

**UNIVERSITÀ DEGLI STUDI DI TRIESTE**

---

XXII CICLO DEL  
DOTTORATO DI RICERCA IN  
SCIENZE E TECNOLOGIE CHIMICHE E FARMACEUTICHE

**LIBRARIES OF SHORT RECEPTOR PEPTIDES AND  
ANTIBODIES FOR THE RECOGNITION OF  
SMALL ORGANIC MOLECULES**

CHIM06

Dottorando:  
Dott. Giampaolo Fontanive

Responsabile Dottorato di Ricerca  
Prof. Enzo Alessio

Relatore  
Dott. Federico Berti

## Abstract

La presente tesi di Dottorato riporta i risultati di ricerche svolte nell'ambito di alcune collaborazioni tra Università ed industria nel campo della progettazione e sviluppo di librerie di trasduttori primari per biosensori. In particolare l'attenzione si è rivolta da un lato a piccoli peptidi come recettori di xanthine (caffaina e teofillina), dall'altro ad una molecola di grandi dimensioni, la palitossina, verso la quale al momento sono stati preparati anticorpi con tecniche più convenzionali. Il contributo di un chimico in questo tipo di progetti è analogo, e consiste in un'attività di sintesi di derivati modificati delle molecole di interesse, nella loro bioconiugazione a proteine ed enzimi, nella sintesi di peptidi, nello studio dell'interazione ligando-recettore attraverso misure di attività, analisi conformazionale e strutturale, modellizzazione. L'attività svolta è brevemente descritta nei punti seguenti.

1) I coiled-coil sono motivi strutturali presenti in natura dati dall'unione di due o più  $\alpha$ -eliche avvolte su loro stesse al fine di generare un'unica super-elica. L'analisi di tali strutture rivela spesso la presenza di sequenze ripetute di 7 amminoacidi ripetuti (*g-a-b-c-d-e-f*) con precise caratteristiche: la posizioni *a* e *d* per esempio sono occupate sempre da amminoacidi con caratteristiche idrofobiche, atti a generare una interfaccia idrofobica interna alla super-elica molto stabile e non esposta al solvente esterno. Parallelamente, le posizioni *e* e *g* sono invece occupate da amminoacidi polari carichi che si trovano sulla superficie esterna della super-elica e possono interagire tra di loro rinforzando l'interazione e provocando effetti di orientamento.

La nostra attenzione si è focalizzata sul coiled-coil denominato EK, sviluppato agli inizi degli anni '90 da Hodges e colleghi. In questa struttura due catene peptidiche denominate appunto E e K si avvolgono su loro stesse al fine di generare una super-elica anfipatica di notevole stabilità. Partendo da tale struttura e sfruttando la tecnica biomolecolare del phage display che permette l'esposizione di strutture a carattere proteico sulla superficie di un batteriofago, sono state create librerie di EK mutati con l'obiettivo di generare all'interno della super-elica siti recettoriali eventualmente utilizzabili all'interno di biosensori come trasduttori primari per caffeina e teofillina, due piccole molecole organiche altamente correlate e ampiamente presenti nella vita quotidiana. È stata prodotta una prima libreria fagica di dimensioni notevoli ( $10^8$  elementi), all'interno della quale sono stati isolati due cloni specifici per le xanthine; i due cloni sono stati denominati E1 ed E8. La caratterizzazione del binding è avvenuta sia con tecniche biomolecolari che prevedono l'immobilizzazione dell'analita su superficie solida (test ELISA e dot-blot), sia con tecniche chimiche con analiti liberi e privi di modificazioni strutturali.

Al fine di ridurre il più possibile le dimensioni dello scaffold di partenza è stata sviluppata una seconda libreria fagica in cui è stata randomizzata soltanto la catena K dell'originario EK proposto da Hodges. La libreria finale presenta una dimensione di  $10^8$  elementi, tra i quali è stato isolato e caratterizzato attraverso test ELISA il clone denominato K<sup>E1</sup>. Le ridotte dimensioni di

tale peptide (35 amminoacidi) hanno permesso la sua sintesi chimica su una buona scala (40 mg), una caratterizzazione strutturale attraverso spettroscopia di dicroismo circolare, un'analisi dell'interazione in soluzione con la teofillina sfruttando sempre la tecnica CD e quella calorimetrica ed infine un'analisi cristallografica attraverso luce di sincrotrone.

2) Una seconda parte del lavoro è stata finalizzata, invece, allo sviluppo di pool anticorpali mono- e policlonali quali trasduttori primari in biosensori per l'individuazione della palitossina.

Tale tossina algale con elevato indice di pericolosità ( $LD_{50}=0.45 \mu\text{g}/\text{kg}$ ) negli ultimi anni ha causato notevoli disagi in campo ambientale, inquinando flora e fauna acquatica di diverse zone marittime, compreso il Mediterraneo e l'alto Adriatico.

Per lo sviluppo di anticorpi specifici contro tale tossina, sono state condotte immunizzazioni su coniglio e su topo con palitossina coniugata a due tipi di proteine carrier (BSA e KLH) al fine di ridurre la sua tossicità e indurre una risposta immunitaria nell'animale. Il pool anticorpi policlonali di siero di coniglio è stato isolato e caratterizzato attraverso sandwich ELISA contro palitossina libera e palitossina chimicamente modificata allo scopo di ottimizzare tale tipo di test e utilizzarlo come possibile dosatore nei confronti dell'analita presente in campioni di acqua oppure in sistemi più complessi quali mitili o ulteriore fauna marina.

## Abstract

This PhD thesis reports the results of research activities carried out at the Dept. of Chemistry of the University of Trieste, in collaboration with several industries, in the field of design and development of libraries of primary transducers for biosensors. We have focused our attention on one side on short peptides as xanthine (caffeine and theophylline) receptors, and on the other side on a very large molecule, palytoxin, which has been considered to raise antibodies by more conventional techniques. The chemist's contribution to such projects is similar, consisting in the synthesis of modified target molecules, in their bioconjugation to protein and enzymes, in peptide synthesis, in studying the receptors-ligand interactions by affinity measurements, conformational and structural analysis, modelling. The activity is briefly described in the following points.

1) Coiled-coil are natural motifs made by bundles of two or more  $\alpha$ -helix that generate a super-helix. Structural analysis of these motifs reveals often the presence of a repeated sequence 7 aminoacids along the helix (*g-a-b-c-d-e-f*). *a* and *d* are hydrophobic aminoacids able to generate an internal hydrophobic core most important for the overall stabilization of the bundle, *e* and *g* are polar or charged aminoacids that can contribute to the final stability of the coil via electrostatic interchains interactions.

Starting from the EK coiled-coil developed by Hodges et al., made by two different peptide chains (E:EVSALEK; K:KVSALKE) and exploiting phage display techniques, two different phage libraries of E and K mutants were generated, in order to obtain receptors molecules suitable as primary transducer in biosensors for caffeine and theophylline.

In the first library, both E and K were randomized, obtaining  $10^8$  different peptide clones; two specific clones were isolated against immobilized caffeine and theophylline, their characterization was carried out both by biomolecular experiments (ELISA test; dot-blot) and by chemical techniques.

In a second library, we have further reduced the scaffold size, using only a single K chain. A new library of  $10^8$  elements was obtained and characterized. The isolation of K<sup>E1</sup> clone as the best xanthine binder, allowed chemical synthesis on a good scale (40 mg); structural characterization and binding analysis of its interaction with theophylline in water was performed by circular dichroism and calorimetry; the X-ray structure of the peptide was obtained exploiting synchrotron light.

2) In a second part of this work, the development of mono- and polyclonal antibodies towards palytoxin was stated. Palytoxin is one of the most dangerous algal toxins in nature ( $LD_{50}=0.4 \mu\text{g}/\text{kg}$  in mice) and its presence in the last years in Mediterranean sea, both in the water sea and in fish or mussels, constitutes a real problem for tourism economy and for the fish market.

In order to obtain specific antibodies as bioreceptors towards palytoxin, a series of immunizations on rabbits were carried out with palytoxin-BSA and palytoxin.KLH conjugates in order to generate immune response in the animals; the pool of high reactivity bleeds of rabbits, were isolated and used to set up ELISA sandwich system towards palytoxin.

# INDEX

## INTRODUCTION

<b>1. Peptide scaffolds and libraries</b>	1
1.1 The interaction between biomolecules and small molecules	1
1.1.1 Scaffolds with $\alpha$ -helices domains	4
1.1.2 Scaffolds with $\beta$ domains	6
1.1.3 Scaffolds with irregular domains	9
1.2 Coiled-coil domains	9
1.2.1 Structure of coiled-coils	10
1.2.2 The hydrophobic core	11
1.3 Biological combinatorial libraries of peptides	12
1.3.1 Phage-display	13
1.3.2 Phage-display libraries	14
1.3.3 Bacterial display	16
1.3.4 Ribosome display	17
<b>2. Biosensors and microarrays</b>	19
2.1 Biosensors	19
2.1.1 The biorecognition element	20
2.1.2 The secondary transducer	22
2.1.3 Immobilisation of biorecognition element	23
2.2 Microarrays	24
2.2.1 Ligand detection	27
2.2.2 Peptide array	29
<b>3. Target molecules</b>	31
3.1 Palytoxin	31
3.2 Theophylline and caffeine	33

<b>AIM OF RESEARCH</b>	36
------------------------	----

<b>RESULTS AND DISCUSSION</b>	37
-------------------------------	----

<b>4. Modified target molecules</b>	38
4.1 Caffeine and theophylline	38
4.1.1 The synthesis of caffeine/theophylline-biotin conjugates	41
4.1.2 The synthesis of caffeine and theophylline amino derivatives	42
4.1.3 Synthesis of polar-linker xanthine	44

4.1.4 Xanthine-protein conjugates	45
4.2 A menthone-biotin conjugate	48
4.3 Palytoxin	50
4.3.1 Retrosynthetic analysis for the C <sub>101</sub> -C <sub>115</sub> synthesis	51
4.3.2 Bioconjugation of palytoxin	54
4.3.3 Biotinylation of palytoxin	55
<b>5. The E/K coiled-coil peptide</b>	57
5.1 The E/K structure	57
5.1.1 Physical characterization of the E/K coiled-coil	59
5.2 The C-EK <sub>inv.</sub> -C scaffold, a first library project	60
5.2.1 Preliminary study on the EK <sub>inv.</sub> folding	62
5.2.2 Binding site design	68
<b>6. C-EK<sub>inv.</sub>-C libraries</b>	69
6.1 Phage-library construction	69
6.2 Selection of xanthine-binding clones	71
6.2.1 The ELISA tests	73
6.2.2 Measurement of the dissociation constant of the peptide/caffeine complexes	77
6.2.3 Characterization of bE1 and bE8 sequence	78
6.3 Expression of E1 and E8 in <i>E. Coli</i>	81
6.3.1 Dot-blot	82
6.3.2 E1 and E8 as affinity chromatography stationary phases	83
<b>7. K library, towards a simplified system</b>	86
7.1 K library construction	87
7.2 Synthesis and chemical characterization of K <sup>E1</sup>	90
7.2.1 Solid phase synthesis of K <sup>E1</sup> peptide	91
7.2.2 The solution behaviour of K <sup>E1</sup>	93
7.2.3 The crystal structure of K <sup>E1</sup>	95
7.3 The interaction between K <sup>E1</sup> and theophylline	98
7.3.1 Binding of theophylline via circular dichroism	98
7.3.2 Binding of theophylline via calorimetry	100
7.4 Towards an electrochemical-sensor: F16H-K <sup>E1</sup>	103
<b>8. Antibodies library</b>	106
8.1 Immunization strategy	106
8.1.1 Production of polyclonal antibodies from rabbits	108
8.1.2 The “sandwich” ELISA	108



# INTRODUCTION

# Chapter 1: Peptide scaffolds and libraries

## 1.1 The interactions between biomolecules and small molecules

The regulation of numerous processes in living systems is often based on the interaction between a large macromolecule as a protein, enzyme or antibody, and a small organic molecule that can interact with the host at a binding site. Inside the binding site, amino acids with hydrophobic or hydrophilic character are able to generate different structural motives in which small organic molecule are complexed by covalent or non-covalent binding.

During the recent years several proteins have been studied with the aim of identifying peculiar peptides sequences that could be able to bind small molecules; the idea was to use such peptides as bioreceptors in a biosensor or sensors array.

A first example was given by olfactory receptors.<sup>1</sup> The mammalian olfactory system has been recognized as one of the most effective sensing system with extraordinary capability. Humans and other mammals can detect some odourants at low ppb levels and discriminate 10000 distinct odours.<sup>2</sup> The initial event in odour discrimination is represented by the odourant over binding inside the receptor molecules, which reside on the cilia of sensory neurons in the olfactory epithelium.<sup>3</sup> It would be desirable to use directly these biomolecules as a bioelement in a sensor with a suitable physical transducer, for the detection of organic molecules present in different samples; several attempts have indicated that such a sensor could effectively discriminate between different odourants molecule; however, the system suffers from many limitations, due to scarce reproducibility and standardization in probe preparation because of the variation in animal species, bacteria degradation and orientation problems in protein immobilization. In order to avoid these problems a series of oligopeptides have been synthesized having their sequences corresponding to a selective portion of a putative binding site between olfactory receptors and small target molecules;<sup>4</sup> these peptides were designed, synthesized and coated onto the electrode surface for target odourant binding assays.<sup>5</sup> By this way a proof of concept was given by Wu, with a sensor that was able to identify trimethylamine with a certain selectivity.<sup>6</sup>

Proteome offers many small-protein motifs organized in a stable conformation: one example is represented by conotoxins, 15-20 aminoacid peptides organized in a small  $\alpha$ -domain

---

<sup>1</sup> Hines E.L., Gardner J.W., Potter C.E.R. *Measure Cont.*, **1997**, 262-268.

<sup>2</sup> Persaud K.C., Khaffad S.M., Payne J.S., Pisaneeli A.M., Lee D.H., Byun H.G. *Sensor Actuators B*, **1996**, 36, 267-273.

<sup>3</sup> a) Sheperd G.M. *Neurobiology*, **1988**, Oxford University Press Inc., 205-221. b) Raming K., Krieger L., Strotmann J., Bockhoff L., Kubick S., Brer H. *Nature*, **1993**, 361, 353. c) Chen Z., Lancet D. *Proc. Natl. Acad. Sci. USA*, **1984**, 81, 1859-1863. d) Chen Z., Ophir D., Lancet D. *Brain Research*, **1986**, 368, 328-338. Linda B., Axel R. *Cell*, **1991**, 65, 345-367.

<sup>4</sup> Ephraim K.K., Shariv I., Eisenstein M., Friesen A.A., Vakser I.A. *Proc. Acad. Natl. Sci. USA*, **1992**, 89, 2195-2199.

<sup>5</sup> Wu T.Z., Lo R.Y., Chan E.C. *Biosensor Bioelectron.*, **2001**, 16, 945-953.

<sup>6</sup> Wu T.Z. *Biosensor Bioelectron.*, **2001**, 16, 945-953.

and joined together by one or two disulfide bonds; this constitute the bottom of a small interacting site which involves also a loop; more than 50000 different types of conotoxins are known and they could be considered as candidates for the development of large libraries of small receptors by molecular biology techniques (Figure 1.1).<sup>7</sup>



Figure 1.1

Cyclic peptides are another example of potential small receptors. To develop a library of cyclic peptides is very difficult due to the cyclization step. However, *E. Coli* produces a cyclotide, the microcine J25 (MccJ25), with an extraordinary lassoed tail structure (Figure 1.2).

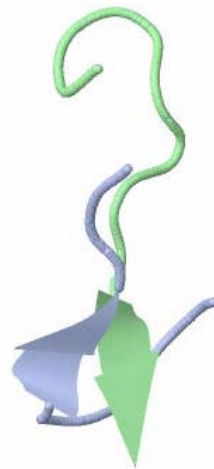


Figure 1.2

These two examples introduce us to a more general insight on the structural features that are required in order to identify and design a small peptide receptor. A conformationally stable peptide is required (scaffold), in order to allow the insertion of a new flexible and randomizable binding sequence.

---

<sup>7</sup> Terlau H., Olivera B.M. *Physiol. Rev.*, **2004**, *84*, 41-68.

As to the scaffold, the following definition has been recently given: “a protein framework that can carry altered amino acids or sequence insertion, that confer on protein variants different functions, usually for binding specific targets”.<sup>8</sup>

Peptide scaffold can be derived from structural proteins, sub-unit peptides from a multimeric structure and also enzymes,<sup>9</sup> but it is possible to design *de novo* structures.<sup>10</sup> They can be selected by *in vivo* and *in vitro* techniques, isolating new molecules with better affinity against specific ligand.<sup>11</sup>

As to the binding area, loops are the most accessible regions for the recognition of the small target molecules, they can link together structural elements like  $\alpha$ -helices and  $\beta$ -sheets of the scaffold. Also  $\alpha$ -helices<sup>12</sup> and internal cavities of protein<sup>13</sup> can be involved in target recognition. However loops are the most accessible region for modification of the sequence, thus integrating an affinity function onto or into the structural framework without changing the overall structure and generating a stable folded peptide. It is important to choose carefully the residues to be randomized because some modifications may cause misfolding or instability of the scaffold itself. Several different type of scaffolds, and their possible modifications, are shown in Figure 1.3:

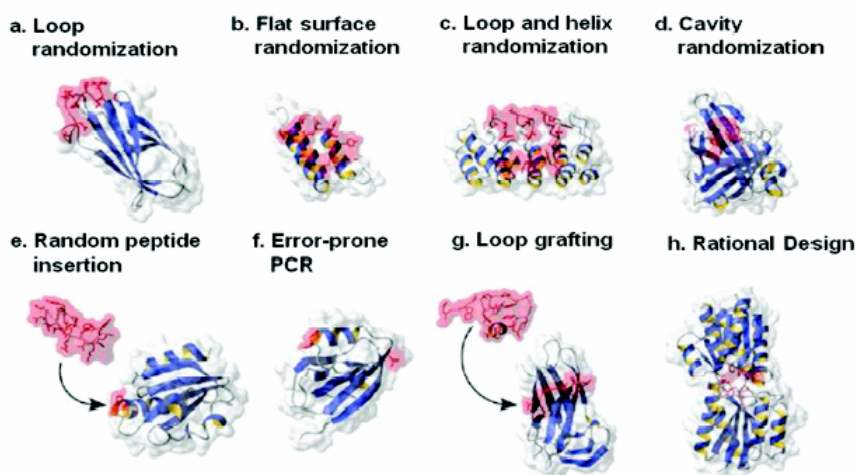


Figure 1.3

Based on the architecture of their scaffold peptide receptors can be classified as

1. based on  $\alpha$ -helices structures
2. based on  $\beta$ -sheets structures
3. small scaffolds with an irregular architecture of  $\alpha$ -helices or  $\beta$ -sheets.

<sup>8</sup> Biniz H.K., Berell G., Hytuani G.F., Santos J.A. *Nat. Biotechnol.*, **2005**, 23(10), 1257-1268.

<sup>9</sup> Vita C. *Curr. Opin. Biotechnol.*, **1997**, 8(4), 429-434.

<sup>10</sup> a) Custon A.S., Sherman J.C. *Bioorg. Med. Chem.*, **1999**, 7(1), 23-27. b) O'Shea E.K. *Curr. Biol.*, **1993**, 3(10), 658-667.

<sup>11</sup> Biniz H.K., Pluckthun A. *Curr. Opin. Biotechnol.*, **2005**, 16(4), 459-469.

<sup>12</sup> Enander K., Sullivan D., Kiyonari N.H. *J. Am. Chem. Soc.*, **2004**, 126(14), 4464-4465.

<sup>13</sup> Lee M. *J. Org. Chem.*, **2006**, 71(14), 5082-5092.

### 1.1.1 Scaffolds with $\alpha$ -helices domains

Although  $\alpha$ -helix is the most important motif present in nature protein, the number of scaffolds based on this backbone is limited respect to the scaffolds based on  $\beta$ -sheet motif.

1. Affibodies are small proteins derived from an engineered version (Z domain)<sup>14</sup> of the five stable three  $\alpha$ -helices bundle domains of the immunoglobulin Fc-binding region of staphylococcal protein A (Figure 1.4).<sup>15</sup>



Figure 1.4

These scaffolds are soluble, proteolytically and thermally stable and they don't contain disulfide bonds. They have been used for construction of phage libraries and selection against several proteins and enzymes, like *Tag* DNA polymerase, human insulin and human apolipoprotein, while there are not reputed examples of an affibodie directed towards small molecules

The small dimension and the strong protein structure are fundamental properties of these molecules for a broad range of biotechnology applications: affibodies have been described for purification, detection and targeting, also suggesting the therapeutic diagnostic potential of this protein scaffold.<sup>16</sup>

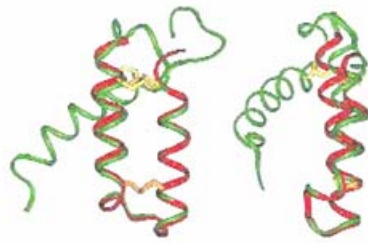
2.  $\alpha$ 2p8 peptide is a polypeptide of 38 amino acids; it is assembled in two  $\alpha$ -helices structure bound together with a disulfide bond in order to obtain a final hairpin structure (Figura 1.5).

---

<sup>14</sup> Nord K. *Protein Eng.*, **1995**, *8(6)*, 601-608.

<sup>15</sup> Nilsson B., Moks T., Jansson B., Abrahmsen L., Elmblad A., Uhlen M. *Protein Eng.*, **1987**, *1(2)*, 107-113.

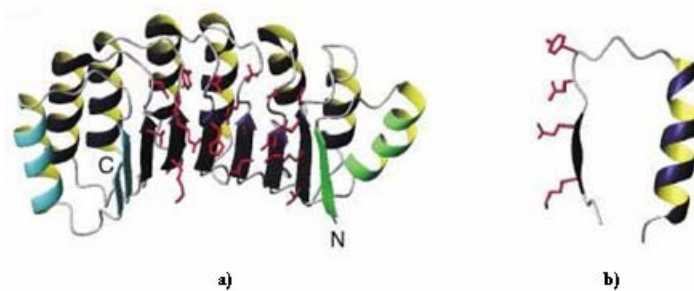
<sup>16</sup> Renberg B., Shiroyama I., Engfeldt T., Nygren P., Karlstrom A.E. *Anal. Biochem.*, **2005**, *341(2)*, 334-343.



**Figure 1.5:** NMR structure of  $\alpha 2p8$  peptide (red) overlying at the  $p8^{MITCP1}$  structure (green).

$\alpha 2p8$  peptide is present in the N-terminal region of human protein  $p8^{MITCP1}$  that also present a hairpin motif in its tridimensional structure. The  $\alpha 2p8$  peptide has not been currently used for the development of combinatorial libraries or for a phage-display selection against any targets molecules.

3. Repeat proteins or proteins with ankyrin repeat domains. The ankyrin repeat is a 33-residue motif in proteins, consisting of two  $\alpha$ -helices separated by loops; they are present in bacterial, archaeal, and eukaryotic proteins (Figure 1.6).<sup>17</sup> Designed ankyrin repeat proteins (DARPin)s are versatile binding proteins which display great advantages over other binding protein. These repeat-motifs protein offer an advantage for libraries construction because the surface of binding interface depends on the number of randomized repeats. In a recent example, libraries of ankyrin repeat proteins have been used to select high affinity inhibitors of aminoglycoside phosphotransferase (3')-IIIa (APH).<sup>18</sup> The selected proteins showed the great DARPin's ability for modulation of intracellular enzyme activity, with implication for therapeutic applications. However also this example refers to the recognition of a large protein domain rather to the recognition of a small molecule.



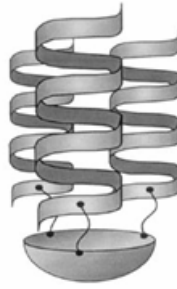
**Figure 1.6:** a) Molecular model of DARPin with red randomized residues. b) Molecular model of a single domain.

4. TASPs (Template-Assembled Synthetic Proteins) are very small structures stable to high temperature, strong acidic conditions and at denaturing agents (Figure 1.7)

Three 12-15 aminoacids  $\alpha$ -helices, are bound on a support molecule (ciclotribenzylene) in order to obtain a particular cylinder scaffold.

<sup>17</sup> Binz H.K., Stumpp M.T., Forrer P., Amstutz P., Pluckthun A. *J. Mol. Biol.*, **2003**, 332(2), 489-503.

<sup>18</sup> Kohl A., Amstutz P., Parizek P., Binz H.K., Briand C., Capitani G., Forrer P., Grutter M.G. *Structure*, **2005**, 13(8), 1131-1141.



**Figure 1.7:** schematic representation of TASPs in which three  $\alpha$ -helix peptides are bound at the beneath on a support molecule.

### 1.1.2 Scaffolds with $\beta$ domains

1. Antibodies and related fragments represent the most important natural binding protein formed by  $\beta$ -sheet motifs.

Antibodies, also known as immunoglobulins (Ig), are globular proteins that are found in blood or other bodily fluids of vertebrates and are used by the immune system to identify and neutralize foreign objects such as bacteria and viruses. They are typically made of basic structural units each with two long heavy chains and two shorter light chains assembled to form different structures as monomers with one unit, dimers with two units or pentamers with five units.

Though the general structure of all antibodies is very similar, a small region at the top of the protein is extremely variable, allowing millions of antibodies with slightly different top structures, or antigen binding sites, to exist. This region is known as the hypervariable region. Each of these variants can bind to a different target, known as antigen. The huge diversity of antibodies allows the immune system to recognize an equally wide diversity of antigens; the unique part of the antigen recognized by an antibody is called epitope. Epitopes bind to their antibody in a highly specific interaction, that allows antibodies to identify and bind only their unique antigen in the midst of the millions different molecules in the organism, with affinities in the sub picoM range.

The antibody dimer is a “Y”-shaped molecule that consists of four polypeptide chains; two identical heavy chain and two identical light chains connected by disulfide bonds. Each chain is composed of structural domains called Ig domains. These domains contain about 70-110 amino acids and are classified into different categories, according to their size and function. They have a characteristic immunoglobulin fold in which two  $\beta$ -sheets create a “sandwich” shape, held together by interactions between conserved cysteines and other charged amino acids (Figure 1.8).



Figure 1.8

The specificity is mediated by the complementary determining regions (CDR), hypervariable loop regions that mediate binding to the antigen. Immunoglobulins can bind proteins, sugars and diverse small proteins, but there are some limitations for biotechnology applications, for example due to the dimensions of antibodies. These are motivations, in some cases, to find alternative binding molecule.

2. Knottins: knottins are small disulfide-rich proteins characterized by a special “disulfide through disulfide knot”. The knottin motif is found in some plant-protease inhibitors, toxins and antimicrobial peptides.

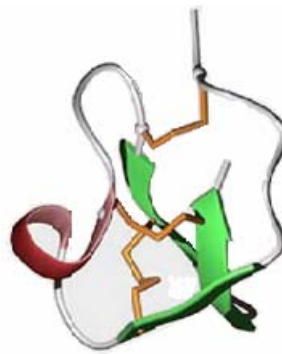


Figure 1.8: NMR structure of a typical knottin motif with particular disulfide bridges.

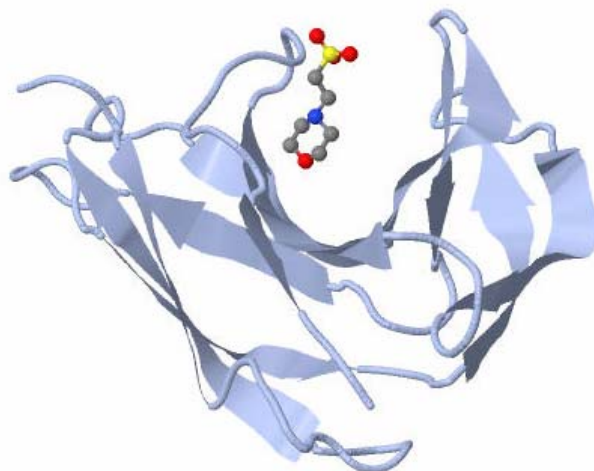
A typical “knotted topology” unit consist in a small triple-stranded  $\beta$ -sheet with at least 3 disulfide bridges and interspersed variable peptide loops.

Residue in the loops can be mutated and with affinity selection methods a better binding activity can be obtained.<sup>19</sup> The knot proteins are an attractive scaffold for drug design and they have been considered a good starting point for grafting recognition sites onto it for the production of new bioactive molecules.<sup>20</sup>

<sup>19</sup> Kolmar H. *Curr. Opin. Pharm.*, **2009**, *9*(5), 608-614.

<sup>20</sup> Souriau C. *Biochemistry*, **2005**, *44*(19), 7143-7155.

3. Statine is a protein made by seven  $\beta$ -sheets strands in order to form a  $\beta$ -sheet sandwich, in a structure similar to that of immunoglobulins. Generally in a statin protein two loops make a deep pocket where is possible to obtain binding towards organic molecules as anticancer drugs.



**Figure 1.9:** neocarzinostatin binds small organic molecule.

Heyd et al. have shown that the binding region of the neocarzinostatin can be engineered in order to obtain a strong binding with organic molecules (Figure 1.9).<sup>21</sup> In another example, by randomizing thirteen amino acids in the pocket of statin, Heyd et al. have obtained three different peptide libraries with one component able to bind testosterone.

4. Lipocalins are a family of proteins that are able to transport small hydrophobic molecules such as steroids, lipids and retinoids in the human organism. They are 160 to 180-residues polypeptides that share a common tertiary architecture, consisting in a  $\beta$ -barrel of six or eight anti-parallel  $\beta$ -strands, connected by four hypervariable loops (Figure 1.10).



**Figure 1.10**

The conical structure encloses an internal ligand binding site which is ideal to randomized residues, allowing generation of protein binders with high affinity and specificity. Some new

---

<sup>21</sup> Heyd B., Castel G., Tordo N. *Virology*, **2009**, 13(8), 93-102.

binders have been isolated from a bilin-binding protein library and selected against fluorescein and digoxigenin with nanomolar  $K_d$ .<sup>22</sup>

### 1.1.3 Scaffolds with irregular domains

1. Defensins are small cysteine-rich cationic protein made by 15-20 amino acids residues. They are present in both vertebrates and invertebrates; they play very important roles against invading bacteria, fungi and viruses. The group comprises peptides with an irregular architecture of  $\alpha$ -helices and  $\beta$ -sheets, or less structured sequences. Insect defensin A has been proposed as a new scaffold for the construction of a mutant library: seven residues were randomized within the two loops of the peptide.<sup>23</sup> The library was screened against tumor necrosis factor A, TNF receptor 1 and TNF receptor 2 (Figure 1.11).



Figure 1.11

2. Kunitz domain inhibitors are a group of short and irregular peptides of about 60 amino acids, stabilized by three disulfide bonds. They act as reversible inhibitors of serine proteases. They possess high stability and, most important for library generation, a loop that can be mutated without destabilizing the structural folding. Therefore, they have been used for modelling specificity to novel protease targets and entered clinical trials.<sup>24</sup>

## 1.2 Coiled-coil domains

The  $\alpha$ -helix is involved in the folding of another important protein domain: the coiled-coil.<sup>25</sup> The coiled-coil motif is found in approximately 10% of all protein sequences and is responsible for

---

<sup>22</sup> Schlehuber S, Skerra A. *Expert Opin. Biol. Ther.*, **2005**, 5(11), 1453-1462.

<sup>23</sup> Zhao W.-M., Chen Y., Thian Y.-C. *Peptides*, **2004**, 25(4), 629-635.

<sup>24</sup> Williams A., Baird L.G. *Transfus. Apher. Sci.*, **2003**, 29(3), 255-258.

<sup>25</sup> Lupas A.N., Gruber M. *Adv. Proteins Chem.*, **2005**, 70, 37-78.

the oligomerization of proteins in a highly specific manner. Coiled-coil proteins exhibit a large diversity of function (gene regulation, cell division, membrane fusion, drug extrusion) thus demonstrating the significance of oligomerization in biological systems.<sup>26</sup>

Its existence was first predicted in a series of seminal papers by Crick<sup>27</sup> and supported by an observation by Perutz and others that the long fibrous protein tropomyosin and other members of the *k-m-e-f* proteins exhibited diffraction peaks with a repeated spacing of 5.15 Å. Crick proposed that this could be explained by an arrangement in which two  $\alpha$ -helices wrapped around each other at an angle of approximately 20° such that their side chains were interlocked together in a motif that repeated itself every 7 residues, or 2 turns of the helix.

Crick referred to this packing as a “knobs into holes” arrangement in which one residue (knob) fits into a space generated by four residues of the opposing helix (hole). The simplest coiled-coil structure is the dimer, but trimers, tetramers, pentamers<sup>28</sup> and heptamers<sup>29</sup> have been reported. In addition, if the helices come from the same protein chain, the resultant coiled-coil is described as a homooligomer. Conversely, if the helices originate from different protein chains, the coiled coil is referred to as a hetero-oligomer.

### 1.2.1 Structure of coiled-coils

The primary sequence of a coiled-coil embodies a combination of simplicity and repetitiveness. A typical coiled-coil sequence is made of seven-residue repeats called heptads. The number of heptads in coiled-coils can vary from up to 200 heptads in fibrous proteins,<sup>27</sup> down to only two heptads in a *de novo* designed synthetic coiled-coil.<sup>30</sup> The seven positions in a heptad are conventionally denoted as *a-b-c-d-e-f-g*. Positions *a* and *d* are always occupied by hydrophobic residues (the most common ones are Leu, Ile and Val). Position *e* and *g* are often occupied by charged residues (the most common are Lys and Glu). Such a distribution of amino acid residues makes the helices amphipathic with residues at *a* and *d* forming the inter-helical hydrophobic core and residues at *e* and *g* forming inter-helical ionic interactions. The arrangements of residues in coiled-coil are most conveniently illustrated by the helical-wheel diagram (Figure 1.12)

---

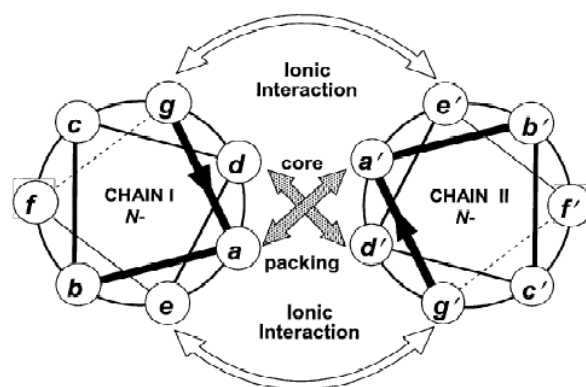
<sup>26</sup> a) Cohen C, Parry A.D.D. *Proteins*, **1990**, *7*, 1-15. b) Kohn W.D., Mant C.T., Hodges R.S. *J. Biol. Chem.*, **1997**, *272*, 2583-2586.

<sup>27</sup> a) Crick F.H.C. *Nature*, **1952**, *170*, 882-883. b) Crick F.H.C. *Acta Crystallogr.*, **1953**, *6*, 685-689.

<sup>28</sup> Efimov V.P., Engel J., Malashkevich V.N. *Proteins*, **1996**, *24*, 259-262.

<sup>29</sup> Liu E., Mui S., Brown J.H., Strand J., Reshetnikova L., Tobacman L.S., Cohen C. *Proc. Natl. Acad. Sci. USA*, **2002**, *99*, 7378-7383.

<sup>30</sup> Burkhard P., Meier M., Lustig A., *Protein Sci.*, **2000**, *9*, 2294-2301.



**Figure 1.12** : Helical-wheel diagram of a dimeric coiled-coil showing the helical cross-section. The seven positions are labeled *a-b-c-d-e-f-g*; generally *a* and *d* positions are occupied by hydrophobic side-chains, *e* and *g* positions are mostly occupied by charged or polar side-chains. Amino acids in *a* and *d* positions form the hydrophobic core while side-chains at the *e* and *g* positions form inter-helical ionic interactions.

Furthermore, the *e* and *g* edge position can shield the hydrophobic core from aqueous surroundings. The characteristic hydrophobic repeat is strongly maintained, it is a fundamental feature for structural reasons, for folding and stability of protein.<sup>31</sup>

### 1.2.2 The hydrophobic core

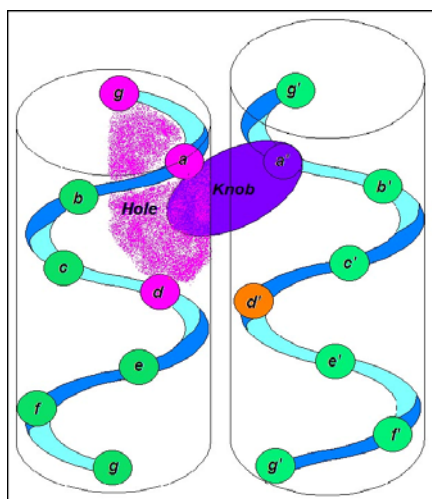
The analysis of several natural and designed coiled-coils showed the relevance of apolar amino acids in the core: in *de novo* designed peptides, leucine or isoleucine in *a* and *d* position showed better stability;<sup>32</sup> *d* position is typically leucine or alanine, contributing to dimerization stability but not to specificity. In contrast the *a* position is more variable and, thus, may contribute to both dimerization stability and specificity.

The amino acids variability can be explained by side-chain residues orientation relative to the axis formed by two  $\alpha$ -helices. The side chains of *a* residues point towards the external super-coil surface, favoring  $\beta$ -branched residues for tight packing (for example isoleucine or valine). The side-chains of *d* residues point straight away the hydrophobic interface, favoring unbranched amino acids;<sup>33</sup> this is the characteristic packing for the “knob-into-holes” described by Crick.<sup>27</sup>

<sup>31</sup> Steinmetz M.O. *Proc. Natl. Acad. Sci. USA*, **2007**, 104(17), 7062-7067.

<sup>32</sup> a) Litowski J.R., Hodges R.S. *J. Pept. Res.*, **2001**, 58(6), 477-492. b) Litowski J.R., Hodges R.S. *J. Biol. Chem.*, **2002**, 207(40), 37272-37279.

<sup>33</sup> a) Zhu B.Y. *Protein Sci.*, **1993**, 2(3), 383-394. b) Chao H., Bautista D.L., Litowski J., Randall Irvin T., Hodges R.S. *J. Chromatogr. B Biomed. Sci. Appl.*, **1998**, 715(1), 307-329.



**Figura** : Schematic representation of “kno-into-hole” described by Crick for two parallel right-handed helices.

### 1.3 Biological combinatorial libraries of peptide

Combination of a relatively small number of amino acids is the strategy used by nature to construct proteins which are the most variable and adaptable molecules of life. No other molecule can be organized in such a large number of different structures and carry out so many different functions as proteins. Proteins are the principal mediators of molecular “crosstalking” and their extreme variability makes them suitable mediators of specific molecular interactions. Interactions between proteins or between proteins and other molecules, mediate all the physiological and pathological phenomena of life and as such can be the targets of specific drugs to modify physiological or pathological events.

Evolution of chemical and biological technologies for the production of peptides and proteins in the laboratory has made it possible to use the same natural molecular approach to construct new peptide sequences which recognize specific target. The intrinsic variability of peptide sequences and their ability to specifically recognize other molecules can therefore be used to design peptide drugs.

A biological library consists of a pool of microorganisms expressing different polypeptides. Each microorganism carries only one encoding DNA or RNA sequence for certain peptide, representing one clone; each single clone of the library can be propagated and it will express the same peptide.<sup>34</sup>

A polypeptide library construction starts with the design of encoding DNA sequence; the source for this insert can be a pool of chemically synthesized degenerated oligonucleotides, cDNA, genomic DNA fragments or mutagenized specific gene fragments. The library will be constituted by viral particles or by cells.

<sup>34</sup> Mersich C., Jungbauer A. *J. Chrom.B.*, **2008**, *861*, 160-170.

The next important step will be the screening of the library against the target molecule. Clones, identified as binders to the target substance, will be sequenced and their coding regions will be translated into the particular peptide sequences. One method of designing a random peptide library is the use of random oligonucleotides. In a fully degenerated oligonucleotide each triplet will code for one of the 64 possible codons and at each coupling reaction an equal mixture of all four nucleotides will be used for all three positions in the triplet. In this way the oligonucleotide will contain all 64 possible codons and all 20 amino acids and three stop codons will be represented.

Peptide libraries can be used to “fish” active ligands from a large variety of structural combinations making it possible to identify peptide mimotopes, which mimic the structure and function of a native protein. The principle for selecting of new ligands for a variety of biological targets by means of combinatorial peptide libraries exploits the natural diversity of protein-protein and protein-ligand interactions.

Molecular biology techniques make it possible to build polypeptide libraries expressed by different biological entities including bacteria, phages and ribosomes.

### 1.3.1 Phage display

Phage display is the most well-studied high-throughput screening method to study protein interactions, based on displaying peptides or proteins on the surface of a bacteriophage.

The origin of the phage display dates to the mid-80s when Smith first expressed a foreign segment of protein on the surface of bacteriophage M13 virus particle.<sup>35</sup> Smith demonstrated that a fusion phage can be isolated from a background of phage particles displaying no antigen with more than 1000-fold enrichment, allowing phages displaying the correct antigen to be affinity purified against the cognate antibody.

Phages are viruses that infect bacterial cells and many of the vectors used in recombinant DNA research are phages that infect the standard recombinant DNA host: the bacterium *Escherichia Coli*. The key feature of recombinant DNA vectors, including phages, is that they accommodate segments of “foreign” DNA-pieces, human DNA, for instance, or even stretches of chemically synthesized DNA. As vector DNA replicates in its *E. Coli* host, the foreign insert replicates along with it as a sort of passenger. An “expression vector”, including a phage-display vector, has an additional feature compared to vectors in general: the foreign DNA is expressed as a protein. Phage display differs from conventional expression system, however, in that the foreign gene sequence is spliced into the gene for one of the phage coat proteins, so that the foreign amino acid sequence is genetically fused to the endogenous amino acids of the coat protein to make a hybrid fusion protein.

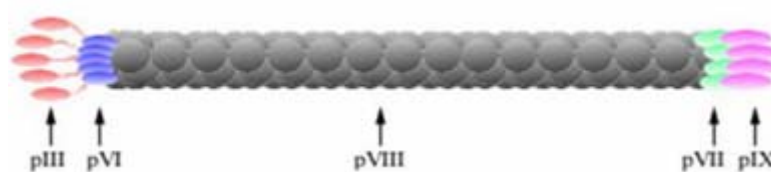
---

<sup>35</sup> Smith G.P., Petrenko V.A. *Chem. Rev.*, **1997**, *97*, 391-410.

The hybrid coat protein is incorporated into phage particles (virions) as they are released from the cell, so that the foreign peptide or protein domain is displayed on the outer surface.

The most investigated phage used for the phage–display are fF phage, including fd, f1 and M13; display systems based on bacteriophage T4 and  $\lambda$  are extremely promising but their use is not so frequently.

Filamentous phages are flexible rods about 1  $\mu\text{m}$  long and 6 nm in diameter, composed mainly of a tube of helically arranged molecules of the 500-residue major coat protein pVIII; there are 2700 copies in wild-type virions, encoded by a single phage gene VIII. Inside this tube lies the single-stranded viral DNA (Figure 1.13).



**Figure 1.13:** phage structure with pIII, pVI, pVIII, pVII and pIX protein.

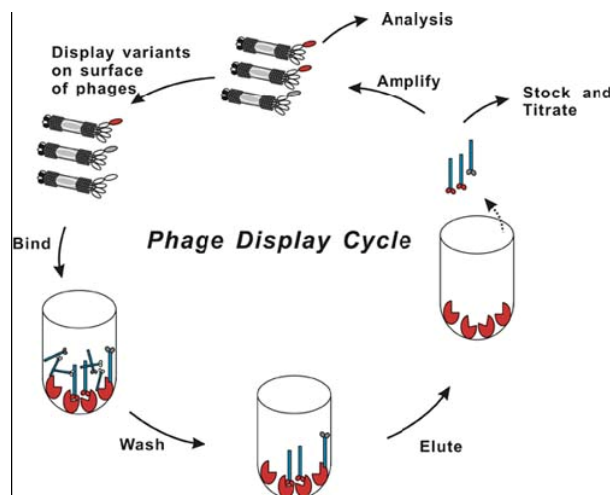
At one tip of the particle there are five copies each of the minor coat proteins pIII and pVI encoded by gene III and gene VI respectively; minor coat proteins pVII and pIX (genes VII and IX) are at the other tip.

### 1.3.2 Phage display libraries

Generally the phage display cycle starts with the creation of diversity at the DNA level, creating a library of different variants. The DNA fragments are usually cloned upstream the gene encoding the protein III or the protein VIII of the phage: in the first case there are 3 to 5 recombinant proteins at one end of the phage; in the second, all 2700 copies of the major coat protein are recombinant. Excluding the display of short peptides, the protein III based system is generally preferred. Although pIII and pVIII have been used to display proteins, pIII is the display protein of choice, having been used to display large numbers of different proteins.

In a typical phage display selection, the phage displaying the protein of interest P is retained on a surface, coated with an antigen or antibody recognizing P, while non-adherent phage are washed away. Bound phages can be removed from the surface, re-infected into bacteria and re-grow for further enrichment and eventually for binding analysis. The possibility to perform successive round of selection permits the isolation of proteins present in very low number in a population of billions different phages. Usually, three or five rounds of panning are sufficient to enrich for binding peptide sequence.

Phagemide vectors have been developed, containing an Ff origin of replication and the gene encoding pIII can be manipulated and propagated as plasmids, and when display is required, bacteria containing the vectors are infected with a helper phage, which provides all other components necessary for the creation of recombinant, selectable phage particles.



**Figure 1.14** – Schematic representation of a typical round selection with phage display

Over the past year, significant advances have been achieved through the use of phage-displayed peptide libraries. A wide variety of bioactive molecules, including antibodies, receptors and enzymes, have been selected high-affinity and/or highly-specific peptide ligands from a number of different types of peptide library. The demonstrated therapeutic potential of some of these peptides, as well as new insights into a protein structure and function that peptide ligands have provided, highlight the progress made within this rapidly expanding field.

A number of different groups have screened peptide libraries with monoclonal antibody (mAbs) and have isolated “peptide mimics” that cross react with the mAbs. Typically, phage libraries are generated to encode small foreign peptides which can range in size from 6 to 45 amino acids, while retaining large diversity. Peptide libraries may be displayed as linear or cysteine-constrained sequences. Traditionally, the process of affinity selection is an iterative process of several rounds of selection, elution and amplification of phage; recently new developments in high throughput methods have allowed for the screening and characterization of large numbers of phage clones simultaneously. Most procedures are performed in multiwell plates and involve fluorimetric and spectrophotometric data measurements.

Demanegl et al. previously isolated clones bearing linear –epitope-mimic peptides by screening a disulfide constrained 6-mer (CX<sub>6</sub>C; single letter amino acid code where X may be any amino acid) peptide library with a malaria-specific mAb. More recently, this group isolated two mAbs from the malaria-binding, anti-phage response and compared their heavy chain variable region and

light chain variable region sequences with those of the parent mAb that was initially used to isolate the phage clone.

There is also evidence that immunogenic-mimic peptides can be isolated by polyclonal serum Abs. In previous work, Mennuni et al. isolated several diabetes-specific peptides by screening two peptide libraries with serum IgGs from among these were bound by the serum Abs from a number of patients having full-blown diabetes. Mennuni et al. showed that, on immunization of rabbits, a phage clone bearing one of these peptides elicited Abs that recognized pancreatic-islet cells. Thus this immunogenic-mimic peptide may reveal new antigens (or epitopes) involved in the evolution of type-1 diabetes, especially if it appears to produce pathogenic Abs as a hapten on its own.

In another case, phage display libraries of small peptides have been used in the chemical catalysis.<sup>36</sup> Peptide possessing aldolase activity, in fact, was produced by Tanaka et al., displayed on the surface of phage and selected by reaction of peptide phage with 1,3-diketone that have been employed for the induction of aldolase antibody.

### 1.3.3 Bacterial display

Surface display of heterologous proteins in bacteria has been employed as a tool for basic and applied research in microbiology, molecular biology, vaccinology and biotechnology.

The most frequently used host for bacterial display is Gram-negative bacterium *Escherichia Coli* owing to rapid growth rate, ease of genetic and physical manipulation, and its suitability for making large libraries of up to  $10^{11}$  members.

Bacterial display scaffolds can be broadly grouped into those allowing N-terminal, C-terminal and sandwich fusions. Display of passenger polypeptides as N-terminal fusions with a surface-exposed N-terminus of the display scaffold can be accomplished via fusion to autotransporter proteins. Autotransporters used for library screening include the IgA protease from *Neisseria gonorrhoeae*, *E. Coli* AIDA-I or *EstA* from *Pseudomonas aeruginosa*.<sup>37</sup>

Display via the scaffold's C-terminus may be beneficial to enhance the diversity of peptide libraries since stop codons arising from common randomization schemes and nonintended errors (primer deletions or PCR errors) can yield functional binders without truncating the carrier protein. C-terminal libraries display have been generated and screened using intimins (EaeA), invasins and the LppOmpA vector.

Sandwich fusion is the most commonly used strategy for the surface display of proteins in Gram-negative bacteria. Three classes of protein have been used as carrier proteins: outer membrane proteins (OMPs), subunit proteins of extracellular appendages and S-layer proteins. OMPs form

---

<sup>36</sup> Tanaka F., Barbas III C.F. *Chem.Comm.*, **2001**, 769-770.

<sup>37</sup> Yang T.H., Pang J.G., Seo Y.S., Rhee J.S. *Appl. Environ. Microbiol.*, **2004**, 70, 6968-6976.

transmembrane  $\beta$ -barrels on the outer membrane. The  $\beta$ -barrels are composed of antiparallel  $\beta$ -strands pair connected by short loops on the periplasmic side and by long loops on the external side. The external loops are generally less conservative and therefore seem to be tolerant to a certain degree of modification such as substitution, insertion and deletion. The external loops can potentially be used as fusion sites for the display of heterologous proteins. It has generally been believed that the external loops of OMPs could only accept foreign peptides of 70 amino acids or less owing to the disruption of membrane integrity of carrier protein. However, it has more recently been demonstrated that *E. Coli* OmpC could be used as a sandwich fusion partner, displaying much longer polypeptides of 162 amino acids, which is the largest peptide inserted to date using the sandwich fusion method.<sup>38</sup>

In many cases for to generate bacterial libraries is possible to use Gram-positive bacteria. Many surface proteins of Gram-positive bacteria are covalently immobilized to the cell wall, typically involving a specific C-terminal sorting signal consisting of 32-38 amino acids.<sup>39</sup> Staphylococcal protein A (SpA) has often been used as a model system to study anchoring mechanisms of surface protein in Gram-positive bacteria. The sorting signal includes an LPXTG sequence motif, where X denotes any amino acid, followed by a stretch of 23 hydrophobic residues at the extreme C-terminus.<sup>39</sup> The LPXTG sequence contains a cleavage site for sortase, which is an enzyme of 206 amino acids that cleaves polypeptides between the treonine and the glycine of the LPXTG motif. The anchoring domain, which includes a standard cell wall sorting signal as well as an appropriate cell wallspanning region, has been used to display various polypeptides and proteins.

#### 1.3.4 Ribosome display

Ribosome display is a totally “cell-free” method of selection and evolution of proteins that is not limited by any cell-based translation steps. It is an *in vitro* selection method for the isolation of proteins and peptides from large libraries that was successfully applied to the affinity maturation of antibodies using eukaryotic or procaryotic systems. The use of ribosome display overcomes current limitations for protein selection technologies as diversity is not limited by transformation efficiency but by the physical number of ribosome present and different mRNA molecules available *in vitro*.

The process results in translated proteins that are associated with their mRNA progenitor that is used, as a complex, to bind to an immobilized ligand in a selection step. The mRNA-protein hybrids that bind well are then reverse-transcribed to cDNA and their sequence amplified using PCR. The end result is a nucleotide sequence that can be used to create tightly binding proteins.

---

<sup>38</sup> Xu Z., Lee S.Y. *Appl. Environ. Microbiol.*, **1999**, *65*, 5142-5147.

<sup>39</sup> Schneewind O. *EMBO J.*, **1993**, *12*, 4803-4811.

