

Virtual insights on G protein inhibition and ion channel block – a computer-based study

Vom Fachbereich Chemie
der Technischen Universität Darmstadt

zur Erlangung des Grades
Doctor rerum naturalium
(Dr. rer. nat.)

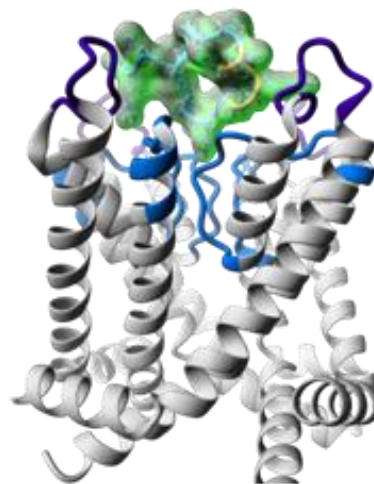
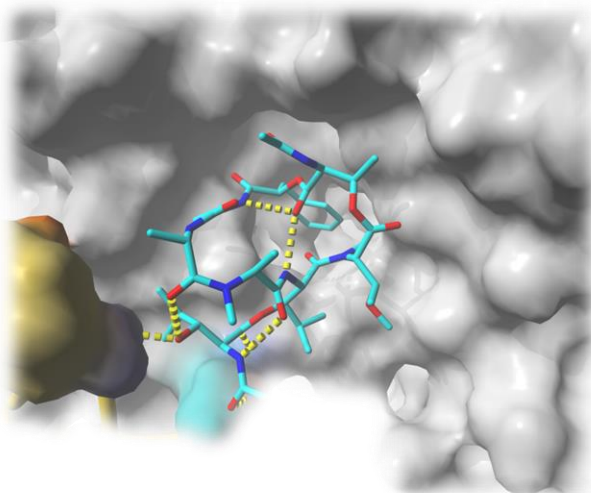
Dissertation
von Desiree Kaufmann

Erstgutachter: Dr. Daniel Tietze
Zweitgutachter: Prof. Dr. Gerd Buntkowsky

Darmstadt 2019



TECHNISCHE
UNIVERSITÄT
DARMSTADT



Kaufmann, Desiree:

Virtual insights on G protein inhibition and ion channel block – a computer-based study

Darmstadt, Technische Universität Darmstadt,

Jahr der Veröffentlichung der Dissertation auf TUprints: 2019

URN: urn:nbn:de:tuda-tuprints-91579

Tag der Einreichung: 16. Mai 2019

Tag der mündlichen Prüfung: 1. Juli 2019

Veröffentlicht unter CC BY-SA 4.0 International

<https://creativecommons.org/licenses/>



Dekan: Prof. Dr. Harald Kolmar

Gutachter:
Erstgutachter: Dr. Daniel Tietze
Zweitgutachter: Prof. Dr. Gerd Buntkowsky

Fachprüfer:
1. Fachprüfer: Prof. Dr. Felix Hausch
2. Fachprüfer: Prof. Dr. Wolf-Dieter Fessner

“This is the essence of science. First you think of something that could be true. Then you see if it is the case and in general it is not the case.”

„Science does not aim at establishing immutable truths and eternal dogmas; its aim is to approach the truth by successive approximations, without claiming that at any stage final and complete accuracy has been achieved.”

--- *Bertrand Russell (1872-1970), British philosopher, mathematician and Nobel Laureate for Literature*

„There are two kinds of truth: the truth that lights the way and the truth that warms the heart. The first of these is science, and the second is art. Neither is independent of the other or more important than the other. Without art, science would be as useless as a pair of high forceps in the hands of a plumber. Without science, art would become a crude mess of folklore and emotional quackery. The truth of art keeps science from becoming inhuman, and the truth of science keeps art from becoming ridiculous.”

--- *"Great Thought" (19 February 1938), published in The Notebooks of Raymond Chandler (1976)*

„Not only is the universe stranger than we think, it is stranger than we *can* think.”

--- *Werner Heisenberg (1901-1976), German scientist, physicist and Nobel Laureate in Physics*

„Science starts getting interesting where it ends.”

--- *Justus von Liebig (1803-1873), German chemist and friend of Emanuel Merck*

Acknowledgements

In the first place I would like to give my great thanks to my supervisor Dr. Daniel Tietze, who I am very glad to know. I am grateful to him for his dedicated supervision and constant support for my research activities in these interesting fields. Daniel knew to respectfully guide me while giving me the sufficient amount of freedom during the whole progress of this work.

My sincere thank is also owed to Prof. Dr. Gerd Buntkowsky from the professorial chair at the Department of Physical Chemistry at the Technical University of Darmstadt as well as to Prof. Dr. Diana Imhof from the professorial chair at the Institute of Pharmaceutical Biochemistry and Bioanalytics at the University of Bonn for making this work possible. In this context, I would like to express my special thanks to Prof. Dr. Gerd Buntkowsky for enabling me the attendance at the Conference of the “Seventh Annual Niels Bohr International Academy Workshop on ESS Science” (Deciphering the hidden dynamics of Soft Matter) in Kopenhagen and thus giving me the opportunity to establish inspiring contacts.

Furthermore my thanks go to David Rosenberger and to Tobias Mongelli from the Computational Physical Chemistry Department at the Technical University of Darmstadt for technical support issues, as well as to Dr. Raphael Reher, Dr. Toni Kühl and to Dr. Pascal Heimer from the Pharmaceutical Biochemistry and Bioanalytics Research Unit at the University of Bonn for the constructive teamwork and the interesting collaborations that resulted in a common publication which is part of this work.

Of course, I would also like to thank Dr. Hergen Breitzke and all other colleagues of Professor Gerd Buntkowsky’s working group as well as the members of Dr. Alesia Tietze’s lab for the help with any unforeseen problems and of course for the pleasant and collegial atmosphere.

It is important to me to extend my thanks to Daniyal Kazempour, an in my opinion outstanding and inspiring scientist and friend I met during my studies. He encouraged and motivated me on staying persistent and genuine during the past years.

Last but not least I would of course like to equally give my thanks to all my other friends, to my family and my fiancé for their constant and devoted support and accompaniment in any stage of my life.

I would like to dedicate this work to family Fleischer.





Jurat

In accordance with the standard scientific protocol, I will use the personal pronoun *we* to indicate the reader and the writer, or my scientific collaborators and myself.
I affirm that I created this work autonomously and only with the aid of the cited literature.

date: _____

signature: _____



Abstract

The present work focuses on understanding the mechanisms of action of two pharmaceutically relevant inhibitor protein systems. Understanding the mechanism of how drugs or drug candidates are affecting their molecular target is of vital importance in order to develop a promising drug candidate into a valuable medicine. Nowadays, such understanding is mainly gained from computational studies, better known as molecular modelling approaches, which are an essential part of every drug development campaign.

Using the example of two distinct protein ligand systems, respectively consisting of a target protein and an appropriate binding molecule, we aimed at elucidating the intrinsic subtype specificity determinants inherent in interacting biological systems by employing up-to-date computational techniques.

The first part of the thesis focuses on the subtype-specific interaction of the cyclic depsipeptides YM-254890 (YM) and FR900359 (FR) with the alpha (α) subunit of a heterotrimeric G protein. Through this interaction with the α subunit of the G_q G protein subtype the immanent G protein function of GDP/GTP exchange is hampered and thus intervenes the associated signal transmissions within and between cells. So far, the direct targeting of intracellular $G\alpha$ subunits by chemical tools represents a novel promising pharmacological approach for e.g. targeting oncogenic $G_{q/11}$ proteins as a potential therapeutic option for cancer treatment of the adult eye.¹⁻⁵

Here, we thoroughly analyzed novel FR900359-based analogues from natural sources, synthetic cyclic peptides as well as all so-far known $G\alpha_q$ inhibitors in a comprehensive computational study. Using molecular docking we have identified unique features in the macrocyclic FR/YM structures responsible for the binding to the target protein and correlating with their inhibitory activity.⁵ We hence could elucidate the FR/YM specific characteristics that determine their interaction with the G_q binding pocket.⁵

Our main findings suggest key regions in G_q , more specifically two major anchor positions at the joints of the G protein α subunit, namely residue Arg60, situated at a protein region called “linker 1”, as well as the two adjacent residues Ile190 and Glu191 located close to a protein region called “switch 1”. The fixation of these two hinge points by the crucial *N*-MeDha5 and the opposite β -hydroxy-leucine residues in YM/FR inhibits their intrinsic mobility and thus counteracts the opening movement of the $G\alpha$ subunit which is relevant for GDP/GTP exchange. An additional feature attributed to the YM/FR inhibitor is the associated stability of its backbone structure, which enables it to fit into the binding pocket of the G protein and thus address said critical positions.

Although our computational data were in agreement with experimental results that no other inhibitor exceeded the biological activity of wild type FR, the question remains whether FR analogues represent valuable chemical tools for the targeted and specific inhibition and whether they constitute suitable natural lead structures for the development of novel compounds to target $G\alpha$ subunits apart from the $G\alpha_q$ subtype.⁵

The second part of this work focuses on the investigation of voltage-gated ion channels and their – equally subtype-specific – conotoxin blockers, venoms that are obtainable from the marine cone snail. We herein mainly focused on voltage-gated sodium (Na_v) and potassium (K_v) channels both of which are mainly found in the central nervous system (CNS). Their selective blockage prevents the reciprocal ion flux across the membrane that separates a neuron from its adjoining synaptic cleft. The flow of ions, in case of resulting in a potential difference between the nerve interior and the outside, finally accounts for the so-called resting potential. This potential difference is a natural property that is usually provided by selectively permeable membranes and makes up an essential prerequisite for the initiation of stimuli conductions. Understanding the subtype-specific ion channel block by such natural peptide-based toxins can therefore provide essential guidance for the further advancement of such compounds into promising drug candidates.⁶

The first associated study comprises a docking study of 15 different disulfide bond isomers of the μ -conotoxin PIIIA to the sodium channel subtype $Na_v1.4$ of the Na_v1 sodium channel family, for which a subtype-specific block by the native μ -PIIIA isomer was observed from experimental data.

Using a sophisticated docking approach, we achieved to demonstrate a correlation between the μ -PIIIA disulfide bridge pattern and the toxin's capability to block $Na_v1.4$. Since the structural conformation of a toxin and the associated flexibility results from its internal disulfide bridging pattern, the extent of the pore block can be directly attributed to the toxins' structural properties.

Here, through our docking approach we were able to gain further information on which toxin parts of the isomers address certain parts of the pore surface. It could be demonstrated that there is an ultimate correlation between the cysteine disulfide bridge pattern of a toxin, which determines its structure and flexibility, and finally its extent of ability to block certain channel subtypes. In contrast to active isomers, the moderate or poorly active isomers either lack the central pore contact and/or the distributed covering contacts to the outer pore surface. In particular, it could be shown that a channel is effectively blocked by a full pore coverage, whereas only partially covered in weaker blocks.

In these respects, the docking experiments were in accordance with laboratory generated data, pointing out a specific binding pose of the most active native μ -PIIIA isomer, comprising coincident contacts with the inner central pore duct, addressing the very well-known “DEKA” selectivity filter motif and the outer located loop regions of the channel. Due to the more globular fold of the native μ -PIIIA isomer, which is caused through its unique disulfide pattern, the acidic pore residues are primarily addressed by the basic arginine residues of the toxin, stabilizing the central pore-covering toxin position on the channel.

The anticipated importance of the toxins’ flexibilities and how they influence the different bioactivities is explored in more detail by our second study on subtype-selective ion channel blockage through μ -conotoxins, where, apart from docking studies, molecular dynamics simulations are employed.

Here, an extensive molecular dynamics analysis of the μ -PIIIA, μ -SIIIA and μ -GIIIA conotoxins on the voltage-gated K_V1 potassium channel family subtypes $K_V1.1$, $K_V1.5$ and $K_V1.6$ and on further artificially created chimeric models is given. In order to more thoroughly understand the dynamics and binding states of these ion channel toxin systems, we correlated them with their experimental activities which were recently reported by Leipold et al.^{6,7} As for all channel-bound toxins investigated herein, a certain conformational flexibility that reflects the individual pore blocking potencies could be observed through our molecular dynamic simulations, the necessity of considering dynamic data for the understanding of subtype-specific blocking events and concomitant activity deviations can be assumed.⁶

Consistent with our previous studies on the $Na_V1.4$ blockage by various μ -conotoxin PIIIA isomers, our data suggest an unambiguous and unique binding mode of active μ -PIIIA on $K_V1.6$ and $K_V1.1$, made up by concomitant hydrogen bonds contacts towards the outer and the inner channel pore loops.⁶ Here, a plethora of hydrogen bonds is formed in particular by the toxin’s centric arginine and lysine residues within the helical μ -PIIIA core region towards the central pore residues.⁶ In this context, a special, symmetrically designed toxin segment in the central μ -PIIIA core region, was identified, which contains the motif Arg12-Ser13-Arg14-Gln15 shown to be crucial for ion channel binding.⁶

Finally, it is worth noting that for all systems considered, we were able to determine structural reasons for the subtype-selective inhibition or blocking processes. Here, the subtype-specific binding of μ -PIIIA to the K_V1 family members does not tolerate positively charged amino acids in both the inner and outer pore loops, as shown by comparison of $K_V1.6$, $K_V1.3$ and $K_V1.1$. Additionally, the lack of the essential third aspartic acid residue in the outer P1 loop of $K_V1.3$ further strengthens the exclusive μ -PIIIA binding to $K_V1.6$ and $K_V1.1$.

However, considering the active inhibition of G proteins as well as the effective blockage of ion channels, the essential specificity determinants can be broken down into definite, unique attributes.

Highly interesting and instructive in this context was the apparent interdependence of all these critical determinants in both systems considered that were decisive for their subtype-specific blockage or inhibition: It was found that in the absence of at least one of these criteria, the functioning of an active system or its subtype-specific working mechanisms can no longer be maintained. Such insights can provide comprehensive support for the development of novel and highly specific drug agents.

Last but not least, our explorations resulted in the successful generation of reliable data sets that are in accordance with to date published literature from laboratory experiments and might even be able to lay the foundation of learning or training data sets required for further computer-based investigations on similar systems.

Zusammenfassung

Die vorliegende Arbeit konzentriert sich auf das Verständnis der Wirkmechanismen von zwei pharmazeutisch relevanten Inhibitor-Protein-Systemen. Das Verständnis des Mechanismus, wie Medikamente oder Medikamentenkandidaten auf ihr molekulares Ziel einwirken, ist von entscheidender Bedeutung für die Entwicklung eines erfolgsversprechenden Medikamentenkandidaten zu einem nützlichen Medikament. Heutzutage wird dieses Verständnis hauptsächlich aus rechnergestützten Studien gewonnen, besser bekannt als molekulare Modellierungsansätze, die ein wesentlicher Bestandteil jeder Medikamentenentwicklungskampagne sind.

Anhand zweier betrachteter Beispiele von Protein-Ligand-Systemen, jeweils bestehend aus einem adressierten Zielprotein und einem entsprechenden Bindungsmolekül, wird die natürlich intrinsische Eigenschaft der Subtypspezifität, welcher biologische interagierende Systeme unterliegen, mit Hilfe moderner computerbasierter Technologien entschlüsselt.

Der erste Teil dieser Arbeit widmet sich den zyklischen Depsipeptiden YM-254890 (YM) und FR900359 (FR) und deren Zielprotein, der alpha- (α) Untereinheit eines heterotrimeren G-Proteins. Ihre subtypspezifische Bindung an besagte α -Untereinheit des G-Protein-Subtyps G_q verhindert dessen eigentliche Funktion; den Austausch von gebundenem GDP mit GTP, womit weitere Signalübertragungswege innerhalb und zwischen Zellen unterbunden werden. Bisher stellt die direkte Adressierung intrazellulärer $G\alpha$ -Untereinheiten mittels chemischer Werkzeuge einen neuartigen, vielversprechenden pharmakologischen Forschungsansatz dar, um z.B. onkogene $G_{q/11}$ -Proteine als potenzielle Therapieansatzpunkte für die Krebsbehandlung des erwachsenen Auges zu nutzen.¹⁻⁵

Im Rahmen einer umfassenden Computerstudie haben wir hier neuartige FR900359-basierte Analoga aus natürlichen Quellen, synthetische zyklische Peptide sowie alle bisher bekannten $G\alpha_q$ Inhibitoren eingehend analysiert. Mit Hilfe molekularer Dockingberechnungen konnten wir einzigartige Merkmale in den makrozyklischen FR/YM-Strukturen identifizieren, die für die Bindung an das Zielprotein verantwortlich sind und in direktem Zusammenhang mit deren hemmender Wirkung stehen.⁵ Wir konnten somit die FR/YM-spezifischen Besonderheiten ergründen, die ihre Interaktion mit der G_q -Bindetasche ausmachen.⁵

Unsere Hauptkenntnisse zeigen für die Hemmung kritische Schlüsselpositionen in G_q auf: insbesondere handelt es sich hierbei um zwei wichtige Ankerpositionen an Gelenkteilen der $G\alpha$ -Untereinheit, nämlich das Residuum Arg60, welches sich am sogenannten „Linker 1“ befindet, sowie die beiden nebeneinanderliegenden Aminosäurereste Ile190 und Glu191, die in der Nähe der Proteinregion namens „Switch 1“ lokalisiert sind. Die Fixierung dieser beiden Scharnierpunkte durch die entscheidenden Residuen *N*-MeDha5 und die gegenüberliegenden β -Hydroxyleucinreste im YM-/FR-Molekül hemmt deren intrinsische Mobilität und wirkt damit der für den GDP/GTP-Austausch relevanten Öffnungsbewegung der $G\alpha$ -Untereinheit entgegen. Ein weiteres Merkmal, welches den Inhibitoren YM und FR zugeschrieben wird, ist ihre stabile Rückgratstruktur, die es ihnen ermöglicht, passgenau die Bindungstasche des G-Proteins zu adressieren und so die kritischen Positionen innerhalb dieser Bindungstasche anzusprechen.

Obwohl unsere Daten in Übereinstimmung mit den experimentellen Ergebnissen nachweisen konnten, dass kein anderer Inhibitor die biologische Aktivität des FR-Wildtyps übertrifft, bleibt dennoch die offene Frage, ob FR-Analoga wertvolle chemische Werkzeuge für die gezielte und spezifische Hemmung darstellen und darüber hinaus auch, ob sie als geeignete natürliche Leitstrukturen für die Entwicklung neuer Substanzen zur Adressierung anderer $G\alpha$ -Untereinheiten als die des G_q -Subtyps fungieren können.⁵

Der zweite Teil dieser Arbeit konzentriert sich auf die Untersuchung spannungsgesteuerter Ionenkanäle und deren – ebenso subtypspezifischen – Kegelschneckenblocker, auch bekannt als Conotoxine, also spezifischen Giften, die aus marinen Kegelschnecken gewonnen werden können. In diesem Teil haben wir uns hauptsächlich auf spannungsgesteuerte Natrium- (Na_v) und Kalium- (K_v) Kanäle fokussiert, die beide hauptsächlich im zentralen Nervensystem (ZNS) vorkommen. Deren selektive Blockierung verhindert den bidirektionalen Ionenfluss durch die Membran, die ein Neuron von seinem angrenzenden synaptischen Spalt trennt. Der transmembranöse Ionenfluss, der zu einer Potentialdifferenz zwischen Nervenzellinnerem und dem außen liegenden synaptischen Spalt führt, trägt schließlich zur Aufrechterhaltung des sogenannten Ruhepotenzials bei. Diese Potentialdifferenz ist eine natürlich gegebene Eigenschaft selektiv-permeabler Membranen und bildet eine wesentliche Grundvoraussetzung für die Initiierung der Reizweiterleitung.

Das Verständnis der subtypspezifischen Ionenkanalblockierung durch diese natürlichen peptidbasierten Toxine kann daher wesentliche Orientierungshilfen für die Weiterentwicklung solcher Peptide zu vielversprechenden potentiellen Medikamenten leisten.⁶

Die erste zugehörige Studie umfasst eine Docking-Studie an 15 verschiedenen Disulfidbindungsisomeren des μ -Conotoxins PIIIA am Natriumkanal-Subtyp Nav1.4 aus der Nav1-Natriumkanalfamilie, für welche experimentelle Messungen eine subtypspezifische Blockierung des Nav1.4-Subtyps durch das native μ -PIIIA-Isomer nachweisen konnten.

Unter Verwendung ausgereifter Docking-Methoden ist es uns gelungen, den eindeutigen Zusammenhang des nativen μ -PIIIA-Disulfidbrückenmusters und der Fähigkeit des Toxins, den Nav1.4-Kanal zu blockieren, nachzuweisen. Da die strukturelle Konformation eines Toxins und seine damit verbundene Flexibilität unmittelbar aus dessen internem Disulfidbrückenmuster resultiert, kann das Ausmaß der Kanalporenblockierung direkt auf die strukturellen Eigenschaften der Toxine zurückgeführt werden.

Unsere Docking-Ergebnisse konnten Aufschluss darüber geben, welche Toxinanteile der Isomere bestimmte Teile der Porenoberfläche ansprechen. Es konnte gezeigt werden, dass eine eindeutige Korrelation zwischen dem Cysteindisulfid-Brückenmuster eines Toxins, das auch seine Struktur und Flexibilität bestimmt, und schließlich seiner Blockierungsfähigkeit bestimmter Kanalsubtypen besteht. Im Gegensatz zu aktiven Isomeren zeigen mäßig oder schwach aktive Isomere keine Kontakte zur Porenmitte und/oder gleichmäßig verteilte, abdeckende Kontakte zum Außenbereich der Porenoberfläche auf. Insbesondere konnte gezeigt werden, dass die Kanalpore bei effektiver Blockierung vollständig und bei schwächerer Blockierung nur teilweise abgedeckt wird. In dieser Hinsicht stimmten die Docking-Experimente mit den laborgenerierten Daten überein und zeigten eine spezifische Bindungsposition des am stärksten blockierenden nativen μ -PIIIA-Isomers auf; diese wird durch gleichzeitige Kontakte zum Poreninneren im Kanalzentrum, in welchem auch das sehr bekannte Selektivitätsfiltermotiv "DEKA" enthalten ist, und zu den außen liegenden Loop-Regionen des Kanals ermöglicht. Durch die eher globulärförmige Toxinfaltung, die das native μ -PIIIA-Isomer durch sein einzigartiges Cysteindisulfidmuster annimmt, werden die sauren Porenresiduen vorwiegend durch basische Argininreste des Toxins adressiert und die zentrale porenabdeckende Toxinpositionierung auf dem Kanal stabilisiert.

Die vorweggenommene Bedeutung der Toxinflexibilitäten und deren Einfluss auf die verschiedenen Bioaktivitäten wird in unserer zweiten Studie zur subtypspezifischen Blockade von Ionenkanälen durch μ -Conotoxine näher beleuchtet, in welcher neben Docking-Studien auch molekulardynamische Simulationen eingesetzt wurden.

Hier wird eine umfassende Molekulardynamik-Analyse der μ -PIIIA, μ -SIIIA und μ -GIIIA Conotoxine auf den spannungsgesteuerten Kaliumkanalsubtypen Kv1.1, Kv1.5 und Kv1.6 der Kaliumkanalfamilie Kv1 und virtuell erstellten Chimärmodellen geliefert. Um die Dynamik und Bindungszustände dieser Ionenkanal-Toxinsysteme besser zu verstehen, korrelierten wir sie mit ihren experimentellen Aktivitäten, die kürzlich von Leipold und Mitarbeitern⁷ berichtet wurden.⁶ Da mittels unserer molekulardynamischen Simulationen für alle hier untersuchten kanalgebundenen Toxine eine gewisse Konformationsflexibilität und die dadurch adäquate Wiedergabe ihrer individuellen Blockierungsvermögen beobachtet werden konnte, ist davon auszugehen, dass die Berücksichtigung dynamischer Daten zum Verständnis subtypspezifischer Blockierungsereignisse und damit einhergehender Aktivitätsabweichungen notwendig ist.⁶

In Übereinstimmung mit unseren vorangegangenen Studien zur Nav1.4-Blockade durch verschiedene μ -PIIIA-Isomere suggerieren unsere Daten einen eindeutigen und einzigartigen Bindungsmodus des auf Kv1.6 und Kv1.1 aktiven μ -PIIIA, der sich aus gleichzeitigen Wasserstoffbrückenkontakten zu den äußeren und inneren Loop-Teilen der Kanalpore zusammensetzt.⁶ Hier wird insbesondere durch die zentrischen Arginin- und Lysinreste des Toxins innerhalb der helikalen μ -PIIIA-Kernregion eine Vielzahl an Wasserstoffbrückenbindungen zu den mittig liegenden Porenresiduen gebildet.⁶ In diesem Zusammenhang wurde ein spezielles, symmetrisch gestaltetes Toxinsegment in der zentralen Kernregion von μ -PIIIA identifiziert, das das Motiv Arg12-Ser13-Arg14-Gln15 enthält, welches sich auch als entscheidend für Ionenkanalbindung erwiesen hat.⁶

Zum Schluss ist anzumerken, dass wir einheitlich für alle betrachteten Systeme strukturelle Faktoren, in welchen die subtypselektiven Hemm- oder Blockierungsprozesse begründet liegen, ermitteln konnten. Dabei werden für die subtypspezifische Bindung von μ -PIIIA an die K_v1 -Kaliumkanalfamilie keine positiv geladenen Aminosäuren, weder in den inneren noch in den äußeren Loop-Teilen der Pore toleriert, was anhand des Vergleichs von $K_v1.6$, $K_v1.3$ und $K_v1.1$ festgestellt werden kann. Die Ausschließlichkeit der Bindung von μ -PIIIA an die Kaliumkanalsubtypen $K_v1.6$ und $K_v1.1$ wird zudem durch das Fehlen des essentiellen dritten Asparaginsäurerests in den äußeren P1-Loopregionen des $K_v1.3$ -Subtyps weiter untermauert.

Unter Berücksichtigung der aktiven Hemmung von G-Proteinen sowie der effektiven Blockierung von Ionenkanälen können die wesentlichen spezifitätsbestimmenden Faktoren jedoch auf konkrete, einzigartige Attribute heruntergebrochen werden.

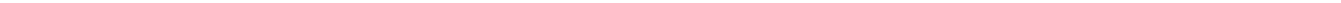
Hochinteressant und lehrreich in diesem Zusammenhang war die scheinbar gegenseitige Abhängigkeit all dieser kritischen Determinanten in beiden betrachteten Systemen, die für ihre subtypspezifische Blockade oder Hemmung ausschlaggebend waren: Es stellte sich heraus, dass bei Fehlen mindestens eines dieser Kriterien die Funktionsfähigkeit eines aktiven Systems bzw. seiner subtypspezifischen Wirkmechanismen nicht mehr aufrechterhalten werden kann. Solche Erkenntnisse können die Entwicklung neuer und hochspezifischer Wirkstoffe ganzheitlich unterstützen.

Nicht zuletzt führten unsere Untersuchungen zur erfolgreichen Generierung zuverlässiger Datensätze, die literarkonform zu bisher veröffentlichten Labordaten sind und möglicherweise sogar den Grundstein für Lern- oder Trainingsdatensätze legen können, mit welchen weitere computergestützte Untersuchungen an ähnlichen Systemen wie sie hier betrachtet wurden möglich sind.



TABLE OF CONTENTS

1.Introduction	1
2.Motivation	10
3.G protein inhibition	12
3.1. G proteins and GPCRs	12
3.1.1. G protein activation	15
3.2. YM and FR – the ultimate G _q inhibitors	17
3.3. In silico studies on G _q inhibition by YM, FR and analogues	20
3.3.1. Conclusions and future perspectives	33
4.Investigations on the mechanism of ion channel block by conesnail toxins	35
4.1. Voltage-gated ion channels	35
4.2. Subtype-selective block of ion channels by conotoxins	37
4.3. In silico binding studies of μ -PIIIA isomers on the sodium channel Nav1.4	41
4.3.1. Conclusions and future perspectives	56
4.4. In silico analysis of the subtype-selective block of KCNA ion channels through the μ -conotoxins PIIIA, SIIIA and GIIIA	57
4.4.1. Homology modelling	59
4.4.2. HADDOCK docking	59
4.4.3. Toxin dynamics and cluster analysis	60
4.4.4. Conclusions and future perspectives	82
5.Methods section	86
5.1. G _q – YM / FR (analogues) docking protocol	86
5.1.1. In silico generation of YM-254890 and FR900359 analogues	87
5.1.2. VINA Docking and energy minimizations	87
5.1.3. MD simulations of selected YM/FR analogues	87
5.2. G _q – YM / FR supplementary docking data	88
5.3. Homology modelling of Nav1.4	91
5.4. μ -PIIIA Nav Docking	92
5.5. Generation of K _v models and chimera	92
5.6. μ -PIIIA, μ -SIIIA and μ -GIIIA docking on K _v 1 channel subtypes	93
5.7. μ -PIIIA, μ -SIIIA and μ -GIIIA simulations on K _v 1 channel subtypes	94
5.8. Combined clustering for structure selection	95
5.8.1. Clustering method 1 – cluster close to the average structure	96
5.8.2. Clustering method 2 – graph-based cluster search	97
5.8.3. Selection of the final representative	99
6.List of publications	i
7.List of references	ii



1. Introduction

Understanding chemical reactions by illustrating their underlying mechanisms on an atomistic level depicts an essential step towards giving an explanation for distinct and measurable observations generated by laboratory experiments. The fact that such dynamic interplays are not visible to the naked eye to date motivates numerous researchers in natural sciences and engineering to develop computer-based representations of biochemical systems. Virtual investigations and the behavioural prediction of such systems are of high interest, especially when trying to explain or substantiate results gained from experimental studies. Particularly in today's new development possibilities for the prevention and therapy of diseases, diverse alternative but also cooperative perspectives with experiments on living animals can arise. This certainly affects not only ethical but also a number of economically important issues.

Intrinsically, a chemical reaction or process depicts a time-dependent mechanism. Hence, time parameters are to be taken into account to be able to provide accurate descriptions of aforesaid processes *in silico*, besides the inclusion of the respective *in vitro* conditions.

A prominent example of application is the deployment of computer-assisted processes in the pharmaceutical research area, where the investigation of drug-target interactions, their dynamics as well as of the resulting *structure activity relationships* (SAR) plays a key role. In these contexts, chemical properties of the agents that make up their subtype-specific binding towards the respective target molecules are important to be deciphered. As a huge variety for the rational composition of more complex molecular structures is naturally provided, numerous tests are required for the assessment of specific impacts on a reaction that may result from particular modifications on the molecule. Thus, it can be determined for instance which atom groups or amino acid residues should definitely be present in a considered agent molecule, intensifying or maintaining its binding affinity towards the target, and which ones that weaken the binding effect should equally be avoided.

Nowadays, there are many ways to address the challenging discovery of novel active substances, for instance by virtually or experimentally mimicking molecules that naturally bind to the target due to their intrinsic subtype specificity. It is obvious in these contexts that the researcher can learn from this nature-provided capacity when trying to revise or optimize particular agents. This approach is also known as *ligand-based drug design*.^{8,9}

A further method is represented by the highly automated testing of vast amounts of molecules from molecular libraries containing different substances which may bind to the target, also known as *high-throughput screening* (HTS).¹⁰ This technique can help for the fast and targeted identification of active biological agents among large molecular datasets.

The third and in this work primarily applied approach depicts the so-called *virtual screening* (VS): here, the target molecule and the inspected ligand substances are artificially recreated as 3D objects and the corresponding interactions are simulated by computers.^{11, 12} Especially in this scope of applications the three-dimensional structural information of the investigated molecules is of crucial importance, because generally speaking most of these computational prediction approaches rely on the known spatial configuration of the input molecules. This explains why these methods are also often referred to as *rational structure-based drug design* approaches.^{9, 13}

One of the best-known and to date widely applied VS techniques is represented by *molecular docking*. The main idea in solving the docking problem is to determine the optimal positioning and the associated mutual orientation of the target and the ligand structure as a bound complex, i.e. in a preferably adopted state of the system's global energy minimum.^{14, 15} Molecular or macromolecular interacting complexes in these contexts can be represented by protein-protein-, protein-ligand-, DNA- or RNA-complexes.¹⁶ – The associated terminologies are designated accordingly, as protein-protein-docking, protein-ligand-docking and docking of DNA- or RNA.

Since then, different approaches for solving the docking problem and thus for predicting preferred target ligand constellations have been established. "DOCK", for example, approaches this optimization problem by clique-search in a graph generated from the input data, which is the first docking approach not based on energy minimization calculations.¹⁷ Another method used by "AutoDock" is based on the principles of *simulated annealing*, a heuristic approximation method for solving optimization problems.¹⁸ "GOLD" is based on an inherent implementation of a genetic algorithm¹⁹ and "FlexX"'s particularly fast procedure predicts the optimal position of the ligand through its incremental reconstruction.^{20, 21} Obviously, the possible approaches can thus be of distinct variability.

The predicted mutual positioning, a so-called *mode of binding* usually holds the corresponding estimated value of the resulting *binding free energy* of the system, depending on the applied prediction method. As the complete molecular structural information is quite complex for computers to handle and requires huge computing capacities, modern search algorithms presently applied implement computational approximations, where only the surface coordinates of the involved molecules is taken into account.^{22, 23} Thus, the huge amount of complex data is massively reduced though having ensured to keep the critical information of the 3D objects. According to the employed algorithm, the related binding free energy of a certain complex mode is estimated by an optionally included *scoring function* that is executed on the previously docked complex.

Both, the docking algorithm as well as the scoring function nowadays hold various employable approaches. The simplest and most straightforward way to address the docking problem is to keep the considered molecules as rigid participants, based on the assumption that the spatial shapes of the molecules are maintained as they are provided by the pure input structures and do not change when interacting with one another. This approach is also known as *rigid docking*.

In contrast to rigid docking methods there are also established *flexible docking* approaches, where both or one of the considered molecules are computationally modelled including movement-related parameters. Allowed variations in bond lengths, torsion or bond angles ultimately affects the overall molecular geometry.

Conformational changes during a binding event are for instance of vital importance in *induced-fit* binding systems.²⁴

Obviously, the complexity of rigid but especially of flexible docking increases in proportion to the size of the considered molecules or more precisely of the considered molecular surface areas. For this reason, flexible docking approaches are predominantly applied in protein-ligand docking issues where a small, “drug-like” ligand-molecule is “flexibly” docked to a larger protein.

However, meanwhile there also exist more recent approaches that are even able to include the flexibility of the larger protein target structures. One of those methods is represented by the so-called *ensemble docking* strategy. Here, the ligand can be simultaneously docked to an *ensemble* of multiple target protein structures and the protein structure of highest fit to the ligand can be found (by optimizing operations on the structural parameters of the ligand and protein ensemble).²⁵

Nowadays, there are numerous implemented docking variations available – be they offered as web services (e.g. Haddock) or also as part of installable software. The same holds true for the variety of the applied algorithms.

However, the previous explanations indicate the fundamental significance of flexibility information in a considered system. Concerning the ensemble docking approach just mentioned, the protein structure ensemble can also be obtained by molecular dynamics simulation.²⁶ It is plausible and obvious, that not only rigid states of the system under consideration are of high interest, but also the dynamic development of the complex resulting from its varying inter- and intramolecular interactions. Especially for the description of the activity and thus the binding capacity of a ligand molecule, the dynamic information of the complex gives far-reaching statements about the ligand’s binding strength: while weak bonds result in high complex dynamics, strong chemical bonds allow the system to remain more rigid (for longer periods of time) in an energetically favourable state. Dynamic information thus does not only result in specific binding modes, but also in complete modes of motion of a considered system, which should aimed to be modelled.

This is where docking and experimental approaches reach their limits in terms of their ability to describe and illustrate a system under consideration, and computer simulations must be used to do justice to the real conditions as far as possible.

The methodologies of *molecular dynamics* (MD) are capable to master the problem by iterative motion calculations of the investigated system. The spatial movements of atoms and molecules in turn result from the particular interactions that occur in each time step. Consequently, the next state results recursively from the respective previous one.

In a preceding artificial equilibration, the *steepest descent minimization*, the complete system is first brought into the state of thermal equilibrium and thus the initial position of each particle is determined for the further simulated movement.

In the actual simulation that follows, *Newton’s equation of motion* is applied iteratively to the position vector of each particle in the system and thus its (subsequent) position is redetermined per time step.

In general the equation of motion for a particle of mass m is given by

$$\vec{F} = m \cdot \vec{a}$$

and with proper arguments applied

$$\vec{F}(\vec{r}(t)) = m \cdot \vec{a}(t),$$

where

\vec{F} denotes the force that acts on the particle and is made up by the sum $\sum_i \vec{F}_i$ of all inner and outer forces \vec{F}_i that affect the particle,

\vec{r} designates the position vector of the particle,

t denotes the time and

\vec{a} the motion acceleration of the particle moved.

Newton's equation of motion thus depicts a descriptive formula that is able to delineate the development process of a physical system on a spatial and temporal level, taking into account the totality of all external and internal forces acting on it. Accordingly, a particle or mass point of mass m moves on a physical orbit $\vec{r}(t)$ under the influence of a *force field* \vec{F} .

Numerous current simulation programs provide different applicable force fields, which allow for additional geometry optimizations of very large macro- or biomolecules. Here, intramolecular geometry parameters such as bond angles, i.e. torsion and dihedral angles and bond lengths between atoms are included.

Generally, the description of the *force field* is given by the *potential energy* E of the whole system. In order to avoid misunderstandings regarding the use of a potential energy equation to describe a force field, it should be noted here that the force results directly from the equation as a derivative of the potential in relation to its position.²⁷ The force field thus directly results from the system's potential energy.

In general, for molecular and macromolecular systems, their total energy E_{total} results from the sum of the partial binding energies E_{bonded} and $E_{nonbonded}$, arising from (covalent) binding interactions and non-covalent interactions of the system, respectively:

$$E_{total} = E_{bonded} + E_{nonbonded}$$

Here, the total energy for covalent interactions (E_{bonded}) is calculated as the sum of the energies, respectively resulting from the bond lengths (E_{bond}) as well as from the bond angles (E_{angle}) and the torsion angles ($E_{dihedral}$) between atoms or atom groups:

$$E_{bonded} = E_{bond} + E_{angle} + E_{dihedral}$$

All electrostatic and non-covalent van der Waals forces are considered as non-binding interactions, yielding

$$E_{nonbonded} = E_{electrostatic} + E_{van\ der\ Waals}$$

to describe the non-binding energy $E_{nonbonded}$ that is calculated as the sum of energies arising from electrostatic interactions ($E_{electrostatic}$) and energies emerging from van der Waals forces ($E_{van\ der\ Waals}$).

In most standard force field programs *harmonic potentials* (derived from *Hooke's law*) are used to characterize the particular potential binding energies E_{bonded} , given by

$$E_{ab} = k_{ab}(r_{ab}^0 - r)^2$$

for a particular potential binding energy term of E_{bonded} between a pair of atoms a and b , where

k_{ab} defines the force constant,

r_{ab}^0 delineates the atomic pair distance regarding the considered parameter, assumed that bound a and b are in the resting position (an imaginary definite position and zero momentum) and

r is the actual (equilibrium) distance (regarding the parameter under consideration).

This approach relies on an assumption of “genuine” bond lengths and angles and can be used, for example, to describe the potential energy resulting from the bond lengths of a bonded pair of atoms and there are analogous description formulas also for the bond angles (torsion and dihedrals).

Using the example of the basic equation of the *AMBER* force field (**A**ssisted **M**odel **B**uilding with **E**nergy **R**efinement) used in one of our MD study, under which it was first published in 1995,²⁸ one can see the concrete application of the above-mentioned interrelationships and thus obtain the potential total energy E , also simply potential, of a system of N particles at particular positions r through:

$$\begin{aligned}
 E(r^N) = & \underbrace{\sum_{bonds} k_b (l - l_0)^2 + \sum_{angles} k_a (\theta - \theta_0)^2 + \sum_{torsions} \sum_n \frac{1}{2} E_n [1 + \cos(n\omega - \gamma)]}_{E_{bonded}} \\
 & + \underbrace{\sum_{j=1}^{N-1} \sum_{i=j+1}^N f_{ij} \left\{ \underbrace{\varepsilon_{ij} \left[\left(\frac{r_{0ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{r_{0ij}}{r_{ij}} \right)^6 \right]}_{E_{van\ der\ Waals} \text{ (Lennard - Jones potential)}} + \underbrace{\frac{q_i q_j}{4\pi \varepsilon_0 r_{ij}}}_{E_{electrostatics} \text{ (Coulomb's law)}} \right\}}_{E_{nonbonded}}
 \end{aligned}$$

28

where

k_a and k_b designate the particular force constants (k_a for the calculation of the energy resulting from the bond length distances and k_b for the energy calculation from the bond angle differences),

l determines the bond length of the bound atomic pair,

θ the bond angle difference between covalent bond pairs (i.e. bonds to two neighboring atoms) of the atom,

ω, γ torsion angles (dihedrals in a protein backbone),

r_{0ij} the equilibrium distance,

ε the well depth and

q the (atomic) charge (q_i the charge of atom i and q_j the charge of atom j accordingly).

It is well known that a system is in the most energetically favourable state if its (potential) energy is minimal. This principle also comes into play when a force field is applied to a molecule: if the sum of the partial energies describing a force field is kept to a minimum and distributed to the molecule in as many small components as possible, this results in a deformation of the molecule in favour of its energetic state.

To come back again to the solution of Newton's equation of motion, the key issue in molecular dynamics, our formula considered, $\vec{F} = m \cdot \vec{a}$, transforms into a complete system of differential equations, since usually many-particle systems are considered in a simulation. Since the computational effort to solve such a system of equations is generally much too high, *numerical approximation* methods must be used at this point to solve the equation of motion (in the form of a differential equation system).

There are several algorithms to master numerical approximation, such as the well-known and frequently used *Velocity Verlet algorithm*. This algorithm utilizes the principles of Taylor series development to approximately calculate an atomic position r_i of a molecule i at a preceding time $t - \delta t$ and at a subsequent time $t + \delta t$, respectively. According to the Taylor approximation by summing up the (derivative) polynomials (in the case of Velocity Verlet, the Taylor development is carried out up to the third degree), one obtains:

$$r_i(t + \delta t) = r_i(t) + \delta t \dot{r}_i(t) + \frac{1}{2} \delta t^2 \ddot{r}_i(t) + \frac{1}{6} \delta t^3 \dddot{r}_i(t) + O(4)$$

for the atomic position at the next point in time and accordingly

$$r_i(t - \delta t) = r_i(t) - \delta t \dot{r}_i(t) + \frac{1}{2} \delta t^2 \ddot{r}_i(t) - \frac{1}{6} \delta t^3 \dddot{r}_i(t) + O(4)$$

for its position at the previous point in time.

By summing up both of the above approximation equations one obtains

$$r_i(t + \delta t) = 2r_i(t) - r_i(t - \delta t) + \delta t^2 \ddot{r}_i(t) + O(4) ,$$

which greatly simplifies the equation by eliminating the third-degree terms. Thus the accuracy is maintained by approximating the motion up to the Taylor development of the third degree, however, without having to consider the corresponding terms again. Thus, the equation of motion can be successively and approximately integrated without serious loss of precision.

The recalculation of such an approximation algorithm (iteration per time step) takes place as long as it is considered reasonable by the user – in the case of macromolecular complexes, the simulation time usually amounts to nano- or even microsecond ranges.

During such an approximation in discrete time steps, a so-called *statistical ensemble* consisting of different possible configuration states of the considered system is generated (under artificially set physical conditions). This ensemble ultimately forms a real subset of the complete system *phase space*, i.e. the set of all possible system states.

Depending on the setting of the external physical parameters, there are different types of ensembles producible in which certain physical quantities are kept constant by respective controlling calculations (while the other parameters have more degrees of freedom).

So for example the *NVE* ensemble (constant-energy, constant-volume ensemble), also known as the *microcanonical ensemble*, which is characterized by the maintenance of constant energy and volume during the solving process of Newton's equation. Here, no pressure or timing parameters are considered in the approximation run to guarantee energy and volume consistency.

Analogously, *NVT* ensembles (constant-temperature, constant-volume ensemble) or *canonical ensembles* provide constant volume and temperature during the simulation. The maintenance of the constant temperature despite "disturbing" external influences is provided by a so-called *heat bath* or *thermostat*.²⁹ The NVT ensemble depicts the most commonly used default-set ensemble in numerous simulation programs.

Also worth mentioning is the *NPT* ensemble (constant-temperature, constant-pressure ensemble), where, in addition to temperature controls, pressure consistency is obviously ensured by volume adjustments, as well as the *NST* ensemble (constant-temperature, constant-stress ensemble) that further allows for an extended pressure control by regulating its particular dimension components.

Finally, the *NPH* ensemble (constant-pressure, constant-enthalpy ensemble), in reverse manner to the NVE ensemble, preserves the stability of pressure by volume adjustments and thus also the persistence of enthalpy.³⁰ Due to technical limitations, only a "restricted" part of reality can be represented in the simulation, so a system of finite size has to be chosen. From this system it is in turn possible to draw conclusions about larger units from reality.

So-called *periodic boundary conditions* are used for this purpose, the intention of which is to describe a system as an infinite number of equivalent finite systems.³⁰ The objects to be modelled are positioned in a *simulation cell*, which represents a finite volume. The simulation cell, in turn, is replicated in the entire simulation space, which concurrently ensures that each system property is represented equally in all replicas.³⁰

Finally, the successive numerical solution of the equation of motion yields a time-dependent *trajectory* within the phase space, which describes, so to speak, a motion path through this phase space that steadily traverses the calculated configurations at the various time steps. Therefore, the resulting trajectories can also be understood as emerging molecular configurations (ensembles) as a function of time.

From the trajectories, the movement paths of the individual atoms through the phase space and thus the overall behavior of a molecule can ultimately be deduced – for example, by reducing the determined overall movement trajectory of a certain system to definite movement modes. The motions can also be analyzed on a number basis, e.g. as changes in the distances of atomic positions of specific (C α -backbone) atoms over time or as changes in the (free) binding energy or enthalpy of the overall system as a function of time. In addition, the trajectories may also provide information on the overall flexibility of the system due to variations in fluctuation values.

All in all, the ultimate goal is to obtain an overall understanding of the essential macroscopic properties of the system under consideration by analysis of the individual atom trajectories. In addition, comparative simulation series should help to gain an extended comprehension of different (macroscopic) system behaviours in correlation with the differences in the respective system compositions.

Two invaluable approaches for the analysis of simulation-generated trajectories constitute the so-called *Principal Component Analysis* (PCA) and the *Normal Mode Analysis* (NMA). Both techniques are able to determine definite or discrete motion modes of any system of consideration. Therefore they are also suitable for the analysis of motion data, such as conformational changes of more complex biomolecular systems from md-generated trajectories.³¹

The basic principle of PCA is to reduce the dimensionality of the space containing the data set of the system-describing properties by a resulting (data) space of lower dimension that should give an equally appropriate picture of the system on the basis of the reduced set of properties. This means that only so much information may be reduced, so that no essential characteristics of the system are lost, or simply: keep as much data as necessary, delete as little data as possible.

Detected redundancy in the data is eliminated by a summarizing operation on correlating data points.

The set of remaining, linearly uncorrelated property variables are designated as principal components. Each component must be orthogonal to all others, so the resulting vectors describing the properties end up representing an orthonormal base set.

In the case of trajectory analysis from an MD simulation, the set or ensemble of constellations traversed by the trajectory describes the full set of system properties that can then be reduced to the essential characteristics of the system using a PCA. For example, different conformations of a protein (e.g. change of the mutual position of its subdomains) can be unraveled and their transition states can be derived.

According to methodologies of *Normal Mode Analysis*, it is assumed that the system under consideration oscillates back and forth. On the assumption that these movements of the system occur uniformly, i.e. in a consistent manner, the term "*mode*" means a sinusoidal harmonic form of motion. Here, the individual movements of the particular system components are *independent* of each other and are then referred to as "*normal*", but are however all carried out *with the same frequency* in a certain form of movement (mode).

An *eigen* or *normal mode* hence designates a certain type of natural vibration movement of a system in which all its parts move together at the *same frequency*.^{32, 33} Thus, each normal mode has its own characteristic frequency, also called *eigenfrequency* or *eigenvalue*.

All possible observable configurations of a system can be derived from its normal modes.

As each movement mode of the system can be described by an equation of motion, for all movement modes of the system an equation system results from the individual equations of motion. The *independent normal modes* are

thus derivable from the *eigenvectors* of this system of equations, where their respective *eigenvalues* describe the particular vibrational frequencies, also *eigenfrequencies*, of the associated normal modes.

In the usual procedure of a normal mode analysis for a protein, the *Hessian matrix* is first calculated after energy minimization and divided by the mass matrix (since $\vec{F} = m \cdot \vec{a}$ and one wants to obtain the values of the accelerations). By diagonalizing the resulting matrix one then obtains its eigenvectors, whose eigenvalues, as already mentioned, in turn indicate the oscillation frequencies. The eigenvectors ultimately indicate the different directions of motion, which can then be visualized again.

From the above statements it follows that a high eigenvalue of an eigenvector indicates a low force constant of the respective normal mode, i.e. a low frequency. Since low frequencies are particularly relevant for the analysis of protein movements, *low normal mode analysis* is often invoked for the determination of protein motion modes. Finally, this form of analysis can provide information on the flexibility of the objects considered in the system. In summary, it can be stated that the PCA concentrates more on determining *representative states* of the moving system, while the NMA focuses more on uncovering *representative motion modes* (forms of movements) of the system and thus drawing conclusions about its intrinsic flexibility.

It is thus apparent that the large amount of data generated by the simulation represents new challenges for their downstream data mining processes, for example in their automated handling or in their statistical evaluation. However, large and reliable simulation data sets may also offer further basic support in general machine learning or corresponding methods of classification, even up to their use as learning data sets (e.g. for the problem of protein structure prediction). Learning data sets for predicting the motion behaviour of specific (not yet tested) systems could also be generated from carefully prepared and evaluated simulation series.

The large amount of simulation data could further represent a useful completion for molecule databases and thus for the entire virtual screening area. This is especially true in view of the fact that different (energetically favourable) states of a molecule can be determined by simulation, which may not yet be available in the databases.

In addition to the numerous analysis possibilities of the results, there are today all the more possibilities for the actual simulation management, which - based on variations of the above-mentioned basic principles - offer differentiated and more sophisticated simulation alternatives.

The above mentioned combinations of defined parameters and the resulting different ensembles allow a wide range of (meanwhile very established) simulation methods. The algorithms used are also constantly revised and optimized. To name a few simulation variants at this point, the best-known ones are briefly listed below.

Ab initio molecular dynamics are primarily used in the field of quantum chemistry. As their underlying principle depicts the approximative solution of the Schrodinger equation³⁴ and the only inputs designate physical constants³⁵ the computational costs are very high so they are only able to treat systems of smaller sizes. As a prescribed potential such as used in classical MD (as described above) is not used in this case, we will not get into further detail at this point, since there are no direct comparison parameters for both methods.

Further, modern simulation programs provide the ability to perform *membrane simulations*, i.e. to simulate the system of consideration within a membrane (of different possible compositions), if desired.

Steered Molecular Dynamics (SMD) allows the use of an additional, external artificial force applied to selected system atoms (also acceleration),³⁶ which in total can lead to an acceleration of the executed simulation, i.e. to faster recognizable effects. The latter is also possible, for example, using higher simulation temperatures.

The *Replica Exchange Molecular Dynamics Simulation* (Replica Exchange MD, REMD) also makes use of variations in the simulation temperatures:^{37, 38} here, several simulation runs are executed in parallel at different temperatures, the algorithm that searches for the global energy minimum state of the system, can more easily overcome local minima when traversing the (thus extended) energy landscape. As the overall search space is expanded, the global minimum state of the system can thereby be reached more likely.

Numerous software (open source and commercial) have already been developed over the past decades which offer different versions of implemented MD applications. A few examples to be mentioned are the software GROMACS,³⁹ LAMMPS,⁴⁰ NAMD,⁴¹ ORAC,⁴² Schrodinger Maestro⁴³ and the YASARA molecular modelling software,⁴⁴ to name but a few. The latter has been used for the studies conducted in this work.

These explanatory notes demonstrate that the field of molecular dynamics is very broad, based on the fundamental principles. New simulation approaches and associated algorithms are still being realized today while the development process is subject to continuous optimizations.

The computational techniques mentioned and their stated extensions however still depict a great time- and cost-saving support for the wet laboratory tests in the *pre-clinical* development phase of active substances (**Figure 1**).

Findings from the associated basic research equally form important pillars for related test and development questions. Therefore, computer-aided methods such as docking and molecular dynamics simulation are very often used today in the development of novel drugs, even in combination. – Here, it is particularly important to be able to assess and evaluate the effects of modifications in a biological system (such as target mutations or modifications of the binding ligand molecule) in advance or in cooperation with the ubiquitously established methods from the wet laboratory. However, as such a modification may bring along an improvement in one molecular property but however also a coincident deterioration in another one, the computer-generated results must also be repeatedly verified by experimental laboratory tests. Despite more cost-effective and faster than conventional wet laboratory methods, not all effects of a specific modification can be registered and assessed via the methodologies offered by computer-based mimicking. Animal experiments are always necessary, for example to determine the length of time a newly created substance remains in the body. In case of a too rapid degradation by the body's own enzymes, the corresponding, degradation-sensitive site in the molecule must be found, modified and re-tested. This stage of development is also referred to as *lead optimization*.

For the reasons stated, the optimization path can end up in a very protracted and time-consuming up and down process and the complete pre-clinical test phase can take up to 12 years until a suitable candidate can ultimately be approved for further clinical studies (**Figure 1**). This fact requires a very close and constructive cooperation between interdisciplinary sciences, on the part of the life sciences as well as the computer sciences.

Figure 1 shows that the entire test phases (pre-clinical and clinical) can take a total of 11 – 16 years before a drug is actually approved for the market.

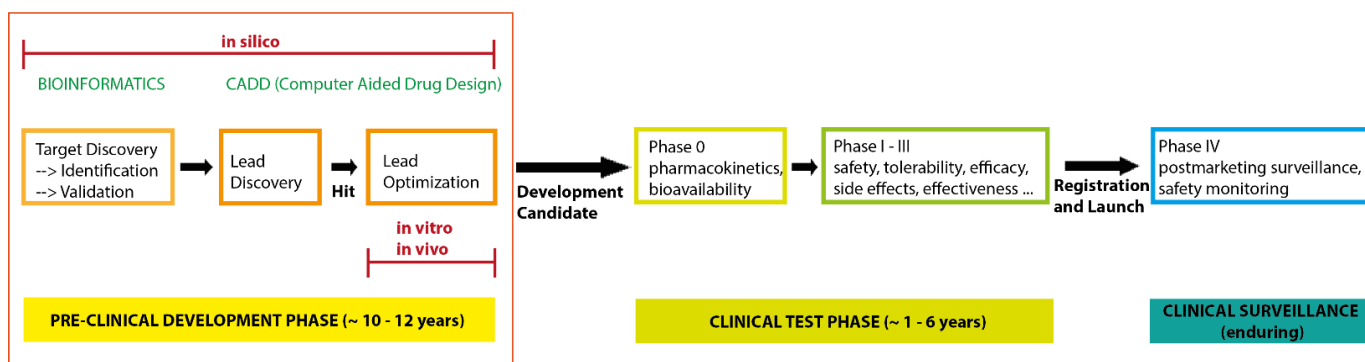


Figure 1: Main stages in drug development; the field of application on which this work focuses is bordered in orange.

During the development phase, particularly during lead optimization, there may even be cases where it can be concluded that the naturally existing lead structure (or a *hit*) already represents the molecule of the highest binding quality (**Figure 1**). This was the case with the study by Reher et al.⁵ which is closely linked to chapters 3.2 and 3.3 of this thesis.

Phase III in particular takes up most of the time in the clinical test phase, in which several hundred to thousand patients (300 – 3000 according to current status) are involved in order to test core aspects with regard to drug efficacy and safety. This phase can last from 1 to 4 years (**Figure 1**). Even when the medication is finally on the market, ongoing controls and surveillances are still necessary (**Figure 1**) – analogously to the continuous optimization efforts regarding the simulation algorithms.

In spite of permanent revision, computer-based methods cannot replace laboratory experiments until today. However, since new computer technologies are constantly evolving and are even approaching this challenges⁴⁵, there is still hope that animal trials will one day be completely superseded.

Moreover, although such a mutual development process between natural and computer sciences may take a long time, *in silico* methods have become an integral part of modern life sciences. Since the decoding of the human genome in 200 and the concomitant high amount of biological data, computer-based use has gained significantly in importance in the life science areas. The application in the natural science fields reach beyond the topics covered in this thesis and ranges from data mining and management issues of gene data to applications in the medical imaging branch.

Overall, *virtual reality* (VR) and simulation techniques have so far been able to achieve economic benefits as well as time and cost savings in many other areas of academic and industrial research. Due to their interdisciplinarity and their transferability of applications to all possible fields they are still successfully in use in various branches: For instance, simulation or *computer-assisted design* (CAD) techniques are used in many test phases for design revision and optimization purposes in the development and manufacturing of diverse industrial products. Further examples of VR applications include, for example, finite element analysis or technical applications such as flow and circuit simulation, building design, simulating production processes, aircraft simulators (or flight simulators for pilot trainings) or even game simulations. Areas of application range up to interfaces with topics of artificial intelligence or image recognition. These overarching advancements could be achieved thanks to the predictive power and the concomitant and error avoidances up to optimizations of entire business processes. All these facts indicate that the broad field of scientific computing will continue to evolve and be applied in all areas of the physical, natural and medical sciences. Here, however, the laws of physics will always lay the essential foundations for concrete applicability.

2. Motivation

Understanding the mechanism of how potential drug candidates or drugs achieve their effects at their molecular target is of vital importance for a guided process to develop a promising drug candidate into a valuable medicine. Nowadays, such understanding is mainly gained from computational studies, better known as molecular modelling approaches, which are an essential part of every drug development campaign.

However, the present work focuses on understanding the mechanisms of action of two pharmaceutically relevant inhibitor protein systems.

First, the mechanism of inhibition of the heterotrimeric G protein (G_q) through the small cyclic depsipeptides YM-254890 (YM) and FR900359 (FR) shall be elucidated in order to guide rational-design based approaches for the development of more potent inhibitors and inhibitors with a different specificity regarding other G protein subtypes.

Like GPCRs, which represent a major drug target class comprising the largest family of membrane protein receptors,⁴⁶ the downstream G proteins represent promising drug targets.⁴⁷ However, due to the lack of selective G protein modulators, their role in disease-related biology remains unclear. Selective G protein modulation will certainly shed light on their role in GPCR signalling.⁴⁸ To date only a very few G protein modulators have been reported in the literature.^{1-3, 49-51} Among them, two natural products YM-254890 and FR900359, which serve as potent modulators for $G_{q/11}$ -induced signalling to, e.g. target oncogenic $G_{q/11}$ proteins as a potential therapeutic option for cancer treatment of the adult eye.¹⁻⁴

As a second focus, this work intends to shed light on the question how peptide-based toxins, which are derived from the venom of marine cone snails block the ion flow through potassium and sodium channels and how these toxins achieve their great subtype specificity.

