

Porównawcza analiza strukturalno-funkcjonalna
in silico białek STAT i IRF,
w celu identyfikacji związków chemicznych,
działających jako specyficzne inhibitory funkcjonalne

rozprawa doktorska w języku angielskim
ze streszczeniem w języku polskim i angielskim

autor:

Małgorzata Szelağ

Zakład Genetyki Molekularnej Człowieka
Instytut Biologii Molekularnej i Biotechnologii
Wydział Biologii
Uniwersytet im. Adama Mickiewicza w Poznaniu

przygotowana pod kierunkiem:

Prof. Dr hab. inż. Johannes A.R. Bluijssen

Zakład Genetyki Molekularnej Człowieka
Instytut Biologii Molekularnej i Biotechnologii
Wydział Biologii
Uniwersytet im. Adama Mickiewicza w Poznaniu

Dr Anna Czerwonec

Zakład Bioinformatyki
Instytut Biologii Molekularnej i Biotechnologii
Wydział Biologii
Uniwersytet im. Adama Mickiewicza w Poznaniu

Poznań 2016

Summary in Polish

Streszczenie po polsku

Wstęp

Sygnalizacja komórkowa

Efektywna komunikacja pomiędzy komórkami jest konieczna dla prawidłowego rozwoju, utrzymania homeostazy tkanek i organizmu, jak również dla odpowiedzi immunologicznej. Aktywacja różnych szlaków sygnalizacyjnych prowadzi do zróżnicowanych odpowiedzi fizjologicznych, takich jak proliferacja komórek, śmierć, różnicowanie i metabolizm.

Transdukcja sygnału jest procesem, w którym zewnętrzne i wewnętrzne bodźce aktywują receptory uruchamiające łańcuch reakcji prowadzących do zmian fizjologicznych wewnątrz komórki. Bodźcami sygnalizacyjnymi są najczęściej małe cząsteczki lub białka, podczas gdy odpowiedź może dotyczyć modyfikacji w ekspresji genów lub zmian niezależnych od genomu. Po odkryciu komponentów molekularnych głównych ścieżek sygnałowych, kontrolujących wzrost i rozwój uważa się obecnie, że reakcja na dany sygnał nie jest wynikiem jednego szlaku sygnalizacji liniowej, ale raczej włączeniem dróg, które podlegają integracji na wielu poziomach. Zjawisko to doprowadziło do używania terminu „**sygnalizacji krzyżowej**” [1]. Zgodnie z tą definicją, za przykład sygnalizacji krzyżowej uważa się złożone interakcje, które istnieją między wewnątrzkomórkowymi szlakami sygnalizacyjnymi przy użyciu sensytyzacji różnych typów komórek układu odpornościowego i naczyniowego w mechanizmie miażdżycy [2].

Układ odpornościowy chroni organizm gospodarza przed atakiem drobnoustrojów i transformacją nowotworową. Składa się on z wrodzonej i nabytej odpowiedzi immunologicznej. **Odpowiedź wrodzona** stanowi pierwszą linię obrony przed patogenami. Aktywacja wrodzonych reakcji polega na rozpoznawaniu zachowanych struktur występujących w patogenach, zwanych molekularnymi wzorcami związanymi z patogenami (PAMPs). Te wzorce molekularne pozwalają odróżnić komórki własne od komórek obcego organizmu. W takim rozpoznaniu pośredniczą głównie receptory Toll-podobne (TLRs), obecne na powierzchni komórek dendrytycznych (DCs) i makrofagów. Monocyty/makrofagi i komórki „natural killer” (NK) zajmują centralne miejsce we wrodzonej odpowiedzi immunologicznej. Makrofagi bezpośrednio fagocytują i niszczą czynniki zakaźne. Komórki NK niszczą komórki zakażone wirusem lub zmienione nowotworowo przez cytolizę. DCs są komórkami prezentującymi antygen (APCs), które wychwytyują i przetwarzają obce antygeny pochodzące z drobnoustrojów. Następnie przemieszczają się do lokalnych węzłów chłonnych, gdzie prezentują antygen naiwnym limfocytom T, aktywując specyficzną odpowiedź immunologiczną. Makrofagi, komórki NK i DCs wytwarzają cytokiny, które są mediatorami uruchomionej wrodzonej odpowiedzi immunologicznej oraz inicjują rozpoczęcie odpowiedzi nabytej przez limfocyty T i B, czyli komórki zdolne do rozpoznawania i reagowania na specyficzne

antygeny [3]. **Odpowiedź nabyta** obejmuje przeciwciała, czyli odpowiedź humoralną i limfocyty będące odpowiedzią komórkową. Limfocyty B, pochodzące ze szpiku kostnego wytwarzają przeciwciała. Limfocyty T dojrzewające w grasicy, różnicują się do komórek pomocniczych, które uczestniczą w dojrzewaniu innych limfocytów lub do komórek zdolnych do bezpośredniego niszczenia zmienionych wirusowo komórek organizmu gospodarza. Odpowiedź humoralna i komórkowa są istotne dla obrony antywirusowej. Udział tych mechanizmów jest różny i zależy zarówno od wirusa jak i gospodarza. Przeciwciała zazwyczaj wiążą się z cząsteczką wirusa we krwi lub na powierzchni błony śluzowej, blokując w ten sposób rozprzestrzenianie się infekcji. Natomiast limfocyty T bezpośrednio rozpoznają i niszczą komórki zainfekowane. Ważną cechą nabytej odpowiedzi immunologicznej jest pamięć. Powtórne zakażenia przez tego samego wirusa inicjuje natychmiastową, silną i ukierunkowaną odpowiedź, która skutecznie zatrzymuje infekcję ze zmniejszeniem w znacznym stopniu odpowiedzi pierwotnej [3].

