prima analisi effettuata durante il lavoro di tesi, è relativa alla sequenza degli amminoacidi degli aplotipi HLA predisponenti e protettivi in Sardegna (Cocco 2012), evidenziati tramite un test di disequilibrio di trasmissione (TDT) in 943 famiglie. Per la difficoltà di estrarre un modello unitario per tutti gli aplotipi, l'analisi delle sequenze degli amminoacidi si è focalizzata sulle proteine del solo serotipo DR2, cui appartengono tre degli aplotipi evidenziati: DRB1*15:02-DQB1*06:01, DRB1*16:01-DQB1*05:02 e DRB1*15:01-DQB1*06:02.

All'interno di tale gruppo, in particolare, si è visto che il polimorfismo Gly/Val in posizione 86 per il DRB1, e specularmente Val/Ala in posizione 38 per il DQB1, consente di classificare rispettivamente l'aplotipo come protettivo o predisponente (Cocco 2012). Per l'allele DRB1 il cambiamento principale è legato al diverso ingombro sterico di Gly e Val, che modifica le specificità della tasca di legame P1, sito di ancoraggio primario per il peptide. In particolare la Glicina è in grado di legare grandi amminoacidi aromatici o alifatici mentre la Valina solo amminoacidi più piccoli. Questa prima osservazione, pur basata su considerazioni di struttura primaria (si vedano le Sezioni 3.8-11 per le analisi successive di carattere strutturale), potenzialmente spiega il diverso trattamento degli antigeni self e non-self operato

dagli alleli predisponente e protettivo. Per quanto concerne il polimorfismo dell'allele DQB1, il sito 38 è anch'esso ritenuto abbastanza importante. In particolare, esso è collocato sotto la tasca di legame P9 e molto vicino all'amminoacido Asp 57, il cui ruolo nella repulsione dei peptidi carichi negativamente è già stato proposto per spiegare la protezione verso alcune forme di diabete di tipo I (Lee 2001), sarcoidosi (Chow 2011) e predisposizione alla celiachia (Hovhannisyan 2008) ed all'artrite infiammatoria (Chow 2012), per quanto attraverso meccanismi funzionali distinti.

3.6. Rassegna di letteratura sul ruolo del sito DRB1 86 e DQB1 38

A conferma dell'importanza del ruolo dei siti DRB1 86 e DQB1 38, si presenta qui una breve rassegna critica della letteratura esistente. Alcuni autori (ul-Haq 2011) hanno osservato che il composto adamantine, in presenza di Glicina 86 nel sito dei legami di una specifica proteina MHC, facilita la transizione da una conformazione non ricettiva ad una ricettiva nei confronti di specifici peptidi, agendo quindi come un catalizzatore per il legame peptide ed MHC. L'effetto catalizzante era invece inibito dalla presenza del polimorfismo Val 86, per motivi di ingombro sterico. Nello stesso studio, gli autori riportano l'importanza di alcuni residui vicini all'86 come sensori cruciali per il riconoscimento del peptide, specificamente Asn 82, Tyr 83 e Thr 90, in grado di realizzare legami diretti o mediati dall'acqua con i rispettivi residui sul peptide. Altri autori (Yaneva 2009) hanno osservato come una mutazione DRB1 Val 86 Tyr producesse un blocco funzionale della tasca di legame P1, con un inatteso effetto a lungo raggio sull'intero sito di legame peptidico, come riportato anche più recentemente (Hov 2011).

In quest'ultimo lavoro, relativo alla Colangite Sclerosante Primitiva (PSC) (Hov 2011), un'altra malattia autoimmune, gli autori trovano altresì che un unico polimorfismo Asn/Tyr nella posizione 37 (tasca di legame P9) del DRB1 è in grado di

caratterizzare l'allele rispettivamente come predisponente o protettivo, conferendo alla tasca P9 una carica positiva o negativa. Gli autori ipotizzano inoltre che questo effetto potrebbe essere mascherato da cambiamenti della tasca P9 indotti dal polimorfismo (con effetti a lungo raggio) Val/Gly 86 della tasca P1. In particolare, la presenza di Val 86 fornisce predisposizione anche in caso della presenza di Tyr 37. Incidentalmente, in questo studio si studiano le stesse tasche di legame e praticamente gli stessi siti e residui che sono stati analizzati nel paragrafo precedente (Sezione 3.5) per la SM (Cocco 2012), questa volta in azione in concerto nello stesso allele DRB1 e non in maniera differenziale negli alleli DRB1 e DQB1 come per la SM. Per quanto riguarda gli alleli HLA-DQ, recentemente molti di essi sono emersi in letteratura come collegati a una varietà di malattie autoimmuni (Larizza 2012). Ad esempio, contrariamente alla SM, l'allele DRB1*15:01-DQB1*06:02 fornisce una forte protezione nei confronti del diabete di tipo I (Noble 2011, Rojas 2010), mentre in maniera simile alla SM (Vrethem 2012), la narcolessia è fortemente associata agli alleli DRB1*15:01, DQA1*01:02 e DQB1*06:02 (Siebold 2004) per molti gruppi etnici (Hong 2007). In particolare, DQB1*06:02 è considerato primariamente associato alla narcolessia (Han 2012, Weiner 2012), come recentemente confermato durante test clinici (Woo 2012, Manzotte 2013). Inoltre, la differenza di nove residui fra l'allele predisponente (DQB1*06:02) e quello protettivo (DQB1*06:01:1) per la narcolessia produce i suoi effetti strutturali principalmente sulla tasca di legame P9, conservandone il carattere polare, e sulla tasca P4, cambiandone il volume. Questi nove polimorfismi includono Ala/Val 38 e Tyr/Asp 37 (tasca P9), Tyr/Leu 26 ed Ala/Gly 13 (tasca P4) (Siebold 2004). Infine, anche la Interstiziopatia Polmonare (ILD) è associata al gruppo serologico DR2 ed ha come fattore di rischio la presenza dell'allele DQB1*06 (Furukawa 2012). L'allele DQB1*06:02 è altresì suggerito come predisponente per la SM, per quanto il suo ruolo diretto sia ancora incerto (Kaushansky 2010, Kaushansky 2012, Sadovnick 2102, Lincoln 2005, Lincoln 2009)

3.7. Antigeni rilevanti per la SM

Prima di passare all'analisi molecolare dell'interazione MHC e peptide, è bene soffermarsi a discutere brevemente quali siano gli antigeni più rilevanti per la SM. Sia la ricerca clinica che gli studi *in vivo* su modelli animali hanno identificato una serie di auto-antigeni: MBP, MOG, PLP, MOBP, MAG ed OSP, tutti derivati dalla mielina, in grado di giocare un ruolo distinto a seconda delle diverse fasi e tipologie di decorso della malattia, ad esempio induzione, ricaduta e successive ricadute cicliche (Oksenberg 2008). L'auto-antigene immuno-dominante nell'uomo, nel contesto del serotipo DR2 è MBP 85-98 (Myelin Basic Protein, Figura 2.5, sequenza ENPVVNFFKNIVTP), mentre nel contesto del gruppo DR4 è MBP 111-129.

Per quanto riguarda gli antigeni provenienti da agenti virali, un candidato largamente proposto in letteratura è il virus di Epstein-Barr (EBV), anche conosciuto come virus della mononucleosi. In particolare, recentemente (Mechelli 2011, Mameli 2013) è stato identificata in pazienti di SM una risposta autoimmune indotta dal peptide EBNA-1 400-413 (EBV Nuclear Antigen 1, Figura 2.5, sequenza PGRPPFFHPVGEAD).

3.8. Analisi molecolare degli alleli DRB1 del gruppo serologico DR2 ✳

Dati i risultati, preliminari ma incoraggianti, ottenuti dalla analisi di sequenza per gli aplotipi del gruppo DR2 (Sezione 3.5)(Cocco 2012), si è deciso di procedere ad una indagine strutturale più fine a livello molecolare (Kumar 2013). Si è realizzato un modello di riferimento, basato su due alleli del gruppo DR2: il predisponente DRB1*15:01 ed il protettivo DRB1*06:01. Il primo è stato scelto perché è nel contempo presente nella popolazione sarda (frequenza 1.5%) ed è il principale allele associato alla SM nella popolazione Caucasica; il secondo è uno dei principali alleli (frequenza 19.1%) associato alla SM in Sardegna. Si sono eseguite delle dinamiche molecolare

estensive di 120 ns delle due proteine espresse, in legame con un antigene derivato da EBV, EBNA-1 400-413, ed uno dalla mielina, MBP 85-98, ed in assenza di peptide. Lo studio analizza principalmente le configurazioni globali assunte dai complessi proteici nel tempo, e la flessibilità del sito di legame peptide:proteina al variare del tempo e della regione all'interno del sito (diviso in 4 zone, come illustrato nella Sezione 2.2.1), tramite grafici della RMSD, RMSF e istogrammi di larghezza delle regioni interne al sito. Vengono inoltre stimate le energie di legame, le entropie di configurazione ed i dettagli della rete di interazione peptide:proteina (per gli aspetti tecnici ed i protocolli per la preparazione dei sistemi, la DM e gli strumenti di analisi si vedano le Sezioni 2.2.3, 2.2.4, 2.2.5).

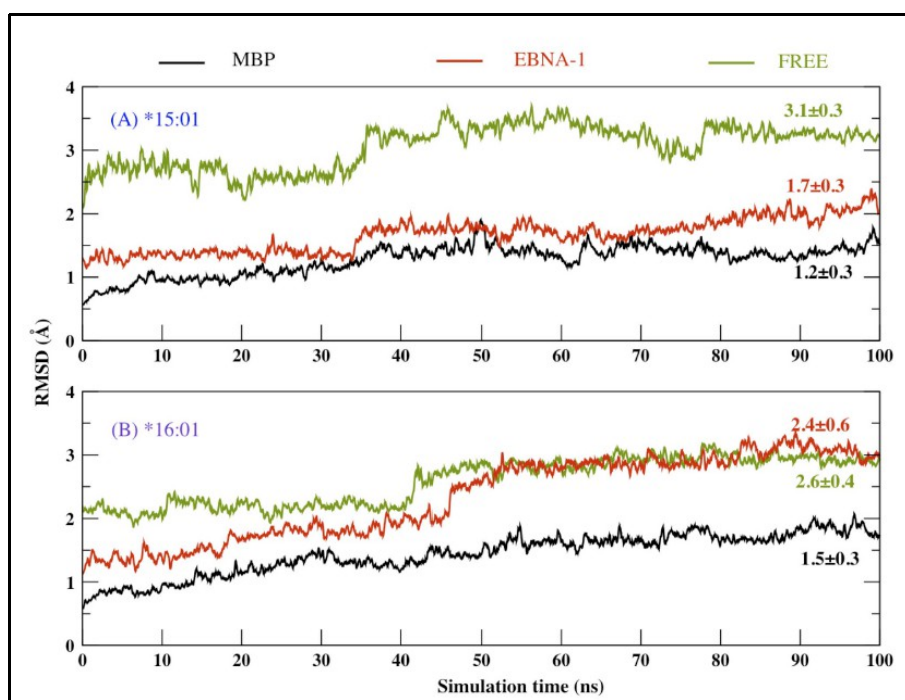


Figura 3.8. Radice della deviazione quadratica media (RMSD) per gli atomi pesanti del backbone peptidico delle proteine MHC (A) predisponente DRB1*15:01 e (B) protettiva DRB1*16:01, al variare del tempo di simulazione, e al variare del peptide presente nel sito di legame (nessuno, MBP ed EBNA1). Su ciascun grafico é riportato il valore numerico con errore della RMSD mediata nel tempo.

Il grafico della RMSD nel tempo (Figura 3.8) evidenzia immediatamente come l'allele protettivo DRB1*16:01 abbia una configurazione più rigida in legame con il peptide MBP che con EBNA-1 (1.5 Å contro 2.6 Å per la RMSD media), mentre l'allele predisponente assume configurazioni di legame molto simili verso entrambi i peptidi (1.2 Å contro 1.7 Å). Si noti come l'RMSD del complesso dell'allele protettivo con il peptide virale tenda verso quello della proteina vuota al progredire della simulazione, primo indizio che il legame in atto risulta essere piuttosto debole.

Tali osservazioni sono confermate dalle energie di legame (Tabella 3.1), che mostrano come l'allele protettivo abbia una affinità nettamente superiore per l'autoantigene, mentre l'allele predisponente abbia affinità simili per l'antigene virale e l'autoantigene. La maggiore rigidità dell'allele protettivo viene altresì confermata dalle stime di entropia configurazionale (Tabella 3.2).

	ΔG (kcal/mol)
DRB1*15:01 : MBP	-16.7±0.6
DRB1*15:01 : EBNA-1	-16.0±0.7
DRB1*16:01 : MBP	-17.5±0.7
DRB1*16:01 : EBNA-1	-12.2±0.7

Tabella 3.1. Energie di legame dei due alleli con i due peptidi.

	$T\Delta S$ (kcal/mol)		
	MBP	EBNA-1	Free
Predisposing (*15:01)	1046	1060	1073
Protective (*16:01)	1060	1075	1083

Tabella 3.2. Contributo all'energia di Gibbs da parte dell'entropia configurazionale ($\Delta G = \Delta H - T\Delta S$).

Osservando in dettaglio i grafici degli istogrammi di RMSD (Figura 3.9), si vede come l'allele predisponente abbia un profilo di RMSD con un picco significativo in comune, quando legato ad entrambi i peptidi, che segnala la presenza di una

configurazione di legame molto simile. Viceversa, i profili per l'allele protettivo sono maggiormente diversi, con quello relativo al legame col peptide virale che si sovrappone principalmente al profilo del caso libero.

Per quanto riguarda i legami idrogeno, si osserva (Figura 3.10) che entrambi gli alleli creano reti di interazione simili col peptide mielinico, con l'unica eccezione di un legame specifico dell'allele protettivo (R71-K93, tasche di legame P4, P7). Viceversa, si hanno pattern di interazione completamente distinti dei due alleli nei confronti del peptide virale (Figura 3.11).

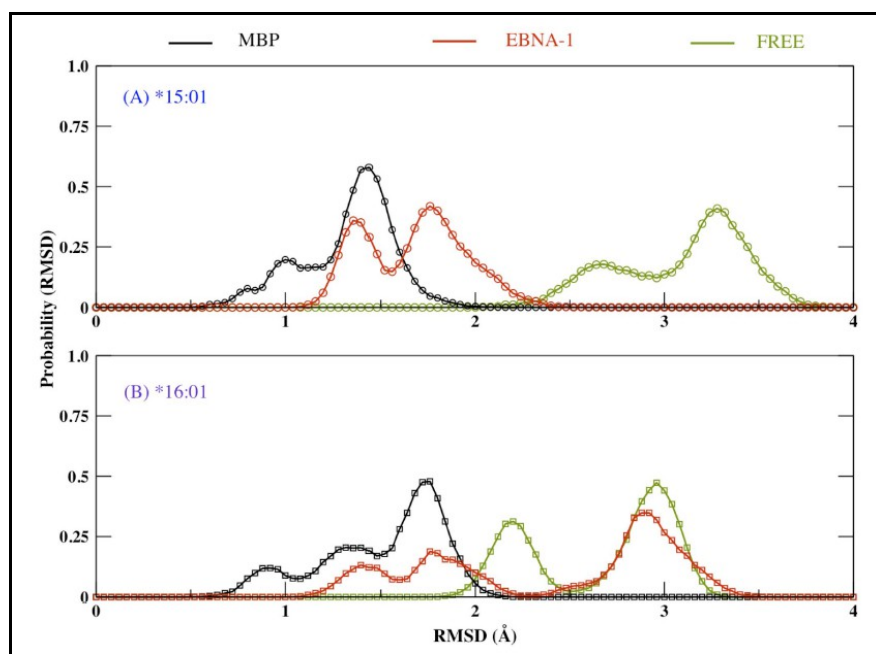


Figura 3.9. Istogrammi di RMSD per per gli atomi pesanti del backbone peptidico delle proteine MHC (A) predisponente DRB1*15:01 e (B) protettiva DRB1*16:01, al variare del peptide presente nel sito di legame (nessuno, MBP ed EBNA1). Gli istogrammi sono ottenuti cumulando i valori di RMSD assunti dal sistema durante il tempo di simulazione.

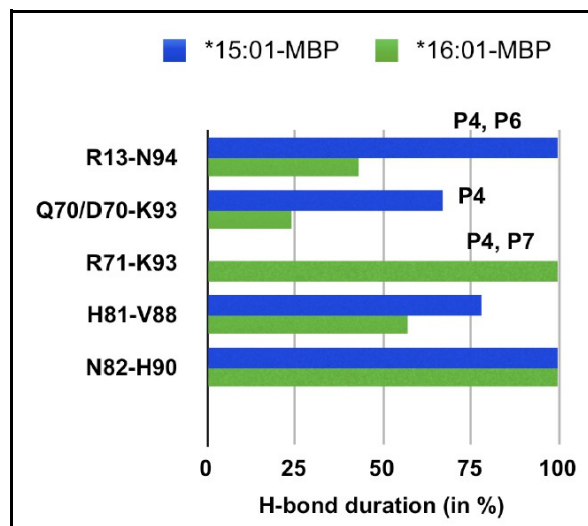


Figura 3.10. Legami idrogeno duraturi fra i residui dei due alleli (blu predisponente e verde protettivo) ed il peptide mielinico.

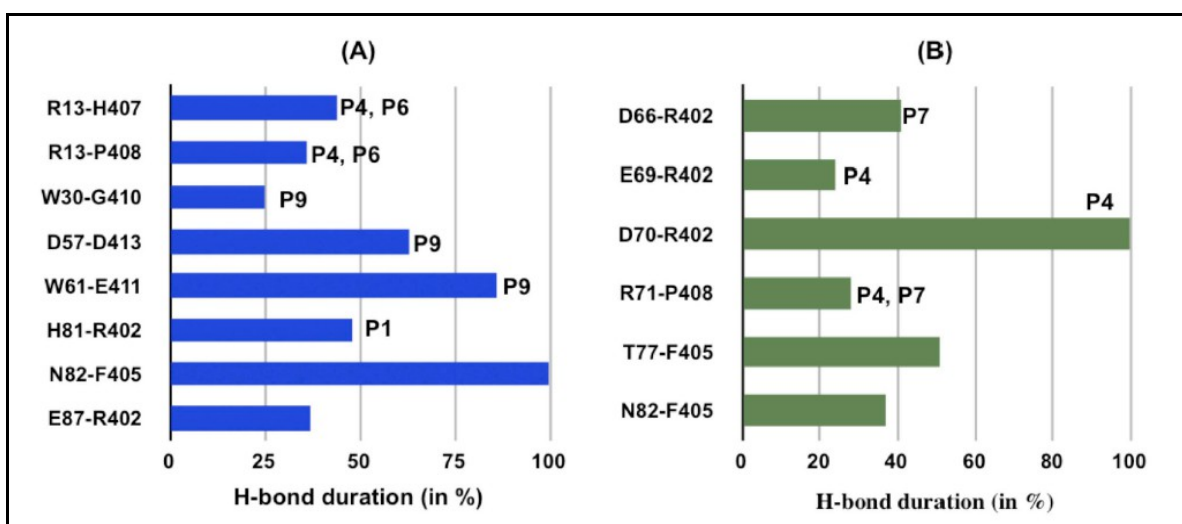


Figura 3.11. Legami idrogeno duraturi fra i residui dei due alleli, (A) predisponente e (B) protettivo, ed il peptide virale.

Andando nel dettaglio degli istogrammi di fluttuazione di larghezza delle varie regioni in cui é stato suddiviso il sito di legame (si veda la Sezione 2.2.1), si osserva che l'allele predisponente ha dei profili molto simili in complesso con entrambi i peptidi nelle regioni D2, D3 e D4, e mostra invece una differenza significativa nella sola regione D1 (Figura 3.12). La Figura 3.13 mostra come l'allele predisponente libero possieda tre principali configurazioni, corrispondenti ai tre picchi del grafico, (10.5 Å, 13 Å, 16.8 Å), la seconda delle quali viene condivisa dall'allele in legame col

peptide virale. É dunque possibile interpretare le tre configurazioni come uno stato chiuso del sito di legame (10.5 Å), ricettivo (13 Å) e aperto (16.8 Å), relativamente al legame col peptide virale. Al contrario, l'allele protettivo mostra dei profili nettamente diversi per il legame col peptide virale o mielinico su tutto il sito di legame, particolarmente nella regione D1 (Figura 3.13) e D2 (Figura 3.14). In particolare nel sito D2 si osserva una differenza di circa 10 Å fra l'apertura media del sito nel caso dei due peptidi, col peptide virale che dà origine all'apertura maggiore.

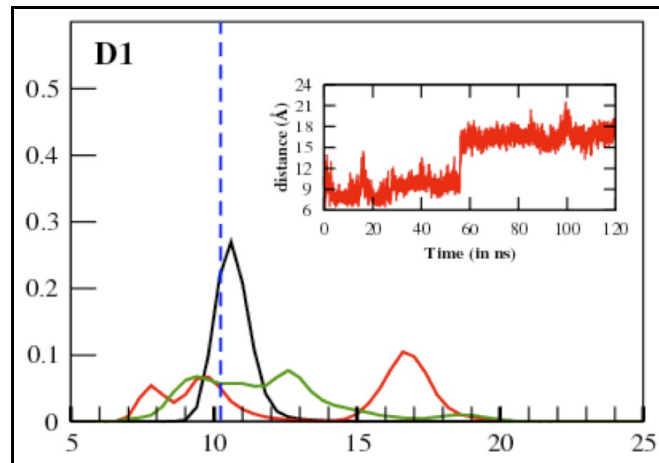


Figura 3.12. Istogramma delle fluttuazioni di larghezza del sito di legame in D1, per l'allele predisponente libero (verde) o in legame con peptide mielinico (nero) o virale (rosso). Sull'asse orizzontale é riportata la larghezza della regione in Å.

