

# Modeling the Ternary Complex TCR-V $\beta$ /CollagenII(261–273)/HLA-DR4 Associated with Rheumatoid Arthritis

Maria Cristina De Rosa<sup>1\*</sup>, Bruno Giardina<sup>2</sup>, Caterina Bianchi<sup>2</sup>, Cristiana Carelli Alinovi<sup>2</sup>, Davide Pirolli<sup>2</sup>, Gianfranco Ferraccioli<sup>3</sup>, Maria De Santis<sup>3</sup>, Gabriele Di Sante<sup>4</sup>, Francesco Ria<sup>4</sup>

**1** Istituto di Chimica del Riconoscimento Molecolare, Consiglio Nazionale delle Ricerche, Rome, Italy, **2** Istituto di Biochimica e Biochimica Clinica, Università Cattolica del Sacro Cuore, Rome, Italy, **3** Dipartimento di Reumatologia, Università Cattolica del Sacro Cuore, Rome, Italy, **4** Istituto di Patologia Generale, Università Cattolica del Sacro Cuore, Rome, Italy

## Abstract

**Background:** It is known that genetic predisposition to rheumatoid arthritis (RA) is associated with the MHC class II allele HLA-DR4 and that residues 261–273 of type II collagen (huCollp261) represent an immunodominant T cell epitope restricted by the DR4 molecule. Despite recent advances in characterization of MHC and T cell receptor (TCR) contacts to this epitope, the atomic details of TCR/huCollp261/HLA-DR4 ternary complex are not known.

**Methodology/Principal Findings:** Here we have used computational modeling to get insight into this interaction. A three-dimensional model of the TCR V $\beta$  domain from a DR4<sup>+</sup> patient affected by RA has been derived by homology modeling techniques. Subsequently, the structure of the TCR V $\beta$  domain in complex with huCollp261/HLA-DR4 was obtained from a docking approach in conjunction with a filtering procedure based on biochemical information. The best complex from the docking experiments was then refined by 20 ns of molecular dynamics simulation in explicit water. The predicted model is consistent with available experimental data. Our results indicate that residues 97–101 of CDR3 $\beta$  are critical for recognition of huCollp261/HLA-DR4 by TCR. We also show that TCR contacts on p/MHC surface affect the conformation of the shared epitope expressed by DR alleles associated with RA susceptibility.

**Conclusions/Significance:** This work presents a three-dimensional model for the ternary complex TCR-V $\beta$ /collagenII(261–273)/HLA-DR4 associated with rheumatoid arthritis that can provide insights into the molecular mechanisms of self reactivity.

**Citation:** De Rosa MC, Giardina B, Bianchi C, Carelli Alinovi C, Pirolli D, et al. (2010) Modeling the Ternary Complex TCR-V $\beta$ /CollagenII(261–273)/HLA-DR4 Associated with Rheumatoid Arthritis. PLoS ONE 5(7): e11550. doi:10.1371/journal.pone.0011550

**Editor:** Anna Kristina Croft, University of Wales Bangor, United Kingdom

**Received:** March 16, 2010; **Accepted:** June 15, 2010; **Published:** July 14, 2010

**Copyright:** © 2010 De Rosa et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported by Consiglio Nazionale delle Ricerche (CNR). Computer simulations were performed at the Caspur Interuniversities Consortium for Supercomputing Applications (grant DPM@CASPUR 2008). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: mariacristina.derosa@icrm.cnr.it

## Introduction

Recognition by T cell receptors (TCRs) of antigenic peptides (p) presented by class I or class II major histocompatibility complex (MHC) protein is central to cellular immune responses [1,2]. The molecular events taking place at the TCR/p/MHC interface are also directly involved in immunomediated diseases. Rheumatoid arthritis (RA) is an autoimmune disease characterized by a chronic inflammation of the synovial joints leading to a progressive destruction of the articular cartilage that progressively invalidates patients [3]. Genetic predisposition to RA has been significantly associated with the HLA class II alleles HLA-DRB1 01 and HLA-DRB1 04 [4,5]. DR  $\beta$ -chain encoded by these RA-related DRB1 genes possess a “shared epitope” formed by conserved amino acids at positions 67–74 [6]. It has been observed that sequence differences in this region, especially in residue 71, strongly influence T cell recognition and immune response [7,8,9] by determining the selection of peptides presented by the DR molecule [8].

Although the etiology of rheumatoid arthritis remains unknown, type II collagen is a strong candidate autoantigen. It is the

predominant protein of articular cartilage, and autoimmunity to type II collagen is commonly detected in patients with RA [10,11]. T cell responses in experimental collagen-dependent RA are directed towards the immunodominant pathogenic epitope encompassing residues 261–273 of collagen II [12,13] (here-after called huCollp261).

Elucidating the mechanism underlying the molecular recognition of huCollp261/HLA-DR4 by TCR requires structural models of the complex. Till now, approximately 20 TCR/p/MHC complex structures have been solved and among these, eight are TCR/p/MHC of class II structures. Most contacts between TCR and peptide occur through the CDR3 loops, which exhibit the greatest degree of variability whereas the preponderance of generally conserved contacts with the MHC  $\alpha$  helices is mediated through CDR1 and CDR2 [14]. So far, however, no crystal structure has been published for a TCR in complex with human type II collagen peptide/MHC assembly.

In our recent study, we used the VB-JB spectratyping (the so called “immunoscope” technique) to identify T cells specific for huCollp261 of DR4<sup>+</sup> subjects in the early phases of RA [15]. This

PCR-based technique divides a bulk T cell population in approximately 3000 groups on the basis of the VB and JB segments recombined and of the length of the CDR3 region that varies according to bases additions and subtractions at the V-D-J joint [16]. In the index DR4<sup>+</sup> RA patient studied thoroughly by immunoscope [15], we identified and sequenced the CDR3 of a TCR  $\beta$ -chain belonging to T cells specifically stimulated by huCollp261 that are present in the blood and spontaneously enriched in the synovial fluid of inflamed joints. This TCR is obtained by recombination of VB1 and JB2.6 and has the following CDR3 region (VB1)CASS DTGS SGAN(BJ2.6). This sequence is termed here VB1<sup>OE</sup>. We also studied a DR4<sup>-</sup>DR1<sup>+</sup> RA patient and found one T cell expanding specifically in response to huCollp261 that used a TCR obtained by recombination of the same VB1 and JB2.6, having a CDR3 region of the same length, but displaying a sequence (VB1)CASS GDRS AGAN(JB2.6). This sequence is termed here VB1<sup>VB</sup>. This TCR recognizes the same peptide but in a conformation different from the one recognized by VB1<sup>OE</sup> since the DR molecule presenting it is certainly not DR4. The availability of two highly homologous sequences for TCR  $\beta$ -chains from DR4<sup>+</sup> and DR4<sup>-</sup> patients affected by RA gave us the opportunity to build a three-dimensional model of the ternary complex TCR-V $\beta$ /huCollp261/HLA-DR4 by computational approaches. The VB1<sup>VB</sup> sequence has been used as a sort of background sample that provides information about the non-specific contacts that can emerge when modeling the interaction of a VB1-JB2.6 TCR and the DR4/huCollp261 complex.

In the work presented here, reasonable structures of TCR V $\beta$  domain and of huCollp261 bound to HLA-DR4 were first constructed by molecular modeling methods. Subsequently, we identified an equilibrated reasonable structure of the TCR-V $\beta$ /huCollp261/HLA-DR4 ternary complex using protein-protein docking and molecular dynamics simulations. On the basis of the final complex structure and simulation results, we elucidated the molecular basis underlying the selectivity of the interaction within this ternary complex.

## Methods

### Ethics Statement

Informed written consent was obtained from all the patients. The research is in compliance with the Helsinki Declaration. The research was approved by the Institutional Ethical Committee of the Catholic University of the Sacred Heart.

### Patients and TCR V $\beta$ domain sequencing

Patients OE and VB satisfied the American College of Rheumatology criteria for RA. Patients were characterized for the HLA-DR haplotype by PCR-SSO, using the Inno-LiPA HLA-DRB1 Amp Plus kit (Immogenetics N.V., 9052 Gent, Belgium), according to manufacturer's instructions.

In our previous work, we described the clinical characteristics of patient OE as well as TCR repertoire analysis and sequencing of the TCR V $\beta$  domains of patient OE [15]. In this study we used the same protocols for TCR repertoire analysis and sequencing of the TCR V $\beta$  domains of patient VB. DNA sequence was translated into protein sequence through the ExpASY Proteomics Server.

### Homology modeling of TCR V $\beta$ domains

Sequences of TCR  $\beta$ -chains for which the three-dimensional structure was known were selected based upon similarity using PSI-BLAST (available on the World Wide Web at blast.ncbi.nlm.

nih.gov/Blast). The homology model was based on the structure of human autoimmune TCR bound to a myelin basic protein self-peptide and a multiple sclerosis-associated HLA class II allele (Protein Data Bank (PDB) code: 1zgl) [17]. VB1<sup>VB</sup> and VB1<sup>OE</sup> were aligned with the crystallographic TCR V $\beta$  domain of 1zgl using ClustalW (available on the World Wide Web at clustalw.genome.ad.jp/) and rendered using ESPript program [18]. Based on ClustalW alignments three-dimensional models of VB1<sup>VB</sup> and VB1<sup>OE</sup> were generated by comparative protein modeling with MODELLER [19] module in Discovery Studio Modeling1.1 (Accelrys Inc.).

