bioRxiv preprint doi: https://doi.org/10.1101/537696; this version posted February 1, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

# Unexpected implications of STAT3 acetylation revealed by genetic encoding of acetyl-lysine

Yael Belo,[†] Zack Mielko,[‡] Hila Nudelman,[†] Ariel Afek,[‡] Oshrit Ben-David,[†,¶] Anat Shahar,[¶] Raz Zarivach,[§,¶] Raluca Gordan,[‡,‖] and Eyal Arbely[*,†,§,¶]

†*Department of Chemistry, Ben-Gurion University of the Negev, Beer-Sheva, 8410501, Israel*

‡*Center for Genomic and Computational Biology, Department of Biostatistics and Bioinformatics, Duke University, Durham, NC 27708, USA*

¶*The National Institute for Biotechnology in the Negev, Ben-Gurion University of the Negev, Beer-Sheva, 8410501, Israel*

§*Department of Life Sciences, Ben-Gurion University of the Negev, Beer-Sheva, 8410501, Israel*

‖*Department of Computer Science, Department of Molecular Genetics and Microbiology, Duke University, Durham, NC 27708, USA*

E-mail: arbely@bgu.ac.il

Phone: +972 (0)8 6428739. Fax: +972 (0)8 6428449

**Abstract**

The signal transducer and activator of transcription 3 (STAT3) protein is activated by phosphorylation of a specific tyrosine residue (Tyr705) in response to various extracellular signals. STAT3 activity was also found to be regulated by acetylation of Lys685. However, the molecular mechanism by which Lys685 acetylation affects the transcriptional activity of STAT3 remains elusive. By genetically encoding the

bioRxiv preprint doi: https://doi.org/10.1101/537696; this version posted February 1, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

co-translational incorporation of acetyl-lysine into position Lys685 and co-expression with the Elk receptor tyrosine kinase, we were able to biochemically characterize site-specifically acetylated, and simultaneously acetylated and phosphorylated STAT3. We measured the effect of acetylation on the crystal structure, and DNA binding affinity and specificity of Tyr705-phosphorylated and non-phosphorylated STAT3. In addition, we monitored the deacetylation of acetylated Lys685 by reconstituting the mammalian enzymatic deacetylation reaction in live bacteria. Surprisingly, we found that acetylation, *per se*, had no effect on the crystal structure, and DNA binding affinity or specificity of STAT3, implying that the previously observed acetylation-dependent transcriptional activity of STAT3 involves an additional cellular component. In addition, we discovered that Tyr705-phosphorylation protects Lys685 from deacetylation in bacteria, providing a new possible explanation for the observed correlation between STAT3 activity and Lys685 acetylation.

Signal transducer and activator of transcription 3 (STAT3) is a member of the STAT protein family of latent transcription factors that are activated in response to the binding of cytokines, growth factors and hormones to extracellular receptors.[1,2] Structurally, STAT3 comprises an N-terminal domain, followed by a coiled-coil domain, a DNA-binding domain, a Src homolodgy 2 (SH2) domain and a C-terminal transactivation domain.[3–5] According to the canonical Janus kinase (JAK)-STAT pathway, receptor tyrosine-phosphorylation, promoted by the binding of signaling molecules to cell surface receptors, is followed by SH2 domain-mediated binding of STAT3, which is then phosphorylated on Tyr705. Tyrosine phosphorylation enables STAT3 dimerization by reciprocal binding between the SH2 domain of one STAT3 monomer and the phosphorylated Tyr705 (pY705) of the other STAT3 monomer. Dimeric STAT3 then accumulates at the nucleus, where it acts as a transcription factor by binding to specific DNA response elements.[6] That being said, it was found that non-phosphorylated STAT3 is transcriptionally active, and that STAT3 may serve non-canonical roles, such as, for example, in mitochondria.[5,7–11]

STAT3 was found to be constitutively activated in various cancer cell lines and tumor

bioRxiv preprint doi: https://doi.org/10.1101/537696; this version posted February 1, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

tissues, where it promotes tumor cell proliferation, invasion, and migration.[12,13] As a transcription factor that mediates extracellular signaling and gene transcription, STAT3 is also involved in the communication between cancer cells and the microenvironment.[14] In addition, STAT3 plays metabolic, developmental and anti-inflammatory roles.[15] These diverse activities of STAT3 are regulated by various mechanisms, including an array of post-translational modifications, such as phosphorylation (e.g., Tyr705), methylation and acetylation. Specifically, acetylation of Lys685 was suggested to be important for STAT3 dimerization and full transcriptional activity.[16,17] Lys685 acetylation was also found to promote STAT3-DNA methyltransferase 1 interactions, and subsequent methylation of tumor-suppressor promoters.[18] However, the exact biochemical mechanism of Lys685 acetylation-dependent transactivation is not fully understood. According to crystal structures of STAT3, the relatively flexible side-chain of Lys685 is not directly involved in mediating dimerization.[3,5] Therefore, it is not clear if and how the intra-dimer interface is affected by Lys685 acetylation. In addition, Stark and co-workers found that acetylation of Lys685 is important for the transcriptional activity of unphosphorylated, but not of phosphorylated, STAT3.[19] Furthermore, acetylation of Lys685 was suggested to increase in response to cytokine-mediated stimulation,[17,20] although Chen and co-workers found that CD44 can mediate Lys685 acetylation in a cytokine- and growth factor-independent manner.[21]

