Università degli Studi di Napoli

"Federico II"

DOTTORATO IN SCIENZE CHIMICHE

XXII CICLO

DESIGN AND CHARACTERIZATION OF MOLECULAR SCAFFOLDS

Tutore:

Ch.mo Prof.
Carlo Pedone

Co-tutore:

Dott.
Luca D. D'Andrea

Relatore:

Ch.ma Prof.ssa
Gabriella D'Auria

Candidata:

Dott.ssa
Barbara Ziaco

INDEX

RIASSUNTO

RIASSUNTO

PROGETTAZIONE E CARATTERIZZAZIONE DI SCAFFOLD MOLECOLARI

Nell'era post-genomica lo studio delle interazioni tra biomolecole ed in particolare quelle proteina-proteina riveste sempre maggiore interesse, in quanto esse sono alla base di tutti quei processi fisiologici mediati dalla formazione di complessi fra biomolecole. Pertanto, la conoscenza dettagliata dei meccanismi molecolari responsabili di queste interazioni, è indispensabile per sviluppare molecole capaci di modulare l'attività biologica delle proteine target e quindi dei relativi processi cellulari.

Attualmente, l'identificazione di molecole capaci di inibire o favorire le interazioni proteina-proteina o proteina-acidi nucleici rappresenta una delle maggiori sfide del drug discovery. Infatti, diversamente dal tradizionale approccio basato sul design di molecole modellate su substrati specifici per siti attivi enzimatici, lo sviluppo di composti capaci di modulare le interazioni proteina-proteina è un processo più complesso.

L'interfaccia molecolare coinvolta è solitamente un'area estesa che comprende più regioni non contigue mancanti di tasche adatte al legame con piccole molecole. Inoltre, queste regioni spesso presentano elementi di struttura secondaria che una volta isolati dal loro contesto proteico non assumono la conformazione nativa.

Finora, diverse classi di composti sono state utilizzate per modulare le interazioni proteina-proteina: anticorpi, peptidi e miniproteine e raramente piccole molecole organiche [1]. Queste ultime, infatti, sebbene costituiscano la maggior parte dei principi attivi dei farmaci attualmente in commercio risultano inadatte per interagire con superfici proteiche molto ampie. Gli anticorpi presentano un elevata specificità e sono ampiamente impiegati ma hanno elevati costi di produzione. I peptidi sono considerati, invece, dei buoni candidati per sviluppare nuovi composti che interferiscono con il riconoscimento proteina-proteina [2]. I principali approcci attualmente utilizzati per sviluppare composti di natura peptidica consistono nello screening di librerie fagiche, sintesi parallele di peptidi su membrana e "rational design". Quest'ultimo richiede che siano disponibili le informazioni strutturali e biochimiche su almeno uno dei due partner interagenti e che siano stati identificati i residui coinvolti nel legame [3]. Nel caso della progettazione di peptidi che mimano l'organizzazione strutturale dei segmenti coinvolti nell'interazione è prevista una prima fase in cui si introducano residui in modo da stabilizzare la struttura secondaria e una seconda nella quale si devono inserire i residui responsabili della interazione nella giusta orientazione spaziale.

Un approccio alternativo è rappresentato dall'utilizzo di scaffold [4] molecole strutturalmente stabili che già presentano la struttura secondaria desiderata, in cui possono essere direttamente introdotti i residui nella giusta orientazione spaziale.

L'obiettivo di questo progetto di dottorato è stato quello di progettare e caratterizzare scaffold molecolari, quali peptidi e mini-proteine, in grado di modulare interazioni tra biomolecole.
Tale scopo è stato affrontato mediante tre diversi approcci:

- Lo sviluppo di una nuova strategia sintetica per l'ottenimento di polipeptidi; funzionalmente attive mediante expressed protein ligation;
- La caratterizzazione bio-fisica di uno scaffold peptidico elicoidale;
- Il trasferimento di epitopi funzionali su uno scaffold proteico noto.

La prima strategia consiste nella messa a punto di una procedura per legare mediante legami covalenti stabili l'estremità C-terminale di due frammenti polipeptidici ottenuti mediante espressione ricombinante in batteri. Tale strategia permette di ottenere sistemi peptidici modello [5] e mini-proteine che conservano la stessa funzionalità delle proteine target ma le cui dimensioni sono notevolmente ridotte. La strategia sintetica prevede la reazione di chemical ligation tra due polipeptidi attivati come tioestere al C-terminale ed un linker bifunzionale caratterizzato dalla presenza di due cisteine in posizione pseudo N-terminale. Il linker è stato sintetizzato a partire dall'etilendiammina alla quale sono state legati i due residui di cisteina, mediante un legame peptidico. Le mini-proteine tioestere al C-terminale sono state ottenute utilizzando vettori di espressione contenenti inteine [6]. La procedura sintetica è stata sviluppata utilizzando come sistema modello per le mini-proteine, la sequenza che codifica per il sito di clonaggio del vettore pTrcHisA. Il vettore è stato modificato con l'inserimento dell'inteina MxeGyrA (N198A). Il costrutto di fusione è stato espresso nelle cellule BL21(DE3) di Escherichia coli e purificato mediante cromatografia di affinità su resina di chitina; la mini-proteina tioestere è stata ottenuta in seguito allo splicing dell'inteina in presenza di tioli. In seguito la mini-proteina tioestere è stata utilizzata in due diverse reazioni di ligation con il linker, per l'ottenimento di omodimeri ed eterodimeri. I prodotti puri, caratterizzati mediante LC-MS, sono stati ottenuti tutti con buone rese [7]. Tale strategia di sintesi offre la possibilità di unire chimicamente due frammenti proteici in modo stabile attraverso l'uso di un linker bifunzionale, che può essere opportunamente modificato variandone la lunghezza e la rigidità dello spacer tra le cisteine; ciò ha un notevole potenziale per applicazioni biotecnologiche. Questa metodologia può essere utilizzata per combinare catene peptidiche vicine nello spazio ma non nella sequenza, per mimare, ad esempio, epitopi discontinui, per sintetizzare scaffold di ridotte dimensioni (mini-anticorpi) oppure in alternativa domini di dimerizzazione come ad esempio le leucin zipper.

Per il secondo approccio si è effettuata la caratterizzazione chimico-fisica di un peptide per un suo possibile utilizzo come scaffold di motivi strutturali elicoidali.

Recentemente è stato descritto un peptide, QK, capace di mimare in vitro ed in vivo l'attività biologica del VEGF [8-10] che assume in soluzione acquosa una conformazione elicoidale ben definita. L'analisi dei dati di dicroismo circolare e NMR indicano che il peptide QK presenta una stabilità termica insolita per un peptide costituito solo da aminoacidi naturali. Per valutare i determinanti strutturali di questa stabilità, i dati sperimentali sono stati integrati con simulazioni di dinamica molecolare.
Gli studi teorici hanno indicato che la regione N-terminale ed un contatto idrofobico tra la Leu7 e la Leu10 sono importanti per la stabilità termica del peptide.
Per verificare tali previsioni sono stati sintetizzati 3 peptidi analoghi di QK: QK1-12 che manca della parte C-terminale; QK4-15 di quella N-terminale e QK10A, in cui la Leu10 è stata sostituita con un'alanina. L'analisi dei peptidi, mediante dicroismo circolare e NMR, ha evidenziato che QK1-12, a differenza di QK4-15, mantiene una struttura elicoidale e stabilità termica analoga a QK, QK10A presenta circa la metà del contenuto elicoidale e non mantiene l'inusuale stabilità termica [11]. Infine, utilizzando una combinazione di tecniche sperimentali quali CD, NMR e MD è stato possibile caratterizzare a livello atomico uno dei possibili pathway di formazione dell'elica del peptide QK10A e fornire informazioni sull'inusuale stabilità termica del peptide QK requisito fondamentale per il suo impiego come scaffold elicoidale.

Il terzo approccio ha previsto lo sviluppo di una mini-proteina biologicamente attiva a partire da uno scaffold noto. Il sistema biologico scelto è stato quello del VEGF ed i suoi recettori. Dall'analisi della struttura tridimensionale del complesso tra VEGF/Flt-1D2 e da studi di mutagenesi sono stati individuati i residui del VEGF importanti per il legame ai recettori [12]; questi dati hanno dimostrato che la regione di binding del VEGF per il recettore comprende l'elica 17-25.
Lo scaffold scelto è stato l'Avian Pancreatic Polipeptide (APP), una miniproteina di 36 amminoacidi, molto stabile, che possiede un'α-elica esposta al solvente ed un'elica di poliprolina di tipo II. Sulla base delle sovrapposizioni dell'elica del VEGF e dello scaffold APP sono state progettate due diverse molecole nelle quali sono stati trasferiti i residui responsabili dell'interazione con i recettori del VEGF e del peptide QK rispettivamente denominate APP1 e APP_QK. Le molecole progettate e quella wilde type sono state ottenute per via ricombinante nelle cellule BL21(DE3) di Escherichia coli.
Le proteine sono state purificate mediante cromatografia di affinità ed analizzate mediante LC-MS. Infine, preliminari saggi biologici in vitro hanno evidenziato per la molecola APP_QK un'attività analoga a quella del peptide QK.
In conclusione, questo lavoro provvedendo con diversi approcci contribuisce significativamente allo sviluppo di nuovi scaffold per il targeting delle interazioni proteina-proteina.

BIBLIOGRAFIA

[1] Cochran A.G., Chem. Bio. 5: 654-659 (2001)

[2] Souroujon M.C. and Mochly-Rosen D., Nat. Biotechnol. 16: 919-924 (1998)

[3] Cochran A.G., Chem. Bio. 7: R85-94 (2000)

[4]Hershberger S.J., Lee S.G. and J. Chmielewski Scaffolds for blocking proteinprotein interactions. Curr. Top Med. Chem. (2007) 7(10):928-42

[5] B.R. Gibney, F. R. and P. L. Dutton, Curr. Opin. Chem. Biol. 1997, 1: 537-542

[6] T. W. Muir, Annu Rev Biochem 2003, 72, 249

[7] B. Ziaco, S. Pensato, L.D.D'Andrea, E.Benedetti and A.Romanelli, Org.Let. 2008, 10(10):1955-58

[8] L.D. D'Andrea, G. Iaccarino, R. Fattorusso, D. Sorriento, C. Carannante, D. Capasso, B. Trimarco and C. Pedone, Proc Natl Acad Sci U S A. 2005, 102, (40):14215-20

[9] Dudar GK, D'Andrea LD, Di Stasi R, Pedone C, Wallace JL. Am J Physiol Gastrointest Liver Physiol. 2008 Aug;295(2):G374-81. Epub 2008 Jun 26.

[10] G Santulli, M Ciccarelli, G Palumbo, A Campanile, G Galasso, B Ziaco, GG Altobelli, V Cimini, F Piscione, LD D'Andrea, C Pedone, B Trimarco and G Iaccarino Journal of Translational Medicine 2009, 7:41

[11]D Diana, B Ziaco, G Colombo, G Scarabelli, A Romanelli, C Pedone, R Fattorusso and LD. D'Andrea Chemistry Eur. J. 2008, 14, 4164 – 4166

[12]Wiesmann C., Fuh G., Christinger H.W., Eigenbrot C., Wells J.A., and De Vos A.M., Cell 91: 695-704 (1997)

SUMMARY

SUMMARY

DESIGN AND CARACTERIZATION OF MOLECULAR SCAFFOLD

In the post-genomic era the study of the interactions between biomolecules and in particular protein-protein interactions is of growing interest, since they are the basis of all the physiological processes mediated by the formation of complexes between biomolecules. Therefore, detailed knowledge of the molecular mechanisms responsible for these interactions is essential to develop molecules capable of modulating the biological activity of the protein target and then its cellular processes.

Currently, the identification of molecules that inhibit or promote protein-protein interactions or protein-nucleic acids is one of the greatest challenges of drug discovery. Unlike the traditional approach based on the design of molecules modeled on specific substrates for enzymatic active sites, the development of compounds able to modulate protein-protein interactions is a more complex process. The interface molecules involved is usually an extended area that includes more non-contiguous regions lacking suitable pockets of binding to small molecules. Moreover, these regions often have elements of secondary structure that once isolated from their context, not having the protein native conformation. So far, several classes of compounds have been used to modulate protein-protein interactions: antibodies, peptides and small organic molecules and rarely miniproteine [1]. The latter, in fact, while constituting the majority of the active ingredients currently on the market are unsuitable to work with very large surface protein. The antibodies show high specificity and are widely used but have high production costs. The peptides are considered, however, good candidates for developing new compounds that interfere with protein-protein recognition [2]. The main approaches currently used to develop compounds of peptidic nature consists in the screening of phage libraries, parallel synthesis of peptides on membrane and rational design.

Latter requires that the structural and biochemical information available on at least one of the two interacting partners and have been identified residues involved in the bond [3]. If the design of peptides that mimic the structural organization of the segments involved in the interaction is expected to introduce first stage of waste in order to stabilize the secondary structure and a second in which you must enter the residues responsible for interaction in the right spatial orientation.

An alternative approach is to use scaffold [4] molecules structurally stable that already have the desired secondary structure, which can be directly introduced into the residues in the correct spatial orientation.

The aim of this PhD project was to design and characterize molecular scaffold, such as peptides and mini-proteins that can modulate the interactions between biomolecules.

This aim was addressed by three different approaches:

• The development of a new synthetic strategy for obtaining polypeptides, functionally active protein expressed by ligation;
• The bio-physical characterization of a helical peptide scaffold;
• The transfer of functional epitopes on a scaffold protein known.

The first strategy consists in developing a procedure for binding by stable covalent bonds the C-terminus of two polypeptide fragments obtained by recombinant expression in bacteria. This strategy allows to obtain peptide model systems [5] and mini-proteins that retain the same functionality of the target protein but whose dimensions are considerably reduced. The synthetic strategy provides for the chemical ligation reaction between two polypeptides activated as the C-terminal thioester and a bifunctional linker is characterized by the presence of two cysteines in position pseudo N-terminal. The linker was synthesized from dall'etilendiammina which they were linked to two cysteine residues via a peptide bond. The mini-protein to the C-terminal thioester was obtained using expression vectors containing inteine [6]. The synthetic procedure was developed using as a model system for the mini-protein, the sequence coding for the cloning site of the vector pTrcHisA. The vector was modified by the insertion dell'inteina MxeGyrA (N198A). The fusion construct was expressed in cells BL21 (DE3) of Escherichia coli and purified by affinity chromatography on chitin resin, the mini-protein thioester was obtained following the dell'inteina splicing in the presence of thiols. Following the mini-protein thioester was used in two separate ligation reactions with the linker, to obtain homodimers and heterodimers. The pure products, characterized by LC-MS, were all obtained with good yields [7]. This synthetic strategy offers the opportunity to unite chemically two protein fragments in a stable manner through the use of a bifunctional linker, which can be suitably modified by varying the length and rigidity of the spacer between the cysteine residues and this has considerable potential for biotechnological applications . This methodology can be used to combine neighboring peptide chains in space but not in sequence, to mimic, for example, discontinuous epitopes, to synthesize scaffolds of small (mini-antibody) as an alternative dimerization domains such as Leucin zipper.

For the second approach has made the chemical and physical characterization of a peptide for its possible use as scaffolds for helical structural reasons.
It was recently described a peptide, QK, able to mimic in vitro and in vivo biological activity of VEGF [8-10] in aqueous solution which assumes a well-defined helical conformation. Analysis of circular dichroism and NMR data indicate that the peptide QK has a thermal stability is unusual for a peptide composed only of natural amino acids. To assess the structural determinants of this stability, the experimental data have been supplemented with molecular dynamics simulations.

Theoretical studies have indicated that the N-terminal region and a hydrophobic contact between the Leu7 and Leu10 are important for the thermal stability of the peptide.
To test these predictions have been synthesized 3 peptides similar QK: QK1-12 that lacks the C-terminal; QK4-15 than the N-terminal and QK10A, in which Leu10 was replaced with un'alanina. The analysis of peptides by circular dichroism and NMR showed that QK1-12, unlike QK4-15, maintains a helical structure and thermal stability similar to QK, QK10A has about half the helical content and does not retain 's unusual thermal stability [11]. Finally, using a combination of experimental techniques such as CD, NMR and MD was possible to characterize at atomic one possible pathway for formation of the peptide helix QK10A and provide information sull'inusuale thermal stability of the peptide QK prerequisite for its use as helical scaffold.

The third approach has included the development of a mini-biologically active protein from a scaffold known. The biological system chosen was that of VEGF and its receptors. An analysis of three-dimensional structure of the complex between VEGF/Flt-1D2 and mutagenesis have identified residues important for binding to VEGF receptors [12], these data demonstrated that the binding region of VEGF receptor includes the helix 17-25.
The scaffold was the chosen Avian Pancreatic Polypeptide (APP), a miniproteina of 36 amino acids, very stable, which owns un'α-helix exposed to solvent and thruster poliprolina type II. Based on the overlap of the propeller of VEGF and the APP scaffold have been designed in which two different molecules have been transferred to the residues responsible for interaction with VEGF receptors and the peptide QK respectively named APP1 and APP_QK. Molecules are designed and the wild type were obtained by recombinant cells BL21 (DE3) of Escherichia coli.
The proteins were purified by affinity chromatography and analyzed by LC-MS. Finally, preliminary in vitro biological assays have shown for the molecule APP_QK a similar activity to that of QK peptide.
In conclusion, this work by providing different approaches contribute significantly to the development of new scaffolds for targeting protein-protein interactions.

BIBLIOGRAFY

[1] Cochran A.G., Chem. Bio. 5: 654-659 (2001)

[2] Souroujon M.C. and Mochly-Rosen D., Nat. Biotechnol. 16: 919-924 (1998)

[3] Cochran A.G., Chem. Bio. 7: R85-94 (2000)

[4]Hershberger S.J., Lee S.G. and J. Chmielewski Scaffolds for blocking proteinprotein interactions. Curr. Top Med. Chem. (2007) 7(10):928-42

[5] B.R. Gibney, F. R. and P. L. Dutton, Curr. Opin. Chem. Biol. 1997, 1: 537-542

[6] T. W. Muir, Annu Rev Biochem 2003, 72, 249

[7] B. Ziaco, S. Pensato, L.D.D'Andrea, E.Benedetti and A.Romanelli, Org.Let. 2008, 10(10):1955-58

[8] L.D. D'Andrea, G. Iaccarino, R. Fattorusso, D. Sorriento, C. Carannante, D. Capasso, B. Trimarco and C. Pedone, Proc Natl Acad Sci U S A. 2005, 102, (40):14215-20

[9] Dudar GK, D'Andrea LD, Di Stasi R, Pedone C, Wallace JL. Am J Physiol Gastrointest Liver Physiol. 2008 Aug;295(2):G374-81. Epub 2008 Jun 26.

[10] G Santulli, M Ciccarelli, G Palumbo, A Campanile, G Galasso, B Ziaco, GG Altobelli, V Cimini, F Piscione, LD D'Andrea, C Pedone, B Trimarco and G Iaccarino Journal of Translational Medicine 2009, 7:41

[11]D Diana, B Ziaco, G Colombo, G Scarabelli, A Romanelli, C Pedone, R Fattorusso and LD. D'Andrea Chemistry Eur. J. 2008, 14, 4164 – 4166

[12]Wiesmann C., Fuh G., Christinger H.W., Eigenbrot C., Wells J.A., and De Vos A.M., Cell 91: 695-704 (1997)

I CONTEST and PROJECT AIMS

CONTEST

The study of proteins impacts directly on human health. Indeed, enzymes, receptors, and key regulator proteins have been targeted for decades for drug discovery. However, a recent and exciting development is the exploitation of a better knowledge of interactions between biomolecules. Protein-protein and protein-nucleic acid interactions are central to many processes in molecular biology. Through such interactions, translation is initiated or terminated, apoptotic signals are stimulated or inhibited, transcription is activated or repressed, and a whole host of other cellular processes relying on recognition, regulation, and signaling are performed. For this reason, understanding protein-protein and protein-nucleic acid interactions is critical for gaining insight into signaling and regulation within biological systems. Knowledge of these interactions might enable the development of new molecules for therapies that could target a multitude of human diseases, thus highlighting the practical importance of understanding and modulate such interactions.

Developing modulators of protein–protein interactions is a far more complicated process that involves a numerous factors. First, the interfacial surface area necessary for specific recognition is typically large (approximately 750–1500 $Å^2$), suggesting that large ligands may be required to compete effectively with the natural protein partner, as opposed to '*druglike*' small molecules that have been successfully utilized in enzyme inhibition. Secondly, interaction surfaces are often shallow and relatively featureless, rather than the well-defined binding pockets present in enzyme active sites, making the design of selective binders difficult. Thirdly, the binding regions of protein–protein interactions are often non-contiguous, so that mimicry of these domains is not possible by simple synthetic peptides or peptidomimetics. In addition, the adaptively of the protein surfaces involved in protein–protein interactions suggests that there may be binding conformations suitable for small molecules that are invisible in a single crystal structure, there by the design of suitable and effective binders can be a very difficult task.

Finally, unlike enzyme activity that may simply be monitored by commercially available assays, novel and efficient screening assays must be developed for the interaction between proteins [Sillerud L.O. and R.S. Larson, 2005; González-Ruiz D. and H. Gohlke, 2006; Moreira I.S. et al., 2007]. Even if few notably examples of small molecule targeting of protein-protein interactions have been reported, peptides still represent the molecule of choice to start to develop new effective protein binders.

Several approaches are currently used to develop this type of compounds, they consist in the screening of phage libraries, parallel synthesis of peptides on membrane and rational design.

Moreover, the identification of peptides based on the amino acid sequences found at protein-protein interfaces can be an excellent and straight forward way to obtain new

leads for antagonist development. Ather convinient strategies are based on the design of molecular scaffolds, protein surface mimetics, alpha-helical mimetics, beta-sheet or beta-strand mimetics, as well as beta-turn mimetics, have been successfully utilized to modulate protein-protein interactions involved in a number of diseases, including cancer and HIV [Che Y. et al., 2006; Hershberger S.J. et al., 2007].

We focused on the design synthesis and characterization of molecular stable scaffolds to address protein-protein interactions conbining three different approaches:

- develop of a synthetic protocol to obtain polypeptide connected through their C-termini;
- biophysical characterization of a short helix peptide as molecular scaffold for targeting of protein-protein interactions mediated by α-helix;
- design and synthesis of scaffold mini-proteins to modulate angiogenesis targeting VEGF-receptors.

Firstly, the intention was to develop the conditions of synthesis with higher reaction yields and greater biocompatibility, to assemble the fragments peptide in order to apply them in the development of molecular scaffolds as mini-antibodies and similar leucine-zipper. Then we have identified the structural determinants responsible of the unusual stability of the QK helical peptide (L.D. D'Andrea et al., 2005, Dudar GK, et al., 2008), for use as a scaffold for protein-protein interactions mediated by alpha helices. Finally, we described the development of a molecular scaffold to modulate the interaction between VEGF and its receptors.

Che Y., Brooks B.R. and G.R. Marshall *J. Comput. Aided Mol. Des.* (2006) 20(2):109-30

D'Andrea L.D., G. Iaccarino, R. Fattorusso, D. Sorriento, C. Carannante, D. Capasso, B. Trimarco and C. Pedone, Proc Natl Acad Sci U S A. 2005, 102, (40):14215-20

Dudar GK, D'Andrea LD, Di Stasi R, Pedone C, Wallace JL. Am J Physiol Gastrointest Liver Physiol. 2008 Aug;295(2):G374-81

González-Ruiz D. and H. Gohlke *Curr. Med. Chem.* (2006)13(22):2607-25.

Hershberger S.J., Lee S.G. and J. Chmielewski *Curr. Top Med. Chem.* (2007) 7(10):928-42

Moreira I.S., Fernandes P.A. and M.J. Ramos *Proteins* (2007) 1;68(4):803-12

Sillerud L.O. and R.S. Larson *Curr. Protein Pept. Sci.* (2005) 6(2):151-69

II SYNTHETIC STRATEGY FOR POLYPEPTIDES ASSEMBLY

SYNTHETIC STRATEGY FOR POLYPEPTIDES ASSEMBLY

II.1 Introduction

Targeting the regions of interaction between proteins is often difficult because these are very broad, spanning more binding domain, and lack of pockets suitable for binding small molecules. Moreover, protein binding regions, very often, are composed of peptide segments with well-defined secondary structures that are distant in sequence but close in space. In recent years there have been significant efforts to obtain minimized versions of naturally occurring proteins, such as dimeric DNA binding proteins, which include the different structural motif for binding and retain their function.

Polypeptides fragments have been assembled both by covalent and non covalent interactions. In the last years chemoselective reactions, such as native chemical ligation, Staudinger ligation, Cu(I)-catalyzed azide–alkyne cycloaddition, and imine ligation (hydrazone and oxime) have been widely employed as way to covalently connect two polypeptides.

The most representative example of reaction yielding a native peptide bond is that between an N-terminal cysteine and a C-terminal α-thioester gives a native peptide bond at the site of ligation (figure II.1) (Dawson PE et al., 1994). In a freely reversible first step, a transthioesterification occurs to yield a thioester-linked intermediate; this intermediate rearranges irreversibly under the usual reaction conditions to form a peptide bond at the ligation site.

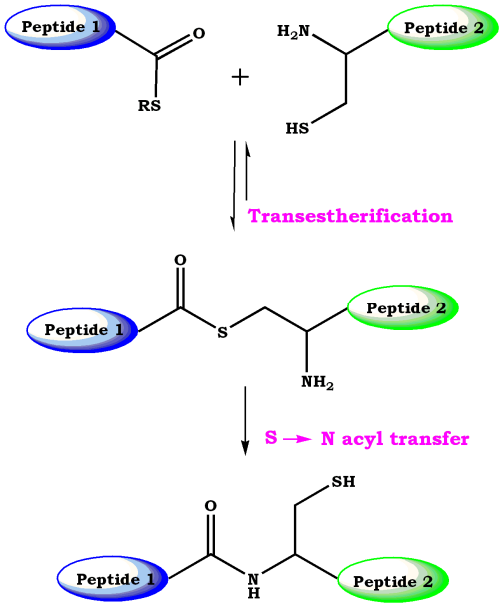

Figure II.1: Schematic meccanism of the native chemical ligation (NCL) reaction.

The most important limitation of native chemical ligation is the fact that a cysteine residue must participate in the reaction as nucleophilic reaction partner. To remove this restriction, auxiliary groups have been developed that imitate the cysteine structure in that they provide a cleavable thiol unit. These auxiliaries are joined synthetically to the N-terminus of a peptide. A second peptide is activated at its C-terminus in a manner that allows capture of the auxiliary. After the capture step, a Y→N acyl transfer reaction produces an amide. The final step is removal of the extraneous atoms of the auxiliary, which often involves an additional synthetic step. The first auxiliaries used for peptide ligation at noncysteine residues were *N*α- ethanethiol and *N*α-oxyethanethiol (Canne L.E. et al., 1996). The use of β-mercaptoamino acids, in place of a cysteine residue, and their subsequent desulfurization offers an alternative to the use of auxiliaries. Recently this strategy was extendely using a β-mercaptophenylalanine building block which allowed the insertion of a phenylalanine at the site of ligation (D. Crich and A. Banerjee, 2007; P. Botti and S. Tchertchian, 2006). Until recently, the necessary use of large amounts of desulfurization reagent was a critical disadvantage of this strategy. A new metal-free method was developed by Danishefky shedding new light onto the coupling–desulfurization strategy (B.Wu et al., 2006).

A powerful extension of native chemical ligation is the expressed protein ligation (EPL) which combines protein expression with peptide synthesis (T.W. Muir Annu. Rev. Biochem. 2003. 72:249–289) and allow the synthesis of modified proteins with unnatural amino acids, biophysical probes, and stable isotopes into specific locations.

Recently, Bertozzi (Saxon, 2000) and Raines (Nilsson, 2000) independently developed modified Staudinger reactions in which a native peptide bond was formed between a thioester and an azide-containing peptide, in the presence of phosphinobenzenethiol or other suitable thiols (Figure III.2). These strategies, based on the Staudinger reaction, although similar to the native chemical ligation reaction described earlier (Dawson, 1994) are characterized by the fact that no cysteine residue is required at the ligating site.

