



JACOBS  
UNIVERSITY

# **Fluorescent Probes in Biomolecular Systems: Information on Peptide Dynamics and Analyte Uptake into Live Cells**

by

**Amir Norouzy**

A Thesis submitted in partial fulfillment of the requirements for the degree of

**Doctor of Philosophy in Biotechnology**

Approved Dissertation Committee

---

Name and title of Chair

Prof. Dr. Werner Nau

---

Name and title of Committee Member

Prof. Dr. Mathias Winterhalter

---

Name and title of Committee Member

Prof. Dr. Dong-Sheng Guo

---

Name and title of Committee Member

Dr. Anke Meyer

---

Date of Defense: 2014/11/28

---

School of Engineering and Science



## Statutory Declaration

I, **Amir Norouzy**, hereby declare that I have written this PhD thesis independently, unless where clearly state otherwise. I have used only the source, the data and support that I have clearly mentioned. This PhD thesis has not been submitted for conferral of degree elsewhere.

Bremen, January 6, 2015

Signature\_\_\_\_\_



Fausto Cercignani<sup>1</sup>:

**“Your will cannot always choose the path; very often the route is determined by chance or by the will of others.”**

---

<sup>1</sup> Fausto Cercignani (born 1941) is an Italian scholar, essayist and poet.



## Abstract

In this thesis two main projects of my PhD is explained the in chapters 2 and 3 following a pertinent introduction for understanding the projects better in Chapter 1.

In chapter 2, the effect of charge repulsion on homo repeat hexapeptides will be explained. The peptides investigated are homo repeats of either acidic (Glu,Asp) or a basic (Arg, His and Lys) amino acid and labeled at two ends for measuring their end-to-end distance and flexibility by methods explained in great detail in chapters 1 and 2. As a function of pH, negative or positive charges appear on the peptide side chains. The repulsion occurs mainly between the N-terminus and the side chains in basic peptides. Astonishingly, no effective negative charge repulsion among the acidic side chains or the C-terminus with the side chains was observed. This can be explained by the difference in binding affinities of water towards carboxylates versus positively charged groups, which results in screening of the negative charges. Charge repulsion between negative charges was recovered when I changed the solvent from water to a 92.5% methanol/water.

In the second project -that was explained in Chapter 3- an *in vivo* host-dye displacement was carried out: In this project V79 and CHO cells were loaded with a host-dye complex called *p*-sulfonatocalix[4]arene (CX4) / lucigenin (LCG). The fluorescence of LCG is quenched inside the CX4. The cationic analytes acetylcholine, choline and protamine were sent into the cells. The dye displaced with analyte and made a fluorescence turn-on signal. The invented method can be used for assaying the analyte in the cells. My results also show that CX4 facilitate the membrane passage of LCG. Therefore, CX4 can be considered as a carrier.



## **Acknowledgements**

I would like to acknowledge my supervisor Prof. Dr. Werner M. Nau for supervising me and for supporting me to work on my ideas, in particular the application of the indicator displacement assay in live cells. During my stay in his lab I learnt many things, not only scientific but also from new techniques to scientific writing which all will be integral part of my future carrier.

I also thank Prof. Dr. Mathias Winterhalter, Prof. Dr. Dong-Sheng Guo and Dr. Anke Meyer for accepting to be members of my thesis committee.

I would also like to acknowledge Prof. Dr. Thomas Nugent for all his support and advices. I learnt from him how to have a realistic analysis of various situations and problem-solving strategies.

I am grateful to the members of the Nau group, both past and present. I want to especially thank Alexandra Irina Lazar for all her support and kind help.

Special thanks go to my effective and very supportive collaborator and wife Zahra Azizi, as well as my collaborator out of Nau's lab Prof. Dr. Uwe Pischel, Ca'tia Parente Carvalho, Prof. Dr. Roland Benz and Narges Abdali.

For providing cell culture facilities, I would like to acknowledge Prof. Dr. Detlef Gabel.

I would like to acknowledge Thomas Schwarzlose for teaching me how to synthesize Fmoc-  
Dbo

Last but not least, I am thankful to DAAD for my PhD stipends. I am especially thankful to Ms. Nora Pietsch, Ms. Carola Seeler and Ms. Dagmar Hosseini-Razi from the Iran and Iraq department of DAAD, for their great support for me as an Iranian DAAD scholarship holder.



## ***Table of Content***

---

Abstract .....	V
Acknowledgements .....	VII
Table of Content .....	IX
<b>Chapter 1</b> .....	<b>1</b>
Introduction .....	1
1.1. Peptide folding .....	1
1.1.1. Proteins vs peptides .....	1
1.1.2. Limitation in studying peptide folding .....	8
1.1.3. Fluorescence methods for determining peptide folding .....	10
1.1.3.1. Fluorescence-Resonance Energy Transfer (FRET) .....	10
1.1.3.2. Collision-induced Fluorescence Quenching (CIFQ) .....	17
1.2. Macrocyclic-based indicator displacement assays in biology and chemistry .....	19
1.2.1. Sensing citrate in beverages using IDA .....	20
1.2.2. Ion sensing in water samples .....	20
1.2.3. Sensing organic and small biomolecuels .....	22
1.2.4. Multicomponent Indicator Displacement Assay (MIDA) .....	24
1.2.5. Indicator displacement assays in medical researches .....	26
1.2.6. Immobilized indicator displacement assay .....	27
1.3. Using macrocycles for drug discovery, delivery and release .....	30
1.3.1. Macrocyclic hosts and their hosting mechanisms .....	30
1.3.2. Macrocyclic hosts as a drug carrier .....	32
1.3.3. Calixarenes as a drug carrier .....	33
1.3.4. Macrocyclic-based enzyme tandem assay for drug discovery .....	34
1.4. References .....	37
<b>Chapter 2</b> .....	<b>46</b>
The Effect of Electrostatic Charge Repulsion on Peptide Folding .....	46
2.1. Abstract .....	46
2.2. Investigation of oligopeptides by fluorescence-based methods .....	47
2.3. Experimental .....	48
2.4. Results .....	49
2.4.1. Fluorescence lifetime of the FTrp as a probe .....	51

2.4.2. Flexibility change as a function of charge repulsion .....	54
2.4.3. End-to-end distance change as a function of charge repulsion .....	59
2.4.4. Charge repulsion recovery in an acidic peptide.....	62
2.4.5. The effect of charge repulsion on looping activation energy .....	63
2.5. Resolving peptide folding by FRET / CIFQ in tandem .....	64
2.5.1. The electrostatic charge repulsion disfavors the basic peptides looping.....	65
2.5.2. Absence of the charge repulsion effect on acidic peptides folding.....	66
2.6. Conclusion.....	70
2.7. References .....	71
<b>Chapter 3</b> .....	75
Sensing Analytes <i>in vivo</i> Using Macrocyclic-Based Reporter Pairs .....	75
3.1. Abstract.....	75
3.2. Antibody-free analyte sensing.....	76
3.2.1. Fluorescent turn-on as a mean to follow <i>in vivo</i> sensing .....	76
3.2.2. Applying calix[4]arene in cells.....	78
3.3. Results.....	79
3.3.1. CX4 as a carrier.....	81
3.3.2. Loading cells with CX4/LCG complex .....	84
3.3.3. The displacement of LCG with cationic bio-analytes from CX4 inside live cells	86
3.3.4. Endosomal dye-release .....	96
3.4. References .....	97
<b>Chapter 4</b> .....	104
Outlook .....	104
Appendix A) .....	106
1- Conclusions and innovations .....	106
2- Own Contribution .....	107
Appendix B) Curriculum Vitae .....	108
Appendix C ) List of Publications .....	113





# Chapter 1

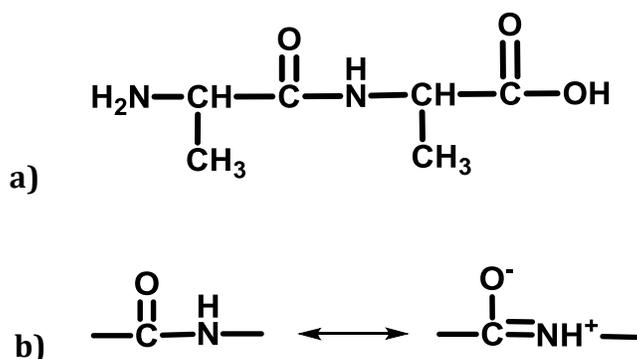
## Introduction

Note: Some parts of this chapter are derived from the review article: "Synthetic Macrocyclic Receptors as Tools in Drug Delivery and Drug Discovery" by Amir Norouzy and Werner M. Nau, published online by *European Drug Target Review*.

### 1.1. Peptide folding

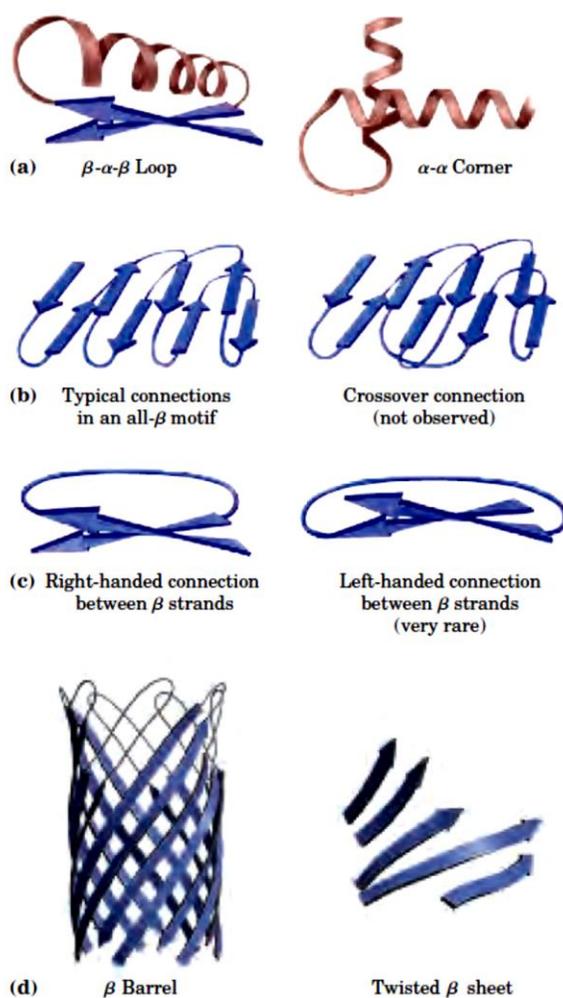
#### 1.1.1. Proteins vs peptides

Proteins are (macro)biomolecules made of small peptide segments. As a part of a protein, a peptide segment has a defined three dimensional-structure, called secondary structure. The peptide segments themselves are made of smaller units, amino acid residues, linked by amide (peptide) bonds (Figure 1.1. a). An amide bond has a dual single/double bond character *i.e.* there is an isomerization between single and double C-N bond (Figure 1.1 b). The amide bond is therefore planar and not free to rotate, while the CH-C and CH-N are single bonds that can rotate freely. The bond angle resulting from rotation around CH-C and CH-N are called  $\psi$  and  $\Phi$  respectively. Peptide and protein are folded by  $\Phi$  and  $\psi$  angle changes.



**Figure 1.1.** a) The structure of a di-alanine peptide. The  $\Phi$  and  $\psi$  angles are varied by rotation around CH-N and CH-C respectively. b) The amid bond isomers.

A description of all covalent bonds (peptide bonds and possible disulfide bonds) linking amino acid residues is called primary structure of a polypeptide chain. As it was mentioned above, secondary structure is a defined arrangement of amino acids in peptide segments.  $\beta$  turns,  $\alpha$  helices and  $\Omega$  loops are examples of secondary structures. Super secondary structures (motifs) are a combination of two or more secondary structures such as:  $\beta$ - $\alpha$ - $\beta$  loop,  $\alpha$ - $\alpha$  corner,  $\beta$  barrel, and twisted  $\beta$  sheets (Figure 1.2).



**Figure 1.2.** Structure of protein motifs; (a) combination of  $\alpha$  helix and  $\beta$  strand. (b) and (c) A bundle of  $\beta$  strands make a layered  $\beta$  sheets. (d)  $\beta$  barrel and twisted  $\beta$  sheets made of two arrangements of  $\beta$  strands. (Reprinted from[1]).

On next levels, tertiary structure refers to the three-dimensional structure of a poly peptide and quaternary structure in a dimer or multi subunit proteins, describes the spatial arrangement of subunits.

While *E-coli* has 3000 proteins, humans have 25000 to 35000 proteins with a unique three-dimensional structure and function. Each type of protein also has a unique amino acid sequence. [1-4]

Unlike proteins, peptides do not have a defined structure. Even peptide segments of a protein in solution do not keep the structure they had in the intact protein. My results represented herein as well as many other studies show the conformation change in short peptides are so fast that two ends of a short peptide can collide as fast as million times per second. This rapid conformation change results from altering  $\Phi$  and  $\psi$  bond angles. The energy for this rotation is simply provided by the internal energy of the peptide solution at room temperature. In the following section, the importance of understanding the protein structure and folding at the peptide level will be discussed. Especially loops and turns which conforming nearly 33% of proteins[1], can be investigated from correspondence peptides.

### **Peptide and Protein Folding**

The forces drive protein folding are the same for peptide folding. Nevertheless, proteins have defined three-dimensional structures while short (oligo) peptides lack any defined structure or even conformation. Amino acids in peptides move rapidly and quite freely in a large spatial area while amino acids in proteins are fixed as part of a defined structure and have some vibrations rather than spatial movement. To understand this, protein folding and the involved forces for the folding and maintaining their three dimensional structure will be explained below and is then compared with peptides.

An amino acid sequence dictates the three-dimensional structure of a protein. These are flexibility, intra protein interactions between neighboring residues and their possible  $\varphi$  and  $\psi$  angels that determine the secondary structure. Therefore, protein structure –and

function– is not irrelevant of its amino acid sequence as well as its peptide segments. There are additional biochemical evidences: First, proteins with different amino acid sequence have different structure and function. Second, many of the diseases are caused by a point mutation, results in loss of protein's function. And finally, proteins with a similar function from different species have often similar amino acid sequence. Ubiquitin is a good example here. Ubiquitin is a 76-residue and highly conservative protein, involved in regulating the degradation of other proteins. This clearly reveals the importance of primary structure in protein function. However, in polymorphic proteins, even up to 30% variant in amino acid sequence has little or no effect on the proteins' function.[1] Although the amino acid sequence in some regions of the primary structure might vary considerably without affecting biological function, most proteins contain crucial regions that are essential to their function and whose sequence is therefore, conserved. The fraction of the overall sequence that is critical, varies from protein to protein, complicating the task of relating sequence to three-dimensional structure, and structure to function.

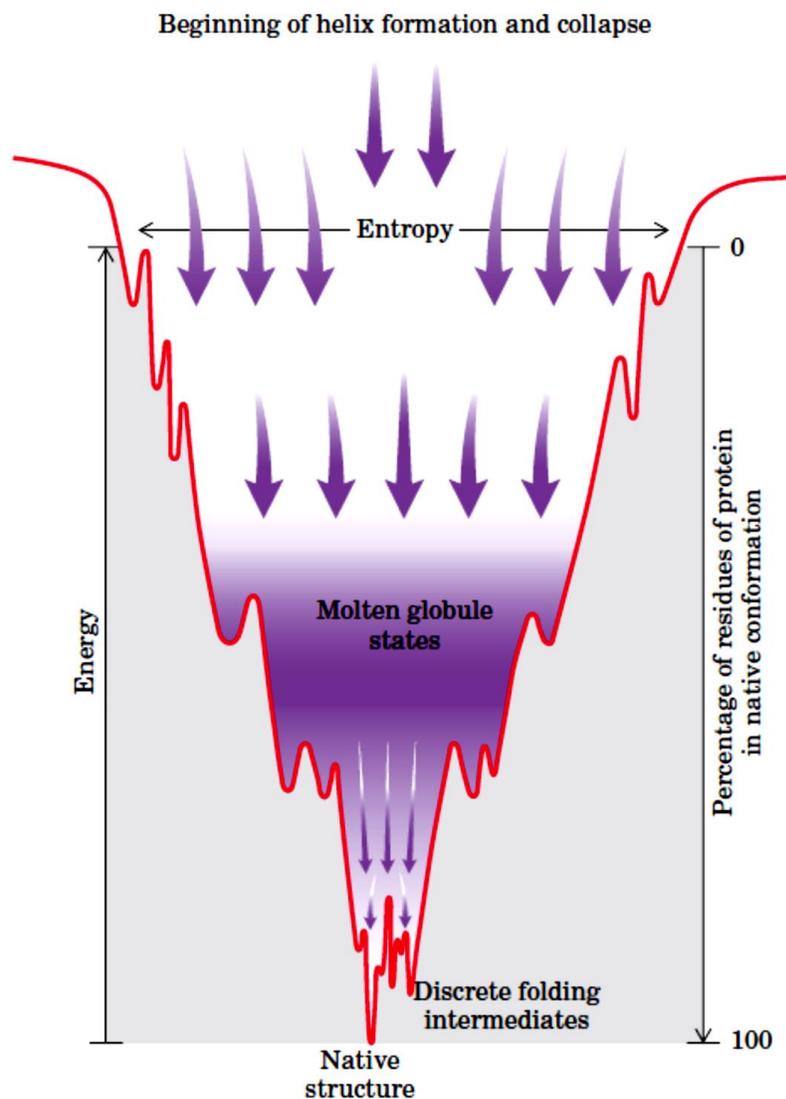
Exactly how an amino acid sequence determines the three-dimensional structure is not yet understood in detail, nor can we always predict function from a sequence. 25 % similarity in amino acid sequence could result in same functional and somehow structural features. Some sequences have discovered dictating a defined structure and function.[4] A number of similar substructures (domains) occur in many functionally unrelated proteins. Therefore, studying protein folding from the peptide level might help to understand this complicated relationship.

The stability of a protein is defined as the tendency to keep its native structure. Proteins are not structurally extremely stable; the  $\Delta G$  difference between the folded and unfolded states in a protein is in the range of only 20 to 65 kJ/mol. Recall that breaking one hydrogen bond needs nearly 23 kJ/mole. Folding of a protein results in a well- defined three-dimensional arrangement. This highly ordered structure is unfavorable entropically. Moreover, theoretically, the maximum number of hydrogen bond with water is in a fully unfolded protein. Therefore unfolded state must had been more favorable. Then the questions are why proteins fold? and what the driving force for protein folding is. Disulfide

binds between cysteine residues, inter molecular hydrogen bonds, ionic and hydrophobic interactions keep a protein folded.

To understand how polypeptides acquire and keep their native secondary and tertiary conformation, we need to understand these weak interactions. Just to give a comparison, 200-460 kJ/mol are required to break a covalent bond on average, whereas a weak interaction takes about 4-30 kJ/mol. However, a disulfide covalent bond can stitch the separated parts in a polypeptide or in a protein, but the numbers of such covalent bonds are negligible compare to number of weak interactions. Therefore, weak interactions play a major role in folding of proteins despite their little strength.

Herein the influence of hydrogen bonds and ionic interaction is compared to the hydrophobic interaction during protein folding in order to understand the driving force of peptide/protein folding. Proteins in their native conformation have the maximum number of weak interactions and the lowest amount of free energy (Figure 1.3)



**Figure 1.3.** A description of protein folding depicted as a free-energy funnel. A protein has maximum entropy when it is 100% unfolded. As a function of folding, moving down the funnel, the protein loses its entropy compensated by lowering its Gibbs free energy. (Reprinted from [5])

In the unfolded state, hydrogen bond donor groups of a protein make hydrogen bonds with water. During folding, these hydrogen bonds break and new ones form within the protein or peptide. The net free energy difference between the folded and unfolded state is therefore; near zero or it has even a positive value when the number of newly formed

hydrogen bonds is less than the number of broken bonds. Therefore the intramolecular hydrogen bonds are not favoring folding significantly.

In the unfolded state, hydrophobic groups form highly ordered of arranged water molecules around them, which is not entropically favorable. Water molecules in pure water make a hydrogen-bonding network. The hydrophobic residues disrupt the network, and the highly ordered H<sub>2</sub>O is structured around them. To avoid such an entropic penalty, proteins tend to fold in such a way the hydrophobic groups compact together inside the protein core and consequently, access to water molecules is minimized. The water molecules that had surrounded the hydrophobic groups therefore release and their entropies increases. This compensates the entropic penalty of protein folding mentioned earlier. Therefore the protein folding is an entropic-driven process.

Salt-bridges –between oppositely charged amino acids– in protein core makes protein significantly stable. But the salt-bridge at protein surface might not have such an effect. The hydration shell around two ionic amino acid side chains should be broken in order to form a salt-bridge. The released energy is less than the broken hydration energy. This is of course not an energetically favored process. The driving force is the increase of entropy due to the release of the structured water molecules around the ionic groups. However the net  $\Delta G$  shows the value of nearly zero.

Hydrophobic interactions are undoubtedly the most important and frequent forces stabilizing a protein conformation. The interior of a protein is generally a densely packed core of hydrophobic amino acid side chains. It is also important that any polar or charged group in the protein interior has a suitable partner(s) for hydrogen bonding or salt-bridge. One hydrogen bond seems to have a negligible effect on stabilizing a native structure, but if it does not find its proper partner in the hydrophobic core of a protein, then it can drastically destabilize the protein structure thermodynamically. Hydrogen bonds are formed cooperatively in proteins; means formation of one hydrogen bond facilitates the formation of other bonds. [1]

*Oligopeptides world* on the other hand is completely different from *protein world*. Compare to polypeptides, the oligopeptides made of mainly hydrophobic amino acids are poorly soluble. The few numbers of amino acids forming an oligopeptide does not allow forming a hydrophobic core surrounded by ionic or polar amino acids. Therefore, all amino acid residues are always partially or fully exposed to water. Consequently, the oligopeptides do not obtain a defined structure based on the hydrophobic collapse discussed for proteins. My results here and many other studies show that end-to-end of a short peptide chain collides million times per second.[5-8] Peptides therefore, are structure-less biomolecules that are changing their conformations swiftly. The term *stability* as a tendency of keeping a natural structure is not applicable to peptides. Nevertheless, amino acid sequence affects peptides dynamic; for example bulky amino acids in a peptide sequence bring some spatial hindrance results in expansion of peptides.[6] Presence of two hydrophobic amino acids in a peptide sequence however do not dictate its structure but may form more dominant sub structure in which the two hydrophobic amino acids form a hydrophobic dimer. These temporal substructures are in the earliest stage of protein folding. Because protein folding starts while it is being synthesized by a ribosome. [9] Therefore, studying protein folding from peptide level is of great importance but also suffers from some limitations:

### **1.1.2. Limitation in studying peptide folding**

Protein structure is experimentally determined by well-known NMR and X-ray crystallography methods, while the average end-to-end distance between two points in peptides or proteins is determined by fluorescence-resonance energy transfer (FRET) and electron paramagnetic resonance (EPR) methods. The dynamics (flexibility) of a protein is investigated by stop-flow and for peptides by fluorescence methods such as the one based on collision-induced fluorescence quenching (CIFQ). [6] Simulation methods on the other hand could suggest both the structure and the dynamics of a protein or peptide theoretically, which alone is not capable of reliably predicting structure yet.

## X-Ray crystallography

According to optical principles, the uncertainty of an object's position is related to the wavelength of radiation.[10] A covalent bond has a length of approximately 1.5 Å and the shortest-wavelength visible beam has a wavelength of 3500 Å. Therefore, molecules are not visible by light microscopes. The X-ray has a wavelength of 1.5 Å. Therefore, X-ray crystallography is used to determine the structure of several proteins. The crystal of a sample is exposed to X-rays and the electron density map is obtained. From the obtained diffraction pattern of electron density (not the position of atom nucleus) the three dimensional structure of the sample is calculated mathematically. The provided crystals of a protein contain 40 to 60% water, which is necessary for the crystal structure to resemble the protein native structure. X-ray crystallography is a powerful and a highly-used technique. However it suffers from low accuracy of determining positions of atoms in low resolution electron density maps, difficulty in providing the crystals of peptide as well as crystalizing the membrane bound protein, and distinguishing between similar amino acids such as Leu, Ile, Thr and Val which have the comparable size and shape. Crystallizing short peptides are rarely successful, which indicates the inappropriate of X-ray crystallography method in studying short peptides' structure. [2] Even if one obtains a crystal of a short peptide, it will be just one conformation at the time out of millions.

## Structure Determining with 2D NMR

2D NMR can be used to determine the structure of polypeptides. In this method, the hydrogen atoms that are less than 5 Å apart can be detected. [10] Together with 2D NMR data and covalent bond distances, angles, group planarity, chirality, and van der Waals radii one can calculate possible structures of a peptide. Since the interproton distance measurements are imprecise, the unique structure of a peptide is not obtained. Due to several possible conformation of an oligopeptide and several hydrogen bonds forming and breaking during peptide folding, determining a dominant conformation for an oligopeptide has also some technical difficulties.

As we will see in the following pages, fluorescence-based methods are widely used for measuring the average end-to-end distance as a structural scale, as well as flexibility of short peptides.

### 1.1.3. Fluorescence methods for determining peptide folding

Since short peptides do not obtain a three dimensional structure at room temperature, the term “peptide folding” refers to rapid conformational changing rather a directed conformational changes toward a defined structure. These conformational changes are quantified by labeling tow points of a peptide with fluorescent probes, followed by measuring the peptide average point-to-point distance and point-to-point collision rate. Usually points of interests are at the two ends of a short peptide chain, therefore; **average end-to-end distance**[11] and **end-to-end collision rate** [6] is known as a gauge for describing peptide folding. The average end-to-end distance represents structural aspect of a peptide while end-to-end collision rate represents its flexibility.

Two ends of a peptide collide within a nanosecond to a maximum of microsecond time-range. [12] Short peptides with different amino acid sequence do not behave identically *i.e.* they have different end-to-end distance and collision rate. I measured average end-to-end distances and end-to-end collision rates based on Fluorescence-Resonance Energy Transfer (FRET) and Collision-induced Fluorescence Quenching (CIFQ) respectively.

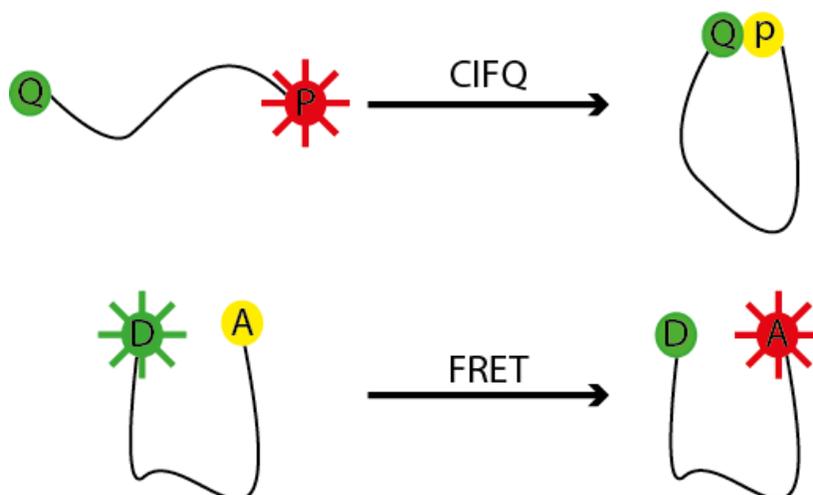
#### 1.1.3.1. Fluorescence-Resonance Energy Transfer (FRET)

In FRET, the energy of an excited **donor** probe is transferred to an **acceptor** probe (Figure 1.4 bottom) [13-15]. The rate of FRET is inversely proportional to the sixth root square of the distance between the donor and the acceptor:

$$K_T = \frac{1}{\tau_D} \left( \frac{R_0}{R} \right)^6 \quad (1.1)$$

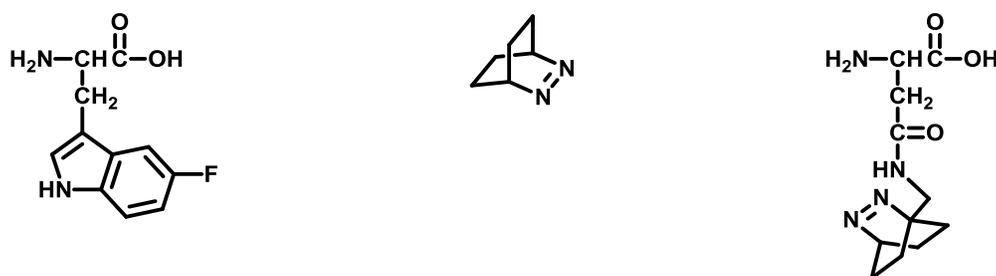
$K_T$  is the rate of FRET and  $\tau_D$  is the fluorescence **lifetime** of the donor and  $R_0$  is the Förster radius, and  $R$  is the distance between the donor and the acceptor.

“The lifetime ( $\tau$ ) of a fluorophore is the average time between its excitation and return to the ground state.”[13]



**Figure 1.4.** Schematic presentation of FRET and CIFQ, Top: The excited probe (P) is quenched by a quencher (Q) upon collision which reveals the dynamics of the chain. Bottom: The excitation energy of the donor (D) is transferred to the acceptor (A)

We attached 5-Fluoro-L-tryptophan (FTrp, Figure 1.5) to the C-terminus of the peptides as the donor. FTrp was excited at 280 nm, and its lifetime was measured at 350 nm. The excited FTrp transfers its resonance energy to the acceptor 2,3-diazabicyclo[2.2.2]oct-2-ene (DBO, Figure 1.5), which results in a decrease of the FTrp lifetime (quenching).



**Figure 1.5.** Chemical structures of 5-Fluoro-L-tryptophan (FTrp, left), 2,3-diazabicyclo[2.2.2]oct-2-ene (DBO, middle) and Asparagine decorated with DBO (Dbo, right).

Förster radius ( $R_0$ ) is a merely constant and defined distance for a donor/acceptor pair, and has a value of angstrom (Equation 1.2).

$$R_0 \propto (K^2 Q_D J)^{\frac{1}{6}} \quad (1.2)$$

$K$  is the angle between the donor emission dipole and acceptor absorption dipole,  $Q_D$  is the quantum yield of the donor, and  $J$  is the spectral overlap integral.[13]

In Förster distance,  $R_0$ , the rate of fluorescence decay equals to rate of FRET and 50% of total donor's energy is being taken by the acceptor. The amount of transferred energy through FRET is inversely proportional to sixth root of the distance between donor and acceptor (Equation 1.3). Thus the amount of FRET is significant when the distance between a donor and its acceptor (end-to-end distance in peptides) ( $R_{DA}$ ) does not exceed  $1.5 R_0$ .

$$E = \frac{R_0^6}{R_0^6 + R_{DA}^6} \quad (1.3)$$

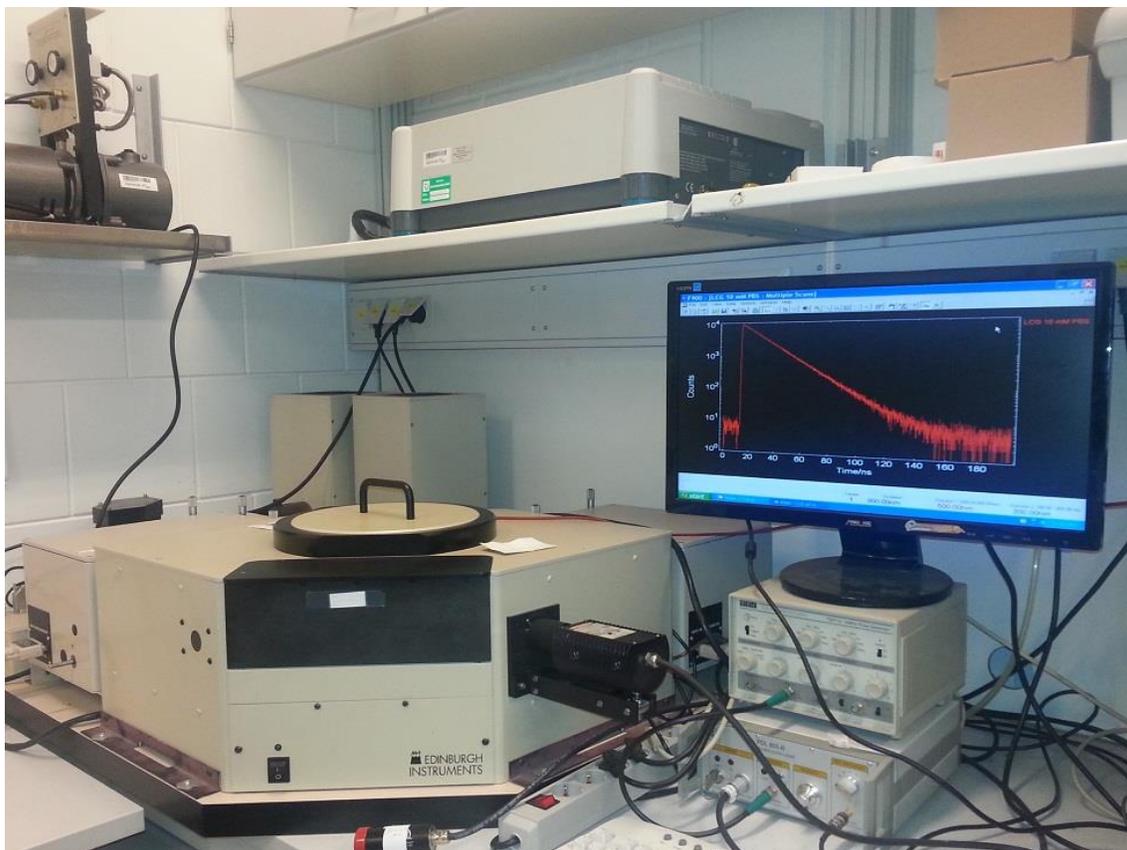
By rearranging equation 1.3 to equation 1.4 the average distance between donor/acceptor is calculated.

$$R = R_0 \left( \frac{1}{E} - 1 \right)^{\frac{1}{6}} \quad (1.4)$$

The transferred energy can be determined experimentally by both time decay (lifetime,  $\tau$ ) and intensity decay ( $I$ ) of a donor:

Fluorescence lifetimes of the fluorophores (FTrp and DBO) were measured for calculating FRET rate and CIFQ rate using a time correlated single photon counting (TCSPC) instrument (Figure 1.6). FTrp was excited by using a PicoQuant pulsed LED (PLS-280,  $\lambda_{ex} = 280$  nm,  $\lambda_{obs} = 350$  nm, fwhm *ca.* 450 ps). The energy of the excited FTrp transfers to DBO via FRET (Figure 1.4). The rate of the energy lost is known as FRET rate. DBO was excited by using a PicoQuant diode laser LDH-P-C 375 ( $\lambda_{exc} = 373$  nm,  $\lambda_{obs} = 450$

nm, fwhm ca. 50 ps). The excited DBO on the other hand is quenched upon collision with the FTrp (Figure 1.4).



**Figure 1.6.** A time correlated single photon counting (TCSPC) for measuring fluorescence lifetimes.

Trp shows a biexponential decay *i.e.* it has two lifetime decays ( $\tau_1 = 3.1$  ns and  $\tau_2 = 0.53$  ns at 25 °C and pH = 7.0). But FTrp shows a monoexponential decay ( $\tau = 1.96$  ns). [7] FTrp was preferred over Trp because mono exponential decay is calculated more accurately. However side chains of some amino acids such as imidazole of His can quench the FTrp and makes it fluorescence decay biexponential. For biexponential decays, **average lifetime** has to be calculated. The average lifetime was calculated by Equation 1.5. The obtained average lifetime was used to calculate the transferred energy by equation 1.6, in which  $\tau_{DA}$  and  $\tau_D$  is the lifetime of FTrp ( $W'$ ) in double-labeled and single labeled, reference, peptides (*i.e.*  $W'X_6Dbo$  and  $W'X_6$ ) respectively .

$$\tau_{avg} = \sum_i^n \alpha_i \tau_i \quad (1.5)$$

$$E = 1 - \frac{\tau_{DA}}{\tau_D} \quad (1.6)$$

For given lifetime values of FTrp:  $\tau_1$ ,  $\tau_2$ ,  $a_1$  and  $a_2$  ( $a$  is the relative percentage of the corresponding  $\tau$  participating in a lifetime decay) from the TCSPC,  $\tau_{avg}$  was calculated by:

$$\alpha_1 = \frac{a_1/\tau_1}{a_1/\tau_1 + a_2/\tau_2} \quad (1.7)$$

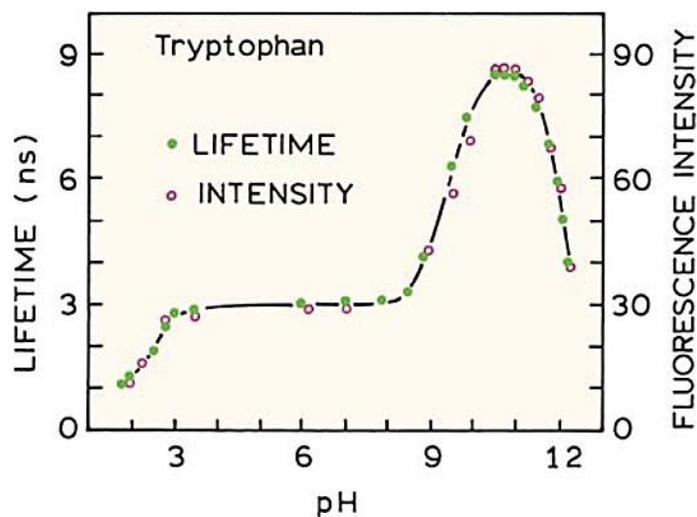
$$\alpha_2 = \frac{a_2/\tau_2}{a_1/\tau_1 + a_2/\tau_2} \quad (1.8)$$

$$\alpha_1 + \alpha_2 = 1 \quad (1.9)$$

$$\tau_{avg} = \alpha_1 \tau_1 + \alpha_2 \tau_2 \quad (1.10)$$

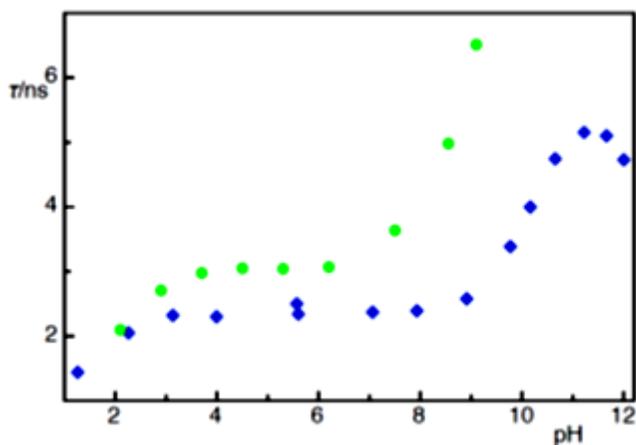
For monoexponential decay,  $a_1 = a_2 = 100\%$ .

Trp or FTrp not only show biexponential decay traces, but their lifetime is also changing as function of pH (Figure 1.7). Below pH 3 the indole ring in tryptophan is quenched by the carboxylic group, from pH 3-8 the zwitterion tryptophan shows a constant average lifetime of approx. 3 ns. Protonated amine group ( $-\text{NH}_3^+$ ) of the tryptophan quenches the indole ring. From pH 9 to 10 the amino group lose proton which results in fluorescence enhancement (lifetime increase). At pH values higher than approx. 10 the hydroxide ion quenches the indole ring.



**Figure 1.7.** Relative fluorescence intensity and mean lifetime of Tryptophan as a function of pH ( Reprinted from [13]).

The lifetime of FTrp or Trp in pure water does not change from pH 3 to 9 (Figure 1.7) but high concentration of salts can change this lifetime trend. For example we found out that in the presence of  $\text{MgSO}_4$ , the lifetime of Trp starts increasing around pH 7 (Figure 1.8). In the presence of  $\text{MgSO}_4$  the  $pK_a$  shifts to a lower pH. Therefore, the possible effect of all solute has to be discovered prior to measuring lifetime or fluorescence study.



**Figure 1.8.** Lifetime of Trp as a function of pH in the presence (green) and absence of  $\text{MgSO}_4$  (blue).

In order to calculate the end-to-end distance of the peptide, the lifetime of a reference peptide (single-labeled with FTrp, *i.e.* W'X6) has to be measured. The average length is given by equation 1.4. The Förster radius ( $R_0$ ) is 9.5 Å for FTrp/Dbo pair in water at pH 7. Because of Trp lifetime variation as a function of pH, the Förster radius has to be corrected at each pH value according to FTrp lifetime in the reference peptide.

In a peptide labeled with a donor and an acceptor, the distance between the donor and the acceptor is changing due to peptide folding. The use of donors with shorter lifetimes, results in a more accurate end-to-end measurement because distance distribution between the donor and the acceptor is limited by the donor's short lifetime. FTrp as a donor has a lifetime near 2 ns while peptide folding has a rate of hundreds of ns to few  $\mu$ s. [12, 16] within 2 ns the distance between DBO and FTrp does not vary significantly. Trp is also a natural amino acid and is not highly occurring in natural proteins,[1] which makes it a good candidate in protein fluorescent studies. Trp can also be inserted into a protein by site-directed mutagenesis. FRET is being used to determine the distance between many sites in proteins. Trp is used as a natural donor for measuring FRET in proteins. [17-19] DBO also has excellent characteristics as an acceptor. It is a stable chromophore against degradation, temperature, light and hydrolysis. DBO is soluble in water and organic solvents. [20] Some organic solvents such as acetonitrile and 1,4 dioxane do not significantly quench the Dbo. Acetonitrile and 1,4 dioxane are also water soluble, therefore: mixing them with water can make a water/organic homogenous solution. The solution has a lower dielectric constant compare to the pure water. Dioxane, due to its high hydrophobicity has already been used to study the solvent effect on fluorescent properties of number of chromophores [21, 22] and immobilized acetylcholinesterase. [23, 24] Tyrosine (Tyr) and Trp are the only two standard amino acids with fluorescent properties. Trp has a higher transition dipole spread than Tyr. Therefore, it is preferred over Tyr to be used with DBO (donor/acceptor pair) for FRET.

The emission spectra of Trp and absorption spectra of DBO overlap largely and by labeling peptide with DBO, the peaks overlapping are still significant (Figure 1.9).

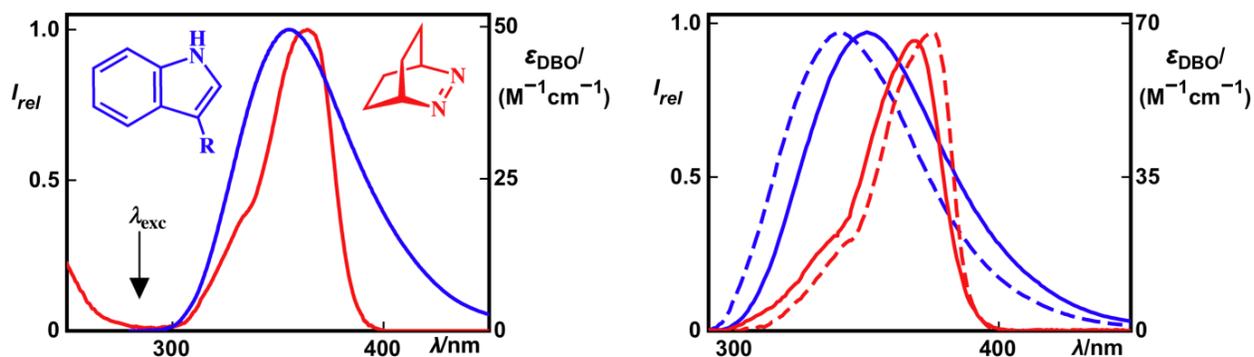


Figure 1.9. Spectral overlap for the Trp emission and DBO absorption as a FRET pairs (left) and for the pair of peptides, W-P<sub>6</sub>-NH<sub>2</sub>/P<sub>4</sub>-Dbo-NH<sub>2</sub>, in water (solid lines) and in propylene glycol (dashed lines).

DBO has additional excellent characteristics to be used in peptides. It is small in size, it has high fluorescence quantum yield, and has a long fluorescence lifetime (505 ns in aerated D<sub>2</sub>O). [5] Many DBO derivatives have already been synthesized for different applications [25] and it has been applied to several peptides. [5, 6]

### 1.1.3.2. Collision-induced Fluorescence Quenching (CIFQ)

The flexibility of peptides plays a crucial role in their biological activity [26], in binding activity, [27] in antigenicity, [28] and in enzymatic activity. [29] The peptides' flexibility is measured through CIFQ.

The same double-labeled peptides used in FRET experiments could be also used together with a proper reference peptide in CIFQ measurements. The reference peptide is Dbo-labeled peptide (X<sub>6</sub>Dbo). The flexibility of some standard amino acids was described in detail by Nau and coworkers in 2003, by applying CIFQ to short peptides. [6]

End-to-end collision of a polypeptide chain, takes between ns to μs [5, 30] and DBO has the lifetime around 300 ns in aerated water, which is long enough for the collision to occur, especially for the short peptides studied herein which have collision rate shorter than DBO lifetime (explained above). The collision rates are reported by collision rate constant ( $k_q$ ).

The  $k_q$  values were calculated from the lifetimes of the double-labeled peptides ( $\tau$ ) and their corresponding reference peptides ( $\tau_0$ ) following equation 1.12.

$$k_q = \frac{1}{\tau} - \frac{1}{\tau_0} \quad (1.12)$$

Huang and *et al* have measured the  $k_q$  of short homo repeat peptides. Amino acids were ordered according to flexibility they brought to the peptides:

Gly > Ser > Asp; Asn; Ala > Thr; Leu > Phe; Glu; Gln > His; Arg > Lys > Val > Ile > Pro [6]

No relation between flexibility and hydrophobicity of the amino acids was reported. Larger amino acids are less flexible; on the contrary, small amino acids such as Gly or Ser are the most Flexible. Branched-amino acids reduce the flexibility, while  $\beta$ -hydroxyl substitutions are remaining flexible. It was revealed that some charged-amino acids could reduce the flexibility because of charge repulsion. The charge repulsion was not studied in great details. The measured  $k_q$  values in this thesis are in good agreement with the reported values. For the first time in our group, I used CIFQ and FRET in parallel to study peptides flexibility and average end-to-end distance as a function of charge repulsion (Chapter 2).

## 1.2. Macrocyclic-based indicator displacement assays in biology and chemistry

Indicator displacement assay (IDA), [31] is known as an analytical method works based on displacement of an indicator (guest) from its host (receptor) by an analyte (equation 1.13):



The indicator guest binds reversibly to its host (equation 1.14)



The binding strength is:

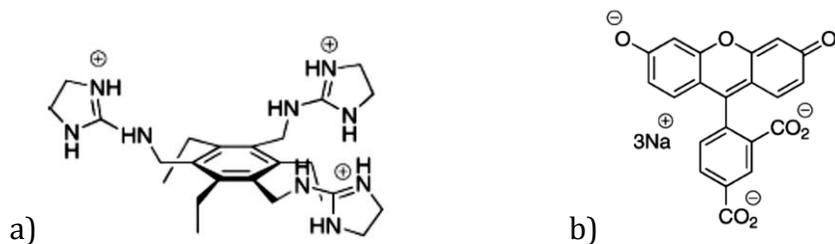
$$k_a = \frac{[\text{Host/Guest}]}{[\text{Host}] \times [\text{Guest}]} \quad (1.15)$$

The binding strength in biology is approximately in the range of micro molar ( $10^6 \text{ M}^{-1}$ ). Therefore, host–indicator (host–dye) binding strength in the micro molar range is suitable for biological applications. If a host–indicator binding constant is greatly larger than this, the binding constant of an analyte has to be even higher for the displacement. Consequently, few analyte could be monitored. On contrary, a very weak binding constant is also prone to competing with large number of non-analyte competitors that are omnipresent in any biological samples like blood, cell cytoplasm or food. For sensing with high sensitivity and selectivity, host-guest binding constant in biological range is preferred and host-analyte binding constant has to be at least two orders of magnitude stronger. The stronger binding constant between host and analyte grants an efficient competition in the favor of analyte and consequently guest (indicator) release (equation 1.13). When an analyte has a weaker binding constant to a host than its indicator, the assay is not impossible but higher concentration of analyte is needed for indicator release which means lower sensitivity for analyte sensing. Binding constant is not the only determinant factor for a successful IDA: The free indicator undergoes an optical signal. The signal strength is also important. For example if an indicator is fluorescence dye, low quantum yield of the dye, low extinction coefficient and its poor photo stability can also be a bottle neck for an

IDA method. In the following part, different aspects and applications of IDA will be reviewed.

### 1.2.1. Sensing citrate in beverages using IDA

There are many reports on applying IDA for sensing ions and other analytes in different samples ranging from water [32] to cells. [33, 34] For this purpose, different reporter pairs (hosts and guests) are needed for sensing the analyte. Host-guest interaction can be ionic or hydrophobic (explained in 1.2.2). For example 1,3,5-trisubstituted-2,4,6-triethylbenzene receptor (host) (Figure 1.10a) binds to carboxylate groups by its imidazole rings through ionic interactions.[35] 5-carboxyfluorescein indicator guest (Figure 1.10 b) bound to the receptor for sensing citrate –that has three carboxylate groups– and is an additive in many beverages. The absorbance of the dye enhanced at around 500 nm upon binding to the host. This is an example of colorimetric IDA, which let citrate recognizing with even naked eyes.

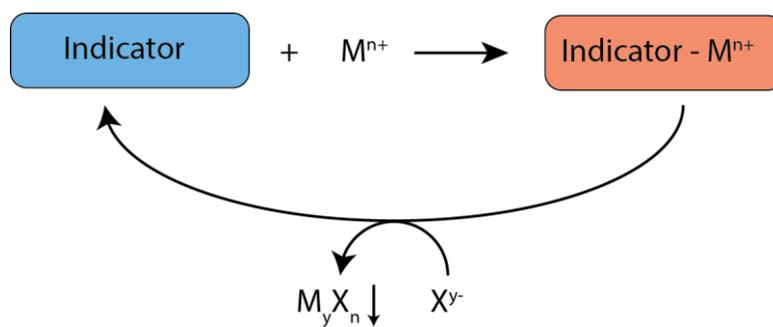


**Figure 1.10.** Structure of a) 1,3,5-trisubstituted-2,4,6-triethylbenzene (TSTHB), a receptor, and b) 5-carboxyfluorescein, an indicator, for sensing carboxylate through IDA

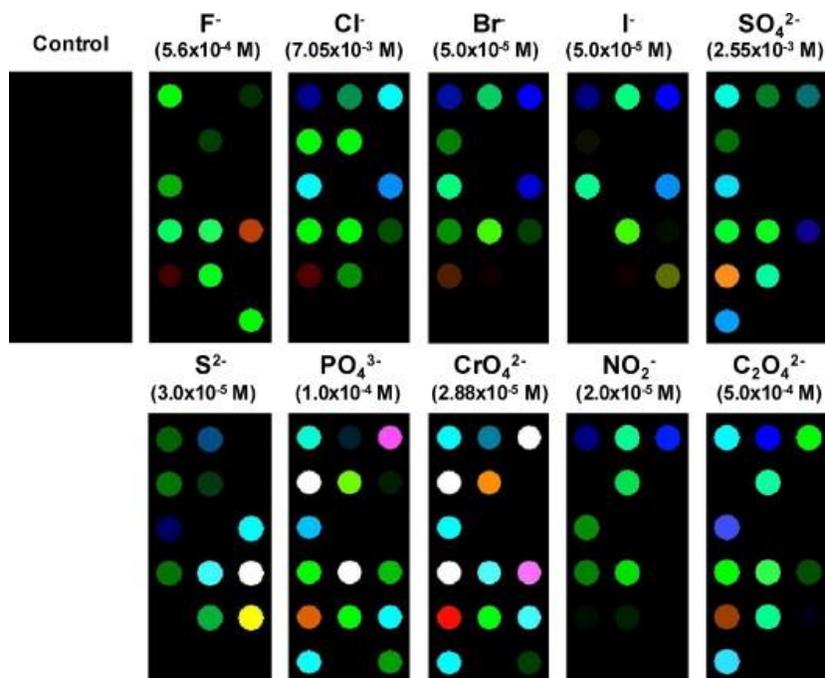
### 1.2.2. Ion sensing in water samples

Analyte competition with a guest for binding to a receptor can also be through hydrogen binding. For example nitrate was sensed using a reporter pair by competition between nitrate molecules and the indicator for making hydrogen bonds with a host. [36] IDA has shown precise and specific analysis results in real world. Sensing anions with a same charge or comparable size is challenging in industry and biology. Some of the common ions are important for cell functioning such as chloride or phosphate. Chloride is the zwitterion of sodium and potassium that keep the cell membrane polarized and

transduce a nerve impulse. Phosphate is crucial for DNA and RNA synthesis and is in the heart of signal transduction through kinase/phosphatase. Nitrate is needed for plant growth and etc. Therefore, measuring the anions concentration accurately with cheap and easy-to-use kits is interesting for biological purposes. Moreover, ions are among industrial pollutants and measuring their concentration has environmental importance. Feng and *et al* have introduced a colorimetric IDA for sensing ten different ions in water samples with micro molar accuracy.[32] The assay method works based on chelating cations with colorimetric indicator and used for a real water sample. The developed method showed high selectivity and sensitivity in micro molar range. This unique assay was not recorded by UV-Vis instrument; instead it worked based on precipitating an indicator in the presence of an ion and re-solubilizing by adding a competitor ion (Figure 1.11). The color change in these solubility variations are a finger print for ions. Figure 1.12 represents some examples of the finger print which analyzed by a standard statistical method. During IDA, no pH change was observed that ensured the integrity of the absorption spectra and revealed that the color changes are solely due to ions competition for binding to the indicator. The detection system worked in different temperatures ranging from 10 to 40 °C.



**Figure 1.11.** Schematic ion sensing by indicator chelating.



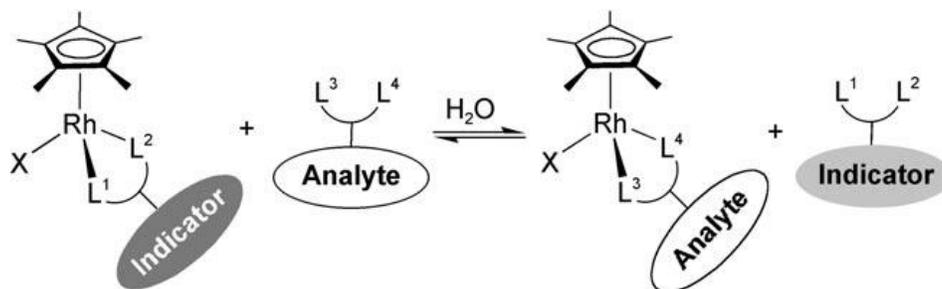
**Figure 1.12.** Examples of fingerprint of ten anions with different size and charge (reprinted by permission from [32]).

Nevertheless of the mentioned advantages, the method suffers from some limitations: First, the anion concentration has to be three fold higher than S/N to overcome background color. Second, however interfering ions such as nitrate, sodium or potassium was not observed but nitrate in the form of nitric acid or sodium in the form of sodium hydroxide may change the pH drastically that can affect the measurement. Therefore, the pH has to be adjusted. The introduced method was also worked in a real swage sample. This is an example of IDA's real application for determining water quality and inorganic contaminations.

### 1.2.3. Sensing organic and small biomolecuels

IDA application is not limited to inorganic ions sensing but it is also being used for organic compounds. Reported by Severin and coworkers, amino acids, sugars and

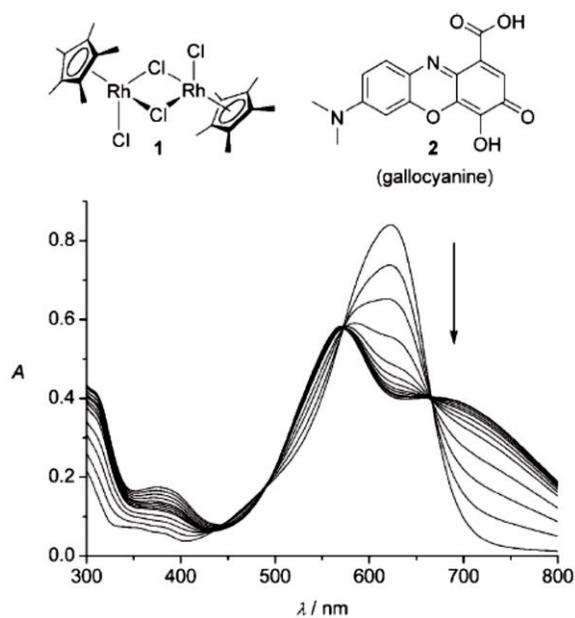
aminoglycosides have been sensed by using  $[(\text{Cp}^*\text{RhCl}_2)_2]$ / gallocyanine<sup>2</sup> reporter pair, where  $[(\text{Cp}^*\text{RhCl}_2)_2]$  worked as a host and gallocyanine as an indicator. [37]



**Figure 1.13.** Schematic representation of an indicator-displacement assay based on an organometallic  $\text{Cp}^*\text{Rh}$  complex. L=donor group. (reprinted by permission from [37] )

This is an example of applying IDA with an organometallic based host ( $\text{Cp}^*\text{Rh}$ ). The assay method is simple but the host donor groups (Figure 1.13) has to be designed for a particular analyte. Nevertheless, the introduced method works selectively for sensing simple amino sugars (galactosamine and mannoseamine) and amino glycosides. Glycoamines have several biological roles in body. The most famous one is cell-cell recontingtion. Amino glycosides are also among important class of antibiotics containing few amino sugars in their structure. Amino sugar based antibiotics are mainly prescribed against aerobic and gram-negative bacteria. The  $\text{Cp}^*\text{Rh}$ -based IDA is a colorimetric method measured by a UV-Vis instrument. Figure 1.14 shows gallocyanine OD change upon addition of the host. OD change at 730 nm was suggested for tracing the complexation optically. The host-indicator binding strength is  $K = 2.3 \times 10^6 \text{ M}^{-1}$ , which is in the biological range, as explained above.

<sup>2</sup>  $\text{Cp}^*$  = pentamethylcyclopentadienyl



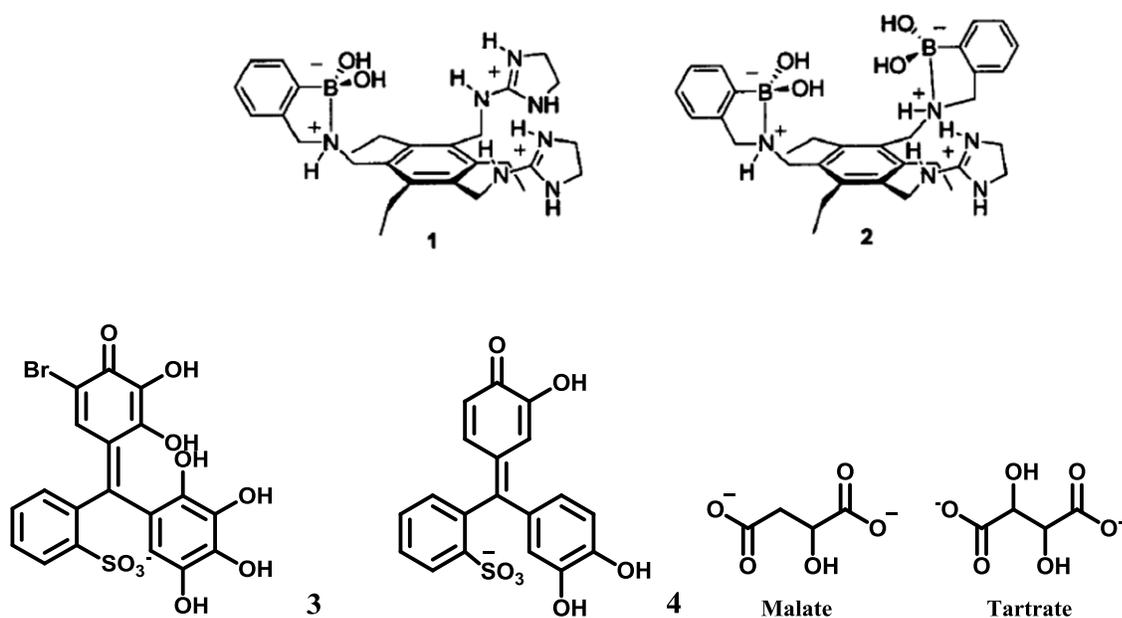
**Figure 1.14.** UV-Vis spectra of **2** titration with different concentration of **1** (0 – 50 mM). (reprinted by permission from [37])

To illustrate the sensitivity of the sensing (analyte and indicator competition), the binding strengths have to be compared. Glucosamine, galactosamine and mannosamine binding constants are  $5.6 (\pm 0.5) \times 10^4 \text{ M}^{-1}$ ,  $6.6 (\pm 0.5) \times 10^4 \text{ M}^{-1}$  and  $6.6 (\pm 0.5) \times 10^5 \text{ M}^{-1}$  respectively. However, the weaker binding constant of the analyte favors the competition for the host-indicator, the analyte by the concentration as low as  $20 \mu\text{M}$  were sensed. The concentration of aminoglycosides can be alternatively determined by a fluorescence probe against their amino group. [38] However fluorescence-based methods are generally more sensitive than absorbance methods but the probe suffers from non-specificity *i.e.* it can react with other primary amines in a real sample like food or blood.

#### 1.2.4. Multicomponent Indicator Displacement Assay (MIDA)

Synthesizing host and indicator guest is the bottle neck in IDA. That is why when a new host-dye is being invented lots of attention is attracted to know their properties, possible binding guests, IDAs and etc. To bypass host-synthesis one can use a host with different guest-dyes for sensing different analytes. This strategy is called multicomponent indicator displacement assay (MIDA). Based on this strategy, Severin and coworkers tried

to use their  $[(\text{Cp}^*\text{RhCl}_2)_2]$  host for sensing different analyte in a single cuvette.[39] Three different dyes (Mordant Yellow 10, Gallocyanine and Evans Blue) were selected for binding to the host. The complexes of the host and the three dyes were put in a cuvette and ADP, GTP and ATP as analytes were sequentially added. Mordant Yellow 10 bound with the lowest affinity to the host compared to others and was displaced by ADP, an analyte with the weakest binding strength. Gallocyanine, who bound stronger to the host was displaced by GTP, the analyte with a stronger binding constant. Finally, Evans Blue was displaced by ATP, the indicator and the analyte with the strongest binding strength. Ansyln and coworkers reported a MIDA in which a combination of two poor-specific receptors (receptors 1 and 2 Figure 1.15) makes a diagnostic tool for analyte sensing rather synthesizing specific receptors for each analyte. [40]



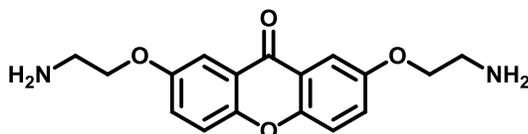
**Figure 1.15.** Structures of hosts (1 and 2), dyes (bromopyrogallol red 3, and pyrocatechol violet 4) and analytes (tartrate and malate) (Structures 1 and 2 were reprinted by permission from [41])

Receptor 1 has equal affinity for tartrate and malate and receptor 2 is more specific for tartrate over malate. Different combinations of 1,2,3,4, tartrate and malate simply change the UV/Vis spectra pattern. Analysis –by artificial neural network analysis (ANN)

method– of the spectral change makes the diagnostic tool for the analysis. The simplicity of the method and obviation for specific receptor synthesis is an advantage for the introduced MIDA. They have used their invented method (MIDA joint with ANN) for measuring Citrate and calcium in flavored vodkas. [41]

### 1.2.5. Indicator displacement assays in medical researches

Nakatani and coworkers have developed an IDA method for assaying some compounds that are capable of binding to RNA. [42] From medical point of view, RNA is the target of many drugs, fluorescence dyes and covalent modification for studying its binding affinity to proteins and other drugs. [43] But any chemical manipulation of RNA alters its functionality. Therefore they invented an IDA method for studying binding properties of a wild type RNA, without any need for a chemical modification. Their fluorescent indicator was 2,7-bis(2-aminoethoxy)-9H-xanthen-9-one derivative (X2S, Figure 1.16) [42] that is quenched when it is bound to the RNA. X2S was released from the RNA by human thrombin, viral protein Rev or neomycin antibiotic.



**Figure 1.16.** Structure of 2,7-bis(2-aminoethoxy)-9H-xanthen-9-one (X2S), a fluorescent indicator

X2S is quenched by double strand RNA (dsRNA) or with a RNA containing bulge. RNA worked as a host. X2S binds quite strongly to RNA with a binding constant of  $10^7 \text{ M}^{-1}$ . RNA-bulge quenches X2S more strongly than dsRNA. The quenching efficiency in a RNA-bulge also depends on the nucleotide composition. It was found that purine nucleotides quench RNA more efficiently than pyrimidines as follow:  $A > G > C > U$ . The assay therefore revealed some structural aspect of RNA.

For entering into a cell, RNA has to bind to a protein carrier. For example in HIV infection, mRNA of HIV-I has to bind to a viral protein Rev for transferring into cytoplasm.

[44] Compounds like neomycin (explained above) that compete with the Rev protein for binding to the RNA are among strong suppressor for HIV virus replication.

X2S was displaced from RNA-X2S complex with Rev and neomycin. The IDA revealed interesting unknown properties of the RNA binding to the Rev and neomycin. X2S release, resulted in a sigmoidal pattern fluorescence recovery. The sigmoidal pattern is a sign of a cooperative binding of Rev to the RNA, a reminiscent of cooperative binding of oxygen to hemoglobin. [1] Neomycin on the other hand releases the X2S by a slight slope linear pattern fluorescence enhancement, suggesting that neomycin not only binds to the RNA non cooperatively but also is not effectively binds to the RNA at low concentrations. In this study the IDA not only represented an assay method for neomycin and Rev but also revealed binding mechanism.

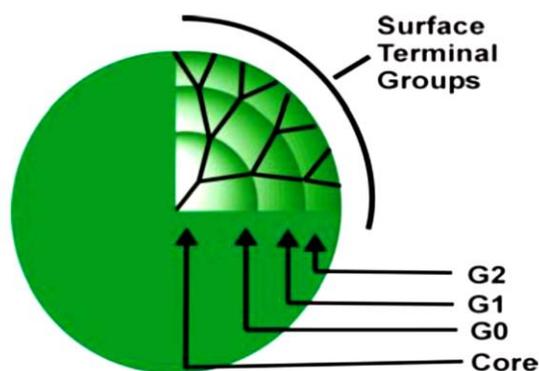
#### **1.2.6. Immobilized indicator displacement assay**

Immobilization is a method for making expensive materials reusable. Similar to enzyme immobilization, IDA immobilization is for making reusable hosts. Enzyme immobilization was a hot topic in 1970 – 1980 therefore it is well known over the past years. Naturally, immobilized IDA inherited the obtained experiences of enzyme immobilization. I would like to write a concise history on enzyme immobilization to understand immobilization capabilities and limitations.

Enzyme immobilization is a method for making reusable enzymes and biosensors. Enzyme is being immobilized on a solid surface (for example on a nanoparticle); a solution of a substrate is passed through the immobilized enzyme. The conversion of substrate is being assayed by a method specific to each enzyme. The immobilized enzyme is easily separated from the solution and is reused. In this way the expensive enzyme is saved for several assays.[23] The concentration of enzymatic product is directly proportional to the substrate concentration; therefore, the substrate sensing is detectable. The enzyme activity can be reduced by enzyme inhibitors. Immobilized acetylcholinesterase was used for sensing carbaryl insecticide. [23, 24] The developed method was further used for sensing dichlorvos, an organophosphorus insecticide. [45]

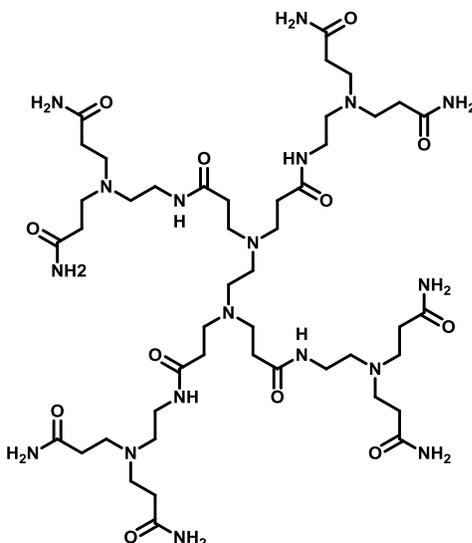
Most enzymes are very specific to their substrates; therefore, immobilized enzyme-based sensing is limited to the enzyme's substrate or their few inhibitors which is an advantage in one hand for specific sensing and is a disadvantage on the other hand because of limitation to few compounds. IDA as explained earlier is sensitive to broad number of analytes. While immobilized enzymes assays are mainly colorimetric methods the IDA is both colorimetric and fluorescence based. In all above examples of different aspects of IDA, the assay is based on soluble, single used hosts. It is an economic interest to design reusable hosts or reporter pair. Margerum and coworkers introduced IDA for sensing amino acids based on immobilized dendrimers.[46]

Dendrimers (Figure 1.17), are made of a core and branching molecules radially derived from the core. Generation (G) or size of dendrimers depends on the number of repeating groups branching from the core. Dendrimers are famous drug carriers.



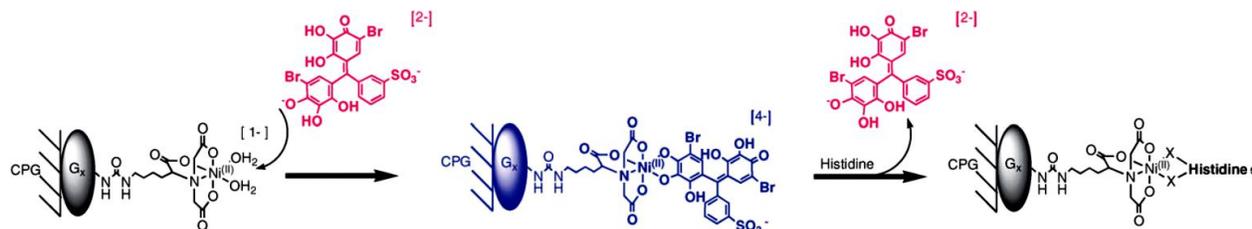
**Figure 1.17.** Schematic structure of dendrimers (awarded by Dr. Jyotirmayee Mohanty)

One of the most famous classes of dendrimers is polyamido amines (PAMAMs) dendrimers (Figure 1.18). PAMAMs provide a large 3D surface containing large number of amino groups prone for chemical modification. These properties make them interesting target for immobilizing for providing a monolayer with an enhanced surface for further chemical modifications.



**Figure 1.18.** Structure of polyamido amine (PAMAM)

Margerum and coworkers immobilized monolayer of PAMAMs G3, G4 and G5 on controlled pore glass (CPG). The immobilized PAMAMs were then modified with Ni-NTA made an immobilized host. The immobilized host then bound to a dye (bromopurogallol red). The dye were then displaced by histidine (Figure 1.19). Hosting capacity for the dye was increased by PAMAM size from G3 to G5. Similarly the threshold for histidine (His) detection for G4 and G5 is 5  $\mu\text{M}$  while it is 40  $\mu\text{M}$  for G3. Finally, due to the hosting capacity, the bigger dendrimers are capable of higher saturation. Immobilized G3, G4 and G5 could detect His up to around 750  $\mu\text{M}$ , 1500  $\mu\text{M}$  and 3500  $\mu\text{M}$  respectively. Among standard amino acids, the immobilized dendrimer are sensitive exclusively to His.



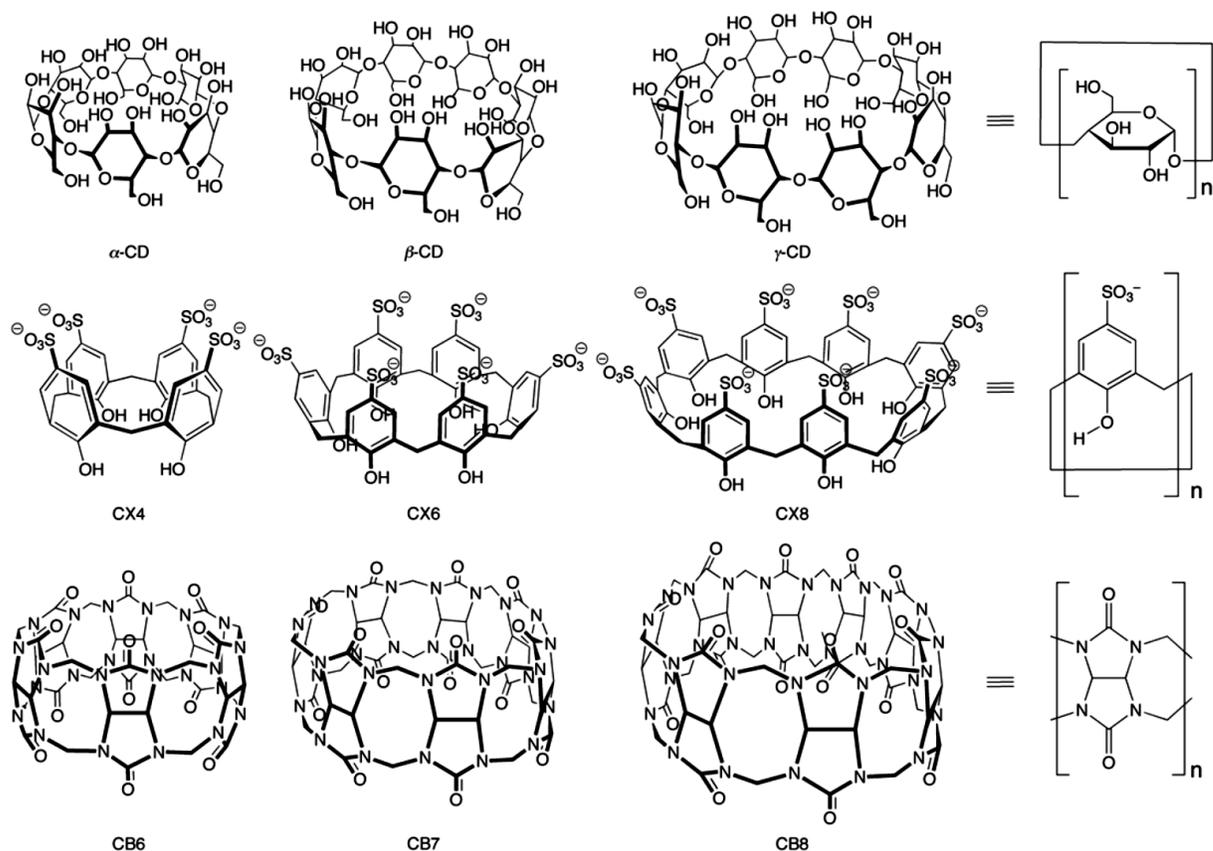
**Figure 1.19.** Immobilized indicator displacement assay for CPG-G<sub>x</sub>-NTA-Ni (reprinted by permission from [46])

### 1.3. Using macrocycles for drug discovery, delivery and release

Macrocylic hosts without any doubt are the most famous and highly-used hosts in IDAs. Herein, I review and consider the possibility of using macrocyclic molecules for hosting drugs and deliver them to cells and release them there by help of liberating competitors. From this point of view, the basic IDA's principles like hosting and competitive binding have applications beyond analyte sensing and can be used for drug delivery and release as well. This is an option when a drug has a poor solubility or low uptake efficiency by cells but upon binding to a macrocycle (macrocycle/drug) is more soluble or gains a better uptake property.

#### 1.3.1. Macrocyclic hosts and their hosting mechanisms

Figure 1.20 depicts three famous classes of macrocycles which are cucurbiturils (CBs), *p*-sulfonateocalixarenes (CXs) and cyclodextrines (CDs). All of these macrocyclic compounds have three properties in common: 1- They are highly water soluble 2- They made of hydrophilic groups on their top and bottom rims and a hydrophobic belt in the middle that makes a hydrophobic cavity 3- They are biocompatible.[47-49] These properties in general and the second one in particular, make them a suitable candidate for hosting hydrophobic compounds or drugs by bindings to the hydrophobic cavity to be shielded from water. Consequently, they can solubilize them. [50, 51]



**Figure 1.20.** Chemical structure of macrocyclic hosts: cyclodextrins, *p*-Sulfonatocalixarenes and cucurbiturils.

Biedermann and *et al* have recently suggested a mechanism explaining tendency of CBs for hosting molecules. [52, 53] Their results show that CBs containing high energy water molecules in their cavity that release upon a guest binding. Obviously releasing the high energy water molecules cannot be the only mechanism explaining the hosting because, this mechanism cannot explain the difference between binding strengths of different guest to the same host. Therefore, there must be complementary mechanisms. Nau and coworkers studies suggest that it is not only the hydrophobic cavity that interacts with guests but also the rims of CBs play a role in hosting by shifting  $pK_a$  of compounds[54, 55] and drugs. [56] Anionic rims of some macrocyclic compounds such as *p*-sulfonateo calix[4]arene (CX4), through complexion-derived  $pK_a$  shift can make cationic parts on the drugs that are neutral

in normal condition. The  $pK_a$  shift enhances their degree of ionization and bioavailability, stabilize the active form of the drugs and more importantly solubilize the drugs.[57]

### 1.3.2. Macrocyclic hosts as a drug carrier

Increasing attention is being paid recently to *in vivo* hosting using macrocycles due to their application in drug delivery, [58-60] increasing cell permeability to fluorescent dyes, [59] DNA transferring into cells, [61] cell immobilizing [62], and etc. Calixarenes and Cyclodextrins by having hydroxyl functional groups on their rims are capable of chemical modification and Nano particles fabrication. A supramolecular gel called CB8/MBCS was synthesized from a chitosan derivative, N-(4-dimethylaminobenzyl) chitosan (MBCS), and CB8. The gel is capable of carrying and releasing 5-Fluorouracil, an anti-cancer drug. [63] By modifying a cyclodextrin one can synthesize self-assembling molecules. Sun and *et al* have modified  $\beta$ -Cyclodextrin with an anthraquinone moiety that self-assembled into vesicles. The vesicles carried Paclitxel (an anti-cancer drug) and Ferrocene (an organometallic compound) into carcinoma cell lines, HepG2 and BEL-7402. Paclitxel/vesicle had more curative effect than Paclitxel itself. [64]

If a macrocycle/drug transfers by blood stream they will be taken up by different cells. Macrocycles cannot distinguish target cells from none target ones. On the other hand, monoclonal antibodies are very specifically recognizing their target cells but they lack hosting capability. Attaching macrocycles to an antibody can make a carrier which is capable of target cell recognizing, drug carrying and releasing. Antibodies were decorated with  $\beta$ -cyclodextrin were attached to biomarkers on different cells. The  $\beta$ -cyclodextrin then hosted adamantane-modified nanoparticles and used for cell imaging.[65] Such a strategy can be considered for directing drugs by hosting them inside macrocycles linked to a specific antibody.

Cucurbiturils however are more difficult to modify. But even none-modified CBs have fascinatingly improved drug properties. CB7 for example, drastically increases the solubility of some benzimidazole-based anthelmintic and fungicides drugs even up to 100 folds by shifting the drug's  $pK_a$ . [66] CB7 binds to the protonated forms of the drug more

strongly than to their neutral form therefore, CB7 derives the  $pK_a$  of the benzimidazole drugs to higher pH values. CB7 further improves the photostability of the drugs. CB7 also have therapeutically application. It stabilizes and activates proton-pump inhibitors lansoprazole and omeprazole (drugs for gastroesophageal reflux) in aqueous solution. It binds to the active form of the drugs and prolongs their degradation from one hour to three weeks. [56]

Macrocyclic hosts not only should be able to carry the drugs into the cells but they should also be able to release them there. The binding of a guest to a macrocycle is usually reversible with a binding strength in the range of  $10^3$  to  $10^6 \text{ M}^{-1}$  in water. [57] Ions compete with guests for binding to the macrocycles, [67] which weakens the binding constant of a guest in biological environment such as in blood or culturing media. There are many reports on fluorescence turn-on upon ions-binding to macrocycles, releasing the guest dye and its application for imaging and detecting ions in different cells. [33, 34]

The macrocycle/guest has a binding constant which is weakened in blood or media:



When the free guest is used by a cell, the macrocycle/guest dissociates in order to replenish the free guest and macrocycle concentration for keeping the binding constant unchanged according to Le Chatelier's principle. If such a hypothesis works, not only the drug release problem will be overcome but also macrocycles can act as a drug reservoir who delivers the drugs to the cells and releases them by rate of its consumption that can prolong the drug effect, as CB7 did for lansoprazole and omeprazole (explained above).

### 1.3.3. Calixarenes as a drug carrier

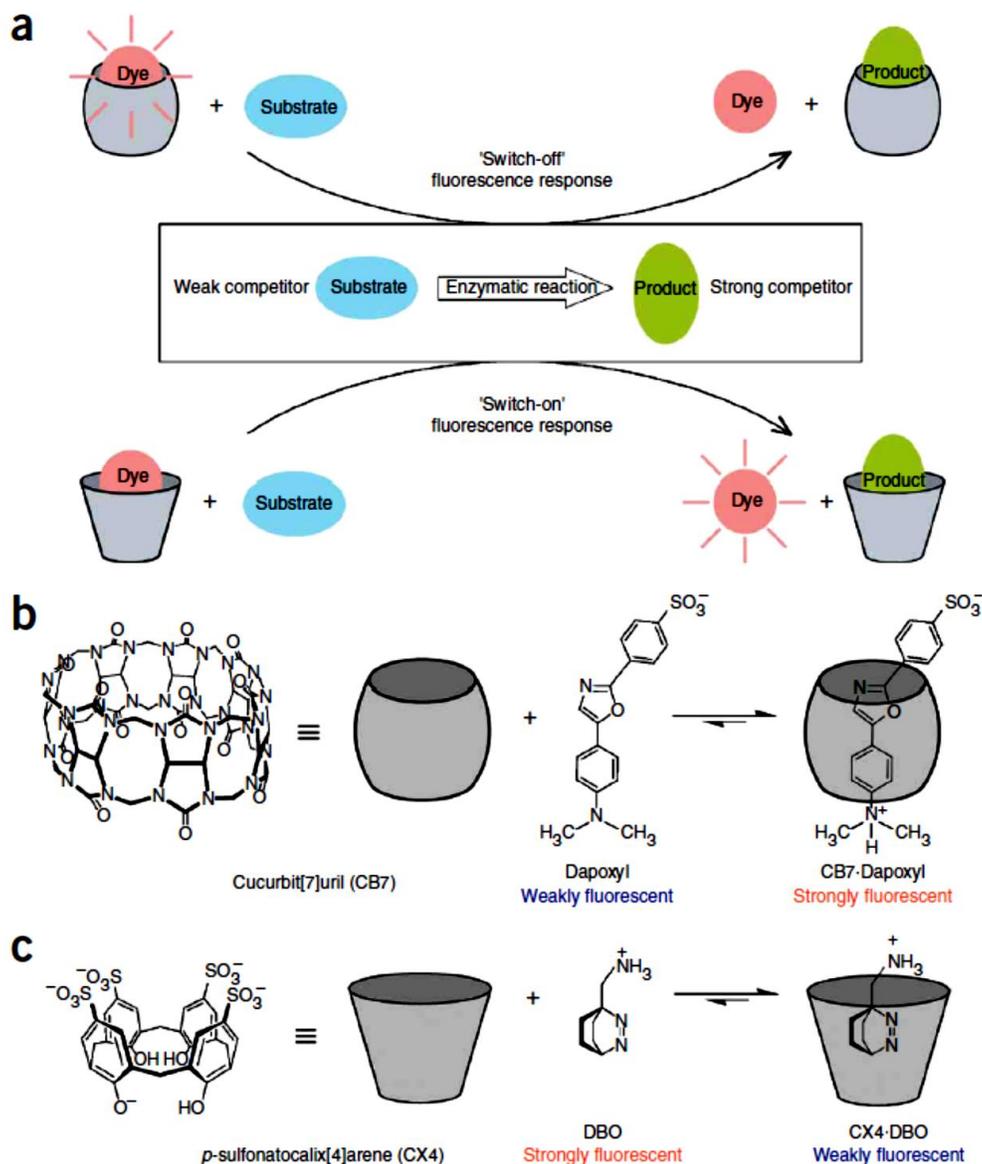
CXs are capable of carrying wider range of molecules in size compare to CBs and CDs. Because, CXs are flower shape in which their aromatic rings are connected just at the bottom (Figure 1. 20) and can have more variable conformations. Little conformation change, make a cavity shape flexible for obtaining an optimum interaction with its guest.

But CBs and CDs on the other hand have more rigid structures. Especially CBs that are barrel shaped in which the glycoluril monomers are connected on the top and bottom. As it was mentioned above functional groups on a calixarenes rims makes it prone for synthesize calixarene derivatives. Calixarenes and their derivatives have been used for carrying several compounds and fluorescent dyes into live cells for ion sensing. [68-70] Calixarene derivatives are also used for carrying biomolecules into live cells. Bagnacani and *et al* have recently suggested using CX4 decorated with arginine as a cargo carrier for DNA into different cells. [71] *para*-amino-calix[4]arene-based solid lipid nanoparticle could transfer DNA into cytoplasm of MDCK cells (Madin-Darby canine kidney cells). [72]

#### 1.3.4. Macrocyclic-based enzyme tandem assay for drug discovery

Many drugs are enzyme inhibitors. Potential drugs or enzyme inhibitors from a library of compounds have to be evaluated by enzyme assays. Conventional enzyme assays are usually colorimetric methods in which the product reacts rapidly with a chemical compound to produce a dye. Enzyme activity and kinetics is measured by recording the absorbance of the produced dye. For example the enzyme acetylcholinesterase (AChE) is being assayed by Elman's method in which the conversion of acetylthiocholine to thiocholine is detected by a chemical reaction between the product (thiocholine) and DTNB (5,5'-Dithio-bis(2-nitrobenzoic acid)). Thiocholine reacts with DTNB and produces a yellow product whose absorbance is recorded at 412 nm. [24]

Nau and coworkers have developed a real-time "tandem enzyme assay" based on IDAs (Figure 1. 21). [73] In the enzyme tandem assay, it is the substrate or the product that acts like an analyte and displaces a fluorescent indicator from its macrocyclic complex.



**Figure 1. 21.** The principle of supramolecular tandem enzyme assay and structure of reporter pairs. **a)** Supramolecular tandem enzyme assay is shown schematically. Conversion of substrate to product can make a fluorescence “switch-on” or “switch-off” using a proper macrocycle/dye as a reporter-pair. **b)** and **c)** shows the structures of macrocycles and dyes for switch-on and off process. Reprinted by permission from [74].

