

Implementation of structural bioinformatics in thromboinflammation studies

Citation for published version (APA):

Liu, X. (2021). *Implementation of structural bioinformatics in thromboinflammation studies*. [Doctoral Thesis, Maastricht University]. Ridderprint. <https://doi.org/10.26481/dis.20211209xl>

Document status and date:

Published: 01/01/2021

DOI:

[10.26481/dis.20211209xl](https://doi.org/10.26481/dis.20211209xl)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Appendix I: English summary

We aim to study cardiovascular diseases by performing a combination of *in silico* and *in vitro* approaches. Over the last two decades, structural bioinformatics provides various computational methods which are powerful and efficient in aiding and supporting experimental methods, and together they overcome the challenges that limit classical approaches. In this thesis, we have shown the successes of their cooperation by focusing on their application in several potential therapeutic targets of interest, including PAD4, coagulation factor Va and CXCR4.

Chapter 1 gives the general overview of this thesis. Frequently used computational tools are described in this chapter. The structure based virtual screening workflow is often used to discover the first-in-class small molecules towards specific target. For a protein without crystal structure, structural bioinformatics provides homology modeling method to predict its secondary structure. The molecular dynamic simulation method can simulate the movement of apo protein state in solvent or its complex with protein partner/inhibitor, followed by the calculation of binding free energies between two components to estimate their binding affinities. The important roles of therapeutic targets of interest in cardiovascular diseases were also described here.

Chapter 2 reviews how to rationally design modulators (activators or inactivators) that target (interrupt or stabilize) protein-protein interactions (PPIs). Protein-protein interactions have received increasing attention from researchers as they play important roles in different biological pathways. However, unlike a typical drug target with binding pocket, protein-protein interaction-based target has a flat and large surface which remain a challenge to be targeted by small chemical molecules. Thus, the workflow of rational design of peptides and peptidomimetics by the application of *in silico* approaches was described in this chapter in detail. The structure based peptidic modulators design requires the knowledge of 3D structure of the protein-protein complexes of targets, we thus summarized current experimental and computational methods used to determine 3D structure of protein-protein complexes.

Chapter 3 reviews the current understanding of PAD4 target, including the structural information and biological functions of PAD4, the crucial role of PAD4 in NET formation, and the role of PAD4 in the pathogenesis of rheumatoid arthritis, sepsis and thrombosis. PAD4 requires high concentration calcium to become active *in vitro*, however, such high concentration is not achievable *in vivo*. Thus, it remains unclear whether there are other cofactors which activate PAD4 and subsequently lowers its requirement to calcium concentration. Although studies on the PAD4 and diseases are increasing, the mechanism of PAD4 activity control is still incomplete. As a promising therapeutic target, PAD4 has been investigated by focusing on the discovery of inhibitors by increasing studies, especially by Thompson's group. However, the overall high conservation in active site between different PAD isoforms forms a medicinal chemistry challenge with respect to the design of specific inhibitors. Finally, we compared and

discussed PAD4 residues that interact with substrate or different types of inhibitors and give our input for the design of a novel class of therapeutics.

Chapter 4 studies the effect of autocitrullination on PAD4 activity by performing *in vitro* and *in silico* approaches. PAD4 can citrullinate itself, which is called autocitrullination or autodeimination. It remains controversial whether autocitrullination reduces PAD4 activity. We hypothesized that the autocitrullination is a self-control mechanism that controls PAD4 activity. To test our hypothesis, we initially generated citrullinated PAD4 by the incubation of PAD4 alone in the presence of calcium and observed 24/27 arginine citrullinated on PAD4 after 2 hours. PAD4 citrullinates itself in a time- and concentration-dependent manner. The activities of both PAD4 and citrullinated PAD4 were evaluated using small-, medium-, and big- sized substrates, and the obtained results showed that no significant difference detected between them. Furthermore, we performed molecular dynamic (MD) simulations to investigate structural dynamics and binding abilities of PAD4 and citrullinated PAD4 to their substrate. The MD simulation results indicated that autocitrullination does not change the shape of substrate binding pocket on PAD4. The calculated binding free energies suggested that PAD4 and citrullinated PAD4 exhibit comparable binding abilities with a H3-derived peptidyl substrate. Overall, our findings suggested that the autocitrullination does not represent a self-control mechanism for PAD4 activity.

