

Suche und Lernen im Immunsystem: Modelle der T-Zell-Zirkulation und der Negativauslese

Johannes Textor

Theoretical Biology & Bioinformatics
Universität Utrecht
Padualaan 8
3584 CH Utrecht, Niederlande
johannes.textor@gmx.de

Abstract: Das Immunsystem ist ein hochkomplexes verteiltes System, das größtenteils aus multifunktionalen und stets mobilen Zellen besteht. Viele wichtige immunologische Vorgänge sind bis heute nur unzureichend verstanden. In den letzten Jahren wurden durch die Entwicklung neuer Technologien wie der Multiphotonenmikroskopie neuartige und reichhaltige Daten gewonnen. Um aus diesen Daten fundierte Erkenntnisse zu gewinnen, werden in der modernen Immunologie zunehmend Computersimulationen und mathematische Modelle eingesetzt. Diese Arbeit befasst sich zum Einen mit der Überwachung des Organismus durch zirkulierende T-Zellen, die als stochastischer Suchprozess modelliert wird, und zum Anderen mit der Generierung von T-Zell-Rezeptoren durch die sogenannte Negativauslese, die als stochastischer Lernprozess modelliert wird. Beide Modelle werden anhand formaler Methoden aus der theoretischen Informatik untersucht, um ein tieferes Verständnis informationsverarbeitender und stochastischer Aspekte der modellierten Prozesse zu erlangen. Zudem werden aus den Modellen quantitative Vorhersagen generiert, die anhand experimenteller Daten validiert werden. Der Erkenntnisgewinn dieser formalen und quantitativen Analysen wird im Kontext aktueller immunologischer Forschung diskutiert und bewertet. Unter Anderem gelingt es, durch das stochastische Zirkulationsmodell eine Vielzahl bisher isolierter experimenteller Daten quantitativ in Beziehung zu setzen, während das algorithmische Modell der Negativauslese die aktuell besten Vorhersagen zur Erkennung von HIV-Peptiden durch sogenannte CD8-T-Zellen liefert.

1 Einleitung

Die Arbeit besteht aus zwei weitgehend unabhängigen Teilen, die allerdings einer gemeinsamen Struktur und einem gemeinsamen wissenschaftlichen Ansatz folgen, nämlich dem der mathematischen Biologie. Ähnlich der theoretischen Physik wird hierbei zunächst ein gutes *Problemverständnis* angestrebt, um dann das zu untersuchende Problem geeignet als mathematisches Modell zu *formalisieren*. Das erhaltene Modell wird dann im Hinblick auf analytische sowie quantitative *Lösungen* untersucht, die dann, und das ist der wichtigste Schritt, in biologisch sinnvolle *Erkenntnisse und Vorhersagen* übersetzt werden müssen. Die Vorhersagen dienen sowohl der Falsifizierbarkeit des Modells als auch der besseren Planung zukünftiger Experimente. Als Besonderheit verwenden wir in der

Dissertation Techniken aus der Wahrscheinlichkeitstheorie, der Lerntheorie und der Algorithmik, wohingegen die “traditionelle” mathematische Biologie vor allem kontinuierliche Modellierungstechniken wie Differentialgleichungen einsetzt. Die Resultate der Arbeit demonstrieren, dass solche Techniken tatsächlich zum Erkenntnisgewinn für die moderne Immunologie von Nutzen sein können.

2 T-Zell-Zirkulation

Die T-Zellen sind neben den B-Zellen eine der zwei Zellarten des Immunsystems, die dem Körper völlig unbekannte Pathogene erkennen können. Jede T-Zelle ist auf die Erkennung einiger weniger körperfremder Muster spezialisiert (*Spezifität*), und kaum eine T-Zelle gleicht der Anderen (*Diversität*). Durch die hohe Anzahl von T-Zellen (ca. 10^{11} beim Menschen, 10^8 bei einer Maus) wird so insgesamt ein umfassender Schutz erreicht. Dies bedeutet aber, dass die Anzahl von T-Zellen, die auf eine bislang unbekannte Infektion reagieren können, extrem klein ist – in einer Maus sind dies unter Umständen nur 20 Zellen [MCP⁺07]! Diese würden natürlich nicht ausreichen, um eine Infektion wirksam zu bekämpfen. Erkennt eine T-Zelle ein Antigen in einem Lymphknoten oder der Milz, beginnt sie sich deshalb zu teilen, wodurch sich die Population spezifischer Zellen innerhalb weniger Tage millionenfach vergrößert. Je früher ein eingedrungenes Antigen durch die wenigen spezifischen T-Zellen gefunden und erkannt wird, desto schneller kann diese Vermehrung beginnen. Daher befinden sich T-Zellen ständig auf Patrouille durch Lymphknoten, Milz und andere lymphatische Organe (Abbildung 1).

