UNIVERZA V LJUBLJANI

FAKULTETA ZA FARMACIJO

TANJA KOKALJ

MAGISTRSKA NALOGA

ENOVITI MAGISTRSKI ŠTUDIJSKI PROGRAM FARMACIJA

Ljubljana, 2016

UNIVERZA V LJUBLJANI

FAKULTETA ZA FARMACIJO

TANJA KOKALJ

DOLOČANJE VEZAVE POTENCIALNIH ZAVIRALCEV STAT3

MEASUREMENT OF POTENTIAL STAT3 INHIBITORS BINDING AFFINITY

ENOVITI MAGISTRSKI ŠTUDIJSKI PROGRAM FARMACIJA

Ljubljana, 2016

This master's thesis was done at the School of Pharmacy, University College London, Department of Pharmaceutical and Biological chemistry, under supervision of Dr. Andy Wilderspin; and under home mentorship of Assoc. Prof. Dr. Marko Anderluh.

Statement

I hereby declare that I carried out my master's thesis work independently under the mentorship of Assist. Prof. Dr. Marko Anderluh and co-mentorship of Dr Andy Wilderspin.

Tanja Kokalj

Ljubljana, 2016

Graduation commission president: Prof. Dr. Mirjana Gašperlin

Graduation commission member: Assist. Dr. Jasna Omersel

Contents

List of figures

List of equations

List of tables

Abstract

STAT3 (Signal Transducer and Activator of Transcription 3) is a protein, involved in cell growth, motility and regulation of apoptosis. Cytokines, growth factors, hormones and other factors trigger its activation and subsequent dimerization. It has been found constitutively activated in many human cancers, which has led to it being investigated as a target for cancer treatment. There is no available crystal structure of the STAT3 protein with a bound inhibitor, so the structure of the binding site is still unknown. In our research we focused on direct inhibitors of STAT3 that bind to the SH2 domain and inhibit dimerization and the resulting processes. We used fluorescence polarization to test the binding percentages of potential inhibitors. The aims of our work were to improve the method of fluorescence polarization, to screen a small library of potential inhibitors and to try to understand how and where the compounds are binding to the protein.

We observed that using pYLPQ and pYLPQTV 6-FAM-labeled peptides is preferable to pYLKTK. 5% glycerol stabilizes proteins and contributes to higher inhibition, and performing the assays at pH 7,4 or higher (8,5; 9,5) results in higher inhibition values, however, the results are then non-applicable for use in cancer treatment in humans. Dithiothreitol, betamercaptoethanol and tris(2-carboxyethyl)phosphine can all be used as reducing agents. We screened known hit compounds (5, 15-diphenylporphyrin, BP-1-102 and cryptotanshinone) and new potential hit compounds (series EL, STK and STL), using a truncated form of STAT3ß and a mutated protein STAT3 C542S with a mutation from cysteine to serine. DTT and/or DNA were added to the assay solutions. EL compounds reached up to 70% inhibition (100 μM), while STK and STL compounds only reached up to 30%. All compounds, except for BP-1-102 showed lower inhibition with added DTT, meaning –SH groups are important for their binding. When using the mutated protein, BP-1-102 and cryptotanshinone's inhibition stayed the same, 5,15-DPP's fell to 50% and EL group members had lower inhibition, therefore, 5,15-DPP and EL compounds rely somewhat on cysteine 542 for binding. In the STK and STL groups there were big differences between different compounds. These two groups are, however, more heterogenous that EL. This method is not suitable for investigating whether or not the inhibitors affect the binding of STAT3 to DNA. While the compounds we screened are not the optimal ligands for STAT3, the results can be used to modify chemical structures and create stronger inhibitors in the future.

Povzetek

STAT3 (prenašalec signalov in aktivator transkripcije 3) je transkripcijski dejavnik, udeležen v vrsti pomembnih celičnih procesov, kot so rast, diferenciacija, gibljivost in uravnavanje apoptoze. Aktivacijo (fosforilacijo) in posledično dimerizacijo sprožijo številni citokini, rastni faktorji, hormoni in drugi faktorji, poteka pa preko Janus kinaz, Src kinaz in receptorskih tirozin kinaz za rastne faktorje. Dimeri se nato vežejo na importine in se tako premestijo v jedro, kjer se vežejo na gama-aktivirano zaporedje genov (GAS) in aktivirajo transkripcijo. V normalnih pogojih je potrebna hitra aktivacija in inaktivacija. Pri velikem deležu rakavih obolenj (50-90%), npr. raku možganov, dojk, glave, vratu, pljuč, jajčnikov, trebušne slinavke, prostate, levkemiji in drugih hematoloških malignih boleznih, pa je STAT3 konstitutivno aktiviran. Vzrok so lahko različne mutacije, ki povzročajo neprekinjeno signaliziranje, ali pa velike količine citokinov in rastnih faktorjev, ki se nahajajo v tumorjih.

Ker je STAT3 validirana tarča, potekajo raziskave za razvoj zaviralcev, ki bi bili uporabni kot zdravila proti raku. Obstaja več načinov ciljanja STAT3 signalne poti: indirekten način, ki zajema zaviralce tirozinske fosforilacije, ter trije direktni načini: zaviranje DNA vezavne domene; N-terminalne domene in SH2 domene oziroma posledične dimerizacije. V nasprotju z nekaterimi drugimi protitumornimi učinkovinami, imajo usmerjeni STAT3 zaviralci minimalen učinek na zdrave celice, zato je njihova toksičnost manjša.

V magistrski nalogi smo se osredotočili na direktne zaviralce STAT3 - SH2 domene in posledično dimerizacije. Ker kristalna struktura STAT3 z vezanim SH2 zaviralcem še ni znana, ne vemo, kakšna je struktura vezavnega mesta, zato lahko za računalniško podprto iskanje novih potencialnih zaviralcev trenutno uporabljamo samo rešetanje na osnovi ligandov. Ker gre za novo in razmeroma neraziskano področje, tudi ni uveljavljenih rutinskih testov za vrednotenje interakcij ligandov. Zato je cilj našega dela bil izboljšati metodo fluorescentne polarimetrije, s katero merimo delež zaviranega proteina v vzorcu. To smo poskušali doseči z dodatki pufrov (Bis Tris pH 5,7; Tris pH 6,8; 7,4; 8,5), reducentov (ditiotreitol, tris(2-karboksietil)fosfin, betamerkaptoetanol, askorbinska kislina, salicilna kislina) in dveh različnih fluorescentno označenih peptidov (6-FAM-pYLPQ in 6-FAM-pYLPQTV) ter povečati zanesljivost metode, testirati manjšo kemijsko knjižnico potencialnih spojin zadetkov (EL, STK in STL spojine), pridobljeno z virtualnim rešetanjem, ter brez kristalne strukture ugotoviti kraj in način vezave znanih in

potencialnih zaviralcev na STAT3β. V ta namen smo preučevali vpliv dodatkov, kot so reducenti, dvoverižna DNA (M67) z visokoafinitetnim vezavnim mestom za STAT3, ter alternativni protein z mutacijo iz cisteina v serin na mestu 542.

Proteina STAT3β in STAT3β C542S smo izrazili v Rosetta celicah in ju očistili s q-sepharose pretočnimi kolonami. Koncentracijo izoliranega proteina smo preverjali z UV-VIS spektroskopijo in čistoto z SDS-PAGE. Za merjenje deleža zavirane SH2 domene v preiskovanih vzorcih smo uporabili fluorescentno polarimetrijo. Rezultat, delež zaviranega proteina v vzorcu, smo dobili s pomočjo fluorescentno označenega peptida, ki se lahko veže samo na nezavirane molekule proteina. Kot glavni protein smo v eksperimentih uporabili skrajšano obliko STAT3β in z različnimi dodatki spreminjali pogoje za vezavo.

Da bi ovrednotili pomen intramolekularnih disulfidnih vezi, smo izvedli dve seriji poskusov: eno brez dodatka močnega reducenta in eno z dodatkom. Ta eksperiment smo v nadaljevanju nadgradili tako, da smo uporabili alternativni protein s točkovno mutacijo STAT3β C542S, pri katerem je cisteinski ostanek 542 zamenjan s serinskim. Tako smo preverili, ali je prisotnost te –SH skupine nujno potrebna za vezavo zaviralcev na protein. Vsem vzorcem smo odčitali intenziteto fluorescence takoj po pripravi in jih shranili pri 4°C. Ponovno smo intenziteto odčitali po 16 urah, da smo lahko presodili, kakšen je vpliv časa na vzorce in poskušali sklepati na procese, ki se v njih dogajajo.

Ugotovili smo, da sta peptida pYLPQ in pYLPQTV, označena s 6-karboksifluoresceinom, bolj primerna za uporabo od do sedaj uporabljanega pYLKTK ter da dodatek 5% glicerola stabilizira proteine in prispeva k višjim rezultatom meritev. Uporaba pufrov s pH, višjim od 7,4 povzroči višje vrednosti zaviranja, vendar s tem izgubimo primerljivost s fiziološkimi pogoji. Uporaba tris(2-karboksietil)fosfina daje boljše rezultate, kot uporaba ditiotreitola, vendar je veliko dražji, zato smo še naprej uporabljali ditiotreitol, saj je že prisoten v vseh stopnjah ekspresije in izolacije proteinov.

Opravili smo teste na treh serijah novih spojin (EL, STK in STL). Do 70% zaviranje pri 100μM smo opazili samo pri seriji EL (EL56 in EL70), ki jo sestavljajo analogi znanega zaviralca Stattic. Samo tri ostale spojine (STK244106, STL064798 in STL065959) so dosegle več kot 30% zaviranje. Dodatno smo opravili teste tudi na treh že znanih STAT3 SH2 zaviralcih (5, 15-difenilporfirin, BP-1-102 ter cryptotanshinone). Čeprav naj bi se vse spojine vezale na SH2 domeno, so bile vidne očitne razlike v vezavi brez ali z dodatkom reducentov, ter ob uporabi mutiranega proteina STAT3β C542S. Vezava 5, 15-

difenilporfirina je delno odvisna od –SH skupine na mestu 542, ni pa nam uspelo najti nobenega iona, ki bi kot dodatek testni raztopini povečal delež vezave na protein. Na BP-1-102 in cryptotanshinone uporaba reducentov in mutiranega proteina nimata nobenega vpliva, torej vezava ni odvisna od –SH skupin. Serija EL je, tako kot njen strukturni analog, Stattic, uspešna pri zaviranju samo, če v raztopini ni reducentov. Prav tako je vezava manjša, če uporabimo mutirani protein, torej je vezava v veliki meri odvisna od –SH skupine na mestu 542. V serijah STK in STL je prisotnost reducenta in mutiranega proteina vplivala različno na posamezne spojine. V testih z dodatkom DNA se je pri nekaterih vzorcih delež zaviranega proteina povečal, pri nekaterih pa zelo zmanjšal, kar kaže na to, da se je označeni peptid vezal na protein, na katerem je bila vezana tudi DNA. Uporabljena metoda pa nam ne more povedati nič o tem, ali se DNA veže tudi na tiste proteinske molekule, ki imajo že vezan zaviralec; tega ne moremo ugotoviti z označenim peptidom, ker se ta veže samo v prosto vezavno mesto.

Pri presejalnih testih nismo bili tako uspešni, kot smo upali, saj so pri 100 μM samo spojine iz serije EL dosegle relativno visoko zaviranje 70%, tri spojine iz serij STK in STL pa do 30% in so zato primerne za nadaljnje raziskave. To vrednost lahko v prihodnosti povečamo s spremembami pogojev testov ali s strukturno modifikacijo spojin. Še vedno pa bo izziv prilagoditi druge lastnosti spojine, da bo omogočena dostava na želeno mesto delovanja, ter izogibanje vplivu na zdrave celice. Testirane spojine niso bile optimalni ligandi za protein, poleg tega pa ne vemo, ali majhne molekule v tem primeru sploh lahko izpodrivajo velike peptidne ligande. Testi so bili opravljeni na izoliranem proteinu, zato ni realno pričakovati, da bodo ti zaviralci bolj učinkoviti v celičnih testih ali celo po aplikaciji v živalski ali človeški organizem. Rezultati, ki smo jih pridobili pri pH vrednostih različnih od fiziološke, so lahko neuporabni za primerjavo z *in vivo* pogoji, saj pH vpliva na ionizacijo spojin in na strukturo vezavnega mesta; kljub temu pa so uporabni za morebitne študije kristalizacije. Presejalni testi so vseeno dobra in poceni metoda, s pomočjo katere se lahko odločamo, katere spojine bomo testirali na dražjih in bolj zapletenih sistemih.

Pri prihodnjem eksperimentalnem delu bi bilo pred začetkom presejalnih testov smiselno testirati topnost vseh spojin, reaktivnost s preostalimi aditivi v testni raztopini ter morebitno vezavo na polipropilensko ploščico. Razlike v rezultatih med serijami proteinov bi lahko zmanjšali tako, da bi izrazili in prečistili večje količine proteinov, ki bi jih nato skladiščili na -80°C in odmrznili po potrebi.

Key words

Cancer, STAT3 inhibitors, SH2 domain inhibitors, small-molecule inhibitors, fluorescence polarization.

Abbreviations

APS: ammonium persulfate

BSA: bovine serum albumin

DMSO: dimethyl sulfoxide

DNA: deoxyribonucleic acid

DTT: dithiothreitol

EDTA: ethylenediamine tetraacetic acid

FP: fluorescence polarization

HTS: high-throughput screening

IFN γ: interferon γ

JAK: Janus kinase

PAGE: polyacrylamide gel electrophoresis

PCR: polymerase chain reaction

pH: $-\log[H^+]$

PMSF: phenylmethylsulphonyl fluoride

SD: standard deviation

SDS: sodium dodecylsulphate

SH2: Src homology 2

STAT: signal transducer and activator of transcription

TBE: tris borate ethylenediamine tetraacetic acid

TEMED: tetramethylethylenediamine

1. Introduction

1.1. Transcription Factors

Transcription factors are specific regulatory proteins that bind to short DNA sequences (regulatory regions) and regulate transcription of genes in a positive (activators), or negative sense (repressors) (1). They enable transcription by assembling on the promoter, positioning the RNA polymerase, pulling the double helix apart and launching the RNA polymerase (2). There are two ways these proteins can become oncogenic: overexpression or mutation (3).

1.2. Signal Transducers and Activators of Transcription (STATs)

The STAT (Signal Transducer and Activator of Transcription) group of proteins is a family of mammalian transcription factors that was discovered as a result of studying transcriptional activation in response to interferons α and γ (4). The name STAT comes from the dual roles these proteins have: they are transducers of signals through the cytoplasm and they act as transcription factors inside the nucleus (5). The family comprises of 7 members: STAT1, 2, 3, 4, 5A, 5B and 6 (6). They exist in the form of monomers, homodimers (STAT1, 3, 4, 5a, 5b, 6) and heterodimers (STAT1, 2 and STAT1, 3) (7). STAT3 is capable of forming tetramers as well (8).

1.2.1. Function

This family of transcription factors has various roles. They are known to possess a number of highly specific functions throughout the human body, from cell differentiation and proliferation to angiogenesis and apoptosis. They also play a role in innate and adaptive immunity (6). Based on their function, they can be divided into two groups.

The first group consists of STAT2, STAT4 and STAT6. They are only activated by a limited number of cytokines and are involved in IFN-γ signalling and T-cell development (9).

The second group consists of STAT1, STAT3 and STAT5a/b, which are activated by many different ligands and, among other functions, involved in control of cell cycle and apoptosis, which is why their abnormal activation can cause oncogenesis (9). Cancers with

constitutively activated STAT1 or STAT5 in addition to STAT3 are listed in table I. The cancers caused by constitutive activation of STAT3 will be discussed in the next chapter.

Table I: Activation of STAT1 and STAT5 (in addition to STAT3) in human cancers (10,11)

Activation of:	**STAT1+STAT3**	**STAT5+STAT3**
Multiple myeloma	**X**	**X**
HTLV-I-dependent leukaemia		X
Erythroleukaemia	X	X
Acute lymphoblastic leukaemia		X (10)
Chronic lymphocytic leukaemia	X (10)	
Acute myelogeneous leukaemia (AML)	X	X
Chronic myelogeneous leukaemia (CML)		X (only STAT5)
Breast cancer	X	X
Head and neck cancer	X	X
Lung cancer	X (10)	
Brain tumours	X (10)	

Researchers have produced mice model deficient in every one of the STAT proteins to study the roles they play in the body; the findings are summarized in table II. The only mice that did not survive embryogenesis were STAT3-deficient mice, showing the importance of this protein as early as in embryonic development (12).

Table II: STAT proteins and their functions

Name	**Function (involvement in)**
STAT1	Activation of the macrophages, defense from pathogens (13)

STAT2	Defense from viral pathogens (11)
STAT3	Cell growth Cell motility Suppression and induction of apoptosis (14)
STAT4	Regulation of T helper cells differentiation Mediation of responses to IL12 in lymphocytes (15)
STAT5	**STAT5a**: breast development, lactation (11)
	STAT5b: apoptosis T-cell receptor signalling (16) Sexually dimorphic gene regulation patterns in the liver (16)
	Both: T-cell and B-cell development (11) Erythropoiesis, Granulopoiesis (6)
STAT6	Exertion of IL-4 mediated biological responses Differentiation of T-helper 2 cells Expression of cell surface markers Class switch of immunoglobulins (17)

1.3. STAT3

1.3.1. Structure

STAT3, like all other members of the STAT family, consists of six structural domains (18), pictured in figure 1:

- **N-terminal domain** (oligomerization domain) - mediates dimer-dimer interactions in the formation of tetramers, which stabilizes the binding of dimers to DNA (19)
- **coiled-coil domain** – necessary for phosphorylation of the tyrosine and the recruitment of STAT3 to the cytokine (IL-6) receptor (20)
- **DNA-binding domain**
- **linker domain** – interacts with tyrosine-phosphorylated STAT residues to form STAT3 homo- or heterodimers (21)

- **Src-homology-2 (SH2) domain** – recognizes and binds phosphorylated tyrosines (in this case, tyrosine 705 in the transactivation domain), resulting in formation of a reciprocal SH2 domain-pTyr interaction – a dimer (19)
- **transactivation domain** – contains the tyrosine 705 residue and a serine 727 residue, the latter of which needs to be phosphorylated to achieve maximal transcriptional activity (22)

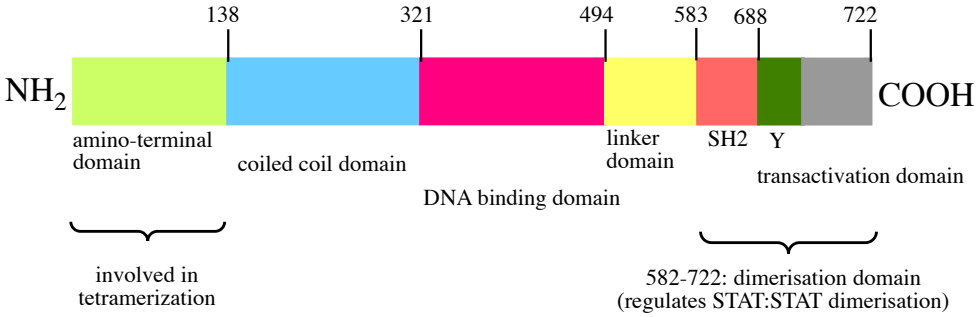


Figure 1: STAT3ß functional domain structure. Adapted from source (23)

The first crystal structure of STAT3ß (bound to DNA) with the resolution of 2,25 Å (24) was published in 1998 and is pictured in figure 2. As of May 2016, no crystal structure exists for STAT3 bound to an inhibitor.

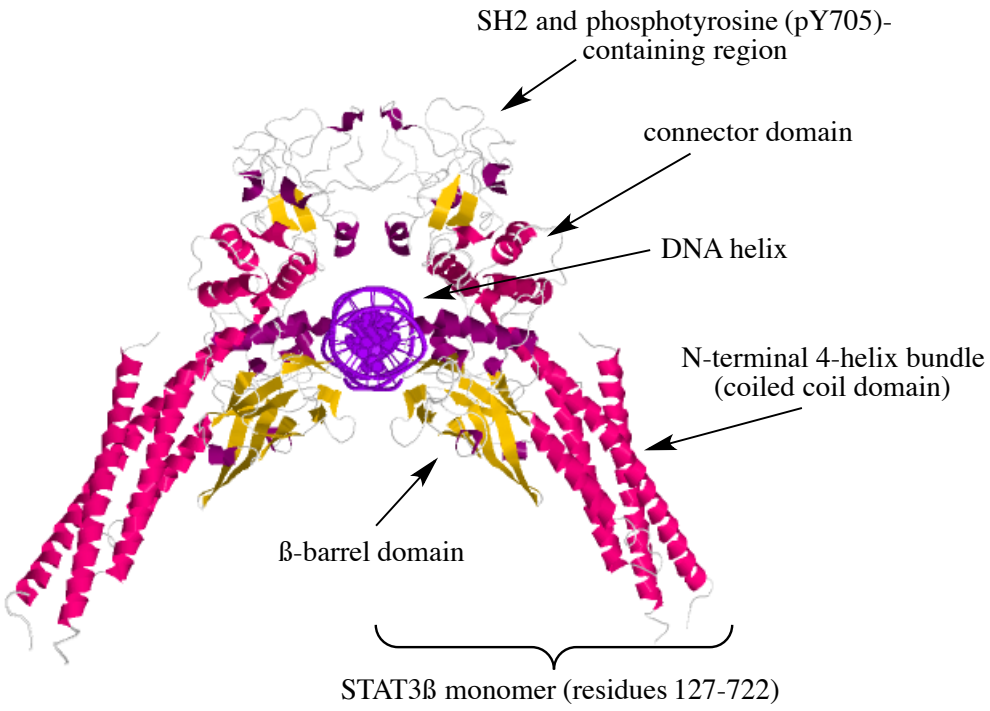


Figure 2: Ribbon diagram of STAT3ß homodimer-DNA complex (top view), PDB ID: 1BG1 (24)

1.3.2. Isoforms

STAT3 exists in four isoforms:

- STAT3α (full-length isoform), expressed in most cells
- STAT3ß, a C-terminally truncated form of STAT3α; 55 C-terminal amino acids of STAT3α are replaced by a unique 7-amino acid sequence (25)
- STAT3γ, a C-terminally truncated form of STAT3α, derived by limited proteolysis (26)
- STAT3∂, a putative isoform, expressed in the early stages of granulocytic differentiation (27)

STAT3α and ß are the two dominant types.