Chapter 5 investigates the structural integrity of factor Va in venom of *Pseudonaja textilis*. The human coagulation factor Va (hFVa) become inactivated after activated protein C (APC)-catalyzed proteolysis. In contrast with this, the *Pseudonaja textilis* venom-derived factor Va (ptFVa) remains cofactor function despite APC cleavage. based on the structural analysis, hFVa contains one more cleavage site at position Arg306 and one less disulfide bond which covalently connect A2 and A3 domain compared with ptFVa. In order to uncover the mechanism principle which determines functional difference between hFVa and ptFVa, we experimentally created chimeric ptFVa by the addition of human Arg306 cleavage site and the deletion of disulfide bond, however, it still exhibits cofactor function. Furthermore, we *in silico* prepared hFVa and chimeric ptFVa structure and performed MD simulations to explore whether non-covalent interactions play a role in their structure-determined activities. The results observed by computational approaches showed that hFVa has a A2 domain dissociation trend along MD simulation time which was not observed in chimeric ptFVa. In addition, the calculated binding free energy between A2 and A1+A3 domains in hFVa increased significantly when comparing its values at t=0ns and t=100ns. Finally, *in silico* approaches proposed a crucial loop which locates at A2 domain and could affects A2 domain association or dissociation state in ptFVa or hFVa.

Chapter 6 rationally designed various peptides which target C-X-C chemokine receptor type 4 (CXCR4) by the application of the workflow of rational design peptides based on PPIs using *in silico* approaches (as described in chapter 2). As CXCR4 is described as the major mediator for neutrophil aging, we aim to develop novel peptides as therapeutics which can activate CXCR4 and limit aged neutrophils for the high risk of patients with vascular diseases. We proposed two strategies by designing peptides which are supposed to bind CXCR4 intracellularly or extracellularly. We adopted a known pepducin and natural

ligand CXCL12 as template through intracellular and extracellular designing path, respectively. All designed peptides were optimized through short MD simulations and their binding affinities to CXCR4 were estimated based on the binding free energy calculation. Finally, we have in total designed 75 peptides with selected three peptides displaying biological activities against CXCR4. Our presented methods here for peptide design can significantly reduce time and financial cost and can be used further for the development of peptidic modulators against other GPCR receptors.

Chapter 7 gives the critical discussion of the results obtained using the combination of computational and experimental approaches in this thesis. We also discussed the challenges and future development of the computational approaches.

Appendix II: Nederlandse samenvatting

Het doel van het onderzoek, beschreven in dit proefschrift, is de bestudering van hart- en vaatziekten door toepassing van een combinatie van *in silico* en *in vitro* benaderingen. De afgelopen twee decennia heeft structurele bio-informatica verschillende computationele methodes opgeleverd die krachtig en efficiënt zijn en die experimentele benaderingen kunnen helpen en ondersteunen. Daarmee bieden ze een oplossing voor de beperkingen die klassieke onderzoeksmethodes met zich meebrengen. In dit proefschrift hebben we het succes van deze combinatie van methodes aangetoond door ons te concentreren op hun toepassing in verschillende potentiële therapeutische doelwitten, waaronder PAD4, stollingsfactor Va en CXCR4.

Hoofdstuk 1 geeft een algemeen overzicht van dit proefschrift. In dit hoofdstuk worden enkele van de meest gebruikte computationele methoden beschreven. De op structuur gebaseerde virtuele screening workflow wordt vaak gebruikt om kleine moleculen voor een specifiek doelwit te ontdekken. Voor een eiwit zonder bekende experimentele 3D structuur biedt structurele bioinformatica een op homologie gebaseerde modelleringsmethode om de secundaire structuur van een eiwit te kunnen voorspellen. De moleculair dynamische simulatiemethode kan de beweging van een apo-eiwit toestand in oplossing of zijn complex met een eiwit partner/remmer simuleren, gevolgd door de berekening van bindingsvrije energieën tussen twee componenten om zo hun bindingsaffiniteiten te kunnen schatten. De rol van therapeutische doelwitten die van belang zijn bij hart- en vaatziekten worden hier ook beschreven.