The key issue of the strategy is dependent on coupling individual nascent proteins to their corresponding mRNA for selection. This selection enables the isolation of proteins against virtually any antigen including self-molecules, certain drugs, potent toxins and haptens, which would normally be impossible to raise with in *vivo*-based systems because of toxicity and lack of immunogenicity.

Ribosome display offers two important advantages: first, large libraries can be made rapidly because there is no need to transform large numbers of mutant plasmid into a host and second, additional mutations can be introduced at every round because a PCR step is included in each selection cycle rather than an in vivo amplification step.

One of the most powerful applications of ribosome display has been the isolation of recombinant antibodies with high affinity specificity. In this regard ribosome display technologies can be used:

1. to isolate human antibodies from patients exposed to certain viral pathogens to understand better the immune response during infection and how protective antibodies are generated,
2. to generate human antibodies, significant for cancer immunotherapy, and
3. elucidate the specificity of autoimmune antibodies

A protocol of production of human single-chain Fv antibodies by ribosome display is established by Hell et al.

Because ribosome display avoids the problems of cytotoxicity, soluble protein expression and secretion bias in cell-based systems, it could be an ideal means by which to display functional proteins for drug discovery applications like target discovery and functional identification. With the completion of genome sequences, it is possible to design general cDNA libraries for ribosome display. In combination with high-throughput proteins arrays, the screening power of ribosome display could be further increased, permitting library-versus-library screening and genome-wide analysis of protein-protein interactions.

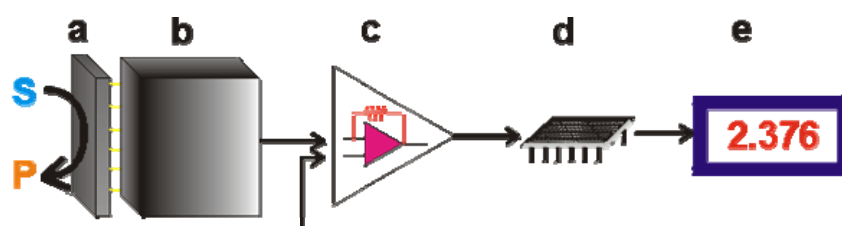
## Chapter 2: Biosensors and microarrays

### 2.1 Biosensors

Good analytical devices are available now but new reliably, cheaply and rapidly devices are urgently needed to detect environmental pollutants and medical-related targets. The conventional techniques of high-pressure liquid chromatography (HPLC), gas chromatography (GC) or enzyme-linked immunoassorbent assay test (ELISA-test) allow for a highly accurate and sensitive determination of the exact chemical composition of many samples, but these methods disadvantages including the following:

- 1-they enable only the detection of a single compound or a group of structurally related compounds at any given time,
- 2-they cannot assure the biological effects of the tested compounds (e.g. cytotoxicity or genotoxicity)
- 3-these test may take a few days and require highly skilled people as well as expensive equipment to conduct.<sup>40</sup>

To analyse a molecular interaction with or without the use of reporter labels, it is necessary to couple a molecular recognition element (e.g. an antigen or target receptor) to a transducer that converts a chemical or biological interaction into an electrical signal. A biosensor is thus defined as a unique combination of a receptor for molecular recognition and a transducer for transmitting the interaction information into an electrical signal.<sup>41</sup>



Finally the role of biosensors is to convert a biological response into an electrical signal: biological material is directly associated with a physicochemical transducer. Only certain substances, that are supposed to be interactors of biological part, will result in an optical or electrical signal of physical transducer. The specificity of the biosensor will depend on the binding abilities of the biorecognition element; the sensitivity on the transducer used for the detection of binding.<sup>42</sup>

<sup>40</sup> Eltzov E., Marks R. *IEEE Instr. Meas. Magazine*, **2009**, 10-16.

<sup>41</sup> Cooper M. A. *Drug Discov. Today*, **2006**, *11*, 1068-1074.

<sup>42</sup> Pohanka J. *Appl. Biomed.*, **2009**, *7*(3), 121-127.

### 2.1.1 The biorecognition element

The biorecognition element is chosen according to the target analyte on the report method. Any of the possible bioreceptor elements of a biosensor will indicate interaction with the target analytes; this capability ensures that the sensor is selective. The biological components of a biosensor are not only responsible for the correct recognition of the analyte but also the generation of the physico-chemical signal monitored on the transducer and, ultimately, the sensitivity of the final device.<sup>43</sup>

They can be divided in two categories: catalytic and non-catalytic. The catalytic group include enzymes, microorganisms and tissues. Devices incorporating these elements are appropriate for monitoring metabolites in the millimolar to micromolar range and can be used for continuous monitoring. The non-catalytic or affinity class biological components comprises antibodies (or antigen), lectins, receptors and nucleic acids which are more applicable to “single use” disposable devices for measuring hormones, steroids, drugs, microbial toxins, cancer markers and viruses to concentration in the micromolar to picomolar range.<sup>44</sup>

From an analytical point of view, the most important class of enzymes are oxidoreductases, which catalyses oxidations using oxygen or NAD, and the hydrolases, which catalyse the hydrolyse.<sup>45</sup> Most successful biosensors exploit enzymes as the biological recognition/response system because of the range of transducible components such as protons, ions, heat, light, electrons and mass that can be exchanged as part of their catalytic mechanism.<sup>46</sup> This catalytic activity is controlled by pH, ionic strength, temperature and the presence of cofactors. Enzyme stability is usually the deciding factor in determining the lifetimes of enzyme based biosensors, typically between one day and one or two months.<sup>47</sup>

In many case organelles, whole cells or tissue section have been used directly as biocatalytic element since they contain, together with numerous enzymes, all the other necessary components needed to convert substrates into products in an environment which has been optimized by evolution. The appropriate use of enzyme inhibitors, activators and stabilizing agent can be used to enhance the selectivity and lifetimes of tissue based biosensors.<sup>45</sup>

Naturally occurring receptors are non-catalytic proteins that span cell membranes, extending into both the extracellular and intracellular spaces; they are involved in the chemical senses such as olfaction and taste as well as in metabolic and neural biochemical pathways. Within the organism they act as links in cell-cell communication and they are also the binding site for many drugs and toxins.

---

<sup>43</sup> Eltzov E., Kushmaro A., Marks R. *Biosensor and Related Techniques for Endocrine Disruptors, Endocrine Disrupting Chemicals in Food*, 2008, ed. I. Snow, Cambridge, U.K.

<sup>44</sup> a) D’Orazio *Clin. Chim. Acta*, 2003, 334(1-2), 41-69. b) Griffiths D., Hall G. *Trends. Biotechnol.*, 2003, 11, 122-130.

<sup>45</sup> Li H., Liu S., Dai Z., Bao J., Yang X. *Sensors*, 2009, 9(11), 8547-8561.

<sup>46</sup> Lowe C.R. *Phil. Trans. R. Soc. Lond.*, 1989, B324, 487-496.

<sup>47</sup> D’Auria S., Lacowikz J.R. *Curr. Opin. Biotech.*, 2001, 15, 201-218.

Two methods have been defined by which binding of a transmitter molecule to the extracellular side of the receptor leads to modification of intracellular processes. The first method is by opening a transmembrane ion channel and the second method is by activation or inhibition of membrane enzyme that generate secondary messenger compounds.<sup>48</sup> All three processes involved in this situation (binding specificity, ion channel opening and enzyme activation) provide possible opportunities for receptor utilization in biosensors. Attempts at using neuroreceptors as the recognition elements in biosensors have largely been restricted to the nicotine acetylcholine receptor which can be isolated from the electric organ or electric eel or ray in relatively large quantities. The unavailability of other receptors for biosensors use is no doubt a reflection of the fact that they are normally present in small amount in tissues and are unstable once removed from their natural lipid membrane environment.<sup>49</sup>

Antibodies are a molecules produced by animals in response to the antigen and which bind to the latter specifically. Antibody to smaller molecular weight environmental contaminants such as pesticides, herbicides, microbial toxins or industrial chemicals can be made attaching the latter to carrier proteins like BSA (Bovine Serum Albumin) or KLH (Keyhole Limpet Haemocyanine)<sup>50</sup>: the small molecular component of the resultant conjugate is known as haptan. Direct measurements of the resultant change in mass can be made with acoustic devices<sup>51</sup> and of the change in refractive index with surface plasmon resonance.<sup>52</sup> Binding can also be detected at the surface of optical fibres by conjugating the antigen or antibody to a fluorescent label,<sup>53</sup> or with electrochemical devices by the use of an enzyme label.<sup>54</sup>

The ability of a single-stranded nucleic acid molecule to recognize and bind to its complementary partner is known like DNA-hybridization and has been used in genetic analyses and can also be used in a biosensor. Over the last decade there has been a movement away from radiometric methods towards the application of electrochemical, microgravimetric, and optical transduction system for the measurement of nucleic acid binding reactions.<sup>55</sup> Microgravimetric methods have been reported to identify the increase in mass due to the hybridization of double single stranded DNA and its complementary strand in a sample<sup>56</sup> and the measurement of surface

---

<sup>48</sup> Cooper M.A. *J. Rec. Signal Transd.*, **2009**, 29 (3-4), 224-233.

<sup>49</sup> Spisak S., Tulassay Z., Molnar B., Guttman A. *Electrophoresis*, **2007**, 28(23), 4261-4273.

<sup>50</sup> a) Harrison R.O., Goodrow M.H., Gee S.J., Hammock B.D. *Immunoassay for Trace Chemical Analysis*, **1991**, ed. Houghton. b) Vanderlan S.H., Stanker B.E., Watkins D.W., Roberts M., **2000**, *American Chemical Society Symposium Series*, 451, 14-27. c) Harris J.R., Markl J. *Micron*, **1999**, 64, 79A-88A.

<sup>51</sup> Ferreira G.N.M., da-Silva A.C., Tomè B. *Trends in Biotech.*, **2009**, 27(12), 689-697.

<sup>52</sup> Mc Loughlin M., Collins D. *J. of Chromat. B*, **2009**, 878(2), 149-153.

<sup>53</sup> a) Anis A.N., Wright J., Rogers K.R., Valdes J.J. *Anal. Lett.*, **2002**, 25, 627-635. b) Gauglitz G. *Anal. Bioanal. Chem.*, **2005**, 381(1), 141-155.

<sup>54</sup> a) Aizawa M. *Phil. Trans. R. Soc. Lond.*, **1997**, B316, 121-134. b) Lee W.E., Thompson H.G., Hall J.G., Fulton R.F., Wong J.P. *J. Immunol. Meth.*, **1993b**, 166, 123-131 c) Ogert R.A., Burans J., Liegler F.S. *Proc. SPIE Int. Soc. Opt. Eng.*, **1994**, 2068, 151-158.

<sup>55</sup> Miranda Castro R. Miranda-Ordieres A.J. *Electroanalysis*, **2009**, 21(19), 2077-2090.

<sup>56</sup> Lippa P.B., Sokoll L.J., Chan D.W. *Clin. Chim. Acta*, **2001**, 314(1-2), 1-26.

plasmon resonance at a liquid-metal film interface has also been used to detect the hybridization of bound DNA.<sup>57</sup>

### 2.1.2 The secondary transducer

Biosensors can also be classified into four different basic groups on the basis of signal transduction: electrochemical, optical, mass sensitive and thermal sensor. Most of the biosensors described in the literature are electrochemical biosensors. Electrochemistry offers high sensitivity, compatibility with modern microfabrication technologies, portability, low cost and minimal power requirements.<sup>58</sup>

Optical biosensors are the most common after amperometric and potentiometric biosensors. The various types of optical transducer exploit properties such as light absorption, fluorescence or phosphorescence, bio-chemiluminescence, reflectance, Raman scattering and refractive index. There are other some example of optical techniques usually utilised in biosensors like fibre optic, optical waveguide structures and surface plasmon resonance (SPR).<sup>59</sup>

Within mass-sensitive biosensors, acoustic wave biosensors operate on the basis of an oscillating crystal that resonates at a fundamental frequency. The crystal element is coated with a layer containing the biorecognition element designed to interact selectively with the target analyte. A measurable change in the resonance frequency occurs after the binding of the analyte on the sensing surface according to the mass change of the crystal. Most of these biosensors utilise piezoelectric materials as the signal transducers<sup>51</sup> and it provides real-time data of binding events but, as the sensitivity is lower than SPR, this method has been used less frequently.<sup>60</sup> An emerging group of mass-sensitive biosensors are the so-called cantilever biosensors, which are based on the bending of microfabricated silicon cantilevers. The mass change, originating from the adsorption of target molecules on the microcantilever surface causes a differential surface stress change, and therefore a bending or deflection of the cantilever; the deflection of a few nanometres can be detected by an electric or optical methods.<sup>61</sup>

Microcantilever biosensors offer various advantages due to their microscopic dimensions (in the range of  $10^{-3}$  mm<sup>2</sup>): only small quantities of receptor and analyte are necessary and limits of detection achieved are often lower than those obtained by classical methods.<sup>62</sup>

Thermometric biosensors, finally, exploit the absorption or evolution of heat in biological reactions. This is reflected as a change in the temperature within the reaction medium and is

---

<sup>57</sup> Scarano S., Mascini M., Minunni M. *Biosens. Bioelect.*, **2009**, 25(5), 957-966.

<sup>58</sup> Wang J. *Anal. Chim. Acta.*, **2002**, 469, 63-71.

<sup>59</sup> a) Collings A.F., Caruso F. *Rep. Prog. Phys.*, **1997**, 60, 585-591. b) Vellusamy V., Arshak K., Adley C. *Biotech. Adv.*, **2010**, 28(2), 232-254.

<sup>60</sup> a) Subrahmanyam S., Piletsky S.A., Turner A.P.F. *Anal. Chem.*, **2002**, 74, 3942-3951. b) Tamayo J., Alvarez M., Lechuga L.M. *Sens. Actuators*, **2003**, 89, 33-39.

<sup>61</sup> Fritz J. *Analyst*, **2008**, 133(7), 855-863.

<sup>62</sup> Lang H.P., Berger P., Battiston F.M., Fornaro P. *Anal. Chim. Acta*, **1999**, 393, 59-65

transducer by a change in the resistance of a thermistor, which acts as a temperature transducer. Thermistor-based calorimetric biosensors have mainly been applied to clinical and industrial process monitoring; in the environmental field, only a few applications have been described up to now: several pesticides, for example, have been measured by thermometric biosensors.<sup>63</sup>

### 2.1.3 Immobilisation of biorecognition element

Immobilisation is one of the key steps in the development of biosensors. It not only helps to achieve the required close proximity between the biomaterial and the transducer, but also to stabilise the biomaterial for reuse.<sup>64</sup> The most generally employed immobilisation procedures are adsorption, cross-linking, covalent binding, entrapment, sol-gel entrapment, self-assembled biomembranes and immobilisation based on bulk modification.

Physical adsorption of the biocomponent, based on hydrophobic, electrostatic and van der Waals attractive forces, is the oldest and simplest immobilisation method. It does not require chemical modification of the biological components and allows the regeneration of the matrix membrane. In spite of its simplicity, this immobilisation method is clearly restricted to a limited number of applications because of its long-term instability. Loss of adsorbed biological components is possible if changes in pH, ionic strength or temperature occur during measurements.<sup>65</sup>

The cross-linking method is based on the formation of covalent binding between molecules of an inert bed and the functional reagents. Cross-linking procedures have been attractive due to their simplicity and to the strong chemical binding achieved for the biomolecules; the main drawback is the possibility of activity losses due to the distortion of the active enzyme conformation during the cross-linking. Cells<sup>66</sup> and enzymes<sup>67</sup> have been successfully immobilised by cross-linking method.

The entrapment of cells in polymer structures may create diffusion barriers, which could hinder the transport of analytes and result in prolonged washing, equilibration and measuring steps.<sup>68</sup>

The other common approach is to retain an enzyme solution, a suspension of cells or a slice of tissue in close proximity to the transducer surface, confined by analyte semi-permeable membranes such as dialysis membranes or ultrafiltration membranes, which permit the free passage of the substrate and product but retain the sensing element.

Sol-gel entrapment is based on inorganic material that is able to give physical rigidity, chemical inertness and high photochemical and thermal stability at the final biosensor structure.<sup>69</sup> Solids prepared by sol-gel chemistry shrink around hosted reagents, thereby blocking them from

---

<sup>63</sup> Ramanathan K., Danielsson B. *Biosens. Bioelectron.*, **2001**, *16*, 417-423.

<sup>64</sup> D'Souza S.F. *Biosens. Bioelectron.*, **2001**, *16*, 337-353.

<sup>65</sup> Mello L.D., Kubota L.T. *Food Chem*, **2002**, *77*, 237-256.

<sup>66</sup> D'Souza S.F. *Biosens. Bioelectron.*, **2001**, *16*, 337-353.

<sup>67</sup> Tyagi R., Batra R., Gupta M.N. *Enzy. Microb. Technol.*, **1999**, *24*, 348-354.

<sup>68</sup> Keusgen M. *Naturwissenschaften*, **2002**, *89*, 433-444.

<sup>69</sup> Dong S., Chen X. *Rev. Molec. Biotech.*, **2002**, *82*, 303-323.

leaching into the sample solution, but possess a porous structure that provides access to small analyte. Sol-gel chemistry yields a matrix that retains the protein's native conformation and reactivity and avoids the denaturation of proteins during the immobilization.<sup>70</sup>

Biomembranes are self-organised structures consisting of amphiphilic lipid molecules. In water solution thermodynamics drive these molecules to assemble into structures such as self-assembled monolayer membranes (SAMs) and self-assembled bilayer membranes (BLMs). SAMs offer promising possibilities in the control over the orientation of the molecule immobilised<sup>71</sup> and BLMs have recently been used in various experimental biosensors as models of actual living cell membranes. Molecular self-assembly occurs in nature and produces nanosystem far more advanced than man-made system. Inspired by the nature those nanostructures are applied to the development of biosensors.<sup>72</sup> Nanotechnology is generating new ideas not only for immobilisation purposes but also in the design of nanostructures with particular optical properties.

At the end, a biocatalyst can be incorporated within the bulk of entire electrode material: a carbon composite matrix or graphite epoxy resin. These modified biosensors offer several advantages such as the close proximity of the biocatalytical and sensing sites, possibility to incorporate other components (cofactors), easy renewing of the surface, economy of fabrication and high stability of the incorporated biocatalysts. Biosensors with these features can be the carbon paste electrodes, screen-printed biosensors and solid binding matrices of defined molecular structure.<sup>73</sup>

## 2.2 Microarrays

Microarray, sometimes referred to as a chips or arrayed libraries, can be classified into two general types: biochips (biomicroarray) and chemical microarray.<sup>74</sup> Biochips are usually generated from biochemical or biological components, such as protein (including enzymes and antibodies), DNA, cells<sup>75</sup> and tissues;<sup>76</sup> chemical microarrays consist of arrays of organic compounds including small organic molecules, peptides and sugars.<sup>77</sup>

In recent years microarrays have become an invaluable research tool for life scientists and their use in diagnostic has emerged as a great promise of medicine. Microarrays consist of immobilized biomolecules spatially addressed on substrates such as planar surfaces (typically coated microscope glass slide), microwells or array of beads. Immobilized biomolecules, here

---

<sup>70</sup> Tess M.J., Cox J.A. *J. Pharm. Biomed. Anal.*, **1999**, *19*, 56-58.

<sup>71</sup> Wink T.H., van Zuilen S.J., Bult A., van Bennekom W.P. *Analyst*, **1997**, *122*, 43-50.

<sup>72</sup> Tiefenauer L.X., Kossek S., Pedeste C., Thiebaud P. *Biosens. Bioelectron.*, **1997**, *12*, 213-223.

<sup>73</sup> Miertus S., Katrik J., Pizzariello A., Strend'ansky M., Svitel J., Svorec J. *Biosens. Bioelectron.*, **1998**, *13*, 911-923.

<sup>74</sup> Quingchai X., Kit L. *J. Biomed. Biotech.*, **2003**, *5*, 257-266.

<sup>75</sup> Ziauddin J., Sabatini D.M. *Nature*, **2001**, *411(6833)*, 107-110.

<sup>76</sup> Moch H., Kononen T., Kallioniemi O.P., Sauter G. *Adv Anat Pathol*, **2001**, *8(1)*, 14-20.

<sup>77</sup> Robinson W.H., Di Gennaro C., Genovese M.C., Smolen J. H. *Nat. Med.*, **2002**, *8(3)*, 295-301.

referred as probes, usually include oligonucleotides, PCR products, proteins, peptides, carbohydrates and other small molecules.

Generally a protein or peptide microarray are made by a large set of this biological molecule arrayed on a solid support; after washing and blocking surface unreacted site the array is probed with a sample containing, among a variety of unrelated molecules, the counterparts of the molecular recognition events under study. If the interaction occurs, a signal is revealed on the surface by a variety of detection techniques; or with label-free methods like mass spectrometry (MS), surface plasmon resonance (SPR), atomic force microscopy (AFM), quartz crystal microbalance (QCM), or directly with a labelled probe.

The aim of peptide and protein microarrays is the study of molecular interactions occurring between two partners: one contained in a liquid sample and one immobilized on a solid support. The key requirements of the surface hosting a protein microarray assay are:

1. Provision of an optimal binding capacity of capture ligands (probes),
2. Retaining of biological activity of capture ligand, in order to allow internal hydrophobic side chains to form hydrophobic bonds with the solid surface,<sup>78</sup>
3. Accessibility of the ligand by the interaction partner (protein-substrate interaction reduce the accessibility of target, possibly leading to false negative results). This issue is particularly important for peptide microarrays due to the small molecular mass of capture ligands.
4. Low degree of aspecific interaction; the achievement of a low degree of aspecific binding is extremely difficult when the sample is a complex mixture of thousand of molecules such a serum.<sup>79</sup>

One of the most important point in the microarray development is the binding of protein or peptide over the solid surface. The simplest way for binding a protein is through surface adsorption. This approach has been used in standard ELISA and Western blot for many years and is based on adsorption of the macromolecules either by electrostatic forces on charged surfaces (poly-Lysine coated slides)<sup>80</sup> or by hydrophobic interactions (nitrocellulose or polyvinylidene fluoride membranes).<sup>81</sup>

The covalent binding of proteins and peptides on the substrate represent a more robust approach. Considerable effort has been made in the last years to improve immobilization of protein on modified glass surface using a number of different strategies which recently have been compared systematically.<sup>82</sup> The covalent mechanism of attachment requires the presence of

---

<sup>78</sup> Buttler J.E. *Methods*, **2000**, 22(1), 4-23.

<sup>79</sup> Kuznesow W., Hoheisel J.D. *J. Mol. Recognit.*, **2003**, 16(4), 165-176.

<sup>80</sup> Haab B.B., Dunham M.J., Brown P.O. *Genome Biol.*, **2001**, 2(2), 4.

<sup>81</sup> Bussow K., Cahill D., Nietfeld W., Bancroft D. *Nucleic Acids Res.*, **1998**, 5007-5008.

<sup>82</sup> Guillaume B., Buness A., Schmidt C., Klimek F., Huber W., Arlt D., Korf U., Wiemann S., Poustka A., *Proteomics*, **2005**, 5(18), 4705-4712.

reactive support such as epoxides,<sup>83</sup> aldehydes,<sup>84</sup> succinimidyl esters or isothiocyanate functionalities<sup>85</sup> able to react with nucleophilic groups (amino, thiol, hydroxyl) on the ligand molecules.

The use of a matrix that embeds the protein in a structured environment is an alternative way to immobilize proteins. This mechanism does not involve the cross linking of the capture molecules with the surface, but is based on the physical entrapment of proteins in gel such as polyacrylamide<sup>86</sup> or agarose.<sup>87</sup> The three-dimensional structure of these substrates generally increases the loading capacity and does not disturb the potential functional sites or regulatory domains of the protein; moreover the aqueous environment of the gel reduces protein or peptide denaturation. However, the gel structure can represent a barrier to the diffusion and the molecular recognition events may require long incubation times. A mixed approach, based on nanoengineered 3D polyelectrolyte thin films (PET) deposited on glass slide by consecutive adsorption of polyelectrolyte via self-assembled technique, has been reported.<sup>88</sup> Proteins were immobilized in the PET-coated glass slides by electrostatic adsorption and entrapment on porous structure on the 3D polymer film.

In addition to non-specific methods, probe molecules can be tagged and immobilized by a specific non-covalent interaction between the tag and an immobilized capturing molecules. The attachment to the surface is typically mediated by a molecular recognition event such as biotin-avidin interaction, by interaction between nickel coated slides and His-tagged proteins or glutathione and GST tags. An innovative affinity-based protein immobilization strategy was developed based on the combined use of the carbohydrate binding module (CBM) and its natural three-dimensional substrate such as cellulose.<sup>89</sup> CBM fused antibodies or peptide were produced, immobilized on cellulose surfaces and used for serodiagnosis of human immunodeficiency virus patients.

Besides the attachment mechanism (adsorption, physical entrapment into gels, covalent binding or oriented molecular recognition), also the molecular architecture of the coating plays an essential role and must be considered. Proteins or peptides, regardless of the attachment mechanism, can be directly in contact with the activated surface on a one dimensional coating; alternatively the probe can be inserted in a simple monolayer architecture such as in SEM or poly ethylen glycol (PEG)<sup>90</sup> or in a more complex three-dimensional structure like a dendrimeric

---

<sup>83</sup> Zhu H., Klemic J.F., Chang S., Bertone P., Casamayor A., Smith D., Gerstein M., Reed M.A., Snyder M. *Nat. Genet.*, **2000**, 26(3), 283-289.

<sup>84</sup> MacBeath G., Schreiber P.L. *Science*, **2000**, 289, 1760-1763.

<sup>85</sup> Benters R., Neimheyer C.M., Wohrle D. *ChemBiochem*, **2001**, 2(9), 686-694.

<sup>86</sup> Arenkov P., Kukhtin A., Gemell A., Voloshchuk S., Chupeeva V., Mirzabekov A. *Anal. Biochem.*, **2000**, 278(2), 123-131.

<sup>87</sup> Afanassiev V., Hanemann V., Wolf S. *Nucleic Acids Res.*, **2000**, 28(12), 1-5.

<sup>88</sup> Zhu H., Klemic J.F., Chang S., Bertone P., Casamayor A., Smith D., Gerstein M., Reed M.A., Snyder M. *Nat. Genet.*, **2000**, 26(3), 283-289.

<sup>89</sup> Ofir K., Berdichevsky Y., Benhar I., Lamed R., Barak Y., Bayer E.A., Morag E. *Proteomics*, **2005**, 5(7), 1806-1814.

<sup>90</sup> Jun Y., Cha T., Guo A., Zhu X.Y. *Biomaterials*, **2004**, 25(17), 3503-3509.

polymer or a random co-polymer<sup>91</sup> like a co-polymer of N,N-dimethylacrylamide, N-acryloyloxysuccinimide, and 3-(trimethoxysilyl)propyl methacrylate.<sup>92</sup> A covalently attach biomolecules on glass slides was firstly reported for the preparation of low-density DNA microarrays and the innovative aspect in this approach relies in the fact that the polymer self-adsorbs onto the glass surface very quickly, simply immersing glass slides in a diluted aqueous solution of the polymer and without time consuming glass pretreatments.

Supramolecular structures are generally considered favorable in retaining protein activity recognition. In addition, due to the decreased reaction rate constant in close vicinity to the surface, recognition of small molecules, such as peptides, is often hindered. The polymeric coating on the surface acts as a linker and moves the bound probe away from the surface resulting in a much faster and efficient interaction with the target. This issue is very important when dealing with small molecule microarrays as peptide arrays.

### 2.2.1 Ligand detection

Current detection strategies for protein microarrays are generally classified as:<sup>93</sup>

1. Label free methods, including mass spectrometry (MS), surface plasmon resonance (SPR) and grating-coupled surface plasmon resonance (GC-SPR), atomic force microscopy (AFM), micro-electromechanical systems (MEMS) cantilevers and quartz-crystal microbalance analysis (QCM);
2. Labelled probe methods, including fluorescence, chemiluminescence, electrochemiluminescence and radioactivity detection.

SPR spectroscopy is a well-know technique to study the kinetics of protein-protein interactions in real-time<sup>57</sup> and it is the leading technology in the field of label-free detection of molecule interactions.<sup>94</sup> In conventional SPR spectroscopy, the capture ligand is immobilized on a thin metal film (gold or silver) and the target is added. The change in the reflection angle of the incident light indicates the amount of target molecule captured on the surface. SPR is a versatile tool but it allows the analysis of only a few channels for a single experiment.

Grating-coupled surface plasmon resonance is a method for the assessment of analyte in a multiplexed format. In GC-SPR the analyte is flowed across specific probes that have been immobilized on a sensor chip. The chip surface is illuminated with p-polarized light that couples to the gold surface's electrons to for a surface plasmon. At a specific angle of incidence, called the GC-SPR angle, the maximum amount of coupling occurs, thus reducing the intensity of

---

<sup>91</sup> Cretich M., Pirri G., Damin F., Solinas I., Chiari M. *Anal. Biochem.*, **2004**, 332(1), 67-74.

<sup>92</sup> Pirri G., Damin F., Chiari M., Bontempi E., Depero L.E. *Anal. Chem.*, **2004**, 76(5), 1352-1358.

<sup>93</sup> Espina V., Woodhouse E.C., Wulfkhule J., Asmussen H.D., Lotta L.A. *J. Immunol. Methods*, **2004**, 290 (1-2), 121-133.

<sup>94</sup> Ramachandran N., Larson D.N., Stark P.R., Hainsworth E., LaBaer J. *FEBS J.*, **2005**, 272(21), 5412-5425.

reflected light. Since regions of the chip can be independently analyzed, this system can assess 400 interactions between analyte and receptor on a single chip.<sup>95</sup>

AFM reveals the change in height of an immobilized protein upon binding with its cognate molecule. It has been introduced for detection at the singular molecular level in protein nanoarrays<sup>96</sup> but is currently widely used for surface characterization in protein microarrays.

Microcantilevers for biosensing are silicon strips attached at one end, with the capture molecule (antibody or protein) bound to the surface. When an analyte binds onto the microcantilever, its bending as a result of surface stress is measured by the deflection of an optical beam or by a change of electrical resistance in a piezoelectric thin film on the cantilever. Alternatively, a change in its mechanical resonant frequency is measured.<sup>97</sup>

All label-free detection methods are promising tools to characterize binding events on surface. They do not require labelling of molecules that may affect protein activity. However, they are generally based on sophisticated equipment not easily available in all types of laboratories.

Labelled probe methods have directly evolved from clinical immunoassays, radioimmunoassays or ELISA protocols where they are widely used. They can be distinguished in:

1. Direct; when a mixture of protein is immobilized and the detection is performed using labelled binding molecules such as antibodies,
2. Indirect; when immobilized antibodies are used as capture ligands and probed by labelled proteins,
3. Sandwich; when a first immobilized antibody acts as a capture agent for the assayed protein which is revealed by a recognition with a secondary labelled antibody.

Labelled probe detection methods rely on fluorescence, chromogenic, chemiluminescence, electrochemiluminescence or radioactive labelling strategies. Detection methods developed for microarrays, due to the miniaturized format, are required to provide high sensitivity and high throughput. The use of fluorescent probes and signal amplification techniques with chromogenic or fluorescent probes usually leads to performances that meet such criteria.

The fluorescence sensing system coupled with antibodies represents the simplest and most versatile strategy for an highly specific and sensitive analysis.<sup>98</sup> Even though the requirement of two antibodies severely constrains the array design, the sandwich method, increases the specificity of antibody arrays and is widely used especially in immunoassays on array format.

---

<sup>95</sup> Unfricht D.W., Colpitts S.L., Fernandez S.M., Lynes M.A. *Proteomics*, **2005**, 5(17), 4432-4442.

<sup>96</sup> Lee K.B., Park S.J., Mirkin C.A., Smith J.C., Mrksich M. *Science*, **2002**, 295(5560), 1702-1705.

<sup>97</sup> Lee J.H., Hwang K.S., Park J., Yoon K.J., Yoon D.S., Kim T.S. *Biosens. Bioelectron.*, **2005**, 20(10), 2157-2162.

<sup>98</sup> Tomizaki K.Y., Usui K., Mihara H. *ChemBiochem*, **2005**, 6(5), 782-799.

Chromogenic detection is based on the use of chromogens. They are substrates for an enzymatic reaction that generates an insoluble coloured product. Commonly used enzymes for chromogenic reactions on microarrays are alkaline phosphatase and horseradish peroxidase. These enzymes act on colourless chemical substrates generating stable coloured product; for example the blue spots generated by 5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium and alkaline phosphatase system has been used for proteomic profiling of cancer cells on nitrocellulose coated slides;<sup>99</sup> tetramethylbenzidine (TMB) and diaminobenzidine (DAB) are widely used substrate for peroxidase.<sup>100</sup>

### 2.2.2 Peptide array

An important difference between peptide arrays and the oligonucleotide arrays is found in the range of applications they can be used to address. DNA-arrays are used almost entirely to profile mRNA populations in cells and rely exclusively on hybridization of sample DNA with the immobilized DNA. Peptide arrays, by contrast, can be used to screen diverse interactions, including the binding of proteins, the action of enzymes, the adhesion of cells, the binding of metals and many others.<sup>101</sup> Further issues that have made peptides arrays more challenging to implement include the expense of preparing large numbers of peptides, the development of chemistries and supports that give good control in presenting immobilized peptides, the prevention of non-specific interactions at the surface and the development of analytical methods to detect activities on the chips.

How can we prepare a peptide array? There are several methods that can be classified by the assembly of peptide on the surface, typical *in situ* peptide synthesis and immobilization of functionalized peptides onto chips,<sup>102</sup> and by the properties of the support.

The *in situ* synthetic approaches have the primary benefit that they avoid conventional synthesis of each of the peptide sequences that are present in the array. The expense and time required to prepare hundreds of peptides, including purification before immobilization, prohibit routine applications. These considerations have motivated the development of the *in situ* approaches, which benefit from using small amounts of reagents and avoiding purification and immobilization of peptides.

The two general approaches for *in situ* peptides synthesis are photolithographic synthesis and SPOT synthesis. The photolithographic approach was first reported by Fodour et al. in 1991.<sup>102</sup> This group translated solid phase methods for peptides synthesis directly on the array support by

---

<sup>99</sup> Knezevic V., Leethanakul C., Bichsel V.E., Worth J.M., Prabhu V.V., Gutkind J.S., Liotta L.A., Munson P.J., Petricoin 3<sup>rd</sup> E.F., Krizman D.B. *Proteomics*, **2001**, 1271-1278.

<sup>100</sup> Paweletz C.P., Charboneau L., Simone N.L., Chen T., Gillespie J.W., Emmert-Buck M.R., Roth M.J., Petricoin 3<sup>rd</sup> E.F., Liotta L.A. *Oncogene*, **2001**, 20(16), 1981-1989.

<sup>101</sup> Dal-Hee M., Mrksich M. *Curr. Opin. Chem. Biol.*, **2004**, 8, 554-558.

<sup>102</sup> Fodour S.P., Read J.L., Pirrung M.C., Stryer L., Lu A.T., Solas D. *Science*, **1991**, 251, 767-773.

developing photolabile protecting groups that would allow synthesis to proceed only at regions of the surface that are illuminated. Since this first report, several improved photolithographic synthesis methods have been reported, including the use of photoresist films and development of maskless array synthesis by Singh-Gasson et al. in 1999.<sup>103</sup> This later development was exciting because the use of a micro-mirror array served as a virtual mask for the photolithographic exposures, leading to lower expense, faster throughput and higher densities in array preparation. The second technique used for *in situ* peptide synthesis is the SPOT method, first reported by Frank in 1992.<sup>104</sup> In the SPOT synthesis, peptides are synthesized by sequential spotting of microliter volumes of activated amino acids to a porous membrane, like the same chemistry used in conventional solid-phase synthesis. SPOT synthesis is one player on the array technology field and fulfills the main characteristics of array technology, at the same time it is able to make peptides as capture molecules. Those are synthesized in a stepwise and parallel manner directly on cellulose support; the basic principle of SPOT technology involves the positionally addressed delivery of small volumes of activated amino acid directly onto a cellulose membrane sheet. The areas wetted by the resulting droplets can be considered microreactors, provided that a non-volatile solvent system is used; the physical properties of the membrane surface, the solvent system and the applied volumes define the size of the resulting spots. Both immobilised and soluble peptides can be prepared by SPOT technology. Generally small-scaled spots are typically used for an immobilised synthesis; in contrast, large-scaled spots are used for preparation of soluble peptides.<sup>105</sup>

The alternative method to *in situ* peptide synthesis relies on immobilization of peptide that have been previously synthesized by chemical or recombinant methods. This strategy is important when small number of peptides will be used to prepare large numbers of identically arrays, situation that can justify the expense of peptide synthesis and purification.

In a way, the peptide microarray is nothing more than a library of immobilized peptides that are spatially addressable, highly miniaturized and therefore requiring only a minimal amount of analytes for analysis. Using various screening techniques, we can identify peptide or peptides with a specific biological, chemical or physical property and isolate them for structural determination.

---

<sup>103</sup> McGall G., Labadie J., Brock P., Wallraff G., Nguyen T., Hinsberg W. *Proc. Natl. Acad. Sci. USA*, **1996**, *93*, 13555-13560.

<sup>104</sup> a) Frank R., Doring R. *Tetrahedron*, **1998**, *44*, 6031-6040. b) Frank R. *Tetrahedron*, **1992**, *48*, 9217-9232

<sup>105</sup> Frank R. *J. Immunol. Methods*, **2002**, *267*, 13-26

## Chapter 3: Target molecules

The development of new biorecognition elements for organic molecules is the aim of this project. In our case, target analyte were represented by three molecules: caffeine and theophylline, that are two small organic xanthines and own similar chemical properties, and palytoxin that is a dangerous algal toxin with very different properties compared with the first two molecules.

Palytoxin is an very large molecule and its structure allow to identify different strategic functional groups useful for chemical modifications or conjugations at proteins; in contrast, caffeine and theophylline are small organic molecules with less number of chemical groups available to a chemical modification.

### 3.1 Palytoxin

Palytoxin is one of the most potent and complex marine toxins. Though first isolated from the zoanthid *Palythoa toxica*, its biogenetic origin was questioned because of marked seasonal and regional variations, but the question was never pursued. Sporadic occurrence of palytoxin in an alga, crabs and herbivorous fish also suggested that the toxin might be produced by a microorganism and trasmitted to the marine food chain, as is the case with ciguatera, where the epiphytic dinoflagellate *Gambierdiscus toxicus* produces ciguatoxins and maitotoxin that are isolated from fish.<sup>106</sup>

The letal marine toxin palytoxin, first extracted from the polyps of soft-bodies zoanths of the genus *Palythoa* is a potentially useful tool for elucidating the molecular reactions by which the Na/K-ATPase stoichiometrically transports Na and K ions in opposite directions across the cell surface membrane: Palytoxin was found to depolarize mammalian cells by causing small-conductance relatively nonselective cation channels to appear in their surface membranes.<sup>107</sup>

Although the palytoxin-containing zoanthid organisms are located in tropical waters, the dinoflagellate of the genus *Ostreopsis* have a worldwide distribution and toxin producing species have been described not only in tropical and subtropical waters, but also in the Mediterranean Sea. Humans show symptoms of intoxication by palytoxin following ingestion of contaminated seafood or exposue to seawater aerosol in bathing areas. Oral poisoning can induce among other symptom vomiting, diarrhea, paresthesia of the extremities, myalgia, respiratory distress and death; exposure to the toxin by aerosol induces rhinorrhea, cough, fever and broncoconstriction.<sup>108</sup>

---

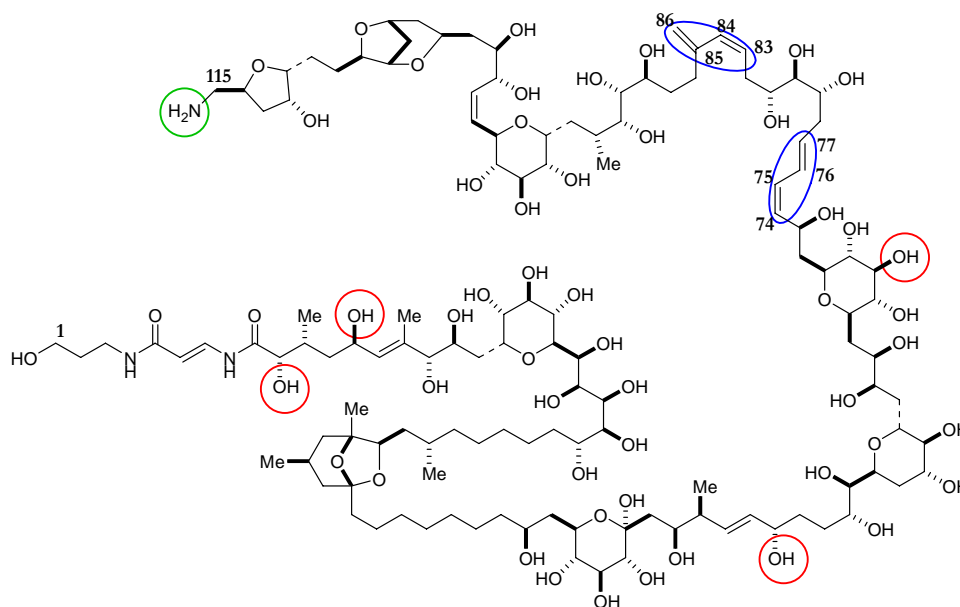
<sup>106</sup> Wang D.Z. *Marine Drugs*, **2009**, 6(2), 349-371.

<sup>107</sup> Horisberger J.D. *Physiology*, **2004**, 19(6), 377-387.

<sup>108</sup> Cagide E., Yasumoto T., Botana L.M. *Cell. Physiol. Biochem.*, **2009**, 23(4-6), 431-440.

Palytoxin is regarded as one of the most potent toxins so far known; the LD<sub>50s</sub> 24 hours after intravenous injection vary from 0.025 µg/kg in rabbits and about the same in dogs, to 0.45 µg/kg in mice, with monkeys, rats and guinea pigs around 0.9 µg/kg.

Chemically, palytoxin is a very large and complicated structure (Figure 3.1):



**Figure 3.1:** palytoxin structure. In red hydroxy groups, in blue diene motifs, in green terminal amino group.

It is a hydrophilic molecule, with 115 carbons atoms and 64 of these stereogenic centers; it is similar to a carbohydrate structure because of the presence of hydroxy groups (42 groups). The molecular weight of the toxin is 2678.5 g/mol; most of the functional groups presents in the molecule are secondary hydroxy groups; unsaturated and hydrophobic moieties are found in diene motifs: the first is found at C<sub>74</sub>-C<sub>77</sub> position and the second between C<sub>83</sub> and C<sub>86</sub>. The most reactive group is the primary amino group at C<sub>115</sub>. This nucleophilic functionality could be used to modify the original structure of palytoxin in order to generate new different conjugate of this molecule with biomolecules as proteins and enzymes, or with other organic molecules.

The structure of palitoxyn has been elucidated by two different groups separately with NMR nuclear Overhauser effects (NOE) in order to define the absolute stereochemistry of all 64 chiral centers;<sup>109</sup> total synthesis of palytoxin have been achieved by Kishi et al. in about 20 years.<sup>110</sup> A convergent strategy was adopted by Kishi, after the identification of eight key building blocks. Coupling reactions between all the key building blocks allow to obtain final complete palytoxin structure; generally, the coupling reactions involve the carbon-carbon bond formation and the most used reactions are, for examples, Wittig reaction for the C<sub>22</sub>-C<sub>23</sub> and C<sub>75</sub>-C<sub>76</sub> bond followed by hydrogenation, Ni(II)/Cr(II) reaction to obtain only *trans* stereoisomer for the carbon-carbon

<sup>109</sup> Kan Y., Uemura D., Hirata Y., Iwashita T. *Tetrahedron Lett.*, **2001**, 42(18), 3197-3202.

<sup>110</sup> Kishi Y., Hirata Y., Uemura D. *J. Am. Chem. Soc.*, **1982**, 104(25), 3776-3779.

bond contiguous to a double bond and Suzuki coupling for the diene structure formation (Figure 3.2).<sup>111</sup>

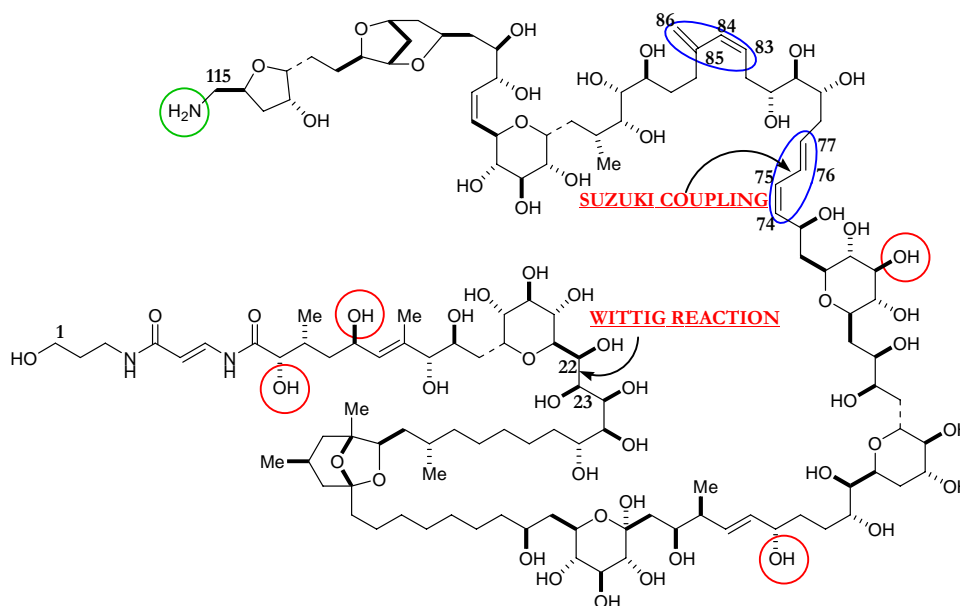


Figure 3.2

### 3.2 Theophylline and caffeine

The xanthine alkaloids caffeine (1) and theophylline (2) (Figure 3.3) are closely related compounds found in a variety of plants indigenous to several continents. They are not true alkaloids, as they are only weakly basic, and possess some acidic properties. Their structure include a purine ring, as found in the adenine and guanine nucleotides. The occurrence of N-methyl groups determines the specific pharmacological profile of each compound. The xanthines are not of great current importance as medicinals, but their history and cultural significance makes them of continuing interest.<sup>112</sup>

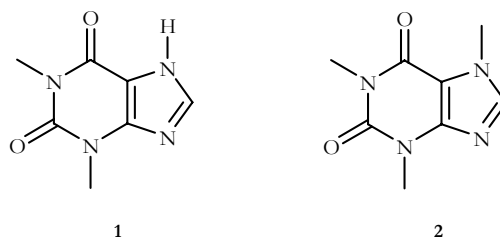


Figure 3.3: caffeine and theophylline structure.

Theophylline, also known as dimethylxanthine, is a methylxanthine drug used in therapy for respiratory diseases such as asthma under a variety of brand names. It is present in tea, although in

<sup>111</sup> Uenishi J. C., Beau J.M., Armstrong R.W., Kishi Y. *J. Am. Chem. Soc.*, **1987**, *109*, 4756-4758.

<sup>112</sup> Franco R. *Atencion Primaria*, **2009**, *41(10)*, 578-581.

trace amounts (about 1 mg/L), significantly less than therapeutic doses; it is found also in cocoa beans and in Criollo cocoa beans in amounts as high as 3.7 mg/g.<sup>113</sup>

Theophylline, first isolated in 1888 from tea extract, was discovered much later than the other xanthines, presumably because its presence was masked by the relatively large amount of caffeine contained in tea.

Theophylline and its bronchodilator action has been of greatest importance in modern medicine. This xanthine in fact relaxes airway smooth muscle, thus exerting a therapeutic effect in asthma. The mechanism of action is based on a non-selective inhibition of cyclic nucleotide phosphodiesterases (PDEs), preventing breakdown of cyclic AMP and cyclic GMP.<sup>114</sup>

Another important action is theophylline's competitive antagonism at adenosine receptors which prevents bronchoconstriction in asthmatic patients. Theophylline also inhibits synthesis and secretion of inflammatory mediators from various type of cells; it has been proposed, but not yet proven, that this anti-inflammatory effect may be relevant to the drug's therapeutic action. The therapeutic range for theophylline bronchodilator action has been established at 10-20 mg/L. Above this value, toxic effects may occur, including tachycardia, severe restlessness, agitation, emesis and occasionally seizures; desired plasma concentrations are obtained by careful dose adjustment and periodic blood testing.

Caffeine is a bitter, white crystalline xanthine alkaloid that acts as a psychoactive stimulant drug. Caffeine was discovered by a German chemist in 1819 and he coined the term "kaffein", a chemical compound in coffee. Caffeine molecule is found in varying quantities in the beans, leaves and fruit of some plants, where it acts as a natural pesticide that paralyzes and kill certain insects feeding on the plants. It is most commonly consumed by humans in infusions extracted from the cherries of the coffee plant and the leaves of the tea bush, as well as from various food and drinks containing products derived from the yerba mate, guaraná berries and Yaupon Holly.<sup>115</sup>

Caffeine's most prominent pharmacologic action is to stimulate the central nervous system (CNS), increasing arousal and vigilance, reducing fatigue and decreasing motorial reaction time for some tasks. It may interfere with sleep and it tends to decrease cerebral blood flow. As to cardiovascular effects, the drug induces a 5-10 mm increase in blood pressure via peripheral vasoconstriction, a slight decrease in heart rate and to cause the systemic release of epinephrine, norepinephrine and renin. Gastric secretion of acid and pepsin is also stimulated by caffeine and may contribute to acid reflux symptoms by decreasing the lower esophageal sphincter pressure. Other effects include caffeine's ability to increase metabolic rate and circulating concentrations of

---

<sup>113</sup> Montuschi P., Pagliari G., Fuso L. *Ther. Adv. Resp. Dis.*, **2009**, 3(4), 175-191.

<sup>114</sup> Simpson J.L., Gibson P.G. *Pharmacol. Ther.*, **2009**, 124(1), 86-95.

<sup>115</sup> Myslobodsky M. *Curr. Drug Target*, **2009**, 10(10), 1009-1020.

free fatty acids, cortisol and glucose.<sup>116</sup> Caffeine has long been used as an adjuvant in analgesic products, where it is believed to enhance gastrointestinal absorption of the analgesic and potentiate their activity. It was a component of an analgesic-antipyretic tablet containing aspirin, phenacetin and caffeine first marketed by Burroughs Wellcome as Empirin Compound; caffeine is also used for the treatment of apnea in newborns. Both caffeine and theophylline are comparably effective in regularizing the breathing, but caffeine is easier to administer, has a wider therapeutic index, and produces fewer peripheral effects.<sup>117</sup>

The interest of our research group in the development of primary transducer towards xanthines dates back to year 2005 as a consequence of a joint research plan between the Dept. of Chemistry of our University and Biosensor Technologies Ltd, a small enterprise based in Trieste, whose core activity was represented by the development of peptide libraries towards molecules of interest for the coffee and flavour industry. More recently a similar collaboration was started with T&B associates, another SME, and a joint research project coordinated by our Dept. and T&B, the Dept. of Life Sciences, the group of Statistical Mechanics and Biophysics at SISSA, the group of Virology at ICEGB.

---

<sup>116</sup> Geleijnse J.M. *Vasc Health Risk Manag*, **2008**, *4*(5), 963-970.

<sup>117</sup> Jankiewicz K., Czuczwar S.J. *Oncogene*, **2007**, *26*(56), 7816-7824.

# AIM OF RESEARCH

Our research group has been recently involved in a research project, granted by the regional government of “Regione Autonoma Friuli-Venezia Giulia” focused on the development of a biosensor for the detection of palytoxin in sea water and sea food.

Due to the affluence of *Ostreopsis* blooms in the Trieste gulf area, monitoring of such toxin has become of high interest for the seafood industry.

Among the different strategies outlined within the project to obtain primary transducers (including mono- and polyclonal antibodies towards palytoxin) and the development of libraries of antibody fragments. We have also planned to obtain peptides that could be interact with the whole molecule, or with fragment of the toxin as well.

