Title: Analysis of the expression and structural differences of STAT3 isoforms on skin cutaneous melanoma and their relation with immune cell infiltration and prognosis

Christian Zevallos-Delgado*

Department of Biochemistry and Molecular Medicine, The George Washington University, Washington, DC 20037

*Corresponding author

Christian Zevallos-Delgado: czevallosd@gwu.edu

ABSTRACT

Skin cutaneous melanoma (SKCM) is one of the most aggressive and resistant forms of melanoma. It has been shown that this type of skin cancer is immunogenic, and that condition has been used to identify new immune-related targets to its treatment. The Signal Transducer and Activator of Transcription 3 (STAT3), a transcription factor that has been involved in the development of immunogenic cancers due to its hyperactivity, is considered a potential target to treat cutaneous melanoma. Also, STAT3 isoforms have shown a regulatory role in its activity. Some of the STAT3 isoforms can be formed by splicing or can be proteolytic isoforms. Previous studies have shown that STAT3β is a negative regulator of STAT3α on SKCM. This study aims to determine the relation of STAT3 isoforms with skin cutaneous melanoma prognosis on cancer patients and study the relation of immune cell infiltration on SKCM with the expression of the STAT3 isoforms. The results showed that six different isoforms of STAT3 are expressed on SKCM (3 types related to STAT3α, one expression of STATβ, and two short versions of STAT3 with a smaller number of nucleotides). The expression of these isoforms showed that patients who express the isoforms STAT3β, STAT3α (missing one amino acid), and STAT3 (without expressing the exon 2) have a significantly better prognosis than the other cases. Also, the deconvolution of RNAseq data from the TCGA-SKCM project showed that these patients have a significantly high population of M0 macrophages, M2 macrophages, and CD8 T cells. The correlation of the expression of the STAT3 isoforms with the population of immune cells demonstrated that STAT3α expression is correlated with the amount of Tcell CD4 and Macrophages M0, and STAT3β with the expression of M2 macrophages. The STAT3 isoform that not expressing the exon2 has a significant correlation with the expression of M1 macrophages. The structural differences among the STAT3 isoforms sowed the conservation of most of the structural domains, with the exception of STAT3β, which structure lack of the TAD domain, a modification which induces an electrostatic change increasing the positive charges in the linker domain, a modification that would be related with the easy loose of dimerization and activity. These results showed that there is a strong correlation between the STAT3 isoforms with the survival of SKCM patients and the infiltration of immune cells on the tumor microenvironment of SKCM. This study proposes the use of STAT3 isoforms as potential targets to develop new therapies against SKCM.

INTRODUCTION

Skin cancer is classified as non-melanoma skin cancer (NMSK) and melanoma skin cancer [6]. 90% of deaths associated with skin cancer are caused by skin cutaneous melanoma (SKCM) [4]. SKCM is developed on humans due to diverse factors such as exposure to UV light, mutations, and other factors [5]. SKCM is curable in the first stage of the disease; however, on metastatic stages, it is potentially dangerous due to its resistance to therapies [7]. The resistance of SKCM is due to its immunogenic behavior and immune cell infiltration [6]. Although SKCM tumors have infiltration of immune cells during their growth, the immune cells cannot stop the tumor growth, suggesting that the SKCM tumor microenvironment (TME) modifies the behavior of immune cells [8]. The Signal Transducer and Activator of Transcription 3 (STAT3) has been identified as a potential target to treat SKCM [9]. STAT3 is continuously activated in SKCM, and its expression is correlated with tumor development, metastasis, and poor patient survival [9]. STAT3 activity is responsible for the activation of immunosuppressive factors such as cytokines and the accumulation of immunoregulatory [10] STAT3 aberrant hyperactivity has been associated with immune suppression modifying immune cells such as macrophages and dendritic cells [11]. The STAT3 protein structure has an amino-terminal domain (NTD), a coiled-coil domain, a DNA-binding domain, a linker domain (LD), an SH2 domain, and a trans-activation domain (TAD) and has to form homodimers to be active [12]. Several STAT3 isoforms have been identified from alternative splicing or proteolytic processing, but only four isoforms have identified to form a protein, these are named STAT3α, STAT3β, STAT3γ, and STAT3δ. STAT3α and STAT3β are isoforms formed by alternative splicing of 23 exons. In the case of STAT3β, due to a frameshift, it has seven different amino acids instead of the TAD domain. STAT3γ and STAT3δ are generated by proteolytic processing of STAT3. Most oncogenic functions have been associated with STAT3α because of its role in the upregulation of genes related to cell proliferation, migration, and survival. On the other hand, STAT3β repress of STAT3α activity on cancer cells [13].

This study aims to determine the presence and relation of STAT3 isoforms on SKCM. As previously described, there is important evidence of the regulatory role of STAT3 isoforms on the aberrant constitutively activated STAT3α, being important to determine the relation of the expression of STAT3 isoforms with the SKCM prognosis and the presence of immune cells infiltrated on SKCM tumors. This is an in-silico study which is expected to introduce a new approach to understand the role of STAT3 isoforms on SKCM development and be considered as potential therapeutic targets against SKCM.

MATERIALS AND METHODS

To understand the effect of the expression of STAT3 isoforms in SKCM, the expression and clinical data was obtained from The Cancer Genome Atlas (TCGA) program through the Broad Institute firehose portal (https://gdac.broadinstitute.org). Since the isoforms are collected under the UCSC nomenclature on TCGA data, to identify only STAT3 isoforms, a mapping file and the sequence of the reads were obtained from the Lineberger Bioinformatics Core (https://webshare.bioinf.unc.edu/public/mRNAseq_TCGA/).

Once the STAT3 isoforms present on SKCM were identified, they were compared with previously human STAT3 isoforms described at Ensembl Genome Browser (https://useast.ensembl.org/index.html). For the matching of the isoforms found in SKCM file and Ensembl isoforms, the sequences were aligned using Clustal Omega [1] in the software Ugene [2]. In the alignment was used the coding STAT3 CDS(CCDS ID: CCDS32656) available at the CCDS Database (https://www.ncbi.nlm.nih.gov/CCDS/CcdsBrowse.cgi), the complete STAT3 genome CDS (exons and introns) available at the UCSC Genome Browser (https://genome.ucsc.edu/), additionally to the exons sequence of the protein-coding STAT3 Isoforms in Ensembl and SKCM STAT3 isoforms. The tool Composed of EMBOSS was used to determine the length of the SKCM STAT3 isoform sequences. BLAST (available at https://blast.ncbi.nlm.nih.gov/Blast.cgi) was used to determine the identity and query cover of the sequence of expressed STAT3 isoforms from SKCM and the STAT3 isoforms found in Ensembl.

The Overall Survival of SKCM patients by STAT3 isoforms expression and the comparison of the expression by cancer stage was done using the clinical data available at the TCGA-SKCM. The analysis of the survival was performed by quartiles comparing Quartile 1 (25% Percent of the Lower expression – Lower Quartile 1) and Quartile 4 (25 % of the Upper Expression – Upper Quartile 4) log-rank test. The analysis of immune cells infiltration of normal RNAseq data was done using CIBERSORTx [3] using the LM22 gene signature and applying 500 permutations. The correlation of the expression of STAT3 isoforms with the immune cells was done using the Spearman correlation method. All the analysis was done using R version 3.6.2 and RStudio. The amino acid sequences for all STAT3 isoforms (including the canonical sequence) were obtained from UniProt (https://www.uniprot.org/), three sequences were reviewed under the same ID (UniProt ID: P40763) and two not reviewed under different ID (UniProt ID: G8JLH9, K7EP08). The sequences were aligned using the same strategy on Clustal Omega by Ugene.