A widely accepted experimental approach in functional studies of lysine acetylation involves the use of Lys-to-Arg or Lys-to-Gln mutants. In a different experimental approach, acetylation levels are modified by deacetylase inhibitors, or by knockdown or over-expression of acetyltransferases or histone deacetylases. However, the former approach may have unknown structural and functional effects, while the latter may indirectly affect cell physiology. Therefore, we decided to study site-specifically acetylated STAT3 by genetically encoding the co-translational incorporation of Nε-acetyl lysine in response to an in-frame TAG stop codon at position 685 of the STAT3 core domain (residues 128–715).[22–26] Lys685-acetylated STAT3 was produced by co-expression of K685-TAG *stat3* with pyrrolysine amber suppressor

bioRxiv preprint doi: https://doi.org/10.1101/537696; this version posted February 1, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

tRNA (PylT) and evolved acetyl lysine synthetase (AcKRS).[22] The Tyr705-phosphorylated protein was obtained by co-expression of STAT3 with the protein-tyrosine kinase domain of the Elk receptor, as previously demonstrated.[27,28] To express STAT3 site-specifically modified by acetylation and phosphorylation, K685-TAG *stat3* was co-expressed with the amber suppression machinery for co-translational incorporation of an acetylated lysine residue at position Lys685, together with the kinase domain of Elk receptor for post-translational phosphorylation of Tyr705.

The incorporation of acetyl-lysine at position 685 was validated by Western blot analysis using specific antibodies against Lys685-acetylated STAT3 (Supplementary Figure S1A), and trypsin digestion followed by tandem MS/MS (Supplementary Figure S2). Similarly, Tyr705-phosphorylation was dependent on co-expression with the Elk kinase domain, and was validated by immunoblotting using specific antibodies against Tyr705-phosphorylated STAT3 (Supplementary Figure S1B) and MS/MS (Supplementary Figure S2). This experimental setup enabled the expression of four STAT3 variants, namely wild type (WT), Lys685-acetylated (AcK685), Tyr705-phosphorylated (pY705), and Lys685-acetylated+Tyr705-phosphorylated (AcK685+pY705) STAT3. Subsequent comparative studies of WT and AcK685 STAT3 should, therefore, report on effects of acetylation on non-phosphorylated STAT3, while studies comparing pY705 and pY705+AcK685 should report on effects of Lys685 acetylation on phosphorylated STAT3.

The effect of Lys685 acetylation on in vitro DNA-binding affinity was measured by fluorescence anisotropy. A 3'-fluorescein-labeled oligonucleotide derived from the binding site of the α2-macroglobulin (α2M) promoter (5'-AGCAGTTCTGGGAAATCT-3') was incubated with increasing concentrations of the four STAT3 variants, and the fluorescence polarization signal as a function of STAT3 concentration was fitted to Equation 1 (Methods). Similar $K_D$ values were measured for pY705 STAT3 and AcK685+pY705 STAT3 (25±2 nM and 20±3 nM, respectively), demonstrating that Lys685 acetylation had little to no effect on the in vitro affinity of Tyr705-phosphorylated STAT3 to the binding site of the α2M promoter

bioRxiv preprint doi: https://doi.org/10.1101/537696; this version posted February 1, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

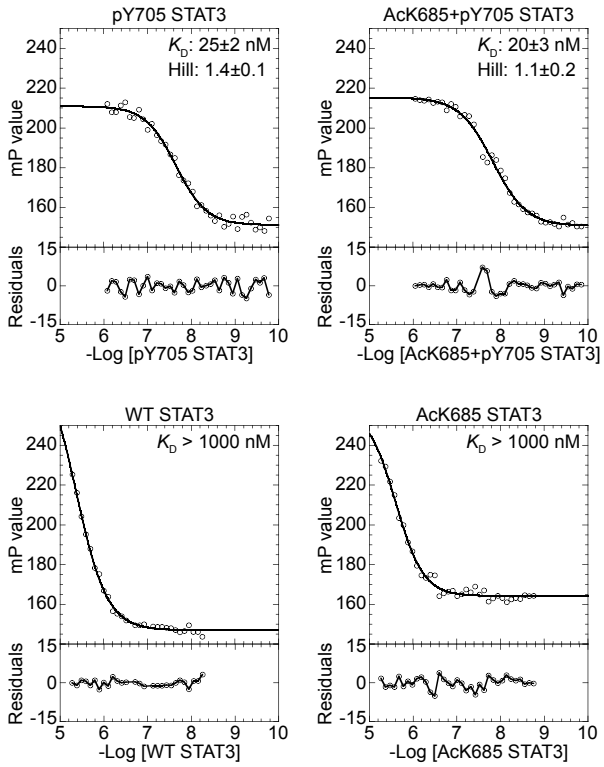



Figure 1: Affinity of STAT3 variants to the binding site of the α2M promoter, measured by fluorescence anisotropy. Increasing concentrations of the indicated STAT3 proteins were incubated with 2 nM of 3'-fluorescein-labeled probe, and data were fitted to Equation 1. Average $K_D$ and Hill slope values are displayed ($n$=3, ±SD). The affinity of non-phosphorylated STAT3 variants was below the detection limit. mP: millipolarization units.

(Figure 1). In addition, the $K_D$ values of WT STAT3 and AcK685 STAT3 were higher than 1000 nM, suggesting that non-phosphorylated STAT3 had low affinity for the binding site of the α2M promoter, with no significant effect of Lys685 acetylation on DNA-binding affinity. Therefore, our data show that the addition of an acetyl group to Lys685, *per se*, had no effect on STAT3 DNA-binding affinity.