The interest of our research group is also oriented in the development of primary transducer towards xanthines, as a consequence of a joint research plan between the Dept. of Chemistry of our University and Biosensor Technologies Ltd, a small enterprise based in Trieste, whose core activity was represented by the development of peptide libraries towards molecules of interest for the coffee and flavour industry. More recently a similar collaboration was started with T&B associates, another SME, and a joint research project coordinated by our Dept. and T&B, the Dept. of Life Sciences at the University of Trieste, the group of Statistical Mechanics and Biophysics at SISSA, Trieste, the group of Virology at ICEGB, Trieste.

# RESULTS AND DISCUSSION

## Chapter 4: Modified target molecules

### 4.1 Caffeine and theophylline

In order to fish out binders towards small organic molecules from biological libraries of receptors, several modified versions of the target molecules are required, that could allow immobilization and eventually bioconjugation.

Immobilization is thus one of the most important step in the identification and isolation of receptors or bioreceptors having affinity to the target analyte. The immobilization of small molecules must allow the full exposure of all the functional groups of the organic compound in order to allow optimal interactions with the receptor macromolecule. In order to avoid masking effect between target molecule and the solid surface used for the immobilization,<sup>118</sup> generally a spacer is introduced (Scheme 4.1):



Scheme 4.1

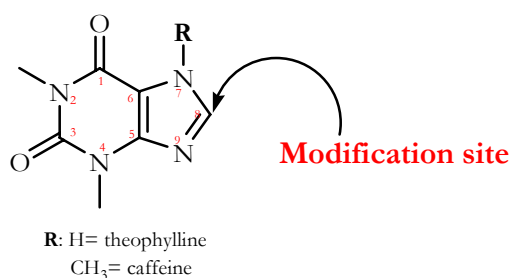
Linker spacer in this case plays a fundamental role for the selection; it must not interfere with the binding between target molecule and receptor macromolecules but only avoid non-specific interactions between solid support and reactive groups of target compound. This conditions is very difficult to obtain, since even very simple linker moieties, as a hydrocarbon chain, are able to interact with some members of the receptor library, via, for example, non specific hydrophobic interactions. Many often this leads to the selection of a rather high number of false positives or to the selection of true binders towards the target, but showing high affinity only in the presence of the linker and thus unable to be used as competition elements between the free and immobilized target<sup>119</sup>. In order to minimize the occurrence of such false positivity phenomena, we have planned to introduce diversity inside the spacer structure by changing its length, its polarity and the final reactive group as well.

<sup>118</sup> a) Lin C., Shen J., Chan W., Wang H. *J. Med. Biol. Engin.*, **2009**, 29(6), 276-283. b) Barbas A.S., White R.R. *Curr. Opin. Investig. Drugs*, **2009**, 10(6), 572-578. c) Bailon P., Won C.Y. *Exp. Opin. Drug Del.*, **2009**, 6(1), 1-16.

<sup>119</sup> Christensen D.J., Gottlin E.B., Hamilton P.T. *Drug. Disc. Tod.*, **2001**, 6(14), 721-727.

The spacer must be introduced on the analyte at carefully selected positions, in order to avoid the modification of reactive groups present in the original structure. When possible, linker is placed at an unreactive site of the molecule, leaving free the other reactive groups.

The first step of this work was thus represented by the synthesis of different typologies of modified caffeine and theophylline in order to introduce in their structure a spacer linker. The structure of caffeine and theophylline does not allow easy modification without loss of potential important interaction groups of this small aromatic compound; in order to avoid chemical modification of carbonyl or nitrogen positions, the only point available for a chemical modification is represented by position 8 (Scheme 4.2):



**Scheme 4.2:** theophylline and caffeine structure with modification site

According to this, a first set of two different modified molecules **8**, **9** (Figure 4.1) has been synthesized, having caffeine and theophylline linked to biotin as the final reactive group<sup>120</sup> (a high affinity streptavidin binder). The biotin-avidin complex has been used in numerous bioanalytical and biotechnological applications because of its high affinity constant, which leads to stable biomolecular construct. The stability of the avidin-biotin complex was first reported by Launer and Fraenkel-Conrat based on dialysis equilibrium measurements and they found that the dissociation constant is  $2.4 * 10^{-10}$  M at 25°C.<sup>121</sup>

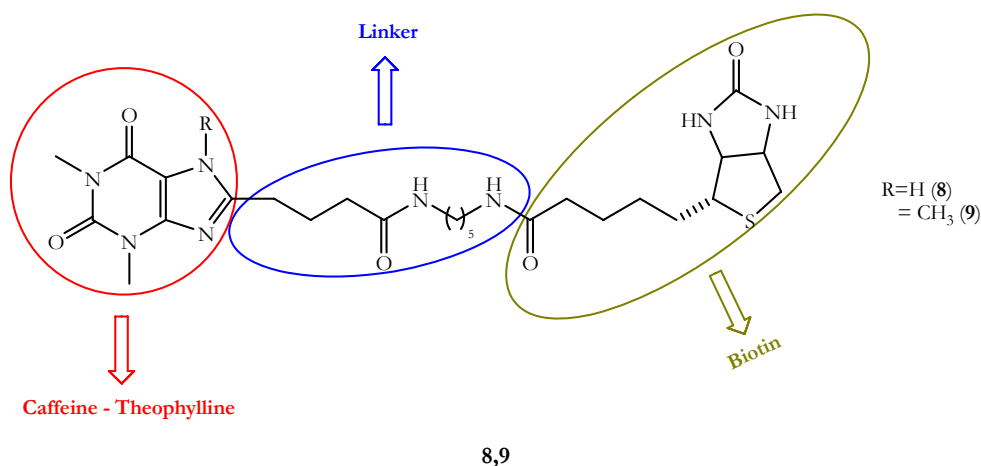
This allows target-biotin conjugates to be immobilized to a solid surface previously coated with streptavidin.

The overall structure of the first set of conjugates is reported in Figure 4.1. A relatively highly hydrophobic ten atoms linker, containing two hydrogen-bond capable group was introduced. Such linker has been used in the past during the production of a monoclonal antibody towards caffeine with good results<sup>122</sup>.

<sup>120</sup> Oh C., Kim J., Kearns B., Cheng A., Dobashi T. *Clin. Chim. Acta*, **1993**, 218(1), 59-71.

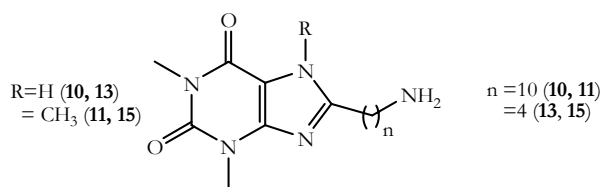
<sup>121</sup> Launer L.J., Fraenkel-Conrat H. *FASEB*, **1994**, 8(6), 452-453.

<sup>122</sup> Del Carlo M., Mascini M. *Field Anal. Chem. Tech.*, **1999**, 3(3), 179-184.



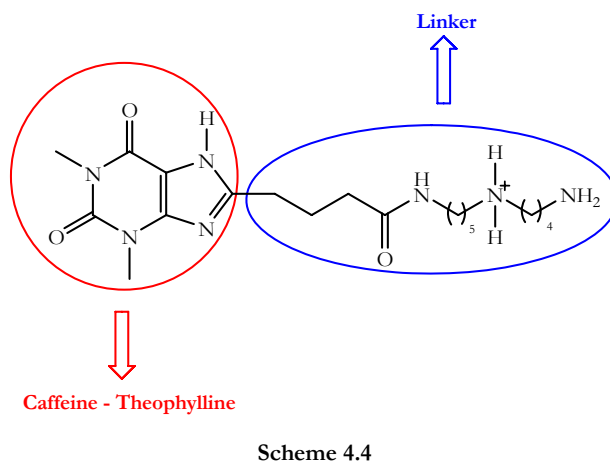
**Figure 4.1:** structure of caffeine/theophylline-biotin conjugates.

We have also considered a set of amino-modified xanthines **10, 11, 13, 15** (Scheme 4.3):



This different linker is shorter than the other in the biotin derivatives and its amino terminal allows both non covalent immobilization to several commercially available dedicated matrices such as Amino-Nunc<sup>®123</sup>, and immobilization by acetylation to carboxy-functionalized surfaces or carrier proteins.

Finally we have also planned a set of modified xanthines bearing a polar and charged linker **20** (Scheme 4.4) in order to introduce a high degree of polarity diversity among the available linkers.

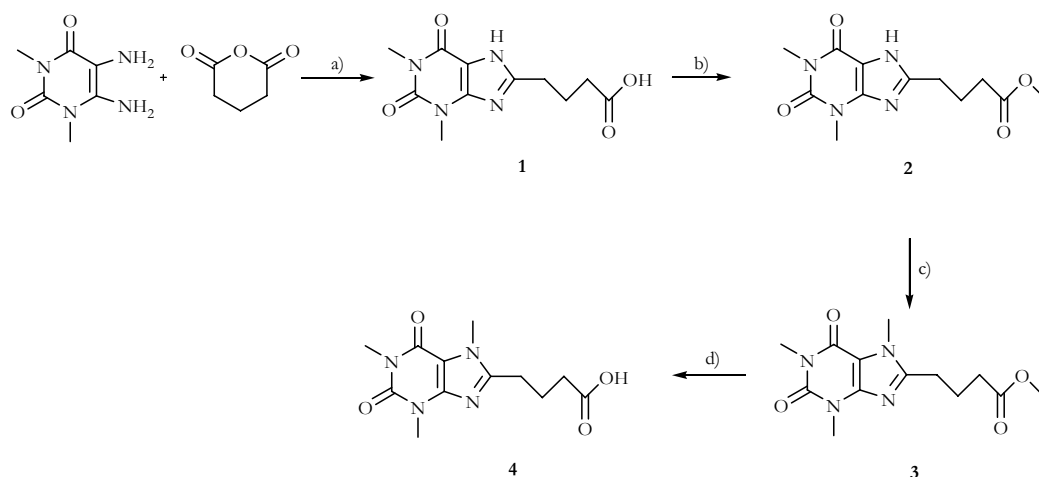


<sup>123</sup> Robin K., Benisty H. *Biosens. Bioelect.*, **2009**, 24(6), 1585-1591.

#### 4.1.1 The synthesis of caffeine/theophylline-biotin conjugates (8, 9).

When synthesizing a modified version of the target molecule, one could start from the original structure and introduce a bond to the linker by direct modification. However, in our case this is not easy, due to the scarce reactivity of xanthines at position 8. We have thus decided to synthesize *de novo* a xanthine scaffold with a built-in linker at position 8.

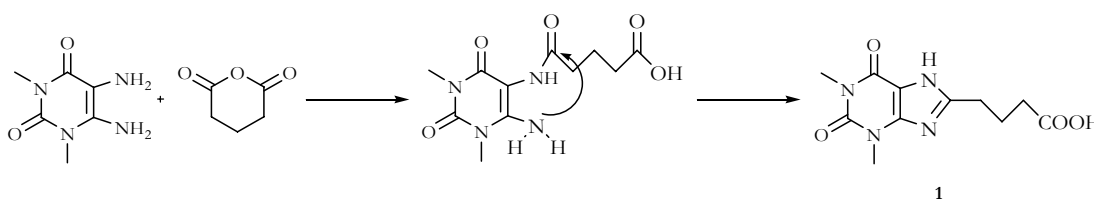
The first synthesis steps for theophylline and caffeine modification are thus outlined in Scheme 4.5:



**Scheme 4.5**

a)NaOH/DMF/75°C; b)SOCl<sub>2</sub>/MeOH/reflux; c)CH<sub>3</sub>I/NaH/THF; d)LiOH/H<sub>2</sub>O-THF

The synthesis started from 4,5-diamino-1,3-dimethyl-uracil hydrate that was reacted with glutaric anhydride under basic conditions;<sup>124</sup> in this step glutaric anhydride undergoes a condensation with one of the uracil amino groups, in the second step sodium hydroxide solution was added in order to allow the cyclization and theophylline-8-butanoic acid **1** was obtained in 50% yield.



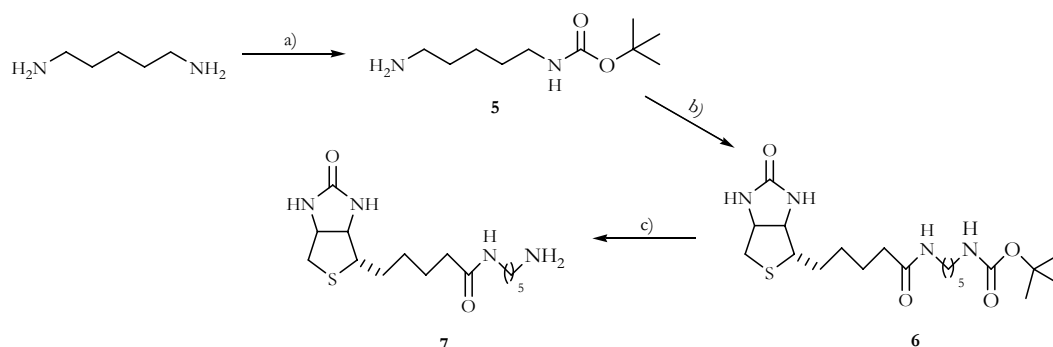
**Scheme 4.5b**

The caffeine derivative was obtained from this theophylline derivative by methylation. The acid **1** was first protected as methyl ester using thionyl chloride in methanol to give compound **2** in excellent 98% yield. Alkylation of the xanthine nitrogen at position 9 was carried on under anhydrous conditions with methyl iodide and sodium hydride as base in tetrahydrofuran; compound **3** was obtained in 70% yield. Finally, caffeine-8-butanoic acid **4** was obtained from

<sup>124</sup> Kim H., Ji X., Melman N., Olah M., Jacobson A. J. *Med. Chem.*, **1994**, *37*, 3373-3382.

ester **3** by base hydrolysis of the methyl ester moiety with lithium hydroxide aqueous solution using a mixture of water/THF=50/50 as solvent (65% yield).

At the same time, the synthesis of a biotin-cadaverine derivative was carried out (Scheme 4.6):

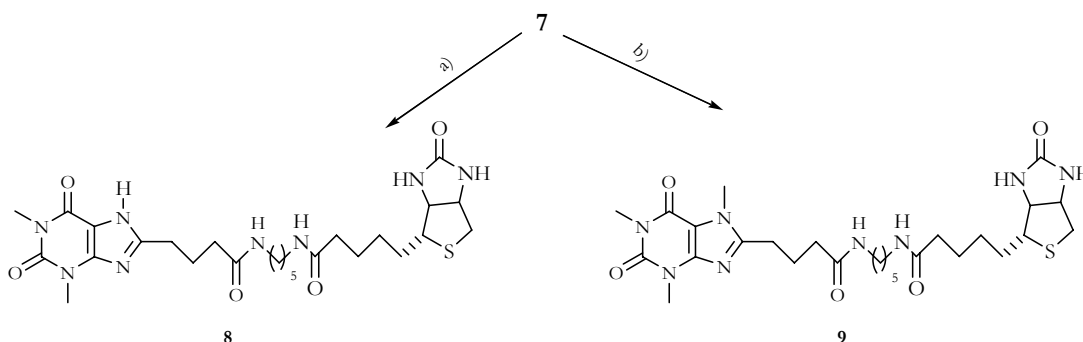


**Scheme 4.6**

a)  $\text{Boc}_2\text{O}/\text{DCM}/0^\circ\text{C}$ ; b) Biotin/HOBT/NMM/EDC-Cl/ $\text{CH}_3\text{CN}$ ; c) TFA/DCM

1,5-diaminopropane (cadaverine) was monoprotected at one of its two amino groups as *t*-butoxycarbonyl carbamate (Boc) using Boc anhydride and dichloromethane as solvent at  $0^\circ\text{C}$ ; compound **5** was obtained in 45% yield and then used directly for the synthesis of amide **6**. The reaction of monoprotected cadaverine with biotin under anhydrous conditions using 1-hydroxybenzotriazole (HOBT) to optimize the condensation, N-methylmorpholine (NMM) as base and ethyl-carbodiimide (EDC-Cl) as activator, gave compound **6** in 64% yield. Removal of Boc-protection with 50% trifluoroacetic acid (TFA)/DCM gave the final product **7** in 98% yield.

Compound **7** was finally reacted with theophylline-8-butanoic acid (**1**) or caffeine-8-butanoic acid (**4**) under the same conditions previously described for the coupling of monoprotected cadaverine with biotin. Final conjugates **8** and **9** were obtained in 54% and 65% yields respectively. (Scheme 4.7):



**Scheme 4.7**

a) **1**/HOBT/NMM/EDC-Cl/DMF; b) **4**/HOBT/NMM/EDC-Cl/DMF

#### 4.1.2 The synthesis of caffeine/theophylline amino derivatives **10**, **11**, **13**, **15**

Another important modification for caffeine and theophylline was obtained with the introduction of a free amino group at the end of linker. The linker is shorter than that used in the biotin

conjugates and this should allow a reduction of linker-binding false positives occurrence by sequential multiple selection rounds. The longer amino linker of compounds **10** and **11** (Figure 4.2) was simply obtained from the coupling of **1** and **4** with monoprotected cadaverine **5** under the same conditions previous described, with HOBT, NMM and EDC-Cl. Removal of Boc protection from the free amino group gave the final products with excellent total yields (80%) (Figure 4.2).

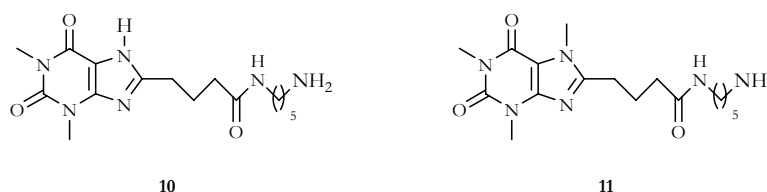
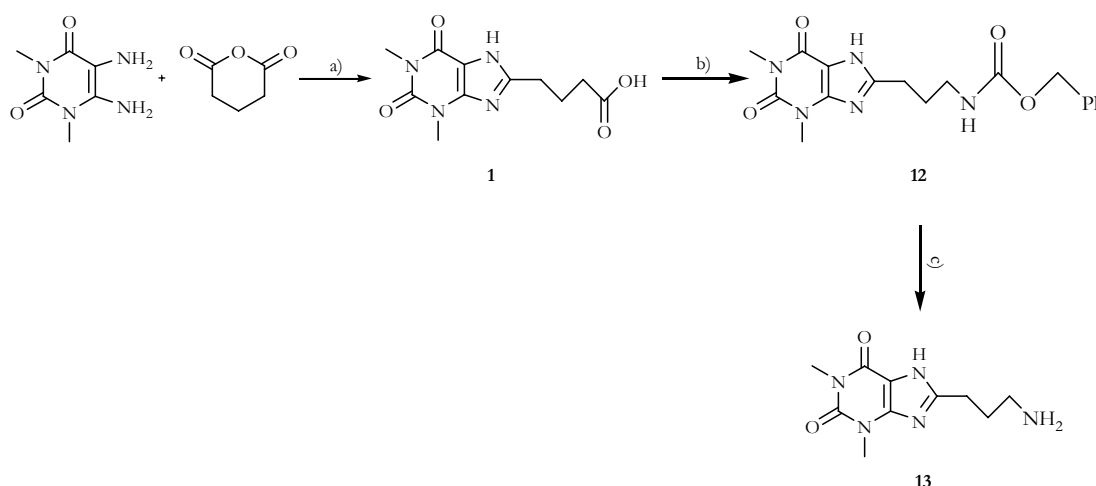


Figure 4.2

Beside these conjugates, the shorter structures **13** and **15** have been synthesized (Scheme 4.8):

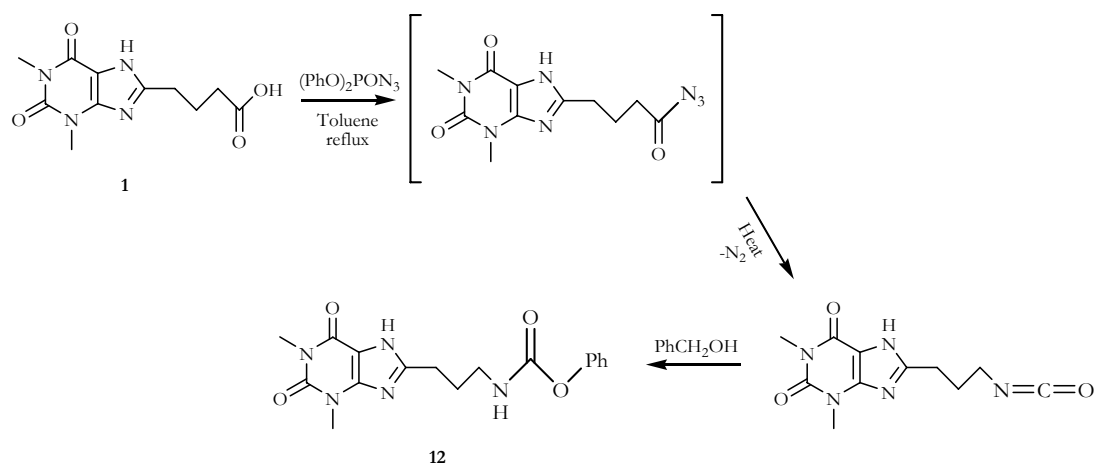


Scheme 4.8

a) NaOH/DMF/75°C; b) (PhO)<sub>2</sub>PON<sub>3</sub>/NEt<sub>3</sub>/benzyl alcohol/toluene; c) Pd(C)/H<sub>2</sub>/MeOH

This synthesis starts also from theophylline-8-butanoic acid (**1**) but after having obtained the xanthine system, in the second step Curtius rearrangement<sup>125</sup> was used in order to convert the carboxy group into an amino group, protected as benzylloxycarbonyl carbamates (Cbz). Theophylline-8-butanoic acid was thus dissolved in dry toluene and under anhydrous conditions and diphenylphosphoryl azide was added. When the solution was heated, the acid was quickly converted in to an acyl azide intermediate, that rearranges to isocyanate giving finally, in the presence of benzyl alcohol, the Cbz-protected amine (Scheme 4.9):

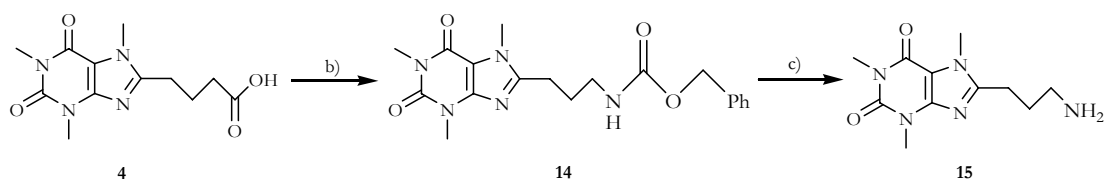
<sup>125</sup> a) Curtius T.H. *Chem. Ber.*, **1890**, *23*, 3023. b) Ma B., Lee W.C. *Tetrahedron*, **2010**, *51*(2), 385-386.



**Scheme 4.9:** Curtius rearrangement mechanism

The N-Cbz protected derivative **13** was thus obtained from theophylline 8-butanoic acid in 56% yield.

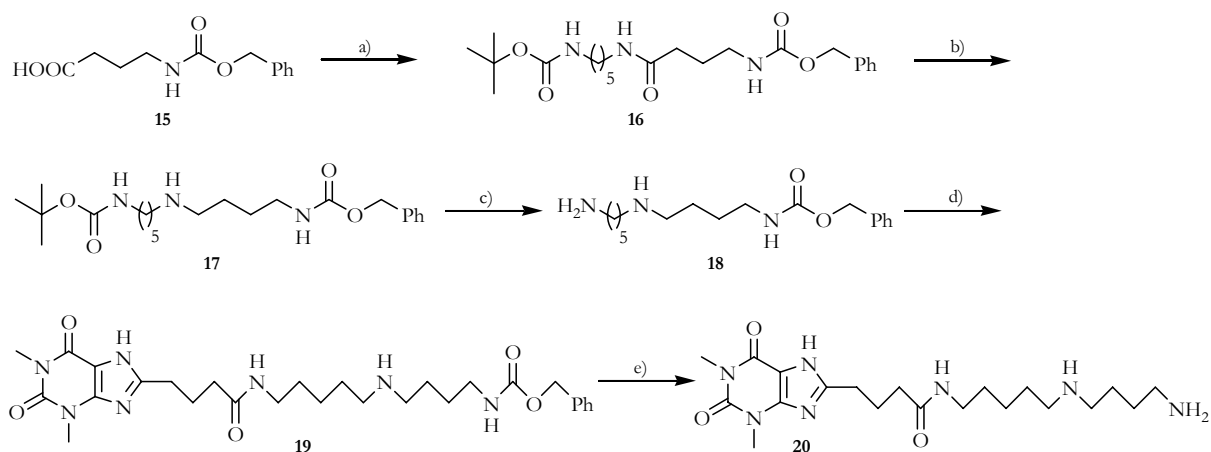
In a similar way the amino linker caffeine derivative **15** was synthesized starting from caffeine-8-butanoic acid **4** (Scheme 4.10):



**Scheme 4.10**

### 4.1.3 Synthesis of polar-linker xanthines

In order to increase the polarity of linker, a new family of conjugates was generated in which a central amino group was introduced; Depending on the environmental pH, such a derivative can be used during the library selection either as a cationic molecule or as a hydrogen-bond acceptor highly polar selection tool. As in the previous modification, at the end of linker a free amino group was introduced to allow binding with suitable groups on the solid surface used for the immobilization (Scheme 4.11):



**Scheme 4.11**

a) Compound **5**/HOBT/NMM/EDC-Cl/CH<sub>3</sub>CN; b) BH<sub>3</sub>-DMS/THF; c) TFA/CH<sub>2</sub>Cl<sub>2</sub>; d) Compound **1**/HOBT/NMM/EDC-Cl/CH<sub>3</sub>CN; e) Pd(C)/H<sub>2</sub>/MeOH

$\gamma$ -aminobutyric acid was protected as Cbz at the free amino group (**15**) and it was reacted with compound **5** in the normal conditions for the aminoacids coupling using acetonitrile as solvent to give the corresponding amide **16**. Compound **16** was reduced with BH<sub>3</sub>-DMS in THF as solvent; in these conditions the reduction takes place only on the amide carbonyl group and not on the Boc or Cbz carbonyl group<sup>126</sup>. Removing Boc protection with TFA in dichloromethane the free amino molecule **18** was obtained and after reaction with theophylline-8-butanoic acid (**1**) and removing of the Cbz protection, final conjugate **20** could be used directly for the immobilization (66%).

#### 4.1.4 Xanthine-protein conjugates.

The target molecules caffeine and theophylline have also been labeled with two proteins: Horseradish peroxidase (HRP) and Albumin from chicken egg white (OVA). These conjugates have been synthesized only as reagents for the ELISA tests and not for the immobilization of organic molecules on solid surface, as the previous cases.

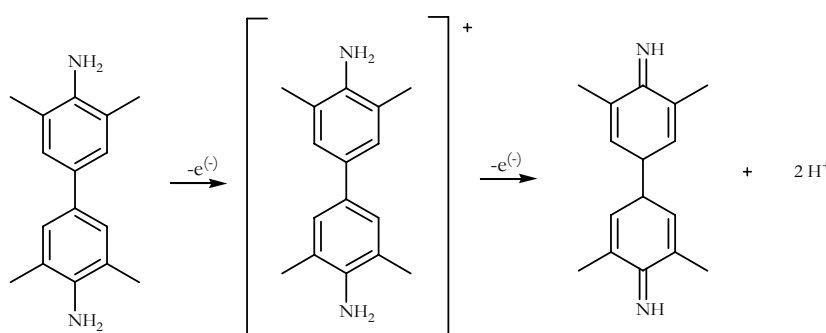
Enzyme-linked immunosorbent assay, also called ELISA, is a biochemical technique used mainly in immunology to detect the presence of an antibody or an antigen in a sample. The ELISA has been used as a diagnostic tool in medicine as well as quality control check in various industries. In simple terms, in ELISA an unknown amount of antigen or other molecules are affixed to a surface, and the a specific receptor is incubated over to the surface so that it can bind to the analyte. The receptor is linked to an enzyme, and in the final step a substrate is added that the enzyme can convert to give some detectable signal.

<sup>126</sup> Felluga F., Gombac V., Pitacco G., Valentin E. *Tetrahedron: Asymmetry*, **2004**, *15*, 3323-3327.

In many cases the enzyme used as label is the Horseradish Peroxidase (HRP) that can react with an appropriate substrate namely 3,3',5,5'-tetramethylbenzidine (TMB) causing a change in color, which is used as a signal.

Horseradish peroxidase is a 44 kDa glycoprotein with four lysine residue suitable for conjugation to a labelled molecules. It produces a colour, fluorimetric or luminescent derivative of the labelled molecule allowing it to be detected and quantified.

HRP is a oxidoreductase that transfers hydrogen from donator molecule to hydrogen peroxide. In ELISA test in acid conditions, TMB was oxidized from HRP to obtain a final blu radical compound that can be converted in a yellow compound ( $\lambda_{\text{max}}=450 \text{ nm}$ ) with high absorbance, by pH decreasing<sup>127</sup> (Scheme 4.12):

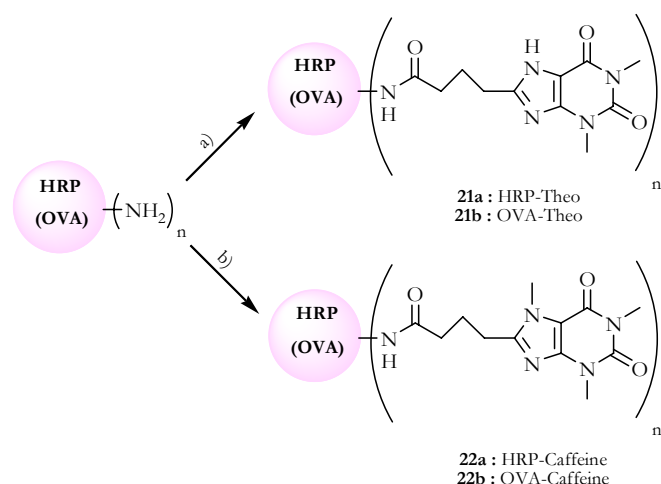


Scheme 4.12

Albumin from chicken egg white (OVA) makes up 75% of the total protein in egg whites. The protein contains 20 lysine residues and 14 aspartic acid and 33 glutamic acid groups. This gives a total of 20  $\epsilon$ -amines, 1 N-terminal amine, 47 side-chain carboxylates, and 1 C-terminal carboxylate for conjugation reactions. The majority of acid groups gives the protein a pI of 4.63. OVA is sensitive to temperature (above 56°C), electric fields and vigorous mixing. Care should be taken in handling the protein to prevent denaturation and subsequent precipitation.

The conjugation of theophylline and caffeine to HRP and OVA was carried out starting from free acids **1** and **4** which were used to acylate the primary amino groups of lysines present on the protein (Scheme 4.13):

<sup>127</sup> Tijssen P. *Pratice and Theory of Enzyme Immunoassay*, 1988, cap. 12, Elsevier Amsterdam.



**Scheme 4.13**

a) Compound **1**/EDC-Cl/MES(pH=4.5)-DMSO/MES(pH=4.5)-DMF; b) Compound **4**/EDC-Cl/MES(pH=4.5)-DMSO/MES(pH=4.5)-DMF

Both the conjugates were purified by dialysis against phosphate buffer for HRP conjugates or water for OVA, and the degree of protein modification was estimated by UV spectroscopy.

The modification index for the two conjugates was calculated in a spectrophotometrical way at  $\lambda=401.4$  nm for HRP;  $\lambda=283.2$  for OVA and  $\lambda=275.4$  for the two xanthines.

By this way it was possible to estimate the concentration of protein and hapten in solution after dialysis according to the Lambert-Beer law (Scheme 4.14):

1)HRP-theophylline conjugate:  $[\text{HRP}]_f=2.27 \mu\text{M}$ ;  $[\text{theophylline}]_f=26.2 \mu\text{M}$

$$\text{IM}=11$$

2)HRP-caffeine:  $[\text{HRP}]_f=6.72 \mu\text{M}$ ;  $[\text{caffeine}]_f=228 \mu\text{M}$

$$\text{IM}=33$$

3)OVA-theophylline :  $[\text{OVA}]_f=19.5 \mu\text{M}$ ;  $[\text{theophylline}]_f=916.3 \mu\text{M}$

$$\text{IM}=47$$

**Scheme 4.14**

Assuming that after dialysis all the hapten presence is bound to the protein, and the ratio  $[\text{hapten}]/[\text{protein}]$  represent the modification index (IM) for the two conjugates, as the number of molecules of hapten bound per molecule of protein.

## 4.2 A menthone-biotin conjugate

Caffeine and theophylline are closely similar aromatic molecules, and the only difference is given by the nitrogen at 7 position that is methylated in caffeine and is not methylated in the theophylline.

Beside these xanthines, another small organic molecule, menthone, has been chosen with different characteristics in order to evaluate the selectivity of the anti-xanthine selected peptides. Menthone is a small non aromatic terpene and it has two stereocenter (Figure 4.3).

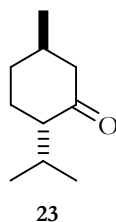


Figure 4.3

Biotinylation of menthone is necessary in order to immobilize this molecule on solid surface; thus a the menthone-biotin conjugate **26** was synthesized, within a similar frame to the theophylline and caffeine-biotin conjugates (Figure 4.4):

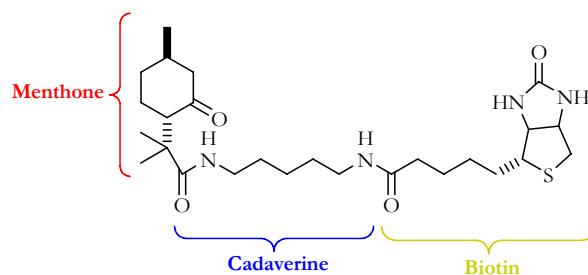
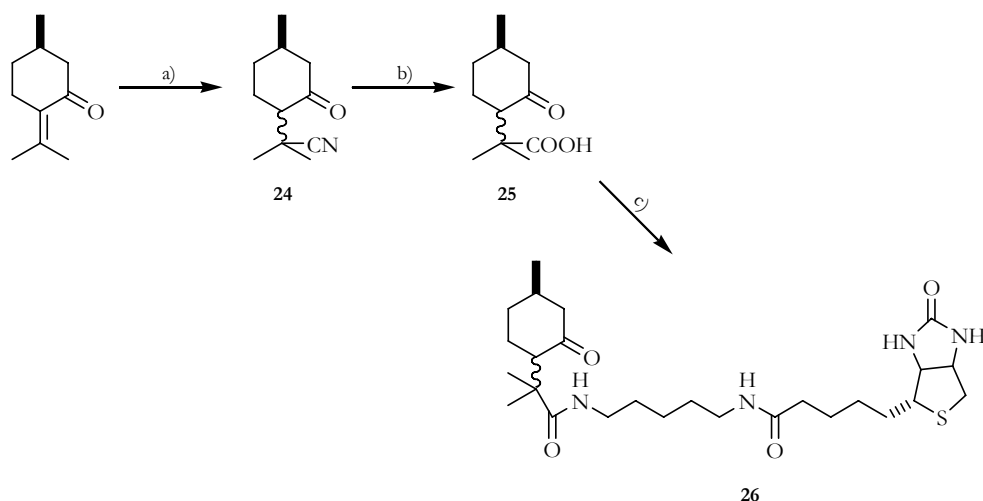


Figure 4.4: the structure of menthone-biotin conjugate

The synthesis of this conjugate are reported in Scheme 4.15:

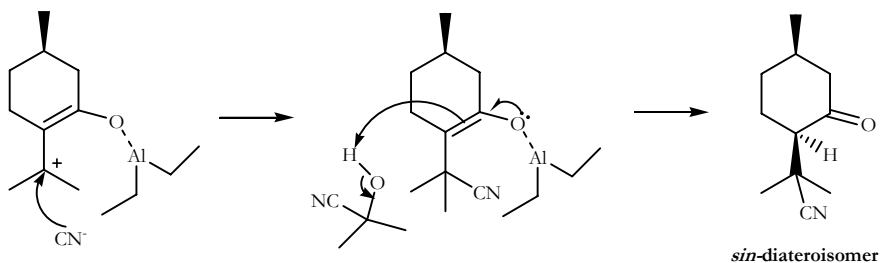


**Scheme 4.15**

a)  $\text{Et}_2\text{AlCN}/(\text{CH}_3)_2\text{C(OH)CN}/\text{toluene}$ ; b)  $\text{H}_2\text{SO}_4:\text{H}_2\text{O}=1:1$ ; c) compound **7**/ $\text{HOBT}/\text{NMM}/\text{EDC-Cl}/\text{DMF}$

The synthesis starts from (+)-pulegone that was converted into cyanoketone **24** by the addition of diethylaluminium cyanide in presence of acetone cyanohydrin with toluene as solvent.<sup>128</sup>

The cyanoketone was obtained as a 3:1 mixture of diastereoisomers, since protonation of the addition intermediate (scheme 1.16) by acetone cyanohydrin is somewhat diastereoselective due to the stereocontrol by the stereocenter in position 1 of pulegone:



**Scheme 4.16**

In our case we have obtained the ratio *sin/anti* = 3/1. For the further steps we have not separated the two diastereoisomers because it was not an objective of our purposes.

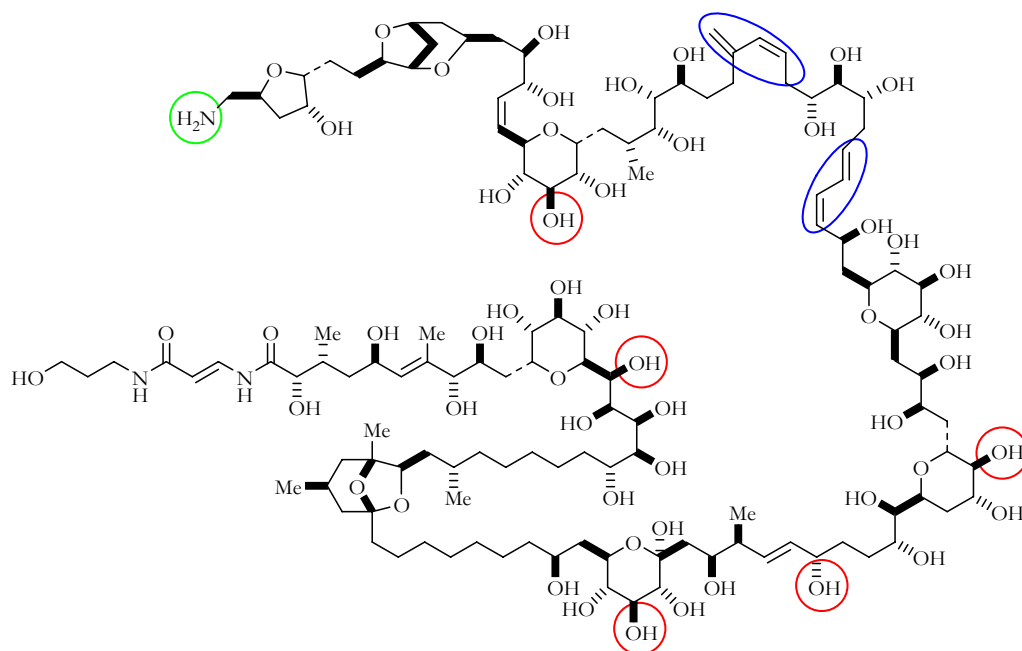
Product **24** was oxidized in aqueous sulphuric acid to give the free acid **25** (75%) that was converted into the final conjugate **26** by the coupling with compound **7** in the standard conditions for coupling of amino acids in 66% yield.

<sup>128</sup> Benedetti F., Berti F., Norbedo S. *Tetrahedron Lett.*, **1999**, *40*, 1041-1044.

### 4.3 Palytoxin

Beside theophylline and caffeine, the other target molecules studied and considered for chemical modifications is palytoxin, an extremely complex organic molecule.

Along its backbone are found 40 secondary hydroxy groups, two diene motifs, a conjugate acrilamide-enamide system, three unsaturations, two hydrophobic hydrocarbon chains, cyclic ether system bicyclic acetal, and a primary amino group at the end of carbon sequence (Figure 4.5):



**Figure 4.5:** palytoxin structure with examples of secondary hydroxy groups (red); diene motifs (blue); free amino group (green).

The primary amino group is essential for both the conformation adopted by the molecule in solution and for its toxicity.

Very recent X-ray scattering and NMR studies suggest that palytoxin exists as a dimer in aqueous environment, while acetylation of the amino group leads to a monomeric structure in solution.<sup>129</sup> Although most of the dimerization interface is hydrophobic and made by the diene and hydrocarbon regions, protonation of the amino group is essential to shift the equilibrium towards the dimer.

As the amino group is likely to be essential, we have decided on one side to attempt the synthesis of a fragment of palytoxin containing this region towards this “epitope”. On the other side, on carrying out bioconjugations of palytoxin, the use of this amino group for conjugations, is almost mandatory.

<sup>129</sup> Inuzuka T., Uemura D., Arimoto H. *Tetrahedron*, **2008**, 64(33), 7718-7723.

As to the synthesis of a small fragment of palytoxin, we have considered the C<sub>101</sub>-C<sub>115</sub> fragment where the most reactive primary amino group is present (Figure 4.6):

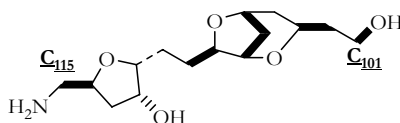
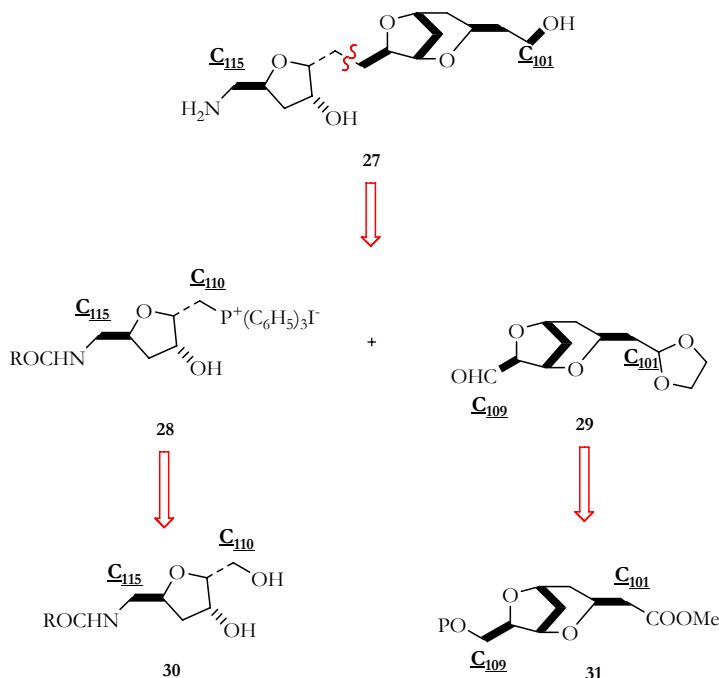


Figure 4.6

#### 4.3.1 Retrosynthetic analysis for the C<sub>101</sub>-C<sub>115</sub> synthesis.

Fragment C<sub>101</sub>-C<sub>115</sub> can be considered as the union of other two smaller fragment: one is given by C<sub>110</sub>-C<sub>115</sub> fragment and one is C<sub>101</sub>-C<sub>109</sub> fragment as reported in the retrosynthesis scheme (Scheme 4.17):



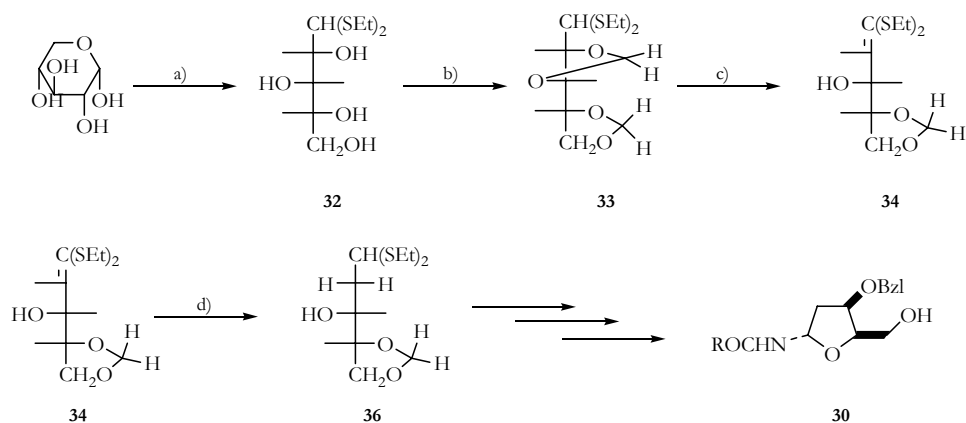
Scheme 4.17: retrosynthesis of C<sub>101</sub>-C<sub>115</sub> fragment

where the C<sub>109</sub>-C<sub>110</sub> carbon bond is obtained by a Wittig reaction between the aldehyde **29** and the triphenyl phosphonium ylide **28**, followed by hydrogen reduction of double bond.

The ylide **28** is easily obtained from the primary alcohol **30** by a nucleophilic substitution and the other partner aldehyde **29**, is obtained by oxidation of protected alcohol **31**.

Synthesis of compounds **30** and **31** has been performed by Kishi<sup>130</sup> during his total synthesis of palytoxin. We have followed the approach proposed by Kishi for the synthesis of **30** as reported in Scheme 4.18a:

<sup>130</sup> Kishi Y., Leder J., Fujioka H. *Tetrahedron Lett.*, **1983**, 24(14), 1463-1466.



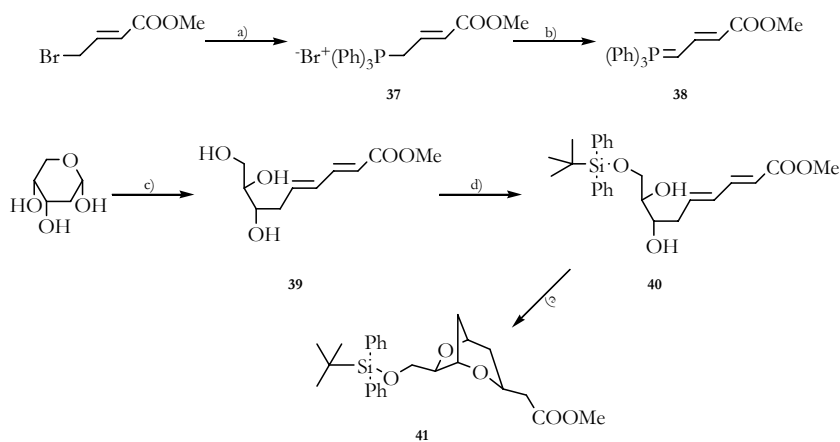
**Scheme 4.18a**

a) EtSH/ZnCl<sub>2</sub>/HCl; b) I<sub>2</sub>/acetone; c) *tert*-BuOK/THF:DMSO=3:1; d) LiAlH<sub>4</sub>/THF.

D-xylose was protected at the aldehyde group as to ether by treatment with ethanethiol (EtSH), zinc chloride (ZnCl<sub>2</sub>) and chloridric acid to give the corresponding ethyl mercaptane **32** with good 70% yield. Protection of hydroxy groups was carried on with I<sub>2</sub> and acetone in order to obtain the diacetone **33** that was converted into the unsaturated compound **34** by potassium *tert*-butoxide (*tert*-BuOK) in THF/DMSO as solvent via acetone elimination.

Reduction of double bond with LiAlH<sub>4</sub> in anhydrous conditions gave the product **34** that can be converted in further five steps into the final product in about 60% overall yield from compound **34**.

The synthetic approach for the fragment **31** is described in the Scheme 4.18b and 4.18c:



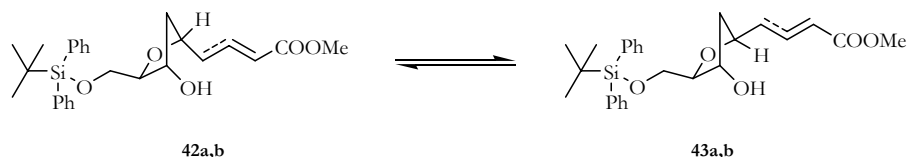
**Scheme 4.18b**

a) P(Ph)<sub>3</sub>/toluene; b) NaOH/water; c) compound **38**/THF/80°C; d) (Ph)<sub>2</sub>Si(t-Bu)Cl/DIMAP/pyridine/DCM/0°C; e) *tert*-BuOK (0.3 eq; 1 M in *tert*-BuOH)/toluene/50°C

Methyl-4-bromocrotonate was converted into the phosphonium bromide **37** by treatment with triphenylphosphine (P(Ph)<sub>3</sub>) in toluene, and then ylide **38** was obtained directly from **37** by base treatment with sodium hydroxide in water solution. Ylide **38** was reacted with 2-deoxy-D-ribose in

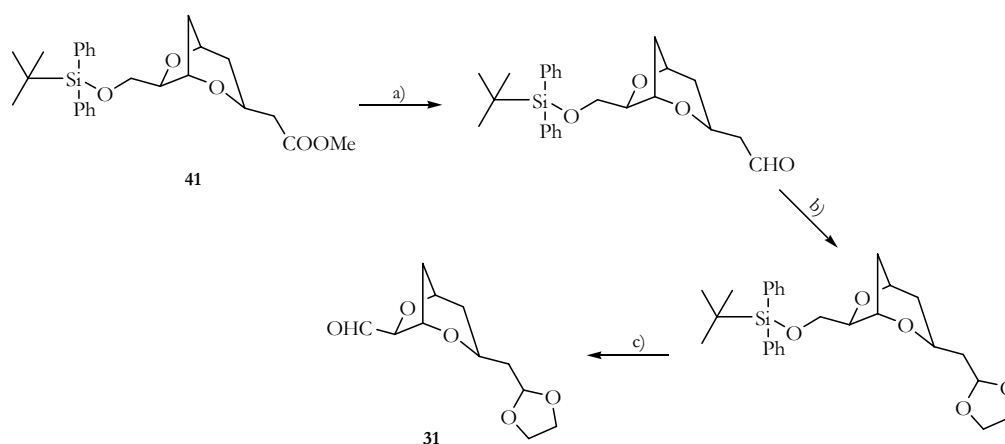
a Wittig reaction, to obtain the diene ester **39** that was protected selectively at the primary hydroxyl group as *tert*-butyldiphenylsilyl ether (**40**).

Compound **40** was then treated with potassium *tert*-butoxide under carefully controlled conditions to give the desired bicyclic ester **41** in addition to a mixture of four monocyclic esters **42a,b** and **43a,b** (Figure 4.7):



**Figure 4.7:** equilibrium of the four monocyclic ester

An equilibrium between species **41**, **42** and **43** was established via the precursor diene ester **40**, and the composition of the equilibrium mixture depends on the base and solvent used; the best conditions were found to be 0.3 equivalents of *tert*-BuOK in a mixture of *tert*-BuOH and toluene. We have followed the same synthetic way described by Kishi, and we have obtained compound **41** in an overall 21% yield from D-xylose. In order to accumulate enough material to reach the end of the synthesis, we have repeated the whole preparation of **41** seven times, thus accumulating 2.3 g of this compound. The synthetic work is still in progress, and further steps (Scheme 4.18c):



**Scheme 4.18c**

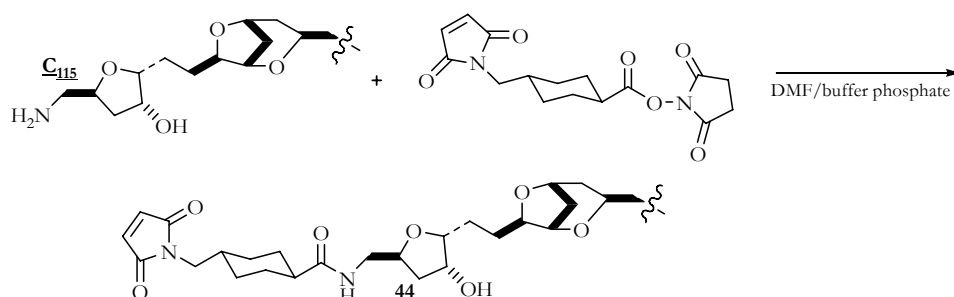
**Scheme 4.18c:** a) DIBAL-H/DCM/-78°C; b) (CH<sub>2</sub>OH)<sub>2</sub>/p-TSA/toluene(reflux); c) (COCl)<sub>2</sub>/DMSO/Net<sub>3</sub>/DCM.

