UNIVERSITÀ DEGLI STUDI DI NAPOLI

"FEDERICO II"

FACOLTA' DI FARMACIA

Dottorato di Ricerca in

"Scienza del Farmaco"

XXIV CICLO

"Identification of peptide inhibitors of protein-protein interactions as new potential drugs"

Tutor
Ch.mo Prof.
Paolo Grieco

Co-Tutor
Dr. Daniela Marasco

Candidata
Dott.ssa Pasqualina Liana Scognamiglio

Coordinatore
Ch. ma Prof.ssa
Maria Valeria D'Auria

INDEX

Chapter I

Chapter II

Chapter III

Chapter IV

Chapter V

Chapter VI

CHAPTER I

1. INTRODUCTION

1.1. Protein-protein interaction

Biological systems are made up of very large numbers of different components interacting at various scales. Most genes, proteins and other cell components carry out their functions within a complex network of interactions and a single component can affect a wide range of other components. Protein–protein interactions (PPIs) are essential events of virtually all cellular processes. A cellular interactome represents the whole set of molecular physical interactions between biological entities in cells and its investigation is essential for the comprehension of how gene functions and regulations are integrated within an organism. Protein complexes are involved in supramolecular assemblies (collagens, elastic fibers, actin filaments), in the building of molecular machines (molecular motors, ribosomes, proteasome) participating to major biological processes such as immunity (antigen–antibody interaction), metabolism (enzyme–substrate interaction), signalling (interaction of messenger molecules, hormones, neurotransmitters with their cognate receptors), and gene expression (DNA–protein interactions). Furthermore, the sequencing of the genome and advances in proteomics lead to the identification of proteins of unknown functions (1).

Commonly PPIs are conceived as physical contacts with molecular docking between proteins that occur in a cell or in a living organism in vivo. The investigation of these functional links requires different experimental techniques designed to find such specific types of relationships. Therefore, the definition of PPI has to consider firstly that the interaction interface should be intentional and not accidental, i.e., the result of specific selected biomolecular events/forces; and then that the interaction interface should be non-generic, i.e., evolved for a specific purpose distinct from totally generic functions such as protein production, degradation, and others. Most drugs act by binding to specific proteins to modulate their biological activities, with affects biological processes. The search tool for interaction of chemicals (Stitch, http:// stitch.embl.de/) database (2) contains protein–chemical interaction data for over 68,000 chemicals, including 2200 drugs, and connects them to 1.5 million genes across 373 genomes. A promising application of these networks is to provide information and to formulate hypotheses on human diseases and therapies (drug discovery and targeting) (3).

The importance of PPIs in cell signalling is now better understood and has been the basis of intense research aiming to generate new pharmacological tools to regulate cellular responses. The dynamic formation of selective protein-protein interactions is an essential part of the cellular signal transduction cascades. A powerful rational approach is to map sequences that participate in protein-protein interactions and many studies have demonstrated that peptides corresponding to these sequences act as highly specific inhibitors or activators of signalling events and can serve as selective drugs, in basic research and in animal models of human diseases.

Stabilization and/or disruption of protein-protein interactions and disruption of intra-protein contacts during folding are all being pursued. However, despite some success stories, discovering small-molecule drugs that modulate protein-protein interactions remains an enormous challenge.

Recent advances in peptide and peptidomimetic synthesis (4,5) are generating renewed enthusiasm in the use of peptides as therapeutics, for example, in vaccine development, angiogenesis-related and auto-immune diseases, diabetes and neuroprotection (6-9).

In parallel, the role of short linear motifs in mediating protein-protein interactions in signaling complexes is being unraveled, and the importance of such motifs (and potentially of their peptidic mimics) in signal channeling has been demonstrated.

Cyclic peptides and other modified peptides can be highly potent as well as selective, providing a compromise between structural preorganization and sufficient flexibility to mold to a target surface and maximize binding interactions.

One popular strategy used to search for peptides that inhibit a protein-protein interaction is to mimic the sequence of one of the interacting proteins (10) (*Fig.* 1 B).

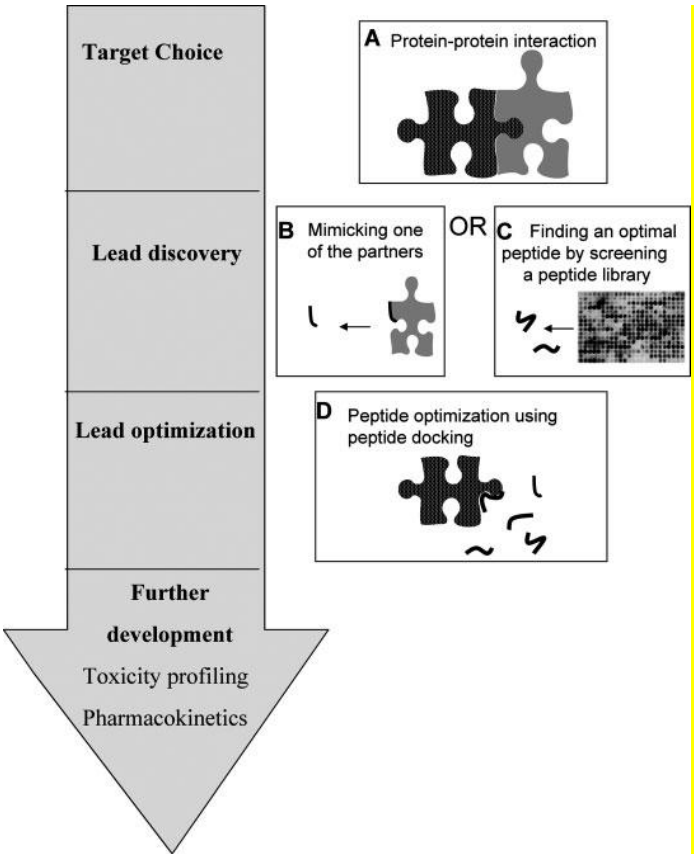


***Fig.* 1: Schematic representation of several approaches in the general context of drug (or modulator of signaling) development.** (**A**) Protein-protein interaction cartoon (**B**) Mimicking the interface sequence of one of the partners by a peptide. (**C**) Finding inhibitory peptides by biochemical screening of peptide libraries. (**D**) Predicting the most probable bound conformation of the peptide in complex with the target protein (peptide docking). This approach can be applied to guide optimization of peptide sequences derived by A or B; or to simultaneously optimize the conformation and the sequence of the peptide

A complementary approach is to derive peptides based on the sequence of the enzyme rather than the substrate. It has been applied with success to several enzymes such as protein kinases and identified well-defined substrate-binding regions consisting of accessible and variable patches distributed between conserved and buried anchors. The variable nature of these regions suggests that they are involved in determining substrate specificity: from these substrate-binding regions several peptides of 7–11 amino acids in length have been derived (11).

The key challenge for computational prediction of protein peptide interactions is the flexibility of the systems under consideration: peptides have many more degrees of freedom than typical small molecules. This flexibility is crucial for the rational design and optimization of peptides and peptidomimetics. In addition, some disruptors of

protein-protein interactions aim to stabilize the inactive form of the protein. Thus, the rational design of such inhibitors must take into consideration the conformational plasticity of the protein and the interplay between different conformations.

Other methods predict the detailed interaction between the protein and the peptide, namely ‘‘peptide docking’’ (*Fig.* 1 D).

As essentially all protein-mediated biological processes are based on specific interactions between proteins and their ligands, the inhibition of underlying molecular interactions is a promising strategy for the controlled interference with biological functions. This approach requires the availability of molecules capable of mimicking conformationally defined binding and/or functional sites of proteins (*Fig.* 1). Furthermore, with respect to the fast growing fields of genomics and proteomics, libraries of molecules, having propensity to mimic conformationally defined protein binding sites, are valuable tools for the investigation and characterization of unknown binding specificities of protein–ligand interactions. The libraries represent copies of protein fragments, enabling the detailed analysis of the interaction at the level of individual amino acid residues. Moreover, when using chemical synthesis methods, diverse modifications, including the incorporation of building blocks other than the proteinogenic amino acids, as well as the alteration of the peptide backbone, can be used to extend the chemical and structural diversity, as well as the metabolic stability, of synthetic peptides.

1.2. Drug Discovery

1.2.1. High-Throughput Screening

High-Throughput Screening (HTS) is a well-established process for lead discovery that comprises the synthesis and the screening of large chemical libraries for activity against biological targets via the use of automated and miniaturized assays and large-scale data analysis. In recent years, high-throughput technologies for combinatorial and multiparallel chemical synthesis and automation technologies for isolation of natural products have tremendously increased the size and diversity of compound collections.

A consequence growing demand for new, highly effective drugs is driven by the identification of novel targets derived from the human genome project and from the understanding of complex protein-protein interactions that contribute to the onset

and maintaining of pathological conditions. To illustrate the dynamics of quantitative and qualitative process approaches to accelerated drug development, an hypothetical pipeline is constructed in *Fig.* 2.

***Fig.* 2**. **Drug Discovery pipelines**. IND: Investigational New Drug, CTA: Clinical Trial Application, NDA: New Drug Application, MAA: Marketing Authorization.

HTS has achieved a dominant role in drug discovery over the past two decades. Its aim is to identify active compounds (hits) by screening large numbers of diverse chemical compounds against selected targets and/or cellular phenotypes. The HTS process consists of multiple automated steps involving compound handling, liquid transfers, and assay signal capture, all of which unavoidably contribute to systematic variation in the screening data. It represents the process of testing a large number of diverse chemical structures against disease targets to identify 'hits'. Screening (HTS) is indeed widely employed to create simultaneously different sets of compounds and rapidly screening them for the identification of active components. This approach is nowadays regarded as a powerful tool for the discovery of new drug candidates, catalysts and materials.

Stimulated by the automation of peptide synthesis on solid support, the concepts now embodied in combinatorial chemistry have completely changed the way in which compound collections are assembled and structure–activity relationships (SARs) are explored. Reactions are accelerated by microwave and other non-traditional techniques; reaction workups are streamlined with solid-phase reagents, and parallel or mix-and-split formats enable chemists to make hundreds of analogues in the time they used to synthesize one or two.

Increases in capacity have been driven by increases in the size of compound libraries, and modern screening technologies have greatly improved the analysis of the interaction of small molecules with novel pharmacological targets. Assay methods are increasingly moving away from the use of radioisotopic labels. Modern

fluorescent technologies now play an important role and are widely employed for the evaluation of more complex target systems.

At the other extreme, there is also interest in identifying low-affinity ligands for pharmacological targets. In fragment-based screening, the goal is to identify pharmacologically active partial structures. If different fragments for adjacent binding sites on the target are found, they can be linked to form ligands with higher affinity. NMR-based techniques have emerged as the most powerful of the methods used to detect low-affinity ligands in biological systems.

In addition to dynamic range and throughput, compound consumption remains one of the major issues in HTS. Miniaturization as well as high-density array formats have not only greatly improved the throughput of HTS campaigns, but also significantly reduced the amount of chemical material needed to determine ligand affinity. The newer screening techniques are able to analyze many different types of ligand–target interactions and have provided research teams with a huge number of hits from HTS campaigns and, in turn, leads for the optimization phase of drug discovery.