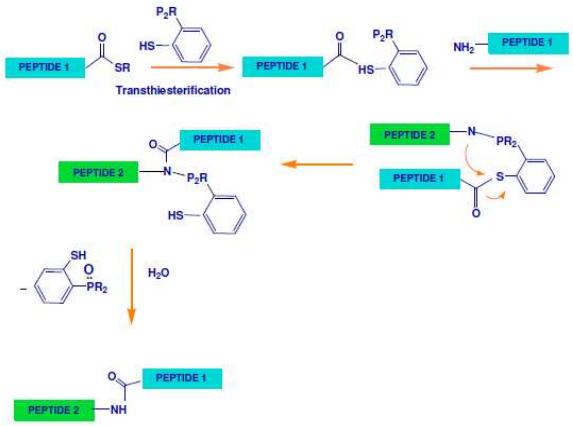

Figure III.2: Modified Staudinger ligation developed by Nilsson et al., 2000.

In the case of reactions yielding a non-native linkage, the imine-ligations and azide–alkyne cycloaddition reactions are widely used. The imine-ligation reactions involve a carbonyl group and a hydrazine or an aminooxy group. They are highly chemoselectives and both reacting groups can be straight forwardly introduced in peptides and proteins (Cordes EH and Jencks WP., 1962; Dirksen A et al., 2006). However, slow reaction rates limited the scope of this ligation method. Instead, the Cu(I)-catalyzed azide–alkyne cycloaddition (click chemistry) is more applied in the labeling of biological macromolecules rather that in peptide ligation (Kolb HC et al., 2001).
The chemoselectivity and improved stability of the Cu(I) catalyst in the presence of ligands suggest a promising future of the azide–alkyne cycloaddition in peptide ligation (Punna S. et al., 2005). Finally, a promising novel approach for amide formation utilizes the selective reaction between hydroxylamines and ketoacids.
We intend to develop a new synthetic strategy to covalently connect polypeptides through their C termini. This strategy has potential biochemical and bioorganic applications, for obtaining minimized and/or modified natural proteins and for the assembly of polypeptides on a same scaffold. The strategy is based on the use of an opportune linker and EPL reactions.

II.1.1 Inteins and Expressed Protein Ligation

Inteins, or “protein introns”, are a segment of protein that catalyze a posttranslational processing event that involves the precise removal of an internal polypeptide segment, from a precursor protein with the concomitant ligation of the flanking polypeptide sequences, termed exteins (Perler, F.B et al., 1994). Since its discovery in 1990, more than 200 inteins have been identified in all three domains of life (Perler, F.B., 2002).
The inteins, ranging from 128 to 1650 amino acids, show a set of highly conserved sequence motifs. The most part of inteins contains the characteristic motifs of a homing endonuclease that confers genetic mobility upon the intein-encoding gene and splits the region required for splicing. An other part of inteins are known as mini-inteins because they lack an endonuclease-coding region. Of special interest are the naturally occurring transsplicing inteins in which a host gene is split into two separate coding regions, each fused to either the N-terminal or C-terminal portion of an intein-coding region (Wu, H. et al., 1998). Formation of the full-length host protein occurs when the N-terminal and C-terminal intein regions come together to reconstitute protein splicing activity. Many inteins have been shown to self-splice in vitro without the requirement of external energy or protein cofactors (Paulus, H., 2000). The mechanism of protein splicing has been elucidated by the identification of key catalytic amino acid residues and intermediates. Most inteins start with a cysteine or serine residue that is responsible for an acyl shift at the N-terminal splice junction.

A study of the oceanic nitrogen-fixing cyanobacterium Trichodesmium erythraeum has showed the presence of three inteins, including one split intein, in the dnaE gene encoding the catalytic domain of DNA polymerase III (Liu, X.Q. and Yang, J., 2003). The study of T. erythraeum has also led to the first report of the coexistence of multiple inteins and introns in a single gene (Yang, J. et al., 2004; Liu, X.Q. et al., 2003).

A new example of a viral intein was recently found in Mimivirus (Ogata, H. et al., 2005).

Most intein properties facilitates their use in diverse protein engineering strategies in diverse protein engineering strategies. The steps that underlie protein splicing consist of two acyl rearrangements, a transesterification and cyclization of an asparagine. The elucidation of the proteinsplicing pathway led to the discovery that catalysis of each of the steps is often relatively independent. Thioester formation, for example, which occurs after the initial acyl rearrangement, can take place even when the the intein C-terminal asparagine is replaced with alanine and the subsequent splicing steps are blocked (Xu, M.Q and Perler, F.B., 1996; Chong, S.et al.., 1996). Thioester formation is the basis for an intein-mediated purification system in which a target protein is fused to the N terminus of an intein and can be released in a thiol-induced reaction. This intein fusion system has been extended to produce recombinant proteins possessing a C-terminal thioester for ligation with synthetic peptides or recombinant proteins carrying a variety of modifications or chemical moieties (Muir, T.W. et al., 1998; Evans, T.C. et al., 1998).

EPL (Expressed Protein Ligation) is a protein engineering approach in which a recombinant protein thioester, generated by thiolysis of an intein fusion protein, is reacted with a synthetic or recombinant peptide with an N-terminal cysteine to produce a native peptide bond (Figure II.3).

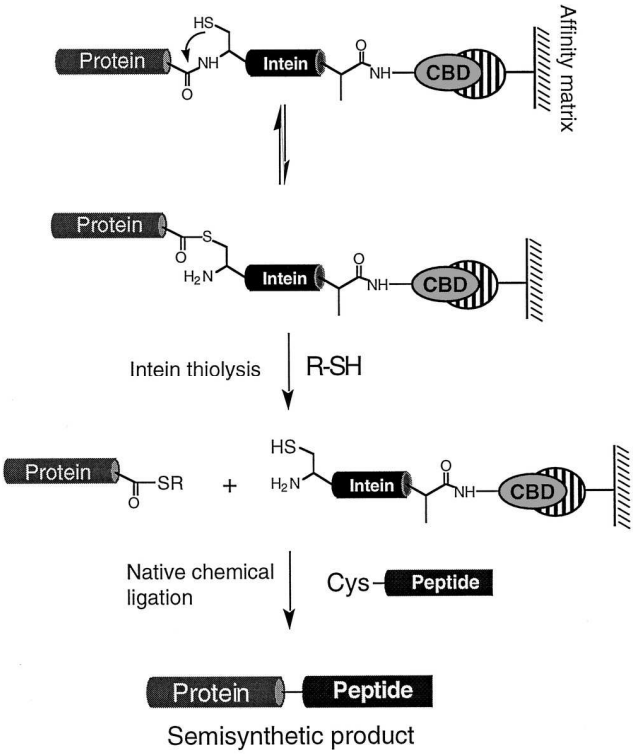

Figure II.3: Recombinant protein α-thioesters and expressed protein ligation.

Several groups demonstrated that inteins can be cut into two pieces that individually have no activity but that when combined will associate non-covalently to give a functional protein (Mills, K.IV. et al., 1998; Wu, H. et al., 1998; Yamazaki, T. et al., 1998). Affinity tags can be fused to the split site for purification of precursor fragments (Southworth, M.W. et al., 1998) and the reconstituted split intein then mediates a normal protein splicing reaction.

II.2 Results

In order to assembly two polypeptides through their C-termini by means of a covalent bond The synthetic strategy we developed consists on the reaction between a bifuncional linker and thioester polypeptides. The bifunctional linker contains two cysteines in a N-terminal- like position, separated by an ethylendiamine spacer; the thioester polypeptides are obtained by intein mediated splicing reactions. Thioesters were generated after splicing of proteins containing the MxeGyrA intein at the C-terminus of the amino acid sequences codified by the multicloning site of vectors such as a modified pTrcHisA (protein A) and pTXB1 (protein B), containing as purification tag a hexa-histidine sequence and a chitin binding domain respectively. In these constructs the MxeGyrA intein contains a single mutation, Asn-198-Ala, which prevents cleavage of the intein–C-extein peptide bond, without affecting the intein N-terminal splicing reaction. After induction of splicing in the presence of thiols, these constructs generate respectively a 36 and a 16 amino acid polypeptides (respectively protein A and protein B) as thioesters as described in materials and methods.

II.2.1 Synthesis of the linker

The linker was obtained reacting ethylendiamine with a two-fold excess of Boc-Cys(Trt)-OH. After acidic deprotection of the Boc and Trt groups and purification, the desired linker was obtained (Figure II.4); its identity was confirmed by ESI-MS and ^{1}H NMR spectrum(Figure II.5).

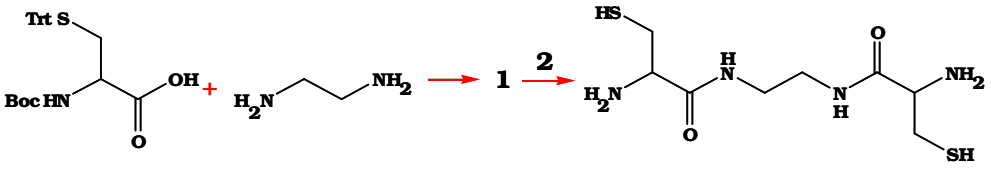

Figure II.4: Reagents and conditions
1: Boc-Cys-(Trt)OH (2eq), $NH_2CH_2CH_2NH_2$ (1eq), PyBOP/HOBT (2eq), DIPEA (6eq), DCM for 2 hours at room temperature;
2: TFA/DCM/TIS (50/49/1) for 5 minutes at room temperature.

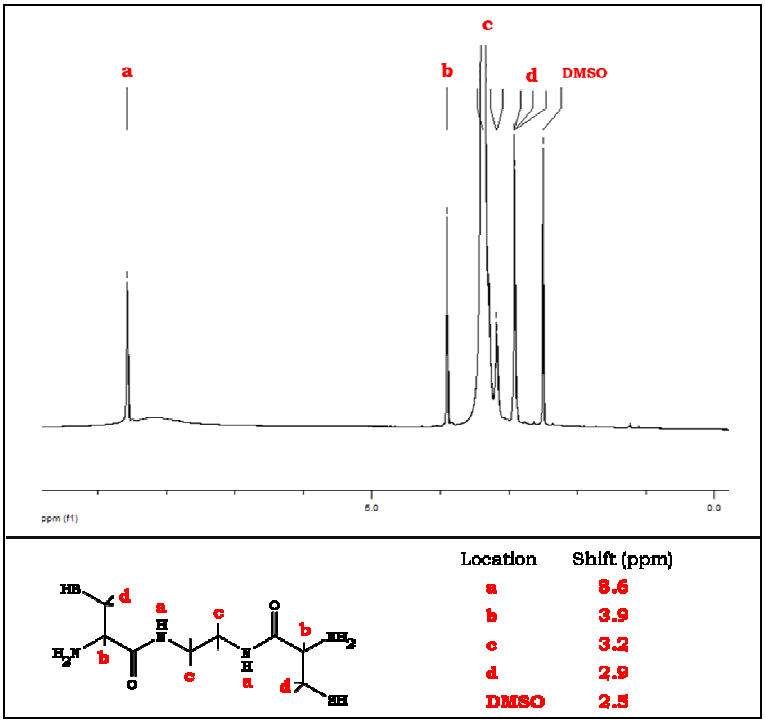

Figure II.5: 1D Proton spectrum (DMSO d_6) of linker and chemical shift and assignment of signals

II.2.2 Synthesis of dimers

The linker synthetized is suitable for two chemical ligation reactions. Proteins were expressed in Escherichia Coli BL21 (DE3) cells, transformed with the appropriate plasmid. The homodimer (A-linker-A) was obtained using the construct protein A-MxeGyrA. The protein was purified by affinity chromatography using a Ni^{2+} NTA resin. The splicing and ligation reactions occurred simultaneously in phosphate buffer 20 mM, 0.18 mM MESNA, 0.23 mM EDT, pH 7 with 0.5 equivalents of linker (Figure II.6).

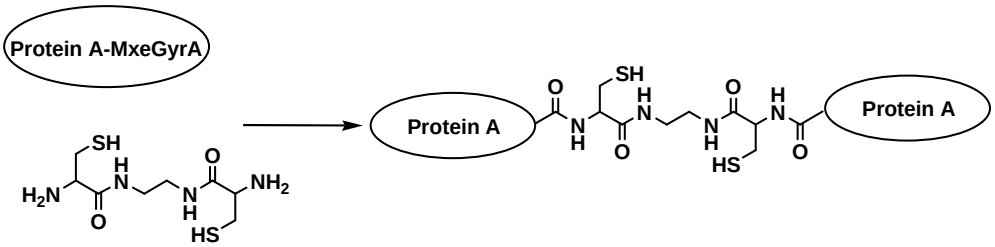

Figure II.6: Reagents and conditions
EDT 2%, MESNA 3% at room temperature overnight.

The homodimer was obtained after an overnight incubation with a 50% yield. The yield of the reaction is strictly dependent on the concentration: when the protein concentration is 50 µM or lower, only a single protein A unit connected at the C terminus with the linker (A-linker), was obtained. At a concentration of 130µM (or higher) the yield of dimer

increased up to 50%. When a threefold excess of linker is used at a protein concentration higher than 130 µM the yield of the homodimer is always 50%; the unreacted protein is converted into the A-linker derivative (Figure II.7).

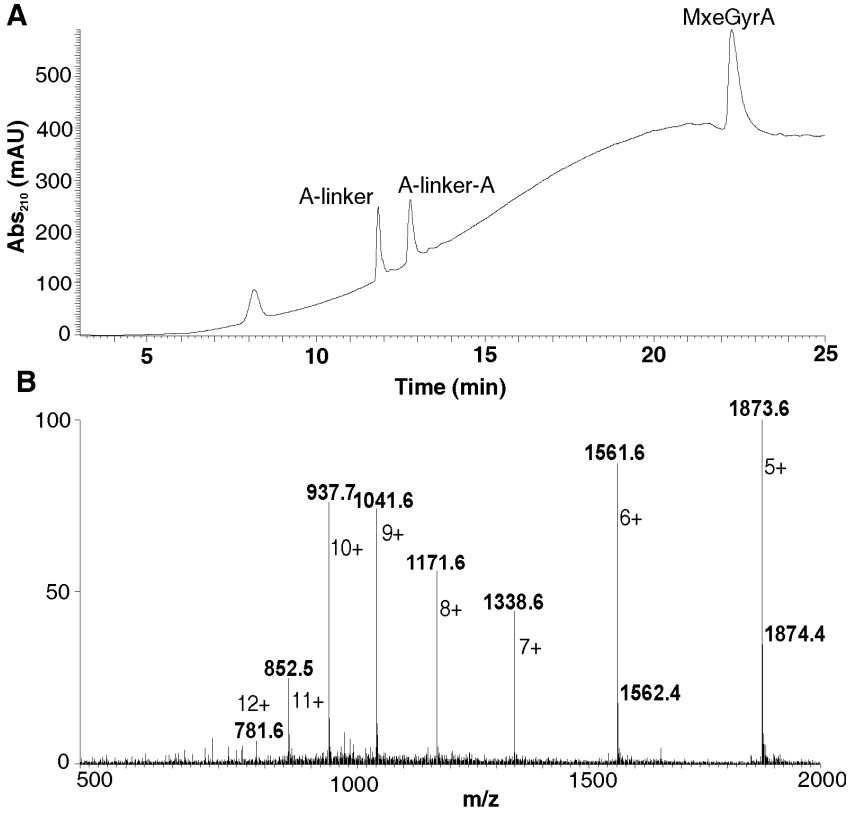

Figure II.7: Homodimer synthesis.
A) RP-HPLC profile of the homodimer reaction formation, peak identity are reported.
B) ESI-MS spectrum of A-linker-A peak.

Reaction did not occur without EDT, probably because the linker easily get oxidized. For the synthesis of the heterodimer (A-linker-B) we used the protein A derivatized with the linker (A-linker) and reacted it with thioester protein B (Figure II.8).

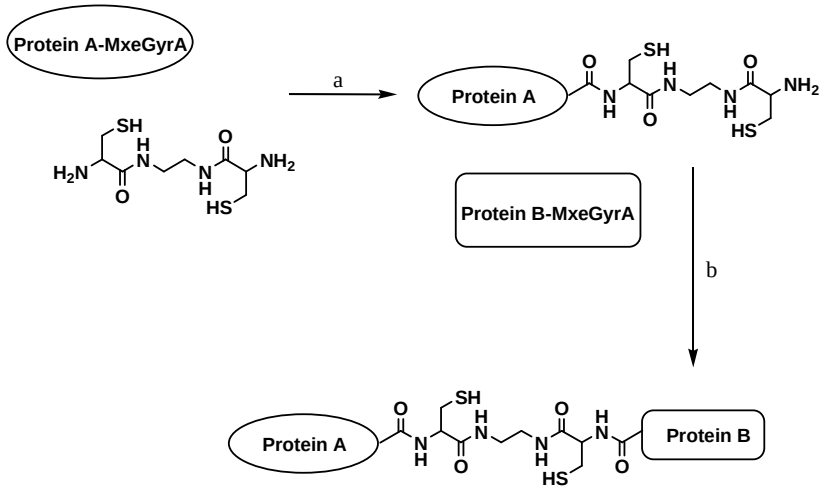

Figure II.8: Reagents and conditions
a) EDT 2%, MESNA 3% at room temperature overnight; **b**) EDT 2%, MESNA 3% at 4°C overnight.

The A-linker derivative was obtained after the splicing of the protein A-MxeGyrA fusion protein (20 μM) with a 3 fold excess of linker. The excess of linker was separated by dyalisis and the protein was identified by LCMS, purified by HPLC and lyophilized.
The A-linker derivative co-elutes with a little amount of Protein A (C terminal carboxyl).
The protein B-MxeGyrA fusion protein was immobilized on the chitin resin, splicing was induced incubating the resin in 20 mM phosphate buffer containing 500 mM NaCl, 50 mM MESNA and 1 mM EDTA, pH7.
The thioester was identified by LCMS, purified by HPLC and lyophilized. The purified A-linker (0.4 mM) derivative was reacted with two equivalents of thioester protein B (0.4 mM). Reaction was carried out in 20 mM phosphate buffer pH7, in the presence of MESNA and EDT overnight. A-linker was all converted into the heterodimer A-linker-B, while the excess protein B was converted into the EDT-thioester (Figure II.9).

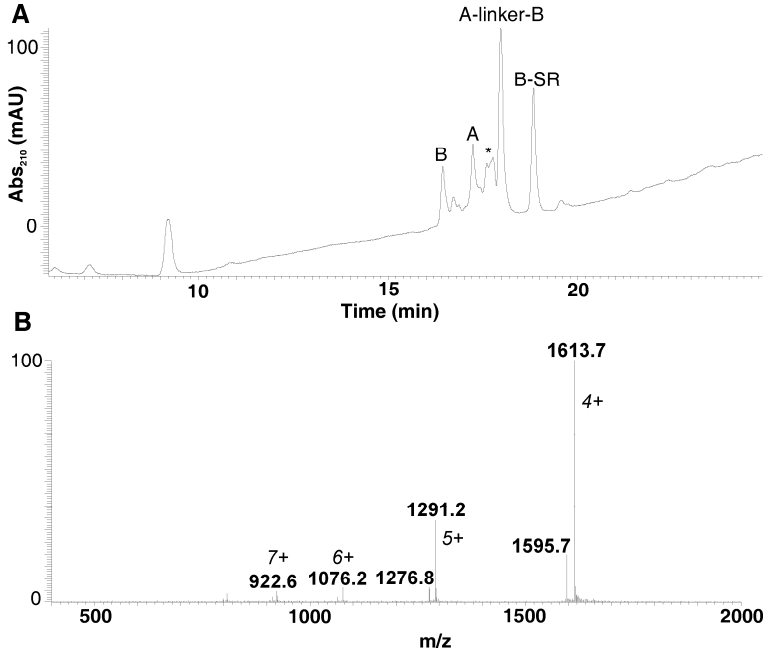

Figure II.9: Heterodimer synthesis
A) RP-HPLC profile of the heterodimer reaction formation, peak identity are reported;
* indicates products by heterodimer oxidation;
B) ESI-MS spectrum of A-linker-B peak.

The protein B is very prone to degradation; in the mass spectra of the protein B and the heterodimer was finded oxidation products (M+16) and products without the first two amino acids (M-202).

II.3 Discussions

In recent years there have been significant efforts to obtain minimized versions of naturally occurring proteins such as dimeric DNA binding proteins which retain their biological function (Lajmi, A. R. et al., 2000; Cuenoud, B. and Schepartz, A., 1993; Woolley, G. A. et al., 2006) and to assembly peptides and proteins by covalent and noncovalent interactions. (Sato, S. et al., 2002; Morii, T. et al., 1997; Canne, L. E. et al., 1995). For example, host-guest inclusion complexes of β-cyclodextrin and adamantane were used for joining two monomers by non covalent interactions (Ueno, M. et al., 1993) and scaffolds were employed for the synthesis of dendrimeric peptides/proteins (Tam, J. P. and Yu, Q., 2002; Van Baal, I. et al., 2005; Tam, J. P. et al., 2002).

In this work we propose a one pot strategy for obtaining protein homodimers covalently connected at the C-terminus by Expressed Protein Ligation (EPL). The synthetic strategy was also extended to the synthesis of heterodimers.

EPL is a protein engineering tool for the chemo and regio-selective modification of proteins based on the use of intein containing constructs (Dawson, P.E et al., 1994; Muir, T.W., 2003; Xu, M.Q. and Evans, T.C.Jr, 2005). Engineered inteins are used today for the production of C-terminal thioester and N-terminal cysteine proteins (Xu, M.Q. and Perler, F.B., 1996; Chong, S.et al., 1996). Reactive thioesters are used in ligation reactions with N-terminal cysteine peptides to give a new protein with all native peptide bonds. The ligation reaction proceeds through a trans thioesterification initiated by the cysteine thiol, followed by a spontaneous N-S acyl shift which affords the peptide bond between the two reacting moieties. EPL allows for the preparation of proteins containing natural and artificial modifications (Hahn, M.E. et al., 2007; Chatterjee, C. et al., 2007; Schwarzer, D. and Cole, P.A., 2005).

The new developed protocols are based on the reactions between thioester polypeptides, obtained via EPL, and a bis-cisteinyl linker. The main advantages of this semisynthetic strategy are its efficency and versatility. Indeed, the reaction spontaneously proceed and the obtained products just depend on the peptide sequences and on the linker selected. Besides, the linker can be easily modified as needed by changing the length and rigidity of the spacer between the cysteines. Another advantage of this strategy, with respect to the other chemoselective approaches, is the absence of size restiction on the polypeptides dimension as the two fragments are expressed and do not need to be synthesized. Moreover, the use of EPL allows easily and with low cost the preparation of labeled derivatives, for example for NMR characterization.

In conclusion, a strategy for the assembly of polypeptides covalently bound through a linker at the C-terminus was developed. This strategy based on EPL reactions between thioester peptides and a new bifunctional linker affords chemically stable fused proteins in a one-pot reaction. The linker can be easily modified as needed by changing the length

and rigidity of the spacer between the cysteines. This strategy has potential in biochemical and bioorganic applications, for obtaining minimized and/or modified natural proteins and for joining two different proteins at the C-terminus.

II.4 Experimental Section

II.4.1 Materials and Methods

Protected N-Fmoc-amino acid derivatives, acetic anhydride, ethylenediamine, coupling reagents and Rink amide MBHA resin have been purchased from Novabiochem.
DIPEA is provided from Applied Biosystem. All other solvents are commercially available from LabScan. Column chromatography was performed on Fluka silica gel 60 (size: 0.04-0.063mm). Reagents used for preparation of buffers and growth media of *Escherichia coli* were supplied by Sigma Aldrich, the reagents for polyacrylamide gels electrophoresis (Acrylamide, APS, TEMED, SDS, Tris, Glycine) by Applichem.
Isopropylbeta- D-thiogalactopyranoside (IPTG) is from Inalco. The pTXB1 vector, the chitin resin and the modification enzymes were purchased at New England Biolabs; pTrcHisA and Ni-NTA resin were from Invitrogen. *Escherichia coli* BL21(DE3) cells used for expression were supplied by Invitrogen.
Preparative purification was carried out on a Shimadzu LC-8A, equipped with a SPD-M10 AV diode array detector. Preparative HPLC was performed on a Phenomex Jupiter 10µ C12 Proteo 90Å 250 x 10.00 mm column.
LCMS analyses have been performed on an LC-MS system (Thermo Electron) comprising an LCQ Deca XP MAX ion trap mass spectrometer equipped with an ESI source and a complete Surveyor HPLC system (including MS pump, autosampler and photo diode array [PDA]). The column used was the Phenomenex Jupiter 5µ C4 300Å, 250 x 2.00mm.
LCMS analyses of the linker have been performed on a LC-MS system (Thermo Electron) comprising an Surveyor MSQ single quadrupole mass spectrometer equipped with a ESI source and a complete Surveyor HPLC system (including MS pump, autosampler and photo diode array [PDA]). The column used was the Phenomenex Jupiter 5µ C18 300Å, 150 x 4.6 mm.
^{1}H- and ^{13}C-NMR spectra were recorded on a Varian Innova instrument (600 MHz) at room temperature. All chemical shifts are expressed in ppm with respect to the signals of the residual protonated solvents ($CDCl_3$ or DMSO d_6).

II.4.2 Synthesis of the linker N-N' bis-cysteinyl- ethylendiamine

To a solution of ethylendiamine (30.1µL, 0.45mmol) and DIPEA (485.5 µL, 2.7 mmol) in dry DCM (300 µL) was added a solution of Boc-Cys-(Trt)-OH (500mg, 1.08 mmol), PyBOP (467 mg, 0.90 mmol) and HOBT (121.4 mg, 0.90 mmol) in dry DCM (2 mL). The reaction was stirred overnight at room temperature.

The mixture was extracted with a 5% $NaHCO_3$ aqueous solution. The organic phase was dried over $MgSO_4$ and the solvent was evaporated under reduced pressure. The residue was purified by silica gel flash chromatography (ethyl acetate/petroleum ether, 7/3 v/v) to afford 427 mg of protected linker N-N' bis[N-tert-butyloxycarbonyl-S-triphenylmethyl-cysteinyl] ethylendiamine as a yellow oil (99% Yield). The Boc/Trt protected linker (159 mg, 0.16 mmol) was dissolved in CH_2Cl_2/TFA/TIS 50/47/3 v/v/v (4 mL). Deprotection of the Trt and Boc groups was complete after 5 minutes. The solution was concentrated under reduced pressure and diluted with cold diethyl ether. The precipitated crude product N-N' bis-cysteinyl-ethylendiamine was dissolved in water, analyzed by LCMS and purified by preparative HPLC with an increasing gradient of CH_3CN (0.1% TFA) in water (0.1% TFA) from 1 to 70% in 30 minutes.

II.4.3 Protein expression and purification

E. coli BL21(DE3) cells, transformed with the appropriate plasmid, were grown to mid-log phase at 37°C in LB medium. Protein expression was induced with 0.4 mM Isopropylbeta- D-thiogalactopyranoside (IPTG) at 37°C for 5 h, after which cells were harvested and lysed by sonication. Protein expression was followed by SDS-PAGE (15%). Protein A-MxeGyrA fusion was purified at room temperature by affinity chromatography on a Ni^{2+}NTA resin. Protein was eluted in 50 mM NaH_2PO_4 , 300 mM NaCl, 250 mM imidazole pH 8. Protein B-MxeGyrA-fusion was immobilized and purified on chitin resin, following standard protocol of IMPACT™ TWIN manual.

II.4.4 Synthesis of the homodimer (A-linker-A)

Purified protein A-MxeGyrA fusion (150µM) was reacted in Phosphate buffer 50mM, 0.2mM MESNA, 0.3mM EDT, pH 7 with 0.5 (or 3) equivalents of linker. Splicing and ligation reactions were performed simultaneously overnight at room temperature.
The reaction leads to the homodimer in a 50% yield.
The crude was analysed by LCMS with an increasing gradient of CH_3CN (0.1% TFA), in water (0.1% TFA) from 5 to 70% in 30 minutes. Mass spectrum shows the presence of the homodimer A-linker-A and the protein A (or protein A-linker derivative).

