

## Samenvatting

Het onderzoek waar ik de afgelopen jaren bij betrokken ben geweest, richt zich op de  $\beta$ -globine genen en in het bijzonder op de vraag hoe de activiteit van deze genen gereguleerd wordt in de celkern van bloedcellen. Allereerst zal ik een korte inleiding geven, vervolgens de doelstelling van het onderzoek en tot slot de behaalde resultaten beschrijven.

De term actief gen betekent dat er transcriptie plaats vindt op dat gen. Transcriptie is het proces waarbij een enkele streng van het genetische materiaal (d.w.z. het DNA) wordt gekopieerd naar een complementaire RNA molecuul. Dit proces wordt uitgevoerd door speciale eiwitcomplexen genaamd RNA polymerases (RNAPs) en de celkern bevat drie verschillende RNAPs. Eén daarvan, RNAP II, is voornamelijk betrokken bij de transcriptie van genen die coderen voor eiwitten, zoals de  $\beta$ -globine genen. Transcriptie is niet een spontaan proces, maar gebeurt onder strikte regulatie in de celkern. Allereerst, begint transcriptie op een specifieke plek op het DNA ten opzichte van het gen om een functioneel RNA product te krijgen. Dit zijn de zogenaamde promoter sequenties die slechts enkele honderden basenparen (d.w.z. DNA bouwstenen) groot zijn, terwijl het gehele DNA in de kern (het genoom) enkele miljarden basenparen beslaat. RNAP II moet als het ware deze kleine promoter elementen herkennen om er vervolgens te “landen” en transcriptie te starten. Ten tweede, elke celkern bevat het complete genoom, maar niet alle genen (een mens heeft ongeveer 30 duizend genen) zijn actief. Tijdens de ontwikkeling van een organisme, zoals mens of muis, wordt bepaald welke genen in welke cellen actief zijn. Bovendien kunnen bepaalde celtypen de expressie hun genen nog veranderen tijdens het volwassen leven. Bijvoorbeeld: de vorming van bloedcellen gebeurt continu, hierbij wordt een zogenaamde stamcel via verschillende celdelingen omgevormd tot een bepaald type bloedcel. Dit gaat gepaard met het activeren en inactiveren van bepaalde genen. Tot slot, het DNA in een celkern is “verpakt” in bepaalde eiwitten en gezamenlijk vormen zij het chromatine (zie figuur 1.2.1). Dankzij deze verpakking zijn RNAP complexen niet in staat zijn transcriptie uit te voeren. Deze drie omstandigheden zorgen ervoor dat RNAP II complexen hulp nodig hebben van andere componenten die promoter elementen herkennen, bepalen wanneer een gen actief is, en de chromatine structuur beïnvloeden, om uiteindelijk transcriptie mogelijk te maken.

Die componenten zijn op te delen in twee categorieën. Ten eerste zijn er componenten die functioneren *in cis*, d.w.z. het zijn DNA elementen die op hetzelfde molecuul gelegen zijn als het gen dat transcriptie ondergaat. Dit zijn de zogenaamde *cis*-regulerende DNA elementen. Deze elementen bestaan naast de al eerder genoemde promoter uit enkele andere elementen allen met verschillende functionele definities. Echter, deze elementen lijken twee eigenschappen gemeen te hebben. Het zijn korte DNA sequenties (enkele honderden basenparen) waarin veel unieke korte sequenties (van enige basenparen groot) voorkomen. Daarnaast is de structuur van hun chromatine is wezenlijk anders dan het omliggende chromatine. Het lijkt meer toegankelijk te zijn voor eiwitten. Naast de *cis*-regulerende DNA elementen zijn er componenten die functioneren *in trans*, d.w.z. het zijn voornamelijk eiwit (maar ook RNA) moleculen die niet deel uitmaken van het gen bevattende DNA molecuul. Hier omschreven als *trans*-werkende factoren. Deze factoren behelzen eiwitten die deel uitmaken van de zogenaamde ‘transcriptie machine’ (waaronder RNAP II), factoren die chromatine structuur kunnen beïnvloeden, en transcriptie factoren die de korte unieke sequenties in *cis*-regulerende DNA elementen kunnen herkennen en binden. Al deze factoren kunnen binden aan elkaar en deze interacties hebben invloed op transcriptie van genen. De samenstelling en concentratie van deze *trans*-werkende factoren kunnen per celtype verschillen. Zo heeft een bloedcel een andere samenstelling van factoren dan een hersencel en daarom zijn er verschillende genen actief in de beide celtypen. Bij veel genen is waargenomen dat een juiste samenstelling van *cis*-regulerende DNA

elementen en *trans*-werkende factoren noodzakelijk is om transcriptie te krijgen. De  $\beta$ -globine genen vormen hierop geen uitzondering.