Combinatorial methods have fundamentally ameliorated the ability to synthesize and screen large numbers of compounds thanks to improvements made in technology, instrumentation, and library design strategies, particularly in the field of oligonucleotides and peptides (12,13). In combinatorial screening approaches, it is relevant to consider the density, the coverage of chemical space, and the diversity and structural similarity of combinatorial libraries compared with other collections, particularly the known drugs; in detail, the three most important properties are size correlated with molecular weights (MWs), flexibility associated with the presence／absence of rigid scaffolds, and molecular polarity depending on functional groups (14–16). Synthetic mixture–based combinatorial libraries in positional scanning format (17) are collections of compounds systematically arranged to have both defined and randomized positions, which can be synthesized in narrow areas of chemical space, and their screening leads to the acquisition of information regarding specific functional groups at each variable position. These groups determine the activity of a specific chemical scaffold or pharmacophore (15–17). The identification of compounds capable to control protein–protein interactions represents a major goal for modern drug discovery (18–20). Both structure-based design and random screening of combinatorial collections have been successfully

applied, in this instance, and often small peptides and peptidomimetics have been selected as outstanding inhibitors of protein complexes (21–24). In many cases, the lack of sufficient structural information on protein-interacting regions renders the combinatorial approach the most suitable method, but the use of completely random peptide libraries often produces redundant chemical species that complicate the deconvolution phases of screening and puzzles the way to the identification of active compounds (25). Several methods for the generation of peptide mixtures have been reported since the introduction of the original idea of Furka and coworkers, in 1991. However most of them are based on two basic methodologies, known as the *Mix and Split* and the *Pre-mix* approaches. The Mix and Split (or portioning-mixing) method is largely accepted as the first synthetic approach used to create combinatorial mixtures of compounds. It consists of three basic steps: splitting, coupling and mixing. Firstly, resin beads (or other microreactors) are split in a number of aliquots that equals the number of building blocks to be utilized in the synthesis (splitting). Then, to each resin aliquot one building block is coupled and the reactions are driven to completion (coupling). Following this step, beads are randomly and thoroughly mixed (mixing) and then re-splitted in the same number of aliquots, achieving a set of homogeneous and equimolar collections of compounds. Repetitive execution of these basic steps for “n” times will thus produce a rapid increase of newly generated molecules, while bead number remains constant. Importantly, since resin beads encounter one reactant at a time, no more than one compound per bead can be generated (OBOC), therefore the total number of final compounds is limited by the number of initial beads. To overcome the limitations posed by OBOC approach and thus to obtain larger libraries, the synthesis of peptide mixtures using pre-mixed solutions of peptides was introduced. By this approach, the amino acids chosen for the library are pre-mixed and then coupled to a single resin batch. Before coupling the last amino acid, the resin is split in a number of aliquots that equals the number of building blocks and to each aliquot, one single amino acid is separately coupled. This procedure leads to the generation of sub-libraries that are labelled by the N-terminal residue. Every single bead contains all the ensemble of peptides of the sub-library (26), as any single bead encounters the reagents under the same conditions. This method is faster than the Mix and Split and allows the preparation of very large libraries, thus producing a wide chemical diversity.

Deconvolution methods evaluate the contribution of each residue to the desired biological activity. There are two main deconvolution approaches: the *iterative process* (27) or the *positional scanning method* (17). In the first approach, a progressive elucidation of the active sequence is carried out by identifying one amino acid at a time for each position. The process involves subsequent cycles of screening and re-synthesis of active pools, thus producing a progressive activity enrichment and a contextual semplification of the peptide pools. The chemical synthesis of the next sublibraries is dictated by the results from the previous one, thus only a sub-population of the library components is re-synthesized, arriving at the last assay where single molecules, differing only for the last amino acid are in parallel prepared and tested. In the positional scanning (PS) approach, all sublibraries needed to determine the final active sequence are prepared at the beginning of the screening and, in this case, the library is replicated a number of times that equals the sequence length. Each variable position is tested independently, at the same time, using a set of sub-libraries with known residues for every individual position. The most active sequence is finally "read" at the end of the process by selecting the most active pools on each known position.

The synthesis of completely random combinatorial libraries of peptides generates a large number of "quasi-duplicates" deriving from the strong similarity between several side chains. In large libraries, the need to manage large arrays of tubes and codes can puzzle the way to the identification of lead compounds and slows down the synthetic and deconvolution steps. In the perspective of simplifying the synthesis and deconvolution procedures without grossly affecting the probability to find out active peptides, a recent approach regards the reasoning on the general properties of L and D-amino acids, reaching a compromise between the need to maintain the highest possible diversity and that of reducing the number of building blocks. These new libraries have been named "Simplified Libraries", intending with this any new ensemble of possible sequences achievable with a reduced and non redundant set of amino acids. The new set of "non redundant" different residues (12 instead of 20) is chosen following several simple rules. First of all "quasi-identical" groups of amino acids are selected: Asp and Glu; Asn and Gln; Ile and Val; Met and Leu; Arg and Lys; Gly and Ala; Tyr and Trp; Ser and Thr. Then, from each group, one is selected trying to keep a good distribution of properties, including hydropathic properties, charges, pKa, aromaticity. The cysteine is converted to the stable

acetamidomethyl-derivative to prevent polymerization and to increase the number of polar residues within the final set. The distribution of residue molecular weights is also an important parameter in cases of deconvolution of the library by mass spectrometry approaches (28).

1.2.2. Rationally Designed Peptide Inhibitors

Protein-protein interactions are known to play a critical role in the normal function of cellular/organelle structure, immune response, protein enzyme inhibitors, signal transduction, and apoptosis. Rational approaches toward the recognition of protein surfaces may provide better insights into exactly how proteins interact with one another, and is an alternative to enzyme inhibitor design as a molecule-based disease therapy. Moreover, rational protein surface recognition is a challenging test of our knowledge of molecular design. The nature of protein-protein interfaces has been the focus of investigation for some time. Particular emphasis in the literature has been placed on structure, energetics, electrostatic complementarity, and kinetics of protein-protein interactions. An important breakthrough has been the identification of “hot spots” on protein surfaces. A hot spot is a defined locale of ca. 600 $Å^2$ on the surface of a protein or near the geometric center of the protein-protein interface. The residues that comprise the hot spot contribute significantly to the stability of the protein-protein complex. Mutating a hot spot residue to alanine results in a lower affinity of the protein for its partner.

The main goals in designing synthetic molecules that disrupt protein-protein interactions through specific recognition take two complementary threads. One is the recognition of protein surfaces in ways that agonize a biological response, mediate some protein dimerization, or stabilize the native oligomerization state of the protein. The second objective in targeting a protein surface would be to selectively bind half of a dimeric interaction in a way that would sterically block association with the natural protein partner.

The practice of molecular design relies on a full understanding of non covalent forces. Often, many of these forces (hydrogen bonding, electrostatics, van der Waals forces, π-π interactions, conformational energy, etc.) work in concert, making it difficult to isolate the contributions of any one component. Model systems can be used to illustrate the contributions of these forces in a systematic manner. The primary goal in this research is the design of molecules with predictable recognition

properties. Among the greatest tests of the knowledge gained from model systems is their application to a natural target. Here the requirements for specificity and tight binding are more rigorous than in the model systems. Protein surfaces represent a challenging target for the field of molecular design. Large surface areas (500-2000 $Å^2$) must be covered and multiple interactions utilized to achieve sufficient binding affinity to observe an effect. Oligopeptides represent an intermediate step toward the recognition of protein surfaces for the field of molecular design. Short peptide sequences (3-20 amino acids) are themselves worthwhile targets for recognition due to their potential applications in separation, diagnostic, or biological areas. The notion of peptide recognition as a stepping stone to protein binding arises because many of the early synthetic receptors involved a small number of functional group interactions. Binding a protein surface should require large arrays of interacting functionality in a conformationally defined context. Also, these arrays may not be confined to one domain or secondary structural element of the protein. Short peptides present an additional factor to binding because they are themselves usually conformationally flexible (29).

1.2.3. Peptides as Drugs

The advantages of peptide-based therapeutics derive from the underlying physics, chemistry and modularity of amino acids combined with the physical and chemical nature of the amide bond and the fact that nature has solved numerous physiological and pharmacological problems using peptides. These advantages include: the capacity for high target affinity, specificity and potency; predictable and relatively low cost and easy synthesis and storage; low and predictable toxicity, and reduced antigenicity. The advantage of low and predictable toxicity is crucial for small molecule drugs that often fail due to human toxicity. Similarly, due to their relatively small size peptide-based drugs are often superior to protein-based therapeutics. For example, immunogenicity, intracellular penetration, and synthesis and storage represent challenges for proteins that are not nearly as daunting as for peptides.

The major challenges confronting the successful development of peptide drugs appear to be the development of successful delivery strategies and the development of optimal pharmacokinetics. But small molecules and proteins must solve these problems, too. Moreover, the study of natural peptides and proteins,

and recent interest in the area of peptide drug discovery indicate that peptides can be engineered for oral or other noninvasive routes of administration and to have favorable pharmacokinetic profiles. Indeed, recent investigations have explored the possibility of improving small molecule drug bioavailability through conjugation with small peptides that are the ligands of specific gut transporters (30).

Protein-protein interactions, however, often involve the burial of large and relatively featureless contacting surfaces and are typically mediated by numerous and spatially distributed hydrogen bonding and ionic interactions that can involve sub-nanomolar binding affinities. This, in part, explains why a herculean effort is sometimes required to develop small molecule inhibitors of protein-protein interactions and why, despite a few notable exceptions, small molecules might prove to be poor choices for targeting such interactions (31). A potentially more fruitful and less costly strategy is to develop peptide inhibitor leads from knowledge of how natural proteins and peptides interact. These peptide leads could then be developed into peptide-based or peptidomimetic compounds with superior drug profiles. Indeed viable peptide candidates, approximately five residues in length, can be obtained from protein-protein interfaces to serve as templates for peptide-based, peptidomimetic and even small molecule drugs that inhibit protein-protein interactions (32).

An equally ‘natural’ strategy for peptide drug discovery but that contemplates the experimental screening of larger peptides has been outlined by Watt (33). Hence, developing peptide-based or peptide-derived drugs for the therapeutic modulation of protein-protein interactions appears to make good scientific and business sense.

A second specific advantage of peptide-based is that peptide ligands can be designed that are more robust to mutational changes at protein binding or active sites. The hypothesized advantage derives from the potential to design peptides that are flexible, that bury relatively large surface areas on binding, that interact with proteins through a large number of spatially distributed contacts, and that interact with target proteins predominantly through main-chain interactions (34). Presumably, a peptide ligand that combined all of these binding properties would retain significant binding, despite the loss or substitution of various side-chains through mutation. There is thus a sound theoretical basis and some suggestive evidence for speculating that peptide drugs can be designed to be more refractory

to the effects of mutations in the target protein. This could have enormous implications for the design of antibacterial and anti-viral agents..

REFERENCES

1. Chautard E., Thierry-Mieg N., Ricard-Blum S. (2009). Interaction networks: From protein functions to drug discovery. A review. Pathologie Biologie 57: 324–333
2. Kuhn M, von Mering C, Campillos M, Jensen LJ, Bork P. STITCH. (2008) interaction networks of chemicals and proteins. Nucleic Acids Res;36(Database issue):D684–8.
3. Ideker T, Sharan R. Protein networks in disease. (2008) Protein networks in disease. Genome Res;18:644–52.
4. Goodman, C. M.; Choi, S.; Shandler, S.; DeGrado, W. F. Nat Chem Biol 2007, 3, 252–262.
5. Vagner, J.; Qu, H. C.; Hruby, V. J. Curr Opin Chem Biol 2008, 12, 292–296.
6. Sulochana, K. N.; Ge, R. Curr Pharm Des 2007, 13, 2074–2086.,
7. Yusuf-Makagiansar, H.; Anderson, M. E.; Yakovleva, T. V.; Murray, J. S.; Siahaan, T. J. Med Res Rev 2002, 22, 146–167.,
8. Jain, R.; Chawrai, S. Mini-Rev Med Chem 2005, 5, 469–477.,
9. Gozes, I.; Divinski, I. Curr Alzheimer Res 2007, 4, 507–509.
10. Peczuh, M. W.; Hamilton, A. D. Chem Rev 2000, 100, 2479–2493.
11. .Niv, M. Y.; Rubin, H.; Cohen, J.; Tsirulnikov, L.; Licht, T.; Peretzman-Shemer, A.; Cna'an, E.; Tartakovsky, A.; Stein, I.; Albeck, S.; Weinstein, I.; Goldenberg-Furmanov, M.; Tobi, D.; Cohen, E.; Laster, M.; Ben-Sasson, S. A.; Reuveni, H. J Biol Chem 2004, 279, 1242–1255.
12. Frank, R., Heikens, W., Heisterberg-Moutsis, G., and Blocker, H. (1983) A new general approach for the simultaneous chemical synthesis of large numbers of oligonucleotides: segmental solid supports. Nucleic Acids Res 11: 4365-4377.
13. Houghten, R.A. (1985) General method for the rapid solid-phase synthesis of large numbers of peptides: specificity of antigen-antibody interaction at the level of individual amino acids. Proc Natl Acad Sci U S A 82: 5131-5135.
14. Balakin, K.V., Kozintsev, A.V., Kiselyov, A.S., Savchuk, N.P. (2006) Rational design approaches to chemical libraries for hit identification. Curr Drug Discov Technol 3:49-65.
15. Ecker, D.J., and Crooke, S.T. (1995) Combinatorial drug discovery: which methods will produce the greatest value? Biotechnology (N Y) 13:351-360.
16. Kennedy, J.P., Williams, L., Bridges, T.M., Daniels, R.N., Weaver, D., and Lindsley, C.W. (2008) Application of combinatorial chemistry science on modern drug discovery. J Comb Chem 10: 345-354.
17. Pinilla, C., Appel, J.R., Blanc, P., and Houghten, R.A. (1992) Rapid identification of high affinity peptide ligands using positional scanning synthetic peptide combinatorial libraries. Biotechniques 13: 901-905.
18. Betzi, S., Restouin, A., Opi, S., Arold, S.T., Parrot, I., Guerlesquin, F., Morelli, X., and Collette, Y. (2007) Protein protein interaction inhibition (2P2I) combining high throughput and virtual screening: Application to the HIV-1 Nef protein. Proc Natl Acad Sci U S A 104: 19256-19261.
19. Wells, J.A., and McClendon, C.L. (2007) Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. Nature 450, 1001-1009.