Die Wege der T-Zell-Zirkulation werden seit den 1960er Jahren erforscht. Bereits früh wurde bekannt, dass T-Zellen etwa einmal pro Tag einen Lymphknoten besuchen, wo sie mehrere Stunden verweilen. Über die Wege der Zellen *innerhalb* des Lymphgewebes konnte man allerdings lange Zeit nur spekulieren. Man nahm an, dass lymphatische Organe auf eine gerichtete, synchronisierte Art durchwandert werden. Im Jahr 2002 gelang es dann mittels der Zweiphotonenmikroskopie erstmals, einzelne T-Zellen beim Durchwandern von Lymphknoten im lebenden Tier zu beobachten [MWPC02]. Das Resultat war für viele Immunologen eine große Überraschung: Die Migration der T-Zellen im Gewebe gleicht einer zufälligen Bewegung, einem *random walk*. Die Pioniere der Zweiphotonenmikroskopie beschrieben daher die T-Zell-Zirkulation erstmals als einen stochastischen Suchprozess [WPMC03], was einen Paradigmenwechsel in der Immunologie auslöste.

In der Arbeit verfolgten wir den naheliegenden Ansatz, ein diskretes mathematisches Modell der T-Zell-Zirkulation aufzustellen und dieses Modell mit Techniken zu untersuchen, die man z.B. auch bei der Analyse randomisierter Suchheuristiken anwendet. Durch die kleine Zahl der spezifischen T-Zellen ist ein diskreter Ansatz sinnvoller als ein kontinuierliches Differentialgleichungsmodell. Beispielsweise kann in einer Maus die Zahl der spezifischen T-Zellen kleiner sein als die Zahl der Lymphknoten (etwa 35), so dass viele Lymphknoten zum Zeitpunkt einer Infektion keine spezifischen T-Zellen enthalten. Hieraus sieht man unmittelbar die Notwendigkeit der Zirkulation. Bei kontinuierlichen Modellen dagegen wird oft davon ausgegangen, dass die Population der spezifischen T-Zellen gleichmäßig auf alle Organe verteilt ist [SPN97], so dass der beschriebene Fall in

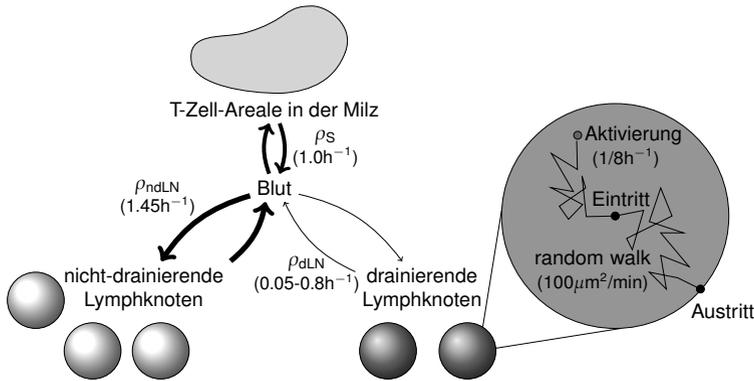


Abbildung 1: Zirkulation von T-Zellen zwischen Blut, Lymphknoten und Milz. Bei systemischen Infektionen wie *listeria*, die sich über das Blut verbreiten, wird die Immunantwort vorwiegend in der Milz gebildet, während bei lokalen Infektionen wie *influenza* oder *herpes simplex* die lokal drainierenden Lymphknoten – dies sind nur etwa 5% aller Lymphknoten – die wichtigste Rolle bei der Abwehr spielen. Vom Blut aus wandern T-Zellen nach dem Zufallsprinzip in die Milz oder einen Lymphknoten ein, wo sie mehrere Stunden nach Antigenen suchen und dann wieder in das Blut zurückkehren. Innerhalb eines Lymphknotens bewegen sich T-Zellen im Wesentlichen zufällig [WPMC03] und werden aktiviert, wenn sie dabei auf Antigenen stoßen.

solchen Modellen nicht berücksichtigt werden kann.

