

# Development of Immune-Specific Interaction Potentials and Their Application in the Multi-Agent-System Vacclmm

Anna Lena Woelke<sup>1\*</sup>, Joachim von Eichborn<sup>1</sup>, Manuela S. Murgueitio<sup>1</sup>, Catherine L. Worth<sup>1</sup>, Filippo Castiglione<sup>2</sup>, Robert Preissner<sup>1</sup>

<sup>1</sup> Institute for Physiology, Charité Universitätsmedizin Berlin, Berlin, Germany, <sup>2</sup> Institute for Computing Applications, National Research Council of Italy, Rome, Italy

## Abstract

Peptide vaccination in cancer therapy is a promising alternative to conventional methods. However, the parameters for this personalized treatment are difficult to access experimentally. In this respect, *in silico* models can help to narrow down the parameter space or to explain certain phenomena at a systems level. Herein, we develop two empirical interaction potentials specific to B-cell and T-cell receptor complexes and validate their applicability in comparison to a more general potential. The interaction potentials are applied to the model Vacclmm which simulates the immune response against solid tumors under peptide vaccination therapy. This multi-agent system is derived from another immune system simulator (C-ImmSim) and now includes a module that enables the amino acid sequence of immune receptors and their ligands to be taken into account. The multi-agent approach is combined with approved methods for prediction of major histocompatibility complex (MHC)-binding peptides and the newly developed interaction potentials. In the analysis, we critically assess the impact of the different modules on the simulation with Vacclmm and how they influence each other. In addition, we explore the reasons for failures in inducing an immune response by examining the activation states of the immune cell populations in detail. In summary, the present work introduces immune-specific interaction potentials and their application to the agent-based model Vacclmm which simulates peptide vaccination in cancer therapy.

**Citation:** Woelke AL, von Eichborn J, Murgueitio MS, Worth CL, Castiglione F, et al. (2011) Development of Immune-Specific Interaction Potentials and Their Application in the Multi-Agent-System Vacclmm. PLoS ONE 6(8): e23257. doi:10.1371/journal.pone.0023257

**Editor:** Jean Peccoud, Virginia Tech, United States of America

**Received:** March 31, 2011; **Accepted:** July 9, 2011; **Published:** August 17, 2011

**Copyright:** © 2011 Woelke 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 study was supported by the Deutsche Forschungsgemeinschaft (DFG) (SFB 449), the European Union (EU) project SynSys, the Deutsche Krebshilfe, and the Federal Ministry of Education and Research (BMBF) project MedSys (0315450). AW was funded by the graduate school "Computational Systems Biology" (GRK1772/1 of the DFG). JvE was funded by the International Research Training Group "Genomics and Systems Biology of Molecular Networks" (GRK1360 of the DFG). The open access charge was funded by DFG, BMBF and EU. 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: anna-lena.woelke@charite.de

## Introduction

Cancer is still one of the major causes of death in industrial nations, although in principle the immune system is able to eradicate a tumor. Bearing that in mind, many studies have tried to trigger an anticancer immune response using different methods, e.g. adoptive cell transfer, cytokine therapy or vaccination schedules [1]. Immune therapy is promising, but its success has been limited so far. The main reason is that the mechanisms of the tumor-immune-interplay are still poorly understood. A huge amount of, sometimes conflicting, data has accumulated, which can be difficult to interpret. Therefore, it is desirable to have a simplified model able to highlight at the system level the main processes of the phenomenon. In addition, *in silico* experiments are far less expensive, less time consuming and a lot more flexible in terms of parameter changes.

We have described the main theoretical modeling techniques, differential equations and rule-based models, and their application to tumor immunology elsewhere [2]. For this project, we have chosen a rule-based model because of its capability to characterize every single cell or molecule in its location, developmental state and specificity.

The aim of our present study is to support peptide vaccination approaches in cancer therapy by modeling the specific tumor-immune interaction in a realistic fashion. For that purpose, we integrated a previously published model of the tumor-immune interplay [3] with a detailed description of the immune receptor-ligand interactions based on structural and sequence information. To our knowledge, this is the first approach simulating peptide vaccination in cancer treatment that takes the peptide sequence into account explicitly. An analogical approach designed for generic infections has been described by Rapin et al. [4].

## Rule-Based Modeling for Simulating the Immune System

Rule-based models are composed of discrete agents identifiable within a spatial environment. The agents interact, move and change their state according to behavioral rules in discrete time steps.

One of the first approaches to simulating the immune system using a cellular automaton was introduced in 1992 by Celada and Seiden [5]. Their cellular automaton called ImmSim used very simple rules but was able to reproduce several phenomena in immunology, e.g. clonal expansion of B- and T-cells after stimulation or the different time-lines of the first and second

immunization. To account for specificity of the immune receptors, they developed a representation in the form of bit-strings that had to be complementary to favor an interaction between the immune cells [6]. Within the model, they examined optimal ranges to induce a sufficient immune response for some generic parameters such as the number of major histocompatibility complexes (MHCs) per individual or the number of self-peptides compared to the whole diversity of protein sequences. The ImmSim model has been extended and improved by several research groups (see below); the present study itself is based on an implementation of the Celada-Seiden ImmSim automaton.

Since 1992 the agent-based modeling community in theoretical biology has grown. Apart from ImmSim, other rule-based immune simulations have recently been developed which are briefly reviewed here.

The Basic Immune Simulator aims at understanding the interplay between the innate and the adaptive immune response [7]. It is an agent-based model composed of three virtual spaces, the parenchymal tissue, the secondary lymph node and the lymphatic/humoral circulation. The immune response to a localized viral infection is simulated, during which an immunity gain or loss or a hyper-response might occur. The Basic Immune Simulator was updated in 2010, now including new cell types and enhanced behavioral rules [8]. Using the new version, a network analysis was applied defining the cell types as nodes and their interaction as edges.

SIMMUNE is more a modeling environment than a model of a certain phenomenon [9]. Its purpose is to investigate how context adaptive behavior might emerge from local cell-cell and cell-molecule interactions. The model is composed of a cellular automaton on a molecular level. Molecules can be defined that interact according to the behavior they get equipped with. These entities move within a discrete lattice, which can be subdivided into different compartments. Since it takes a generic approach to the modeling of cell biology, SIMMUNE is able to simulate a wide area of signal cascades, not only those related to immunology.

SIMISYS is focused on the different stages of bacterial infection [10]. The model is based on a cellular automaton and composed of two different grids representing a blood vessel and a lymph node. By means of the simulation, the authors want to learn more about the interplay between the innate and adaptive immune system.

Another rule-based study is the tumor-immune model by Mallet and De Pillis [11]. They use a hybrid form of a cellular automaton with partial differential equations. The purpose of this model is to examine the effects of cytotoxic lymphocytes and natural killer cells in the early stages of tumor growth under limited nutrient supply.

The ImmSim model has also been enlarged, refined and applied to a variety of phenomena. Castiglione et al. translated ImmSim to ANSI C language and applied their refined model, C-ImmSim, to several immunological phenomena, such as Epstein-Barr virus infection [12], TH1/2 differentiation [13], HIV-infection [14] and immune therapy of cancer [3]. C-ImmSim applied to cancer simulates a solid tumor under immune therapy, including all main mechanisms of the humoral and cellular adaptive immune response. Lollini et al. used the architecture of C-ImmSim to build SimTriplex, a model for HER2/neu transgenic mice treated with the Triplex cell vaccine [15]. This model was extended and applied to lung cancer metastasis in MetastaSim, which is a hybrid agent-based/differential equation model [16].