In the fluorescence switch-off IDA, the dye’s fluorescence is enhanced inside a macrocycle and upon displacing with an enzyme product its fluorescence intensity decreases (Figure 1.

21 a, on the top). While in the fluorescence switch-on the dye is quenched while it is inside the macrocycle and upon displacing with the product it fluoresce (Figure 1. 21a, at the bottom). The enzyme tandem assay has been used for assaying AChE which is a good example for comparing the conventional enzyme assay with the tandem enzyme assay. [75] In the conventional AChE assay (explained above) the artificial substrate acetylthiocholine produces an artificial product (thiocholine) while the tandem enzyme assay works with natural enzyme substrate and product. In the conventional method, concentration of the dye is measured with a UV-visible instrument while the tandem assay is measured with the fluorometer. Fluorescence spectroscopy is usually more sensitive than absorbance methods and with a fluorometer one can detect much lower concentration of a dye, consequently lower concentration of substrate and product is measurable. For example assaying decarboxylase with conventional methods, including colorimetric needs substrate concentration of 0.25-20 mM while tandem enzyme assay needs only 20–100  $\mu\text{M}$  of the substrate.[74] This is promising for detecting very low concentrations of enzyme because a low concentration of an enzyme produce little amount of product than may not be detectable by a UV-visible instrument. Secondly IDA using macrocyclic reporter pairs can be applied to a wide range of biological analytes that can compete with the guest in a same macrocycle. Therefore a particular reporter pair can be applied for variety of analytes while assaying using the conventional methods is designed only for the particular product. The final and the most important advantage is the possibility of using tandem assay inside a live cell for assaying range of analytes, which can be a substrate or a product of an enzyme. The macrocycles have been used in live cells, as explained above. In one study, *p*-Sulfonated calix[6]arene (CX6) were immobilized on Graphene oxide nanoparticle. The obtained nanoparticle quenched the fluorescence of *safranin T* in HepG2 cells. The fluorescence turn-on was observed by incubating the cells with L-carnitine who can displace dye form the immobilized CX6. [70] Ghale and *et al* have also recently reported a possibility of applying IDA in liposomes.[76] In this work, a fluorescent dye, lucigenin (LCG) binds to its host CX4 to form CX4/LCG. The CX4/LCG enters the liposomes through a protein channel (OmpF). LCG was then displaced by different analytes passing through the

protein channel into the liposome. The IDA was in real time. Membrane and cellular biocompatibility of macrocycles call for using them for *in vivo* enzyme tandem assays.

#### 1.4. References

- [1] Nelson DL, Cox MM, Lehninger AL. *Lehninger principles of biochemistry*. 5th ed. New York ; Basingstoke: *W.H. Freeman*; **2008**.
- [2] Berman HM. The past and future of structure databases. *Curr Opin Biotech*, **1999**;10:76-80.
- [3] Brenner SE, Chothia C, Hubbard TJP. Population statistics of protein structures: Lessons from structural classifications. *Curr Opin Struc Biol*, **1997**;7:369-376.
- [4] Fuchs E, Cleveland DW. A structural scaffolding of intermediate filaments in health and disease. *Science*, **1998**;279:514-519.
- [5] Hudgins RR, Huang F, Gramlich G, Nau WM. A fluorescence-based method for direct measurement of submicrosecond intramolecular contact formation in biopolymers: An exploratory study with polypeptides. *J Am Chem Soc*, **2002**;124:556-564.
- [6] Huang F, Nau WM. A conformational flexibility scale for amino acids in peptides. *Angew Chem Int Ed*, **2003**;42:2269-2272.
- [7] Jacob MH, Dsouza RN, Ghosh I, Norouzy A, Schwarzlose T, Nau WM. Diffusion-Enhanced Forster Resonance Energy Transfer and the Effects of External Quenchers and the Donor Quantum Yield. *J Phys Chem B*, **2013**;117:185-198.
- [8] Sahoo H, Roccatano D, Zacharias M, Nau WM. Distance distributions of short polypeptides recovered by fluorescence resonance energy transfer in the 10 angstrom domain. *J Am Chem Soc*, **2006**;128:8118-8119.

- [9] Cabrita LD, Dobson CM, Christodoulou J. Protein folding on the ribosome. *Curr Opin Struc Biol*, **2010**;20:33-45.
- [10] Voet D, Voet JG, Pratt CW. Fundamentals of biochemistry : life at the molecular level. 2nd ed. Hoboken, N.J.: *Wiley*; **2006**.
- [11] Sahoo H, Roccatano D, Hennig A, Nau WM. A 10-angstrom spectroscopic ruler applied to short polyprolines. *J Am Chem Soc*, **2007**;129:9762-9772.
- [12] Eaton WA, Munoz V, Hagen SJ, Jas GS, Lapidus LJ, Henry ER, et al. Fast kinetics and mechanisms in protein folding. *Annu Rev Bioph Biom*, **2000**;29:327-359.
- [13] Lakowicz JR. Principles of fluorescence spectroscopy. 3rd ed. New York, N.Y.: *Springer*; **2006**.
- [14] May V, Kühn O. Charge and energy transfer dynamics in molecular systems : a theoretical introduction. Berlin ;: *Wiley-VCH*; **2000**.
- [15] Meer BWvd, Coker G, Chen SYS. Resonance energy transfer : theory and data. New York ; Cambridge: *VCH*; **1994**.
- [16] Hardin C, Luthey-Schulten Z, Wolynes PG. Backbone dynamics, fast folding, and secondary structure formation in helical proteins and peptides. *Proteins*, **1999**;34:281-294.
- [17] Sinha KK, Udgaonkar JB. Dissecting the non-specific and specific components of the initial folding reaction of barstar by multi-site FRET measurements. *J Mol Biol*, **2007**;370:385-405.
- [18] Magg C, Kubelka J, Holtermann G, Haas E, Schmid FX. Specificity of the initial collapse in the folding of the cold shock protein. *J Mol Biol*, **2006**;360:1067-1080.

- [19] Amir D, Haas E. Estimation of Intramolecular Distance Distributions in Bovine Pancreatic Trypsin-Inhibitor by Site-Specific Labeling and Nonradiative Excitation Energy-Transfer Measurements. *Biochemistry-U.S.*, **1987**;26:2162-2175.
- [20] Nau WM, Greiner G, Rau H, Wall J, Olivucci M, Scaiano JC. Fluorescence of 2,3-diazabicyclo[2.2.2]oct-2-ene revisited: Solvent-induced quenching of the  $n,\pi^*$ -excited state by an aborted hydrogen atom transfer. *J Phys Chem A*, **1999**;103:1579-1584.
- [21] Tamaki T. The Photoassociation of 1-Ylanthracene and 2-Acetylanthracene with Methanol. *B Chem Soc Jpn*, **1982**;55:1761-1767.
- [22] Singh MK, Pal H, Bhasikuttan AC, Sapre AV. Dual solvatochromism of neutral red. *Photochem Photobiol*, **1998**;68:32-38.
- [23] Norouzy A, Habibi-Rezaei M, Qujeq D, Vatani M, Badiei A. Adsorptive Immobilization of Acetylcholine Esterase on Octadecyl Substituted Porous Silica: Optical Bio-analysis of Carbaryl. *B Korean Chem Soc*, **2010**;31:157-161.
- [24] Norouzy A, Qujeq D, Habibi-Rezaei M. The inhibitory effect of dissolved carbaryl in dioxane on physically adsorbed acetylcholinesterase. *React Kinet Catal L*, **2009**;98:391-401.
- [25] Nau WM, Wang XJ. Biomolecular and supramolecular kinetics in the submicrosecond time range: The fluorazophore approach. *Chemphyschem*, **2002**;3:393-398.
- [26] Huber R. Flexibility and Rigidity in Protein-Protein Protein-Pigment and Protein-Ligand Interaction. *Biological Chemistry Hoppe-Seyler* **1987**. p. 749.
- [27] Braddock DT, Louis JM, Baber JL, Levens D, Clore GM. Structure and dynamics of KH domains from FBP bound to single-stranded DNA. *Nature*, **2002**;415:1051-1056.

- [28] Taschner N, Muller SA, Alumella VR, Goldie KN, Drake AF, Aebi U, et al. Modulation of antigenicity related to changes in antibody flexibility upon lyophilization. *J Mol Biol*, **2001**;312:579-579.
- [29] Zavodszky P, Kardos J, Svingor A, Petsko GA. Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins. *P Natl Acad Sci USA*, **1998**;95:7406-7411.
- [30] Lapidus LJ, Steinbach PJ, Eaton WA, Szabo A, Hofrichter J. Effects of chain stiffness on the dynamics of loop formation in polypeptides. Appendix: Testing a 1-dimensional diffusion model for peptide dynamics. *J Phys Chem B*, **2002**;106:11628-11640.
- [31] Nguyen BT, Anslyn EV. Indicator-displacement assays. *Coordination Chemistry Reviews*, **2006**;250:3118-3127.
- [32] Feng L, Li H, Li X, Chen L, Shen Z, Guan Y. Colorimetric sensing of anions in water using ratiometric indicator-displacement assay. *Analytica Chimica Acta*, **2012**;743:1-8.
- [33] Que EL, Domaille DW, Chang CJ. Metals in neurobiology: Probing their chemistry and biology with molecular imaging (vol 108, pg 1517, 2008). *Chem Rev*, **2008**;108:4328-4328.
- [34] Domaille DW, Que EL, Chang CJ. Synthetic fluorescent sensors for studying the cell biology of metals. *Nat Chem Biol*, **2008**;4:168-175.
- [35] Wiskur SL, Ait-Haddou H, Lavigne JJ, Anslyn EV. Teaching old indicators new tricks. *Accounts Chem Res*, **2001**;34:963-972.
- [36] Niikura K, Bisson AP, Anslyn EV. Optical sensing of inorganic anions employing a synthetic receptor and ionic colorimetric dyes. *J Chem Soc Perk T 2*, **1999**:1111-1114.
- [37] Zaubitzer F, Buryak A, Severin K. Cp\*Rh-based indicator-displacement assays for the identification of amino sugars and aminoglycosides. *Chem-Eur J*, **2006**;12:3928-3934.

- [38] Kusnir J, Barna K. Fluorimetric determination of some basic antibiotics at very low concentrations. *Fresenius Zeitschrift Fur Analytische Chemie*, **1974**;271:288-288.
- [39] Buryak A, Pozdnoukhov A, Severin K. Pattern-based sensing of nucleotides in aqueous solution with a multicomponent indicator displacement assay. *Chem Commun*, **2007**:2366-2368.
- [40] Wiskur SL, Floriano PN, Anslyn EV, McDevitt JT. A multicomponent sensing ensemble in solution: Differentiation between structurally similar analytes. *Angew Chem Int Ed*, **2003**;42:2070-2072.
- [41] McCleskey SC, Floriano PN, Wiskur SL, Anslyn EV, McDevitt JT. Citrate and calcium determination in flavored vodkas using artificial neural networks. *Tetrahedron*, **2003**;59:10089-10092.
- [42] Zhang JH, Umemoto S, Nakatani K. Fluorescent Indicator Displacement Assay for Ligand-RNA Interactions. *J Am Chem Soc*, **2010**;132:3660-+.
- [43] Chow CS, Bogdan FM. A structural basis for RNA-ligand interactions. *Chem Rev*, **1997**;97:1489-1513.
- [44] Pollard VW, Malim MH. The HIV-1 Rev protein. *Annu Rev Microbiol*, **1998**;52:491-532.
- [45] Qujeq D, Roushan T, Norouzy A, Habibi-Rezaei M, Mehdinejad-Shani M. Effects of dichlorvos and carbaryl on the activity of free and immobilized acetylcholinesterase. *Toxicol Ind Health*, **2012**;28:291-295.
- [46] Greaney MJ, Nguyen MA, Chang CC, Good A, Margerum LD. Indicator displacement assays for amino acids using Ni-NTA tethered to PAMAM dendrimers on controlled pore glass. *Chem Commun*, **2010**;46:5337-5339.

- [47] Rowe RC, Sheskey PJ, Quinn ME. Handbook of Pharmaceutical Excipients. London . Chicago: *Pharmaceutical Press and American Pharmacists Association*; **2009**.
- [48] Coleman AW, Jebors S, Cecillon S, Perret P, Garin D, Marti-Battle D, et al. Toxicity and biodistribution of para-sulfonato-calix[4] arene in mice. *New J Chem*, **2008**;32:780-782.
- [49] Uzunova VD, Cullinane C, Brix K, Nau WM, Day AI. Toxicity of cucurbit 7 uril and cucurbit 8 uril: an exploratory in vitro and in vivo study. *Org Biomol Chem*, **2010**;8:2037-2042.
- [50] Brewster ME, Loftsson T. Cyclodextrins as pharmaceutical solubilizers. *Adv Drug Deliver Rev*, **2007**;59:645-666.
- [51] Dong N, Xue S-F, Zhu Q-J, Tao Z, Zhao Y, Yang L-X. Cucurbit n urils (n=7,8) binding of camptothecin and the effects on solubility and reactivity of the anticancer drug. *Supramolecular Chemistry*, **2008**;20:659-665.
- [52] Biedermann F, Uzunova VD, Scherman OA, Nau WM, De Simone A. Release of High-Energy Water as an Essential Driving Force for the High-Affinity Binding of Cucurbit n urils. *J Am Chem Soc*, **2012**;134:15318-15323.
- [53] Biedermann F, Vendruscolo M, Scherman OA, De Simone A, Nau WM. Cucurbit 8 uril and Blue-Box: High-Energy Water Release Overwhelms Electrostatic Interactions. *J Am Chem Soc*, **2013**;135:14879-14888.
- [54] Mohanty J, Bhasikuttan AC, Nau WM, Pal H. Host-guest complexation of neutral red with macrocyclic host molecules: Contrasting pK(a) shifts and binding affinities for cucurbit 7 uril and beta-cyclodextrin. *J Phys Chem B*, **2006**;110:5132-5138.

- [55] Shaikh M, Mohanty J, Singh PK, Nau WM, Pal H. Complexation of acridine orange by cucurbit 7 uril and beta-cyclodextrin: photophysical effects and pK(a) shifts. *Photochemical & Photobiological Sciences*, **2008**;7:408-414.
- [56] Saleh Ni, Koner AL, Nau WM. Activation and stabilization of drugs by supramolecular pK(a) shifts: Drug-delivery applications tailored for cucurbiturils. *Angew Chem Int Ed*, **2008**;47:5398-5401.
- [57] Ghosh I, Nau WM. The strategic use of supramolecular pK(a) shifts to enhance the bioavailability of drugs. *Adv Drug Deliver Rev*, **2012**;64:764-783.
- [58] Duan QP, Cao Y, Li Y, Hu XY, Xiao TX, Lin C, et al. pH-Responsive Supramolecular Vesicles Based on Water-Soluble Pillar[6]arene and Ferrocene Derivative for Drug Delivery. *J Am Chem Soc*, **2013**;135:10542-10549.
- [59] Li ZY, Sun SG, Yang ZG, Zhang S, Zhang H, Hu MM, et al. The use of a near-infrared RNA fluorescent probe with a large Stokes shift for imaging living cells assisted by the macrocyclic molecule CB7. *Biomaterials*, **2013**;34:6473-6481.
- [60] Yin JJ, Sharma S, Shumyak SP, Wang ZX, Zhou ZW, Zhang YD, et al. Synthesis and Biological Evaluation of Novel Folic Acid Receptor-Targeted, beta-Cyclodextrin-Based Drug Complexes for Cancer Treatment. *Plos One*, **2013**;8:0062289.
- [61] Hu Y, Chai MY, Yang WT, Xu FJ. Supramolecular Host-Guest Pseudocomb Conjugates Composed of Multiple Star Polycations Tied Tunably with a Linear Polycation Backbone for Gene Transfection. *Bioconjugate Chem*, **2013**;24:1049-1056.
- [62] Neiryneck P, Brinkmann J, An Q, van der Schaft DWJ, Milroy LG, Jonkheijm P, et al. Supramolecular control of cell adhesion via ferrocene-cucurbit[7]uril host-guest binding on gold surfaces. *Chem Commun*, **2013**;49:3679-3681.

- [63] Lin YW, Li LF, Li GW. Preparation and Properties of Temperature and pH Sensitive Chitosan Supramolecular Gel Based on Host-guest Interaction of Cucurbit 8 uril. *Acta Chim Sin*, **2012**;70:2246-2250.
- [64] Sun T, Guo Q, Zhang C, Hao JC, Xing PY, Su J, et al. Self-Assembled Vesicles Prepared from Amphiphilic Cyclodextrins as Drug Carriers. *Langmuir*, **2012**;28:8625-8636.
- [65] Agasti SS, Liong M, Tassa C, Chung HJ, Shaw SY, Lee H, et al. Supramolecular Host-Guest Interaction for Labeling and Detection of Cellular Biomarkers. *Angew Chem Int Ed*, **2012**;51:450-454.
- [66] Koner AL, Ghosh I, Saleh Ni, Nau WM. Supramolecular encapsulation of benzimidazole-derived drugs by cucurbit 7 uril. *Canadian Journal of Chemistry-Revue Canadienne De Chimie*, **2011**;89:139-147.
- [67] Bakirci H, Koner AL, Nau WM. Binding of inorganic cations by p-sulfonatocalix[4]arene monitored through competitive fluorophore displacement in aqueous solution. *Chem Commun*, **2005**:5411-5413.
- [68] Pathak RK, Hinge VK, Rai A, Panda D, Rao CP. Imino-Phenolic-Pyridyl Conjugates of Calix[4]arene (L-1 and L-2) as Primary Fluorescence Switch-on Sensors for Zn<sup>2+</sup> in Solution and in HeLa Cells and the Recognition of Pyrophosphate and ATP by [ZnL<sub>2</sub>]. *Inorg Chem*, **2012**;51:4994-5005.
- [69] Pandey S, Chawla HM. New dansylated calix(4)arene compounds useful as chemosensing agent for testing for the presence of mercury ion in sample e.g. blood, cells, tissue, saliva, sweat, drink, food and water. *Pandey S; Chawla H M; Indian Inst Technology Delhi*.

- [70] Mao XW, Tian DM, Li HB. p-Sulfonated calix[6]arene modified graphene as a 'turn on' fluorescent probe for L-carnitine in living cells. *Chem Commun*, **2012**;48:4851-4853.
- [71] Bagnacani V, Franceschi V, Bassi M, Lomazzi M, Donofrio G, Sansone F, et al. Arginine clustering on calix[4]arene macrocycles for improved cell penetration and DNA delivery. *Nat Commun*, **2013**;4.
- [72] Nault L, Cumbo A, Pretot RF, Sciotti MA, Shahgaldian P. Cell transfection using layer-by-layer (LbL) coated calixarene-based solid lipid nanoparticles (SLNs). *Chem Commun*, **2010**;46:5581-5583.
- [73] Dsouza RN, Hennig A, Nau WM. Supramolecular Tandem Enzyme Assays. *Chem-Eur J*, **2012**;18:3444-3459.
- [74] Hennig A, Bakirci H, Nau WM. Label-free continuous enzyme assays with macrocycle-fluorescent dye complexes. *Nat Methods*, **2007**;4:629-632.
- [75] Guo DS, Uzunova VD, Su X, Liu Y, Nau WM. Operational calixarene-based fluorescent sensing systems for choline and acetylcholine and their application to enzymatic reactions. *Chem Sci*, **2011**;2:1722-1734.
- [76] Ghale G, Lanctot AG, Kreissl HT, Jacob MH, Weingart H, Winterhalter M, et al. Chemosensing Ensembles for Monitoring Biomembrane Transport in Real Time. *Angew Chem Int Ed*, **2014**;53:2762-2765.

## Chapter 2

# The Effect of Electrostatic Charge Repulsion on Peptide Folding

Note: Some parts of this chapter are derived from contents of an article: "Coulomb Repulsion in Short Polypeptides" by Amir Norouzy, Khaleel I. Assaf, Shuai Zhang, Maik H. Jacob and Werner M. Nau accepted for publishing by *The Journal of Physical Chemistry B*.

### 2.1. Abstract

In this chapter the effect of Coulomb charge repulsion on the folding of 5-Fluorotryptophan-(X)<sub>6</sub>-Dbo peptides will be explained, where X is either an acidic (Glu,Asp) or a basic (Arg, His and Lys) amino acid, and 5-Fluorotryptophan and Dbo are fluorophores. End-to-end distances or structural aspect and end-to-end collision rates or flexibility aspect of peptides were measured at different pH values ranging from 1 to 12 by Förster resonance energy transfer (FRET) and collision-induced fluorescence quenching (CIFQ) respectively. By varying pH, charges appear on the peptides. The repulsion occurs only between the positively charged groups –N-terminus and the side chains in basic peptides– except for (His)<sub>6</sub> where there is also charge repulsion between side chains. Astonishingly, no effective charge repulsion among the negatively charged acidic side chains or the C-terminus with the side chains was observed. This can be explained by the difference in binding affinities of water towards carboxylates than the positively charged groups, which results in screening of the negative charges. The charge repulsion was recovered for the negatively charged peptides by changing the solvent from water to a 92.5% methanol/water. My results also show that the charge repulsion can affect peptides end-to-end collision rate more than their average end-to-end distance. The size of the charged groups on the side chains and their distances from the backbone can affect the flexibility of the peptides.

## 2.2. Investigation of oligopeptides by fluorescence-based methods

Although short peptides are believed to occur as random coils without a defined structure, but the folding of a protein already starts at early stage while it is being synthesized in a ribosome.[1] Therefore, peptide folding might mimic the early stages of protein folding. As it was explained in chapter 1, in spite of several theoretical suggestions,[2-4] The study of interactions involved in peptide folding is experimentally challenging by the well-established X-ray crystallography and 2D-NMR methods due to crystallization failure and flexibility of peptides. Even the fastest methods in stopped-flow analysis, by dead time of 50-200  $\mu s$  [5] is not fast enough for studying the dynamics of a peptide whose end-to-end collides in ns time range. But fluorescence-based methods can address these problems.[6-8]

In this project, we presented the structural and flexibility changes (dynamic) in ionic hexa-peptides caused by side chain–side chain and N-terminus–side chain charge repulsion using the fluorescence-based methods explained below. Charged amino acids, however, do not drive the protein folding, but they can solubilize [9] and structurally stabilize some proteins[10] as well as destabilize the structure of some others by electrostatic interactions in their denatured state ensembles (DSE).[11] In a DSE  $\leftrightarrow$  folded equilibrium, stabilizing the DSE favors unfolding. We measured the end-to-end distance of charged peptide by FRET. The flexibility of the peptides was measured by collision-induced fluorescence quenching (CIFQ). The flexibility of the peptides can coherently correlate to their end-to-end distances and also resembles the energy barrier between looped and straight (fully unfolded) peptide conformation.

End-to-end collision rate of peptides were measured by attaching a Dbo to their N-terminus and a 5-Fluorotryptophan to the C-terminus (Figure 2.2). DBO (Figure 2.1 a), is the fluorophore part of the Dbo molecule (Figure 2.1b) and is being quenched upon collision with a Trp (quencher) attached to their N-terminus of a peptide (Figure 2.2 b). End-to-end collision of hexa-peptides takes about 25 ns to more than 250 ns, depending on peptide rigidity.[7] Normally, fluorophores have lifetimes of about 1- 10 ns, [12] which is not an enough time for all collisions to occur. DBO by a long lifetime of about 300 ns provide enough time for measuring all end-to-end collisions in hexa-peptides to occur.[13]



**Figure 2.1.** Chemical structures of **a)** 2,3-diazabicyclo[2.2.2]oct-2-ene (DBO) and **b)** DBO-labeled asparagine (Dbo)

DBO not only can be used as a probe but can also be used as an acceptor in FRET. The resonance energy of an excited tryptophan can be absorbed by DBO (Figure 2.2. a).[8] Therefore, determination of transferred energy and consequently the end-to-end distance is possible by using the same double-labeled peptide.

All peptides were synthesized by Biosyntan company by purity more than 95% and amidated at C-terminus except for Ac-(Asp)<sub>6</sub>-OH, which had been acetylated at the N-terminus instead of C-terminus. Dbo was synthesized according to a method previously described. [13] All other chemicals were purchased from Sigma-Aldrich. The pH was adjusted by adding HCl and NaOH.

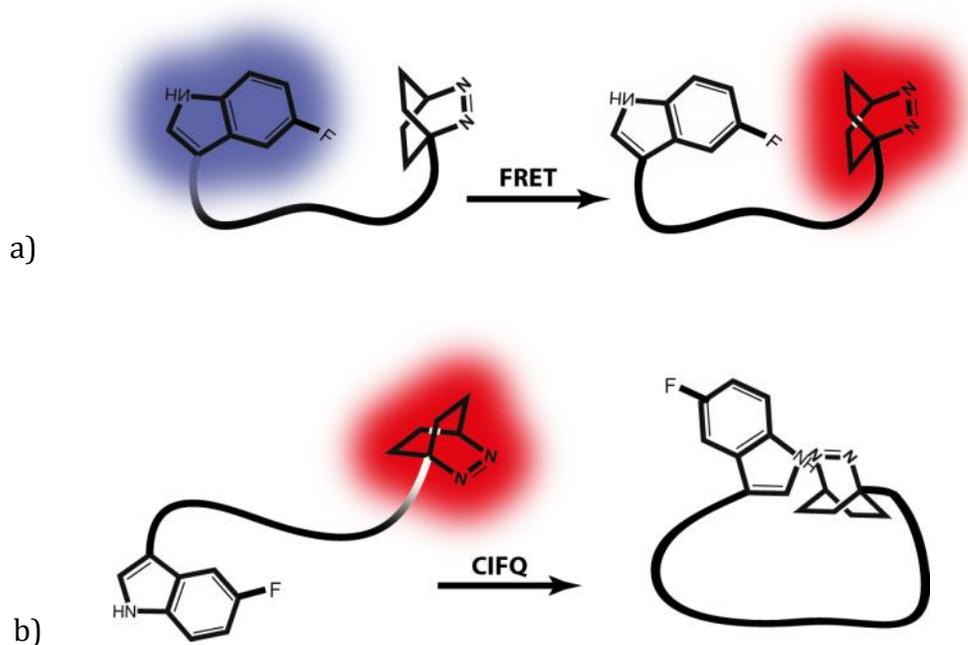
### 2.3. Experimental

The fluorescence lifetimes of all peptides (concentration < 100  $\mu$ M) were measured by the time-correlated single-photon counting instrument (FLS920, Edinburgh Instruments Ltd. Figure 1.6). The temperature was fixed to 25 °C by using a circulating water bath (Julabo F25/HD thermostat). Trp was excited by using a PicoQuant pulsed LED (PLS-280,  $\lambda_{exc}$  = 280 nm,  $\lambda_{obs}$  = 350 nm, fwhm ca. 450 ps). DBO was excited by using a PicoQuant diode laser LDH-P-C 375 ( $\lambda_{exc}$  = 373 nm,  $\lambda_{obs}$  = 450 nm, fwhm ca. 50 ps). Methanol makes the average fluorescence lifetime of tryptophan to decrease from 2.77 ns to 2.01 ns and of DBO from 305.25 to 21.48 ns. Nevertheless, DBO lifetime traces remains mono exponential in methanol (Figure 2.13 b). The CD spectra were recorded on a Jasco J-810 circular dichrograph (1nm resolution, 10 accumulations) in 0.1-cm cuvettes. Peptide concentrations for CD measurements were 200  $\mu$ M in D<sub>2</sub>O instead of water because of hydroxide absorbance in UV range.[14] CD spectra were measured at pD 2 and 12. The pH values were measured by a SenTix<sup>®</sup> Mic micro electrode pH meter.

## 2.4. Results

All peptides were labeled with 5-Fluorotryptophan at the N-terminus and Dbo at the C-terminus as mentioned above. We used double-labeled peptides: hexaaspartate (Asp)<sub>6</sub>, hexaglutamate (Glu)<sub>6</sub> and N-terminally acetylated (capped), C-terminally free hexaaspartate Ac-(Asp)<sub>6</sub>-OH as acidic peptides and hexaarginine (Arg)<sub>6</sub>, N-terminally capped hexaarginine Ac-(Arg)<sub>6</sub>, hexalysine (Lys)<sub>6</sub>, hexahistidine (His)<sub>6</sub> and N-terminally capped hexahistidine Ac-(His)<sub>6</sub> as basic peptides. Double-labeled hexaasparagine (Asn)<sub>6</sub> was used as a control peptide because it has no ionizable side chains.

As it was extensively explained in chapter 1, by FRET and CIFQ, the average end-to-end distance ( $R$ ) and the end-to-end collision rate constant ( $k_q$ ) of peptides were measured. Average end-to-end distance reveals structural properties of peptides[6] while  $k_q$  reveals flexibility information about peptide chain dynamics. [7] In CIFQ, excited Dbo is quenched upon collision with the Trp (Figure 2. 2. b) whereas in FRET, the energy of the excited Trp can be transferred to the Dbo (Figure 2. 2. a).



**Figure 2.2.** Schematic quenching of a fluorophore-probed hexapeptide by: a) FRET, where excited 5-Fluorotryptophan transfers its energy to Dbo and b) collision-induced fluorescence quenching (CIFQ), where excited Dbo is quenched upon collision with the 5-Fluorotryptophan.

CIFQ rate constants ( $k_q$ ) for each peptide were calculated by equation 2.1, where  $\tau$  is the average fluorescence lifetime of a double-labeled peptide (FTrpX<sub>6</sub>Dbo, X being any amino acid) and  $\tau_0$  is the average fluorescence lifetime of the corresponding single-labeled X<sub>6</sub>Dbo peptide (as a reference peptide for CIFQ measurement).

$$k_q = \frac{1}{\tau} - \frac{1}{\tau_0} \quad (2.1)$$

$\tau$  and  $\tau_0$  values were calculated from the lifetime traces (Figure 2.3. a) and were used for calculating  $k_q$  (Figure 2.3. b). The lifetime of DBO is not affected by pH in the range between 1 and 12.5 (single labeled peptides in Figure 2.3 a).

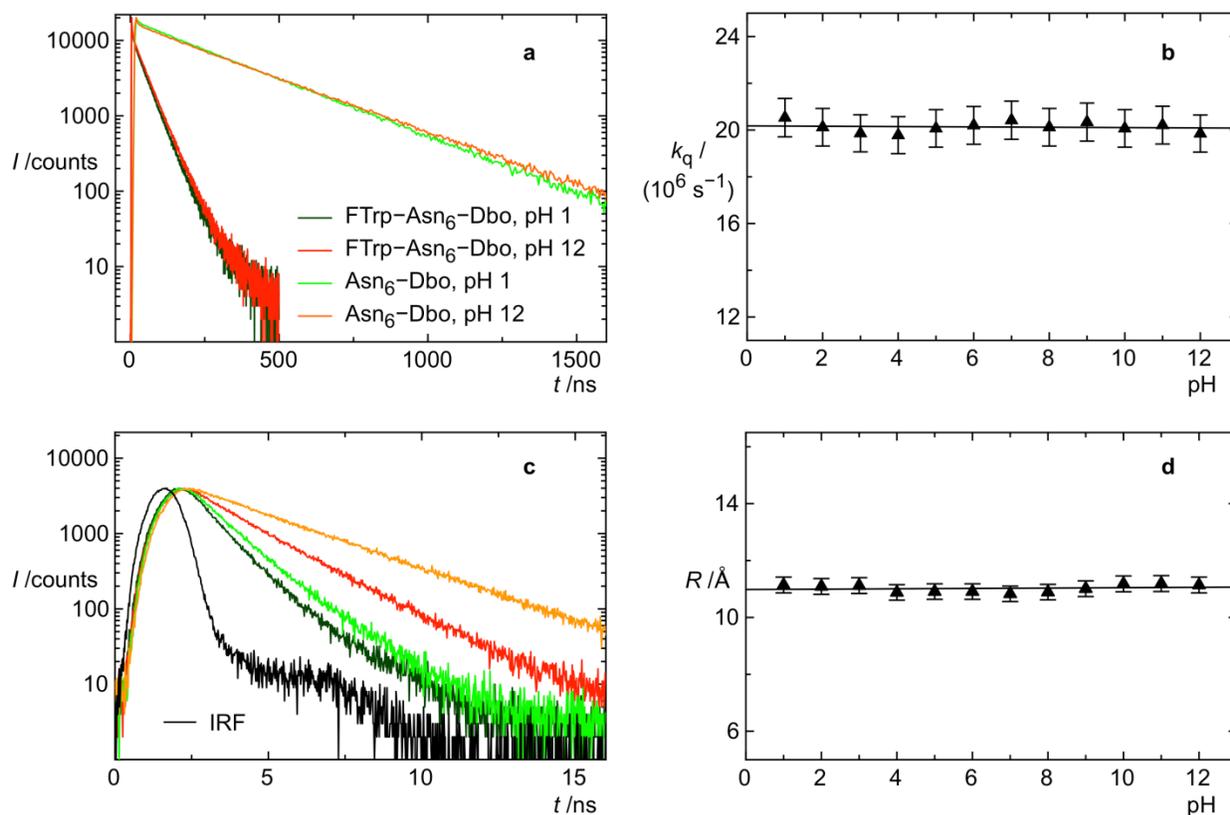
The average end-to-end distances of the peptides were calculated using FRET efficiencies ( $E$ ), which can be obtained from equation 2.2, where  $\tau_{DA}$  is the lifetime of a double-labeled peptide (FTrpX<sub>6</sub>Dbo), where FTrp is the energy donor and DBO is the energy acceptor and  $\tau_D$  is the lifetime of a single-labeled peptide with the donor only (FTrpX<sub>6</sub>).

$$E = 1 - \frac{\tau_{DA}}{\tau_D} \quad (2.2)$$

The average distance between a donor and an acceptor can be calculated by equation 2.3, where  $R_0$  is the Förster radius and  $E$  is the FRET efficiency from equation 2.2.  $R_0$  values were corrected according to the corresponding  $\tau_D$  values. The FRET analysis is explained in more detail in chapter 1.

$$R = R_0 \left( \frac{1}{E} - 1 \right)^{\frac{1}{6}} \quad (2.3)$$

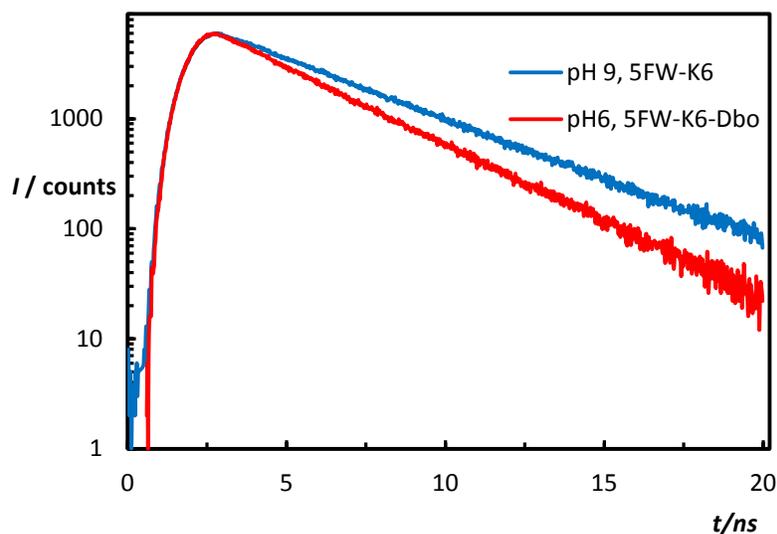
From lifetime decay traces (Figure 2.3.c) and equations 2.2 and 2.3, the end-to-end distances were calculated at different pH values (Figure 2.3.d).



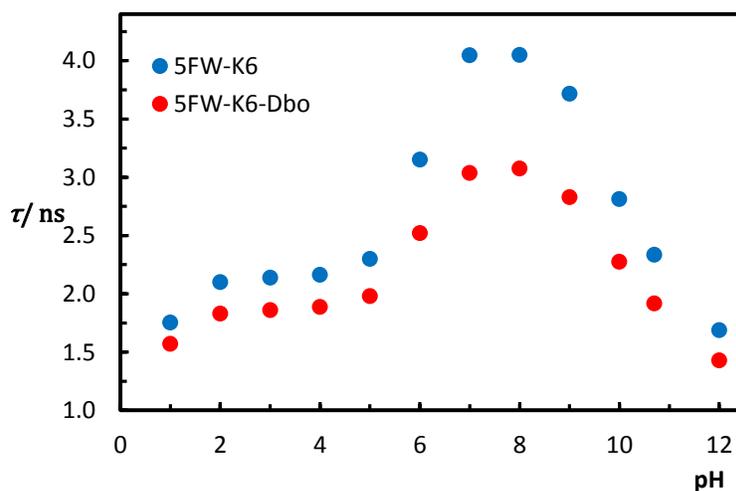
**Figure 2.3.** Fluorescence decays of  $(\text{Asn})_6\text{-Dbo}$  and  $\text{FTrp}-(\text{Asn})_6\text{-Dbo}$  (a) were analysed for plotting  $k_q$  values as a function of pH (b). Fluorescence decays of  $\text{FTrp}-(\text{Asn})_6$  and  $\text{FTrp}-(\text{Asn})_6\text{-Dbo}$  (c) were analysed by deconvolution fitting by using the instrument response function (IRF) for plotting end-to-end distances as a function of pH (d).

#### 2.4.1. Fluorescence lifetime of the FTrp as a probe

Unlike for the DBO, the lifetime of tryptophan varies by pH (Chapter 1 and [12]). Trp also shows a biexponential decay at pH 7 ( $\tau_1=3.1$  ns and  $\tau_2=0.53$  ns at 25 °C and pH=7.0), but we selected FTrp because –whether it is free or attached to the peptides– it shows a monoexponential decay [12] by the average lifetime of about 3 ns at pH 7 in single-labeled,  $\text{FTrp-X}_6$  peptides. Lifetime decays of FTrp at pH 9 in  $\text{FTrp}-(\text{Lys})_6$  and  $\text{FTrp}-(\text{Lys})_6\text{-Dbo}$  was represented in Figure 2.4, in which the lifetime of the single labeled peptide is longer than the double-labeled peptide. However, in some peptides like  $\text{H}_6$ , the side chains can quench the FTrp together with DBO. Consequently the FTrp shows a biexponential decay (Figure 2.6).



**Figure 2.4.** Fluorescence lifetime decays of 5-Fluorotryptophane in a single labeled (FTrpK<sub>6</sub>) and double-labeled (FTrpK<sub>6</sub>Dbo) peptide at pH 9.

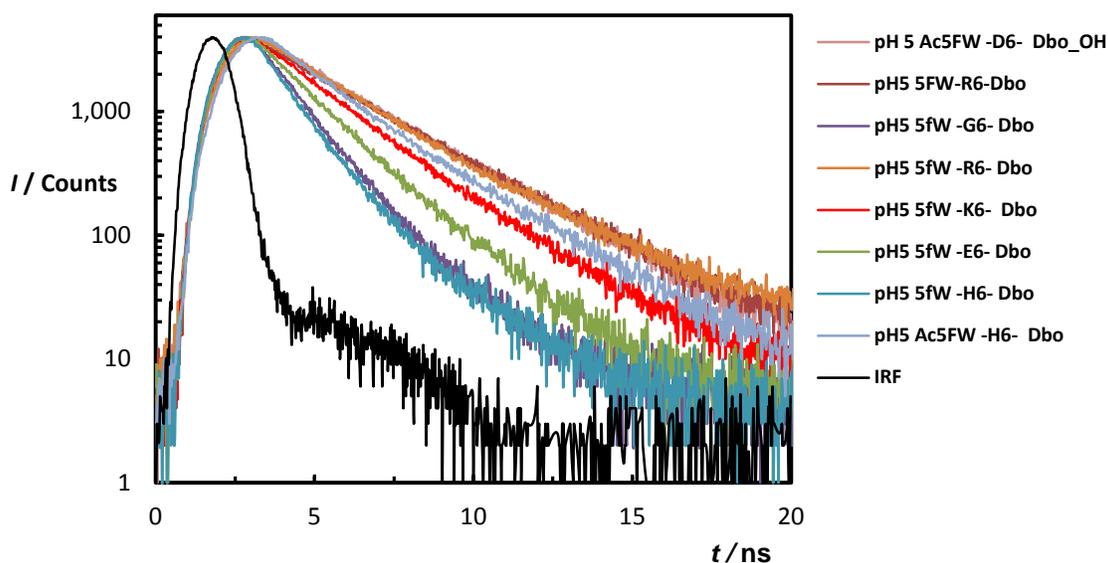


**Figure 2.5.** pH dependency of FTrp lifetime in both double-labeled (FTrpK<sub>6</sub>Dbo) and single labeled (FTrpK<sub>6</sub>) peptides.

Compare to Trp [12], FTrp attached to peptides N-terminus (FTrp-X<sub>6</sub>) shows shorter lifetime at almost any pH. For example compare the 5FWK6 life time in Figure 2.5 with Trp lifetimes in Figure 1.7. The shorter lifetime is due to FTrp quenching by the peptide backbone. [12]

Lifetime of FTrp like tryptophan is pH depended (Figure 2.5). The lifetime traces (raw data) of the double-labeled peptides were compared at a constant pH 5 of in Figure 2.6. Different peptides have different lifetimes depending on their end-to-end distance (FTrp quenching by DBO) or quenching of FTrp by basic amino acid side chains. Imidazoles of histidine sidechains quench the FTrp most efficiently.

As a consequence of lifetime change by pH, the Förster radius is also changing and has to be corrected for calculating FRET rate (Chapter 1).



**Figure 2.6.** 5-Fluorotryptophan lifetime decay traces in FTrp- $X_6$ -Dbo peptides at pH 5

$R_0$  is proportional to the sixth root of the lifetime of the reference peptides  $\tau_D$  (equation 2.4).

$$R_0 \propto \sqrt[6]{\tau_D} \quad (2.4)$$

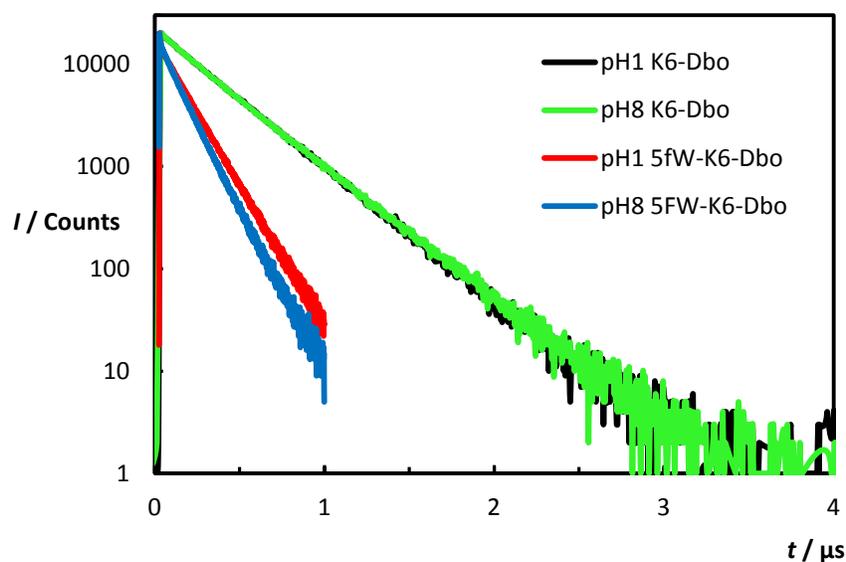
$R_0$  value for different  $W-(GS)_n$  peptides were reported [8].  $R_0$  and Trp lifetime for  $W-(GS)_4$  was assumed as a reference peptide ( $R_{0(R)}=9.2 \text{ \AA}$  and  $\tau_{D(R)}= 1.32 \text{ ns}$ ) and by calculating the measured lifetime of the single labeled peptide ( $\tau_{D(m)}$ ),  $R_0$  for each measurement ( $R_{0(m)}$ ) can be calculated by equation 2.5.

$$\frac{R_{0(m)}}{R_{0(R)}} = \sqrt[6]{\frac{\tau_{D(m)}}{\tau_{D(R)}}} \quad (2.5)$$

### 2.4.2. Flexibility change as a function of charge repulsion

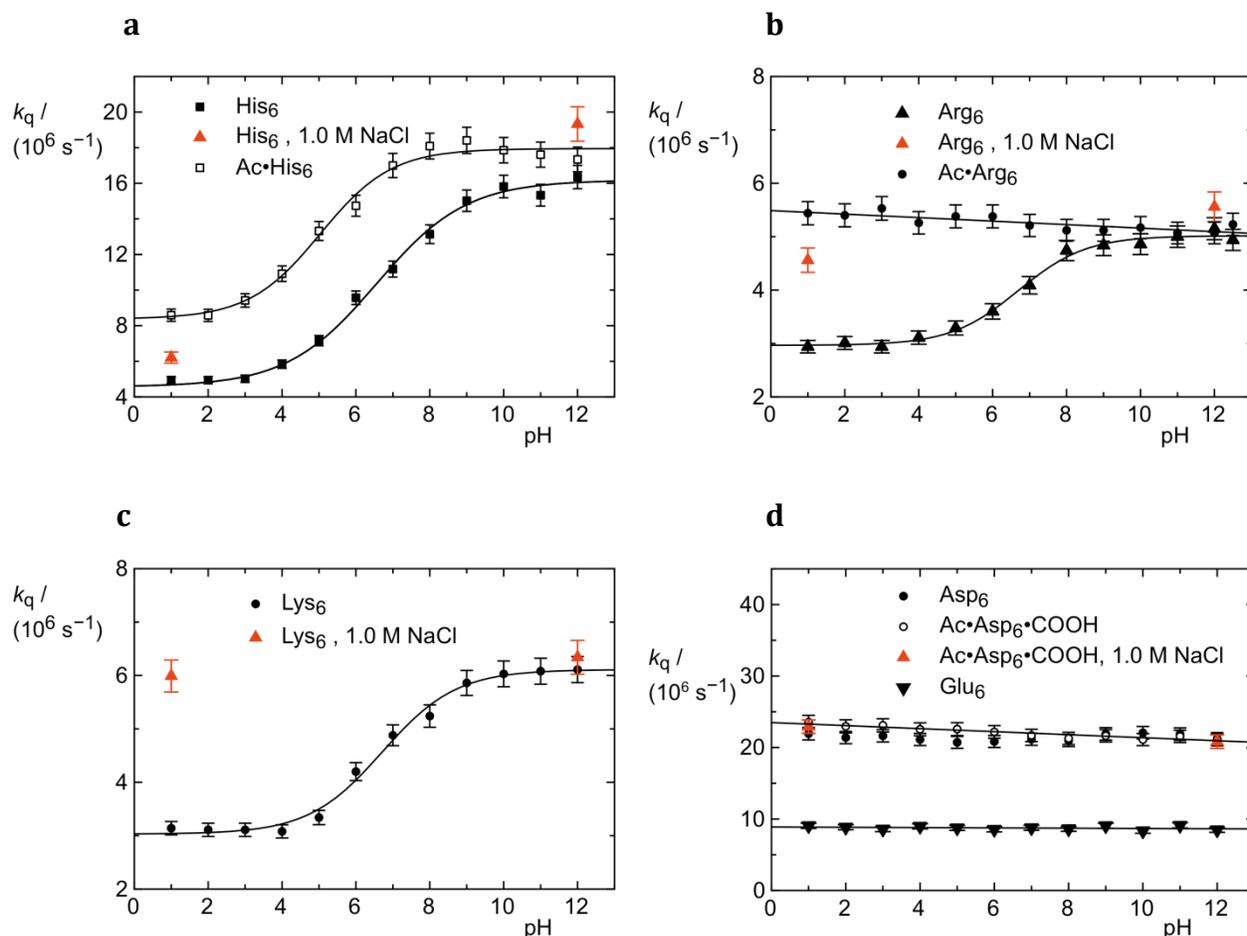
The same double-labeled peptides used in the FRET experiments could be also used together with a relevant reference peptide in CIFQ measurements. The reference peptide in CIFQ is a single Dbo labeled peptide. The peptide flexibility depends on the amino acid kind. The flexibility scale of 16 standard amino acids was described in detail by Nau and coworkers in 2003, by applying the CIFQ to short peptides [7].

End-to-end collision of a polypeptide chain, takes between ns to  $\mu\text{s}$  [13, 15] and DBO has the lifetime around 300 ns in aerated water, which is long enough for the collision. The collision rate was reported by collision rate constant ( $k_q$ ). The  $k_q$  (equation 2.1) can be calculated from the lifetimes of double-labeled peptide  $\tau$  and a reference peptide  $\tau_0$  (Figure 2.7)



**Figure 2.7.** Fluorescence Lifetime decays of DBO in double-labeled and single labeled  $K_6$  peptide at pH 1 and 8.

Collision quenching rate constants ( $k_q$ ) were measured for all peptides at different pH values ranging from 1 to 12 (Figure 2.8). The measurements are in good agreement with the previously reported  $k_q$  values. [7]



**Figure 2.8.** Flexibility of double-labeled, ionic peptide chains are represented by end-to-end collision rate constant ( $k_q$ ). Error bars show 8% error for measured  $k_q$  values. a)  $(\text{His})_6$  and N-terminally acetylated hexa-histidine  $\text{Ac}-(\text{His})_6$  b)  $(\text{Arg})_6$  and  $\text{Ac}-(\text{Arg})_6$ , c)  $(\text{Lys})_6$ , d) negatively charged peptides.  $(\text{Asp})_6$ ,  $(\text{Glu})_6$  and N-terminally acetylated C-terminally free hexa-aspartate  $\text{Ac}-(\text{Asp})_6\text{-OH}$

For simplicity in synthesizing, the C-terminus in all peptides except for  $\text{Ac}-(\text{Asp})_6\text{-OH}$  was amidated; therefore, the ionizing groups are the N-terminus and the side chains. Charge repulsion is expected between protonated ionizing groups in basic peptides (at pH values less than their  $\text{pK}_a$ ). By increasing the pH, there is a charge repulsion turn-off *i.e.* the basic side chains and the N-terminus become uncharged.

At pH values less than four, the basic (positively charged) peptides show minimum flexibility, as the charge repulsion among their ionizing groups disfavors looping (Figure

2.8 a,b and c). Upon deprotonation at higher pH values, the charge repulsion is neutralized and consequently, the flexibility increases. For relatively comparing the flexibility changes, the maximum  $k_q$  of a peptide in Figure 2.8 were normalized to 100% (Table 2.1).

The highest flexibility change is observed for (His)<sub>6</sub> by 69% (Table 2.1, Figure 2.8 a). Acetylation of the N-terminus in Ac-(His)<sub>6</sub> can cause approximately 18% decrease (from 69% in basic (His)<sub>6</sub> to 51% in Ac-(His)<sub>6</sub>). A highly flexibility increase of 41% is also observed for (Arg)<sub>6</sub> as a function of pH. But by acetylation of the N-terminus in Ac-(Arg)<sub>6</sub>, the peptide shows a constant and maximum flexibility at any pH value between 1 and 12 (Figure 2.8 b).

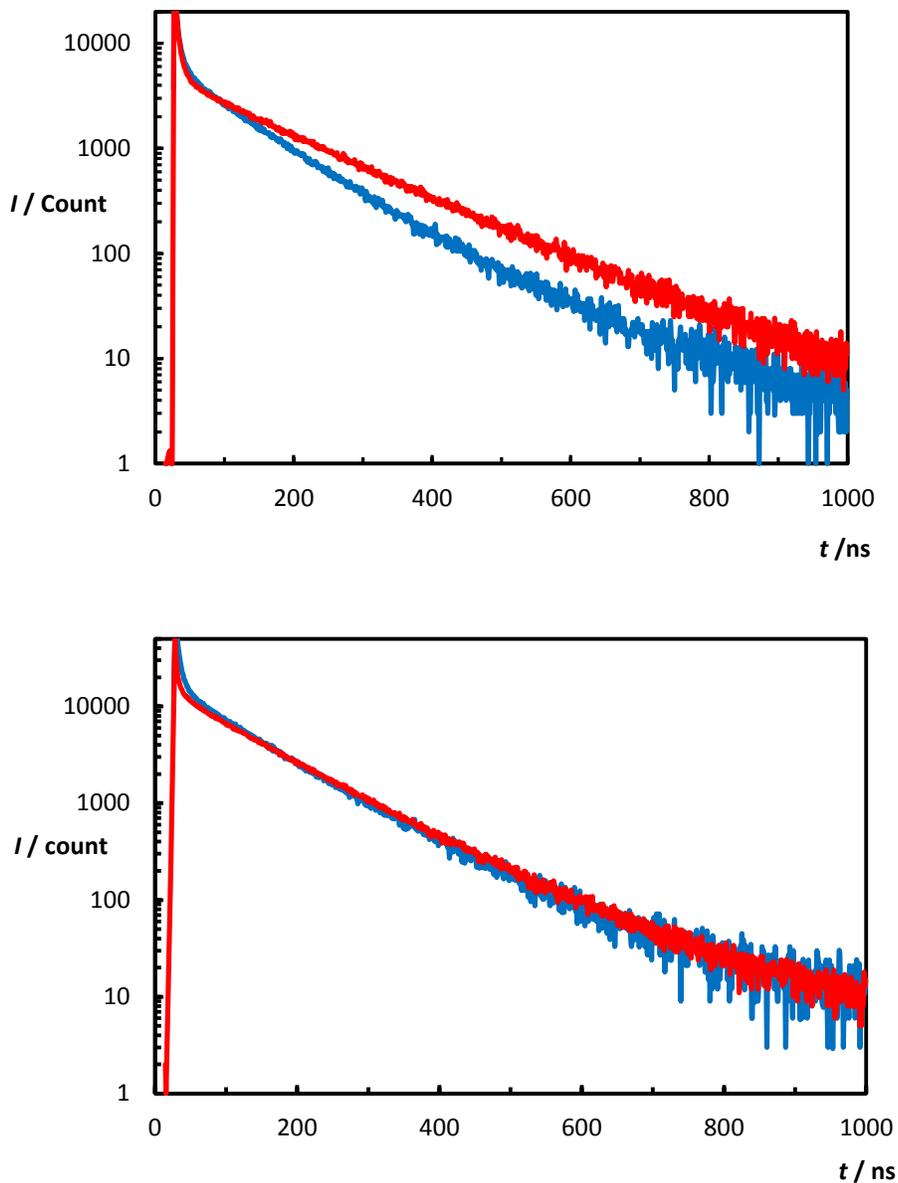
**Table 2.1.** Average flexibility ( $\bar{k}_q$ ) and average end-to-end distance ( $\bar{R}$ ) at pH values 1 and 2 for acidic and pH values 11 and 12 for basic solutions.  $|\Delta k_q|$  are  $|\Delta R|$  are the absolute value of relative flexibility and average end-to-end distance change respectively.

	$\bar{k}_q (10^6 s^{-1})$			$\bar{R} (\text{Å})$		
	pH<2 <sup>a</sup>	pH>11 <sup>b</sup>	$ \Delta k_q $ (%) <sup>c</sup>	pH<2 <sup>a</sup>	pH>11 <sup>b</sup>	$ \Delta R $ <sup>d</sup> (%)
(Arg) <sub>6</sub>	3.0	5.1	41	17.1	13.8	12
Ac-(Arg) <sub>6</sub>	5.4	5.1	7	12.9	13.7	3
(Lys) <sub>6</sub>	3.1	6.1	49	13.5	13.0	2
(His) <sub>6</sub>	4.9	16	69	13.5	11.8	6
Ac-(His) <sub>6</sub>	8.6	19	51	13.2	11.6	6
(Asp) <sub>6</sub>	22	22	3	10.9	11.8	3
(Glu) <sub>6</sub>	8.9	8.8	2	12.0	12.2	0
Ac-(Asp) <sub>6</sub> -OH	23	21	8	11.2	11.8	2
(Asn) <sub>6</sub>	20	20	2	11.1	11.2	0

<sup>a,b</sup> The mean values at pH 1 & 2 and 11 & 12. <sup>c</sup> The flexibility change in peptides as function of pH is represented in relative absolute value of  $k_q$  change ( $|\Delta k_q|$ ), by normalizing the maximum  $\bar{k}_q$  at pH>11 to 100% . <sup>d</sup> experimentally-obtained relative absolute values of  $R$

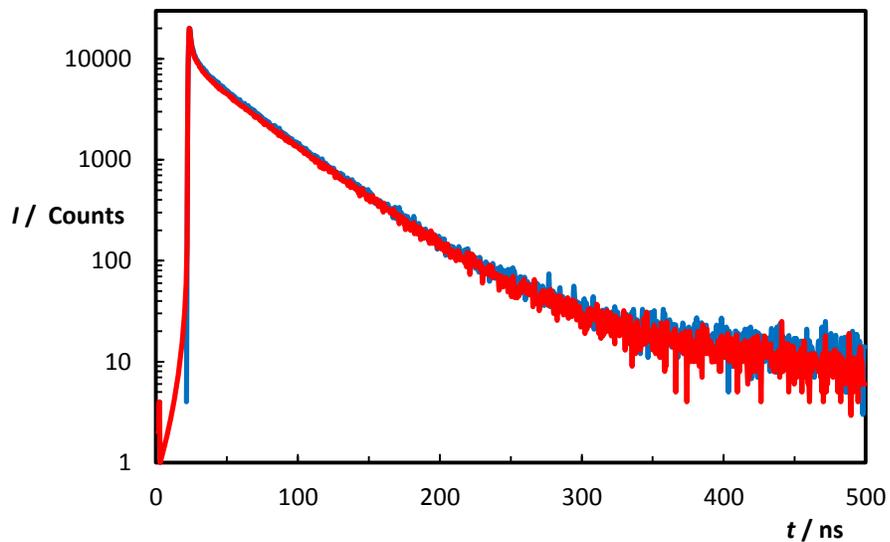
change,  $|\Delta R|$ , were calculated by normalizing 27.8 Å (end-to-end distance of a fully extended octamer peptide) to 100 %.

The lifetime traces of Arg<sub>6</sub> and Ac-Arg<sub>6</sub>'s peptide at extreme pHs (Figure 2.9) clearly endorse the observation of N-terminus acetylation effect represented in Table 2.1 and 28.b.



**Figure 2.9.** Lifetime traces for double-labeled Arg<sub>6</sub> peptide's (top) is pH dependent but Ac-Arg<sub>6</sub> 's lifetimes (bottom) is not. pH 1 and 12 are depicted in red and blue respectively.

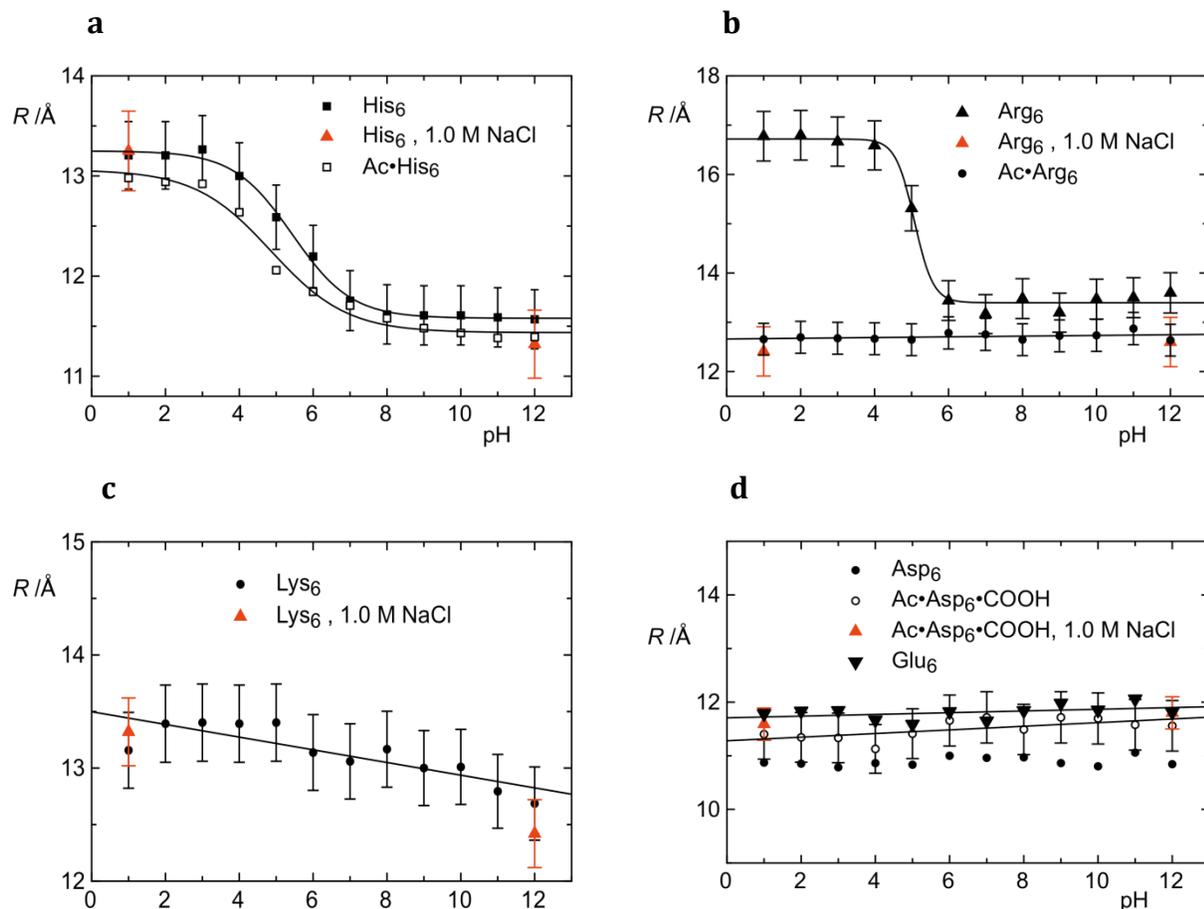
In (Lys)<sub>6</sub>, the deprotonated state indicates slightly higher flexibility than (Arg)<sub>6</sub> (Table 2.1). Unlike for the basic peptides, no charge effect was observed for acidic and (Asn)<sub>6</sub> peptide (Table 2.1 and Figures 2.3 and 2.8d, 2.10).



**Figure 2.10.** Lifetime traces for Ac-D<sub>6</sub>-COOH are identical at pH 1 (blue) and 8 (red).

### 2.4.3. End-to-end distance change as a function of charge repulsion

The average end-to-end distances for all peptides are represented in Figure 2.11.

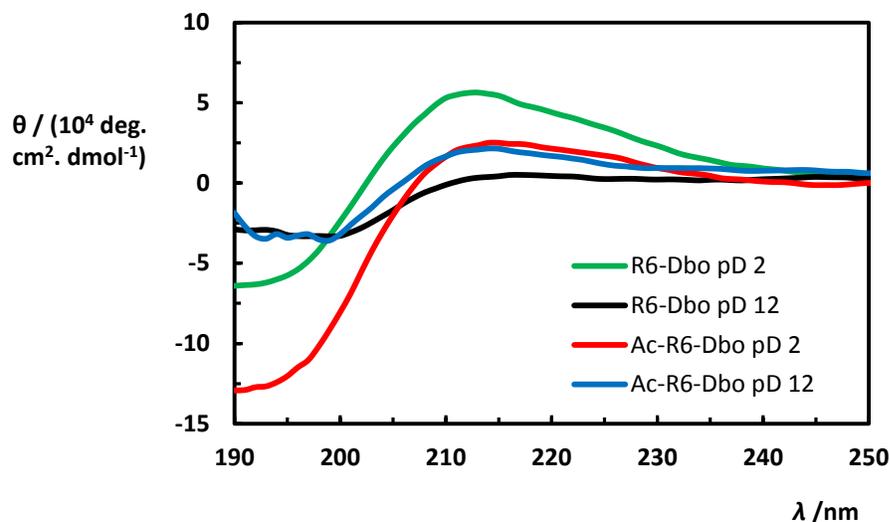


**Figure 2.11.** End-to-end distance ( $R$ ) of double-labeled ionic peptides as a function of pH. Error bars show 5% for the  $R$  values. a)  $(\text{His})_6$  and N-terminally acetylated hexa-histidine  $\text{Ac}\cdot(\text{His})_6$  b)  $(\text{Arg})_6$  and  $\text{Ac}\cdot(\text{Arg})_6$ , c)  $(\text{Lys})_6$ , d) The negatively charged peptides.  $(\text{Asp})_6$ ,  $(\text{Glu})_6$  and N-terminally acetylated / C-terminally free hexa-aspartate  $\text{Ac}\cdot(\text{Asp})_6\cdot\text{OH}$

The maximum end-to-end distances values for fully expanded double-labeled peptides (eight amino acids) is  $R = 27.8 \text{ \AA}$  calculated by Avogadro program with  $\vartheta = \psi = +180^\circ$  where side chains are furthest apart. The relative absolute value of end-to-end distance change ( $|\Delta R|$ ) where calculated by normalizing the maximum end-to-end distance value of an octamer peptide ( $R = 27.8 \text{ \AA}$ ) to 100%.

Carboxylate group in the side chains of glutamate and aspartate have a  $pK_a$  value of 4.25 and 3.65, respectively.[16] Therefore, charge repulsion among side chains is expected at pH values higher than their  $pK_a$  values. Nevertheless, the negatively charged peptides do not show end-to-end distance change as a function of pH (Figure 2.11 d).  $(Arg)_6$  shows the maximum end-to-end distance change as function of pH (Table 2.1). Capping the N-terminus causes end-to-end distance change to disappear (Figure 2.11 b). This observation was further confirmed with CD measurements:

CD was measured for singly labeled  $R_6$ -Dbo and Ac- $R_6$ -Dbo, because they do not contain FTrp which can make more complication in CD spectra. [17] The CD spectra were measured in  $D_2O$ . Water was replaced by  $D_2O$  because at high pH values water shows very high OD around 200 nm[14]. The CD spectra were measured at two extreme pD values (pD 2 and 12), where the ionizing groups are protonated and deprotonated, respectively (Figure 2.12). The CD spectra were analyzed with Spectra Manager software from the Spectropolarimeter (Table 2.2). Peptides do not have a defined structure; therefore,  $\beta$  strand or  $\alpha$  helix component is 0% at either pD values. However,  $R_6$  shows less randomness and more turned-structure in its protonated state (pD 2) than its deprotonated state (pD 12). Unlike for  $R_6$ , Ac- $R_6$  does not show structural change as a function of pD (i.e. for Ac- $R_6$  ratio of approximately 50% random/turn remains unchanged at two pDs). The CD measurements for all other peptides endorse the randomness of the peptide chains (results not shown).



**Figure 2.12.** The CD spectra of Ac-R<sub>6</sub>-Dbo and R<sub>6</sub>-Dbo in D<sub>2</sub>O and at pD 2 and 12

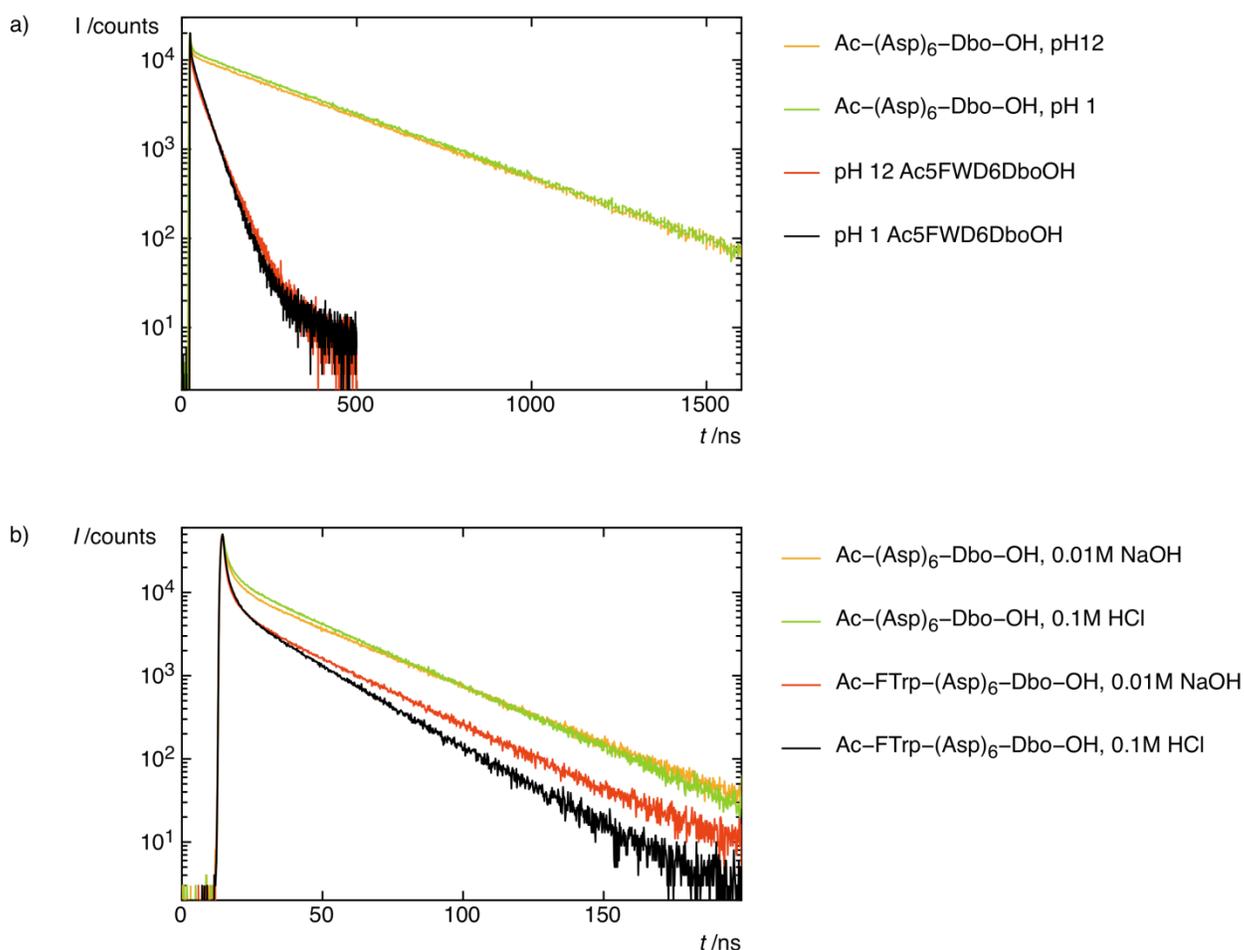
**Table 2.2.** Structural component of R<sub>6</sub> and Ac-R<sub>6</sub> at pD 2 and 12

Peptide, pD	Helix (%)	Beta (%)	Turn (%)	Random (%)	Sum
R6, pD 2	0.0	0.0	86.0	14.0	100.0
R6, pD 12	0.0	0.0	37.2	62.8	100.0
Ac-R6 pD 2	0.0	0.0	46.9	53.1	100.0
Ac-R6 pD 12	0.0	0.0	49.8	50.2	100.0

The end-to-end distance of (His)<sub>6</sub> and Ac-(His)<sub>6</sub> are both similarly decreasing from pH 3 to pH 7 (Figure 2.11 b). Capping the N-terminus in (His)<sub>6</sub> peptide has not a discernible effect on the end-to-end distance change. The (Asn)<sub>6</sub> and (Lys)<sub>6</sub> do not show any end-to-end distance change (Figures 2.3 d and 2.11 d ). The ionizable group in both peptides is an amino group.

#### 2.4.4. Charge repulsion recovery in an acidic peptide

The end-to-end distance and  $k_q$  of Ac-(Asp)<sub>6</sub>-OH were measured in acidic and basic 92.5% methanol/water. The mixture contained either 0.1M HCl for pH 1 or 0.01M NaOH for pH 12. Charge repulsion between carboxylate groups in basic solution causes less end-to-end collision rate as well as larger end-to-end distance (Table 2.3).



**Figure 2.13.** Fluorescence decays of Ac-FTrp-(Asp)<sub>6</sub>-Dbo-OH (red and black) and its single labeled peptide, Ac-(Asp)<sub>6</sub>-Dbo-OH (green and orange), in water at acidic and basic pHs (a) and in 92.5% methanol/water mixture in the presence of 0.1M HCl for pH 1 and 0.01M NaOH for pH 12 (b).

**Table 2.3.** The end-to-end collision rate constant values (obtained from Figure 2.13) and average end-to-end distance of Ac-(Asp)<sub>6</sub>-OH in acidic and basic 92.5% methanol/water solution mixture.

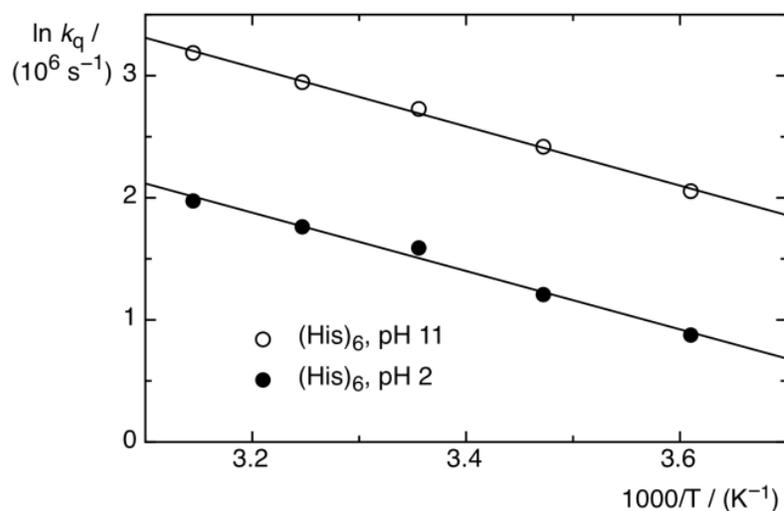
pH	$k_q (10^6 \text{ s}^{-1})^a$	$R (\text{\AA})^b$
Acidic	$10.0 \pm 0.4$	$11.2 \pm 0.3$
Basic	$3.4 \pm 0.2$	$9.3 \pm 0.2$

<sup>a</sup>  $k_q$  values were calculated with 8% errors. <sup>b</sup> end-to-end distance values were calculated with 5% error.

#### 2.4.5. The effect of charge repulsion on looping activation energy

The collision rate values of (His)<sub>6</sub> peptide at pH 2 and 11 were measured at 4, 15, 25, 35 and 45 °C and presented in Figure 2.14 with an Arrhenius plot. (His)<sub>6</sub> was selected because it has the largest change in  $k_q$  as a function of pH among all the studied peptides (Table 2.1). The activation energy and the pre-exponential factor were extrapolated from the Arrhenius plot and equation 2.4 (Table 2.4).

$$\ln k = \frac{E_a}{RT} + \ln A \quad (2.4)$$



**Figure 2.14.** Arrhenius plot of (His)<sub>6</sub> peptide at pH 2 and 11. Natural logarithm of  $k_q$  was plotted against  $1000/T$ . Where,  $T$  is the absolute temperature.

**Table 2.4.** The activation energy ( $E_a$ ) and the pre-exponential factor ( $A$ ) of (His)<sub>6</sub> looping at pH 2 and 11.

pH	$E_a$ (kJ/mol)	$A(10^3 \text{ s}^{-1})$
2	$19.9 \pm 0.1$	$13.8 \times 10^3$
11	$20.1 \pm 0.1$	$50.0 \times 10^3$

There is not a major difference between activation energies in protonated and deprotonated (His)<sub>6</sub>, which means there is not a strong intrachain hydrogen bond at two pH values. Therefore, the largest change in  $k_q$  is due to elimination of charge repulsion at high pH values, and not because of hydrogen bonds at higher pHs. The pre-exponential factor ( $A$ ) has significantly a higher value at pH 12. Higher pre-exponential factor (collision factor) means higher number of successful collisions,[18] which further supports the higher collision quenching rate at higher pH values.

## 2.5. Resolving peptide folding by FRET / CIFQ in tandem

All positively-charged peptides (basic peptides) except for Ac-(Arg)<sub>6</sub> show flexibility change as a function of pH (Figure 2.8,a,b and c). The more flexible the peptide is, the smaller the average end-to-end distance is expected. For example, (His)<sub>6</sub> at basic pH is more flexible than at acidic pH while its end-to-end distance is longer at acidic pH values than basic pH values (Figures 2.8 and 2.11).

Peptides can be ordered according to their end-to-end distances as follows.

R (Å): (Asp)<sub>6</sub>=Ac-(Asp)<sub>6</sub><(Glu)<sub>6</sub>=(His)<sub>6</sub>=Ac-(His)<sub>6</sub><Ac-(Arg)<sub>6</sub>= (His)<sub>6</sub><sup>+</sup>=Ac-(His)<sub>6</sub><sup>+</sup><(Lys)<sub>6</sub>=(Arg)<sub>6</sub>=(Arg)<sub>6</sub><sup>+</sup>

On the other hand, according to the flexibility results, the above order is reversed. Overall, the results indicate that the peptide flexibility is affected more than its end-to-end distance as a function of charge repulsion. For example, (His)<sub>6</sub> shows 69% flexibility change between its protonated and deprotonated state while it can show only 6% end-to-end distance change (Table 2.1).

\* The sign (+) means the peptide is in its protonated state.

### 2.5.1. The electrostatic charge repulsion disfavors the basic peptides looping.

The collision rate in (Arg)<sub>6</sub> gradually rises from pH 5 to 9 (Figure 2c). By the fact that the ionizing groups in (Arg)<sub>6</sub> are the N-terminus and the side chains, there are three possibilities for this observation: i) Charge repulsion between the protonated side chains (guanidino groups) makes the (Arg)<sub>6</sub> chain rigid at pH values less than 5 and by deprotonation (removal of the charges starting at pH ~5) the peptide chain become more flexible. This means that the  $pK_a$  of guanidine would shift from 12.48 [16] to approximately 7 (middle of 5-9). Although, a subtle electrostatic-induced  $pK_a$  shift was already reported in synthetic polypeptides, [19] a  $pK_a$  shift by more than 5 units seems to be unlikely. ii) The charge repulsion is between the N-terminus and the side chains. The N-terminus amine in arginine has a  $pK_a = 9.04$  and it would shift to 7, which is more believable. iii) The observed flexibility change is due to a combination of side chain-side chain (partial deprotonation of guanidino at pH ~9) and N-terminus-side chain charge repulsion. Capping the N-terminus makes the  $k_q$  to remain unchanged and jumps to the maximum at any pH value (Figure 2.8 b). This clearly proves that the charge repulsion is only between the N-terminus and the side chains. The average end-to-end distance of (Arg)<sub>6</sub> was also decreased by increasing the pH and again acetylation of the N-terminus caused the end-to-end distance to fall to minimum (Figure 2.11 b). Therefore, both FRET and CIFQ methods are mutually supportive and strongly indicate that the charge repulsion is between the N-terminus and the side chains in (Arg)<sub>6</sub>. Repulsion between the N-terminus and the side chains in (Arg)<sub>6</sub> was also greatly weakened by adding one molar of sodium chloride. The effect of ions at high concentration on protein refolding has already been reported.[20] The counter ions can screen the charges and reduce the charge repulsion. Although the repulsion between the positively charged N-terminus and the side chain can be generalized to other basic peptides as well, but in the case of (His)<sub>6</sub>, acetylation of the N-terminus Ac-(His)<sub>6</sub> can partially decrease the charge repulsion effect on the flexibility by merely 18% (from 69% for (His)<sub>6</sub> to 51% for Ac-(His)<sub>6</sub> in Table 2.1). (His)<sub>6</sub> shows the maximum flexibility change as function of pH among all peptides which is due to both the N-terminus-side chain and the side chain-side chain repulsion. The enormous change in (His)<sub>6</sub> flexibility as a function of pH (Table 2.1) can be explained by the bulkiness of the imidazole group and its proximity to the

peptide backbone. There is more evidence regarding the effect of side chain' size and proximity to the backbone on flexibility of peptides: i)  $G_6$  is the most flexible peptide because of its smallest side chain[7], ii) on the other hand, branched amino acids brings rigidity to a peptide. [7] iii) Glutamic acid's side chain is longer than Aspartic acid's by one  $-CH_2-$  which makes it twice as rigid (Figure 2.8 d). iv) The contour length of the side chain is another determining factor in the peptides flexibility. The basic peptides side chains lengths are as follow: Arg>Lys>His, the flexibility is reversed (deprotonated state in Figure 2.8) because the longer peptide side chain can make some steric hindrance during the loop formation. Adding NaCl by concentration of up to 1 M did not change the flexibility. A number of recent studies suggested that two histidines' imidazole rings can form a contact pair with one ring stacked upon the other. [21-23] In the absence of charge repulsion, the deprotonated imidazoles make a better ring compact. The average end-to-end distance is measured in a limited distance ( $0.5R_0 < R < 1.5R_0$ ) and limited time range (lifetime of FTrp in FTrp- $X_6$  peptides) in FRET. Similarity of end-to-end distance change in Ac-(His) $_6$  and (His) $_6$  (Figure 2.11 a) could be due to lack of efficient charge repulsion between the end terminus and the side change during the FRET distance-time limit.

(Lys) $_6$  shows a flexibility change as a function of pH (Figure 2.8 c) but unlike all other basic peptides no end-to-end distance change is observable for it (Figure 2.11 c). Among basic peptides ( $R_6$ ,  $H_6$  and  $K_6$ ),  $K_6$  has the smallest ionizing group in its side chain (an amino group). Ammonium ( $NH_4^+$ ) by inaccessible volume of  $24.5 \text{ \AA}^3$  (calculated by the HyperChem program) is 2.4 times smaller than guanidine ( $59 \text{ \AA}^3$ ) and 3 times smaller than imidazole group ( $74 \text{ \AA}^3$ ). The  $K_6$ 's amino group is separated by four  $-CH_2-$  from its backbone. The small size and the long distance from the backbone might explain the lack of measurable end-to-end distance change as a function of pH for  $K_6$  due to lack of side chain – side chain and the N-terminus – side chain interaction.

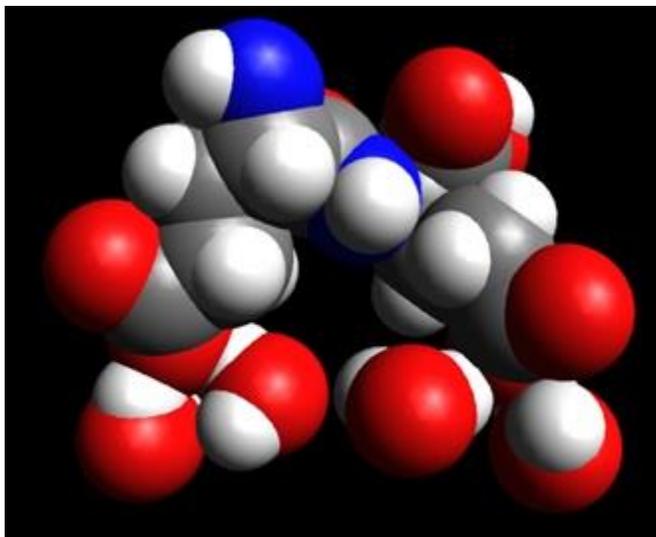
### 2.5.2. Absence of the charge repulsion effect on acidic peptides folding

The ionizing group in acidic peptides' side chain is a carboxylate with  $pK_a \sim 4$ . Neither  $k_q$  nor end-to-end distance change was observed for the acidic peptides at any pH value (Figures 2.8 d, and 2.11 d). However the shorter side chain in aspartic acid causes

(Asp)<sub>6</sub> and Ac-(Asp)<sub>6</sub>-OH to be two times more flexible than (Glu)<sub>6</sub> but the length of the side chain did not make any difference in average end-to-end distance between two peptides. Therefore, steric hindrance of the side chains can only affect the flexibility of the peptides not their average end-to-end distances. In another word, the length of a side chain has more flexibility effect than structural effect on peptides. For testing the possibility of C-terminus - side chain repulsion and eliminating the possible N-terminus-side chain attraction we used Ac-FTrp-(Asp)<sub>6</sub>-Dbo-OH (for simplicity; Ac-(Asp)<sub>6</sub>-OH) in which the C-terminus is in the form of -COOH not amidated and the N-terminus is capped. As a result, no charge effect was recovered (Figure 2.8 d, 2.10 and 2.11 d). Surprisingly, the negative repulsion is so subtle that neither it can make any change in the peptides' flexibility or in its end-to-end distance, while positive repulsion can do so. Therefore, we are challenging the conventional belief that negative charges on peptide chains repulse each other as strongly as the positive ones do. Another example for the lack of charge repulsion among negative charges was reported by Soranno and *et al* in 2009. [24] They have replaced three glutamine residues with three glutamates in C(AGQ)<sub>4</sub>W (*agq*) → C(AGE)<sub>3</sub>AGQW (*age*) to comparatively study the charge repulsion among carboxylates in *age*. "probability of end-to-end contact formation", "end-to-end root mean-square distance" and "intrachain diffusion coefficient" were measured by applying CIFQ (quenching of the triplet state of tryptophan by collision with a cysteine), FRET (using Trp as a donor and dinitrophenyl as an acceptor), and computational techniques respectively. A significant end-to-end distance and flexibility difference was not observed between *age* and *agq* from experimental results. But calculations show a two time increase in the intrachain dynamics for *age*. Their experimental results can be confirmed with our results. The lack of charge repulsion was explained by the separation of the glutamates by three peptide bonds. The distance between the carboxylates was assumed to be far more than Bjerrum length  $\sim 7 \text{ \AA}$  of two electrons in water.

We believe that water is responsible for lack of charge repulsion in the acidic peptides and for the *age* as well. Water has dielectric constant of 80 which is among the highest dielectric constants in solvents. Water molecules can bind to carboxylate groups more strongly than to the positive charges on positively charged amino acids, reviewed by

Collins and Chaplin. [25, 26] Moreover, the apparent dynamic hydration number (ADHN), or the number of water molecules tightly bound to an ion, for ammonium and carboxylate are 0 and 2 respectively, determined by Collins and Kiriukhin.[27] It can be hypothesized that the carboxylates in  $(\text{Glu})_6$ ,  $(\text{Asp})_6$ ,  $\text{Ac}-(\text{Asp})_6\text{-OH}$  and *age* have been screened by water molecules (Figure 2.15).



**Figure 2.15.** Carboxylate groups of diglutamate bound to two water molecules. The Water molecules screen the negative charge on the carboxylates.