Isoforms α and ß have been intensely studied and it has been determined that they have different functions in the body. STAT3α can activate transcription on its own. STAT3ß, however, lacks the transcription activation domain. They can both activate transcription through participation in the formation of higher hierarchy complexes, which can then recruit coactivators. There might even be some specific target genes that are responsive specifically to STAT3ß (28). It is also now known that STAT3ß is not a dominant negative factor, as it was previously thought (28). Intracellular dynamics are different between the two isoforms. Compared to STAT3α, STAT3ß remains in the nucleus for longer periods of time, which is caused by its unique C-terminal sequence of 7 amino acids. When inside the nucleus, STAT3ß's mobility is reduced after ligand stimulation (29). The function of STAT3γ is not yet clear (27). Figure 3 shows the sequences of STAT3α, STAT3ßtc and a further truncated form, STAT3ß 688stop, which was used for screening in this research group.

The STAT3α sequence is 770 residues long (30) and pictured in figure 3. The STAT3ßtc sequence, which runs from residue 127 to 722 and was used in previous experiments, is highlighted yellow. This is an N-terminally truncated version of STAT3ß, different from the wild type. The shorter, STAT3ß 688stop sequence (127-688), which we used for screening, is underlined.

Also pictured:

- tyrosine 705 residue (red)

- LKTK sequence (residues 706-709), which was used in phosphopeptide design (green)
- cysteine 542, mutated to serine in C542S 688s protein (blue)

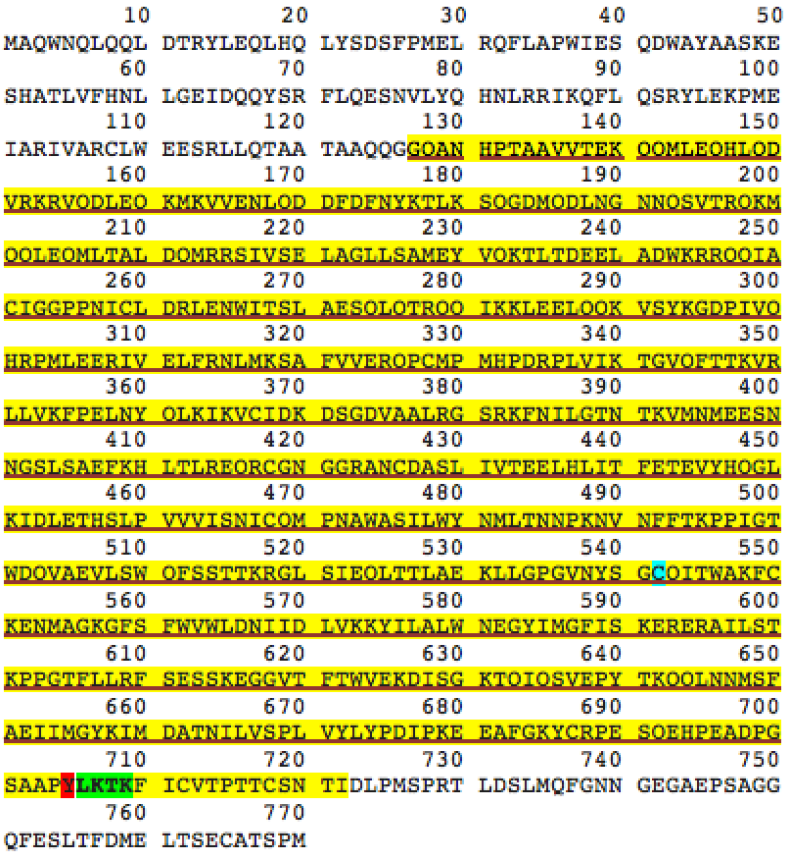


Figure 3: Amino acid sequence of the human STAT3 protein, isoform α. The sequence was obtained from UniProt (P40763 STAT3_HUMAN).

1.3.3. Activation

STAT3 exists in an inactive form in the cytoplasm. After activation by a vast array of ligands (cytokines, growth factors, hormones and other factors), phosphorylation of tyrosine residues (namely tyrosine 705) occurs (4), mediated by cytoplasmic kinases - Janus kinases (JAKs), Src family kinases - and growth factor receptor tyrosine kinases (19). The activation is a fast and temporary process in normal cells (31), lasting from a few minutes to several hours (32). However, in many cancers, STAT3 is persistently activated (11). It was first thought that phosphorylation was a prerequisite for STAT3 to be able to bind to DNA, but it has since been discovered that the unphosphorylated form (U-STAT3)

can bind to DNA as well (23). For maximal transcriptional activity, phosphorylation on a single serine (727) is required (22), but that is only possible on the STAT3α isoform, as the STAT3ß only has the tyrosine 705 phosphorylation site and STAT3γ doesn't have any of the two. In addition, when activated by cytokines, STAT3 is acetylated on lysine 685 by histone acetyltransferase p300, which is critically important for the formation of stable dimers (33).

STAT3 is activated in response to:

- cytokines: cardiotrophin-1, CNTF (ciliary neurotrophic factor), CCL5/RANTES (chemokine ligand 5), IFN-γ(interferon γ), IL-5, IL-6, IL-9, IL-10, IL-11, IL-12, IL-21, IL-22, IL-27, LIF (leukaemia inhibitory factor), LIGHT (a member of the TNF superfamily), MCP-1 (monocyte chemotactic protein-1), MIP-1α (macrophage inflammatory protein-1α), OSM (oncostatin M), stem cell factor, TNF-α (tumour necrosis factor α),
- growth factors: EGF (epidermal growth factor), G-CSF (granulocyte colony-stimulating factor), GM-CSF (granulocyte-macrophage colony-stimulating factor), IGF-1 (insulin-like growth factor 1), M-CSF (macrophage colony-stimulating factor), PDGF (platelet-derived growth factor), TGFα (transforming growth factor α),
- other factors: bile acids and low pH, black soy peptides, diazoxide, diesel exhaust particles, HCV (hepatitis C virus) core protein, HIV-1 Nef protein, isoliquiritigenin, leptin, LPS (lipopolysaccharides), nicotine, osmotic shock, heat shock, oxidative stress, EBV (Epstein-Barr virus) oncoprotein LMP1 (latent membrane protein 1), CaMKIIγ (CaM kinase II), olanzapine, UV light,
- oncogenic proteins through phosphorylation of tyrosine 705 (34).

The activation is pictured in figure 4.

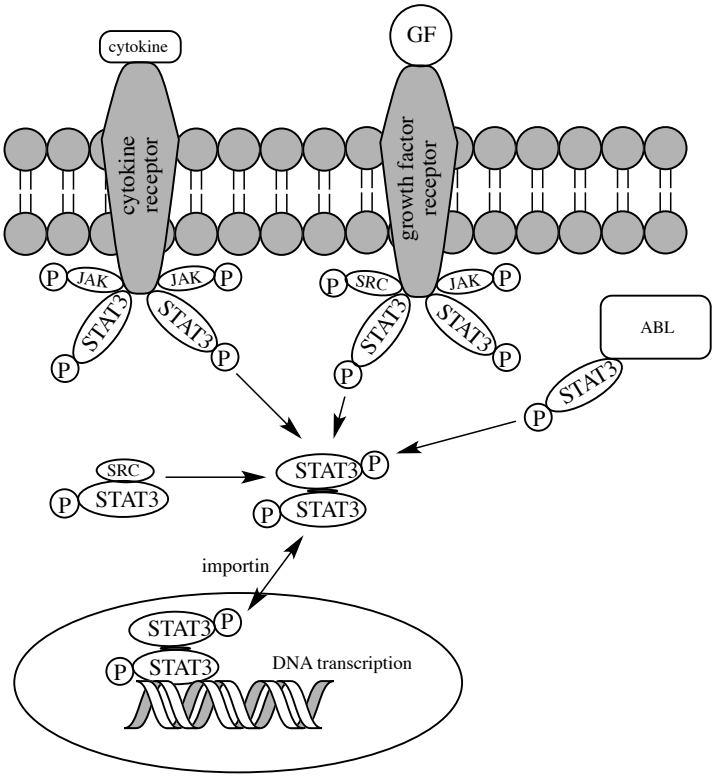


Figure 4: Activation of STAT3, adapted from source (34)

After phosphorylation, homodimers are formed via a reciprocal SH2 domain-pTyr interaction (3,19). The dimers are then translocated to the nucleus by binding to importins (35). There, they bind to the interferon γ(Gamma)-Activated Sequence (GAS motifs) of genes and activate their transcription (36). STAT3 mediates a large complexity of responses (37).

The unphosphorylated form of STAT3 (U-STAT3) can bind to the same GAS DNA binding site in the form of monomers and dimers. It also binds to AT-rich sequences and recognizes sequences, important for the organization of chromatin. This suggests that the unphosphorylated STAT3 could have a role as a genome organizer (36).

1.3.1. Roles of STAT3 in different tissues

When first discovered, it was thought that the only role STAT3 had in the body was the role in inflammation after being activated in response to one particular cytokine, IL-6 (37). However, studies have shown that IL-6 is only one part of the wide variety of activating factors, and many of the STAT3 functions have still not been explained or linked with a particular activating molecule. To study the function of STAT3 in different tissues, a

conditional gene ablation approach had to be used because of the early embryonic lethality in STAT3-deficient mice (37). The results are summarized in table III.

Table III: Roles of STAT3 in different tissues, in mice (7,37)

Tissue	Phenotype of STAT3-deficient mice
Granulocytes	Enhanced proliferation
Liver	Impaired acute phase response
Mammary epithelium	Defective apoptosis, delayed mammary involution
Monocites/neutrophils	Chronic colitis Enhanced inflammatory responses, T_H1 differentiation
Neurons	Impaired cell survival (37)
Skin	Impaired cell migration, second hair cycle, wound repair
T lymphocytes	Impaired IL-6-dependent survival and IL-2rα expression
Thymic epithelium	Hypersensitivity to stress, age-dependent thymic hypoplasia

1.3.2. Inactivation

In normal conditions, STAT3 needs to be activated and inactivated quickly. Nuclear ubiquitin E3 ligases, cytoplasmic tyrosine phosphatases, protein inhibitors of activated STATs (PIAS) and suppressors of cytokine signalling (SOCS) are the proteins that prevent further signalling (31).

1.3.3. STAT3 Target Genes

The activated STAT3 dimers activate the transcription of genes involved in cancer growth, which are listed in table IV:

Table IV: Target genes of STAT3 (34)

Process	Target gene
Angiogenesis	VEGF, HIF1ALPHA, bFGF
Inflammation	IL-1, IL-6, M-CSF, COX-2
Invasion	E-cadherin, FAK, ICAM-1, integrin ß4, integrin ß6, MUC1, stathmin

Metastasis	MMP-1, MMP-2, MMP-9, Twist
Proliferation	Cyclin B, Cyclin D1, cdc2, cdc25A, p21, c-Myc, Pim-1, Pim-2, j-jun, c-fos
Survival	Bcl-xL, Bcl-2, Mcl-1, survivin, RegIIIß, Hsp70, p53

1.3.4. Role in Cancer

Cancer is the common term for the group of diseases that include abnormal cells with two properties: uncontrollable proliferation and the ability to invade and colonize other tissues. A benign tumour is formed if the cells can only divide without stopping. These cells remain in a single mass. If the cells have both properties, it is called a malignant tumour. These cells can invade other places in the body (form metastases) by invading the bloodstream or lymph system (2).

Evidence suggests that activated STAT3 plays a critical role in malignant transformations (37). It is now known that STAT3 is activated constitutively in numerous human cancers, such as brain, breast, head and neck, leukaemia and other hematological malignancies, lung, lymphoma, melanoma, multiple myeloma, ovarian, pancreas, prostate, and renal carcinoma (11). The cause for this can be mutations that cause continued signalling or mutations that affect negative upstream regulation (7). There are large amounts of cytokines and growth factors present within the tumours as well. At least one study has shown that constitutive activation of STAT3 is associated with poor prognosis, in this case in human colorectal cancer (38).

1.4. Stat3 as a Drug Target

STAT3 has been validated as a cancer drug target by many studies (19), which has made it a popular subject of research. In addition, it fulfills all four main criteria for being a good target for cancer therapy, according to Yu and Jove (10):

- it is overly active in different tumour types,
- its activity determines patterns of gene expression that promote malignant properties in addition to survival and proliferation of cancer cells,
- it can be inhibited by small molecule inhibitors,
- target activity affects tumour cells more than normal cells.

The STAT3 signalling pathway can theoretically be targeted:

- indirectly (targeting of upstream components of the signalling pathway)
    - tyrosine phosphorylation inhibitors
- directly
    - inhibition of the DNA-binding domain
    - inhibition of the N-terminal domain
    - inhibition of the SH2 domain/dimerization.

A number of different classes of molecules that block STAT3 signalling have been studied, including antisense oligonucleotides, RNA interference, phosphopeptides and small molecular weight inhibitors. Evidence shows that blocking the signalling in tumour cells results in suppressed cell growth and apoptosis (39). In contrast, blocking the signalling in normal cells does not lead to apoptosis. This could mean that normal cells do not depend on STAT3 activity as much, while tumour cells need large amounts of activated STAT3 to survive (10).

In comparison with other compounds used to treat cancer, direct STAT3 inhibitors might be less toxic overall, because their effects on mature cells are minimal (35).

1.4.1. Direct targeting – SH2 domain inhibitors

The focus of this thesis is on discovering new small molecular weight inhibitors of the SH2 domain and, subsequently, dimerization, nuclear translocation and activation of transcription. The SH2 domain is important for STAT3:STAT3 dimerization. Blocking the domain results in disrupted dimerization, which means there are no phosphorylated dimers available to translocate to the nucleus and activate transcription.

These inhibitors need to be selective against STAT1, because the structures of STAT1 and STAT3 are similar, but their functions are opposite (35).

With the exception of BP-1-102, they have not been tested *in vivo* yet, and may require structural modifications to improve their activity, efficacy and selectivity (19). They might not all be binding to the same site.

Peptides and peptidomimetics

Coleman et al. (2005) used a peptidomimetic approach and screened a series of phosphopeptides, based on sequences of the STAT3 receptor docking sites (phosphotyrosine 705 of STAT3, glycoprotein130/gp130, epidermal growth factor

receptor/EGFR, interleukin 10 receptor/IL-10R and granulocyte-macrofage colony-stimulating factor/GM-CSF). The lead phosphopeptide they found was Ac-Tyr(PO_3H_2)-Leu-Pro-Gln-Thr-Val-NH_2 or pYLPQTV (40), which is also the phosphopeptide that we used as a fluorescently labeled molecule to measure the binding of other investigated inhibitors. Peptides and peptidomimetics are not optimal as drugs because their stability and low membrane permeation (due to size) cause low activity in vivo (19). Nonpeptidic inhibitors bypass this problem by being smaller, more cell-permeable and more stable.

Nonpeptidic inhibitors

Potential small-molecule inhibitors were identified by structure-based computational modelling of the SH2 domain with the phosphopeptide bound.

STA-21 (figure 5) was the first compound found to disrupt STAT3 dimerization in vitro (19).

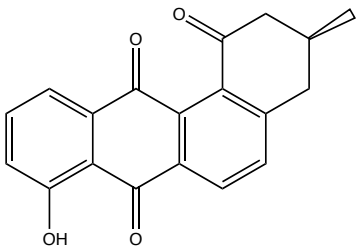


Figure 5: Structure of STA-21

S3I-201 (figure 6) was also found to disrupt dimerization in vitro. It is 3-fold more selective for STAT3 over STAT1 (19).

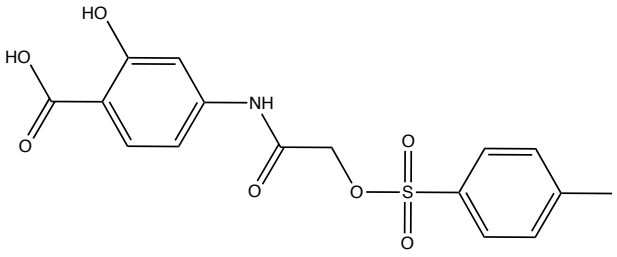


Figure 6: Structure of S3I-201

Stattic (Stat Three Inhibitory Compound, figure 7) was discovered using high throughput screening. It is an irreversible inhibitor of STAT3 activation (and, subsequently, dimerization and nuclear translocation). It was the first known selective STAT3 SH2

domain function inhibitor of nonpeptidic origin (41), inhibiting the activated as well as inactivated form.

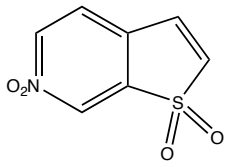


Figure 7: Structure of Stattic

BP-1-102 (figure 8) is an orally bioavailable STAT3 SH2 domain ligand that was discovered after optimizing the dimerisation-disrupting lead compound S3I-201.1066. It inhibits dimerization and, subsequently, activation, both in vitro and in vivo (42).

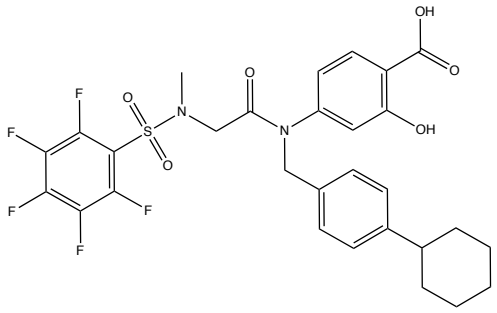


Figure 8: Structure of BP-1-102

Cryptotanshinone (figure 9) is a potent, selective inhibitor. It binds to the STAT3 monomer and inhibits STAT3 Tyr705 phosphorylation, independently of JAK2, which, subsequently, blocks the dimerization of STAT3 monomers. It is a natural compound, structurally a quinoid diterpene (43), that has been isolated from Salvia milthiorrhiza bunge (Danshen), a plant used in traditional oriental medicine.

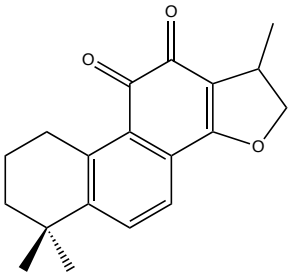


Figure 9: Structure of cryptotanshinone

5-15 DPP (5,15-diphenylporphyrin, figure 10) is a selective STAT3 inhibitor and a selective STAT3-SH2 antagonist. It prevents SH2-domain mediated ligand binding and dimerization (44).

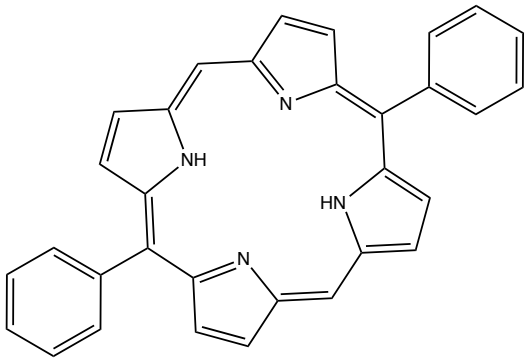


Figure 10: Structure of 5,15-diphenylporphyrin

1.5. Drug Design

Rational drug design is based on a promising biological target. There are two different approaches of virtual screening: ligand-based and structure (receptor) based drug design (45).

Ligand-based drug design is the option we choose if the structure of the target is not available. This consists of predicting the activity of compounds based on their similarity to the molecule whose activity and potency has been proven. This is what was done to discover the existing STAT3 SH2 domain inhibitors.

If the three-dimensional structure of a target (binding site) has already been obtained, either by using nuclear magnetic resonance (NMR) or X-ray crystallography, we can use that knowledge to find drug candidates by using structure-based drug design methods. Our options are *de novo* design (designing a new molecule, based on the structure of the active site), virtual screening (using docking programs to screen a chemical library *in silico*) and optimization of known ligands (45).

1.5.1. X-ray crystallography

Knowing the structure of the binding site gives us the possibility to discover new compounds, completely different than the known ligands. X-ray crystallography is the most popular method of obtaining three-dimensional structures of macromolecules. The

protein is purified and crystalized. The crystals are then exposed to an X-ray beam and the structure is calculated from the diffraction pattern (46).

The most difficult part and the rate-limiting step of this method is the crystallization. Proteins crystallize only in certain conditions and only at the right concentration. It is usually best to cover the widest possible range of conditions (buffer, pH, temperature, additives, protein concentration, choice of precipitant, choice of crystallization technique) when setting up crystallizations (46). Because this trial-and-error process can be wasteful when trying to crystallize a protein with a ligand, it is more efficient to try and optimize the conditions beforehand, using cheaper and faster methods.

1.6. Choice of Screening Method

Binding to the STAT3 SH2 domain has previously been measured by electrophoretic mobility shift assays and enzyme-linked immunosorbent assays (47). These methods are slow and labor-intensive. In the search of a faster method, fluorescence polarization was found to be useful in measuring protein-ligand interactions.

Fluorescence polarization is a non-destructive technique, specially applied to study molecular interactions. It gives a direct, nearly instantaneous measure of a labeled compound's bound/free ratio and has a low limit of detection (sub-nanomolar range). It uses inexpensive reagents and equipment; and is a high-throughput method that produces results in real time. The measurement happens in solution, and there is no need for separation of the bound and free ligand. The method produces no hazardous radioactive waste (48). Because the samples are not affected by the measurement, reagents can be added to them later and reanalyzed (49).

However, the method requires a large change in molecular volume to produce the maximum signal. In addition, the fluorescence of ligands themselves can cause artifacts, and the lifetime of the dye and size of the ligand are important in relation to the size of the protein (49). These problems can be avoided, or at least minimized, by careful planning of the chemicals involved in the assay.

2. Aim of the Study

STAT3 is constitutively activated in 50 to 90% of human cancers (50). This, along with the ability of being inhibited by small molecules, makes it a promising target for cancer treatment. The Src-homology 2 (SH2) domain is a good target for inhibition, because by inhibiting it, we block activation by phosphorylation, and subsequently dimerization, nuclear translocation and activation of transcription as well. Schust et al. have developed an assay based on fluorescence polarization, specifically for measuring the inhibition of the SH2 domain (51). This makes it possible to quickly screen small molecules and measure their percentage of inhibition of the STAT3-SH2 domain. There is no crystal structure of STAT3 with a bound inhibitor available yet. Because the structure of the binding site is unknown, we are limited to the ligand-based drug design.

