

Abstract

In this computational work, I developed models that describe the interactions of proteases and random copolymers. For the development of the models, I created the framework *PolyLibScan* with which new models can be set up, validated and improved.

The modulation of proteins with copolymers is an emerging field that has sparked much interest but as of now, there are only few computational contribution to the field of copolymer-protease interactions.

Because there are still many open questions, it is important to develop new tools that can help the understanding in this field.

The high conformational flexibility of the copolymers and also their large variability of their random sequences pose a challenge to successfully explore conformational- and sequence-space. Because of this challenge, the model relies on coarse grained simulations to reduce complexity and make computation feasible.

For the calibration and validation of the models, I use the results of the inhibition experiments of Gilles et al., who found numerous inhibitors among 49 synthesized random copolymer-types for six proteases.

For the coarse graining, the most minimalistic approach is chosen as a starting point by using a single united atom per monomer and amino acid. The inhibition property of each copolymer-type is determined by multiple simulations.

The first model relies on two computational methods to identify binding sites for the monomers and the proteases, which are used to parameterize the model. Because this model is not accurate enough, a second model is developed that relies on the charges of monomers and amino acids for the interactions. This model has a better agreement with the experiments for three of the proteases. Based on these results, the third model adds the hydrophobic interaction, more precisely takes into account the implicit solvent and makes the protonation of the particles pH-dependent. With these changes, the third model is a further improvement over the second model.

The last model changes the coarse grained structure of the copolymers to more closely resemble the all-atom copolymers. Even though the results are different to the last model, it is similar in its accuracy to the third model.

Despite their minimalistic structure, the last two models accomplish a good agreement with four of the six protease, but also show the limits of single bead models with the other cases.

Zusammenfassung

Im Rahmen dieser Arbeit wurden vier Modelle entwickelt, welche die Interaktion von Proteasen und statistischen Copolymeren beschreiben. Für die Entwicklung der Modelle habe ich das Framework *PolyLibScan* erschaffen, in welchem die Modelle erstellt, validiert und verbessert werden können. Die Veränderung der Wirkungsweise von Proteinen und Copolymeren ist ein aufkommendes Forschungsfeld, in dem es noch wenige bioinformatische Beiträge gibt. Weil die Wirkungsweise von Copolymeren-Protease-Wechselwirkungen viele offene Fragen hat, ist es wichtig neue Werkzeuge für die Verbesserung des Verständnisses zu entwickeln. Die hohe Beweglichkeit der Copolymere und die große Variabilität der statistischen Sequenzen stellen eine Herausforderung dar, den großen Konformations- und den Sequenzraum genügend zu sampeln. Aufgrund dieses Problems nutzen die Modelle coarse grained (vergrößerte) Simulationen, um die Komplexität des Systems zu verringern und die Berechnung zu ermöglichen. Zur Kalibrierung und Validierung der Modelle verwende ich die Ergebnisse der Inhibitionsexperimente von Gilles et al., welche mehrere Inhibitoren aus 49 synthetisierten Copolymerarten für sechs verschiedene Proteasen identifizierten. Für das Coarse-Graining benutzte ich einen möglichst minimalistischen Ansatz, indem jedes Monomer und jede Aminosäure durch ein einzelnes vereinigt Atom ersetzt wird. In den folgenden Simulationen interagiert ein Copolymer mit einer Protease und kann diese inhibieren, wenn das Copolymer in der Nähe des aktiven Zentrums bindet. Eine Vielzahl an Simulationen wird für viele statistische Sequenzen und Startbedingungen durchgeführt.

Das erste Modell basiert auf zwei computergestützten Verfahren, mit denen die Bindestellen der Monomere auf den Proteaseoberflächen bestimmt werden. Basierend auf diesen Bindestellen werden die Wechselwirkungen im Modell festgelegt. Aufgrund der unzureichenden Genauigkeit des Modells wird im zweiten Modell die elektrostatische Wechselwirkung als Grundlage gewählt. Dieses Modell liefert eine bessere Übereinstimmung mit den Inhibitionsexperimenten von drei Proteasen. Beim dritten Modell wird die hydrophobe Wechselwirkung hinzugefügt und das implizite Lösungsmittel besser berücksichtigt. Zusätzlich wird die Protonierung der Teilchen abhängig vom pH gemacht. Diese Änderungen führen zu einer weiteren Verbesserung des Modells. Das letzte Modell ändert die Struktur der Copolymere, damit sie ähnlicher der atomistischen Struktur der Copolymere wird, was zu weiteren Verbesserungen führt. Trotz der minimalistischen Struktur der Monomere und Aminosäuren konnten die Modelle eine gute Übereinstimmung mit den experimentellen Daten liefern. Auf der anderen Seite haben sich auch die Grenzen dieses Ansatzes bei den zwei anderen Proteasen gezeigt.