For our research interest of studying peptide vaccination in cancer treatment, the model C-ImmSim was the most suitable. Indeed it provides all the important behavioral rules for immune

and cancer cells. The only major drawback of C-ImmSim is that the immune receptors are modeled as bit-strings, having no direct translation to amino acid sequences. Therefore, we decided to extend C-ImmSim, producing a sequence-based version, VaccImm, where all calculations depend on the amino acid sequence of immune receptors, MHCs and antigens. In VaccImm, experimental and clinical data of cancer targets, expression data or MHC-genotypes are directly integrated into the model and the simulation predicts the success of peptide vaccination taking the amino acid sequence into account.

In order to understand the architecture of the model, C-ImmSim is described in more detail in the next chapter. Afterwards, the adoption of amino acid sequences in place of bit-strings is discussed.

## C-ImmSim

Our present model VaccImm is based on C-ImmSim [3], an agent-based model written in ANSI C language. C-ImmSim is composed of a Cartesian 3D lattice that contains cancer cells and several types of immune-cells moving from one lattice point to another in discrete time steps corresponding to 8 hours of real life. The different types of immune cells can be classified as helper T-cells (TH), cytotoxic T-cells (TC), B-cells, macrophages and dendritic cells (DCs). In addition, the model includes different classes of molecules comprising antibodies, antigens, Interleukin-2 (IL-2) and a danger signal that represents a general activator signal for macrophages. These cells and molecules can interact according to behavioral rules simulating the phenomena of an adaptive immune response against cancer.

The lymph node close to the tumor is modeled explicitly by the lattice that contains all cell types and molecules. In contrast, the thymus selection is simulated implicitly, as all T-cells are probed for their reactivity against MHC and self-peptides in positive and negative selection before being introduced into the simulation.

All specific interactions of the adaptive immune response depend on the complementarity of bit-strings in C-ImmSim. To be more explicit, each immune cell receptor, MHC, antigen or MHC-presented peptide is represented by a string composed of zeros and ones, and whenever an immune receptor meets a possible ligand, the complementarity of the bit-strings defines their interaction probability. Whereas this way of representing the molecular specificity is very generic and computationally easy, it is very distant from real protein-protein interactions. The development of a sequence-based definition of molecules is therefore an important step toward the generation of more realistic models.

## From Bit-Strings to Amino Acids

Rapin et al. have extended C-ImmSim to a sequence-based version for simulating generic infectious diseases [4]. In contrast, our extension, VaccImm, is focused on peptide vaccination in cancer therapy and therefore has to fulfill several different needs.

To turn the bit-string model into a sequence-based model, amino acid sequences need to be introduced for all specific interactions between an immune receptor and a ligand. In addition, in C-ImmSim the interaction strength is based on the complementarity of bit-strings. In VaccImm, the interaction probabilities dependent on amino acid sequence had to be developed. Rapin et al. used the Miyazawa-Jernigan potential [17] for that purpose. In the present work we demonstrate that this approach does not clearly differentiate observed immune receptor-ligand complexes from random complexes. For that reason, we developed new interaction potentials specific to B-cell and T-cell receptors that are able to clearly differentiate experimentally observed from random complexes.

## Results

### Development of an Interaction Potential for Immune Receptor-Ligand Binding

An essential step in the adaptive immune response is the recognition of a target cell by the immune cell mediated by the immune receptor binding to its matching ligand. Both the interaction of a T-cell receptor with a peptide-MHC complex and a B-cell receptor with an antigen are very complex in nature and still a matter of research [18,19]. Therefore, it is desirable to have a universal function differentiating binding from non-binding immune complexes.

In the present study, we develop a separate empirical pair-wise interaction potential (IP) between the immune receptor and its ligand for both the B-cell and the T-cell receptors. The IP is based on an analysis of crystal structures from the Protein Data Bank [20] using a similar method to that used by von Eichborn et al. [21]. We compared the amino acid pairs within the interface of observed crystal structures of immune receptor complexes to those of random structures built to represent non-binding receptor ligand pairs. From this calculation, we gained an IP specific for the type of immune receptor-ligand-complex;  $IP_T$  for T-cell receptors and  $IP_B$  for B-cell receptors (Fig. 1 and Fig. S1).

Before the analysis, structures were separated into training (90%) and validation (10%) sets and then all complexes were scored using  $IP_T$  or  $IP_B$ . The scores of training and validation sets of the crystal and artificial structures were compared to scores that were gained using the well known Miyazawa-Jernigan potential [17], as Rapin et al. [4] used the Miyazawa-Jernigan potential to score the immune receptor-ligand interaction. Our analysis shows that for the peptide/MHC-TCR interaction the Miyazawa-Jernigan potential is not able to differentiate between crystal and artificial structures, while the sets can be clearly differentiated using  $IP_T$  (Fig. 2A). The area under curve (AUC) was calculated with the Mann-Whitney test for the structures scored with  $IP_T$  yielding an AUC of 0.93 and a P-value below 0.01, so the differentiation is highly significant.

For the BCR-antigen interaction, the Miyazawa-Jernigan potential again does not differentiate between the crystal and the artificial structures while  $IP_B$  does (Fig. 2B). As the standard deviation is very large when using  $IP_B$ , we looked for any

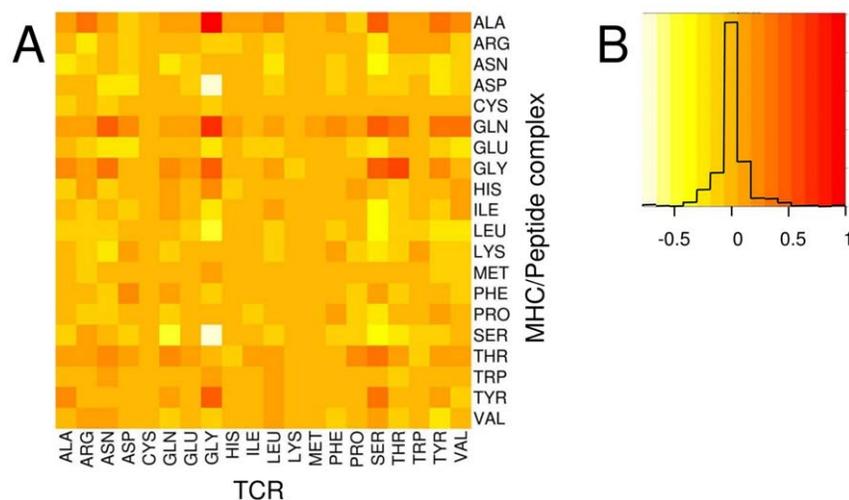
characteristics of the structures to explain this phenomenon. We found that the standard deviation can be considerably decreased if the data sets are subdivided into sets with high and low glycine frequencies within the antigen interface (Fig. 2C+D and Fig. S2,S3). The limit was set at 6.9% glycine within the interface of the antigen, as this is the frequency of glycine on protein surfaces. The separation resulted in two data sets of nearly equal size. The Mann-Whitney test yields an AUC of 0.93 for the high glycine data set ( $IP_{B \text{ high}}$ ) and 0.88 for the low glycine data set ( $IP_{B \text{ low}}$ ), while the P-value is below 0.01 for both. The reason for this difference in high and low glycine content in the interface of an antigen might be that the higher glycine content allows sharper turns in the backbone and that could lead to a slightly different mechanism of recognition by antibodies.