RNA transcripten van  $\beta$ -globine genen worden omgezet in het eiwit  $\beta$ -globine. Samen met  $\alpha$ -globine vormt  $\beta$ -globine het eiwitcomplex hemoglobine dat er voor zorgt dat zuurstof van de longen naar de verschillende organen in ons lichaam wordt vervoerd. Indien er iets fout gaat bij het maken van het globine eiwit ontstaat er een bloedcel die deze taak niet naar behoren kan uitvoeren. Er zijn twee groepen patiënten met zulke bloedcellen: patiënten met een niet-functioneel globine eiwit en patiënten die helemaal geen globine eiwitten maken. In beide gevallen is aangetoond dat dit wordt veroorzaakt door foutjes in het DNA. Dit heeft geleid tot veel onderzoek naar de regulatie van  $\beta$ -globine transcriptie.

De  $\beta$ -globine genen zijn gelegen in een cluster (een locus). De mens heeft vijf globine genen en de muis vier (weergegeven als rode blokjes in figuur 1.4.2) en elk gen heeft zijn eigen promotor. Tevens bevat zowel het menselijk- als het muizenlocus enkele andere *cis*-regulerende DNA elementen (pijlen in figuur 1.4.2). Ook is er een heel scala aan *trans*-werkende factoren betrokken bij de transcriptie van deze genen. De genen zijn alleen actief in rode bloedcellen en tijdens de ontwikkeling verandert het expressie patroon van de genen. Voor de mens houdt dit in dat tijdens de eerste acht weken van de embryonale ontwikkeling het  $\epsilon$ -globine gen actief is. Vervolgens zijn tot aan de geboorte de genen  $\epsilon\gamma$  en  $\delta\gamma$  actief. Na de geboorte worden de  $\delta$  en  $\beta$  genen geactiveerd. Een aantal jaren geleden is aangetoond dat de globine genen naast de promotor ook afhankelijk zijn van enkele *cis*-regulerende DNA elementen die relatief ver van de genen liggen. Deze combinatie van *cis*-regulerende DNA elementen wordt ook wel de "Locus Control Region" (LCR) genoemd. Een vraag die hierbij ontstond was hoe relatief ver uit elkaar gelegen *cis*-regulerende DNA elementen transcriptie kunnen beïnvloeden? Bovendien suggereert het dat de LCR en de  $\beta$ -globine genen op een of andere manier met elkaar communiceren.

De doelstelling van het onderzoek waar ik bij betrokken ben geweest was het verkrijgen van meer inzicht in de regulatie van de  $\beta$ -globine genen. Hierbij werd er voornamelijk geconcentreerd op de manier van communicatie tussen de LCR en de genen. Bij het onderzoek gingen we uit van een model dat verklaarde hoe de LCR de verschillende genen van het  $\beta$ -globine locus zou kunnen activeren. In dit model functioneren de afzonderlijk *cis*-regulerende DNA elementen van de LCR als een eenheid, een "holocomplex", dat bijeen gehouden wordt door *trans*-werkende factoren. Het "holocomplex" gaat een directe interactie aan met de *trans*-werkende factoren die gebonden zijn aan een promotor van een globine gen. Door deze interactie wordt dit gen geactiveerd. Het DNA dat tussen de LCR en het gen ligt vormt dan als het ware een lus en heeft geen functie. Communicatie tussen LCR en genen is in dit model gebaseerd op een directe interactie tussen beiden (zie figuur 1.4.5). Het onderzoek beschreven in dit proefschrift beproeft dit model aan de hand van experimenten met het menselijke  $\beta$ -globine locus in muizen, maar ook door studie van het  $\beta$ -globine locus van de muis zelf.