The reduction of **41** with diisobutyl aluminiumhydride (DIBAL-H) in order to obtain the aldehyde **42** that will be protected as cyclic acetals **43** using ethylene glycol, p-toluensulfonic acid and toluene as solvent.

Final Swern oxidation allows to obtain fragment **31**.

### 4.3.2 Bioconjugation of palytoxin.

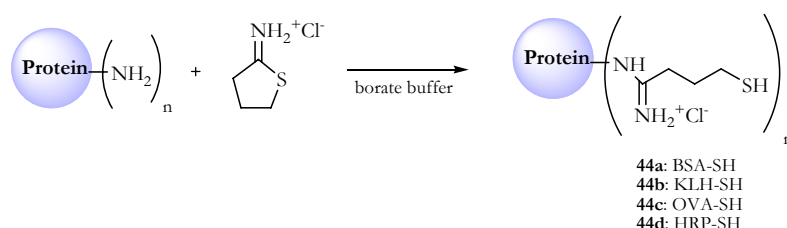
The conjugation of palytoxin to proteins as KLH, BSA, OVA, HRP was performed in order to use the bioconjugates as immunogens or as coating antigens for immunoassay development, or as an enzyme marker for competition ELISA. All the bioconjugations were carried out by modifying palytoxin at its amino terminal. Proteins were binded via a maleimido-linker that was bound to palytoxin by acylation (Scheme 4.19):



**Scheme 4.19:** palytoxin modification

Treatment of palytoxin with commercially available bifunctional linker succinimidyl-4-(N-maleimidomethyl)-cyclo-hexane-1-carboxylate (SMCC) yielded the corresponding amide derivative **44** (PTX-MCC). The reaction was carried out in DMF/buffer phosphate on a 1 mg scale and the modified palytoxin **44** was not purified.

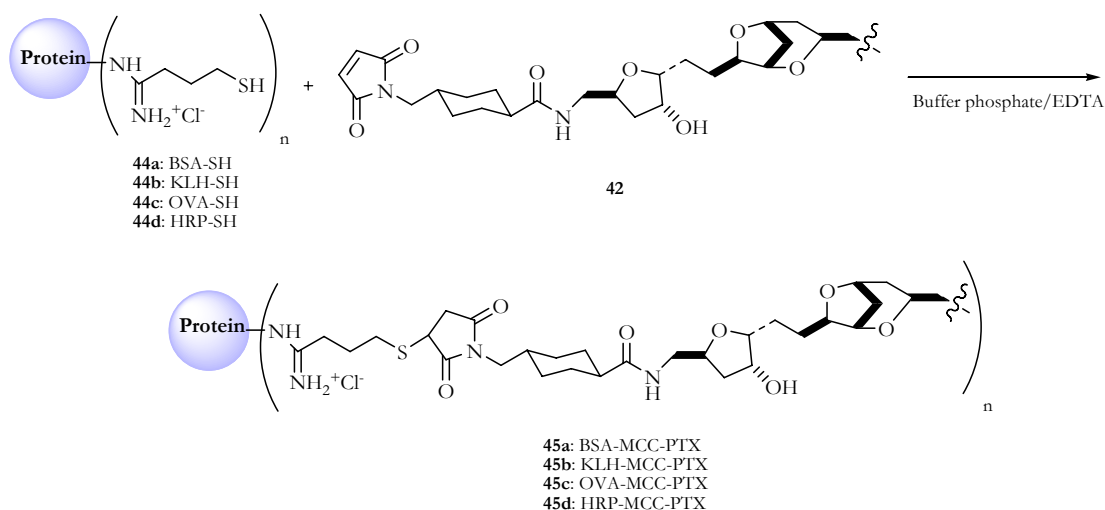
At the same time, a modification of the lysine primary amino groups of proteins was necessary in order to convert them into thiol groups. Three different proteins and one enzyme have been thiolated using 2-iminothiolane (2-IMT); the three proteins were albumin from bovine serum (BSA), keyhole limpet hemocyanin (KLH) and albumin from chicken egg white (OVA); the enzyme was horseradish peroxidase (HRP). BSA, OVA and KLH are proteins frequently used for immunization treatment as carrier for different aptens; HRP is typical enzyme used in immunoassay (Scheme 4.20):



**Scheme 4.20:** protein modification

The reactions were carried out in borate buffer (pH=9.0) and the modified proteins was not purified.

PTX-MCC **42** can react with protein thiol residues to form thioether bond to give final thioether conjugates (Scheme 4.21):



Scheme 4.21

The resulting conjugates were purified by dialysis against buffer phosphate. UV spectroscopy was used as previously described for the caffeine/theophylline conjugates to obtain the modification index of the BSA and KLH conjugates. These conjugates are infact used to immunize animals, and thus in function is necessary in order to evaluate the amount of conjugate to be administred (Table 4.1):

Conjugate	n PTX / n protein
PTX-BSA	34
PTX-KLH	133

Table 4.1

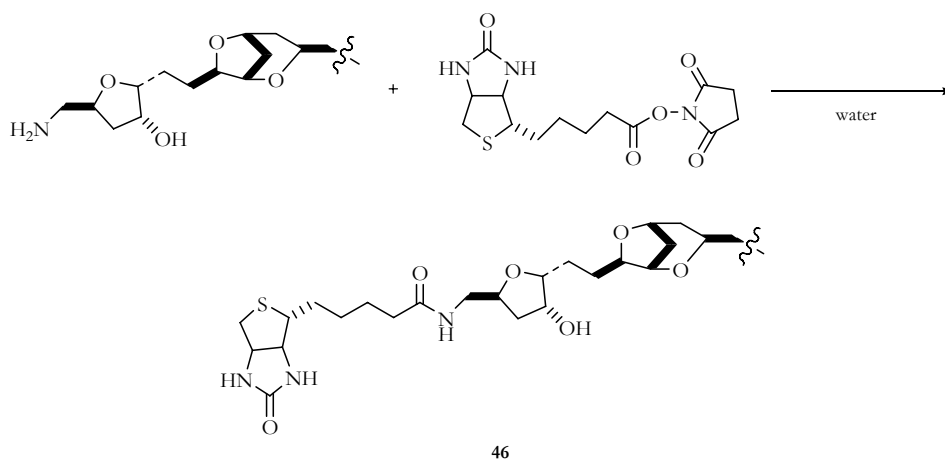
The value n PTX /n protein is the index of modification (IM) and represents the number of PTX molecules bound at one molecule of protein.

The IM value is high in both the PTX-BSA and KLH-PTX and the conjugates can be used for the immunization runs.

#### 4.3.3 Biotinylation of palytoxin.

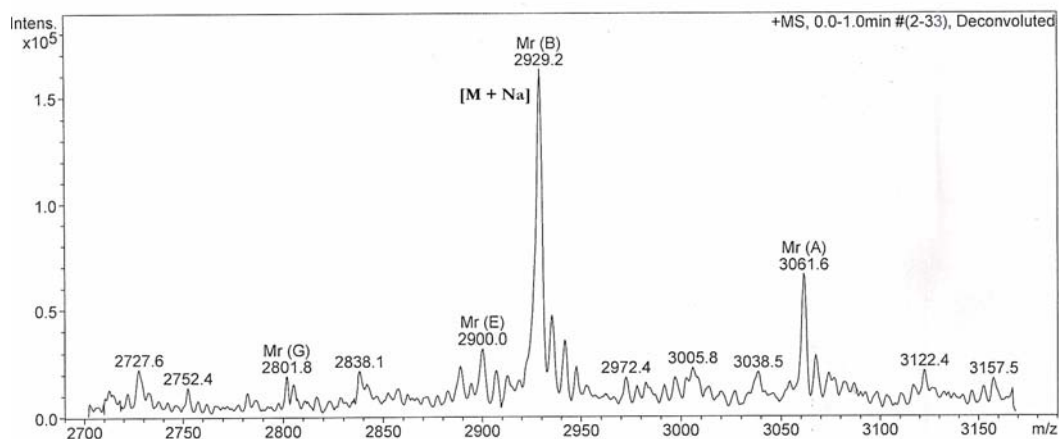
An other important modification of palytoxin was introduced by biotinylation of its amino terminal; palytoxin-biotin conjugate can be used for the immobilization of toxin on a solid system exploiting streptavidin

Likewise the other palytoxin modification, synthesis of palytoxin-biotin involves a modification on the primary amino group of toxin (Scheme 4.22):



**Scheme 4.22:** palytoxin biotinylation

Reaction was carried out on a 1 mg scale directly between palytoxin and commercial biotin N-hydroxysuccinimide ester using water as solvent. The product was well characterized by mass spectroscopy and the mass spectrum is reported in Figure 4.8:



**Figure 4.8**

As this spectrum clearly shows the presence of  $[M+Na]^+$  molar ion of biotinylated-palytoxin plus  $Na^+$  and there are no other products. The conjugate has been used without further purification.

## Chapter 5 : The E/K coiled-coil peptide

Coiled-coils have been used in numerous applications, including affinity purification, direct assembly of extracellular receptor domains, creation of miniaturized antibodies, microarray and peptide libraries.<sup>131</sup> The small size and defined structure of the coiled-coil make it an extremely used motif for a growing number of *de novo* design, protein engineering and biotechnological application.

### 5.1 The E/K structure

Based on typical coiled-coil features described in the “chapter 1”, in 1998 Hodges et al. characterized the E/K coiled-coil, a *de novo* designed heterodimeric parallel coiled-coil domain, for biosensor-based affinity and chromatographic applications.<sup>132</sup> They analyzed the structure of several natural coiled-coils, and some requirements were established to permit the correct folding of the new protein (see Figure 2.1)

- 1- to select amino acids that are non-destructive for the helical structure;
- 2- the placement of non-polar residues at *a* and *d* positions for tighten packing in the hydrophobic core;<sup>133</sup>
- 3- to chose appropriately polar or charged residues at *e* and *g* positions that will promote the helix associations in a heterodimeric form;<sup>134</sup>
- 4- the inclusion of polar or charged residues at position *f*, to permit the protein solubilization in aqueous solution;<sup>135</sup>
- 5- the polar or charged residue at position *f* must be opposite charge respect to aminoacids in *e* and *g* position, in order to balance the total charge on individual  $\alpha$ -helix;
- 6- to select an optimal lenght of  $\alpha$ -helix to permit the protein production with synthetic means, and to avoid alteration of structure or stability of the coiled-coil;
- 7- balance between stability of hydrophobic core that permit heterodimer folding and electrostatic repulsion that prevent the assembly of homodimers.<sup>136</sup>

---

<sup>131</sup>a)Mc Farlane A. A., Orriss G., Stetefeld J. *Eur. J. Pharmac.*, **2009**, *625*, 101-107. b)Lupas, A. *Trends Biochem. Sci.*, **1996**, *21*, 375-382.

<sup>132</sup> Chao H., Bautista D. L., Litowky J., Irvin T. R., Hodges R. S. *J. Chromatogr B. Biomed. Sci. Appl.*, **1996**, *715(1)*, 307-329.

<sup>133</sup>a)O'Shea E.K., Albert T., Kim P.S. *Science*, **1991**, *254*, 539. b)Woofson D.V., Alber T. *Protein Sci.*, **1996**, *4*, 1596. c)Hodges R.S., Zhu B.Y., Zhou C.M. *J. Biol. Chem.*, **1984**, *259*, 13253. d)Hodges R.S., Sodek J., Smillie L.B., Jurasek L. *Proc. Natl. Acad. Sci USA*, **1972**, *69*, 3800.

<sup>134</sup>a)Hodges R.S., Kohn W.D., Kay C.M. *J. Mol. Biol.*, **1997**, *267*, 1039. b)Hodges R.S., Kay C.M., Zhou N.E. *Protein Eng.*, **1994**, *7*, 13655. c)Krylov D., Mikhailenko I. *EMBO J.*, **1994**, *13*, 2849. d)Lumb K.J., Kim P.S. *Science*, **1995**, *268*, 436.

<sup>135</sup> Lyu P.C., Gans P.J., Kallenback N.R. *J. Mol. Biol.*, **1992**, *223*, 343-351.

<sup>136</sup> Lavigne P., Kondejewski L.H., Houston M.E., Sonnichsen F.D., Lix B., Sykes B.D., Hodges R.S. Kay C.M. *J. Mol. Biol.*, **1995**, *254*, 505-517.

Based on the statements given above, the E/K coiled-coil developed by Hodges et al. results in the couple of sequences reported in Figure 5.1:

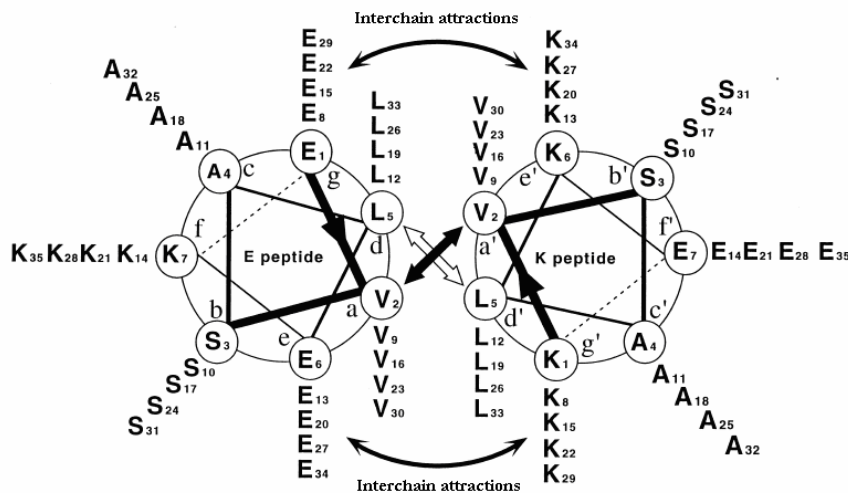


Figure 5.1: helical wheel representation of E/K coiled-coil system viewed from N-terminus.

The E/K coiled-coil is made of two distinct chains namely the E and K peptides. Each peptide is composed by five repeats of a single heptad; in the E peptide, the heptad sequence is E-V-S-A-L-E-K (*g-a-b-c-d-e-f*) and in the K peptide the sequence is K-V-S-A-L-K-E (*g-a-b-c-d-e-f*).

The hydrophobic core is made by valine in *a* position and leucine in *d* position; these aminoacids are common in the natural and synthetic coiled-coils. Valine is often replaced by isoleucine in order to generate a leucine-isoleucine core but the choice of Hodges et al. tends to favour the dimeric parallel association.<sup>137</sup>

Complementary lysine and glutamate residues at *e* and *g* position were placed in order to favour electrostatic interchain attractions; these aminoacids are also frequently present in natural coiled-coil motifs. In addition, at neutral pH values, the repulsive interactions between the charged side-chains prevent formation of homodimeric E/E and K/K coiled-coils, allowing the formation of E/K heterodimeric coiled-coil.<sup>138</sup>

The *c* position was assigned to alanine, because of its particular ability to stabilize helical structure; serine in *b* position increases coiled-coil solubility in a polar solvent and glutamate on the K-peptide and lysine on the E-peptide in the *f* positions were used to balance the total charge of coiled-coil.

<sup>137</sup> Harbury P.B., Zhang T., Kim P.S. Alber T. *Science*, **1993**, 262, 1401-1403.

<sup>138</sup> a) Kohn W.D., Kay C.M., Hodges R.S. *Protein Sci.*, **1995**, 4, 237. b) Yu Y., Monera O.D., Hodges R.S., Privalov P.L. *Biophys Chem.*, **1996**, 59, 299.

E and K associate in a heterodimeric protein of 70 aminoacids with a molecular weight of 7682 Da. Each peptide is amidated and acetylated in order to increase the helix formation and to reduce the likelihood of proteolysis.

### 5.1.1 Physical characterization of the E/K coiled-coil

The EK peptide prefers to form a heterodimeric coiled-coil at moderate protein concentrations.<sup>139</sup> The CD spectra of E- or K-peptides alone shown a largely random coil character; in contrast, a solution containing both peptides displays a CD profile that is characteristic of a helical structure, with spectral minima near 208 and 222 nm<sup>132</sup> (Figure 5.2). The ratio of molar ellipticities at 222 nm ( $[\theta]_{222}$ ) over 208 nm ( $[\theta]_{208}$ ) is 1.07, a typical value for coiled coils, indicating the presence of interacting helices in a coiled-coil conformation (a ratio of 0.85-0.95 is conversely expected for single  $\alpha$ -helices).<sup>140</sup>

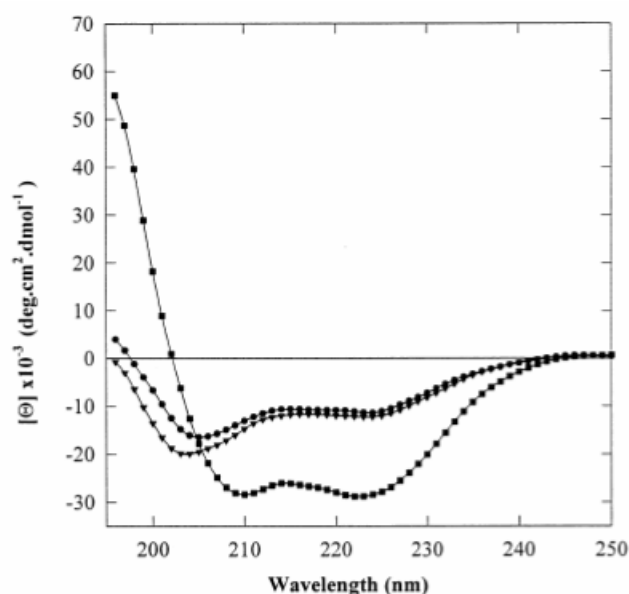


Figure 5.2: CD spectra of the E peptide ( $\blacktriangledown$ ), K peptide ( $\bullet$ ) and the E/K coiled-coil ( $\blacksquare$ ) recorded at 20°C. taken from reference 10.

The stability of E/K coiled-coil was also studied by both chemical and temperature denaturation experiments as monitored by CD spectroscopy.<sup>132</sup> The E/K coiled coil is resistant to heat denaturation up to 85°C or urea denaturation. Guanidinium HCl (GdnHCl) was the only agent which could denature the E/K coiled-coil completely<sup>132</sup> (Figure 5.3):

<sup>139</sup> Chao H., Houston Jr., Grothe S., Kay C.M., Irvin R.J., Hodges R.S. *Biochemistry*, **1996**, *35*, 12175.

<sup>140</sup> Litowski J., Hodges R. S. *J. Chem. Biol.*, **2002**, *277(40)*, 37272-37279.

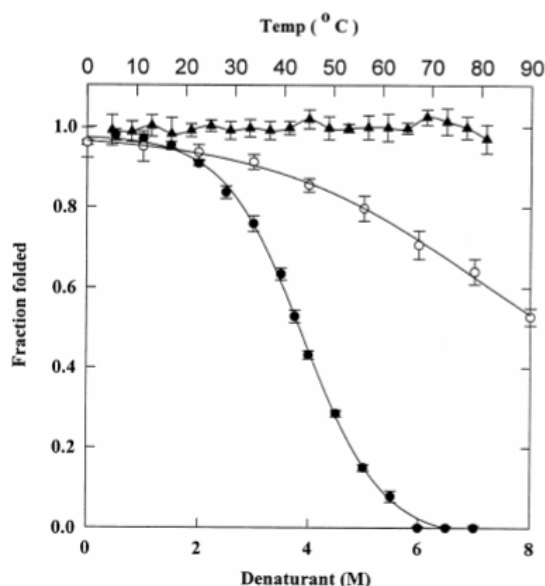


Figure 5.3: Urea (○), GdnHCl (●), heat denaturation (▲) of the E/K coiled-coil taken from reference 10 .

This high association strength is highlighted by the dissociation constant for this dimer that is 63 pM.<sup>141</sup>

High stability and affinity of dimerization make E and K peptides as good candidates for biosensor applications in which the biorecognition element must demonstrate capacity of binding with analyte and particular resistance to environmental stress.

## 5.2 The C-EK<sub>inv</sub>-C scaffold, a first library project

Starting from the E/K coiled-coil, the aim of this part of the project was to generate libraries of modified E and K peptides with the following features:

- 1- ability to bind our small target molecules caffeine and theophylline;
- 2- short length, in order to allow both chemical synthesis and biological expression;
- 3- good stability, under different pH or temperature conditions;
- 4- small size of the binding site, constituted by a small number of aminoacids to allow rational affinity improvement after a first selection round.

As a first version of an E/K-based artificial receptor, we have chosen to adopt a scaffold design that could in principle allow a maximum stability due to a circular structure. In order to achieve this kind of stability we have introduced a flexible linker that joins E and K peptide chains, consisting of eight aminoacids: a threonine and seven glycine that assure high flexibility and low steric hindrance. This kind of Thr/Gly linker is commonly used in molecular biology technique in order to build up proteins constructs arising from the fusion of different domains.<sup>142</sup>

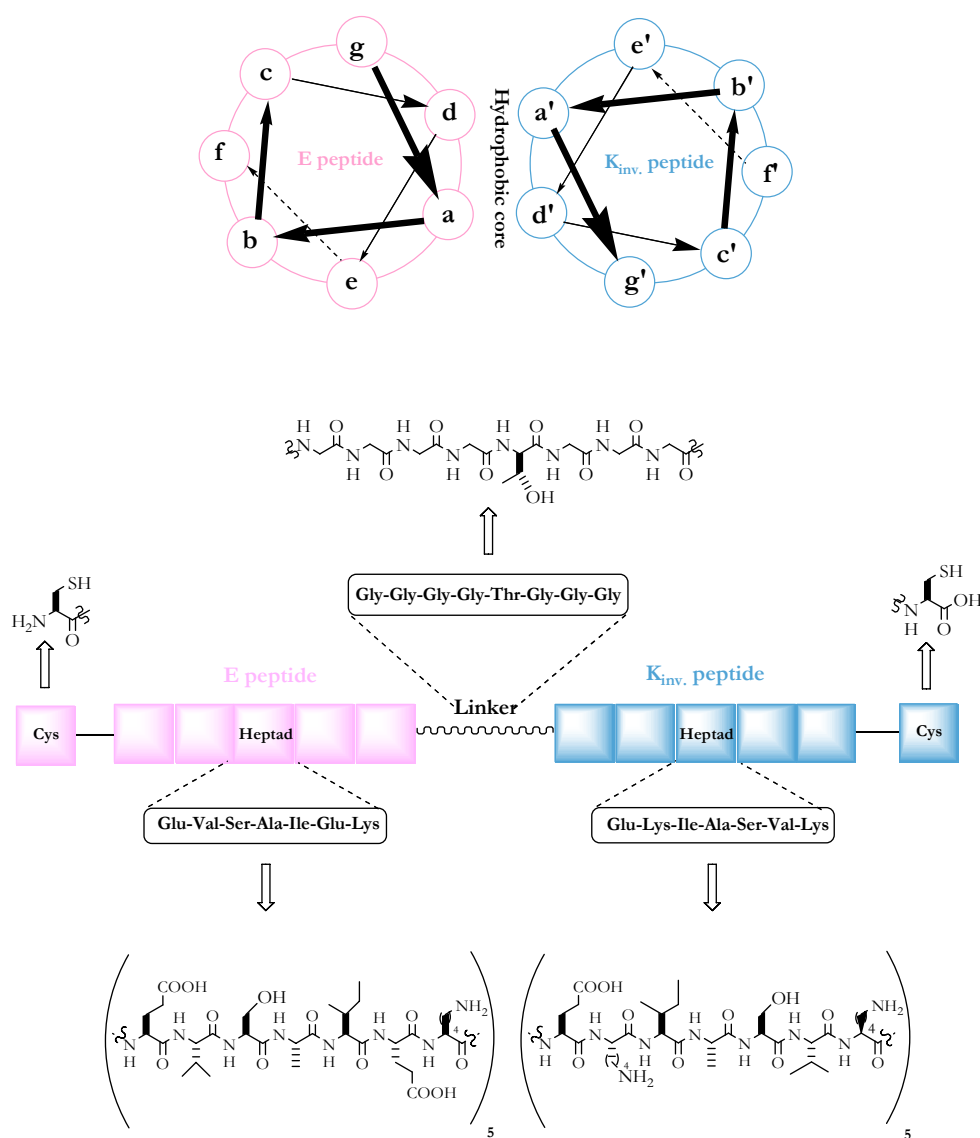
<sup>141</sup> De Crescenzo G., Litowski J.R., Hodges R.S. O'Connor-Mc Court M.D. *Biochemistry*, **2003**, 42(6), 1754-1763.

<sup>142</sup> Zoldak G., Aumuller T., Lucke C., Hritz J., Schmid F.X. *Biochemistry*, **2009**, 48(43), 10423-10426.

Reverting the K sequence ( $K_{inv}$ ) in order to keep the E/K, we have planned coiled-coil association even if the coil is shifted from a parallel to antiparallel orientation.

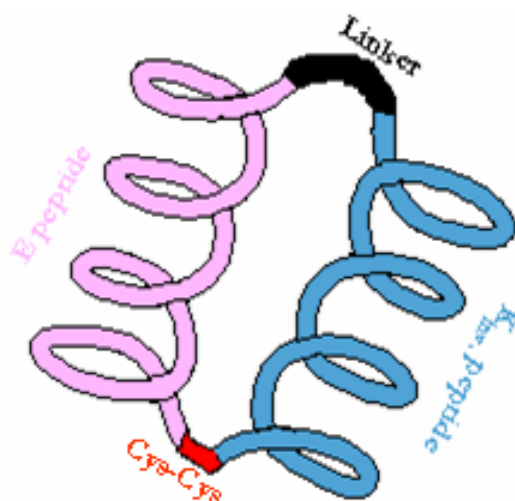
Finally, we have planned the introduction of two cysteines as amino and carboxy terminals in the hope of generating an intramolecule disulfide bond.

As to the  $\alpha$ -helical domains, both the two  $\alpha$ -helices were constituted by a repeat five times of a single heptad; in the E domain the heptad sequence is the same of the original coiled-coil: E-V-S-A-L-E-K (positions  $g$ - $a$ - $b$ - $c$ - $d$ - $e$ - $f$ ); in the  $K_{inv}$  domain the new sequence is E-K-L-A-S-V-K (positions:  $f'$ - $e'$ - $d'$ - $c'$ - $b'$ - $a'$ - $g'$ ). The new scaffold was thus dubbed C-E $K_{inv}$ -C (Figure 5.4-5.5):



**Figure 5.5:** helical wheel representation and linear structure of E $K_{inv}$  peptide

With the flexible linker connecting the helical domains, and the disulfide bond at the terminal, the scaffold C-E $K_{inv}$ -C can in principle adopt a circular closed tertiary structure (Figure 5.6):



**Figure 5.6:** EK<sub>inv</sub> scaffold closed in double way. Flexible linker was curved in order to make a loop for E and K<sub>inv</sub> chains assembly. Disulfide bond blocks final structure.

### 5.2.1 Preliminary study on the EK<sub>inv</sub> folding

Having thus designed the overall structure of C-EK<sub>inv</sub>-C, we have started several preliminary studies aimed at determining its conformation in water.

A theoretical model of the coiled-coil was built first by homology modeling. The NMR structure of the C-Jun homodimer, one of the few available in the Protein Data Bank (Pdb id: 1JUN) was chosen as the starting topology.<sup>143</sup>

The C-Jun homodimer is a six heptad parallel homodimeric coiled-coil consisting of a non repeating sequence (Table 5.1-entry 1):

1. C-Jun	G	R	I	A	R	L	E	E	K	V	K	T	L	K	A	Q	N	S	E	L	A	S	T	A	N	M	L	R	E	Q	V	A	Q	L	K	Q	K	V	M	N	Y
2. Mut1		R	V	A	R	L	E	E	K	V	K	T	L	K	A	Q	V	S	E	L	A	S	T	V	N	M	L	R	E	Q	V	A	Q	L	K	Q					
3. Mut2		E	V	A	R	L	E	K	E	V	K	T	L	E	K	E	V	S	E	L	E	K	E	V	N	M	L	E	K	E	V	A	Q	L	E	K					
		K	V	A	R	L	K	E	K	V	K	T	L	K	E	K	V	S	E	L	K	E	K	V	N	M	L	K	E	K	V	A	Q	L	K	E					
4. E/K		E	V	S	A	L	E	K	E	V	S	A	L	E	K	E	V	S	A	L	E	K	E	V	S	A	L	E	K	E	V	S	A	L	E	K					
		K	V	S	A	L	K	E	K	V	S	A	L	K	E	K	V	S	A	L	K	E	K	V	S	A	L	K	E	K	V	S	A	L	K	E					

**Table 5.1**

The initial C-Ju homodimeric structure was relaxed first into a periodic box of TIP water molecules by a 100 ns molecular dynamics run in the NTV ensemble, carried out with the AMBER force field as implemented in Sybyl 8.2 in iT'S cornell version.<sup>144</sup> The relaxed structure was then gradually mutated to that of the E/K parallel heterodimer, by first replacing all the residues at position *a* and *d* with valine and leucine respectively, in both the chains (Table 5.1, entry 2, mut1). The charged residues at positions *g*, *e*, *f*, where then mutated into E, K, E in one chain and into K, E, K in the other, as in the E/K heterodimer (Table 5.1, entry 3, mut2). Finally, also

<sup>143</sup> Jomins F.K., O'Donoghe S.I., Nilges M., Weiss A.S., King G.F. *J.Biol.Chem.*, **1996**, 271, 13663-13667.

<sup>144</sup> Cornell W.D., Gelpack P., Bayly C.I., Gould I.R., Merz K.M.S., Ferguson D.M., Sjellmeyer D.C., Fox T., Caldwell J.W. Kollman P.A. *J. Am. Chem. Soc.*, **1995**, 117, 5179.

positions *b* and *c* were mutated to S and A in both the chains, giving by this way a model of the E/K heterodimer. Full MD relaxation was allowed at each stage, to verify the stability of the mutated models. The final model adopts the typical superhelical conformation of a coiled-coil

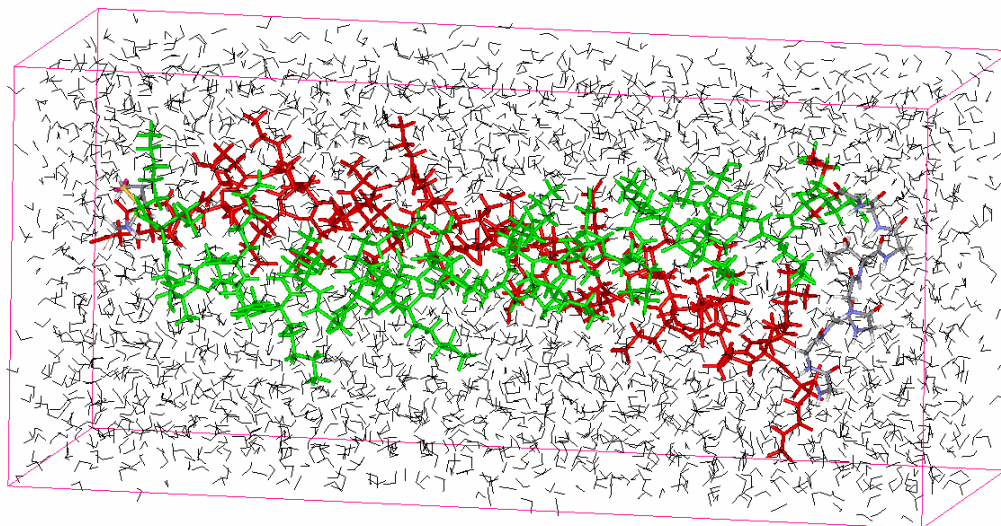


Figure 5.6

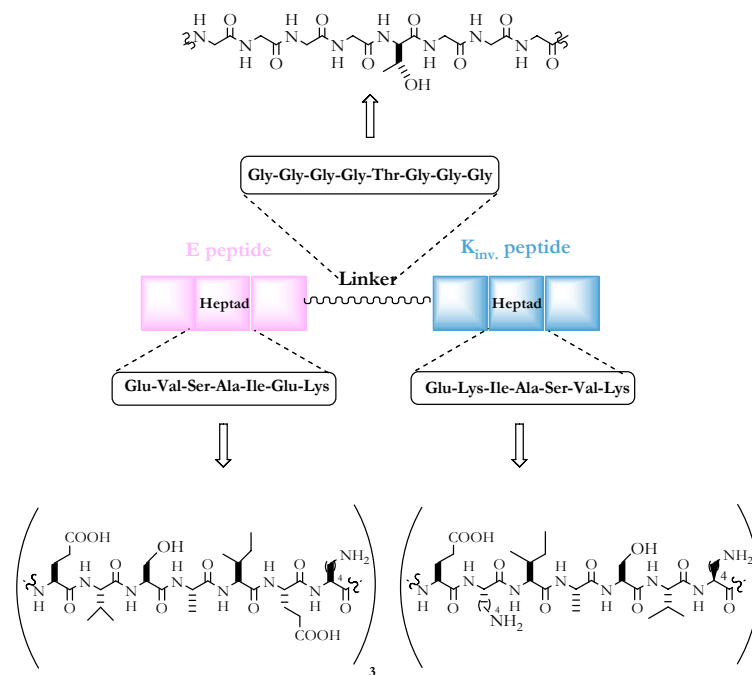
Further mutation to the E/K<sub>inv.</sub> model required a more complex approach, and the K chains was temporarily broken at each heptad end, and each heptade was separately inverted and relaxed. Finally, the heptades were joined again and the whole model was submitted to 300 ns MD run. The resulting antiparallel dimer (Figure 5.6) resulted stable and no unfolding or separation of the two chains were detected during this very long molecular dynamics symulation.

The number of hydrophobic contacts at the coil interface is conserved, and also the buried hydrophobic surface is similar in the E/K and in the antiparallel EK<sub>inv.</sub> models (Table 5.2).

	EK	EK <sub>inv.</sub>
<b>Buried hydrophobic surface (Å<sup>2</sup>)</b>	1290	1253
<b>Hydrophobic contacts</b>	30	28

Table 5.2

Starting from this favourable theoretical prediction, we have synthesized the E<sub>3</sub>K<sub>inv.3</sub> peptide on a 10 mg scale, as a reduced model of the full C-EK<sub>inv.</sub>-C scaffold (Figure 5.7).



**Figure 5.7:** linear structure of  $EK_{inv}$  synthesized

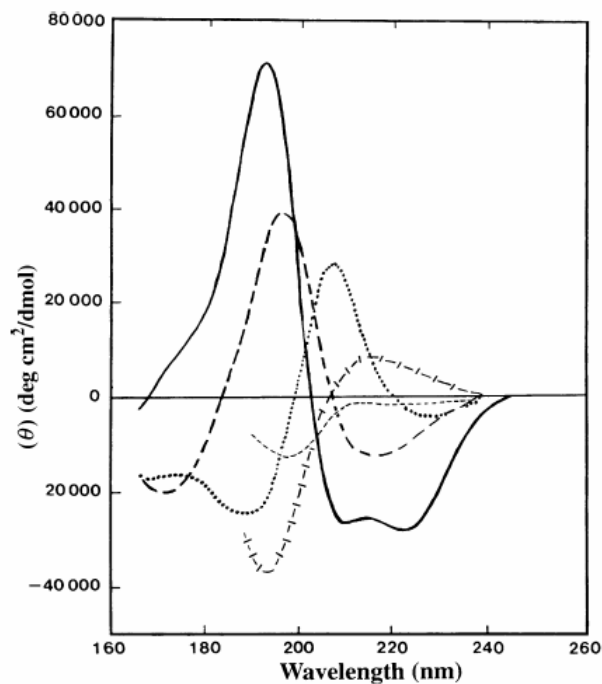
Due to its synthetic cost,  $E_3K_{inv,3}$  is shorter than the full scaffold and the absence of final cysteine can influence the real peptide behaviour; however, it can mimic with a certain reliability the behaviour of  $EK_{inv}$  and a preliminary study of its folding has been carried out.

The folding of previous peptide was studied by CD spectroscopy.

Circular dichroism (CD) is one of the most used technique to investigate secondary structure of proteins or peptides, as CD spectra can be readily used to estimate the fraction of  $\alpha$ -helix,  $\beta$ -sheet or random-coil conformation.

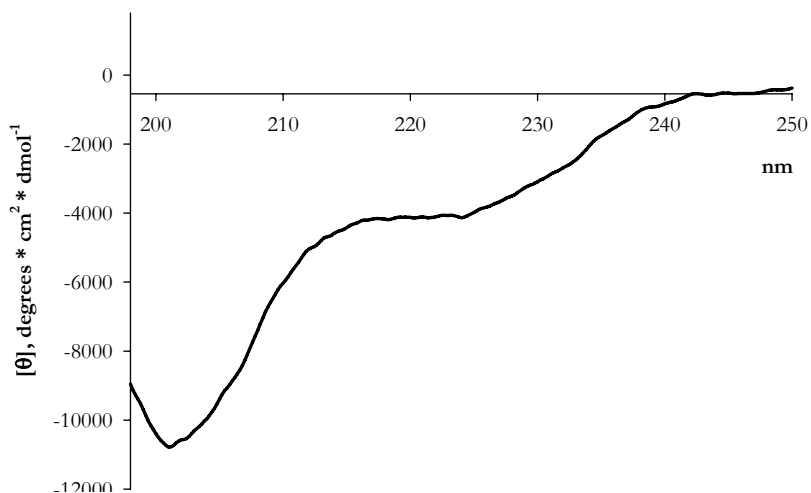
Protein or peptide secondary structure can be determined by CD spectroscopy in the far-UV spectral region (190-250 nm), where the chromophore is the peptide bond and the CD signal arises when the bond interact with the chiral environment of the folded peptide. The lowest energy transition in the peptide chromophore is an  $n \rightarrow \pi^*$  transition observed at 210-220 nm, and the highest energy process is the  $\pi \rightarrow \pi^*$  transition that gives rise to a band at 120 nm. The ellipticity associated to such transition depends on the local conformation of the peptide chain and  $\alpha$ -helices,  $\beta$ -sheets and random-coils each gives rise to a characteristic shape and magnitude CD spectrum<sup>145</sup> (Figure 5.8).

<sup>145</sup> Kelly S.M., Jess T.J., Price N.C. *Biochim. Biophys. Acta*, **2005**, 1751, 119-139.



**Figure 5.8:** CD spectra of various types of secondary structures. Solid line,  $\alpha$ -helix; long dashed line, antiparallel  $\beta$ -sheet; dotted line, type I  $\beta$ -turn; cross dashed line, extended  $3_1$ -helix or poly(Pro) II helix; short-dashed line, irregular structure. Data taken from ref. 145 .

The circular dichroism spectrum of a freshly prepared 5  $\mu$ M peptide solution in water is shown in Figure 5.9:



**Figure 5.9**

In our case the previous spectrum is typical for a random-coil conformation due to the presence of a minimum 202 nm, and to the lack of Cotton effect inversion along the whole spectral

range.<sup>146</sup> A fraction of folded peptide can be responsible for the shoulder around 220 nm spectra. A new spectrum was registered after four days (Figure 5.10):

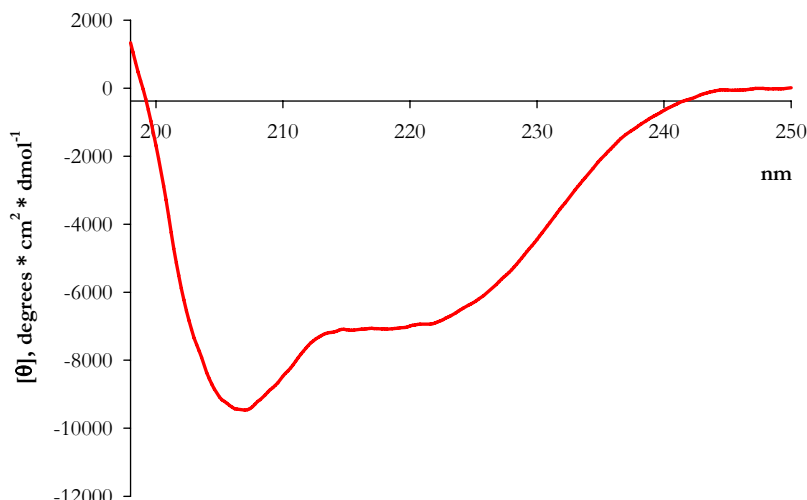


Figure 5.10

Major changes are observed after this time, and now the spectrum is typical for a folded structure.

The  $\alpha$ -helix content of this structure has been calculated according to the empirical equation 2.1<sup>147</sup>

$$f_{\alpha} = ([\theta]_{222} - [\theta]_R) / ([\theta]_{\alpha} - [\theta]_R) \quad \text{eq.5.1}$$

which is commonly used for short peptides, where  $f_{\alpha}$  is the fractional  $\alpha$ -helix content,  $[\theta]_R$  is the ellipticity at 222 nm for a disordered chain and  $[\theta]_{\alpha}$  is the ellipticity at 222 nm for a 100%  $\alpha$ -helix  $n$  long peptide.  $[\theta]_{\alpha}$  is obtained from this equation(4.2):<sup>146</sup>

$$[\theta]_{\alpha} = [\theta]_{\infty} (1 - k/n) \quad \text{eq. 5.2}$$

where  $[\theta]_{\infty}$  is the molar ellipticity per residue for a chain of infinite length with maximal  $\alpha$ -helix content and  $k$  is an empirically determined factor interpreted as the average number of peptide bonds in the chain that cannot participate in  $\alpha$ -helix formation. There is general agreement for an  $[\theta]_{\infty}$  value of 40000 deg dmol<sup>-1</sup>cm<sup>2</sup> but the  $k$  value remains in dispute. Estimated value of  $k$  have

<sup>146</sup> Fasman G.D. *Circular dichroism and the conformational analysis of biomolecules*, New York U.S.A.; Plenum Publishing Corp.

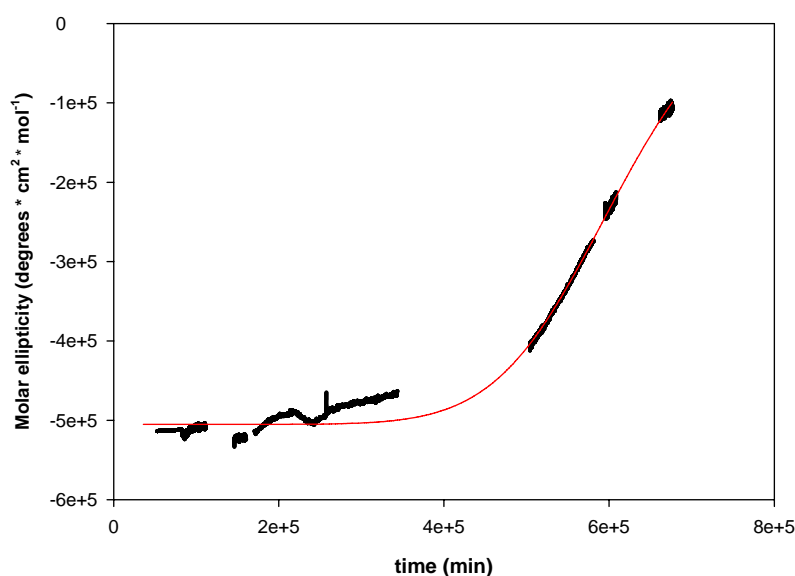
<sup>147</sup> Ahmad B., Ahmed Zulfazal M., Haq S., Hasan Khan R. *Biochim. Biophys. Acta*, **2005**, 1750, 93-102.

ranged from  $2.5^{148}$  to  $4.3^{148}$ , and even up to  $6.3^{149}$ ; for our work we have used the commonly value of  $4^{146}$ ; for the  $[\theta]_R$  value we have used the estimate  $0 \text{ deg dmol}^{-1} \text{ cm}^2$ .

A 40% helix content was obtained from eq. 2.1 for our case. This value is not in agreement with the expected  $\alpha$ -helical character of  $E_3K_{inv.3}$  if it were correctly folded.

Conversely, the  $[\theta]_{222}/[\theta]_{208}$  ratio is consistent with the presence of a single  $\alpha$ -helix rather than with a coiled-coil, or, more likely, the coiled-coil is too short to induce the superhelical deformation to the  $\alpha$ -helix which are responsible for the enhancement in 208 nm helipticity that is observed in longer coils.<sup>140</sup>

The folding process of  $E_3K_{inv.3}$  is apparently extremely slow. This point is perhaps remarkable, and would require a more detailed study. Unfortunately, due the small amount of available peptide, we were able to obtaining a single very preliminar kinetic run by monitoring at differents times the molar ellipticity change at 222 nm (Figure 5.11). The resulting sigmoidal behaviour is consistent with a cooperative process with a very long histeresis followed by an autocatalytic phase.



**Figure 5.11:** kinetic folding of  $EK_{inv.}$  peptide.

Although very unusual, this behaviour is not unprecedented.

The four-helix bundle Rop, also known as Rom, is a homodimer of two helix-loop-helix monomers, each of which is 63 aminoacid long and its folding time after guanidinium denaturation, is in the order of days.<sup>150</sup> The mystery of Rop folding was addressed by many authors, and several theoretical explanations have been proposed. In a recent, general study on the molecular basis of coiled-coil formations, Steinnetz, Kammerer and colleagues have

<sup>148</sup> Chen Y. H., Chau K.H. *Biochemistry*, **1974**, *13*, 3350-3359.

<sup>149</sup> Ozdemir A., Lednev I.K., Asher S.A. *Biochemistry*, **2002**, *41*, 1893-1896.

<sup>150</sup> Monson M., O'Beein R., Startevant J.M., Regan L. *Protein Sci.*, **1994**, *3*, 2015-2022.

suggested that , in order to allow a correct folding of a coiled-coil, a trigger sequence must be present.<sup>151</sup> However, the E/K sequence in the native heterodimer leads to very usual and fast folding kinetics<sup>152</sup> and the unusual folding of E<sub>3</sub>K<sub>inv.3</sub> seems rather connected with the presence of the loop. The presence of a trapping minimum in the protein energy landscape has been proposed for the Rop dimer, and viewed in a series of Rop mutants, as a *syn* association of the two monomer instead of the correct anti topology, or in other cases as a molten globule.<sup>153</sup> In an attempt to generate crystals for an X-ray analysis of E<sub>3</sub>K<sub>inv.3</sub> structure, we have also submitted the peptide to a GPC analysis, that revealed the existence of three fractions, at molecular weight corresponding to a monomer, a trimer and an octamer respectively. The octamer appears as the predominant fraction. Again, the CD spectrum of a freshly isolated octamer fraction, reveals a largely unfolded population and again the CD slowly changes and after 48 hours the folded structure is observed, thus supporting the possibility that the octameric form could represent the free energy trap that makes slow the folding of E<sub>3</sub>K<sub>inv.3</sub>.

### 5.2.2 Binding site design

In order to create a specific binding site for our target molecules, a further modification at the EK<sub>inv.</sub> scaffold was introduced: sequence randomization of a single heptad of both E and K<sub>inv.</sub> domain could in principle allow to generate a pocket in which random aminoacids can interact with our xanthine targets and bind them

We have chosen to generate the binding site in the central heptads of the C-EK<sub>inv.</sub>-C: the central heptades of E and K<sub>inv.</sub> were randomly mutagenized by substituting each residue with any natural aminoacid; the remains heptads were kept unchanged (E<sub>WT</sub> and K<sub>inv.WT</sub>) (Figure 5.12).

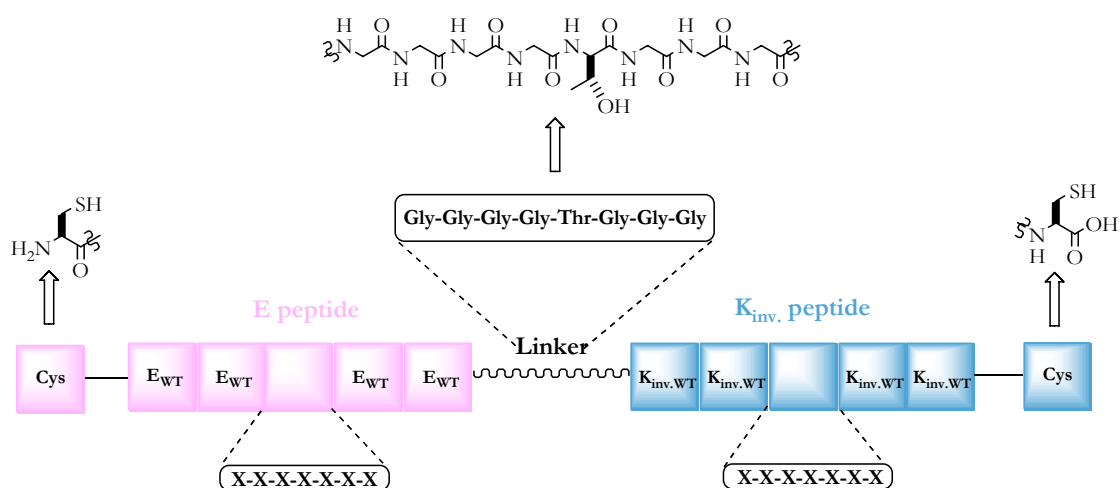


Figure 5.12

<sup>151</sup> Steinmetz M.O., Selesorov I., Matanjek W.M., Honnappa S., Janka W., Missimer J.H., Frank S., Alexandrescu A., Kammerer R.A. *PNAS*, **2007**, *104*(17), 7062-7067.

<sup>152</sup> Salwiczek M., Koksich B. *ChemBioChem*, **2009**, *10*, 2867-2870.

<sup>153</sup> Leuy Y., Cho S., Shen T., Onuchic J.N. Wolynas P.G. *PNAS*, **2005**, *102*(7)

## Chapter 6: C-EK<sub>inv.</sub>-C libraries

After having designed the topology of the C-EK<sub>inv.</sub>-C peptide receptor, we have started the development of a library of receptors.

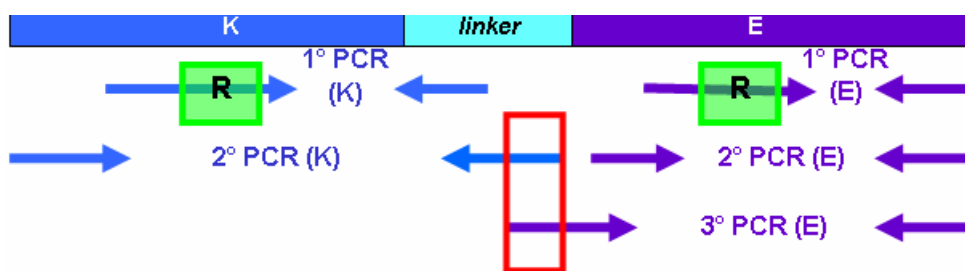
The libraries were developed by phage-display technique in collaboration with the group of prof. Marzari at the Dept. of Life Sciences of our University; using this biomolecular technique a large collection of cloned peptides was obtained and selected against immobilized theophylline-biotin (8) and caffeine-biotin (9) by ELISA-tests in order to identify and characterize the best binders.

Further characterization of several binders was obtained by sequencing the peptides, by them in proper expression system or by expressing. Dot blot experiments, in which peptide receptors were immobilized on a nitrocellulose membrane and the binding with target analytes was identify using theophylline-HRP (21a) and caffeine-HRP (22a) and the peroxidase catalyzed chemiluminescent reactions.

### 6.1 Phage-library construction.

Randomized peptide libraries were obtained by phage-display technique. In principle by randomizing a the central heptad of both E and K domain, one could obtain  $1,63 * 10^{18}$  different peptides receptors, and of course any available technique while allow to obtain only an extremely small fraction of such a chemical space.

The starting point for library construction is represented by the primer design. In our case the randomization of central heptades was obtained separately for E and K peptides, by PCR (Polimerase Chain Reaction) using E<sub>wt</sub> and K<sub>wt</sub> as template (Figure 6.1):



**Figure 6.1:** representation of amplification in PCR strategy. K chain needs of 2 cycling and E chain 3. In green, randomized primers were introduced in the central heptad; in red, common portion in the final product of each chain.