Protein A sequence:
MGGSHHHHHH GMASMTGGQQ MGRDLYDDDD KDRWGSGHIE GR

II.4.5 Synthesis of the heterodimer (A-linker-B)

Purified protein A-MxeGyrA fusion (50μM) was reacted in Phosphate buffer 20mM, 0.18mM MESNA, 0.23mM EDT, pH 7 with a three fold excess of linker. Splicing and ligation reactions were performed simultaneously overnight at room temperature to give the protein A-linker derivative in quantitative yield. The crude was dyalized against deionized water; A-linker was purified by preparative HPLC with an increasing gradient of CH_3CN (0.1% TFA), in water (0.1% TFA) from 10 to 45% in 38 minutes. Protein B-MxeGyrA fusion was immobilized on the chitin resin, splicing was induced incubating the resin in 20 mM Phosphate buffer containing 300 mM NaCl, 50 mM MESNA and 1 mM EDTA, pH7. The thioester was purified by preparative HPLC with an increasing gradient of CH_3CN (0.1% TFA), in water (0.1% TFA) from 10 to 40% in 30 minutes. The purified protein A-linker (0.36mM) derivative was reacted with one equivalent of thioester protein B (0.36mM). Reaction was carried out in 20 mM Phosphate buffer containing 0.18 mM MESNA and 0.23 mM EDT, pH 7 overnight. The crude was analyzed by LCMS with an increasing gradient of CH_3CN (0.1% TFA), in water (0.1% TFA) from 5 to 70% in 30 minutes. Protein A-linker was all converted into the homodimer A-linker-B, while the protein B excess was converted into the EDT-thioester.

Protein B sequence:
MASSRVDGGR EFLEGSS

II.5 References

- Botti P., S. Tchertchian, WO 2006/133962, 2006
- Canne L.E.; Bark S.J.; Kent S.B.H.; J. Am. Chem. Soc. (1996) 118: 5891–96.
- Canne, L. E.; Ferre- D'Amare, A. R.; Burley, S. K.; Kent, S. B. H.; J. Am. Chem. Soc. (1995) 117: 2998.
- Chatterjee, C.; McGinty, R. K.; Pellois, J. P.; Muir, T. W.; Angew. Chem., Int. Ed. (2007) 46: 2814–2818.
- Chong, S.; Shao, Y.; Paulus, H.; Benner, J.; Perler, F. B.; Xu, M. Q.; J. Biol. Chem. (1996) 127: 22159–22168.
- Chong, S.; Shao, Y.;Paulus, H.; Benner, J.; Perler, F.B.; Xu, M.Q.; J Biol Chem (1996) 271:22159-22168.
- Cordes EH, Jencks WP. J Am Chem Soc 1962, 84:826-831
- Crich D., A. Banerjee, J. Am. Chem. Soc. 2007, 129, 10064 – 10065;
- Cuenoud, B.; Schepartz, A.; Science (1993) 259: 510.
- Dawson, P.E.; Muir,T.W.; Clark-Lewis; Kent, S.B.H.; Science (1994) 266: 776-779
- Dirksen A, Hackeng TM, Dawson PE: Angew Chem Int Ed 2006, 45:7581-7584
- Dirksen A. and Dawson P.E., Curr.Op.Chem. Biology 2008, 12:760–766
- Evans, T.C.; Benner, J:; Xu, M.Q.;Protein Sci (1998) 7:2256-2264.
- Hahn, M. E.; Pellois, J. P.; Vila-Perello, M.; Muir, T. W.; ChemBioChem (2007) 8, 2100–2105.
- Kolb HC, Finn MG, Sharpless KB. Angew Chem Int Ed 2001, 40:2004-2021
- Lajmi, A. R.; Lovrencic, M. E.; Wallace, T. R.; Thomlinson, R. R.; hin, J. A.; J. Am. Chem. Soc. (2000) 122: 5638.
- Liu, X.Q.; Yang, J.; Meng, Q.; J Biol Chem (2003) 278:46826-46831.
- Mills, K.IV.; Lew, B.M,; Jiang, S.Q.; Paulus, H.; Proc. Natl. Acad. Sci. USA (1998) 95:3543–48.
- Morii, T.; Saimei, Y.; Okagami, M.; Makino, K.; Sugiura, Y.; J. Am. Chem. Soc. (1997) 119: 3649.
- Muir, T. W.; Annu. ReIV. Biochem. (2003) 72: 249–289.
- Muir, T.W.; Sondhi, D.;,Cole, P.A.; Proc Natl Acad Sci USA (1998) 95:6705- 6710.
- Nilsson, B.L.; Kiessling, L.L.; Raines, R.T.; Org. Lett. (2000) 2:1939 – 1941.
- Ogata, H.; Raoult, D.; Claverie, J.M.;: Virol J (2005) 2:8.
- Paulus, H.; Annu Rev Biochem (2000) 69:447-496
- Perler, F.B.; Davis, E.O.; Dean, G.E.; Gimble, F.S.; Jack, W.E.; Neff, N.; Noren, C.J.; Thorner, J.; Belfort, M.; Nucleic Acids Res (1994) 22:1125-1127
- Perler, F.B.; Nucleic Acids Res (2002) 30:383-384
- Punna S, Kuzelka J, Wang Q, Finn MG. Angew Chem Int Ed 2005, 44:2215-2220
- Sato, S.; Hagihara, M.; Sugimoto, K.; Morii, T. Chem. Eur. J. (2002) 8: 5066.

- Saxon, E.; Armstrong, J.I.; Bertozzi, C.R.; Org. Lett. (2000) 2: 2141–2143.
- Schwarzer, D.; Cole, P. A. Curr. Opin. Chem. Biol. (2005) 9: 561–569.
- Smart, O. S.; Kumita, J. R.; Biochemistry (2006) 45: 6075.
- Southworth, M.W.; Adam, E.; Panne, D.; Byer, R.; Kautz,R.; Perler, F.B.; EMBO J. (1998) 17:918–26.
- Tam, J. P.; Lu, Y. A.; Yang, J. L.; Eur. J. Biochem. (2002) 269: 923.
- Tam, J. P.; Yu, Q.; Org. Lett.(2002) 4: 4167.
- Ueno, M.; Murakami, A.; Makino, K.; Morii, T.; J. Am. Chem. Soc. (1993) 115: 12575.
- Van Baal, I.; Malda, H.; Sinowski, S. A.; van Dongen, J. L. J.; Hackeng, T. M.; Maarten, M.; Meijer, E. W.; Angew. Chem., Int. Ed. (2005) 44: 5052
- Woolley, G. A.; Jaikaran, A. S.; Berezovski, M.; Calarco, J. P.; Krylov, S. N.;
- Wu B., J. H. Chen, J. D.Warren, G. Chen, Z. H. Hua, S. J. Danishefsky, Angew. Chem. 2006, 118, 4222 – 4231
- Wu, H.; Hu, Z.; Liu, X.Q.; Proc Natl Acad Sci USA (1998)95:9226-9231.
- Xu, M. Q.; Evans, T. C. Jr; Curr. Opin. Biotechnol. (2005) 16: 440–446.
- Xu, M. Q.; Perler, F. B.; EMBO J. (1996) 15: 5146–5153.
- Yamazaki, T.; Otomo, T.; Oda, N.; Kyogoku, Y.; Uegaki, K.; J. Am. Chem. Soc. (1998) 120:5591–92.
- Yang, J.; Meng, Q.; Liu, X.Q.; Mol Microbiol (2004) 51:1185-1192.

III BIOPHYSICAL CHARACTERIZATION OF A α-HELIX SCAFFOLD PEPTIDE

BIOPHYSICAL CHARACTERIZATION OF A α-HELIX SCAFFOLD PEPTIDE

III.1 Introduction

The a-helix is the most abundant secondary structure in proteins and is frequently involved in functionally important protein–protein and protein–nucleic acid interactions (R.Lavery, 2005; A.Klug, 2005; J. M. Davis et al., 2007). However, the removal of these polypeptide recognition motifs, which are typically 15–25 residues in length, from the stabilising tertiary structure of proteins generally results in peptides that adopt only random coil structures or low populations of conformations containing a-helical secondary structure in water. In addition, the efficacy of short polypeptides corresponding to the a-helical regions in proteins is compromised in vivo due to an increased susceptibility to proteolytic degradation and a reduction in cell wall permeability (M. J. Kelso and D. P. Fairlie, 2003).

Synthetic strategies to develop peptides presenting stable α-helical conformation have elicited considerable interest as a means of generating potential therapeutic agents or probes for exploring protein–protein interactions (Gamer, J. and Harding, M.M., 2007; Henchey, L. K. et al., 2008).

The majority of the a-helix mimetic-based molecules reported to date can be subdivided into two categories. The first includes molecules that adopt helical structures similar to natural helices and most commonly include peptidomimetics such as β-peptides and peptoids (C. M. Goodman et al., 2007; J. A. Kritzer et al., 2005; J. D. Sadowsky et al., 2005; T. Hara et al., 2006). The second ignores the helical structure of natural helices and instead uses organic scaffolds to project functionality in a similar spatial arrangement to the natural conformation.[J. M. Davis et al., 2007; Yin H. et al., 2005 a; S. M. Biros et al., 2007 ; Yin H. et al., 2005 b; Yin H. et al., 2005 c; J. Becerril et al., 2007). In both cases, it is the strategic functionalization of the scaffolds that enables the development of compounds that effectively recognize biomacromolecules (Rodriguez JM et al., 2009).

Verdine and its group described the design and synthesis of stabilized peptide helices (staple) to bind to a protein–protein interface. In particular, peptide stapling is accomplished by synthetic incorporation of two α-methylated amino acids, bearing olefinic side-chains of varying length and configured with either R or S stereochemistry, along one face of the α-helix, followed by ring-closing olefin metathesis (Schafmeister et al., 2000). This strategy find application to inhibit growth of human leukemia xenografts (Walensky, L. D et al., 2004; Kutchukian, P. S. et al., 2009).

Brines et al., instead, developed a peptides mimetic of helix B of the erythropoietin. They showed that this peptide simulating the aqueous surface of helix B also exhibits EPO's trophic effects by accelerating wound healing and augmenting cognitive function in rodents.

Li et al. reported the interaction of a single short helical segment of the tumor-suppressor protein p53 with two regulators of p53 activity, mouse double minute 2 (MDM2) and MDMX. In this work, the grafting of four residues of p53 critical for MDM2/MDMX binding to the N-terminal alpha-helix of BmBKTx1, a scorpion toxin isolated from the venom of the Asian scorpion Buthus martensi Karsch, converts the miniature protein into an effective inhibitor of p53 interactions with MDM2 and MDMX. Additional mutations enable the 27-residue miniprotein inhibitor to traverse the cell membrane and selectively kill tumor cells in a p53 dependent manner (Li et al., 2008; Li et al., 2009).

Recently, unnatural oligomers with aromatic repeating units were developed by Hamilton and coworkers, these small molecule scaffolds can mimic the structural and recognition binding features of an R-helix (Figure III.1)

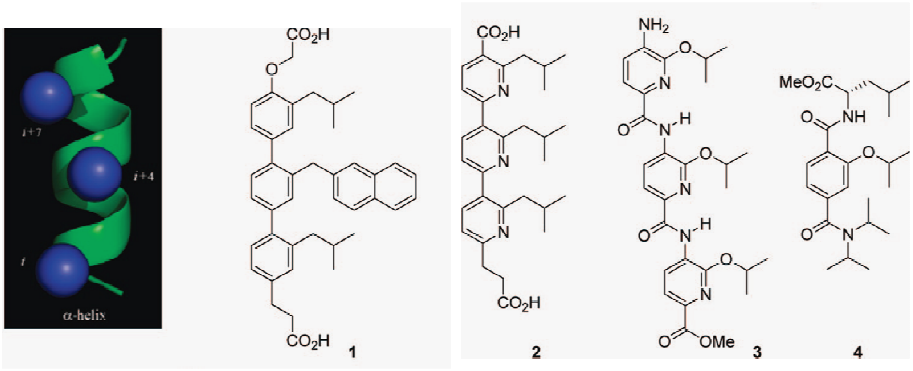

Figure III.1: Unnatural oligomers with aromatic repeating units developed by the group of Hamilton.

A functionalized terphenyl scaffold **1** was found to provide a rigid framework from which aryl o-substituents are projected to mimic the side chains at the i, i+4, and i+7 positions of an R-helix (Fairlie, D. P. et al., 1998). This design was extended to other closely related structures including terpyridine **2** (Davis, J. M., 2005), oligoamide **3** (Ernst, J. T., 2003) and terephthalamide **4** derivatives (Yin H. et al., 2005; Yin, H. and Hamilton A., 2004). These rationally designed compounds were shown to effectively inhibit protein-protein interactions featuring R-helix-mediated binding and recognition including Bcl-xL/Bak and p53/HDM2 thus validating the design (A. Shaginian et al., 2009).

The purpose of this work was the development of a molecular scaffold to be modulate biomolecules interactions mediated by α-helix.

Recently, the design, biological and structural characterization of a short peptide mimicking the Vascular endothelial growth factor (VEGF) was described (D'Andrea et al., 2005). This 15-mer peptide, (named QK), composed only of natural amino acids was designed to reproduce the VEGF helix region 17–25. NMR conformation analysis of QK revealed that it adopts a well defined helical conformation in water, moreover, it was

demonstrated that QK has the same biological properties of VEGF *in vitro* (D'Andrea et al., 2005) and *in vivo* (Dudar et al., 2008; Santulli et al., 2009).

This project focus on the biophysical characterization of peptide QK has potential peptidic scaffold to modulate interactions mediate by α-helix.

III.2 Results

NMR experiments were carried out in collaboration with the laboratory of Prof. R. Fattorusso the Second University of Naples.
MD experiments were performed in collaboration with Dr. G. Colombo CNR.

III.2.1 Analysis of the thermal stability of QK

The thermal stability of the QK peptide was investigated by CD and NMR spectroscopy. In particular, the helical content was investigated by monitoring the temperature-dependent change in the intensity of the 222 nm band in the CD spectra. The analysis of the spectra and the dependence of ellipticity at 222 nm with the temperature showed a gradual increase in the signal intensity at 222 nm with the temperature, which indicated helix unwinding at high temperature. Nevertheless, QK helix shown a remarkable degree of thermal stability, as the peptide retains 79% and 65% of its room temperature helix content at 343 K and 368 K, respectively (Figure III.2).

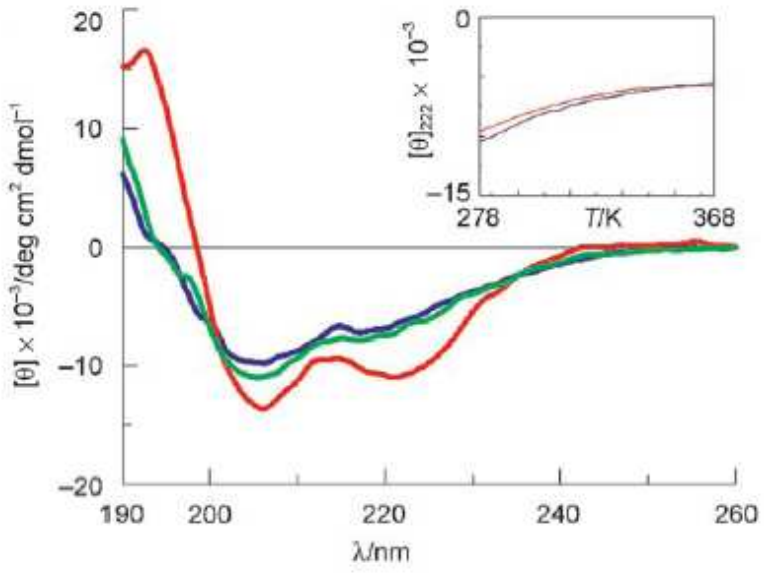

Figure III.2: CD spectra of the QK peptide at 298K (red), 343K (green) and 368K (blue).
In the inset the unfolding (blue) and refolding (red) curves are shown.
[θ] is the molar ellipticity per residue.

The analysis of the spectra and the dependence of ellipticity at 222 nm with the temperature showed that the peptide loose reversibly part of its helical structure but neither at 368K appears to assume a complete random coil conformation.
Similar experiments were carried out using NMR, the variations of the Hα protons with the temperature were reported for each amino acid (Figure III.3).

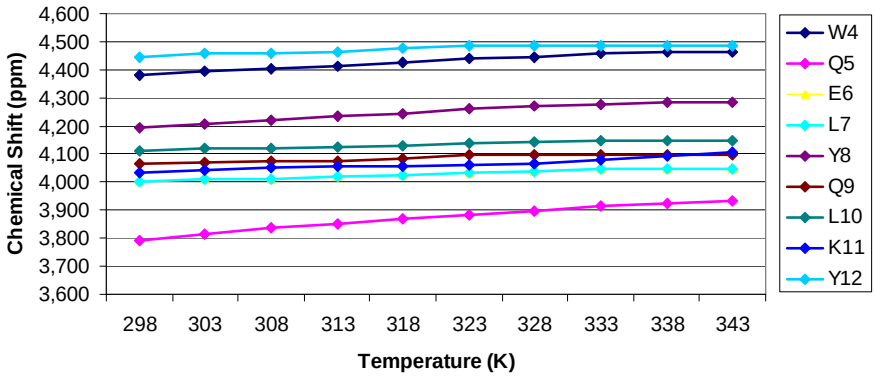

Figure III.3: Plot of the backbone protons chemical shifts against temperature for nine representative residues.

The plot of Ha chemical shifts again showed that only minor changes were observed in analyzed range suggesting a significant thermal stability. QK structure variations upon temperature increase were followed also by ^{2}D TOCSY NMR experiments (Figure III.4).

a) b) c)

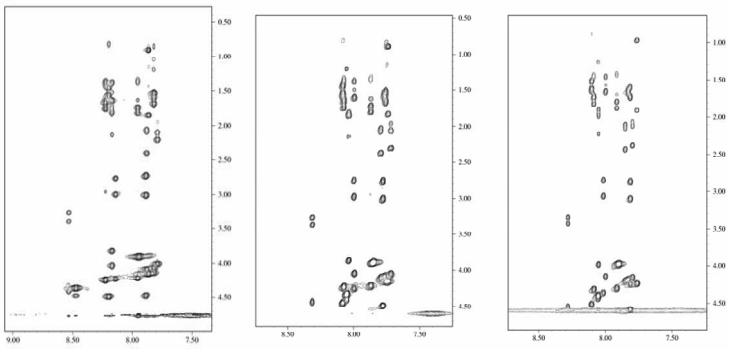

Figure III.4: Fingerprint region of the 2D [^{1}H, ^{1}H] TOCSY spectra of QK peptide at a) 298K, b) 323K, and c) 343K

Table III.I in Appendix III.6 reports the ^{1}H chemical shift assignment of the Hα and H_N protons and clearly shows only minor changes in the ^{1}H chemical shifts for backbone protons collected at temperature increments; then, it was possible to track with confidence the identities of many peaks (approximately 70%) previously assigned at room temperature.

Subsequently, the temperature dependence of the Hα chemical shift deviations from random coil values (ΔδHα), which are an indicator of the secondary structure formation, were used as a measure of stability at high temperature. The Figure III.5 shows a continuous stretch of negative values encompassing residues 4-12, strongly suggesting the presence of helical conformation at high temperature.

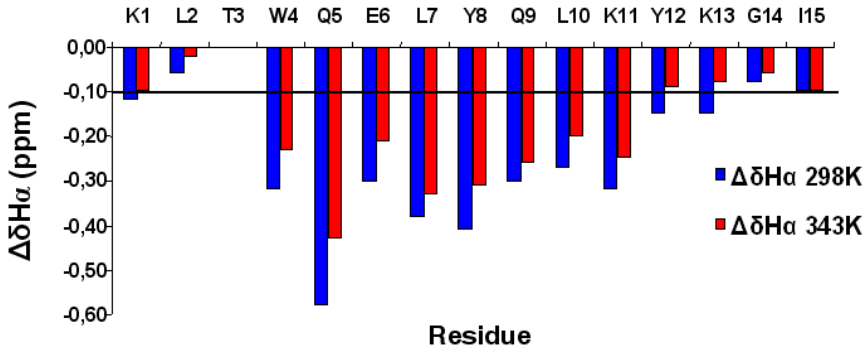

Figure III.5: Hα chemical shift deviations from random coil values (ΔδHα) of the QK at 298K and 343K. The continuous line represents the CSI threshold for amino acid in helical conformation.

Moreover, QK helix has the property that at higher temperatures ΔδHα values decrease, reflecting a slight reduction in the amount of folded structure present in solution. In particular, moving from 298 K to 343 K the peptide retains the 80% of its secondary structure. The striking observation is that all residues, in particular 4-12 region of the sequence, behave in an analogous manner, reflecting similar temperature changes in the population of the folded state.

III.2.2 Molecular Dynamics Analysis

To get insight on the molecular determinant of this unusual thermal stability the conformational and dynamic properties of the QK peptide were studied with Molecular Dynamics simulations in explicit water.

Extensive MD simulations were carried out for a total of 2.4 microseconds exploring different temperature conditions (300K, 320K, 340K, 380K).

Five α-helical structures corresponding to five different models present in the NMR ensemble were used as starting structures for the simulations, to assess the determinants of the stabilization of the helix in solution. Moreover, in order to shed light on the possible folding mechanism, four different simulations with lengths ranging from 50 to 100ns, at 350K, were run from a completely extended structure of the polypeptide.

All the simulations starting from the helical structures show a clear, unusual stability of the helix that is maintained at high temperatures (Figure III.6) for most of the simulation time, consistently with NMR observations.

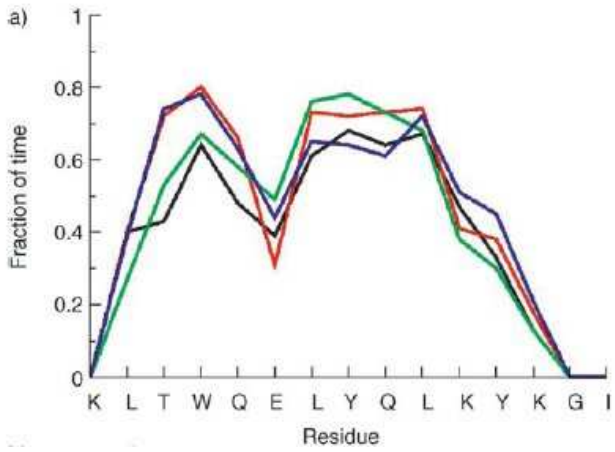

Figure III.6: Percentage of time that each residue spends in helical conformation at 300 K (black), 320 K (red), 340 K (green) and 380 K (blue).

Cluster analysis of the trajectories, and the evaluation of stabilizing contacts, show the presence of a network of contacts always involving the hydrophobic side chains of residues Leu7 and Leu10 (Figure III.7).

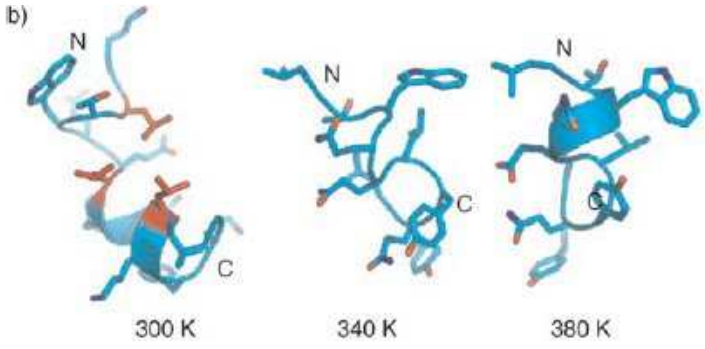

Figure III.7: Representative conformations of main cluster obtained from the analysis of all trajectories at different temperatures.

On the other hand, analysis of the refolding simulations shown a higher tendency for residues located at the N-terminal or central part of the sequence to adopt a helical conformation in the first events of the QK folding. (Figure III.8).

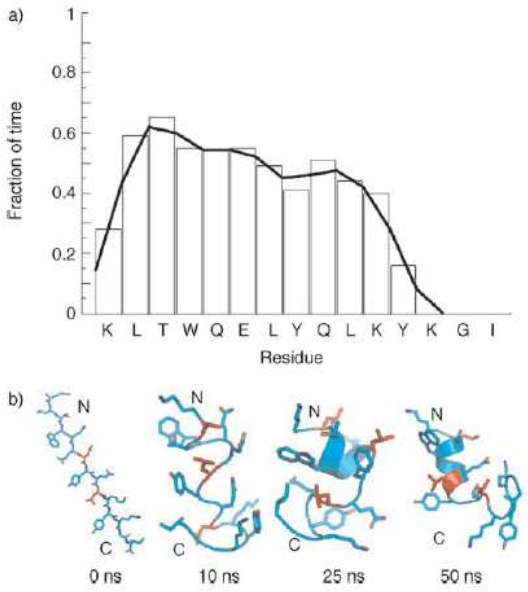

Figure III.8: a) Percentage of helical conformation attained by each residue during the refolding process. b) Selected structures along the refolding trajectories. Leu7 and Leu10 are highlighted in red.

III.2.3 Peptide analogues studies

To confirm experimentally the suggestions coming from MD simulations we designed three QK analogues which amino acid sequences are:

QK	**Ac-KLT WQELYQLKY KGI-NH2**
QK1-12	**Ac-KLT WQELYQLKY -NH2**
QK4-15	**Ac-K WQELYQLKY KGI-NH2**
QK10A	**Ac-KLT WQELYQAKY KGI-NH2**

To test the role of the capping regions we synthesized QK1-12 and QK4-15 peptides which present the deletion of the C-capping and N-capping sequence respectively.
In order to verify the importance of the hydrophobic contact between Leu 7 and Leu 10 for the helix stability, we replaced Leu 10 with Ala (QK10A).
Peptides were synthesized by SPPS and purified by RP-HPLC. All peptides were efforded in good yields, in high pure and homogenous forms as assessed by LC-MS (Figure III.9, III.10 and III.11). Following the identity was verified by ESI spectrometry and a comparison of experimental and calculated MW is reported below.

Peptide	**MW_{Th}**	**MW_{Exp}**
QK 1-12	1655.8	1655.2
QK 4-15	1738.9	1738.2
QK10A	1912.1	1910.9

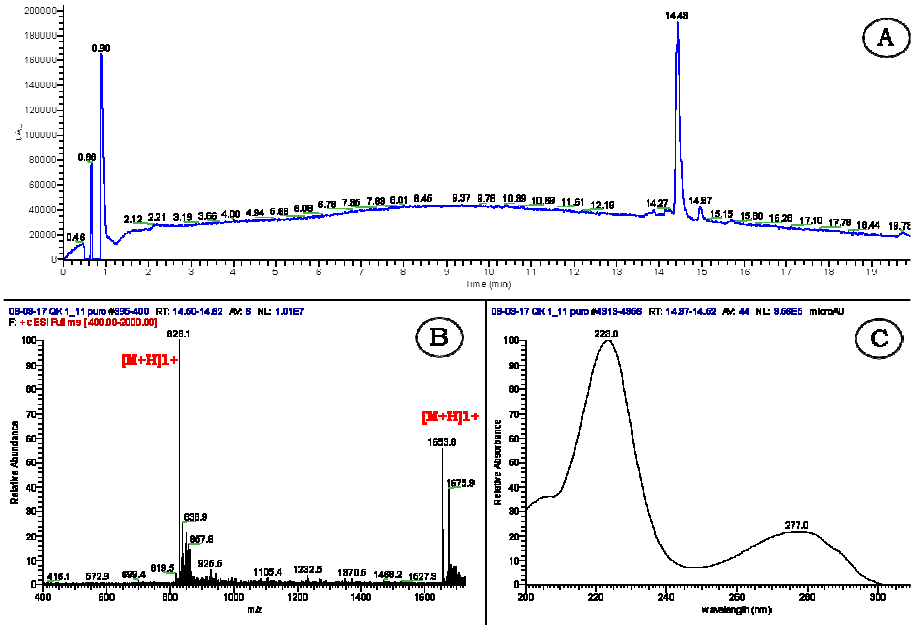

Figure III.9: LC-MS analysis of QK1-12 peptide
A) RP-HPLC profile revealed at 210nm; **B**) ESI-MS spectrum of the peak at RT:14.43 min;
C) UV absorption spectra of the peak at RT:14.43.

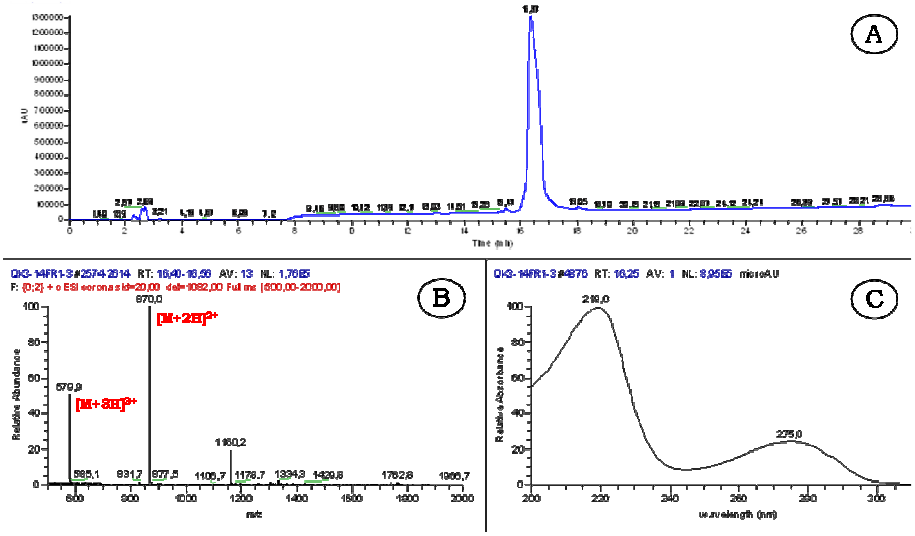

Figure III.10: LC-MS analysis of QK4-15 peptide
A) RP-HPLC profile revealed at 210nm; **B**) ESI-MS spectrum of the peak at RT:16.37 min;
C) UV absorption spectra of the peak at RT:16.37 min.