It has been demonstrated that post-translational modifications can affect the DNA-binding specificity of transcription factors, and consequently, their transcriptional activity.[30] Therefore, we asked whether Lys685 acetylation affected the DNA-binding specificity of Tyr705 phosphorylated STAT3. To answer this question, we comprehensively characterized the DNA-binding specificities of pY705 STAT3 and AcK685+pY705 STAT3 in an unbiased manner, using the well-established protein-binding microarray (PBM) technology

bioRxiv preprint doi: https://doi.org/10.1101/537696; this version posted February 1, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

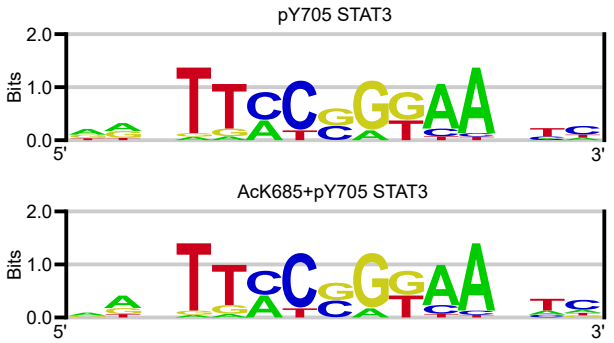



Figure 2: In vitro STAT3 DNA-binding specificity. Position weight matrix logos for pY705 STAT3 and AcK685+pY705 STAT3, obtained from a universal protein-binding microarray containing all possible 10mer combinations on a single chip.[29] The measurements provide two indistinguishable PWM logos, suggesting an identical DNA-binding specificity for pY705 STAT3 and AcK685+pY705 STAT3.

(Figure 2).[31] Briefly, we applied 6×His-tagged purified STAT3 to a dsDNA chip containing all possible 10-mer nucleotides. Next, STAT3 proteins on the chip were labeled with fluorophore-conjugated anti-6×His antibodies, and fluorescent signals were quantified using a high-resolution microarray scanner. Position weight matrix (PWM) logos, the most common way to represent protein-DNA binding specificity,[32] were then generated for the two STAT3 variants using the most enriched DNA sequences (Figure 2). We found that both pY705 STAT3 and AcK685+pY705 STAT3 showed the highest affinity to the same 9 nucleotide-long sequence (5'-TTCC(G/C)GGAA-3'), with essentially indistinguishable PWM logos. These PWM logos strongly suggest that acetylation of Lys685 has no effect on the in vitro DNA-binding specificity of Tyr705-phosphorylated STAT3.

To study the effect of Lys685 acetylation on protein-protein and protein-DNA interactions, we co-crystallized AcK685+pY705 STAT3 with double-stranded DNA (5'-AAGATT-TACGGGAAATGC-3'). The complex of AcK685+pY705 STAT3 with DNA was crystallized in the $P4_1$ space group and the structure was solved to a resolution of 2.85 Å (PDB: 6QHD; statistics of data collection and model refinement are listed in Supplementary Table S1). The asymmetric unit was composed of a STAT3 dimer bound to a double-stranded DNA molecule, in contrast to other STAT3 crystal structures, which contain only one STAT3

bioRxiv preprint doi: https://doi.org/10.1101/537696; this version posted February 1, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

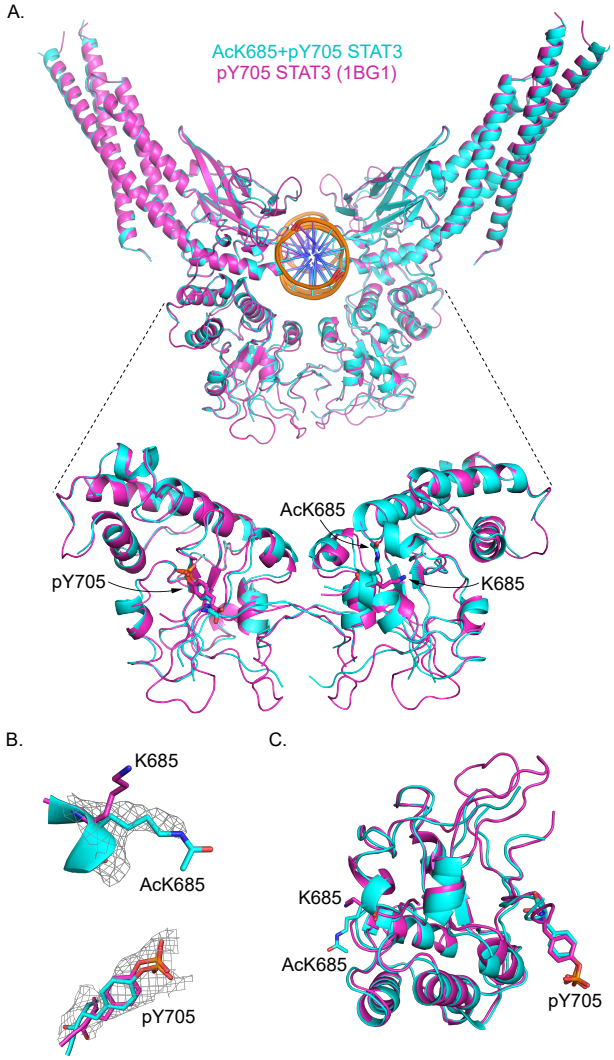



Figure 3: Crystal structure of AcK685+pY705 STAT3 in a complex with DNA. **A.** Superposition of AcK685+pY705 STAT3 (cyan) and pY705 STAT3 (magenta, PDB: 1BG1). Top: overall structure. Bottom: residues 500–715, with residues AcK685, K685, and pY705 displayed in sticks model. **B.** Electron density map around residues AcK685 (top) and pY705 (bottom) ($2F_o - F_c$ at $1.0\sigma$ and $0.7\sigma$ level, respectively), displayed relative to the position of the same residues in non-acetylated pY705 STAT3 (magenta). **C.** Position and orientation of residues AcK685, K685, and pY705 (in sticks model) within the SH2 domains of Lys685-acetylated and non-acetylated STAT3.

monomer with single-stranded DNA within the asymmetric unit.[3,5]