Der grundlegende Aufbau des vorgeschlagenen Modells wird in Abbildung 1 gezeigt. T-Zellen migrieren zwischen dem Blut, der Milz, sowie (bei lokalen Infektionen) drainierenden und nicht-drainierenden Lymphknoten hin und her. Die Verweilzeit der Zellen im Blut wird durch eine Exponentialverteilung modelliert, während für die lymphatischen Organe sowohl konstante Verweilzeiten als auch die sich durch einen *random walk* ergebende Verweilzeitverteilung untersucht wurden. Die Parameter des Modells werden anhand der Literatur festgelegt. Zur grundlegenden Validierung werden quantitative Vorhersagen generiert und mit experimentellen Daten abzugleichen. Einige Resultate dieses Vergleichs werden in Abbildung 2 (A,B) gezeigt. Das Modell nähert experimentell ermittelte Kinetiken der T-Zell-Rekrutierung für verschiedene Infektionstypen (*listeria*, *influenza*, und *herpes simplex*) gut an. Aus den Ergebnissen kann geschlossen werden, dass bei lokalen Infektionen mit wenigen drainierenden Lymphknoten wie *herpes simplex* ohne eine erhöhte Zufuhr von Lymphozyten zu drainierenden Lymphknoten eine effektive Immunabwehr nicht möglich wäre (Abbildung 2B).

Nach der quantitativen Validierung des Modells erfolgt eine tiefer gehende formale Analyse der grundlegenden Modelleigenschaften. Aus der Perspektive stochastischer Suche interessiert uns dabei besonders folgendes Optimierungsproblem: Wie lange sollte eine T-Zelle idealerweise in jedem Lymphknoten bleiben? Wird jeder Lymphknoten nur kurz durchsucht, besteht das Risiko, ein vorhandenes Antigen nicht zu finden. Wird umgekehrt zu viel Zeit auf die Durchsuchung verwendet, so könnte sich unterdessen anderswo eine Infektion ausbreiten, die dann mitunter erst zu spät erkannt wird. Dieses Problem ist

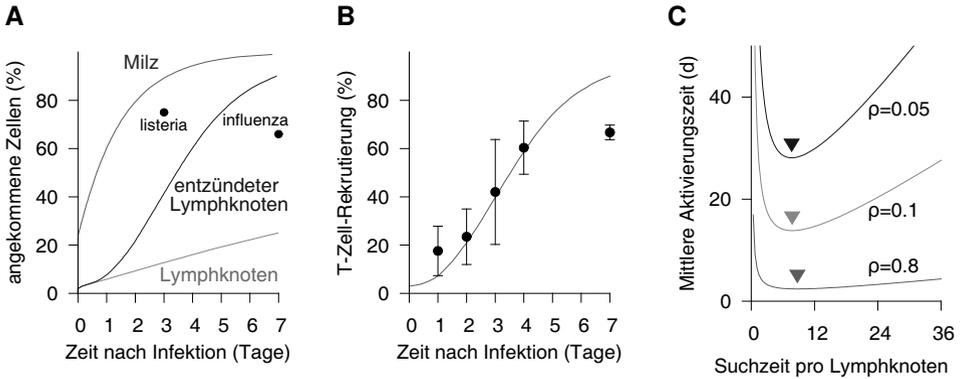


Abbildung 2: Quantitative Vorhersagen des Zirkulationsmodells im Vergleich mit experimentellen Daten in Mäusen. (A) Ankomstgeschwindigkeit der zirkulierenden T-Zellen in der Milz (rote Linie), in einem Lymphknoten (grüne Linie) und in einem entzündeten, anschwellenden Lymphknoten (blaue Linie) verglichen mit T-Zell-Rekrutierungsdaten für *listeria* (Abwehr in der Milz) und *influenza* (Abwehr in Lymphknoten) [vHGS+09]. (B) Verschwinden spezifischer T-Zellen aus der Zirkulation durch Aktivierung in den drainierenden Lymphknoten für *herpes simplex* (Punkte und Fehlerbalken: experimentelle Daten [SJHC11]; Linie: Modellvorhersage). (C) Die erwartete Zeit der Aktivierung einer T-Zelle als Funktion der Suchzeit pro Lymphknoten (Linien) und die sich daraus ergebende optimale Suchzeit (Dreiecke) für verschiedene Einflussraten in drainierende Lymphknoten, wie sie bei einer typischen Entzündung auftreten [SPS+05], und einer Aktivierungsrate von 8 Stunden [MHvA04].

eng verwandt mit der Frage des *optimalen Restarts* bei Las-Vegas-Algorithmen mit bekannter Laufzeitverteilung. Für diese wurde bereits gezeigt, dass keine Strategie besser sein kann als das Neustarten des Algorithmus in festen Zeitintervallen [LSZ93]. Da auch das Erkennen von Antigenen durch zirkulierende T-Zellen als Las-Vegas-Algorithmus aufgefasst werden kann, lässt sich diese Erkenntnis unmittelbar übertragen. In der Arbeit untersuchen wir darauf aufbauend die konkrete “Laufzeitverteilung”, die sich durch das T-Zell-Zirkulationsszenario ergibt.