Cytokiny, mianowicie **interferony (IFNs)** i **interleukiny (ILs)** są grupą międzykomórkowych białek sygnalizacyjnych, które odgrywają ważną rolę w regulacji homeostazy komórkowej i odpowiedzi odpornościowej. Są one uwalniane przez komórki zarówno w mechanizmie odpowiedzi wrodzonej (monocyty, makrofagi, DCs) jak i nabytej (limfocyty T i B) układu odpornościowego. Cytokiny wpływają na różnicowanie i funkcję komórek układu odpornościowego, przez regulację aktywacji i ekspresję specyficznych czynników transkrypcyjnych. Zewnątrzkomórkowe sygnały są przesyłane z powierzchni komórki do jądra poprzez wewnątrzkomórkowe szlaki sygnalizacyjne sprzężone ze specyficznymi receptorami cytokin. Szlak kinaz Janusowych oraz przekaźników sygnałów i aktywatorów transkrypcji (JAK-STAT) to ścieżka najczęściej używana przez cytokiny. Inne szlaki sygnalizacyjne i czynniki transkrypcyjne aktywowane przez cytokiny obejmują: szlak zależny od kinaz białkowych aktywowanych mitogenem (MAPK), system mediowany przez jądrowy czynnik κ B (NF- κ B) i czynniki regulujące interferon (IRFs). Sygnały z tych wszystkich ścieżek zbiegają się w jądrze i w efekcie powodują zmiany w komórkowych profilach transkrypcyjnych [4-6].

Czynniki transkrypcyjne to nie jedyne elementy wewnątrzkomórkowe regulujące ekspresję genów. W komórkach eukariotycznych interakcja białek i DNA tworzy dynamiczną strukturę zwaną chromatyną. Dostęp elementów regulatorowych DNA do czynników transkrypcyjnych i polimerazy II RNA, w kontekście chromatyny regulowana jest poprzez modyfikacje kowalencyjne białek związanych z DNA, a w niektórych przypadkach samego DNA.

Rodziny białek STAT i IRF – czynniki transkrypcyjne

Przekaźniki sygnału i aktywatory transkrypcji (STAT) są zaangażowane w kontrolowanie ekspresji genów w procesach biologicznych tj. rozwój embrionalny, programowana śmierć komórki, organogeneza, odporność wrodzona i nabyta, regulacja wzrostu komórek [7]. U ssaków rodzina STAT składa się z siedmiu zachowanych ewolucyjnie białek:

STAT1, 2, 3, 4, 5A, 5B i 6. STAT-y są strukturalnie zbudowane z 6 zachowanych domen: α -helikalna N-końcowa (ND), zwój czterech α -helis określane jako „coiled-coil” (CC), centralna Ig-podobna wiążąca DNA (DBD), α -helikalny łącznik (LK), domena Src-homologiczna typu 2 (SH2) i C-końcowa transaktywacyjna (TAD). Aktywacja STAT-ów może być osiągnięta przez bodźce, takie jak cytokiny, w tym interferony (IFNs) i interleukiny (ILs), oraz czynniki wzrostu (np. EGF i PDGF), hormony i onkoproteiny (np. Abl i Src). Stymulowana IFN i IL aktywacja STAT jest mediowana fosforylacją kluczowej reszty tyrozynowej przez związane jej z receptorami kinaz Janusa (JAK). Prowadzi to do dimeryzacji STAT-ów, poprzez bilateralne oddziaływanie monomerów z udziałem fosfotyrozyn (pTyr) i domen SH2. Homo- i heterodimery STAT przemieszczają się do jądra komórkowego, gdzie regulują transkrypcję licznych genów docelowych, wiążąc się do ich elementów promotorowych zawierających palindromowy motyw DNA zwany elementem GAS (miejsce aktywacji IFN γ), różny dla poszczególnych STAT-ów. Szlak sygnalizacyjny JAK-STAT to jeden z najprostszych architektonicznie paradygmatów, pozwalających na bezpośrednią komunikację pomiędzy receptorami błonowymi a jądrem komórkowym [8].

Czynniki transkrypcyjne określane jako **czynniki regulujące interferon (IRF)** są modulatorami transkrypcji w wielu procesach biologicznych, począwszy od immunologicznej odpowiedzi wrodzonej i nabytej organizmu, regulacji wzrostu komórek i ich apoptozy po hematopoezę. IRF-y są pierwotnie związane z wrodzoną odpowiedzią układu odpornościowego zależną od receptorów rozpoznających wzorce (PRRs), zawierających grupę receptorów Toll-podobnych (TLRs). Białka IRF uczestniczą w mediowanej przez PRRs indukcji cytokin prozapalnych, odpowiedzi indukowanej uszkodzeniem DNA czy regulacji onkogenezy. Ponadto IRF-y pełnią kluczową rolę w rozwoju komórek układu odpornościowego, tj. DCs, NK, limfocyty T i B [9]. Występujące u ssaków IRF-y to rodzina 9 homologicznych białek (IRF1-9) o budowie wielodomenowej. N-koniec stanowi domenę wiążącą DNA (DBD) i charakteryzuje się obecnością 5 reszt tryptofanowych tworzących „klastery tryptofanowy”. Białka IRF cechuje ewolucyjnie zachowana homologia sekwencji i struktury 2- i 3-rzędowej w rejonie DBD. C-koniec wszystkich IRF-ów zawiera domenę asocjacyjną IRF (IAD), za pomocą której dochodzi do tworzenia homo- i heterodimerów IRF oraz wiązania różnorodnych czynników transkrypcyjnych, co jest niezwykle istotne w regulacji rozpoznawania motywów DNA. Tego typu interakcje modulują aktywność IRF-ów oraz wiązanie do wielu genów. IRF-y rozpoznają podobny motyw DNA – element regulacyjny IFN (IRE), który jest obecny w rejonach regulatorowych genów IFNs i genów stymulowanych przez IFNs (ISGs). IRF-y mogą również wiązać się do elementu odpowiedzi stymulowanej interferonem (ISRE), będącego tandemowym powtórzeniem kilku IRE [10].

