



Universiteit  
Leiden  
The Netherlands

## **Biophysics of disordered nuclear receptors and their DNA binding regulation**

Heling, L.W.H.J.

### **Citation**

Heling, L. W. H. J. (2026, June 25). *Biophysics of disordered nuclear receptors and their DNA binding regulation*. Retrieved from <https://hdl.handle.net/1887/4306978>

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/1887/4306978>

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

## Nederlandse samenvatting

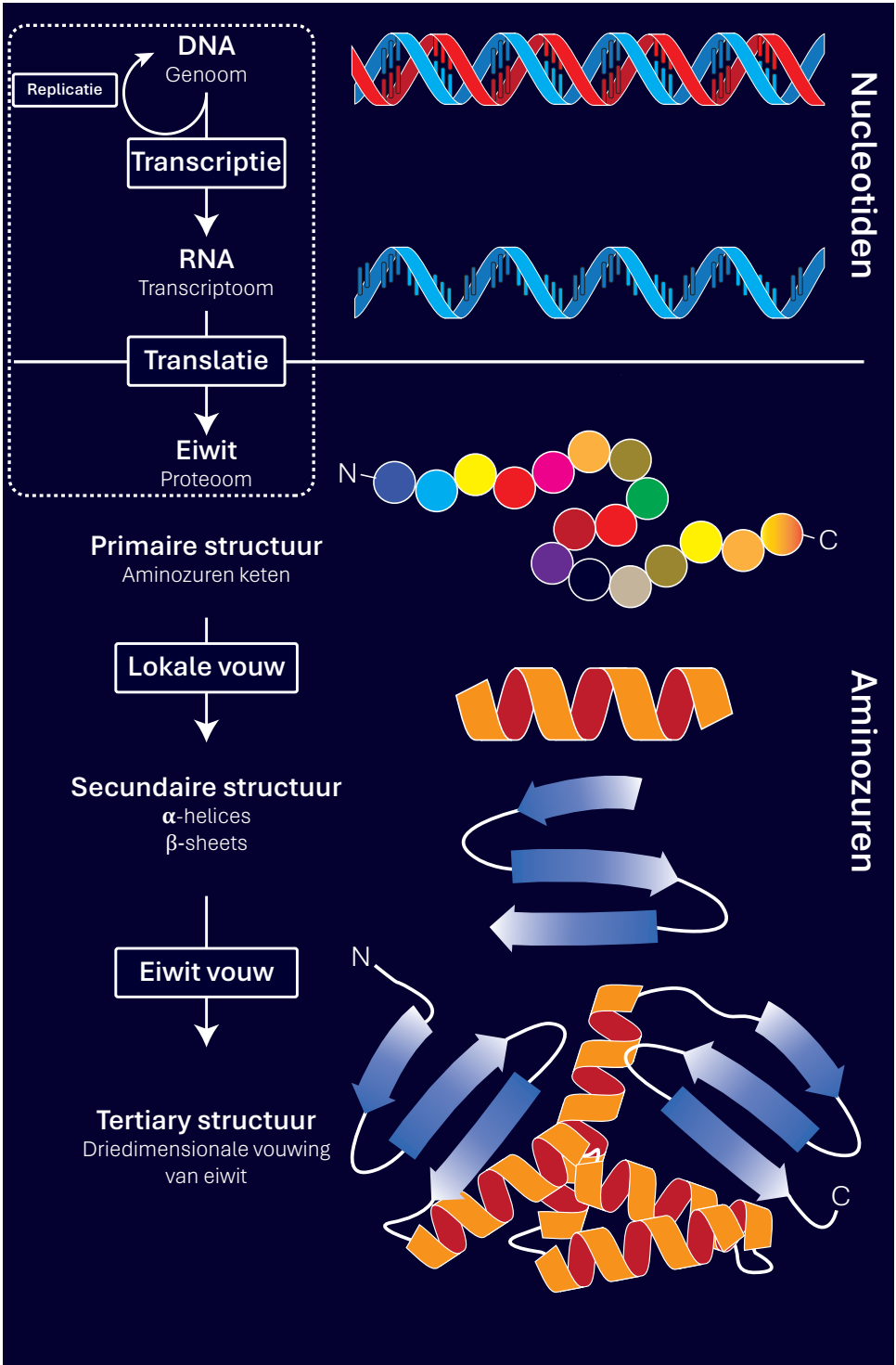
Om te overleven, moeten alle cellen—van eencellige organismen tot de cellen in ons eigen lichaam—complexe vitale functies uitvoeren. Van celdeling, DNA-replicatie en metabolisme tot celcommunicatie en genexpressie: al deze essentiële processen in de cel worden aangedreven door eiwitten, die daarom vaak worden aangeduid als moleculaire machines. Door deze fundamentele rol zijn eiwitten tevens cruciaal in ziekteprocessen. Een diepgaand begrip van eiwitfunctie is dus niet enkel belangrijk om de basisprincipes van de biologie te begrijpen, maar vormt ook de basis voor de ontwikkeling van nieuwe medicijnen en behandelmethoden.