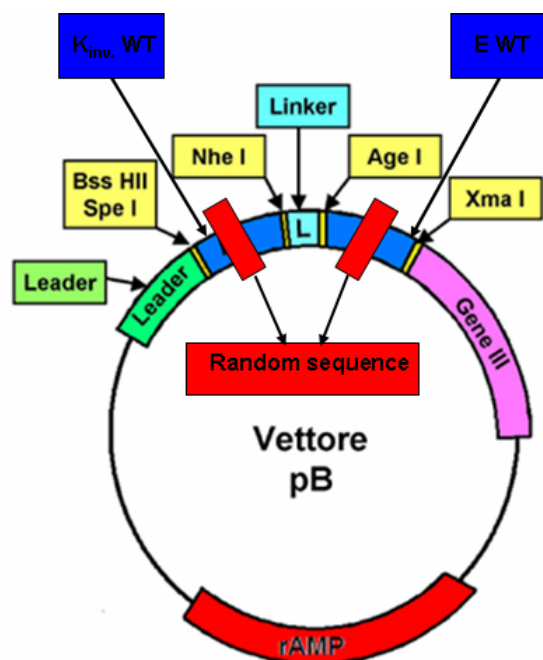
Primers used for randomization are made by a region for the pairing with E<sub>wt</sub> and K<sub>wt</sub> and a randomized region; in the first PCR reaction, a “random sense” primer and a reverse primer were used in order to introduce the random region with the first and to allow the pairing at the last non randomized heptad of the chain with the second primer.

The random sense primer was introduced in the third heptad of each chain and its random region contains the triplet NNK (N=A,T,C,G; K=G,T) in order to generate any amino acid code and increase the variety of random region.

After having synthesized the two chains separately, the E and K random domain were assembled in the common region of the linker which was previously inserted in both the strains, in the absence of the external primers. After nine PCR cycles, external primers are added in order to allow the amplification of the whole construct, which is 300 bp long. Assembling was obtained by a PCR carried out in the presence of a mixture of a 1:1 E and K randomized chains.

For the generation of the random library, a phagemidic vector (pB vector) was used (Figure 6.2)<sup>154</sup>. It is composed by 7 different parts:

- 1- a plasmidic replication origin (ORI),
- 2- a cloning site before phage gene III,
- 3- a lacZ promoter for the expression of the construct,
- 4- phage gene III which codes for protein 3 on the phage capsid,
- 5- the leader sequence for the secretion of scaffold in the periplasmic region,
- 6- the gene for  $\beta$ -lactamase which carries ampicillin resistance to the infected bacteria,
- 7- the sequence of C-EK<sub>inv.</sub>-C.



**Figure 6.2:** schematic representation of pB vector used for library construction.

<sup>154</sup> Chang C.Y., Kemp P., Molineux I.J. *Virology*, 2010, 398(2), 176-186.

In the last step of library construction, about 320 ng of the fragment obtained through PCR reaction were purified and then cloned in to the purified pB vector using T4 DNA ligase as enzyme for the ligation step.

This final library was composed by  $2 * 10^8$  elements and such a numerosity was considered satisfactory.

The library was checked first for the proper length of C-EK<sub>inv</sub>-C. 23 clones were randomly selected, amplified and composed by agarose electrophoresis run. 18 out of 23 clones were expressing a correct length fragment. The library diversity was also evaluated by random selection of 16 clones that were sequenced (Table 6.1).

clone	K Central heptade	E Central heptade	K	E
1	WT	R	E K L A S V K	V L T W W P L
2	R	R	L L L G R R S	L R N S R Y Y
3	R	R	P G G L R R W	R A F P S I R
4	WT	R	E K L A S V K	S F L Y V F W
6	R	R	F L T L A R I	S W K S Y V Y
7	WT	R	E K L A S V K	C W Q A C F F
9	R	R	L L I C L G G	A V V W V S E
10	WT	R	E K L A S V K	R S G M K L K
11	R	R	S L V G G N W	Q C G V F V F
12	R	R	P S S V G F C	C V V K M V F
13	WT	R	E K L A S V K	F I L F A L F
14	R	R	C F R W M V F	D L L G G I I
15	R	R	V L L F D M R	F V C F G T W
16	R	R	V F Q C V L G	R A C S L V C

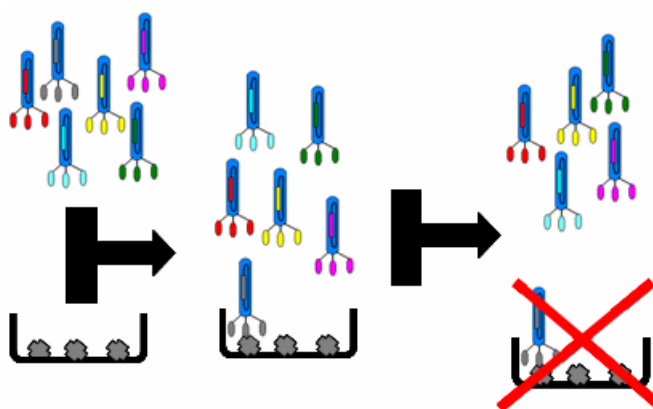
Table 6.1

Five clones resulted as correctly randomized in E domain, but not in K, while all the other clones were randomized in both domains. The presence of non randomized K domains is likely due to a contamination from K<sub>WT</sub> that has been used as template. However, since also peptides randomized of the E domain only are in principle capable of generating binding areas for our small molecules, we have decided to keep the whole library and to begin selection runs towards xanthine.

## 6.2 Selection of xanthine-binding clones.

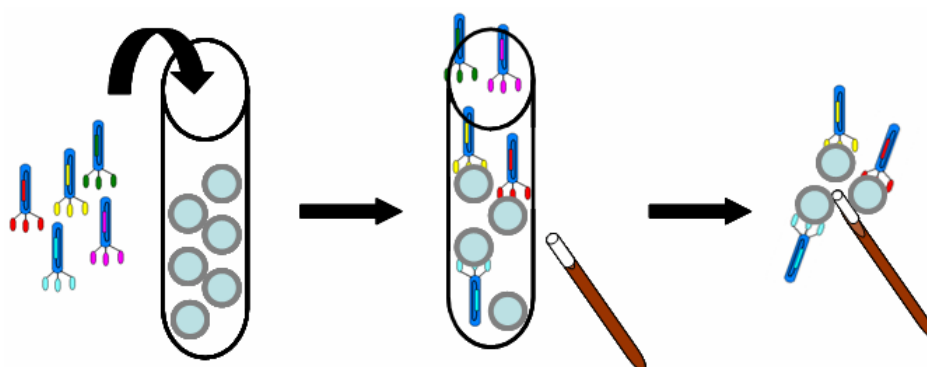
This library was thus selected towards immobilized target molecules caffeine and theophylline. The selections were carried out using both the caffeine/theophylline-biotin conjugates (8, 9) and the immobilization was obtained using the biotin-streptavidin system.

Generally, phages can show a certain background reactivity against streptavidin or other reagents used for the selection test, and give a high level of false positivity during the selection runs. In order to reduce the occurrence of such false-positives, non specific phages were eliminated via a pre-selection step (Figure 6.3) over free streptavidin beads. The fraction of unreactive phages was then used to select binders.



**Figure 6.3:** Initial pre-selection. Phages were incubated on streptavidin only and the other reagents used in the experiment; unblocked phages were recovered for the selection, while blocked phage was eliminated.

The selection of phages was performed using streptavidin-coated magnetic beads<sup>155</sup>; beads were incubated with biotin analyte in the first step and then with the phages. The phages blocked on the beads surface are able to recognize and bind the analyte and can be separated from the other non specific phage through the use of a magnetic rod (Figure 6.4).



**Figure 6.4:** selection of phages with magnetic beads covered of streptavidin and biotinilated analyte.

Final elution with unbiotinylated caffeine allows to recover phages and to sequence cloned peptides in order to obtain the primary sequence of binders. In our case, four cycle of selection have been made; at the end of fourth cycle, 86 clones were isolated and tested against biotinilated analytes (8, 9) and free amino modified xanthines (10, 11, 13) in different ELISA experiments.

<sup>155</sup> Shlyapnikov Y.M., Shlyapnikova E.M., Morozov V.M. *Anal. Biochem.*, 2010, 50(2), 235-239.

The selection required an initial amount of  $10^{12}$  phages in 250  $\mu\text{L}$  of phosphate buffer. Preselection was achieved also towards defatted milk in order to remove most of non specific protein reactivity. The phage suspension was then incubate over 10 ng/mL immobilized xanthine. After repeated washing cycles, binders were eluted with a 100 ng/mL caffeine solution, and the eluted phages were then used to infect DH5 $\alpha$ F<sup>'</sup> bacterial strains. Bacteria were then grown up over plates and recovered to start a new selection cycle. Four selection cycle were carried out, and at the end of fourth cycle, 86 clones were selected.

### 6.2.1 The ELISA tests.

Enzyme-linked immunosorbent assay, also called ELISA, is a widely used<sup>156</sup> immunochemical technique to detect the presence of a receptor towards a target molecule. The target molecule caffeine or theophylline in our case is immobilized over a solid surface, and a solution of the potential receptor/binder (an antibody, a membrane protein, a whole phage displaying a binder on its surface as in our case) is incubated. After washing, a secondary enzyme-labelled reagent is added (typically an antibody, in phage display an antibody towards one of the phage proteins, labelled with horseradish peroxidase). Finally a substrate for an enzyme-catalyzed colorimetric reaction is added.

In our case, phage-ELISA test was used for a first screening with the 86 isolated clones. Polystyrene microtiter plate were coated with streptavidin; caffeine-biotin (9) was immobilized on solid surface exploiting streptavidin-biotin system and a solution of each selected phage clone displaying randomized peptides was added.

After washing, the concentration of bound phage in a single well was determined with an antibody specific for an external protein of phage (Figure 6.5).

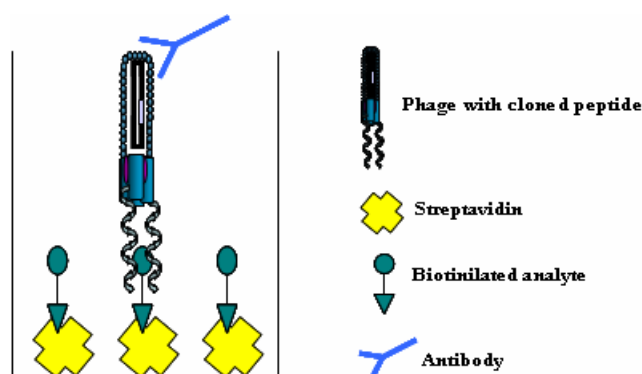


Figure 6.5: schematic representation of phage-ELISA test.

Six out the 86 tested clones demonstrated a good affinity to caffeine, although with a rather high cross-reactivity towards streptavidin (Figure 6.6). Figure 6.6 reports a typical ELISA result

<sup>156</sup> Nielsen U.B., Geierstanger B.H. J. *Immunol. Meth.*, 2004, 290(1-2), 107-120.

obtained towards biotin/avidin immobilized caffeine. The same concentration of phage was used in all cases; the negative control is represented by the Wild-Type non mutated C-EK<sub>inv.</sub>-C scaffold, while the positive control is given by a caffeine-binding peptide obtained in previous studies.<sup>157</sup> Despite the occurrence of a significant amount of reactivity towards streptavidin alone, specific caffeine binding can be observed.

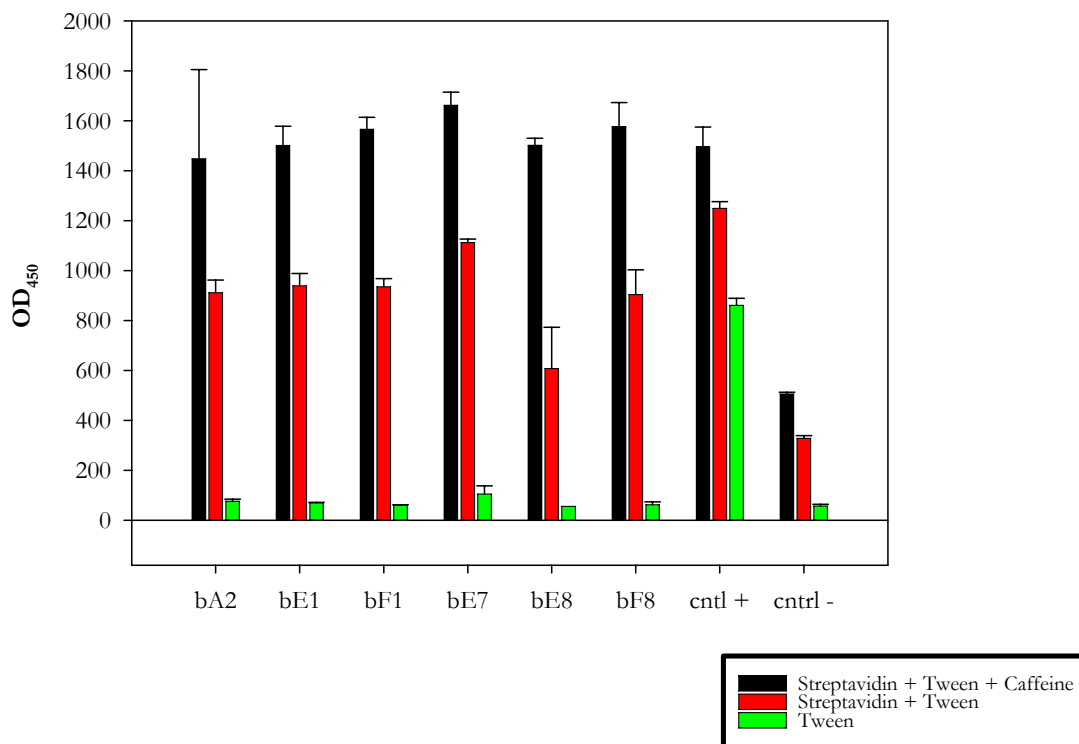
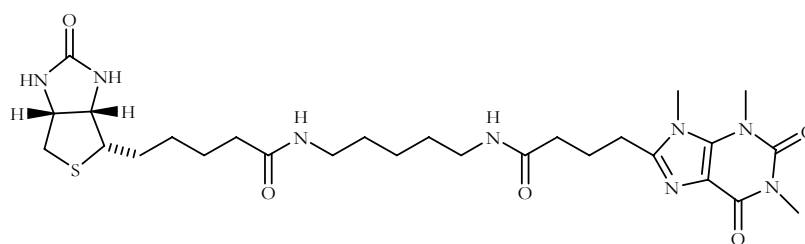
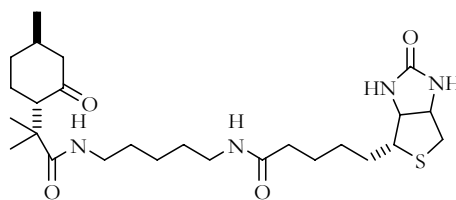


Figure 6.6

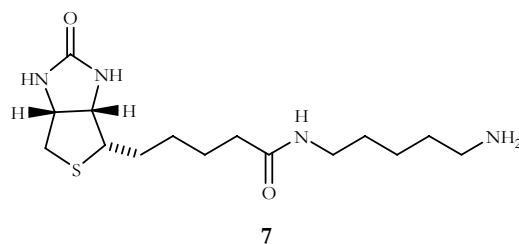
In parallel, clones were tested also in another phage-ELISA system using immobilized caffeine-biotin (**9**), menthone-biotin (**26**) and linker-biotin (**7**) (Scheme 6.1) in order to analyse the cross-reactivity of clones against unrelated molecules and against the linker alone (Figure 6.7):



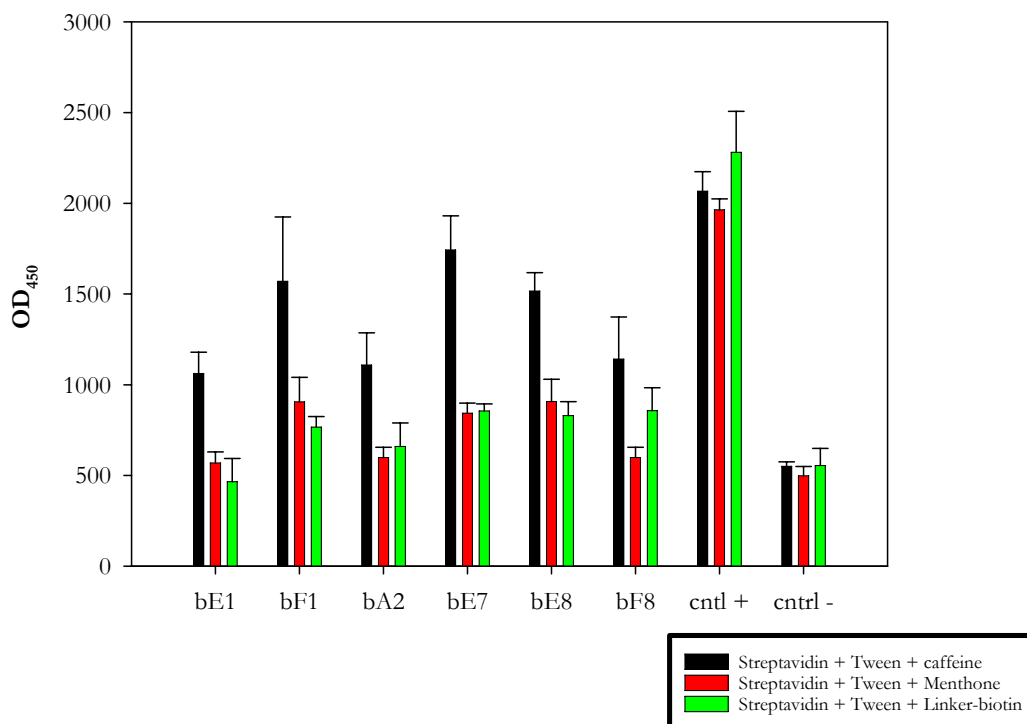
9



26



Scheme 6.1



**Figure 6.7:** comparison of three different phage-ELISA using caffeine-biotin, menthone-biotin and linker-biotin as target analytes.

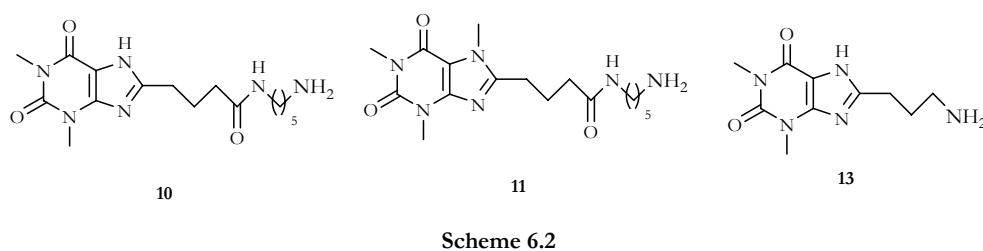
All the six clones show a partial cross-reactivity towards menthone but the interaction with menthone leads to signals that are comparable to those obtained with the linker alone, and thus it can be considered negligible. This is an important information, as it means a potentially good selectivity of our peptides in favour of aromatic xanthenes; however, clones show cross-reactivity also against linker-biotin that represents a portion of the overall conjugates.

Such a level of interfering interaction with the linker used to immobilize target molecules, is a very common phenomenon which is observed many times also during the selection of antibodies from phage display libraries, and more generally during any kind of generation of biological binders that require immobilization or bioconjugation of the target.<sup>158</sup>

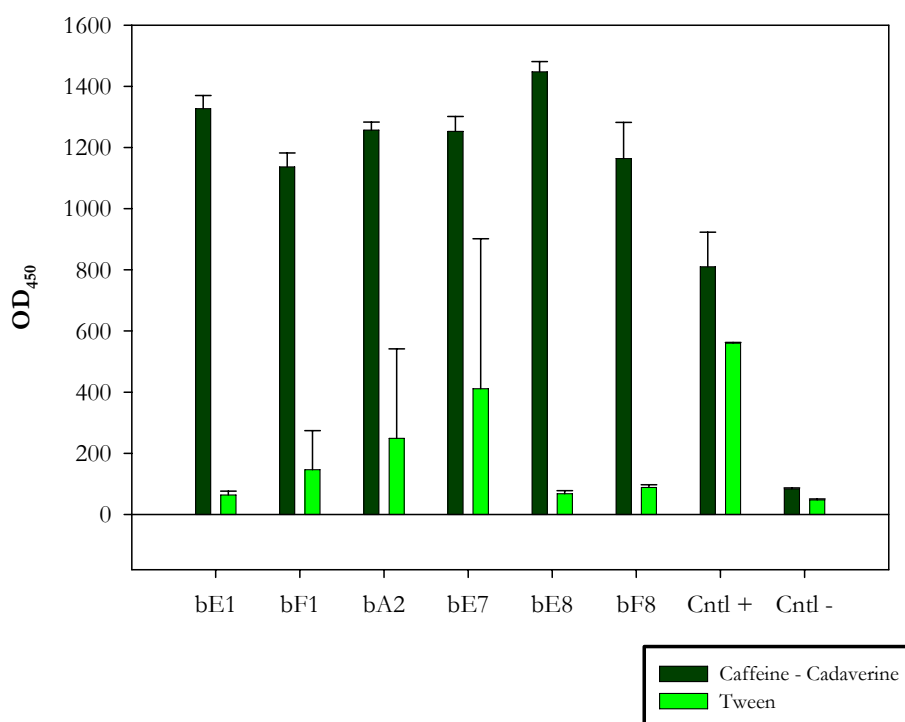
In order to clean out this cross-reactivity toward linker, the clones were tested again against caffeine derivatives with different linker having a free terminal amino group (**10**, **11**, **13**).

<sup>158</sup> a) de Villiers S.H.L., Lindblan N., Kalayanov G., Gordon S., Boraznekok I., Malmfelt A., Marcus M.M., Svesson T.H. *Vaccine*, **2010**, *23*(7), 34-37. b) Stanberg J., El Ayed M., Wisztorski M., Day R., Fournier I., Salzet M. *Anal. Chem.*, **2009**, *81*(22), 9512-9521.

In this case modified caffeine and theophylline were directly linked onto a properly activated solid surface under (Scheme 6.2). In a first run, the clones were tested towards the caffeine-cadaverine derivative **11** (Figure 6.8):

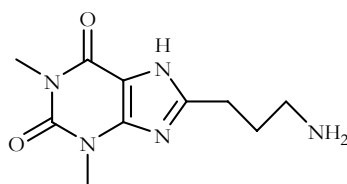


As expected, under the new ELISA conditions the clones show a clean recognition of caffeine with a high difference over the background signal given by Tween, the surfactant used to saturate the solid plane in the ELISA test. This means more safely that the observed signals are due to specific recognition only for the aromatic target compound (Figure 6.8):



**Figure 6.8:** new phage-ELISA obtained from amino conjugate. Cntl+ and cntl- were the same used in the previous ELISA experiment.

In order to further reduce any linker effect, another phage-ELISA was developed towards compound **13** (Scheme 6.3); in this case linker dimensions were reduced from eleven atoms to four and the clones tested were only bE1 and bE8, the best isolated receptors from the previous experiment, with a good affinity toward caffeine and with very low background (Figure 6.9):



Scheme 6.3

This test was carried out using a theophylline derivative instead of a caffeine use, in order to join insight also on the caffeine/theophylline cross-reactivity.

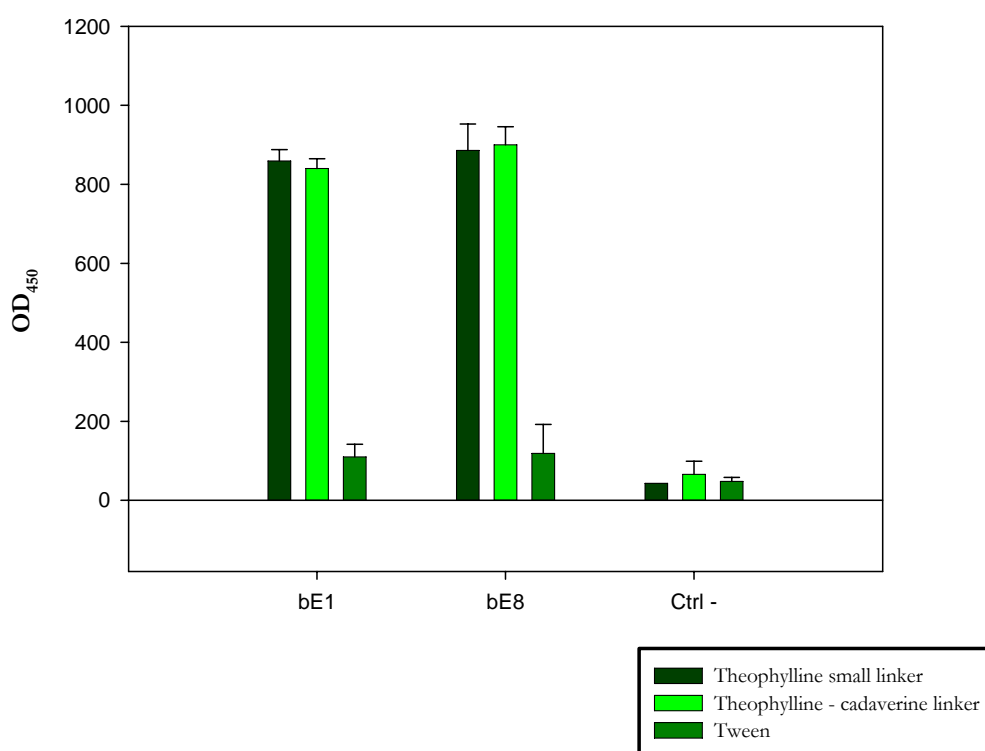


Figure 6.9: comparison results between ELISA test.

Recognition of xanthines by E1 and E8 is thus not affected by the linker length, no by the methyl group at position 3 of the xanthine ring.

### 6.2.2 Measurement of the dissociation constant of the peptide/caffeine complexes

After having selected the E1 and E8 binders, we have measured the dissociation constant of their caffeine complexes, using an ELISA methodology described by Loomans and colleagues.<sup>159</sup>

The dissociation constant is measured by coating a series of ELISA plates with different concentrations of caffeine (namely 2, 5, 5, 10 and 20  $\mu\text{g}/\text{mL}$ ) and then by incubating different concentrations of peptides over any concentration of immobilized caffeine. A peptide

<sup>159</sup> Loomans E.E., Roelen A.J. *J. Immunol. Meth.*, **1995**, *184*(2), 207-217.

concentration/response plot is obtained at any caffeine concentration (Figure 6.10: the series of curves given by E1, Figure 6.11: the response obtained with E8):

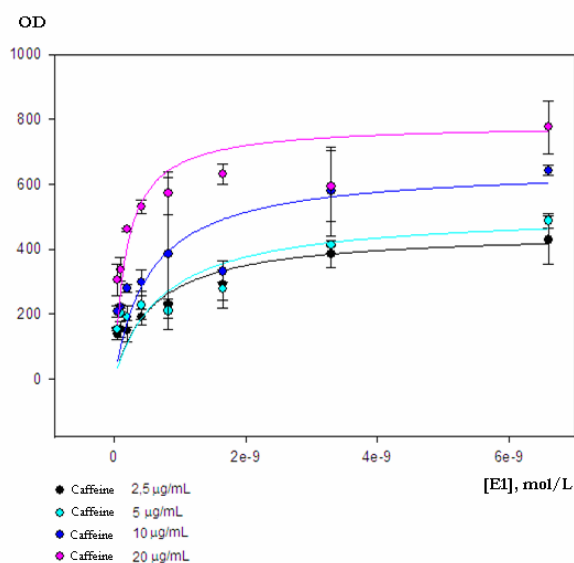


Figure 6.10

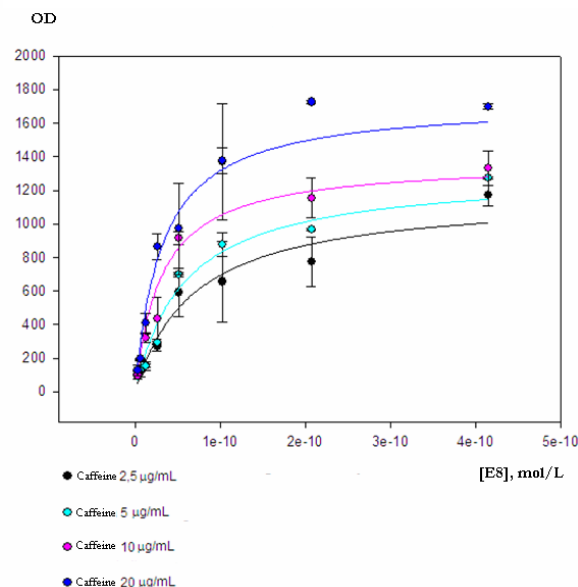


Figure 6.11

The experimental points are then submitted to a nonlinear regression analysis that yield hyperbolic fittings and  $OD_{50}$  calculated values. A series of  $K_d$  values is then obtained by submitting any possible couple of peptide concentrations corresponding to the calculated  $OD_{50}$  values, to equation 6.1:

$$K_d = (n-1)/2(n[B]-[A]) \quad \text{eq. 6.1}$$

where  $[B]$  and  $[A]$  are the peptide concentrations giving the  $OD_{50}$  signal at two different caffeine concentrations, and  $n$  is their ratio. The average of such calculated  $K_d$  values gives finally a  $K_d$  of  $0.70 \pm 0.12$  nM and of  $6.90 \pm 0.17$  nM for E8 and E1 caffeine complexes, respectively.

A remarkable affinity towards immobilized caffeine is thus measured by the ELISA experiments, and the nanomolar dissociation constants are comparable with these usually observed in antibody-hapten complexes.

With this series of phage-ELISA experiments we have identified two new potential peptides receptors for caffeine and theophylline molecules.

### 6.2.3 Characterization of bE1 and bE8 sequence.

By PCR amplification and sequencing of the DNA strands contained into the phage expressing E1 and E8, their aminoacid sequence was obtained (Table 6.2):

	<u>E random</u>	<u>K<sub>inv.</sub> random</u>
<b>E1</b>	<b>FLRRRIR</b>	<b>EKLASVK</b>
<b>E8</b>	<b>QSRRLKR</b>	<b>EKLASVK</b>

Table 6.2

In both the peptides the K<sub>inv.</sub> random sequence correspond to that of the wild-type scaffold, and the whole of diversity is found into the E randomized heptad only. The sequence of E1 and E8 at this level are nevertheless similar and very acidic: four arginine residues are found in E1, three arginine and one lysine residue are found in E8. The presence of such a high number of arginines is rather surprising. However, an analysis of xanthine binding motives inside enzyme binding sites reveals that strong interactions between arginine and xanthine carbonyls are often found. This is the case of xanthine oxidase, where Arg 880 plays a key role in orientating xanthine and stabilizing its complex with the catalytic molybdenum ion of the enzyme.<sup>160</sup>

Also in xanthine dehydrogenase from *Rhodobacter Capsulata*, arginine 310 is essential, and its mutation to methionine leads to a 10<sup>3</sup> reduction in both the thermodynamic and kinetic parameters for a series of purine substrate<sup>161</sup>; arginine is essential also in caffeine inhibition of ionotropic glycine receptors.<sup>162</sup> With the sequence of E1 and E8 in hand, we have tried to build up also a model of the two caffeine complexes.

A Predict-Protein<sup>163</sup> sequence analysis carried out on the two peptides yields a random-coil secondary structure prediction at the mutated E sequences in both the peptides, while  $\alpha$ -helix conformation is confirmed as highly probable along the remaining sequence. This allow to take as a starting hypothesis the possible creation of a small binding pocket onto a largely conserved coiled-coil skeleton. As the starting geometry for the coiled-coil scaffold, we have chosen the homology model described in chapter 5 page 58. The linker chain, the terminal cysteines and their disulfide bridge were added first, and the topology was submitted again to a relaxation MD cycle as we have previously described. The sequence were then mutated to obtain the first models of E1 and E8. The models were submitted again to a final-multistage relaxation session<sup>164</sup>, and then docking of caffeine was performed using a GRID-flexible docking protocol.<sup>165</sup> We have also docked the linker-modified caffeine and thephylline into the peptide binding site, in order to evaluate also the interaction s towards the linker, and its cooperation in recognizing immobilized xanthines.

<sup>160</sup> Pauff J.M., Zhang J., Bell C.E. Hille R. *J.Biol.Chem.*, **2008**, 283(8), 4818-4824.

<sup>161</sup> Pauff J.M., Heman C.F., Junemann N., Leimkubler S., Hille R. *J.Biol.Chem*, **2007**, 282(17), 12785-12790.

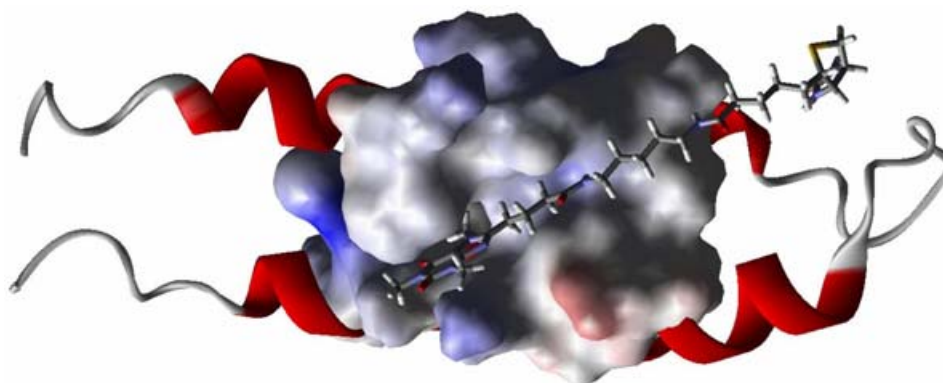
<sup>162</sup> Duan L., Yang J., Slaughter M.M. *J. Physiol.*, **2009**, 587(16), 4063-4075.

<sup>163</sup> Rost B., Yachdav J. *Nucl. Acids Res.*, **2004**, 32, W321-W326.

<sup>164</sup> Berti F., Forzato C. *Tetrahedron Asymmetry*, **2005**, 16, 1091-1102.

<sup>165</sup> Goodford P.J. *J. Med. Chem.*, **1985**, 28, 849-857.

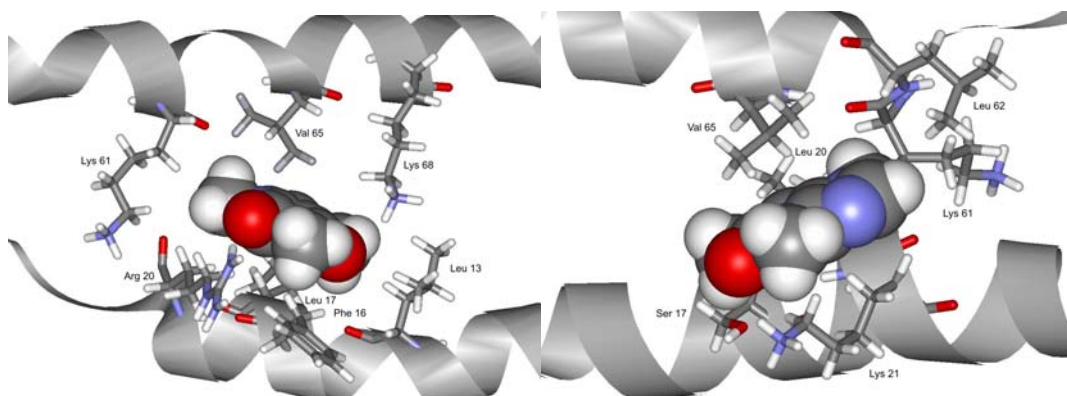
The more stable docking solution obtained after clustering of the docking poses were then optimized again in a periodic box of water molecules to obtain the final models. The final geometry of the E8-linker-caffeine complex is reported as an example in Figure 6.12:



**Figure 6.12:** E8 and caffeine-biotin complex.

A rather deep caffeine-binding site is observed. Caffeine is buried to a deeper extent at its imidazole system and at methyl in position 7. Carbonyl 2 and partly methyl group 1 and 3 are conversely exposed to the solvent. The site is contiguous to a hydrophobic cleft that corresponds to the zipper interface hosting the hydrocarbon chain of the linker.

A detailed view of the caffeine interactions is given in Figure 6.13 for E1 and in Figure 6.14 for E8 complexes.



**Figure 6.13:** Detail of E1 peptide and caffeine interaction. **Figure 6.14:** Detail of E8 peptide and caffeine interaction.

In the E1 complex, hydrophobic contacts are established between caffeine and the hydrocarbon chains of Lys 61, Val 65, Lys 68, Leu, 13, Leu 17, Arg 20. Phenylalanine 16 is at  $\pi$ -methyl distance to caffeine methyl 1 while Arg 20 is at hydrogen bond distance to carbonyl 2.

In the E8 complex the side chains of Val 61, Leu 62, Leu 20, Lys 61, Lys 21, Ser 17 are at hydrophobic contact distance with caffeine, while the distances between caffeine carbonyl 2, the

hydroxy group of Ser 17 and the cationic head of Lys 21 are consistent with an hydrogen bond network.

We have calculated complexation energy values, obtained by adding a solvation energy term (obtained from a discrete water model) to the conformational, non bonded and electrostatic energy terms arising from the AMBER force field (equation 6.2)

$$\Delta E_{\text{compl.}} = \Delta E_{\text{conf.}} + \Delta E_{\text{elast.}} + \Delta E_{\text{nb}} + \Delta \Delta E_{\text{sol}} \quad (\text{equation 6.2})$$

The complexation energy are reported in Table 3.2, relatively to the calculated complexation energy for the E1/caffeine complex.

	E1	E8
	$\Delta E_{\text{compl}} / \text{Kcal mol}^{-1}$	$\Delta E_{\text{compl}} / \text{Kcal mol}^{-1}$
Caffeine- <i>linker</i>	-44.4	-62.3
Caffeine	0	8.5
Teophylline	3.9	7.1

Table 3.3

In the lack of a calibration set of experimental data, such complexation energies can not be used to predict binding constants; however, the model predicts a better interaction for the E1 complexes, with a very large gain in binding energy for the linker-modified target molecules.

### 6.3 Expression of E1 and E8 in *E. Coli*

After the first ELISA characterization of E1 and E8, the peptides have been expressed also in a soluble form suitable for large scale bacterial production and for simpler purification protocols. According to commonly adopted techniques, the peptides were produced in fusion with the *E. Coli* maltose binding protein, which is a part of maltose/maltodextrin system of *Escherichia Coli*. Maltose binding protein is a part of maltose/maltodextrin system of *Escherichia Coli*. As different, MBP has an approximate molecular weight of 42.5 kDa, and it widely used to increase the solubility of recombinants proteins expressed in *E. Coli*.<sup>166</sup> When the protein of interest is expressed as a MBP-fusion protein its solubility is often enhanced, and by this way, relatively high concentration of peptide can be found in the bacterial culture supernatant. In addition, MBP itself can be used as an affinity tag for purification of recombinant proteins. The fusion protein binds to amylose columns while other proteins flow through; the MBP-protein fusion can be

<sup>166</sup> Chater K.F., Birò S., Palmer T., Schrempf H. *FEMS*, **2010**, *34(2)*, 171-198.

purified by eluting the column with maltose. Once the fusion protein is obtained in purified form, the protein of interest is often cleaved from MBP with a specific protease.<sup>167</sup>

Cloning of the peptides at MBP was performed with a pMAL vector (Figure 6.15), E1 and E8 fragment were produced by PCR reactions and then transferred in pMAL vector using T4 DNA ligase.<sup>168</sup>

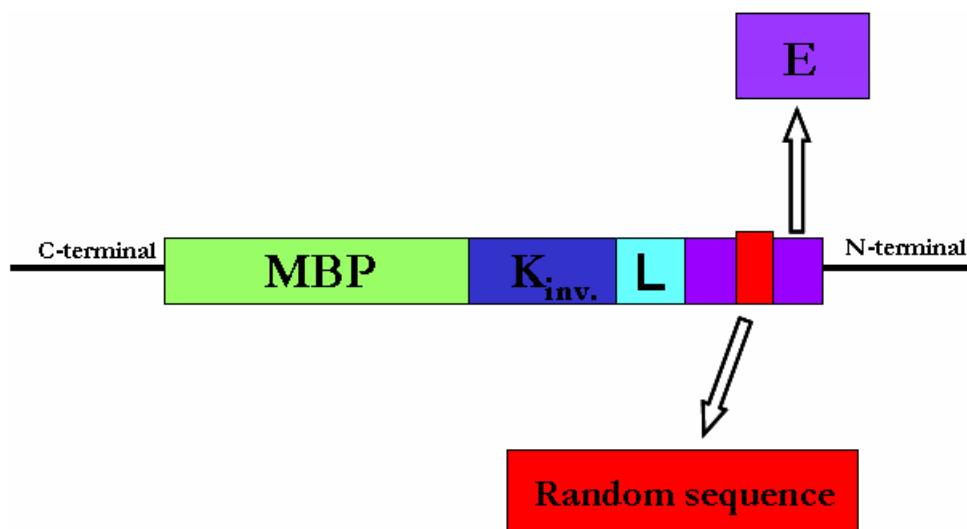


Figure 6.15: E1, E8 in pMAL vector

The MBP-E1 and MBP-E8 fusion proteins were then produced in NM522 *E.Coli* cultures and purified by affinity chromatography, followed by dialysis in order to eliminate the maltose used for the elution; the purified MBP-peptides were then tested in two different systems to evaluate the reactivity towards caffeine and theophylline.

### 6.3.1 Dot-Blot

Dot Bolt is a molecular biology technique used to detect binding between biomolecules.<sup>169</sup> In a Dot Blot experiment the biomolecule to be detected is applied directly on a membrane as a dot. This step is followed by incubation with a labelled binding partner.

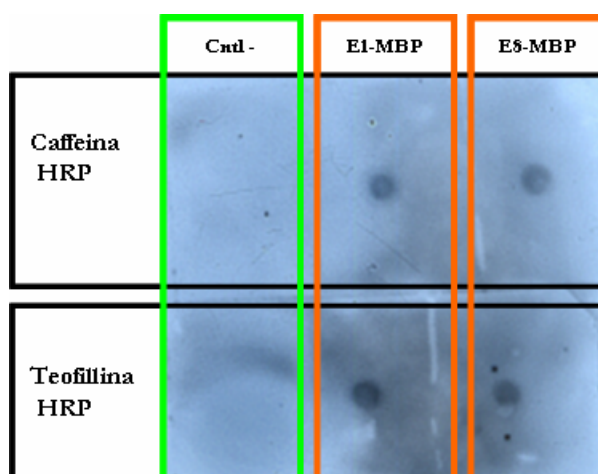
In our case, a small amount of MBP-E1 and MBP-E8 (about 2 µg of protein) was immobilized over a nitrocellulose membrane and then incubated with caffeine-HRP and theophylline-HRP, the conjugates of xanthine with horseradish peroxidase described in the previous sections. After a washing step, membranes were incubated with an enhanced chemiluminescence system (ECL) in which horseradish peroxidase oxidizes luminol to 3-aminophthalate in the presence of several

<sup>167</sup> Malhotra A. *Meth.Enzym.*, **2009**, 463(C), 239-258.

<sup>168</sup> Riggs P. *Molecular Biotech.*, **2000**, 15, 51-63.

<sup>169</sup> a) Vincitelli R., Cambillau C. *Bio Tech International*, **2005**, 6, 14-16. b) Andoh A., Benno Y., Kanauchi O., Fujiyama Y. *Curr. Pharm. Design*, **2009**, 15(18), 2066-2073.

cofactors.<sup>170</sup> The oxidation of luminol is accompanied by emission of low intensity light at 428 nm; this emission is then boosted up to 1000-fold by the cofactor, making the light easier to detect and increasing the sensitivity of the reaction.



**Figure 6.16:** E1-MBP and E8-MBP clones tested with HRP labeled analyte. EK<sub>inv.</sub> wild-type was used as Cntl-.

Figure 6.16 reports a typical Dot-Blot output, obtained by impression of photographic paper by the emitted light. A clean interaction between E1 and E8 peptides and xanthines analytes is detected with the appearance of a blu spot; EK<sub>inv.</sub> wild-type, used as negative control (Cntl-), does not bind caffeine or theophylline.

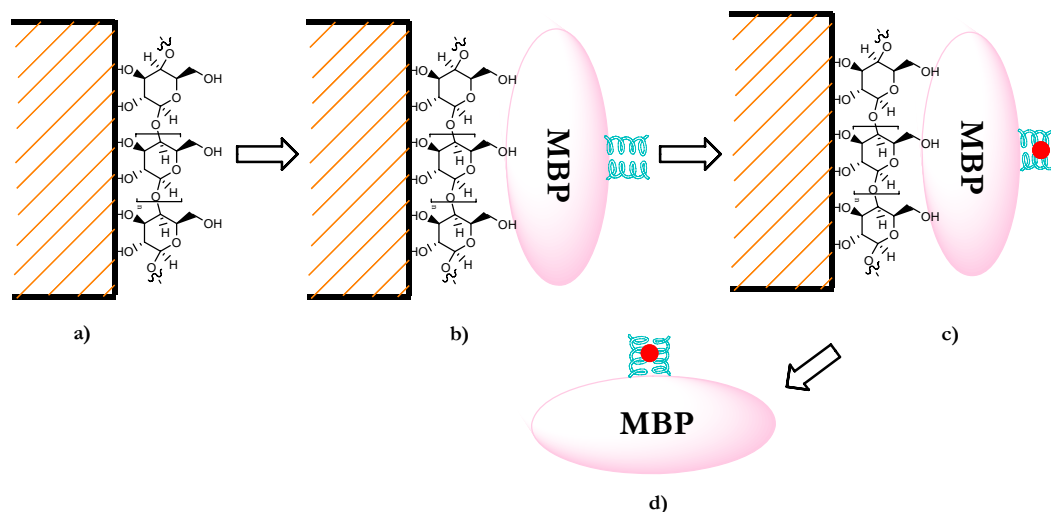
With this important result, we have confirmed that interaction between E1 and E8 peptides and caffeine or theophylline is not affected by the expression system of the binders peptides.

### 6.3.2 E1 and E8 as affinity chromatography stationary phases

The high affinity of E1 and E8 towards caffeine, lead us to the idea of testing their ability as stationary phases in an affinity chromatography system for the purification/removal of caffeine from solution.

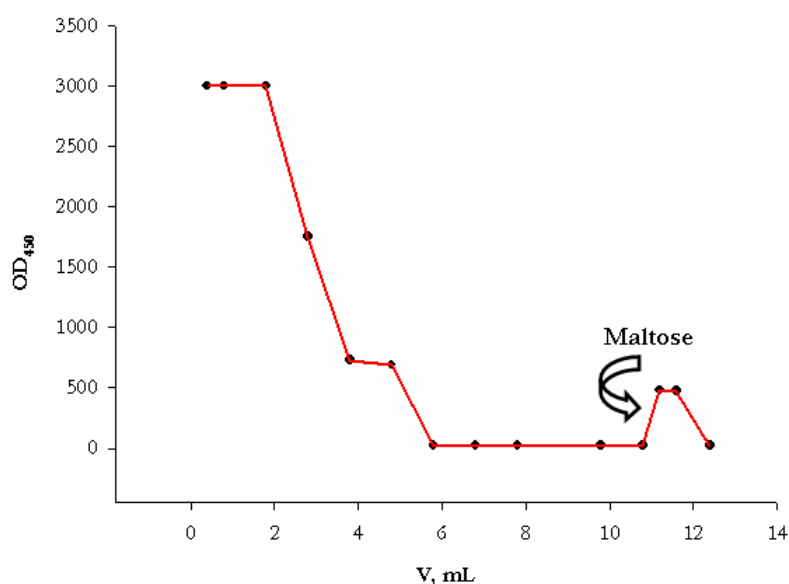
The MBP expression system provides an ideal way to set up a stationary phase, due to the ease of immobilization of MBP constructs over amylose beads. We have thus exploited the affinity of MBP for amylose in order to block the MBP-E1 and MBP-E8 on a amylose resin in a chromatography column. In order to generate an analytical instrument for caffeine and theophylline detection, water solutions of these xanthine were flowed through the column and the xanthine column retention after washing was estimated in different ways (Figure 6.17).

<sup>170</sup> Newmann G.R., Jasani B. *J. Pathology*, **1998**, 186(2), 119-125.



**Figure 6.17:** schematic representation of affinity column. Resin of amylose was coated on surface of column (a), MBP-peptides was blocked through amylose-MBP interaction (b) and a solution with our analyte (red) was flowed under the column. E1 and E8 bound the analyte in the randomized region (c) and at the end MBP-peptides-analyte was removed from the column using maltose solution (d).

In a first series of experiments, the theophylline-HRP conjugate was used as analyte; without taking care, in this stage, of proper elution conditions for caffeine (i.e. of conditions allowing the breakup of the peptide –caffeine complex without destroying the stationary phase) we have initially simply eluted the whole MBP-peptide-caffeine-HRP complex by adding maltose to the mobile phase after a proper washing volume, and then we have measured peroxidase activity in any fraction by a colorimetric test, using 3',3',5',5' tetramethylbenzidine (TMB) as substrate. The resulting chromatograms are reported in Figure 6.18:



**Figure 6.18**

The results are excellent, with a release of caffeine well over the background given by the MBP-WT peptide in both cases. We have also tested the effect of a pH shock on the E8 column, and

we have observed that lowering of the pH down to 3 with glycine buffer does not affect the stability of the system, and no caffeine is eluted under these conditions.

The loading capacity of the E8-MBP system was also evaluated by measuring with a reference HPLC methodology.<sup>171</sup> The amount of bound caffeine over one ml of amylose resins loaded with 50 µg of the MBP-E8 protein. 8.42 µg of caffeine were dissolved in TRIS buffer and the solution was then eluted over the MBP-E8 column. The column was then extensively washed and the amount of recovered caffeine was measured in any fraction. 8.04 µg of caffeine were present in the eluted fraction. A control experiment carried out on a resin loaded with the wild type MBP construct gave 8.16 µg of caffeine recovery. A specific loading capacity of 0.12 µg caffeine per 50 µg of MBP-E8 can thus be estimated. This corresponds roughly to a realistic ratio of 0.6 moles of caffeine per mole of binder peptides.

---

<sup>171</sup> Novakova L., Vlckova H. *Anal. Chim. Acta*, **2009**, 656(1-2), 8-35.

## Chapter 7: K library, towards a simplified system.

The development of C-EK<sub>inv.</sub>-C library has represented the first step for the achievement of small receptor peptides able to identify caffeine and theophylline. E1 and E8 peptides are promising structures; the affinity for our two target analytes were tested in different way and in all cases a relevant interaction takes place. However, based on statements on usage facility and structure stability given in Chapter 2, we have considered the possibility of further reduction of C-EK<sub>inv.</sub>-C scaffold to a very simplified and short structure.

In order to reduce as more as possible the scaffold length and with the intention to keep nevertheless the  $\alpha$ -helix as principal motif, we have decided to use as scaffold only the K chain of the EK coil. Although nor E neither K wild type chain are reported to be able of organized folding alone (see chapter 2), and indeed they have been designed in order to avoid self-organization and homodimerization, we have reasoned that the energy balance between homo- and hetero-organization in coiled-coils, is generally very subtle, and for this reasons, mutations in the central heptade of one of the two chains could in principle lead to self-organizing structures, including folded monomers, dimers or even more complicated aggregates (Figure 7.1):

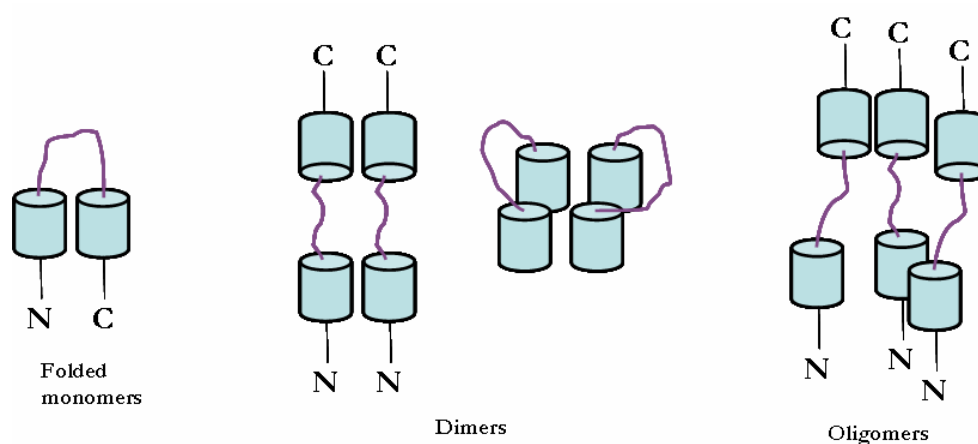


Figure 7.1

Following an approach similar to that used for C-EK<sub>inv.</sub>-C randomized library, we have chosen to generate a library of K randomized chains using the phage-display technique. The K chain was chosen instead of the E chain because, on carrying out several synthetic trials on automated peptide synthesizers, we have always encountered huge difficulties in the synthesis of the E chain, while the K<sub>WT</sub> chain resulted relatively easy to be prepared. Randomization was generated only in the central heptad of K chain, again in order to identify a region of overall randomized peptide able to bind caffeine or theophylline.