Badania genetyczne prowadzone na myszach dostarczyły różnorodnych informacji jak funkcjonują czynniki transkrypcyjne STAT i IRF oraz dowiodły ich kluczowego znaczenia we wrodzonej i nabytej odpowiedzi immunologicznej. Cytokiny i PRRs uczestniczą w sygnalizacji

krzyżowej poprzez współdzielenie tych samych czynników transkrypcyjnych – STAT, IRF, NF-κB, AP1 – mających zdolność do aktywacji ekspresji genów w różnych kombinacjach. Sygnalizacja ta może doprowadzić do synergistycznej amplifikacji całych zestawów genów poprzez wiele kombinacji wiązania czynników transkrypcyjnych do genów docelowych. Prozapalne działanie IFNs, IL-6 i TLR4 prowadzi do sensytyzacji różnych komórek układu odpornościowego i naczyniowego w mechanizmie miażdżycy, co stanowi jedno z zagadnień rozwijanych w niektórych publikacjach wchodzących w skład przedstawionej pracy doktorskiej.

Rola białek STAT i IRF w organizmie. Wpływ na rozwój i progresję chorób

Białka STAT i IRF mają wpływ na fundamentalne procesy komórkowe, w tym wzrost komórek i ich różnicowanie, procesy rozwojowe, apoptozę, odpowiedź immunologiczną i stan zapalny. Poprzez utrzymanie równowagi fizjologicznej organizmu, regulację wrodzonej i nabytej odpowiedzi immunologicznej gospodarza oraz supresję nowotworową STAT-y i IRF-y są „węzłami komunikacyjnymi” dla wielu onkogennych i zapalnych dróg przekazywania sygnałów. Nieprawidłowa aktywacja szlaków sygnalizacyjnych STAT i IRF, a w konsekwencji ich deregulacja przyczynia się do rozwoju i progresji wielu chorób.

Niektóre typy zaburzeń pierwotnego niedoboru odporności połączono ze znanymi mutacjami STAT. Mutacje w genie STAT1 wiążą się z podatnością na infekcje wirusowe, przewlekłą śluzówkowo-skórną kandydozą (CMC), autoimmunizacją, tętniakami mózgu, rakiem kolczystokomórkowym i nawracającymi zakażeniami *Mycobacterium*. Mutacje genu STAT2 mogą powodować podatność na infekcje wirusowe. Aberracje genu STAT3 są podłożem zespołu hiper-IgE (HIES, zespół Joba). Mutacje w genie STAT5B prowadzą do złożonego zespołu objawów charakteryzującego się karłowatością, niedoborami immunologicznymi i autoimmunizacją [11].

Pod koniec XX wieku konstytutywna aktywacja szlaku JAK-STAT została powiązana z nowotworami złośliwymi. Nieprawidłową aktywację STAT stwierdzono w wielu schorzeniach onkologicznych. STAT3 jest konstytutywnie aktywny w nowotworach litych (rak piersi i nosogardzieli) i hematologicznych, (chłoniak B-komórkowy). Nieprawidłowa sygnalizacja STAT5 również jest zaangażowana w patogenezę nowotworów narządów stałych i hematologicznych (przewlekła białaczka szpikowa) [11].

Całogenomowe badania asocjacyjne wykazały udział dysfunkcję szlaku JAK-STAT w mniej inwazyjnych chorobach człowieka. Polimorfizmy STAT1 są związane ze zwiększonym ryzykiem wystąpienia nowotworu, a polimorfizmy STAT3 z chorobą Crohn'a i łuszczycą. Polimorfizmy pojedynczych nukleotydów STAT4 przyczyniają się do rozwoju reumatoidalnego zapalenia stawów (RA) i toczenia rumieniowatego układowego (SLE), zaś w obrębie STAT6 astmy i alergii [11].

Rodzinę czynników transkrypcyjnych IRF można podzielić ze względu na udział w onkogenezie na dwie grupy: supresyjne w stosunku do nowotworów (IRF1, 5, 8 i 9) i onkogenne

(IRF2 i 4). Mutacje genowe i nieprawidłowa ekspresja IRF-ów mają związek z niektórymi nowotworami układu krwiotwórczego: AML – ostra białaczka szpikowa (IRF1 i IRF8), ALL – ostra białaczka limfoblastyczna (IRF5) i CML – przewlekła białaczka szpikowa (IRF1 i IRF8). Zaburzenie sygnalizacji IRF obserwuje się także w innych nowotworach: rak piersi (IRF1, IRF2 i IRF9), rak płaskonabłonkowy przełyku (IRF1 i IRF2) NSCLC – niedrobnokomórkowy rak płuc (IRF3) i guzy przewodu pokarmowego (IRF5) [12].

Aktualnie prowadzone liczne badania nad schorzeniami serca pokazały, że niewielkiego stopnia stan zapalny jest jednym z objawów chorób układu sercowo-naczyniowego (CVDs). U pacjentów z miażdżycą tętnic i nadciśnieniem stwierdzono udział odpowiedzi immunologicznej w mechanizmach, które przyczyniają się do stanu zapalnego w przebiegu chorób sercowo-naczyniowych [13]. Aktywacja JAK-STAT jest zarówno charakterystyczna i unikalna dla sygnalizacji zależnej od cytokin i czynników wzrostu, pełniących centralną rolę w fizjologii serca. Deregulacja przekazywania sygnałów JAK-STAT, jest związana z różnymi chorobami układu sercowo-naczyniowego będących następstwem miażdżycy, tj. choroba wieńcowa serca, choroba naczyń mózgowych, choroba tętnic obwodowych, zwężenie tętnicy nerkowej, nadciśnieniowa choroba serca, oraz z częstymi powikłaniami w postaci udaru i zawału mięśnia sercowego (MI), [14]. IRF-y są niezwykle ważne jako nowe potencjalne czynniki biorące udział w regulacji genów pacy serca podczas rozwoju patologicznej hipertrofii i przebudowy mięśnia sercowego. Jednakże nadal pozostaje niejasne czy białka IRF wpływają na choroby układu sercowo-naczyniowego pośrednio – poprzez infiltrujące naczynia krwionośne komórki układu odpornościowego i/lub bezpośrednio – zgodnie z mechanizmem komórkowo-autonomicznym [9].