- We will further develop and optimize the existing method of fluorescence polarization assay, by experimenting with different buffers (Bis Tris pH 5,7; Tris pH 7,4; 8,5; 9,5), reducing agents (dithiothreitol, tris(2-carboxyethyl)phosphine, ß-mercaptoethanol, ascorbic acid, salicylic acid), fluorescently labeled peptides (6-FAM-pYLPQ and 6-FAM-pYLPQTV) and salts.
- We will screen a small library of potential hit compounds (EL, STK and STL series) to see whether or not they are binding well enough to justify further study and/or crystallization.
- We will try to understand, without the crystal structure, how and where known inhibitors (5,15-diphenylporphyrine, BP-1-102 and cryptotanshinone) and potential hit compounds (EL, STK and STL series) are binding to the STAT3 protein, using added reducing agents (dithiothreitol, tris(2-carboxyethyl)phosphine, ß-mercaptoethanol), double stranded M67 DNA with a high affinity binding site for STAT3, and an alternative protein, in which cysteine 542 is mutated to serine, as additives during screening.

This work is part of a wider research, using a grant from the Association of International Cancer Research (AICR). The ultimate aim is to obtain a crystal structure of STAT3ß bound to an SH2 domain inhibitor.

3. Materials and Methods

3.1. Materials

3.1.1. Devices

Device	Manufacturer, Model
High-speed centrifuge	Avanti J-E Beckman Coulter
Rotor	Beckman JA-25.50 25,000 RPM
Centrifuge	Sorvall ST 40R
Swinging bucket rotor	Thermo Scientific (75003608) Swinging Bucket Rotor
Micro centrifuge	Eppendorf 5415 C
Test tube heater	Stuart Scientific Test Tube Heater SHT1
PCR thermal cycler	Biometra Personal Cycler
Microplate reader	BMG Labtech PHERAstar
Pump	Pharmacia LKB Pump P-1
Column	GE Healthcare Life Sciences HiTrap Q Sepharose FF 5 mL
pH meter	Mettler Toledo Five EasyPlus
Electrophoresis power supply	Bio-Rad PowerPac Basic Power Supply
Electrophoresis chamber	BIO-RAD Mini Protean Tetra Cell
Sonicator	MSE Soniprep 150
UV-VIS spectrophotometers	Perkin-Elmer Lambda 15 UV/VIS Pharmacia Biotech Ultrospec 2000
Orbital shaker	Stuart SSM1
Pipettes	Gilson Pipetman P2, P10, P100, P1000
Water Purifier	Elga PURELAB Option-R 7/15

3.1.2. Chemicals

Chemical	Manufacturer
Ammonium persulfate	Sigma, 98%+
Ammonium sulphate	Fisher BioReagents, 99,5% minimum, protease free
BIO-RAD Protein Assay	Bio-Rad Laboratories GmbH
Bovine serum albumin	SIGMA
DMSO	SIGMA
DTT	SIGMA
Glycerol	SIGMA
Lysogeny broth (LB)	SIGMA
Lysozyme from hen egg white	Fluka
MW marker for SDS-PAGE	Thermo Scientific Unstained Protein MW marker, #26610
PMSF BioChemica	AppliChem
Protease Inhibitor	Thermo Scientific Halt™ Protease Inhibitor Cocktail (100x)
Salicylic acid	Alfa Aesar
ß-mercaptoethanol	SIGMA
TBE buffer 10x	Invitrogen
Tris base	Fisher BioReagents, 99,8% minimum

3.1.3. Biological materials

Agilent Tech QuikChange II site directed mutagenesis kit was used to introduce the mutation of cysteine 542 to serine.

3.1.4. Other equipment

Equipment	Manufacturer, model
Concentrator tubes	Sartorius Vivaspin 20
Microtitre well plates	Greiner Bio-One Microplate 96-well, polypropylene, flat bottom, black

3.1.5. Molecules, used for HTS

5,15-DPP, BP-1-102 and cryptotanshinone were purchased from Sigma. EL compounds were synthesized at the UCL School of Pharmacy. STK and STL compounds were purchased from Vitas-M Laboratory, LTD.

3.1.6. Software

Chemdraw Professional 15.0 (CambridgeSoft) was used to draw structural formulas and other graphics.

3.2. Materials used in the assay

3.2.1. Fluorescently labeled phosphopeptides

The phosphopeptides pYLKTK, pYLPQ and pYLPQTV were labeled with 6-carboxyfluorescein (6-FAM, figure 11), chosen because of its relative photostability (52). The labels were added to the N-terminal position, as the C-terminal is important for tight binding (53). These phosphopeptides bind to the SH2 domain of STAT3 (54), which is why they are used to monitor whether the binding site is inhibited by another compound or not.

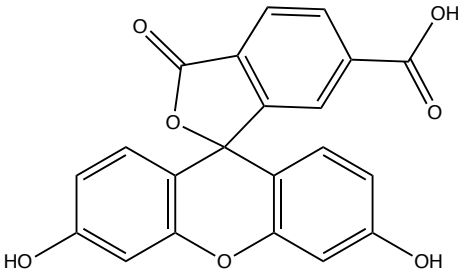


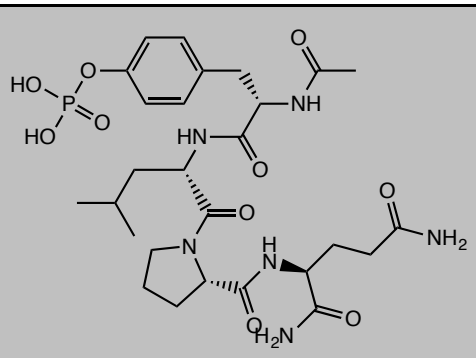
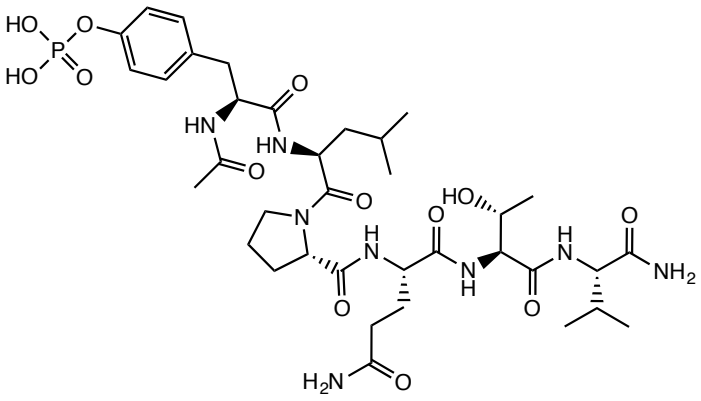
Figure 11: Structure of 6-carboxyfluorescein (6-FAM)

The fluorescently labeled phosphopeptides (6-FAM-pYLPQ and 6-FAM-pYLPQTV) had previously been dissolved in DMSO at the concentration of 10 mM, frozen and stored in 200 µL aliquots at -80°C, protected from light. On the day of use in the assay, 0,2 µL was diluted to 200 nM in 10 mL purified water for final assay concentration of 20 nM. It was vortexed and stored at 4°C protected from light sources until the time of the assay.

Although data from previous assay development studies suggests that the best conditions for storage of FAM-peptide are 10 mM Tris pH 7,4 buffer, 5% glycerol and distilled water to total 10 mL volume and 200 nM concentration, because the activity is consistent even after 48 hours (53), we decided to use distilled water only and to make the solution fresh every day, because of the minimal usage of FAM-peptide (0,2 µL daily) and better reproducibility of results.

The IC_{50} of two of the used peptides, shown in table V, were determined beforehand (53).

Table V: Structures and IC_{50} of the phosphopeptides used in the assay

Name	**Chemical structure**	**IC_{50}**
pYLPQ		1,235 µM
pYLPQTV		0,1504 µM

3.2.2. Bovine Serum Albumin (BSA)

BSA was used at a concentration of 0,1 mg/mL to minimize nonspecific binding to the plate (55). Other concentrations were tested but it was discovered that while higher concentrations could provide a better masking effect of nonspecific binding, BSA can then form colloidal aggregates and block the light path and also bind to the test compounds (53).

3.2.3. Ethylenediaminetetraacetic acid (EDTA)

EDTA was used as a metal ion scavenger, because there were multiple sources of metal ions throughout the process of protein purification (ion-exchange column, ammonium sulphate precipitation). EDTA was not added to the buffer when testing the effect of metal ions on the assay.

3.2.4. Buffer

Depending on the desired pH, we used the following buffers:

- Bis Tris pH 5,7
- Tris pH 7,4; 8,5 and 9,5.

3.2.5. Glycerol

Glycerol was added to the buffer mixture as a cosolvent (5%) to enhance the stability of the protein in the solution (56).

3.2.6. Proteins

The protein we used in assays (except where noted) was STAT3ß 688stop (688s). This is a C-terminally truncated form of STAT3ß with amino acids 688 to 722 removed. The initial experiments were done with STAT3ßtc (residues 127-722) and STAT3 688stop (residues 127-688); the results were comparable. However, since 688s has the interlinking arms removed, which is better for the kinetic simplicity of the assay, it was decided that it was the better choice (53). We also used the STAT3ß (127-722) C542S mutated protein, in which the cysteine 542 was mutated to serine. We produced both proteins, STAT3ß 688stop and STAT3ß C542S, as described on page 25.

The concentration of the protein was 2,5 μM, and 250 nM in the assay solution.

3.2.7. Reducing agents

Reducing agents are usually added to crystallization trays to prevent aggregation of the protein. It is therefore crucial to screen potential inhibitors with the addition of a reducing agent, to make sure they will bind to the protein in those conditions as well. It is also preferable to use a reducing agent in the assay solution for the same reason – aggregation.

Dithiothreitol (DTT)

The compounds were tested without and with DTT (0,5 mM), because it has been reported that Stattic only inhibits STAT3 when no DTT is present. Since DTT is added to the protein in all stages of the purification, it has to be removed before the assay by a few cycles of dilution and concentration. The reduction of a disulfide bond is pictured in figure 12.

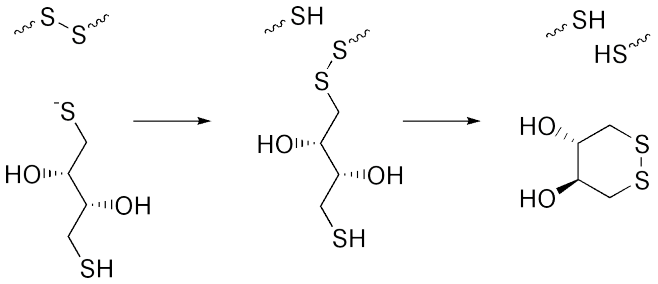


Figure 12: Reduction of a disulfide bond with DTT

Tris(2-carboxyethyl)phosphine (TCEP)

TCEP is a strong reducing agent and was tested as an alternative to DTT. The main advantages over DTT are better stability, more powerful and irreversible reduction of disulfides (57). It is, however, more expensive. The concentration in the assay solution was 0,5 mM. The reduction of a disulfide bond is pictured in figure 13.

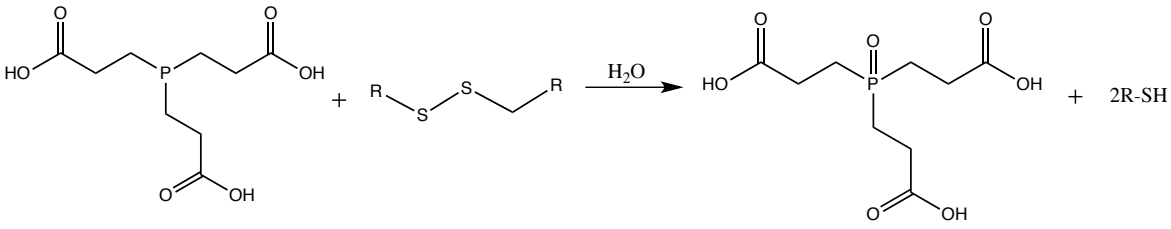


Figure 13: Reduction of a disulfide bond with TCEP

ß-mercaptoethanol

This reducing agent is usually used in SDS-PAGE sample buffers. It is effective in a smaller pH range than TCEP. The concentration in the assay solution was 1 mM. The reduction of a disulfide bond is pictured in figure 14.

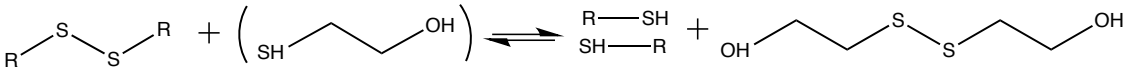


Figure 14: Reduction of a disulfide bond with ß-mercaptoethanol

3.2.8. DMSO

The DMSO concentration in the assay solution was 4,01% (4% from the tested compound and 0,01% from the 6-FAM-labeled peptide).

3.2.9. DNA

The DNA we used was a 17-base pair double-stranded M67 DNA. M67 contains a high affinity binding site for STAT3 (23).

M67 dsDNA was prepared previously, by dissolving two oligonucleotides 50-TGCATTTCCCGTAAATCT- 30 and 50-AAGATTTACGGGAAATGC-30 in 100mM NaCl at 100 nM each, then annealing at 95°C for 5 min and slowly cooling to room temperature (23).

When M67 DNA was used, it was added to the buffer mixture (5µL/mL protein) and incubated for 10 minutes before use. The nominal concentration was 0,66 mM, but the purity was only 70%, so the actual DNA concentration is closer to 0,5 mM.

3.2.10. Compounds

There have been numerous studies into finding a good nonpeptidic, selective STAT3 inhibitor. We used three known inhibitors in the assay to try and optimize the conditions for crystallization: 5,15 diphenylporphyrin (5,15-DPP), BP-1-102 and cryptotanshinone.

We screened EL, STK and STL groups of compounds. They were initially dissolved in DMSO and then diluted with DMSO as needed to 2,5 mM, for final assay concentration of 100 µM (except where noted, because of some compounds' low solubility).

3.3. Methods

The workflow is presented in figure 15.

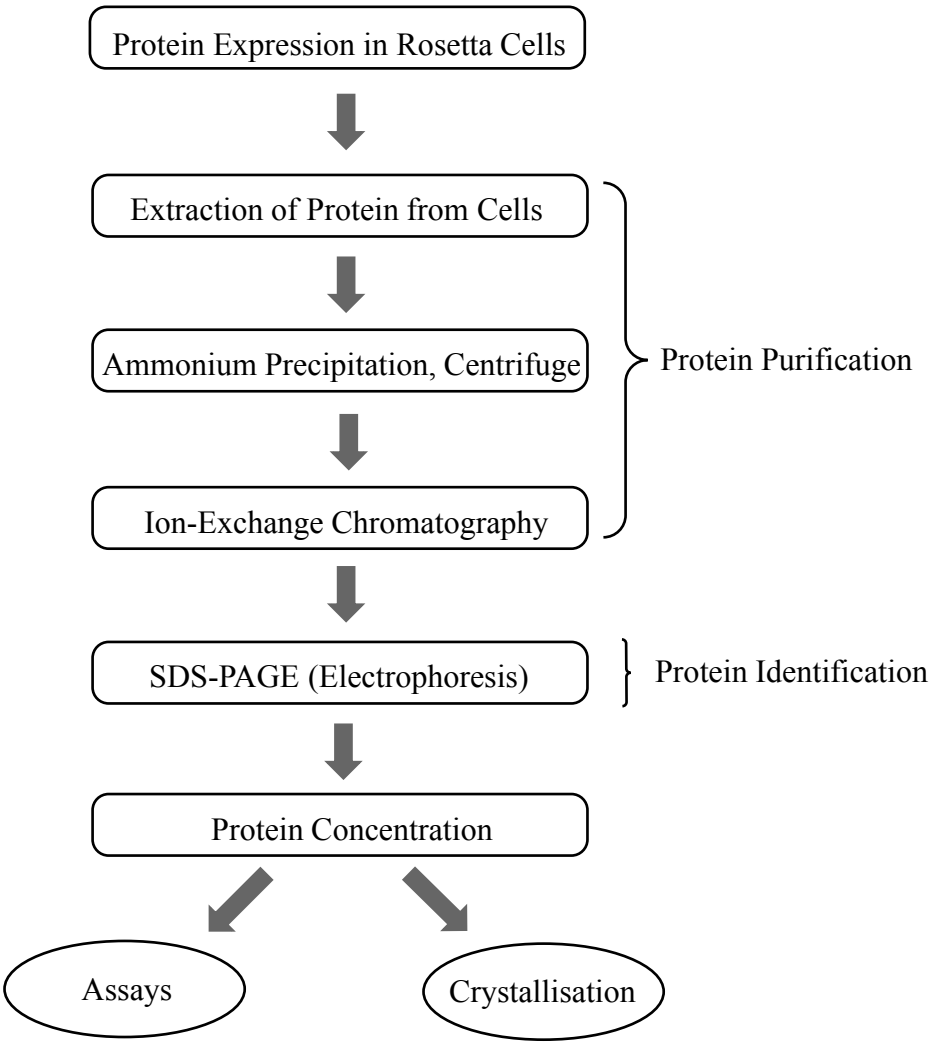


Figure 15: Workflow of experiments

3.3.1. Production of STAT3 C542S

To introduce the mutation of cysteine 542 to serine, QuikChange II Site Directed Mutagenesis Kit was used.

The primers used were: forward 5'CCTGGTGTGAACTACTCAGGGAGCCAGATC ACATGGGCTAAATTC3', reverse 5'GAATTTAGCCCATGTGATCTGGCTCCCT GAGTAGTTCACACCAGG3'.

3.3.2. Protein expression

The method of protein production was adapted from Becker et al. (58). The expressed proteins were STAT3ß 688s and C542S and the protocol was the same for both of them. STAT3ß 688s has the native human codon sequence, which is not suitable for expression in E. coli. We used Rosetta strain competent cells, which encode tRNA, important for expression of human sequences (53).

The cells were grown to OD600 of 2-3 in 4 500 mL flasks of LB media (16g LB/500 mL) at 30°C. The temperature was chosen to prevent excessive protein degradation (53). Then, another 500 mL of LB media were added (16g LB/500 mL), the temperature reduced to 21°C and left for 20 minutes. Isopropyl-beta-D-thiogalactopyranoside was added, the concentration made to 1 mM and incubated for 6 hours.

The suspension was centrifuged at 4°C for 15 minutes at 4000 g. The supernatant was discarded. The samples were kept on ice to prevent host protease breakdown of the protein.

3.3.3. Extraction of protein from cell pellet

Each of the four pellets was resuspended in 50 mL of the extraction buffer consisting of: (makes 200 mL, enough for 4 pellets)

- 200 mM KCl
- 20 mM $MgCl_2$
- 50 mM Tris (pH 8.5)
- 10 mM DTT
- 30% glycerol
- 1/2 spatula benzamidine
- 1/2 spatula lysozyme

- 1/2 small spatula PMSF (phenylmethanesulfonylfluoride)
- 0.3 mL halt protease inhibitor cocktail

The suspension was transferred to an ice bath and sonicated five times for one minute, with 30-second breaks to allow for cooling. The cell homogenate was then centrifuged for 1 hour at 48254 g (20000 rpm); the pellet was discarded. Saturated ammonium sulphate solution was added to the supernatant to final concentration of 40% and centrifuged at 48254 g (20000 rpm) for 40 minutes. The supernatant was discarded. The pellet was resuspended in 60 mL 20 mM Tris, 1 mM DTT, 50 mM NaCl; 60 µL Halt protease inhibitor cocktail (100x) was added. This was transferred into two 50 mL falcon tubes and centrifuged at 4000 g for 20 minutes. 30 mL saturated ammonium sulphate was added to the supernatant and centrifuged at 48254 g (20000 rpm) for 1 hour. The supernatant was discarded. The pellet was dissolved in 72 mL of 100 mM Tris pH 8,5; 10 mM magnesium chloride, 10 mM DTT and 70 µL Halt protease mixture. This was transferred into two 50 mL falcon tubes and centrifuged at 4000 g for 20 minutes. 8 mL saturated ammonium sulphate solution was added to the supernatant and centrifuged at 12064 g (10000 rpm) for 20 minutes. The pellet was discarded and 16 mL saturated ammonium sulphate solution added to the supernatant. This was centrifuged at 48254 g (20000 rpm) for 1 hour. The supernatant was discarded. The pellet was redissolved in 50 mL 25 mM Tris pH 8,5; 5 mM magnesium chloride, 5 mM DTT and 50 µL Halt protease mixture.

3.3.4. Ion-exchange Protocol

Four 5 mL fast flow q-sepharose columns (GE Healthcare HiTrap Q Sepharose FF) were used to clean the protein. Between uses, the columns were washed with 200 mL distilled water, 200 mL 20% ethanol, 300 mL 1M NaCl, 50 mM Tris; 0,1% N-lauroyl sarcosine, followed by another 200 mL distilled water, 200 mL 0,5M NaOH, 200 mL distilled water and 500 mL of the running buffer.

The extract was loaded onto the column at the flow rate of 1 ml/min. The columns were then washed with 50 mL of running buffer (20 mM Tris pH 8.5/9.5), and elutes of 50 mL + 8 mL were collected after every buffer fraction (E100, E200, E300, E400, E1000 – consisting of running buffer plus 100/200/300/400/1000 mM NaCl). Ice-cold saturated ammonium sulphate solution was added to fractions E100 and E200 (they contained most of the protein) to 40% final concentration.

3.3.5. SDS-PAGE

Samples for SDS-PAGE were taken throughout the extraction and purification process (before the addition of saturated ammonium sulphate solution). The gel was prepared according to the recipe in table VI.

Table VI: 10% polyacrylamide gel recipe for SDS-PAGE

Resolving gel	**Volume (µL)**	**Stacking gel**	**Volume (µL)**
ddH_2O	1875	ddH_2O	1775
1M Tris pH 8,5	3000	1M Tris pH 6,8	338
40% acrylamide /bisacrylamide	2500	40% acrylamide /bisacrylamide	433,5
10% SDS	60	10% SDS	26
TEMED	6	TEMED	2,6
10% APS	60	10% APS	26

*APS: ammonium persulfate

*TEMED: tetramethylethylenediamine

The glass plates were thoroughly washed with purified water and wiped with 70% ethanol. The resolving gel was prepared and pipetted between two glass plates in the gel caster first, left to solidify for 20 minutes and then the stacking gel was pipetted on top and left to solidify for at least 30 minutes at room temperature before using it.

The TBE 0,5x buffer was prepared by diluting the TBE (Tris/borate/EDTA) buffer 10x.

Well 1 was used for 10 uL of marker. Protein samples (20 µL) were placed into test tubes and mixed with 10 µL of sample buffer, which denatured the proteins and reduced disulfide bonds. They were heated to 95°C for 10 minutes in a test tube heater and then centrifuged for one minute. 16 µL of each sample were pipetted into the wells. The gel was run at 70 V until the dye reached the resolving gel. Then, the voltage was raised to 150 V.

Gels were stained with Coomassie blue and analyzed to identify the protein and assess the level of purity after ion exchange chromatography.