Very recently, drug discovery research identified the voltage gated sodium channel subtype $Nav1.7$ as a major target for pain treatment.⁵² As venom-derived peptides are known to be very specific ion channel blockers, they *per se* represent promising drug candidates. However, in order to develop these kinds of compounds into a promising painkiller, it is important to understand how these peptides interact with their targets and how they achieve their subtype specificity.⁵³

With respect to the first major focus of the present thesis, the group of Stromgaard recently succeeded in developing the total synthesis of the two aforementioned natural products YM and FR and synthesized a set of approx. 40 analogues and experimentally determined their binding affinities towards G_q . This set of compounds will serve as a starting point for deciphering the specific features of recognition of the peptide-based ligands at their target site.

As a first step for the anticipated structure relationship study, structural models for these analogues need to be built and a reliable docking protocol needs to be developed in order to investigate the ligand protein interactions in detail. Additionally, the structural integrity (conformational stability) of the inhibitor analogues should be verified through molecular dynamics-based approaches.

Lastly, a pharmacophore model for this type of G protein inhibitors shall be created and be visualized.

The second focus of this thesis, first aimed to understand how the disulfide bond pattern of the μ -conotoxin PIIIA alters the ability to block the voltage gated sodium channel $Nav1.4$. As μ -PIIIA harbors six cysteine residues, it can form 15 disulfide bond isomers. These isomers were recently synthesized by the Imhof group (University of Bonn) and their disulfide bond pattern was verified through a combined approach of mass spectrometry and NMR spectroscopy. Furthermore, their ability to block the $Nav1.4$ was accessed through electrophysiological experiments by the Heinemann lab (University of Jena) and their three-dimensional structures were elucidated (Tietze lab, TU Darmstadt)⁵⁴ and will now serve as the starting point for a comparative docking study perusing to correlate their structure and bioactivity. As there is no crystal structure available, a suitable molecular model of $Nav1.4$ must be developed. In a second step a reliable docking protocol will be established and validated with respect to the bioactivity of the μ -PIIIA isomers.

As μ -PIIIA shows some inhibitory activity towards $Nav1.7$ it was considered for having a potential as analgesics as long as it is specific.^{55, 56} Unlike, the conotoxin μ -GIIIA, which exclusively blocks skeletal muscle voltage-gated sodium channel $Nav1.4$, the conopeptides μ -PIIIA and μ -SIIIA additionally inhibit the neuronal sodium channels $Nav1.2$ but were recently shown to be not exclusively specific for Nav channels.^{7, 57-59} As μ -SIIIA only

partially inhibited $K_v1.1$ and $K_v1.6$ μ -PIIIA blocked both channels, but were both inactive on the subtypes $K_v1.2$ to $K_v1.5$ and $K_v2.1$.⁷

Consequently, the last part of this thesis sought to unveil the mechanism of subtype-specific inhibition of potassium channels by μ -conotoxins. Based on the experimental data published by Leipold et al.,⁷ docking and subsequent MD simulation experiments will be performed, which are intended to illustrate the dynamic interplay between μ -conotoxins PIIIA, SIIIA and GIIIA at the potassium channels $K_v1.1$, $K_v1.5$, $K_v1.6$ and two chimeric channel constructs. For reasons of comparison, μ -SIIIA was investigated as a partially channel blocking semi-active system while μ -GIIIA as an inactive candidate. In particular, the different pore blocking modes – full or partial pore coverage – will be of special interest intending to achieve more insights about the origin of remaining currents, especially observable for the partly pore blocking toxins.

As for the previous systems, a molecular model for the potassium ion channels will be established first employing the crystal structure of $K_v1.2$ - $K_v2.1$ paddle chimera channel.⁶⁰ The NMR structures of μ -PIIIA, μ -SIIIA and μ -GIIIA will be used for the docking and simulation approach.

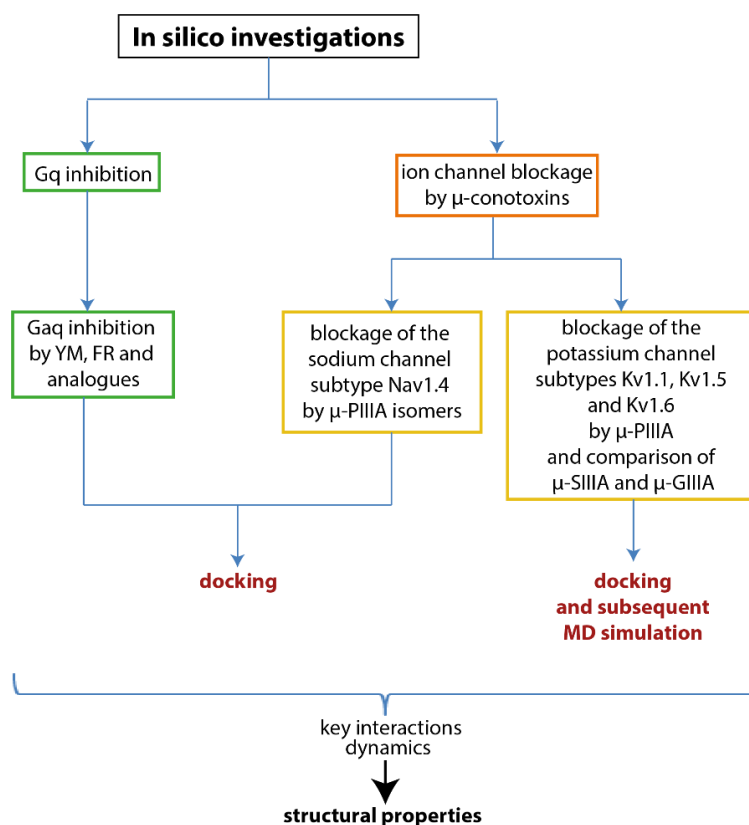


Figure 2: Overview of in silico studies performed in this work; the aim of the studies was to investigate two different systems in terms of detecting key structural and dynamic properties in their interacting mechanisms.

3. G protein inhibition

3.1. G proteins and GPCRs

G proteins, guanosine nucleotide binding proteins undergo a close interplay with their G protein coupled receptors (GPCRs) since their mechanism of action is directly initiated by GPCRs. Due to their close cooperation, they both play an equally important role in many cell signalling processes, a large number of which is involved in the development and growth of cells. Such energy consuming processes of signalling are subject to the exchange of the G protein's nucleotide GDP by GTP which, in simple terms, is triggered by the activation of G proteins through GPCRs.

GPCRs themselves are heptahelical receptor proteins, meaning that they are constituted by seven transmembrane helices (TMs), wherein each transmembrane domain is composed of an about 20 amino acid long α -helical segment.⁶¹ The helical strands are in turn connected by three alternating intracellular and three extracellular loops.⁶¹ As the name of their bundle of transmembrane helices suggests, GPCRs are located in the cell membrane environment and therefrom they mediate numerous receptor-dependent downstream signalling pathways, being for instance responsible for the initiation of olfaction, taste, vision, slow neurotransmission or cell division processes.^{62, 63}

The activation of the majority of these cellular physiological responses in our living systems is due to their stimulation by external signals from heterogeneous ligands that bind to the receptors or by environmental stimulators.⁶² In particular these can be endogenous ligands such as hormones, drugs, odors, neurotransmitters, pheromones, photons, growth factors, ions or even proteases.^{62, 63}

The spectrum of activators is thus visibly large – no wonder, because the GPCRs constitute the largest and most diverse superfamily of membrane receptors represented in every eukaryotic cell, comprising more than 1000 different members, encoded by just as many genes.^{62, 63} Further, it has been reported that the human and mouse GPCR repertoire for endogenous ligands comprises around 400 receptors.⁶⁴ Due to their obviously huge variety, GPCRs are important targets for drugs and are therefore the subject of great research interest.^{64, 65}

Vertebrate GPCRs can be further categorized into six main classes on the basis of their sequence homology and their structural or functional similarity: Here, the rhodopsin-like receptor family constitutes Class A (GPCRA), secretin receptors represent Class B (GPCRB), the metabotropic glutamate or pheromone receptors rank among Class C (GPCRC), fungal mating pheromone receptors represent Class D (GPCRD), cAMP receptors define Class E (GPCRE) and finally frizzled/smoothed receptors constitute Class F(GPCRF).⁶⁶⁻⁷³

Among all classes, the rhodopsin family depicts by far the largest and most diverse family, including 683 members in humans^{62, 74} and was the first GPCR whose crystal structure was determined to high resolution.⁷⁵ Since then, further progresses concerning the understanding and functional characterization of GPCRs have been achieved, meanwhile providing high-resolution structural studies and insights into the molecular mechanisms for instance of ligand-activated GPCRs, such as the human β 2 adrenergic receptor (β 2AR), the avian β 1AR and the human A2A adenosine receptor, or of opsin and an active form of rhodopsin.⁶²

Despite their structural uniformity, GPCRs bear individual combinations of signal transduction activities tailored to their multiplicity of addressed G-protein subtypes.⁶² Therefore, they can be classified alternatively by the properties of the bound G protein. The classification is then based on the equally diverse G proteins, namely into G_s -coupled receptors triggering a stimulating cascade, G_i -coupled receptors that induce an inhibitory cascade, G_q -coupled receptors, which activate protein kinase C via the second messenger pathway (phospholipase C mediated) via inositol trisphosphate (IP3) and diacylglycerol (DAG)⁷⁶ and lastly G_t -coupled receptors that activate c-GMP-dependent phosphodiesterase in rod cells via the α subunit of the G protein, called transducin.⁷⁷

Despite suggested by the huge variety of possible activation pathways by GPCRs just outlined, they exhibit structural uniformities, as do the GPCRs described above. Likewise, those structural features make a significant contribution to the basic function of G proteins. Generally, G proteins are heterotrimeric proteins, i.e. they consist of three different separate subunits α , β and γ .⁷⁸ The α subunit has a GTPase-like function and contains a GDP/GTP-binding site, where the GDP-bound form constitutes the inactivated form of the G protein. Here the α subunit stays conjugated with the subunits $\beta\gamma$.⁷⁹ The exchange of GDP with GTP, which is an important

part of any signal transduction cascade, takes place at this protein domain. More precisely, the GTPase function of the G α subunit is fulfilled by its partial “Ras-like” domain, where the second part is made up by a unique α -helical domain with five helices.^{79, 80} As the names imply, the Ras domain in contrast to the purely helical domain, consists of mixed β -sheet and α -helical segments and is conserved in all members of the GTPase superfamily.⁸¹ Also, the α -helical domain is unique to heterotrimeric G-proteins.⁸¹ As will be outlined later in more detail, the G α subunit is considered to be of central importance with regard to the activation state of the G protein. In this respect, it has been shown, that especially the helical domain of the G α subunit enhances the efficiency of G protein signalling by serving as an intrinsic GTPase-activating protein.^{80, 81} More specifically, it showed accelerating effects on GTP hydrolysis⁸¹ and it is reported that changes in the helical domain, even far from the nucleotide-binding pocket, may profoundly affect the nucleotide-binding function of the G protein in general.⁸² Due to the importance of the G α subunit outlined before, the determination of the G protein subtypes is mostly determined on the basis of this subunit.

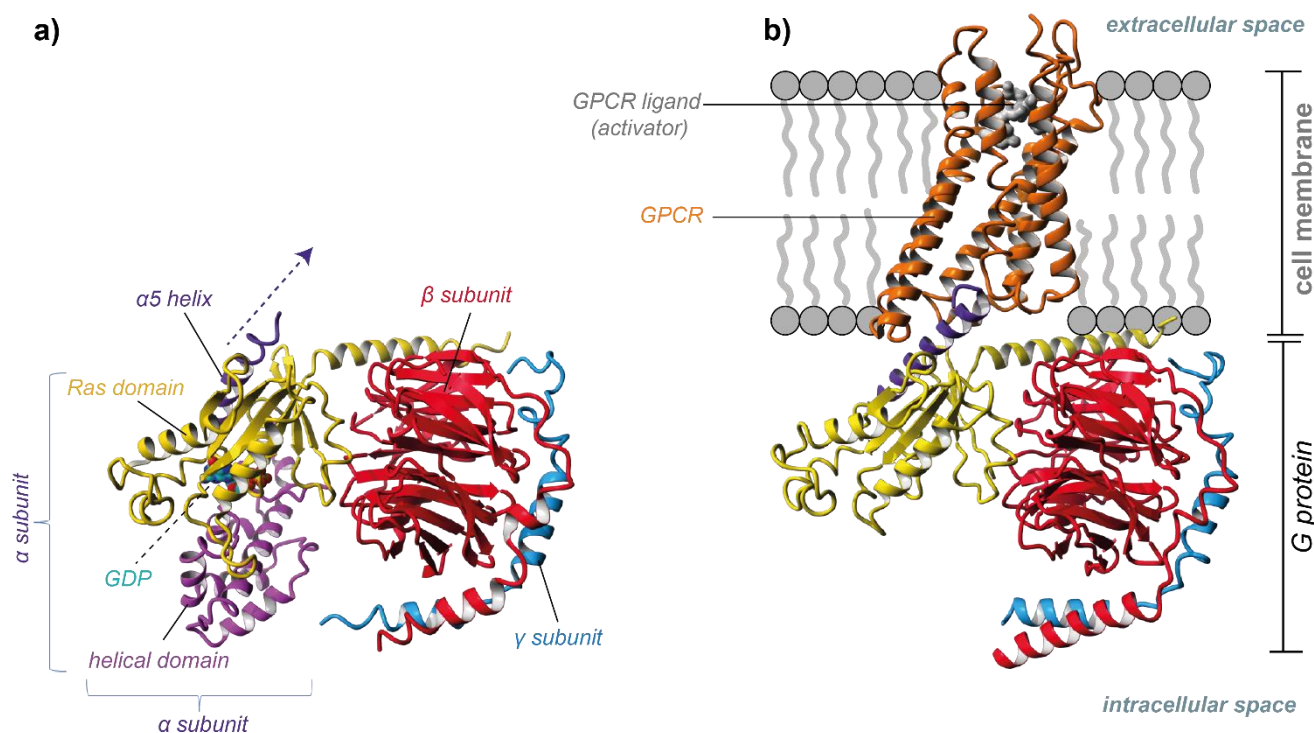


Figure 3: 3D representation of heterotrimeric G protein structures (a) including all subunits α , β and γ (PDB ID: 3AH8⁸¹) and (b) in complex with the membrane-bound G protein coupled receptor (GPCR) and detached helical domain (PDB ID: 6DDF⁸³). Colour scheme: G α subunit Ras domain – yellow, G α subunit $\alpha 5$ helix – deep purple, G α subunit helical domain – magenta, G β subunit – red, G γ subunit – blue, GPCR – orange, GPCR activator and cell membrane – gray.

In addition to the proteins G_s, G_i, G_q and G_t, there are other G protein subtypes such as G₁₂ and G₁₆. Their activation is subject to similar mechanisms, however, not their recognition reactions with regard to effector molecules.

All of these are similar in their sequence homology only to a limited extent, although there is a high degree of structural homology (as is the case with GPCRs) due to their equivalent functions for GDP/GTP exchange (Table 1). In terms of sequence similarity, G β and G γ subunits can vary even more than the G α subunits, further substantiating the importance and uniform function of the G α subunit.

Table 1: percentage of pairwise sequence identities resulting from pairwise alignments; alignments were calculated with the SIM Alignment tool for protein sequences⁸⁴ of the ExPASy Bioinformatics Resource Portal (<https://web.expasy.org/sim/>)

Organism	ID Uniprot		Family						
			G_i	G_s	G_q		G_{12}		G_{16}
			(α) subtype						
			$G\alpha_i$	$G\alpha_{s-i2}$	$G\alpha_q$	$G\alpha_{11}$	$G\alpha_{12}$	$G\alpha_{13}$	$G\alpha_{16}$ (GNA15)
Human	P63096	$G\alpha_i$	100,00%	45,30%	52,80%	51,70%	42,00%	39,90%	43,50%
Human	P63092-2	$G\alpha_{s-i2}$		100,00%	40,70%	40,90%	39,70%	37,40%	37,20%
Human	P50148	$G\alpha_q$			100,00%	90,30%	44,10%	45,70%	55,50%
Human	P29992	$G\alpha_{11}$				100,00%	43,10%	45,10%	55,50%
Human	Q03113	$G\alpha_{12}$					100,00%	66,70%	39,10%
Human	Q14344	$G\alpha_{13}$						100,00%	39,10%
Human	P30679	$G\alpha_{16}$							100,00%

Certainly, GPCRs depict attractive therapeutic targets for numerous diseases and still bear a great potential as such due to their high diversity and involvement in a broad spectrum of signalling pathways, as mentioned before. Not surprising that about one third of prescription drugs work by GPCR modulation.^{85, 86} Examples of drugs based on interaction with GPCRs include beta-blockers and a subgroup of thrombocyte aggregation inhibitors.^{87, 88} Notwithstanding, this represents a challenging endeavor concerning their specific and targeted addressing. However, in this thesis we will discuss an alternative method to prevent unwanted downstream signalling processes based on the targeting of their directly coupled G proteins by specific ligands.

Usually, GPCRs were preferred to their associated G proteins in drug development because it was hoped that their addressing would lead to the targeted manipulation of physiological processes controlled by individual receptors. However, such an approach can fail in more complex pathologies, for example when the malfunction in question is controlled by more than one receptor and its related signalling circuits. Such occurrences have been observed for several lung diseases, metabolic disorders as well as for certain types of pain and cancer.⁸⁹⁻⁹⁴

Since G proteins – due to the close interaction – are just as relevant in physiological events as their associated GPCRs, an equally high potential for them to represent drug targets can be inferred. It is not without reason that the discovery of G proteins was awarded the Nobel Prize for Medicine in 1994. As furthermore G proteins, compared to the GPCRs, represent the actuators of the subsequent stage within the overall signalling cascade, a potentially superior approach of more targeted downstream regulatory modulation per G protein inhibition is emerging at this point. For example, it could be shown that among the four families of the $G\alpha$ proteins $G\alpha_{i/o}$, $G\alpha_s$, $G\alpha_{q/11}$ and $G\alpha_{12/13}$ only members of the $G\alpha_{i/o}$ family can be specifically inhibited by pertussis toxin (PTX).^{95, 96} Such insights still serve today as a valuable basis for the analysis of GPCR signalling mechanisms and G protein-mediated cell reactions. The high pharmaceutical potential of G protein inhibitors is further underpinned by the crucial role that has been shown being played by GPCR-activated G proteins concerning tumor diseases.^{97, 98} Likewise, the controlled and specified inhibition and/or activation of G proteins is also motivated by their found contexts with apoptotic signalling pathways.⁹⁹

In all such approaches, however, it is always important to keep in mind that inhibitors of this type should have both cell permeability and selectivity for different G protein families in order to achieve potent efficacy.

All in all, both GPCRs and G proteins show that understanding and taking advantage of their subtype specificity is of great importance for their activation, functioning and their controlled inhibition.

3.1.1. G protein activation

The activation of G proteins by GPCRs depicts a fascinating and complex regulatory process that cyclically proceeds and can be broken down into several interdependent sub-steps. The propagation of the signal pulse is initiated by GPCR stimulation through extracellular actuators, such as the binding of an effector molecule to the GPCR. Upon this first event of activation a conformational change of the GPCR is triggered allowing for the transduction of the inducted stimulus to the G proteins which are attached to the inner membrane surface and thus to the GPCR. The G protein, still in its GDP-bound inactivated state, is also induced to undergo a massive conformational change, i.e. a large separation of the helical and Ras domain. This opening movement again promotes the release of GDP which was previously bound to the concerned α subunit. It has been shown through simulation studies that this internal unclasping motion of the α subunit, in particular the speed at which the separation movement of the helical and Ras domains takes place, is essentially dependent on the presence of a GPCR, which shows a catalyzing effect on the subdomains' movement.¹⁰⁰ Considered in more detail, the contact between a $G\alpha$ subunit and an activated receptor is a major factor responsible for the displacement of the $G\alpha$ C-terminal $\alpha 5$ helix.¹⁰⁰ This change of position of the $\alpha 5$ helix is made up by a slight protrusion towards the receptor, so the position changes relative to the rest of the Ras domain.¹⁰⁰ The direction of movement of the $\alpha 5$ helix is indicated in **Figure 3** by the comparison of two G protein structures once in and once without the presence of a GPCR (**Figure 3a** and **b**). The docking of the helix into the GPCR promotes an internal structural rearrangement of the Ras domain that is directly accompanied by weakening effects of the binding affinity of the Ras domain towards the nucleotide. More precisely, the Ras domain's structural reorganization ensures for the prevention of its (former) key interactions to GDP and thus for the accelerated nucleotide release.¹⁰⁰ Furthermore, it could be shown that in case of binding to the GPCR and the simultaneous absence of the nucleotide (within 25 microseconds of simulation), the domains stay separated from each other and the α subunit thus remains open – obviously thanks to the anchoring of the $\alpha 5$ helix with the receptor.¹⁰⁰ However, even in the absence of a GPCR, an intrinsic natural opening and closing movement of the GDP-bound α subunit or a separation of the helical and Ras domains from each other can be observed, which could also allow for GDP release.¹⁰⁰ Nevertheless, the extent of opening and its velocity are far from being as high as in the catalytic GPCR-bound situation.¹⁰⁰ If both receptor and nucleotide are missing, this results (by simulation) in a strong, sudden opening movement starting from the initial closed state, due to a spontaneous adaptation movement of the $\alpha 5$ helix to a receptor-bound-like state.¹⁰⁰ If GDP is now released and the $G\alpha$ subunit is opened, the GDP to GTP exchange, i.e. the binding of GTP to the $G\alpha$ subunit can thus take place more easily, finally bring the G protein into its activated (GTP-bound) state. The resulting activation leads to the dissociation of the $G\beta\gamma$ dimer from the $G\alpha$ subunit and inducing their individual downstream signalling processes. After the signal propagation, the $G\alpha$ bound GTP is hydrolyzed to GDP, resulting again in the inactive, GDP-bound form of the $G\alpha$ subunit and coincidentally leading to re-association of $G\alpha$ with the $G\beta/\gamma$ dimer.^{98, 101-107} The starting situation is now restored and a new activation cascade can begin.

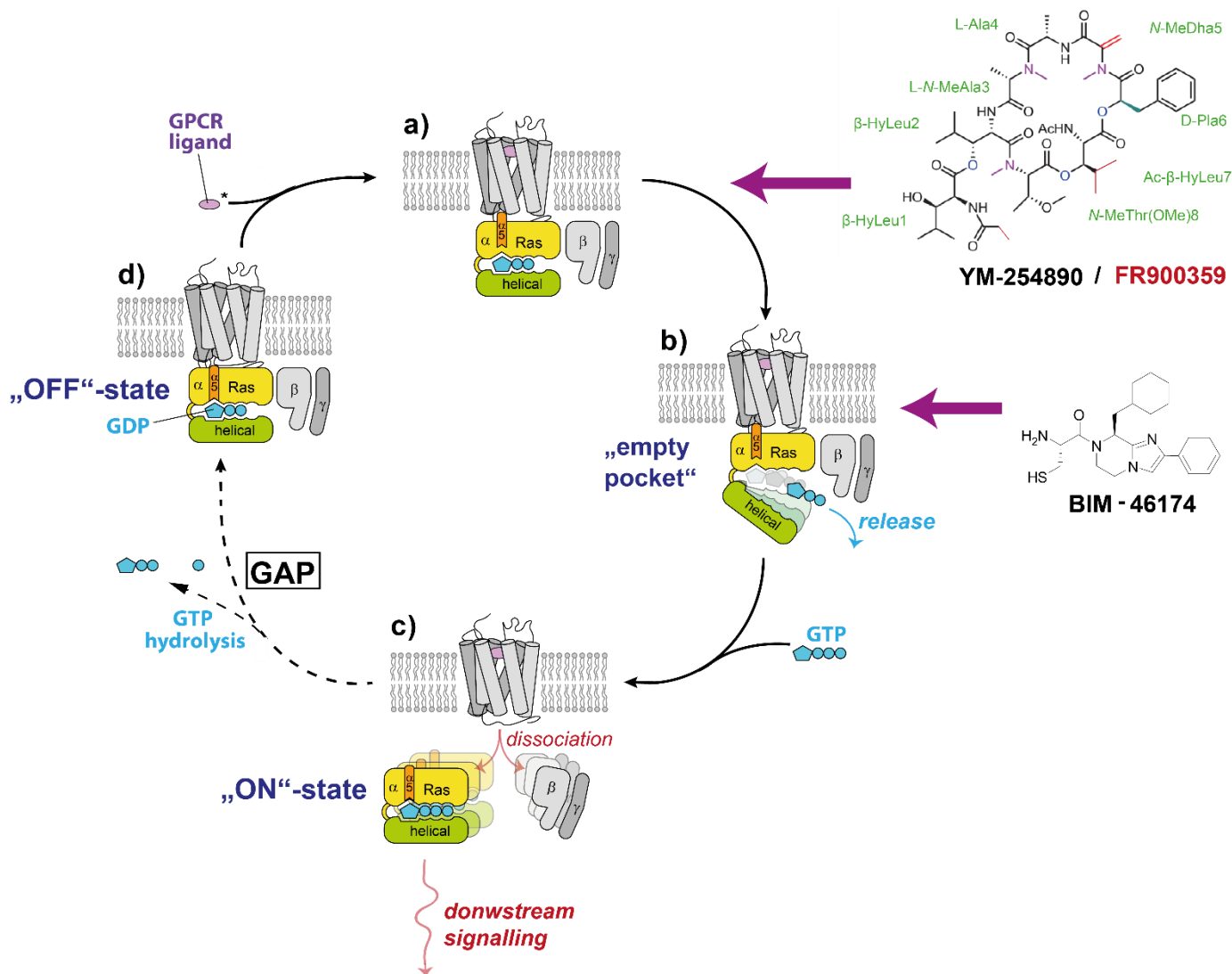


Figure 4: Cyclic scheme of G protein activation by a GPCR and associated GDP/GTP exchange; inhibition points for preventing GDP/GTP exchange and downstream signalling cascades by the molecules YM-254890/FR900359 (YM/FR) or BIM-46174 (BIM) are indicated by purple arrows. **(a)** Activator ligand and inactive (GDP-bound) G protein bind to the GPCR; $\alpha 5$ helix docks into the receptor; pharmacological inhibition of the $G\alpha_{q/11}$ subtype can be achieved at this point by YM/FR through inhibiting GDP release. **(b)** Ligand-bound activated GPCR triggering GDP release of the receptor-bound G protein; pharmacological inhibition of the $G\alpha_{q/11}$ subtype can be achieved at this point by BIM through trapping the "empty pocket" conformation after GDP release and thus preventing exchange with GTP. **(c)** GDP/GTP exchange induced G protein activation; dissociation of the activated (GTP-bound) $G\alpha$ subunit (in "on"-state) from the $G\beta/\gamma$ dimer complex and full dissociation of the G protein subunits from the receptor; initiation of further $G\alpha$ -induced signastranduction cascades. **(d)** Re-association to heterotrimeric G protein complex and GTP hydrolysis to GDP induced by GAP (GTP-ase activating protein), resetting the G protein into its inactivated, GDP-bound "off"-state for renewed reactivatability.

Finally, two exemplary physiological processes which are triggered by G protein activation through a GPCR will shortly be outlined.

An illustrative signalling pathway mediated by GPCRs is the slowing of heartbeat by activation of the muscarinic acetylcholine receptor, which is a GPCR of the heart muscle cells (M2 receptor).¹⁰⁸ As the name suggests, it is stimulated by acetylcholine, which is released when the parasympathetic nervous system is activated.¹⁰⁹ The GPCR in turn stimulates the β/γ complex of the G_i protein.¹¹⁰ This complex opens the potassium channels and thus enables the flow of K^+ ions into the extracellular space.^{111, 112}

As previously indicated, G proteins also play an important role in the visual process: the excitation of the G protein transducin by the GPCR rhodopsin leads to phosphodiesterase activation.^{113, 114} Rhodopsin itself is

located in the rod cells of the retina and is activated by light stimuli.^{115, 116} The stimulated phosphodiesterase hydrolyses cGMP to GMP, which leads to the closure of cGMP-dependent sodium and calcium channels and thus to a decreasing of the cGMP concentration.^{113, 117} The membrane potential for stimulus transmission in the optic nerve is then predominantly determined by potassium diffusion, which accounts for an equilibrium potential of about -90mV, i.e. the initiation and transmission of the stimulus in the ocular nerve is triggered.^{118, 119}

Both examples show that there is a connection between G protein activation and downstream regulated ion channels, which comprise the second system considered in this thesis. Complex signal transduction cascades can therefore be stopped, blocked or even redirected at different points of action.

3.2. YM and FR – the ultimate G_q inhibitors

As mentioned previously, the inhibition of G α -induced GDP/GTP exchange to silence G protein signalling has a high pharmaceutical potential with respect to numerous diseases (for instance cancer¹²⁰) associated with GPCR- and their related G protein activation cascades. However, only a few compounds that are able to modulate G protein activity are known these days and the targeting of G proteins unequivocally depends on their subtype.¹⁰⁴ For instance, pertussis toxin and cholera toxin can modulate G_i and G_s proteins, respectively,^{49, 50, 121} where in addition, a small molecule, BIM-46187 was shown more recently to trap the G α subunit of the G_q subtype in its empty-pocket (nucleotide-free) conformation, i.e. after its activation and GDP release.⁵¹

Further innovative discoveries concerning G α modulators have also been made, including the discovery of the small depsipeptides FR900359 (FR) and YM-254890 (YM) both of which are known to prevent G protein-induced signalling by acting as guanosine nucleotide exchange inhibitors.

BIM-46187 and FR/YM thus show their inhibitory effect on the G protein function by attaching at different action points in the activation cycle (**Figure 4**).

In 1988, Fujioka and co-workers first extracted the cyclic depsipeptide FR900359 (FR) from the evergreen plant *Ardisia Crenata sims*, which showed inhibitory effects on platelet aggregation and decrease of blood pressure.¹⁰⁷ In 2003 the chemically and structurally almost identical platelet aggregation inhibitor YM-254890 (YM) was isolated from bacterial strain *Chromobacterium species* QS3666 by Masatoshi Taniguchi et al.^{2, 122} It has been shown that the inhibitory effect of YM and FR is underlying the immobilization of the G α subunit opening motion of the G protein subtype G_{q/11}.^{1-3, 101, 104, 123} YM-254890 is currently accepted as the most efficient and selective inhibitor by G α _{q/11}.⁵ Since its discovery in 2003, YM-254890 has continued to attract great interest as this specific G α _{q/11} inhibitor facilitates the understanding of the role of G α _{q/11} in the GPCR transactivation of protein kinase receptors. Various studies have shown that YM-254890 could be a potential antihypertensive agent as it showed antithrombotic and thrombotic effects.¹²⁴ For this reasons YM, FR and analogue structures thereof bear a great pharmacological potential.

In 2016 and 2017, Xiong, Zhang and co-workers succeeded in producing YM, FR and analogues by total chemical syntheses starting first structure-activity relationships studies that were further continued by Reher and colleagues in 2018.^{5, 102, 104} However, any attempts to improve the inhibitory activity of FR, YM or analogues derived thereof and to achieve some inhibitory activity towards any other G protein subtype failed, although the crystal structure of YM-bound G_q was solved by Nishimura et al. in 2010.¹ In turn, we hypothesis, that the mechanism of G protein inhibition might not fully be understood further motivating our in silico analyses of G_q-inhibition by YM and FR.

Both molecules are structurally very similar and only differ by three additional methyl groups in FR (**Figure 5**). They are cyclic depsipeptides, i.e. ring-shaped, macrocyclic peptides consisting of eight amino acids each (**Figure 5**).

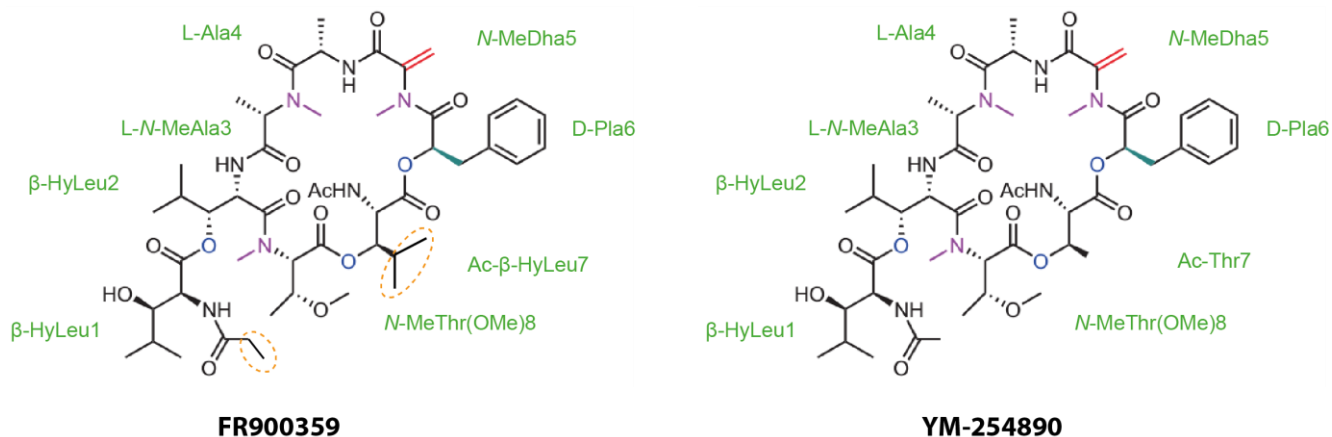


Figure 5: Chemical structures of FR900359 (left) and YM-254890 (right); FR additional methyl groups compared to YM are orbited in orange dotted lines. (Illustration modified from Xiong et al.¹⁰⁴)

FR and YM are small to medium-sized compounds, but with a high chemical complexity that should not be underestimated: they can assume specific three-dimensional conformations, which, together with a concomitant intrinsic rigidity of the molecule, results in a precise and targeted alignment of its backbone and side chain atoms. The typical depsipeptide structure is composed of a macrocyclic backbone surrounded by bound side chain residues and a β -hydroxy-leucine tail (**Figure 6**).

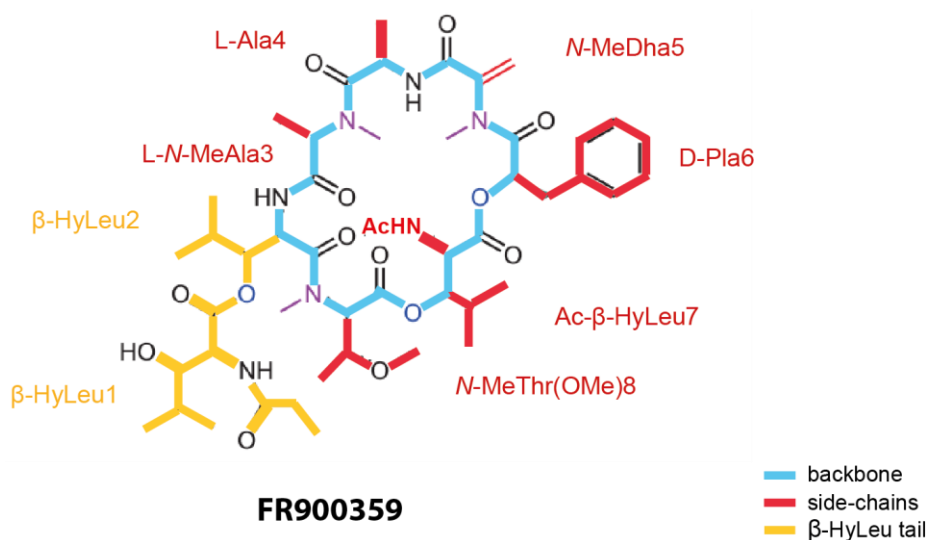


Figure 6: schematic view of the cyclic structure of the depsipeptides YM and FR shown on the structural basis of the FR molecule; the different molecular bond groups are indicated as bold lines. Colour scheme: cyclic backbone – blue, β -hydroxy-leucine tail – yellow, side chains – red. (Illustration modified from Xiong et al.¹⁰⁴)

Due to the high similarity of YM and FR and their respective capacities to selectively interfere with the function of $G\alpha_{q/11}$, it can be concluded that their inhibitory mechanism at the $G\alpha_q$ subunit is identical and it makes them exceptional among G protein inhibitors.⁵ Ultimately, G_q inhibition by YM or FR is achieved by impairing the domain opening motion of the helical and the Ras domain of the $G\alpha_q$ subunit which prevents the release of GDP.^{1, 3}

According to the crystal structure of G_q -bound YM, the inhibitor resides in a hydrophobic cleft at the interface of the α helical and Ras domain (**Figure 7**). The helical and the Ras domain of the $G\alpha$ subunit are linked together by two short flexible loop segments, called "Linker" and "Switch I", which enables the separation of the two domains from each other like through a hinge (**Figure 7**). In turn, YM most likely stabilizes the linker and switch conformation not allowing both domains to separate from each other preventing the nucleotide to escape. Therefore, FR and YM act as nucleotide dissociation inhibitors (GDIs).^{1, 3, 5}

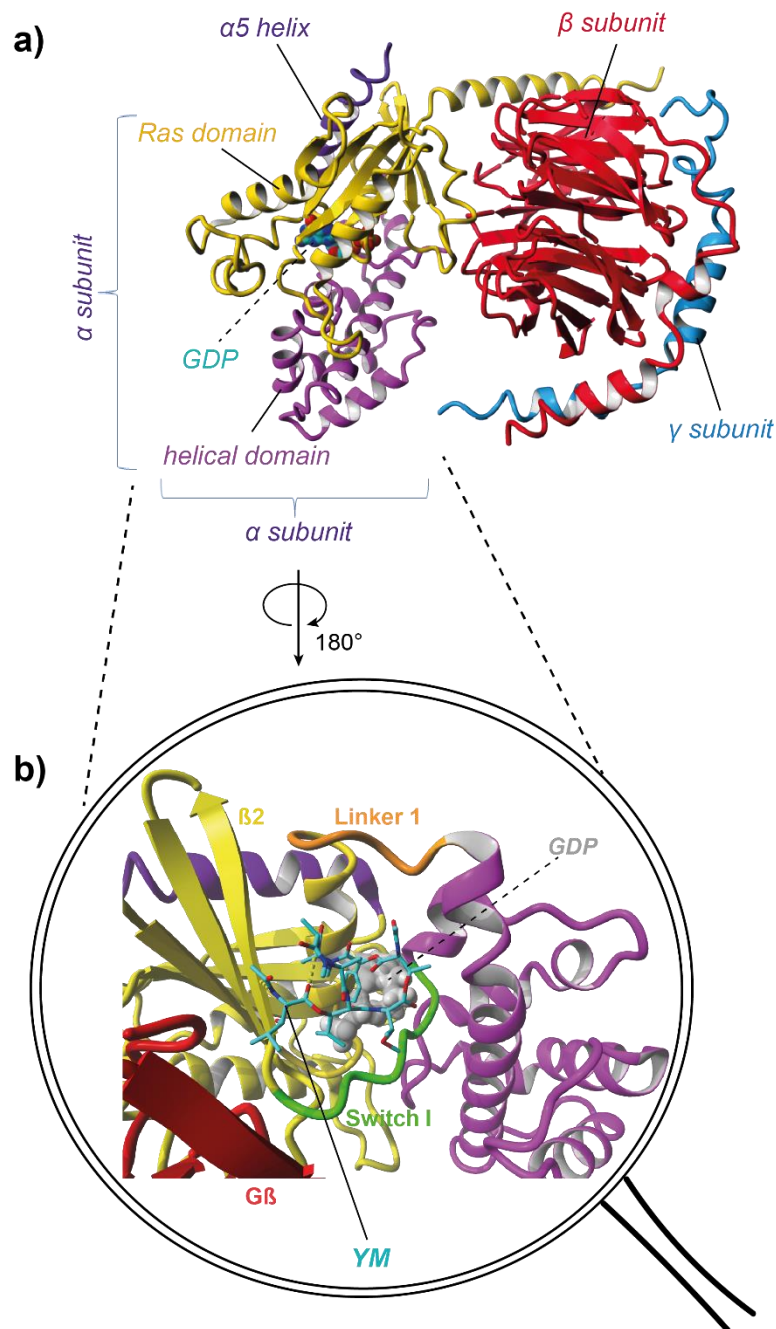


Figure 7: Localization of the G protein inhibitor binding site by the example of $G\alpha_q$ and interacting YM-254890 (PDB ID: 3AH8¹). (a) Complete heterotrimeric complex. (b) Rotated and zoomed view pointing out the inhibitor binding site connecting the helical and the Ras domain of $G\alpha_q$ in an articulated way. Colour scheme: $G\alpha$ subunit Ras domain – yellow, $G\alpha$ subunit $\alpha 5$ helix – deep purple, $G\alpha$ subunit helical domain – magenta, $G\beta$ subunit – red, $G\gamma$ subunit – blue, linker – orange, switch – green.

In the following chapter the structural and functional determinants of G protein inhibition by FR and YM will be computationally analyzed in detail in order to guide future rational design approaches on the development of FR/YM-derived G protein modulators.

3.3. In silico studies on G_q inhibition by YM, FR and analogues

Molecular modelling and docking studies

In order to accurately decipher the binding modes and ligand-G_q protein interactions, molecular modelling and docking studies for a large set of YM/FR analogues were performed (**Figure 8**) and analyzed with respect to their experimental biological activities. Therefore, we first grouped 14 new FR- and YM-derived compounds,⁵ which were synthesized and characterized in the labs of Imhof, König and Kostenis at the University of Bonn as well as all 23 already published analogues^{102, 104} according to their potency to identify tolerated, partially tolerated, non-tolerated minor, and non-tolerated major modifications in the original sequence (**Figure 8**). To account for the slightly variable values of bioactivity obtained in different labs, the YM/FR analogues were grouped into compounds of high (IC₅₀ < 5.0µM), medium (5.0µM ≤ IC₅₀ ≤ 20.0µM) and low (IC₅₀ > 20.0µM) binding affinity (**Figure 8**).

Molecular modelling and docking studies were applied for the analogues **15**, **17–27**, and **29–36** (**Figure 8**) and were constructed from the G_q-bound conformation of YM-254890.¹ The analogues were reconstructed according to their respective reported molecular chemical formulas, as they are described in their published syntheses performed by Taniguchi et al. in 2004,¹⁰³ Kaur et al. in 2015,¹⁰⁵ Rensing et al. in 2015,¹⁰⁶ too, Xiong together with Stromgaard and co-workers in 2016¹⁰⁴ and Zhang together with Stromgaard and co-workers in 2017.¹⁰²

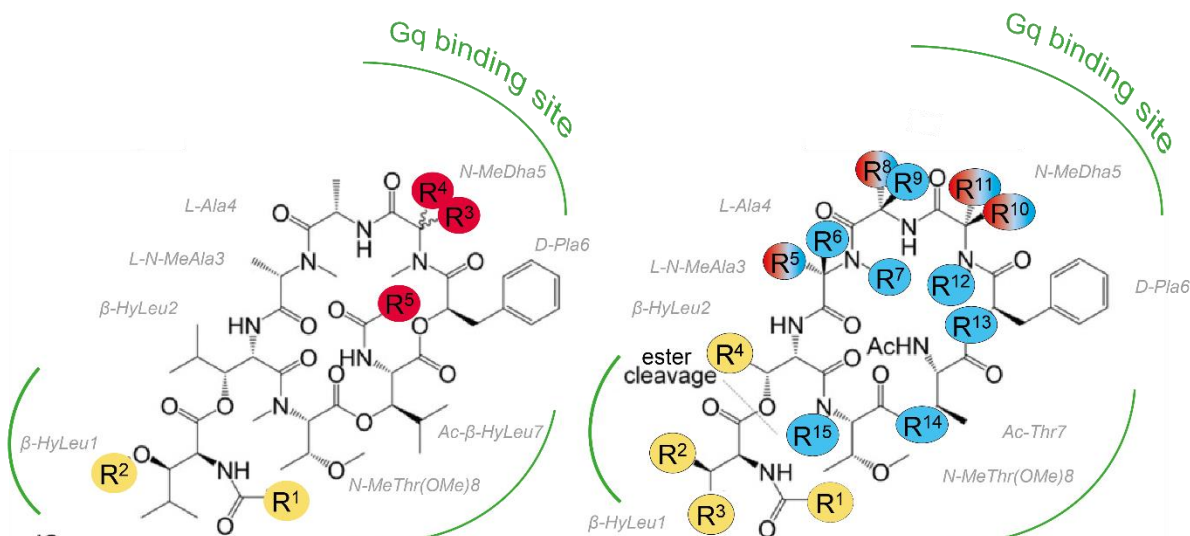
For FR-derived analogues **1–4**, two conformations were built employing the solution NMR structure of FR in water, which was recently determined in the Tietze lab and the G_q-bound conformation of YM (further details are provided in chapter 5.1 in the methods section). The resulting structures were then docked to the heterotrimeric G_q protein in its original conformation, limiting the docking site to the interface of the α helical and Ras domain.

For the analysis and scoring of the docking results, calculated binding energies which are expressed by VINAs scoring function as well as the structural deviation of the YM/FR analogues compared to the orientation of YM in pdb 3ah8 was considered. The later criterium was rationalized from the assumption that G protein inhibition can only be achieved when the relative orientation of the inhibitor in the binding pocket is similar to the orientation of YM as found in PDB 3AH8.¹

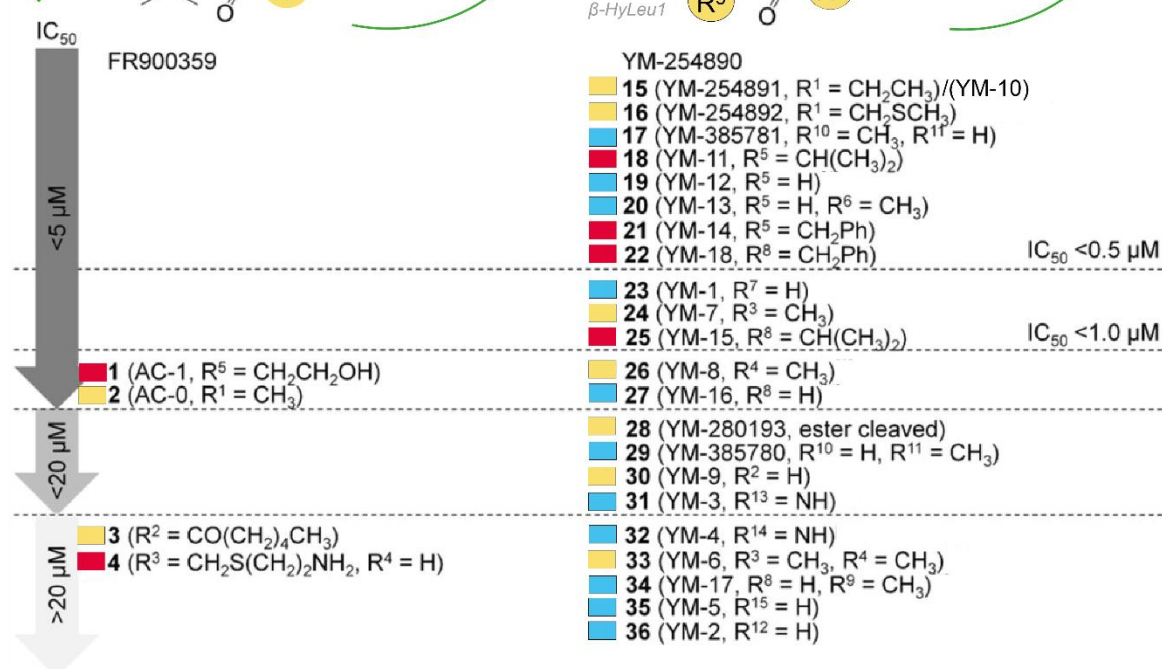
a)

FR900359 analogues

YM-254890 analogues



b)



- modifications effecting backbone alterations, eventually also influencing the intramolecular hydrogen bond network*
- modifications effecting upper side-chain alterations*
- modifications effecting β-HyLeu side chain alterations*

Figure 8: Individual FR (left) and YM (right) analogues investigated. (a) Molecular structure view; positions of residue modifications are marked in color according to their type of structural modification they can affect (see legend, *:= according to Reher et al.⁵); G_q binding pocket interacting regions are indicated by green lines. (b) Sorted representation of the individual analogues according to their activity levels together with the corresponding indications of the individual modifications; sorted analogues are colour-labelled according to the type of structural alteration affected by their individual modifications (see legend, *:= according to Reher et al.⁵). (Illustration adapted from Reher et al.⁵)

After conducting a first series of docking experiments, it emerged that for some YM/FR analogues the predicted binding energy from the docking runs were either "too good" or "too bad" with respect to their experimental activity assuming a somewhat incorrect bound state (**Table 10**).

However, there are relatively few outliers among the high affinity compounds, whereas medium and low affinity compounds almost always required additional revisory steps to find a suitable representative structure. Altogether, this shows the tendency of a rather high prediction performance of the docking program for high affinity structures and a rather low prediction accuracy for structures with low binding affinity, which may be due to the intrinsic purpose of the docking program for ideal spatial positioning of ligand and target structures.

Therefore, we refined our methodology to the effect that the selected outlier structures were equilibrated through a 100 ns molecular dynamics simulation to probe for their conformational integrity (**Figure 9**, **Figure 10**) and to obtain a more accurate, docking result with respect to the experimentally obtained IC₅₀ values. (Detailed information on the equilibrating simulations are given in chapter 5.1.3 in the methods section.)

Therefore, an equilibrated structure of the inhibitor was extracted from the simulation data for each outlier, focusing on somehow stable conformations within the simulations (considering C α backbone atom RMSD, (**Figure 9**, **Figure 10**). The selected inhibitor conformation was then re-docked as described above (see also chapter 5.1.2).

The refinement together with the equilibration data of the individual structures shows us the relevance of including the intrinsic dynamics and the conformational integrity of the analogues in the evaluation of the docking experiments. For instance, for some structures two different but stable conformers are observed (**Figure 9e**, **Figure 10c**). Furthermore, overall higher structural flexibilities of the outliers are especially observable for moderate and low affinity outlier compounds (**Figure 9**, **Figure 10**).

Especially for all high affinity outlier compounds our pre-equilibration step consistently resulted in an improvement, i.e. an adjustment of the overall fit of the binding energy values with respect to the experimental IC₅₀ values (**Table 10**, **Table 2**).

The stronger impact of the equilibration step on the numerical values within the high affinity outliers compared to the medium and low affinity "outliers" also suggests the higher structural dynamics of the medium and low affinity compounds when bound to G_q and their more difficult reproducibility by the docking experiment (**Table 10**). Particularly in these cases, subsequent simulation experiments would reveal a higher flexibility docked state – expectedly slower for moderately active compounds and faster for inactive compounds, i.e. accompanied by. Such insights on concomitant dynamics would compensate for the higher binding energies of docked low affinity cases that had been obtained even after the intermediate equilibration (**Table 10**, 35 (YM-5) and 33 (YM-6)). This also shows that the inclusion of dynamic information is particularly important to achieve a more accurate in silico description, especially for medium and low affinity systems.

However, the overall picture of in silico reflection per affinity group was not misrepresented and was altogether improved by the previous equilibrations. **Table 2** lists the docking results and calculated binding energies of the final selected clusters, including structures with previously equilibrated compounds.

