

VU Research Portal

Multi-Scale Investigation of Protein-Protein Interactions

Hou, Q.

2017

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Hou, Q. (2017). *Multi-Scale Investigation of Protein-Protein Interactions*.

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 ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

Samenvatting

Eiwit-eiwit interacties zijn essentieel voor de meeste biologische activiteit. Onderzoek naar eiwit-eiwit interactie is daarom een belangrijke taak voor structurele bioinformatica. De hoeveelheid eiwit structuurgegevens, verkregen door high-throughput meettechnieken, neemt tegenwoordig snel toe. Echter, de hoeveelheden eiwitsequentie gegevens groeit nog veel sneller als gevolg van Next-Generation Sequencing (NGS)-technieken. Het detecteren van de verborgen boodschappen in deze gegevens en blootleggen van de gegevens van eiwit-eiwit interacties en daarmee de mechanismen van eiwitfunctie komen steeds meer binnen bereik als gevolg van de bovengenoemde explosie van gegevens.

In dit proefschrift hebben we een aantal instrumenten ontwikkeld en toegepast, waarmee we signalen uit verschillende soorten eiwit-gerelateerde informatie kunnen halen, om daarmee effectief eiwit-eiwit interacties te kunnen identificeren. Het proefschrift is als volgt ingedeeld:

Hoofdstuk 1 geeft een algemene inleiding op de onderzoeksvragen, daarmee zorgend voor de nodige achtergrond informatie. Identificatie van native bindingspose is een van de belangrijkste taken in de structurele bioinformatica. Meten van de vrije energie van interactie maakt het mogelijk om de sterkte van die interacties in te schatten.

In hoofdstuk 2 geven we een proefconcept van het gebruik van bindingsvrije-energie van grofkorrelige MD simulaties voor het detecteren van de native bindingspose uit duizenden dockings-oplossingen van dockings programma's. Ten opzichte van de benchmark dataset, kan onze aanpak het percentage van bijna-native bindingsoriëntaties verrijken voor het grootste deel van de eiwit complexen die hier bekeken zijn. Het grofkorrelig krachtveld stelt ons in staat om de interacties te interpreteren op een fysieke basis. Bovendien is onze methode nauwkeurig voor vele eiwitcomplexen, en vinden we bijna-native bindingsposes in onze Top 1 en Top 5 selectie. Dit geeft aan dat onze methode de sterke potentie heeft voor het 'screenen' voorafgaand aan experimenten, waarmee mogelijk hoge kosten vermeden worden, en dat het zelfs mogelijk is om te worden gebruikt voor in silico voorspelling van complete 'interactomen'.

In hoofdstuk 3 laten we zien dat sequentie specificiteit tussen interacterende en niet-interacterende homologen enig signaal draagt waarmee interactieplaatsen van sequenties voorspeld kunnen worden. Veel eiwit families die betrokken zijn bij eiwit-eiwit interacties hebben verschillende sub-families die zich binden aan verschillende partners, of die de interactie verliezen. De specificiteit van deze interacties is vaak bepalend voor de functie van de eiwitten. In dit hoofdstuk laten we zien dat de specificiteit tussen interacterende versus niet-interacterende groepen op se-

quentieniveau kan worden gebruikt voor het herkennen van interactieplaatsen. De resultaten geven aan dat het mogelijk is interactieplaatsen te voorspellen vanuit niets meer dan de sequentie en groeps-specificiteit informatie. Voor zover wij weten is dit de eerste keer is dat deze specificiteit tussen interacterende versus niet-interacterende homologen wordt gebruikt om interacterende plaatsen te voorspellen op grote schaal.

In hoofdstuk 4 ontwikkelen we een 'Random Forest' voorspeller om de interactie sites te identificeren vanuit sequentie informatie met gebruik van evolutionaire informatie, voorspelde structurele eigenschappen en de eiwitlengte als eigenschappen. Voorspellen van eiwit-eiwit interactie plaatsen vanuit sequentie-informatie wordt steeds belangrijker met de toenemende beschikbaarheid van grote hoeveelheden genoom gegevens. Conservering van 'meervoudig uitgelijnde sequenties' ('multiple sequence alignments'), voorspelde oppervlakte toegankelijkheid en voorspelde secundaire structuur werden als kenmerken gebruikt door machine leertechnieken voor het identificeren van eiwit interfaces. Anders dan de 'oude' features, laten we zien dat flexibiliteit van de eiwit-ruggengraat, afgeleid van sequentie, en de lengte van de sequentie kunnen worden gebruikt om interagerende posities te lokaliseren. Bovendien, met het combineren van al deze eigenschappen hebben we een sequentie gebaseerde interface voorspeller ontwikkeld die uitstekende prestaties bereikt op zowel homodimere als heterodimere eiwitten. Het succes van de predictie geeft aan dat onze methode de gemeenschappelijke eigenschappen van zowel homodimere als heterodimere interfaces heeft kunnen vangen.

In hoofdstuk 5 bouwen we voort op de aanpak beschreven in Hoofdstuk 4 met het creëren van een webserver genaamd SeRenDIP: 'SEquence-based Random forest predictor with IENgth and Dynamics for Interacting Proteins' – SEquentie-gebaseerde Random forest voorspeller met IENgte en Dynamica voor Interacterende Proteïnes. Dit is een gebruiksvriendelijke webinterface voor het flexibele voorspellen van eiwit-eiwitbindingsinterfaces.

Tenslotte trekken we hoofdstuk 6 enkele conclusies en bespreken we mogelijke toekomstige richtingen.

Samenvattend, in dit proefschrift onderzoeken we eiwit-eiwit interacties door middel van biofysische en bioinformatica benaderingen en valideren de bereikte voortgang met verschillende soorten eiwit gegevens. De resulterende methoden kunnen op grote schaal worden toegepast om een compleet landschap van eiwit-eiwit interactie te helpen schetsen.