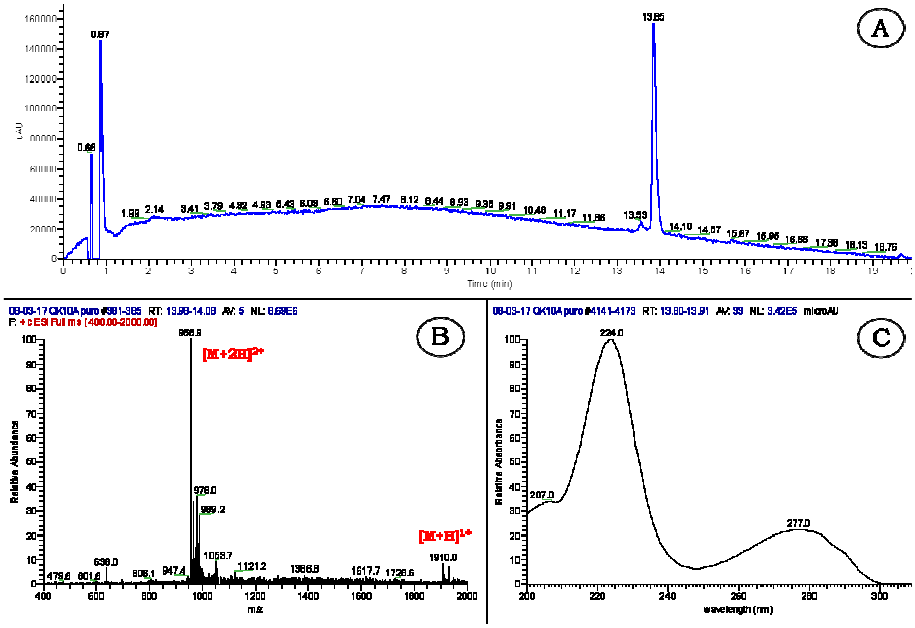

Figure III.11: LC-MS analysis of QK10A peptide
A) RP-HPLC profile revealed at 210nm; **B**) ESI-MS spectrum of the peak at RT:13.85 min;
C) UV absorption spectra of the peak at RT:13.85 min.

The helical content of each peptide was analyzed by means of CD spectroscopy. Peptides were dissolved in 10 mM phosphate buffer, pH 7.1 at a concentration of 50 µM and spectra were recorded at 5°C (Figure III.12 and III.13).

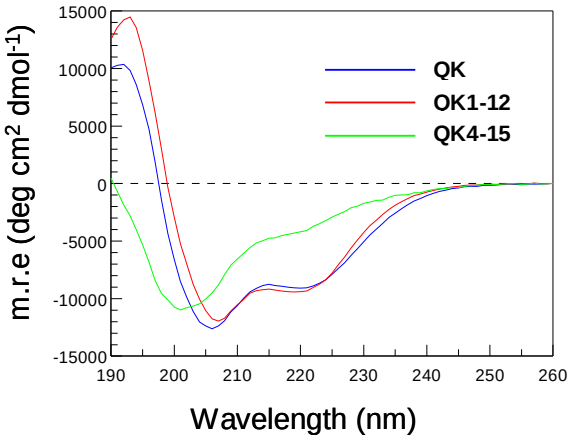

Figure III.12: Overlap of CD spectra of QK, QK 1-12 and QK 4-15

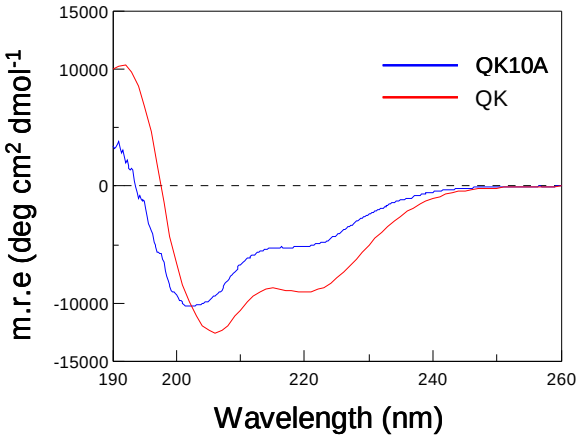

Figure III.13: Overlap of CD spectra of QK and QK10A

All peptides, except QK4-15, showed the two minima at 207nm and 222nm characteristic of α-helical conformation. In particular, CD spectra clearly showed that QK4-15 loose about half of the QK helicity while QK1-12 entirely retains the helical content; this result suggests the importance of the N-terminal in stabilizing the helix conformation. The CD spectra in figure III.13 showed that QK10A loose about 50% of the QK helical content. This result is quite unusual because alanine is an amino acid that tends to stabilize the helical conformation with respect to leucine, whereas for QK10A there was a loss, this suggests that the hydrophobic contact between the two leucine is really important for stability of helix.

III.2.4 Conformational analysis and thermal unfolding on QK1-12

^{1}H NMR spectra have been fully assigned, using a combination of DQF-COSY, TOCSY and NOESY data sets, following the same procedure described in Chapter III.3.
^{1}H chemical shift assignment and $^{3}J_{HNH\alpha}$ measured coupling constant values are listed in table III.2 in Appendix III.6. DOSY measurements has given a value of $2.30*10^{-10}$ m^2 s^{-1}, which suggested that QK1-12 is monomeric in solution. The 2D NOESY and of the DOSY spectra are shown in Figure III.14 and III.15.

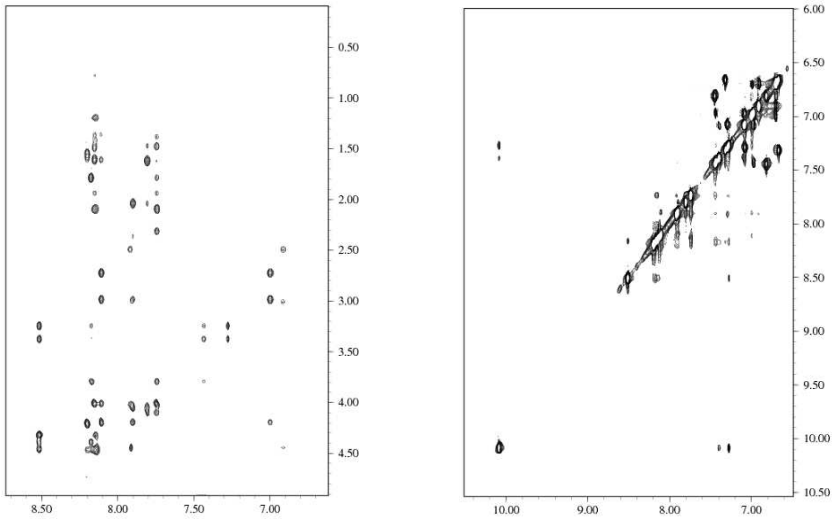

Figure III.14: 2D [^{1}H, ^{1}H] of 250 ms NOESY spectrum of QK1-12 peptide

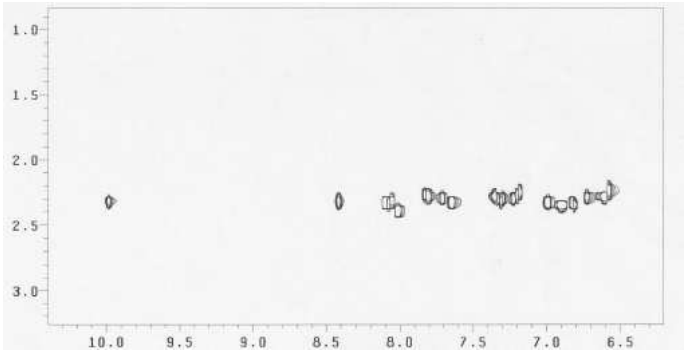

Figure III.15: 600 MHz ^{1}H-detected DOSY spectrum of 1mM of QK1-12 in H_2O/D_2O mixture at 298K

ΔδHα NMR analysis shown a characteristic appearance of a helix, encompassing the residue 4-12 (Figure III.16).

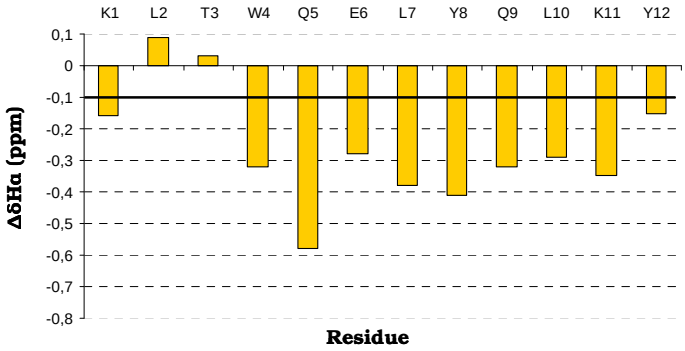

Figure III.16: : Hα chemical shift deviations from random coil values (ΔδHα) of the QK1-12 at 298K. The continuous line represents the CSI threshold for amino acid in helical conformation.

The most conclusive evidence for helical folded structure was the observation of a pattern of NOEs, HN-$HN_{(i,\ i+1)}$, Hα-$HN_{(i,\ i+3)}$ and Hα-$H\beta_{(i,\ i+3)}$, which is well consistent with the α-helix conformation. A total of 167 NOEs were assigned and quantitated for use in structure calculations. The NOEs were broken down into 37 intra-residue, 38 short- and

34 medium-range. A total of 71 torsional angle restraints for Φ and Ψ angles were determined from the $^{3}J_{HNH\alpha}$ coupling constants.

The final 20 annealed structures were shown in Figure III.17; the whole range of structures was a helix shaped with a good pairwise RMSD from the backbone atoms and from all heavy atoms of the mean structure, 0.12 (±0.01)Å and 0.98 (±0.02)Å, respectively.

NMR Structural Statics of the QK1-12	
NMR Distance Constraints	
Distance Constraints	109
Total NOE	167
Intra residue	37
Short range	38
Medium range	34
Total Diehedral Angles	71
Φ (Phi)	35
Ψ (Psi)	36
Avarage Pairwise RMSD (Å)	
Heavy (3..12)	0.98
Backbone (3..12)	0.12

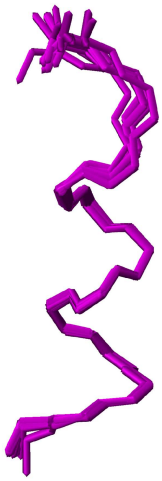

Figure III.17: Superposition of the backbone of the best 20 QK1-12 structures.

Also for QK1-12 were recovered the variations of the Hα protons with the temperature were reported for each amino acid (Figure III.18).

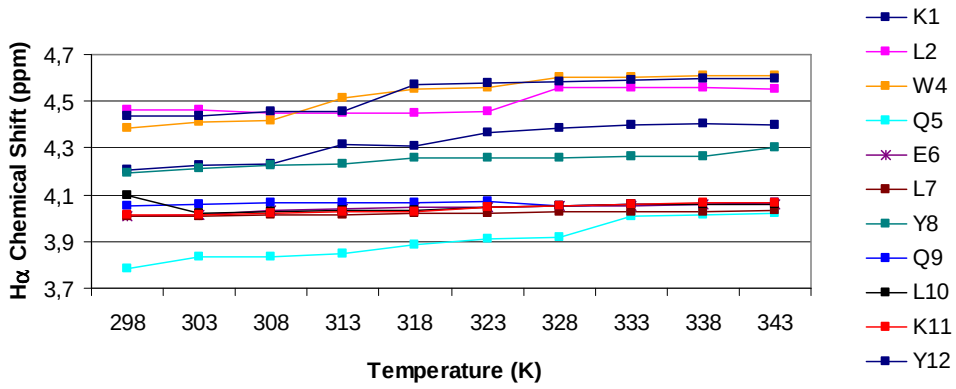

Figure III.18: Plot of the backbone protons chemical shifts against temperature for nine representative residues

The plot of Ha chemical shifts again showed that only minor changes were observed in analyzed range suggesting a significant thermal stability as QK peptide. QK1-12 structure variations upon temperature increase were followed also by 2D TOCSY NMR experiments. Table III.3 in Appendix III.6 reports the 1H chemical shift assignment of the Hα and HN protons and clearly shows only minor changes in the 1H chemical shifts for backbone protons collected at temperature increments.

In addition, the temperature dependence of the Hα chemical shift deviations from random coil values (ΔδHα), which are an indicator of the secondary structure formation, were used as a measure of stability at high temperature. The Figure III.19 shows a continuous stretch of negative values encompassing residues 4-12, suggesting, also for QK1-12, the presence of helical conformation at high temperature.

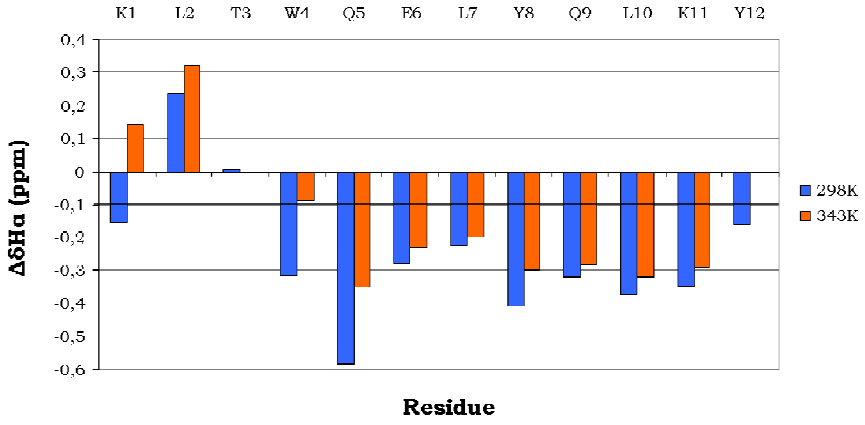

Figure III.19: Hα chemical shift deviations from random coil values (ΔδHα) of the QK1-12 at 298K and 343K.

III.2.5 Conformational analysis and thermal unfolding on QK10A

To verify the importance for the helix stability of the hydrophobic contact between Leu7 and Leu10 suggested by the MD simulation, residue Leu10 was replace with Ala. Similar

helical propensity of Ala and Leu allowed to relate the changes of the mutant structural stability to the absence of the side chains interaction.

ΔδHα NMR analysis (Figure III.20) at room temperature showed that QK10A looses about 35-40% of the QK helical content.

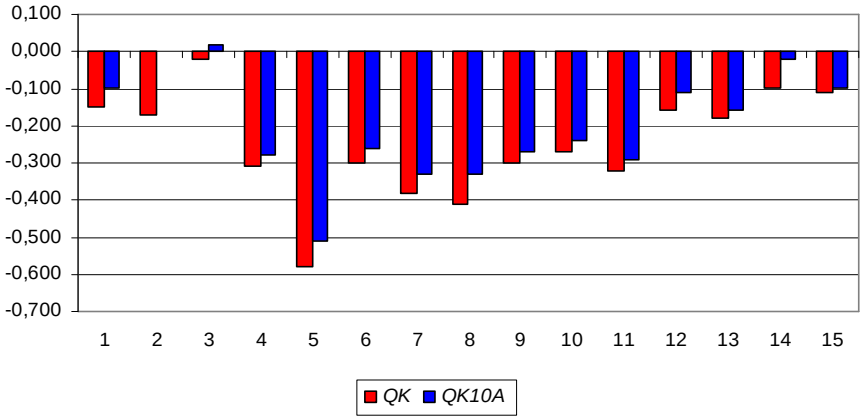

Figure III.20: Plot of ΔδHα for QK (red) and QK10A (blue) at 298 K

The NMR structure of QK10A in water solution at 298 K were obtained by a careful inspection of a combinations of DQF-COSY, TOCSY and NOESY data sets, following the same procedure described in Chapter III.3.

^{1}H chemical shift assignment and ^{3}JHNHα measured coupling constant values are listed in table III.4 reported in Appendix III.6. DOSY measurements has given a value of 1.98*10-10 m^2s^{-1}, the same diffusion coefficient value measured for QK.

The DOSY and of the NOESY spectra are reported in Figure III.21 and III.22.

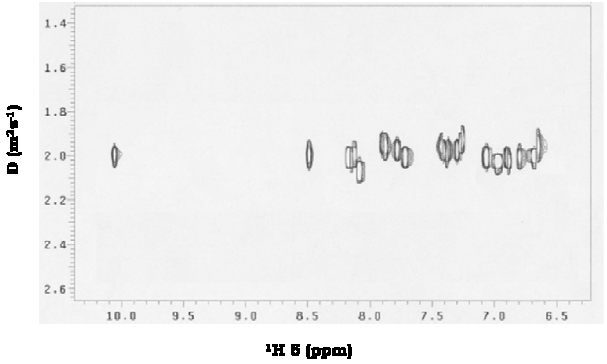

Figure III.21: 600 MHz 1H-detected DOSY spectrum of 1mM of QK10A in H_2O/D_2O mixture at 298K.

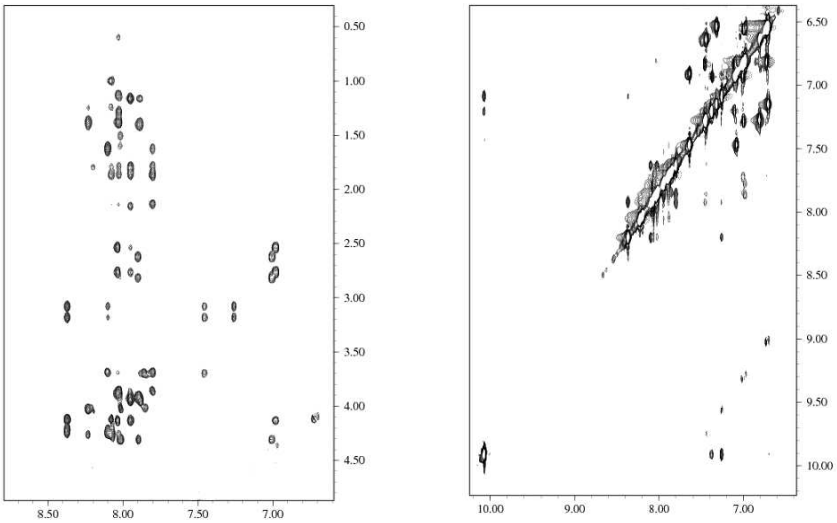

Figure III.22: 2D [^{1}H, ^{1}H] of 250 ms NOESY spectrum of QK10A peptide at 298 K.

To calculate the NMR structure of the QK10A mutant, a total of 137 NOE constraints (47 intra-residue, 54 short- and medium-range) and 71 torsional angle restraints for Φ and Ψ angles were applied.

NMR Structural Statics QK10A at 298 K	
NMR Distance Constraints	
Distance Constraints	206
Total NOE	137
Intra residue	47
Short range	54
Medium range	36
Total Diedral Angles Restraints	72
Φ (Phi)	36
Ψ (Psi)	36
Avarage Pairwise RMSD (Å)	
Heavy	1.05
Backbone	0.15

The NMR structure of QK10A showed that it is predominantly in helical conformation such as the peptide QK (Figure III.23).

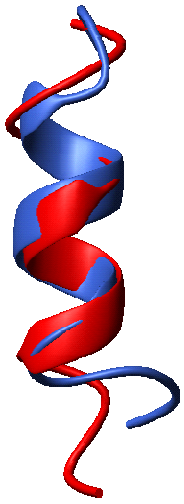

Figure III.23: Comparison of the NMR-derived structures, as ribbon model, of QK (blue) and QK10A (red) in water at 298 K.

To better investigate the helix destabilization, due to the absence of the interaction between Leu7 and Leu10, the QK10A thermal unfolding has been performed via NMR, acquiring 2D TOCSY at different temperature.

Table III.5 in Appendix III.6 showed the lists of the chemical shift of the Hα protons at increasing temperature.

ΔδHα analysis at 298 K and 343 K clearly indicated that at high temperature QK10A is predominantly in random coil conformation, loosing the unusual thermal stability of QK (Figure III.24).

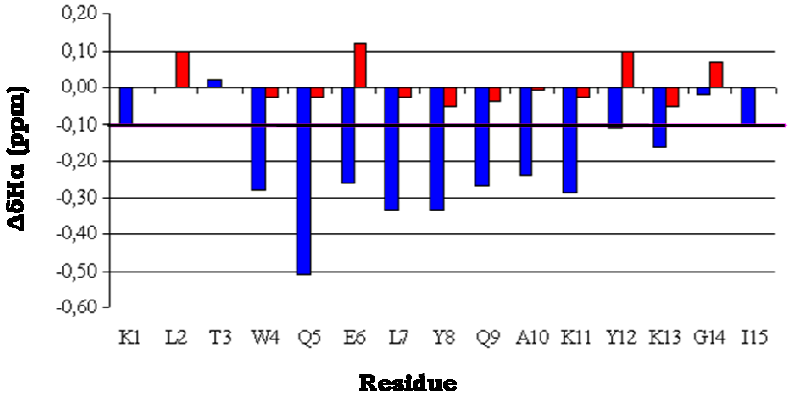

Figure III.24: Plot of ΔδHα for QK10A at 298 K (blue) and 343 k (red).

Moreover, the variations of the Hα protons with the temperature were reported for each amino acid. The curves of the amino acids at the N- and C-terminal region present a sigmoidal-like behavior, whereas curves of the central region present a bi-phasic behavior, suggesting, in both cases, a melting temperature of about 313K (Figure III.25).

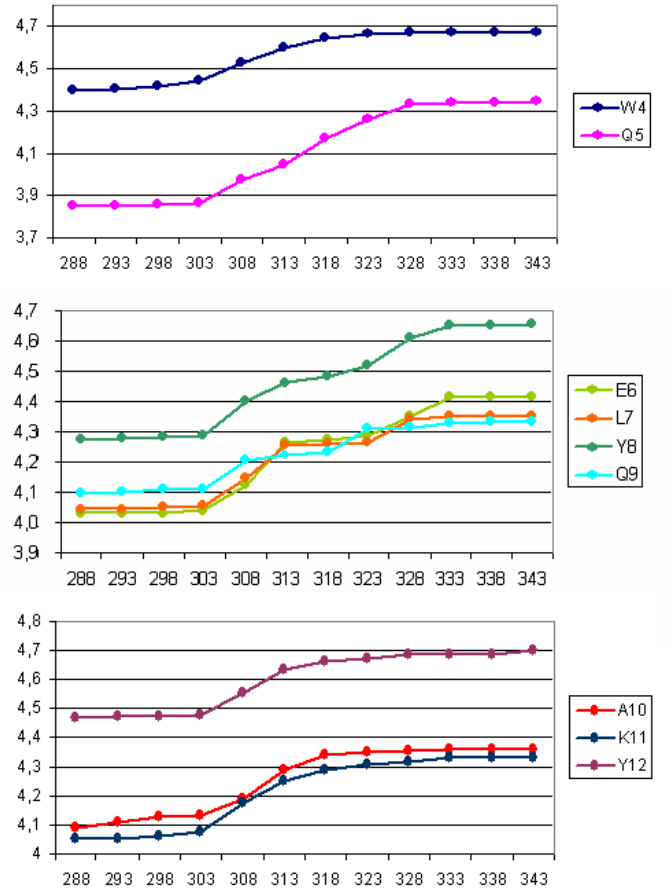

Figure III.25: Temperature-dependent chemical shifts for Hα protons.

Then, to verify the presence of an intermediate structure, the solution structure of QK10A at 313K was obtained. Different mixing times were used to evaluate the linear build-up of NOE and to find the mixing time appropriate at 313 K; NOESY spectrum recorded with a mixing time of 350 ms was chosen for obtaining the distance constraints. The complete ^{1}H assignment at 313 K is reported in Appendix III.6 (Table III.6), while a section of the ^{2}D NOESY is reported in Figure III.26.

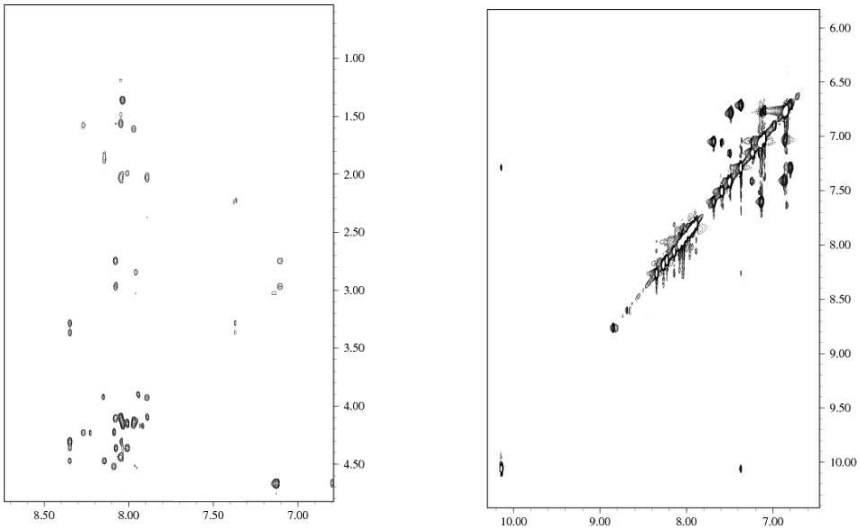

Figure III.26: 2D [^{1}H, ^{1}H] of 350 ms NOESY spectrum of QK10A peptide at 313 K.

On the basis of observation of a relatively small number of inter-residue cross peaks in the NOESY spectra, it was concluded that QK10A is essentially a random coil conformation; however, a grouping of NOEs, dNN (i, i+1), dαN (i, i+3) and dαβ (i, i+3) supported the presence of a residual helical structure in central region of peptide (Figure III.27). This is also confirmed by CSI analysis. The final input file for the CYANA structure calculation software contained 83 meaningful distance constraints (32 intra-residue, 30short- and 20 medium-range) and 72 angle constraints.

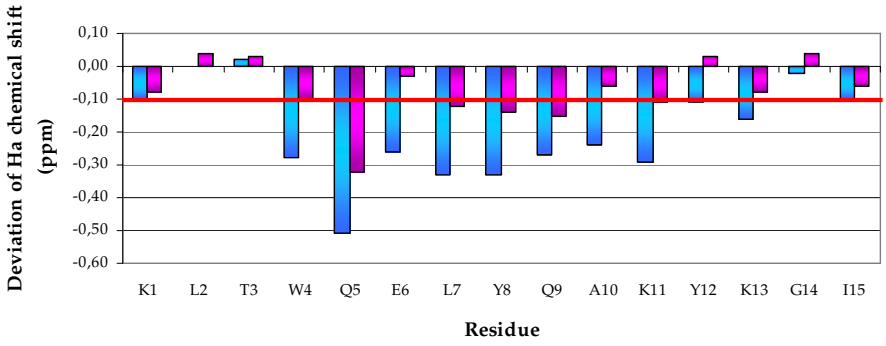

Figure III.27: Hα chemical shift deviations from random coil values (ΔδHα) of the QK10A at 298K (magenta) and at 313 K (cyana). The continuous line represents the CSI threshold for amino acid in helical conformation.

NMR Structural Statics of the QK10A at 313 K	
NMR Distance Constraints	
Distance Constraints	140
Total NOE	83
Intra residue	32
Short range	30
Medium range	20
Total Diehedral Angles Restraints	72
Φ (Phi)	36
Ψ (Psi)	36
Avarage Pairwise RMSD (Å)	
Heavy	0.90
Backbone	0.17

As expected, the final NMR structure of QK10A at 313 K (Figure III.28) showed that the C-terminal region is completely in random coil conformation while the N-terminal region, due to the lack of medium range cross peaks and the large ^{3}JHNHα, is not more in helix conformation, but presents an open turn. Instead, in the central region is still present a characteristic helical turn.