### Application of the Interaction Potential in VaccImm

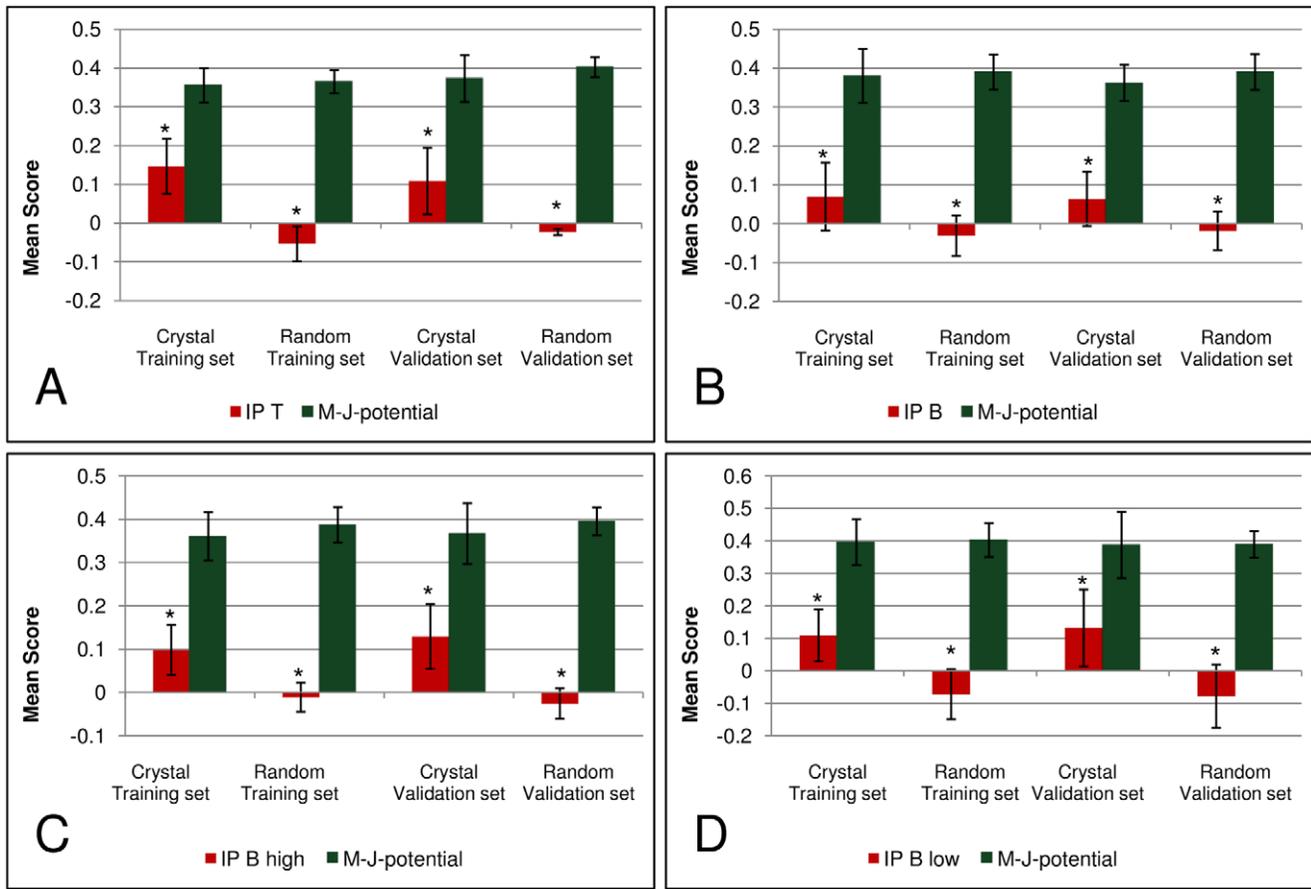
The newly developed IP was applied to the agent-based simulation VaccImm, an extension of C-ImmSim [3], along with approved prediction methods for MHC-peptide binding (see Materials and Methods).

VaccImm accounts for the most important immune cell populations in their different activation states. In this section, we analyze these output-curves of the cell counts over time. Figure 3 depicts example curves of cell counts resulting from a successful (right panel) and a failing (left panel) peptide vaccination therapy. The difference between the two simulations is that the peptides of the successful treatment were emulsified in adjuvant while the peptides of the failed treatment were not. To initiate or boost the immune response, adjuvant is often added for immunization [22]. Cancer epitopes are naturally of low immunogenicity as they result from the body's own proteins that are usually not recognized as foreign.

Without adjuvant, the cancer cells double within one year (Fig. 3A), while a successful treatment is able to eradicate the tumor within less than four months (Fig. 3B). In successful immune therapy, matching T-cell and B-cell receptors recognize the antigen and TH-cells, TH-cells and B-cells get activated and replicate, as seen in Figure 3. In contrast, peptide vaccination without adjuvant is not able to activate TH-cells, TC-cells nor B-cells for proliferation. Also the antigen presenting cells (APCs) behave differently in the two settings. In both cases, DCs present



**Figure 1. Interaction Potential  $IP_T$  for MHC-Peptide/TCR Complexes.** A: The newly developed interaction potential  $IP_T$  for MHC-peptide/TCR complexes. B: Color code and color frequency for interaction potential map. doi:10.1371/journal.pone.0023257.g001



**Figure 2. Evaluation of the Miyazawa Jernigan Potential and the Newly Developed Potentials  $IP_T$ ,  $IP_B$ ,  $IP_{B\ high}$  and  $IP_{B\ low}$ .** Mean scores of MHC/peptide-TCR crystal and random structures for training (90% of structures) and validation set (10% of structures) are depicted. Scores observed using the newly developed interaction potentials and the Miyazawa-Jernigan interaction potential [17] are compared. A: Scores of MHC/peptide-TCR complexes using  $IP_T$ . B: Scores of antibody/antigen complexes (all) using  $IP_B$ . C: Scores of antibody/antigen complexes with antigens of high glycine frequency (>6.9%) within the interface using  $IP_{B\ high}$ . D: antibody/antigen complexes with antigens of low glycine frequency (<6.9%) within the interface using  $IP_{B\ low}$ . \* indicates P-value of crystal and random set below 0.01 using Mann-Whitney test. M-J-potential: Miyazawa-Jernigan potential.