Twenty models, optimized by a short simulated annealing refinement protocol available in MODELLER, were generated for each V $\beta$  domain. The simulated annealing procedure was carried out *in vacuo* (dielectric constant = 1) considering that in MODELLER, the effect of solvation, like electrostatics, is assumed to be encoded in the template structure and thus in the distance restraints derived from the template. The temperatures used in the simulated annealing procedure were 150 K, 250 K, 400 K, 1000 K for heating and 800 K, 600 K, 500 K, 400 K, 300 K for cooling.

The geometrical consistency of the model was evaluated based on PDF violations provided by MODELLER. The models were then evaluated using the programs VERIFY3D [20], PROCHECK [21] and by visual inspection using the computer graphics program Discovery Studio 2.1 (Accelrys Inc.).

### Molecular modeling of DR4/huCollp261 complex

Twenty conformations of huCollp261 (AGFKGEQGPKEGEP, position 1, P1 = F) bound to HLA-DR4 were generated using the simulated annealing protocol of MODELLER and HLA-DR4 in complex with human collagen peptide 1168–1180 (PDB code: 2seb) as starting structure [22]. The objective function used for structure generation in MODELLER is referred to as a molecular PDF (probability density function). This is a combination of all the feature PDFs used to restrain particular geometric features of the protein model. Molecular PDF values are collected to Table 1 and the best-ranked model based on PDF violations (Model 8) was selected.

The quality of the structures was assessed using VERIFY3D [20] and PROCHECK [21] (Table 1).

The  $\beta$ 105– $\beta$ 112 and  $\beta$ 165– $\beta$ 168 sequences, which were not resolved in 2seb x-ray crystal structure, were built with an *ab initio* modeling routine of MODELLER [19].

### Protein-protein docking

Docking of all protein pairs was performed with the FTDOCK program [23]. FTDOCK is based on rigid-body geometric docking method originally proposed by Katchalski-Katzir et al. [24]. In this approach the relatively larger protein is held fixed while the smaller protein translates and rotates on a defined grid surface so as to establish the best geometric fit. FTDOCK uses fast Fourier transformation (FFT) grids to rapidly evaluate shape complementarity. A simple Coulombic model is then used as a binary filter, leaving only those complexes with attractive electrostatic interactions. Large positive FTDOCK scores denote complex formation with good surface complementarity, a score of zero indicates that the molecules do not interact at all and the score is negative if the smaller molecule significantly overlaps with the larger one. In all our calculations FTDOCK was run with electrostatics on using default parameters. For all protein pairs the rigid-body docking approach of FTDOCK resulted in 10,000 possible docked complexes which were reduced by a filtering procedure, based on biochemical knowledge. Interaction energies

**Table 1.** Molecular PDF and protein structure analysis for the twenty models of DR4/huCollp261.

Model	Molecular PDF	Ramachandran plot quality (%)			
		Core	Allowed	General	Disallowed
Model 8	2189.39	92.2	6.6	1.2	0.0
Model 11	2209.43	92.5	6.6	0.9	0.0
Model 12	2263.19	93.1	5.7	1.2	0.0
Model 15	2278.74	91.9	6.9	1.2	0.0
Model 10	2283.98	92.2	6.9	0.6	0.3
Model 4	2287.99	91.6	6.9	1.5	0.0
Model 16	2329.97	93.1	5.4	1.5	0.0
Model 20	2352.77	92.5	6.6	0.9	0.0
Model 9	2356.94	92.2	6.6	1.2	0.0
Model 3	2359.49	94.0	5.4	0.6	0.0
Model 5	2373.85	91.6	6.6	1.5	0.3
Model 13	2388.94	92.2	6.9	0.9	0.0
Model 2	2397.23	91.0	7.2	1.8	0.0
Model 19	2403.73	92.5	6.0	1.2	0.3
Model 7	2409.73	92.8	6.6	0.3	0.3
Model 18	2435.51	91.9	6.6	1.5	0.0
Model 17	2529.67	91.6	6.9	1.5	0.0
Model 14	2540.33	92.5	6.3	1.2	0.0
Model 6	3074.24	92.2	6.9	0.9	0.0
Model 1	3235.73	91.9	6.9	0.9	0.3

doi:10.1371/journal.pone.0011550.t001

were computed using CHARMM [25] as implemented in DiscoveryStudio (Accelrys Inc.). The NACCESS program [26] was used to calculate the interface Accessible Surface Area (ASA).

### Validation of docking method

The methodology was first tested on the complex of the human TCR HA1.7 specific for the hemagglutinin antigen peptide (HA) from influenza A virus bound to class II molecule HLA-DR4 (PDB code: 1j8h) [27]. The system was digitized onto a 246×246×246 grid with the resolution of 0.69 Å. The distance constraint applied required the separation between the following residues and protein chains to be <4.5 Å: 30D:A 30D:B 50D:A 50D:B (97–98)D:P where A and B stand for  $\alpha$  and  $\beta$  chain of HLA-DR4, respectively, D stands for  $\beta$  chain of TCR, P stands for HA peptide. This biological filter was selected based on the hypothesis of the two-step binding mechanism for T-cell receptor recognition of p/MHC complex for which the TCR binds p/MHC in a conserved diagonal orientation that first positions the CDR1 and the CDR2 loops mainly over the HLA-DR4 and then the CDR3 loops over the peptide [28]. The two-step binding mechanism is consistent with TCR/p/MHC crystal structures which show that the CDR1 and CDR2 loops primarily contact the MHC, whereas the highly diverse CDR3 loops mainly interact with the peptide [29].

The first part of the notation (30D:A 30D:B 50D:A 50D:B), therefore, which refers to the first step of the mechanism, means that we first isolated the diagonal orientations in which residues  $\beta$ 30 and  $\beta$ 50 of CDR1 and CDR2, respectively, are within 4.5 Å distance of any residue of MHC  $\alpha$  and  $\beta$  chains. Subsequently, according to the second part of the notation ((97–98)D:P), related to the second step of the mechanism, we isolated the complexes for

which residues  $\beta$ 97 and  $\beta$ 98, at the apex of CDR3, are closer than 4.5 Å to atoms of the antigen peptide.

Analysis of TCR/p/MHC crystal structures clearly indicates that residues  $\beta$ 30 and  $\beta$ 50 of CDR1 and CDR2, respectively, and  $\beta$ 97 and  $\beta$ 98 of CDR3 make the most frequent contacts for TCRs to p/MHC [29].

### Docking of TCR-VB1<sup>VB</sup> and TCR-VB1<sup>OE</sup> to huCollp261/HLA-DR4 complex

The systems were digitized onto a 206×206×206 (VB1<sup>VB</sup>) and 202×202×202 (VB1<sup>OE</sup>) grid with the resolution of 0.69 Å. The same filtering procedure of 1j8h complex was applied, which corresponded to: 32D:A 32D:B 52D:A 52D:B (100–101)D:P where A, B stand for  $\alpha$  and  $\beta$  chain of HLA-DR4, respectively, D stands for  $\beta$  chain of TCR, P stands for huCollp261.

### Molecular dynamics simulations

The best structural model of the ternary complexes TCR-VB1<sup>OE</sup>/huCollp261/HLA-DR4 obtained from docking experiments, i.e. the model with the largest binding energy, largest interface ASA and that is likely to have biological meaning, was subjected to aqueous-phase MD simulations using GROMACS v.3.3.1 and employing the GROMOS96 force field [30]. The structure was immersed in a triclinic box with periodic boundary conditions and was solvated with explicit SPC water molecules. The system was then neutralized by 20 Na<sup>+</sup> counterions that were added at random positions to the bulk solvent. The dimensions of the box (9.0 nm×9.3 nm ×11.4 nm) were set to allow at least 1.2 nm between protein and box faces on each side. The final system consisted of 5182 protein atoms surrounded by 30000 water molecules. Before running simulation, the system was energy minimized for 1000 iterations of steepest descents and then equilibrated for 20 ps, during which the protein atoms were restrained. All restraints were then removed from the complexes and the temperature of system was brought to 300 K in a stepwise manner: 10-ps long MD runs were carried out at 50, 100, 200 and 250 K before the production runs were started at 300 K. The total length of simulation was 20 ns. Berendsen coupling was employed to maintain a constant temperature of 300 K with a coupling constant  $\tau$  of 0.1 ps. van der Waals interactions were modeled using 6–12 Lennard-Jones potentials with a 1.4 nm cutoff. Long-range electrostatic interactions were calculated using Particle Mesh Ewald method, with a cutoff for the real space term of 1.2 nm. Covalent bonds were constrained using LINCS algorithm. The time step employed was 2 fs and the coordinates were saved every 5 ps for analysis of MD trajectories which was carried out using the standard GROMACS tools *g\_rms*, *g\_rmsf* and *g\_hbond*. In the use of *g\_hbond*, a cutoff radius of 0.35 nm between donor and acceptor and a cutoff angle of 30° as geometric criteria were employed for the existence of a hydrogen bond.

This same MD protocol was applied to all VB1<sup>OE</sup> models emerging from the docking filtering procedure in order to calculate the interface ASA at the end of the MD run.