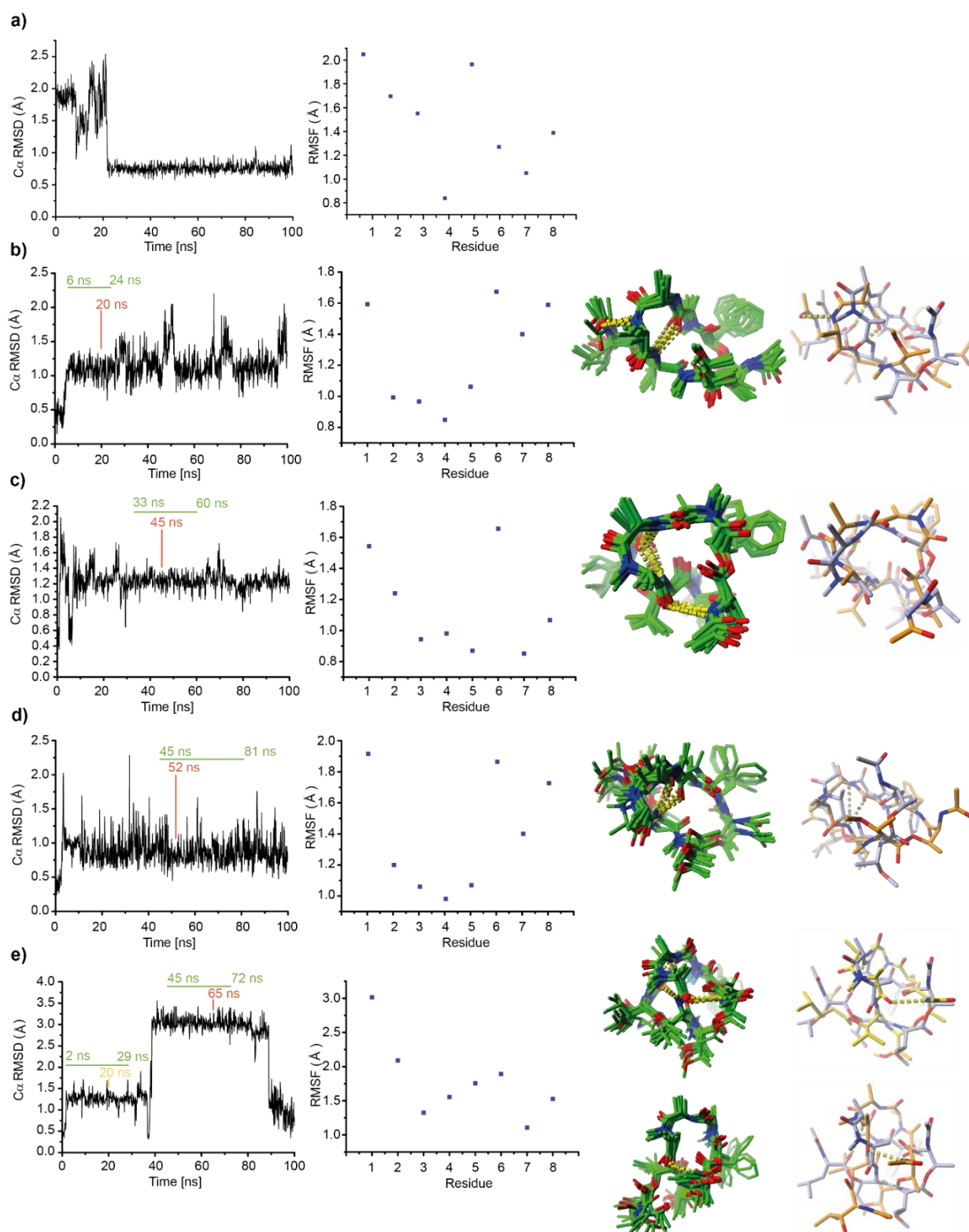


Figure 9: Overview of 100ns MD simulations of (a) YM and high/medium affinity outlier analogues (b) 23 (YM-1), (c) 30 (YM-9), (d) 31 (YM-3) and (e) 29 (YM-385780). RMSD time trace of C α atoms and per residue RMSF (first two panels from left) and structural ensemble (third panel, 10 structures, coloring scheme: carbon – green, nitrogen – blue, oxygen – red, hydrogens are omitted for clarity, hydrogen bonds are shown as yellow dashed lines) derived from the time interval indicated (green line) at the C α RMSD-time trace. Fourth panel: Superposition of YM-254890 (carbon – light purple) and energy-minimized representative structure (carbon – orange/yellow) extracted from the simulation after a certain equilibration time, which is indicated by a yellow/orange label at the C α RMSD-time trace and which was used for (re-)docking. (Illustration modified from Reher et al.⁵)

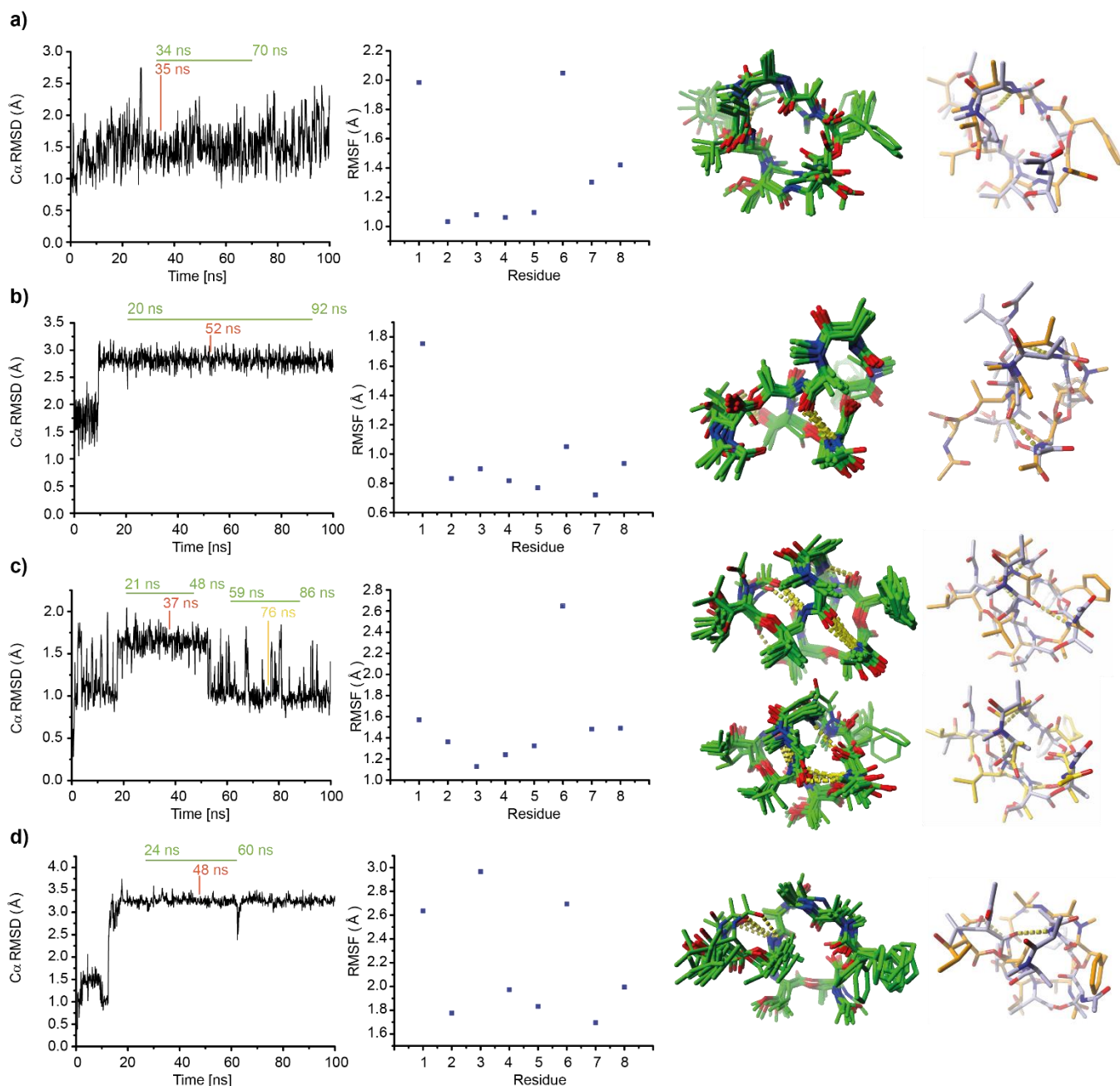


Figure 10: Overview of 100ns MD simulations of low affinity outlier analogues (a) 32 (YM-4), (b) 33 (YM-6), (c) 34 (YM-17) and (d) 35 (YM-5). RMSD time trace of C α atoms and per residue RMSF (first two panels from left) and structural ensemble (third panel, 10 structures, coloring scheme: carbon – green, nitrogen – blue, oxygen – red, hydrogens are omitted for clarity, hydrogen bonds are shown as yellow dashed lines) derived from the time interval indicated (green line) at the C α RMSD-time trace. Fourth panel: Superposition of YM-254890 (carbon – light purple) and energy-minimized representative structure (carbon – orange/yellow) extracted from the simulation after a certain equilibration time, which is indicated by a yellow/orange label at the C α RMSD-time trace and which was used for (re-)docking. (Illustration modified from Reher et al.⁵)

Table 2: Overview of the final docking results for the FR and YM analogues. The AutoDock Vina (referred to as Vina here) score is given for the best scoring cluster of each analogue. The structural deviation (all atom RMSD) of the selected cluster member of the $G\alpha_q$ -bound YM/FR analogue with respect to the orientation of $G\alpha_q$ -YM complex (PDB ID 3AH8¹) and the resulting conformation of the YM/FR analogue with respect to the YM conformation are calculated from the docking experiments. MD simulation derived structures of YM/FR-analogues are labeled with #. Alternative NMR-derived conformations are indicated with *. (Table source: Reher et al.⁵)

	Vina score	RMSD	RMSD
	[kcal/mol]	Gαq-YM analog/ Gαq-YM [Å]	YM analog/ YM [Å]
FR900359	8.63 ± 0.67	1.608	1.26
cis FR900359	7.62 ± 0.60	2.681	2.19
YM-254890	8.58 ± 0.54	0.000	0.0000
Best Fit			
15 (YM-10/ YM-254891)	8.64 ± 0.71	2.191	1.58
21 (YM-14) [#]	8.83 ± 0.38	2.199	1.51
20 (YM-13)	8.64 ± 0.42	1.810	1.49
22 (YM-18)	9.16 ± 0.56	2.366	1.93
18 (YM-11) [#]	9.01 ± 0.38	2.374	1.76
19 (YM-12)	8.34 ± 0.58	2.035	1.51
23 (YM-1) [#]	9.18 ± 0.51	2.094	1.29
25 (YM-15)	8.44 ± 0.54	1.870	1.67
24 (YM-7)	9.02 ± 0.75	1.631	1.27
1	8.51 ± 0.81	2.160	1.74
1 cis*	7.24 ± 0.00	2.457	2.01
17 (YM-385781)	8.30 ± 0.41	2.236	1.69
2	8.80 ± 0.63	2.164	1.90
2 cis*	7.55 ± 0.22	4.336	2.46
26 (YM-8)	8.94 ± 0.76	2.187	1.52
27 (YM-16)	8.57 ± 0.70	2.019	1.62
Moderate Fit			
30 (YM-9) [#]	9.55 ± 0.50	3.99	2.80
31 (YM-3) [#]	8.80 ± 0.79	8.307	3.28
29 (YM-385780) [#]	8.04 ± 0.30	6.662	3.00
Least Suitable Fit			
3	8.22 ± 0.17	8.432	2.02
3 cis*	7.51 ± 0.49	8.399	2.21
32 (YM-4) [#]	8.50 ± 0.34	5.226	3.17
33 (YM-6) [#]	9.83 ± 0.34	5.612	2.46
34 (YM-17) [#]	8.14 ± 0.44	4.533	3.37
35 (YM-5) [#]	9.28 ± 0.14	7.281	4.48
36 (YM-2)	7.96 ± 0.60	4.367	1.53
4	7.67 ± 0.26	3.878	2.29
4 cis*	7.57 ± 0.25	7.157	2.30

It can be stated that among all the analogues investigated, compound 23 depicts the only analogue containing a modification on its backbone (namely *N*-MeAla3 to Ala) (**Figure 6, Figure 8**). All other considered YM-derived analogues 15, 17-22, 24-27 and 29-36 and also FR-derived analogues 1-4, include modifications within the side chain groups, more precisely at *N*-MeAla3, Ala4, β -HyLeu1/2 or at the acetyl group of Thr7/ β -HyLeu7 (**Figure 6, Figure 8**). Closer inspection of the respective G_q-bound inhibitor conformations yielded by the docking calculations showed only minor changes and the intramolecular hydrogen bonds as found in YM and FR remained intact (**Figure 11a**).⁵ This also resulted in a similar backbone conformation of the analogues and thus in a similarly high binding capacity as YM and FR indicating that the conformation and stability of the backbone structure forms an essential basis for a proper binding or fit of the molecule into the binding pocket (**Figure 11a**). As has already been concluded for the similar binding behaviour of FR and YM due to the small structural differences, a similar positioning in the G_q binding pocket can also be deduced for the high affinity analogues 1, 2, and 15-23 (**Figure 8, Figure 11a, b**).⁵

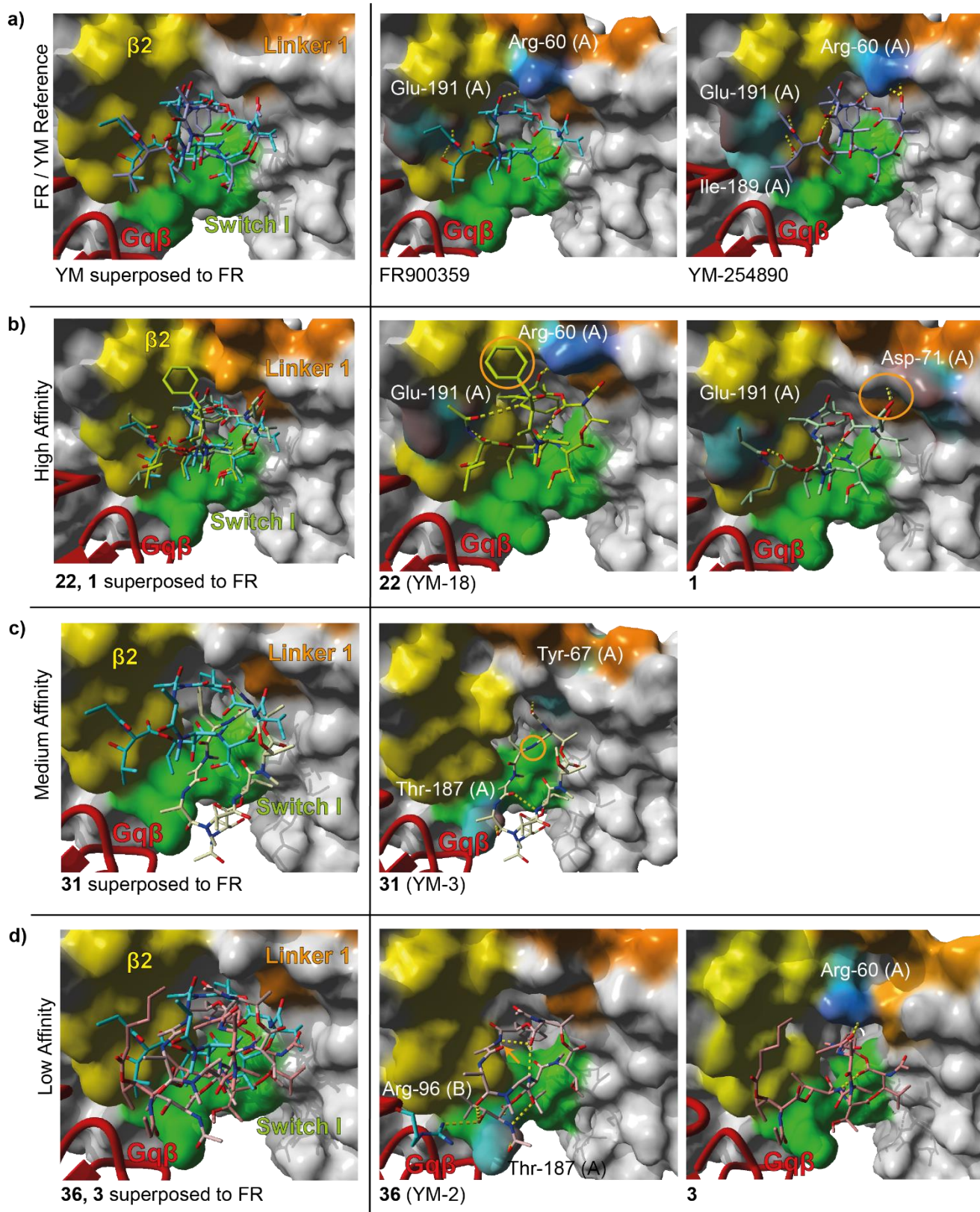


Figure 11: Representative docking results. (a) Superposition of reference structures YM-254890 (PDB ID 3AH8,¹ carbon – light purple) and FR900359³ (carbon – cyan) on the left, separate views of their individual binding modes on the right (coloring scheme: nitrogen – blue, oxygen – red, hydrogens are omitted for clarity, hydrogen bonds are shown as yellow dashed lines; Linker 1 surface – orange, Switch I surface – green, G α_q Ras domain $\beta 2$ sheet surface – yellow, remaining G α_q surface – gray, hydrogen bond interacting residues – labeled and highlighted by element color, G β – red). (b) Superposition of FR900359 (carbon - cyan) with the docked conformation of two selected best inhibitory analogues 22 (YM-18, carbon – yellow-green) and 1 (carbon – mint-green) (left) as well as separate views of their individual binding

modes (right); particular modifications are tagged by light orange circles or arrows. (c) Superposition of FR900359 (carbon - cyan) and the docked conformation of the selected representative moderately active analogue 31 (YM-3, carbon - pale yellow) (left) as well as separate view of its individual binding mode (right). (d) Superposition of FR900359 (carbon - cyan) and the docked conformation of two selected less active analogues 36 (YM-2, carbon - pale rose) and 3 (carbon - salmon pink) (left) as well as separate views of their individual binding modes (right). (Illustration modified from Reher et al.⁵)

Moreover, for these compounds our data indicate almost identical hydrogen bond interactions with the target protein (**Table 11**). Similar orientations as found for YM within the YM-binding pocket on G_q are found when the alternative NMR-derived structures for FR and compound **1** (referred to as *cis* FR and *cis* **1** thereafter) were used for the docking experiments (**Figure 12**). A slightly different G_q-bound orientation was unveiled for the NMR derived structure of **2** (referred to as *cis* **2** thereafter, **Figure 12d**).⁵

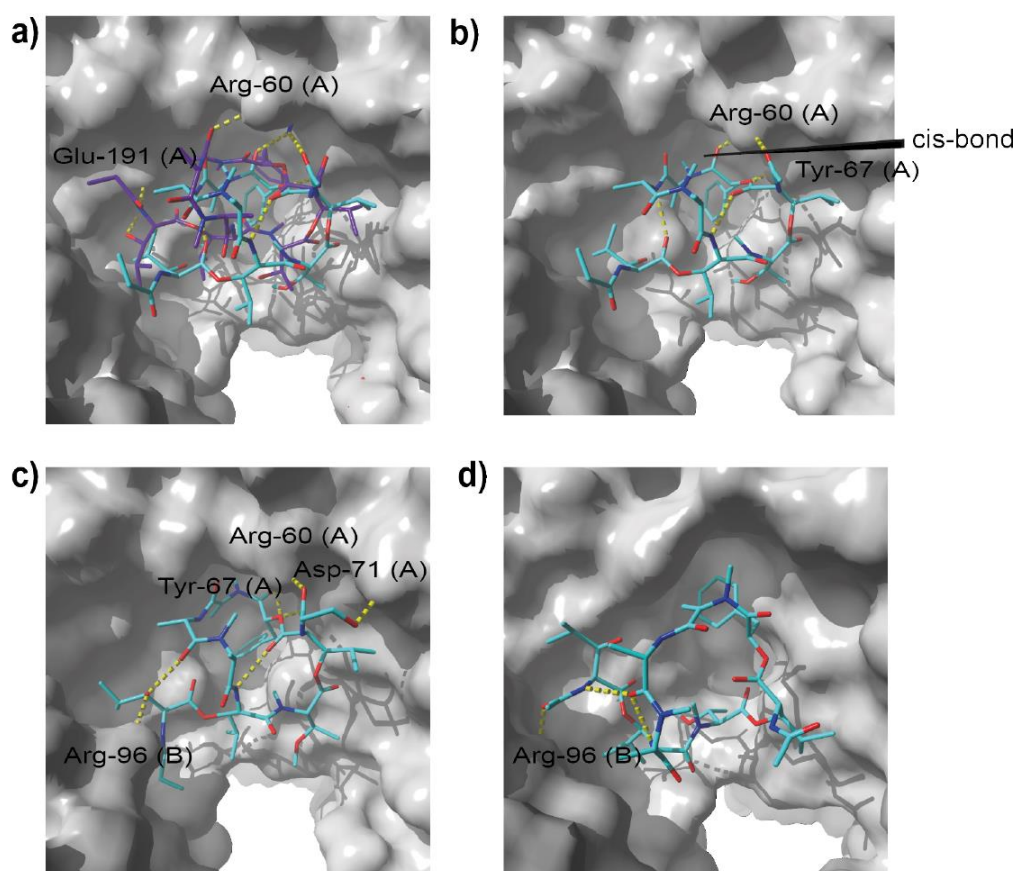


Figure 12: (a) Superposition of FR and *cis* FR bound to G_q as revealed from the docking studies. Orientation of (b) *cis* FR (c) *cis* **1** and (d) *cis* **2** bound to the G_q-FR/YM binding pocket. (Image source: Reher et al.⁵)

In conclusion, our docking experiments indicate that the high affinity compounds are strongly oriented towards linker 1 and β sheet 2 of the G α_q subunit. Especially, the hydrogen bonds of D-Pla6 or O-acetyl L-Thr7 of the YM analogue to Arg60 and Glu191 within the G α_q linker region and the hydrophobic interactions of the inhibitor residues Thr7 (YM) or β -HyLeu7 (FR) towards segments α 1 (I56, K57), α A (F75) and switch I (V184), and inhibitor residues β -HyLeu1 to Ala4 with the protein residues (I190, Y192, P193) located at the β sheet 2 region are essential features for an effective G_q inhibitor (**Figure 13**, **Table 11**).⁵

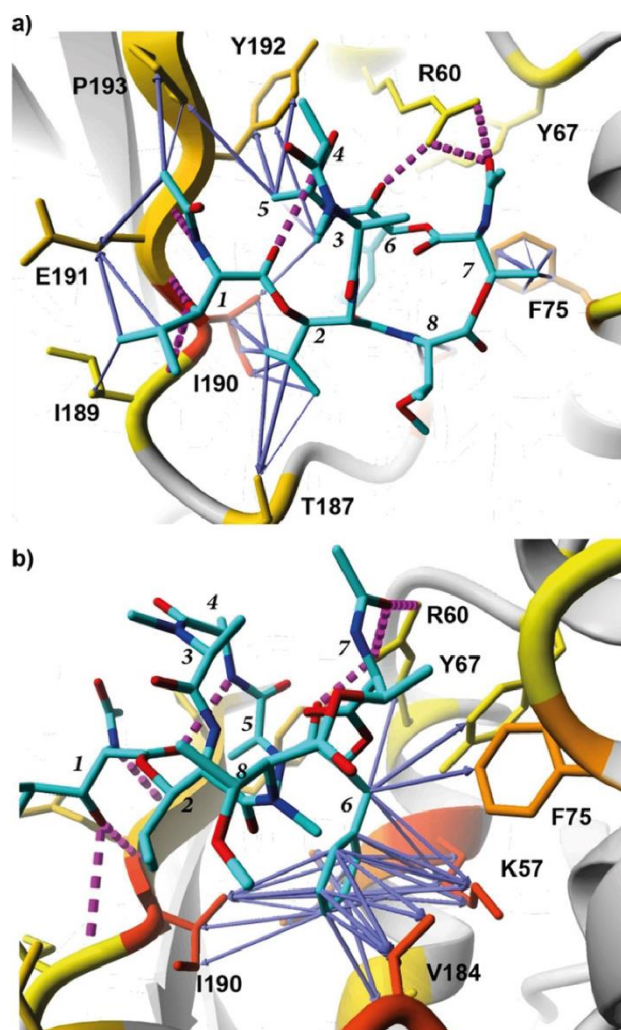


Figure 13: (a) and (b) Illustration of key interactions of the best inhibitory (high-affinity) compounds exemplarily shown for YM-254890 (residue numbers – italic; coloring: carbon – cyan, nitrogen – blue, oxygen – red) with G_q (residue numbers – regular, bold). (b) Network of hydrophobic interactions of D-Pla of YM-254890 within the hydrophobic cleft at the interface of the helical and Ras domain of the G protein. Intra- and intermolecular hydrogen bonds are indicated by magenta dotted lines; hydrophobic interactions between G_q and inhibitor are depicted by blue arrows between the interacting atom pairs (arrow thickness indicates the strength of interaction); G_q residues are coloured according to the sum of their hydrophobic interaction energy with the inhibitor, ranging from yellow (1 kJ mol⁻¹, weak) to red (10 kJ mol⁻¹, strong). (Figure adapted from Ref.⁵)

Interestingly, all analogues of medium and low binding affinity, except of 32 (YM-4), lack the H-bonds with the β sheet 2 region (Glu191) (**Table 11**) suggesting an important anchor point for the inhibitor in order to effectively silence the G protein motion. In particular, the hydrogen bonds formed between D-Pla6 and Arg60 in high affinity compounds seem to favor the hydrophobic interactions of D-Pla6 arising with other residues of the linker region around Arg60 and thus stabilize the inhibitory binding positioning of the ligand molecule within the pocket (**Table 11, Figure 13b**). This is exemplarily illustrated by G_q-bound YM-254890 and holds also true for β -HyLeu tail groups interacting with the β 2 sheet region close to the switch 1 region (**Figure 13**). In particular, the fundamentally higher total fluctuation (RMSF) of Residue 6 (D-Pla6) for almost all outlier structures in comparison to YM is noticeable in this context when considering our equilibration simulations (**Figure 9, Figure 10**).

As already indicated, for the evaluation of our docking studies of moderately or weakly active compounds with regard to the reduced binding affinity and the presumably associated higher inter- and intramolecular dynamics, the respective flexibilities should also be considered, which is also reflected by the particularly large number of outlier structures in these groups (**Table 2, Table 10**).

For all moderately active analogues 29-31 our MD equilibration simulations yielded higher molecular flexibilities compared to YM/FR and thus a generally higher tendency to adopt different conformations (**Figure 9**). This can also be explained by a higher flexibility of the backbone, which more likely allows for a transition into another conformational state including a different intramolecular H-bond pattern. The resulting different global conformations compared to the original starting conformation of course lead to altered H-bond formations compared to the high affinity interaction pattern and associated different positionings in the docking where the overall orientation of the analogue in the G_q binding pocket differs from that of YM or FR (**Figure 11, Table 2, Table 11**).

Equally noteworthy in these respects are the hydrogen bond contacts of the inhibitor compounds with Thr187 of switch I and with Arg96 of the $G_q\beta$ subunit, which preferably arose within docked medium and low affinity analogues (**Table 11**) mainly formed by the β -HyLeu tail (**Table 11**). Moreover, our data indicate, that the ligands protrudes more from the binding pocket to the direction of the $G\beta$ subunit and/or switch I simultaneously and are thus located further away from the previously addressed linker 1 and the $\beta 2$ sheet of the G_q Ras domain, as seen for the high affinity compounds (**Figure 11c, d**), clearly explaining their lower potency.

Considering the low affinity compounds, i.e. analogues 3-14 and 32-36 an increased structural flexibility as revealed by our MD equilibrations (except of analogues 3 and 4) significant alterations of the global conformation and in the bound conformations compared to YM/FR were observed (**Figure 9, Figure 10, Figure 11d, Table 2**).⁵

Upon closer inspection of the orientation of the low affinity analogues in the FR/YM-binding pocket of G_q , it becomes apparent that the phenyl ring of D-Pla altogether resides in the hydrophobic cleft, as consistently observable for the high and moderate active analogues (**Figure 11d**).⁵ However, the overall binding affinity towards linker 1 is weakened for the low affinity compounds¹, whereas the tendency of establishing hydrogen bonds with residues of the switch I region or the $G\beta$ subunit seemed to be increased resulting in an even larger deviation from the high affinity binding pose (**Table 2, Table 11, Figure 11d**). Except from compounds 3 and 32, none of the low affinity analogues still shows H-bond contacts to Arg60 and the major partner for hydrogen bond formations is now represented by Arg96 of subunit $G\beta$ (**Table 11**).⁵

The modifications introduced into the low affinity analogues were sufficient to result in significant alterations of the molecular conformation of the inhibitor and thus to weaken its ability for a proper accessing of the rather cavernous G_q binding pocket.⁵

Finally, for the design and synthesis of novel, YM/FR-like G protein inhibitors, these insights teach us to preferably preserve simultaneously formed contacts towards Arg60- and Glu191-corresponding residues at proportionate positions of other G protein subtypes as well as to avoid interactions with residues affecting the switch region or the $G\beta$ subunit.

Deviation of recognition determinants for inhibitor binding to $G\alpha_q$

Based on the previous analyses, another ultimate goal was to delineate distinctive recognition determinants that were characteristic for the inhibitory binding to $G\alpha_q$. In this respect, we first approached our aim by identifying characteristic attributes of YM/FR that account for high levels of inhibitory activity and whose lack or modification would drastically reduce the effect of inhibition. In this respect we sampled the chemical space of the inhibitor compounds especially in order to determine molecular sites conceivable for eventual optimizations. As were seen from previous illustrations, YM and FR represent small to medium-sized macrocyclic compounds of particular three-dimensional conformations and a coincident high degree of chemical complexity.⁵ However, their intrinsic rigidity involves a quite concise overall backbone orientation together with a distinctive arrangement of the side chain groups (**Figure 6, Figure 8**).⁵

Altogether our analyses could highlight molecular properties of three distinct categories whose changes affect the inhibitory capacity compared to the lead compounds.⁵ More specifically, these concern alterations of I) the backbone conformation, II) the backbone conformation that in turn influence the intramolecular hydrogen bond network, and III) the side chain constitution and orientation (**Figure 14, Figure 8**).⁵

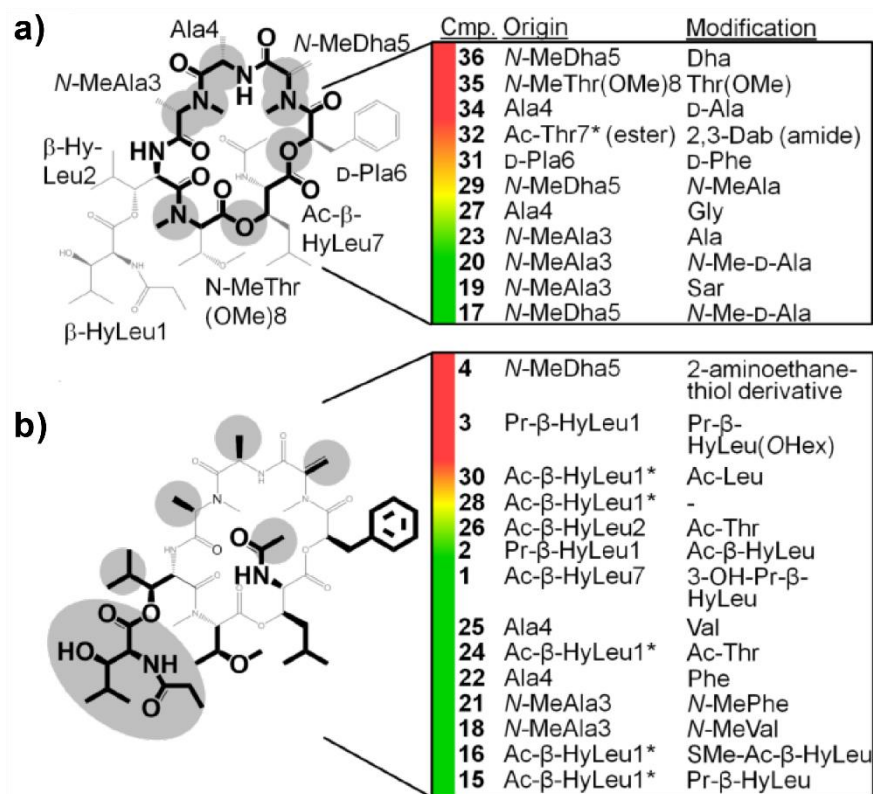


Figure 14: Structural classification of compounds with modifications at one site of FR and/or YM according to their impact on (a) backbone and intramolecular hydrogen bonds and (b) side-chain modifications; that is, inactive analogues 5–14 carrying more complex modifications at different sites are not considered here. Coloured bars next to compounds indicate acceptance of modifications (red: inactive compounds with minor non-tolerated substitutions, yellow: moderately active compounds with partially tolerated substitutions, green: active compounds with tolerated substitutions). *Nomenclature refers to YM-254890. 2,3-Dab=2,3-diaminobutyric acid; Sar=sarcosine; SMe-Ac=methylthioacetyl.¹⁰³ (scheme from Ref.⁵)

Among analogues ranking among categories I or II, i.e. mainly containing alterations that cause a change in the backbone, which strongly decreases or in some extreme cases completely diminishes the activity are attributable to compounds 36, 35, 34, and 32 (Figure 8, Figure 14).⁵ Even minimal structural changes of a single atom or only a small functional group at positions 8 (*N*-MeThr(OMe)), 7 (Ac-Thr in YM), 5 (*N*-MeDha) or 4 (Ala) of the parent YM/FR structure could achieve those strong effects (Figure 8, Figure 14a).⁵ As already mentioned, among the general backbone-modified analogues, a further distinction can be made between compounds that only carry a pure (sole) conformational change of the backbone (category I) and compounds whose backbone functional impairment is caused by their intramolecular hydrogen bonding network (category II).⁵ For example, for analogue 34 in category I, the backbone conformational change is caused by inversion of the configuration of L-Ala4 to D-Ala, while the category II analogues 32, 35 and 36 show alterations within their intramolecular hydrogen bonds (Figure 9, Figure 10).⁵ These aspects were also revealed in our equilibration simulations for outlier structures 34 of category I and the two outliers 32 and 35 of category II (Figure 9, Figure 10).

With respect to the individual binding modes of category I and II analogues, their structural integrity compared to YM and FR was generally disturbed.⁵ Nevertheless, these effects can be further broken down with regard to the position of the modified site in the molecule and its efficiency: for example, when comparing the inactive compound 32 with the still moderately active compound 31, it turns out that the substitution of an ester by an amide bond, can have different effects depending on the modification site. Whereas D-Pla6 has been modified to D-Phe6 in compound 31, retaining moderate activity, introduction of an amide bond at the proceeding amino acid position (Ac-β-HyLeu7) yielded only weakly active compounds (32 and 34–36, Figure 8, Figure 14b).⁵ This shows that the ester bond at position 7 is more prioritized with respect to the bioactivity of the macrocyclic conformation than at position 6.⁵ In contrast, modifications at opposite side of the macrocycle did not show such a large effect in view of activity attenuation, as can be seen by the example of compound 23, in which *N*-MeAla3 was substituted by Ala (Figure 8, Figure 14).⁵ However, this backbone modification is also associated with a change in the internal hydrogen bond network.⁵

Concerning backbone modifications only the reduction of *N*-MeDha5 to *N*-MeAla had a significant effect on the biological activity (**Figure 8**, **Figure 14**).⁵

All other structural changes as present in 27, 23, 20, 19, and 17, including configuration inversions (*N*-MeAla3 to *N*-Me-D-Ala as seen in compound 20) or an increase in the molecular flexibility (e.g. by the modification Ala4 to Gly in compound 27) were tolerated.⁵

In summary, it can be stated that modifications within the region of the YM/FR molecule, comprised by residues *N*-MeAla3 and Ala4 were more accepted compared to modifications between *N*-MeDha5 to *N*-MeThr(OMe)8.⁵ Probably because the latter are located closer to the protein surface and face the binding pocket to a larger extent, even small or marginal modifications at these positions are not tolerated (**Figure 8**, **Figure 14**).⁵

The situation was somewhat different for analogues from category III. Detailed analyses revealed that all changes except from the modifications at the *N*-MeDha5 exomethylene group as seen in compound 4 and the Pr- β -HyLeu1 side chain in analogue 3 were partially or almost fully tolerated (**Figure 8**, **Figure 14**).⁵ Noteworthy, the above-mentioned backbone-related criteria are preserved in these analogues which may be a major reason for the higher acceptance of this type of modifications.⁵ Moreover, equilibration simulations support these aspects, unveiling lower intrinsic per residue flexibilities for the side-chain modified analogues 30 (**Figure 9c**) which is moderately active and 33 (**Figure 10b**) which ranks among the low affinity compounds. In both cases, their higher rigidity is caused by stabilizing internal hydrogen bonds of the unaltered backbone similar to that of YM/FR (**Figure 9**, **Figure 10**).

In addition, regarding the protein-facing residues 6, 7, and 8 in YM/FR no alterations were made or identified whereas at both ends spanning the sequence between residues 5 (*N*-MeDha) and 1, 2 (*Ac*- β -HyLeu) modifications maybe critical as can be proven with compounds 3, 4, 28, and 30 (**Figure 8**, **Figure 14**).

For side-chain modified analogues, such influences can especially be proven when considering compounds 3, 4, 28, and 30.⁵ In particular, the low and medium affinity analogues 3 and 30 lack the crucial hydrogen bondings between the hydroxyl function of *Ac*- β -HyLeu1 and the G_q backbone residues Ile190 and Glu191⁵ that would provide a stabilizing effect of the compounds' binding mode. On the other hand, but achieving a comparable effect, the prolongation of the *N*-MeDha5 residue in compound 4 leads to an unfavourable bulkiness of the affected residue and thus prohibits the equally important interactions for D-Pla6 at the binding site.⁵

Possibly for similar reasons simultaneous exchange of *Ac*- β -HyLeu1 and *Ac*- β -HyLeu2 by the smaller *Ac*- β -Thr resulted in total loss of activity for the side-chain modified analogue 33.⁵ Regarding our equilibrating simulations and docking data (**Figure 10b**) this modification may facilitate their reorientation towards the Arg96 residue of subunit G β ,⁵ ending up in a similar situation as seen for the low affinity compound 36 (**Figure 11**). Moreover, this reorientation would not only result in the loss of essential contacts to Ile190 and Glu191, as is the case with the moderately active compound 30 and the weakly active FR-analogue 3 (both side-chain modified), but also in the disappearance of hydrophobic interactions within this corresponding region.⁵ In the case of the side-chain modified compounds 24 and 26, each containing a separately mutation to Thr at one of the two β -HyLeu tail groups, the aforementioned steric reductions were not sufficiently large to achieve such reorientations and could thus preserve the critical hydrogen bond interactions with the protein, allowing for their higher activity (**Table 11**).⁵

In conjunction with the analyses carried out on category I and II backbone-modified analogues it emerges that except of the two residues *N*-MeAla3 and Ala4, the remaining part of the FR/YM molecule can altogether be considered as "pharmacophore" scaffold for addressing the G_q binding pocket (**Figure 13**).⁵

3.3.1. Conclusions and future perspectives

In conclusion, we were able to characterize the binding mode of the most potent inhibitors together with the detection of interaction patterns being critical for the inhibition of G_q . Based on our modelling data of the macrocyclic depsipeptide structures of YM-254890 and FR900359 and related analogues, we could also obtain comprehensive and accurate insights into the underlying structure-activity relationships associated with effective G_q inhibitory mechanisms.

In these respects we were able to identify at least two essential binding regions for YM and FR at G_q , namely residue Arg60 in the linker 1 region and Glu191 in the $\beta 2$ sheet region of the $G\alpha$ subunit, close to switch I region.

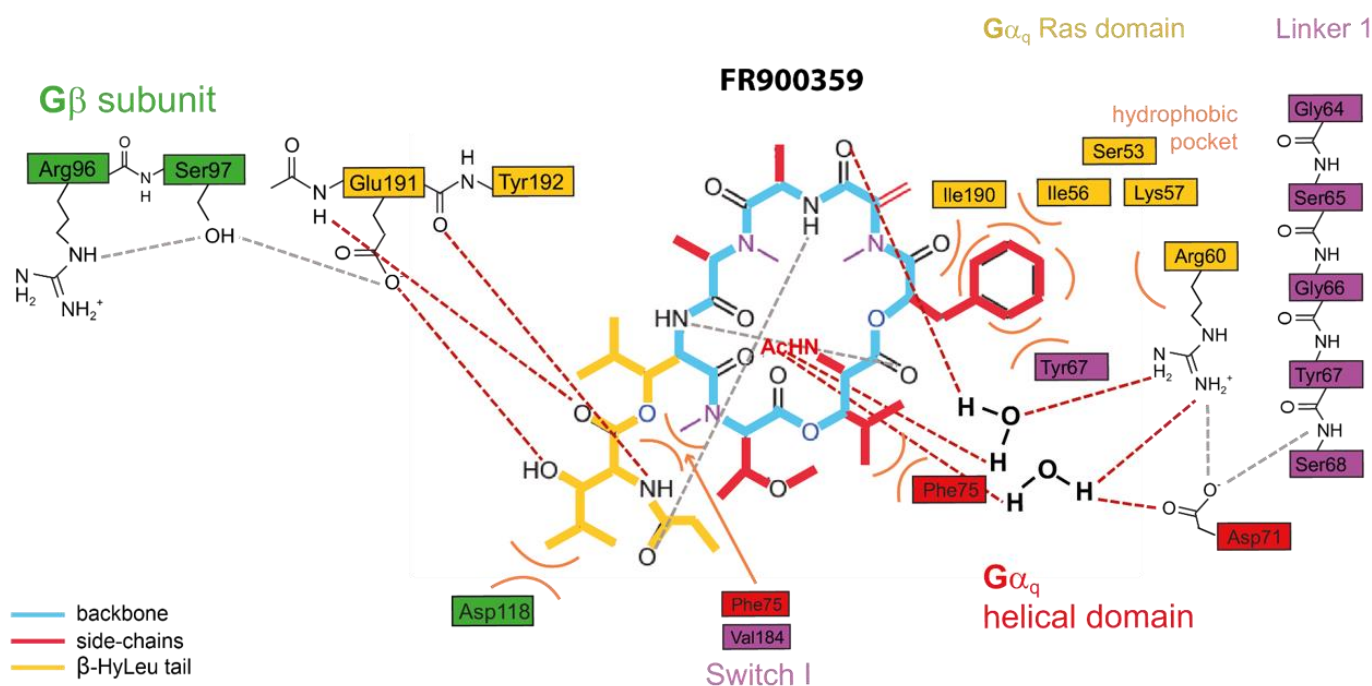


Figure 15: Scheme of main inhibitory YM/FR interactions to the essential positions in the G_q binding pocket using the example of FR900359; hydrogen bonds are displayed as dashed lines (intermolecular – red, intramolecular – gray), hydrophobic interactions as orange curved lines.

For the second essential binding site, namely Glu191 and also the neighbouring Ile189, it was shown that the shape, length and mobility of the β -HyLeu residues interacting with it is of essential importance for an inhibitory effect. It is absolutely necessary to mention that only the presence of the interactions with both important anchor points, Arg60 and Glu191, can achieve efficient inhibition. Since it could be shown that the inhibitory effect drops drastically to only moderate activity or even to complete inactivity as soon as at least one of these important contacts is missing, the necessity of the existence of both of these contacts must be assumed.

Recently, data from our collaborators suggested a third essential binding site, which is formed mainly through hydrophobic interactions of Thr7 (YM) or β -HyLeu7 (FR) with Phe75 also play an important role (Figure 15).¹²⁵

Thus our data suggested, that except of two residues, *N*-MeAla3 and Ala4, the remaining part of the YM/FR molecule represents the “pharmacophore” for this specific binding pocket on G_q (Figure 15).

At the same time, the combination of docking and molecular dynamics simulation studies also showed that it is particularly important for medium and poorly binding compounds to consider dynamic information compared to active compounds for a more accurate in silico description. This is due to the fact that the computer-generated reproduction of a “representative binding” state by sole docking experiments is altogether more difficult for systems of higher dynamics.

Lastly, our computational studies underpin the insight that both molecules YM and FR obviously represent the optimally designed compounds in respects of G_q inhibition: since none of the introduced modifications in the natural product structures yielded more powerful inhibitors – neither from an artificial perspective as seen in this work nor from a natural synthesis point of view as conducted by Reher and co-workers⁵ – YM and FR can be declared as the highest efficient and specific compounds for G_q inhibition known so far.⁵

In view of the extensive evolutionary history that natural products bring with them and the associated recurring adaptations towards an ideal mode of action, the high specificity of YM and FR against G_q can also be made plausible: since YM and FR as natural products are equally subject to such an evolutionary optimization process of millions of years, even in co-evolution with their target protein,^{126, 127} their high specificity, which they have developed over time against G_q, can be conclusively explained.

Hence, it seems unlikely that a significant improvement regarding G_q-inhibitors can be obtained with the limited opportunities provided by the YM/FR template(s), which already serve as suitable chemical tools in basic pharmacological studies.¹²⁸⁻¹³⁰ In terms of perspective, one conceivable task thus could be the implementation of strategies to target other G α protein subunits based on the YM/FR natural models.

Finally, regarding the conformation of FR in water, which differs from the G_q-bound conformation of YM as found in PDB 3AH8 it remained unclear, which isomer is preferably bound or whether they exhibit different target affinities. However, attempts to address these questions are currently ongoing and were not completed when the thesis was written.

4. Investigations on the mechanism of ion channel block by conesnail toxins

4.1. Voltage-gated ion channels

Voltage-gated ion channels, especially specific for the flow of calcium potassium or sodium ions through cell membranes, are important components for the activation of excitable cells like neurons or muscle cells. They ensure for the *selective permeability* of biological membranes, a particular essential feature in synaptic nerve cell membranes. Generally, such channels can be in a closed, or open, inactivated state. Within nerve synapses, they directly contribute to the tension gradient between the nerve cell interior and the synaptic cleft, the border area between two neighboring nerve cells. This gradient is also known as *resting potential* and depicts an utmost prerequisite for the triggering and transmission of nerve stimuli. The maintenance of this basal voltage of approx. -70 mV is regulated by sodium-potassium pumps, i.e. sodium- and potassium-permeable ion channels in the cell membrane. The reciprocally controlled channel opening and closing is also indispensable for the actual transmission of stimuli between neurons, i.e. during *depolarization*, *repolarization* and *hyperpolarization*. The ion flow itself occurs along the specific concentration gradient, either from the extracellular space into the cytosol or in the opposite direction. Regarding stimuli initiations and transmissions, the ion flow through sodium channels into the nerve cell interior is crucial for the impulse and propagation of action potentials and the flow of potassium ions out to the extracellular space has a hyperpolarizing influence on the membrane voltage at the end of the transmission.⁷ This interplay clearly points out the physiological relevance of ion channels.

The channels themselves are distinguishable by their selectivity towards different ion types, which they are able to transport through the cell membrane. The so-called "*selectivity filter*", named thereafter, predominantly consists of charged amino acid residues in the inside of the annular channel pore. The charged amino acids of the selectivity filter exhibit *highly conserved* patterns, for example the sodium-permeable channel is equipped with a "DEKA-motif" in its pore centre.¹³¹ Just like these amino acid residues, also the ionic radius as well as the charge of an ion represent important determinants for the channels' selectivity. The selectivity filter thus serves, so to say, as narrow passage inside the channel.

Correspondingly, the formal designation of ion channels is based on their selectivity for specific ions; for instance they are called calcium, potassium or sodium channels (Cav, Kv and Nav channels), according to their respective permeability for calcium, potassium or sodium ions. As already indicated, they are very well-known channel types whose physiological relevance has already been proven by numerous scientific studies and they are still of great interest especially in the neuroscience area. Due to their high relevance in the initiation and propagation of action potentials outlined above, they are predominantly found in the presynaptic membrane, more specifically in the membranes of the axon terminal of a neuron. This area ranks among the so-called synaptic cleft, the intermediate space of two neighboring neurons, where the transmission of stimuli takes place. The following figure schematically illustrates a neural synapse, the point of junction of two neurons, including all associated entities (**Figure 16**, left side).

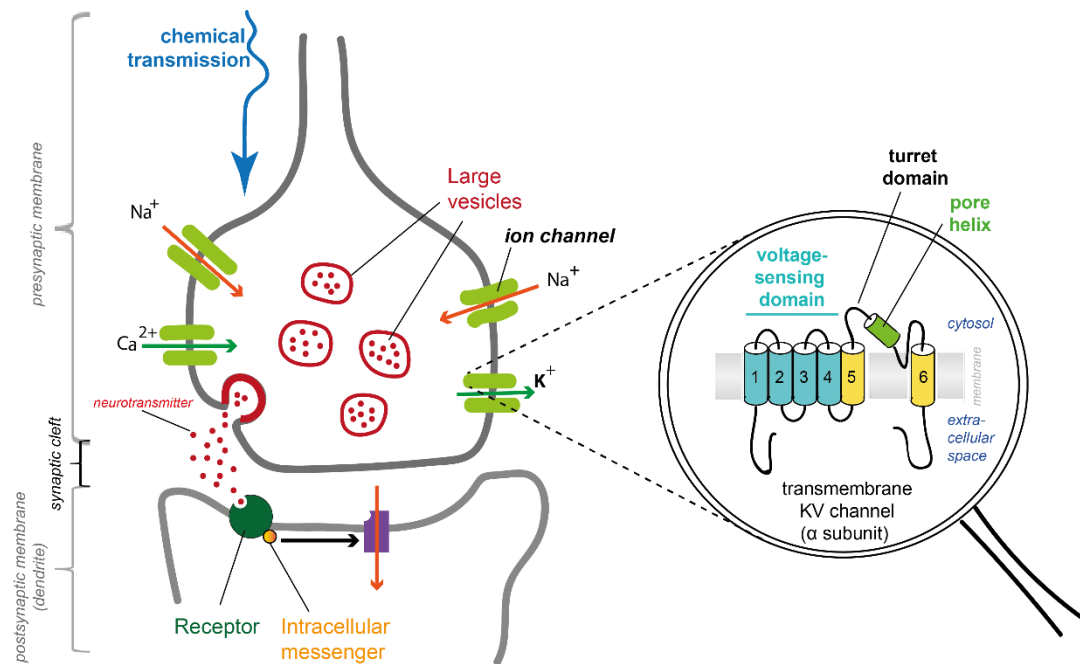


Figure 16: Neural synapse in schematic representation (left) and zoomed view of a Kv channel α subunit monomer in transmembrane topology representation (right).

The individual channel types can be subdivided into subfamilies according to their individual compositions of amino acid sequences. Hence, potassium channels can be divided into 12 subfamilies $K_V1 - K_V12$ and further into subtypes according to their inactivation mode.⁷

In structural terms, ion channels are proteins that form a ring-shaped channel pore through the cell membrane. These can be opened temporarily, whereby the opening and closing process ultimately results from a conformational change of the channel. It is assumed that the conformational change is triggered by a charge shift within or at the channel pore.¹³²

The ring-shaped complex of a channel is formed by four equally constructed homologous subunits, the *monomers*, and thus depicts a tetramer. The complex can occur as a tetramer of (sequentially) different monomers, i.e. as *heterotetramer* or of (sequentially) the same monomers, i.e. as *homotetramer*. Structurally, however, the subunits are always almost identical. As an example of a heterotetrameric complex the Na_V channel can be named, whereas the K_V channel is a homotetramer. Despite this difference and although not completely symmetric, a general similar basic structure of Na_V and K_V channels is presumed.⁷ In this context it has to be noted that X-ray structural data is explicitly available for K_V channels, but not for the mammalian Na_V channels. Nevertheless, cryo-EM structures of the latter are present these days.^{6, 133}

Using the example of the K_V channel, the general voltage-gated ion channel structure can be described in more detail: one monomer is formed by six *transmembrane helices*.⁷ According to their order arising in the channel they are designated as S1 – S6 (**Figure 16**, right side). Hereby the first four helices S1 – S4 constitute the *voltage-sensing domain*, and are, considering the channel complex as a whole, located at the outside of the channel, adjacent to the surrounding cell membrane (**Figure 16**, right side). A change in the electrical field of the cell membrane leads to the opening of the channel. In this context, studies showed that the more outer and closer to the membrane located starting helical segments S1 – S3 seem control the adjacent S4 helix^{134, 135} which is further responsible for performing the opening movement of the channel.^{7, 134, 135} In this respect the term *voltage-gated ion channels* is used.

The highly conserved selectivity filter mentioned above is directly located within the center of the pore duct, more precisely it is embedded in the *loop-like linker segment* which forms a connection between helices S5 and S6.⁷ However, the linker segment also contains a short helical section, the so-called *pore helix* (**Figure 16**, right side), which is why the term "*loop-like*" is used here.

The following **Figure 17** shows the transmembrane domain's three-dimensional structure from above and in side view using the example of the cryo-EM solved $K_V1.2-2.1$ paddle chimera channel (PDB ID: 6EBK¹³⁶ – full channel including transmembrane and cytosolic domain or PDB 6EBM – transmembrane domain only).¹³⁶

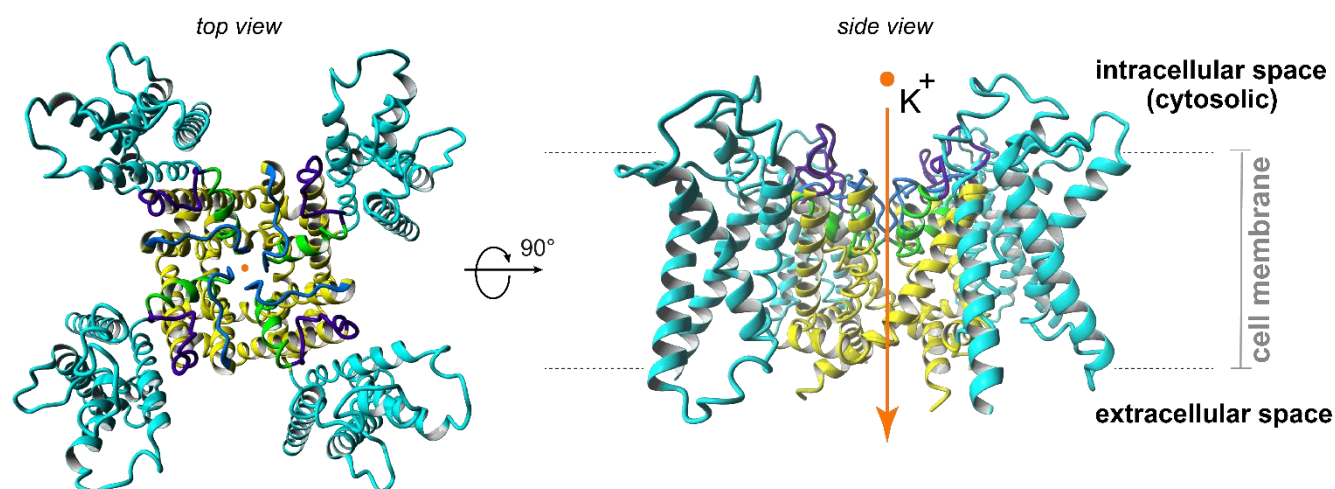


Figure 17: Voltage-gated ion channel tetrameric structure in 3D representation including all domains in top (left) and side view (right). The channel structure is shown by the example of the voltage-activated $K_V1.2-2.1$ paddle chimera potassium channel (illustration created from PDB ID: 6EBK¹³⁶). Colour scheme: voltage-sensing domain (S1-S4 helices) – cyan, S5 + S6 helices – yellow, pore helices – green, turret domain (outer loop connecting S5 and pore helix) – dark blue, selectivity filter domain (loop connecting pore helix and S6) – light blue.

4.2. Subtype-selective block of ion channels by conotoxins

Conotoxins depict small polypeptide chains, mostly being composed of around 22—30 amino acid residues,¹³⁷ obtainable from the venom of marine cone snails. Despite being obviously small, they are quite effective substances: As for the majority of drugs their main target of interaction depict membrane proteins like transporters and receptors, but also voltage-gated sodium and potassium channels.^{138, 139} The latter in turn form key players in many physiological processes such as in the electrical excitability of neural tissue. Here they control interneural signal transmission by regulating the balance between action and rest potentials in the synaptic cleft (see chapter 4.1).

Physically and specifically blocking those channels,^{137, 140} for instance by pore occlusions¹³⁷ they offer great potential for their use as analgesics.¹³⁹ Alternatively to the direct occlusions, channel deactivation can also be induced by affecting the kinetics of channel opening movements^{137, 141-145} which prevents the controlled flow of ions necessary for many physiological processes (see chapter 4.1). Neither any voltage potential, nor any de- or repolarization effects on the neural membrane are enabled which however lastly are determined by this controlled ionic flux (chapter 4.1). Disabling these essential prerequisites will stop the initiation or prevent the entire transmission of any stimuli.

Such deactivations of the natural channel function can take place in an agonistic or antagonistic manner and ultimately form the basis of analgesic effects.

It has been suggested by French and colleagues¹⁴⁶ that Na^+ currents are inhibited by μ -conotoxins in at least two ways: While on the one hand the channel vestibule is sterically blocked by the physical mass of a μ -conotoxin, basic residues on different sides of the toxin reduce the effective absorption volume of ions.

Due to their known deactivating and blocking effects on ion channels, conotoxins are already widely applied in the area of neurosciences. Their use as local anaesthetics, antiarrhythmics or anticonvulsives¹⁴⁷⁻¹⁴⁹ for instance bears witness to their neurotoxic functions. Also, various conotoxins are currently in the clinical trial test phase.¹⁵⁰

The wide range of potential conotoxin applications is, without doubt, a consequence of their intrinsic structural diversity, mainly formed by their pattern of internal disulfide bonds as a result of their high amount of cysteines.^{54, 151} For this reason, disulfide-rich conotoxins can be subdivided into superfamilies (α , μ , ω , δ and κ) based on their pattern of cystein-disulfide-bonds and further into subfamilies with respect to their interacting targets.¹⁵² Not surprisingly, the structure and flexibility of a conotoxin is also ultimately subject to its superfamily-specific disulfide bridge pattern. In view of the generally known structure-activity relationships of biomolecular complexes, it is now known that different disulfide binding patterns of one and the same amino acid primary sequence lead to different biological activities.^{151, 153-155}

The fact that the structure given by the disulfide bridge pattern is more essential in this context than the underlying sequence was further confirmed by Olivera et al. in 1991.¹⁵⁵ This study suggested degenerated pathways for marine cone snail toxins to obtain congruent conformations, whereby it turned out that the amino acid identity at certain positions is not obligatory.¹⁵⁵

It is therefore well known that the identification of an individual disulfide connectivity represents an essential part in conotoxin research. Superfamilies containing more than four cysteine residues, i.e. more than two disulfide bridges, are particularly affected, since the structural variability increases with the number of possible disulfide bridge combinations, i.e. patterns. In addition, the occurrence of several conformational isomers of proteins and peptides with several disulfide bonds has also been reported, irrespective of their origin, e.g. synthetic production, recombinant expression or isolation from biological material.¹⁵⁶⁻¹⁶¹

Conotoxins from the μ superfamily, for instance, contain six cysteines and thus a special native disulfide bridge pattern ascribed to them, which consists of three disulfide bridges. **Figure 18** gives an example of three representatives of the μ -conotoxin superfamily, showing μ -PIIIA, μ -SIIIA and μ -GIIIA as aligned amino acid sequences which will also be in the focus of our studies.



Figure 18: Sequence alignment of the μ -conopeptides μ -GIIIA, μ -SIIIA and μ -PIIIA as published by Leipold et al.⁷ Dashes indicate gaps; the cysteine-formed (native) μ -conotoxin disulfide bridge pattern present in each of the three toxins is indicated by connecting lines at the bottom of the alignment; Z = pyroglutamate, O = hydroxyproline, * = amidation; amino acid single letters are coloured by type: alkaline – light blue, acidic – magenta, polar / neutral – green, non-polar / hydrophobic – orange. (Illustration modified from Leipold et al.⁷)

These typically 16 to 25 amino acids long μ -conotoxins contain three disulfide bridges, which connecting Cys1 – Cys4, Cys2 – Cys5 and Cys3 – Cys6 (cysteines numbered in order of their occurrence in the amino acid sequence).^{162, 163} They are known for their subtype-specific blockage of Na_v channels.⁷

The venom of each cone snail species is composed by up to 200 pharmacologically active components^{152, 164} and there are indications that each species has developed its own conotoxin mixture over the course of its 55 million years old evolution.^{165, 166} Currently there are ~9.000 entries stated in the conoserver database, covering ~242 species (“<http://www.conoserver.org/>”). This manifoldness makes them at the same time so highly specific towards their targeted macromolecules.^{55, 56}

Likewise, figures confirm diversity and heterogeneity as well for the targeted potassium channels, of which more than 70 genes are known to encode for,¹⁶⁷ 40-50 of them known in the meantime to be present in humans and equally subdivided into 12 families.¹⁶⁸ To affirm the importance of specificity also on the part of the targeted ion channels, there are known numerous diseases caused by channel mutations and their concomitant malfunctioning – so-called channelopathies.^{169, 170} The hereditary disorder cystic fibrosis, but also certain types of epilepsy, deafness or cardiac arrhythmia can be mentioned as examples here.¹⁷⁰ Moreover, it is known that altered expression of $\text{K}_v1.3$ and $\text{K}_v1.5$ channels has been detected in several types of tumors and cancer cells¹⁷¹ and furthermore mutations in the $\text{Ca}_v2.1$ gene cause familial hemiplegic migraine.¹⁷²

Despite the potassium channels' high number, accompanied by their huge variety, there are only known few conotoxin family members of the equally diverse conopeptides interacting with or blocking those channels and much less their detailed mode of action (**Table 3**).

Table 3: Overview of blockage activity levels of the μ -conotoxins μ -GIIIA, μ -SIIIA and μ -PIIIA on different potassium and sodium channel subtypes of the K_v1 , K_v2 and Na_v1 family and generated chimera $K_v1.6-5P1$ and $K_v1.6-5P2$; active \triangleq subtype-specific inhibition or blockage, semi-active \triangleq weaker inhibition but partial blockage occurs, inactive \triangleq no blockage; an example of occurrence is given for each channel (CNS: central nervous system, PNS: peripheral nervous system); the last column displays the availability of experimental structural data published for the scientific community; systems treated in this work are highlighted yellow.

	μ -PIIIA	μ -SIIIA	μ -GIIIA	channel occurrence	channel structure data available
K_v1.1	active	semi-active	inactive	CNS	yes (PDB ID: 2AFL)
K_v1.2	inactive	inactive	inactive	CNS	yes (PDB ID: 2A79)
K_v1.3	inactive	inactive	inactive	CNS	yes (PDB ID: 4BGC)
K_v1.4	inactive	inactive	inactive	CNS	yes(PDB IDs: 1KN7, 1ZTO)
K_v1.5	inactive	inactive	inactive	CNS	no (model)
K_v1.6-5P2	semi-active	(presumably inactive or less active than on K _v 1.6)	(presumably inactive)	-	no (model)
K_v1.6-5P1	semi-active	(presumably inactive or less active than on K _v 1.6)	(presumably inactive)	-	no (model)
K_v1.6	active	semi-active	inactive	CNS	no (model)
K_v2.1	inactive	inactive	inactive	CNS	yes (PDB ID: 2R9R)
Na_v1.2	semi-active (no specificity but better antagonists)	semi-active (no specificity but better antagonists)	inactive	CNS	no
Na_v1.4	active (no specificity but better antagonists)	active (no specificity but better antagonists)	active (high specificity)	skeletal muscle	yes (PDB ID: 5XSY)
Na_v1.5	inactive	inactive	inactive	cardiac	yes (PDB IDs: 2KBI, 2L53, 4DCK, 4DJC, 4JQ0, 4OVN)
Na_v1.7	semi-active, (no specificity but better antagonists)	semi-active (no specificity but better antagonists)	inactive	PNS	yes (PDB ID: 5EK0)

Despite the known high specificity and diversity associated with conotoxins, the underlying mechanisms of mutual interactions with their targets are not yet fully understood. **Table 3** points out the still perpetual lack of knowledge considering subtype-specific channel blockage by μ -conopeptides and leaves very demanding and costly issues for the researcher when trying to explicitly explain a blockage event or the inactivity of a system by reasons of distinct interactions. Also and as a result, trying to detect regularities or similarities of blockage or inactive systems would require an all-system comparison with respect to individual binding modes and the concomitant dynamics. However, a more general determination of different classes of blockage mechanisms could be provided through this approach. **Table 3** shows that this is very difficult to be unambiguously grasped: readily apparent, a particular conotoxin agent does not necessarily have to block a single channel subtype only, but is for instance also able to semi-actively bind to other channel subtypes than to its original target. The mutual blockage activities are therefore not precisely delimitable and there are fluent transitions observable in these respects. Unfortunately, various and severe adverse events may be caused by such unspecific inhibitory effects on diverse ion channel subtypes apart from the “actual” target channels⁷ or even by other off-target

interactions. Due to the high analgesic efficacy of the conotoxins mentioned above, these relationships are particularly relevant for the field of analgesic treatment.

The representative ω -MVIIA from the ω -conotoxin superfamily, also known as ziconotide, can be cited as an example of painkilling drug associated with various side-effects that are suspected to be caused by such “off-target” effects. Severe neurological side effects¹⁷³⁻¹⁷⁵ like dizziness, hallucinations, abnormal gait, ataxia, confusion, nystagmus, somnolence¹⁷⁶⁻¹⁷⁸, but also nausea and/or vomiting¹⁷⁶ and even potential psychiatric side effects^{178, 179} are thus far known as side effects associated with ω -MVIIA treatment. However, besides its intrinsic use as blocker of the N-type calcium channel Cav2.2 it thus far is only known to additionally show slight binding activities to the P/Q-type calcium channel Cav2.1^{177, 180-182} and, as analgesic it is still approved and applied as a drug for the treatment of severe chronic pain.^{183, 184}

4.3. In silico binding studies of μ -PIIIA isomers on the sodium channel Nav1.4

As already mentioned, the disulfide bond pattern is suggested to have a dramatic impact on the fold and the global shape of the conotoxin. As one can imagine, changing the disulfide pattern will also change the affinity towards the toxins target. As this was demonstrated for a conotoxin harbouring two disulfide bridges, it remained unclear, how the structure and activity of a conotoxin with three disulfide bonds is altered.^{151, 153-155} As μ -PIIIA, which contains three disulfide bonds was suggested to represent a potential drug lead the structure and bioactivity of its 15 possible isomers was of special interest and the focus of our investigations.

Studies on natural peptides with at least three, or broadly speaking more than two disulfide bridges and one pair of adjacent cysteines are particularly interesting because they have been shown to have a high therapeutic potential or important biological functions. Examples are the vascular endothelial growth factor A,¹⁸⁵ resistin,¹⁸⁶ Norrin,¹⁸⁷ insulin-like growth factor I,¹⁸⁸ placenta growth factor,¹⁸⁹ sarcospan,¹⁹⁰ hepcidin,¹⁹¹ defensins¹⁹² and in part Kunitz-type protease inhibitors.¹⁹³⁻¹⁹⁵

The elucidation of disulfide connectivity is therefore also one of the main objectives in current proteomics and bioanalysis studies of protein and peptide therapeutics, such as monoclonal antibodies.¹⁹⁶

Of course, the complexity of the possible structural arrangements also increases with the number of disulfide bridges which is a challenge for related research efforts: it was found that for about 30% of all NMR structures stored in the PDB database there is no nearer specified disulfide pattern present¹⁹⁷ and furthermore, only studies on structural and functional effects of disulfide binding exclusively on peptides with only two disulfide bridges are known to date (albeit in large numbers).^{153, 154, 198}

In-depth insights into the correlations of cysteine-disulfide patterns with their carrying conopeptides' activity – especially that of more complex bonding patterns and architecture – will therefore provide important foundations for the rational design process of subtype-selective Nav channel addressing blockers. Since, experimental strategies addressing the structural analysis of cysteine-rich peptide isomers of intrinsic high conformational variability are largely missing,⁵⁴ computer-based approaches will provide vital supports in these regards.