3.3.6. Protein precipitation

The precipitated protein from the column was centrifuged and the supernatant was discarded. The pellet was dissolved in 20 mL 25 mM Tris pH 8.5/9.5, 20 mM NaCl. This was transferred to a 20 mL centrifugal concentrator tube and centrifuged until the volume reduced to 0.5 mL. The solution was rediluted to 20 mL and centrifuged at 1800 g until the volume reduced to 0.5 ml again. Then, the remaining solution was topped up with 20 mM Tris to 5 mL.

The protein was left on ice for at least five minutes. Then, its concentration was measured, using a spectrophotometer (wavelength: 595 nm) and BIO-RAD Protein Assay (1 part reagent, 4 parts water), using 20 µL of the protein solution 1 mL of the reagent. The protein solution was then diluted accordingly to 0.165 mg/mL (2,5 µM).

3.3.7. Storage of protein

If not used on the same day, the (undiluted) protein was frozen. A volume corresponding to approximately 1 mg of the STAT3ß 688stop protein was pipetted into 2 mL tubes and precipitated using a saturated ammonium sulphate solution. After allowing one hour at 4°C for precipitation, the pellets were snap frozen in liquid nitrogen and stored at -80°C.

The protein activity was tested before and after freezing to ensure there was no significant change in activity.

3.3.8. Preparation of Protein for FP Assay (after freezing)

A small quantity of the protein was thawed and processed every day, as needed for the assay. The dilutions and centrifugations were necessary to bring the DTT concentration down to a minimum.

Frozen 1 mg aliquots were thawed and spun at 23644 g (14000 rpm) in a microcentrifuge. The supernatant was discarded and the pellet resuspended in 1 mL of 200 mM NaCl/20mM Tris. This was centrifuged for 10 minutes at 23644 g (14000 rpm). The supernatant was transferred to a Vivaspin 20 concentrator tube and topped up to 20 mL with 200 mM NaCl/20 mM Tris. This was centrifuged at 1800 g until volume reduced to 1 mL, rediluted to 20 mL and centrifuged again, rediluted and this time, centrifuged until the volume reduced to 0,3 mL. Protein concentration was measured with a UV/VIS spectrophotometer and diluted to 0,165 mg/mL with 20 mM Tris.

3.3.9. FP Assay Protocol

Microtitre plate preparation

96-well plates were washed thoroughly: filled with distilled (15 MΩ) water twice and distilled (18 MΩ) water once, each time placed on a reciprocating lab shaker for five minutes. Afterwards they were dried with pressurized nitrogen. Following this procedure reduced background variation significantly.

4 μL of (typically) 2,5 mM compound dissolved in DMSO was pipetted into the wells first. Then, 86 μL of the buffer-protein/water mixture was added. Final well composition was as follows: 5% glycerol, 20 mM Tris, 0,1 mM EDTA, 0,1 mg/mL BSA (+0,5 mM DTT, when used). The plates were incubated on a lab shaker for one hour. 10 μL of FAM-labeled peptide was then added to each well and the mixture allowed to equilibrate for 30 minutes on the lab shaker. The steps of the assay protocol are listed in table VII.

Table VII: Assay protocol

Step	Value	Description
1	4 μL	Compound in DMSO
2	86 μL	Buffer solution+water/protein
3	60 minutes	Incubation
4	10 μL	Addition of FAM-labeled peptide
5	30 minutes	incubation

There were three wells without protein and three wells with protein for every condition being tested. This ensured we always had triplicate results on every plate. The wells without protein were used to test for intrinsic fluorescence of the compounds. The first row was always blank; we used 4 μL of DMSO in place of compounds. This was done to take into account the small inconsistencies in plate geometry, which can affect the results. The layout of the 96-well plate is pictured in figure 16.

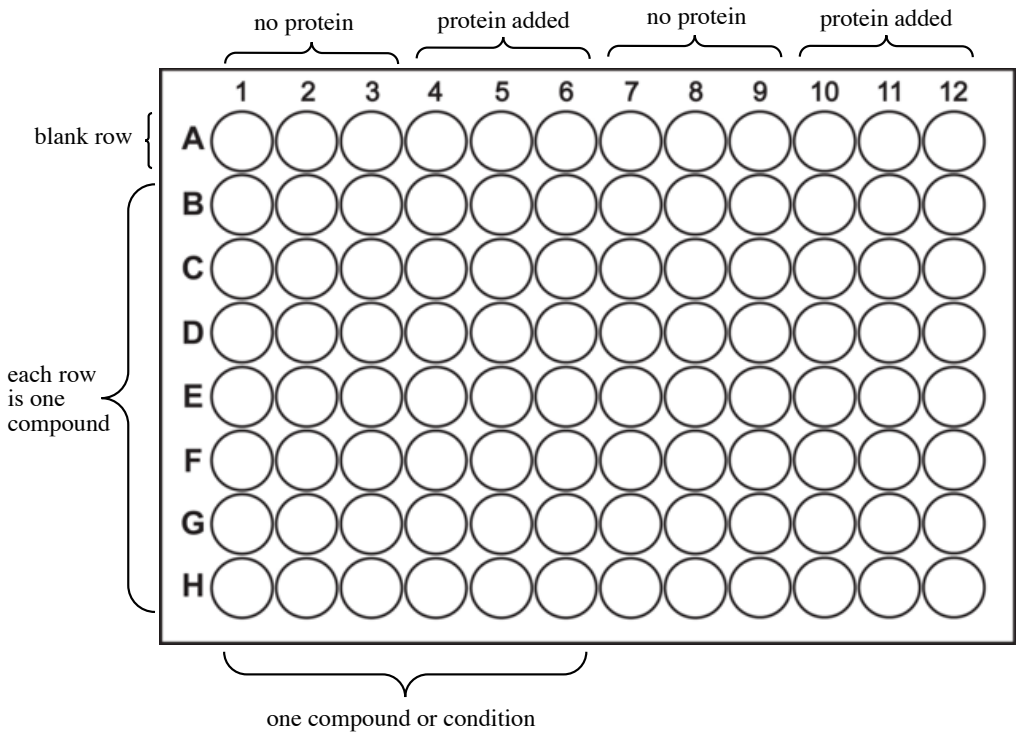


Figure 16: 96-well plate layout

Fluorescence Polarization Assay

At the beginning, the light passes through a polarizing filter. When the fluorescently labeled molecule is excited by this plane-polarized light of the right (absorption) wavelength, it absorbs the light and becomes excited, rotates and tumbles in the solution and, in the end, emits light at a different, emission wavelength. If the molecule didn't move during the excitation, the light would be emitted back into a fixed plane. However, because all the molecules/complexes rotate and tumble, light is emitted in different planes. The measured polarization is inversely correlated to the speed of rotation in the solution, which depends on the molecular mass of the complex (49,55). The principle is depicted in figure 17.

The assay protocol we used was developed by Schust and Berg (47) to measure the ability of small molecules to bind to the SH2 domain of STAT3 (47).

Phosphopeptides pYLKTK, pYLPQ and pYLPQTV, which are known to bind to the SH2 domain of STAT3 (54), were labeled with a fluorescent probe: 6-carboxyfluorescein (6-FAM). 100% DMSO was used as negative control. Measurement of the percentage bound fluorescently labeled phosphopeptide to the protein was carried out by the PHERAstar Plus BMG Labtech plate reader. The excitation wavelength was set to 480 nm and the emission

wavelength to 530 nm; well depth was 5 mm. The measurements were carried out at room temperature and repeated after 16 hours of storage at 4°C.

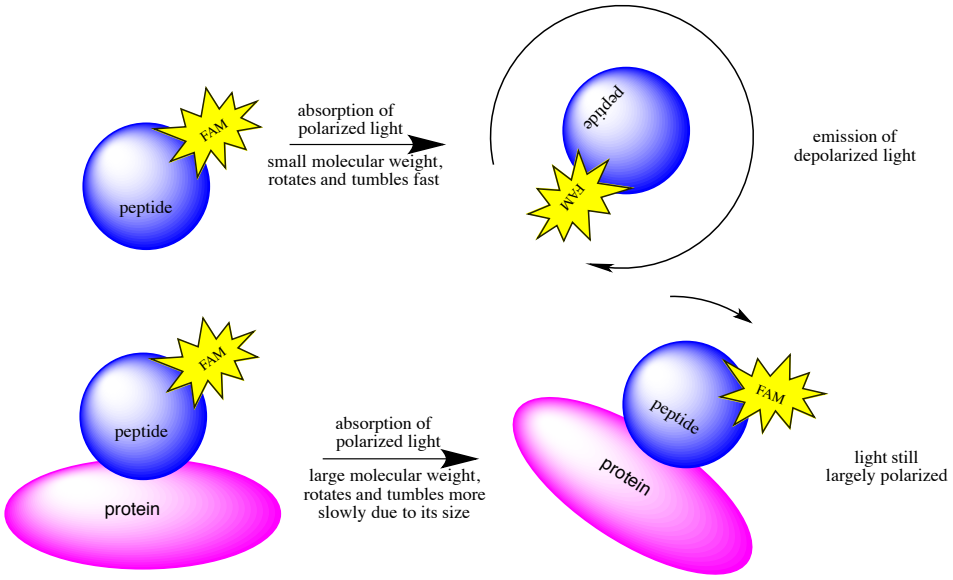


Figure 17: Principle of fluorescence polarization, adapted from Moerke (55)

Analysis of results

The software calculated the fluorescence polarization of each well from raw data using equation 1:

$$P = 1000 * \frac{Iparallel - Iperpendicular}{Iparallel + Iperpendicular}; [mP]$$

Equation 1: calculation of fluorescence polarization from intensity of emitted light

P: fluorescence polarization
I: intensity of the emitted light
mP: milipolarization unit

P is a dimensionless number that can be expressed in polarization or milipolarization units (1 polarization unit=1000 milipolarization units).

The percentage of inhibition was calculated in Excel using equation 2:

$$\%inh. = 100 - \left(\frac{P(sample + protein - P(sample - protein)}{avg(3x\ P(blank + protein)) - avg(3xP(blank - protein))}\right) * 100$$

Equation 2: calculation of the percentage of inhibition of a protein

Then, standard deviation was calculated from the three or more values.

4. Results and discussion

All the assays were performed in triplicate. The preliminary screening was performed on one plate, generating three measurements for each condition. The best results were obtained when using the non-physiological pH 8,5, so that was studied further, as the very same pH could be used in crystallography studies. Assays at pH 8,5 with and without DTT, DNA and the mutated protein (STAT3ß C542S) were performed on three plates simultaneously, generating nine measurements for each compound/condition. The results are average values, along with the standard deviations, with not more than 30% of outliers removed. The measurements where human error occurred (failure to add the fluorescently labeled peptide) are omitted. The values represent the percentage of inhibited STAT3ß (SH2 domain) in the assay solution, except where noted. Read 1 and Read 2 represent the measurements at t=0 h and t=16 h.

The inhibition is high if the signal was low. When an inhibitor was already bound to the SH2 domain, the labeled peptide remained in the solution. Because of its small size, it rotated and tumbled fast, so the light was largely depolarized and the signal was lower. When the peptide did bind to the SH2 domain, the newly formed complex was a lot larger, so the speed of tumbling and rotation was slower and the light stayed more polarized, causing the higher signal.

The error bars represent the standard deviation (SD) of the results (calculated by using equation 2). Standard deviation is high for some results. However, most of the time these results do not specifically point to human error, especially when the SD is high for the same compound in different conditions, or when more repeats of the same assay show similar SD. There is a number of possible reasons:

- low solubility of the compound (as later discovered for 5,15-diphenylporphyrin),
- aggregation on a visible or invisible scale,
- binding of the compound to the polypropylene plate,
- incompatibility of the compound with the additives,
- unknown behavior that changes the signal independently of the labeled peptide (addition of DNA).

The human error only happened when adding the fluorescently labeled peptide, because this addition does not result in visible changes to the sample. It could be avoided by using

a multichannel pipette instead of standard pipettes, or by using a black light to make it visible if the fluorescently labeled peptide has already been added to a well. The low solubility and subsequent aggregation could be avoided by using a larger percentage of DMSO in the assay. However, we would first have to confirm it does not affect the inhibition in any way, since previous findings by this group indicate that using more than 4% DMSO results in a decrease of the ligands' inhibitory potential. Binding of the compound to the plate can be avoided by measuring FP as soon as possible after the pipetting and incubation.

The timeline of our screening is:

- Preliminary screening at pH 7,4
- Screening at pH 5,7/6,8/8,5/9,5 (effect of pH)
- Screening with and without DTT (effect of reduced groups)
- Screening with 688s and C542S (effect of cysteine 542 – possible binding site)
- Screening with DNA (interaction between protein and DNA in the presence of inhibitors; how DNA affects inhibition)

4.1. Method Development

The method we used in our experiments has been developed by different research groups (47,52) and further modified here to increase reliability.

Presence of glycerol and different pH levels

We tested two different 6-FAM-peptides that had been used in this research group's experiments before, using four different pH levels and the addition of 5% glycerol. No ligands were used in this experiment. The goal was to see which 6-FAM-peptide and which pH level are best for reaching the largest difference in polarization between the samples with and without protein. We concluded it would be best to continue the work using the 6-FAM-pYLPQ peptide. The results (pictured in figure 18) show it is preferable to work at pH 7,4 or higher. The signal difference at higher pH was also higher after the addition of 5% glycerol, which had been expected, because glycerol stabilizes proteins (56).

Even though DTT wasn't used in this particular assay, it is worth noting that it would not function as a reducing agent at pH 5,7, because its reducing power is limited to pH>7.

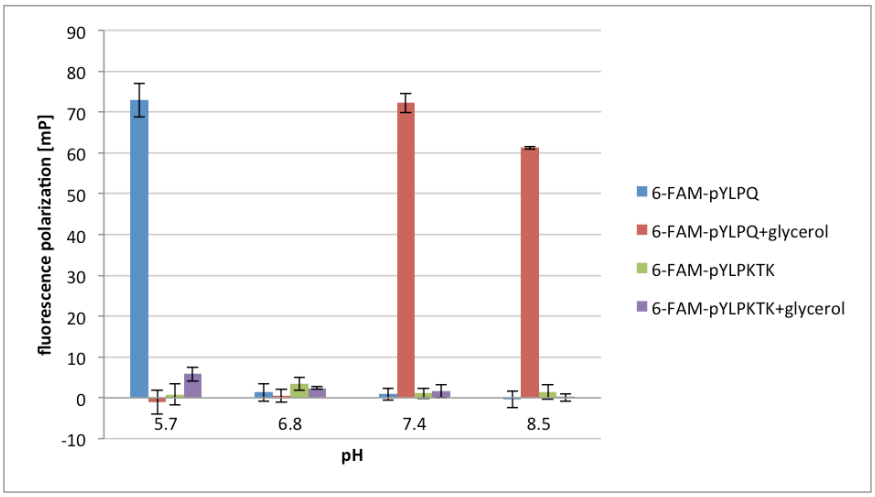


Figure 18: Difference in FP signal between protein and no protein samples without and with 5% glycerol. STAT3ß 688s, Bis Tris pH 5,7; Tris pH 6,8/7,4/8,5; 6-FAM-pYLPQ/pYLPQTV. Error bars represent the SD of measurements, n=3.

Choice of fluorescently labeled peptide

After eliminating 6-FAM-pYLKTK from the assay, we were left with 6-FAM-pYLPQ and 6-FAM-pYLPQTV. The observed inhibition values (pictured in figure 19), when using 6-FAM-pYLPQ or 6-FAM-pYLPQTV, are similar for many of the tested compounds. The measured IC_{50} values, however, show a large difference between phosphopeptides pYLPQ and pYLPQTV (1,235 µM and 0,1504 µM, respectively). This means that the 6-FAM part of the molecule is somehow influencing the affinity of STAT3 for both peptides, especially 6-FAM-pYLPQ. At this point, it is difficult to say which one of the labeled peptides would be preferable for use in the assay. Most of our assays were therefore done with both.

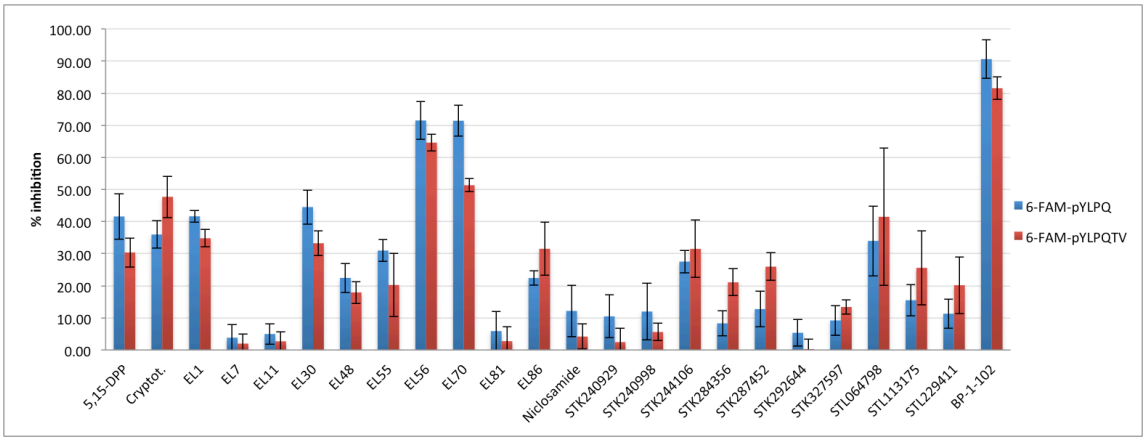


Figure 19: Comparison of % inhibition using 6-FAM-pYLPQ and 6-FAM-pYLPQTV; STAT3ß 688s, Tris pH 8,5. Error bars represent the SD of measurements, n=9.

Addition of ascorbic and salicylic acid

Ascorbic and salicylic acid (antioxidants/reducing agents, structures in figure 21) were tested to see if they would make a good additive for stabilization of the assay solution and production of a stronger signal. The inhibition was compared to inhibition with 10 µM and 0,5 mM DTT. The tested compounds from the STK and STL groups showed significantly higher inhibition when using ascorbic or salicylic acid as an additive instead of the usual concentration of DTT (pictured in figure 20). Further experimentation is needed because of the high SD, but for now we can conclude that, at least ascorbic acid could have a beneficial effect on the assays of STK and STL compounds.

Salicylic acid has a structure similar to a known hit compound, niclosamide, which is why it could interfere with the labeled phosphopeptide binding by blocking the active site and producing false higher inhibition.

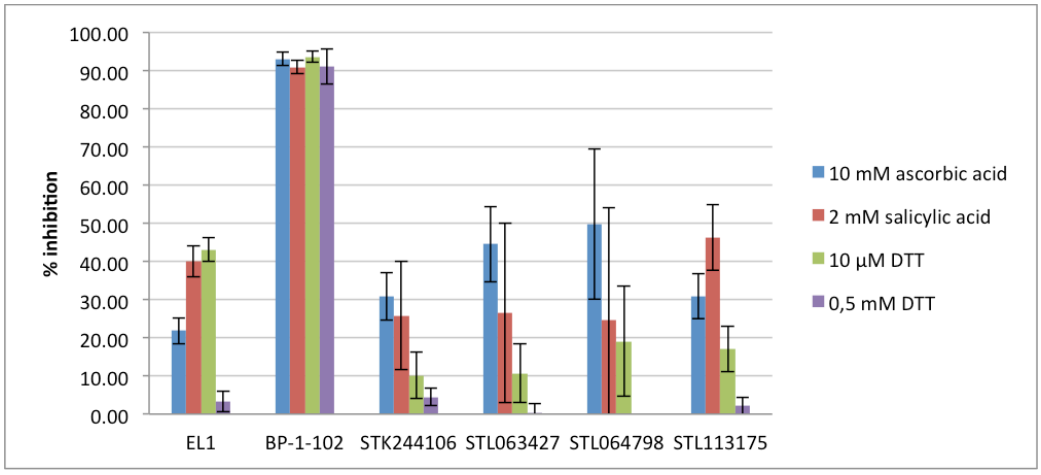


Figure 20: Inhibition after the addition of ascorbic acid, salicylic acid and DTT. STAT3ß 688s, Tris pH 8,5; 6-FAM-pYLPQTV. Error bars represent the SD of measurements, n=9.

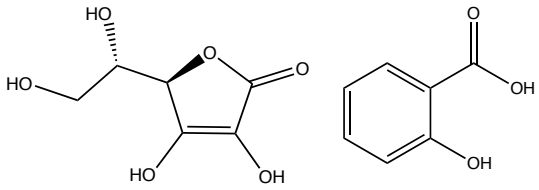


Figure 21: Structures of ascorbic acid and salicylic acid

Reducing agents – alternatives to DTT

Here, we used EL1 as an inhibitor. It is an analog of Stattic, a good inhibitor whose inhibition gets significantly lower after the addition of a reducing agent (usually DTT). ß-mercaptoethanol (BME) and tris(2-carboxyethyl)phosphine (TCEP) hydrochloride were tested as possible alternatives to DTT. Most compounds show a lower percentage of inhibition when DTT is added, however its addition in the assay is preferable, as it keeps –SH groups from forming disulfide bonds. This can lead to aggregation or inactivity of the protein. Results, pictured in figure 22, show that the inhibition is higher when using BME and TCEP, with TCEP being the preferable reducing agent, but also the most expensive. It is also worth noting that, while the samples with DTT show higher inhibition the next day, the ones with BME and TCEP show lower inhibition.

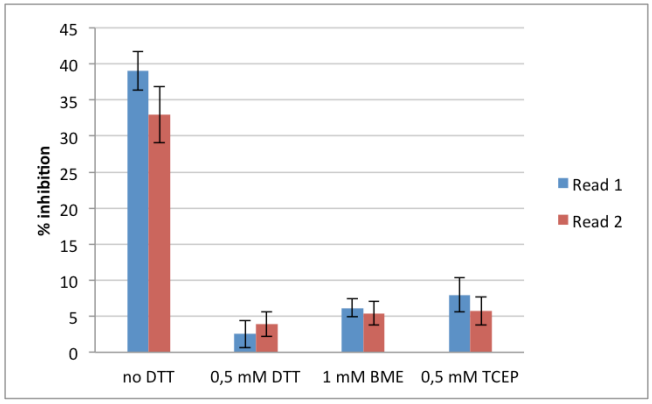


Figure 22: Comparison of inhibition of EL1 after the addition of DTT, BME and TCEP. EL1, STAT3ß 688s, Tris pH 8,5; 6-FAM-pYLPQTV. Error bars represent the SD of measurements, n=9. Read 1: t=0h, read 2: t=16h.

4.2. Screening

4.2.1. 5, 15-diphenylporphyrin

5, 15-diphenylporphyrin (figure 10) is a selective STAT3 inhibitor and a selective STAT3-SH2 antagonist. It is less soluble in DMSO than other screened compounds.

