Development and validation of new approaches for quantitative intermolecular interaction analysis by MicroScale Thermophoresis under near-native experimental conditions

Dissertation

zur Erlangung des Doktorgrades
der Naturwissenschaften
(Dr. rer. nat.)
der Fakultät für Biologie und Vorklinische Medizin
der Universität Regensburg

vorgelegt von

Tanja Astrid Bartoschik

aus

Darmstadt

im Jahr 2018

Das Promotionsgesuch wurde eingereicht am: 28.03.2018

Die praktischen Arbeiten wurden von März 2015 bis März 2018 bei NanoTemper Technologies GmbH in München durchgeführt. Hier wurde die Arbeit angeleitet von Dr. Dennis Breitsprecher und Dr. habil. Nuska Tschammer. Die universitäre Betreuung oblag Prof. Dr. Gernot Längst.

Unterschrift:

II

Acknowledgements

Abstract

The analysis of molecular binding events is of great importance for both basic research and drug development. Several biophysical methods can be applied for the characterization of interactions, regarding affinity, kinetics or thermodynamic parameters. One commonly applied biophysical method is label-free MicroScale Thermophoresis (MST), which provides K_d determination under label-free and in-solution conditions. This method utilizes the intrinsic fluorescence of tryptophan, an amino acid incorporated in the vast majority of proteins. Up to now, label-free MST was restricted to quantification of interactions in which only one binding partner exhibits fluorescence in the detection wavelength. This excludes proteins as second binding partners, along with a significant number of small molecules or fragments. The reason for this is that preferred scaffolds used for the synthesis of small molecules and fragments often include indole or similar ring systems, which lead to fluorescence interference in label-free MST assays. Because of these reasons, the main goal of this thesis was to explore approaches which would enable a broader applicability of label-free MST and facilitate the quantification of intermolecular interactions under close-to-native MST-based experimental conditions.

In a first approach, a modified emission filter was tested to potentially cut off any unwanted signal arising from interfering compounds to a higher extent compared to the established filter. The modified emission bandwidth indeed decreased the extent of fluorescence interference caused by compounds. However, as many compounds exhibit so-called privileged structures, such as indole motifs which are also present in tryptophan residues of proteins, the number of compounds that still interfere in label-free assays highly depends on the compound library used and remains hard to predict, as different chemical substituents can already drastically alter the emission spectrum of compounds. As this modification of the device´s optical system did not provide an overall solution for interfering compounds and in addition could not be used for the analysis of protein-protein interactions (PPIs), a second strategy was developed. Here, a composition gradient titration strategy in combination with data analysis based on a least-mean-square approximation was applied for the quantification of PPIs in a label-free MST approach. The obtained K_d values were in good agreement with data obtained from standard (= non-label-free) MST experiments. Although in general, this approach was suitable for the quantification of PPIs, simulations of various experimental conditions revealed several limitations and restrictions regarding proteins´ fluorescence intensities, F_{norm} values and start concentrations in general.

Due to the limitations of both strategies, a compromise strategy between preserving the proteins´ native structure as much as possible, while at the same time making use of the

advantages of a fluorescent tag was applied. Therefore, target proteins were site-specifically labeled at their His$_6$-tag using tris-NTA fluorophores, which was found to provide robust results for both, protein-small molecule and PPIs. Using such a site-specific labeling approach, the proteins´ native structure is highly preserved and interference of fluorophores with ligand binding is prevented. Among three different fluorophores tested, RED-tris-NTA proved to be the most suitable dye for this purpose. Furthermore, this approach offered the possibility to directly measure MST in crude cell lysate, which further increased the close-to-native format. In addition, such measurements wouldn´t be possible using label-free MST, which further highlights the advantages of the site-specific labeling approach.

Zusammenfassung

Die Analyse molekularer Interaktionen ist von großer Bedeutung für die Grundlagenforschung und für die Entwicklung neuer Arzneistoffe. Für ihre Charakterisierung stehen verschiedene Methoden zur Verfügung, mit denen Affinität, Kinetik oder auch thermodynamische Parameter analysiert werden können. Eine dieser Methoden ist markierungsfreie mikroskalierte Thermophorese (MST), mit der K_ds markierungsfrei und frei in Lösung bestimmt werden können. Dafür verwendet diese Methode die Fluoreszenz von Tryptophan, einer Aminosäure, die in den meisten Proteinen vorkommt. Bisher war diese Methode auf die Analyse von Interaktionen beschränkt, bei denen nur einer der beiden Bindungspartner in dem detektierten Wellenlängenbereich fluoresziert. Damit waren nicht nur Protein-Protein Interaktionen (PPI) von der Analyse ausgeschlossen, sondern auch eine große Anzahl an kleinen Molekülen und Fragmenten. Der Grund dafür ist, dass viele dieser kleinen Moleküle und Fragmente sogenannte privilegierte Strukturen wie Indole oder ähnliche Ringstrukturen aufweisen, was zur Fluoreszenzüberschneidung führt. Daher war das Hauptziel dieser Dissertation Möglichkeiten zu finden, die ein breiteres Anwendungsspektrum der markierungsfreien MST ermöglichen und dabei weiterhin die Quantifizierung von Interaktionen in einem nahezu nativen Zustand ermöglichen.

In einem ersten Ansatz wurde ein veränderter Emissionsfilter getestet, der die ungewollte Detektion von Fluoreszenzsignalen der Interaktionspartner reduzieren sollte. Tatsächlich konnte durch diesen Filter der Grad an Fluoreszenzüberschneidung gesenkt werden. Allerdings weisen viele chemische Verbindungen privilegierte Strukturen wie Indole auf, die auch Teil des Tryptophans sind. Daher ist der Grad an Signalüberlappung stark von der Substanzbibliothek abhängig und die Vorhersagbarkeit über das Ausmaß an Signalüberschneidung bleibt schwierig, da kleinste Modifikationen der chemischen Struktur dieser Substanzen bereits drastisch deren Emissionsspektrum verändern können.

Da der neue Filter also keine generelle Lösung darstellte und auch nicht für die Analyse von PPI verwendet werden kann, wurde eine zweite Strategie entwickelt. Dabei wurde eine Mischungsgradienten-Titration (Englisch: composition gradient titration (CGT)) zusammen mit einer auf den kleinsten mittleren Quadraten basierten Datenanalyse für die Charakterisierung von PPI angewandt. Obwohl dieser Ansatz prinzipiell dafür geeignet war PPI zu quantifizieren, zeigten Simulationen einer Vielzahl an experimentellen Bedingungen, dass dieser Ansatz Limitationen in Hinsicht auf Fluoreszenzintensität, F_{norm} Wert und Startkonzentration der verwendeten Proteine hat.

Aufgrund der Nachteile beider Ansätze, wurde ein Kompromiss zwischen der Erhaltung der nativen Struktur des Proteins einerseits und den Vorteilen einer Fluoreszenzmarkierung andererseits gesucht. Um das zu erreichen wurde das Protein orts-spezifisch mit einem tris-NTA Farbstoff fluoreszenzmarkiert, was robuste Messungen von sowohl PPI als auch Protein-Ligand Interaktionen ermöglichte. Durch die orts-spezifische Markierung bleibt die native Struktur des Proteins erhalten und eine mögliche Beeinträchtigung der Ligandbindung durch Farbstoffmoleküle ist weitestgehend ausgeschlossen. Unter verschiedenen getesteten Fluorophoren erwies sich RED-tris-NTA als für diese Anwendung am besten geeignet. Darüber hinaus konnte dieser Ansatz für MST Messungen in Zelllysat verwendet werden, was zusätzlich den nahezu nativen Charakter dieses Ansatzes erhöht. Dies wäre mit markierungsfreier MST nicht möglich gewesen und stellt damit einen weiteren Vorteil dar.

Table of Contents

VIII

X

List of Abbreviations

°C	degree Celsius
µg	microgram
µL	microliter
µm	micrometer
µM	micromolar
A	adenine
ATAD-3	ATPase family AAA domain-containing protein 3
ATP	adenosine triphosphate
BLIP	beta-lactamase inhibitor protein
BSA	bovine serum albumin
C	cytosine
Ca II	carbonic anhydrase II
Cap	capillary
CGT	composition gradient titration
Cm	centimeter
CREB	cAMP response element-binding protein
Cy-5	cyanine 5
ddH$_2$O	double-distilled water
DLS	Differential Light Scattering
DMF	dimethylformamide
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DTT	1.4-Dithiothreitol
E.coli	Escherichia coli
EDIPA	ethyldiisopropylamine
EDTA	ethylenediaminetetraacetic acid
em	emission
ex	excitation
F$_{cold}$	fluorescence in the cold region
F$_{hot}$	fluorescence in the hot region
F$_{norm}$	normalized fluorescence
FP	Fluorescence Polarization
G	guanine
g	times gravity
GlcNAc	N-Acetylglucosamine
GPCR	G protein coupled receptor
H	hydrogen atom
HeLa	human epithelioid carcinoma
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His$_6$	6 histidine amino acids
His-tag	histidine tag
HPLC	high performance liquid chromatography
I	iodine atom

IDH1	isocitrate dehydrogenase 1
IgG	immunoglobulin G
IL6	interleukin 6
IR	infrared
ITC	isothermal calorimetry
K_a	equilibrium association constant
KCl	potassium chloride
K_d	equilibrium dissociation constant
kDa	kilodalton
KH_2PO_4	potassium phosphate monobasic
k_{off}	dissociation rate constant
k_{on}	association rate constant
LB medium	Luria-Bertani medium
LED	light emitting diode
PPI	protein-protein interaction
M	molar
MAPK	mitogen-activated protein kinase
MBP	maltose binding protein
MCAD	medium-chain acyl-CoA dehydrogenase
MEK1 (MAP2K1)	see MAPK
Mg	miligram
min	minute
mL	mililiter
mm	milimeter
mM	milimolar
MS-ESI	electrospray ionization mass spectrometry
MST	MicroScale Thermophoresis
n	number of experiments
Na_2HPO_4	disodium hydrogen phosphate
NaCl	sodium chloride
NHS	N-hydroxysuccinimidyl
NI	nickel
NiCl2	nickel chloride
nm	nanometer
nM	nanomolar
nmol	nanomolar
ns	nanoseconds
NT	NanoTemper
NTA	nitrilotriacetic acid
OD	optical sensity
P	phosphorus
PAH	phenylalanine-4-hydroxylase
PBS	phosphate buffered saline
PDB	protein data base
pH	potentia hydrogenii

XII

PHGDH	D-3-phosphoglycerate dehydrogenase
pM	picomolar
RNA	ribonucleic acid
S/N	signal to noise
SDS	sodium dodecyl sulfate
sec	second
SPR	Surface Plasmon Resonance
T	thymine
TANK	TRAF family member-associated NF-kappa-B activator
TCEP	Tris (2-Carboxyethyl) Phosphine
BRD	bromodomain
CREBBP	CREB binding protein
TFA	trifluoroacetic acid
TRIC	temperature related intensity change
Tris	tris(hydroxymethyl) aminomethane
Trp	tryptophan
Tyr	tyrosine
UV-Vis	ultra-violet visible
V_HH	variable domain of heavy chain only antibody
λ	lambda / wavelength

List of Figures

List of Tables

General Introduction

Every intra- and extracellular process depends on the interaction between different molecules, such as the binding of hormones to receptors, protein-protein interactions as part of a signaling cascade, protein-DNA interactions for gene regulation or antibody-antigen interactions as part of the immune response. Thereby these molecular interactions are highly specific, mostly reversible and fulfill strictly defined biological functions. Miscommunication in any of these processes can lead to a diseased state. Thus, robust and reliable determination of the affinity between a target molecule and its interaction partner is a critical step in many areas of biological, biochemical and biomedical research and technology[1]. Here, a deep understanding of the physical forces governing molecular recognition, affinity and specificity, is a prerequisite for the development of new and effective drugs[2]. As the drug discovery process is a time consuming (10 or more years from a hit to the drug on the market) and highly expensive (currently around 1 billion €) process, it is of great importance that biophysical methods provide reliable and robust results already during the early stages of drug development[3]. Nowadays there are various methods available for the investigation of intermolecular interactions and for the determination of underlying binding affinities (Figure 1)[4,5]. All methods can be categorized into either label-based, surface-immobilization based, or label-free and in-solution techniques[4,5]. Additionally, all available methods have their advantages and disadvantages regarding sensitivity, sample-, and time-consumption and provided information content. Thus, it is in general recommended to use more than one technique to verify the data outcome from a single analytical binding assay[4].
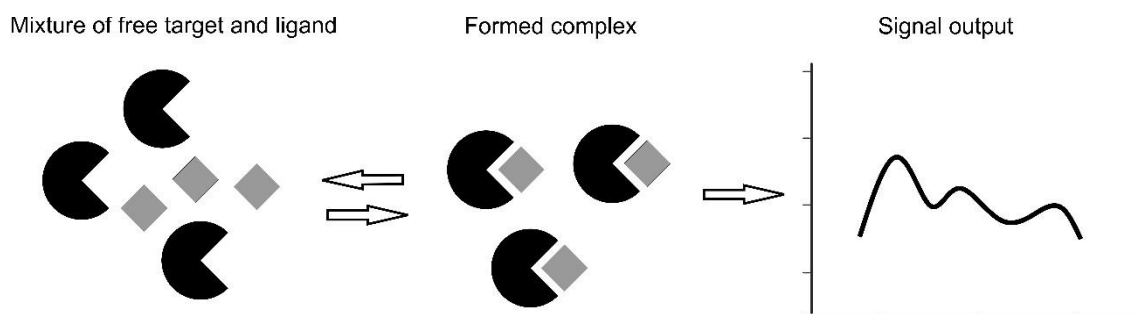
Figure 1: Schematic principle of ligand-binding assays. Free target and ligand molecules are present. Molecules start to bind to each other and form complexes. This binding event is detected and transduced into a signal output that can further be quantified.

A comprehensive listing of all available assays is beyond the scope of this thesis. However, the most important techniques for the quantitative characterization of intermolecular interactions are briefly summarized below and categorized according to their basic measurement principles.

1   Labeling-based methods

For a label-based approach for the quantification of biomolecular interactions, different types of labels and methods are available. The earliest label-based experiments were carried out using radiolabels for radioimmunoassays[6]. Such radio-isotopic labels, like ^{3}H, ^{125}I and ^{32}P, exhibit several disadvantages like radioactive waste and its disposal, high costs, the requirement for special licenses, and general significant health and environmental hazards[7]. This led to the development of other technologies, which are not harmful but still provide high sensitivity, such as fluorescence-based methods[8]. For such fluorescence-based assay, the attachment of a specific fluorophore to the molecule of interest is required. The behavior of the labeled molecule and the changes observed upon ligand binding will then be recorded. These changes can be detected e.g., as a quenching of the fluorescence intensity upon binding, as a change in the movement of molecules in a temperature gradient, or as changes in fluorescence anisotropy[9–12] (Table 1). However, such assays typically require time-consuming fluorescent labeling steps that in addition can interfere with the native conformation of the target molecule and might alter binding energetics[13]. The fact, that most molecules are labeled with more than one fluorophore per molecule to increase sensitivity, can further contribute to destabilization or alteration of the labeled molecules. To exclude any unspecific binding and negative influence of the labeling approach, carefully designed control experiments are required. In addition, fluorescence-based assays have to deal with potential fluorescence-interference of compounds, for which alternative methods have to be applied[14].

Table 1: Fluorescence-based methods.

Method	Strengths	Limitations	Sample consumption	Affinity range
MST* **MicroScale Thermophoresis**	Fast, low sample consumption, additional information on protein aggregation/precipitation, no instrument maintenance required, measurements in cell lysate and plasma[15,16].	Fluorescence label required	10 µL at nM concentration per data point	pM to mM
TSA** **Thermal shift assay**	Fast[4]	Fluorophore binding involved, quenching or aggregation can cause experimental artifacts[4], not applicable for hydrophobic proteins[4]	40 µL of 2 µM[4]	1 nM to 100 µM[4]
FP/FA*** **Fluorescence polarization/ anisotropy**	Highly reproducible[17], high throughput/fast[5], low cost[18]	Fluorescence label required, requires large change in size upon binding[5], autofluorescence/quenching/light scattering interference[18]	Several µL at nM concentration per data point[5]	nM to mM[5]
Protein-observed NMR**** **Nuclear magnetic resonance**	High structural resolution (identification of binding epitopes)[19]	Protein labeling required, high sample consumption[4], complex data analysis[5], low throughput, not possible for large proteins[20]	Several mg per data point[5]	100 nM to mM[4]

*Quantifies the difference in thermophoresis and TRIC of the unbound and the bound state of a molecule.
**Quantifies the shift in thermal unfolding of a protein upon ligand binding[11].
***Measures the polarization or anisotropy of light caused by changes in molecular size[18].
**** Follows $^1H/^{15}N/^{13}C$ chemical shifts of specific residues in the protein[21].

2 Surface-immobilization based techniques

Another class of biophysical methods for the investigation of binding affinities are surface-immobilization based techniques. These methods require the immobilization of one binding partner to a solid surface, which often is specifically functionalized for the immobilization procedure. The potential interaction partner will then bind to the immobilized molecule, whereby the binding event can either be detected due to changes in refractive index of light (surface plasmon resonance, back-scattering interferometry), changes in the vibration frequency (quartz crystal microbalance), or due to changes in the amplitude and phase of acoustic waves (surface acoustic wave)[22–25]. While surface-immobilization techniques are highly sensitive and have provided valuable knowledge in the study of molecular interactions, they do not reflect a physiologically relevant

environment as one binding partner is fixed to the surface and cannot diffuse free in solution. This can, on the one hand, alter the molecular structure or it can also sterically hinder the ligand from binding if the active site is in close proximity to the surface[26]. In addition, immobilization-based assays are often time-consuming and might require challenging surface immobilization strategies. However, once a capture molecule is immobilized to the solid surface, it can be used for the quantification of several binding events, as surface regeneration strategies can be applied[27]. The most important immobilization-based techniques, their strengths, limitations, sample consumption and affinity range are listed in Table 2.

Table 2: Immobilization-based methods.

Method	Strengths	Limitations	Sample consumption	Affinity range
SPR* **Surface plasmon resonance**	Kinetics and thermodynamics[5], high sensitivity[22]	Surface-immobilization, long experimental time, signals altered by solvent effects[4]	Sub-μg to μg[5]	Sub-nM to low mM[5]
QCM** **Quarz crystal microbalance**	Real-time determination of interactions with proteins and cells[4]	Surface-immobilization	15 nmol protein (0.5 mg of 30 kDa)[4]	1 nM to 500 μM[4]
BSI*** **Backscattering interferometry**	Simple and low-cost hardware and high sensitivity[28], kinetics[23]	Surface-immobilization	1μL of 5 μM protein[28]	pM to μM[23,28]

*Molecular interactions on the chip change the absorbed mass on the surface which directly lead to changes in the intensity of the reflected light. The intensity values are then converted into resonance signals by the optical detection unit[22].
**Monitoring vibration frequency of a quartz crystal upon interaction between a ligand and an immobilized protein or cell[29].
***Detects changes in refraction pattern of light, which passes through the sample in the channel of a microchip[23].

3 Label-free in-solution techniques

The techniques described above require labeling (attachment of an isotope or fluorophore) or surface immobilization of one of the binding partners, which can lead to misleading results, as the incorporated dye or the solid surface can impair protein function. This can occur due to destabilization of the proteins upon the immobilization or labeling procedure, or due to a close proximity of the modifications to the active site, whereby the latter might sterically hinder the ligand from binding[13,26]. Hence, label-free in-solution determination of binding strength is more favorable because the risk of perturbing the system is significantly reduced. Some methods exist which use inherent properties of the molecules for binding detection, without the need for any label or surface immobilization. Examples are isothermal titration calorimetry, analytical ultracentrifugation, dynamic light scattering and label-free MicroScale Thermophoresis.

These techniques provide different advantages and disadvantageous regarding costs, efficiency and the type of information they provide (Table 3)[4,5]. Generally, label-free detection systems are not as sensitive as label-based methods and thus often require a higher quantity of sample, but they provide physiologically relevant results.

Table 3: Label-free and in-solution based methods.

Method	Strengths	Limitations	Sample consumption	Affinity range
ITC* Isothermal titration calorimetry	Label-free and in-solution, Thermodynamic parameters, high precision and reproducibility[17]	Low throughput[5], long preparation time[17], buffer limitations[5], high sample consumption[4]	Several hundred µg per binding assay[5]	nM to sub-mM[5]
AUC** Analytical ultra-centrifugation	Label-free and in-solution, size-independent[5]	Sedimentation equilibrium needs to be reached[5]	Several hundred µL at nM to µM concentration per data point[5]	nM to mM[5]
DLS*** Dynamic light scattering	Label-free and in-solution	Requires defined difference in hydrodynamic radius of free and bound molecules[5]	Several µL at pM concentration per data point[5]	pM to mM[5]
Label-free MST****	Label-free and in-solution, fast, additional information on protein aggregation/precipitation, no instrument maintenance required	Requires strong intrinsic fluorescence, problems with autofluorescence, only one binding partner can fluoresce	10 µL of nM concentration per data point	nM to mM

* Measures the heat which is released or absorbed during combining of two substances by titration[30]
**Detects the separation of unbound and bound molecules using centrifugal force[5].
***Processes the time-dependent fluctuations in scattered light to yield the hydrodynamic radius of particles in solution, which will change upon binding[31].
****Quantifies the difference in thermophoresis and TRIC of the unbound and the bound state of a molecule.

4   Theoretical background

The following chapter provides theoretical background information on MicroScale Thermophoresis (MST) for K_d determination. As this method is based on fluorescence, the physical principles behind fluorescence will be described as well.

4.1   Basic principles of fluorescence

Fluorescence, whether those of fluorophores or intrinsic fluorescence of tryptophan (Trp) and tyrosine (Tyr), is the basis for MST investigations. For this reason, the phenomenon and the underlying physical principles of fluorescence will be described first.

Fluorescence occurs when an electron is excited by light with a distinct wavelength and then falls back to its ground state. More in detail, light excitation will cause a transition of an electron from its ground state to a higher energy electronic state, while the energy difference between the ground and the excited state reflects the required excitation energy[32]. In this excited state, the system will only remain for a few nanoseconds before the excited electron loses some of its energy due to heat or vibrational rotation[32]. The electron can then fall back to its ground state while emitting energy in form of a photon, a phenomenon called fluorescence[32]. Thereby, the fluorescence emission wavelength is always longer than the excitation wavelength, since it contains less energy[33]. The difference in excitation and emission wavelength is called Stokes shift and is of importance for all fluorescence-based assays, as it enables the excitation of fluorescence at one wavelength and its emission at another. Thereby, a large Stokes shift minimizes potential cross-talk between the excitation and emission of fluorescence[32,34].

Figure 2 shows a simplified Jablonski diagram, in which the process of fluorescence is illustrated.
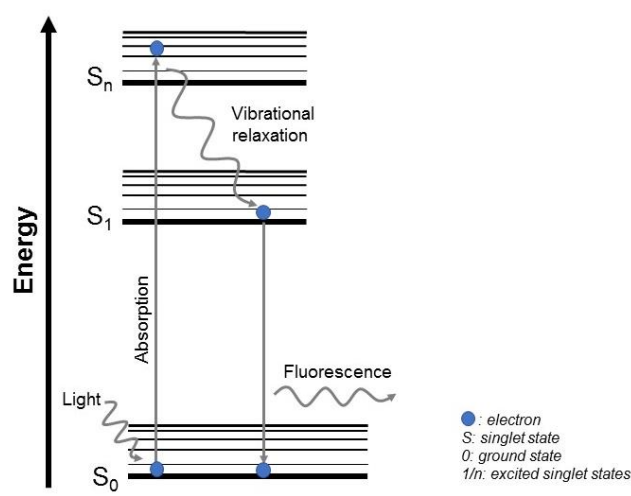
Figure 2: Jablonski diagram. Jablonski diagram for schematic illustration of fluorescence. Electronic states of a molecule are present as horizontal lines, while thicker lines being electronic energy levels and thinner lines are vibrational energy states. Absorption of photons from UV-light leads to a transition of the electron (blue) from the ground state to an excited state (here S_n). Loss of energy due to vibrational relaxation leads to an electron transition towards a lower energetic singlet state. Transition to its ground state leads to the emission of fluorescence. Figure was modified from Pingoud et al.[35] and Sasaki et al.[36].

4.2 Binding affinity of biomolecular interactions

For the comprehensive understanding of the molecular mechanisms behind biomolecular interactions, different parameters can be quantified. Besides binding kinetics and thermodynamic parameters, binding affinity can be quantified by measuring the equilibrium dissociation constant (K_d), or the equilibrium association constant (K_a), respectively. Binding affinity is influenced by non-covalent intermolecular interactions such as hydrogen bonding, electrostatic interactions, hydrophobic and Van der Waals forces between involved molecules[17]. Thereby, the binding process itself is regarded as an equilibrium condition, which results from a balance between association and dissociation of two interacting molecules[17]. The equilibrium state is dynamic, as molecules constantly form a complex while in the same time complex dissociates again[37].

In case of an interaction, two mixed proteins A and B will at some point colloid and bind together to form a complex. Most biological binding reactions can be described by a 1:1 interaction scheme. For such an interaction, the time-dependent association and dissociation can be expressed as shown in equation 1[17].

$$A + B \underset{k_{off}}{\overset{k_{on}}{\rightleftarrows}} AB$$

(1)

AB represents the protein–ligand complex and k_{on} [M^{-1}min^{-1}, number of association events per minute and molar] and k_{off} [min^{-1}, number of dissociation event per minute] are the kinetic rate constants[17].

At equilibrium, the association reaction of the monomers towards the complex and the dissociation reaction of the complex are equal, as given by equation 2[17],

$$k_{on}[A][B] = k_{off}[AB]$$

(2)

where the square brackets represent the equilibrium concentration of the single molecules ([A], [B]) or the complex ([AB]), respectively. The binding constants, K_d and K_a, are then defined by the law of mass action and can be expressed as[17]

$$K_d = \frac{[A][B]}{[AB]} = \frac{k_{off}}{k_{on}}$$

$$K_a = \frac{[AB]}{[A][B]} = \frac{k_{on}}{k_{off}}.$$

(3)

The relation between K_a and K_d can be expressed as[17]

$$K_a = \frac{1}{K_d}.$$

(4)

The equations above form the basis for determining their values experimentally. Thereby the equilibrium constant characterizes the affinity of molecules for each other, by calculating how much ligand is bound by the protein at equilibrium. An interaction with a fast association and a slow dissociation rate will result in a high binding affinity and in a low dissociation constant. Thus, the lower the dissociation constant the stronger the binding. In biological systems "tight binding" corresponds to a dissociation constant in the order of 1 nM or less[38]. For example, therapeutics should bind to their targets with high affinity as this will not only decreases costs but also the risk of potential side effects as lower doses can be applied[39]. The different dimensions of K_d-values are illustrated in Figure 3.
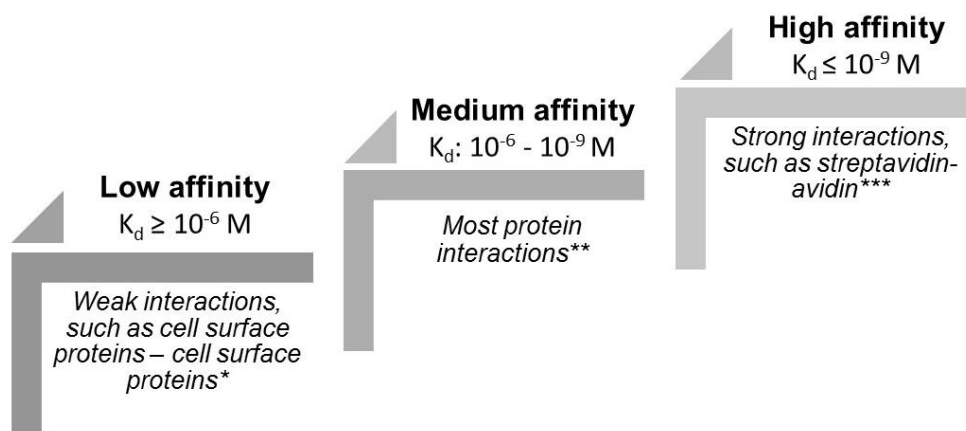
Figure 3: K_d dimensions. Overview of K_d dimensions, ranging from low, medium to high affinity[38]. *[40]; **[41]; ***[42]

4.3   MicroScale Thermophoresis

Thermophoresis was first described in the 19th century by Carl Ludwig and Charles Soret. Here, thermophoresis was described as a directed movement of particles along a temperature gradient, typically from a hot to a cold region[43]. This leads to a change in concentration, which can be quantified by the Soret coefficient S_T (the percentage of the concentration change per kelvin)[44].

$$S_T: \frac{c_{hot}}{c_{cold}} = \exp(-S_T \Delta T)$$

(5)

Several decades later, thermophoresis was further investigated by Philipp Baaske and Stefan Duhr, who used this thermophoretic effect for the quantification of molecular interactions on a micro scale[45,46]. Since the foundation of NanoTemper Technologies GmbH in 2008, MST developed into a well established biophysical tool for the rapid, sensitive and immobilization-free quantification of biomolecular interactions, ranging

from oligonucleotide interactions, like protein-DNA or RNA interactions, to protein-protein, protein-small molecule and protein-lipid interactions free in-solution[47–50]. Thereby, both the binding of ions, as well as the interaction of high molecular weight molecules and multidomain complexes can be quantified[51–53]. In addition, MST experiments can be carried out in any buffer, even in plasma and cell lysate, which allows for the evaluation of interactions under close to native conditions[15,16]. Another advantage of MST is the low sample consumption and a large dynamic range of affinities that can be quantified (pM to mM).

4.3.1 Theoretical background

Theoretical background information was obtained from the User Starting Guide for the Monolith NT.115_V23 if no other reference information is provided.

MST is used for the affinity quantification of molecular interactions. It is carried out in a Monolith instrument (NanoTemper Technologies GmbH), which consists of a fluorescence excitation and detection unit and an IR laser beam with 1480 nm wavelength (Figure 4 A)[1]. For the experiments, a serial dilution of the ligand is prepared, and a constant amount of a fluorescent target molecule is added to all dilution steps. Samples are then filled into glass capillaries with a maximal volume of 10 µL, placed on a tray and loaded into the Monolith device. At the beginning of the experiment, the fluorescence intensity of each capillary is recorded (capillary scan), to ensure sample homogeneity and accurate pipetting. Next, the IR laser is used to locally increase the temperature of the sample within a region spanning ~200 µm in diameter by 2 – 6 °C, depending on the MST power used (Table 4)[1]. This increase in temperature leads to a decrease of the observed fluorescence in the observation window, which comes from a change in fluorescence due to TRIC (temperature related intensity change) during the first seconds after laser activation and from a change in the concentration of the fluorescent molecule due to thermophoresis. Both parameters together, the thermophoretic movement and the TRIC are analyzed for MST quantification. Thus, the overall change in fluorescence $\frac{\partial}{\partial T}(cF)$ after IR laser activation can be expressed as:

$$\frac{\partial}{\partial T}(cF) = F\frac{\partial c}{\partial T} + c\frac{\partial F}{\partial T}$$

$F\frac{\partial c}{\partial T}$ *change in concentration (thermophoresis)*

$c\frac{\partial F}{\partial T}$ *change in fluorescence (TRIC)*

(6)

As the temperature sensitivity of the dyes, which are used for MST analysis, is highly sensitive towards changes in their local environment, ligand binding can be detected due to changes in the target conformation or in the intramolecular dynamics or because of a close proximity of the binding site and the dye. But not only TRIC, but also thermophoresis will change upon binding, as it depends on the size, charge and hydration shell[44]. Since at least one of these parameters will change upon binding, it is possible to quantify the thermophoretic change from an unbound to a bound molecular state.

4.3.2   Calculation of normalized fluorescence and binding affinities

MST is quantified by dividing the fluorescence intensity in the hot region by the fluorescence intensity in the cold region, which refers to the time after a defined MST-on time (Table 4) and the time before laser-on time, respectively. This value is normalized (F_{norm}), converted into promille and plotted against the logarithmic ligand concentration, which then yields a dose-response curve, from which the affinity constant can be obtained, by fitting the following equation

$$f(c) = Unbound + (Bound - Unbound) * \frac{c + c_{target} + K_d - \sqrt{(c + c_{target} + K_d)^2 - 4\,c\,c_{target}}}{2\,c_{target}} \quad , \qquad (7)$$

where f(c) is the fraction bound at a given ligand concentration c, unbound is the F_{norm} signal of the target, bound is the F_{norm} signal of the complex, K_d is the dissociation constant, and c_{target} is the final target concentration in the assay.

In case the fluorescence of the target molecule changes upon ligand binding without IR laser activation, this ligand-induced fluorescence change can be quantified as well. For this, it needs to be determined, if the fluorescence change is either a result of unspecific material loss or is ligand-binding specific. This can be done using the SD test, in which the samples are denatured using 4 % SDS and 40 mM DTT and heat, to then compare the fluorescence intensity of the samples prior and after denaturation. If the fluorescence counts are equal, the change in fluorescence was caused by the ligand-binding and thus the fluorescence scan can be evaluated.

Figure 4 represents some general principles of the MST measurement. Thereby, the optical system and the sample containing glass capillaries (A), a standard MST trace together with illustrations of the processes inside the capillaries during the MST measurement (B), MST traces of a 16-step MST binding experiment, as well as the resulting dose-response curve are shown (D).
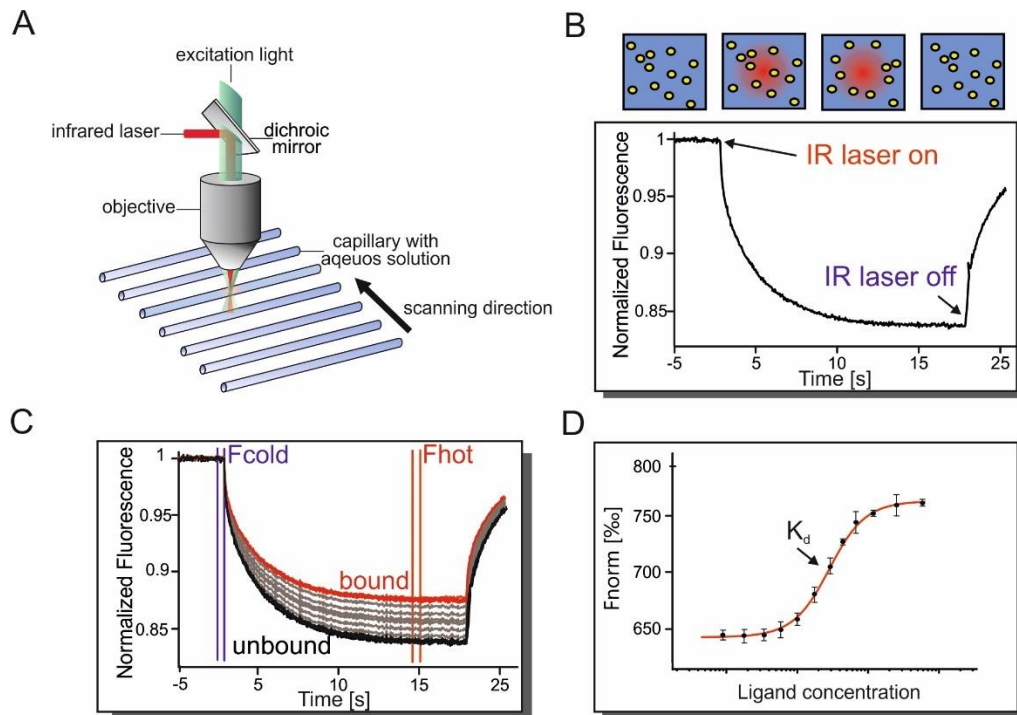
Figure 4: General MST measurement. A) Schematic illustration of MST optical system. Samples are loaded into glass capillaries and sample fluorescence is excited and recorded with the same objective. An implemented IR laser is used to heat a defined spot, leading to a temperature gradient in the capillaries. B) Illustration of an MST trace. At the beginning of the experiment, molecules are homogeneously distributed. After activation of the IR laser, molecules move out of the heated spot, leading to a steep decrease in detected fluorescence intensity due to the TRIC effect and thermophoresis. After around 10 sec, molecules reach a steady state, before they diffuse back ones the laser is switched off again. C) MST binding experiment. Normalized MST traces of unbound (black), intermediate state (grey) and bound molecule (red) are shown. Marked regions of F_{cold} and F_{hot} refer to fluorescence before (F_{cold}) and fluorescence after laser activation (F_{hot}). Cursor positions define fluorescence values that are used for data analysis. D) Dose-response curve. Normalized fluorescence counts are plotted against the logarithmic ligand concentration. F_{norm} refers to the normalized ratio of F_{hot} / F_{cold} *1000. The inflection point represents the K_d value. The figure was modified from Jerabek-Willemsen et al.[47].

The following table summarizes basic MST background information and measurement parameters.

Table 4: Overview of general MST measurement parameters.

Capillary type	Glass capillaries are used for sample loading. Standard capillaries are physically treated to obtain high-quality surface properties. Premium capillaries are covalently coated with a polymer to prevent surface adsorption of molecules.
LED power	Is used to excite the fluorophores. Intensity can be selected and should be high enough to detect sufficient fluorescence intensity.
Capillary Scan	At the beginning of the MST experiment, the fluorescence intensity of each capillary is detected. This provides already valuable information about sample adhesion to the capillary wall, inhomogeneous sample and incorrect pipetting.
MST power	Refers to the IR laser power and thus to the extent of the temperature gradient. Intensity can be selected, while the higher the MST power, the larger is the temperature increase.
MST trace	Defines the typical motion of a fluorescent molecule in a temperature gradient, that is recorded for 5 sec before laser-on time, 20 sec laser-on time and 5 sec after laser-on time.
MST-on time	Defines the time of IR laser activation, which is used for data quantification.
F_{norm} value	Describes the normalized fluorescence of MST traces.

4.3.3   Sample quality control using MST

MST data will not only provide binding affinity data, but will also deliver direct feedback on sample quality regarding aggregation or sticking of the fluorescent molecule to the surface of the capillaries. As aggregates can be observed by irregular MST traces, broader shapes of scanned capillary fluorescence are an indication for sample adsorption to the capillary wall. To improve bad sample quality, an addition of detergents or centrifugation of the samples can be used to get rid of these effects. Figure 5 illustrates both phenomena.
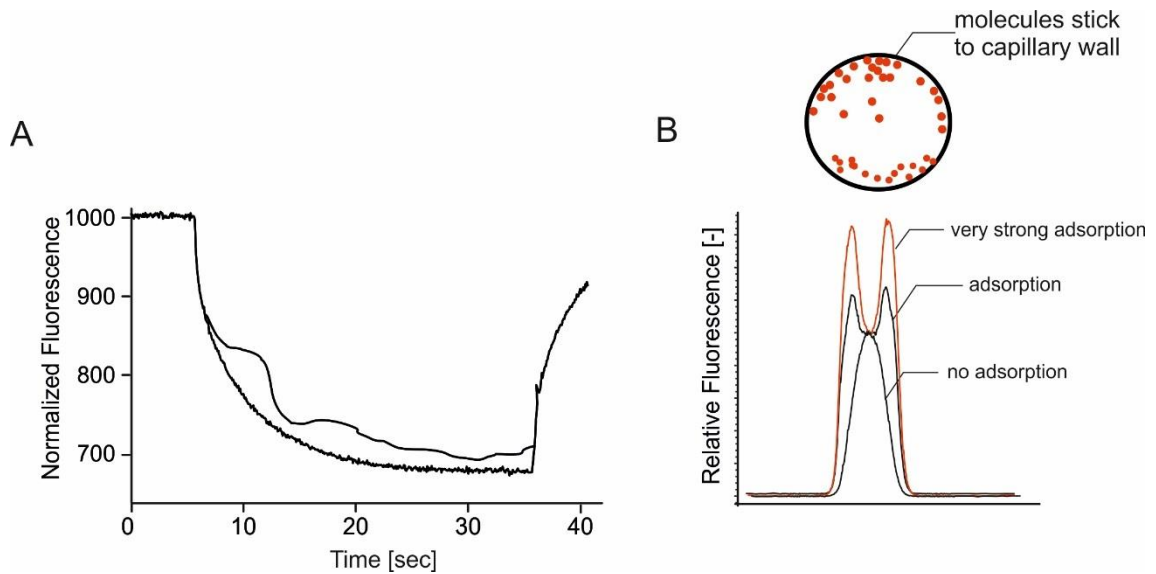
Figure 5: MST sample quality control. A) Schematic illustration of MST traces. The upper trace shows an irregular bulky shape because of the presence of aggregates. B) Different shapes of capillaries are overlaid: very strong adsorption of samples to the capillary wall (red), adsorption and no adsorption. Capillary cross-section is illustrated above and shows adsorption of molecules (red) to the capillary wall. Figure was modified from the MO.Control software.