Coulombic charge repulsion between carboxylates in benzene dicarboxylate in a solution containing methanol/water was reported by Wang and *et al.*[28] The 90% methanol/water mixture has dielectric constant of approximately 35 reported in the literature[29] which is two times less than the dielectric constant of water. By changing the solvent from water to an acidic and a basic 92.5% methanol/water mixture the charge repulsion was observed for  $\text{Ac}-(\text{Asp})_6\text{-OH}$  as a decrease in flexibility and a slight increase in its average end-to-end distance (Table 2.3 and Figure 2.13 b). Solvent replacement can change the folding of peptides and proteins, therefore; the flexibilities and end-to-end distances are not comparable in water and methanol/water. Nevertheless, the charge repulsion was recovered in the basic methanol/water mixture. As explained above, applying high concentration of salts can reduce the charge repulsion. By using 1 molar of

NaCl solution for screening the charges, neither the  $k_q$  nor the end-to-end distance changed (Figures 2.8 d and 2.11 d), which further proves the lack of efficient charge repulsion among carboxylates. To the best of our knowledge, except for the mentioned study[24], there is no report about the effect of charge repulsion on peptide folding and this effect was not specifically studied. Among different intrachain interactions in protein folding (hydrogen bonding, hydrophobic interaction, etc.), charge repulsion was studied robustly and for the first time in this project from a peptide level.

It should be noticed that lack of charge repulsion in short peptides may not be generalizable to proteins. Reported by Müller-Späth and *et al* in 2010[30], charge repulsion exists among negatively charged residues in unfolded state of N-terminal domain of HIV-1 integras (IN) and human prothymosin  $\alpha$  (ProT $\alpha$ ) proteins. But whether the carboxylate groups are exposed to be hydrated properly or buried in a hydrophobic pocket is not known. Most of the charged residues in ProT $\alpha$  [31] and IN are situated between hydrophobic amino acid residues in its sequence. These hydrophobic residues may have made a hydrophobic pocket which bans water molecules to hydrate carboxylate groups next to them. Therefore charges were not screened by water and showed repulsive effect.

Generally, protein cores are made of hydrophobic residues making a microenvironment with a low dielectric constant which can greatly enhance the charge interaction.

Any repulsion between charged side chains depends critically on the effective dielectric constant to be assigned to the microenvironment of the charge.[32] In a useful first approximation, one can evaluate the fractions of water molecules and protein atoms in the charge vicinity to arrive at a dielectric constant between 4 (protein interior) and 80 (water) as dependent on the fractions [33, 34]. In absence of asymmetric forces, a long random chain will always curve in such a way that it fits homogeneously into a spherical shape. Consequently, there will always be less water molecules and more chain atoms surrounding a charge than in the much shorter peptides studied here[35]; the dielectric constant will be smaller and charge-charge repulsion more effective.

However, Schweitzer Stenner and coworkers studied the tripeptide DDD and observed a pH-induced conformational response of the central aspartate side chain to pH

variation described by a changed set of preferred dihedral angles. [36] These measurements did not imply a changing average end-to-end distance.

## 2.6. Conclusion

We found out that the charge repulsion between the N-terminus and the side chains (cationic groups) can affect peptide structure (end-to-end distance) and dynamics but there is not meaningful charge repulsion among negatively charged groups. My explanation is: negative charges on peptides (carboxylates) are strongly hydrated, whereas the positive charge groups on peptides (N-terminus and side chains) are weakly hydrated. Therefore, negative charges are screened. I approved my hypothesis by changing the solvent from water to a 92.5% methanol/water mixture solution. Peptides seem to be more affected dynamically than structurally due to charge repulsion. Side chain size and approximation to the backbone can affect the peptide flexibility. To the best of my knowledge the above findings and given explanations especially lack of charge effect in negatively charged peptides is novel and counterintuitive. These results and the given interpretation above solely, could change people believes and way of thinking about peptide folding.

## 2.7. References

- [1] Cabrita LD, Dobson CM, Christodoulou J. Protein folding on the ribosome. *Curr Opin Struc Biol*, **2010**;20:33-45.
- [2] Daura X, Jaun B, Seebach D, van Gunsteren WF, Mark AE. Reversible peptide folding in solution by molecular dynamics simulation. *J Mol Biol*, **1998**;280:925-932.
- [3] Gnanakaran S, Nymeyer H, Portman J, Sanbonmatsu KY, Garcia AE. Peptide folding simulations. *Curr Opin Struc Biol*, **2003**;13:168-174.
- [4] Favrin G, Irbäck A, Wallin S. Folding of a small helical protein using hydrogen bonds and hydrophobicity forces. *Proteins: Structure, Function and Genetics*, **2002**;47:99-105.
- [5] Eaton WA, Munoz V, Hagen SJ, Jas GS, Lapidus LJ, Henry ER, et al. Fast kinetics and mechanisms in protein folding. *Annu Rev Bioph Biom*, **2000**;29:327-359.
- [6] Sahoo H, Roccatano D, Hennig A, Nau WM. A 10-angstrom spectroscopic ruler applied to short polyprolines. *J Am Chem Soc*, **2007**;129:9762-9772.
- [7] Huang F, Nau WM. A conformational flexibility scale for amino acids in peptides. *Angew Chem Int Ed*, **2003**;42:2269-2272.
- [8] Sahoo H, Roccatano D, Zacharias M, Nau WM. Distance distributions of short polypeptides recovered by fluorescence resonance energy transfer in the 10 angstrom domain. *J Am Chem Soc*, **2006**;128:8118-8119.
- [9] Voet D, Voet JG, Pratt CW. Fundamentals of biochemistry : life at the molecular level. 2nd ed. Hoboken, N.J.: Wiley; **2006**.
- [10] Xiao L, Honig B. Electrostatic contributions to the stability of hyperthermophilic proteins. *J Mol Biol*, **1999**;289:1435-1444.

- [11] Cho JH, Sato S, Horng JC, Anil B, Raleigh DP. Electrostatic interactions in the denatured state ensemble: Their effect upon protein folding and protein stability. *Arch Biochem Biophys*, **2008**;469:20-28.
- [12] Lakowicz JR. Principles of fluorescence spectroscopy. 3rd ed. New York, N.Y.: Springer; **2006**.
- [13] Hudgins RR, Huang F, Gramlich G, Nau WM. A fluorescence-based method for direct measurement of submicrosecond intramolecular contact formation in biopolymers: An exploratory study with polypeptides. *J Am Chem Soc*, **2002**;124:556-564.
- [14] Zuman P, Szafranski W. Ultraviolet-Spectra of Hydroxide, Alkoxide, and Hydrogen-Sulfide Anions. *Anal Chem*, **1976**;48:2162-2163.
- [15] Lapidus LJ, Steinbach PJ, Eaton WA, Szabo A, Hofrichter J. Effects of chain stiffness on the dynamics of loop formation in polypeptides. Appendix: Testing a 1-dimensional diffusion model for peptide dynamics. *J Phys Chem B*, **2002**;106:11628-11640.
- [16] Nelson DL, Cox MM, Lehninger AL. Lehninger principles of biochemistry. 4th ed. New York: Freeman; **2005**.
- [17] Kelly SM, Jess TJ, Price NC. How to study proteins by circular dichroism. *Bba-Proteins Proteom*, **2005**;1751:119-139.
- [18] Lister T, Renshaw J. New understanding chemistry for advanced level. 4th [i.e. 3rd] ed. Cheltenham: *Thornes*; **2000**.
- [19] Urry DW, Peng SQ, Parker TM, Gowda DC, Harris RD. Relative Significance of Electrostatic-Induced and Hydrophobic-Induced Pk(a) Shifts in a Model Protein - the Aspartic-Acid Residue. *Angewandte Chemie-International Edition in English*, **1993**;32:1440-1442.

- [20] Goto Y, Calciano LJ, Fink AL. Acid-Induced Folding of Proteins. *P Natl Acad Sci USA*, **1990**;87:573-577.
- [21] Heyda J, Mason PE, Jungwirth P. Attractive Interactions between Side Chains of Histidine-Histidine and Histidine-Arginine-Based Cationic Dipeptides in Water. *J Phys Chem B*, **2010**;114:8744-8749.
- [22] Marsili S, Chelli R, Schettino V, Procacci P. Thermodynamics of stacking interactions in proteins. *Phys Chem Chem Phys*, **2008**;10:2673-2685.
- [23] Haghani K, Khajeh K, Naderi-Manesh H, Ranjbar B. Evidence regarding the hypothesis that the histidine-histidine contact pairs may affect protein stability. *Int J Biol Macromol*, **2012**;50:1040-1047.
- [24] Soranno A, Longhi R, Bellini T, Buscaglia M. Kinetics of Contact Formation and End-to-End Distance Distributions of Swollen Disordered Peptides. *Biophys J*, **2009**;96:1515-1528.
- [25] Collins KD. Ions from the Hofmeister series and osmolytes: effects on proteins in solution and in the crystallization process. *Methods*, **2004**;34:300-311.
- [26] Chaplin MF. The Importance of Cell Water. *Science in Society*, **2004**:42-45.
- [27] Kiriukhin MY, Collins KD. Dynamic hydration numbers for biologically important ions. *Biophys Chem*, **2002**;99:155-168.
- [28] Wang XB, Nicholas JB, Wang LS. Intramolecular Coulomb repulsion and anisotropies of the repulsive Coulomb barrier in multiply charged anions. *J Chem Phys*, **2000**;113:653-661.
- [29] Kvamme B. Thermodynamic properties and dielectric constants in water/methanol mixtures by integral equation theory and molecular dynamics simulations. *Phys Chem Chem Phys*, **2002**;4:942-948.

- [30] Müller-Späth S, Soranno A, Hirschfeld V, Hofmann H, Ruegger S, Reymond L, et al. Charge interactions can dominate the dimensions of intrinsically disordered proteins. *P Natl Acad Sci USA*, **2010**;107:14609-14614.
- [31] Pan LX, Haritos AA, Wideman J, Komiyama T, Chang M, Stein S, et al. Human Prothymosin-Alpha - Amino-Acid-Sequence and Immunological Properties. *Arch Biochem Biophys*, **1986**;250:197-201.
- [32] Warshel A, Dryga A. Simulating electrostatic energies in proteins: perspectives and some recent studies of pKas, redox, and other crucial functional properties. *Proteins*, **2011**;79:3469-3484.
- [33] Warshel A, Sharma PK, Kato M, Parson WW. Modeling electrostatic effects in proteins. *Biochimica et biophysica acta*, **2006**;1764:1647-1676.
- [34] Jacob MH, Amir D, Ratner V, Gussakowsky E, Haas E. Predicting reactivities of protein surface cysteines as part of a strategy for selective multiple labeling. *Biochemistry-U S*, **2005**;44:13664-13672.
- [35] Zhou HX. Direct test of the Gaussian-chain model for treating residual charge-charge interactions in the unfolded state of proteins. *J Am Chem Soc*, **2003**;125:2060-2061.
- [36] Duitch L, Toal S, Measey TJ, Schweitzer-Stenner R. Triaspartate: a model system for conformationally flexible DDD motifs in proteins. *J Phys Chem B*, **2012**;116:5160-5171.

## Chapter 3

# Sensing Analytes *in vivo* Using Macrocyclic-Based Reporter Pairs

Note: Some parts of this chapter are derived from contents of a published paper: "Indicator Displacement Assays inside Live Cells" by Amir Norouzy, Zahra Azizi, and Werner M. Nau, *Angewandte Chemie International Edition*.

### 3.1. Abstract

In this chapter I am explaining a project that I independently designed and in which an indicator displacement assay was applied inside live cells. The p-sulfonatocalix[4]arene (CX4) host quenches the fluorescence of the cationic fluorescent guest lucigenin dye (LCG) upon forming the CX4/LCG complex. In this work, V79 and CHO cells were first incubated with the CX4/LCG complex. The CX4/LCG complexes passed through the cell membrane and existed in the cytoplasm of the cells despite the presence of many biomolecules of various size and charge. The loaded cells were then incubated with cationic analytes acetylcholine, choline and protamine. The cationic analytes passed through the cell membrane with different mechanism, bound to CX4 and displaced LCG, whose fluorescence recovery was followed by fluorescence microscopy. Additionally, to measure the LCG uptake, the cells were also incubated with CX4/LCG or LCG separately and lysed. The cells incubated with CX4/LCG were found to have a higher dye uptake. Therefore, CX4 can facilitate the cellular uptake of LCG and potentially be used as a carrier. This is a proof-of-concept for using CX4 as a carrier for those (bio)molecules who binds to the CX4 but having a poor cellular uptake.

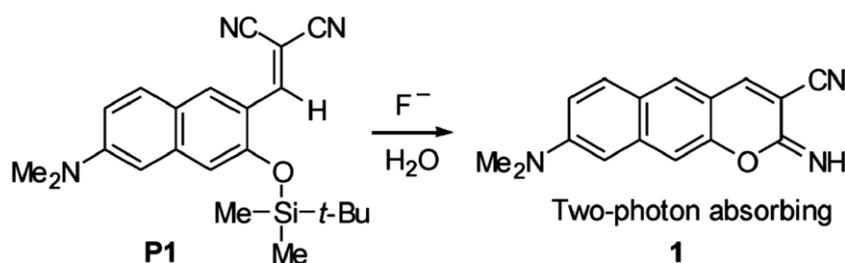
### 3.2. Antibody-free analyte sensing

Using macrocyclic hosts *in vivo* is a hot topic in science due to their various application in drug delivery, [1-3] fluorescent dyes encapsulation for uptake, [4] genes delivery, [5] and etc. See Chapter 1 for more information on different macrocyclic compounds and their miscellaneous applications.

Sensing bio analytes using indicator displacement assay (IDA) has wooed increasing attentions for replacing expensive antibody sensing methods like ELIZA. Jeon and *et al* have very recently introduced modified graphene sensor for antibody-free sensing of bio analytes: glutamic acid, ATP, ADP, dopamine and norepinephrine[6]. In IDA-based methods, low molecular weight analytes can be sensed that their detection with antibody-based methods is impossible.

#### 3.2.1. Fluorescent turn-on as a mean to follow *in vivo* sensing

The fluorescence “turn-on” technique has several applications in biological sciences. Fluorescence turn-on might happen by a chemical reaction inside a cell. For example, fluoride triggered a ring closure of a compound (*tert*-butyldimethylsilyl ether Figure 3.1) and turned it into a fluorescent dye from a none-fluorescence component in B16F10 cells (mouse skin cancer cell line, Figure 3.2). [7]



**Figure 3.1.** Ring closure reaction of *tert*-butyldimethylsilyl ether(P1) with  $\text{F}^-$  results in **1** which is a fluorescent. Reprinted by permission from [7].

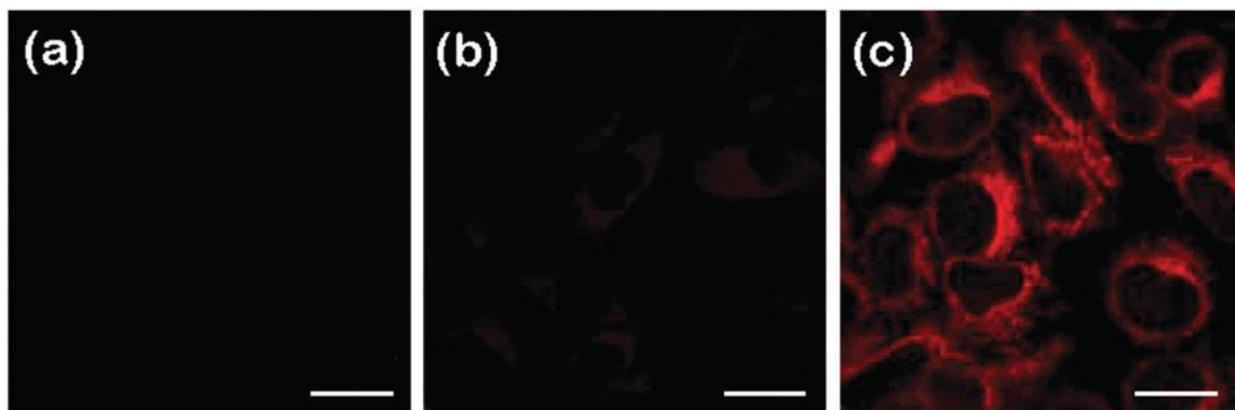
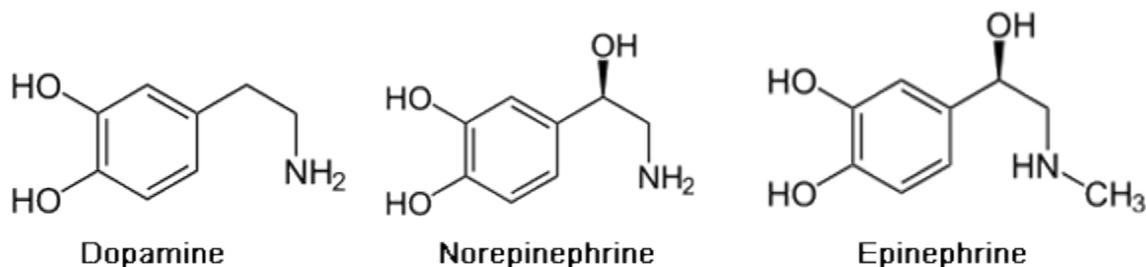


Figure 3.2. Fluorescent images of B16F10 cells (mouse skin cancer cell line). (a) B16F10 cells only; (b) the cells incubated with P1; (c) image of B16F10 cells incubated sequentially with P1 and  $F^-$ . The scale bar is 20  $\mu\text{M}$  long. Reprinted by permission from [7].

There are many reports on triggering fluorescence turn-on upon ions-binding to a fluorophore and their applications for imaging and detecting ions in different cells [8, 9]. There is also a special attention on detecting inorganic cations in cells using fluorescence turn-on based on ion displacement. For example a synthetic fluorophore was quenched upon binding to  $\text{Hg}^{2+}$  and recovered by adding  $\text{KI}$ .  $\text{I}^-$  displaced  $\text{Hg}^{2+}$ . [10] In another study, a dye was quenched in MCF-7 cells by  $\text{Cu}^{2+}$  and turned-on by adding  $\text{S}^{2-}$  who displaces the  $\text{Cu}^{2+}$ . [11]  $\text{Cu}^{2+}$  could make a redox and hydrolysis reaction on a Rhodamine B hydrazide inside HeLa cells to turn it into a fluorescent probe. [12] Copper was sensed in another study by a synthetic fluorescent dye that showed fluorescence turn-on upon binding to  $\text{Cu}^{2+}$  in HeLa cells. [13] Not only cations but thiole groups also could trigger a fluorescence turn-on *in vivo* by reacting with non-fluorescent compounds [14, 15].

Fluorescence turn-on revealed apoptosis. Caspase-3/7 becomes active when cells undergo apoptosis. The caspase-3/7 activity in cells were monitored by fluorescence turn-on. [16] This could be a proof-of-principle for manufacturing antibody-free kits for detecting apoptosis.

Neurosensor 521 invented by Glass and coworkers could selectively stain the cells containing norepinephrine. [17] Neurosensor 521 turned into a fluorescent dye by binding to a norepinephrine molecule. The introduced neurosensor can differentiate norepinephrine from epinephrine.



Using nanoparticles for biological and medical purposes is a hot topic in science. Antibodies decorated with  $\beta$ -cyclodextrin attached to biomarkers on different cells who then hosted adamantane-modified nanoparticles.[18] Due to high quenching properties of Graphene (monolayer carbon sheet), Graphenes Nano particles have been used as a matrix for immobilizing different molecules *in vivo* [19-21] as well as *in vitro* [22, 23] for bio sensing. In one study, *p*-Sulfonated calix[6]arene (CX6, Figure 1.20) were immobilized on Graphene oxide nanoparticle. The obtained nanoparticle quenched the fluorescence of safranin T in HepG2 cells. The fluorescence turn-on was observed by incubating the cells with L-carnitine who can displace the dye form the immobilized CX6.

### 3.2.2. Applying calix[4]arene in cells

Bagnacani and *et al* have recently suggested using CX4 decorated with arginine as a cargo carrier for DNA into different cells. [24] *para*-amino-calix[4]arene-based solid lipid nanoparticle could also transfer DNA into cytoplasm of MDCK cells (Madin-Darby canine kidney cells). [25] CX4 has been used in fluorescence turn-on detection: A sensor was synthesized by decorating CX4 with a pyridyl-based triazole compound. The sensor was used for detecting  $Zn^{2+}$  in HeLa cells by a fluorescence turn-on response. [26] Calix[4]arenes have a tangible application: Dansylated calix[4]arene compounds was reported as a patent to be used for sensing mercury ion in veriaty biological samples like

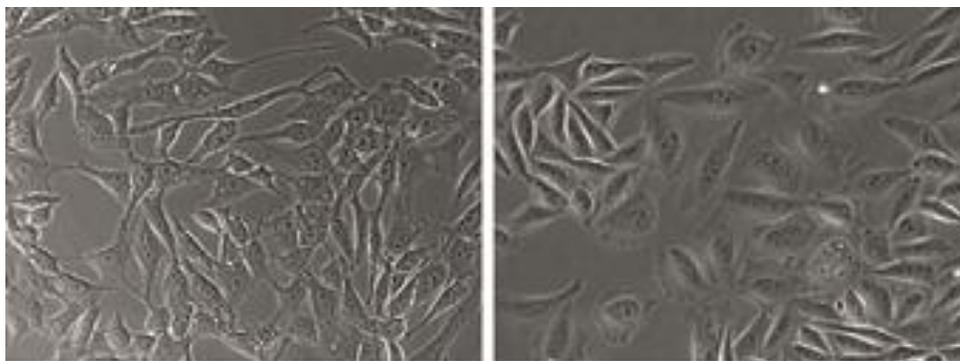
blood, cells, tissue, saliva and food.[27]

### 3.3. Results

CX4 quenches the fluorescence of the cationic fluorescent guest lucigenin (LCG) upon forming the CX4/LCG complex. The LCG can be displaced with cationic analytes showing a fluorescence response [28]:



The analyte can be Choline (Ch), Acetylcholine (ACh) or protamine. I showed that the displacement depicted in equation 3.1 can be also done inside live cells in the presence of numerous biomolecules with different size and charge. We used Chinese hamster fibroblast, V79, and Chinese hamster ovary, CHO, cells (Figure 3.3) for this purpose.

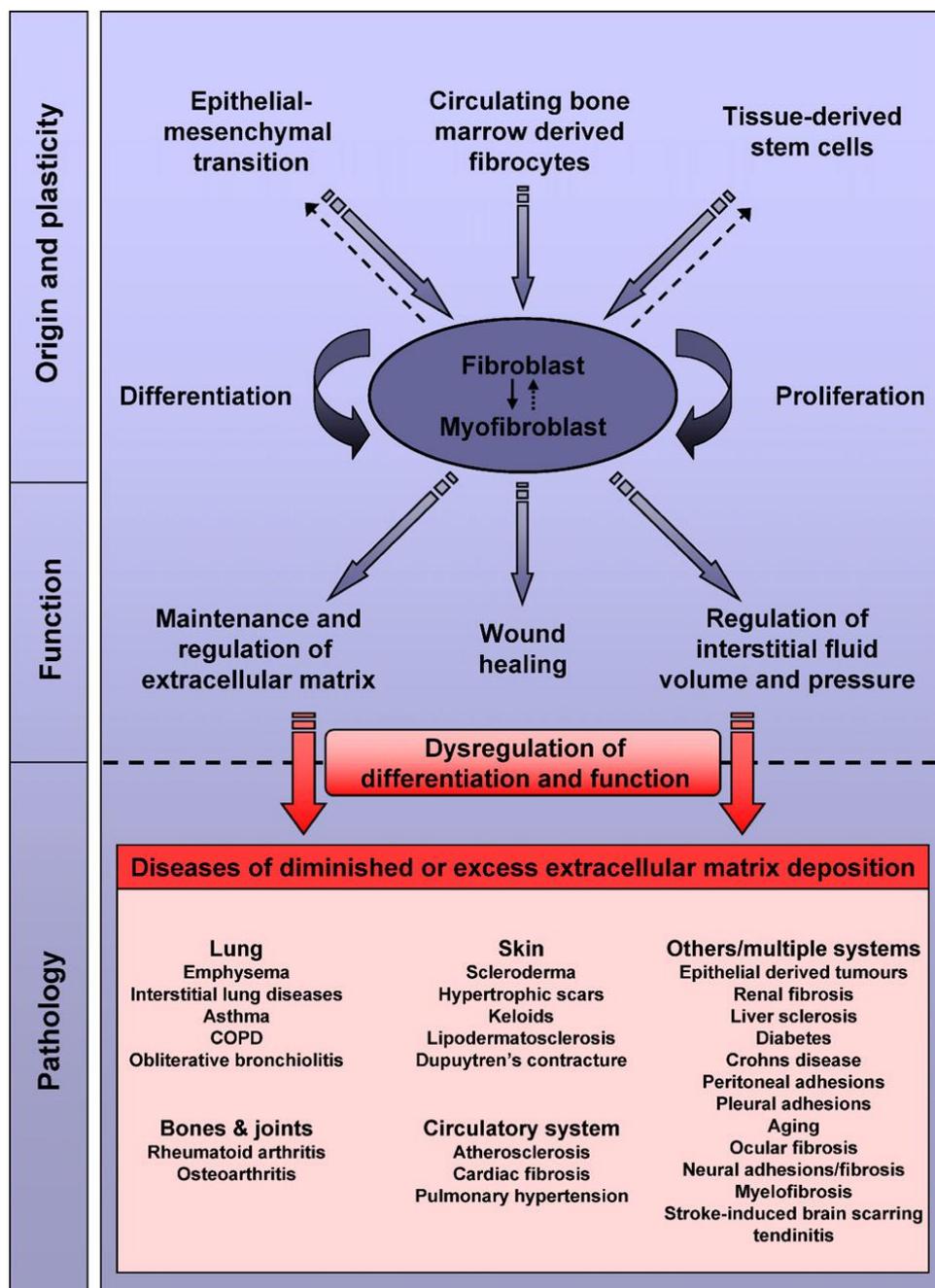


**Figure 3.3.** Bright-field images of CHO (left) and V79 (right) cells, with 40X magnification.

V79 is a fibroblast cell line. Fibroblast cells are found in most tissues. Fibroblasts cells are playing critical roles such as regulation of extracellular matrices, interstitial fluid volume and pressure and also wound healing. Fibroblasts dysfunction results in many diseases. [29] Figure 3.4 introduce the fibroblasts sources, functions and the related diseases.

In this chapter the used cells, cellular uptake of the reporter pair, supramolecular chemistry of the displacement and finally cytoplasmic and endosomal

displacement of Lucigenine with analytes is explained.



**Figure 3.4.** Schematically representation of fibroblast cell sources, function and the diseases related to their dysfunction. Reprinted by permission from [29].

CHO cells are also widely used in biology. 70% of all recombinant proteins are produced by CHOs! Recombinant proteins for therapeutically purposes have to be expressed in eukaryotic cells to obtain all post-translational modifications such as glycosylation. [29]

History and use of CHO as host for protein expression was reviewed by Jayapal and *et al.* [29] CHO and V79 therefore are two cell lines that are commercially available and are broadly in use. The cells have different function and origin therefore, the reproducibility of the host–dye displacement in these two cell lines can quite confidently generalizes to any cell.

### 3.3.1. CX4 as a carrier

My results show that both cells tend to uptake bound LCG (CX4/LCG) more than LCG. The uptake efficiencies of LCG in free and bound form are presented in Table 3.1.

**Table 3.1.** Cellular concentrations and uptake efficiency of LCG in the free and in the complex with CX4<sup>a</sup>.

Cell line	[CX4]/mM	[LCG in cells]/ $\mu\text{M}^b$	Uptake/% <sup>b</sup>
V79 cells	0	75	9.4
	4.5	150	19
CHO cells	0	10	1.3
	4.5	20	2.5

<sup>a</sup> Quantified with the lysis method. <sup>b</sup> Incubation conditions: 800  $\mu\text{M}$  LCG and 30 min.

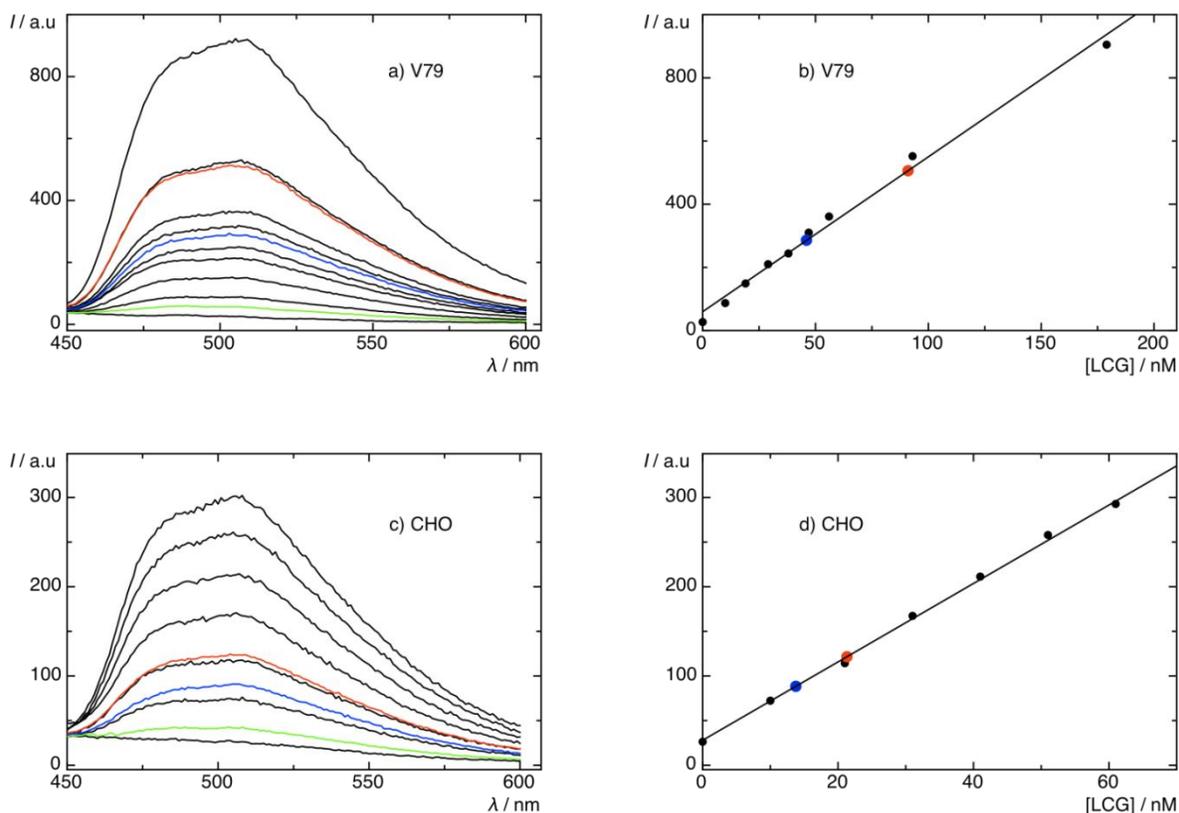
Uptake efficiency of LCG is almost double in the form of CX4/LCG compare to uptake efficiency of the free LCG. In other word, CX4 could facilitate the uptake of LCG by the cells. CX4 therefore might be used as a carrier. CB7 can also facilitate a dye uptake into the cytoplasm. [4] The uptake efficiency of LCG or CX4/LCG is significantly higher in V79 than in CHO cells (Table 3.1). This might be due to the fact that V79 by the size of 421  $\mu\text{m}^3$  (in the suspended, spherical, form) is almost 4 times smaller than CHO by average size of 1636  $\mu\text{m}^3$ . The larger size of CHOs in the adherent form is detectable even with naked eye (Figure 3.3) Therefore, area-to-volume ratio in V79 ( $272 \mu\text{m}^2/421 \mu\text{m}^3=0.65 \mu\text{m}^{-1}$ ) is almost as twice as in CHO ( $618 \mu\text{m}^2/1636 \mu\text{m}^3=0.37 \mu\text{m}^{-1}$ ). The bigger area to volume ratio

in V79 might explain the higher uptake efficiency by it because larger area-to-volume increases the diffusion probability and consequently higher accumulation of the diffused substance.

### **Quantification of LCG and CX4/LCG uptake**

For calculating the uptake amount and efficiencies reported in table 3.1: Almost the same number of V79 or CHO cells was cultivated in three 60.1 cm<sup>2</sup> dishes. In Dish 1 the cells were incubated with 800 μM LCG for 30 minutes, in Dish 2 they were incubated with CX4/LCG by host-dye ratio 5/1 and [LCG] = 800 μM (H/D = 5/1, [D] = 800 μM) for 30 minutes. The cells in Dish 3 were just used for plotting standard curves by a method explained following (Figures 3.5 b,d).

After the incubations, the dishes 1&2 were washed three times with the warm medium (37 °C) to remove un-uptake LCG and CX4/LCG molecules and followed by trypsinization. The number of the suspended cells in the homogenous fraction was counted on a cytometer for estimating the total number of cells on each dish. For collecting their cytoplasmic component, the suspended cells were then lysed by triton X-100 following a protocol described in the literature. [30] All the lysed products were equally diluted by adding water. The fluorescent spectrum of the obtained solution from Dish 2 were recorded by exciting at 369 nM and scanning from 450 – 600 nm (green spectra in Figures 3.5a,c). Then, excess amount of protamine was added to the obtained solutions from all dishes. Protamine releases the LCG from CX4/LCG in the lysed product from Dish 2, whose fluorescence spectrum shows a significant increase (red spectra in Figure 3.5a,c). This experiment further proves that both CX4 and LCG are coexisting in the cells' cytoplasm. The spectra of the solutions from the Dish 1 (free LCG uptake) were simply recorded (the blue spectra). Protamine was added into the solution from other Dishes (1, 3) only for having identical components in all solutions which in turn makes all the recorded spectra comparable.



**Figure 3.5.** Emission spectra of LCG on left and the derived standard curves on right for quantifying the LCG and CX4/LCG uptake by V79 and CHO cells. Green spectrum was recorded from the lysed product of Dish 2 containing uptake CX4/LCG, red spectrum is from the liberated LCG from the intracellular CX4/LCG. The blue spectrum is from the lysed product from Dish 1 containing uptake LCG. Black spectra were obtained by adding LCG into cellular lysed product from the Dish 3 contained only cells without CX4 or LCG incubation.

Usually for estimating an analyte concentration, a standard curve is needed. Obviously a standard curve has to be obtained from a solution that is identical as the solution where analytes dissolved in. Therefore, LCG had to be added gradually to the lysed products for obtaining emission spectra and it could not be done simply in water. Therefore, Dish 3 with the same number of cells was lysed and diluted with water and protamine by the same way that was done for Dishes 1 and 2. The standard curves for CHO and V79 cells were obtained

by adding known amount of LCG to the lysed solutions of the Dishes 3, recording the fluorescence intensities (black spectra in Figure 3.5a,c) and finally plotting the fluorescent intensities at  $\lambda_{\max} = 505$  nm against the LCG concentrations.

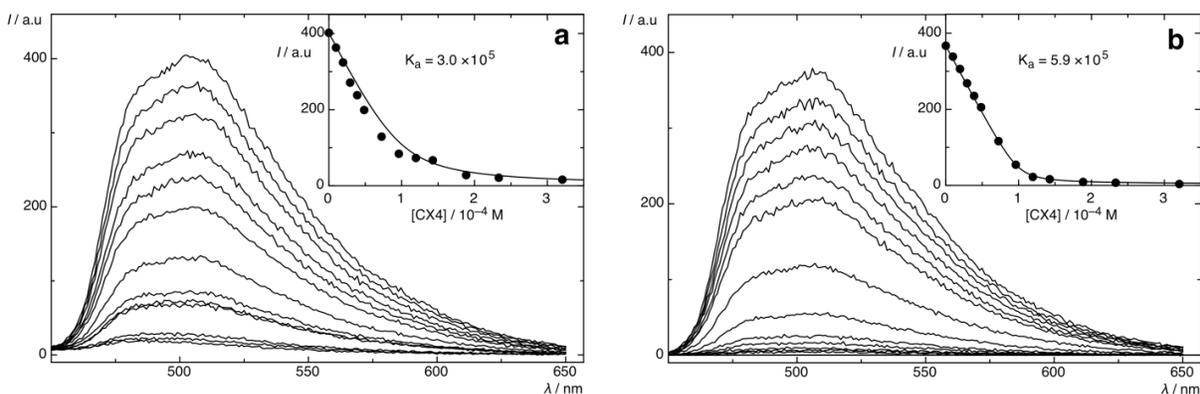
Cells volumes were estimated for calculating the concentration of uptake LCG in each cell. I assumed that volume of a cell remains the same whether it is suspended or it is attached to a surface. By this assumption, the volume of spherical suspended cells on average were estimated by knowing the diameter of the suspended cells ( $d=9.3$   $\mu\text{M}$ , of V79 and  $d= 14.02$   $\mu\text{M}$  of CHO) from the literature. [31, 32] Data in table 3.1 were calculated by using the LCG concentration in the lysed product using the standard curve, number of cells in each dish, and their average volume, all explained above.

### 3.3.2. Loading cells with CX4/LCG complex

There are countless biomolecules in the medium as well as in the cytoplasm of the cells with various size and charge. Some of these molecules are competing with LCG for binding to CX4. Nevertheless, this competition does not abolish LCG binding to CX4. The binding constant value of LCG to CX4 in medium is  $3.0 \times 10^5 \text{ M}^{-1}$  and it is just half of the binding constant value in Tyrode's solution which is  $5.9 \times 10^5 \text{ M}^{-1}$  (Figure 3.6a,b). The cations of the Tyrode's solution<sup>3</sup> by a total concentration of 155.1 mM, can compete with LCG for binding to CX4 [33]. The Ham's medium has cations by a total concentration of 148.5 mM and additional 20 standard amino acids, vitamins, 10% fetal bovine serum and *etc.* [34]. The salts concentration in both Tyrode's and the medium are nearly the same and dominates weakening the binding constant from  $10^7 \text{ M}^{-1}$  in water to about  $10^5 \text{ M}^{-1}$  in both Tyrode's solution and the media. Together with the fluorescence recovery results (Figure 3.12) this evidences that CX4 could keep it host in the presence of many biomolecule with different size and charges.

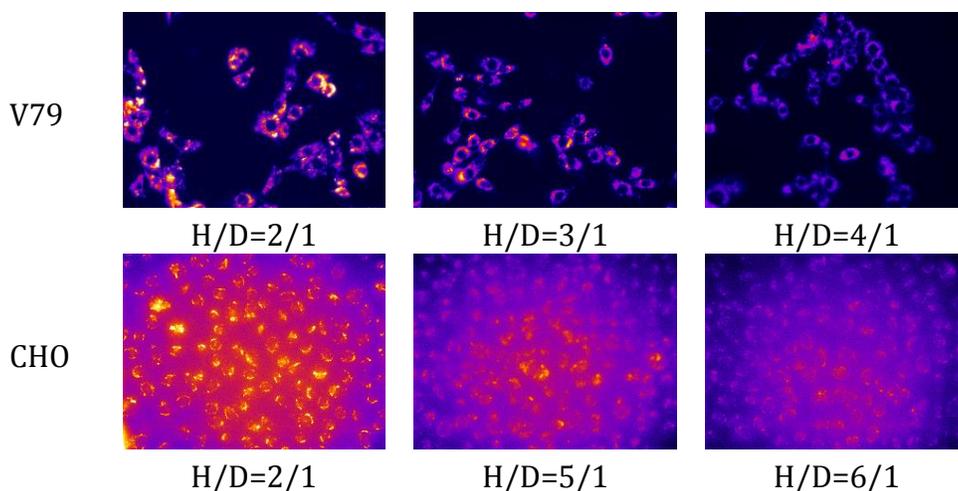
<sup>3</sup> 137 mM NaCl, 2.7 mM KCl, 1 mM MgCl<sub>2</sub>, 1.8 mM CaCl<sub>2</sub>, 0.2 mM Na<sub>2</sub>HPO<sub>4</sub>, 12 mM NaHCO<sub>3</sub>, 5.5 mM D-glucose. Made according to the receipt NO 10479 from Cold Spring Harbor Protocolled (doi:10.1101/pdb.rec10479)

CX4 by the concentrations 4 times higher than LCG in the medium ( $CX4/LCG > 4/1$ ), quenches LCG down to 10% of its initial fluorescence intensity. The inset curve in the Figure 3.6a also reaches to a plateau at CX4 concentrations about this ratio.



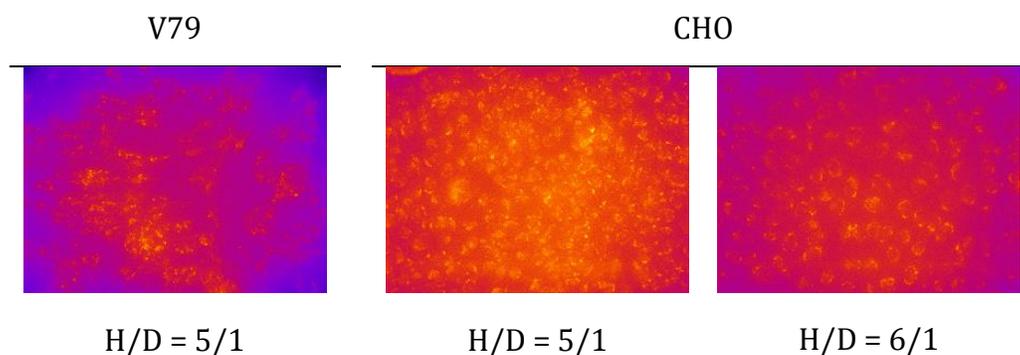
**Figure 3.6.** Direct fluorescence titration of LCG (100  $\mu$ M) with CX4 in V79's medium (a) and in Tyrode's solution (b), pH = 7,  $\lambda_{ex} = 369$  nm. The inset shows the associated titration curve,  $\lambda_{max} =$  average 503-507 nm, and fit according to a 1:1 binding stoichiometry.

V79 and CHO cells loaded with different host–dye ratios are shown in Figure 3.7. By increasing the host ratio, both cells show less fluorescence.



**Figure 3.7.** Fluorescence images of V79 and CHO cells. From left to right, the cells were incubated with increasing host–dye ratio and a constant dye concentration of 50  $\mu$ M.

However the CHO cells need more host for a proper quenching. For example CHO cells by  $H/D = 5/1$  are brighter than V79 cells with the same ratio of host–dye (Figure 3.8), and they need higher concentration of the host ( $H/D = 6/1$ ), for having the same brightness. Compare to V79 cells, the CHO cells might contain more LCG-releasing biomolecules like ACh, Ch or even arginine rich (positively-charged) proteins in their cytoplasm. The arginine or lysine can also release LCG from CX4 by weakly bind to CX4 ( $K_d = 10^3$  in water) [35]. Based on these adjustments (Figure 3.7 and 3.8), V79 cells were incubated with CX4/LCG 5:1 and CHO cells were incubated with CX4/LCG 6:1 and compared to the cells that showed fluorescent recovery followed by adding analytes in the Figure 3.12.



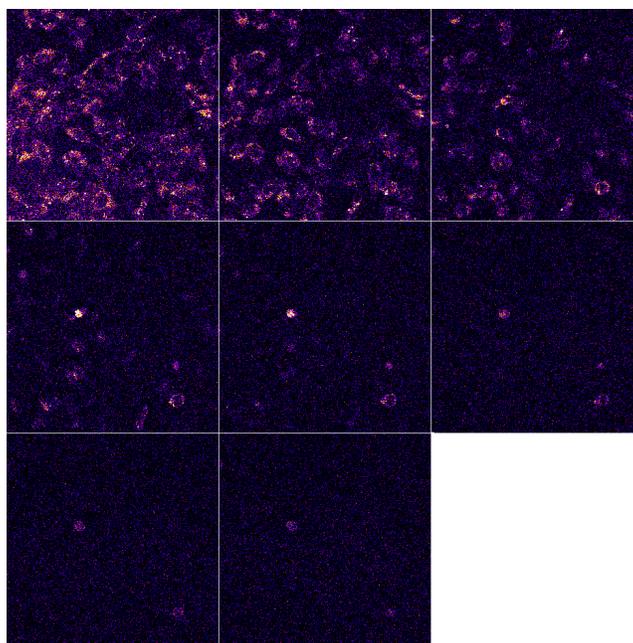
**Figure 3.8.** Fluorescence images of V79 and CHO cells incubated with host–dye complex. CHO cells need more host for proper quenching.

### 3.3.3. The displacement of LCG with cationic bio-analytes from CX4 inside live cells

The displacement of LCG with different cationic bio-analytes from CX4/LCG is represented schematically in Figure 3.11. For loading with CX4/LCG, V79 or CHO cells were first incubated with 250  $\mu\text{M}$  of CX4 and 50  $\mu\text{M}$  of LCG for 10 to 15 min at 37  $^{\circ}\text{C}$ ,  $\text{pH} \approx 7.2$ . The cells were then washed three times with fresh and warm medium and incubated with 50mM acetylcholine or 50mM choline dissolved in the medium for 10 minutes at 37  $^{\circ}\text{C}$  or with 200  $\mu\text{M}$  protamine dissolved in the medium for 15 - 20 minutes at 37  $^{\circ}\text{C}$ . For having the constant temperature during the incubation, the analyte solutions were warmed up to 37 $^{\circ}\text{C}$  prior to incubation. Finally the analytes solutions were collected and cells were washed 3 times with the Tyrode's solution and subjected to imaging under the Tyrode's solution.

Analytes release LCG from CX4 instantly (in milliseconds) in cuvette, [36] but the analyte solution did not show any fluorescence emission by exciting at 369 nm. Consequently, there was not host–dye complexes physically adsorbed on the cell membrane surface or reconstituted into the outer cell membrane bilayer. Moreover, a fluorescent CX4 derivative was not found in the cell membrane.[37] Our Z-stacking imaging results by using a confocal laser scanning microscopy (CLSM, Figure 3.9) also shows distribution of released dye in the whole volume of the cells. This last result completely renders the accumulation of the CX4/LCG complex in the cell membranes. However due to photo bleaching property of LCG the images show fluorescence lost after 3 times scanning.

For obtaining the CLSM images, monolayer of V79 cells were incubated with CX4/LCG by in a mixture of 50  $\mu$ M [LCG] and 250  $\mu$ M [CX4]. The LCG was released by incubating cells with 50 mM Ch following a method explained above. The cells were subjected to 3D scanning using a confocal laser scanning microscope by exciting LCG with a 405 nm laser (Figure 3.9). Vertical scanning from bottom to top shows distribution of the dye in the entire volume of the cells. Because of the photo bleaching property of LCG using CLSM is not an advantage here. Nevertheless Figure 3.9 clearly shows that the dye is not accumulated in a particular region of a cell and it is rather distributed in the cytoplasm with a punctate pattern. LCG stains cells with a punctate pattern.

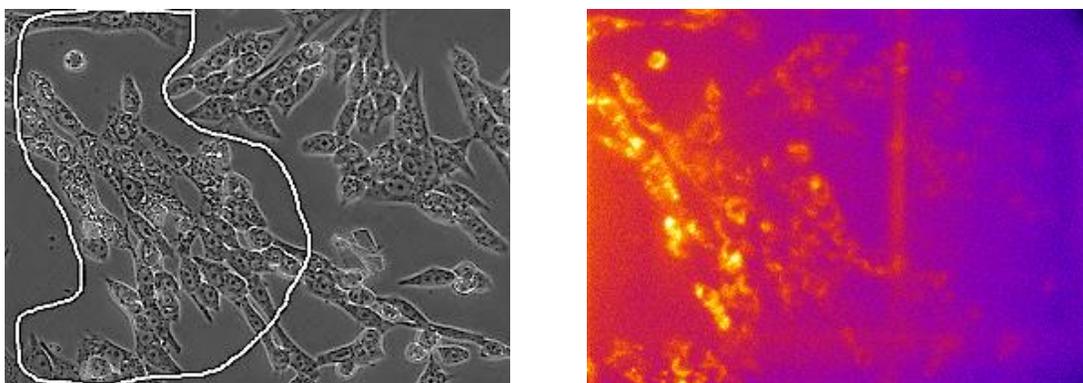


**Figure 3.9.** Fluorescence images of the released dye (LCG) from CX4/LCG in a cell monolayer recorded by CLSM 3D scanning. From left to right in each row, one layer of the cells was scanned from bottom of the dish to the top (Z stacking).

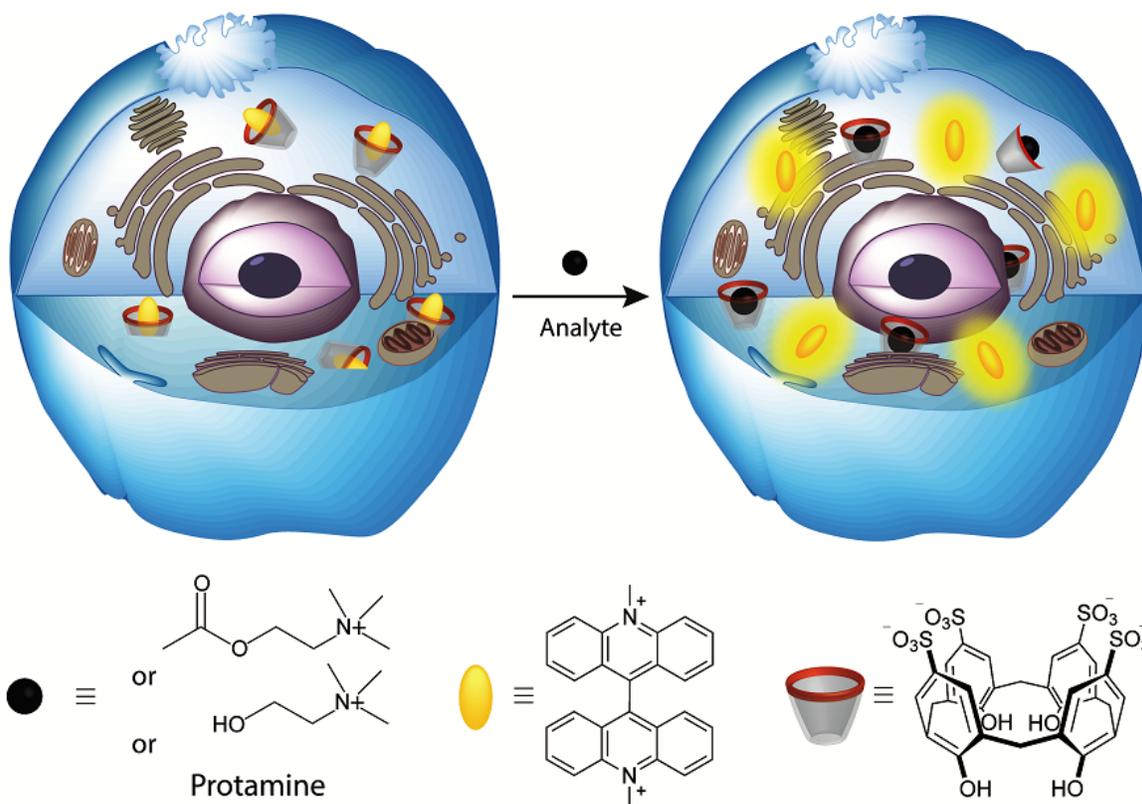
### **Dye-release by Micro Injection**

V79 cells were incubated with HD 5:1, [D]=10  $\mu\text{M}$  dissolved in medium for 10 min. Then the cells were washed with Tyrod's and objected to cytoplasmic injection with ACh 100 mM using a FemtoJet Eppendorf. The capillary and injection pressure were 60 and 190 hpa respectively and ACh has been injecting for 3 seconds. The injections were about 15 minutes after washing. The injected cells were encircled in the bright field image of Figure 3.10. A fluorescence response is observable on almost all of the injected cells.

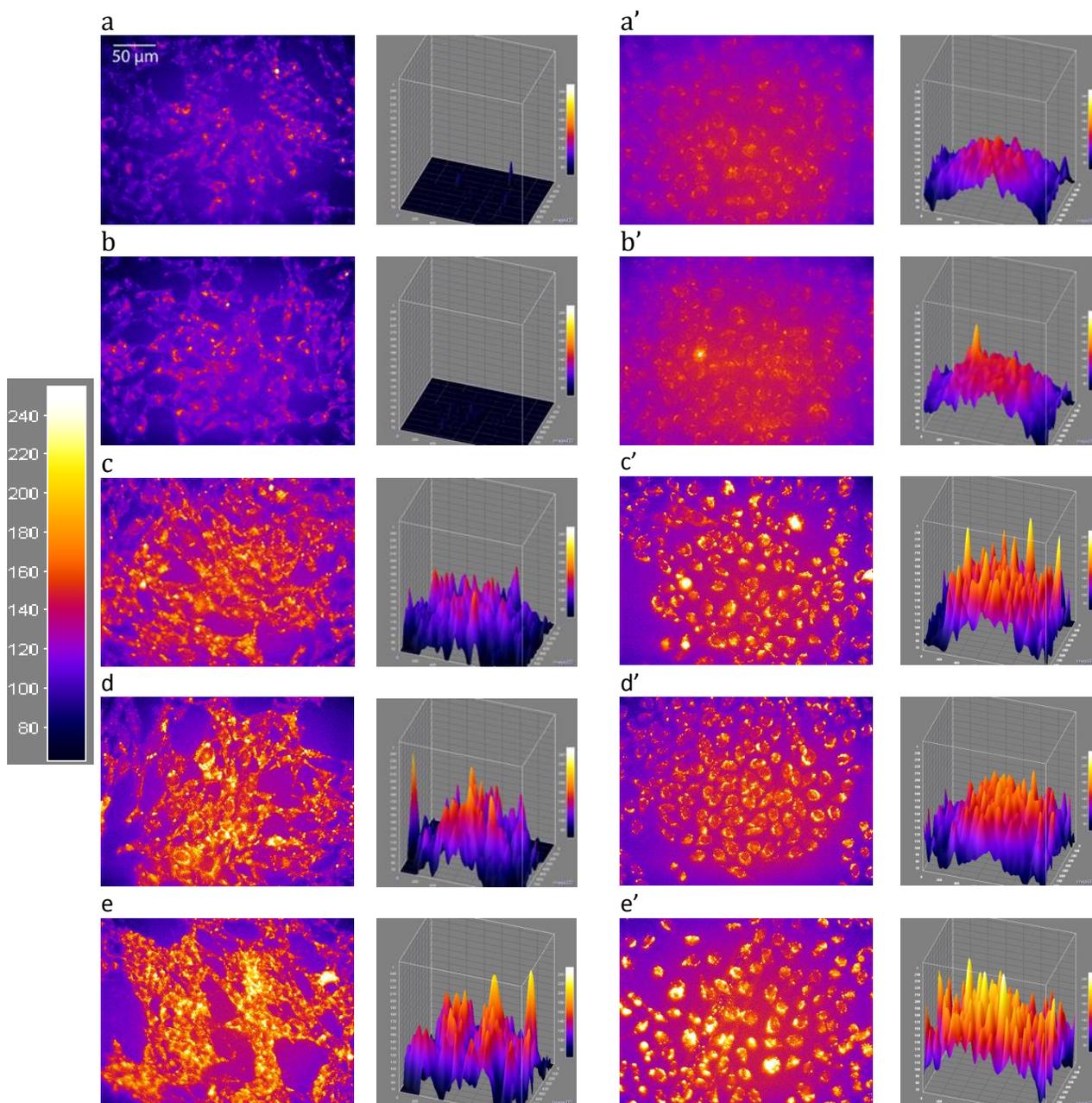
The host/dye complex is not free in the cytoplasm for a long time and micro injection is time consuming and invasive (can break the cell membrane), therefore I did not have reproducible results. Nevertheless, occasionally a weak and immediate fluorescence recovery was observed after injecting analytes (Figure 3.10) which further proves that CX4/LCG molecules were free in the cytoplasm.



**Figure 3.10.** Bright field on the left and fluorescence image on the right of the injected cells. ACh was injected into the cytoplasm of the cells previously loaded with CX4/LCG. The bright field image on the left shows the injection area encircled and fluorescence recovery image is on the right.



In this project we clearly showed that, CX4/LCG complex passed through the cell membrane and remained complex in the cytoplasm of the cells despite the presence of many biomolecules of various size and charge. The analytes passed through the cell membranes with different mechanisms explained below, and displace the LCG from CX4/LCG complex, whose fluorescence recovery was followed by fluorescence microscopy (Figure 3.12). Control experiments showed that the displacement also worked when the reporter pair and analytes incubation is reversed *i.e.* the cells were preloaded by means of analytes incubation (ACh and Cho for 2 hours and Protamine for 30 minutes) followed by washing the cells three times with warm medium and CX4/LCG incubation for 15 minutes. LCG itself can stain the cells in a punctate pattern [38] similar to images represented here. For preparing the images in Figure 3.12, the 16-bit fluorescence images from the fluorescence microscope were added to an imageJ stack for having a comparable and homogenized background for further processing. The stacked-images were painted “fire”. Using stacks make any brightness and contrast adjustment unbiased and equal on each image in the stack. The stack split back into images and saved as tiff format (Figure 3.12).



**Figure 3.12.** V79 Cells (left column) were incubated with a mixture of 50  $\mu\text{M}$  LCG and 250  $\mu\text{M}$  CX4 and CHO cells (right column) with 50  $\mu\text{M}$  LCG and 300  $\mu\text{M}$  CX4 at  $\theta = 37^\circ\text{C}$  for 15 min, followed by incubation with medium (control experiment a and a'), 50 mM betaine (non-analyte but cell permeable control b, b') 50 mM choline (c, c'), 50 mM acetylcholine (d, d') for 10 min and with 0.9  $\text{mg}\cdot\text{ml}^{-1}$  of protamine for 15-20 min (e, e'). The 3D surface plot quantifies the fluorescence intensity of images.

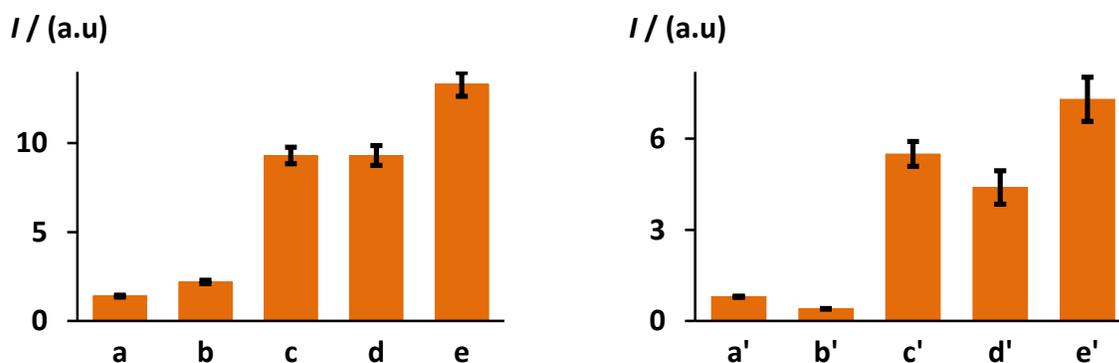
In non-neuronal cells Ch is being uptake mainly by a Ch transporter via facilitated diffusion mechanism which is sodium-independent and is with low-affinity. [39, 40] However HEK-293 cell line which is a kidney-derived fibroblast cell also shows high affinity sodium-dependent uptake for Ch. [41] ACh can cross membrane in both directions by organic cation transporters (OCTs). [42] OCTs are expressed in almost all cells. [43] While both ACh and Ch simply pass the membrane via a facilitated diffusion, LCG displacement with protamine inside cells seems be more complicated. However protamine is known since 1960s for in taking RNA into cells [44, 45] but its uptake mechanism, kinetics and its final fate has recently been reported by Nagai and coworkers. [46] Protamine is being uptake by a receptor-mediated clathrin- and caveolin-dependent endocytosis and appeared in the cytoplasm and nucleus of Opossum Kidney cells.[46] Over 80% of protamine is being uptake in 15 minutes and in 30 min all of the protamine is taken up by the cells. Our results also indicate that protamine notably shows a strong fluorescence turn-on in 15 minutes. Therefore, sufficient amount of protamine can be endocytosed in 15 minutes in both V79 and CHO cells. But the question is how a protamine meets CX4/LCG molecules while it is entrapped in an endosome? Do CX4/LCG molecules diffuse to an endosome? Or do protamines escape the endosome? Protamine transported RNA into cells in about 15 – 30 min and the transported RNA have been translated into the corresponding protein. [45] The transferred RNA had to escape the endosome for translation – to proteins by ribosomes – therefore, we believe that there is an endosomal escape mechanism for protamine. Endosomal escape for DNA delivery using protamine and other polycations have also been reviewed. [47, 48]

Specificity of sensing in biological systems is challenging due to presence of many biomolecules with different charge and size which could be potentially a competitor for sensing.[49] For testing the specificity of the analytes for releasing LCG we incubated the cells with 50 mM betaine (in a same condition for analytes) followed by host–dye incubation explained above. Betaine cannot release LCG from CX4, [28] however it could actively (mainly Na<sup>+</sup>-dependent) being transported into SV-3T3 cells. [50]

Mechanism of some CX4 derivatives uptake by CHO cells was studied by Mueller and coworkers [37, 51, 52]. They found out that the uptake is non-specific not connected to the caveolar or clathrin-dependent endocytosis process. The calixarene derivative was accumulated in the cytoplasm and it was not observed in the nucleus and cell membrane. Further studies by using CHO, THP-1 and HeLa cells showed that uptake of CX4 inked to a dye is fast and after uptake by help of a plasma membrane carrier they travel to Golgi and then continue moving into lysosomes.[52] Therefore the CX4 is not free in the cytoplasm for a long time. We did not also observe a meaningful fluorescent recovery after 5 hours from initial exposure of the cells to the CX4/LCG complex. Moreover, shorter incubation time with CX4/LCG (10 – 15 min) gave a better fluorescent recovery results followed by analyte incubation. From above results we can conclude that CX4/LCG is not available to the analytes in the cytoplasm for a long time.

### Quantifying fluorescence intensity of a cell

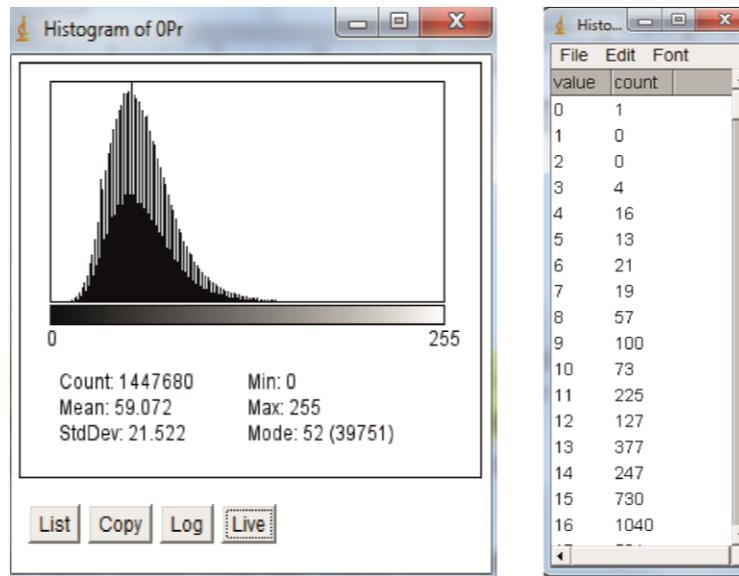
An innovative method is being introduced here for quantifying the average fluorescence intensity of each single cell in focus presented in figure 3.12. The obtained results are compared in Figure 3.13, which clearly shows the fluorescence recoveries in entities c-e and c'-e'. The fluorescence recoveries are higher with protamine albeit its lower concentration.



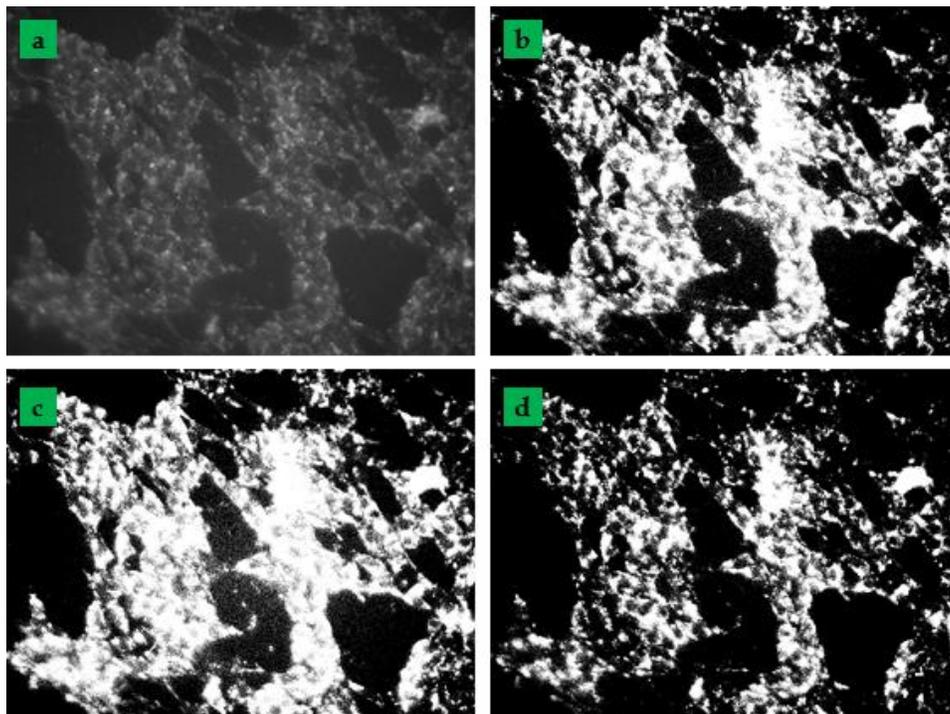
**Figure 3.13** Average fluorescence intensities of each cell images presented in Figure 3.12.

I went on the following steps using imageJ and excel for calculating the average fluorescence of a cell: The original images (16-bits) were converted to 8-bits type by

imageJ. A histogram of an image from imageJ shows the distribution of fluorescence values of spots (Fluorescence images taken by the digital cameras are made of spots). In 8-bits images, black and white colors have the value of 0 and 255 respectively. The fluorescence intensity of a spot therefore has a value between 0 and 255 depending on its intensity. The histogram (Figure 3.14 left) shows the distribution of the spots intensity. It is possible to extract the list of spots and their relative value (Figure 3.14 right). Magnitude of each value was obtained by multiplying the value to its count (the number of the repeated spots in the whole image disregards of its location) and I call it here *value-count (VC)*. The obtained VC includes the background of course, which needs to be eliminated. For eliminating the background, the threshold of the meaningful fluorescence has to be determined. Several methods have been introduced for thresholding but none of them works better than determining a threshold manually. Figure 3.15 represents a comparison between manual thresholding (d) and Otsu (b) and Li (c) methods. Manual thresholding leaves the least white spots on the background; consequently the cell images are crisper and clearer. After determining the thresholds, the VC values corresponding for the backgrounds were deleted from of the list of VCs. The sum of the remaining VCs arbitrarily shows a total fluorescence obtained from the cells and by dividing it by the number of cells, average fluorescence per cells is obtained. Finally the obtained average fluorescence is shown as bar graphs in the Figure 3.13. Error bars individually calculated from the standard deviations from the mean fluorescence value. The mean fluorescence value had been calculated by ImageJ.



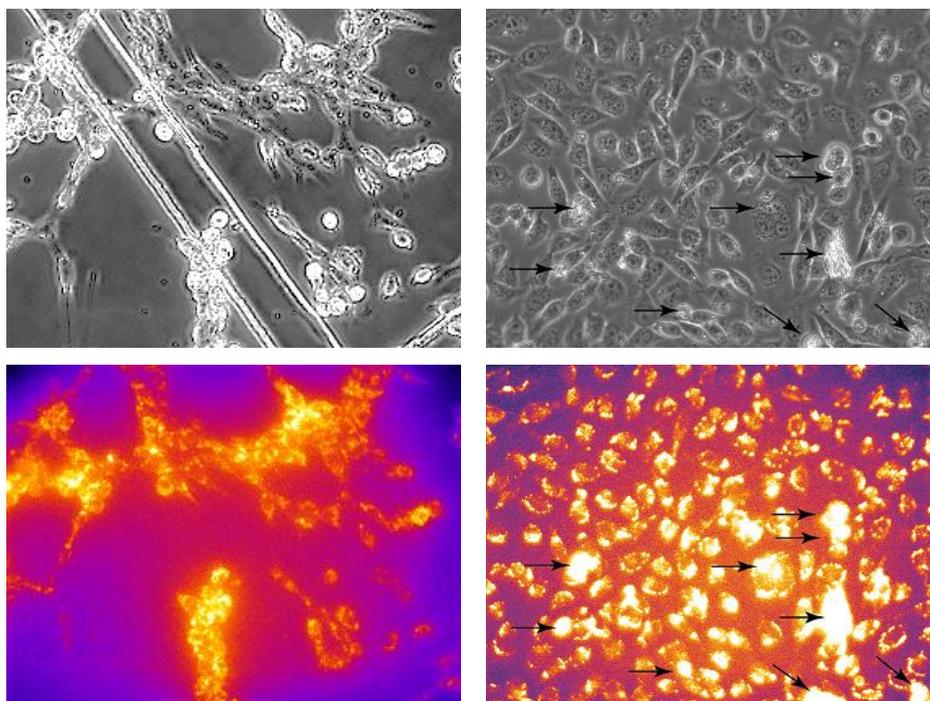
**Figure 3.14.** A histogram of an 8-bits image (Figure 3.15, a) on the right and its digital value vs counts listed on the left.



**Figure 3.15** Fluorescence images of the V79 cells. An 8-bit image (a) was subjected to different thresholding methods: Otsu (b), Li (c) and manually (d).

### 3.3.4. Endosomal dye-release

Higher concentration of protamine can damage the cells by disrupting all of their membrane bilayers, release all of the enclosed CX4/LCG from acidic vesicles (explained above) and makes the strongest fluorescence on the dead cells debris (Figure 3.16). While almost all of the V79s were went necrosis only few of the CHO cells were affected therefore, CHO cells seem to be more resistance to toxic concentration of protamine than V79 cells.



**Figure 3.16.** Toxic concentration of protamine ( $18 \text{ mg}\cdot\text{ml}^{-1}$ ) disrupts the cell organelles and additionally releases LCG from entrapped CX4/LCG. On the left V79 cells and on the right CHO cells. Some of the damaged CHO cells were pointed by arrows, show stronger fluorescence intensity compare to intact cells.

### 3.4. References

- [1] Duan QP, Cao Y, Li Y, Hu XY, Xiao TX, Lin C, et al. pH-Responsive Supramolecular Vesicles Based on Water-Soluble Pillar[6]arene and Ferrocene Derivative for Drug Delivery. *J Am Chem Soc*, **2013**;135:10542-10549.
- [2] Zhang J, Yuan ZF, Wang Y, Chen WH, Luo GF, Cheng SX, et al. Multifunctional Envelope-Type Mesoporous Silica Nanoparticles for Tumor-Triggered Targeting Drug Delivery. *J Am Chem Soc*, **2013**;135:5068-5073.
- [3] Yin JJ, Sharma S, Shumyak SP, Wang ZX, Zhou ZW, Zhang YD, et al. Synthesis and Biological Evaluation of Novel Folic Acid Receptor-Targeted, beta-Cyclodextrin-Based Drug Complexes for Cancer Treatment. *Plos One*, **2013**;8:0062289.
- [4] Li ZY, Sun SG, Yang ZG, Zhang S, Zhang H, Hu MM, et al. The use of a near-infrared RNA fluorescent probe with a large Stokes shift for imaging living cells assisted by the macrocyclic molecule CB7. *Biomaterials*, **2013**;34:6473-6481.
- [5] Hu Y, Chai MY, Yang WT, Xu FJ. Supramolecular Host-Guest Pseudocomb Conjugates Composed of Multiple Star Polycations Tied Tunably with a Linear Polycation Backbone for Gene Transfection. *Bioconjugate Chem*, **2013**;24:1049-1056.
- [6] Jeon S-J, Kwak S-Y, Yim D, Ju J-M, Kim J-H. Chemically-Modulated Photoluminescence of Graphene Oxide for Selective Detection of Neurotransmitter by "Turn-On" Response. *J Am Chem Soc*, **2014**;136:10842-10845.
- [7] Kim D, Singha S, Wang T, Seo E, Lee JH, Lee SJ, et al. In vivo two-photon fluorescent imaging of fluoride with a desilylation-based reactive probe. *Chem Commun*, **2012**;48:10243-10245.

- [8] Domaille DW, Que EL, Chang CJ. Synthetic fluorescent sensors for studying the cell biology of metals. *Nat Chem Biol*, **2008**;4:168-175.
- [9] Que EL, Domaille DW, Chang CJ. Metals in neurobiology: Probing their chemistry and biology with molecular imaging (vol 108, pg 1517, 2008). *Chem Rev*, **2008**;108:4328-4328.
- [10] Mahapatra AK, Roy J, Sahoo P, Mukhopadhyay SK, Chattopadhyay A. Carbazole-thiosemicarbazone-Hg(II) ensemble-based colorimetric and fluorescence turn-on toward iodide in aqueous media and its application in live cell imaging. *Org Biomol Chem*, **2012**;10:2231-2236.
- [11] Wang JL, Long LP, Xie D, Zhan YW. Highly selective fluorescence turn-on sensor for hydrogen sulfide and imaging in living cells. *J Lumin*, **2013**;139:40-46.
- [12] Yu MX, Shi M, Chen ZG, Li FY, Li XX, Gao YH, et al. Highly sensitive and fast responsive fluorescence turn-on chemodosimeter for Cu(2+) and its application in live cell imaging. *Chem-Eur J*, **2008**;14:6892-6900.
- [13] Sivaraman G, Anand T, Chellappa D. Quick accessible dual mode turn-on red fluorescent chemosensor for Cu(II) and its applicability in live cell imaging. *Rsc Adv*, **2013**;3:17029-17033.
- [14] Cui K, Chen ZL, Wang Z, Zhang GX, Zhang DQ. A naked-eye visible and fluorescence "turn-on" probe for acetyl-cholinesterase assay and thiols as well as imaging of living cells. *Analyst*, **2011**;136:191-195.
- [15] Kwon H, Lee K, Kim HJ. Coumarin-malonitrile conjugate as a fluorescence turn-on probe for biothiols and its cellular expression. *Chem Commun*, **2011**;47:1773-1775.
- [16] Shi HB, Wang G, Liang J, Liu B. Fluorescence turn-on detection of live cell apoptosis using a hyperbranched conjugated polyelectrolyte. *Medchemcomm*, **2013**;4:554-558.

- [17] Hettie KS, Liu X, Gillis KD, Glass TE. Selective Catecholamine Recognition with NeuroSensor 521: A Fluorescent Sensor for the Visualization of Norepinephrine in Fixed and Live Cells. *Acs Chem Neurosci*, **2013**;4:918-923.
- [18] Agasti SS, Liong M, Tassa C, Chung HJ, Shaw SY, Lee H, et al. Supramolecular Host-Guest Interaction for Labeling and Detection of Cellular Biomarkers. *Angew Chem Int Ed*, **2012**;51:450-454.
- [19] Wang Y, Li ZH, Hu DH, Lin CT, Li JH, Lin YH. Aptamer/Graphene Oxide Nanocomplex for in Situ Molecular Probing in Living Cells. *J Am Chem Soc*, **2010**;132:9274-9276.
- [20] Sun X, Liu Z, Welsher K, Robinson JT, Goodwin A, Zaric S, et al. Nano-Graphene Oxide for Cellular Imaging and Drug Delivery. *Nano Research*, **2008**;1:203-212.
- [21] Mao XW, Tian DM, Li HB. p-Sulfonated calix[6]arene modified graphene as a 'turn on' fluorescent probe for L-carnitine in living cells. *Chem Commun*, **2012**;48:4851-4853.
- [22] Chang HX, Tang LH, Wang Y, Jiang JH, Li JH. Graphene Fluorescence Resonance Energy Transfer Aptasensor for the Thrombin Detection. *Anal Chem*, **2010**;82:2341-2346.
- [23] Lu C-H, Yang H-H, Zhu C-L, Chen X, Chen G-N. A Graphene Platform for Sensing Biomolecules. *Angew Chem Int Ed*, **2009**;48:4785-4787.
- [24] Bagnacani V, Franceschi V, Bassi M, Lomazzi M, Donofrio G, Sansone F, et al. Arginine clustering on calix[4]arene macrocycles for improved cell penetration and DNA delivery. *Nat Commun*, **2013**;4.
- [25] Nault L, Cumbo A, Pretot RF, Sciotti MA, Shahgaldian P. Cell transfection using layer-by-layer (LbL) coated calixarene-based solid lipid nanoparticles (SLNs). *Chem Commun*, **2010**;46:5581-5583.

- [26] Pathak RK, Hinge VK, Rai A, Panda D, Rao CP. Imino-Phenolic-Pyridyl Conjugates of Calix[4]arene (L-1 and L-2) as Primary Fluorescence Switch-on Sensors for Zn<sup>2+</sup> in Solution and in HeLa Cells and the Recognition of Pyrophosphate and ATP by [ZnL<sub>2</sub>]. *Inorg Chem*, **2012**;51:4994-5005.
- [27] Pandey S, Chawla HM. New dansylated calix(4)arene compounds useful as chemosensing agent for testing for the presence of mercury ion in sample e.g. blood, cells, tissue, saliva, sweat, drink, food and water. *Pandey S; Chawla H M; Indian Inst Technology Delhi*.
- [28] Guo DS, Uzunova VD, Su X, Liu Y, Nau WM. Operational calixarene-based fluorescent sensing systems for choline and acetylcholine and their application to enzymatic reactions. *Chem Sci*, **2011**;2:1722-1734.
- [29] McAnulty RJ. Fibroblasts and myofibroblasts: Their source, function and role in disease. *Int J Biochem Cell B*, **2007**;39:666-671.
- [30] Ji H. Lysis of cultured cells for immunoprecipitation. *Cold Spring Harbor protocols*, **2010**;2010:pdb.prot5466.
- [31] Bettega D, Calzolari P, Doglia SM, Dulio B, Tallone L, Villa AM. Cell thickness measurements by confocal fluorescence microscopy on C3H10T1/2 and V79 cells. *Int J Radiat Biol*, **1998**;74:397-403.
- [32] Han Y, Liu XM, Liu H, Li SC, Wu BC, Ye LL, et al. Cultivation of recombinant Chinese hamster ovary cells grown as suspended aggregates in stirred vessels. *J Biosci Bioeng*, **2006**;102:430-435.

- [33] Bakirci H, Koner AL, Nau WM. Binding of inorganic cations by p-sulfonatocalix[4]arene monitored through competitive fluorophore displacement in aqueous solution. *Chem Commun*, **2005**:5411-5413.
- [34] Ham RG. An Improved Nutrient Solution for Diploid Chinese Hamster and Human Cell Lines. *Exp Cell Res*, **1963**;29:515-&.
- [35] Nau WM, Ghale G, Hennig A, Bakirci H, Bailey DM. Substrate-Selective Supramolecular Tandem Assays: Monitoring Enzyme Inhibition of Arginase and Diamine Oxidase by Fluorescent Dye Displacement from Calixarene and Cucurbituril Macrocycles. *J Am Chem Soc*, **2009**;131:11558-11570.
- [36] Ghale G, Lanctot AG, Kreissl HT, Jacob MH, Weingart H, Winterhalter M, et al. Chemosensing Ensembles for Monitoring Biomembrane Transport in Real Time. *Angew Chem Int Ed*, **2014**;53:2762-2765.
- [37] Lalor R, Baillie-Johnson H, Redshaw C, Matthews SE, Mueller A. Cellular uptake of a fluorescent calix[4]arene derivative. *J Am Chem Soc*, **2008**;130:2892-2893.
- [38] Rembish SJ, Trush MA. Further Evidence That Lucigenin-Derived Chemiluminescence Monitors Mitochondrial Superoxide Generation in Rat Alveolar Macrophages. *Free Radical Bio Med*, **1994**;17:117-126.
- [39] Muller SA, Holzapfel K, Seidl C, Treiber U, Krause BJ, Senekowitsch-Schmidtke R. Characterization of choline uptake in prostate cancer cells following bicalutamide and docetaxel treatment. *Eur J Nucl Med Mol I*, **2009**;36:1434-1442.
- [40] Lockman PR, Allen DD. The transport of choline. *Drug Dev Ind Pharm*, **2002**;28:749-771.

- [41] Schloss P, Mayser W, Niehuis A, Betz H. Na<sup>+</sup>-Dependent High-Affinity Uptake of Choline into Cultured Fibroblasts. *Biochem Biophys Res Commun*, **1994**;199:1320-1325.
- [42] Lips KS, Volk C, Schmitt BM, Pfeil U, Arndt P, Miska D, et al. Polyspecific cation transporters mediate luminal release of acetylcholine from bronchial epithelium. *Am J Respir Cell Mol Biol*, **2005**;33:79-88.
- [43] Wessler I, Kirkpatrick CJ. Acetylcholine beyond neurons: the non-neuronal cholinergic system in humans. *Brit J Pharmacol*, **2008**;154:1558-1571.
- [44] Amos H. Protamine Enhancement of Rna Uptake by Cultured Chick Cells. *Biochem Biophys Res Commun*, **1961**;5:1-&.
- [45] Amos H, Kearns KE. Influence of Bacterial Ribonucleic Acid on Animal Cells in Culture .2. Protamine Enhancement of Rna Intake. *Exp Cell Res*, **1963**;32:14-&.
- [46] Nagai J, Komeda T, Katagiri Y, Yumoto R, Takano M. Characterization of Protamine Uptake by Opossum Kidney Epithelial Cells. *Biological & Pharmaceutical Bulletin*, **2013**;36:1942-1949.
- [47] Guy J, Drabek D, Antoniou M. Delivery of DNA into Mammalian-Cells by Receptor-Mediated Endocytosis and Gene-Therapy. *Mol Biotechnol*, **1995**;3:237-248.
- [48] Cho YW, Kim JD, Park K. Polycation gene delivery systems: escape from endosomes to cytosol. *Journal of Pharmacy and Pharmacology*, **2003**;55:721-734.
- [49] Shorthill BJ, Avetta CT, Glass TE. Shape-selective sensing of lipids in aqueous solution by a designed fluorescent molecular tube. *J Am Chem Soc*, **2004**;126:12732-12733.
- [50] Petronini PG, Deangelis E, Borghetti AF, Wheeler KP. Osmotically Inducible Uptake of Betaine Via Amino-Acid-Transport System-a in Sv-3t3 Cells. *Biochem J*, **1994**;300:45-50.

[51] Redshaw C, Elsegood MRJ, Wright JA, Baillie-Johnson H, Yamato T, De Giovanni S, et al. Cellular uptake of a fluorescent vanadyl sulfonylcalix[4]arene. *Chem Commun*, **2012**;48:1129-1131.

[52] Mueller A, Lalor R, Cardaba CM, Matthews SE. Stable and Sensitive Probes for Lysosomes: Cell-Penetrating Fluorescent Calix[4]arenes Accumulate in Acidic Vesicles. *Cytom Part A*, **2011**;79A:126-136.

# Chapter 4

## Outlook

---

I am planning to build up my future scientific research projects based on the (bio)chemistry knowledge I learned during my bachelor (in chemistry) and master (in biochemistry) and the experiences I obtained during my master and PhD. For reaching such a goal, I am interested in design and performing biotechnological and biochemical projects dealing with chemical synthesis or peptide science and applying them in live cells for therapeutic or basic science purposes.

Cell-penetrating peptides can bridge the projects explained in chapters 2 and 3. These peptides can carry biomolecules with different molecular weights even the ones that are several times bigger than themselves. Peptide synthesis has been established very well over the past years both manually and machinery. I have some experience in manually synthesizing peptides. Cell-penetrating peptides are capable of solubilizing poorly soluble drugs and –as it is clear from their name– penetrating into cells. For solubilizing its bio-cargo the peptide should have some aromatic amino acids in its sequence. The peptide/cargo complex on the other hand has to remain water soluble. Therefore, the peptide needs to be designed according to structure of its cargo. Peptides fortunately are capable of fluorescent labeling. That is how their cellular uptake can be tracked by a fluorescent microscope. The effectiveness of the carried cargo in the presence or absence of its peptide vehicle represents the efficiency of the transfer. Peptide acting as a host is

reminiscent of transferring LCG by CX4 into live cells shown in Chapter 3. Unlike macrocycles –that have determined cavity sizes– peptides can vary in size from very few to tens amino acids. The aim is the peptide wrap around a cargo of interest and facilitates its uptake. Applying peptide as a macromolecule host, bypasses challenging synthesizing macrocyclic host for the desired cargo. Changing one amino acid in a peptide sequence can alter its adverse property such as toxicity, poor-solubility or inefficient cellular uptake. Peptides therefore make several opportunities for hosting.

I would like to make my try in chemical biology as well by modifying or fluorescent labeling biomolecules for using in live cells. I am quite familiar with synthesis. So far I have modified meso porous silica particles with octadecyl groups in my master and synthesized Fmoc-Dbo in a 12 step process. Such experiences may not be sufficient for challenging chemical synthesis but definitely is helpful when a few-steps chemical modification is needed or for understanding collaborative projects with an organic chemist for more challenging synthesis.

## Appendix A)

### 1- Conclusions and innovations

In Chapter 2, a counter intuitive result was observed for negatively charged peptides that is lack of charge effect in them: By an extensive study, it was clearly shown that when the charges on the peptides were turned on by changing the pH, the negatively charged peptides behave differently from the positively charged peptides. The flexibilities and end-to-end distances of the positively charged peptides were affected when the Coloumb charges turned-on. The negatively charged peptides on the other hand did not react to charge turn-on. This was explained by high affinity of carboxylate sidechain to water molecules results in negative charge screening. The charge repulsion in the negatively charged peptides was recovered by changing the solvent from water to methanol/water. The size of the sidechains, their proximity of the backbone and their interaction with N-terminus were also found important in boosting charged repulsion. The obtained results not only shed more light on the effect of charge repulsion on peptide folding but also can help to understand charge repulsion in homo repeat proteins and protein folding at early stage.

In Chapter 3, an indicator displacement assay (IDA) method based on calix[4]arene (CX4) macro cycle was performed inside live cells. In the IDA, biocationic analytes displaced with Lucigenin (LCG) indicator from CX4/LCG reporter pair. The freed LCG then showed a fluorescence signal. The introduced method can semi-quantitatively detect different bio cationic analyte from none analyte compound. The IDA however was established previously in cuvette but it was the first time we tried it in live cells. The obtained results are a proof-of-concept for using calixarenes and other macrocyclic receptors in biosystems for detecting low-molecular-weight analytes. Such analytes are difficult to sense by antibody based assays. But with the introduced IDA they can be sensed without being worried about bio competitors presents in biological samples. The calixarene receptor also facilitated the dye uptake into live cells. This result additionally calls for using macrocycles as carrier.