This computational study was part of a highly collaborative project involving various groups at the University of Bonn (Imhof, Biswas, synthesis and MS characterization), Jena (Heinemann, electrophysiology), Lübeck (Leipold, Schmitz, electrophysiology) and from the Technische Universität Darmstadt (Tietze, modelling and NMR characterization). Recently, Heimer and co-workers⁵⁴ succeeded to synthesize the fifteen possible disulfide isomers of the 22mer conopeptide μ -PIIIA.⁵⁴ Applying a combination of various analytical methods such as 2D NMR spectroscopy and tandem MS further gave remarkable insights in the impacts on disulfide connectivity on a toxin's global fold provided by this enlightening research study.⁵⁴ The 3D structures of these 15 isomers, represented by their corresponding NMR ensembles (**Figure 19b**), strongly suggested that the global fold of a toxin is a direct consequence of its disulfide bond pattern.

a)

ZRLCC GFOKS CRSRQ CKOHR CC-NH₂

b)

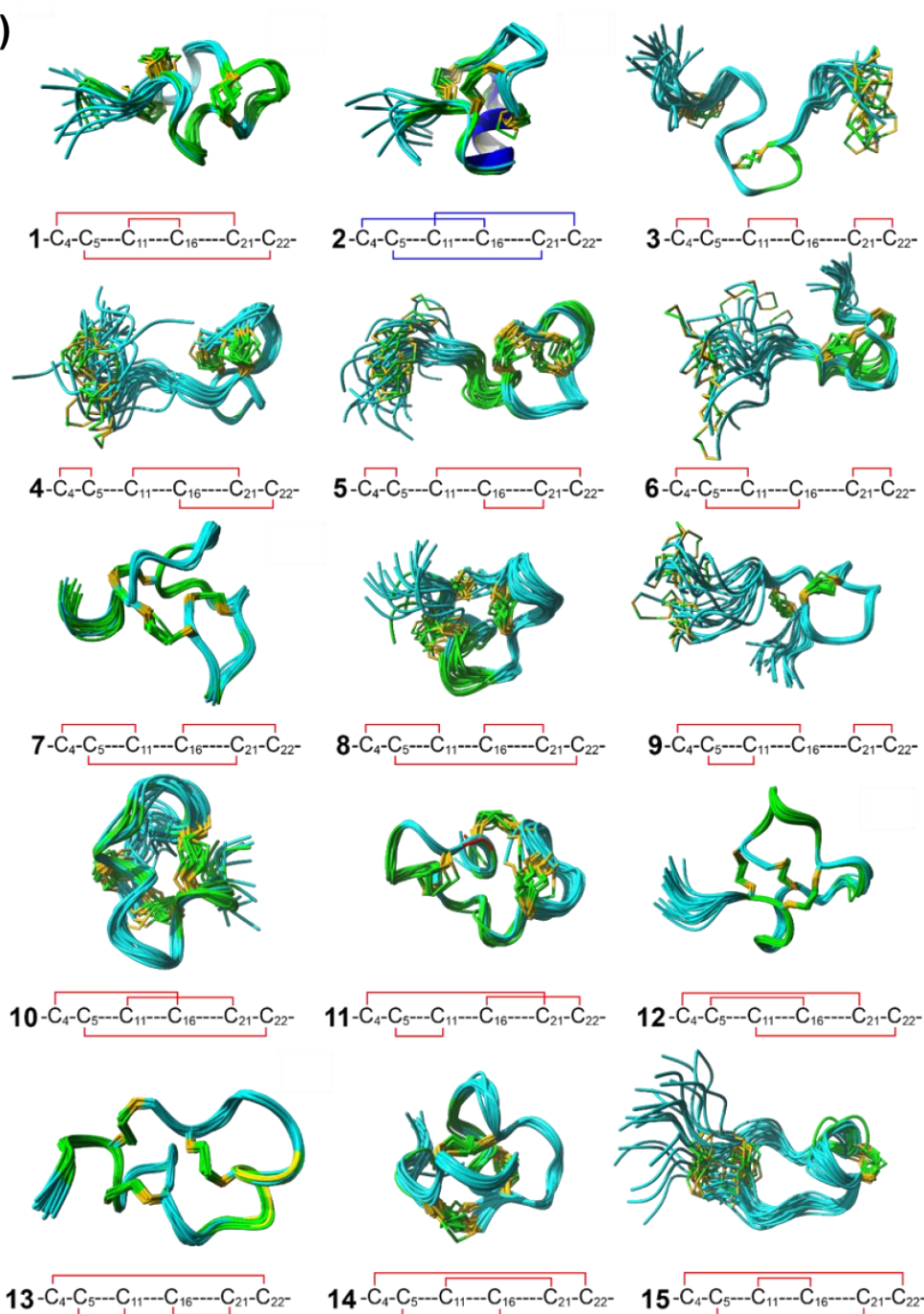


Figure 19: (a) μ-PIIIA sequence. (b) NMR structures of all 15 μ-PIIIA isomers. For each isomer with the indicated disulfide-bond pattern, the solution NMR structures are respectively presented as ensembles of the 20 lowest-energy structures; structure 2 represents the so-called *native* μ-PIIIA isomer.¹⁹⁹ (Image source: Heimer et al.⁵⁴)

The capability of the different μ-PIIIA isomers to block the Nav1.4 channel was assessed by current measurements with the whole-cell patch clamp method (data kindly provided by E. Leipold and S. S. Heinemann). The impact of the toxin isomers on depolarization-elicited Na⁺ inward currents is expressed as the time course of current block and the maximal block after long toxin exposure (**Figure 20**). Accordingly, currents were diminished by all 15 isomers, but to a different extend (**Figure 20**).

At a concentration of 10 μM , isomer **2**, which harbours the native disulfide bond pattern of $\mu\text{-PIIIA}$ (**Figure 19b**, blue marked disulfide pattern indications), was the most potent in inhibiting Nav1.4 channels, followed by isomers **15**, **11**, **14**, **1**, **7**, and **12**.

Interestingly, all of these pore blocking isomers have a disulfide bridge pattern in which (seen over the sequence) an overlap of all (three) possible disulfide bridging pairs is shown (**Figure 19b**).

This may allow a particularly high stability in the spatial arrangement, which in turn can favour an efficient block, provided that a good fit is given on the channel surface.

Further concentration dependence analyses yielded an apparent IC_{50} of 105.3 ± 29.9 nM for the block of Nav1.4 by isomer **2** (**Figure 20b**). However, even at a saturating concentration of 100 μM , a small current component of $7.2 \pm 2.2\%$ remained for some isomers, showing that even by blockage through isomer **2**, the channel is not covered to such an extent that the sodium ion flow is completely eliminated (**Figure 20a, b**). Additionally, the start time for the onset of the pore block was estimated with single exponential functions and yielded the fastest inhibition reaction for isomer **2** compared to the other, much slower interacting isomers **15**, **11**, **14**, **1**, **7**, and **12** of lower blocking potency (**Figure 20c**). The possibility of a precise evaluation of the channel block at concentrations lower than 10 μM and associated IC_{50} values was thus excluded. However, sticking thus to the assessment at higher concentrations, it turned out that, despite their different disulfide connectivity patterns (**Figure 19b**), isomers **15** and **7** were especially interesting from the measurements' point of view due to their almost identical total current block at 10 and 100 μM (**15**: $48.2 \pm 5.9\%$ and $51.1 \pm 2.2\%$, respectively; **7**: 32.3 ± 4.3 and 35.1 ± 4.6 , respectively) (**Figure 20c**).

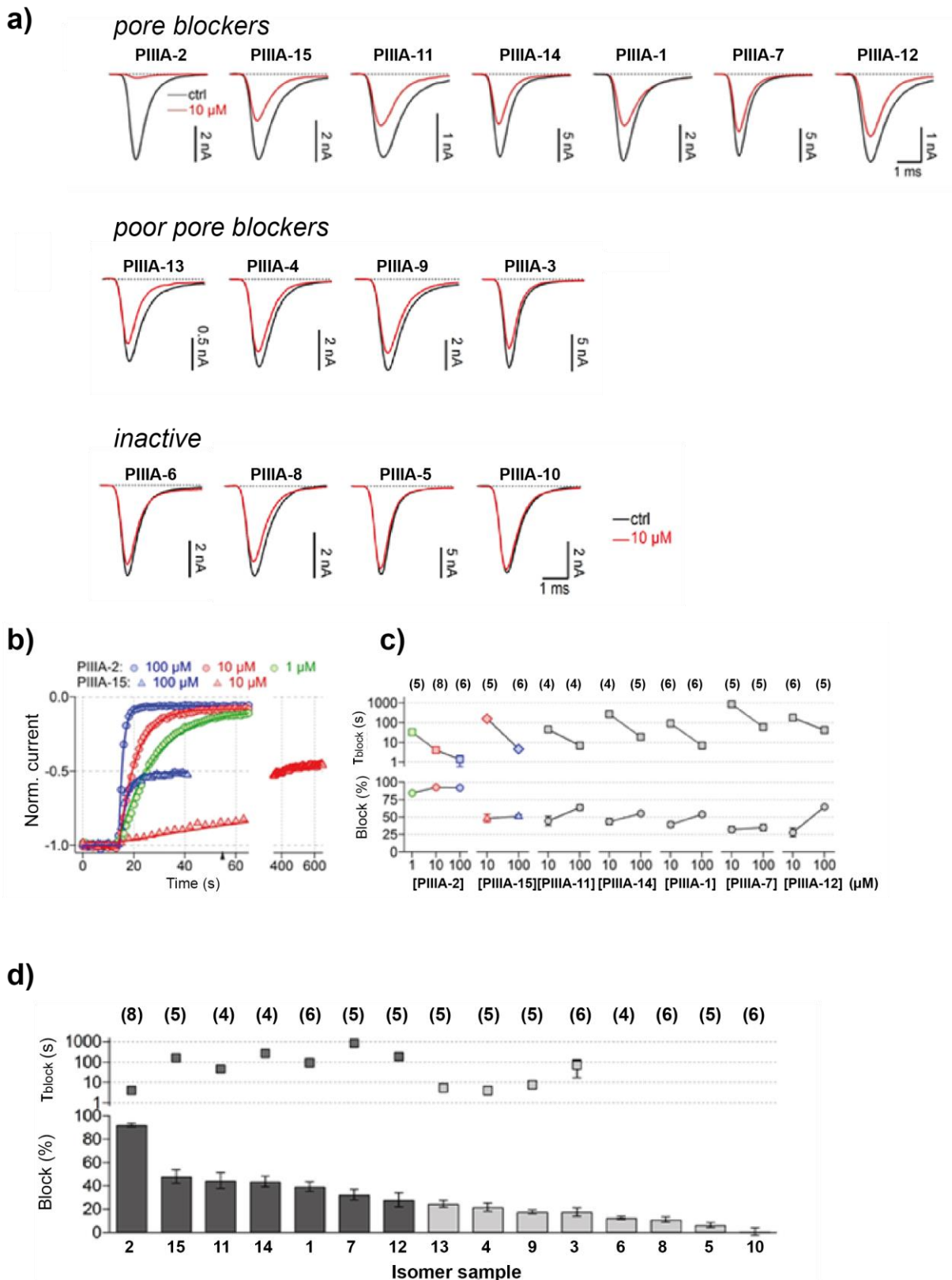


Figure 20: Block of Nav1.4-mediated currents by disulfide isomers of μ -PIIIA. **(a)** Representative current traces of transiently expressed Nav1.4 channels evoked at a test potential of -20 mV before (black, control) and after (red) application of 10μ M of the indicated μ -PIIIA isomers. **(b)** Normalized peak current amplitudes obtained from repetitively evoked current responses were plotted as a function of time to follow the time course of current block mediated by various concentrations of μ -PIIIA isomers 2 and 15. Continuous lines are single exponential data fits used to characterize the onset of channel block. The time axis was split to illustrate that channel block by isomer 15 saturates at about 50% suggesting that isomer 15 seals the channel pore only partially. In contrast, isomer 2-mediated channel inhibition saturates at about

95% with an IC_{50} of 105.3 ± 29.9 nM. **(c)** Steady-state block (bottom) as well as the associated single exponential time constant (top), describing the onset of channel block, for various concentrations of the indicated isomers. Lines connect data points for clarity. Symbols and colour-coding of data obtained with isomers 2 and 15 are as in b). **(d)** Steady-state block of $Nav1.4$ -mediated currents by $10 \mu M$ of the indicated μ -PIIIA isomers (bottom) and the associated single exponential time constant, τ_{block} , describing the onset of block (top). Numbers of individual experiments, n , are provided in parentheses. (Data kindly provided by E. Leipold and S. Heinemann)

Analyses of the isomers 3-6, 8, 10, and 13, performed at $10 \mu M$, identified isomer 13 as the still most active isomer among this group, showing an pore block of $24.6 \pm 3.0\%$ (**Figure 20**). Isomers 6, 8, 5, and 10 reduced $Nav1.4$ -mediated currents by less than 15% (**Figure 20**), and were thus considered as inactive peptides.

In summary, isomers 2, 15, 11, 14, 1, 7, and 12 form the group of seven active isomers, isomers 13, 4, 9 and 3 constitute another group of weaker but still moderately active pore blockers, and finally the remaining isomers 6, 5, 8 and 10 represent the group of inactive compounds (**Figure 20**).

Accordingly, the group of active isomers generally shows a more complex disulfide bridge patterns in which all three disulfide bridges either overlap at the same time (as for example with isomer 14) or at least still transitively a certain connection among all three exists via a disulfide bridge which crosses the other two separate bridges (as e.g. in the example of isomer 11) (**Figure 19b**). This connectivity pattern, which affects the entire peptide sequence, is supposed to be the reason for the low overall flexibility of the toxin as can be accessed from the NMR structural ensemble (**Figure 19b**). This is further supported by the fact that all these isomers contain at least one disulfide bond that connects the N- and C-terminal sections of μ -PIIIA (**Figure 19b**). (Isomer 10, which nevertheless exhibits a moderately high flexibility with simultaneously low activity, can be regarded as a special case in this context.)

In the moderately active isomers 13, 4, 9 and 3, isolated or separated disulfide bridges occur more frequently, i.e. they do not overlap with any other disulfide bridges (**Figure 19b**). This is particularly the case with the more (terminal) flexible isomers 4, 9 and 3 (**Figure 19b**). Nevertheless, the existing overlap of isomer 13 with the two other internal disulfide bridges, which, similar to the interconnectivity of the more active isomers described above, extends over the entire toxin sequence. This also demonstrates that isomer 13 is the only semi-active isomer that has an N- to C-terminal disulfide bridge, causing a higher rigidity of the toxin compared to the other semi-active representatives (**Figure 19b**, **Figure 20**). This also gives a possible and plausible explanation for the isomer 13 blocking activity, which is still highest within all moderate (and inactive) isomers.

The low active or inactive isomers 6, 5, and 8 (with the exception of the special case isomer 10) continue this pattern of containing more isolated, i.e. non-overlapping disulfide bridges as far as possible together with the occurrence of higher intrinsic flexibilities (**Figure 19b**). Isomers 6 and 5 have particularly flexible N- and C-terminal regions (compared to isomers 8 and 10) due to a connection between consecutive cysteines. In summary, these structure-activity correlations give evidence that the inhibitory capacity of μ -PIIIA isomers for the blockage of $Nav1.4$ channels increases with their rigidity (and decreases with increasing isomer flexibility), which can directly be derived from the structural deviation of the NMR structural ensemble (**Figure 19b**).

In order to gain deeper insights in how the pore block is achieved by the different μ -PIIIA isomers, a detailed computational analysis was performed.

Development of a $Nav1.4$ homology model

Closer inspections of the particular μ -PIIIA isomers' interactions with the $Nav1.4$ channel were performed on the basis of a homology model of the channel which, similar to another model that has been previously established by us¹⁵¹ and others,²⁰⁰ merely encompassed the central unit of the pore. For the homology model a recently published structure of the voltage-gated calcium channel $Cav1.1$ (PDB ID 3JBR²⁰¹) was defined as a suitable template for the creation of the $Nav1.4$ model. In this respect, it must also be mentioned that the recently resolved structure of the human voltage-gated $Na1.4$ channel (PDB ID 6AGF²⁰²) was not yet available at the beginning of this study. However, the chosen template also enabled us to model even the large pore-surrounding extracellular loops, which is usually a rather demanding task in homology modelling of proteins and was not included in the models generated so far. The inclusion of the extracellular loop parts also provided an advantageous basis for further analyses, since especially the extracellular loops are known to contain important interaction and recognition sites for various μ -conotoxins. (For more detailed information on the modeling process, please refer to chapter 5.3 in the methods section.)

Figure 21: Alignment of the pore modules of the channel template $Ca_v1.1$ (PDB ID 3JBR²⁰³, uniprot ID P07293) and the target sequence of $Nav1.4$ (uniprot ID P35499) of all four domains. Conservation of the aligned residues is indicated in the histogram below the sequence and given as a numerical index reflecting the conservation of physico-chemical properties in the alignment employing the AMAS method²⁰⁴: identities score highest and the next most conserved group contain substitutions to amino acids lying in the same physico-chemical class. Regions from the template structure that were part of the final homology model are indicated by secondary structure elements, schematically outlined above (transmembrane segments – black box, loops - straight black line) the aligned amino acid sequences. Homology-modeled regions within the template structures are color-coded. Sequence alignments were created with Clustal Omega²⁰⁵⁻²⁰⁷ on www.uniprot.org. The alignment figure was created with JALVIEW 2.8.²⁰⁸

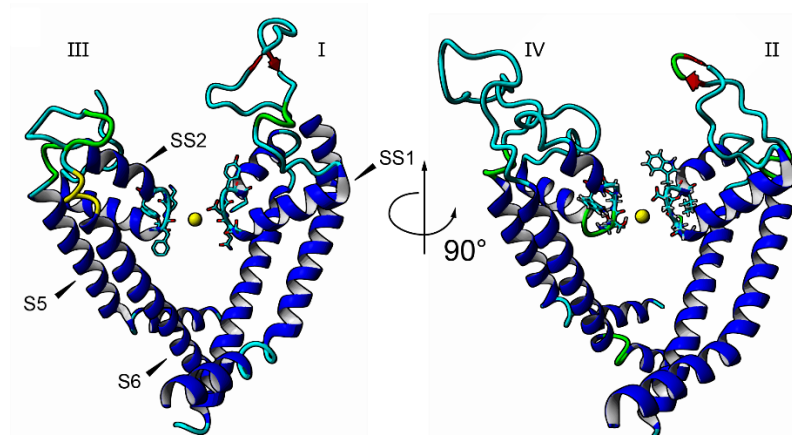


Figure 22: Homology model of $Nav1.4$, selectivity filter residues are shown in stick representation (coloring scheme: carbon – cyan, nitrogen – blue, oxygen – red, calcium – yellow), subunits are indicated with roman letters.

Docking of the μ -PIIIA isomers to $Nav1.4$

Employing the previously generated homology model of the voltage-gated sodium channel $Nav1.4$, the 15 μ -PIIIA isomers (**Figure 19b**) were docked to the channel model using the HADDOCK easy web interface.^{209, 210} (For more details see chapter 5.4 the methods section.) The docking clusters were generated on the basis of the interface-ligand RMSDs of the individual docking results and then scored according to the HADDOCK scoring function yielding individual Z- and HADDOCK scores assigned to each cluster (**Table 4**). The individual best scoring channel-toxin complexes of the respective highest scoring cluster were subjected to further analysis (**Table 4**, bold lines).

Table 4: Summary of the HADDOCK docking runs. For each isomer docking run the three best-scoring clusters are outlined, starting with the highest scored (i.e. lowest Z- and HADDOCK score) cluster. Docking results relating to highly active isomers are highlighted in light green (very highly active isomer 2 - darker green), moderately active isomer results are highlighted in light yellow and docking results of weak or inactive toxin isomers are highlighted in light orange.

Isomer	Z-score	HADDOCK score	Cluster size	RMSD from the overall lowest-energy structure [Å]	Buried surface area [Å ²]
1	-1.2	174.1 ± 6	25	3.2 ± 0.0	3114.0 ± 168
	-1.2	175.3 ± 17	19	0.7 ± 0.5	2888.9 ± 212
	-0.9	181.9 ± 4	60	2.8 ± 0.1	2909.7 ± 72
2	-1.4	162.9 ± 20	41	1.3 ± 0.8	2877.3 ± 134
	-0.9	182.3 ± 5	18	2.5 ± 0.1	2961.9 ± 146
	-0.8	183.6 ± 13	26	2.2 ± 0.2	2817.6 ± 121
3	-1.6	127.2 ± 11	45	1.2 ± 0.7	3349.9 ± 140
	-1.5	129.7 ± 2	92	3.7 ± 0.1	3435.1 ± 51
	-0.8	151.2 ± 24	9	3.1 ± 0.0	3185.7 ± 201
4	-2.1	139.5 ± 5	89	1.4 ± 0.9	3239.6 ± 43
	-0.9	181.0 ± 3	20	3.5 ± 0.3	3078.0 ± 98
	-0.3	202.0 ± 5	33	3.3 ± 0.1	2906.4 ± 142
5	-0.9	158.8 ± 3	41	3.2 ± 0.0	3237.1 ± 51
	-0.6	168.9 ± 11	55	3.1 ± 0.2	3038.7 ± 97
	-0.6	169.6 ± 35	54	0.9 ± 0.6	3040.3 ± 44
6	-1.8	132.2 ± 8	42	0.5 ± 0.3	3360.5 ± 98
	-1.0	156.1 ± 16	19	3.0 ± 0.1	3120.9 ± 124
	-0.8	163.2 ± 11	19	2.9 ± 0.0	3254.4 ± 35
7	-1.2	171.8 ± 7	38	0.7 ± 0.5	2913.6 ± 123
	-1.0	177.3 ± 10	40	2.5 ± 0.0	3028.1 ± 123
	-0.6	187.5 ± 7	18	2.5 ± 0.1	3056.5 ± 164
8	-1.4	151.5 ± 8	14	2.6 ± 0.1	3083.6 ± 141
	-1.0	162.9 ± 16	14	0.8 ± 0.5	3004.7 ± 138
	-0.4	180.4 ± 8	58	2.4 ± 0.0	3120.0 ± 59
9	-1.3	162.1 ± 20	19	1.2 ± 0.8	3262.9 ± 243
	-0.9	169.3 ± 7	28	2.3 ± 0.2	3168.8 ± 80
	-0.9	170.1 ± 23	40	3.0 ± 0.1	2912.1 ± 270
10	-1.6	142.2 ± 28	25	0.9 ± 0.7	3134.1 ± 52
	-0.9	165.1 ± 4	10	2.3 ± 0.1	3022.5 ± 110
	-0.8	169.6 ± 18	28	2.2 ± 0.1	3057.6 ± 79
11	-2.2	131.2 ± 21	28	0.5 ± 0.3	3019.3 ± 99
	-0.8	164.1 ± 14	24	3.0 ± 0.0	3123.3 ± 62
	-0.4	174.4 ± 14	30	2.2 ± 0.1	2967.6 ± 34
12	-1.9	123.5 ± 11	82	0.8 ± 0.5	3220.7 ± 55
	-0.7	169.6 ± 11	37	2.4 ± 0.0	3021.8 ± 110
	-0.7	171.4 ± 14	21	2.8 ± 0.1	2903.0 ± 155
13	-1.0	193.0 ± 19	22	2.8 ± 0.0	2913.5 ± 108
	-0.7	201.4 ± 29	78	1.1 ± 0.9	2858.1 ± 161
	-0.6	204.5 ± 4	21	2.1 ± 0.0	2891.9 ± 44
14	-1.8	181.0 ± 12	89	0.7 ± 0.5	2830.2 ± 41
	0	211.4 ± 6	45	2.5 ± 0.0	2713.1 ± 98
	0	211.7 ± 21	34	2.0 ± 0.2	2679.1 ± 126
15	-1.5	162.2 ± 21	65	1.2 ± 0.7	2966.4 ± 186
	-1.3	168.1 ± 12	28	2.0 ± 0.1	3096.8 ± 29
	-0.9	179.5 ± 13	9	3.4 ± 0.1	3146.3 ± 147

At first glance, a pore-covering situation on the channel appears for all docked isomers, also with respect to the central pore interface (**Figure 23**, **Figure 24**), impeding a straight forward rationalization of the toxins' individual bioactivities solely from the docked states. Moreover absolute pore block values under saturating concentrations at 100 μM were only determined for isomers **2**, **15** and **7**, while all other isomers did not show

saturation at 100 μM (**Figure 20**). (Higher concentrations could technically not be measured.) We will therefore orient our analyses to the docking predictions of isomers **2**, **15** and **7** and use these as references.

Upon closer inspection of the respective binding poses, the commonality of the arginines of these isomers, which simultaneously protrude into the pore and bind to the outer regions, can be determined (**Figure 23a, b, f**). Thus, our docking results suggest the presence of an arginine that protrudes more deeply into the central pore duct, together with a simultaneous contact of the outer pore area (also by arginines or basic residues) to be critical for a proper pore-blocking formation. The latter, outer located pore surface area mainly concerns the acidic residues in the S5-SS1 and SS2-S6 pore loops that might, together with the high number of acidic residues within these loop regions (**Figure 21**), indicate a key role for toxin affinity.

When viewing all complexes from the top it can be noted that the binding of the toxin can be determined not only by the H bond contacts with the channel, but also by the orientation of its arginine residues. Here generally, arginine residues (Arg2, Arg12, Arg14, Arg20) turn out to be the key interacting residues, predominantly binding to aspartic acid in the loop regions (**Figure 23, Figure 24**). Also important are positions of the equally basic lysines (Lys9, Lys17) and Pyr1 with respect to the channel (**Figure 23, Figure 24**). The orientation of these residues is however not least a consequence of the spatial isomer structure predetermined by the inherent disulfide bridges.

In particular isomer **2** as the most active blocker can be observed to reside closest to the selectivity filter residues located at the center of the pore, occluding it by Arg20, while Arg14 and Arg2 contact the outer residues of the subunits (**Figure 23a**). Moreover, isomer **2** shows an even distribution of all its arginines over the whole surface of the channel (**Figure 23a**).

In contrast, bound arginines of the majority of moderately active and inactive isomers can be stated to point away from the pore centre, being rather directed to the channel interfaces and no simultaneous contacts to the pore duct and the outer pore loops arise (**Figure 24**).

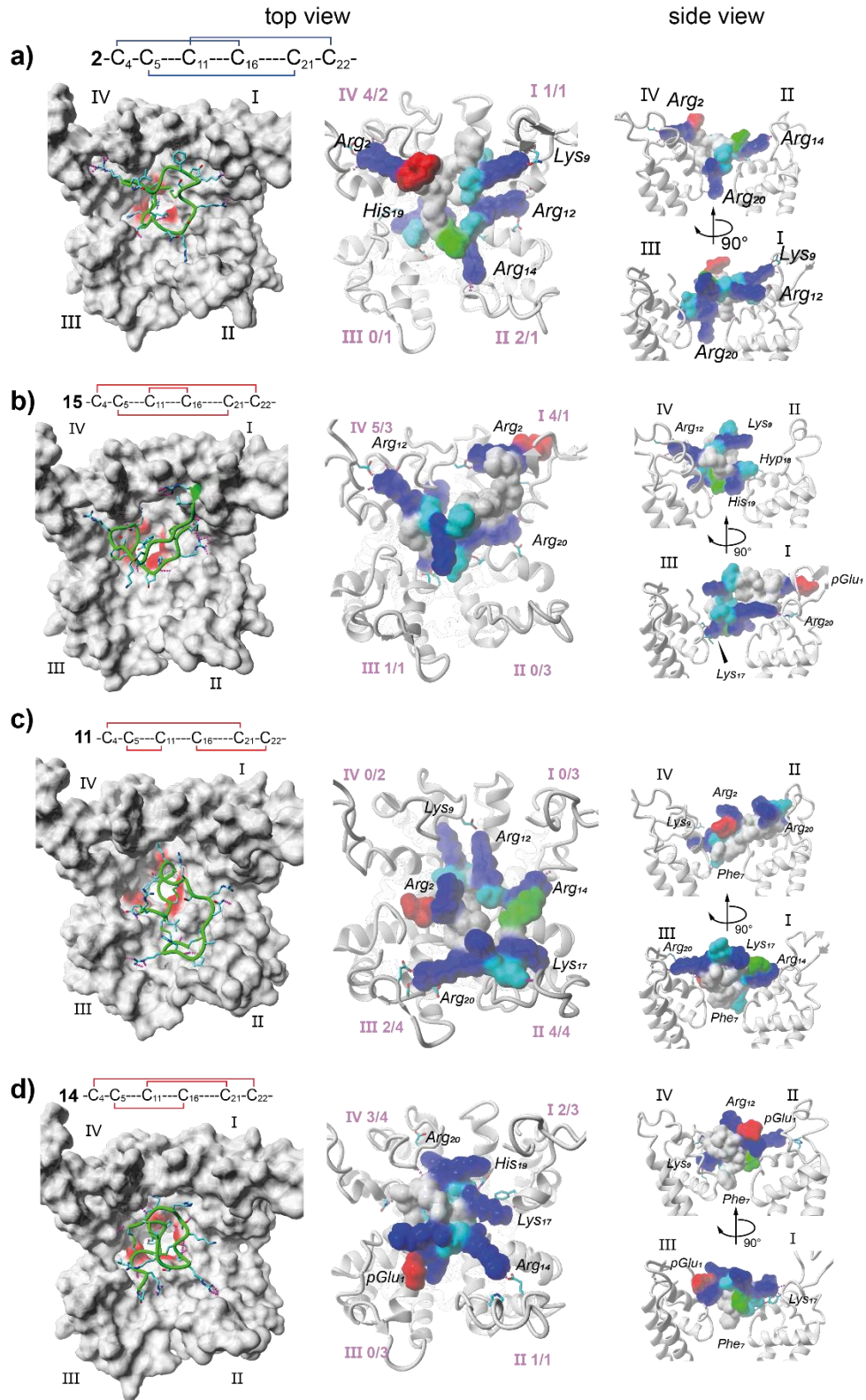
The predicted binding poses of the high potency isomers **2**, **15** and **7** might explain the decreasing blocking ability by showing less hydrogen bonds with residues close to the selectivity filter for **15** and **7** compared to the isomer **2** (**Figure 23, Table 5**). At the same time, for isomer **15** and **7** a smaller and less even spread of the outer binding arginine residues compared to the native isomer **2** can be observed (**Figure 23**). Here, isomer **15** is oriented a bit more towards domains I and IV and isomer **7** clearly points more towards the opposite moiety (domains II and III) of the channel (**Figure 23**), which nicely explains their partial pore block.

Generally, for all pore blocking isomers a fairly even distribution of their basic or arginine residues across the pore entrance of the channel was observed (**Figure 23**), which was somehow lacking for the less active and inactive isomers (**Figure 24**).

Interestingly, among the remaining moderately blocking isomers (**11**, **14**, **1** and **12**), similar formations of simultaneous pore-diving and outer contacts as observed for isomer **2** can also be stated for isomers **1** and **12** (**Figure 23**). Although isomers **11** and **14** still have an even pore coverage on the surface, they lack the contact of a basic residue to the inner pore center. (**Figure 23**).

Among the the poor pore blockers (**13**, **4**, **9** and **3**) isomer **13** is observed to be the only isomer exhibiting simultaneous contacts to the pore outside and the centre (**Figure 24**). Notably, isomer **13** represents the still most potent isomer within this group (**Figure 24**). However, no even distribution of arginine or basic residues is observable in this group (**Figure 24**). Moreover, while isomers **4**, **9** and **3** all lack the contact to the central pore duct, it can further be noted, that all their central and terminal arginine or basic residues (especially Arg2, Lys9, Arg12, Arg14, His19 and Arg20) address the outer pore loops of the channel surface, preventing the toxin to establish the simultaneous contact to the pore centre (**Figure 24**). Particularly, isomer **4** shows an inhomogeneous one-way orientation for its majority of arginine residues towards the pore loops linking transmembrane segments S5 with SS1 and SS2 with S6 of domains I and IV (**Figure 24**). It further shows a separate positioning of Arg2 together with Pyr1 away from the other arginines and towards pore domains III and II (**Figure 24**).

As the remaining isomers **6**, **8**, **5** and **10** are considered as inactive, a bound state in this sense is not assumed and therefore the docked binding poses of these isomers are not further discussed in more detail.



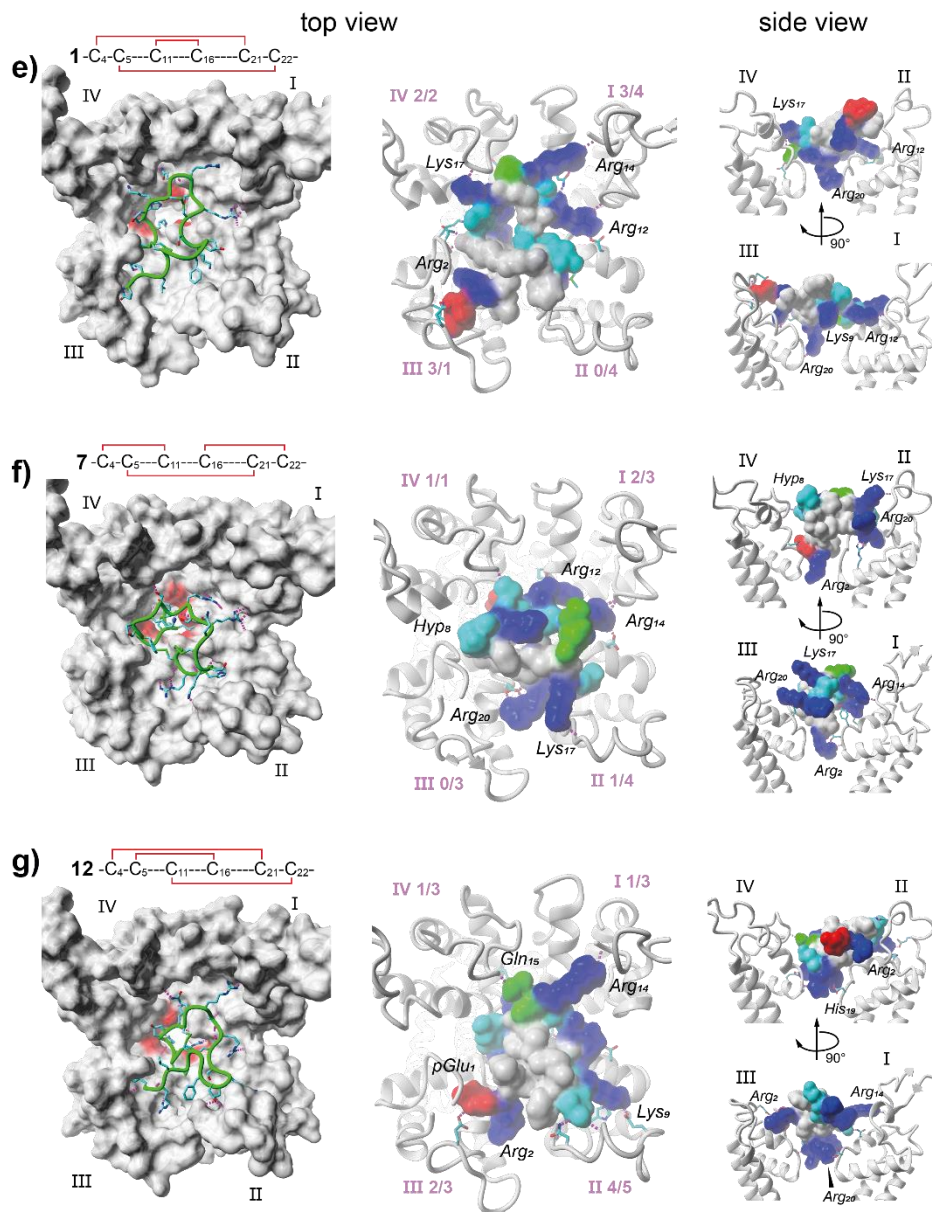
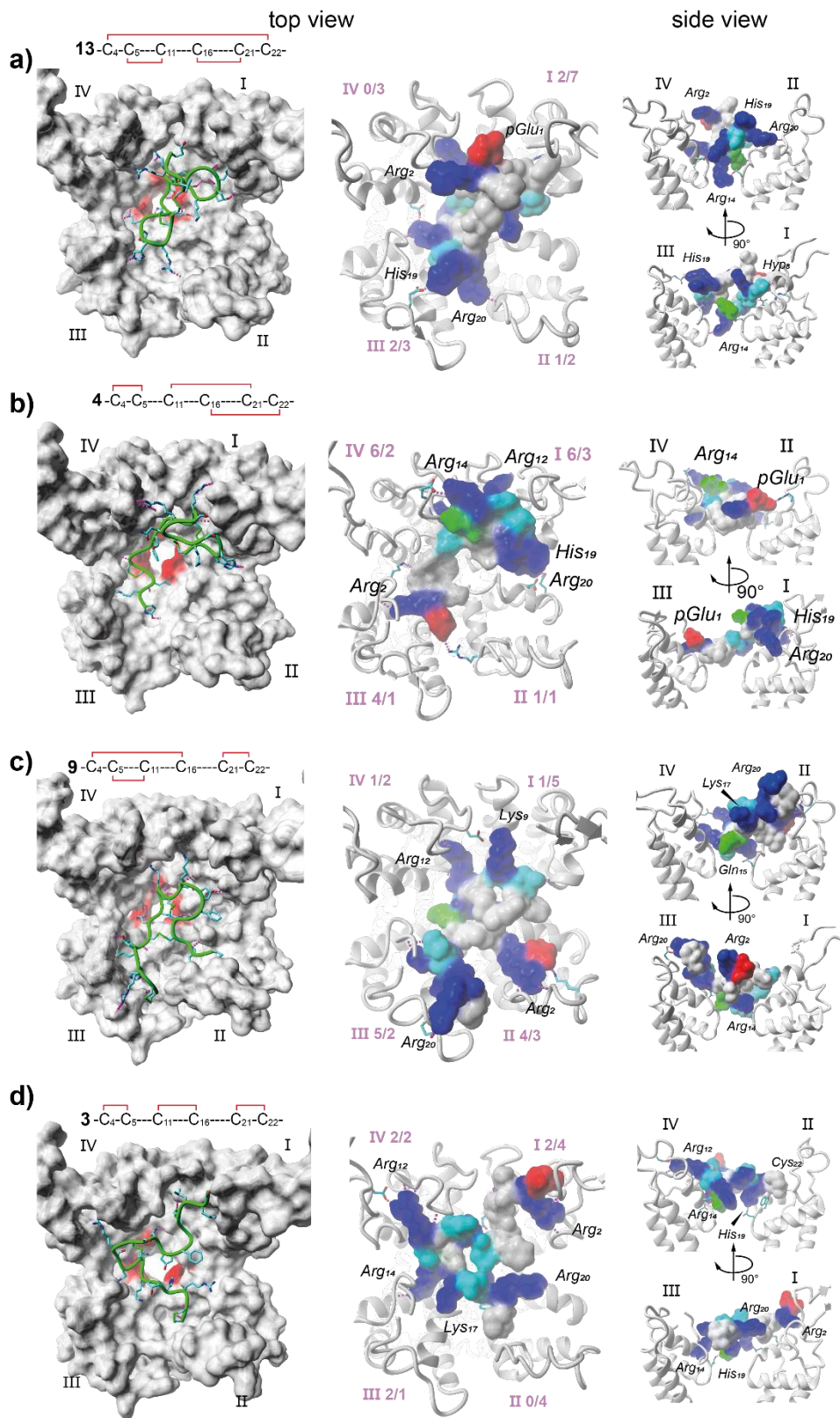


Figure 23: Visualization of the best scoring μ -PIIIA–Nav1.4 complex conformations obtained from our docking experiments for pore blocker isomers (**a**, **b**, **f** – high potency; **c** – **e**, **g** – moderate potency). The four Nav1.4 domains are indicated by purple Roman numbers. Left panel – top view of the toxin-channel complex. The Nav1.4 channel surface (Van der Waals) is given in gray, the selectivity filter motif (DEKA) is highlighted in red. The toxin is shown in green with side chain atoms present (coloring scheme: carbon – cyan, nitrogen – blue, oxygen – red, sulfur – green). Middle panel – top view, right panel – side view of the toxin-channel complex. The toxin is shown in surface representation (Van der Waals) and colored according to amino acid properties (coloring scheme: basic – blue, acidic – red, polar – cyan, non-polar – gray). The number of hydrogen bonds formed by the toxin with each channel domain (loops/SS1-SS2) is highlighted in purple distinguishing between hydrogen bonds formed with the TM bridging loops (S5-SS1, SS2-S6) and the pore helices (SS1-SS2).



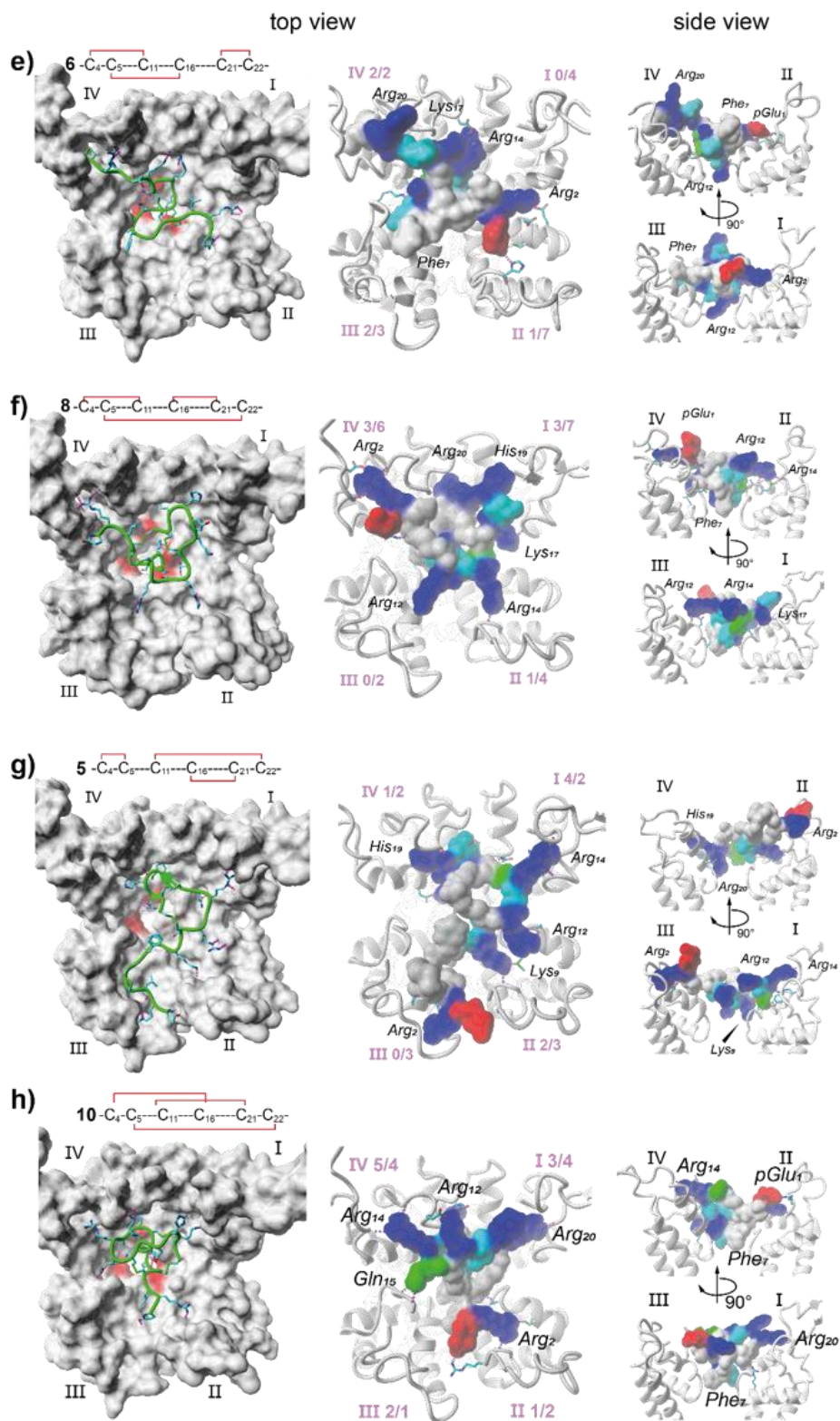


Figure 24: Visualization of the best scoring μ -PIIIA–Nav1.4 complex conformations obtained from our docking experiments for poor pore blockers (a - d) and inactive isomers (e - h). The four Nav1.4 domains are indicated by purple Roman numbers. Left panel – topview of the toxin-channel complex. The Nav1.4 channel surface (Van der Waals) is given in gray, the selectivity filter motif (DEKA) is highlighted in red. The toxin is shown in green with side chain atoms present (coloring scheme: carbon – cyan, nitrogen – blue, oxygen – red, sulfur – green). Middle panel – top view, right panel – side view of the toxin-channel complex. The toxin is shown in surface representation (Van der Waals) and colored according to amino acid properties (coloring scheme: basic – blue, acidic – red, polar – cyan, non-polar – gray). The number of hydrogen bonds formed by the toxin with each channel domain (loops/SS1-SS2) is highlighted in purple distinguishing between hydrogen bonds formed with the TM bridging loops (S5-SS1, SS2-S6) and the pore helices (SS1-SS2).

However, our observations on the variety of binding poses are also consistent with further H bond analyses that suggest an individual distribution of hydrogen bonding contacts towards the channel. Upon calculating the percentage hydrogen bond distributions over the different halves of the channel – assuming that one half is defined by the adjacent domain pair with the largest proportion of hydrogen bondings – the general tendency of the toxin to be positioned more to one side on the pore for more inactive toxin isomers can be derived (**Table 5**), which is compatible with the above stated findings.

Conclusively, in order to rationalize the amount of pore block and bioactivity through computer-aided representations, information about the flexibility of the individual toxin segments is necessary. Although such approaches would be possible with simulation studies, one can already derive from our docking data that for some isomers (especially the poor binders), the pore is only covered by the flexible molecule part (**Figure 19**, **Figure 24**). Thus, an estimation how tightly a toxin part is able to bind to the channel can therefore be attributed to its particular flexibility.

Finally, the unique isomer **2** and its noticeable binding mode should be mentioned again in particular, as its bound conformation (**Figure 23a**) is also consistent with the significance of distinct μ -PIIIA residues, e.g. Arg14 and Arg2^{211, 212} (although Arg14 does not protrude into the pore). Another interesting aspect is the rationalization of available mutation data^{211, 212} by our docked isomer **2** on the Nav1.4 model, involving all basic residues in hydrogen bonding and keeping them distributed among the entire channel surface. This in turn agrees with the findings on the impact of basic residues in ion channel blocks, as formulated by French and colleagues.¹⁴⁶ Furthermore, it can be concluded from this that the inhibitory effect of μ -PIIIA is not exclusively achieved by a single residue critical for a block, but always a combination of the interaction of several residues with the channel is essential for blocking. At the same time, this equally means that the inhibition of the current does not follow an all-or-nothing principle, but is rather gradually determined by distinct interaction modes. The special fact that the inhibitory binding of the μ -PIIIA isomer **2** to Nav1.4 (as well as the blockage of Kv1.1 and Kv1.6 by μ -PIIIA) is determined by simultaneously existing contacts to the inside and outside pore area could not be fully rationalized by previous models so far, as in the respective homology model proposals the external pore loops were not considered.^{200, 211, 213, 214}

The coincident contacts of μ -PIIIA with the outer surface area of the Nav1.4 channel beyond the contact protruding into the pore through Arg20 is also shown by the slightly asymmetric pore coverage that μ -PIIIA exhibits on Nav1.4. In particular this can be seen by its predominance in hydrogen bond contacts to domains I and IV (8 in total) compared to domains II and III (4 in total) (**Figure 23a**, **Table 5**). In detail, Arg14 being hydrogen bound to the S5-SS1 loop of the Nav1.4 domain II suggests that this contact point (or at least a corresponding binding position within another domain since the channel is tetrameric) might provide an important anchor for efficient pore occlusions being lost upon mutation.

From all these aspects it is certainly be deducible that the slightly asymmetric position of μ -PIIIA isomer **2** on Nav1.4 is an ultimate result of its simultaneous contact with the inner central pore duct, from which further contacts towards the outer surface area is possible for the toxin rather towards only one direction. The compact globular and relatively rigid toxin structure required for such a formation is further provided by the μ -PIIIA native isomer's unique disulfide connectivity. The fact that the orientation to the side appears up to an only slight extent (**Table 5**) is still plausible from the same point of view, i.e. the further contact of the toxin to the inside of the channel. This typical binding mode can also be verified by simulations (chapter 4.4).

Table 5: H-bond analysis (H-bonds predicted by the HADDOCK docking calculations); valid "channel halves" are always defined by adjacent domains (I-II, II-III, III-IV or IV-I as possibilities); the halves and thus the proportion of contacting H-bridges are defined by that adjacent domain pair which proportionately has the majority of H-bridges (the proportion of H-bonds to the other half is thereby determined); docking results relating to highly active isomers are highlighted in light green (very highly active isomer 2 - darker green), moderately active isomer results are highlighted in light yellow and docking results of weak or inactive toxin isomers are highlighted in light orange.

Isomer	total number of H-bonds	domain				percentage of H-bonds to channel half 1 [%] (domain/domain)	percentage of H-bonds to channel half 2 [%] (domain/domain)
		I H-bonds to loops / SS1-SS2 / sum	II H-bonds to loops / SS1-SS2 / sum	III H-bonds to loops / SS1-SS2 / sum	IV H-bonds to loops / SS1-SS2 / sum		
1	19	3 / 4 / 7	0 / 4 / 4	3 / 1 / 4	2 / 2 / 4	57.89	42.11
2	12	1 / 1 / 2	2 / 1 / 3	0 / 1 / 1	4 / 2 / 6	66.66	33.34
7	17	2 / 3 / 5	1 / 4 / 5	0 / 3 / 3	1 / 1 / 2	58.82	41.18
11	19	0 / 3 / 3	4 / 4 / 8	2 / 4 / 6	0 / 2 / 2	73.68	26.32
12	22	1 / 3 / 4	4 / 5 / 9	2 / 3 / 5	1 / 3 / 4	63.63	36.37
14	17	2 / 3 / 5	1 / 1 / 2	0 / 3 / 3	3 / 4 / 7	70.59	29.41
15	18	4 / 1 / 5	0 / 3 / 3	1 / 1 / 1	5 / 3 / 8	72.22	27.78
					average:	66.21	33.79
3	17	2 / 4 / 6	0 / 4 / 4	2 / 1 / 3	2 / 2 / 4	58.82	41.18
4	23	6 / 3 / 9	1 / 1 / 2	4 / 1 / 5	6 / 2 / 8	73.91	26.09
9	23	1 / 5 / 6	4 / 3 / 7	5 / 2 / 7	1 / 2 / 3	60.87	39.13
13	20	2 / 7 / 9	1 / 2 / 3	2 / 3 / 5	0 / 3 / 3	60.00	40.00
					average:	63.4	36.6
5	17	4 / 2 / 6	2 / 3 / 5	0 / 3 / 3	1 / 2 / 3	64.71	35.29
6	21	0 / 4 / 4	1 / 7 / 8	2 / 3 / 5	2 / 2 / 4	61.90	38.1
8	26	3 / 7 / 10	1 / 4 / 5	0 / 2 / 2	3 / 6 / 9	73.08	26.92
10	21	3 / 4 / 7	1 / 2 / 3	2 / 1 / 3	5 / 4 / 9	76.19	23.81
					average:	68.97	31.03

4.3.1. Conclusions and future perspectives

With the docking studies conducted here, we were able to gain detailed insights into the binding patterns of the disulfide bond isomers of the μ -conotoxin PIIIA on the sodium channels such as the channel Nav1.4. It could be demonstrated that there is an ultimate correlation between the cysteine disulfide bridge pattern of a toxin, which determines its structure and flexibility, and finally its extent of ability to block certain channel subtypes.

The observations from the docking experiments were in line with the experimental observations pointing out that the native and most active isomer 2 exhibits coincident pore-diving and outer located contacts, evenly spread on the pore surface (**Figure 23a**). In contrast, the moderate or poorly active isomers either lack the central pore contact and/or the distributed covering contacts to the outer pore surface (**Figure 23, Figure 24**). Moreover, the corresponding H-bond distributions of the isomers with the four channel domains, suggests a more sideward directed positioning for the more inactive isomers (**Table 5**).

More precisely considering the native isomer 2 typical binding mode and the prerequisites for efficient pore blockages, the toxin's arginine (and other basic) residues play a key role. In the case of native isomer 2, they centrally immerse into the channel pore and simultaneously establish a fixation of the bound conformation by hydrogen bond contacts to the more outer located pore loops (**Figure 23a**). The coincident presence of pore-diving and surface contacts possibly leads to a stabilizing position of the toxin on the channel. Further, these contact formations may allow the toxin to span over the central pore on the one hand and to establish additional interactions with the selectivity filter region on the other hand. The basic toxin residues predominantly interact with acidic channel residues that are located in particular on the pore loops and are represented by the amino acids Asp (D) and Glu (E) within the known "DEKA" motif of the selectivity filter. In the case of native μ -PIIIA, the apparently neutralizing influence of Arg14 on the native acidic pore residues (such as Glu758), which serve to increase the effective Na⁺ concentration at the outer pore mouth, is particularly noteworthy.

Another feature supporting efficient pore block by μ -conotoxins is their particular intrinsic topology of cysteine disulfide bridging which can facilitate the formation of compact globular toxin structures.

Upon inspection of each isomer's structure (**Figure 19**) under consideration of the respective block efficacies (**Figure 20**), the limitation of disulfide bridge combinations showing potent inhibitory effects on Nav1.4 to a subset among all bridging patterns clearly shows up. This is not least due to the rigidity of the toxins or parts thereof, which follows directly from their disulfide bridging pattern. Here it becomes clear that those patterns of disulfide connectivity which result in lower toxin (backbone) flexibility, i.e. higher rigidity and thus in a more globular toxin structure rather promote an increased rate of pore block of the toxin.

In contrast, previous modelling studies on μ -PIIIA do exist,²¹⁵ in which not a certain and fixed toxin constellation, but several conformations for the binding and blocking of the pore centre are suggested. However, these findings are to be considered from the point of view of differing prerequisites, i.e. the missing loop areas in the study-related Nav1.4 models used. This can explain why the contacts to the outside area have not yet been discovered in those studies, but have turned out to be important anchor points for enabling the covering toxin span across the central pore area in our proposed models.

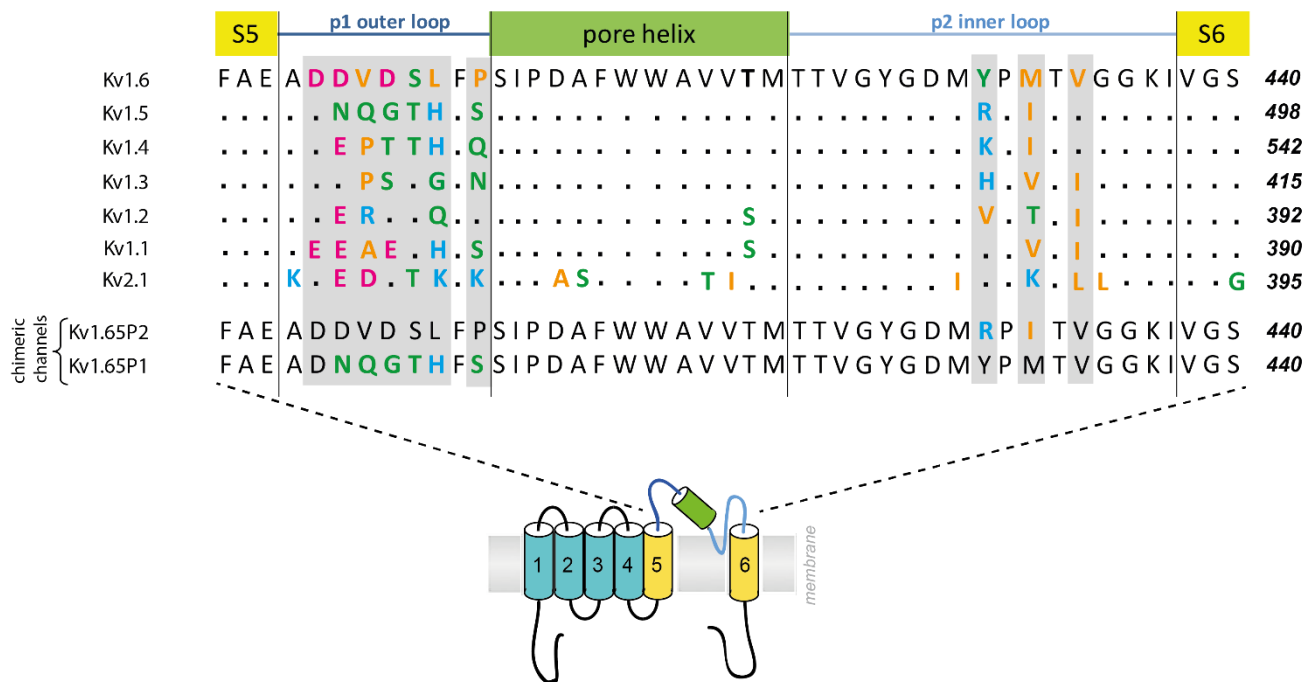
However, the connection between disulfide topologies and (segmental) toxin flexibilities and the impact on their bioactivities underpins the necessary consideration of flexibility aspects for a reasonable explanation of the blocking potencies on a virtual level. In these respects, it would be especially important to complement the docked representations by further simulation experiments which would allow for a full interpretation of the docked images.

