

6.3 Samenvatting

Antilichamen zijn een essentieel onderdeel van ons immuun systeem en behoren tot de meest polymorfe eiwitten in het menselijk lichaam. Dit polymorfisme zorgt voor de ongelooflijke flexibiliteit van het verworven immuunsysteem en maakt ons antilichamen repertoire tot een "gepersonaliseerd proteoom", uniek voor elk individu en aangepast aan diens behoeften. Deze diversiteit maakt het analyseren van antilichamen echter ook zeer uitdagend. De hoofdstukken in dit proefschrift beschrijven mijn contributies aan het ontwikkelen van generaliseerbare computationale analyse strategieën voor het bestuderen van antilichaam repertoires met behulp van massaspectrometrie.

Hoofdstuk 2 is uniek in dit proefschrift, omdat we het belang benadrukken van algemene computationele hulpmiddelen om grote en complexe crosslinking proteomics datasets effectief te analyseren. Crosslinking MS is een aantrekkelijke methode geworden om eiwitinteracties te onderzoeken, maar levert zeer grote datasets op, omdat de complexiteit van eiwitinteracties exponentieel stijgt ten opzichte van individuele eiwitten. Dit heeft geleid tot de productie van grote hoeveelheden complexe datasets waarvan interpretatie moeilijk is, vooral voor experimenten gericht op het gehele proteoom. Om dit probleem te verhelpen hebben wij een interactief programma ontwikkeld dat de analyse en visualisatie van deze grote datasets vergemakkelijkt. Het programma, CrossID, verwerkt de output van XlinkX voor Proteome Discoverer, maar kan ook worden gebruikt met output van andere platforms door een flexibele import-module. Het kan spectra annoteren en visualiseren en ondersteunt de analyse van replicaten, zodat grote hoeveelheden gegevens gemakkelijk kunnen worden verwerkt. We hebben ook gegevens geïntegreerd uit de EggNOG en Uniprot databases, die geïntegreerde genontologieverrijkingsanalyse, groepering op basis van eiwitfunctie, en visualisatie van bekende posttranslationale modificatieplaatsen, domeinen en secundaire structuren mogelijk maken. Een andere functie

is validatie van gedetecteerde crosslinks middels hun lengte door het plaatsen van de gecrosslinkede peptiden op gevalideerde structurele modellen van eiwitten of eiwitcomplexen. In situaties waarin geen structuur beschikbaar is, kunnen structuren verkregen door homologiëmodelleren worden gebruikt. In dergelijke gevallen worden gecrosslinkede peptiden met behulp van sequentiealignering (Engels: sequence alignment) geplaatst op de homologe sequentie om een betrouwbare plaatsing van de gekoppelde residuen te verkrijgen waarna de afstand tussen deze residuen wordt berekend aan de hand van de 3D-structuur. Crosslinks tussen twee eiwitten met bekende structuren waar geen structuur van het complex beschikbaar is kunnen ook direct worden ingediend bij DisVis voor visualisatie en kwantificering van de informatie in de crosslinks. Hierbij worden de crosslinks, en het feit dat deze een bekende min- en maximum lengte hebben, gebruikt om een structuur te genereren die zo goed mogelijk overeenkomt met de crosslinking data.

In **hoofdstuk 3** laten we zien dat ontwikkelingen in eiwitmassaspectrometrie de detectie van individuele IgG1-moleculen in menselijk serum mogelijk maakt, met klonale resolutie. Dit stelde ons in staat om gepersonaliseerde klonale IgG1-repertoires te construeren. Deze serologische repertoires werden gedomineerd door enkele honderden klonen, een onverwachts laag aantal wanneer het enorme aantal theoretisch mogelijke klonen in acht wordt genomen. Bovendien waren deze profielen grotendeels stabiel: Het merendeel van de gedetecteerde klonen was langdurig aanwezig. We zagen echter ook aanzienlijke veranderingen in de repertoires na een sepsis-episode. Tevens hebben we aangetoond dat een combinatie van peptide- en eiwitgerichte massaspectrometrie kan worden gebruikt om de sequentie van individuele klonen uit het serum *de novo* te bepalen. De peptide-centrische benadering levert hierbij een uitgebreide dekking op, terwijl de eiwit-centrische benadering sequentie-informatie oplevert die per definitie gegroepeerd is per kloon. De synergie tussen deze technieken werd gebruikt om de sequentie van een enkele zeer abundante