Diese “Laufzeitverteilung” wird durch folgende Parameter definiert: Sei ρ die Wahrscheinlichkeit, in einen drainierenden Lymphknoten einzuwandern; T die mittlere Zeit, die zwischen zwei Lymphknotenbesuchen im Blut, der Milz und anderswo verbracht wird; sowie α der Parameter einer Exponentialverteilung, die die Aktivierungszeit in einem drainierenden Lymphknoten beschreibt¹. Sei H der Zeitpunkt der Aktivierung der T-Zelle in einem drainierenden Lymphknoten, gerechnet vom Beginn der Infektion. Welche Verweilzeit R pro Lymphknoten minimiert den Erwartungswert $E[H]$? Man kann nun zeigen, dass gilt:

$$R = \frac{1}{\ln(1 - \alpha)} \left(W_{-1} \left(-\frac{(1 - \alpha)^{\frac{T}{1-\rho}}}{e} \right) + 1 \right) - \frac{T}{1 - \rho}$$

Hier bezeichnet W_{-1} den Nebenast der Lambertschen W -Funktion. Die durch diese Gleichung

¹Die Verwendung der Exponentialverteilung ergibt sich anhand wohlbekannter mathematischer Eigenschaften des dreidimensionalen random walks, und approximiert den random walk erstaunlich genau.

chung vorhergesagten Optima (Abbildung 2C) befinden sich in einem Bereich von etwa 9 Stunden, was zu experimentellen Daten über die tatsächliche Verweilzeit [SPN97] konsistent ist. Mittels asymptotischer Analysen lässt sich außerdem zeigen, dass das Optimum recht stabil gegenüber Parameterschwankungen ist, wie sie im biologischen Kontext zu erwarten sind. Als weiteres, recht überraschendes Ergebnis zeigt sich, dass der Parameter ρ im biologisch sinnvollen Bereich ($\rho \ll 1$) kaum Einfluss auf das Optimum hat. Die entzündungsvermittelte Erhöhung der Rekrutierungsrate und die Verweilzeit im Lymphknoten tragen also weitgehend unabhängig voneinander zur Effizienz der stochastischen T-Zell-Suche bei.

Zusammenfassend kann gesagt werden, dass das vorgeschlagene Modell trotz seiner Einfachheit in der Lage ist, experimentelle Daten aus verschiedensten Techniken wie Lymphdrainage, Histologie, Durchflusszytometrie und Zweiphotonenmikroskopie quantitativ in Verbindung zu bringen. Die Simulationen und die formale Analyse des stochastischen Modells lieferten quantitative, qualitative und funktionale Erkenntnisse, die durch experimentelle Ansätze nur schwer oder gar nicht hätten gewonnen werden können.

3 Negativauslese

Der zweite Teil der Arbeit befasst sich mit der Generierung und Negativauslese von T-Zellen. Diese Prozesse stellen sicher, dass immunokompetente T-Zellen normale Bestandteile des Organismus (Selbst) tolerieren und nur fremdartige Substanzen (Nichtselbst) angreifen. Im Wesentlichen wird dies erreicht, indem T-Zellen durch stochastische DNA-Rearrangements “zufällig generiert” werden, und dann normalen Proteinen aus dem Selbst ausgesetzt werden. Reagieren die neu erzeugten T-Zellen auf solche harmlosen Substanzen, werden sie wieder getötet. Die Arbeit greift ein etabliertes *algorithmisches* Modell [FPAC94] der Negativauslese auf. Der grundlegende Aufbau des Algorithmus wird in Abbildung 3 verdeutlicht: Als Grundlage definiert man eine Pattern-Matching-Funktion, die jedem Element einer Musterklasse Π eine Menge von Elementen eines Universums \mathcal{U} zuordnet (z.B. $\Pi = \mathcal{U} = \mathbb{R}^2$ mit dem euklidischen Abstand als Matching-Funktion). Der Algorithmus generiert zunächst per Zufall eine Pattern-Menge P und gleicht diese dann mit der Eingabemenge S ab. Alle Pattern, die ein Element von S matchen, werden gelöscht. Im Idealfall beschreiben die dann übrigen Pattern ungefähr die dem Algorithmus unbekanntes Bipartitionierung von \mathcal{U} , anhand derer die Beispiele generiert wurden.

