

## Summary

“The role of GAF, ISGF3 and IRF1 complexes in time-dependent IFN $\alpha$ - and IFN $\gamma$ -activated transcriptional responses and functional overlap”

Interferons (IFN) belong to the superfamily of cytokines discovered as antiviral proteins, but they are also involved in cell proliferation, apoptosis, inflammation, and adaptive immunity.

IFN $\alpha$  and IFN $\gamma$  induce the expression of IFN-stimulated genes (ISG) by phosphorylating Signal Transducer and Activator of Transcription (STAT)1 and/or STAT2 proteins. Transcription factors (TF) as STAT1 homodimers (GAF, IFN $\alpha$  and IFN $\gamma$ ) or STAT1/2 heterodimers (GAF-like, IFN $\alpha$ ) directly activate genes containing the IFN $\gamma$ -activation site (GAS) DNA element. STAT1/2 heterodimer with interferon regulatory factor (IRF)9 (known as ISGF3) in response to IFN $\alpha$  expands the range of regulatory elements that can be targeted by the JAK-STAT pathway to IFN-stimulated response element (ISRE). Both IFNs induce the expression of IRF1, which can regulate the expression of ISGs in response to IFN $\alpha$  and IFN $\gamma$  by binding the ISRE independently. Although the core signaling pathways of IFN $\alpha$  and IFN $\gamma$  are distinct, their transcriptional responses greatly overlap.

To explain this phenomenon, we define the hypothesis that differential binding of ISGF3, IRF1 and GAF/GAF-like complexes to ISRE, GAS and ISRE+GAS composite sites regulates positive feedback of the STAT1, STAT2, IRF9 and IRF1 genes and long-term IFN $\alpha$  and IFN $\gamma$  responsiveness and explains functional overlap.

Using the Huh7.5 cell line as a model, and genome-wide, integrative approach, which combines binding and expression data, we identified 220 overlapping genes induced by IFN $\alpha$  and IFN $\gamma$ . Moreover, we characterized the regulatory elements in their promoter and divided them into GAS-, ISRE-, or composite GAS+ISRE-containing gene groups. Next, we assessed the TFs which can activate each gene group in response to IFN $\alpha$  or IFN $\gamma$ . While being responsive to both stimuli, ISRE-containing genes were higher induced in response to IFN $\alpha$ , because of the potency of ISGF3. On the contrary, GAS-containing genes were more responsive to the GAF complex in response to IFN $\gamma$ . We identified a group of 47 GAS+ISRE composite-containing genes, which by co-recruitment of GAF/GAF-like complexes to the GAS and ISGF3/IRF1 to the ISRE site are highly responsive to both stimulations.

Furthermore, using site-directed mutagenesis, we showed that composite-containing genes could switch between regulatory elements and, to some extent, preserve the expression when one of the sites is mutated. What is more, we proved that in IFN $\alpha$ - and IFN $\gamma$ -treated STAT1-, STAT2-, IRF9-, IRF1- and IRF9/IRF1-mutant Huh7.5 cells confirmed the existence of the alternative signaling pathways, which regulate the ISG expression in the absence of the canonical TFs.

Among composite-containing genes, we identified *STAT1* (composite in distal regulatory element), *STAT2* and *IRF9*, which, together with GAS-containing *IRF1* gene, were up-regulated in response to IFN $\alpha$  and IFN $\gamma$ , thus enhancing the new complex formation and accelerating the positive feedback loop, which controls the long-term IFN-activated cellular response.

Together, we propose a novel, detailed model of IFN $\alpha$ - and IFN $\gamma$ -induced cellular response, which depend on unphosphorylated and phosphorylated ISGF3, GAF/GAF-like complexes and IRF1. In a combinatorial and timely fashion, these complexes mediate the early and prolonged IFN response through GAS-, ISRE- and composite-containing genes. The core set of these genes and activated TFs are commonly regulated by IFN $\alpha$  and IFN $\gamma$  what underlies the transcriptional and functional overlap between these two stimulations. In the heart of this system are placed the composite genes, which on the one hand regulate the *STAT1*, *STAT2* and *IRF9* expression, thus enhancing the new TF complex formation and long-term IFN response, on the other hand, control many diverse ISG expressions shaping the cellular IFN responsiveness and its biological activity. We combined the consecutive activation of GAS, composite- and ISRE containing gene activation with possible engagement in transcriptional, then innate, towards adaptive immune response. To complete this model, we incorporated the potential role of STAT1, STAT2, IRF9 and IRF1 in constitutive ISG expression and rebalancing the homeostasis.