## **2- Own Contribution**

All results presented in this thesis were obtained solely by myself unless they were cited accordingly. The peptides sequences and their terminus modifications mentioned in Chapter 2 were designed by me or Prof. Nau. The project presented in chapter 3 was designed by myself. Generally during my PhD I made my best try to explain the obtained results and having an innovative idea toward designing complementary experiments for better understanding the obtained results rather than simply harvesting results. Trying the IDA inside live cells, independently trouble-shooting during the projects or charge repulsion recovery in negatively charged peptide by changing the solvent from water to methanol/water mixture are few examples of these ideas and innovative manner of doing science.

## Appendix B) Curriculum Vitae

### Amir Norouzy



**Postal address:** Research III, Jacobs University Bremen,  
Campus Ring 1, 28759 Bremen, Germany

**Tel:** +49-421-2003625

**Mobile phone:** +4917663383314

**E-mails:** [a.norouzy@jacobs-university.de](mailto:a.norouzy@jacobs-university.de)

[amir.norouzy@yahoo.com](mailto:amir.norouzy@yahoo.com)

---

### Education

-**2010-2014** Doctor of Philosophy (PhD) in biotechnology at Jacobs University Bremen, Germany.

-**2004-2007** Master of Science in medical biochemistry at Babol University of medical sciences (MUBabol), Babol, Iran. Total GPA: **17.0/20.0**.

-**1999-2003** Bachelor of Science in applied chemistry at Azad University Center of Tehran Branch (IAUCTB), college of basic sciences, Tehran, Iran. Total GPA: **16.01/20.0**

### Publications

[1] **Norouzy A**, Assaf K.I, Jacob M, Nau W. M. Coulomb repulsion in short polypeptides. *J Phys Chem B*, **2014** (DOI: 10.1021/jp508263a).

[2] **Norouzy A**, Azizi Z, Nau W. M. Indicator-displacement assays inside live cells. *Angew. Chem. Int. Ed*, **2014** (DOI: 10.1002/anie.201407808).

[3] **Norouzy A** and Nau, W. M. Synthetic macrocyclic receptors as tools in drug delivery and drug discovery. *European Drug Target Review*, **2014**. (<http://www.drugtargetreview.com/1850/z-homepage-promo/synthetic-macrocyclic-receptors/#.VlsmqDGUc1Y>)

- [4] Abdali N, Barth E, **Norouzy A**, Schulz R, Nau W. M, Kleinekathöfer U, *et al.* Corynebacterium jeikeium jk0268 constitutes for the 40 amino acid long PorACj, which forms a homooligomeric and anion-selective cell wall channel. *Plos one*, **2013**; 8: e75651.
- [5] Jacob M, Dsouza RN, Ghosh I, **Norouzy A**, Schwarzlose T, Nau W. M. Diffusion-enhanced Förster resonance energy transfer and the effects of external quenchers and the donor quantum yield. *J Phys Chem B*, **2013**;117:185-198.
- [6] Qujeq D, Roushan T, **Norouzy A**, Habibi-Rezaei M, Mehdinejad-Shani M. Effects of dichlorvos and carbaryl on the activity of free and immobilized acetylcholinesterase. *Toxicol Ind Health*, **2012**;28:291-295.
- [7] **Norouzy A**, Habibi-Rezaei M, Qujeq D, Vatani M, Badiei A. Adsorptive rmmobilization of acetylcholine esterase on ectadecyl substituted porous silica: optical bio-analysis of carbaryl. *B Korean Chem Soc*, **2010**;31:157-161.
- [8] **Norouzy A**, Qujeq D, Habibi-Rezaei M. The inhibitory effect of dissolved carbaryl in dioxane on physically adsorbed acetylcholinesterase. *React Kinet Catal L* **2009**;98:391-401.

#### **Accepted Manuscript:**

- [9] Parente Carvalho C, **Norouzy A**, Ribeiro V, Nau W. M, Pischel U. Cucurbiturils as Supramolecular Inhibitors of DNA restriction by type II endonucleases. *Org Biomol Chem*; **2014**

#### **Projects**

##### **2004 M.Sc. Thesis:**

- Preparation & characterization of structural and catalytical properties of immobilized acetylcholinesterase. (*Department of biochemistry and biophysics, Babol University of medical sciences, Babol/Iran*)

##### **2009, Researcher:**

- Studying scorpion's toxin effect on metabolism and histology of Balb-C mice. (*Pasteur Institute of Iran*)

##### **2010, PhD Thesis (Two Main Projects):**

- 1- Indicator Displacement Assays in Live Cells.

2- Polypeptide Dynamics and Structure Studies Simultaneously by Collision-induced Fluorescence Quenching and Resonance Energy Transfer in the 10-Å Domain.

### **Communications:**

- **A. Norouzy**, Z. Azizivarzaneh, W. M. Nau “Host–Guest Reporter Pair with ‘Turn-on’ Fluorescence Response to Monitor the Uptake of Cationic Analytes into Live Cells ” Molecular life sciences, 3<sup>rd</sup> to 6<sup>th</sup> October 2013 in Frankfurt, Germany
- **A. Norouzy** ”Host-guest Reporter Pairs with ‘Turn-on’ Fluorescence Response to Signal the Uptake of Cationic Analytes into Live Cells” Retreat Molife center, 25<sup>th</sup> and 26<sup>th</sup> of March 2013 in Bad Bevensen Germany.
- “Hybrid in Microscopy” **Workshop** from 30<sup>th</sup>-31<sup>st</sup> January 2013 at Jacobs University Bremen.
- **A. Norouzy**, I. Ghosh, M. H. Jacob and W. M. Nau, “Simultaneous short-distance FRET and collision-Induced Fluorescence Quenching to quantify how peptide conformation and dynamics depend on side-chain charge” Förster resonance energy transfer, 27<sup>th</sup> to 30<sup>th</sup> of March, 2011 Göttingen, Germany
- **A. Rei**, I. Ghosh, R. D'Souza, **A. Norouzy**, M. H. Jacob, W. M. Nau, 2011, “Short-distance FRET applied to the polypeptide chain” MAF 12<sup>th</sup> conference on methods and applications of fluorescence spectroscopy, Imaging and Probes, Strasbourz, France, September, **2011**
- **Norouzy, A.** Qujeq, D., 2008. “Effect of carbaryl and dioxan on physically adsorbed acetylcholinesterase kinetic”. The 8<sup>th</sup> Iran biophysical chemistry conference, University of Sistan-O Baluchestan, Zahedan, Iran.
- **Norouzy, A.** Qujeq, D,.... 2007. “Preparation and characterization of immobilized Acetylcholinesterase”. The second International Congress of Biochemistry and Molecular Biology. Oral presentation, Shiraz University of medical sciences, Shiraz, Iran.

- **Norouzy, A** , Theoretical and Experimental workshop of “immobilization of enzymes” Main director and lecturer, Babol University of Medical Sciences, Babol, Iran.
- **Norouzy, A**, Qujeq, D.2007. “A New Approach to Protein Analysis through Bound-Alkyl/Thioflavin T Competition procedure”. First Iranian proteomics congress, University of Tehran, Tehran, Iran.
- Qujeq, D. **Norouzy, A**,... 2005.”Determination Of Serum Iron, TIBC and Ferritin of patients with hepatitis B”. The First International Congress of Biochemistry and Molecular Biology. Oral presentation, Tarbiat Modarres University, Tehran, Iran.

### **English proficiency:**

-1990-present Studying English language, last diploma is pre FCE-2 from Kalam English Institute, Tehran, Iran. Total GPA: 92/100.

-1998 ITC (Instructors Training Course) Diploma, From Sadegh English Institute, Signed by M.S Hamzepoor, PhD

### **Awards and Honors**

**2010-2013** DAAD fellowship: *Research Grants for Doctoral Candidates and Young Academics and Scientists*

**2008** Awarded for excellent teaching at Zabol University of medical sciences.

**2002, 2003** accepted among top-five students for three serial semesters in (IAUCTB), college of basic science.

**2004** 13<sup>th</sup> placed among thousands of volunteers in the national University entrance examination in the field of medical biochemistry.

**2004** 29<sup>th</sup> Placed among thousands of volunteers in the national University entrance examination in the field of biochemistry.

### **Technical Skills**

#### **1- Spectroscopic**

- Time-Resolved Fluorescence Spectroscopy
- Steady-stated Fluorescence Spectroscopy

- Fluorescence Studying of Proteins and Peptides
- UV-Visible Spectroscopy
- Circular Dichroism (CD) Spectroscopy for Studying Proteins structure
- Dynamic Light Scattering

## **2- Chemical Synthesis**

- Chemical Synthesis of Peptides (SPPS method)
- Organic Synthesis (DBO fluorophore was synthesized in 13 serially steps)
- Functionalizing of Inorganic Matrixes

## **3- Enzymes and Proteins**

- Enzyme Handling (both pure or extracted from animal tissues)
- Immobilization of Enzymes and Proteins
- Protein Purification by HPLC
- Protein-based Biomineralization

## **4- Laboratory animal handling & Rat dissecting**

## **5- Cellular**

- Cell Culture and other Routine Cell Handling Skills
- Cell Manipulation (Microinjection)
- *in vivo* Macromolecular Host/Dye Displacement Assay
- Fluorescence and CLSM (Confocal Laser Scanning Microscopy) Imaging

## **Teaching Experiences**

- 2000-2006:** Teaching chemistry at many institute and schools to high school students.
- 2002-2007:** Teaching English in many language institutes.
- 2005-2006:** Teaching biochemistry and biochemistry laboratory in Babol University of medical science, professor assistant.
- 2007-2008:** Faculty of science at Zabol University of medical science.

## Appendix C ) List of Publications

[1] Abdali N, Barth E, Norouzy A, Schulz R, Nau W. M, Kleinekathöfer U, Tauch A, Benz R. Corynebacterium jeikeium jk0268 constitutes for the 40 amino acid long PorACj, which forms a homooligomeric and anion-selective cell wall channel. *Plos One*, **2013**; 8: e75651.

[2] Jacob MH, Dsouza RN, Ghosh I, Norouzy A, Schwarzlose T, Nau W. M. Diffusion-enhanced Förster resonance energy transfer and the effects of external quenchers and the donor quantum yield. *J. Phys. Chem. B*, **2013**;117:185-198.

[3] Norouzy A, Azizi Z, Nau W. M. Indicator-displacement Assays inside Live Cells. *Angew. Chem. Int. Ed*, **2014**. (DOI: 10.1002/anie.201407808)

[4] Norouzy A, Assaf K, Zhang S, Jacob M, Nau W. M. Coulomb Repulsion in Short Polypeptides. *J. Phys. Chem. B*, **2014**. (DOI: 10.1021/jp508263a)

[5] Norouzy A and Nau W. M. Synthetic macrocyclic receptors as tools in drug delivery and drug discovery. *European Drug Target Review*, **2014**. (published online)

### Accepted Manuscript:

[6] Parente Carvalho C, Norouzy A, Ribeiro V, Nau W. M, Pischel U. Cucurbiturils as Supramolecular Inhibitors of DNA restriction by type II endonucleases. *Org Biomol Chem*; **2014**.



# *Corynebacterium jeikeium* jk0268 Constitutes for the 40 Amino Acid Long PorACj, Which Forms a Homooligomeric and Anion-Selective Cell Wall Channel

Narges Abdali<sup>1,9</sup>, Enrico Barth<sup>2,9</sup>, Amir Norouzy<sup>1</sup>, Robert Schulz<sup>1</sup>, Werner M. Nau<sup>1</sup>, Ulrich Kleinekathöfer<sup>1</sup>, Andreas Tauch<sup>3</sup>, Roland Benz<sup>1,2\*</sup>

**1** School of Engineering and Science, Jacobs University Bremen, Bremen, Germany, **2** Rudolf Virchow Center, DFG-Research Center for Experimental Biomedicine, University of Würzburg, Würzburg, Germany, **3** Institute for Genome Research and Systems Biology Center for Biotechnology (CeBiTec), Bielefeld University, Bielefeld, Germany

## Abstract

*Corynebacterium jeikeium*, a resident of human skin, is often associated with multidrug resistant nosocomial infections in immunodepressed patients. *C. jeikeium* K411 belongs to mycolic acid-containing actinomycetes, the mycolata and contains a channel-forming protein as judged from reconstitution experiments with artificial lipid bilayer experiments. The channel-forming protein was present in detergent treated cell walls and in extracts of whole cells using organic solvents. A gene coding for a 40 amino acid long polypeptide possibly responsible for the pore-forming activity was identified in the known genome of *C. jeikeium* by its similar chromosomal localization to known *porH* and *porA* genes of other *Corynebacterium* strains. The gene *jk0268* was expressed in a porin deficient *Corynebacterium glutamicum* strain. For purification temporarily histidine-tailed or with a GST-tag at the N-terminus, the homogeneous protein caused channel-forming activity with an average conductance of 1.25 nS in 1M KCl identical to the channels formed by the detergent extracts. Zero-current membrane potential measurements of the voltage dependent channel implied selectivity for anions. This preference is according to single-channel analysis caused by some excess of cationic charges located in the channel lumen formed by oligomeric alpha-helical wheels. The channel has a suggested diameter of 1.4 nm as judged from the permeability of different sized hydrated anions using the Renkin correction factor. Surprisingly, the genome of *C. jeikeium* contained only one gene coding for a cell wall channel of the PorA/PorH type found in other *Corynebacterium* species. The possible evolutionary relationship between the heterooligomeric channels formed by certain *Corynebacterium* strains and the homooligomeric pore of *C. jeikeium* is discussed.

**Citation:** Abdali N, Barth E, Norouzy A, Schulz R, Nau WM, et al. (2013) *Corynebacterium jeikeium* jk0268 Constitutes for the 40 Amino Acid Long PorACj, Which Forms a Homooligomeric and Anion-Selective Cell Wall Channel. PLoS ONE 8(10): e75651. doi:10.1371/journal.pone.0075651

**Editor:** Eric Cascales, Centre National de la Recherche Scientifique, Aix-Marseille Université, France

**Received:** February 25, 2013; **Accepted:** August 16, 2013; **Published:** October 8, 2013

**Copyright:** © 2013 Abdali et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** Funds by Jacobs University. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: roland.benz@uni-wuerzburg.de

<sup>9</sup> These authors contributed equally to this work.

## Introduction

Members of the genus *Corynebacterium* are of considerable interest because some are potent producers of glutamate, lysine and other amino acids through fermentation processes on an industrial scale. Prominent examples of amino acid producers are *Corynebacterium glutamicum* or *Corynebacterium callunae* [1–6]. These bacteria belong to the family Corynebacteriaceae that is a distinctive suprageneric actinomycete taxon, the mycolata, which also includes mycobacteria, nocardiae, rhodococci and closely related genera. The mycolata share with the genus *Corynebacterium* the property of having an unusual cell envelope composition and architecture [7]. The mycolata have a thick peptidoglycan layer, covered by lipids in form of mycolic acids and other lipids [8–10]. The mycolic acids are covalently linked to the arabinogalactan, which is in turn attached to the murein of the cell wall [11]. The chain length of these 2-branched, 3-hydroxylated fatty acids varies considerably within the mycolata. Long mycolic acids have been found in Mycobacteria (60–90 carbon atoms), but they are short in

Corynebacteria (22–38 carbon atoms) [12–16]. This means that the cell wall of the mycolata forms a permeability barrier and probably has the same function as the outer membrane of gram-negative bacteria. These membranes contain channel-forming proteins for the passage of hydrophilic solutes [17–19]. Similarly, channels are present in the mycolic acid layer of the mycobacterial cell wall and the cell walls of a variety of Corynebacteria, such as *C. glutamicum*, *Corynebacterium efficiens*, *C. callunae*, and *Corynebacterium diphtheriae* [20–25]. In all these cases it seems conceivable that PorA and PorH form heterooligomers responsible for the cation-selective major cell wall channel besides a smaller anion-selective channel [26,27]. Cell wall channels define the mycolic acid layer as a permeability barrier on the surface of the Corynebacteria similar as has been found in recent years by the investigation of cell wall channels in different members of the mycolata [24,28–32].

The genus *Corynebacterium* contains on the other hand only a few pathogens. The main pathogen is *C. diphtheriae* [33], well known as the cause of diphtheria which is an acute, communicable respiratory disease. Other possible pathogens are only *Corynebac-*

*terium urealyticum* and *Corynebacterium jeikeium* [34,35]. *C. jeikeium* is part of the normal microflora of the human skin. It is a lipid-requiring pathogen that is associated with severe nosocomial infections recognized first in 1970 by Johnson and Kaye [36]. *C. jeikeium* is a strictly aerobic gram-positive rod that causes bioprosthetic valve endocarditis with a high mortality rate (33%) [37,38]. The bacterium may be multidrug-resistant and needs vancomycin for its treatment. Today the knowledge on the complete genome sequence of *C. jeikeium* K411, a clinical isolate originally recovered from the axilla of a bone marrow transplant patient, provides the basis for an in-depth understanding of the physiology of this medically important bacterium [39]. The chromosome of *C. jeikeium* K411 has a size of 2.46 Mbp and comprises 2104 predicted coding regions, of which 68 most likely represent pseudogenes. The chromosomal architecture of *C. jeikeium* K411 revealed a moderate number of genomic rearrangements when compared to other sequenced corynebacterial genomes [39]. These structural differences of the chromosome have been attributed very recently to the phylogenetic position of *C. jeikeium* within the taxonomic tree of the genus *Corynebacterium* [40]. Annotation of the genomic data revealed that the lipophilic phenotype of *C. jeikeium* is caused by the absence of a gene coding for a fatty acid synthase and linked to pathogenicity, and that events of horizontal gene transfer are responsible for multidrug resistance [39]. The annotated genome sequence can be regarded as starting point for comprehensive post-genomic studies at the transcriptomic and proteomic levels [41,42], but also for the detailed functional analysis of predicted coding regions, for instance the putative porin gene locus of *C. jeikeium* K411.

In this study, we extended the search for cell wall channels to the *C. jeikeium* strain K411 that is a clinical isolate with a known genome [39]. Using lipid bilayer experiments we could demonstrate that the extracts of whole *C. jeikeium* cells contain a protein that forms wide and water-filled channels similar to the porins found in gram-negative bacteria and in other Corynebacteria [17,20,43]. The channel-forming protein, named PorAC<sub>j</sub>, was identified within the accessible genome of *C. jeikeium* K411 [39] by using its homology to PorA of *C. glutamicum*. PorAC<sub>j</sub> was expressed in a PorA/PorH-deficient strain of *C. glutamicum* ATCC13032 [20,44] and purified to homogeneity. The protein is active as a homooligomer in contrast to PorA/PorH of most Corynebacteria, which form heterooligomeric channels [27]. We present in this

study the characterization of the first homooligomeric channel-forming protein of the PorA type of a strain within the genus *Corynebacterium*, which is formed in contrast to other cell wall channels from the mycolata by alpha-helical stretches. A phylogenetic tree suggests that PorAC<sub>j</sub> could be the ancestor of all known PorA/H proteins from *Corynebacterium* strains.

## Experimental Procedures

### Bacterial Strains and Growth Conditions

The *Corynebacterium* strains *C. glutamicum* ATCC13032 and *C. jeikeium* K411 (obtained from the National Collection of Type Cultures, NCTC, London, UK) were grown in 1000 ml baffled Erlenmeyer flasks containing 250 ml of brain-heart infusion (BHI) media (Becton) and 250 ml Erlenmeyer flasks containing 25 ml BYT medium [45]. Former cultures were stirred on a rotary shaker at 140 rpm and 30°C, latter at 280 rpm and 37°C. *Escherichia coli* NEB5 $\alpha$  (New England Biolabs), used for cloning, was grown under standard conditions in Luria broth (LB). If required agar plates and liquid media were supplemented with 20 and 40  $\mu$ g/ml chloramphenicol, respectively.

### Cloning of *jk0268*

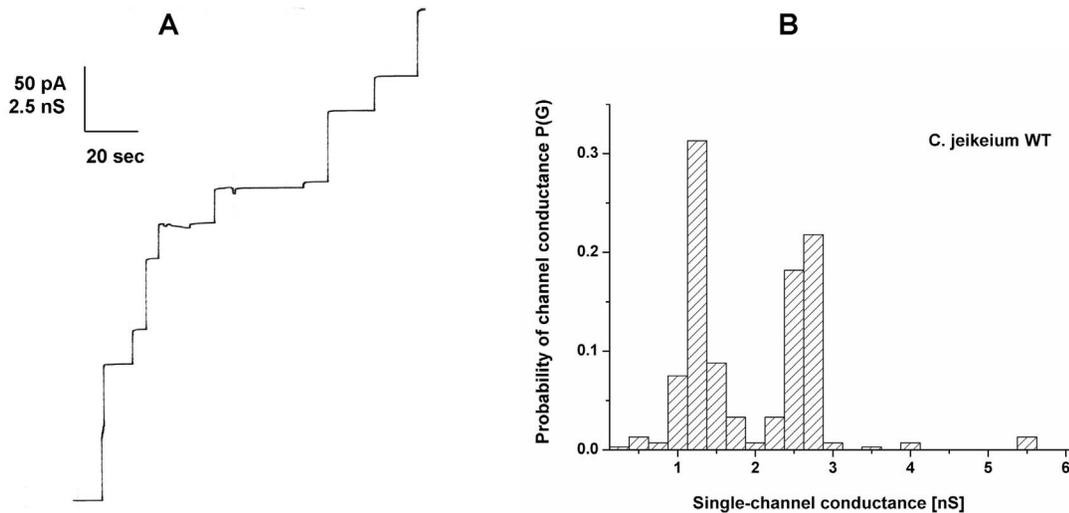
The gene *jk0268* and its putative ribosome binding site was PCR amplified from genomic *C. jeikeium* DNA. It was isolated in a 50  $\mu$ l reaction volume containing 1  $\times$  Taq buffer, 0.2 mM dNTPs, 3 mM MgCl<sub>2</sub>, 1 U Taq DNA polymerase (Fermentas) and 0.4  $\mu$ M primers FP JK0268XbaI and RP JK0268EcoRI (Table 1). Used PCR conditions were: initial denaturing at 95°C for 5 minutes, 30 cycles at 95°C for 1 min, 45°C for 1 minutes, 72°C for 1 min and a final extension at 72°C for 10 minutes. A PCR product of ~ 200 bp was cut out after agarose gel size separation, ligated into TOPO2.1 vector (Invitrogen) and heat-shock transformed into *E. coli* NEB5 $\alpha$  competent cells according to the instructions of the manufacturers. One plasmid containing the amplification was *EcoRI* and *XbaI* (Fermentas) digested and the 200 bp fragment was ligated in the backbone of *EcoRI* and *XbaI* linearized vector pXMJ19 [46] eventually resulting in the expression plasmid pXJK0268.

**Table 1.** Oligonucleotides used in this study.

Oligonucleotides	Sequence 5'→3'	Position
FP JK0268XbaI	GGAACCTGGCGCTCTAGATCTCTTAAGAGGA	329071–329101
RP JK0268EcoRI	GAAGCCGGGGTTTGAATCTTAAGCGGAAGC	329232–329262
RP JK0268KpnI	TAAGCGGAGGTACCCTTAGCAGCGGTCCACTTAACG	–
FP KpnIXa8HisEcoRI	CATCGAGGGCCGCGCGGCCACCACCACCACCACCACCAATAGG	–
RP KpnIXa8HisEcoRI	AATTCCTATTAGTGGTGGTGGTGGTGGTGGTGGCCCGCGCCCTCGATGGTAC	–
FP pXMJ19Seq	GTGAGCGGATAACAATTCAC	–
RP pXMJ19Insert	CTCTCATCCGCAAACAGC	–
Fwd GST-jk Seq	CAC TCC CGT TCT GGA TAA TG	–
Rev GST-jk Seq	CAC TCC GCT ATC GCT ACG TGA C	–
Fwd R927G	CTGGTCCGGGTGGATCCC	–
Rev R927G	GGGATCCACCGGAACCCAG	–

The sequences of the primers were derived from the prospective gene *jk0268* of the cell wall channel and its flanking regions taken from the genome of *C. jeikeium* K411 [39]. Primer binding positions in the chromosome of the accessible genome of *C. jeikeium* K411 (reference sequence NC\_007164) are provided.

doi:10.1371/journal.pone.0075651.t001

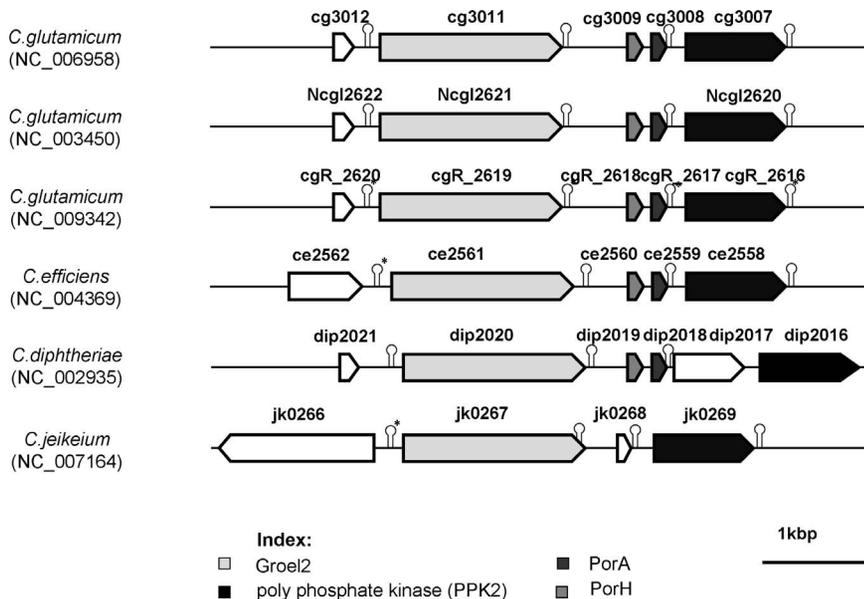


**Figure 1. Study of pore-forming capacity of detergent extracts of *C. jeikeium*.** (A) Single-channel recording of a PC/*n*-decane membrane in the presence of the detergent extract of whole *C. jeikeium* K411 cells. The aqueous phase contained 1 M KCl, pH 6 and 50 ng/ml protein extract. The applied membrane potential was 20 mV;  $T = 20^{\circ}\text{C}$ . (B) Histogram of the probability  $P(G)$  for the occurrence of a given conductivity unit observed with membranes formed of 1% PC dissolved in *n*-decane. It was calculated by dividing the number of fluctuations with a given conductance rise by the total number of conductance fluctuations in the presence of detergent extracts of whole *C. jeikeium* K411 cells. Two frequent conductive units were observed for 307 single events taken from 13 individual membranes. The average conductance of the steps corresponding to the left-side maximum was 1.25 nS and that of the right-side maximum was 2.5 nS. The aqueous phase contained 1 M KCl, pH 6 and 50 ng/ml protein extract, the applied membrane potential was 20 mV,  $T = 20^{\circ}\text{C}$ . doi:10.1371/journal.pone.0075651.g001

### Construction of a C-terminal His<sub>8</sub>-tag

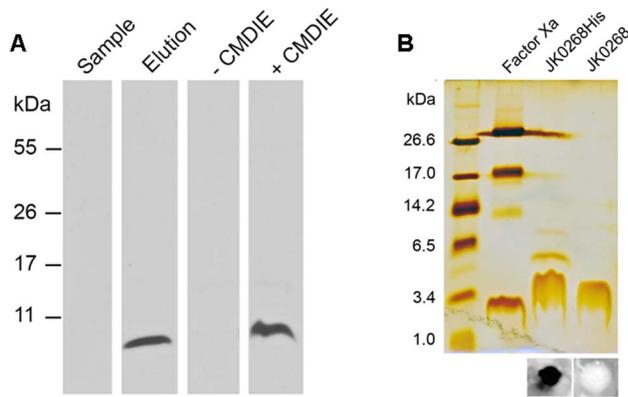
For immobilized metal ion affinity purification (IMAC) the vector pXMJ19 was upgraded by introduction of a DNA cassette coding a C-terminal factor Xa (I-E-G-R) linked octa-histidine tag. Therefore, single-stranded oligonucleotides FP KpnIXa8HisEcoRI and RP KpnIXa8HisEcoRI were first 5' phosphorylated, then

annealed by a temperature gradient step to provide double-stranded DNA with *Kpn*I and *Eco*RI overhangs and finally ligated in *Kpn*I and *Eco*RI cut pXMJ19 vector (T4 DNA ligase, Fermentas). The resulted plasmid was designated pXHis. To apply the C-terminal tag of plasmid pXHis to *jk0268* the native stop codon was mutated by PCR amplification. Using PCR



**Figure 2. Analysis of the accessible genomes from *C. glutamicum*, *C. efficiens*, *C. diphtheriae* and *C. jeikeium*.** The homologous genes of the chaperonin GroEL2 and a poly phosphate kinase PPK2 enclose a presumed conserved porin domain. The operon covering the genes *CgporH* and *CgporA* whose proteins build the bicomponental main cell wall channel of *C. glutamicum* is inferred to exist in all strains except for *C. jeikeium*. Possible terminator sequences of mRNA transcripts were predicted with TranstermHP (indicated by hairpins [62]; or were identified manually (marked by asterisk). doi:10.1371/journal.pone.0075651.g002





**Figure 4. Analysis of PorACj purification.** (A) Western blot analysis illustrating IMAC purification of his-tagged PorACj protein. The protein was expressed in *C. glutamicum* ATCC13032  $\Delta$ porH $\Delta$ porA and purified by Ni<sup>2+</sup> affinity from the supernatant of detergent extracted whole cells. CMDIE represents chloroform-methanol treated cells in which the crude protein content was concentrated around 8 fold by diethyl-ether precipitation of pJK0268His transfected (+) or non-transfected (–) *C. glutamicum*  $\Delta$ porH $\Delta$ porA cells. Subsequent to tricine (12%) SDS-PAGE the gel was blotted on a nitrocellulose membrane and PorACj-His was visualized by Anti-His antibodies and a chemiluminescent reaction. All samples were boiled for 5 minutes in Redmix before loading. (B) Silver stained tricine (16.5%) SDS-PAGE of Ni<sup>2+</sup>-purified and factor Xa digested PorACj-His protein. Lanes: 1, 3 units of protease Xa (control); 2, 10  $\mu$ l of three pooled Ni-NTA elution containing PorACj-His; 3, 10  $\mu$ l of protease Xa treated and purified PorACj protein (for details see text). The dot blot immunosays pictures underneath lanes 2 and 3 show cleavage of the histidine tail using anti-his antibody of 5  $\mu$ l of the corresponding protein samples. Before loading all samples were boiled for 5 minutes in Redmix.

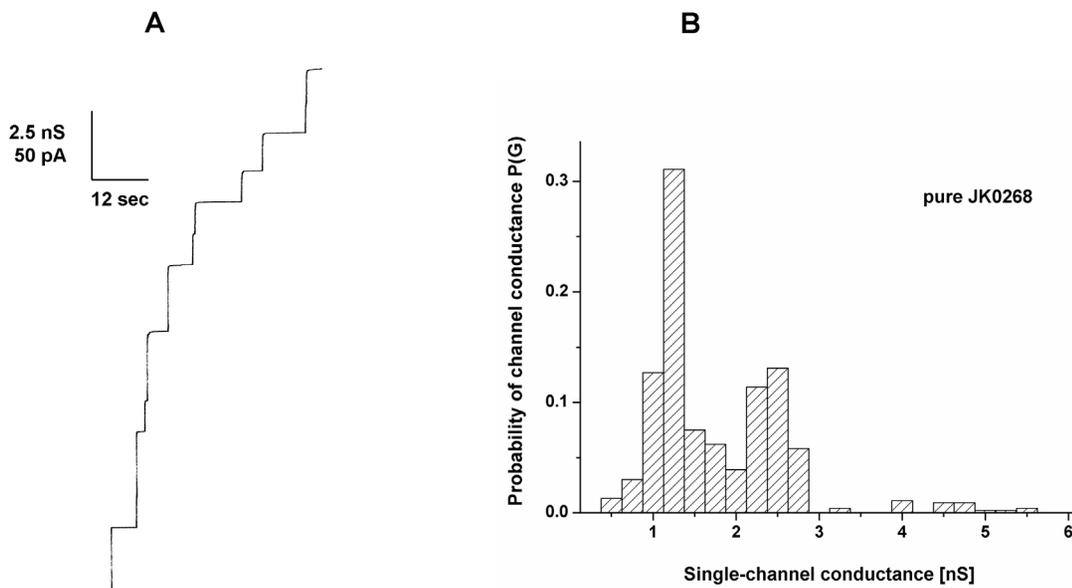
doi:10.1371/journal.pone.0075651.g004

**Table 3. Average single-channel conductance, G, of purified PorACj in different salt solutions.**

Salt	Concentration [M]	Single-channel conductance G [pS]	
		left-side maximum	right-side maximum
KCl	3.0	3,250 $\pm$ 230	6,500 $\pm$ 410
	1.0	1,250 $\pm$ 115	2,500 $\pm$ 150
	0.3	350 $\pm$ 25	750 $\pm$ 55
	0.1	175 $\pm$ 15	325 $\pm$ 25
	0.03	50 $\pm$ 4	120 $\pm$ 8
0.01	33 $\pm$ 2	70 $\pm$ 5	
LiCl	1.0	1,000 $\pm$ 85	2,000 $\pm$ 140
KCH <sub>3</sub> COO	1.0	500 $\pm$ 33	1,100 $\pm$ 70
pH 7	0.1	50 $\pm$ 6	115 $\pm$ 7
KBr	0.1	170 $\pm$ 12	340 $\pm$ 19
KNO <sub>3</sub>	0.1	140 $\pm$ 11	260 $\pm$ 16
KClO <sub>3</sub>	0.1	115 $\pm$ 9	235 $\pm$ 22
KF	0.1	70 $\pm$ 5	180 $\pm$ 12
KCHOO	0.1	65 $\pm$ 4	155 $\pm$ 10

The membranes were formed of 1% PC/*n*-decane. The aqueous solutions were unbuffered and had a pH of about 6 if not indicated otherwise. The applied voltage was 20 mV and the temperature 20°C. The single values represent the means ( $\pm$  SD) of at least 100 single-channel events derived from at least four individual membranes.

doi:10.1371/journal.pone.0075651.t003



**Figure 5. Study of pore-forming capacity of purified PorACj.** (A) Single-channel recording of a PC/*n*-decane membrane in the presence of pure PorACj. The aqueous phase contained 1 M KCl, pH 6 and 10 ng/ml protein. The applied membrane potential was 20 mV; T = 20°C. (B) Histogram of the probability P(G) for the occurrence of a given conductivity unit observed with membranes formed of 1% PC dissolved in *n*-decane. It was calculated by dividing the number of fluctuations with a given conductance rise by the total number of conductance fluctuations in the presence of pure PorACj. Two frequent conductive units were observed for 295 single events taken from eight individual membranes. The average conductance of the steps corresponding to the left-side maximum was 1.25 nS and that of the right-side maximum was 2.5 nS. The aqueous phase contained 1 M KCl, pH 6 and 10 ng/ml protein, the applied membrane potential was 20 mV, T = 20°C.

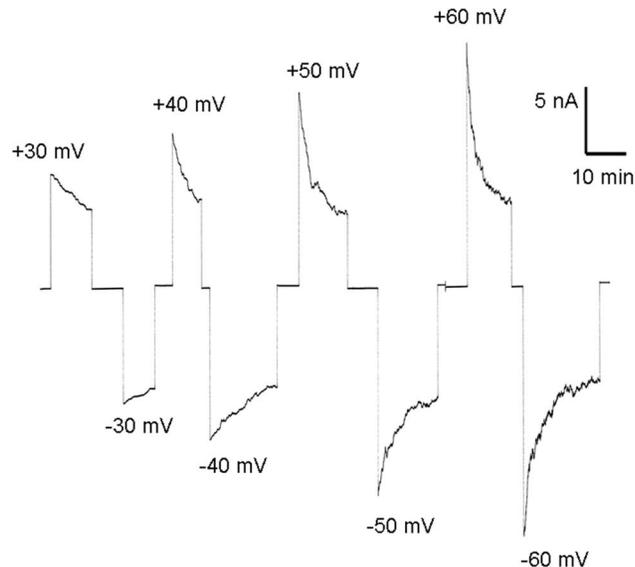
doi:10.1371/journal.pone.0075651.g005

analysis was done via double strand DNA sequencing into pCR2.1 vector.(Eurofins MWG Operon, Germany) One pCR2.1-jk0268 plasmid containing the amplified modified DNA fragment was EcoRI and BamHI (NEB) digested and the 200 bp fragment was ligated in the backbone of EcoRI and BamHI linearised vector eventually resulting in the expression plasmid pGEX-2TJK0268. Subsequently, site-directed mutagenesis was utilized in pGEX-2T expression vector prior to usage, to convert R927 of the plasmid to G in the thrombin cleavage sequence (LVPR|GS CTG GTT CCG CGT GGATCC) to avoid any other cleavage site for Factor Xa protease.

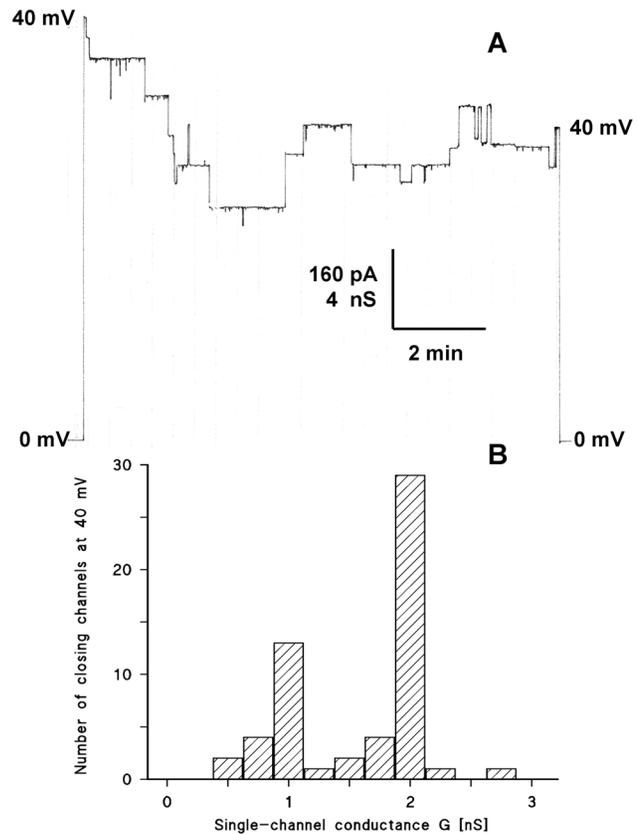
The sequence of all expression vectors were verified by sequencing (Seqlab, Göttingen Germany) prior to transformation of the plasmids into competent porin deficient *C. glutamicum* ATCC 13032 ΔAH or porin deficient *BL21 DE3 Omp8 E. coli* strains.

### Isolation of Cell Wall Proteins

Cell wall-associated proteins were isolated by methods described in details previously [28,30]. A liquid culture of grown cells was centrifuged (6,000 rpm, 15 minutes, 4°C in Heraeus Minifuge RF centrifuge). The cell pellet was washed twice with 10% culture volume (10 mM Tris, pH8) before cell wall proteins were extracted either by shaking the cells in detergent solution or in a 1:2 (v/v) mixture of the organic solvents chloroform and methanol. For both extraction methods one part cells (0.3 g wet weight bacterial pellet) was resuspended in five parts detergent solution (1.5 ml 1% LDAO (lauryldimethylamine-oxide), 10 mM Tris, pH8) or organic solvent (1.5 ml chloroform/methanol). After 3 hours agitation at RT cells were sedimented in a table top centrifuge (10 minutes, 4°C, 10,000 rpm) and the pellet was discarded. The detergent supernatant was immediately applied to IMAC purification. The chloroform-methanol mixture had first to be precipitated with 9 times the volume of ice-cold diethyl ether (16 h, -20°C) before the obtained pellet was either resolved in



**Figure 6. Investigation of the voltage-dependence of PorACj in a multi-channel experiment.** The purified protein was added to the *cis*-side of a PC membrane (100 ng/ml) and the reconstitution of channels was followed until equilibrium. Then increasing positive (upper traces) and negative voltages (lower traces) were applied to the *cis*-side of the membrane, and the membrane current was measured as a function of time. The aqueous phase contained 1 M KCl; T=20°C. doi:10.1371/journal.pone.0075651.g006



**Figure 7. Investigation of the voltage-dependence of PorACj in single-channel experiments.** A: The purified protein was added to the *cis*-side of a PC membrane (10 ng/ml) and the reconstitution of channels was followed until about 10 PorACj-channels inserted into the membrane. Then 40 mV were applied to the *cis*-side of the membrane, and the membrane current was measured as a function of time. The aqueous phase contained 1M KCl; T=20°C. B: Histogram of 56 closing events of the experiment in A and similar experiments. The closing events were plotted in a bargraph as a function of the conductance of the closing events. 1 M KCl; T=20°C. Note that the PorACj channels closed in two distinct conductance values of 1 and 2 nS. doi:10.1371/journal.pone.0075651.g007

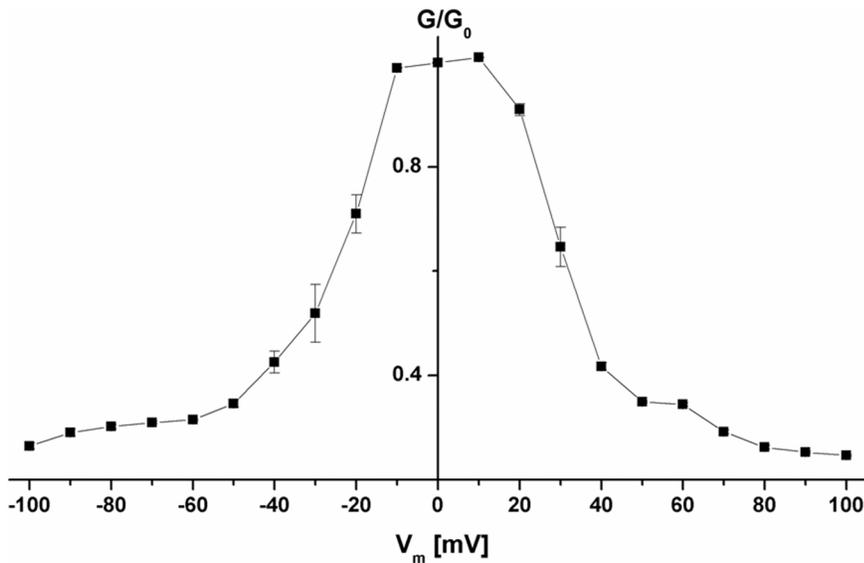
detergent solution (1% LDAO, 10 mM Tris, pH8) or in loading buffer for gel electrophoresis [28].

### IMAC Purification

Histidine-tailed *C. jeikeium* protein was purified to homogeneity by utilization of immobilized metal ion affinity chromatography (IMAC). From detergent treated cells 5 ml of the 1% LDAO supernatant were loaded on Ni-NTA spin columns (Qiagen) equilibrated with buffer 1 (20 mM Tris, 50 mM NaCl, 0.4% LDAO, pH8). After ten washing steps using each 650 μl of buffer 2 (= buffer1 with 10 mM imidazol) bound protein was eluted from the column with 200 μl buffer 3 (=buffer 1 with 300 mM imidazol).

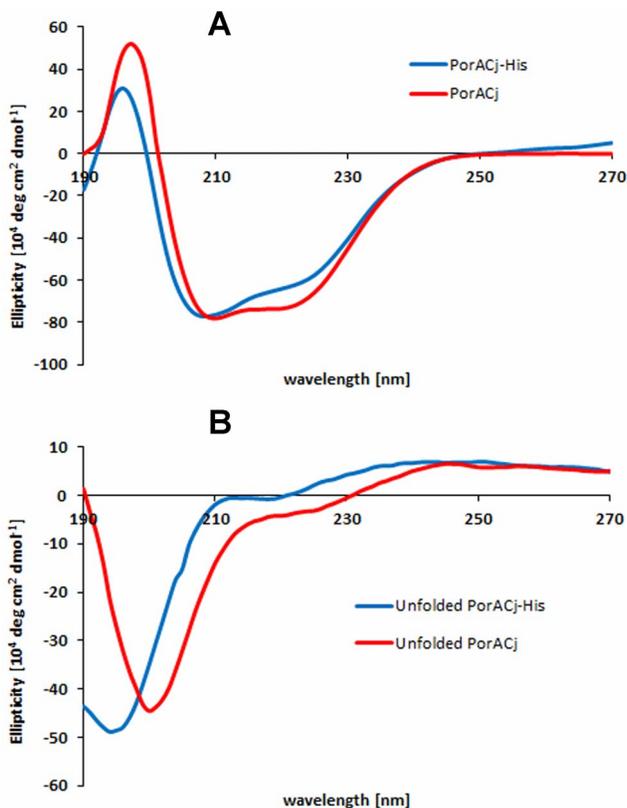
### Expression of Recombinant GST-jk0268

The plasmid including desired gene was sequenced afterward and transformed into the porin deficient *BL21 DE3 Omp8 E. coli*. Cells were grown at 37°C in LB medium and induced by 1 mM IPTG. The culture media incubated over night at 26°C after induction. Subsequently, cells were harvested by centrifugation at 4,000 ×g for 20 minutes at 4°C and were resuspended in PBS



**Figure 8. Conductance ( $G$ ) at a given membrane potential ( $V_m$ ) divided by the conductance at 10 mV ( $G_0$ ) expressed as a function of the membrane potential.** The symbols represent the mean ( $\pm$  SD) of six measurements, in which pure PorACj protein was added to the *cis*-side of the membranes. The aqueous phase contained 1 M KCl and 100 ng/ml porin. The membranes were formed from PC/*n*-decane at a temperature of 20°C.

doi:10.1371/journal.pone.0075651.g008



**Figure 9. Analysis of secondary structure of PorACj using CD-spectrometry.** A: CD spectra of recombinant PorACj (69  $\mu$ M) and PorA-His<sub>8</sub> (12  $\mu$ M) solubilized in 0.5% Genapol, 100 mM NaCl, 50 mM TrisHCl and 1 mM CaCl<sub>2</sub>, pH 8 measured at room temperature. B: CD-spectra of the same protein samples as in (A). The aqueous solutions of the proteins was supplemented with 4 M urea to destroy the secondary structure of the proteins.

doi:10.1371/journal.pone.0075651.g009

phosphate buffered saline (pH 7.4) then lysed with a high-pressure-homogenizer (2×1500 bar). Unbroken cells removed by centrifugation, supernatant used for purification.

#### GST-PorACj Fusion Protein Purification

Purification of GST-PorACj was performed using glutathione sepharose 4B medium, (following batch protocol GST Gene Fusion System Handbook, GE Healthcare). After 5 times washing with buffer A (0.5% Genapol, 100 mM NaCl, 50 mM Tris-HCl, 1 mM CaCl<sub>2</sub>, 2.5 mM DTT, pH 7.4) to remove non-bound sample components, the purified GST-fusion protein was eluted by addition of buffer A supplemented with 10 mM Glutathione, pH 8. The protein sample was concentrated using amicon ultra 3 kDa [Millipore] to a concentration of 3 mg/ml GST-PorACj determined using OD280. Uses 30  $\mu$ l of sample highly specific detect only GST fusion protein. Western blotting was carried out using anti-GST antibody. (Data not shown).

#### Protease Xa Cleavage of PorACj-His<sub>8</sub>

Subsequent to IMAC purification the sample contained high imidazol concentrations which strongly inhibit protease Xa (Qiagen) activity. Removal of imidazol was performed by dialysing the sample over night against cleavage buffer (20 mM Tris, 50 mM NaCl, 1 mM CaCl<sub>2</sub>, 0.4% LDAO, pH6.5) using a cellulose membrane with a MWCO of 2 kDa (Spectra/Por 6, Carl Roth, Germany). For cleavage of the poly-histidine-tag 4 units protease Xa (Qiagen) were added to the sample (37°C, over night). The enzyme was removed with the factor Xa removal Kit according Qiagen instructions while the cleaved JK0268 (PorACj) protein was separated from uncleaved PorACj-His<sub>8</sub> by a second passage through a Ni-NTA filter.

#### Protease Xa cleavage of PorACj-GST

The purified GST-PorACj contained 10 mM of glutathione. PorACj was dialysed for 36 h at 4°C with constant stirring in glutathione and DTT free buffer (100 mM NaCl, 50 mM Tris,

**Table 4.** Radii of the anions and relative permeability of PorACj from *C. jeikeium* in different salt solutions.

Anion	Radii of the hydrated anions $a$ [nm]	Limited molar conductivity $\lambda_i$ [mS/M]	Permeability relative to 0.1 M KBr	
			left-side maximum	right-side maximum
Br <sup>-</sup>	1.05	78.14	1.0	1.0
Cl <sup>-</sup>	1.07	76.35	1.0	0.96
NO <sub>3</sub> <sup>-</sup>	1.15	71.46	0.82	0.76
ClO <sub>3</sub> <sup>-</sup>	1.27	64.60	0.68	0.69
F <sup>-</sup>	1.48	55.40	0.41	0.53
HCOO <sup>-</sup>	1.50	54.59	0.38	0.46
CH <sub>3</sub> COO <sup>-</sup>	2.00	40.90	0.29	0.34

The data for the limiting conductivities of the different ions were taken from ref. [66]. The radii of the hydrated anions were calculated using the Stokes equation [67]. The single channel conductance of PorACj for the different salts at 0.1 M was taken from Table 3. The relative permeability of the single anions was calculated by dividing the single-channel conductance of the individual anion by that of 0.1 M KBr. The relative permeability for 0.1 M KBr was set to unity.  
doi:10.1371/journal.pone.0075651.t004

0.5% Genapol, 1 mM CaCl<sub>2</sub>, pH 8.0 at room temperature). Then 2 units FXa protease (NEB) was added for 16 h at 22°C to the protein solution. No DTT was added to the dialysis buffer, because of its absorption for CD measurement. The sample was loaded again onto glutathionsepharose 4B medium and the flow through was collected. A concentration of 0.65 mg/ml pure PorACj was determined using OD280.

### Protein Electrophoresis and Immunoblotting

Protein samples were size separated subsequent to a denaturation step (5 minutes, 95°C) with Tris-Tricine 12% or 16.5% polyacrylamide gels [48]. After electrophoresis gels were either stained with Coomassie Brilliant Blue G-250, or by silver stain [49] or electroblotted [50]. In the latter case, proteins were transferred

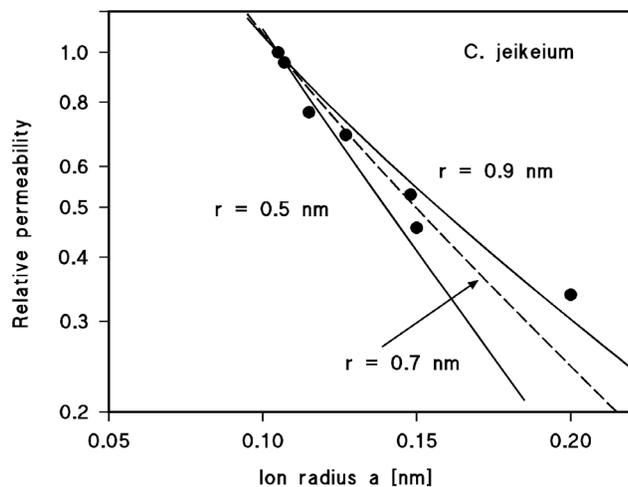
to a 0.1 μm nitrocellulose membrane (Protran, BA79, Whatman). The blotting was performed for 4–5 minutes in a wet tank blot system (Biorad) with Towbin buffer (25 mM Tris, 192 mM glycine, 20% methanol) at 350 mA current. Unspecific binding sites on the membrane were blocked with 5% skimmed milk in TBS-T buffer (20 mM Tris, 0.01 M NaCl, 0.1% Tween, pH7.5) before probing with the first 1:5000 diluted monoclonal mouse Anti-his antibody (Amersham Biosciences, UK). Subsequent to multiple TBS-T washing steps the second peroxidase-conjugated Anti-Mouse antibodies (DAKO, Denmark) were added at the same dilution. Attending to the manufacturer's instructions use of the ECL Western blotting detection system (GE Healthcare, UK) resulted in light emission recorded on autoradiography films (Hyperfilm<sup>TM</sup>MP, GE Healthcare, UK). Dot blot immunodetection was carried out identically without prior SDS-PAGE. Exposure times varied between 10 seconds to 5 minutes as required by the sample.

### Test for Susceptibility to Antibiotics

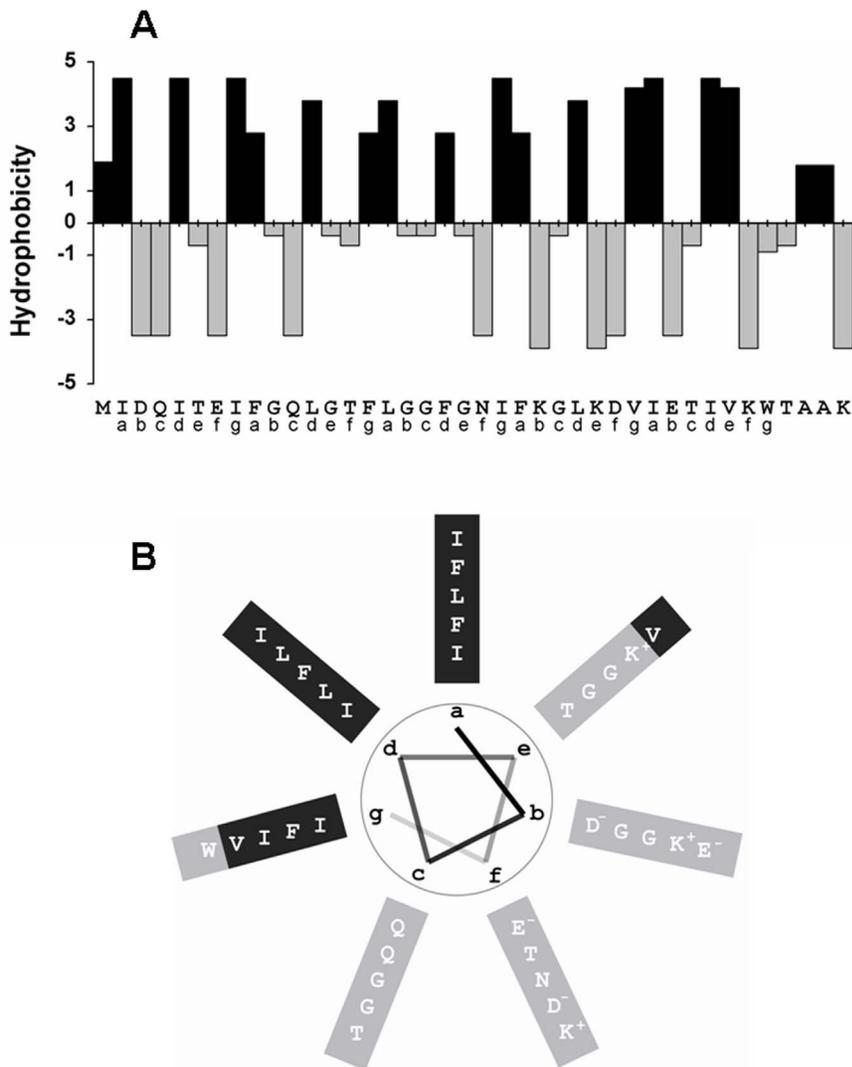
Sterilized Whatman filter disks with 5 mm diameter were used for qualitative tests. Overnight cultures in the suitable medium were diluted 1:1,000; then 1 ml of each culture containing approximately 10<sup>7</sup> cells/ml was spread onto BHI agar medium (Difco) containing 1 mM IPTG. Five-microliter portions of the diluted 1:1,000 stock solutions of the different antibiotics were deposited to each disk. The diameters of the growth inhibition zones were measured after 16 and 24 h. The concentrations of the antibiotics in the growth inhibition zone experiments were: ampicillin, 100 μg/ml; penicillin G, 100 μg/ml; carbenicillin, 100 μg/ml; ceftazidime 100 μg/ml; ertapenem, 20 μg/ml; imipenem, 10 μg/ml; gentamycin, 25 μg/ml; tetracycline 25 μg/ml; Besides the qualitative evaluation of the antibiotic susceptibility also the minimum concentration (MIC) of antimicrobial agents was measured, which inhibits the growth of the different *C. glutamicum* strains used in this study. The stock concentration of the antibiotics in these experiments were the same as above for the growth inhibition zones.

### CD Measurements

CD measurements were performed on a Jasco J-810 circular dichroism spectropolarimeter using 0.1 cm light pathway cells at room temperature while flushing the cuvette chamber with nitrogen gas. Spectra were recorded from 190–260 nm with a resolution of 1 nm and an acquisition time of 200 nm/minutes.



**Figure 10.** Fit of the single-channel conductance data of PorACj by using the Renkin correction factor times the aqueous diffusion coefficients of the different anions [67]. The product of both numbers was normalized to 1 for  $a = 1.05$  nm (Br<sup>-</sup>). Single-channel conductance was normalized to the one of 0.1 M KBr and plotted versus the hydrated ion radii taken from Table 3. The single-channel conductance correspond to Br<sup>-</sup>, Cl<sup>-</sup>, NO<sub>3</sub><sup>-</sup>, ClO<sub>3</sub><sup>-</sup>, F<sup>-</sup>, HCOO<sup>-</sup> and CH<sub>3</sub>COO<sup>-</sup> which were all used for the pore diameter estimation. The fit (solid lines) is shown for  $r = 0.5$  nm (lower line) and  $r = 0.9$  nm (upper line). The best fit was achieved with  $r = 0.7$  nm (diameter = 1.4 nm) corresponding to the broken line.  
doi:10.1371/journal.pone.0075651.g010



**Figure 11. Analysis of PorACj secondary structure.** (A) The panel shows the hydrophobicity indices of the individual amino acids of PorACj according to ref [80]. (B) The secondary structure of PorACj was predicted using a consensus method [83] at the Pole Bioinformatique Lyonnaise network ([http://npsa-pbil.ibcp.fr/cgi-bin/npsa\\_automat.pl?page=/NPSA/npsa\\_seconcs.html](http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_seconcs.html)); the protein was suggested to form  $\alpha$ -helices. Amino acid residues arranged on basis of heptameric repeats (a–g) showing distinct separation in a hydrophobic domain supposable surrounded by lipid molecules (dark grey) while the hydrophilic domain (light grey) is suggested to represent the component orientated to the water-filled lumen in the presumed oligomeric PorACj.

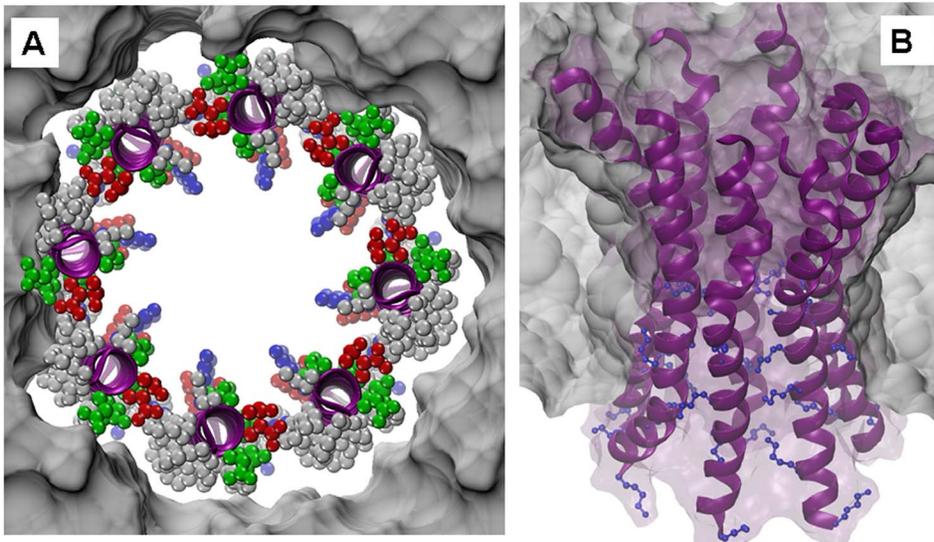
doi:10.1371/journal.pone.0075651.g011

Ten scans were taken for each sample, and the average of these scans was smoothed and stored for further analysis. Blanks of the respective protein-free sample were recorded under the same conditions and subtracted from the protein spectrum before further analysis. Analysis of the CD spectra was performed by expressing the spectra as linear combinations of 4 reference spectra (alpha-helix, beta-sheet, beta-turns and random-coil) as described elsewhere by using public domain programs [51].

### Black Lipid Bilayer Membranes

The methods used for the lipid bilayer experiments have been previously described in detail [52]. In the experimental setup two compartments of a Teflon cell filled with electrolyte solution are connected by a small circular hole with an area of about 0.2 mm<sup>2</sup>. The black lipid bilayer membrane was made by painting a 1% (w/v) diphytanoyl phosphatidylcholine (PC)/*n*-decane solution (Avanti

Polar Lipids, Alabaster, USA) across the hole. Ag/AgCl electrodes were used connected in series to a voltage source and a home-made current-to-voltage converter for the electrical measurements. The bandwidth of the instrument was set to 100–300 Hz dependent on the magnitude of the output signal. The amplified signal was monitored with a digital oscilloscope and recorded with a strip chart recorder. All salts were obtained from Merck (Germany) or Sigma-Aldrich (Germany) at analytical grade. The aqueous salt solutions were unbuffered and, if not explicit mentioned, had a pH of around 6. The temperature during all experiments was maintained at 20°C. The zero-current membrane potential measurements were performed as it has been described earlier [53] by establishing a 5-fold salt gradient across membranes containing 10–1000 cell wall channels. Zero-current potentials were measured using a high impedance electrometer (Keithley 617).



**Figure 12. Analysis of the putative quaternary structure of PorACj.** (A) Model of the octameric form of the PorACj channel in a lipid bilayer PorACj seen perpendicular to the membrane surface. Top view describes the initial setup of eight straight helices arranged in a circular manner to form a tube with a diameter of about 1.4 nm as derived from the experimental measurements. While the secondary structure is colored in purple, the individual amino-acid side chains are depicted as ball chains and colored according to their electrophysiological nature, i.e., neutral/hydrophobic in grey, polar in green, and charged in red (negative) and blue (positive), respectively. The surrounding bilayer is drawn as a grey surface. (B) Side view of the model of the octameric PorACj channel. After a few tenths of nanoseconds of unbiased molecular dynamics simulations, the helices kink in the central region - forming an hourglass shape - where several short side chains of the amino acids are located. The blue ball-stick side chains represent the lysines in the lower region, which are presumably responsible for the ion selectivity of PorACj and which form some kind of constriction zone. doi:10.1371/journal.pone.0075651.g012

## Results

### Cell Wall Proteins Effect the Conductance of Lipid Bilayer Membranes

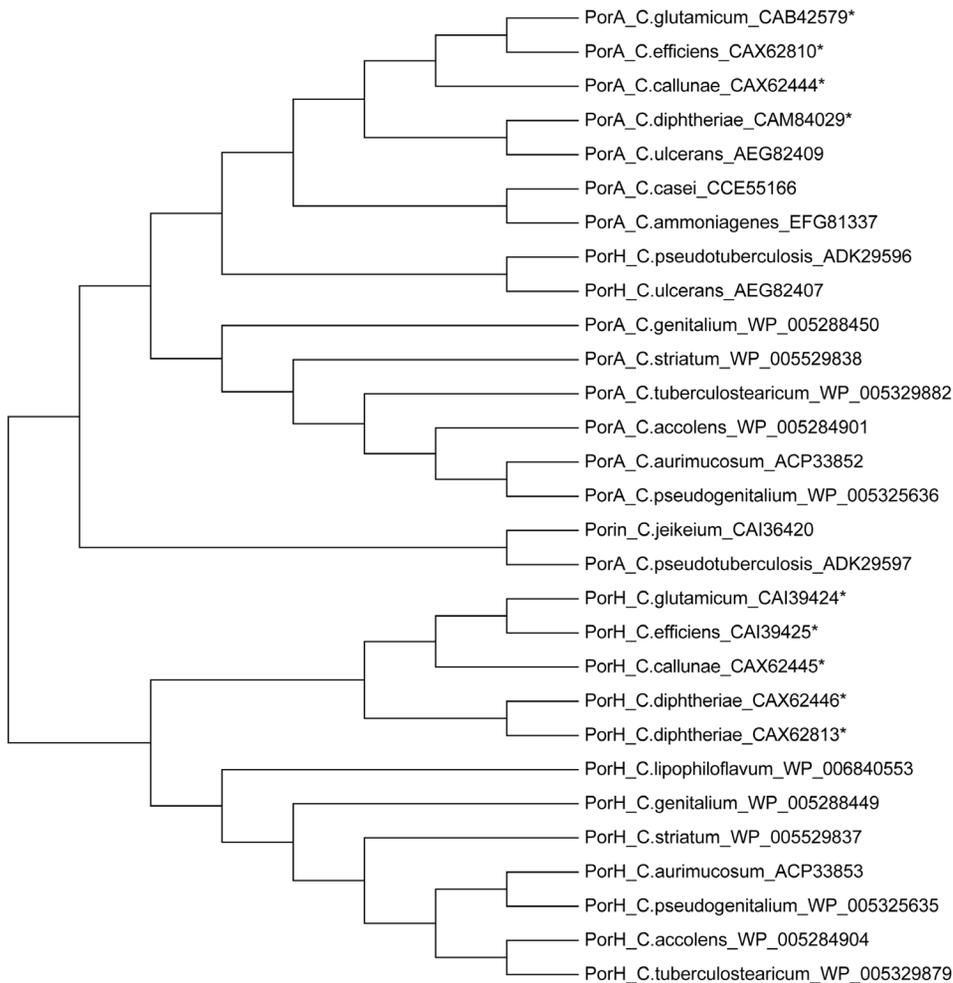
Cells of an overnight grown *C. jeikeium* culture were extracted with 1% LDAO. A few  $\mu\text{l}$  of the crude cell wall extract were tested in the lipid bilayer assay for pore-forming activity (see Figure 1A). Independently, if added to one or both sides of the lipid membranes two discrete conductance steps with 1.25 and 2.5 nS were observed at 20 mV applied membrane potential in 1 M KCl solution (see Figure 1B). Furthermore, the conductance increase caused by the detergent extract was a function of time after the addition of the protein to membranes in the black state. Within about 20 to 30 minutes the membrane conductance increased by several orders of magnitude above that of membranes without the extracts (from about  $0.05 \mu\text{S}/\text{cm}^2$  to  $150 \mu\text{S}/\text{cm}^2$ ). Control experiments with LDAO alone at the same concentration as in the experiments with extracts demonstrated that the membrane activity was caused by the presence of the extracts and not by the detergent. This result suggested that the channel-forming activity was an intrinsic property of the detergent extracts of whole *C. jeikeium* cells.

### Identification of the Gene Coding for the Cell Wall Channel of *C. jeikeium* K411

The extracts from whole *C. jeikeium* cells contained too many bands in tricine containing SDS-PAGE that it was impossible to relate one single band to the channel-forming activity although it showed a strong band in the low molecular mass region. Excision of the low molecular mass bands from preparative SDS-PAGE suggested that they contained the channel-forming proteins (data not shown). We looked for an alternative method to identify the channel-forming protein. Previously we could identify PorA of *C.*

*diphtheriae* based on its homology with PorA of *C. glutamicum* and other *Corynebacteriae* [20,30,44]. Therefore, we performed a similar approach here. A NCBI BLAST-translation tool search [54,55] using *porA* of *C. glutamicum* in the known genome of *C. jeikeium* K411 [39] did not lead to a clear indication for a homologous gene. However, search within the genome suggested that it contained an open reading frame (ORF) between the genes coding for GroEL2 (*jk0267*; [56]) and a hypothetical protein (polyphosphate kinase (PKK2), *jk0269*) that could code for a low molecular mass cell wall protein similar to PorA (see Figure 2). This means that an ORF (named *jk0268*) is localized within a region homologous to that of the *C. glutamicum* genome containing *porA* and *porH* [27]. Primers were designed to clone the whole region between the two genes *jk0267* (GROEL2) and *jk0269* (polyphosphate kinase) using DNA of *C. jeikeium* as a template (see Table 1). The PCR-product was cloned into the TOPO 2.1 vector and was sequenced. It contained *jk0268* that could code for a PorA-like protein.

The involvement of JK0268 (named in the following PorACj for PorA of *C. jeikeium*) in the observed channel formation of *C. jeikeium* extracts was examined by expressing the corresponding gene *jk0268* in a *C. glutamicum* mutant deficient of its main cell wall channel [44]. Whereas in terms of identical treatment cell wall samples of the *C. glutamicum* mutant did not show any channels reconstituted in a planar bilayer membrane, samples of pXJK0268 transformed and IPTG induced mutant cells caused well-defined and step-wise channel events that were very similar to those shown in Figure 1 (data not shown). These channels were identical to those observed after addition of the *C. jeikeium* extracts to lipid bilayer membranes, which means that two maxima of 1.25 and 2.5 nS were observed in 1 M KCl (see Figure 1). This result suggested that PorACj is the cell wall channel of *C. jeikeium*.



**Figure 13. Dendrogram representing the phylogenetic relationships of PorA and PorH of different *Corynebacterium* species obtained by the neighbor-joining method.** The tree was derived from the alignments of corresponding gene sequences. The support of each branch, as determined from 1,000 bootstrap samples, is indicated by the value at each node (in percent). The software used to construct alignment and tree was MEGA5.1. The sequence was aligned by ClustalW. Parameters: Multiple Alignment: Gap Opening Penalty: 10; Gap Extension Penalty: 0.2; Protein Weight Matrix: Gonnet; Gap Separation Distance: 4; Delay Divergent Cutoff (%): 30; The phylogenetic tree of corynebacterial species was constructed using the Maximum Likelihood statistical method; Substitution Model; Substitutions Type: Amino acid; Model/Method: Jones-Taylor-Thornton (JTT) model [81].

doi:10.1371/journal.pone.0075651.g013

### Comparison of PorACj with PorA and PorH of Other *Corynebacteriae*

Figure 3 shows a comparison of the sequences of the major cell wall proteins PorA and PorH of different *Corynebacteriae* with that of PorACj derived with Clustal W Protein sequences multiple alignments from Pole Bioinformatique Lyonnaise Network Protein Sequence Analysis (<http://npsa-pbil.ibcp.fr>). The multiple alignments were controlled with AliBee ([www.genebee.msu.su](http://www.genebee.msu.su)). The protein has similar to PorA and PorH known to date only the inducer methionine but no presequence similar to PorA and PorH of *C. glutamicum* [21,23], which means that the proteins are exported to the cell wall by a not yet identified mechanism. Similar is also the length of PorACj (40 amino acids) as compared to that of the different PorA (on average 43 amino acids) and PorH (57 amino acids) proteins. Otherwise, the homology is very poor in particular to PorH because only a limited number of amino acids are preserved (see Figure 3). Nevertheless, it is clear that the sequences of the different proteins are related, which means that

they are presumably descendents of a common ancestral protein (see Discussion).

### Antibiotic Resistance

The antibiotic susceptibility of *C. glutamicum*  $\Delta$ HA cells expressing PorACj was examined qualitatively by using the filter disk method and measuring the diameter of growth inhibition zones (see Table 2). The latter method was performed with *C. glutamicum* ATCC13032 wildtype as control, *C. glutamicum* ATCC13032  $\Delta$ HA and *C. glutamicum*  $\Delta$ HA pXHis PorACj. Whereas a considerable decrease of the antibiotic susceptibility was observed between wildtype and the  $\Delta$ HA mutant similar to that observed previously [20], expression of PorACj had an only minor influence on the antibiotic susceptibility of *C. glutamicum*  $\Delta$ HA pXHis PorACj. Despite expression of the channel former PorACj in *C. glutamicum*  $\Delta$ HA, the results of antibiotic susceptibility of this mutant strain showed that it was the same for antibiotics such as penicilline G, carbenicillin, ceftazidime, ertapenem and imipenem. Only for ampicillin, gentamicin and tetracycline the *C.*

*glutamicum*  $\Delta$ HA pXHis PorACj strain was slightly more susceptible indicating some minor increase of cell wall permeability in the presence of PorACj.

To determine the minimal concentration of antimicrobial agents, which inhibits the growth of the different *C. glutamicum* strains used in this study, we performed dilution susceptibility tests. This was achieved by dilution of antimicrobial agents in broth media. Antimicrobials were tested in series of two-fold dilutions. The minimum inhibitory concentration (MIC) was measured for *C. glutamicum*  $\Delta$ HA as a control and *C. glutamicum*  $\Delta$ HA pXHis PorACj. The results are summarized in Table 2. In general, reduction of antibiotic transport in bacteria using several classes of antibiotics such as  $\beta$ -lactams, aminoglycosides or fluoroquinolones, is mainly caused by changes of membrane permeability [19,57]. Loss of porins or expression of modified porin structure can cause low uptake of fluoroquinolones and  $\beta$ -lactams. The results of our study demonstrated that *C. glutamicum*  $\Delta$ HA was still susceptible to all antibiotics. There did also not exist significant differences between *C. glutamicum*  $\Delta$ HA and *C. glutamicum*  $\Delta$ HA pXHis PorACj for all antibiotics except for tetracycline, gentamycin and penicillin G. This suggests that *C. glutamicum*  $\Delta$ HA may have some leakage in the cell wall caused by deletion of its major cell wall channels and therefore did not provide a proper control for the change of cell wall permeability for uptake of antibiotics when PorACj is expressed.

### Heterologous Expression and Purification of the C-terminal His<sub>8</sub>-tagged and the N-terminal GST-tagged Channel-forming Protein PorACj

Cell wall preparations taken either from *C. jeikeium* or *C. glutamicum* expressing PorACj indicated a significant contribution of this protein to the observed channels. For purification of the protein we used a PCR based mutagenesis approach to introduce a DNA sequence coding a protease Xa cleavable, 8 histidine residues comprising protein tail which was added to the C-terminus of the wild-type gene. The constructed DNA cassette, accounting for the *C. glutamicum* codon usage frequencies [58], was introduced in the pXMJ19 expression vector and the stop codon (TAA) of *jk0268* was substituted (GGT) to fuse the peptide tail.

From IPTG-induced and detergent-extracted *C. glutamicum* cells expressing PorACj it was not possible to directly observe expression of neither the wild-type nor the modified *C. jeikeium* protein in SDS-PAGE. Only higher concentrated samples of chloroform/methanol precipitates combined with immunoassay using an anti-His antibody revealed expression of a small-sized protein as revealed by Western blots with detergent extracts of induced and not induced cells (see Figure 4A). Its molecular mass matched well with the calculated MW of PorACj-His (6.2 kDa). In a next step, the protein was purified from detergent extracts by Ni<sup>2+</sup> affinity chromatography (see Figure 4B). The pure and still histidine-tagged protein was able to form channels after trace amounts were added to the aqueous phase in the lipid bilayer setup (not studied in detail). To approach the situation of the native protein the His-tag was cleaved by protease Xa and the truncated PorACj protein (containing the C-terminal linker residues G-T-I-E-G-R) was again purified to homogeneity (Figure 4B). Similar results were obtained from lipid bilayer experiments with PorACj as obtained from expression in *E. coli* with the exception that uncleaved GST-PorACj did not form channels (data not shown).

The yield of the expression of the N-terminal GST-PorACj was at least ten times higher than described above for the His<sub>8</sub>-tagged protein. Purification of the GST-tagged PorACj was performed using glutathione Sepharose 4B medium. The protein was eluted

with an elution buffer containing 10 mM glutathione. The protein sample was concentrated to a concentration of 3 mg/ml pure GST-PorACj protein as judged by SDS-PAGE and Western blotting. (Data not shown). The GST-tag of the pure protein was removed by cleavage with FXa protease followed by treatment with glutathione Sepharose 4B medium to remove trace amounts of uncleaved GST-PorACj. The final concentration of pure PorACj was 0.65 mg/ml as judged by measurements at OD280. All further biophysical measurements were performed with pure PorACj obtained by cleavage of the His<sub>8</sub>-tag of PorA-His<sub>8</sub> or by cleavage of the GST-tag of GST-PorACj.

### Single-channel Analysis of PorACj

Addition of small amounts of the purified PorACj (~ 10 ng/ml) to one or both sides of a black lipid membrane made of PC/*n*-decane resulted after a few minutes delay in observation of step-like conductance increases. These channels had the same size as described above for detergent-solubilized material from *C. jeikeium* and *C. glutamicum* pXJK0268. This means that most of the steps were directed upwards indicating that the channels were for long time in an open state under low voltage conditions (20 mV; see Figure 5A). Only few channels showed some flickering indicating transitions between open and closed states. The statistical analysis similar to that shown in Figure 1B indicated that most of the channels (more than 40% of all fluctuations) caused conductivity steps with 1.25 or 2.5 nS in 1 M KCl (20 mV applied membrane potential; see Figure 5B). This means that beside a major conductance step of about 1.25 nS we observed also channels with a higher single-channel conductance, in particular channels with a single-channel conductance of about 2.5 nS. It is possible that the two different channels represent two different arrangements of the PorACj monomers (see discussion). Interestingly, we found a 1:2 relationship for the two maxima within the histograms under all conditions used here, which were denoted as left-side and right-side maximum in the histograms. Single-channel experiments were also performed with salts containing ions other than K<sup>+</sup> and chloride. These experiments were done to get some insight in the biophysical properties of the cell wall porin of *C. jeikeium*.

The results summarized in Table 3 suggested that the channel may be moderately selective for anions. This conclusion could be derived from experiments in which KCl was replaced by LiCl or KCH<sub>3</sub>COO. The exchange of the mobile ions K<sup>+</sup> and Cl<sup>-</sup> by the less mobile ions Li<sup>+</sup> and acetate<sup>-</sup> indicates that cations and anions have certain permeability through the channel of *C. jeikeium*, although the effect of anions on the single-channel conductance was more substantial. The permeability of the anions through the channels followed approximately their mobility sequence in the aqueous phase. This probably means that the channels formed by PorACj are wide and water-filled and have only a small field strength inside and no small selectivity filter (i.e. no binding site) as is suggested by the fact that also large organic anions could also penetrate the channel (see Discussion).

Table 3 shows also the average single-channel conductance, *G*, as a function of the KCl concentration in the aqueous phase. The values for *G* always corresponded to those of the two maxima in the histograms, i.e. to the 1.25 and 2.5 nS peaks in the case of 1 M KCl. Measurements were performed down to 0.01 M KCl. In contrast to other cell wall channels of the mycolata [25,28,30,59], we observed a linear relationship between single-channel conductance and KCl-concentration, which would be expected for wide water-filled channels that do not contain point charges similar to those formed by gram-negative bacterial porins [17,43]. This means the cell wall channels of *C. jeikeium* are together with those of *C. diphtheriae* [44] the first ones without point charge effects on the

channel properties within the taxon *Corynebacteriae* (see also Discussion).

### Selectivity of the Cell Wall Channel of *C. jeikeium*

Zero-current membrane potential measurements were performed to obtain further information on the structure of the *C. jeikeium* cell wall channel. The experiments were performed in the following way. After the incorporation of 100 to 1000 channels into the PC membranes bathed in 100 mM salt solution, the salt concentration on one side of the membranes was raised fivefold beginning from 100 mM and the zero-current potentials were measured 5 minutes after every increase of the salt gradient across the membrane. For KCl, LiCl and KCH<sub>3</sub>COO the more diluted side of the membrane (100 mM) always became negative, which indicated for all three salts preferential movement of the anions. This result indicates that the channel functions as a general diffusion pore for negative solutes. Analysis of the membrane potential using the Goldman-Hodgkin-Katz equation [53] confirmed the assumption that anions and cations are permeable through the channel. The ratios of the permeability  $P_{\text{cation}}$  and  $P_{\text{anion}}$  were 0.34 (KCl), 0.25 (LiCl) and 0.40 (potassium acetate), which means that the selectivity followed the mobility sequence of anions and cations in the aqueous phase, i.e. it is indeed water-filled.

### The Cell Wall Channel of *C. jeikeium* is Voltage-dependent

In single-channel recordings, the cell wall porin exhibited some flickering at higher voltages, i.e. it showed rapid transitions between open and closed configurations. This could be caused by voltage-dependent closure of PorACj, which was studied in detail in multi-channel and single-channel experiments. In the first set of experimental conditions, PorACj was added in a concentration of 100 ng/ml to one side of a black PC/*n*-decane membrane (the *cis*-side). After 30 minutes the conductance reached a stationary state. At this time, different positive and negative potentials were applied to the *cis*-side of the membrane. For negative and for positive potentials at the *cis*-side of the membrane the current decreased in an exponential fashion (see Figure 6 for voltages between  $\pm 30$  mV and  $\pm 60$  mV). This result indicated symmetrical voltage-dependence of the cell wall channel. The addition of the protein to the *trans*-side of the membrane or to both sides of the membrane also resulted in a symmetric response to the applied voltage (data not shown).

In a second set of experiments, PorACj was added in a concentration of about 10 ng/ml to one side of a black PC/*n*-decane membrane. After the reconstitution of about 10 channels into the membrane we applied +40 mV to the *cis*-side of the membrane and followed the time course of the current for about 10 minutes as it is shown in figure 7A. Because of the limited lifetime of the PorACj channel at higher voltages, the channels switched to substates. Figure 7B shows the distribution of the channel closures as a function of the conductance. Two main peaks of conductance were observed. One had a conductance of 1 nS and the other one 2 nS in 1 M KCl. When we keep the onset of conductance with 1.25 and 2.5 nS under the same conditions in mind (see figures 1 and 5), then it is possible that the 1.25 nS channel closes with 1 nS, whereas the 2.5 nS channel closes with 2.0 nS.

The data of the multi-channel experiments similar to that shown in figure 6 were analyzed in the following way: the membrane conductance ( $G$ ) as a function of voltage,  $V_m$ , was measured when the opening and closing of channels reached an equilibrium, i.e. after the exponential decay of the membrane current following the voltage step  $V_m$ .  $G$  was divided by the initial

value of the conductance  $G_o$ , (which was a linear function of the voltage) obtained immediately after the onset of the voltage. The data of figure 8 correspond to the symmetric voltage-dependence of the cell wall porin (mean of four membranes) when the protein was added to the *cis*-side (closed squares). To study the voltage-dependence in more detail the data of Figure 8 were analyzed assuming a Boltzmann distribution between the number of open and closed channels,  $N_o$  and  $N_c$ , respectively [60]:

$$N_o/N_c = \exp(nF(V_m - V_o)/RT) \quad (1)$$

$F$ ,  $R$  and  $T$  are standard symbols (Faraday constant, gas constant and absolute temperature, respectively),  $n$  is the number of charges moving through the entire transmembrane potential gradient for channel gating and  $V_m = V_o$  is the potential at which 50% of the total number of channels are in the closed configuration. The open-to-closed ratio of the channels,  $N_o/N_c$ , may be calculated from the data in Figure 8 according to

$$N_o/N_c = (G - G_{\text{min}})/(G_o - G) \quad (2)$$

$G$  is in this equation the conductance at a given membrane potential  $V_m$ ,  $G_o$  and  $G_{\text{min}}$  are the conductance at 10 mV (conductance of the open state) and at very high potentials, respectively. The data of Figure 8 could be fitted with combination of eqs. (1) and (2). The fit allowed the calculation of the number of gating charges  $n$  (number of charges involved in the gating process) and the midpoint potential  $V_o$  (potential at which the numbers of open and closed channels are identical). The midpoint potential for the addition of the protein to the *cis*-side was for applied positive voltages about +25 mV and for applied negative voltages about -24 mV. The gating charge was in both cases close to 2 (1.9 elementary charges).

### Investigation of the Secondary Structure of PorACj by CD Measurements

Secondary structure predictions ([http://npsa-pbil.ibcp.fr/cgi-bin/npsa\\_automat.pl?page=/NPSA/npsa\\_secons.html](http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_secons.html)) suggested that the channel formed by PorACj contained alpha-helices in contrast to the known structure of gram-negative bacterial porins [17] and MspA of the gram-positive *Mycobacterium smegmatis* [61], which are formed by beta-strands. Therefore, we considered it as a good model for studies, in which the folding characteristics and the structure of this protein under different conditions were investigated. PorACj and PorACj-His<sub>8</sub> were solubilized in a solution containing 100 mM NaCl, 50 mM Tris-HCl and 1 mM CaCl<sub>2</sub>, pH8 supplemented with 0.5% Genapol. The concentrations were 69  $\mu$ M for PorACj and 12  $\mu$ M for PorACj-His<sub>8</sub>, which was high enough to obtain a reasonable signal to noise ratio in the CD spectra. For both protein samples we found peaks in the CD-spectra around 208 nm and shoulders at about 220 nm (Figure 9A), which is typical for a high content of alpha-helical structures in proteins. This was confirmed by the fit procedure of the CD spectra that resulted in an estimate of about 80–100% alpha-helical structure, somewhat dependent on the type of public domain program used for the fit procedure. For control purpose, we performed CD-spectra with the same protein samples at the same concentrations to which urea was added in a concentration of 4 M. Figure 9B demonstrates that the spectra for both PorACj species changed completely. The peaks of the spectra are no

between 190 nm and 206 nm indicating that the spectra could be explained by random coil structure.

## Discussion

### The Genome of *C. jeikeium* Contains Only One Gene Coding for a PorA-like Cell Wall Channel

In this work, we extended our study of channel-forming proteins within the family of *Corynebacteriaceae* to the species *C. jeikeium*. Methodologies used previously for the isolation and characterization of cell wall-associated, channel-forming proteins showed that the supernatant of detergent extracted *C. jeikeium* cells also contained channel-forming units that form in the lipid bilayer assay channels with a conductance of 1.25 and 2.5 nS in 1 M KCl (Figure 1). Preparative SDS-PAGE suggested that a small sized protein present in *C. jeikeium* was responsible for the pore-forming activity similar to that of PorH and PorA porins of *C. glutamicum*, *C. callunae*, *C. efficiens* and *C. diphtheriae* [21,22,28,44]. Therefore, the localizations of *porH* and *porA* genes within the chromosomes were compared with the homologous region of *C. jeikeium*. The results suggested that these genes are located in a conserved region flanked by genes coding a chaperonin (GroEL2) and a polyphosphate kinase (PPK2). Applied to *C. jeikeium*, the region between the genes *jk0267* and *jk0269* was astonishingly smaller than that of the other *Corynebacteriae*. It contained only one open reading frame (ORF), *jk0268* containing 123 nucleotides and coding for a 40AS long polypeptide with a (calculated) MW of 4401 Da. The gene contains 6 bp upstream of the inducer methionine a putative ribosome-binding site (5'-AGGAG-3'). Furthermore, various predicted rho-independent terminator sequences suggested that the gene *jk0268* similar to the situation in the genome of *C. glutamicum* transcribes autonomously of the surrounding genes. Substantiated downstream of *jk0268* by a high-scored (100) stem-loop structure (5'-CCCCGGCCTTCGGCCGGGG-3') upstream structures are probably able to end GroEL2 mRNA transcription (score value 39–66) [62].