Within each monomer, electron density map for most of our protein model was well defined, yet as with the crystal structure of pY705 STAT3 (PDB: 1BG1),[3] several residues, including loops within the SH2 domain, were poorly defined and consequently were not in-

bioRxiv preprint doi: https://doi.org/10.1101/537696; this version posted February 1, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

cluded in the final model; i.e., the loop connecting α-helices 1 and 2 (185–193), residues 419–427 between β-sheets e and f, a loop at the end of α-helix 7 (536–538), and several residues within the SH2 domain (626–632, 658–665, 689–702, 705, all numbers refer to monomer A). Nevertheless, the overall SH2 domain backbone could be traced, and the relative orientation of the SH2 domains was highly similar to that found in the crystal structure of non-acetylated STAT3 (Figure 3A). Moreover, the overall crystal structure of AcK685+pY705 STAT3 in complex with DNA was essentially identical to the structure of pY705 STAT3 (PDB: 1BG1).[3] Superposition of these two structures revealed a $C_\alpha$ root mean square deviation (RMSD) value of 0.47 Å, indicative of the high similarity between them, with only minor conformational changes (Figure 3A). In addition, according to the electron density map, the positions of AcK685 and pY705 backbone atoms were not affected by Lys685-acetylation (Figure 3B and C). Taken together, we found no significant effect of Lys685 acetylation on the crystal structure of Tyr705-phosphorylated STAT3 in a complex with DNA.

An important aspect of any reversible post-translational modification, such as acetylation, is its regulation by enzymes that catalyze its removal. There are currently 18 known lysine deacetylases (KDACs) in the human genome, 7 nicotinamide adenine dinucleotide ($NAD^+$)-dependent sirtuins (SIRT1–SIRT7) and 11 $Zn^{2+}$-dependent histone deacetylases (HDAC1–HDAC11). To gain insight into the interaction between Lys685-acetylated STAT3 and mammalian KDACs, we used a semi-quantitative assay in bacteria to follow the deacetylation of AcK685 STAT3 and AcK685+pY705 STAT3 (Figure 4).[33] In this assay, a KDAC is co-expressed in *Escherichia coli* (*E. coli*) together with a C-terminal 6×His-tagged acetylated substrate, produced by genetically encoding the incorporation of an acetylated lysine. As such, the mammalian enzymatic deacetylation reaction is reconstituted in bacteria that serve as a 'living test tube'. Deacetylase activity is then evaluated from Western blot analyses, by calculating the ratio between anti-acetyl-lysine immunoblot intensity (proportional to acetylation levels) and anti-6×His immunoblot intensities (proportional to total protein levels). Using this methodology, we monitored the catalytic activity of SIRT1–7, HDAC6 (residues

bioRxiv preprint doi: https://doi.org/10.1101/537696; this version posted February 1, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

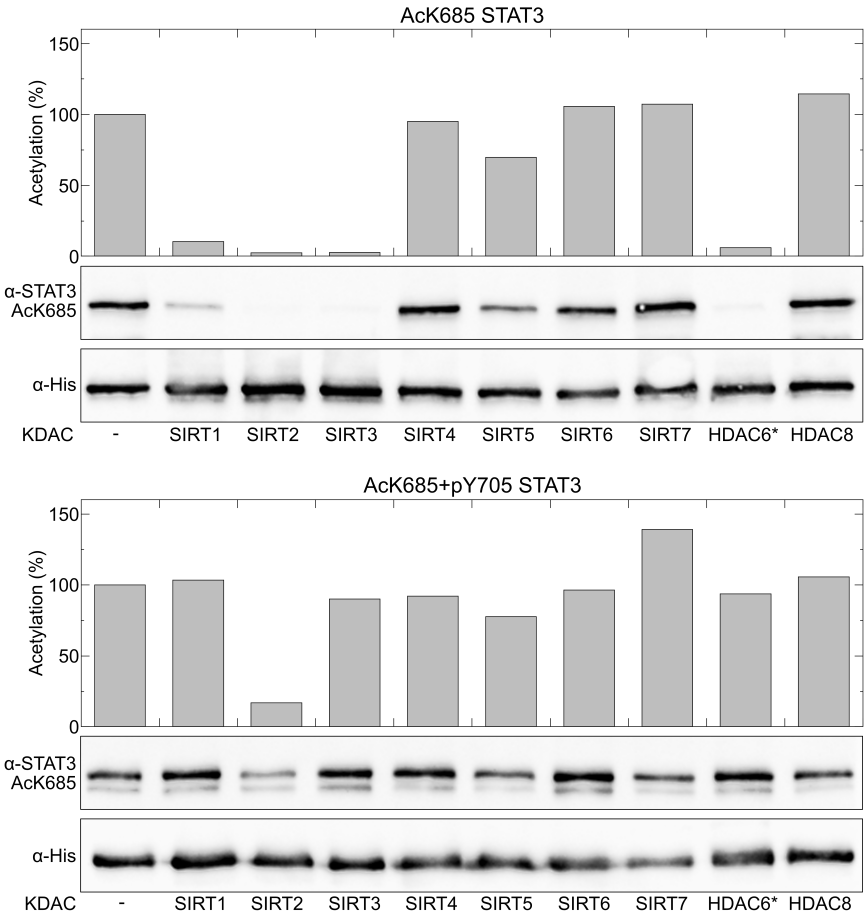



Figure 4: Deacetylation of STAT3 AcK685 in bacteria. Indicated KDACs were co-expressed in bacteria with AcK685 STAT3 (top panel) or AcK685+pY705 STAT3 (bottom panel). Acetylation levels were calculated from the ratio between anti-STAT3 AcK685 and anti-6×His immunoblot intensities and presented relative to the ratio calculated for the negative control (expression without KDAC). In this semi-quantitative assay, a given KDAC was considered active if the averaged ratio, calculated from at least three independent measurements, was less than 50% of the ratio calculated for the negative control. *–HDAC6 was expressed as a truncated protein (residues 479–835), encompassing the second catalytic site.