The amino acid sequences were used to analyze the structural and electrostatic differences among the STAT3 isoforms. The STAT3 sequences, three reviewed and one unreviewed (UniProt ID: P40763, G8JLH9) were introduced in the protein structure predictor I-TASSER (available at https://zhanglab.ccmb.med.umich.edu/I-TASSER/), using the condition by default. Once the four structures were obtained, POCASA (http://altair.sci.hokudai.ac.jp/g6/Research/POCASA_e.html) was used to determine the most probable cavities for a binding site based on their volume-depth algorithm. POCASA was used with the stablished parameters on the server. The predicted models obtained in I-TASSER were compared with a STAT3 experimentally obtained model (PDB ID: 6QHD) using PyMOL (version 2.3.3 under Educational License) alignment and RMSD tools. The electrostatic potential iso-surfaces were generated using the APBS plugin in PyMOL in a rank of ± 5.0 kbT/e_c (equivalent to 25.85 mV). All the protein structural models were obtained using PyMOL. The secondary structure analysis was done using PDBsum (http://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbcode=index.html)

RESULTS AND DISCUSSION

To identify the STAT3 isoforms expressed on SKCM, the TCGA-SKCM expression nomenclature was filtered using the STAT3 isoform mapping file. Six STAT3 isoforms were identified from the SKCM data (uc002hzk.1, uc002hzl.1, uc002hzm.1, uc002hzn.1, uc010cyf.1, uc010wgh.1), and they were compared with previously human STAT3 isoforms from Ensembl Genome Browser (Table 1) using sequence alignment together with STAT3 CDS sequence (CDS_STAT3) and STAT3 exon-intron sequence (STAT3_complete_cds) as is shown in Figure 1. The alignment with exon1 (Figure 1.A) confirmed that most of the SKCM isoforms are similar to STAT3-213, except for uc002hzk.1. In the alignment of exon2 (Figure 1.B), uc010wgh.1 had a similarity with STAT3-212, and uc002hzl.1 had a similar sequence with STAT3-201. STAT3-213 was similar to uc002hzn.1, and STAT3-203 was similar to uc002hzm.1. In the result of the alignment of the exon3 (Figure 1.C), the STAT3-212 sequence end prematurely. The alignment of the exon 21-22 (Figure 1.D) showed the similarity of STAT3-203 with uc002hzn.1 (Both sequences had a deletion of 3 nucleotides). Finally, in the alignment of the exon 23-24, uc010cyf.1 showed to be part of both exons (Figure 1. E); however, it also was aligned with the intron 23 (Retained intron was not shown in Figure 1.E), and STAT3-211 showed a similarity with the sequence uc002hzl.1. All the results are shown in the phylogenetic tree, showing the distances and among all sequences. All the results were summarized in Table 2, which according to the BLAST result showed high similarity and according to the alignment results. The isoforms found in the SKCM were renamed as STAT3α – 1(uc002hzl.1), STAT3α – 2(uc002hzn.1), STAT3α - M1(uc002hzn.1), STAT3β (uc002hzk.1), STAT3 – ME2(uc010wgh.1), and STAT3 – RI (uc010cyf.1). In the analysis of SKCM patients based in the lower survival (Quartile 1) and upper survival (Quartile 4) shown in Figure 2, the patients that showed high expression of STAT3β (p-value: 0.056), STAT3α-M1(p-value: 0.0088), and STAT3-ME2(p-value: 0.023) had a significantly better prognosis compared with the patients with the lowest expression (Figure 2. A, F, B). These results are comparable with the proposed by Aigner et al [13], and Yin et al [14]. Also, in the analysis of the expression by stage in SKCM, STAT3α-1 showed a higher expression in the four stages compared with the other isoforms, and STAT3-RI showed the lowest expression in SKCM progression (Figure 3). The result from the deconvolution of RNAseq data from TCGA-SKCM in CIBERSORTx showed the high percentage of Macrophages M0, with an important presence of M2 macrophages and much less presence of M1 phenotype. The high percentage of M2 macrophages have been seen in in situ and in vivo studies, suggesting the role of M2 macrophages as pro inflammatory and pro-oncogenic immune cells [14] . Another predominant presence of cells is CD8 T cells (Figure 4.A). The analysis of the correlation of the expression of STAT3 isoforms with the presence of immune cells in the SKCM TME showed the correlation of STAT3α-1 and STAT3α-2 with CD4 T cells, dendritic cells activated, and Macrophages M0. STAT3α-M1 expression showed a strong correlation with M2 macrophages presence. On the other hand, STAT3β showed a correlation with the expression of T Cells follicular. Some of the correlations have a relation with the previous results that found that STAT3 isoforms increase can regulate the activity of immune cells [14],

however, in this case STAT3β is showing a contradictory behavior being strongly correlated with M2 macrophages which are pro tumorigenic immune cells [14]. STAT3-ME expression showed a significantly strong correlation with B cells of memory and Macrophages M1 (Figure 4.B) The protein sequences of STAT3 isoforms available at UniProt by cross-reference with Ensembl were aligned with Clustal Omega using Ugene. The result showed in Figure 5 showed the similitude of these amino acid sequences, and only the sequence with UniProt ID: P40763 was found on SKCM expression (supporting data is available in Table 2). The results of the alignment showed that the sequence P40763-1 belonged to STAT3α-1 and STAT3α-2. The P40763-2 was the result of the expression STAT3α - M1, P40763-3 was STAT3β, G8JLH9 was STAT3 – ME2, and K7EP08 was not present on the SKCM expression; however, it can be a product of proteolytic splicing. Only STAT3 – RI was not translated to protein. The phylogenetic tree comparing the amino acid sequences (Figure 5 B) showed the similarity of the P40763 sequences; however, the other two sequences are more distant from this distribution.

The structural analysis of the STAT3 isoforms found in melanoma, there is clear conservation in the secondary and tertiary structures (Figures 6, 7, and 8). The STAT3 structures obtained by prediction using I-TASSER showed a high prediction value (Table 3). According to the I-TASSER values, the C-score (confident score) should be between -5 to 2. The Estimated TM-score and RMSD, a value to measure the similarity of the predicted protein with the models used and the root-mean-square-deviation respectively, were obtained having high structural standards for a protein prediction according to Yang & Zhang [15]. PDBsum was used to analyze the secondary structure of the predicted STAT3 isoforms, as is shown in Figure 6 and Figure 8 (A-D). The models showed a conserved secondary structure among them; however, when comparing the amino acids TAD domain between the model P40763-1 and 2 (Figure 6, A and B), there is an increment in beta-turn structure in the P40763-2. The comparison of the P40763-3 secondary structure (Figure 6 C) showed the total loss of the TAD domain. The STAT3 isoforms predicted structures were compared with a STAT3 structure model obtained by X-ray crystallography (PDB ID: 6QHD, Resolution: 2.85 Å) that possesses in its structure two post-translational modification (acetylation in the Lys 685 and phosphorylation in Try 705) (Figure 7). According to the obtained RMSD values, there is not a significant difference between the crystal structure and the STAT3 isoforms. The prediction of the binding cavities present in the STAT3 isoforms would let determine which zone is suitable to bind other structures (small ligands or other proteins). In this case, as it is shown in Figure 8 (E-H) and Table 4, the cavities are changing in the structures. Comparing the isoforms from the model P40763, the loos of the TAD reduce the capacity of binding cavities. However, in the model G8JLH9, there is a bigger cavity due to the absence of the NTD domain (Figure 8 – H). The comparison of the electrostatic potential iso-surface shown the increment of the negative electrostatic potential surface in the STAT3α (P40763 and P40763-2) in the linker domain, and on the other hand, the increment of the positive electrostatic iso-surface STAT3β (P40763-3) in the linker domain (Figure 9). The high differences in the STAT3 isoforms electrostatic iso-surfaces could play an important role in the STAT3 dimerization and activity.

