



UvA-DARE (Digital Academic Repository)

Path-metadynamics: A computational study of conformational transitions in proteins

Díaz Leines, G.

Publication date
2013

[Link to publication](#)

Citation for published version (APA):

Díaz Leines, G. (2013). *Path-metadynamics: A computational study of conformational transitions in proteins*. [Thesis, fully internal, Universiteit van Amsterdam].

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Samenvatting

Eiwitten worden beschouwd als de machinerie van het leven. Ze zijn van belang voor vrijwel alle fundamentele functies in een cel. Bijvoorbeeld perceptie en signalering, de catalyse van stofwisselings reacties, het geven van structuur, het transport van moleculen van één plaats naar een andere en de replicatie van DNA. Om deze functies uit te voeren dienen eiwitten structurele of conformationele veranderingen te ondergaan. Dit was de aanleiding dat in de laatste decennia is voorgesteld dat de werking van eiwitten bepaald wordt door het dynamisch karakter in plaats van de statische structuur. In deze benadering kunnen eiwitten een grote hoeveelheid van conformaties aftasten rond de evenwichtsstructuur als gevolg van de thermale energie. Dit beeld komt voort uit het multi-dimensionaal vrije-energy landschap dat de respectievelijke kansen bepaald van de conformaties (thermodynamica) en de energie barrières daartussen (kinetica). Een van de meest fundamentele voorbeelden van structurele veranderingen dat het dynamische gedrag van eiwitten illustreert is het vouwen van eiwitten in hun drie dimensionale structuur. Nadat eiwitten als lineaire ketens gesynthetiseerd zijn door de ribosomen, vouwen ze, voordat ze actief worden, in een specifieke drie dimensionale structuur. Indien het eiwit niet correct gevouwen wordt of gevouwen blijft, resulteert dit normaal gesproken in niet of incorrect werkende eiwitten die aanleiding kunnen geven tot verschillende neurodegeneratieve ziekten. Andere voorbeelden van conformationele veranderingen, waar het eiwit verschillende functies heeft, zijn onder andere, allosterische overgangen in enzymen, de generatie van kracht door motor eiwitten, het openen en sluiten van ionkanalen en tot slot de conformationele veranderingen die geïnduceerd worden door de ligand binding van enzymen en receptoren. Dit betekent dat eiwitten zeer dynamische objecten zijn en een gedetailleerde moleculaire beschrijving van de structurele veranderingen zou ons in staat stellen om de onderliggende mechanismes en hun functies te begrijpen en hiermee zouden we modellen met voorspellende waarde kunnen ontwikkelen die helpen in het begrijpen van bijvoorbeeld de oorsprong van de verschillende neurodegeneratieve ziekten of die ons in staat zouden stellen om liganden te ontwikkelen die de locatie van stabiele toestanden van conformationele veranderingen en hun snelheden kunnen beïnvloeden. Moleculaire dynamica (MD) simulaties zijn uitgegroeid tot een krachtig

gereedschap voor de theoretische bestudering van complexe overgangen, die inzicht geven in de stabiliteit en de dynamica van vele moleculaire systemen. MD simulaties kunnen in het bijzonder een aanvulling zijn op experimenten door de dynamische tijdsevolutie van biomoleculaire systemen op atomair niveau te modelleren. Hoewel traditionele MD nu milliseconde tijdschalen kan modelleren, vinden veel conformationele veranderingen plaats op tijdschalen die zeer lang zijn in vergelijking met de tijdschalen die toegankelijk zijn voor traditionele MD. Hierdoor is de simulatie van de kinetica van het vouwen en van grote conformationele veranderingen in eiwitten beperkt gebleven tot relatieve kleine systemen en korte tijdschalen. Deze zogenaamde rare event processen worden vaak geassocieerd met hoge vrije-energie barrières tussen metastabiele toestanden. Dit verhindert zijn beurt het effectief sampelen van de reactie paden met traditionele MD. Het bestuderen van rare events vereist de ontwikkeling van speciale rare event methodologieën die toestaan om grote vrije-energie barrières te passeren en zo complexe overgangen te bestuderen. Het doel van dit proefschrift is de ontwikkeling en toepassing van een nieuwe rare event methode, genaamd padmetadynamica, om vrije-energie landschappen te onderzoeken en om transitie paden tussen metastabiele toestanden te vinden. Padmetadynamica is een alternatieve benadering om de mechanica van dynamische processen te begrijpen die plaats vinden op tijdschalen die zeer lang zijn in vergelijking met moleculaire tijdschalen. Hoewel er zeer veel voorbeelden zijn van rare event processen, zoals conformationele veranderingen in biomoleculen, nucleatie gebeurtenissen in fase overgangen en chemische reacties, zijn we in dit proefschrift vooral geïnteresseerd in het bestuderen van het vouwen en ontvouwen van eiwitten en hebben we padmetadynamica toegepast om het vrije-energie landschap van verschillende eiwitten te onderzoeken.