4.3.4   Different MST instruments

NanoTemper Technologies GmbH provides Monolith instruments for the analysis of interactions in different spectral regions: blue (excitation 460 nm – 480 nm, emission 515 nm – 530 nm), green (excitation 515 nm – 525 nm, emission 560 nm – 585 nm) and red (excitation 605 nm – 645 nm, emission 680 nm – 685 nm). For the measurements, covalently attached fluorescent dyes or fluorescent fusion proteins are required. The Monolith NT.115^{Pico} detects low concentrations of red-emitting fluorophores, which allows analysis of high-affinity interactions with K_ds in the low pM range. The Monolith NT.LabelFree detects intrinsic fluorescence of proteins, which originates from Trp and Tyr residues. Thus, label-free MST provides a label-free in-solution analysis method.

5   Aims of the thesis

Since MST quantifies the transition from an unbound to a bound state of a protein, by recording changes in the proteins thermophoretic signal, it cannot be addressed to interactions, in which both binding partners fluoresce in the detection wavelength. Regarding an interaction, in which both binding partners fluorescence in the same spectral region, it is no longer possible to distinguish between an unbound and a bound state of a protein, as all recorded MST time traces would be a mixture of ligand, target

and complex time trace with unknown distributions. In addition, the serial dilution of the ligand might exceed the detection range, resulting in nondetectable low ligand concentrations and in detector saturation for high ligand concentrations, as illustrated in Figure 6.
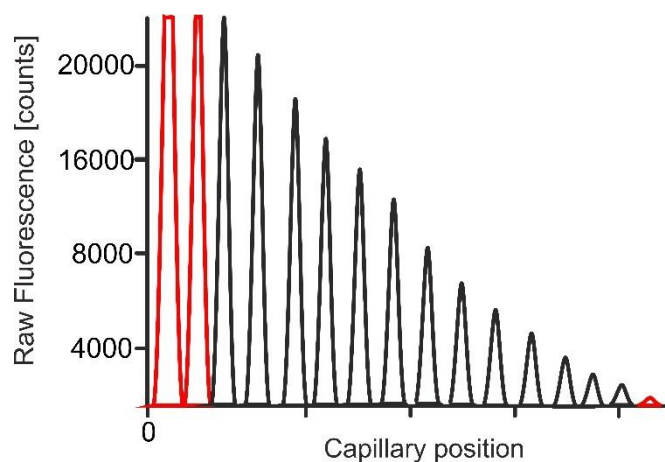
Figure 6: Capillary Scan. Schematic representation of a capillary scan of a 16-step standard serial dilution of two fluorescent molecules. Red peaks indicate the lowest ligand concentration (right) that cannot be detected by the optical system and the highest ligand concentrations (two left), which lead to detector saturation.

Fluorescence interference of compound is of special concern when planning label-free MST experiments. Many small molecules are based on privileged structures with indole motif, which is also present in the Trp residues of proteins. In addition, this method is not accessible to the analysis of protein-protein interactions (PPIs), which is a large scientific field of interest, comprising antibody-antigen interactions, among others.

The main aim of this thesis was to establish new approaches that would expand the application range of label-free MST and near-native approaches for MST in general. This thesis is divided into three major aims, which are discussed in three separate chapters. Each chapter deals with a distinct strategy to overcome the limitations of label-free MST and to find options for MST affinity analysis under close to native physiological conditions without the limitations described above.

The first aim was to increase the application range of the Monolith NT.LabelFree device for fragment and small molecules screenings. The assumption was that with the careful selection of an emission filter one could cut-off unwanted fluorescence signal from compounds, which otherwise leads to fluorescence interference in the assay. In Chapter 1 detailed analysis of factors determining the extent of fluorescence interference is presented and the influence of the emission filter selection is discussed.

The second aim was to establish an experimental approach that would enable the analysis of PPIs with label-free MST (Chapter 2). This chapter deals with the applicability of a composition gradient titration, combined with a least-mean-square fitting algorithm, for K_d determination without eliminating parts of the detected signal. Using this approach, label-free MST could, for the first time, be used to quantify PPIs. However, this approach was limited to proteins that exhibit similar fluorescence intensities and comparable F_{norm} values. In addition, low affine interactions might require high protein concentrations, whereas ultra-low sensitivities are not easy to determine as the device´s sensitivity might not be high enough.

Because the first two aims have not improved the applicability of label-free MST, an additional approach was considered. The main aim of Chapter 3 was to investigate and develop an MST approach that would enable the quantification of intermolecular affinities under near-native conditions. The combination of a site-specific labeling targeting the oligohistidine tag of proteins and an MST optimized fluorophore was investigated as a tool to achieve near-native experimental conditions.
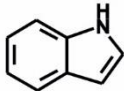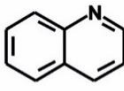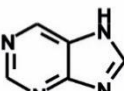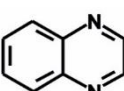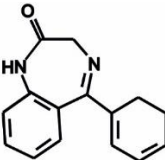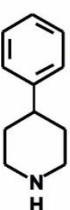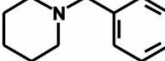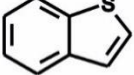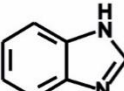
Chapter 1

1 Introduction

The main goal of the pharmaceutical research is the specific manipulation or inhibition of disease-related targets using small molecular compounds. During early stages of drug discovery often very large compound libraries (up to 4 billion compounds) are screened against a target molecule in a single-point binding experiment[54]. In case the target and the natural ligand are known, these libraries mostly contain small molecules with highly similar structures, which match the targets active site and the structure of the natural ligand[55]. However, if this detailed information is missing, a highly diverse library is more favorable, as it increases the chance for a successful screening campaign[55]. Around 30 years ago, Evans and his colleagues observed the potential of certain regularly occurring structural motifs as templates for defined modifications to generate novel and potent drugs[56,57]. These organic scaffolds are known as "privileged structures"[56]. Often, these scaffolds are derived from natural products and exhibit similar motifs than the peptide backbone[56,58,59]. In 2010, Welsch et al. published a broad list of privileged structures, such as indoles, quinolines, purines, and benzimidazoles, among others[58–60]. Table 5 lists different privileged structures together with their therapeutic effects as structurally modified drugs.

Among all privileged structures, the indole scaffold probably represents one of the most important structural subunits for the discovery of new drug candidates[56]. It is widely distributed in biological systems as it is an important constituent of the amino acid Trp, of alkaloids and of the neurotransmitter serotonin[56]. Furthermore, the indole scaffold is present in many drugs, such as GPCR agonists and antagonists, ion channel blockers and enzyme inhibitors[56].

Table 5: Overview of different privileged structures. Modified from Welsch et al.[59]

Name	Privileges structure	Drug class
Indole		Anti-cancer, serotonin-reuptake inhibitors, anti-depressives
Quinoline		Anti-malarial, immunosuppressants, anti-cancer
Purine		Anti-viral, anti-bacterial
Quinoxaline		Anti-glaucoma, anti-bacterial
Benzodiazepine		Anxiolytics, sedatives
Arylpiperidine		Anti-parkinsonian, anti-diabetic
Benzylpiperidine		Anti-parkinsonian, vasodilators
Benzothiophene		Anti-asthmatic
Dihydropyridines		Vasodilators
Benzimidazole		Anti-helmintic, anti-histaminic

The aromatic systems incorporated in the privileged structure can lead to the fluorescence interference in any label-free technique that uses the intrinsic Trp fluorescence to quantify intermolecular interactions. This can cause misinterpretation of the signal output, leading to potential false negatives and false positives[61–64]. Different strategies can be applied to overcome this problem. One way is simply to increase the concentration of fluorescently labeled target in an assay because the degree of fluorescence interference is directly related to the ratio between the concentration of the

compound and the concentration of the labeled target[61,63]. However, the concentration of labeled target used in many assays is in the nM range, whereas screened compounds are typically tested at µM – mM concentrations, making this approach not generally applicable[65]. Especially in MST experiments the target concentration used in the assay needs to be below the K_d, whereas the ligands are used in concentrations 20 to 50-fold above the K_d. With this, the free ligand concentration is similar to the total ligand concentration and ligand depletion is avoided[66]. This ensures maximum resolution and highest precision when determining dissociation constants[44].

Another described strategy for the avoidance of fluorescence interference results from the observation that lower wavelength dyes, like Fluorescein, Cyanine 3 or Rhodamine 110, are particularly prone for compound interference, since the percentage of fluorescent compounds is on average inversely proportional to the recorded emission wavelength[63,67]. Thus, many fluorescent compounds show emission below 530 nm[61,62,68,69]. Consequently, longer wavelength, red-shifted dyes, like Alexa 647 or Cyanine 5, are often preferred and can be successfully applied to reduce fluorescence interference[61–63]. Simeonov et al. analyzed a large chemical library with compound concentrations that are typically used in HTS for their fluorescence interference in different spectral regions[14]. About 5 % of the library was more fluorescent than 10 nM of the fluorophore 4-methylumbelliferone, but by red-shifting the emission wavelength, this number could be decreased[14]. Again, this solution strategy cannot be applied for label-free MST, as the fluorescence of Trp is recorded, which cannot be red-shifted. Another solution strategy would be to separate each recorded time trace into all underlying time traces, to then only analyze the transition from an unbound protein to its bound state, while excluding the time trace that arises from the ligand. A similar approach for signal separation can be found in literature, in which linear unmixing is used for fluorescence multichannel microscopy images, with which regions of overlapping fluorescence signals are reassigned to the different fluorophores used, regarding both their color and their intensity[70]. However, this is only possible as reference spectra of all underlying dyes can be recorded, which is not the case for label-free MST, for which the time trace of a pure complex is missing. Hence, another strategy to overcome fluorescence interference is required. The simplest possibility would be to separate the fluorescent signal which arises from the protein from the signal that arises from the compound upon detection. In principle, this is possible when protein and compound exhibit significant differences in any detectable output signal, ranging from fluorescence excitation, emission to fluorescence lifetime. Signal separation using different fluorescent lifetimes is already used in TR-FRET (time-resolved fluorescence resonance energy transfer) assays and in FCS (fluorescence correlation spectroscopy) measurements[71,72]. These techniques are

based on the observation that compound fluorescence often displays very short lifetimes, whereas that of the fluorophores such as Trp is longer[65]. Using such a time-resolved method can reduce interference by measuring the emission of the fluorophore right after the compound autofluorescence has vanished[65,73,74]. The main issue of this approach is that Trp itself can have very different fluorescent lifetimes. Thus, depending on its local environment, if it is in contact with other residues or if it is exposed to the solvent, its fluorescence lifetime can vary to a significant extent[75]. Thus the fluorescence lifetime of proteins can range from 0.5 to 5.5 ns[75,76]. Moreover, compounds can show diverse fluorescence lifetimes ranging from less than 0.2 to 1.4 ns and thus not always exhibit a shorter fluorescence lifetime compared to proteins[77]. To conclude, the variation of fluorescence lifetimes is too high to be used for detection in label-free MST.

In contrast, the emission of proteins in the UV-Vis spectral region is highly conserved, while depending on the compound library used, emission profiles of compounds can vary to a significant extent. Thus, the fluorescence emission could be used for a spectral separation of protein and compound.

The following will describe the efforts to separate the fluorescence signal that arises from the protein from that of the compound using appropriate emission filter. The spectrofluorometric profiling of several proteins and compounds was performed and the applicability of a new emission filter for label-free MST was tested. Obtained data showed that although the new filter can significantly reduce the fluorescence interference, the success rate highly depends on the chemical composition of the compound library used.

2   Materials

Table 6: Chemicals

Chemicals	Company
DMSO (≥ 99.8 %)	Carl Roth GmbH + Co. KG, Karlsruhe, DE, cat.no. A994.2
DTT	Carl Roth GmbH + Co. KG, Karlsruhe, DE, cat. no. 6908.2
GlcNAc	Sigma-Aldrich Co. LLC, St. Louis, US, cat. no. T2144
H_2O AnalaR NORMAPUR®	VWR Prolabo Chemicals, Darmstadt, DE, cat. no. 102927G
HEPES	Carl Roth GmbH + Co. KG, Karlsruhe, DE, cat. no. HN78.2
Maltose	Sigma-Aldrich Co. LLC, St. Louis, US, cat. no. M5885
MST buffer	NanoTemper Technologies GmbH
NaCl	Carl Roth GmbH + Co. KG, Karlsruhe, DE, cat. no. 0962.1
Pluronic® F-127	Sigma-Aldrich Co. LLC, St. Louis, US, cat. no. P2443
Roti®-Stock 10x PBS	Carl Roth GmbH + Co. KG, Karlsruhe, DE cat. no. 1058.1
SDS	Carl Roth GmbH + Co. KG, Karlsruhe, DE, cat. no. 4360.1
Tween®-20	Sigma-Aldrich Co. LLC, St. Louis, US, cat. no. 93773

Table 7: Buffers and solutions

Buffers and solutions	Composition
MST buffer	50 mM Tris-HCl, 250 mM NaCl, 10 mM $MgCl_2$, pH 7.8
MSTP buffer	0.1 % Pluronic® F-127 in MST buffer
MSTT buffer	0.05 % Tween® 20 in MST buffer
PBSP	0.1 % Pluronic® F-127 in PBS buffer
PBST	0.05 % Tween® 20 in PBS buffer
SD mixture	4 % SDS, 40 mM DTT

Table 8: Proteins

Proteins	Company
BI09 IgG	Boehringer Ingelheim, Ingelheim, DE
BLIP DBV1	CRELUX GmbH, Munich, DE
BRD4	CRELUX GmbH, Munich, DE

CREBBP	CRELUX GmbH, Munich, DE
His$_6$-MBP	Was kindly provided by Susanna v. Gronau and Dr. Sabine Suppmann from the biochemistry core facility of the Max Planck Institute of Biochemistry, Munich, DE
His$_6$-p38α MAPK	CRELUX GmbH, Munich, DE
His$_6$-Ca II	Sino Biological Inc., Beijing, CN, cat.no. 10478-H08E-50
MEK1	CRELUX GmbH, Munich, DE
Protein A	Thermo Fisher Scientific, Rockford, US, cat. no. 21181
SNF	CRELUX GmbH, Munich, DE
TEM DBV1 and TEM P107A	CRELUX GmbH, Munich, DE
α-Amylase	Sigma-Aldrich Co. LLC, St. Louis, US, cat. no. 10102814001

For the characterization of compound emission spectra, in total 1513 compounds were used. This set comprised compounds out of 7 different screening campaigns and 8 commercially available inhibitors. All compounds were diluted in 100 % DMSO upon arrival and were stored at -20 °C.

Table 9: Compound libraries

Target (number of compounds)	Supplier
(±)-Sulpiride against carbonic anhydrase	Sigma-Aldrich Co. LLC, St. Louis, US, cat. no. S8010
AGI-6780 against IDH2	Sigma-Aldrich Co. LLC, St. Louis, US, cat. no. SML0895
BIRB 796 against p38α	LC Laboratories®, Woburn, US, cat. no. D2744
Commercially available compounds (8):	
Compounds against MCAD (245)	LMU Medical Center, Dr. von Hauner Children's Hospital, Munich, DE
Fragments against Fyn (18)	CRELUX GmbH, Munich, DE
Fragments against MEK1 (193)	Sanofi Aventis, Paris, FR
Fragments against PAH (56)	LMU Medical Center, Dr. von Hauner Children's Hospital, Munich, DE
Furosemide against Ca II	Sigma-Aldrich Co. LLC, St. Louis, US, cat. no. F4381
I-CBP112 against CREBBP	Sigma-Aldrich Co. LLC, St. Louis, US, cat. no. SML1134
IDH-C227 against IDH1	Xcess Biosciences, Inc., San Diego, US, cat. no. M60043
IDH-C35 against IDH1	Sigma-Aldrich Co. LLC, St. Louis, US, cat. no. SML0839
PHGDH inhibitors (349)	Astra Zeneca, Cheshire, UK
SB203580 against p38α	Sigma-Aldrich Co. LLC, St. Louis, US, cat. no. S8307

small molecules against ATAD-3 (298)	Astra Zeneca, Cheshire, UK
TANK inhibitors (346)	Astra Zeneca, Cheshire, UK

Table 10: Devices

Device	Company
Calibration Thermostat (Ecoline ER 207)	LAUDA GmbH & CO. KG, Lauda-Königshofen, DE
FP-8300 Fluorescence Spectrometer	JASCO Germany GmbH, Gross-Umstadt, DE
NanoDrop™ One	Thermo Fisher Scientific Inc., Wilmington, US
NanoPhotometer	Implen GmbH, Munich, DE
NT.LabelFree (339 – 380 nm and 320 – 340 nm emission filter)	NanoTemper Technologies GmbH, Munich, DE

Table 11: Centrifuges and rotors

Centrifuges and rotors	Company
Centrifuge 5424	Eppendorf AG., Hamburg, DE
Centrifuge 5430 R	Eppendorf AG., Hamburg, DE
Rotor FA-45-24-11-HS	Eppendorf AG., Hamburg, DE

Table 12: Online tools and software

Online tools and software	Company
NT.Control 2.0.2.29	NanoTemper Technologies GmbH, Munich, DE
PyMOL(TM) 1.7.4.5 Edu - Educational Product	Schrödinger, LLC, New York, US
MO.Control_x86_1.5.3.6096	NanoTemper Technologies GmbH, Munich, DE
MO.AffinityAnalysis_x86_2.2.7.6056	NanoTemper Technologies GmbH, Munich, DE
Chemograph Plus 6.4	DigiLab Software GmbH, Altenholz, DE

Table 13: Further material

Material	Company
320 – 340 nm filter (AHF F47-330)	AHF analysetechnik AG, Tübingen, DE
330 – 380 nm filter (Semrock FF01-357 / 44-25)	Semrock Inc., New York, US
Glass capillaries with autofluorescence	NanoTemper Technologies GmbH, Munich, DE
Label-free emission filters:	
Monolith NT.LabelFree Zero Background MST Premium Coated Capillaries	NanoTemper Technologies GmbH, Munich, DE
Monolith NT.LabelFree Zero Background Standard Treated Capillaries	NanoTemper Technologies GmbH, Munich, DE

| SPROUT® MINI CENTRIFUGE 12V | Heathrow Scientific®, Vernon Hills, US |
| Ultra-micro-cuvette (quartz glass SUPRASIL®), 1.5x1.5 mm | Hellma GmbH & Co. KG, Müllheim, DE |

3   Methods

3.1   Fluorometric profiling of proteins and compound libraries

A three-dimensional fluorescence spectra of the protein p38α was recorded on a Jasco FP-8300 spectrofluorometer equipped with a Xenon lamp. The protein p38α was diluted in MSTP buffer to a final concentration of 168 µM. 20 µL of the solution was then filled into a quartz cuvette and loaded into the device. Fluorescence was excited using excitation wavelengths that range from 230 to 350 nm using a wavelength interval of 2 nm. The emission wavelengths were recorded between 260 and 600 nm. The excitation and emission bandwidths were both fixed at 5 nm, and the scan speed was 5000 nm per min. Measurements were performed at room temperature in the quartz cuvettes with a path length of 1.5 mm. Quartz cuvettes were cleaned between sample measurements using ddH_2O and dried with compressed air. Obtained spectral data were analyzed using Microsoft Excel 2016.

The same device was used to record two-dimensional fluorescence emission spectra of selected compounds (fragment against Fyn kinase, CBP112, fragment against MEK1, SB203580) and proteins (IgG BI09, BRD4, amylase, carbonic anhydrase, CREBBP, MBP, p38α). Therefore, samples were diluted in MSTP buffer to a final concentration of 1 µM for the proteins or of 100 µM or 132 µM for the compounds. 20 µL of the samples were then filled into quartz cuvettes and loaded into the device. A 280 nm excitation wavelength was used for the fluorescence measurements, while emission spectra were recorded between 290 nm and 750 nm using a wavelength interval of 0.2 nm. The excitation and emission bandwidths were both set to 5 nm, and the scan speed was 1000 nm per min. Measurements were done at room temperature in quartz 1.5 mm path length cuvettes. The quartz cuvettes were cleaned as described above. Obtained spectral data were analyzed using Microsoft Excel.

The PMT voltage was kept at 500 in all measurements; the intensity values of the overall spectra were in between 200 and 10000 fluorescence counts.

To estimate the proportion to which the given compound interferes with the fluorescence intensity readout, the data of two-dimensional compound emission spectra were used and all fluorescence intensity values were added to a sum, which lie between the filter bandpass borders determined. To enable better comparison between the compounds, their fluorescence intensity was first normalized by defining the maximum fluorescence intensity as 100 %.

3.2   Determination of fluorescence interference by compounds

The fluorescence intensities of in total 1513 compounds were determined with two different NT.LabelFree devices, one equipped with a 330 – 380 nm emission filter and one with a 320 – 340 nm emission filter, respectively. Data were recorded with the cap scan routine in the NT.Control 2.0.2.29 software. Depending on the stock concentration, the samples were diluted with MSTP buffer to reach a final concentration of 0.1 to 1 mM. More in detail, compounds were diluted in MSTP buffer to final concentrations of 1 mM (fragments against MEK1) or of 0.1 mM (PHGDH-, TANK-, and ATAD-3 inhibitory compounds). Compounds against PAH and MCAD were diluted to a final concentration of 500 µM using 20 mM HEPES, 200 mM NaCl, 0.1 % Pluronic, pH 7.0. All other compounds were diluted 1:200 in MSTP buffer to reach a final concentration of 0.1 to 0.75 mM. In addition, three different concentrations of MEK1 were prepared using MSTP buffer: 0.5 µM, 1 µM and 2 µM. Those served as reference samples and were used for later data analysis.

All samples were loaded into Monolith NT.LabelFree Zero Background MST Premium Coated Capillaries and loaded into the NT.LabelFree instrument. Probes were excited at 280 nm using an LED power of 15 %. MST experiments were recorded using 1 sec before IR-laser-on time, 1 sec IR-laser-on time and 1 sec IR-laser off time as a workaround to record the cap scan data. Measurements were carried out at 25 °C.

For data analysis, fluorescence intensities of the compounds were compared to the fluorescence counts of MEK1 (reference samples). Therefore, the fluorescence intensity of the three MEK1 measurements was set to 100 % and the compounds proportion to this fluorescence intensity was calculated. Obtained values were then distributed into two groups: Less than 30 % MEK1 fluorescence and more than 31 % MEK1 fluorescence.

3.3   Proteins fluorescence intensity in both emission filters

A set of different proteins (p38α, CREBBP, BRD4, IgG BI09, carbonic anhydrase, MBP, ProtA, SNF, BLIP, TEMP107A, MEK1, amylase) was diluted in MSTP buffer to a final concentration of 1 µM. Samples were then filled into Monolith NT.LabelFree Zero Background MST Premium Coated Capillaries and loaded into the Monolith NT.LabelFree device, equipped with the 330 – 380 nm or with the 320 – 340 nm emission filter, respectively. The probes were excited at 280 nm using 10 % LED power. MST experiments were recorded using 1 sec before IR-laser-on time, 1 sec IR-laser-on time

and 1 sec IR-laser off time to record the cap scan data. Measurements were carried out at 25 °C. Raw fluorescence counts were compared using Microsoft Excel.

3.4 Label-free MicroScale Thermophoresis Assay

For MST affinity analysis of MBP against maltose or $(GlcNAc)_3$, a 16-step serial dilution of maltose or $(GlcNAc)_3$ was prepared, starting from 1 mM highest concentration with a final volume of 10 µL in each titration step. MSTP buffer was used as assay buffer and was sterile filtered prior usage. Afterward 10 µL of 500 nM MBP was added to all vials and the reaction was incubated for 15 min at room temperature in the dark. Samples were then filled into Monolith NT.LabelFree Zero Background MST Premium Coated Capillaries and measurements were carried out at 25 °C, using 20 % LED and 80 % MST power. Data was evaluated after 5 sec laser-on time using the 330 – 380 nm emission filter, and after 10 sec laser-on time using the 320 – 340 nm emission filter.

For the binding affinity quantification of p38α against SB203580, a 16-step serial dilution of SB203580 was prepared using sterile filtered MSTP buffer (supplemented with 2 % DMSO). Thereby, the highest ligand concentration was set to 40 µM, with 10 µL in each dilution step. Afterward 10 µL of 220 nm p38α was added to all dilution steps and the reaction was incubated for 15 min at room temperature in the dark. Samples were then filled into Monolith NT.LabelFree Zero Background Standard Treated Capillaries and loaded into the two Monolith NT.LabelFree devices (equipped with 320 – 340 nm or with 330 – 380 nm emission filter). MST was recorded at 25 °C using 20 % LED power. After running the SD test (see below), data were analyzed using the raw fluorescence counts.

3.5 SDS denaturation (SD) test

If a ligand-dependent fluorescence change is detected during the cap scan, the SD test needs to be performed to identify the reason for this observation. This test allows discriminating between binding-specific fluorescence quenching and nonspecific loss of material. For this, the vials 1 to 3, which contain the highest ligand concentrations, and the vials 14 to 16, which contain the lowest ligand concentrations, are mixed 1:1 with the 2 x SD mixture (4 % SDS and 40 mM DTT in ddH_2O). Samples are then incubated for 5 min at 95 °C for protein denaturation. Afterward samples are filled into the same type of glass capillaries as the original samples were measured in. The SD test routine of the MO.Control software is then used to record the fluorescence intensities of each sample and to compare it to the fluorescence counts prior addition of the SD mixture. In case of specific ligand-induced fluorescence change, the fluorescence intensity of all six

samples should be nearly identical after denaturation. Consequently, the fluorescence intensity can be used for K_d determination. No changes in fluorescence intensity between the probes before and after the SD test are indicators of nonspecific loss of material. In this case, the experimental data is not valid; further optimization of assays conditions is required (like the use of low-binding tubes, a different labeling strategy or the addition of buffer additives).

As the interaction analysis of p38α against SB203580 showed ligand-induced fluorescence change in the device equipped with the 320 – 340 nm emission filter, the SD test was performed. For this, samples were prepared as described above, filled into Monolith NT.LabelFree Zero Background Standard Treated Capillaries and loaded into the device. The fluorescence was then recorded at 20 % LED power and 25 °C.

4   Results

With the intention to select the most appropriate emission filter for the reduction of fluorescence interference originating from the compounds used in the screening, the spectrofluorometric profiling of several proteins and compounds was performed and the applicability of new emission filter in the label-free MST was tested. Overall, 7 proteins and 7 different compound libraries were analyzed, which contained 1513 compounds in total. Moreover, label-free MST experiments were carried out using both emission filter sets. Obtained data were compared regarding fluorescence interference, S/N ratio, fluorescence intensity and determined dissociation constant.

4.1   Fluorescence profiling of proteins and compounds

Fluorescence profiling of proteins and compounds was performed to determine the spectral overlap. The regions which do not overlap can be used for detection by a selection of a filter with an appropriate bandwidth. To determine the typical fluorescence excitation and emission wavelengths of a protein, first a three-dimensional spectrum of p38α was recorded. For this, p38α was diluted in MSTP buffer and a three-dimensional protein spectrum and the corresponding contour map, showing the bird´s eye view of the spectrum, were recorded using spectrofluorometry (Figure 7). The protein spectra show two emission peaks, one at λ_{ex} = 280 nm and λ_{em} = 350 nm and the second peak at λ_{ex} = 230 nm and λ_{em} = 350 nm. The first peak refers to the Trp residues of the protein, while the second peak shows the peptide bond (absorption below 240 nm) [78].
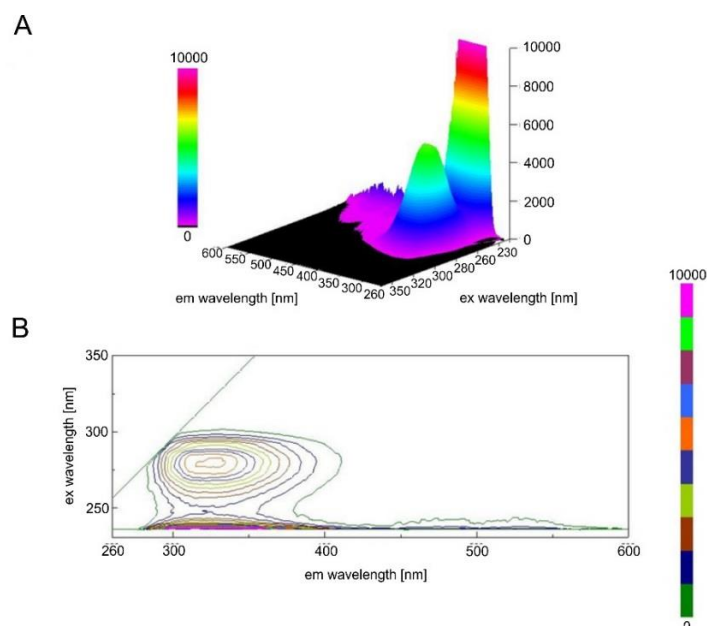
Figure 7: Fluorimetric profiling of p38α. Three-dimensional fluorescence spectrum (A) and corresponding contour diagram (B) were recorded on a JASCO spectrofluorometer, using 168 µM p38α diluted in MSTP buffer. Emission wavelengths were recorded between 260 nm and 600 nm using excitation wavelengths of 230 nm to 350 nm, with 10 nm steps in between.

The Monolith NT.LabelFree device excites the Trp fluorescence of the protein at 280 nm and detects it at 330 – 380 nm. As the three-dimensional spectra of p38α exhibits maximum emission at 330 nm after excitation at 280 nm, the protein can efficiently be detected using label-free MST. However, as the fluorescence properties of Trp are highly sensitive towards its local environment, Trp emission can vary among different proteins. Thus, Trp residues which are directed towards the hydrophobic core of a protein will exhibit emission maximum at 330 nm, while Trp amino acids that are directed towards the solvent will have a maximum emission at 350 nm. To gain insides into the variation of protein emission profiles, a set of representative proteins was chosen, and the fluorescence spectroscopic profile was recorded. For this, proteins were diluted in MSTP buffer and emission was recorded at an excitation of 280 nm, using spectrofluorometry (Figure 8). Obtained emission profiles were normalized to their maximum fluorescence intensity of the Trp residues. The normalized emission spectra show only slight variations in the maximum peak of emission, ranging from 320 to 350 nm. These slight deviations arise from the Trp sensitivity described above. Although there are few exceptions that show different emission spectra, such as an emission maximum of ~ 308 nm for azurin or ~ 355 nm for glucagon, the pharmaceutically most relevant proteins, like enzymes or antibodies show similar spectra as shown in Figure 8[79].
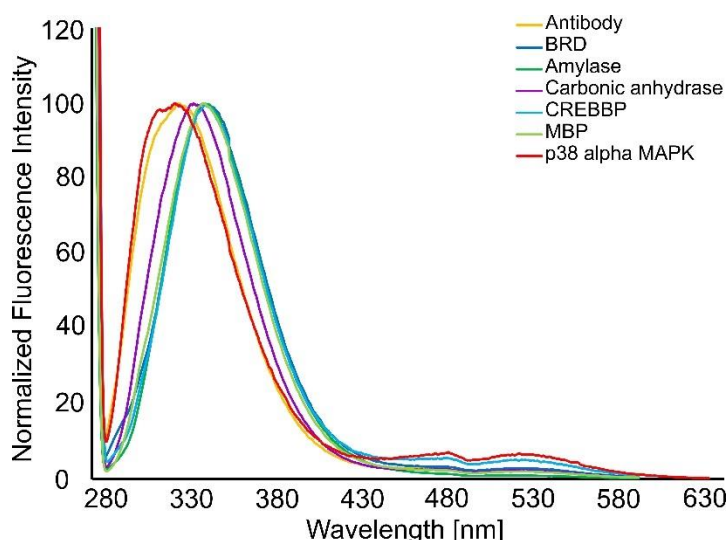
Figure 8: Emission spectra of different proteins. Spectra were recorded on a JASCO spectrofluorometer. Excitation wavelength was set to 280 nm, while the emission wavelength ranged from 270 nm to 600 nm.

Knowing that the emission maximum of proteins typically varies between 320 – 350 nm, a new emission filter can be designed covering this part of the spectrum. However, before the exact bandwidth of the filter could be selected, emission properties of a representative set of compounds were analyzed. To record the spectra, samples were diluted in MSTP buffer and spectrofluorometric analysis was carried out on a JASCO spectrofluorometer. Figure 9 shows the normalized emission spectra of ten compounds. As the structural motifs and scaffolds of those compounds are more diverse than the peptide backbone of proteins, their emission profiles vary to a significantly higher extent than those of proteins.
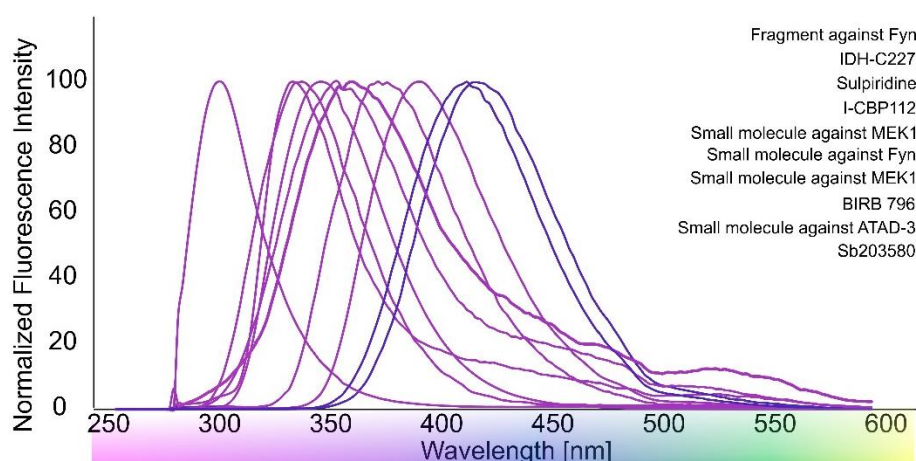
Figure 9: Emission spectra of compounds after excitation at 280 nm. Fluorescence spectra were recorded at a fixed excitation wavelength of 280 nm and an emission range from 280 nm to 600 nm using a JASCO spectrofluorometer. Names of compounds are listed on the right (from the top down refers to from left to right). Their emission maxima range from 310 nm to 425 nm.

New emission filter should thus enable the transmission of Trp fluorescence to guarantee strong and stable signal originating from the protein and at the same time cut-off the signal which arises from the compounds.

To fulfill mentioned requirements, the emission filter with the bandwidth of 320 – 340 nm was chosen, as most proteins emit in this spectral region (Figure 8). At the same time, the chosen bandwidth cuts-off most of the interfering fluorescence originating from compounds. As a comparison, the bandwidth of the current filter is 330 – 380 nm. The transmission spectra of both emission filters are illustrated in Figure 10.
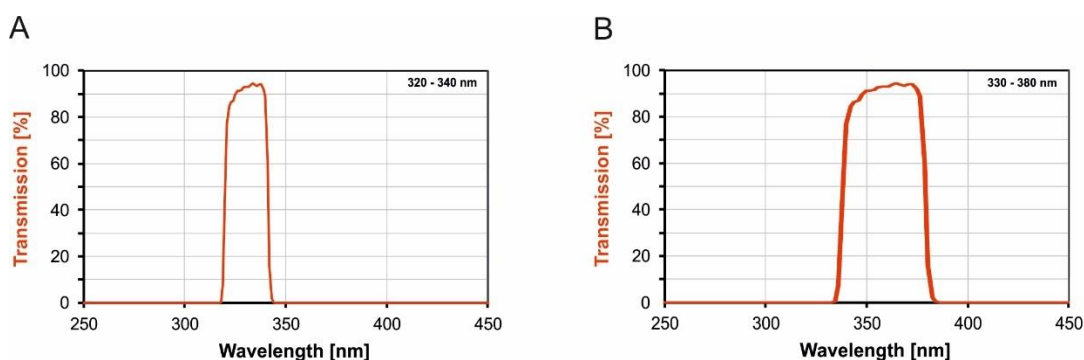
Figure 10: Two different filter sets for the NT.LabelFree device. A) The emission filter is implemented in NT.LabelFree devices for detection of Trp fluorescence. It spans a bandwidth from 330 – 380 nm. B) The emission filter was designed to reduce fluorescence interference from compounds in label-free MST assays. It spans a bandwidth from 320 – 340 nm.

While both filter sets can detect sufficient protein fluorescence, compounds interference is significantly higher for the broad emission filter (Figure 11). Here, all four compounds overlap the region of detection, while less compound interference can be observed for the filter with narrower bandwidth.
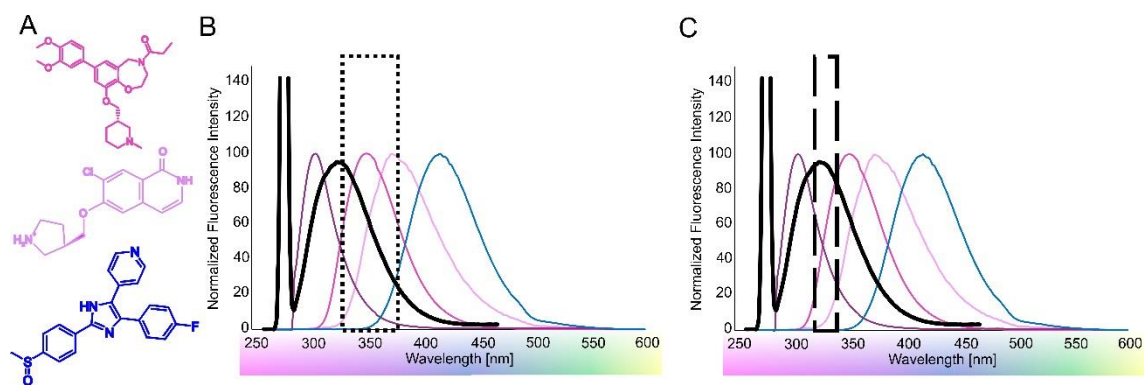
Figure 11: Emission spectra of compounds, IgG BI09 and position of emission filters. A) Chemical structures of CBP112, fragment against MEK1 and SB203580 (from top to bottom). Structure for fragment against Fyn was not available. Structures are illustrated in the same color code as corresponding emission profiles in B and C. If available, compounds chemical structures were re-drawn using Chemograph Plus 6.4 software. B and C) Fluorescence emission spectra of four compounds (purple to blue) and IgG BI09 (black) are shown. Spectra were recorded at a fixed excitation wavelength of 280 nm and an emission range from 280 to 600 nm using a JASCO spectrofluorometer. Spectral profiles are overlapped with a 330 – 380 nm (B) or with a 320 – 340 nm (C) emission filter, illustrated as black boxes.