K random peptides were tested against the same modified targets and conjugates used for the first library: namely theophylline-biotin **8**, caffeine-biotin **9** for phage-ELISA experiment and ELISA test with K random chains expressed in fusion with MBP.

### 7.1 K library construction.

The starting point for library construction was again represented by the choice of primers. As in the C-EK<sub>inv</sub>-C library, the randomization of K<sub>wt</sub> was achieved via PCR using K<sub>wt</sub> sequence as template. Two primers were used: the first to allow pairing at one end of K chain and the second to allow pairing at the central heptad of K. The fragment obtained by PCR reaction was purified and cloned into the pBIOS 1 phagemidic vector using T4 DNA ligase for the reaction of insertion; pBIOS 1 allows to express peptide in fusion with protein 3 of phage capsid (Figure 7.2):<sup>172</sup>

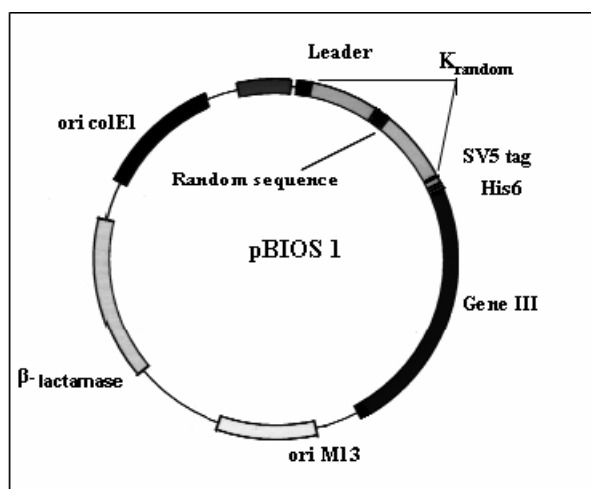


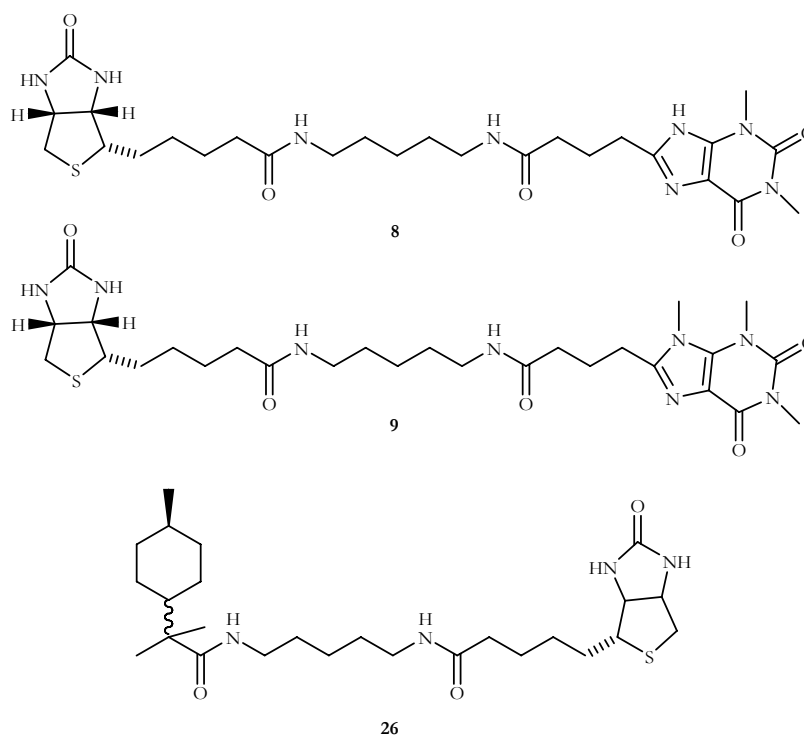
Figure 7.2

Randomization of the K central heptad lead to a  $1.1 * 10^8$  members library.

The selection of specific phages against caffeine was carried on in three rounds, again using streptavidin coated beads with caffeine-biotin as target. 96 clones were isolated from the library and tested by phage-ELISA experiments in order to check their reactivity against caffeine and theophylline.

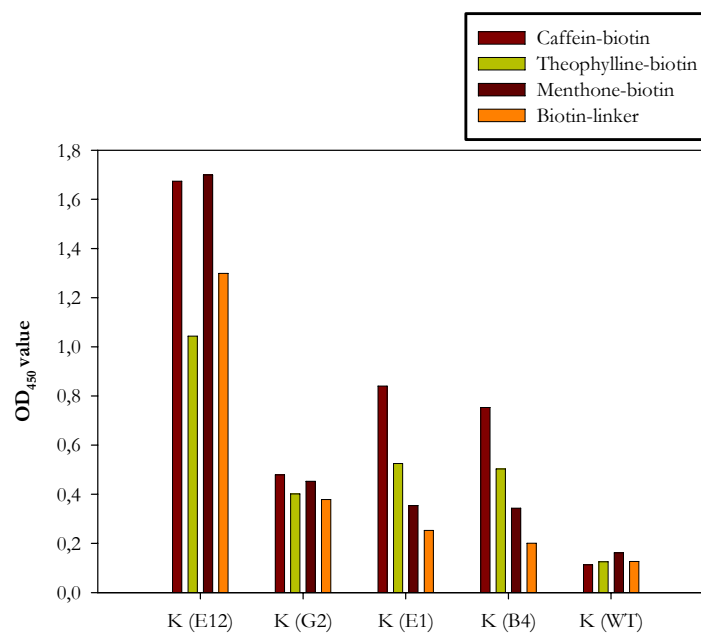
In this case, analyte was immobilized in the ELISA wells exploiting streptavidin-biotin system using modified molecules **8**, **9** and **26** (Scheme 7.1):

<sup>172</sup> Chassagne S., Laffly E., Thullier P. *Mol. Immunol.*, 2004, 41(5), 539-546.



**Scheme 7.1**

The performance of the best four clones with a good affinity toward caffeine and theophylline is reported in Figure 7.3 and compared with the background given by  $K_{WT}$ :



**Figure 7.3**

Clones named  $K^{E1}$  and  $K^{B4}$  have demonstrated a high and specific reactivity towards caffeine and theophylline and a reduced cross-reactivity towards menthone and biotin-linker; on the contrary  $K^{E12}$  lacks of specificity, as well as  $K^{G2}$ .

The four clones were then sequenced (Table 7.1):

Clone name	Aminoacidic sequence
K <sup>E12</sup>	(KVSALKE) <sub>2</sub> SERRCQQ(KVSALKE) <sub>2</sub>
K <sup>G2</sup>	(KVSALKE) <sub>2</sub> RYFRCQS(KVSALKE) <sub>2</sub>
K <sup>E1</sup>	(KVSALKE) <sub>2</sub> QFLCLMF(KVSALKE) <sub>2</sub>
K <sup>B4</sup>	(KVSALKE) <sub>2</sub> QFLCLMF(KVSALKE) <sub>2</sub>

**Table 7.1:** randomized central heptad is marked in yellow

The best performing peptides K<sup>E1</sup> and K<sup>B4</sup> resulted to be identical clones of the same sequence. K<sup>E1</sup> was thus the only peptide submitted to further characterization.

Clone K<sup>E1</sup>, K<sub>WT</sub> and also E<sub>WT</sub> were thus expressed in fusion with MBP, as it was done for E1 and E8 in the previous library.

A cysteine is present in the K<sup>E1</sup> sequence as the fourth aminoacid in the randomized sequence (position 18). In order to reduce the occurrence of high background due to disulfide bond between this cysteine and other cysteine that could be present on the phage capsid on expression protein, we have chosen to mutate the cysteine into a methionine; the new peptide sequence (scheme 7.2) was for this reason expressed in fusion with MBP in E.Coli.



**Scheme 7.2**

After purification on amylose column, the peptide was tested for its ability to bind caffeine (Figure 7.4):

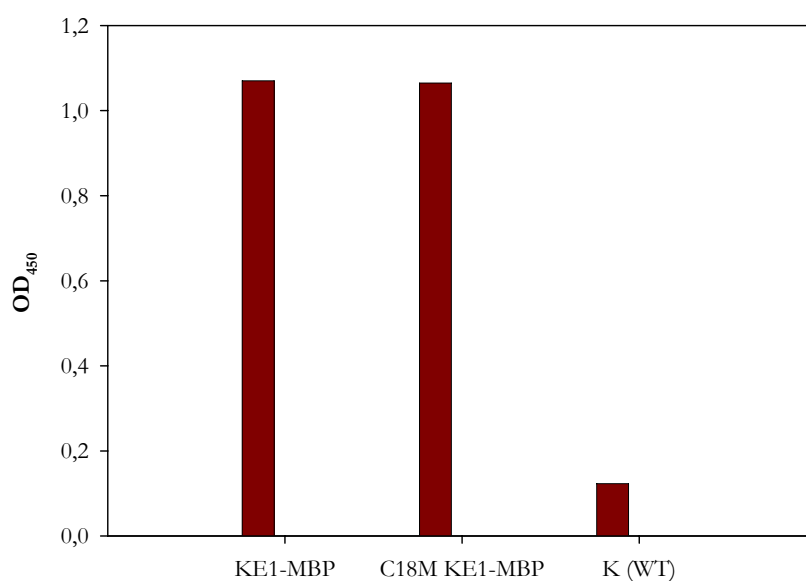


Figure 7.4

There is no difference between the reactivity of  $K^{E1}$  and its C18M mutant. Based on this finding, the original sequence of  $K^{E1}$  was abandoned, and any further characterization was carried out on the C18M mutant.

## 7.2 Synthesis and chemical characterization of $K^{E1}$

According to the design of the K library, the short length of  $K^{E1}$  allows its full chemical synthesis: the peptide was synthesized on a 40 mg scale, and this allowed a more detailed characterization of its chemical and binding properties; we have thus obtained informations on the peptide folding and on the crystal structure of the peptide that were unavailable for the members of the C-EK<sub>inv.</sub>-C library.

Currently, peptide synthesis includes a wide range of techniques and protocols that enable the preparation of materials ranging from small peptides to large proteins. The pioneering work of Bruce Merrifield who introduced solid phase synthesis (SPPS)<sup>173</sup>, dramatically changed the strategy of peptide synthesis and simplified the step of purification associated with solution phase synthesis. Moreover, solid phase peptide synthesis also permitted the development of automation and the extensive range of robotic instrumentation now available. After defining a synthesis strategy and programming the amino acid sequence of peptides, machines can automatically perform all the synthesis steps required to prepare multiple peptide sample.<sup>174</sup>

<sup>173</sup> Merrifield R.B. *J. Am. Chem. Soc.*, **1963**, *85*, 2149-2154.

<sup>174</sup> a) Sabatino G., Papini A.M. *Curr. Opin. Drug Develop.*, **2008**, *11(6)*, 762-770. b) Lam K.S., Renil M. *Curr. Opin. Chem. Biol.*, **2002**, *6(3)*, 353-358.

Solid support synthesis has obvious advantages. Large excess of reagents at high concentration can drive coupling reactions to completion. Excess reagents and side products can be separated from the growing insoluble peptide simply by filtration and washings and all the synthesis steps can be performed in the same vessel without any transfer of material. In the SPPS approach, the N-protected C-terminal amino acid residue is anchored via its carboxyl group to a hydroxy, chloro, or amino resin to yield respectively an ester or amide linked peptide. After loading the first amino acid, the desired peptide sequence is assembled in a linear fashion from the C-terminus to the N-terminus by repetitive cycles of N<sup>α</sup> deprotection and amino acid coupling reactions. Side chain functional groups of amino acids must be masked with permanent protecting groups (Pn) that are stable in the reaction conditions used during peptide elongation. After coupling, the excess of reactants is removed by filtration and washings. Temporary N-terminal protecting group is removed allowing the addition of the next N-urethane protected amino acid residue by activation of its carboxylic acid. This process of deprotection and coupling is repeated until the desired sequence is obtained. In a final step, the peptide is released from the resin and the side chain protecting groups concomitantly removed.

In SPPS, two strategies of protection are commonly used: the Boc/Bzl and the Fmoc/tBu approaches for protection of terminal amino group/side chain functional group.

Fmoc-based SPPS is now the method of choice for the routine synthesis of peptides.<sup>175</sup>

### 7.2.1 Solid phase synthesis of K<sup>E1</sup> peptide

The sequence of K<sup>E1</sup> was evaluated first with the Peptide Companion software<sup>176</sup>, that allow to predict the ease of synthesizing a certain sequence, and the resulting yield, based on bioinformatics knowledge of the conformational behaviour of sequence motifs, and of their tendency to form aggregates. The level of difficulty for the synthesis of the K<sup>E1</sup> sequence is reported in Figure 7.5:

---

<sup>175</sup> Mizhiriistki M., Shpernat Y., Askinazi B. *Chimica Oggi*, **2003**, 21(6), 22-25.

<sup>176</sup> CSPS Pharmaceuticals, Inc. Copyright© 1994-2004 All. Rights Reserved.

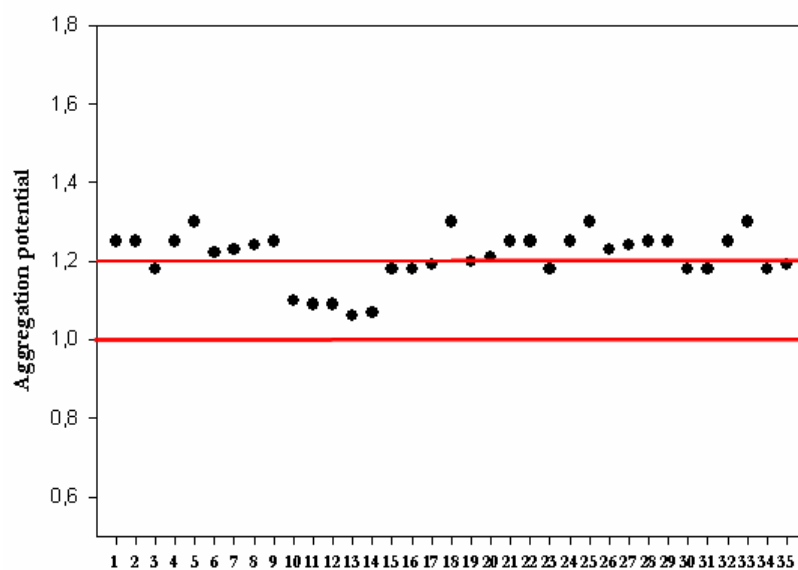


Figure 7.5

A high degree of difficulty is predicted for the constant K regions of the peptide, which is in fact designed in order to form aggregates, and the occurrence of such phenomena, that lowers the yield by masking the reactive terminals of the growing peptide chains, can not be avoided by the presence of side chain protecting groups (for lysine, serine and glutamic acid we have used Boc, TMS and *t*-butyl ester respectively). We have thus considered the use of microwave synthesis to overcome this problem, since microwave greatly improve the yield of solid phase reactions.<sup>174</sup> The synthesis of K<sup>E1</sup> peptide was thus performed with an automated CEM microwave synthesizer on a NovaSyn© TGA resin<sup>177</sup> (Figure 7.6)

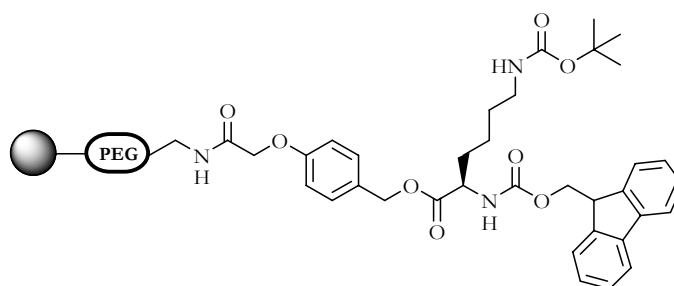


Figure 7.6

This peg-based solid support is based on Tentagel, a composite of polyethylene oxide grafted onto a low cross-linked polystyrene gel-type matrix, which has been amino functionalized and derivatized with the TFA-labile 4-hydroxymethylphenoxyacetic acid linker. The 90  $\mu\text{m}$  beads have a narrow size distribution, excellent pressure stability and swelling properties, and high diffusion rates, making them ideally suited for both batch or continuous flow peptide synthesis. In situ carboxy activation was achieved by PyBop activation in N-methyl pyrrolidone<sup>178</sup>.

<sup>177</sup> Bouajila J, Raffin G, Waton H, Sanglar C. *Polymers and Polymer Composites*, 2003, 11(4), 263-276.

Complete peptide synthesis was obtained in 36 hours with 36 cycle of deprotection and coupling reaction; final treatment with trifluoroacetic acid (TFA), triisopropylsilane (TIPS) and water allow to detach the peptide from the solid support. 40 mg of peptide were by this way obtained in high purity as it can see from its mass spectrum (Figure 7.7):

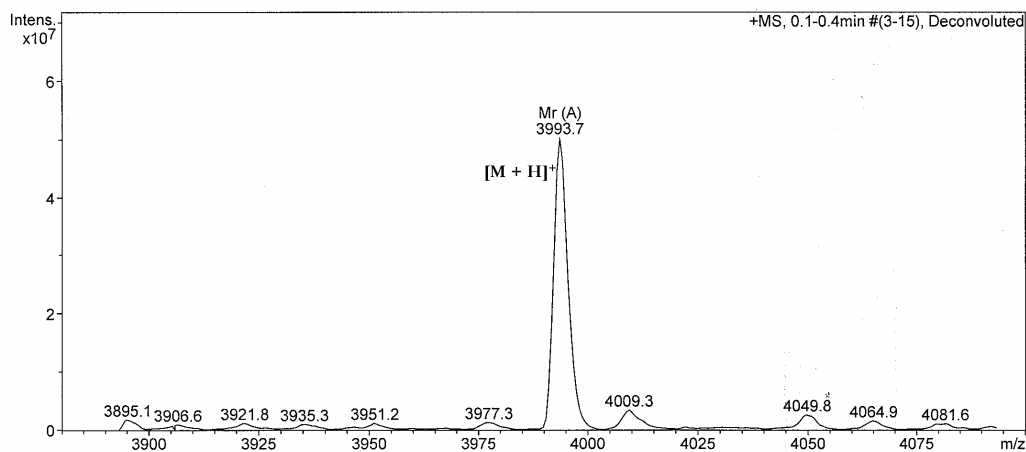


Figure 7.7

### 7.2.2 The solution behaviour of K<sup>E1</sup>.

The CD spectrum of a 50  $\mu$ M solution of K<sup>E1</sup> phosphate buffer is reported in Figure 7.8:

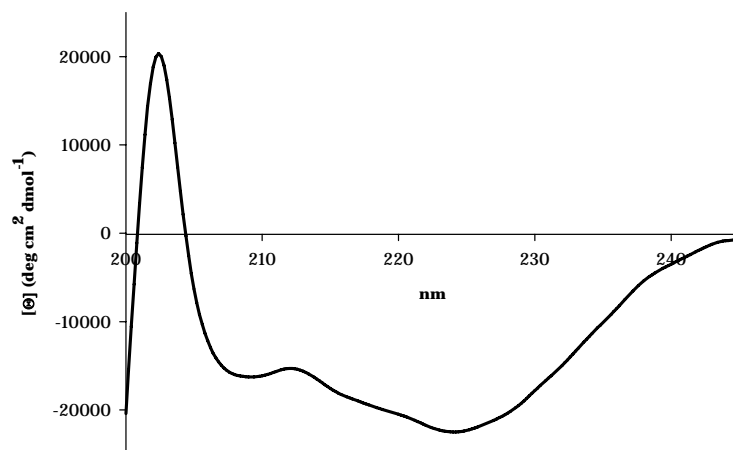


Figure 7.8

The spectrum is consistent with a highly  $\alpha$ -helical structure, with a CD minima at 222 nm and 208 nm and with positive molar ellipticity near 200 nm. This result is very important; the original K

chain, in fact, in the absence of its complementary E chain, shows a typical random-coil spectra and only in the presence of E chain generates an  $\alpha$ -helix structure<sup>179</sup> (Figure 7.9):

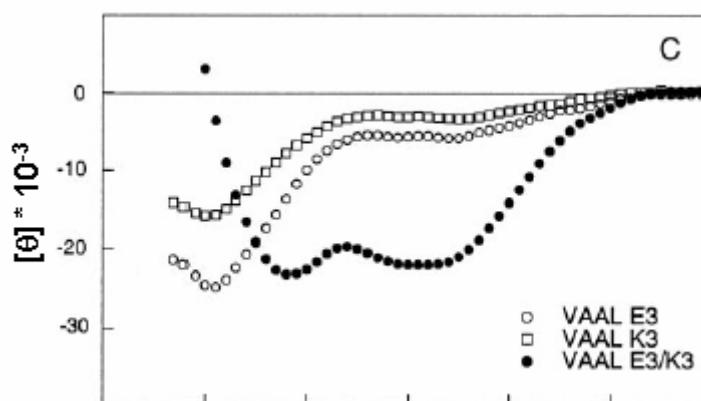


Figure 7.9 from reference 179.

The  $[\theta]_{222}/[\theta]_{208}$  ratio is 1.4 and this suggests a multimeric aggregate in solution as a dimeric, trimeric or higher order coiled coil.

Using equation 5.1 showed in (Chapter 5) we have calculated a 0.6 fraction of  $\alpha$ -helix content for  $K^{E1}$ , corresponding to about 21 amino acids. This is roughly consistent with a helical arrangement conserving a rather high degree of conformational freedom at its terminals and perhaps also at the central heptade region.

The thermal denaturation of  $K^{E1}$  was then studied by measuring the ellipticity value at 222 nm in the 0°-80°C temperature range for a 50  $\mu$ M peptide solution in phosphate buffer (Figure 7.10):

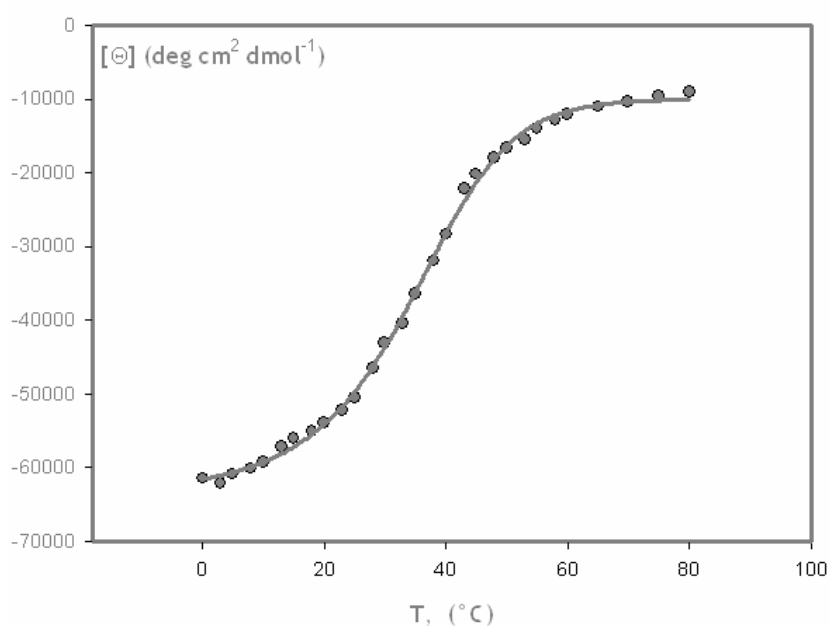


Figure 7.10

<sup>179</sup> Litowski J.R., Hodges R.S. *J. Biol. Chem.*, **2002**, 277(40), 37272-37279.

The peptide exhibited a single sigmoidal thermal unfolding curve, with a melting point at 41°C. The simple sigmoidal behaviour is consistent with a single two-states unfolding process, with the folded structure transforming directly into the final unfolded state without the occurrence of intermediate structures. A dissociation enthalpy of 95.7 kJ/mol is found for this transition by estimating the dissociation constant at any temperature from the  $[\theta]_{222}$  value.<sup>180</sup> This enthalpy magnitude is consistent with that of other trimeric coiled-coil of such length.<sup>181</sup> K<sup>E1</sup> is rather stable upon thermal unfolding and the ellipticity value at 80°C is still corresponding to a 10% helix fraction.

If we compare this behaviour with that of the wild-type K sequence, which is unable of any folding even at room temperature, we are led to look at the key role played by the central heptade sequence, that strongly cooperates in folding up on an aggregate structure, due to its enhanced hydrophobicity.

### 7.2.3 The crystal structure of K<sup>E1</sup>

K<sup>E1</sup> peptide was submitted to crystallization trials in collaboration with the group of prof. Geremia at the Dept. of Chemistry of our University using an automated screening system for crystallization conditions (Tecan robot Technologies); the peptide was finally crystallized at best in 0.1 M HEPES buffer, 1.8-2.5 M sodium sulphate and 1.5-3% w/v PEG400.

X-ray diffraction analysis was carried out using synchrotron radiation light source of ELETTRA Trieste; data set was collected with  $\lambda=1.0$  nm and a resolution of 1.10 Å on a small fragment of a non geminated crystals whose cell parameters are reported in Table 7.2:

<i>KE1 Cell Parameters of non twinned crystal</i>						
System	Trigonal					
Space group	P 3 (P 3 <sup>-</sup> )					
M.V.	22192.62Å <sup>2</sup> - 1.64 - Single molecule					
Solvent	24.36%					
Elementary cell	a (Å)	b (Å)	c (Å)	a(°)	b(°)	g(°)
	22.427	22.427	50.949	90	90	120

Table 7.2

The refined structure allows to individuate a perfect primitive trigonal elementary cell in which K<sup>E1</sup> is structured as an unusual parallel trimeric coiled-coil (Figure 7.11):

<sup>180</sup> Chapeaurouge A., Johansson J., Ferreira S.T. *J. Biol. Chem.*, **2002**, 277(19), 16478-16483.

<sup>181</sup> Dragan A.I., Potekhin S.A., Sivolob A., Lu M., Privalov P. *Biochemistry*, **2004**, 43(47), 14891-14900.

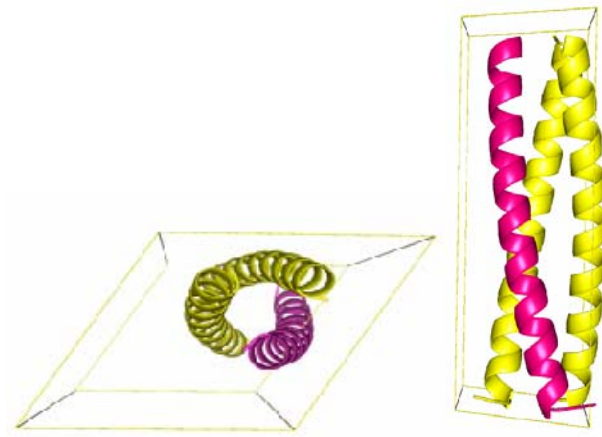


Figure 7.11

The trimer core interface is composed prevalently by hydrophobic residues of valine, leucine and phenylalanine; the external side of the bundle is characterized by polar residue of lysines, glutamic acid or glutamine (Figure 7.12):

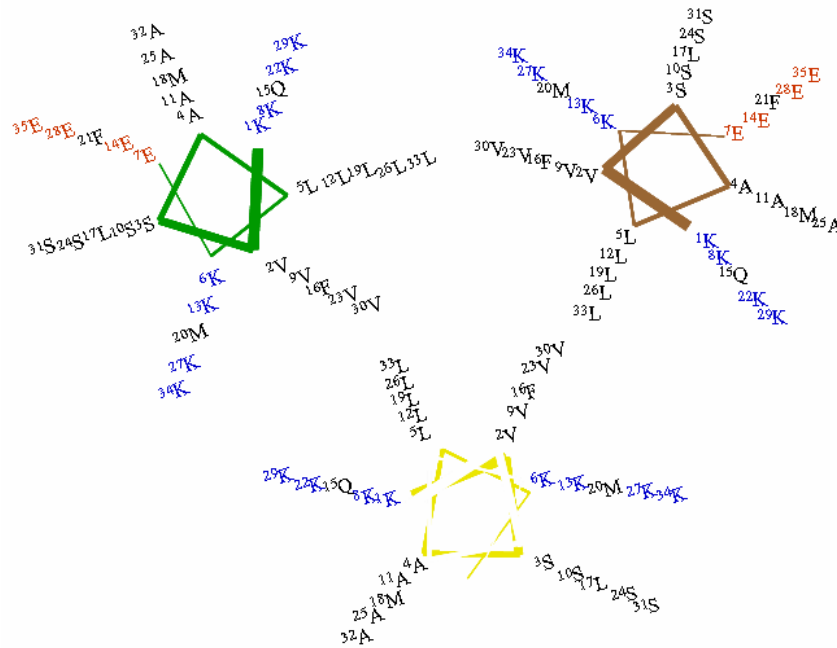


Figure 7.12: the interface of trimer.

A more detailed view of the central heptade shows hydrophobic intrachain contacts between the aromatic side chain of phenylalanine 16 (i) and the hydrocarbon chain of methionine 20 (i+4) and an interchain contact between the same Phe 16 and the hydrocarbon chain of glutamine 15 (i-1), with consequent stabilization of the coiled-coil bundle via *trans* rotamer conformation of phenylalanine (Figure 7.13):

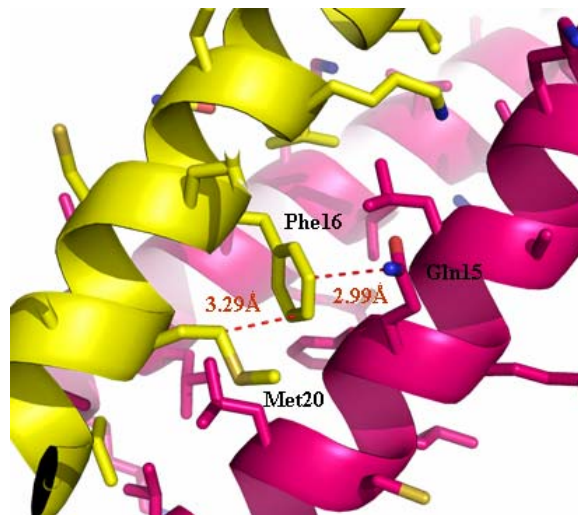


Figure 7.13

An inter cell analysis of trigonal  $K^{E1}$  packing shows a very compact structure. In particular, phenylalanine 21 points out the trimeric bundle and interacts with the corresponding residues of other two bundles of through  $\pi$ -stacking, and a sulphate ion contributes to structure stability through hydrogen bonding and polar interaction with lysine 27 of three different bundles again (Figure 7.14):

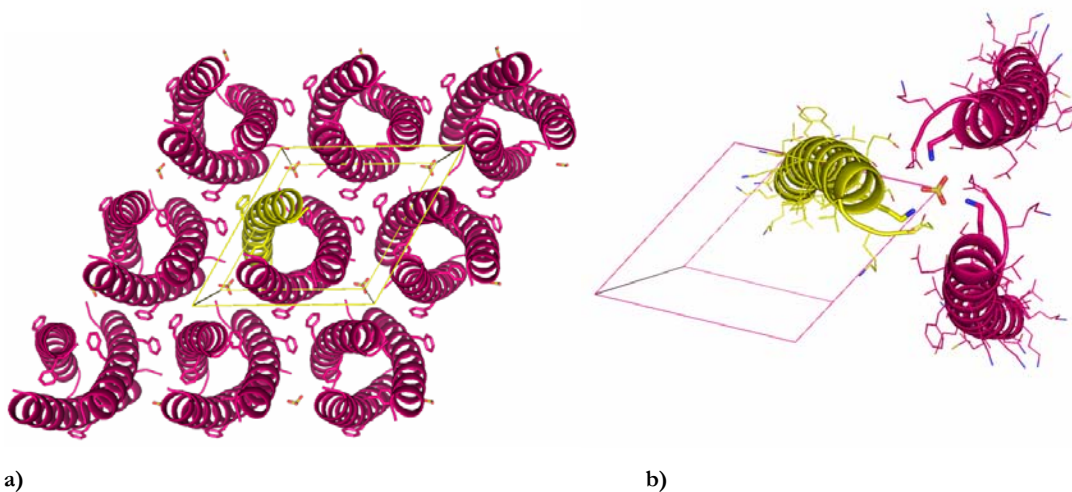


Figure 7.14

As a general comment on this structure, we have to note first that a fully parallel homotrimeric coiled-coil is a very unusual and rare structure, and to our knowledge there are no examples of such an arrangement in the protein databank. In trimeric coiled-coils, one of the chains is usually found in an antiparallel orientation with respect to the other two, in order to reduce the very high

dipole moment given by the coherent orientation of all the peptide carbonyl groups within any  $\alpha$ -helical strand<sup>182</sup>

The strong inter coil interactions between Phe 21 rings and between Lys 27 and sulphate, which can clearly occur only in a full parallel coil, are likely to stabilize this structure in the solid phase. This structure might also be different from that adopted by the peptide in solution and upon binding of xanthenes. It is in fact rather difficult to identify a xanthine binding site in the frame of such a compact coiled-coil which preserves the helical structure also at the level of the central heptade. An attempted crystallization run carried out in the presence of caffeine or theophylline gave no results, and the only structure obtained was that of the trimeric coil. However, the solution aggregation state of  $K^{E1}$  could be very different not only due to the lack of interactions, but also due to the presence of caffeine itself.

### 7.3 The interaction between $K^{E1}$ and theophylline

We have thus started a series of experiments aimed at obtaining a more detailed view of binding of xanthenes by  $K^{E1}$ . As we have seen, binding has been detected on both the phage display and MBP expression version of the peptide. The synthetic peptide, available in high amounts and with its high purity, allows more precise measurements, that were carried out again by CD and by calorimetry.

#### 7.3.1 Binding of theophylline via circular dichroism

We have recently developed a CD methodology that allows to detect binding of small organic molecules to proteins from the appearance of ligand induced bands in the CD spectrum.<sup>183</sup>

We have thus adopted the same approach to study the binding of theophylline and  $K^{E1}$ . The UV spectrum of theophylline in the 250-300 nm range is reported in Figure 4.15b and it shows an absorption band at 272 nm, in a spectral region that is free from peptide absorption. Both the CD spectra of peptide and theophylline alone do not show any CD activity in this region.

A mixture of peptide solution and theophylline, however, shows an induced band at the same wavelength of theophylline absorption in the UV spectrum due to interaction of the non chiral xanthine with the chiral peptide environment, which result in an overall chiral structure (Figure 7.15):

---

<sup>182</sup> Roy L., Core M.A. *Chem. Commun.*, **2009**, 192-194.

<sup>183</sup> a) Berti F., Bincoletto S., Donati I., Fontanive G., Fregonese M., Benedetti F. *Eur. J. Org. Chem.* Submitted. b) Benedetti F., Berti F., Donati I., Fregonese M. *Chem. Comm.*, **2002**, 8, 828-829.

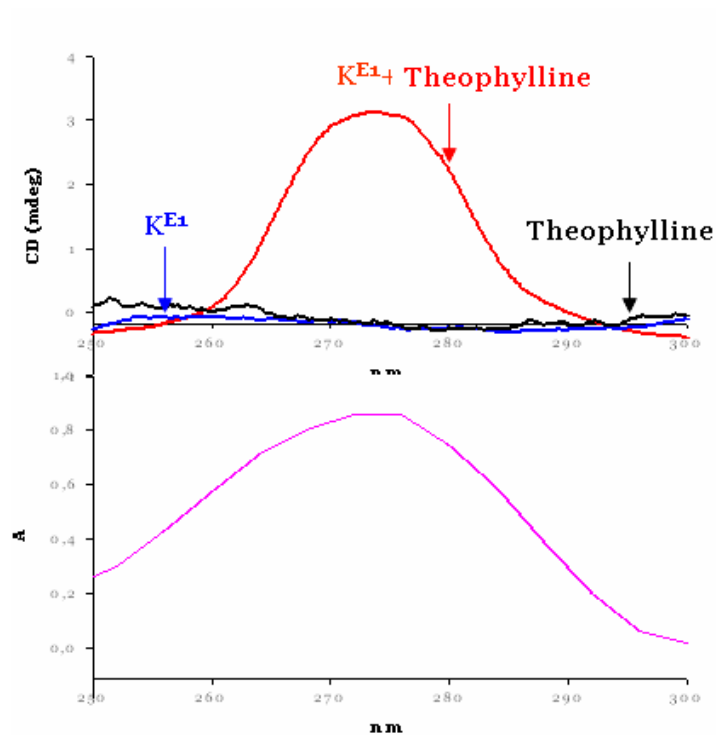


Figure 7.15

The appearance of the CD induced band is again a clean evidence of the recognition of theophylline by  $K^{E1}$  in solution.

The ellipticity of the induced band was found to be proportionally dependent upon the concentration of theophylline in solution and thus it can be related to the amount of bound ligand at any theophylline concentration.

By plotting the measured ellipticity versus the total theophylline concentration, a sigmoidal behaviour is observed (Figure 7.16):

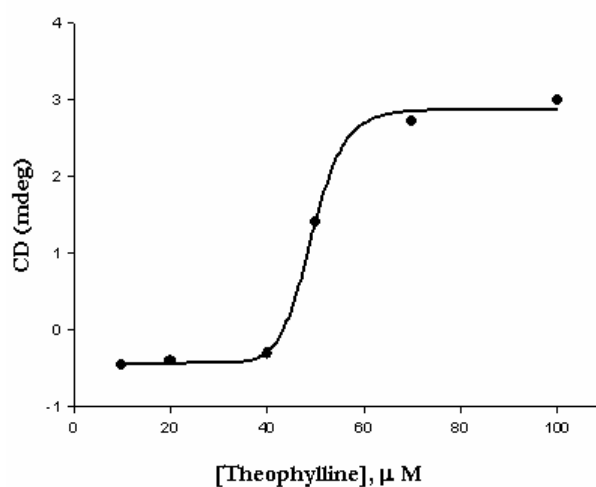


Figure 7.16

For the theophylline- $K^{E1}$  complex, it was possible to estimate a dissociation constant of about 50  $\mu\text{M}$ .

The dissociation constant is thus seen by far less favourable than that obtained via ELISA techniques for the E1 and E8 peptides (pag 77). However, the value is realistic for a small peptides as  $K^{E1}$ , and it can be compared with the few data available in the literature.

In the mid-1990s, Alber reported the design of a coiled coil with an Ala mutation in each helix at the hydrophobic core of the structure.<sup>184</sup> This system was found to bind hydrophobic molecules such as benzene and cyclohexane as a trimer, but in absence of guest, it was present as a mixture of dimers and trimers. More recently, McLendon has simplified the analysis of these coiled-coil receptors through metal coordination, thus eliminating equilibration between dimeric and trimeric complexes.<sup>185</sup> In addition, he demonstrated the use of hexafluorobenzene as a probe to measure the competitive binding of several guest small molecules.  $^{19}\text{F}$  NMR was used to measure binding of benzene to the bundle, giving a  $K_D$  of 110  $\mu\text{M}$ . Using toluene instead of hexafluorobenzene, McLendon found a  $K_D$  of 3  $\mu\text{M}$ ; binding affinities appeared to correlate with ligand size as well as hydrophobicity, and 1,3,5-trimethylbenzene binds more weakly than toluene ( $K_D \approx 50 \mu\text{M}$ ).

The  $K_D$  value found for our equilibrium is comparable with the values found by McLendon for aromatic molecules similar to theophylline.

The relatively low affinity of  $K^{E1}$  for theophylline could also explain the unsuccessful crystallization trial in the presence of xanthenes, since a very delicate balance between inter coil stabilization by sulphate and intra coil stabilization by xanthine is likely to occur during crystal formation.

On the other side, the sigmoidal behaviour reported in Figure 4.16 is suggestive for a cooperative binding process, requiring a critical concentration of theophylline. This could be explained with a shift in the aggregation state similar to that observed by McLendon. Unfortunately, the peptide region of the CD spectrum is entirely masked by very strong theophylline bands, and nothing can be said about changes in peptide conformations upon theophylline binding.

### 7.3.2 Binding of theophylline via calorimetry.

In order to evaluate enthalpic contributions in the  $K^{E1}$  theophylline interaction, calorimetric studies have also been performed, during a visit at the laboratory of prof. Ksenija Kogej, in the dept. of Physical Chemistry of the University of Ljubljana.

---

<sup>184</sup> Gonzales L., Plecs J.J., Alber T. *Nat. Struct. Biol.*, **1996**, *3*, 510-515.

<sup>185</sup> Doer A.J., Case M.A., Pelczar I., McLendon G.L. *J. Am. Chem. Soc.*, **2004**, *126*, 4192-4198.

We have thus carried out an isothermal titration calorimetry (ITC) run. ITC is a methodology to measure the heat change in a reaction at constant temperature.<sup>186</sup> In ITC analysis, the measurement of heat is made at a given temperature, which allows direct determination of binding constant (K), reaction stoichiometry (n), enthalpy change ( $\Delta H$ ), free energy change ( $\Delta G$ ) and entropy change ( $\Delta S$ ). In addition, varying the temperature of the experiment allows the determination of the heat capacity ( $\Delta C_p$ ) for the reaction, thereby providing a complete thermodynamic profile of the molecular interaction.

ITC is mostly used to measure strengths of interaction between proteins and ligands on either small molecules like drugs, or large ones such as nucleic acids or other proteins.<sup>187</sup>

A typical isothermal titration calorimeter is composed of two identical cells (a reference and a sample cell) made of highly efficient thermal conducting material surrounded by an adiabatic jacket. Sensitive thermopile/thermocouple circuits are used to detect temperature differences between the reference cell (filled with buffer or water) and the sample cell containing the macromolecule. Known concentrations of degassed sample are taken in the injector and sample cell. Prior to the injection of the titrant, a constant power (<1 mW) is applied to the reference cell. This signal directs the feedback circuit to activate the heater located on the sample cell and this represents the baseline signal. During the experiment the titrant is added into the sample cell in precisely known aliquots, causing heat to be either taken up or evolved. For an exothermic reaction, the temperature in the sample cell will increase, and the feedback power will be deactivated to maintain equal temperatures between the two cells; for the endothermic reactions the reverse will occur and the feedback circuit will increase power to the sample cell to maintain the temperature (Figure 7.17).

---

<sup>186</sup> Peljhan S., Zagar E., Cerkovnik J., Kogej K. *J. Phys. Chem. B*, **2009**, *113*(8), 2300-2309.

<sup>187</sup> a) Bertini I., Calderone V., Fragai M., Giachetti A. *J. Am. Chem. Soc.*, **2007**, *127*, 2466-2475. b) Haq I. *Arch. Biochem. Biophys.*, **2002**, *403*, 115. c) Chatterjee A., Moulik S.P., Sanyal S.K. *Biophys. Chem.*, **2002**, *98*, 313-327. d) Houtmann J.C., Brown P.H. *Protein Sci.*, **2007**, *16*, 30-42.

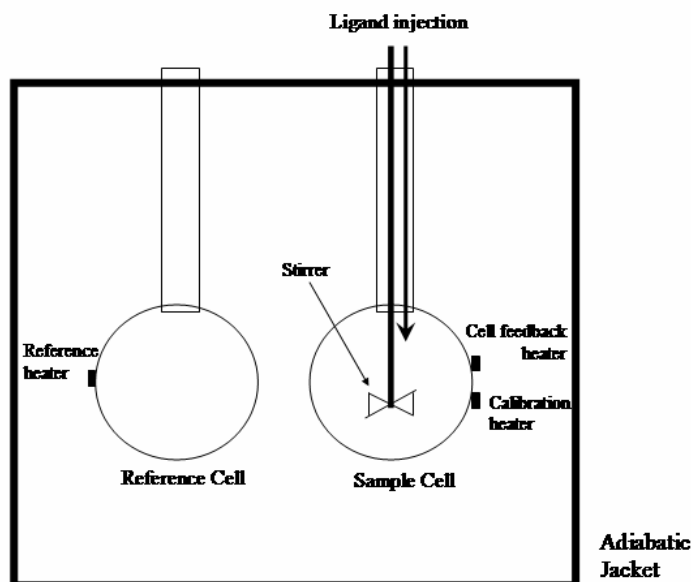


Figure 7.17: schematic representation of ITC calorimeter

The raw data for an experiment consists of a series of heat peaks, with every peak corresponding to a ligand injection. These heat flow peaks are integrated with respect to time, giving the total heat effect per injection. This heat released or absorbed is in direct proportion to the amount of binding that occurs. Additionally, a volume correction is also performed due to dilution in the sample cell due to each injection.

The pattern of these heat effects as a function of the molar ratio [ligand]/[macromolecule] can then be derived to give the thermodynamic parameters. They are derived as follows (eq. 7.1):

$$\Delta G = \Delta H - T\Delta S = \Delta G^0 + RT \ln K_a \quad (\text{eq. 7.1})$$

At the equilibrium  $\Delta G = 0$ , so (eq.7.2):

$$\Delta G^0 = -RT \ln K_a = \Delta H^0 - T\Delta S^0 \quad (\text{eq. 7.2})$$

where  $\Delta G$ ,  $\Delta G^0$ ,  $\Delta H^0$ ,  $\Delta S^0$  and  $K_a$  are free energy change, standard free energy change, standard enthalpy change, standard entropy change and affinity constant, respectively.  $K_a$  is derived from isotherm study. As  $\Delta H$  is dependent on concentration and is not considered as standard state, the  $\Delta H$  value is directly measured by ITC analysis.

Titration of a 53  $\mu\text{M}$   $\text{K}^{\text{E1}}$  solution with increasing amounts of theophylline from a 23 mM mother solution, allowed to obtain important informations (Figure 7.18). The first was that, in agreement with circular dichroism measures, the free molecules  $\text{K}^{\text{E1}}$  and theophylline in solution do interact, with an endothermic process ( $\Delta H = 27.5 \text{ kcal} \cdot \text{mol}^{-1}$ ) and a calorimetric constant for the association process of about  $2 \cdot 10^5 \text{ L} \cdot \text{mol}^{-1}$  that for the inverse process becomes  $K_D = 5 \cdot 10^{-6} \text{ mol} \cdot \text{L}^{-1}$ .

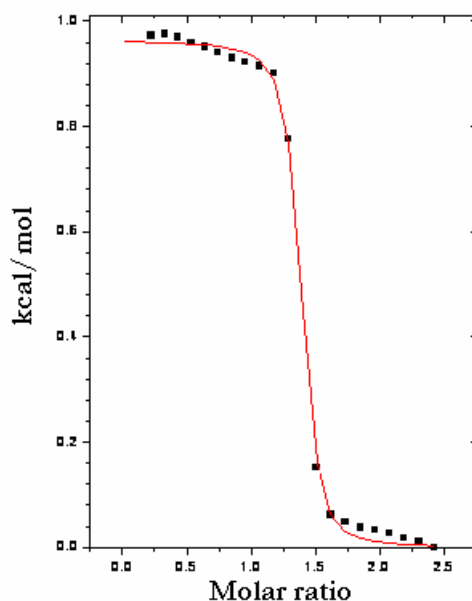


Figure 7.18

#### 7.4 Towards an electrochemical-sensor: F16H-K<sup>E1</sup>

After having characterized the binding of xanthines by K<sup>E1</sup>, we have also designed a further mutant version of the peptide that could be useful for electrochemical sensing of xanthines. The X-ray structure of the peptide lead us to the idea of creating a metal-binding site in the middle of the bundle by mutating phenylalanine 16 to histidine, (Figure 7.19):

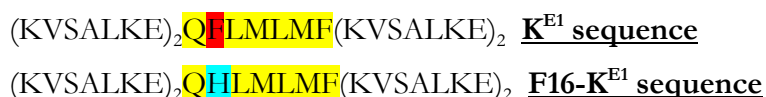
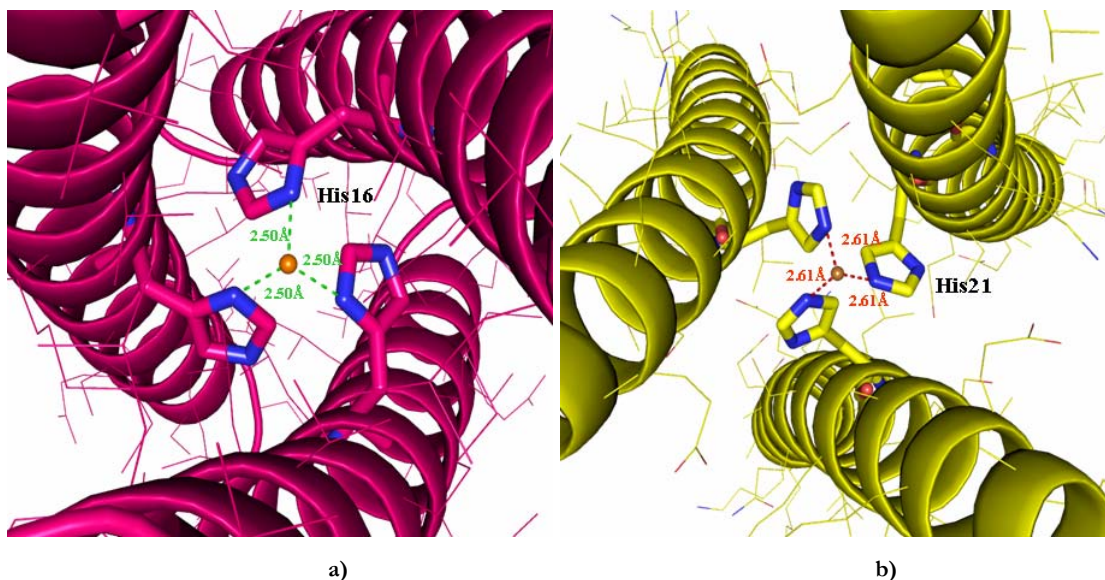


Figure 7.19

The histidine side chain with its imidazole moiety is a key binding site for a number of metals ions including Cu<sup>2+</sup>; Zn<sup>2+</sup> and Ni<sup>2+</sup>. The metal ion coordination generally takes place via N(1) or N(3) atoms of imidazole as is found in a great variety of nickel, zinc or copper proteins including carbonic anhydrase, carboxypeptidase and blue copper proteins.<sup>188</sup> The metal complexes of peptides containing two or more imidazole residues are also frequently used to mimic the structural and catalytic features of the active sites of metalloproteins. Our idea was thus to mutate phenylalanine 16 with histidine in order to generate a binding site for a metal and to use this structure as an immobilized receptor on glassy carbon electrodes for a biosensor system for theophylline (Figure 7.20a); such a system could in fact provide a metal electrochemical signal that could be modulated by binding of xanthines occurring in close proximity.