REFERENCES

1. Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W. & Thompson, J. D. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular systems biology, 7(1).
2. Okonechnikov, K., Golosova, O., Fursov, M., & Ugene Team. (2012). Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics, 28(8), 1166-1167.
3. Newman, A. M., Steen, C. B., Liu, C. L., Gentles, A. J., Chaudhuri, A. A., Scherer, F.& Diehn, M. (2019). Determining cell type abundance and expression from bulk tissues with digital cytometry. Nature biotechnology, 37(7), 773-782.
4. Hübner, J., Waldmann, A., Geller, A. C., Weinstock, M. A., Eisemann, N., Noftz, M.& Breitbart, E. (2017). Interval cancers after skin cancer screening: incidence, tumour characteristics and risk factors for cutaneous melanoma. British journal of cancer, 116(2), 253-259.
5. Fink, C., & Haenssle, H. A. (2017). Non-invasive tools for the diagnosis of cutaneous melanoma. Skin Research and Technology, 23(3), 261-271.
6. Paulson, K. G., Lahman, M. C., Chapuis, A. G., & Brownell, I. (2019). Immunotherapy for skin cancer. International immunology, 31(7), 465-475.
7. Pasquali, S., Hadjinicolaou, A. V., Sileni, V. C., Rossi, C. R., & Mocellin, S. (2018). Systemic treatments for metastatic cutaneous melanoma. Cochrane Database of Systematic Reviews, (2).
8. Obeid, J. M., Erdag, G., Smolkin, M. E., Deacon, D. H., Patterson, J. W., Chen, L., ... & Slingluff, C. L. (2016). PD-L1, PD-L2 and PD-1 expression in metastatic melanoma: correlation with tumor-infiltrating immune cells and clinical outcome. Oncoimmunology, 5(11), e1235107.
9. Emeagi, P. U., Maenhout, S., Dang, N., Heirman, C., Thielemans, K., & Breckpot, K. (2013). Downregulation of Stat3 in melanoma: reprogramming the immune microenvironment as an anticancer therapeutic strategy. Gene therapy, 20(11), 1085-1092.
10. Ashizawa, T., Iizuka, A., Maeda, C., Tanaka, E., Kondou, R., Miyata, H. & Hayashi, N. (2019). Impact of combination therapy with anti-PD-1 blockade and a STAT3 inhibitor on the tumor-infiltrating lymphocyte status. Immunology letters, 216, 43-50.
11. Wang, Y., Shen, Y., Wang, S., Shen, Q., & Zhou, X. (2018). The role of STAT3 in leading the crosstalk between human cancers and the immune system. Cancer letters, 415, 117-128.
12. Sgrignani, J., Garofalo, M., Matkovic, M., Merulla, J., Catapano, C. V., & Cavalli, A. (2018). Structural biology of STAT3 and its implications for anticancer therapies development. International journal of molecular sciences, 19(6), 1591.
13. Aigner, P., Just, V., & Stoiber, D. (2019). STAT3 isoforms: Alternative fates in cancer? Cytokine, 118, 27-34.

14. Yin, W., Cheepala, S., Roberts, J. N., Syson-Chan, K., DiGiovanni, J., & Clifford, J. L. (2006). Active Stat3 is required for survival of human squamous cell carcinoma cells in serum-free conditions. Molecular cancer, 5(1), 15.
15. Yang, J., & Zhang, Y. (2015). I-TASSER server: new development for protein structure and function predictions. *Nucleic acids research*, *43*(W1), W174-W181.

Figure Legends

Figure 1. Alignment of nucleotides. Alignment of the STAT3 isoforms. Isoforms found in the SKCM isoforms expression data (uc002hzk.1, uc002hzl.1, uc002hzm.1, uc002hzn.1, uc010cyf.1, and uc010wgh.1), the protein-coding isoforms found in Ensembl Genome Browser (STAT3-201, STAT3-202, STAT3-203, STAT3-211, STAT3-212, and STAT3-213), coding STAT3 CDs (CDS_STAT3), and intro-exon STAT3 CDS(STAT3_complete_cds). **A**. Alignment of Exon 1. **B**. Alignment of Exon2. **C**. Alignment of Exon3. **D**. Alignment of Exon 21-22. **E**. Alignment of Exon 23-24. **F**. Phylogenetic tree of the sequences in the aliment describing the distance among them.

Figure 2. Overall Survival plots of the STAT3 isoforms from SKCM. **A.** STAT3-ME2. **B**. STAT3α-M1. **C**. STAT3α-1. **D**. STAT3-ME2. **E**. STAT3α-2. **F**. STAT3β

Figure 3. Expression of the STAT3 isoforms by stages in SKCM.

Figure 4. Immune cell infiltration on SKCM and correlation with the expression of STAT3 isoforms. **A**. Immune cells deconvolution from RNAseq data obtained from CIBERSORTx p-value <0.05. **B**. Heatmap of the correlation of STAT3 isoform expression with the immune cell's presence in the samples from the deconvolution from RNA seq data.

Figure 5. STAT3 isoforms amino acid sequences alignment. **A**. Amino acid alignment of STAT3 isoforms found in UniProt (UniProt ID: P40763, G8JLH9, and K7EP08). **B**. Phylogenetic tree of the amino acid sequences showing the distances among them.

Figure 6. Secondary structure of the predicted STAT3 isoforms. **A.** P40763 (STAT3α). **B.** P40763-2 (STAT3α-1) **C.** P40763-3 (STAT3β) **D.** G8JLH9. The red rectangles are shown the structural differences among the two STAT3α and the STAT3β on the TAD domain.

Figure 7. Structural alignment of the three-dimension structure of the STAT3 isoforms with the STAT3 crystal structure (PDB ID: 6QHD) (red structure), showing the RMSD values in Å. **A.** P40763 (STAT3α). **B.** P40763-2 (STAT3α-1) **C.** P40763-3 (STAT3β) **D.** G8JLH9.

Figure 8. Identification of the STAT3 domains and binding cavity prediction **A-E.** P40763 (STAT3α). **B-F.** P40763-2 (STAT3α-1) **C-G.** P40763-3 (STAT3β) **D-H.** G8JLH9. **A-D** STAT3 structural domain identification. **E-F** binding cavities identified in the STAT structures.

Figure 9. STAT3 isoforms electrostatic iso-surface done by APBS (Red – Negative -5 kbT/ec and Blue – Positive 5 kbT/ec) **A.** P40763 (STAT3α). **B.** P40763-2 (STAT3α-1) **C.** P40763-3 (STAT3β) **D.** G8JLH9

Tables

Table 1 – Protein coding isoforms found on the Ensembl Genome Browser.

Name	bp	CCDS	Protein	UniProt ID	Exons	Coding Exons
STAT3-201	5047	CCDS32656	770aa	P40763	24	23
STAT3-213	2772	CCDS32656	770aa	P40763	24	23
STAT3-203	2633	CCDS32657	769aa	P40763	24	23
STAT3-211	3319	CCDS59288	722aa	P40763	24	22
STAT3-202	2615	-*	672aa	G8JLH9	23	21
STAT3-212	571	-*	84aa	K7EP08	4	3

*Not available CCDS code for this isoform

Table 2 - Summary of the comparison among STAT3 isoforms found on the TCGA-SKCM and Ensembl Genome Browser data.

UCSC ID	Sequence length (bp)	Comparison Ensembl	BLAST Identity	Query	Notes
uc002hzk.1	4815	STAT3-211	100%	100%	STAT3β
uc002hzl.1	4973	STAT3-201	100%	100%	STAT3α - 1
uc002hzm.1	4948	STAT3-203	100%	100%	STAT3α - M1 (Missing 1 amino acid)
uc002hzn.1	4951	STAT3-201	100%	100%	STAT3α - 2
uc010cyf.1	1980	STAT3-201	5%	100%	STAT3 – RI (Alignment with Exon 23, Intron 23, and Exon 24)
uc010wgh.1	4822	STAT3-201	94%	100%	STAT3 – ME2 (Missing Exon 2)

Table 3 – STAT3 isoforms structure prediction results produced by I-TASSER.