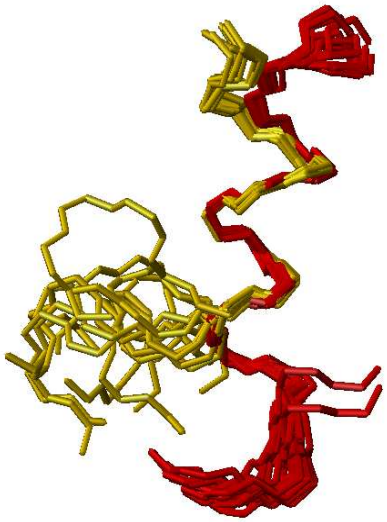

Figure III.28: Converged NMR-derived structures of the residues 4-12 of QK10A peptide at 298 K (red) and 313 K (yellow). RMSD for the mean structures is 0.32 Å (±0.02).

III.3 Discussions

α-Helical sequences of amino acids make up to the 30% of protein structure, play crucial roles in stabilizing tertiary structure, and mediate important biological processes through interactions with proteins, DNA, or RNA (Cochran, 2000). However, short peptides corresponding to these regions are not stable α-helices in water (Zimm, B. H. and Bragg, 1959). In order to develop α-helix mimetics to be to specifically recognize protein target helical peptides need to be stabilized.

Recently, it was described a short peptide, QK, adopting a well-defined helical conformation (D'Andrea et al., 2005). We would evaluate if peptide QK has the potential to be useful helical scaffold. To this aim we undertook a biophysical characterization of QK. To gain a more detailed understanding of the molecular basis of the QK helical structure we carried out a QK stability characterization. QK showed an unusual thermal stability up to 368 K that prevented us to obtain a structural depiction of the folding/unfolding pathway of the helix. To identify the structural determinants of QK thermal stability, we performed a MD analysis which highlighted the contribution to the peptide stability of the N-terminal region and the hydrofobic contact between Leu7 and Leu10. CD and NMR studies confirmed the theoretical results. Successively, we have obtained more information about the structural determinants of QK stability through thermal stability analysis of QK10A mutant.

In particular, on the basis of the temperature dependencies of QK10A H_α chemical shifts we show that the first phase of the thermal helix unfolding, having T_m of 308-309K, involves principally the C-terminal end (residues 10, 11 and 12) and in a minor part also the N-terminal end (residues 4). A second phase of the transition comprises the central helical turn with T_m of 325-327K. The determination of high resolution QK10A conformational preferences in water at 313 K permits to identify, at atomic resolution, one intermediate of the folding-unfolding pathway. MD simulations corroborate experimental observations, detecting a stable central helical turn, which represents the most probable site for the helix nucleation. The data here presented allow to draw a folding/unfolding picture for the small peptide QK compatible with the nucleation/propagation model. This study, besides contributing to the basic field of peptide helix folding, get insights into the design of stable helical peptides which could find applications as molecular scaffold to target protein-protein interactions.

III.4 Experimental Section

III.4.1 Materials and Methods

Protected N-Fmoc-amino acid derivatives, acetic anhydride, coupling reagents and Rink amide MBHA resin have been purchased from Novabiochem. DIPEA is provided from Applied Biosystem. All other reagents are commercially available from Sigma-Aldric and all solvents are commercially available from LabScan.

Preparative purification of synthetic peptides has been carried out on a Shimadzu LC-8A, equipped with a SPD-M10 AV detector. The column used was the Phenomex Jupiter Proteo 90Å 10µ C12 250 x 10.00 mm. Analytical characterization of synthetic peptides was performed on an LC-MS apparatus LCQ Deca XP MAX (Thermo Scientific) spectrometer equipped with an ESI source and ion trap mass completed to a Surveyor HPLC system (with photo diode array detector). The column used was the Phenomenex Jupiter Proteo C12 4µ 90Å, 50 x 2.00 mm. UV-VIS spectra were performed on a Jasco V-550 UV-VIS spectrophotometer, using 1 cm length quartz cell (Hellma).

Far-UV circular dichroism spectra recorded on a Jasco J-810 spectropolarimeter, equipped with a PTC-423S/15 Peltier temperature controller, using a 0.1 cm quartz cell (Hellma) in the range 190-260 nm.

To obtain all peptide structures decribed bi-dimensional homonuclear experiments, such as TOCSY and NOESY spectra, were recorded at 298 K on a Varian Unity Inova 600 MHz spectrometer, equipped with a cold-probe.

Thermal unfolding of peptides was studied by 2D [^{1}H, ^{1}H] TOCSY acquired at 298, 303, 308, 313, 318, 323, 328, 333, 338 and 343 K, acquiring all the spectra consecutively. NMR experiments were carried out on a Varian Inova 400 MHz spectrometer, where the probe temperature was regularly calibrated by using methanol and ethylenglycol.

III.4.2 Peptide synthesis

All peptides were synthesized by solid phase peptide synthesis as C-terminally amidated and N-terminally acetylated derivatives following standard Fmoc chemistry protocol.

A Rink-amide MBHA resin (substitution 0.53 mmol/g).

Cycles of coupling for each amino acid involves the following steps:

- Deprotection of the N-terminal function (2 step by 10 min) with a solution to a 30% Piperidine in DMF;
- Coupling of 10 equivalents of Fmoc-AA, 9.9 equivalent of HOBT/HBTU (0.45 M solution in DMF) and 20 equivalents of DIPEA (2 M solution in NMP) compared with 0.1 mmol scale synthesis. The reaction time for each pair was 45 min;
- Acetylation two treatments of 5 min with the appropriate volume of a solution of Acetic anydride (2M)/DIPEA (0.55M)/ HOBt (0.06M).

After each module were performed three washes with DMF for 1min. Cleavage from the solid support was performed by treatment with a TFA/TIS/water (95:2,5:2,5, v/v/v) mixture for 3 hours at R.T..
Crude peptides were precipitated in cold diethyl-ether, dissolved in a water/acetonitrile (9:1, v/v) mixture and lyophilized. Peptide purification was carried out by RP-HPLC applying a linear gradient of acetonitrile (0.1% TFA) from 20% to 80% in 30 min. at 20 mL/min. Peptide purity and identify were confirmed by LC-MS system previously described; with A gradient of solvent B (0.05% TFA in CH_3CN) from solvent A (0.08% TFA in H_2O) of 5% to 70% was applied over 30 min.

III.4.3 CD analysis

Spectra were acquired using a band width of 1 nm, a response of 8 sec, a data pitch of 0.1 nm and a scanning speed of 10 nm/min. Each spectrum was the average of three scans with the background of the buffer solution subtracted. Spectra were performed at 20°C in 10 mM Sodium phosphate buffer at pH 7.1. CD data were expressed as mean residue ellipticity (θ). Spectra processing was obtained by using the Spectra Manager software. Thermal denaturation was performed in 10 mM Sodium phosphate buffer, pH 7.1. The temperature was increased from 5°C to 95°C at a rate of 30°C/h. The CD signals were acquired at 1°C intervals at 222 nm.

III.4.4 Nuclear Magnetic Resonance

2D [^{1}H,^{1}H]-TOCSY (Griesinger, C.; Otting, G.; Wuethrich, K.; Ernst, Richard R. Clean Journal of the American Chemical Society (1988), 110(23), 7870-2) experiments (1mM in aqueous solution at pH 5.5) were acquired at 298, 303, 308, 313, 318, 323, 328, 333, 338 and 343K, acquiring all the spectra consecutively. Each spectra contained 2048 data

points and 64 scans, and had a sweep width of 7000 Hz in both dimensions. Water suppression was achieved through Double Pulsed Field Gradient Spin Echo sequences. The diffusion-ordered NMR spectroscopy (DOSY) (Morris, Kevin F.; Johnson, Charles S., Jr. Journal of the American Chemical Society (1992), 114(8), 3139-41) was carried out at 298 K using a 5-mm triple resonance xyz-gradient probe. The strength of the 3 ms gradient pulses was incremented in 15 experiments, with a diffusion time of 100 ms and a longitudinal eddy currents delay of 5 ms. The water resonance was suppressed by low-power presaturation during the relaxation delay of 1.5 msec. Diffusion coefficient of QK is obtained by fitting peaks intensities for most well resolved peaks across the spectra, using equation $I=I_0 \exp[-\gamma^2 g^2 D\delta^2(\Delta-\delta/3)]$, where γ is the gyromagnetic ratio and g, δ, Δ are the amplitude, duration and separation of the single pair of gradient pulses, respectively. For each measurements, a large number of cross peaks of the spectrum was used for the fitting and an average value was utilized.

III.4.5 Molecular Dynamics Simulations

Multiple Molecular Dynamics MD simulations of the QK peptide were started from five different α-helical NMR-derived models, after fitting to NOE constraints. The five starting structures correspond to Models 1, 5, 10, 15 and 20 in the NMR derived structure ensemble. Each model was used as a starting structure for four simulations at the four different temperatures of 300K, 320K, 340K, and 380K. Each 300K MD simulation was 200ns long, while all the higher temperature ones were 100ns long. In total 20 MD trajectories were generated for a total sampling time of approximately 2.4 microseconds. Each simulation was labelled with the first number indicating the NMR model and the second one indicating the temperature at which the simulation was run: for instance, simulation 5_340 indicates the simulation starting from model 5 from the NMR ensemble and run at 340K. Moreover, in order to shed light on the folding mechanism, four different simulations with lengths ranging from 50 to 100ns, at 350K, were run from a completely extended structure of the polypeptide. In each simulation, the isolated peptide structure was first solvated with water in a periodic truncated octahedron, large enough to contain the peptide and 0.9 nm of solvent on all sides. The protonation and charge states of the sidechains of the ligand and of the receptor were chosen to be consistent with the solution conditions of the experiments: NH groups were considered with a +1 charge and Carboxylic groups were considered to bear a -1 charge.
The system resulted to have a total charge of +2. All solvent molecules within 0.15 nm of any peptide atom were removed. Two Cl- counterions were added to the system. Different sets of initial velocities obtained from a Maxwellian velocity distribution at the desired temperatures were used to start production runs. In each case, the system was initially energy minimized with a steepest descent method for 1000 steps. In all simulations the

temperature was maintained close to the intended value of 300K by weak coupling to an external temperature bath (*Berendsen, H.J.C.; Postma, J.P.M.; van Gunsteren, W.F.; Di Nola, A.; Haak, J.R. J. Chem. Phys. (1984), 81; 3684-3690*) with a coupling constant of 0.1 ps. The QK peptide and the rest of the system were coupled separately to the temperature bath. The AMBER force field (*Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, R. I.; Merz, K.M.Jr.; Ferguson, D. M.; Spellmeyer, D. C. ; Fox, T.; Caldwell, J. W.; Kollman, P. A.; J. Am. Chem. Soc. (1995), 117; 5179-5197*) was used. The TIP3P water model (*Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.; Impey, R.W.; Klein, M.L.; J. Chem. Phys. (1983) 79; 926-935*) was used. The LINCS algorithm (*Hess, B.; Bekker, H.; Fraaije, J. G. E. M.; Berendsen, H. J. C. A J. Comp. Chem. (1997) 18; 1463-1472*) was used to constrain all bond lengths.
A dielectric permittivity, ε=1, and a time step of 2 fs were used. A cut-off was used for the calculation of the non-bonded Van der Waals interactions.
The cutoff radius was set to 0.9 nm. The calculation of electrostatic forces utilized the PME implementation of the Ewald summation method. In each simulation, the density of the system was adjusted performing the first equilibration runs at NPT condition by weak coupling to a bath of constant pressure (P_0 = 1bar, coupling time τ_P= 0.5 ps) (*Berendsen, H.J.C.; Postma, J.P.M.; van Gunsteren, W.F.; Di Nola, A.; Haak, J.R. J. Chem. Phys., (1984) 81; 3684-3690*). All simulations were equilibrated by 50 ps of MD runs with position restraints on the protein and ligand to allow relaxation of the solvent molecules.
These first equilibration runs were followed by other 50 ps runs without position restraints on the solute. The production runs using NVT conditions, after equilibration, covered the simulation lengths discussed at the beginning of this paragraph. All the MD runs and the analysis of the trajectories were performed using the GROMACS software package (*Lindahl, E.; Hess, B.; van der Spoel, D.; J. Mol. Mod. (2001) 7; 306-317*).
Configurations of the peptide were saved every 4ps for subsequent statistical analysis.
Conformational cluster analysis of the trajectories was performed using the method described in Daura et al. (*X. Daura; K. Gademann; B. Jaun; D. Seebach; W. F. van Gunsteren; A. E. Mark Angew. Chemie Intl. Ed.; (1999) 38; 236-240*): count number of neighbors using a cut-off of 0.15 nm RMSD between the optimal backbone superposition of different structures, take structure with largest number of neighbors with all its neighbors as cluster and eliminate it from the pool of clusters. This procedure is repeated for the remaining structures in the pool. This procedure was applied separately to each trajectory. To obtain a more global view on the peptide behavior, the clustering analysis was performed on a single ensemble obtained from all the structures obtained from all the simulations. The representative structure of the most populated cluster, representative of the most visited structures in the MD simulations, was used as a template for defining which residues to mutate to alanine.
The trajectories were also analyzed in terms of the time evolution of root mean square deviation (rmsd) after backbone-backbone superposition, flexibility (root mean square fluctuations, rmsf), secondary structure evolution, and in terms of stabilizing

interactions. Trajectories were also analyzed to define the principal components of peptide motions. Finally, the first 20 ns of each refolding trajectory were analyzed in terms of the percentage of time spent by each residue in a helical conformation, to define the presence of a preferred folding direction.

III.5 References

- Becerril J., A. D. Hamilton, Angew. Chem. 2007, 119, 4555; Angew. Chem. Int. Ed. 2007, 46, 4471.
- Biros S. M., L. Moisan, E. Mann, A. Carella, D. Zhai, J. C. Reed, J. Rebek, Bioorg. Med. Chem. Lett. 2007, 17, 4641.
- Bracken, C.; Guylas, J.; Taylor, J. W.; Baum, J. J. Am. Chem. Soc. 1994, 116, 6431-32
- Cochran, A. G. Chem. Biol. 2000, 7, R85
- D'Andrea LD, Iaccarino G, Fattorusso R, Sorriento D, Carannante C, Capasso D, Trimarco B, Pedone C. Proc Natl Acad Sci U S A. 2005 Oct 4;102(40):14215-20. Epub 2005 Sep 26.
- Davis J. M., L. K. Tsou and A.D. Hamilton, Chem. Soc. ReIV., 2007, 36, 326
- Davis, J. M.; Truong, A.; Hamilton, A. D. Org. Lett. 2005, 7, 5405–5408
- Dudar GK, D'Andrea LD, Di Stasi R, Pedone C, Wallace JL. Am J Physiol Gastrointest Liver Physiol. 2008 Aug;295(2):G374-81. Epub 2008 Jun 26.
- Ernst, J. T.; Becerril, J.; Park, H. S.; Yin, H.; Hamilton, A. D. Angew. Chem., Int. Ed. 2003, 42, 535–539.
- Fairlie, D. P.; West, M. L.; Wong, A. K. Curr. Med. Chem. 1998, 5, 29–62
- Gamer, J.; Harding, M. M. Org. Biomol. Chem. 2007, 5, 3577–3585.
- Goodman C. M., S. Choi, S. Shandler, W. F. DeGrado, Nat. Chem. Biol. 2007, 3, 252.
- Hara T., S. R. Durell, M. C. Myers, D. H. Appella, J. Am. Chem. Soc. 2006, 128, 1995
- Henchey, L. K.; Jochim, A. J.; Arora, P. S. Curr. Opin. Chem. Biol. 2008, 12, 692–697
- Kelso M. J. and D. P. Fairlie, in Mol. Pathomechanisms and New Trends in Drug Research, ed. G. Keri and I. Toth, Taylor and Francis, London, 2003, pp. 578–598
- Klug A., FEBS Lett., 2005, 579, 892.
- Kritzer J. A., O. M. Stephens, D. A. Guarracino, S. K. Reznik, A. Schepartz, Bioorg. Med. Chem. 2005, 13, 11.
- Kutchukian, P. S.; Yang, J. S.; Verdine, G. L.; Shakhnovich, E. I.; *J. Am. Chem. Soc.* 2009, **131**, 4622-4627
- Lavery R., Q. ReIV. Biophys., 2005, 38, 339.
- Li C, Liu M, Monbo J, Zou G, Li C, Yuan W, Zella D, Lu WY, Lu W. J Am Chem Soc. 2008 Oct 15;130(41):13546-8.
- Li C, Pazgier M, Liu M, Lu WY, Lu W. Angew Chem Int Ed Engl. 2009;48(46):8712-5.
- Moellering RE, Cornejo M, Davis TN, Del Bianco C, Aster JC, Blacklow SC, Kung AL, Gilliland DG, Verdine GL, Bradner JE. Nature. 2009 Nov 12;462(7270):182-8.
- Rodriguez JM, Nevola L, Ross NT, Lee GI, Hamilton AD. Chembiochem. 2009 Mar 23;10(5):829-33.
- Sadowsky J. D., M. A. Schmitt, H. S. Lee, N. Umezawa, S. M. Wang, Y. Tomita, S. H. Gellman, J. Am. Chem. Soc. 2005, 127, 11966.

- Santulli G, Ciccarelli M, Palumbo G, Campanile A, Galasso G, Ziaco B, Altobelli GG, Cimini V, Piscione F, D'Andrea LD, Pedone C, Trimarco B, Iaccarino G. J Transl Med. 2009 Jun 8;7:41.
- Shaginian A., L.R. Whitby, S. Hong, I. Hwang, B. Farooqi, M. Searcey, J. Chen, P.K. Vogt and D.L. Boger, J.A.C.S. 131 (15) (2009), pp. 5564–5572.
- Schafmeister C.E, Po J., and G.L. Verdine *J. Am. Chem. Soc.* 2000, *122,* 5891-5892
- Yin H., A. D. Hamilton, Angew. Chem. 2005, 117, 4200; Angew. Chem. Int. Ed. 2005, 44, 4130. **A**
- Yin H., G. Lee, H. S. Park, G. A. Payne, J. M. Rodriguez, S. M. Sebti, A. D.Hamilton, Angew. Chem. 2005, 117, 2764; Angew. Chem. Int. Ed. 2005, 44, 2704. **B**
- Yin, H.; Hamilton, A. D. Bioorg. Med. Chem. Lett. 2004, 14, 1375–1379.
- Yin, H.; Sedey, K. A.; Rodriguez, J. M.; Wang, H.-G.; Sebti, S. M.; Hamilton, A. D J. Am. Chem. Soc. 2005, 127, 5463–5468. **C**
- Walensky, L. D.; Kung, A. L.; Escher, I.; Malia, T. J.; Barbuto, S.; Wright, R.; Wagner, G.; Verdine, G. L.; Korsmeyer, S. J.*Science*. 2004 September 3; 305(5689): 1466–1470.
- Zimm, B. H.; Bragg, J. K. J. Chem. Phys. 1959, 31, 526-535

III.6 Appendix

TABLE III.I

Temperature dependence of QK backbone protons chemical shift

Residue	*Atom*	*Temperature (K)*									
		298	**303**	**308**	**313**	**318**	**323**	**328**	**333**	**338**	**343**
K1	**Hα**	4,24	4,242	4,245	4,248	4,252	4,254	4,257	4,259	4,26	4,262
	H_N	8,22	8,218	8,215	8,214	8,21	8,208	8,202	8,198	8,189	8,184
L2	**Hα**	4,22	4,241	4,264	4,268	4,279	4,284	4,295	4,324	4,355	4,36
	H_N	8,19	8,187	8,183	8,18	8,175	8,168	8,161	8,154	8,149	8,146
T3	**Hα**	4,355	4,355	4,355	4,356	4,358	4,36	4,36	4,36	4,36	4,361
	H_N	8,47	8,464	8,459	8,453	8,447	8,437	8,43	8,425	8,416	8,412
W4	**Hα**	4,382	4,394	4,404	4,415	4,427	4,438	4,445	4,457	4,461	4,461
	H_N	8,5	8,466	8,375	8,312	8,233	8,192	8,137	8,109	8,023	7,978
Q5	**Hα**	3,793	3,816	3,836	3,85	3,87	3,885	3,898	3,915	3,923	3,935
	H_N	8,147	8,137	8,064	8,029	7,974	7,96	7,928	7,924	7,903	7,839
E6	**Hα**	4,001	4,01	4,013	4,021	4,025	4,036	4,037	4,046	4,046	4,047
	H_N	7,773	7,763	7,745	7,734	7,741	7,741	7,735	7,733	7,733	7,688
L7	**Hα**	4,001	4,01	4,013	4,021	4,025	4,036	4,037	4,046	4,046	4,047
	H_N	8,129	8,115	8,04	7,996	7,932	7,916	7,88	7,871	7,862	7,773
Y8	**Hα**	4,194	4,208	4,219	4,234	4,245	4,261	4,269	4,278	4,285	4,285
	H_N	8,096	8,095	8,034	8,003	7,946	7,936	7,905	7,901	7,83	7,803
Q9	**Hα**	4,066	4,072	4,074	4,076	4,085	4,097	4,097	4,097	4,1	4,1
	H_N	7,843	7,837	7,812	7,797	7,756	7,768	7,767	7,769	7,755	7,737
L10	**Hα**	4,11	4,121	4,121	4,123	4,131	4,14	4,142	4,147	4,147	4,147
	H_N	7,784	7,83	7,771	7,749	7,686	7,728	7,725	7,723	7,715	7,66
K11	**Hα**	4,035	4,045	4,051	4,057	4,057	4,061	4,066	4,081	4,094	4,107
	H_N	7,782	7,809	7,774	7,763	7,727	7,732	7,738	7,743	7,74	7,722
Y12	**Hα**	4,445	4,458	4,46	4,461	4,475	4,484	4,485	4,486	4,486	4,486
	H_N	7,851	7,861	7,807	7,782	7,736	7,73	7,704	7,698	7,695	7,626
K13	**Hα**	4,212	4,217	4,225	4,229	4,236	4,24	4,249	4,258	4,265	4,278
	H_N	7,95	7,947	7,933	7,925	7,904	7,865	7,843	7,75	7,72	7,615
G14	**Hα**	3,89	3,892	3,897	3,898	3,902	3,905	3,907	3,907	3,91	3,911
	H_N	7,931	7,928	7,907	7,9	7,897	7,855	7,832	7,75	7,72	7,61
I15	**Hα**	4,128	4,128	4,128	4,128	4,128	4,128	4,129	4,13	4,13	4,13
	H_N	7,854	7,85	7,778	7,696	7,625	7,551	7,458	7,387	7,329	7,284

Table III.2:
^{1}H assignment of the QK1 -12

Residue	*Atom*	δ (ppm)	$^{3}J_{HNH\alpha}$ (Hz)
K1	HA	4.196	6.9
K1	QD	1.611	
K1	QE	2.944	
L2	H	8.197	6.5
L2	HA	4.463	
L2	HB2	1.551	
L2	HB3	1.433	
L2	QD1	0.792	
T3	H	8.151	9.0
T3	HA	4.396	
T3	HB	4.372	
T3	QG2	1.203	
W4	H	8.508	5.5
W4	HA	4.381	
W4	HB2	3.371	
W4	HB3	3.256	
W4	HD1	7.276	
W4	HE3	7.436	
W4	HE1	10.081	
W4	HZ3	7.008	
W4	HZ2	7.405	
W4	HH2	7.067	
Q5	H	8.169	5.1
Q5	HA	3.786	
Q5	HB2	1.791	
Q5	HB3	1.700	
Q5	QG	2.101	
Q5	HE21	6.670	
Q5	HE22	7.323	
E6	H	7.744	5.3
E6	HA	4.006	
E6	HB2	2.089	
E6	HB3	1.946	
E6	QG	2.319	
L7	H	8.147	5.3
L7	HA	4.002	
L7	HB2	1.607	
L7	HB3	1.485	
L7	QD1	0.793	
L7	HG	1.373	
Y8	H	8.109	6.1
Y8	HA	4.186	
Y8	HB2	2.979	
Y8	HB3	2.730	
Y8	QD	6.999	

Y8	QE	6.717	
Q9	H	7.897	5.9
Q9	HA	4.049	
Q9	QB	2.047	
Q9	QG	2.367	
Q9	HE21	6.822	
Q9	HE22	7.447	
L10	H	7.803	6.0
L10	HA	4.093	
L10	HB2	1.644	
L10	HB3	1.483	
L10	HG	1.089	
L10	QD1	0.799	
L10	QD2	0.854	
K11	H	7.740	5.6
K11	HA	4.011	
K11	HB2	1.479	
K11	HB3	1.384	
K11	QD	1.449	
K11	QE	2.807	
Y12	H	7.915	7.0
Y12	HA	4.446	
Y12	HB2	3.013	
Y12	HB3	2.493	
Y12	QD	6.916	
Y12	QE	6.709	

Table III.3
Temperature dependence of QK 1-12 backbone protons chemical shift

Residue	*Atom*	*Temperature (K)*									
		298	**303**	**308**	**313**	**318**	**323**	**328**	**333**	**338**	**343**
K1	**Hα**	4,207	4,223	4,229	4,313	4,312	4,369	4,388	4,400	4,402	4,400
	H_N	8,200	8,136	8,094	8,061	7,953	7,874	7,822	7,777	7,701	7,689
L2	**Hα**	4,466	4,460	4,449	4,450	4,452	4,454	4,556	4,558	4,558	4,550
	H_N	8,203	8,131	8,096	8,132	7,968	7,893	7,856	7,822	7,791	7,788
T3	**Hα**	4,316	4,335	4,332	4,337	4,340	4,347	4,353	4,355	4,357	4,358
	H_N	8,129	8,238	8,187	8,180	8,177	8,173	8,167	8,164	8,160	8,161
W4	**Hα**	4,384	4,414	4,420	4,516	4,552	4,562	4,601	4,605	4,611	4,611
	H_N	8,515	8,409	8,343	8,273	8,146	8,043	8,040	8,036	8,033	8,025
Q5	**Hα**	3,786	3,833	3,833	3,849	3,886	3,912	3,918	4,010	4,013	4,019
	H_N	8,176	8,082	8,039	8,000	7,902	7,830	7,804	7,774	7,768	7,762
E6	**Hα**	4,010	4,011	4,032	4,040	4,045	4,048	4,050	4,056	4,058	4,060
	H_N	7,744	7,743	7,739	7,729	7,710	7,694	7,677	7,601	7,592	7,588
L7	**Hα**	4,007	4,006	4,011	4,016	4,019	4,021	4,026	4,027	4,029	4,031
	H_N	8,156	8,085	8,036	7,990	7,874	7,803	7,753	7,742	7,735	7,731
Y8	**Hα**	4,191	4,213	4,227	4,229	4,255	4,259	4,260	4,261	4,262	4,300
	H_N	8,109	8,063	8,022	7,982	7,882	7,874	7,869	7,862	7,856	7,848
Q9	**Hα**	4,050	4,057	4,063	4,067	4,068	4,069	4,051	4,058	4,061	4,068
	H_N	7,902	7,857	7,835	7,811	7,731	7,616	7,582	7,577	7,503	7,488
L10	**Hα**	4,098	4,023	4,025	4,031	4,031	4,048	4,052	4,057	4,061	4,061
	H_N	7,809	7,783	7,763	7,745	7,667	7,611	7,602	7,588	7,569	7,555
K11	**Hα**	4,012	4,014	4,019	4,029	4,030	4,048	4,052	4,058	4,065	4,066
	H_N	7,747	7,738	7,732	7,733	7,661	7,651	7,647	7,637	7,630	7,626
Y12	**Hα**	4,012	4,014	4,019	4,029	4,030	4,048	4,052	4,058	4,065	4,066
	H_N	7,920	7,870	7,840	7,811	7,723	7,651	7,641	7,638	7,627	4,618

Table III.4
^{1}H Assignment of QK10A peptide at 298 K

Residue	Atom	δ (ppm)	$^{3}J_{HNH\alpha}$ (Hz)
K1	H	8.132	7.02
K1	HA	4.256	
L2	H	8.211	7.01
L2	HA	4.371	
L2	QB	1.412	
L2	HG	1.310	
L2	QD1	0.893	
T3	H	8.044	6.00
T3	HA	4.375	
T3	HB	4.199	
T3	QG2	1.098	
W4	H	8.310	7.05
W4	HA	4.414	
W4	HB2	3.238	
W4	HB3	3.133	
W4	HD1	7.206	
W4	HE3	7.425	
W4	HE1	10.027	
W4	HZ3	6.926	
W4	HZ2	7.291	
W4	HH2	7.007	
Q5	H	8.035	-
Q5	HA	3.856	
Q5	QB	1.653	
Q5	QG	1.991	
Q5	HE21	6.735	
Q5	HE22	7.366	
E6	H	7.767	7.07
E6	HA	4.030	
E6	HB2	1.908	
E6	HB3	1.827	
E6	QG	2.198	
L7	H	7.706	6.05