kloon uit het serum van een van onze donoren te bepalen.

Hoofdstuk 4 beschrijft hoe wij de antilichaam repertoire-profileringsstechniek uit Hoofdstuk 3 gebruikt hebben om individuele donoren binnen een groep te vergelijken en te karakteriseren. We hebben SIgA1-profielen samengesteld voor moedermelk van een cohort van zes moeders, die twee identieke SARS-CoV-2 vaccinaties hadden ontvangen. De betreffende profielen vormden een aanvulling op bevindingen van eerdere ELISA-gebaseerde analyse van deze monsters, waarbij een bifasische stijging van spike-specifiek IgA werd gevonden. In alle donoren detecteerden wij een heterogene polyklonale populatie van tussen de 100 en 200 klonen die na vaccinatie onstonden. Deze populatie werd gedomineerd door populatie van persistente klonen die kort na de eerste vaccinatie verschijnt en blijft bestaan tot ten minste vijf dagen na de tweede vaccinatie. Wij observeerden bij elke donor ook een populatie klonen die pas meer dan drie dagen na de tweede vaccinatie ontstaan. Diepgaande analyse van een sterke responder, geselecteerd door ELISA en bevestigd door onze analyse, toonde aan dat de tweede stijging van spike-specifiek IgA van deze donor werd gedreven door een zeer abundante populatie van door de tweede dosis geïnduceerde klonen, wat de uiteenlopende klonale samenstelling benadrukt van wat aanvankelijk leek op een symmetrische bifasische respons. Bovendien zagen we enkele zeer abundante klonen verschijnen en vervolgens verdwijnen uit het repertoire in een periode van ~40 dagen, waaruit blijkt dat zeer overvloedige klonen niet noodzakelijk langdurig dominant blijven.

In **hoofdstuk 5** bouwen we op in hoofdstuk 3 gelegde funderingen voor antilichaam-sequentiebepaling, door een meer gestandaardiseerde en algemeen toepasbare computationele workflow te creëren voor de *de novo* sequentiebepaling van antilichaamketens in mengsels met behulp van een combinatie van peptide- en eiwitgerichte massaspectrometrie. Onze aanpak bepaalt de sequentie van een geselecteerde eiwitketen in een modulair, driedelig proces op basis van sequentiedomeinen. We

beginnen met het bepalen van de sequentie van de geconserveerde domeinen, de zogeheten “framework” regio's. Vervolgens bepalen we de sequentie van de complementariteitsbepalende regio's en de aangrenzende framework regio's, voordat we uiteindelijk de sequentie van de volledige eiwitketen bepalen. Door het gebruik van middle-down fragmentatie konden we de ambiguïteit in de hypervariabele complementariteitsbepalende regio's oplossen. Daartoe filterden wij kandidaat-sequenties door hun theoretische massa te vergelijken met het “gat” tussen de aangrenzende framework regio's. We hebben de effectiviteit van deze aanpak aangetoond door de sequentie van een antilichaamketen te bepalen in drie monsters: Een zuiver monoklonaal antilichaammonster, een equimolair mengsel van drie monoklonale antilichamen en een polyklonaal serummonster. Deze aanpak biedt een generaliseerbare workflow die in toekomstige studies kan worden gebruikt om complexe monsters met hoge nauwkeurigheid te sequencen en brengt ons dichterbij volledige automatisering van dit proces.