It is now clear that the major cell wall channels of most *Corynebacteriae*, i.e. those of *C. glutamicum*, *C. callunae*, *C. efficiens* and *C. diphtheriae*, are formed by oligomers of two small polypeptides, PorA and PorH [27]. However, this study demonstrates that in contrast to this the major cell wall channel of *C. jeikeium* is formed by an oligomer of a single PorA-like polypeptide. There are two clear findings supporting this. (i) There exists only a single gene (*jk0268*) between *jk0267* (GroEL2) and *jk0269* (PPK2), whereas in the same region within the genomes of the other *Corynebacteriae* two genes are localized that are transcribed together. (ii) We cloned *jk0268* in expression plasmids. Its expression in a *C. glutamicum* mutant that lacked the genes coding for PorA and PorH resulted in the same channels observed from detergent [or organic solvent] extracts of *C. jeikeium* K411 cells. Similarly, PorACj expressed in *E. coli* resulted in the same channels as the recombinant protein expressed in *C. glutamicum* ΔAH. This result revealed that oligomers of a short 40 amino acids long polypeptide are sufficient to form channels in the cell wall of *Corynebacteriae*.

Analogous to the situation of PorH and PorA proteins of *C. glutamicum*, *C. callunae*, *C. efficiens* and *C. diphtheriae* the *C. jeikeium* PorACj polypeptide does not contain N- or C-terminal or twin-arginine sorting signals commonly used by gram-positive bacteria for Sec- or TAT-transport [63–66]. This means that PorACj uses the same, but still unknown transport mechanism for cell wall proteins of *Corynebacteria* to overcome the cytoplasmic membrane to reach the cell wall. The single PorACj polypeptide monomers themselves are undoubtedly not large enough to form

the observed channels. On the other hand, oligomerization was frequently observed in the field of porin research within mycolate actinomycetes [28,31,44,61].

### PorACj Forms Wide Water-filled Channels with No Indication for Point Charges

We observed two conductance values in the reconstitution experiments with PorACj in all salt solutions (see Table 3). The single channel conductance showed a relation of about 1:2 and reflected most likely two different channel configurations. A different number of monomers in a PorACj oligomer may cause the two configurations because the pore-forming unit may partially or completely dissociate in detergents or organic solvents during the isolation and purification process. Two configurations of channel closures were also observed in experiments where we studied the voltage dependence of PorACj. However, it is clear from the experiments that the channel did not close completely at higher voltages (see figure 7). This means presumably that the two maxima in the histograms do not reflect the reconstitution of one or two channels at once but reflect two conformers of PorACj.

The conductance of the channels formed by PorACj was a linear function of the bulk aqueous concentration (see Table 3). Similarly, the selectivity of the PorACj channels was dependent on the mobility of the ions in the aqueous phase. This means that the PorACj channels sort mainly according to the molecular mass of the solutes similar to the function of general diffusion pores in gram-negative bacteria [17]. This result is very surprising because up to date many cell wall channels within the taxon mycolata were identified that contained charges in or near the channel opening. The channel presented in this study is the second channel within the *Corynebacterineae* besides PorA/PorH of *C. diphtheriae* [44] that does not contain point charges. This means also that the single channel analysis does not allow estimation of the channel size on the basis of the existence of point net charges as was performed for cell wall channels of *Mycobacteria*, *Nocardia* and *Corynebacteriae* [25,28,31]. On the other hand, a rough estimate of the size of the PorACj channel could be provided from a fit of the single channel data for salts of different anions using the Renkin equation [67]. This was possible because the channel was anion selective under all conditions, which means that the single-channel conductance was mostly limited by the permeability of the anions through PorACj. Furthermore, the permeability of the anions moving through the channel followed approximately their mobility sequences in the aqueous phase ( $\text{Br}^- > \text{Cl}^- > \text{NO}_3^- > \text{ClO}_3^- > \text{F}^- > \text{HCOO}^- > \text{CH}_3\text{COO}^-$ ). This means that the anions passing through the channel interior do not interact much with the channel. Table 4 contains the limiting molar conductivity adopted from ref. [68], the hydrated anion radii calculated according the Stokes equation [67] and the single-channel conductance of the *C. jeikeium* porin which was set to unity relative to the conductance in 0.1 KBr. The fit of the normalized single-channel conductance (mean of the left- and right-hand relative permeability) of PorACj with the Renkin correction factor (eqn. (2) of [67]) times the aqueous diffusion coefficient of the corresponding anion is shown in Figure 10. The best fit of the relative permeabilities was obtained with  $r = 0.7$  nm (diameter 1.4 nm). Thus, the *C. jeikeium* porin is ranking into known channel diameters varying from 1.4 over 2.0 to 3.0 nm of different mycolata, such as *Nocardia farcinica*, *Rhodococcus erythropolis* and *Mycobacterium smegmatis* [30,31,61]. However, it was considerably smaller than that of the main PorA/PorH cell wall channel from *C. glutamicum* with 2.2 nm [28].

### Putative Structure of the Channel Formed by PorACj

The comparison of the sequences of different PorA and PorH proteins with PorACj of *C. jeikeium* (Figure 3) demonstrated that the latter one is more comparable in size to the different PorA proteins. In addition, the alignments do not allow a distinctive allocation to PorA or PorH because each class affiliated members show with 13.6% (PorH) and 11.1% (PorA) a remarkable low degree of conserved residues as compared to PorACj. Nevertheless, there existed something like a structural homology between PorACj and the other two channel-forming proteins (PorA/H). Secondary structure predictions of all three proteins suggested that they contain heptameric repeat motive (abcdefg, see Figure 11A) indicating the existence of large  $\alpha$ -helical structures with hydrophobic and hydrophilic residues localized on different sides of the helices. Figure 11B shows the possible arrangement of the amino acids in PorACj in an  $\alpha$ -helix. This means that this protein could form an amphipathic helix similar to the possible secondary structure in the monomeric PorH and PorA proteins [21,22,30]. In agreement with the experimental data (demanding a water filled pore with 1.4 nm diameter), the *C. jeikeium* channel is postulated to consist of oligomeric  $\alpha$ -helical subunits. The number of subunits is an open question but this could influence channel conductance. Fig. 12A shows a possible arrangement of PorACj as an octamer seen perpendicular to the axis of the channel. Fig. 12B shows a side view of the octamer. The latter figure was created after 50 ns of unbiased molecular dynamics simulations. A fully stable conformation was not yet achieved at this stage and therefore the present structure suggestion is certainly preliminary (work in progress). Besides an octamer, it is also possible that the channel is formed by a hexamer, which would lead to a considerably smaller single-channel conductance, i.e. two maxima in the channel distribution as we found. We consider the possibility of an uneven number of monomers of PorACj in a channel as rather unlikely because the subunit of the PorA/H channel is presumably a PorA/PorH dimer. The present structure of the channel is definitely in contradiction to the 3D-structure of gram-negative bacterial porins [69] and that of the mycobacterial MspA channel [61] that both form  $\beta$ -barrel cylinders. On the other hand, it represents a similar structure as those of antibiotic channels, such as alamethicin [70], the cell wall porin PorB of *C. glutamicum* [71] and the ligand-gated ion channel in the inner membrane of *Erwinia chrysanthemi* (ELIC) [72].

This could mean that the PorACj monomer is orientated with the leucine, isoleucine and phenylalanine residues (Figure 12B, (g d a)) to the lipid phase, while glycine rich interfaces (c e) allow close contact to neighboring units. The four negative aspartates (D3, D28) and glutamates (E7, E31) together with the four lysines (K24, K27, K35, K40) are oriented to the channel lumen (f b e). At least one of the positive residues must take a dominant position in the otherwise charge-balanced protein causing the determined anion selectivity. The number of PorACj monomers in the homooligomeric channel is not known yet and need further experimental and

structural information about the channel. Five monomers are suggested to form the PorB channel of *C. glutamicum* [71] and the ligand-gated channel of prokaryotes [72]. Thus is very likely that the PorACj channel is formed by six to eight monomers because of its larger diameter.

### Is PorACj of *C. jeikeium* the Subunit of the Ancestral Cell Wall Channel of *Corynebacteriae*?

The many similarities between PorA/PorH and PorACj suggest that these channel-forming proteins form a family of proteins analogous to the MspABCD cell wall channel family of *M. smegmatis* and related species [58,73–75]. Even though PorACj structurally differs from the cell wall channel composition of the heterooligomeric channels within in the genus *Corynebacterium* there is a clear evidence for phylogenetic relationships of the investigated species. The different channel characteristics (e.g. diameter and selectivity) may indicate mirror adaptation to the wide-spread habitats of *Corynebacterium* species ranging from soil to skin and tissue of plants, animals as far as to man. In the first run mycolata were mainly classified according to properties of the phenotype and the chemical composition of their cell wall (containing meso-diaminopimelic acid, arabinose and glucose as major sugars [76]. Analysis of 16S rRNA provided deeper insight into separation of species within the monophyletic *Corynebacteria*-*Mycobacteria*-*Nocardia*-group although it cannot claim absolute classification accuracy [77–79]. Hence, strains yet known to be concomitant of *porH* and *porA* genes, namely *C. glutamicum*, *C. efficiens*, *C. diphtheriae*, *C. callunae* and many others likely evolved from an ancestor of *C. jeikeium* [78]. A similar picture may also be derived from a phylogenetic tree of the PorA/H family of channel-forming proteins together with PorACj from the genus *Corynebacterium* shown in Figure 13. The distance between PorACj of *C. jeikeium* and PorA/H of many *Corynebacterium* species is similar. This could indeed mean that *jk0268* (*porACj*) of *C. jeikeium* could be related to the ancestor of the genes *porH* and *porA* that may have evolved by gene duplication. The interesting point in this relationship is that the PorA family of proteins is more closely related to PorACj than the PorH family (see Figure 13).

### Acknowledgments

The authors would like to thank Luis Guimaraes, Institute for Genome Research and Systems Biology Center for Biotechnology (CeBiTec), Bielefeld University, for the design of the phylogenetic tree of the PorA/H family and Christian Andersen for helpful discussions.

### Author Contributions

Conceived and designed the experiments: EB NA AN WN RB. Performed the experiments: EB NA AN RS. Analyzed the data: EB NA AN RB. Contributed reagents/materials/analysis tools: WN UK AT. Wrote the paper: NA EB RS UK AT RB.

### References

- Eggeling L, Sahn H (1999) L-glutamate and L-lysine: traditional products with impetuous developments. *Appl Microbiol Biotechnol* 52: 146–153.
- Keilhauer C, Eggeling L, Sahn H (1993) Isoleucine synthesis in *Corynebacterium glutamicum*: molecular analysis of the *ilvB-ilvN-ilvC* operon. *J Bacteriol.* 175: 5595–603.
- Kinoshita S, Udaka S, Shimono M (1957) Studies on the amino acid fermentation. Part 1. Production of L-glutamic acid by various microorganisms. *J Gen Appl Microbiol* 50: p. 331–43.
- Leuchtenberger W (1996) Amino acids - technical production and use. Products of primary metabolism, in *Biotechnology*, Rehm HJ, et al., Editors., Wiley-VCH: Weinheim. p. 465–502.
- Sahn H, Eggeling L, Eikmanns B, Krämer R (1996) Construction of L-lysine-, L-threonine-, and L-isoleucine-overproducing strains of *Corynebacterium glutamicum*. *Ann N Y Acad Sci.* 782: 25–39.
- Udaka S (1960) Screening method for microorganisms accumulating metabolites and its use in the isolation of *Micrococcus glutamicus*. *J Bacteriol* 79: 754–5.
- Daffe M, Draper P (1998) The envelope layers of mycobacteria with reference to their pathogenicity. *Adv Microb Physiol* 39: 131–203.
- Barksdale L (1981) The genus *Corynebacterium*, in *The prokaryotes*, Starr MP et al., Editors. Springer-Verlag: Berlin. p. 1827–1837.
- Goodfellow M, Collins MD, Minnikin DE (1976) Thin-layer chromatographic analysis of mycolic acid and other long-chain components in whole-organism methanolysates of coryneform and related taxa. *J Gen Microbiol* 96: 351–8.

10. Ochi K (1995) Phylogenetic analysis of mycolic acid-containing wall-chemotype IV actinomycetes and allied taxa by partial sequencing of ribosomal protein AT-L30. *Int J Syst Bacteriol* 45: 653–60.
11. Minnikin DE (1987) Chemical targets in cell envelopes, in *Chemotherapy of tropical diseases*, Hopper M., Editor., John Wiley & Sons Ltd.: Chichester. p. 19–43.
12. Brennan PJ, Nikaido H (1995) The envelope of mycobacteria. *Annu Rev Biochem* 64: 29–63.
13. Daffé M, Brennan PJ, McNeil M (1990) Predominant structural features of the cell wall arabinogalactan of *Mycobacterium tuberculosis* as revealed through characterization of oligoglycosyl alditol fragments by gas chromatography/mass spectrometry and by <sup>1</sup>H and <sup>13</sup>C NMR analyses. *J Biol Chem* 265: 6734–43.
14. Holt JG, Kreig NR, Sneath PH, Stanley JT, Williams ST (1994). *Nocardioform actinomycetes*, in *Bergey's manual of determinative biology*, 9th edn. 1994, The Williams and Wilkins Co.: Baltimore. p. 625–650.
15. Minnikin DE, Patel P, Goodfellow M (1974) Mycolic acids of representative strains of *Nocardia* and the 'rhodochrous' complex. *FEBS Lett* 39: 322–324.
16. Yano I, Saito K (1972) Gas chromatographic and mass spectrometric analysis of molecular species of corynomycolic acids from *Corynebacterium ulcerans*. *FEBS Lett.* 23(3): p. 352–356.
17. Benz R (1994) Solute uptake through the bacterial outer membrane, in *Bacterial cell wall*, Ghuyens JM, Hakenbeck R Editors., Elsevier Science B.V.: Amsterdam. p. 397–423.
18. Nikaido H (1992) Porins and specific channels of bacterial outer membranes. *Mol Microbiol* 6: 435–42.
19. Nikaido H (2003) Molecular basis of bacterial outer membrane permeability revisited. *Microbiol Mol Biol Rev* 67: 593–656.
20. Costa-Riu N, Burkovski A, Krämer R, Benz R (2003) PorA represents the major cell wall channel of the gram-positive bacterium *Corynebacterium glutamicum*. *J Bact* 185: 4779–4786.
21. Hunten P, Costa-Riu N, Palm D, Lottspeich F, Benz R (2005) Identification and characterization of PorH, a new cell wall channel of *Corynebacterium glutamicum*. *Biochim Biophys Acta.* 1715: 25–36.
22. Hunten P, Schiffler B, Lottspeich F, Benz R (2005) PorH, a new channel-forming protein present in the cell wall of *Corynebacterium efficiens* and *Corynebacterium callunae*. *Microbiology.* 151: 2429–2438.
23. Lichtinger T, Rieß F, Burkovski A, Engelbrecht F, Hesse D, et al. (2001) The cell wall channel of the Gram-positive *Corynebacterium glutamicum*: immunological localization and molecular cloning and sequencing of its gene *porA*. *Eur J Biochem* 268: 462–469.
24. Trias J, Jarlier V, Benz R (1992) Porins in the cell wall of mycobacteria. *Science* 258: 1479–81.
25. Trias J, Benz R (1993) Characterization of the channel formed by the mycobacterial porin in lipid bilayer membranes. Demonstration of voltage gating and of negative point charges at the channel mouth. *J Biol Chem* 268: 6234–40.
26. Costa-Riu N, Maier E, Burkovski A, Krämer R, Lottspeich F, et al. (2003) Identification of an anion-specific channel in the cell wall of the gram-positive bacterium *Corynebacterium glutamicum*. *Mol Microbiol* 50: 1295–1308.
27. Barth E, Barcelo M, Kläckta C, Benz R (2010) Reconstitution experiments and gene deletions reveal the existence of two-component major cell wall channels in the genus *Corynebacterium*. *J Bacteriology*, 192: 786–800.
28. Lichtinger T, Burkovski A, Niederweis M, Krämer R, Benz R (1998) Biochemical and biophysical characterization of the cell wall channel (porin) of *Corynebacterium glutamicum*: the channel is formed by a low molecular mass subunit. *Biochemistry* 37, 15024–15032.
29. Lichtinger T, Heym B, Maier E, Eichner H, Cole ST, et al. (1999) Evidence for a small anion-selective channel in the cell wall of *Mycobacterium bovis* BCG besides a wide cation-selective pore. *FEBS Lett.* 454: 349–355.
30. Lichtinger T, Reiss G, Benz R (2000) Biochemical identification and biophysical characterization of a channel-forming protein from *Rhodococcus erythropolis*. *J Bacteriol.* 182: 764–70.
31. Rieß FG, Lichtinger T, Cseh R, Yassin AF, Schaal KP, et al. (1998) The cell wall channel of *Nocardia farcinica*: Biochemical identification of the channel-forming protein and biophysical characterization of the channel properties. *Mol Microbiol* 29: 139–150.
32. Riess FG, Benz R (2000) Discovery of a novel channel-forming protein in the cell wall of the non-pathogenic *Nocardia corynebacteroides*. *Biochim Biophys Acta* 1509: 485–95.
33. MacGregor RR (1995) *Corynebacterium diphtheriae*, in *Principles and practices of infectious diseases*, 4th edn., Mandell GL, Bennett JE, Dolin R, Editors., Churchill Livingstone: New York. p. 1865–1872.
34. Jackman PJ, Pelczynska S (1986) Characterization of *Corynebacterium* group JK by whole-cell protein patterns. *J Gen Microbiol* 132: 1911–5.
35. Oteo J, Aracil B, Ignacio Alós J, Luis Gómez-Garcés J (2001) [Significant bacteremias by *Corynebacterium amycolatum*: an emergent pathogen]. *Enferm Infecc Microbiol Clin* 19: 103–6.
36. Johnson WD, Kaye D (1970) Serious infections caused by diphtheroids. *Ann N Y Acad Sci*, 1970. 174(2): p. 568–76.
37. Funke G, von Graevenitz A, Clarridge JE 3rd, Bernard KA (1997) Clinical microbiology of coryneform bacteria. *Clin Microbiol Rev.*, 10(1): p. 125–59.
38. Mookadam F, Cikes M, Baddour LM, Tleyjeh IM, Mookadam M. (2006) *Corynebacterium jeikeium* endocarditis: a systematic overview spanning four decades. *Eur J Clin Microbiol Infect Dis* 25: 349–53.
39. Tauch A, Kaiser O, Hain T, Goesmann A, Weisshaar B, et al. (2005) Complete genome sequence and analysis of the multiresistant nosocomial pathogen *Corynebacterium jeikeium* K411, a lipid-requiring bacterium of the human skin flora. *J Bacteriol* 187: 4671–82.
40. Tauch A, Trost E, Tilker A, Ludewig U, Schneider S, et al. (2008) The lifestyle of *Corynebacterium urealyticum* derived from its complete genome sequence established by pyrosequencing. *J Biotechnol* 136: 11–21.
41. Brune I, Becker A, Paarmann D, Albersmeier A, Kalinowski J, et al. (2006) Under the influence of the active deodorant ingredient 4-hydroxy-3-methoxybenzyl alcohol, the skin bacterium *Corynebacterium jeikeium* moderately responds with differential gene expression. *J Biotechnol* 127: 21–33.
42. Hansmeier N, Chao TC, Daschke S, Müsken M, Kalinowski J, et al. (2007) A comprehensive proteome map of the lipid-requiring nosocomial pathogen *Corynebacterium jeikeium* K411. *Proteomics* 7: 1076–96.
43. Benz R (1988) Structure and function of porins from gram-negative bacteria. *Annu Rev Microbiol.* 42: 359–93.
44. Schiffler B, Barth E, Daffé M, Benz R (2007) *Corynebacterium diphtheriae*: identification and characterization of a channel-forming protein in the cell wall. *J Bacteriol* 189: 7709–19.
45. Tauch A, Bischoff N, Pühler A, Kalinowski J. (2004) Comparative genomics identified two conserved DNA modules in a corynebacterial plasmid family present in clinical isolates of the opportunistic human pathogen *Corynebacterium jeikeium*. *Plasmid* 52: 102–18.
46. Schäfer A, Schwarzer A, Kalinowski J, Pühler A (1994) Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. *Gene* 145: 69–73.
47. van der Rest ME, Lange C, Molenaar D. (1999) Heat shock following electroporation induces highly efficient transformation of *Corynebacterium glutamicum* with xenogenic plasmid DNA. *Appl Microbiol Biotechnol* 52: 541–5.
48. Schagger H, von Jagow G (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem* 166: 368–79.
49. Blum H, Beier H, Gross HJ (1987) Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. *Electrophoresis* 8: 93–99.
50. Towbin H, Staehelin T, Gordon J (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci U S A* 76: p. 4350–4.
51. Greenfield N, Fasman GD. (1969) Computed circular dichroism spectra for the evaluation of protein conformation. *Biochemistry* 8: 4108–4116.
52. Benz R, Janko K, Boos W, Läger P (1978) Formation of large, ion-permeable membrane channels by the matrix protein (porin) of *Escherichia coli*. *Biochim Biophys Acta* 511: 305–19.
53. Benz R, Janko K, Läger P (1979) Ionic selectivity of pores formed by the matrix protein (porin) of *Escherichia coli*. *Biochim Biophys Acta* 551: 238–47.
54. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215: 403–10.
55. Zhang J, Madden TL (1997) PowerBLAST: a new network BLAST application for interactive or automated sequence analysis and annotation. *Genome Res* 7: 649–56.
56. Barreiro C, González-Lavado E, Pátek M, Martín JF (2004) Transcriptional analysis of the *groES-groEL1*, *groEL2*, and *dnaK* genes in *Corynebacterium glutamicum*: characterization of heat shock-induced promoters. *J Bacteriol* 186: 4813–7.
57. De E, Basle A, Jaquinod M, Saint N, Mallea M, et al. (2001) A new mechanism of antibiotic resistance in Enterobacteriaceae induced by a structural modification of the major porin. *Mol. Microb.* 41: 180–198.
58. Hallin PF, Ussery DW (2004) CBS Genome Atlas Database: a dynamic storage for bioinformatic results and sequence data. *Bioinformatics* 20: 3682–6.
59. Riess FG, Dörner U, Schiffler B, Benz R (2001) Study of the properties of a channel-forming protein of the cell wall of the gram-positive bacterium *Mycobacterium phlei*. *J Membr Biol* 182: 147–57.
60. Ludwig O, De Pinto V, Palmieri F, Benz R (1986) Pore formation by the mitochondrial porin of rat brain in lipid bilayer membranes. *Biochim Biophys Acta* 860: 268–76.
61. Faller M, Niederweis M, Schulz GE. (2004) The structure of a mycobacterial outer-membrane channel. *Science* 303: 1189–1192.
62. Kingsford CL, Ayanbule K, Salzberg SL (2007) Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. *Genome Biol* 8: R22.
63. Freudl R (1992) Protein secretion in gram-positive bacteria. *J Biotechnol* 23: 231–40.
64. Berks BC, Sargent F, Palmer T (2000) The Tat protein export pathway. *Mol Microbiol* 35: 260–74.
65. van Wely KH, Swaving J, Freudl R, Driessen AJ (2001) Translocation of proteins across the cell envelope of Gram-positive bacteria. *FEMS Microbiol Rev* 25: 437–54.
66. Ton-That H, Marraffini LA, Schneewind O (2004) Protein sorting to the cell wall envelope of Gram-positive bacteria. *Biochim Biophys Acta* 1694: 269–78.
67. Trias J, Benz R (1994) Permeability of the cell wall of *Mycobacterium smegmatis*. *Mol Microbiol.* 14: 283–90.
68. Castellan GW (1983) The ionic current in aqueous solutions, in *Physical Chemistry*. Reading, MA: Addison-Wesley, 769–780.

69. Schulz GE (2004) The structures of general porins, in Bacterial and Eukaryotic Porins, Benz R, Editor, Wiley-VCH: Weinheim. p. 25–40.
70. Chugh JK, Wallace BA (2001) Peptaibols: models for ion channels. *Biochem Soc Trans* 29: 565–70.
71. Ziegler K, Benz R, Schulz GE (2008) A putative alpha-helical porin from *Corynebacterium glutamicum*. *J Mol Biol.* 379: 482–91.
72. Hilf RJ, Dutzler R (2008) X-ray structure of a prokaryotic pentameric ligand-gated ion channel. *Nature* 452: 375–9.
73. Niederweis M, Ehart S, Heinz C, Klöcker U, Karosi S, et al. (1999) Cloning of the mspA gene encoding a porin from *Mycobacterium smegmatis*. *Mol Microbiol* 33: 933–45.
74. Stephan J, Bender J, Wolschendorf F, Hoffmann C, Roth E, et al. The growth rate of *Mycobacterium smegmatis* depends on sufficient porin-mediated influx of nutrients. *Mol Microbiol* 2005. 58: 714–30.
75. Dörner U, Maier E, Benz R (2004) Identification of a cation-specific channel (TipA) in the cell wall of the gram-positive mycolata *Tsukamurella inchenensis*: the gene of the channel-forming protein is identical to mspA of *Mycobacterium smegmatis* and mppA of *Mycobacterium phlei*. *Biochim Biophys Acta* 1667: 47–55.
76. Lechevalier HA, Lechevalier MP, Gerber NN (1971) Chemical composition as a criterion in the classification of actinomycetes. *Adv Appl Microbiol* 14: 47–72.
77. Ruimy R, Riegel P, Boiron P, Monteil H, Christen R (1995) Phylogeny of the genus *Corynebacterium* deduced from analyses of small-subunit ribosomal DNA sequences. *Int J Syst Bacteriol* 45: 740–746.
78. Pascual C, Lawson PA, Farrow JA, Gimenez MN, Collins MD (1995) Phylogenetic analysis of the genus *Corynebacterium* based on 16S rRNA gene sequences. *Int J Syst Bacteriol* 45: 724–728.
79. Khamis A, Raoult D, La Scola B (2004) rpoB gene sequencing for identification of *Corynebacterium* species. *J Clin Microbiol* 42: 3925–31.
80. Kyte J, Doolittle RF. (1982) A simple method for displaying the hydropathic character of a protein. *J Mol Biol* 157: 105–32.
81. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol.* 28: 2731–2739.

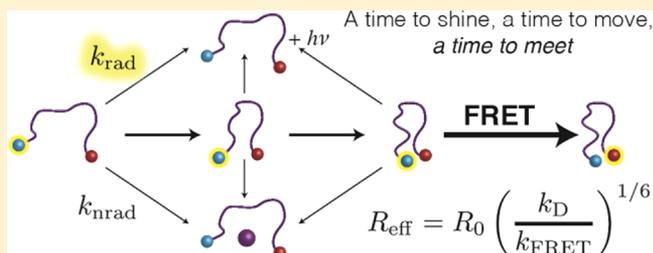
# Diffusion-Enhanced Förster Resonance Energy Transfer and the Effects of External Quenchers and the Donor Quantum Yield

Maik H. Jacob,\* Roy N. Dsouza,<sup>†</sup> Indrajit Ghosh,<sup>†</sup> Amir Norouzy, Thomas Schwarzlose, and Werner M. Nau\*

School of Engineering and Science, Jacobs University Bremen, Campus Ring 1, D-28759, Bremen, Germany

## S Supporting Information

**ABSTRACT:** The structural and dynamic properties of a flexible peptidic chain codetermine its biological activity. These properties are imprinted in intrachain site-to-site distances as well as in diffusion coefficients of mutual site-to-site motion. Both distance distribution and diffusion determine the extent of Förster resonance energy transfer (FRET) between two chain sites labeled with a FRET donor and acceptor. Both could be obtained from time-resolved FRET measurements if their individual contributions to the FRET efficiency could be systematically varied. Because the FRET diffusion enhancement (FDE) depends on the donor-fluorescence lifetime, it has been proposed that the FDE can be reduced by shortening the donor lifetime through an external quencher. Benefiting from the high diffusion sensitivity of short-distance FRET, we tested this concept experimentally on a (Gly–Ser)<sub>6</sub> segment labeled with the donor/acceptor pair naphthylalanine/2,3-diazabicyclo[2.2.2]-oct-2-ene (NAla/Dbo). Surprisingly, the very effective quencher potassium iodide (KI) had no effect at all on the average donor–acceptor distance, although the donor lifetime was shortened from ca. 36 ns in the absence of KI to ca. 3 ns in the presence of 30 mM KI. We show that the proposed approach had to fail because it is not the experimentally observed but the radiative donor lifetime that controls the FDE. Because of that, any FRET ensemble measurement can easily underestimate diffusion and might be misleading even if it employs the Haas–Steinberg diffusion equation (HSE). An extension of traditional FRET analysis allowed us to evaluate HSE simulations and to corroborate as well as generalize the experimental results. We demonstrate that diffusion-enhanced FRET depends on the radiative donor lifetime as it depends on the diffusion coefficient, a useful symmetry that can directly be applied to distinguish dynamic and structural effects of viscous cosolvents on the polymer chain. We demonstrate that the effective FRET rate and the recovered donor–acceptor distance depend on the quantum yield, most strongly in the absence of diffusion, which has to be accounted for in the interpretation of distance trends monitored by FRET.



## INTRODUCTION

Many bioactive peptides, proteins, and nucleic acids are too flexible to be fully accessible to traditional methods of structure elucidation. Examples are antimicrobial and cell-penetrating peptides, amyloidogenic peptides,<sup>1,2</sup> globular proteins in diverse stages of enzymatic catalysis as well as in diverse phases of refolding and unfolding and in the unfolded state,<sup>3–10</sup> and the large class of intrinsically disordered or natively unfolded proteins.<sup>1,2,11</sup> Nevertheless, every polypeptide chain shows distinct structural and dynamic patterns, which—if they were known—could assist in comprehending or even predicting biological activity.<sup>12,13</sup>

In this context, Förster resonance energy transfer (FRET) has increasingly come into focus. FRET from an energy donor dye to an acceptor dye, both conjugated to the chain at selected positions, is distance dependent and can thereby bring the sensitivity and speed of fluorescence spectroscopy to distance and structure determination of polypeptides (and nucleic acids).<sup>3,4,14–24</sup> Time-resolved FRET on the ensemble and also single-molecule FRET<sup>25</sup> offer the additional possibility to

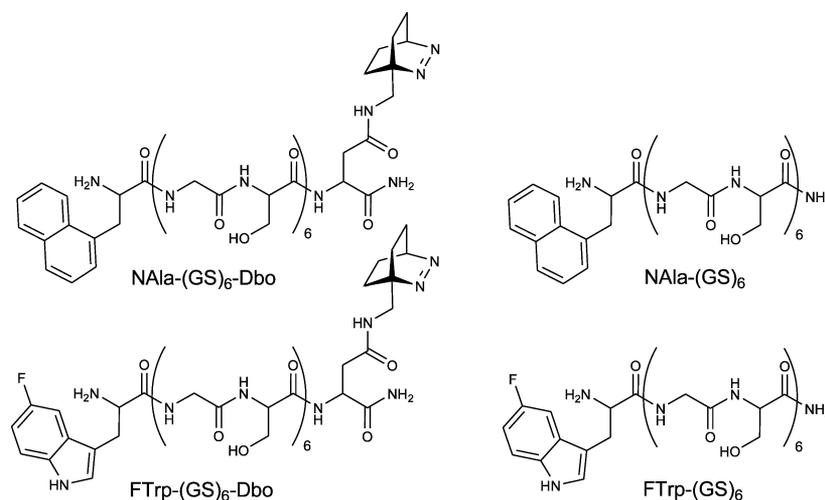
obtain information on the dynamic properties of the peptide chain, i.e., on intrachain diffusion; in general, this diffusion enhances FRET.<sup>2,25</sup> Unfortunately, although the analysis of time-resolved FRET data has been pioneered by Steinberg and Haas,<sup>26–30</sup> and expanded by Lakowicz and Kusba,<sup>31–33</sup> both experimental and computational challenges have arisen which have prevented it from becoming routine. The main roadblock is that the efficiency of FRET is determined by *both* chain structure and dynamics, that is, by the distance distribution of and by the mutual diffusion between two labeled chain positions. They can only be reliably delineated from time-resolved-fluorescence measurements if their relative contributions can be independently determined or, at least, systematically modified, in a series of experiments.<sup>27</sup>

Lakowicz and co-workers proposed an elegant idea to systematically reduce the FRET diffusion enhancement, the

Received: October 19, 2012

Revised: December 5, 2012

Published: December 6, 2012



**Figure 1.** Selected donor–acceptor and donor-only labeled peptides composed of Gly–Ser units. Donors are naphthyl-1-L-alanine (NAla) and 5-fluoro-L-tryptophan (FTTrp); the acceptor is Dbo, whose optically active group is the azo group of the bicyclic chromophore.

FDE, by addition of external quenchers.<sup>34</sup> Their idea was inspired by the random-walk perspective,<sup>35</sup> for which Lakowicz gave the following numerical example:<sup>36</sup> The possible root-mean-square (rms) donor–acceptor displacement,  $\Delta R_{DA}$ , that can occur during the lifetime of the donor,  $\tau_D$ , is obtained from  $\Delta R_{DA} = (2D\tau_D)^{1/2}$ .<sup>36</sup>  $\Delta R_{DA}$  becomes 3.2 Å when the diffusion coefficient,  $D$ , is set to 1 Å<sup>2</sup>/ns and the donor lifetime is set to 5 ns—a representative value for the experimental lifetimes of customary dyes. For a chain ensemble that displays an average distance of 20 Å, a displacement of 3.2 Å seems to be of limited significance. However, whenever a large displacement critically distorts the distance analysis, an external quencher can be added to the measurement solutions that, as proposed, reduces the donor lifetime and, by that, the displacement and the FDE.<sup>34</sup>

We will demonstrate that while the random-walk perspective grasps essential physical aspects of the FDE, it requires several modifications from the original idea. A minor correction, as already pointed out by others,<sup>37</sup> considers that mutual donor–acceptor diffusion takes place in three dimensions, leading to eq 1,<sup>35</sup> and results in an even larger rms displacement of 5.5 Å.

$$\Delta R_{\text{rms}} = (6D\tau_D)^{1/2} \quad (1)$$

More intriguingly, we will demonstrate that any donor lifetime reduction by a quencher cannot reduce the FDE. In fact, the donor lifetime to be inserted into eq 1 is the radiative and not the experimentally measured lifetime, which is often shorter by an order of magnitude. In the above example, the donor–acceptor displacement can then become as large as the average distance, exposing diffusion as an often dominant contributor to the FRET efficiency. As the impact of diffusion on FRET is much larger than has previously been assumed, other propositions of how to separately determine the distance distribution and the diffusion coefficient have to be revisited, too.<sup>34,38</sup>

To measure FDE experimentally, we will employ short-distance FRET,<sup>39–43</sup> i.e., FRET employing a donor–acceptor pair with a very short critical transfer distance of only 10 Å. This will allow us to follow the FDE with the highest resolution presently available.<sup>40,42</sup> On the theoretical level, we use known and introduce novel tools of elementary FRET analysis that will enable us to corroborate and generalize the experimental results

by numerical simulations based on the Haas–Steinberg equation (HSE).<sup>3,6,27,28,34,38</sup> On our way, we have discovered critical aspects of FRET analysis that were never pointed out before.

## MATERIALS AND METHODS

Donor-only and donor–acceptor labeled (GS)<sub>6</sub> peptides were obtained in higher than 95% purity from Biosyntan (Berlin). All peptides were amidated at the C terminus to exclude attractive interactions between the terminal positions. Fmoc-Dbo and 5-fluoro-L-tryptophan (FTTrp) used in peptide synthesis were prepared according to literature methods.<sup>44,45</sup> Other chemicals were from Sigma-Aldrich. The ethylene glycol concentration was determined by refractive index measurement as previously described.<sup>46,47</sup> Absorption and fluorescence spectra were recorded on a Varian Cary 4000 UV–vis spectrometer and a Varian Cary Eclipse fluorometer. Time-resolved-fluorescence decays were followed by time-correlated single-photon counting with a FLS920 lifetime spectrometer (Edinburgh Instruments). A pulsed diode laser (Picoquant) was used for selective donor excitation at 280 nm. Peptide concentrations were determined by using extinction coefficients at 280 nm of  $5.7 \times 10^3 \text{ M}^{-1} \text{ cm}^{-1}$  (FTTrp)<sup>48</sup> and  $5.5 \times 10^3 \text{ M}^{-1} \text{ cm}^{-1}$  (naphthylalanine, NAla)<sup>49</sup> and were adjusted to about 10 μM in aerated solutions at 25 °C and pH 5.0. The donor quantum yields were determined by comparison with *N*-acetyltryptophanamide (0.14, pH 7.0).<sup>39,50,51</sup> The Förster radii of NAla/Dbo and FTTrp/Dbo were determined by using the absorption and emission spectra of the single-labeled peptides as previously described.<sup>39,41</sup> Fluorescence decays were analyzed by using the instrument software as described<sup>39</sup> and by using ProFit (Quantumsoft). The reproducibility of the reported fluorescence lifetimes for the monoexponential decays was ±3%. In the simulations, the HSE was solved numerically by using the finite element method implemented in the PDE toolbox of Matlab (MathWorks).

## RESULTS

**Short-Distance FRET Grants Maximal Diffusion Enhancement.** Short-distance FRET employs donor–acceptor pairs with Förster radii of about 10 Å.<sup>39–42,52</sup> Thus, already a 1-Å diffusional displacement from a 10-Å donor–acceptor

distance increases the FRET efficiency and the donor decay rate to an extent that can be precisely determined.<sup>40</sup> Short-distance FRET is based on the acceptor Dbo, an asparagine residue conjugated to 2,3-diazabicyclo[2.2.2]oct-2-ene (DBO). DBO's exceptionally small extinction coefficient causes a small spectral overlap of acceptor absorption and donor emission, which causes the observed short Förster radii.<sup>39,42,52</sup> The point-dipole approximation used in Förster's theory requires that the transition dipole vectors of donor and acceptor are much shorter than the distance between them.<sup>53</sup> The transition dipole of DBO is located on two atoms only, on the N=N group, which constitutes the azo chromophore. Quantum mechanical calculations on the Trp/DBO FRET pair yielded the conclusion that DBO as acceptor is a superior choice for short-distance detection.<sup>52</sup>

Previously, we have only used Trp as donor, but its intrinsic lifetime heterogeneity complicates data analysis.<sup>39,54</sup> In this work, we introduce two new FRET donors with mono-exponential fluorescence kinetics. The first, naphthyl-1-L-alanine (NALa), has an intrinsic lifetime of about 35 ns in aerated water and, thus, allows ample time for mutual donor-acceptor diffusion; the second, 5-fluoro-L-tryptophan (FTrp, Figure 1) has a lifetime of 2.0 ns, comparable to that of Trp with 1.4 ns.<sup>55,56</sup> Because all three donors, NALa, FTrp, and Trp, are virtually isosteric, the peptide chains (which differ only slightly in the donor structure) were expected to show identical structure and dynamics. Since Gly-Ser peptides have been shown to be highly flexible and have been investigated by a variety of methods, including FRET and collision-induced quenching,<sup>38,39,44,57-61</sup> we have selected a (GS)<sub>6</sub> sequence as the principal model, which afforded sizable FRET for all investigated peptides. For consistency, peptides of the sequence Donor-(GS)<sub>6</sub>-Dbo were used (N-terminal donor and C-terminal acceptor) and the donor-only peptide, Donor-(GS)<sub>6</sub>, was always included as a reference peptide for the donor decay in the absence of FRET (Figure 1).

**Diffusion-Enhanced FRET Monitored through the Effective Donor-Acceptor Distance.** Essential aspects of the impact of diffusion can already be accessed through a conventional FRET analysis. After optical excitation, the donor transfers its energy to the acceptor with a rate constant,  $k_T$ , which decreases with the sixth power of the donor-acceptor distance,  $R_{DA}$  (eq 2). Equation 2 is a central result of Förster theory and shows that the FRET rate depends also on the rate of donor deactivation,  $k_D$ , in the absence of the acceptor, that is, without competition by FRET. The Förster radius,  $R_0$ , is the donor-acceptor distance at which the FRET rate is as fast as deactivation by all non-FRET radiative and radiationless decay processes, including donor deactivation by an external quencher. The Förster radii were obtained from eq 3.<sup>62</sup> Here, the numerical constant  $c = 9 \ln 10 / (128\pi^5 N)$ , with  $N$  being Avogadro's number.<sup>63-65</sup> The orientation factor,  $\kappa^2 = 2/3$ , assumes fast and random sampling of donor-acceptor orientations; this has been shown by molecular dynamics calculations to be a good approximation for the peptides under investigation.<sup>39,41</sup> The overlap integral,  $J$ , is determined by the extent by which the donor emission spectrum, normalized to an area of 1, i.e.,  $F_D(\lambda)$ , overlaps with the spectrum,  $\varepsilon(\lambda)$ , of the acceptor absorption coefficient:  $J = \int d\lambda F_D(\lambda) \varepsilon(\lambda) \lambda^4$ . The Förster radius decreases with the refractive index,  $n$ , and increases with the donor quantum yield,  $\Phi_D$ .

$$k_T = k_D(R_0/R_{DA})^6 \quad (2)$$

$$R_0^6 = c\kappa^2 J \Phi_D / n^4 \quad (3)$$

The analysis is centered around the energy transfer efficiency (ETE),  $E$ , which is simply the fraction of donor-deactivation events,  $\Delta N^*$ , caused by FRET;  $E = \Delta N^*_{\text{FRET}} / \Delta N^*$ . Thus, it is also the ratio of the FRET rate over the sum of all rates of donor deactivation (eq 4). The ETE depends only on the Förster radius and the donor-acceptor distance (eq 5), which is seen when  $k_T$  in eq 4 is substituted by the FRET distance law (eq 2). Once the Förster radius has been determined, the measured ETE gives access to  $R_{DA}$ .

$$E = k_T / (k_T + k_D) \quad (4)$$

$$E = R_0^6 / (R_0^6 + R_{DA}^6) \quad (5)$$

The ETE itself is determined from time-resolved-fluorescence measurements by assuming that the observed rate of deactivation in the donor-acceptor peptide,  $k_{DA}$ , is composed of the FRET rate and that of donor deactivation in the donor-only peptide:  $k_{DA} = k_T + k_D$ . Rate and time constants such as the donor lifetime are reciprocally related ( $k_D = \tau_D^{-1}$ ), and both quantities are henceforth used interchangeably. This leads to eq 6, which is directly applicable in the case of monoexponential kinetics. In the case of more complex kinetics, amplitude-weighted lifetimes are used or, most straightforwardly, the areas under the decay curves (eq 7) are used as direct measures of the total fluorescence emitted by the donor-only or the donor-acceptor peptide.<sup>39,54</sup> The ETE can also be accessed through steady-state measurements of the fluorescence intensities,  $I$  (eq 8), at equal concentrations of donor-only and donor-acceptor labeled peptides. Of course, in the limit of a perfect experiment with infinitely fast temporal resolution (where no fluorescence can escape by being conceived as static in nature), the relative steady-state-fluorescence intensities (eq 8) and relative areas under the time-resolved-fluorescence decay curves (eq 7) should be the same.

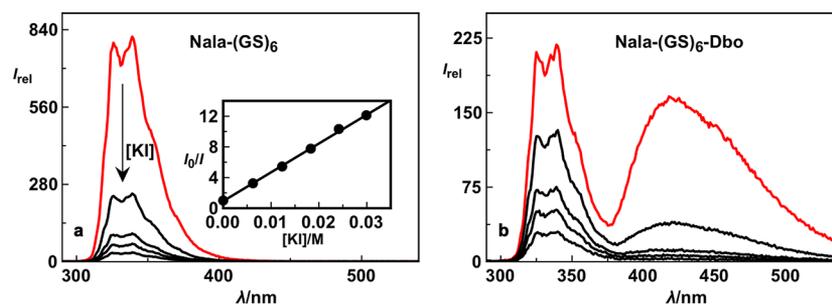
$$E_{tr} = (\tau_D - \tau_{DA}) / \tau_D \quad (6)$$

$$E_{tr} = (A_D - A_{DA}) / A_D \quad (7)$$

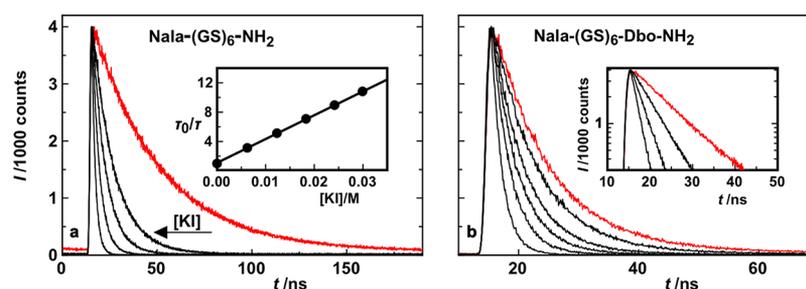
$$E_{ss} = (I_D - I_{DA}) / I_D \quad (8)$$

The  $R_{DA}$  value determined from the ETE (eq 5) is a realistic physical distance only if the donor-acceptor separation does not vary among the molecules of the ensemble and when it does not vary with time. These conditions are approximately met in folded single-domain proteins,<sup>37</sup> in short polyproline chains of rigid helical PPII conformation,<sup>38</sup> or in the solid state.<sup>41,42</sup> However, when the same elementary procedure is used to determine the donor-acceptor distance for a chain ensemble characterized by a probability distribution of static distances,  $p(r)$ , the observed ETE includes FRET contributions from the entire range of distances and follows  $E_{obs} = \int E(r) p(r) dr$ . Because the efficiency at a specific distance,  $E(r)$ , is much larger at short distances than at long distances, the effective distance,  $R_{eff}$ , calculated from the observed ETE (cf. eq 5) will always be shorter than the average ensemble distance.<sup>42</sup>

A much larger apparent distance distortion is usually caused by chain mobility, that is, by mutual donor-acceptor motion. While the donor resides in its excited state, its distance to the acceptor is lengthened and shortened by mutual diffusion. The measured ETE and the effective FRET rate derived from it (cf. eq 4) is raised because any motion toward shorter distance



**Figure 2.** (a) Steady-state-fluorescence spectra of NAla-(GS)<sub>6</sub>-NH<sub>2</sub> ( $\lambda_{\text{exc}} = 280$  nm) in the absence (red trace) and presence (black traces) of increasing concentrations of potassium iodide (KI). Inset: The Stern–Volmer plot resulted in a quenching rate constant of  $(10 \pm 1) \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$ . (b) Steady-state-fluorescence spectra of NAla-(GS)<sub>6</sub>-Dbo in the absence (red trace) and presence (black trace) of KI. Note that the Dbo emission near 440 nm decreases even more rapidly with KI concentration than the Trp emission near 340 nm, owing mainly due to the longer fluorescence lifetimes of the acceptor chromophore. The Dbo emission was not employed in the present analysis.



**Figure 3.** Donor fluorescence decay traces ( $\lambda_{\text{exc}} = 280$  nm,  $\lambda_{\text{obs}} = 350$  nm) for (a) donor-only and (b) donor–acceptor labeled (GS)<sub>6</sub> peptide in the absence (red traces) and presence (black traces) of increasing concentrations of KI. The inset in (a) shows the kinetic plot for the determination of the quenching rate constant ( $(9 \pm 1) \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$ ), and that in (b) shows the (monoexponential) decays on a logarithmic intensity scale.

enhances FRET significantly, while any movement apart has a comparably smaller effect; this is due to the characteristic  $R^6$  dependence of FRET.<sup>27,66</sup> Consequently, the  $R_{\text{eff}}$  value determined by FRET in the presence of donor–acceptor diffusion is always shorter than it would be in its absence. However, this also means that variations of the  $R_{\text{eff}}$  value as they occur when an experimental parameter is systematically varied with the initial distribution remaining unchanged are diagnostic for variations in the contribution of diffusion to FRET, i.e., the FDE.

If the proposition was correct that the FDE can be reduced through donor lifetime reduction by an external quencher, the effective donor–acceptor distance should increase with quencher concentration.

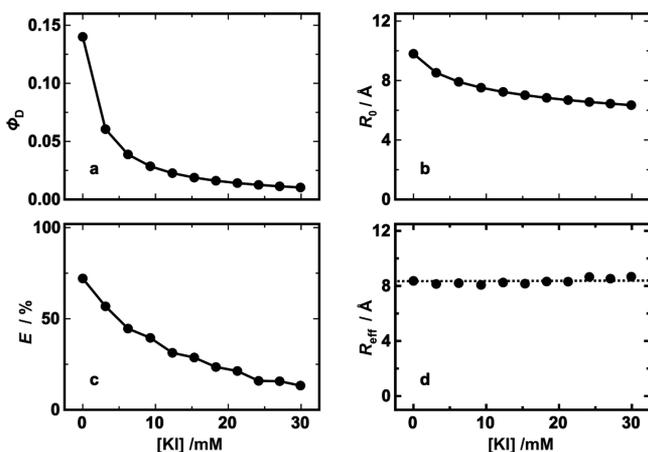
**Effect of External Quencher on Steady-State and Time-Resolved Fluorescence.** We found potassium iodide (KI) to be a highly effective quencher of NAla fluorescence. The required small concentrations (0–30 mM) posed no risks of chain perturbation due to salt effects, which, in the absence of specific binding, is known to occur only at high millimolar or molar concentrations of additives.<sup>67–69</sup>

Upon going from 0 to 30 mM KI, the NAla steady-state fluorescence of the donor-only labeled peptide dropped by more than 1 order of magnitude (factor 12.6, Figure 2a). In the time-resolved measurements, the lifetime of the donor in the donor-only labeled peptide decreased from 35.8 ns (Figure 3a, red trace) to 3.31 ns (lowest black trace). The KI quenching rate constant obtained from the corresponding Stern–Volmer or kinetic plots (insets in Figures 2a and 3a) revealed iodide as a diffusion-controlled quencher of donor fluorescence. The effect of external quencher on the steady-state- and time-resolved-fluorescence emission from the donor in the donor–

acceptor labeled peptide (Figure 2b and 3b) was relatively smaller (e.g., from 10.3 to 2.87 ns) due to the intrinsically shorter fluorescence lifetimes in the presence of FRET.

All donor emission decay traces followed monoexponential decay kinetics. This was expected for the donor-only peptide as NAla was specifically chosen for its simple monoexponential decay kinetics, but it was surprising for the double-labeled peptide. As the FRET rate is distance dependent, the kinetics are always expected to be multiexponential in the absence of mutual donor–acceptor diffusion (eq 11). Thus, the observed monoexponentiality of the kinetics provided a first hint for fast diffusion: Short-distance chains (i.e., those with a short donor–acceptor or end-to-end separation) are deactivated faster than long-distance chains but they are replenished sufficiently fast, such that the excited-state distribution decreases in height but retains its shape. A time-invariant shape of the distribution means that the entire excited-state ensemble decays with a single observed rate constant even if the individual chain conformations decay with individual distance-dependent “local” rate constants  $k_{\text{T}}(r)$ .<sup>70</sup>

**The Apparent Donor–Acceptor Distance Is Independent of External Quenching.** To evaluate the apparent donor–acceptor distance,  $R_{\text{eff}}$  at each KI concentration, the Förster radius was determined at each quencher concentration, because it depends (and decreases) with the donor fluorescence quantum yield,  $\Phi_{\text{D}}$  (eq 3). Going from 0 to 30 mM KI, the quantum yield of NAla measured for NAla-(GS)<sub>6</sub> and calculated from the steady-state-fluorescence spectra relative to *N*-acetyltryptophanamide as reference ( $\Phi = 0.14$ , pH 7.0)<sup>39,50,51</sup> decreased from 14 to 1.1% (Figure 4a), affecting a marked decrease in the Förster radius from 9.8 to 6.3 Å (Figure 4b). Note that the overlap integral,  $J$ , and the refractive index,  $n$ ,



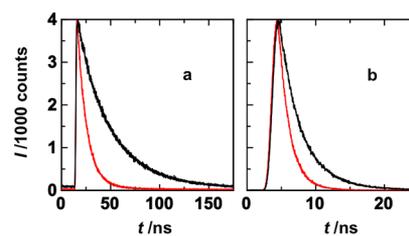
**Figure 4.** (a) Donor quantum yield, (b) Förster radius, (c) ETE, and (d) recovered effective donor–acceptor distance plotted against KI concentration for the NAla-(GS)<sub>6</sub>-Dbo peptide. (d) The dotted horizontal line at 8.3 Å is drawn at the average of the  $R_{eff}$  values.

were independent of the low millimolar KI concentration. The ETE values were subsequently determined from the donor-only and donor–acceptor fluorescence decay kinetics (Figure 3) via eq 6 by using the  $R_{DA}$  values from eq 5 (Figure 4c). It is important to note that the fluorescence decays of the donor–acceptor labeled peptide remained monoexponential even in the presence of quencher.

The presence of quencher—to our great astonishment as experimentalists—led not to the slightest modification of the recovered effective donor–acceptor distance. It remained at about 8.3 Å (Figure 4d), regardless of the dramatic shortening in donor lifetime by more than 1 order of magnitude, which should intuitively limit “the time for diffusion” and therefore the probability to enhance FRET by diffusion, which in turn should result in longer effective distances as quencher is added. The experimental choices we made had permitted highly precise measurements; the conclusion is inevitable that an external quencher that reduces the donor lifetime does not reduce the impact of diffusion on FRET. If it would, the  $R_{eff}$  value should have increased as quencher was added. To support this last assumption, we carried out two control experiments, again inspired by the random-walk equation (eq 1), which suggests that the effective distance should increase with either a decreasing diffusion coefficient or with a shorter-lived donor.

**Diffusion-Enhanced FRET in Dependence on Solvent Viscosity and Donor Lifetime.** Gly–Ser peptides including (GS)<sub>6</sub> display no internal friction;<sup>60,71</sup> their end-to-end diffusion coefficient is expected to be directly proportional to the reciprocal macroscopic solvent viscosity,  $D \propto \eta^{-1}$ , given that the macroviscosity and the relevant microviscosity of the solutions are identical.<sup>38,46,60,61</sup> To demonstrate the influence of viscosity, we measured the  $R_{eff}$  value for NAla-(GS)<sub>6</sub>-Dbo in water and in 90% (v/v) ethylene glycol, that is, at a 12.7-fold higher viscosity;<sup>46,47</sup> we observed an increase in  $R_{eff}$  value by about 3 Å (Figure 5, Table 1). This increase is as expected when diffusion is limited. Of course, it cannot be simply ruled out that substituting water for an aqueous solution of a viscogen modifies not only chain dynamics but also structure, as we previously observed for polyproline peptides dissolved in propylene glycol.<sup>41,42</sup>

Conversely, to demonstrate the influence of the donor lifetime, we compared NAla-(GS)<sub>6</sub>-Dbo and FTTrp-(GS)<sub>6</sub>-Dbo.



**Figure 5.** Donor fluorescence decay traces of (a) NAla-(GS)<sub>6</sub>-Dbo ( $\lambda_{exc} = 280$  nm,  $\lambda_{obs} = 350$  nm) and (b) FTTrp-(GS)<sub>6</sub>-Dbo ( $\lambda_{exc} = 280$  nm,  $\lambda_{obs} = 350$  nm) in water (red traces) and in 90% (v/v) ethylene glycol (black traces).

The isosteric donors, NAla and FTTrp, are not expected to have a significantly different influence on the structure and dynamics of the common Donor-(GS)<sub>6</sub>-Dbo peptide chain. When the lifetime was decreased about 18-fold, from 35.8 ns for NAla to 1.96 ns for FTTrp,  $R_{eff}$  increased again by about 3 Å. Thus, the variation of both parameters by more than 1 order of magnitude, donor lifetime and diffusion coefficient (viscosity), yielded the expected result: The apparent distance increased significantly (by ca. 3 Å).<sup>72</sup>

This result contradicted the absence of an observed effect of quencher: If the shortening of the fluorescence lifetime by more than 1 order of magnitude (from NAla to FTTrp) has a sizable effect if different donors are utilized, why is the same effect not detectable if the fluorescence lifetime of NAla itself is varied by more than 1 order of magnitude through addition of quencher? To answer this question, we decided to couple the tools of elementary FRET analysis to simulations based on the Haas–Steinberg equation.

**Diffusion-Enhanced FRET Studied by Coupling Elementary and Haas–Steinberg Analysis.** With one notable exception,<sup>37</sup> the FDE has not been studied on the level of elementary FRET analysis before. To understand, corroborate, and generalize the experimental results, we needed to link the traditional evaluation of FRET measurements to the more complex analysis based on the HSE (eq 9).<sup>1,2,27,29,73</sup> We did that with three questions in mind: (i) Does the HSE predict a symmetrical dependence of the FDE on the donor lifetime and the donor–acceptor diffusion coefficient as demanded by eq 1? (ii) Which donor lifetime should be used in eq 1 if it is not the experimental one? (iii) Can the HSE explain the inefficiency of external quenchers in reducing the FDE?

The HSE relates the rate of donor deactivation to the sum of three terms accounting for the donor decay in the absence of FRET, for the donor decay caused by FRET, and for diffusion.

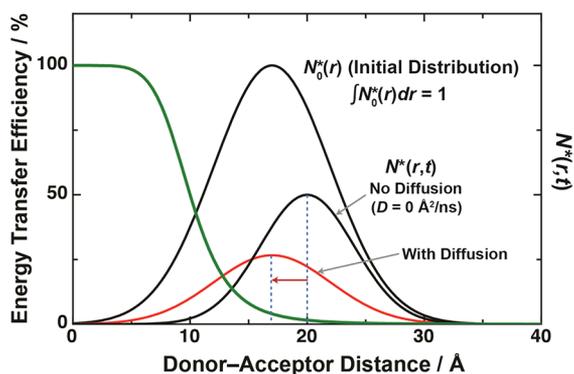
$$\frac{\partial N^*(r, t)}{\partial t} = -k_D N^*(r, t) - k_D \frac{R_0^6}{r^6} N^*(r, t) + \frac{\partial}{\partial r} \left( N_0^*(r) D \frac{\partial (N^*(r, t) / N_0^*(r))}{\partial r} \right) \quad (9)$$

A visual illustration of the HSE, first derived in ref 27, is given in Figure 6, which shows an exemplary distance distribution,  $N_0^*(r)$  or  $N_0^*(r, t=0)$ , instantly after short-pulse donor excitation. Because the probability of donor excitation should be independent of the conformation of the peptide chain (and because there is no indication for exciplexes between the selected donors and the acceptor), the initial distance probability distribution of excited states,  $N_0^*(r)$ , normalized to  $\int N_0^*(r) dr = 1$ , is identical to the probability

Table 1. FRET Analysis of Donor-(GS)<sub>6</sub>-Dbo Peptides

donor	solvent	$\Phi_D$	$\tau_D$ /ns	$\tau_0$ /ns	$\tau_{DA}$ /ns	$R_0/\text{\AA}$	$E^a$ /%	$R_{\text{eff}}^b/\text{\AA}$
NAla	water	0.14	35.8	256	10.3	9.8	70.5	8.4
	ethylene glycol <sup>c</sup>	0.20	51.9	260	33.1	10.2	36.2	11.2
FTrp	water	0.10	1.96	19.6	1.50	9.6	24.8	11.7
	ethylene glycol <sup>c</sup>	0.23	3.68	16.0	2.96 <sup>d</sup>	11.1	19.7	14.1

<sup>a</sup>Energy transfer efficiency from time-resolved-fluorescence measurements (eq 6). <sup>b</sup>Effective distance determined from  $E$  and  $R_0$  (cf. eq 5). <sup>c</sup>With 90% (v/v) ethylene glycol, pH 5.0, 25 °C. <sup>d</sup> $\tau_{DA}$  is an amplitude-weighted lifetime obtained from  $\tau_{DA} = \alpha_1\tau_1 + \alpha_2\tau_2 = 0.62 \cdot 2.26 \text{ ns} + 0.38 \cdot 4.11 \text{ ns}$ .



**Figure 6.** Diffusion-enhanced FRET. The initial ( $t = 0$ , solid black line) distance distribution of chains with excited donor is identical to the ground-state probability distance distribution in the equilibrium chain ensemble,  $N_0^*(r) = N^*(r, t=0) = p(r)$  with  $\int p(r) dr = 1$ . After excitation, chains with proximate donor and acceptor are rapidly deactivated and the distribution is down-sized and continuously shifted to larger distances (dotted line). However, if the labeled sites can approach each other by mutual diffusion, short-distance chains are continuously replenished (red line) such that the observed ETE,  $E_{\text{obs}}$ , is raised. The green line marks the local ETE,  $E(r)$ , for a particular distance.

distribution in the ground-state equilibrium ensemble of chains, i.e.,  $N_0^*(r) = p(r)$  (Figure 6, solid black line). Subsequently, more chains with a short donor–acceptor distance undergo FRET than long-distance chains, such that the distribution  $N^*(r, t)$ , is shifted toward longer distances (Figure 6, dotted line). Mutual donor–acceptor motion counteracts this trend by constantly replenishing short-distance chains (Figure 6, red line). From the HSE, the donor decay kinetics can be generated for any distribution, diffusion coefficient, and donor lifetime.

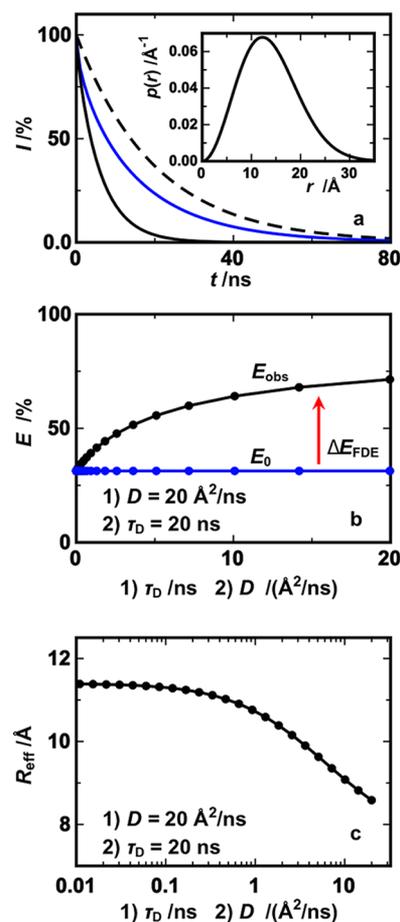
How can we quantify the FDE? Central to traditional FRET analysis is the ETE as the fraction of donor deactivation events caused by FRET. The ETE can be viewed as being composed of a contribution,  $E_0$ , from the initial distance distribution as it would develop in the absence of diffusion and of a contribution from diffusion that yields the additional efficiency,  $\Delta E_{\text{FDE}}$  (eq 10).

$$E_{\text{obs}} = E_0 + \Delta E_{\text{FDE}} \quad (10)$$

To obtain these quantities from the HSE, one needs to generate and evaluate the following kinetic traces of donor fluorescence intensity (Figure 7 a).

(i) The donor decay in the absence of FRET, in the donor-only peptide (Figure 7 a, black dashed curve): The associated fluorescence intensity,  $I_D(t)$ , follows eq 11. The emitted fluorescence is given by the area under the curve,  $A_D$ . The initial intensity is set to 100%, such that the values of the area and donor lifetime become identical:  $A_D = k_D^{-1} = \tau_D$ .

$$I_D(t) = I_0 \exp(-k_D t) \quad (11)$$



**Figure 7.** HSE simulations (at constant donor quantum yield). (a) Simulated decay traces of donor fluorescence in the donor-only peptide (dashed line), the donor–acceptor peptide in the absence of diffusion,  $D = 0$  (blue), and the donor–acceptor peptide in the presence of diffusion with  $D$  set to  $20 \text{ \AA}^2/\text{ns}$  (black). Further parameters were  $\tau_{DA} = 20 \text{ ns}$ ,  $R_0 = 10 \text{ \AA}$ , and the minimal and maximal distances of approach set to 4 and 50 Å, respectively. The areas under the curves were obtained by numerical integration and used to determine the ETE values. Inset: The initial (equilibrium) probability density distribution in an ideal Gaussian chain (eq 14) with  $\langle (r^2) \rangle^{1/2}$  set to 15 Å. (b) (1) The observed ETE,  $E_{\text{obs}}$  (black), and the ETE in the absence of diffusion,  $E_0$  (blue) were plotted against the donor lifetime, while  $D$  was kept fixed at  $20 \text{ \AA}^2/\text{ns}$ . (2)  $E_{\text{obs}}$  and  $E_0$  were plotted against  $D$  while the lifetime was kept fixed at 20 ns. The ETE courses in both simulations coincide. (c) The effective donor–acceptor distance calculated from  $E_{\text{obs}}$  (eq 5) plotted (1) against the donor lifetime when the diffusion coefficient was fixed at  $20 \text{ \AA}^2/\text{ns}$  and (2) against the diffusion coefficient when the donor lifetime was set to 20 ns. Both courses coincide.

(ii) The donor decay in the presence of FRET, in the donor–acceptor peptide, but in the absence of diffusion: Its intensity,  $I_{DA0}$ , is described by a closed expression (eq 12),

where  $p(r) = N_0^*(r)$ . Equation 12 is a consequence of the relations  $k_{\text{DA}}(r) = k_{\text{D}} + k_{\text{F}}(r)$  and  $k_{\text{F}}(r) = k_{\text{D}}(R_0/r)^6$  (eq 2). An identical decay can be simulated by setting the diffusion coefficient in the HSE to zero. The area under the curve,  $A_{\text{DA0}}$ , measures the fluorescence emitted in the absence of diffusion.

$$I_{\text{DA0}}(t) = I_0 \int \exp(-k_{\text{D}}t - k_{\text{D}}(R_0/r)^6 t) p(r) dr \quad (12)$$

(iii) The donor decay in the presence of FRET and diffusion as generated through the HSE: With the distribution  $N^*(r,t)$  being determined, the decay follows eq 13 and yields the fluorescence emitted in the presence of diffusion,  $A_{\text{DAD}}$ .

$$I_{\text{DA}}(t) = I_0 \int N^*(r, t) dr \quad (13)$$

The efficiencies  $E_{\text{obs}}$ ,  $E_0$ , and  $\Delta E_{\text{FDE}}$  are then determined from the simulated areas via  $E_{\text{obs}} = (A_{\text{D}} - A_{\text{DAD}})/A_{\text{D}}$ ,  $E_0 = (A_{\text{D}} - A_{\text{DA0}})/A_{\text{D}}$ , and  $\Delta E_{\text{FDE}} = (A_{\text{DA0}} - A_{\text{DAD}})/A_{\text{D}}$  (cf. eq 7).

To test whether  $E_{\text{obs}}$  and  $\Delta E_{\text{FDE}}$  depend symmetrically on donor lifetime and diffusion coefficient, we assumed a constant quantum yield and ran two series of simulations. In the first series, the diffusion coefficient was kept fixed at  $20 \text{ \AA}^2/\text{ns}$  and the donor lifetime was varied from 0.1 to 20 ns to cover a representative range.<sup>54,65</sup> In the second series, the donor lifetime was held constant at 20 ns and the diffusion coefficient was varied from 0.1 and  $20 \text{ \AA}^2/\text{ns}$ , a range which encompasses published diffusion coefficients.<sup>1,2,27,74–76</sup> In the case of a symmetrical dependence, plots of  $E_{\text{obs}}$  versus lifetime and versus diffusion coefficient would coincide. The HSE simulations used the model of an ideal Gaussian chain for the distribution (eq 14) with a root-mean-square distance of  $15 \text{ \AA}$  (Figure 4a, inset). In subsequent controls, two-parameter distributions were also employed: In the wormlike chain model (eq 15),<sup>77</sup> the parameters  $l_{\text{p}}$  and  $l_{\text{c}}$  are the persistence and contour length of the chain; in the widely used skewed Gaussian distribution (eq 16)  $a$  and  $b$  determine the width and position of the distribution, whereas  $c$  is determined from the normalization condition.

$$p(r) = 4\pi r^2 \left( \frac{3}{2\pi \langle r^2 \rangle} \right)^{3/2} \exp\left( -\frac{3}{2} \frac{r^2}{\langle r^2 \rangle} \right) \quad (14)$$

$$p(r) = \frac{4\pi r^2 N}{l_{\text{c}}^2 (1 - (r/l_{\text{c}})^2)^{9/2}} \exp\left( \frac{-3l_{\text{c}}}{4l_{\text{p}} ((1 - (r/l_{\text{c}})^2)^2)} \right) \quad (15)$$

$$p(r) = \pi r^2 c \exp(-a(r - b)^2) \quad (16)$$

Indeed, the ETs,  $E_{\text{obs}}$ , and  $\Delta E_{\text{FDE}}$  ( $=E_{\text{obs}} - E_0$ ) grew in an identical fashion with the diffusion coefficient as well as with the donor lifetime (Figure 7b). Thus, the HSE predicts the diffusion enhancement to depend symmetrically on diffusion coefficient and donor lifetime, that is, to be a function of their product. This symmetry depended neither on the choice of the distance-distribution model (eqs 14–16) nor on the specific parameter choices. The results also confirmed that  $E_0$  was suitably defined and robustly reproduced by the programmed software: It depends neither on the donor lifetime nor on the diffusion coefficient but was constant throughout the simulations (Figure 7b, blue trace). However, does  $E_0$  really measure the information from the equilibrium distance distribution, even under diffusion? This is certainly not the case as this piece of information is gradually lost with increasing

diffusion. Later we will introduce a more suitable parameter, fully within the spirit of traditional FRET analysis.

In the two simulation series, the effective distance grew with a decreasing diffusion coefficient in the same manner as with a decreasing donor lifetime (Figure 7c), which confirms that  $R_{\text{eff}}$  reflects the FDE. Throughout the simulations, the Förster distance was held constant at  $10.0 \text{ \AA}$ . As  $R_0$  depends on the donor quantum yield (eq 2), the simulations cover only the case of a *constant quantum yield*, that is, of intrinsic donor-lifetime variations as they occur for two chromophores with different fluorescence lifetimes but the same quantum yield. Qualitatively, the simulations agree with the intuitive expectation and the experimental results in that both, a donor-lifetime reduction and a raised solvent viscosity, lead to an increased effective distance. We had to explain why the same does not hold when the donor lifetime is reduced by addition of an external quencher. In attempting to do so, it is important to realize that an external quencher reduces the fluorescence lifetime to the same extent that it reduces the fluorescence quantum yield.

**The Radiative Donor Lifetime Dominates the FDE.** The measured donor lifetime in the donor-only peptide,  $\tau_{\text{D}}$ , is not simply the lifetime that controls the FDE. The lifetime  $\tau_{\text{D}}$ , or its reciprocal quantity,  $k_{\text{D}}$ , can be dissected into two components (eq 17): The radiative decay rate  $k_{\text{rad}}$  is the rate of donor emission when the donor quantum yield,  $\Phi_{\text{D}}$ , equals unity, that is, when each photon absorbed by the donor leads to one being emitted. The rate  $k_{\text{nrad}}$  captures nonradiative decay processes including quenching by a coagent. Addition of an external quencher raises only  $k_{\text{nrad}}$  and reduces, by that, the donor quantum yield defined by eq 18. What the experiments suggest, therefore, is that the nonradiative decay rate has no or little influence on the FDE and that it is the radiative lifetime,  $\tau_0 = k_{\text{rad}}^{-1}$ , that has to be changed to alter the FDE and the effective distance. In the context of FRET, it is, in fact, well-known and has often been emphasized that  $k_{\text{T}}$  only depends on  $\tau_0$ ,<sup>22,78</sup> but in the context of FDE, where a correlation between FDE and  $\tau_{\text{D}}$  seems intuitive, this conclusion is more difficult to reach.

$$k_{\text{D}} = k_{\text{rad}} + k_{\text{nrad}} \quad (17)$$

$$\Phi_{\text{D}} = k_{\text{rad}} / (k_{\text{rad}} + k_{\text{nrad}}) \quad (18)$$

$$k_{\text{rad}} = \Phi_{\text{D}} k_{\text{D}} \quad (19)$$

$$\tau_{\text{D}} = \Phi_{\text{D}} \tau_0 \quad (20)$$

An impact of the nonradiative decay rate or of the quantum yield on the FDE could be expected if the rate of FRET at a specific distance would depend on the quantum yield. At first glance, this appears to be the case because  $k_{\text{T}}$  depends on  $k_{\text{D}}$  (eq 2,  $k_{\text{T}} = k_{\text{D}}(R_0/R_{\text{DA}})^6$ ), which itself depends on  $\Phi_{\text{D}}$  (cf. eq 19,  $k_{\text{D}} = k_{\text{rad}}/\Phi_{\text{D}}$ ). Yet, according to the fundamental but less frequently employed equation that Förster derived (eq 21),<sup>78</sup> the distance-dependent FRET rate is independent of  $\Phi_{\text{D}}$ . In the next step, we needed to confirm by HSE simulations that the same held for the diffusion-enhanced FRET rate. At first, we determined what had to be done differently in this set of simulations than in the previous one (Figure 7): Collecting all constants in eq 21 into one,  $cJ\kappa^2/n^4 = R_{\text{F}}^6$ , yields eq 22.

$$k_{\text{T}} = (cJ\kappa^2/n^4)(k_{\text{rad}}/R_{\text{DA}})^6 \quad (21)$$

$$k_{\text{T}} = k_{\text{rad}} R_{\text{F}}^6 / R_{\text{DA}}^6 \quad (22)$$

Thus,  $R_F$  is the donor–acceptor distance ( $R_{DA} = R_F$ ) at which the donor emits a photon with an equal probability as it transfers its excitation energy to the acceptor,  $k_T = k_{rad}$ . The distance  $R_F$  equals the Förster radius only at a quantum yield of unity. In contrast, the Förster radius  $R_0$  is the distance at which donor deactivation by FRET is as likely as donor deactivation by all alternative processes including photon emission but also collision-induced quenching. The condition that  $k_T = k_D$  when  $R_{DA} = R_0$  allows rewriting eq 21 as the familiar eq 2,  $k_T = k_D R_0^6 / R_{DA}^6$ . By using eq 19, we can rewrite eq 22 as  $k_T = \Phi_D k_D R_F^6 / R_{DA}^6$  to obtain the relationship between  $R_0$  and  $R_F$  (eq 23). It follows that, although both the Förster radius and the measured donor lifetime depend on the quantum yield,  $k_T$  does not because the product  $k_D R_0^6$  in eq 2 remains constant as the quantum yield is changed:  $k_D R_0^6 = k_{rad} \Phi_D^{-1} \Phi_D R_F^6 = k_{rad} R_F^6$  (eq 24; cf. eqs 19 and 23).

$$R_0^6 = \Phi_D R_F^6 \quad (23)$$

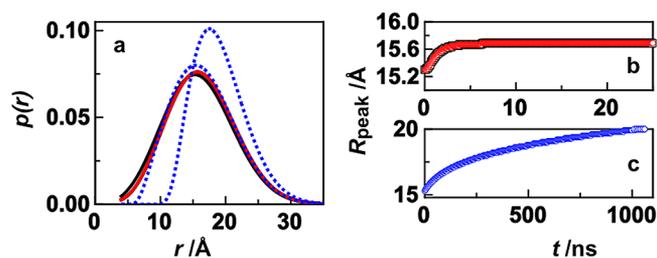
$$k_D R_0^6 = k_{rad} R_F^6 \quad (24)$$

In the foregoing series of simulations, we left the Förster radius unchanged while we varied the lifetime. Thus, in essence, we varied the radiative lifetime and learned that the FRET diffusion enhancement depends symmetrically on the radiative lifetime and the diffusion coefficient. However, when we simulate the effect of external quenching by changing the quantum yield and, by that, the donor lifetime, we must keep the product  $k_D R_0^6$  constant (eq 24). The HSE can be appropriately rewritten for this set of simulations (eq 25).

$$\frac{\partial N^*(r, t)}{\partial t} = -\frac{k_{rad}}{\Phi_D} N^*(r, t) - \frac{k_{rad} R_F^6}{r^6} N^*(r, t) + \frac{\partial}{\partial r} \left( N_0^*(r) D \frac{\partial N^*(r, t) / N_0^*(r)}{\partial r} \right) \quad (25)$$

**HSE and Traditional FRET Analysis of the FDE Quantum-Yield Dependence.** We set the radiative donor lifetime,  $\tau_0 = k_{rad}^{-1}$ , to 250 ns and the radiative Förster radius,  $R_F$ , to 14 Å to mimic the radiative lifetime of NAla and the radiative Förster radius of the NAla/Dbo donor/acceptor pair (13.7 Å). The quantum yield was varied from 1 to 0.01, thereby mimicking a donor lifetime variation ( $\tau_D = k_D^{-1}$ ) from 250 to 2.5 ns. Further simulation details are described in the legend of Figure 8.

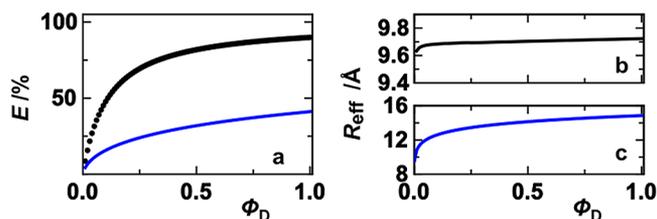
For the selected distance distribution (Figure 8a, black line), a donor–acceptor pair with a Förster radius of 14 Å would be a proper experimental choice. In the absence of diffusion (Figure 8a, blue dotted lines), the probability density shifts significantly with time, as expected. In contrast, in the presence of diffusion, the initial probability distribution rapidly develops into a stationary distribution. This is seen when the distributions  $N^*(r, t)$  obtained after 5, 100, and 200 ns are normalized and compared. The obtained  $p(r, t)$  profiles coincide well (Figure 8a, red line). It is remarkable how fast the stationary state is reached and how closely the ground-state equilibrium and the stationary excited-state distribution resemble each other. When the distance at the distribution peak is plotted against time, a stable plateau is reached after about 2 ns (Figure 8b). The same courses are obtained for quantum yields of 1.0 and 0.1, as only quenching by FRET is distance-dependent and able to alter the shape of the distribution. When diffusion is absent, a plateau is



**Figure 8.** HSE results for  $\tau_0 = 250$  ns,  $R_F = 14$  Å, and a skewed Gaussian distance distribution (eq 16) with  $a = 0.0123$  Å<sup>-2</sup> and  $b = 10$  Å. (a) Black line: equilibrium probability distribution. Red line: probability distance distribution with diffusion ( $D = 20$  Å<sup>2</sup>/ns) after 5, 100, and 200 ns. The curves virtually coincide. Also shown are the probability distance distributions in the absence of diffusion after 5 ns (left-hand dotted blue line) and 100 ns (right-hand line). Note that, due to the normalization utilized in this graph ( $\int p(r) dr = 1$ ), the narrower distribution is higher. (b) Distance of the distribution peak in the presence of diffusion ( $D = 20$  Å<sup>2</sup>/ns) plotted against time for  $\Phi_D = 1$  (red) and 0.1 (black). The courses coincide. (c) Time course of the distribution-peak distance in the absence of diffusion.

never reached (Figure 8c), and the distribution continues to shift to higher distances.

With quenching, i.e., decreasing quantum yield and donor lifetime, the efficiencies,  $E_{obs}$  and  $E_0$ , decreased and approached zero (Figure 9a), whereas the effective distance calculated from



**Figure 9.** (a) Observed ETE,  $E_{obs}$  (black), in the presence of diffusion ( $D = 20$  Å<sup>2</sup>/ns), and ETE in the absence of diffusion,  $E_0$  (blue), plotted against the quantum yield. Quantum-yield dependence of the effective distance (b) in the presence of diffusion ( $D = 20$  Å<sup>2</sup>/ns) and (c) in the absence of diffusion (blue), as calculated from (b)  $E_{obs}$  and (c)  $E_0$  (cf. eq 5), and  $R_0 = \Phi_D^{1/6} R_F$ .

$E_{obs}$  (cf. eq 5) and from  $R_0$  ( $R_0 = \Phi_D^{1/6} R_F$ ; cf. eq 23) remained virtually constant (Figure 9b). Only at very small values of the quantum yield was a minimal decrease observed (Figure 9b). Recall that a successful minimization of the FRET diffusion enhancement should result, according to the original intuitive anticipation, in an increase and not in a decrease of the effective distance. Thus, both experiment and simulation indicate that an external quencher cannot increase the effective distance in contrast to the two other strategies of FDE reduction that we tested—raising the viscosity and shortening the radiative lifetime by the choice of a different donor.

It is important to realize the consequences of how strongly the effective distance in the absence of diffusion decreased with decreasing quantum yield (Figure 9c, blue line). A multitude of studies have been aimed at extracting effective distances and distance trends from FRET ensemble measurements,<sup>10,14,40,79,80</sup> but none has taken into account that the quantum yield does not merely alter  $R_0$ , but that it indeed changes the measured effective distance. This conclusion reaches far: For FRET to yield reliable information on distance trends in an experimental series of polydisperse polymers or

biopolymers, particularly in the absence of diffusion, a constant quantum yield needs to be maintained, or corrections according to those outlined herein would need to be made. Note that the donor quantum yield can easily increase or decrease in an experiment when, for instance, a protein folds,<sup>10,14</sup> when the temperature is varied, or when a coadditive is added.<sup>40</sup> The  $R_{\text{eff}}$  trend might then be interpreted as a shortening of the monitored distance where, in reality, none is taking place. Note that, in reaching this result, we do not exclusively rely on HSE simulations; we obtain an identical quantum-yield dependence of the effective distance from a closed analytic expression derived in the Supporting Information.

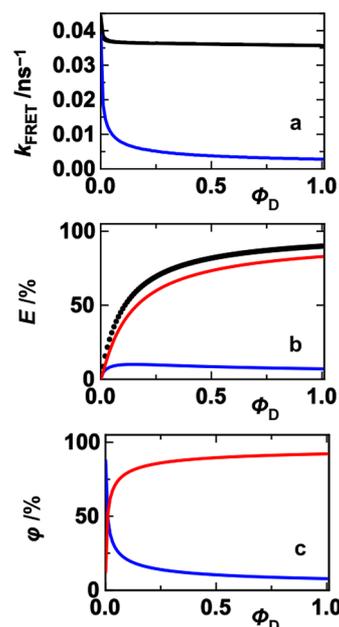
$$R_{\text{eff}} = \Phi_D^{1/6} R_F \left( 1 / \left( 1 - \int p(r) r^6 / (r^6 + \Phi_D R_F^6) dr \right) - 1 \right)^{1/6} \quad (26)$$

The most appropriate perspective to understand the results on not merely the mathematical but also the physical level is to analyze them in terms of the effective FRET rate constants,  $k_{\text{FRET}}$  and  $k_{\text{FRET}0}$ , in the presence and absence of diffusion. Earlier, we demonstrated that the Förster rate at a specific distance,  $k_T(r)$ , is independent of the quantum yield. However, the effective FRET rate obtained by measurement is defined through  $E_{\text{obs}}$  ( $k_{\text{FRET}} = k_D(1/E_{\text{obs}} - 1)^{-1}$ ; cf. eq 4); that is, it accounts for FRET at all distances accessible to the chain. It now turns out that the effective  $k_{\text{FRET}}$  value for the entire distribution (which is experimentally accessible) *does vary with the quantum yield*, not because the microscopic FRET rates (at a fixed distance) depend on the quantum yield, but because the quantum yield (quenching) dictates which distances contribute more highly to the effective FRET rate. This variation is more pronounced in the absence of diffusion.

Note Figure 10a.  $k_{\text{FRET}}$  remains almost constant with the quantum yield in the presence of diffusion. That it shares this behavior with the effective distance (cf. Figure 9b) is expected as the equation  $k_{\text{FRET}} = k_D(R_0/R_{\text{eff}})^6$  holds for the effective parameters. In the absence of diffusion, for  $k_{\text{FRET}0}$ , it is again possible to obtain an analytic expression for the quantum-yield dependence (eq 27) that can be derived, for instance, from the fluorescence of the donor–acceptor peptide as given by the area under the donor–acceptor decay,  $A_{\text{DA}0} = \iint \exp(-(k_D + k_{\text{FRET}})t) p(r) dr dt$  (cf. eqs 5 and 14). Two alternative derivations are given in the Supporting Information.

$$k_{\text{FRET}0} = k_{\text{rad}} \Phi_D^{-1} \left( \left( \int p(r) / (1 + \Phi_D R_F^6 / r^6) dr \right)^{-1} - 1 \right) \quad (27)$$

Thus, the dependence of  $k_{\text{FRET}}$  on the quantum yield is strongest in the absence of diffusion and vanishes with increasing diffusion. In the limit of very fast diffusion, the equilibrium or initial distance probability distribution is maintained at all times, the observed kinetic traces are monoexponential, and the effective FRET rate is equal to the distance average of the Förster rate,  $k_T$ .<sup>81</sup> One has  $k_{\text{FRET}} = \langle k_T \rangle = \int k_T(r) p(r) dr = \int k_D (R_0^6 / r^6) p(r) dr = k_{\text{rad}} R_F^6 \langle r^{-6} \rangle$ , i.e., an obvious independence of the quantum yield. We are close to this case in both experiments—where we obtained monoexponential kinetics for NAla-(GS)<sub>6</sub>-Dbo and the simulations (Figure 8a)—where a stationary distribution, hardly distinguishable from the equilibrium distribution, is quickly attained. More details are discussed in the Supporting Information.



**Figure 10.** (a) Observed FRET rate,  $k_{\text{FRET}}$  (black line), calculated from  $E_{\text{obs}}$  (cf. eq 4), and FRET rate for the static distribution,  $k_{\text{FRET}0}$  (blue), calculated from  $E_0$  (and  $R_0 = \Phi_D^{1/6} R_F$ ) plotted against the quantum yield. (b)  $E_{\text{obs}}$  (black line),  $E_{\text{EQ}}$  the fraction of donor deactivation events caused by FRET due to the equilibrium distance distribution (blue), and  $E_M$ , the fraction of deactivation events caused by FRET due to donor–acceptor motion (red), plotted against the donor quantum yield. (c) Fraction of FRET events,  $\phi_0 = E_{\text{EQ}}/E_{\text{obs}} = k_{\text{FRET}0}/k_{\text{FRET}}$ , caused by the initial distance distribution (blue) and the fraction of FRET events,  $\phi_M = E_M/E_{\text{obs}} = 1 - \phi_{\text{EQ}}$  caused by donor–acceptor motion (red) plotted against the donor quantum yield. Simulations in the presence of diffusion were performed with  $D = 20 \text{ \AA}^2/\text{ns}$ .