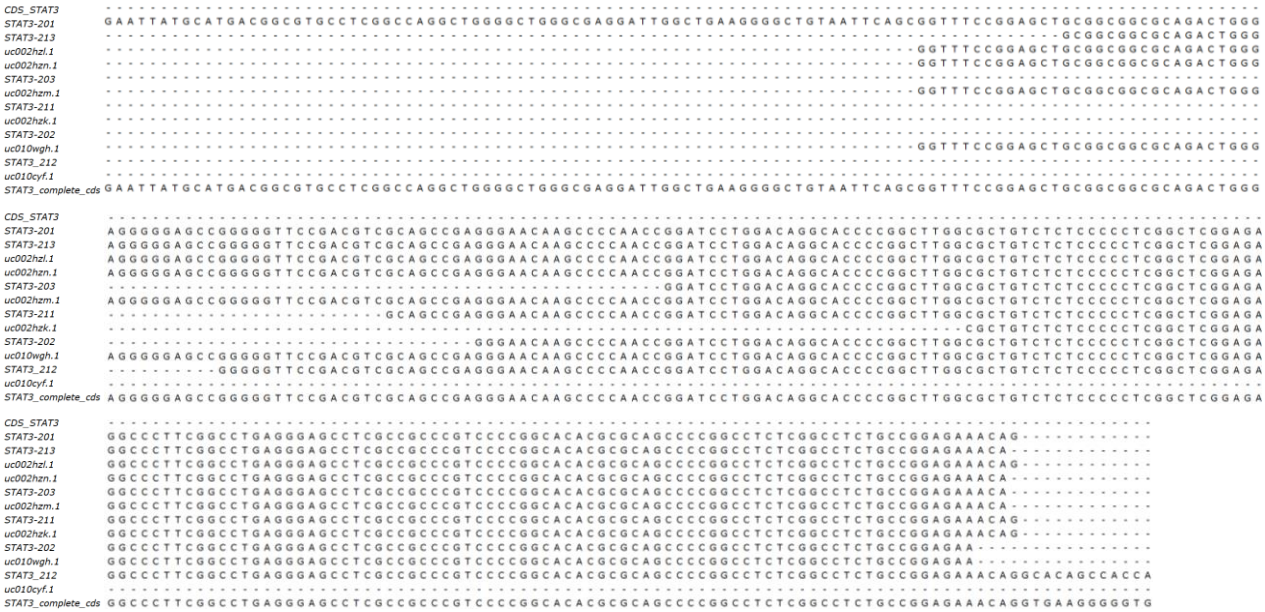
Structure	C-score	Estimated TM Score	Estimated RMSD (Å)
P40763	-0.1	0.70 ± 0.12	8.5 ± 4.5Å
P40763-2	-0.17	0.69 ± 0.12	8.7 ± 4.5
P40763-3	0.78	0.82 ± 0.08	6.4 ± 3.9
G8JLH9	-0.61	0.64 ± 0.13	9.4 ± 4.6Å

Table 4 – Results obtained in the cavity prediction obtained for each predicted STAT3 isoforms using POCASA.

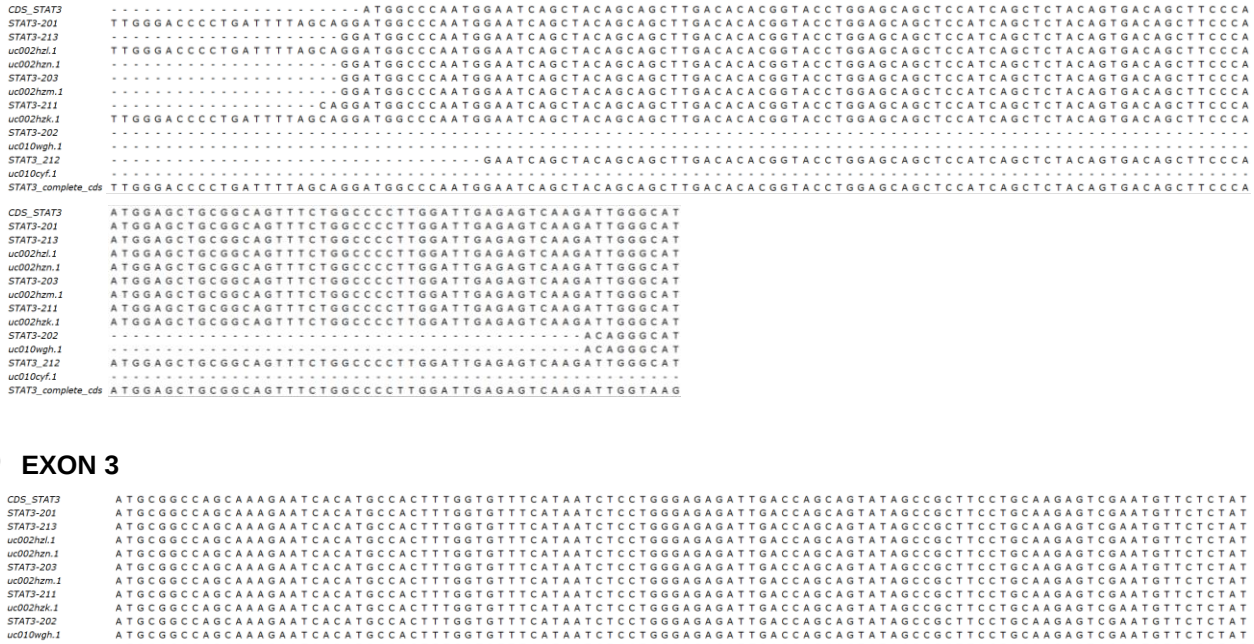
P40763		
Cavity	**Volume(Å)**	**Volume-Depth (VD) (Å)**
1	195	1072
2	79	253
3	75	219
4	68	159
5	57	151
P40763-2		
Cavity	**Volume(Å)**	**Volume-Depth (VD) (Å)**
1	337	2243
2	170	412
3	78	190
4	69	179
5	51	135
P40763-3		
Cavity	**Volume(Å)**	**Volume-Depth (VD) (Å)**
1	214	598
2	188	502
3	82	206
4	74	183
5	51	117
G8JLH9		
Cavity	**Volume(Å)**	**Volume-Depth (VD) (Å)**
1	1382	3927
2	122	226
3	54	167
4	49	120
5	32	81