doi:10.1371/journal.pone.0023257.g002

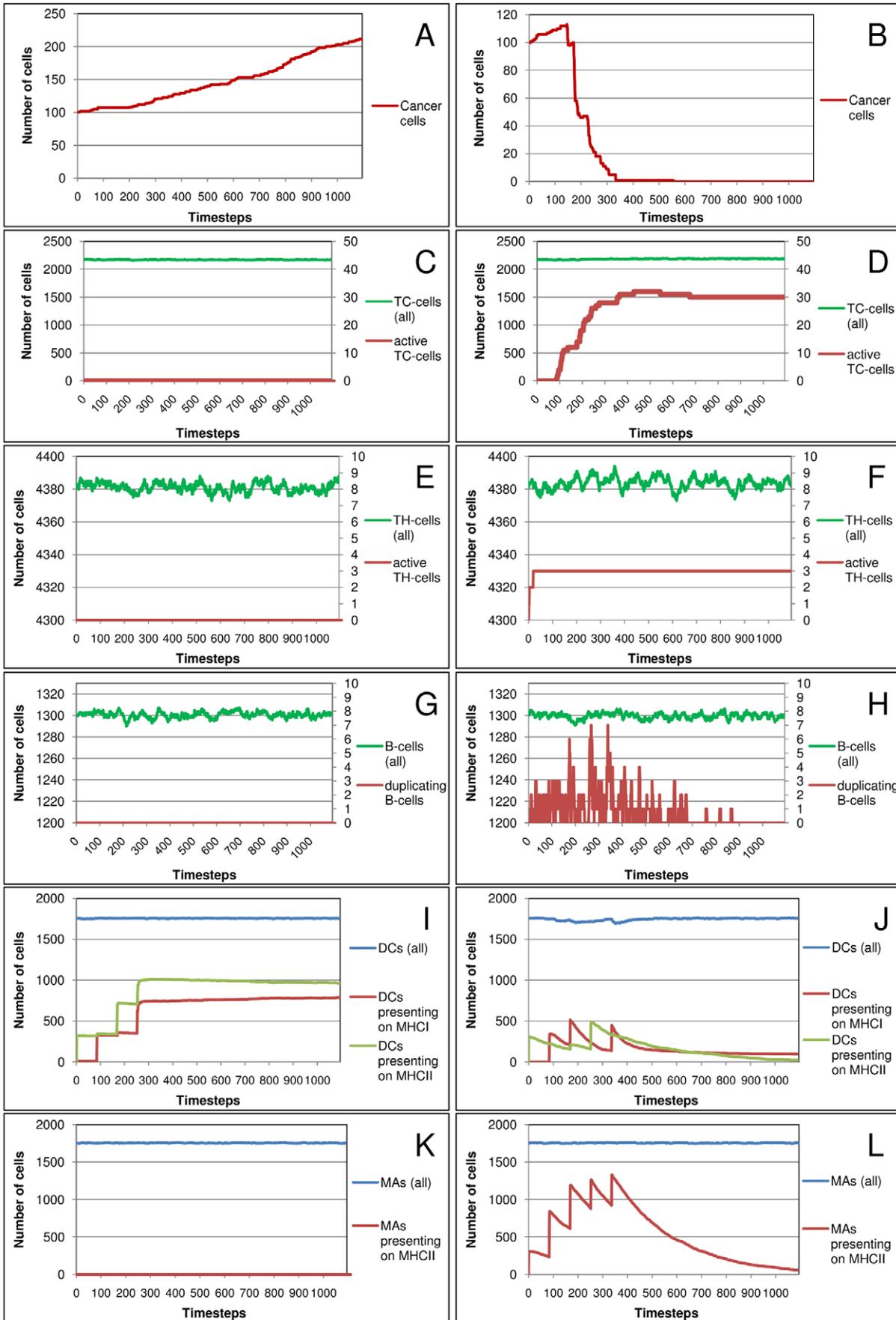
peptides on their MHC I and MHC II, but without a stimulating signal from the adjuvant this leads to anergy rather than to activation of the T-cells. Macrophages are not recruited to present peptides on their MHC II without adjuvant stimulation. In successful treatment, presentation of both APC types on MHC peaks when the antigen is injected but cell counts decrease immediately afterwards because activated TC-cells kill the APC.

### Impact of $IP_T$ on the Simulation with VaccImm

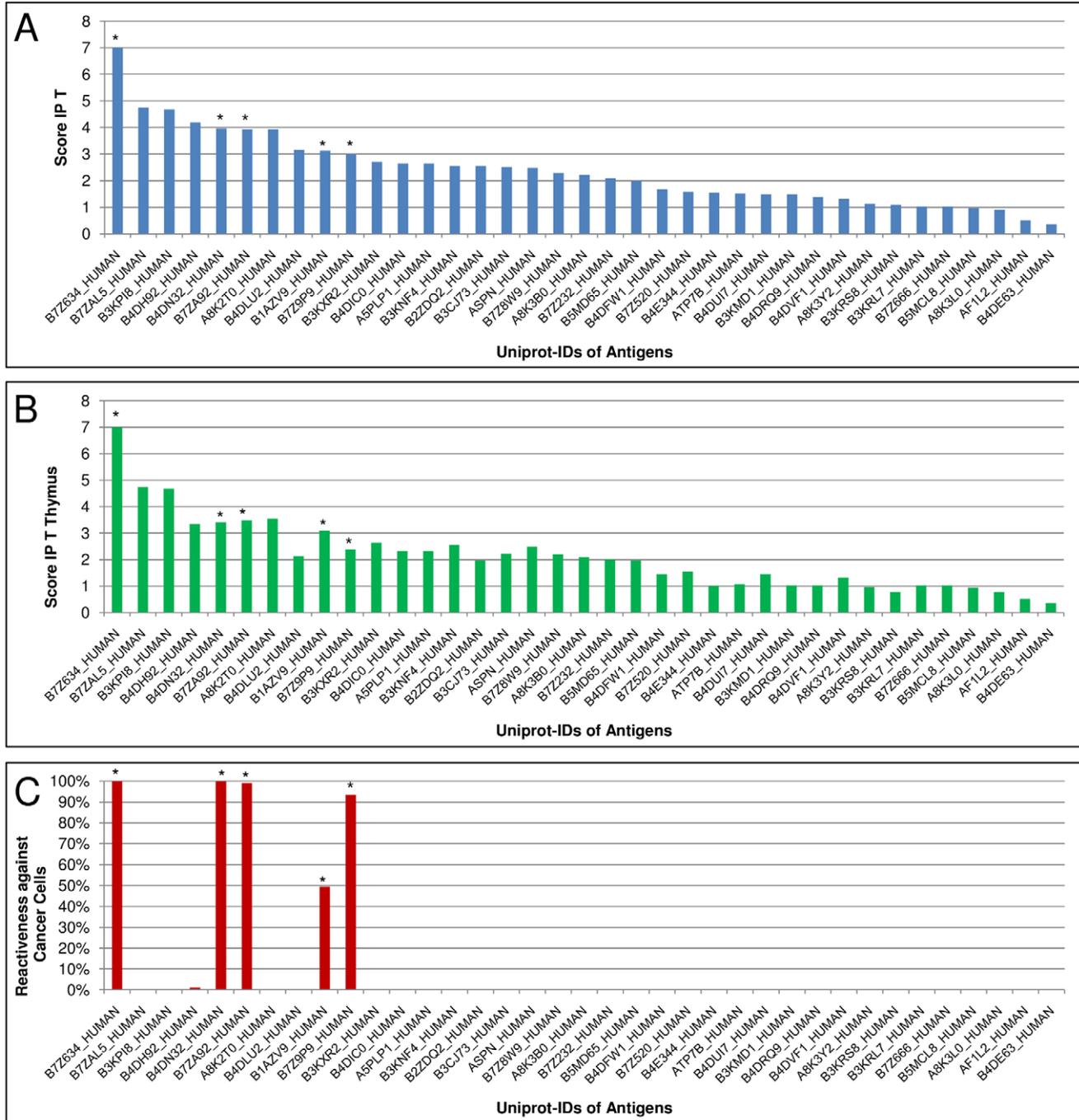
The most important step of the immune response against a tumor is the recognition of a cancer epitope on MHC I by TC-cells that will result in the lysis of tumor cells under activating conditions. However, the bottleneck of cancer immunotherapy is that most of the TC-cells that could possibly recognize the cancer cells are eliminated in the thymus as they are reactive against self-peptides. If a matching TC-cell survives thymus selection, lowering of the tumor burden by the TC-cell clone depends on several other factors e.g. the MHC-binding properties of the cancer epitope, the way of antigen encounter by the TC-cell, its spatial environment and the cytokine interplay. To analyze the impact of  $IP_T$  and thymus selection on the simulation, we compared their calculated reactiveness.