Strategie inhibicji białek STAT i IRF

W ostatnich latach terapia lekowa celująca w szlaki sygnalizacyjne stała się jednym z najważniejszych obszarów współczesnych badań farmakologicznych. W zdrowym organizmie procesy wzrostu komórek i ich różnicowania podlegają ścisłej kontroli. W stanie patologicznym procesy te podlegają deregulacji, prowadzącej do przekazywania sygnałów wywołujących dalsze uszkodzenia oraz rozwoju nieprawidłowych komórek. Proliferacja uszkodzonych lub nieprawidłowych komórek stanowi podłoże licznych chorób, m.in. nowotworów, chorób zakaźnych, stanu zapalnego, miażdżycy, chorób reumatycznych i neurodegeneracyjnych.

Strategie manipulacji białkami STAT są intensywnie badane przez naukowców, i dotyczą inhibicji pośredniej jak i bezpośredniej. Poszukiwanie STAT3 specyficznych ligandów, celujących w blokowanie dimeryzacji, są liczne i zaowocowały odkryciem ponad 100 związków pochodzenia naturalnego lub syntetycznego. Niewiele znamy natomiast inhibitorów innych STAT-ów (1, 4, 5A/B i 6), oraz żadnego dla STAT2 [15-17]. Strategie inhibicji IRF są ograniczone do pośredniego hamowania ekspresji i funkcji, głównie w celach badawczych nad ich aktywnością antywirusową czy korelacją z chorobami nowotworowymi. Opisane do tej pory oddziaływania syntetycznych i

naturalnych substancji na funkcję IRF cechuje obecność niepożądanych reakcji niespecyficycznych (inhibicja kinaz tyrozynowych, wiązanie ligandów do receptorów, negatywny wpływ na postawanie kompleksów sygnalizacyjnych). Do tej pory nie opisano bezpośrednich strategii inhibicji IRF, wykorzystujących wirtualne badania przesiewowe i/lub eksperymenty wysokoprzepustowe [2].

Symulacje *in silico* inhibitorów STAT1 i STAT3 wskazują na istnienie krzyżowej specyficzności wiązania do domeny SH2

Białka STAT1 i STAT3 znajdują się w centrum uwagi badaczy, ponieważ ich aktywność została powiązana z rozwojem wielu chorób. Poszukiwanie związków chemicznych, blokujących aktywność STAT3 poprzez inhibicję kompetencyjną oddziaływania pTyr⁷⁰⁵ z domeną STAT3-SH2, zaowocowało wieloma ligandami o wysokiej specyficzności do STAT3, tj. STATTIC [18]. Pomimo istnienia struktur krystalicznych STAT1 i STAT3, nie wykonano szczegółowej analizy porównawczej domen SH2. Ponadto zjawisko krzyżowej specyficzności wiązania znanych inhibitorów STAT3 do innych białek z tej rodziny nie zostało we właściwy sposób zweryfikowane eksperymentalnie. Cytotoksyczny lek Fludarabina (Flu) jest rzadkim przykładem inhibitora STAT1. Udowodniono, iż Flu hamuje ekspresję STAT1 oraz jego fosforylację [19]. W komórkach śródbłonna naczyń krwionośnych (ECs) Flu znacząco obniżała fosforylację STAT1 jak również ekspresję genów docelowych tego białka, doprowadziła do drastycznej redukcji adhezji monocytów do ECs, co udowodnili Sikorski i Bluijssen z Zakładu Genetyki Molekularnej Człowieka na Wydziale Biologii Uniwersytetu im. Adama Mickiewicza w Poznaniu [20].

W publikacji zatytułowanej „***In silico* simulations of STAT1 and STAT3 inhibitors predict SH2 domain cross-binding specificity**” (**European Journal of Pharmacology, 2013, 720:38-48, DOI: 10.1016/j.ejphar.2013.10.055**), wykorzystując stworzone przez nas modele 3D dla ludzkiego (h)STAT1, hSTAT2 i hSTAT3, wykonaliśmy dokowanie STATTIC i fosforanów Flu do domeny STAT-SH2, w celu analizy specyficzności ich wiązania do STAT. Strategia porównawczego dokowania *in silico* została połączona z analizą inhibicji fosforylacji STAT *in vitro*. Odkryliśmy, że STATTIC wykazywał podobne powinowactwo do kieszeni wiążącej pY+0 w domenie SH2 dla hSTAT1, -2 i -3 oraz hamował ich fosforylację indukowaną IFN α . Natomiast fosforany Flu, wykazały podobieństwo strukturalne do pTyr oraz wysokie powinowactwo wiązania do kieszeni wiążącej pTyr w obrębie STAT1-SH2. Flu hamowała indukowaną IFN γ i LPS fosforylację hSTAT1 i ekspresję IRF1. Okazało się, iż fosforany Flu miały również wysokie powinowactwo wiązania do kieszeni wiążącej pTyr w hSTAT3-SH2, ale brak preferencji do interakcji z hSTAT2-SH2. Flu hamowała indukowaną IFN α fosforylację hSTAT1 i hSTAT3, ale nie hSTAT2 w warunkach *in vitro*.