Decennialang werd de moleculaire biologie gedomineerd door centrale hypothesen omtrent de stroom van genetische informatie en eiwitvouwing (**Box 1**). Samen vormden deze klassieke postulaten het structuur–functie denkmodel, dat stelt dat de aminozuursequentie van een eiwit opvouwt tot een stabiele, drie-dimensionale conformatie die absoluut vereist is voor biologische activiteit. Hoewel dit model decennialang stand heeft gehouden en accuraat blijft voor een aanzienlijk deel van het proteoom, wordt het inmiddels fundamenteel uitgedaagd door structurele waarnemingen die niet binnen dit klassieke kader passen. Intrinsiek ongeordende eiwitten (intrinsically disordered proteins; IDPs) en eiwitregio's (intrinsically disordered regions; IDRs) (**Box 2**) tonen aan dat een gedefinieerde 3D-structuur geen universeel vereiste is. Deze domeinen kenmerken zich juist door een uitgesproken conformationele flexibiliteit en bestaan onder fysiologische omstandigheden als een dynamisch ensemble van metastabiele toestanden. Het is inmiddels duidelijk dat deze eiwitten, juist dóór deze structurele plasticiteit, onmisbare functies vervullen in complexe cellulaire processen zoals signaaltransductie en genregulatie.

### Box 1. Klassieke postulaten van de moleculaire biologie

De functie van verschillende componenten in de cel, van eiwitten en polysachariden tot nucleïnezuuren, was decennialang onderworpen aan heftig wetenschappelijk debat. In 1953 publiceerden Francis Crick en James Watson, op basis van röntgendiffractiefoto's van Rosalind Franklin, hun model voor de dubbele helixstructuur van DNA, en later dat jaar hun direct daaropvolgende theorie over genen en genetisch materiaal. Dat DNA de erfelijkheid bepaalt, was een aantal jaar daarvoor al verondersteld. De vraag bleef echter wat de exacte functie van genen is, en hoe deze informatie in het DNA ligt opgeslagen. In een lezing in 1957 stelde diezelfde Crick, als theoretisch moleculair bioloog, dat "informatie" in deze context simpelweg neerkwam op "de bepaling van een reeks eenheden", oftewel de exacte nucleotidenvolgorde in het DNA. Dit inzicht stelde Crick in staat om de link tussen genen en eiwitten te conceptualiseren. Hij noemde deze link de "stroom van informatie".