Different antigens were evaluated for their score using  $IP_T$  and the T-cell reactivity they induced in the agent-based model simulation. For that purpose, a set of about 15,000 TCR sequences was randomly generated. In the first step, the antigens were sorted by the highest  $IP_T$ -score they gained when probed against the whole TCR set (Fig. 4A). In the second step, all TCRs that would not survive thymus selection, because of their complementarity to self peptides, were eliminated and the antigens were scored again with the reduced TCR set (Fig. 4B). In the third step, the entire simulation was executed for all the antigens using VaccImm (Fig. 4C). The reactiveness against cancer cells was defined as the reduction of tumor size compared to unlimited growth, with 100% reactiveness leading to complete tumor eradication.

In thymus selection, strongly self-reactive TC-cells are eliminated, that is why several scores are lowered after thymus selection (Fig. 4A+B). Nevertheless, the same overall trend was observed for stronger and weaker immunogenic antigens. In the simulation, only very few antigens are able to induce a sufficient immune response to reduce or eradicate the tumor (Fig. 4C). Antigens having a low score from  $IP_T$  are not recognized by TC-cells because of their low interaction probability. However, some antigens with high interaction probabilities calculated from  $IP_T$  are also not able to induce a



**Figure 3. Peptide Vaccination With and Without Adjuvant.** Cell counts are depicted for cancer cells (A/B), cytotoxic T-cells (C/D), helper T-cells (E/F), B-cells (G/H), dendritic cells (I/J) and macrophages (K/L) over one year of simulated time. One time step equals 8 hours. Left panel: No adjuvant was added; right panel: Adjuvant was included. For both experiments one hundred initial cancer cells were simulated within 5  $\mu$ l of blood that were treated with peptide injections starting at time point zero and being repeated five times at an interval of 28 days. The MHC-genotype of this virtual individual is HLA\_A80/HLA\_B56/HLA\_DQA2\_DQB2/HLA\_DRB3. The antigen A2SUH6\_HUMAN (Survivin variant 3 alpha) was chosen from over-expression data in prostate tumors. Two MHC I-binding peptides and two MHC II-binding peptides were injected.  
doi:10.1371/journal.pone.0023257.g003



**Figure 4. Comparison of  $IP_T$ , Thymus Selection and Simulation.** Scores for different antigens using  $IP_T$  are compared with (A) a random sample of about 15,000 TCRs, (B) all TCRs from the set that survived the thymus selection and (C) with the reactivity against cancer cells in the simulation with Vacclmm. The antigens in all three diagrams are sorted by their  $IP_T$ -score including all TCRs. Antigens inducing an immune response against the tumor, are marked with \*. Simulation data were obtained from 100 simulations per column. The MHC-genotype is HLA\_A02/HLA\_DRB3. For all experiments one hundred initial cancer cells were simulated within 5  $\mu$ l of blood that were treated with peptide injections starting at time point zero and being repeated five times at an interval of 28 days. The antigens were chosen from over-expression data in prostate tumors. One MHC I-binding peptide and one MHC II-binding peptide were injected.  
doi:10.1371/journal.pone.0023257.g004

sufficient immune response. This result underlines that the complete immune cell and signal molecule interplay needs to be investigated for understanding the tumor-immune-interaction.

In the next section, we analyze the failing antigen treatments in more detail.

### Reasons for Failures in Treatment

In order to better understand the reasons why the immune system failed to successfully fight the tumor for most of the antigen injections, we took a closer look at the immune cell populations (Fig. 5). A successful treatment was defined as a simulation with a smaller tumor size after one year of simulated time with respect to the beginning of treatment.

Interestingly, we found that the resulting patterns can be separated into two groups: In the first group, a sufficient immune response was induced in less than 50% of the simulations (frequent failures, Fig. 5a) and in the second group in more than 50% (rare failures, Fig. 5b). In both groups, peptide presentation on MHC I and II by DCs and macrophages was successful in 75% to 100% of the simulations. Consequently, this part of the immune response was not the main bottleneck in the treatment. In the frequent failures, no active TC-cells and almost no active TH-cells were observed. For the rare failures, active TC-cells were observed in more than 80% of the failing simulations. We did take a closer look at these TC-cell populations, searching for a reason why they were not able to eradicate the tumor. We found that these TC-cell populations were either very small (composed of only one or two cells) or started dividing at a later stage of the simulation such that the tumor size could not be decreased before the end of the simulation (data not shown).

These observations indicate that the rare failures would occur with an even lower frequency if the simulations were carried out over a longer time period. In contrast, the frequent failures are prone to failing to induce an immune response either because their antigens are too closely related to self-peptides or because the TC-cells did not encounter the antigen in a sufficiently activating environment.

### Simulation Run Time and Error Propagation from the Interaction Potential

VaccImm is able to simulate up to one million initial cancer cells treated with a maximum of ten peptides. One run of the simulation takes from several seconds to a few minutes on a single

core of a 3 GHz Intel Core 2 Duo CPU depending on the input parameters chosen.

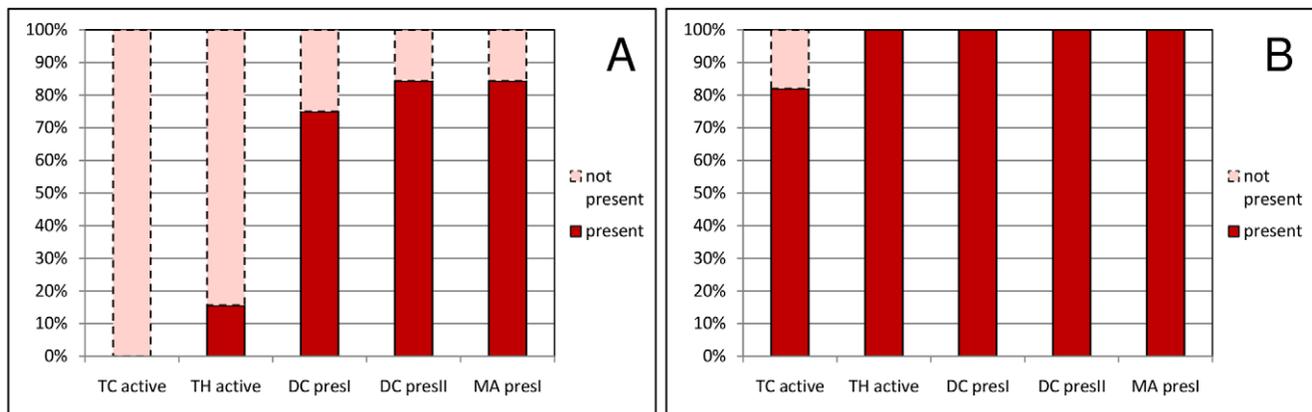
Before starting the simulation, peptide sequences binding to MHC I and MHC II are predicted within a few seconds. The VaccImm simulation itself can be divided into two parts, thymus selection of initial population of T lymphocytes and spatial dynamics within the agent-based model accounting for the whole immune response. Having a simulation with standard configuration (see Table 1), the two parts require about the same amount of CPU time. In general, however, while the time for thymus selection increases with the number of self-peptides and the number of different MHC-alleles, the time for the agent-based simulation increases with cell numbers, simulated time and volume. Some benchmark experiments with their simulation times are given in Table 1.

To investigate the error propagation from the IP to the simulation, an example was chosen where the tumor is almost eradicated. The interaction probability calculated using  $IP_T$ ,  $IP_B$  high and  $IP_B$  low was increased or decreased in a stepwise manner and the impact on the final number of cancer cells and the time it takes to reduce the cancer cell number compared to the initial tumor size was analyzed (Fig. 6). Increasing the interaction probability led to a gradual decrease in the time to reduce the tumor size while the tumor was almost or completely eradicated with an increase in the interaction probability of 10% and more. Decreasing the interaction probability increased the time to reduce the tumor and the relative number of cancer cells at the end of the simulation. At a 15% decrease of the interaction probability, the mean number of cancer cells was not reduced compared to the beginning of the simulation. Decreasing the interaction probability by 20% or more leads to tumor reduction in less than 50% of the simulations.