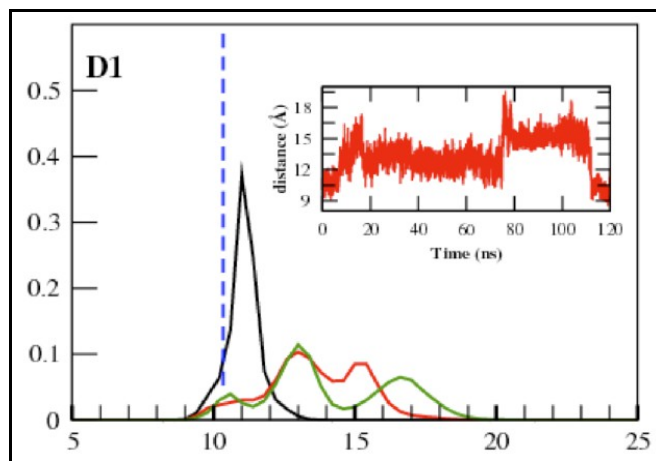


Figura 3.13. Istogramma delle fluttuazioni di larghezza del sito di legame in D1, per l'allele protettivo libero (verde) o in legame con peptide mielinico (nero) o virale (rosso). Sull'asse orizzontale é riportata la larghezza della regione in Å.

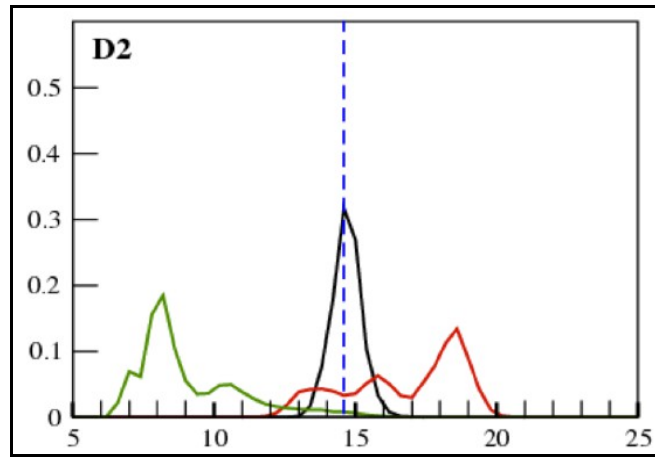


Figura 3.14. Istogramma delle fluttuazioni di larghezza del sito di legame in D2, per l'allele protettivo libero (verde) o in legame con peptide mielinico (nero) o virale (rosso). Sull'asse orizzontale é riportata la larghezza della regione in Å.

Infine, la Tabella 3.3 riporta le interazioni idrofobiche e di stacking osservate fra i due alleli ed i due peptidi. Si osserva che entrambi gli alleli formano le stesse interazioni di stacking nell'interazione col peptide mielinico, mentre solo quello predisponente forma una interazione di stacking col peptide virale. Lo stesso fenomeno si verifica per le interazioni idrofobiche.

	Interazioni Aromatiche di Stacking	Interazioni Idrofobiche
1501-MBP	α F54-F91, β F26-F92, β H81-H90	β N82-V89, β V85-V89
1501-EBNA-1	β Y78-H407	β V85-P404
1601-MBP	α F54-F91, β F26-F92, β R71-F92	β N82-V89, β V85-V89
1601-EBNA-1	Absent	Absent

Tabella 3.3. Reti di interazione di stacking aromatiche ed idrofobiche fra alleli e peptidi.

In sommario: si è osservato che mentre l'allele predisponente DRB1*15:01 possiede una rete di interazione conservata con entrambi gli antigeni EBNA-1 400-413 ed MBP 85-98, l'allele protettivo ha la peculiarità di distinguere i due antigeni, attraverso interazioni idrofobiche e di stacking uniche e specifiche per l'auto-antigene

mielinico. In tal modo, la modellistica conferma che per gli alleli DRB1 del gruppo serologico DR2 ci sia un mimetismo molecolare fra i due specifici peptidi virale e mielinico, che può pertanto condurre ad una reazione autoimmune ed alla SM.

Proprio recentemente, alcuni autori hanno sperimentalmente confermato le diverse caratteristiche dinamiche degli aplotipi predisponenti e protettivi qui discusse, a livello di riconoscimento anticorpale (Mameli 2014). In dettaglio, la risposta umorale al peptide MBP 85-98 dei pazienti di SM positivi al virus EBV, risulta essere molto elevata e confrontabile a quella nei confronti del peptide EBNA-1 400-413. Inoltre, tramite saggi di binding competitivo, gli autori hanno verificato che gli auto-anticorpi reattivi verso MBP 85-98 erano altresì in grado di reagire con il peptide EBNA-1 400-413. Gli autori, in linea col lavoro di modellistica qui svolto, concludono osservando che, “dal momento che i due peptidi non condividono omologia di sequenza, è lecito supporre che essi condividano omologia conformazionale”, quando presentati dalla proteina MHC. Per quanto nel lavoro di modellistica si sia considerata la risposta immunitaria mediata dai linfociti T, e non dai linfociti B (responsabili della generazione degli anticorpi), il lavoro sperimentale appena pubblicato rappresenta una prima conferma dei risultati delle simulazioni.

3.9. Uno studio più ampio, a livello di sequenza e funzione, di nove aplotipi

predisponenti e protettivi ✳

In seguito alle ricerche effettuate dal gruppo clinico col quale si collabora, tramite test di disequilibrio di trasmissione (TDT) (Cocco 2012) e studi genetici casi-controllo (Cocco 2013), sono stati caratterizzati i profili di resistenza e di permissività (rispetto alla malattia) degli aplotipi HLA-DRB1-DQB1. Attraverso una classificazione filogenetica (Figura 3.15) si osserva immediatamente che i nove aplotipi si raggruppano all'interno di un cluster principale, composto da tre aplotipi

appartenenti al serotipo DR2 (DRB1*15:01–DQB1*06:02, DRB1*15:02–DQB1*06:01 e DRB1*16:01–DQB1*05:02), ed un cluster composto dai rimanenti cinque aplotipi DRB1–DQB1 (*13:03–*03:01, *04:05–*03:01, *03:01–*02:01 , *08, *14:01–*05:03). Dal punto di vista della struttura primaria, si veda la Tabella 3.4, i due gruppi sono caratterizzati dal polimorfismo dei residui numero 9 (W/E) e 133 (L/R). All'interno del nuovo gruppo (non-DR2), i cambiamenti più rilevanti (volume o caratteristiche polari) si verificano alle posizioni 74 (tasca di legame P4) e 57 (P9). Come base per le analisi successive, si è pertanto ipotizzato che i meccanismi molecolari collegati alla malattia fossero in generale diversi all'interno di ciascun raggruppamento, DR2 o 'non-DR2', il che spiegherebbe le difficoltà nel trovare un comune denominatore del ruolo degli alleli nella genesi e nel mantenimento della malattia.

La seconda osservazione è la posizione atipica del DRB1*14:01, unico allele protettivo all'interno del nuovo gruppo 'non-DR2' così identificato. Pertanto è stata svolta una analisi strutturale comparata del DRB1*14:01 e dell'allele più vicino, il DRB1*03:01, con particolare attenzione alle tasche di legame P4 e P9.

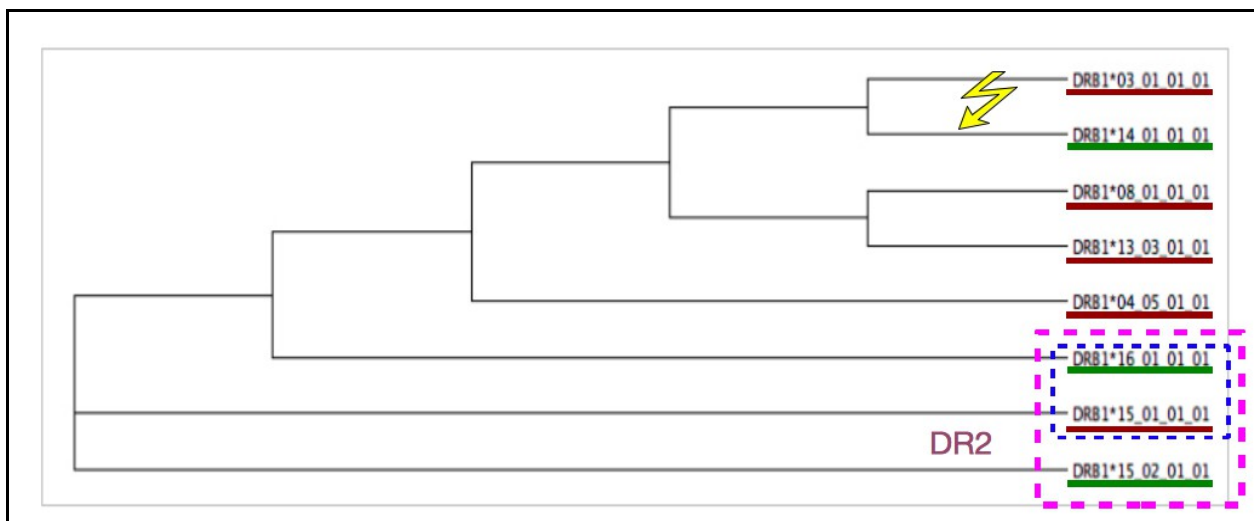


Figura 3.15. Raggruppamento degli alleli DRB1 associati alla SM in Sardegna (rosso = predisponente, verde = protettivo). Allineamento delle sequenze con BLASTP (blast.ncbi.nlm.nih.gov), albero generato con Clustal-W phylogeny tool (www.ebi.ac.uk/Tools/phylogeny/) e visualizzato con Dendroscope (ab.inf.uni-tuebingen.de/software/dendroscope/).

Pos	9	10	11	12	13	26	32	33	37	47	57	60	67	70	71	73	74	77	86	96	98	104	112	120	133	140	142	149
*16:01	W	Q	P	K	R	F	Y	N	S	Y	D	Y	F	D	R	A	A	T	G	Q	K	S	H	S	L	A	M	Q
*15:02	W	Q	P	K	R	F	Y	N	S	F	D	Y	I	Q	A	A	A	T	G	Q	K	S	H	S	L	A	M	Q
*14:01	E	Y	S	T	S	F	H	N	F	Y	A	H	L	R	R	A	E	T	V	H	K	S	Y	S	R	T	V	H
*04:05	E	Q	V	K	H	F	Y	H	Y	Y	S	Y	L	Q	R	A	A	T	G	Y	E	A	H	N	R	T	V	Q
*13:03	E	Y	S	T	S	F	Y	N	Y	Y	S	Y	I	D	K	A	A	T	G	H	K	S	H	S	R	T	V	H
*03:01	E	Y	S	T	S	Y	H	N	N	F	D	Y	L	Q	K	G	R	N	V	H	K	S	H	S	R	T	V	H
*15:01	W	Q	P	K	R	F	Y	N	S	F	D	Y	I	Q	A	A	A	T	V	Q	K	S	H	S	L	A	M	Q
*08:01	E	Y	S	T	G	F	Y	N	Y	Y	S	Y	F	D	R	A	L	T	V	H	K	S	H	S	R	T	V	H
Pock	9	6	46						7	9	7	4	47	4	1													

Tabella 3.4. Allineamento degli otto alleli DRB1 in analisi. Sono evidenziate le sole posizioni con un diverso residuo.

La struttura iniziale DRB1*03:01 é stata presa dalla banca dati PDB (id 1a6a, in complesso col CLIP) ed é stata anche utilizzata per il modello per omologia del DRB1*14:01. Entrambi gli alleli sono stati legati con il peptide MBP 85-98 e hanno seguito il protocollo usuale per la DM (si vedano le Sezioni 2.2.3-5). Sono state quindi eseguite delle brevi dinamiche di 3 ns. La prima peculiarità che emerge dalla DM riguarda la presenza di un legame idrogeno relativo alla tasca di legame P4 per il solo DRB1*14:01 (protettivo), si veda la Figura 3.16, che contribuisce a rendere più stabile il legame col peptide mielinico.

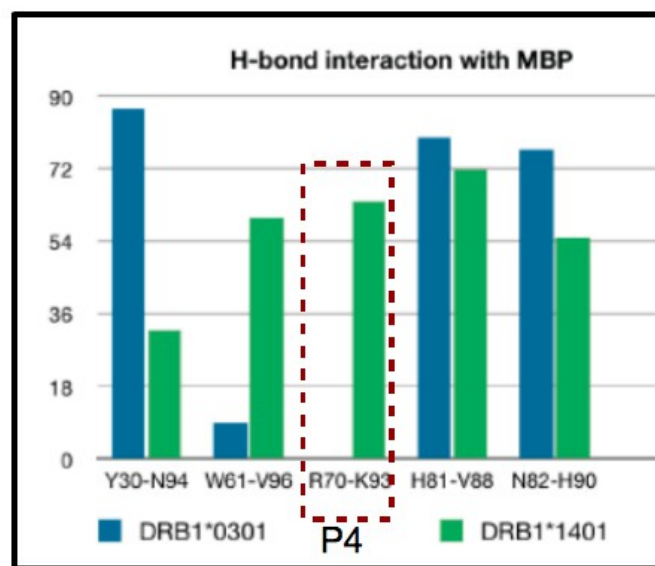


Figura 3.16. Interazioni idrogeno fra gli alleli DRB1*03:01 e DRB1*04:01 ed il peptide mielinico.

Osservando inoltre l'area disponibile per il legame col peptide nei residui della tasca P4 e P9 (Figura 3.17), si osserva come i due alleli abbiano caratteristiche abbastanza diverse nei siti 70 – 74 appartenenti alla tasca P4. Questa serie di posizioni é stata ampiamente studiata (Ou 1998, Fu 1995), in particolare in relazione all'artrite reumatoide (Wuncherpfenning 1995) ed alla SM ed alla sua prognosi (Gilbert 2003, Greer 2005).

Andando infine nel dettaglio della P4 e del suo intorno, si osserva una differenza nel rapporto fra area polare ed apolare disponibile per l'interazione, fra i due alleli nella P4 stessa, e una differenza più marcata vicino al residuo 60 (Figura 3.18).

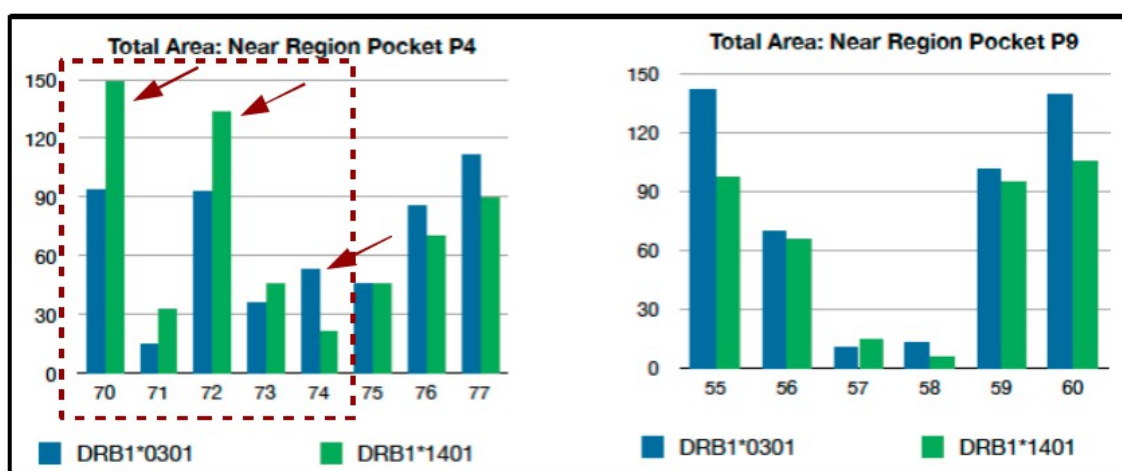


Figura 3.17. Area totale disponibile per il legame fra gli alleli DRB1*03:01 e DRB1*04:01 ed il peptide mielinico, nelle tasche di legame P4 e P9.

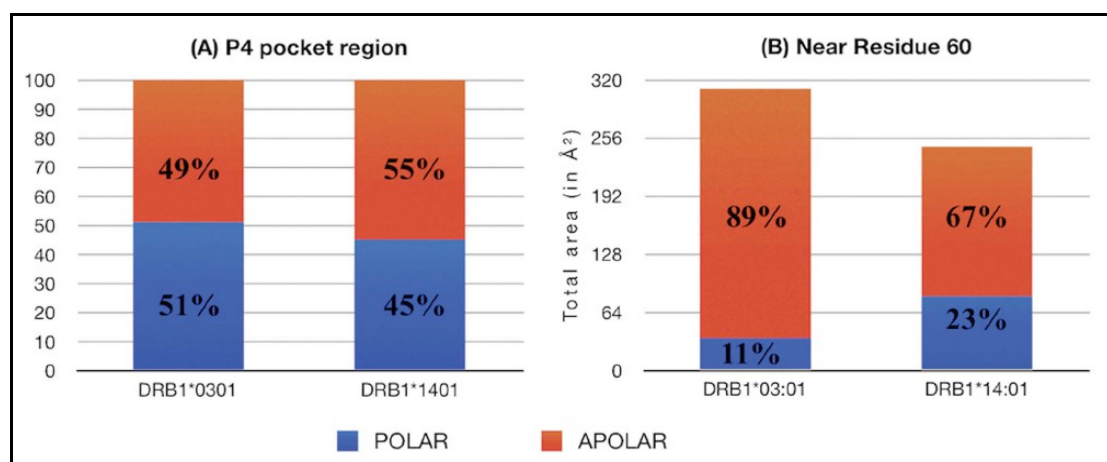


Figura 3.18. Area polare ed apolare disponibile per il legame, in vicinanza della tasca di legame P4 e del residuo 60, per i due alleli DRB1*03:01 e 04:01.

Le diverse caratteristiche polari dei due alleli, qui riscontrate, decisamente contribuiscono a spiegare il loro diverso ruolo nella malattia. Alcuni autori (Barcelos 2006) hanno ipotizzato che il pattern di amminoacidi all'interno della tasca P9 (specificatamente il residuo 60 H/T) potesse avere un ruolo importante, portando ad un docking sub-ottimale di peptidi encefalitogenici per l'allele DRB1*14:01 e conferendogli dunque caratteristiche protettive in confronto al 15:01. I risultati qui presentati confermano questa osservazione, in questo caso con la tasca di legame P9 che svolge un ruolo funzionale più rilevante nel distinguere l'allele protettivo *14:01 dal predisponente *03:01.

Infine, per quanto riguarda l'interazione con i linfociti T, alcuni autori (Wuncherpfenning 1995) hanno suggerito che alcuni residui del peptide MBP 85-98 svolgessero un ruolo primario nel legare il TCR. Fra di essi si trova proprio il residuo K93, che interagisce con l'R70 nel solo caso dell'allele protettivo *14:01 (Figura 3.16).

Pertanto, i risultati qui presentati sono in linea con le caratteristiche che gli alleli devono possedere nel legare il peptide mielinico e nel presentarlo al TCR, al fine di proteggere o predisporre nei confronti della SM. In dettaglio, i risultati consentono di ipotizzare che le tasche di legame P4 e P9 agiscano in sinergia nel conferire all'allele le sue specifiche caratteristiche di legame peptidico, dando resistenza all'allele in grado di: i) legarsi al peptide mielinico più debolmente in P9 e ii) con maggiore affinità in P4, rispetto all'allele predisponente. Questa duplice caratteristica consente all'allele protettivo sia di legare l'antigene e svolgere dunque le sue funzioni fisiologiche nel contesto del sistema immunitario, sia di privare il TCR di un possibile punto di ancoraggio primario, che risulterebbe maggiormente internalizzato all'interno del sito di legame della proteina MHC.

3.10. *Analisi molecolare degli aplotipi DRB1-DQB1 del gruppo DR2* ✳

Coerentemente con l'approccio adottato in precedenza per l'analisi molecolare degli alleli DRB1 del gruppo serologico DR2 (Sezione 3.8), si è proceduto ad allargare l'ambito di ricerca allo spettro degli aplotipi completi DRB1-DQB1 del gruppo serologico DR2. Sono stati pertanto considerati due aplotipi: DRA1*01:01-DRB1*15:01 – DQA1:01*02-DQB1*06:02 (predisponente) e DRA1*01:01-DRB1*16:01 – DQA1:01*02-DQB1*05:02 (protettivo), sempre con l'idea di creare un quadro di riferimento semplice ma non banale per comprendere come emergano le caratteristiche di suscettibilità della SM a livello di aplotipo. Sono stati qui riportati anche gli alleli delle catene A, dato che essi sono necessari per la costruzione del modello dell'intera proteina, ma si ricorda (Sezione 3.4) che essi non sono associati alla SM, e non sono polimorfici (come è il caso del DRA1), oppure sono molto meno variabili della catena B (caso del DQA1).

Come discusso ampiamente nelle Sezioni 2.2.2 e 3.3, gli elementi chiave per l'interpretazione delle dinamiche molecolari nel contesto della sclerosi multipla sono che: i) l'attivazione del linfocita T è collegata alle caratteristiche dinamiche del complesso ternario TCR:p:MHC, ii) e dunque indirettamente alle caratteristiche dinamiche del complesso binario p:MHC; iii) l'auto-reattività del linfocita T è connessa alla incapacità di discriminare specifici antigeni virali e auto-antigeni nel contesto della presentazione effettuata da specifiche proteine MHC, iv) e dunque indirettamente alla abilità delle proteine MHC di distinguere peptidi virali e self (mimetismo molecolare). In maniera simile a quanto osservato nella precedente analisi (Sezione 3.8) dei due alleli DRB1 del gruppo DR2, anche gli alleli DQB1 predisponente e protettivo mostrano delle caratteristiche dinamiche (flessibilità) specifiche del legame col peptide virale e mielinico: l'allele DQB1 protettivo è in grado di operare una blanda discriminazione fra peptide mielinico e virale, mentre

l'allele DQB1 predisponente opera una distinzione più fine. Questo effetto è l'opposto di quanto osservato per gli alleli DRB1, suggerendo quindi che il riconoscimento molecolare operato dall'aplotipo DRB1-DQB1, nel suo complesso, sia basato su meccanismi complementari e distinti operati dai singoli alleli DRB1 e DQB1 che lo compongono.

	MBP	EBNA1	$\Delta G^{\text{EBNA1}} - \Delta G^{\text{MBP}}$
	ΔG (kcal/mol)	ΔG (kcal/mol)	(kcal/mol)
DQB1*05:02 (prot)	-18.5±0.6	-17.0±0.7	1.5±1.3
DQB1*06:02 (pred)	-14.1±0.8	-17.5±0.8	-3.4±1.6
DRB1*16:01 (prot)	-17.5±0.7	-12.2±0.8	5.3±1.5
DRB1*15:01 (pred)	-16.7±0.6	-16.0±0.7	0.7±1.3
DQB1*06:02–	4.4±1.4	-0.5 ±1.5	kcal/mol
DQB1*05:02			
DRB1*15:01–	0.8±1.3	-3.8±1.4	
DRB1*16:01			

Tabella 3.5. Energie di legame dei vari aplotipi DRB1-DQB1 con i due peptidi in analisi. L'ultima colonna a destra riporta la differenza fra l'energia di legame con i peptidi EBNA1 ed MBP per ciascun allele (DRB1 o DQB1); le ultime due righe in basso riportano la differenza di energia fra l'allele (DRB1 o DQB1) predisponente e quello protettivo nel legare lo stesso peptide.