Na podstawie otrzymanych wyników wywnioskowaliśmy, że wykorzystując jako pierwotne miejsce wiązania ewolucyjnie zachowaną kieszeń pY+0 w domenie SH2, STATTIC nie jest specyficznym inhibitorem hSTAT3. Wiąże się on równie silnie z hSTAT1 i hSTAT2. Biorąc pod

uwagę jego prostą strukturę chemiczną, STATTIC może wykazywać brak selektywności w stosunku do STAT-ów. Fosforany Flu oddziałują z hSTAT1 i hSTAT3 poprzez ich dwie zachowane ewolucyjnie kieszenie wiążące – pY+0 i pY-X. Mechanizm ten odpowiada za wysokie powinowactwo wiązania Flu, która jest w stanie konkurować z wiązaniem receptora do STAT-SH2 i/lub hamować dimeryzację STAT. Znaczny stopień homologii pomiędzy tymi miejscami w hSTAT1 i hSTAT3 determinuje krzyżowe wiązanie Flu, jak również sugeruje znaczne utrudnienia w znalezieniu specyficznych inhibitorów hSTAT1 lub hSTAT3. W połączeniu z wynikami doświadczeń *in vitro*, nasze symulacje *in silico* oferują podstawy molekularne wyjaśniające wiązanie krzyżowe STATTIC i fosforanów Flu do różnych STAT-ów i otwierają drogę do znalezienia nowej strategii selekcji specyficznych inhibitorów STAT1 i STAT3.

Porównawcze wirtualne badania przesiewowe oraz ich walidacja jako nowe narzędzie do identyfikacji specyficznych inhibitorów STAT

Znalezienie specyficznych inhibitorów STAT spotyka się z wieloma wyzwaniem, począwszy od niedostatecznej ilości danych doświadczalnych na temat ich struktury po słabo zdefiniowany mechanizm działania tych białek, mało wydajną aktywność *in vivo* i w konsekwencji brak specyficzności do konkretnego STAT-a wybranych związków chemicznych. Ponadto pojawiają się problemy w odniesieniu do badań *in silico* oraz walidacji i selekcji *in vitro*, obejmujące brak ujednoczonych protokołów dla wirtualnych badań przesiewowych, które opisywałyby powinowactwo związków i ich siłę wiązania do STAT-ów; niestosowanie jednakowych czynników indukujących aktywację STAT w doświadczeniach porównawczych *in vitro*; brak istniejących znanych inhibitorów STAT, które można by wykorzystać jako pozytywną kontrolę.

W publikacji na zaproszenie zatytułowanej „**Comparative screening and validation as a novel tool to identify STAT-specific inhibitors**” (**European Journal of Pharmacology, 2014 740:417-420, DOI: 10.1016/j.ejphar.2014.05.047**) podsumowaliśmy aktualną wiedzę na temat inhibicji STAT-ów z wykorzystaniem związków niskocząsteczkowych pochodzenia naturalnego oraz otrzymanych drogą syntezy. Podjęliśmy się również analizy przyczyny porażek dotychczas podejmowanych strategii modulacji aktywności STAT-ów. Wykorzystując własne doświadczenie z substancjami STATTIC i Flu, proponujemy zastosowanie ujednoczonego protokołu opartego jednocześnie na porównawczych wirtualnych badaniach przesiewowych i walidacji, w celu identyfikacji specyficznych inhibitorów STAT. Wykorzystując nowo opracowane modele strukturalne 3D dla wszystkich ludzkich STAT-ów, sugerujemy połączenie porównawczego dokowania *in silico* do modeli STAT-SH2 z testem fosforylacji STAT *in vitro*, jako nowe narzędzie do przeszukiwania wielomilionowych bibliotek związków, aby zidentyfikować specyficzne inhibitory dla różnych STAT-ów. Identyfikacja specyficznych i wydajnych inhibitorów STAT może stać się nie tylko strategią terapeutyczną w leczeniu nowotworów, stanu zapalnego i

chorób autoimmunologicznych, ale również stanowić narzędzie do zrozumienia funkcjonowania tych białek w różnych stanach chorobowych.

Identyfikacja STAT1 i STAT3 specyficznych inhibitorów z wykorzystaniem porównawczych wirtualnych badań przesiewowych i dokowania jako metody walidacji

Biorąc pod uwagę liczne wyzwania towarzyszące identyfikacji nowych inhibitorów STAT, podjęliśmy decyzję o przygotowaniu ujednoliconego protokołu wirtualnych badań przesiewowych. Podstawowym założeniem naszej publikacji, zatytułowanej „**Identification of STAT1 and STAT3 specific inhibitors using comparative virtual screening and docking validation**” (PLoS ONE, 2015, 10(2):e0116688, DOI: 10.1371/journal.pone.0116688) było opracowanie nowej strategii bioinformatycznej służącej selekcji STAT-specyficznych inhibitorów w oparciu o modele STAT-SH2 w połączeniu z porównawczymi badaniami przesiewowymi *in silico*. Na wstępie wygenerowaliśmy nowe modele 3D dla wszystkich ludzkich STAT-ów (1, 2, 3, 4, 5A, 5B i 6) oraz ich peptydów-pTyr, w oparciu o przyrównanie wielu sekwencji (MSA) należących do białek STAT wśród kręgowców oraz dostępne struktury krystalograficzne STAT1, STAT3 i STAT5A. Przeanalizowaliśmy szczegółowo oddziaływanie STAT-SH2/linker-pTyr w odniesieniu do wcześniej opisanych aktywnych reszt aminokwasowych w domenie STAT3-SH2 oraz superpozycję kieszeni wiążących pY+0 i pY-X dla wszystkich modeli STAT. Nasze symulacje wykazały iż te „miejsca aktywne” zbudowane są przede wszystkim z zachowanych ewolucyjnie aminokwasów oraz ich pozycja w domenie SH2 również jest niezmienna. Udało nam się również potwierdzić obecność bardziej zróżnicowanych rejonów w domenie SH2, wyraźnie odmiennych w zależności od danego typu STAT-a. Sugerujemy, iż jednoczesne wiązanie liganda do kieszeni pY+0 i pY-X oraz wykorzystanie tego podejścia we wszystkich modelach STAT do wirtualnych badań przesiewowych, pozwoli na selekcję STAT1 i STAT3 specyficznych inhibitorów.