From these results we can deduce that a bias in  $IP_T$ ,  $IP_B$  high and  $IP_B$  low will affect the results of the simulation. A qualitative difference in the outcome, as having or not having an induction of the immune response, will probably occur from a bias larger than 10–15%.

### Discussion

In this work, we developed new IPs between an immune receptor and its ligand from a statistical analysis of crystal



**Figure 5. Reasons for Failures in Treatment.** The propensity to obtain certain activation states of the immune cell populations is depicted. Experiments are grouped into frequent failures (A) that induce a sufficient immune response in less than 50% of the simulations and rare failures (B) with more than 50% sufficient responses. Abbreviations: TC active: active cytotoxic T-cells. TH active: active helper T-cells. DC presl/DC presII: dendritic cells presenting peptides on MHC I/MHC II. MA presl: macrophages presenting peptides on MHC II. doi:10.1371/journal.pone.0023257.g005

**Table 1.** Computational Cost of the Simulation.

Configuration	Computational Time [s]
Standard*	21
Initial tumor size [cells]: 10 000	33
Initial tumor size [cells]: 100 000	129
Initial tumor size [cells]: 1 000 000	855
Simulated Space [ $\mu$ l]: 10	82
Simulated Space [ $\mu$ l]: 20	165
Simulated Space [ $\mu$ l]: 20, Initial tumor size [cells]: 100 000	250
2 MHC-I binding peptides, 2 MHC-II binding peptides	40
3 MHC-I binding peptides, 3 MHC-II binding peptides	73
4 MHC-I binding peptides, 4 MHC-II binding peptides	100
5 MHC-I binding peptides, 5 MHC-II binding peptides	117
2 MHC-I alleles, 2 MHC-II alleles	31
3 MHC-I alleles, 3 MHC-II alleles	57
6 MHC-I alleles, 6 MHC-II alleles	45
2190 time steps	25
4380 time steps	34

\*The standard configuration for the simulations is: Uniprot-ID 2B1D\_HUMAN, 1 MHC-I allele, 1 MHC-II allele, 1 MHC-I binding peptide, 1 MHC-II binding peptide, 5  $\mu$ l simulated space, 1095 time steps, 100 cancer cells as initial tumor size. All changes to this configuration are mentioned in the table.  
doi:10.1371/journal.pone.0023257.t001

structures for B-cell and T-cell receptors. It was shown that the newly developed IPs are able to differentiate naturally observed from randomly generated immune complexes, in contrast to the Miyazawa-Jernigan potential [17]. We believe that our empirically

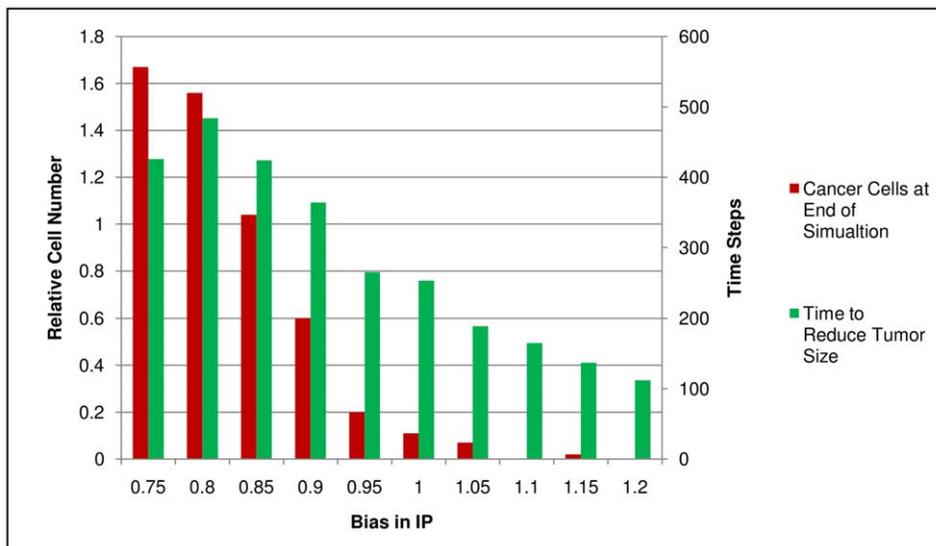
developed potentials account for specific properties of the respective immune receptor interactions, that are not included in general potentials like the Miyazawa-Jernigan potential.

In the second step, the newly developed IPs were applied to the agent-based model VaccImm, an extension of C-ImmSim [3]. There is one other agent-based model that includes the amino acid sequence in calculation; the model of Rapin et al. [4] simulates the immune response against infections and was developed independently from VaccImm. While Rapin et al. took a more general approach to infectious diseases; we focused on the specific needs of peptide vaccination in cancer therapy.

Prediction methods for complex protein-protein interactions, such as our newly developed IPs, always have a limited accuracy. Therefore, we investigated what impact a systematic error in the IPs would have on the VaccImm simulation. As demonstrated, a bias in the IPs will surely affect the simulation, but qualitative changes in the outcome are not expected as long as the bias in the interaction prediction stays below 10%.

In contrast to the former versions of C-ImmSim, and similarly to the new published version [4], VaccImm predicts reactivity of the immune system against real amino acid sequences of any injected peptide. The model includes several core phenomena of the immune system, e.g., the humoral and cellular branch, clonal expansion of single immune cells, thymus selection of T-cells and spatial interactions of autonomous cells with distinct specificity. Still, some properties of the immune-cancer interplay might be underrepresented in the current version of the model and we are planning to extend VaccImm in that direction. The most important features we are planning to include in the future are the different types of cytokines, the influence of regulatory T-cells ( $T_{reg}$ s) and the mutation of cancer cell epitopes.

Cytokines have an essential role in facilitating communication between immune cells, most of them acting within short ranges [23]. Thus far, only IL-2 and a general danger signal are included in the simulation, but agent-based models are particularly tailored



**Figure 6. Error Propagation from IP to Simulation.** The relative number of cancer cells after one year of treatment compared to the tumor size at the beginning of the simulation is depicted (red bars) along with the average number of time steps needed to reduce the tumor size (green bars). All interaction probabilities in the simulation calculated with IP were multiplied with the bias in IP shown on the y-axis. Data were obtained from 100 simulations per column. The MHC-genotype is HLA\_A02/HLA\_DRB1-1. For all experiments one hundred initial cancer cells were simulated within 5  $\mu$ l of blood that were treated with peptide injections starting at time point zero and being repeated five times at an interval of 28 days. The antigen B7Z8B0\_HUMAN was chosen from over-expression data in prostate tumors. One MHC I-binding peptide and one MHC II-binding peptide were injected.

doi:10.1371/journal.pone.0023257.g006

to take into account local effects of molecules in a spatial environment. Consequently, we are planning to include a more detailed representation of cytokines in the model in the future. Examples for important cytokines in the cancer-immune interplay to be included soon are the transforming growth factor- $\beta$  [24] or interferon- $\gamma$  [25].

$T_{reg}$ s are an immune cell population suppressing the immune response, presumably with the main function to prevent autoimmune diseases [26]. Their immunosuppressive effect hinders the immune response against cancer [27]. Therefore, our plan is to include this cell type in the next version of VaccImm.