Figure 1

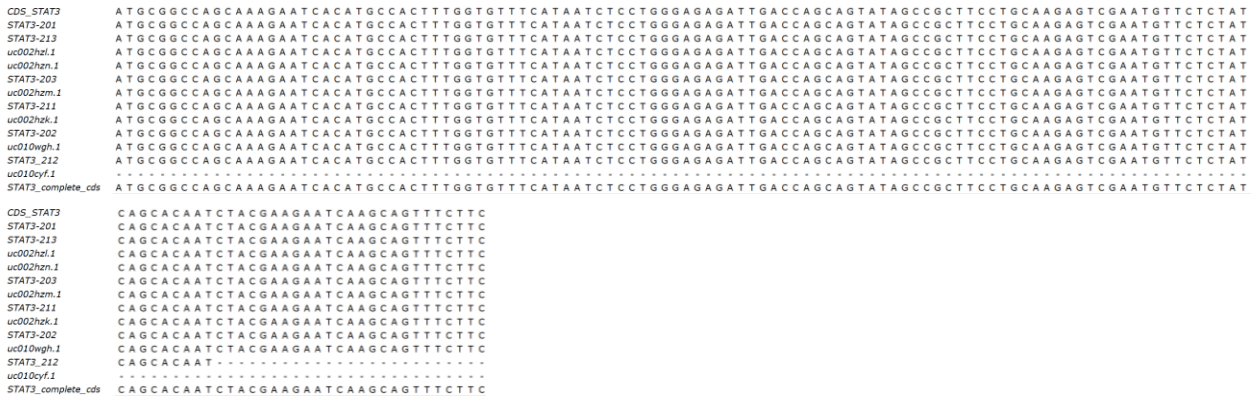
A) EXON 1



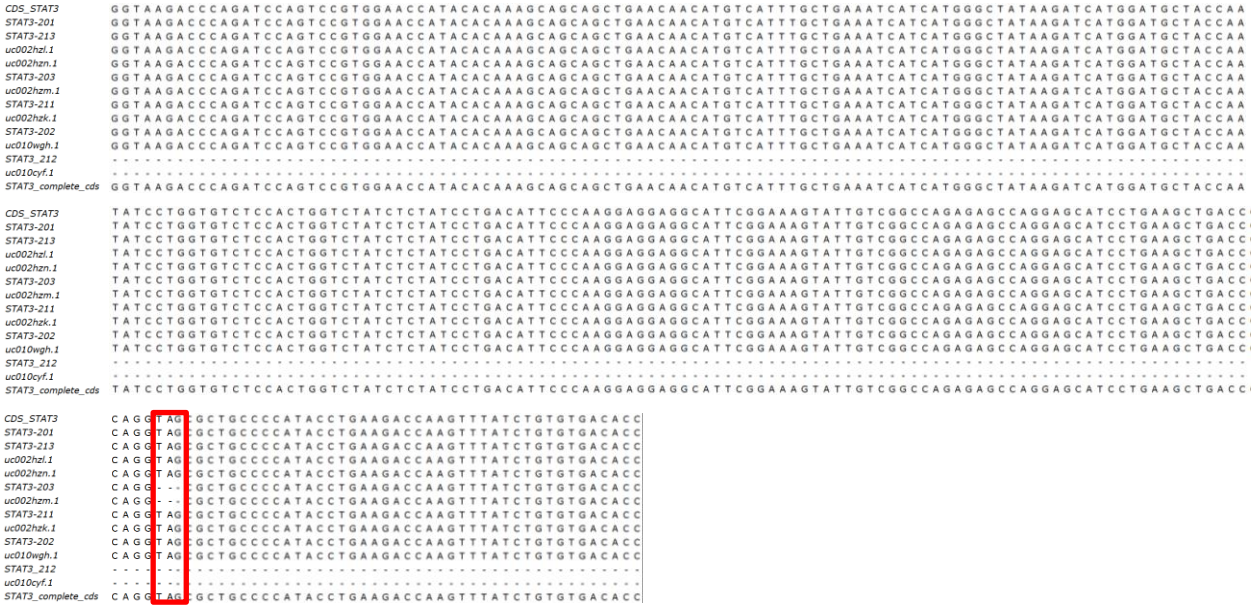
B) EXON 2



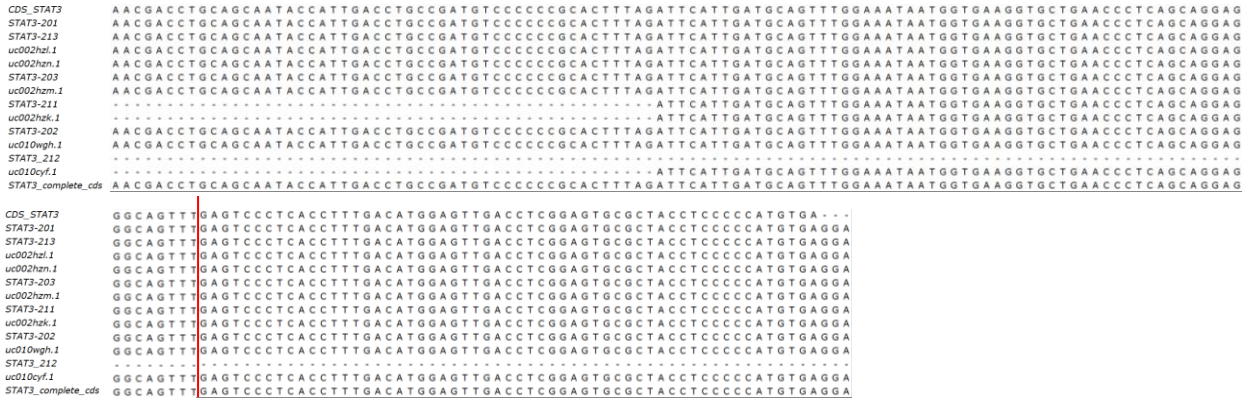
C) EXON 3



D) EXON 21-22



E) EXON 23-24



F)

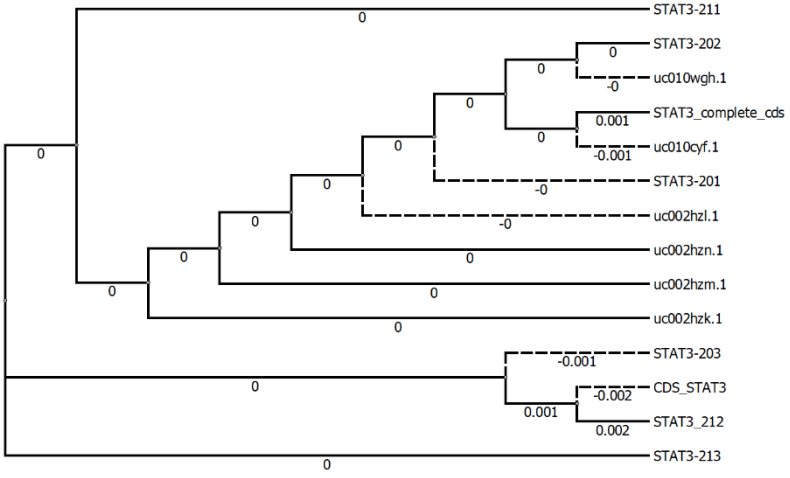
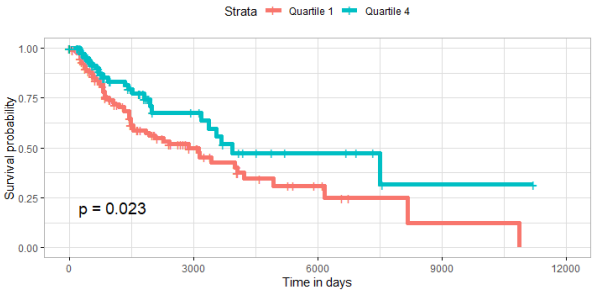
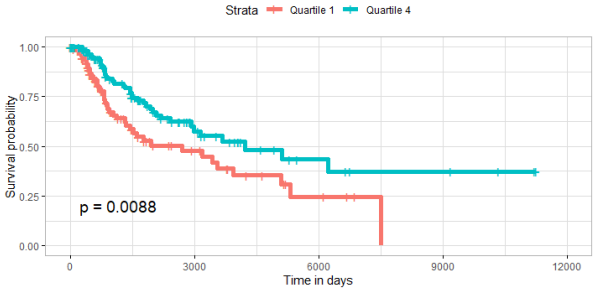