Effect of ions

Ions can affect the inhibition of 5,15-DPP, because the centre of a porphyrin molecule can capture a metal ion, resulting in a stable organometallic complex. In order to exclude a possible negative effect of the ions on the inhibition, we tested different ions as additions

to the assay. First, we tested the influence of ions on the protein alone, using DMSO in place of a compound. The results are pictured in figure 23. While the low "inhibition" shown by the ions themselves should be good news, the high standard deviation makes the results unreliable, apart from the addition of Cu^{2+}.

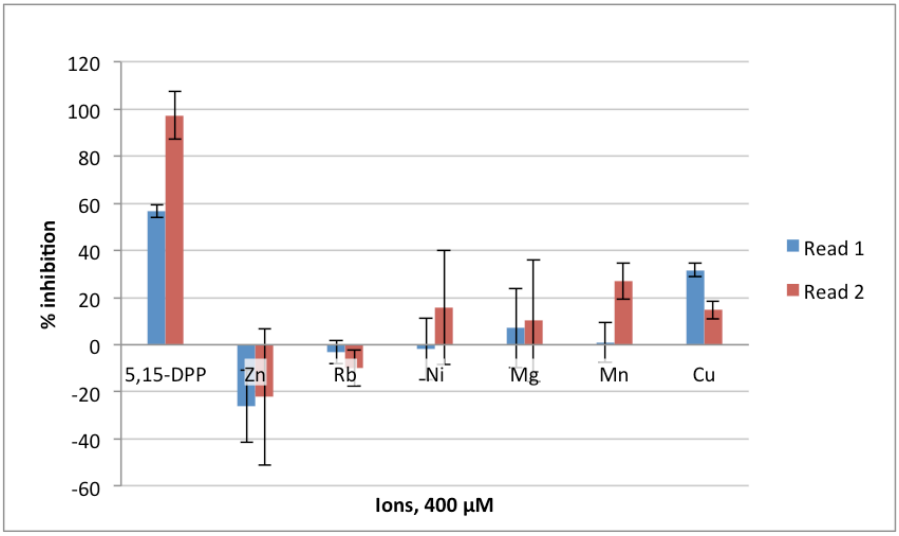


Figure 23: Effect of different metal ions on the inhibition. DMSO, STAT3ß 688s, Tris pH 7,4; 6-FAM-pYLPQ, ions (400 µM). Error bars represent the SD of measurements, n=3. Read 1: t=0h, read 2: t=16h.

On the second plate, we also tested iron in addition to the previous ions, while adding 5,15-DPP to the mix. The results are pictured in figure 24. The inhibition of 5,15-DPP without any ions is around 10% higher after 16 hours, which could be explained as slow bond formation. Zn^{2+}, Ni^{2+} and Mg^{2+} produced three extremely different results, both on the plate without and with 5,15-DPP present. The only two ions that appear to have significantly enhanced the binding were Fe^{2+} and Mn^{2+}. The sizes of these ions do not seem to have a great effect on the results; Rb^{2+} ions, for example, are twice as large as the others, but there is no notable difference between the inhibition percentages.

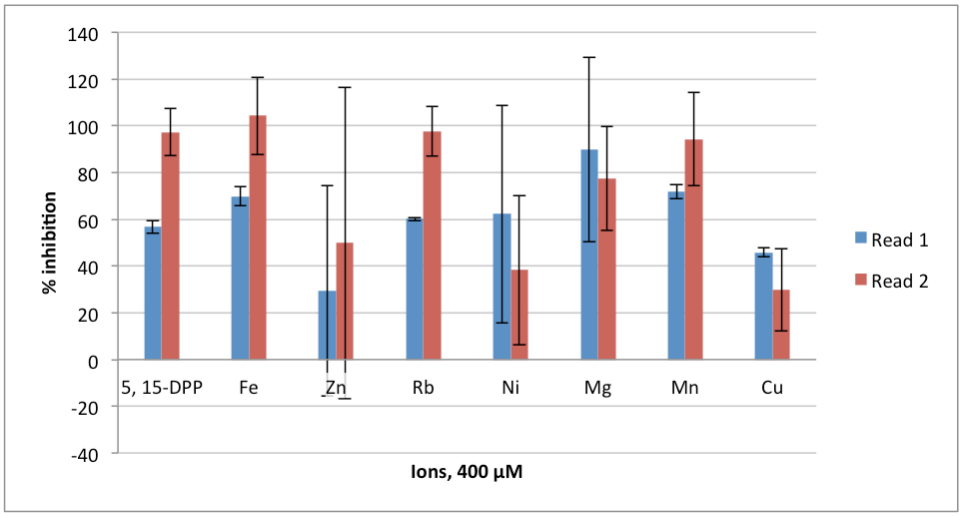


Figure 24: Effect of different metal ions on the inhibition. 5,15-DPP, STAT3ß 688s, Tris pH 7,4; 6-FAM-pYLPQ, ions (400 µM). Error bars represent the SD of measurements, n=3. Read 1: t=0h, read 2: t=16h.

We tested a few other ions, this time using 6-FAM-pYLPQTV and a 4x smaller concentration (figure 25). Again, most of the results are variable, resulting in a high standard deviation and therefore the ions are not useful as a potential additive to the assay. Only Ni^{2+} and Cu^{2+} show a relatively high "inhibition" percentage, which tells us they somehow affect the binding of the labelled peptide even when no inhibitors are present.

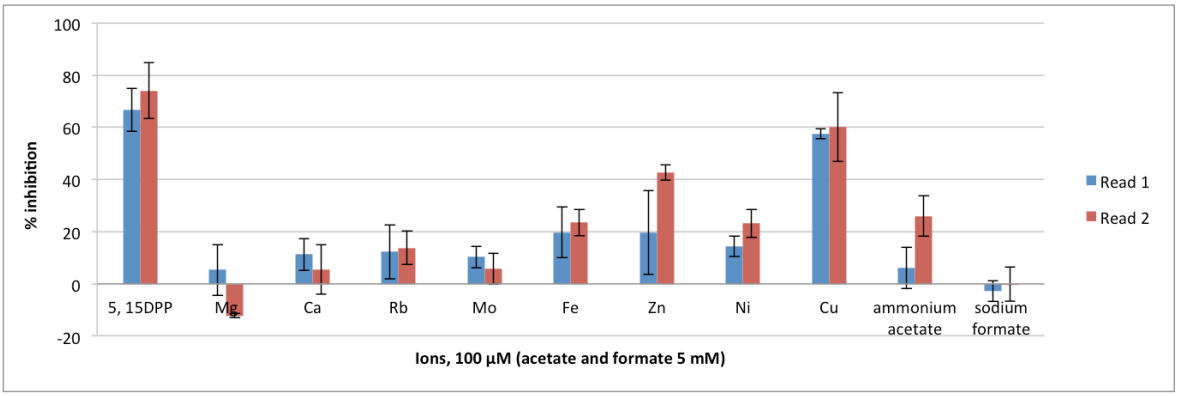


Figure 25: Effect of different ions on the inhibition. DMSO, STAT3ß 688s, Tris pH 7,4, 6-FAM-pYLPQTV, ions (100 µM). Error bars represent the SD of measurements, n=3. Read 1: t=0h, read 2: t=16h.

In the second group of ions there are some promising ones that seem to enhance inhibition (shown in figure 26). Since the inhibition of 5,15-DPP alone is 59,35%, the useful ions to add to the assay would be Mo^{2+}, Fe^{2+}, Zn^{2+}, Ni^{2+}, Cu^{2+}, ammonium acetate and sodium

formate. Looking at the previous graph we see that some of these (Fe^{2+}, Zn^{2+}, Ni^{2+}, Cu^{2+} and ammonium acetate) show high "inhibition" even when no inhibitor is present in the solution. We do not know if these results mean that the binding of the inhibitor to the protein is enhanced, or that the additional part is actually the ions blocking the binding site from the labeled peptide, so we remove these ions from further experiments. We are left with Mo^{2+} and sodium formate as the only two promising additives.

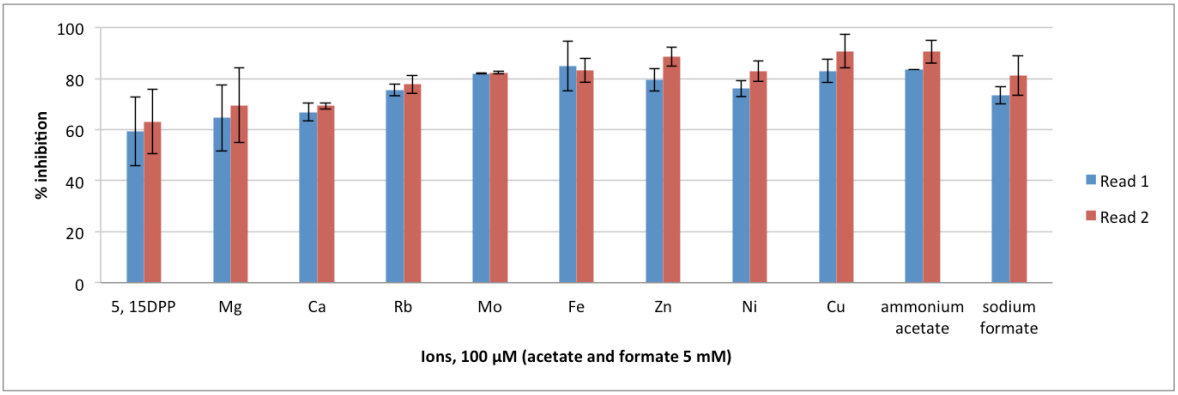


Figure 26: Effect of different ions on the inhibition. 5,15-DPP, STAT3ß 688s, Tris pH 7,4, 6-FAM-pYLPQTV, ions (100 µM). Error bars represent the SD of measurements, n=3. Read 1: t=0h, read 2: t=16h.

As a result of the previous experiments, the two promising additives, molybdic acid and sodium formate, were both added to the usual buffer (results are shown in table VIII). The final concentration of 5,15-DPP was 48µM, because it is poorly soluble in the assay solution. The addition didn't have the synergistic effect we were hoping for, as it did not enhance inhibition at all. In fact, the inhibition was two thirds lower than without any additives.

Table VIII: Inhibition in the presence of molybdic acid and sodium formate. 5,15-DPP, STAT3ß 688s, Tris pH 7,4; 6-FAM-pYLPQTV. SD was calculated using 3 measurements. Read 1: t=0h, read 2: t=16h.

5,15-DPP	Read 1	SD	Read 2	SD
10mM molybdic acid+ 100mM sodium formate	21,12	9,82	24,09	7,70

Effect of DNA and DTT

At this point we were still using 100 µM 5,15-DPP, which was poorly soluble in our assay solution. The results are shown in table IX. The negative percentages and relatively low

SD point to another process happening in the solution, affecting the signal. Because of the low solubility, this could be a sign of compound aggregation. There have been reports of a similar compound (porphyrin) that forms long, DNA-templated aggregates (59), which could also be the case here.

Table IX: Inhibition in the presence of DTT and DNA. 5,15-DPP, STAT3ß 688s, Tris pH 8,5; 6-FAM-pYLPQTV. SD was calculated using 3 measurements. Read 1: t=0h, read 2: t=16h.

100 µM 5,15-DPP	Read 1	SD	Read 2	SD
No DNA	-20,23	4,43	-24,72	7,82
No DNA, DTT	-4,14	1,77	-4,41	2,16
DNA	-45,54	5,15	-44,17	5,47
DNA, DTT	-52,61	10,86	-51,81	8,53

Effect of DNA

The values, shown in table X, were calculated using a blank sample without DNA and a blank sample containing the same amount of DNA as the samples with 5,15-DPP. The large difference between them is probably a consequence of the DNA molecule having an effect on the fluorescence polarization signal.

Table X: Inhibition in the presence of DNA. 5,15-DPP, STAT3ß 688s, Tris pH 8,5; 6-FAM-pYLPQTV, DNA. SD was calculated using 3 measurements. Read 1: t=0h, read 2: t=16h.

5, 15 DPP	Read 1	SD	Read 2	SD
Blank without DNA	58,13	5,11	40,86	4,24
Blank with DNA	2,55	11,46	1,78	7,35

Effect of mutated protein

As shown in table XI, the mutation lowers the inhibition for ~50%. This points to the binding being somewhat, but not entirely dependent on the –SH group on residue 542.

Table XI: Inhibition of normal and mutated protein. 5,15-DPP, STAT3ß 688s and C542S, Tris pH 8,5; 6-FAM-pYLPQTV. SD was calculated using 3 measurements. Read 1: t=0h, read 2: t=16h.

	Read 1	SD	Read 2	SD
688s	30,38	4,54	5,58	2,32
C542S	14,74	6,48	8,27	2,73

4.2.2. BP-1-102

BP-1-102 (figure 8) inhibits STAT3 activation and peptide interactions, both in vitro and in vivo.

Effect of DTT

DTT has almost no effect on the inhibition (results shown in table XII). Contrary to most other compounds, the inhibition is higher when DTT is added, and in both cases, it is higher after 16 hours.

Table XII: Inhibition in the presence of DTT. BP-1-102, STAT3ß 688s, Tris pH 7,4; 6-FAM-pYLPQTV. SD was calculated using 9 measurements. Read 1: t=0h, read 2: t=16h.

BP-1-102	Concentration	Read 1	SD	Read 2	SD
no DTT	100 µM	77,26	5,96	84,47	3,56
	10uM	21,54	4,41	27,69	8,82
DTT	100 µM	82,09	4,27	88,26	4,44
	10uM	23,92	6,17	34,47	8,60

At a higher pH (table XIII), the inhibition is also higher with DTT than without, and is in general higher than at pH 7,4. It is also higher when using 6-FAM-pYLPQ, which is the smaller of the two labelled peptides.

Table XIII: Inhibition in the presence of DTT. BP-1-102, STAT3ß 688s, Tris pH 8,5; 6-FAM-pYLPQ and pYLPQTV. SD was calculated using 9 measurements. Read 1: t=0h, read 2: t=16h.

pYLPQ		Read 1	SD	Read 2	SD
no DTT	BP-1-102	90,66	6,07	95,96	5,24
DTT	BP-1-102	95,58	6,14	96,06	1,96
pYLPQTV					
no DTT	BP-1-102	81,58	3,54	83,04	1,78
DTT	BP-1-102	90,97	4,62	88,52	1,56

Effect of pH

As shown in figure 27, BP-1-102 shows higher inhibition when the pH is higher than 7,4. It is equally high at pH 5,7, but this is not optimal for the protein and for DTT.

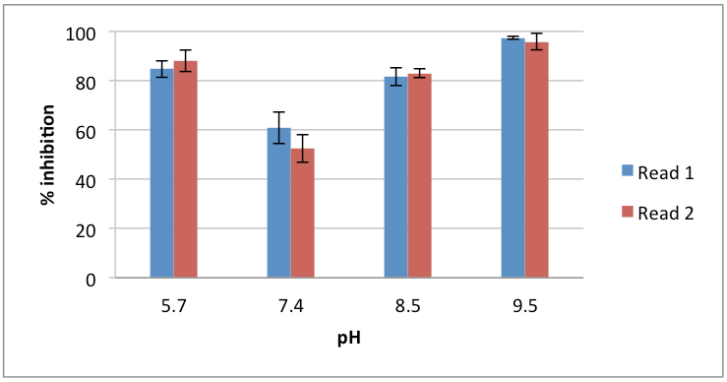


Figure 27: Effect of pH on the inhibition. BP-1-102, STAT3ß 688s, Bis Tris pH 5,7; Tris pH 7,4/8,5/9,5; 6-FAM-pYLPQTV. Error bars represent the SD of measurements, n=3. Read 1: t=0h, read 2: t=16h.

Effect of DNA and DTT

As shown in figure 28, the inhibition of an otherwise good inhibitor is lowered by the addition of DNA. These results are not reliable, because they were calculated using the blank without DNA. It is likely that the real inhibition is lower.

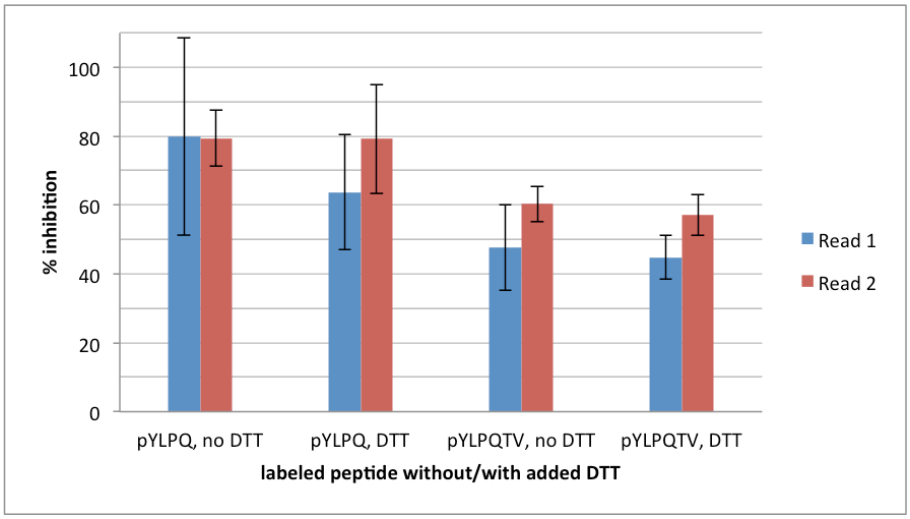


Figure 28: Inhibition in the presence of DTT and DNA. BP-1-102, STAT3ß 688s, Tris pH 8,5; 6-FAM-pYLPQ/ pYLPQTV, DNA. Error bars represent the SD of measurements, n=9. Read 1: t=0h, read 2: t=16h.

Effect of DNA

The inhibition after adding the DNA is only a few percent lower, suggesting that DNA has almost no effect on the labeled peptide binding to the SH2-domain after all (table XIV).

Table XIV: Inhibition in the presence of DNA. BP-1-102, STAT3ß 688s, Tris pH 8,5; 6-FAM-pYLPQTV, DNA. SD was calculated using 9 measurements. Read 1: t=0h, read 2: t=16h.

BP 1-102	Read 1	SD	Read 2	SD
Blank no DNA	89,92	3,30	90,77	3,05
Blank DNA	77,66	7,59	85,13	5,03

Effect of mutated protein

The mutation has no significant effect on BP-1-102, meaning the binding is not dependent on the presence of the –SH group on residue 542 (table XV).

Table XV: Inhibition of normal and mutated protein. BP-1-102, STAT3 688s, C542S, Tris pH 8,5; 6-FAM-pYLPQTV. SD was calculated using 9 measurements. Read 1: t=0h, read 2: t=16h.

BP-1-102	Read 1	SD	Read 2	SD
688s	81,58	3,54	83,04	1,78
C542S	86,78	3,56	85,22	1,87

4.2.3. Cryptotanshinone

Cryptotanshinone (figure 9) is a potent, selective inhibitor, isolated from a plant, Salvia milthiorrhiza bunge (Danshen).

Effect of DTT

When using 6-FAM-pYLPQ, the inhibition is higher with added DTT, but when using 6-FAM-pYLPQTV, the inhibition is lower (table XVI).

Table XVI: Inhibition in the presence of DTT. Cryptotanshinone, STAT3ß 688s, Tris pH 8,5; 6-FAM-pYLPQ and pYLPQTV. SD was calculated using 9 measurements. Read 1: t=0h, read 2: t=16h.

Cryptotanshinone				
pYLPQ	**Read 1**	**SD**	**Read 2**	**SD**
no DTT	35,99	4,23	47,25	4,82
DTT	60,27	2,00	61,03	1,90
pYLPQTV				
no DTT	47,74	6,48	50,91	8,19
DTT	36,11	3,69	34,09	3,24

Effect of DNA

The inhibition with added DNA is only 32 percent, which is 15% lower than in normal conditions (table XVII).

Table XVII: Inhibition in the presence of DNA. Cryptotanshinone, STAT3ß 688s, Tris pH 8,5; 6-FAM-pYLPQTV, DNA. SD was calculated using 9 measurements. Read 1: t=0h, read 2: t=16h.

cryptotanshinone	Read 1	SD	Read 2	SD
Blank no DNA	72,48	5,03	62,47	8,15
Blank DNA	32,39	12,01	31,70	15,27

Effect of mutated protein

The mutation at 542 seems to have no effect on the binding of cryptotanshinone, from which we can conclude that the binding is not dependent on the cysteine 542 residue (table XVIII). There is a notable difference, however, in the inhibition percentages after 16 hours. It is possible that the C542S protein has a higher affinity for cryptotanshinone than 688s.

Table XVIII: Inhibition of normal and mutated protein. Cryptotanshinone, STAT3ß 688s, Tris pH 8,5; pYLPQTV. SD was calculated using 9 measurements. Read 1: t=0h, read 2: t=16h.

Crypto-tanshinone	Read 1	SD	Read 2	SD
688s	47,74	6,48	36,11	3,69
C542S	45,85	6,78	46,69	4,85

4.2.4. EL group of compounds

This is a group of analogues of Stattic (figure 29), which is a known SH2-domain inhibitor. We hypothesized that we would measure high inhibition because of the similarity.

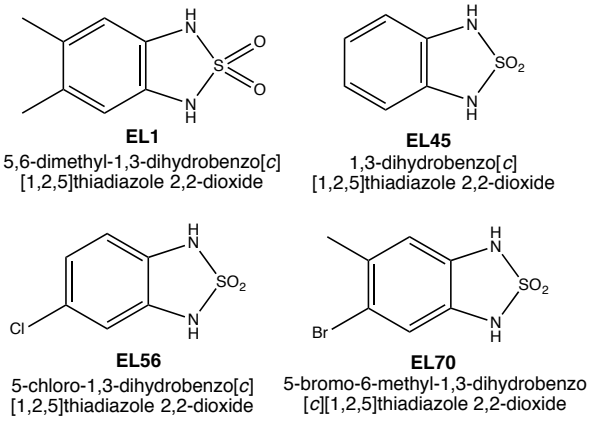


Figure 29: Structures of EL1, EL45, EL56 and EL70

Preliminary testing

As displayed in figure 30, screening at pH 7,4 reveals relatively low inhibition values; the highest is EL56 with 54%. Most of the compounds' inhibitions are lower after 16 hours, but for EL45 and 70, the inhibition is higher.