La Tabella 3.5 e la Figura 3.19 riportano le energie di legame dei complessi analizzati. Il confronto dei dati tabellari, pur considerando gli errori sulle energie, suggerisce un quadro complesso caratterizzato dai seguenti elementi:

(1) L'aplotipo protettivo è in grado di distinguere peptide MBP (mielinico) ed EBNA-1 (virale) attraverso due modalità distinte: il DRB1 consente un elevato livello di discriminazione dei peptidi virale/mielinico (differenza di 5.3 kcal/mol), con un livello di affinità intermedio per entrambi i peptidi; mentre il DQB1 integra questa

caratteristica con una abilità di discriminazione intermedia (1.5 kcal/mol) ed una elevata affinità per entrambi i peptidi.

(2) L'aplotipo predisponente presenta una abilità di discriminazione di livello più basso per gli alleli DRB1 (differenza 0.7 kcal/mol), e di livello da intermedio ad alto per gli alleli DQB1 (-3.4 kcal/mol) e con affinità superiore verso il peptide virale.

(3) L'affinità di entrambi gli alleli DRB1 e DQB1 per il peptide mielinico è superiore per l'aplotipo predisponente (0.8 e 4.4 kcal/mol), mentre vale l'opposto per il peptide virale (-3.8 e -0.5 kcal/mol), con un scambio della magnitudine.

Il primo punto può essere interpretato in termini della efficacia di presentazione del peptide da parte delle cellule specializzate per tale funzione (APC). In dettaglio: l'APC dotata di aplotipo protettivo DRB1*16:01-DQB1*05:02 è in grado di presentare complessi MBP:MHC stabili e dunque presumibilmente con una emivita lunga, sia attraverso il DRB1 che il DQB1, mentre presenta complessi EBNA-1:MHC di emivita intermedia solo tramite il DQB1. Pertanto, tale APC potrebbe essere in grado di legarsi al TCR con due modalità diverse, dunque attivare potenzialmente i linfociti T tramite distinti meccanismi, ed in maniera più efficace.

Per quanto riguarda il secondo punto, l'aplotipo predisponente DRB1*15:01-DQB1*06:02 ha perso in parte la normale abilità discriminatoria, invertendo il ruolo dei peptidi virale e mielinico nel processo di discriminazione. I primi due punti, considerati assieme, suggeriscono che la predisposizione alla SM sia collegata alla perdita, da parte dell'aplotipo DRB1-DQB1, della capacità di discriminare i peptidi MBP ed EBNA-1 e nel contempo ad una forma di complementarità DRB1/DQB1 nel meccanismo di riconoscimento molecolare

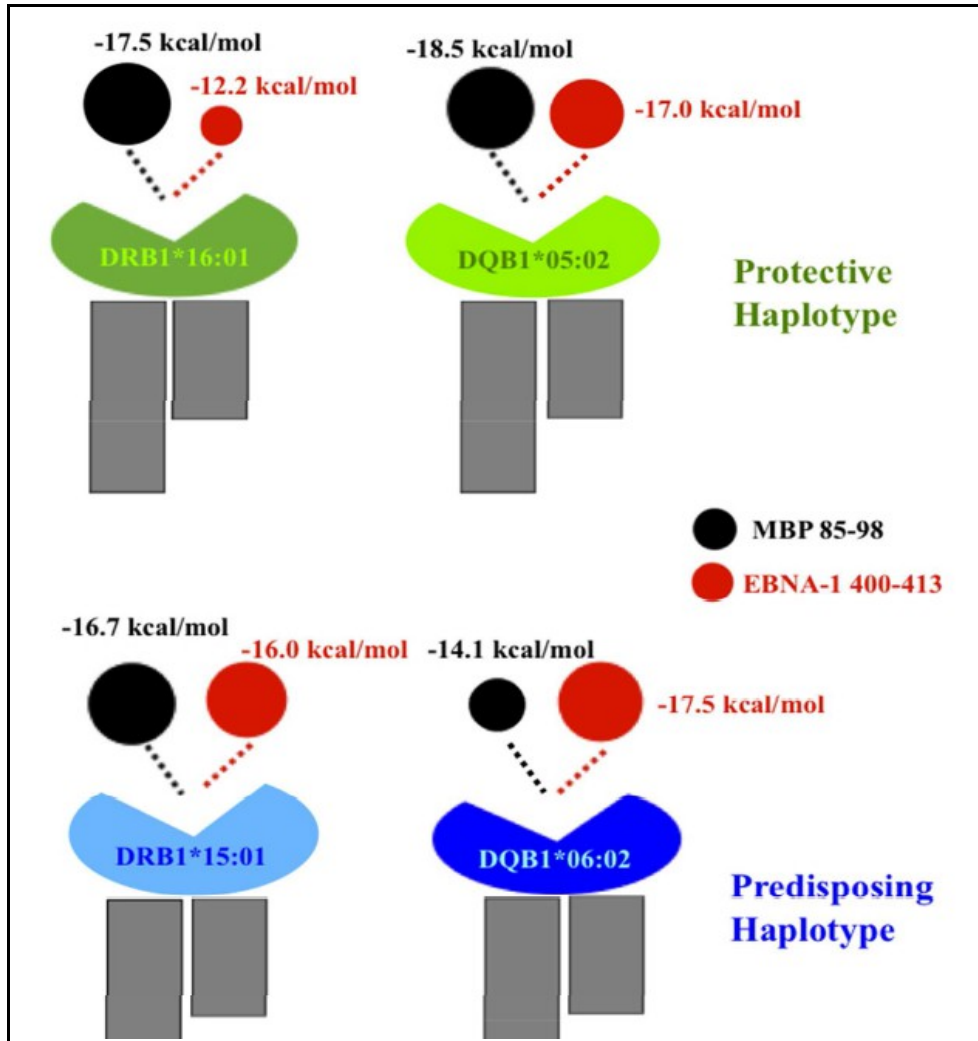


Figura 3.19. Rappresentazione grafica della Tabella 3.5. Per ogni complesso sono riportate le energie di legame dell'aplotipo coi due peptidi in analisi. La dimensione delle sfere (rossa per l'antigene viale e nera per quello mielinico), rappresenta il valore dell'energia di legame. In alto a sinistra si ha il DRB1*16:01, a destra *05:02, in basso a sinistra *15:01, a destra *06:02.

Come precedentemente discusso nel contesto della SM (Sezione 3.3), il mimetismo molecolare richiede una serie di eventi con una adeguata sequenza temporale: a) dapprima il linfocita T incontra un complesso EBNA-1:MHC, in grado di legare il TCR ed eventualmente attivare il linfocita stesso; e b) dopo differenziazione, espansione clonale e migrazione nel SNC, c) il repertorio di linfociti T EBNA-1-reattivi incontra i complessi MBP:MHC e, in caso di linfocita T cross-reattivo (ed auto-reattivo), esso inizia o coadiuva l'attacco contro le proteine mieliniche. I

passaggi a) e c) possono anche avvenire in tessuti distinti, con diversi profili di espressione delle proteine di membrana DRB1 e DQB1. Pertanto, il terzo punto (3) suggerisce che il ruolo del DRB1 sia più rilevante nella prima fase di attivazione (a), quando il peptide derivato dalla proteina virale EBNA-1 viene presentato nei tessuti periferici; mentre il ruolo del DQB1 dovrebbe essere più importante nella seconda fase (c), quella dell'infiammazione autoimmune nel SNC.

In sintesi: nel caso dei peptidi derivati da EBV ed MBP e dei due specifici alplotipi qui investigati, le proteine DRB1 e DQB1 presentano complementarità ed integrazione, tali da rendere il meccanismo di mimetismo molecolare particolarmente efficace, con l'aplotipo predisponente capace di una blanda discriminazione e con una affinità maggiore per l'antigene virale rispetto all'aplotipo protettivo. Questa strategia potrebbe essersi evoluta attraverso duplicazione dei geni e successiva divergenza evolutiva, col risultato finale di un processo di riconoscimento dell'agente patogeno più robusto e flessibile, in grado altresì di agire diversamente nei diversi tessuti.

Questo risultato concorda con la restrizione promiscua dimostrata per la prima volta da alcuni autori nel contesto della SM (Sospedra 2006), ossia il fatto che lo stesso TCR sia in grado di distinguere lo stesso peptide nel contesto di due differenti proteine MHC coesprese, DRB e DQB. Gli autori ipotizzano che questo meccanismo si sia evoluto per fornire un vantaggio durante l'infezione, dato che un livello di riconoscimento e presentazione superiore implica una più elevata efficienza di attivazione dei linfociti T. Secondo gli autori, proprio questo aumento di efficienza, aumentando il rischio di cross-reattività, potrebbe essere responsabile della risposta autoimmunitaria nella SM.

Esistono inoltre evidenze di letteratura, da modelli animali di SM, che gli alleli DRB1*15:01 e DQB1*06:02 esercitino un ruolo indipendente nell'induzione della malattia, anche se nel contesto di due auto-antigeni diversi (PLP, MOBP) da quello qui studiato (MBP) (Kaushansky 2012, Kaushansky 2010). In particolare i due alleli

sono in grado di attivare i linfociti T in maniera distinta, attivando profili di citochine molto diversi, che conducono alla differenziazione verso i fenotipi cellulari Th2 oppure Th1/Th17. Questi due fenotipi sono noti per avere esiti molto diversi sui processi di infiammazione. L'abilità degli alleli DRB1 e DQB1 di attivare indipendentemente ed in maniera differenziale i linfociti T, viene anche osservata nel modello qui proposto ed analizzato. Altre simulazioni sono naturalmente necessarie per poter procedere ad un confronto puntuale, includendo nell'analisi altri sistemi pMHC e l'interazione col TCR in casi selezionati. Infine, il modello qui proposto potrebbe fornire un meccanismo di attivazione differenziale dei linfociti T che risulti specifico del tessuto, contribuendo a chiarire alcuni effetti di interazione epistatica dei geni DRB1 e DQB1, non ancora spiegati a livello funzionale (Lincoln 2009).

3.11. Analisi molecolare dei mutanti DRB1-DQB1 del gruppo DR2 ✳

Nella precedente analisi (Sezione 3.8) (Cocco 2012), si é osservato come il polimorfismo di un singolo sito (DRB1 Gly/Val 86 o DQB1 Val/Ala 38) consentisse di classificare gli aplotipi predisponenti o protettivi relativamente alla SM nei pazienti sardi. Per valutare l'aspetto funzionale del residuo e la sua rilevanza nel conferire all'aplotipo il suo caratteristico legame con lo stato di malattia, si é pertanto valutato l'effetto di una mutazione di quel sito che invertisse il ruolo dei due valori del polimorfismo. Sono stati generati gli istogrammi di larghezza del sito di legame diviso nelle solite quattro regioni D1, D2, D3 e D4 (si veda la Sezione 2.2.1) per tutti i complessi originali (wild-type) ed i loro mutati. Dato l'elevato numero di grafici così ottenuti, si é operata una selezione di quelli contenenti le informazioni più significative.

Per quanto concerne le energie di legame del peptide alla proteina MHC, in seguito a mutazione (Tabella 3.6), gli unici cambiamenti significativi sono osservati per il predisponente DRB1*15:01 : EBNA1 ed il protettivo DQB1*06:02 : MBP, che

rappresentano quindi quelli più colpiti dagli effetti della mutazione. Negli altri casi l'errore di predizione non consente di estrapolare nessuna informazione utile, ed è quindi necessario rivolgersi ad analisi maggiormente dettagliate della rete di interazione peptide-proteina.

In generale si è osservato che, per quanto concerne gli alleli DQB1:

- 1) la mutazione Val 38 Ala nel complesso DQB1*05:02 protettivo – MBP ha una influenza a lungo raggio (Figura 3.20 (A)), suggerendo l'esistenza di un effetto cooperativo (Ferrante 2012) fra la regione dove è avvenuta la mutazione (D4) e quella dove è stato osservato il massimo effetto (D1);
- 2) la mutazione Val 38 Ala nel complesso DQB1*05:02 protettivo – EBNA1 ha un effetto maggiormente locale, con la massima perturbazione osservabile nella regione D2 (Figura 3.20 (C));
- 3) la mutazione Ala 38 Val per l'allele DQB1*06:02 predisponente ha anch'essa un impatto in generale minore, che si estende sino alla regione D2 per il complesso con entrambi i peptidi (Figura 3.20 (C)), suggerendo un blando effetto cooperativo fra le regioni D4 e D2.

	MBP		EBNA1	
	ΔG_{MUT} (kcal/mol)	$\Delta G_{MUT} - \Delta G_{WT}$ (kcal/mol)	ΔG_{MUT} (kcal/mol)	$\Delta G_{MUT} - \Delta G_{WT}$ (kcal/mol)
V38A (prot)-DQB1*05:02	-18.3±0.6	0.2±1.2	-17.9±0.6	-0.9±1.3
A38V (pred)-DQB1*06:02	-15.6±0.7	-1.5±1.5	-18.3±0.8	-0.8±1.6
G86V (prot)-DRB1*16:01	-18.0±0.6	-0.5±1.3	-13.0±0.8	-0.8±1.5
V86G (pred)-DRB1*15:01	-16.0±0.7	0.7±1.5	-18.4±0.8	-2.4±1.5

Tabella 3.6. Energie di legame fra i due peptidi in analisi e le proteine mutanti dei quattro alleli studiati. La notazione V38A (prot)-DQB1*05:02 indica, ad esempio, la proteina mutante ottenuta dall'allele protettivo DQB1*05:02 trasformando l'amminoacido Val nel sito 38 in Ala.

Come osservato precedentemente (Sezione 3.5), le variazioni della dinamica locale nella regione D4, in seguito a mutazione Val 38 Ala o Ala 38 Val, possono essere imputate ad un effetto sterico locale in seguito alla mutazione.

Per quanto riguarda gli alleli DRB1, si ha che:

1) l'effetto della mutazione Gly 86 Val sull'allele DRB1*16:01 protettivo si osserva principalmente per il solo complesso con peptide virale (Figura 3.21), e produce i) un restringimento locale del sito di legame in D1 e D2 ed ii) un allargamento in D3, D4.

Il primo effetto si spiega con la comparsa di nuovi legami idrogeno fra i siti H81, E87 ed il peptide EBNA-1 (Figura 3.22 B), ed il secondo con la perdita dei contatti D66, E69 e D70 (Figura 3.22 B).

2) La mutazione Val 86 Gly sull'allele predisponente DRB1*15:01 ha un impatto principalmente locale per entrambi i peptidi, nella sola zona D1. In dettaglio si osserva nel caso del peptide virale un restringimento della regione (a causa di nuovi legami idrogeno H81-R403 e R13-P408) (Figura 3.22 A) e nel caso del peptide mielinico un aumento di flessibilità (a causa di una interazione idrogeno transiente R13-N94) (Figura 3.23 A).

In sintesi, l'azione delle mutazioni é quella di distruggere o modificare sensibilmente i pattern di interazione esistenti fra proteina MHC e peptide, cambiando drasticamente le caratteristiche dell'allele e le sue peculiarità nel processo di riconoscimento molecolare dell'antigene.

Dall'analisi dell'intero pattern di interazione peptide:DQB1 emerge che le differenze dell'allele predisponente e protettivo nei siti 57, 74 e 77 sono rispecchiate nella presenza di interazioni idrogeno in seguito al legame col peptide, sia esso virale o mielinico (Figure 3.22 e 3.23). Come discusso nella sezione 3.10, il ruolo di questi residui polimorfici nell'interazione col TCR, nell'attivazione dei linfociti T e nella genesi della malattie autoimmuni é stato oggetto di ampie speculazioni (Rudolph 2006). Anche in questa simulazione sono stati osservati dei legami idrogeno con D57

(P9) ed E74 (P4) come caratteristica dell'allele DQB1 predisponente, mentre nelle stesse posizioni per l'allele DQB1 protettivo è presente l'amminoacido polare Ser. Pertanto, le simulazioni suggeriscono che anche l'ambiente polare delle tasche di legame P4 e P9 sia importante nel conferire all'allele DQB1 la caratteristica di essere predisponente verso la SM.

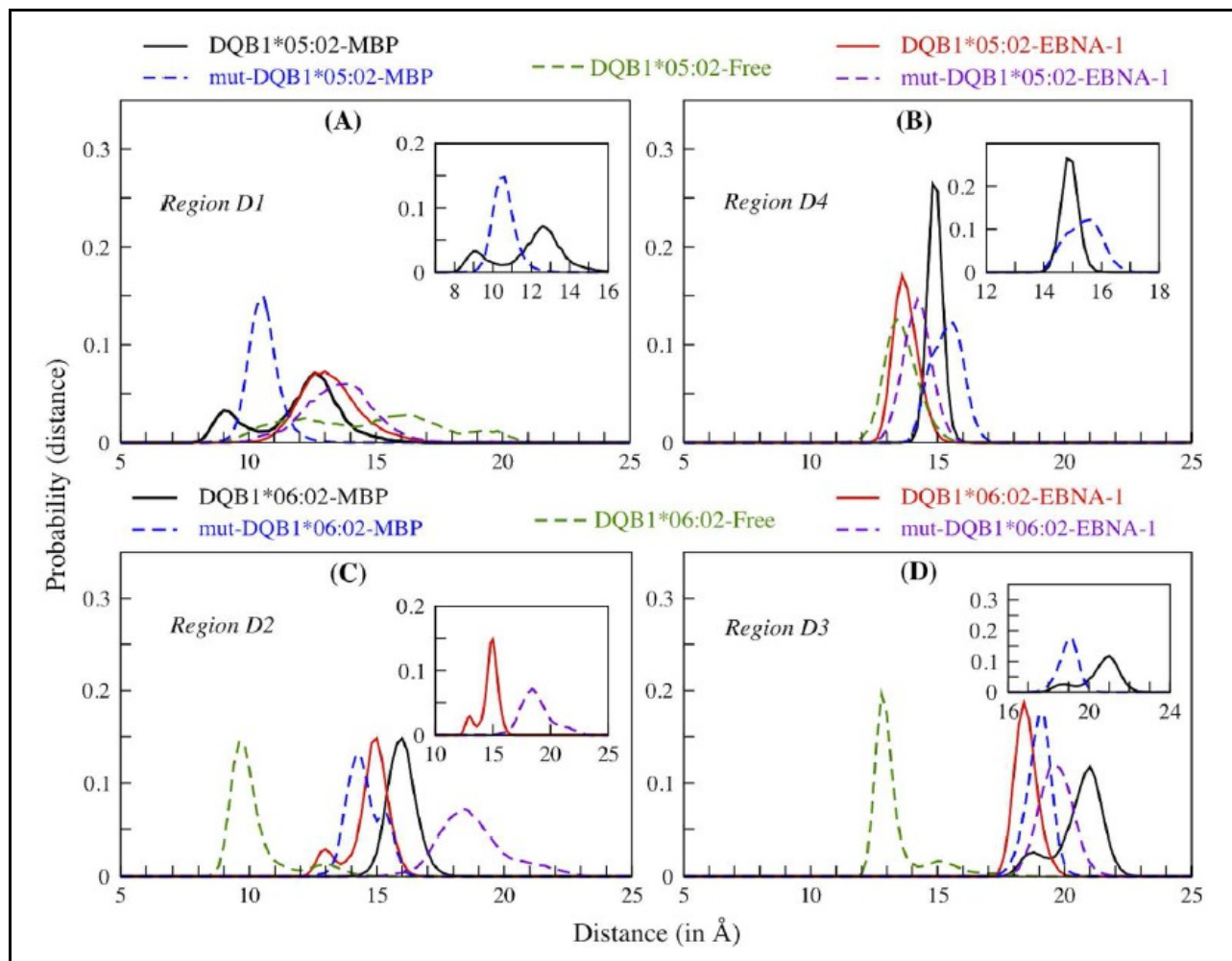


Figura 3.20. Istogrammi della larghezza del sito di legame nelle quattro aree D1, ..., D4, per gli alleli DQB1, wild-type e mutanti, in complesso con i due peptidi in analisi, e per gli alleli DQB1 non legati. Gli inserti mostrano gli istogrammi per specifici alleli DQB1 (in accordo con la convenzione dei colori riportata in Figura) nelle stesse regioni del grafico principale in cui sono inseriti.