### Computational alanine scanning

Residues important for the stabilization of the complex were identified using Baker's alanine scanning procedure [31,32] and the Robetta web server (<http://www.robetta.org>). This approach calculates van der Waals' and electrostatic contributions to the free energy of binding. Positive values of  $\Delta\Delta G$  means that the alanine mutation is predicted to destabilize the complex and negative values indicate a stabilizing effect. Binding energy "hot spots" are defined for residues at the subunit-subunit interface, whose Ala

mutation causes a loss of free energy greater or equal to 1 kcal/mol.

## Results and Discussion

### Choice of TCR $\beta$ -chains

The expansion of T cells specific for huCollp261 in DR4<sup>+</sup> subjects was studied by “immunoscope”, in the peripheral blood mononuclear cells (PBMC), as previously reported for the only patient OE [15]. T cells carrying a rearrangement of VB1 and JB2.6 of 134b length were particularly interesting. This TCR is in fact enriched spontaneously in the inflamed synovia. In addition, its usage appeared specifically linked to DR4 haplotype and RA development. In fact, p139-specific cells using a rearrangement of this type were present in 4/6 DR4<sup>+</sup> patients during acute presentation of the disease (3 samples) and remission (2 samples). Meanwhile, we did not find cells carrying this rearrangement in any of 5 DR4<sup>+</sup> healthy donors.

In a separate set of experiments, not previously reported, we tested 4 more subjects, 2 DR1<sup>+</sup> DR4<sup>-</sup> RA patients, 1 DR1<sup>-</sup> DR4<sup>-</sup> RA patient and 1 patient suffering from acute arthritis of other origin. Out of this latter group of patients, one DR4<sup>-</sup> DR1<sup>+</sup> RA patient (patient VB) displayed the usage of a VB1-JB2.6 of 134b length in response to stimulation with huCollp261. Since T cells from this patient recognized huCollp261 most likely in the context of DR1, we sequenced and modeled the VB1-JB2.6 134b TCR  $\beta$ -chain from this patient to obtain a control for the specificity of the modeled interaction between the TCR  $\beta$ -chain obtained from DR4<sup>+</sup> patient OE, huCollp261 and the DR4 molecule itself.

### Modeling of TCR V $\beta$ domains and of huCollp261/HLA-DR4 complex

Homology models of VB1<sup>VB</sup> and VB1<sup>OE</sup> domains were constructed with the program MODELLER which implements an automated approach to comparative protein structure modeling by satisfaction of spatial restraints. The program automatically generates a set of restraints that includes the CHARMM forcefield, statistical preferences mined from PDB, and distance and dihedral restraints extracted from aligned templates, and generates a set of models which are consistent with all restraints. To find template structures, a specific BLAST sequence search of the Protein Data Bank (PDB) using default parameters was performed. The convergence matrix used to search the PDB produced significant alignments with the autoimmune TCR bound to a myelin basic protein self-peptide and a multiple sclerosis-associated HLA class II molecule (PDB code: 1zgl; identity 62%, similarity 75% with VB1<sup>VB</sup>; identity 63%, similarity 76% with VB1<sup>OE</sup>), the TCR V $\beta$ 5.1 in complex with staphylococcal enterotoxin K (SEK) (PDB code: 2nts; identity 62%, similarity 75% with VB1<sup>VB</sup>; identity 63%, similarity 76% with VB1<sup>OE</sup>), the TCR in complex with an anti-TCR Fab fragment derived from a mitogenic antibody (PDB code: 1nfd; identity 58%, similarity 72% with VB1<sup>VB</sup>; identity 58%, similarity 72% with VB1<sup>OE</sup>). The 1zgl complex in which human TCR is bound to HLA-DR2 protein appeared the appropriate template for comparative modeling of VB1<sup>VB</sup> and VB1<sup>OE</sup>  $\beta$  chains. The alignment of VB1<sup>VB</sup> and VB1<sup>OE</sup> domains with 1zgl generated by ClustalW is reported in Fig. 1A. The twenty different models generated for each V $\beta$  domain by MODELLER were carefully analyzed for energy value and probability density function (PDF) violations. Among the first three original models, which were quite similar with respect to violations and energy values, the model possessing the least root mean square deviation (RMSD) value with the backbone atoms of

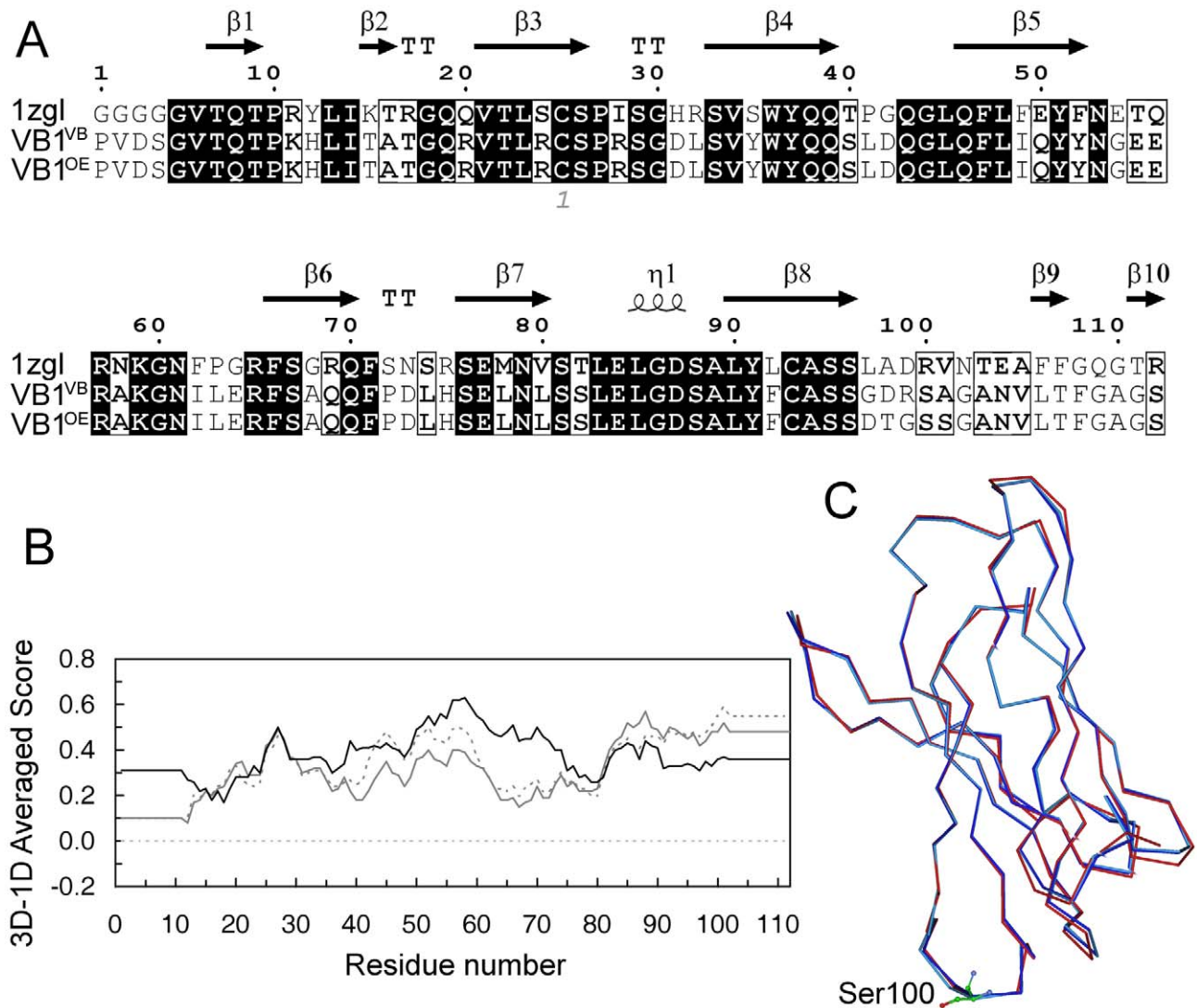
1zgl (0.23 Å for VB1<sup>VB</sup> and 0.26 Å for VB1<sup>OE</sup>), was selected for further study. The goodness of the folding was assessed by VERIFY-3D, which evaluates the compatibility of a given residue in a certain three-dimensional environment. As shown in Fig. 1B, the three/one-dimensional scores of our models are always positive and are similar to those obtained with the template structure of 1zgl. PROCHECK analysis indicates that the quality of the Ramachandran plots (97.9% of the residues in the allowed regions) were equivalent to those of the template structure. The superimposition of the modeled TCR V $\beta$  structures is reported in Fig. 1C. CDR3 Ser100 of both VB1<sup>VB</sup> and VB1<sup>OE</sup> is positioned at the apex and appears most accessible for interaction with peptide/HLA-DR4 complex. The flanking residues (Arg99 and Ala101 in VB1<sup>VB</sup>, Gly99 and Ser101 in VB1<sup>OE</sup>) may play important roles in the packing of the CDR3 loop structures.