Padmetadynamica breidt de bestaande metadynamica methode uit met nieuwe algoritmes om: 1 Het systeem een bias te geven langs een geparmetriseerde kromme, die een functie is van andere relevante order parameters die ook wel collectieve variabelen worden genoemd en 2. om on-the-fly deze kromme te evolueren naar het meest waarschijnlijke reactie pad. Dit wordt gedaan door een nieuwe variabele te definiëren die een functie is van alle andere collectieve variabelen en die een schatting is van het reactie pad op het vrije-energie oppervlak. Door de dynamica langs deze variabele te sturen met een groeiende bias potentiaal van metadynamica evolueert het pad efficiënt naar het lokaal meest waarschijnlijke reactie pad, terwijl tegelijkertijd de bias potentiaal een goede schatting wordt van het vrije-energie profiel langs het geconvergeerde pad. De computer tijd van deze gecombineerde methode hangt niet langer af van het aantal collectieve variabelen. Vooral deze laatste eigenschap stelt ons in staat om vrij complexe overgangen te bestuderen die meer collectieve variabelen nodig hebben dan een standaard metadynamica simulatie aan kan. De computer tijd van standaard metadynamica simulaties schalen namelijk exponentieel met het aantal collectieve variabelen, waardoor er hoogstens 3 of 4 collectieve variabelen gebruikt kunnen worden. In het eerste gedeelte van dit proefschrift, hoofdstuk 3, presenteren we de padmetadynamica methode en tonen hoe deze gebruikt kan worden om in een

simulatie gelijktijdig het reactie mechanisme en het vrije-energie profiel te vinden tussen twee stabiele toestanden. Ik illustreer de methode met een toepassing op een klassieke moleculaire dynamica simulatie op een conformationele overgang in alanine-dipeptide.

In hoofdstuk 4 onderzoeken we in meer detail de theorie en efficiëntie van padmetadynamica. We bereiden de methode uit met een extra voortgangs parameter met als span de hele ruimte van de collectieve variabelen in de verzameling van hypervlakken loodrecht op het pad. Deze variabele representeert de transitie buis tussen twee stabiele toestanden. Door een buis potentiaal toe te passen op deze variable tonen we dat de padmetadynamica methode twee bestaande limieten heeft: het gemiddelde transitie pad (de gemiddelde plaatsing ook wel eindige temperatuur limiet) en het minimum vrije-energie pad (gradient methode of de limiet van geen temperatuur). Bovendien geven we een voorschrift om de padmetadynamica methode te gebruiken. Dit voorschrift bestaat uit verschillende tests om parameters te optimaliseren zodat het pad en het vrije-energie profiel convergeren.

Het tweede gedeelte van dit werk, hoofdstukken 5 en 6 tonen de toepassing van padmetadynamica op het bestuderen van conformationele veranderingen in eiwitten. In hoofdstuk 5 tonen we dat padmetadynamica succesvol kan worden toegepast om het ontvouwings proces te bestuderen van een relevant fotoreceptor, het Photoactive Yellow Protein (PYP), wat in de order van milliseconden duurt. Hier vind de methode niet alleen het vrije-energie profiel langs het reactie pad van het ontvouwen, maar het geeft ook uniek moleculair inzicht in de verkregen reactie paden. De resultaten geven een aanvullend beeld op de conclusie van Vreede et. al. over het bestaan van tussengelegen metastabiele toestanden in het ontvouwings proces van de PYP en de relevante collectieve variabelen in de beschrijving van de overgangen. Bovendien vinden we dat de hoogste barriere van het proces ligt bij het blootstellen aan het oplossingsmiddel van het Glu46 en dat de relevante collectieve variabelen een belangrijke rol toekennen aan het residu ASN43 in de overgang naar de signaaltoestand van het eiwit. We hebben de padmetadynamica methode in de ontwikkel versie van het pakket PLUMED geïmplementeerd. PLUMED is een plugin voor vrije-energie berkeningen zoals metadynamica en wordt gebruikt in combinatie met een moleculair dynamica simulatie pakket, zoals bijvoorbeeld Gromacs, wat we hebben gebruikt voor onze simulaties, en het stelt ons in staat om de mechanismes en reactie snelheden te bepalen als een functie van alle relevante collectieve variabelen die het ontvouwing proces van het PYP te beschrijven. Bovendien stellen we een strategie voor om met padmetadynamica, in combinatie met een opeenvolgende analyse van de vrije-energie profielen, relevante collectieve variabelen te vinden en om nauwkeurige barrieres te schatten.