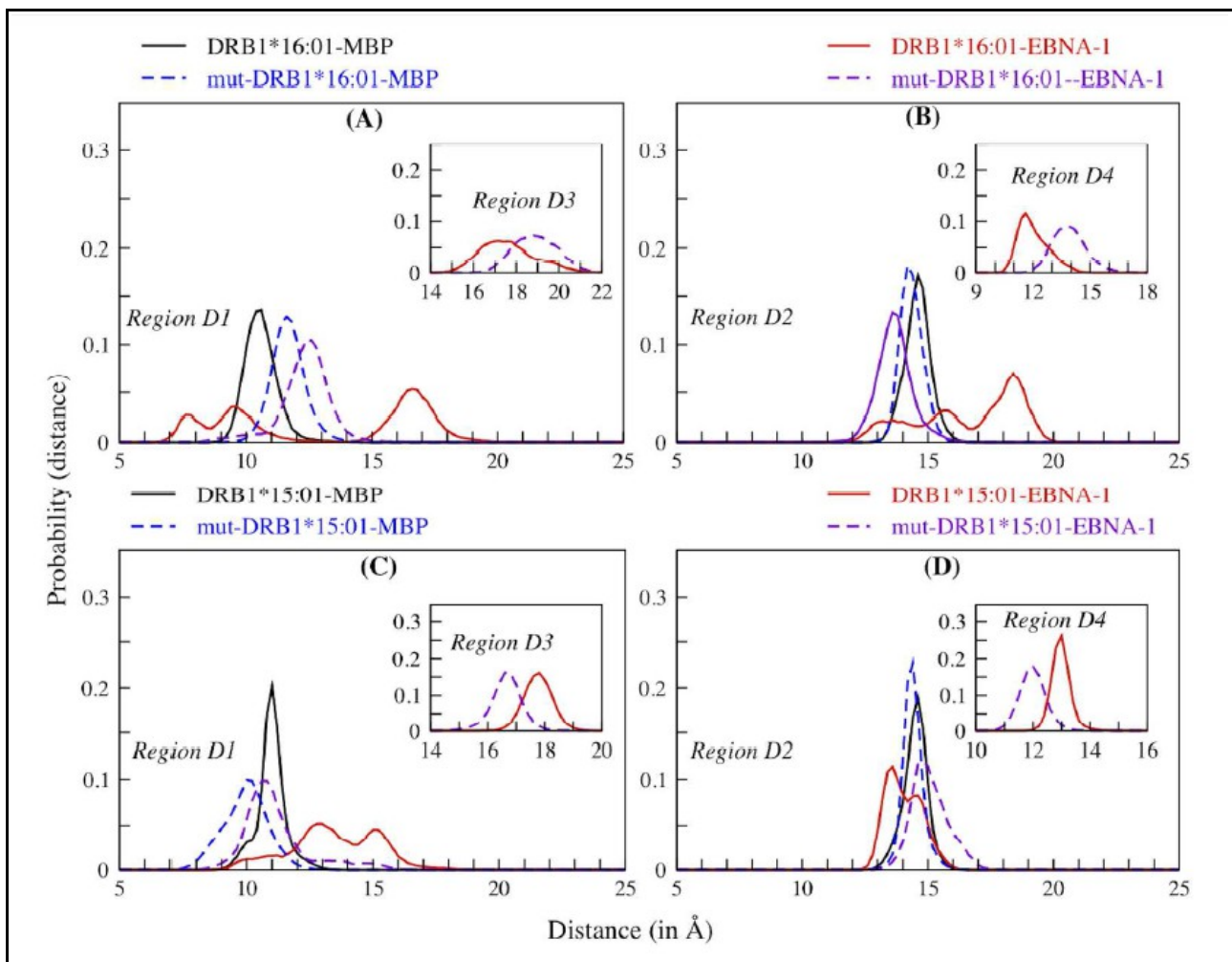


Figura 3.21. Istogrammi della larghezza del sito di legame nelle quattro aree D1, ..., D4, per gli alleli DRB1, wild-type e mutanti, in complesso con i due peptidi in analisi. Gli inserti mostrano gli istogrammi per i soli DRB1*16:01 wild-type e mutante, in complesso col peptide EBNA-1, nelle regioni indicate.

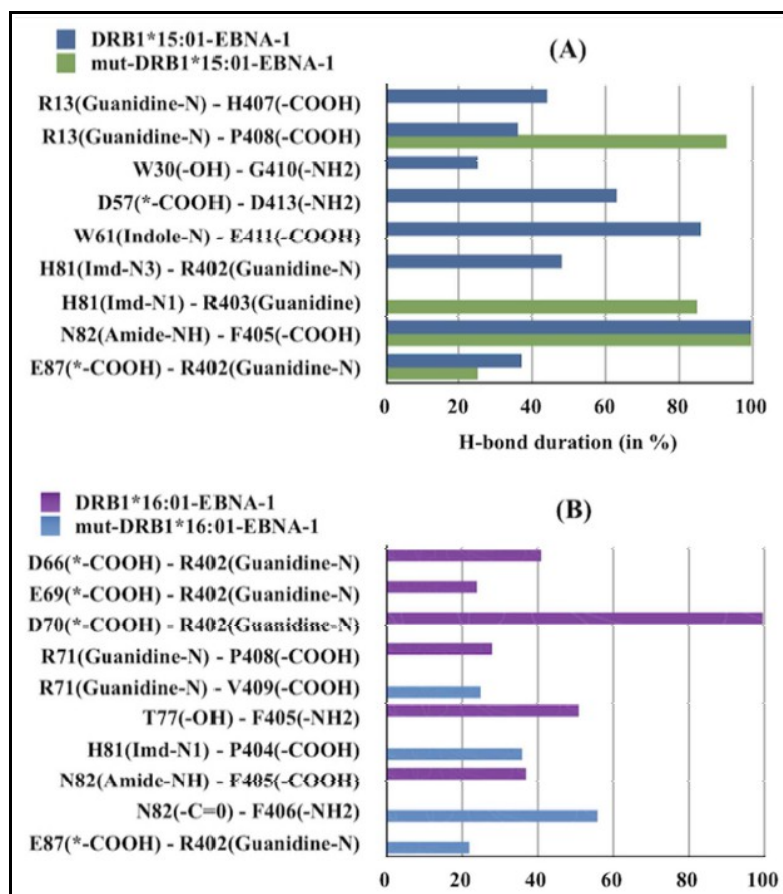


Figura 3.22. Legami idrogeno fra alleli DRB1 e mutanti, e peptide virale.

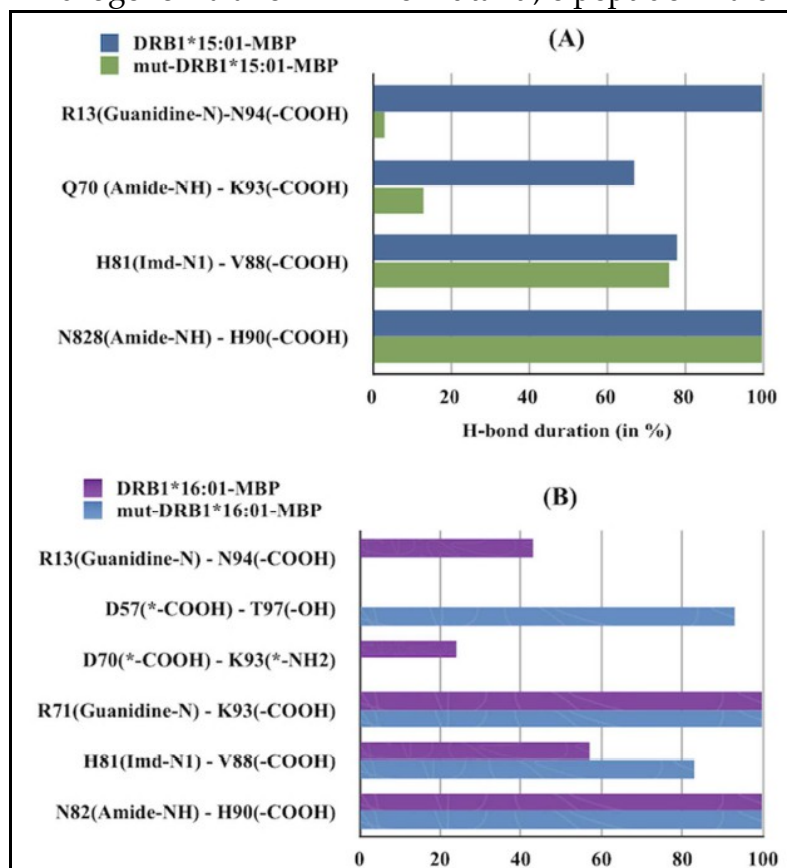


Figura 3.23. Legami idrogeno fra alleli DRB1 e mutanti, e peptide mielinico.

3.12. Limiti e prospettive dell'analisi svolta ✨

Il lavoro qui presentato sul meccanismo molecolare di riconoscimento antigenico effettuato dalla proteina MHC, deve essere ulteriormente esteso ad includere:

- a) il terzo alplotipo del serotipo DR2, DRB1*15:02-DQB1*06:01, in complesso con i peptidi derivati da MBP ed EBNA1;
- b) gli alplotipi DRB1-DQB1 non appartenenti al serotipo DR2, sul quale ci si è principalmente focalizzati;
- c) l'interazione con altri peptidi *self*, particolarmente PLP e MOBP (Kaushansky 2009, 2010), e *non self*, particolarmente MAP (Mameli 2014);
- d) l'interazione fra alplotipi all'interno dei genotipi DRB1-DQB1/DRB1-DQB1, come studiato recentemente dal solo punto di vista genetico (Cocco 2013);
- e) l'interazione fra complessi pMHC ed TCR, per alcuni sistemi pMHC selezionati, dato l'elevato numero dei gradi di libertà;

In prospettiva, si ritiene che il lavoro svolto possa risultare d'aiuto nel:

- a) comprendere alcuni dei meccanismi funzionali di base che contribuiscono alla malattia all'interno di una determinata predisposizione genetica;
- b) design di peptidi terapeutici (Steinman 2007), con un impatto minimo sul funzionamento fisiologico del sistema immunitario, ma in grado di agire selettivamente verso tessuti e proteine obiettivo nella fase di danno autoimmune;
- c) integrare informazioni fenotipiche (incluso il profilo metabolico), cliniche e modellistiche per pervenire a modelli individuali per la predizione del rischio, del decorso della malattia e della risposta ai farmaci.

3.13. Conclusioni ✨

Integrando tutte le ricerche svolte in questo lavoro nel contesto della SM, si può tratteggiare un quadro che colleghi le proprietà funzionali delle proteine immunologiche coinvolte e dei peptidi analizzati (derivati da MBP ed EBNA-1) con la caratteristica di predisporre o meno alla malattia, tramite un meccanismo di mimetismo funzionale-molecolare.

In dettaglio:

1) per gli alleli DRB1 del serotipo DR2:

- il polimorfismo del DRB1 86 nella tasca di legame P1 svolge un ruolo fondamentale (Gly = protettivo, Val = predisponente);
- l'allele predisponente è segnato da un legame con entrambi i peptidi che ha delle caratteristiche (energia di legame, RMSD sito, reti interazioni, configurazione chimica) molto simili;
- l'allele protettivo è caratterizzato da un robusto legame col peptide mielinico, mediato da reti di interazione uniche, e da un legame con il peptide virale molto debole;
- l'allele predisponente si caratterizza per una diversa configurazione di legame con i due peptidi analizzati limitata alla regione D1 (P1);
- l'allele protettivo invece ha differenze significative, nel legare i due peptidi, estese a tutto il sito di legame, ed in particolare alle regioni D1 e soprattutto D2 (P2);
- queste osservazioni sono state confermate a livello di riconoscimento anticorpale da un lavoro di immunologia sperimentale (Mameli 2014).

2) Per gli aplotipi DRB1-DQB1 del serotipo DR2:

- il polimorfismo del sito DQB1 38, vicino alla tasca di legame P9, svolge un ruolo fondamentale (Val = protettivo, Ala = predisponente);
- l'importanza dei polimorfismi DRB1 86 e DQB1 38 nella SM è confermata da studi di mutazione virtuale, in accordo con la letteratura esistente;
- il DQB1 contribuisce indipendentemente ed in sinergia col DRB1 nel rendere il meccanismo di riconoscimento antigenico maggiormente efficace;
- il DQB1 contribuisce dunque indipendentemente al rischio di sviluppare la malattia;
- l'aplotipo protettivo opera una interazione differenziata con peptide virale e mielinico, sia attraverso il DRB1 (discriminazione elevata e affinità intermedia per i due peptidi) che il DQB1 (discriminazione intermedia e affinità elevata);
- l'aplotipo predisponente presenta una abilità di discriminazione dei due peptidi minore, con una affinità maggiore verso il peptide virale;
- tessuti con diversi profili di espressione delle proteine DRB1 e DQB1 potrebbero attuare meccanismi di riconoscimento antigenico distinti, con diversi esiti di attivazione dei linfociti T;
- il ruolo del DRB1 potrebbe essere più rilevante nella fase in cui il peptide derivato dalla proteina virale EBNA-1 viene presentato nei tessuti periferici; mentre il ruolo del DQB1 potrebbe essere più importante nella seconda fase, dell'infiammazione autoimmune nel SNC;
- queste osservazioni sono coerenti con la restrizione promiscua dimostrata di recente per la SM (Sospedra 2006) e con l'attivazione di diversi profili di citochine da parte di alleli DRB1 e DQB1 predisponenti (Kaushansky 2012).

3) Per gli alplotipi DRB1 del gruppo "non-DR2" (DRB1-DQB1: *13:03-*03:01, *04:05-*03:01, *03:01-*02:01, *08:01-*03:01, *14:01-*05:03):

- la distinzione fra alplotipi protettivi e predisponenti è legata ai polimorfismi DRB1 9 (W/E) e 133 (L/R), con le differenze funzionalmente più rilevanti nelle posizioni DRB1 74 (P4) e 57 (P9);
- le caratteristiche peculiari del DRB1*14:01, unico protettivo del gruppo, possono essere spiegate dalle differenze strutturali e di legame col peptide mielinico nelle tasche di legame P4 e P9, rispetto agli altri alleli predisponenti del gruppo;
- la tasca P4 (siti 70-74) è stata studiata in altri contesti, ma collegata in questo lavoro alla predisposizione verso la SM, dove assume nel gruppo in analisi un ruolo di maggior rilievo rispetto alla P9, già ampiamente studiata;
- si ipotizza un meccanismo nel quale le due tasche P4 e P9 agiscano in sinergia: l'allele protettivo lega il peptide mielinico più debolmente in P9 e con maggiore affinità in P4, rispetto all'allele predisponente, combinando in tal modo la funzione di riconoscimento antigenico (legame col peptide mielinico in P4, P9) con quella di internalizzare un punto di ancoraggio primario per il successivo riconoscimento da parte del TCR;
- queste osservazioni sono in linea con quanto ipotizzato da altri autori per quanto concerne altre malattie autoimmuni (Wucherpfenning 1995, Ou 1998, Fu 1995), la SM (Gilbert 2003, Greer 2005) e l'attivazione del linfocita T (Wucherpfenning 1995).

4. Modellistica dei meccanismi biologici del dolore, applicata alla Sindrome Fibromialgica

La sindrome fibromialgica é un condizione di malattia cronica, di eziologia sconosciuta, caratterizzata da dolore diffuso in varie sedi del corpo, unitamente ad altri sintomi di natura generica. L'insieme dei sintomi é piuttosto vasto e presenta notevoli sovrapposizioni con altre malattie, con conseguente difficoltà di diagnosi e trattamento. Al momento attuale non esistono test di laboratorio o marcatori biologici che possano confermare lo stato di malattia. Grazie alla caratterizzazione del profilo metabolico di una serie di pazienti selezionati, effettuata dai collaboratori medici e biochimici, l'approccio modellistico qui adottato é stato quello di studiare se alcuni dei lipidi sovra-rappresentati nel sangue dei pazienti fossero in grado di interagire, attivandolo, con un recettore deputato alla modulazione dei meccanismi biologici del dolore.

4.1. Introduzione: il sistema PAF / PAFR

Il Platelet Activating Factor (PAF) è un mediatore lipidico coinvolto in molte condizioni fisiologiche e patologiche (Sezione 2.3.1)(Ishii 1997, Gelbard 1994, Prescott 2000), che quindi richiede una fine regolazione quantitativa a livello ematico e nelle cellule. L'azione del PAF viene mediata dall'interazione col recettore PAFR (Figura 4.1 e 2.12), appartenente alla categoria dei recettori accoppiati alla proteina G (GPCR), ed espresso sulla membrana di svariati tipi di cellule (Ishii 2000, Nakamura 1991, Ye 1991). Il recettore PAFR è quindi in grado di dare inizio ad una vasta tipologia di eventi cellulari, svolgendo un gran numero di funzioni biologiche (Prescott 2000). Il sistema PAF:PAFR è coinvolto nella modulazione dei percorsi molecolari connessi alla segnalazione del dolore. In particolare, è stato visto che l'amministrazione di PAF in tessuti periferici, ad esempio la cute, aumenta la sensibilità al dolore nell'uomo e negli animali (Tsuda 2011). In maniera simile, l'iniezione di PAF nel midollo spinale produce allodinia tattile nei topi (Morita 2004), suggerendo dunque un legame col dolore di natura neuropatica in seguito ad un danno nervoso (Morita 2008). Il dolore neuropatico viene considerato una componente della sindrome fibromialgica (Dworkin 2005) e alcuni autori (Hasegawa 2010) propongono che l'attivazione del PAFR sia un evento rilevante nello sviluppo e mantenimento di allodinia tattile e nella produzione di citochine pro-infiammatorie, quali $TNF\alpha$ ed $IL-1\beta$.

4.2. La Sindrome Fibromialgica e l'approccio metabolomico

La sindrome fibromialgica (FMS) è una malattia cronica, caratterizzata da dolore diffuso e diagnosticata clinicamente dalla presenza di risposta dolorosa in 11 su 18 punti critici, secondo le prescrizioni dell'American College of Rheumatology (Wolfe 1990). Al dolore si accompagnano svariati altri sintomi di natura generica, quali

stanchezza, ansietà, depressione e disturbi della funzionalità intestinale, di recente inclusi nei criteri diagnostici (Wolfe 2010). Pertanto, la malattia rimane complessa da diagnosticare e da trattare, e in assenza di una comprensione della sua patogenesi è altresì difficile identificare bio-marcatore oggettivi che possano aiutare ad identificare gli individui a rischio o convalidare la diagnosi, per cui al momento attuale non esistono test clinici di laboratorio per la FMS. Attraverso tecniche combinate di spettrometria di massa e cromatografiche (LC-TOF/MS) sono stati analizzati i campioni di sangue di 22 pazienti e 21 controlli sani, riscontrando la presenza di molti composti lipidici, principalmente lisofosfoline, fosfoline e ceramidi. L'approccio metabolomico ha già dato prova di essere affidabile e predittivo in ambito clinico (Weljie 2007, Young 2013, Griffin 2011, Nicholson 2012) e in questa ottica è stato sperimentato a livello clinico, probabilmente per la prima volta, nel contesto della sindrome fibromialgica. L'analisi statistica multivariata ha poi consentito di isolare da questi metaboliti quelli maggiormente in grado di discriminare i pazienti dai controlli, che sono risultati essere le lisofosfoline PC(14:0/0:0) e PC(16:0/0:0) (Figure 2.13 e 2.14, Sezione 2.3.2), denotate qui per brevità PC(14:0) e PC(16:0).

4.3. Modellistica molecolare ✨

Come descritto nella Sezione 2.3.2, attraverso una simulazione di docking, sono stati dapprima evidenziati i siti di legame più rilevanti di ciascuno dei tre lipidi PC(14:0), PC(16:0) e PAF (PC(0-16:0/2:0)) nei confronti del recettore PAFR (Sezione 2.3.1). I risultati ottenuti suggeriscono che in tutti e tre i casi il sito di legame si trovi all'interno del recettore (Figura 4.2), in accordo coi dati e le simulazioni (Gui 2007). I risultati del docking mostrano inoltre (Tabella 4.1) come il legando PC(14:0) abbia una energia di legame favorevole rispetto agli altri due, mentre PC(16:0) abbia una energia di legame molto simile a quella del legando fisiologico PAF.

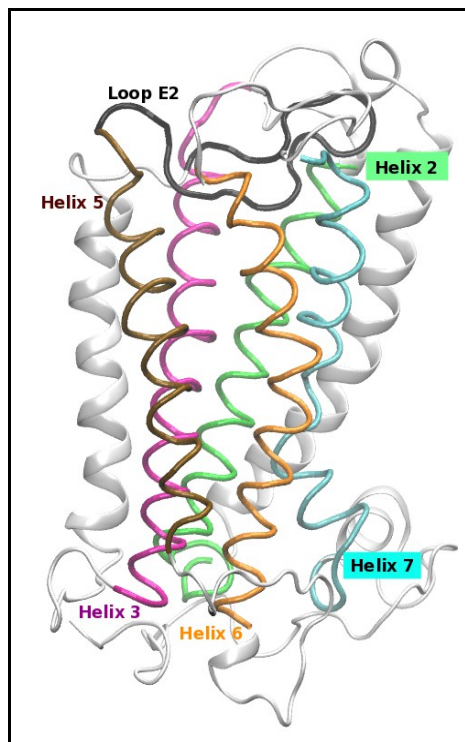


Figura 4.1. Struttura 3D del PAFR e localizzazione degli elementi di struttura secondaria.

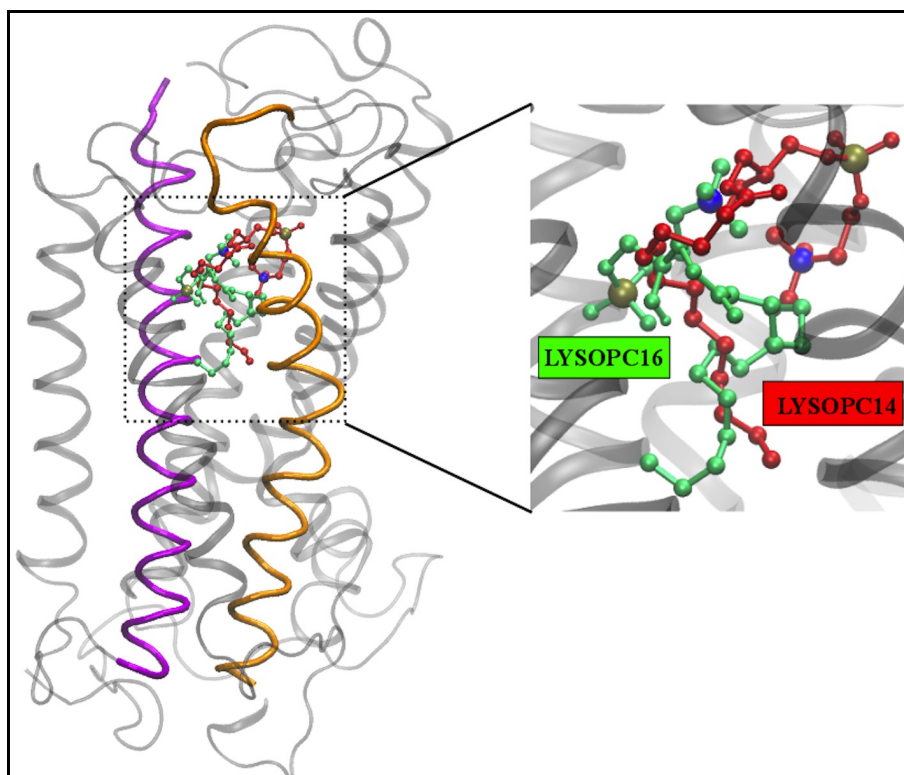


Figura 4.2. Pose migliori ottenute tramite docking per il legame fra PC(14:0), PC(16:0) ed il PAFR.