Modeling of huCollp261 peptide (AGFKGEQGPKEP, position 1, P1 = F) bound to HLA-DR4 was performed using the simulated annealing protocol of MODELLER and the crystallographic complex between human collagen peptide (CII) 1168–1180 and HLA-DR4 as starting structure. It is worth noticing that the automatically generated three-dimensional model (Fig. 2) displays the same structural features proposed for the DR4 recognition of huCollp61 by Dessen et al. [22] simply on the basis of the analysis of the MHC class II structures determined to date. In fact, observing that the peptides in complexes with human and murine MHC class II molecules have remarkably similar conformations, with P1, P4, P6, P7 and P9 partially buried in pockets and P-2, P-1, P2, P3, P5, P8, P10 and P11 substantially exposed to solvents, Dessen et al. hypothesize how the CII(261–273) peptide could be aligned and modeled from the DR4/CII(1168–1180) structure. In agreement with Dessen et al. working hypothesis, in our computationally-generated model Phe263 (P1) of huCollp261 is bound in the large nonpolar pocket 1 which was occupied by Met in the CII(1168–1180)/DR4 complex; Glu266 (P4) is hydrogen bonded to Lys71 $\beta$  as does P4 Asp in CII(1168–1180); Pro269 (P7) fits in the shallow pocket 7 (Fig. 2). As expected, Gln267 (P5) and Lys270 (P8), extend prominently into solvent where can be contacted by TCR (Fig. 2). In addition, as hypothesized [22], CII(1168–1180) main-chain hydrogen bonds to DR4 are maintained in the model.

### Docking validation

To validate the docking method, we first applied our technique to the rebuilding of a known TCR/p/MHCII crystallographic structure. The complex of TCR V $\beta$  domain specific for the hemagglutinin antigen peptide (HA) with HLA-DR4 (PDB code: 1j8h) was selected for this purpose. FTDOCK was used to scan the relative orientations of the molecular complex in a systematic way, and produced 10,000 docking orientations. For the 10,000 docked models, we calculated the root mean square deviation (RMSD) of C $\alpha$  atoms of each model structure from the native complex structure (1j8h). We observed that there are only two complexes with RMSD < 4.5 Å (2.43 Å and 4.27 Å) and 17 complexes with 4.5 Å < RMSD < 8 Å. The rest have very high RMSD.

The 10,000 structures were then filtered by the distance constraints using available biological information. We first took into account the two-step binding mechanism [28] for TCR recognition as a means of filtering the docking results. In this model, initial TCR-MHC interactions, aided by minor contributions from TCR contacts with the peptide, guide the TCR to its ligand. This is followed by a final folding of the CDR3 loops of the TCR over the peptide [28]. Analysis of TCR/p/MHC crystallographic structures [29], which is consistent with the two-step binding mechanism, indicated the constraints for filtering procedure. We initially eliminated the models



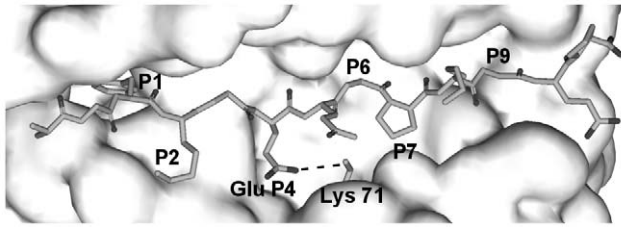
**Figure 1. Modeling of TCR-VB1<sup>VB</sup> and TCR-VB1<sup>OE</sup> domains.** A) Sequence alignment of VB1<sup>VB</sup> and VB1<sup>OE</sup> with 1zgl Vβ domain. Alignments were performed with ClustalW algorithm and ESPrInt software. Identical and similar amino acids are in dark and white boxes, respectively. The secondary structural elements are shown aligned to their respective sequences. Numbering of amino acids begins with the first amino acid residue of sequenced VB1<sup>VB</sup> and VB1<sup>OE</sup>; B) Residue-based quality assessment results obtained by the Verify3D program using the coordinates of 1zgl crystal structure (black) and the coordinates of the three-dimensional models of VB1<sup>VB</sup> and VB1<sup>OE</sup> (continuous grey and dashed grey, respectively). X axis: amino acid numbering of VB1<sup>VB</sup> and VB1<sup>OE</sup> starting from the first sequenced residue. Y axis: average three/one-dimensional scores for residues in a 21-residue sliding window; C) Cα trace of VB1<sup>VB</sup> (dark blue) and VB1<sup>OE</sup> (light blue) model structures superimposed onto the Cα trace of template 1zgl Vβ domain (red). Ser100, at the apex of CDR3 loop of VB1<sup>VB</sup> and VB1<sup>OE</sup>, is shown in ball and stick representation.  
 doi:10.1371/journal.pone.0011550.g001

in which residues Asnβ30 and Aspβ50 of CDR1 and CDR2, respectively, were positioned away (>4.5 Å) from MHC α and β chains; then, among the remaining predictions, only those that had a distance less than 4.5 Å between Leuβ97 and Proβ98 of CDR3 and any residue of the HA peptide, were kept for further analyses. The eight remaining models were examined using the molecular graphics display program DiscoveryStudio (Accelrys Inc.) and a further biological filter was applied following the structural analysis of Deng and Mariuzza [33] which points out how the three CDR loops of Vβ contact the central and C-terminal portion of peptide. Therefore, complexes for which the CDR3β does not focus on the central portion of the MHC-bound peptide (P5-P6) were eliminated [33]. This left four models which were energy-minimized. Among these, the complex possessing the largest interaction energy between TCR Vβ and p/MHC and the greatest buried surface area is the same

model structure exhibiting the lowest RMSD among the original 10,000 docked models (Cα RMSD = 2.43 Å). We were thus confident that FTDOCK can generate TCR/p/MHC model complexes close to native structures and this procedure was chosen to study the interaction of VB1<sup>VB</sup> and VB1<sup>OE</sup> TCR domains with huCollp261/HLA-DR4 complex.

#### Docking of VB1<sup>VB</sup> and VB1<sup>OE</sup> domains with huCollp261/HLA-DR4 complex

The same docking protocol and filtering procedure successfully used for method validation was then employed to investigate the docking of VB1<sup>VB</sup> and VB1<sup>OE</sup> domains with huCollp261/HLA-DR4 complex. The selection of residues for filtering constraints based on the two-step binding mechanism for TCR recognition [28] and on the analysis of TCR/p/MHC crystal structures [29].



**Figure 2. Model of huCollp261 peptide in the HLA-DR4 binding cleft.** The computationally generated model closely resembles the hypothetical model suggested by Dessen et al. (22) for the DR4 recognition of huCollp261. Phe263 fits into the P1 pocket and Glu266 (P4) hydrogen bonds to Lys71 $\beta$ . The MHC peptide-binding groove is represented with Connolly solid surface, whereas the ligand peptide is shown in stick representation.  
doi:10.1371/journal.pone.0011550.g002

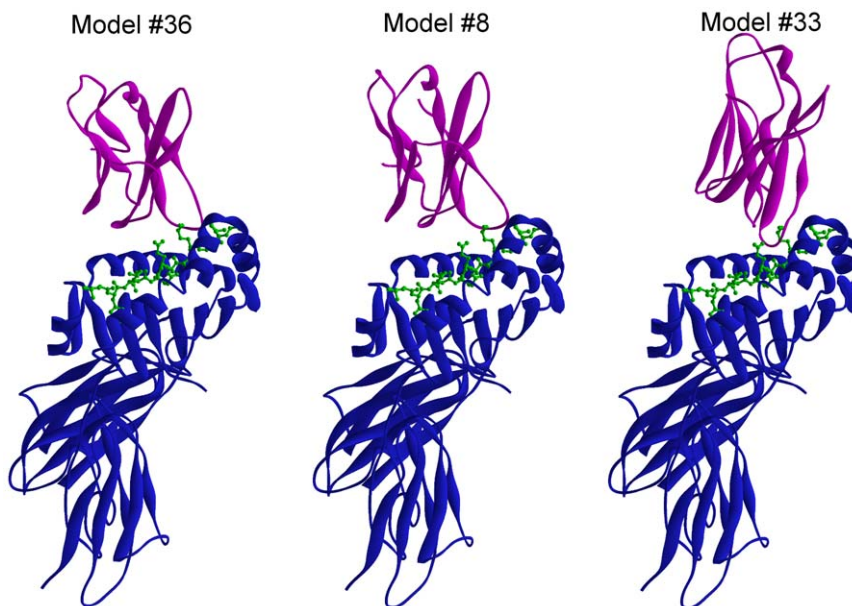
Therefore, to narrow down the predicted 10,000 conformations of FTDOCK to a manageable number we first requested Leu $\beta$ 32 and Tyr $\beta$ 52 (corresponding to TCR HA1.7 Asn $\beta$ 30 and Asp $\beta$ 50, respectively) to be positioned at a distance less than 4.5 Å from MHC  $\alpha$  and  $\beta$  chains. Then, according to the second step of the mechanism, we included only orientations where the peptide was within 4.5 Å distance from VB1<sup>VB</sup> Ser100 and Ala101 and from VB1<sup>OE</sup> Ser100 and Ser101 (corresponding to TCR HA1.7 Leu $\beta$ 97 and Pro $\beta$ 98, respectively). In doing so, seven complexes remained for both VB1<sup>VB</sup> and VB1<sup>OE</sup>. Interaction energies and interface Accessible Surface Area (ASA) for the selected minimized complexes are reported in Table 2. In the case of VB1<sup>VB</sup>, only three candidate complexes (namely models #17, #21 and #36 in FTDOCK ranking) show the V $\beta$  domain located over the central and C-terminal portions of the peptide. Analogously, of the seven candidate conformations for VB1<sup>OE</sup> complex, only three (namely model #8, model #33 and model #37 in FTDOCK ranking) display the V $\beta$  domain poised above the C-terminal half of the peptide. Considering that a higher interface ASA is an indication of a higher shape complementary [26] between the molecules, the

**Table 2.** Interaction energy and interface Accessible Surface Area (ASA) in the minimized complexes emerging from docking calculations.