479–835, marked HDAC6*), and HDAC8. We found that under these conditions, SIRT1–3 and HDAC6* were able to recognize AcK685 STAT3 and hydrolyze the acetyl group from position AcK685 (Figure 4, top panel). However, phosphorylation of position Tyr705 protected AcK685 STAT3 from deacetylation by SIRT1, SIRT3, and HDAC6*; AcK685+pY705 STAT3 could only be deacetylated by SIRT2 (Figure 4, bottom panel). Hence, our data suggest that Tyr705 phosphorylation stabilizes the acetyl group on Lys685 by hindering the

bioRxiv preprint doi: https://doi.org/10.1101/537696; this version posted February 1, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

interactions with potential deacetylases.

To summarize, previous in vivo studies demonstrated that acetylation of Lys685 promotes the dimerization and transcriptional activity of STAT3, suggestive of an acetylation-dependent mode of DNA binding and transcriptional activity. However, in the current study, we found no direct effect of Lys685 acetylation on STAT3 DNA-binding affinity or specificity. These results suggest that the acetylation-dependent STAT3 transcriptional activity observed in vivo may depend on other factors or conditions found in the complex cellular environment, such as additional post-translational modifications, protein-protein interactions, sub-cellular compartmentalization, etc. Thus, our understanding of the role of Lys685 acetylation in STAT3 regulation could benefit from studies in cultured mammalian cells that utilize amber suppression technology to genetically encode the co-translational incorporation of an acetyl-lysine at STAT3 Lys685. However, such studies are still technically challenging.

We found that Lys685 acetylation alone had no effect on the crystal structure of Tyr705-phosphorylated STAT3 in a complex with DNA. Several crystal structures of the STAT3 core domain have been determined, including the structure of non-phosphorylated STAT3 dimer in complex with DNA.[3,5] These crystal structures and the structure presented here demonstrate essentially identical modes of DNA binding by STAT3. A possible explanation for the lack of any acetylation-dependent structural differences is that the crystal structure of STAT3 in complex with DNA represents only one DNA-binding mode, one which is not sensitive to Lys685 acetylation.[34,35]

Using deacetylation assay in bacteria, we found that SIRT1–3 and HDAC6* are capable of recognizing Lys685-acetylated STAT3 and catalyzing the hydrolysis of the acetyl group. Lys685-acetylated STAT3 is a known substrate of SIRT1, which predominantly resides in the nucleus and can also be found in cytoplasm.[36–38] HDAC6 shuttles between the nucleus and cytoplasm and was found to form a complex with the core domain of STAT3.[39] SIRT3 is a major mitochondrial deacetylase, and STAT3 can translocate into mitochondria, where it regulates the mitochondrial respiratory chain via transcription-independent activ-

bioRxiv preprint doi: https://doi.org/10.1101/537696; this version posted February 1, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

ity.[9,10,40] Interestingly, STAT3 residues Lys707 and Lys709 (but not Lys685) were found to be deacetylated by SIRT3 in mitochondria.[11] We found that deacetylation of AcK685 by SIRT1, SIRT3 and HDAC6* was dependent on the phosphorylation state of Tyr705. Considering that Tyr705-phosphorylation promotes STAT3 dimerization and that Lys685 is positioned close to the dimer interface, the observed phosphorylation-dependent protection might be explained by steric hindrance. This observation provokes the question of whether Lys685 acetylation promotes STAT3 dimerization and activation, or rather is the result of STAT3 Tyr705 phosphorylation-dependent dimerization. Moreover, considering the observed effect of acetylation on STAT3 activity in vivo, phosphorylation of Tyr705 may activate STAT3 synergistically by increasing Lys685 acetylation levels.

Taken together, our data show no direct effect of Lys685 acetylation on the DNA-binding mode of STAT3. Numerous in vivo studies reported positive correlation between acetylation and transcriptional activity, dimerization, and nuclear translaocation of STAT3. Generally, in vivo studies are critical to our understanding of cellular processes, yet in vivo measurements of acetylation-dependent protein function are usually based on mutational analyses, gene knockdown or knockout, or use of deacetylase inhibitors. Consequently, any observed acetylation-dependent protein activity may be biased by indirect effects, such as altered histone modifications or transcription. Therefore, our work highlights the advantage of methodologies based on the site-specific incorporation of modified amino acids, as well as the importance of complementing in vivo studies with data obtained from in vitro measurements using homogeneously modified protein samples.

# Methods

## Protein expression

For expression of Lys685-acetylated STAT3, *E. coli* BL21(DE3) cells were co-transformed with a pBK vector for constitutive expression of evolved acetyl-lysine synthetase and the

bioRxiv preprint doi: https://doi.org/10.1101/537696; this version posted February 1, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

appropriate pCDF vector for expression of either non-phosphorylated or phosphorylated STAT3 (see Supplementary Information file for detailed description of plasmid construction). Transformed cells were incubated overnight in 2×TY medium, supplemented with 50 $\mu$g/mL spectinomycin and 50 $\mu$g/mL kanamycin. The next day, overnight cultures were diluted to $OD_{600}$=0.02 into 4 L of 2×TY medium supplemented with the same antibiotics. The cultures were incubated at 37°C until $OD_{600}$=0.3, when they were supplemented with 10 mM acetylated lysine and 20 mM nicotinamide. At $OD_{600}$=0.6, protein expression was induced with 0.5 mM IPTG, and the incubation temperature was reduced to 18°C. After 18 h, the cells were harvested by centrifugation, and the pellet was stored at -80°C. Expression of non-acetylated STAT3 was performed in a similar manner, except that cells were only transformed with the pCDF vector for expression of STAT3 (or co-expression of STAT3 and Elk), and the bacteria were cultured in media supplemented with 50 $\mu$g/mL spectinomycin without nicotinamide or acetylated lysine.