20. Chene, P. (2006) Drugs targeting protein-protein interactions. ChemMedChem 1: 00-411.
21. Shultz, M.D., Ham, Y.W., Lee, S.G., Davis, D.A., Brown, C., and Chmielewski, J. (2004) Small-molecule dimerization inhibitors of wild-type and mutant HIV protease: a focused library approach. J Am Chem Soc 126: 9886-9887.
22. Zhao, H.F., Kiyota, T., Chowdhury, S., Purisima, E., Banville, D., Konishi, Y., and Shen, S.H. (2004) A mammalian genetic system to screen for small molecules capable of disrupting protein-protein interactions. Anal Chem 76: 2922-2927.
23. Ponticelli, S., Marasco, D., Tarallo, V., Albuquerque, R.J., Mitola, S., Takeda, A., Stassen, J.M., Presta, M., Ambati, J., Ruvo, M., and De Falco, S. (2008) Modulation of angiogenesis by a tetrameric tripeptide that antagonizes vascular endothelial growth factor receptor 1. J Biol Chem 283: 34250-34259.
24. Sillerud, L. O., Larson, R.S. (2005) Design and structure of peptide and peptidomimetic antagonists of protein-protein interaction.Curr Protein Pept Sci. 6:151-69.
25. Martin, E. J., Blaney J. M, Siani, M. A., Spellmeyer, D. C., Wong, A. K., Moos W. H. (1995) Measuring Diversity: Experimental Design of Combinatorial Libraries for Drug Discovery .J. Med. Chem.,38: 1431-1436.
26. Geysen H.M. (1986) Mol. Immunol., 23, 709.
27. Houghten R.A., Pinilla C., Blondelle S.E., Appel, J.R., Dooley T.C., Cuervo J.H. (1991) Nature, 354, 84.
28. Marasco, D., Perretta, G., Sabatella, M., and Ruvo, M. (2008) Past and future perspectives of synthetic peptide libraries. Curr Protein Pept Sci 9 : 447-467.
29. Torneiro, M.; Still, W. C. J. Am. Chem. Soc. 1995, 117, 5887- 5888.
30. Nielsen CU, Brodin B. Di/tri-peptide transporters as drug delivery targets: regulation of transport under physiological and pathophysiological conditions. Curr Drug Targets 2003; 4(5): 373-88
31. Sawyer TK. AILERON Therapeutics. Chem Biol Drug Des 2009; 73(1): 3-6.
32. Sillerud LO, Larson RS. Design and structure of peptide and peptidomimetic antagonists of protein-protein interaction. Curr Protein Pept Sci 2005; 6(2): 151-69.
33. [Watt PM. Screening for peptide drugs from the natural repertoire of biodiverse protein folds. Nat Biotechnol 2006; 24(2): 177-83.
34. Rao GS, Bhatnagar S. In silico structure-based design of a potent, mutation resilient, small peptide inhibitor of HIV-1 reverse transcriptase. J Biomol Struct Dyn 2003; 21(2): 171-8.

CHAPTER II

1. INTRODUCTION

1.1. Diabetes mellitus

Obesity, as metabolic syndrome and diabetes mellitus are actually frequent disorders affecting the Western civilization. Diabetes is one of the most common disorders characterized by hyperglycemia resulting from defects in insulin secretion, action, or both. The chronic hyperglycemia of diabetes is associated with long-term damage, dysfunction, and failure of various organs, especially the eyes, kidneys, nerves, heart, and blood vessels.

Several pathogenic processes are involved in the development of diabetes. These range from autoimmune destruction of the β-cells of the pancreas with consequent insulin deficiency to abnormalities that result in resistance to insulin action. The basis of the abnormalities in carbohydrate, fat, and protein metabolism in diabetes is a deficient action of insulin on target tissues. Irregular insulin action results from inadequate insulin secretion and/or diminished tissue responses to insulin at one or more points in the complex pathways of hormone action.

Symptoms of marked hyperglycemia include polyuria, polydipsia, weight loss, sometimes with polyphagia, and blurred vision. Impairment of growth and susceptibility to certain infections may also accompany chronic hyperglycemia. Acute, life-threatening consequences of uncontrolled diabetes are hyperglycemia with ketoacidosis or the nonketotic hyperosmolar syndrome (1)

Long-term complications of diabetes include retinopathy with potential loss of vision; nephropathy leading to renal failure; peripheral neuropathy with risk of foot ulcers, amputations, and Charcot joints; and autonomic neuropathy causing gastrointestinal, genitourinary, and cardiovascular symptoms and sexual dysfunction. Patients with diabetes have an increased incidence of atherosclerotic cardiovascular, peripheral arterial, and cerebrovascular disease. Hypertension and abnormalities of lipoprotein metabolism are often found in people with diabetes.

According to the classification recommended by the American Diabetes Association (2010), diabetes mellitus can be classified as type 1 (T1DM), type 2 (T2DM), other specific types, and gestational diabetes mellitus.

In type 1 diabetes, the cause is an absolute deficiency of insulin secretion due to autoimmune or idiopathic destruction of insulin-producing β-cells in the pancreas by CD4 and CD8 T cells and macrophages infiltrating the islets. Insulin injection is

always necessary for survival because of the complete lack of endogenous insulin (2). In type 2 diabetes, the most common, the cause is a combination of insulin deficiency, reduced insulin action, and insulin resistance of glucose transport in skeletal muscle and adipose tissue. The manifestation of type 2 diabetes is a continuum of insulin resistance culminating in the failure of insulin secretion to compensate for insulin resistance, and the progression to full diabetes ensues when pancreatic β-cell hypersecretion of insulin fails to compensate for insulin resistance (3) (*Fig*. 1.1).

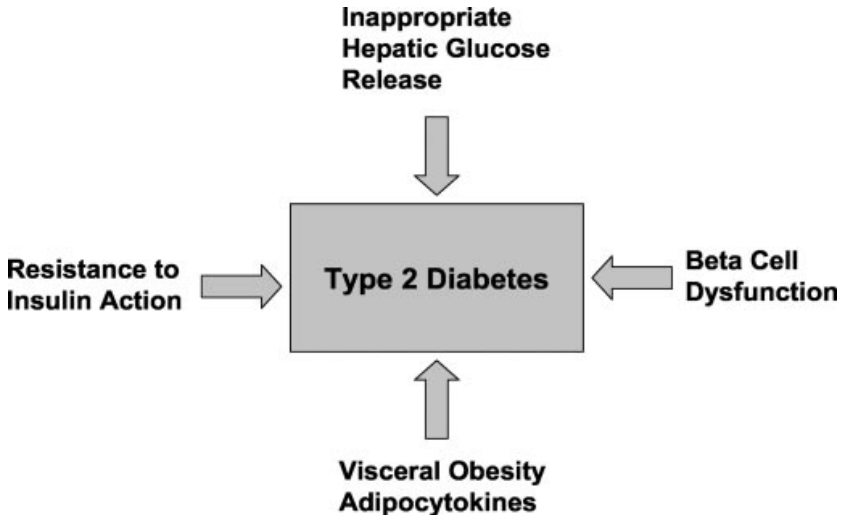


***Fig.* 1.1: Pathophysiology of T2DM**

1.1.1. Insulin signaling and pathway in insulin actions

Insulin is the most potent anabolic hormone known and is essential for appropriate tissue development, growth, and maintenance of whole-body glucose homeostasis. It is secreted by the β-cells of the pancreatic islets of Langerhans in response to increased circulating levels of glucose and amino acids after a meal. Insulin regulates glucose homeostasis at many sites by reducing hepatic glucose output (gluconeogenesis and glycogenolysis) and by increasing the rate of glucose uptake in skeletal muscle and fat to regulate blood glucose concentration (*Fig.* 1.2). In liver and fat cells, insulin also promotes the synthesis and storage of carbohydrates,

lipids and proteins, inhibits their degradation and release into the circulationattenuating fatty acid release from triglycerides in fat and muscle (4).

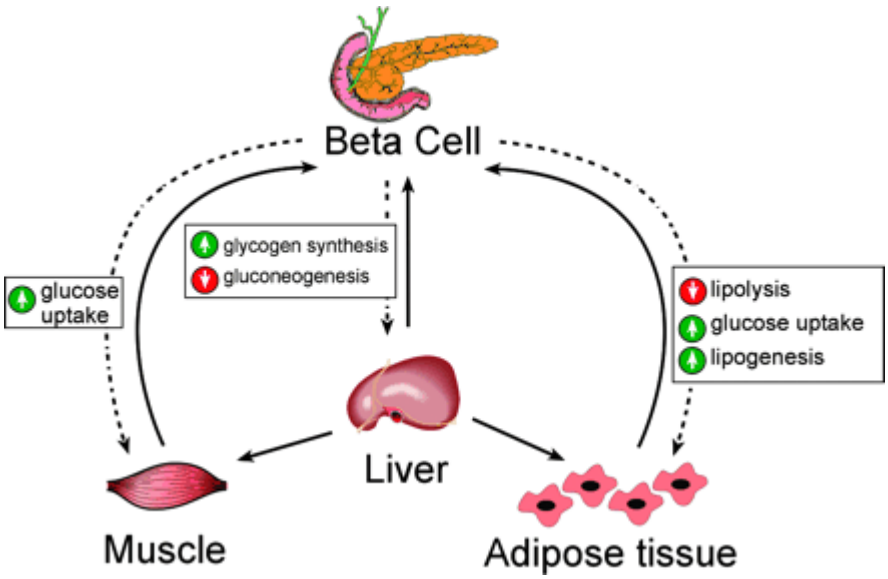


***Fig*.1.2: Glucose homeostasis**

The first step by which insulin increases energy storage or utilization involves the regulated transport of glucose into the cell, mediated by the glucose transporter GLUT4. Insulin stimulates the translocation of a pool of GLUT4 to the plasma membrane, through a process of regulated recycling, in which endocytosis, sorting into specialized vesicles, exocytosis, tethering, docking, and fusion are tightly regulated (5). The precise intracellular events that mediate insulin action are initiated through the binding to and activation of its cell-surface receptor, the insulin receptor substrates (IRS), and the molecules that interact with these substrates (*Fig.* 1.3).