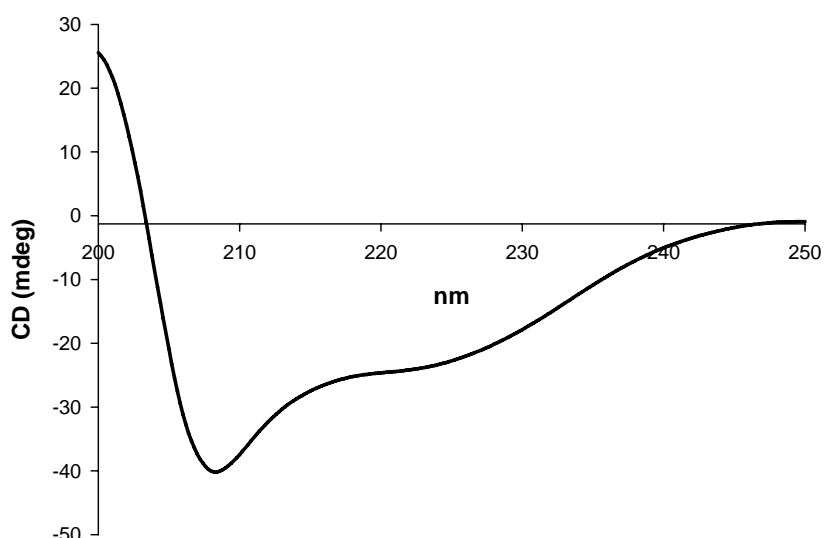
<sup>188</sup> Bertini I, Cavallaro G., Mc Greevy K.S. *Coord. Chem. Rev.*, **2010**, 254(5-6), 506-524.



**Figure 7.20:** a) model structure of mutated K<sup>E1</sup>peptide with metal ion coordination and b) the metal binding site of the catalytic site of carbonic anhydrase.<sup>189</sup>

The synthesis of F16-K<sup>E1</sup> peptide was performed of K<sup>E1</sup>; the complete synthesis was obtained in 40 hours with 40 cycle of deprotection and coupling reaction and by this way, 30 mg of peptide were obtained.

Circular dichroism analysis of the mutant peptide shows that (Figure 7.21). The peptide adopts an  $\alpha$ -helix conformation, but the ratio  $[\theta]_{222}/[\theta]_{208}$  is less than 1 and this is consistent with the presence of a single  $\alpha$ -helix rather than with a coiled-coil.



**Figure 7.21**

<sup>189</sup> Alterio V., Hilvo M., Di Fiore A., Supuran C.T., Pan P., Parkkila S., Scaloni A., De Simone G. *PNAS*, **2009**, *106*(38), 16233-16238.

F16H-K<sup>E1</sup> was studied also as a theophylline receptor in solution, using again circular dichroism. The experimental conditions were the same used as described for K<sup>E1</sup>-theophylline and also the mutant peptide results able of interaction with the xanthine. The concentration-dependent theophylline CD induced band is reported in Figure 7.22.

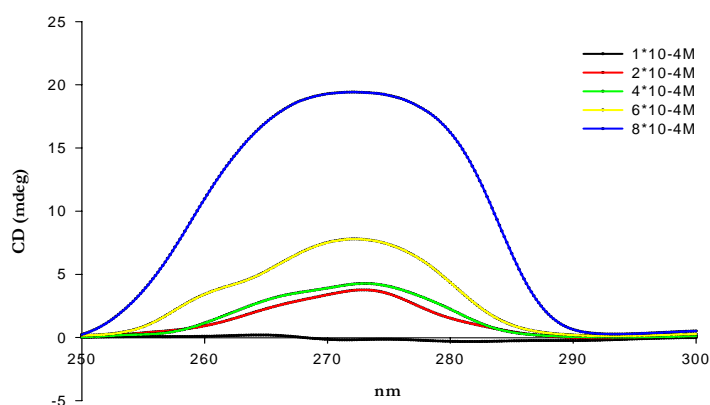


Figure 7.22: titration of peptide K<sup>E1</sup>-His with increasing amount of theophylline.

We are currently carried out a series of preliminary studies on the metal coordinating ability of F16H-K<sup>E1</sup>, and once optimal complexation conditions will be found, we are planning to develop a F16H-K<sup>E1</sup> coated glassy carbon electrode.

## **Chapter 8: Antibodies library**

In this last chapter we report briefly a series of preliminary results obtained using the palytoxin conjugates that have been prepared as described in Chapter 3.

As we have outlined, the palytoxin project is very ambitious, including the development of bioreceptors towards palytoxin by many different approaches, ranging from conventional to highly innovative methodologies, on one side, and the preparation of several sensing devices ranging from standard ELISAs, to stick “on the field” rapid analysis systems, to electrochemical sensor based on carbon nanotubes.

Among the primary transducers, were planned:

- conventional polyclonal antisera and monoclonal antibodies raised against palytoxin bioconjugates with immunogenic carrier proteins,
- phage display single-chain Fv antibodies from both immune and nonimmune populations of T-lymphocytes selected towards the whole palytoxin molecule and its fragment,
- phage display libraries of peptide receptors.

Up to date, polyclonal and a monoclonal antibody have been obtained and used to set up a reference sandwich ELISA.

### **8.1 Immunization strategy**

Monoclonal antibodies (mAb) are monospecific antibodies pool that are identical because as they are the progenies of a unique parent lymph cell.<sup>190</sup>

Monoclonal antibodies are typically made by fusing myeloma cells with the spleen cells from a mouse that has been immunized with the desired antigen. Polyethylene glycol is used to fuse adjacent plasma membranes, but the success rate is low and a selective medium is used in which only fused cell can grow. In the myeloma cell line the hypoxanthine-guanine-phosphoribosyl transferase (HGPRT) activity has been depleted.

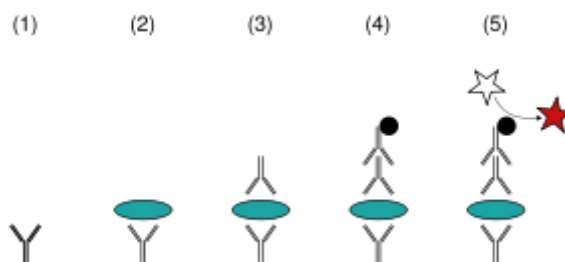
Unfused myeloma cells cannot survive as they lack HGPRT while unfused spleen cells are also unable to survive indefinitely because of their limited life span. Only fused hybrid cells, referred to as hybridomas, are able to grow indefinitely in the media because the spleen cell partner supplies HGRPT. The mixture of cells is then diluted and clones are grown from single parent cells on microtitre wells. The antibody secreted by the different clones are then assayed for their ability to bind the antigen by a test such ELISA or immuno-Dot Blot. The most productive and stable clones is then selected are future use.

---

<sup>190</sup> Tonegawa S., Hozumi N., Matthyssens G., Schuller R. *Cold Spring Harbor Symp. Quant. Biol.* **41**, 877-889.

Polyclonal antibodies (or antisera) are antibodies mixture that are obtained from different B-cells. They are a combination of immunoglobulin molecules raised against a specific antigen, each identifying a different epitope. These antibodies are typically produced by immunization of a suitable mammal, such as mouse, rabbit or goat.

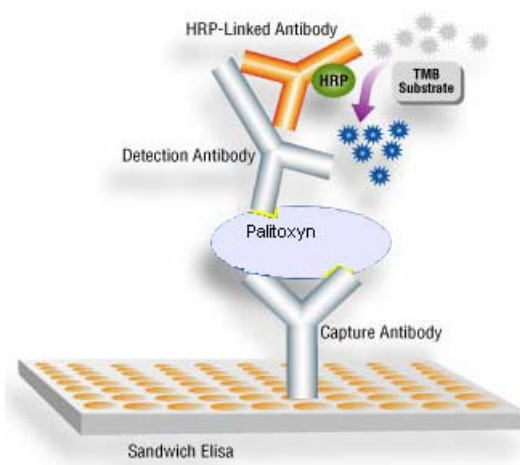
Using mono- and polyclonal antibodies in analytical systems represents one of the most established methodologies of biotechnologies; typically antibodies are used in ELISA test and widely used in the “sandwich” ELISA. (Figure 8.1):<sup>191</sup>



**Figure 8.1:** a general scheme of sandwich ELISA.

In this type of ELISA test, a solid phase is coated with capture antibody (1), the sample is added and the antigen binds to the capture antibody (2). An enzyme linked capture antibody, used as detecting antibody is added, and binds again to the antigen via another epitope (3); substrate is finally added and is converted by the enzyme to a detectable product (4).

The ELISA test we are focusing on, along to this class of immunoenzymatic systems (Figure 8.2). In our case, a mouse monoclonal antibody was used as capture antibody and a rabbit polyclonal antibody is the second detecting antibody. Binding between palytoxin and the detecting antibody was determined by a secondary goat antirabbit antibody-HRP conjugate.



**Figure 8.2**

<sup>191</sup> Velusamy V., Arshak K., Korostynska O., Oliwa K., Adley C. *Biotech. Adv.*, **2010**, 28(2), 232-254.

### 8.1.1 Production of polyclonal antibodies from rabbits.

The monoclonal antibody required as capture antibody was provided by Hawaii Biotechnology Group, Inc.; the polyclonal antibody was produced by the group prof. Marzari at the Dept. of Life Science by immunization of two New Zealand rabbits using the PTX-BSA conjugate as immunogenic antibodies production.

Generally the first immunization causes a weak response in the immune system of the animal (primary response); subsequent expositions to the same antigen, however, induce a more intense response and consequently a more intense production of antibodies in the serum.

For these reasons, the two rabbits were treated respectively with ten and seven subcutaneous inocules of 200 µg of PTX-BSA.

A pool of high reactivity bleedings of the two rabbits was recovered and tested in ELISA against a coating of PTX-BSA conjugate, BSA, PTX-OVA and OVA, in dilution of pool bleedings (Figure 8.3):

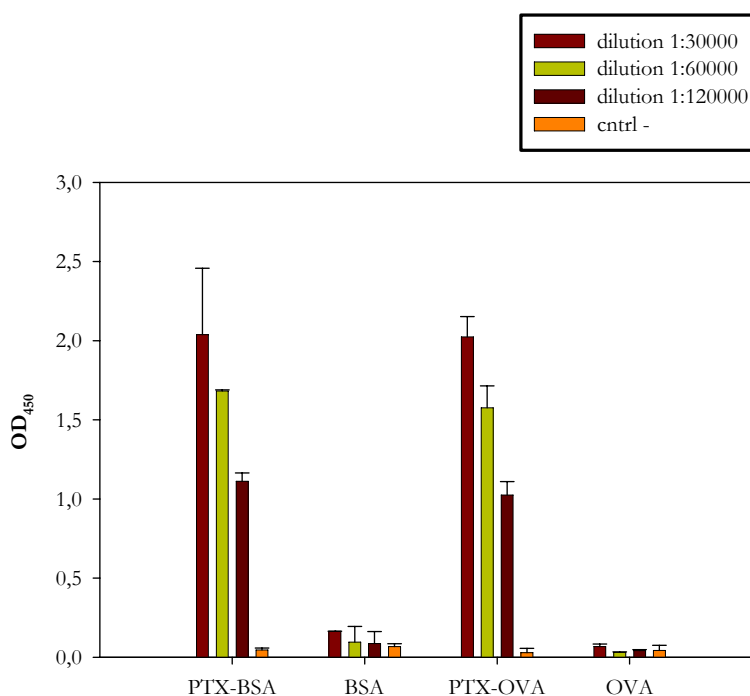


Figure 8.3

The anti serum shows a good reactivity against the conjugate molecules and a low reactivity against the proteins. For this reason it can be used as detection antibody in the sandwich ELISA.

### 8.1.2 The “sandwich” ELISA.

The development of any ELISA analytical system requires a careful and time-consuming optimization of the most important parameters that can influence sensitivity, accuracy and precision of the methodology, with the aim of obtaining the maximum sensitivity, intra assay precision in the order of 5%, inter assay precision around 10%, accuracy within ± 10% in the

whole dynamic range, that must be as wider as possible, and also with the aim of reducing the time required for the whole procedure at best. Among the most important variables, are often found the temperature and time required for the immobilization of the capture antibody on the solid surface, the conditions for the subsequent saturation step, the time, pH, ionic strength, temperature for the binding of the analyte to the capture antibody, the binding of the second and third antibody, the number of washing cycle.

In our case, the ELISA sandwich for free palytoxin was optimized under the following conditions:

- 1- coating of monoclonal antibody anti-PTX (mAb-antiPTX) at 4°C, overnight in phosphate buffer,
- 2- saturation with defatted milk at room temperature for one hour,
- 3- incubation with free palytoxin in phosphate buffer at different dilution;
- 4- washing for three times with phosphate buffer,
- 5- incubation with 1:4000 dilution of the polyclonal antiserum in defatted milk for 1h at room temperature;
- 6- washing for three times with phosphate buffer and three time with phosphate buffer and tween,
- 7- incubation with antirabbit-HRP conjugate for 45 minutes at room temperature in defatted milk,

Under these conditions it was possible to generate the calibration curve for standard palytoxin samples reported in Figure 8.4:

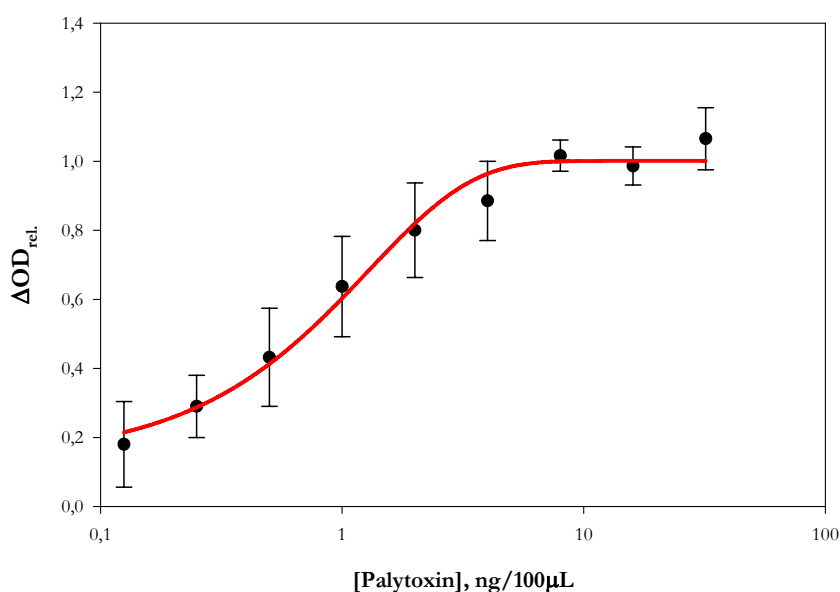


Figure 8.4

As expected, the calibration curve obeys to a sigmoidal behaviour. The interassay precision is 10.4%: this satisfactory level of variability allows to identify the dynamic range of the test between 0.25 (LOD) and 2.5 ng/100 $\mu$ L. Although this range is rather narrow, it is none the less sufficient to begin the further steps in the development of the ELISA system, which will include the determination of the accuracy from the recovery of spiked sample at LOD, medium and high concentrations of palytoxin, the comparison with a reference LC-MS analytical method.

# EXPERIMENTAL SECTION

Moisture sensitive reactions were carried out in oven-dried vessels under a positive argon pressure. THF was pre-dried over KOH, fractioned and redistilled from sodium benzophenone before use. Dichloromethane was dried over calcium chloride and fractioned .

Flash column chromatography was performed on silica gel 60 (230-400 mesh); silica gel 60<sub>F254</sub> coated plastics sheets were used for TLC and developed with I<sub>2</sub> or with an aqueous solution of KMO<sub>4</sub> and H<sub>2</sub>SO<sub>4</sub>.

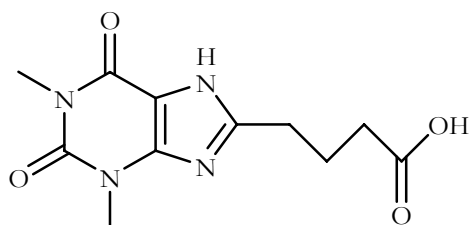
Melting point were determined in an open capillary apparatus and are uncorrected.

IR spectra was obtained with ATAVAR 320 FT/IR spectrophotometer dispersing the sample in Nujol.

CD spectra was obtained with JASCO J-710 spectrometer under nitrogen flux.

Mass spectra ESI was obtained with a API 1 Perkin-Elmer SICIEX spectrometer.

<sup>1</sup>H-NMR spectra and <sup>13</sup>C NMR spectra was recorded with Jeol EX400 (400 MHz) and VARIANT 500 (500 MHz) spectrometer. CDCl<sub>3</sub> solvent solution contain Me<sub>4</sub>Si as an internal standard.



**4-(2,3,6,7-tetrahydro-1,3-dimethyl-2,6-dioxo-1H-purin-8-yl)butanoic acid (1)**

5 g of 5,6-diamino-1,3-dimethyl uracil (29.38 mmol) were dissolved under anhydrous atmosphere in dry dimethylformamide (50 mL). 5.02 g of glutaric anhydride were added (44.07 mmol). The mixture was stirred at room temperature for 17 hours. After this time the solvent was eliminated by distillation at reduced pressure and the solid residue was dissolved into 80 mL of 2N NaOH; the mixture was heated at 75°C for 2 hours.

HCl 6N was added until pH=5, the volume was reduced by evaporation and the red precipitate was filtered, washed with water (3 times) and dried under vacuum for 12 hours to obtain 5.61 g (71%) of compound **1**.

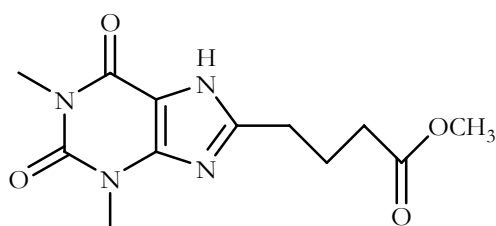
**<sup>1</sup>H NMR** ( $\delta$ -DMSO- $d_6$ ): 1.89 (quint., 2H, CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>COOH, J=7.4 Hz), 2.25 (t, 2H, CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>COOH, J=7.2 Hz), 2.69 (t, 2H, CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>COOH, J=7.2 Hz), 3.19 (s, 3H, CON(CH<sub>3</sub>)CO), 3.40 (s, 3H, N(CH<sub>3</sub>)CO), 8.40 (s, 1H, COOH).

**<sup>13</sup>C NMR** ( $\delta$ -DMSO- $d_6$ ): 23.40 (CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>COOH), 28.01 (CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>COOH), 28.28 (N(CH<sub>3</sub>)CO), 31.56 (CON(CH<sub>3</sub>)CO), 33.87 (CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>COOH), 106.64 (NHC=CN), 148.53 (NHC=CN), 151.71 (N(CH<sub>3</sub>)CON(CH<sub>3</sub>)), 154.12 (C=N), 154.53 (N(CH<sub>3</sub>)CO), 174.55 (COOH).

**MS** m/z: 267.1 [M+H]<sup>+</sup>; 289.1 [M+Na]<sup>+</sup>; 305.2 [M+K]<sup>+</sup>.

**IR** (cm<sup>-1</sup>): 3025 (NH); 1715 (C=O).

**m.p.:** 78-80 °C.



**Methyl 4-(2,3,6,7-tetrahydro-1,3-dimethyl-2,6-dioxo-1H-purin-8-yl)butanoate (2)**

3 g of compound **1** (11.42 mmol) were dissolved in 100 mL of methanol and cooled in a ice bath. 1.66 mL of thionyl chloride (2 g, d=1.638, 22.84 mmol) were added dropwise; the solution was stirred for 10 minutes and then heated under reflux for 24 hours.

The solvent was then eliminated by distillation at reduced pressure to give 3.2 g (99%) of ester **2**.

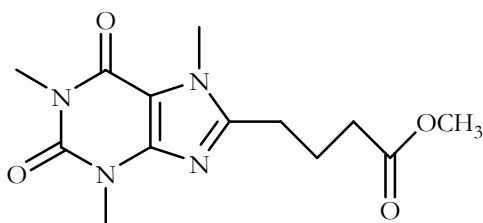
**<sup>1</sup>H NMR** ( $\delta$ -DMSO- $d_6$ ): 1.93 (quint., 2H, CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>COOH, J=7.5 Hz), 2.36 (t, 2H, CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>COOH, J=7.4 Hz), 2.72 (t, 2H, CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>COOH, J=7.4 Hz), 3.20 (s, 3H, CON(CH<sub>3</sub>)CO), 3.40 (s, 3H, N(CH<sub>3</sub>)CO), 3.56 (s, 3H, COOCH<sub>3</sub>).

**<sup>13</sup>C NMR** ( $\delta$ -DMSO- $d_6$ ): 25.20 (CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>COOH), 30.30 (CONH(CH<sub>3</sub>)CO), 31.40 (CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>COOH), 32.30 (CON(CH<sub>3</sub>)C=C), 33.20 (CH<sub>2</sub>COOCH<sub>3</sub>), 51.91 (COOCH<sub>3</sub>), 108.52 (C=C=CN), 150.06 (CH<sub>3</sub>NC=CN), 151.38 (CH<sub>3</sub>NCONCH<sub>3</sub>), 153.90 (C=N), 158.99 (CH<sub>3</sub>NCOC=C), 174.15 (COOCH<sub>3</sub>).

**MS** m/z: 281.1 [M+H]<sup>+</sup>, 303.3 [M+Na]<sup>+</sup>.

**IR** (cm<sup>-1</sup>): 3021 (NH); 1732 (C=O).

**m.p.:** 92-95 °C.



### Methyl 4-(2,3,6,7-tetrahydro-1,3,7-trimethyl-2,6-dioxo-1 *H* purin-8-yl)butanoate (3)

2.0 g of compound **2** (3.04 mmol) were dissolved under anhydrous atmosphere in 60 mL of anhydrous tetrahydrofuran. 1.1 mL of methyl iodide (2.5 g, d=2.28, 17.7 mmol) were added and the reaction was stirred for 16 hours.

After this time 0.57 g of sodium hydride (14.16 mmol) were added and a precipitate appeared white. After 45 minutes the precipitate redissolved; 10% citric acid was then added dropwise until pH=5.

The organic phase was collected, the aqueous phase was extracted with ethyl acetate (3 \* 40 mL), washed with saturated solution of NaHCO<sub>3</sub> and with brine. The combined extracts were dried over anhydrous sodium sulphate, the solvent was evaporated and the crude product was purified by flash chromatography on silica gel using dichloromethane/methanol=6/4 as eluent mixture.

Final purification give 1.46 g of pure yellow solid product (73%).

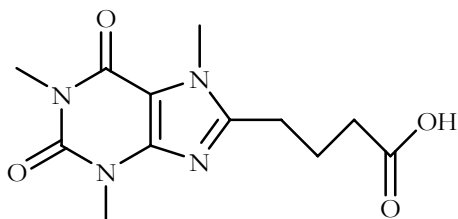
**<sup>1</sup>H NMR** ( $\delta$ -CDCl<sub>3</sub>): 1.91 (quint., 2H, CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>COOCH<sub>3</sub>, J=7.1 Hz), 2.26 (t, 3H, CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>COOCH<sub>3</sub>, J=7.1 Hz), 2.59 (t, 3H, CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>COOCH<sub>3</sub>, J=7.2 Hz), 3.19 (s, 3H, CON(CH<sub>3</sub>)CO), 3.35 (s, 3H, N(CH<sub>3</sub>)CO), 3.48 (s, 3H, COOCH<sub>3</sub>), 3.72 (s, 3H CN(CH<sub>3</sub>)C).

**<sup>13</sup>C NMR** ( $\delta$ -CDCl<sub>3</sub>): 22.49 (CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>COOCH<sub>3</sub>), 25.87 (CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>COOCH<sub>3</sub>), 27.93 (N(CH<sub>3</sub>)CO), 29.77 (CON(CH<sub>3</sub>)CO), 31.77 (CN(CH<sub>3</sub>)C), 32.86 (CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>COOCH<sub>3</sub>), 51.78 (COOCH<sub>3</sub>), 107.48 (NHC=CN), 147.95 (NHC=CN), 151.73 (N(CH<sub>3</sub>)CON(CH<sub>3</sub>)), 153.32 (C=N), 155.34 (N(CH<sub>3</sub>)CO), 173.36 (COOCH<sub>3</sub>).

**MS** m/z: 294.1 [M+H]<sup>+</sup>, 333.1 [M+K]<sup>+</sup>.

**IR** (cm<sup>-1</sup>): 1745 (C=O); 3010 (C-H).

**m.p.:** 88-90 °C.



**4-(2,3,6,7-tetrahydro-1,3,7-trimethyl-2,6-dioxo-1H-purin-8-yl)butanoic acid (4)**

0.5 g of compound **3** were mixed in 20 mL of a 3N LiOH checking the pH value (pH=12).

The solution was stirred for 12 hours until complete dissolution was achieved.

The solution was then acidified with 3N hydrochloric acid to pH=2, and extracted with ethyl acetate (5 \* 20 mL). The combined organic layers were washed with brine and dried over anhydrous sodium sulphate. The solvent was evaporated to give 0.4 g (87%) of pure product that can be used directly for next step.

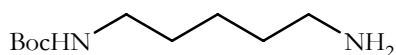
**<sup>1</sup>H NMR** (δ-CDCl<sub>3</sub>): 1.89 (quint., 2H, CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>COOH, J=7.2 Hz), 2.35 (t, 2H, CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>COOH, J=7.1 Hz), 2.75 (t, 2H, CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>COOH, J=7.2 Hz), 3.18 (s, 3H, CON(CH<sub>3</sub>)CO), 3.37 (s, 3H, N(CH<sub>3</sub>)CO), 3.78 (s, 3H, CN(CH<sub>3</sub>)C), 8.83 (s, 1H, COOH).

**<sup>13</sup>C NMR** (δ-CDCl<sub>3</sub>): 23.04 (CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>COOH), 26.07 (CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>COOH), 28.10 (N(CH<sub>3</sub>)CO), 29.80 (CON(CH<sub>3</sub>)CO), 31.77 (CN(CH<sub>3</sub>)C), 33.10 (CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>COOH), 108.44 (NHC=CN), 147.86 (NHC=CN), 151.82 (N(CH<sub>3</sub>)CON(CH<sub>3</sub>)), 154.11 (C=N), 155.42 (N(CH<sub>3</sub>)CO), 174.60 (COOH).

**MS** m/z: 281.1 [M+H]<sup>+</sup>, 303.2 [M+Na]<sup>+</sup>, 319.1 [M+K]<sup>+</sup>.

**IR** (cm<sup>-1</sup>): 1783 (C=O); 3017 (C-H).

**m.p.:** 121-122 °C.



**(5).**

1.2 g of di-*tert*-butyl dicarbonate (5.87 mmol) dissolved in 20 mL of dichloromethane were added dropwise to a solution of 3.45 mL of 1,5-diamino propane (3g, d=0.869, 29.36 mmol) in 40 mL of dichloromethane at 0°C. The solution were stirred for 1 hour at 0°C and then for 1 hour again at room temperature. After this time the resulting suspension was filtered under vacuum in order to remove the white precipitate; the solution was concentrated with distillation at reduce pressure to give a yellow oil. The residue was redissolved in ethyl acetate (20 mL) and the organic solution was washed with a saturated solution of NaHCO<sub>3</sub> and brine.

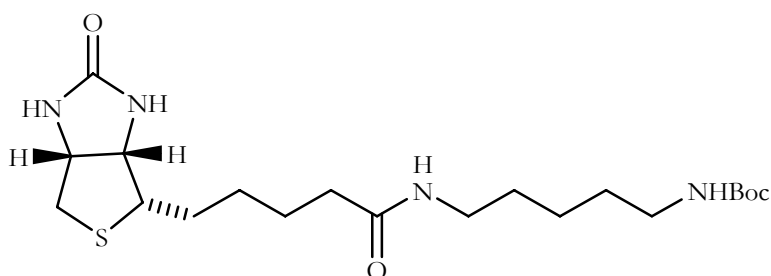
The solution was dried on anhydrous sodium sulfate, the solvent was eliminated to give an oil product (5.28 g, 89%).

**<sup>1</sup>H NMR** ( $\delta$ -CDCl<sub>3</sub>): 1.22-1.40 (m, 6H, CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>), 1.44 (s, 9H, CH<sub>3</sub> (Boc)), 2.65-2.68 (m, 2H, CH<sub>2</sub>NH<sub>2</sub>, J<sub>3</sub>=7.1 Hz, J<sub>5</sub>=3.1 Hz), 3.10 (m, 2H, CH<sub>2</sub>NHBoc, J<sub>3</sub>=6.1 Hz), 4.57 (s, 1H, NH).

**<sup>13</sup>C NMR** ( $\delta$ -CDCl<sub>3</sub>): 23.83 (CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>), 28.43 (C(CH<sub>3</sub>)<sub>3</sub> Boc), 29.79 (CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>), 30.29 (CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>), 40.26 (BocNHCH<sub>2</sub>), 41.56 (CH<sub>2</sub>NH<sub>2</sub>), 77.54 (C(CH<sub>3</sub>)<sub>3</sub> Boc), 156.17 (CO Boc).

**MS** m/z: 203.1 [M+H]<sup>+</sup>; 241.2 [M+K]<sup>+</sup>.

**IR** (cm<sup>-1</sup>): 1771 (C=O (Boc)); 3088 (N-H).



**tert-butyl 5-(5-((3aR,6S,6aS)-hexahydro-2-oxo-1H-thieno[3,4-d]imidazol-6-yl)pentanamido)pentylcarbamate (6).**

0.3 g of biotin (1.24 mmol) were dissolved under anhydrous atmosphere in 10 mL of dry acetonitrile. The solution was cooled to 0°C and then 0.17 g of HOBT (1.24 mmol) were added; the pH of solution was adjusted to 8.5 of N-methylmorpholine. 0.25 g of compound **5** (1.24 mmol) were added, 0.29 g of EDC-Cl (1.49 mmol) were finally added and the mixture was stirred for 1 hour at 0°C and then for 16 hours at room temperature.

After this time the solvent was removed, the solid residue was divided between ethyl acetate and water and the organic layer was washed with 10% citric acid, saturated NaHCO<sub>3</sub> and brine.

The organic phase was dried with anhydrous sodium sulfate, the solvent was removed and the crude product was recrystallized from 2-propanol/*di-iso*-propyl ether to give 0.092 gr (60%) of pure brown solid product.

**<sup>1</sup>H NMR** ( $\delta$ -DMSO-d<sub>6</sub>): 1.28-1.65 (m, 21H, CH<sub>3</sub> (Boc)+CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>+CH<sub>2</sub>CH<sub>2</sub>(CH<sub>2</sub>CHS), 2.18 (t, 2H, NHCOCH<sub>2</sub>, J=7.2 Hz), 2.69 (m, 1H, CHS), 2.91 (m, 1H, CH<sub>2</sub>S), 3.02 (t, 2H, CH<sub>2</sub>NHBoc, J<sub>3</sub>=6.4 Hz), 3.17 (m, 3H, CH<sub>2</sub>S+CH<sub>2</sub>CH<sub>2</sub>NHCO), 4.30 (m, 1H, CHCHS), 4.48 (m, 1H, CHCH<sub>2</sub>S), 6.36 (s, 1H, NHCONH), 6.42 (s, 1H, NHCONH), 7.73 (s, 1H, NHBoc), 7.78 (NHCO).

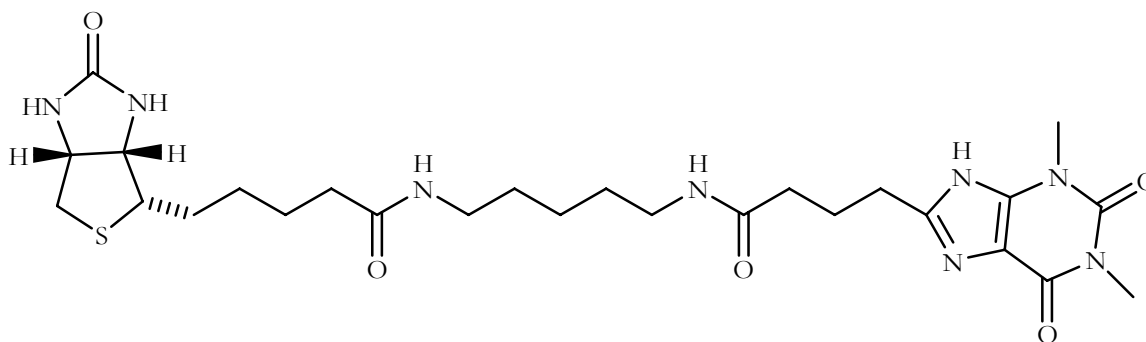
**<sup>13</sup>C NMR** ( $\delta$ -DMSO-d<sub>6</sub>): 23.45 (CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>), 24.40 (NHCOCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CHS), 24.47 (NHCOCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CHS), 25.90 (NHCOCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CHS), 28.60 (CH<sub>3</sub> Boc),

29.41-29.44 (NH<sub>2</sub>BocCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NHCO), 31.91 (NHCOCH<sub>2</sub>), 40.4 (CH<sub>2</sub>NHCO), 42.9 (CH<sub>2</sub>NHBoc), 55.99 (CHS), 59.75 (NHCHCH<sub>2</sub>S), 60.51 (NHCHCHS), 61.60 (CH<sub>2</sub>S), 79.5 (C(CH<sub>3</sub>)<sub>3</sub> Boc), 156.0 (CO Boc), 164.7 (NHCONH), 172.7 (NHCOCH<sub>2</sub>).

**MS** m/z: 429.45 [M+H]<sup>+</sup>.

**IR** (cm<sup>-1</sup>): 3320 (N-H); 1790 (C=O).

**m.p.:** 138-140 °C.



**N-(5-(4-(2,3,6,9-tetrahydro-1,3-dimethyl-2,6-dioxo-1H-purin-8-yl)butanamido)pentyl)-5-((3aR,6S,6aS)-hexahydro-2-oxo-1H-thieno[3,4-d]imidazol-6-yl)pentanamide (8).**

0.31 g of theophylline-8-butanamide **1** (1.15 mmol) were dissolved in anhydrous atmosphere in 10 mL of N,N-dimethylformamide. The solution was cooled at 0°C and then 0.15 g of HOBT (1.15 mmol) were added; the pH value of solution was adjusted with N-methylmorpholine at the final of 8.5.

After this, 0.41 g of compound **7** were added (1.15 mmol) and then 0.26 g (1.38 mmol) of EDC-Cl. The mixture was stirred at room temperature for 18 hours.

After this time, the solvent was evaporate and the crude oil product was recrystallized from 2-propanol/*di-iso*-propyl ether to give 0.22 g (33%) of final brown solid product **8**.

**<sup>1</sup>H NMR** (δ-DMSO-d<sub>6</sub>): 1.23-1.75 (m, 12 H, CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>+NHCOCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CHS), 1.93 (q, 3H, CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CO, J<sub>3</sub>=6.6 Hz), 2.08 (m, 4H, CH<sub>2</sub>CONH+NHCOCH<sub>2</sub>), 2.56 (d, 1H, CHS, J<sub>3</sub>=6.7 Hz), 2.66 (t, 3H, N=CCH<sub>2</sub>, J<sub>3</sub>=6.8 Hz), 2.81 (dd, 1H, CH<sub>2</sub>S, J<sub>2</sub>=10.1 Hz, J<sub>3</sub>=6.6 Hz), 2.99 (m, 4H, NHCOCH<sub>2</sub> + CH<sub>2</sub>NHCO), 3.09 (dd, 1H, CH<sub>2</sub>S, J<sub>2</sub>=10.1 Hz, J<sub>3</sub>=6.6 Hz), 3.22 (s, 3H, CON(CH<sub>3</sub>)C=C), 3.42 (s, 3H, CON(CH<sub>3</sub>)CO), 4.12 (m, 1H, NHCHCHS), 4.29 (dd, 1H, NHCHCH<sub>2</sub>S, J<sub>3</sub>=6.6 Hz), 6.36 (s, 1H, NHCONH), 6.42 (s, 1H, NHCONH), 7.76 (s, 2H, NH (linker)), 12.1 (s, 1H, NH (theophylline)).

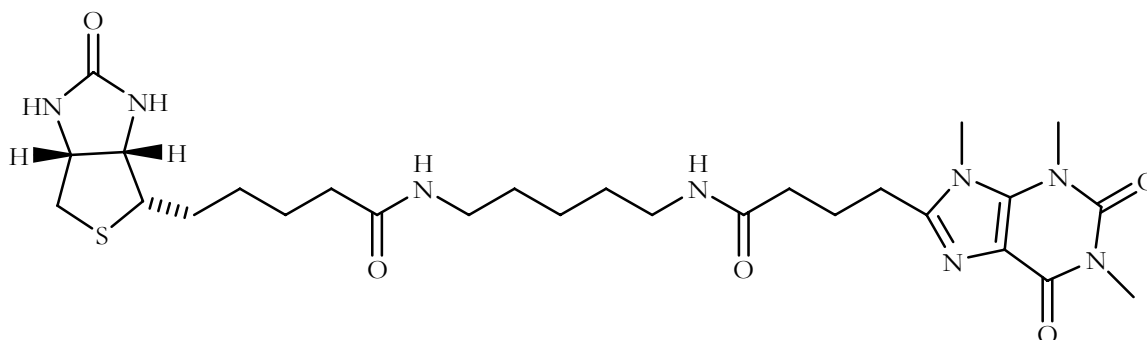
**<sup>13</sup>C NMR** (δ-DMSO-d<sub>6</sub>): 23.50 (CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>); 24.53 (NHCOCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CHS); 24.98 (NHCOCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CHS), 26.81 (NHCOCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CHS), 29.11 (N(CH<sub>3</sub>)), 30.13 (NH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NHCO), 31.30 (C=NCH<sub>2</sub>), 32.34 (N(CH<sub>3</sub>)), 36.82 (CH<sub>2</sub>CONH), 40.47 (CH<sub>2</sub>NHCO), 55.55 (CHS), 57.75 (CH<sub>2</sub>S), 59.75 (NHCHCH<sub>2</sub>S), 61.61 (NHCHCHS), 106.43 (C=CNH), 148.75 (C=CNH), 154.55 (N(CH<sub>3</sub>)CON(CH<sub>3</sub>)), 153.67

(N=CCH<sub>2</sub>), 154.55 (N(CH<sub>3</sub>)COC=C), 163.28 (NHCONH), 171.79 (CH<sub>2</sub>CONH), 172.36 (CH<sub>2</sub>NHCO).

**MS** m/z: 577.3 [M+H]<sup>+</sup>; 599.3 [M+Na]<sup>+</sup>.

**IR** (cm<sup>-1</sup>): 3338 (N-H); 1820 (C=O).

**m.p.**: 178-180 °C.



**N-(5-(4-(2,3,6,9-tetrahydro-1,3,9-trimethyl-2,6-dioxo-1H-purin-8-yl)butanamido)pentyl)-5-((3aR,6S,6aS)-hexahydro-2-oxo-1H-thieno[3,4-d]imidazol-6-yl)pentanamide (9).**

0.17 g of caffeine-8-butanoic acid **4** (0.624 mmol) were dissolved in anhydrous atmosphere in 5 mL of dry N,N-dimethylformamide. The solution was cooled to 0°C and 0.13 g of HOBT (0.624 mmol) were added. N-methyl morpholine was also added until pH=9 was reached. Then, 0.22 g of compound **7** (0.624 mmol) and 0.18 g of EDC-Cl were added, and the mixture was stirred at room temperature for 16 hours.

After this time the solvent was evaporated and the crude product was recrystallized from 2-propanol/di-*iso*-propyl ether to give final compound **9**. (0.19 g, 51%).

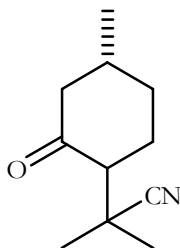
**<sup>1</sup>H NMR** (δ-DMSO-d<sub>6</sub>): 1.22-1.48 (m, 12H, CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>+NHCOCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CHS), 1.90 (quint., 2H, CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CO, J<sub>3</sub>=7.3 Hz), 2.08 (m, 4H, CH<sub>2</sub>CONH+NHCOCH<sub>2</sub>), 2.54 (m, 1H, CHS), 2.71 (t, 2H, CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CO, J<sub>3</sub>=7.4 Hz), 2.81 (dd, 1H, CH<sub>2</sub>S, J<sub>2</sub>=12.0 Hz, J<sub>3</sub>=6.5 Hz), 3.05 (m, 4H, CONHCH<sub>2</sub>+CH<sub>2</sub>NHCO), 3.14 (dd, 1H, CH<sub>2</sub>S, J<sub>2</sub>=12.0 Hz, J<sub>3</sub>=6.5 Hz), 3.21 (s, 3H, CON(CH<sub>3</sub>)C=C), 3.34 (s, 3H, CON(CH<sub>3</sub>)CO), 3.81 (s, 3H, CN(CH<sub>3</sub>)C), 4.11 (m, 1H, NHCHCHS), 4.32 (d, 1H, NHCHCH<sub>2</sub>S, J<sub>3</sub>=6.7 Hz), 6.36 (s, 1H, NHCONH), 6.42 (s, 1H, NHCONH).

**<sup>13</sup>C NMR** (δ-DMSO-d<sub>6</sub>): 20.77 (N=CN(CH<sub>3</sub>)) 23.90 (CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>), 24.61 (NHCOCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CHS), 25.02 (NHCOCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CHS), 26.63 (NHCOCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CHS), 29.13 (N(CH<sub>3</sub>)), 30.44 (NH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NHCO), 31.22 (C=NCH<sub>2</sub>), 32.41 (N(CH<sub>3</sub>)), 37.01 (CH<sub>2</sub>CONH), 40.43 (CH<sub>2</sub>NHCO), 55.46 (CHS), 57.32 (CH<sub>2</sub>S), 59.69 (NHCHCH<sub>2</sub>S), 61.69 (NHCHCHS), 106.40 (C=CNH), 148.80 (C=CNH), 155.02 (N(CH<sub>3</sub>)COC=C), 164.64 (NHCONH), 171.67 (CH<sub>2</sub>CONH), 172.41 (CH<sub>2</sub>NHCO).

**MS** m/z: 591 [M+H]<sup>+</sup>.

IR (cm<sup>-1</sup>): 3345 (N-H); 1825 (C=O).

m.p.: 183-185 °C.



**2-methyl-2-((4R)-4-methyl-2-oxocyclohexyl)propanenitrile (24).**

2 g of (+)-pulegone (18.39 mmol, d=0.936) were dissolved in anhydrous atmosphere in dry toluene. 36.7 mL of 1M diethylaluminium cyanide (in hexane, 36.78 mmol) were added, and 1,56 g of acetone cyanohydrine (18.39 mmol, d=0.932) were then immediately added. The mixture was stirred for 1 hour at room temperature. After, 6 g of NaF (73.56 mmol) and 2.6 mL of water (73.56 mmol) were added carefully. The mixture was stirred for 20 minutes again and 6 g of anhydrous sodium sulfate were finally added.

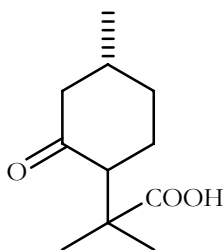
The resulting solid was eliminated by filtration under vacuum, solvent was evaporated and the crude product was purified by flash chromatography on silica with diethyl ether/petroleum ether=8/2 as eluent mixture obtaining 2.3 g (70%) of final oily pure product.

<sup>1</sup>H NMR (δ-CDCl<sub>3</sub>): 1.06 (d, 3H, CH<sub>3</sub>, J=7.7 Hz), 1.46 (s, 6H, C(CH<sub>3</sub>)<sub>2</sub>CN), 1.63-1.88 (m, 4H, CHCH<sub>2</sub>CH<sub>2</sub>CH), 1.97 (m, 1H, CH<sub>3</sub>CHCH<sub>2</sub>), 2.06-2.31 (m, 2H, CHCH<sub>2</sub>CO), 2.2 (m, 1H, CHC(CH<sub>3</sub>)<sub>2</sub>CN).

<sup>13</sup>C NMR (δ-CDCl<sub>3</sub>): 22.23 (CH<sub>3</sub>CH), 24.72 (CH<sub>3</sub>CCN), 28.90 (CHCH<sub>2</sub>CH<sub>2</sub>CH), 32.07 (C(CH<sub>3</sub>)<sub>2</sub>CN) 33.83 (CH<sub>3</sub>CH), 35.84 (CHCH<sub>2</sub>CH<sub>2</sub>CH), 51.24 (CHCH<sub>2</sub>CO), 56.45 (COCHCH<sub>2</sub>), 128.26 (CN), 208.76 (CO).

MS m/z: 180.1 [M+H]<sup>+</sup>.

IR (cm<sup>-1</sup>): 2251 (CN); 1667 (C=O).



**2-methyl-2-((4R)-4-methyl-2-oxocyclohexyl)propanoic acid (25)**

0.1 g of compound **24** were treated with a mixture of water/sulfuric acid=20/1. The solution was stirred at room temperature for 48 hours. After this time, 4N NaOH was added until pH=3. The

resulting suspension was extracted with ethyl acetate (8 \* 15 mL); the combined organic phases were washed with brine, dried over anhydrous sodium sulfate and evaporated to give 0.1 g (83%) of pure product that can be directly used for further step.

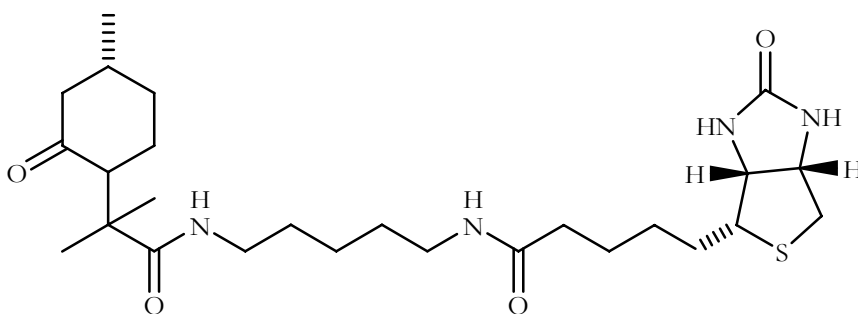
**<sup>1</sup>H NMR** ( $\delta$ -CDCl<sub>3</sub>): 1.08 (d, 3H, CH<sub>3</sub>, J<sub>3</sub>=6.9 Hz), 1.43 (s, 6H, C(CH<sub>3</sub>)COOH), 1.60-1.90 (m, 4H, CHCH<sub>2</sub>CH<sub>2</sub>CH), 1.97 (m, 1H, CH<sub>3</sub>CHCH<sub>2</sub>), 2.05-2.31 (m, 2H, CHCH<sub>2</sub>CO), 2.20 (m, 1H, CHC(CH<sub>3</sub>)<sub>2</sub>COOH).

**<sup>13</sup>C NMR** ( $\delta$ -CDCl<sub>3</sub>): 21.85 (CH<sub>3</sub>CH), 22.26 (CH<sub>3</sub>CCOOH), 28.97 (CHCH<sub>2</sub>CH<sub>2</sub>CH), 32.62 (C(CH<sub>3</sub>)<sub>2</sub>COOH), 33.79 (CH<sub>3</sub>CH), 36.01 (CHCH<sub>2</sub>CH<sub>2</sub>CH), 46.88 (CHCH<sub>2</sub>CO), 52.14 (COCHCH<sub>2</sub>), 182.96 (COOH), 208.11 (CO).

**MS** m/z: 199.1 [M+H]<sup>+</sup>; 197.1 [M-H]<sup>+</sup>.

**IR** (cm<sup>-1</sup>): 3541 (COOH); 1725 (C=O).

**m.p.:** 80-82 °C.



**N-(5-(2-methyl-2-((4R)-4-methyl-2-oxocyclohexyl)propanamido)pentyl)-5-((3aR, 4R, 6aS)-hexahydro-2-oxo-1H-thieno[3,4-d]imidazol-4-yl)pentanamide (26)**

0.5 g of compound **25** (1.22 mmol) were dissolved in anhydrous acetonitrile. 0.12 g of HOBT (1.22 mmol) were added with N-methylmorpholine (0.2 mL, 2.44 mmol) until pH=8.5. 0.8 g of compound **6** were then added and the mixture was cooled at 0°C. 0.21 g of EDC-Cl were finally added. The mixture was stirred under an anhydrous atmosphere in anhydrous for 16 hours.

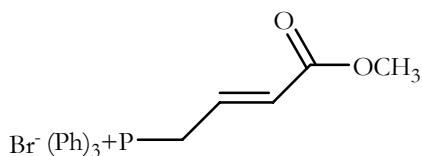
After this time, solvent was distilled with distillation at reduce pressure and the final crude product was crystallized with acetone/n-hexane, to give 0.81 g (63.2%) of the final pure product.

**<sup>13</sup>C NMR** ( $\delta$ -DMSO-d<sub>6</sub>): 20.22 (CH<sub>3</sub>CH), 21.14 (CH<sub>2</sub>CHCO), 23.11 (CH<sub>3</sub>C), 24.03 (NHCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NH), 25.37 (NHCOCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH), 26.81 (CH<sub>2</sub>CHS), 30.01 (NHCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NH), 32.33 (CH<sub>3</sub>CHCH<sub>2</sub> menthone), 33.91 (CH<sub>3</sub>CH), 36.55 (NHCOCH<sub>2</sub>), 40.47 (NHCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NH), 40.99 (C(CH<sub>3</sub>)<sub>2</sub>CO), 41.03 (NHCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NH), 41.12 (CH<sub>2</sub>S), 49.33 (CH<sub>3</sub>CHCH<sub>2</sub>CO menthone), 55.55 (CHS), 60.44 (CHCH<sub>2</sub>S), 61.21 (COCH menthone), 61.75 (CHCHS), 164.70 (NHCONH), 172.70 (CH<sub>2</sub>NHCOCH<sub>2</sub>), 178.05 (CCONHCH<sub>2</sub>), 211.51 (CH<sub>2</sub>COCH menthone).

**MS** m/z: 509.3 [M+H]<sup>+</sup>.

IR (cm<sup>-1</sup>): 3334 (N-H); 1713 (C=O).

m.p.: 165-167 °C.



**(37)**

To a solution of P(Ph)<sub>3</sub> (1.4 g, 5.5 mmol) in 6 mL of toluene, was added bromo-methyl ester (1g, 5.5 mmol) were added dropwise monitoring the temperature below 40°C. The reaction was stirred for 24 hours. At the end of this time a white solid precipitates was collected and washed with toluene and then with petroleum ether to give the final product white salt **37** (2.33 g; 95%).

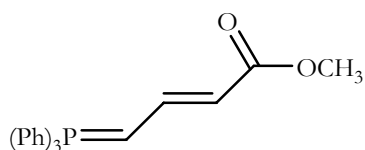
<sup>1</sup>H NMR (δ-CD<sub>3</sub>OD): 3.66 (s, 3H, COOCH<sub>3</sub>); 4.74-4.83 (m, 2H, CH<sub>2</sub>P, J<sub>H-P</sub>=22.0 Hz, J<sub>3</sub>=7.1 Hz); 6.24 (d, 1H, CHCOOCH<sub>3</sub>, J<sub>trans</sub>=15.4 Hz); 6.80 (m, 1H, CHCH<sub>2</sub>P); 7.76-7.92 (m, 15H, CH<sub>aromatic</sub>).

<sup>13</sup>C NMR (δ-CD<sub>3</sub>OD): 26.30 (d, CH<sub>2</sub>P, J<sub>C-P</sub>=131 Hz); 47.79 (OCH<sub>3</sub>); 118.12 (PC<sub>aromatic</sub> J<sub>C-P</sub>=130 Hz); 129.43 (CHCOOCH<sub>3</sub>); 130.46 (CH<sub>aromatic</sub>); 133.76 (CHCH<sub>2</sub>P); 135.41 (CH<sub>aromatic</sub>); 165.27 (COOCH<sub>3</sub>).

MS m/z: 441.1 [M+H]<sup>+</sup>.

IR (cm<sup>-1</sup>): 1762 (C=O); 1425 (P-CH<sub>2</sub>).

m.p.: 143-145 °C.



**(38).**

1 g (2.27 mmol) of the previous compound **37** was dissolved in 40 mL of distilled water. NaOH (2.27 mmol) were added under stirring; immediately, a yellow precipitate is found.

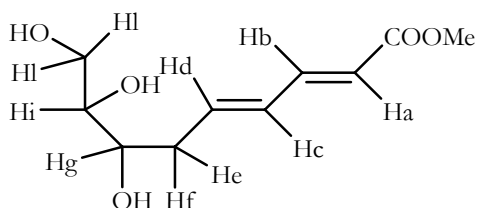
The precipitate was filtered under vacuum and washed with water (3 x 10 mL), dried under vacuum and recrystallized from toluene to give 0.74 g (90%) of pure yellow solid product **38**.