Come descritto nelle Sezioni 2.3.2 e 2.3.3, i complessi migliori ottenuti dal protocollo di docking sono stati utilizzati come strutture iniziali per le simulazioni di DM (Figura 4.2). Dai valori della RMSD di ciascun sistema (Tabella 4.1, Figura 4.3), si nota una differenza modesta fra PC(14:0) e PC(16:0), entrambi caratterizzati da un istogramma di RMSD con un unico picco, corrispondente ad una unica e ben definita configurazione chimica di legame. Il grafico di RMSD del PAF risulta invece spostato verso sinistra di $\sim 2\text{\AA}$, rispetto agli altri due. Dai grafici si evince come i due legandi esogeni, PC(16:0) e PC(14:0), siano leggermente più flessibili di quello endogeno PAF, ma nel complesso tutti e tre i legandi mostrano simili configurazioni chimiche di legame all'interno del recettore.

Legando	ΔG (kcal/mol) MD	ΔG (kcal/mol) Docking	RMSD (\AA)
PC(14:0)	-11.0 ± 0.4	-7.1	1.8 ± 0.4
PC(16:0)	-11.2 ± 0.4	-5.1	1.9 ± 0.4
PAF	-11.1 ± 0.4	-7.0	1.3 ± 0.4

Tabella 4.1. Energie di legame dei vari complessi PC:PAFR ottenute tramite MD e Docking. L'ultima colonna mostra anche la RMSD media durante la simulazione.

L'analisi dei legami idrogeno mostra una interazione duratura che coinvolge il singolo residuo R172 (Figura 4.4) nel loop E2 (Figura 4.1) del PAFR ed entrambi i legandi esogeni. Il legando PC(14:0) è inoltre coinvolto in una interazione idrogeno coi residui dell'elica 2, 6 ed il loop E2, mentre il PC(16:0) con residui dell'elica 3, 5, 6 ed il loop E2. Questa rete di interazioni è coerente con le energie di legame ottenute tramite DM (Tabella 4.1), che mostrano come, tenuto conto dell'errore di simulazione, tutti e tre i legandi abbiano complessivamente affinità molto simili per il PAFR. La differenza di ordinamento dei valori ottenuti tramite MD e docking è da

attribuirsi al fatto che nel docking i cambiamenti conformazionali del recettore non sono ammessi (si veda la Sezione 2.1).

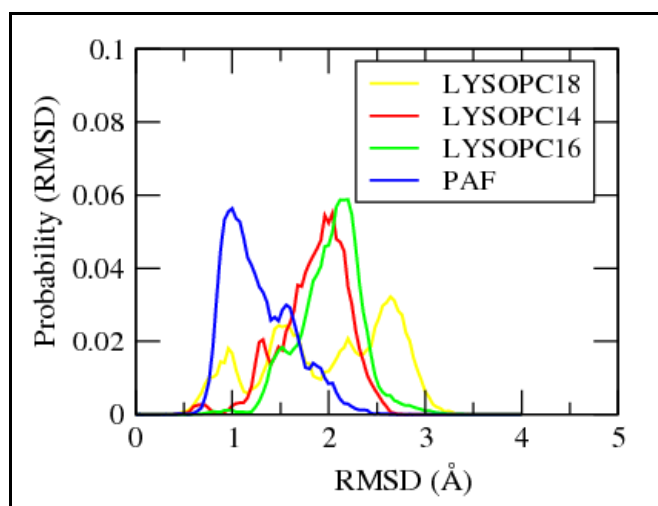


Figura 4.3. Istogramma della RMSD per le due PC in analisi (indicate come LYSOPC14 e LYSOPC16) ed il PAF in legame col PAFR. Il LYSOPC18 è rimasto in quanto oggetto di analisi ulteriori da sviluppare, ma non va' considerato in questo lavoro.

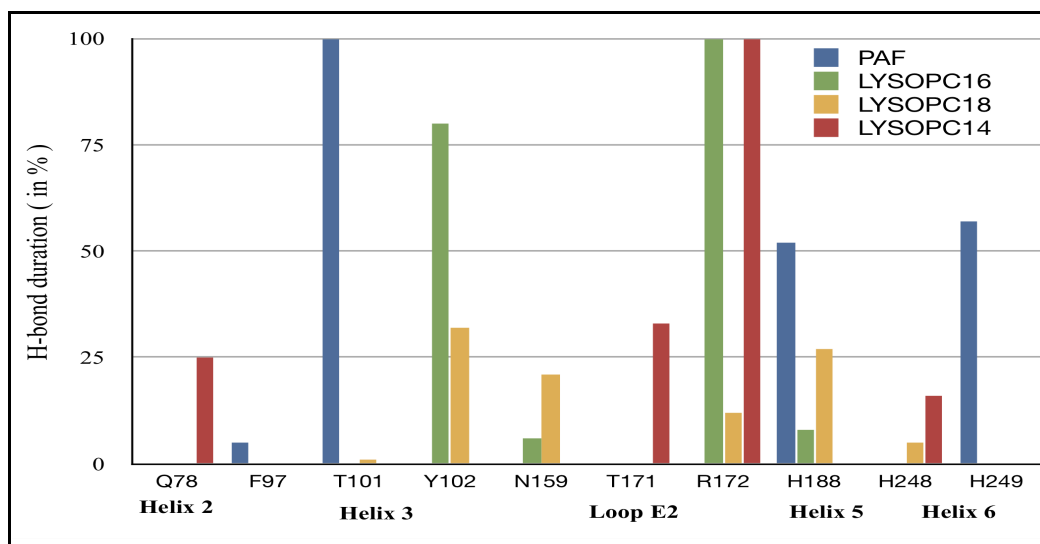


Figura 4.4. Rete di interazioni idrogeno fra PAFR e le tre PC in analisi indicate come (LYSOPC14, LYSOPC16 e PAF). La riga inferiore indica anche la localizzazione dei residui all'interno degli elementi della struttura secondaria del recettore. Per il LYSOPC18 vale quanto detto nella didascalia di Figura 4.3.

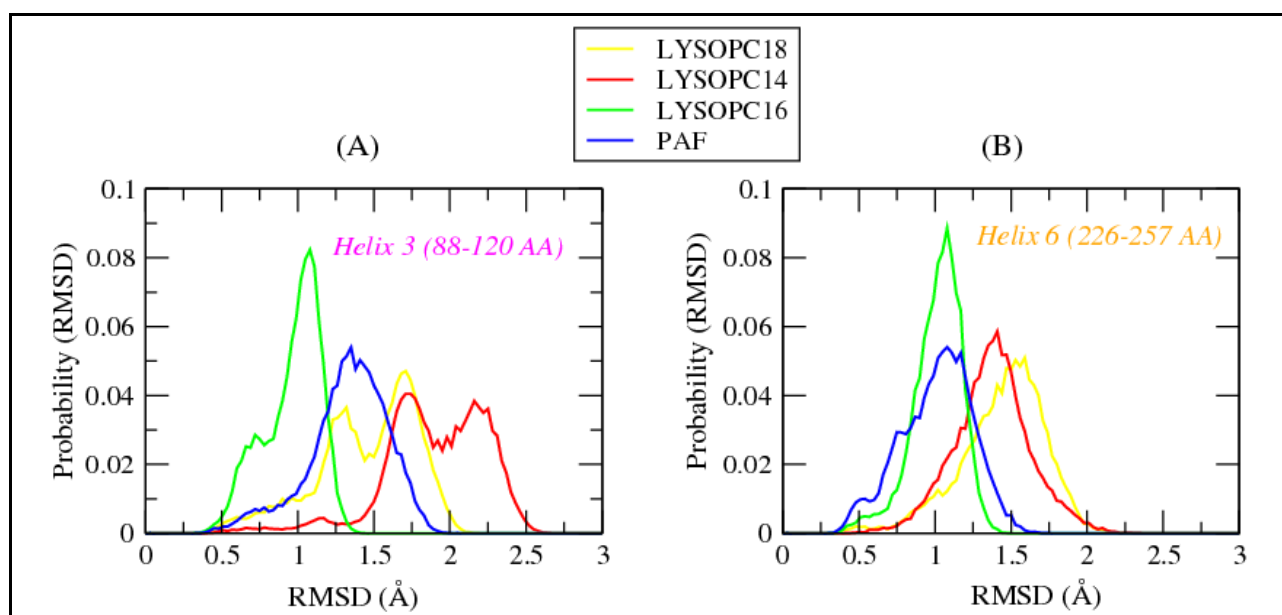


Figura 4.5. Istogrammi di RMSD per le Eliche 3 (A) e 6 (B) del PAFR in legame con le PC in esame (LYSOPC14, LYSOPC16 e PAF). Per il LYSOPC18 vale quanto detto nella Figura 4.3.

Studi precedenti, sia sperimentali (Ishii 1997, Fais 2013) che di modellistica (Gui 2007), suggeriscono un ruolo rilevante nel legame coi legandi per i residui del recettore His 188 (elica 5) ed His 249 (elica 6). Tale osservazione risulta confermata dalle simulazioni qui effettuate, che mostrano His 188 coinvolto in un legame idrogeno sia col PAF che col PC(16:0), mentre His 249 forma un legame idrogeno col solo PAF (Figura 4.4). La variazione di RMSD osservata per l'elica 6, nel caso dei legami con tutti e tre i legandi, è più bassa di quanto riportato precedentemente in letteratura per il legame con un'altra molecola antagonista del PAFR (Gui 2007). Comunque le simulazioni mostrano un istogramma di RMSD per l'elica 6 con un identico picco nel caso di PAF e PC(16:0) (Figura 4.5 B). Il secondo legando esogeno PC(14:0), mostra invece valori più elevati di RMSD nella elica 3 nella seconda parte della simulazione (Figura 4.5 A), caratteristica che rappresenta l'inizio di un

cambiamento conformazionale che potenzialmente è collegato all'attivazione del recettore.

In sommario: il fatto che tutti e tre i legandi i) abbiano configurazioni chimiche di legame col recettore simili; ii) possiedano energie di legame molto vicine, e iii) mostrino i primi segni di cambiamenti allosterici che precludono l'attivazione del recettore, dimostra che tutti e tre i legandi hanno simili potenzialità di indurre la configurazione attiva del recettore PAFR, dando così inizio alla cascata di segnalazione attraverso l'accoppiamento del dominio intracellulare del recettore con la proteina G.

4.4. Conclusioni ✨

La presenza delle due PC sovra-rappresentate nel profilo metabolico dei pazienti, PC(14:0) e PC(16:0), ed i risultati della modellistica, autorizzano ad ipotizzare che i) la frammentazione ossidativa dei lipidi operata dalle cellule generi lisofosfocoline in abbondanza, ii) alcune delle quali dotate di caratteristiche chimiche simili a quelle del PAF, che dunque possono agire come bio-attivatori del PAFR in maniera simile al PAF.

Questa ipotesi è supportata da alcuni lavori (Oestvang 2011) che hanno mostrato come le lisofosfocoline ed il PAF siano in grado di stimolare: i) il rilascio di acido arachidonico in una specifica linea cellulare umana (THP-1); ii) l'attività dell'enzima fosfolipasi D (PLD), tramite meccanismi indipendenti; iii) l'attività della fosfolipasi citosolica A2 (cPLA2), sempre tramite meccanismi indipendenti.

Pertanto le due PC esogene qui analizzate, PC(14:0) e PC(16:0), hanno un ruolo potenziale nelle manifestazioni cliniche della sindrome fibromialgica, e rappresentano una nuova categoria di potenziali marcatori biologici ed obiettivi terapeutici della patologia.

5. Modellistica della proteina CASQ1, applicata al funzionamento muscolare

La proteina CASQ1 è in grado di legare sino a circa 70-80 ioni Calcio, ma il meccanismo di riconoscimento e legame, i siti di legame ed i cambiamenti strutturali indotti dal legame, non sono stati ancora sufficientemente caratterizzati a livello molecolare. In questo studio viene costruito un modello di dimero della proteina CASQ1, col quale vengono effettuate delle DM a varie concentrazioni di ioni Calcio. La dinamica mostra che: i) esistono una varietà di siti che presentano sia affinità che geometria diverse, alcuni dipendenti dalla concentrazione degli ioni; ii) il dominio C-terminale arriva a saturazione prima del resto della proteina; iii) il modello proposto per il dimero risulta maggiormente stabilizzato da interazioni idrofobiche all'aumentare della concentrazione di Calcio; iv) la proteina assume all'aumentare della concentrazione una struttura via via più compatta e stabile. I risultati ottenuti consentono una migliore interpretazione dei dati fisiologici relativi al funzionamento dei tessuti muscolari, consentono di ipotizzare un modello di polimerizzazione ed interazione con gli ioni Calcio della proteina, ed infine rappresentano il primo passo per lo studio della variante 'cardiaca' della proteina, con prospettive interessanti di applicazioni al design di molecole per la modulazione del battito cardiaco.

5.1. Stabilità della proteina al variare della $[Ca^{2+}]$ ✨

L'esamero utilizzato per costruire il modello dimerico della proteina CASQ1 è mostrato in Figura 5.1 (si veda anche la sezione 2.4.1). Come usuale la stabilità della proteina dimerica in complesso con gli ioni Calcio è stata studiata attraverso i grafici di RMSD e l'entropia configurazionale del sistema (Tabella 1). I dati mostrano una diminuzione progressiva sia della RMSD che dell'entropia all'aumentare della $[Ca^{2+}]$, suggerendo quindi che il legame con gli ioni abbia un effetto di progressiva stabilizzazione della struttura. Questo è in accordo con i dati cristallografici, che mostrano una regione piuttosto estesa senza alcuna struttura secondaria ben definita (sezione 2.4.1).

Ioni Calcio	RMSD degli atomi del backbone (Å)			Entropia (J/molK)
	Monomero 1	Monomero 2	Dimero	C-Terminal (residui 350-364)
120	2.8+/-0.5	2.7+/-0.7	3.3+/-0.7	1715
80	3.2+/-0.7	2.8+/-0.4	3.5+/-0.7	1705
40	3.0+/-0.8	3.9+/-0.8	4.9+/-1.2	1859
20	2.7+/-0.5	2.9+/-0.6	4.5+/-1.4	1896
0	3.5+/-0.6	4.0+/-0.9	5.7+/-1.4	1875

Tabella 5.1. RMSD ed Entropia configurazionale di ciascun monomero e del dimero di CASQ1 al variare della concentrazione di ioni Calcio nel sistema.

Si è successivamente realizzata una mappa di carica superficiale per ciascun sistema alle varie concentrazioni ioniche, che evidenzia come la superficie della proteina, inizialmente piuttosto acida, venga progressivamente neutralizzata all'aumentare della concentrazione di ioni Calcio (Figura 5.2). Si osservi come sino ad una

concentrazione corrispondente a circa 40 ioni Calcio, il sistema rimanga essenzialmente acido con alcuni effetti di neutralizzazione di carica superficiale a carattere locale, e solo a concentrazioni superiori avvenga una transizione verso una neutralizzazione su scala spaziale maggiore.

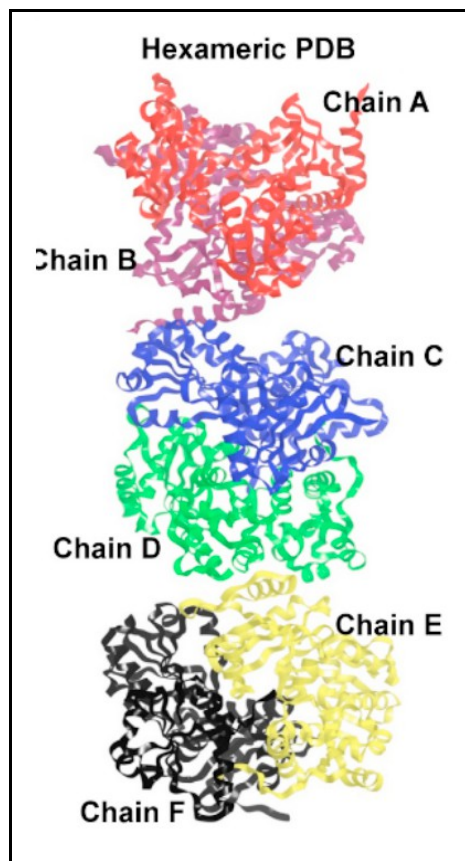


Figura 5.1. La struttura dell'esamero utilizzata per realizzare un modello di dimero della proteina, estraendo le catene C (Monomero 1) e B (Monomero 2). PDB id 3uom.

Si è pertanto analizzato il numero di ioni Calcio che si legano alla proteina durante la simulazione al variare del numero totale di ioni presenti nel sistema (Figura 5.3). Si osserva che, mentre la superficie complessiva della proteina viene progressivamente neutralizzata senza arrivare a saturazione (per l'intervallo di concentrazioni studiato)(Figura 5.3 B), la regione C-terminale (dominio CAS, si veda la sezione 2.4.1), ha un rapido aumento del numero di interazioni con gli ioni proprio a partire da circa 40 ioni, e raggiunge un plateau a circa 80 ioni (Figura 5.3 A). La saturazione 'precoce' della regione C-terminale è quindi in grado di spiegare l'andamento della

neutralizzazione superficiale osservata in Figura 5.2, ed è l'indizio di un cambiamento conformazionale del dominio CAS che consente la transizione osservata nella capacità di legare gli ioni da parte dell'intera proteina.

Queste osservazioni suggeriscono che i siti di legame degli ioni Calcio sulla proteina debbano essere di differente affinità, spazialmente organizzati, dipendenti dalla concentrazione di ioni Calcio, ed in grado di indurre cambiamenti conformazionali della proteina. Il processo delineato può portare alla neutralizzazione progressiva e sistematica osservata, sulla superficie, con la proteina che arriva eventualmente ad assumere una forma maggiormente compatta e strutturata (globulare), tale da consentire il processo di polimerizzazione.

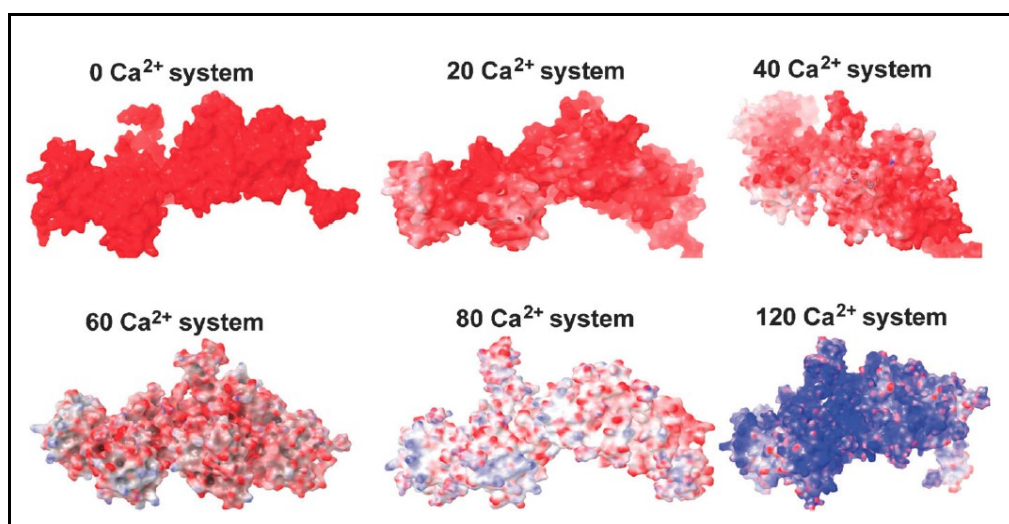


Figura 5.2. Mappe di carica superficiale del dimero CASQ1 alle varie concentrazioni ioniche.

5.2. Organizzazione e struttura dei siti di legame ionico al variare della $[Ca^{2+}]$ ✪

Per verificare quanto suggerito dalle precedenti analisi, si è pertanto proceduto alla caratterizzazione dettagliata dei siti di legame degli ioni Calcio. La struttura depositata nella banca dati PDB, ed utilizzata come modello (id 3uom), mostra solo 15 siti di legame per gli ioni Ca^{2+} , mentre la dinamica mostra come ciascuno dei due

monomeri sia in grado di legare sino a circa 55 ioni, per un totale di 110 per il dimero, su 120 disponibili alla massima concentrazione analizzata (Figura 5.3, 8 ioni nella regione C-terminale e poco più di 45 nel resto della proteina). I siti di legame sono stati identificati come descritto nella Sezione 2.4.4, essenzialmente valutando i legami (*non bonded*) di Van der Waals ed elettrostatici fra i residui del sito e gli ioni Calcio, rilevati per un tempo sufficientemente prolungato sul totale della simulazione, ed all'interno di una sfera di interazione di raggio dato.

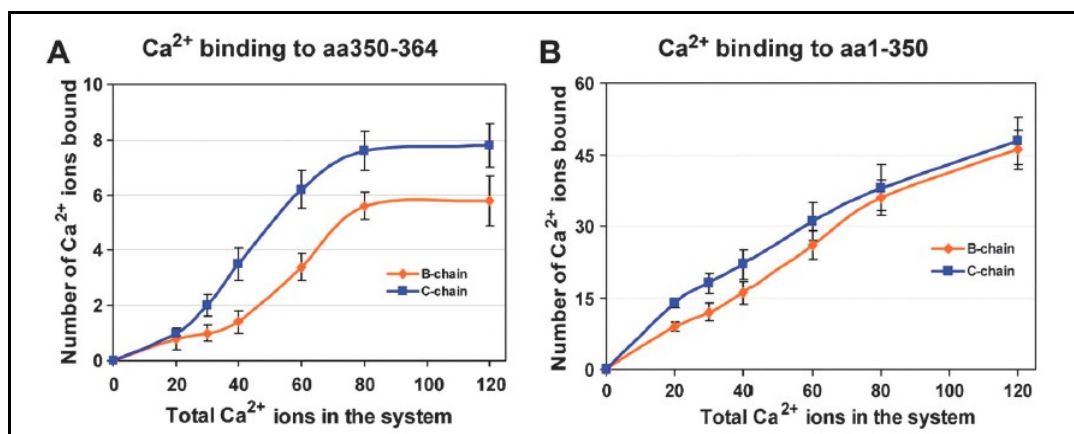


Figura 5.3. Numero di ioni Calcio legati alla proteina nel suo complesso (B) ed alla regione C-terminale (A), al variare della concentrazione di ioni nel sistema.

