



Universiteit
Leiden
The Netherlands

Single-molecule fluorescence in sequence space

Severins, I.W.H.

Citation

Severins, I. W. H. (2024, May 8). *Single-molecule fluorescence in sequence space*. Retrieved from <https://hdl.handle.net/1887/3753605>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3753605>

Note: To cite this publication please use the final published version (if applicable).

Samenvatting

Het ontwerpen en aanpassen van biologische systemen, zoals het creëren van nieuwe medicijnen om ziekten te genezen en het maken van enzymen om chemische processen te versnellen, vereist dat we de fysische processen begrijpen die ten grondslag liggen aan de biologie. Deze processen zijn uitgebreid bestudeerd gedurende vele decennia en op de schaal van hele ecosystemen, tot organismen, organen, weefsels en cellen; maar uiteindelijk komt het gedrag op al deze schalen voort uit de eigenschappen van, en interacties tussen, individuele biomoleculen. Een groot deel van deze biomoleculen - DNA, RNA en eiwitten - bestaat uit een specifieke sequentie van bouwstenen die de moleculaire structuur en functie bepaalt. Om het gedrag van deze moleculen te begrijpen moeten we dus weten wat de invloed is van de sequentie.

De relatie tussen sequentie en functie wordt gewoonlijk onderzocht met bulk studies, waar een groot aantal moleculen gezamenlijk wordt bestudeerd. In eerste instantie werd dit alleen gedaan met een sequentiebibliotheek in oplossing. Hierbij werden sequenties met bepaalde eigenschappen geselecteerd en geïdentificeerd, bijvoorbeeld de DNA sequenties die aan een specifiek eiwit binden. Deze experimenten leverden echter vooral kwalitatieve informatie over de sequentie-specifieke eigenschappen. Met de introductie van een nieuwe generatie van sequentiebepalingstechnieken werd het echter mogelijk om meer kwantitatieve informatie te verkrijgen. De sequentiebepaler gebruikt een sequentiebibliotheek om gescheiden clusters van DNA moleculen te creëren op het oppervlak van een microscoop glas. Door het bepalen van de sequentie wordt de identiteit van elk cluster onthuld, wat het mogelijk maakt om de clusters te gebruiken voor biochemische en biofysische studies. Deze studies zijn toegepast om eiwit-DNA interacties, zoals die van het CRISPR-Cas eiwit complex, te bestuderen. Het DNA van elk cluster kan daarnaast door middel van transcriptie worden omgezet in RNA, en door middel van translatie in eiwitten. Daardoor is het ook mogelijk om de invloed van RNA sequentie op RNA- of eiwitbinding te onderzoeken, en zelfs om de invloed van eiwitsequentie op enzymatische reacties te bestuderen. De kwantitatieve sequentie-afhankelijke data verkregen met deze technieken heeft ons nieuwe kennis gegeven over moleculaire mechanismen en maakt het mogelijk om thermodynamische modellen te bouwen en te verifiëren.

Ondanks de interessante eigenschappen die hiermee zijn blootgelegd, leveren de grootschalige studies op cluster niveau slechts gelimiteerde informatie over de moleculaire kinetiek. Variaties binnen populaties en over de tijd zijn bijvoorbeeld niet waar te nemen. Ook zijn zulke studies vaak gelimiteerd tot simpele systemen met slechts twee toestanden. Deze beperkingen kunnen worden verholpen door moleculen apart te bestuderen, door middel van fluorescentie experimenten op individuele moleculen. Deze "single-molecule" technieken maken het mogelijk om complexe dynamische processen met meerdere toestanden zichtbaar te maken. Tot nu toe kan echter slechts een klein aantal sequenties worden onderzocht, aangezien elke sequentie apart moet worden verkregen en gemeten, wat kostbaar en tijdrovend is. Voor het uitvoeren van zulke experimenten is een parallelle aanpak dus essentieel. Daarom hebben we een nieuwe techniek, genaamd SPARXS, ontwikkeld om parallelle analyse van individuele moleculen uit te voeren voor snelle verkenning van de sequentieruimte. SPARXS maakt het mogelijk om duizenden sequenties op moleculair niveau te karakteriseren door fluorescentie experimenten op individuele moleculen te koppelen aan sequentiebepaling.