To quantify the extent of fluorescence interference of each compound, their two-dimensional emission spectra were used. First, their spectra were normalized to their maximum emission. Then, all fluorescence intensity values that lie either between 320 – 340 nm or between 330 – 380 nm emission wavelength, were added to a sum. This way, a comparison of both filter sets was obtained (Table 14). Here, the relative fluorescence of each compound in both filter sets is listed separately. Thereby, only one compound (fragment against Fyn kinase) has higher fluorescence intensity with the filter 320 – 340 nm than with the filter 330 – 380 nm. Thus, all the other compounds exhibit significantly higher fluorescence in the 330 – 380 nm emission filter.

Table 14: Relative fluorescence intensity of each compound in both filter sets. The fluorescence interference of each compound in the emission filter was defined as the sum of all normalized intensity values in between 320 – 340 nm and 330 – 380 nm, respectively.

Compound	Sum of normalized fluorescence intensities	
	320 – 340 nm	*330 – 380 nm*
Fragment against Fyn	513.8	430.9
IDH-C227	770.9	1886.6
Sulpiridine	850.0	2048.4
I-CBP112	571.0	2163.5
Small molecule against MEK1	506.7	2191.0
Small molecule against Fyn	386.1	2084.4
Small molecule against MEK1	506.8	2190.1
BIRB 796	3.0	576.1
Small molecule against ATAD-3	0.6	173.9
SB203580	2.1	386.1

In detail, the fluorescence interference of nine compounds was lower in the narrower emission filter. In addition, three compounds showed no interference for the 320 – 340 nm emission filter. For the other compounds, up to 5-fold less interference was observed. Only the fragment against Fyn shows slightly stronger interference in the smaller emission filter set. The emission profile of this compound exhibits maximum close to 300 nm as seen in Figure 11.

4.2 Quantitative comparison of emission filters for the NT.LabelFree device

Based on the results obtained above, it was assumed that the narrower emission filter should significantly reduce the autofluorescence from compounds in label-free MST experiments. In order to verify this assumption, the fluorescence intensity of compounds and chemical fragments out of 7 compound libraries (in total 1513 compounds and fragments) was analyzed using the cap scan routine in the Monolith NT.LabelFree device equipped with the 330 – 380 nm emission filter and in a prototype label-free device equipped the 320 – 340 nm emission filter. Therefore, all compounds were diluted in MSTP buffer, filled into glass capillaries and separately loaded into both devices. The fluorescence intensity of each sample was then recorded using the cap scan routine of the device. As reference for data analysis the fluorescence intensity of two different MEK1 concentrations (0.5 µM and 2 µM) was recorded. For this, the fluorescence intensity of MEK1 was set to 100 %, while the fluorescence counts of each compound

were then calculated as their percental proportion to this. All compounds exhibiting less than 30 % MEK1 fluorescence were counted and are summarized in Table 15.

Table 15: Summary of fluorescence intensity screening of 1513 compounds. Seven compound libraries and eight commercially available substances were analyzed for their fluorescence intensity compared to the fluorescence counts of two different MEK1 concentrations, using two emission filters. Number of compounds, showing less than 30 % of MEK1 fluorescence are listed as absolute numbers and as percentage values.

Target	Number of compounds	320 – 340 emission filter		330 – 380 nm emission filter	
		0.5 µM MEK1	2 µM MEK1	0.5 µM MEK1	2 µM MEK1
PAH	56	14 (25.0 %)	34 (60.7 %)	4 (7.1 %)	24 (42.9 %)
ATAD-3	298	162 (54.4 %)	257 (86.2 %)	80 (26.8 %)	212 (71.1 %)
Fyn	18	2 (11.1 %)	10 (55.6 %)	2 (11.1 %)	7 (38.9 %)
MEK1	193	74 (38.3 %)	124 (64.2 %)	14 (7.3 %)	76 (39.4 %)
PHGDH	349	229 (65.6 %)	311 (89.1 %)	187 (53.6 %)	280 (80.2 %)
TANK	346	222 (64.2 %)	295 (85.3 %)	102 (29.5 %)	222 (64.2 %)
MCAD	245	5 (2.0 %)	14 (5.7 %)	15 (6.1 %)	109 (44.5 %)
diverse targets	8	3 (37.5 %)	6 (75.0 %)	1 (12.5 %)	4 (50.0 %)

As seen in Table 15, in seven out of eight libraries, the narrower filter reduced the number of interfering compounds to a significant extent. Especially for the screening against TANK, 35 % less compounds interfered with the detection, which refers to around 120 compounds for a screening of this size. Interestingly, compounds of the screening campaign against MCAD are probably more prone to interfere in the region of 320 – 340 nm, making the narrower emission filter less suitable for this screening project, compared to a broader filter bandwidth.

In general, data from Table 15 illustrate, that the number of interfering compounds could be lowered with higher concentrations of MEK1. Thus, depending on the K_d of the interaction and with this on the required compound concentrations, more or less fluorescence interference of compounds will be observed.

4.3 Determination of fluorescence interference by compounds

In order to investigate the influence of the narrower filter bandwidth on the fluorescence intensity emitted by proteins, a set of different proteins was diluted in MSTP buffer and loaded into the Monolith NT.LabelFree devices, equipped with either the 330 – 380 nm or with the 320 – 340 nm emission filter. Protein fluorescence was recorded after excitation at 10 % LED power. Results are illustrated in Figure 12. For all proteins, the narrower emission filter reduced the overall fluorescence counts to ~ 50 %.
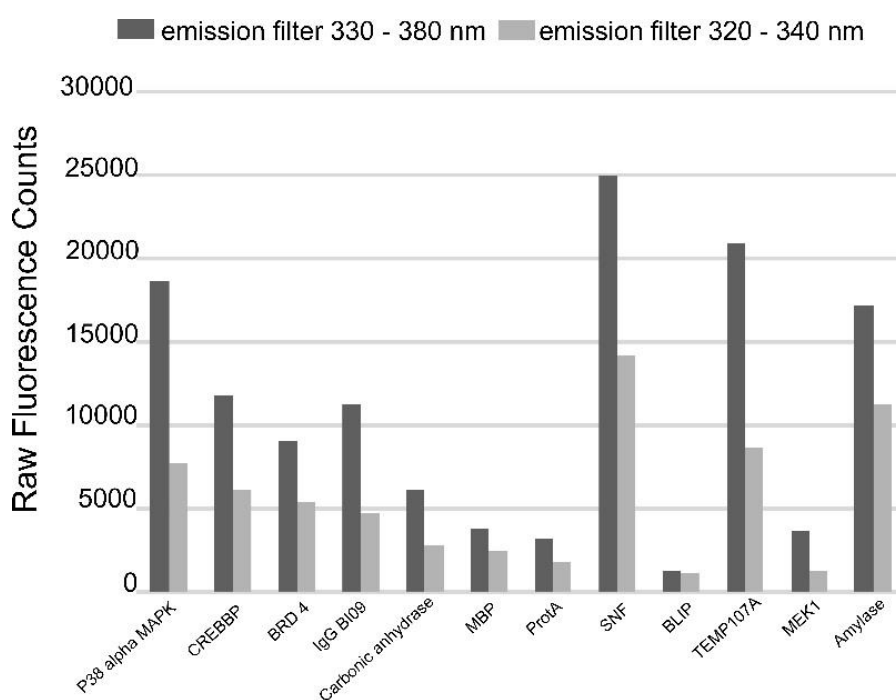
Figure 12: Fluorescence intensity of proteins in both emission filters. Proteins fluorescence was excited at 280 nm and at 10 % LED power using two Monolith NT.LabelFree devices (emission filter 330 – 380 nm (black), 320 – 340 nm (grey)). Fluorescence was recorded at 25 °C. Measurements were done in duplicates.

4.4 Influence of filter bandwidth on the quality of label-free MST measurements

As the narrower bandwidth of the emission filter reduces the sensitivity of the device to around 50 %, the impact on data quality of label-free MST measurements was investigated. Therefore, affinity analysis of two already MST approved interactions were carried out on both devices. Therefore, the interaction between MBP and maltose is known to show high binding amplitudes and high S/N ratios using standard label-free

MST, whereas the interaction of p38α against SB203580 was so far not accessible by label-free MST due to fluorescence interference caused by SB203580.

MBP is part of the periplasmic transport system of *E.coli*. Here, MBP binds the disaccharide and translocate it across the inner membrane [80]. Figure 13 shows the structure of MBP together with maltose.
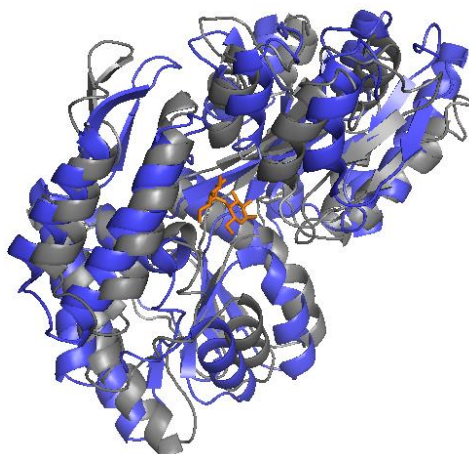
Figure 13: Crystal structure of MBP and maltose. Once MBP binds to maltose (colored orange), it undergoes a conformational change. Structural alignment of unbound structure (blue; PDB 1OMP) and bound conformation (grey; PDB 1ANBF) was performed using PyMol.

The re-purified MBP (Supplementary data Figure 2) was used for the label-free MST experiments. Briefly, a serial dilution of maltose was prepared, starting from 1 mM as the highest concentration. MSTP buffer served as assay buffer. Afterward 500 nM of MBP was added to all dilution steps. As a negative control, N,N',N''-triacetylchitotriose, also known as $(GlcNAc)_3$, was titrated against MBP using the same experimental conditions. Samples were measured in both NT.LabelFree devices, equipped either with the emission filter 330 – 380 nm (Figure 14 A), or with the emission filter 320 – 340 nm (Figure 14 B). Measurements with the 330 – 380 nm emission filter exhibit a higher S/N ratio and more homogeneous MST traces than those obtained by the narrower filter (Figure 14). However, obtained binding affinities of 9.7 ± 1.2 µM and 6.6 ± 1.0 µM for the broader and narrower emission filter, respectively, are in good agreement with literature values, where the K_d is around 2 µM[81].
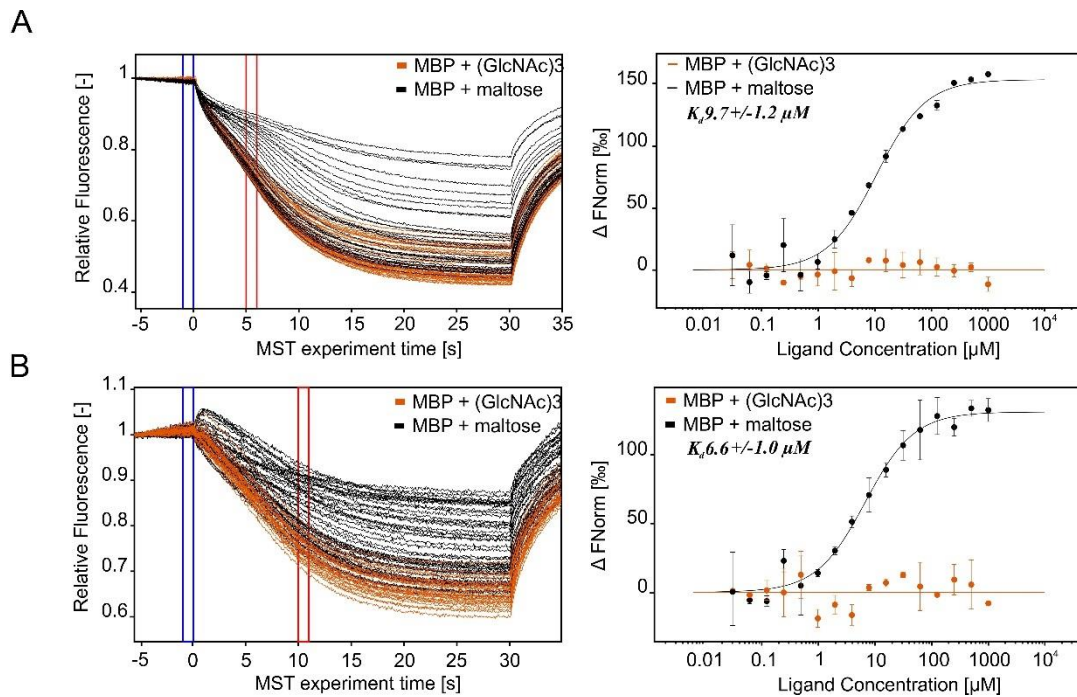
Figure 14: MST affinity analysis of MBP against maltose. A & B): Comparison of the MST traces (left) and the dose-response curves (right) of maltose (black) and $(GlcNAc)_3$ (orange) titrated against MBP using the NT.LabelFree device equipped with the 330 – 380 nm emission filter (A) or with the 320 – 340 nm emission filter (B). All graphs display merged data from three independent experiments. MST experiments were carried out at 20 %LED and 80 % MST power at 25 °C.

Next, the interaction between p38α and a commercially available inhibitor, SB20358, was analyzed. P38α belongs to the class of mitogen-activated protein kinases (MAPK) and plays a major role in the regulation of cell cycle processes and in the inflammatory response, what makes this kinase a well-known therapeutic target for many pathologies[82]. The mentioned interaction was previously not accessible by label-free MST (Figure 17 A).
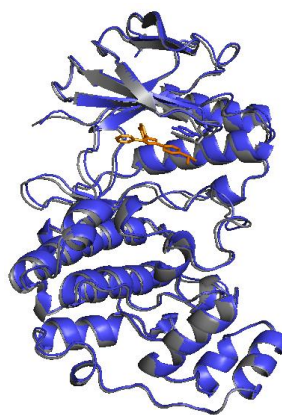
Figure 15: Crystal structure of p38α and SB203580. Structure of p38α alone (grey) and in a complex (blue) with the SB203580 inhibitor (orange) (PDB code: 1A9U).

Figure 16 shows the emission spectra of p38α (black) and the compound SB203580 (blue). The transmission width of both filters is depicted with black boxes. With the use of the broader 330 – 380 nm filter, the fluorescence interference of SB203580 negatively influences the MST measurements (Figure 16, red region). The use of a narrower filter should cut-off the fluorescence of SB203580 and thus enables determination of binding affinity in the label-free MST assay.
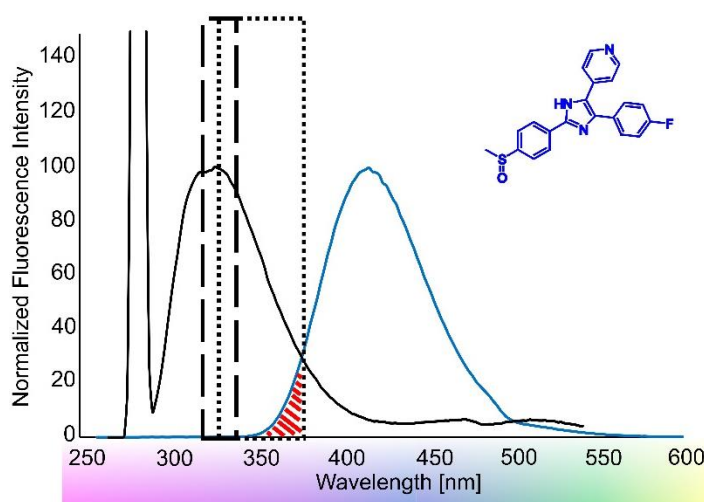
Figure 16: Emission profiles of p38α and SB203580. Emission spectra were recorded using a JASCO spectrofluorometer. The emission spectrum of p38α shows two emission peaks at an extinction of 280 nm. The first peak at 280 nm is residual excitation light, while the second is caused by aromatic amino acids, especially Trp (black). In blue, the emission spectrum of SB203580 inhibitory compound is shown, with its chemical structure illustrated on the right. The two black boxes show the emission filters that were used for label-free MST affinity analysis. Using the 320 – 340 nm emission filter, the fluorescence signal of the compound is eliminated, whereas there is fluorescent interference with the 330 – 380 nm filter. Region of interference is marked in red.

Figure 17 shows the MST results for the interaction described above. Here, with the broad emission filter (Figure 17 A) the fluorescence interference of SB203580 prevented determination of the binding affinity. However, the same interaction could be analyzed using the narrower emission filter (Figure 17 B). The SD test confirmed that the ligand-induced fluorescence change is binding-specific (Figure 18). Thus, this change in fluorescence was used for K_d quantification, which was in good agreement with literature values[83].
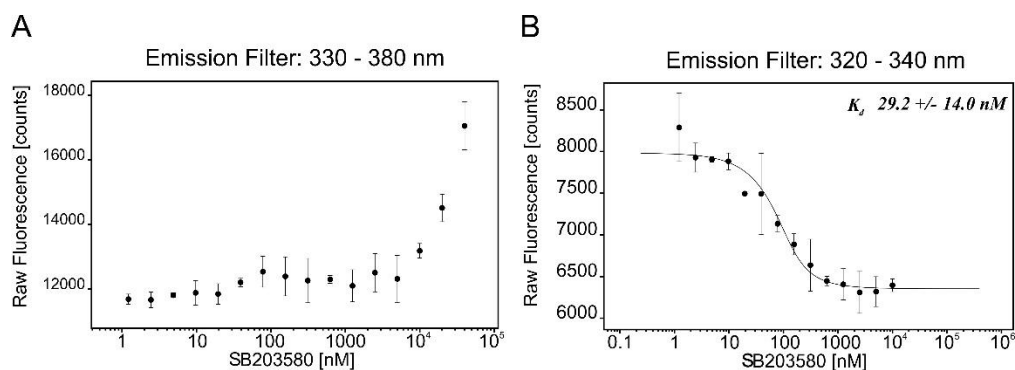
Figure 17: Binding of SB203580 to p38α. In both experiments, the concentration of p38α was kept constant at 110 nM, while the concentration of SB203580 was varied from 1.2 nM – 40 000 nM. Experiments were performed at 20 % LED power and at 25 °C (n = 2). A) Raw fluorescence counts of the measurements using the 330 – 380 nm emission filter. Quantification of the K_d was not possible. B) Raw fluorescence counts of Monolith device with a narrower emission filter (320 – 340 nm). Binding specific ligand-induced fluorescence change was used for data analysis.
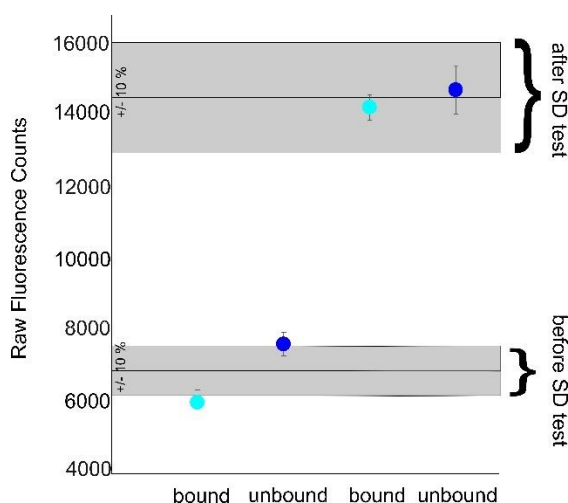
Figure 18: SD test for p38α against SB203580. Fluorescence intensity of samples was recorded prior and after denaturation using the cap scan routine in the Monolith NT.LabelFree device.

The observed reduction in fluorescence upon binding can be explained by the existence of a Tyr residue in the SB203580 binding site of p38α, as illustrated in Figure 19. As not only Trp, but also Tyr residues contribute to the signal in label-free MST measurements, the binding of ligands near those amino acid residues can lead to changes in the detected fluorescence intensity. This decrease in fluorescence cannot be detected using the broader emission filter, as here the fluorescence of SB203580 overlays any change in Trp or Tyr fluorescence. In addition, Tyr emission maximum lies at around 330 nm, which makes the 320 – 340 nm emission filter more prone to changes in the Tyr fluorescence[84].
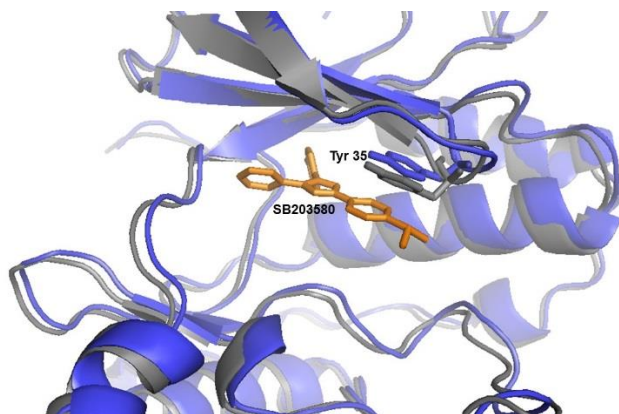
Figure 19: Crystal structure of p38α binding pocket and SB203580. Structure of the binding site of p38α alone (grey) and in a complex (blue) with the SB203580 inhibitor (orange) (PDB code: 1A9U). Tyr35 amino acids are illustrated as sticks.

5   Discussion

Nowadays, drug discovery mostly starts with the screening of thousand potential drug candidates against a disease-related target molecule in an HTS campaign. Thereby, label-free and in-solution methods are the techniques of choice, as they provide close to native experimental conditions. One example of such a technique is label-free MST. It uses the proteins intrinsic Trp and Tyr fluorescence to track the movement of molecules in a temperature gradient. As all other fluorescence-based techniques, label-free MST must deal with fluorescence interference caused by certain compounds. To increase the applicability of the NT.LabelFree device, I therefore first tested the assumption that it is possible to reduce the fluorescence interference by separating the fluorescence signal that arises from the protein from that of the compound by using appropriate emission filter.

To address the first question, I analyzed the fluorescence interference of 1513 small molecules and fragments out of seven different libraries and 8 commercially available inhibitors. Experimental data showed that this simple optical modification could decrease the percentage of interfering compounds from 70 to 36 %, depending on the compound library. However, it is important to note that the number of interfering compounds highly depends on the composition of the compound library. Although many libraries are based on privileged structures like indoles, that are also part of Trp. These compounds don´t necessarily interfere with the detection as already slight modifications of chemical substitutes in the compound structure can significantly alter the respective emission profile. Thus, it remains hard to predict how many compounds will interfere with the MST assay in an HTS, even if their basic structure is known.

An additional disadvantage of the narrower filter bandwidth is the reduced fluorescence intensity (down to ~ 50 %), depending on the location and amount of Trp residues in the protein. As this reduction in overall sensitivity can cause lower S/N ratio, as shown for MBP maltose interaction, the implementation of a smaller emission filter requires additional adjustments in the optical system of the device to compensate this loss of intensity. This could be achieved using different detection systems or stronger LED power.

Nevertheless, the use of narrower emission filter allowed characterization of selected interactions, that were previously not accessible to label-free MST because of high fluorescence interference. As an example, the interaction between p38α and SB203580 was used, as the narrow emission filter completely blocked the compound´s fluorescence. Here, the determination of a K_d was possible by analyzing binding specific

ligand-induced changes in the fluorescence of the titration series. This measurement revealed an additional weak point of the narrower emission filter. With the use of this filter, one can easily see slight changes in the Trp or Tyr fluorescence shift caused by the ligand binding. For the analysis of this data, the ligand-induced fluorescence change is used to determine the binding affinity. Such data mostly exhibit lower S/N ratios, because data normalization is missing. Also, an additional verification step by performing SD test is required to determine the specificity of observed fluorescence change.

Based on these findings it was concluded that although changes in the bandwidth of the emission filter can improve the number of compounds analyzed by label-free MST, this improvement does not generally suffice. Also, this approach is only applicable for the quantification of interactions between small molecule compounds and fragments. The analysis of interactions where both of the interacting species are proteins is still not possible by label-free MST, because of the too high similarity in the emission spectra of proteins. Thus, for the analysis of protein-protein interactions by label-free MST, a novel approach is required.

In the following chapter, I describe the development and validation of a label-free approach for the determination of the binding affinity between two proteins.

Chapter 2

1   Introduction

Most biological functions in a cell are controlled by proteins. Over 80 % of proteins do not function as independent cellular components but interact with other proteins to fulfill a large variety of cellular functions[85]. Those range from gene expression, cell proliferation and apoptosis, to signal transduction and transport[31,86]. Thus, protein-protein interactions (PPIs) play a major role in nearly every biological process. Regarding the human body, it has been estimated that around 170 000 PPIs exist, of which most are still unexplored[5]. Since abnormal PPIs, such as loss of interactions or gain of inappropriate interactions, can cause several diseases, the study of PPIs is of huge interest for pharmaceutical research[87,88]. In this regard, the inhibition of PPIs, using small molecular inhibitors has emerged as a field of interest during the last few years, leading to the investigation of PPI modulators as potential therapeutic agents[89–92].

Molecular forces, that are involved in PPIs, are a combination of hydrophobic bonding, van der Waals forces and salt bridges, where the surface which is involved in the binding process can span only view amino acids, or can comprise large surface areas[93]. In addition, there is a direct correlation between binding affinity and the amount of surface area buried at the interface, while no relationship between binding affinity and the chemical composition of the interface could be observed[93]. Binding affinities can span a wide range from pM to the mM range[93]. To study binding affinities, different methods can be applied, as already discussed in the general introduction of this thesis. MST offers the advantage of being a fast and reliable technique for the in-solution quantification of molecular binding events, with only low amounts of sample required. In addition, it offers the possibility for label-free measurements, that so far is restricted to interactions, in which only one binding partner exhibits intrinsic fluorescence. As the highly conserved nature of proteins does not allow for signal separation as it was applied in the previous chapter, another solution strategy is required. Thus, it was tested whether the ligands fluorescence contribution can be used for data analysis rather than being eliminated from the overall signal. Therefore, the concept of data analysis in dynamic light scattering (DLS) was chosen as a model system, as there also all involved particles contribute to the detected signal when molecular interactions are analyzed[94].

In general, DLS analyses the diffusion behavior of macromolecules to establish the size, shape, and molecular weight of macromolecules in solution[95]. Therefore, it records time-

dependent fluctuations in scattered light, which arise from Brownian motion of scattering molecules[31]. This information is used to calculate the diffusion coefficient and the hydrodynamic diameter, using the Stokes-Einstein equation[31]. Thereby, the diameter that is measured refers to how a particle diffuses within a fluid. The larger a particle is, the slower the diffusion speed will be[96]. As both, the diffusion coefficient and hence the hydrodynamic radius depend on the size and shape of macromolecules, changes in the average molecular radius or molecular mass provide a direct indication of the formation or dissociation of complexes[97].

For sample preparation for DLS binding studies, a variation of the experimental method known in the literature as the method of continuous variation, or the Job Plot, is used, referred to as composition gradient titration (CGT)[31,98]. Thereby, the total molar concentration is kept constant, while their ratio of interacting partners is varied[31]. In DLS, the average hydrodynamic radius is then plotted against the mole fraction of two analyzed molecules. For the case of an interaction, measured values will increase to a higher extent than expected for a solution of two non-interacting monomers[31]. This increase can then be quantified and used for K_d determination.

At first sight, this approach is not intuitive, as usually binding studies are carried out by holding the concentration of one binding partner fixed while varying the concentration of the other molecule[99]. Such experiments can then be illustrated as saturation curves, from which the K_d can easily be gathered, as illustrated in Figure 20. In contrast, instead of varying the concentration of one component at a time, the method of continuous variation holds the total concentration of both molecules constant, while varying the relative proportions of both species[31]. Having such a titration scheme, the K_d cannot easily be taken from the graph but can be quantified with a nonlinear least mean square fit[37]. However, the shape of the curve provides already qualitative information about the binding affinity of the regarded interaction, as strong binding results in a more triangle-shaped graph, whereas lower affinities show a more curved form (Figure 20)[100]. Moreover, this approach can be used to determine binding stoichiometry, if both binding partners exhibit the same starting concentration. Here, the position of the peak maximum is then used for binding stoichiometry determination, whereby a simple 1:1 interaction will have a maximum increase in the measured parameter at equal amount of both proteins, whereas 1:2 or 2:1 interactions exhibit maximum change at below or above equal concentration, respectively[100].
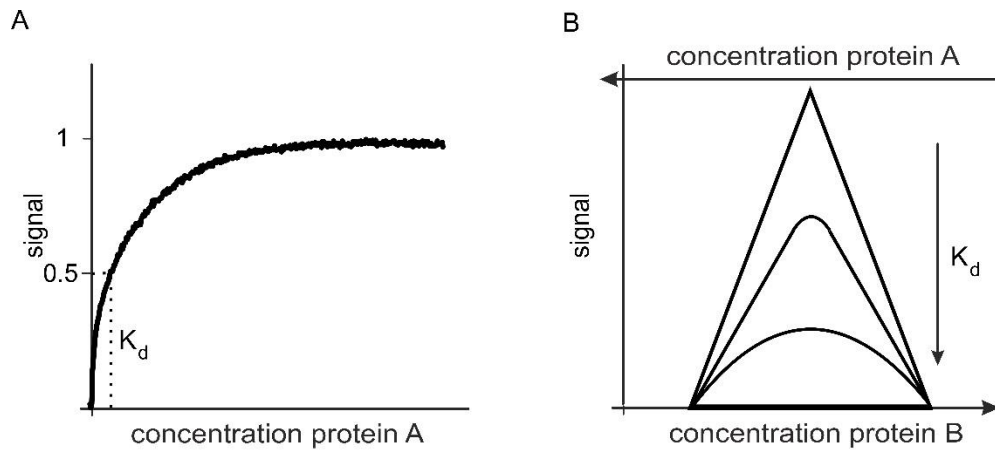
Figure 20: Saturation curve and CGT data. Results of a protein-protein interaction analysis from a saturation experiment (A) and from a DLS experiment, in which the method of continuous variation was applied (B). While the K_d of the interaction can be obtained from the plot in A, only a qualitative value for the underlying K_d can be gained from the experiment in B. Here, quantification of the interaction requires further fitting of the measurement to simulated data.

In the following chapter, I examined whether the same approach, namely CGT for K_d determination, can also be applied for label-free MST affinity analysis of two fluorescent species. Therefore, the general applicability of this titration for label-free affinity analysis of PPIs will be investigated and limitations of this approach will be discussed. In summary, a novel approach to analyze PPIs by MST in a label-free format will be implemented.

2 Materials

Table 16: Chemicals

Chemicals	Company
EDTA	Sigma-Aldrich Co. LLC, St. Louis, US, cat. no. 1233508
H_2O AnalaR NORMAPUR®	VWR Prolabo Chemicals, Darmstadt, DE, cat. no. 102927G
HEPES	Carl Roth GmbH + Co. KG, Karlsruhe, DE, cat. no. HN78.2
MST buffer	NanoTemper Technologies GmbH
NaCl	Carl Roth GmbH + Co. KG, Karlsruhe, DE, cat. no. 0962.1
Pluronic® F-127	Sigma-Aldrich Co. LLC, St. Louis, US, cat. no. P2443
Roti®-Stock 10x PBS	Carl Roth GmbH + Co. KG, Karlsruhe, DE cat. no. 1058.1
TRIS PUFFERAN®	Carl Roth GmbH + Co. KG, Karlsruhe, DE 4855.1
Tween®-20	Sigma-Aldrich Co. LLC, St. Louis, US, cat. no. 93773

Table 17: Buffers and solutions

Buffers and solutions	Composition
MST buffer	50 mM Tris-HCl, 250 mM NaCl, 10 mM $MgCl_2$, pH 7.8
MSTP buffer	0.1 % Pluronic® F-127 in MST buffer
MSTT buffer	0.05 % Tween® 20 in MST buffer
NHS labeling buffer	130 mM $NaHCO_3$, 50 mM NaCl, pH 8.2–8.3
PBSP	0.1 % Pluronic® F-127 in PBS buffer
PBST	0.05 % Tween® 20 in PBS buffer

Table 18: Proteins and oligonucleotides

Proteins and oligonucleotides	Company
BLIP DBV1	CRELUX GmbH, Munich, DE
DNA oligo strands: 5´-Cy5-ATAT TTA CGA TCT GAT CCT T -3´) 3´-AAT GCT AGA CTA GGA A (TATA-Cy5)-5´ 3´-AAT GCT A*C*A CTA GGA A (TATA-Cy5)-5´ 3´-AAT GCT ACT CTA GGA A (TATA-Cy5)-5´	Metabion, Planegg, DE
IL6 antigen	Was kindly provided by Dr. Patrick Kunz from the functional genome analysis department of the german cancer research center, Heidelberg, DE

IL6 nanobody	Was kindly provided by Dr. Patrick Kunz from the functional genome analysis department of the german cancer research center, Heidelberg, DE
MBP	antibodies-online GmbH, Atlanta, US, cat. no. MBP0801
MBP binding protein	ChromoTek GmbH, Planegg, DE
TEM DBV1	CRELUX GmbH, Munich, DE

Table 19: Fluorophores

Fluorophores	Company
ATTO647-NHS ester MW:811 g/mol; ε: 120 000 M^{-1}cm^{-1}	Sigma-Aldrich Co. LLC, St. Louis, US, cat. no. 07376
NT647-NHS ester MW: n.a.; ε: 250 000 M^{-1}cm^{-1}	From NanoTemper Technologies MO-L001 labeling kit, NanoTemper Technologies GmbH, Munich, DE

Table 20: Kits

Kits	Company
Monolith Protein Labeling Kit RED-NHS (Amine Reactive)	MO-L001, NanoTemper Technologies GmbH, Munich, DE

Table 21: Devices

Machines and devices	Company
Monolith NT.115 (RED/BLUE, BLUE/GREEN, RED/GREEN), Monolith NT.115.^{Pico}	NanoTemper Technologies GmbH, Munich, DE
NanoDrop^{TM} One	Thermo Fisher Scientific Inc., Wilmington, US
Monolith NT.LabelFree	NanoTemper Technologies GmbH, Munich, DE

Table 22: Centrifuges and rotors

Centrifuges and rotors	Company
Centrifuge 5424	Eppendorf AG., Hamburg, DE
Centrifuge 5430 R	Eppendorf AG., Hamburg, DE
Rotor FA-45-24-11-HS	Eppendorf AG., Hamburg, DE

Table 23: Online tools and software

Online tools and software	Company
MO.AffinityAnalysis_x86_2.2.7.6056	NanoTemper Technologies GmbH, Munich, DE
MO.Control_x86_1.5.3.6096	NanoTemper Technologies GmbH, Munich, DE

Table 24: Further materials

Further material	Company
10 mL (12 mL) NORM-JECT®,	Henke Sass Wolf, Tuttlingen, DE
CHROMAFIL® RC-20/25, pore size: 0.2 µm	MACHEREY-NAGEL GmbH & Co. KG, Düren, DE
Monolith NT.115 MST Premium Coated Capillaries	NanoTemper Technologies GmbH, Munich, DE
Monolith NT.115 Standard Treated Capillaries	NanoTemper Technologies GmbH, Munich, DE
Monolith NT.LabelFree Zero Background MST Premium Coated Capillaries	NanoTemper Technologies GmbH, Munich, DE
Monolith NT.LabelFree Zero Background Standard Treated Capillaries	NanoTemper Technologies GmbH, Munich, DE
SPROUT® MINI CENTRIFUGE 12V	Heathrow Scientific®, Vernon Hills, US

3 Methods

3.1 Discrimination between binder and non-binder molecules

To investigate, whether an interaction can be distinguished from a non-binding event when both molecules emit in the detection wavelength, binding check experiments were carried out using exemplary samples for both devices, NT.115 and NT.LabelFree. For experiments on the NT.115, two red-emitting fluorophores (ATTO647-NHS ester and NT647-NHS ester) were chosen as an example for a non-binding, whereas RED-tris-NTA against NT647-labeled His_6-p38α was selected as an example of an interaction. For label-free measurements, His_6-p38α against TEM1 was chosen as an example for no-binding, while BLIP against TEM1 was selected for two interacting molecules.

3.1.1 No interaction on NT.115

NT647-NHS ester and ATTO647-NHS ester were diluted to 100 nM using 20 mM Tris buffer pH 7.5, supplemented 100 mM NaCl and 0.05 % Tween. Solutions were stored overnight at room temperature to let the NHS ester groups hydrolyze. Un-reactive dyes were then mixed in a 1:1 ratio. Afterward both dyes alone and as a 1:1 mixture were filled into Monolith NT.115 Capillaries and MST traces were recorded at 25 °C, 20 % LED and medium MST power. Measurements were carried out in triplicates.

3.1.2 Interaction on NT.115

RED-tris-NTA and NT647-labeled His_6-p38α were diluted to 50 nM in PBST buffer and then mixed 1:1. After incubating the reaction mixture for 30 min at room temperature away from light, RED-tris-NTA, NT647-labeled-His_6-p38α and the mixture of both samples were filled into Monolith NT.115 Premium Capillaries and loaded into the NT.115 device. MST traces were recorded at 25 °C, 20 % LED and medium MST power. Samples were measured in triplicates.

3.1.3 No interaction on NT.LabelFree

His_6-p38α and TEM1 were both diluted to 1 µM using sterile filtered MSTP buffer. After centrifugation at 14 000 g, 4 °C for 10 min, samples were mixed 1:1. After incubation for 15 min at room temperature, both proteins separately and as 1:1 mixture were filled into Monolith NT.LabelFree Premium Capillaries and loaded into the NT.LabelFree device.

MST traces were recorded at 20 % LED and high MST power at 25 °C. Samples were measured in triplicates.

3.1.4   Interaction on NT.LabelFree

BLIP and TEM1 were both diluted to 1 µM using sterile filtered MSTP buffer. After centrifugation at 14 000 g, 4 °C for 10 min, samples were mixed 1:1 and reaction was incubated for 15 min at room temperature. Afterward both proteins alone and the 1:1 mixture were filled into Monolith NT.LabelFree Premium Capillaries and loaded into the device. MST traces were recorded at 25 °C, 20 % LED and high MST power. Measurements were carried out in triplicates.

3.2   Correlation between fluorescence intensity and F_{norm} value

To test, whether the F_{norm} value is independent of the fluorescence intensity of a sample, three different proteins were titrated against PBSP or MSTP buffer using the CGT pipetting scheme below. Therefore, TEM1 was diluted to 300 nM using MSTP buffer, while IL6 nanobody and IL6 antigen were both diluted to 300 nM using PBSP buffer. After titration of the CGT against the corresponding buffer, samples were filled into Monolith NT.LabelFree Premium Capillaries and MST was recorded at 20 % or 40 % LED power for TEM1 and IL6 antigen and nanobody, respectively and at high MST power. Measurements were carried out at 25 °C. Data was evaluated after 20 sec laser-on time.