Tijdens diezelfde lezing introduceerde hij de theorie die deze stroom beschrijft: het "Centrale Dogma" (gearceerd met een stippellijn in de figuur). Het fundamentele principe hiervan is dat deze informatiestroom strikt directioneel is. Informatie kan worden gekopieerd van DNA naar DNA (replicatie), of worden overgeschreven van DNA via RNA (transcriptie) naar eiwitten



(translatie), maar nooit in de omgekeerde richting. Zoals Crick stelde: "zodra informatie in een eiwit is overgegaan, kan het er niet meer uit." Kortom, DNA fungeert als de blauwdruk voor de synthese van eiwitten, waarbij veranderingen in het resulterende eiwit de originele genetische code niet met terugwerkende kracht kunnen wijzigen. Crick gaf later toe spijt te hebben van zijn keuze voor het woord "dogma". Hij was zich destijds niet volledig bewust van de absolute, haast religieuze betekenis van die term en gaf aan dat hij achteraf gezien liever het woord "hypothese" had gebruikt.

Een lineaire polypeptideketen is doorgaans pas biologisch actief wanneer deze een specifieke driedimensionale (3D) conformatie aanneemt. Al in de vroege jaren '50 pionierde Linus Pauling in dit veld door aan te tonen dat eiwitketens zich lokaal opvouwen in fundamentele, rigide structuren zoals de alfa-helix en de bèta-plaat. Voortbouwend op de aanname dat deze ketens zich op specifieke manieren vouwen, stelde het latere dogma van Anfinsen (jaren '60) dat de uiteindelijke, unieke 3D-structuur van een eiwit volledig en autonoom wordt bepaald door de exacte opeenvolging van zijn aminozuren. Dit vormde decennialang het sluitstuk van de moleculaire biologie: DNA levert de lineaire code, RNA vertaalt dit naar een aminozuurketen, en deze keten vouwt zich op tot een vaste, functionele 3D-structuur. Samen leidden het werk van Pauling en Anfinsen tot het klassieke denkmodel dat een welgedefinieerde, rigide structuur een absolute vereiste is voor biologische functie.

Dit proefschrift bestudeert de biofysica van intrinsiek ongeordende kernreceptoren en de functionele implicaties hiervan voor de regulatie van hun DNA-interacties. Kernreceptoren zijn transcriptiefactoren die zich, na activering door ligandbinding of post-translationele modificaties, verplaatsen naar de celkern waar ze aan het chromatine (het complex van DNA en eiwitten) binden en co-regulators rekruteren om genexpressie te sturen. De menselijke kernreceptor superfamilie deelt een geconserveerde moleculaire architectuur met gestructureerde DNA- en ligandbindende domeinen. Daarnaast bezitten zij echter een uitgesproken en sterk variabel intrinsiek ongeordend N-terminaal domein (NTD), waarvan de dynamiek en functie in dit onderzoek centraal staan.

In **Hoofdstuk 1** wordt de kenniskloof verkend door de gevestigde principes van intrinsieke wanorde te bespreken binnen de context van de kernreceptor superfamilie en stellen we het centrale thema vast: dat de inherente flexibiliteit van het NTD fungeert als een essentieel reguleringsmechanisme. Het ontrafelen van de precieze werking van deze domeinen wordt echter ernstig belemmerd door de beperkingen van klassieke structuur- en analysemethoden. Zowel technieken die primair ontworpen zijn voor rigide eiwitten, als biochemische proeven die afhankelijk zijn van bulkmetingen (**Box 3**), schieten tekort in het vastleggen van de uiterst dynamische, kortstondige conformaties die IDR's kenmerken. Zelfs recente doorbraken in kunstmatige intelligentie (zoals AlphaFold) bieden hier nog geen volledige uitkomst, aangezien de voorspelbaarheid van structuur van dergelijke flexibele ensembles fundamenteel beperkt is. Om deze methodologische kloof te overbruggen, combineert dit proefschrift geavanceerde moleculaire modellering met single-moleculaire biofysica (**Box 3**). Deze vernieuwende aanpak wordt in de daaropvolgende hoofdstukken toegepast op twee belangrijke kernreceptoren: de androgeenreceptor (AR) en de orphan receptor Nur77.