4.4. In silico analysis of the subtype-selective block of KCNA ion channels through the μ -conotoxins PIIIA, SIIIA and GIIIA

As previously outlined the understanding of subtype specificity in the ion channel blockage by peptide-based toxins is crucial for developing such compounds into promising drug candidates.⁶ In this context, knowledge of which conotoxins are active on which channel subtypes and which are inactive is beneficial. However, the number of addressed channels can widely vary: μ -GIIIA, for instance, exclusively inhibits the skeletal muscle voltage-gated sodium channel Nav1.4, whereas μ -PIIIA and μ -SIIIA additionally block the neuronal sodium channels Nav1.2 and Nav1.7 (Table 3).^{6, 7} Furthermore, μ -PIIIA and μ -SIIIA were recently shown to not be exclusively specific for Nav channels.^{6, 7, 57-59}

More precisely, Leipold and co-workers were able to prove that μ -PIIIA and μ -SIIIA are inactive on the potassium channel subtypes Kv1.2 to Kv1.5 and also on Kv2.1, while however μ -SIIIA at least partially inhibited Kv1.1 and Kv1.6 and μ -PIIIA even blocked both, the Kv1.1 and Kv1.6 subtypes, revealing a nanomolar affinity (Figure 25b, Table 3).^{6, 7} Additional evaluations on constructed chimeric channels obtained from Kv1.5 and Kv1.6 yielded the inclusion of pore regions into the channel block by μ -PIIIA whereby the subtype-specific interactions are mainly determined by the so-called turret domain (P1 outer loop, Figure 25a) and partly by the sequence close to the selectivity filter (P2 inner loop, Figure 25a).^{6, 7}

a)



b)

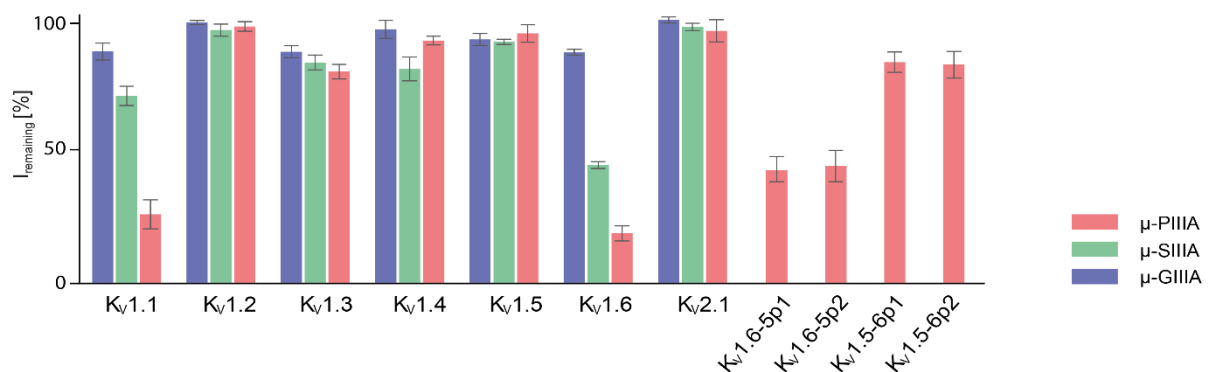


Figure 25: (a) Aligned amino acid sequences of the central, toxin-interacting channel region shown for one subunit of the Kv1 family (Kv1.1–Kv1.6), Kv2.1, and of the chimeric channels Kv1.6-5P2 and Kv1.6-5P1, which were all tested against μ -PIIIA, μ -SIIIA, and/or μ -GIIIA, respectively, by Leipold et al.⁷ Amino acids are coloured according to their

physicochemical properties (basic – light blue, acidic – magenta, polar/neutral – green, non-polar polar/hydrophobic – orange). The secondary structure elements are indicated above the alignment. **(b)** Activity rates [%] of μ -PIIIA, μ -SIIIA, and μ -GIIIA on potassium channels $K_V1.1$ – $K_V1.6$, $K_V2.1$, and $K_V1.5$ – $K_V1.6$ chimera channels, as published by Leipold et al.⁷ Lower percentage values of $I_{\text{remaining}}$ correspond to higher blocking activities. (Illustration modified from Kaufmann et al.⁶)

Further examples outlining the obscure venom specificity on K_V1 channel subtypes are given by a research study of the Iranian scorpion toxin OdK1 showing its selective inhibitory effects on $K_V1.2$ channels but ineffectiveness on $K_V1.3$, $K_V1.4$ and $K_V1.5$ ²¹⁶ and a further work on scorpion Hemitoxin which selectively blocks the rat $K_V1.1$, $K_V1.2$, but also $K_V1.3$ channels.²¹⁷

Unfortunately, the analgesic potential of μ -conotoxins is clearly limited by such special behaviors.⁶ Despite these disappointing findings, the high number of well-resolved potassium channel crystal structures present these days would certainly allow for deep and detailed in silico analyses of the binding of μ -conotoxins to K_V channels.⁶ In addition, the experimental data provided by Leipold and his team could enable us to clarify the question of how toxins attain their subtype specificity at a much higher qualitative level than it would be currently possible for μ -conotoxin binding on Na_V channels.⁶

Thus, in this section we will show on how powerful state-of-the-art in silico approaches can be used to unravel the subtype-specific inhibition of potassium channels by μ -conotoxins. Since an adequate understanding of the observed bioactivity differences of the toxins on the particular channels is only provided when dynamic information is taken into account, molecular dynamics simulations were performed in addition to docking experiments in order to elucidate the distinct specificities.⁶ (A summarizing workflow scheme is given in **Figure 26**.) We particularly addressed the three conotoxins μ -PIIIA, μ -SIIIA and μ -GIIIA in combination with different representatives of the K_V1 voltage-gated potassium channel family and artificially created chimeric channels. Correlations of conformational flexibilities and the experimental activities as recently reported by Leipold et al.⁷ could be determined. In this context and for reasons of comparison, μ -SIIIA was investigated as a partially channel blocking semi-active system, while μ -GIIIA was investigated as an inactive candidate. In particular, the different pore blocking modes – full or partial pore coverage – were of special interest, and were intended to achieve more insight into the origin of the remaining currents, especially observable for semi-active systems.⁶

In detail, our dataset of consideration was set up as follows: in silico investigations were performed respectively on μ -PIIIA, μ -SIIIA and μ -GIIIA on virtual models of the K_V1 -channel subtypes $K_V1.1$, $K_V1.5$, $K_V1.6$ and two further artificially constructed $K_V1.6/1.5$ chimera channels. The chimeric channels are further designated as “ $K_V1.6$ -5P1”, having exchanged the outer P1 loop of $K_V1.6$ by the $K_V1.5$ P1 loop, and as “ $K_V1.6$ -5P2”, having exchanged the inner P2 loop of $K_V1.6$ by the inner P2 loop of $K_V1.5$, respectively (see **Figure 25a**). More precisely, we studied μ -PIIIA on $K_V1.1$, $K_V1.5$ and $K_V1.6$ as well as on the chimera channels $K_V1.6$ -5P1 and $K_V1.6$ -5P2, while μ -SIIIA- and μ -GIIIA-binding was analyzed on $K_V1.6$, respectively, in order to decipher subtype-specific features with respect to progressive alterations from the perspective of the targeted channel as well as with respect to the toxin applied. We thus also made sure to cover at least one case of active, semi-active and inactive representative among our data set of investigated systems.

Especially, different pore blocking modes, ranging from partial to full pore coverage (named active systems thereafter), were of particular interest as being intended to provide insights into the origin of the remaining currents that are observable for the partly blocked ion channel pores (named semi-active systems thereafter).⁶

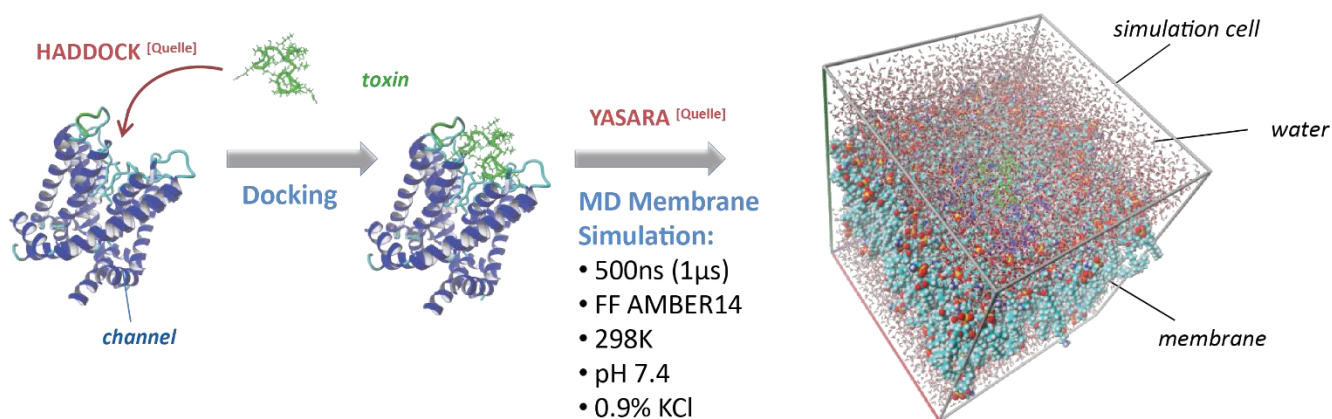


Figure 26: Workflow overview showing the docking outputs being subjected to subsequent MD calculations.

4.4.1. Homology modelling

In a first step a homology model of the K_v1.6 potassium channel was created. Starting from the K_v1.6 model, the model structures for K_v1.1 and K_v1.5 as well as the chimera channels “K_v1.6-5P1” and “K_v1.6-5P2” were reconstructed.⁶ The loop segments of the particular models were respectively substituted according to the sequences as shown in **Figure 25a**, while maintaining the backbone secondary structure.⁶ For more information on the modelling process please refer to chapter 5.5 in the methods section. Finally, all models were energy-minimized for further use.⁶ The voltage-sensor domains K_v1 channel models were omitted in the further steps, as they are not involved in the μ -conotoxin binding area focused in this work.⁶

4.4.2. HADDOCK docking

In order to obtain an adequate starting point for our simulations, we first performed HADDOCK docking using the toxin NMR structures applied on our homology models of the target channels.⁶ The homology modelling and docking processes are described in more detail in the methods section (chapters 5.5 and 5.6).

On the basis of the user-specified active and passive residues, the HADDOCK algorithm generates a set of so-called Ambiguous Interaction Restraints (AIRs).²¹⁸ Those restraints are meant to drive the docking by incorporation of experimental information – just in a similar manner as ambiguous distance restraints based on NOE data that drive the structural calculation of an NMR structure (see e.g. ²¹⁹),²¹⁸

In HADDOCK, for every active-set residue, a single AIR restraint is defined between that residue and all active and passive residues on the partner molecule.²¹⁸ The docking process then models and predicts the three-dimensional conformation of the complex based on the known single constituents in their free, unbound form²¹⁸ and under consideration of the predefined restraints. The ideal positioning of the ligand molecule into the target however describes a local minimum state in most cases.

Re-scoring of the best scoring HADDOCK clusters using VINA Autodock scoring function yielded binding energy values of noticeably high correlation regarding the toxin activity rates reported by Leipold et al.⁷ (**Table 6**),⁶ further encouraging our decision to take these docked conformations as starting structures for subsequent equilibration simulations.²²⁰

Table 6: Scoring summary for the best scoring μ -conotoxin-channel systems as obtained from HADDOCK docking²¹⁸ and Vina re-scoring²²⁰ calculations; assigned blockage activity levels as reported by Leipold et al.⁷ are given. (Table source: Kaufmann et al.⁶)

		HADDOCK Z-score	HADDOCK score	Vina score [kcal/mol]
μ -PIIIA on Kv1.6	<i>active</i>	-1.0	174.3 \pm 8.7	10.5
μ -PIIIA on Kv1.1	<i>active</i>	-1.4	202.1 \pm 5.9	9.5
μ -PIIIA on Kv1.6-5P1	<i>semi-active</i>	-1.4	178.2 \pm 14.0	9.7
μ -PIIIA on Kv1.6-5P2	<i>semi-active</i>	-0.9	196.0 \pm 12.7	10.0
μ -SIIIA on Kv1.6	<i>semi-active</i>	-1.3	231.1 \pm 14.7	10.2
μ -PIIIA on Kv1.5	<i>inactive</i>	-1.6	202.6 \pm 10.5	8.5
μ -GIIIA on Kv1.6	<i>inactive</i>	-1.7	187.1 \pm 14.0	8.0

4.4.3. Toxin dynamics and cluster analysis

Since all docked toxin structures were centred in the middle of the pore (**Figure 27**), which complicated the interpretation of their bioactivity, the best scoring docked structures were equilibrated through molecular dynamic simulations in a membrane environment for 0.5 to 1.0 μ s (for more information considering the MD experiments see chapter 5.7 in the methods section).⁶ Such subsequent MD simulation applications will take any kind of residue contacts into account and equilibrate the whole system without wrongly omitting any potential interactions. At this point it should briefly be mentioned that MD equilibrations are to be considered as ensuing refining step, anyway “filtering” essential contacts from the false-positive ones, eventually being present in the initial state.

The simulations thus aimed to provide more accurate descriptions of the toxin binding and dynamics when bound to a target channel, especially since the toxins did not only show a full pore block, but also partly blocked ion-channel pores.^{6,7} At this point, it has to be noted that the corresponding docking results and simulation data of inactive systems (toxins which do not block the pore) might be difficult to construe.⁶ Nevertheless, these data were also shown and analyzed for the sake of comparisons and completeness.⁶

Molecular dynamics simulations as a complementary measure to docking experiments

To attain more detailed descriptions of protein-ligand binding, dynamic information of such systems is of utmost importance, especially in the sense of completing the imagination of a rigid docked representation of an interacting target-ligand complex. In particular, meaningful target-ligand constellations which differ from the docked start position can occur, that would not have been recognizable or predictable by simple docking studies.

This can already be recognized by comparison of the docked initial structures with the calculated average structures from the respective simulations, revealing a toxin displacement over time and the corresponding direction of movement (**Figure 27**). The different degrees of deviation from the initial state further lead to the assumption that varying rates of bioactivities can be mirrored by their corresponding dynamics. In this sense, a stronger deviation from the initial state generally indicates a higher detachment from the pore surface and thus suggests more inactivity.

We thus aimed to gain more informative states beyond the docked starting structure by trapping the systems in an energetically favourable local minima, i.e. *equilibrated states*. Based on the occurrence of such states we aspire to figure out essential residue contacts and structural features ultimately accounting for subtype-specific recognition events. Such crucial contacts can hardly be determined from a rigid state alone and without any supporting dynamic data.

We anyhow consider the docked states proposed as appropriate starting point from which inferences about further blocking or unbinding events can be drawn by subsequent MD calculations.

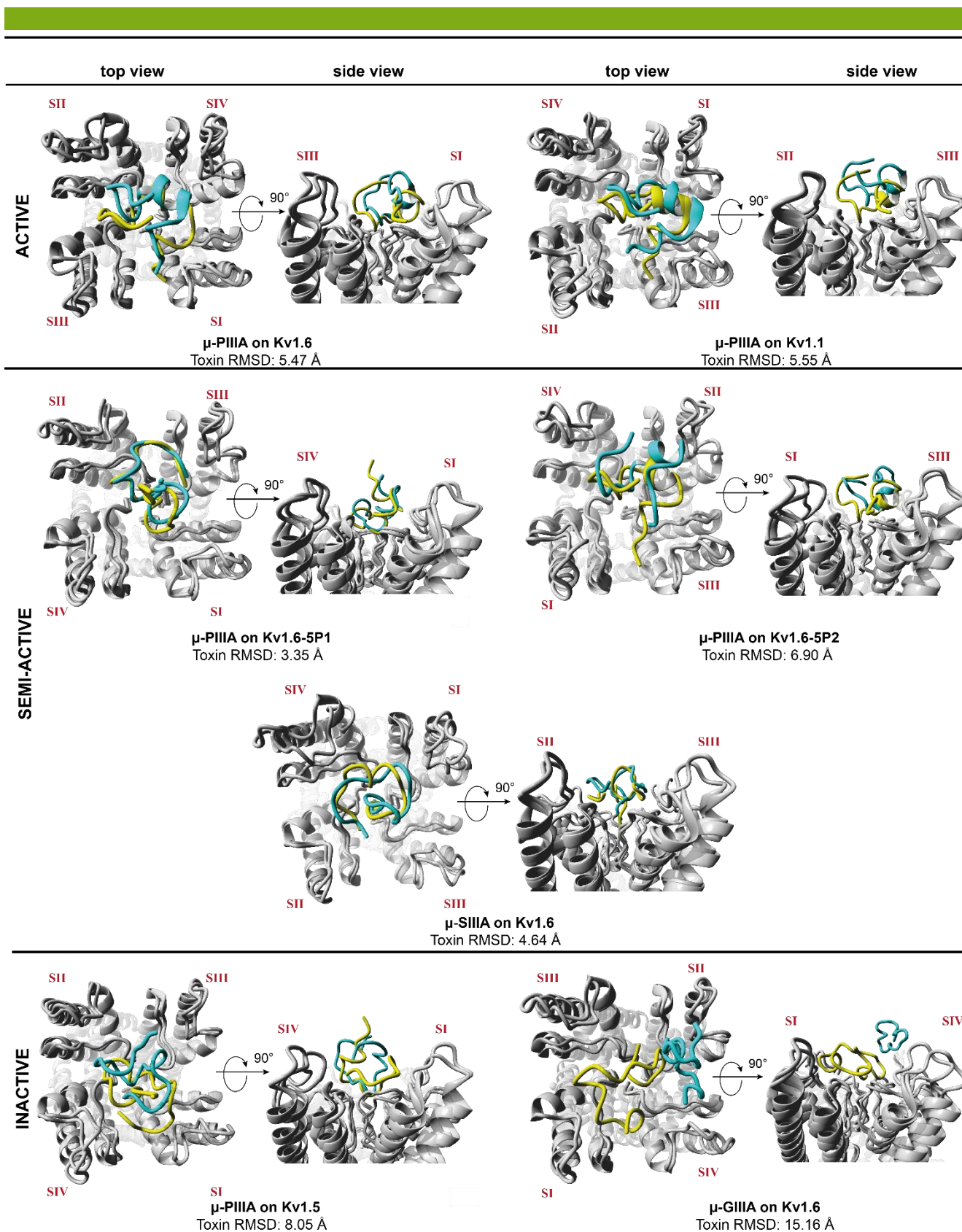


Figure 27: System-wise comparison of start and average simulation structure. Toxin-channel systems are shown in 3D secondary structure representation with docked toxin starting structure and superimposed simulation average structure. Colour scheme: start toxin structure – yellow, simulation average toxin structure – cyan, channels – gray. The respective toxin RMSD values are indicated below each system. The systems are arranged according to their activity level (active – top, semi-active – middle, inactive – bottom) and are respectively shown from above (left) and from the side (right). Subunits are designated by Roman numerals (red, SI – SIV).

Cluster analysis

Figure 31 illustrates the system-varying dynamics using representative structures of calculated simulation data clusters, each of which differs by 4Å RMSD over the total simulation run (**Figure 28**). At the same time, the start structure of the docked state is shown for each system. According to our simulation data, less active systems (exhibiting less pore coverage) show higher toxin dynamics, i.e. a higher number of representative cluster structures, while semi-active or active systems show lower toxin dynamics by less cluster representatives, i.e. an overall higher stability (**Figure 31**).⁶ For all systems, structural deviations from the initial docked state were observed (**Figure 27, Figure 28, Figure 31**). Here we could identify the toxin as the main moving part in a channel-toxin system by further comparison of the complete systems' C α root mean square deviation (RMSD, data not shown) and the toxin C α RMSDs, revealing negligible variations.

Moreover, toxin dynamics are equally reflected by the toxins' C α RMSD over time, altogether showing higher deviations for more flexible (inactive) toxin-channel complexes and less deviations for better pore blocking systems (more active) (**Figure 28, Figure 31**). As the simulations were performed until a nearly linear behavior of the RMSDs was achieved, simulation times varying between 0.5 and 1.0 μ s were reached.⁶

Altogether, dynamics analyses suggested a unique, stable channel blocking mode, while revealing a more diverse behavior for semi- and in-active systems (**Figure 28, Figure 31**).⁶ In this context it is noticeable, that active representative systems μ -PIIIA on K_V1.6 and K_V1.1 show a very similar course development over the whole simulation time, indicating their tightly bound condition, whereas semi- and inactive systems exhibit differences in these respects (**Figure 28**). These findings substantiate the uniqueness of subtype-specific active channel blockage.⁶

On the other hand, recurring movement patterns in semi- and inactive systems allow us for a further classification or differentiation within these systems. More precisely, two different binding modes respectively describable for semi- and inactive systems: in the semi-active systems, one group is built by the less fluctuating and more stable systems μ -PIIIA on K_V1.6-5P1 and μ -SIIIA on K_V1.6 and another, more dynamic and less stable one only represented by μ -PIIIA on K_V1.6-5P2 (**Figure 28**).⁶

Likewise, two different modes of toxin movement were identified for the two inactive toxin-channel systems. For K_V1.5-bound μ -PIIIA, the toxin displacement occurred more gradually, whereas for μ -GIIIA, it was moving more rapidly on the channel surface (**Figure 28, Figure 31**).⁶

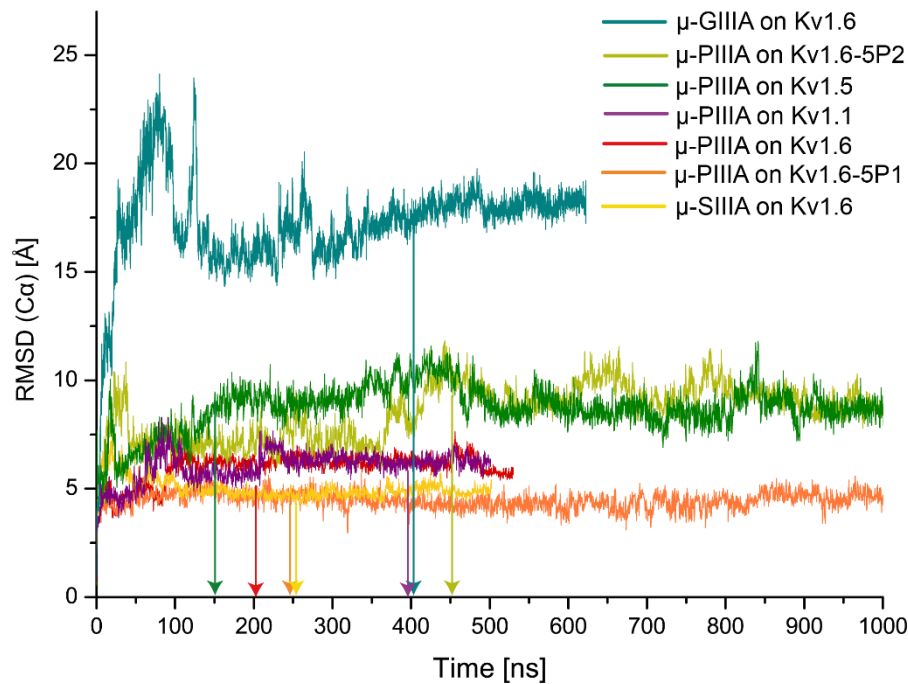


Figure 28: Toxin C α RMSD over the whole simulation. Toxin RMSD traces up to 1 μ s for semi-active systems μ -PIIIA on Kv1.6-5P1, μ -PIIIA on Kv1.6-5P2 and for the inactive system μ -PIIIA on Kv1.5. Arrows indicate simulation times at which our representative simulation structures were selected for subsequent analyses. (Illustration modified from Kaufmann et al.⁶)

Closer inspection of the toxin movements with respect to the four channel subunits (**Figure 29, Figure 30**) further supported the stated differences within the semi- and inactive systems: C α RMSD traces of μ -PIIIA on Kv1.6-5P2 (semi-active) and μ -GIIIA on Kv1.6 (inactive) clearly unveils a toxin movement towards the side of a channel subunit, in contrast to the other semi- or inactive systems (**Figure 29, Figure 30**).⁶ More specifically, μ -PIIIA on Kv1.6-5P2 and μ -GIIIA on Kv1.6 both move towards the channels' P1 outer loop of subunit II (SII) and coincidentally away from the P1 loop of the opposite located subunit I (SI) (**Figure 29, Figure 30**).⁶ For the case of μ -PIIIA on Kv1.6-5P2 this drift away seem to be triggered by the Y429R mutation in the P2 loop that is closely located to the selectivity filter.⁶ This would entail a reduced covering or blockage of the pore.⁶ μ -PIIIA thus moves towards the outer P1 loops of Kv1.6-5P2 subunits SII and SIII and at the same time away from the opposite SI and SIV outer and inner loops (**Figure 29, Figure 30**).⁶ μ -GIIIA moves towards SII and SIV of Kv1.6 and at coincidentally drifts away from the SI and SIII loops (**Figure 29, Figure 30**).

In contrast, semi-active systems μ -PIIIA Kv1.6-5P1 and μ -SIIIA Kv1.6 and the inactive μ -PIIIA Kv1.5 system do not show toxin displacements towards any subunit.⁶ Rather, similar to the active μ -PIIIA Kv1.1 and Kv1.6 systems, a more center-positioned state of the on the pore is maintained (**Figure 27, Figure 28, Figure 29, Figure 30**).⁶

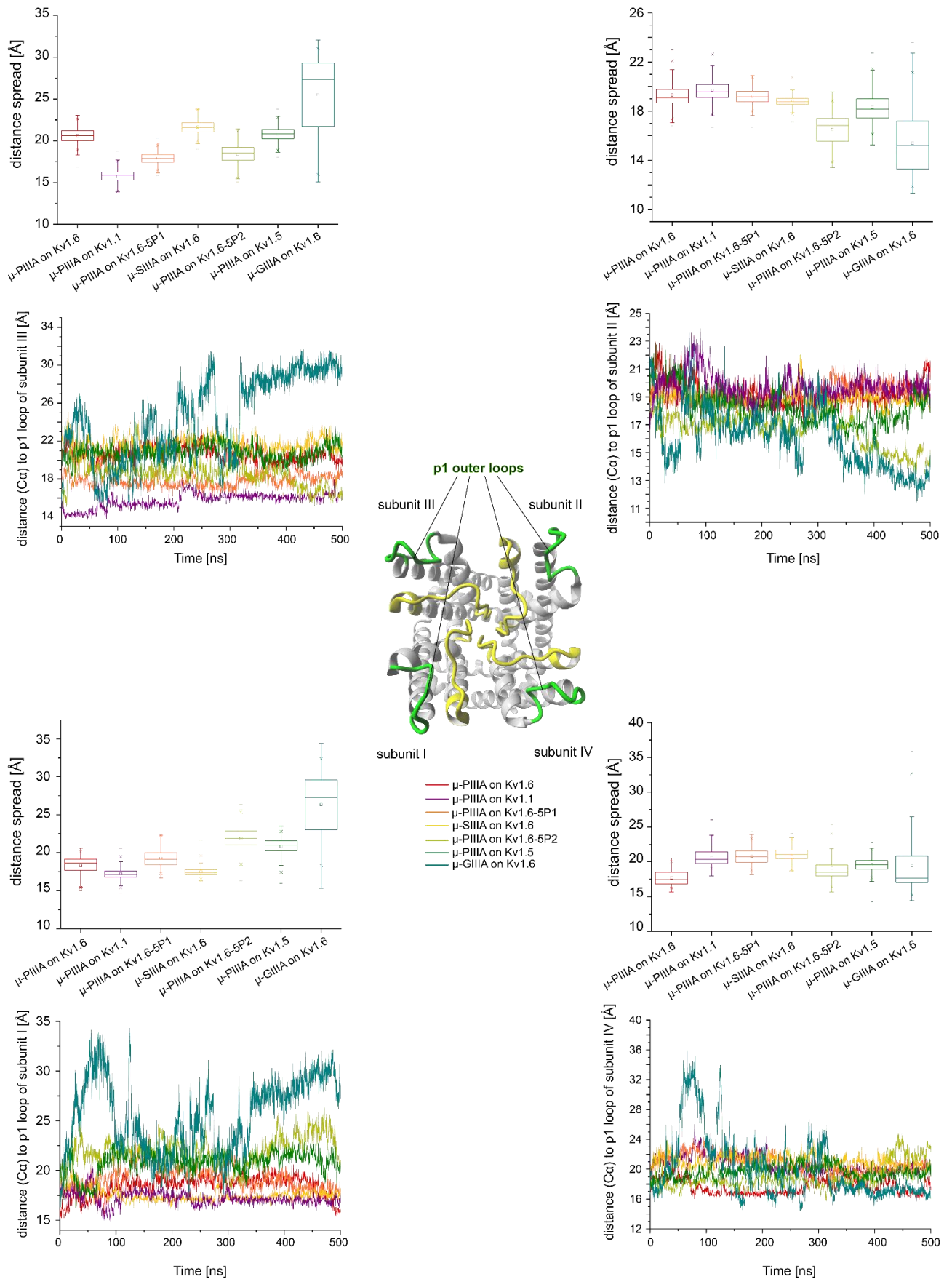


Figure 29: C α RMSD traces showing the movement, i.e. changing distance of the toxin mass centre with respect to the outer loops (P1) of the individual channel subunits (P1 outer loops – light green, P2 inner loops – light yellow). (Illustration modified from Kaufmann et al.⁶)

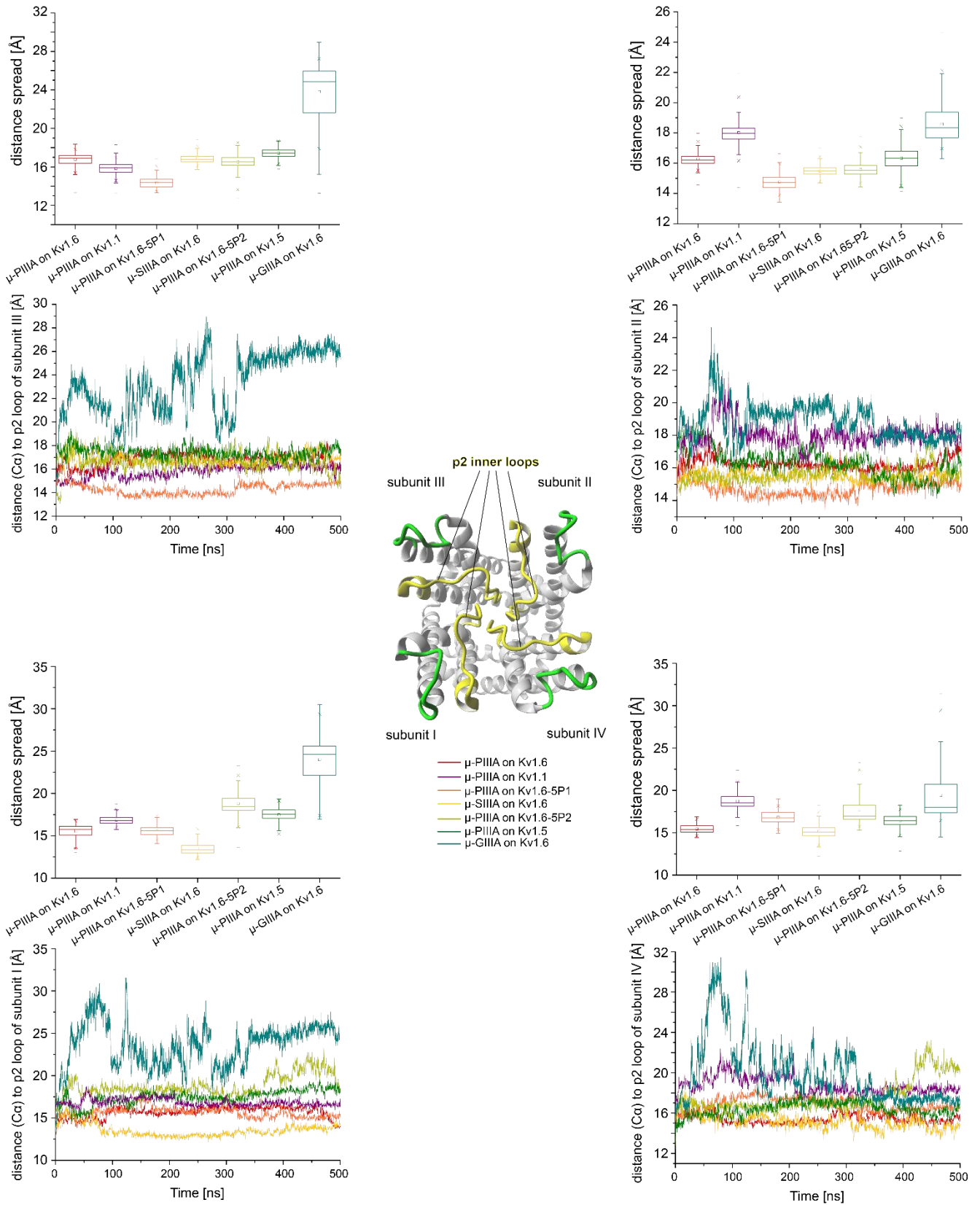


Figure 30: C α RMSD traces showing the movement, i.e. changing distance of the toxin mass centre with respect to the inner loops (P2) of the individual channel subunits (P1 outer loops – light green, P2 inner loops – light yellow). (Illustration modified from Kaufmann et al.⁶)

These differences within semi- and inactive systems with regard to the movement direction of the toxins allow for a general differentiation between toxin movement modes where the toxin either stays central with respect to the channel pore – hereinafter referred to as “*centrally located*” mode (a) – and systems where the toxin moves towards the outside of a channel moiety – hereinafter referred to as “*side-directed*” mode (b). As for the two considered active systems μ -PIIIA on K_V1.1 and μ -PIIIA on K_V1.6 there were no such movement differences observable, we assume a unique binding mode for active potassium channel blockage by μ -PIIIA. It can further be noted that the movement of μ -PIIIA on K_V1.6-5P2 towards a subunit moiety is reasonable because of mutations in the central inner P2 loops triggering the toxin to drift away from the center to the outside (**Figure 25a**).

Apart from the RMSD analyses, the intrinsic higher flexibilities exhibited by toxin-channel systems with a toxin moves away from a central starting position was quite obvious, as also extracted from the simulations per atom b-factors (**Figure 31, Figure 32**).⁶ (Here, it should be considered that the docked start position is already central by default, as far as the pore is concerned.) Further examinations of the different simulation cluster representatives (**Figure 31**) equally confirm this observation, pointing out a higher number of cluster representatives for the more flexible side-directed systems and a lower number of cluster representatives for centrally equilibrated systems (**Figure 31, Figure 32**). Movement patterns can thus be directly correlated with the overall dynamics and flexibility of the toxin. Additionally, the equal numbers of resulting cluster representatives for μ -PIIIA on K_V1.1 and K_V1.6 confirm our abovementioned proposal of a unique binding mode for active systems (**Figure 31**).⁶

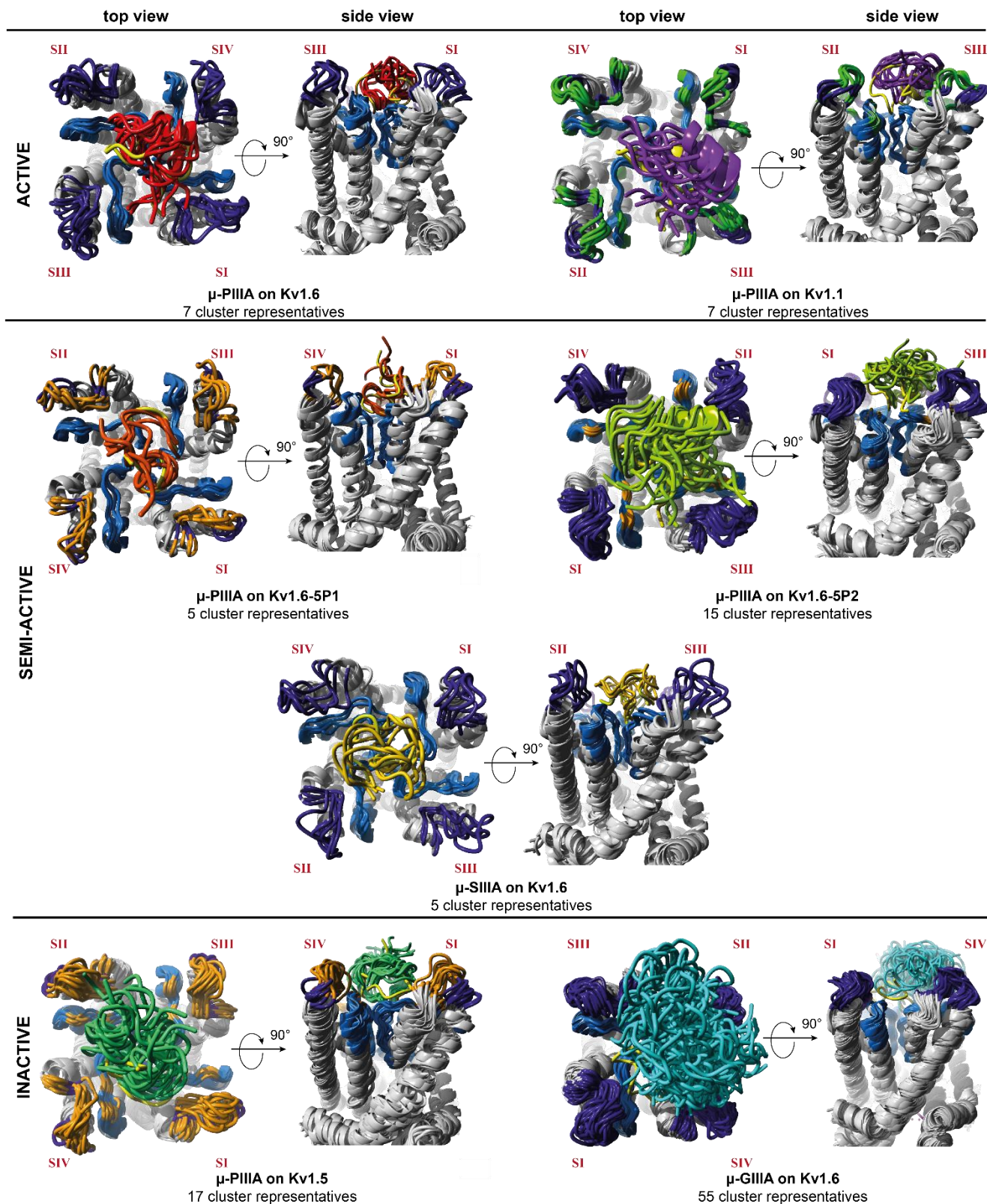


Figure 31: Toxin dynamics as cluster representatives. Toxin-channel systems are shown in secondary structure representation. The identified main simulation clusters are designated by respective 3D structures (4 Å toxin RMSD threshold) for active (top), semi-active (middle) and inactive (bottom) systems. The total numbers of identified clusters are given below each toxin-channel system. The colour scheme was set as follows: toxins: μ -PIIIA on Kv1.6 – red, μ -PIIIA on Kv1.1 – violet, μ -PIIIA on Kv1.6-5P1 – dark orange, μ -PIIIA on Kv1.6-5P2 – yellow green, μ -SIIIA on Kv1.6 – dark yellow, μ -PIIIA on Kv1.5 – dark green, μ -GIIIA on Kv1.6 – cyan, docked toxin start structures – light yellow; channels: P1 outer loops – dark blue, P2 inner loops – light blue, Kv1.6/5 differing channel mutations – orange, Kv1.6/1 differing channel mutations – green, remaining channel – gray. Channel subunits are designated by Roman numerals (red, SI – SIV). (Illustration modified from Kaufmann et al.⁶)

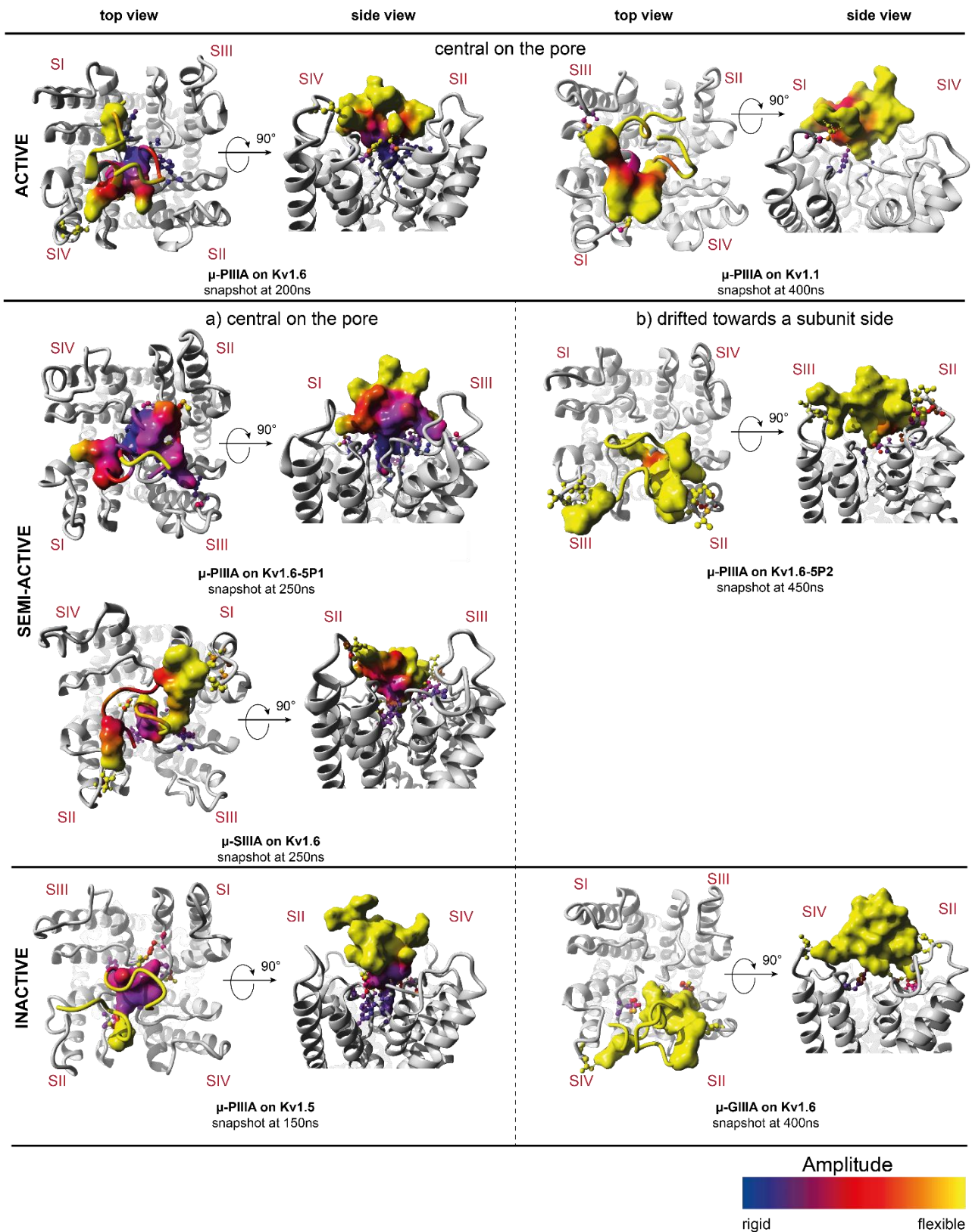


Figure 32: b-Factor coloring of the interacting toxin and channel residues (flexible – yellow, rigid – blue); the toxin is shown partly in molecular surface (interacting residues) and in secondary structure representation (remaining residues) in the top view and in complete molecular surface representation in the side view; interacting channel residues are shown as ball and sticks; remaining channel is shown as secondary structure (gray); the displayed structures are energy-minimized and depict the selected representative snapshots from the simulation data; channel subunits are indicated by Roman numerals (red, SI – SIV). (Illustration modified from Kaufmann et al.⁶)

Nevertheless, one may ask why, in comparison to the active systems μ -PIIIA-K_V1.6/K_V1.1, the semi-active systems μ -SIIIA-K_V1.6 and μ -PIIIA-K_V1.6-5P1 show a similar low overall fluctuation, suggesting even better binding stabilities pretended by the lower number of calculated cluster representatives over the whole simulation (**Figure 28**, **Figure 31**).

However, whereas μ -PIIIA on K_V1.6 and K_V1.1 both exhibit a more pronounced pore coverage, forming contacts that are equally distributed towards all four channel subunits, the toxins in μ -SIIIA-K_V1.6 and μ -PIIIA-K_V1.6-5P1 seem to be more compressed by a higher number of contacts towards the centre of the pore (**Figure 32**). In the case of μ -PIIIA-K_V1.6-5P1, this effect was certainly caused by the mutations in the outer loop. Concurrently, when the P2 loop of K_V1.6 was replaced by the corresponding part of the toxin insensitive K_V1.5 channel, an even more lateral orientation of the toxin can be noticed, presumably being triggered by the newly inserted positively charged Arg429 (**Figure 27**, **Figure 32**).⁶

Altogether, the experimental activities could be rationalized from the simulation data, even though one would expect somehow higher flexibilities or dynamics for μ -SIIIA-K_V1.6 and μ -PIIIA-K_V1.6-5P1 in order to understand the remaining currents for these semi-active systems.⁶

Most likely, the rather short simulation times (up to 1 μ s) were insufficient to fully unveil all critical pore blocking determinants.⁶ This was especially true for the binding of μ -PIIIA towards the toxin insensitive channel K_V1.5.⁶ Here, one would expect no binding, or at least a relatively fast toxin detachment movement, as observable for the case μ -GIIIA on K_V1.6.⁶ Despite unbinding events were not observed even during our extended 1 μ s simulation, μ -PIIIA remained only bound to the center of the K_V1.5 pore, only showing very poor contact towards the pore surface, being clearly different from the situation of toxin binding at the active and semi-active systems, indicating that μ -PIIIA was about to dissociate from the pore soon.⁶

For subsequent analyses however, we stick on the assumption that a stable and (partially) blocking conformation can be observed within a simulation period where the system is in a steady and energetically favorable state that has stabilized over time. Our further examinations will therefore be focused on the corresponding interactions between toxin and channel arising at such equilibria.⁶ These are thought to hold the toxin in a specific positioning on the channel and thus to stabilize the different orientations on the pore.⁶ It seems to be straightforward to conclude the level of pore coverage or blockage from the interactions of the toxin with the channel surface.⁶ The adopted toxin positioning on the pore will therefore influence the channel's overall permeability. The identification of such representative states thus forms an essential basis for the further detection of key interactions responsible for the resulting toxin positions and the blocking activities.

For the identification of such system-wise representative snapshots, we developed and applied a selection protocol,⁶ which is supposed to filter out predominantly assumed system states or modes. For this purpose, we have used two automated clustering methods in combination, each finding a stable main cluster and forming the intersection (or union) of both results (for more details see chapter 5.8 in the methods section). The selection protocol respectively parses a periodically pre-filtered subset of eleven simulation snapshots (in the interval from 0ns up until 500ns, spaced at 50ns).⁶

Further, our clustering computations drastically reduced the amount of data to be parsed and hence also brought a lot of savings in computing time. We count on this approach based on the presumption that a stable, equilibrated state can be represented as a cluster of maximum size out of the eleven pre-selected snapshots and – consequently – a representative snapshot candidate is probably present in this main stable cluster. At the same time we took into account the distances of the individual cluster members to the average simulation structure, ensuring a smallest possible RMSD of a representative candidate to the average.

In addition, we included the observation of principal detachments towards a subunit side for the semi-active system μ -PIIIA on K_V1.6-5P2 and μ -GIIIA on K_V1.6; these sideways-drifting systems are further differentiable by the different points in time of toxin detachment. To take this into consideration, we included a final revising step to check the instances falling under category semi-active and side-directed mode (b) and adapted our selection methodology accordingly. For more information see methods section, chapter 5.8.

For μ -PIIIA on $K_{V1.6-5P2}$ in particular, our latest checks resulted in a further revised final representative replacing the first proposed representative snapshot at 200ns by snapshot at 450ns of the latest simulation quarter (**Table 7**). All other result snapshots of our selection procedure are given in **Table 7** based on their simulation times (including intermediate steps of the selection procedure). For details regarding the partial results, please also refer to chapter 5.8 in the methods section.

Immediately noticeable, the identic main clusters are detected by clustering method 1 and 2 for the active system μ -PIIIA on $K_{V1.6}$ (**Table 7**, 1st row, 2nd and 4th column), actually considerable as a number-based confirmation of the stability and specificity the unique active blocking mode entails. The conclusiveness of our methodical approach can also be confirmed by looking at the corresponding results of μ -GIIIA on $K_{V1.6}$, which show almost the opposite in this respect (**Table 7**, last row, 2nd and 4th column), demonstrating the high fluctuation and instability of the inactive system.

Further to be noted is the constant decrease of the main clusters' sizes obtained from both clustering methods, considered over the entire activity levels and starting from the active systems. (This also includes the total number of clusters detected by method 2.) Thus matching the performed clustering calculations shown in **Figure 31**, the systems' different flexibilities were also adequately reflected with our own method.

For more detailed information on the intermediate clustering results please refer to **Table 12** – **Table 14** in the methods section.

Table 7: Cluster result summary showing intermediate and final results of our clustering selection protocol (chapter 5.8); main cluster obtained by cluster analysis 2 is highlighted yellow (4th row). (Table adapted from Kaufmann et al.⁶)

System	(main cluster) result cluster analysis 1	cluster analysis 1 representative: min. toxin RMSD to avg. toxin [Å]	result cluster analysis 2 (main cluster)	cluster analysis 2 representative: min. toxin RMSD to avg. toxin [Å]	intersection (or union) cluster	intersection ("revised") representative min. toxin RMSD to avg. toxin [Å]	RMSD of revised representative to toxins of last three sim. snapshots [Å]	final representative:
μ-PIIIA Kv1.6	{3,4,5,6,7,8,9,10}	snapshot {5} at 200ns: 1.5248	{1},{2},{3,4,5,6,7,8,9,10},{11}	snapshot {5} at 200ns: 1.5248	{3,4,5,6,7,8,9,10}	snapshot {5} at 200ns: 1.5248	to snapshot {9} (at 400ns): 2.1996 to snapshot {10} (at 450ns): 2.3839 to snapshot {11} (at 500ns): 4.0344	snapshot {5} at 200ns
μ-PIIIA Kv1.1	{2,3,4,5,6,7,8,9,10,11}	snapshot {10} at 450ns: 1.7046	{1},{2},{3,5},{4,6,10},{7,8,9,11}	snapshot {9} at 400ns: 1.8812	{7,8,9,11}	snapshot {9} at 400ns: 1.8812	to snapshot {9} (at 400ns): 2.0698 to snapshot {10} (at 450ns): 0.0000 to snapshot {11} (at 500ns): 2.5402	snapshot {9} at 400ns
μ-PIIIA Kv1.6-5P1	{2,5,6,9,10,11}	snapshot {6} at 250ns: 2.5522	{1},{2,3},{4,5,6,7},{8,9,10},{11}	snapshot {6} at 250ns: 2.5522	{5,6}	snapshot {6} at 250ns: 2.5522	to snapshot {9} (at 400ns): 4.8724 to snapshot {10} (at 450ns): 4.3513 to snapshot {11} (at 500ns): 5.0838	snapshot {6} at 250ns
μ-SIIIA Kv1.6	{2,3,4,5,6,7,8,9,10,11}	snapshot {6} at 250ns: 1.2155	{1},{2,3},{4,5,6,7,8,9,10,11}	snapshot {6} at 250ns: 1.2155	{4,5,6,7,8,9,10,11}	snapshot {6} at 250ns: 1.2155	to snapshot {9} (at 400ns): 1.6569 to snapshot {10} (at 450ns): 2.4387 to snapshot {11} (at 500ns): 1.9899	snapshot {6} at 250ns
μ-PIIIA Kv1.6-5P2	{2,3,4,5,6}	snapshot {5} at 200ns: 2.2403	{1},{2,6},{3},{4,5},{7},{8},{9},{10},{11}	snapshot {5} at 200ns: 2.2403	{2,4,5,6}	snapshot {5} at 200ns: 2.2403	to snapshot {9} (at 400ns): 5.0467 to snapshot {10} (at 450ns): 7.7670 to snapshot {11} (at 500ns): 8.5439	snapshot {10} at 450ns
μ-PIIIA Kv1.5	{4,5,6,7}	snapshot {4} at 150ns: 2.4267	{1},{2},{3},{4,7},{5},{6},{8},{9},{10},{11}	snapshot {4} at 150ns: 2.4267	{4,7}	snapshot {4} at 150ns: 2.4267	to snapshot {9} (at 400ns): 5.0881 to snapshot {10} (at 450ns): 5.5301 to snapshot {11} (at 500ns): 4.3945	snapshot {4} at 150ns
μ-GIIIA Kv1.6	{}	snapshot {6} at 250ns: 5.7302 (lowest RMSD among all)	{1},{2},{3},{4},{5},{6},{7},{8},{9,11},{10}	snapshot {9} at 400ns: 6.5590	{9,11}	snapshot {9} at 400ns: 6.5590	to snapshot {9} (at 400ns): 0.0000 to snapshot {10} (at 450ns): 4.2314 to snapshot {11} (at 500ns): 2.7594	snapshot {9} at 400ns

Altogether, our systematized selection protocol constitutes a useful addition for further interaction analyses since arbitrariness for the determination of representative states out of large volume of data is restricted by a profound and computationally underpinned method which combines two different clustering approaches and a final revising step.⁶ In addition, our implemented selection procedure could mirror the systems' differing flexibilities similarly to the independent clustering analysis applied previously (**Figure 31**).⁶ This is reflected in the comparable distribution of the respective resulting numbers of clusters that equally correlate with the flexibilities of the particular systems (**Table 7, Figure 31**).

As will finally be seen, our selected snapshots further confirm the presumed binding modes, subdivided into to central (a) and side-directed systems (b), coupled with the activity levels. Appropriate RMSD values, indicating the distance from the docked starting structure also substantiates this subdivision, revealing larger distance from the respective starting position for sideward moving systems (b) and generally smaller movement distances for toxins which remain in a central position on the channel pore (a) (**Table 8**).

Table 8: RMSD of finally selected representative snapshots to the corresponding docked structure used as start for MD calculations (Table adapted from Kaufmann et al.⁶)

	mode of activity	selected snapshot at time [ns]	RMSD from start [Å]
μ -PIIIA on Kv1.6 at 200ns	<i>active (unique)</i>	200	6.066
μ -PIIIA on Kv1.1 at 400ns	<i>active (unique)</i>	400	6.077
μ -PIIIA on Kv1.6-5P1 at 250ns	<i>semi-active (a)</i>	250	4.462
μ -SIIIA on Kv1.6 at 250ns	<i>semi-active (a)</i>	250	4.652
μ -PIIIA on Kv1.6-5P2 at 450ns	<i>semi-active (b)</i>	450	10.064
μ -PIIIA on Kv1.5 at 150ns	<i>inactive (a)</i>	150	8.914
μ -GIIIA on Kv1.6 at 400ns	<i>inactive (b)</i>	400	17.471

Analysis of ion channel-toxin interactions

As previously outlined the extent of the channel blockage is a direct consequence of the different adopted toxin orientations on the pore, motivating further investigations on the corresponding interactions arising in equilibrated states.⁶ In particular, we aim to focus on emerging key interactions which are able to stabilize the toxin on a specific place on the channel and thus allow for either more or less blocking conformations.⁶

Figure 34 shows the chosen representative snapshots as energy-minimized 3D structures resulting from our cluster-analysis. It provides a preliminary indication of the correlation between the extent of overall channel blockage, structural characteristics and the toxin positioning on the pore which is stabilized by the related interactions.⁶

In the following sections we will provide a detailed description of the toxin-channel interactions and resulting structural properties that occur in the representative constellations.

μ -PIIIA on Kv1.1/6 – a symmetric core motif accounts for a unique binding mode

Structural superposition of both representative structures reveals a very similar conformation of μ -PIIIA on Kv1.6 and on Kv1.1, confirming the uniqueness of the active μ -PIIIA blocking mode on the potassium channels (**Figure 33**, **Figure 34a**).⁶ This emerges as a pore-protruding and simultaneously channel covering situation, which is made possible by the central toxin positioning and coincident contacts to outer and inner pore loops (**Figure 33**, **Figure 34a**, **Figure 37**).

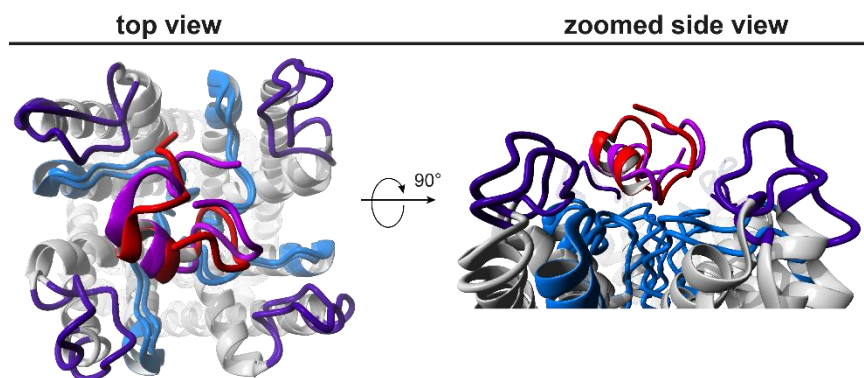


Figure 33: Superposed representative bound states of μ -PIIIA on Kv1.1 (purple) and on Kv1.6 (red) in top and side view.

This highly specific blocking mode is primarily stabilized by central toxin residues that evenly address at least three of the four channel subunits (**Figure 34a**) through an intensive hydrogen bond pattern.⁶ In the case of μ -PIIIA bound to Kv1.6 the blocking constellation is further strengthened by the N-terminal Arg2 of μ -PIIIA that forms hydrogen bonds towards residues from the inner P2 loops of the Kv1.6 subunit I (SI) (**Figure 34a**).⁶ In case of μ -PIIIA on Kv1.6, hydrogen bonds are primarily formed towards the inner (P2) pore loop residues, whereas for μ -PIIIA on Kv1.1 more residues of the outer (P1) loops are involved in toxin channel interactions (**Figure 34a**).⁶

In these respects, the central Arg12 and Arg14 seem to be essential for the pore block of Kv1.1 and Kv1.6, respectively forming hydrogen bonds towards the outer located (P1) loop by either interacting acidic glutamic acid (Glu353/SI) in the case of Kv1.1 or aspartic acid (Asp401/SIV) in the case of Kv1.6 (**Figure 34a**).⁶ Equally, Asp427/SII of the inner (P2) loop is also addressed in the case of μ -PIIIA on Kv1.6 (**Figure 34a**).⁶ Consequently, there seem to be no major impact on the physicochemical composition of the channel surface, when aspartic acid (Asp) in Kv1.6 is converted to glutamic acid (Glu) in Kv1.1 (and also from hydrophobic valine to hydrophobic alanine), enabling the similar binding modes for μ -PIIIA on both channels (**Figure 25a**, **Figure 33**, **Figure 34a**).⁶

Neutral serine Ser13 which is also located in the centre of the toxin contacts the equally neutral glycine or larger tyrosine residues (Gly426/SIV and Tyr429/SIV in Kv1.6 and Tyr379/SI in Kv1.1) of the inner P2 loops that are closely located to the selectivity filter (**Figure 34a**).⁶

Gln15 also appears to be of relevance for the specific block of Kv1.6 and Kv1.1 as well, as it forms hydrogen bonds with the inner glycine residues of the selectivity filter (Gly424/SIII, Gly424/SIV, Gly426/SIII of Kv1.6 and Gly376/SIV of Kv1.1) (**Figure 34a**).⁶ Thus, the pore-protruding orientation of Gln15 occluding the central cavity of the channel is equally present in both bound states of μ -PIIIA to Kv1.1 and Kv1.6 (**Figure 34a**, **Figure 37**). In Kv1.6, Gln15 further stabilizes its pore blocking orientation by another hydrogen bond to Tyr429/SII (**Figure 34a**).⁶ As the cavity addressed by Gln15 is intrinsically located in the centre of the pore, Gln15 might be of vital importance for the positioning of μ -PIIIA on the channel.