Table 25: CGT pipetting scheme

Vial number	MBP-3xMyc	PBSP
1	25.2 µL	0.0 µL
2	23.4 µL	1.8 µL
3	21.6 µL	3.6 µL
4	19.8 µL	5.4 µL
5	18.0 µL	7.2 µL
6	16.2 µL	9.0 µL
7	14.4 µL	10.8 µL
8	12.6 µL	12.6 µL
9	10.8 µL	14.4 µL
10	9.0 µL	16.2 µL
11	7.2 µL	18.0 µL
12	5.4 µL	19.8 µL
13	3.6 µL	21.6 µL
14	1.8 µL	23.4 µL
15	0.0 µL	25.2 µL

3.3   Standard MST experiments

To verify obtained K_d values using the CGT approach, the same interactions were quantified using standard MST affinity analysis, as described below:

3.3.1   DNA hybridization

DNA hybridization was quantified using a Cy5-labeled single-stranded DNA template (5´-Cy5 – ATAT TTA CGA TCT GAT CCT T -3´) against one of three unlabeled single stranded DNA molecules: perfect match: 3´-AAT GCT AGA CTA GGA A-5´, mismatch 1: 3´-AAT GCT A*C*A CTA GGA A-5´, or mismatch 2: 3´-AAT GCT ACT CTA GGA A-5´. For the experiments, unlabeled DNA strands were diluted in 20 mM HEPES buffer, pH 7.4, supplemented with 100 mM NaCl, 5 mM EDTA and 0.05 % Tween. Starting from 2 µM highest concentration, a 16-step serial dilution was prepared. Afterward 2 nM Cy5-labeled template strand was then added to the dilution series and the reaction was incubated for 10 min at room temperature in the dark. Samples were then filled into Monolith NT.115 Capillaries and loaded into the Monolith NT.115[Pico] device. Measurements were carried out at 5 % LED, medium MST power and at 25 °C. Data was quantified after 5 sec laser on-time. Measurements were carried out in triplicates.

3.3.2 IL6 nanobody against IL6 antigen or against BIRC75 antigen

Labeling of IL6 and BIRC75 antigen was carried out according to the Monolith Protein Labeling Kit RED-NHS (Amine Reactive) from NanoTemper Technologies GmbH. Briefly, 20 µM of antigen was mixed with 60 µM of RED-NHS dye using NHS labeling buffer. Reaction mixture was incubated for 30 min at room temperature before unreacted dye was removed using size-exclusion chromatography. For this, a B column was equilibrated using MSTT buffer, which was also used for elution of labeled antigen. Protein was then aliquoted and stored at -80 °C.

For the MST experiment of RED-NHS IL6 antigen or RED-NHS BIRC75 antigen against IL6 nanobody, a serial dilution of L6 nanobody was prepared, starting from 10 µM highest concentration using MSTT buffer. Afterward 40 nM RED-NHS-labeled antigen was added to all dilutions and reaction was incubated for 15 min at room temperature away from light. Samples were then filled into Monolith NT.115 Premium Capillaries and loaded into the Monolith NT.115 device. The MST measurement was carried out at 40 % LED and at 80 % MST power at 25 °C. Data was quantified after 10 sec laser-on time and measurements were carried out in triplicates.

3.3.3 MBP binding protein against MBP

His_6-MBP-binding protein was diluted to 200 nM in PBST buffer. RED-tris-NTA dye was diluted in PBST buffer to a final concentration of 100 nM. 100 µL of protein was then mixed with 100 µL of dye and the reaction mixture was incubated for 30 min at room temperature in the dark.
For the serial dilution, 5 µM of MBP diluted in PBST buffer was used as the highest ligand concentration of the 16-step dilution series, with a final volume of 10 µL in each reaction tube.
Then 10 µL of 100 nM labeled protein was added to all 16 vials and samples were mixed by pipetting up and down. Reactions were incubated for 30 min at room temperature away from light and then filled into Monolith NT.115 Capillaries. Using the Monolith NT.115 device, MST was carried out at 40 % LED and high MST power. Data were evaluated at 15 sec MST-on time. Measurements were carried out in triplicates.

3.4   Composition gradient titration (CGT)

3.4.1   Pretest to check for fluorescence difference between both molecules

For the CGT experiment the molecules concentrations need to be adjusted in a way, that both exhibit the same fluorescence intensities. Therefore, a pretest was carried out, in which molecules are diluted in MSTP buffer to a final concentration of 100 nM and were then filled into Monolith NT.LabelFree Premium Capillaries. Then capillaries were scanned to record the fluorescence intensity of both molecules, using 20 % LED power.

3.4.2   Titration scheme

For the CGT, both proteins were concentrated at least ~10 to 100-fold above the K_d, in a volume of 200 µL each. Prior titration of the CGT, samples were centrifuged at 14 000 g for 10 min at 4 °C. Afterward samples were titrated according to the following pipetting scheme:

Table 26: CGT pipetting scheme

Vial number	Molecule A	Molecule B
1	25.2 µL	0.0 µL
2	23.4 µL	1.8 µL
3	21.6 µL	3.6 µL
4	19.8 µL	5.4 µL
5	18.0 µL	7.2 µL
6	16.2 µL	9.0 µL
7	14.4 µL	10.8 µL
8	12.6 µL	12.6 µL
9	10.8 µL	14.4 µL
10	9.0 µL	16.2 µL
11	7.2 µL	18.0 µL
12	5.4 µL	19.8 µL
13	3.6 µL	21.6 µL
14	1.8 µL	23.4 µL
15	0.0 µL	25.2 µL

Samples were then filled into Monolith NT.LabelFree (Premium) Capillaries, and loaded into the NT.LabelFree device to record F_{norm} values for each capillary.

3.4.3 CGT data analysis

Analysis was done using Microsoft Excel. First, measurements were transferred to the MO.Affinity Analysis software, in which an appropriate laser-on time was selected. Corresponding F_{norm} values were then exported to Microsoft Excel and plotted against the capillary position. Afterward F_{norm} values for an interaction and for the case of no-interaction were simulated and plotted into the same graph. By minimizing the difference between the measured data and the simulated data for an interaction, while altering the K_d and the complex F_{norm} value using the solver function, the K_d was determined.

In principle, this fitting consists of four parts, which will be described below:

1) Modeling the complex- and the free concentrations in solution:

Regarding the reversible interaction between protein A and B,

$$A + B \longleftrightarrow AB \quad , \tag{8}$$

the equilibrium dissociation constant (K_d) can be expressed using the law of mass action

$$K_d = \frac{[A][B]}{[AB]} \quad \text{or} \quad [AB] = \frac{[A][B]}{K_d} \quad \text{(I)} \quad , \tag{9}$$

whereby [A] and [B] are the free concentrations of protein A and B, respectively and [AB] is the concentration of the complex.

The known total molar concentrations of two proteins [A_{tot}] and [B_{tot}] can be expressed as

$$[B_{tot}] = [B] + [AB]$$

$$[A_{tot}] = [A] + [AB] \tag{10}$$

or as

$$[AB] = [A_{tot}] - [A] \quad \text{(II)}$$

$$[AB] = [B_{tot}] - [B] \quad \text{(III)} \quad . \tag{11}$$

Equalizing the two equations (I) and (II) or (I) and (III) and solving them for [A] will give the equations (IV) and (V):

$$[A] = \frac{[A_{tot}]}{(1 + \frac{[B]}{K_d})} \quad \text{(IV)}$$

$$[A] = \frac{K_d([B_{tot}] - [B])}{[B]} \quad \text{(V)}$$

. (12)

Equalizing the two equations (IV) and (V) and solving it for [B] with the quadratic equation yields:

$$[B] = -\frac{1}{2}\{([A_{tot}] - [B_{tot}] + K_d) - \{([A_{tot}] - [B_{tot}] + K_d)^2 + 4[B_{tot}]K_d\}^{\frac{1}{2}}\}$$

. (13)

Knowing the total concentrations of A and B in each titration step, their free concentrations and the complex concentration could be calculated, using an assumed K_d and the following equations, derived from the law of mass action as described above:

$$[AB] = \frac{[A][B]}{K_d}$$

(14)

$$[B] = -\frac{1}{2}\{([A_{tot}] - [B_{tot}] + K_d) - \{([A_{tot}] - [B_{tot}] + K_d)^2 + 4[B_{tot}]K_d\}^{\frac{1}{2}}\}$$

. (15)

2) Modeling F_{norm} values for an interaction and for the case of no-interaction, using the following equations with an estimated complex F_{norm} and an assumed K_d.

Fnorm for no interaction: $\dfrac{FnormA * [A_{tot}] + FnormB * [B_{tot}]}{[A_{tot}] + [B_{tot}]}$

Fnorm for interaction: $\dfrac{FnormA * [A] + FnormB * [B] + FnormAB * [AB]}{[A] + [B] + [AB]}$

(16)

3) Fitting modeled results to measurements by minimizing the difference between both graphs, using the solver function in Microsoft Excel.

Therefore, for each F_{norm} value of the titration series, the difference between the measured and the simulated F_{norm} value for the case of an interaction was calculated and squared. All squared deviations were then taken to a sum. This sum was then minimized, by adjusting the K_d and the complex F_{norm} value, using the solver function. To estimate the quality of the fit, a quadratic error was calculated as well. Therefore, the square root of the minimized sum was calculated and divided by the number of titration steps, in this case,15.

3.5 Quantified interactions using the CGT

3.5.1 DNA Hybridization

All DNA single strands (template strand: 5´-Cy5 – ATAT TTA CGA TCT GAT CCT T -3´, perfect match single strand: 3´-AAT GCT AGA CTA GGA A TATA-Cy5-5´, mismatch 1 single strand: 3´-AAT GCT A*C*A CTA GGA A TATA-Cy5-5´, mismatch 2 single strand: 3´-AAT GCT ACT CTA GGA A TATA-Cy5-5´), were diluted in 25 mM HEPES buffer, pH 7.4, supplemented with 100 mM NaCl, 5 mM EDTA, 0.1 % Pluronic, to reach a final concentration of 100 nM. These stock solutions were centrifuged at 14 000 g for 10 min at 4 °C before they were used for CGT. After titrating the samples, they were filled into Monolith NT.115 Capillaries and loaded into the device. MST traces were recorded at 20 % LED and medium MST power. F_{norm} values after 5 sec laser-on time were taken for analysis. Measurements were carried out in triplicates.

To investigate the influence of the start concentration on the obtained K_d, perfect match DNA hybridization was used. Therefore, 100 nM, 50 nM and 5 nM solutions were prepared for CGT, as described above. F_{norm} values were recorded using 2 % / 20 % and 100 % LED power of 100 nM / 50 nM / 5 nM, respectively and medium MST power at 25 °C. Data was evaluated after 10 sec laser-on time.

3.5.2 IL6 Nanobody against IL6 antigen or against BIRC75 antigen

Proteins were diluted in MSTP buffer to reach a final volume of 100 nM. Samples were then filled into Monolith NT.LabelFree Premium Capillaries and loaded into the NT.LabelFree device. Fluorescence intensity was recorded using the expert function of the MO.Control software and a LED power of 20 %. As the fluorescence intensity of IL6 nanobody was 1.4 times higher than that of the antigens, the concentrations were adjusted to exhibit the same fluorescence intensity. Therefore, 250 nM of IL6 nanobody

and 350 nM antigens were prepared using MSTP buffer and samples were centrifuged at 14 000 g, for 10 min at 4 °C, before CGT was carried out. Samples were then incubated for 15 min at room temperature and were filled into Monolith NT.LabelFree Premium Capillaries and loaded into the NT.LabelFree device. MST was recorded at 25 % LED and medium MST power. Data was evaluated after 5 sec MST-on time. F_{norm} values were exported into Microsoft Excel, for data analysis as described above.

3.5.3   MBP binding protein against MBP

Both proteins were diluted in MSTP buffer to reach a final volume of 100 nM. Samples were then filled into Monolith NT.LabelFree Premium Capillaries and loaded into the Monolith NT.LabelFree device. Fluorescence intensity was recorded using the expert function of the MO.Control software and a LED power of 20 %. As the fluorescence intensity of MBP was 2 times higher than that of MBP binding protein, the concentrations were adjusted to exhibit the same fluorescence intensity. Therefore, 100 nM and 200 nM were prepared using MSTP buffer and samples were centrifuged at 14 000 g, for 10 min at 4 °C, before CGT was carried out. Samples were then incubated for 15 min at room temperature and were filled into Monolith NT.LabelFree Premium Capillaries and loaded into the Monolith NT.LabelFree device. MST was recorded at 10 % LED and high MST power. Data was evaluated after 15 sec MST-on time. F_{norm} values were exported into Microsoft Excel, for data analysis as described above.

To investigate the influence of the start concentration on the obtained K_d using MBP against MBP binding protein, 100 nM / 200 nM / 400 nM MBP was titrated against 200 nM / 400 nM / 800 nM of MBP binding protein, respectively as described above. F_{norm} values were recorded using 10 % LED and high MST power at 25 °C. Data was evaluated after 15 sec laser-on time. Experiments were carried out in triplicates.

Excel sheets with the complete CGT data analysis for all investigated interactions are provided in the supplementary data.

3.6   Simulations to determine the limitations of the CGT

Simulations were carried out using Microsoft Excel. For the simulations, F_{norm} values for the case of an interaction were simulated, using the following equation:

$$\text{Fnorm for interaction:} \quad \frac{FnormA * [A] + FnormB * [B] + FnormAB * [AB]}{[A] + [B] + [AB]} \tag{17}$$

The F_{norm} value of A was fixed to 725 and the F_{norm} value of the complex was fixed to 825. All other values were varied. Parameters are listed in the table below:

Table 27: Parameters for CGT simulation. Parameters for simulation of CGT label-free MST measurements of the interaction of A and B.

Parameter	Values used for simulation
K_d	0.01 nM, 0.1 nM, 1 nM, 10 nM, 100 nM, 1 µM, 10 µM, 100 µM, 1 mM, 10 mM
F_{norm} value of A / B	725 / 725, 725 / 750, 725 / 775, 725 / 825
Start concentration of A / B	10 nM / 10 nM, 10 nM / 15 nM, 10 nM / 25 nM, 10 nM / 50 nM, 10 nM / 100 nM
	100 nM / 100 nM, 100 nM / 150 nM, 100 nM / 250 nM, 100 nM / 500 nM, 100 nM / 1000 nM
	1 µM / 1 µM, 1 µM / 1.5 µM, 1 µM / 2.5 µM, 1 µM / 5 µM, 1 µM / 10 µM
	10 µM / 10 µM, 10 µM / 15 µM, 10 µM / 25 µM, 10 µM / 50 µM, 10 µM / 100 µM
	100 µM / 100 µM, 100 µM / 150 µM, 100 µM / 250 µM, 100 µM / 500 µM, 100 µM / 1000 µM
	1 mM / 1 mM, 1 mM / 1.5 mM, 1 mM / 2.5 mM, 1 mM / 5 mM, 1 mM / 10 mM
	10 mM / 10 mM, 10 mM / 15 mM, 10 mM / 25 mM, 10 mM / 50 mM, 10 mM / 100 mM

3.7 Fluorescence intensity of p38α and SB203580

To compare the intrinsic fluorescence intensity of p38α and SB203580, both molecules were diluted in PBSP buffer to reach a final concentration of 1 µM. Samples were then filled into Monolith NT.LabelFree Premium Capillaries and loaded into the Monolith NT.LabelFree device. Fluorescence was excited at 20 % LED power.

4   Results

This chapter deals with the quantification of PPIs using label-free MST. Therefore, composition gradient titration will be used, which on the one hand prevents detector saturation, while on the other hand, it allows the determination of K_d using a least mean square approximation. The fitting algorithm is based on the assumption that one can clearly distinguish a binding from a non-binding event. Therefore, it was tested whether binder and non-binder molecules can be discriminated using MST, if both molecules exhibit fluorescence in the same spectral region. This should be possible as in theory the measured time trace of a 1:1 mixture of two not-interacting molecules is simply the meantime trace of the two single time traces. This principle can further be applied to all different molecule to molecule ratios, as the recorded time trace would be a mixture of the single time traces, each weighted by the relative amount of each molecule in the mixture. In case of an interaction, the measured time trace would be a mixture of the single time traces plus the unknown time trace of the formed complex. The extent to which all three single time trace species contribute to the resulting time trace is unknown as long as the dissociation constant is not known. Figure 21 illustrates the two scenarios, in which (A) shows the case for no interaction and (B) represents two interacting molecules.
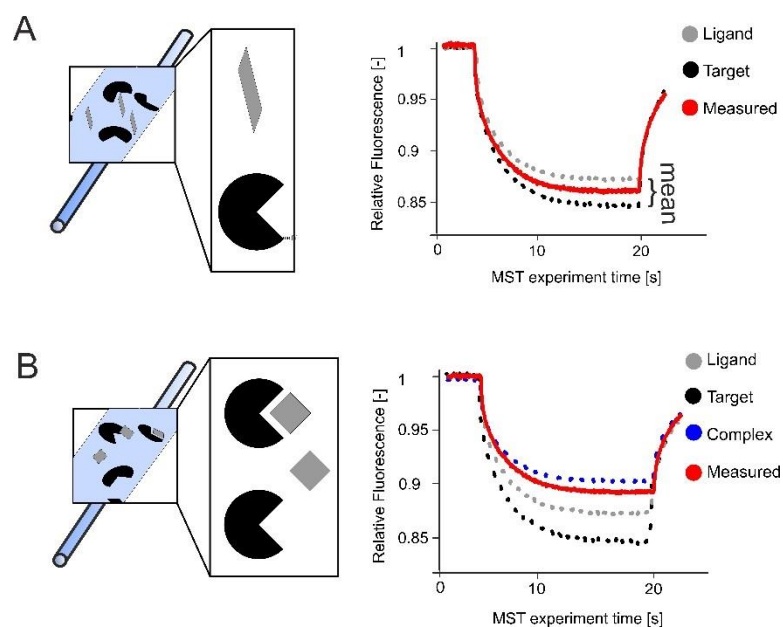
Figure 21: MST traces for binder and non-binder molecules. Illustration of sample filled capillaries with corresponding MST time traces. A) Two not-interacting molecules are shown together with their MST time traces, which are represented in the same color. The red time trace represents the recorded one from a 1:1 mixture of both molecules. B) Two interacting molecules are shown as single molecules as well as formed complex. The MST time traces on the right illustrate the time traces of each single species. The red time trace represents the recorded time trace.

To confirm these theoretical ideas, MST time traces of two single molecules as well as of a 1:1 mixture of both molecules were recorded. This was done for different sets of samples, including binder and non-binder molecules. According to the theory described above, no interaction occurred if there is an overlap between the measured "mixed" time trace and the "mean" time trace of the single-molecule time traces. If there is a difference between both, then an interaction occurred.

For these experiments, I first used the NT.115 device to prevent any issues from dust or other fluorescent artifacts. Therefore, two not-reactive red fluorescent dyes, ATTO647 and NT647, were chosen as an example for "no interaction". As expected, the measured time trace clearly is the mean of both single time traces, and thus no interaction occurred (Figure 22). As example for an interaction, red-NHS labeled His_6-p38α was mixed with RED-tris-NTA. Here, the measured time trace of the mixture of both molecules can clearly be distinguished from a mean time trace of the single molecules, and thus an interaction took place.
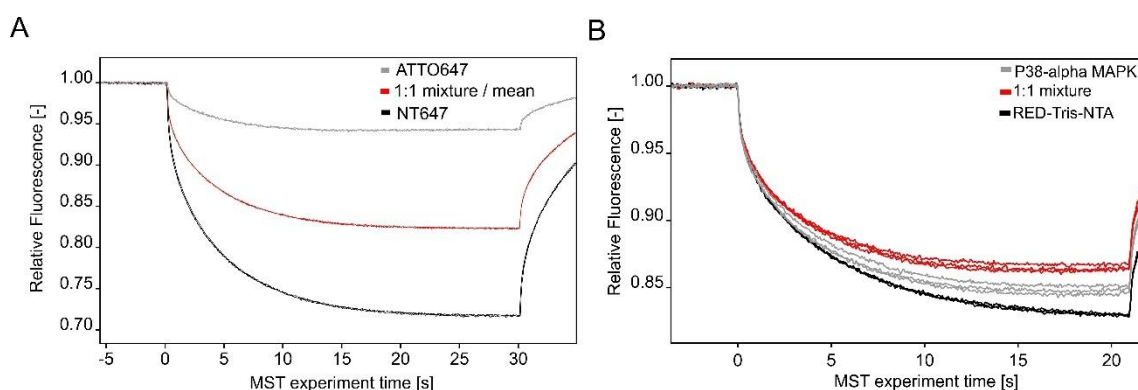
Figure 22: Differentiation of binder and non-binder molecules. Differentiation of binder and non-binder molecules on the NT.115. A) MST time traces of ATTO647, NT647 and a 1:1 mixture of both. Single dyes and the mixture were loaded into Monolith NT.115 Capillaries. Thermophoresis was recorded at 20 % LED, medium MST power and 25 °C. B) MST time traces of red-NHS labeled His_6-p38α, RED-tris-NTA and both molecules together, in a 1:1 ratio. Samples were loaded into Monolith NT.115 Premium Capillaries separately and as 1:1 mixture. Thermophoresis was recorded at 20 % LED, medium MST power and at 25 °C. Legend) The grey and black traces represent the single molecular time traces, while the red one shows the time trace of the mixture and of the mean value of the single time traces. Measurements were carried out in triplicates.

Next, similar experiments were done using the NT.LabelFree device. For this, TEM1 and His_6-p38α was chosen as an example for no interaction, while TEM1 and BLIP served as an example for an interaction. Figure 23 shows the results for these experiments. The grey and black traces represent the single molecules, while the red trace shows the 1:1 mixture. Here, the MST trace of the combined sample of TEM1 and His_6-p38α lie in-

between the single time traces, indicating that no interaction took place. In contrast, the time trace of TEM1 and BLIP mixture can nicely be distinguished from the mean of both single traces. This is an indication for an underlying binding event.
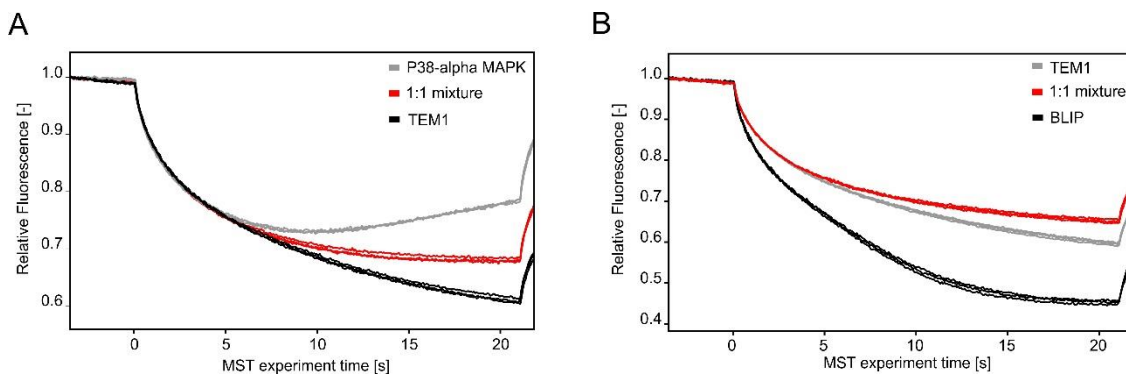
Figure 23: Discrimination of binder and non-binder molecules. Discrimination of binder and non-binder molecules on the NT.LabelFree device. A) MST time traces of TEM1, His6-p38α and of both molecules together. B) MST time traces of TEM1, BLIP and of both molecules together. A&B) Molecules were loaded into Monolith NT.LabelFree Premium Capillaries separately and as 1:1 mixture. Each sample was measured in triplicates at 20 % LED, high MST power and at 25 °C. Legend) The grey and black traces represent the single molecular time traces, while the red one shows the time trace of the mixture and of the mean value of the single time traces. Measurements were carried out in triplicates.

The referencing approach described above, offers an easy and straightforward way to discriminate between binder and non-binder molecules. However, it has the potential to generate false positives, because small changes in the thermophoretic time trace of both molecules already implies that an interaction occurred. To prevent such false positives and as a prove of value it is best to run several different compositions to increase the number of samples and to validate the data.

4.1 From discriminating binder and non-binder molecules to K_d quantification

The fact that it is possible to discriminate binder from non-binder molecules can be applied to a whole titration series, were different compositions of both molecules are used to then later extract the dissociation constant. This is possible as the time traces for every single titration step would be a superposition of the time traces of the two unbound proteins and of the complex time trace, all weighted by their relative amount in the sample.

After recording F_{norm} values from the titration series, cases for no interaction and for interaction were simulated using the following equations for each capillary:

Fnorm for no interaction:
$$\frac{FnormA*[A_{tot}] + FnormB*[B_{tot}]}{[A_{tot}]+[B_{tot}]}$$

Fnorm for interaction:
$$\frac{FnormA*[A]+FnormB*[B]+FnormAB*[AB]}{[A]+[B]+[AB]}$$

(18)

The difficulty for this simulation was the unknown F_{norm} value of the formed complex and the unknown K_d value, which is required for calculating the free and the complex concentrations. Hence, a K_d from a standard experiment was used as a first guess. For the complex F_{norm} value an assumption was generated. These formulas were then used to calculate a preliminary F_{norm} value for each capillary, for the case of an interaction and for the case of no interaction. These values were then plotted in one graph, together with the measured data. The dissociation constant was then obtained by a least mean square fitting approach that minimizes the square difference between the measured F_{norm} values and the calculated ones for the case of an interaction. Therefore, the deviation between the calculation and the measurement was squared and all squared deviations were added to a sum. This sum of squared deviations was minimized by a least mean square approximation through varying the initial guess of the complex F_{norm} value and the K_d. The quadratic error of this least mean square approximation, describes the quality of the fit and is defined as the square root of the sum of squared deviations, divided by the number of composition gradient steps, here 15.

4.2   Implementation of the CGT for label-free MST

Regarding an interaction with two fluorescent molecules, the application of a standard 16-step serial dilution for MST would be problematic because of the high fluorescent range that needs to be covered by the optical system. This explains the need for a different titration strategy, that keeps the overall fluorescence intensity constant. For this, the composition gradient titration, short CGT, was used. This method is based on the method of continuous variation, also known as the Job Plot[100]. This technique keeps the total molar concentration of the two molecules constant, while systematically varying the ratio of both molecules[31]. Therefore, different compositions of both molecules are obtained by titrating the molecules against each other, whereby the overall fluorescence intensity is kept constant. Figure 24 illustrate this titration procedure.
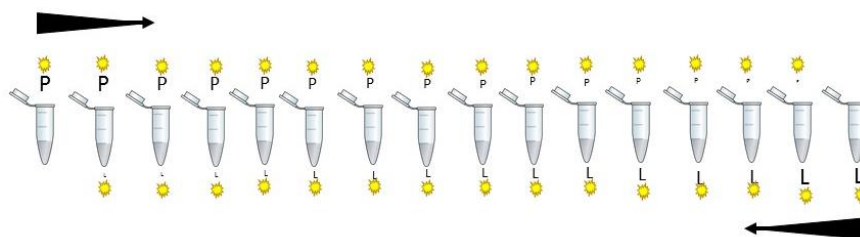
Figure 24: Illustration of CGT. Schematic representation of the composition gradient titration of two fluorescent molecules. The tube on the left only contains molecule P (protein), while the tube on the right only contains molecule L (ligand). The amount of P and L decreases by 1.8 µL per titration step, from left to right and reverse, respectively. The overall fluorescence intensity is kept at one level.

Regarding a CGT series over 15 different compositions with non-interacting molecules, the resulting F_{norm} values will be the sum of the single F_{norm} values, weighted by their relative amounts in the mixture (Figure 25 red). Whereas in case of an interaction, there will be a greater change in the measured F_{norm} values, due to the presence of complexed species with a different F_{norm} value. Schematically, the sample with equal molarity will result in the highest change in F_{norm}, leading to a peak in the measured F_{norm} values (Figure 25 blue). This is illustrated in Figure 25.
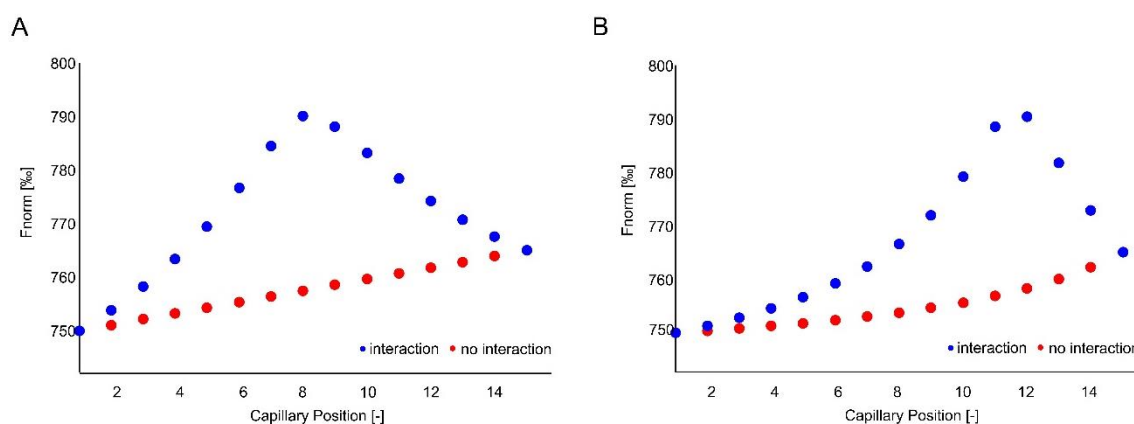
Figure 25: F_{norm} values after CGT. Schematic representation of F_{norm} values after a 15-step CGT of interacting molecules (blue) and of non-interacting (red) molecules, if both molecules start with the same concentration (A) or if one molecule exhibits a 3-fold higher starting concentration, than the other one (B). For this simulation, complex F_{norm} was set to 800, and K_d was fixed at 0.1 nM, with starting concentrations of 10 nM / 10 nM (A) or 10 nM / 30 nM (B).

4.3 Correlation between fluorescence intensity and F_{norm} value

Before preparing whole CGTs, it is important to clarify to which extend the fluorescence counts of a CGT can vary without generating different background fluorescence contributions. Therefore, the correlation between the fluorescence intensity and the F_{norm}

value was investigated. In theory, the F_{norm} value is independent from the fluorescence intensity, but especially in the label-free device different contributions from the background are possible and need to be taken into account.

In order to define the degree of fluorescence, which exhibit the same F_{norm} values, different proteins were titrated against buffer, using a 15-step CGT. Fluorescence was detected using a LED power of 20 % for TEM1 or of 40 % for the IL6 proteins to cover a broad spectrum of fluorescence intensities. F_{norm} values were recorded at high MST power. Results are shown in Figure 26. Here, the grey values are below the recommended fluorescence intensity of about 2500 fluorescence counts and thus the F_{norm} values need to be neglected from the analysis. As illustrated in Figure 26, the F_{norm} values change upon variations in the fluorescence intensity, although the fluorescence is well above the detection minimum. Hence, the fluorescence intensities of a CGT need to be constant throughout the whole titration series, to prevent capillary and buffer artifacts.
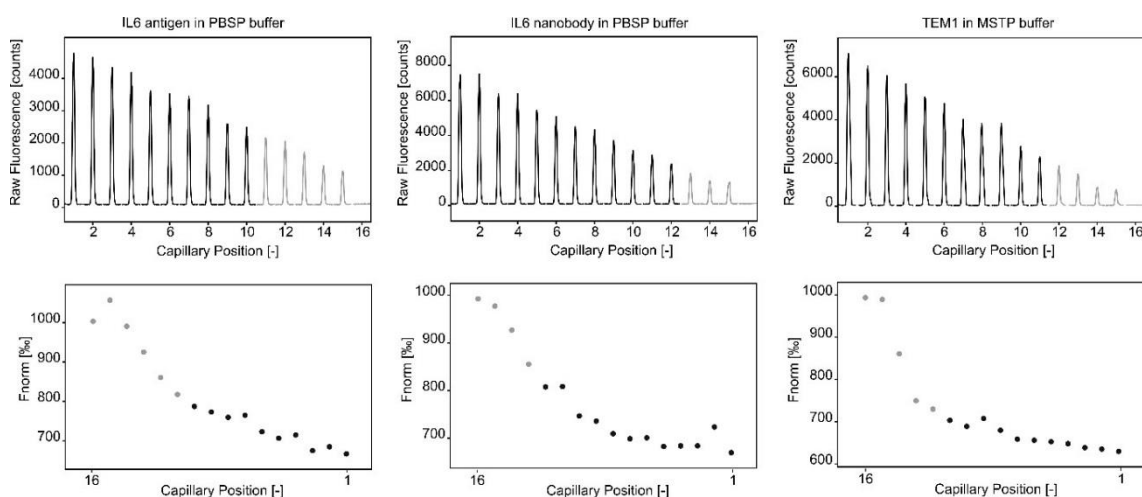
Figure 26: Fluorescence intensity and F_{norm} value. Correlation between fluorescence intensity and F_{norm} value for different proteins, IL6 antigen, IL6 nanobody and TEM1 (from left to right). CGT of proteins against PBSP or MSTP buffer was prepared. Fluorescence intensities (top) over 15 capillaries and corresponding F_{norm} values (bottom) are presented. Capillaries with fluorescence intensities below the detection minimum of 2500 fluorescence counts are illustrated in grey.

4.4 Composition gradient titration for K_d determination

In a first approach the general applicability of the CGT for MST experiments with two fluorescent molecules was tested. I decided to first run an exemplary measurement on the NT.115 device, to prevent any fluorescent artifacts from sample impurities, for which the NT.LabelFree device is highly sensitive. Thus, the first CGT experiment for K_d determination was done using DNA hybridization, with two fluorescent single stranded

DNA molecules. This interaction was chosen as this DNA hybridization is already deeply characterized by MST and is a robust and high affine interaction[101]. Furthermore, DNA hybridization allows for systematic alterations in the affinity, by introducing base pair mismatches[102]. Hence, the template strand was kept the same for all measurements, while the complementary single strand was either a perfect match oligo (Figure 27 black), a single-strand comprising one mismatch (Figure 27 grey) or a single strand with two mismatches (Figure 27 purple). These three interactions were quantified using the CGT approach on the NT.115 instrument. Obtained data out of three independent experiments are shown in Figure 28. Simulated K_d values are illustrated in the Table below. Here, the K_d decreases with higher number of mismatched base pairs. In addition, the height of the curve decreases with lower affinities. In general, obtained curves exhibit low quadratic error of the fits (0.14 – 1.16).

Template:	✹ **5´-Cy5-ATAT TTA CGA TCT GAT CCT T-3´**	
Perfect match:	**3´-**	**AAT GCT AGA CTA GGA A TATA-Cy5-5´** ✹
Mismatch 1:	**3`-**	**AAT GCT ACA CTA GGA A TATA-Cy5-5´** ✹
Mismatch 2:	**3´-**	**AAT GCT ACT CTA GGA A TATA-Cy5-5´** ✹
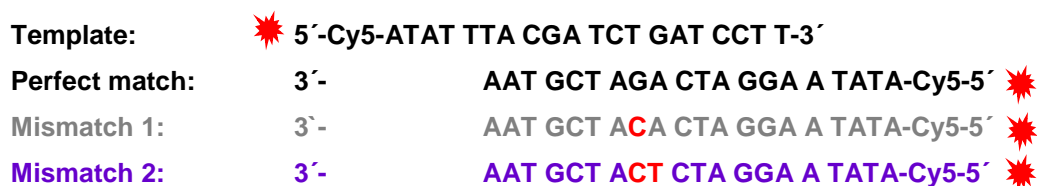
Figure 27: Nucleotide sequences. Nucleotide sequences of single stranded DNA molecules. Template DNA strand and perfect match (black), mismatch 1 (grey) and mismatch 2 (purple). Mismatch nucleotides are shown in red. Location of Cy5-fluorophore is illustrated as red star.

• interaction   • no interaction   • measured
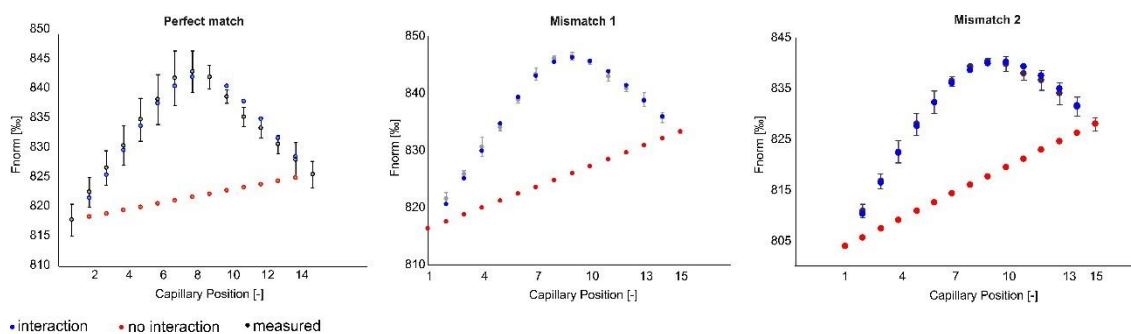
DNA Oligo Pair	Complex F_{norm} value for simulation	Simulated K_d	Quadratic error
Perfect match	874.9	17.9 nM	1.16
Mismatch 1	880.3	19.3 nM	0.14
Mismatch 2	946.8	85.1 nM	0.15

Figure 28: CGT of DNA Hybridization. Composition Gradient titration of two Cy5-labeled DNA strands using the NT.115 device. A) Perfect match DNA oligo pair, mismatch 1 DNA oligo pair and mismatch 2 DNA oligo pair (from left to right). Experiments were carried out using starting concentrations of 100 nM for both single-stranded DNA molecules, 20 % LED and medium MST power. All interactions were measured in triplicates. Values obtained from the least mean square fitting of the corresponding CGT experiments are stated in the table below.

To next validate obtained K_d values, all three interactions were quantified using a standard MST approach, with only the template strand being fluorescently labeled. Results are shown in Figure 29. Here, data exhibit high S/N ratios and a shift in K_d with higher amounts of mismatch base pairs to lower K_d values.
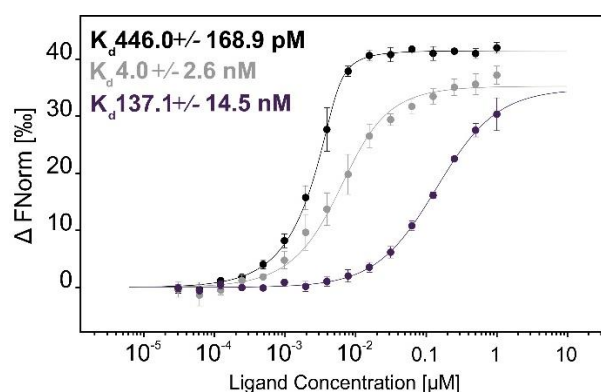
Figure 29: Standard MST for DNA Hybridization. Standard MST analysis of Cy5-labeled template DNA strand against unlabeled perfect match DNA single-strand (black), mismatch 1 DNA single-strand (grey) and mismatch 2 DNA single-strand (purple). Therefore, unlabeled DNA single-strands were diluted from 1 µM highest concentration, while labeled DNA template strand was added at 1 nM constant concentration. MST was carried out at 5 % LED and medium MST power. Data were quantified after 5 sec laser-on time. Interactions were measured in triplicates.

In general, for both, the standard MST data and the CGT results, the K_d decreases with higher amounts of mismatches. However, the K_ds for the perfect match and mismatch 1 are over-estimated in the CGT experiment, while the obtained affinity for mismatch 2 is in good agreement with data gained by standard-MST experiments. Despite slight deviations in the obtained K_d, data showed that the general principle of the CGT can also be applied for MST affinity analysis, in which two fluorescent molecules are present.