## Box 2. Intrinsiek ongeordende eiwitten

Hoewel intrinsiek ongeordende eiwitten (IDP's) en regio's (IDR's) een vaste vouwing missen, zijn ze alomtegenwoordig; in eukaryote proteomen bevindt 30 tot 40% van alle aminozuren zich in een ongeordende staat. Omdat deze eiwitketens niet in een rigide structuur zijn gefixeerd, bezitten zij een hoge conformationele entropie: een biofysische eigenschap die resulteert in een uitgebreide structurele vrijheid en de continue fluctuatie tussen diverse conformaties. De werking van deze eiwitten en domeinen volgt het principe van de sequence-ensemble-function relatie: de aminozuurvolgorde bepaalt het ensemble, en het ensemble bepaalt de uiteindelijke functie. Specifieke eigenschappen van dit ensemble—zoals hoe compact het is en of er kortdurende lokale structuren ontstaan—worden direct gestuurd door de chemische samenstelling van de eiwitketen. Vooral de verdeling van elektrische ladingen en aromatische bouwstenen speelt hierbij een doorslaggevende rol.

Deze hoge conformationele entropie biedt fundamentele voordelen voor moleculaire communicatie. Het stelt IDP's en IDR's in staat om een aanzienlijk groter cellulair volume ruimtelijk af te tasten, wat leidt tot sterk verhoogde associatiesnelheden met doelwitmoleculen. Bovendien stelt deze structurele plasticiteit hen in staat om effectief te interacteren met een grote diversiteit aan bindingspartners. Hierdoor breiden zij het cellulaire repertoire aan moleculaire interacties substantieel uit via een continuüm aan interactiemodi. Dit spectrum varieert van coupled folding and binding—waarbij de ongeordende keten (deels) opvouwt tijdens de interactie met een doelwit—tot fuzzy complexes, waarbij aanzienlijke structurele heterogeniteit in de gebonden toestand behouden blijft. In specifieke gevallen binden sterk tegengesteld geladen IDP's of IDR's met picomolaire affiniteit aan elkaar om een structureel ongedefinieerd complex te vormen, waarbij de ketenflexibiliteit volledig intact blijft.

Ondanks deze uitgesproken flexibiliteit is deze vorm van moleculaire herkenning uiterst doelgericht. Deze specificiteit wordt overwegend gerealiseerd via Short Linear Motifs (SLiMs): korte, geconserveerde aminozuursequenties. De integratie van meerdere SLiMs binnen één ongeordende keten resulteert in multivalentie, wat een hoge bindingspecificiteit combineert met de snelle associatie-dissociatie kinetiek die vereist is voor adequate signaaltransductie. Bovendien zijn deze dynamische ensembles in hoge mate responsief op hun biochemische micro-omgeving; fysiologische fluctuaties (zoals veranderingen in pH en zoutconcentratie) of de eerder genoemde post-translationele modificaties kunnen het ensemble structureel herprogrammeren. Hierdoor functioneren IDP's en IDR's niet alleen als passieve structurele linkers, maar treden zij op als responsieve sensoren van de cellulaire omgeving en als fundamentele regulatoren binnen genexpressienetwerken en de formatie van biomoleculaire condensaten.

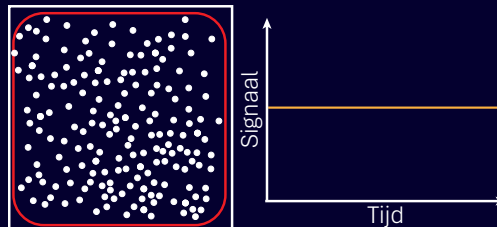
In **Hoofdstuk 2** worden deze methoden toegepast op een klinisch relevant ziektemodel. Hier presenteren we het eerste gedetailleerde moleculaire model van de polyglutamine (pQ)-expansie binnen het AR-NTD, de mutatie die verantwoordelijk is voor de neurodegeneratieve aandoening spinobulbaire musculaire atrofie (SBMA). Onze bevindingen onthullen dat de pQ-expansie een dynamische misvouwing van het NTD induceert, wat de globale structuur verandert en de delicate balans van interacties met partnereiwitten verstoort. Dit werk biedt de eerste verklaring op atomair niveau van hoe een mutatie in een ongeordende regio het oligomerisatieproces verstoort en leidt tot toxische eiwitaggregatie. Hiermee wordt de