## Fluorescence polarization assay

A DNA probe (5'-AGCAGTTCTGGGAAATCT-3') modified by fluorescein at the 3' end was purchased from Integrated DNA Technologies (Coralville, IA). Fluorescein-labeled single-stranded DNA was mixed with the non-labeled complement sequence at a 1:1.1 ratio in phosphate-buffered saline (PBS). The solution was heated to 90°C and the DNA was annealed by a slow temperature gradient (1°C/min). The solution of 3'-fluorescein-labeled double-stranded probe was filtered (0.2 $\mu$m) and stored in small aliquots at a final concentration of 20 $\mu$M. Fluorescence polarization signals were recorded on a Spark multimode microplate reader (Tecan, Männedorf, Switzerland) in a 384-well format. Aliquots (50 $\mu$l) were prepared by mixing 2 nM DNA probe and the indicated STAT3 variant at increasing concentrations, starting at 0.5 $\mu$M for Tyr705-phosphorylated variants or 5 $\mu$M for non-phosphorylated variants. For each variant, 40 samples were prepared with a 1.3-fold increase in protein concentration between samples, in PBS supplemented with 1 mM DTT, 1 mg/ml

bioRxiv preprint doi: https://doi.org/10.1101/537696; this version posted February 1, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

bovine serum albumin and 50 $\mu$g/ml salmon sperm DNA (R&D Systems, Minneapolis, MN). Fluorescence polarization data were fitted to Equation 1 using KaleidaGraph (Synergy Software, Reading, PA), where $Y$ is the fluorescence polarization signal, $X$ is the STAT3 concentration, $a$ and $b$ are the minimal and maximal fluorescence polarization signals, respectively, $K_D$ is the dissociation constant, and $H$ is the Hill slope.

$$Y = \frac{a + (b - a)}{1 + 10^{(logK_D - X) \cdot H}} \tag{1}$$

## In vitro DNA binding specificity

In vitro DNA-binding specificities were measured using the standard PBM protocol.[31] Double-stranded DNA oligonucleotides were first obtained by primer extension. Next, the microarray was incubated 1 h with blocking solution (2% milk in PBS), followed by a 1 h incubation with PBS-based protein-binding mixture, and a 1 h incubation with Alexa488-conjugated anti-6×His antibodies (1:20 dilution, Qiagen 35310). The array was washed and scanned using a GenePix 4400A scanner (Molecular Devices) at 2.5 µm resolution. Data were normalized as previously described.[31]

## In vitro deacetylation

Deacetylation assay was performed as previously described.[33] Briefly, *E. coli* ΔCobB BL21(DE3) cells were transformed with pBK vector for constitutive expression of an evolved acetyl-lysine synthetase. The transformed bacteria were made competent and co-transformed with a pACYC-Duet vector encoding one of the human deacetylases (Sirt1–7, HDAC6$_{479–835}$, and HDAC8) and a pCDF vector for the expression of AcK685 STAT3, or AcK685+pY705 STAT3. Cells were recovered in 1 mL SOC (37°C, 600 rpm) and incubated for 16-18 h (37°C, 220 rpm) in LB media supplemented with 50 $\mu$g/mL kanamycin, spectinomycin, and chloramphenicol. Overnight cultures were diluted to $OD_{600}$=0.05 into 5 mL of pre-warmed (37°C) auto-induction medium,[41] supplemented with 10 mM AcK and 25 $\mu$g/mL chloram-

bioRxiv preprint doi: https://doi.org/10.1101/537696; this version posted February 1, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

phenicol and spectinomycin each, and 50 $\mu$g/mL kanamycin. Cultures were incubated at 37°C (220 rpm) for 6 h, when the temperature was lowered to 22°C for an additional 42 h. Western blot analysis was performed as described in the Supplementary Methods, using antibodies against the C-terminal 6×His tag and K685-acetylated STAT3.

## Acknowledgement

The structural studies were performed on beamline ID29 at the European Synchrotron Radiation Facility (ESRF), Grenoble, France. We are grateful to the beamline scientists for providing assistance in using this beamline. We would also like to thank CCP4 staff for their contribution to the structure determination during the 1st CCP4/BGU Structure Solution Workshop held at Ben-Gurion University in February, 2018.This work was funded by the Israel Science Foundation (grant number 807/15 to EA) and by the U.S. National Institutes of Health (grant number 1R01GM117106 to RG).

## Supporting Information Available

The following files are available free of charge.

- Supporting information: Supplementary tables, figures and methods.

## References

1. Zhong, Z., Wen, Z., and Darnell, J. (1994) Stat3: a STAT family member activated by tyrosine phosphorylation in response to epidermal growth factor and interleukin-6. *Science 264*, 95–98.

2. Schindler, C., Shuai, K., Prezioso, V., and Darnell, J. (1992) Interferon-dependent tyrosine phosphorylation of a latent cytoplasmic transcription factor. *Science 257*, 809–813.

bioRxiv preprint doi: https://doi.org/10.1101/537696; this version posted February 1, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

3. Becker, S., Groner, B., and Müller, C. W. (1998) Three-dimensional structure of the Stat3$\beta$ homodimer bound to DNA. *Nature 394*, 145–151.

4. Ren, Z., Mao, X., Mertens, C., Krishnaraj, R., Qin, J., Mandal, P. K., Romanowski, M. J., McMurray, J. S., and Chen, X. (2008) Crystal structure of unphosphorylated STAT3 core fragment. *Biochemical and Biophysical Research Communications 374*, 1–5.