Our current model does not account for mutation or changes in gene expression of the tumor. The genome of a tumor is often unstable and many different mutations or gene rearrangements result in a huge diversity of tumor cells that do not present the same epitopes on their MHCs. It is frequently observed, particularly under treatment, that some tumor cells change their behavior or their expression pattern, thereby circumventing eradication; a phenomenon called tumor escape [28]. For this purpose, we plan to include tumor epitope mutation and changes in the antigen expression levels.

The present model is a first step towards *in silico* experiments predicting T-cell reactivity taking into account the amino acid sequence. Expanding knowledge in tumor immunology will help to further improve our model. As VaccImm is very flexible and its architecture is separated into different modules, it can be easily extended or refined with new experimental data.

## Materials and Methods

### Peptide Binding to MHC Complexes

The first step necessary for T-cell recognition is the processing of a protein to a peptide and its presentation on a MHC complex by the target cell. This process has been evaluated in great detail by several research groups. It has been found that key residues exist at distinct places within the peptide that are most important for binding to the MHC [29]. From the accumulated MHC binding and elution data, several methods have been generated to predict which part of the protein sequence will be presented by a certain MHC. For simulation, we used two well-known position-specific scoring matrices to predict peptide-MHC binding, namely *smmpmbec* (<http://tools.immuneepitope.org>) for MHC I binding and *arb* [30] for MHC II binding.

### Development of an Interaction Potential for Immune Receptor-Ligand Binding

A separate empirical pairwise interaction potential between the immune receptor and its ligand was developed for both the B- and T-cell receptor, based on crystal structures from the Protein Data Bank [20] using a similar method to that used by von Eichborn et al. [21]. A statistical analysis was drawn from 237 antibody-antigen and 33 MHC-TCR non-redundant crystal structures. Interacting amino acid pairs between receptor and ligand were defined as all residues having no more than 8 Å distance between their C $\alpha$  atoms. We compared the observed crystal structures to random structures built to represent non-binding receptor ligand pairs. To generate the random structures, the peptide sequences of the MHC-peptide complex or the antigen were replaced by random sequences with the same amino acid distribution that was observed on general protein surfaces. The number of interacting pairs was counted for both the crystal and the random structures and the numbers were subtracted using the formula:

$$M(x, y) = C_{crystal}(x, y) - C_{random}(x, y) \quad \forall x, y \in AA$$

where  $M$  is the interaction potential scoring matrix and  $C_{crystal}$  and  $C_{random}$  are the counts for crystal and random structures, respectively. From this calculation, we gained an interaction potential (IP) specific for the type of immune receptor-ligand-complex;  $IP_T$  for T-cell receptors and  $IP_B$  for B-cell receptors. Within  $IP_T$  and  $IP_B$ , positive values represent an increased likelihood of observing the corresponding amino acid pair in the interface of the complex, while negative values represent a decreased likelihood.  $IP_T$  and  $IP_B$  were normalized to a maximum value of 1.0 and then used to score the crystal and artificial structures for validation.

### Application of the Interaction Potential to the Simulation

Whenever an active T-cell meets an APC presenting a peptide on a matching MHC type, their interaction probability is calculated using  $IP_T$ . To define which residues of the immune receptor and the ligand are in contact with each other, we created contact matrices from crystal structures analogously to the work of Rapin [4]. Again, the interacting amino acid pairs between receptor and ligand were defined as all residues having at most 8 Å distance between their C $\alpha$  atoms. For each crystal structure, a matrix composed of ones and zeros representing interacting and non-interacting residues was created. In a next step the interaction probability was calculated as the sum of the  $IP_T$  matrix values for all interacting residues that is normalized by the number of interactions.

As the number of MHCs with known structure is limited, MHC sequences not having been crystallized were mapped to the closest related MHC structures for the definition of the contact residues. Contact matrices were built from MHC-peptide-TCR complexes containing HLA\_A02 (pdb-ID: 1OGA), HLA\_B08 (1MI5), HLA\_B35 (3MV7), HLA\_B44 (3DXA), and HLA\_DRB1-4 (1J8H). For the interaction between a BCR or an antibody and its antigen, we used one high-resolution structure of an antibody binding to a peptide (2B1H) to create the contact matrix, as the injected antigens are peptides and their tertiary structure can probably be neglected.

### Steps of the Simulation

Starting the simulation, the different behaviors and movements of the cells are executed in discrete time steps. For the purpose of modeling peptide vaccination to treat cancer with VaccImm, the timeline of observed phenomena is as follows:

1. The peptides are injected emulsified in adjuvant. The adjuvant is modeled as a danger signal activating macrophages.
2. APCs take up the peptides by phagocytosis. In the case of macrophages and DCs, phagocytosis is unspecific, while B-cells have to recognize the antigen for ingestion. The probability of recognition is calculated using the B-cell contact matrix and  $IP_B$  high or  $IP_B$  low depending on the antigen.
3. Peptides are processed and presented by the APCs. The probability of presenting a certain peptide by a specific MHC is calculated using the position-specific scoring matrices described above.
4. The peptides presented on MHC I and MHC II can be recognized by TC- and TH-cells, respectively. The probability for recognition is calculated using  $IP_T$  together with the contact matrix for the respective MHC. Reacting T-cells will be activated to duplicate and create memory cells.

5. The humoral and the cytotoxic immune response begin. Activated TC-cells might recognize the cancer cells, while the probability of interaction is calculated again via  $IP_T$  and the respective contact matrix. Successful interactions between TC-cells and cancer cells will kill the cancer cells and therefore the tumor will be eliminated. B-cells stimulated by TH-cells start duplicating into memory cells and antibody-producing plasma cells. Plasma cells secrete antibodies that clear the antigen.

This model is quite generic and flexible and captures the main mechanisms in the tumor-immune interaction. As an output of the model, cell lines and their respective activation states can be followed over time.

### Input Data

For VaccImm, all sequence-based input parameters needed to be changed with respect to C-ImmSim. Within the present analysis, we focus on urological tumors of kidney and prostate, but the program is very general and able to simulate any solid tumor. As input, the model needs the amino acid sequences of the cancerous proteins, the MHC alleles, and the self-peptides used in thymus selection.

Cancerous proteins were selected based on expression data from patients suffering from prostate or kidney tumors. [31]. Any gene that is over-expressed at least 2 times in the analyzed cancerous tissue with respect to healthy control tissue was chosen as a cancer target. The respective protein sequences were collected from the UniProt. [32]. Peptides presented on MHC I or MHC II originating from the protein sequences were predicted using prediction algorithms (*consensus* [33] for MHC I, *smm\_align* [34] for MHC II) from the Immune Epitope Database [35]. These predicted peptides are presented by the MHCs in the simulation. The same peptide sequences are used for injection in immune therapy within the simulation. As short peptides usually do not fold into complex secondary structures, the injected peptides are used directly as epitopes for recognition by B-cells and antibodies.

One part of the simulation is the selection of T-cells within the thymus. If any T-cell receptor is strongly reactive to self-peptides, the T-cell will be eliminated. To define the self-peptides, we downloaded all peptide sequences eluted or known to bind to MHC I or MHC II from the IEDB [35]. To decrease the number of self-peptides and therefore increase the speed of the simulation, we used only the 50 self-peptide sequences most similar to each

injected peptide. All other self-peptide sequences having a lower similarity will not interfere with the injected peptides and thus are neglected.

The MHC sequences for all different alleles were downloaded from the UniProt [32], 61 for MHC I and 20 for MHC II. All TCR and BCR sequences were generated at random when the respective cell was introduced into the simulation.