Porównawcza strategia dokowania *in silico* do wszystkich domen hSTAT-SH2 została zastosowana w celu przeanalizowania krzyżowego wiązania STATTIC, uważanego za inhibitor STAT3. Symulacje dokowania STATTIC do domeny SH2 wszystkich hSTAT-ów połączyliśmy z parametrami CBAV (wartość porównawczego powinowactwa wiążącego) i LBPV (zróżnicowanie pozycji wiążących liganda) i wykazaliśmy, że STATTIC rzeczywiście wiąże się jednakowo do wszystkich hSTAT. Mechanizm ten polega na oddziaływaniu z jedną z kieszeni wiążących w domenie SH2, mianowicie pY+0 lub pY-X, co odzwierciedlają wartości CBAV i LBPV. Podobną strategię zastosowaliśmy do analizy specyficzności wiązania wyselekcjonowanych 14 inhibitorów STAT3, w tym związków naturalnych (kurkumina, kukurbitacyna E i Q) oraz syntetycznych (LLL12, FLLL32, Cpd188, Cpd30-12, STX-0119, HJC1023, S3I-201, S3I-201.1066, BP-1-102, OPB-31121 i WP1066). Cechą wspólną tych związków jest fakt, iż zostały przebadane pod względem specyficzności do hSTAT3 oraz udowodniono poprzez dokowanie *in silico* ich powinowactwo do kieszeni wiążących w domenie hSTAT3-SH2. Podobnie jak w przypadku

STATTIC wszystkie 14 ligandów oddziaływało z kieszenią wiążącą pTyr w domenie SH2 wszystkich hSTAT-ów. Ponadto, w oparciu o wartości STAT3-CBAV i LBPV oraz graficzną reprezentację wyników, możemy stwierdzić, że związki te nie są STAT3-specyficzne. Symulacje dokowania wybranych 14 inhibitorów hSTAT3 w domenie SH2 wszystkich hSTAT-ów w połączeniu z potwierdzonym w doświadczeniach *in vitro* krzyżowym wiązaniem, pozwoliło nam powiązać parametry CBAV i LBPV z „krzyżowym powinowactwem do STAT” i „specyficznością względem STAT”. W oparciu o te kryteria rozwinęliśmy nową metodologię porównawczych badań przesiewowych *in silico* przeszukiwania komercyjnych, wielomilionowych bibliotek ligandów, aby znaleźć potencjalne, specyficzne inhibitory STAT-ów. Rzeczywiście zastosowanie tego podejścia do przeszukiwania grupy związków naturalnych i wielomilionowej biblioteki ‘clean leads’, umożliwiło nam z sukcesem zidentyfikować potencjalnie specyficzne ligandy STAT1 i STAT3.

CAVS – nowa strategia *in silico* selekcji specyficznych inhibitorów STAT

Metody obliczeniowe stanowią istotny dodatek do badań doświadczalnych *in vitro*, mających na celu selekcję specyficznych związków chemicznych. Strategia ta polega na zaprojektowaniu cząsteczek oraz wykonaniu wirtualnych badań przesiewowych z wykorzystaniem programu komputerowego, zanim zostaną one użyte w eksperymentach laboratoryjnych. Wymagane jest więc podejście precyzyjnie przewidyujące strukturę, funkcję, dynamikę jak również oddziaływanie badanego układu biologicznego z ligandami. Z myślą o wydajności doświadczeń *in vitro* stworzyliśmy narzędzie CAVS do selekcji *in silico* ligandów wiążących się z białkami STAT. W naszej metodologii zastosowaliśmy CAVS w połączeniu z programem Surflex-Dock 2.6 do iteracyjnego przeszukiwania bibliotek ligandów w oparciu o ich powinowactwo wiążące. Wybrane związki chemiczne podlegały wizualnej inspekcji programem PyMOL zanim trafiły do walidacji eksperymentalnej.

W publikacji zatytułowanej „**CAVS – Novel *in silico* selection strategy of specific STAT inhibitory compounds**” (**Journal of Computational Science, 2015, 10:186-194, DOI: 10.1016/j.jocs.2015.03.001**) szczegółowo opisujemy nasze nowe narzędzie do wyszukiwania inhibitorów STAT *in silico* – CAVS (ang. Comparative Approach for Virtual Screening), stworzone w trakcie dokowania ligandów do wszystkich białek STAT. CAVS stanowi istotne usprawnienie analizy wyników dokowania w porównaniu z podejściem zastosowanym w naszych poprzednich badaniach nad STATTIC i fosforanami Flu [21]. Zgodnie z naszą wiedzą jest to pierwsze narzędzie pozwalające na sortowanie i filtrowanie wyników otrzymanych z wykorzystaniem programu Surflex-Dock 2.6. W CAVS zaimplementowane zostały kryteria selekcji inhibitorów białkowych CBAV i LBPV. CAVS pozwala na sortowanie i analizę rezultatów wirtualnych badań przesiewowych wykonanych w Surflex-Dock 2.6. CAVS to program o szerokim zastosowaniu, który można wykorzystać do analizy innych rodzin białkowych czynników sygnalizacyjnych,

uwzględniając kinazy tyrozynowe Janusa (JAK), supresory sygnalizacji cytokin (SOCS) oraz czynniki regulujące interferon (IRF), które spełniają kluczową rolę w mechanizmach odporności i stanu zapalnego, jak również stanowią bardzo interesujące cele terapeutyczne.