L7	HA	4.052	
L7	HB2	1.354	
L7	HB3	1.175	
L7	HG	1.428	
L7	QD1	0.863	
L7	QD2	0.883	
Y8	H	7.993	6.05
Y8	HA	4.280	
Y8	HB2	2.818	
Y8	HB3	2.598	
Y8	QD	6.951	
Y8	QE	6.698	
Q9	H	7.933	6.08
Q9	HA	4.107	
Q9	HB2	1.954	
Q9	HB3	1.851	
Q9	QG	2.255	
Q9	HE21	6.632	
Q9	HE22	7.237	
A10	H	7.897	-
A10	HA	4.129	
A10	QB	1.486	
K11	H	7.933	6.06
K11	HA	4.228	
K11	QB	1.327	
K11	HG2	1.204	
K11	HG3	1.095	
K11	QD	1.506	
K11	QE	2.795	
Y12	H	7.829	8.02
Y12	HA	4.476	
Y12	HB2	2.872	
Y12	HB3	2.672	
Y12	QD	6.971	
Y12	QE	6.683	
K13	H	7.983	9.00

K13	HA	4.228	
K13	HB2	1.671	
K13	HB3	1.636	
K13	QD	1.535	
K13	QE	2.868	
G14	H	7.826	7.07
G14	QA	3.886	
I15	H	7.850	-
I15	HA	4.130	
I15	HB	1.461	
I15	HG12	1.275	
I15	HG13	1.094	
I15	QD1	0.771	

Tab III.5
QK10A temperature dependence of the Ha protons chemical shifts

Temperature (K)

Residue	288	293	298	303	308	313	318	323	328	333	338	343
K1	4.256	4.256	4.256	4.256	4.256	4.282	4.297	4.301	4.347	4.354	4.355	4.355
L2	4.371	4.379	4.382	4.394	4.401	4.425	4.438	4.444	4.450	4.459	4.465	4.476
T3	4.363	4.370	4.375	4.378	4.381	4.380	4.385	4.378	4.375	4.364	4.365	4.365
W4	4.395	4.395	4.414	4.439	4.525	4.597	4.646	4.665	4.668	4.671	4.673	4.672
Q5	3.847	3.849	3.856	3.861	3.976	4.045	4.170	4.257	4.334	4.339	4.340	4.342
E6	4.029	4.031	4.030	4.038	4.124	4.264	4.274	4.287	4.347	4.410	4.412	4.412
L7	4.044	4.044	4.052	4.056	4.146	4.255	4.261	4.263	4.340	4.347	4.348	4.350
Y8	4.275	4.278	4.280	4.288	4.397	4.462	4.483	4.516	4.606	4.645	4.648	4.651
Q9	4.097	4.099	4.107	4.109	4.200	4.219	4.230	4.309	4.315	4.327	4.331	4.331
A10	4.089	4.111	4.129	4.131	4.191	4.290	4.343	4.348	4.355	4.355	4.355	4.355
K11	4.050	4.050	4.063	4.077	4.176	4.251	4.289	4.308	4.315	4.329	4.329	4.329
Y12	4.469	4.469	4.476	4.476	4.576	4.634	4.662	4.671	4.685	4.685	4.686	4.700
K13	4.180	4.198	4.228	4.257	4.279	4.284	4.297	4.304	4.310	4.310	4.310	4.310
G14	3.821	3.859	3.886	3.896	3.951	4.010	4.025	4.033	4.036	4.040	4.040	4.040
I15	4.130	4.130	4.130	4.143	4.150	4.171	4.170	4.179	4.202	4.228	4.234	4.235

Table III.6
^{1}H Assignment of QK10A
Peptide at 313 K

Residue	Atom	δ (ppm)	$^{3}J_{HNH\alpha}$ (Hz)
K1	H	8.155	-
K1	HA	4.329	
L2	H	8.261	8.02
L2	HA	4.507	
L2	QB	1.641	
L2	HG	1.473	
L2	QD1	-	
T3	H	8.074	8.00
T3	HA	4.380	
T3	HB	4.402	
T3	QG2	1.284	
W4	H	8.305	8.09
W4	HA	4.597	
W4	HB2	3.397	
W4	HB3	3.314	
W4	HD1	7.333	
W4	HE3	7.545	
W4	HE1	10.102	
W4	HZ3	7.080	
W4	HZ2	7.447	
W4	HH2	7.153	
Q5	H	8.121	-
Q5	HA	4.045	
Q5	QB	1.971	
Q5	QG	2.163	
Q5	HE21	6.824	
Q5	HE22	7.451	
E6	H	7.907	8.06
E6	HA	4.264	
E6	HB2	2.034	
E6	HB3	1.940	
E6	QG	2.287	
L7	H	8.041	6.06

L7	HA	4.255	
L7	HB2	1.589	
L7	HB3	1.324	
L7	HG	1.569	
L7	QD1	0.887	
L7	QD2	-	
Y8	H	8.051	6.05
Y8	HA	4.462	
Y8	HB2	3.060	
Y8	HB3	2.899	
Y8	QD	7.052	
Y8	QE	6.788	
Q9	H	8.000	6.09
Q9	HA	4.290	
Q9	HB2	2.076	
Q9	HB3	2.011	
Q9	QG	2.425	
Q9	HE21	6.742	
Q9	HE22	7.331	
A10	H	8.009	-
A10	HA	4.290	
A10	QB	1.379	
K11	H	7.885	6.06
K11	HA	4.251	
K11	QB	1.480	
K11	HG2	1.431	
K11	HG3	1.206	
K11	QD	1.986	
K11	QE	2.989	
Y12	H	7.978	8.05
Y12	HA	4.634	
Y12	HB2	3.102	
Y12	HB3	2.933	
Y12	QD	7.076	
Y12	QE	6.791	
K13	H	7.964	9.00

K13	HA	4.284	
K13	HB2	1.643	
K13	HB3	1.565	
K13	QE	3.203	
K13	QE	-	
G14	H	7.932	7.07
G14	QA	4.010	
I15	H	7.882	9.02
I15	HA	4.171	
I15	HB	1.940	
I15	HG12	1.439	
I15	HG13	1.172	
I15	QD1	-	

IV APPLICATION OF HELICAL SCAFFOLD: DESIGN OF VEGF RECEPTOR BINDERS

APPLICATION OF HELICAL SCAFFOLD: DESIGN OF VEGF RECEPTOR BINDERS

IV.1 Introduction

There are important classes of proteins whose role lies in the tight complex formation with other polypeptides and whose architecture is mainly determined by rigid secondary structure. In these cases loop regions play a lesser role in the binding function but often adopt regular turn structures and as such ensure the dense packing of secondary structural elements. In this manner, several a-helices or β-strands are brought together and side chains protruding from two or more of these elements form the interface for complex formation with the target protein.

The *de novo* design of a peptide sequence adopting a well-defined secondary structure includes a first step of introducing residues in order to stabilize the secondary structure and a second in which you enter the residues responsible for interaction in the correct spatial orientation. An alternative approaches is grafting the interacting residues on protein scaffolds that have the desired secondary structure exposed to the solvent. In this case the protein scaffold serve to structurally fix the grafted peptide segment and, indeed, achievable affinities seem to be higher if compared to the corresponding conformationally flexible peptide. Another advantage of the scaffolding approach is the possibility to obtain reagents with high physicochemical and possibly proteolytic stability. In fact, it is reasonably assumed that the beneficial properties of a robust wild-type scaffold protein will be largely retained in its engineered descendants. Such stability aspects may not only include the capability of maintaining a functional conformation during various types of 'environmental stress' but also relate to the ability of reversibly regaining the native structure after potential denaturation.

Several examples of scaffolds have been reported in the literature: natural scaffolds by animal toxins, which present α/β or β-structures cross-linked by several disulfide bridges, able to bind different ion channels with high affinity and specificity (Kellenberger, E., et al., 1999; Chong Li, et al., 2008); zinc fingers and Leu zippers (Domingues, H. et al., 1999), which have α/β or all-α structures, respectively, and represent structural modules of more complex protein ensembles that can bind specific DNA sequences. (S. E. Rutledge et al., 2003; T. L. Schneider, et al., 2005; J. A. Kritzer et al., 2006) and IgG domain a natural biomolecular scaffold for various applications in basic science and medicine.

Our intent was engineering an appropriate α-helical scaffold to develop a mini-protein as VEGF receptors binder.

IV.1.1 Angiogenesis and VEGF

Angiogenesis is the process of remodeling of the vascular tissue characterized by the branching out of a new blood vessel from a pre-existing one. The angiogenesis is particularly active during embryogenesis, while during adult life it is quiescent and limited to particular physiological phenomena. In the last years, the study of molecular mechanisms of angiogenesis has stirred renewed interest due to the recognition of the role played by angiogenesis in several pathologies of large social impact, such as tumors and cardiovascular diseases, but also to the pharmacological interest rising from the possibility of modulating this phenomenon (Carmeliet, 2003). Antibodies, peptides and small molecules targeting active endothelial cells (ECs) represent an innovative tool in therapeutic and diagnostic fields. In the process of angiogenesis, vascular endothelial growth factor (VEGF) is essential for growth, mitogenesis, and tube formation of ECs. VEGF binds to two tyrosine kinase receptors, fms-like tyrosine receptor (VEGFR1) and kinase insert domain containing receptor (VEGFR2), on the surface of ECs, thereby activating signal transduction and regulating physiological and pathological angiogenesis. Several VEGF structures have been reported: free (Muller, 1997a and 1997b), bound to peptide inhibitors (Wiesmann, C., 1998; Pan, B., 2002), to a neutralizing antibody (Muller, 1998) and to the domain 2 of the extracellular region of VEGFR1 receptor, VEGFR1D2, (Figure IV.1) (Wiesmann, C., 1997).

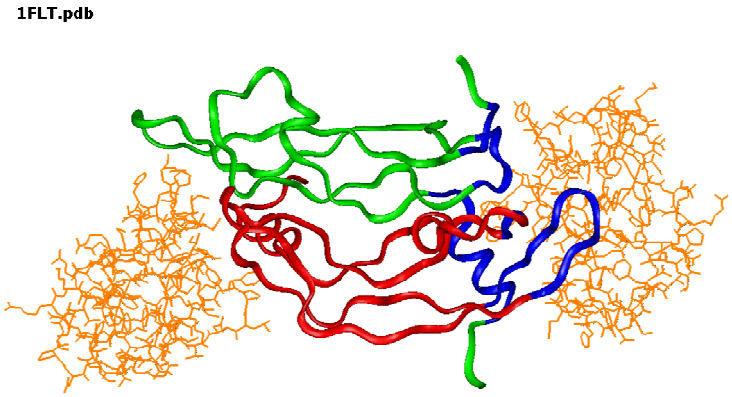

Figure IV.1: VEGF (green and red) and the domain 2 of its receptor VEGFR1 (orange); in blue the residues involved in the binding to the receptors.

VEGF is an antiparallel homodimer, covalently linked through two disulfide bonds. It is characterized by a cystine knot motif. The knot consists of two disulfide bridges, with a third disulfide bond passing through them. Domain deletion studies on VEGFR1 have shown that the ligand binding site resides within the first three domains. The second domain binds VEGF about 60-times less tight than wild type protein, but its removal

abolish the binding. VEGFR1D2 is a member of the immunoglobulin family. It consists of a sandwich formed by two sheets, one consisting of five strands and the other of three.
The overall structure of the VEGF/VEGFR1D2 complex possesses approximately a two-fold simmetry. The VEGF recognition interface is divided about 65% and 35% between both monomers. The analysis of structural and mutagenesis data allowed to identify residues involved in the binding to the receptors. They are distributed over a discontinuous surface which include residues from the Nterminal helix (17-25), the loop connecting strand β3 to β4 (61-66) and strand β7 (103-106) of one monomer, as well as residues from strand β2 (46-48) and from strand β5 and β6 together with the connecting turn (79-91) of the other monomer. The recognition interface is manly hydrophobic, except for the polar interaction between Arg224 (VEGFR1) and Asp63 (VEGF).
VEGFR2 and VEGFR1 share the same VEGF binding region, in fact 5 out of 7 most important VEGF binding residues are present in both interfaces (Muller 1997a; Muller, 1997b; Wiesmann, C., 1997).

IV.2 Results

IV.2.1 Molecular Design

The analysis of three-dimensional structure of the complex between VEGF/VEGFR1D2 showed that the VEGF N-terminal helix (residue 17-25) is involved in receptor recognition. The role of this region for interaction with the receptors was also demonstrated using a peptide, QK, mimicking the VEGF helix 17-25 (D'Andrea *et al.*, 2005). We decided to graft the interacting residues of VEGF helix and QK on a scaffold that has a helix exposed to the solvent order to make stable molecules capable of interacting with receptors of VEGF with potential applications in vivo.

The Avian pancreatic polypeptide (APP) was chosen as protein scaffold. It is very stable, small in size and also because it has already showed its ability to modulate protein-DNA (Zondlo and Schepartz, 1999) and protein-protein interactions. APP is a stable mini-protein of 36 amino acids (Blundell et al., 1981), which presents a helix of poly proline type II and an α-helix exposed to solvent, which interact to define a small hydrophobic core is also present at the interface between the two helices (Figure IV.2).

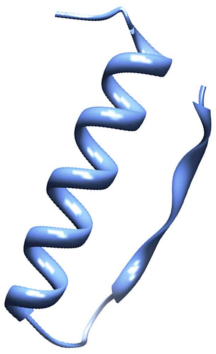

Figure IV.2: X-ray structure of the Avian Pancreatic Polypeptide (PDB ref).

Based on the structure of the complex VEGF/VEGFR-1D2 receptor we identified the VEGF residues those located in the region 17-25 at less than 4.5 Å from the receptor: Phe17, Met18, Tyr21 and Gln22 and Tyr25. The corrisponding residue on the peptide QK (see Figure IV.3) are Trp4, Gln5, Tyr8, Gln9 and Tyr12. They occupy one face of the helix and establish hydrophobic interactions with the receptor.

```
VEGF(17-25)      FMDVYQRSY
QK                KLTWQELYQLKYKGI
```

Figure IV.3: Sequence of VEGF helix 17-25 and QK peptide .

QK and VEGF helix have only two rounds of helix, while APP has four, so we have made all possible structural alignments the helices can align with follow residues of APP: Val14, Ile18 e Asn22.

In the last two cases occur steric encumbrance between the turn preceding the helix and the receptor (Figure IV.4).

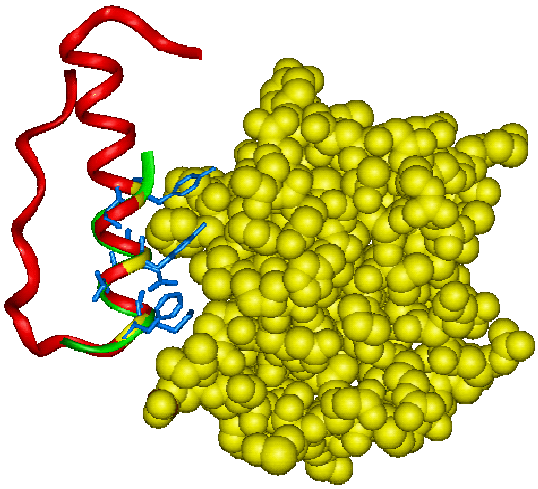

Figure IV.4: Overlapping of APP structure (red) with the VEGF helix 17-25 (green) complexed to the receptors (yellow); the interacting residue are highlighed in light blue.

Thus, the residues of the wild type sequence (APPwt) replaced are in the first mutated scaffold (APP1): V14F, E15M, I18Y, R19Q and N21Y; in second scaffold mutated (APP_QK): V14W, E15Q, I18Y, R19Q e N21Y (Figure IV.5).

```
APPwt          GPSQPTYPGD DAPVEDLIRF YNDLQQYLNV VTRHRY

VEGF(17-25)                  FMDVYQR SY
APP1           GPSQPTYPGD DAPFMDLYQF AYNLQQYLNV VTRHRY

QK                         KLTWQELYQL KYKGI
APP_QK         GPSQPTYPGD DAPWQDLYQF AYNLQQYLNV VTRHRY
```

Figure IV.5: Alignment of sequences between APP and two helix of VEGF and QK.

In addiction, the Tyr20 was mutated in Ala to avoid the interferences with bond to receptor since in positions 21 was insert the Tyr responsible of interactions with it. Ala was choice to not perturb the helix.

IV.2.2 Gene synthesis and cloning

The gene coding for the mini proteins, APPwt, APP1 and APP_QK, were synthesized starting from short oligonucleotide sequences (Figure IV.6).
The full length gene double strand was divided in four short oligonucleotide sequence (A, B, C, D) which were ligated, after hybridization, to afford two half gene (AB, CD), a final ligation reaction afford the final gene product. Color code refer to oligonucleotide sequence reported in paragraf IV.4.2.

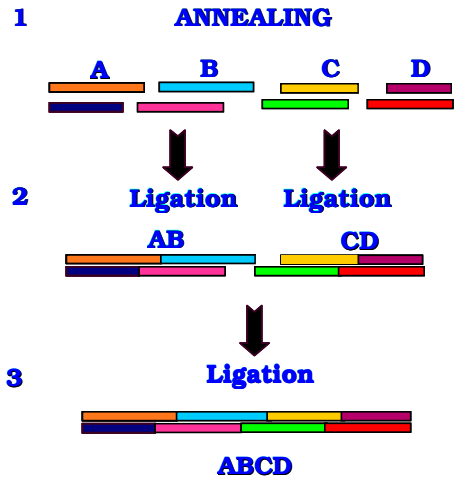

Figure IV.6: APP gene synthesis startegy;
the colours were congruent with the sequences reported in paragraph IV.4.2

In figure IV.7 is reported the agarose gel of the costruction of gene APPwt before last ligation; the full length APPwt gene purified and analyzed by electrophoresis on a agarose gel (Figure IV.8).

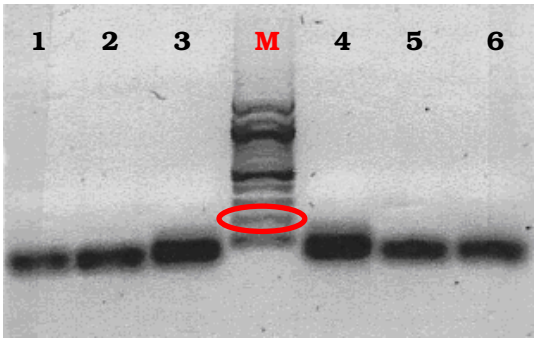

Figure IV.7: Partial synthesis of APPwt gene; in red marker band of 100 bp,
lane 1 annealing A (30bp), lane 2 annealing B (29bp), lane 3 ligase reactions AB (59bp),
lane 4 ligase reaction CD (59bp), lane 5 annealing C (29bp), lane 6 annealing D (30bp).

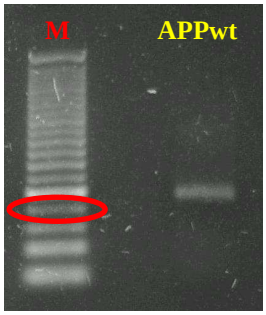

Figure IV.8: APPwt gene (118bp) purified, in red marker band of 100 bp

Succesively, it was amplified by PCR, product and ligated into the expression vectors pPROEXHTa between NcoI and XhoI restriction site.

The APP1 and APP_QK genes were synthetized in following the same procedures as APPwt. In figure IV.9 is reported the agarose gel of the costruction of genes before last

ligation; the full length APP1 and APP_QK genes purified and analyzed by electrophoresis on a agarose gel (Figure IV.10).

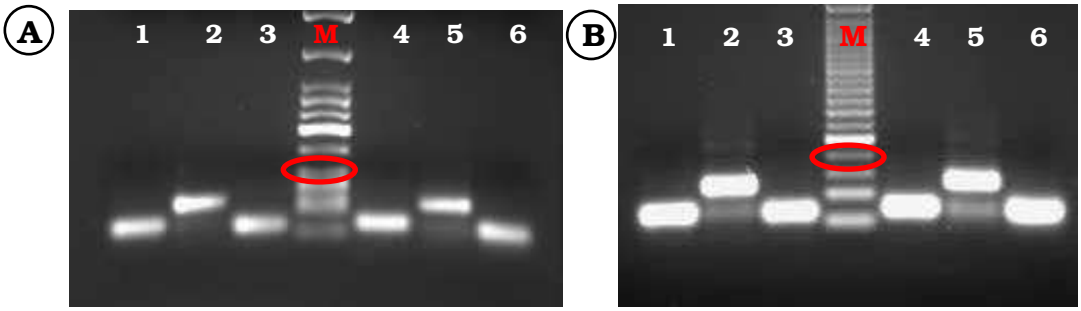

Figure IV.9: **A)** Partial synthesis of APP_QK gene; in red marker band of 100 bp, lane 1 annealing A (30bp), lane 2 ligase reaction AB (59bp), lane 3 annealing B (29bp), lane 4 annealing C (29bp), lane 5 ligase reaction CD (59bp), lane 5 annealing D (30bp).
B) Figure IV.2.4a: Partial synthesis of APP1 gene; in red marker band of 100 bp, lane 1 annealing A (30bp), lane 2 ligase reaction AB (59bp), lane 3 annealing B (29bp), lane 4 annealing C (29bp), lane 5 ligase reaction CD (59bp), lane 6 annealing D (30bp).

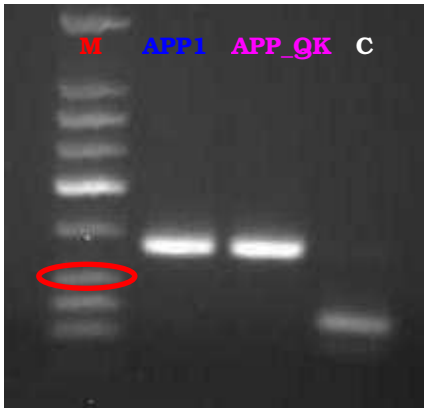

Figure IV.10: APP1 and APP_QK genes (118bp) purified, in red marker band of 100 bp, lane **C** PCR control.

Succesively, it was amplified by PCR, product and ligated into the expression vectors pETM11, pETM20 and pETM30 between NcoI and XhoI restriction site. The features of expression vectors and the relative maps was reported in paragraf IV.4.3.
The presence of the inserts was tested by PCR screening (Figure IV.11) and confirmed by DNA sequencing.

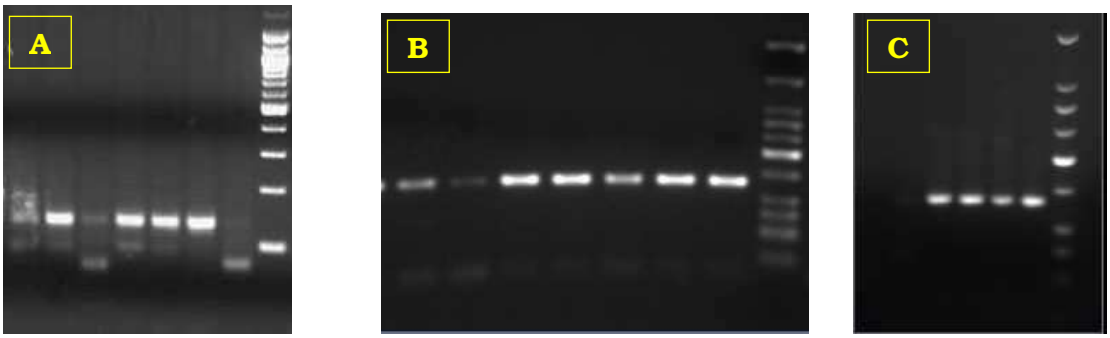

Figure IV.11: **A)** screening PCR of APPwt; **B)** screening PCR of APP1; **C)** screening PCR of APP_QK

IV.2.3 Protein expression

Optimal expression conditions were selected for each recombinant protein, by performing a complete screening in BL21(DE3), BL21(DE3)STAR, BL21*codon plus*(DE3)RIL and Rosetta GAMI(DE3) *E.coli* strains, using different induction time, temperatures and inducing agent concentration. The result are summarized in Tables IV.1-7 in Appendix IV.6. The optimal conditions selected to obtain the desidered protein soluble in large yields were illustrated in table IV.8.

Table IV.8	**STRAIN**	**IPTG concentration**	**Temperature**	**Induction time**
APPwt	BL21*codon plus*(DE3)RIL	1mM	37°C	16 hour
APP1	BL21(DE3)	0.4mM	22°C	5 hour
APP_QK	BL21(DE3)	0.4mM	22°C	5 hour

Analysis by SDS–PAGE of bacterial lysates showed the presence the high level expression of the proteins APPwt, APP1 and APP_QK in the soluble fraction. Molecular weights rivealed by gel are comparable with those expected, riported in Table IV.9.

Table IV.9	**MW**
APPwt	7,46 KDa
APP1	33,15 KDa
APP_QK	33,19 KDa

The protein over-expression are riported in figure IV.12.

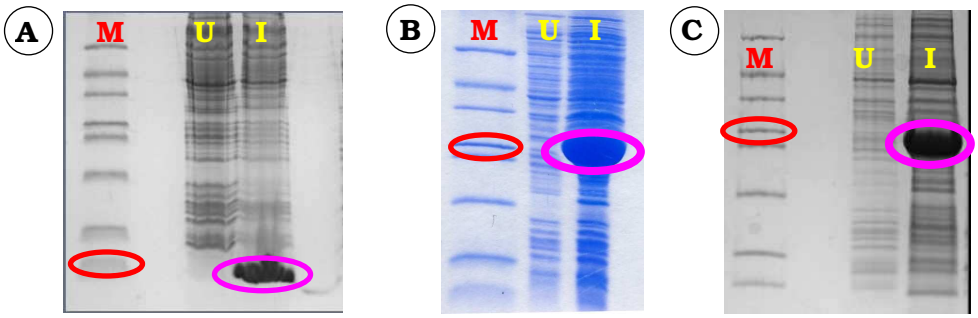

Figure IV.12: SDS-PAGE of protein expression: lane **M=**marker, lane **U**=uninduced, lane **I**=induced.
A) in magenta His6x-APPwt 7,5Kda, in red marker band of 10KDa;
B) in magenta His6x-APP1 33 Kda, in red marker band of 29KDa;
C) in magenta His6x-APP_QK 33 Kda, in red marker band of 29Kda.

IV.2.4 Protein purification

The protein 6xHis-APPwt was purified by affinity chromatography on Ni^{2+}-NTA resin.
The lysate was then loaded in presence of 100 mM NaCl to avoid non-specific binding of *E. coli* contaminants and all fractions were eluted increasing imidazole concentration.
The analysis of the SDS–PAGE gels revealed that the protein was well purified (Figure IV.13).

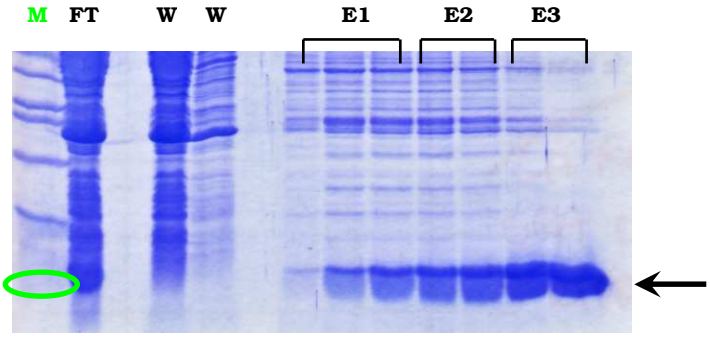

Figure IV.13: Elettrophoresis analysis on SDS-PAGE the affinity purification of His6x-APPwt;
lane **M** marker, in green band of 10 KDa, **FT**= flow through,
W=Wash resine with 20mM Imid., E1= elution to 50mM Imid., E2= elution to 150mM Imid.
E3 elution to 250 mM Imid.; arrow indicate His6x-APPwt .

Protein samples eluited were dialyzed in buffer of the protease TEV for the subsequently reaction.
The purification of 6xHis-GST-APP1 and 6xHis-GST-APP_QK proteins were loaded on His-Trap HP column, equilibrated with buffer A described in paragraf IV.4.4.
APP1 and APP_QK are eluited at 50% and 70% buffer B, described in paragraf IV.4.4, respectively. The chromatography profiles are showed in Figure IV.14.