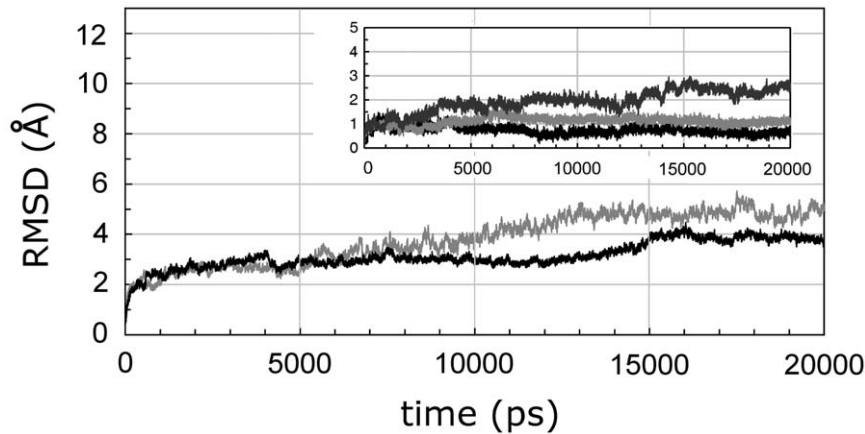
BV1 <sup>VB</sup>			BV1 <sup>OE</sup>		
FTDOCK score	Eint (kcal/mol)	ASA (Å <sup>2</sup> )	FTDOCK score	Eint (kcal/mol)	ASA (Å <sup>2</sup> )
8	-46.6	1176.5	8	-45.3	1181.1
17	-53.8	1224.6	13	-46.7	851.4
21	-56.6	1298.6	14	-54.5	1247.9
33	-34.6	1201.3	29	-43.5	977.0
36	-62.5	1370.5	33	-63.7	1384.8
52	-41.5	1065.3	34	-56.0	1198.1
56	-50.4	1235.2	37	-54.4	1077.0

doi:10.1371/journal.pone.0011550.t002

models with the largest binding energy and interface ASA were chosen as the most reasonable orientation. From the analysis of Table 2, model #36 in FTDOCK ranking was the most reasonable model for VB1<sup>VB</sup> and model #33 was the most reasonable for VB1<sup>OE</sup>. Visual inspection of candidate conformations reported in Table 2 also showed that the binding conformations #36 (VB1<sup>VB</sup>) and #8 (VB1<sup>OE</sup>) are very similar with overall RMSD of 1.36 Å when the C $\alpha$  atoms are optimally aligned. The three models emerging from the docking studies, namely #36<sup>VB</sup>, #8<sup>OE</sup> and #33<sup>OE</sup> are reported in Figure 3. If we look at the regions of the proteins involved in the interaction, we see an important difference among the models. CDR3 $\beta$  of model #36<sup>VB</sup> and #8<sup>OE</sup> focus on the peptide C-terminus, whereas in model #33<sup>OE</sup> CDR3 $\beta$  is positioned above P5-P6 of peptide as observed in binding topologies of autoimmune complexes [33]. As stated in the introduction, VB1<sup>VB</sup> recognizes huCollp261 bound to an HLA molecule different from DR4. Model #36<sup>VB</sup>, therefore,



**Figure 3. Molecular docking results.** Structural comparison of overall structures of TCR-V $\beta$ /huCollp261/HLA-DR4 complex models #36 (VB1<sup>VB</sup>), #8 (VB1<sup>OE</sup>) and #33 (VB1<sup>OE</sup>) emerging from the docking study. Color coding is as follows: magenta, TCR-V $\beta$  domain; blue, HLA-DR4; green, peptide.  
doi:10.1371/journal.pone.0011550.g003



**Figure 4. C $\alpha$ -RMSDs versus time.** The time evolution of the RMSD values for huCollp261/HLA-DR4 complex (grey line) and VB1<sup>OE</sup> TCR (black line). Inset shows individuals CDR loops: CDR1 (black line), CDR2 (light grey) and CDR3 (dark grey) computed through the 20 ns MD simulation of VB1<sup>OE</sup> TCR/huCollp261/HLA-DR4 ternary complex.  
doi:10.1371/journal.pone.0011550.g004

provides a sort of background model of the non-specific interaction that any TCR  $\beta$ -chain, generated by recombination of VB1 and BJ2.6 of the same length, engages with the huCollp261/HLA-DR4 complex. Since model #8<sup>OE</sup> is very similar to this non-specific interaction, we selected model #33<sup>OE</sup> as the complex structure for TCR-VB1<sup>OE</sup>/huCollp261/HLA-DR4 complex most likely provided of biological meaning.

### Molecular dynamics simulation

The model emerging from the docking was further investigated by MD simulation. Evaluation of structural drift is provided by analysis of the C $\alpha$  atom RMSDs from the initial structures as a function of time. The RMSDs of the VB1<sup>OE</sup> TCR and of the huCollp261/HLA-DR4 complex through the 20 ns trajectory were computed with respect to their corresponding initial structures. The C $\alpha$  RMSD values fluctuated for the last 5 ns of the MD run around values of  $3.9 \pm 0.1$  and  $4.8 \pm 0.2$  Å for VB1<sup>OE</sup> TCR and huCollp261/HLA-DR4, respectively (Fig. 4). The moderately large RMSD of the bound-MHC and, to a lesser extent, of the bound TCR V $\beta$  domain indicates the existence of conformations that are somewhat different from the starting structure. A closer inspection of the RMSD for the HLA-DR4 molecule indicates that its variation is mainly due to the contribution of the several loops in the  $\beta$ 2 domain. Analysis of RMSD for the TCR's CDRs (Fig. 4, inset) indicates that CDR1

and CDR2, which are positioned almost exclusively over the DR4 helix  $\alpha$ 1, show high stability reaching a plateau of 0.7 Å and 1.1 Å already after 5 ns. The CDR3 shows a higher deviation (C $\alpha$  RMSD values fluctuating around values of  $2.4 \pm 0.2$  Å in the last 5 ns) indicating that it significantly contributes to the overall deviation of the TCR  $\beta$ -chain. The RMSDs of the CDRs correspond to a structural feature of TCR/p/MHC complexes in which the CDR1 and CDR2 loops are significantly more rigid than CDR3 loop and show little or no rearrangement upon binding to the p/MHC complex. In contrast, the CDR3 loop appears highly flexible and mobile and undergoes substantial repositioning upon binding the p/MHC complex. Such a behavior of CDRs is consistent with the reported observation that CDR loops of the TCR display different conformations in the free and bound states [34].

### Interactions at the protein surface

The average structure over the last 5 ns of simulation was used to analyze the interactions at the protein surface. The interface

**Table 3.** Interaction energy and interface Accessible Surface Area (ASA) for the VB1<sup>OE</sup>/huCollp261/HLA-DR4 complexes averaged over the MD simulation runs.

FTDOCK score	Eint (kcal/mol)	ASA (Å <sup>2</sup> )
8	-114.1	1645.4
13	-67.4	1104.9
14	-108.7	1340.2
29	-90.3	1517.2
33	-126.8	1686.5
34	-93.4	1483.6
37	-66.4	964.5

doi:10.1371/journal.pone.0011550.t003

**Table 4.** van der Waals contacts<sup>†</sup> between VB1<sup>OE</sup> and huCollp261/HLA-DR4.

VB1 <sup>OE</sup>	HLA-DR4	huCollp261
Ser29	Ala61 $\alpha$ , Ala64 $\alpha$ , Val65 $\alpha$ , Ala68 $\alpha$	Pro269
Gly30	Ala61 $\alpha$ , Val65 $\alpha$	
Leu32	Gln57 $\alpha$ , Ala61 $\alpha$	
Tyr52	Gln57 $\alpha$ , Leu60 $\alpha$	
Arg57	Gln57 $\alpha$	
Thr98	Gln57 $\alpha$ , Gly58 $\alpha$ , Ala61 $\alpha$	
Gly99		Gln267
Ser100	Gln70 $\beta$	Gln267, Gly268, Pro269
Ser101	Leu67 $\beta$ , Gln70 $\beta$ , Lys71 $\beta$	Glu266, Gln267, Gly268, Pro269
Gly102	Leu67 $\beta$ , Gln70 $\beta$	Pro273
Ala103	Gln70 $\beta$	Pro269, Pro273
Asn104	Asp66 $\beta$	

<sup>†</sup>van der Waals contacts are  $\leq 4.0$  Å.

doi:10.1371/journal.pone.0011550.t004

**Table 5.** Intermolecular hydrogen bonds between VB1<sup>OE</sup> and huCollp261/HLA-DR4.

VB1 <sup>OE</sup>	huCollp261	HLA-DR4	Frequency (%) <sup>†</sup>
Tyr52 (OH)		Gln57 $\alpha$ (NE2)	52.24
Arg57 (NH2)		Gln57 $\alpha$ (NE2)	17.22
Arg57 (NH1)		Gln57 $\alpha$ (NE2)	20.74
Ser100 (O)		Gln70 $\beta$ (NE2)	52.24
Ser101 (OG)		Gln70 $\beta$ (O)	17.22
Ser101 (OG)	Glu266 (OE1)		20.74
Ser101 (N)	Gln267 (O)		52.36
Gly102 (N)		Gln70 $\beta$ (OE1)	81.62