It is obvious that the decomposition of the ETE by  $E_{\text{obs}} = E_0 + \Delta E_{\text{FDE}}$  cannot help to quantify the FDE under quantum-yield variation. The ETE in the presence and absence of diffusion decreases strongly as the quantum yield decreases (Figure 9a) while with diffusion the effective distance and FRET rate are virtually constant (Figures 9b and 10a). To apprehend when and why a quantum-yield dependence of the effective parameters can occur and to quantify the information that a time-resolved FRET experiment can yield on the equilibrium distance distribution and on diffusion, we have to use the effective FRET rates,  $k_{\text{FRET}}$  and  $k_{\text{FRET}0}$ , to decompose the ETE.

The ETE is the fraction of donor deactivation events caused by FRET,  $E_{\text{obs}} = \Delta N_{\text{FRET}}^* / \Delta N^*$  or  $E_{\text{obs}} = k_{\text{FRET}} / (k_{\text{FRET}} + k_D)$ . We are interested in the fraction of donor-deactivation events that can be attributed to FRET from the initial (or the equilibrium) distance distribution,  $E_{\text{EQ}}$ , as well as to FRET from diffusional motion,  $E_M$  (read “ $E_{\text{Motion}}$ ”). We are specifically interested in the *fraction of FRET events* that can be attributed to the equilibrium distribution,  $\phi_{\text{EQ}}$  and to diffusion,  $\phi_M$ . With  $k_{\text{FRET}}$  and  $k_{\text{FRET}0}$  obtained from  $E_{\text{obs}}$  and  $E_0$ , we can define and determine these quantities (eqs 28–32).

$$E_{\text{obs}} = E_{\text{EQ}} + E_M \quad (28)$$

$$E_{\text{EQ}} = k_{\text{FRET}0} / (k_{\text{FRET}} + k_D) \quad (29)$$

$$E_M = (k_{\text{FRET}} - k_{\text{FRET}0}) / (k_{\text{FRET}} + k_D) \quad (30)$$

$$\varphi_{\text{EQ}} = k_{\text{FRET0}}/k_{\text{FRET}} \quad (31)$$

$$\varphi_{\text{M}} = (k_{\text{FRET}} - k_{\text{FRET0}})/k_{\text{FRET}} \quad (32)$$

It follows that  $\varphi_{\text{EQ}} = E_{\text{EQ}}/E_{\text{obs}}$ ,  $\varphi_{\text{M}} = E_{\text{M}}/E_{\text{obs}}$ , and  $\varphi_{\text{EQ}} + \varphi_{\text{M}} = 1$ .

As Figure 10b shows, the diffusion contribution,  $E_{\text{M}}$ , contributes the overwhelming part to the ETE and decreases in parallel to it while the quantum yield decreases. The FRET contribution of the equilibrium distribution,  $E_{\text{EQ}}$ , stays small and is much smaller than the ETE in the absence of diffusion,  $E_0$  (Figure 9a), despite a Förster radius of 14 Å that would appear to be an excellent experimental choice as it is close to the average distance of the distribution (Figure 8a). As  $E_{\text{EQ}}$  mirrors the absolute signal or information from the equilibrium distribution, it clearly transpires that suppressing diffusion might often be the only means to access the sought-after equilibrium distances. However, adding an external quencher, to dwell on this once more, leads only to a loss of information.

When the quantum yield is small, the activated donor has little chance to emit a photon and to contribute to the measured fluorescence signal. However, if the distance to the acceptor is too large, the donor's chance is also small to transfer its energy to the acceptor and to contribute to the measured difference between the emission of the donor–acceptor peptide and that of the donor-only peptide. Thus, the lower the quantum yield, the more chains of the equilibrium distribution become literally invisible; the higher the distance, the more chains are simply not monitored as if they were not present in the first place. This explains why, in the absence of diffusion, the effective FRET rate increases and the effective distance decreases with decreasing quantum yield, with both giving the impression of shorter distances present in the equilibrium distribution than is in reality the case.

At a very small quantum yield, it is only at a very short distance where the FRET rate is sufficiently high to compete with the rate of external quenching. However, when the distance is so short, FRET occurs before the donor and acceptor can considerably change their distance. This explains the observed variation of  $\varphi_{\text{EQ}}$  and  $\varphi_{\text{M}}$  (Figure 10c). The fraction of FRET events caused by the equilibrium distribution,  $\varphi_{\text{EQ}}$ , tends to go toward 100% as the quantum yield approaches zero. In this extreme, all FRET events that still take place have to be attributed to the short-distance tail of the initial ( $t = 0$ ) distribution. The fraction of FRET events due to diffusion,  $\varphi_{\text{M}}$ , decreases first very weakly with decreasing quantum yield to drop dramatically at very low values ( $\varphi_{\text{EQ}} + \varphi_{\text{M}} = 1$ ).

Are there any circumstances under which FRET measurements at a reduced quantum yield (high quencher concentration) can add new information to that obtained from a measurement at a high quantum yield? For time-resolved-fluorescence measurements, this is likely not the case. The kinetic trace of the donor–acceptor peptide starts with an initial steep decline that is caused by the short-distance tail of the distribution, where FRET is fastest. With increasing quencher concentration, any “later” signal from the bulk of the distribution is gradually eliminated. Thus, quantum-yield variation by adding an external quencher cannot add new information but is merely equivalent to evaluating only the beginnings of the kinetic traces obtained at a high quantum yield. A global analysis of the decay traces obtained at successively increasing quencher concentration can only distort the result as it puts an increasing weight on the information that

stems from the short-distance parts of the equilibrium distribution. This approach is therefore inferior to simply repeating the measurements at the highest quantum yield to so benefit from an increased signal-to-noise ratio of the averaged kinetics. The target to modulate the FDE of the whole equilibrium distribution, which is a precondition for a meaningful global analysis, cannot be attained through an external quencher. Of course, quantum-yield variations can be of benefit for other spectroscopic methods aimed at characterizing the dynamics of a polymer chain, e.g., when they are based on excimer formation instead of FRET.<sup>82</sup>

## DISCUSSION

**Diffusion-Enhanced FRET Is Independent of the Donor Quantum Yield.** The random-walk equation captures the intuition that the impact of internal chain motions on FRET could be minimized by minimizing the donor lifetime through an external quencher. Lakowicz and co-workers predicted that “...collisional quenching can be used to decrease the decay time of the donor. Under these conditions of shortened donor lifetime, there is less time for diffusion, and the data contain more information on the initial  $t = 0$  distance distribution.”<sup>34</sup> This would indeed be the case if the donor–acceptor displacement that can occur during the donor lifetime,  $\tau_{\text{D}}$  ( $\Delta R_{\text{rms}} = (6D\tau_{\text{D}})^{1/2}$ , eq 1), would be a proper measure of the FRET diffusion enhancement. As it is otherwise hard to escape from this thinking pattern—similar reasoning has found its way into a fundamental textbook<sup>54</sup>—we decided to demonstrate its incorrectness not merely by theoretical arguments but also by experiment and simulation. For further discussion, however, we formulate a simple analogy that can convey why external quenching can hardly affect the FDE: The effect of adding a quencher to the measurement solutions of the donor-only and donor–acceptor labeled peptides is identical to the effect that it would have when the solutions would be continuously diluted during measurement. The donor deactivated by a quencher or removed by dilution neither can emit a photon (to so contribute to the measured fluorescence signal) nor can it transfer its excitation energy to the acceptor (to so contribute to the signal difference between single- and double-labeled peptides). The quencher reduces the observed ETE but affects the FDE defined by  $\varphi_{\text{M}} = E_{\text{M}}/E_{\text{obs}}$  only weakly and only by eliminating the information from the bulk of the distribution. Reducing the donor lifetime by reducing the donor quantum yield merely compromises the signal-to-noise ratio. At this stage, these conclusions might appear almost self-evident, but they imply surprising and important consequences.

Lakowicz and co-workers studied the dynamics and structure of the chain (CH<sub>2</sub>)<sub>22</sub> labeled by the FRET donor–acceptor pair tryptamine (TMA) and dansyl.<sup>34</sup> The Förster radius decreased from 24.9 Å in the absence to 20.1 Å in the presence of 9.3 mM acrylamide (cf. eq 20). The time-resolved-fluorescence responses of the donor-only and donor–acceptor chains in the absence and presence of quencher were globally fitted to the HSE used with a Gaussian distribution model. The recovered diffusion coefficient (126 Å<sup>2</sup>/ns) was about half that of the free probes (264 Å<sup>2</sup>/ns), and the average distance and full width at half-maximum were determined to be 18.9 Å and 17.1 Å. Whereas the analysis of the unquenched samples could not yield such a defined set of values, a remarkable resolution enhancement was achieved when the results from the quenched and unquenched samples were simultaneously fitted. This observation was attributed to the lower impact of

diffusion on FRET and to the higher impact of the initial distance distribution in the presence of quencher. As we demonstrated by experiment, by simulation, and by theoretical arguments, this intuitive reasoning could not have been the cause for the observed resolution enhancement: The data in the presence of quencher do not add new information. The problem—expressed as mathematical analogy—remains that two unknowns cannot be determined from a single equation and also not from a multitude of linearly dependent equations that all contain the same information. The apparent resolution enhancement was more likely a consequence of an (unintended) increased weighing of the information from the short-distance parts of the distribution.

### The Random-Walk Perspective Revisited and Applied to Global Analysis.

In other cases, the conditions for a successful global analysis are only apparently met. In a recent investigation on how the denaturant guanidinium chloride (GdmCl) influences the peptidic chain (GS)<sub>16</sub>, the two donors naphthylalanine (NALa) and pyrene (Py) were alternately used in combination with dansyl as acceptor (Dans).<sup>38</sup> The Förster radii of the pairs are similar (NALa/Dans, 22.3 Å; Py/Dans, 20.5 Å) and the experimental donor lifetimes of Py (226 ns) and NALa (37 ns) seemed to guarantee a much larger FDE in the Py/Dans chain than in the NALa/Dans chain. In the original article, the diffusion coefficient in water was reported as 4 Å<sup>2</sup>/ns and the average end-to-end distance was reported as 19 Å.<sup>38</sup> These values were later drastically corrected by 1 order of magnitude to  $D = 50 \text{ Å}^2/\text{ns}$  and by a factor of 2 to  $R_{\text{av}} = 38 \text{ Å}$  in water and to  $D = 60 \text{ Å}^2/\text{ns}$  and  $R_{\text{av}} = 50 \text{ Å}$  in the presence of 8 M GdmCl.<sup>61</sup> We base the following considerations on the corrected results.

At the average distance of 38 Å, the probability and rate of FRET are negligibly small in relation to the fluorescence decays (NALa,  $k_{\text{T}} = 0.0011 \text{ ns}^{-1}$ ; Py,  $k_{\text{T}} = 0.00011 \text{ ns}^{-1}$ ). FRET becomes fast only at distances shorter than  $R_0$  (e.g.,  $k_{\text{T,NALa}} (12 \text{ Å}) = 1.11 \text{ ns}^{-1}$ ,  $k_{\text{T,Py}} (12 \text{ Å}) = 0.11 \text{ ns}^{-1}$ ). In water, the bulk of the equilibrium distance distribution contributes negligibly to the ETE (NALa,  $E_0 \sim 4\%$ ; Py,  $E_0 \sim 2\%$ ; eq 4), and even less at 8 M GdmCl (NALa,  $E_0 \sim 1\%$ ; Py,  $E_0 \sim 0.5\%$ ). If we next apply the corrected random-walk equation<sup>54</sup> (eq 1,  $\Delta R_{\text{rms}} = (6D\tau_{\text{D}})^{1/2}$ ), we obtain the displacement that can occur in the absence of FRET until the donor is deactivated by photon emission or quenching,  $\tau_{\text{D}} = (\tau_0^{-1} + \tau_{\text{rad}}^{-1})^{-1}$ . The  $\Delta R_{\text{rms}}$  value is 105 Å for NALa and 260 Å for Py. If these values would properly indicate the FDE, Py would indeed lead to a much larger FDE than NALa. However, the Results demonstrate beyond doubt that the FDE does not depend on the quantum yield but only depends on the radiative lifetime. Hence, when we want to use the displacement to quantify the FDE, we have to calculate it from  $\tau_0$ . The displacements are then 277 Å for NALa ( $\Phi_{\text{D}} \sim 0.14$ ,  $\tau_0 \sim 256 \text{ ns}$ ) and 313 Å for Py ( $\Phi_{\text{D}} \sim 0.69$ ,  $\tau_0 \sim 328 \text{ ns}$ ). The radiative lifetimes are similar and, as a consequence, also the displacement values are similar. While the donor lifetimes differ by more than 1 order of magnitude, the displacement values relevant for assessing the FDE differ by merely 18%, a small difference considering the 50% variation in size of the two aromatic donors. The two donor pairs (NALa and Py) have been, in retrospect, only apparently ideal for the intended FRET analysis.

Be this as it may, the similar, very large displacement values (ca. 300 Å) ascertain that donor and acceptor are always able to traverse through regions of negligible FRET probability to reach distances at which FRET becomes fast, i.e., distances

shorter than the Förster radii. For a global analysis, this situation is far from ideal, as FRET is almost completely caused by diffusion:  $\varphi_{\text{M}}$ , the fraction of FRET events due to diffusion, approaches 100%, while  $\varphi_{\text{EQ}}$  vanishes, no matter whether NALa or Py is used as donor. Such measurements can at best provide a possible lower limit for the dominant distances present in solution and can inform on neither the kind nor the shape of the distribution, certainly not at distances larger than the  $R_0$  value. As becomes obvious from Figure 7c, any variation in the product of  $\tau_0$  and  $D$  would need to cover at least 2 orders of magnitude difference to cover the regimes of fast diffusion (the region with negative slope in Figure 7c) to negligible diffusion (plateau region). In any case, a reliable delineation of diffusion coefficient and distribution requires at least a minimum of FRET, of information, from the equilibrium distances present in solution. The optimal basis for a global analysis would be measurements where  $\varphi_{\text{EQ}}$  varies gradually from 100 to 0%. In a minimalist approach, at least two conditions should be studied: one with  $\varphi_{\text{EQ}}$  significantly above and the other with  $\varphi_{\text{EQ}}$  significantly below 50%. The NALa/Dans and Py/Dans pairs fall short of this requirement, because  $\varphi_{\text{EQ}}$  is virtually zero ( $\varphi_{\text{M}} = 100\%$ ) for both. We have to ask about ways to realize improved conditions.

**Conditions for a Successful Delineation of Polypeptide Structure and Dynamics.** A modest FDE, a modest possible displacement,  $\Delta R_{\text{rms}}$ , requires that the equilibrium distance distribution is mainly defined by distances at which FRET is fast: The FDE-relevant donor lifetime at a specific distance obeys  $\tau_{\text{DA}}(r) = (k_{\text{rad}} + k_{\text{T}}(r))^{-1}$  with high values of  $k_{\text{T}}$  implying small displacements. Indeed, it is common practice in FRET experiments to choose the FRET pairs such that their Förster radii,  $R_0$ , match the anticipated equilibrium distances. To further keep the FDE at a minimum, donors with short radiative lifetimes are required since  $k_{\text{T}} \propto \tau_0^{-1}$  (cf. eq 21).

For a flexible chain in nonviscous solution, however, the FDE is always large and has to be included in the analysis. The short-lived donor FT<sub>rp</sub>, for instance, has an experimental lifetime of only 2 ns but a radiative lifetime of 20 ns ( $\Phi_{\text{D}} \sim 0.10$ ). At negligible FRET probability, the displacement is 70 Å for  $D = 40 \text{ Å}^2/\text{ns}$  and is still high, 11 Å, when the chain is less flexible,  $D = 1 \text{ Å}^2/\text{ns}$ .

A robust and reliable route to reduce and vary the FDE in preparation of a global analysis is by external means, by adding a viscogen that reduces the diffusion coefficient. The remaining problem is that high concentrations of a viscogen such as ethylene glycol, propylene glycol, or sucrose could affect the equilibrium distance distribution; the analysis would again become cumbersome, ambiguous, and unsuitable for a large-scale investigation on polypeptide conformation and dynamics as a function of amino acid composition and sequence. The HSE simulations point to a simple test that can establish whether a viscogen affects only chain dynamics and not structure, a test that can be carried out by simple steady-state-fluorescence measurements.

**The FDE Depends Symmetrically on Diffusion and Radiative Donor Lifetime.** The random-walk perspective led us to speculate that the FDE might depend symmetrically on the diffusion coefficient and the donor lifetime and that both parameters could be used independently to control the FDE. Naturally, the time until the FRET donor–acceptor pair encounters regions of high FRET probability decreases with increasing diffusion, and the average donor lifetime,  $\langle \tau_{\text{DA}} \rangle$ , depends itself on the diffusion coefficient such that  $\Delta R_{\text{rms}} =$

$(6D\langle\tau_{\text{DA}}\rangle(D))^{1/2}$ . As confirmed by simulation, the actual symmetry exists between the radiative lifetime and the diffusion coefficient and is a direct consequence of the specific form of the FRET distance law,  $k_{\text{T}} = k_{\text{rad}}R_{\text{F}}^6/r^6$ : The FDE-relevant donor lifetime at a specific distance follows  $\tau_{\text{DA}}(r) = (k_{\text{rad}} + k_{\text{T}})^{-1} = (k_{\text{rad}} + k_{\text{rad}}R_{\text{F}}^6/r^6)^{-1} = \tau_0(1 + R_{\text{F}}^6/r^6)$ . Thus, also the average donor lifetime  $\langle\tau_{\text{DA}}\rangle$  is proportional to the radiative lifetime:  $\langle\tau_{\text{DA}}\rangle = (k_{\text{rad}} + \langle k_{\text{T}} \rangle)^{-1} = (k_{\text{rad}} + k_{\text{rad}}R_{\text{F}}^6\langle R_{\text{DA}}^{-6} \rangle)^{-1} = \tau_0(1 + R_{\text{F}}^6\langle R_{\text{DA}}^{-6} \rangle)$ . We realize that changing the radiative lifetime leads to the same possible displacement, that is, the same FDE, as changing the diffusion coefficient in identical proportion. By substituting NAla for FTrp in the  $(\text{GS})_6$  chain, we changed the radiative lifetime from 256 to 19.6 ns, a decrease by a factor of 13.0. If we, instead, decrease the diffusion coefficient by the same factor (12.7) through a viscogen, we should monitor an identical  $R_{\text{eff}}$  and FDE value as after donor substitution ( $R_{\text{eff}} = 11.6 \text{ \AA}$ , Table 1). This is indeed the case, within error, when we used ethylene glycol (90% v/v) as a viscogen ( $R_{\text{eff}} = 11.2 \text{ \AA}$ , Table 1). If, upon coagent addition, the chain would collapse or expand, the effective distance would be much smaller or larger than after donor substitution. It seems that, even at high concentrations, ethylene glycol has only a very limited impact on chain structural propensities, which has already been of value in earlier studies on protein folding,<sup>46,83</sup> and provides now an experimental perspective for the exploratory results reported in Table 1.

Finally, we should relate time-resolved short-distance FRET to two complementary photophysical methods employed to assess the biopolymer chain dynamics: single-molecule FRET<sup>25,84</sup> and collision-induced quenching.<sup>57,60,85,86</sup> The donors used in short-distance FRET (FTrp and NAla) are unsuitable for single-molecule FRET due to their far-UV absorbance and low brightness, while the probes commonly used in single-molecule FRET are too large to be useful in monitoring short distances.<sup>41</sup> Collision-induced quenching, on the other hand, which is also limited by intrachain diffusion, differs formally from short-distance FRET in that an immediate contact between probe and quencher is a prerequisite, and it is a future challenge to compare these two methods.<sup>87</sup>

## CONCLUSIONS

The structural and dynamic properties of *flexible* polypeptide chains have seen an immense increase in interest in regard to the refolding of globular proteins, the misfolding of amyloidogenic proteins, and the nonfolding of the intrinsically disordered proteins. Time-resolved FRET spectroscopy is a tool predestined to access dynamic and structural information on the chain. The decisive challenge is to delineate distance distribution and diffusion. We ventured into this project by working on the assumption, made on the basis of suggestions in the literature, that a systematic variation of the concentration of an external quencher could vary the relative proportion of diffusion-enhanced FRET. However, as demonstrated by experiment, theory, and simulations, this is not the case: The information content of FRET experiments depends on the radiative donor lifetime, yet the addition of external quencher does not alter this critical photophysical parameter. The insights which we have obtained are nevertheless of much experimental significance and fundamental for assessing mutual intrachain diffusion by FRET. First, experimental conditions to accurately characterize the initial distribution (plateau region in Figure 7c) and the chain mobility (linear region with negative slope in Figure 7c) would ideally need to involve variations of

either the diffusion coefficient (viscosity) or the radiative donor lifetime, or the product of both, by 2 orders of magnitude. A variation in the observed donor lifetime or quantum yield alone, as it can be readily affected by the addition of an external quencher, varies neither of both critical parameters. Second, the effect of decreased donor quantum yield or quenched donor lifetime is opposite to what was originally expected: The effective donor–acceptor distance extracted from FRET decreases rather than increases in the presence of quencher (Figure 9c), while in the presence of diffusion, this effect becomes only obvious at very large quencher concentrations (or very low quantum yields). The large sensitivity of the effective donor–acceptor distance to the quantum yield of the employed donor for FRET pairs embedded in polydisperse polymer or biopolymer chains of limited or no mobility (or even of any mixture of two noninterconverting FRET systems of varying distance) has not been noted previously.

## ASSOCIATED CONTENT

### Supporting Information

Matlab files for HSE simulations are available from the authors upon request. Derivations of the quantum-yield dependence of the effective FRET parameters for a static equilibrium distance distribution, a rationalization for the pseudomonoeponential kinetics in time-resolved FRET measurements, and a list of symbols and abbreviations are available as Supporting Information. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## AUTHOR INFORMATION

### Corresponding Author

\*E-mail: [m.jacob@jacobs-university.de](mailto:m.jacob@jacobs-university.de) (M.H.J.); [w.nau@jacobs-university.de](mailto:w.nau@jacobs-university.de) (W.M.N.).

### Author Contributions

<sup>†</sup>Both authors contributed equally to this work.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

We would like to thank the Deutsche Forschungsgemeinschaft (DFG, NA 686/6) and the DAAD (doctoral fellowship for A.N.) for financial support. We are grateful to Attila Szabo and Eitan Lerner for their encouraging interest and for fruitful discussions. We respectfully dedicate this article to Izchak Z. Steinberg.

## REFERENCES

- Grupi, A.; Haas, E. *J. Mol. Biol.* **2011**, *405*, 1267–1283.
- Grupi, A.; Haas, E. *J. Mol. Biol.* **2011**, *411*, 234–247.
- Ratner, V.; Kahana, E.; Haas, E. *J. Mol. Biol.* **2002**, *320*, 1135–1145.
- Haas, E. *ChemPhysChem* **2005**, *6*, 858–870.
- Jacob, M. H.; Amir, D.; Ratner, V.; Gussakowsky, E.; Haas, E. *Biochemistry* **2005**, *44*, 13664–13672.
- Ratner, V.; Amir, D.; Kahana, E.; Haas, E. *J. Mol. Biol.* **2005**, *352*, 683–699.
- Ittah, V.; Haas, E. *Biochemistry* **1995**, *34*, 4493–4506.
- Navon, A.; Ittah, V.; Landsman, P.; Scheraga, H. A.; Haas, E. *Biochemistry* **2001**, *40*, 105–118.
- Sinev, M. A.; Sineva, E. V.; Ittah, V.; Haas, E. *Biochemistry* **1996**, *35*, 6425–6437.
- Orevi, T.; Ben Ishay, E.; Pirchi, M.; Jacob, M. H.; Amir, D.; Haas, E. *J. Mol. Biol.* **2009**, *385*, 1230–1242.
- Wright, P. E.; Dyson, H. J. *J. Mol. Biol.* **1999**, *293*, 321–331.

- (12) Avbelj, F.; Grdadolnik, S. G.; Grdadolnik, J.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 1272–1277.
- (13) Schweitzer-Stenner, R. *Mol. Biosyst.* **2012**, *8*, 122–133.
- (14) Magg, C.; Schmid, F. X. *J. Mol. Biol.* **2004**, *335*, 1309–1323.
- (15) Sinha, K. K.; Udgaonkar, J. B. *J. Mol. Biol.* **2007**, *370*, 385–405.
- (16) Dasgupta, A.; Udgaonkar, J. B. *J. Mol. Biol.* **2010**, *403*, 430–445.
- (17) Sarkar, S. S.; Udgaonkar, J. B.; Krishnamoorthy, G. *J. Phys. Chem. B* **2011**, *115*, 7479–7486.
- (18) Schuler, B.; Lipman, E. A.; Eaton, W. A. *Nature* **2002**, *419*, 743–747.
- (19) Schuler, B.; Eaton, W. A. *Curr. Opin. Struct. Biol.* **2008**, *18*, 16–26.
- (20) Lipman, E. A.; Schuler, B.; Bakajin, O.; Eaton, W. A. *Science* **2003**, *301*, 1233–1235.
- (21) Schuler, B.; Lipman, E. A.; Steinbach, P. J.; Kumke, M.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 2754–2759.
- (22) Gopich, I. V.; Szabo, A. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 7747–7752.
- (23) Klostermeier, D.; Millar, D. P. *Biopolymers* **2001**, *61*, 159–179.
- (24) Kahra, D.; Kovermann, M.; Low, C.; Hirschfeld, V.; Haupt, C.; Balbach, J.; Hübner, C. G. *J. Mol. Biol.* **2011**, *411*, 781–790.
- (25) Nettels, D.; Gopich, I. V.; Hoffmann, A.; Schuler, B. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 2655–2660.
- (26) Steinberg, I. Z.; Katchalski-Katzir, E. *J. Chem. Phys.* **1968**, *48*, 2404–2410.
- (27) Haas, E.; Katchalski-Katzir, E.; Steinberg, I. Z. *Biopolymers* **1978**, *17*, 11–31.
- (28) Beechem, J. M.; Haas, E. *Biophys. J.* **1989**, *55*, 1225–1236.
- (29) Steinberg, I. Z. *J. Theor. Biol.* **1994**, *166*, 173–187.
- (30) Haas, E.; Wilchek, M.; Katchalski-Katzir, E.; Steinberg, I. Z. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 1807–1811.
- (31) Lakowicz, J. R.; Gryczynski, I.; Kusba, J.; Wiczak, W.; Szmajdzinski, H.; Johnson, M. L. *Photochem. Photobiol.* **1994**, *59*, 16–29.
- (32) Kusba, J.; Lakowicz, J. R. *Methods Enzymol.* **1994**, *240*, 216–262.
- (33) Kusba, J.; Piszczek, G.; Gryczynski, I.; Johnson, M. L.; Lakowicz, J. R. *Chem. Phys. Lett.* **2000**, *319*, 661–668.
- (34) Lakowicz, J. R.; Kusba, J.; Gryczynski, I.; Wiczak, W.; Szmajdzinski, H.; Johnson, M. L. *J. Phys. Chem.* **1991**, *95*, 9654–9660.
- (35) Berg, H. C. *Random Walks in Biology*; Princeton University Press: Princeton, NJ, 1993.
- (36) Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*, 3rd ed.; Kluwer Academic/Plenum: New York, 2006; p 498.
- (37) Makarov, D. E.; Plaxco, K. W. *J. Chem. Phys.* **2009**, *131*, No. 085105.
- (38) Möglich, A.; Joder, K.; Kiefhaber, T. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 12394–12399.
- (39) Sahoo, H.; Roccatano, D.; Zacharias, M.; Nau, W. M. *J. Am. Chem. Soc.* **2006**, *128*, 8118–8119.
- (40) Sahoo, H.; Nau, W. M. *ChemBioChem* **2007**, *8*, 567–573.
- (41) Sahoo, H.; Roccatano, D.; Hennig, A.; Nau, W. M. *J. Am. Chem. Soc.* **2007**, *129*, 9762–9772.
- (42) Jacob, M. H.; Nau, W. M. In *Folding, Misfolding and Nonfolding of Peptides and Small Proteins*; Schweitzer-Stenner, R., Ed.; John Wiley & Sons: Hoboken, NJ, 2011.
- (43) Verbaro, D.; Ghosh, I.; Nau, W. M.; Schweitzer-Stenner, R. *J. Phys. Chem. B* **2010**, *114*, 17201–17208.
- (44) Hudgins, R. R.; Huang, F.; Gramlich, G.; Nau, W. M. *J. Am. Chem. Soc.* **2002**, *124*, 556–564.
- (45) Nau, W. M.; Huang, F.; Wang, X.; Bakirci, H.; Gramlich, G.; Marquez, C. *Chimia* **2003**, *57*, 161–167.
- (46) Jacob, M.; Schindler, T.; Balbach, J.; Schmid, F. X. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 5622–5627.
- (47) Jerome, F. S.; Tseng, J. T.; Fan, L. T. *J. Chem. Eng. Data* **1968**, *13*, 496.
- (48) Schuler, B.; Kremer, W.; Kalbitzer, H. R.; Jaenicke, R. *Biochemistry* **2002**, *41*, 11670–11680.
- (49) Naganathan, A. N.; Perez-Jimenez, R.; Sanchez-Ruiz, J. M.; Munoz, V. *Biochemistry* **2005**, *44*, 7435–7449.
- (50) Eftink, M. R.; Jia, Y. W.; Hu, D.; Ghiron, C. A. *J. Phys. Chem.* **1995**, *99*, 5713–5723.
- (51) Joshi, S.; Ghosh, I.; Pokhrel, S.; Madler, L.; Nau, W. M. *ACS Nano* **2012**, *6*, 5668–79.
- (52) Khan, Y. R.; Dykstra, T. E.; Scholes, G. D. *Chem. Phys. Lett.* **2008**, *461*, 305–309.
- (53) Förster, T. *Ann. Phys.* **1948**, *2*, 55–75.
- (54) Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*, 3rd ed.; Kluwer Academic/Plenum: New York, 2006; p 142.
- (55) Liu, T.; Callis, P. R.; Hesp, B. H.; de Groot, M.; Buma, W. J.; Broos, J. *J. Am. Chem. Soc.* **2005**, *127*, 4104–4113.
- (56) Sarkar, S. S.; Udgaonkar, J. B.; Krishnamoorthy, G. *J. Phys. Chem. B* **2011**, *115*, 7479–7486.
- (57) Bieri, O.; Wirz, J.; Hellrung, B.; Schutkowski, M.; Drewello, M.; Kiefhaber, T. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9597–601.
- (58) Krieger, F.; Fierz, B.; Bieri, O.; Drewello, M.; Kiefhaber, T. *J. Mol. Biol.* **2003**, *332*, 265–274.
- (59) Fierz, B.; Joder, K.; Krieger, F.; Kiefhaber, T. *Methods Mol. Biol.* **2007**, *350*, 169–187.
- (60) Neuweiler, H.; Lollmann, M.; Doose, S.; Sauer, M. *J. Mol. Biol.* **2007**, *365*, 856–869.
- (61) Möglich, A.; Joder, K.; Kiefhaber, T. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 6787.
- (62) Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*, 3rd ed.; Kluwer Academic/Plenum: New York, 2006; p 445.
- (63) Steinberg, I. Z.; Haas, E.; Katchalski-Katzir, E. In *Time Resolved Spectroscopy in Biology*; Cundall, R. B., Dale, R. E., Eds.; Plenum Publishing: New York, 1983; pp 411–415.
- (64) Haas, E. In *Protein Folding Handbook*; Buchner, J., Kiefhaber, T., Eds.; John Wiley & Sons: Hoboken, NJ, 2008.
- (65) Klan, P.; Wirz, J. *Photochemistry of Organic Compounds*; Wiley: New York, 2009.
- (66) Plaxco, K. W.; Uzawa, T.; Cheng, R. R.; Cash, K. J.; Makarov, D. E. *Biophys. J.* **2009**, *97*, 205–210.
- (67) Collins, K. D.; Neilson, G. W.; Enderby, J. E. *Biophys. Chem.* **2007**, *128*, 95–104.
- (68) Santoro, M. M.; Bolen, D. W. *Biochemistry* **1988**, *27*, 8063–8068.
- (69) Jacob, M. H.; Saudan, C.; Holtermann, G.; Martin, A.; Perl, D.; Merbach, A. E.; Schmid, F. X. *J. Mol. Biol.* **2002**, *318*, 837–845.
- (70) Bicout, D. J.; Szabo, A. *Protein Sci.* **2000**, *9*, 452–465.
- (71) Roccatano, D.; Sahoo, H.; Zacharias, M.; Nau, W. M. *J. Phys. Chem. B* **2007**, *111*, 2639–2646.
- (72) The combination of both “diffusion-suppressing” strategies, i.e., the use of the short-lived donor in a highly viscous solution, led to an increase in  $R_{\text{off}}$  by ca. 6 Å (see last entry in Table 1); the “true”  $R_{\text{off}}$  value (in the absence of any FDE) is still expected to be larger, and we are trying to reach the limit in an ongoing experimental project.
- (73) Huang, F.; Rajagopalan, S.; Settanni, G.; Marsh, R. J.; Armoogoo, D. A.; Nicolaou, N.; Bain, A. J.; Lerner, E.; Haas, E.; Ying, L. M.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 20758–20763.
- (74) Nettels, D.; Hoffmann, A.; Schuler, B. *J. Phys. Chem. B* **2008**, *112*, 6137–6146.
- (75) Lapidus, L. J.; Eaton, W. A.; Hofrichter, J. *J. Mol. Biol.* **2002**, *319*, 19–25.
- (76) Soranno, A.; Longhi, R.; Bellini, T.; Buscaglia, M. *Biophys. J.* **2009**, *96*, 1515–1528.
- (77) Thirumalai, D.; Ha, B.-Y. In *Theoretical and Mathematical Models in Polymer Research*; Grossberg, A., Ed.; Academic: New York, 1998; pp 1–35.
- (78) May, V.; Kühn, O. *Charge and Energy Transfer Dynamics in Molecular Systems*; Wiley-VCH: Berlin, 2000.
- (79) Sridevi, K.; Udgaonkar, J. B. *Biochemistry* **2003**, *42*, 1551–1563.
- (80) Sinha, K. K.; Udgaonkar, J. B. *J. Mol. Biol.* **2005**, *353*, 704–718.
- (81) Lapidus, L. J.; Steinbach, P. J.; Eaton, W. A.; Szabo, A.; Hofrichter, J. *J. Phys. Chem. B* **2002**, *106*, 11628–11640.
- (82) Ingratta, M.; Duhamel, J. *J. Phys. Chem. B* **2008**, *112*, 9209–9218.

(83) Jacob, M.; Geeves, M.; Holtermann, G.; Schmid, F. X. *Nat. Struct. Biol.* **1999**, *6*, 923–926.

(84) Neuweiler, H.; Schulz, A.; Bohmer, M.; Enderlein, J.; Sauer, M. *J. Am. Chem. Soc.* **2003**, *125*, 5324–5330.

(85) Huang, F.; Nau, W. M. *Angew. Chem., Int. Ed.* **2003**, *42*, 2269–72.

(86) Lapidus, L. J.; Eaton, W. A.; Hofrichter, J. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 7220–7225.

(87) When the monoexponential FTrp decay lifetimes (Table 1) are employed to calculate an apparent quenching rate constant from the short-distance FRET experiments, the resulting values are up to 1 order of magnitude faster than the rate constants extracted from collision-induced quenching experiments for the same peptides (ref 44) or the same peptide backbone (ref 58).

## Supporting Information

# Diffusion-enhanced FRET and the Effect of External Quenchers and the Donor Quantum Yield

*Maik H. Jacob* \*, *Roy Dsouza*, *Indrajit Ghosh*, *Amir Norouzy*,

*Thomas Schwarzlose*, and *Werner M. Nau* \*

*School of Engineering and Science, Jacobs University of Bremen, Campus Ring 1,*

*D-28759, Bremen, Germany*

*E-mail: m.jacob@jacobs-university.de, w.nau@jacobs-university.de*

### Table of Contents

- 1. The quantum-yield dependence of effective FRET parameters for a static equilibrium distance distribution**
  - 1.1. The quantum-yield dependence of the effective distance**
  - 1.2. The quantum-yield dependence of the effective FRET rate**
- 2. Pseudo-monoexponential kinetics in time-resolved FRET measurements**
- 3. Symbols and Abbreviations**

## 1. The quantum-yield dependence of effective FRET parameters for a static equilibrium distance distribution.

### 1.1. The quantum-yield dependence of the effective distance

The quantum yield dependence of the effective distance was found to be:

$$R_{eff}(q) = \Phi_D^{1/6} R_F \left( \left( 1 - \int p(r) \frac{r^6}{r^6 + \Phi_D R_F^6} dr \right)^{-1} - 1 \right)^{1/6}.$$

The derivation begins with the definition of the effective distance.

$$E_0 = \frac{R_0^6}{R_0^6 + R_{eff}^6} \quad \text{or} \quad R_{eff} = R_0 \left( \frac{1}{E_0} - 1 \right)^{1/6}.$$

Given a distance probability density distribution,  $p(r)$ , the ETE in the absence of diffusion obeys

$$E_0 = \langle E_0 \rangle = \int p(r) E_0(r) dr$$

(Recall the definition of the average of an observable,  $A$ ,  $\langle A \rangle = \int A(x) p(x) dx$ , the observed ETE has to be identical to the distance average of the local efficiency,  $E_0(r)$ .)

Using  $E_0(r) = \frac{R_0^6}{R_0^6 + r^6}$  and that the average of a constant is the constant itself, we can form

$$E_0 = \int p(r) \frac{R_0^6}{R_0^6 + r^6} dr = \int p(r) \left( 1 - \frac{r^6}{r^6 + R_0^6} \right) dr = 1 - \int p(r) \frac{r^6}{r^6 + R_0^6} dr.$$

Inserting this expression into the definition of the effective distance, we obtain

$$R_{eff} = R_0 \left( \left( 1 - \int p(r) \frac{r^6}{r^6 + R_0^6} dr \right)^{-1} - 1 \right)^{1/6}.$$

The quantum-yield dependence is obtained after the replacements  $R_0^6 = \Phi_D R_F^6$  and  $R_0 = \Phi_D^{1/6} R_F$ .

For a single, static distance,  $R_{DA}$ , the distribution is close to a  $\delta$ -function and the integral  $\int p(r) \frac{r^6}{r^6 + R_0^6} dr$  becomes  $\frac{R_{DA}^6}{R_{DA}^6 + R_0^6}$ . Insertion yields correctly that  $R_{eff} = R_{DA}$ . We see that the

quantum-yield dependence of the effective distance is a consequence of the distance distribution:

It does not apply to the case of a single, static distance that describes, for instance, a short, inflexible polyproline peptide.

## 1.2. The quantum-yield dependence of the effective FRET rate

Although the elementary FRET rate constant,  $k_T(r)$  does not depend on the quantum yield, the effective FRET rate,  $k_{\text{FRET}}$ , shows a weak dependence, because it is experimentally defined through the observed efficiency,  $E_{\text{obs}}$ . In the absence of diffusion, an analytical treatment is possible. The quantum yield dependence of the effective FRET rate was found to be:

$$k_{\text{FRET0}}(\Phi_D) = \frac{k_{\text{rad}}}{\Phi_D} \left( \left( \int p(r) \frac{r^6}{r^6 + \Phi_D R_F^6} dr \right)^{-1} - 1 \right).$$

We recognize the consistency of the obtained expressions with  $k_{\text{FRET0}} = k_D \frac{R_0^6}{R_{\text{eff}}^6}$ .

To instill trust in and familiarity with this formula we offer two alternative derivations.

1) The effective FRET rate in the absence of diffusion ( $E_{\text{obs}} = E_0$ ) is defined by

$$E_0 = \frac{k_{\text{FRET0}}}{k_{\text{FRET0}} + k_D} \text{ or } k_{\text{FRET0}} = k_D \left( \frac{1}{E_0} - 1 \right)^{-1}.$$

Using  $E_0 = \int p(r) \frac{R_0^6}{R_0^6 + r^6} dr$ , we obtain

$$k_{\text{FRET0}} = k_D \left( \left( \int p(r) \frac{R_0^6}{R_0^6 + r^6} dr \right)^{-1} - 1 \right)^{-1} \text{ or } k_{\text{FRET0}} = \frac{k_{\text{rad}}}{\Phi_D} \left( \left( 1 - \int p(r) \frac{r^6}{\Phi_D R_F^6 + r^6} dr \right)^{-1} - 1 \right)^{-1}.$$

2) In the second derivation we use again the definition  $E_0 = k_{\text{FRET0}}/(k_{\text{FRET0}}+k_D)$  but also that  $E_0$  is given by  $E_0 = (A_D - A_{\text{DA0}})/A_D$ , the areas under the kinetic traces, where  $(A_D - A_{\text{DA0}})$  is the fluorescence quenched by FRET. Both equations are consequences of the ETE being the fraction of donor deactivation events caused by FRET. Eliminating  $E_0$ , we obtain for  $k_{\text{FRET}}$

$$k_{\text{FRET}} = k_D \left( \frac{A_D}{A_{\text{DA0}}} - 1 \right) = \frac{1}{A_{\text{DA0}}} - k_D,$$

where we used that  $A_D = \tau_D = k_D^{-1}$ .

The time-course of the donor emission intensity of the donor-acceptor peptide in the absence of diffusion is given by

$$I_{\text{DA0}}(t) = \int p(r) \exp(-(k_D + k_T(r))t) \cdot dr,$$

where  $I_{\text{DA0}}(t=0)$  is 100%.

The area, which presents the measure of the total emitted fluorescence, becomes

$$A_{\text{DA0}} = \int_t \int_r p(r) \exp(-(k_D + k_T)t) dr dt$$

or, when we switch the order of integration

$$A_{\text{DA0}} = \int_r p(r) \int_t \exp(-(k_D + k_T(r))t) dt dr.$$

The solution of the inner integral is  $1/(k_D + k_T)$ .

$$A_{\text{DA0}} = \int_r p(r) \frac{1}{k_D + k_T(r)} dr.$$

We use the Förster distance law  $k_T(r) = k_D R_0^6 / r^6$  and  $k_D^{-1} = \tau_D$  to obtain

$$A_{\text{DA0}} = \int p(r) \frac{1}{k_D + k_D \frac{R_0^6}{r^6}} dr = \tau_D \int p(r) \frac{1}{1 + \frac{R_0^6}{r^6}} dr = \tau_D \int p(r) \frac{r^6}{r^6 + R_0^6} dr.$$

Insertion into  $k_{\text{FRET}} = I/A_{\text{DA0}} - k_D$  yields

$$k_{\text{FRET}} = \frac{1}{A_{\text{DA0}}} - k_D = \frac{1}{\tau_D \int p(r) \frac{r^6}{r^6 + R_0^6} dr} - k_D \text{ or}$$

$$k_{\text{FRET}} = \left( \tau_D \int p(r) \frac{r^6}{r^6 + R_0^6} dr \right)^{-1} - k_D = k_D \left( \left( \int p(r) \frac{r^6}{r^6 + R_0^6} dr \right)^{-1} - 1 \right).$$

Both,  $k_D$  and  $R_0$ , depend on the quantum yield (eqs. 19 and 23 in main text), and the dependencies do not cancel out completely:

$$k_{\text{FRET0}}(\Phi_D) = \frac{k_{\text{rad}}}{\Phi_D} \left( \left( \int p(r) \frac{r^6}{r^6 + \Phi_D R_F^6} dr \right)^{-1} - 1 \right)$$

When the distribution narrows and approaches a  $\delta$ -function, the Förster distance law is recovered and the quantum-yield dependence vanishes:

$$k_{\text{FRET0}} = k_{\text{D}} \left( \left( \int p(r) \frac{r^6}{r^6 + R_0^6} dr \right)^{-1} - 1 \right) = k_{\text{D}} \left( \left( \frac{r^6}{r^6 + R_0^6} \right)^{-1} - 1 \right) = k_{\text{D}} \left( \left( \frac{r^6 + R_0^6}{r^6} \right) - 1 \right) = k_{\text{D}} \frac{R_0^6}{r^6} = k_{\text{F}} \frac{R_{\text{F}}^6}{r^6}$$

The results of both derivations are identical (use  $\int p(r) \frac{R_0^6}{R_0^6 + r^6} dr = 1 - \int p(r) \frac{r^6}{R_0^6 + r^6} dr$ ):

$$k_{\text{FRET0}} = k_{\text{D}} \left( \left( \int p(r) \frac{R_0^6}{R_0^6 + r^6} dr \right)^{-1} - 1 \right)^{-1} = k_{\text{D}} \left( \left( \int p(r) \frac{r^6}{R_0^6 + r^6} dr \right)^{-1} - 1 \right).$$

## 2. Pseudo-monoexponential kinetics in time-resolved FRET measurements

We observed in the experiments that the donor decay kinetics of the donor-acceptor peptides NAla-(GS)<sub>6</sub>-Dbo and FTrp-(GS)<sub>6</sub>-Dbo could agreeably be fitted to monoexponential decay functions. Secondly, we noted in the simulations that the effective rate constant,  $k_{\text{FRET}}$ , is virtually independent of the quantum yield, in contrast to the more pronounced dependency in the absence of diffusion. Both observations are related.

When the donor emission decay in the donor-acceptor peptide proceeds monoexponentially, its time-dependent intensity has to follow  $I_{\text{DA}} = I_0 \exp(-(k_{\text{D}} + k'_{\text{FRET}})t)$ . Neither  $k_{\text{D}}$  nor  $k'_{\text{FRET}}$  depend on the distance. The monoexponential time course is a consequence of a time-invariant shape of the distribution  $N^*(r,t)$ . Whereas  $N^*(r,t)$  decreases with time, the normalized distribution  $p(r,t)$  is time invariant. For this invariancy to be possible, the rate constant of the emission decay cannot change with distance. We show shortly that  $k'_{\text{FRET}}$  is identical to the effective FRET rate,  $k_{\text{FRET}}$ , as defined by  $E_{\text{obs}} = k_{\text{FRET}}/(k_{\text{FRET}} + k_{\text{D}})$  and calculated by  $k_{\text{FRET}} = k_{\text{D}} E_{\text{obs}}/(1 - E_{\text{obs}})$ . The area  $A_{\text{DA}}$  under the kinetic trace obeys  $A_{\text{DA}} = \int_t \exp(-(k_{\text{D}} + k'_{\text{FRET}})t) dt = (k_{\text{D}} + k'_{\text{FRET}})^{-1}$ . With that result used in the equation  $k_{\text{FRET}} = 1/A_{\text{DA}} - k_{\text{D}}$ , we obtain  $k_{\text{FRET}} = (k_{\text{D}} + k'_{\text{FRET}}) - k_{\text{D}} = k'_{\text{FRET}}$ .

The shape invariance of the distribution also means that the rate constant  $k'_{\text{FRET}}$  is identical to the average of the distance-dependent FRET rate,  $\langle k_T \rangle$ ,<sup>78</sup> which excludes any quantum-yield dependency:

$$k'_{\text{FRET}} = \langle k_T \rangle = \int k_T(r) p(r) dr = \int k_D (R_0^6 / r^6) p(r) dr = k_{\text{rad}} R_F^6 \langle r^{-6} \rangle$$

It requires unrealistically large diffusion coefficients for the initial shape of the distance distribution to undergo not the slightest change. But an already sufficient condition for the experimental decays to appear monoexponentially is that the stationary probability distribution is attained very soon after excitation, and that the initial and stationary distribution differ only slightly (Fig. 8a in main text).

### 3. Symbols and Abbreviations

$A$	Index that denotes the FRET acceptor
$A_D$	Area under the time course of donor emission in the donor-only peptide
$A_{DA}$	Area under the time course of donor emission in the donor-acceptor peptide
$A_{DA0}$	Area under the time course of donor emission in the donor-acceptor peptide in the absence of diffusion ( $D = 0$ )
$D$	Diffusion coefficient
$D$	Index that denotes the FRET donor
<b>ETE</b>	Energy transfer efficiency
$E_{\text{obs}}$	Observed energy transfer efficiency
$E(r)$	Distance-dependent energy transfer efficiency
$E_0$	Observed energy transfer efficiency in absence of diffusion
$\Delta E_{\text{FDE}}$	Increase of $E_{\text{obs}}$ caused by diffusion
$E_{\text{EQ}}$	Fraction of donor-deactivation events attributable to equilibrium distance distribution
$E_{\text{M}}$	Fraction of donor-deactivation events attributable to donor-acceptor diffusional motion
<b>FDE</b>	FRET diffusion enhancement

<b>FRET</b>	Förster resonance energy transfer
$\varphi_{\text{EQ}}$	Fraction of FRET events attributable to the equilibrium distance distribution
$\varphi_{\text{M}}$	Fraction of FRET events attributable to the equilibrium distance distribution
$\Phi_{\text{D}}$	Donor-quantum yield
<b>HSE</b>	Haas-Steinberg equation
$I_{\text{D}}$	Intensity of the donor fluorescence in the donor-only peptide.
$I_{\text{DA}}$	Intensity of the donor fluorescence in the donor-acceptor peptide.
$k_{\text{T}}(\mathbf{r})$	Distance-dependent rate constant of Förster transfer
$k_{\text{FRET}}$	Effective FRET rate constant
$k_{\text{FRET0}}$	Effective FRET rate constant in absence of diffusion
$k_{\text{rad}}$	Rate constant of donor deactivation at a quantum yield of unity and in absence of FRET
$N^*(\mathbf{r},t)$	Distance distribution after donor excitation in the donor-acceptor peptide
$N_0^*(\mathbf{r})$	Normalized initial distance distribution at $t = 0$
$p(\mathbf{r})$	Normalized equilibrium distance distribution
$p_s(\mathbf{r})$	Normalized stationary distance distribution attained after donor excitation
$r$	Donor-acceptor distance in a distribution
$R_{\text{DA}}$	Single donor-acceptor distance in a rigid system
$R_{\text{eff}}$	Effective donor-acceptor distance
$R_0$	Förster radius at a specific donor quantum yield
$R_{\text{F}}$	Förster radius at a quantum yield of unity, radiative Förster radius
$\tau_0$	Radiative donor lifetime at a quantum yield of unity
$\tau_{\text{D}}$	Experimentally determined donor lifetime in the donor-only peptide
$\tau_{\text{DA}}$	Experimentally determined donor lifetime in the donor-acceptor peptide

## Indicator Displacement Assays Inside Live Cells\*\*

Amir Norouzy, Zahra Azizi, and Werner M. Nau\*

**Abstract:** The macrocycle *p*-sulfonatocalix[4]arene (CX4) and the fluorescent dye lucigenin (LCG) form a stable host-guest complex, in which the dye fluorescence is quenched. Incubation of live V79 and CHO cells with the CX4/LCG chemosensing ensemble resulted in its spontaneous uptake. Subsequent addition of choline, acetylcholine, or protamine, which have a high affinity for CX4 and are capable of entering cells, resulted in a fluorescence switch-on response. This can be traced to the displacement of LCG from CX4 by the analytes. The results establish the principal functionality of indicator displacement assays with synthetic receptors for the detection of the uptake of bioorganic analytes by live cells.

Indicator displacement assays (IDAs) with synthetic receptors have gained increasing attention in (bio)analytical chemistry<sup>[1]</sup> because they offer a supramolecular approach for the sensitive detection and differentiation of analytes. IDAs bypass both the need to construct highly specific antibodies for immunoassays and the design of nanotechnological sensing systems, for example those based on graphene oxide<sup>[2]</sup> or capped mesoporous<sup>[3]</sup> architectures. The measurement principle is based on the use of an indicator dye bound to a receptor (the “reporter pair”). The receptor, frequently a macrocyclic host, is chosen to have a high affinity for the target analyte, such that competitive binding leads to a release of the dye, which in turn results in a change of its photo-physical properties, preferably its fluorescence. A characteristic of such assays is the low selectivity with which the receptor binds different analytes, which can be an advantage because a broad range of analytes can be targeted without the need to synthesize a specific receptor for each of them. On the other hand, it is a disadvantage because the assays are very sensitive to competitive binding by other species with complementary charge and size, which are omnipresent in many fluid matrices and are almost always present in

biological samples.<sup>[4]</sup> Nevertheless, several biologically relevant applications of macrocycles have been brought forward, including the delivery of drugs,<sup>[5]</sup> improved membrane passage of fluorescent dyes<sup>[6]</sup> and DNA,<sup>[7]</sup> the immobilization of cells,<sup>[8]</sup> protein recognition,<sup>[9]</sup> and cell imaging.<sup>[10]</sup>

We have previously designed time-resolved variants of displacement assays, which we refer to as supramolecular tandem assays because they are capable of monitoring enzymatic reactions in real time.<sup>[11]</sup> Interestingly, the assays performed well not only with purified enzymes but also with crude enzyme extracts<sup>[11c]</sup> and even with dried cells<sup>[11a]</sup> expressing a particular enzyme. Although the presence of salts, metabolites, and proteins in biological samples may lead to a sizable instantaneous change of the absolute signal owing to a competitive binding,<sup>[11c,12]</sup> it is the temporal response of the assay that serves as the robust fingerprint of the enzymatic activity, and this temporal response is what is analyzed.<sup>[13]</sup> This insight has encouraged the development of additional methods in which the spatiotemporal response of indicator displacement can be put to work. For example, we recently introduced tandem membrane assays, in which the reporter pair is spatially isolated inside the liposomes; in this manner, after addition of a target analyte to the outside bulk solution, its protein-mediated translocation through the biomembranes can be monitored in real time.<sup>[14]</sup> The combined results, namely a sizable tolerance towards competitive binders and a compartmentalized response, encouraged us to proceed towards the penultimate biological challenge: the transfer of IDAs with synthetic receptors to live cells in order to monitor the cellular uptake of biomolecular analytes. In fact, the application of artificial receptors for monitoring bioorganic analytes or drugs in cellular systems has been identified as a challenge in its own right;<sup>[1b,4,10,15]</sup> it goes beyond the well-established detection of certain inorganic ions in cells<sup>[16]</sup> or the use of functional group-specific fluorogenic sensors, such as those introduced for biothiols.<sup>[17]</sup>

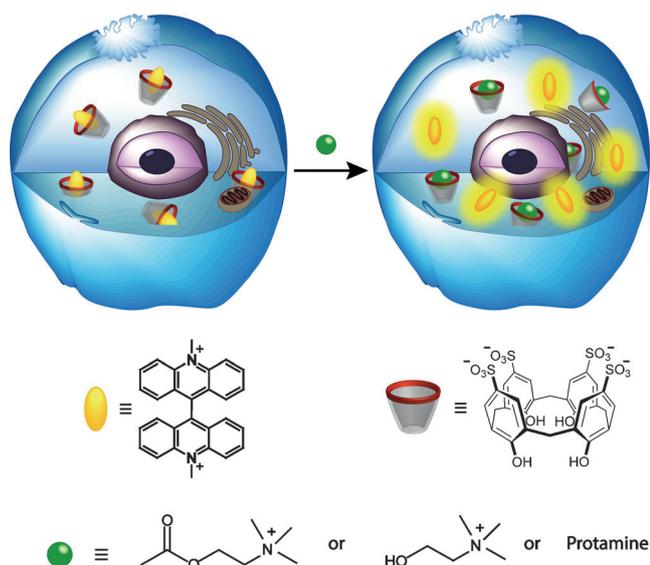
We tested various macrocycle/dye combinations for their compatibility with biomembranes and found that the combination of *p*-sulfonatocalix[4]arene (CX4) and *N,N'*-dimethyl-9,9'-biacridinium dinitrate (lucigenin, LCG) gave a reporter pair that showed spontaneous uptake into live Chinese hamster ovary (CHO) and fibroblast cells (V79). The uptake of the reporter pair should allow us to subsequently perform cellular IDAs, while the high biocompatibility and low toxicity of both components, LCG and calixarenes in general,<sup>[18]</sup> as well as CX4 in particular,<sup>[18e,f,19]</sup> should be beneficial for live-cell measurements. Indeed, incubation with LCG (250  $\mu$ M, 24 h) had no adverse effect on the cells. The cellular IDA, which we have now been able to realize, is schematically depicted in Scheme 1. The results relating to the successful uptake of analytes, monitored through the switch-on fluorescence response, are shown in Figure 1.

[\*] Dr. A. Norouzy, Prof. Dr. W. M. Nau  
Department of Life Sciences and Chemistry, Jacobs University  
Bremen  
Campus Ring 1, 28759 Bremen (Germany)  
E-mail: w.nau@jacobs-university.de

Z. Azizi  
Centre for Biomolecular Interactions, University of Bremen  
Leobener Straße NW2, 28359 Bremen (Germany)

[\*\*] We appreciate the help of D. Gabel (Jacobs University), who provided the V79 cells and cell culture facilities, K. Brix, M. Rehders, J. Fritz, and M. Winterhalter (Jacobs University), K. Maedler (University of Bremen), as well as R. Amann, B. Fuchs and A. Ellrott (Max-Planck-Institute für Marine Mikrobiologie) for instrument access. The authors acknowledge project support from the DFG (W.M.N.) and the DAAD (doctoral fellowship for A.N.).

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/anie.201407808>.



**Scheme 1.** The uptake of analytes into cells preloaded with the macrocycle–dye complex results in displacement of the dye from the macrocycle and an associated fluorescence increase.

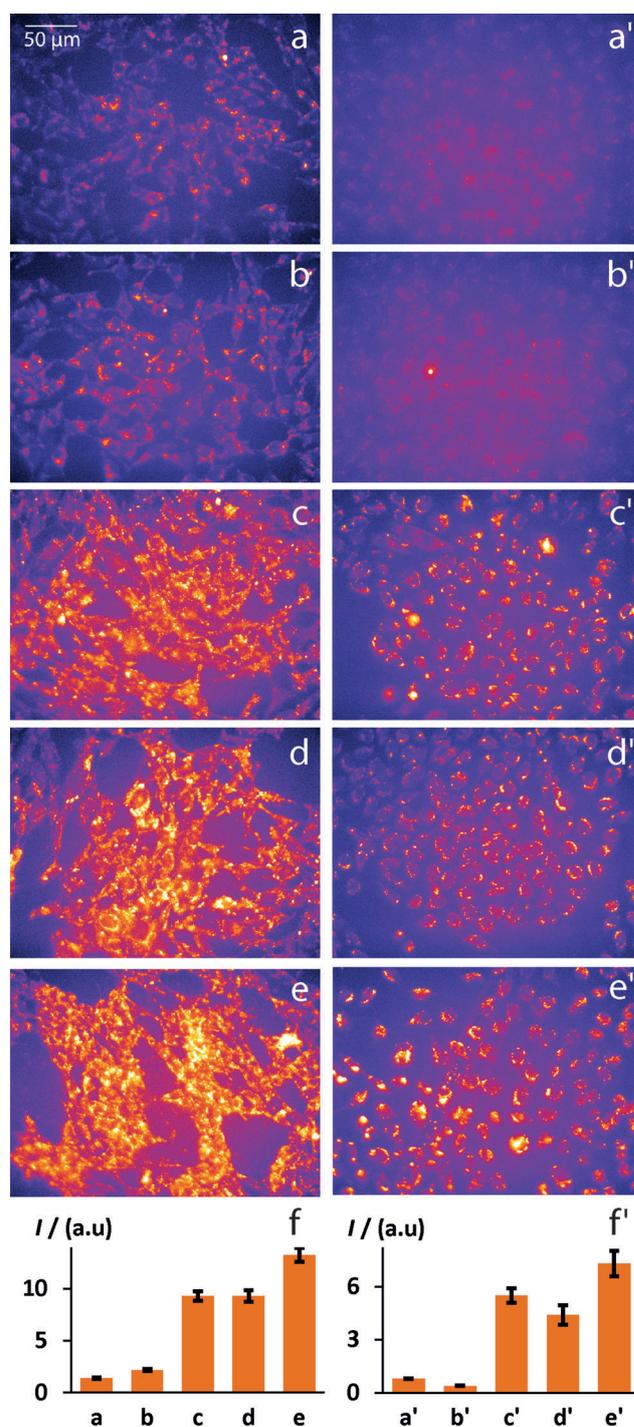
The uptake of dye was confirmed and quantified through incubation of the dye (in the absence and presence of macrocycle), subsequent lysis of the cells, and measurement of the fluorescence after dilution and addition of competitor (to ensure dissociation of the host–dye complex). It should be noted that the fluorescence of the dye is quenched in the supramolecular complex, thus resulting in a switch-off response upon the addition of CX4 and a switch-on response upon addition of competitor.<sup>[20]</sup> The uptake of macrocycle (in the presence of dye) was confirmed indirectly from no less than four independent observations: 1) Cells incubated with LCG (800  $\mu\text{M}$ , 30 min) in the presence of CX4 (up to 5-fold excess) displayed an up to twice as efficient uptake of LCG (Table 1),<sup>[21]</sup> thus suggesting that the host additionally functions as a carrier (reminiscent of DNA transfection with cationic calixarenes).<sup>[7]</sup> 2) Fluorescence images of cells incubated with LCG in the presence of CX4 afforded much weaker fluorescence (Figure S7 in the Supporting Information). 3) The fluorescence of the lysate of cells incubated with LCG in the presence of CX4 increased substantially upon the addition of competitors, thereby confirming the dissociation of the host–dye complex and thus the presence of CX4 (Figure S4). 4) The observed switch-on fluorescence upon the addition of analytes to the live cells (Figure 1), as opposed to

**Table 1:** Uptake efficiency and absolute cellular concentrations of LCG in the absence and presence of CX4.<sup>[a]</sup>

Cell line	CX4 concentration [mM]	LCG concentration in cells [ $\mu\text{M}$ ] <sup>[b]</sup>	Uptake [%] <sup>[b]</sup>
V79 cells	0	75	9.4
	4.5	150	19
CHO cells	0	10	1.3
	4.5	20	2.5

[a] Quantified with the lysis method, see the Supporting Information.

[b] Incubation conditions: 800  $\mu\text{M}$  LCG, 30 min.



**Figure 1.** Fluorescence images of V79 cells (left column) and CHO cells (right), both incubated with 50  $\mu\text{M}$  LCG and 250–300  $\mu\text{M}$  CX4 at 37 °C for 15 min, followed by a 10 min incubation with medium (a and a', as blank controls), 50 mM betaine (b and b', as negative controls), 50 mM choline (c, c'), 50 mM acetylcholine (d, d'), and 15 min incubation with 200  $\mu\text{M}$  protamine (e, e'). The bar graphs (f, f') show the relative averaged fluorescence intensity per cell.

a simple quenching, requires the liberation of the dye from a quenched state, the macrocycle–dye complex.<sup>[22]</sup>

The surprising finding that the IDA remains operational inside live cells (Figure 1) can be explained by the favorable

characteristics of the specific host/dye system. First, from experiments in pure water, LCG binds CX4 with a high affinity ( $K_a = 2.8 \times 10^7 \text{ M}^{-1}$ )<sup>[20]</sup> for a synthetic macrocycle,<sup>[23]</sup> thus facilitating the detection of competitive binders that also display high effective binding constants, such as acetylcholine (ACh,  $K_a = 1.0 \times 10^5 \text{ M}^{-1}$ ), choline (Ch,  $K_a = 1.0 \times 10^5 \text{ M}^{-1}$ ),<sup>[20]</sup> and the polycationic peptide protamine ( $K_a = 1.24 \times 10^9 \text{ M}^{-1}$ ).<sup>[14]</sup> Second, CX4 quenches the fluorescence of LCG very efficiently (by a factor of up to 140),<sup>[20]</sup> which ensures a readily detectable response upon competitive binding. Preliminary experiments in Tyrode's solution and in the media (Figure S6), showed that both properties of the reporter pair are diminished as expected owing to the presence of large amounts of competitive salts, nutrients, and other biomolecules,<sup>[12b]</sup> but remained at a sufficiently high level ( $K_a = 3.0 \times 10^5 \text{ M}^{-1}$ , fluorescence enhancement factor of 90 in Tyrode's solution and 25 in the media) to retain a response to the target analytes. This encouraged the implementation of the CX4/LCG reporter pair to monitor analyte uptake into live cells.

If the CX4/LCG reporter pair was only physically adsorbed on the cell membrane surface or accessible in the outer cell membrane layer, the incubation of the cells with competitors such as choline, acetylcholine (50 mM),<sup>[20]</sup> and protamine (200  $\mu\text{M}$ )<sup>[14]</sup> should lead to an effective release of uncomplexed LCG into the medium because the displacement process itself occurs within milliseconds.<sup>[14]</sup> However, the analyte solution collected after 10–20 min incubation did not show any LCG fluorescence, thus suggesting that the reporter pair was contained within the cells. By contrast, the cells exposed to the analyte solutions showed significantly enhanced fluorescence, which varied depending on the cell line and the analyte type (Figure 1). These fluorescence enhancements demonstrate that the analytes were taken up within 10–15 min into the cells, where they displaced LCG from the macrocyclic complex to result in enhanced cellular fluorescence. Incubation times longer than 20 min led to relatively smaller fluorescence recoveries, and no significant increase was observed after 5 h of incubation, thus suggesting that the reporter pair proceeds from the cytoplasm into different cellular compartments.<sup>[18b]</sup> The staining pattern of the released LCG is punctate, in agreement with previous cellular imaging studies of the dye alone.<sup>[18d]</sup> Although LCG turned out to be photobleached in confocal laser scanning microscopy experiments, the corresponding Z-stacking imaging results (Figure S2) corroborated distribution of the released dye (shortly after analyte addition) through the entire volume of the cell. The fact that dye displacement could also be achieved through cytoplasmic microinjection of the analyte (Figure S8) supports the idea that the initial localization of the CX4/LCG reporter pair is in the cytoplasm.<sup>[18a–d]</sup>

Indeed, the analytes were selected not only to act as strong competitors, but also because of their efficient cellular uptake. Ch is taken up into non-neuronal cells mainly by specific and nonspecific choline transporters,<sup>[24]</sup> while ACh crosses the membrane in both directions through organic cation transporters (OCTs),<sup>[25]</sup> which are expressed in almost all cells.<sup>[26]</sup> The cellular uptake of protamine is established but

mechanistically more diverse, taking place predominantly through rapid endocytosis within 15–30 min,<sup>[27]</sup> the same time scale as that of the fluorescence recovery in our live-cell IDAs. After endosomal escape,<sup>[27b,28]</sup> protamine can interact with the reporter pair in the cytoplasm. It should also be noted that betaine was used as a negative control because it is also known to enter cells by a  $\text{Na}^+$ -dependent active transport mechanism,<sup>[29]</sup> but does not bind to CX4.<sup>[20]</sup> Indeed, this analyte (50 mM) afforded no significant fluorescence response in the live cells (Figure 1, parts b and b'), which demonstrates that the method is analyte-selective. The fluorescence enhancements could be semi-quantitatively analyzed (Figure 1, parts f and f'), and the analysis confirmed that, for both cell lines, betaine caused no significant fluorescence response, ACh and Ch caused about the same fluorescence enhancement, and protamine caused the largest fluorescence recovery, even at much lower concentrations. Although the absolute fluorescence responses are composite effects that depend on several factors (including the concentration of the analyte and its binding constant, see above), live-cell-based IDAs in principle offer the possibility for monitoring uptake at different incubation times.

In conclusion, we have established a receptor/dye IDA system that gives an easily measurable response to bioorganic analytes inside live cells. Live-cell-based IDAs could be employed as a simple, economic screening tool to monitor the uptake efficiencies of closely related compounds with similar affinities, such as different trimethylammonium ions or a library of polycationic peptides, the bioactivity of which is presently under intense discussion.

Received: July 31, 2014

Revised: October 14, 2014

Published online: ■ ■ ■ ■ ■, ■ ■ ■ ■ ■

**Keywords:** bioanalytical chemistry · fluorescence · host-guest complexes · macrocycles · supramolecular chemistry

- [1] a) B. T. Nguyen, E. V. Anslyn, *Coord. Chem. Rev.* **2006**, *250*, 3118–3127; b) E. V. Anslyn, *J. Org. Chem.* **2007**, *72*, 687–699; c) S. K. Sun, K. X. Tu, X. P. Yan, *Analyst* **2012**, *137*, 2124–2128; d) X. J. Liu, H. T. Ngo, Z. J. Ge, S. J. Butler, K. A. Jolliffe, *Chem. Sci.* **2013**, *4*, 1680–1686; e) D. Leung, J. F. Folmer-Andersen, V. M. Lynch, E. V. Anslyn, *J. Am. Chem. Soc.* **2008**, *130*, 12318–12327; f) R. G. Hanshaw, E. J. O'Neil, M. Foley, R. T. Carpenter, B. D. Smith, *J. Mater. Chem.* **2005**, *15*, 2707–2713.
- [2] a) Y. Wang, Z. H. Li, D. H. Hu, C. T. Lin, J. H. Li, Y. H. Lin, *J. Am. Chem. Soc.* **2010**, *132*, 9274–9276; b) X. Sun, Z. Liu, K. Welscher, J. T. Robinson, A. Goodwin, S. Zaric, H. Dai, *Nano Res.* **2008**, *1*, 203–212; c) X. W. Mao, D. M. Tian, H. B. Li, *Chem. Commun.* **2012**, *48*, 4851–4853; d) X. W. Mao, H. B. Li, *J. Mater. Chem. B* **2013**, *1*, 4267–4272.
- [3] a) S. Angelos, Y. W. Yang, K. Patel, J. F. Stoddart, J. I. Zink, *Angew. Chem. Int. Ed.* **2008**, *47*, 2222–2226; *Angew. Chem.* **2008**, *120*, 2254–2258; b) A. Agostini, L. Mondragón, A. Bernardos, R. Martínez-Mañez, M. D. Marcos, F. Sancenón, J. Soto, A. Costero, C. Manguan-García, R. Perona, M. Moreno-Torres, R. Aparicio-Sanchis, J. R. Murguía, *Angew. Chem. Int. Ed.* **2012**, *51*, 10556–10560; *Angew. Chem.* **2012**, *124*, 10708–10712; c) Y. Zhou, L. L. Tan, Q. L. Li, X. L. Qiu, A. D. Qi, Y. C. Tao, Y. W. Yang, *Chem. Eur. J.* **2014**, *20*, 2998–3004; d) D. G. He,

- X. X. He, K. M. Wang, Z. Zou, X. Yang, X. C. Li, *Langmuir* **2014**, *30*, 7182–7189.
- [4] a) B. J. Shorthill, C. T. Avetta, T. E. Glass, *J. Am. Chem. Soc.* **2004**, *126*, 12732–12733; b) S. H. Zhang, T. E. Glass, *Tetrahedron Lett.* **2010**, *51*, 112–114.
- [5] a) Q. P. Duan, Y. Cao, Y. Li, X. Y. Hu, T. X. Xiao, C. Lin, Y. Pan, L. Y. Wang, *J. Am. Chem. Soc.* **2013**, *135*, 10542–10549; b) J. Zhang, Z. F. Yuan, Y. Wang, W. H. Chen, G. F. Luo, S. X. Cheng, R. X. Zhuo, X. Z. Zhang, *J. Am. Chem. Soc.* **2013**, *135*, 5068–5073; c) J. J. Yin, S. Sharma, S. P. Shumyak, Z. X. Wang, Z. W. Zhou, Y. D. Zhang, P. X. Guo, C. Z. Li, J. R. Kanwar, T. X. Yang, S. S. Mohapatra, W. Q. Liu, W. Duan, J. C. Wang, Q. Li, X. J. Zhang, J. Tan, L. Jia, J. Liang, M. Q. Wei, X. T. Li, S. F. Zhou, *PLoS One* **2013**, *8*, e62289.
- [6] Z. Y. Li, S. G. Sun, Z. G. Yang, S. Zhang, H. Zhang, M. M. Hu, J. F. Cao, J. Y. Wang, F. Y. Liu, F. L. Song, J. L. Fan, X. J. Peng, *Biomaterials* **2013**, *34*, 6473–6481.
- [7] a) V. Bagnacani, V. Franceschi, M. Bassi, M. Lomazzi, G. Donofrio, F. Sansone, A. Casnati, R. Ungaro, *Nat. Commun.* **2013**, *4*, 1721; b) L. Nault, A. Cumbo, R. F. Pretôt, M. A. Sciotti, P. Shahgaldian, *Chem. Commun.* **2010**, *46*, 5581–5583.
- [8] a) P. Neirynek, J. Brinkmann, Q. An, D. W. J. van der Schaft, L. G. Milroy, P. Jonkheijm, L. Brunsveld, *Chem. Commun.* **2013**, *49*, 3679–3681; b) Q. An, J. Brinkmann, J. Huskens, S. Krabbenborg, J. de Boer, P. Jonkheijm, *Angew. Chem. Int. Ed.* **2012**, *51*, 12233–12237; *Angew. Chem.* **2012**, *124*, 12399–12403.
- [9] a) D. Bier, R. Rose, K. Bravo-Rodriguez, M. Bartel, J. M. Ramirez-Angueta, S. Dutt, C. Wilch, F. G. Klarner, E. Sanchez-Garcia, T. Schrader, C. Ottmann, *Nat. Chem.* **2013**, *5*, 234–239; b) D. W. Lee, K. M. Park, M. Banerjee, S. H. Ha, T. Lee, K. Suh, S. Paul, H. Jung, J. Kim, N. Selvapalam, S. H. Ryu, K. Kim, *Nat. Chem.* **2011**, *3*, 154–159; c) S. A. Minaker, K. D. Daze, M. C. F. Ma, F. Hof, *J. Am. Chem. Soc.* **2012**, *134*, 11674–11680.
- [10] S. S. Agasti, M. Liong, C. Tassa, H. J. Chung, S. Y. Shaw, H. Lee, R. Weissleder, *Angew. Chem. Int. Ed.* **2012**, *51*, 450–454; *Angew. Chem.* **2012**, *124*, 465–469.
- [11] a) A. Hennig, H. Bakirci, W. M. Nau, *Nat. Methods* **2007**, *4*, 629–632; b) D. M. Bailey, A. Hennig, V. D. Uzunova, W. M. Nau, *Chem. Eur. J.* **2008**, *14*, 6069–6077; c) W. M. Nau, G. Ghale, A. Hennig, H. Bakirci, D. M. Bailey, *J. Am. Chem. Soc.* **2009**, *131*, 11558–11570.
- [12] a) V. Francisco, A. Pineiro, W. M. Nau, L. Garcia-Rio, *Chem. Eur. J.* **2013**, *19*, 17809–17820; b) K. D. Daze, C. E. Jones, B. J. Lilgert, C. S. Beshara, F. Hof, *Can. J. Chem.* **2013**, *91*, 1072–1076.
- [13] a) R. N. Dsouza, A. Hennig, W. M. Nau, *Chem. Eur. J.* **2012**, *18*, 3444–3459; b) D. A. Jose, M. Elstner, A. Schiller, *Chem. Eur. J.* **2013**, *19*, 14451–14457; c) B. Vilozny, A. Schiller, R. A. Wesling, B. Singaram, *Anal. Chim. Acta* **2009**, *649*, 246–251; d) T. Desmet, W. Soetaert, P. Bojarova, V. Kren, L. Dijkhuizen, V. Eastwick-Field, A. Schiller, *Chem. Eur. J.* **2012**, *18*, 10786–10801.
- [14] G. Ghale, A. G. Lanctot, H. T. Kreissl, M. H. Jacob, H. Weingart, M. Winterhalter, W. M. Nau, *Angew. Chem. Int. Ed.* **2014**, *53*, 2762–2765; *Angew. Chem.* **2014**, *126*, 2801–2805.
- [15] K. S. Hettie, X. Liu, K. D. Gillis, T. E. Glass, *ACS Chem. Neurosci.* **2013**, *4*, 918–923.
- [16] a) D. Kim, S. Singha, T. Wang, E. Seo, J. H. Lee, S. J. Lee, K. H. Kim, K. H. Ahn, *Chem. Commun.* **2012**, *48*, 10243–10245; b) E. L. Que, D. W. Domaille, C. J. Chang, *Chem. Rev.* **2008**, *108*, 4328–4328; c) D. W. Domaille, E. L. Que, C. J. Chang, *Nat. Chem. Biol.* **2008**, *4*, 507–507; d) A. K. Mahapatra, J. Roy, P. Sahoo, S. K. Mukhopadhyay, A. Chattopadhyay, *Org. Biomol. Chem.* **2012**, *10*, 2231–2236; e) G. Sivaraman, T. Anand, D. Chellappa, *RSC Adv.* **2013**, *3*, 17029–17033; f) J. L. Wang, L. P. Long, D. Xie, Y. W. Zhan, *J. Lumin.* **2013**, *139*, 40–46; g) M. X. Yu, M. Shi, Z. G. Chen, F. Y. Li, X. X. Li, Y. H. Gao, J. Xu, H. Yang, Z. G. Zhou, T. Yi, C. H. Huang, *Chem. Eur. J.* **2008**, *14*, 6892–6900.
- [17] a) K. Cui, Z. L. Chen, Z. Wang, G. X. Zhang, D. Q. Zhang, *Analyst* **2011**, *136*, 191–195; b) H. Kwon, K. Lee, H. J. Kim, *Chem. Commun.* **2011**, *47*, 1773–1775; c) C. Huang, T. Jia, M. Tang, Q. Yin, W. Zhu, C. Zhang, Y. Yang, N. Jia, Y. Xu, X. Qian, *J. Am. Chem. Soc.* **2014**, *136*, 14237–14244.
- [18] a) R. Lalor, H. Baillie-Johnson, C. Redshaw, S. E. Matthews, A. Mueller, *J. Am. Chem. Soc.* **2008**, *130*, 2892–2893; b) A. Mueller, R. Lalor, C. M. Cardaba, S. E. Matthews, *Cytometry Part A* **2011**, *79A*, 126–136; c) C. Redshaw, M. R. J. Elsegood, J. A. Wright, H. Baillie-Johnson, T. Yamato, S. De Giovanni, A. Mueller, *Chem. Commun.* **2012**, *48*, 1129–1131; d) S. J. Rembish, M. A. Trush, *Free Radical Biol. Med.* **1994**, *17*, 117–126; e) A. W. Coleman, S. Jebors, S. Cecillon, P. Perret, D. Garin, D. Marti-Battle, M. Moulin, *New J. Chem.* **2008**, *32*, 780–782; f) D.-S. Guo, Y. Liu, *Acc. Chem. Res.* **2014**, *47*, 1925–1934.
- [19] a) F. Perret, A. W. Coleman, *Chem. Commun.* **2011**, *47*, 7303–7319; b) M. H. Paquet, C. F. Rousseau, C. Yannick, F. Morel, A. W. Coleman, *J. Inclusion Phenom. Macrocyclic Chem.* **2006**, *55*, 353–357; c) K. Wang, D. S. Guo, X. Wang, Y. Liu, *ACS Nano* **2011**, *5*, 2880–2894.
- [20] D. S. Guo, V. D. Uzunova, X. Su, Y. Liu, W. M. Nau, *Chem. Sci.* **2011**, *2*, 1722–1734.
- [21] V79 cells showed more efficient uptake of the dye than CHO cells (Table 1), which could be due, among other factors, to different surface-to-volume ratios of the two cell types ( $0.65 \mu\text{m}^{-1}$  for V79 and  $0.37 \mu\text{m}^{-1}$  for CHO cells, see the Supporting Information for size estimations).
- [22] While LCG has been demonstrated to undergo cellular uptake by adsorptive endocytosis: I. Braakman, T. Pijning, O. Verest, B. Weert, D. K. F. Meijer, G. M. M. Groothuis, *Mol. Pharmacol.* **1989**, *36*, 537–542, there are several potential uptake routes for the CX4/LCG complex (see the Supporting Information).
- [23] R. N. Dsouza, U. Pischel, W. M. Nau, *Chem. Rev.* **2011**, *111*, 7941–7980.
- [24] a) S. A. Müller, K. Holzappel, C. Seidl, U. Treiber, B. J. Krause, R. Senekowitsch-Schmidtke, *Eur. J. Nucl. Med. Mol. Imaging* **2009**, *36*, 1434–1442; b) P. R. Lockman, D. D. Allen, *Drug Dev. Ind. Pharm.* **2002**, *28*, 749–771; c) P. Schloss, W. Mayser, A. Niehuis, H. Betz, *Biochem. Biophys. Res. Commun.* **1994**, *199*, 1320–1325.
- [25] K. S. Lips, C. Volk, B. M. Schmitt, U. Pfeil, P. Arndt, D. Miska, L. Ermert, W. Kummer, H. Koepsell, *Am. J. Respir. Cell Mol. Biol.* **2005**, *33*, 79–88.
- [26] I. Wessler, C. J. Kirkpatrick, *Br. J. Pharmacol.* **2008**, *154*, 1558–1571.
- [27] a) H. Amos, *Biochem. Biophys. Res. Commun.* **1961**, *5*, 1–4; b) H. Amos, K. E. Kearns, *Exp. Cell Res.* **1963**, *32*, 14–32; c) J. Nagai, T. Komeda, Y. Katagiri, R. Yumoto, M. Takano, *Biol. Pharm. Bull.* **2013**, *36*, 1942–1949.
- [28] a) J. Guy, D. Drabek, M. Antoniou, *Mol. Biotechnol.* **1995**, *3*, 237–248; b) Y. W. Cho, J. D. Kim, K. Park, *J. Pharm. Pharmacol.* **2003**, *55*, 721–734.
- [29] P. G. Petronini, E. Deangelis, A. F. Borghetti, K. P. Wheeler, *Biochem. J.* **1994**, *300*, 45–50.

## Communications

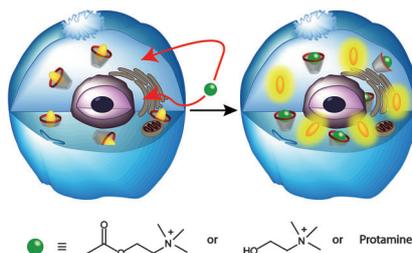


### Bioanalytical Chemistry

A. Norouzy, Z. Azizi,  
W. M. Nau\* —————



Indicator Displacement Assays Inside  
Live Cells



**Live coverage:** Macrocyclic host–dye complexes can be used to detect the presence of an analyte inside live cells through dye displacement. *p*-Sulfonato-calix[4]arene (CX4) and the fluorescent dye lucigenine (LCG, yellow) are both taken up into CHO and V79 cells to form a weakly fluorescent CX4–LCG complex in the cytoplasm. Subsequent addition of the corresponding analytes (acetylcholine, choline, and protamine; green) results in a clear increase in fluorescence.

## Supporting Information

### Indicator Displacement Assays inside Live Cells

Amir Norouzy,<sup>[a]</sup> Zahra Azizi,<sup>[b]</sup> and Werner M. Nau<sup>[a],\*</sup>

<sup>[a]</sup> School of Engineering and Science, Jacobs University Bremen, Campus Ring 1, D-28759 Bremen (Germany).

<sup>[b]</sup> Centre for Biomolecular Interactions, University of Bremen, Leobener Straße NW2, D-28359 Bremen (Germany).

\* To whom correspondence should be addressed. E-mail: w.nau@jacobs-university.de

#### • Table of Contents

---

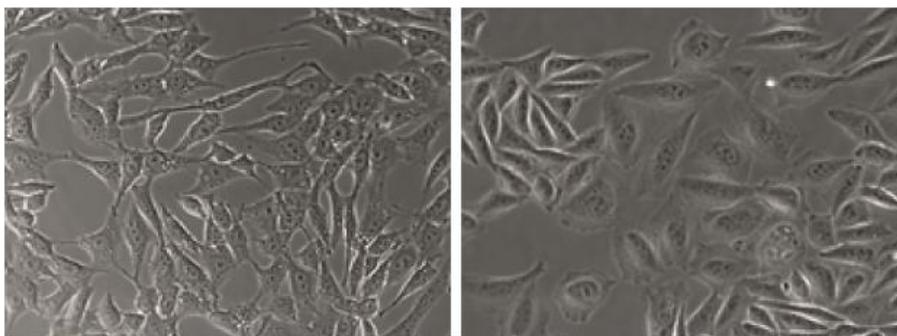
1.	Materials .....	S2
2.	Cell Cultivation and Preparation for Imaging.....	S2
3.	Loading with Host and Dye and Subsequent Dye Displacement by Analytes .....	S2
4.	Microscopy .....	S3
4.1.	Fluorescence microscopy.....	S3
4.2.	Confocal Laser Scanning Microscopy (CLSM) .....	S3
5.	Image Processing and Analysis .....	S4
6.	Calculation of Uptake of LCG in the Absence and Presence of CX4 .....	S5
7.	Potential Uptake Routes of LCG in the Absence and Presence of CX4.....	S6
8.	Gradual Decrease of LCG Fluorescence in Cells .....	S7
9.	Dye Release by Microinjection.....	S8
10.	Effects at High Protamine Concentrations.....	S9
11.	References.....	S10

## 1. Materials

Frozen Chinese hamster ovary (CHO) cells were purchased from Sigma-Aldrich. Chinese hamster fibroblast (V79) cells were kindly made available by Detlef Gabel, Jacobs University Bremen. Cultivation dishes (60 cm<sup>2</sup>) and Falcon tubes as well as penicillin/streptomycin and trypsin (1000 U ml<sup>-1</sup>) were purchased from Biochrom AG. Ham's F10 and F12 media, fetal bovine serum (FBS), sterile glutamine solution, acetylcholine perchlorate, and protamine sulfate (from herring) were purchased from Sigma-Aldrich. Imaging  $\mu$ -dishes (35 mm, high wall and low wall, for micro injection) were purchased from *ibidi*. Choline chloride was from Fluka.

## 2. Cell Cultivation and Preparation for Imaging

V79 cells (Figure S1) were cultivated in Ham's F10 medium, 10% FBS, while CHO cells (Figure S1) were grown in Ham's F12 medium, 10% FBS, and 2 mM L-Glutamine. Dishes were kept in a humid, 5% CO<sub>2</sub> incubator at 37 °C until the cells reached a confluence of about 90%. The cells were subsequently trypsinized and diluted by 1:10 into the *ibidi* dishes followed by 2-3 days incubation until 60-80% confluent monolayers were obtained.