Celowana inhibicja STAT i IRF jako potencjalna strategia leczenia chorób układu sercowo-naczyniowego

Miażdżycza jest wiodącą przyczyną zachorowalności i śmiertelności na świecie, szczególnie w krajach wysoko rozwiniętych, pomimo znacznego postępu w zrozumieniu patogenezy tej choroby i strategii leczenia [22]. Najnowsze doniesienia dostarczają dowodów potwierdzających ideę traktującą STAT1, STAT2 i STAT3, oraz IRF1 i IRF8 jako wiodących modulatorów stanu zapalnego, zwłaszcza w komórkach układu odpornościowego i naczyniowego w rozwoju i progresji miażdżycy.

W publikacji zatytułowanej „**Targeted inhibition of STATs and IRFs as a potential treatment strategy in cardiovascular disease**” (**Oncotarget, 2016, May 5, DOI: 10.18632/oncotarget.9195 [Epub ahead of print]**) przedstawiamy podsumowanie tych doniesień. Postulujemy, iż STAT1, STAT2 i STAT3, oraz IRF1 i IRF8 są interesującymi celami terapeutycznymi. Celowana inhibicja tych białek może być potencjalną strategią leczenia chorób układu sercowo-naczyniowego (CVDs). Ponadto, po raz pierwszy prezentujemy nowe modele homologiczne ludzkich czynników transkrypcyjnych, a mianowicie domeny wiążącej DNA IRF1 i IRF8 (hIRF-DBDs). Modele te zostały wykorzystane do badań przesiewowych *in silico* i selekcji potencjalnych, specyficznych inhibitorów. Wykorzystując porównawcze dokowanie *in silico* związków chemicznych jednocześnie do modeli STAT-SH2 i IRF-DBD wraz z walidacją *in vitro* inhibicji aktywacji STAT i IRF, przedstawiamy nowy protokół przeszukiwania wielomilionowych bibliotek ligandów, w celu znalezienia specyficznych inhibitorów STAT i IRF.

Aby lepiej zrozumieć oddziaływanie IRF1 i IRF8 z DNA opracowaliśmy modele strukturalne 3D kompleksów hIRF1- i hIRF8-DBD z docelową sekwencją DNA – IRE, jak również IRF2-DBD dla porównania miejsca oddziaływania białko-DNA. W oparciu o metodologię opisaną dla porównawczego dokowania STAT-specyficznych inhibitorów, zastosowaliśmy to samo podejście do znalezienia specyficznych inhibitorów IRF1 i IRF8 poprzez dokowanie biblioteki związków naturalnych z bazy danych ZINC do hIRF-DBD. W tym celu wykonaliśmy również modele form apo dla hIRF1, hIRF2 i hIRF8-DBD (niezwiązane z DNA), które zostały wykorzystane jako cele molekularne w opracowanej przez nas procedurze wirtualnych badań przesiewowych.

Białka te stanowią interesujące cele terapeutyczne, które mają kluczowe znaczenie w interakcji pomiędzy uszkodzonymi naczyniami a układem immunologicznym w kontrolowaniu miażdżycy wywołanej przez różne czynniki prozapalne. Tak więc STAT1, STAT2 i STAT3, oraz IRF1 i IRF8 mogą stanowić nową metodę CVDs.

Perspektywy

Identyfikacja specyficznych inhibitorów STAT i IRF o wysokiej skuteczności, wydajności i biodostępności stanowi narzędzie do zrozumienia funkcjonowania tych białek w wielu chorobach. Nasza strategia porównawczych wirtualnych badań przesiewowych otwiera nowe możliwości w doświadczeniach klinicznych przez zastosowanie chemicznych modyfikacji, optymalizację i testowanie znanych i dostępnych niespecyficznych inhibitorów STAT.

Począwszy od roku 2000 bezpośrednia inhibicja STAT-ów, celująca w mechanizm dimeryzacji tych białek oparty o oddziaływania peptyd-pTyr-SH₂, jest przedmiotem badań naukowych prowadzonych na szeroką skalę na całym świecie. Jak wiele czynników transkrypcyjnych, białka STAT i IRF nie posiadają typowych kieszeni wiążących związki małowcząsteczkowe. Opierając się na strukturach białkowych otrzymanych drogą eksperymentalną i stworzonych przez nas modelach 3D *in silico*, stwierdza się obecność we wszystkich domenach STAT i IRF rozległych i planarnych powierzchni. Spełniają one funkcję płaszczyzn oddziaływania typu białko-białko, białko-peptyd czy białko-DNA. Te miejsca kontaktu są odpowiedzialne za podstawowe funkcje fizjologiczne białek STAT i IRF, takie jak: oligomeryzacja dimerów STAT mediowana przez domeny N-końcowe, formowanie homo- i heterodimerów STAT poprzez bilateralne oddziaływanie pTyr-SH₂, formowanie homo- i heterodimerów IRF przez oddziaływanie IAD, wiązanie STAT i IRF do DNA mediowane przez DBD. Postulujemy, iż nasze modele 3D białek STAT przyczynią się do rozwoju projektowania leków, opartego o badania struktury. Ponadto pozwolą na syntezę STAT-ND oraz STAT-DBD selektywnych związków chemicznych dzięki eksploracji nowych możliwości inhibicji celowanej, odmiennych od wyeksploatowanej inhibicji domeny STAT-SH₂. Podobne podejście może zostać zaimplementowane do syntezy selektywnych inhibitorów IFR-DBD czy IRF-IAD.