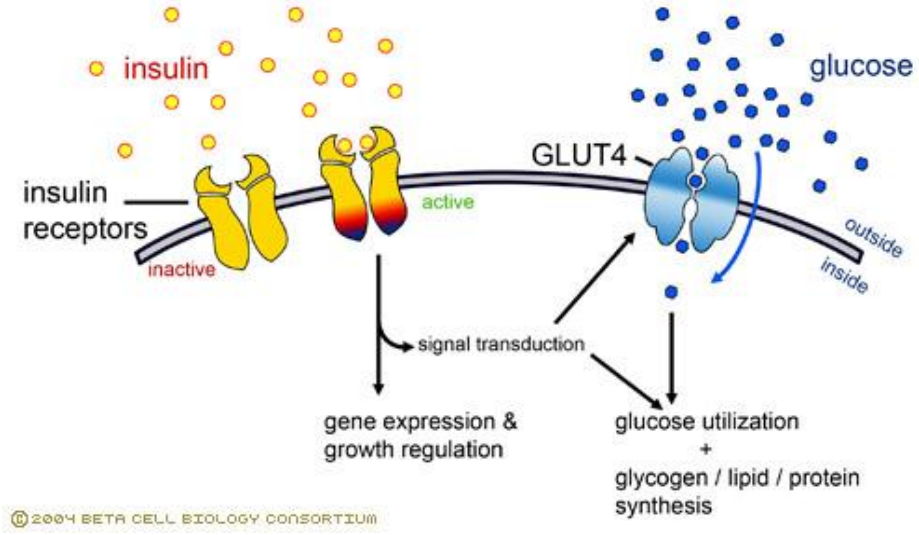


***Fig.* 1.3: Insulin signal transduction**

1.1.2. Insulin resistence

Insulin resistance (IR) is a physiological condition where the natural hormone insulin becomes less effective at lowering blood sugars. The resulting increase in blood glucose may raise levels outside the normal range and cause adverse health effects, depending on dietary conditions. Certain cell types such as fat and muscle cells require insulin to absorb glucose. When these cells fail to respond adequately to circulating insulin, blood glucose levels rise. The liver helps regulate glucose levels by reducing its secretion of glucose in the presence of insulin. This normal reduction in the liver's glucose production may not occur in people with insulin resistance (6).

Insulin resistance in muscle and fat cells reduces glucose uptake (and also local storage of glucose as glycogen and triglycerides, respectively), whereas insulin resistance in liver cells results in reduced glycogen synthesis and storage and a failure to suppress glucose production and release into the blood. Insulin resistance normally refers to reduced glucose-lowering effects of insulin. However, other functions of insulin can also be affected. For example, insulin resistance in fat cells reduces the normal effects of insulin on lipids and results in reduced uptake of circulating lipids and increased hydrolysis of stored triglycerides. Increased mobilization of stored lipids in these cells elevates free fatty acids in the blood plasma. Elevated blood fatty-acid concentrations (associated with insulin resistance

and diabetes mellitus Type 2), reduced muscle glucose uptake, and increased liver glucose production all contribute to elevated blood glucose levels. High plasma levels of insulin and glucose due to insulin resistance are a major component of the metabolic syndrome. If insulin resistance exists, more insulin needs to be secreted by the pancreas. If this compensatory increase does not occur, blood glucose concentrations increase and type 2 diabetes rises up.

1.1.3. Mechanisms of Insulin resistence

Insulin resistance may arise through a variety of mechanisms (*Fig*. 1.4) can be caused by structural or functional changes in any of the insulin regulatory actions on intracellular signalling mediator molecules, from the insulin receptor to the genes (7). The defect of glucose utilization may therefore be induced in any phase of the signalling process: from the binding of insulin to its receptor, through the activity of GLUT transporters and of proteins responsible for the translocation of these transporters to the cellular membrane, to the structural or functional abnormalities of other signaling molecules, and transcriptional factors on the route of the insulin signal to genes. Disorders at the level of insulin receptor have been thoroughly examined, especially the mutations disturbing its function as a tyrosine phosphatase.

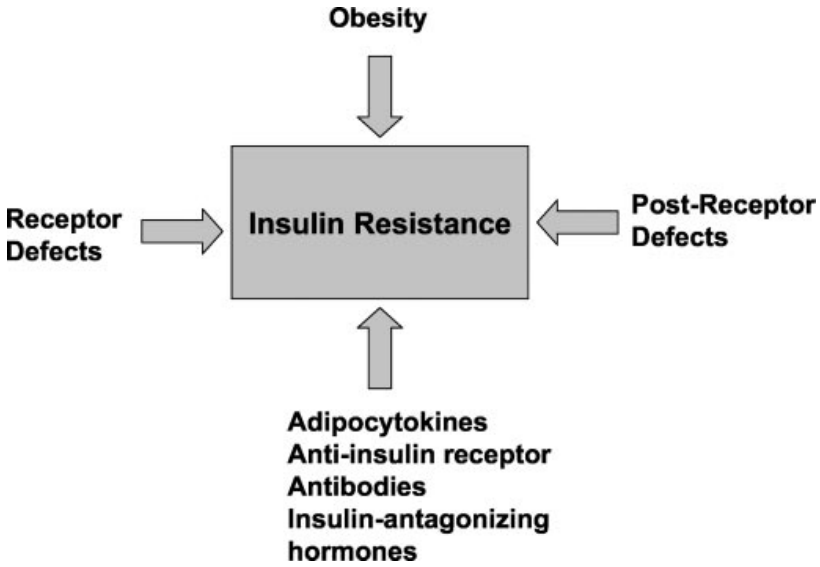


***Fig.* 1.4: Mechanisms connected to insulin resistance.**

The utilization of glucose by the muscles and adipose tissue can be impaired by many other disturbances in the signalling processes. In patients with type 2 diabetes, different abnormalities concerning the signal transduction cascade below the insulin receptor have already been found. For example, in patients with insulin resistance (not necessarily with type 2 diabetes), impaired IRS 1 phosphorylation in skeletal muscles is frequently encountered (8). In the adipose tissue of patients with type 2 diabetes mellitus and obesity, a decrease in glucose transport is connected with changes in the expression of the GLUT4 protein. In this way, a decrease of the glucose transport contributes to hyperglycaemia. The “transport” hyperglycaemia in diabetes mellitus may arise in two ways: 1) as a result of a decrease in the GLUT transporters’ expression and concentration, or 2) as an effect of functional disturbances in the process of the GLUT transporters’ intercellular translocation. Studies of the impairment of cellular glucose transport in diabetes mellitus, particularly in type 2, should now be seen as a new potential sphere of clinical activity (9). In abdominal obesity — after normal body mass is restored the insulin receptor regains its lost tyrosine phosphatase activity. Disorders of the expression and translocation of glucose transporter GLUT4 could be normalized and insulin sensitivity restored.

1.2. The Role of PED/PEA-15 in Glucose Tolerance and in Type 2 Diabete

Several studies aimed at the identification of genetic variants associated with type 2 diabetes and to understand their significance to and their role in the disease. This effort led to the identification of a number of diabetes susceptibility loci, which have been widely replicated in different and large populations (10). In expression profile studies, *phosphoprotein enriched in diabetes/ phosphoprotein enriched in astrocytes-15* (*ped/pea-15*) emerged as overexpressed a gene in adipose and skeletal muscle tissues and in skin fibroblasts from type 2 diabetic individuals independently of obesity (11) and severalstudies investigate the mechanisms that may lead to dysregulating the expression of this gene in type 2 diabetic patients and in their first-degree relatives.

In a initial study by Condorelli et al. (11), it was shown that PED/PEA-15 overexpression represents a common trait among type 2 diabetics, because approximately one-third of the individuals diagnosed with the disorder exhibit

PED/PEA-15 expression levels higher than 2 SD above the mean in the control subjects (12). Functional analysis in cultured skeletal muscle and adipose cells demonstrated that increased expression of PED/PEA-15 causes resistance to insulin action in glucose uptake (11). This abnormality depends largely on PED/PEA-15-induced dysfunction of the Protein Kinase C (PKC)signaling system. Indeed, several studies evidenced that PED/PEA-15 binds phospholipase D (PLD) isoforms, increasing their stability and intracellular diacylglycerol levels (13), thereby activating the dyacylglycerol-sensitive PKCα isoform. These studies also revealed that the active PKCα in PED/PEA-15-overexpressing cells and tissues prevents insulin induction of the PKCζ isoform. This is distinct from α and serves as a major activator of glucose transporter 4 (GLUT4) vesicle translocation toward the plasma membrane. Thus, studies in isolated muscle and adipose cells have shown that the overexpression of PED/PEA-15 may determine peripheral resistance to insulin action by dysregulating the signaling of PKC (*Fig*. 1.5) (14).

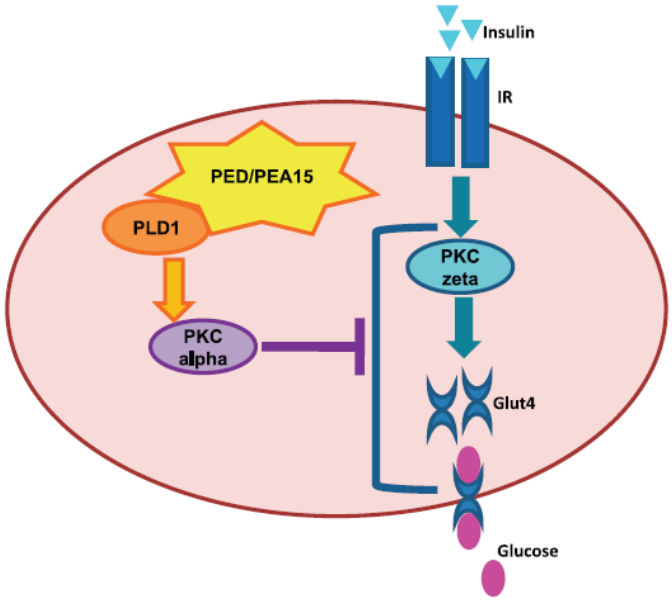


***Fig*.1.5: Molecular mechanism of action of PED/PEA-15 in Insuline Resistance**

To assess the consequences of *PED/PEA-15* overexpression on glucose tolerance, Vigliotta et al. (15) generated transgenic mice overexpressing *ped/pea-15* ubiquitously, because this same situation occurs in many individuals with type 2 diabetes. Phenotyping these mice revealed that overexpression of *PED/ PEA-15*

impairs glucose tolerance and may lead to diabetes under appropriate environmental conditions. Indeed, upon high-fat diet feeding, transgenic mice become overtly diabetic. The impaired glucose tolerance observed in the *PED/PEA-15* transgenic mice was accompanied by significant resistance to insulin action on glucose disposal, reduced insulin-stimulated glucose transport in fat and skeletal muscles, and inhibition of insulin effect on GLUT4 membrane translocation. The mechanistic significance of these findings has been further corroborated by additional studies in *ped/pea-15*-null mice. In these animals, ablation of even a single copy of the *ped/pea-15* gene causes a remarkable increase in insulin sensitivity, which is paralleled by enhanced PKCζ activity upon feeding. Thus, the additional concept has emerged that the control of insulin sensitivity is a physiological function of the *PED/PEA-15* gene.

1.2.1. PED/PEA-15 protein

The Phosphoprotein Enriched in Diabetes/ Phosphoprotein Enriched in Astrocytes-15 (PED/PEA-15) is a 15 kDa cytosolic protein widely expressed in different tissues and highly conserved among mammals, whose gene maps on human chromosome 1q21-22 (16). The structure of PED/PEA-15 includes a canonical N-terminal Death Effector Domain (DED) of 80 amino acids that regulate apoptotic signaling pathways (*Fig.* 1.6 A). Within the N-terminus there is a nuclear export sequence (NES) that mediate its cytoplasmic localization. In particular, PED/PEA-15 regulates the actions of ERK-MAP kinase cascade by binding to ERK in the nucleus, exporting it into the cytoplasm and preventing the entrance into the cell cycle caused by sustained phospho-ERK nuclear accumulation (17). In the C-terminal domain, PED/PEA-15 protein can be phosphorylated on two different residues: the Ser104, identified as a site of regulation for protein kinase C (PKC) (16), and the Ser116 regulated by calcium–calmodulin kinase 2 (CaMKII) (18) and PKB/Akt (19).