Figure 2

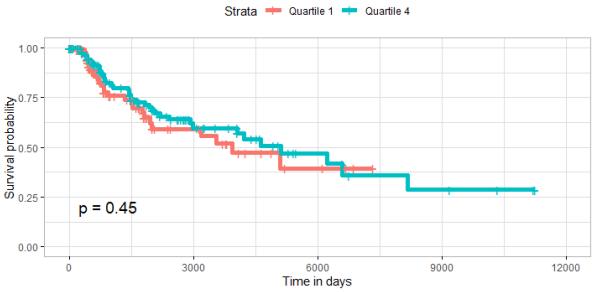
A) STAT3-ME2



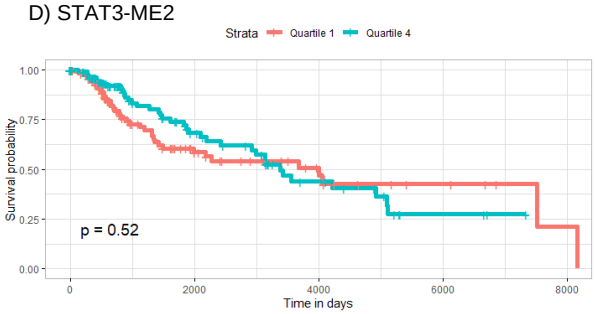
B) STAT3α-M1



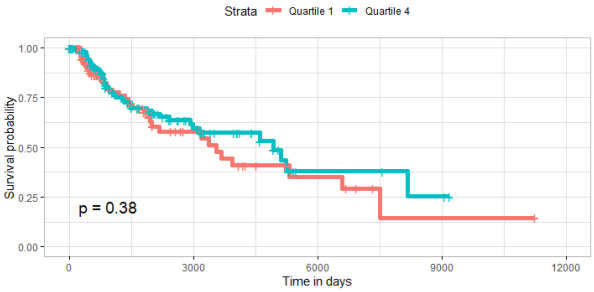
C) STAT3α-1



D) STAT3-ME2



D) STAT3α-2



F) STAT3β

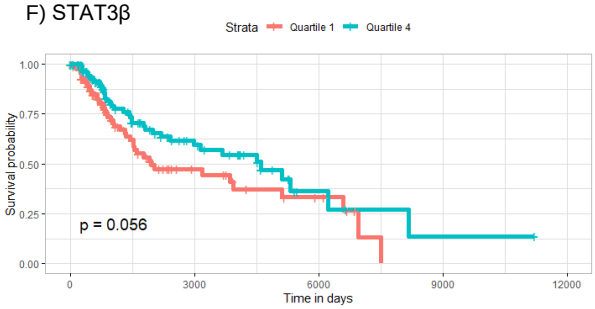


Figure 3

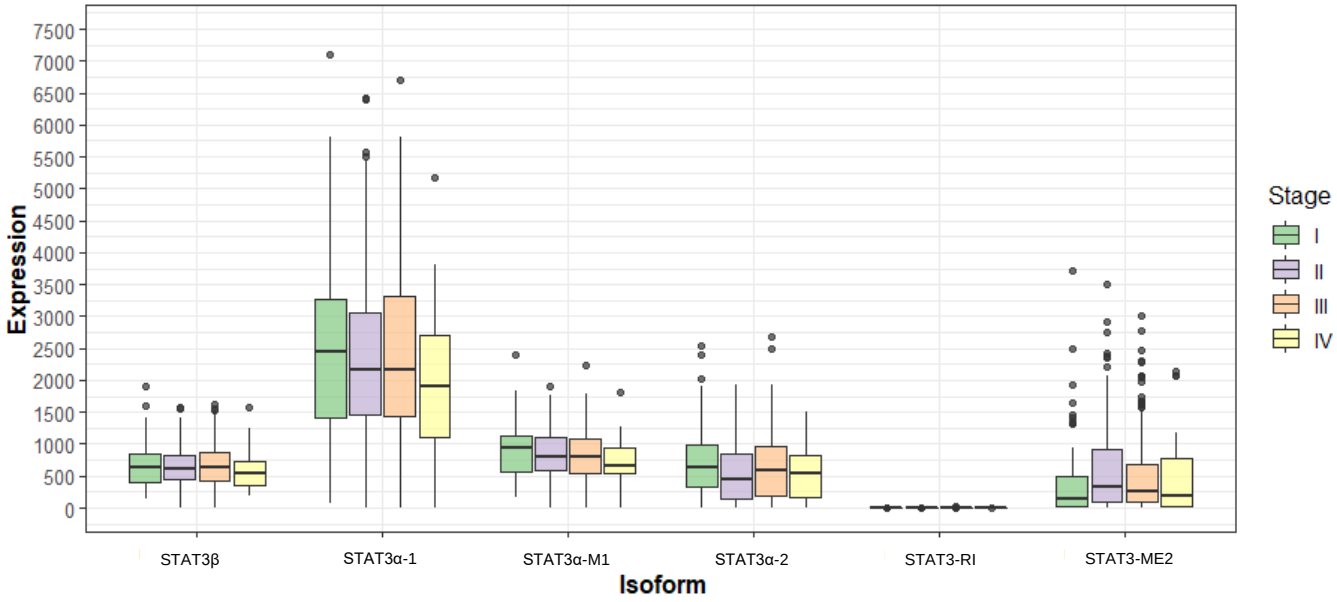

Expression
7500
7000
6500
6000
5500
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
STAT3β
STAT3α-1
STAT3α-M1
STAT3α-2
STAT3-RI
STAT3-ME2
Isoform
Stage
I
II
III
IV

Figure 4

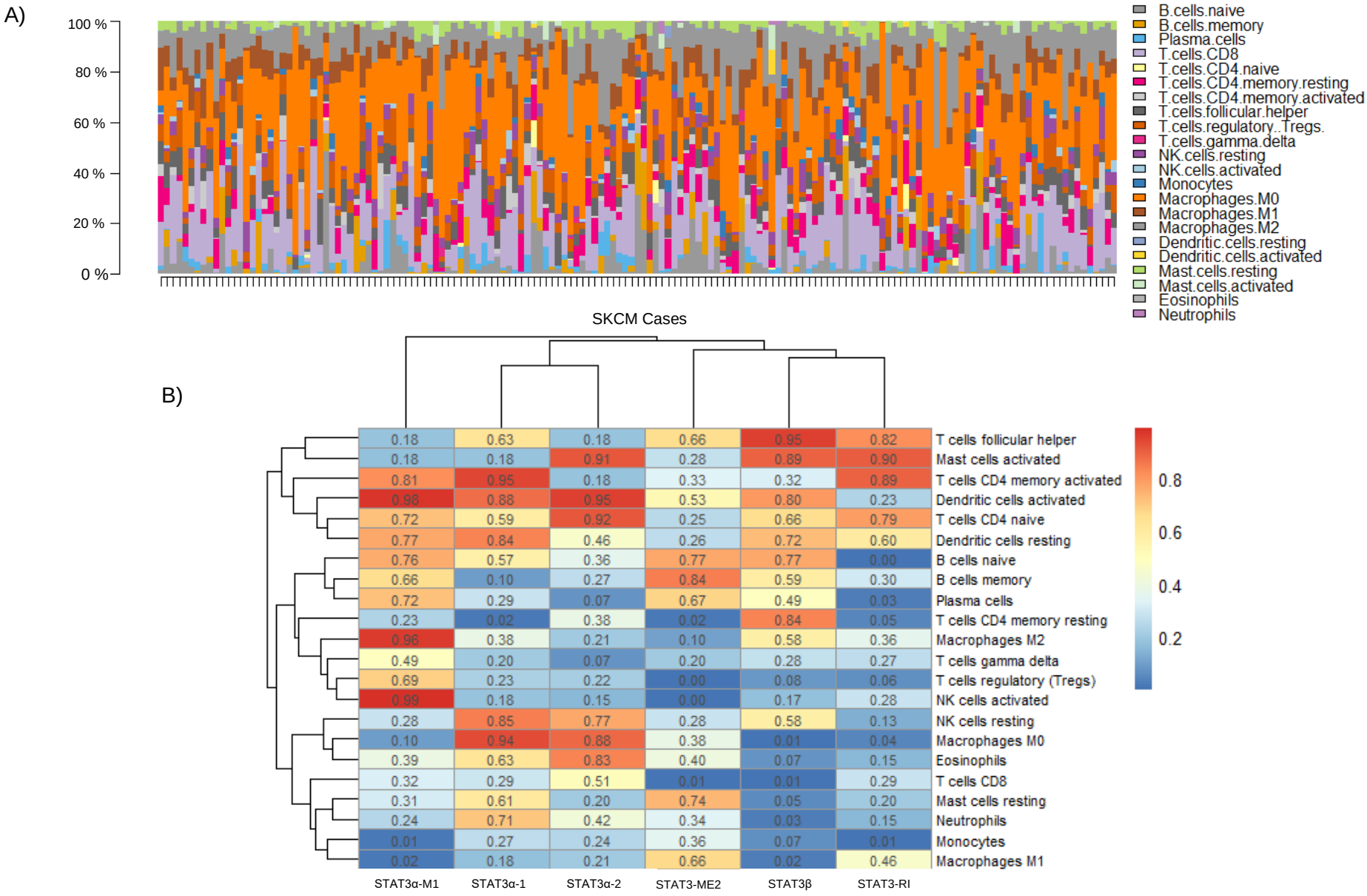
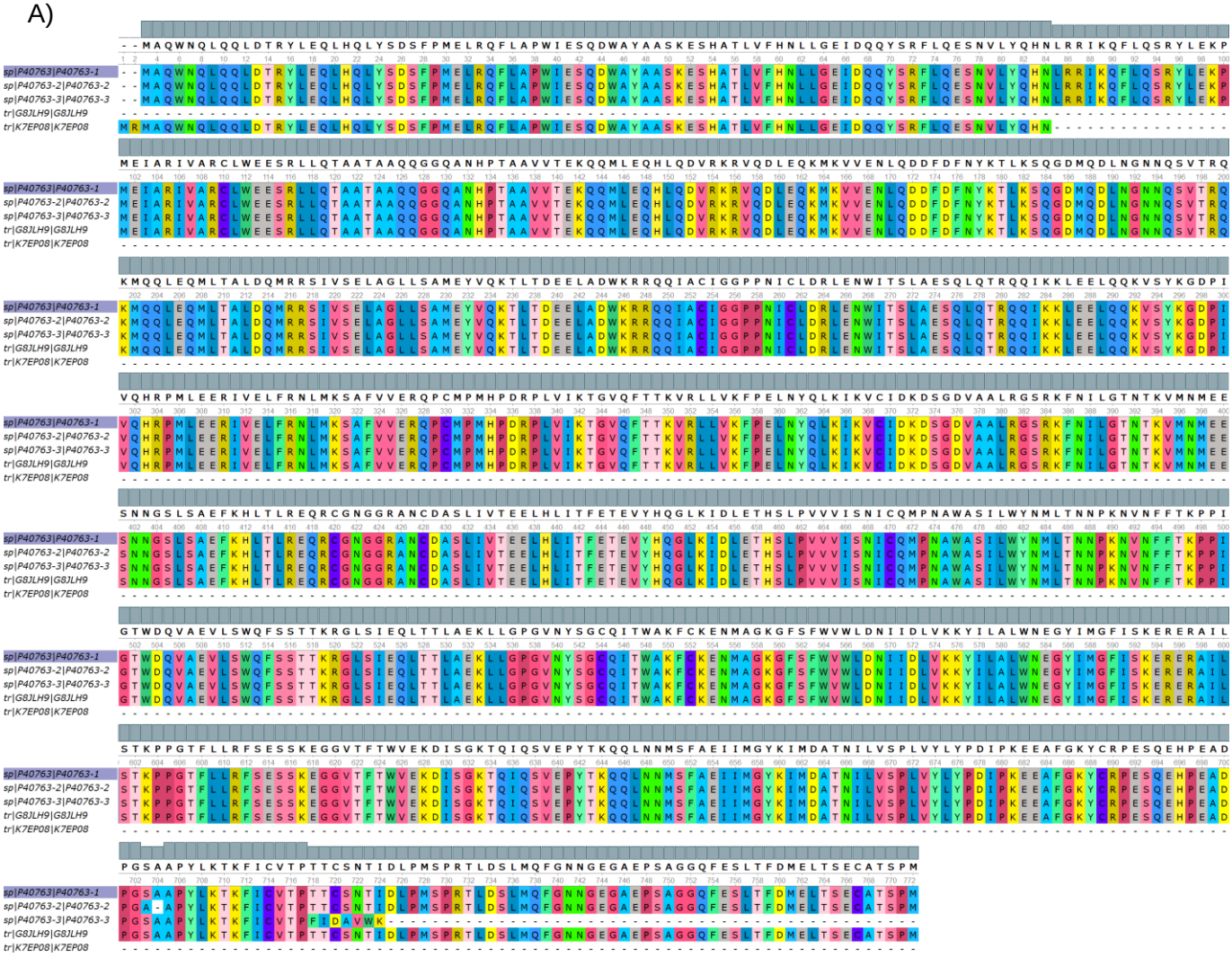