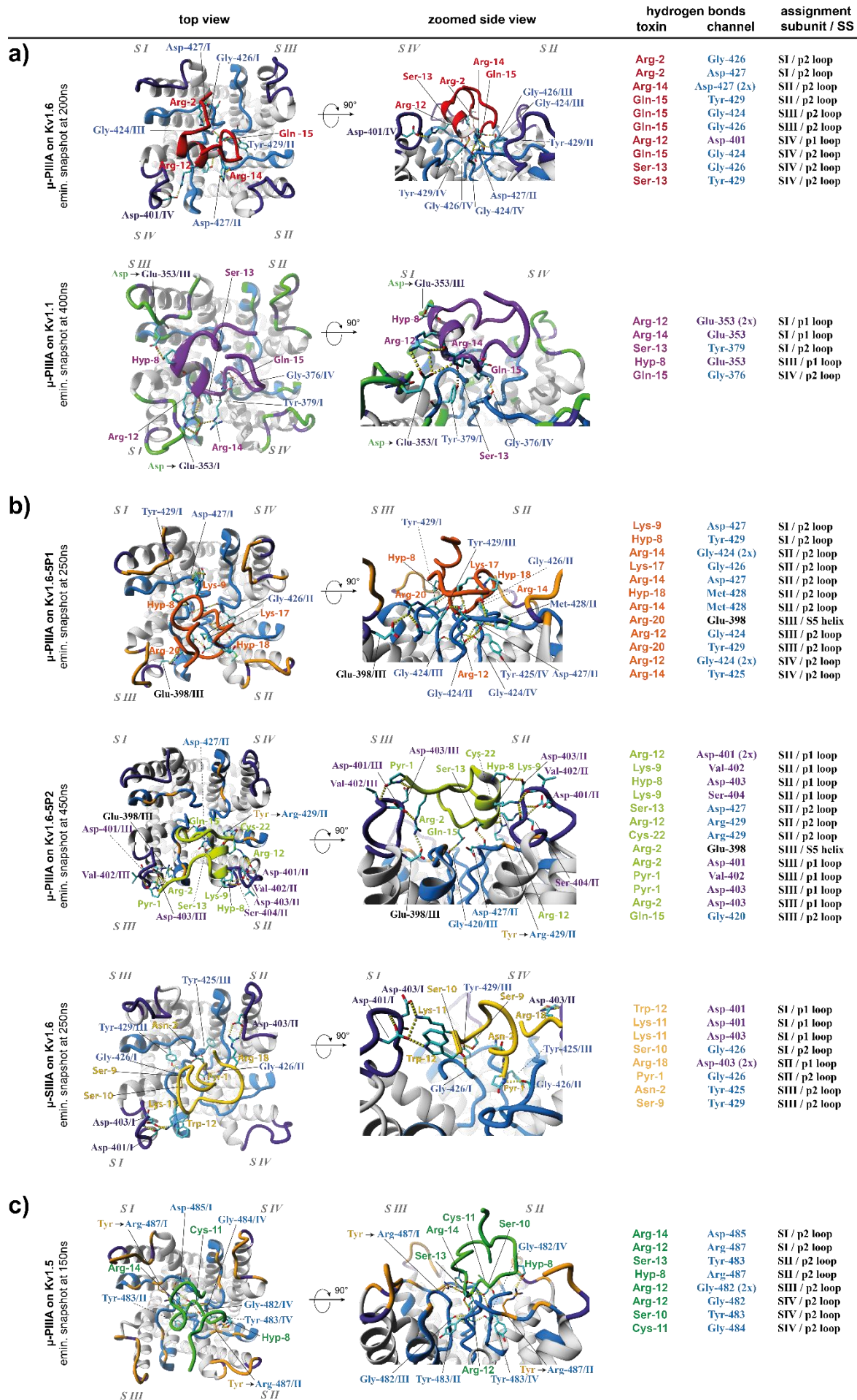


Figure 34: 3D visualization of the representative binding poses and corresponding hydrogen bond interactions shown for (a) active, (b) semi-active and (c) inactive systems. The shown snapshots are energy-minimized and were obtained from our clustering calculations. Left panel – top view, middle panel – side view, right panel – table listing the particular interactions. Toxin-channel colour scheme is adopted as indicated in **Figure 31**. Channel subunits are designated by Roman numerals (*italic gray*, SI – SIV). (Illustration modified from Kaufmann et al.⁶)

In summary, the key interactions are primarily built with neutral residues (tyrosine and/or glycine) of the inner P2 loops that reside close to the selectivity filter and coincidentally stabilized by hydrogen bonds to glutamic acid (for the case of Kv1.1) or aspartic acid (for the case of Kv1.6) in the P1 loops (**Figure 34a**).⁶ The increased number of hydrogen bonds of μ -PIIIA towards the P1 outer loop in Kv1.1 compared to Kv1.6 is mainly due to the replacement of aspartic acid of Kv1.6 by glutamic acid in Kv1.1 at the corresponding positions (**Figure 25a**, **Figure 34a**).⁶ For both cases the key binding motif consists of the adjacent residues Arg12, Ser13, Arg14 and Gln15, located in the central part of the toxin.⁶ In the binding mode to Kv1.1, the interacting segment of μ -PIIIA is expanded by the H-bond forming Hyp8, whereas in Kv1.6 Arg2 forms an additional stabilizing H-bond with the inner P2 residues of subunit I (SI) (**Figure 34a**).

Closer inspection of the central toxin region reveals that it lies within the α -helical section of μ -PIIIA whose amino acid sequence is composed as alternating symmetric pattern considering the physicochemical amino acid properties, primarily consisting of alternating basic or neutral amino acids, reaching from Hyp8 to Hyp18 (**Figure 35a**).⁶ Thereby, the Kv1.1 and Kv1.6 binding motif of the toxin corresponds to the central K–R–R–K-like motif (**Figure 35a**).⁶ Interestingly, the μ -SIIIA and μ -GIIIA lack such or a similar symmetric motif in the corresponding central regions, i.e. the regions that are aligned with Hyp8 – Hyp18 of μ -PIIIA (**Figure 35a**).⁶ However, at least μ -GIIIA features a remotely similar motif KK–K–R–K at corresponding positions (**Figure 35a**),⁶ but obviously not sufficient to block Kv1.6 or Kv1.1. Hence, as this distinct sequence pattern is exclusively carried by μ -PIIIA, it can be hypothesized that it is preserved by evolution as specific feature and may provide an explanation for the high μ -PIIIA specificity towards Kv1.1 and Kv1.6.

Further investigations of the central segment yielded that especially its densely packed helical conformation contributes to the pore-blocking position. Apparently, the α -helical central structure ensures (in the bound state) that the positively charged residues are directed towards the channel surface (**Figure 34a**, **Figure 35b**).⁶ Compared to μ -SIIIA and μ -GIIIA that lack the corresponding sequence and structural characteristics, the affected residues are uniformly oriented and are not buried or hidden by the non-helical fold of the toxin (**Figure 35b**).⁶ We assume that these features further strengthen the Kv1.1- or Kv1.6-bound state of μ -PIIIA by stabilizing the pore-blocking position of the interacting residues.⁶ For both cases however, Kv1.1 and Kv1.6, a horizontal, crosswise direction of the central symmetric toxin region (Hyp8 – Hyp18) parallel to the pore surface is enabled, supporting the overall coverage of the pore (**Figure 35**).

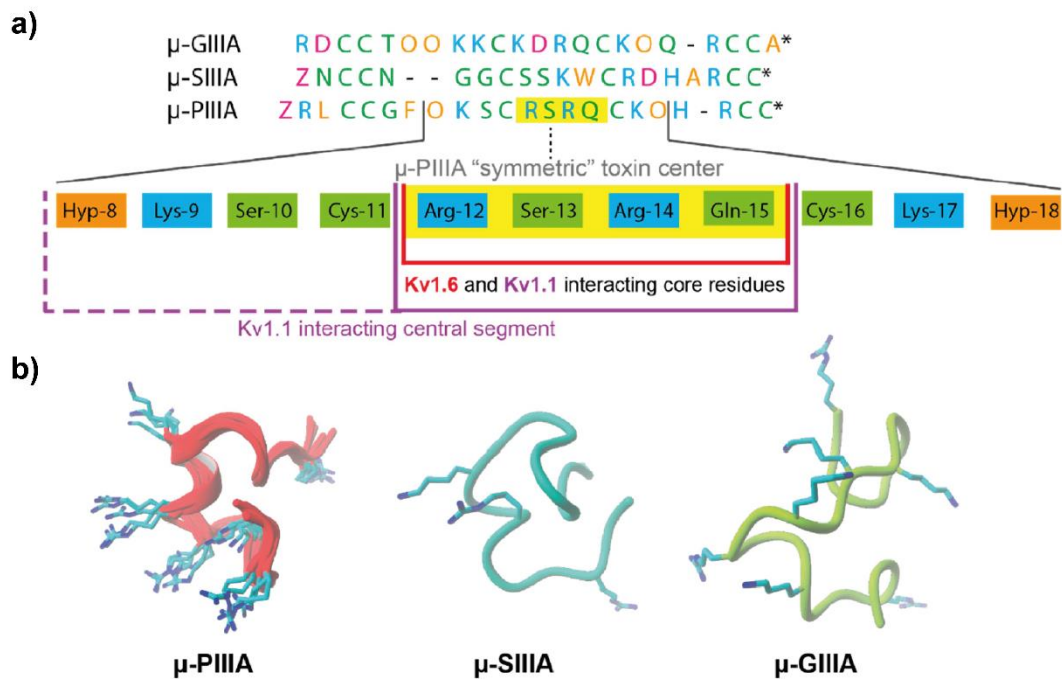


Figure 35: (a) Alignment of the μ -conotoxins μ -GIIIA, μ -SIIIA and μ -PIIIA. Residues are coloured by their physicochemical properties (acidic – purple, basic – cyan, nonpolar/hydrophobic – orange, polar/neutral – green). The “symmetric” μ -PIIIA centre is given below the alignment, highlighting the central interacting motif of μ -PIIIA in yellow. (b) K_V channel-bound toxin structures of μ -PIIIA (superposition of all docked structures), μ -SIIIA and μ -GIIIA after HADDOCK docking. (Image source: Kaufmann et al.⁶)

A similar constitution of a straight toxin centre, horizontally located on the pore with simultaneous pore-diving contacts cannot be found in semi- and inactive systems (**Figure 37**), again underpinning the uniqueness of the μ -PIIIA blocking mode towards $K_V1.1$ and $K_V1.6$ ⁶ and the presence of both of these attributes as a premise for achieving effective pore blockage.

The “kinked” deformation of the overall central segment (Hyp8 – Hyp18 corresponding residues), i.e. being not “in line” with the motif-corresponding segment (Arg12 – Gln15), either results in a mere pore protruding state of the core residues or in a crosswise positioning onto a channel moiety, such as can be seen for semi-active and inactive systems (**Figure 37**).⁶ The overall pore coverage is thus significantly reduced.⁶ For μ -GIIIA even the whole toxin centre remains completely flexible, impeding any blockage or binding event (**Figure 32**).

However, a general tendency of the toxins to maintain a more rigid and stable central part can be observed for all systems, which especially becomes evident when analyzing the per residue RMSFs, i.e. their flexibility over the toxin sequence when bound to the channel (**Figure 32**, **Figure 37**).⁶ In particular the symmetric toxin centre exhibits the lowest fluctuation rates for μ -PIIIA on $K_V1.1$ and on $K_V1.6$ being located in the global minimum area regarding the complete toxin sequences (**Figure 37**).⁶

On this basis one can also conclude for the other systems and in general that not the quantity of the interactions, but their “quality”, i.e. their stability and rigidity, but also the provoked toxin positioning, essentially constitutes the extent of a blocking event (**Figure 32**, **Figure 34**).

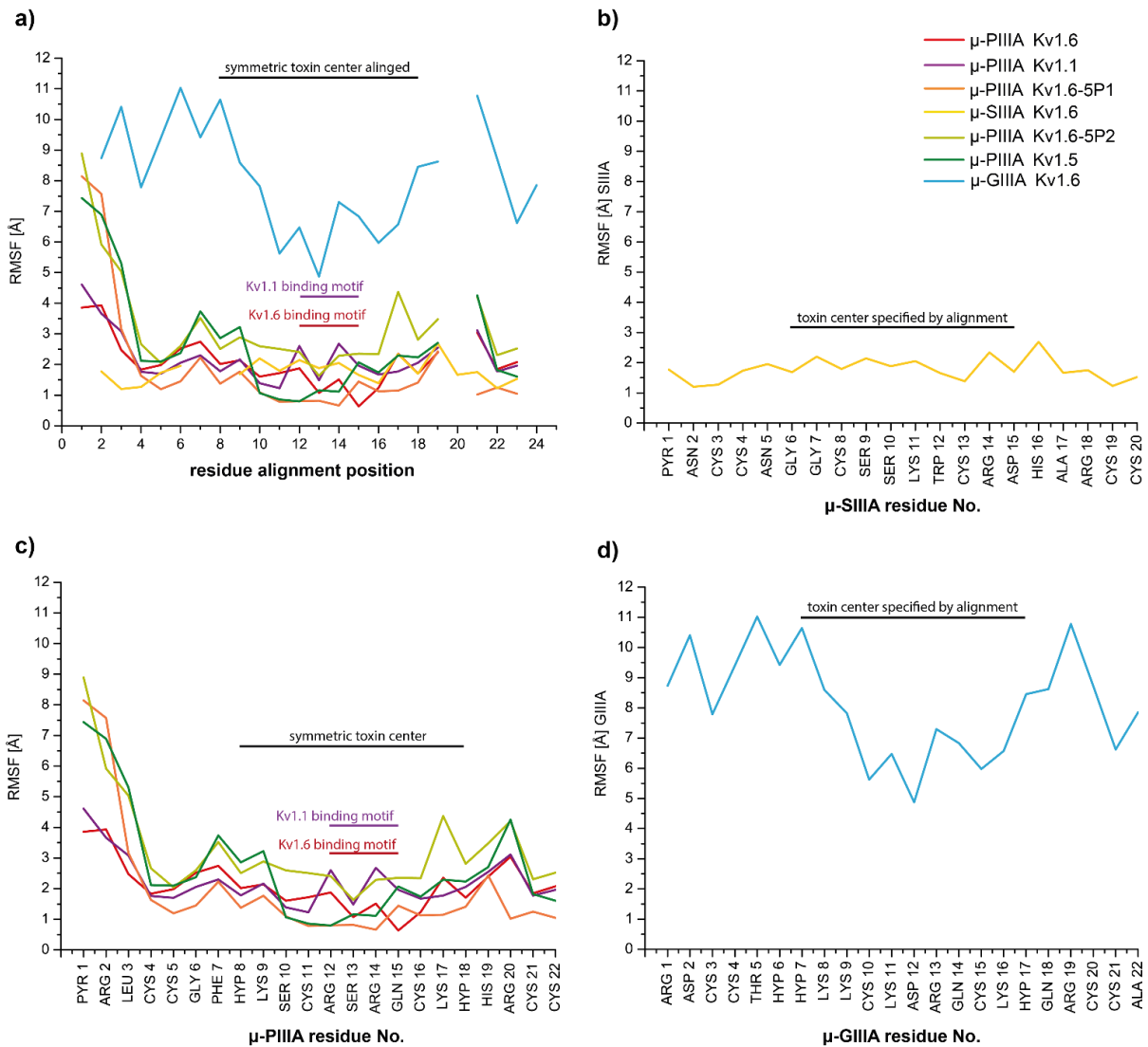


Figure 36: Per residue RMSF summary of the toxins shown per simulation (a) according to the toxin alignment as shown in Figure 35a, (b) for simulation μ -SIIIA on Kv1.6 – yellow, (c) for μ -PIIIA simulations (on Kv1.6 – red, on Kv1.1 – purple, on Kv1.6-5P1 – orange, on Kv1.6-5P2 – yellow-green, on Kv1.5 – dark green) and (d) for μ -GIIIA on Kv1.5 – cyan. (Image source: Kaufmann et al.⁶)

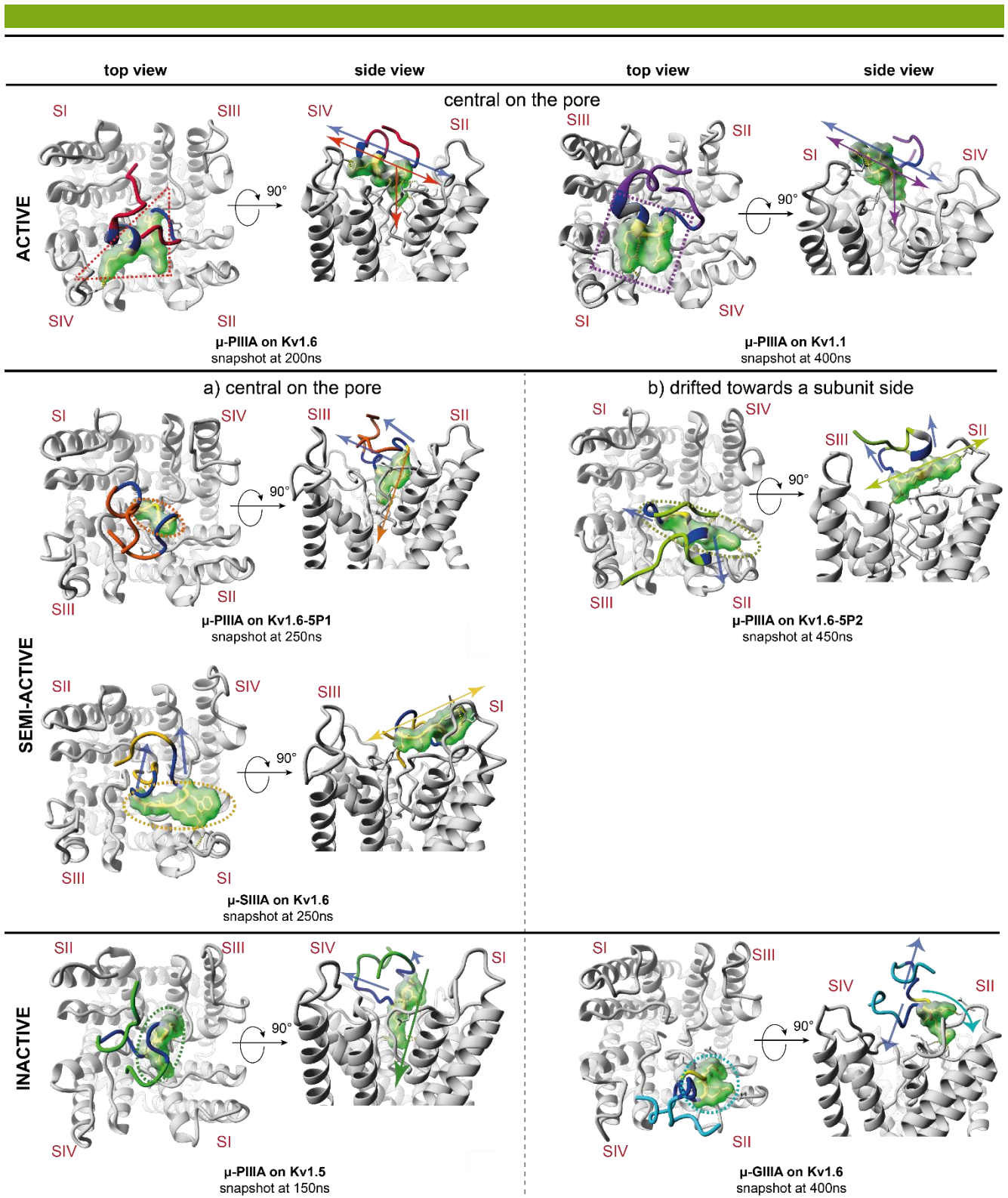


Figure 37: Overview of the channel bound toxin states regarding the orientation of the symmetric toxin centre for active (top), semi-active (middle) and inactive systems (bottom). The representative snapshots are shown in secondary structure representation (channels are coloured grey, residues which account for the μ -PIIIA core motif Arg12 – Gln15 are shown as yellow sticks, together with their molecular surface (transparent light green). Dashed lines indicate the overall channel coverage by the (binding) motif-corresponding core. The motif's surrounding area, corresponding to the symmetric μ -PIIIA central segment Hyp8–Hyp18 (see **Figure 35a**), is coloured dark blue, and the remaining toxin endings are coloured correspondingly to their individual systems (μ -PIIIA on Kv1.6 – red, μ -PIIIA on Kv1.1 – violet, μ -PIIIA on Kv1.6-5P1 – dark orange, μ -PIIIA on Kv1.6-5P2 – light green, μ -SIIIA on Kv1.6 – sun-yellow, μ -PIIIA on Kv1.5 dark green, μ -GIIIA on Kv1.6 – cyan). Delineated arrows indicate the orientations of the correspondingly coloured toxin centre regions with respect to the channel surface. Channel subunits are indicated by Roman numerals (red, SI–SIV). (Illustration modified from: Kaufmann et al.⁶)

Our observations considering the significance and structure of the symmetric toxin centre within μ -PIIIA and its positioning regarding channel blockage events are further supported by mutagenesis and molecular dynamics simulation data published by Chen et al.^{200, 221} that point out the highest significance of Arg12, Arg14, Lys17 and Arg20 for μ -PIIIA in the blockage of Nav1.4.⁶ Three out of these four (Arg12, Arg14 and Lys17) account for our stated the symmetric motif (**Figure 35a**) whereby the two residues Arg12 and Arg14 have also been found by our representative interaction analysis to form key players considering the blockage of Kv1.1 and Kv1.6 (**Figure 34a, Figure 35a**).

A further study suggests that Lys9 and Arg2 protrude into the Nav1.4 channel pore (beyond Arg14, Arg12, Lys17 and Arg20) as concluded from two equally populated binding modes,²⁰⁰ of which Lys9 and Arg2 also rank among the key interactive residues for the block of Kv1.6 according to our data (**Figure 34a, Figure 35, Figure 38**). The striking similarity between the related binding poses, comparing the binding towards Kv1.1 and a proposed pore blocking mode on Nav1.4 (including contacts from Lys9 and Arg14),²⁰⁰ further supports the crucial role of the symmetric segment considering the blockage of Kv and Nav channels by μ -PIIIA (**Figure 38**).⁶ All these common features and especially the demonstrated significance of the central Arg12 and Arg14 residues underpin the functional importance of the central toxin segment.⁶ In this context it is noteworthy that only a small part of μ -PIIIA, represented by the four amino acid long centric motif Arg12 – Gln15, seem to be sufficient for maintaining a stable blockage and an according suitable toxin positioning on the pore.

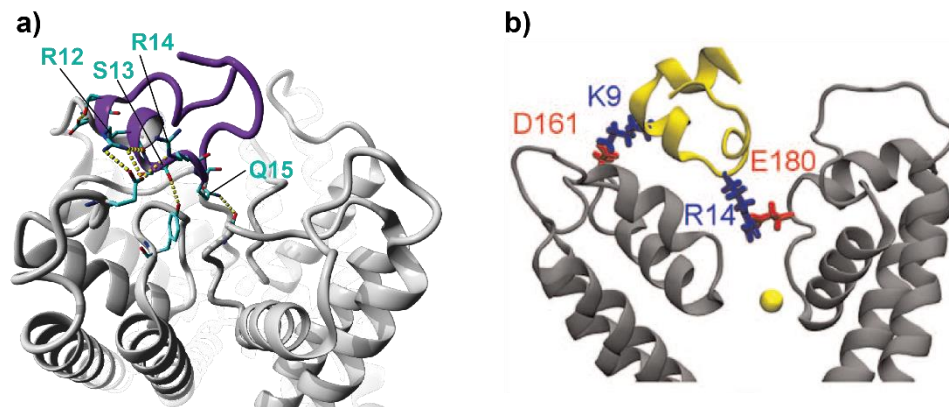


Figure 38: Comparison of predicted μ -PIIIA binding mode to (a) Kv1.1 as revealed in this work and to (b) Nav1.4 (Fig. adapted from Chen et al.²⁰⁰). (Illustration modified from Kaufmann et al.⁶)

μ -PIIIA on Kv1.6-5P1

Even though the case of Kv1.6-5P1 the selected snapshot pretends a blocked pore when μ -PIIIA was bound to the channel, it was completely lacking the necessary and stabilizing interactions of the μ -PIIIA residues with the P1 loops, as they were exchanged with P1 loops of Kv1.5 (**Figure 25a, Figure 34a and b, Figure 37**).⁶ More precisely, aspartic acid and the non-polar/hydrophobic valine and proline residues of the outer P1 loops are mutated to polar/neutral amino acids – except from leucine, which is mutated into basic histidine and except for the position of serine that remains neutral by a mutation to threonine, none of the physicochemical features from the outer Kv1.6 loop are maintained (**Figure 25a**). The outer loop modification into a mainly polar/neutral environment obviously triggers a toxin displacement and a coincident toxin deformation that differs from the equilibrating binding mode towards Kv1.1 or Kv1.6 (**Figure 25a, Figure 34a and b, Figure 37**).⁶

The resulting binding pose is accompanied with an increase of interactions with the center of the pore (**Figure 34b**), similar to the low affinity situation at the toxin-insensitive Kv1.5 channel. However, sole contacts to the inner P2 loops do not satisfy to reach a complete blocking condition, since the crucial simultaneous contact to the P1 loops of the outer channel loops is missing. Rather the symmetric toxin segment in its central part is tilted by $\sim 90^\circ$ with respect to the channel surface plane compared to the binding towards Kv1.1 and Kv1.6 (**Figure 34a and b, Figure 37**) resulting in a partial channel block as observed by Leipold et al.^{7.6}

The resulting toxin deformation is also reflected by the lack of contacts from the central Ser13 residue to the channel and also the “diving” key contact usually performed by Gln15 which is excluded from any interactions with K_V1.6-5P1 (**Figure 34a and b, Figure 37**).⁶

μ-PIIIA on K_V1.6-5P2

When polar/neutral Tyr429 in K_V1.6 – or the corresponding Tyr379 of K_V1.1 – is mutated into a positively charged arginine as for the case of K_V1.6-5P2 (**Figure 25a**), some of the crucial Ser13 and Gln15 interactions as observed for the binding on K_V1.6 and K_V1.1 are eliminated (**Figure 34b**).⁶

Likewise, key residue Arg14 is not involved in the representative interactions, however basic Arg12 forms an additional bond to the mutated position of the equally basic Arg429/SII, previously being placed by hydrophobic methionine (**Figure 25a, Figure 34b**).

As anticipated previously, positively charged Arg429 forces or pushes the toxin into another binding pose compared to K_V1.1 and K_V1.6 (**Figure 34a and b**).⁶ As a result, Ser13 and Gln15 form H-bonds with Asp427/SII and polar/neutral Gly420/SIII of the non-altered P1 loops (**Figure 34b**).⁶

Further on, the preferred addressing of μ-PIIIA towards the more acidic P1 outer loop is demonstrated by the interactions of the central Arg12, Lys9 and Hyp8 residues with Asp401/SII, Val402/SII, Asp403/SII and Ser404/SII (**Figure 34b**).⁶ Hence, a pore moiety is still partly covered.⁶ The observations indicating a partial pore coverage are further substantiated by the high conformational flexibility rates on the pore surface (**Figure 31, Figure 32**).⁶

The said modifications result in a complete shift of the toxin compared to its positioning on K_V1.1 or K_V1.6, such that the helical segment points towards the outside of the channel instead of a more central orientation (**Figure 37**). We assume that this partial covering state of the channel maintains for our pre-defined simulation conditions appearing as a possible circling movement of the toxin onto the channel surface. Most likely, this phenomenon accounts for the remaining currents in this system.

μ-PIIIA on K_V1.5

As could be shown previously with the cases of μ-PIIIA on K_V1.6-5P1 and μ-PIIIA on K_V1.6-5P2, the stepwise modification of the channel from K_V1.6 towards K_V1.5 consistently leads to a lack of most of the interactions observed for μ-PIIIA on K_V1.6: As neither the physicochemical composition of the outer nor the inner loops is maintained from K_V1.6 (**Figure 25a**), a different orientation of the toxin is induced compared to on K_V1.6/1, resulting in an upright positioning of the toxins core segment (Arg12 – Ser13 – Arg14) (**Figure 34c**).⁶ Consequently, instead of forming contacts to the outer loops like in K_V1.6 and K_V1.1 the central Arg12 residue interacts with inner residues at the pore cavity resulting in a plethora of further toxin centre H-bonds towards residues next to the selectivity filter motif (**Figure 34c**).⁶ Thus a much smaller portion of the pore surface is covered which suggests a low affinity binding pose (**Figure 34c, Figure 37**).⁶

Consistently with the inactivity of μ-PIIIA at K_V1.5, VINA rescoring suggests the second lowest binding energy among all toxin-channel systems analysed in this work (**Table 6**).⁶ Equally, the high degree of local per residue flexibility rates outside the central core of μ-PIIIA (**Figure 32**) underpin the poor pore coverage.⁶

All in all, the cases of μ-PIIIA binding to various K_V channels considered so far teach us that the composition of the residues in the inner P2 loops contributes significantly to the μ-PIIIA binding and its orientation on K_V channels.⁶

μ -SIIIA on Kv1.6

Since μ -SIIIA shows a similar moderate activity as the semi-active instances of μ -PIIIA on the Kv1.6-1.5 chimera, we also investigated the binding of μ -SIIIA on Kv1.6 (**Figure 34b**).⁶

Despite the identical disulfide-bond patterns, sequence comparison of μ -SIIIA and μ -PIIIA reveals overall different compositions and a somewhat different central segment, lacking the symmetric sequence pattern (**Figure 18, Figure 35a**).⁶

Interestingly, our analysis unveiled that the central region of μ -SIIIA (Ser9 to Trp12), which corresponds to the symmetric centre of μ -PIIIA, was mainly forming H-bonds with the central pore region, as it was observed for μ -PIIIA binding at Kv1.6-5P2 (**Figure 34b, Figure 35a**).⁶ Similar to Gln15 within μ -PIIIA, polar/neutral Ser9 and Ser10 respectively address Tyr429/SIII and Gly426/SI of the inner P2 loops (**Figure 34b**).⁶

The further contacts to the outer P1 loops of the opposite subunits SI and SII altogether stabilize an overall centrally oriented positioning of the toxin at the pore, despite involving a sideward oriented central core of μ -SIIIA (**Figure 34b, Figure 37**). More precisely, these concern the symmetric hydrogen bonds of basic Lys11 and acidic Trp12 to the outer located acidic Asp401/SI and Asp403/SI and of the equally basic Arg18 to the proportionate residue Asp403/SII of the opposite subunit SII (**Figure 34b**).⁶

Remarkably, μ -PIIIA's centrally located Arg12 residue equally connects to Asp401 in Kv1.6 (**Figure 34a**), which also endorses the overall central positioning of μ -SIIIA.⁶ The resulting positioning allows for the interaction of the starting residues with the pore centre, more specifically of Pyr1 to Gly426/SII and Asn2 to Tyr425/SIII at the inner P2 loops of the pore center, however resulting in an incomplete pore block (**Figure 34b**).⁶ This is due to a more lateral position of the toxin core on the pore compared to μ -PIIIA on Kv1.6 despite the overall central toxin position given by the contacts to the opposing subunits SI and SII (**Figure 34, Figure 37**).⁶

The incomplete block can further be substantiated by the conformational and local flexibilities of the outer P1 loops provoking a decrease of the binding pose's stability (**Figure 32**).⁶

μ -GIIIA on Kv1.6

In accordance with experimental studies, all our data show clearly that μ -GIIIA is not capable of binding to Kv1.6. This is especially reflected by the low binding energy of μ -GIIIA towards Kv1.6 revealed from our VINA resoring calculations as well as the high motion and dynamics of this system during the equilibration simulations.⁶ Therefore, in this case no further statements about a bound state can be made.

4.4.4. Conclusions and future perspectives

Our *in silico* data could provide thorough insights into the dynamics and different adopted binding states of the μ -conotoxins μ -PIIIA, μ -SIIIA and μ -GIIIA on the voltage-gated potassium channels of the K_V1 family.⁶

The data clearly indicate that different system activities can only adequately be understood under consideration of their dynamic information and the conformational space the toxin can sample in its channel-bound state.⁶ Since for all of the channel-toxin systems investigated a certain flexibility of their conformational arrangement as well as deviations from the initial docked state could be shown through our molecular dynamics simulations (**Figure 31**, **Table 8**), we hypothesize that these flexibilities ultimately explain remaining currents as especially observed for semi- and inactive systems.⁶ In particular, the semi-active and inactive systems μ -PIIIA/ $K_V1.6$ -5P2 and μ -GIIIA/ $K_V1.6$ showed significantly higher conformational flexibilities than the other semi- and inactive systems, further rationalizing their incomplete pore coverage.⁶ More precisely, this is demonstrated by the clearly visible sideward orientation of their toxins.⁶ In contrast, less flexible semi-active systems μ -PIIIA/ $K_V1.6$ -5P1 and μ -SIIIA/ $K_V1.6$ and inactive μ -PIIIA/ $K_V1.5$ whose toxins remain more centrally located cannot be fully rationalized by this criterion, however do not exhibit the interactions identified in the active systems, resulting in a more detached toxin positioning.⁶ In contrast, the unique active systems are generally less flexible and dynamic, exhibiting more stabilized and rigidly bound toxins.

Altogether the two different modes of interaction observable for semi- and inactive systems can be distinguished by more centered toxin positions (in less flexible systems) and more sideways drifting toxin movements (in more flexible systems). Our findings further suggest a similar and unique binding mode for the active blockage of the channels $K_V1.1$ and $K_V1.6$ through the μ -conotoxin μ -PIIIA.⁶ This is made up by a “pore-diving” and coincident “crosswise” positioning of the central toxin core residues onto the channel surface, achieved by toxin contacts to the inner and the outer pore loops.⁶ The resulting central toxin position is made up by a plethora of hydrogen bonds formed by the μ -PIIIA α -helical core residues with the central pore duct and with outer surface residues.⁶ In these respect it was found that for μ -PIIIA binding at $K_V1.6$, $K_V1.5$ and the respective chimeras the inner P2 loop composition mainly destines the resulting toxin orientation where further contacts to the outer P1 loops stabilize the pose.⁶ Further, the crucial role of the “physicochemically symmetric” μ -PIIIA helical core segment Hyp8 – Hyp18 could be pointed out: in particular the arginines within the centric Arg12 – Ser13 – Arg14 – Gln15 binding motif included (**Figure 35a**) show a stabilizing effect on the positioning of the toxin onto the pore.⁶ The importance of the central μ -PIIIA region on $K_V1.1/6$ has also been shown by further studies suggesting comparable bound states for μ -PIIIA on $Nav1.4$ (**Figure 38**); here, stabilizing contacts of Arg12 and Arg14 to the channel are also reported to strongly influence the channel blocking efficacy.^{6, 200, 221}

In addition, the predicted μ -PIIIA orientation on $K_V1.6$ was also similar to the experimentally determined binding pose of the structurally related μ -conotoxin KIIIA on $Nav1.2$ (PDB ID 6J8E) further suggesting a unique pore blocking motif included in μ -conotoxins (data not shown).^{6, 222}

Interestingly, partial mutations such as the insertion of the μ -PIIIA-insensitive $K_V1.5$ channel pore loop residues into $K_V1.6$ could be shown to trigger a reorientation of the toxin in the sense of either clearly sideward or pore-diving toxin cores (**Figure 37**).⁶ A similar sideward oriented toxin centre could also be observed for the corresponding region within μ -SIIIA (**Figure 37**).⁶

The interactions which stabilize the toxin placement are mainly defined by positively charged Arg or Lys toxin residues contacting the negatively charged Asp or Glu residues of $K_V1.6$ or $K_V1.1$, respectively, being located in the outer P1 loops.⁶ Here in particular the third Asp401 in the case of $K_V1.6$ and the corresponding Glu353 in $K_V1.1$, both being located closer to the pore centre compared to the other Asp and Glu residues, are addressed.⁶ Related analyses pointed out that the subtype-specific binding of μ -PIIIA to the K_V1 family members does not tolerate positively charged amino acids in both the inner and outer pore loops, as shown by comparison of $K_V1.6$, $K_V1.3$ and $K_V1.1$ (**Figure 25a**).⁶ Rather, a strong reorientation and repulsion of the toxin is enforced by such amino acid constellations which is deducible from our data provided for μ -PIIIA/ $K_V1.6$ -5P2.⁶ Especially the lack of the essential third aspartic acid residue in the outer P1 loop of $K_V1.3$ points out the exclusive μ -PIIIA binding to $K_V1.6$ and $K_V1.1$.⁶ However, as the Asp- and Glu-rich $K_V1.6/1$ outer P1 loops are rather small in size, a very compact toxin fold is required to achieve a proper pore blockage where the toxin's Arg and Lys residues can simultaneously orient towards the outer channel surface and the inner P2 loops.⁶ The lack of such residue constellations in μ -GIIIA can, among others, provide an explanation for its inactivity (**Figure 35b**).⁶

Lastly, even though our MD simulations were not fully able to unveil a toxin's conformational landscape, the importance of system dynamics even for the channel-bound toxin state could be proven with our *in silico* studies after all, encouraging our strong recommendation of combined docking and MD simulation approaches concerning *in silico* analyses of protein-ligand systems and their coincident evaluation.⁶

Future perspectives

The new insights into subtype-specific channel blockage by μ -conotoxins can provide a basis and new impulses for future rational peptide-based design approaches. Here, especially the more targeted and specific addressing of distinct ion channel subtypes should be focused, for instance for the purpose of preventing unintended side effects.

Additional discoveries considering recurrent modes of toxin movement events in semi-active and inactive systems (that either exhibit centrally or sideways oriented toxins) allow us to take a subsequent step towards a proposal of a preliminary classification model for the description of toxin movement patterns in channel-toxin systems: unique binding modes, being assumed in an equilibrated state as observed for active systems can be distinguished from two further types of toxin detachment movements attributable to semi- and inactive systems. In these respects we formulated the corresponding recognition determinants for each mode of activity based on our simulation data analyses (**Table 9**). Although too few determinants may still be defined for an extensive automated learning or classification procedure, we assume that our stated modes can also be transferred to systems that haven't been investigated in more detail yet (**Table 3**).

Table 9: Preliminary classification model for the categorization of channel-toxin systems established from our simulation data evaluations; assignments can be done according to our suggested classification scheme (to active, semi-active (a) / (b) and inactive (a) / (b) modes) and on the basis of different determinants: cluster-specific features – highlighted gray, interaction-specific features – highlighted pale blue and core-position and coincident shape-specific features – highlighted light turquoise; for unique classification, each mode is assigned an individual combination of determinant settings.

Mode of activity	Main cluster 2 size ≥ 3	Index of preliminary representative is located in \sim the first three sim. quarters (e [0ns, 350ns])	Revising step positive	Representative interactions to at least three subunits	interactions to P1 loops / interactions to P2 loops > 0	Position and shape of residues corresponding, i.e. aligned to μ -PIIIA core motif (Arg12 – Gln15): "diving" / "horizontal" (central or lateral w.r.t. channel) Please refer to Figure 37	Resulting shape of toxin central segment (corresponding to μ -PIIIA Hyp8 – Hyp18) w.r.t. core motif residues (corresponding to μ -PIIIA core motif Arg12 – Gln15) Please refer to Figure 37
active (unique)	yes	yes	no	yes	yes	diving AND horizontal (central)	stretched out / "in line"
semi-active a)	yes	yes	no	yes	no	diving OR horizontal (lateral)	bending
semi-active b)	no	yes	yes	no	yes	horizontal (lateral)	bending
inactive a)	no	yes	no	yes	no	diving	bending
inactive b)	no	no	no	no	yes	none	bending

Considering all evaluations conducted, we conclude that the active systems μ -PIIIA on Kv1.6 and Kv1.1 can serve as prominent model systems and will form an instructive basis regarding further developments and the testing of newly designed ion channel blockers. In consequence, μ -PIIIA holds a very high capacity to serve as precursor for the detection of preliminary μ -conotoxin-pharmacophores, being equipped by its unequivocal core-motif embedded in a physicochemically symmetric central toxin segment that lack its competitors μ -SIIIA and μ -GIIIA. Its additional potential of being adjusted for an exclusive targeting of Nav and various other channels by pointed modifications further encourages its employment as lead structure.

Therefore, motivated by our findings and by a further reported correlation of the overexpression of Kv1.5 channels in some human cancer cell lines,²²³ we elaborated a further preliminary proposal for a putative Kv1.5 blocker, derived from μ -PIIIA as precursor structure and its discovered binding mode to Kv1.6 (**Figure 39**). Additional research on various studies indicating the sole presence of reported small molecule compound and no larger polypeptide blockers of the Kv1.5 channel^{7, 216, 224–226} further motivated our ideas of hypothesized μ -PIIIA redesigns.

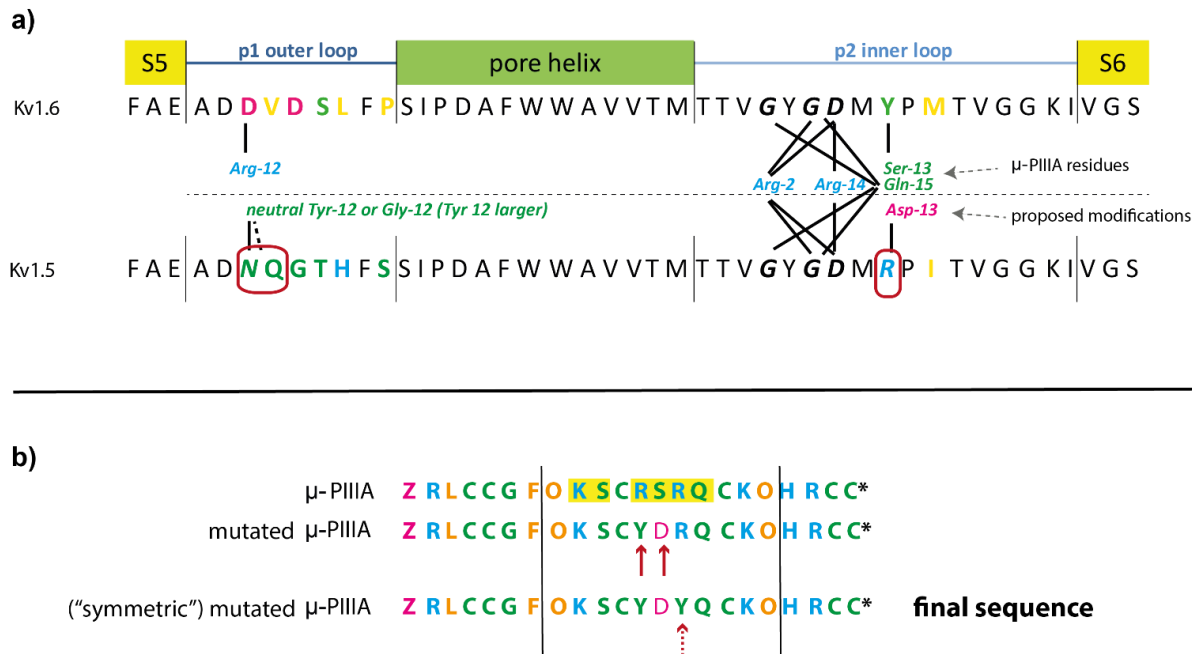


Figure 39: Attempt to derive a Kv1.5 blocker toxin based on the active model system μ -PIIIA on Kv1.6. **(a)** Kv1.5 aligned to Kv1.6 as also obtained from **Figure 25a**; the differing residues are colored by physicochemical properties (alkaline – light blue, acidic – magenta, polar/neutral – green, non-polar/hydrophobic – orange); μ -PIIIA-Kv1.6 interacting residues are equally coloured and indicated below the Kv1.6 sequence; interactions to Kv1.6 (and possible interactions to Kv1.5) are shown as black lines between the channel and the toxin residues; proposed modifications for μ -PIIIA are stated above the Kv1.5 sequence, separated from the original μ -PIIIA residues by a dashed line. **(b)** Resulting modified sequence of μ -PIIIA, also including a symmetric physicochemical sequence pattern (and unmodified cysteines); amino acids are equally coloured by amino acid type.

Thereby we propose the replacement of μ -PIIIA residues that interact with Kv1.6 residues of mutated positions arising by the modification from Kv1.6 to Kv1.5 and causing the ineffective binding mode (**Figure 25a**, **Figure 34**, **Figure 39**). For the replacement, residues empirically known to contact the concerning channel positions would be considered. By coincidentally ensuring the maintenance of the outlined essential features such as the presence of an internal physicochemical symmetry and the μ -conotoxin disulfide bridge pattern, we obtain a modified sequence of μ -PIIIA as shown in **Figure 39b**.

However, it should be kept in mind that other arising interactions such as Hyp8 to Arg487/SII of μ -PIIIA to Kv1.5 should be considered when performing such mutations (**Figure 34**).

In addition it should be recalled from these perspectives, that in recent protein structure prediction research there has been a great interest using amino acid interaction preferences (e.g. contact potentials or potentials of mean force) to align or “thread” a protein sequence to a known structural motif. In these respects it could be shown that finding the globally optimal protein threading is NP-hard²²⁷ and we therefore stucked to a more intuitive approach considering our outlined proposal.

To address the open questions outlined – peptide-based ligand optimizations and the extension of the predefined determinants in our established classification model (**Table 9**) – further in silico investigations of not yet examined systems would be required, especially in view of possible training data setups. In addition, the presumed significance of found internal toxin symmetries or sequence motifs for the blockage of specific ion channels could be verified or better understood.

A further topic of interest considering these respects and the elucidation of its severe side effect causes would be provided by distinct investigations of the ω -conotoxin ω -MVIIA on Kv1.1, Kv1.6 and Nav1.4 as targeted channels.

Lastly, further in silico analyses and the ongoing increase of knowledge with the assistance of modern technologies will hopefully contribute to the specification and optimization of therapeutic pain or even cancer treatment towards a substantial reduction of unintended side-effects. Our study results uncovering the binding of μ -PIIIA on Kv1.1 as well as on Kv1.6 and others, being aware of its intrinsic specificity towards Nav1.4, may already provide essential contributions at least for some of these aspects.

5. Methods section

5.1. G_q – YM / FR (analogues) docking protocol

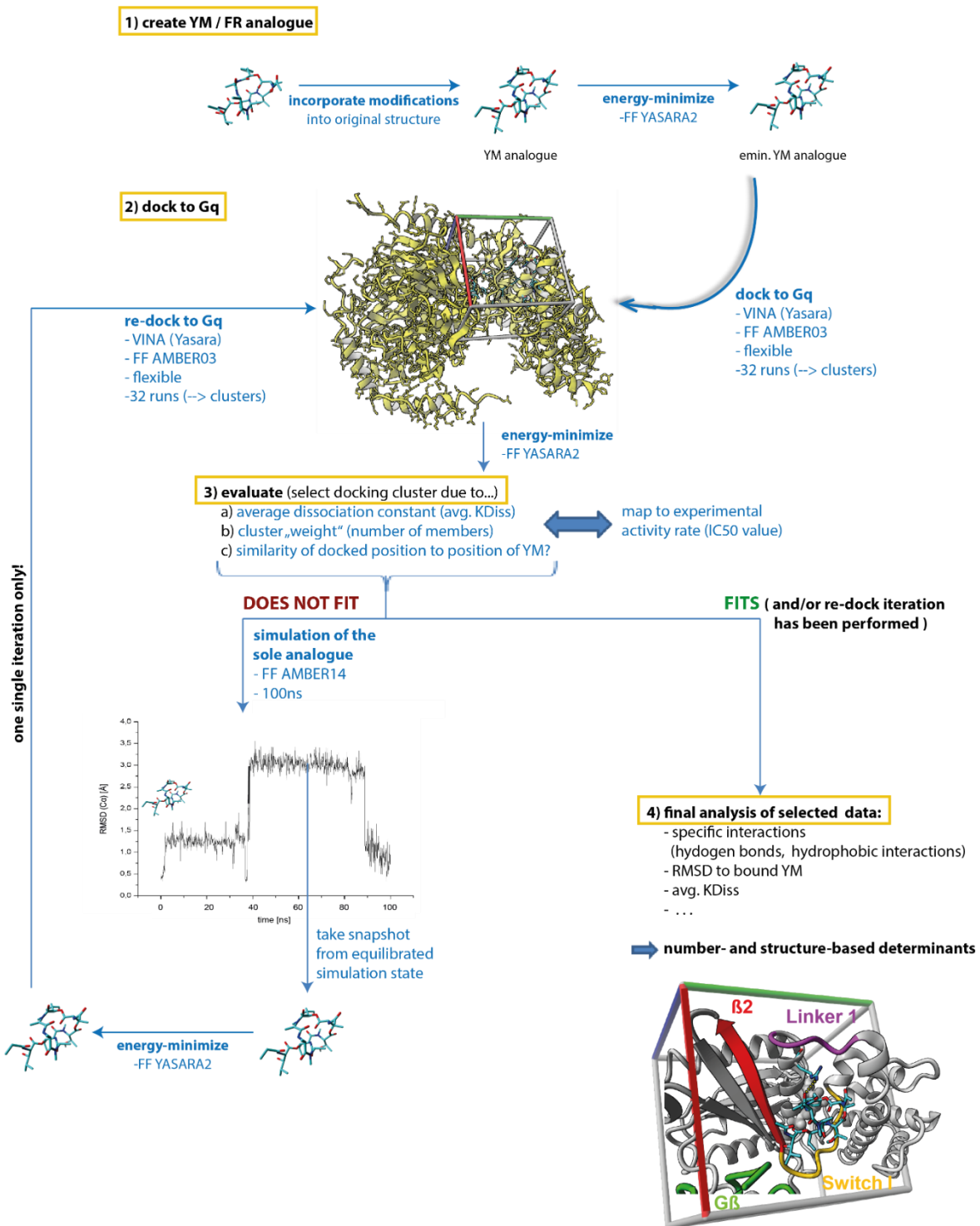


Figure 40: Schematic overview of the complete docking workflow as applied for the analysis of docked YM/FR analogues.

5.1.1. In silico generation of YM-254890 and FR900359 analogues

The molecular structures of the YM-254890 and FR900359 analogues were built on the basis of the G_q-bound YM-254890 crystal structure (PDB ID 3AH8¹), employing the YASARA molecular modelling program^{44, 228} (**Figure 40**, step 1). Each structure was geometry optimized using the YASARA YAPAC module. Therefore, we employed semi-empirical quantum chemical methods followed by an energy minimization step in explicit water (**Figure 40**, step 1). The energy minimization was performed under physiological conditions (0.9% NaCl, 298 K, pH 7.4²²⁹) using the PME method²²⁹ and a cut-off distance of 8 Å to describe long-range electrostatics as well as making use of the YASARA2 force field²³⁰ (**Figure 40**, step 1).

5.1.2. VINA Docking and energy minimizations

The G_{α_q} target structure was extracted again from the YM-bound G_q structure (PDB ID 3AH8¹). The docking setup was accomplished with the aid of the YASARA molecular modelling program,^{44, 228} applying default parameter settings for the actual docking process of the prepared analogues which was then carried out employing the Vina docking algorithm (**Figure 40**, step 2).²²⁰ The virtual docking cell was placed around the YM(/FR) binding epitope as reported by Nishimura et al.¹ in order to optimally guide the docking runs. Moreover, ligands and receptor residues were kept flexible during the docking.

The respective best scoring result of 32 individual docking runs was subjected to succeeding energy minimization runs prior to subsequent G_{α_q}-inhibitor interaction analyses (**Figure 40**, step 3 case “fits”, step 4). The individual energy minimization steps were carried out by employing steepest-descent minimizations followed by simulated annealing minimizations until convergence (<0.05 kJ/mol/200/steps). Ewald summation²²⁹ was used to assign charged amino acids according to their individual predicted side chain pK_a values. Further neutralization of the (previously) charged amino acids was achieved by the adding of counter ions (NaCl). The further assignment of the ligands' force field parameters was performed applying YASARA's AutoSMILES approach.^{231, 232}

Unless otherwise stated TIP3P water model as well as the Yasara2 force field²³⁰ was used to conduct the energy minimizations^{233, 234} employing the PME method²³⁵ to describe long-range electrostatics at a cut-off distance of 8 Å at physiological conditions (0.9% NaCl, 298 K, pH 7.4²²⁹).

5.1.3. MD simulations of selected YM/FR analogues

From all best scoring docking results a real subset of outlier structures was determined that, regarding the average dissociation constant given by the according docking experiment, did not appropriately reflect their corresponding IC₅₀ values (**Figure 40**, step 3 case “does not fit”). If the initial docking result didn't properly fit the experimentally expected outcomes, we undertook further searches for a more reliable and stable conformation of the structure, being able to be re-docked for further analyses. To that end, we performed molecular dynamic simulations experiments of 100ns duration on all these selected outlier structures, employing the AMBER14 force field.²³⁶

For re-docking the most stable conformations from the simulations were selected as indicated by the C_α backbone atom RMSD (**Figure 40**, step 3 case “does not fit”). To allow for approach consistency, the selected structures were energy-minimized prior to and after the re-dockings. To carry out the latter, we made use of the approaches as described in the previous section (5.1.2). Interaction analysis (hydrophobic interactions and hydrogen bond analysis, **Table 11**) was performed on the representative ligand-protein complex as outlined in **Table 11** after each structure was energy minimized as described above.

5.2. G_q – YM / FR supplementary docking data

Table 10: Binding energy calculated from the docking experiments before (1st docking run) and after the 100ns equilibration step.