In a next step, its applicability for label-free MST experiments was investigated. Therefore, I first analyzed the interaction between IL6 nanobody against its antigen. BIRC75 antigen served as negative control. As the fluorescence intensity of the nanobody was 1.4 times higher than the intensity of the antigens, the concentrations had to be adjusted for the GCT. Thus, 250 nM IL6 nanobody was titrated against 350 nM antigen.

Figure 30 shows the results of the CGT experiments for IL6 nanobody against IL6 antigen (A) and for IL6 nanobody against BIRC75 antigen (B), respectively. As can be seen, the binding event in A can nicely be discriminated from the non-binding event in B. As the fitting algorithm could precisely determine the K_d of this interaction, the quadratic error of the fit is with 0.11 low. This indicates the good quality of the fit. The table below summarizes the data obtained with the least mean square fitting approach.
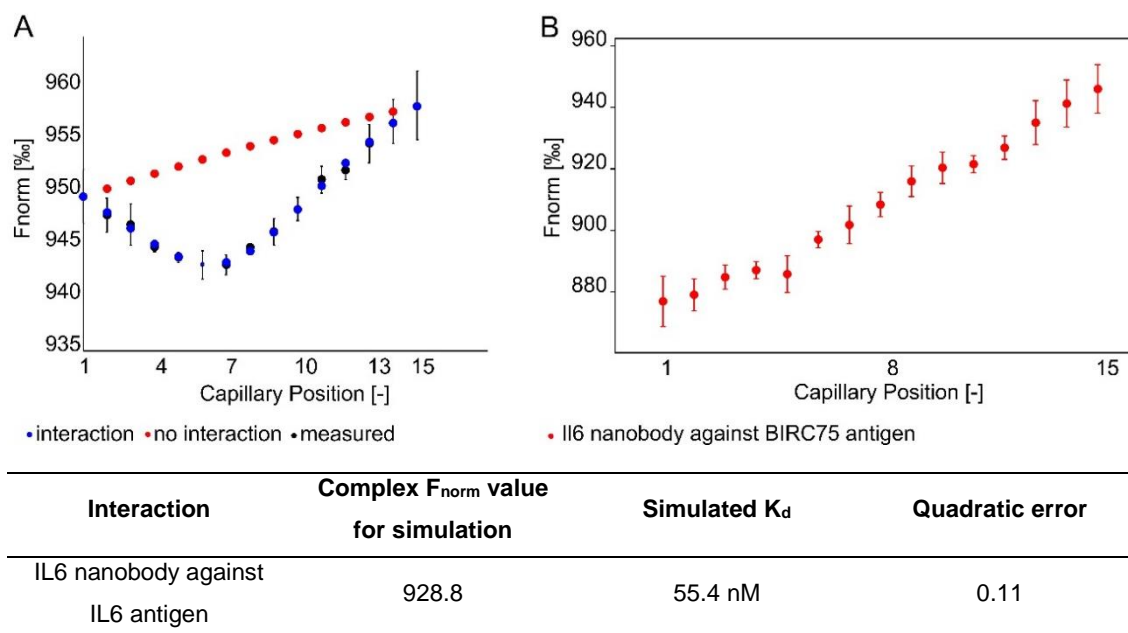
Figure 30: CGT of IL6 nanobody antigen. Results of a CGT of IL6 nanobody against IL6 antigen or against BIRC75 antigen. A) CGT of IL6 nanobody against its specific antigen (A) or against BIRC75 antigen (B) was carried out. F_{norm} values were recorded at 25 % LED and medium MST power. F_{norm} values after 5 sec laser-on time were exported to Microsoft Excel for data analysis. A) Measured F_{norm} values are illustrated in black, with corresponding error bars indicating the standard deviation out of three independent experiments. F_{norm} values in blue represent the fitted values for the case of an interaction and red illustrate the F_{norm} values for the case of no interaction. Experiments were carried out in triplicates. B) Measured F_{norm} values are shown. Error bars indicate the standard deviation from three independent experiments. The Table below summarizes the values that were obtained using the least mean square fit for the interaction between IL6 nanobody against IL6 antigen.

The interaction was verified using standard MST. Therefore, both the IL6 antigen and the BIRC75 antigen were analyzed against IL6 nanobody. Results are shown in Figure 31. Here, only for IL6 nanobody against IL6 antigen binding could be detected, exhibiting a K_d of 26.7 +/- 9.4 nM. Obtained K_d values for IL6 nanobody against its specific antigen were in good agreement with data acquired in the CGT experiment.
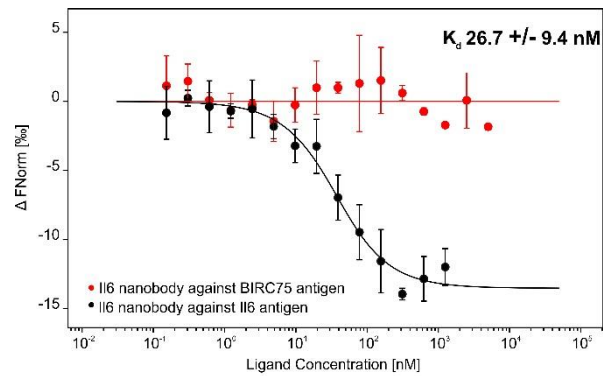
Figure 31: Standard MST for IL6 nanobody antigen. Standard MST experiment of IL6 nanobody against NT647-IL6 antigen (black) or against NT647-BIRC75 antigen (red). MST was carried out using 40 % LED and 80 % MST power. Data were analyzed after 10 sec laser-on time. Experiments were performed in triplicates.

The next interaction that was analyzed was MBP binding protein against MBP. Therefore, 200 nM MBP binding protein was titrated against 100 nM MBP. MST experiments of this CGT approach are illustrated in Figure 32 A. Here, the noise level of triplicate measurements is quite high, however, the fitting algorithm could nicely solve the underlying K_d of 5.5 nM, which is in good agreement with data gained from a standard MST approach (Figure 32 B).
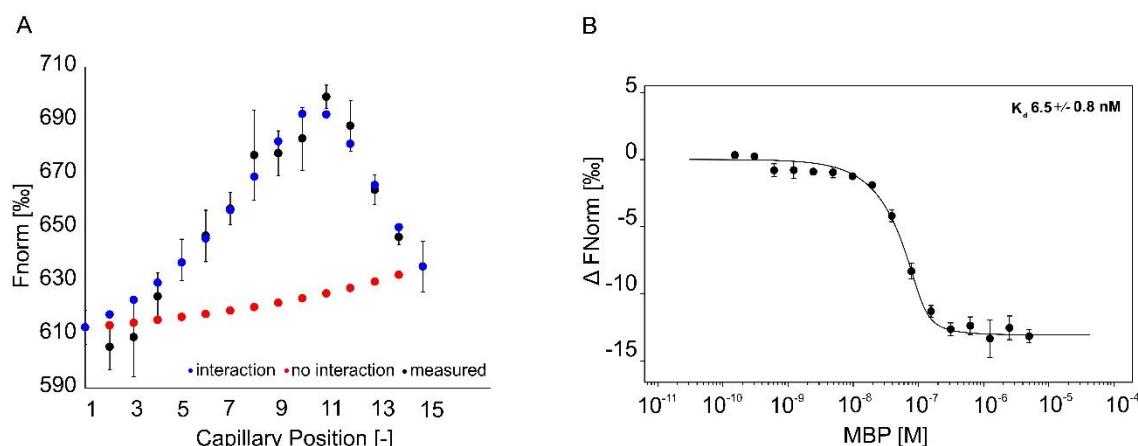
Figure 32: CGT of MBP binding protein against MBP. MBP binding protein against MBP. A) CGT of

Interaction	Complex F_{norm} value for simulation	Simulated K_d	Quadratic error
MBP binding protein against MBP	739.7	5.5 nM	1.69

unlabeled MBP binding protein (200 nM) and unlabeled MBP (100 nM). MST traces were recorded at 10 % LED and high MST power. F_{norm} values after 5 sec laser-on time are shown in black, with corresponding error bars out of three independent experiments. Data for no-interaction (red) and for interaction (blue) were simulated using chemical equilibrium relations and least mean square fitting. Table below summarizes the values obtained by the least mean square fitting of the GCT data. B) Standard MST experiment of RED-tris-NTA labeled MBP binding protein against titrated unlabeled MBP. MST was carried out at 40 % LED and high MST power, n=3. Data were analyzed after 15 sec laser-on time.

4.5 Effect of the starting concentration on the simulation output

Since molecular interactions are commonly concentration-dependent, it is important to estimate the appropriate start concentrations that must be high enough to produce a significant amount of complex and at the same time low enough to still leave some free monomer (avoiding the saturation).

In order to clarify to which extend the starting concentration affects the calculated K_d, two well-characterized model systems were used, MBP binding protein against MBP and perfect match DNA hybridization. For latter interaction, the starting concentration could be reduced to 5 nM, as the sensitivity of the NT.115 device allows such low concentrations for labeled DNA oligos. In contrast, the range of starting concentrations that could be tested in the NT.LabelFree device was limited through the detection system of the device, which does not allow the sample concentrations below 100 nM, for proteins with an average amount of Trp residues, which refers to around 2.2 % of Trp amino acids in a eukaryotic protein[103]. Thus, for the DNA hybridization CGT experiments, 5 nM, 50 nM and 100 nM were used as starting concentrations for both molecules. For the MBP experiments, 100 nM, 200 nM and 400 nM of MBP was titrated against 200 nM,

70

400 nM and 800 nM of MBP binding protein, respectively. Results from those six experimental conditions are illustrated in the Tables below. Here, the higher the starting concentration is, the higher the obtained K_d value will be. However, regarding the complex F_{norm} value, that was selected by the solver tool in Microsoft Excel, variable numbers are obtained, although the reaction that is regarded is the same and thus should exhibit the same complex F_{norm} value. Hence in a next approach, the complex concentration was fixed for all three different starting concentrations and the corresponding K_ds were calculated. With this, obtained K_d values were independent of the starting concentration. In addition, the quadratic error of the fit reduced to an acceptable extent, except for 400 nM MBP against 800 nM MBP binding protein. Here the error of the fit increased from 1.2 to 2.13, which can be explained by the starting concentration being ~160-fold higher than the K_d.

Table 28: Effect of start concentration on CGT data analysis.

Perfect match DNA						
Start concentrations	F_{norm} complex	K_d	Quadratic error	Fix complex F_{norm} to: 816.3	K_d	Quadratic error
100 nM both	821.7	10.2 nM	0.30		6.7 nM	0.32
50 nM both	809.7	4.9 nM	0.26		6.9 nM	0.27
5 nM both	816.3	1.7 nM	0.20		-	-

MBP binding protein against MBP						
Start concentrations	F_{norm} complex	K_d	Quadratic error	Fix complex F_{norm} to: 739.7	K_d	Quadratic error
400 vs 800 nM	794.5	46.7 nM	1.20		4.3 nM	2.13
200 vs 400 nM	783.6	22.6 nM	1.72		3.7 nM	1.79
100 vs 200 nM	739.7	5.5 nM	1.69		-	-

As illustrated by the experiments above, the obtained K_d is independent of the start concentrations, as long as they are close enough to the underlying K_d and as long as the freedom for complex F_{norm} value selection is restricted. Thus, the differences in K_d does not result from different starting concentrations but arise from the too high tolerance in the fitting algorithm for the complex F_{norm} selection.

4.6   Simulation of different experimental conditions for the CGT

As only a small set of interactions does permit the in-depth analysis of limitations of the CGT approach, a set of 1400 experiments was virtually simulated and findings are summarized in the matrix below (Figure 35).

For the simulations, an interaction between two molecules A and B was selected. The F_{norm} value of free A was set to 725 and the F_{norm} value of the complex was set to 825. All other parameters, like the K_d, the starting concentrations, the difference in starting concentration and the difference in the F_{norm} value of free molecules, were systematically varied (Figure 35). This provides a general idea of the limitations of the CGT for label-free MST. In addition, when the conditions are met, this analysis shows the great value of the CGT approach for label-free MST, which enables reliable estimation of the K_d value for a given PPI.

In order to explain the results illustrated in the matrix in Figure 35, exemplary simulations are shown in Figure 33. Here, the interaction between 2 arbitrary molecules A and B with a K_d of 10 nM is analyzed. Figure 33 shows the simulated CGT MST results of this interaction with a starting concentration for both molecules of 10 nM (A), 100 nM (B) or 100 µM (C) for both proteins. The case was chosen in a manner at which both proteins have the same fluorescence intensity at the same concentration. Knowing that the F_{norm} value of free A and the F_{norm} value of the complex AB is 825, F_{norm} values for 10 different K_ds and for the case of no interaction, were simulated (Figure 35). As can be seen, the yellow traces belong to a K_d of 10 nM. In Figure 33, the starting concentrations chosen in B are best suited to analyze the interaction, as the yellow graph can nicely be distinguished from a no-binding flat line and from the more triangular shaped curves of higher affine interactions. In contrast, the starting concentration in A is too low to clearly separate the trace from a non-binding event, especially if higher noise levels of a real experiment are taken into account. Further, the starting concentrations in C are too high. Thus, they are suitable to discriminate between K_d values between 1 µM and 100 µM but are not conclusive for K_d values lower than 1 µM.

All simulations were then divided into the following categories: distinct K_d determination, no distinct K_d determination, distinct K_d determination for high S/N ratio and false negative. This is illustrated in Figure 34. Using this color code, data for all simulations were summarized in a matrix, which is presented in Figure 35.

A

B

C

● 0.01 nM ● 0.1 nM ● 1 nM ● 10 nM ● 100 nM ● 1 µM ● 10 µM ● 100 µM ● 1 mM ● 10 mM ● no interaction
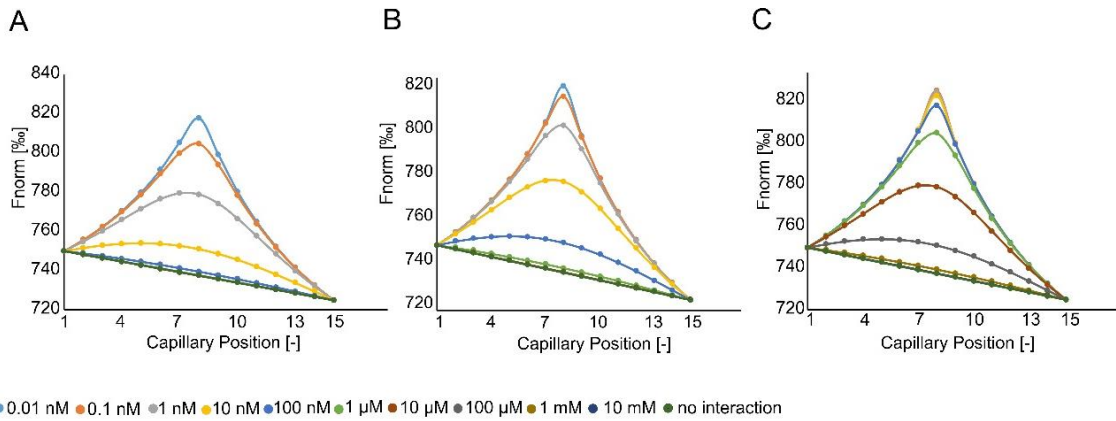
Figure 33: CGT simulation with different start concentrations. Simulations of the same interaction with three different starting concentrations. In this case study, both molecules have the same fluorescence intensity at the same concentration, and thus can be used in equal molar starting concentrations: 10 nM (A), 100 nM (B) or 100 µM (C). F_{norm} value of A was set to 725, F_{norm} value of the complex was fixed to 825. Both proteins exhibit a difference in F_{norm} value of 25 counts. For each K_d, a plot was simulated. Simulated K_d values are illustrated below in the corresponding colors.

● 0.01 nM ● 0.1 nM ● 1 nM ● 10 nM ● 100 nM ● 1 µM ● 10 µM ● 100 µM ● 1 mM ● 10 mM ● no interaction
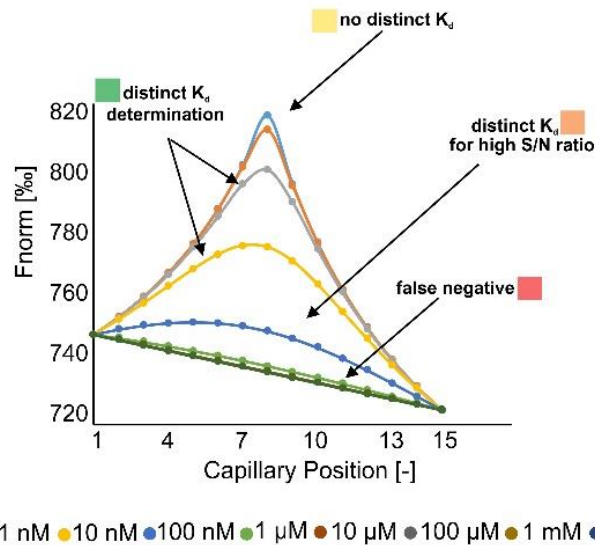
Figure 34: Ranking of simulated data. These simulated CGT data arise from an interaction between A and B, with a starting concentration of 100 nM for both proteins. Different K_d values were simulated, using a fixed F_{norm} value for A of 725 counts and a fixed complex F_{norm} of 845 counts. Each curve is then assigned to "distinct K_d determination", "no distinct K_d determination", distinct K_d for high S/N ratio" or "false negative".
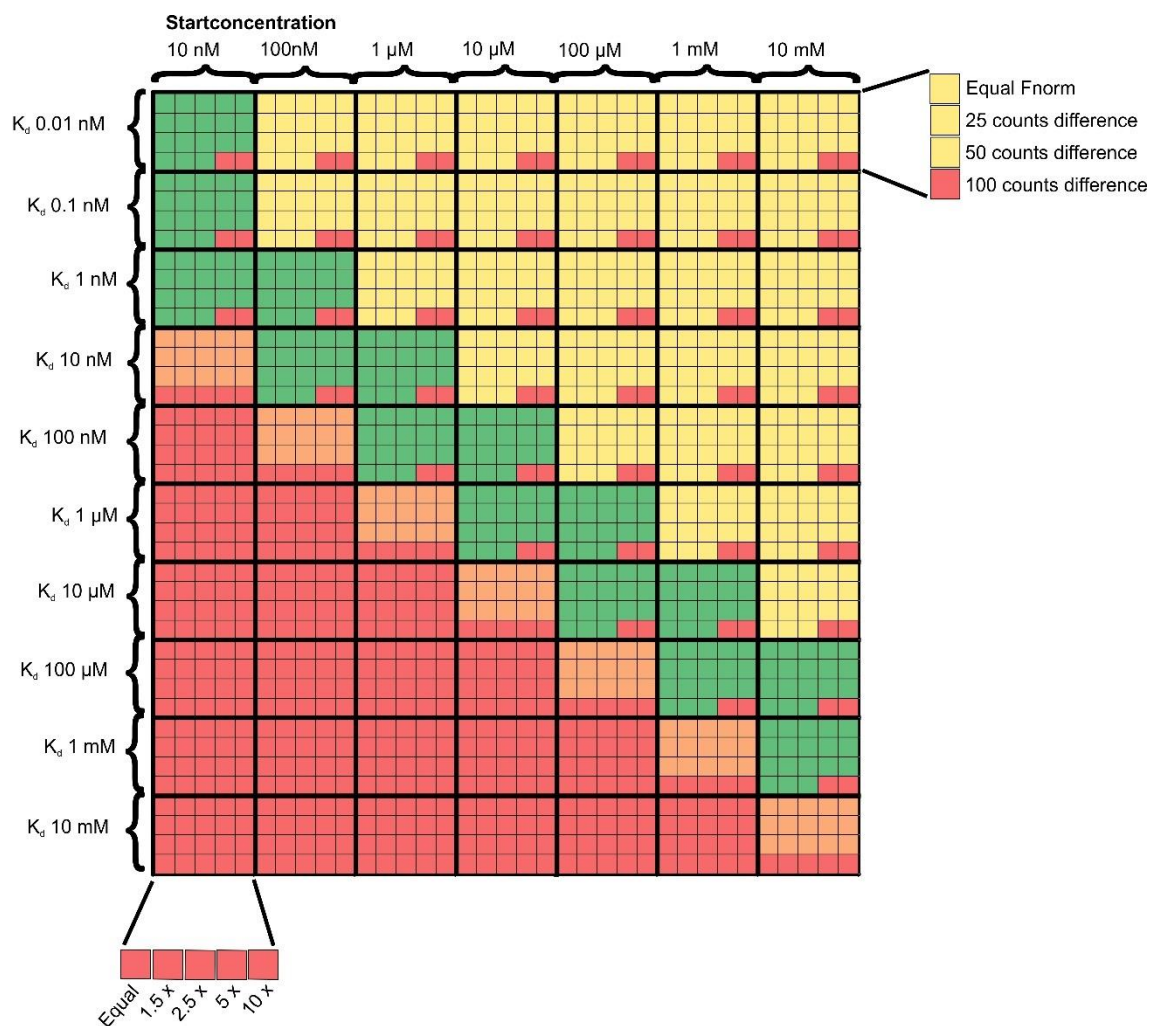
Figure 35: Matrix. Summary of 1400 simulated CGT experiments for a PPI on the NT.LabelFree device. Red squares represent false negative results, orange squares illustrate interactions that can be quantified if the S/N ratio is high enough, yellow squares represent interactions, that can be detected as binding events, but distinct K_d determination is not possible, and green squares represent those interactions, that can be assigned to a distinct K_d value.

As can be seen in the matrix above, the K_d quantification requires starting concentrations of ~10 to 100-fold above the K_d. Whereas lower concentrations can lead to false negative results, as the amount of complex is too low to change the recorded F_{norm} value to a significant extent. However, also too high starting concentrations should be avoided, as the underlying K_d might be overestimated.

In general, red areas also belong to those interactions, in which the starting concentrations of both proteins and / or the F_{norm} values of both vary to a too high extent. Figure 36 illustrates the GCT data for a high affine interaction of 0.1 nM with a 10-fold difference in starting concentration. This high difference shifts the peak maximum to the left, resulting in less data points in which the complex concentration is high enough to lead to a quantifiable change in the F_{norm} value. Furthermore, a difference in the F_{norm}

values between both binding partners will, in addition, lower the amplitude of the binding curve, which is illustrated in Figure 36 B.
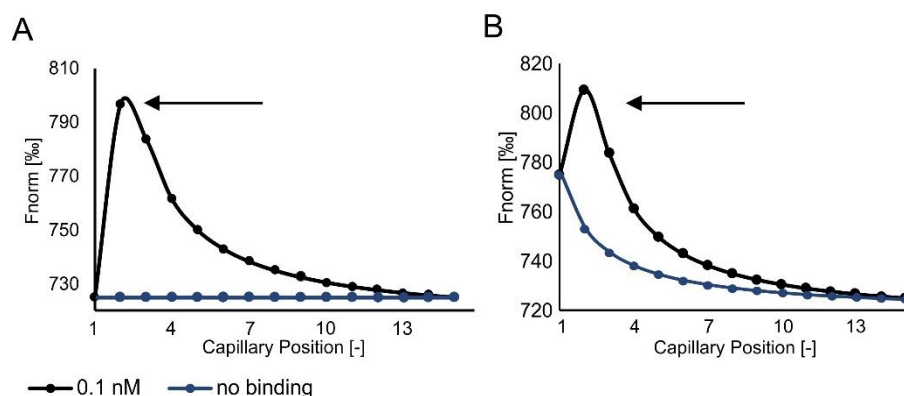
Figure 36: Effect of fluorescence intensity discrepancy for CGT. Simulation of CGT experiment with a K_d of 0.1 nM and with a 10-fold difference in starting concentration. A) Both binding partners exhibit the same F_{norm} value in the unbound state. B) The F_{norm} values of both interacting molecules exhibit a difference of 50 counts.

As a too high difference in start concentrations can make data quantification and interpretation more difficult, as illustrated above, the GCT is not easy to implement for protein – small molecule interaction analysis. The reason is, that the fluorescent intensities of proteins and compounds can be highly different, as shown in Figure 37. Here, p38α and SB203580 were diluted to 1 µM, leading to a difference in fluorescence intensity of 20-fold. Hence, regarding a CGT experiment in which the concentration of both molecules needs to be adjusted to exhibit the same fluorescence intensity, this interaction could not be quantified, as nearly all titration steps would contain high amounts of SB203580 and only negligible amounts of p38α. Thus, the signal of free compound would dominate in mostly all capillaries, leaving no detectable change in F_{norm} due to complex formation. Thus, although the compound shows fluorescence interference in standard MST experiments, as shown in the previous chapter, its fluorescence intensity for a CGT experiment is still too low.
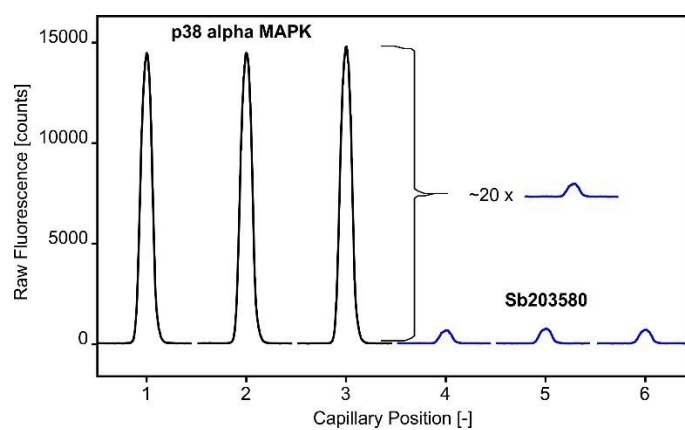
Figure 37: Fluorescence intensity of protein and small molecule. Fluorescence intensity of p38α and its inhibitory compound SB203580. Both samples were diluted to 1 µM using PBSP buffer. Fluorescence was excited with 20 % LED power.

5   Discussion

Several methods are available for the characterization of PPIs, such as surface-immobilization based techniques (e.g. SPR), fluorescence based methods (e.g. FP) or truly label-free and in-solution methods (e.g. MST)[5]. Label-free MST allows determination of dissociation constants in a label-free and free in-solution manner and thus under close-to-native conditions. However, this approach was so far not applicable for the determination of binding affinities between two proteins, as in established methodology only one binding partner may fluoresce in the detection wavelength.

Here, for the first-time, label-free MST was applied for the quantification of PPIs. To that end, the method of continuous variation (Job´s method), that is also applied for PPI analysis in DLS, was used[100]. This method keeps the total concentration of both molecules constant, while systematically varying the relative proportions of both species[31]. To achieve this, both molecules are titrated against each other, which is referred to as CGT. Applying this titration strategy together with data analysis based on a least mean square approximation, first hybridization of two Cy5-labeled DNA single strands was used as proof of concept experiments. Results showed that the CGT is capable for the dissection of hybridization affinities, while K_d values of analyzed single-stranded DNA pairs spanned around three orders of magnitude. In addition, the same experimental workflow allowed for affinity quantification of two nanobody antigen interactions in the NT.LabelFree device, while obtained K_d values were in good agreement with data obtained from standard MST experiments. Further, limitations of the CGT for label-free MST experiments were systematically simulated. This analysis revealed that in theory the CGT and the applied fitting algorithm are capable to determine a broad range of K_d values, ranging from pM to mM affinity. However, the algorithm cannot be applied if the start concentrations and F_{norm} values of both interacting partners simultaneously differ around 10-fold and 100 counts or more, respectively. Thus, it works best at nearly equal F_{norm} values and similar start concentrations of both molecules. Furthermore, the chosen absolute start concentrations were shown to highly affect the data output. Thus, in case the start concentrations are too high, the signal of free monomers will always overlay the signal of formed complex. On the other hand, using too low start concentrations, the amount of formed complex is too low to change the F_{norm} value to a quantifiable extend. In addition, as the detection system of the label-free device requires a minimum protein concentration of ~100 nM, K_ds in the pM range cannot be precisely determined. A general rule of thumb would be to use start concentrations of ~10 to 100-fold above the K_d. This also implies that weak interactions will require a high amount of sample, which might cause protein aggregation or unspecific self-association,

which can lead to false results[104]. In general, a pretest, in which the fluorescence intensities and F_{norm} values of both molecules are recorded, can be used to simulate GCT traces using an assumed K_d and complex F_{norm} value. This will help to estimate the most suitable start concentrations for the binding assay. Notably, the listed considerations and limitations are not MST specific, but arise from the CGT approach in general. Thus, not only MST, but also DLS must deal with the challenge to find the most suitable start concentrations. For DLS, the fitting algorithm works best if both molecules exhibit a similar light scattering signal at the chosen start concentrations, while at the same time the start concentrations are not allowed to vary too much, shifting the peak of equal concentrations too far to one side of the titration curve[97]. In addition, the selection of appropriate start concentrations becomes further challenging if the molar masses of both proteins differ more than 3-fold[97]. Furthermore, not only the ratio of start concentration needs to be carefully selected for DLS measurements, but also the absolute concentrations of both molecules, to allow for precise K_d determination[97]. Thus, the provided matrix of limitations for label-free MST analysis of PPIs, could be transferred to DLS measurements, in which similar restrictions are observed. Nevertheless, CGT experiments in DLS allow for the analysis of binding stoichiometries and protein self-association, which is not the case for label-free MST experiments, as here the fluorescence intensities of all samples need to be the same. Thus, neither information about binding stoichiometries can be obtained, for which equal start concentrations are required, nor a dilution series of one single protein can be analyzed for self-association. Despite these drawbacks in this approach, MST provides advantages like no instrument maintenance and small sample consumption. DLS requires tedious instrument maintenance and higher amounts of sample (around 2 – 20 mL compared to 200 µL for MST, with being concentrated around 10-fold the K_d)[104].

In conclusion, this chapter showed that label-free MST analysis of PPIs is possible if certain conditions are met. In addition, the fitting algorithm turned out to require further improvements regarding its too high tolerance for the selection of the complex F_{norm} value and the K_d. Here, defined limitations for the two unknown parameters need to be implemented. However, as this chapter is focused on the general implementation of the CGT approach for MST affinity analysis and on the definition of its overall limitations, fitting optimization was beyond the scope of this dissertation. Because of observed restrictions of the CGT approach regarding the fluorescence intensities and F_{norm} values of involved molecules, this approach is not generally applicable for the quantification of PPIs. Thus, a need for an approach combining the advantages of the close-to-native conditions and red-shifted fluorescence tag was obvious.

To that end, I implemented a site-specific labeling approach for the analysis of both, protein-protein and protein-small molecule interactions, in which the protein structure is kept in a close-to-native state. This strategy will be presented in the following chapter.

Chapter 3

The following chapter contains results from recently published data in Scientific Reports. Images are used with permission of the respective Journal. The synthesis of analyzed dyes was performed in the working group of Prof. Dr. Jacob Piehler (University Osnabrück), the experiments with the cell lysates were performed by Dr. Katarzyna Walkiewicz and Dr. Christian Kleusch. The cell lysates were a generous gift from Prof. Dr. Yves Müller (University Erlangen/Nürnberg) and from Prof. Dr. Dirk Daelemans (Rega Institute for Medical Research, Leuven).

1 Introduction

Reliable determination of binding affinity between a target molecule and its interaction partner is a critical step in many areas of biological, biochemical and biomedical research and technology. For example, early phases of drug discovery include steps such as target identification and validation, hit discovery and lead optimization. During all of these steps, quantitative characterization of intermolecular interaction affinity is highly necessary to develop novel and effective drugs for therapeutic interventions[105]. MST is a versatile method to quantify binding affinities in solution that is increasingly applied for interaction analysis[4]. In this technique an infrared laser is used to induce a local temperature gradient, causing molecules to migrate out of the heated spot[47]. This phenomenon is termed thermophoresis and strongly depends on molecular properties such as size, charge, hydration shell and conformation[10]. As at least one of these parameters will change upon binding, MST can be used to quantify the interaction and to determine equilibrium dissociation constants (K_d). One notable advantage of MST over other routinely used methods for the quantification of molecular binding events, such as SPR and ITC, is that it can also be used for the determination of K_d values in complex sample matrices like cell lysate and serum[16,47,106]. Although MST measurements can be performed using intrinsic fluorescence of proteins, labeling of the target proteins with a suitable fluorophore is required when using such complex samples. Unfortunately, in routine labeling techniques, the fluorophore is covalently attached to lysine residues using NHS- or to cysteine residues using maleimide chemistry. These labeling strategies are limited to purified proteins and cannot be applied in a mixture of several proteins or in complex biological matrices such as cell lysate or blood serum[107]. The generation of purified protein can be challenging, time-consuming and expensive, sometimes not even

applicable for the protein of interest[108]. Moreover, it is not possible to predict where the fluorophore will bind to the protein. Hence, covalent labeling of a protein with NHS or maleimide conjugated dye can lead to inhomogeneous protein-dye conjugates, some of which might even display destabilization or loss of functionality[109]. Fortunately, in contrast, site-specific protein modification strategies allow structurally and stoichiometrically well-defined labeling with minimal perturbation of structural and functional integrity. Two things that have conquered modern life cell fluorescence imaging are the genetic fusion of fluorescent proteins and enzymes specifically engineered for posttranslational labeling[110], but such relatively large tags are not always desired for quantitative interaction analysis. With the use of bioorthogonal conjugation reactions, labeling of non-purified proteins with high selectivity is possible, allowing rapid and cost-effective labeling[111]. Different site-specific labeling strategies have been proposed and applied, including the co-translational introduction of unnatural or modified amino acids, or labeling via specific amino acid sequences, such as His-tag sequences and tetracysteine motifs[107,112–116]. Among these sequences, the His-tag is the most popular and widely used affinity tag for purification, immobilization or detection of proteins[117–120].

The tris-NTA / His-tag system comprises one of the smallest high-affinity recognition elements known to date[121]. This interaction is based on the capacity of the histidine's imidazole groups to form coordinative bonds with transition metal ions such as Ni(II). Chelators such as nitrilotriacetic acid (NTA)[122] stably bind Ni(II) ions via three oxygen atoms and one nitrogen atom. The two remaining coordination sites of Ni(II) can each bind one histidine moiety of a His-tag (Figure 38)[107], yielding a molecular binding affinity of ~10 µM[123]. Tris-NTA is comprised of three NTA moieties coupled to a cyclic scaffold and thus can simultaneously bind six Histidine residues of a His$_6$-tag, yielding subnanomolar binding affinity and a well-defined 1:1 stoichiometry[123]. Fast, stoichiometric binding of tris-NTA conjugates enabled *in situ* protein labeling of His-tagged proteins[124–127] that was compatible with complex sample matrices including living cells[121,124,128–130]. These unique features make tris-NTA / oligohistidine interaction labeling an attractive candidate for quantitative protein interaction analysis by MST.

In this work, the application of tris-NTA-based labeling of His-tagged proteins for MST measurements will be presented. For this purpose, tris-NTA was conjugated to three different fluorescent dyes: RED (NT647), GREEN (NT547) and BLUE (Oregon Green® 488), providing fluorophores from different spectral regions (Figure 38 A). The dyes NT647 and NT547 are MST-optimized dyes, which are commercially available as NHS

or maleimide derivatives for MST measurements. Our data highlights the versatility, robustness and superiority of the novel tris-NTA labeling approach for MST. Overall, the RED-tris-NTA conjugate (NT647-tris-NTA) arose as the optimal dye conjugate for this approach. This conjugate is characterized by a high affinity for His-tags, a high fluorescence signal and the best S/N ratio of all investigated DYE-tris-NTA conjugates. Owing to its red emission spectrum, it enables reliable measurements in complex biological matrices such as cell lysates, which display substantial autofluorescence in the blue and green part of the spectrum.
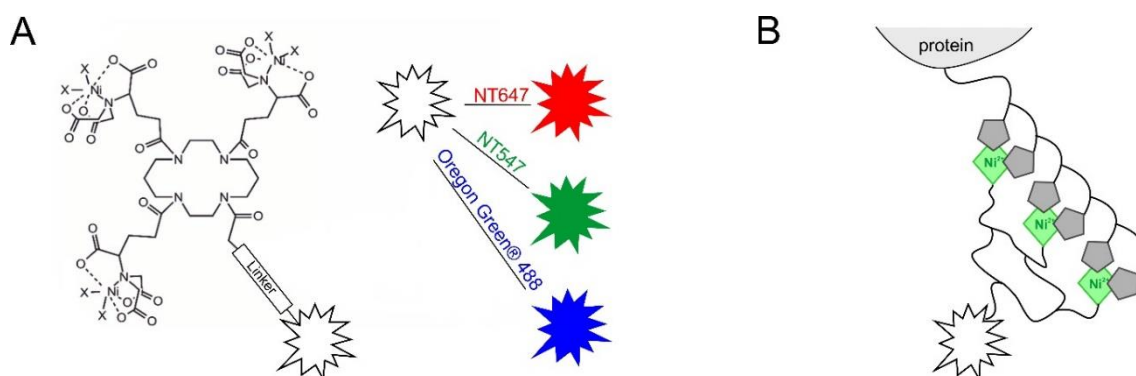
Figure 38: Labeling of His-tagged proteins via tris-NTA conjugates. A) Chemical structure of the tris-NTA moiety conjugated to a fluorophore via a linker. Fluorophores are illustrated on the right: RED (NT647), GREEN (NT547) and BLUE (Oregon Green® 488). B) Schematic representation of DYE-tris-NTA bound to a His-tagged protein. The conjugate is loaded with Ni(II) ions for the site-specific labeling of histidine-tagged proteins. Two remaining coordination sites of the NTA-complexed Ni(II) ions interact with histidine moieties of the protein´s His-tag. Figure adapted with permission from ref.[131], Scientific Reports.

2   Materials

Table 29: Chemicals

Chemicals	Company
AGI-5198 / IDH-C35	Sigma-Aldrich Co. LLC, St. Louis, US, cat. no. SML0839
DMSO (≥ 99.8 %)	Carl Roth GmbH + Co. KG, Karlsruhe, DE, cat.no. A994.2
H_2O AnalaR NORMAPUR®	VWR Prolabo Chemicals, Darmstadt, DE, cat. no. 102927G
HEPES	Carl Roth GmbH + Co. KG, Karlsruhe, DE, cat. no. HN78.2
PD169316	Sigma-Aldrich Co. LLC, St. Louis, US, cat. no. P9248
Roti®-Stock 10x PBS	Carl Roth GmbH + Co. KG, Karlsruhe, DE cat. no. 1058.1
SB203590	Sigma-Aldrich Co. LLC, St. Louis, US, cat. no. S8307
TCEP	Sigma-Aldrich Co. LLC, St. Louis, US, cat. no. C4706
Tween®-20	Sigma-Aldrich Co. LLC, St. Louis, US, cat. no. 93773

Table 30: Fluorophores

Coupling of listed fluorophores to tris-NTA moiety was done by Piehler et al.