L'analisi ha confermato tutti e 15 i siti della struttura cristallografica e ne ha rilevato molti altri, di cui una selezione rappresentativa è mostrata in Tabella 5.2. A seconda dell'entità dell'energia di interazione (Tabella 5.2) i siti sono stati classificati ad alta e bassa affinità, corrispondenti rispettivamente ad energie di legame < -400 kcal/mol e > -400 kcal/mol, su tutti i sistemi alle varie concentrazioni ioniche. I siti attivi principalmente alle concentrazioni $[Ca^{2+}]$ intermedie, sono stati classificati come siti di affinità intermedia. I siti sono stati ulteriormente classificati come dipendenti o indipendenti dalla $[Ca^{2+}]$, a seconda che l'energia di legame cambi in maniera significativa o meno al variare della $[Ca^{2+}]$.

I siti di legame più interessanti sono proprio quelli dipendenti dalla $[Ca^{2+}]$, ad esempio i siti 5, 6 e 7 in Tabella 5.2, che mostrano affinità crescenti all'aumentare

della concentrazione ionica, con affinità quasi nulla nel sistema con 20 ioni Calcio e massima nel sistema con 120 ioni. Lo studio delle geometrie di legame di questi siti mostra una configurazione che migliora progressivamente all'aumentare della concentrazione $[Ca^{2+}]$, dimostrando quindi che questi siti vengono indotti in seguito al legame della proteina con gli ioni Calcio (Figura 5.4). I dati della Tabella 5.2 sono in accordo con un modello che preveda prima il legame degli ioni Ca^{2+} ai siti ad alta affinità ed indipendenti dalla $[Ca^{2+}]$, il che induce una serie di progressive modifiche strutturali della proteina, che portano quindi i residui dei siti ad alta affinità $[Ca^{2+}]$ dipendenti ad attuare geometrie di legame maggiormente favorevoli per i successivi legami ionici.

N. di ioni Ca^{++}	Alta affinità $[Ca^{2+}]$ -indipendenti				Alta affinità $[Ca^{2+}]$ -dipendenti			Intermedia affinità $[Ca^{2+}]$ -dipendenti			Bassa affinità $[Ca^{2+}]$ -indipendenti	
	1	2	3	4	5	6	7	8	9	10	11	12
	E39 E54 D93	E102 E169	E194 D196	D356 D357	E39 P40 E42	D259 D261	D328 E331 D332	D306 S308	E326 E158	E319 E350 E354	E66	E199 T229
120	-495	-458	-371	-506	-481	-389	-905	-	-279	-313	-	-256
80	-560	-533	-386	-590	-355	-291	-785	-221	-274	-818	-198	-
40	-471	-483	-375	-370	-158	-214	-312	-230	-545	-599	-	-207
20	-479	-551	-393	-530	-	-	-	-	-480	-	-195	-276

Tabella 5.2. Siti di legame rappresentativi per gli ioni Calcio identificati dalla DM, al variare del numero di ioni complessivamente presenti nel sistema.

I siti ad affinità intermedia hanno un comportamento piuttosto particolare: ad esempio il sito 10 (Tabella 5.2), come si vede in Figura 5.5, presenta i residui con una orientazione inappropriata per il riconoscimento degli ioni nel sistema con concentrazione corrispondente a 20 Ca^{2+} , mentre nel sistema con 80 Ca^{2+} tutti e tre i residui acidi si dispongono col corretto orientamento per una interazione ottimale con gli ioni Ca^{2+} , ed infine concentrazioni superiori portano alla distruzione della geometria ottimale ed al rapido declino dell'energia di interazione.

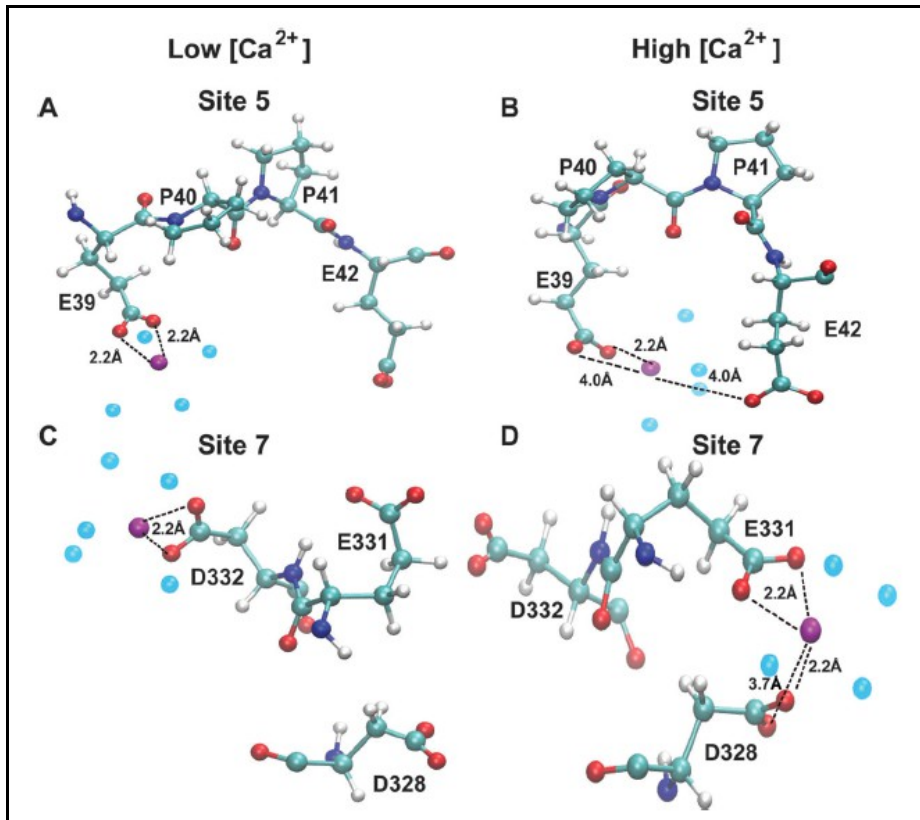


Figura 5.4. Siti di legame ad alta affinità e dipendenti dalla $[Ca^{2+}]$.

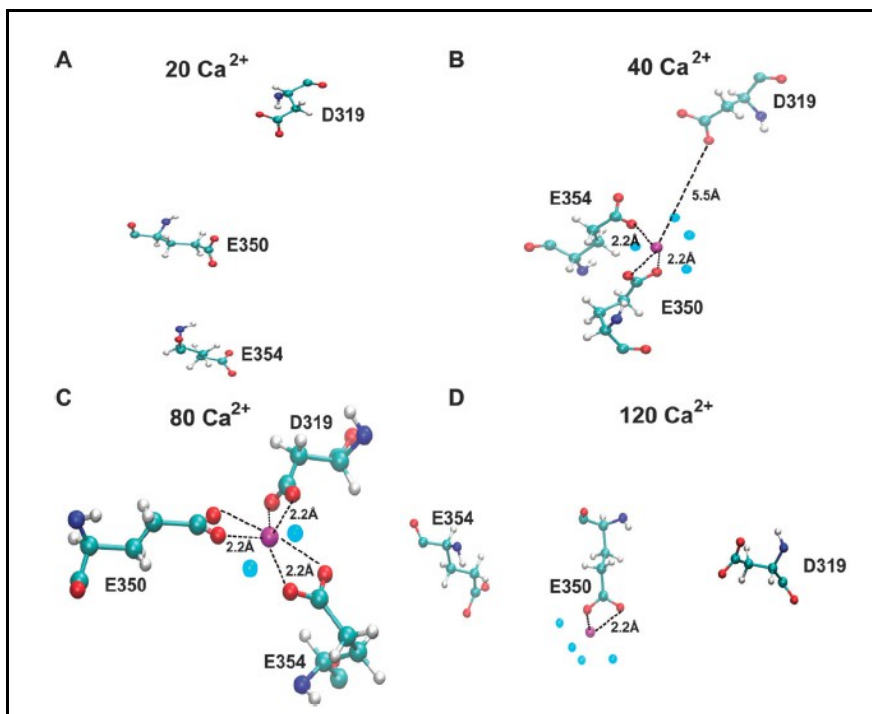


Figura 5.5. Sito di legame ad affinità intermedia, alle varie concentrazioni di Ca^{2+} .

Complessivamente, sono stati trovati 25 siti di legame dipendenti dalla $[Ca^{2+}]$, di cui i più rappresentativi sono riportati in Tabella 5.2. Infine, sono stati osservati svariati siti a bassa infinità $[Ca^{2+}]$ -indipendenti, tipicamente formati da un residuo acido (Asp o Glu) da solo o in combinazione con un qualunque altro residuo di supporto.

5.3. Interazioni proteina-proteina che stabilizzano il dimero ✨

Osservando l'interfaccia fra i due monomeri si trova una unica interazione idrogeno, mentre il numero delle interazioni idrofobiche aumenta progressivamente all'aumentare degli ioni Ca^{2+} (Figura 5.6).

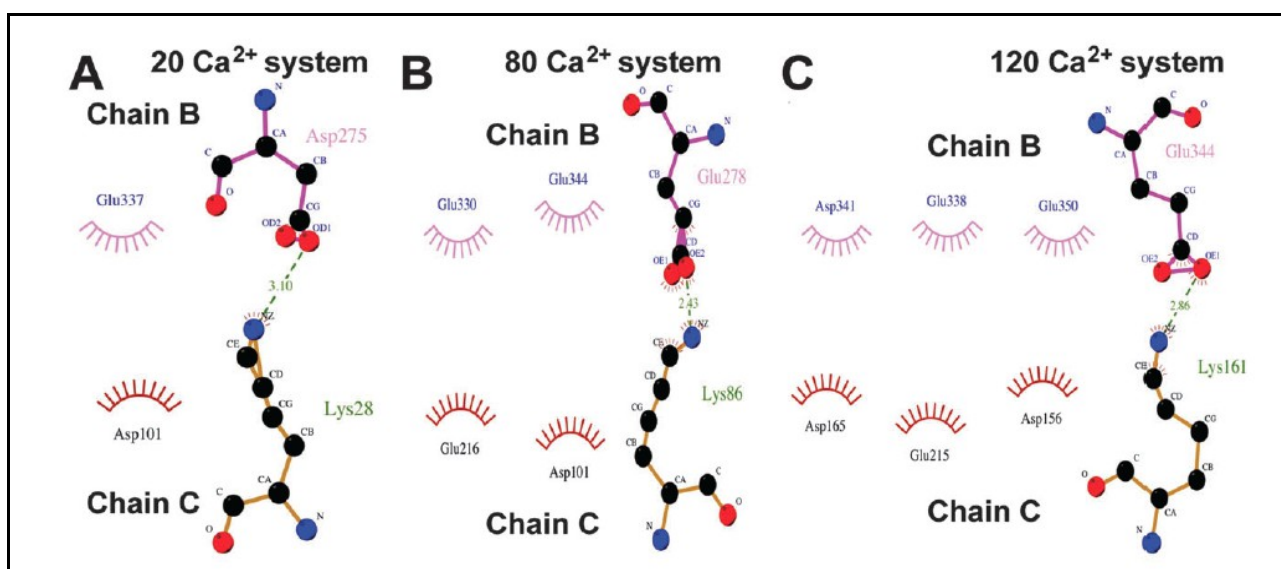


Figura 5.6. Interazioni idrofobiche all'interfaccia monomero-monomero, all'interno del dimero della proteina CASQ1, all'aumentare della $[Ca^{2+}]$.

Tutte le interazioni idrofobiche avvengono fra residui acidi (Gln ed Asp), ma in genere senza alcun residuo di interazione specifico conservato al variare della $[Ca^{2+}]$, formando dunque una rete di interazioni idrofobiche promiscua. Questa osservazione suggerisce che proprio la neutralizzazione superficiale dei residui acidi, mediata dagli ioni Ca^{2+} , consenta ai residui di stabilire interazioni idrofobiche stabili.

5.4. Alterazioni del dominio CAS ✨

Il dominio C-terminale, CAS, acquista una struttura 3D via via più compatta all'aumentare della $[Ca^{2+}]$. L'analisi delle traiettorie di DM mostra infatti che il CAS assume nel tempo tre tipologie di configurazione: lineare, intermedia e compatta (Figura 5.7).

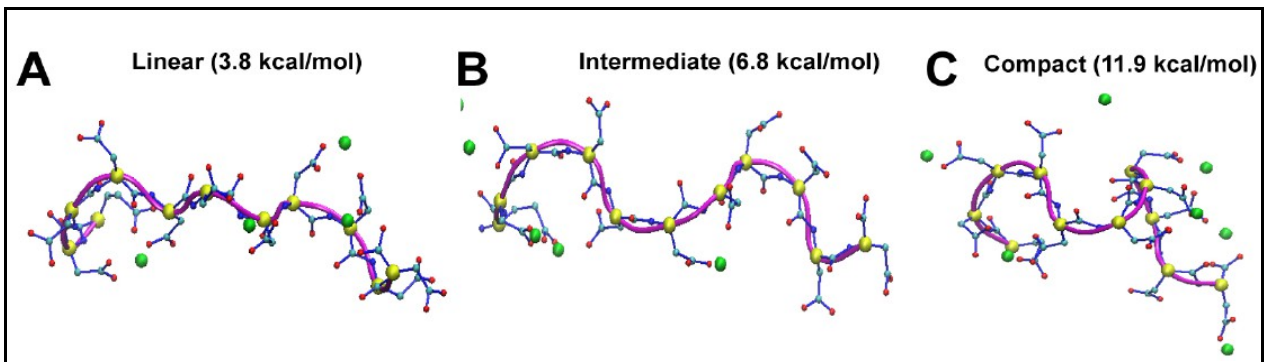


Figura 5.7. Configurazioni della CAS osservate durante la MD: lineare, intermedia e compatta. L'energia viene valutata tramite il numero di legami idrogeno formati: 2-3 per la configurazione lineare, 4-5 per l'intermedia e 6-8 per la compatta.

La frequenza di ciascuna conformazione (popolazione) durante la simulazione di MD varia a seconda della $[Ca^{2+}]$. L'analisi delle traiettorie di DM consente di verificare come la configurazione lineare sia presente principalmente a $[Ca^{2+}]=0$, e che il rapporto tra le frequenze della configurazione compatta e quella intermedia aumenta gradualmente all'aumentare della $[Ca^{2+}]$, arrivando al massimo per concentrazione di 120 ioni Calcio, quando la configurazione compatta diviene dominante. Queste osservazioni suggeriscono che il dominio CAS, inizialmente non strutturato, vada incontro ad un processo graduale di riorganizzazione e strutturazione tridimensionale modulato dalla $[Ca^{2+}]$, che lo porta ad assumere una configurazione ottimale come sensore di ioni Calcio, e che potrebbe rappresentare un nuovo motivo strutturale.

5.5. Aspetti biologici ✨

Le proteine CASQ, ritrovate esclusivamente nei muscoli striati, costituiscono una famiglia non-convenzionale del più ampio gruppo di proteine in grado di interagire e riconoscere i metalli (Ikemoto 1974, Bal 2011, Scriven 2013, Gaburjakova 2012). La proteina CASQ1 possiede due caratteristiche importanti: i) alta selettività per gli ioni Calcio, a fronte dell'assenza di un motivo strutturale noto per il riconoscimento degli ioni Calcio, e ii) la presenza di una regione C-terminale composta da soli residui di acido aspartico (CAS).

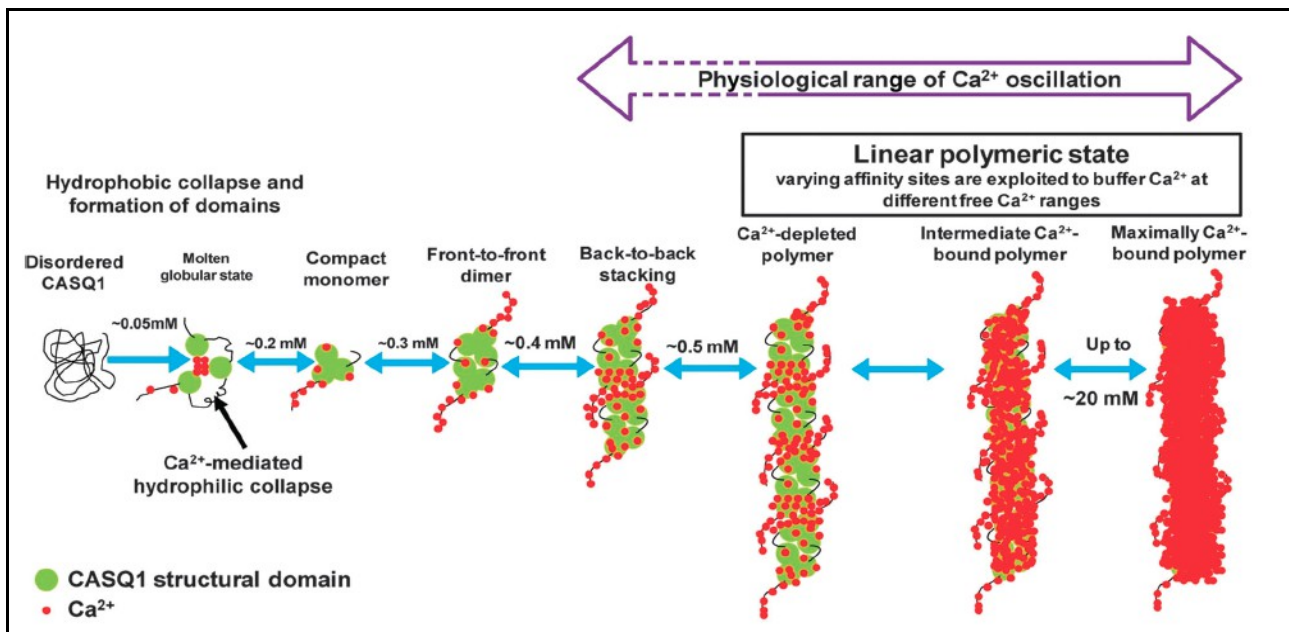


Figura 5.8. Modello di polimerizzazione ed interazione CASQ1:Ca²⁺

Come già osservato, studi di biofisica hanno dimostrato come la proteina CASQ1 sia in grado di legare un elevato numero di ioni, oltre 70 (Park 2004). L'analisi qui presentata delucida le caratteristiche uniche delle geometrie di legame richieste per espletare questo compito, e consente di avere una visione qualitativa e quantitativa delle peculiarità dell'interazione. La presenza di siti con affinità diverse e indotti dalla [Ca²⁺] consente alla proteina di immagazzinare gli ioni Calcio in maniera ottimale all'interno di un ampio intervallo di oscillazioni della concentrazione [Ca²⁺]. Come si è visto, il dominio CAS viene neutralizzato molto prima del resto della

proteina, pertanto si può ipotizzare che la polimerizzazione avvenga a concentrazioni più elevate di quelle sufficienti a neutralizzare il solo CAS. La presenza di siti di legame ad affinità variabile sulla superficie proteica potrebbe infatti consentire al polimero CASQ di immagazzinare e distribuire ioni Calcio senza la necessità di subire un processo di depolimerizzazione, per lo meno all'interno delle oscillazioni di $[Ca^{2+}]$ fisiologiche nell'apparato muscolo-scheletrico. Inoltre, si è osservato che la rete di interazioni idrofobiche promiscue proteina-proteina stabilizza il dimero in modalità crescente all'aumentare della $[Ca^{2+}]$. Questo suggerisce che la sola contrapposizione dei monomeri, pur geometricamente ottimale, in assenza di ioni Calcio non è sufficiente a fornire al dimero una stabilità intrinseca. Il dimero, pertanto, deve essere considerato una struttura in equilibrio instabile, fatto che potrebbe contribuire ad una depolimerizzazione dinamica in caso di richiesta elevata di ioni Calcio per il funzionamento muscolare. L'ipotesi di una depolimerizzazione dinamica dell'isoforma CASQ2 della proteina CASQ è stata suggerita di recente (Bal 2010, Lee 2012) ed è sperimentalmente sostenuta per la CASQ1 da recenti misure delle proprietà di immagazzinamento del calcio del reticolo endoplasmatico in un modello animale (Manno 2013)

5.6. Modello di polimerizzazione ed interazione CASQ1:Ca²⁺ ✨

Sulla base delle osservazioni fatte nei paragrafi precedenti, si può ipotizzare un modello di polimerizzazione modulata dalla interazione proteina-Ca²⁺ (Figura 5.8). Come visto (Sezione 2.4.1), il singolo monomero è composto da tre domini, ciascuno dei quali ha una zona centrale idrofobica che deve collassare così da formare i domini strutturali. L'interfaccia fra i tre domini, nel cuore della proteina CASQ1, è di natura acida (come da natura della proteina stessa). Pertanto, affinché la struttura inizialmente disordinata della proteina CASQ 1 possa formare un monomero compatto, occorre che il core della proteina venga neutralizzato tramite gli ioni Calcio per diminuire la repulsione elettrostatica.

6. Prospettive

In questo lavoro si sono affrontati tre sistemi di elevato interesse clinico e biologico: la sclerosi multipla, la sindrome fibromialgica ed il funzionamento delle fibre muscolari. Per ciascuno di essi è stato identificato un meccanismo da studiare in dettaglio a livello molecolare, e che rappresenta un punto chiave del funzionamento biologico sottostante a ciascun sistema:

- 1) la presentazione dei peptidi antigenici da parte della proteina HLA nella sclerosi multipla, perché essa è in grado di avviare una risposta immune ed autoimmune,
- 2) l'interazione fra lipidi derivati da grassi endogeni o esogeni ed il recettore PAFTR, in quanto collegata ai meccanismi biologici del dolore nella sindrome fibromialgica;
- 3) il legame ed il rilascio degli ioni Calcio da parte di una specifica proteina, CASQ1, per il funzionamento delle fibre muscolari.

Le specifiche prospettive terapeutiche per questi tre sistemi sono legate al design di:

- a) peptidi terapeutici, in grado di interferire nella presentazione antigenica e modulare la reazione autoimmune nella sclerosi multipla;
- b) un modello di rischio e progressione di malattia individualizzato, per lo specifico paziente di sclerosi multipla;
- c) molecole in grado di interferire con il recettore responsabile dei meccanismi biologici del dolore nella sindrome fibromialgica;
- d) test clinici di laboratorio per la sindrome fibromialgica, basati su marcatori nel profilo metabolico individuale;
- e) molecole in grado di interferire con l'immagazzinamento ed il rilascio del Calcio da parte della proteina CASQ2 per la modulazione del battito cardiaco;
- f) diete personalizzate, che bilancino i lipidi di natura esogena che possono contribuire ai meccanismi biologici del dolore nella sindrome fibromialgica.