Any other parameter not pertaining to the amino acid sequence implementation are left unchanged with respect to C-ImmSim [3].

### Supporting Information

**Figure S1 Interaction Potential  $IP_B$  for All Antibody/Antigen Complexes.** A: The newly developed interaction potential  $IP_B$  for all antibody/antigen complexes. B: Color code and color frequency for interaction potential map. (TIFF)

**Figure S2 Interaction Potential  $IP_B$  high for Antibody/Antigen Complexes, High Glycine Frequency.** A: The newly developed interaction potential  $IP_B$  high for antibody/antigen complexes with antigens of high glycine frequency (>6.9%) within the interface. B: Color code and color frequency for interaction potential map. (TIFF)

**Figure S3 Interaction Potential  $IP_B$  low for Antibody/Antigen Complexes, Low Glycine Frequency.** A: The newly developed interaction potential  $IP_B$  low for antibody/antigen complexes with antigens of low glycine frequency (<6.9%) within the interface. B: Color code and color frequency for interaction potential map. (TIFF)

### Acknowledgments

We would like to thank Hellmuth-Alexander Meyer for analyzing the expression data of kidney and prostate tumors.

### Author Contributions

Conceived and designed the experiments: AW RP. Performed the experiments: AW JvE. Analyzed the data: AW JvE CW. Contributed reagents/materials/analysis tools: AW MM FC. Wrote the paper: AW JvE CW.

### References

- Palucka K, Ueno H, Banchereau J (2011) Recent developments in cancer vaccines. *J Immunol* 186: 1325–1331.
- Waelke AL, Murgueitio MS, Preissner R (2010) Theoretical modeling techniques and their impact on tumor immunology. *Clin Dev Immunol* 2010: 271794.
- Castiglione F, Toschi F, Bernaschi M, Succi S, Benedetti R, et al. (2005) Computational modeling of the immune response to tumor antigens. *J Theor Biol* 237: 390–400.
- Rapin N, Lund O, Bernaschi M, Castiglione F (2010) Computational immunology meets bioinformatics: the use of prediction tools for molecular binding in the simulation of the immune system. *PLoS One* 5: e9862.
- Celada F, Seiden PE (1992) A computer model of cellular interactions in the immune system. *Immunol Today* 13: 56–62.
- Farmer JD, Packard NH, Perelson AS (1986) The immune system, adaptation, and machine learning. *Physica D: Nonlinear Phenomena* 22: 187–204.
- Folcik VA, An GC, Orosz CG (2007) The Basic Immune Simulator: an agent-based model to study the interactions between innate and adaptive immunity. *Theor Biol Med Model* 4: 39.
- Folcik VA, Broderick G, Mohan S, Block B, Ekbote C, et al. (2010) Using an agent-based model to analyze the dynamic communication network of the immune response. *Theor Biol Med Model* 8: 1.
- Meier-Schellersheim M, Mack G (1999) SIMMUNE, a tool for simulating and analyzing immune system behavior. Available: <http://www.wicetbase.org/abstract?id=oiarXivorgcs/9903017>.
- Kalita JK, Chandrashekar K, Hans R, Selvam P (2006) Computational modelling and simulation of the immune system. *Int J Bioinform Res Appl* 2: 63–88.
- Mallet DG, De Pillis LG (2006) A cellular automata model of tumor-immune system interactions. *J Theor Biol* 239: 334–350.
- Castiglione F, Duca K, Jarrah A, Laubenbacher R, Hochberg D, et al. (2007) Simulating Epstein-Barr virus infection with C-ImmSim. *Bioinformatics* 23: 1371–1377.
- Santoni D, Pedicini M, Castiglione F (2008) Implementation of a regulatory gene network to simulate the TH1/2 differentiation in an agent-based model of hypersensitivity reactions. *Bioinformatics* 24: 1374–1380.
- Baldazzi V, Castiglione F, Bernaschi M (2006) An enhanced agent based model of the immune system response. *Cell Immunol* 244: 77–79.
- Lollini PL, Motta S, Pappalardo F (2006) Discovery of cancer vaccination protocols with a genetic algorithm driving an agent based simulator. *BMC Bioinformatics* 7: 352.
- Pennisi M, Pappalardo F, Palladini A, Nicoletti G, Nanni P, et al. (2010) Modeling the competition between lung metastases and the immune system using agents. *BMC Bioinformatics* 11 Suppl 7: S13.
- Miyazawa S, Jernigan RL (1999) An empirical energy potential with a reference state for protein fold and sequence recognition. *Proteins* 36: 357–369.
- van der Merwe PA, Dushek O (2011) Mechanisms for T cell receptor triggering. *Nat Rev Immunol* 11: 47–55.

19. Tolar P, Sohn HW, Liu W, Pierce SK (2009) The molecular assembly and organization of signaling active B-cell receptor oligomers. *Immunol Rev* 232: 34–41.
20. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, et al. (2000) The Protein Data Bank. *Nucleic Acids Res* 28: 235–242.
21. von Eichborn J, Gunther S, Preissner R (2010) Structural features and evolution of protein-protein interactions. *Genome Inform* 22: 1–10.
22. Pichichero ME (2008) Improving vaccine delivery using novel adjuvant systems. *Hum Vaccin* 4: 262–270.
23. Siebert JC, Walker EB (2010) Monitoring cytokine profiles during immunotherapy. *Immunotherapy* 2: 799–816.
24. Meulmeester E, Ten Dijke P (2011) The dynamic roles of TGF-beta in cancer. *J Pathol* 223: 205–218.
25. Saha B, Jyothi Prasanna S, Chandrasekar B, Nandi D (2010) Gene modulation and immunoregulatory roles of interferon gamma. *Cytokine* 50: 1–14.
26. Sakaguchi S (2011) Regulatory T cells: history and perspective. *Methods Mol Biol* 707: 3–17.
27. Klebanoff CA, Acquavella N, Yu Z, Restifo NP (2011) Therapeutic cancer vaccines: are we there yet? *Immunol Rev* 239: 27–44.
28. Miller AM, Pisa P (2007) Tumor escape mechanisms in prostate cancer. *Cancer Immunol Immunother* 56: 81–87.
29. Appella E, Padlan EA, Hunt DF (1995) Analysis of the structure of naturally processed peptides bound by class I and class II major histocompatibility complex molecules. *Exs* 73: 105–119.
30. Bui HH, Sidney J, Peters B, Sathiamurthy M, Sinichi A, et al. (2005) Automated generation and evaluation of specific MHC binding predictive tools: ARB matrix applications. *Immunogenetics* 57: 304–314.
31. Jones J, Otu H, Spentzos D, Kolia S, Inan M, et al. (2005) Gene signatures of progression and metastasis in renal cell cancer. *Clin Cancer Res* 11: 5730–5739.
32. Leinonen R, Diez FG, Binns D, Fleischmann W, Lopez R, et al. (2004) UniProt archive. *Bioinformatics* 20: 3236–3237.
33. Moutafisi M, Peters B, Pasquetto V, Tschärke DC, Sidney J, et al. (2006) A consensus epitope prediction approach identifies the breadth of murine T(CD8+)-cell responses to vaccinia virus. *Nat Biotechnol* 24: 817–819.
34. Nielsen M, Lundegaard C, Lund O (2007) Prediction of MHC class II binding affinity using SMM-align, a novel stabilization matrix alignment method. *BMC Bioinformatics* 8: 238.
35. Peters B, Sidney J, Bourne P, Bui HH, Buus S, et al. (2005) The immune epitope database and analysis resource: from vision to blueprint. *PLoS Biol* 3: e91.