In hoofdstuk 6 ronden we het proefschrift af met de toepassing van padmetadynamica op het voorspellen van het formatie/dissociatie mechanisme van een van de meest bestudeerde opgerolde spoel in globulaire eiwitten het leucine zipper eiwit domein factor van de transcriptie factor GNC4 in gist. De opgerolde spoel is de motivatie

geweest voor vele studies naar de fundamentele relatie tussen aminozuur sequenties en het vouwen van eiwitten. We hebben een aantal relevante vrijheidsgraden geïdentificeerd die deelnemen aan de vorming van het complex. Onze resultaten tonen dat de transcriptie niet langs één robuust pad plaats vindt, maar met heterogeniteit van de overgangstoestanden. Bovendien wijzen de vrije-energie profielen, die verkregen zijn langs het gemiddelde reactie pad, uit dat het meest waarschijnlijke mechanisme gebeurd door een intermediair dat gekarakteriseerd wordt door de dissociatie van de N-terminal en het gedeeltelijk verlies van de helicische structuur van het dimeer. Experimentele studies hebben de sterke stabiliteit van de C-terminal bevestigd en hebben een waarschijnlijk pad gesuggereerd dat door deze tussen gelegen toestand leidt. Tot slot merken we op dat de padmetadynamica methode een weg opent om rare event transities te bestuderen op hoog dimensionale vrije-energie landschappen. De computer tijd die de simulaties in dit proefschrift kostten waren ordes van grootte kleiner dan met traditionele MD. Dit terwijl de verzameling van collectieve variabelen die gebruikt zijn om de transities te beschrijven een orde groter waren dan die gebruikt zijn bij de traditionele metadynamica benadering. Dit is een stap vooruit in het bestuderen van complexe rare event transities. We benadrukken dat hoewel de methode de selectie van een hoog dimensionale ruimte van collectieve variabelen toelaat, het vinden van een voorschrift voor het a priori kiezen van de relevantie reactie coördinaten de grootste uitdaging blijft in het bestuderen van complexe rare event transities. Dit blijft een van de grootste beperkingen van padmetadynamica. Bij ons weten bestaat er op dit moment geen exact voorschrift voor het bepalen van relevante reactie coördinaten, ook al kunnen methodes zoals likelihood maximization analysis in combinatie met methodes als transition path sampling licht werpen op de initiële selectie van de verzameling van collectieve variabelen. Het is belangrijk om op te merken dat hoewel padmetadynamica het probleem van een a priori selectie van collectieve variabelen niet op kan lossen, het mogelijk is om een strategie te gebruiken om te testen of de verzameling van collectieve variabelen voldoende is. Deze strategie werkt op basis van de één dimensionale vrije-energie die verkregen wordt uit een intrinsiek multi-dimensionale gebeurtenis. Ondanks deze beperking heeft het padmetadynamica algoritme bewezen dat het lokaal waarschijnlijke transitie paden op hoog dimensionale landschappen redelijk robuust kan bestuderen. Het is belangrijk om op te merken dat hoewel het doel van dit proefschrift het bestuderen van conformationele transities in eiwitten was, padmetadynamica ook toegepast kan worden op een breed spectrum aan disciplines, van materiaalkunde tot biophysica, overal waar reactieve transities op lange tijdschalen plaats kunnen vinden. Hoewel de huidige versie van het padmetadynamica algoritme lokaal is in de zin dat het beperkt is tot de verkenning van één enkele reactie buis tussen twee stabiele toestanden, zou de methode in de toekomst uitgebreid kunnen worden met een gestuurde verwisseling van paden in verschillende reactie buizen, wat de verkenning van transities met verschillende overgangstoestanden mogelijk zou maken. Dit voor het geval dat er geen unieke oplossing is voor het gemiddelde transitie pad. Deze uitbereiding zou het verkennen

van nog complexere vrije-energie landschappen mogelijk maken en een optie zijn om grotere eiwitten te bestuderen.