7. Conclusioni

Senza la pretesa di universalità, cosa che la complessità dei sistemi biologici all'interno degli organismi viventi renderebbe decisamente inappropriata, ritengo di aver contribuito a dimostrare come la modellistica computazionale a scala atomica possa essere un ausilio efficace nella comprensione dei meccanismi funzionali che portano al processo degenerativo della malattia.

Essendomi formato all'interno del paradigma Galileiano della Scienza, ritengo che nessun modello, per quanto esteticamente bello o stimolante, possa costituire da solo un reale avanzamento della conoscenza scientifica. Affinché ciò avvenga occorre che per prima sia verificata una *conditio sine qua non*: la **verifica sperimentale** del modello. Una seconda condizione, è che il modello sia **predittivo**, ossia che fornisca previsioni sul comportamento del sistema in analisi non ancora osservate o collegamenti inattesi con altri sistemi. La prima condizione esige che il modello spieghi *in primis* tutte le evidenze sperimentali note, e congiunta alla seconda condizione, che tutte le predizioni del modello siano sottoposte a verifica sperimentale sul sistema in analisi, con esito positivo.

I modelli qui esposti hanno cercato appunto, con difficoltà, di rimanere all'interno di questa strada, tentando di spiegare alcuni fatti che emergono in un contesto clinico (come per la SM e la FMS) o biologico (come per la proteina CASQ1), e proponendo alcune previsioni del comportamento dei sistemi in analisi, da sottoporre a verifica.

Nel contempo, continuando a fare ricerca in un contesto contemporaneo, sono ben consapevole che la rapidissima generazione di dati sperimentali offre una occasione molto importante, anche se in parte sfugge al paradigma Galileiano, perlomeno come esso viene inteso in senso classico. Infatti, la presenza di enormi quantità di

informazione (dati) generata ad elevate velocità esige, per motivi di fattibilità, una analisi non supervisionata e la generazione di modelli empirici e predittivi, ma dei quali talvolta non si conoscono le basi funzionali sottostanti. Questo nuovo paradigma, di modellistica guidata dai dati, è oggi fondamentale e nel contesto clinico offre comunque delle prospettive molto importanti, perché alla fine dei conti, qualsiasi terapia che operi un compromesso positivo fra effetti terapeutici ed effetti collaterali, anche se non si sa il perché con sufficiente dettaglio, rappresenta comunque un avanzamento per la salute umana o animale. Per quanto detto, si ha quindi attualmente una enorme potenzialità di sviluppo per la modellistica applicata, in quanto essa si propone fra gli strumenti più economici in grado di proporre una analisi funzionale dei problemi biologici.

In conclusione, spero di aver dimostrato come la modellistica possa aiutare la clinica ed il laboratorio, ma sono ben consapevole che essa non possa dire 'l'ultima parola', anche se sicuramente può guidare la scelta ed il design degli esperimenti, fornire supporto funzionale a dati sperimentali e proporre meccanismi funzionali da verificare sperimentalmente. Tutto questo ha un impatto importante non solo sulla salute umana ed animale e sull'avanzamento scientifico, ma anche in ambito etico ed economico, ad esempio contribuendo a ridurre progressivamente il numero ed il costo dei test da effettuare in vitro ed in vivo.

8. Bibliografia

- (Aaron 1984) B. M. Aaron, K. Oikawa, R. A. Reithmeier and B. D. Sykes, *J Biol Chem*, 1984, **259**, 11876-11881.
- (Adams 2011) Adams JJ et al., T cell receptor signaling is limited by docking geometry to peptide-major histocompatibility complex, *Immunity*. 2011 Nov 23;35(5):681-93. doi: 10.1016/j.immuni.2011.09.013.
- (Agudelo 2009) Agudelo WA, Galindo JF, Ortiz M, Villaveces JL, Daza EE, et al. (2009) Variations in the electrostatic landscape of class II human leukocyte antigen molecule induced by modifications in the myelin basic protein peptide: a theoretical approach. *PLoS One* 4: e4164.
- (Andricioaei 2001) Andricioaei I, Karplus M (2001) On the calculation of entropy from covariance matrices of the atomic fluctuations. *J Chem Phys* 115: 6289.
- (Armstrong 2008) Armstrong KM, Piepenbrink KH, Baker BM (2008) Conformational changes and flexibility in T-cell receptor recognition of peptide-MHC complexes. *Biochem J* 415: 183-196.
- (Bal 2010) N. C. Bal, A. Sharon, S. C. Gupta, N. Jena, S. Shaikh, S. Gyorke and M. Periasamy, *J. Biol. Chem.*, 2010, 285, 17188–17196.
- (Bal 2011) N. C. Bal, N. Jena, D. Sopariwala, T. Balaraju, S. Shaikh, C. Bal, A. Sharon, S. Gyorke and M. Periasamy, *Biochem. J.*, 2011, 435, 391–399.
- (Balaraju 2012) Balaraju T et al., (2012) Aromatic interaction profile to understand the molecular basis of raltegravir resistance, *Structural Chemistry*
- (Baranzini 2011) SE Baranzini, *Curr Opin Genet Dev*, 2011, 21, 317-324
- (Barcelos 2006) Barcellos LF, Sawcer S, Ramsay,P.P., Baranzini,S.E., Thomson,G, et al. (2006) Heterogeneity at the HLA-DRB1 locus and risk for multiple sclerosis. *Hum Mol Genet* 15: 2813-2824.
- (Berendsen 1995) H.J.C. Berendsen, et al., GROMACS: A message-passing parallel molecular dynamics implementation, *Comp. Phys. Comm.*, Vol 91, 1–3, 1995

- (Berg 2007) van den Berg HA, Rand DA (2007) Quantitative theories of T-cell responsiveness. *Immunol Rev* 216: 81-92.
- (Borbulevych 2009) Borbulevych OY, Piepenbrink KH, Gloor BE, Scott DR, Sommese RF, et al. (2009) T cell receptor cross-reactivity directed by antigen-dependent tuning of pMHC molecular flexibility. *Immunity* 31: 885-896.
- (Brown 1993) Brown JH, Jardetzky TS, Gorga JC, Stern LJ, Urban RG, et al. (1993) Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature* 364: 33-39.
- (Call 2011) Call MJ (2011) Small molecule modulators of MHC class II antigen presentation: mechanistic insights and implications for therapeutic application. *Mol Immunol* 48: 1735-1743.
- (Chastain 2012) Chastain EM, Miller SD (2012) Molecular mimicry as an inducing trigger for CNS autoimmune demyelinating disease. *Immunol Rev* 245: 227-238.
- (Chen 2008) Chen C, Xia SH, Chen H, Li XH, Therapy for acute pancreatitis with platelet-activating factor receptor antagonists, *World J Gastroenterol*. 2008 Aug 14;14(30):4735-8.
- (Chow 2012) I. T. Chow, E. A. James, V. Tan, A. K. Moustakas, G. K. Papadopoulos and W. W. Kwok, *Mol Immunol*, 2012, 50, 26-34
- (Cocco 2013) Cocco E, et al. (2013) Interaction between HLA-DRB1-DQB1 haplotypes in Sardinina multiple sclerosis population, *PloS ONE* 8(4)
- (Cocco 2012) Cocco E, Sardu C, Pieroni E, Valentini M, Murru R, et al. (2012) HLA-DRB1-DQB1 haplotypes confer susceptibility and resistance to multiple sclerosis in Sardinia. *PLoS One* 7: e33972.
- (Corina) http://www.molecular-networks.com/online_demos/corina_demo.
- (Cozens 1984) B. Cozens and R. A. Reithmeier, *J Biol Chem*, 1984, **259**, 6248-6252.
- (Craig 2003) Craig R. Roy-Nature 425, 351-352, September 2003

- (Csermely 2010) Csermely P, Palotai R, Nussinov R (2010) Induced fit, conformational selection and independent dynamic segments: an extended view of binding events. *Trends Biochem Sci* 35: 539-546.
- (Cui 2008) Cui Q, Sulea T, Schrag JD, Munger C, Hung MN, et al. (2008) Molecular dynamics-solvated interaction energy studies of protein-protein interactions: the MP1-p14 scaffolding complex. *J Mol Biol* 379: 787-802.
- (Darden 1999) Darden T, Perera L, Li L and Pedersen L. (1999) "New tricks for modelers from the crystallography toolkit: the particle mesh Ewald algorithm and its use in nucleic acid simulations", *Structure* 7, R55–R60.
- (Dror 2011) Dror RO, Pan AC, Arlow DH, Borhani DW, Maragakis P, et al. (2011) Pathway and mechanism of drug binding to G-protein-coupled receptors. *Proc Natl Acad Sci U S A* 108: 13118-13123.
- (Dworkin 2005) Dworkin RH, Fields HL. Fibromyalgia from the perspective of neuropathic pain. *J. Rheumatol* 2005;75(s):1-5
- (Edwards 2009) Edwards LJ, Constantinescu CS., Platelet activating factor/platelet activating factor receptor pathway as a potential therapeutic target in autoimmune diseases, *Inflamm Allergy Drug Targets*. 2009 Jul;8(3):182-90.
- (Eswar 2006) N. Eswar, et al. Comparative Protein Structure Modeling With MODELLER. *Current Protocols in Bioinformatics*, John Wiley & Sons, Inc., Supplement 15, 5.6.1-5.6.30, 2006
- (Ewald 1921) Ewald P. (1921) "Die Berechnung optischer und elektrostatischer Gitterpotentiale", *Ann. Phys.* 369, 253–287
- (Fais 2013) Fais A, Cacace E, Corda M., Era B, Peri M, Utzeri S, Ruggiero V. Purine metabolites in fibromyalgia syndrome. *Clinical Biochemistry*. 2013; 46: 37–39.
- (Fenwick 2011) Fenwick RB, Esteban-Martin S, Salvatella X (2011) Understanding biomolecular motion, recognition, and allostery by use of conformational ensembles. *Eur Biophys J* 40: 1339-1355.
- (Ferrante 2012) A. Ferrante and J. Gorski, *J Mol Biol*, 2012, 417, 454-467.

- (Foloppe 2000) Foloppe N, MacKerell AD, All-atom empirical force field for nucleic acids. II :Application to molecular dynamics simulations of DNA and RNA in solution., *J Comput Phys* 151:283–312 (2000)
- (Frisch 2004) Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb, MA, Cheesemanet JR al. Gaussian 03, Revision C.02. Wallingford CT. 2004.
- (Fu 1995) Fu XT, Bono CP, Woulfe SL, Swearingen C, Summers NL, et al. (1995) Pocket 4 of the HLA-DR (alpha,beta 1*0401) molecule is a major determinant of T cells recognition of peptide. *J Exp Med.* 181: 915-26.
- (Fugger 2009) L. Fugger, M. A. Friese and J. I. Bell, *Nat Rev Immunol*, 2009, 9, 408-417.
- (Furukawa 2012) Hiroshi Furukawa, et al., Association of Human Leukocyte Antigen with Interstitial Lung Disease in Rheumatoid Arthritis: A Protective Role for Shared Epitope, *PlosONE*, May 07, 2012 DOI: 10.1371/journal.pone.0033133
- (Gaburjakova 2012) M. Gaburjakova, N. C. Bal, J. Gaburjakova and M. Periasamy, *Cell. Mol. Life Sci.*, 2012, DOI: 10.1007/s00018-012-1199-7.
- (Gagnon 2006) Gagnon SJ, Turner RV, Shiue MG, Damirjian M, Biddison WE (2006) Extensive T cell receptor cross-reactivity on structurally diverse haptenated peptides presented by HLA-A2. *Mol Immunol* 43: 346-356.
- (Gelbard 1994) Gelbard HA, Nottet HS, Swindells S, Jett M, Dzenko KA, Genis P, et al. Platelet-activating factor: a candidate human immunodeficiency virus type 1-induced neurotoxin. *J Virol* 1994;68:4628-4635.
- (Ghosh 1995) Ghosh P, Amaya M, Mellins E, Wiley DC (1995) The structure of an intermediate in class II MHC maturation: CLIP bound to HLA-DR3. *Nature* 378: 457-462.
- (Gilbert 2003) Gibert M, Balandraud N, Touinssi M, Mercier P, Roudier J, et al (2003) Functional categorization of HLA-DRB1 alleles in rheumatoid arthritis: the protective effect. *Hum Immunol.* 64: 930-5.

- (Greer 2005) Greer JM, Pender MP. (2005) The presence of glutamic acid at positions 71 or 74 in pocket 4 of the HLA-DRbeta1 chain is associated with the clinical course of multiple sclerosis. *JNNP* 76: 656-62.
- (Grigg 2012) Grigg J., The platelet activating factor receptor: a new anti-infective target in respiratory disease? *Thorax*. 2012 Sep;67(9):840-1. doi: 10.1136/thoraxjnl-2012-202206. Epub 2012 Jul 16.
- (Glykos 2006) Glykos NM (2006) Software news and updates. Carma: a molecular dynamics analysis program. *J Comput Chem* 27: 1765-1768.
- (Griffin 2011) Griffin JL, Atherton H, Shockcor J, Atzori L. Metabolomics as a tool for cardiac research. *Nat Rev Cardiol* 2011; 8:630-43.
- (Guex 1997) Guex, N. and Peitsch, M.C. (1997), SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling, *Electrophoresis* 18, 2714-2723
- (Gui 2007) Gui C, et al. Understanding the Regulation Mechanisms of PAF Receptor by Agonists and Antagonists: Molecular Modeling and Molecular Dynamics Simulation Studies. *PROTEINS: Structure, Function, and Bioinformatics* 2007;67:41–52.
- (Guvench 2008) O. Guvench, A.D. MacKerell Jr., Comparison of Protein Force Fields for Molecular Dynamics Simulations, in *Molecular Modeling of Proteins, Methods Molecular Biology Volume 443*, 2008, pp 63-88, Humana Press
- (Gunther 2010) Gunther S, Schlundt A, Sticht J, Roske Y, Heinemann U, et al. (2010) Bidirectional binding of invariant chain peptides to an MHC class II molecule. *Proc Natl Acad Sci U S A* 107: 22219-22224.
- (Hammer 1993) Hammer J, Valsasnini P, Tolba K, Bolin D, Higelin J, et al. (1993) Promiscuous and allele-specific anchors in HLA-DR-binding peptides. *Cell* 74: 197-203.
- (Hammond) B. L. Hammond, William A., Jr. Lester, Peter J. Reynolds, Monte Carlo Methods in Ab Initio Quantum Chemistry, World Scientific

- (Han 2012) Han F, et al., HLA-DQ association and allele competition in Chinese narcolepsy, *Tissue Antigens*. 2012 Oct;80(4):328-35. doi: 10.1111/j.1399-0039.2012.01948.x. Epub 2012 Aug 4.
- (Harkiolaki 2009) Harkiolaki M, Holmes SL, Svendsen P, Gregersen JW, Jensen LT, et al. (2009) T cell-mediated autoimmune disease due to low-affinity crossreactivity to common microbial peptides. *Immunity* 30: 348-357.
- (Hasegawa 2010) Hasegawa S, Kohro Y, Shiratori M, Ishii S, Shimizu T, Tsuda M, Inoue K. Role of PAF Receptor in Proinflammatory Cytokine Expression in the Dorsal Root Ganglion and Tactile Allodynia in a Rodent Model of Neuropathic Pain. *PLoS ONE*. 2010;5:e10467.
- (Heldin) C-H Heldin, M Purton Editors, *Signal Transduction*, Garland Science
- (Holland 2012) Holland SJ et al., The T-cell receptor is not hardwired to engage MHC ligands, *Proc Natl Acad Sci U S A*. 2012 Nov 6;109(45):E3111-8. Doi: 10.1073/pnas.1210882109. Epub 2012 Oct 17.
- (Hong 2006) Hong SC et al., DQB1*0301 and DQB1*0601 modulate narcolepsy susceptibility in Koreans, *Hum Immunol*. 2007 Jan;68(1):59-68. Epub 2006 Oct 30.
- (Hov 2011) J. R. Hov, V. Kosmoliaptsis, J. A. Traherne, M. Olsson, K. M. Boberg, A. Bergquist, E. Schrumpf, J. A. Bradley, C. J. Taylor, B. A. Lie, et al., *Hepatology*, 2011, 53, 1967-1976
- (Hovhannisyanyan 2008) Z. Hovhannisyanyan, A. Weiss, A. Martin, M. Wiesner, S. Tollefsen, K. Yoshida, C. Ciszewski, S. A. Curran, J. A. Murray, C. S. David, et al., *Nature*, 2008, 456, 534-538
- (Humphrey 1996) Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. *J Mol Graph* 14: 33-38, 27-38.
- (Ikemoto 1989) N. Ikemoto, M. Ronjat, L. G. Meszaros and M. Koshita, *Biochemistry*, 1989, 28, 6764-6771.

- (Ikemoto 1974) N. Ikemoto, B. Nagy, G. M. Bhatnagar and J. Gergely, *J. Biol. Chem.*, 1974, 249, 2357–2365.
- (IMSG 2011) C. International Multiple Sclerosis Genetics, C. Wellcome Trust Case Control, S. Sawcer, G. Hellenthal, M. Pirinen, C. C. Spencer, N. A. Patsopoulos, L. Moutsianas, A. Dilthey, Z. Su, et al., *Nature*, 2011, 476, 214-219.
- (Iovino 2013) Iovino F, Brouwer MC, van de Beek D, Molema G, Bijlsma JJ., Signalling or binding: the role of the platelet-activating factor receptor in invasive pneumococcal disease, *Cell Microbiol.* 2013 Jun;15(6):870-81. doi: 10.1111/cmi.12129. Epub 2013 Mar 20
- (Ishii 1997) Ishii I, Izumi T, Tsukamoto H, Umeyama H, Ui M, Shimizu T. Alanine exchanges of polar amino acids in the transmembrane domains of a platelet-activating factor receptor generate both constitutively active and inactive mutants. *J Biol Chem* 1997;272:7846-7854.
- (Ishii 2000) Ishii S, Shimizu T. Platelet-activating factor (PAF) receptor and genetically engineered PAF receptor mutant mice. *Prog Lipid Res* 2000;39:41-82.
- (Jansson 2011) Andreas Jansson, Kinetic proofreading and the search for nonself-peptides, *Self Nonself.* 2011 Jan-Mar; 2(1): 1–3, Published online Jan 1, 2011. doi: 10.4161/self.2.1.15362, PMID: PMC3136897
- (Jorgensen 1983) W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J Chem Phys*, 1983, **79**, 926-935.
- (Kalé 1995) L. Kalé et al., NAMD2: greater scalability for parallel molecular dynamics, *Journal of Computational Physics*, 1995
- (Kasperska-Zajac 2008) Kasperska-Zajac A, Brzoza Z, Rogala B, Platelet-activating factor (PAF): a review of its role in asthma and clinical efficacy of PAF antagonists in the disease therapy, *Recent Pat Inflamm Allergy Drug Discov.* 2008 Jan;2(1):72-6.

- (Kaushansky 2012) N. Kaushansky, D. M. Altmann, C. S. David, H. Lassmann and A. Ben-Nun, *J Neuroinflammation*, 2012, 9, 29.
- (Kaushansky 2010) N. Kaushansky, M. Eisenstein, R. Zilkha-Falb and A. Ben-Nun, *Autoimmun Rev*, 2010, 9, 233-236.
- (Kaushansky 2009) N. Kaushansky, D. M. Altmann, S. Ascough, C. S. David, H. Lassmann and A. Ben- Nun, *J Immunol*, 2009, 183, 3531-3541.
- (Kim 2007) E. Kim, B. Youn, L. Kemper, C. Campbell, H. Milting, M. Varsanyi and C. Kang, *J Mol Biol*, 2007, **373**, 1047-1057.
- (Klein 2000) J. KLEIN, A. SATO, *NEJM* 2000
- (Knapp 2010) Knapp B, Omasits U, Schreiner W, Epstein MM (2010) A comparative approach linking molecular dynamics of altered peptide ligands and MHC with in vivo immune responses. *PLoS One* 5: e11653.
- (Kuby Immunology) RA Goldsby, TJ Kindt, BA Osborne, Kuby Immunology, 2000, W H Freeman & Co
- (Kumar 2007) Kumar N, Mohanty D (2007) MODPROPEP: a program for knowledge-based modeling of protein-peptide complexes. *Nucleic Acids Res* 35: W549-555.
- (Kumar 2010a) Kumar A, Hajjar E, Ruggerone P and Ceccarelli M. Structural and dynamical properties of the porins OmpF and OmpC: insights from molecular simulations. *J Phys Condens Matter* 2010;22:454125.
- (Kumar 2010b) Kumar A, Hajjar E, Ruggerone P, Ceccarelli M (2010) Molecular simulations reveal the mechanism and the determinants for ampicillin translocation through OmpF. *J Phys Chem B* 114: 9608-9616.
- (Kumar 2013) A. Kumar, E. Cocco, L. Atzori, M.G. Marrosu, E. Pieroni. "Structural and Dynamical insights on HLA-DR2 complexes that confer susceptibility to Multiple Sclerosis in Sardinia: A Molecular Dynamics simulation study", *PLoS ONE* (2013).

- (Larizza 2012) D. Larizza, V. Calcaterra, C. Klersy, C. Badulli, C. Caramagna, A. Ricci, P. Brambilla, L. Salvaneschi and M. Martinetti, *Autoimmunity*, 2012, 45, 470-475.
- (Lawrence 2011) Lawrence E. Samelson Cold Spring Harb. Perspect. Biol. 2011]
- (Leach) Andrew Leach, *Molecular Modelling: Principles and Applications* (2nd Edition), Prentice Hall
- (Lee 2001) K. H. Lee, K. W. Wucherpfennig and D. C. Wiley, *Nat Immunol*, 2001, 2, 501-507.
- (Lee 2012) K. W. Lee, J. S. Maeng, J. Y. Choi, Y. R. Lee, C. Y. Hwang, S. S. Park, H. K. Park, B. H. Chung, S. G. Lee, Y. S. Kim, H. Jeon, S. H. Eom, C. Kang, H. Kim do and K. S. Kwon, *J. Biol. Chem.*, 2012, 287, 1679–1687.
- (Lesk 2005) Lesk Arthur M., *Introduzione alla Bioinformatica*, McGraw-Hill
- (Li 2005) H. Li, A. D. Robertson and J. H. Jensen, *Proteins*, 2005, 61, 704-721.
- (Lincoln 2005) Lincoln MR, Montpetit A, Cader MZ, Saarela J, Dyment DA, et al., A predominant role for the HLA class II region in the association of the MHC region with multiple sclerosis. *Nat Genet* 37: 1108–1112.
- (Lincoln 2009) Lincoln MR, Ramagopalan SV, Chao MJ, Herrera BM, Deluca GC, et al. (2009) Epistasis among HLA-DRB1, HLA-DQA1, and HLA-DQB1 loci determines multiple sclerosis susceptibility. *Proc Natl Acad Sci USA* 106: 7542–7547.
- (MacKerell 2000) A. D. MacKerell, Jr., N. Banavali and N. Foloppe, *Biopolymers*, 2000, 56, 257-265
- (Madden 1995) Madden DR (1995) The three-dimensional structure of pMHC complexes. *Annu Rev Immunol* 13: 587-622.
- (Mameli 2013) Mameli G, et al., EBNA-1 IgG titers in Sardinian multiple sclerosis patients and controls., *J Neuroimmunol*. 2013 Nov 15;264(1-2):120-2. doi: 10.1016/j.jneuroim.2013.07.017. Epub 2013 Aug 3.