Die Arbeit betrachtet zunächst grundlegende Eigenschaften dieses Algorithmus, die von der verwendeten Pattern-Matching-Funktion weitgehend unabhängig sind, aus der Perspektive der algorithmischen Lerntheorie. Dieser Ansatz zeigt Verbindungen zum konsistenten Lernen und dem Versionsraumlernen auf. Darauf aufbauend untersuchen wir verschiedene Mechanismen der Verallgemeinerung, die eine wesentliche Eigenschaft lernender Systeme ist. Verallgemeinerung kann zum Beispiel durch die Verwendung sogenannter *restringierter Pattern* erfolgen, bei denen ein zusätzlicher Parameter die Anzahl der gematchten Elemente pro Pattern kontrolliert (z.B. könnte man bei den Kreis-Pattern aus Abbildung 3 die Verallgemeinerung über den Radius einstellen). Weiterhin wird eine Verallgemeinerung des Negativauslese-Algorithmus vorgeschlagen, bei der nicht nur

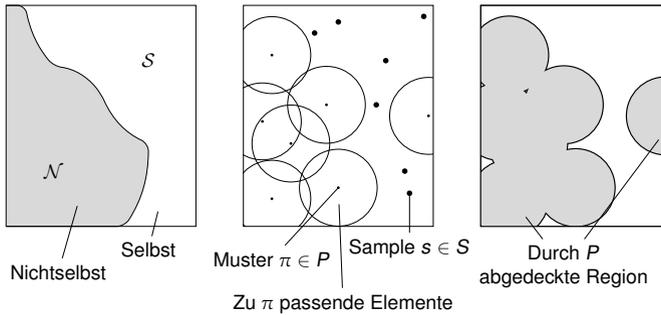


Abbildung 3: Illustration des Negativauslese-Algorithmus (siehe Text).

betrachtet wird *ob* ein zu klassifizierendes Element x durch ein Pattern aus P gematcht wird, sondern *wie viele* Pattern aus P das Element x matchen. Diesem Wert wird der Name *Sampling-Distanz* gegeben. Über einen Schwellwert für die Sampling-Distanz kann ebenfalls der Grad der Verallgemeinerung eingestellt werden.

In der Literatur wurden verschiedene konkrete Pattern-Matching-Funktionen zur Modellierung der Antigenerkennung durch T-Zellen vorgeschlagen. Diese basieren meist auf der Tatsache, dass T-Zellen nicht mit kompletten Proteinen interagieren, sondern nur mit kleinen Abbauprodukten von Proteinen (Peptiden). Abbildung 4 zeigt dieses Prinzip für die CD8-T-Zellen, die der Abwehr intrazellulärer Pathogene wie Viren dienen, und in deren Fall die Peptide typischerweise lineare Ketten von 9 Aminosäuren (*nonamere*) sind. Diese Beobachtung motivierte eine Klasse stringbasierter Modelle von T-Zell-Rezeptoren, bei denen T-Zellen als Strings und die Peptiderkennung als String-Matching formalisiert werden. Stringbasierte Modelle haben in der immunologischen Literatur bereits eine recht lange Tradition, und haben in jüngerer Zeit vielbeachtete qualitative Erkenntnisse geliefert [KRQ⁺10]. Allerdings ist die Nutzbarkeit dieser Modelle durch das Problem beschränkt, dass eine direkte Implementierung des Negativauslese-Algorithmus meist exponentielle Laufzeit aufweist, da die Pattern-Menge P exponentielle Größe hat [Sti09].

Angesichts dieses Problems stellen wir die Frage, unter welchen Bedingungen die Negativauslese effizient durch einen Ein-Ausgabe-äquivalenten Algorithmus simuliert werden kann. Wir zeigen zunächst allgemein, wie die Komplexität dieser Aufgabe anhand formaler Entscheidungs- und Zählprobleme charakterisiert werden kann, und analysieren dann nach dieser Methodik mehrere konkrete Pattern-Matching-Funktionen. Für einige davon erhalten wir tatsächlich effiziente Simulationsalgorithmen. Dies gelingt vor allem durch Anwendung von effizienten Datenstrukturen wie Präfix- und Suffixbäumen (die allerdings teilweise technisch komplex ist). Für andere Pattern-Matching-Funktionen führen wir dagegen Reduktionsbeweise durch, die zeigen, dass eine effiziente Simulation unter der Annahme $P \neq NP$ nicht möglich ist.