## Streszczenie

### “Rola kompleksów białkowych GAF, ISGF3 oraz IRF1 w rozwoju transkrypcyjnej odpowiedzi indukowanej IFN $\alpha$ oraz IFN $\gamma$ oraz ich funkcjonalnym podobieństwie”

Interferony (IFN) są cytokinami odkrytymi dzięki swoim właściwościom przeciwwirusowym, jednak również są one zaangażowane w inne procesy np. w regulację proliferacji komórek, apoptozy czy komórkowej odpowiedzi zapalnej oraz w funkcjonowanie odporności wrodzonej i nabytej.

IFN $\alpha$  i IFN $\gamma$  indukują ekspresję genów stymulowanych przez IFN (ISG) poprzez fosforylację białek STAT1 i/lub STAT2. Czynniki transkrypcyjne jako homodimery STAT1 (zwane dalej GAF, w odpowiedzi na IFN $\alpha$  i IFN $\gamma$ ) lub heterodimery STAT1/2 (GAF-podobne, tylko IFN $\alpha$ ) bezpośrednio aktywują geny zawierające element DNA - GAS. Heterodimer białek STAT1/2 wraz z białkiem IRF9 tworzą kompleks ISGF3, który w odpowiedzi na IFN $\alpha$  poszerza zestaw elementów regulacyjnych, które mogą być wykorzystane przez szlak JAK-STAT o element ISRE. Oba IFN indukują ekspresję białka IRF1, które może następnie indukować ekspresję ISG w odpowiedzi na IFN $\alpha$  i IFN $\gamma$  poprzez niezależne wiązanie się do ISRE. Chociaż klasyczne szlaki sygnałowe w odpowiedzi na IFN $\alpha$  i IFN $\gamma$  są różne, ich odpowiedzi transkrypcyjne, jak i rola biologiczna w dużym stopniu się pokrywają.

Aby wyjaśnić to zjawisko, postawiliśmy hipotezę, że zróżnicowane wiązanie się kompleksów ISGF3, IRF1 i GAF/GAF-podobnych do miejsc regulatorowych ISRE, GAS i ISRE+GAS powoduje oparte na genach *STAT1*, *STAT2*, *IRF9* i *IRF1* pozytywne sprzężenie zwrotne, a także jest podstawą odpowiedzi komórkowej na IFN $\alpha$  i IFN $\gamma$  wyjaśniając ich funkcjonalne podobieństwo.

Wykorzystaliśmy integrację danych pochodzących z całogenomowych eksperymentów ChIP-seq oraz RNA-seq przeprowadzonych na linii Huh7.5. Dostarczyły one informacji o ekspresji genów oraz interakcji chromatyny z czynnikami transkrypcyjnymi indukowanych stymulacją IFN $\alpha$  lub IFN $\gamma$ .

Dzięki temu podejściu zidentyfikowaliśmy 220 genów indukowanych zarówno przez IFN $\alpha$  i IFN $\gamma$ . Ponadto scharakteryzowaliśmy elementy regulatorowe w ich promotorach i podzieliliśmy je na grupy genów ze względu na posiadanie GAS, ISRE lub kompozytowego