5. Nkansah, E., Shah, R., Collie, G. W., Parkinson, G. N., Palmer, J., Rahman, K. M., Bui, T. T., Drake, A. F., Husby, J., Neidle, S., Zinzalla, G., Thurston, D. E., and Wilderspin, A. F. (2013) Observation of unphosphorylated STAT3 core protein binding to target ds DNA by PEMSA and X-ray crystallography. *FEBS Letters 587*, 833–839.

6. Darnell Jr., J. E. (1997) STATs and Gene Regulation. *Science 277*, 1630–1635.

7. Timofeeva, O. A., Chasovskikh, S., Lonskaya, I., Tarasova, N. I., Khavrutskii, L., Tarasov, S. G., Zhang, X., Korostyshevskiy, V. R., Cheema, A., Zhang, L., Dakshanamurthy, S., Brown, M. L., and Dritschilo, A. (2012) Mechanisms of Unphosphorylated STAT3 Transcription Factor Binding to DNA. *Journal of Biological Chemistry 287*, 14192–14200.

8. Yang, J., Liao, X., Agarwal, M. K., Barnes, L., Auron, P. E., and Stark, G. R. (2007) Unphosphorylated STAT3 accumulates in response to IL-6 and activates transcription by binding to NF B. *Genes and Development 21*, 1396–1408.

9. Gough, D. J., Corlett, A., Schlessinger, K., Wegrzyn, J., Larner, A. C., and Levy, D. E. (2009) Mitochondrial STAT3 Supports Ras-Dependent Oncogenic Transformation. *Science 324*, 1713–1716.

10. Wegrzyn, J. et al. (2009) Function of mitochondrial Stat3 in cellular respiration. *Science 323*, 793–797.

bioRxiv preprint doi: https://doi.org/10.1101/537696; this version posted February 1, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

11. Xu, Y. S., Liang, J. J., Wang, Y., Zhao, X.-Z. J., Xu, L., Xu, Y.-Y., Zou, Q. C., Zhang, J. M., Tu, C.-E., Cui, Y.-G., Sun, W.-H., Huang, C., Yang, J.-H., and Chin, Y. E. (2016) STAT3 Undergoes Acetylation-dependent Mitochondrial Translocation to Regulate Pyruvate Metabolism. *Scientific Reports 6*, 39517.

12. Teng, Y., Ross, J. L., and Cowell, J. K. (2014) The involvement of JAK-STAT3 in cell motility, invasion, and metastasis. *JAK-STAT 3*, e28086.

13. Bromberg, J. F., Wrzeszczynska, M. H., Devgan, G., Zhao, Y., Pestell, R. G., Albanese, C., and Darnell, J. E. (1999) Stat3 as an Oncogene. *Cell 98*, 295–303.

14. Yu, H., Kortylewski, M., and Pardoll, D. (2007) Crosstalk between cancer and immune cells: role of STAT3 in the tumour microenvironment. *Nature Reviews Immunology 7*, 41–51.

15. Yu, H., Pardoll, D., and Jove, R. (2009) STATs in cancer inflammation and immunity: a leading role for STAT3. *Nature Reviews Cancer 9*, 798–809.

16. Wang, R., Cherukuri, P., and Luo, J. (2005) Activation of Stat3 Sequence-specific DNA Binding and Transcription by p300/CREB-binding Protein-mediated Acetylation. *Journal of Biological Chemistry 280*, 11528–11534.

17. Yuan, Z.-l. (2005) Stat3 Dimerization Regulated by Reversible Acetylation of a Single Lysine Residue. *Science 307*, 269–273.

18. Lee, H., Zhang, P., Herrmann, A., Yang, C., Xin, H., Wang, Z., Hoon, D. S. B., Forman, S. J., Jove, R., Riggs, A. D., and Yu, H. (2012) Acetylated STAT3 is crucial for methylation of tumor-suppressor gene promoters and inhibition by resveratrol results in demethylation. *Proceedings of the National Academy of Sciences 109*, 7765–7769.

19. Dasgupta, M., Unal, H., Willard, B., Yang, J., Karnik, S. S., and Stark, G. R. (2014)

bioRxiv preprint doi: https://doi.org/10.1101/537696; this version posted February 1, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Critical Role for Lysine 685 in Gene Expression Mediated by Transcription Factor Unphosphorylated STAT3. *Journal of Biological Chemistry 289*, 30763–30771.

20. Ohbayashi, N., Ikeda, O., Taira, N., Yamamoto, Y., Muromoto, R., Sekine, Y., Sugiyama, K., Honjoh, T., and Matsuda, T. (2007) LIF- and IL-6-Induced Acetylation of STAT3 at Lys-685 through PI3K/Akt Activation. *Biological and Pharmaceutical Bulletin 30*, 1860–1864.

21. Lee, J.-L., Wang, M.-J., and Chen, J.-Y. (2009) Acetylation and activation of STAT3 mediated by nuclear translocation of CD44. *The Journal of Cell Biology 185*, 949–957.

22. Neumann, H., Hancock, S. M., Buning, R., Routh, A., Chapman, L., Somers, J., Owen-Hughes, T., van Noort, J., Rhodes, D., and Chin, J. W. (2009) A Method for Genetically Installing Site-Specific Acetylation in Recombinant Histones Defines the Effects of H3 K56 Acetylation. *Molecular Cell 36*, 153–163.

23. Dumas, A., Lercher, L., Spicer, C. D., and Davis, B. G. (2015) Designing logical codon reassignment – Expanding the chemistry in biology. *Chemical Science 6*, 50–69.

24. Chin, J. W. (2014) Expanding and Reprogramming the Genetic Code of Cells and Animals. *Annual Review of Biochemistry 83*, 379–408.