Glücklicherweise ist unser effizienter Simulationsansatz für einige wichtige stringbasierte Modelle durchführbar, unter anderem für das sogenannte *r-contiguous-Modell* [PPP93], für das bereits mehrere Jahre nach einem Polynomialzeitverfahren gesucht worden war

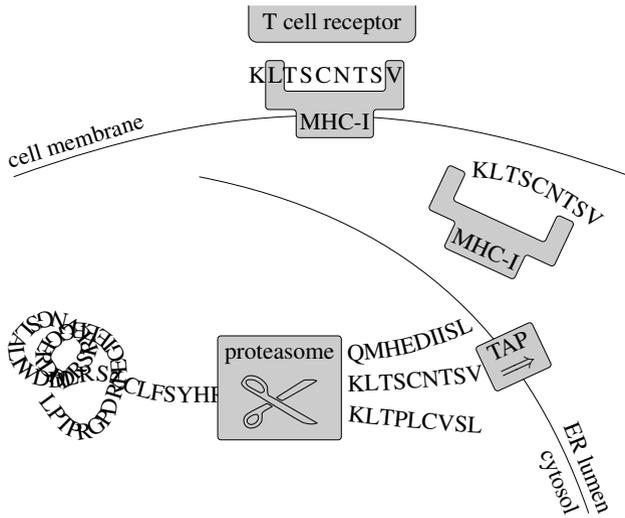


Abbildung 4: Schema der Antigenerkennung durch sogenannte CD8-T-Zellen (oder auch T-Killerzellen). Innerhalb jeder Zelle des menschlichen Organismus werden Stoffwechselprodukte kontinuierlich durch das Proteasom abgebaut und dabei in kleine Stücke aus wenigen Aminosäuren (Peptide) zerlegt. Diese Peptide werden durch das Molekül TAP in das endoplasmatische Retikulum transportiert, wo einige Peptide (meistens der Länge 9) an MHC-Klasse-I-Moleküle binden. Die Komplexe aus MHC und Peptid werden dann an die Zelloberfläche transportiert, wo sie den CD8-T-Zellen präsentiert werden. Durch das Prinzip der Negativauslese sind diese Zellen so generiert, dass sie normale Stoffwechselprodukte der Körpers nicht erkennen können, und daher im Normalfall nicht auf die präsentierten Peptide reagieren. Werden aber nun in der Zelle anormale Proteine synthetisiert, z.B. durch einen Virusbefall, so kann dies durch eine CD8-T-Zelle mit der richtigen Spezifität erkannt werden. Dies führt dann zum programmierten Zelltod der infizierten Zelle.

[Sti09]. Hierbei wird ein T-Zell-Rezeptor als Tupel aus einem Strings s der Länge ℓ und einer natürlichen Zahl $r \leq \ell$ modelliert, z.B. $(\text{KLTSVCNTSV}, 3)$. Eine solches Pattern matcht alle Strings der Länge ℓ , die zu s in mindestens r aufeinanderfolgenden Positionen identisch sind. Durch unsere algorithmischen Verbesserungen ist es nun möglich, quantitative Vorhersagen aus diesem Modell auch für große Eingabedatensätze zu generieren.

Zum Abschluss der Arbeit wenden wir die neu entwickelten algorithmischen Verfahren zur Untersuchung der Immunogenität von HIV-Peptiden an. Wie bei den meisten Viren können bei HIV nicht alle viralen Peptide, die den T-Zellen gezeigt werden, auch durch diese erkannt werden. Konkret werden von HIV-infizierten Zellen etwa 90 verschiedene Peptide² präsentiert, wovon nur etwa die Hälfte von T-Zellen erkannt wird. Die andere Hälfte unterteilt sich wiederum in Peptide, die nur selten erkannt werden (kryptisch), solche, die manchmal erkannt werden (subdominant), und einige wenige, die meistens erkannt werden (dominant). Es ist bislang kaum verstanden, welche Eigenschaften der Peptide die Immunogenität bestimmen, und dies ist eines der wichtigsten ungelösten Pro-

²Die genaue Zahl hängt vom entsprechenden Mutanten des Virus ab.