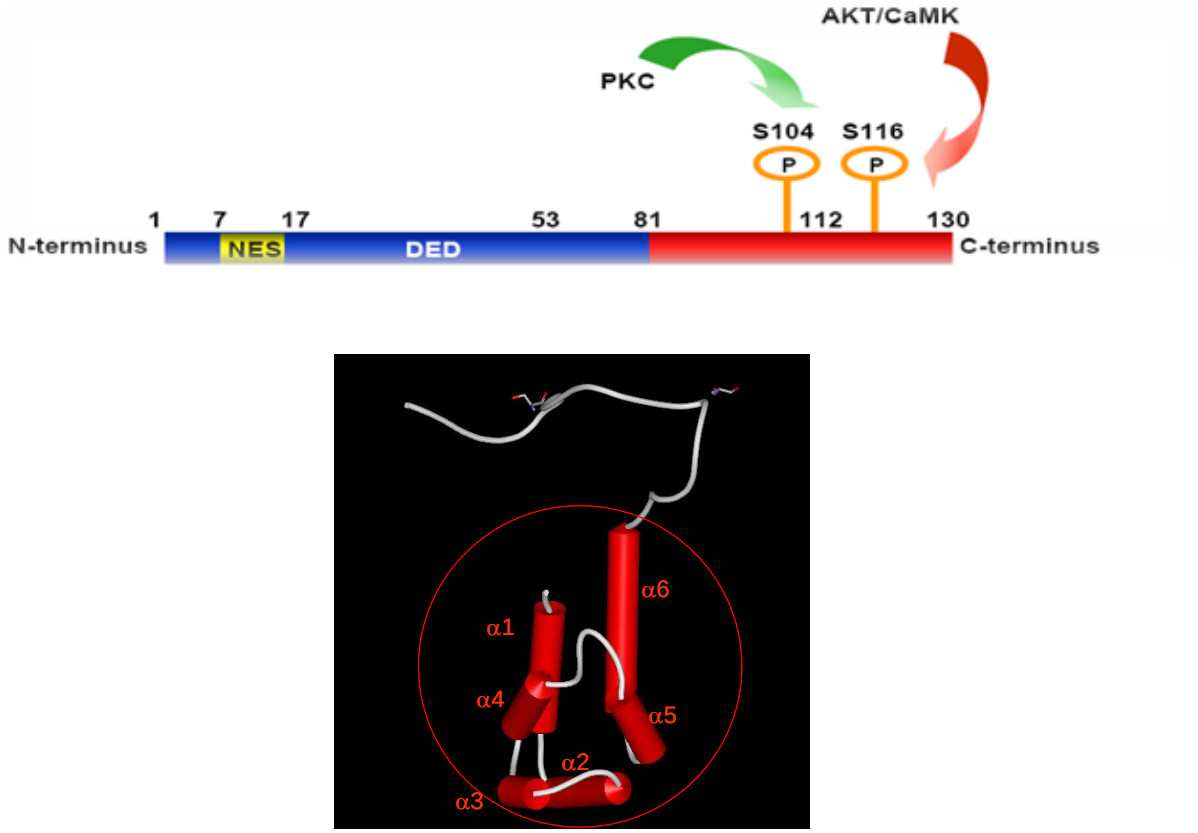


***Fig.*1.6: Schematic representation of modular structure of PED/PEA-15 (upper) and 3D-Structure of PED/PEA-15 protein determined by NMR (lower)**

The PED protein structure, shown in *Fig.* 1.6 B, was determined by NMR spectroscopy (20); it shows the presence in the region N-terminal of a DED domain (Death Effector Domain) while the C-terminal portion appears to be quite unstructured. The DED domain belongs to the superfamily of so-called "death domain" (Death Domains), which also includes the domains DD (Death Domain), the CARD (Caspase Recruitment Domain) and PYRIN domains; they act as primary mediators in the protein-protein interactions necessary for the transmission and regulation of the apoptotic signal, and both regions appear to be necessary to make PED interaction with ERK proteins. The DED domain is composed of 6 antiparallel amphipathic α-helices grouped around a central hydrophobic core, such α-helices in the DED domain are linked together by small loops, two of which (α2-α3 and α4-α5) contain short β-structures and help to take the domain to a spatial arrangement as a "Greek cross", typical of DD domains, with the helices α1 and α2 localized at the center, and the helices α3 and α4 on one side and the α5 and α6 on the other one

(*Fig.* 1.6 B). This structure is very similar to those reported for other DED-containing proteins: the protein FADD (21) and the FLIC-like inhibitory protein (FLIP) (22), whose structures are known at the atomic level by X-ray diffraction. The DED domain present on its surface several charged residues that suggest different interactions of electrostatic nature are responsible for the recognition of ERK1/2. The protein contains two other important sites for regulation: a domain of phosphorylation by PKC (residues 99-107) and a phosphorylation site of type II calcium calmodulin-dependent (110-122). The antiapoptotic function is regulated by phosphorylation of residues Ser104 and Ser116.

The amino acid sequence of the protein PED/PEA15 is shown in *Fig.* 1.7 while Table 1.1 shows the protein segments that include the six alpha helices of the DED.

^{1}MAEYGTLLQDLTNNITLEDLEQLKSACKEDIPSEKSEEITTGSAWFSFLE^{50}
^{51}SHNKLDKDNLSYIEHIFEISRRPDLLTMVVDYRTRVLKISEEDELDTKLT^{100}
^{101}RIPSAKKYKDIIRQPSEEEIIKLAPPPKKA^{130}

***Fig*.1.7: Primary sequence of h-PED/PEA15**

Segment	Helix	Sequence
PED 2-14	α1	AEYGTLLQDLTNN
PED 17-27	α2	LEDEQLKSAC
PED 33-37	α3	SEKSE
PED 42-51	α4	GSAWFLESHN
PED 61-69	α5	SYIEHIFE
PED 73-89	α6	PDLLTMVVDYRTRVLKI

Table 1.1: Helical regions of PED/PEA15 as assessed by NMR study

In addition, this protein is a interactor of phospholipase D1 (PLD1) isoforms, which increases stability and therefore their activities. This effect is accompanied by changes of some metabolic mechanisms (23).

1.2.2. Phospholipase D protein

As already reported, PED/PEA-15 interaction with PLD may represent a crucial event in dysregulating insulin action (15). Classical phospholipase D (PLD) enzymes hydrolyze the principal membrane lipid phosphatidylcholine (PC) to generate phosphatidic acid (PA) and free choline (Cho) and performs a transphosphatidylation reaction using water or primary alcohols (ethanol or 1-butanol) as the nucleophile to generate PA, phosphatidylethanol or phosphatidylbutanol (PBut), respectively (24). (*Fig*. 1.8)

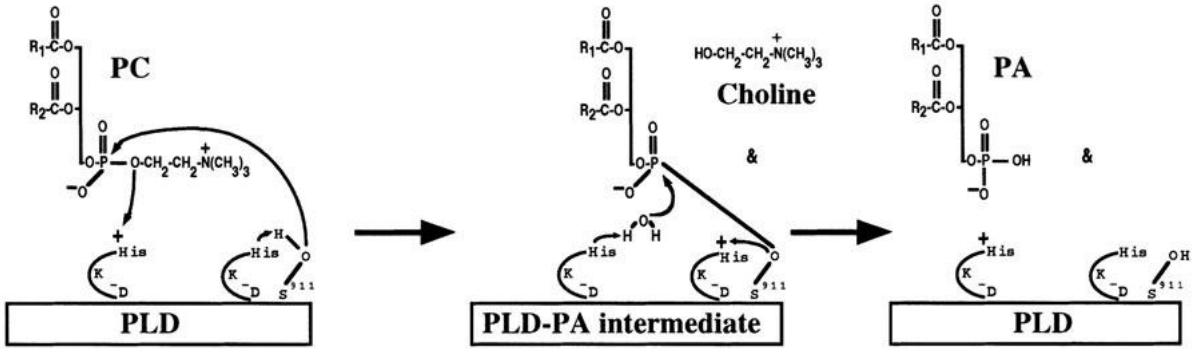


***Fig*.1.8: Schematic representation of enzymatic reaction of PLD**

PLD is a member of a superfamily, which includes phosphoinositidespecific phospholipase C, phospholipase A2 and sphingomyelinase; all are phospholipids-degrading enzymes that generate biologically active products which are assumed to play important functions in cell regulation. In the context of these specific roles, PLD has been proposed to function in regulated secretion, cytoskeletal reorganization, membrane trafficking, transcriptional regulation and cell-cycle control (25). PLD superfamily members are found in organisms ranging from viruses to bacteria,

yeast, plants and animals, where PLD activity is present in a wide variety of cell types including blood platelets, hepatocytes, lymphocytes, fibroblasts, neuronal cells, muscle cells and endothelial cells. Cloning of the plant and yeast PLDs led ultimately to the cloning of two mammalian isoforms: PLD1 and PLD2, each of which is expressed as two splice variants and localized the first, on the Golgi or perinuclear vesicular structures, whereas the second, appears to be associated with the plasma membrane (26). (*Fig*. 1.9)

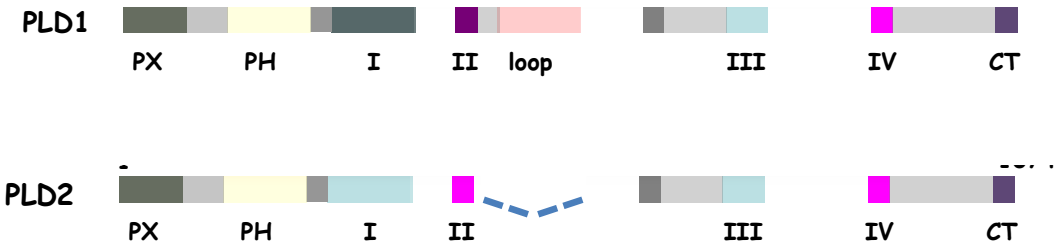


***Fig*.1.9: Schematic representation of modular structure of h-PLD1 and PLD2**

The common catalytic domain of PLD enzymes designated as HKD, bears a consensus HxxxxKxD sequence, where histidine (H), lysine (K) and aspartic acid (D)are critical for enzymatic activity both in vitro and in vivo, as evidenced by site mutations studies (27).

Other highly conserved regions the PLD genes are the phox consensus sequence (PX), the pleckstrin homology (PH) domain and the PI4,5P2 binding site at their N-terminal which are implicated in phospholipids and protein binding (26). Further, only in PLD1 is present a conserved loop region proposed to function as a possible negative regulatory element (28). The C-terminal four amino acids of mammalian PLDs are completely conserved and any change of these residues causes loss of catalytic activity (29). Mammalian PLD activity is regulated by many factors, including phosphoinositides, Protein kinase C-α (PKCα), ADP ribosylation factor (ARF), Rho GTPases such as RhoA, Cdc42, Rac1 and non-PKC protein phosphorylation. In comparison with PLD1, recombinant PLD2 exhibits high basal

activity in vitro, so it can be activated mildly by ARF, exhibiting 1,5-2 fold activation (30).

1.2.3. PED/PLD1 interaction: D4α domain

Several studies showed that PLD1 associated with GLUT4-containing membranes acted in a constitutive manner to promote the mechanisms of GLUT4 translocation by insulin (31). Given the proposed roles for PLD1 in regulation of GLUT4 translocation, in signal transduction, apoptosis and its regulation by PKC, PED/PEA-15 is an interacting partner of crucial biological significance. Interestingly, observations suggested that PED/PEA-15/PLD1 binding resulted in increased PLD1 levels, thus increasing PLD activity both in vitro and in vivo.
The minimum PLD1 region required for PED/PEA-15 interaction has been identified by Zhang et al. (32) and termed D4 (*Fig.* 1.10).