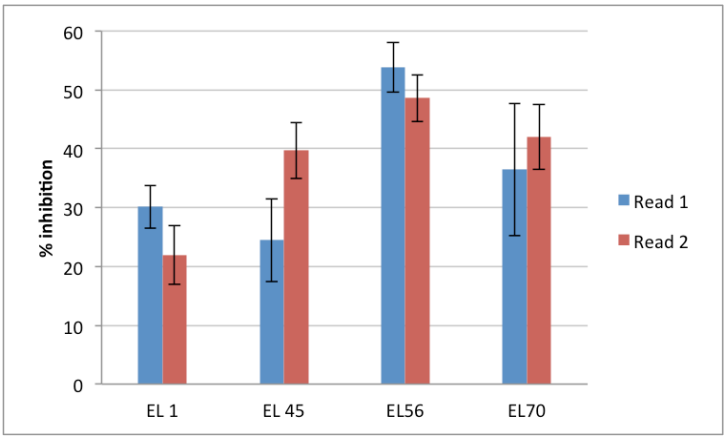


Figure 30: Inhibition at pH 7,4. EL compounds, STAT3ß 688s, Tris pH 7,4; 6-FAM-pYLPQTV. Error bars represent the SD of measurements, n=3. Read 1: t=0h, read 2: t=16h.

Effect of DTT

The EL group of compounds shows almost no inhibition when DTT is added (figure 31), just like it has been discovered for Stattic. There are some promising Stattic analogues in this group (EL1, 30, 56, 70). All four of them have similar structures: the 1,3-dihydrobenzo[*c*][1,2,5]-thiadiazole 2,2-dioxide, substituted on the benzene ring on sites 5 and/or 6 with small groups (CH_3, $2xCH_3$, Cl or Br). Larger groups, like NO_2 and CF_3

(EL48, EL55) are not beneficial for the binding, and neither are substitutions on the thiadiazole ring.

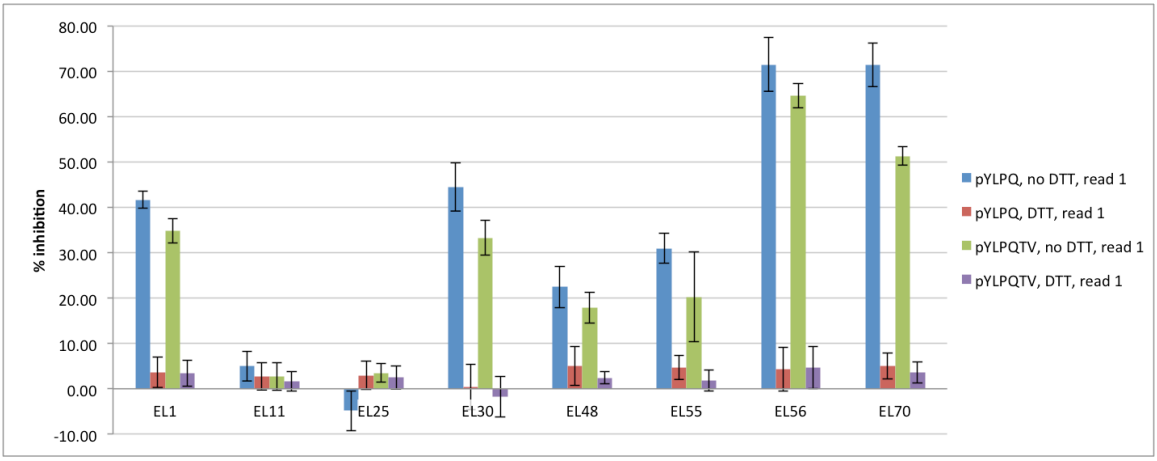


Figure 31: Inhibition in the presence of DTT. EL compounds, STAT3ß 688s, Tris pH 8,5; 6-FAM-pYLPQTV, DTT. Error bars represent the SD of measurements, n=9. Read 1: t=0h.

Effect of DNA

The inhibition of the only tested compound, EL56, is about 15% lower after adding DNA (table XIX).

Table XIX: Inhibition in the presence of DNA. EL56, STAT3ß 688s, Tris pH 8,5; 6-FAM-pYLPQTV, DNA. SD was calculated using 9 measurements. Read 1: t=0h, read 2: t=16h.

EL56	Read 1	SD	Read 2	SD
Blank no DNA	62,71	3,97	5,70	9,47
Blank DNA	52,40	4,89	15,98	6,91

Effect of mutated protein

The inhibition of all the EL compounds with the C542S protein is lower, showing that the –SH group at residue 542 is important for their binding (figure 32). This is expected, because the inhibition drops significantly after adding DTT, which is a sign of the importance of –SH groups. After 16 hours, all the 688s inhibition percentages are lower, suggesting that the protein has higher affinity for 6-FAM-pYLPQTV than the inhibitors.

The inhibition of the mutated protein only reached more than 50% of the normal protein's inhibition when using EL1 (23% and 35%, respectively).

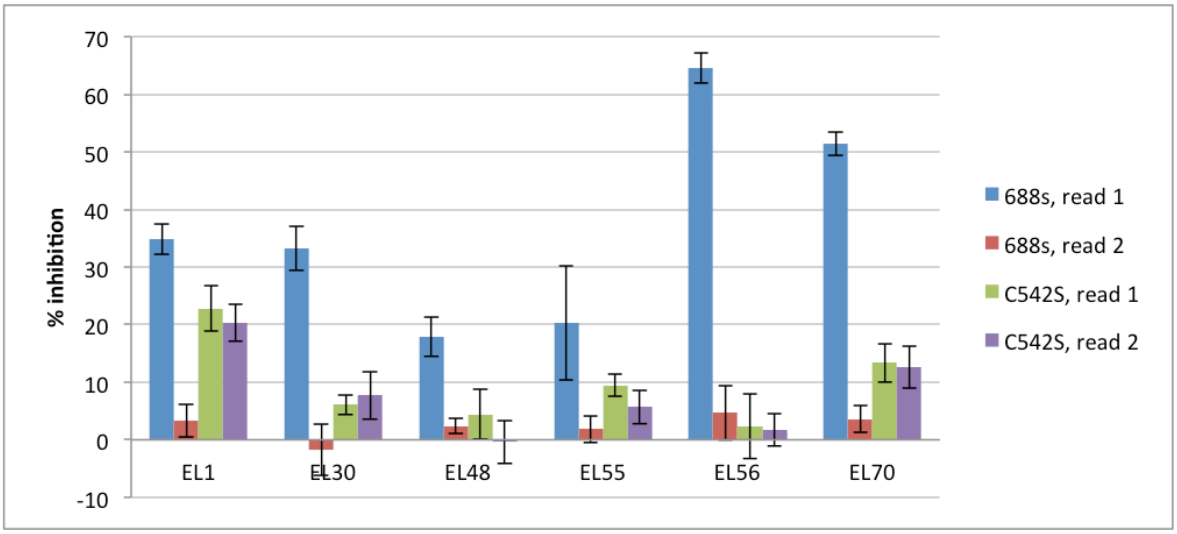


Figure 32: Inhibition of normal and mutated protein. EL compounds, STAT3ß 688s and C542S, Tris pH 8,5; 6-FAM-pYLPQTV. Error bars represent the SD of measurements, n=9. Read 1: t=0h, read 2: t=16h.

4.2.5. STK group of compounds

Preliminary test at Tris pH 7,4

The compound with the highest measured inhibition was STK244106 (20%). The graph (figure 33) shows all the compounds except STK240998 have lower inhibitions when adding DTT, showing the possible importance of –SH groups - cysteines (disulfide bonds). The notable structural difference between STK240998 and other compounds is the $–NO_2$ at site 4 of the phenyl ring.

All compounds except STK708561 have higher inhibition after 16 hours. However, since the standard deviations are so high, the results are not reliable.

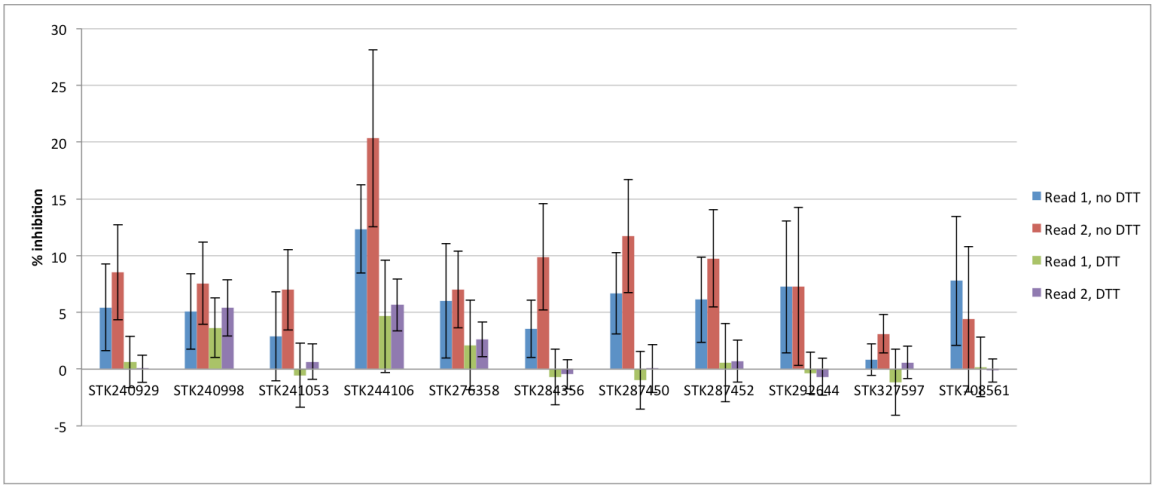


Figure 33: Inhibition at pH 7,4. STK compounds, STAT3ß 688s, Tris pH 7,4; 6-FAM-pYLPQTV. Error bars represent the SD of measurements, n=3. Read 1: t=0h, read 2: t=16h.

Effect of pH level

The highest signals were produced at pH 9,5 (results in figure 34 show read 1 at t=0h and in figure 35 read 2 at t=16h). Only at pH 9,5 the inhibition remained the same after 16 hours. Some of the results show negative inhibition at pH 5,7; further confirming these are not the optimal conditions for the protein. Measurements at both pH 7,4 and 8,5 showed signals that were also acceptable, the major difference being that using pH 7,4 resulted in higher signals after 16 hours. This could be caused by the pH difference, which could affect protein behaviour or the ionisation of the compounds.

Only three of the five compounds were tested with Tris pH 8,5.

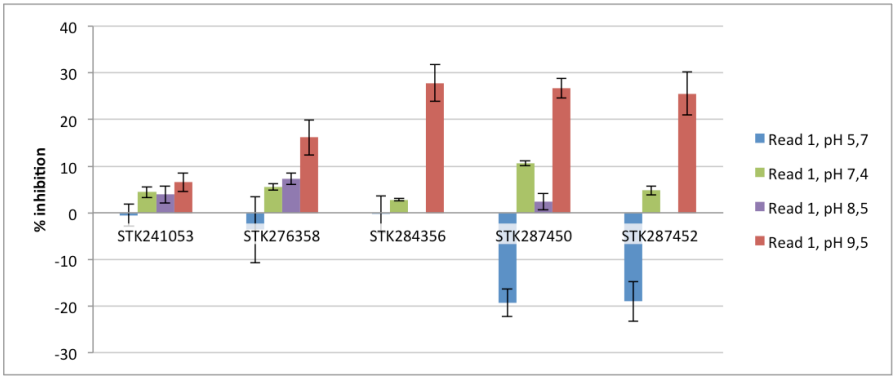


Figure 34: Inhibition at different pH levels, read 1 (t=0). STK compounds, STAT3ß 688s, Bis Tris pH 5,7; Tris pH 7,4; Tris pH 8,5; Tris pH 9,5; 6-FAM-pYLPQTV. Error bars represent the SD of measurements, n=3. Read 1: t=0h.

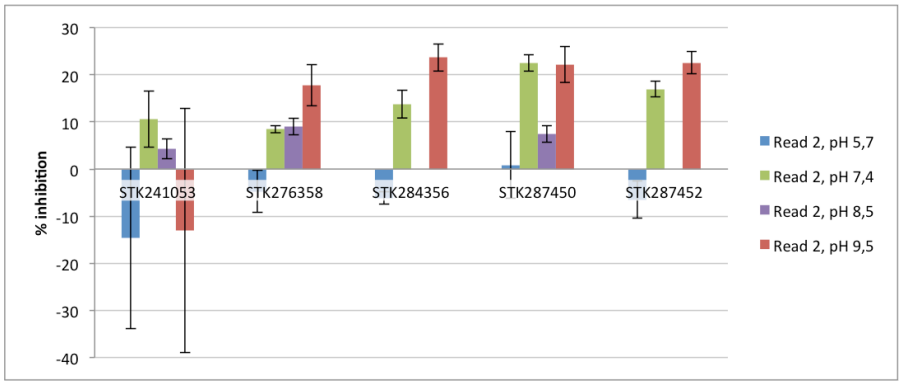


Figure 35: Inhibition at different pH levels, read 2. STK compounds, STAT3ß 688s, Bis Tris pH 5,7; Tris pH 7,4; Tris pH 8,5; Tris pH 9,5; 6-FAM-pYLPQTV, +16 hours. Error bars represent the SD of measurements, n=3. Read 2: t=16h.

Effect of DTT

DTT lowers the inhibition of all three tested compounds to 2% or less (with high standard deviations), meaning these STK compounds likely depend on –SH groups for binding (figure 36). Adding DTT has the same effect on other compounds in this group as well, as seen in figure 33.

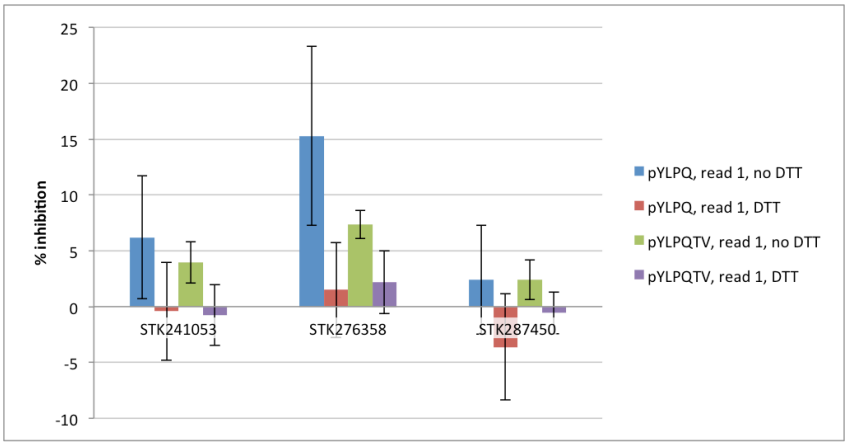


Figure 36: Inhibition in the presence of DTT. STK compounds, STAT3ß 688s, Tris pH 8,5; 6-FAM-pYLPQ and pYLPQTV, DTT. Error bars represent the SD of measurements, n=9. Read 1: t=0h.

Effect of DTT -/+ DNA

As seen previously, in figure 36, the addition of DTT lowers the inhibition substantially. Furthermore, adding DNA results in much larger standard deviation as usual (figure 37). DNA could be forming aggregates or blocking the binding site, although the crystal

structure suggests the latter is not possible. However, we have to keep in mind that the assays were performed at pH 9,5, which is far from the physiological pH, and even further away from the pH 7,0 that was used to obtain the crystal structure of STAT3ß bound to the DNA (58).

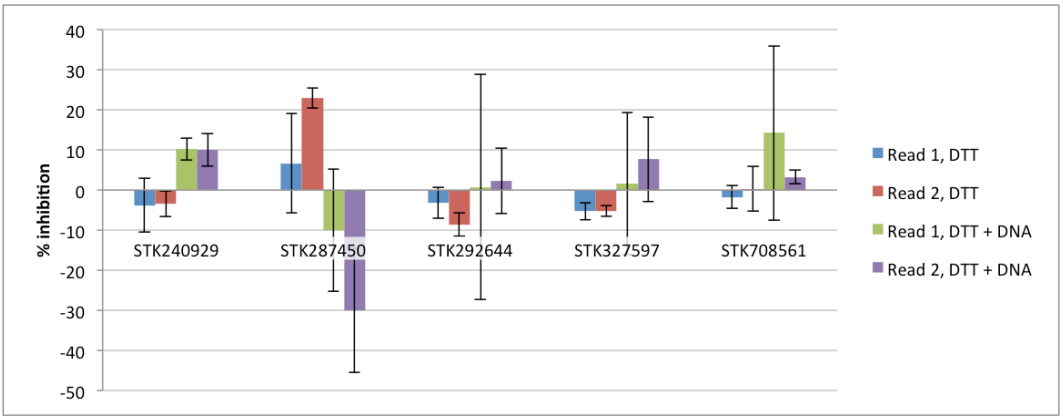


Figure 37: Inhibition in the presence of DTT and DNA. STK compounds, STAT3ß 688s, Tris pH 9,5; 6-FAM-pYLPQTV, DNA (6µL/ml protein). Error bars represent the SD of measurements, n=3. Read 1: t=0h, read 2: t=16h.

The results (figure 38) are not reliable because of the high standard deviations. However, it is visible that the inhibition values are lower when using 6-FAM-pYLPQTV, which is the largest molecule of the two.

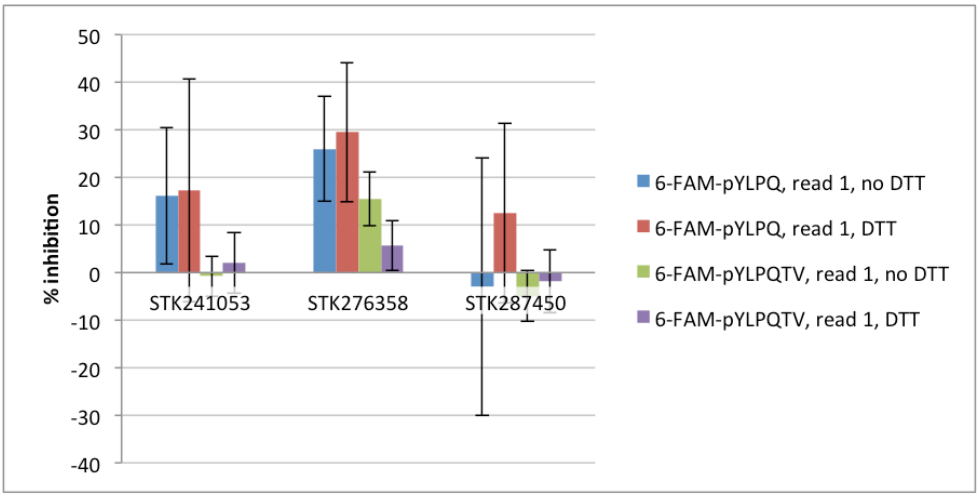


Figure 38: Inhibition in the presence of DTT and DNA. STK compounds, STAT3ß 688s, Tris pH 8,5;6-FAM-pYLPQ/pYLPQTV, DNA (6µL/mL protein). Error bars represent the SD of measurements, n=9. Read 1: t=0h.

Effect of DNA

All compounds that initially had more than 5% inhibition were tested by adding M67 DNA to the assay (results are pictured in figure 39). There were two blank rows, the first one which did not contain DNA and the second one that did. The results were calculated using the blank sample without and with DNA; the "with DNA" results were used. From results calculated using blank samples without DNA we could conclude that the DNA increases inhibition a lot. However, by calculating the results using blank samples with an appropriate DNA concentration, we learned that the DNA molecule must have changed the conditions of the assay and influenced the readings.

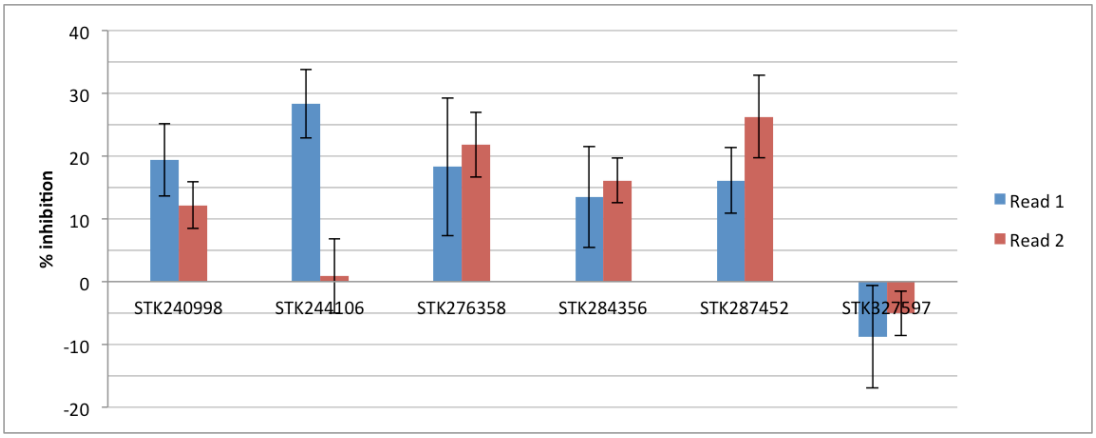


Figure 39: Inhibition in the presence of DNA. STK compounds, STAT3ß 688s, Tris pH 8,5; 6-FAM-pYLPQTV, DNA (blank+DNA). Error bars represent the SD of measurements, n=9. Read 1: t=0h, read 2: t=16h.

The different inhibition of STK327597 could be attributed to its different structure (figure 40), as it is the only molecule of the STK group that contains an 1,3-benzodioxole group.

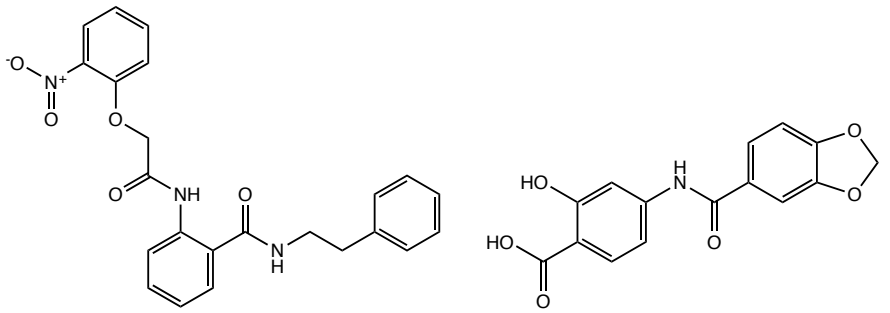


Figure 40: Structures of STK244106 and STK327597

Effect of mutated protein

Two compounds, STK240998 and STK276358 show a higher inhibition with C542, showing that the lack of the –SH group is actually preferable for their binding (figure 41). They have similar structures (figure 42). STK244106, which also shows relatively high inhibition of the C542S protein is different from STK240998 only by the different placement of the nitro group on the phenyl ring. The other three compounds have different, smaller structures.