Fluorophores	Company
NT547 NHS ester	NanoTemper Technologies GmbH, Munich, DE, cat. no. MO-L002
NT647 NHS ester	NanoTemper Technologies GmbH, Munich, DE, cat. no. MO-L001
Oregon Green® 488 NHS ester	Thermo Fisher Scientific Inc., Wilmington, US, cat. no. 06149

Table 31: Proteins and peptide

Proteins and peptide	Company
His_6-MBP-binding protein	ChromoTek GmbH, Munich, DE, cat. no. mbt-250
His_6-p38α MAPK	CRELUX GmbH, Munich, DE
His_6-peptide	Apexbt Technology LLC, Houston, US, cat. no. A6006.
His_6-pUL53 containing *E.coli* cell lysate and purified His_6-pUL53	Was kindly provided by Sebastian Weigert, Prof. Dr. Yves Muller, Division of Biotechnology, Department of Biology Friedrich-Alexander University Erlangen/Nuremberg, Erlangen, DE

| MBP | antibodies-online GmbH, Atlanta, US, cat. no. MBP0801 |
| p38α-mNeonGreen-His$_6$ and mNeonGreen-His$_6$ in HeLa cell lysate | Was kindly provided by Prof. Dr. Dirk Daelemans, Dr. Thomas Vercruysse, KU Leuven, Department of Immunology and Microbiology, Laboratory of Virology and Chemotherapy, Rega Institute for Medical Research, Leuven, BEL |

Table 32: Devices

Devices	Company
Monolith NT.115 (RED/GREEN or BLUE)	NanoTemper Technologies GmbH, Munich, DE

Table 33: Further material

Material	Company
Monolith NT.115 Capillaries	NanoTemper Technologies GmbH, Munich, DE
Monolith NT.115 MST Premium Capillaries	NanoTemper Technologies GmbH, Munich, DE
SPROUT® MINI CENTRIFUGE 12V	Heathrow Scientific®, Vernon Hills, US

Table 34: Software

Software	Company
MO.AffinityAnalysis_x86_2.2.7.6056	NanoTemper Technologies GmbH, Munich, DE
MO.Control_x86_1.5.3.6096	NanoTemper Technologies GmbH, Munich, DE

3   Methods

3.1   Synthesis and preparation of tris-NTA conjugated fluorophores

OregonGreen® 488-tris-NTA was prepared as previously published[124]. NT647 and NT547-tris-NTA conjugates were synthesized and characterized as following: tris-NTA modified by an aminocaproic acid (tris-NTA, MW: 1048 g/mol) was synthesized as previously described[123] to yield tris-NTA-ACA. 5.9 mg tris-NTA-ACA was dissolved in 100 µL dry DMF, followed by addition of 12 µL EDIPA. 3.0 mg NHS ester of fluorophores was separately dissolved in 100 µL dry DMF. Solutions were mixed and stirred overnight (20 h) at room temperature under the protection of N_2 atmosphere in darkness. After addition of 100 µL H_2O, the reaction mixture was continuously stirred for 1 h to quench unreacted NHS groups. The mixture was diluted in 20 mL of 0.1 % TFA / water and was loaded on a C_{18} reverse phase HPLC column (Vydac 218TP, 250 mm × 4.6 mm) for purification using a 0–70 % acetonitrile gradient in 0.1 % TFA / water (1/6 of the reaction per run). The DYE-tris-NTA conjugates were eluted at ~45 % acetonitrile. The purified DYE-tris-NTA fractions of each dye were pooled together and lyophilized as blueish, orange or pink powder and stored at -20 °C. Obtained products were verified by the MS-ESI analyses.

For the nickel loading, the obtained DYE-tris-NTA- were dissolved in 20 to 50 mL of 10 mM HEPES buffer pH 7.5 to a final concentration of 0.1 mM or less. $NiCl_2$, in a final concentration of 5 mM was added to this solution for loading Ni(II) ions on the NTA groups. After 15 min, the solution was loaded onto a 1 mL anion-exchange column (Hitrap Q, GE Healthcare) and eluted with a gradient of 0-600 mM sodium chloride in 10 mM HEPES buffer, pH 7.5. The Ni(II)-loaded tris-NTA-fluorophores were eluted at ~300 mM sodium chloride. For each dye, the fractions from the elution peak were combined, and the concentrations were determined photometrically at 647 nm and 547 nm using an extinction coefficient of 250000 $M^{-1}cm^{-1}$ and 150000 $M^{-1}cm^{-1}$, respectively. The final products were aliquoted in black Eppendorf tubes and stored at -20 °C.

3.2   Determination of DYE-tris-NTA binding affinity for oligohistidine tags

His_6-peptide or His_6-p38α (expression construct CJA3) were diluted in PBST buffer to a final concentration of 10 µM and 2 µM, respectively. This solution was used for a 16-step serial dilution in PBST buffer (137 mM NaCl, 2.5 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4, pH 7.4, 0.05 % Tween-20) with 10 µL volume in each sample. Next, 10 µL of

50 nM dye, dissolved in PBST buffer, was added to all vials of the serial dilution. Samples were mixed by pipetting up and down and the reaction was incubated for 30 min at room temperature in the dark before samples were loaded into Monolith NT.115 Capillaries. Samples were then transferred into the Monolith NT.115 device and MST experiments were carried out at 40 % (RED) / 20 % (BLUE) / 100 % (GREEN) LED and medium MST power for the His$_6$-peptide measurements and at 40 % (RED) / 60 % (BLUE) / 100 % (GREEN) LED and medium MST power for the His$_6$-p38α studies.

3.3 Labeling and MST measurements of purified His$_6$-tagged proteins

Proteins (His$_6$-p38α and His$_6$-MBP-binding protein were diluted to 200 nM in PBST buffer (137 mM NaCl, 2.5 mM KCl, 10 mM Na$_2$HPO$_4$, 2 mM KH$_2$PO$_4$, pH 7.4, 0.05 % Tween-20). Tris-NTA dyes were diluted in PBST buffer to a final concentration of 100 nM. 100 µL of protein was then mixed with 100 µL of each dye separately and the reaction mixtures were incubated for 30 min at room temperature in the dark.

Ligand dilution series: Small molecule inhibitors were stored in 100 % DMSO at -20 °C. For the dilution series, a 10 µM solution of PD169316 or of AGI-5198 (IDH-C35) was prepared using PBST buffer (2 % final DMSO concentration). This stock solution was used for the preparation of a 16-step serial dilution in PBST buffer, supplemented with 2 % DMSO, with a final volume of 10 µL in each vial of the dilution series. For the protein-protein interaction, 5 µM of MBP diluted in PBST buffer was used as the highest ligand concentration of the 16-step dilution series, with a final volume of 10 µL in each reaction tube.

Then 10 µL of 100 nM RED / GREEN or BLUE labeled protein (p38α or MBP-binding protein) was added to all 16 vials and samples were mixed by pipetting up and down. Reactions were incubated for 30 min at room temperature away from light and then loaded into Monolith NT.115 Capillaries. Using the Monolith NT.115 device, MST was carried out at 20 % (RED) / 60 % (BLUE) / 100 % (GREEN) LED and high MST power for p38α, and at 40 % (RED) / 60 % (BLUE) / 100 % (GREEN) LED and high MST power for MBP-binding protein.

3.4 Labeling and MST measurements of oligohistidine-tagged proteins in crude cell lysate

3.4.1 p38α against SB203580

The p38α protein sequence was obtained through reverse transcription on mRNA from A549 cells. Using In-Fusion Cloning technology (Takara Bio USA, Inc.) this p38α coding sequence was cloned into a pcDNA3.1 mammalian expression vector behind a CMV promotor and separated from the mNeongreen-His$_6$-tag by the linker sequence ESGSGS. A pcDNA3.1 vector coding for only mNeongreen-His$_6$ was used as a control. These two plasmids expressing mNeongreen-His$_6$ with and without the p38α sequence were transfected into $3*10^6$ HeLa cells using separate T-75 flasks. Cells were grown for 24 h reaching approximately $10*10^6$ cells. Cells were pelleted by centrifugation and resuspended in 1 mL PBST buffer, supplemented with protease inhibitors. At this step, the cells were disrupted using a Dounce homogenizer and centrifuged again at 14 000 x g for 30 min at 4 °C to remove cell debris. Obtained supernatant was diluted 1:10 in PBST, supplemented with protease inhibitors.

Concentration of His$_6$-tagged protein in cell lysate was determined by the MST experiment as described in the Supplemental information, Figure 4. The labeling of p38α-mNeonGreen-His$_6$ and mNeonGreen-His$_6$ in HeLa cells was carried out by mixing 100 µL of about 100 nM p38α-mNeonGreen-His$_6$ or mNeonGreen-His$_6$ with 100 µL of 100 nM NT647-tris-NTA dye in PBST buffer. Reaction mixture was incubated for 30 min at room temperature.

For the MST binding experiment, the stock solution of SB203580 (stored at 2.65 mM in 100 % DMSO at -20 °C) was diluted 1:50 in PBST, reaching a concentration of 53 µM with 2 % DMSO. This solution was used for a 1:1 serial dilution using 16 dilution steps and a final volume of 10 µL for each point of the dilution series. Afterward 10 µL cell lysate was added to all steps of the dilution series, giving a final ligand concentration of 26.5 µM with 1 % DMSO. Reaction was incubated for 30 min at room temperature, centrifuged at 14 000 g for 10 min at 4 °C and loaded into Monolith NT.115 MST Premium Capillaries. MST experiment was carried out using 100 % or 20 % LED power for the p38α containing sample and for the negative control, respectively, and high MST power for the NT.115 RED instrument. For the NT.115 blue device, 20 % LED and high MST power was used.

3.4.2 pUL53 against pUL50

For protein expression, a plasmid encoding a His$_6$-tagged protein variant of pUL53 (residues 50 to 292 of human cytomegalovirus ORF-UL53 1-376)[132] was transformed into BL21(DE3) cells and grown in LB medium in the presence of 100 mg/mL ampicillin and 32 mg/mL kanamycin at 33 °C until OD$_{600}$ of 0.4. When the required OD was reached, 0.25 mM isopropyl -D-thiogalactopyranoside was added to induce protein expression. The culture was further incubated overnight at 20 °C. Cells were harvested by centrifugation, disrupted by sonication and resuspended in lysis buffer (50 mM phosphate buffer, pH 7.4, 300 mM NaCl) containing protease inhibitors, lysozyme, and DNase.

For affinity determination of His$_6$-pUL35 against RED-tris-NTA, pUL53-containing cell lysate was diluted 1:10 in PBST buffer and a 16-step serial dilution was prepared. RED-tris-NTA was then added to all dilution steps with a final concentration of 25 nM. Samples were incubated for 30 min at room temperature in the dark before they were loaded into Monolith NT.115 MST Premium Capillaries and loaded into the Monolith NT.115 device. MST experiment was carried out at 40 % LED and high MST power.

Labeling of His$_6$-pUL53 in *E. coli* lysate was carried out by diluting the lysate 1:10 in PBST buffer and adding RED-tris-NTA dye at a final concentration of 50 nM. The mixture was incubated for 30 min at room temperature. For the labeling of purified His$_6$-pUL53, the protein was first purified from the *E.coli* cell lysate *via* a Ni-NTA affinity chromatography followed by a size exclusion chromatography step. After purification, His$_6$-pUL53 was labeled by mixing100 µL of a 200 nM protein solution with 100 µL of 50 nM RED-tris-NTA using PBST as reaction buffer. Mixture was incubated for 30 min at room temperature in the dark.

For affinity analysis of purified and non-purified RED-tris-NTA His$_6$-pUL53 against pUL50 (obtained as previously described)[132], HEPES buffer (200 mM, 25 mM HEPES, 1 mM TCEP, pH 8.0) was used. The highest ligand concentration in the 16-step serial dilution series was 1 µM, with 10 µL volume in each titration step. 10 µL labeled protein was then added to all dilutions at a final concentration of 100 nM. Samples were mixed and loaded in Monolith NT.115 MST Premium Capillaries. MST experiments were carried out at 40 % LED and high MST power.

3.5   Data acquisition and analysis

The data were acquired with MO.Control 1.5.3 (NanoTemper Technologies GmbH). Recorded data were analyzed with MO.Affinity Analysis 2.2.7 (NanoTemper Technologies GmbH). The MST on-time yielding the highest S/N ratio was used for the K_d determination. The data were fitted using a K_d fit model that describes a molecular interaction with a 1:1 stoichiometry according to the law of mass action. The K_d is estimated by fitting the equation 19:

$$f(c) = \text{Unbound} + (\text{Bound} - \text{Unbound}) \times \frac{c + c_{target} + K_d - \sqrt{(c + c_{target} + K_d)^2 - 4cc_{target}}}{2c_{target}} \qquad (19)$$

Where *f(c)* is the fraction bound at a given ligand concentration *c, Unbound* is the F_{norm} signal of the target, *Bound* is the F_{norm} signal of the complex, K_d is the dissociation constant or binding affinity, and the *c_{target}* is the final concentration of target in the assay.

4   Results

For the MST experiments in this study, the MST optimized dyes NT647 and NT547 were conjugated to tris-NTA. For the blue channel, the dye Oregon Green® 488[124] was chosen. BLUE (OregonGreen® 488) is a derivate of fluorescein, the dyes RED (NT647) and GREEN (NT547) have distinct structures. The synthesis of DYE-tris-NTA constructs and the loading of the dyes with Ni(II) ions was performed according to previously described protocols[123,124]. Concentrations of DYE-tris-NTA conjugates were determined photometrically and the dyes were stored at -20 °C until further use.

4.1   The affinity of DYE-tris-NTA for oligohistidine sequences

A high affinity of DYE-tris-NTA for oligohistidines is a prerequisite for labeling of proteins for the quantification of intermolecular interactions with MST. I therefore determined the affinity between the DYE-tris-NTA and two oligohistidine sequences: a His_6-peptide and the MAP kinase p38α fused to an N-terminal His_6-tag. As expected[123], BLUE-, GREEN- and RED-tris-NTA displayed high affinity for the His_6-peptide (6.7 ± 4.1 nM, 4.4 ± 3.7 nM and 3.8 ± 0.5 nM, respectively, Figure 39). Comparable affinities were measured for the His_6-tag fused to the N-terminus of p38α, yielding K_d values of 2.7 ±1.7 nM for BLUE-tris-NTA, 6.3 ± 1.7 nM for GREEN-tris-NTA and 2.1 ± 0.8 nM for RED-tris-NTA, respectively (Figure 40). These binding affinities are in excellent agreement with previously published K_d values of tris-NTA / His-tag interaction[123]. Slight differences in the K_d values between His_6-peptide and His_6-p38α can be explained with differences in the accessibility and the electrostatics due to structural characteristics of the protein and the fluorophores.

Notably, the binding of BLUE-tris-NTA and GREEN-tris-NTA to His_6-p38α resulted in a ligand-induced fluorescence change. To exclude the possibility of nonspecific interactions between these dyes and the ligand, I performed an EDTA / His_6-peptide (ECP) test, which quantifies ligand-induced fluorescence changes while using DYE-tris-NTA labeling. It consists of two subtests that must be performed to unambiguously distinguish between fluorescence changes caused by interaction and those caused by non-specific effects. In the case of His-tag labeling, non-specific effects can be caused by the interaction of a ligand with the His-tag bound tris-NTA dye (His_6-peptide test) or by ligand-induced aggregation or adsorption to labware (EDTA test). The nonspecific interaction between BLUE-tris-NTA and GREEN-tris-NTA and the ligand was excluded based on this test (Supplementary data Figure 3). Hence, the fluorescence signal was

used for evaluation of binding data (Figure 40 B and C). In addition, for the p38α protein lacking His$_6$-tag, no binding was detected with any DYE-tris-NTA (Figure 40).
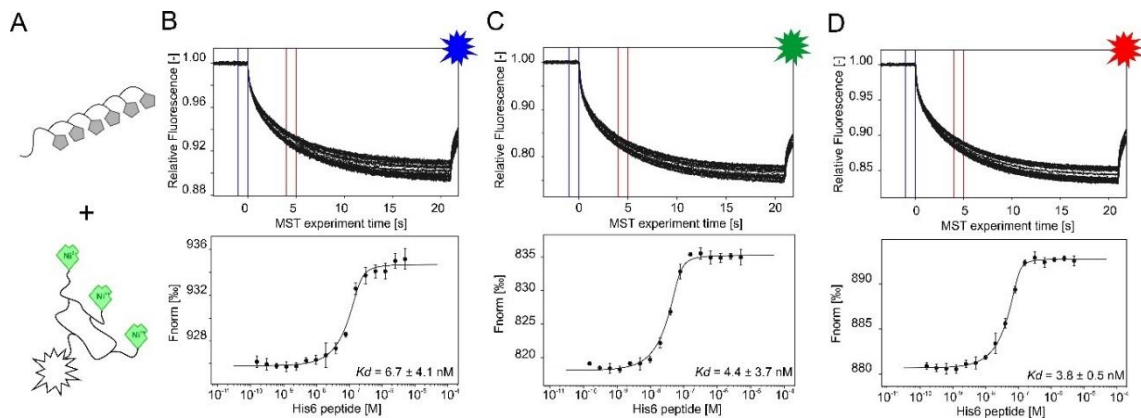
Figure 39: MST interaction analysis of His6-peptide against different tris-NTA fluorophores. A) Schematic representation of DYE-tris-NTA interaction with His$_6$-peptide. B) C) D) MST traces (top) and dose-response curves (bottom) of His$_6$-peptide titrated against tris-NTA conjugated dyes. B) BLUE-tris-NTA C) GREEN-tris-NTA and D) RED-tris-NTA. 25 nM of DYE-tris-NTA was added to a 16-step serial dilution of His$_6$-peptide. Mean values of three independent experiments are shown. Error bars indicate the standard deviation. MST experiments were carried out at medium MST power at 25 °C. LED power was set to 20 % (B), 100 % (C) and 40 % (D). The resulting dose-response curves were fitted to a one-site binding model to extract K_d values; the standard deviation was calculated using the K_d values from each independent experiment. F_{norm} = normalized fluorescence. Figure adapted with permission from ref.[131], Scientific Reports.
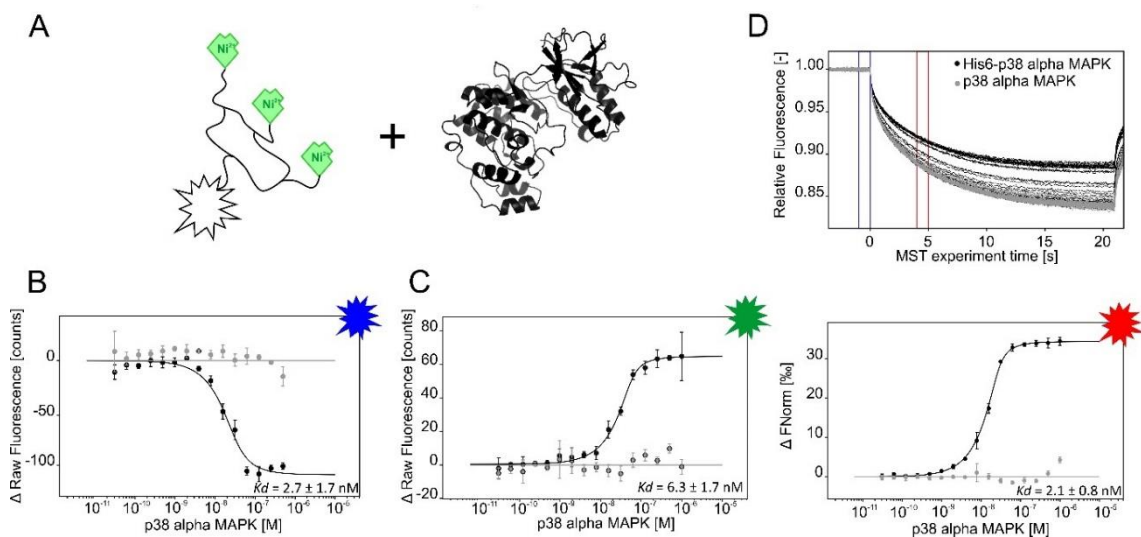
Figure 40: MST analysis of tris-NTA fluorophores interacting with p38α. A) Schematic representation of tris-NTA-fluorophores and p38α (PDB: 1A9U). B) C) Dose-response curves for BLUE-tris-NTA and GREEN-tris-NTA against His$_6$-p38α (black) and p38α (grey) (n=2). D) MST traces (top) and dose-response curve (bottom) of His$_6$-p38α (black) and p38α (grey) towards RED-tris-NTA (n=3). All resulting dose-response curves were fitted to a one-site binding model to obtain K_d values. Error bars indicate the standard deviation. MST experiments were performed at a LED power of 60 % (A), 100 % (C) and 40 % (D) and at medium MST power. F_{norm} = normalized fluorescence. Figure adapted with permission from ref.[131], Scientific Reports.

4.2 Stability of DYE-tris-NTA binding to His-tags

The Ni(II) / His-tag interaction is reversible upon addition of competitor substances such as histidine or imidazole. This feature is useful in some contexts since it enables removal of immobilized molecules from microarray surfaces or elution of purified proteins from column-chromatography[133,134]. However, in the context of protein labeling for binding studies, disruption of the Ni(II) / His-tag interaction should be avoided, since it results in dissociation of the dye. To investigate potentially interfering buffer components, I systematically screened a set of common additives with respect to their effects on tris-NTA labeling. To this end, His_6-peptide was titrated against RED-tris-NTA, while varying the concentration of different additives in the assay buffer PBST. These additives are listed in Supplementary Table 20, together with the highest concentration tested and their maximum tolerable concentration for the tris-NTA labeling approach. The maximum allowed assay concentration was defined as the highest concentration that did not alter the K_d value, whereas a slight decrease in the S/N ratio or in the binding amplitude was tolerated. In general, chelating agents such as ethylenediaminetetraacetic acid (EDTA) and ethylene glycol-bis(aminoethyl ether) (EGTA) or the ionic detergent sodium dodecyl sulfate (SDS) should be avoided. In addition, Tris-based assay buffers are known to exhibit some ion-complexing capacities[135,136], therefore a caution and additional pretests are recommended when used for DYE-tris-NTA labeling. The use of divalent metal ions as Zn(II), Co(II) and Ni(II) as additives is not recommended because they interfere with the labeling. Bovine serum albumin (BSA) and various proteins without His-tag showed no interference at any of the tested concentrations. In general, the tris-NTA labeling method turned out to be highly robust toward buffer additives, even regarding competitor substances such as imidazole or histidine. Here, only concentrations higher than 1 mM showed interference with the labeling reaction. Additionally, reducing agents such as tris(2-carboxyethyl)phosphine hydrochloride (TCEP) and dithiothreitol (DTT) can be used in the labeling buffer as well. This insensitivity and robustness of DYE-tris-NTA thus mostly allows labeling of proteins directly in their storage buffers, without the need for buffer exchange.

4.3 Determination of ligand binding affinity using DYE-tris-NTA labeled target proteins

A high affinity of DYE-tris-NTA for His-tags is a prerequisite for efficient stoichiometric labeling of His-tagged proteins for MST measurements. Because of all our DYE-tris-NTA candidates showed an affinity in the low nM range, I proceeded with the determination

of ligand binding affinity using DYE-tris-NTA labeled target proteins. For a protein-small molecule interaction, His_6-p38α was labeled with all three DYE-tris-NTA candidates separately. To ensure that virtually all dye is bound to the protein, I labeled the protein at a ratio of 1:2 (dye:protein) and incubated the protein / dye mixture for 30 min at room temperature to enable complete binding of the DYE-tris-NTA to the protein. As evident from the experiment depicted in Figure 40, when I use 25 nM of DYE-tris-NTA, the binding of the DYE-tris-NTA reaches the saturation at the concentration of p38α of about 50 nM. With further increase of the protein concentration, no additional increase in the binding of the dye is observed. This means that all dye is bound to the protein at a ratio of about 1:2. Based on this finding, efficient labeling was achieved and no separation of unbound dye was required. I proceeded immediately to the next step and added the labeled target protein to a dilution series of PD169316, a known selective inhibitor of p38α[137,138]. The ligand AGI-5198 (IDH-C35), a potent inhibitor of isocitrate dehydrogenase 1, served as a negative control. For PD169316 I measured K_d values of 16.7 ± 1.2 nM, 35 ± 5 nM and 24 ± 9 nM for BLUE-, GREEN- and RED-tris-NTA, respectively, which is in accordance with published values[51] (Figure 41).
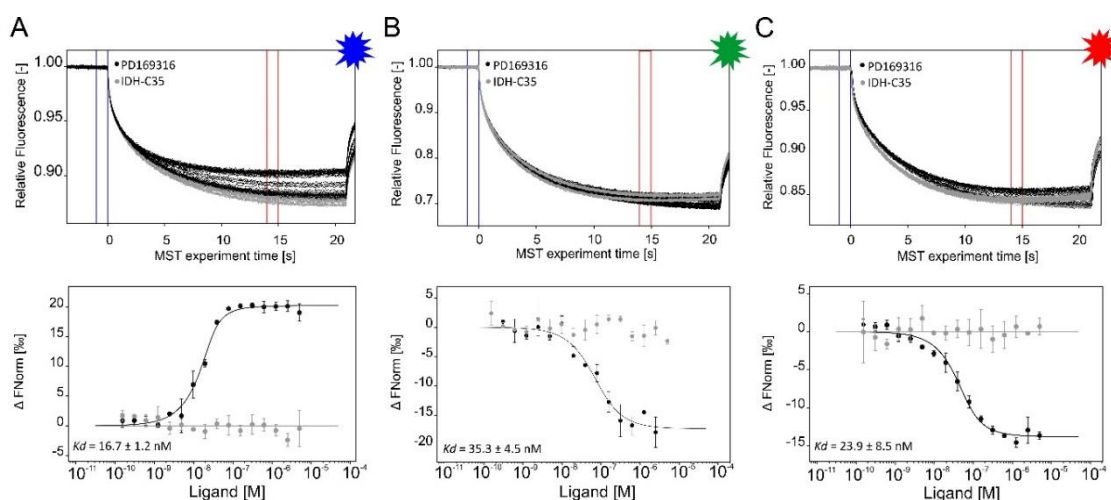
Figure 41: The interaction of p38α protein with small molecule inhibitors. MST traces (top) and dose-response-curves (bottom) for labeled His_6-p38α (BLUE-tris-NTA (B), GREEN-tris-NTA (C) and RED-tris-NTA (A)) against PD169316 (black) or IDH-C35 (grey). The resulting dose-response curves were fitted to a one-site binding model for K_d determination. All measurements were done in triplicates, error bars indicate the standard deviation. MST experiments were performed at high MST power and LED power of 60 %,100 % and 20 % (from left to right). F_{norm} = normalized fluorescence. Figure adapted with permission from ref.[131], Scientific Reports.

As a second example, the interaction between maltose-binding-protein (MBP) and MBP-binding protein was analyzed. This 15 kDa V_HH binds MBP of *E.coli* with high affinity (ChromoTek GmbH, unpublished data). For MST affinity analysis, His_6-tagged MBP-

binding protein was labeled with all three DYE-tris-NTA conjugates separately and added to a serial dilution of MBP in PBST buffer. The K_d values measured were 6 ± 2 nM for BLUE-tris-NTA, 5 ± 4 nM for GREEN-tris-NTA and 7 ± 1 nM for RED-tris-NTA (Figure 42). No binding was detected for labeled His_6-peptide against titrated MBP, which underscores the high specificity of the MBP-binding protein for its ligand MBP. Among all DYE-tris-NTA candidates, RED-tris-NTA showed the best S/N ratio.
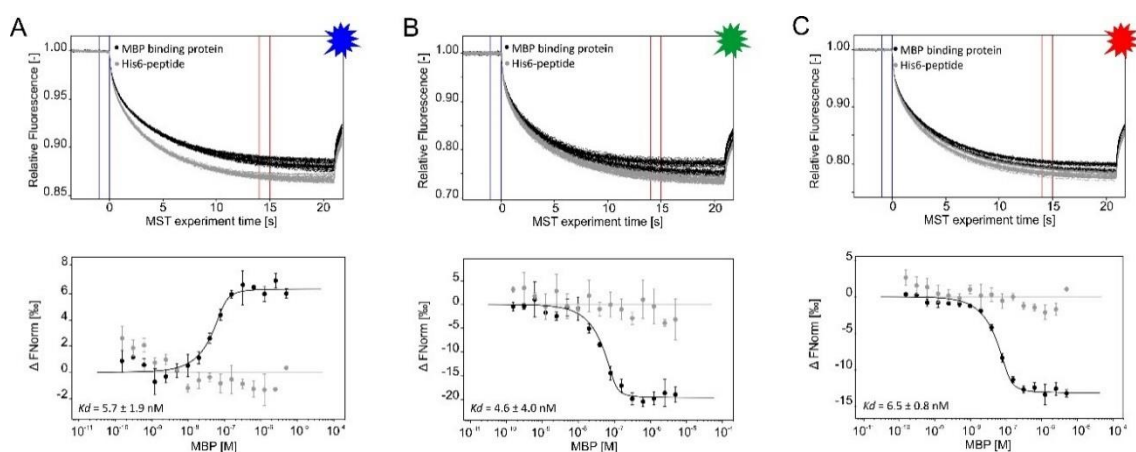
Figure 42: MST measurements of MBP towards MBP binding protein. MST traces (top) and dose-response curves (bottom) of MBP towards BLUE-tris-NTA-MBP-binding protein (A), GREEN-tris-NTA-MBP-binding protein (B) and RED-tris-NTA-MBP-binding protein (C). The resulting dose-response curves were fitted to a one-site binding model to extract K_d values. All experiments were done in triplicates. Error bars indicate the standard deviation. MST experiments were performed at a LED power of 60 % (A), 100 % (B) and 40 % (C) and at high MST power. F_{norm} = normalized fluorescence. Figure adapted with permission from ref.[131], Scientific Reports.

As is evident from Figure 40 to Figure 42, differences in MST binding curve direction can be observed between the three DYE-tris-NTA candidates. For RED-tris-NTA and GREEN-tris-NTA, the unbound state of the protein exhibits the smallest changes in fluorescence; while for BLUE-tris-NTA, the bound state of the protein exhibits the smallest changes in the fluorescence. As mentioned previously, RED- and GREEN-tris-NTA are highly similar regarding their chemical structures, whereas BLUE-tris-NTA belongs to a different family of dyes. These molecular properties likely result in different MST signals of labeled proteins in the unbound and bound state and thus reverse the direction of the sigmoidal binding curve.

4.4 Determination of ligand binding affinity by labeling of His$_6$-tagged target proteins in crude cell lysate

The determination of ligand binding affinity directly in complex sample matrices like crude cell lysate, cell-free expression systems and blood serum offers several advantages. This approach is not only faster and more cost-effective, but it also enables studies with proteins in their natural environment. As the degree of autofluorescence from cell lysate components is higher for green and blue spectral regions, only RED-tris-NTA was used for all ligand binding studies in cell lysates.

As a first example, p38α-mNeonGreen-His$_6$ and mNeonGreen-His$_6$ were both expressed in mammalian cells for direct labeling in cell lysate using RED-tris-NTA. Labeling products were then used to quantify the interaction between p38α and SB203580. Thereby, mNeonGreen-His$_6$ served as a negative control to verify the high specific binding of the small inhibitory compound to p38α. Further, the same interaction was quantified using mNeonGreen-His$_6$-p38α as a target. Fluorescent proteins like mNeonGreen can be used as a fluorescent label for MST affinity analysis, which is performed directly in the cell lysate.

To first determine the optimal dye-to-protein ratio for this labeling approach, first the concentration of p38α in cell lysate was experimentally determined. Therefore, cell lysate was titrated against a constant concentration of RED-tris-NTA and MST was carried out. Knowing the K$_d$ value of this interaction from the experiments described above (Figure 40), the concentration of p38α in the cell lysate could be determined as described in the Method section. Assuming the expression level of mNeonGreen-His$_6$ to be like that of mNeonGreen-His$_6$-p38α, the same amount of fluorophore was added to cell lysate containing mNeonGreen-His$_6$, which served as negative control. For the MST experiment, a serial dilution of SB203580 was prepared using PBST buffer and a constant amount of labeled p38α-mNeonGreen-His$_6$ or labeled mNeonGreen-His$_6$ was added to all dilution steps. Only cell lysate containing labeled p38α showed clear binding towards SB203580 with a K$_d$ of 116 ± 0.84 nM, while no binding could be detected for the p38α-free measurement (Figure 43 A). When mNeonGreen-His$_6$-p38α was used as the target, much higher noise level was observed at simultaneously smaller binding amplitude (Figure 44 B), which consequently resulted in a less reliable fit of the K$_d$ value. Nevertheless, the estimated K$_d$ of 56.8 ± 39 nM is comparable to the K$_d$ determined with RED-tris-NTA labeled His$_6$-p38α.
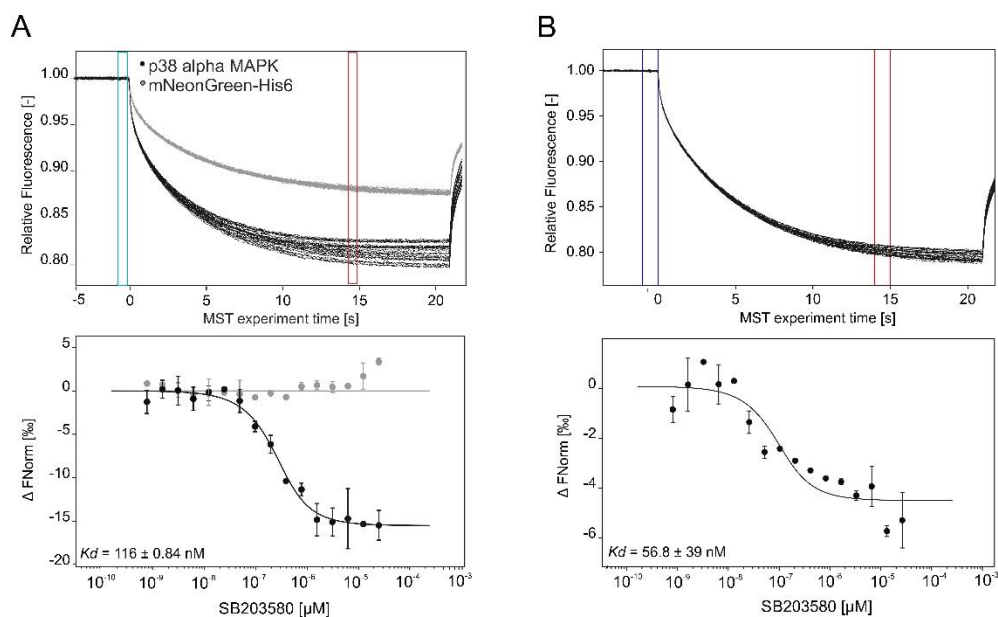
Figure 43: MST analysis of p38α MAPK against SB203580 in HeLa cell lysate. MST traces (top) and dose-response curve (bottom) of the interaction analysis of p38α against SB203580 in HeLa cell lysate in two different detection channels. A) His$_6$-p38α was labeled in HeLa cell lysate to determine its binding affinity toward SB203580 (black, n=3). MST experiments were carried out at 100 % LED and high MST power. mNeonGreen-His$_6$ served as negative control and did not yield a binding curve (grey, n=2). Here, 20 % LED and high MST power were used. B) p38α-mNeonGreen-His$_6$ against SB203580 (n=2). Experiments were carried out at 20 % LED and high MST power. The extracted K_d value was 56.8 ± 39 nM. F_{norm} = normalized fluorescence. Figure adapted with permission from ref.[131], Scientific Reports.

In the next experimental settings, the MST measurements, in which either the purified protein or the protein in crude cell lysate was employed, were compared. Therefore, two proteins His$_6$-pUL53 and pUL50 were used, which form the core nuclear egress complex of human cytomegalovirus (HCMV)[132]. First, the binding affinity of RED-tris-NTA towards His$_6$-pUL53 in *E.coli* cell lysate was tested. For this purpose, protein-containing cell lysate was titrated and a constant amount of RED-tris-NTA was added to all dilution steps. Figure 44 A shows the MST data for this interaction. The binding curve was further used to roughly estimate the concentration of His$_6$-pUL53 in the cell lysate to further determine the optimal dye concentration for protein labeling. To quantify the interaction between RED-tris-NTA labeled His$_6$-pUL53 and pUL50, a serial dilution of the ligand was prepared using PBST buffer and a constant amount of labeled target protein was added to all dilution steps. As a control, pUL53 was purified after expression in *E.coli*, labeled using RED-tris-NTA and added to a serial dilution of pUL50 in HEPES buffer. K_d values of 1.2 ± 0.5 μM for the purified protein and 1.8 ± 0.2 μM for the measurements in crude cell lysate were obtained. These two values are in good agreement with each other and differ only slightly from the K_d value reported by Sam et al. using ITC measurements (K_d = 0.29 μM)[139]. The likely reason for this is that formation of the heterodimeric His$_6$-

pUL53:pUL50 complex is preceded by the dissociation of homodimeric His$_6$-pUL53 into monomers. Hence, the K$_d$ values measured here represent apparent affinities and are thus concentration-dependent.
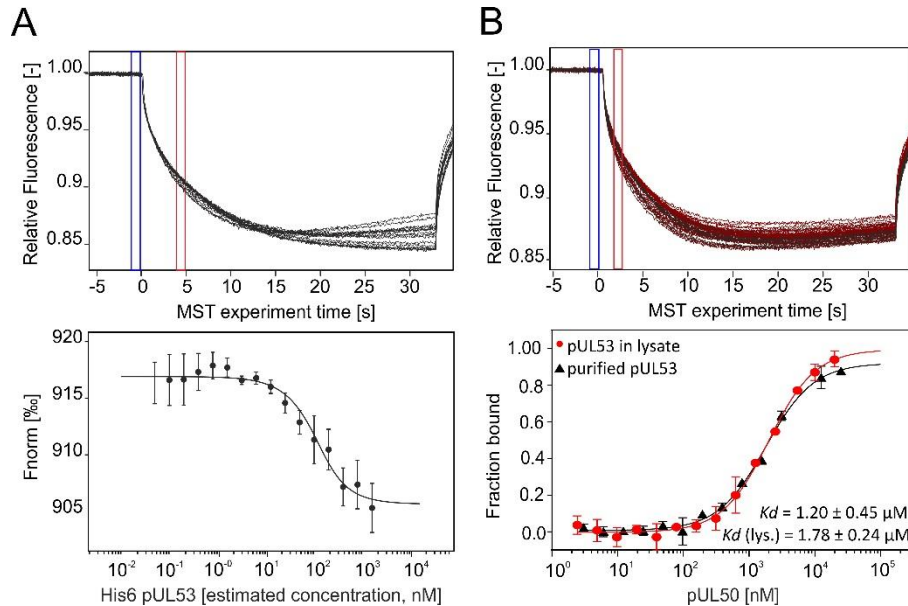
Figure 44: MST analysis of pUL53 against pUL50 in E.coli cell lysate. MST affinity analysis of RED-tris-NTA toward His$_6$-pUL53 containing cell lysate and of labeled pUL53 toward pUL50 in *E.coli* lysate. MST traces (top) and dose-response curves (bottom) for His$_6$-pUL53 against RED-tris-NTA (A) and RED-tris-NTA labeled pUL53 against pUL50 (B) (n=3). A) The interaction between RED-tris-NTA and His$_6$-pUL53 was measured directly in *E.coli* lysate at LED 40 % and high MST power. B) MST traces of His$_6$-pUL53 RED-tris-NTA in *E. coli* lysate (top) and comparison of the binding affinity of pUL50 toward labeled His$_6$-pUL53 measured either with purified His$_6$-pUL53 (black) or with His$_6$-pUL53 in crude lysate (red) (bottom). Measurements were performed at 40 % LED and medium MST power. F$_{norm}$ = normalized fluorescence. Figure adapted with permission from ref.[131], Scientific Reports.

These two examples highlight the versatility, specificity and robustness of the RED-tris-NTA / His-tag system for the use in complex sample matrices such as cell lysate. Measured K$_d$ values were in good agreement with published values and indistinguishable from the binding affinities determined in MST experiments using purified proteins.