**Figure S1.** Bright-field images with 40X object magnification of V79 cells (left) and CHO cells (right).

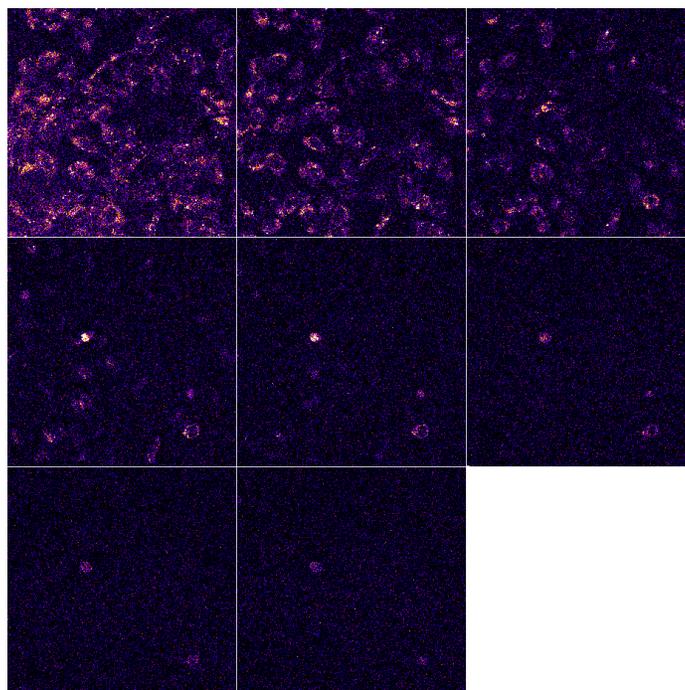
## 3. Loading with Host and Dye and Subsequent Dye Displacement by Analytes

For loading with CX4/LCG, the V79 and CHO cells in the *ibidi* dishes were incubated for 15 min at 37 °C with a mixture of 50  $\mu$ M LCG and 250-300  $\mu$ M CX4, dissolved in medium. An excess of host was needed for efficient fluorescence quenching, as judged from the cell images (Figure S6). The cells were then washed 3 times with fresh medium and incubated for 10-15 minutes at 37 °C with 50 mM ACh, 50 mM Ch, or 200  $\mu$ M (0.9 mg ml<sup>-1</sup>) protamine, dissolved in medium pre-warmed to 37°C. Subsequently, the supernatant analyte solutions were collected and tested for LCG to confirm that the dye was retained in the cells (and not, for example, weakly adsorbed on the outer cell membrane). Finally, the cells were washed 3 times with Tyrode's solution and immediately used for imaging under Tyrode's solution (Figure 1).

## 4. Microscopy

### 4.1. Fluorescence microscopy

Cell images were taken in a dark room on an Axiovert 200 microscope (Zeiss) equipped with an Evolution™ QEi camera with 1.4 megapixels and 1392×1040 dimension. Filter set 44 (Zeiss) in combination with a 500-nm beam splitter was used for excitation at 435-500 nm and emission at 480-580 nm. To reduce photobleaching of LCG, the snapshot was taken immediately upon exposure to the excitation light, using identical light exposure (500 ms, unless stated differently) and gain settings.



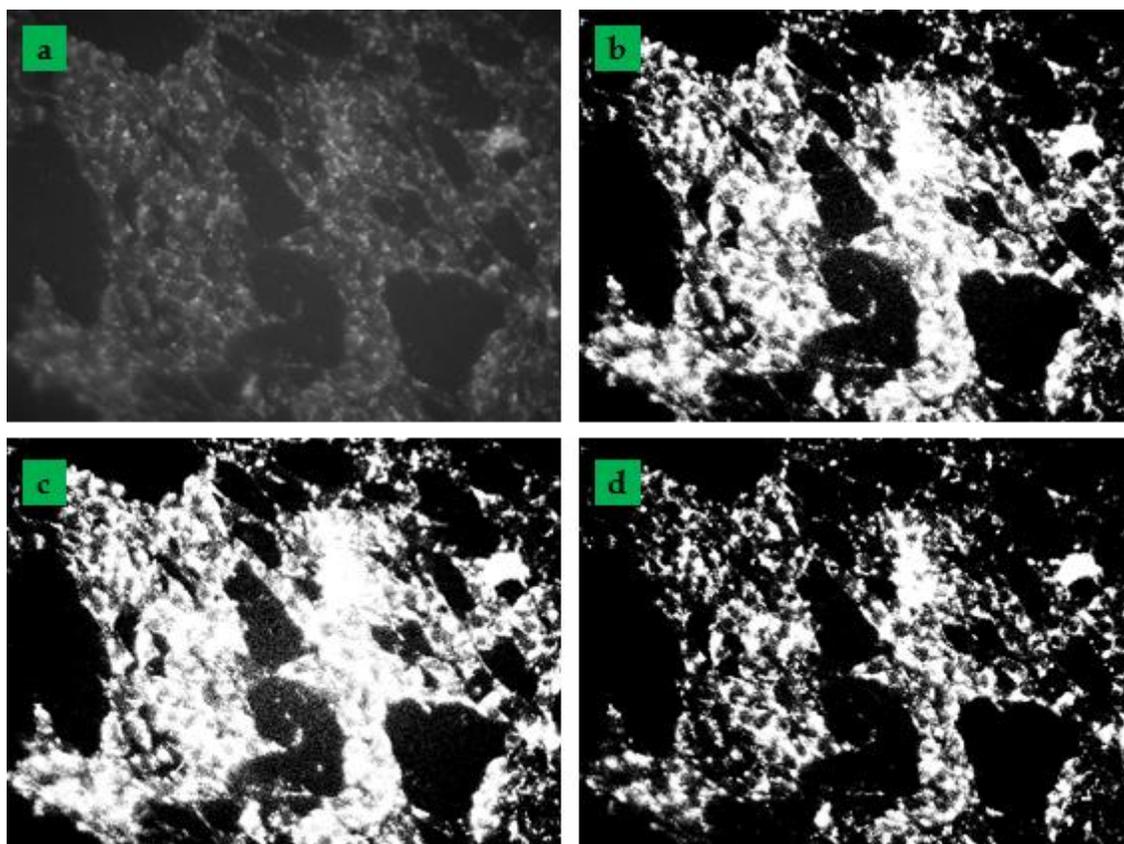
**Figure S2.** Fluorescence images of a V79 cell monolayer were recorded by CLSM 3D scanning after incubation with CX4/LCG and subsequently 50 mM Ch. From left to right in each row, one layer of the cells was scanned from bottom of the dish to the top (Z stacking).

### 4.2 Confocal Laser Scanning Microscopy (CLSM)

The uptake of Ch was also monitored by 3D scanning using a confocal laser scanning microscope (Zeiss LSM 780) by 405-nm laser excitation (Figure S2). Since the dye was significantly photobleached already after the third scan, the confocal measurements were not preferred for the comparative measurements of analyte uptake. Nevertheless, vertical scanning, from bottom to top, showed the distribution of the dye in the entire volume of the cells. This provided circumstantial evidence that the dye does not accumulate in the cell membrane but rather distributes in the cytoplasm with a punctate pattern.

## 5. Image Processing and Analysis

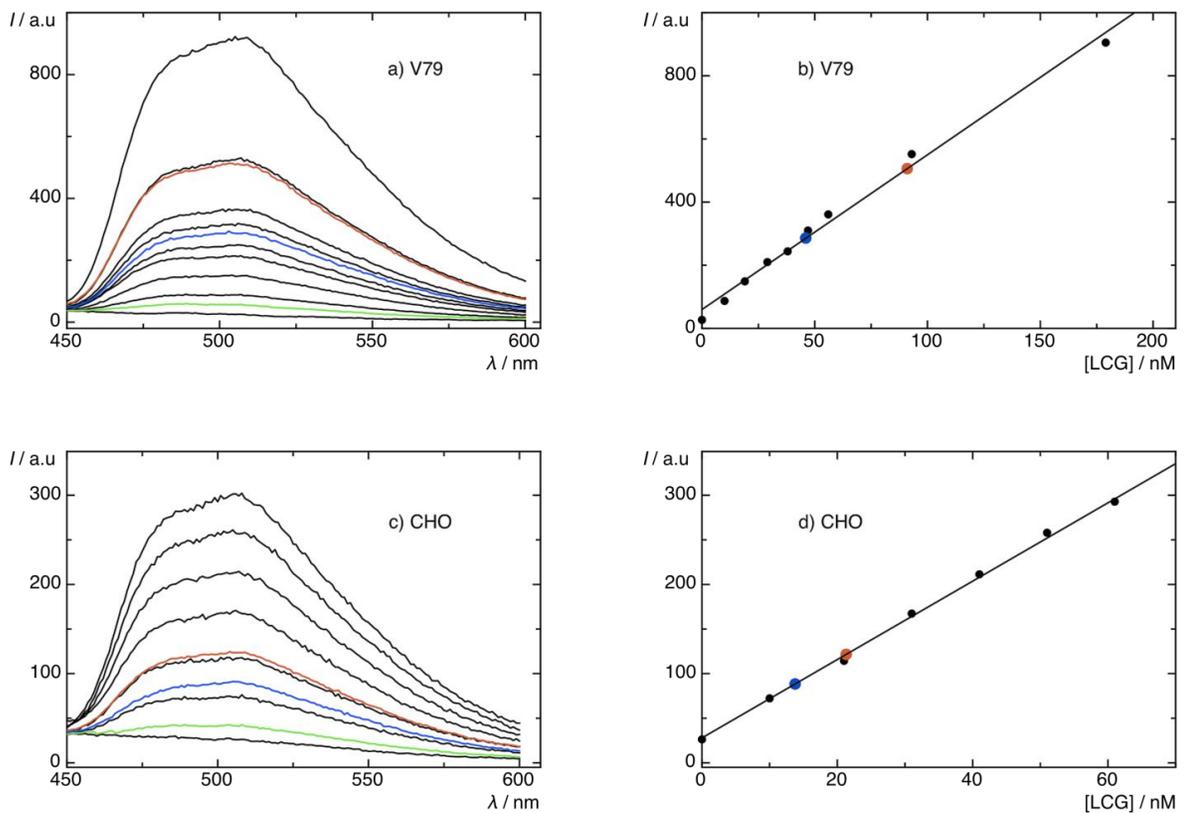
The 16-bit fluorescence-microscopic images were added to an imageJ stack<sup>[1]</sup> with normalized background for further processing (Figure 1, selected color was “fire”). The average fluorescence intensities (Figure 1 f and f') were obtained by binning of the entire fluorescence images in 8-bit type, and by choosing the threshold manually (the selection of automated thresholding procedures provided qualitatively comparable results, but with lower contrast, see Figure S3). The abundance of each histogram bar was multiplied with the relative intensity (between the threshold value and the maximum value of 255), and finally, divided by the number of cells (counted manually). The error bars correspond to the standard deviation from the mean value calculated by the imageJ program for each individual image.



**Figure S3.** Fluorescence images of V79 cells. An 8-bit image (a, raw data) was subjected to different thresholding methods: Otsu<sup>[2]</sup> (b), Li<sup>[3]</sup> (c), and manually (d).

## 6. Calculation of Uptake of LCG in the Absence and Presence of CX4

Comparable numbers of V79 or CHO cells were cultivated in three 60 cm<sup>2</sup> dishes. In Dish 1 the cells were incubated with 800 μM LCG, in Dish 2 they were incubated with CX4/LCG ([LCG] = 800 μM, [CX4]/[LCG] = 5/1), and Dish 3 was used as a reference, without additive incubation (Figures S4b,d). After 30 min incubation, Dish 1 and 2 were washed three times with warm medium (37 °C) to remove residual LCG and CX4/LCG. Following trypsinization, the homogenously suspended cells were counted on a cytometer to estimate the total number of cells on each dish. For dye content analysis, the suspended cells were lysed with Triton X-100 to collect their cytoplasmic components.<sup>[4]</sup>



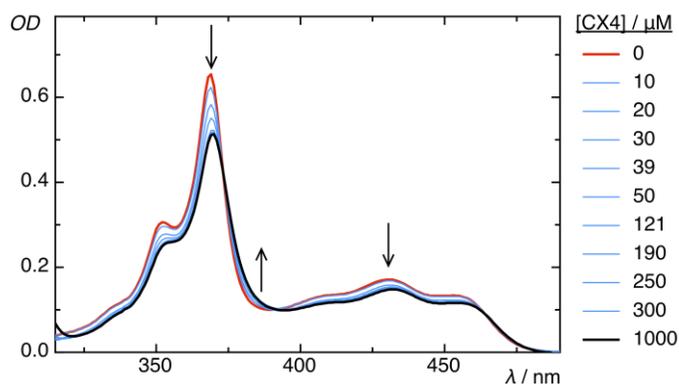
**Figure S4.** a,c) Fluorescence spectra of lysed V79 and CHO cells after incubation with LCG (800 μM) in the absence (blue) and presence of a 5-fold excess of CX4 (green) and after addition of excess protamine (4 mM, red). The black spectra are for different LCG concentrations in the lysed reference solutions. b,d) Calibration curves, obtained from the fluorescence maxima ( $\lambda_{\max} = 505$  nm) of the black spectra in a,c); the blue and red data points mark the interpolated LCG concentrations in the absence (blue) and presence of a 5-fold excess of CX4 after addition of excess protamine (4 mM, red), respectively.

Fluorescence spectra of all lysates ( $\lambda_{\text{exc}} = 369 \text{ nm}$ ) were obtained after dilution to 3 ml (to allow cuvette measurement) and filtering (0.20  $\mu\text{m}$  syringe filter, tested not to adsorb the dye). In a second step, an excess of protamine (4 mM) was added to the obtained lysed solutions from all dishes in order to affect a complete dissociation of the CX4/LCG complex, resulting in a significant increase due to dye release for Dish 2 only (from green to red spectra in Figure S4a,c); incidentally, this experiment corroborates the presence of CX4 in the cytoplasm. In contrast, the fluorescence intensity in Dish 1, which contained no CX4, remained unchanged, because control experiments showed that the direct addition of competitors (up to 10 mM Ch or ACh and 4 mM protamine) cause no significant fluorescence quenching. To determine the absolute uptake of LCG, different LCG concentrations were added to the lysed *reference* solutions (Dish 3, without additives). This allowed us to construct calibration curves (Figure S4, right) for the dye fluorescence intensity in this medium (black spectra in Figure S4a,c). The absolute concentrations of LCG after uptake into the cells (Table 1 of main text) were estimated by employing spherical cell volumes derived from reported average cell diameters (9.3  $\mu\text{m}$  for V79 and 14.0  $\mu\text{m}$  for CHO).<sup>[5]</sup>

## 7. Potential Uptake Routes of LCG in the Absence and Presence of CX4

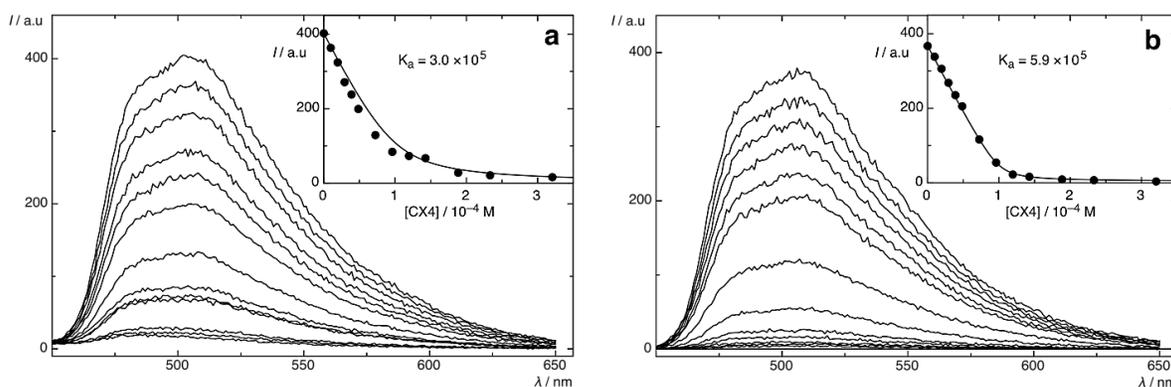
Due to the established use of LCG, among others for the detection of mitochondrial superoxide generation in cells<sup>[6]</sup> its cellular uptake mechanism has already been investigated in detail and identified as proceeding by adsorptive endocytosis.<sup>[7]</sup> The presence of CX4 results not only in an uptake of the macrocycle but also an enhanced uptake of LCG (see main text), the mechanism for which is yet not clarified in detail.

Although CX4 by itself does not permeate through lipid bilayer membranes,<sup>[8]</sup> it has been demonstrated that it can pass through anion channels under certain conditions.<sup>[9]</sup> Moreover, especially in the presence of hydrophobic guests,<sup>[10]</sup> CX4 is able to form charge-neutralized host-guest aggregates, which have been shown to facilitate cellular uptake.<sup>[10b]</sup> To evaluate a possible aggregate formation, we have investigated the host-guest system by UV titrations, which, however, did not reveal any light scattering in the typical concentration range, but, instead, clearly defined isosbestic points at 373, 392 and 464 nm (Figure S5). Similarly, dynamic light scattering experiments did not afford any evidence for the formation of large aggregates. We therefore have at present no evidence for the involvement of CX4/LCG aggregates. Frequently, the formation of such aggregates is facilitated by  $\pi$ - $\pi$  stacking interactions between guest molecules (for a review of guests promoting aggregate formation see ref.<sup>[10a]</sup>). For LCG, this possibility is prevented by the twisted arrangements of the aryl groups, which is also evident from the experimental crystal structure.<sup>[11]</sup>



**Figure S5.** UV-Vis titration of 50  $\mu\text{M}$  LCG with CX4 (up to 1 mM); the arrows depict the trend of the  $OD$  changes in the different spectral regions.

It should also be noted that the uptake mechanism of several labeled calixarene derivatives into different cells, including CHO, THP-1 and HeLa, has been extensively studied by Mueller and coworkers.<sup>[12]</sup> A non-specific, membrane carrier-mediated uptake with an initial homogeneous distribution in the cytoplasm, followed by appearance in the Golgi and, finally, in the acidic vesicles was suggested.

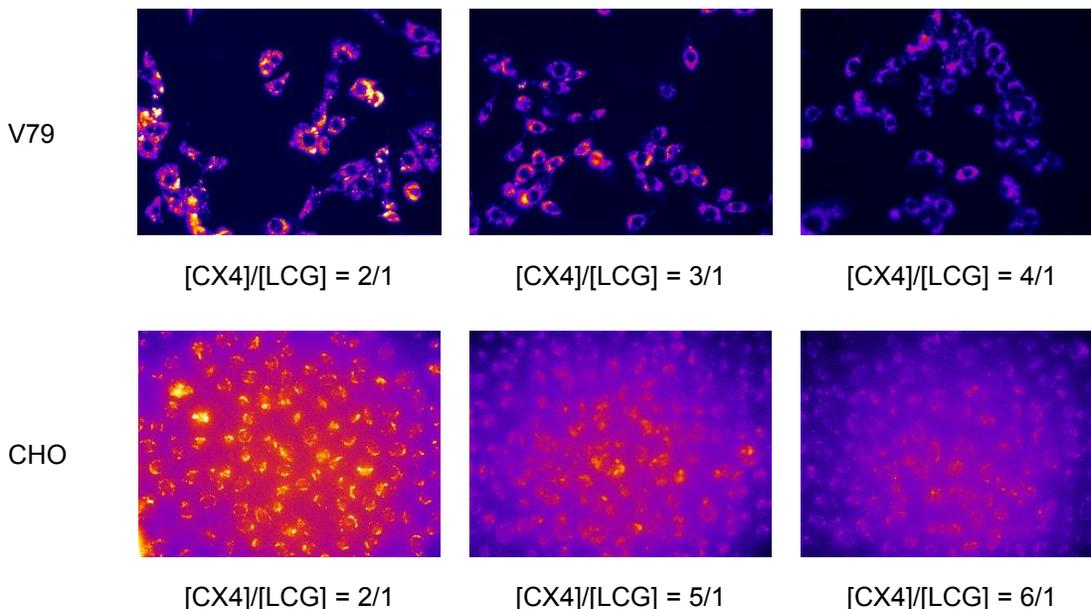


**Figure S6.** Direct fluorescence titration of LCG (100  $\mu\text{M}$ ) with CX4 in Ham's medium (a) and in Tyrode's solution (b)  $\text{pH}=7$ ,  $\lambda_{\text{ex}}=369$  nm. The inset shows the associated titration curve, and nonlinear fitting according to a 1:1 binding stoichiometry.

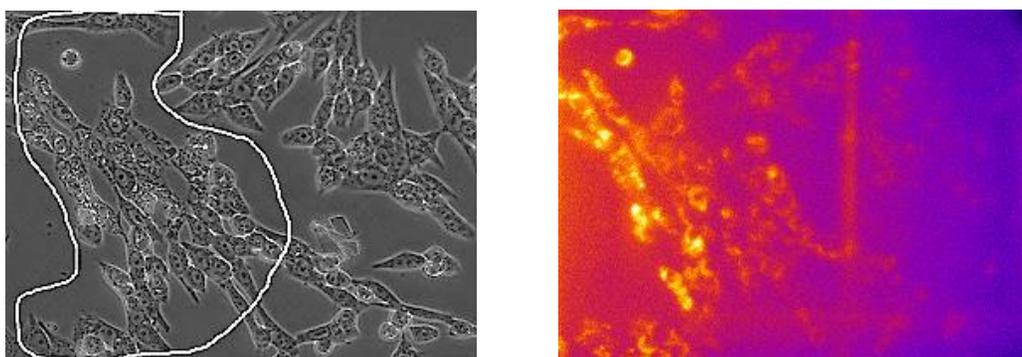
## 8. Gradual Decrease of LCG Fluorescence in Cells

Figure S6 shows direct titrations of LCG by CX4 in Ham's medium<sup>[13]</sup> and in Tyrode's solution<sup>[14]</sup> and the determinations of the apparent binding constants (insets). Due to the presence of ca. 150 mM salts, nutrients, amino acids, vitamins, 10% fetal bovine serum, etc. in Ham's medium, the host-dye binding constant is reduced in the medium compared to neat water (see main text), but remains large ( $> 3 \times 10^5 \text{ M}^{-1}$ ). Based on the titrations, an excess of CX4 needed to be employed to ensure efficient quenching of LCG fluorescence and, thus, to facilitate a readily

detectable fluorescence recovery upon analyte uptake. Figure S7 shows the fluorescence of V79 and CHO cells incubated with different host/dye ratios. A 4:1 ratio for V79 and 6:1 ratio for CHO was preferable to affect a strong (> 50%) fluorescence quenching.



**Figure S7.** Fluorescence of V79 and CHO cells incubated with different superstoichiometric amounts of host at a constant dye concentration (50 μM); light exposure was 100 ms for V79 cells and 500 ms for CHO cells.



**Figure S8.** Bright-field (left) and fluorescence image (right) of V79 cells incubated with CX4/LCG solution ([LCG] = 10 μM, [CX4]/[LCG] = 5/1, 10 min) and after microinjection with ACh (area with injected cells encircled on the left).

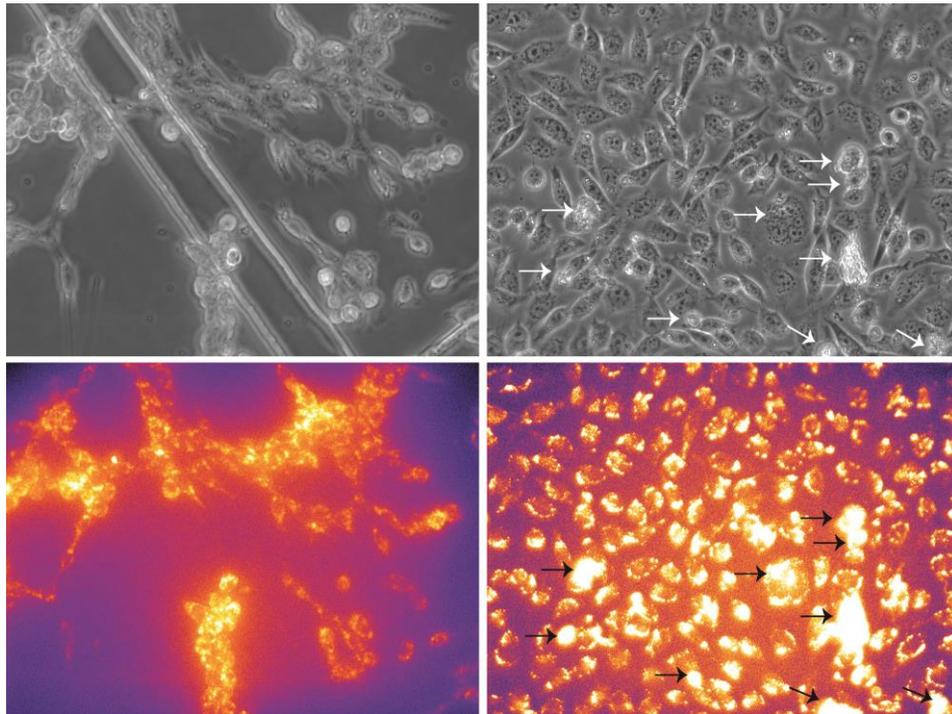
## 9. Dye Release by Microinjection

For microinjection experiments, V79 cells were incubated with CX4/LCG solution, washed with Tyrode's solution, and 100 mM ACh solution was injected into the cytoplasm of selected cells by

using a FemtoJet Eppendorf microinjector (60 hpa capillary pressure, injection pressure 190 hpa, 3 seconds). A strong fluorescence recovery was observed (Figure S8). Since this method was prone to cellular damage and because the timing between incubation and injection turned out to be critical but difficult to ensure by this method, it was not preferred for follow-up experiments.

## 10. Effects at High Protamine Concentrations

It is known that high (mM) concentrations of protamine can damage cells by causing membrane rupture. Therefore, we also applied high concentrations and, indeed, dead cells were observed (Figure S9). Interestingly, these cells showed the largest fluorescence enhancements (see arrows in Figure S9), presumably due to additional release of dye from the endosomes.<sup>[12b]</sup> The comparison of two cell lines showed further that V79 cells are more sensitive to protamine than CHO cells, since almost all V79 cells were adversely affected. Analytes causing cell lysis do therefore need to be investigated with particular care to avoid false positive results.



**Figure S9.** Bright-field (top) and fluorescence images of V79 cells (left column) and CHO cells (right), both incubated with 50  $\mu$ M LCG and 250-300  $\mu$ M CX4 at 37  $^{\circ}$ C for 15 min, followed by a 20-min incubation with 4 mM protamine. Dead CHO cells resulting from protamine-induced cell damage are marked by arrows.

## 11. References

- [1] C. A. Schneider, W. S. Rasband, K. W. Eliceiri, *Nat. Methods* **2012**, *9*, 671-675.
- [2] J. H. Xue, Y. J. Zhang, *Pattern Recognit. Lett.* **2012**, *33*, 793-797.
- [3] C. H. Li, P. K. S. Tam, *Pattern Recognit. Lett.* **1998**, *19*, 771-776.
- [4] H. Ji, *Cold Spring Harb. Protoc.* **2010**, *2010*, pdb.prot5466.
- [5] a) D. Bettega, P. Calzolari, S. M. Doglia, B. Dulio, L. Tallone, A. M. Villa, *Int. J. Radiat. Biol.* **1998**, *74*, 397-403; b) Y. Han, X. M. Liu, H. Liu, S. C. Li, B. C. Wu, L. L. Ye, Q. W. Wang, Z. L. Chen, *J Biosci Bioeng* **2006**, *102*, 430-435.
- [6] S. J. Rembish, M. A. Trush, *Free Radical Biol. Med.* **1994**, *17*, 117-126.
- [7] I. Braakman, T. Pijning, O. Verest, B. Weert, D. K. F. Meijer, G. M. M. Groothuis, *Mol. Pharmacol.* **1989**, *36*, 537-542.
- [8] a) Y. L. Ying, J. J. Zhang, F. N. Meng, C. Cao, X. Y. Yao, I. Willner, H. Tian, Y. T. Long, *Sci. Rep.* **2013**, *3*, 1662; b) B. J. Laventie, C. Potrich, C. Atmanene, M. Saleh, O. Joubert, G. Viero, C. Bachmeyer, V. Antonini, I. Mancini, S. Cianferani-Sanglier, D. Keller, D. A. Colin, T. Bourcier, G. Anderluh, A. van Dorselaer, M. Dalla Serra, G. Prevost, *Biochem. J.* **2013**, *450*, 559-571; c) G. Ghale, A. G. Lanctot, H. T. Kreissl, M. H. Jacob, H. Weingart, M. Winterhalter, W. M. Nau, *Angew. Chem. Int. Ed.* **2014**, *53*, 2762-2765.
- [9] G. Droogmans, J. Prenen, J. Eggermont, T. Voets, B. Nilius, *Am. J. Physiol.-Cell Ph.* **1998**, *275*, C646-C652.
- [10] a) D.-S. Guo, Y. Liu, *Acc. Chem. Res.* **2014**, *47*, 1925-1934; b) Patents: FR2904782-A1; WO2008020126-A2; WO2008020126-A3; EP2073783-A2; CA2660295-A1; US2010292272-A1.
- [11] D. S. Guo, V. D. Uzunova, X. Su, Y. Liu, W. M. Nau, *Chem. Sci.* **2011**, *2*, 1722-1734.
- [12] a) R. Lalor, H. Baillie-Johnson, C. Redshaw, S. E. Matthews, A. Mueller, *J. Am. Chem. Soc.* **2008**, *130*, 2892-2893; b) A. Mueller, R. Lalor, C. M. Cardaba, S. E. Matthews, *Cytom. Part A* **2011**, *79A*, 126-136; c) C. Redshaw, M. R. J. Elsegood, J. A. Wright, H. Baillie-Johnson, T. Yamato, S. De Giovanni, A. Mueller, *Chem. Commun.* **2012**, *48*, 1129-1131.
- [13] R. G. Ham, *Exp. Cell Res.* **1963**, *29*, 515-526.
- [14] 137 mM NaCl, 2.7 mM KCl, 1 mM MgCl<sub>2</sub>, 1.8 mM CaCl<sub>2</sub>, 0.2 mM Na<sub>2</sub>HPO<sub>4</sub>, 12 mM NaHCO<sub>3</sub>, 5.5 mM D-glucose. The Tyrode's solution was prepared according to the receipt NO 10479 from Cold Spring Harbor Protocols (doi:10.1101/pdb.rec10479).

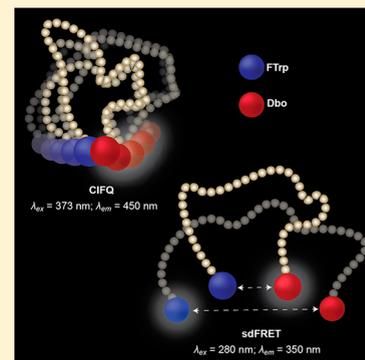
# Coulomb Repulsion in Short Polypeptides

Amir Norouzy, Khaleel I. Assaf, Shuai Zhang, Maik H. Jacob,\* and Werner M. Nau\*

Department of Life Sciences and Chemistry, Jacobs University Bremen, Campus Ring 1, D-28759 Bremen, Germany

## Supporting Information

**ABSTRACT:** Coulomb repulsion between like-charged side chains is presently viewed as a major force that impacts the biological activity of intrinsically disordered polypeptides (IDPs) by determining their spatial dimensions. We investigated short synthetic models of IDPs, purely composed of ionizable amino acid residues and therefore expected to display an extreme structural and dynamic response to pH variation. Two synergistic, custom-made, time-resolved fluorescence methods were applied in tandem to study the structure and dynamics of the acidic and basic hexapeptides Asp<sub>6</sub>, Glu<sub>6</sub>, Arg<sub>6</sub>, Lys<sub>6</sub>, and His<sub>6</sub> between pH 1 and 12. (i) End-to-end distances were obtained from the short-distance Förster resonance energy transfer (sdFRET) from N-terminal 5-fluoro-L-tryptophan (FTrp) to C-terminal Dbo. (ii) End-to-end collision rates were obtained for the same peptides from the collision-induced fluorescence quenching (CIFQ) of Dbo by FTrp. Unexpectedly, the very high increase of charge density at elevated pH had no dynamical or conformational consequence in the anionic chains, neither in the absence nor in the presence of salt, in conflict with the common view and in partial conflict with accompanying molecular dynamics simulations. In contrast, the cationic peptides responded to ionization but with surprising patterns that mirrored the rich individual characteristics of each side chain type. The contrasting results had to be interpreted, by considering salt screening experiments, N-terminal acetylation, and simulations, in terms of an interplay of local dielectric constant and peptide-length dependent side chain charge–charge repulsion, side chain functional group solvation, N-terminal and side chain charge–charge repulsion, and side chain–side chain as well as side chain–backbone interactions. The common picture that emerged is that Coulomb repulsion between water-solvated side chains is efficiently quenched in short peptides as long as side chains are not in direct contact with each other or the main chain.



## INTRODUCTION

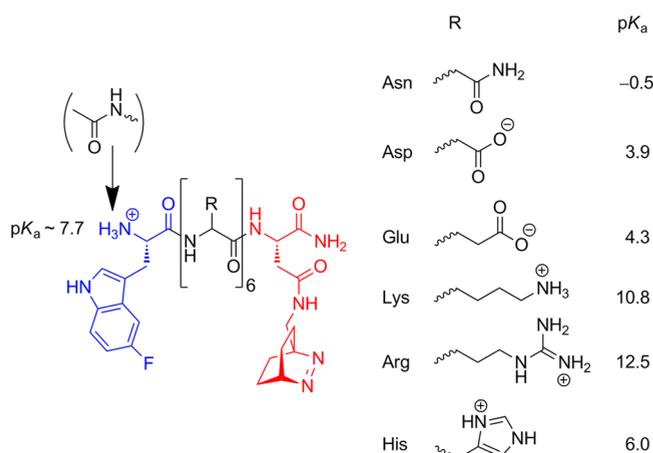
Electrostatic repulsion between like-charged side chains is considered to determine the dimensions of natively unfolded or intrinsically disordered peptides and proteins (IDPs), and to influence their biological function.<sup>1–7</sup> For instance, the radius of gyration of protamine peptides, cationic IDPs rich in arginine, correlates with the positive net charge per residue.<sup>1</sup> The radius of gyration of human prothymosin  $\alpha$ , an acidic IDP rich in aspartate and glutamate and with a net charge per residue of minus 0.4, is effectively reduced from  $\sim 40$  to  $\sim 30$  Å, when the charges are screened in the presence of 1 M KCl.<sup>2</sup> Repulsive Coulomb forces can also control chain dimensions during the refolding of regular proteins.<sup>8,9</sup> These new prospective insights encourage systematic studies on how Coulomb repulsion affects the unstructured polypeptide chain. Particularly, the short chain should be understood before any simulation-based prediction on the behavior of the long chain can be trusted. However, as the reservoir of biophysical methods to study the structure and dynamics of short polypeptides is limited, investigations on short chains have remained scarce.

We have previously developed two time-resolved fluorescence methods tailored to investigate polypeptide chains between 2 and 22 residues.<sup>10–18</sup> Here, we demonstrate that, when used in tandem, the methods provide synergistic information on chain structure and dynamics. The methods

are based on the two unique photophysical properties of 2,3-diazabicyclo[2.2.2]oct-2-ene (DBO) incorporated into peptides as asparagine derivative Dbo (Figure 1).<sup>16</sup> DBO has the longest fluorescence lifetime among organic fluorophores, 325 ns in aerated water.<sup>18</sup> Yet, its fluorescence is effectively quenched upon collision with tryptophan (Trp) or with the equally effective Trp derivative 5-fluoro-L-tryptophan (FTrp, Figure 1), which we have chosen for this work, because of its advantageous photophysical characteristics.<sup>19,20</sup> When FTrp and Dbo are appended to the termini of a polypeptide, they can rapidly diffuse and collide during the long fluorescence lifetime of Dbo. The frequency of end-to-end collision, the collision rate, is then easily obtained by comparing the Dbo lifetimes in the presence and absence of FTrp (Figure 2).<sup>17</sup> This rate provides a direct measure of chain end-to-end flexibility, and depends on the equilibrium distance distribution and intrachain diffusion but also, as emphasized in this work, on local interaction forces when probe-labeled regions come close.<sup>17,18</sup> The second unique property of DBO is its small oscillator strength with an extinction coefficient of only  $50 \text{ M}^{-1} \text{ cm}^{-1}$  at 350 nm in water. The DBO absorption spectrum and the FTrp fluorescence spectrum extend over the same range of

Received: August 15, 2014

Revised: November 6, 2014



**Figure 1.** Chemical structures of the N- and C-terminally conjugated fluorophores, 5-fluoro-L-tryptophan (FTrp, blue), 2,3-diazabicyclo[2.2.2]oct-2-ene (Dbo), and DBO-labeled amidated asparagine (Dbo, red) in acidic (Glu, Asp) and basic (Lys, Arg, His) peptides.

wavelengths such that optically excited FTrp can be deactivated by Förster resonance energy transfer (FRET) to DBO that now acts as energy acceptor (Figure 2). The small DBO extinction coefficient is responsible for the short critical radius ( $R_0$  ca. 10 Å), which ensures that energy transfer becomes effective only at distances below 20 Å and that an effective distance can be diagnosed with an unprecedented sub-Ångstrom resolution. As we previously discussed in detail,<sup>10–14</sup> the effective distance depends not only on the equilibrium distance distribution but also on intrachain diffusion. In contrast to the collision rate, however, it should show little sensitivity to local repulsive interactions between probe-labeled chain regions.

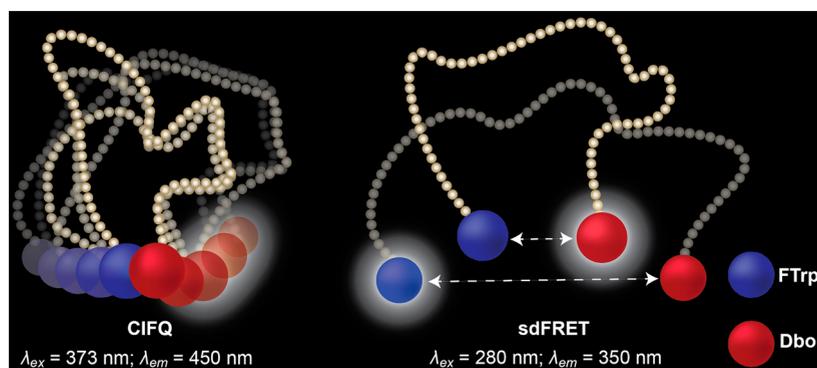
We will refer to the first method as collision-induced fluorescence quenching (CIFQ) and to the second as short-distance FRET (sdFRET). To switch from one method to the other merely requires adjustment of the excitation wavelength (Figure 2). We have already used CIFQ to establish a flexibility scale for the side chains of natural amino acids in peptides.<sup>17</sup> This scale depended strongly on side chain type and bulk and apparently less on side chain charge status, but already these preliminary experiments pointed to the question of Coulomb-force influence. In a follow-up study, we analyzed the  $\beta$ -sheet and found its  $\beta$ -turn segments to be considerably more flexible than its  $\beta$ -strand segments, while substitution of charged side

chains did not measurably alter segment flexibility.<sup>21</sup> Therefore, we ask here for the impact of repulsion in sequences with six ionizable residues in a row: Such homorepeat peptides have important biological functions<sup>22</sup> and warrant a maximal variation of charge status when the solution pH is varied from 1 to 12. In the acidic hexapeptides, Asp<sub>6</sub> and Glu<sub>6</sub>, electrostatic repulsion is “switched on” upon deprotonation and was expected to stretch the chains. In the basic chains, His<sub>6</sub>, Arg<sub>6</sub>, and Lys<sub>6</sub>, Coulomb repulsion is “switched off” upon deprotonation and was expected to be accompanied by profound contraction. MD simulations are already trusted to give insight into the behavior of long chains, of the refolding protein,<sup>23</sup> and proved also here to be invaluable in the interpretation of the experiments.

## RESULTS

**Design of Peptide Sequences.** A peptide chain composed of nonionizable residues should display no pH-induced dynamical or structural response. Coulomb effects could vary only when the charge status of the free N-terminal amino group and C-terminal carboxylate group is altered. We therefore used, with one exception, peptides with an amidated, neutral C-terminus. We further included chains with an acetylated, neutral N-terminus in our study to pinpoint potential charge interactions inflicted by the N-terminus. The different types of chains are surveyed in Table 1, where we also introduce a convenient short notation and list the single-labeled reference peptides that the CIFQ or sdFRET analyses require when utmost precision is the goal.

**Absence of pH Effects in Peptides with Nonionizable Side Chains (Asparagine).** As an initial negative control experiment, and as a general test of whether the CIFQ/sdFRET method tandem is suited for the investigation of pH-induced Coulomb effects, we performed the required sets of experiments (double-labeled peptide and the two single-labeled reference peptides) on Asn<sub>6</sub>. Because its side chains were neutral, and because the C-terminus was amidated, no change of attractive or repulsive Coulomb effects was expected over the pH range from 1 to 12. Indeed, in CIFQ measurements, the fluorescence kinetic traces of the double-labeled peptide measured between pH 1 and 12 coincided, as did the kinetic traces of the reference peptide (Figure 3a). The collision rate constant,  $k_q$ , was determined from the amplitude-weighted fluorescence lifetime of the double-labeled peptide ( $\tau_D$ ) and of the single-labeled reference peptide ( $\tau_S$ ) by using the equation

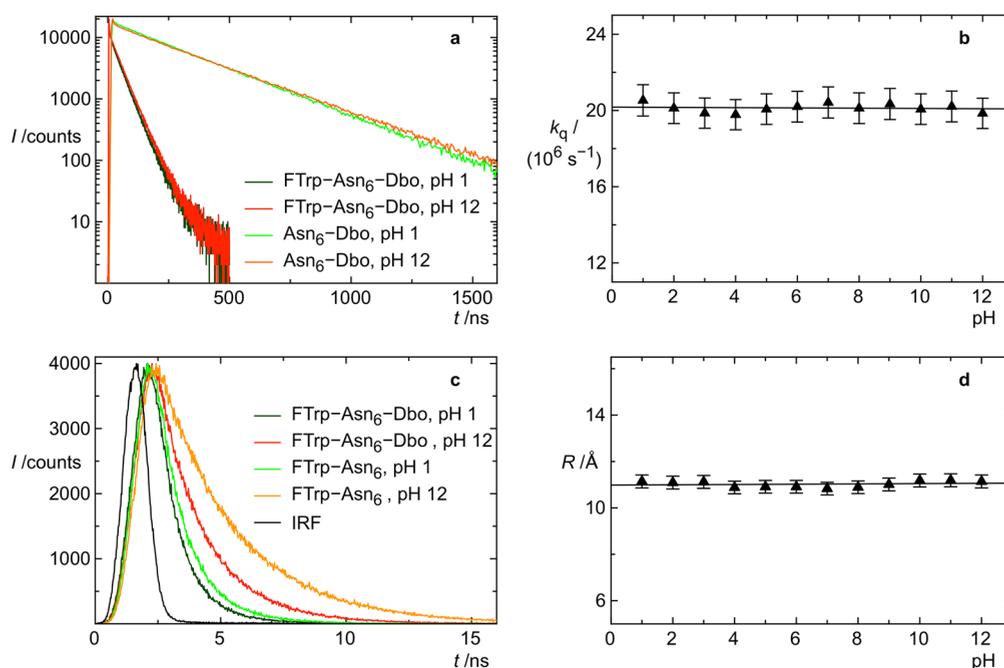


**Figure 2.** CIFQ: Following excitation of Dbo, quenching by FTrp occurs upon collision. sdFRET: Following excitation of FTrp, energy transfer to Dbo occurs at distances between 5 and 20 Å.

**Table 1. Experimentally Investigated Peptide Sequences and Terminal Protection Groups**

peptide [short notation] <sup>a</sup>	reference for CIFQ <sup>b</sup> [short notation]	reference for sdFRET <sup>c</sup> [short notation]
Peptides with Amidated C-Terminus (Only N-Terminal Charges Possible)		
H <sub>2</sub> N-FTTrp-Xaa <sub>6</sub> -Dbo-CONH <sub>2</sub> [Xaa <sub>6</sub> ]	H <sub>2</sub> N-Xaa <sub>6</sub> -Dbo-CONH <sub>2</sub> [Xaa <sub>6</sub> -Dbo]	H <sub>2</sub> N-FTTrp-Xaa <sub>6</sub> -CONH <sub>2</sub> [FTTrp-Xaa <sub>6</sub> ]
Peptides with Amidated C-Terminus and Acetylated N-Terminus (No Terminal Charges Possible)		
AcNH-FTTrp-Xaa <sub>6</sub> -Dbo-CONH <sub>2</sub> [Ac-Xaa <sub>6</sub> ]	AcNH-Xaa <sub>6</sub> -Dbo-CONH <sub>2</sub> [AcNH-Xaa <sub>6</sub> -Dbo]	AcNH-FTTrp-Xaa <sub>6</sub> -CONH <sub>2</sub> [AcNH-FTTrp-Xaa <sub>6</sub> ]
Peptides (Only Asp <sub>6</sub> ) with Acetylated N-Terminus (Only C-Terminal Charges Possible)		
AcNH-FTTrp-Asp <sub>6</sub> -Dbo-COOH [Ac-Asp <sub>6</sub> -COOH]	AcNH-Asp <sub>6</sub> -Dbo-COOH [- - -]	AcNH-FTTrp-Asp <sub>6</sub> -COOH [- - -]

<sup>a</sup>Short notation used in the main text and in Figure 4. <sup>b</sup>Dbo-only reference peptide required for CIFQ. <sup>c</sup>FTTrp-only reference peptide required for sdFRET.



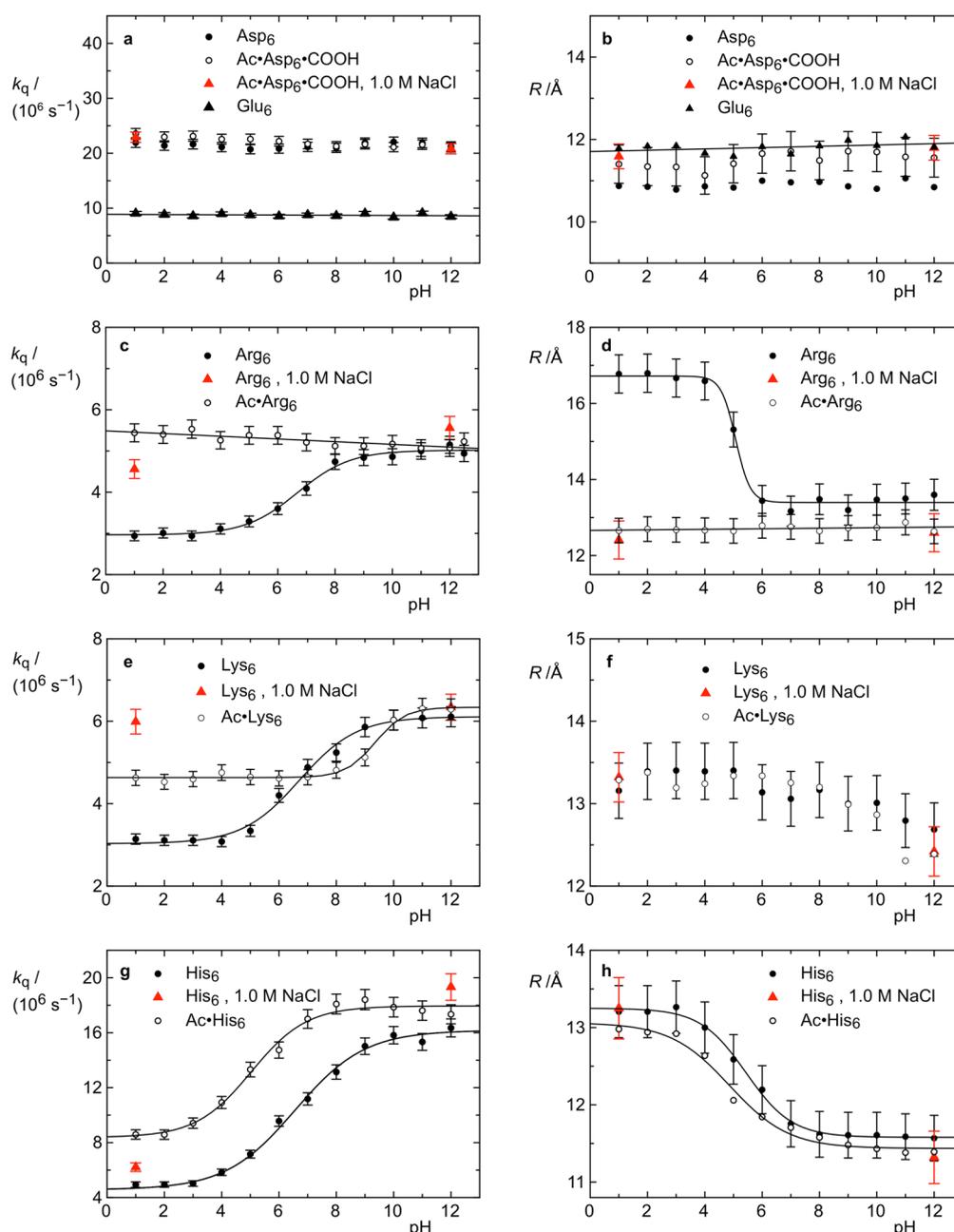
**Figure 3.** (a) CIFQ: Fluorescence decay profiles of Dbo in FTTrp-Asn<sub>6</sub>-Dbo at pH 1 (dark green) and pH 12 (red) and in the reference peptide Asn<sub>6</sub>-Dbo at pH 1 (green) and pH 12 (orange). (b) The collision rate of FTTrp-Asn<sub>6</sub>-Dbo plotted against the solution pH. (c) Fluorescence decay profiles of FTTrp in FTTrp-Asn<sub>6</sub>-Dbo at pH 1 (dark green) and pH 12 (red) and in the reference peptide FTTrp-Asn<sub>6</sub> at pH 1 (green) and pH 12 (yellow). The instrument response function (IRF, black trace) was routinely recorded for data deconvolution. (d) The resulting effective donor-acceptor distance in FTTrp-Asn<sub>6</sub>-Dbo plotted against the solution pH; see text for analysis.

$k_q = 1/\tau_D - 1/\tau_S$ .<sup>18</sup> The collision rate remained constant at ca.  $20 \times 10^6 \text{ s}^{-1}$ , the time constant ( $1/k_q$ ) at ca. 50 ns (Figure 3b).

In sdFRET measurements, the kinetic traces at pH 1 and 12 could not coincide (Figure 3c) because the quantum yield and fluorescence lifetime of the donor, FTTrp, depend on pH (Supporting Information Figure S2). We obtained the rate constant of FRET,  $k_{\text{FRET}} = 1/\tau_D - 1/\tau_S$ , by using now the FTTrp-only labeled peptide as reference. On the basis of Förster's theory, this rate can be directly translated into an effective donor-acceptor distance,  $R$ , by using his central equation,  $k_{\text{FRET}} = (1/\tau_S) \times (R_0/R)^6$ . Here, both the natural donor lifetime  $\tau_S$ , in the absence of acceptor, and the Förster radius,  $R_0$ , depend on the donor quantum yield, while the FRET rate constant does not as the combined dependencies cancel out.<sup>10</sup> Thus, the factor  $R_0^6/\tau_S$  does not change with the quantum yield, and the value of  $R_0$  can be obtained at any pH from the reported values at pH 7.0 ( $R_{0,\text{ref}} = 9.6 \text{ \AA}$ ,  $\tau_{S,\text{ref}} = 2.0 \text{ ns}$ ) and from the measured lifetime,  $\tau_S$ , via  $R_0 = R_{0,\text{ref}} \times (\tau_S/\tau_{S,\text{ref}})^{1/6}$ .<sup>10</sup>

Despite the pH-dependency of the fluorescence kinetics, the extracted effective distances between FTTrp and Dbo in Asn<sub>6</sub> displayed a remarkable constancy with a value of ca. 11 Å over the entire range of pH values (Figure 3d), consistent with the pH-independent collision dynamics obtained by CIFQ (Figure 3c). This observed absence of pH effects for nonionizable side chains demonstrates that any pH-induced response of ionizable chains has to be traced back to an altered side chain charge status.

**Disparity in the Response to Ionization of Acidic and Basic Chains.** The pH range from 1 to 12 ensures the complete transition from uncharged to charged side chains in all peptides with the exception of Arg<sub>6</sub> and Lys<sub>6</sub> (pK<sub>a</sub>: Asp 3.9, Glu 4.3, Arg 12.5, Lys 10.8, His 6.0).<sup>24,25</sup> The results from all chains are shown in Figure 4 with CIFQ measurements on the left and sdFRET measurements on the right, going from the acidic peptides, Asp<sub>6</sub> and Glu<sub>6</sub>, to the basic chains, Arg<sub>6</sub>, Lys<sub>6</sub>, and His<sub>6</sub>. The acidic and basic chains displayed profoundly different patterns. The collision rate and effective distance of



**Figure 4.** Collision rate,  $k_q$ , obtained from CIFQ measurements (left) and the effective distance,  $R$ , obtained from sdFRET measurements (right), plotted against the solution pH (a–h) for Xaa<sub>6</sub> chains, for Ac•Xaa<sub>6</sub> chains, for the Ac•Asp<sub>6</sub>•COOH peptide (a,b), and in the presence of salt (1.0 M NaCl, red). See also Supporting Information Figures S4–6.

the acidic peptides remained constant over the entire pH range, while the basic chains showed a rich variety of CIFQ and sdFRET profile combinations.

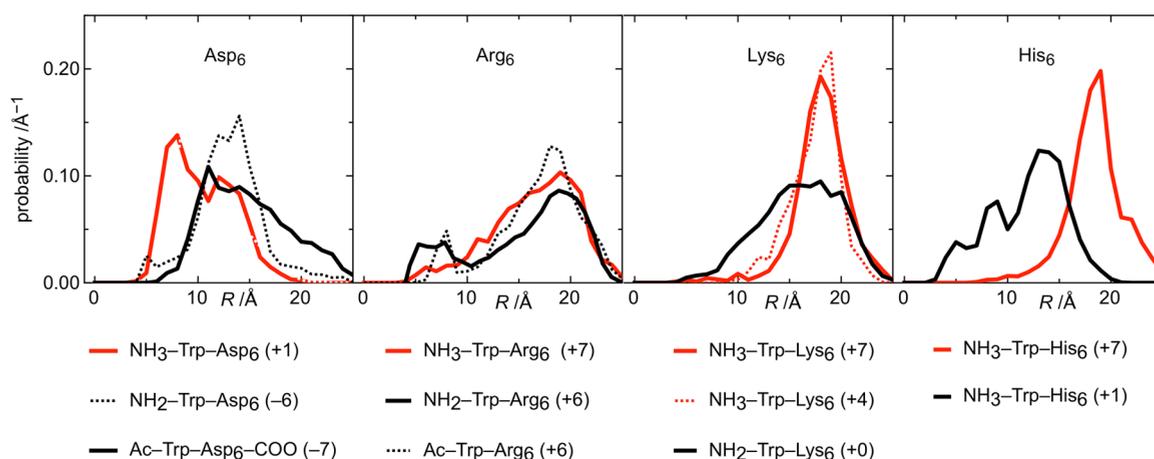
The absence of any charge effect in the acidic chains came as another surprise, and we had to assume an involvement of the N-terminus. Coulomb repulsion between negatively charged side chains could have been offset by favorable interactions between the side chains and a positively charged N-terminus. We disproved this possibility by using a peptide with blocked N- and unblocked C-terminus, AcNH-FTrp-Asp<sub>6</sub>-Dbo-COOH (Ac•Asp<sub>6</sub>•COOH, Table 1), again expecting that it could make a difference when seven negative charges appear on seven adjacent residues. Yet, again, the pH profiles remained horizontal (Figure 4a,b; Table 2).

**Salt Effects.** A common test for Coulomb repulsion between like-charged groups is whether the putative effects disappear in the presence of salt. A clear indication that repulsion controls chain dimensions is chain compaction in salt presence.<sup>2,8</sup> Like-charged side chains can be nearest neighbors along the chain but can still come closer when any strain caused by repulsion is neutralized by salt. Following Debye–Hückel theory, salt addition can shorten the Debye length, that is, the distance above which repulsion becomes ineffective, to less than  $2 \text{ \AA}$ ,<sup>26</sup> a distance that is short even for neighboring charged residues. Interestingly, salt effects (with 1.0 M NaCl, at low pH for the basic chains and high pH for the acidic chains) afforded quite different responses; that is, the collision rate showed a factor of 2 increase for Lys<sub>6</sub>, 50% increase for Arg<sub>6</sub>, 25%

Table 2. Collision Rate Constants and Effective Donor–Acceptor Distances at Low and High pH

peptide chain	$k_q/(10^6 \text{ s}^{-1})^{a,b}$		$R/\text{\AA}^{a,c}$		$R_{\text{MD}}/\text{\AA}$ (charge status) <sup>d</sup>	
	pH < 2	pH > 11	pH < 2	pH > 11	protonated	unprotonated
Asn <sub>6</sub>	20	20	10.9	11.0	13.0 (+1)	12.4 (+0)
Asp <sub>6</sub>	22	22	10.9	11.0	10.0 (+1)	12.5 (−6)
Ac-Asp <sub>6</sub> -COOH <sup>e</sup>	23	21	11.5	11.6	13.4 (0)	14.4 (−7)
Glu <sub>6</sub>	8.9	8.8	11.8	12.0	9.3 (+1)	12.1 (−6)
Arg <sub>6</sub>	3.0	5.1	16.8 <sup>f</sup>	13.5	16.0 (+7)	15.6 (+6)
Arg <sub>6</sub> /1.0 M NaCl	4.6	5.6	12.4	12.6		
Ac-Arg <sub>6</sub>	5.4	5.1	12.7	12.8	16.3 (+6)	
Lys <sub>6</sub>	3.1	6.1	13.2	12.7	17.5 (+7)	15.5 (+0)
Lys <sub>6</sub> /1.0 M NaCl	6.0	6.3	13.3	12.4		
Ac-Lys <sub>6</sub>	4.6	6.3	13.3	12.3		
His <sub>6</sub>	4.9	16	13.2	11.6	17.9 (+7)	12.4 (+1)
His <sub>6</sub> /1.0 M NaCl	6.2	19	13.3	11.3		
Ac-His <sub>6</sub>	8.6	19	12.9	11.4		

<sup>a</sup>Average of values measured below pH 2 and above pH 11, respectively. <sup>b</sup>±5% error. <sup>c</sup>The error increases exponentially from ±0.15 Å at  $R = 9.6$  Å ( $R_0$ ) to ±0.8 Å at  $R = 12.0$  Å and ±2.5 Å at  $R = 17.0$  Å. <sup>d</sup>Average end-to-end distance in simulated peptides; charge status given in brackets. <sup>e</sup>The experimental values in the absence and presence of salt (1.0 M NaCl) are virtually identical. <sup>f</sup>Distances outside the range between  $0.5R_0$  and  $1.5R_0$ , where  $R_0$  is 9.6 Å, cannot be determined with high accuracy; the given value constitutes a lower limit.



**Figure 5.** Distance distributions of the acidic and basic peptides with varying charge status. The distributions were obtained from end-to-end distance trajectories of at least 200 ns duration. The end-to-end distance was measured from the N-terminus (center of the central C–C bond of the indol ring) to the carbon atom of the C-terminal carboxyl group.

increase for His<sub>6</sub>, but no effect for Asp<sub>6</sub> and Glu<sub>6</sub> (Table 2, Figure 4, Supporting Information Figures S3–5). The effective distance, in contrast, was affected only for Arg<sub>6</sub> (Figure 4d).

**MD Simulations.** Confronted with the apparent absence of any charge repulsion between negatively charged side chains in homorepeat peptides in water, we questioned common chemical intuition and tested whether the physical models used in MD simulations (MD) could effortlessly predict the absence of any Coulomb effects in acidic chains. Thus, we simulated representative peptide sequences that we studied experimentally by using Amber 11 and the Amber ff99SB force field,<sup>27,28</sup> routine but state-of-the-art MD tools that were, for instance, successfully employed in protein-folding thermodynamic predictions.<sup>29</sup>

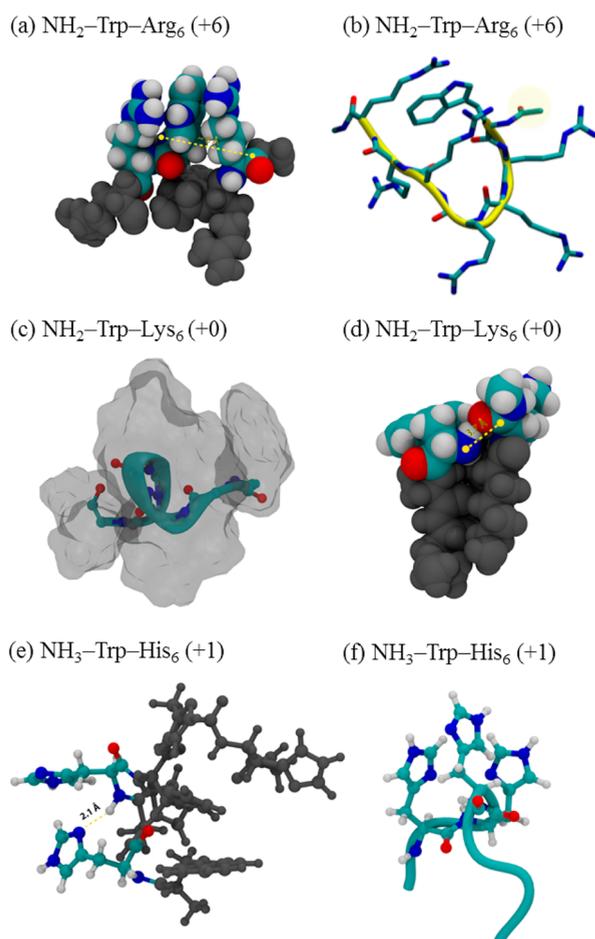
Because absolute quantitative agreement could not be expected (see below), but at best an agreement between experimental and theoretical trends with pH and side chain charge status, we accepted two simplifications to bypass parametrization inconsistencies; the simulated chains contained Trp instead of FTrp, and they contained no C-terminal Dbo. Akin to our previous studies,<sup>12,15</sup> water was modeled explicitly,

and the simulations were run for at least 200 ns to ensure convergence. The average end-to-end distances,  $R_{\text{MD}}$ , obtained from the distance distributions (Figure 5), and in particular their pH dependence, can be compared to the effective distances extracted from the sDFRET experiments (Table 2). Note that these experimental distances are systematically slightly smaller than the real average distance because of the higher contribution of short distances to FRET.<sup>10,11</sup> Thus, also the MD-calculated average distances tend to be larger than the experimental values.

The MD simulations predicted an impact of side chain charges for all peptides (Figure 5). When we performed a comprehensive cluster analysis of all systems based on ref 30 (Supporting Information Appendix) and also analyzed the probabilities of transient intrachain hydrogen bonds (Supporting Information Table S1), we were led to critical short-distance conformations (Figure 6a–f).

## DISCUSSION

While it is broadly accepted that electrostatic repulsion can impact the thermodynamic stability of globular proteins,<sup>31,32</sup> its



**Figure 6.** Short-distance structures of basic peptides. (a,b) The  $\pi$ -cation interactions in Arg<sub>6</sub> induce short distances (9 Å) by stabilizing an arginine-tryptophan-arginine sandwich conformation. (c) Hydrophobic, neutral lysine side chains constitute a low-permittivity environment that stabilizes main chain hydrogen bonds in Lys<sub>6</sub> and (D) favors conformations of short end-to-end distance (6–10 Å). (e) Side chain–main chain and (f) side chain–side chain hydrogen bonds in His<sub>6</sub>.

dominant role in determining the spatial dimensions of flexible IDPs including arginine-rich antimicrobial peptides, protamine peptides, and human prothymosin  $\alpha$  has been investigated only recently.<sup>1,2</sup> We have now tackled the fundamental understanding of charge effects in proteins in a bottom-up-approach, by investigating short polypeptides exclusively composed of the same ionizable side chains, Asp, Glu, Lys, Arg, Lys, and His. The experimental parameters that we recorded were the effective end-to-end distance and collision rate, as a measure of chain structure and flexibility. Foregoing to any interpretation of pH-caused charge effects, it is important to discuss how distances and collision rates vary between the different types of amino acid residues.

**Dependence of End-to-End Distances and Collision Rates on the Amino Acid Type.** Obvious factors that dominate differences in end-to-end distances of homorepeat peptides are side chain–side chain excluded-volume effects and the maximization of side chain entropy that is achieved when the main chain expands.<sup>17,33–36</sup> Accordingly, the average length of a peptide increases when the steric bulk and the length of a side chain increase. The same reasons dictate the end-to-end collision rates. The side chains of Arg, Lys, and His occupy a

larger volume than the acidic side chains, Asp and Glu, and those of Arg and Lys are also “longer” in the sense that they possess 3–4 instead of 1–2 flexible methylene groups. Indeed, the effective end-to-end distances extracted from the experiments (Table 2, Figure 4) for the “bigger” side chains, Arg (ca. 17 Å at pH 2) and His and Lys (ca. 13 Å at pH 2), are significantly longer than those for the “smaller” ones, Asn and Asp (ca. 11 Å) and Glu (ca. 12 Å). Note that even the small difference between Glu and Asp/Asn (one additional methylene group for Glu) is nicely reflected in a ca. 1 Å larger distance. The MD calculations afforded consistent results, that is, shorter average distances for the acidic amino acids (Table 2). The trends of the end-to-end collision rates (Table 2, Figure 4) can also be roughly rationalized in part through steric effects;<sup>17</sup> for example, basic chains display smaller rates than acidic chains, and the collision rate for Glu is a factor of 2 less than that for Asp/Asn.

While the consistency of the trends of the end-to-end distances and collision rates with the type of amino acid is reassuring for the employed experimental and theoretical approaches, they have also important implications for the charge effects that could be expected. If side chain bulk and entropy cause larger end-to-end distances and slower end-to-end collision rates of the chain, electrostatic repulsion between the side chains should also become less effective. Therefore, the most prominent effects of charge repulsion were expected for the shortest peptides, those composed of the acidic Asp<sub>6</sub> and Glu<sub>6</sub> backbone, but, strikingly, exactly these side chains turned out to be immune, as far as their dynamics and end-to-end distances were concerned, toward charge and pH effects (Table 2, Figure 4).

**pH Independence for Acidic Hexapeptides.** The Asp and Glu hexamers displayed almost constant end-to-end collision rates (within  $\pm 5\%$ ) and effective end-to-end distances (within  $\pm 0.5$  Å) at every pH between 1 and 12 (Figure 4a,b). This result, obtained by two photophysically independent kinetic methods, CIFQ and sFRET, and pointing to the absence of charge repulsion in highly negatively charged short peptide chains, could not have been anticipated: In MD simulations, deprotonation stretched Asp<sub>6</sub>/Glu<sub>6</sub> from 10.0 Å/9.3 Å to 12.5 Å/12.1 Å by 2.5 Å/2.8 Å (Table 2), which is a significant fraction of the maximal 10-Å stretching that we obtained, when we placed artificially high charges on the Asp<sub>6</sub> carboxylate groups, to obtain a reference value. On the other hand, when the N-terminus was neutralized, acetylated, as in Ac-Asp<sub>6</sub>-COOH, the pH-induced deprotonation led to a stretching of merely 1.0 Å, from 13.4 to 14.4 Å. This larger value of 13.4 Å in contrast to the value of 10.0 Å in protonated Asp<sub>6</sub> suggests that it is not so much Coulomb repulsion that was poorly represented in the MD simulations but, instead, the strength of interactions between the charged N-terminal ammonium group of Asp<sub>6</sub> and its chain. Indeed, it is exactly this idea, which is reinforced, when we later turn to the basic chains.

In pertinent studies, Buscaglia and co-workers replaced three glutamine residues of a 14-mer peptide for glutamate, Cys-(Ala-Gly-Gln)<sub>4</sub>-Trp  $\rightarrow$  Cys-(Ala-Gly-Glu)<sub>4</sub>-Trp, and measured the Förster energy transfer between Trp and Dansyl-labeled Cys.<sup>37</sup> Also here, the additional charges, but on residues well separated along the chain, did not affect the end-to-end intrachain distance. Schweitzer-Stenner and co-workers studied the extremely short trimer Asp<sub>3</sub> at varying pH. The required methods, vibrational CD, UV-CD, and <sup>1</sup>H NMR spectroscopy,

allowed a multitude of conclusions, among them a conformational response of the central aspartate side chain, but no simple extrapolation to intrachain collision rates and distances.<sup>38</sup>

We also considered whether the horizontal pH profiles of the acidic chains could be caused by different influences that exactly compensate each other. When the chain becomes charged, it might expand, but intrachain diffusion might increase. While chain expansion would lead to a reduced collision frequency and increased effective distance, enhanced diffusion would effect the reverse. It is admittedly improbable that the net result would be zero over the entire pH range, but, to address this remote possibility, we included a study of salt effects (1.0 M NaCl, Table 2, Supporting Information Figures S4–6), because the end-to-end collision rates as well as effective distances should be systematically affected if Coulomb repulsion was dominant. Salt addition raises the solution viscosity by merely 2% (1.0 M NaCl) such that viscosity effects on intrachain diffusion, that is, on the collision rate and the effective distance, can be ignored. It is all the more intriguing that the acidic peptides displayed exactly the same rates and distances in the absence and presence of salt. With the combined results of pH and salt independence, we could now conclude with confidence that Coulomb repulsion in short acidic chains affects neither their end-to-end dimensions nor their end-to-end dynamics.

The observed insensitivity toward pH is consistent with the classical concept suggested by Flory, that long linear polymers can be viewed as a sequence of short segments, as “blobs”, whose conformational equilibria do not respond to exterior conditions, neither to the chemical nature of the solvent nor to changes of the solution pH.<sup>39,40</sup> While this approximation, the inertia of a blob, is valuable in the treatment of long chains,<sup>41,42</sup> it is insufficient when we try to understand the many facets of peptide chain segments, as our work demonstrates. It would be equally justified to adopt the opposite point of view and argue with Manning’s theory of counterion condensation.<sup>43–47</sup> Here, a polyelectrolyte, for instance, the DNA duplex, is modeled as a rigid rod. If the Manning parameter,  $q_0 = l_B/l$ , that is, the ratio of the Bjerrum length ( $l_B$ ,  $\sim 7.0$  Å at room temperature) and the distance between adjacent charges,  $l$ , is greater than unity ( $q_0 > 1$ ), the system is only predicted to be stable in the presence of counterions that, at low salt concentration, would condense onto the chain.<sup>46</sup> The condensation of bivalent counterions is predicted already for  $q_0 > 0.5$ . However, our experimental procedures warranted a total concentration of mono- and bivalent counterions, even at extreme pH (pH 1.0), of maximally  $0.6 \mu\text{M}$  (cationic) and  $6.5 \mu\text{M}$  (anionic) as compared to  $50 \mu\text{M}$  of peptide. In our simulations, the average distance between adjacent aspartate carboxylates in neutral and charged chains was 7 and 8 Å, respectively, such that condensation of monovalent ions is, according to the theory, not enforced ( $q_0 \approx 1$ ). Measurements in the presence of 1.0 mM EDTA at low and high pH afforded the same results as in its absence (results not shown). When we intentionally added 100 mM calcium chloride to an Asp<sub>6</sub> solution at pH 12.0, we did not observe a chain collapse. Instead, we observed an increased effective distance, possibly as the consequence of repulsion between carboxylate-bound calcium ions (data not shown). This is in line with the investigations of Haber-Pohlmeier et al. on the Ca<sup>2+</sup> affinity of glutamate-rich sequences, including polyglutamic acid. These authors also detected no counterion condensation at very low Ca<sup>2+</sup> concentrations and only low-affinity binding at higher ones.<sup>48</sup>

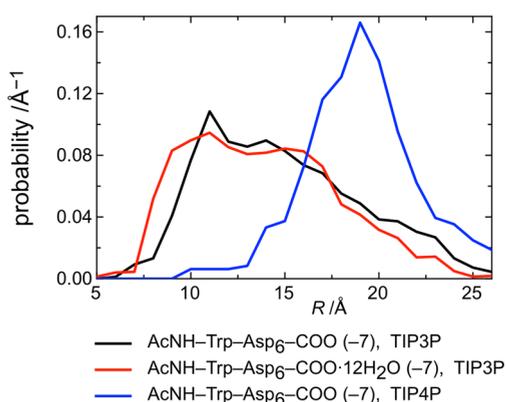
**Electrostatic Effects Increase with Increasing Chain Length.** How can Coulomb repulsion determine the spatial dimension of an IDP of 20–100 residues, for instance, of the acidic chain of human prothymosin  $\alpha$ ,<sup>2</sup> when it has no effect on short acidic peptides? Prothymosin  $\alpha$  (110 aa) contains not only Asp and Glu but also polar and hydrophobic residues ( $\sim 50\%$ ). The result is a reduced dielectric permittivity of the charge-surrounding space, which could be a precondition for sufficiently strong and chain-stretching Coulomb forces. However, Duhamel and co-workers studied electrostatic repulsion versus hydrophobicity, varied the fractions of Asp and Phe in long Asp/Phe polymers, and could clearly identify effective Coulomb repulsion in the pure Asp polymer chain.<sup>49</sup>

With increasing length, the chain itself begins, by mere stochastics, to constitute a microenvironment of its own charged side chains. The number of charge-surrounding and solvating water molecules is reduced and, by that, the dielectric permittivity. The repulsive Coulomb energy between the charges increases, as does the desolvation “penalty”, to accommodate a charge at low permittivity.<sup>50–52</sup> As a consequence, the chain is forced to expand. To strengthen this hypothesis, we evaluated equilibrium structures obtained from simulations of neutral chains containing 6 or 30 protonated aspartate residues. An approximate model<sup>42</sup> afforded average apparent dielectric constants of 33 for Asp<sub>6</sub> and 10 for Asp<sub>30</sub>, that is, a 3 times higher Coulomb energy and desolvation penalty for Asp<sub>30</sub> (see the Supporting Information for details). This large difference guarantees that, after deprotonation and complete side chain ionization, Asp<sub>30</sub> expands dramatically, in contrast to Asp<sub>6</sub>.

In recent simulations, Ramachandran et al. compared the neutral and charged states of the Asp<sub>10</sub> decamer and Asp<sub>30</sub> tricosamer.<sup>53</sup> The end-to-end distance in Asp<sub>10</sub> almost doubled upon deprotonation, while it increased about 5-fold in Asp<sub>30</sub> (see p 13909 in Table 1 of ref 53). These results underline our point, although it is slightly surprising that the decamer expands significantly while the hexamer that we studied shows no response in the experiments and only a modest one in the simulations. Our production runs extended over 200 ns, while the runs of Ramachandran et al. extended over only 50 ns, which might not be sufficient, particularly, when we consider that the experimental time constant of end-to-end collision ( $1/k_q$ ) is already 45 ns (Table 2) in the shorter Asp<sub>6</sub> chain and that it will be even larger in the Asp<sub>10</sub> chain.<sup>18</sup>

In conclusion, dielectric-permittivity arguments are likely a dominant cause of the strong response to charge repulsive effects in longer acidic peptides or IDPs, but they cannot explain why the MD simulations predict modest but sizable differences for short chains, while experimentally no effect is found. We therefore evaluated the robustness of the MD simulations and varied the explicit water solvent model, going from the TIP3P<sup>54</sup> to the succeeding TIP4P model.<sup>55</sup> However, the chain expanded even more (from 14.4 to 18.5 Å, Figure 7). Although this trend can be rationalized by the discrepant dielectric constants of water predicted by the models, 82 for TIP3P and 50 for TIP4P,<sup>56</sup> it impressively illustrates that systematic shortcomings of MD methods can still not be ruled out.

Searching for factors that are potentially critical but ignored in MD simulations, we focused on specific-ion effects. It has, for instance, been established that the carboxylate group binds water molecules tightly, with an apparent hydration number of two.<sup>57–61</sup> When we suppressed the exchange of carboxylate-



**Figure 7.** MD-simulated end-to-end distance distributions of hexaaspartate in its deprotonated state ( $-7$ ) in the absence (black) and presence (red) of constraints: In the simulation of the peptide AcNH-Trp-Asp<sub>6</sub>-COO·12H<sub>2</sub>O ( $-7$ ), the exchange of side chain-bound water molecules with bulk water was suppressed. Also shown is the effect of the explicit water solvent models TIP3P (black, red) and TIP4P (blue) on the distance distributions.

bound water molecules with bulk water in the simulations (Supporting Information Figure S8), the distance distribution shifted toward shorter distances (Figure 7), and the resulting average distance of 13.4 Å did indeed coincide with that of the neutral Ac-Asp<sub>6</sub>-COOH chain. This distance is still significantly larger than that of the Asp<sub>6</sub> chain with neutral side chains but charged N-terminus (10.0 Å, Table 2). There are likely to be additional factors, which are slightly misrepresented in the force-field calculations. In the neutral Asp<sub>6</sub> chain, the strength of intrachain interactions between the ammonium group and the chain is probably overestimated, while, in the charged chain, the charge distribution within the carboxylate–water clusters is ignored, because the negative charge is a priori localized on the carboxylate oxygen atoms. As it stands, the proper treatment of carboxylate hydration is still under debate.<sup>62,63</sup>

**pH Dependence for Basic Hexapeptides.** In sharp contrast to the acidic hexapeptides, the basic ones (Arg, Lys, His) all displayed a significant pH dependence of the end-to-end collision rates and distances in the experiments as well as the MD simulations (Table 2, Figure 4c–h). In detail, their end-to-end collision rates in their low-pH, positively charged forms, were lower and their effective end-to-end distances were larger. While it is tempting to attribute this trend to Coulomb forces, its absence in the acidic peptides called for extensive control experiments.

The end-to-end flexibility of hexaarginine and -lysine almost doubled when the pH was raised (Figure 4c,e). The very similar transitions started at about pH 6 and had to be due to the deprotonation of the free N-terminal ammonium group ( $pK_a \approx 7.7$ ); the side chains titrate much later (Arg  $pK_a \approx 12.5$ , Lys  $pK_a \approx 10.5$ ).<sup>64</sup> Indeed, when we acetylated the N-terminus of Arg<sub>6</sub>, the collision rate remained constantly at high values over the entire pH range. When we acetylated the N-terminus of Lys<sub>6</sub>, a low-amplitude transition at higher pH was still detectable and could only be caused by deprotonation of lysine side chains. Moreover, adding salt increased the collision rate of free-N-terminus Arg<sub>6</sub> and Lys<sub>6</sub> as effectively as raising the pH and had no effect at high pH, where N-terminal neutrality is guaranteed (pH > 9.0, Figure 4c–e; Supporting Information Figures S5, 6).

As has been mentioned and is known from other contact-quenching methods,<sup>37</sup> the end-to-end flexibility of chains is

highly sensitive to attractive/repulsive interaction forces that act on probes in proximity. Hence, there is no surprise in that a charged N-terminal FTrp can energetically favor or disfavor chain structures with contacting ends. Such looped conformations are necessarily more compact than relaxed and better solvated conformations and display an accordingly smaller dielectric permittivity. The increased effectiveness of Coulomb repulsion between the N-terminal ammonium group and the arginine and lysine side chains or, simply, the penalty to be paid for a charged group to stay in a low-permittivity environment could explain the low end-to-end flexibility observed at low pH.

Förster energy transfer, on the other hand, takes place at all distances; even when the probabilities of distances, at which probes are almost in direct contact, vary strongly, if they remain small in comparison to the probabilities of nonstrained conformations, they can hardly be detected by sFRET. Indeed, the FRET profile of Lys<sub>6</sub> (as well as N-acetylated Lys<sub>6</sub>) indicated only a marginal structural change, which became more visible at high pH values (Figure 4f), at which the lysine side chains are expected to become neutral. MD prediction almost coincided with these experimental observations. Upon going from the protonated to the neutral hexalysine chain, the distance distribution broadened due to an increase of short-distance probability (Figure 5), as was experimentally reflected by a collision-rate increase. The peak and average distances, however, changed only mildly (from 17.5 to 15.5 Å, Table 2), which was reasonably reflected by the small (0.7 Å) decrease in the sFRET measurements reporting on the entire distribution of distances. Side chain neutralization favored hydrophobic as well as intramain chain interactions due to the considerable hydrophobicity of the lysine butano spacers and the accompanying decrease of dielectric permittivity, at which transient helical segments and short-distance conformations became slightly more stable (Figure 6c,d).

In view of the foregoing discussion, the second surprise of our study was the large and sharp structural transition of the Arg<sub>6</sub> peptide (Figure 4d). Its chain displayed the largest effective end-to-end distance among the peptides (16.8 Å at low pH, Table 2), which strictly depended on the presence of a charged N-terminus. When the N-terminus was neutralized (by acetylation or at high pH) or when salt was added, the chain collapsed by about 3 Å (Figure 4d). Hence, again, the experiments yielded only little indication for side chain–side chain Coulomb repulsion (13.5 Å in the absence and 12.6 Å in the presence of 1.0 M NaCl, Table 2) but rather for a repulsion between the N-terminal ammonium group and the charged side chain units or, simply, for the penalty to be paid for a charged group located in a low-permittivity environment.

The MD simulations of Arg<sub>6</sub> with charged guanidinium units and neutral N-terminus were the only ones that did not converge within 500 ns. The distance trajectory pointed to a remarkably stable short-distance subpopulation (Supporting Information Figure S9). In a representative compact structure with an end-to-end distance of only 9 Å, two arginine residues were stacked to the two sides of the tryptophan indol ring (Figure 6c,d). Note that the indole–guanidinium interaction is the strongest among the familiar cation- $\pi$  interactions between basic and aromatic side chains.<sup>65</sup> The stability of the cation- $\pi$  stacked complexes decreases when the N-terminal FTrp becomes itself positively charged, for reasons of competitive binding of the terminal  $-\text{NH}_3^+$  group as well as charge repulsion of adjacent charges, resulting in a more expanded chain distribution, as observed experimentally and theoretically.

The Arg<sub>6</sub> experiments expose once again the N-terminus as a potentially important denominator of the structural and dynamic properties of peptides and of their biological activity; another prominent example is  $\alpha$ -synuclein, abundant in the human brain and involved in Parkinson's disease.<sup>66–68</sup>

**pH Dependence for Hexahistidine.** Among the basic residues, histidine is closest to glutamate and aspartate with side chains of six, five, and four heavy atoms. Yet, in contrast to the acidic peptides, the collision rate of His<sub>6</sub> increased and the effective distance decreased between pH 4 and 8, which could easily be ascribed to a reduced Coulomb repulsion and concomitant chain contraction. The corresponding MD simulations were qualitatively in agreement but exaggerated the compaction: The average distance decreased from 17.9 to 12.4 Å, by more than 5 Å (Table 2; a reasonable explanation is suggested in the Supporting Information, section 10).

While pH-dependent Coulomb forces seemed to dominate the CIFQ and sFRET pH profiles of His<sub>6</sub>, the observed effective distances remained unchanged when salt was added. These apparently contradictory observations can be reconciled: Instead of focusing on neutral and spatially separated side chains that, when they become charged, separate even more and stretch the chain, we can consider neutral side chains that interact and bind to each other and to the backbone and keep the chain in compact conformations. Such contacts disappear when the contact partners become charged by protonation, but in a salt-independent manner. Recent studies are in support of this concept and suggest that the imidazole rings of two neutral histidine side chains can form transient hydrogen bonds and also  $\pi$ – $\pi$  stabilized contact pairs with one ring stacked upon the other.<sup>69–72</sup> In addition, the MD simulations indicated significantly increased main chain–side chain interactions in neutral His<sub>6</sub> (Supporting Information Table S1); only the neutral imidazole ring can act as both hydrogen-bond donor and acceptor for the peptide backbone. In summary, the His<sub>6</sub> peptide experiences an apparent compaction upon deprotonation due to increased specific intrachain interactions, which accounts for the fact that the effective end-to-end distance is rather insensitive to the presence of salts, but nevertheless sensitive to pH.

## CONCLUSIONS

Coulomb repulsion plays no or a remarkably small role in highly charged but well solvated peptide chains such as hexaaspartate and -glutamate. It is, however, effective at contact distance when it disrupts side chain–side chain and side chain–main chain interactions in hexahistidine and when it disfavors the looped and compact conformations of hexaarginine and -lysine probed by CIFQ. The charged N-terminus is kept apart from the low-permittivity volume of the chain body. This result is counterintuitive for the human but less counterintuitive for the machine; MD simulations predict sizable but modest effects. We demonstrated the synergy and efficiency of the fluorescence methods sFRET and CIFQ that can test the structure and flexibility of the short peptide, but can potentially diagnose any developing short-distance relationship, for instance, during protein folding or enzymatic activity.

## MATERIALS AND METHODS

**Materials.** All peptides were synthesized in cooperation with Biosyntan (Berlin) and were obtained in >95% purity. We prepared the Fmoc derivative of Dbo and of 5-fluoro-L-

tryptophan as previously described.<sup>18</sup> In their standard form, the peptides had a free N-terminus (NH<sub>2</sub>–/NH<sub>3</sub><sup>+</sup>–) and amidated C-terminus (–CONH<sub>2</sub>); variations from this pattern are indicated in the main text. All chemicals were purchased from Sigma-Aldrich.

**Spectroscopy.** The fluorescence lifetimes of all peptides were measured by time-correlated single-photon counting (FLS920, Edinburgh Instruments Ltd.) at a fixed temperature of 25 °C (Julabo F25/HD thermostat) and peptide concentrations of ca. 50  $\mu$ M. The solution pH was adjusted by adding HCl or NaOH; the pH values remained as initially set, as was routinely confirmed by pH determination after measurement.

In sFRET measurements, the tryptophan residue was excited at 280 nm by using a pulsed LED (PicoQuant, PLS-280, fwhm ca. 300 ps), and tryptophan fluorescence was followed at 350 nm. The Förster radius of the FTrp/Dbo pair was determined as previously described.<sup>10</sup> The instrument response function was routinely recorded and used in deconvolution. In CIFQ measurements, Dbo was excited at 373 nm by using a diode laser (PicoQuant, LDH-P-C 375, fwhm ca. 50 ps), and Dbo fluorescence was followed at 450 nm.

**Molecular Dynamics Simulations.** The simulated peptides contained a C-terminal Trp residue (instead of FTrp) and contained no N-terminal Dbo (to bypass parametrization inconsistencies). In their standard form, peptides had a free N-terminus (NH<sub>2</sub>–/NH<sub>3</sub><sup>+</sup>–) and a methylamidated C-terminus (–CONHCH<sub>3</sub>); variations from this pattern are indicated in the main text. We used the AMBER 11 program with the Amber ff99SB force field.<sup>27,28</sup> The trajectories were visualized by using the VMD software package<sup>73</sup> and were analyzed by using the PTRAJ module of AMBER, followed by a cluster analysis based on the average-linkage algorithm with a 3.0 Å cutoff value on a total of 2500 conformations sampled from the simulation, one per every 80 ps.<sup>30</sup> See the Supporting Information for further details.