A)
100 %
80 %
60 %
40 %
20 %
0 %
B.cells.naive
B.cells.memory
Plasma.cells
T.cells.CD8
T.cells.CD4.naive
T.cells.CD4.memory.resting
T.cells.CD4.memory.activated
T.cells.follicular.helper
T.cells.regulatory..Tregs.
T.cells.gamma.delta
NK.cells.resting
NK.cells.activated
Monocytes
Macrophages.M0
Macrophages.M1
Macrophages.M2
Dendritic.cells.resting
Dendritic.cells.activated
Mast.cells.resting
Mast.cells.activated
Eosinophils
Neutrophils
SKCM Cases
B)
T cells follicular helper
Mast cells activated
T cells CD4 memory activated
Dendritic cells activated
T cells CD4 naive
Dendritic cells resting
B cells naive
B cells memory
Plasma cells
T cells CD4 memory resting
Macrophages M2
T cells gamma delta
T cells regulatory (Tregs)
NK cells activated
NK cells resting
Macrophages M0
Eosinophils
T cells CD8
Mast cells resting
Neutrophils
Monocytes
Macrophages M1
STAT3α-M1
STAT3α-1
STAT3α-2
STAT3-ME2
STAT3β
STAT3-RI
0.8
0.6
0.4
0.2

Figure 5

A)


sp|P40763|P40763-1
sp|P40763-2|P40763-2
sp|P40763-3|P40763-3
tr|G8JLH9|G8JLH9
tr|K7EP08|K7EP08

B)

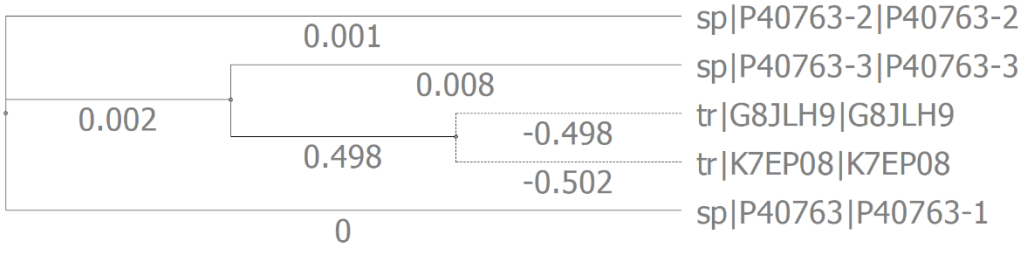

0.001
0.008
0.002
-0.498
0.498
-0.502
0
sp|P40763-2|P40763-2
sp|P40763-3|P40763-3
tr|G8JLH9|G8JLH9
tr|K7EP08|K7EP08
sp|P40763|P40763-1