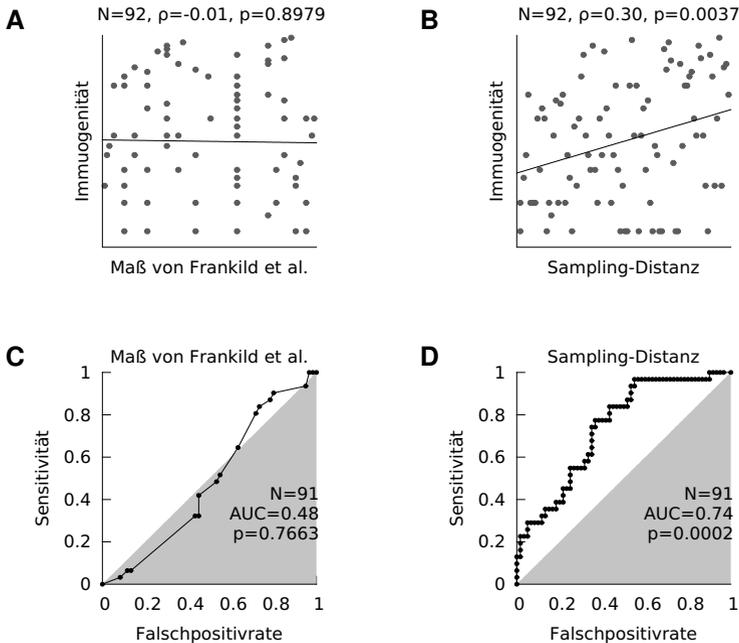


Abbildung 5: Quantitative Vorhersagen eines Modells aus der Literatur (A,C; [FdBL⁺08]) und des Negativauslesemodells aus der Dissertation (B,D) im Vergleich mit experimentellen Daten zur HIV-Infektion. (A,B) Für 92 Peptide der Länge 9 wurde anhand veröffentlichter Daten von 120 HIV-Patienten [FKA⁺04] die Erkennungshäufigkeit jedes Peptids bestimmt. Die Abbildungen zeigen die Korrelation (ρ) zwischen diesem Wert und den Vorhersagen des Literaturmodells (A) und des Negativauslesemodells (B). (C,D) Anhand der HIV-Datenbank des Los Alamos National Laboratory [YKB⁺09] wurde für 91 Peptide der Länge 9 bestimmt, welche davon prinzipiell immunogen sind. Die Abbildung zeigen die sich ergebenden ROC-Kurven, wenn man die Immunogenität der 91 Peptide mit der Literaturmethode (C) und dem Negativauslesemodell (D) vorhersagt. Zum Vergleich: Zufälliges Raten (graue Fläche) entspricht einer diagonalen ROC-Kurve mit einer Fläche (AUC) von 0,5, während ein perfekter Klassifikator eine AUC von 1,0 erreicht. Die p -Werte geben die Wahrscheinlichkeit an, dass entsprechende Ergebnisse durch zufälliges Raten zu Stande kommen.

bleme der modernen Immunologie. In der Arbeit stellen wir die Hypothese auf, dass einige Peptide unter anderem deswegen nicht oder nur schwer erkennbar sein könnten, weil die meisten dafür spezifischen T-Zellen die Negativauslese nicht überleben.

Um diese Hypothese zu testen, simulieren wir die Generierung und Negativauslese eines Repertoires künstlicher T-Zellen nach dem r -contiguous-Modell (mit einigen biochemisch motivierten Erweiterungen). Als Eingabemenge werden alle 140,000 bekannten Nonamere des menschlichen Proteoms verwendet. Tatsächlich korrelieren die Vorhersagen des Modells signifikant mit der Immunogenität der untersuchten Peptide (Abbildung 5), und erklären diese sogar weitaus besser als das bislang beste bekannte Modell aus der Literatur [FdBL⁺08]. Somit kann erstmals gezeigt werden, dass das r -contiguous-Modell trotz seiner Einfachheit tatsächlich quantitative Vorhersagekraft aufweist. Gleichzeitig wird eine neuartige Erklärung für die Immunogenität von HIV-Peptiden gefunden.

4 Ausblick

Wie zu Beginn betont wurde, lag der Hauptfokus der Arbeit auf der Generierung neuer Erkenntnisse für die Immunologie mit Methoden der (vor allem theoretischen) Informatik. Jedoch zeigt die Arbeit auch Anknüpfungspunkte zu aktuellen Fragestellungen der Informatik selbst auf. Diese ergeben sich vor allem zu bioinspirierten randomisierten Suchheuristiken (im ersten Teil) sowie zu sogenannten „künstlichen Immunsystemen“ für die Computersicherheit (zweiter Teil). Besonders hervorzuheben ist, dass der Algorithmus zur Negativauslese ursprünglich als Grundlage für Intrusion-Detection-Systeme konzipiert wurde [FPAC94]. Nach anfangs vielversprechenden Resultaten wurde die Forschung in diese Richtung nach etwa 10 Jahren weitgehend aufgegeben, was vor allem an der unzureichenden Effizienz des Algorithmus lag [Sti09]. Da die Arbeit dieses Effizienzproblem löst, erscheint weitere Forschung in diese Richtung nun wieder sinnvoll. Von allgemeinem Interesse dürfte auch die Grundlage für eine lerntheoretische Betrachtung des Immunsystems sein, die im zweiten Teil geschaffen wird. Das Immunsystem ist neben dem zentralen Nervensystem eines der beiden wichtigen kognitiven Systeme des menschlichen Körpers, wurde aber bisher kaum lerntheoretisch untersucht. Eine Vertiefung dieser Forschung birgt das Potential, gemeinsame Mechanismen des Lernens und der Gedächtnisbildung in beiden Systemen zu identifizieren.