- (Mameli 2014) Mameli G, et al., Epstein-Barr virus and Mycobacterium avium subsp. paratuberculosis peptides are cross recognized by anti-myelin basic protein antibodies in multiple sclerosis patients, *J Neuroimmunol.* 2014
- (Manno 2013) C. Manno, M. Sztretye, L. Figueroa, P. D. Allen and E. Rios, *J. Physiol.*, 2013, 591, 423–442.
- (Manzotte 2013) Manzotte T, Guindalini C, Mazzotti DR, Palombini L, de Souza AL, Poyares D, Bittencourt LR, Tufik S., The human leucocyte antigen DQB1*0602 allele is associated with electroencephelograph differences in individuals with obstructive sleep apnoea syndrome, *J Sleep Res.* 2013 Apr;22(2):217-22. doi: 10.1111/jsr.12005. Epub 2012 Nov 9.
- (Marsh) SGE. Marsh, , P Parham, LD Barber, *The HLA FactsBook*, Academic Press
- (Marsh 2010) SGE Marsh, et al, *Nomenclature for factors of the HLA system*, 2010, *Tissue Antigens*
- (Maurer 1985) A. Maurer, M. Tanaka, T. Ozawa and S. Fleischer, *Proc Natl Acad Sci U S A*, 1985, **82**, 4036-4040.
- (Mazza 2007) Mazza C, Auphan-Anezin N, Gregoire C, Guimezanes A, Kellenberger C, et al. (2007) How much can a T-cell antigen receptor adapt to structurally distinct antigenic peptides? *EMBO J* 26: 1972-1983.
- (McGaughey 1998) McGaughey, GB; Gagné, M; Rappé, AK (1998). "Pi-Stacking interactions. Alive and well in proteins". *J. Biol. Chem.* 273 (25): 15458–15463
- (Mechelli 2011) R. Mechelli, J. Anderson, D. Vittori, G. Coarelli, V. Annibali, S. Cannoni, F. Aloisi, M. Salvetti, J. A. James and G. Ristori, *Mult Scler*, 2011, 17, 1290-1294.
- (Montrucchio 2000) *Role of Platelet-Activating Factor in Cardiovascular Pathophysiology*, Giuseppe Montrucchio , Giuseppe Alloatti , Giovanni Camussi, *Physiological Reviews*, 2000, Vol. 80, no. 1669-1699

- (Morris 2009) Morris, G. M., et al., (2009) Autodock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Computational Chemistry* 2009, 16: 2785-91
- (Morita 2004) Morita K, Morioka N, Abdin J, Kitayama S, Nakata Y, Dohi T. Development of tactile allodynia and thermal hyperalgesia by intrathecally administered platelet-activating factor in mice. *Pain*. 2004;111:351-359.
- (Morita 2008) Morita K, Kitayama T, Morioka N, Dohi T. Glycinergic mediation of tactile allodynia induced by platelet-activating factor (PAF) through glutamate-NO-cyclic GMP signalling in spinal cord in mice. *Pain*. 2008;138:525-536.
- (Murthy 1997) Murthy VL, Stern LJ (1997) The class II MHC protein HLA-DR1 in complex with an endogenous peptide: implications for the structural basis of the specificity of peptide binding. *Structure* 5: 1385-1396.
- (Naim 2007) Naim M, Bhat S, Rankin KN, Dennis S, Chowdhury SF, et al. (2007) Solvated interaction energy (SIE) for scoring protein-ligand binding affinities. 1. Exploring the parameter space. *J Chem Inf Model* 47: 122-133.
- (Nakamura 1991) Nakamura M, Honda ZI, Izumi T, Sakanaka C, Mutoh H, Minami M et al. Molecular cloning and expression of platelet-activating factor receptor from human leukocytes. *J Biol Chem* 1991;266:20400-20405.
- (Nicholson 2012) Nicholson JK, Holmes E, Kinross JM, Darzi AW, Takats Z, Lindon JC. Metabolic phenotyping in clinical and surgical environments. *Nature* 2012; 15,49:384-92.
- (Noble 2011) Noble JA, Valdes AM., Genetics of the HLA region in the prediction of type 1 diabetes, *Curr Diab Rep*. 2011 Dec;11(6):533-42. doi: 10.1007/s11892-011-0223-x.
- (Oestvang 2011) Oestvang J, Anthonsen MW, Johansen B. LysoPC and PAF trigger arachidonic acid release by divergent signaling mechanisms in monocytes. *J Lipids*. 2011, Article ID 532145, 11 pages ; doi:10.1155/2011/532145

- (Oksenberg 2008) Oksenberg JR, Baranzini SE, Sawcer S, Hauser SL (2008) The genetics of multiple sclerosis: SNPs to pathways to pathogenesis. *Nat Rev Genet* 9: 516-526.
- (Ou 1998) Ou D, Mitchell LA, Tingle AJ (1998) A new categorization of HLA DR alleles on a functional basis. *Hum Immunol* 59: 665-76.
- (Painter 2008) Painter CA, Cruz A, Lopez GE, Stern LJ, Zavala-Ruiz Z (2008) Model for the peptide-free conformation of class II MHC proteins. *PLoS One* 3: e2403.
- (Painter 2011) Painter CA, Negroni MP, Kellersberger KA, Zavala-Ruiz Z, Evans JE, et al. (2011) Conformational lability in the class II MHC 310 helix and adjacent extended strand dictate HLA-DM susceptibility and peptide exchange. *Proc Natl Acad Sci U S A* 108: 19329-19334.
- (Park 2003) H. Park, S. Wu, A. K. Dunker and C. Kang, *J Biol Chem*, 2003, **278**, 16176-16182.
- (Park 2004) H. Park, I. Y. Park, E. Kim, B. Youn, K. Fields, A. K. Dunker and C. Kang, *J Biol Chem*, 2004, **279**, 18026-18033.
- (Phillips 2005) Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E et al. Scalable molecular dynamics with NAMD. *J Comput Chem* 2005;26:1781-1802.
- (Pierce 2011) Pierce BG, Hourai Y, Weng Z, 2011, Accelerating Protein Docking in ZDOCK Using an Advanced 3D Convolution Library. *PLoS One*
- (Ponder 2003) J.W. Ponder, D.A. Case. Force fields for protein simulations. *Adv. Prot. Chem.* 66, 27-85 (2003)
- (Prescott 2000) Prescott SM, Zimmerman GA, Stafforini DM, McIntyre TM: Platelet-activating factor and related lipid mediators. *Annu Rev Biochem* 2000, 69:419-445.
- (Quarteroni 2006) A. Quarteroni, et al., Eds., *Complex Systems in Biomedicine*, Springer Italia, 2006

- (Ramagopalan 2009) SV Ramagopalan, JC Knight and GC Ebers, *Curr Opin Neurol*, 2009, 22
- (Rammensee 1995) Rammensee HG (1995) Chemistry of peptides associated with MHC class I and class II molecules. *Curr Opin Immunol* 7: 85-96.
- (Rapaport) D. C. Rapaport , *The Art of Molecular Dynamics Simulation*, Cambridge University Press
- (Reboul 2012) Reboul CF, Meyer GR, Porebski BT, Borg NA, Buckle AM (2012) Epitope flexibility and dynamic footprint revealed by molecular dynamics of a pMHC-TCR complex. *PLoS Comput Biol* 8: e1002404.
- (Rojas-Villarga 2010) Rojas-Villarraga A, Botello-Corzo D, Anaya JM., HLA-Class II in Latin American patients with type 1 diabetes, *Autoimmun Rev.* 2010 Aug;9(10):666-73. doi: 10.1016/j.autrev.2010.05.016. Epub 2010 May 23.
- (Ryckaert 1977) Ryckaert, J-P; Ciccotti G, Berendsen HJC (1977). "Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes". *Journal of Computational Physics* 23 (3): 327–341
- (Rudolph 2006) M. G. Rudolph, R. L. Stanfield and I. A. Wilson, *Ann Rev Immunol*, 2006, 24, 419-466.
- (Rupp 2011) Rupp B, Gunther S, Makhmoor T, Schlundt A, Dickhaut K, et al. (2011) Characterization of structural features controlling the receptiveness of empty class II MHC molecules. *PLoS One* 6: e18662
- (Sadegh-Nasseri 2010) Sadegh-Nasseri S, Natarajan S, Chou CL, Hartman IZ, Narayan K, et al. (2010) Conformational heterogeneity of MHC class II induced upon binding to different peptides is a key regulator in antigen presentation and epitope selection. *Immunol Res* 47: 56-64.
- (Sadovnick 2012) Sadovnick AD., Genetic background of multiple sclerosis, *Autoimmun Rev.* 2012 Jan;11(3):163-6. doi: 10.1016/j.autrev.2011.05.007. Epub 2011 May 18.

- (Salomon 2013) R. Salomon-Ferrer, D.A. Case, R.C.Walker. An overview of the Amber biomolecular simulation package. *WIREs Comput. Mol. Sci.* 3, 198-210 (2013)
- (Sanchez 2012) Sanchez, E.J., Lewis, K.M., Danna, B.R., Kang, C., High-capacity Ca²⁺ binding of human skeletal calsequestrin, (2012) *J.Biol.Chem.* 287: 11592-11601
- (Schlick 2002) Schlick T. (2002). *Molecular Modeling and Simulation: An Interdisciplinary Guide* Springer-Verlag Interdisciplinary Applied Mathematics, Mathematical Biology, Vol. 21. New York, NY
- (Scott 2011) Scott DR, Borbulevych OY, Piepenbrink KH, Corcelli SA, Baker BM (2011) Disparate degrees of hypervariable loop flexibility control T-cell receptor cross-reactivity, specificity, and binding mechanism. *J Mol Biol* 414: 385-400.
- (Scriven 2013) D. R. Scriven, P. Asghari and E. D. Moore, *Cardiovasc. Res.*, 2013, 98, 169–176.
- (Siebold 2004) C. Siebold, B. E. Hansen, J. R. Wyer, K. Harlos, R. E. Esnouf, A. Svejgaard, J. I. Bell, J. L. Strominger, E. Y. Jones and L. Fugger, *Proc Natl Acad Sci U S A*, 2004, 101, 1999-2004
- (Smith 1998) Smith KJ, Pyrdol J, Gauthier L, Wiley DC, Wucherpfennig KW (1998) Crystal structure of HLA-DR2 (DRA*0101, DRB1*1501) complexed with a peptide from human myelin basic protein. *J Exp Med* 188: 1511-1520.
- (Sospedra 2006) M. Sospedra, et al., *J Immunol*, 2006, 176, 1951-1961.
- (Southwood 1998) Southwood S, Sidney J, Kondo A, del Guercio MF, Appella E, et al. (1998) Several common HLA-DR types share largely overlapping peptide binding repertoires. *J Immunol* 160: 3363-3373.
- (Spoel 1998) David van der Spoel, et al., A systematic study of water models for molecular simulation: Derivation of water models optimized for use with a reaction field, *J. Of Chem. Phys.* Vol. 108, 24, 1998
- (Steinman 2007) L. Steinman, *Neurotherapeutics*, 2007, 4, 661-665.

- (Stern 1994) Stern LJ, Brown JH, Jardetzky TS, Gorga JC, Urban RG, et al. (1994) Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. *Nature* 368: 215-221.
- (Sulea 2012) Sulea T, Purisima EO, The solvated interaction energy method for scoring binding affinities, *Meth Mol Biol.* 2012; 819:295-303
- (Tramontano 2006) Anna Tramontano, *Protein Structure Prediction: Concepts and Applications.*, Arthur M. Lesk, J. Wiley & Sons, 2006
- (Tsoulos 2011) Tsoulos IG, Stavrakoudis A (2011) Eucb: A C++ program for molecular dynamics trajectory analysis. *Computer Physics Communications* 182: 834-841.
- (Tsuda 2011) Tsuda M, Tozaki-saitoh H and Inoue K. Platelet-Activating Factor and pain. *Biol. Pharm. Bull.*2011;34:1159-1162.
- (ul-Haq 2011) Zaheer-ul-Haq, Waqasuddin Khan, Molecular and structural determinants of adamantyl susceptibility to HLA-DRs allelic variants: an in silico approach to understand the mechanism of MLEs, *Journal of Computer-Aided Molecular Design*, January 2011, Volume 25, Issue 1, pp 81-101
- (Venkatakrishnan 2013) Venkatakrishnan AJ, Deupi X, Lebon G, Tate CG, Schertler GF, Babu MM (2013). "Molecular signatures of G-protein-coupled receptors". *Nature* 494 (7436): 185–194.
- (Vrethem M 2012) M. Vrethem, K. Malmgren and J. Lindh, *J Neurol Sci*, 2012, 321, 89-91.
- (Wallace 1995) A. C. Wallace, R. A. Laskowski and J. M. Thornton, *Protein Eng*, 1995, 8, 127-134.
- (Wang 1998) Wang S, Trumble WR, Liao H, Wesson CR, Dunker AK, Kang CH (1998). "Crystal structure of calsequestrin from rabbit skeletal muscle sarcoplasmic reticulum". *Nat. Struct. Biol.* 5 (6): 476–83. doi:10.1038/nsb0698-476. PMID 9628486

- (Weiner 2012) Weiner Lachmi K1, Lin L, Kornum BR, Rico T, Lo B, Aran A, Mignot E., DQB1*06:02 allele-specific expression varies by allelic dosage, not narcolepsy status., *Hum Immunol.* 2012 Apr;73(4):405-10. doi: 10.1016/j.humimm.2012.01.004. Epub 2012 Jan 31.
- (Weljie 2007) Weljie AM, Dowlatabadi R, Miller BJ, Vogel HJ, Jirik FR. An inflammatory arthritis-associated metabolite biomarker pattern revealed by 1H NMR spectroscopy. *J Proteome Res.* 2007; 6:3456-64.
- (Wolfe 1990) Wolfe F, et al, The American College of Rheumatology 1990 criteria for the classification of fibromyalgia. Report of the Multicenter Criteria Committee. *Arthritis Rheum.* 1990;33:160-72.
- (Wolfe 2010) Wolfe et al.,. The American College of Rheumatology Preliminary Diagnostic Criteria for Fibromyalgia and Measurement of Symptom Severity. *Arthritis Care & Research* 2010;62:600-610.
- (Woo 2012) Woo HI1, Joo EY, Hong SB, Lee KW, Kang ES., Use of PCR with sequence-specific primers for high-resolution human leukocyte antigen typing of patients with narcolepsy., *Ann Lab Med.* 2012 Jan;32(1):57-65. doi: 10.3343/alm.2012.32.1.57. Epub 2011 Dec 20.
- (Wuncherpfenning 1995) Wuncherpfenning KW Strominger JL (1995) Molecular mimicry in T cell-mediated autoimmunity: viral peptides activate human T cell clones specific for myelin basic protein. *Cell* 80: 695-705.
- (Yaneva 2009) Yaneva R, Springer S, Zacharias M (2009) Flexibility of the MHC class II peptide binding cleft in the bound, partially filled, and empty states: a molecular dynamics simulation study. *Biopolymers* 91: 14-27
- (Ye 1991) Ye RD, Prossnitz ER, Zou A, Cochrane CG. Characterization of a human cDNA that encodes a functional receptor for platelet activating factor. *Biochem Biophys Res Commun* 1991;180:105-111.
- (Young 2013) Young SP, et al., The impact of inflammation on metabolomic profiles in patients with arthritis. *Arthritis & Rheumatism* 2013; 5:2015-23.

Ringraziamenti

Il ringraziamento più sentito va' al Dott. Amit Kumar, che ha svolto un ruolo energizzante nella nostra collaborazione, e si è accollato il maggior carico di lavoro ripetitivo, nell'esecuzione e nel controllo della correttezza esecuzioni dei software paralleli utilizzati per la modellistica.

Ringrazio la mia relatrice, Prof.ssa Marisa Marrosu, e la sua collaboratrice Dott.ssa Eleonora Cocco, per avermi fornito stimoli interessanti ed interdisciplinari, ed avermi iniziato ad un argomento tanto affascinante per la sua complessità quanto tragico per gli effetti sulla salute umana.

Ringrazio il Dott. Pierluigi Caboni ed il Prof. Luigi Atzori per avermi fornito una ulteriore traccia di lavoro sul collegamento modellistica dei modelli biologici e salute umana, e per aver creduto nella modellistica più di quanto tipicamente avviene in ambiti così complessi.

Ringrazio il Dott. Ashoke Sharon e la Dott.ssa Harapriya Chakravarty, per aver lavorato con tenacia ed originalità, e suggerito un problema biologico di estremo interesse, e le cui prospettive sono anch'esse legate al miglioramento della salute umana.

Ringrazio i pazienti di sindrome fibromialgica e di sclerosi multipla, che direttamente ed indirettamente hanno contribuito a fornire le informazioni ed i dati iniziali dai quali i lavori qui discussi sono scaturiti.

Ringrazio infine la mia istituzione, il CRS4, ed in particolare il Prof. Giaime Cao, il Dott. Alessandro Bulfone, lo staff del Centro di Calcolo ad Alte Prestazioni ed i colleghi Maria Valentini e Matteo Floris, coi quali ho mosso i primi passi su questa strada.

A1. Lavori Scientifici Pubblicati, Presentazioni a Conferenze e Partecipazioni a Scuole, durante il triennio del dottorato

A1.1 Elenco delle pubblicazioni

- (1) E. Cocco, C. Sardu, E. Pieroni, M. Valentini, R. Murru, G. Costa, S. Tranquilli, J. Frau, G. Coghe, N. Carboni, M. Floris, P. Contu, M. G. Marrosu, "HLA-DRB1-DQB1 Haplotypes conferring susceptibility and resistance to multiple sclerosis in Sardinia", PLoS ONE 7(4): e33972. doi:10.1371/journal.pone.0033972 (2013).
- (2) A. Kumar, E. Cocco, L. Atzori, M.G. Marrosu, E. Pieroni. "Structural and Dynamical insights on HLA-DR2 complexes that confer susceptibility to Multiple Sclerosis in Sardinia: A Molecular Dynamics simulation study", PLoS ONE (2013).
- (3) E. Cocco, R. Murru, G. Costa, A. Kumar, E. Pieroni, C. Melis, L. Barberini, C. Sardu, L. Loreface, G. Fenu, J. Frau, G. Coghe, N. Carboni, M.G. Marrosu. "Interaction between HLA-DRB1-DQB1 haplotypes in Sardinia multiple sclerosis population", PLoS ONE (2012).
- (4) A. Kumar, H. Chakravarty, N.C. Bal, T. Balaraju, N. Jena, G. Misra, C. Bal, E. Pieroni, M. Periasamy, A. Sharon. "Identification of Calcium binding sites on calsequestrin1 and its implications to polymerization", Molecular BioSystems (2013).
- (5) A Kumar, P Melis, V Genna¹, E Cocco, MG Marrosu, and Enrico Pieroni, Antigenic Peptide Molecular Recognition by DRB1 - DQB1 Haplotype modulates Multiple Sclerosis Susceptibility, under review by Molecular BioSystems, April 2014
- (6) P Caboni, B Liori, A Kumar, ML Santoru, S Asthana, E Pieroni, A Fais, B Era, E Cacace, V Ruggiero, L Atzori, Metabolomics analysis and modeling suggest lysophosphocholines-PAF Receptor interaction in fibromyalgia, submitted to Bioanalysis, April 2014

A1.2 Elenco dei lavori presentati a conferenze

(1) Structural and dynamical insights on HLA DRB1-peptide complexes which confer resistance and susceptibility to Multiple Sclerosis in Sardinia: a Molecular Dynamics simulation study, A. Kumar, E. Cocco, L. Atzori, M.G. Marrosu, E. Pieroni, poster presented at 22nd IUBMB - 37th FEBS, Seville (2012)

(2) E. Pieroni, A. Kumar, M. Pisu, V. Genna, A. Concas, G. Cao, "Combination of classical and molecular modeling approaches to investigate the effect of various types of drugs on cell proliferation kinetics", with G. Cao invited speaker to ISCRE 22 (International Symposium on Chemical Reaction Engineering), Maastricht, The Netherlands (2012)

(3) Analysis of the inhibitory effect of antipsychotic drugs during in vitro cell proliferation through classical and molecular modeling approaches, Enrico Pieroni, Amit Kumar, Massimo Pisu, Vito Genna, Alessandro Concas and Giacomo Cao, submitted in 2012 and accepted for presentation to ECCE-ECAB, 2nd European Congress of Applied Biotechnology, The Hague, The Netherlands, April 21-25 (2013)

A1.3 Elenco lavori su proceedings di conferenza con referaggio

(1) Combination of classical and molecular modeling approaches to investigate the effect of antipsychotic drugs on cell proliferation kinetics, Enrico Pieroni, Amit Kumar, Massimo Pisu, Vito Genna, Alessandro Concas and Giacomo Cao, Presentation at Icheap 11, Milan 2-5 June (2013) with publication of the extended abstract on Chemical Engineering Transactions (2013).

A1.4 Partecipazioni a scuole

(1) XXX Scuola Annuale GNB, Neuroinfarmatica, Bressanone, 19 - 23 settembre 2011.

(2) XXXI Scuola Annuale GNB, Dalla ricerca al mercato: trasformare il risultato della ricerca in un prodotto, Bressanone, 17 - 21 settembre 2012.

(3) XXXII Scuola Annuale GNB, Approccio Integrato per la Medicina Rigenerativa, Bressanone, 16 - 20 settembre 2013.