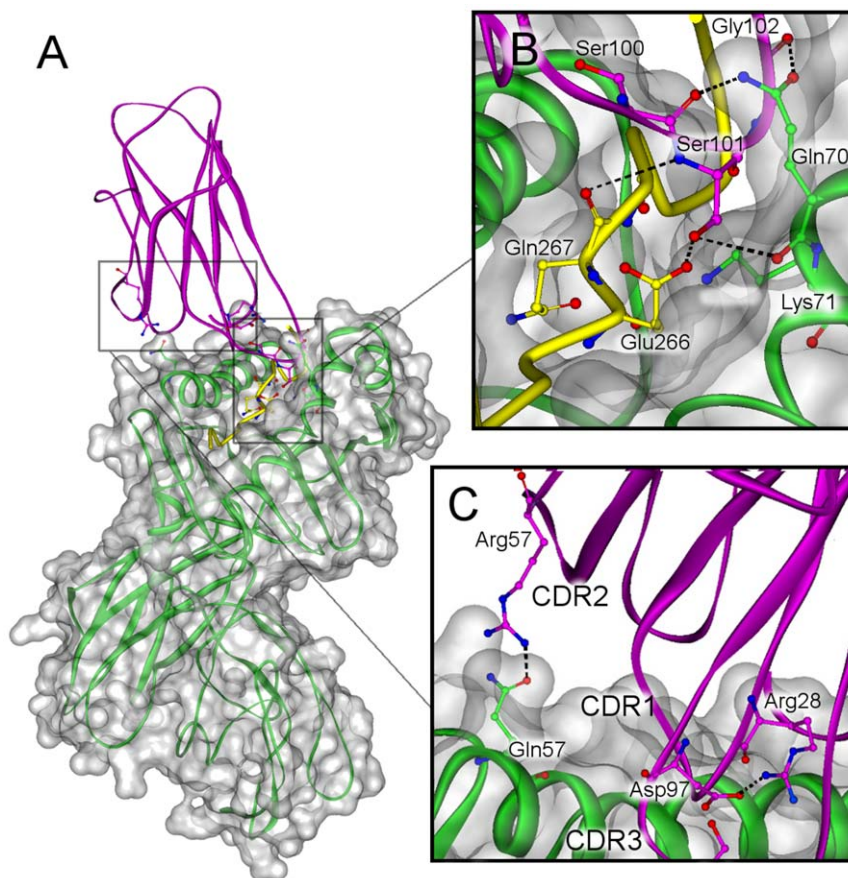
<sup>†</sup>Interactions were statistically monitored throughout the MD trajectory for a total of 5000 conformations.  
doi:10.1371/journal.pone.0011550.t005

ASA value resulted significantly increased (1686.5 Å<sup>2</sup>) indicating that the molecular dynamics run induced a significant improvement in the surface matching. All the VB1<sup>OE</sup> models emerging from the docking (Table 2) were further investigated by MD simulations to the aim of better evaluating the enhancement of

interface ASA in the selected model #33<sup>OE</sup>. The interaction energy and interface ASA for the seven VB1<sup>OE</sup>/huCollp261/HLA-DR4 complexes averaged over the MD simulation runs are reported in Table 3. By comparing Tables 2 and 3, we realize that the ASA value resulted essentially increased in models #8, #13, #29, #33 and #34, while a less pronounced increase in the ASA is observed for model #14; a decrease being instead registered for model #37. Notably, the selected model #33 proved to be the model with the largest binding energy and interface ASA also at the end of the MD simulations, thus confirming the validity of the model selection.

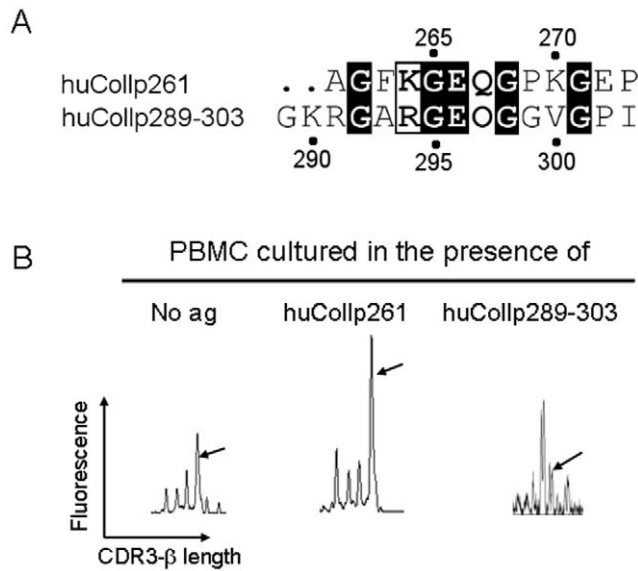
Interactions between TCR and huCollp261/DR4 complex are mainly restricted to van der Waals contacts, with limited juxtaposition of hydrophobic surfaces (Table 4). Of note, several TCR V $\beta$  residues make contacts to HLA-DR4 Gln57 $\alpha$ , Ala61 $\alpha$  and Gln70 $\beta$ , which stand out as MHC class II conserved contact residues from the analysis of TCR/p/MHC crystallographic complexes [29].

Identity of residues 67–74 of MHCII  $\beta$ -chain in DR4 and DR1 (“shared epitope” region) has been correlated with increased risk for RA [35]. In the predicted model we observe that some residues of the CDR3 loop 97–101 of VB1<sup>OE</sup> (DTGSS) form van der Waals contacts with this “shared epitope” region. As found for recognition of a myelin basic protein self-peptide by TCR 3A6



**Figure 5. Positioning of the TCR VB1<sup>OE</sup> domain over the huCollp261/HLA-DR4 complex.** A) Overall three-dimensional structure of TCR-VB1<sup>OE</sup>/huCollp261/HLA-DR4 ternary complex generated using molecular dynamics simulations. The backbone structure of VB1<sup>OE</sup> (magenta) and HLA-DR4 (green) are displayed in solid ribbon representation; huCollp261 (yellow) is shown in tube representation. HLA-DR4 is represented with Connolly transparent solid surface; (B) Zoom view of the peptide binding cleft showing the hydrogen-bonding interactions involving CDR3 $\beta$ ; (C) Zoom view of the binding surface showing TCR $\beta$  residues important for the stabilization of the complex. Atoms are shown in ball and stick representation and colored by atom type with the exception of C atoms, colored by subunit.  
doi:10.1371/journal.pone.0011550.g005





**Figure 6. T cells carrying the BV1-BJ2.6 TCR  $\beta$  chain do not expand in response to stimulation with peptide huCollp289–303.** A) Sequence alignment of peptide huCollp261 with huCollp289–303 performed with ClustalW algorithm and ESPrpt software. Identical and similar amino acids are in dark and white boxes, respectively. B) BV1-BJ2.6 spectra obtained from peripheral blood mononuclear cells (PBMC) of patient OE stimulated *in vitro* in the absence of added antigen (background) or in the presence of 20  $\mu$ g/ml of peptides huCollp261 or huCollp289–303. PBMC were obtained from patient OE during a clinical relapse of disease. They were cultured and cDNA was obtained. BV-BJ spectratyping for rearrangements of BV1 and BJ2.6 was performed as described [15]. The spectra report the distribution of each TCR  $\beta$ CDR3 as a function of its length, where peaks are separated by a 3-base, i.e. one amino acid, difference. The fluorescence of each peak is a function of the amount of CDR3 of each length. Arrows indicate the 134b peak corresponding to the TCR  $\beta$ -chain under study. doi:10.1371/journal.pone.0011550.g006

(PDB code: 1zgl) [17], few hydrogen bonds are observed between TCR V $\beta$  domain and the huCollp261 peptide (Table 5), a feature that likely contributes to low affinity binding and the observed crossreactivity in autoimmune TCRs [36,37].

Our MD simulation showed that peptide contacts are made primarily through the CDR3 loop (Table 4, Table 5) which overlays the central region of the peptide-binding groove (Fig. 5A). The side chain of peptide Glu266 (P4) is the first CDR3-contact residue and forms a hydrogen bond with Ser101 OG atom (Fig. 5B). At various points during the molecular dynamics run, Ser101 OG atom alternately forms hydrogen bonds with Glu266 OE1 atom and the carbonyl oxygen atom of DR4 Gln70 $\beta$  (Fig. 5B). Due to the binding of TCR Ser101 $\beta$  to Glu266, Lys71 $\beta$  of MHC “shared epitope” moves away from Glu266 causing a separation of 8.8 Å between the side chains of Lys71 and Glu266, a distance that appears to disfavor a direct contact. Thus, the Lys71 $\beta$ -Glu266 interaction, suggested by Dessen [22] and by our modeling (Fig. 2) is significantly affected by TCR binding. The second CDR3-contact residue of huCollp261 is Gln267 (P5) whose backbone oxygen atom is hydrogen bonded to Ser101 N atom (Fig. 5B). Gly268 (P6), Pro269 (P7) and Pro273 (P11) provide much weaker interactions to the modeled TCR, with Pro269 and Pro273 loosely interacting with the hydrophobic Ala103 (Table 4).