5   Discussion

MST is routinely used for the quantification of molecular binding events and can even be used with complex biological matrices like cell lysate and serum. With the intention to develop a near-native labeling strategy for MST measurements, the tris-NTA / His-tag system was exploited, which comprises one of the smallest high-affinity recognition elements known to date. The conjugation of BLUE (OregonGreen® 488), GREEN (NT547) and RED (NT647) to tris-NTA resulted in fluorescence probes with high affinity and selectivity for oligohistidine tags, an ideal tool for site-selective labeling of proteins for quantitative characterization of intermolecular interactions by MST. All investigated DYE-tris-NTA conjugates displayed a high affinity for oligohistidine-tags. The RED-tris-NTA conjugate was identified as the optimal dye conjugate, requiring only low LED power and yielding the best S/N ratios. This conjugate was also successfully used for the labeling of oligohistidine-tagged proteins in crude cell lysate, which allowed the determination of binding affinity for a binding partner directly in lysate. As outlined in this study, the compatibility of RED-tris-NTA with complex sample matrices such as cell lysate has two major benefits: firstly, it permits the study of biomolecular interactions in a near-native environment, allowing a more physiologically realistic assessment. And secondly, it may eliminate the need for protein purification for many MST assay setups, enabling shorter workflows and easier investigation of difficult-to-purify proteins. Combined with its high specificity for His-tags, this labeling strategy offers numerous advantages for protein labeling.

Conclusion

Molecular interactions play a major role in nearly every biological process and are therefore of enormous interest for basic and for drug discovery research[14,97]. Several methods are available for the characterization of molecular interactions, including surface-immobilization based methods, such as SPR, fluorescence-based techniques, like FP, and label-free and in-solution methods, such as DLS or label-free MST[5]. The latter method was so far not applicable for affinity quantification of two molecules that both fluoresce in the same spectral region. This does not only exclude label-free MST from the analysis of PPIs, but in addition restrict its application range to small inhibitory compounds that lack privileged structures, such as indole motifs. Thus, so far MST affinity analysis of such interactions required the attachment of a fluorophore to one of the binding partners. However, applying NHS- or maleimide-based labeling strategies can alter the native structure of a target molecule or can interfere with the ligand binding. These potential negative effects of a covalent labeling let to a strong demand for strategies to increase the application range of label-free MST[59,75]. Hence, in this dissertation three different strategies were presented, to overcome the problem of interfering compounds and to allow measurements of PPIs using the Monolith NT.LabelFree device.

Regarding protein-small molecule interactions, differences in the emission profile of proteins and compounds were used to choose a smaller emission filter that can detect enough protein fluorescence while at the same time cuts-off interfering fluorescence from compounds. This was possible, as the emission spectra of proteins are highly conserved, while the emission profile of compounds can significantly differ from that of proteins. Thus, with the use of a smaller emission filter bandwidth, label-free MST was for the first time accessible to interactions that were so-far not measurable using the broader emission filter. However, spectral separation of proteins and compounds is no longer possible if the compound structure is based on privileged scaffolds like indole-motifs, which are also present in the Trp residues of proteins[59]. Furthermore, prediction of the degree of interference is difficult, as already slight changes in the chemical structure of the substances can change its emission profiles to a significant extent. Besides the fact that the smaller filter bandwidth is not an overall solution for interfering compounds, this filter decreases the sensitivity of the device and thus increases the required amount of sample. To compensate for this effect, further modifications of the optical system would be required. Taken all these limitations together, a second solution strategy was

developed and presented in Chapter 2. Here, a CGT strategy together with data analysis based on a least-mean-square approximation was applied for the quantification of PPIs. Obtained K_d values were verified using standard MST experiments, in which one of the binding partners was fluorescently labeled. Data from both approaches were in good agreement for all tested interactions, while slight deviations in the obtained K_d values were shown to arise from the too high freedom of the fitting algorithm for the selection of the complex F_{norm} value. To define overall limitations of the CGT strategy, various experimental conditions were simulated and showed that distinct K_d determination is restricted to several requirements that need to be fulfilled, such as the requirement for similar fluorescent intensities and F_{norm} values of both unbound proteins. In addition, the mentioned limitation regarding the similarity of the fluorescence intensities of both molecules made this approach not applicable for protein-small molecule interactions, as their fluorescence characteristics differ to a too high extent.

Although a completely label-free assay is preferred as it provides close-to-native experimental conditions, it will always have some limitations, regarding binding partners or buffer components that interfere with the detection system. Thus, the third Chapter describes a compromise between preserving the native structure of the target molecule, while taking the advantage of a fluorescence tag, which allows for highly sensitive binding assays without signal interference of ligands or buffer additives. To achieve this, site-specific labeling of His_6-proteins was applied. Using such a site-specific labeling approach the protein remains in its native conformation and dye-interference with the ligand binding is unlikely, as the tag is mostly located far away from the active site. Another advantage of this labeling strategy is that it provides a fast and gentle way to label a protein, as no removal of unreacted dye is needed. In addition, it can be used to direct label histidine tagged proteins in cell-lysate, what wouldn't be possible using label-free MST as many cell components show intrinsic fluorescence in the UV region. Consequently, the presented site-specific labeling of proteins using tris-NTA-fluorophores is a valuable tool for MST affinity analysis under close-to-native experimental conditions.

References

1. Bartoschik, T. *et al.* Microscale Thermophoresis in Drug Discovery. in *Applied Biophysics for Drug Discovery* (eds. Huddler, D. & Zartler, E. R.) 73–87 (John Wiley and Sons, 2017). doi:10.1002/9781119099512.ch5

2. Kenakin, T. P. Drug Affinity and Efficacy. in *Pharmacology in drug discovery - understanding drug response* 19–35 (Academic Press, 2012).

3. Dickson, M. The Cost of New Drug Discovery and Development. *Discov. Med.* (2009).

4. Renaud, J.-P. *et al.* Biophysics in drug discovery: impact, challenges and opportunities. *Nat. Rev. Drug Discov.* **15,** 679–698 (2016).

5. Zhou, M., Li, Q. & Wang, R. Current Experimental Methods for Characterizing Protein-Protein Interactions. *ChemMedChem* **11,** 738–756 (2016).

6. Yalow;, R. S. & Berson, S. A. Immunoassay of endogenous plasma insulin in man. *J Clin Invest.* 1157–1175 (1960). doi:10.1172/JCI104130

7. Salvador, J.-P., Adrian, J., Galve, R. & Pinacho, D. G. Application of bioassays/biosensors and analysis of pharmaceuticals. in *Comprehensive analytical chemistry* 293 (Elsevir, 2007).

8. Eggeling, C., Brand, L., Ullmann, D. & Jäger, S. Highly sensitive fluorescence detection technology currently available for HTS. *Drug Discov. Today* **8,** 632–641 (2003).

9. Samworth, C. M., Esposti, M. D. & Lenaz, G. Quenching of the intrinsic tryptophan fluorescence of mitochondrial ubiquinol - cytochrome-c reductase by the binding of ubiquinone. *Eur. J. Biochem* **86,** 81–86 (1988).

10. Seidel, S. A. *et al.* Microscale Thermophoresis Quantifies Biomolecular Interactions under Previously Challenging Conditions. *Methods* **59,** 301–315 (2013).

11. Bruylants, G., Wouters, J. & Michaux, C. Differential Scanning Calorimetry in Life Science: Thermodynamics, Stability, Molecular Recognition and Application in Drug Design. *Curr. Med. Chem.* **12,** 2011–2020 (2005).

12. Rossi, A. M. & Taylor, C. W. Analysis of protein-ligand interactions by fluorescence polarization. *Nat. Protoc.* **6,** 365–387 (2011).

13. Sun, Y. *et al.* Effect of Fluorescently Labeling Protein Probes on Kinetics of Protein-Ligand Reactions. *Langmuir* **24,** 13399–13405 (2008).

14. Simeonov, A. *et al.* Fluorescence spectroscopic profiling of compound libraries. *J. Med. Chem.* **51,** 2363–2371 (2008).

15. Ramakrishnan, M. *et al.* Probing cocaine-antibody interactions in buffer and human serum. *PLoS One* **7,** (2012).

16. Lippok, S. *et al.* Direct Detection of Antibody Concentration and Affinity in Human Serum Using Microscale Thermophoresis. *Anal. Chem.* **84,** 3523–3530 (2012).

17. Du, X. *et al.* Insights into Protein–Ligand Interactions: Mechanisms, Models, and Methods. *Int. J. Mol. Sci.* **17,** 144 (2016).

18. Owicki, J. C. Fluorescence Polarization and Anisotropy in High Throughput Screening: Perspectives and Primer. (2000).

19. Pellecchia, M. *et al.* Perspectives on NMR in drug discovery: A technique comes of age. *Nat. Rev. Drug Discov.* **7,** 738–745 (2008).

20. Clore, G. M. & Gronenborn, A. M. NMR structure determination of proteins and protein complexes larger than 20 kDa. *Curr. Opin. Chem. Biol.* **2,** 564–570 (1998).

21. Goldflam, M., Tarragó, T., Gairí, M. & Giralt, E. NMR Studies of Protein-Ligand Interactions. in *Shekhtman A., Burz D. (eds) Protein NMR Techniques* (ed. Springer) 233–259 (Humana Press, 2012). doi:https://doi.org/10.1007/978-1-61779-480-3_14

22. Huber, W. & Mueller, F. Biomolecular interaction analysis in drug discovery using surface plasmon resonance technology. *Curr. Pharm. Des.* **12,** 3999–4021 (2006).

23. Bornhop, D. J. *et al.* Molecular Interactions Studied by Back-scattering Interferometry. *Science (80-.).* **317,** 1732–1736 (2007).

24. Todorova-Balvay, D., Stoilova, I., Gargova, S. & Vijayalakshmi, M. a. An efficient two step purification and molecular characterization of beta-galactosidases from Aspergillus oryzae. *J. Mol. Recognit.* **19,** 299–304 (2007).

25. Hirst, E. R., Yuan, Y. J., Xu, W. L. & Bronlund, J. E. Bond-rupture immunosensors- A review. *Biosens. Bioelectron.* **23,** 1759–1768 (2008).

26. Secundo, F. Conformational changes of enzymes upon immobilisation. *Chem. Soc. Rev.* **42,** 6250 (2013).

27. Tang, Y., Mernaugh, R. & Zeng, X. Nonregeneration Protocol for Surface Plasmon Resonance: Study of High-Affinity Interaction with High-Density Biosensors. *Anal. Biochem.* 1841–1848 (2006). doi:10.1007/s10955-011-0269-9.Quantifying

28. Baksh, M. M., Kussrow, A. K., Mileni, M., Finn, M. G. & Bornhop, D. J. Quantitation of Membrane-Ligand InteractionsUsing Backscattering Interferometry. **29,** 357–360 (2011).

29. OGI, H. Wireless-electrodeless quartz-crystal-microbalance biosensors for studying interactions among biomolecules: A review. *Proc. Japan Acad. Ser. B* **89,** 401–417 (2013).

30. Chaires, J. B. Calorimetry and Thermodynamics in Drug Design. *Annu. Rev. Biophys.* **37,** 135–151 (2008).

31. Hanlon, A. D., Larkin, M. I. & Reddick, R. M. Free-solution, label-free protein-protein interactions characterized by dynamic light scattering. *Biophys. J.* **98,** 297–304 (2010).

32. Lakowicz, J. R. Introduction to Fluorescence. in *Principles of fluorescence spectroscopy* 5–6 (Springer, 2006).

33. Sanderson, M. J., Smith, I., Parker, I. & Bootman, M. D. Fluorescence microscopy. *Cold Spring Harb Protoc* **2014,** (2016).

34. Gao, Z., Hao, Y., Zheng, M. & Chen, Y. A fluorescent dye with large Stokes shift and high stability: synthesis and application to live cell imaging. *RSC Adv.* **7,** 7604–7609 (2017).

35. Pingoud, A. & Urbanke, C. Physikalisch-Chemische Methoden. in *Arbeitsmethoden der Biochemie* (ed. Gruyter, W. de) 238 (1997).

36. Sasaki, S., Drummen, G. P. C. & Konishi, G. Recent advances in twisted intramolecular charge transfer (TICT) fluorescence and related phenomena in materials chemistry. *J. Mater. Chem. C* **4,** 2731–2743 (2016).

37. Fielding, L. NMR methods for the determination of protein-ligand dissociation constants. *Prog. Nucl. Magn. Reson. Spectrosc.* **51,** 219–242 (2007).

38. Erijman, A., Rosenthal, E. & Shifman, J. M. How structure defines affinity in protein-protein interactions. *PLoS One* **9,** (2014).

39. An, Z. Therapeutic RabMAbs Development. in *Therapeutic Monoclonal Antibodies From Bench to Clinic* (John Wiley and Sons, 2009).

40. Jiang, L. & Barclay, A. N. New Assay to Detect Low Affinity Interaction and Characterisation of Leukocyte Receptors for Collagen including LAIR-1. *Eur. J. Immunol* **39,** 1167–1175 (2009).

41. Kastritis, P. L. *et al.* A structure-based benchmark for protein-protein binding affinity. *Protein Sci.* **20,** 482–491 (2011).

42. Howarth, M. *et al.* A monovalent streptavidin with a single femtomolar biotin binding site. *Nat. Methods* **3,** 267–273 (2008).

43. Ludwig, C. Diffusion zwischen ungleich erwärmten Orten gleich zusammengesetzter Lösungen. *Sitzungsbericht Bayer Akad Wiss Wien Math-Naturwiss* 20 (1856).

44. Seidel, S. A. I. *et al.* Microscale thermophoresis quantifies biomolecular interactions under previously challenging conditions. *Methods* **59,** 301–315 (2013).

45. Wienken, C. J., Baaske, P., Rothbauer, U., Braun, D. & Duhr, S. Protein-binding

assays in biological liquids using microscale thermophoresis. *Nat. Commun.* **1,** 100 (2010).

46. Baaske, P., Wienken, C. J., Reineck, P., Duhr, S. & Braun, D. Optical thermophoresis for quantifying the buffer dependence of aptamer binding. *Angew. Chemie - Int. Ed.* **49,** 2238–2241 (2010).

47. Jerabek-Willemsen, M. *et al.* MicroScale Thermophoresis: Interaction analysis and beyond. *J. Mol. Struct.* **1077,** 101–113 (2014).

48. Dong, D. *et al.* The crystal structure of Cpf1 in complex with CRISPR RNA. *Nature* **532,** 1–16 (2016).

49. Linke, P. *et al.* An Automated Microscale Thermophoresis Screening Approach for Fragment-Based Lead Discovery. *J. Biomol. Screen.* **21,** 414–21 (2016).

50. Danne, L. *et al.* Membrane-binding mechanism of a bacterial phospholipid N-methyltransferase. *Mol. Microbiol.* **95,** 313–331 (2015).

51. Seidel, S. A. I. *et al.* Label-free microscale thermophoresis discriminates sites and affinity of protein-ligand binding. *Angew. Chemie - Int. Ed.* **51,** 10656–10659 (2012).

52. Schulze, R. J. *et al.* Membrane protein insertion and proton-motive-force-dependent secretion through the bacterial holo-translocon SecYEG–SecDF–YajC–YidC. *Proc. Natl. Acad. Sci.* **111,** 4844–4849 (2014).

53. Godinic-Mikulcic, V. *et al.* Archaeal aminoacyl-tRNA synthetases interact with the ribosome to recycle tRNAs. *Nucleic Acids Res.* **42,** 5191–5201 (2014).

54. VanLang, C. Quora - What is the largest small molecule library assembled? (2014). Available at: https://www.quora.com/What-is-the-largest-small-molecule-library-assembled.

55. Galloway, W. R. J. D., Isidro-Llobet, A. & Spring, D. R. Diversity-oriented synthesis as a tool for the discovery of novel biologically active small molecules. *Nat. Commun.* **1,** 1–13 (2010).

56. de Sá Alves, F. R., Barreiro, E. J. & Fraga, C. A. M. From nature to drug discovery: the indole scaffold as a 'privileged structure'. *Mini Rev. Med. Chem.* **9,** 782–793 (2009).

57. Evans, B. E., Rittle, K. E. & Bock, M. G. Methods for drug discovery: development of potent, selective, orally effective cholecystokinin antagonists. *J. Med. …* **31,** 2235–2246 (1988).

58. Schneider, P. & Schneider, G. Privileged Structures Revisited. *Angew. Chemie - Int. Ed.* **56,** 7971–7974 (2017).

59. Welsch, M. E., Snyder, S. a & Stockwell, B. R. Privileged Scaffolds for Library Design and Drug Discovery. *Curr Opin Chem Biol.* **14,** 347–361 (2010).

60. DeSimone, R. W., Currie, K. S., Mitchell, S. A., Darrow, J. W. & Pippin, D. A. Privileged structures: applications in drug discovery. *Comb. Chem. High Throughput Screen.* **7,** 473–94 (2004).

61. Turek-Etienne, T. C. *et al.* Evaluation of Fluorescent Compound Interference in 4 Fluorescence Polarization Assays: 2 Kinases, 1 Protease, and 1 Phosphatase. *J Biomol Screen* **8,** 176–184 (2003).

62. Banks, P., Gosselin, M. & Prystay, L. Impact of a Red-Shifted Dye Label for High Throughput Fluorescence Polarization Assays of G Protein-Coupled Receptors. *Biomol. Screen.* **5,** (2000).

63. Simeonov, A. & Davis, M. I. Interference with Fluorescence and Absorbance Assay Guidance Manual. *Assay Guid. Man.* 1–12 (2015).

64. Fowler, A. *et al.* An evaluation of fluorescence polarization and lifetime discriminated polarization for high throughput screening of serine/threonine kinases. *Anal. Biochem.* **308,** 223–231 (2002).

65. Gribbon, P. & Sewing, A. Fluorescence readouts in HTS: No gain without pain? *Drug Discov. Today* **8,** 1035–1043 (2003).

66. Invitrogen Corporation. Theory of Binding Data Analysis. in *Fluorescence Polarization Technical Resource Guide* (2008).

67. Jadhav, A. *et al.* Quantitative analyses of aggregation, autofluorescence, and reactivity artifacts in a screen for inhibitors of a thiol protease. *J. Med. Chem.* **53,** 37–51 (2010).

68. Turconi, S. *et al.* Real Experiences of uHTS: A Prototypic 1536-Well Fluorescence Anisotropy-Based uHTS Screen and Application of Well-Level Quality Control Procedures. *J. Manage.* (1989). doi:10.1177/0021955X9903500605

69. Allen, M., Reeves, J. & Mellor, G. High throughput fluorescence polarization: a homogeneous alternative to radioligand binding for cell surface receptors. *CyberPsychology Behav.* **24,** 247–277 (1989).

70. Zimmermann, T. Spectral imaging and linear unmixing in light microscopy. *Adv. Biochem. Eng. Biotechnol.* **95,** 245–265 (2005).

71. Imbert, P.-E. *et al.* Recommendations for the Reduction of Compound Artifacts in Time-Resolved Fluorescence Resonance Energy Transfer Assays. *Assay Drug Dev. Technol.* **5,** 363–372 (2007).

72. Böhmer, M., Wahl, M., Rahn, H.-J., Erdmann, R. & Enderlein, J. Time-resolved fluorescence correlation spectroscopy. *Chem. Phys. Lett.* **353,** 439–445 (2002).

73. Falconer, M. *et al.* High-throughput screening for ion channel modulators. *J. Biomol. Screen. Off. J. Soc. Biomol. Screen.* **7,** 460–465 (2002).

74. Bazin, H., Trinquet, E. & Mathis, G. U. Time resolved amplification of cryptate

emission : a versatile technology to trace biomolecular interactions. *Mol. Biotechnol.* (2002).

75. Formoso, C. & Forster, S. Tryptophan Fluorescence. *Biol. Chem.* **250,** 3738–3745 (1975).

76. Pan, C.-P. & Barkley, M. D. Conformational effects on tryptophan fluorescence in cyclic hexapeptides. *Biophys. J.* **86,** 3828–35 (2004).

77. Moger, J., Gribbon, P., Sewing, A. & Winlove, C. P. The application of fluorescence lifetime readouts in high-throughput screening. *J. Biomol. Screen.* **11,** 765–72 (2006).

78. Kelly, S. M., Jess, T. J. & Price, N. C. How to study proteins by circular dichroism. *Biochim. Biophys. Acta - Proteins Proteomics* **1751,** 119–139 (2005).

79. Vivian, J. T. & Callis, P. R. Mechanisms of tryptophan fluorescence shifts in proteins. *Biophys. J.* **80,** 2093–2109 (2001).

80. Nikaido, H. Maltose transport system of Escherichia coli: An ABC-type transporter. *FEBS Lett.* **346,** 55–58 (1994).

81. Telmer, P. G. & Shilton, B. H. Insights into the Conformational Equilibria of Maltose-binding Protein by Analysis of High Affinity Mutants. *J. Biol. Chem.* **278,** 34555–34567 (2003).

82. Olson, J. M. & Hallahan, A. R. p38 MAP kinase: A convergence point in cancer therapy. *Trends Mol. Med.* **10,** 125–129 (2004).

83. Thurmond, R. L., Wadsworth, S. A., Schafer, P. H., Zivin, R. A. & Siekierka, J. J. Kinetics of small molecule inhibitor binding to p38 kinase. *Eur. J. Biochem.* **268,** 5747–5754 (2001).

84. Holtzhauer, M. Optische Spektroskopie. in *Methoden in der Proteinanalytik* 136 (Springer, 1996).

85. Berggård, T., Linse, S. & James, P. Methods for the detection and analysis of protein-protein interactions. *Proteomics* **7,** 2833–2842 (2007).

86. Braun, P. & Gingras, A. C. History of protein-protein interactions: From egg-white to complex networks. *Proteomics* **12,** 1478–1498 (2012).

87. Stendel, C. *et al.* SH3TC2, a protein mutant in Charcot-Marie-Tooth neuropathy, links peripheral nerve myelination to endosomal recycling. *Brain* **133,** 2462–2474 (2010).

88. Jellinger, K. A. Interaction between pathogenic proteins in neurodegenerative disorders. *J. Cell. Mol. Med.* **16,** 1166–1183 (2012).

89. Reznick, D. N., Shaw, F. H., Rodd, F. H. & Shaw, R. G. American Association for the Advancement of Science. *Source Sci. New Ser.* **275,** 1934–1937 (2016).

90. Shangary, S. & Wang, S. Small-Molecule Inhibitors of the MDM2-p53 Protein-

Protein Interaction to Reactivate p53 Function: A Novel Approach for Cancer Therapy. *Annu. Rev. Pharmacol. Toxicol.* 223–241 (2010). doi:10.1146/annurev.pharmtox.48.113006.094723.Small-Molecule

91. Arkin, M. R. & Wells, J. A. Small-molecule inhibitors of protein–protein interactions: progressing towards the dream. *Nat. Rev. Drug Discov.* **3,** 301–317 (2004).

92. Yi, F. & Regan, L. A Novel Class of Small Molecule Inhibitors of Hsp90. **33,** 645–654 (2018).

93. Chen, J., Sawyer, N. & Regan, L. Protein-protein interactions: General trends in the relationship between binding affinity and interfacial buried surface area. *Protein Sci.* **22,** 510–515 (2013).

94. Margeat, E., Boukari, H. & Royer, C. A. The Characterization of Biomolecular Interactions using Fluorescence Fluctuation Techniques. in *Protein Interactions: Biophysical approaches for the study of complex reversible systems* (ed. Schuck, P.) 3 (Springer, 2007).

95. Stetefeld, J., McKenna, S. A. & Patel, T. R. Dynamic light scattering: a practical guide and applications in biomedical sciences. *Biophys. Rev.* **8,** 409–427 (2016).

96. Pettibone, J. M., Elzey, S. & Grassian, V. H. An integrated approach toward understanding the environmental fate, transport, toxicity, and health hazards of nanomaterials. in *Nanoscience and Nanotechnolgiy environmental and health impacts* (ed. Grassian, V. H.) 53–54 (Wiley, 2008).

97. Some, D., Kenrick, S. & Corp, W. T. Characterization of Protein-Protein Interactions via Static and Dynamic Light Scattering. *Protein Interact.* 401–426 (2012). doi:10.5772/2336

98. Huang, C. Y. Determination of binding stoichiometry by the continuous variation method: The job plot. *Methods Enzymol.* **87,** 509–525 (1982).

99. Hulme, E. C. & Trevethick, M. A. Ligand binding assays at equilibrium: Validation and interpretation. *Br. J. Pharmacol.* **161,** 1219–1237 (2010).

100. Renny, J. S., Tomasevich, L. L., Tallmadge, E. H. & Collum, D. B. Method of Continuous Variations: Applications of Job Plots to the Study of Molecular Associations in Organometallic Chemistry. *Angew. Chemie - Int. Ed.* **v,** 265–275 (2013).

101. André, T. & Breitsprecher, D. DNA-DNA Interaction Analysis Thermodynamic characterization of DNA hybridization. *Appl. Note* 1–4 (2013).

102. Naiser, T. *et al.* Impact of point-mutations on the hybridization affinity of surface-bound DNA/DNA and RNA/DNA oligonucleotide-duplexes: Comparison of single base mismatches and base bulges. *BMC Biotechnol.* **8,** 1–23 (2008).

103. Rehm, H. & Letzel, T. Gele färben. in *Der Experimentator Proteinbiochemie Proteomics* 21 (Springer, 2016).

104. Wyatt Technology Corporation. *Calypso II ®.* (2011).

105. Hughes, J. P., Rees, S. S., Kalindjian, S. B. & Philpott, K. L. Principles of early drug discovery. *Br. J. Pharmacol.* **162,** 1239–1249 (2011).

106. Khavrutskii, L. *et al.* Protein Purification-free Method of Binding Affinity Determination by Microscale Thermophoresis. *J. Vis. Exp.* **12,** 1–6 (2013).

107. Krishnan, B., Szymanska, A. & Gierasch. Site-specific Fluorescent Labeling of Poly-histidine Sequences Using a Metal-chelating Cysteine. *Chem. Biol. Drug Des.* **76,** 211–220 (2007).

108. Nfor, B. K. *et al.* Design strategies for integrated protein purification processes: challenges, progress and outlook. *J. Chem. Technol. Biotechnol.* **82,** 1115–1121 (2007).

109. Lindhoud, S., Westphal, A. H., Visser, A. J. W. G., Borst, J. W. & van Mierlo, C. P. M. Fluorescence of Alexa Fluor Dye Tracks Protein Folding. *PLoS One* **7,** 1–8 (2012).

110. Crivat, G. & Taraska, J. W. Imaging proteins inside cells with fluorescent tags. *Trends Biotechnol.* **30,** 8–16 (2012).

111. Chen, I., Howarth, M., Lin, W. & Ting, A. Y. Site-specific labeling of cell surface proteins with biophysical probes using biotin ligase. *Nat. Methods* **2,** 99–104 (2005).

112. Nienberg, C. *et al.* Site-specific labeling of protein kinase CK2: Combining surface display and click chemistry for drug discovery applications. *Pharmaceuticals* **9,** 1–15 (2016).

113. Xie, J. & Schultz, P. G. Adding amino acids to the genetic repertoire. *Curr. Opin. Chem. Biol.* **9,** 548–554 (2005).

114. Stromgaard, A., Jensen, A. A. & Stromgaard, K. Site-specific incorporation of unnatural amino acids into proteins. *ChemBioChem* **5,** 909–916 (2004).

115. Kapanidis, A. N., Ebright, Y. W. & Ebright, R. H. Site-specific incorporation of fluorescent probes into protein: Hexahistidine-tag-mediated fluorescent labeling with (Ni2+:Nitrilotriacetic Acid)n - Fluorochrome conjugates. *J. Am. Chem. Soc.* **123,** 12123–12125 (2001).

116. Griffin, B. A. *et al.* Specific Covalent Labeling of Recombinant Protein Molecules Inside Live Cells. **281,** 269–272 (2016).

117. Janknecht, R. *et al.* Rapid and efficient purification of native histidine-tagged protein expressed by recombinant vaccinia virus. *Biochemistry* **88,** 8972–8976 (1991).

118. Lai, Y.-T. *et al.* Rapid labeling of intracellular His-tagged proteins in living cells. *Proc. Natl. Acad. Sci.* **112,** 201419598 (2015).

119. Chen, H. M., Wang, W. C. & Chen, S. H. A metal-chelating piezoelectric sensor chip for direct detection and oriented immobilization of polyhis-tagged proteins. *Biotechnol. Prog.* **20,** 1237–1244 (2004).

120. Nieba, L. *et al.* BIACORE Analysis of Histidine-Tagged Proteins Using a Chelating NTA Sensor Chip. *Anal. Biochem.* **252,** 217–228 (1997).

121. Braner, M., Kollmannsperger, A., Wieneke, R. & Tampé, R. 'Traceless' tracing of proteins – high-affinity trans-splicing directed by a minimal interaction pair. *Chem. Sci.* **7,** 2646–2652 (2016).

122. Hochuli, E., Döbeli, H. & Schacher, A. New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues. *J. Chromatogr. A* **411,** 177–184 (1987).

123. Lata, S., Reichel, A., Brock, R., Tampé, R. & Piehler, J. High-affinity adaptors for switchable recognition of histidine-tagged proteins. *J. Am. Chem. Soc.* **127,** 10205–10215 (2005).

124. Lata, S., Gavutis, M., Tampé, R. & Piehler, J. Specific and stable fluorescence labeling of histidine-tagged proteins for dissecting multi-protein complex formation. *J. Am. Chem. Soc.* **128,** 2365–2372 (2006).

125. Van Der Does, C., Presenti, C., Schulze, K., Dinkelaker, S. & Tampé, R. Kinetics of the ATP hydrolysis cycle of the nucleotide-binding domain of Mdl1 studied by a novel site-specific labeling technique. *J. Biol. Chem.* **281,** 5694–5701 (2006).

126. Strunk, J. J. *et al.* Probing Protein Conformations by in Situ Non-Covalent Fluorescence Labeling Probing Protein Conformations by in Situ Non-Covalent Fluorescence Labeling. *Bioconjug. Chem.* 41–46 (2009). doi:10.1021/bc8002088

127. DeRocco, V. C., Anderson, T., Piehler, J. & Erie, D. A. Four-color single molecule fluorescence with noncovalent dye labeling to monitor dynamic multimolecular complexes. **49,** 807–816 (2011).

128. Reichel, A. *et al.* Noncovalent , Site-Specific Biotinylation of Histidine-Tagged Proteins. **79,** 8590–8600 (2007).

129. You, C. *et al.* Self-controlled monofunctionalization of quantum dots for multiplexed protein tracking in live cells. *Angew. Chemie - Int. Ed.* **49,** 4108–4112 (2010).

130. Giannone, G. *et al.* Dynamic superresolution imaging of endogenous proteins on living cells at ultra-high density. *Biophys. J.* **99,** 1303–1310 (2010).

131. Bartoschik, T. *et al.* Near-native, site-specific and purification-free protein labeling for quantitative protein interaction analysis by MicroScale Thermophoresis. *Sci.*

Rep. **8,** 4977 (2018).

132. Walzer, S. A. *et al.* Crystal structure of the human cytomegalovirus pUL50-pUL53 core nuclear egress complex provides insight into a unique assembly scaffold for virus-host protein interactions. *J. Biol. Chem.* **290,** 27452–27458 (2015).

133. Bornhorst, J. A. & Falke, J. J. Purification of Proteins Using Polyhistidine Affinity Tags. **37,** 62–70 (2012).

134. Liu *et al.* Specific and reversible immobilization of histidine-tagged proteins on functionalized silicon nanowires. *Nanotechnology* **21,** 245105 (2010).

135. McPhail, D. B. & Goodman, B. A. Tris buffer - a case for caution in ints use in copper-containing systems. *Biochem. J.* **1,** 559–560 (1977).

136. Fischer, B. E., Häring, U. K., Tribolet, R. & Sigel, H. Metal Ion/Buffer Interactions: Stability of Binary and Ternary Complexes Containing 2-Amino-2(hydroxymethyl)-1,3-propanediol (Tris) and Adenosine 5′-Triphosphate (ATP). *Eur. J. Biochem.* **94,** 523–530 (1979).

137. Fu, Y., O'Connor, L. M., Shepherd, T. G. & Nachtigal, M. W. The p38 MAPK inhibitor, PD169316, inhibits transforming growth factor beta-induced Smad signaling in human ovarian cancer cells. *Biochem. Biophys. Res. Commun.* **310,** 391–397 (2003).

138. Zhang, Z. *et al.* PD169316, a specific p38 inhibitor, shows antiviral activity against Enterovirus71. *Virology* **508,** 150–158 (2017).

139. Sam, M. D., Evans, B. T., Coen, D. M. & Hogle, J. M. Biochemical, biophysical, and mutational analyses of subunit interactions of the human cytomegalovirus nuclear egress complex. *J. Virol.* **83,** 2996–3006 (2009).

Supplementary Data

Table of Contents

1. nanoDSF for protein stability determination

nanoDSF is a label-free differential scanning fluorimetry method, that is used to quantify protein stability. Therefore, it determines the unfolding transition midpoint T_m (°C), which is the point at where half of the protein is unfolded. For the unfolding procedure, either thermal or chemical unfolding of the protein is used. While thermal unfolding experiments use a temperature ramp to unfold proteins, chemical unfolding experiments use chaotropes such as urea or guanidinhydrochlorid. During both measurement modes, the device monitors changes in the intrinsic fluorescence of proteins Trp and Tyr residues, as those strongly depend on their local environment[1]. Trp residues which are located in the hydrophobic core of proteins show maximum fluorescence emission at 330 nm, while the same amino acid exposed to the solvent, will show maximum emission at 350 nm[2]. To detect both signals, nanoDSF uses and dual-UV detector, giving the option to monitor the unfolding event either at 330 nm, 350 nm or at the ratio of 350 nm / 330 nm. This fluorescence signal is then plotted against the temperature or against the concentration of the denaturant. An example for a resulting unfolding curve is shown in Figure 1.
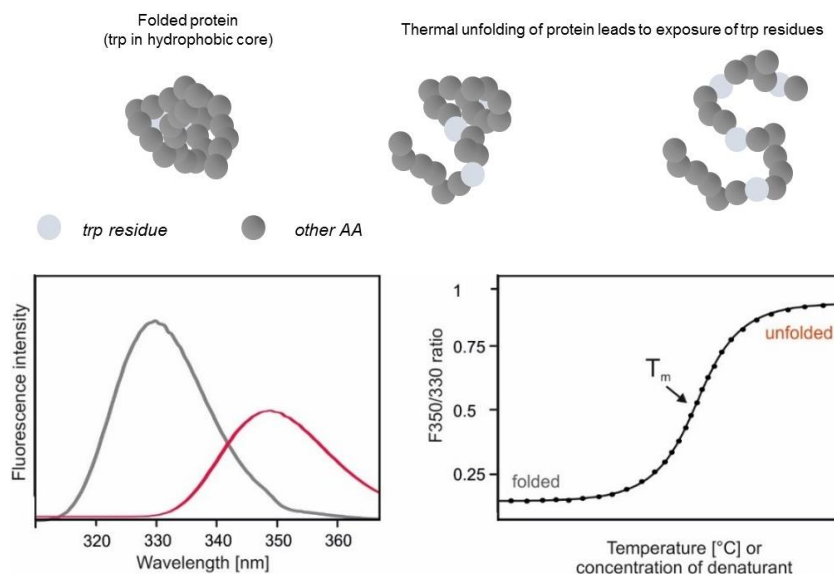
Figure 1: The principle of protein unfolding using nanoDSF. Regarding a folded protein, Trp residues are mostly located in the hydrophobic core, exhibiting an emission maximum at 330 nm (grey peak). Upon unfolding, Trp residues are more and more oriented towards the solvent, which shifts the emission maximum towards 350 nm (red peak). For data analysis, fluorescence intensity can be recorded at the ratio of 350 nm / 330 nm, which is plotted against the temperature or against the denaturant concentration. The inflection point of the resulting unfolding curve gives the melting temperature (T_m).

In addition, nanoDSF is used to determine colloidal stability using a backreflection optics. Here, light passes through the sample-containing glass capillary. In case of no

aggregates in the solution, all incident light passes through the capillary and is then reflected by the capillary tray and detected by the optical system. If the sample contains aggregates, the incident light will be scattered by the particles, leading to loss of detected light. This loss of light can be quantified and provides information on protein aggregation formation.

nanoDSF typically is measured in the Prometheus NT.48 (NanoTemper Technologies GmbH), which consists of a capillary tray with the option to load 48 glass capillaries for one measurement. Thereby, the capillaries are placed on a heatable tray for the thermal unfolding of proteins. For the measurement, 10 µL of sample is filled into each glass capillary. Thereby the concentration of the sample can range from 5 µg/mL up to 250 mg/mL (concentrations refer to a standard IgG) in one run. For thermal unfolding, samples will be heated from 15 to 95 °C, in either 1 °C/min or faster. Additionally, the device can be equipped with a high-temperature upgrade, which allows a maximum temperature of 110 °C.

2. Thermal shift assay for maltose binding protein

For the label-free thermal shift assay, the Prometheus NT.48 instrument was used. This assay was used to determine if the binding pocket of maltose binding protein (MBP) contains maltose as an impurity left from column purification of MBP. First, size exclusion chromatography was used to remove any residual maltose from the shipped MBP. Briefly, column B was equilibrated using 9 mL MSTP buffer. After equilibration 200 µL of 25 µM solution was placed onto the column. After the sample entered the resin, the total volume was adjusted to 500 µL using assay buffer. For protein elution 600 µL of assay buffer was used. The first few µL did not contain protein and were discarded, while the residual flow-through was collected (~ 500 µL). The purified MBP was diluted further with the MSTP buffer to a final concentration of 5 µM. For the samples in which MBP is bound to maltose, 10 mM maltose (dissolved in ddH$_2$O) was mixed 1:1 with 10 µM MBP (purified and non-purified) to reach experimental concentrations of 5 µM MBP and 5 mM maltose. Samples were filled into Prometheus NT.48 Series nanoDSF Grade Standard Capillaries and loaded into the Prometheus NT.48 instrument. Before running the thermal unfolding experiment, capillaries were sealed using the sealing paste. Thermal unfolding was recorded using 70 % LED power and a heating ramp of 20 °C to 100 °C with 1°C / min. Thermal unfolding data were exported with the chart export function in the PR Control v1.12.3 software.

Because it is known that residual maltose remains bound to MBP after the purification process, MBP was purified using size exclusion chromatography prior the MST

experiment to remove residual maltose from the binding pocket. Thermal shift assay was used to determine the success rate of the purification. Therefore, nanoDSF technology was used to compare the thermal unfolding data of purified and non-purified MBP, both with and without the addition of maltose (Figure 2). After successful removal of maltose with size exclusion chromatography, the melting temperature of MBP decreased (green), whereas after addition of maltose, the unfolding curve of non-purified MBP nicely overlaps with the data of re-purified MBP supplemented with maltose (yellow and red) These unfolding data show that binding of maltose stabilizes MBP and thus increases its melting temperature. Furthermore, data show that the size-exclusion purification was sufficient to reduce any maltose from the MBP binding pocket and that the protein can further be used for maltose-binding quantification using MST.
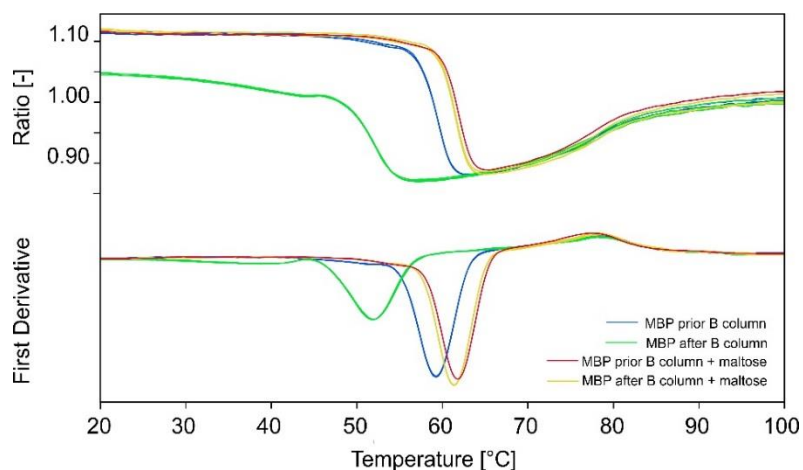
Figure 2: Thermal unfolding of MBP with and without the addition of maltose. 350 / 330 nm ratio and first derivatives of 5 µM MBP in MSTP buffer prior (blue) and after (green) size exclusion chromatography, re-purified MBP with 5 mM maltose (yellow), and non-purified MBP with 5 mM maltose (red), n=3. Samples were loaded into Prometheus NT.48 Series nanoDSF Grade Standard Capillaries and sealed. Thermal unfolding was run using 70 % LED power and a heating ramp from 20 °C to 100 °C and 1.0 °C / min.