## ASSOCIATED CONTENT

### Supporting Information

(1) Collisional quenching of DBO by FTrp (Figure S1). (2) pH-dependent fluorescence lifetime of 5-fluoro-L-tryptophan (Figure S2) and time-resolved fluorescence decays of FTrp-Lys<sub>6</sub>-Dbo and FTrp-Lys<sub>6</sub> (Figure S3). (3) CIFQ measurements in the presence of salt (Figures S4–6). (4) Temperature dependence of CIFQ in hexahistidine (Figure S7). (5) Apparent dielectric constants in Asp<sub>6</sub> and Asp<sub>30</sub>. (6) Details of MD simulations. (7) Transient hydrogen bonds in MD simulations (Table S1). (8) Simulation of an acidic chain with tightly bound water (Figure S8). (9) Indications from MD simulations of two-state folding behavior in hexaarginine (Figure S9). (10) The hexahistidine collapse exaggerated in MD simulations. Appendix: Representative structures from cluster analysis. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## AUTHOR INFORMATION

### Corresponding Authors

\*Phone: +49-421-2003233. E-mail: [m.jacob@jacobs-university.de](mailto:m.jacob@jacobs-university.de).

\*E-mail: [w.nau@jacobs-university.de](mailto:w.nau@jacobs-university.de).

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

We thank Indrajit Ghosh for preliminary measurements, and Achim Geleßus, Danilo Roccatano, and Ulrich Kleinekathöfer for MD discussions. The Deutsche Forschungsgemeinschaft (DFG, NA 686/6) supported this work. A.N., S.Z., and K.I.A. thank the DAAD for financial support.

## REFERENCES

- (1) Mao, A. H.; Crick, S. L.; Vitalis, A.; Chicoine, C. L.; Pappu, R. V. Net Charge Per Residue Modulates Conformational Ensembles of Intrinsically Disordered Proteins. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 8183–8188.
- (2) Müller-Spätth, S.; Soranno, A.; Hirschfeld, V.; Hofmann, H.; Ruegger, S.; Reymond, L.; Nettels, D.; Schuler, B. Charge Interactions Can Dominate the Dimensions of Intrinsically Disordered Proteins. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 14609–14614.
- (3) Marsh, J. A.; Forman-Kay, J. D. Sequence Determinants of Compaction in Intrinsically Disordered Proteins. *Biophys. J.* **2010**, *98*, 2383–2390.
- (4) Uversky, V. N.; Oldfield, C. J.; Dunker, A. K. Intrinsically Disordered Proteins in Human Diseases: Introducing the D2 Concept. *Annu. Rev. Biophys.* **2008**, *37*, 215–246.
- (5) Uversky, V. N.; Dunker, A. K. Understanding Protein Non-Folding. *Biochim. Biophys. Acta* **2010**, *1804*, 1231–1264.
- (6) Ganguly, D.; Zhang, W. H.; Chen, J. H. Electrostatically Accelerated Encounter and Folding for Facile Recognition of Intrinsically Disordered Proteins. *PLoS Comput. Biol.* **2013**, *9*, e1003363.
- (7) Hofmann, H.; Nettels, D.; Schuler, B. Single-Molecule Spectroscopy of the Unexpected Collapse of an Unfolded Protein at Low pH. *J. Chem. Phys.* **2013**, *139*.
- (8) Hofmann, H.; Golbik, R. P.; Ott, M.; Hübner, C. G.; Ulbrich-Hofmann, R. Coulomb Forces Control the Density of the Collapsed Unfolded State of Barstar. *J. Mol. Biol.* **2008**, *376*, 597–605.
- (9) Huang, F.; Lerner, E.; Sato, S.; Amir, D.; Haas, E.; Fersht, A. R. Time-Resolved Fluorescence Resonance Energy Transfer Study Shows a Compact Denatured State of the B Domain of Protein A. *Biochemistry* **2009**, *48*, 3468–3476.
- (10) Jacob, M. H.; Dsouza, R. N.; Ghosh, I.; Norouzy, A.; Schwarzlose, T.; Nau, W. M. Diffusion-Enhanced Förster Resonance Energy Transfer and the Effects of External Quenchers and the Donor Quantum Yield. *J. Phys. Chem. B* **2013**, *117*, 185–198.
- (11) Jacob, M. H.; Nau, W. M. In *Folding, Misfolding and Nonfolding of Peptides and Small Proteins*; Schweitzer-Stenner, R., Ed.; John Wiley & Sons: NJ, 2012.
- (12) Sahoo, H.; Roccatano, D.; Hennig, A.; Nau, W. M. A 10-Angstrom Spectroscopic Ruler Applied to Short Polyprolines. *J. Am. Chem. Soc.* **2007**, *129*, 9762–9772.
- (13) Sahoo, H.; Nau, W. M. Phosphorylation-Induced Conformational Changes in Short Peptides Probed by Fluorescence Resonance Energy Transfer in the 10 Angstrom Domain. *ChemBioChem* **2007**, *8*, 567–573.
- (14) Sahoo, H.; Roccatano, D.; Zacharias, M.; Nau, W. M. Distance Distributions of Short Polypeptides Recovered by Fluorescence Resonance Energy Transfer in the 10 Angstrom Domain. *J. Am. Chem. Soc.* **2006**, *128*, 8118–8119.
- (15) Roccatano, D.; Sahoo, H.; Zacharias, M.; Nau, W. M. Temperature Dependence of Looping Rates in a Short Peptide. *J. Phys. Chem. B* **2007**, *111*, 2639–2646.
- (16) Nau, W. M.; Huang, F.; Wang, X. J.; Bakirci, H.; Gramlich, G.; Marquez, C. Exploiting Long-Lived Molecular Fluorescence. *Chimia* **2003**, *57*, 161–167.
- (17) Huang, F.; Nau, W. M. A Conformational Flexibility Scale for Amino Acids in Peptides. *Angew. Chem., Int. Ed.* **2003**, *42*, 2269–2272.
- (18) Hudgins, R. R.; Huang, F.; Gramlich, G.; Nau, W. M. A Fluorescence-Based Method for Direct Measurement of Submicrosecond Intramolecular Contact Formation in Biopolymers: An Exploratory Study with Polypeptides. *J. Am. Chem. Soc.* **2002**, *124*, 556–564.
- (19) Liu, T.; Callis, P. R.; Hesp, B. H.; de Groot, M.; Buma, W. J.; Broos, J. Ionization Potentials of Fluoroindoles and the Origin of Nonexponential Tryptophan Fluorescence Decay in Proteins. *J. Am. Chem. Soc.* **2005**, *127*, 4104–4113.
- (20) Sarkar, S. S.; Udgaonkar, J. B.; Krishnamoorthy, G. Reduced Fluorescence Lifetime Heterogeneity of 5-Fluorotryptophan in Comparison to Tryptophan in Proteins: Implication for Resonance Energy Transfer Experiments. *J. Phys. Chem. B* **2011**, *115*, 7479–7486.
- (21) Huang, F.; Hudgins, R. R.; Nau, W. M. Primary and Secondary Structure Dependence of Peptide Flexibility Assessed by Fluorescence-Based Measurement of End-to-End Collision Rates. *J. Am. Chem. Soc.* **2004**, *126*, 16665–16675.
- (22) Jorda, J.; Kajava, A. V. Protein Homorepeats Sequences, Structures, Evolution, and Functions. *Adv. Protein Chem. Struct. Biol.* **2010**, *79*, 59–88.
- (23) Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Shaw, D. E. How Fast-Folding Proteins Fold. *Science* **2011**, *334*, 517–520.
- (24) Arnold, M. R.; Kremer, W.; Lüdemann, H. D.; Kalbitzer, H. R. <sup>1</sup>H-NMR Parameters of Common Amino Acid Residues Measured in Aqueous Solutions of the Linear Tetrapeptides Gly-Gly-X-Ala at Pressures between 0.1 and 200 MPa. *Biophys. Chem.* **2002**, *96*, 129–140.
- (25) Bundi, A.; Wüthrich, K. <sup>1</sup>H-NMR Parameters of the Common Amino-Acid Residues Measured in Aqueous-Solutions of the Linear Tetrapeptides H-Gly-Gly-X-L-Ala-OH. *Biopolymers* **1979**, *18*, 285–297.
- (26) Dill, K. A.; Bromberg, S. *Molecular Driving Forces*; Garland Science: New York, 2011.
- (27) Simmerling, C.; Strockbine, B.; Roitberg, A. E. All-Atom Structure Prediction and Folding Simulations of a Stable Protein. *J. Am. Chem. Soc.* **2002**, *124*, 11258–11259.
- (28) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins: Struct., Funct., Bioinf.* **2006**, *65*, 712–725.
- (29) Day, R.; Paschek, D.; Garcia, A. E. Microsecond Simulations of the Folding/Unfolding Thermodynamics of the Trp-Cage Mini-protein. *Proteins: Struct., Funct., Bioinf.* **2010**, *78*, 1889–1899.
- (30) Shao, J. Y.; Tanner, S. W.; Thompson, N.; Cheatham, T. E. Clustering Molecular Dynamics Trajectories: 1. Characterizing the Performance of Different Clustering Algorithms. *J. Chem. Theory Comput.* **2007**, *3*, 2312–2334.
- (31) Perl, D.; Müller, U.; Heinemann, U.; Schmid, F. X. Two Exposed Amino Acid Residues Confer Thermostability on a Cold Shock Protein. *Nat. Struct. Biol.* **2000**, *7*, 380–383.
- (32) Pace, C. N.; Alston, R. W.; Shaw, K. L. Charge-Charge Interactions Influence the Denatured State Ensemble and Contribute to Protein Stability. *Protein Sci.* **2000**, *9*, 1395–1398.
- (33) Chellgren, B. W.; Creamer, T. P. Side-Chain Entropy Effects on Protein Secondary Structure Formation. *Proteins: Struct., Funct., Bioinf.* **2006**, *62*, 411–420.
- (34) Trbovic, N.; Cho, J. H.; Abel, R.; Friesner, R. A.; Rance, M.; Palmer, A. G. Protein Side-Chain Dynamics and Residual Conformational Entropy. *J. Am. Chem. Soc.* **2009**, *131*, 615–622.
- (35) Creamer, T. P. Side-Chain Conformational Entropy in Protein Unfolded States. *Proteins: Struct., Funct., Genet.* **2000**, *40*, 443–450.
- (36) Farhangi, S.; Weiss, H.; Duhamel, J. Effect of Side-Chain Length on the Polymer Chain Dynamics of Poly(Alkyl Methacrylate)<sub>S</sub> in Solution. *Macromolecules* **2013**, *46*, 9738–9747.
- (37) Soranno, A.; Longhi, R.; Bellini, T.; Buscaglia, M. Kinetics of Contact Formation and End-to-End Distance Distributions of Swollen Disordered Peptides. *Biophys. J.* **2009**, *96*, 1515–1528.
- (38) Duitch, L.; Toal, S.; Measey, T. J.; Schweitzer-Stenner, R. Triaspartate: A Model System for Conformationally Flexible Ddd Motifs in Proteins. *J. Phys. Chem. B* **2012**, *116*, 5160–5171.
- (39) Rubinstein, M.; Colby, R. H. *Polymer Physics*; Oxford University Press: New York, 2003.

- (40) Dobrynin, A. V.; M, R. Flory Theory of a Polyampholyte Chain. *J. Phys. II France* **1995**, *5*, 677–695.
- (41) Tran, H. T.; Mao, A.; Pappu, R. V. Role of Backbone - Solvent Interactions in Determining Conformational Equilibria of Intrinsically Disordered Proteins. *J. Am. Chem. Soc.* **2008**, *130*, 7380–7392.
- (42) Das, R. K.; Pappu, R. V. Conformations of Intrinsically Disordered Proteins Are Influenced by Linear Sequence Distributions of Oppositely Charged Residues. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 13392–13397.
- (43) Manning, G. S. Limiting Laws and Counterion Condensation in Polyelectrolyte Solutions. 3. An Analysis Based on Mayer Ionic Solution Theory. *J. Chem. Phys.* **1969**, *51*, 3249–3252.
- (44) Manning, G. S.; Ray, J. Counterion Condensation Revisited. *J. Biomol. Struct. Dyn.* **1998**, *16*, 461–476.
- (45) Stigter, D. Evaluation of the Counterion Condensation Theory of Polyelectrolytes. *Biophys. J.* **1995**, *69*, 380–388.
- (46) O'Shaughnessy, B.; Yang, Q. Manning-Oosawa Counterion Condensation. *Phys. Rev. Lett.* **2005**, *94*.
- (47) Manning, G. S. Excess Counterion Condensation on Polyelectrolyte Kinks and Branch Points and the Interaction of Skewed Charged Lines. *Soft Matter* **2014**, *10*, 3738–3747.
- (48) Haber-Pohlmeier, S.; Abarca-Heidemann, K.; Korschen, H. G.; Dhiman, H. K.; Heberle, J.; Schwalbe, H.; Klein-Seetharaman, J.; Kaupp, U. B.; Pohlmeier, A. Binding of Ca<sup>2+</sup> to Glutamic Acid-Rich Polypeptides from the Rod Outer Segment. *Biophys. J.* **2007**, *92*, 3207–3214.
- (49) Fowler, M.; Siddique, B.; Duhamel, J. Effect of Sequence on the Ionization Behavior of a Series of Amphiphilic Polypeptides. *Langmuir* **2013**, *29*, 4451–4459.
- (50) Schaefer, M.; Karplus, M. A Comprehensive Analytical Treatment of Continuum Electrostatics. *J. Phys. Chem.* **1996**, *100*, 1578–1599.
- (51) Olsson, M. H. M. Improving the Desolvation Penalty in Empirical Protein pK<sub>a</sub> Modeling. *J. Mol. Model.* **2012**, *18*, 1097–1106.
- (52) Jacob, M. H.; Amir, D.; Ratner, V.; Gussakowsky, E.; Haas, E. Predicting Reactivities of Protein Surface Cysteines as Part of a Strategy for Selective Multiple Labeling. *Biochemistry* **2005**, *44*, 13664–13672.
- (53) Ramachandran, S.; Katha, A. R.; Kolake, S. M.; Jung, B.; Han, S. Dynamics of Dilute Solutions of Poly(Aspartic Acid) and Its Sodium Salt Elucidated from Atomistic Molecular Dynamics Simulations with Explicit Water. *J. Phys. Chem. B* **2013**, *117*, 13906–13913.
- (54) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (55) Horn, H. W.; Swope, W. C.; Pitner, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an Improved Four-Site Water Model for Biomolecular Simulations: Tip4p-Ew. *J. Chem. Phys.* **2004**, *120*, 9665–9678.
- (56) Kusalik, P. G.; Svishchev, I. M. The Spatial Structure in Liquid Water. *Science* **1994**, *265*, 1219–1221.
- (57) Kiriukhin, M. Y.; Collins, K. D. Dynamic Hydration Numbers for Biologically Important Ions. *Biophys. Chem.* **2002**, *99*, 155–168.
- (58) Collins, K. D. Ions from the Hofmeister Series and Osmolytes: Effects on Proteins in Solution and in the Crystallization Process. *Methods* **2004**, *34*, 300–311.
- (59) Collins, K. D.; Neilson, G. W.; Enderby, J. E. Ions in Water: Characterizing the Forces That Control Chemical Processes and Biological Structure. *Biophys. Chem.* **2007**, *128*, 95–104.
- (60) Collins, K. D. Why Continuum Electrostatics Theories Cannot Explain Biological Structure, Polyelectrolytes or Ionic Strength Effects in Ion-Protein Interactions. *Biophys. Chem.* **2012**, *167*, 43–59.
- (61) Rahman, H. M. A.; Hefter, G.; Buchner, R. Hydration of Formate and Acetate Ions by Dielectric Relaxation Spectroscopy. *J. Phys. Chem. B* **2012**, *116*, 314–323.
- (62) Fedotova, M. V.; Kruchinin, S. E. Hydration of Acetic Acid and Acetate Ion in Water Studied by 1D-RISM Theory. *J. Mol. Liq.* **2011**, *164*, 201–206.
- (63) Liang, T.; Walsh, T. R. Molecular Dynamics Simulations of Peptide Carboxylate Hydration. *Phys. Chem. Chem. Phys.* **2006**, *8*, 4410–4419.
- (64) Grimsley, G. R.; Scholtz, J. M.; Pace, C. N. A Summary of the Measured pK Values of the Ionizable Groups in Folded Proteins. *Protein Sci.* **2009**, *18*, 247–251.
- (65) Gallivan, J. P.; Dougherty, D. A. Cation- $\pi$  Interactions in Structural Biology. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9459–9464.
- (66) Trexler, A. J.; Rhoades, E. N-Terminal Acetylation Is Critical for Forming  $\alpha$ -Helical Oligomer of  $\alpha$ -Synuclein. *Protein Sci.* **2012**, *21*, 601–605.
- (67) Grupi, A.; Haas, E. Segmental Conformational Disorder and Dynamics in the Intrinsically Disordered Protein  $\alpha$ -Synuclein and Its Chain Length Dependence. *J. Mol. Biol.* **2011**, *405*, 1267–1283.
- (68) Maltsev, A. S.; Ying, J. F.; Bax, A. Impact of N-Terminal Acetylation of Alpha-Synuclein on Its Random Coil and Lipid Binding Properties. *Biochemistry* **2012**, *51*, 5004–5013.
- (69) Heyda, J.; Mason, P. E.; Jungwirth, P. Attractive Interactions between Side Chains of Histidine-Histidine and Histidine-Arginine-Based Cationic Dipeptides in Water. *J. Phys. Chem. B* **2010**, *114*, 8744–8749.
- (70) Marsili, S.; Chelli, R.; Schettino, V.; Procacci, P. Thermodynamics of Stacking Interactions in Proteins. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2673–2685.
- (71) Haghani, K.; Khajeh, K.; Naderi-Manesh, H.; Ranjbar, B. Evidence Regarding the Hypothesis That the Histidine-Histidine Contact Pairs May Affect Protein Stability. *Int. J. Biol. Macromol.* **2012**, *50*, 1040–1047.
- (72) Vondrasek, J.; Mason, P. E.; Heyda, J.; Collins, K. D.; Jungwirth, P. The Molecular Origin of Like-Charge Arginine-Arginine Pairing in Water. *J. Phys. Chem. B* **2009**, *113*, 9041–9045.
- (73) Humphrey, W.; Dalke, A.; Schulten, K. Vmd: Visual Molecular Dynamics. *J. Mol. Graphics Modell.* **1996**, *14*, 33–38.

# Supporting Information

## Coulomb Repulsion in Short Polypeptides

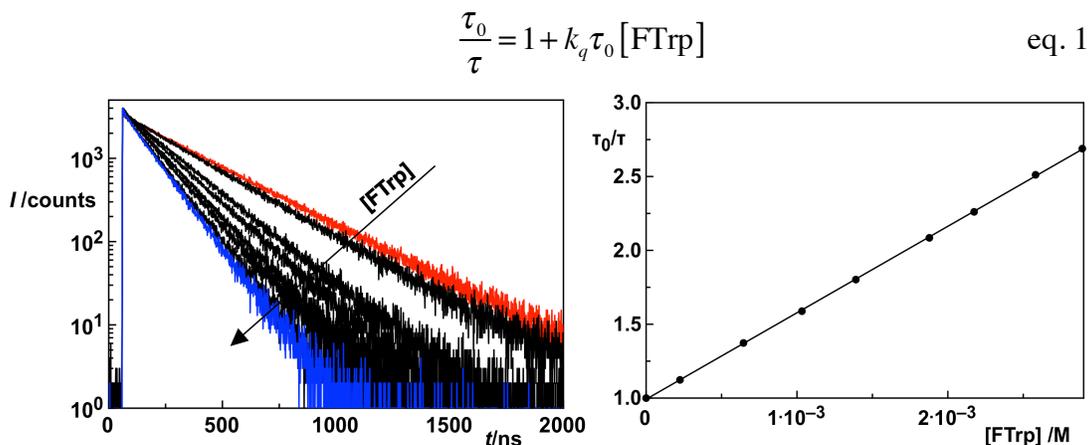
Amir Norouzy, Khaleel I. Assaf, Maik H. Jacob\* and Werner M. Nau\*  
School of Engineering and Science, Jacobs University Bremen, Campus Ring 1, D-28759 Bremen, Germany

### Table of Contents

1. Collisional quenching of DBO by FTrp .....	S-2
2. pH-dependent fluorescence lifetime of 5-fluoro-L-tryptophan .....	S-2
3. CIFQ measurements in the presence of salt .....	S-3
3.1 Hexaaspartate.....	S-3
3.2 Hexaarginine.....	S-4
3.3 Hexalysine .....	S-5
4. Temperature dependence of CIFQ in hexahistidine .....	S-6
5. Apparent dielectric constant in Asp <sub>6</sub> and Asp <sub>30</sub> .....	S-7
6. Details of MD simulations.....	S-7
7. Transient hydrogen bonds in MD simulations .....	S-8
8. Simulation of an acidic chain with tightly bound water .....	S-9
9. Indications from MD simulations for two-state folding behavior in hexaarginine .....	S-9
10. The hexahistidine collapse exaggerated in MD simulations .....	S-10
11. Representative structures from cluster analysis.....	S-11
12. References .....	S-16

## 1. Collisional quenching of DBO by FTrp

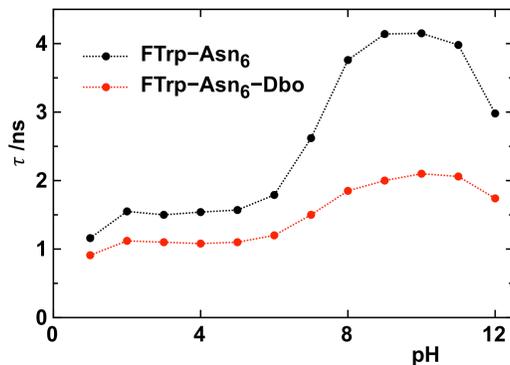
The time-resolved fluorescence of parent DBO was measured in the presence of increasing amounts of FTrp (Figure S1, left). The DBO lifetime decreased from  $\tau_0 = 311$  ns in the absence of FTrp (red trace) to  $\tau = 116$  ns in the presence of 2.9 mM FTrp (blue trace) at pH 7.0 and 25.0 °C. The kinetic analysis was based on equation 1 (Figure S1, right) and yielded a bimolecular quenching rate constant of  $k_q = 1.9 \times 10^9 \text{ M}^{-1}\text{s}^{-1}$ , almost identical to the quenching rate constant  $k_q = 2.1 \times 10^9 \text{ M}^{-1}\text{s}^{-1}$ , reported for DBO fluorescence quenching by Trp.<sup>1</sup>



**Figure S1.** Time-resolved fluorescence of parent DBO in the presence of increasing amounts of FTrp (left), and kinetic analysis (right)

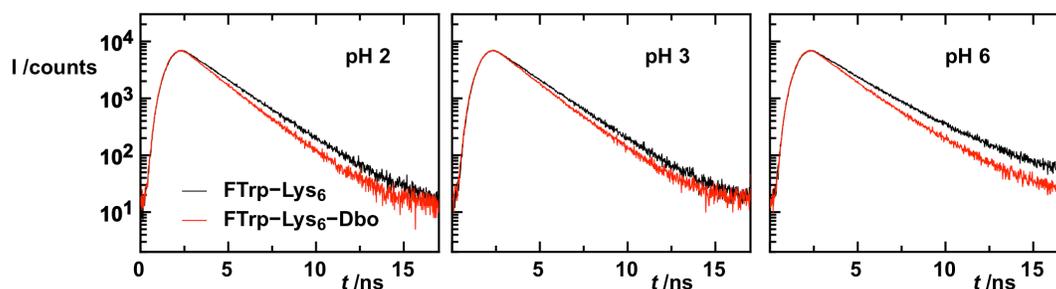
## 2. pH-dependent fluorescence lifetime of 5-fluoro-L-tryptophan

The quantum yield of Trp and also FTrp (Figure S2) depend on the solution pH.



**Figure S2.** The fluorescence lifetime of 5-fluoro-L-tryptophan in FTrp-Asn<sub>6</sub> (black) and FTrp-Asn<sub>6</sub>-Dbo plotted against the solution pH

This dependency was the main motivation to always include the donor-only reference peptide into the analysis. Figure S3 illustrates the capability of sdFRET measurements to resolve minute distance changes. In an independent control measurement on the hexalysine peptide, the effective distance decreased from 13.6 Å at pH 3 to 12.7 Å at pH 6. The lifetimes of donor-acceptor peptide and donor-only reference peptide were obtained, after deconvolution, with an accuracy of  $\pm 0.1\%$ , which corresponds to an error of the effective distance of  $\pm 0.05$  Å. The larger error given in the main text takes the limited sample purity ( $> 95\%$ ) into account.

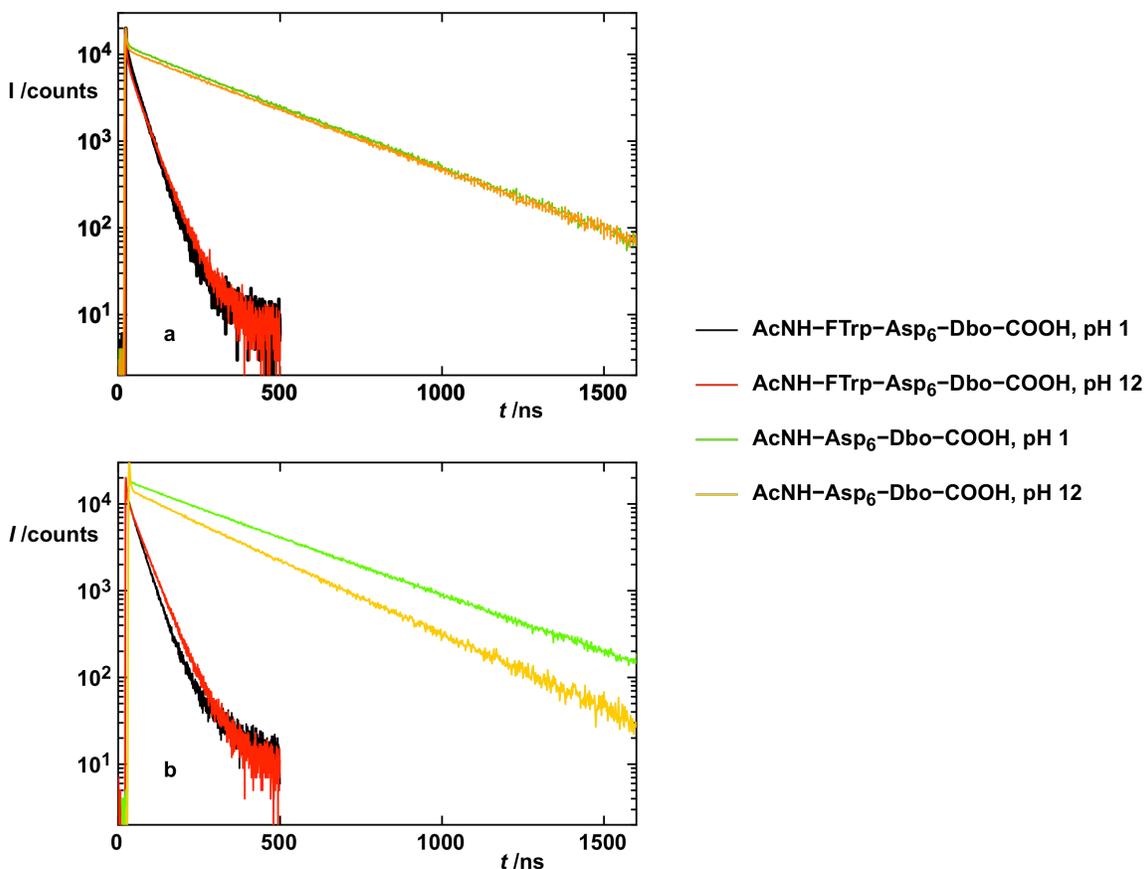


**Figure S3.** Time-resolved fluorescence decays of FTrp-Lys<sub>6</sub>-Dbo and FTrp-Lys<sub>6</sub> (reference peptide) at pH 2, 3 and 6.

### 3. CIFQ measurements in the presence of salt

#### 3.1 Hexaaspartate

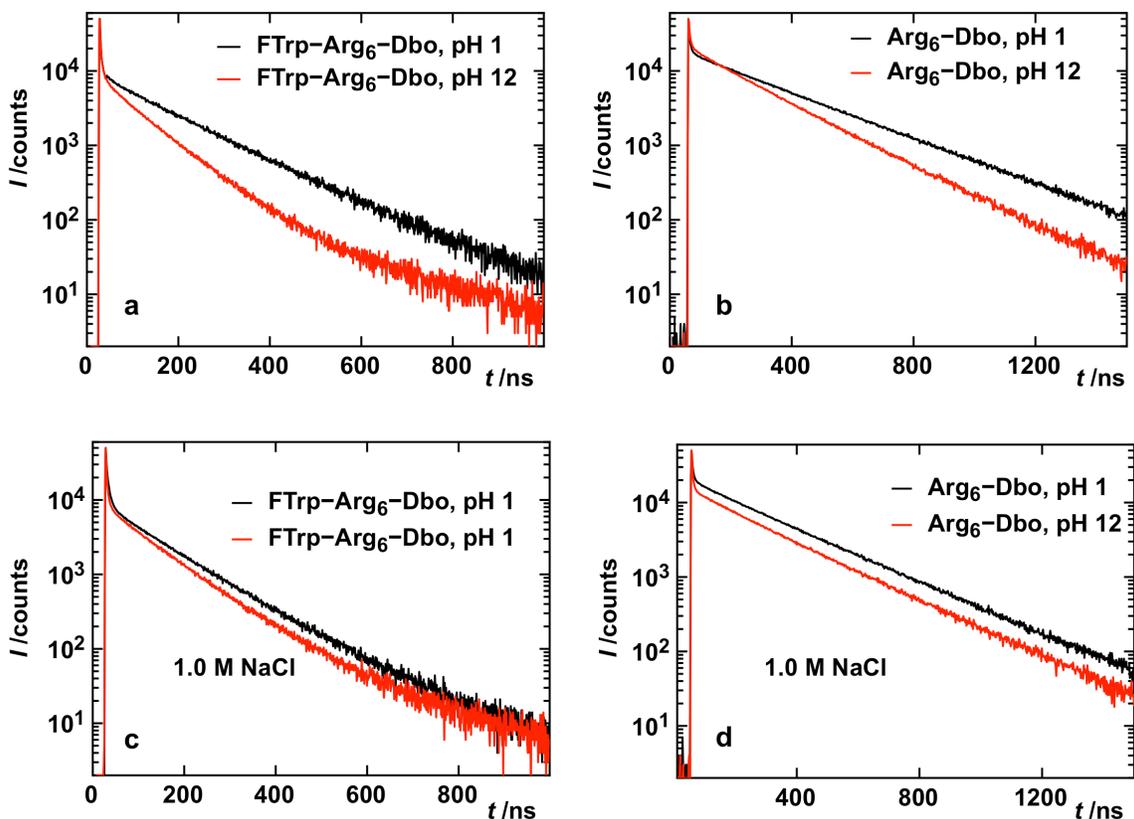
Neither the collision rate nor the effective distance of hexaaspartate chains displayed any dependence on the charge status of the chain, neither in the absence nor in the presence of salt. This is instantly visible from CIFQ measurements, owing to the pH independence of the Dbo fluorescence signal (Figure S4).



**Figure S4.** (a) CIFQ measurements on the hexaaspartate polypeptide in the absence of salt. The fluorescence kinetics of Asp<sub>6</sub> and of its reference peptide is pH independent. (b) CIFQ measurements on the hexaaspartate polypeptide in the presence of salt (1.0 M NaCl) at pH 1 and 12.

### 3.2 Hexaarginine

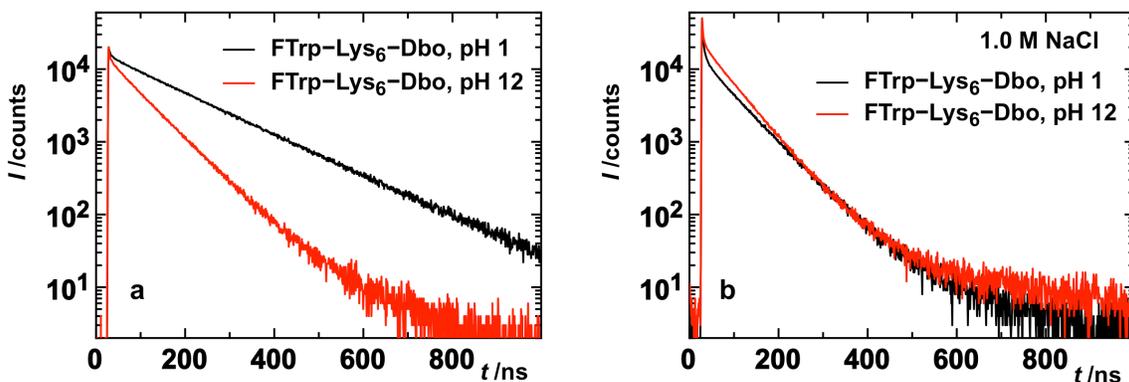
Dbo in the hexaarginine chain in water (Figure S5a) is faster deactivated at pH 12 than at pH 1. The larger collision rate at high pH points to a compact subpopulation that allows for rapid end-to-end contacts. This difference was not observed in the presence of salt (1 M NaCl, Figure S5c). When the arginine guanidinium groups become neutralized around pH 12.5, they act as benign quenchers of Dbo fluorescence (Figure S5b,d).



**Figure S5.** CIFQ measurements on the hexaarginine polypeptide in the absence (a, b) and presence (c, d) of 1.0 M NaCl. Fluorescence decay traces of the double-labeled peptide (a, c) and the single-labeled reference peptide (b, d).

### 3.3 Hexalysine

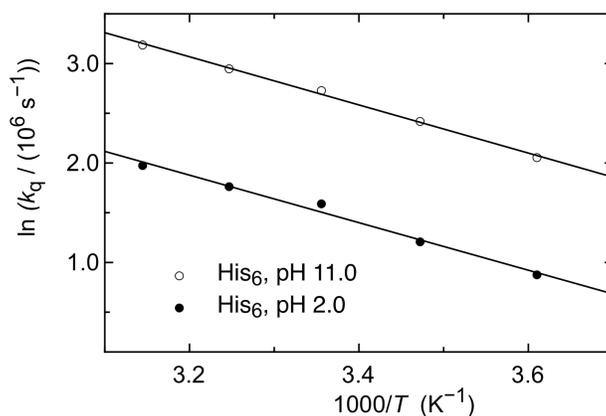
Dbo in the hexalysine chain in water (Figure S6a) is faster deactivated at pH 12 than at pH 1. The larger collision rate at high pH is caused by a higher stability of short-distance subpopulations that allow for rapid end-to-end contact. In contrast, in the presence of salt, 1 M NaCl, the fluorescence decay kinetic traces at low and high pH were virtually identical (Figure S6b).



**Figure S6.** CIFQ measurements on the hexalysine polypeptide at pH 1 and 12 in the absence (a) and presence (b) of 1.0 M NaCl.

#### 4. Temperature dependence of CIFQ in hexahistidine

The collision rate of His<sub>6</sub> was measured between 4.0 and 45 °C at pH 2.0 and 11.0 (Figure S7). The activation energy computed from the Arrhenius plots remained at ca. 20 kJ/mol, when going from pH 2.0 to pH 11.0—from the protonated to the neutral chain, yielding no indications for more than marginally stable interactions or hydrogen bonds.



**Figure S7.** Arrhenius plots for His<sub>6</sub> at pH 2.0 and pH 11.0. Activation energies and pre-exponential factors are  $(19.9 \pm 1.3)$  kJ/mol and  $(13.8 \pm 5.4) \cdot 10^6$  s<sup>-1</sup> at pH 2.0 as well as  $(20.1 \pm 0.6)$  kJ/mol and  $(50.0 \pm 11.1) \cdot 10^6$  s<sup>-1</sup> at pH 11.0.

## 5. Apparent dielectric constant in Asp<sub>6</sub> and Asp<sub>30</sub>

The transfer of a charge from an aqueous high to a low-permittivity environment is accompanied by an unfavorable desolvation energy also called Born penalty, which depends on the apparent dielectric permittivity  $\epsilon_{app}$ , on the permittivity of water  $\epsilon_{water}$ , and on the radius of the charged atom  $b/\text{\AA}$ ,  $W_{Born} = 1.39 \times 10^6 (\epsilon_{app}^{-1} - \epsilon_{water}^{-1})/2b$  [J/mol]<sup>2</sup>. For instance, an extremely large Born penalty of 165 kJ/mol prohibits charge transfer from water with a dielectric constant of 78.5 at room temperature into the hydrophobic core of a protein with a local dielectric constant of  $\epsilon_{core} \approx 4$ . As we did in reference<sup>3</sup>, the literature basis for the Cys pK<sub>a</sub> predictor software Cyspka<sup>4</sup>, we assumed that the apparent dielectric constants,  $\epsilon_{app}$ , at the locations of chain atoms adopt intermediate values between  $\epsilon_{core}$  and  $\epsilon_{water}$ . To roughly estimate  $\epsilon_{app}$  at the location of a specified chain atom, a surrounding spherical volume,  $V_{sphere}$ , with radius  $r = 7.0 \text{\AA}$  was considered. All chain heavy atoms occupying this space were counted and yielded the “neighbor count”  $N_{NC}$ . Chain heavy atoms and water molecules were modeled as spheres of radius  $1.4 \text{\AA}$  and volume  $V = 4/3\pi \cdot 1.4^3 \text{\AA}^3$ , and the number of water molecules  $N_{water}$  was calculated by using equation 1. The factor 0.74 considers excluded-volume effects; solid spheres of equal size can fill a space with a maximal packing efficiency of 74%.

$$(N_{water} + N_{NC}) \cdot V = 0.74 \cdot V_{sphere} \quad (1)$$

Water molecules and chain atoms were assumed to contribute to  $\epsilon_{app}$  in proportion to their numbers (equation 2).

$$\epsilon_{app} = (N_{water} \cdot \epsilon_{water} + N_{NC} \cdot \epsilon_{core}) / (N_{water} + N_{NC}) \quad (2)$$

Taking the average of all local neighbor counts and substituting it into equation 1 and 2 afforded the average dielectric constants in the Asp<sub>6</sub> and Asp<sub>30</sub> chain, 10 and 33. The coordinates of chain atoms were obtained from conformations of protonated Asp<sub>6</sub> and Asp<sub>30</sub> encountered after 100 ns of MD simulation.

## 6. Details of MD simulations

The starting structures were modelled as extended conformations. Each peptide system was solvated in a truncated octahedral box of TIP3P water molecules with a closeness parameter of  $15 \text{\AA}$  to the peptide boundary. Counter ions were added to maintain

neutrality. Periodic boundary conditions were adopted, and the Particle Mesh Ewald's method (PME) was used for the treatment of long-range electrostatic interaction. The non-bonded cut-off value was set to 9.0 Å. Energy minimization was carried out by using the conjugate gradient algorithm, heating up to 298 K for 600 ps, followed by 5 ns of equilibration. An NPT ensemble was assured by coupling the system to a Berendsen thermostat set to 298 K and to a barostat set to 1 atm. Production runs were carried out for 200 ns or 500 ns (arginine hexamer with acetylated N-terminal FTrp).

## 7. Transient hydrogen bonds in MD simulations

Required was a comprehensive parameter that allowed us to compare and follow the extent of intra-mainchain and mainchain-sidechain interactions (Table S1). We translated the generally small individual probabilities,  $p_i$ , of MD-detected transient hydrogen bonds into the probability that at least one hydrogen bond exists at any given time, calculated from  $p_{\text{alo}} = 1 - \prod_i (1 - p_i)$ . Here, it is required to assume that the hydrogen bonds can form independently from each other.

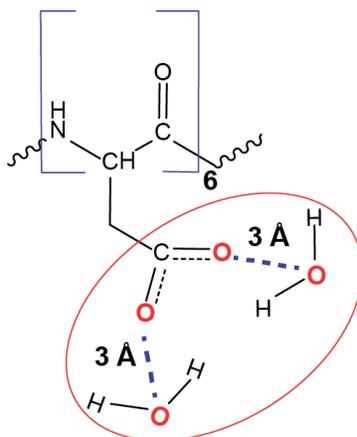
**Table S1.** Transiently stable mainchain (M) and sidechain (S) hydrogen bonds<sup>a</sup>

Peptide (charge status)	M---M	M---S	S---M	S---S
NH <sub>3</sub> -Trp-Asp <sub>6</sub> (+1)	22.6	12.5	2.0	4.2
NH <sub>2</sub> -Trp-Asp <sub>6</sub> (-6)	7.0	22.1	0.6	1.2
NH <sub>2</sub> -Trp-Asp <sub>6</sub> -COO (-7)	3.6	14.4	0.2	2.8
NH <sub>3</sub> -Trp-Arg <sub>6</sub> (+7)	0.6	-	11.8	-
NH <sub>2</sub> -Trp-Arg <sub>6</sub> (+6)	3.6	-	14.6	-
NH <sub>2</sub> -Trp-Arg <sub>6</sub> (+0)	14.2	2.5	21.4	14.1
NH <sub>3</sub> -Trp-Lys <sub>6</sub> (+7)	0.4	-	4.9	-
NH <sub>2</sub> -Trp-Lys <sub>6</sub> (+0)	8.7	1.4	3.5	0.9
NH <sub>3</sub> -Trp-His <sub>6</sub> (+7)	-	-	4.9	-
NH <sub>3</sub> -Trp-His <sub>6</sub> (+1)	1.6	9.4	3.1	0.4

<sup>a</sup> Values of the probability,  $p_{\text{alo}}$  (%), that, at any given time, at least one (alo) hydrogen bond is in place, as calculated from  $p_{\text{alo}} = 1 - \prod_i (1 - p_i)$  from the individual probabilities (occupancies),  $p_i$ , of the hydrogen bonds in the hydrogen donor-acceptor classes, M---M, M---S, S---M and S---S; values smaller than 0.2 % are not listed.

## 8. Simulation of an acidic chain with tightly bound water

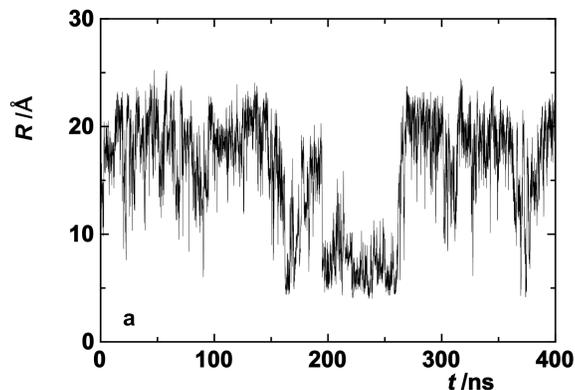
We modeled the strong hydration of the carboxylate group by preventing the exchange of bound and bulk water through a harmonic potential between water and carboxylate oxygen atoms (Figure S8).



**Figure S8.** Carboxylate-bound water molecules were hindered to exchange with bulk water by constraining the distance between water oxygen and carboxylate oxygen through a harmonic potential at 3 Å and a force constant of  $105 \text{ kJ}\cdot\text{mol}^{-1}\text{Å}^{-1}$ .

## 9. Indications from MD simulations for two-state folding behavior in hexaarginine

The distance trajectory of  $\text{H}_2\text{N-Trp-Arg}_6(+6)$  displayed a pattern reminiscent of two-state protein folding (Figure S9). After 200 ns the chain became compact and remained in a folded state for more than 60 ns stabilized by arginine–tryptophan interaction (Figure 6a, b in the main text). Such a two-state behavior was not found in any of the other simulations. Simulation and experiment agreed well: Both CIFQ and sFRET indicated chain compaction when the pH was raised and the charge status of the N-terminus was changed from charged to neutral (Figure 4c, d in the main text).



**Figure S9.** MD simulations on the hexaarginine system. The end-to-end distance in H<sub>2</sub>N-Trp-Arg<sub>6</sub>(+6) has been plotted against time. The C-terminus was methylamidated.

### 10. The hexahistidine collapse exaggerated in MD simulations

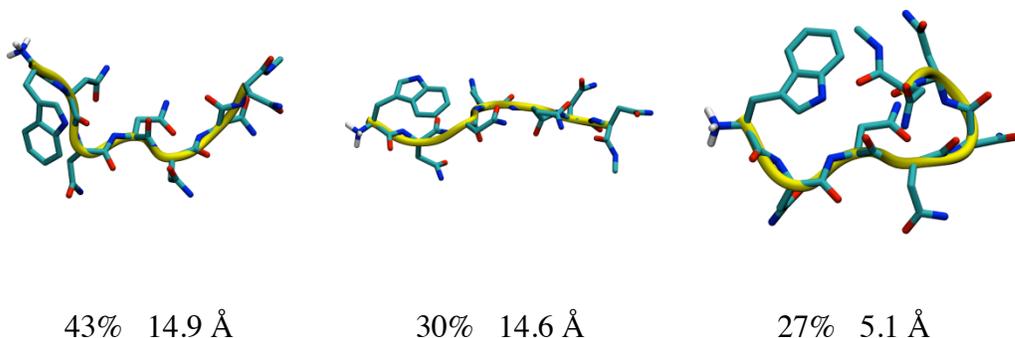
When going from the charged (+7) to the deprotonated (+1) hexahistidine chain, the mainchain-sidechain interaction probability,  $p_{\text{alo}}$ , increases from 4.9 % to 12.5% because the imidazole ring can now also act as hydrogen-bond acceptor. Accordingly, the MD simulations indicated a distance reduction of 5.5 Å, which is severely larger than the experimentally detected change of the effective distance of 1.6 Å. The neutral imidazole ring can simultaneously act as hydrogen donor and acceptor for the peptide backbone such that interaction probability and duration are strongly increased. In any MD simulation, hydrogen bonds are formed through electrostatic interactions between partial charges fixed on electronegative and -positive atoms. In reality, however, the hydrogen-donor and acceptor functions of the two nitrogen moieties in the neutral imidazole ring can change places. This flexibility is lost when the ring binds to the backbone, which offers only a fixed hydrogen-donor (-NH) and acceptor (-CO) functionality. This transient “tautomer trapping” causes an energy penalty that no pure MD simulation takes into account.

## 11. Representative structures from cluster analysis

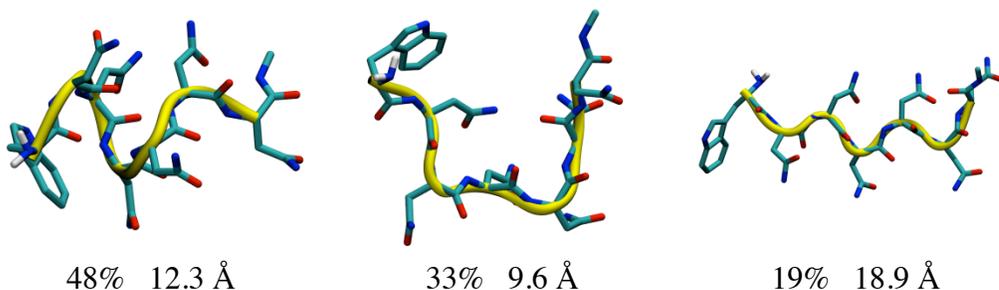
The MD trajectories for each peptide were clustered into three distinct groups based on the average linkage method, as recommended by Shao et al.<sup>5</sup>. The coordinates of each snapshot are compared with the average coordinates of each group, and the main-chain RMSD value is evaluated. The snapshot conformation joins a cluster if the corresponding RMSD value is smaller than the given cutoff (3 Å in our case). For reference, we display representative structures of the three largest clusters, their populations (%) and average distances (Å).

### A. Hexaasparagine systems

#### NH<sub>3</sub>-Trp-Asn<sub>6</sub> (+1)

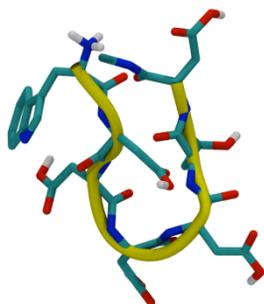


#### NH<sub>2</sub>-Trp-Asn<sub>6</sub> (+0)

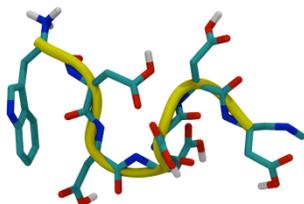


## B. Hexaaspartate Systems

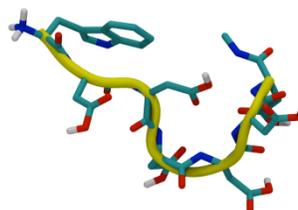
### NH<sub>3</sub>-Trp-Asp<sub>6</sub> (+1)



48% 6.7 Å

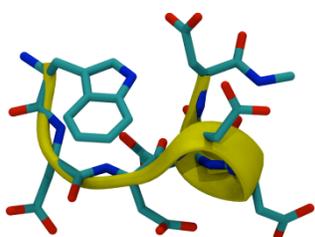


28% 13.7 Å

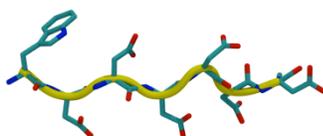


24% 8.7 Å

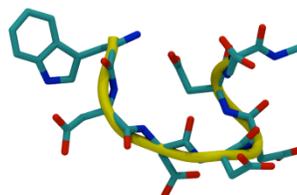
### NH<sub>3</sub>-Trp-Asp<sub>6</sub> (-5)



54% 8.1 Å

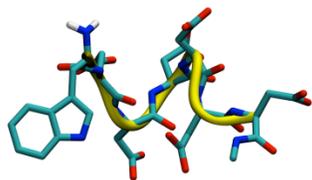


26% 18.5 Å

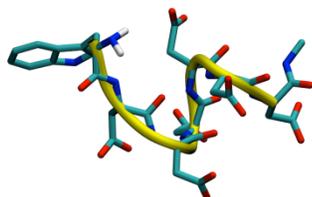


20% 11.9 Å

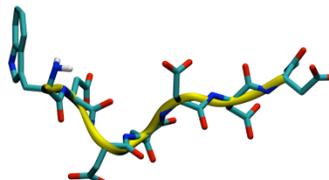
### NH<sub>2</sub>-Trp-Asp<sub>6</sub> (-6)



52% 13.7 Å

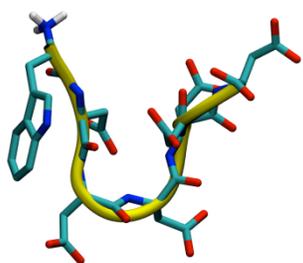


37% 11.4 Å

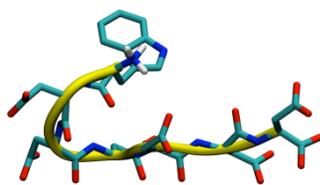


11% 18.4 Å

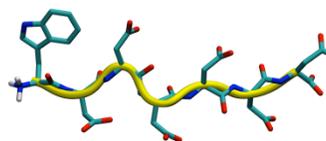
**NH<sub>3</sub>-Trp-Asp<sub>6</sub>-COO (-6)**



67% 6.8 Å

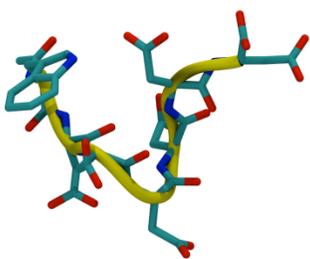


17% 10.0 Å

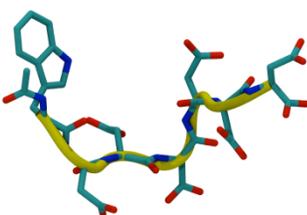


16% 17.7 Å

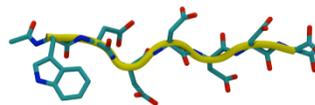
**AcNH-Trp-Asp<sub>6</sub>-COO (-7)**



39% 10.4 Å

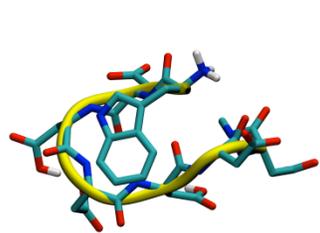


34% 15.1 Å

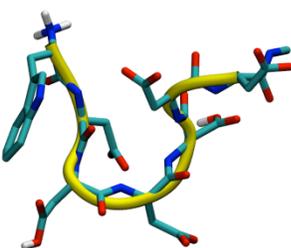


27% 20.6 Å

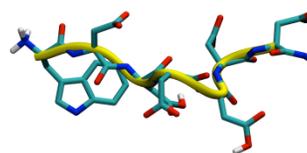
**NH<sub>3</sub>-Trp-Asp(-1)-Asp(+0)-Asp(-1)-Asp(+0)-Asp(-1)-Asp(+0)**



46% 8.5 Å



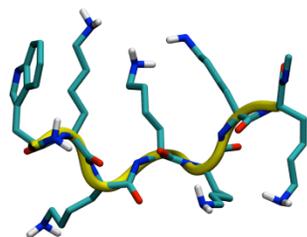
27% 12.0 Å



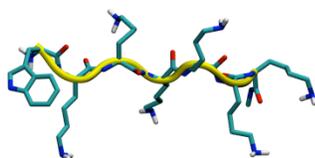
27% 13.8 Å

### C. Hexalysine systems

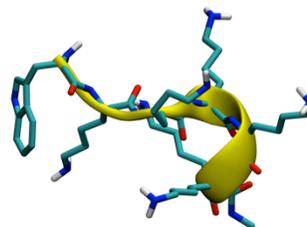
#### $\text{NH}_2\text{-Trp-Lys}_6$ (+0)



39% 14.9 Å

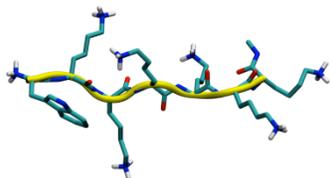


36% 19.8 Å

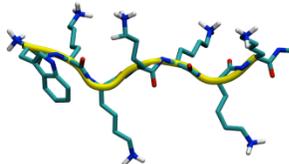


25% 13.6 Å

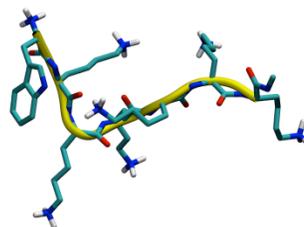
#### $\text{NH}_3\text{-Trp-Lys}_6$ (+7)



48% 17.0 Å

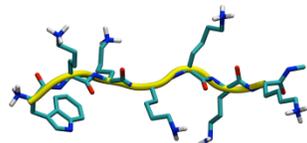


43% 18.6 Å

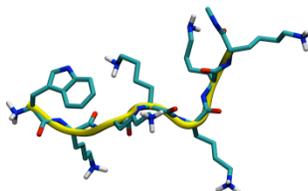


9% 16.1 Å

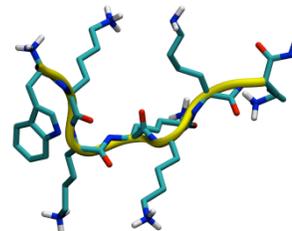
#### $\text{NH}_3\text{-Trp-Lys(+1)-Lys(+0)-Lys(+1)-Lys(+0)-Lys(+1)-Lys(+0)}$



61% 19.3 Å



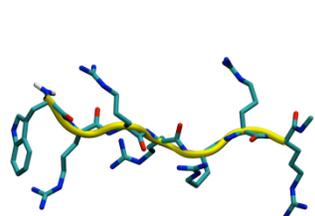
26% 14.3 Å



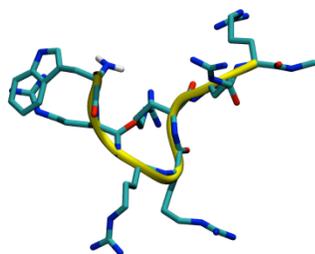
13% 15.2 Å

## D. Hexaarginine systems

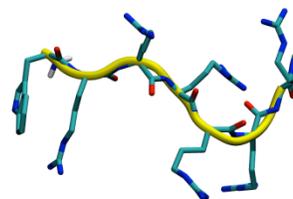
### $\text{NH}_2\text{-Trp-Arg}_6$ (+6)



51% 22.4 Å

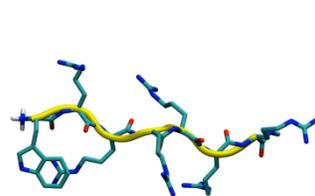


33% 15.8 Å

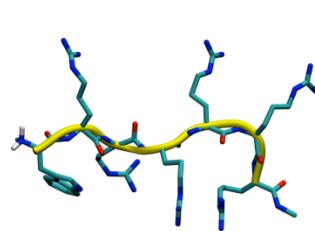


16% 18.9 Å

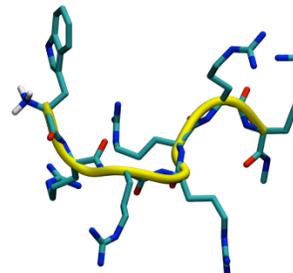
### $\text{NH}_3\text{-Trp-Arg}_6$ (+7)



52% 20.4 Å

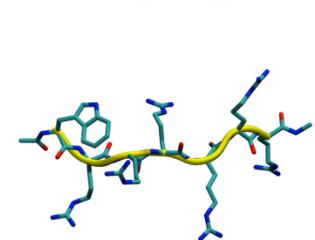


34% 16.0 Å

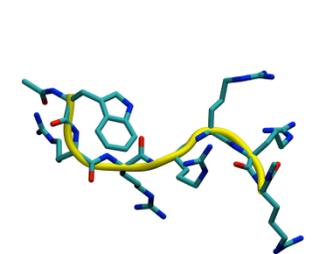


14% 15.6 Å

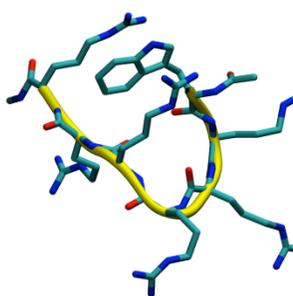
### $\text{AcNH-Trp-Arg}_6$ (+6)



57% 18.3 Å

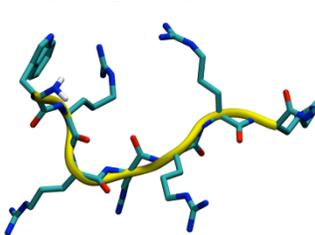


28% 12.1 Å

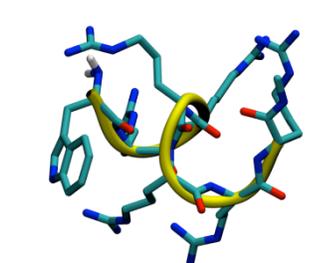


15% 7.9 Å

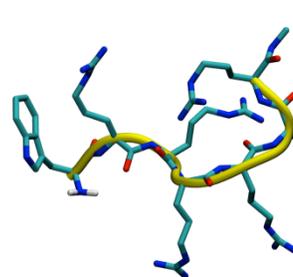
### $\text{NH}_2\text{-Trp-Arg}_6$ (+0)



41% 16.9 Å



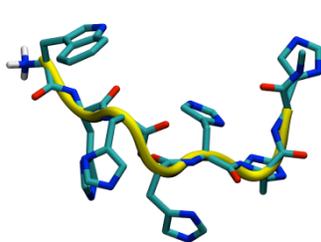
30% 10.7 Å



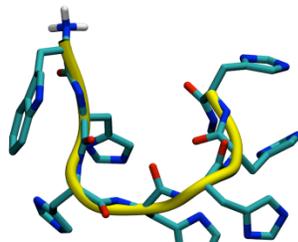
29% 16.2 Å

## E. Hexahistidine Systems

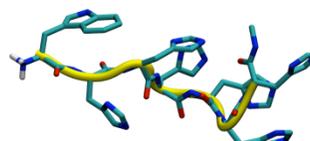
### NH<sub>3</sub>-Trp-His<sub>6</sub> (+1)



46% 12.3 Å

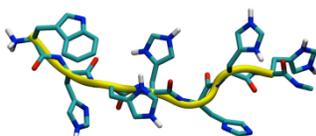


34% 11.9 Å

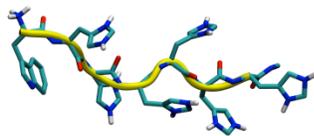


20% 11.8 Å

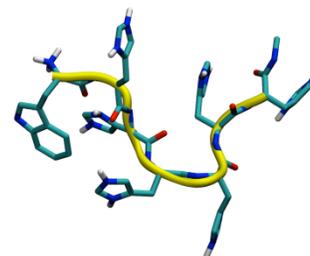
### NH<sub>3</sub>-Trp-His<sub>6</sub> (+7)



50% 16.7 Å



42% 19.0 Å



8% 15.5 Å

## 12. References

- (1) Huang, F.; Hudgins, R. R.; Nau, W. M. *J. Am. Chem. Soc.* **2004**, *126*, 16665-16675.
- (2) Pokala, N.; Handel, T. M. *Protein Sci.* **2004**, *13*, 925-936.
- (3) Jacob, M. H.; Amir, D.; Ratner, V.; Gussakowsky, E.; Haas, E. *Biochemistry* **2005**, *44*, 13664-72.
- (4) Linnet, T. E. Cyspka. [Online Early Access].  
<http://www.pymolwiki.org/index.php/Cyspka> (accessed May 7, 2014).
- (5) Shao, J. Y.; Tanner, S. W.; Thompson, N.; Cheatham, T. E. *J. Chem. Theory Comput.* **2007**, *3*, 2312-2334.

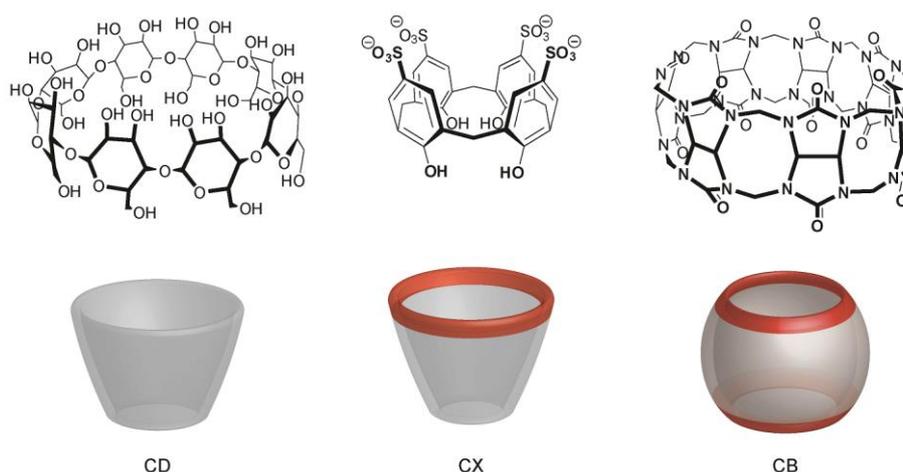
# Synthetic Macrocyclic Receptors as Tools in Drug Delivery and Drug Discovery

Amir Norouzy and Werner M. Nau

Jacobs University Bremen, Campus Ring 1, 28759 Bremen, Germany

E-mail: w.nau@jacobs-university.de

Macrocyclic compounds (Figure 1) have fascinated medicinal chemists since generations. The first ones which have been characterized are the naturally occurring cyclodextrins (CDs), cyclic oligomers of glucose, which can be obtained by bacterial synthesis from starch. Calixarenes (CXs) were among the first man-made macrocycles, followed by crown ethers (Pedersen, Nobel Prize 1987) and cucurbiturils (CBs). While initially challenges related to their efficient synthesis and derivatization needed to be tackled, their functions as molecular receptors have subsequently moved into the focus, which is the basis for the pharmaceutical and medicinal-diagnostic applications summarized in the following.

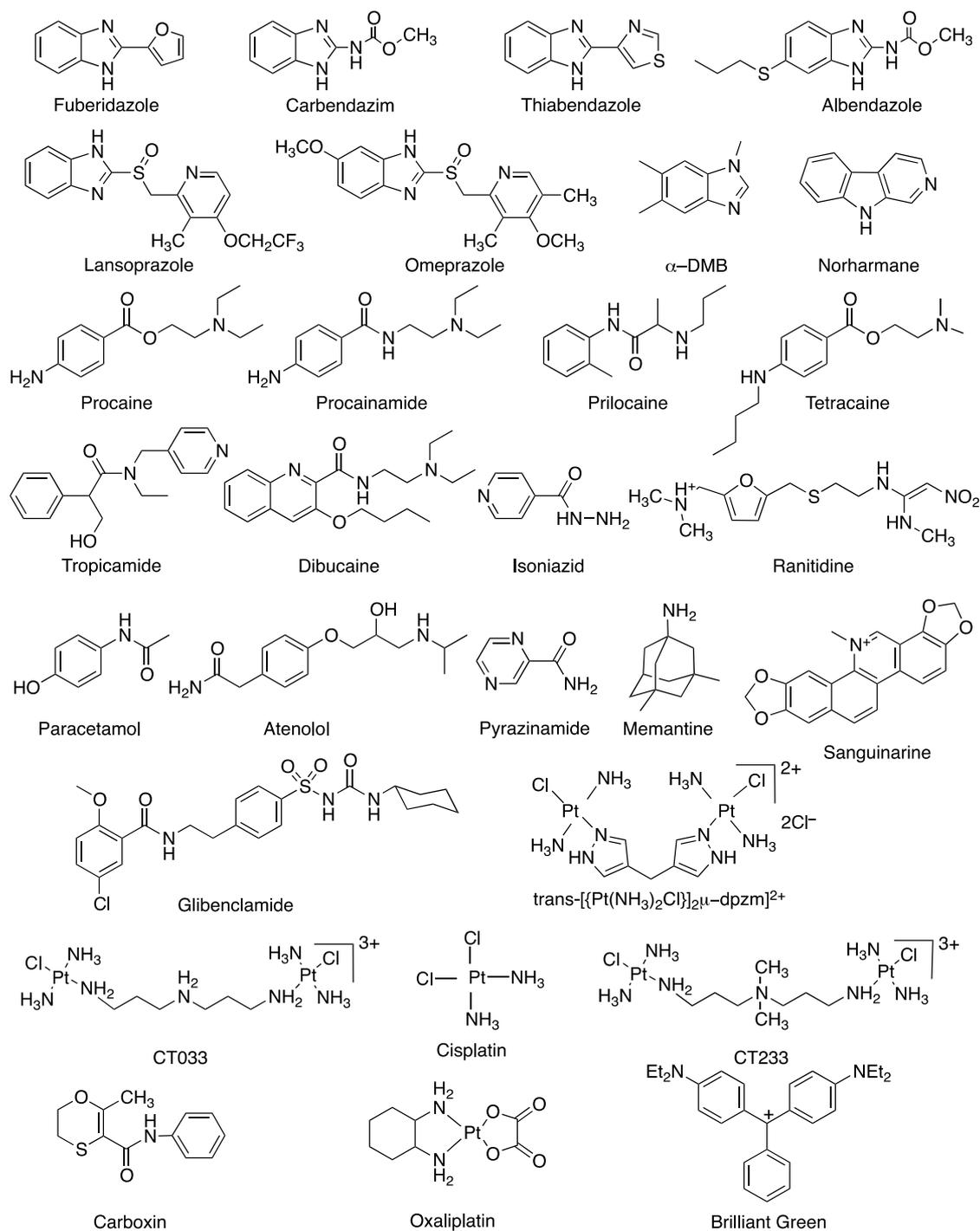


**Figure 1.** Chemical structures of representative macrocyclic hosts: cyclodextrins (CDs), calixarenes (CXs), and cucurbiturils (CBs); the specific chemical structures shown on top are those for  $\gamma$ -cyclodextrin, *p*-sulfonatocalix[4]arene, and cucurbit[7]uril.

The depicted CDs, CXs, and CBs have in common that they are highly water-soluble because they are laced by hydrophilic groups at their top and bottom rims but they possess a hydrophobic, cone- or barrel-shaped cavity predestined for the encapsulation of suitably sized hydrophobic guests. It is important to note that these types of macrocycles are available in different cavity dimensions, that is, different homologues with at least 4 and up to more than 10 monomeric repeat units can be synthesized. Note that macrocyclic receptors are superior to flexible acyclic ones, because their shape is pre-organized for the formation of a so-called “host-guest” complex, nicely fulfilling the requirement for Fischer's lock-key principle. Not surprisingly, macrocycles are considered as mimics of enzymatic binding pockets or biological receptors. This biomimetic receptor functionality, along with their biocompatibility,<sup>1-3</sup> make them suitable for a large range of biological and medicinal applications which range from drug delivery to drug discovery described herein.

## **DRUG DELIVERY**

Owing to the intrinsic water solubility of many macrocycles, their resulting host-guest complexes with different pharmaceutical drugs retain also a high water solubility. Moreover, the encapsulated drugs have different  $pK_a$  values than the uncomplexed ones.<sup>4-6</sup> In particular, the guest molecules can become protonated when immersed in CXs or CBs, which greatly assists their solubilization. This is of particular interest when the drug themselves have an insufficient water solubility. In these cases, macrocycles can be used to solubilize drugs,<sup>7,8</sup> e.g., when intravenous administration is desirable. To illustrate, the solubility of several benzimidazole-based anthelmintic and fungicides drugs increases by up to a factor of 100 upon encapsulation by CBs.<sup>9</sup> In contrast to liposomal or polymeric drug formulations, macrocyclic ones constitute discrete, supramolecular systems, whose pharmacokinetic properties can be readily predictively tuned, for example, to achieve a slow, sustained release. As an additional advantage, due to the spatial isolation of the encapsulated drug from the aqueous environment, enzymatic and other chemical degradation reactions can be effectively suppressed, in part assisted by their different protonation state.<sup>10</sup> This feature is already routinely exploited commercially in CD-based drug formulations, e.g., for oral testosterone delivery.



**Figure 2.** Representative drug molecules with potential in macrocyclic drug delivery, shown for CBs as an emerging class of synthetic macrocycles.

Drug delivery applications of CBs, purely synthetic macrocycles, are presently emerging. A selection of drug molecules which have already been investigated in the context of potential CB-based formulations are shown in Figure 2. Their potential can be gauged from the encapsulation of proton-pump inhibitor pro-drugs of the

omeprazole and lansoprazole type, whose encapsulation results in a 15-times faster activation and 500-times higher stabilization towards unwanted self-dimerization.<sup>6</sup> Related to drug delivery, the removal of transiently applied drugs such as neuromuscular blockers or of xenobiotics from the body is also of great biomedical interest. Sugammadex is a CD derivative already broadly applied for the removal of the general anesthetic agent rocuronium; recently, a CB-derived synthetic receptor has been introduced as an alternative agent for rocuronium complexation.<sup>11</sup> CB formulations are also being considered for the delivery of several platinum-based anticancer drugs (see Figure 2).<sup>10</sup> CXs, on the other, have been found to complex methylviologen-derived xenobiotics, effectively reducing their toxicity.<sup>12</sup>

Increasing attention is also being paid to *in vivo* microheterogeneous applications of macrocycles ranging from drug delivery<sup>13-15</sup> to DNA transfection.<sup>16</sup> For example, a supramolecular gel from a chitosan derivative and a CB has been shown to be capable of carrying and releasing 5-fluorouracil, an anti-cancer drug.<sup>17</sup> Vesicles assembled from amphiphilic CDs were shown to serve as carriers of Paclitxel (an anti-cancer drug) into carcinoma cell lines, HepG2 and BEL-7402, thereby enhancing curative effects.<sup>18</sup> CX-decorated nanoparticles have been shown to effectively transfer DNA into the cytoplasm of MDCK (Madin-Darby canine kidney) cells.<sup>19</sup> DNA transfection into different cells can alternatively be achieved by using arginine-modified CX derivatives.<sup>20</sup> In order to achieve a targeted drug delivery, macrocycles need to be derivatized with suitable cell-specific recognition elements. This can be achieved by tethering them with monoclonal antibodies or other biomarkers, and several examples of this strategy have been reported.

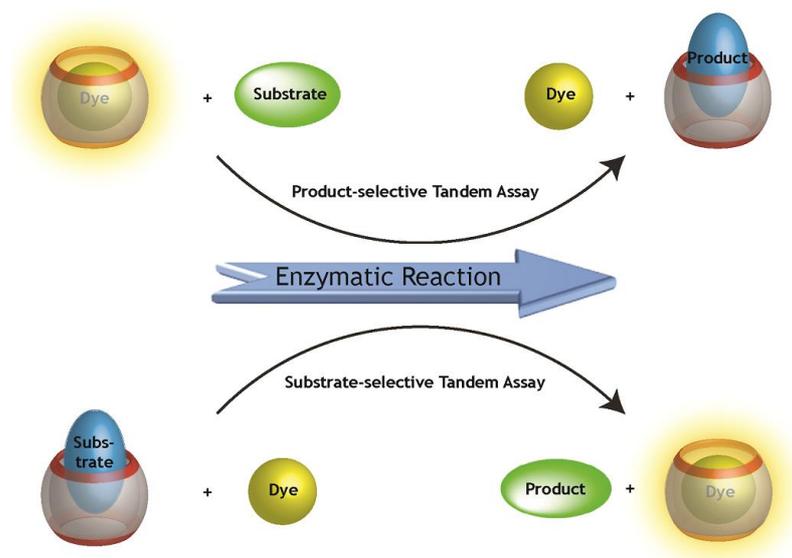
In regard to the release of the drugs from their macrocyclic complexes, some design aspects need to be taken into account. First, the complexation equilibrium between a drug and a macrocycle is rapidly reversible, which generally affects a much faster uptake and release than, for example, that observed for antibodies as biological receptors. The affinity constants of macrocycles lie in the millimolar to nanomolar range,<sup>10,21</sup> but attomolar affinity, exceeding the femtomolar affinity of the avidin-biotin pair, can be achieved by using CBs as synthetic receptors.<sup>22</sup> It is also interesting to note that the affinities depend quite strongly on pH and the presence of

salts, particularly cations.<sup>23</sup> This peculiarity provides, in fact, an interesting tool to affect an accelerated, potentially targeted, drug release.

## **DRUG DISCOVERY**

Since macrocycles have the potential to recognize drug molecules, or classes of drug molecules, other applications, apart from drug delivery, come to mind. For example, enzymes are among the top drug targets such that assays for inhibitor, less frequently activator, screening remain a prime challenge to the pharmaceutical industry. Macrocycles can be elegantly utilized in the drug discovery process as economic and versatile substitutes of antibodies in enzyme assays. The underlying idea is that molecular recognition by macrocycles is considerably less selective and by no means specific as that by antibodies, but can nevertheless become very useful as a complementary method in cases, where either the high cost and time demand of cloning antibodies is discouraging, or where the low molecular weight and ubiquitous presence of certain analytes (e.g., short-chain alkanes, simple amines, or metal ions) entirely prevents antibodies to be employed.

In order for macrocycles to be used for sensing applications, including enzyme assays, the molecular recognition event must be coupled to a signal transduction mechanism. The simplest way to achieve this involves indicator displacement. Thus, a fluorescent dye is used as guest molecule of a macrocycle to set up a so-called reporter pair or chemosensing ensemble. The addition or *in-situ* formation of a competitive guest, that is, any biological analyte which displays also binding to the macrocycle, results subsequently in the displacement of the fluorescent dye from the macrocycles, which is almost always accompanied by a readily detectable change in its fluorescence properties (intensity, spectrum, lifetime, polarization). Decarboxylases, for example, convert L-amino acids through CO<sub>2</sub> cleavage into the corresponding biogenic amines, such as histamine, simple natural metabolites for which no antibodies are available.



**Figure 3.** Working principle of supramolecular tandem enzyme assays, shown for a CB as macrocycle and a generic fluorescent dye. When an enzymatic reaction converting a substrate into a product proceeds in the presence of a macrocycle/fluorescent dye reporter pair, either the product serves as a stronger competitor and displaces the dye from the macrocycle as the enzymatic reaction occurs (product-coupled assay, top), or the substrate acts as stronger competitor, which is consumed by the enzyme, allowing the fluorescent dye to be taken up by the macrocycle (substrate-coupled assay, bottom). In both cases, a fluorescence response is obtained, either a switch-off fluorescence (top), or a switch-on response (bottom).

When macrocycles with cation-receptor properties, such as the CX or CB derivatives shown in Figure 1, are being employed, the enzymatic products of amino acid decarboxylases (doubly protonated and positively charged biogenic amines, at physiological pH) show a 2-3 orders of magnitude larger affinity to the macrocycles than the zwitterionic amino acid substrates. If such enzymatic reactions are carried out in the presence of a fluorescent guest, which is initially encapsulated inside the macrocycle, a continuous displacement of the indicator dye results, which can be gauged through a change in fluorescence. This provides the basis of a new class of enzyme assays, supramolecular tandem enzyme assays, which operate according to the principle shown in Figure 3. Tandem assays can be realized in many variants, including product- and substrate-coupled ones, and they have since then designed for many different enzymes and essentially all enzyme classes.<sup>24,25</sup> The advantages of

tandem enzyme assays, especially when compared with the complementary antibody-based assays are manifold: 1) They are highly economic. 2) The individual reporter pairs can be applied to a range of related enzymatic reactions and are not limited to a single enzyme. 3) The assays allow continuous, real-time monitoring and operate in homogeneous solution. 4) The assays exploit fluorescence for detection and can accordingly be readily adapted to high-throughput screening on a pharmaceutical-industrial scale.

To provide a specific example, acetylcholinesterase (AChE) activity has been traditionally monitored colorimetrically by Ellman's method. Inhibitors for AChE are important, among others, as Anti-Alzheimer drugs. This traditional method not only requires an artificial substrate (acetylthiocholine) but also a chemical follow-up reaction of the enzymatic product (thiocholine) with a reactive chromophore (5,5'-dithio-bis(2-nitrobenzoic acid)) to allow an indirect spectrophotometric detection.<sup>26</sup> The tandem enzyme assay which has been developed<sup>27</sup> for AChE by using CX as macrocycle and lucigenin as fluorescent dye not only allows the use of the unmodified substrate acetylcholine, but allows real-time monitoring by fluorescence instead of the less sensitive absorption measurements; this enables micromolar instead of millimolar concentrations to become readily detectable.

Most recently, tandem assays have also been modified for the monitoring of analyte uptake through channel proteins constituted in liposomal biomembrane models (tandem membrane assays),<sup>28</sup> which allows also this third-most important class of drug targets to be tackled. In the first example, models of antimicrobial peptides (protamine) have been tested with respect to their ability to translocate through the bacterial outer membrane protein F (OmpF).<sup>28</sup>

In summary, synthetic macrocyclic receptors are powerful tools in drug delivery and drug discovery. They provide viable alternatives to liposomal or polymeric formulations of hydrophobic drugs, and they can be implemented, in concert with fluorescent dyes, in a large variety of assays for monitoring enzymatic activity as well as membrane translocation in dependence on different inhibitors,

activators, and analytes/substrates. The resulting supramolecular tandem enzyme and membrane assays are complementary to classical ones involving specific antibodies.

## References

- (1) Rowe, R. C.; Sheskey, P. J.; Quinn, M. E. (2009) Handbook of Pharmaceutical Excipients; Pharmaceutical Press and American Pharmacists Association: London. Chicago.
- (2) Coleman, A. W.; Jebors, S.; Cecillon, S.; Perret, P.; Garin, D.; Marti-Battle, D.; Moulin, M. (2008) Toxicity and biodistribution of para-sulfonato-calix[4]arene in mice. *New. J. Chem.* 32, 780-782.
- (3) Uzunova, V. D.; Cullinane, C.; Brix, K.; Nau, W. M.; Day, A. I. (2010) Toxicity of cucurbit[7]uril and cucurbit[8]uril: an exploratory *in vitro* and *in vivo* study. *Org. Biomol. Chem.* 8, 2037-2042.
- (4) Mohanty, J.; Bhasikuttan, A. C.; Nau, W. M.; Pal, H. (2006) Host-guest complexation of neutral red with macrocyclic host molecules: Contrasting  $pK_a$  shifts and binding affinities for cucurbit[7]uril and beta-cyclodextrin. *J. Phys. Chem. B* 110, 5132-5138.
- (5) Shaikh, M.; Mohanty, J.; Singh, P. K.; Nau, W. M.; Pal, H. (2008) Complexation of acridine orange by cucurbit[7]uril and beta-cyclodextrin: photophysical effects and  $pK_a$  shifts. *Photochem. Photobiol. Sci.* 7, 408-414.
- (6) Saleh, N. I.; Koner, A. L.; Nau, W. M. (2008) Activation and stabilization of drugs by supramolecular  $pK_a$  shifts: Drug-delivery applications tailored for cucurbiturils. *Angew. Chem. Int. Ed.* 47, 5398-5401.
- (7) Brewster, M. E.; Loftsson, T. (2007) Cyclodextrins as pharmaceutical solubilizers. *Adv. Drug Deliv. Rev.* 59, 645-666.
- (8) Dong, N.; Xue, S.-F.; Zhu, Q.-J.; Tao, Z.; Zhao, Y.; Yang, L.-X. (2008) Cucurbit[n]urils (n = 7,8) binding of camptothecin and the effects on solubility and reactivity of the anticancer drug. *Supramol. Chem.* 20, 659-665.
- (9) Koner, A. L.; Ghosh, I.; Saleh, N. I.; Nau, W. M. (2011) Supramolecular encapsulation of benzimidazole-derived drugs by cucurbit[7]uril. *Can. J. Chem.* 89, 139-147.
- (10) Ghosh, I.; Nau, W. M. (2012) The strategic use of supramolecular  $pK_a$  shifts to enhance the bioavailability of drugs. *Adv. Drug Deliv. Rev.* 64, 764-783.
- (11) Ma, D.; Zhang, B.; Hoffmann, U.; Grosse Sundrup, M.; Eikermann, M.; Isaacs, L. (2012) Acyclic cucurbit[n]uril-type molecular containers bind neuromuscular blocking agents *in Vitro* and reverse neuromuscular block *in Vivo*. *Angew. Chem. Int. Ed.* 51, 11358-11362.
- (12) Wang, K.; Guo, D. S.; Zhang, H. Q.; Li, D.; Zheng, X. L.; Liu, Y. (2009) Highly effective binding of viologens by *p*-Sulfonatocalixarenes for the treatment of viologen poisoning. *J. Med. Chem.* 52, 6402-6412.
- (13) Duan, Q. P.; Cao, Y.; Li, Y.; Hu, X. Y.; Xiao, T. X.; Lin, C.; Pan, Y.; Wang, L. Y. (2013) pH-responsive supramolecular vesicles based on water-soluble pillar[6]arene and ferrocene derivative for drug delivery. *J. Am. Chem. Soc.* 135, 10542-10549.

- (14) Zhang, J.; Yuan, Z. F.; Wang, Y.; Chen, W. H.; Luo, G. F.; Cheng, S. X.; Zhuo, R. X.; Zhang, X. Z. (2013) Multifunctional envelope-type mesoporous silica nanoparticles for tumor-triggered targeting drug delivery. *J. Am. Chem. Soc.* 135, 5068-5073.
- (15) Yin, J. J.; Sharma, S.; Shumyak, S. P.; Wang, Z. X.; Zhou, Z. W.; Zhang, Y. D.; Guo, P. X.; Li, C. Z.; Kanwar, J. R.; Yang, T. X.; Mohapatra, S. S.; Liu, W. Q.; Duan, W.; Wang, J. C.; Li, Q.; Zhang, X. J.; Tan, J.; Jia, L.; Liang, J.; Wei, M. Q.; Li, X. T.; Zhou, S. F. (2013) Synthesis and biological evaluation of novel folic acid receptor-targeted, beta-cyclodextrin-based drug complexes for cancer treatment. *PLoS One* 8, 5, e62289.
- (16) Hu, Y.; Chai, M. Y.; Yang, W. T.; Xu, F. J. (2013) Supramolecular host-guest pseudocomb conjugates composed of multiple star polycations tied tunably with a linear polycation backbone for gene transfection. *Bioconj. Chem.* 24, 1049-1056.
- (17) Lin, Y. W.; Li, L. F.; Li, G. W. (2012) Preparation and properties of temperature and pH sensitive chitosan supramolecular gel based on host-guest interaction of cucurbit[8]uril. *Acta Chim. Sin.* 70, 2246-2250.
- (18) Sun, T.; Guo, Q.; Zhang, C.; Hao, J. C.; Xing, P. Y.; Su, J.; Li, S. Y.; Hao, A. Y.; Liu, G. C. (2012) Self-assembled vesicles prepared from amphiphilic cyclodextrins as drug carriers. *Langmuir* 28, 8625-8636.
- (19) Nault, L.; Cumbo, A.; Pretot, R. F.; Sciotti, M. A.; Shahgaldian, P. (2010) Cell transfection using layer-by-layer (LbL) coated calixarene-based solid lipid nanoparticles (SLNs). *Chem. Commun.* 46, 5581-5583.
- (20) Bagnacani, V.; Franceschi, V.; Bassi, M.; Lomazzi, M.; Donofrio, G.; Sansone, F.; Casnati, A.; Ungaro, R. (2013) Arginine clustering on calix[4]arene macrocycles for improved cell penetration and DNA delivery. *Nature Commun.* 4, 1721.
- (21) Dsouza, R. N.; Pischel, U.; Nau, W. M. (2011) Fluorescent dyes and their supramolecular host/guest complexes with macrocycles in aqueous solution. *Chem. Rev.* 111, 7941-7980.
- (22) Cao, L.; Šekutor, M.; Zavalij, P. Y.; Mlinarić-Majerski, K.; Glaser, R.; Isaacs, L. (2014) Cucurbit[7]uril-guest pair with an attomolar dissociation constant. *Angew. Chem. Int. Ed.* 53, 1521-3773.
- (23) Bakirci, H.; Koner, A. L.; Nau, W. M. (2005) Binding of inorganic cations by *p*-Sulfonatocalix[4]arene monitored through competitive fluorophore displacement in aqueous solution. *Chem. Commun.* 5411-5413.
- (24) Dsouza, R. N.; Hennig, A.; Nau, W. M. (2012) Supramolecular tandem enzyme assays. *Chem. Eur. J.* 18, 3444-3459.
- (25) Ghale, G.; Nau, W. M. (2014) Dynamically analyte-responsive macrocyclic host-fluorophore Systems. *Acc. Chem. Res.* in press.
- (26) Norouzy, A.; Qujeq, D.; Habibi-Rezaei, M. (2009) The inhibitory effect of dissolved carbaryl in dioxane on physically adsorbed acetylcholinesterase. *React. Kin. Cat. Lett.* 98, 391-401.
- (27) Guo, D. S.; Uzunova, V. D.; Su, X.; Liu, Y.; Nau, W. M. (2011) Operational calixarene-based fluorescent sensing systems for choline and acetylcholine and their application to enzymatic reactions. *Chem. Sci.* 2, 1722-1734.
- (28) Ghale, G.; Lanctôt, A. G.; Kreissl, H. T.; Weingart, H.; Winterhalter, M.; Nau, W. M. (2014) Chemosensing ensembles for monitoring biomembrane transport in real time. *Angew. Chem. Int. Ed.* 53, 2762-2765.