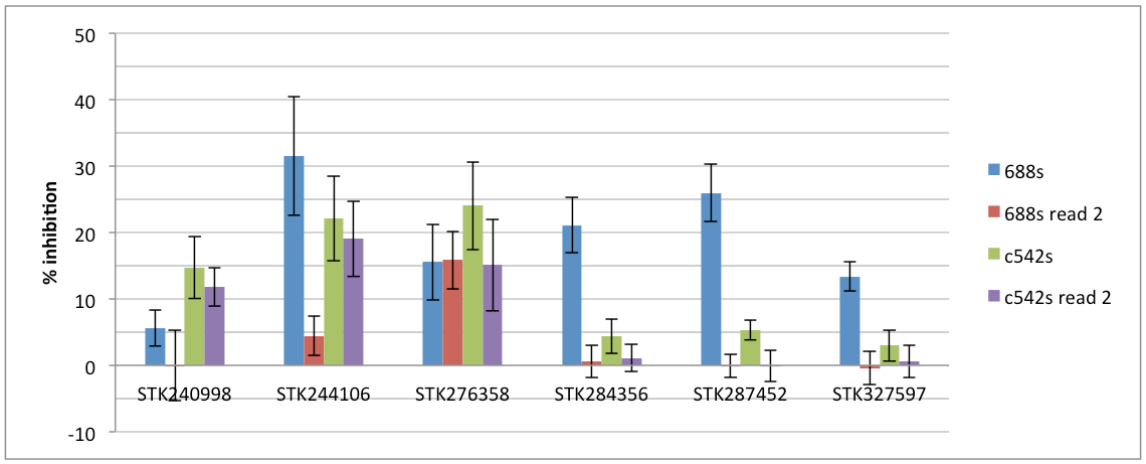
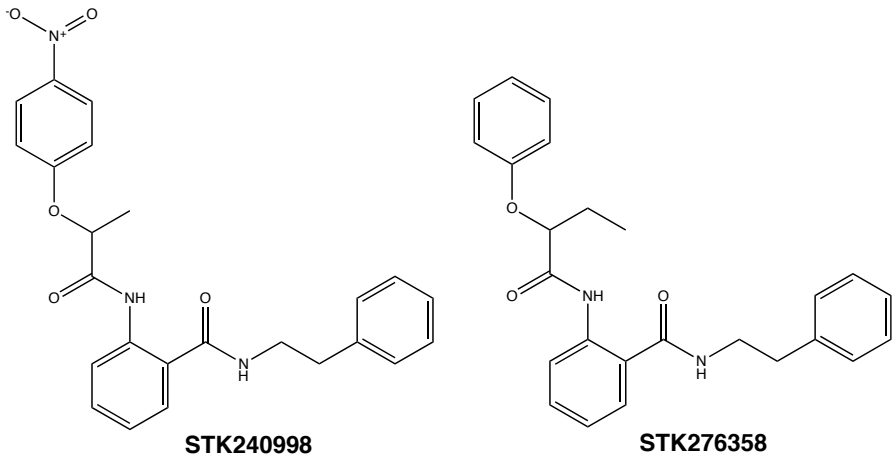


Figure 41: Inhibition of normal and mutated protein. STK compounds, STAT3ß 688s and C542S, Tris pH 8,5; 6-FAM-pYLPQTV. Error bars represent the SD of measurements, n=9. Read 1: t=0h, read 2: t=16h.



STK240998
(2-{[2-(4-nitrophenoxy)propanoyl]amino}
-N-(2-phenylethyl))benzamide

STK276358
2-[(2-phenoxybutanoyl)amino]
-N-(2-phenylethyl)benzamide

Figure 42: structures of STK240998 and STK276358

4.2.6. STL group of compounds

Preliminary test at Tris pH 7,4

Preliminary testing of all STL compounds shows similar results as the STK group – low inhibition, high standard deviations (figure 43). The addition of DTT lowers the inhibition.

After 16 hours, the inhibition is higher for all compounds in the no DTT group and for three compounds (STL065959, STL118659 and STL229411) in the DTT group.

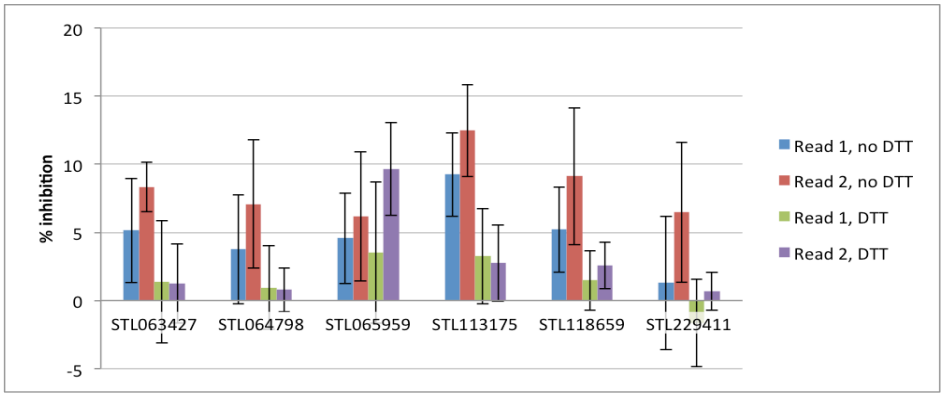


Figure 43: Inhibition at pH 7,4. STL compounds, STAT3ß 688s, Tris pH 7,4; 6-FAM-pYLPQTV. Error bars represent the SD of measurements, n=3. Read 1: t=0h, read 2: t=16h.

Effect of pH

The chart confirms what we have noticed in other groups of compounds as well: higher pH results in higher inhibition (figure 44). This is why we used Tris pH 8,5 for further experimentation.

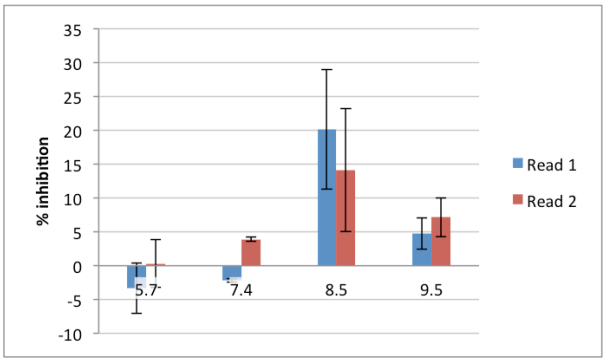


Figure 44: Inhibition at different pH levels. STL229411, STAT3ß 688s, Bis Tris/Tris, 6-FAM-pYLPQTV, no DTT. Error bars represent the SD of measurements, n=3. Read 1: t=0h, read 2: t=16h.

Effect of DTT

With pYLPQTV the first observation is the same as with pYLPQ: all the compounds show higher inhibition when DTT is not present (figure 45). STL063427 again shows an extremely high standard deviation, hinting at low solubility/aggregation.

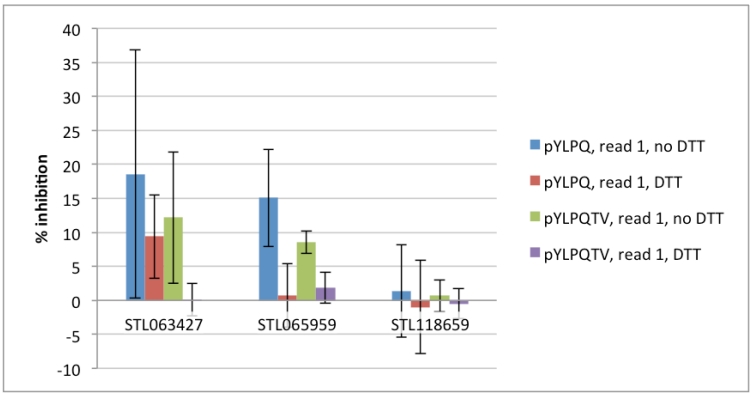


Figure 45: Inhibition in the presence of DTT. STL compounds, STAT3ß 688s, Tris pH 8,5; 6-FAM-pYLPQ and pYLPQTV, DTT. Error bars represent the SD of measurements, n=9. Read 1: t=0h.

Effect of DTT+DNA

Comparing these values with values from samples with no DNA shows that most of the values went up with the addition of DNA (figure 46). The exception is STL063427, where the standard deviation is too high for the value to be reliable.

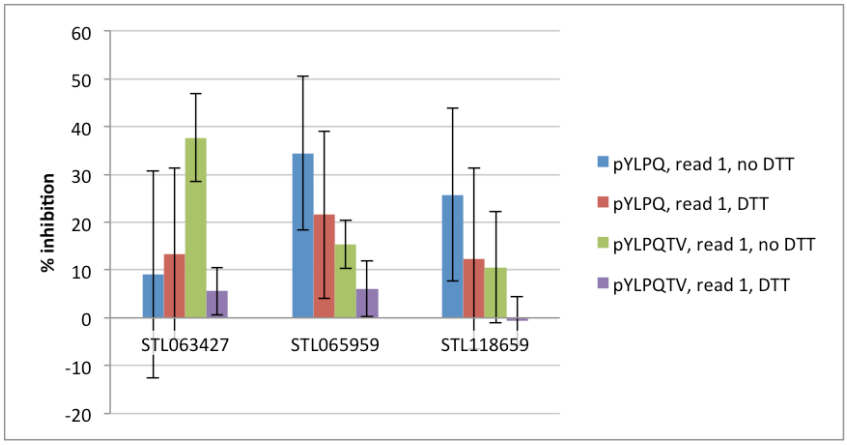


Figure 46: Inhibition in the presence of DTT and DNA. STL compounds, STAT3ß 688s, Tris pH 8,5; 6-FAM-pYLPQ and pYLPQTV, DTT, DNA. Error bars represent the SD of measurements, n=9. Read 1: t=0h.

Effect of DNA

The results were calculated using the blank with DNA. The standard deviations are relatively high, Compared to results of previous experiments, the compounds show higher inhibition of STAT3 when DNA is added (figure 47).

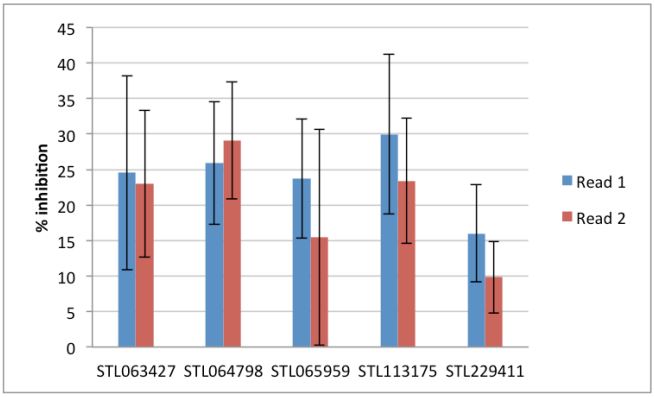


Figure 47: Inhibition in the presence of DNA. STL compounds, STAT3ß 688s, Tris pH 8,5; 6-FAM-pYLPQTV, DNA. Error bars represent the SD of measurements, n=9. Read 1: t=0h, read 2: t=16h.

Effect of mutated protein

Compounds STL064798 and STL229411 show a significantly lower inhibition with C542S (pictured in figure 48). They likely need the –SH group of the cysteine 542 residue to bind to the protein. Their structures are not similar.

The inhibition stays low even after 16 hours, meaning the speed of the bond formation is not important in this case. STL113176 shows roughly the same inhibition of both proteins (standard deviation is very high in both cases, meaning there are probably problems with solubility or aggregation), which shows the –SH group on residue 542 has no significant influence on binding. However, after 16 hours, the inhibition is dramatically different between the two proteins; 23% with C542S and dropping to only 2,09% with 688s. This suggests the compound binds more tightly to the mutated protein than to the original, STAT3ß 688s.

STL063427 and STL065959 (figure 49) show a higher inhibition of C542S than 688s, which could mean the lack of the –SH group somehow makes it easier to bind. They have

similar structures and the phenyl ring is substituted with –Br and –CH3 (STL063427) and –Br and –Cl (STL065959).

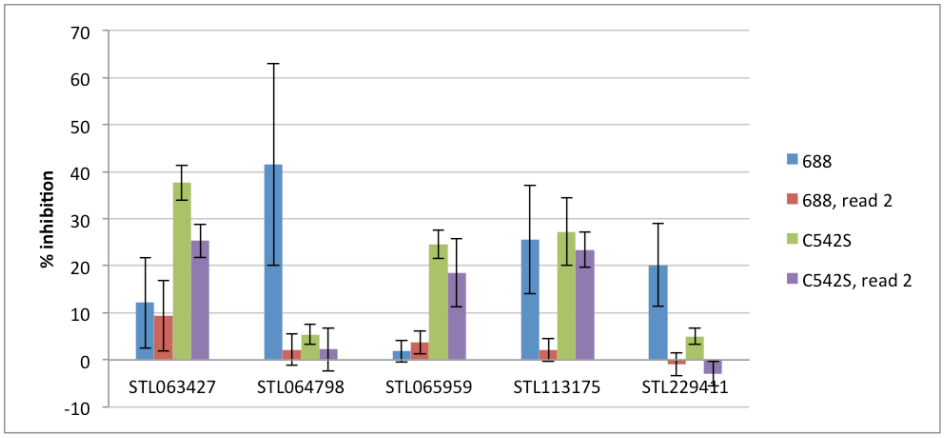
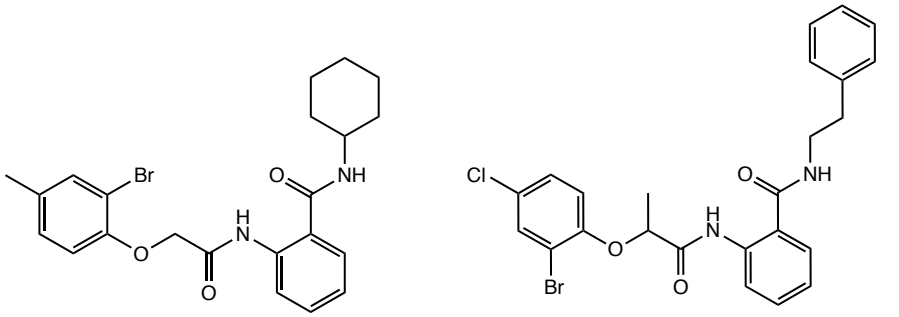


Figure 48: Inhibition of normal and mutated protein. STL compounds, STAT3ß 688s and C542S, Tris pH 8,5; 6-FAM-pYLPQTV. Error bars represent the SD of measurements, n=9. Read 1: t=0h, read 2: t=16h.



STL063427
2-{[(2-bromo-4-methylphenoxy)acetyl]amino}-N-cyclohexylbenzamide

STL065959
2-{[2-(2-bromo-4-chlorophenoxy)propanoyl]amino}-N-(2-phenylethyl)benzamide

Figure 49: structures of STL063427 and STL065959

Altogether in the STL series only 100 µM STL064798 and STL065959 reached higher than 30% inhibition. These are the only two compounds of the series, in which the phenyl ring is substituted with two halogen atoms.

4.3. Discussion

The screened compounds are hypothesized to inhibit the SH2 domain of STAT3. Based on the obtained data they could have different binding sites, however, to confirm that, displacement studies would be necessary.

Results using the physiological pH (7,4) would be applicable for use in humans in cancer treatment. Generally, we have seen that using higher pH levels results in higher inhibition for the two proteins we used. Using pH 8,5 works better during protein isolation as well. The behaviour or proteins during expression, purification and assays is pH-dependent and changes of the pH affect the ligands' binding: they can affect the structure of the protein and binding site. Assays at these pH levels could therefore be relevant for adjusting the crystallography protocol, but not necessarily for future use in therapy, as the inhibition for most of the EL, STK and STL compounds at the physiological pH is quite low under current conditions. Furthermore, these assays were performed using isolated proteins, so it is not reasonable to expect higher values in tests on cell lines or in the human body. The structures of the compounds need to be further optimized for the possibility of different *in vitro* and *in vivo* trials. Our results only show that the compounds we used are not optimal ligands of STAT3; it is not yet clear whether or not large peptidic ligands can be displaced by small molecules at all.

There are differences between inhibition values when using 6-FAM-pYLPQ and 6-FAM-pYLPQTV. Some compounds show lower inhibition with one of the labeled peptides. 6-FAM-pYLPQTV is the largest of the two and it can bind to the protein more strongly, so that could explain lower values when using it. In some cases we saw a decrease in inhibition after 16 hours. This could mean that more 6-FAM-pYLPQ/pYLPQTV is bound because the protein has a higher affinity to it than to the screened compound. Changes in the inhibition could also point to aggregation of the protein.

We added DTT or other reducing agents to discover whether or not the protein's intact -SH groups have any effect on the binding of our compounds. Where the inhibition dropped significantly after the addition, we can say –SH groups do affect the binding. This means at least one of the cysteins is essential for binding.

The mutated protein (STAT3ß C542S) was produced to see if the cysteine 542 is part of the binding site. If the inhibition drops significantly when using C542S, we can say this

cysteine is essential for binding. This happened with 5,15DPP (50% drop), EL1, EL70, STK244106, 284356, 287452, 327597, and STL064798, 229411. On the other hand, some compounds (STK240998, STK276358, STL063427, STL065959) showed higher inhibition when using the C542S mutant, meaning the lack of the –SH group is actually beneficial to their binding. These are therefore not optimal candidates for use in crystallization studies or therapy.

Percentages of inhibition drop significantly after adding DNA; sometimes the values are even negative. The negative values could mean that the DNA is bound to the protein, which is not inhibited by one of our compounds. This makes the volume of the new protein-labeled phosphopeptide complex much larger than before, so the signal is very high. When used in our formula, having a signal this high means the end result will be negative. This tells us that the compounds are not bound to the protein. However, when the result is not negative, we cannot conclude anything from these data because all we know is that the binding site of STAT3 is filled with one of our compounds, so the labeled phosphopeptide cannot bind to it. Therefore, with this method, it is not possible to find out whether or not our compounds are affecting the binding of DNA.

When adding DNA to the samples, there is a change in fluorescence polarization values, so it is always necessary to measure a blank sample with the same amount of added DNA. This is the way to calculate the real inhibition percentage. The fluorescence intensity of a sample can also be decreased by fluorescence quenching, which cannot be measured by measuring the fluorescence polarization of a blank sample.

DNA binds to the DNA-binding domain of STAT3, which is independent of the SH2 domain. It has been shown that the unphosphorylated STAT3 can bind to M67 DNA almost quantitatively, using a very similar interaction to the one of phosphorylated STAT3 (explained with the similarity to STAT1 unphosphorylated dimers); however, its affinity for DNA is lower than the phosphorylated dimers' (23). DNA binding shows a decrease of the inhibition. It has been suggested that the decrease could be a consequence of dimerization and fewer labeled peptides binding to STAT3. However, this does not seem probable, because unphosphorylated DNA binds to DNA as either a monomer or a dimer, connected through N-terminal domain interactions. These interactions are supposed to be structurally far from the SH2 domain; therefore, they should not interfere with the binding of the labeled peptide.

5. Conclusions

The first aim of our work was to improve the fluorescence polarization assay. We changed it by using both 6-FAM-pYLPQ and 6-FAM-pYLPQTV instead of 6-FAM-pYLKTK as labeled peptides in the assay; we continued adding 5% glycerol to the samples and performing the assays at a higher pH than before, at pH 8,5, because it contributed to more potent inhibition. However, this also meant the data would not necessarily be applicable to the physiological conditions. We kept using dithiothreitol (DTT) as a reducing agent because of its use in all steps of the protein extraction process and because of the low price.

Only a few of our new compounds reached inhibition higher that 50% at pH 8,5 (100 μM): the highest inhibition, up to 70% was shown by the EL series (EL 56, EL 70), which is not surprising, as they are analogues of Stattic, a known STAT3 inhibitor. Members of the STK and STL series only reached up to 30% inhibition, the best inhibitors were STK244106, STL064798 and STL065959. The screening of our small compound library was therefore not as successful as we had hoped. These numbers could be increased by further modifying the assay conditions, or, preferably, by modifying the compounds' structures. Nevertheless, from the experiments we learned how we should change the structures of compounds to develop new, better derivatives and which functional groups are most beneficial for binding.

By using different additives we were able to better understand how and where the compounds are binding to the SH2 domain. They might not bind to exactly the same site, because there are visible differences in binding with and without added reducing agents and by binding or not binding to the mutated protein C542S. However, displacement studies would be necessary to confirm this.

Cysteine residues are important for the binding of the EL series and most members of the STK and STL series. Cysteine 542 in particular is important for the EL series and partly for binding of 5,15-diphenylporphyrin; at the same time it has no influence on the binding of BP-1-102 and cryptotanshinone. It is the opposite for STK and STL series, where various compounds show higher inhibition when using the mutated protein.

The method we were using is not suitable for investigating whether or not the inhibitors affect the binding of STAT3 to DNA. When an inhibitor is bound to the protein, there is no way of knowing whether DNA is bound too.

As the assays were conducted with isolated proteins, it is not realistic to expect the compounds with low inhibition in these conditions to have higher inhibition when faced with tests on cell lines or the ADME processes in the body. However, the screening is a useful method for deciding which potential inhibitors are worth testing on more elaborate and expensive systems.

In future experiments it would be beneficial to first test all the compounds for solubility, reactivity with the assay solution, and binding to the polypropylene plate. Expressing and purifying larger amounts of the used proteins and then freezing them would eliminate the variability of the protein between different batches. It would also be better to refrain from using non-physiological pH values, as this makes the results non-comparable to the *in vivo* conditions and do not correlate the possible use in humans.

6. References

1. Latchman D S: Transcription factors: An overview. Int. J. Biochem. Cell Biol. 1997; 29: 1305–1312.
2. Alberts B et al.: Essential Cell Biology, 3rd edition, Garland Science, New York, USA, 2009: 236-240, 717-724.
3. Darnell J E: Transcription factors as targets for cancer therapy. Nat. Rev. Cancer 2002; 2: 740–749.
4. Darnell J E, Kerr I M & Stark G R: Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. Science 1994; 264: 1415–1421.
5. Yu H, Pardoll D & Jove R: STATs in cancer inflammation and immunity: a leading role for STAT3. Nat. Rev. Cancer 2009; 9: 798–809.
6. Horvath C M: STAT proteins and transcriptional responses to extracellular signals. Trends Biochem. Sci. 2000; 25: 496–502.
7. Levy D E & Darnell J E: Stats: transcriptional control and biological impact. Nat. Rev. Mol. Cell Biol. 2002; 3: 651–662.
8. Zhang X & Darnell J E: Functional Importance of Stat3 Tetramerization in Activation of the α2-Macroglobulin Gene. J. Biol. Chem. 2001; 276: 33576–33581.
9. Calò V et al.: STAT Proteins: From Normal Control of Cellular Events to Tumorigenesis. J. Cell. Physiol. 2003; 197: 157–168.
10. Yu H & Jove R: The STATs of cancer--new molecular targets come of age. Nat. Rev. Cancer 2004; 4: 97–105.
11. Bromberg J F: Stat proteins and oncogenesis. J. Clin. Invest 2002; 109: 1139–1142.
12. Takeda K et al.: Targeted disruption of the mouse Stat3 gene leads to early embryonic lethality. Proc. Natl. Acad. Sci. U. S. A. 1997; 94: 3801–3804.
13. Durbin J E, Hackenmiller R, Simon M C & Levy D E: Targeted disruption of the mouse Stat1 gene results in compromised innate immunity to viral disease. Cell 1996; 84: 443–450.