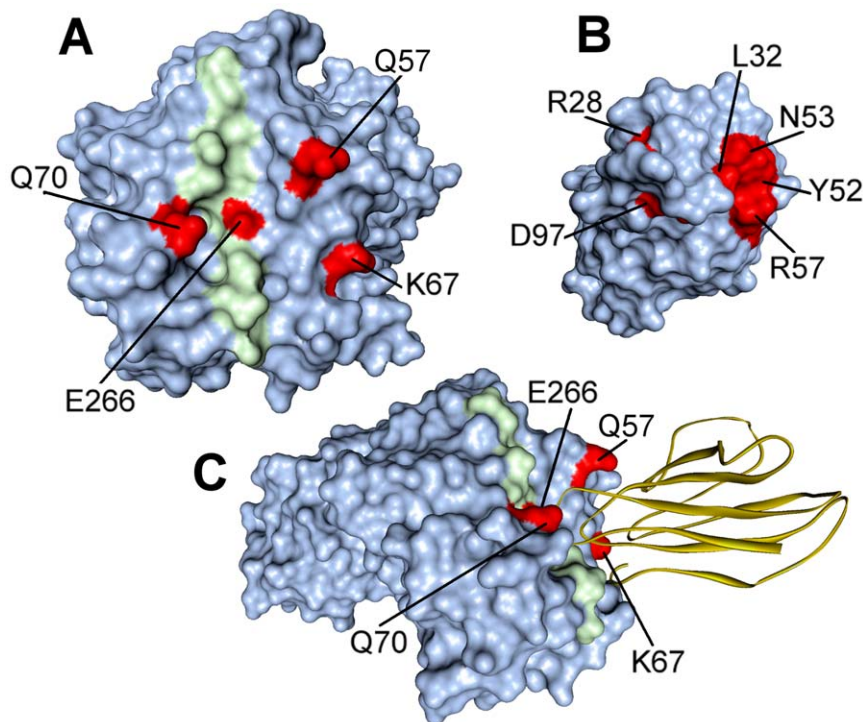
We had the opportunity to test the hypothesis that Gln267 may play a relevant role in the recognition of huCollp261 by the TCR under study, by stimulating in parallel PBMC from patient OE with a peptide encompassing a subdominant epitope of the same

protein, namely peptide huCollp289–303 (sequence GKRGAR-GEQGGVGP, where O is hydroxyproline, Hyp). We showed that 25% of huCollp261-specific T cells recognize also an epitope contained in peptide huCollp289–303 [15]. In Figure 6, panel A, the sequence alignment between huCollp261 and huCollp289–303 is shown. A functional study [38] identified the “core” epitope recognized by T cells within peptide 259–273 of human collagen II in the region encompassing residues 263–268. We can observe that there are three amino acid residues that are conserved between huCollp261 and huCollp289–303 within this area, and that Phe263 of huCollp261 (the most functionally relevant residue for binding to HLA-DR4 according to the same study [38]) is aligned with Ala293 of huCollp289–303. TCR of T cells cross-recognizing huCollp261 and huCollp289–303 will thus possibly contact conserved residues (i.e. Gly265/295, Glu266/296, Gly268/298) all of which are indicated by the above-mentioned study as residues involved in the contact with the TCR. On the contrary, TCRs of T cells selectively recognizing huCollp261 may contact residues that are different between the two peptides, namely Lys264 (replaced by Arg294 in huCollp289–263), whose functional role however appears more relevant in the binding to DR4 than in contacting the TCR [38], and Gln267 (replaced by Hyp297 in huCollp289–263) that is indicated as a main TCR contacting residue in Ref. [38] in agreement with our modeling results.

If the TCR studied here (VB1<sup>OE</sup>) needs Gln267 for recognition of huCollp261, as suggested by our model, peptide huCollp289–303 will fail to stimulate and expand T cells carrying this receptor. This is actually the case.

We cultured PBMC from patient OE in the presence of huCollp261 or huCollp289–303, following the protocol described [15]. Results are shown in panel B of Figure 6, where arrows indicate the peaks corresponding to the product of the VB1<sup>OE</sup> chain in the immunoscope spectra. T cells carrying the VB1<sup>OE</sup> chain proliferate in response to huCollp261, and expansion of the corresponding peak (indicated by the arrow) can be observed, as expected. On the contrary, no proliferation is observed when the same cells are stimulated with huCollp289–303. These experimental findings are in line with the results of the computational modeling proposed.

The role of individual amino acids in stabilizing the complex was inspected by computational alanine scanning [31]. Residues with  $\Delta\Delta G > 1$  kcal/mol are called “hotspots” (Fig. 7) and are listed in Table 6. Recent computational alanine scanning studies [39] have shown that location of the “hotspots” may vary among the various TCR/p/MHC structures. Remarkably, two “hotspots” on TCR VB1<sup>OE</sup> (Leu32 in CDR1 and Tyr52 in CDR2, see Table 6) interact with DR4 $\alpha$  Gln57 and Ala61 (Table 4) which form conserved contacts with TCR V $\beta$  domain [29]. This finding is in close agreement with the alanine scanning analyses of TCR HA1.7 bound to HA/HLA-DR4 and to HA/HLA-DR1 (PDB codes: 1j8h and 1fyt, respectively) [39]. Also computational mutation of CDR3 Asp97 has unfavorable effects on the stability of the complex (Table 6). In the dynamically equilibrated model, Asp97 forms a salt bridge with Arg28 of CDR1 that may be important for maintaining a correctly oriented loop structure (Fig. 5C). We also find that DR4 $\alpha$  Gln57Ala is able to cause destabilization at the interface (Table 6, Fig. 5C) and this result is consistent with the above-mentioned structural role of Gln57 $\alpha$ . Alanine scanning results also indicate DR4 Gln70 $\beta$  as a residue whose mutation dramatically affects TCR/p/MHC complex. The network of hydrogen bonds established by Gln70 $\beta$  with Ser100, Ser101 and Gly102 of the TCR V $\beta$  domain is shown in Fig. 5B. Taken together, these findings lead to the suggestion that the “shared epitope” region plays an essential role in influencing the



**Figure 7. Hotspots predictions in the TCR VB1<sup>OE</sup>-huCollp261/HLA-DR4 interface.** Residues predicted to be hotspots ( $\Delta\Delta G > 1.0$  kcal/mol) are shown in red, the huCollp261 peptide in green and the TCR is represented as a yellow ribbon. A) Unbound huCollp261/HLA-DR4 surface; B) Unbound TCR-VB1<sup>OE</sup> surface; C) TCR-VB1<sup>OE</sup>/huCollp261/HLA-DR4 complex.  
doi:10.1371/journal.pone.0011550.g007

strength of T cell recognition. In turn, the strength of the interaction between TCR and p/MHC complex influences polarization of T cells, since a strong stimulation leads to acquisition of the pathogenic Th1 phenotype [40,41]. Thus the direct engagement of the “shared epitope” by the TCR would contribute to differentiation of T cell specific for huCollp261 towards a pathogenic phenotype and promote the development of RA.

## Conclusions

The molecular mechanism of collagenII(261–273)/HLA-DR4 recognition by a TCR V $\beta$  domain characteristic of a DR4<sup>+</sup> patient affected by rheumatoid arthritis was investigated using molecular

modeling, protein-protein docking, and molecular dynamics simulations. It is clear that the possibility of controlling the clonotypic expansion strictly derives from the knowledge of the three-dimensional structure of the complex of TCR with the putative antigen. It is also true that post-transcriptional modifications of collagen can occur that modify the peptide bound to the DR4 molecule. Yet the putative natural antigen unmodified by posttranslational events likely represents the very early initial trigger of the loss of tolerance occurring under genetic control in RA, as observed in the collagen type II induced model of arthritis, possibly along with the posttranslationally modified antigen [42].

Herein, the proposed model finds a correspondence with a large body of existing experimental data and allows the identification of key residues involved in complex stability and specificity. As expected, key residues belong to the region 97–101 of V $\beta$  that distinguishes the TCR of the DR4<sup>+</sup> patient from that of a DR4<sup>-</sup> patient. Furthermore, the simulations presented here suggest that the “shared epitope”, common to the RA-predisposing alleles HLA-DR4 and HLA-DR1, directly contributes to the engagement of the TCR itself.

Nowadays, PCR based methods can produce large numbers of sequences of candidate antigen-specific TCR, specially for the  $\beta$ -chain.

The molecular modeling method we describe will prove useful to examine e. g. the variability of the recognition for the same p/MHC complex by different TCRs.

Although the presented strategy should be validated by comparison with mutagenesis experiments involving variations in either peptide or the TCR-V $\beta$  molecule, knowledge of the interactions and key binding residues at the interface between TCR and p/MHC complexes, obtained by pooling information from several of these models, will produce a detailed map of the

**Table 6.** Computational alanine scanning-based free energies for the TCR-VB1<sup>OE</sup>/huCollp261/HLA-DR4 complex.

Mutation	Protein	$\Delta\Delta G_{\text{bind}}$ (kcal/mol)
Arg28Ala	TCR CDR1 $\beta$	3.92
Leu32Ala	TCR CDR1 $\beta$	1.33
Tyr52Ala	TCR CDR2 $\beta$	1.61
Asn53Ala	TCR CDR2 $\beta$	2.16
Arg57Ala	TCR CDR2 $\beta$	1.18
Asp97Ala	TCR CDR3 $\beta$	2.11
Glu266Ala	huCollp261	1.55
Gln57Ala	DR4 $\alpha$	1.96
Lys67Ala	DR4 $\alpha$	1.50
Gln70Ala	DR4 $\beta$	2.14

doi:10.1371/journal.pone.0011550.t006

recognized surface, thereby providing insights into the processes of self- and allo-recognition.