SPARXS begint met het immobiliseren van een fluorescent-gelabelde DNA-bibliotheek op het oppervlak van een commercieel verkrijgbare Illumina sequentiebepalings-chip. Vervolgens wordt een experiment uitgevoerd op individuele moleculen door een groot oppervlak te scannen met een geautomatiseerde fluorescentiemicroscoop. De chip wordt daarna overgebracht naar de sequentiebepaler, waar de individuele moleculen worden vermenigvuldigd tot clusters, waarvan uiteindelijk de sequenties worden bepaald. De coördinatenstelsels van de fluorescentie- en

sequentiedatasets worden daarna uitgelijnd, zodat de kinetiek van de moleculen kan worden gekoppeld aan hun sequenties. Het eindproduct is een kinetisch of energetisch landschap in sequentieruimte.

Het concept om sequentiebepaling te combineren met experimenten op individuele moleculen lijkt simpel, maar de implementatie bracht verschillende uitdagingen met zich mee. Bij de sequentiebepaling zijn er bijvoorbeeld verschillende criteria voor de homogeniteit en de samenstelling van de DNA-bibliotheek. Daarnaast waren er voor de fluorescentie experimenten extra stappen nodig om de van nature aanwezige achtergrond op de sequentiebepalings-chips te verwijderen. Voor succesvolle sequentiebepaling na deze experimenten was het verder nodig om verschillende stappen in het standaard protocol van de sequentiebepaler aan te passen. Ook het uitlijnen van de twee datasets was ingewikkeld door de grote hoeveelheid datapunten en het grote aantal ontbrekende sequenties. Als oplossing hebben we een driestapsaanpak ontwikkeld, waarbij eerst een globale uitlijning van de instrumenten wordt gedaan met een geometrisch hash-algoritme, gevolgd door een experiment-specifieke uitlijning op basis van een kruiscorrelatie-algoritme en, als laatste, een precieze uitlijning voor elke afbeelding met een kernelcorrelatie-algoritme. Daarnaast hebben we een methode ontwikkeld om de afstandsrempel te bepalen voor het koppelen van kinetische en sequentiedata.

Om de bruikbaarheid en effectiviteit van SPARXS te demonstreren hebben we onze techniek toegepast op het Holliday DNA-kruispunt, een structuur waar vier DNA strengen samenkomen en die onder meer voorkomt tijdens homologe recombinatie. Deze structuur wisselt tussen twee toestanden, waarbij de snelheid van de wisselingen afhangt van de sequentie in het centrum. Door twee armen van het kruispunt te voorzien van fluoroforen kan met behulp van Förster resonantie energie overdracht (FRET) de kinetiek van de toestandswisselingen zichtbaar worden gemaakt. Op deze manier hebben we het dynamische gedrag van miljoenen Holliday-kruispunten met duizenden afzonderlijke sequenties gemeten, een resultaat dat zonder SPARXS onmogelijk zou zijn geweest. Hiermee kon een verscheidenheid aan kwantitatieve parameters worden bepaald: de fractie moleculen met dynamisch gedrag, de FRET efficiëntie voor de waargenomen toestanden, en de evenwichtsconstante en transitiesnelheden voor moleculen met meerdere toestanden. De verkregen transitiesnelheden konden worden gecorreleerd met de purine-pyrimidine verdeling en het GC-gehalte in het centrum van de structuur, en ook met de theoretische energieën voor nucleotide stapeling. Eenzelfde correlatie was echter niet zichtbaar voor de evenwichtsconstante, wat wijst op de invloed van andere structurele componenten die waarschijnlijk afhankelijk zijn van de identiteiten van specifieke baseparen.

Naast de experimenten op DNA kan SPARXS ook worden toegepast op andere biomoleculen. Door gebruik te maken van transcriptie en translatie zouden de invloed van RNA- en eiwitsequentie ook met SPARXS kunnen worden bestudeerd op het niveau van individuele moleculen. Naast intramoleculaire structuren kunnen ook intermoleculaire binding en chemische reacties worden onderzocht door DNA, RNA, eiwitten of andere biomoleculen in de oplossing te introduceren. Het brede toepassingsgebied, de uitgebreide datasets en het grote aantal moleculen tonen de bruikbaarheid van deze techniek. SPARXS opent dus een compleet nieuwe en tot nu toe onverkende onderzoeksrichting, die kan leiden tot nieuw inzichten in de werking van biologische systemen en mogelijk kan bijdragen aan de volgende stappen om de biologie verder naar onze hand te zetten.