Przedstawiona praca doktorska zapoczątkowuje podjęcie działań badawczych w celu lepszego zrozumienia białek STAT i IRF w odniesieniu do ich roli i funkcji fizjologicznych – wyłączając kontekst medyczny i farmakologiczny. Wiązanie STAT-ów do elementów promotorowych zawierających palindromową sekwencję DNA – zwaną motywem GAS, oraz dimeryzacja STAT, mediowana bilateralnym oddziaływaniem pTyr-SH₂, to znane paradygmaty funkcjonowania tych białek. Pomimo to, do dziś nie wyjaśniono jak specyficzność tych oddziaływań różni się pomiędzy poszczególnymi STAT-ami. Podobny dylemat dotyczy dimeryzacji IRF i rozpoznawania przez nie sekwencji DNA – motywów IRE i ISRE. Początkowe badania nad STAT-ami i IRF-ami były przeprowadzane dla pojedynczych genów. Stanowiło to czynnik limitujący eksplorację zagadnienia oddziaływania STAT i IRF z innymi białkami. Obecnie wiadomo, iż białka STAT i IRF działają nie tylko jako samodzielne czynniki transkrypcyjne, ale również wchodzą w skład złożonych kompleksów makromolekularnych wiążących się z DNA, co potwierdzają przykłady istnienia ISGF3 (heterodimer STAT1-STAT2 i IRF9), PU.1/Spi-B-

IRF4/IRF8 czy p53-NF- κ B/RelA-STAT3. Wykorzystując techniki sekwencjonowania całego genomu dedykowane badaniu czynników transkrypcyjnych (ChIP-Seq i ChIP-exo) możliwe stało się prowadzenie bardziej zaawansowanych doświadczeń nad tym, jak STAT i IRF tworzą kompleksy wielkocząsteczkowe. Modele 3D STAT-ów i IRF-ów opisane w przedstawionej pracy doktorskiej mogą przyczynić się do zrozumienia w jaki sposób te czynniki transkrypcyjne wiążą motywy DNA oraz oddziałują z innymi białkami w mechanizmie regulacji ekspresji genów.

Literatura

1. Vert G and Chory J. Crosstalk in cellular signaling: background noise or the real thing? *Dev Cell*. 2011; 21(6):985-991.
2. Szelag M, Piaszyk-Borychowska A, Plens-Galaska M, Wesoly J and Bluysen HA. Targeted inhibition of STATs and IRFs as a potential treatment strategy in cardiovascular disease. *Oncotarget*. 2016.
3. Abbas AK, Lichtman AH and Pillai S. (2015). Properties and Overview of Immune Responses. In: Abbas AK, ed. *Cellular and molecular immunology*. (Philadelphia: Elsevier Saunders), pp. 1-13.
4. Dy M, Vazquez A, Bertoglio J and Thèze J. (1999). General aspects of cytokine properties and functions. In: Thèze J, ed. *The cytokine network and immune functions*. (Oxford: Oxford University Press), pp. 1-13.
5. Vilcek J and Le J. (1994). Immunology of cytokines: An introduction. In: Thomson AW and Lotze MT, eds. *The Cytokine Handbook*. (London: Academic Press Limited), pp. 1-20.
6. Schindler C, Levy DE and Decker T. JAK-STAT signaling: from interferons to cytokines. *The Journal of biological chemistry*. 2007; 282(28):20059-20063.
7. Horvath CM. STAT proteins and transcriptional responses to extracellular signals. *Trends Biochem Sci*. 2000; 25(10):496-502.
8. Levy DE and Darnell JE, Jr. Stats: transcriptional control and biological impact. *Nature reviews Molecular cell biology*. 2002; 3(9):651-662.
9. Zhao GN, Jiang DS and Li H. Interferon regulatory factors: at the crossroads of immunity, metabolism, and disease. *Biochim Biophys Acta*. 2015; 1852(2):365-378.
10. Honda K and Taniguchi T. IRFs: master regulators of signalling by Toll-like receptors and cytosolic pattern-recognition receptors. *Nature reviews Immunology*. 2006; 6(9):644-658.
11. O'Shea JJ, Schwartz DM, Villarino AV, Gadina M, McInnes IB and Laurence A. The JAK-STAT pathway: impact on human disease and therapeutic intervention. *Annu Rev Med*. 2015; 66:311-328.
12. Takaoka A, Tamura T and Taniguchi T. Interferon regulatory factor family of transcription factors and regulation of oncogenesis. *Cancer Sci*. 2008; 99(3):467-478.
13. Berg AH and Scherer PE. Adipose tissue, inflammation, and cardiovascular disease. *Circ Res*. 2005; 96(9):939-949.
14. Kishore R and Verma SK. Roles of STATs signaling in cardiovascular diseases. *Jak-Stat*. 2012; 1(2):118-124.
15. Debnath B, Xu S and Neamati N. Small molecule inhibitors of signal transducer and activator of transcription 3 (Stat3) protein. *Journal of medicinal chemistry*. 2012; 55(15):6645-6668.
16. Dutzmann J, Daniel JM, Bauersachs J, Hilfiker-Kleiner D and Sedding DG. Emerging translational approaches to target STAT3 signalling and its impact on vascular disease. *Cardiovascular research*. 2015; 106(3):365-374.
17. Miklossy G, Hilliard TS and Turkson J. Therapeutic modulators of STAT signalling for human diseases. *Nature reviews Drug discovery*. 2013; 12(8):611-629.
18. Schust J, Sperl B, Hollis A, Mayer TU and Berg T. Stattic: a small-molecule inhibitor of STAT3 activation and dimerization. *Chem Biol*. 2006; 13(11):1235-1242.
19. Frank DA, Mahajan S and Ritz J. Fludarabine-induced immunosuppression is associated with inhibition of STAT1 signaling. *Nat Med*. 1999; 5(4):444-447.
20. Sikorski K, Czerwoniec A, Bujnicki JM, Wesoly J and Bluysen HA. STAT1 as a novel therapeutic target in pro-atherogenic signal integration of IFN γ , TLR4 and IL-6 in vascular disease. *Cytokine & growth factor reviews*. 2011; 22(4):211-219.
21. Szelag M, Sikorski K, Czerwoniec A, Szatkowska K, Wesoly J and Bluysen HA. In silico simulations of STAT1 and STAT3 inhibitors predict SH2 domain cross-binding specificity. *European journal of pharmacology*. 2013; 720(1-3):38-48.
22. Mathers CD and Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS medicine*. 2006; 3(11):e442.