structurele dynamiek van de receptor direct gekoppeld aan het pathologische fenotype. Onze eerdere ontdekking dat het AR-NTD uit afzonderlijke dynamische subregio's bestaat, riep de vraag op of deze domeinen ook regiospecifieke functies bezitten. In **Hoofdstuk 3** onderzoeken we hoe deze subregio's—de N-terminale regio (NR) en C-terminale regio (CR)—het vermogen van de receptor om DNA te binden reguleren. Met behulp van technieken om deze interacties in real-time te visualiseren op een enkel DNA-molecuul, onthullen we een complexe regulerende wisselwerking. Onze bevindingen tonen aan dat deze twee ongeordende regio's functioneren als een modulair controlesysteem met differentiële en synergetische effecten: de CR fungeert als een remmer die de bindingsaffiniteit van AR voor DNA vermindert, terwijl de NR functioneert als een kinetische modulator die de dissociatiesnelheid van de receptor versnelt. Gezamenlijk verfijnen zij op uiterst dynamische wijze de interactie van de receptor met het DNA.

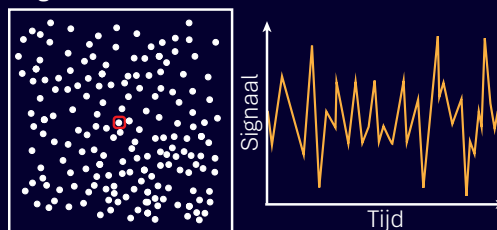
### Box 3. Bulk versus single-molecule biofysica

Traditioneel worden moleculaire processen bestudeerd via 'bulk'-experimenten, waarbij signalen van miljoenen moleculen tegelijkertijd worden gemeten. Dit levert per definitie een signaal op dat het gemiddelde van het ensemble representeert. Hoewel deze benadering een optimale signaal-ruisverhouding oplevert, maskeert het structurele heterogeniteit en kortdurende kinetische toestanden. Zo toont een klassieke bulkmeting weliswaar succesvol een gemiddelde eiwitstructuur aan, of óf een eiwit aan DNA bindt, maar blijven de onderliggende, vaak van korte duur zijnde processen die de daadwerkelijke conformationele en bindingsdynamiek bepalen in zo'n statische meting volledig onzichtbaar omdat de individuele moleculen in het ensemble niet synchroon handelen.

#### Bulk technieken



#### Single-molecule technieken



Om deze verborgen conformationele dynamiek bloot te leggen, is een overstap naar "single-molecule" technieken noodzakelijk. Door het observatievolume te verkleinen tot een individueel molecuul, wordt de ensemble-middeling omzeild. Door één enkel molecuul langdurig te volgen, ontstaat een veel gedetailleerder beeld van dynamisch gedrag. Het

cruciale voordeel hiervan is dat kortstondige tussenfasen niet langer in de massa verdwijnen, maar daadwerkelijk als afzonderlijke stappen in de tijd zichtbaar worden. Het single-molecule veld omvat zowel experimentele biofysische methoden als computationele single-molecule technieken, zoals Molecular Dynamics (MD) simulaties. Waar een bulkmeting slechts een statische eindstaat toont, zoals een gevouwen conformatie of een stabiel gebonden complex, onthullen deze experimentele en computationele single-molecule technieken de daadwerkelijke kinetische paden in de tijd. Ze visualiseren bijvoorbeeld niet alleen de stapsgewijze opvouwing van een flexibel eiwitdomein, maar ook de complexe dynamiek waarmee transcriptiefactoren hun sequentie op het DNA lokaliseren.