De experimenten beschreven in hoofdstuk 3 meten de 3-dimensionale structuur van het muizen  $\beta$ -globine locus in de celkern van hersencellen en rode bloedcellen. De resultaten laten zien dat het locus in hersencellen, waar  $\beta$ -globine genen niet actief zijn, een lineaire structuur aanneemt. In de kern van een rode bloedcel bevinden LCR elementen en twee actieve  $\beta$ -globine genen zich dicht bij elkaar in de 3-dimensionale ruimte. Ofschoon de LCR en deze genen gescheiden worden door ongeveer 40 tot 60 duizend basenparen. De inactieve  $\beta$ -globine genen, die gelegen zijn tussen de LCR en de actieve genen, zijn niet dichtbij gelokaliseerd, maar maken deel uit van een lus. Deze resultaten leveren direct bewijs voor het model waarbij communicatie tussen LCR en genen is

gebaseerd op een directe interactie tussen beiden. Bovendien zijn er twee andere *cis*-regulerende DNA elementen die ook dichtbij de LCR en actieve genen gelegen zijn. Deze elementen zijn aan de buitenzijden van het locus gelegen. De laatste vinding wijst er op dat *cis*-regulerende DNA elementen een cluster vormen in een locus waar actieve transcriptie plaatsvindt. Deze cluster vorming hebben we "Active Chromatin Hub" (ACH) genoemd (zie figuur 3.7).

In hoofdstuk 4 wordt aangetoond dat niet alleen het muizen  $\beta$ -globine locus een ACH vormt, maar ook het menselijke locus gemeten in transgene muizen. De ACH wordt gevormd tijdens het differentiatie proces van rode bloedcellen. In de voorlopers van de rode bloedcellen vindt geen transcriptie plaats, maar er bestaat al wel een structuur in het locus. Het lijkt er op dan de *cis*-regulerende elementen aan de buitenkanten van het locus en de voorkant van de LCR met elkaar interacteren (zie figuur 4.5). Tevens nemen we waar dat tijdens verschillende stadia van de ontwikkeling er een soort kernstructuur in de ACH aanwezig is. Deze kernstructuur is een interactie tussen alle elementen van de LCR en de *cis*-regulerende DNA elementen aan de buitenzijden van het locus. De  $\beta$ -globine genen gaan alleen een directe interactie met de kernstructuur aan wanneer zij transcriptieel actief zijn. De kernstructuur wordt alleen gevormd in rode bloedcellen en is stabiel tijdens de ontwikkeling. De ruimtelijke interactie tussen de kernstructuur en een gen leidt tot een lokaal verhoogde concentratie van consensus-sequenties waar transcriptie factoren kunnen binden. Uiteindelijk leidt dit tot een ophoping van *trans*-werkende factoren binnen de ACH structuur (zie figuur 4.5). Lokale ophoping (of hoge concentratie) van eiwitten is een fenomeen wat vaker wordt waargenomen in de ingewikkelde structuur van een celkern en wordt compartiment vorming genoemd. Een van die compartimenten, de nucleolus, is noodzakelijk voor transcriptie van de zogenaamde ribosmale RNA genen. Compartiment vorming leidt in dit geval tot efficiënte transcriptie. Dit kan worden verklaard uit het feit dat de reactiesnelheid van een (bio)chemische reactie afhankelijk is van de concentratie van de betrokken moleculen. In hoofdstuk 5 worden de mogelijke gevolgen van een ACH compartiment vorming bediscussieerd.

Interessant is wel dat de meeste compartimenten in de celkern in eerste instantie altijd microscopisch werden waargenomen. De ACH is echter te klein om microscopisch waar te nemen. Hoewel in hoofdstuk 2 een techniek beschreven wordt die het mogelijk moet maken om een menselijk  $\beta$ -globine locus in levende celkernen te volgen met een microscoop. Dit werk is nog niet af, maar er wordt nog aan gewerkt om het mogelijk te maken. Mocht dit uiteindelijk lukken dan heeft deze techniek nog wat toe te voegen aan de techniek beschreven in hoofdstukken 3 en 4. Eerder is namelijk aangetoond dat transcriptie van de  $\beta$ -globine genen één voor één plaatsvindt en dat transcriptie tussen verschillende actieve genen voortdurend afwisselt. Dit toont aan dat transcriptie van de genen een dynamisch proces is. Echter, de techniek beschreven in hoofdstukken 3 en 4 is niet in staat om deze dynamiek waar te nemen. Het volgen van de verschillende *cis*-regulerende DNA elementen in levende celkernen zou dit wel kunnen. Dit kan waardevolle informatie opleveren over regulatie van transcriptie, onder andere door te testen hoe stabiel de interacties in de ACH zijn en welke *trans*-werkende factoren essentieel zijn voor stabiliteit.