25. Liu, C. C., and Schultz, P. G. (2010) Adding New Chemistries to the Genetic Code. *Annual Review of Biochemistry 79*, 413–444.

26. Chen, H., Venkat, S., McGuire, P., Gan, Q., and Fan, C. (2018) Recent Development of Genetic Code Expansion for Posttranslational Modification Studies. *Molecules 23*, 1662.

27. Becker, S., Corthals, G. L., Aebersold, R., Groner, B., and Müller, C. W. (1998) Expression of a tyrosine phosphorylated, DNA binding Stat3$\beta$ dimer in bacteria. *FEBS Letters 441*, 141–147.

bioRxiv preprint doi: https://doi.org/10.1101/537696; this version posted February 1, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

28. Ren, Z., and Schaefer, T. S. (2001) Isopropyl-beta-D-thiogalactosidase (IPTG)-inducible tyrosine phosphorylation of proteins in E. coli. *BioTechniques 31*, 1254, 1256, 1258.

29. Workman, C. T., Yin, Y., Corcoran, D. L., Ideker, T., Stormo, G. D., and Benos, P. V. (2005) enoLOGOS: a versatile web tool for energy normalized sequence logos. *Nucleic Acids Research 33*, W389–W392.

30. Arbely, E., Natan, E., Brandt, T., Allen, M. D., Veprintsev, D. B., Robinson, C. V., Chin, J. W., Joerger, A. C., and Fersht, A. R. (2011) Acetylation of lysine 120 of p53 endows DNA-binding specificity at effective physiological salt concentration. *Proceedings of the National Academy of Sciences 108*, 8251–8256.

31. Berger, M. F., Philippakis, A. A., Qureshi, A. M., He, F. S., Estep, P. W., and Bulyk, M. L. (2006) Compact, universal DNA microarrays to comprehensively determine transcription-factor binding site specificities. *Nature Biotechnology 24*, 1429–1435.

32. Stormo, G. D. (2000) DNA binding sites: representation and discovery. *Bioinformatics 16*, 16–23.

33. Avrahami, E. M., Levi, S., Zajfman, E., Regev, C., Ben-David, O., and Arbely, E. (2018) Reconstitution of Mammalian Enzymatic Deacylation Reactions in Live Bacteria Using Native Acylated Substrates. *ACS Synthetic Biology 7*, 2348–2354.

34. Sikic, K., Tomic, S., and Carugo, O. (2010) Systematic Comparison of Crystal and NMR Protein Structures Deposited in the Protein Data Bank. *The Open Biochemistry Journal 4*, 83–95.

35. Andrec, M., Snyder, D. A., Zhou, Z., Young, J., Montelione, G. T., and Levy, R. M. (2007) A large data set comparison of protein structures determined by crystallography and NMR: Statistical test for structural differences and the effect of crystal packing. *Proteins: Structure, Function, and Bioinformatics 69*, 449–465.

bioRxiv preprint doi: https://doi.org/10.1101/537696; this version posted February 1, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

36. Nie, Y., Erion, D. M., Yuan, Z., Dietrich, M., Shulman, G. I., Horvath, T. L., and Gao, Q. (2009) STAT3 inhibition of gluconeogenesis is downregulated by SirT1. *Nature Cell Biology 11*, 492–500.

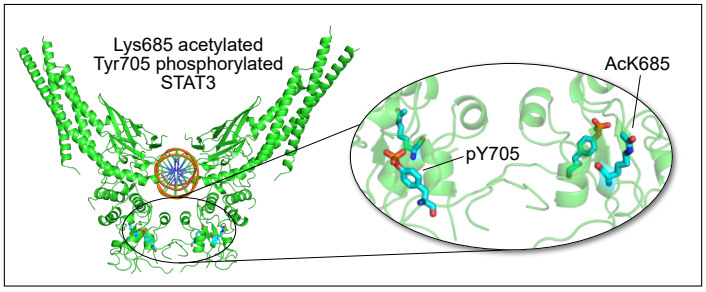
37. Sestito, R., Madonna, S., Scarponi, C., Cianfarani, F., Failla, C. M., Cavani, A., Girolomoni, G., and Albanesi, C. (2011) STAT3-dependent effects of IL-22 in human keratinocytes are counterregulated by sirtuin 1 through a direct inhibition of STAT3 acetylation. *The FASEB Journal 25*, 916–927.

38. Limagne, E., Thibaudin, M., Euvrard, R., Berger, H., Chalons, P., Végan, F., Humblin, E., Boidot, R., Rébé, C., Derangère, V., Ladoire, S., Apetoh, L., Delmas, D., and Ghiringhelli, F. (2017) Sirtuin-1 Activation Controls Tumor Growth by Impeding Th17 Differentiation via STAT3 Deacetylation. *Cell Reports 19*, 746–759.

39. Cheng, F. et al. (2014) A Novel Role for Histone Deacetylase 6 in the Regulation of the Tolerogenic STAT3/IL-10 Pathway in APCs. *The Journal of Immunology 193*, 2850–2862.

40. Tammineni, P., Anugula, C., Mohammed, F., Anjaneyulu, M., Larner, A. C., and Sepuri, N. B. V. (2013) The Import of the Transcription Factor STAT3 into Mitochondria Depends on GRIM-19, a Component of the Electron Transport Chain. *Journal of Biological Chemistry 288*, 4723–4732.

41. Muzika, M., Muskat, N. H., Sarid, S., Ben-David, O., Mehl, R. A., and Arbely, E. (2018) Chemically-defined lactose-based autoinduction medium for site-specific incorporation of non-canonical amino acids into proteins. *RSC Advances 8*, 25558–25567.

bioRxiv preprint doi: https://doi.org/10.1101/537696; this version posted February 1, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

# Graphical TOC Entry