3. CGT titration

Table 1: CGT pipetting scheme.

Vial number	Molecule A	Molecule B
1	10.0 µL	0.0 µL
2	23.4 µL	1.8 µL
3	21.6 µL	3.6 µL
4	19.8 µL	5.4 µL
5	18.0 µL	7.2 µL
6	16.2 µL	9.0 µL
7	14.4 µL	10.8 µL
8	12.6 µL	12.6 µL
9	10.8 µL	14.4 µL
10	9.0 µL	16.2 µL
11	7.2 µL	18.0 µL
12	5.4 µL	19.8 µL
13	3.6 µL	21.6 µL
14	1.8 µL	23.4 µL
15	0.0 µL	10.0 µL

Formulas used for the calculation of free concentrations of [A] and [B] (only shown for B), the complex concentration [AB] and for the simulation of F_{norm} values for the case of an interaction and for the case of no interaction:

$$[B] = -\frac{1}{2}\{([A_{tot}] - [B_{tot}] + K_d) - \{([A_{tot}] - [B_{tot}] + K_d)^2 + 4[B_{tot}]K_d\}^{\frac{1}{2}}\} \tag{20}$$

$$[AB] = \frac{[A][B]}{K_d} \tag{21}$$

Fnorm for no interaction:
$$\frac{FnormA*[A_{tot}] + FnormB*[B_{tot}]}{[A_{tot}]+[B_{tot}]}$$

Fnorm for interaction:
$$\frac{FnormA*[A]+FnormB*[B]+FnormAB*[AB]}{[A] +[B]+[AB]} \tag{22}$$

4. CGT of DNA Hybridization experiments

Table 2: CGT of DNA Hybridization. Absolute concentration in each tube.

Vial number	Absolute conc. template [nM]	Absolute conc. perfect match / mismatch 1 / mismatch 2 [nM]
1	0.00000000	100.0000000
2	7.14285714	92.8571429
3	14.2857143	85.7142857
4	21.4285714	78.5714286
5	28.5714286	71.4285714
6	35.7142857	64.2857143
7	42.8571429	57.1428571
8	50.0000000	50.0000000
9	57.1428571	42.8571429
10	64.2857143	35.7142857
11	71.4285714	28.5714286
12	78.5714286	21.4285714
13	85.7142857	14.2857143
14	92.8571429	7.14285714
15	100.0000000	0.00000000

Table 3: Simulated data summary. Simulated data, using the following values: F_{norm} template: 825.635335, F_{norm} perfect match: 817.940166, F_{norm} complex: 874.859182, K_d: 17.8530193, quadratic error: 1.16443823

Vial number	Free conc. template [nM]	Free conc. perfect match [nM]	Complex conc. [nM]	F_{norm} for interaction	F_{norm} for no interaction	Measured F_{norm} values	St.dev.
1	0.00000	100.0000	0.0000000	817.94016	817.9401	817.94016	2.71958
2	1.21699	86.93127	5.9258659	821.62512	818.4898	822.63069	2.49096
3	2.77063	74.19920	11.515079	825.58833	819.0394	826.78102	2.88454
4	4.79461	61.93747	16.633956	829.73974	819.5891	830.52058	3.31007
5	7.48031	50.33745	21.091113	833.88320	820.1387	834.93548	3.63513
6	11.0866	39.65811	24.627600	837.67010	820.6884	838.37233	4.23097
7	15.9205	30.20625	26.936605	840.60153	821.2380	841.95515	4.59336
8	22.2557	22.25575	27.744246	842.16573	821.7877	843.10298	3.51749
9	30.2062	15.92053	26.936605	842.10612	822.3374	842.14825	1.99491
10	39.6581	11.08668	24.627600	840.58711	822.8870	838.82567	1.16394
11	50.3374	7.480315	21.091113	838.06262	823.4367	835.37182	1.62850
12	61.9374	4.794615	16.633956	835.01435	823.9863	833.51641	1.69436
13	74.1992	2.770635	11.515079	831.80018	824.5360	830.75661	1.56357
14	86.9312	1.216991	5.9258659	828.63646	825.0856	828.15897	2.85521
15	100.0000	0.000000	0.0000000	825.63533	825.6353	825.63533	2.23997

Table 4: Simulated data summary. Simulated data, using the following values: F_{norm} template: 833.295444, F_{norm} mismatch 1: 816.408961, F_{norm} complex: 880.293179, K_d: 19.2906922, quadratic error: 0.1403165

Vial number	Free conc. template [nM]	Free conc. mismatch 1 [nM]	Complex conc. [nM]	F_{norm} for interaction	F_{norm} for no interaction	Measured F_{norm} values	St.dev.
1	0.0000	100.00000	0.0000000	816.40896	816.40896	816.40896	0.6511364
2	1.2962	87.010514	5.8466285	820.60845	817.61513	821.58908	1.0251395
3	2.9423	74.370880	11.343405	825.14322	818.82131	825.85065	0.6338866
4	5.0717	62.214578	16.356850	829.92576	820.02749	830.67616	1.6884190
5	7.8715	50.728712	20.699858	834.76097	821.23367	834.03066	0.5595524
6	11.588	40.160066	24.125648	839.30127	822.43984	838.93687	0.6806615
7	16.507	30.792972	26.349885	843.04967	823.64602	843.35432	1.0668842
8	22.874	22.874913	27.125086	845.48812	824.85220	846.10832	0.4384865
9	30.792	16.507257	26.349885	846.32509	826.05838	846.47944	0.6397446
10	40.160	11.588637	24.125648	845.66009	827.26455	845.49431	0.4372318
11	50.728	7.8715698	20.699858	843.88714	828.47073	842.97256	0.7490935
12	62.214	5.0717208	16.356850	841.46217	829.67691	840.91747	0.6631446
13	74.370	2.9423093	11.343405	838.74827	830.88308	838.93532	1.2250487
14	87.010	1.2962285	5.8466285	835.98137	832.08926	835.77715	0.8381949
15	100.00	0.0000000	0.0000000	833.29544	833.29544	833.29544	0.3060108

Table 5: Simulated data summary. Simulated data, using the following values: F_{norm} template: 827.872274, F_{norm} mismatch 2: 803.861025, F_{norm} complex: 946.806993, K_d: 85.0505851, quadratic error: 0.15111039

Vial number	Free conc. template [nM]	Free conc. mismatch 2 [nM]	Complex conc. [nM]	F_{norm} for interaction	F_{norm} for no interaction	Measured F_{norm} values	St.dev.
1	0.000000	100.00000	0.0000	803.86102	803.86102	803.86102	0.0565020
2	3.486369	89.200655	3.6564	810.15508	805.57611	810.76565	1.3514143
3	7.413442	78.842013	6.8722	816.32100	807.29120	816.64699	1.4986118
4	11.83250	68.975360	9.5960	822.17695	809.00629	822.38911	2.2025124
5	16.79335	59.650498	11.778	827.51566	810.72138	827.83170	2.1547540
6	22.34081	50.912246	13.373	832.12158	812.43647	832.12501	2.1714705
7	28.51081	42.796527	14.346	835.79579	814.15156	836.25987	0.9939388
8	35.32668	35.326680	14.673	838.38398	815.86664	839.13284	0.1333003
9	42.79652	28.510812	14.346	839.80050	817.58173	840.29171	0.5398805
10	50.91224	22.340818	13.373	840.04104	819.29682	839.67665	1.4309383
11	59.65049	16.793355	11.778	839.18003	821.01191	837.88272	1.3478174
12	68.97536	11.832503	9.5960	837.35407	822.72700	836.44815	1.923156
13	78.84201	7.4134425	6.8722	834.73753	824.44209	833.84912	2.1363912
14	89.20065	3.4863693	3.6564	831.51726	826.15718	831.25164	1.8914184
15	100.0000	0.0000000	0.0000	827.87227	827.87227	827.87227	1.2767998

5. Perfect match DNA Hybridization using different starting concentrations

Table 6: CGT of DNA Hybridization. Absolute concentration in each tube.

Vial number	Absolute conc. template [nM]	Absolute conc. perfect match [nM]
1	0.00000000	100.0000000
2	7.14285714	92.8571429
3	14.2857143	85.7142857
4	21.4285714	78.5714286
5	28.5714286	71.4285714
6	35.7142857	64.2857143
7	42.8571429	57.1428571
8	50.0000000	50.00000000
9	57.1428571	42.8571429
10	64.2857143	35.7142857
11	71.4285714	28.5714286
12	78.5714286	21.4285714
13	85.7142857	14.2857143
14	92.8571429	7.14285714
15	100.0000000	0.00000000

Table 7: Simulated data summary. Simulated data, using the following values: F_{norm} template: 773.26655, F_{norm} perfect match: 758.20794, Complex F_{norm} 821.66787, K_d: 10.219274, quadratic error: 0.3024734

Vial number	Free conc. template [nM]	Free conc. perfect match [nM]	Complex conc.	F_{norm} for interaction	F_{norm} for no interaction	Measured F_{norm} values
1	0.00000000	100.000000	0.000000000	758.207939	758.207939	758.207939
2	0.75562359	87.1556236	6.444376408	762.662769	759.283554	762.831529
3	1.75246507	73.7524651	12.64753493	767.612116	760.359169	768.825269
4	3.11197549	60.7119755	18.48802451	773.030943	761.434784	773.455082
5	5.03496552	48.2349655	23.76503448	778.769343	762.510399	777.536562
6	7.84945266	36.6494527	28.15054734	784.424704	763.586014	782.825506
7	12.0419499	26.4419499	31.15805008	789.203951	764.661635	787.053121
8	18.1532043	18.1532043	32.24679571	792.046496	765.737245	791.541288
9	26.4419499	12.0419499	31.15805008	792.317655	766.812862	795.197098
10	36.6494527	7.84945266	28.15054734	790.394302	767.888475	791.339404
11	48.2349655	5.03496552	23.76503448	787.213976	768.964092	787.905469
12	60.7119755	3.11197549	18.48802451	783.568609	770.039705	783.975059
13	73.7524651	1.75246507	12.64753493	779.911489	771.115321	779.805644
14	87.1556236	0.75562359	6.444376408	776.451711	772.190936	776.123565
15	100.000000	0.00000000	0.000000000	773.266551	773.266551	773.266551

Table 8: CGT of DNA Hybridization. Absolute concentration in each tube.

Vial number	Absolute conc. template [nM]	Absolute conc. perfect match [nM]
1	0.0	50.0
2	3.6	46.8
3	7.2	43.2
4	10.8	39.6
5	14.4	36.0
6	18.0	32.4
7	21.6	28.8
8	25.2	25.2
9	28.8	21.6
10	32.4	18.0
11	36.0	14.4
12	39.6	10.8
13	43.2	7.2
14	46.8	3.6
15	50.0	0.0

Table 9: Simulated data summary. Simulated data, using the following values: F_{norm} template: 760.939848, F_{norm} perfect match: 744.752061, Complex F_{norm} 809.738121, K_d: 4.85606702, quadratic error: 0.2588723

Vial number	Free conc. template [nM]	Free conc. perfect match [nM]	Complex conc.	F_{norm} for interaction	F_{norm} for no interaction	Measured F_{norm} values
1	0.00000000	50.0000000	0.000000000	744.752061	744.752061	744.752061
2	0.36106724	43.5610672	3.238932755	749.339115	745.908332	748.639451
3	0.83856555	36.8385656	6.361434449	754.447633	747.064602	752.962325
4	1.49212576	30.2921258	9.307874242	760.060013	748.220873	761.924103
5	2.42151075	24.0215108	11.97848925	766.032685	749.377143	765.441753
6	3.79240745	18.1924075	14.20759255	771.959042	750.533414	771.039652
7	5.85591514	13.0559151	15.74408486	777.010345	751.689684	775.406297
8	8.89752323	8.89752323	16.30247677	780.046852	752.845955	779.819703
9	13.0559151	5.85591514	15.74408486	780.373467	754.002225	781.708092
10	18.1924075	3.79240745	14.20759255	778.399734	755.158496	779.655696
11	24.0215108	2.42151075	11.97848925	775.133217	756.314766	774.928314
12	30.2921258	1.49212576	9.307874242	771.405454	757.471037	772.574006
13	36.8385656	0.83856555	6.361434449	767.680588	758.627307	767.382229
14	43.5610672	0.36106724	3.238932755	764.167287	759.783578	763.683941
15	50.0000000	0.00000000	0.000000000	760.939848	760.939848	760.939848

Table 10: CGT of DNA Hybridization. Absolute concentration in each tube.

Vial number	Absolute conc. template [nM]	Absolute conc. perfect match [nM]
1	0.0	5.00
2	0.36	4.68
3	0.72	4.32
4	1.08	3.96
5	1.44	3.60
6	1.80	3.24
7	2.16	2.88
8	2.52	2.52
9	2.88	2.16
10	3.24	1.80
11	3.6	1.44
12	3.96	1.08
13	4.32	0.72
14	4.68	0.36
15	5.00	0.0

Table 11: Simulated data summary. Simulated data, using the following values: F_{norm} template: 768.593033, F_{norm} perfect match: 755.21791, Complex F_{norm} 816.311303, K_d: 1.70864076, quadratic error: 0.20404395

Vial number	Free conc. perfect match [nM]	Free conc. template [nM]	Complex conc.[nM]	F_{norm} for interaction	F_{norm} for no interaction	Measured F_{norm} values
1	0.00000000	5.00000000	0.000000000	755.217912	755.217912	755.217912
2	0.10036067	4.42036067	0.259639333	758.816924	756.173276	759.343884
3	0.22242049	3.82242049	0.497579511	762.565033	757.128642	761.643245
4	0.37199506	3.25199506	0.708004939	766.351326	758.084008	765.763879
5	0.55606875	2.71606875	0.883931248	770.001074	759.039374	770.048525
6	0.78238029	2.22238029	1.017619714	773.275445	759.994742	773.219142
7	1.05839562	1.77839562	1.101604377	775.900677	760.950106	776.228173
8	1.38970235	1.38970235	1.130297646	777.634235	761.905471	776.509737
9	1.77839562	1.05839562	1.101604377	778.345858	762.860837	778.195833
10	2.22238029	0.78238029	1.017619714	778.063699	763.816203	778.369157
11	2.71606875	0.55606875	0.883931248	776.952419	764.771569	777.539189
12	3.25199506	0.37199506	0.708004939	775.243385	765.726935	777.488501
13	3.82242049	0.22242049	0.497579511	773.165206	766.682301	772.364057
14	4.42036067	0.10036067	0.259639333	770.903989	767.637667	770.291217
15	5.00000000	0.00000000	0.000000000	768.593033	768.593033	768.593033

6. CGT of IL6 nanobody against IL6 antigen

Table 12: CGT of nanobody antigen interaction. Absolute concentration in each tube.

Vial number	Absolute conc. antigen [nM]	Absolute conc. nanobody [nM]
1	0.00	250.00
2	25.00	232.14
3	50.00	214.29
4	75.00	196.43
5	100.00	178.57
6	125.00	160.71
7	150.00	142.86
8	175.00	125.00
9	200.00	107.14
10	225.00	89.29
11	250.00	71.43
12	275.00	53.57
13	300.00	35.71
14	325.00	17.86
15	350.00	0.00

Table 13: Simulated data summary. Simulated data, using the following values: F_{norm} nanobody: 958.223794, F_{norm} antigen: 949.578126, Complex F_{norm}: 928.800717, K_d: 55.3873329, quadratic error: 0.11457584

Vial number	Free conc. nanobody [nM]	Free conc. antigen [nM]	Complex conc. [nM]	F_{norm} for interaction	F_{norm} for no interaction	Measured F_{norm} values	St.dev.
1	0.000000	100.0000	0.0000000	949.57812	949.57812	949.57812	3.6989619
2	5.172467	212.3153	19.827532	948.03062	950.41867	948.06535	1.8987948
3	11.95604	176.2417	38.043959	946.54116	951.21379	946.94831	2.2721378
4	20.99961	142.4281	54.000384	945.25287	951.96706	945.69665	0.9039912
5	33.14539	111.7168	66.854604	944.37069	952.68169	945.01979	0.4512768
6	49.30852	85.0228	75.69147	944.11982	953.36060	944.12029	1.2208516
7	70.16446	63.0216	79.83553	944.63892	954.00639	944.06134	1.0380502
8	95.79511	45.7951	79.20488	945.87578	954.62143	946.08533	0.1044339
9	125.6388	32.7817	74.36110	947.60718	955.20786	947.03045	2.0446483
10	158.8026	23.0883	66.19734	949.56823	955.76763	949.32097	0.8070402
11	194.4092	15.8378	55.59073	951.55589	956.30253	951.82826	1.8641302
12	231.7618	10.3332	43.23819	953.45204	956.81417	952.57822	1.6735065
13	270.3583	6.07258	29.64169	955.20277	957.30404	955.88684	3.1565727
14	309.8508	2.70798	15.14915	956.79219	957.77349	957.09468	2.7785391
15	100.0000	0.000000	0.0000000	958.22379	958.22379	958.22379	4.2853042

7. CGT of MBP binding protein against MBP

Table 14: CGT of MBP binding protein against MBP. Absolute concentration in each tube.

Vial number	Absolute conc. MBP [nM]	Absolute conc. MBP binding protein [nM]
1	0.00000000	200.000000
2	7.14285714	185.714286
3	14.2857143	171.428571
4	21.4285714	157.142857
5	28.5714286	142.857143
6	35.7142857	128.571429
7	42.8571429	114.285714
8	50.0000000	100.000000
9	57.1428571	85.7142857
10	64.2857143	71.4285714
11	71.4285714	57.1428571
12	78.5714286	42.8571429
13	85.7142857	28.5714286
14	92.8571429	14.2857143
15	100.0000000	0.00000000

Table 15: Simulated data summary. Simulated data, using the following values: F_{norm} MBP: 635.398104, F_{norm} MBP binding protein: 612.664769, Complex F_{norm}: 739.658359, K_d: 5.46473725, quadratic error: 1.69408839

Vial number	Free conc. MBP [nM]	Free conc. MBP binding protein [nM]	Complex conc.	F_{norm} for interaction	F_{norm} for no interaction	Measured F_{norm} values	St.dev.
1	0.0000000	200.0000	0.0000000	612.66476	612.66476	612.66476	6.4000821
2	0.2118548	178.7832	6.93100228	617.42477	613.50674	605.32272	8.4266407
3	0.4786894	157.6215	13.8070248	622.92777	614.41348	608.99876	14.700579
4	0.824637	136.538	20.6039341	629.34740	615.39276	624.242	8.6975715
5	1.289944	115.575	27.2814844	636.90319	616.45365	637.84466	7.8318194
6	1.946468	94.8036	33.7678171	645.85980	617.60679	646.85403	9.6458207
7	2.933875	74.3624	39.9232672	656.48590	618.86476	657.08161	6.0869066
8	4.552629	54.5526	45.4473700	668.85676	620.24254	677.05541	16.755838
9	7.515278	36.0867	49.6275781	682.09801	621.75810	677.70054	8.4517613
10	13.47099	20.6138	50.8147195	692.28100	623.43319	683.15627	11.819692
11	24.65377	10.3680	46.7747930	692.13700	625.29439	698.79341	4.4100603
12	40.77883	5.06455	37.7925901	681.13357	627.37457	687.91907	9.5417692
13	59.54459	2.40173	26.1696910	665.74285	629.71477	664.04428	5.6127766
14	79.49035	0.91892	13.366785	650.03653	632.36699	646.30674	2.8519700
15	100.00000	0.000000	0.0000000	635.39810	635.39810	635.39810	9.5434050

Table 16: CGT of MBP binding protein against MBP. Absolute concentration in each tube.

Vial number	Absolute conc. MBP [nM]	Absolute conc. MBP binding protein [nM]
1	0.00000000	400.000000
2	14.2857143	371.428571
3	28.5714286	342.857143
4	42.8571429	314.285714
5	57.1428571	285.714286
6	71.4285714	257.142857
7	85.7142857	228.571429
8	100.000000	200.000000
9	114.285714	171.428571
10	128.571429	142.857143
11	142.857143	114.285714
12	157.142857	85.7142857
13	171.428571	57.1428571
14	185.714286	28.5714286
15	200.000000	0.00000000

Table 17: Simulated data summary. Simulated data, using the following values: F_{norm} MBP: 656.529406, F_{norm} MBP binding protein: 599.601676, Complex F_{norm} 783.63247, K_d: 22.6235537, quadratic error: 1.17179075

Vial number	Free conc. MBP [nM]	Free conc. MBP binding protein [nM]	Complex conc. [nM]	F_{norm} for interaction	F_{norm} for no interaction	Measured F_{norm} values	St.dev.
1	0.000000	400.0000	0.00000000	599.60167	599.60167	599.60167	10.872361
2	0.849134	357.9919	13.4365802	606.37372	601.71011	604.69699	20.614431
3	1.907776	316.1934	26.6636519	614.14938	603.98073	606.64063	12.088545
4	3.261142	274.6897	39.5960005	623.13374	606.43300	621.80241	9.2011455
5	5.045167	233.6165	52.0976895	633.56373	609.08963	631.74827	9.6376562
6	7.48739	193.201	63.941179	645.67874	611.97726	647.85796	5.3840424
7	10.9886	153.845	74.725625	659.61741	615.12742	657.31628	6.0801058
8	16.2864	116.286	83.713520	675.11734	618.57758	676.86210	2.4931291
9	24.7404	81.8833	89.545247	690.78579	622.37276	683.95987	4.8412432
10	38.5481	52.8338	90.023289	703.02488	626.56744	696.94659	3.2091541
11	59.9017	31.3303	82.955351	706.82183	631.22819	710.66638	4.3381924
12	88.8277	17.3991	68.315102	700.60224	636.43726	707.51716	8.1116311
13	123.153	8.86814	48.274716	687.76137	642.29747	692.10629	1.6718194
14	160.669	3.52652	25.044902	672.28989	648.93904	680.44973	15.039229
15	200.000	0.00000	0.0000000	656.52940	656.52940	656.52946	22.745405

Table 18: CGT of MBP binding protein against MBP. Absolute concentration in each tube.

Vial number	Absolute conc. MBP [nM]	Absolute conc. MBP binding protein [nM]
1	0.00000000	800.000000
2	28.5714286	742.857143
3	57.1428571	685.714286
4	85.7142857	628.571429
5	114.285714	571.428571
6	142.857143	514.285714
7	171.428571	457.142857
8	200.000000	400.000000
9	228.571429	342.857143
10	257.142857	285.714286
11	285.714286	228.571429
12	314.285714	171.428571
13	342.857143	114.285714
14	371.428571	57.1428571
15	400.000000	0.00000000

Table 19: Simulated data summary. Simulated data, using the following values: F_{norm} A: 663.150366, F_{norm} B: 592.785, Complex F_{norm} 794.533639, K_d: 46.6522985, quadratic error: 1.20411805

Vial number	Free conc. MBP [nM]	Free conc. MBP binding protein [nM]	Complex conc.	F_{norm} for interaction	F_{norm} for no interaction	Measured F_{norm} values	St.dev.
1	0.0000000	800.00000	0.00000000	592.785	592.785	592.785	13.33373
2	1.7476699	716.03338	26.8237586	600.21798	595.39112	595.74808	7.648838
3	3.9252735	632.49670	53.2175835	608.75389	598.19772	603.91453	9.161531
4	6.7069084	549.56405	79.0073773	618.61866	601.22884	622.66000	14.65067
5	10.369616	467.51247	103.916098	630.07386	604.51256	622.98971	0.824207
6	15.37551	386.8040	127.481630	643.38554	608.08181	644.15719	7.637243
7	22.53454	308.2488	148.894022	658.71434	611.97555	667.15868	4.025270
8	33.32569	233.3256	166.674309	675.79713	616.24012	672.81596	13.87884
9	50.44563	164.7313	178.125792	693.18160	620.93114	690.84148	7.358567
10	78.19382	106.7652	178.949031	707.11292	626.11596	702.47628	8.6557062
11	120.8166	63.67379	164.89763	712.33453	631.87687	713.69207	8.7669567
12	178.3943	35.53723	135.89133	707.039018	638.31553	715.94844	6.829136
13	246.7437	18.17233	96.113380	694.58543	645.55902	689.59310	2.415099
14	321.5263	7.240630	49.902226	679.11899	653.76831	680.66891	12.04590
15	400.00000	0.000000	0.0000000	663.15036	663.15036	663.15036	15.44031

Raw data of all 1400 simulated interactions can be obtained upon request.

8. EDTA test

In MST experiments exhibiting ligand-dependent changes in initial fluorescence, the cause of this effect needs to be determined through specificity tests. In some cases, the interaction itself is causing the fluorescence changes, allowing data evaluation via initial fluorescence. In other cases, the effect is due to material loss such as adsorption of the fluorescent molecule to labware or protein aggregation.

The EDTA (ethylenediaminetetraacetic acid) test is used for assays with DYE-tris-NTA labeled samples and was performed according to instructions given in MO.Control software. For this, 7 µL of samples 1-3 and 14-16 were centrifuged for 10 min at 14 000 g and 4 °C, before 7 µL 50 mM EDTA (pH 7.4) solution was added to all six samples. Solutions were mixed by pipetting up and down and incubated for 30 min at 37 °C using a heating block. Afterward samples were loaded into Monolith NT.115 Capillaries and sample fluorescence was recorded at 25 °C and 60 % or 100 % LED power, for BLUE and GREEN respectively. Fluorescence intensities of duplicate measurements for the GREEN and the BLUE channel are illustrated in Figure 3. Here, the fluorescence intensities of samples 1-3 and 14-16 are presented before and after the addition of EDTA. As the high affinity of this interaction is dependent on the presence of Ni(II) ions complexed with the NTA molecule, the addition of a chelating agent like EDTA removes the Ni(II) ions from the tris-NTA dye, causing dissociation of the dye from the His-tagged protein. In case of a non-specific fluorescence decrease, the difference in initial fluorescence intensity will remain after addition of EDTA. In case of a binding specific fluorescence decrease, the initial fluorescence of all samples will be nearly identical after EDTA addition, as seen in Figure 3.

In situations where interactions of DYE-tris-NTA labeled proteins with a third molecule are analyzed, the EDTA test is followed by the Control Peptide test. This is to detect fluorescence changes caused by the direct interaction of the third molecule with either the tris-NTA dye or the labeled target protein's His-tag. In this particular assay, only the interaction of the dye to the His-tagged protein was investigated, and the Control Peptide test was therefore not needed.
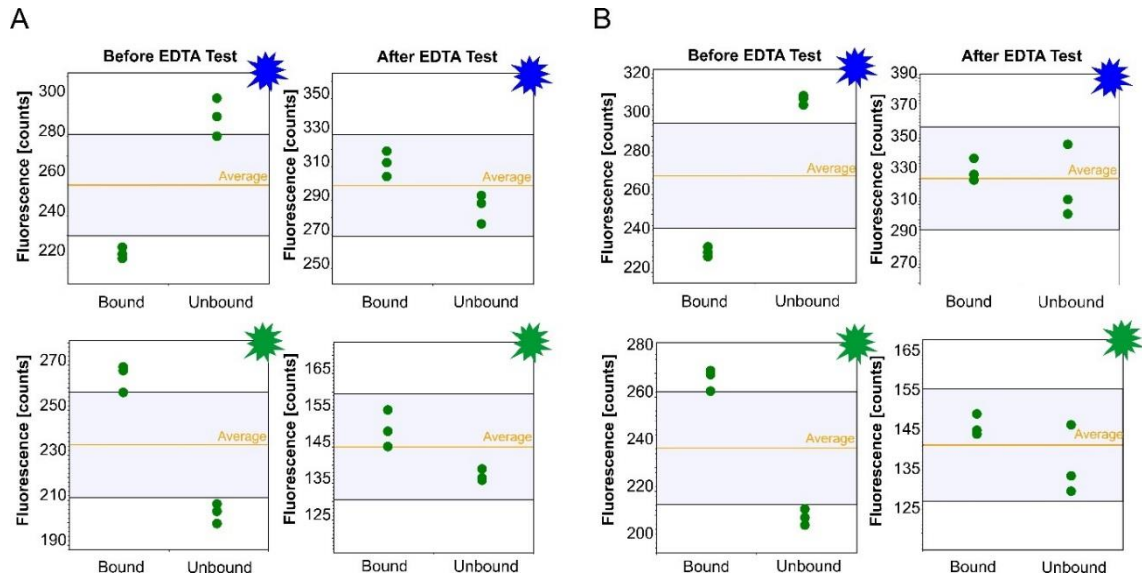
Figure 3: EDTA test. EDTA test to analyze the ligand-induced fluorescence changes observed for BLUE- and GREEN-tris-NTA binding to His6-p38α. Fluorescence intensities before and after addition of EDTA solution are illustrated for both fluorophores, each measured in two biological replicates (A and B). Figure adapted with permission from ref.[131], Scientific Reports.

9. Tested buffers and additives

Table 20: Buffers and additives. Tested buffer additives to investigate robustness of tris-NTA labeling approach. Table adapted with permission from ref.[131], Scientific Reports.

Buffer additive	Tested concentrations	Maximum allowed assay concentration
Histidine	0.25 mM/0.5 mM/1 mM	1 mM
Imidazole	0.25 mM/0.5 mM/1 mM	1 mM
EDTA, EGTA	0.25 mM/0.5 mM	0.5 mM
TCEP	0.25 mM/0.5 mM	0.5 mM*
DTT	0.25 mM/0.5 mM/1 mM/5 mM	5 mM
ß-Mercapto-ethanol	0.25 mM/0.5 mM/1 mM	1 mM
GSH	0.25 mM/0.5 mM/1 mM/5 mM/10 mM	10 mM
GTP, GDP	0.25 mM/0.5 mM/1 mM	1 mM
AMP, ADP, ATP	0.25 mM/0.5 mM/1 mM/5 mM	5 mM
Glycerol	10%	10 %
Zn(II), Co(II), Cu(II)	100 nM	preloaded protein only**
Mg(II)	10 mM	5 mM***
Polyhistidine-tagged ligand	none	none

*NanoTemper Technologies GmbH recommends to avoid the use of TCEP with the red fluorophores in general.

**Zn^{2+}, Co^{2+}, Cu^{2+} ions compete for the binding with tris-NTA fluorophores. Thus, only very low nanomolar concentrations are tolerated in the assay buffer. Additional pretests are required.

***Caution is required when using Mg(II), because the magnesium salts might be contaminated with significant amounts of divalent heavy metal ions (like Zn(II), Co(II), Cu(II)), which might interfere with the labeling.

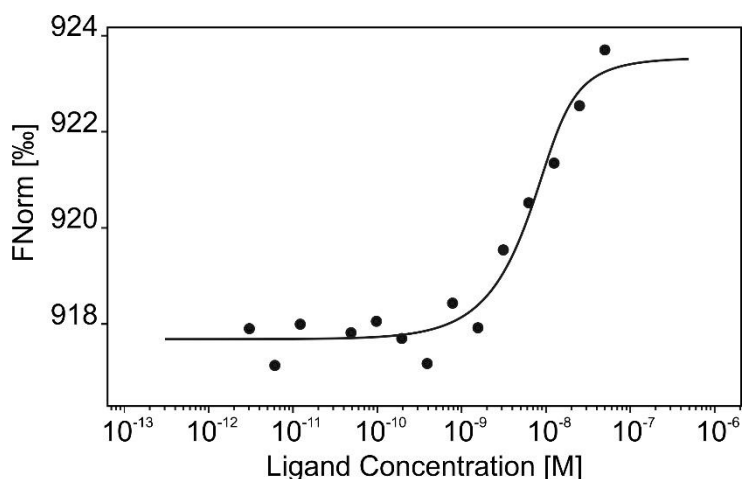## 10. Determination of His$_6$ p38α concentration in cell lysate

Figure 4: His6-p38α against RED-tris-NTA in HeLa cell lysate. Dose-response curve of the interaction analysis of His$_6$-p38α against RED-tris-NTA in HeLa cell lysate, for the determination of protein concentration in the cell lysate. A serial dilution of His$_6$-p38α containing cell lysate was prepared and RED-tris-NTA was added at a constant concentration of 50 nM. Samples were then filled into Monolith NT.115 MST Premium Capillaries and MST experiment was carried out using 50 % LED and medium MST power. Figure adapted with permission from ref.[131], Scientific Reports.

Concentration of His$_6$-tagged p38α in cell lysate was determined as described. Briefly, a 1:1 serial dilution of p38α-mNeonGreen-His$_6$ containing HeLa cell lysate was prepared using non-transfected HeLa cell lysate as dilution buffer. 50 nM of RED-tris-NTA dye was added to all 16 dilution steps, followed by incubation for 30 min at room temperature. Samples were then filled into Monolith NT.115 MST Premium Capillaries and MST experiment was carried out using 50 % LED and medium MST power. MST data was evaluated after 10 sec MST-on time. Resulting dose-response curve is illustrated in Figure 4. The data were fitted using a K_d fit model that describes a molecular interaction with a 1:1 stoichiometry according to the law of mass action. The K_d is estimated by fitting the equation below:

$$f(c) = Unbound + (Bound - Unbound) \times \frac{c+c_{target}+K_d-\sqrt{(c+c_{target}+K_d)^2-4cc_{target}}}{2c_{target}} \qquad (23)$$

Where *f(c)* is the fraction bound at a given ligand concentration *c*, *Unbound* is the F_{norm} signal of the target, *Bound* is the F_{norm} signal of the complex, K_d is the dissociation constant or binding affinity, and the *c_{target}* is the final concentration of target in the assay. For the data set shown in Figure 4, the K_d of RED-tris-NTA for the His-tagged p38α was set to 2.1 nM (as measured for the purified protein) and the *c_{target}* to 25 nM. Taking into consideration the 1:1 dilution step, the concentration of the His-tagged p38α was estimated to be around 50 nM.

11. Overview of experimental parameters

Table 21: Overview of experimental parameters. Table adapted with permission from ref.[131], Scientific Reports.

Interaction	K_d with standard deviation	Amplitude*	S / N ratio**	Reduced χ^{2***}	Number of experiments
BLUE-tris-NTA against His$_6$-peptide	6.7 ± 4.1 nM	8.8	24.9	1.1	3
GREEN-tris-NTA against His$_6$-peptide	4.4 ± 3.7 nM	17.2	34.1	5.2	3
RED-tris-NTA against His$_6$-peptide	3.8 ± 0.5 nM	12.1	61.1	0.2	3
BLUE-tris-NTA against His$_6$-p38α	2.7 ±1.7 nM	100.3	18.6	3.1	3
GREEN-tris-NTA against His$_6$-p38α	6.3 ± 1.7 nM	63.6	34.4	0.3	3
RED-tris-NTA against His$_6$-p38α	2.1 ± 0.8 nM	34.4	91.0	1.4	3
His$_6$-p38α-BLUE-tris-NTA against PD169316	16.7 ± 1.2 nM	19.3	29.6	1.5	3
His$_6$-p38α-GREEN-tris-NTA against PD169316	35 ± 5 nM	17.1	13.5	1.5	3
His$_6$-p38α-RED-tris-NTA against PD169316	24 ± 9 nM	13.8	20.1	3.4	3
BLUE-tris-NTA-MBP-binding protein against MBP	6 ± 2 nM	6.3	12.7	14.9	3
GREEN-tris-NTA-MBP-binding protein against MBP	5 ± 4 nM	18.8	18.4	0.9	3
RED-tris-NTA-MBP-binding protein against MBP	7 ± 1 nM	13.0	42.9	3.8	3
His$_6$-p38α-RED-tris-NTA against SB203580 in HeLa cell lysate	116 ± 0.84 nM	16.1	22.4	3.0	3
mNeonGreen-His$_6$-p38α against SB203580 in HeLa cell lysate	56.8 ± 39 nM	4.6	6.3	26.4	2
RED-tris-NTA toward His$_6$-pUL53 containing cell lysate	not determined	11.3	16.2	0.4	4
RED-tris-NTA labeled pUL53 against pUL50 in *E.coli* lysate	1.8 ± 0.2 μM	10.1	42.3	1.5	3
RED-tris-NTA labeled pUL53 against pUL50 (purified)	1.2 ± 0.5 μM	12.2	26.6	3.8	2

*Response Amplitude $=$ |unbound $-$ bound| $\hspace{4cm}$ (24)

Where unbound and bound are the respective estimated values from the fit. "Unbound" is the plateau at very low concentrations of ligand (also called baseline), while "bound" is the plateau at very high concentrations of ligand (also called saturation).

** The S/N ratio is calculated by dividing the response amplitude by the noise (25). The noise is calculated as the standard deviation of the residuals from the fit.

$$S/N = \frac{\text{Response Amplitude}}{\sqrt{\frac{\Sigma_j(r_j - \bar{r})^2}{n-1}}} \qquad (25)$$

Where r_i and $\bar{r}$ are the residuals of the fit at a given data point or at an average of all residuals, respectively. The number of data points is given by n.

The S/N ratio is a good parameter from which to judge data quality. A value of more than 5 is desirable while a value of more than 12 corresponds to an excellent assay.

*** This value is only calculated for merge sets that contain two or more replicates (26).

$$\chi^2 = \Sigma_i \frac{(m_i - y_i)^2}{\sigma_i} \qquad (26)$$

Where m_i denotes the y-values of the fitted curve, y_i denotes the averaged raw-data y-values and σ_i denotes the standard deviation of the averaged raw-data y-values.

The reduced χ^2 is then defined as (27)

$$\chi^2_{\text{red}} = \frac{\chi^2}{\nu} \qquad (27)$$

With the residual degree of freedom $\nu = n - m$; n is the number of data points and m is the number of parameters that are fitted (four parameters for both, K_d- and Hill-model, except any parameters are fixed).

In MO.AffinityAnalysis the reduced χ^2 can become quite large. The reason for this is that replicates are often very similar. This yields a small standard deviation. Since one divides by these small values, the number can become quite high. Therefore, the absolute value of the reduced χ^2 alone is not a useful parameter from which to judge data quality. It is however very useful for comparing data quality between replicates or comparable samples. In such cases, the smaller χ^2 for one particular dataset in comparison to other datasets, the better the data quality.

12. References:

1.   Vivian, J. T. & Callis, P R. Mechanisms of tryptophan fluorescence shifts in proteins. *Biophysics.* J. **80**, 2093–2109 (2001).

2.   Jiskoot, W. & Crommelin, D. Applications using intrinsic protein fluorescence in *Methods for structural analysis of protein pharmaceuticals* 60 (American Association of Pharmaceutical Scientists, 2005).