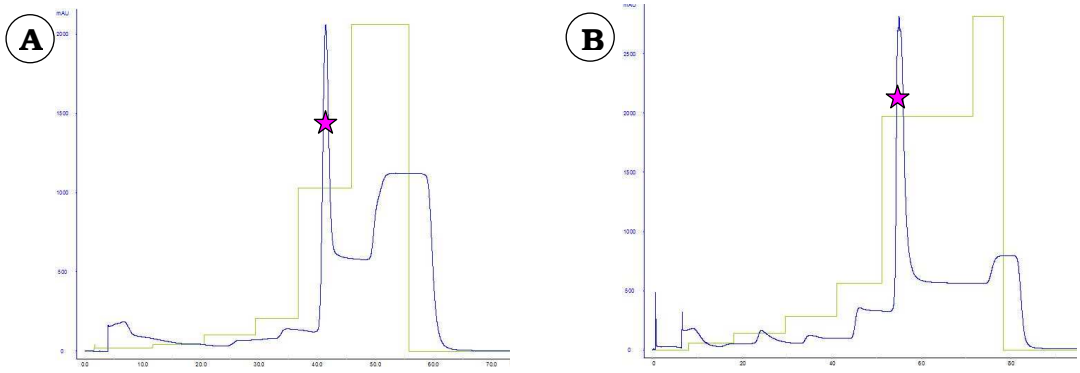

Figure IV.14: Purification on Histrap column,
A) His6x-GST-APP1 * indicate the peack of protein elution at 50% buffer B.
B) His6x-GST-APP_QK * peack of protein elution at 70% buffer B.

All fraction were analyzed by 15% SDS-PAGE (Figure IV.15 and IV.16).

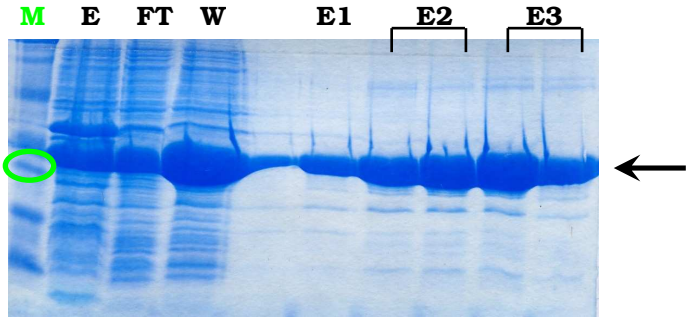

Figure IV.15: Elettrophoresis analysis on SDS-PAGE the affinity purification of His6x-GST-APP1; lane **M** marker, in green band of 29KDa, **E**= protein extract, **FT**= flow through, **W**=Wash resine with 20mM Imid., E1= elution to 50mM Imid., E2= elution to 250mM Imid. E3 elution to 500 mM Imid.; arrow indicate His6x-GST-APP1.

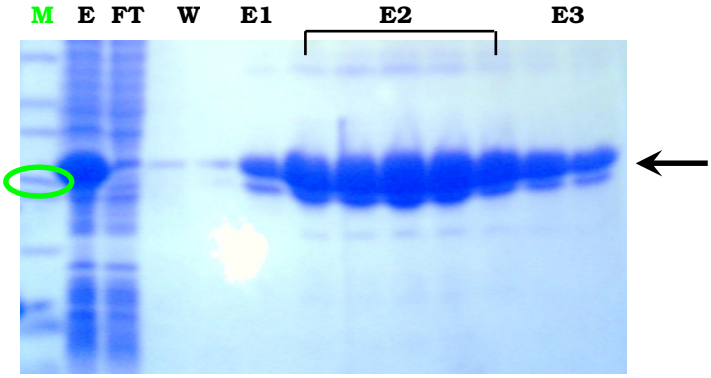

Figure IV.16: Elettrophoresis analysis on SDS-PAGE the affinity purification of His6x-GST-APP_QK; lane **M** marker, in green band of 29KDa, **E**= protein extract, **FT**= flow through, **W**=Wash resine with 20mM Imid., E1= elution to 50mM Imid., E2= elution to 350mM Imid. E3 elution to 500 mM Imid.; arrow indicate His6x-GST-APP_QK.

Protein samples eluited were dialyzed in buffer of the protease TEV for the subsequently reaction.

IV.2.5 TEV digestion and second purifications

Cleavage reaction was performed adding TEV protease to protein solution in a molar ratio of 1:50 the reaction was kept at 30°C for 4 hours.

TEV-cleaved proteins were analyzed by gel electrophoresis (Figure IV.17).

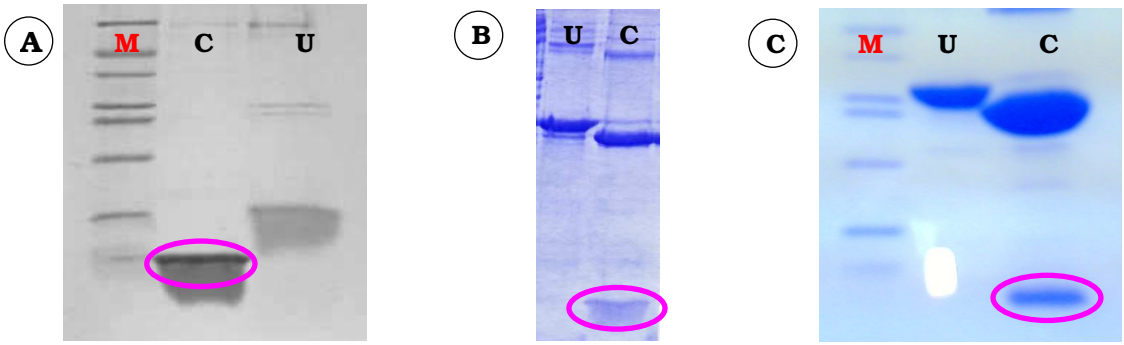

Figure IV.17: Elettrophoresis analysis on SDS-PAGE of the TEVproteolysis reaction; M marker; **C**= cut TEV,**U**= uncut; **A)** in magenta APPwt protein 4,5 KDa; **B)** in magenta APP1 4,5 KDa; **C)** in magenta APP_QK protein 4,5 KDa.

The cleaved APPwt and APP1 proteins were purified by affinity chromatografy used a Ni^{2+}-NTA column in presence of 10mM imidazole; Figure IV.18 showed SDS-PAGE of APPwt after purification and Figure IV.19 affinity profile of APP1 on HisTrap column.

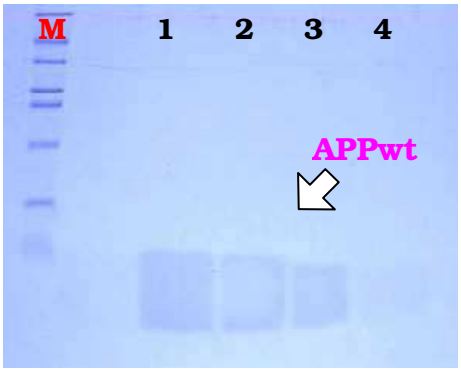

Figure IV.18: Elettrophoresis analysis on SDS-PAGE of APPwt. of the elution (lane1-4) by resin.

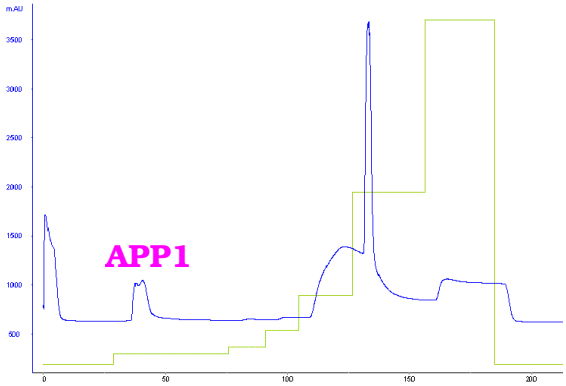

Figure IV.19: Purification of APP1 on Hitrap column by FPLC, eluted at 10mM imidazole; chromatografy profile revealed at 280nm.

Instead, APP_QK was purified by gel filtration (Figure IV.20) because on Ni^{2+}-NTA resin it was eluted with 6xHis-GST at high imidazole concentration probably because of aspecific interactions. The R.T. of 100ml is congruent with molecular weight of protein (4,5 Da).

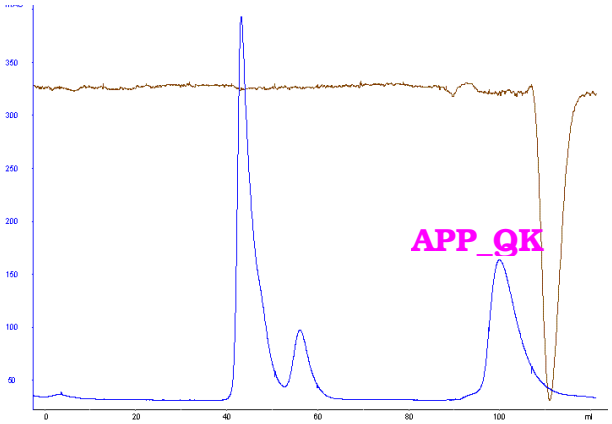

Figure IV.20: Gel filtration purification of APP_QK on Superdex S30 column, chromatografy profile revealed at 280nm.

Finally all pure proteins were concentrated until 0.3 mM.

Purity and identity of all proteins were assessed by liquid chromatography mass spectrometry (Figure IV.21a, 21b and 21c). The experimental molecular weights are in accordance with theoretical ones as showed in following table IV.9.

Table IV.9	**MW_{Th}**	**MW_{Exp}**
APPwt	4497.9	4497.3
APP1	4526.04	4523.3
APP_QK	4562.0	4559.4

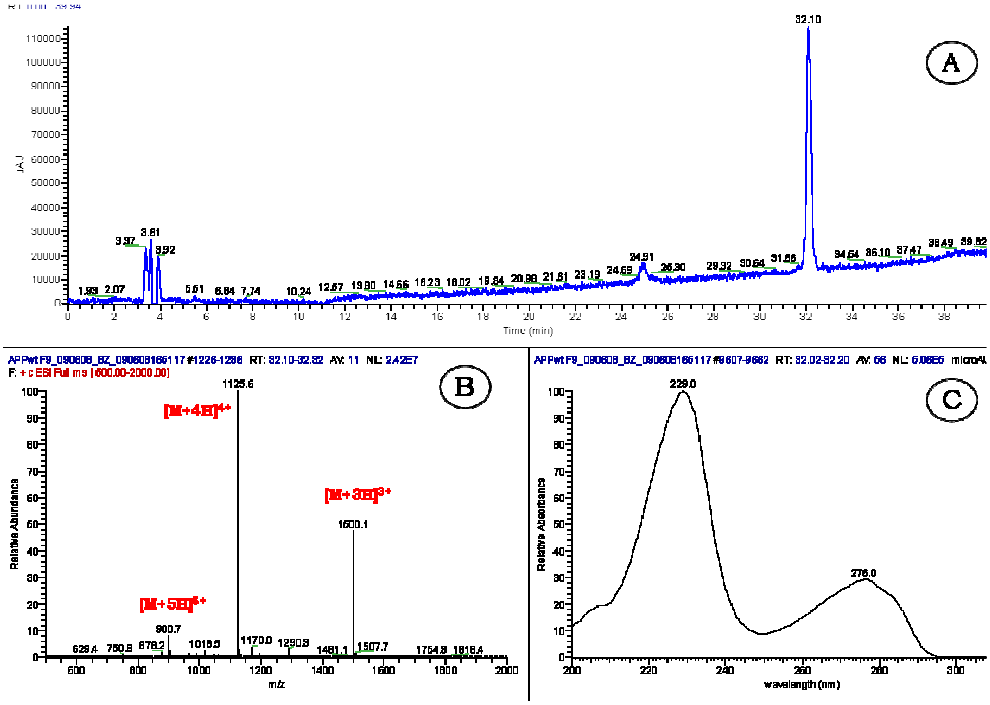

Figure IV.21a: LC-MS of APPwt purified
A) RP-HPLC profile revealed at 210nm; **B**) ESI-MS spectrum of the peak at RT: 32.10 min;
C) UV absorption spectra of the peak at RT: 32.10min.

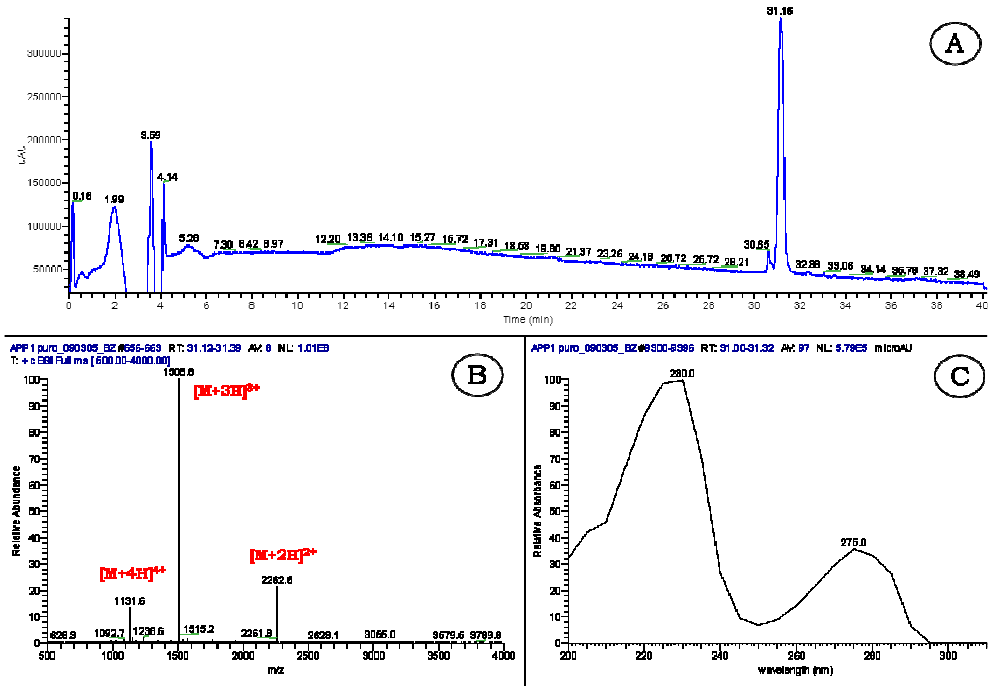

Figure IV.21b: LC-MS of APP1 purified
A) RP-HPLC profile revealed at 210nm; **B**) ESI-MS spectrum of the peak at RT: 31.16 min;
C) UV absorption spectra of the peak at RT: 31.16min.

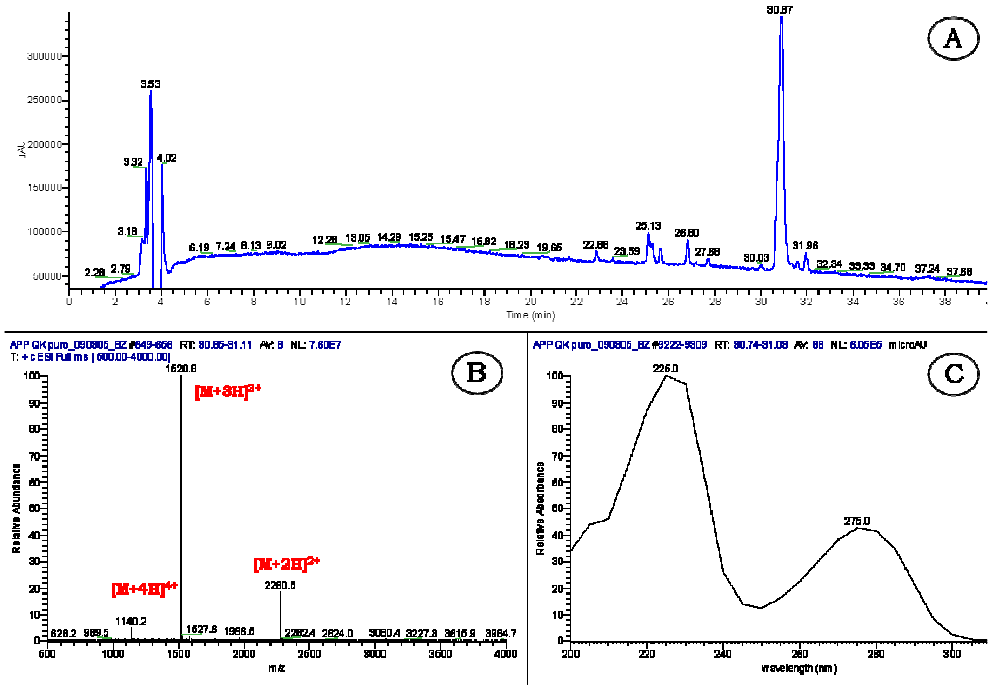

Figure IV.21c: LC-MS of APP_QK purified
A) RP-HPLC profile revealed at 210nm; **B**) ESI-MS spectrum of the peak at RT: 30.87min;
C) UV absorption spectra of the peak at RT: 30.87min.

The three proteins were obteined as a homogeneous product at high purity (Figure IV.22a, 22b and 22c). The final expression yields were about 5mg for APPwt and 3mg for APP1 and APP_QK.

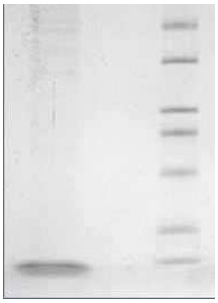

Figure IV.22a: APPwt pure.

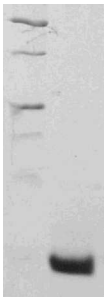

Figure IV.22b: APP1 pure.

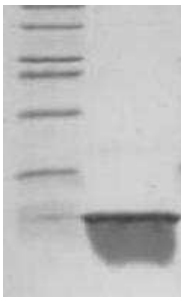

Figure IV.22c: APP_QK pure.

IV.2.6 Caspase 3 fluorimetric Assay

The literature shows that VEGF is able to inhibit approximately 40% induced apoptosis (Yilmaz A, Kliche S, Mayr-Beyrle U Fellbrich G, Waltenberger J., Biochem. Biophys. Res. 2003 306 730 -736), for that reason it has developed a fluorimetric test for an early event, the activity of caspase 3, specific enzyme of apoptosis (JMAdams, Genes. DeIV. 2003, 20 2481-95).

The APP_QK and APPwtas control were tested with *Caspase 3* assay (Figure IV.23).

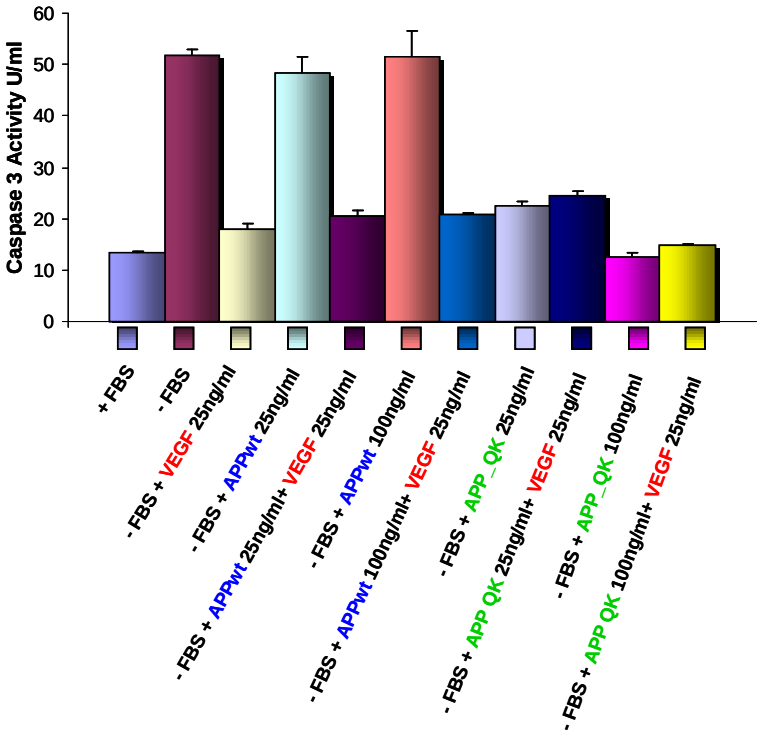

Figure IV.23: Caspase 3 assay on APPwt and APP_QK.

This experiment was show an activity of APP_QK similar to VEGF, while APPwt as expected has not activity.

Caspase 3 experiments with the laboratory of Prof. MC Turco University of Salerno.

IV.3 Discussions

There are important classes of proteins whose role lies in the tight complex formation with other polypeptides and whose architecture is mainly determined by rigid secondary structure. In these cases loop regions play a lesser role in the binding function but often adopt regular turn structures and as such ensure the dense packing of secondary structural elements. In this manner, several a-helices or β-strands are brought together and side chains protruding from two or more of these elements form the interface for complex formation with the target protein.

The *de novo* design of a peptide sequence adopting a well-defined secondary structure includes a first step of introducing residues in order to stabilize the secondary structure and a second in which you enter the residues responsible for interaction in the correct spatial orientation. An alternative approaches is grafting of the interacting residues on protein scaffolds that have the desired secondary structure exposed to the solvent. In this case the protein scaffold serve to structurally fix the grafted peptide segment and, indeed, achievable affinities seem to be higher if compared to the corresponding conformationally flexible peptide. Another advantage of the scaffolding approach is the possibility to obtain reagents with high physicochemical and possibly proteolytic stability. Infact, a "scaffold" proteins are it is reasonably assumed that the beneficial properties of a robust wild-type scaffold protein will be largely retained in its engineered descendants. Such stability aspects may not only include the capability of maintaining a functional conformation during various types of 'environmental stress' but also relate to the ability of reversibly regaining the native structure after potential denaturation.

In this work we chosen APP as starting scaffold to design an optimized miniprotein able to mimic VEGF in the interaction with its membrane receptors. To this aim we mutated 5 amino acids of the original APP, introducing the interacting residues of the VEGF helix 17-25 and of QK epitope binding. The new APP proteins expressed in *E.coli* and preliminary biological assay performed on APP_QK demonstrated that it activity is comparable with that of native VEGF in terms of inhibition of caspase 3.

IV.4 Experimental Section

IV.4.1 Materials and Methods

Reagents used for preparation of buffers and growth media of *Escherichia coli* were supplied by Sigma Aldrich, the reagents for polyacrylamide gels electrophoresis (Acrylamide, APS, TEMED, SDS, Tris, Glycine) by Applichem. The molecular weight markers for proteins were from Sigma Aldrich. Oligonucleotides were synthesized by Sigma-Genosys (Sigma-Aldrich). *Pfu Turbo* DNA polymerase is from Stratagene. Restriction enzymes and the "modification enzymes" (Calf Intestine Phosphatase (CIP), T4 DNA Ligase, *Taq* DNA polymerase (5 U/μL) and T4 DNA Polynucleotide Kinase) are from New England Biolabs. The molecular weight markers for nucleic acids were supplied by NEB and Roche. pETM vectors for the expression of recombinant proteins are from EMBL (Heidelberg), while the expression vector pPROEXHTa was supplied by Invitrogen.
All molecular biology kits are from Qiagen. *E.*coli TOP10F' strain, used for cloning, was supplied by Invitrogen; *E. coli* BL21 *Codon Plus* (DE3) RIL cells, used for over expression, were supplied by Stratagene. BL21(DE3) and BL21(DE3) STAR, also used for expression, were supplied by Invitrogen while RosettaGAMI(DE3) were from Novagen.
Isopropylbeta- D-thiogalactopyranoside (IPTG) is from Inalco. Complete Protease Inhibitor Cocktail Tablets were supplied by Roche and used as a mixture of protease inhibitors, according to manufacturer's instruction.
Chromatography columns and AKTA FPLC are from GE HealthCare. Ethanol, propanol and acetic acid were supplied by J.T. Baker. TFA was from Fluka. Purity and identity of protein samples were assessed by LC-MS system (Thermo Electron) comprising an LCQ Deca XP MAX ion trap mass spectrometer equipped with an ESI source and a complete Surveyor HPLC system (including MS pump, autosampler and photo diode array [PDA]).
UV-VIS spectra were performed by using a Jasco V-550 UV-VIS spectrophotometer, in an 1 cm quartz cell (Hellma).
CD spectra were recorded using a Jasco J-810 spectropolarimeter (JASCO Corp) equipped with a Peltier-type temperature control system.

IV.4.2 Synthesis of gene

The gene coding for the mini proteins, APPwt, APP1 and APP_QK, were synthesized starting from short oligonucleotide sequences showed following.

Gene APPwt:

A1 5′-CGCATATGGGTCCGTCTCAGCCGACC

A2 5′GGGTAGGTCGGCTGAGACGGACCCATATGCG

A3 5′-TACCCGGGTGATGATGCGCCGGTTGAAG

A4 5′-TCAGATCTTCAACCGGCGCATCATCACCC

A5 5′-ATCTGATCCGTTTCTACAACGATCTGCAG

A6 5′-TACTGCTGCAGATCGTTGTAGAAACGGA

A7 5′-CAGTACCTGAACGTTGTTACCCGTCACCGTTAC

A8 5′-GTAACGGTGACGGGTAACAACGTTCAGG

Gene APP1:

A1 5′ -CGCATATGGGTCCGTCTCAGCCGACC

A2 5′ -GGGTAGGTCGGCTGAGACGGACCCATATGCG

A10 5′ -TACCCGGGTGATGATGCGCCGTTCATG

A11 5′ -CCCACTAGTACGCGGCAAGTACCTAGA

A12 5′ -GATCTGTACCAGTTCGCGTACAACCTGCAG

A13 5′ -CATGGTCAAGCGCATGTTGGACGTCGTCAT

A7 5′ -CAGTACCTGAACGTTGTTACCCGTCACCGTTAC

A8 5′ –GTAACGGTGACGGGTAACAACGTTCAGG

Gene APP_QK:

A1 5′ -CGCATATGGGTCCGTCTCAGCCGACC

A2 5′ -GGGTAGGTCGGCTGAGACGGACCCATATGCG

A14 5′ -TACCCGGGTGATGATGCGCCGTGGCAG

A15 5′ -CCCACTAGTACGCGGCACCGTCCTAGA

A12 5′ -GATCTGTACCAGTTCGCGTACAACCTGCAG

A13 5′ -CATGGTCAAGCGCATGTTGGACGTCGTCAT

A7 5′ -CAGTACCTGAACGTTGTTACCCGTCACCGTTAC

A8 5′ –GTAACGGTGACGGGTAACAACGTTCAGG

The previous oligonucleotides were used to obtein the fragment as showed in figure IV.2.1., the colours were congruent with the picture.

Figure IV.2.1: methods of synthesis gene

The methodology was composed of three steps; in step 1 the complementares oligonucleotides before were denatured at the temperature of 95°C for 2 minutes and after were annealing at 4°C , in step 2 the dna doble strand generated were ligated with T4 DNA Ligase (400 U/µL) 20 U/µg for 1 hours at 16°C and in final step 3 the two

fragment were ligated for obtainment the ligase of complete gene with same conditions ligase of step two.

The complete sequence are :

APPwt

```
5′-CGCATATGGGTCCGTCTCAGCCGACCTACCCGGGTGATGATGCGCCGGTTGAAGATCT-
   GCGTATACCCAGGCAGAGTCGGCTGGATGGGCCCACTACTACGCGGCCAACTTCTAGA

  -GATCCGTTTCTACAACGATCTGCAGCAGTACCTGAACGTTGTTACCCGTCACCGTTAC-3′
   CTAGGCAAAGATGTTGCTAGACGTCGTCATGGACTTGCAACAATGGGCAGTGGCAATG
```

APP1

```
5′-CGCATATGGGTCCGTCTCAGCCGACCTACCCGGGTGATGATGCGCCGTTCATGGATCT-
   GCGTATACCCAGGCAGAGTCGGCTGGATGGGCCCACTACTACGCGGCAAGTACCTAGA

  -GTACCAGTTCGCGTACAACCTGCAGCAGTACCTGAACGTTGTTACCCGTCACCGTTAC-3′
   CATGGTCAAGCGCATGTTGGACGTCGTCATGGACTTGCAACAATGGGCAGTGGCAATG
```

APP_QK

```
5′-CGCATATGGGTCCGTCTCAGCCGACCTACCCGGGTGATGATGCGCCGTGGCAGGATCT-
   GCGTATACCCAGGCAGAGTCGGCTGGATGGGCCCACTACTACGCGGCACCGTCCTAGA

  -GTACCAGTTCGCGTACAACCTGCAGCAGTACCTGAACGTTGTTACCCGTCACCGTTAC-3′
   CATGGTCAAGCGCATGTTGGACGTCGTCATGGACTTGCAACAATGGGCAGTGGCAATG
```

IV.4.3 Cloning of gene

The genes after the synthesis were amplified with PCR. All amplification reactions were performed in a final volume of 100 µL, using 50 ng of template DNA.