Outlier	Activity level	Docked binding energy from 1 st docking run [kcal/mol]	Docked binding energy after previous equilibration [kcal/mol] / (sim. time of selected structure)
YM-254890 (YM)	high	8.58 ± 0.54	-
FR900359 (FR)	high	8.63 ± 0.67	-
21 (YM-14)	high	7.70 ± 0.49	8.83 ± 0.38 (40ns)
18 (YM-11)	high	7.97 ± 0.51	9.01 ± 0.38 (79ns)
23 (YM-1)	high	8.22 ± 0.73	9.18 ± 0.51 (20ns)
30 (YM-9)	medium	8.33 ± 0.69	9.55 ± 0.50 (45ns)
31 (YM-3)	medium	9.02 ± 0.24	8.80 ± 0.79 (52ns)
29 (YM-L / YM-385780)	medium	8.16 ± 0.50	8.04 ± 0.30 (65ns)
32 (YM-4)	low	8.50 ± 0.47	8.50 ± 0.34 (35ns)
33 (YM-6)	low	8.79 ± 0.78	9.83 ± 0.34 (52ns)
34 (YM-17)	low	8.40 ± 0.78	8.14 ± 0.44 (37ns)
35 (YM-5)	low	8.62 ± 0.78	9.28 ± 0.14 (48ns)

Table 11: Overview of G_q-inhibitor interactions. Hydrophobic interaction cutoffs: strength ≥ 1.0 ; length $\leq 5\text{\AA}$. (Table source: Reher et al.⁵)

G _q – inhibitor interactions						
	Hydrogen bonds		Hydrophobic interactions strength ≥ 1.0 length $\leq 5\text{\AA}$			
	Gq residue	Inhibitor residue	Gq residue	Inhibitor residue	Gq residue	Inhibitor residue
FR900359	NH2 ARG 60 (A) N GLU 191 (A) O GLU 191 (A)	O N-MeDha5 O γ Ac- β -HyLeu1 N Ac- β -HyLeu1	ILE 56 (A) LYS 57 (A) TYR 67 (A) PHE 75 (A) PHE 75 (A)	D-Pla6 D-Pla6 D-Pla6 N-MeThr(OMe)8 Ac- β -HyLeu7	VAL 184 (A) VAL 184 (A) ILE 190 (A) TYR 192 (A) PRO 193 (A)	N-MeThr(OMe)8 D-Pla6 β -HyLeu2 N-MeDha5 Ac- β -HyLeu1
cis-FR900359	NH1 ARG 60 (A) NH1 ARG 60 (A) OH TYR 67 (A)	O D-Pla 6 O β -HyLeu7 O D-Pla 6	SER 53 (A) ILE 56 (A) LYS 57 (A) PHE 75 (A) Leu 78 (A) VAL 184 (A)	D-Pla6 D-Pla6 D-Pla6 β -HyLeu7 N-MeThr(OMe)8 D-Pla 6	THR 186 (A) ILE 190 (A) ILE 190 (A) TYR 192 (A) ARG 96 (B)	D-Pla 6 D-Pla 6 N-MeDha5 N-MeDha5 β -HyLeu2
YM-254890	NH1 ARG 60 (A) NH1 ARG 60 (A) NH2 ARG 60 (A) O ILE 189 (A) N GLU 191 (A) O GLU 191 (A)	O α D-Pla6 O-acetyl L-Thr7 O-acetyl L-Thr7 O γ Ac- β -HyLeu1 O γ Ac- β -HyLeu1 N Ac- β -HyLeu1	SER 53 (A) ILE 56 (A) LYS 57 (A) PHE 75 (A) VAL 184 (A)	D-Pla6 D-Pla6 D-Pla6 L-Thr7 N-MeThr(OMe)8	VAL 184 (A) THR 187 (A) ILE 190 (A) PRO 193 (A) ASP 118 (B)	D-Pla6 β -HyLeu2 D-Pla6 Ac- β -HyLeu1 Ac- β -HyLeu1
Best Fit						
15 (YM-10 / YM-254891)	NH1 ARG 60 (A) N GLU 191 (A) OE2 GLU 191 (A) NH2 ARG 96 (B)	O α D-Pla6 O Prop β -HyLeu1 O γ Prop β -HyLeu1 O γ Prop β -HyLeu1	SER 53 (A) ILE 56 (A) LYS 57 (A) ARG 60 (A) TYR 67 (A) PHE 75 (A) PHE 75 (A)	D-Pla6 D-Pla6 D-Pla6 D-Pla6 D-Pla6 N-MeThr(OMe)8 D-Pla6	VAL 184 (A) VAL 184 (A) ILE 190 (A) ILE 190 (A) TYR 192 (A) ILE 120 (B)	N-MeThr(OMe)8 D-Pla6 β -HyLeu2 D-Pla6 N-MeDha5 Prop β -HyLeu1
21 (YM-14)	NH1 ARG 60 (A) NH1 ARG 60 (A) OH TYR 67 (A) O ILE 189 (A) N GLU 191 (A)	O α D-Pla6 O-acetyl L-Thr7 O D-Pla6 O γ β -HyLeu1 O β -HyLeu1	ILE 56 (A) LYS 57 (A) ARG 60 (A) PHE 75 (A)	D-Pla6 D-Pla6 D-Pla6 D-Pla6	VAL 184 (A) ILE 190 (A) TYR 192 (A) ILE 120 (B)	N-MeThr(OMe)8 β -HyLeu2 N-MeDha5 Ac- β -HyLeu1
20 (YM-13)	NH1 ARG 60 (A) NH2 ARG 60 (A) OH TYR 67 (A) N GLU 191 (A) OE1 GLU 191 (A) OE1 GLU 191 (A)	O α D-Pla6 O-acetyl L-Thr7 O D-Pla6 O Ac- β -HyLeu1 N Ac- β -HyLeu1 O γ Ac- β -HyLeu1	SER 53 (A) ILE 56 (A) LYS 57 (A) ARG 60 (A) TYR 67 (A) PHE 75 (A)	D-Pla6 D-Pla6 D-Pla6 D-Pla6 D-Pla6 N-MeThr(OMe)8	PHE 75 (A) VAL 184 (A) VAL 184 (A) ILE 190 (A) ILE 190 (A) TYR 192 (A)	D-Pla6 N-MeThr(OMe)8 D-Pla6 β -HyLeu2 D-Pla6 N-MeDha5
22 (YM-18)	NH1 ARG 60 (A) N GLU 191 (A)	O α D-Pla6 O Ac- β -HyLeu1	SER 53 (A) ILE 56 (A) LYS 57 (A) ARG 60 (A) VAL 184 (A)	D-Pla6 D-Pla6 D-Pla6 D-Pla6 N-MeThr(OMe)8	VAL 184 (A) ILE 190 (A) TYR 192 (A) TYR 192 (A) PRO 193 (A)	D-Pla6 D-Pla6 N-MeDha5 Phe4 Phe4
18 (YM-11)	NH2 ARG 60 (A) OE1 GLU 191 (A) OE1 GLU 191 (A) NH2 ARG 96 (B)	O-acetyl L-Thr7 N Ac- β -HyLeu1 O γ Ac- β -HyLeu1 O γ Ac- β -HyLeu1	ILE 56 (A) LYS 57 (A) TYR 67 (A) PHE 75 (A)	D-Pla6 D-Pla6 D-Pla6 D-Pla6	VAL 184 (A) ILE 190 (A) TYR 192 (A) ILE 120 (B)	N-MeThr(OMe)8 β -HyLeu2 N-MeDha5 Ac- β -HyLeu1
19 (YM-12)	NH1 ARG 60 (A) NH1 ARG 60 (A) N GLU 191 (A) OE1 GLU 191 (A)	O α D-Pla6 O-acetyl L-Thr7 O Ac- β -HyLeu1. N Ac- β -HyLeu1	SER 53 (A) ILE 56 (A) LYS 57 (A) TYR 67 (A) PHE 75 (A)	D-Pla6 D-Pla6 D-Pla6 D-Pla6 N-MeThr(OMe)8	PHE 75 (A) VAL 184 (A) ILE 190 (A) ILE 190 (A) TYR 192 (A) ILE 120 (B)	D-Pla6 N-MeThr(OMe)8 β -HyLeu2 D-Pla6 N-MeDha5 Ac- β -HyLeu1
23 (YM-1)	NH1 ARG 60 (A) NH1 ARG 60 (A) O ILE 189 (A) N GLU 191 (A)	O α D-Pla6 O-acetyl L-Thr7 O γ Ac- β -HyLeu1 O Ac- β -HyLeu1	ILE 56 (A) LYS 57 (A) PHE 75 (A) PHE 75 (A) VAL 184 (A)	D-Pla6 D-Pla6 D-Pla6 L-Thr7 N-MeThr(OMe)8	VAL 184 (A) PRO 185 (A) ILE 190 (A) TYR 192 (A)	D-Pla6 N-MeThr(OMe)8 β -HyLeu2 N-MeDha5
25 (YM-15)	NH1 ARG 60 (A) N GLU 191 (A) OE2 GLU 191 (A) OE2 GLU 191 (A)	O α D-Pla6 O Ac- β -HyLeu1 N Ac- β -HyLeu1 O γ Ac- β -HyLeu1	SER 53 (A) ILE 56 (A) LYS 57 (A) TYR 67 (A) PHE 75 (A) PHE 75 (A)	D-Pla6 D-Pla6 D-Pla6 D-Pla6 L-Thr7 D-Pla6	VAL 184 (A) VAL 184 (A) ILE 190 (A) ILE 190 (A) TYR 192 (A) ARG 96 (B)	N-MeThr(OMe)8 D-Pla6 β -HyLeu2 D-Pla6 N-MeDha5 Ac- β -HyLeu1

24 (YM-7)	NE ARG 60 (A) NH2 ARG 60 (A) OH TYR 67 (A) N GLU 191 (A) OE1 GLU 191 (A) OE1 GLU 191 (A)	O α D-Pla6 O N-MeDha5 O-acetyl L-Thr7 O Thr1 N Thr1 O γ Thr1	ILE 56 (A) LYS 57 (A) PHE 75 (A) VAL 184 (A)	D-Pla6 D-Pla6 D-Pla6 D-Pla6	ILE 190 (A) ILE 190 (A) TYR 192 (A)	β -HyLeu2 D-Pla6 N-MeDha5
1	OD2 ASP 71 (A) N GLU 191 (A) OE2 GLU 191 (A) OE2 GLU 191 (A)	O-acetyl Thr7 O Ac- β -HyLeu1 N Ac- β -HyLeu1 O γ Ac- β -HyLeu1	SER 53 (A) ILE 56 (A) LYS 57 (A) ARG 60 (A) TYR 67 (A) ASP 71 (A) PHE 75 (A)	D-Pla6 D-Pla6 D-Pla6 D-Pla6 L-Thr7 L-Thr7	PHE 75 (A) VAL 184 (A) VAL 184 (A) ILE 190 (A) ILE 190 (A) TYR 192 (A) PRO 193 (A)	D-Pla6 N-MeThr(OMe)8 D-Pla6 β -HyLeu2 D-Pla6 N-MeDha5 Ac- β -HyLeu1
1 <i>cis</i> (AC-1)	NH1 ARG 60 (A) NH1 ARG 60 (A) OH TYR 67 (A) OD1 ASP 71 (A) NH2 ARG 96 (B)	O D-Pla6 O β -HyLeu7 O α D-Pla6 OH β -HyLeu7 O γ Prop- β -HyLeu1	SER 53 (A) ILE 56 (A) LYS 57 (A) ASP 71 (A) PHE 75 (A) LEU 78 (A)	D-Pla6 N-MeDha5 D-Pla6 β -HyLeu7 D-Pla6 N-MeThr(OMe)8	VAL 184 (A) THR 186 (A) ILE 190 (A) TYR 192 (A) TYR 192 (A) SER 97 (B)	D-Pla6 D-Pla6 D-Pla6 N-MeDha5 Ala4 Prop- β -HyLeu1
17 (YM-385781)	N GLU 191 (A) O GLU 191 (A)	O Ac- β -HyLeu1 N Ac- β -HyLeu1	ILE 56 (A) ARG 60 (A) PHE 75 (A) PHE 75 (A) VAL 184 (A) THR 187 (A)	D-Pla6 D-Pla6 N-MeThr(OMe)8 L-Thr7 N-MeThr(OMe)8 β -HyLeu2	ILE 190 (A) ILE 190 (A) GLU 191 (A) TYR 192 (A) TYR 192 (A) ARG 96 (B)	β -HyLeu2 D-Pla6 Ac- β -HyLeu1 D-Pla6 Ala4 Ac- β -HyLeu1
2	N GLU 191 (A) OE2 GLU 191 (A)	O Ac- β -HyLeu1 O γ Ac- β -HyLeu1	SER 53 (A) ILE 56 (A) LYS 57 (A) TYR 67 (A) GLY 74 (A) PHE 75 (A) PHE 75 (A)	D-Pla6 D-Pla6 D-Pla6 D-Pla6 L-Thr7 L-Thr7 D-Pla6	VAL 184 (A) THR 187 (A) ILE 190 (A) ILE 190 (A) TYR 192 (A) ARG 96 (B)	D-Pla6 β -HyLeu2 N-MeThr(OMe)8 D-Pla6 N-MeDha5 Ac- β -HyLeu1
2 <i>cis</i> (AC-0)	ARG 96 (B)	OAc Ac- β -HyLeu1	LYS 57 (A) LEU 78 (A) THR 186 (A) TYR 192 (A)	D-Pla6 β -HyLeu7 N-MeThr(OMe)8 D-Pla6		
26 (YM-8)	NH1 ARG 60 (A) NH1 ARG 60 (A) N GLU 191 (A) OE1 GLU 191 (A) NH2 ARG 96 (B)	O α D-Pla6 O-acetyl L-Thr7 O Ac- β -HyLeu1 N Ac- β -HyLeu1 O γ Ac- β -HyLeu1	SER 53 (A) ILE 56 (A) LYS 57 (A) ARG 60 (A) TYR 67 (A) PHE 75 (A)	D-Pla6 D-Pla6 D-Pla6 D-Pla6 D-Pla6 L-Thr7	PHE 75 (A) VAL 184 (A) VAL 184 (A) ILE 190 (A) ILE 190 (A) TYR 192 (A)	D-Pla6 N-MeThr(OMe)8 D-Pla6 Thr2 D-Pla6 N-MeDha5
27 (YM-16)	N GLU 191 (A) OE1 GLU 191 (A) OE1 GLU 191 (A)	O Ac- β -HyLeu1 N Ac- β -HyLeu1 O γ Ac- β -HyLeu1	ILE 56 (A) LYS 57 (A) TYR 67 (A) PHE 75 (A) PHE 75 (A) PHE 75 (A) VAL 184 (A)	D-Pla6 D-Pla6 D-Pla6 N-MeThr(OMe)8 L-Thr7 D-Pla6 N-MeThr(OMe)8	VAL 184 (A) ILE 190 (A) ILE 190 (A) TYR 192 (A) ARG 96 (B) SER 97 (B)	D-Pla6 β -HyLeu2 D-Pla6 N-MeDha5 Ac- β -HyLeu1 Ac- β -HyLeu1
Moderate Fit						
30 (YM-9) [#]	NH1 ARG 60 (A) NH1 ARG 60 (A)	O N-MeDha5 O-acetyl L-Thr7	SER 53 (A) ILE 56 (A) LYS 57 (A) PHE 75 (A) PHE 75 (A)	D-Pla6 N-MeDha5 D-Pla6 L-Thr7 D-Pla6	VAL 184 (A) THR 186 (A) ILE 190 (A) ILE 190 (A) TYR 192 (A)	D-Pla6 D-Pla6 Leu D-Pla6 Ala4
31 (YM-3) [#]	OH TYR 67 (A) OG1 THR 187 (A)	O-acetyl L-Thr7 O L-Ala4	SER 53 (A) ILE 56 (A) LYS 57 (A) TYR 67 (A) PHE 75 (A) LEU 78 (A)	D-Phe6 D-Phe6 D-Phe6 L-Thr7 L-Thr7 β -HyLeu2	VAL 184 (A) PRO 185 (A) PRO 185 (A) THR 186 (A) ILE 190 (A) ILE 120 (B)	D-Phe6 Ac- β -HyLeu1 β -HyLeu2 D-Pla6 D-Phe6 Ala4
29 (YM-385780) [#]	OD1 ASP 118 (B)	O γ Ac- β -HyLeu1	SER 53 (A) ILE 56 (A) LYS 57 (A) TYR 67 (A) PHE 75 (A)	D-Pla6 D-Pla6 D-Pla6 Ala4 D-Pla6	VAL 184 (A) ILE 190 (A) ILE 190 (A) GLU 191 (A) TYR 192 (A)	D-Pla6 L-Thr7 D-Pla6 N-MeThr(OMe)8 D-Pla6
Least Suitable Fit						
3	NH2 ARG 60 (A)	O N-MeDha5	ILE 56 (A) LYS 57 (A) PHE 75 (A) LEU 78 (A) VAL 184 (A) THR 186 (A) THR 187 (A)	D-Pla6 D-Pla6 D-Pla6 Thr7 D-Pla6 β -HyLeu2 N-MeThr(OMe)8	ILE 190 (A) ILE 190 (A) GLU 191 (A) PRO 193 (A) SER 97 (B) ASP 118 (B)	β -HyLeu2 D-Pla6 Ac- β -HyLeu1 Ac- β -HyLeu1 Ac- β -HyLeu1 Ac- β -HyLeu1

3 <i>cis</i>	NE ARG 96 (B) NH1 ARG 96 (B) NH2 ARG 96 (B) OG SER 97 (B)	O D-Pla6 OAc β -HyLeu7 OAc β -HyLeu7 O D-Pla6	ILE 56 (A) LYS 57 (A) VAL 184 (A) PRO 185 (A) ILE 190 (A) GLU 191 (A)	Ac- β -HyLeu1 Ac- β -HyLeu1 Ac- β -HyLeu1 Ac- β -HyLeu1 Ac- β -HyLeu1 D-Pla6	TYR 192 (A) TYR 192 (A) TYR 192 (A)	Ac- β -HyLeu1 D-Pla6 N-MeDha5
32 (YM-4) [#]	NE ARG 60 (A) NH2 ARG 96 (B)	O α D-Pla6 O γ N-MeThr(OMe) ₈	SER 53 (A) ILE 56 (A) LYS 57 (A) PHE 75 (A) VAL 184 (A)	D-Pla6 D-Pla6 D-Pla6 D-Pla6 D-Pla6	ILE 190 (A) TYR 192 (A) PRO 193 (A) GLU 138 (B)	D-Pla6 N-MeDha5 Ac- β -HyLeu1 β -HyLeu2
33 (YM-6) [#]	NH2 ARG 96 (B)	O-acetyl Thr1	ILE 56 (A) LYS 57 (A) LYS 57 (A) ARG 60 (A) TYR 67 (A) THR 186 (A)	D-Pla6 D-Pla6 N-MeDha5 N-MeDha5 Ala4 D-Pla6	ILE 190 (A) ILE 190 (A) GLU 191 (A) TYR 192 (A) ASP 118 (B) ILE 120 (B)	N-MeThr(OMe) ₈ D-Pla6 N-MeThr(OMe) ₈ N-MeDha5 N-MeThr(OMe) ₈ Thr1
34 (YM-17) [#]	NH1 ARG 96 (B)	O-acetyl Ac- β -HyLeu1	ILE 56 (A) LYS 57 (A) ARG 60 (A) PHE 75 (A)	D-Pla6 D-Pla6 D-Pla6 L-Thr7	PHE 75 (A) ILE 190 (A) TYR 192 (A)	D-Pla6 N-MeDha5 D-Pla6
35 (YM-5) [#]	O PRO 185 (A) N THR 187 (A) OG1 THR 187 (A) O ILE 189 (A) OH TYR 192 (A)	N O-MeThr ₈ O β -HyLeu2 O Ac- β -HyLeu1 O γ Ac- β -HyLeu1 O Ala4	ARG 73 (A) GLY 74 (A) PHE 75 (A) ILE 190 (A)	D-Pla6 D-Pla6 N-MeDha5 Ac- β -HyLeu1	ILE 190 (A) TYR 192 (A) TYR 192 (A) ASP 118 (B)	N-MeAla3 β -HyLeu1 N-MeAla3 Ac- β -HyLeu1
36 (YM-2)	N THR 187 (A) OG1 THR 187 (A) NH2 ARG 96 (B)	O Ac- β -HyLeu1 O-acetyl Ac- β -HyLeu1 O γ Ac- β -HyLeu1	LYS 57 (A) VAL 184 (A) THR 186 (A) ILE 190 (A)	D-Pla6 D-Pla6 β -HyLeu2 Ac- β -HyLeu1	ILE 190 (A) ILE 190 (A) ILE 120 (B)	D-Pla6 Dha5 Ac- β -HyLeu1
4	OE1 GLU 191 (A) O ARG 96 (B) NE ARG 96 (B) NH2 ARG 96 (B)	N ζ N-MeDha5 ^[a] N ζ N-MeDha5 ^[a] O γ Ac- β -HyLeu1 O Ala4	ILE 56 (A) LYS 57 (A) ARG 60 (A) TYR 67 (A) PHE 75 (A) THR 186 (A)	D-Pla6 D-Pla6 D-Pla6 L-Thr7 N-MeThr(OMe) ₈ β -HyLeu2	ILE 190 (A) GLU 191 (A) TYR 192 (A) PRO 193 (A) ILE 120 (B) GLU 138 (B)	β -HyLeu2 N-MeDha5 ^[a] D-Pla6 N-MeDha5 ^[a] Ac- β -HyLeu1 Ac- β -HyLeu1
4 <i>cis</i>	NE ARG 60 (A) NH2 ARG 60 (A)	O α N-MeThr(OMe) ₈ O α N-MeThr(OMe) ₈	ARG 60 (A) GLY 74 (A) PHE 75 (A) VAL 184 (A) ILE 190 (A) ILE 190 (A)	Ac- β -HyLeu7 D-Pla6 D-Pla6 Ac- β -HyLeu7 Ac- β -HyLeu7 N-MeAla3	TYR 192 (A) TYR 192 (A)	N-MeThr(OMe) ₈ Ac- β -HyLeu7

^[a]N-MeDha5 was converted into N-methyl-((2-aminoethyl)thio)-2-alanine in this compound.

5.3. Homology modelling of Nav1.4

We used the YASARA molecular modelling program^{230, 237} to create the homology model of the Nav1.4 channel pore module to perform molecular dynamics simulations and energy minimization steps. Further, the crystal structure of Cav1.1 (PDB ID 3JBR²⁰¹), showing an overall sequence identity of 21 %, turned out to serve as suitable template to model the Nav1.4 sodium channel. Loop regions that were hypothesized to rank among the toxin binding epitope were modelled either as a whole or in part, however always with respect to their sequence conservation score (**Figure 21**)²⁰¹ and also according to the UniProt sequence alignment²⁰⁵ of Nav1.4 (UniProt ID P35499) and the template sequence (UniProt ID P07293, **Figure 21**). To properly identify loop-regions and transmembrane segments, we created a structural alignment of the the Cav1.1 template structure and the crystal structure of the bacterial sodium channel NavAB (PDB ID 3RVY²³⁸) employing the MUSTANG algorithm.²³⁹ If necessary, Yasara's build loop command was used to extend loops by insertion of the particular residues into the loop sequences of the template as outlined in **Figure 21**.²⁴⁰ If it was required to shorten loops, the relevant residues were eliminated and the resulting gap was closed. To finalize the loop modelling step, an energy minimization was applied on the remodelled loops in vacuum employing the NOVA²³⁷ force field. During the loop energy minimization, the remaining part of the protein was kept fixed using the LINCS²⁴¹ method. In two final refining preparation steps to perform the docking, the whole homology model was first energy minimized an equilibrated through an unrestrained all-atom molecular dynamics simulation of 2 ns duration in a membrane environment (PEA) and explicit water, and secondly energy minimized again using the AMBER14^{234, 236} force field.

5.4. μ -PIIIA Nav Docking

To perform the docking, we employed the Easy Interface of the HADDOCK web service (<https://haddock.science.uu.nl/services/HADDOCK2.2/haddockserver-easy.html>),^{209, 218}

As required for the individual docking of the μ -PIIIA isomers to the homology model of Nav1.4, we first defined all the specific residues within the channel's surface as active as they are assumed to interact with the toxin (analogously to **Figure 25a**). The lowest energy conformation of the NMR ensemble of the respective μ -PIIIA isomer^{54, 151} was selected for the docking process. The lowest energy conformation of the predicted structural ensemble of the μ -PIIIA isomer 12, 13 and 7 was selected for the docking process. Further, the whole toxin, i.e. all 22 residues were also set as active, as being considered to be equally involved in any interaction. We extracted each best-scoring cluster from the HADDOCK results, i.e. those exhibiting the lowest Z-score respectively, to perform further analyses. For these purposes, a subsequent energy minimization on each best scoring HADDOCK cluster was carried out using the Yasara2²³⁰ force field after the N-terminal glutamate was transformed into γ -pyroglutamate. The resulting Z-scores are listed in **Table 4**.

5.5. Generation of Kv models and chimera

For the creation of our Kv1.6 homology model, we employed the YASARA molecular modelling program YASARA “structure” Vers. 18.3.23.⁴⁴ We built the model of the complete potassium channel, including the naturally associated voltage-sensor domains for the sake of sequence data completeness, necessary for the accurate detection of high amounts of related homologues. The modelling parameters for the complete process were set as follows:

In order to identify a structure template as a first step, three subsequent PSI-BLAST²⁴² iterations were run on the complete target sequence of Kv1.6, yielding 89 possible templates identified. In a second step the Protein Data Bank (PDB) was searched for a match for a position specific scoring matrix (PSSM), that has been extracted from UniRef90²⁴³ in advance. (Hits were detected by an E-value below the cutoff 0.5.) A total of four hits (PDB-IDs 2R9R-B, 3LNM-B, 3LUT-B and 2A79-B) among the 89 structures could be found as possible suitable templates for Kv1.6.

A secondary structure prediction for the target sequence was used for purposes of alignment correction and loop modelling supports. The predicted secondary structure was obtained by again running PSI-BLAST²⁴² to first get a target sequence profile and subsequently feeding the profile to the PSI-Pred secondary structure prediction algorithm.²⁴⁴ A target sequence profile has been created from a multiple sequence alignment in order to aid aligning the target and template sequences. The multiple sequence alignment was built from related sequences from the aforementioned PSI-BLAST-obtained UniRef90 sequences. Alternative alignments of the target and the respective template protein sequences yielded 29 generated models for the four template PDBs in a total. Side chains were added using YASARA's SCWRL3²⁴⁰ implementation and further fine-tuned by taking into consideration solvation effects and electrostatic, knowledge-based packing iterations. In addition, the hydrogen bond network was optimized²⁴⁵ before subjecting each model to an unrestrained energy minimization with explicit water molecules. The energy minimization was performed by simulated annealing employing the Yasara2 force field.²³⁰

All 29 models were ranked according to their overall quality Z-scores and in addition a YASARA-designed hybrid model was constructed by combining the best (scoring) parts of each model. Due to its highest rank among all models, the hybrid model was accepted as the final channel model for Kv1.6.

A more detailed description of the complete YASARA homology modelling protocol can be seen online (<http://yasara.org/homologymodeling.htm>).

On the basis of the obtained Kv1.6 homology model, further model structures were also built for Kv1.1, Kv1.5 and for the chimera channels “Kv1.6-5P1” and “Kv1.6-5P2” by incorporating the respective mutations: individual substitutions of Kv1.6-differing residues in the “P1”-loops in the outer surface area of the Kv1.6 model yielded the structure for Kv1.6-5P1 and, correspondingly, in the “P2” inner pore-loops of the Kv1.6 model yielded our final Kv1.6-5P2 channel structure. Analogously we mutated both, the P1 and the P2 loops in order to obtain Kv1.5 or Kv1.1. The model for Kv1.1 was finished by an additional mutation within the pore helical segment of the Kv1.6 model. This mutation is contained in Kv1.1 compared to Kv1.6 (see **Figure 25a**). For further use, all five models were energy-minimized in a finalizing step.

5.6. μ -PIIIA, μ -SIIIA and μ -GIIIA docking on Kv1 channel subtypes

As the outer located voltage-sensor domains of the modelled Kv1 channels are not considered to be involved in μ -conotoxin binding, they do not account for our focused region of interest and were therefore declared as being negligible for any further investigations. We hence omitted these parts in any further work step, additionally in view of data reduction and time-saving aspects.

In order to start the pursued simulations from an appropriate initial state where both, channel and toxin, are present in a bound form, predicted conformations from docked clusters of the respective systems were used. These were obtained by docking the lowest energy member of the NMR ensemble of μ -PIIIA¹⁵¹ (PDB ID: 1R9I²⁴⁶), μ -SIIIA (BMRB ID: 20023) and μ -GIIIA (PDB ID: 1TCG²⁴⁷) to the potassium channel models of Kv1.1, Kv1.5 and Kv1.6 and to the chimera homology models. The dockings were performed using the Easy Interface of the HADDOCK online platform^{209, 218, 248} (<https://haddock.science.uu.nl/services/HADDOCK2.2/haddockserver-easy.html>), a web service known to be suitable for the handling of more complex peptide ligand structures, apart from small molecule compounds.²⁴⁸

For technical reasons concerning recognition problems of γ -pyroglutamic acid (Pyr) by the HADDOCK program, γ -pyroglutamic acid at toxin position one (Pyr1) of any system containing μ -SIIIA or μ -PIIIA was replaced by glutamic acid (Glu1) before performing the docking routine (and re-converted afterwards for further work steps).

As requested by the HADDOCK user interface and equally due to computing power-saving reasons, “active”, i.e. contact-forming residues within the complex were predefined, distinguished from “inactive” or “passive” residues not being part of the interactive area.

Residues within regions that are part of the channel’s upper surface as well as all toxin residues were defined as “active”. More precisely, the channel’s upper surface region comprises the last three residues of the outer S5 helix, the subsequent outer P1 loop on the channel’s outer surface, the pore helix, connecting the outer P1 and the inner P2 loop of the pore duct followed by the first three residues of the outer S6 helix (see segment shown in **Figure 25a**). Considering the docked μ -conotoxins, all toxin residues were set as active part in the system to ensure equal preconditions i.e. an equal general probability for each toxin residue to be predicted to form interactions with the channel’s surface.

According to the HADDOCK protocol, passive residues can either be limited to residues that surround the active-defined ones or generally that are part of the overall protein’s surface.²¹⁸ In order to avoid data redundancy and to provide further time-saving benefits, we decided to restrict the quantity of passive-set residues by choosing them to be automatically defined around the active-specified ones within the system. This option takes all residues into account that are both, on the surface or within a radius of maximum 6.5 Å of any active-set residue.²¹⁸

From each docking result, the individual best scoring structure from the respective highest scoring cluster (exhibiting the lowest Z-Score within the set of resulting clusters) was respectively selected as starting structure for further analyses (**Table 6**).

In addition, each selected structure was rescored employing the AutoDock Vina algorithm²²⁰ under default settings.²⁴⁹ The corresponding setup was performed using the YASARA molecular modelling software (**Table 6**).⁴⁴

5.7. μ -PIIIA, μ -SIIIA and μ -GIIIA simulations on Kv1 channel subtypes

All further energy minimization and MD simulation steps described below were performed using the YASARA molecular modelling software.⁴⁴

Our previous modifications from γ -pyroglutamic acid to glutamic acid for reasons of meeting the requirements of the HADDOCK web interface,²¹⁸ were retransformed by reconversion of the concerning residues of the chosen docked complexes to γ -pyroglutamic acid for μ -SIIIA and μ -PIIIA. Prior to further MD simulations, the reconversion was followed by a subsequent energy minimization step in water under application of the AMBER14 Force Field^{234, 236} to ensure the proper formation and positioning of the reconverted γ -pyroglutamic acid until system convergence is achieved as soon as the energy improves by less than 0.05 kJ/mol per atom during 200 time steps (<0.05 kJ/mol/200/steps).

MD simulations of the particular selected lowest energy structure from the best-scoring docking clusters were performed under the use of YASARA's incorporated routine for the simulation of membrane proteins in a lipid membrane environment.

The macro firstly scans the protein for secondary structure elements with hydrophobic surface residues, orients it accordingly and embeds it in the membrane of adjustable lipid composition. Finally a 250 ps restrained equilibration simulation is run, which ensures that the membrane can adapt to the newly embedded protein, before the actual simulation starts.²²⁸

The simulations were performed within a cuboid simulation cell, containing the membrane 30 Å larger than the protein and of a height 20 Å higher than the protein (rest of the cell was filled with an aqueous potassium chloride solution of 0.9% concentration). The membrane was chosen to be completely composed of phosphatidylethanolamine (PEA), the most stable lipid membrane,²²⁸ in order to mimic the native membrane environment during our simulations.

The energy minimizations and refinement simulations of the complexes were performed as unrestrained all-atom molecular dynamics simulations in explicit water (TIP3P²⁵⁰) using the PME method²³⁵ to describe long-range electrostatics at a cut-off distance of 8 Å. In addition, physiological conditions (0.9% KCl, pH 7.4²²⁹) were maintained during our simulations as well as a constant temperature of 298 K using a Berendsen thermostat and a constant pressure of 1 bar.

Charged amino acids were assigned according to the predicted pKa of the amino acid side chains by Ewald summation and were neutralized by adding counter ions (KCl).²²⁹

A multiple time step algorithm together with a simulation time step interval of 5 fs was chosen in order to increase the simulation performance.²²⁸ Further, the AMBER14 Force Field^{234, 236} was applied and high frequency bond and angle vibrations of hydrogen atoms were removed employing constraints through the LINCS²⁴¹ and SETTLE²⁵⁰ approach.

Simulation times ranged from 0.5 to 1 μ s, whereby simulation snapshots were saved every 250000 fs, i.e. 250 ps (employing the simulation macro in fast speed mode).

For energy minimization by simulated annealing the Yasara²³⁰ force field was used including optimization of the hydrogen bond network²⁴⁵ and equilibration of the water shell until convergence of the system was achieved.

All molecular graphics were created with YASARA (www.yasara.org) and POVRay (www.povray.org).

5.8. Combined clustering for structure selection

In order to provide a more sophisticated than just an arbitrary snapshot selection for the representative toxin-channel structures, we implemented and applied two clustering approaches and combined them for the final selection of representative snapshots at appropriate simulation times. The particular clustering approaches are described below. Prior to the clustering applications, each data set was subdivided into eleven snapshots respectively, equally spaced at a distance of 50 ns. This preselection defined the starting point for the clustering applications. Correspondingly, snapshot 1 represents the structure at the simulation start time of 0ns, snapshot 2 contains the simulated structure at 50 ns and so on. Finally snapshot 11 represents the simulated structure at 500ns simulation time.

Figure 41 shows the overall selection process as a whole workflow.

We endorsed both of the presented clustering approaches and especially a complementary combination of them, as each of the methods relies on the assumption that an equilibrated state of stable condition can be represented as a detected cluster of maximum size. Consequently, the wanted representative candidate is most probably present within this cluster. As the first method is however more specific than the second one considering its selection criterion, we intersect both methods' results for the further selection process to provide a more sensitive approach. In these respects, clustering method 2 depicts the more sensitive supplementation for the first one, also taking the overall data dispersion as well as the location of the main data focus into account.

The intersection cluster yields a subset out of the 11 pre-selected snapshots, having ensured to contain members satisfying the criteria given in the two clustering approaches. For the selection of the final representative out of the intersection it is thus secured that it designates a snapshot of preferably minimum distance to the average simulation structure and simultaneously being located within an area of high data density. (The preferably low distance to the average structure is even considered until the proposal step for the preliminary representative.)

In these respects, the second clustering method can be considered as completion or refinement revision of the first one. We thus recommend the use of a combination of both approaches in terms to achieve an adequate, i.e. enough sensitive and specific selection process for the detection of representatives.

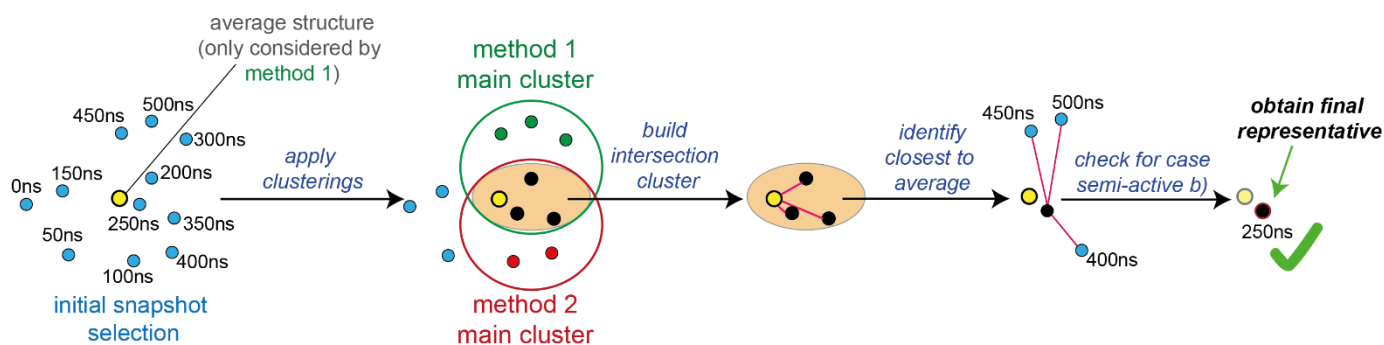


Figure 41: Cluster computation workflow summarizing the combined usage of clustering methods 1 and 2 up to the selection of the final representative snapshot. (Illustration adapted from Kaufmann et al.⁶)

5.8.1. Clustering method 1 – cluster close to the average structure

This method extracts all snapshots of toxin distance $< 3\text{\AA}$ RMSD to the average simulation toxin structure and defines them as members of the main cluster. This approach of “clustering” only yields a single cluster of filtered snapshots as the data subset searched needs to satisfy just one criterion. As the rest of the data is not further considered, the yielded main cluster size directly represents the system’s degrees of stability and dynamics. Employing clustering method 1 to our data yields the particular snapshot clusters as shown in **Table 12** (3rd column). In a worst case scenario like for the highly fluctuating System μ -GIIIA Kv1.6, the algorithm yields an empty set, i.e. a non-existing main cluster (**Table 12**, 8th row, 3rd column).

We will further use the term "MCL1" to denote the main cluster obtained by our first clustering approach.

Table 12: Cluster analysis 1 results. Detection of a main cluster by threshold RMSD [avg. structure] $< 3.0\text{\AA}$ (3rd column); 7th column contains the snapshot exhibiting the lowest toxin RMSD to the average toxin structure among all, i.e. out of the main cluster members.

System	Mode of activity	Main cluster	main cluster size	main cluster weight	Snapshot interval(s) main cluster	Main cluster snapshot of lowest toxin RMSD to avg: RMSD[Å]
μ -PIIIA Kv1.6	active	{3,4,5,6,7,8,9,10}	8	8 / 11	[100ns,450ns]	snapshot 5 (at 200ns): 1.5248
μ -PIIIA Kv1.1	active	{2,3,4,5,6,7,8,9,10,11}	10	10 / 11	[50ns,500ns]	snapshot 10 (at 450ns): 1.7046
μ -PIIIA Kv1.6-5P1	semi-active a)	{2,5,6,9,10,11}	6	6 / 11	[50ns], [200ns, 250ns], [400ns,500ns]	snapshot 6 (at 250ns): 2.5522
μ -SIIIA Kv1.6	semi-active a)	{2,3,4,5,6,7,8,9,10,11}	10	10 / 11	[50ns,500ns]	snapshot 6 (at 250ns): 1.2155
μ -PIIIA Kv1.6-5P2	semi-active b)	{2,3,4,5,6}	5	5 / 11	[50ns,250ns]	snapshot 5 (at 200ns): 2.2403
μ -PIIIA Kv1.5	inactive a)	{4,5,6,7}	4	4 / 10	[150ns,300ns]	snapshot 4 (at 150ns): 2.4267
μ -GIIIA Kv1.6	inactive b)	{}	-	-	-	snapshot 6 (at 250ns): 5.7302 (\rightarrow lowest RMSD among all)

5.8.2. Clustering method 2 – graph-based cluster search

In order to provide some refinement step for our first clustering approach considering the sensitivity of selection, we implemented and applied a further clustering approach as illustrated in **Figure 42**. The algorithm used is explained by means of a graph representation of the snapshot preselection: the eleven nodes (or snapshots) are successively traversed by calculation of the pairwise distance, i.e. their respective toxin RMSD, to the other nodes. In theory, this would require an all-pairwise distance calculation between each pair of nodes, i.e. pair of toxins within the snapshots. The computing effort of such a calculation would result in a polynomial run time of $O(n^2)$ complexity in a worst case scenario, due to the number of edges $= \binom{n}{k} = \frac{n(n-1)}{2}$ in a complete graph (yielding 55 node distances in our case of eleven snapshots).

We try to circumvent this calculation effort by omitting the consideration of already compared and assigned nodes (to a cluster), similar to a decision tree: The nodes are considered chronologically and, if not yet assigned, a node's distances to all remaining nodes is calculated. At this, a node will be assigned to the cluster of the currently considered node (that defines the actual cluster instance) up to a threshold of $< 3 \text{ \AA}$ toxin RMSD to the currently considered node. (The threshold was set to maintain consistencies with the threshold set for clustering method 1.) **Figure 42** illustrates the course of such an iteration. The next node that is has not yet been assigned to any cluster as a member will depict the further node instantiating a new cluster for the subsequent iteration. Consequently, nodes that are already assigned to a cluster, are not considered in any further iteration. In all scenarios better than the worst case, redundant calculations of pairwise distances are avoided (due to the decision tree model). In a best case scenario the number of pair-wise distance calculations is even reduced to $n-1$ (according to $O(n)$) distance calculations, corresponding to a linear effort. (A transitive relation considering a close distance of two pairs of nodes sharing one cluster, but of unknown pairwise distance as they are already assigned, is given by their common close distance to the node of the cluster they were assigned to.)

Dealing with data from molecular dynamics as in our case, a chronological way of data inspection seems to be useful and obvious since the observed data intrinsically holds an even continuity (describing the transitions between the different states.)

Results yielded by our second clustering approach are listed in **Table 13**. The implementation was done using YASARA's Yanaconda macro language (<http://www.yasara.org/yanaconda.htm>).

In contrast to cluster method 1 the second clustering approach takes into account information about the internal data spread, seen by the number of clusters generated (**Table 13**, 4th column **Table 13**), why we wanted to profit from those benefits despite theoretical algorithmic complexities. Information about the location of the data focus and the intensity of the overall data scattering, not necessarily given by method 1, is provided.

As each toxin snapshot is considered and compared to all other ones, instead of to the toxin average only, the main cluster is not detected only by comparison with a theoretically calculated size, but rather with respect to the "real" quality or nature of the inspected data. This approach offers considerable advantages, for instance in the case of precise analyses of data sets of high unequal distribution.

The cluster of largest size or, in case if same maximum sizes occur, the union of the clusters of largest sizes is finally determined as the main cluster out of all generated clusters (**Table 13**, 4th column, yellow highlighted set). A system's stability and dynamics are therefore deducible by the overall size of its calculated main cluster, as well as by the actual number of clusters generated (and obviously again by their correlated respective sizes). In a worst case scenario however, the algorithm yields eleven (or the number of existing snapshot nodes) clusters of respective size one.

We will further use the term "MCL2" to designate main cluster 2, obtained by our second clustering approach.

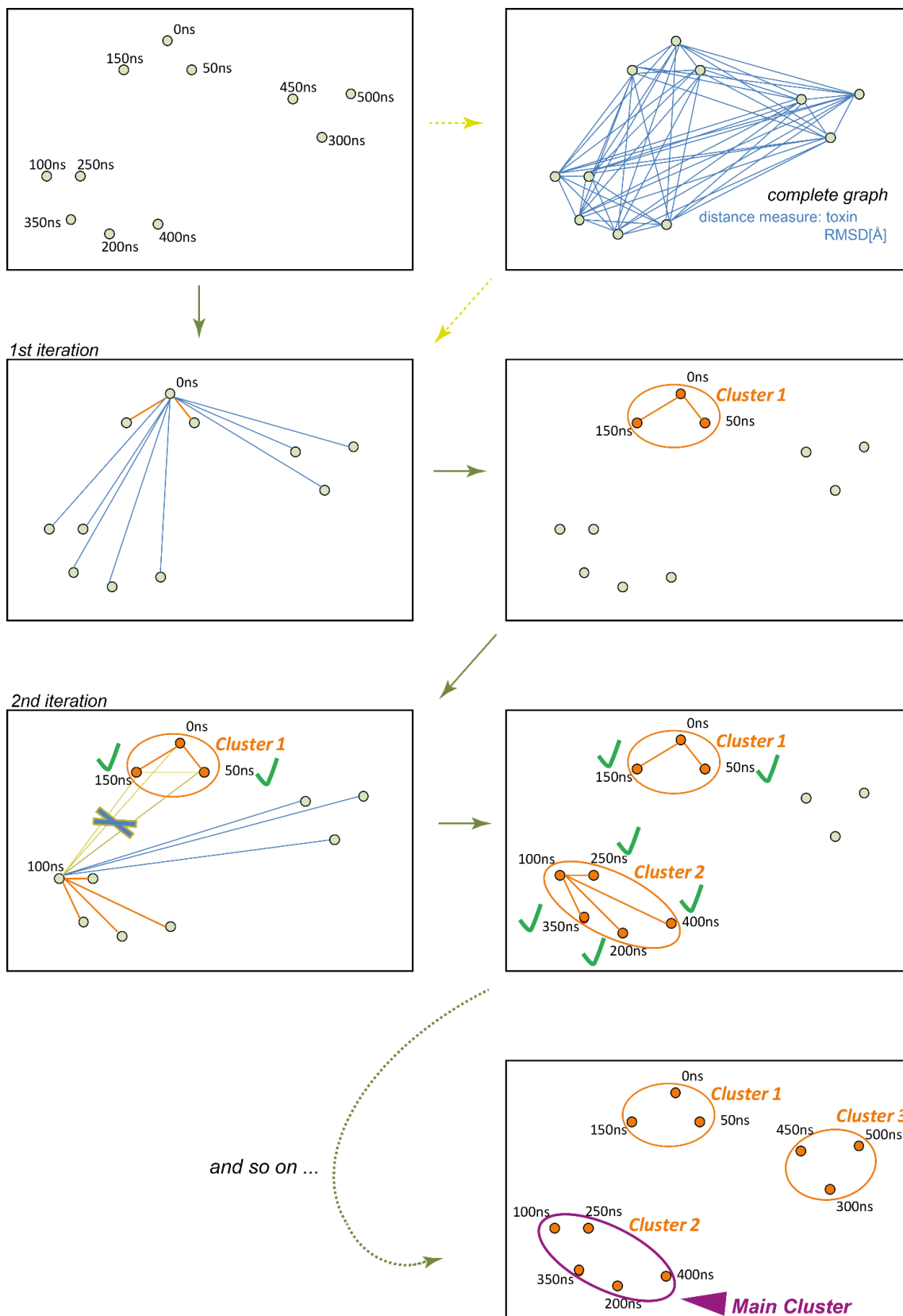


Figure 42: Clustering approach 2 workflow diagram. A graph-based clustering calculation is applied on a preselection of eleven snapshots out of a 500ns simulation; the preselection snapshots are equally spaced at 50ns; each snapshot is represented by a vertex node while the edge lengths designate the pair-wise toxin RMSD values [\AA] of the particular snapshot pair considered; the clustering process is finished when each node is assigned to a cluster; the main cluster(s) is (are) determined by the highest number of members (vertices) or by the union of the clusters of largest sizes (in case if same maximum size clusters occur). (Illustration adapted from Kaufmann et al.⁶)

Table 13: Results yielded by cluster analysis 2. The main cluster(s) out of the preselection of eleven snapshots is (are) respectively determined by the highest cluster size among all clusters detected (highlighted yellow); 8th column contains the snapshot out of the detected main cluster(s) exhibiting the lowest toxin RMSD to the average toxin structure.

System	Mode of activity	Number of detected clusters	Clusters including main cluster(s)	Size of main cluster	Weight of main cluster	Snapshot interval(s) of main cluster(s)	Main cluster snapshot of lowest toxin RMSD to avg: RMSD[Å]
μ -PIIIA Kv1.6	active	4	{1},{2},{3,4,5,6,7,8,9,10},{11}	8	8 / 11	[100ns,450ns]	snapshot 5 (at 200ns): 1.5248
μ -PIIIA Kv1.1	active	6	{1},{2},{3,5},{4},{6,10},{7,8,9,11}	4	4 / 11	[300ns,400ns],[500ns]	snapshot 9 (at 400ns): 1.8812
μ -PIIIA Kv1.6-5P1	semi-active a)	5	{1},{2,3},{4,5,6,7},{8,9,10},{11}	4	4 / 11	[150ns,300ns]	snapshot 6 (at 250ns): 2.5522
μ -SIIIA Kv1.6	semi-active a)	3	{1},{2,3},{4,5,6,7,8,9,10,11}	8	8 / 11	[150ns,500ns]	snapshot 6 (at 250ns): 1.2155
μ -PIIIA Kv1.6-5P2	semi-active b)	9	{1},{2,6},{3},{4,5},{7},{8},{9},{10},{11}	2	2 / 11	[50ns],[250ns],[150ns,200ns]	snapshot 5 (at 200ns): 2.2403
μ -PIIIA Kv1.5	inactive a)	10	{1},{2},{3},{4,7},{5},{6},{8},{9},{10},{11}	2	2 / 11	[150ns],[300ns]	snapshot 4 (at 150ns): 2.4267
μ -GIIIA Kv1.6	inactive b)	10	{1},{2},{3},{4},{5},{6},{7},{8},{9,11},{10}	2	2 / 11	[400ns],[500ns]	snapshot 9 (at 400ns): 6.5590

5.8.3. Selection of the final representative

a) *Generating the final preselection set*

In a third finalizing step regarding cluster formations, we computed the intersection cluster out of both methods' main clusters to generate a final preselection set out of which the final representative should (preferably) be chosen. The intersection was built to make sure to only maintain snapshot members that concomitantly fulfil the following two criteria considering their position within the space of simulation data: first, being located in close proximity to the average simulation structure and second, being located at a position of high data density, if possible.

For cases where the intersection of the two obtained main clusters would result in an empty set, such as for instance in the case of μ -GIIIA on Kv1.6, we took the union cluster to be able to perform further investigations and in order to take at least one resulting main cluster out of the two methods into consideration.

We will further use the term "*PreSel*" to designate the final preselection set resulting from the operations on the two main clusters from method 1 (MCL1) and method 2 (MCL2).

A formal description of building the final preselection set is given by

$$PreSel = \begin{cases} MCL\ 1 \cap MCL\ 2, & \text{if } MCL\ 1 \neq \emptyset \text{ and } MCL\ 2 \neq \emptyset \\ MCL\ 1 \cup MCL\ 2, & \text{else} \end{cases}$$

The resulting preselection sets are listed in **Table 14** (3rd column).

b) Determine the preliminary representative

For further analyses we picked the snapshot out of the preselection set exhibiting the *minimum* RMSD distance to the respective average structure, being denoted as “preliminary” representative snapshot (see **Table 14**, 4th column).

It has to be noted that this preliminary representative is not necessarily identical to the snapshot of the generally smallest distance to the average structure regarding the complete simulation data set, since at least method 2 is able to yield a main cluster *not* containing this snapshot. As the result obtained by clustering method 2 defines a highly weighted prerequisite in the context of generating the preselection set, the aforementioned overall closest to the average snapshot consequently does not necessarily have to be included in the preselection set as well. Such cases may for instance occur when dealing with very one-sided, not evenly distributed data.

In our dataset this case could be observed by the example of μ -PIIIA at Kv1.1 (**Table 12**, **Table 13**, **Table 14**, 3rd row): snapshot 10, which according to method 1 has the lowest RMSD to the average structure was eliminated during the selection process by the intersecting operation and replaced by snapshot 9, yielded by method 2. Actually the updated preliminary representative (snapshot 9) has a higher distance to the average structure than snapshot 10, but interestingly (in contrast to snapshot 10) contained the crucial interaction Gln15 to Gly376/SIV (**Figure 34a**).

This additionally supports the above mentioned facts considering the method’s (clustering method 2) sensitivity, especially considering that the preliminary snapshot is revised again by a further final revision step, which is described in the following.

c) Revision and final selection – check for case “semi-active b)”

As among the systems μ -GIIIA on Kv1.6 and μ -PIIIA on Kv1.6-5P2 a principal detachment towards a channel moiety or a side of a channel’s subunit could be observed, we aimed to trap a snapshot of such already detached state as representative structure for these two instances.

For both cases the detaching movement occurs at different simulation times (for the semi-active μ -PIIIA on Kv1.6-5P2 within around the last simulation quarter, also referred to as mode “*semi-active b)*” as per our own designation and for inactive μ -GIIIA on Kv1.6 within the first simulation quarter, also called “*inactive b)*”, see **Figure 28**, **Table 9**). As the state of detachment is however adopted at later simulation times for semi-active “ μ -PIIIA on Kv1.6-5P2”-like cases, it would not be represented by the majority of simulation data.

We therefore decided to perform a final revising step on any preliminary representative for checking a later detaching behaviour in the last third to quarter of a simulation. For this purpose, we counterchecked the toxin RMSDs between our preliminary representatives and the corresponding toxin structures of the associated last three simulation snapshots (i.e. at simulation times 400ns, 450ns and 500ns). The cutoff was set to 6.0Å. For RMSD values below and up to and including this threshold we retained the preliminary representative and took it as final snapshot. In cases of higher RMSDs the preliminary representative was rejected and substituted by the snapshot out of the last three compared ones which still exhibited the lowest RMSD to the preliminary representative. (If more than one values >6.0Å of identical minimum size have occurred, one of these concerned snapshots would have been randomly selected as representative). For μ -PIIIA on Kv1.6-5P2 our latest checks resulted in a revision and replacement of the proposed preliminary representative at 200ns simulation time by snapshot 10 at 450ns (**Table 14**, 6th row). The “inactive-alike” high fluctuation rate of μ -PIIIA on Kv1.6-5P2, shown by the results of clustering method 2, i.e. the comparatively high number of generated clusters (**Table 13**, 4th column, 6th row) further underpinned our revision steps and the concomitant replacement of the preliminary representative by a snapshot of later simulation time.

All results of our revising step are listed in **Table 14** (6th column).

Last but not least, it should be noted that, in the (unlikely but still possible) worst-case of an empty preselection set, out of the set of all eleven snapshots generated at the beginning, the structure of general lowest RMSD to the average structure would be selected as final representative. This is to maintain consistency with regard to the entire approach.

Table 14: Detection of the final representative snapshot. Combination (intersection or union) of cluster analyses 1 and 2 main clusters is applied (3rd column), followed by a final check for case semi-active (b) by reconciliation with the snapshots of the last simulation quarter (5th column); final representatives are listed in the 6th column.

System	Mode of activity	preselection (intersection or union)	preliminary representative	Revision: Toxin RMSD of the preliminary representative and the last three simulation snapshots (at 400ns, 450ns and 500ns) [Å] (threshold: 6.0Å as check for case semi-active (b))	Final representative(s):
μ-PIIIA Kv1.6	<i>active</i>	{3,4,5,6,7,8,9,10}	snapshot 5 (at 200ns) 1.5248	to snapshot 9 (at 400ns): 2.1996 to snapshot 10 (at 450ns): 2.3839 to snapshot 11 (at 500ns): 4.0344	snapshot 5 (at 200ns)
μ-PIIIA Kv1.1	<i>active</i>	{7,8,9,11}	snapshot 9 (at 400ns) 1.8812	to snapshot 9 (at 400ns): 2.0698 to snapshot 10 (at 450ns): 0.0000 to snapshot 11 (at 500ns): 2.5402	snapshot 9 (at 400ns)
μ-PIIIA Kv1.6-5P1	<i>semi-active a)</i>	{5,6}	snapshot 6 (at 250ns) 2.5522	to snapshot 9 (at 400ns): 4.8724 to snapshot 10 (at 450ns): 4.3513 to snapshot 11 (at 500ns): 5.0838	snapshot 6 (at 250ns)
μ-SIIIA Kv1.6	<i>semi-active a)</i>	{4,5,6,7,8,9,10,11}	snapshot 6 (at 250ns) 1.2155	to snapshot 9 (at 400ns): 1.6569 to snapshot 10 (at 450ns): 2.4387 to snapshot 11 (at 500ns): 1.9899	snapshot 6 (at 250ns)
μ-PIIIA Kv1.6-5P2	<i>semi-active b)</i>	{2,4,5,6}	snapshot 5 (at 200ns) 2.2403	to snapshot 9 (at 400ns): 5.0467 to snapshot 10 (at 450ns): 7.7670 to snapshot 11 (at 500ns): 8.5439	snapshot 10 (at 450ns)
μ-PIIIA Kv1.5	<i>inactive a)</i>	{4,7}	snapshot 4 (at 150ns) 2.4267	to snapshot 9 (at 400ns): 5.0881 to snapshot 10 (at 450ns): 5.5301 to snapshot 11 (at 500ns): 4.3945	snapshot 4 (at 150ns)
μ-GIIIA Kv1.6	<i>inactive b)</i>	{9,11}	snapshot 9 (at 400ns) 6.5590	to snapshot 10 (at 450ns): 4.2314 to snapshot 11 (at 500ns): 2.7594	snapshot 9 (at 400ns)

6. List of publications

Per review articles:

accepted for publication:

1. Kaufmann, D.; Tietze, A. A.; Tietze, D., **In Silico Analysis of the Subtype Selective Blockage of KCNA Ion Channels through the micro-Conotoxins PIIIA, SIIIA, and GIPIA**. *Mar Drugs* **2019**, *17* (3)
2. Reher, R.; Kuhl, T.; Annala, S.; Benkel, T.; Kaufmann, D.; Nubbemeyer, B.; Odhiambo, J. P.; Heimer, P.; Bauml, C. A.; Kehraus, S.; Crusemann, M.; Kostenis, E.; Tietze, D.*; König, G. M.*; Imhof, D.*, **Deciphering Specificity Determinants for FR900359-Derived Gq alpha Inhibitors Based on Computational and Structure-Activity Studies**. *ChemMedChem* **2018**, *13* (16), 1634-1643

in preparation:

3. D. Tietze,^{#*} D. Kaufmann,[#] A. A. Tietze, A. Voll, R. Reher,⁴ G. König,⁴ F. Hausch³, **Structural and dynamical basis of G protein inhibition by YM-254890 and FR900359: an inhibitor in action**, in preparation for *Journal of Medicinal Chemistry*
4. A. A. Paul George[#], P. Heimer[#], E. Leipold, T. Schmitz, D. Kaufmann, D. Tietze, A. Biswas, S. H. Heinemann, and D. Imhof*, **Linking conformational diversity with bioactivity of μ -conotoxin PIIIA disulfide isomers**, in preparation for *Marine Drugs*
([#] equal contribution)

7. List of references

1. Nishimura, A.; Kitano, K.; Takasaki, J.; Taniguchi, M.; Mizuno, N.; Tago, K.; Hakoshima, T.; Itoh, H., Structural basis for the specific inhibition of heterotrimeric G(q) protein by a small molecule. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (31).
2. Taniguchi, M.; Suzumura, K.; Nagai, K.; Kawasaki, T.; Saito, T.; Takasaki, J.; Suzuki, K.; Fujita, S.; Tsukamoto, S., Structure of YM-254890, a novel G(q/11) inhibitor from *Chromobacterium* sp QS3666. *Tetrahedron* **2003**, *59* (25), 4533-4538.
3. Schrage, R.; Schmitz, A.-L.; Gaffal, E.; Annala, S.; Kehraus, S.; Wenzel, D.; Buellesbach, K. M.; Bald, T.; Inoue, A.; Shinjo, Y.; Galandrin, S.; Shridhar, N.; Hesse, M.; Grundmann, M.; Merten, N.; Charpentier, T. H.; Martz, M.; Butcher, A. J.; Slodczyk, T.; Armando, S.; Effern, M.; Namkung, Y.; Jenkins, L.; Horn, V.; Stoessel, A.; Dargatz, H.; Tietze, D.; Imhof, D.; Gales, C.; Drewke, C.; Mueller, C. E.; Hoelzel, M.; Milligan, G.; Tobin, A. B.; Gomeza, J.; Dohlman, H. G.; Sondek, J.; Harden, T. K.; Bouvier, M.; Laporte, S. A.; Aoki, J.; Fleischmann, B. K.; Mohr, K.; Koenig, G. M.; Tuetting, T.; Kostenis, E., The experimental power of FR900359 to study Gq-regulated biological processes. *Nature Communications* **2015**, *6*.
4. Annala, S.; Feng, X.; Shridhar, N.; Eryilmaz, F.; Patt, J.; Yang, J.; Pfeil, E. M.; Cervantes-Villagrana, R. D.; Inoue, A.; Häberlein, F.; Slodczyk, T.; Reher, R.; Kehraus, S.; Monteleone, S.; Schrage, R.; Heycke, N.; Rick, U.; Engel, S.; Pfeifer, A.; Kolb, P.; König, G.; Bünemann, M.; Tüting, T.; Vázquez-Prado, J.; Gutkind, J. S.; Gaffal, E.; Kostenis, E., Direct targeting of G α and G α 11 oncoproteins in cancer cells. *Sci Signal* **2019**, *12* (573), eaau5948.
5. Reher, R.; Kuhl, T.; Annala, S.; Benkel, T.; Kaufmann, D.; Nubbemeyer, B.; Odhiambo, J. P.; Heimer, P.; Bauml, C. A.; Kehraus, S.; Crusemann, M.; Kostenis, E.; Tietze, D.; König, G. M.; Imhof, D., Deciphering Specificity Determinants for FR900359-Derived Gq alpha Inhibitors Based on Computational and Structure-Activity Studies. *ChemMedChem* **2018**, *13* (16), 1634-1643.
6. Kaufmann, D.; Tietze, A. A.; Tietze, D., In Silico Analysis of the Subtype Selective Blockage of KCNA Ion Channels through the micro-Conotoxins PIIIA, SIIIA, and GIIIA. *Mar Drugs* **2019**, *17* (3).
7. Leipold, E.; Ullrich, F.; Thiele, M.; Tietze, A. A.; Terlau, H.; Imhof, D.; Heinemann, S. H., Subtype-specific block of voltage-gated K⁺ channels by mu-conopeptides. *Biochem. Biophys. Res. Commun.* **2017**, *482* (4), 1135-1140.
8. Moro, S.; Braiuca, P.; Deflorian, F.; Ferrari, C.; Pastorin, G.; Cacciari, B.; Baraldi, P. G.; Varani, K.; Borea, P. A.; Spalluto, G., Combined target-based and ligand-based drug design approach as a tool to define a novel 3D-pharmacophore model of human A3 adenosine receptor antagonists: pyrazolo[4,3-e]1,2,4-triazolo[1,5-c]pyrimidine derivatives as a key study. *J. Med. Chem.* **2005**, *48* (1), 152-62.
9. Bohacek, R. S.; McMartin, C.; Guida, W. C., The art and practice of structure-based drug design: a molecular modeling perspective. *Med. Res. Rev.* **1996**, *16* (1), 3-50.
10. Renom-Carrasco, M.; Lefort, L., Ligand libraries for high throughput screening of homogeneous catalysts. *Chem. Soc. Rev.* **2018**, *47* (13), 5038-5060.
11. Lionta, E.; Spyrou, G.; Vassilatis, D. K.; Cournia, Z., Structure-based virtual screening for drug discovery: principles, applications and recent advances. *Curr. Top. Med. Chem.* **2014**, *14* (16), 1923-38.
12. Zoete, V.; Grosdidier, A.; Michielin, O., Docking, virtual high throughput screening and in silico fragment-based drug design. *J. Cell. Mol. Med.* **2009**, *13* (2), 238-48.
13. Rognan, D., Chemogenomic approaches to rational drug design. *Br. J. Pharmacol.* **2007**, *152* (1), 38-52.
14. Trosset, J. Y.; Scheraga, H. A., Reaching the global minimum in docking simulations: a Monte Carlo energy minimization approach using Bezier splines. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95* (14), 8011-5.
15. W.F. van Gunsteren, P. K. W., A.J. Wilkinson, *Computer Simulation of Biomolecular Systems: Theoretical and Experimental Applications*. 1997.
16. Ji, Z. L.; Chen, X.; Zhen, C. J.; Yao, L. X.; Han, L. Y.; Yeo, W. K.; Chung, P. C.; Puy, H. S.; Tay, Y. T.; Muhammad, A.; Chen, Y. Z., KDBI: Kinetic Data of Bio-molecular Interactions database. *Nucleic Acids Res.* **2003**, *31* (1), 255-7.
17. Good, A. C., Novel DOCK clique driven 3D similarity database search tools for molecule shape matching and beyond: adding flexibility to the search for ligand kin. *J. Mol. Graph. Model.* **2007**, *26* (3), 656-66.
18. Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J., AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **2009**, *30* (16), 2785-91.
19. Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R., Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267* (3), 727-48.