^{712}RSLSYPFLLPKSQTTAHELRYQVPGSVHANVQLLRSAADWSAGIKYHEES^{761}

^{762}IHAAYVHVIENSRHYIYIENQFFISCADDKVVFNKIGDAIAQRILKAHRE^{811}

^{812}NQKYRVYVVIPLLPGFEGDISTGGGNALQAIMHFNYRTMCRGENSILGQL^{861}

^{862}KAELGNQWINYISFCGLRTHAELEGNLVTELIYVHSKLLIADDNTVIIGS^{911}

^{912}ANINDRSMLGKRDSEMAVIVQDTETVPSVMDGKEYQAGRFARGLRLQCFR^{961}

***Fig*.1.10: Primary sequence of D4-PLD1**

The D4 region is located at the PLD1 C-terminus and contains one of the two phosphodiesterase domains responsible for PLD1 enzymatic activity. My research group has demonstrated through SPR and ELISA-like assays that PED/PEA-15 binds in vitro the D4 domain with high affinity (33,34) and, furthermore, that a PED/ PEA-15 region spanning residues 1–24 and termed is the mainly involved in D4 recognition (33,34). When loaded into L6 cells overexpressing PED/PEA-15 and into myocytes derived from PED/PEA-15 overexpressing transgenic mice, PED-(1–24) abrogates the PED/PEA-15-PLD1 interaction and reduces PKC activity to levels similar to controls. Importantly, the peptide restored insulin-stimulated glucose

uptake, and similar results were obtained by expression of the D4 domain in L6 cells overexpressing PED/PEA-15, indicating that targeting PED/PEA-15- PLD1 interaction may represent a novel strategy to improve sensitivity to insulin action (33).

In a recent study, Doti et al (35), by using D4 deletion mutants (*Fig*. 1.11), have restricted the PLD1 region involved in PED/PEA15 interaction to the N-terminal fragment named D4α (residues 712–818). This region binds PED/PEA15 with the same efficacy as D4 and, when transfected in different PED/PEA15-overexpressing cells, it is able to reduce PKC-α activity and to restore the sensitivity of PKC-ζ to insulin stimulation, independently of the PI3K/Akt signalling. This results suggest that a restricted binding interface (107 residues) on PLD1 can be used for the design and identification of effective antagonists.

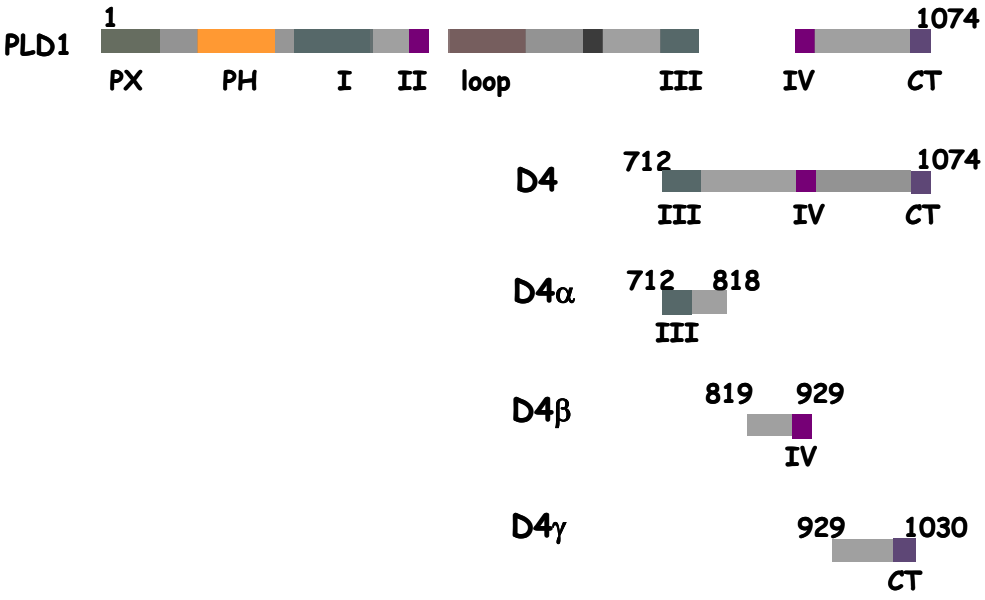


***Fig*. 1.11*:* Schematic representation of PLD1 and D4 sub-domains: Designed truncated domains D4α, D4β and D4γ**

1.3. Aim of the study

Understanding the physiological role of PED/PEA-15 in glucose tolerance has been followed by further efforts aimed at generating novel molecules interfering with PED/PEA-15 function and featuring potential pharmacological activity. Since PED/PEA-15 interaction with PLD1 appears to represent the initial event leading to PED/PEA-15 effect on glucose disposal, efforts were devoted to the identification of molecules inhibiting PED/PEA-15-PLD1 interaction.

Here simplified peptide libraries have been prepared and screened in a positional scanning format for the identification of small peptides able to inhibit the PED/PEA15-D4α interaction.

By using an-ELISA-based screening, performed on a fully automated platform, two simplified peptide libraries were analyzed in a positional scanning format. This screening led to the identification of small peptides able to inhibit PED/PEA15-D4α interaction. The selection of inhibitors was carried out employing combined competitive and direct experiments, through ELISA and SPR techniques, providing peptides with IC_{50} values in the micromolar range.

2. EXPERIMENTAL SECTION

2.1. Materials

pETM vectors were from EMBL (Heidelberg, Germany), oligonucleotides were synthesized by Sigma-Genosys (Sigma-Aldrich, Milano, Italy). Phusyon DNA polymerase was from Finnenzyme (Milano, Italy). Restriction enzymes were from New England Biolabs (Milano, Italy). All molecular biology kits were from Qiagen (Milano, Italy). Escherichia coli bacterial strains were from Novagen (Milano, Italy). Reagents for bacterial medium were from Becton-Dickinson (Milano, Italy). All reagents for SDS-PAGE, chromatographic columns and AKTA FPLC, were from GE Healthcare (Milano, Italy). EZ-Link NHS-LC-Biotin Reagent and Pro-JectTM Protein Transfection Reagent kit were from PIERCE (Rockford, US). STRV-HRP, Sigma-fast o-Phenylenediamine dihydrochloride Tablet Sets were from Sigma-Aldrich (Steinheim, Germany).

Reagents for peptide synthesis (Fmoc-protected amino acids and resins, activation and deprotection reagents) were from Novabiochem (Laufelfingen, Switzerland) and InBios (Napoli, Italy). Solvents for peptide synthesis and HPLC analyses were from Romil (Dublin, Ireland); reversed phase columns for peptide analysis and the LC-MS system were from ThermoFisher (Milano, Italy).

2.2. Methods

2.2.1. ELISA

D4 deletion mutants, PED/PEA15 recombinant proteins were expressed and purified as already reported (33,35). Also direct binding ELISA were carried out as described in the same papers (33,35). All ELISAs were performed using an integrated platform for High-Throughput Screening (Hamilton Robotics, Bonaduz, CH) comprising a fully equipped Starlet 8 channel liquid handler, a robotic arm, a washer and a multi-wavelength plate reader. Competition ELISA experiments were carried out by coating D4α (0.10 μM) and pre-incubating mono-biotinylated PED/PEA15 at 1.0 μM with peptide sub-libraries or single purified peptides at fixed concentrations (100 and 200 μM); dose-response experiments were performed with the single purified peptides at concentrations ranging between 3-400 μM. Competition results were reported as % inhibition = (B/B0)X100, where B means

the average OD from the triplicate data points for a given analyte and B0 is the average OD determined without competitor. Experimental data were fitted with GraphPad Prism, vers. 4.00, GraphPad Software (San Diego, California).

2.2.2. Peptide Synthesis

Peptide libraries and single peptides were prepared by the solid phase method on a 50 µmol scale following the Fmoc strategy and using standard Fmoc-derivatized amino acids (36). Briefly, synthesis were performed on a fully automated multichannel peptide synthesizer Syro I (Multisynthech, Germany). RINK AMIDE resin (substitution 0.5 mmol/g) was used as solid support. Activation of amino acids was achieved using HBTU-HOBt-DIEA (1:1:2), whereas Fmoc deprotection was carried out using a 40% (v/v) piperidine solution in DMF. All couplings were performed for 15 minutes and deprotections for 10 minutes. Randomized positions were coupled using equimolar mixtures of 12 building blocks employing a 100-fold excess of each amino acid. Peptides were removed from the resin by treatment with a TFA:TIS:H2O (90:5:5, v/v/v) mixture, then they were precipitated in cold diethylether and lyophilized.

2.2.3. Library design

The libraries used in this study, named PS-SSPLs (Positional Scanning-Simplified Synthetic Peptide Libraries) consisted of four positions libraries, each with one position defined by one of 12 amino acids chosen as building blocks. It relies in employing 12 different amino acids (instead of 20) that are accurately selected in order to preserve a wide range of hydrophobicity, aromaticity, charges and pKa values and MWs (Table 2.1). Further, randomized positions were set to only four residues in order to approach the size of small molecules to which they, successively, could be converted to improve their pharmacological profile (37).

Aaa	Hydrophobic	Polar	Charged	Aromatic or Aliphatic
Ala	X	-	-	-
Arg	-	X	basic	-
Asn	-	X	-	-
Asp	-	X	acidic	-
Cys	X	-	acidic	-
Gln	-	X	-	-
Glu	-	X	acidic	-
Gly	-	-	-	-
His	-	X	weak basic	Aromatic
Ile	X	-	-	Aliphatic
Leu	X	-	-	Aliphatic
Lys	-	X	basic	-
Met	X	-	-	-
Phe	X	-	-	Aromatic
Pro	X	-	-	-
Ser	-	X	-	-
Thr	-	X	weak acidic	-
Trp	X	-	-	Aromatic
Tyr	X	X	-	Aromatic
Val	X	-	-	Aliphatic

→

	Selected amino acid	Side chain
1	Gln	amide
2	Ser	polar
3	Arg	basic
4	Ala	hydrophobic
5	Tyr	hydrophobic, polar, aromatic
6	Pro	hydrophobic
7	Met	hydrophobic
8	Cys(Acm)	aliphatic
9	Phe	hydrophobic, aromatic
10	Leu	aliphatic
11	His	hydrophobic, polar
12	Asp	acidic

Table 2.1: Selection of 12 building blocks

In details, for residues with very similar properties, only one was chosen (for example Gln instead of Asn, and Asp instead of Glu). Arg and His were chosen instead of Lys because, though they all have a net positive charge at neutral pH, His also retains an aromatic character and Arg contains a unique guanidine group. Among other aromatic side chains, Phe and Tyr were preferred to Trp, due to their less tendency to oxidation. The aim of this approach is to simplify library synthesis and deconvolution, thus rapidly selecting lead peptides. Then residues excluded in the first screening are subsequently evaluated by preparing new libraries of single compounds where they replace the previous ones.

The chemical diversity of each library consisted of 12(number of building blocks)^4(randomized position)=20736 different sequences, each of one repeated four times, arranged in 12x4=48 sub-libraries each containing 1728 peptides (=123). PS-SSPL were prepared following pre-MIX methodologies (38). Two heptapeptide PS-SSPLs (Table 2.2) of general formula KGGXXXX and EGGXXXX (X indicates random residues), with four randomized positions, were used to identify specific inhibitors of the binding between D4α and PED/PEA15. The two charged residues at the N-terminus were chosen in order to substantiate and mutually confirm the data obtained with the two libraries and, at the same time, to make more homogeneous the solubility properties of the mixture compounds. Opposite charges

were introduced at the N-terminus in order to evaluate, and eventually subtract, the contribution of the tag charge itself. Since two distinct libraries were prepared in this format, a total of 96 different sub-libraries were prepared and analyzed. Libraries were screened in parallel on multiple plates with the automated platform and each single assay lasted on average 3h, therefore the screening was carried out in a Low Throughput mode (LT).

A

Peptide mixture number	Peptide sequence
1-12	H-$K_1G_2G_3O_4X_5X_6X_7$-NH_2
12-24	H-$K_1G_2G_3X_4O_5X_6X_7$-NH_2
24-36	H-$K_1G_2G_3X_4X_5O_6X_7$-NH_2
36-48	H-$K_1G_2G_3X_4X_5X_6O_7$-NH_2

B

Peptide mixture number	Peptide sequence
1-12	H-$E_1G_2G_3O_4X_5X_6X_7$-NH_2
12-24	H-$E_1G_2G_3X_4O_5X_6X_7$-NH_2
24-36	H-$E_1G_2G_3X_4X_5O_6X_7$-NH_2
36-48	H-$E_1G_2G_3X_4X_5X_6O_7$-NH_2

Table 2.2: Schematic representation of the two heptapeptide PS-SSPL employed in this study: The 96 peptide mixtures of two PS-SSCLs were synthesized with specific amino acids at one of four randomized positions in the peptide (O4, O5, O6, O7) and the random amino acids (X) using an equimolar amino acid mixture.