14. Akira S: Roles of STAT3 defined by tissue-specific gene targeting. Oncogene 2000; 19: 2607–2611.

15. STAT4 signal transducer and activator of transcription 4 [Homo sapiens (human)]. http://www.ncbi.nlm.nih.gov/gene/6775 [Accessed 29/08/2015]

16. STAT5B signal transducer and activator of transcription 5B [Homo sapiens (human)]. http://www.ncbi.nlm.nih.gov/gene/6777 [Accessed 29/08/2015]

17. STAT6 signal transducer and activator of transcription 6, interleukin-4 induced (Homo sapiens (human)). http://www.ncbi.nlm.nih.gov/gene/6778 [Accessed 29/08/2015]

18. Furqan M et al.: STAT inhibitors for cancer therapy. J. Hematol. Oncol. 2013; 6: 90.

19. Yue P, Turkson J: Targeting STAT3 in cancer: how successful are we? Expert Opin Investig Drugs. Author manuscript. 2009; 18: 45–56.

20. Zhang T, Kee W H, Seow K T, Fung W & Cao X: The coiled-coil domain of Stat3 is essential for its SH2 domain-mediated receptor binding and subsequent activation induced by epidermal growth factor and interleukin-6. Mol. Cell. Biol. 2000; 20: 7132–7139.

21. Haghikia A, Ricke-Hoch M, Stapel B, Gorst I & Hilfiker-Kleiner D: STAT3, a key regulator of cell-to-cell communication in the heart. Cardiovasc. Res. 2014; 102: 281–9.

22. Wen Z, Zhong Z & Darnell J E: Maximal activation of transcription by Stat1 and Stat3 requires both tyrosine and serine phosphorylation. Cell 1995; 82: 241–250.

23. Nkansah E et al.: Observation of unphosphorylated STAT3 core protein binding to target dsDNA by PEMSA and X-ray crystallography. FEBS Lett. 2013; 587: 833–9.

24. Becker S, Groner B & Müller C W: Three-dimensional structure of the Stat3β homodimer bound to DNA. Nature 1998; 394: 145–152.

25. Schaefer T S, Sanders L K & Nathans D: Cooperative transcriptional activity of Jun and Stat3 beta, a short form of Stat3. Proc. Natl. Acad. Sci. U. S. A. 1995; 92: 9097–9101.

26. Chakraborty A & Tweardy D J: Granulocyte colony-stimulating factor activates a 72-kDa isoform of STAT3 in human neutrophils. J. Leukoc. Biol. 1998; 64: 675–

680.

27. Hevehan D L, Miller W M & Papoutsakis E T: Differential expression and phosphorylation of distinct STAT3 proteins during granulocytic differentiation. Blood 2002; 99: 1627–1637.

28. Maritano D et al.: The STAT3 isoforms alpha and beta have unique and specific functions. Nat. Immunol. 2004; 5: 401–409.

29. Huang Y et al.: Stat3 isoforms, α and β, demonstrate distinct intracellular dynamics with prolonged nuclear retention of Stat3β mapping to its unique C-terminal end. J. Biol. Chem. 2007; 282: 34958–34967.

30. Lim C P & Cao X: Structure, function, and regulation of STAT proteins. Mol. Biosyst. 2006; 2: 536–550.

31. Bar-Natan M, Nelson EA, Xiang M, Frank DA: STAT signaling in the pathogenesis and treatment of myeloid malignancies. JAKSTAT 2012; 2: 55-64.

32. Bromberg J & Darnell J E: The role of STATs in transcriptional control and their impact on cellular function. Oncogene 2000; 19: 2468–2473.

33. Yuan Z-L, Guan, Y-J, Chatterjee D & Chin Y E: Stat3 dimerization regulated by reversible acetylation of a single lysine residue. Science 2005; 307: 269–273.

34. Siveen K S et al.: Targeting the STAT3 signaling pathway in cancer: role of synthetic and natural inhibitors. Biochim. Biophys. Acta 2014; 1845: 136–54.

35. Fagard R, Metelev V, Souissi I & Baran-Marszak F: STAT3 inhibitors for cancer therapy: Have all roads been explored? JAK-STAT 2 2013; 2: e22882-1–9.

36. Timofeeva O et al.: Mechanisms of unphosphorylated STAT3 transcription factor binding to DNA. J. Biol. Chem. 2012; 287: 14192–200.

37. Levy D E & Lee C K: What does Stat3 do? J. Clin. Invest. 2002; 109: 1143–1148.

38. Kusaba T et al.: Activation of STAT3 is a marker of poor prognosis in human colorectal cancer. Oncol. Rep. 2006; 15: 1445–51.

39. Deng J, Grande F & Neamati N: Small molecule inhibitors of Stat3 signaling pathway. Curr. Cancer Drug Targets 2007; 7: 91–107.

40. Coleman D R et al.: Investigation of the binding determinants of phosphopeptides targeted to the Src homology 2 domain of the signal transducer and activator of

transcription 3. Development of a high-affinity peptide inhibitor. J. Med. Chem. 2005; 48: 6661–6670.

41. Schust J, Sperl B, Hollis A, Mayer T U & Berg T: Stattic: A Small-Molecule Inhibitor of STAT3 Activation and Dimerization. Chem. Biol. 2006; 13: 1235–1242.

42. Zhang X et al.: Orally bioavailable small-molecule inhibitor of transcription factor Stat3 regresses human breast and lung cancer xenografts. Proc. Natl. Acad. Sci. 2012; 109: 9623–9628.

43. Sigma Aldrich: Cryptotanshinone. http://www.sigmaaldrich.com/catalog/product/sigma/c5624?lang=en&region=SI [Accessed 01/09/2015]

44. Uehara Y, Mochizuki M, Matsuno K, Haino T & Asai A: Novel high-throughput screening system for identifying STAT3-SH2 antagonists. Biochem. Biophys. Res. Commun. 2009; 380: 627–631.

45. Scapin G: Structural biology and drug discovery. Curr. Pharm. Des. 2006; 12: 2087–2097.

46. Smyth M S & Martin J H: X Ray Crystallography. Mol. Pathol. 2000; 53: 8–14.

47. Schust J & Berg T: A high-throughput fluorescence polarization assay for signal transducer and activator of transcription 3. Anal. Biochem. 2004; 330: 114–118.

48. Rossi A M & Taylor C W: Analysis of protein-ligand interactions by fluorescence polarization. Nat. Protoc. 2001; 6: 365–387.

49. Technical Resource Guide: Fluorescence Polarization, 4th edition. Invitrogen Corporation 2006.

50. Kortylewski M, Jove R & Yu H: Targeting STAT3 affects melanoma on multiple fronts. Cancer Metastasis Rev. 2005; 24: 315–327.

51. Schust J & Berg T: A high-throughput fluorescence polarization assay for signal transducer and activator of transcription 3. Anal. Biochem. 2004; 330: 114–118.

52. Wu P, Brasseur M & Schindler U: A high-throughput STAT binding assay using fluorescence polarization. Anal. Biochem. 1997; 249: 29–36.

53. Burgoyne R, Wilderspin A F & Nichols C: Mechanism of action of Stattic

analogues is by binding to Cysteine residue 542 of the Signal Transducer and Activator of Transcription 3 (STAT3) Protein. Unpublished manuscript, University College London, School of Pharmacy, UK 2014.

54. Ren Z, Cabell L A, Schaefer T S & McMurray J S: Identification of a high-affinity phosphopeptide inhibitor of Stat3. Bioorganic Med. Chem. Lett. 2003; 13: 633–636.

55. Moerke N J: Fluorescence Polarization (FP) Assays for Monitoring Peptide-Protein or Nucleic Acid–Protein Binding. Curr. Protoc. Chem. Biol 2009; 1: 1–15.

56. Vagenende V, Yap M G S & Trout B L: Mechanisms of protein stabilization and prevention of protein aggregation by glycerol. Biochemistry 2009; 48: 11084–96.

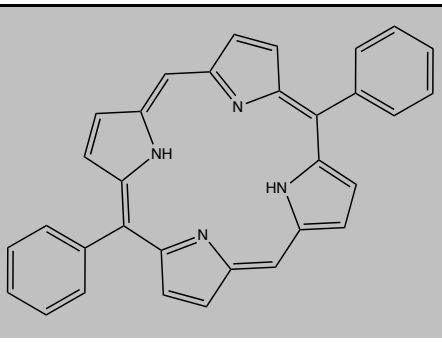
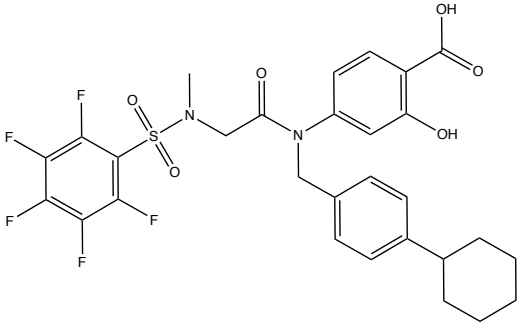
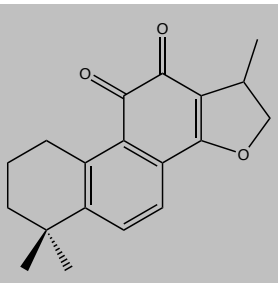
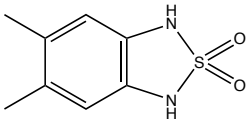
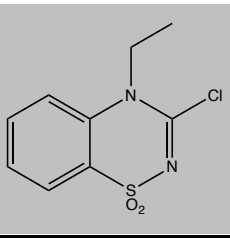
57. Getz E B, Xiao M, Chakrabarty T, Cooke R & Selvin P R: A comparison between the sulfhydryl reductants tris(2-carboxyethyl)phosphine and dithiothreitol for use in protein biochemistry. Anal. Biochem. 1999; 273: 73–80.

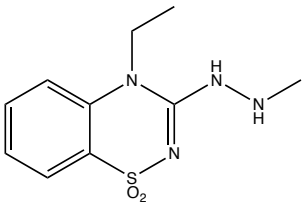
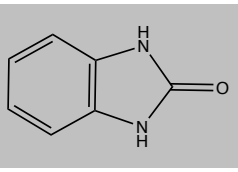
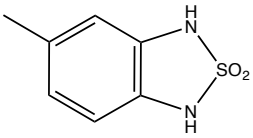
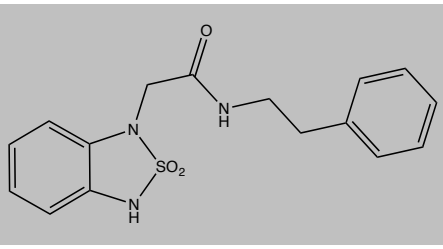
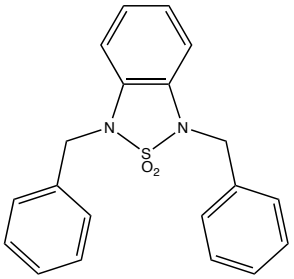
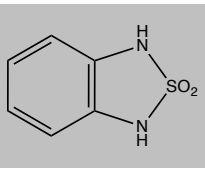
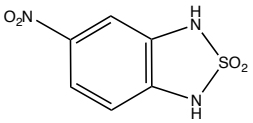
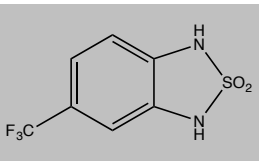
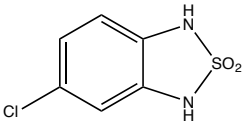
58. Becker S, Groner B, Müller C: Three-dimensional structure of the Stat3beta homodimer bound to DNA. Nature 1998; 394: 145–151.

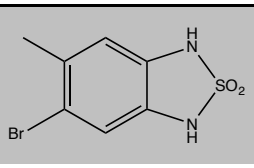
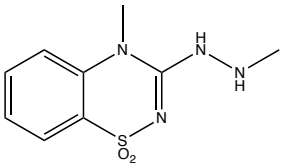
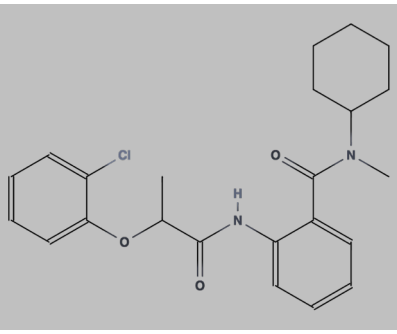
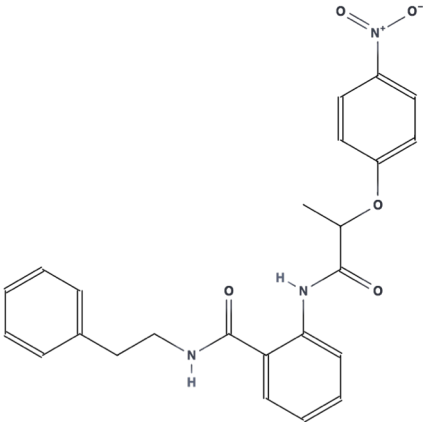
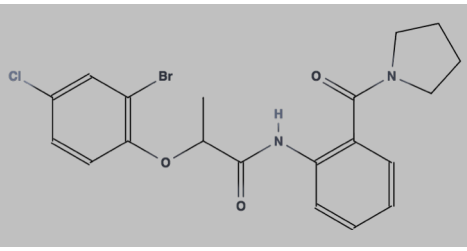
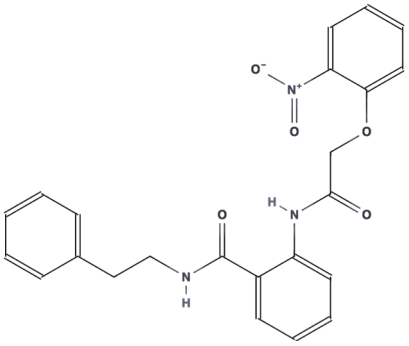
59. Sazanovich I, Panarin A, Stupak A, Terekhov S & Chirvony V: Fluorescence properties of the dicationic porphyrin 5,15-DiMPyP orderly aggregated along DNA surface. Photochem. Photobiol. Sci. 2008; 7: 1091–1098.

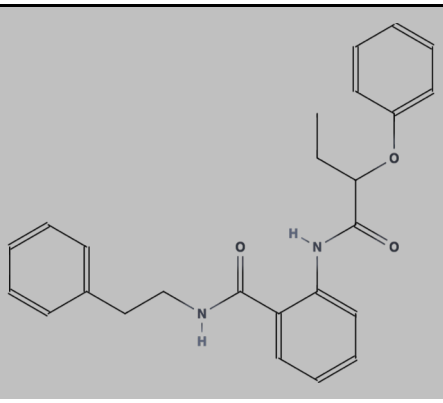
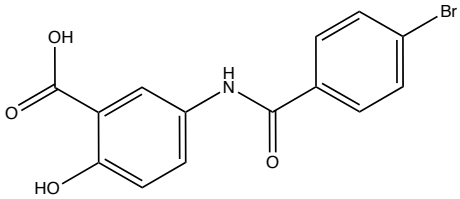
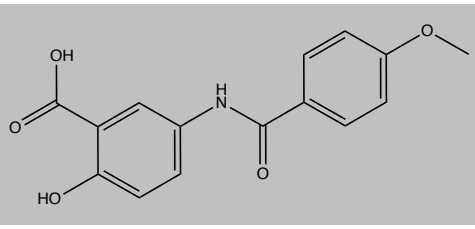
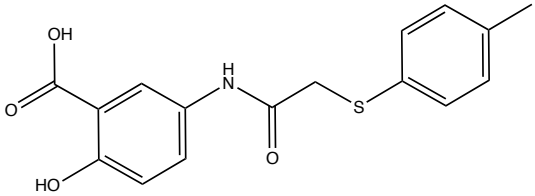
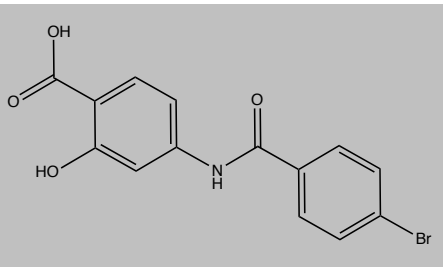
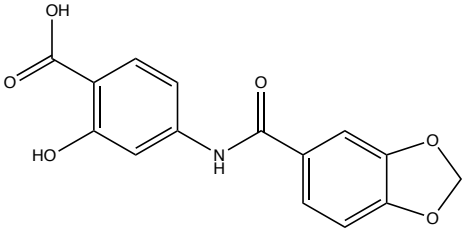
7. Supplemental information

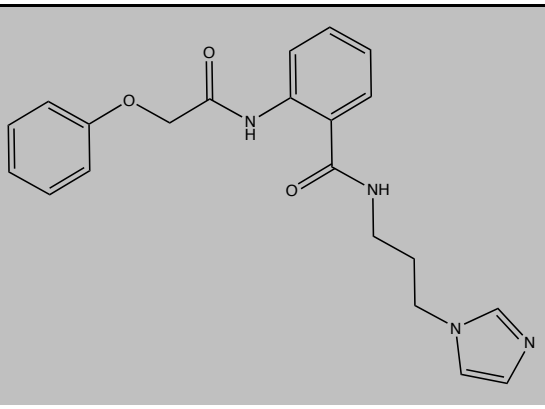
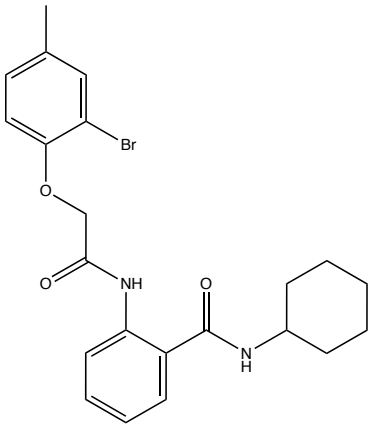
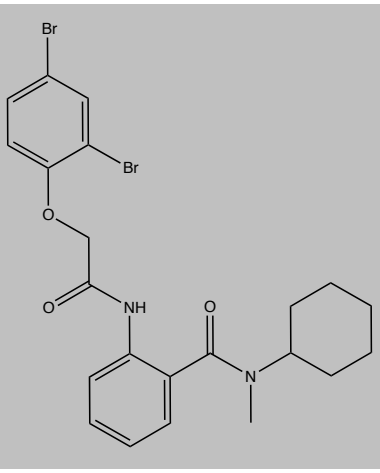
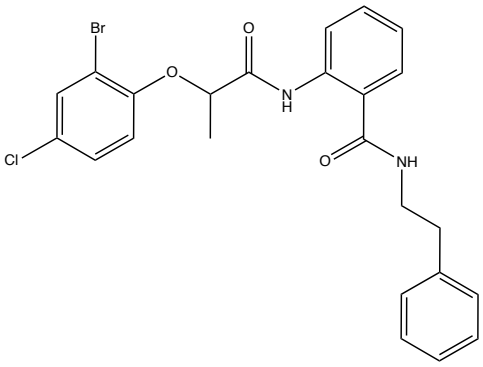
7.1. Compounds used for screening

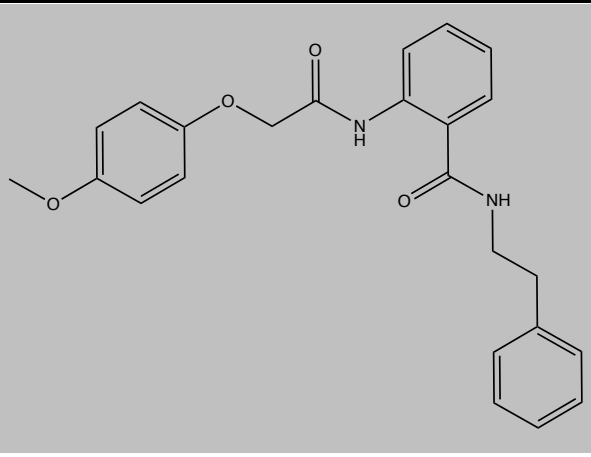
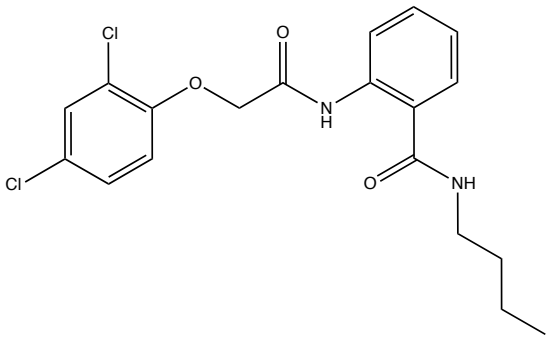
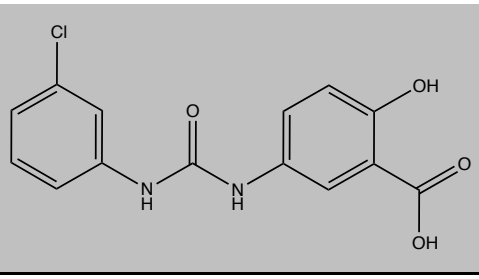
Name	Structure	MW (g/mol)
5, 15-DPP		462,54
BP-1-102		626,59
Cryptotanshinone		296,36
EL1		198,24
EL10		244,70

EL11		254,31
EL25		134,13
EL30		184,22
EL38		319,38
EL44		350,43
EL45		170,19
EL48		215,19
EL55		238,19
EL56		204,64

EL70		264,12
EL8a		240,28
STK240929		414,93
STK240998		433,46
STK241053		451,74
STK244106		419,43

STK276358		402,49
STK284356		336,14
STK287450		287,27
STK287452		317,36
STK292644		336,14
STK327597		301,25

STK708561		378,42
STL063427		445,35
STL064798		524,25
STL065959		501,80

STL113175		404,46
STL118659		395,28
STL229411		306,70