20. Rarey, M.; Kramer, B.; Lengauer, T., The particle concept: placing discrete water molecules during protein-ligand docking predictions. *Proteins* **1999**, *34* (1), 17-28.
21. Bursulaya, B. D.; Totrov, M.; Abagyan, R.; Brooks, C. L., 3rd, Comparative study of several algorithms for flexible ligand docking. *J Comput Aided Mol Des* **2003**, *17* (11), 755-63.
22. Siavoush Dastmalchi, M. H.-M., Babak Sokouti, *Methods and Algorithms for Molecular Docking-Based Drug Design and Discovery*. 2016.
23. by W. Eric Wong, T. Z., *Computer Engineering and Networking: Proceedings of the 2013 International Conference on Computer Engineering and Network (CENet2013)*. 2014.
24. Morando, M. A.; Saladino, G.; D'Amelio, N.; Pucheta-Martinez, E.; Lovera, S.; Lelli, M.; Lopez-Mendez, B.; Marenchino, M.; Campos-Olivas, R.; Gervasio, F. L., Conformational Selection and Induced Fit Mechanisms in the Binding of an Anticancer Drug to the c-Src Kinase. *Sci Rep* **2016**, *6*, 24439.
25. Huang, S. Y.; Zou, X., Ensemble docking of multiple protein structures: considering protein structural variations in molecular docking. *Proteins* **2007**, *66* (2), 399-421.
26. Amaro, R. E.; Baudry, J.; Chodera, J.; Demir, O.; McCammon, J. A.; Miao, Y.; Smith, J. C., Ensemble Docking in Drug Discovery. *Biophys. J.* **2018**, *114* (10), 2271-2278.
27. Peter Pulay, G. F., Frank Pang and James E. Boggs, Systematic ab Initio Gradient Calculation of Molecular Geometries, Force Constants, and Dipole Moment Derivatives *J. Am. Chem. Soc.* **1979**
28. Wendy D. Cornell, P. C., Christopher I. Bayly, Ian R. Gould, Kenneth M. Merz, David M. Ferguson, David C. Spellmeyer, Thomas Fox, James W. Caldwell, and Peter A. Kollman, A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**.
29. Gibbs, J. W., *Elementary Principles in Statistical Mechanics*. 1902.
30. Lesar, R., *Introduction to Computational Materials Science: Fundamentals to Applications*. 2013.
31. Skjaerven, L.; Martinez, A.; Reuter, N., Principal component and normal mode analysis of proteins; a quantitative comparison using the GroEL subunit. *Proteins* **2011**, *79* (1), 232-43.
32. Swapnil Mahajan, Y.-H. S., On the relationship between low-frequency normal modes and the large-scale conformational changes of proteins. *Arch. Biochem. Biophys.* **2015**.
33. Blevins, R. D., *Formulas for Natural Frequency and Mode Shape*. 1995.
34. Friesner, R. A., Ab initio quantum chemistry: methodology and applications. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (19), 6648-53.
35. Leach, A. R., *Molecular Modelling: Principles and Applications*. 2001.
36. Park, S.; Schulten, K., Calculating potentials of mean force from steered molecular dynamics simulations. *J. Chem. Phys.* **2004**, *120* (13), 5946-61.
37. Earl, D. J.; Deem, M. W., Parallel tempering: theory, applications, and new perspectives. *Phys. Chem. Chem. Phys.* **2005**, *7* (23), 3910-6.
38. Zhou, R., Replica exchange molecular dynamics method for protein folding simulation. *Methods Mol. Biol.* **2007**, *350*, 205-23.
39. Pronk, S.; Pall, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E., GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **2013**, *29* (7), 845-54.
40. Plimpton, S., Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* **1995**.
41. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K., Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26* (16), 1781-802.
42. Marsili, S.; Signorini, G. F.; Chelli, R.; Marchi, M.; Procacci, P., ORAC: a molecular dynamics simulation program to explore free energy surfaces in biomolecular systems at the atomistic level. *J. Comput. Chem.* **2010**, *31* (5), 1106-16.
43. Beard, H.; Cholleti, A.; Pearlman, D.; Sherman, W.; Loving, K. A., Applying physics-based scoring to calculate free energies of binding for single amino acid mutations in protein-protein complexes. *PLoS One* **2013**, *8* (12), e82849.
44. Krieger, E.; Vriend, G., YASARA View - molecular graphics for all devices - from smartphones to workstations. *Bioinformatics* **2014**, *30* (20), 2981-2.
45. Chien, J. Y.; Friedrich, S.; Heathman, M. A.; de Alwis, D. P.; Sinha, V., Pharmacokinetics/Pharmacodynamics and the stages of drug development: role of modeling and simulation. *AAPS J* **2005**, *7* (3), E544-59.
46. Hauser, A. S.; Attwood, M. M.; Rask-Andersen, M.; Schiöth, H. B.; Gloriam, D. E., Trends in GPCR drug discovery: new agents, targets and indications. *Nat. Rev. Drug Discovery* **2017**, *16*, 829.

47. Holler, C.; Freissmuth, M.; Nanoff, C., G proteins as drug targets. *Cell. Mol. Life Sci.* **1999**, *55* (2), 257-70.
48. Shukla, A. K., Biasing GPCR signaling from inside. *Sci Signal* **2014**, *7* (310), pe3.
49. West, R. E., Jr.; Moss, J.; Vaughan, M.; Liu, T.; Liu, T. Y., Pertussis toxin-catalyzed ADP-ribosylation of transducin. Cysteine 347 is the ADP-ribose acceptor site. *J. Biol. Chem.* **1985**, *260* (27), 14428-30.
50. Gill, D. M.; Meren, R., ADP-ribosylation of membrane proteins catalyzed by cholera toxin: basis of the activation of adenylate cyclase. *Proc. Natl. Acad. Sci. U. S. A.* **1978**, *75* (7), 3050-4.
51. Schmitz, A.-L.; Schrage, R.; Gaffal, E.; Charpentier, T. H.; Wiest, J.; Hiltensperger, G.; Morschel, J.; Hennen, S.; Haeussler, D.; Horn, V.; Wenzel, D.; Grundmann, M.; Buellesbach, K. M.; Schroeder, R.; Brewitz, H. H.; Schmidt, J.; Gomeza, J.; Gales, C.; Fleischmann, B. K.; Tueting, T.; Imhof, D.; Tietze, D.; Guetschow, M.; Holzgrabe, U.; Sondek, J.; Harden, T. K.; Mohr, K.; Kostenis, E., A Cell-Permeable Inhibitor to Trap G alpha(q) Proteins in the Empty Pocket Conformation. *Chem. Biol.* **2014**, *21* (7), 890-902.
52. Vetter, I.; Deuis, J. R.; Mueller, A.; Israel, M. R.; Starobova, H.; Zhang, A.; Rash, L. D.; Mobli, M., Nav1.7 as a pain target - From gene to pharmacology. *Pharmacol. Ther.* **2017**, *172*, 73-100.
53. Murray, J. K.; Wu, B.; Tegley, C. M.; Nixey, T. E.; Falsey, J. R.; Herberich, B.; Yin, L.; Sham, K.; Long, J.; Aral, J.; Cheng, Y.; Netirojjanakul, C.; Doherty, L.; Glaus, C.; Ikotun, T.; Li, H.; Tran, L.; Soto, M.; Salimi-Moosavi, H.; Ligutti, J.; Amagasu, S.; Andrews, K. L.; Be, X.; Lin, M.-H. J.; Foti, R. S.; Ilch, C. P.; Youngblood, B.; Kornecook, T. J.; Karow, M.; Walker, K. W.; Moyer, B. D.; Biswas, K.; Miranda, L. P., Engineering Nav1.7 Inhibitory JzTx-V Peptides with a Potency and Basicity Profile Suitable for Antibody Conjugation To Enhance Pharmacokinetics. *ACS Chem. Biol.* **2019**, *14* (4), 806-818.
54. Heimer, P.; Tietze, A. A.; Bauml, C. A.; Resemann, A.; Mayer, F. J.; Suckau, D.; Ohlenschlager, O.; Tietze, D.; Imhof, D., Conformational mu-Conotoxin PIIIA Isomers Revisited: Impact of Cysteine Pairing on Disulfide-Bond Assignment and Structure Elucidation. *Anal. Chem.* **2018**, *90* (5), 3321-3327.
55. Norton, R. S., Mu-conotoxins as leads in the development of new analgesics. *Molecules* **2010**, *15* (4), 2825-44.
56. Khoo, K. K.; Feng, Z. P.; Smith, B. J.; Zhang, M. M.; Yoshikami, D.; Olivera, B. M.; Bulaj, G.; Norton, R. S., Structure of the analgesic mu-conotoxin KIIIA and effects on the structure and function of disulfide deletion. *Biochemistry* **2009**, *48* (6), 1210-9.
57. Leipold, E.; Markgraf, R.; Miloslavina, A.; Kijas, M.; Schirmeyer, J.; Imhof, D.; Heinemann, S. H., Molecular determinants for the subtype specificity of mu-conotoxin SIIIA targeting neuronal voltage-gated sodium channels. *Neuropharmacology* **2011**, *61* (1-2), 105-11.
58. Yao, S.; Zhang, M. M.; Yoshikami, D.; Azam, L.; Olivera, B. M.; Bulaj, G.; Norton, R. S., Structure, dynamics, and selectivity of the sodium channel blocker mu-conotoxin SIIIA. *Biochemistry* **2008**, *47* (41), 10940-9.
59. Bulaj, G.; West, P. J.; Garrett, J. E.; Marsh, M.; Zhang, M. M.; Norton, R. S.; Smith, B. J.; Yoshikami, D.; Olivera, B. M., Novel conotoxins from *Conus striatus* and *Conus kinoshitai* selectively block TTX-resistant sodium channels. *Biochemistry* **2005**, *44* (19), 7259-7265.
60. Long, S. B.; Tao, X.; Campbell, E. B.; MacKinnon, R., Atomic structure of a voltage-dependent K⁺ channel in a lipid membrane-like environment. *Nature* **2007**, *450* (7168), 376-82.
61. Zhang, D.; Zhao, Q.; Wu, B., Structural Studies of G Protein-Coupled Receptors. *Mol. Cells* **2015**, *38* (10), 836-42.
62. Rosenbaum, D. M.; Rasmussen, S. G.; Kobilka, B. K., The structure and function of G-protein-coupled receptors. *Nature* **2009**, *459* (7245), 356-63.
63. Vroling, B., *Data integration in the life sciences - A protein family-based approach*. 2012.
64. Vassilatis, D. K.; Hohmann, J. G.; Zeng, H.; Li, F.; Ranchalis, J. E.; Mortrud, M. T.; Brown, A.; Rodriguez, S. S.; Weller, J. R.; Wright, A. C.; Bergmann, J. E.; Gaitanaris, G. A., The G protein-coupled receptor repertoires of human and mouse. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (8), 4903-8.
65. Heifetz, A.; Schertler, G. F.; Seifert, R.; Tate, C. G.; Sexton, P. M.; Gurevich, V. V.; Fourmy, D.; Cherezov, V.; Marshall, F. H.; Storer, R. I.; Moraes, I.; Tikhonova, I. G.; Tautermann, C. S.; Hunt, P.; Ceska, T.; Hodgson, S.; Bodkin, M. J.; Singh, S.; Law, R. J.; Biggin, P. C., GPCR structure, function, drug discovery and crystallography: report from Academia-Industry International Conference (UK Royal Society) Chicheley Hall, 1-2 September 2014. *Naunyn Schmiedebergs Arch Pharmacol* **2015**, *388* (8), 883-903.
66. Fredriksson, R.; Lagerstrom, M. C.; Lundin, L. G.; Schioth, H. B., The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol. Pharmacol.* **2003**, *63* (6), 1256-72.
67. Krishnan, A.; Almen, M. S.; Fredriksson, R.; Schioth, H. B., The origin of GPCRs: identification of mammalian like Rhodopsin, Adhesion, Glutamate and Frizzled GPCRs in fungi. *PLoS One* **2012**, *7* (1), e29817.

68. Clemmensen, C.; Smajilovic, S.; Wellendorph, P.; Brauner-Osborne, H., The GPCR, class C, group 6, subtype A (GPCR6A) receptor: from cloning to physiological function. *Br. J. Pharmacol.* **2014**, *171* (5), 1129-41.
69. Attwood, T. K.; Findlay, J. B., Fingerprinting G-protein-coupled receptors. *Protein Eng.* **1994**, *7* (2), 195-203.
70. Kolakowski, L. F., Jr., GCRDb: a G-protein-coupled receptor database. *Receptors Channels* **1994**, *2* (1), 1-7.
71. Foord, S. M.; Bonner, T. I.; Neubig, R. R.; Rosser, E. M.; Pin, J. P.; Davenport, A. P.; Spedding, M.; Harmar, A. J., International Union of Pharmacology. XLVI. G protein-coupled receptor list. *Pharmacol. Rev.* **2005**, *57* (2), 279-88.
72. Harmar, A. J.; Hills, R. A.; Rosser, E. M.; Jones, M.; Buneman, O. P.; Dunbar, D. R.; Greenhill, S. D.; Hale, V. A.; Sharman, J. L.; Bonner, T. I.; Catterall, W. A.; Davenport, A. P.; Delagrangre, P.; Dollery, C. T.; Foord, S. M.; Gutman, G. A.; Laudet, V.; Neubig, R. R.; Ohlstein, E. H.; Olsen, R. W.; Peters, J.; Pin, J. P.; Ruffolo, R. R.; Searls, D. B.; Wright, M. W.; Spedding, M., IUPHAR-DB: the IUPHAR database of G protein-coupled receptors and ion channels. *Nucleic Acids Res.* **2009**, *37* (Database issue), D680-5.
73. Bjarnadottir, T. K.; Gloriam, D. E.; Hellstrand, S. H.; Kristiansson, H.; Fredriksson, R.; Schioth, H. B., Comprehensive repertoire and phylogenetic analysis of the G protein-coupled receptors in human and mouse. *Genomics* **2006**, *88* (3), 263-73.
74. Nordstrom, K. J.; Sallman Almen, M.; Edstam, M. M.; Fredriksson, R.; Schioth, H. B., Independent HHsearch, Needleman-Wunsch-based, and motif analyses reveal the overall hierarchy for most of the G protein-coupled receptor families. *Mol. Biol. Evol.* **2011**, *28* (9), 2471-80.
75. Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C. A.; Motoshima, H.; Fox, B. A.; Le Trong, I.; Teller, D. C.; Okada, T.; Stenkamp, R. E.; Yamamoto, M.; Miyano, M., Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **2000**, *289* (5480), 739-45.
76. Billups, D.; Billups, B.; Challiss, R. A.; Nahorski, S. R., Modulation of Gq-protein-coupled inositol trisphosphate and Ca²⁺ signaling by the membrane potential. *J. Neurosci.* **2006**, *26* (39), 9983-95.
77. Lobanova, E. S.; Finkelstein, S.; Herrmann, R.; Chen, Y. M.; Kessler, C.; Michaud, N. A.; Trieu, L. H.; Strissel, K. J.; Burns, M. E.; Arshavsky, V. Y., Transducin gamma-subunit sets expression levels of alpha- and beta-subunits and is crucial for rod viability. *J. Neurosci.* **2008**, *28* (13), 3510-20.
78. Wall, M. A.; Coleman, D. E.; Lee, E.; Iniguez-Lluhi, J. A.; Posner, B. A.; Gilman, A. G.; Sprang, S. R., The structure of the G protein heterotrimer Gi alpha 1 beta 1 gamma 2. *Cell* **1995**, *83* (6), 1047-58.
79. Wall, M. A.; Posner, B. A.; Sprang, S. R., Structural basis of activity and subunit recognition in G protein heterotrimers. *Structure* **1998**, *6* (9), 1169-83.
80. Liu, W.; Northup, J. K., The helical domain of a G protein alpha subunit is a regulator of its effector. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95* (22), 12878-83.
81. Shnerb, T.; Lin, N.; Shurki, A., What is the role of the helical domain of Galpha in the GTPase reaction? *Biochemistry* **2007**, *46* (38), 10875-85.
82. Dohlman, H. G.; Jones, J. C., Signal Activation and Inactivation by the G Helical Domain: A Long-Neglected Partner in G Protein Signaling. *Sci Signal* **2012**, *5* (226).
83. Koehl, A.; Hu, H.; Maeda, S.; Zhang, Y.; Qu, Q.; Paggi, J. M.; Latorraca, N. R.; Hilger, D.; Dawson, R.; Matile, H.; Schertler, G. F. X.; Granier, S.; Weis, W. I.; Dror, R. O.; Manglik, A.; Skiniotis, G.; Kobilka, B. K., Structure of the micro-opioid receptor-Gi protein complex. *Nature* **2018**, *558* (7711), 547-552.
84. Huang, X.; Miller, W.; Schwartz, S.; Hardison, R. C., Parallelization of a local similarity algorithm. *Comput Appl Biosci* **1992**, *8* (2), 155-65.
85. Overington, J. P.; Al-Lazikani, B.; Hopkins, A. L., How many drug targets are there? *Nat Rev Drug Discov* **2006**, *5* (12), 993-996.
86. Rask-Andersen, M.; Almen, M. S.; Schioth, H. B., Trends in the exploitation of novel drug targets. *Nat Rev Drug Discov* **2011**, *10* (8), 579-90.
87. Belmonte, S. L.; Blaxall, B. C., Conducting the G-protein Coupled Receptor (GPCR) Signaling Symphony in Cardiovascular Diseases: New Therapeutic Approaches. *Drug Discov Today Dis Models* **2012**, *9* (3), e85-e90.
88. Gurbel, P. A.; Kuliopulos, A.; Tantry, U. S., G-protein-coupled receptors signaling pathways in new antiplatelet drug development. *Arterioscler. Thromb. Vasc. Biol.* **2015**, *35* (3), 500-12.
89. Dorsam, R. T.; Gutkind, J. S., G-protein-coupled receptors and cancer. *Nat. Rev. Cancer* **2007**, *7* (2), 79-94.
90. Druey, K. M., Regulation of G-protein-coupled signaling pathways in allergic inflammation. *Immunol. Res.* **2009**, *43* (1-3), 62-76.

91. Deshpande, D. A.; Penn, R. B., Targeting G protein-coupled receptor signaling in asthma. *Cell. Signal.* **2006**, *18* (12), 2105-20.
92. Stone, L. S.; Molliver, D. C., In search of analgesia: emerging roles of GPCRs in pain. *Mol Interv* **2009**, *9* (5), 234-51.
93. Takefuji, M.; Wirth, A.; Lukasova, M.; Takefuji, S.; Boettger, T.; Braun, T.; Althoff, T.; Offermanns, S.; Wettschureck, N., G(13)-mediated signaling pathway is required for pressure overload-induced cardiac remodeling and heart failure. *Circulation* **2012**, *126* (16), 1972-82.
94. Blad, C. C.; Tang, C.; Offermanns, S., G protein-coupled receptors for energy metabolites as new therapeutic targets. *Nat Rev Drug Discov* **2012**, *11* (8), 603-19.
95. Itoh, Y.; Kawamata, Y.; Harada, M.; Kobayashi, M.; Fujii, R.; Fukusumi, S.; Ogi, K.; Hosoya, M.; Tanaka, Y.; Uejima, H.; Tanaka, H.; Maruyama, M.; Satoh, R.; Okubo, S.; Kizawa, H.; Komatsu, H.; Matsumura, F.; Noguchi, Y.; Shinohara, T.; Hinuma, S.; Fujisawa, Y.; Fujino, M., Free fatty acids regulate insulin secretion from pancreatic beta cells through GPR40. *Nature* **2003**, *422* (6928), 173-6.
96. Denis, C.; Sauliere, A.; Galandrin, S.; Senard, J. M.; Gales, C., Probing heterotrimeric G protein activation: applications to biased ligands. *Curr. Pharm. Des.* **2012**, *18* (2), 128-44.
97. Vallar, L.; Spada, A.; Giannattasio, G., Altered Gs and adenylate cyclase activity in human GH-secreting pituitary adenomas. *Nature* **1987**, *330* (6148), 566-8.
98. Smrcka, A. V., Molecular targeting of Galpha and Gbetagamma subunits: a potential approach for cancer therapeutics. *Trends Pharmacol. Sci.* **2013**, *34* (5), 290-8.
99. Yanamadala, V.; Negoro, H.; Denker, B. M., Heterotrimeric G proteins and apoptosis: intersecting signaling pathways leading to context dependent phenotypes. *Curr. Mol. Med.* **2009**, *9* (5), 527-45.
100. Dror, R. O.; Mildorf, T. J.; Hilger, D.; Manglik, A.; Borhani, D. W.; Arlow, D. H.; Philippsen, A.; Villanueva, N.; Yang, Z.; Lerch, M. T.; Hubbell, W. L.; Kobilka, B. K.; Sunahara, R. K.; Shaw, D. E., Structural basis for nucleotide exchange in heterotrimeric G proteins. *Science* **2015**, *348* (6241), 1361-1365.
101. Crusemann, M.; Reher, R.; Schamari, I.; Brachmann, A. O.; Ohbayashi, T.; Kuschak, M.; Malfacini, D.; Seidinger, A.; Pinto-Carbo, M.; Richarz, R.; Reuter, T.; Kehraus, S.; Hallab, A.; Attwood, M.; Schioth, H. B.; Mergaert, P.; Kikuchi, Y.; Schaberle, T. F.; Kostenis, E.; Wenzel, D.; Muller, C. E.; Piel, J.; Carlier, A.; Eberl, L.; Konig, G. M., Heterologous Expression, Biosynthetic Studies, and Ecological Function of the Selective Gq-Signaling Inhibitor FR900359. *Angew Chem Int Ed Engl* **2018**, *57* (3), 836-840.
102. Zhang, H.; Xiong, X. F.; Boesgaard, M. W.; Underwood, C. R.; Brauner-Osborne, H.; Stromgaard, K., Structure-Activity Relationship Studies of the Cyclic Depsipeptide Natural Product YM-254890, Targeting the Gq Protein. *ChemMedChem* **2017**, *12* (11), 830-834.
103. Taniguchi, M.; Suzumura, K.; Nagai, K.; Kawasaki, T.; Takasaki, J.; Sekiguchi, M.; Moritani, Y.; Saito, T.; Hayashi, K.; Fujita, S.; Tsukamoto, S.; Suzuki, K., YM-254890 analogues, novel cyclic depsipeptides with Galpha(q/11) inhibitory activity from *Chromobacterium* sp. QS3666. *Bioorg. Med. Chem.* **2004**, *12* (12), 3125-33.
104. Xiong, X. F.; Zhang, H.; Underwood, C. R.; Harpsoe, K.; Gardella, T. J.; Woldike, M. F.; Mannstadt, M.; Gloriam, D. E.; Brauner-Osborne, H.; Stromgaard, K., Total synthesis and structure-activity relationship studies of a series of selective G protein inhibitors. *Nat Chem* **2016**, *8* (11), 1035-1041.
105. Kaur, H.; Harris, P. W.; Little, P. J.; Brimble, M. A., Total synthesis of the cyclic depsipeptide YM-280193, a platelet aggregation inhibitor. *Org. Lett.* **2015**, *17* (3), 492-5.
106. Rensing, D. T.; Uppal, S.; Blumer, K. J.; Moeller, K. D., Toward the Selective Inhibition of G Proteins: Total Synthesis of a Simplified YM-254890 Analog. *Org. Lett.* **2015**, *17* (9), 2270-3.
107. Fujioka, M.; Koda, S.; Morimoto, Y.; Biemann, K., Structure of FR900359, a cyclic depsipeptide from *Ardisia crenata* Sims. *The Journal of Organic Chemistry* **1988**, *53* (12), 2820-2825.
108. Harvey, R. D.; Belevych, A. E., Muscarinic regulation of cardiac ion channels. *Br. J. Pharmacol.* **2003**, *139* (6), 1074-84.
109. Scott, G. D.; Fryer, A. D., Role of parasympathetic nerves and muscarinic receptors in allergy and asthma. *Chem. Immunol. Allergy* **2012**, *98*, 48-69.
110. Ivanova-Nikolova, T. T.; Nikolov, E. N.; Hansen, C.; Robishaw, J. D., Muscarinic K⁺ channel in the heart. Modal regulation by G protein beta gamma subunits. *J. Gen. Physiol.* **1998**, *112* (2), 199-210.
111. Gijsbers, L.; Molenberg, F. J.; Bakker, S. J.; Geleijnse, J. M., Potassium supplementation and heart rate: A meta-analysis of randomized controlled trials. *Nutr Metab Cardiovasc Dis* **2016**, *26* (8), 674-82.
112. Kline, R. P.; Morad, M., Potassium efflux in heart muscle during activity: extracellular accumulation and its implications. *J Physiol* **1978**, *280*, 537-58.

113. Arshavsky, V. Y.; Dumke, C. L.; Zhu, Y.; Artemyev, N. O.; Skiba, N. P.; Hamm, H. E.; Bownds, M. D., Regulation of transducin GTPase activity in bovine rod outer segments. *J. Biol. Chem.* **1994**, *269* (31), 19882-7.
114. Chabre, M.; Vuong, T. M., Kinetics and energetics of the rhodopsin-transducin-cGMP phosphodiesterase cascade of visual transduction. *Biochim. Biophys. Acta* **1992**, *1101* (2), 260-3.
115. Okada, T.; Sugihara, M.; Bondar, A. N.; Elstner, M.; Entel, P.; Buss, V., The retinal conformation and its environment in rhodopsin in light of a new 2.2 Å crystal structure. *J. Mol. Biol.* **2004**, *342* (2), 571-83.
116. Govorunova, E. G.; Sineshchekov, O. A.; Janz, R.; Liu, X.; Spudich, J. L., NEUROSCIENCE. Natural light-gated anion channels: A family of microbial rhodopsins for advanced optogenetics. *Science* **2015**, *349* (6248), 647-50.
117. Bauer, P. J., The complex of cGMP-gated channel and Na⁺/Ca²⁺, K⁺ exchanger in rod photoreceptors. *Adv. Exp. Med. Biol.* **2002**, *514*, 253-74.
118. Fu, Y., Phototransduction in Rods and Cones. In *Webvision: The Organization of the Retina and Visual System*, Kolb, H.; Fernandez, E.; Nelson, R., Eds. Salt Lake City (UT), 1995.
119. Beech, D. J.; Barnes, S., Characterization of a voltage-gated K⁺ channel that accelerates the rod response to dim light. *Neuron* **1989**, *3* (5), 573-81.
120. Onken, M. D.; Makepeace, C. M.; Kaltenbronn, K. M.; Kanai, S. M.; Todd, T. D.; Wang, S.; Broekelmann, T. J.; Rao, P. K.; Cooper, J. A.; Blumer, K. J., Targeting nucleotide exchange to inhibit constitutively active G protein alpha subunits in cancer cells. *Sci Signal* **2018**, *11* (546).
121. Katada, T.; Ui, M., ADP ribosylation of the specific membrane protein of C6 cells by islet-activating protein associated with modification of adenylate cyclase activity. *J. Biol. Chem.* **1982**, *257* (12), 7210-6.
122. Taniguchi, M.; Nagai, K.; Arao, N.; Kawasaki, T.; Saito, T.; Moritani, Y.; Takasaki, J.; Hayashi, K.; Fujita, S.; Suzuki, K.; Tsukamoto, S., YM-254890, a novel platelet aggregation inhibitor produced by *Chromobacterium* sp QS3666. *J. Antibiot.* **2003**, *56* (4), 358-363.
123. Kawasaki, T.; Taniguchi, M.; Moritani, Y.; Hayashi, K.; Saito, T.; Takasaki, J.; Nagai, K.; Inagaki, O.; Shikama, H., Antithrombotic and thrombolytic efficacy of YM-254890, a G_q/11 inhibitor, in a rat model of arterial thrombosis. *Thromb Haemost* **2003**, *90* (3), 406-13.
124. Kamato, D.; Thach, L.; Bernard, R.; Chan, V.; Zheng, W.; Kaur, H.; Brimble, M.; Osman, N.; Little, P. J., Structure, Function, Pharmacology, and Therapeutic Potential of the G Protein, Galpha/q,11. *Front Cardiovasc Med* **2015**, *2*, 14.
125. Malfacini, D.; Patt, J.; Annala, S.; Harpsoe, K.; Eryilmaz, F.; Reher, R.; Crusemann, M.; Hanke, W.; Zhang, H.; Tietze, D.; Gloriam, D. E.; Brauner-Osborne, H.; Stromgaard, K.; Konig, G. M.; Inoue, A.; Gomeza, J.; Kostenis, E., Rational design of a heterotrimeric G protein alpha subunit with artificial inhibitor sensitivity. *J. Biol. Chem.* **2019**, *294* (15), 5747-5758.
126. Koehn, F. E.; Carter, G. T., The evolving role of natural products in drug discovery. *Nat Rev Drug Discov* **2005**, *4* (3), 206-20.
127. Harvey, A. L.; Edrada-Ebel, R.; Quinn, R. J., The re-emergence of natural products for drug discovery in the genomics era. *Nat Rev Drug Discov* **2015**, *14* (2), 111-29.
128. Klepac, K.; Kilic, A.; Gnad, T.; Brown, L. M.; Herrmann, B.; Wilderman, A.; Balkow, A.; Glode, A.; Simon, K.; Lidell, M. E.; Betz, M. J.; Enerback, S.; Wess, J.; Freichel, M.; Bluher, M.; Konig, G.; Kostenis, E.; Insel, P. A.; Pfeifer, A., The G_q signalling pathway inhibits brown and beige adipose tissue. *Nat Commun* **2016**, *7*, 10895.
129. Matthey, M.; Roberts, R.; Seidinger, A.; Simon, A.; Schroder, R.; Kuschak, M.; Annala, S.; Konig, G. M.; Muller, C. E.; Hall, I. P.; Kostenis, E.; Fleischmann, B. K.; Wenzel, D., Targeted inhibition of G_q signaling induces airway relaxation in mouse models of asthma. *Sci Transl Med* **2017**, *9* (407).
130. Grundmann, M.; Merten, N.; Malfacini, D.; Inoue, A.; Preis, P.; Simon, K.; Ruttiger, N.; Ziegler, N.; Benkel, T.; Schmitt, N. K.; Ishida, S.; Muller, I.; Reher, R.; Kawakami, K.; Inoue, A.; Rick, U.; Kuhl, T.; Imhof, D.; Aoki, J.; Konig, G. M.; Hoffmann, C.; Gomeza, J.; Wess, J.; Kostenis, E., Lack of beta-arrestin signaling in the absence of active G proteins. *Nat Commun* **2018**, *9* (1), 341.
131. Lipkind, G. M.; Fozzard, H. A., Voltage-gated Na channel selectivity: the role of the conserved domain III lysine residue. *J. Gen. Physiol.* **2008**, *131* (6), 523-9.
132. Catterall, W. A., From ionic currents to molecular mechanisms: the structure and function of voltage-gated sodium channels. *Neuron* **2000**, *26* (1), 13-25.
133. Xu, H.; Li, T.; Rohou, A.; Arthur, C. P.; Tzakoniati, F.; Wong, E.; Estevez, A.; Kugel, C.; Franke, Y.; Chen, J.; Ciferri, C.; Hackos, D. H.; Koth, C. M.; Payandeh, J., Structural Basis of Nav1.7 Inhibition by a Gating-Modifier Spider Toxin. *Cell* **2019**, *176* (5), 1238-1239.

134. Groome, J. R.; Winston, V., S1-S3 counter charges in the voltage sensor module of a mammalian sodium channel regulate fast inactivation. *J. Gen. Physiol.* **2013**, *141* (5), 601-18.
135. Jensen, M. O.; Jogini, V.; Borhani, D. W.; Leffler, A. E.; Dror, R. O.; Shaw, D. E., Mechanism of Voltage Gating in Potassium Channels. *Science* **2012**, *336* (6078), 229-233.
136. Matthies, D.; Bae, C.; Toombes, G. E.; Fox, T.; Bartesaghi, A.; Subramaniam, S.; Swartz, K. J., Single-particle cryo-EM structure of a voltage-activated potassium channel in lipid nanodiscs. *Elife* **2018**, *7*.
137. Bourinet, E.; Zamponi, G. W., Block of voltage-gated calcium channels by peptide toxins. *Neuropharmacology* **2017**, *127*, 109-115.
138. Abraham, N.; Lewis, R. J., Neuronal Nicotinic Acetylcholine Receptor Modulators from Cone Snails. *Mar Drugs* **2018**, *16* (6).
139. Adams, D. J.; Callaghan, B.; Berecki, G., Analgesic conotoxins: block and G protein-coupled receptor modulation of N-type (Ca(V) 2.2) calcium channels. *Br. J. Pharmacol.* **2012**, *166* (2), 486-500.
140. Ellinor, P. T.; Zhang, J. F.; Horne, W. A.; Tsien, R. W., Structural determinants of the blockade of N-type calcium channels by a peptide neurotoxin. *Nature* **1994**, *372* (6503), 272-5.
141. McCleskey, E. W.; Fox, A. P.; Feldman, D. H.; Cruz, L. J.; Olivera, B. M.; Tsien, R. W.; Yoshikami, D., Omega-conotoxin: direct and persistent blockade of specific types of calcium channels in neurons but not muscle. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84* (12), 4327-31.
142. Feng, Z. P.; Doering, C. J.; Winkfein, R. J.; Beedle, A. M.; Spafford, J. D.; Zamponi, G. W., Determinants of inhibition of transiently expressed voltage-gated calcium channels by omega-conotoxins GVIA and MVIIA. *J. Biol. Chem.* **2003**, *278* (22), 20171-8.
143. Terlau, H.; Boccaccio, A.; Olivera, B. M.; Conti, F., The block of Shaker K⁺ channels by kappa-conotoxin PVIIA is state dependent. *J. Gen. Physiol.* **1999**, *114* (1), 125-40.
144. Bähring, R.; Boland, L. M.; Varghese, A.; Gebauer, M.; Pongs, O., Kinetic analysis of open- and closed-state inactivation transitions in human Kv4.2 A-type potassium channels. *J Physiol* **2001**, *535* (Pt 1), 65-81.
145. Beyl, S.; Kugler, P.; Kudrnac, M.; Hohaus, A.; Hering, S.; Timin, E., Different pathways for activation and deactivation in CaV1.2: a minimal gating model. *J. Gen. Physiol.* **2009**, *134* (3), 231-41; S1-2.
146. Hui, K.; Lipkind, G.; Fozzard, H. A.; French, R. J., Electrostatic and steric contributions to block of the skeletal muscle sodium channel by mu-conotoxin. *J. Gen. Physiol.* **2002**, *119* (1), 45-54.
147. Zhorov, B. S., Structural Models of Ligand-Bound Sodium Channels. *Handb. Exp. Pharmacol.* **2018**, *246*, 251-269.
148. Dudley, S. C., Jr.; Chang, N.; Hall, J.; Lipkind, G.; Fozzard, H. A.; French, R. J., mu-conotoxin GIIIA interactions with the voltage-gated Na⁽⁺⁾ channel predict a clockwise arrangement of the domains. *J. Gen. Physiol.* **2000**, *116* (5), 679-90.
149. Wang, S.; Ding, M.; Wu, D.; Zhan, J.; Chen, Z., omega-Conotoxin MVIIA inhibits amygdaloid kindled seizures in Sprague-Dawley rats. *Neurosci. Lett.* **2007**, *413* (2), 163-7.
150. Alonso, D.; Khalil, Z.; Satkunanathan, N.; Livett, B. G., Drugs from the sea: conotoxins as drug leads for neuropathic pain and other neurological conditions. *Mini Rev Med Chem* **2003**, *3* (7), 785-7.
151. Tietze, A. A.; Tietze, D.; Ohlenschlager, O.; Leipold, E.; Ullrich, F.; Kuhl, T.; Mischo, A.; Buntkowsky, G.; Gorch, M.; Heinemann, S. H.; Imhof, D., Structurally Diverse mu-Conotoxin PIIIA Isomers Block Sodium Channel Na(V) 1.4. *Angew. Chem., Int. Ed.* **2012**.
152. Terlau, H.; Olivera, B. M., Conus venoms: a rich source of novel ion channel-targeted peptides. *Physiol. Rev.* **2004**, *84* (1), 41-68.
153. Dutton, J. L.; Bansal, P. S.; Hogg, R. C.; Adams, D. J.; Alewood, P. F.; Craik, D. J., A new level of conotoxin diversity, a non-native disulfide bond connectivity in alpha-conotoxin AuIB reduces structural definition but increases biological activity. *J. Biol. Chem.* **2002**, *277* (50), 48849-57.
154. Carstens, B. B.; Berecki, G.; Daniel, J. T.; Lee, H. S.; Jackson, K. A.; Tae, H. S.; Sadeghi, M.; Castro, J.; O'Donnell, T.; Deiteren, A.; Brierley, S. M.; Craik, D. J.; Adams, D. J.; Clark, R. J., Structure-Activity Studies of Cysteine-Rich alpha-Conotoxins that Inhibit High-Voltage-Activated Calcium Channels via GABA(B) Receptor Activation Reveal a Minimal Functional Motif. *Angew Chem Int Ed Engl* **2016**, *55* (15), 4692-6.
155. Olivera, B. M.; Rivier, J.; Scott, J. K.; Hillyard, D. R.; Cruz, L. J., Conotoxins. *J. Biol. Chem.* **1991**, *266* (33), 22067-70.
156. Weissman, J. S.; Kim, P. S., Reexamination of the folding of BPTI: predominance of native intermediates. *Science* **1991**, *253* (5026), 1386-93.
157. Creighton, T. E., The disulfide folding pathway of BPTI. *Science* **1992**, *256* (5053), 111-4.
158. Weissman, J. S.; Kim, P. S., Kinetic role of nonnative species in the folding of bovine pancreatic trypsin inhibitor. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89* (20), 9900-4.

159. Weissman, J. S.; Kim, P. S., A kinetic explanation for the rearrangement pathway of BPTI folding. *Nat. Struct. Biol.* **1995**, *2* (12), 1123-30.
160. Fersht, A. R., From the first protein structures to our current knowledge of protein folding: delights and scepticisms. *Nat. Rev. Mol. Cell Biol.* **2008**, *9* (8), 650-4.
161. Chang, J. Y., Diverse pathways of oxidative folding of disulfide proteins: underlying causes and folding models. *Biochemistry* **2011**, *50* (17), 3414-31.
162. Terlau, H.; Olivera, B. M., Conus venoms: A rich source of novel ion channel-targeted peptides. *Physiol. Rev.* **2004**, *84* (1), 41-68.
163. Jacob, R. B.; McDougal, O. M., The M-superfamily of conotoxins: a review. *Cell. Mol. Life Sci.* **2010**, *67* (1), 17-27.
164. Becker, S.; Terlau, H., Toxins from cone snails: properties, applications and biotechnological production. *Appl. Microbiol. Biotechnol.* **2008**, *79* (1), 1-9.
165. Huang, Y.; Peng, C.; Yi, Y.; Gao, B.; Shi, Q., A Transcriptomic Survey of Ion Channel-Based Conotoxins in the Chinese Tubular Cone Snail (*Conus betulinus*). *Mar Drugs* **2017**, *15* (7).
166. Peng, C.; Yao, G.; Gao, B. M.; Fan, C. X.; Bian, C.; Wang, J.; Cao, Y.; Wen, B.; Zhu, Y.; Ruan, Z.; Zhao, X.; You, X.; Bai, J.; Li, J.; Lin, Z.; Zou, S.; Zhang, X.; Qiu, Y.; Chen, J.; Coon, S. L.; Yang, J.; Chen, J. S.; Shi, Q., High-throughput identification of novel conotoxins from the Chinese tubular cone snail (*Conus betulinus*) by multi-transcriptome sequencing. *Gigascience* **2016**, *5*, 17.
167. Salkoff, L.; Wei, A. D.; Baban, B.; Butler, A.; Fawcett, G.; Ferreira, G.; Santi, C. M., Potassium channels in *C. elegans*. *WormBook* **2005**, 1-15.
168. Gutman, G. A.; Chandy, K. G.; Grissmer, S.; Lazdunski, M.; McKinnon, D.; Pardo, L. A.; Robertson, G. A.; Rudy, B.; Sanguinetti, M. C.; Stuhmer, W.; Wang, X., International Union of Pharmacology. LIII. Nomenclature and molecular relationships of voltage-gated potassium channels. *Pharmacol. Rev.* **2005**, *57* (4), 473-508.
169. Kim, J. B., Channelopathies. *Korean J Pediatr* **2014**, *57* (1), 1-18.
170. Gargus, J. J., Unraveling monogenic channelopathies and their implications for complex polygenic disease. *Am. J. Hum. Genet.* **2003**, *72* (4), 785-803.
171. Comes, N.; Bielanska, J.; Vallejo-Gracia, A.; Serrano-Albarras, A.; Marruecos, L.; Gomez, D.; Soler, C.; Condom, E.; Ramon, Y. C. S.; Hernandez-Losa, J.; Ferreres, J. C.; Felipe, A., The voltage-dependent K(+) channels Kv1.3 and Kv1.5 in human cancer. *Front Physiol* **2013**, *4*, 283.
172. Tottene, A.; Conti, R.; Fabbro, A.; Vecchia, D.; Shapovalova, M.; Santello, M.; van den Maagdenberg, A. M.; Ferrari, M. D.; Pietrobon, D., Enhanced excitatory transmission at cortical synapses as the basis for facilitated spreading depression in Ca(v)2.1 knockin migraine mice. *Neuron* **2009**, *61* (5), 762-73.
173. Backryd, E., Do the potential benefits outweigh the risks? An update on the use of ziconotide in clinical practice. *Eur J Pain* **2018**.
174. Nair, A. S.; Poornachand, A.; Kodisharapu, P. K., Ziconotide: Indications, Adverse Effects, and Limitations in Managing Refractory Chronic Pain. *Indian J Palliat Care* **2018**, *24* (1), 118-119.
175. Patel, R.; Montagut-Bordas, C.; Dickenson, A. H., Calcium channel modulation as a target in chronic pain control. *Br. J. Pharmacol.* **2018**, *175* (12), 2173-2184.
176. Rauck, R. L.; Wallace, M. S.; Burton, A. W.; Kapural, L.; North, J. M., Intrathecal ziconotide for neuropathic pain: a review. *Pain Pract* **2009**, *9* (5), 327-37.
177. Wang, F.; Yan, Z.; Liu, Z.; Wang, S.; Wu, Q.; Yu, S.; Ding, J.; Dai, Q., Molecular basis of toxicity of N-type calcium channel inhibitor MVIIA. *Neuropharmacology* **2016**, *101*, 137-45.
178. Poli, P.; Ciaramella, A., Psychiatric predisposition to autonomic and abnormal perception side-effects of ziconotide: a case series study. *Neuromodulation* **2011**, *14* (3), 219-24; discussion 224.
179. Thompson, J. C.; Dunbar, E.; Laye, R. R., Treatment challenges and complications with ziconotide monotherapy in established pump patients. *Pain Physician* **2006**, *9* (2), 147-52.
180. Ramirez, D.; Gonzalez, W.; Fissore, R. A.; Carvacho, I., Conotoxins as Tools to Understand the Physiological Function of Voltage-Gated Calcium (CaV) Channels. *Mar Drugs* **2017**, *15* (10).
181. Lee, S.; Kim, Y.; Back, S. K.; Choi, H. W.; Lee, J. Y.; Jung, H. H.; Ryu, J. H.; Suh, H. W.; Na, H. S.; Kim, H. J.; Rhim, H.; Kim, J. I., Analgesic effect of highly reversible omega-conotoxin FVIA on N type Ca²⁺ channels. *Mol Pain* **2010**, *6*, 97.
182. Favreau, P.; Gilles, N.; Lamthanh, H.; Bournaud, R.; Shimahara, T.; Bouet, F.; Laboute, P.; Letourneux, Y.; Menez, A.; Molgo, J.; Le Gall, F., A new omega-conotoxin that targets N-type voltage-sensitive calcium channels with unusual specificity. *Biochemistry* **2001**, *40* (48), 14567-75.
183. Safavi-Hemami, H.; Brogan, S. E.; Olivera, B. M., Pain therapeutics from cone snail venoms: From Ziconotide to novel non-opioid pathways. *J Proteomics* **2018**.

184. McGivern, J. G., Ziconotide: a review of its pharmacology and use in the treatment of pain. *Neuropsychiatr Dis Treat* **2007**, *3* (1), 69-85.
185. Lyons, M. S.; Bell, B.; Stainier, D.; Peters, K. G., Isolation of the zebrafish homologues for the tie-1 and tie-2 endothelium-specific receptor tyrosine kinases. *Dev. Dyn.* **1998**, *212* (1), 133-40.
186. Steppan, C. M.; Bailey, S. T.; Bhat, S.; Brown, E. J.; Banerjee, R. R.; Wright, C. M.; Patel, H. R.; Ahima, R. S.; Lazar, M. A., The hormone resistin links obesity to diabetes. *Nature* **2001**, *409* (6818), 307-12.
187. Chang, T. H.; Hsieh, F. L.; Zebisch, M.; Harlos, K.; Elegheert, J.; Jones, E. Y., Structure and functional properties of Norrin mimic Wnt for signalling with Frizzled4, Lrp5/6, and proteoglycan. *Elife* **2015**, *4*.
188. Sitar, T.; Popowicz, G. M.; Siwanowicz, I.; Huber, R.; Holak, T. A., Structural basis for the inhibition of insulin-like growth factors by insulin-like growth factor-binding proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (35), 13028-33.
189. Christinger, H. W.; Fuh, G.; de Vos, A. M.; Wiesmann, C., The crystal structure of placental growth factor in complex with domain 2 of vascular endothelial growth factor receptor-1. *J. Biol. Chem.* **2004**, *279* (11), 10382-8.
190. Marshall, J. L.; Holmberg, J.; Chou, E.; Ocampo, A. C.; Oh, J.; Lee, J.; Peter, A. K.; Martin, P. T.; Crosbie-Watson, R. H., Sarcospan-dependent Akt activation is required for utrophin expression and muscle regeneration. *J. Cell Biol.* **2012**, *197* (7), 1009-27.
191. Hunter, H. N.; Fulton, D. B.; Ganz, T.; Vogel, H. J., The solution structure of human hepcidin, a peptide hormone with antimicrobial activity that is involved in iron uptake and hereditary hemochromatosis. *J. Biol. Chem.* **2002**, *277* (40), 37597-603.
192. Szyk, A.; Wu, Z.; Tucker, K.; Yang, D.; Lu, W.; Lubkowski, J., Crystal structures of human alpha-defensins HNP4, HD5, and HD6. *Protein Sci.* **2006**, *15* (12), 2749-60.
193. Ranasinghe, S.; McManus, D. P., Structure and function of invertebrate Kunitz serine protease inhibitors. *Dev. Comp. Immunol.* **2013**, *39* (3), 219-27.
194. Shigetomi, H.; Onogi, A.; Kajiwara, H.; Yoshida, S.; Furukawa, N.; Haruta, S.; Tanase, Y.; Kanayama, S.; Noguchi, T.; Yamada, Y.; Oi, H.; Kobayashi, H., Anti-inflammatory actions of serine protease inhibitors containing the Kunitz domain. *Inflamm. Res.* **2010**, *59* (9), 679-87.
195. Zhao, R.; Dai, H.; Qiu, S.; Li, T.; He, Y.; Ma, Y.; Chen, Z.; Wu, Y.; Li, W.; Cao, Z., SdPI, the first functionally characterized Kunitz-type trypsin inhibitor from scorpion venom. *PLoS One* **2011**, *6* (11), e27548.
196. Leader, B.; Baca, Q. J.; Golan, D. E., Protein therapeutics: a summary and pharmacological classification. *Nat Rev Drug Discov* **2008**, *7* (1), 21-39.
197. Poppe, L.; Hui, J. O.; Ligutti, J.; Murray, J. K.; Schnier, P. D., PADLOC: a powerful tool to assign disulfide bond connectivities in peptides and proteins by NMR spectroscopy. *Anal. Chem.* **2012**, *84* (1), 262-6.
198. Gehrmann, J.; Alewood, P. F.; Craik, D. J., Structure determination of the three disulfide bond isomers of alpha-conotoxin GI: a model for the role of disulfide bonds in structural stability. *J. Mol. Biol.* **1998**, *278* (2), 401-15.
199. Allerton, C., *Pain Therapeutics: Current and Future Treatment Paradigms*. 2013.
200. Chen, R.; Robinson, A.; Chung, S.-H., Mechanism of μ -Conotoxin P111A Binding to the Voltage-Gated Na⁺ Channel Nav1.4. *PLoS One* **2014**, *9* (3), e93267.
201. Wu, J.; Yan, Z.; Li, Z.; Yan, C.; Lu, S.; Dong, M.; Yan, N., Structure of the voltage-gated calcium channel Cav1.1 complex. *Science* **2015**, *350* (6267), aad2395.
202. Pan, X.; Li, Z.; Zhou, Q.; Shen, H.; Wu, K.; Huang, X.; Chen, J.; Zhang, J.; Zhu, X.; Lei, J.; Xiong, W.; Gong, H.; Xiao, B.; Yan, N., Structure of the human voltage-gated sodium channel Nav1.4 in complex with beta1. *Science* **2018**, *362* (6412).
203. Wu, J.; Yan, Z.; Li, Z.; Qian, X.; Lu, S.; Dong, M.; Zhou, Q.; Yan, N., Structure of the voltage-gated calcium channel Cav1.1 at 3.6 Å resolution. *Nature* **2016**, *537* (7619), 191-196.
204. Livingstone, C. D.; Barton, G. J., Protein sequence alignments: a strategy for the hierarchical analysis of residue conservation. *Comput Appl Biosci* **1993**, *9* (6), 745-56.
205. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; Thompson, J. D.; Higgins, D. G., *Fast, scalable generation of high - quality protein multiple sequence alignments using Clustal Omega*. 2011; Vol. 7.
206. McWilliam, H.; Li, W.; Uludag, M.; Squizzato, S.; Park, Y. M.; Buso, N.; Cowley, A. P.; Lopez, R., Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Res.* **2013**, *41* (W1), W597-W600.
207. Goujon, M.; McWilliam, H.; Li, W.; Valentin, F.; Squizzato, S.; Paern, J.; Lopez, R., A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res.* **2010**, *38* (suppl 2), W695-W699.

208. Waterhouse, A. M.; Procter, J. B.; Martin, D. M.; Clamp, M.; Barton, G. J., Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **2009**, *25* (9), 1189-91.
209. Dominguez, C.; Boelens, R.; Bonvin, A. M., HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.* **2003**, *125* (7), 1731-7.
210. de Vries, S. J.; van Dijk, A. D.; Krzeminski, M.; van Dijk, M.; Thureau, A.; Hsu, V.; Wassenaar, T.; Bonvin, A. M., HADDOCK versus HADDOCK: new features and performance of HADDOCK2.0 on the CAPRI targets. *Proteins* **2007**, *69* (4), 726-33.
211. McArthur, J. R.; Singh, G.; O'Mara, M. L.; McMaster, D.; Ostroumov, V.; Tieleman, D. P.; French, R. J., ORIENTATION OF {micro}-CONOTOXIN PIIIA IN A SODIUM CHANNEL VESTIBULE, BASED ON ITS VOLTAGE DEPENDENCE OF BINDING. *Mol. Pharmacol.* **2011**.
212. McArthur, J. R.; Ostroumov, V.; Al-Sabi, A.; McMaster, D.; French, R. J., Multiple, distributed interactions of mu-conotoxin PIIIA associated with broad targeting among voltage-gated sodium channels. *Biochemistry* **2011**, *50* (1), 116-24.
213. Mahdavi, S.; Kuyucak, S., Molecular dynamics study of binding of micro-conotoxin GIIIA to the voltage-gated sodium channel Na(v)1.4. *PLoS One* **2014**, *9* (8), e105300.
214. Mahdavi, S.; Kuyucak, S., Systematic study of binding of mu-conotoxins to the sodium channel NaV1.4. *Toxins (Basel)* **2014**, *6* (12), 3454-70.
215. Chen, R.; Robinson, A.; Chung, S. H., Mechanism of mu-conotoxin PIIIA binding to the voltage-gated Na⁺ channel NaV1.4. *PLoS One* **2014**, *9* (3), e93267.
216. Abdel-Mottaleb, Y.; Clynen, E.; Jalali, A.; Bosmans, F.; Vatanpour, H.; Schoofs, L.; Tytgat, J., The first potassium channel toxin from the venom of the Iranian scorpion *Odonthobuthus doriae*. *FEBS Lett.* **2006**, *580* (26), 6254-8.
217. Srairi-Abid, N.; Shahbazzadeh, D.; Chatti, I.; Mlayah-Bellalouna, S.; Mejdoub, H.; Borchani, L.; Benkhalifa, R.; Akbari, A.; El Ayeb, M., Hemitoxin, the first potassium channel toxin from the venom of the Iranian scorpion *Hemiscorpius lepturus*. *FEBS J.* **2008**, *275* (18), 4641-50.
218. de Vries, S. J.; van Dijk, M.; Bonvin, A. M., The HADDOCK web server for data-driven biomolecular docking. *Nat. Protoc.* **2010**, *5* (5), 883-97.
219. Linge, J. P.; Habeck, M.; Rieping, W.; Nilges, M., ARIA: automated NOE assignment and NMR structure calculation. *Bioinformatics* **2003**, *19* (2), 315-6.
220. Trott, O.; Olson, A. J., AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31* (2), 455-61.
221. Chen, F.; Huang, W.; Jiang, T.; Yu, R., Determination of the mu-Conotoxin PIIIA Specificity Against Voltage-Gated Sodium Channels from Binding Energy Calculations. *Mar Drugs* **2018**, *16* (5).
222. Pan, X.; Li, Z.; Huang, X.; Huang, G.; Gao, S.; Shen, H.; Liu, L.; Lei, J.; Yan, N., Molecular basis for pore blockade of human Na(+) channel Nav1.2 by the mu-conotoxin KIIIA. *Science* **2019**, *363* (6433), 1309-1313.
223. Bielanska, J.; Hernandez-Losa, J.; Perez-Verdaguer, M.; Moline, T.; Somoza, R.; Ramon, Y. C. S.; Condom, E.; Ferreres, J. C.; Felipe, A., Voltage-dependent potassium channels Kv1.3 and Kv1.5 in human cancer. *Curr. Cancer Drug Targets* **2009**, *9* (8), 904-14.
224. Beshore, D. C.; Liverton, N. J.; McIntyre, C. J.; Claiborne, C. F.; Libby, B.; Culberson, J. C.; Salata, J. J.; Regan, C. P.; Lynch, J. J.; Kiss, L.; Spencer, R. H.; Kane, S. A.; White, R. B.; Yeh, S.; Hartman, G. D.; Dinsmore, C. J., Discovery of triarylethanolamine inhibitors of the Kv1.5 potassium channel. *Bioorg. Med. Chem. Lett.* **2010**, *20* (8), 2493-6.
225. Brendel, J.; Peukert, S., Blockers of the Kv1.5 channel for the treatment of atrial arrhythmias. *Curr Med Chem Cardiovasc Hematol Agents* **2003**, *1* (3), 273-87.
226. Bachmann, A.; Gutcher, I.; Kopp, K.; Brendel, J.; Bosch, R. F.; Busch, A. E.; Gogelein, H., Characterization of a novel Kv1.5 channel blocker in *Xenopus* oocytes, CHO cells, human and rat cardiomyocytes. *Naunyn Schmiedebergs Arch Pharmacol* **2001**, *364* (5), 472-8.
227. Lathrop, R. H., The protein threading problem with sequence amino acid interaction preferences is NP-complete. *Protein Eng.* **1994**, *7* (9), 1059-68.
228. Krieger, E.; Vriend, G., New ways to boost molecular dynamics simulations. *J. Comput. Chem.* **2015**, *36* (13), 996-1007.
229. Krieger, E.; Nielsen, J. E.; Spronk, C. A.; Vriend, G., Fast empirical pKa prediction by Ewald summation. *J. Mol. Graph. Model.* **2006**, *25* (4), 481-6.
230. Krieger, E.; Joo, K.; Lee, J.; Lee, J.; Raman, S.; Thompson, J.; Tyka, M.; Baker, D.; Karplus, K., Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. *Proteins: Struct., Funct., Bioinf.* **2009**, *77* (S9), 114-122.

231. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A., Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25* (9), 1157-74.
232. Jakalian, A.; Jack, D. B.; Bayly, C. I., Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comput. Chem.* **2002**, *23* (16), 1623-41.
233. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C., ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11* (8), 3696-713.
234. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C., Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **2006**, *65* (3), 712-25.
235. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G., A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103* (19), 8577-8593.
236. Case, D. A.; Babin, V.; Berryman, J. T.; Betz, R. M.; Cai, Q.; Cerutti, D. S.; Cheatham, T. E.; Darden, T. A.; Duke, R. E.; Gohlke, H.; Goetz, A. W.; Gusarov, S.; Homeyer, N.; Janowski, P.; Kaus, J.; Kolossváry, I.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Luchko, T.; Luo, R.; Madej, B.; Merz, K. M.; Paesani, F.; Roe, D. R.; Roitberg, A.; Sagui, C.; Salomon-Ferrer, R.; Seabra, G.; Simmerling, C. L.; Smith, W.; Swails, J.; Walker, Wang, J.; Wolf, R. M.; Wu, X.; Kollman, P. A., {Amber 14}. 2014.
237. Krieger, E.; Koraimann, G.; Vriend, G., Increasing the precision of comparative models with YASARA NOVA—a self-parameterizing force field. *Proteins: Struct., Funct., Bioinf.* **2002**, *47* (3), 393-402.
238. Payandeh, J.; Scheuer, T.; Zheng, N.; Catterall, W. A., The crystal structure of a voltage-gated sodium channel. *Nature* **2011**, *475* (7356), 353-8.
239. Konagurthu, A. S.; Whisstock, J. C.; Stuckey, P. J.; Lesk, A. M., MUSTANG: a multiple structural alignment algorithm. *Proteins* **2006**, *64* (3), 559-74.
240. Canutescu, A. A.; Shelenkov, A. A.; Dunbrack Jr, R. L., A graph theory algorithm for rapid protein side chain prediction. *Protein Sci.* **2003**, *12* (9), 2001-2014.
241. Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M., LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18* (12), 1463-1472.
242. Altschul, S. F.; Madden, T. L.; Schaffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J., Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25* (17), 3389-402.
243. Suzek, B. E.; Wang, Y.; Huang, H.; McGarvey, P. B.; Wu, C. H.; UniProt, C., UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches. *Bioinformatics* **2015**, *31* (6), 926-32.
244. Jones, D. T., Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* **1999**, *292* (2), 195-202.
245. Krieger, E.; Dunbrack, R. L., Jr.; Hooft, R. W.; Krieger, B., Assignment of protonation states in proteins and ligands: combining pKa prediction with hydrogen bonding network optimization. *Methods Mol. Biol.* **2012**, *819*, 405-21.
246. Nielsen, K. J.; Watson, M.; Adams, D. J.; Hammarstrom, A. K.; Gage, P. W.; Hill, J. M.; Craik, D. J.; Thomas, L.; Adams, D.; Alewood, P. F.; Lewis, R. J., Solution structure of mu-conotoxin PIIIA, a preferential inhibitor of persistent tetrodotoxin-sensitive sodium channels. *J. Biol. Chem.* **2002**, *277* (30), 27247-55.
247. Wakamatsu, K.; Kohda, D.; Hatanaka, H.; Lancelin, J. M.; Ishida, Y.; Oya, M.; Nakamura, H.; Inagaki, F.; Sato, K., Structure-activity relationships of mu-conotoxin GIIIA: structure determination of active and inactive sodium channel blocker peptides by NMR and simulated annealing calculations. *Biochemistry* **1992**, *31* (50), 12577-84.
248. van Zundert, G. C. P.; Rodrigues, J.; Trellet, M.; Schmitz, C.; Kastiris, P. L.; Karaca, E.; Melquiond, A. S. J.; van Dijk, M.; de Vries, S. J.; Bonvin, A., The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J. Mol. Biol.* **2016**, *428* (4), 720-725.
249. Trott, O.; Olson, A. J., Software News and Update AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2010**, *31* (2), 455-461.
250. Miyamoto, S.; Kollman, P. A., Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* **1992**, *13* (8), 952-962.

Datum: 15.05.2019

Erklärung

Ich erkläre hiermit, dass ich meine Dissertation selbstständig und nur mit den angegebenen Hilfsmitteln angefertigt und noch keinen Promotionsversuch unternommen habe.

Diana Desirée Kaufmann

Unterschrift

Erklärung der Übereinstimmung

Datum: 15.05.2019

Ich erkläre hiermit, dass die elektronische Version der Doktorarbeit mit der schriftlichen Version übereinstimmt. Die elektronische Version liegt dem Prüfungssekretariat vor.

Diana Desirée Kaufmann

Unterschrift