A rough characterization of peptide mixtures was performed by pool amino acid analysis of sub-libraries, as reported elsewhere (39) and by LC-ESI mass spectrometry. For this purpose, a narrow bore 50x2 mm ID C18 Biobasic column, equilibrated at 0.2 mL/min, was used. A gradient from 2% to 95% of acetonitrile (0.1% TFA) over H2O (0.1% TFA) in 170 minutes was applied and most peptides were eluted in a relatively narrow chromatographic window (from about 25% to about 55%). Averaging the mass spectra (collected between 400 and 2000 amu) over the retention time window where peptides were eluted, we obtained a gross estimation of the MWs distribution of mixture components. This distribution was then compared with the theoretical one expected for an equimolar mixture of the peptides generated in the library under analysis.

Single peptides were purified by preparative RP-HPLC using a Shimadzu LC-8A, equipped with a SPD-M10 AV detector and with a Phenomenex C18 Jupiter column (50x22 mm ID; 10 µm). LC-MS analyses were carried out on an LCQ DECA XP Ion Trap mass spectrometer equipped with an OPTON ESI source, operating at 4.2 kV needle voltage and 320 °C with a complete Surveyor HPLC system, comprised of MS pump, an autosampler and a photo diode array (PDA). Narrow bore 50x2 mm C18 BioBasic LC-MS columns were used for these analyses. Peptide products were lyophilized and stored at -20 °C until use.

2.2.4. SPR analyses

The BIAcore 3000 SPR system for Real Time kinetic analysis was from GE Healthcare (Milano, Italy). All other reagents and chemicals were commercially available by Sigma-Aldrich or Fluka (Steinheim, Germany). D4 deletion mutants were diluted in GndHCl 0.5 M phosphate buffer 50 mM pH7.2 and gradually dialyzed against HBS (10 mM Hepes, 150 mM NaCl, 3 mM EDTA, pH 7.4) 0.1 mM TCEP, prior to use. All proteins, PED/PEA15, D4α-MBP, D4β-MBP and MBP, were immobilized in 10 mM acetate buffer pH 4.0 (as indicated by previous pH-scouting procedures) (flow rate 5 µL/min, time injection 7 min) on a CM5 Biacore sensor chip, using EDC/NHS chemistry according to the manufacturer's instructions (40). Residual reactive groups were deactivated by treatment with 1 M ethanolamine hydrochloride, pH 8.5. Reference channels were prepared simply activating with EDC/NHS and deactivating with ethanolamine, in the case of PED/PEA15 as ligand. MBP was instead immobilized on the reference channels on the biosensors used to analyze the D4 deletion mutants. Competitive SPR experiments were carried out pre-incubating analytes with competitors at different molar ratios. When employed as analytes, D4 deletion mutants, PED/PEA15, MBP and single peptides stock solution were diluted in HBS buffer, 0.1 mM TCEP. Analyte injections of 90 µL were performed at a flow rate of 20 µL/min at the indicated concentrations. The BIAevaluation analysis package (version 4.1, GE Healthcare, Milano, Italy) implemented on the instrument software was used to subtract the signal of the reference channel and for K_{DS} value estimation.

3. RESULTS

3.1. ELISA-based LTS: identification of PED/PEA15-D4α inhibitors employing PS-SSPLs

The identification of PED/PEA15-D4α complex inhibitors was carried out employing SSPLs in a PS format. An automated Low Throughput Screening (LTS) procedure based on ELISA competitive experiments was employed to select the inhibitors.

In a first screening cycle, the most active peptide mixture(s) were selected for each of the four positions. In figure 1 the inhibition percentages of the 96 sub-libraries are reported. Each position was analyzed separately bringing to the selection of Ser for position P4, Pro for P5, Phe for P6, Tyr and Gln for P7 for the library bearing Lys at its N-terminus (*Fig*. 3.1 A). From the library having Glu in P1 were identified Tyr for P4, Pro and Asp for P5, His for P6 and Tyr for P7 (*Fig*. 3.1 B).

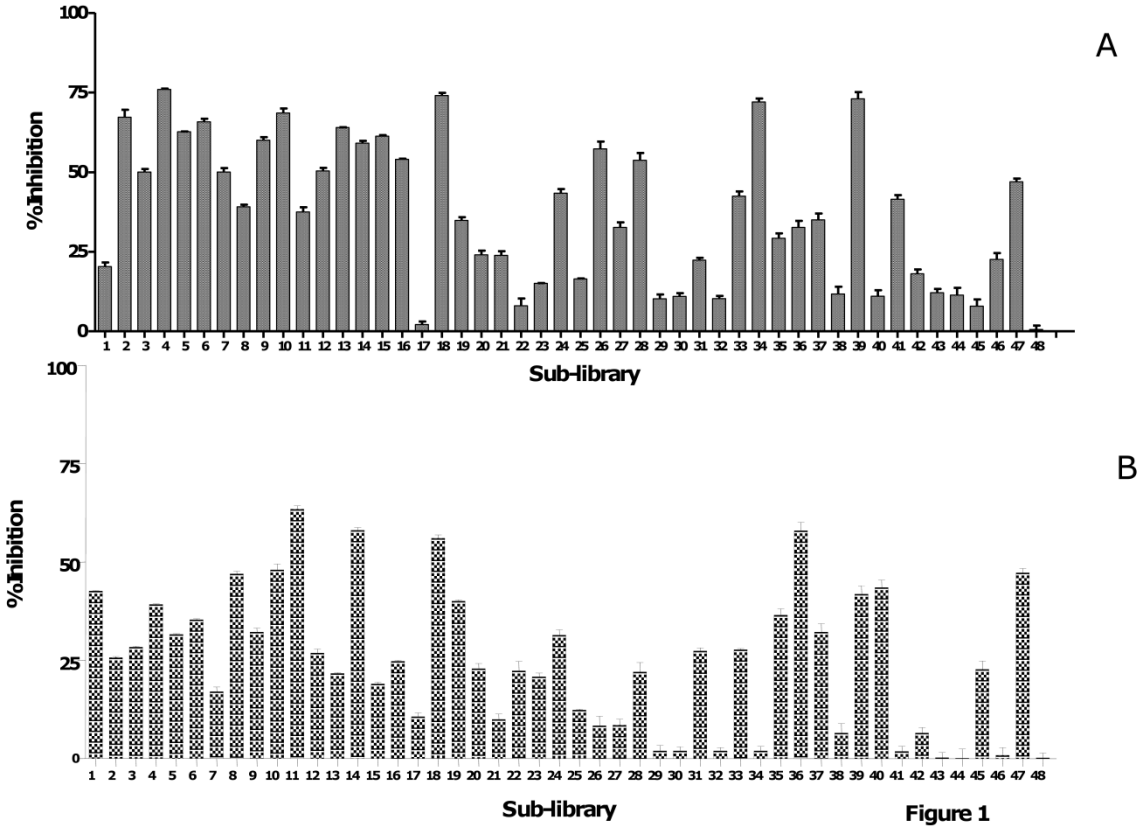


***Fig*. 3.1: ELISA-based screening of two simplified peptide libraries.** Screening of (**A**) KGGXXXX and (**B**) EGGXXXX libraries in positional scanning format, by competitive ELISA. Each library was made of 48 pools each of them was composed by 1728 peptides. Sub-libraries were tested in an ELISA-based competition assay, reporting % inhibition for each PS-mixture.

In order to confirm the screening data competition experiments of individual peptides were performed. By combining the selected amino acids for P4-P7, two single peptides for the first library (KGGSPFQ, KGGSPFY) and two sequences for the second (EGGYPHY, EGGYDHY) were designed. As negative control two sequences were selected, one for each library (KGGMAML and EGGMC(Acm)AL). The corresponding peptides were synthesized, purified and identified by LC-MS analysis (data not shown). Their inhibitory capacities were evaluated by competitive ELISAs, at two different concentrations (100 and 200 µM) as reported in *Fig*. 3.2 A-B.

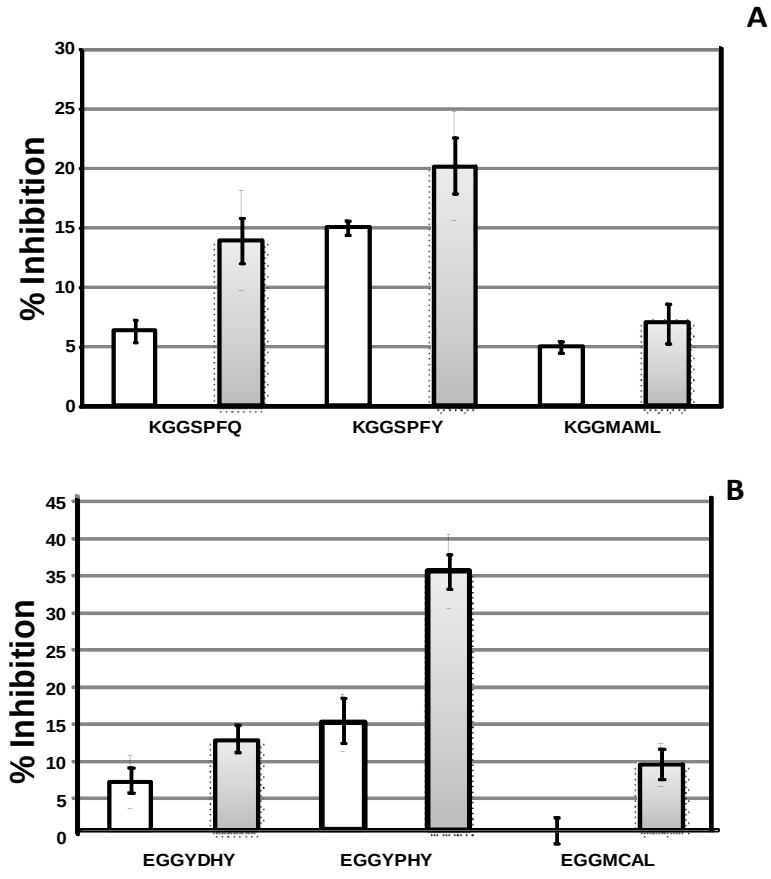


***Fig.* 3.2: Competitive ELISA experiments of lead peptides:** Competitive ELISAs of single peptides deriving from the screening of (**A**) KGGXXXX and (**B**) EGGXXXX, carried out employing peptides at 100 and 200 µM.

After these preliminary experiments, selected sequences were assayed in a dose-dependent competitive test, at a concentration range 3.00-400 µM, (*Fig*. 3.3 A-B), and by a non-linear regression fitting of experimental data, we obtained an estimation of IC_{50} values for three out of the four peptides (EGGYDHY gave a poor dose-response and the IC_{50} value couldn't be calculated).