<sup>1</sup>H NMR (δ-CDCl<sub>3</sub>): 3.56 (s, 3H, COOCH<sub>3</sub>); 3.75 (m, 1H, PCH, J<sub>H-P</sub>=17.4 Hz); 5.12 (d, 1H, CHCOOMe); 7.25 (m, 1H, CHCHCOOMe); 7.57 (m, 15H, CH<sub>aromatic</sub>).

<sup>13</sup>C NMR (δ-CDCl<sub>3</sub>): 48.52 (P=CH, J<sub>C-P</sub>=133 Hz); 90.10 (CHCOOMe); 126.17 (CH<sub>aromatic</sub>); 133.27 (CH<sub>aromatic</sub>); 138.32 (C<sub>aromatic</sub>); 152.04 (CHCHCOOMe); 170.85 (COOMe).

**MS** m/z: 361.1 [M+H]<sup>+</sup>.

**IR** (cm<sup>-1</sup>): 1783 (C=O).



**(2E,4E)-methyl 7,8,9-trihydroxynona-2,4-dienoate (39).**

3.10 g (8.60 mmol) compound **38** were dissolved at anhydrous conditions in 20 mL of anhydrous tetrahydrofuran, and 1.22 gr of 2-deoxy-D-ribose (9.12 mmol) were quickly added. The suspension was heated at 80°C for 6 hours.

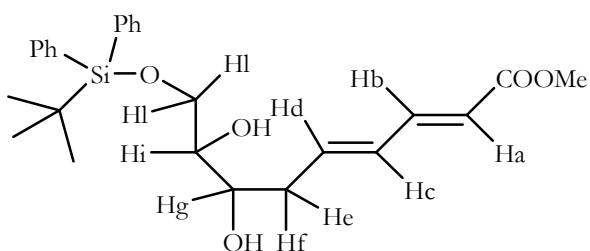
After this time the solvent was removed and the crude solid product was purified by flash chromatography (dichlorometane/methanol=9/1) to give triol **39** (0.85 gr, 37.6%).

**<sup>1</sup>H NMR** (δ-CD<sub>3</sub>OD): 2.33 (m, 1H, H<sub>l</sub>); 2.55 (m, 1H, H<sub>i</sub>); 3.45 (m, 1H, H<sub>l</sub>); 3.58 (m, 3H, H<sub>g</sub>+H<sub>i</sub>); 3.71 (s, 3H, COOCH<sub>3</sub>); 5.83 (dd, 1H, H<sub>b</sub>, J<sub>trans</sub>=14.8 Hz); 6.30 (m, 2H, H<sub>d</sub>+H<sub>a</sub>); 7.27 (dd, 1H, H<sub>c</sub>, J<sub>trans</sub>=15.2 Hz).

**<sup>13</sup>C NMR** (δ-CD<sub>3</sub>OD): 38.16 (CH<sub>e-i</sub>); 52.25 (COOCH<sub>3</sub>); 65.00 (CH<sub>2</sub>OH); 73.24 (CH<sub>g</sub>); 76.12 (CH<sub>i</sub>); 122.25 (CH<sub>b</sub>); 132.03 (CH<sub>d</sub>); 141.83 (CH<sub>a</sub>); 147.19 (CH<sub>c</sub>); 169.91 (COOCH<sub>3</sub>).

**MS** m/z: 592.9 [M+H]<sup>+</sup>; 590.3 [M-H]<sup>+</sup>.

**IR** (cm<sup>-1</sup>): 3212 (OH); 1775 (C=O).



**(40).**

0.2 gr of previous compound **39** (1.12 mmol) were dissolved in 300 μL of pyridine and 5 mL of dichlorometane at anhydrous conditions. DIMAP (4-dimethyl aminopyridine) was added to the reaction (0.01 g, 0.08 mmol) and the suspension was stirred at 0°C for 15 min. After this time, 0.38 mL of tert-Butyl-diphenylsilylchloride (0.4 gr, 1.46 mmol) were added to the suspension.

The reaction was kept under stir at 0°C for 15 minutes and then for 48 hours at room temperature.

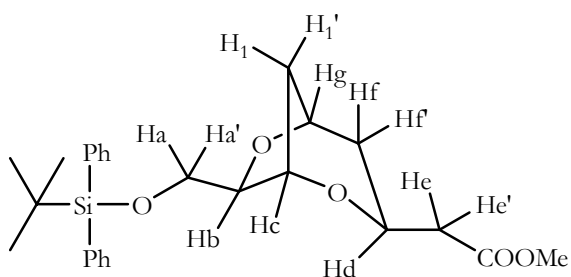
After this time 1 M chloridric acid was added until pH 3; the aqueous layer was extracted with ethyl acetate (3 x 10 mL) and the combined organic layers were washed with brine and dried over anhydrous sodium sulphate. The solvent was eliminated by distillation at reduce pressure to give pure mono-protected alcohol **40**.

**<sup>1</sup>H NMR** ( $\delta$ -CDCl<sub>3</sub>): 1.07 (s, 9H, SiC(CH<sub>3</sub>)<sub>3</sub>); 2.19 (m, 1H, H<sub>e</sub>); 2.21 (m, 2H, H<sub>i</sub>); 2.35 (m, 2H, H<sub>f</sub>); 3.59 (m, 1H, H<sub>j</sub>); 3.72 (s, 3H, COOCH<sub>3</sub>); 3.80-3.82 (m, 3H, H<sub>g</sub>+ 2OH); 5.81 (d, 1H, H<sub>a</sub>, J<sub>trans</sub>=14.8 Hz); 6.13 (m, 1H, H<sub>d</sub>); 6.21 (m, 1H, H<sub>c</sub>); 7.26 (m, 1H, H<sub>b</sub>); 7.39-7.65 (m, 10H, CH<sub>aromatic</sub>).

**<sup>13</sup>C NMR** ( $\delta$ -CDCl<sub>3</sub>): 19.07 (CH<sub>f</sub>); 26.75 (CH<sub>i</sub>); 36.47 (CH<sub>g</sub>); 51.49 (COOCH<sub>3</sub>); 64.65 (CH<sub>j</sub>); 71.74 (SiC(CH<sub>3</sub>)<sub>3</sub>); 100.56 (CH<sub>e</sub>); 121.91 (CH<sub>d</sub>); 128.00 (CH<sub>a</sub>); 132.73 (CH<sub>b</sub>); 133.97 (SiC<sub>aromatic</sub>); 135.63 (SiC<sub>aromatic</sub>); 139.72 (CH<sub>aromatic</sub>); 144.68 (CH<sub>aromatic</sub>); 167.70 (COOCH<sub>3</sub>).

**MS** m/z: 831.3 [M+H]<sup>+</sup>; 829.2 [M-H]<sup>+</sup>.

**IR** (cm<sup>-1</sup>): 3229 (OH); 1754 (C=O).



**(41).**

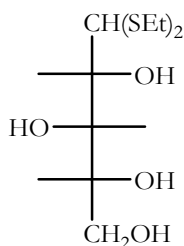
0.4 g of compound **40** (0.89 mmol) were dissolved under anhydrous conditions. 0.034 g of potassium tert-butoxide (0.26 mmol) were dissolved in 0.55 mL anhydrous tert-butyl alcohol.

This solution was added to the solution of **40** in toluene and the resulting suspension was heated at 50°C for 24 hours. After this time the solvent was eliminated and the final crude product was purified on a short flash silica column with a mixture ethyl acetate/petroleum ether = 2/8, to give the final pure product (0.089 gr; 22%).

**<sup>1</sup>H NMR** ( $\delta$ -CDCl<sub>3</sub>): 1.07 (s, 9H, SiC(CH<sub>3</sub>)<sub>3</sub>); 1.40 (m, 1H, H<sub>f</sub>); 1.69 (d, 1H, H<sub>e</sub>, J<sub>3</sub>=7.7 Hz); 1.83 (m, 1H, H<sub>e</sub>); 2.51 (m, 2H, H<sub>a</sub>+ H<sub>a'</sub>); 3.42 (m, 1H, H<sub>i</sub>); 3.63 (m, 1H, H<sub>i</sub>); 3.69 (s, 3H, COOCH<sub>3</sub>); 4.37 (m, 1H, H<sub>d</sub>); 4.45 (m, 1H, H<sub>b</sub>); 4.50 (m, 1H, H<sub>c</sub>); 4.58 (m, 1H, H<sub>j</sub>); 7.40 (m, 5H, CH<sub>aromatic</sub>); 7.63 (m, 5H, CH<sub>aromatic</sub>).

**<sup>13</sup>C NMR** ( $\delta$ -CDCl<sub>3</sub>): 19.28 (SiC(CH<sub>3</sub>)<sub>3</sub>); 26.90 (SiC(CH<sub>3</sub>)<sub>3</sub>); 38.48 (CH<sub>f</sub>); 39.26 (CH<sub>i</sub>); 41.11 (CH<sub>e</sub>); 51.84 (COOCH<sub>3</sub>); 64.72 (CH<sub>a</sub>); 67.57 (CH<sub>d</sub>); 73.42 (CH<sub>g</sub>); 74.34 (CH<sub>e</sub>); 127.80 (CH<sub>b</sub>); 133.26 (CH<sub>aromatic</sub>); 135.65 (C<sub>aromatic</sub>); 171.51 (COOCH<sub>3</sub>).

**MS** m/z: 455.2 [M+H]<sup>+</sup>.



**5,5-bis(ethylthio)-2,3,4-trimethylpentane-1,2,3,4-tetraol (32).**

5 g of D-(+)-xylose (3.33 mmol) were dissolved in 1.25 mL of ethanethiol (0.78 gr, 3.33 mmol) and the mixture was cooled with a ice bath to 0°C. 0.83 gr of zinc chloride were added and, carefully, 5 mL of concentrated hydrochloridric acid. The reaction was kept under stirring at room temperature for 20 min., then 20 mL of water were added.

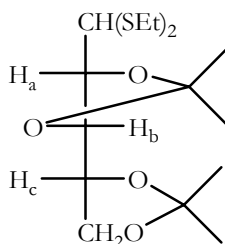
The solution was concentrated and neutralized with 5M NaOH. After this the solvent was evaporated to give 0.694 gr of protected sugar **32** (70 %).

**<sup>1</sup>H NMR** (δ-D<sub>2</sub>O): 1.28 (t, 6H, SCH<sub>2</sub>CH<sub>3</sub>); 2.74 (m, 4H, SCH<sub>2</sub>CH<sub>3</sub>); 3.71 (m, 2H, CH<sub>2</sub>OH); 3.85 (m, 1H, CHCH<sub>2</sub>OH); 3.91 (t, 1H, CHCHCH<sub>2</sub>OH, J<sub>3</sub>=7.3 Hz); 4.02 (t, 1H, CH(SEt)<sub>2</sub>CH, J<sub>3</sub>=7.2 Hz); 4.13 (d, 1H, CH(SEt)<sub>2</sub>, J<sub>3</sub>=7.3 Hz).

**<sup>13</sup>C NMR** (δ-D<sub>2</sub>O): 14.39 (SCH<sub>2</sub>CH<sub>3</sub>); 27.39 (SCH<sub>2</sub>CH<sub>3</sub>); 53.15 (HC(SEt)<sub>2</sub>); 65.94 (CH<sub>2</sub>OH); 75.28 (CHCH(SEt)<sub>2</sub>); 78.72 (CHCH<sub>2</sub>OH); 80.11 (CHCHCH<sub>2</sub>OH).

**MS** m/z: 298.1 [M+H]<sup>+</sup>; 296.1 [M-H]<sup>+</sup>.

**IR** (cm<sup>-1</sup>): 3522 (OH); 2548 (C-S).



**(33).**

300 mg of iodine was dissolved in 50 mL of acetone; 1 g of previous compound **32** was added and the solution was kept under stirring for 3 hours at room temperature. After this time the mixture was heated over reflux for 20 minutes until the solution turns to brown from initial violet. 0.5N NaOH was then added in order to destroy iodine. The mixture was extracted with dichloromethane (3 x 20 mL) and the combined organic layer were washed with water (4 x 20 mL) and dried over anhydrous Na<sub>2</sub>SO<sub>4</sub>.

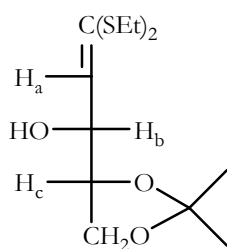
The solvent was evaporated to give 1.03 g of acetone **33** (82%).

$^1\text{H NMR}$  ( $\delta$ - $\text{CDCl}_3$ ): 1.23-1.47 (m, 18H,  $\text{S}(\text{CH}_2\text{CH}_3)_2 + \text{C}(\text{CH}_3)_2$ ); 2.64 (m, 4H,  $\text{S}(\text{CH}_2\text{CH}_3)_2$ ); 4.05 (m, 3H,  $\text{CH}_2\text{O} + \text{CH}$ ); 4.27 (m, 1H,  $\text{CH}_b$ ); 4.49 (m, 1H,  $\text{CH}_a$ ); 5.99 (m, 1H,  $\text{CH}(\text{SEt})_2$ )

$^{13}\text{C NMR}$  ( $\delta$ - $\text{CDCl}_3$ ): 14.74 ( $\text{SCH}_2\text{CH}_3$ ); 26.21- 26.23 ( $\text{C}(\text{CH}_3)_2$ ); 26.85 ( $\text{SCH}_2\text{CH}_3$ ); 53.82 ( $\text{CH}(\text{SEt})_2$ ); 71.67 ( $\text{CH}_2\text{O}$ ); 73.25 ( $\text{CH}_c$ ); 84.70 ( $\text{CH}_b$ ); 85.75 ( $\text{CH}_a$ ); 105.24 ( $\text{C}(\text{CH}_3)_2$ ); 111.79 ( $\text{C}(\text{CH}_3)_2$ ).

**MS** m/z: 337.2  $[\text{M}+\text{H}]^+$ .

**IR** ( $\text{cm}^{-1}$ ): 2555 (C-S).



**(34).**

A solution of 2.0 gr (5.95 mmol) of compound **33** in 20 mL of THF was added dropwise at anhydrous conditions to a solution of 1.0 gr (8.92 mmol) of potassium tert-butoxide in THF/DMSO=3/1 as solvent; the reaction was kept for 1 hour at 23°C.

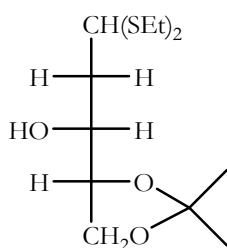
After this time the mixture was poured in 25 mL of ice and extracted (6 x 15 mL) with ethyl acetate. The combined organic layers were washed with water (2 x 20 mL) and dried with sodium sulfate; the solvent was evaporated to give final 1.02 gr (62%) of pure product **34**.

$^1\text{H NMR}$  ( $\delta$ - $\text{CDCl}_3$ ): 1.00-1.37 (m, 12H,  $\text{S}(\text{CH}_2\text{CH}_3)_2 + \text{C}(\text{CH}_3)_2$ ); 2.60-2.84 (m, 4H,  $\text{S}(\text{CH}_2\text{CH}_3)_2$ ); 3.69 (m, 1H,  $\text{CH}_2\text{O}$ ); 3.88 (m, 1H,  $\text{CH}_2\text{O}$ ); 4.01 (m, 1H,  $\text{CH}_c$ ); 4.68 (m, 1H,  $\text{CH}_b$ ); 5.81 (d, 1H,  $\text{CH}_a$ ;  $J_3 = 8.2$  Hz).

$^{13}\text{C NMR}$  ( $\delta$ - $\text{CDCl}_3$ ): 14.12 ( $\text{S}(\text{CH}_2\text{CH}_3)_2$ ); 26.38- 26.70 ( $\text{C}(\text{CH}_3)_2$ ); 27.59 ( $\text{S}(\text{CH}_2\text{CH}_3)_2$ ); 65.81 ( $\text{CH}_2\text{O}$ ); 70.69 ( $\text{CH}_b$ ); 78.87 ( $\text{CH}_c$ ); 109.80 ( $\text{C}(\text{CH}_3)_2$ ); 132.78 ( $\text{CH}_a$ ); 135.94 ( $\text{C}(\text{SCH}_2\text{CH}_3)_2$ ).

**MS** m/z: 279.1  $[\text{M}+\text{H}]^+$ ; 277.1  $[\text{M}-\text{H}]^+$ .

**IR** ( $\text{cm}^{-1}$ ): 3467 (OH).



**(35).**

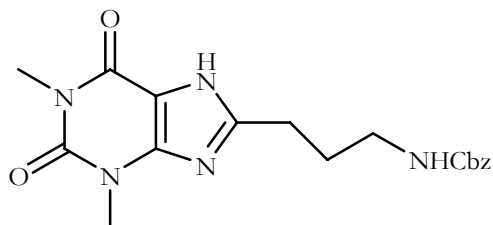
0.17 g of  $\text{LiAlH}_4$  (4.68 mmol) were dissolved in 20 mL of anhydrous THF and the solution was cooled at  $-60^\circ\text{C}$ . A solution of 0.32 g (1.17 mmol) of compound **34** in 10 mL of anhydrous THF was added dropwise and the reaction was kept at  $-60^\circ\text{C}$  for 3 hours. After this time the mixture was stirred at room temperature for 1 hour. After this were added at the organic saturated ammonium chloride until  $\text{pH}=6$ ; the aqueous phase was extracted with ethyl acetate (4 x 20 mL), the combined organic layers were dried with sodium sulphate and the solvent was eliminated to give a crude brown oil that was purified with flash chromatography on silica with the eluent mixture petroleum ether/diethyl ether = 1/1 to obtain the desired reduced product (0.174 gr, 53%).

**$^1\text{H NMR}$**  ( $\delta\text{-CDCl}_3$ ): 1.22-1.45 (m, 12H,  $\text{S}(\text{CH}_2\text{CH}_3)_2 + \text{C}(\text{CH}_3)_2$ ); 1.73 (m, 1H,  $\text{CH}_2\text{CH}(\text{SEt})_2$ ); 2.01 (m, 1H,  $\text{CH}_2\text{CH}(\text{SEt})_2$ ); 2.56-2.72 (m, 4H,  $\text{S}(\text{CH}_2\text{CH}_3)_2$ ); 3.79 (m, 1H,  $\text{CHOH}$ ); 3.88 (s, 1H,  $\text{CH}(\text{SEt})_2$ ); 4.01 (m, 2H,  $\text{CH}_2\text{O}$ ); 4.08 (m, 1H,  $\text{CHO}$ ).

**$^{13}\text{C NMR}$**  ( $\delta\text{-CDCl}_3$ ): 14.56 ( $\text{S}(\text{CH}_2\text{CH}_3)_2$ ); 24.44- 25.28 ( $\text{C}(\text{CH}_3)_2$ ); 28.97 ( $\text{S}(\text{CH}_2\text{CH}_3)_2$ ); 40.47 ( $\text{CH}_2\text{CH}(\text{SEt})_2$ ); 47.81 ( $\text{CH}(\text{SEt})_2$ ); 66.05 ( $\text{CH}_2\text{O}$ ); 69.70 ( $\text{CHOH}$ ); 78.66 ( $\text{CHO}$ ); 109.62 ( $\text{C}(\text{CH}_3)_2$ ).

**MS**  $m/z$ : 281.1  $[\text{M}+\text{H}]^+$ .

**IR** ( $\text{cm}^{-1}$ ): 3423 (OH).



**benzyl 3-(2,3,6,7-tetrahydro-1,3-dimethyl-2,6-dioxo-1H-purin-8-yl)propylcarbamate (13).**

A solution of 0.36 g of theophylline-8-butanoic acid **1** (1.33 mmol) dissolved in 50 mL of anhydrous toluene were added in anhydrous conditions to a solution of 0.15 g of diphenyl phosphorylazide (0.11 mL, 1.622 mmol) dissolved in 20 mL of anhydrous toluene. The solution was stirred at room temperature for 10 minutes and then 0.24 gr of triethylamine (0.23 mL, 1.62 mmol) was added. The suspension was heated under reflux for 1 hour.

The solution was then cooled down to room temperature, and then 0.26 gr of benzil alcohol (0.28 mL, 2.704 mmol,  $d=1.042$ ) added. The solution was refluxed again for 12 hours.

After 12 hours the solvent was eliminated; the crude solid product was dissolved in ethyl acetate (20 mL) and the organic layer was washed with saturated solution of  $\text{NaHCO}_3$ , brine, and then dried on  $\text{Na}_2\text{SO}_4$ .

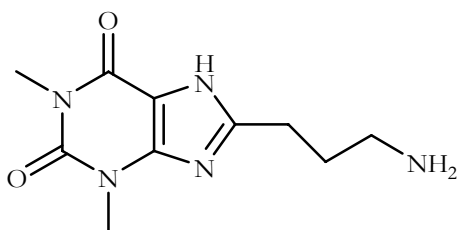
The solvent was removed and the crude product was purified with flash chromatography on silica gel using as eluent mixture ethyl acetate/petroleum ether = 2/8 to give 0.102 gr (61%) of pure oil colourless product.

**<sup>1</sup>H NMR** (δ-DMSO-d<sub>6</sub>): 2.20 (quint., 2H, CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NHCbz, J<sub>3</sub>=7.1 Hz), 2.42 (t, 2H, CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NHCbz, J<sub>3</sub>=7.2 Hz), 2.94 (t, 2H, CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NHCbz, J<sub>3</sub>=7.2 Hz), 3.43 (s, 3H, CON(CH<sub>3</sub>)CO), 3.60 (s, 3H, N(CH<sub>3</sub>)CO), 5.09 (s, 2H, OCH<sub>2</sub>Ph), 7.31 (m, 5H, CH<sub>aromatic</sub> (Cbz)).

**<sup>13</sup>C NMR** (δ-DMSO-d<sub>6</sub>): 28.70 (CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NHCbz), 29.10 (CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NHCbz), 29.82 (N(CH<sub>3</sub>)CO), 32.30 (CON(CH<sub>3</sub>)CO), 41.50 (CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>COOH), 65.90 (OCH<sub>2</sub>Ph), 114.52 (NHC=CN), 127.21-129.1 (CH<sub>aromatic</sub> (Cbz)), 149.91 (NHC=CN), 151.44 (N(CH<sub>3</sub>)CON(CH<sub>3</sub>)), 154.96 (N(CH<sub>3</sub>)CO), 156.80 (CO (Cbz)), 157.12 (C=N).

**MS** m/z: 372.17 [M+H]<sup>+</sup>; 394.17 [M+Na]<sup>+</sup>.

**IR:** 3067 (NH); 1787 (C=O (Cbz)).



**8-(3-aminopropyl)-1,3-dimethyl-1H-purine-2,6(3H,7H)-dione**

**(14).**

80 mg of 10% Pd/C were added to a solution of 0.89 g of previous compound **13** in 70 mL of methanol. The suspension was stirred under a hydrogen flux for 24 hours.

Ater this time the catalyst was removed over Celite, methanol was evaporate to give 0.559 gr (98%) of pure solid white product.

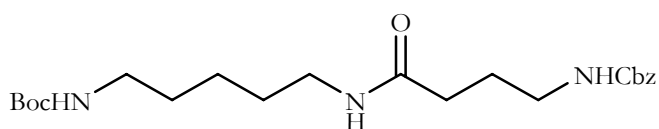
**<sup>1</sup>H NMR** (δ-DMSO-d<sub>6</sub>): 1.87 (m, 2H, CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NH<sub>2</sub>), 2.51 (t, 3H, CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NH<sub>2</sub>, J<sub>3</sub>=6.2 Hz), 2.69 (m, 2H, CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NH<sub>2</sub>), 3.19 (s, 3H, CON(CH<sub>3</sub>)CO), 3.40 (s, 3H, N(CH<sub>3</sub>)CO).

**<sup>13</sup>C NMR** (δ-DMSO-d<sub>6</sub>): 28.4 (CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NH<sub>2</sub>), 29.20 (CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NH<sub>2</sub>), 29.54 (N(CH<sub>3</sub>)CO), 30.60 (CON(CH<sub>3</sub>)CO), 41.77 (CH<sub>2</sub>NH<sub>2</sub>), 114.65 (NHC=CN), 150.02 (NHC=CN), 152.05 (N(CH<sub>3</sub>)CON(CH<sub>3</sub>)), 155.21 (N(CH<sub>3</sub>)CO), 157.38 (C=N).

**MS** m/z: 274.1 [M+H]<sup>+</sup>; 296.1 [M+Na]<sup>+</sup>.

**m.p.:** 100-103 °C

**IR** (cm<sup>-1</sup>): 3106 (N-H); 1715 (C=O).



**(16).**

0.32 g of  $\gamma$ -aminobutyric acid (1.173 mmol) dissolved in anhydrous conditions in 40 mL of anhydrous  $\text{CH}_3\text{CN}$ . 0.16 g of HOBT (1.173 mmol) and 0.258 mL of N-methylmorpholine (1.408 mmol,  $d=0.92$ ) checking that the pH value should be 8.

The mixture was cooled to  $0^\circ\text{C}$  and 0.24 g (1.173 mmol) of N-Boc-cadaverine and 0.27 gr (1.235 mmol) of EDC-Cl were added. The reaction was stirred at room temperature for 16 hours. After of this time the solvent was removed, the solid residue was dissolved with ethyl acetate (20 mL), the organic layer was washed 10% citric acid, saturated  $\text{NaHCO}_3$  solution and dried with sodium sulphate.

The solvent was eliminated to give 0.48 gr (58%) of pure solid product.

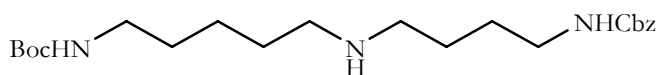
**$^1\text{H}$  NMR** ( $\delta$ - $\text{CDCl}_3$ ): 1.25-1.56 (m, 17H,  $\text{CH}_3(\text{Boc})+\text{CH}_2(\text{cadaverine})$ ), 1.79 (quint., 2H,  $\text{CH}_2\text{CH}_2\text{CH}_2\text{NHCbz}$ ,  $J_3=6.7$  Hz), 2.18 (t, 2H,  $\text{CH}_2\text{CO}$ ,  $J_3=6.7$  Hz), 3.06-3.21 (m, 6H,  $\text{CH}_2\text{NH}(\text{cadaverine})+\text{CH}_2\text{NHCbz}$ ), 4.67 (s, 1H,  $\text{NH}$ ), 5.05 (s, 1H,  $\text{CH}_2(\text{Cbz})$ ), 5.29 (s, 1H,  $\text{NH}$ ), 6.26 (s, 1H,  $\text{NH}$ ), 7.29 (s, 5H,  $\text{CH}_{\text{aromatic}}(\text{Cbz})$ ).

**$^{13}\text{C}$  NMR** ( $\delta$ - $\text{CDCl}_3$ ): 23.98 ( $\text{CH}_2\text{CH}_2\text{CH}_2(\text{cadaverine})$ ), 27.54 ( $\text{CH}_2\text{CH}_2\text{CH}_2\text{NHCbz}$ ), 28.88 ( $\text{CH}_3(\text{Boc})$ ), 29.71-30.30 ( $\text{CH}_2\text{CH}_2\text{CH}_2$ ), 33.64 ( $\text{COCH}_2\text{CH}_2\text{CH}_2\text{NHCbz}$ ), 41.45 ( $\text{CH}_2\text{NH}$ ), 65.31 ( $\text{CH}_2(\text{Cbz})$ ), 79.37 ( $\text{C}(\text{CH}_3)_3(\text{Boc})$ ), 127.26 ( $\text{CH}_{\text{aromatic}}(\text{Cbz})$ ), 141.28 ( $\text{C}_{\text{aromatic}}(\text{Cbz})$ ), 160.01 ( $\text{CO}$ ), 172.24 ( $\text{CO}$ ).

**MS**  $m/z$ : 238.2  $[\text{M}+\text{H}]^+$ ; 260  $[\text{M}+\text{Na}]^+$ .

**IR**: ( $\text{cm}^{-1}$ ): 3111 (N-H); 1762 (C=O).

**m.p.**: 87-89  $^\circ\text{C}$ .



**(17).**

1 gr of compound **16** (2.374 mmol) was dissolved in 45 mL of anhydrous tetrahydrofuran (THF) and cooled at  $-10^\circ\text{C}$ . 2.48 mL of 2 M solution in THF of borane-dimethylsulphur ( $\text{BH}_3\text{-DMS}$ ) (5 mmol) were carefully added and the mixture was stirred at room temperature for 12 hours.

After this time the solvent was removed, 60 mL of water were added and the aqueous layer was extracted with ethyl acetate (6 x 20 mL).

The combined organic layers were washed with brine and dried over anhydrous sodium sulphate; the solvent was then removed to give 0.48 gr (51%) of white solid product.

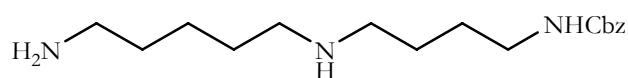
**<sup>1</sup>H NMR** ( $\delta$ -CDCl<sub>3</sub>): 1.24-1.85 (m, 19H, CH<sub>3</sub>(Boc)+CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NH+CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NHCbz), 2.74 (quint., 2H, CH<sub>2</sub>NHCO, J<sub>3</sub>=7.7 Hz), 3.06-3.18 (m, 4H, CH<sub>2</sub>NHCH<sub>2</sub>), 3.70 (s, 1H, NH), 4.68 (s, 1H, NH), 5.10 (s, 2H, CH<sub>2</sub>(Cbz)), 5.48 (s, 1H, NH) 7.32-7.34 (m, 5H, CH<sub>aromatic</sub> (Cbz)).

**<sup>13</sup>C NMR** ( $\delta$ -CDCl<sub>3</sub>): 24.04 (CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NH), 27.58 (CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub> NHCbz), 28.91 (CH<sub>3</sub> (Boc)), 29.68 (CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NH), 30.61 (CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NH), 36.46 (CH<sub>2</sub>NHBoc) 40.81 (CH<sub>2</sub>NHCbz) 41.61 (CH<sub>2</sub>NHCH<sub>2</sub>), 50.01 (CH<sub>2</sub>NH), 65.52 (CH<sub>2</sub> (Cbz)), 79.53 (C(CH<sub>3</sub>)<sub>3</sub> (Boc)), 126.33 (CH<sub>aromatici</sub> (Cbz)), 142.42 (C<sub>aromatic</sub> (Cbz)), 162.01 (CO), 170.24 (CO).

**MS** m/z: 408.2 [M+H]<sup>+</sup>; 430.4 [M+Na]<sup>+</sup>; 446.2 [M+K]<sup>+</sup>.

**IR:** (cm<sup>-1</sup>): 3120 (N-H); 1765 (C=O).

**m.p.:** 106-108 °C



**(18).**

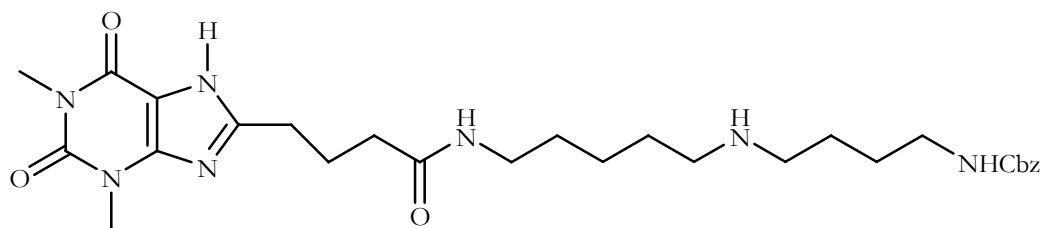
100 mg of compound **18** was dissolved in 1 mL of a mixture dichloromethane/trifluoroacetic acid = 6/4; the solution was stirred at room temperature for 1 hour. After this time of 10 mL of dichloromethane were added and the solvent was removed. In order to remove all excess of trifluoroacetic acid the resulting oil was coevaporated with dichloromethane (5 \* 10 mL) to give **18** as pure yellow oil (0.074 mg; 99%).

**<sup>1</sup>H NMR** ( $\delta$ -CDCl<sub>3</sub>): 1.24-1.87 (m, 10H, CH<sub>2</sub>(cadaverine)+CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NHCbz), 2.84-3.40 (m, 6H, CH<sub>2</sub>NHCH<sub>2</sub>+CH<sub>2</sub>NHCbz), 4.71-5.14 (m, 4H, NH<sub>2</sub>CH<sub>2</sub>+CH<sub>2</sub>(Cbz)), 7.18-7.35 (m, 5H, CH<sub>aromatic</sub> (Cbz)), 7.63 (s, 1H, NH), 8.26 (s, 1H, NH), 9.08 (s, 1H, NH).

**<sup>13</sup>C NMR** ( $\delta$ -CDCl<sub>3</sub>): 18.23 (NH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NH), 27.41-27.72 (CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NHCbz), 30.34 (NH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NH), 32.59 (NH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NH), 41.66 (CH<sub>2</sub>NHCbz), 42.15 (CH<sub>2</sub>NHCH<sub>2</sub>), 49.63 (CH<sub>2</sub>NH<sub>2</sub>), 65.41 (CH<sub>2</sub> (Cbz)), 127.22-141.23 (CH<sub>aromatici</sub> (Cbz)), 156.87 (CO (Cbz)).

**MS** m/z: 308.4 [M+H]<sup>+</sup>, 330.4 [M+Na]<sup>+</sup>.

**IR** (cm<sup>-1</sup>): 3320 (N-H); 1752 (C=O).



**(19).**

97 mg of theophylline-8-butanoic acid **1** (0.36 mmol) were dissolved in 5 mL of anhydrous acetonitrile. Mixture was cooled to 0°C and 0.049 g (0.36 mmol) of HOBt was added to the suspension; after removing the ice-bath, 0.08 mL of N-methylmorpholine (0.073 gr, 0.72 mmol, d=0.92) were added to the suspension to pH=8. 0.111 g of previous compound **18** (0.36 mmol) and 0.083 g of EDC-Cl (0.432 mmol) were added. The mixture was stirred for 16 hours.

At the end of time, the solvent was eliminated and the oily crude product was redissolved in ethyl acetate (15 mL) and washed with 10% citric acid, saturated NaHCO<sub>3</sub>, brine and dried over sodium sulphate.

Solvent was removed to give 0.212 g of crude product that was purified by crystallization (acetone/hexane) to obtain 0.096 gr (48%) of pure final solid product.

**<sup>1</sup>H NMR** (δ-CDCl<sub>3</sub>): 1.46 (m, 4H, NHCH<sub>2</sub>CH<sub>2</sub>NH (Gaba)), 1.93 (m, 6H, CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>), 2.09 (quint., 2H, CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CO, J<sub>3</sub>=6.5 Hz), 2.29 (m, 2H, CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CO), 2.81 (m, 2H, CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CO), 2.97 (m, 4H, CH<sub>2</sub>NHCH<sub>2</sub>), 3.14 (m, 2H, CH<sub>2</sub>NHCbz), 3.32 (m, 5H, CONHCH<sub>2</sub>+CONCH<sub>3</sub>), 3.54 (s, 3H, CON(CH<sub>3</sub>)CO), 5.01 (s, 2H, CH<sub>2</sub> (Cbz)), 6.70 (s, 1H, NH), 7.34 (m, 5H, CH<sub>aromatici</sub> (Cbz)), 9.22 (s, 1H, NH), 11.17 (s, 1H, NH).

**<sup>13</sup>C NMR** (δ-CDCl<sub>3</sub>): 19.22 (NHCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NH (cadaverina)), 25.65 (CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CO), 28.74-30.08 (CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub> (Gaba)+CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CO), 29.06 (N(CH<sub>3</sub>)CO), 33.52 (NHCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NH (cadaverina)), 35.00 (NHCH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>NH (cadaverina)), 36.14 (CH<sub>2</sub>CONH) 36.54 (CON(CH<sub>3</sub>)CO), 42.78-43.61 (CH<sub>2</sub>NHCH<sub>2</sub>), 55.35 (CONHCH<sub>2</sub>), 55.22 (CH<sub>2</sub>NHCbz), 63.77 (CH<sub>2</sub> (Cbz)), 110.75 (NHC=CN), 117.91-141.28 (CH<sub>aromatici</sub> (Cbz)), 150.43 (NHC=CN), 151.29 (N(CH<sub>3</sub>)CON(CH<sub>3</sub>)), 157.82 (C=N), 160.31 (N(CH<sub>3</sub>)CO), 162.72 (CO (Cbz)), 165.65 (CONH).

**MS** m/z: 556.32 [M+H]<sup>+</sup>

**IR** (cm<sup>-1</sup>): 3163 (N-H); 1779 (C=O).

**m.p.:** 131-133 °C

## CD measuerments

### 1- Folding of EK<sub>inv.</sub>

2.3 mg of peptide were dissolved in 5 mL of phosphate buffer (pH=7.3) in order to obtain a mother solution of peptide  $4.6 * 10^{-4} * 10^{-4}$ M. From this, 1.09 mL were diluted to 10 mL with buffer phosphate in order to obtain the final solution of work  $5 * 10^{-5}$ M.

3 mL of this solution were filled in a circular dichroism cell (light path 10 mm) and the data were recovered at 25°C in four accumulations, with data pitch of 0.05 nm, with a scanning speed of 100 nm/min and a band width of 1 mm.

The kinetic measure was carried out at fixed wavelenght of 222 nm with an identical previous peptide solution. The measure was carried on in five days at 25°C with a time measurement of 8 minutes.

### 2- Folding of K<sup>E1</sup>

8.0 mg of K<sup>E1</sup> were dissolved in 10 mL of phosphate buffer in order to obtain a mother solution  $2.053 * 10^{-4}$ M. From this, 2,43 mL were diluted to 10 mL with buffer phosphate in order to obtain a final solution of work  $5 * 10^{-5}$ M. 3 mL of this solution were filled in a circular dichroism cell (light path 10 mm) and data were recovered at 25°C in four accumulations, with data pitch of 0.05 nm, a scanning speed of 50 nm/min and a band width of 1 mm

### 3- Folding of F16-K<sup>E1</sup>

7.60 mg of F16-K<sup>E1</sup> were dissolved in 10 mL of phosphate buffer in order to obtain a mother solution  $1.93 * 10^{-4}$  M. From this, 777 µL were diluted to 3 mL in order to obtain a final solution of work  $5 * 10^{-5}$ M. 3 mL of final solution were filled in a circular dichroism cell (light path 10 mm) and data were recovered at 25°C in four accumulations with data pitch of 0.05 nm, a scanning speed of 50 nm/min and a band width of 1 mm.

### 4- Thermal denaturation of K<sup>E1</sup>

8.0 mg of K<sup>E1</sup> were dissolved in 10 mL of phosphate buffer in order to obtain a mother solution  $2.053 * 10^{-4}$ M. From this, 2,43 mL were diluted to 10 mL with buffer phosphate in order to obtain a final solution of work  $5 * 10^{-5}$ M. 3 mL of this solution were filled in a circular dichroism cell (light path 10 mm) and data were recovered between 0 to 80 °C with a solution time of stabilization at the new temperature of 20 minutes.

Spectra was recovered in four accumulation with data pitch of 0.05 nm, a scanning speed of 50 nm/min and a band width of 1 mm.

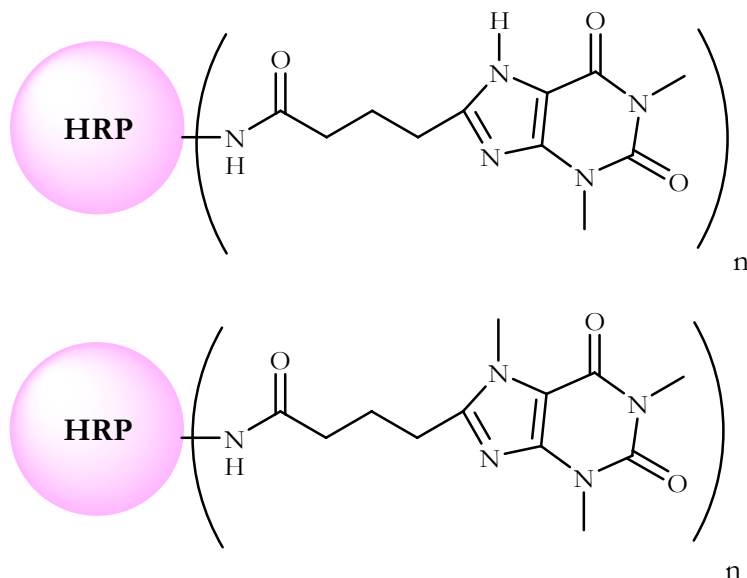
#### 5- $K^{E1}$ titolation with theophylline

73  $\mu\text{L}$  of a mother solution of  $K^{E1}$  ( $2.053 \times 10^{-4}\text{M}$ ) in phopspate buffer were diluted to 3 mL with theophylline solution in water and buffer phosphate, in order to obtain a final solution of peptide  $5 \times 10^{-6}\text{M}$  and final concentrations of theophylline solution in the range of  $1.0 \times 10^{-4}\text{ M}$  and  $1.0 \times 10^{-5}\text{ M}$ . The final solution of peptide and theophylline was kept under stirred for 10 minutes and then the solution was filled in circular dichroism cell (light path 10 mm).

The spectrum was recoreved in four accmulation, with a data pitch of 0.05 nm, a scanning speed of 100 nm/min and a band width of 1 mm.

## Theophylline/Caffeine-protein conjugation

### 1. HRP-Theophylline/Caffeine



2.1 mg of theophylline-8-butanoic acid **1** or 3.6 mg of caffeine-8-butanoic acid **4** were dissolved in 200  $\mu\text{L}$  of a 1:1 mixture of DMSO and of 100mM MES; NaCl 400mM; EDTA 10mM; pH=5.5.

2.1 or 3.6 mg for haptene **1** and **4** respectively of Horseradish peroxidase (HRP) were dissolved in 100  $\mu\text{L}$  of mQ water and the haptene solution was added slowly. 4.2 mg of EDC-Cl were added and the final solution was stirred at room temperature for 3 hours. After this time the mixtures were centrifugated at 4000 RPM for 5 minutes and the resulting supernatants were purified by dialysis against 500 mL of buffer phosphate at 0°C with regular changing of buffer every 4 hours for 8 times. Sodium azide was added to the dialyzed solution at a final concentration of 0.02%.

#### **HRP-Theo:**

200  $\mu\text{L}$  of the solution were diluted 1 times in mQ water to a final volume of 1 mL. The UV spectrum of the resulting solution was of 0.55 at 283.2 nm and of 1.22 at 401.4 nm. From these data, an apparent protein concentration of 2.27  $\mu\text{M}$  and an apparent haptene concentration of 26.2  $\mu\text{M}$  were calculated. The modification index was thus 11.

$$[\text{HRP}]_f = 2.27 * 10^{-6}\text{M}$$

$$[\text{Theophylline}]_f = 2.62 * 10^{-5}\text{M}$$

$$\text{IM} = 11$$

### HRP-Caffeine:

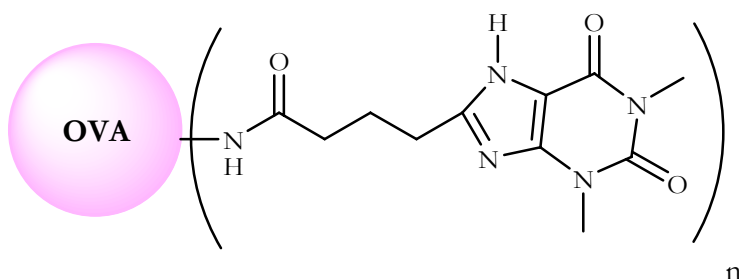
250  $\mu\text{L}$  of the solution were diluted in mQ water to a final volume of 1 mL. The UV spectrum of the resulting solution was 1.01 at 283.2 nm and of 1.23 at 401.4 nm. From these data, an apparent concentration of 6.72  $\mu\text{M}$  and an apparent hapten concentration of 228  $\mu\text{M}$  were calculated. The modification index was thus 34.

$$[\text{HRP}]_f = 6.72 * 10^{-6}\text{M}$$

$$[\text{Caffeine}]_f = 2.28 * 10^{-4}\text{M}$$

$$\text{IM} = 34$$

### 2. OVA-Theophylline



10 mg of theophylline-8-butanoic acid were dissolved in 2 mL of 1:1 mixture of DMF and 100mM MES; NaCl 900mM; 0.02%  $\text{NaN}_3$ ; pH=5.5.

10 mg of Albumin from chicken egg white (OVA) were dissolved in 1 mL of milliQ water.

Hapten solution and 10 mg of EDC-Cl were slowly added to the OVA solution. The mixture was stirred at room temperature for 2 hours. After this time the mixture was centrifugate at 4000 RPM for 5 minutes and the resulting supernatant was purified by dialysis against 500 mL of milliQ water with regular changes of water every 6 huors.

### OVA-Theophylline

300  $\mu\text{L}$  of the solution were diluted 1 times in mQ water to a final volume of 600  $\mu\text{L}$ . The absorbance in the UV spectrum of the resulting solution was of 0.82 at 275.4 nm and 0.44 at 255.8 nm. From these data an apparent protein concentration of 19.1  $\mu\text{M}$  and an apparent hapten concentration of 401 $\mu\text{M}$  were calculated. The modification index was thus 21.

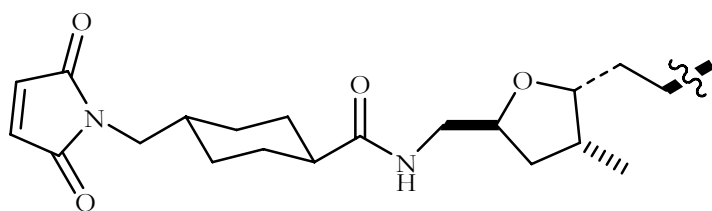
$$[\text{OVA}]_f = 1.91 * 10^{-5}$$

$$[\text{Theophylline}]_f = 4.01 * 10^{-4}$$

$$\text{IM} = 21.$$

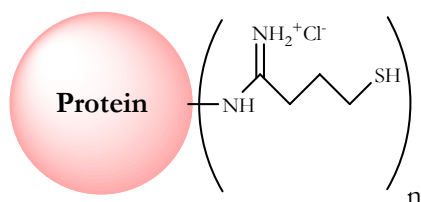
## Protein-palytoxin conjugation

### 1) Palytoxin activation



0.125 mg of 4-(N-Maleimidomethyl)-cyclohexane-carboxylic acid N-hydroxy-succinimide ester (SMCC) ( $3.73 \times 10^{-4}$  mmol) were dissolved in 0.1 mL of N,N-Dimethylformamide (DMF) (Solution **a**). In another flask, 0.2 mg of palytoxin ( $7.47 \times 10^{-5}$  mmol) were dissolved in 1.7 mL of phosphate buffer ( $\text{Na}_2\text{HPO}_4$  0.1M; pH=7.5) (Solution **b**). Solution **a** was added to solution **b** and the mixture was stirred for 8 hours at room temperature. After this the solution was directly used for the next step.

### 2) Protein-modification



#### **BSA-2-IMT:**

2 mg of BSA (Albumin from bovine serum) ( $3.03 \times 10^{-5}$  mmol) were dissolved in 0.8 mL of borate buffer and 0.21 mg of 2-IMT ( $1.52 \times 10^{-3}$  mmol) were added. The reaction was kept for 1 hour at room temperature and then used for the next step.

#### **OVA-2-IMT:**

1 mg of OVA (Albumin from chicken egg white) ( $2.27 \times 10^{-5}$  mmol) were dissolved in 0.4 mL of borate buffer and 0.22 mg of 2-IMT ( $1.14 \times 10^{-3}$  mmol) were added. The reaction was kept at room temperature for 1 hour and then used for the next step.

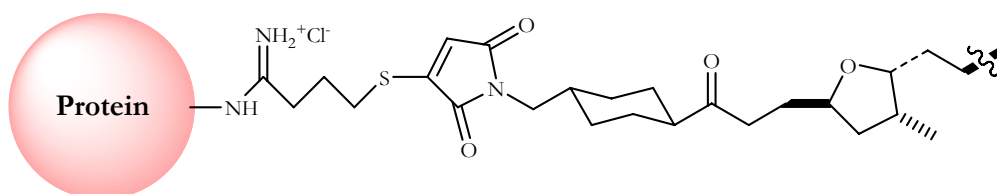
### KLH-2-IMT:

1mg of KLH (Hemocyanin from *Megathura Crenulata*) ( $1.25 * 10^{-7}$  mmol) was dissolved in 0.4 mL of borate buffer and 0.086 mg of 2-IMT ( $6.25 * 10^{-5}$  mmol) was added in the previous solution. The reaction was kept at room temperature for 1 hour.

### HRP-2-IMT:

1 mg of HRP (Horseradish peroxidase) ( $2.27 * 10^{-5}$  mmol) was dissolved in 0.4 mL of borate buffer at room temperature. 0.22 mg of 2-IMT ( $1.14 * 10^{-3}$  mmol) were added at and the reaction was kept under stirring for 1 hour.

### 3) Protein-palytoxin conjugation



#### General procedure

Activated palytoxin was added in 5 fold-molar excess to the activated protein dissolved in phosphate buffer ( $\text{Na}_2\text{HPO}_4$  0.1M; EDTA 1mM; pH=6.6). The solution was stirred for 1 hour at room temperature and then purified by dialysis against 500 mL of phosphate buffer ( $\text{NaCl}$  0.14M;  $\text{KCl}$   $2.7 * 10^{-3}$ M;  $\text{Na}_2\text{HPO}_4$  0.01M;  $\text{KH}_2\text{PO}_4$   $1.76 * 10^{-3}$ M; pH=7.3) for 3 days with regular changes of buffer every 4 hours.

UV analysis give the followin apparent protein and PTX concentrations:

#### **BSA-MCC-PTX:**

$$[\text{BSA}]_f = 9.23 * 10^{-8}\text{M}$$

$$[\text{PTX}]_f = 3.16 * 10^{-6}\text{M}$$

$$\text{IM} = 34$$

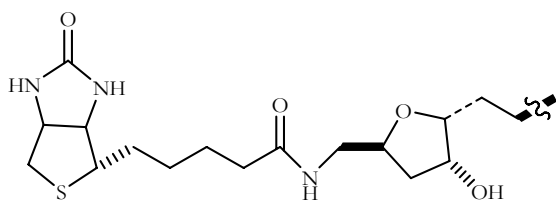
#### **KLH-MCC-PTX:**

$$[\text{KLH}]_f = 6.23 * 10^{-8}\text{M}$$

$$[\text{PTX}]_f = 1.7 * 10^{-5}\text{M}$$

$$\text{IM} = 133$$

#### 4) Biotinylation of palytoxin



4.8 mg of palytoxin ( $1.79 \times 10^{-3}$  mmol) were dissolved in 1 mL of water. 0.5 mg of biotin N-hydroxysuccinimide ester ( $1.43 \times 10^{-3}$  mmol) were added and the final solution was stirred at room temperature for 24 hours.

After this time the water solution was washed with dichloromethane (3 x 0.5 mL) and the aqueous solution was used directly.

**MS** m/z: 2929.2 [M+Na]<sup>+</sup>

## Acknowledgments

We are grateful to:

- Prof. Alessandro Tossi, Dr. Nina Ancheva of the Dept. of Life Sciences for the help given in peptide synthesis and CD spectroscopy.
- Prof. Roberto Marzari, Dr. Elisabetta Azzoni, Dr. Enrica Frare, Dr. Marta Liciulli, Dr. Paola Curto, Dr. Immacolata Luisi, Dr. Sabrina Boscolo of the Dept. Of Life Sciences, for having developed the phage libraries and for the ELISA tests.
- Prof. Aurelia Tubaro, Dr. Silvio Sosa, Dr. Marco De Bortoli, of the Dept. of Materials and Natural Resources for the palytoxin ELISA.
- Prof. Silvano Geremia and Dr. Matteo De March of the Dept. of Chemistry for the X-Ray structure of peptide KE1.
- Prof. Daniele Sblattero, Università del Piemonte Orientale “Amedeo Avogadro”, for his contribution to the design of peptide scaffolds and to the development of phage libraries.
- Prof. Ksenija Kogej, Dept. of Physical Chemistry of the University of Ljubljana, for her supervision in calorimetric measurements and for her kind hospitality.
- Dr. Cristina Busatto, Sincrotrone Trieste, for peptide crystallization trials.
- Dr. Adriano Savoini, Biosensor Technologies srl for his fundamental contribution to the project and his constant presence during any step of this work.