Hoofdstuk 2 bespreekt hoe men weloverwogen modulatoren (activatoren of inactivatoren) kan ontwerpen die zich richten op het onderbreken of stabiliseren van eiwit-eiwitinteracties. Eiwit-eiwit interacties hebben steeds meer aandacht gekregen van onderzoekers omdat ze een belangrijke rol spelen in verschillende biologische processen. In tegenstelling tot een typisch geneesmiddel doelwit met een bindingsplaats, heeft een op eiwit-eiwit interactie gebaseerd doelwit een groot en egaal oppervlak dat een uitdaging blijft om door kleine chemische moleculen te worden gebonden. Zo wordt de workflow van het rationeel ontwerpen van peptiden en peptidomimetica door de toepassing van *in silico*-benaderingen in detail beschreven in dit hoofdstuk. Het op structuur-gebaseerde ontwerp van peptidische modulatoren vereist de kennis van de 3D-structuur van de eiwit-eiwitcomplexen van onze doelwitten, dus hebben we de huidige experimentele en computationele methoden samengevat die worden gebruikt om de 3D-structuur van eiwit-eiwitcomplexen te bepalen.

Hoofdstuk 3 geeft een overzicht van de huidige kennis van het PAD4 eiwit, inclusief de structurele informatie en biologische functies, de cruciale rol van PAD4 in NET vorming, en de rol van PAD4 in de pathogenese van reumatoïde artritis, sepsis en trombose. PAD4 vereist een hoge concentratie calcium om actief te worden *in vitro*, maar een dergelijke hoge concentratie is *in vivo* niet haalbaar. Het blijft dus onduidelijk of er andere cofactoren zijn die PAD4 activeren en vervolgens de calciumdrempel verlagen. Hoewel het aantal gepubliceerde studies over het verband tussen PAD4 en ziekten toeneemt, is het

exacte mechanisme van PAD4 activatie nog steeds onduidelijk. PAD4 staat daarnaast in de belangstelling als een veelbelovend therapeutisch doelwit door de ontdekking van PAD4 remmers, met name door de groep van Prof. Thompson in Worcester, Verenigde Staten. De algehele hoge conservering van de actieve site, het katalytische hart van het enzym, tussen verschillende PAD-isovormen vormt echter een uitdaging voor de medicinale chemie met betrekking tot het ontwerp van specifieke remmers. Ten slotte hebben we PAD4-residuen die een interactie aangaan met het substraat of verschillende soorten remmers vergeleken en besproken en hebben we voorstellen gegeven voor het ontwerp van een nieuwe klasse van therapieën.

Hoofdstuk 4 bestudeert het effect van autocitrullinatie op PAD4-activiteit door *in vitro* en *in silico* benaderingen uit te voeren. PAD4 is in staat zichzelf citrullineren, wat autocitrullinatie of autodeïminatie wordt genoemd. Het blijft controversieel of autocitrullinatie de PAD4-activiteit wel of niet vermindert. We veronderstelden dat de autocitrullinatie een zelfbeheersingsmechanisme is dat de PAD4-activiteit reguleert. Om onze hypothese te testen, hebben we in eerste instantie gecitrullineerd PAD4 gemaakt door de incubatie van enkel PAD4 in de aanwezigheid van calcium en vonden we na 2 uur dat 24 van de aanwezige 27 arginine residuen gecitrullineerde waren. PAD4 citrullineert zichzelf op een tijds- en concentratie-afhankelijke manier. De activiteiten van zowel PAD4 als gecitrullineerde PAD4 werden geëvalueerd met behulp van kleine, middelgrote en grote substraten, en de verkregen resultaten toonden aan dat er geen significant verschil tussen beide soorten PAD4 kon worden aangetoond. Verder hebben we moleculaire dynamische (MD) simulaties uitgevoerd om de structurele dynamiek en bindingscapaciteiten van PAD4 en gecitrullineerd PAD4 aan hun substraat te onderzoeken. De MD-simulatie resultaten gaven aan dat autocitrullinatie de vorm van de substraat bindende pocket op PAD4 niet verandert. De berekende bindingsvrije energieën suggereerden dat PAD4 en gecitrullineerd PAD4 vergelijkbare bindingscapaciteiten vertonen met een van H3 afgeleid peptidylsubstraat. In het algemeen suggereerden onze bevindingen dat de autocitrullinatie geen zelfbeheersingsmechanisme van PAD4-activiteit vertegenwoordigt.