The reaction mixture contained the 2 primers (0.25 µM each), dNTPs (0.25 mM each) and the *Pfu turbo* polymerase (5U) with its buffer 1X.

PCR was performed using an *Eppendorf Mastercycler personal* apparatus, following the protocol indicated below:

- Initial denaturation (step 1) 3 min at 95°C
- Denaturation (step 2) 1 min at 95°C
- Annealing (step 3) 1 min at the appropriate temperature for each gene amplified
- Elongation (step 4) 1 min at 72°C for 30 cycles, from step 2.

All amplification products were analyzed by 2% agarose gel electrophoresis performed in TAE buffer (18.6 g/L EDTA, 242 g/L Tris base, add acetic acid until pH 7.8).

IV.4.3.1 APPwt in pPROEX-HTa

The amplification was performed by using the following couples of primers:

FORWARD

A-NcoI 5'-C ATC GCC ATG GGT CCG TCT CAG CCG ACC TAC CCG 34mer (NcoI)

REVERSE

A-XhoI 5'- AA CAA TGG GCA GTG GCA ATG **ATC ACT** GAGCTC ATA 35mer (XhoI)

The *forward* primers indicated above contained the *NcoI* restriction site, while the *reverse* primers were designed with the *Xho*I restriction site positioned downstream to two stop-translation codons (bold in the sequences). This cloning strategy was used to obtain the gene products fused (at the N-terminal) with the 6x-Histidine-tag of the pPROEXHTa expression vector. The pPROEXHTa plasmid is designed to allow inducible, high level intracellular expression of genes; it contains a *tac* promoter for chemical induction, a multiple cloning site (MCS), an internal *lac Iq* gene compatible to any *E. coli* host and a TEV protease recognition site for cleaving the fusion protein (Figure IV.2.2).

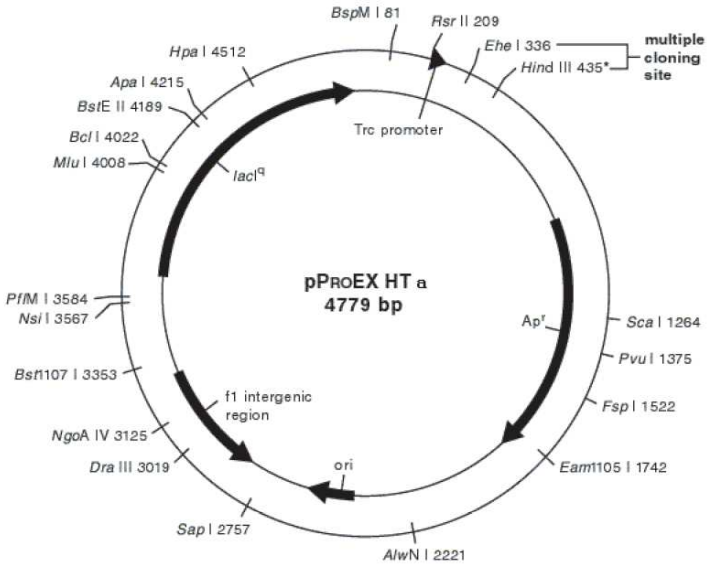

Figure IV.2.2 :Structural organization of pPROEXHTa expression vector

PCR products were purified by using the *QIAquick PCR Purification Kit*, and digested with *Nco*I (20 U/µL) and *Xho*I (10 U/µL) restriction enzymes.

Each amplified fragment (1µg) was digested with 3U of restriction enzymes for 2 hours at 37ºC in a buffer containing 50 mM NaCl, 10 mM Tris-HCl, 10 mM $MgCl_2$, 1mM DTT pH 7.9 supplemented with BSA 100 µg/mL. Following the digestion, each fragment was

cloned into the corresponding sites of the pPROEXHTa expression vector, downstream to the His-tag sequence.

To this purpose, the expression vector was previously digested with the same restriction enzymes (3 U/µg), and treated with CIP enzyme (10 U/µg) for 30 min at 37ºC. CIP was then inactivated at 75°C for 10 min. After digestion, PCR amplifications were purified by *QIAquick PCR Purification Kit*, while the *Nco*I/*Xho*I pPROEXHTa was purified by *QIAquick Gel Extraction Kit*. For ligation reactions was used a 1:6 molar ratio (vector/insert DNA). The reactions were performed using 20 U/µg DNA of the T4 DNA Ligase (400 U/µL), in a final volume of 10 µL, for 1 hours at 16°C. The identity of the inserts in the resulting recombinant plasmids was confirmed by DNA sequencing (MWG-Biotech). *E. coli* TOP F'10 strain was used for cloning.

IV.4.3.2 APP1 and APP_QK in pETM

APP1 and APP_QK amplified gene were cloned initially in pETM11 and pETM20 expression vectors and subsequently in pETM30 (figure IV.2.3, IV.2.4 and IV.2.5).

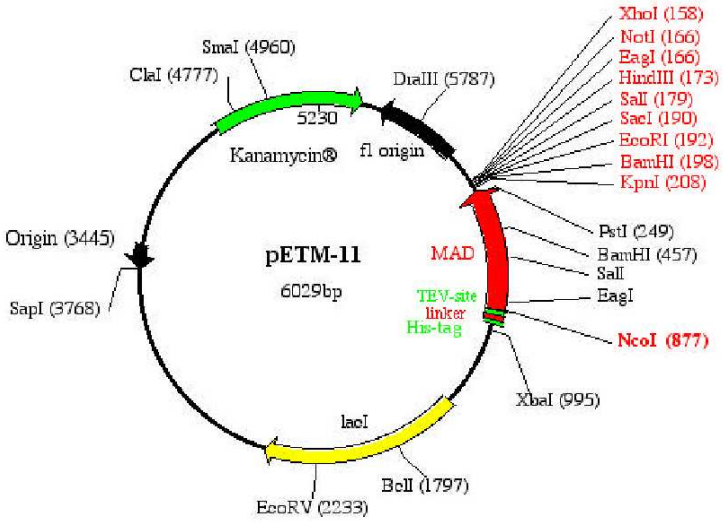

Figure IV.2.3: Structural organization of pETM11 expression vector

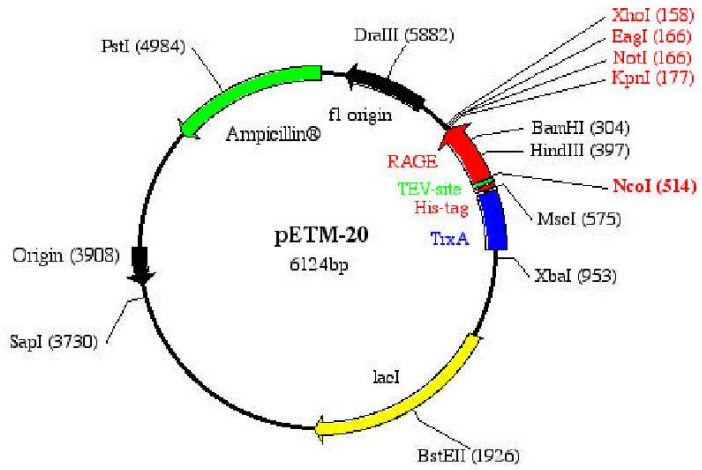

Figure IV.2.4: Structural organization of pETM20 expression vector

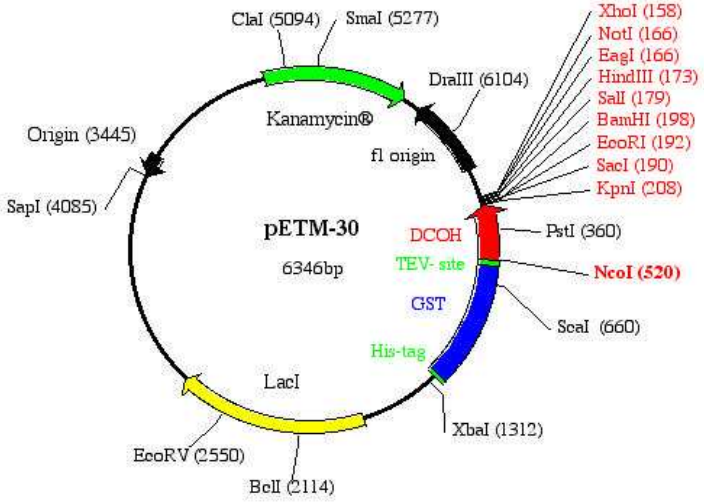

Figure IV.2.5: Structural organization of pETM30 expression vector

pETM11 contains a 6x-Histidine-tag, the pETM20 system contains a 6x-Histidine-TrxA fusion-tag for the increase of the solubility of gene products, while pETM30 system contains a 6xHistidine-GST fusion-tag for the increase of the solubility of gene products.
The primers used for amplification by PCR of gene in these systems were the same used for the amplification of APPwt gene; digestions and purifications of plasmids were performed in the same way as the APPwt in pPROEX-HTa.
For ligation reactions was used a 1:6 molar ratio (vector/insert DNA). The reactions were performed using 20 U/µg DNA of the T4 DNA Ligase (400 U/µL), in a final volume of 10 µL, for 1 hours at 16°C. The identity of the inserts in the resulting recombinant plasmids was confirmed by DNA sequencing (MWG-Biotech). *E. coli* TOP10F' strain was used for cloning.

IV.4.4 Determination of the concentration and Electrophoretic analysis

The concentration of the proteins in solution was determined according to the Bradford's method (Bradford, 1976). The Coomassie Brilliant (Bio-Rad) reagent was added to the samples and the absorbance at 595 nm was monitored.
A solution of 1 μg/μL of bovine serum albumin (BSA) was used as standard. Protein concentration was also measured by UV spectroscopy, reading the tryptophan absorbance at 280 nm, using a Jasco V-550 UV-VIS spectrophotometer, in an 1 cm quartz cell. The electrophoresis on polyacrylamide gel in the presence of SDS (SDS-PAGE) was performed according to Laemmli's protocol (Laemmli, 1970). The samples were denaturized at 95°C for 10 min in 1% SDS (Applichem), 5% β-mercaptoethanol (Sigma), 0.001% bromophenol blue (ICN Biomedicals) and 10% glycerol (Applichem). The samples were then loaded on a 15% or 18% polyacrylamide gel and electrophoresed in 0.025 M Tris-HCl, 0.2 M glycine pH 8.3 and 0.1% SDS. The electrophoresis was performed at 30 mA for ~60 minutes. The proteins were then revealed by Coomassie Brilliant-Blue (Applichem) staining; the gel was submerged in the staining solution (0.1% Coomassie Brilliant-Blue R250, 25% isopropilic alcohol and 10% acetic acid) for ~30 min with gentle agitation. The gel was washed in a solution containing 30% ethanol and 10% acetic acid to remove the excess of Coomassie and then stored in 10% acetic acid.

IV.4.5 Protein expression and purification

The host strain chosen for the expression of 6xHis-APPwt was the BL21 *Codon plus* (DE3) RIL strain. BL21 *Codon Plus* (DE3) RIL competent cells are efficient for the high-expression of eukaryotic proteins in *E. coli* since they are engineered to contain extra copies of the genes that encode the tRNAs that most frequently limit translation of heterologous proteins in E. coli. In fact, these strains supply tRNAs for AGG, AGA, AUA, CUA, CCC and GGA on a compatible chloramphenicol-resistant plasmid.

100 ng of expression vector were chemically transformed into *E. coli* competent cells from the following *E. coli* strains: Bl21 (DE3), Bl21 (DE3)star, Bl21 codon plus (DE3)RIL and Rosetta GAMI (DE3). Recombinant colonies were grown in 5 mL LB medium.
Single clones, previously transformed with each recombinant expression vector and grown at 37ºC on LB agar containing antibiotics (100 μg/mL ampicillin for pPROEXHTa recombinant vectors or 50 μg/mL kanamycin for pETM recombinant vectors), were inoculated in 20 mL of LB medium, containing the same antibiotics. The overnight cultures were inoculated into 1L of prewarmed LB medium supplemented with antibiotics; the culture was grown up to 0.7 OD/mL at 37°C this is achieved, the expression was induced with different IPTG concentrations (0.3 mM, 0.5 mM, 0.7 mM

and 1 mM) for different times (3 hours and 16 hours) and temperature (22°C and 37°C). So the cells were harvested by centrifugation (7000 rpm, 15 min, 4°C).
Cell pellets were resuspended in 50 mM Tris–HCl, pH 8 containing protease inhibitors, to avoid proteins degradation, and the suspension was sonicated in ice using 30" on/30" off cycles for 10 min, by using a Misonix Sonicator 3000 apparatus with a macro tip probe and an impulse output of 4.5/5 (=40/55 Watt). Bacterial lysates were then centrifuged (18000 rpm, 45 min, 4°C). 15 µL total and soluble fractions from the lysates were re-suspended in 5 µL SDS loading buffer (Tris-HCl 50 mM, SDS 1%, blue bromophenol 0.1%, glycerol 10%, pH 6.8) and analyzed on SDS-PAGE.
For batch purification of 6xHis-tagged APPwt, the surnatant was added to 50% Ni^{2+}-NTA agarose slurry resin (Qiagen) for the affinity chromatography purification. The binding capacity of Ni^{2+}-NTA resins is between 5-10 mg/mL for 6xHis-tagged proteins. Before loading the protein extract, the resin was extensively washed with water and finally equilibrated in the lysis buffer. The lysate was then loaded in presence of 100 mM NaCl to avoid non-specific binding of *E. coli* contaminants and the column was gently shaken for 45 min-1h. Then, the flow-through (F.T.), containing all unbound proteins, was collected and the resin was washed three times with the lysis buffer containing 100 mM NaCl, and finally the His-tagged proteins were eluted with concentrations of imidazole (50/100 mM). All fractions eluted from the resin (flow-through, washes and elutions) were analyzed on SDS–PAGE gels and stained with Coomassie Brilliant Blue R-250. Protein samples were dialyzed at 20°C by using Spectra/Por membranes with the appropriate MWCO in 50 mM Tris-HCl, 150 mM NaCl, 1 mM DTT and 0.5 mM EDTA pH 8. buffer. For purification of fusion proteins APP1 and APP_QK with both Trx than with GST, instead the supernatant was loaded on a 5 mL His-Trap HP column, previously equilibrated with buffer A (50 mM Tris-HCl, 300 mM NaCl, pH 8.0), using an AKTA FPLC chromatography system. The column was washed with buffer 1 and the bound protein was eluted using a gradient to step of imidazole from 0 mM to 500 mM in buffer B (50 mM Tris-HCl, 300 mM NaCl, 500 mM imidazole, pH 8) in 4 steps. Protein elution was monitored by measuring the absorbance at 280 nm and 214 nm. The proteins were eluted at 70% buffer B (350 mM imidazole). The total cell protein fraction, the soluble fraction (supernatant of cell lysate) and the purified fraction were analyzed by 15% SDS-PAGE. The proteins recovered were dialyzed at 4°C in 50 mM Tris-HCl, 150 mM NaCl, 1 mM DTT and 0.5 mM EDTA pH 8 buffer by using Spectra/Por membranes with the appropriate MWCO.

IV.4.6 TEV digestion

All 6xHis-tagged fusion proteins, after purification by affinity chromatography were dialyzed overnight against TEV buffer. To protein substrates was added TEV protease,

using a molar ratio (protease:substrate) of 1:50 and the cleavage was allowed to proceed at 30°C for 4 hours. Cleaved products were analyzed by 18% polyacrylamide gel electrophoresis performed in Laemmli buffer, then, mixture was loaded onto a Ni^{2+}-NTA affinity column equilibrated in binding buffer (the same employed for digestion reaction) containing 10 mM imidazole; the proteins without His-tag were collected in the flow-through, while His-tag costruction and TEV protease remained bound to the column.
TEV protease was expressed and purified in our laboratory, after transforming the BL21 *pLysis*S cells (Invitrogen) with the pET24a-TEV recombinant vector kindly provided by Dr. Nina Dathan.
LC-MS analyses were performed to estimate the protein purity and assess the molecular weight. For this analysis, 0.5 µg of protein were loaded on a 300 Å narrow bore 250x2mm C18 Jupiter column (Phenomenex) coupled to the LC-MS system previously described. A gradient of solvent B (0.05% TFA in CH3CN) from solvent A (0.08% TFA in H2O) of 5% to 70% was applied over 30 min. Mass spectra were recorded continuously in the mass interval 400-2000 amu, (LC-MS, condition 1).

IV.4.7 Casapase 3 fluorimetric assay

Determination of caspase-3 activity was performed by a fluorometric assay based on the proteolytic cleavage of the AMC-derived substrate N-acetil- DEVD-AMC, which yields a fluorescent product.
HUVEC cells were plated in 6-well dishes at 1x105 cells/cm^2. On the next day, cells were treated, in starvation medium (EBM-2, heparin 0.1%, BSA 0.1%), with APPwt (25-100 ng/ml) or APP_QK (25-100 ng/ml) peptides for 8h at 37°C. VEGF165 (R & D Systems, Minneapolis, MN, USA), 25 ng/ml, was used as positive control.
After 8h the cells were processed with 150 µl of Caspase-3 reaction buffer (HEPES pH 7.5 50 mM, EDTA 0.1 mM, NP-40 0.1%, CHAPS 0.1%, DTT 1 mM) and cell proteins collected after centrifugation at 13000 rpm for 15 min at 4°C. Protein concentrations were determined by Bradford method (Bio-Rad, Hercules, CA) and 20 µg of lysates were incubated in 96-well plates with 20 µM N-acetil-DEVD-AMC at 37°C for 3h.
Samples were analyzed using a microplate reader (L55 Luminescence Spectometrer Perkin Elmer Instruments) (excitation: 360 nm, emission: 440 nm).

IV.5 References

- Carmeliet, P. Nat. Med. 9, 653-660, 2003
- Chong Li, Min Liu, Juahdi Monbo, Guozhang Zou, Changqing Li, Weirong Yuan, Davide Zella, Wei-Yue Lu, and Wuyuan Lu, J. AM. CHEM. SOC. 9 VOL. 130, NO. 41, 2008
- Domingues, H., Gregut, D., Sebald, W., Oschkinat, H. and Serrano, L. (1999) Nature Struct. Biol. 6, 652-656
- Kellenberger, E., Mer, G., Kellenberger, C., Marguerie, G. and Lefèvre, J.F. (1999) Eur. J. Biochem. 260, 810-817.;
- Kritzer J. A., R. Zutshi, M. Cheah, F. A. Ran, R. Webman, T. M.; Wongjirad and A. Schepartz, ChemBioChem, 2006, 7, 29.
- Maneesh Jain, Neel Kamal and Surinder K. Batra TRENDS in Biotechnology Vol.25 No.7 (307-316);
- Muller, Y.A., Chen, Y., Christinger, H.W., Li, B., Cunningham, B.C., Lowman, H.B., de Vos, A.M. (1998) Structure 6: 1153-1167.
- Muller, Y.A., Christinger, H.W., Keyt, B.A., de Vos, A.M. (1997b) Structure 5: 1325-1338.
- Muller, Y.A., Li, B., Christinger, H.W., Wells, J.A., Cunningham, B.C., de Vos, A.M. (1997a) Proc Natl Acad Sci U S A 94: 7192-7197.
- Pan, B., Li, B., Russell, S.J., Tom, J.Y., Cochran, A.G., Fairbrother, W.J. (2002) J Mol Biol 316: 769-787.
- Rutledge S. E., H. M. Volkman and A. Schepartz, J. Am. Chem. Soc., 2003, 125, 14336.
- Schneider T. L., R. S. Mathew, K. P. Rice, K. Tamaki, J. L. Wood and A. Schepartz, Org. Lett., 2005, 7, 1695
- Wiesmann, C., Christinger, H.W., Cochran, A.G., Cunningham, B.C., Fairbrother, W.J., Keenan, C.J., Meng, G., de Vos, A.M. (1998) Biochemistry 37: 17765-17772.
- Wiesmann, C., Fuh, G., Christinger, H.W., Eigenbrot, C., Wells, J.A., de Vos, A.M. (1997) Cell 91: 695-704.

V CONCLUSIONS

CONCLUSIONS

One of frontiers of the drug discovery is the capacity to develop new molecules able to modulate protein-protein interaction. One promising approach concerns to the use of scaffold where to graft the functional groups required for the molecular interaction.

Our intent was to contribute to this exciting fild with several approaches.

In new easy and one-pot synthetic procedure to covalent link two polypeptides through their C-termini was developed. This method relies on the use of EPL an appropiete linker. This methodology will allow to easily synthetize, without any restrictions on the molecular size, molecules composed of polypeptide fragments belonging to the discontinous recognition interface which are closed in the space but distance in sequence or prepare minimized version of known scaffold as antibody or as alternative dimerization such as leucin zipper.

In the case of protein interaction mediated by secondary structure motif, such as α-helix, it will be useful to dispose of stable helical peptides or mini-proteins. This as been the topic of the other two approaches described in this work.

Starting from a 15-mer peptide, QK, which adopts in water a well-defined helical conformation, we performed a chemico-physical characterization to evaluate if QK could be a suitable helical scaffold. The biophysical characterization showed that peptide QK presents an unusual thermal stability and combining experimental and theoretical techniques we understand, at molecular level, the structural determinants of this extra stability. This study suggests that peptide QK is a potential helical scaffold and furnishes structural information about the role of each residue of the amino acid sequence. This information is a fundamental for the scaffold purification.

Finally, we provide an example functional scaffold. The α-helix of the mini-protein APP was modified introducing amino acid belonging to the VEGF receptor binding epitopes of VEGF and peptide QK. The new generated mini-protein showed a promising biological activity.

In conclusion, this work, providing chemical tools to target protein-protein interactions, contributes to the development of new scaffold for potential application in chemical biology or medicine.

VI ABBREVIATIONS

ABBREVIATIONS

ε	Epsilon
APS	ammonium persulfate
Bp	base pair
BSA	albumin from bovine serum
CD	circular dichroism
DCM:	dichloromethane
DIPEA	diisopropil-etilendiammina
DMF	N,N-dimethylformamide
DNA	deoxyribonucleic acid
dNTP	deoxy nucleotide tri-phosphate
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	ethylene-diamino-tetraacetic acid
ESI	electron spry ionization source
Flt-1	fms-like tyrosine kinase-1
F.T.	flow-through
h	Hour
HPLC	high performance liquid chromatography
HUVEC	Human umbilical vein endothelial cell
IPTG	isopropyl-beta-D-thiogalactopyranoside
KDa	Kilo Dalton
KDR	kinase domain region
LB	Luria-Bertani Broth (10g/L bacto-triptone, 5g/L yeast exstract, 10 g/L NaCl)
LC-MS	Liquid Chromatography Mass Spectrometry
MCS	multi cloning site
min	Minute
ms	Millisecond
MWCO	molecular weight cut off
nm	Nanometer
NMR	Nuclear Magnetic Resonance
NOESY	Nuclear Overhauser Effect SpectroscopY
Ni-NTA	nickel-nitrilotriacetic acid
O.D.	optical density
orf	open reading frame

PBS	phosphate buffer saline
PCR	polymerase chain reaction
pI	isoelectric point
RT	room temperature
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
sec	Second
TAE	Tris acetate EDTA
TE	10 mM Tris/HCl pH 8, 1 mM EDTA pH 8
TEMED	N,N,N',N'-tetramethyl ethylene diamine
TEV	tobacco etch virus
TFA	trifluoroacetic acid
TIC	total ion current
TIS	Triisopropylsilane
Tris	Tris (hidroxy methyl) amino methane
***trx*B**	thioredoxin reductase gene
TrxA	thioredoxin A
U	Units
UV	ultra violet
VEGF	Vascular Endothelial Growth Factor
VEGFR	VEGF receptor

The one letter code is used for amino acids.

SCIENTIFIC PRESENTETIONS

SCIENTIFIC PRESENTATIONS

- 11th Chianti workshop on Magnetic Resonance, Vallombrosa (FI), Italy, 3-8 June 2007
  Structural characterization of the helix stability of a de novo engineered VEGF mimicking peptide with relevant biological activity
  D. Diana, M. Serino, G. Scartabelli, G. Colombo, B. Ziaco, C. Pedone, L.D. D'Andrea, R. Fattorusso

- "Giornate Scientifiche del Polo delle Scienze e delle Tecnologie per la vita", Napoli, Italy, 20/21 September 2007
  Semisintesi di dimeri di proteine
  B. Ziaco, S. Pensato, L.D. D'Andrea, E. Benedetti e A. Romanelli

- Convegno Nazionale della DCSB della SCI, Montagnana (PD), Italy, 8-9 November 2007
  Structural characterization of the helix stability of a de novo engineered VEGF mimicking peptide
  D. Diana, M. Serino, G. Scartabelli, B. Ziaco, G. Colombo, C. Pedone, L.D. D'Andrea, R. Fattorusso

- Convegno Nazionale della DCSB della SCI, Montagnana (PD), Italy, 8-9 November 2007
  Semisynthesis of protein dimers by expressed protein ligation.
  Ziaco B., Pensato S., D'Andrea L.D., Benedetti E., Romanelli A.

- 30th EPS, Helsinki, Finland, 31August-5 september 2008
  Semisynthesis of protein dimers by Expressed protein ligation.
  B. Ziaco, S. Pensato, L.D. D'Andrea, E. Benedetti, A. Romanelli

- 30th EPS, Helsinki, Finland, 31August-5 september 2008
  The unusual helix stability of a VEGF mimetic peptide
  B. Ziaco, S. Pensato, L.D. D'Andrea, E. Benedetti, A. Romanelli

- VIII European Symposium of the Protein Society, Zurich, Switzerland, 14-18 June 2009
  VEGFR1D2 expression and NMR analysis of its interaction with VEGF mimicking peptides
  Di Stasi R., Diana D., Ziaco B., Capasso D., Fattorusso R., Pedone C., D'Andrea L.D.

- XXXIII Congresso Nazionale della SCI, Sorrento, Italy, 5-10 July 2009
  Peptides modulators of VEGF-dependent angiogenesis
  Ziaco B., Diana D., Pedone C., Fattorusso R., D'Andrea L.D.

- XXXIII Congresso Nazionale della SCI, Sorrento, Italy, 5-10 July 2009
  Expression and NMR analysis of the interaction between VEGF mimicking peptides and VEGFR-1D2 receptor
  Diana D., Di Stasi R., Ziaco B., Capasso D., Pedone C., D'Andrea L.D., Fattorusso R.

PUBLICATIONS

PUBLICATIONS

- Org Lett. 2008 Apr 15; : 18410123 (P,S,G,E,B,D)
  Semisynthesis of Dimeric Proteins by Expressed Protein Ligation.
  Barbara Ziaco, Soccorsa Pensato, Luca D'Andrea, Ettore Benedetti, Alessandra Romanelli

- Chemistry Eur. J. 2008, 14, 4164 – 4166
  Structural Determinants of the Unusual Helix Stability of a De Novo Engineered Vascular Endothelial Growth Factor (VEGF) Mimicking Peptide
  Donatella Diana, Barbara Ziaco, Giorgio Colombo, Guido Scarabelli, Alessandra Romanelli, Carlo Pedone, Roberto Fattorusso and Luca D. D'Andrea

- Journal of Translational Medicine 2009, 7:41
  In vivo properties of the proangiogenic peptide QK
  Gaetano Santulli, Michele Ciccarelli, Gianluigi Palumbo, Alfonso Campanile, Gennaro Galasso, Barbara Ziaco, Giovanna Giuseppina Altobelli, Vincenzo Cimini, Federico Piscione, Luca Domenico D'Andrea, Carlo Pedone, Bruno Trimarco and Guido Iaccarino

ACKNOWLEDGEMENTS

ACKNOWLEDGEMENTS

This work would not have been completed without the help and the support of many people to which I would like to say thank you.

Primarily, I would like to thanks Prof. Carlo Pedone for giving me the opportunity to work at Istituto di Biostrutture e Bioimmagini (IBB) of Naples. The friendly and supportive atmosphere inherent to whole IBB group contributed essentially to final outcome of my studies.

My sincerest thanks to Ph.D Luca D. D'Andrea for teaching me so much. His ideas and critical manner mode made his supervision very constructive and also very pleasant.

Many thanks to Ph.D Rossella Di Stasi and to other colleagues in the laboratory who have shared with me a lot of the time of my PhD studies. I would like to thanks them for their friendship, for the uncounted fruitful and unfruitful discussions and for the many other funny things we did.