miejsca GAS+ISRE. Następnie zidentyfikowaliśmy czynniki transkrypcyjne, które aktywują każdą grupę genów w odpowiedzi na poszczególny IFN. Mimo, że powyższy zestaw genów był aktywowany przez oba IFN, to geny zawierające ISRE były mocniej indukowane w odpowiedzi na IFN $\alpha$ , ze względu na wiązanie silniejszego czynnika - ISGF3. Z drugiej strony, geny zawierające GAS były bardziej wystymulowane przez kompleks GAF w odpowiedzi na IFN $\gamma$ , niż GAF-like w odpowiedzi na IFN $\alpha$ . Ponadto, zidentyfikowaliśmy grupę 47 genów zawierających kompozytowe miejsce wiązania GAS+ISRE, które poprzez wiązanie każdego rodzaju kompleksów, to jest GAF/GAF-podobnych do miejsca GAS oraz ISGF3/IRF1 do miejsca ISRE, wykazywały porównywalną indukcję w odpowiedzi na obie stymulacje. Co więcej, stosując ukierunkowaną mutagenezę, wykazaliśmy, że geny zawierające kompozyt GAS+ISRE mogą wykorzystywać jedno z elementów regulatorowymi i, do pewnego stopnia, zachowywać ekspresję, gdy drugie z miejsc jest nieaktywne. Dodatkowo, udowodniliśmy, że w traktowanych IFN $\alpha$  i IFN $\gamma$  mutantach komórek Huh7.5 które nie posiadają jednego z białek: STAT1, STAT2, IRF9, IRF1 lub obu IRF9/IRF1 istnieją alternatywne szlaki sygnałowe, które regulują ekspresję ISG w przypadku braku aktywności podstawowych czynników transkrypcyjnych regulowanych w szlaku odpowiedzi na IFN.

Wśród genów zawierających miejsca kompozytowe zidentyfikowaliśmy geny dla białek STAT1 (kompozyt nie w promotorze, lecz w dystalnym elemencie regulatorowym), STAT2 i IRF9, które wraz z genem dla białek IRF1 (zawierającym miejsce GAS w promotorze), ulegają podwyższonej ekspresji w odpowiedzi na stymulację IFN $\alpha$  i IFN $\gamma$ , powodując w ten sposób nagromadzenie nowych komponentów stymulując formowanie się nowych kompleksów czynników transkrypcyjnych i przyspieszając dodatnie sprzężenie zwrotne, które kontroluje przedłużoną odpowiedź komórkową na IFN.

Jako wynik tej pracy proponujemy nowy, szczegółowy model odpowiedzi komórkowej indukowanej przez IFN $\alpha$  i IFN $\gamma$ , które oparty jest zarówno na niefosforylowanych jak i fosforylowanych kompleksach ISGF3, GAF/GAF-podobnych oraz IRF1. W różnych kombinacjach, a także w zmienny w czasie sposób, kompleksy te pośredniczą w kształtowaniu się wczesnej i przedłużonej odpowiedzi na IFN, działając poprzez geny zawierające miejsca wiązania GAS, ISRE oraz kompozytowe. IFN $\alpha$  oraz IFN $\gamma$  indukują wspólną pulę genów przez podobny zestaw czynników transkrypcyjnych, co leży u podstaw ich transkrypcyjnego oraz funkcjonalnego podobieństwa. W sercu tego systemu znajdują się geny posiadające oba miejsca wiązania, GAS i ISRE, które z jednej strony regulują ekspresję STAT1, STAT2 i IRF9, wzmacniając w ten sposób tworzenie nowych czynników transkrypcyjnych i przedłużają

odpowieź na IFN, z drugiej strony kontrolują ekspresję wspólnej puli genów o zróżnicowanych funkcjach, kształtując w ten sposób biologiczną rolę IFN $\alpha$  oraz IFN $\gamma$ . Dodatkowo w modelu tym, następujące po sobie aktywacje konkretnych genów udało nam się powiązać z odpowiedzią biologiczną i tak wczesne geny posiadające GAS odpowiedzialne są za rozpędzenie maszynerii transkrypcyjnej oraz rozbudzenie wczesnej immunologicznej odpowiedzi wrodzonej, podczas gdy późniejsze geny posiadające dodatkowo lub wyłącznie miejsce ISRE, wzmacniają odpowiedź wrodzoną oraz stymulują odpowiedź nabytą, jednocześnie wykazując coraz silniejsze działanie antywirusowe. Aby dopełnić ten model, uwzględniliśmy potencjalną rolę niefosforylowanych STAT1, STAT2, IRF9 i IRF1 zarówno w utrzymywaniu konstytutywnej ekspresji ISG jak i przywracaniu homeostazy.