Figure 6

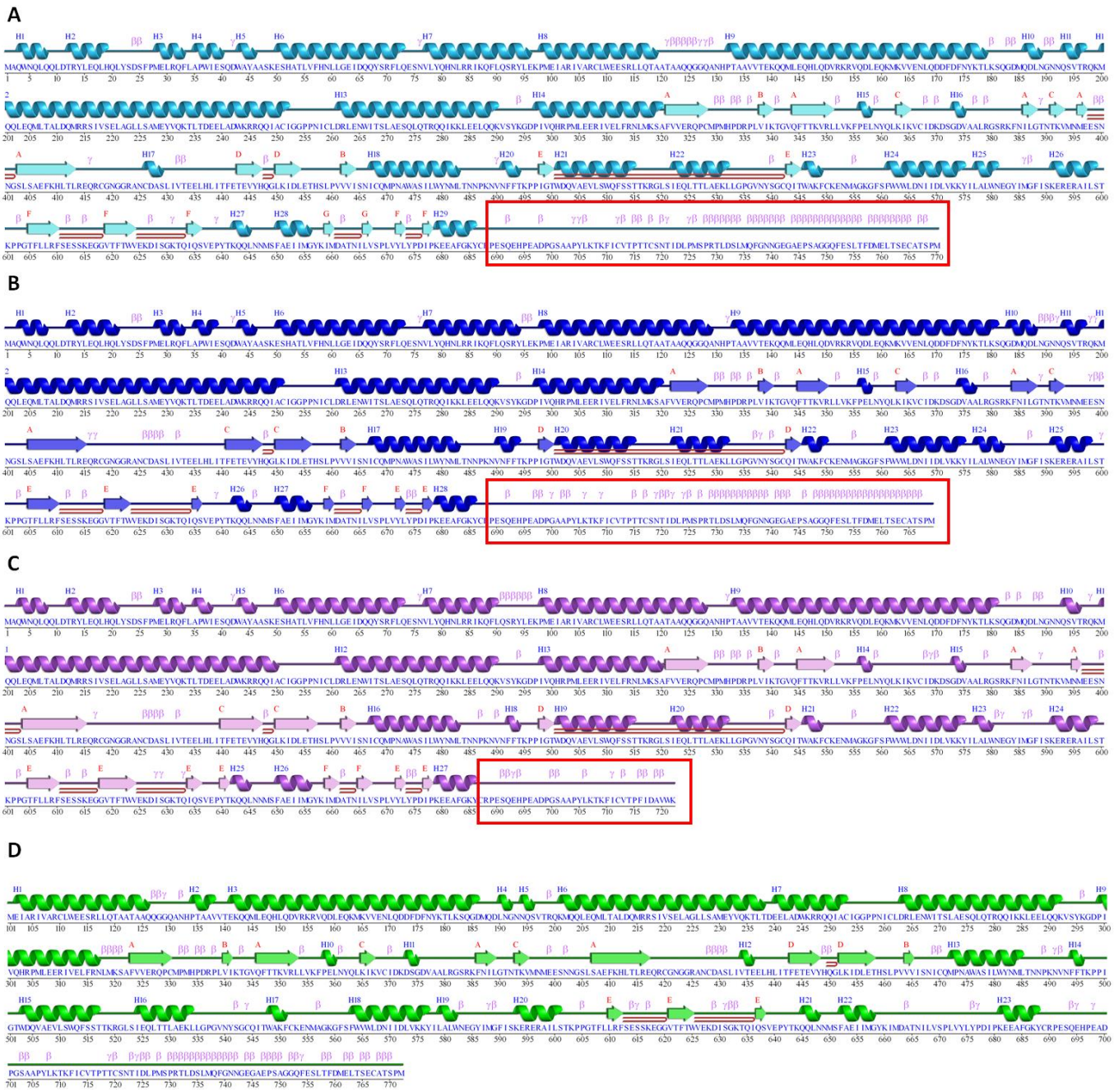


Figure 7

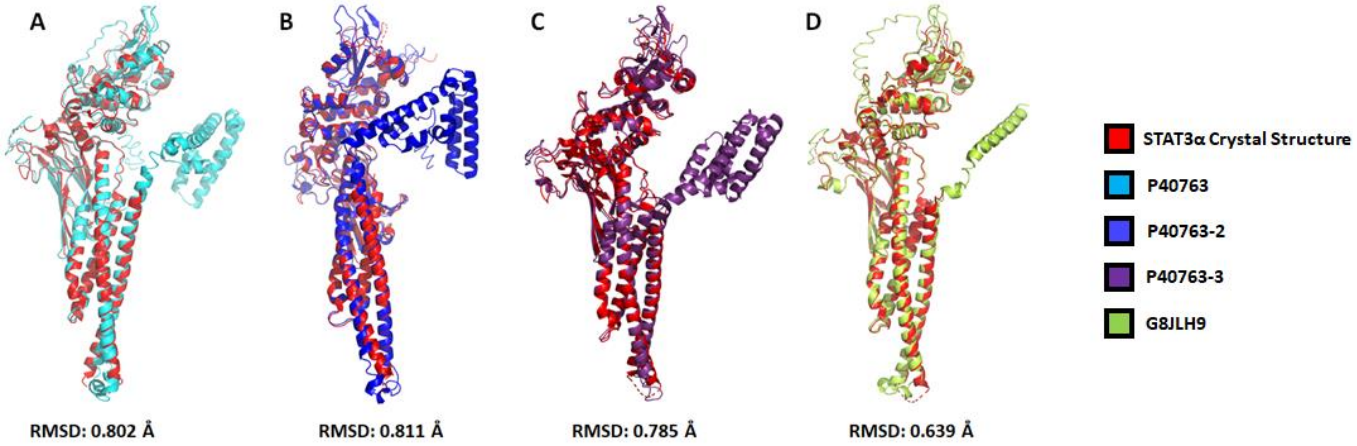


Figure 8

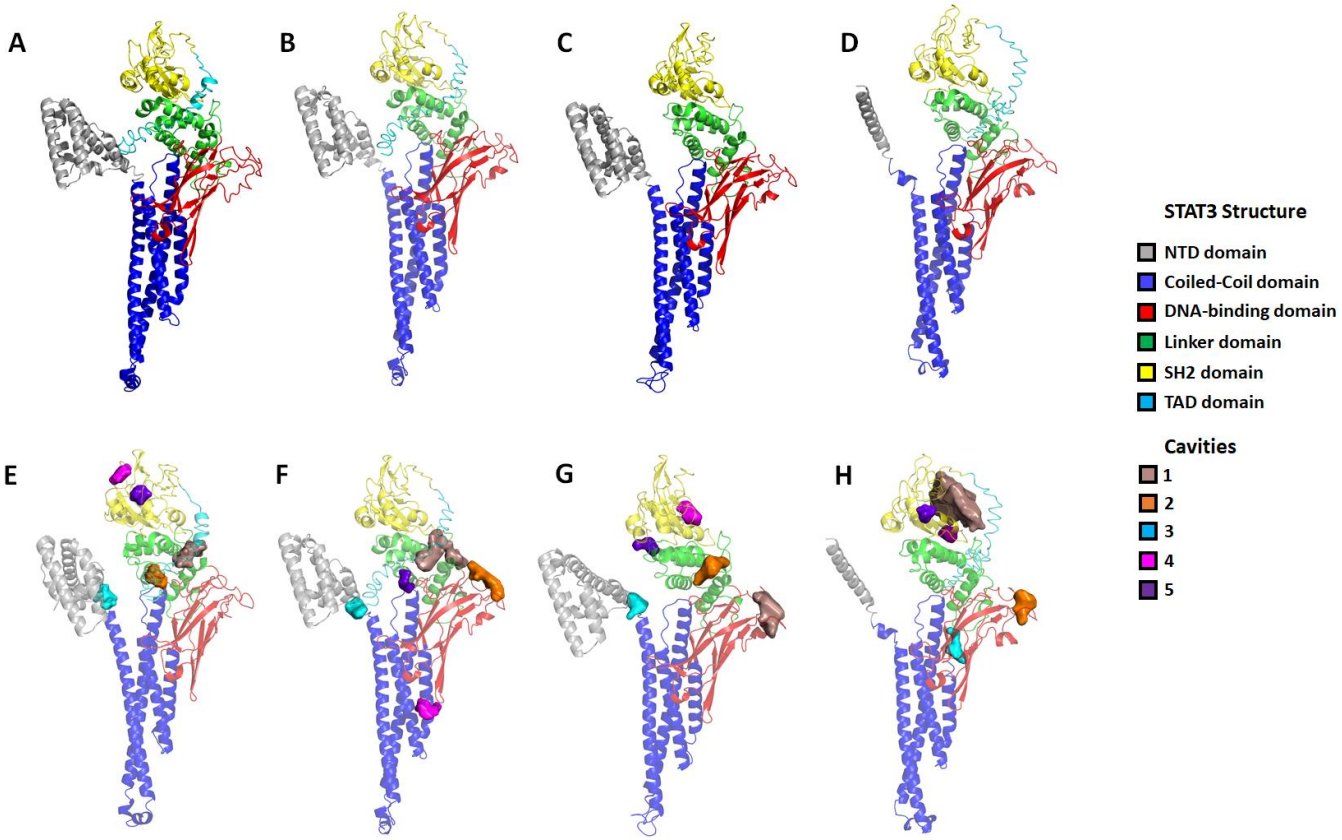
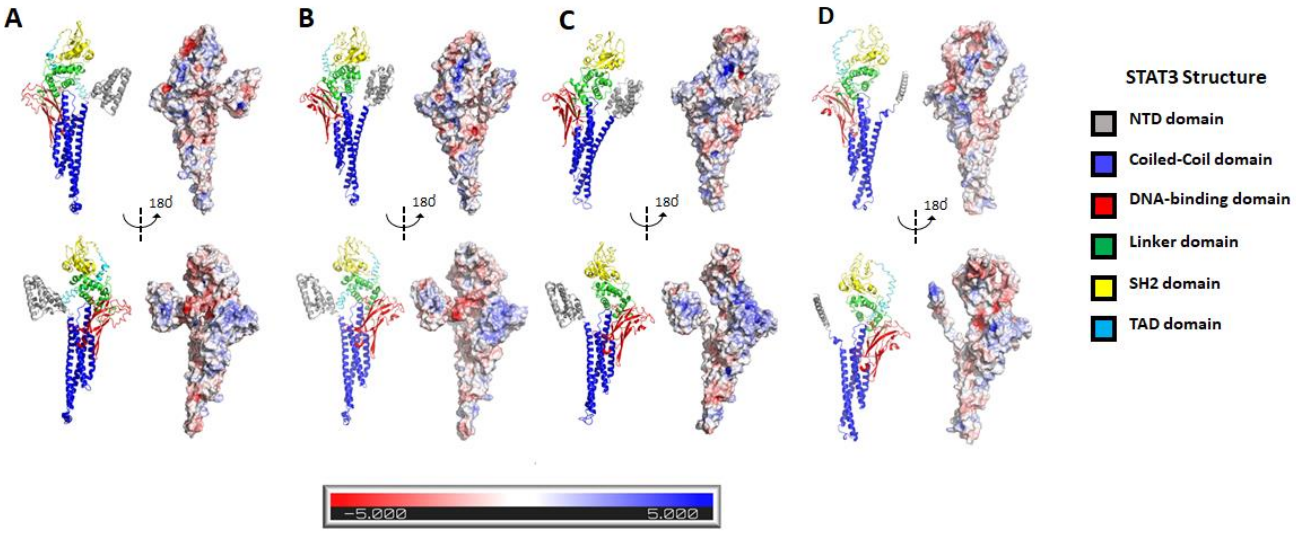


Figure 9


A
B
C
D
180
180
180
180
−5.000
5.000
STAT3 Structure
NTD domain
Coiled-Coil domain
DNA-binding domain
Linker domain
SH2 domain
TAD domain