Hoofdstuk 5 onderzoekt de structurele integriteit van factor Va in het gif van de slang *Pseudonaja textilis*. De menselijke stollingsfactor Va (hFVa) wordt geïnactiveerd na geactiveerde proteïne C (APC)-gekatalyseerde proteolyse. In tegenstelling hiermee blijft de van *Pseudonaja textilis* afkomstige factor Va (ptFVa) cofactorfunctie ondanks APC-splitsing behouden. Op basis van de structurele analyse bevat hFVa in vergelijking met ptFVa één splitsingsplaats op positie Arg306 en één disulfide binding minder die het A2- en A3-domein covalent verbindt. Om het mechanistisch principe te ontdekken dat het functionele verschil tussen hFVa en ptFVa verklaart, hebben we experimentele chimere van ptFVa gemaakt door toevoeging van de menselijke Arg306-splitsingsplaats en de verwijdering van disulfide bindingen; echter deze chimere vertonen, na splitsing door APC, nog steeds een cofactorfunctie. Verder hebben we *in silico* de hFVa- en chimera ptFVa-structuur gemodelleerd en MD-simulaties uitgevoerd om te onderzoeken of niet-covalente interacties een rol spelen in hun structuur-bepaalde activiteiten. De resultaten die werden verkregen door computationele benaderingen toonden aan dat hFVa een neiging tot dissociatie van het A2-domein heeft langs de MD-simulatietijd die niet werd

waargenomen in chimera ptFVa. Bovendien nam de berekende vrije bindingsenergie tussen A2- en A1+A3-domeinen in hFVa significant toe bij vergelijking van de waarden op $t=0$ ns en $t=100$ ns. Ten slotte stelden *in silico*-benaderingen een cruciale lus voor die zich op het A2-domein lokaliseert en de associatie- of dissociatiestatus van het A2-domein in ptFVa of hFVa zou kunnen beïnvloeden.

Hoofdstuk 6 beschrijft het rationele ontwerp van verschillende peptiden die zich richten op de C-X-C chemokine receptor type 4 (CXCR4). Dit ontwerp is gedaan door toepassing van de workflow voor het rationeel ontwerpen van peptiden op basis van eiwit-eiwit interacties met behulp van *in silico*-benaderingen (zoals beschreven in hoofdstuk 2). Aangezien CXCR4 wordt beschreven als de belangrijkste mediator voor veroudering van neutrofielen, streven we ernaar nieuwe peptiden te ontwikkelen als therapieën die CXCR4 kunnen activeren en veroudering van neutrofielen kunnen terugdringen in patiënten met een verhoogd hoge risico op hart- en vaatziekten. We hebben twee strategieën voorgesteld door peptiden te ontwerpen waarvan wordt verondersteld dat ze intracellulair of extracellulair aan CXCR4 binden. We hebben een bekend peptidomimetic en natuurlijke ligand CXCL12 als sjabloon aangenomen via respectievelijk een intracellulaire en extracellulaire ontwerproute. Alle ontworpen peptiden werden geoptimaliseerd door middel van korte MD-simulaties en hun bindingsaffiniteiten voor CXCR4 werden geschat op basis van de berekening van de bindingsvrije energie. Tot slot hebben we in totaal 75 peptiden *in silico* ontworpen waarvan uiteindelijk drie geselecteerde peptiden zijn gesynthetiseerd en waarvan de biologische activiteit tegen CXCR4 is getest, en die activiteit vertonen. De hier gepresenteerde methoden voor peptideontwerp kunnen de benodigde tijd en financiële kosten aanzienlijk verminderen en verder worden gebruikt voor de ontwikkeling van peptidische modulators tegen andere GPCR-receptoren.

Hoofdstuk 7 geeft een kritische bespreking van de resultaten die zijn verkregen in dit proefschrift met de combinatie van computationele en experimentele benaderingen. We bespreken ook de uitdagingen en toekomstige ontwikkelingen van de beschreven computationele methodes.