This represents the basis to envision any future strategy to develop tools capable of damping the autoreactivity or to switch-off the autoreactive signal occurring from the interaction.

## References

- Davis MM, Bjorkman PJ (1988) T-cell antigen receptor genes and T-cell recognition. *Nature* 334: 395–402.
- Starr TK, Jameson SC, Hogquist KA (2003) Positive and negative selection of T cells. *Annu Rev Immunol* 21: 139–176.
- Russell AS (2008) Quality-of-life assessment in rheumatoid arthritis. *Pharmacoeconomics* 26: 831–846.
- Nepom GT, Byers P, Seyfried C, Healey LA, Wilske KR, et al. (1989) HLA genes associated with rheumatoid arthritis. Identification of susceptibility alleles using specific oligonucleotide probes. *Arthritis Rheum* 32: 15–21.
- Wordsworth BP, Lanchbury JS, Sakkas LI, Welsh KI, Panayi GS, et al. (1989) HLA-DR4 subtype frequencies in rheumatoid arthritis indicate that DRB1 is the major susceptibility locus within the HLA class II region. *Proc Natl Acad Sci U S A* 86: 10049–10053.
- Gregersen PK, Silver J, Winchester RJ (1987) The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. *Arthritis Rheum* 30: 1205–1213.
- 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: 915–926.
- Signorelli KL, Watts LM, Lambert LE (1995) The importance of DR4Dw4 beta chain residues 70, 71, and 86 in peptide binding and T cell recognition. *Cell Immunol* 162: 217–224.
- Hiraiwa A, Yamanaka K, Kwok WW, Mickelson EM, Masewicz S, et al. (1990) Structural requirements for recognition of the HLA-Dw14 class II epitope: a key HLA determinant associated with rheumatoid arthritis. *Proc Natl Acad Sci U S A* 87: 8051–8055.
- Watson WC, Tooms RE, Carnesale PG, Dutkowsky JP (1994) A case of germinal center formation by CD45RO T and CD20 B lymphocytes in rheumatoid arthritic subchondral bone: proposal for a two-compartment model of immune-mediated disease with implications for immunotherapeutic strategies. *Clin Immunol Immunopathol* 73: 27–37.
- Londei M, Savill CM, Verhoef A, Brennan F, Leech ZA, et al. (1989) Persistence of collagen type II-specific T-cell clones in the synovial membrane of a patient with rheumatoid arthritis. *Proc Natl Acad Sci U S A* 86: 636–640.
- Kjellen P, Brunsberg U, Broddefalk J, Hansen B, Vestberg M, et al. (1998) The structural basis of MHC control of collagen-induced arthritis; binding of the immunodominant type II collagen 256–270 glycopeptide to H-2Aq and H-2Ap molecules. *Eur J Immunol* 28: 755–767.
- Fridkis-Hareli M, Rosloniec EF, Fugger L, Strominger JL (2000) Synthetic peptides that inhibit binding of the collagen type II 261–273 epitope to rheumatoid arthritis-associated HLA-DR1 and -DR4 molecules and collagen-specific T-cell responses. *Hum Immunol* 61: 640–650.
- Garcia KC, Teyton L, Wilson IA (1999) Structural basis of T cell recognition. *Annu Rev Immunol* 17: 369–397.
- Ria F, Penitente R, De Santis M, Nicolo C, Di Sante G, et al. (2008) Collagen-specific T-cell repertoire in blood and synovial fluid varies with disease activity in early rheumatoid arthritis. *Arthritis Res Ther* 10: R135.
- Ria F, van den Elzen P, Madakamutil LT, Miller JE, Maverakis E, et al. (2001) Molecular characterization of the T cell repertoire using immunoscope analysis and its possible implementation in clinical practice. *Curr Mol Med* 1: 297–304.
- Li Y, Huang Y, Lue J, Quandt JA, Martin R, et al. (2005) Structure of a human autoimmune TCR bound to a myelin basic protein self-peptide and a multiple sclerosis-associated MHC class II molecule. *EMBO J* 24: 2968–2979.
- Gouet P, Courcelle E, Stuart DI, Metz F (1999) ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* 15: 305–308.
- Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 234: 779–815.
- Luthy R, Bowie JU, Eisenberg D (1992) Assessment of protein models with three-dimensional profiles. *Nature* 356: 83–85.
- Morris AL, MacArthur MW, Hutchinson EG, Thornton JM (1992) Stereochemical quality of protein structure coordinates. *Proteins* 12: 345–364.
- Dessen A, Lawrence CM, Cupo S, Zaller DM, Wiley DC (1997) X-ray crystal structure of HLA-DR4 (DRA\*0101, DRB1\*0401) complexed with a peptide from human collagen II. *Immunity* 7: 473–481.
- Gabb HA, Jackson RM, Sternberg MJ (1997) Modelling protein docking using shape complementarity, electrostatics and biochemical information. *J Mol Biol* 272: 106–120.
- Katchalski-Katzir E, Shariv I, Eisenstein M, Friesem AA, Aflalo C, et al. (1992) Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques. *Proc Natl Acad Sci U S A* 89: 2195–2199.
- Brooks BR, Brooks CL, 3rd, Mackerell AD, Jr., Nilsson L, Petrella RJ, et al. (2009) CHARMM: the biomolecular simulation program. *J Comput Chem* 30: 1545–1614.
- Jones S, Thornton JM (1996) Principles of protein-protein interactions. *Proc Natl Acad Sci U S A* 93: 13–20.
- Hennecke J, Wiley DC (2002) Structure of a complex of the human alpha/beta T cell receptor (TCR) HA1.7, influenza hemagglutinin peptide, and major histocompatibility complex class II molecule, HLA-DR4 (DRA\*0101 and DRB1\*0401): insight into TCR cross-restriction and alloreactivity. *J Exp Med* 195: 571–581.
- Wu LC, Tuot DS, Lyons DS, Garcia KC, Davis MM (2002) Two-step binding mechanism for T-cell receptor recognition of peptide MHC. *Nature* 418: 552–556.
- Rudolph MG, Stanfield RL, Wilson IA (2006) How TCRs bind MHCs, peptides, and coreceptors. *Annu Rev Immunol* 24: 419–466.
- Stocker U, van Gunsteren WF (2000) Molecular dynamics simulation of hen egg white lysozyme: a test of the GROMOS96 force field against nuclear magnetic resonance data. *Proteins* 40: 145–153.
- Kortemme T, Baker D (2002) A simple physical model for binding energy hot spots in protein-protein complexes. *Proc Natl Acad Sci U S A* 99: 14116–14121.
- Kortemme T, Kim DE, Baker D (2004) Computational alanine scanning of protein-protein interfaces. *Sci STKE* 2004: 1–8.
- Deng L, Mariuzza R (2007) Recognition of self-peptide-MHC complexes by autoimmune T-cell receptors. *Trends Biochem Sci* 32: 500–508.
- 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.
- Hammer J, Gallazzi F, Bono E, Karr RW, Guenet J, et al. (1995) Peptide binding specificity of HLA-DR4 molecules: correlation with rheumatoid arthritis association. *J Exp Med* 181: 1847–1855.
- Wilson D, Pimilla C, Wilson D, Schroder K, Boggiano C, et al. (1999) Immunogenicity. I. Use of peptide libraries to identify epitopes that activate clonotypic CD4+ T cells and induce T cell responses to native peptide ligands. *J Immunol* 163: 6424–6434.
- Hausmann S, Martin M, Gauthier L, Wucherpfennig K (1999) Structural features of autoreactive TCR that determine the degree of degeneracy in peptide recognition. *J Immunol* 162: 338–344.
- Rosloniec E, Whittington K, Zaller D, Kang A (2002) HLA-DR1 (DRB1\*0101) and DR4 (DRB1\*0401) use the same anchor residues for binding an immunodominant peptide derived from human type II collagen. *J Immunol* 168: 253–259.
- Collins EJ, Riddle DS (2008) TCR-MHC docking orientation: natural selection, or thymic selection? *Immunol Res* 41: 267–294.
- Iezzi G, Scotet E, Scheidegger D, Lanzavecchia A (1999) The interplay between the duration of TCR and cytokine signaling determines T cell polarization. *Eur J Immunol* 29: 4092–4101.
- Badou A, Savignac M, Moreau M, Leclerc C, Foucras G, et al. (2001) Weak TCR stimulation induces a calcium signal that triggers IL-4 synthesis, stronger TCR stimulation induces MAP kinases that control IFN-gamma production. *Eur J Immunol* 31: 2487–2496.
- Sakurai Y, Brand D, Tang B, Rosloniec E, Stuart J, et al. (2006) Analog peptides of type II collagen can suppress arthritis in HLA-DR4 (DRB1\*0401) transgenic mice. *Arthritis Res Ther* 8: R150.

## Author Contributions

Conceived and designed the experiments: MCDR BG FR. Performed the experiments: MCDR CB CCA DP GDS. Analyzed the data: MCDR BG GF MDS. Wrote the paper: MCDR FR.