Literatur

- [FdBL⁺08] Sune Frankild, Rob J. de Boer, Ole Lund, Morten Nielsen und Can Keşmir. Amino Acid Similarity Accounts for T Cell Cross-Reactivity and for “Holes” in the T Cell Repertoire. *PLoS one*, 3(3):e1831, 2008.
- [FKA⁺04] N. Frahm, B. T. Korber, C. M. Adams et al. Consistent cytotoxic-T-lymphocyte targeting of immunodominant regions in human immunodeficiency virus across multiple ethnicities. *Journal of Virology*, 78:2187–2200, 2004.
- [FPAC94] Stephanie Forrest, Alan S. Perelson, Lawrence Allen und Rajesh Cherukuri. Self-Nonself Discrimination in a Computer. In *Proceedings of the IEEE Symposium on Research in Security and Privacy*, Seiten 202–212. IEEE Computer Society Press, 1994.
- [KRQ⁺10] Andrej Košmrlj, Elizabeth L. Read, Ying Qi et al. Effects of thymic selection of the T-cell repertoire on HLA class I-associated control of HIV infection. *Nature*, 465:350–354, 2010.
- [LSZ93] Michael Luby, Alistair Sinclair und David Zuckerman. Optimal speedup of Las Vegas algorithms. *Information Processing Letters*, 47:173–180, 1993.
- [MCP⁺07] J. J. Moon, H. H. Chu, M. Pepper et al. Naive CD4+ T cell frequency varies for different epitopes and predicts repertoire diversity and response magnitude. *Immunity*, 27:203–213, 2007.
- [MHvA04] Thorsten R. Mempel, Sarah E. Henrickson und Ulrich H. von Andrian. T-cell priming by dendritic cells in lymph nodes occurs in three distinct phases. *Nature*, 427:154–159, 2004.

- [MWPC02] M. J. Miller, S. H. Wei, I. Parker und M. D. Cahalan. Two-photon imaging of lymphocyte motility and antigen response in intact lymph node. *Science*, 296:1869–1873, 2002.
- [PPP93] Jerome K. Percus, Ora E. Percus und Alan S. Perelson. Predicting the size of the T-cell receptor and antibody combining region from consideration of efficient self-nonself discrimination. *Proceedings of the National Academy of Sciences of the United States of America*, 90(5):1691–1695, 1993.
- [SJHC11] A. T. Stock, C. M. Jones, W. R. Heath und F. R. Carbone. Rapid recruitment and activation of CD8+ T cells after herpes simplex virus type 1 skin infection. *Immunology and Cell Biology*, 89:143–148, 2011.
- [SPN97] Dov J. Stekel, Claire E. Parker und Martin A. Nowak. A model of lymphocyte recirculation. *Immunology Today*, 18(5):217–221, 1997.
- [SPS⁺05] Kelly A. Soderberg, Geoffrey W. Payne, Ayuko Sato et al. Innate control of adaptive immunity via remodeling of lymph node feed arteriole. *Proceedings of the National Academy of Sciences of the USA*, 102(45):16315–16320, 2005.
- [Sti09] Thomas Stibor. Foundations of R-Contiguous Matching in Negative Selection for Anomaly Detection. *Natural Computing*, 8:613–641, 2009.
- [vHGS⁺09] J. W. van Heijst, C. Gerlach, E. Swart et al. Recruitment of antigen-specific CD8+ T cells in response to infection is markedly efficient. *Science*, 325:1265–1269, 2009.
- [WPMC03] Sindy H. Wei, Ian Parker, Mark J. Miller und Michael D. Cahalan. A stochastic view of lymphocyte motility and trafficking within the lymph node. *Immunological Reviews*, 195:136–159, 2003.
- [YKB⁺09] Karina Yusim, Bette T. M. Korber, Christian Brander et al. HIV Molecular Immunology 2009. Bericht LA-UR 09-05941, Los Alamos National Laboratory, Theoretical Biology and Biophysics, Los Alamos, New Mexico, 2009.



Johannes Textor, geboren 1979 in Göttingen, studierte von 1999 bis 2006 Informatik an der Universität zu Lübeck und promovierte dort von 2006 bis 2011 am Institut für Theoretische Informatik. Seit September 2011 ist er als Postdoktorand im Institut für Theoretische Biologie und Bioinformatik der Universität Utrecht in den Niederlanden tätig. Dort forscht er in verschiedenen Kooperationen mit dem niederländischen Krebsforschungsinstitut, der Harvard Medical School, den National Institutes of Health, und der Washington University St. Louis weiter schwerpunktmäßig zur Migration von T-Zellen.