Aan single-molecule technieken kleven ook praktische uitdagingen. De focus op slechts één molecuul levert inherent een veel zwakker signaal op, wat resulteert in relatief veel ruis en sterke fluctuaties in de meting. Dit vereist zeer gevoelige detectieapparatuur en maakt de experimenten tijdrovend. Om te garanderen dat het waargenomen gedrag representatief is voor de gehele populatie, moeten alsnog tientallen tot honderden individuele moleculen, of lange computationele simulatietijden, geanalyseerd worden om betrouwbare statistiek op te bouwen. Tot slot bepaalt de gekozen methode deels welke details zichtbaar worden: de ene techniek pikt bijvoorbeeld vooral globale vormveranderingen op, terwijl de andere juist lokale interacties vastlegt. De integratie van dergelijke gedetailleerde single-molecule data met robuuste bulkmetingen vormt daarom vaak de krachtigste strategie in modern biofysisch onderzoek.

Nadat de principes van IDR-gemedieerde regulatie in een goed bestudeerde receptor waren vastgesteld, verbreedt dit proefschrift de focus om te testen of deze principes eveneens van toepassing zijn op minder goed begrepen receptoren. **Hoofdstuk 4** onderzoekt de regulerende wisselwerking tussen de weesreceptor Nur77 en de glucocorticoïde receptor (GR). Onze single-molecule experimenten onthullen dat het Nur77 op het DNA stabiliseert in het bijzijn van GR, waardoor het er langer op verblijft. Opvallend is dat we ook ontdekten dat dexamethason, een veelgebruikt synthetisch glucocorticoïde, het vermogen van Nur77 om DNA te binden direct verstoort, geheel onafhankelijk van GR. Deze studie onthult een mogelijk duaal reguleringsmechanisme—één gemedieerd door het GR-eiwit en een ander direct door de steroïde ligand zelf—wat een nieuwe laag van complexiteit toevoegt aan hun antagonistische relatie.

De onverwachte bevinding dat een klein molecuul de DNA interacties van Nur77 direct kan beïnvloeden, was de aanzet tot het onderzoek in **Hoofdstuk 5**. Hier bestuderen we hoe Cytosporone B (Csn-B), een molecuul waarvan bekend is dat het aan Nur77 kan binden, de DNA-bindingsactiviteit moduleert. Onze analyse onthult dat Csn-B fungeert als een allosterische modulator die het DNA-zoekmechanisme van Nur77 fundamenteel herprogrammeert. Het verbetert de specificiteit van Nur77 voor zijn dimere DNA-motief, terwijl het de affiniteit voor niet-specifiek DNA en monomere motieven vermindert. Dit suggereert een mechanistisch model waarbij een molecuul een conformationele verandering in een weesreceptor induceert, wat leidt tot een specifieke oligomere toestand en selectiviteit voor specifieke DNA-motieven.

Tot slot wordt het werk in dit proefschrift in **Hoofdstuk 6** samengevat en wordt de vertaalslag gemaakt van de precieze, reductionistische modellen uit de voorgaande hoofdstukken naar de complexe omgeving van de levende cel. Het bespreekt hoe de fundamentele biofysische eigenschappen die in dit proefschrift zijn blootgelegd—de dynamische (mis)vouwing, de allosterische modulatie en de veranderde kinetiek—een cruciale basis vormen voor het begrijpen van een opwindend nieuw onderzoeksgebied in de huidige celbiologie: biomoleculaire condensaten. Dit hoofdstuk stelt dat de IDR-gedreven interacties en de modulatie van DNA-interacties die hier zijn waargenomen, waarschijnlijk sleutelfactoren zijn in de vorming en regulatie van de transcriptionele complexen waar genexpressie wordt georkestreerd. Het positioneert dit proefschrift niet als een eindpunt, maar als een essentieel startpunt, dat de moleculaire gegevens levert die nodig zijn om het complexe samenspel van krachten te ontcijferen dat genregulatie binnen deze dynamische compartimenten beheerst.