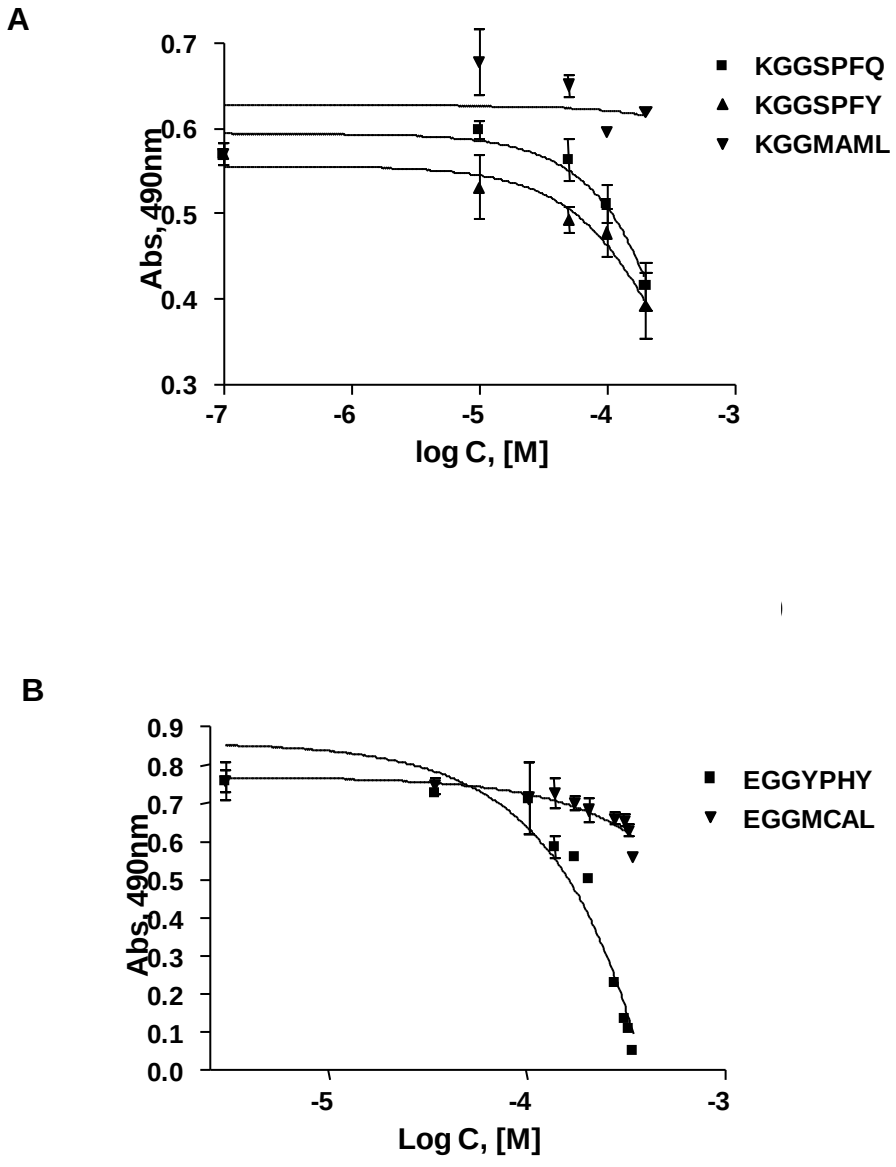


***Fig.3.3*: Competitive ELISA experiments of selected peptides:** Competitive ELISAs of selected peptides deriving from the screening of (**A**) KGGXXXX and (**B**) EGGXXXX in the concentration range (3.00-400 µM)

As reported in Table 3.1, all IC_{50} values were in the micromolar range, suggesting that no remarkable differences occur in the inhibition mechanism of Glu- and Lys-peptides. Given these similarities, for further investigations, we chose the peptide EGGYPHY which has the lowest KD value (79 μM), while the sequence EGGMC(Acm)AL was selected as negative control.

Peptide sequence	**IC_{50} (μM)**
EGGYPHY	79.2 ±1.4
KGGSPFY	105.5 ±1.6
KGGSPFQ	261.5±1.4

Table 3.1: IC_{50} values of most active peptide inhibitors of the complex PED/PEA15-D4α selected from competitive ELISA screening.

On the basis of the amino acidic composition of this lead inhibitor, seven single new peptides were designed and synthesized: they contained amino acids analogous to those of the lead peptide that were excluded for the construction of initial libraries.
Since the consensus sequence was EGGYPHY both tyrosine and histidine were replaced by tryptophan leading to the corresponding sequences: EGGWPHY, EGGYPWY, EGGYPHW, EGGWPWY, EGGWPHW, EGGWPWW (Table 3.2). They were synthesized and assayed in the same competition ELISAs, but results showed that the inclusion of Trp residues did not improve inhibition properties of the selected lead compound (*Fig*. 3.4)

	New derived peptides
1	EGGWPHY
2	EGGYPWY
3	EGGYPHW
4	EGGWPWY
5	EGGWPHW
6	EGGYPWW
7	EGGWPWW

Table 3.2. **New peptide sequences designed on the basis of consensus peptide EGGYPHY**

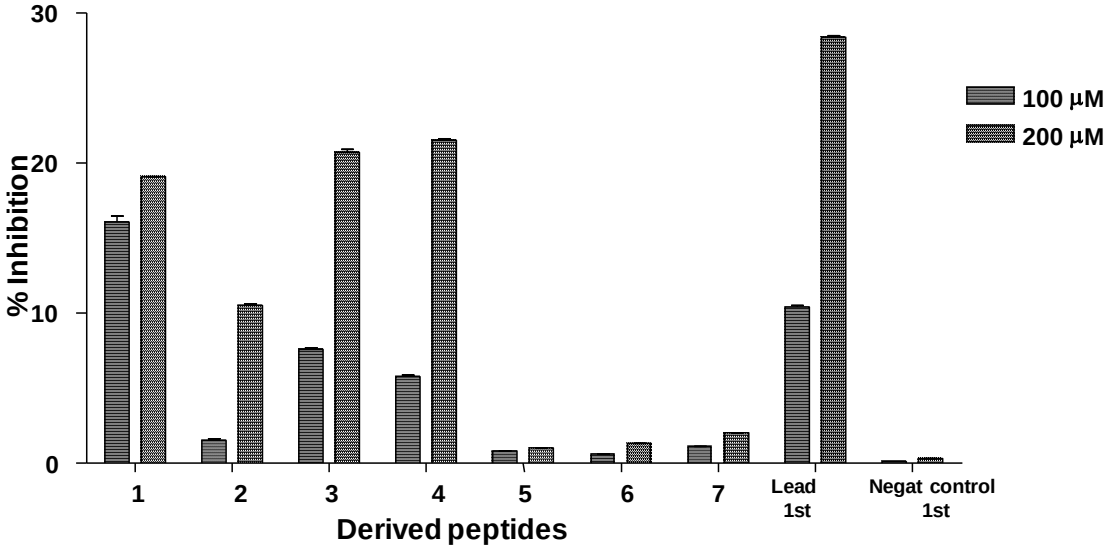


***Fig* 3.4**. **Competitive ELISA experiments:** Competitive ELISAs of single peptides deriving from the second generation, carried out employing peptides at 100 and 200 µM.

3.2. SPR experiments: identification of protein ligand

In order to further characterize the activity of the selected peptide by SPR experiments, we primarily analyzed the interaction PED/PEA15-D4α by the same technique. Since this interaction has been so far only assessed using a direct binding ELISA experiment employing Biotin-Streptavidin as detection system, we confirmed it by a different technique that employs unlabelled bio-molecules. For this purpose, PED/PEA15, D4α-MBP and MBP proteins were immobilized on the chip surfaces, achieving, under the reported conditions (see Experimental Section for details), 350, 100 and 80 RU immobilization levels, respectively.

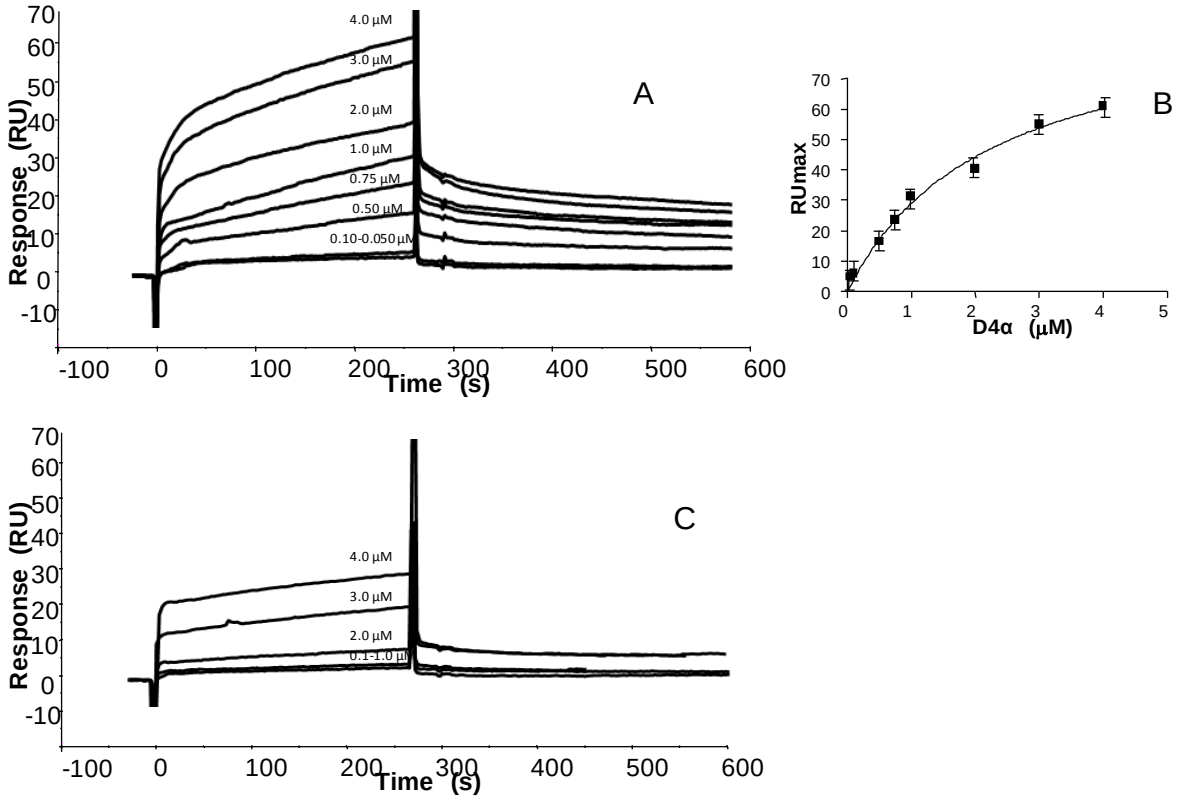


***Fig.* 3.5: SPR experiments of PED/PEA15 D4 deletion mutants.** Overlay of sensorgrams relative to the binding to immobilized PED/PEA15 of (**A**) D4α protein, concentrations were between 0.050 and 4.0 μM; (**B**) plot of RUmax from each experiment versus D4α concentration (μM); (**C**) D4β protein, concentration range was 0.10-4.0 μM. Experiments were carried out in duplicate at a 25° C, at a constant flow rate of 20 μL/min using HBS 0.1mM TCEP, as running buffer (90 μL injected for each experiment), signals deriving from MBP were subtracted from values obtained for D4 deletion mutants fused to MBP.

Direct binding between PED/PEA15 and D4 variants were performed by injecting D4-proteins solutions at increasing concentrations from 0 to 4.0 μM on immobilized PED protein.

As shown in *Fig.* 3.5 A-B, D4α bound to PED/PEA15 in a dose-response manner with a dissociation constant of 2.2±0.4 μM, that was obtained both by kinetic data and by plotting RUmax values, from each binding experiment, as a function of D4α concentrations (*Fig.* 3.5 B). This K_D is in very good agreement with that previously reported (35) and supports the view that the binding of PLD1 to PED/PEA15 is essentially mediated by the region D4α spanning residues 712-818.

To further confirm this view, we also carried out SPR experiments employing D4β region (encompassing residues of PLD1 819-920) (*Fig.* 3.5 C) and the estimated value of K_D was 100.2 ± 0.1 μM, about 50-fold lower than that of D4α. This experiment validate our approach to find out potential inhibitors of PED/PEA15-PLD1 physiological interaction focusing mainly on PED/PEA15-D4α complex.

In order to corroborate our findings, we carried out competitive SPR assays on PED/PEA15-D4α complex, employing the selected peptides as competitors.

These experiments were carried out with increasing amounts of peptides (equivalents range 0-200) pre-incubated with a fixed concentration of D4α (1.0 μM), and injecting the mixture on a PED/PEA15 derivatized sensor-chip. The overlays of obtained sensorgrams are reported in *Fig.*3.6.

As expected, we observed a clear and progressive decrease of the RUmax at increasing peptide concentration (*Fig.* 3.6 A), whereas no effects were observed with the negative control sequence (*Fig*. 3.6 B); a fitting of RUmax values against peptide concentration provided IC_{50} of 149 ± 2 μM (*Fig.* 3.6 C) in fairly good agreement with that estimated through ELISAs.

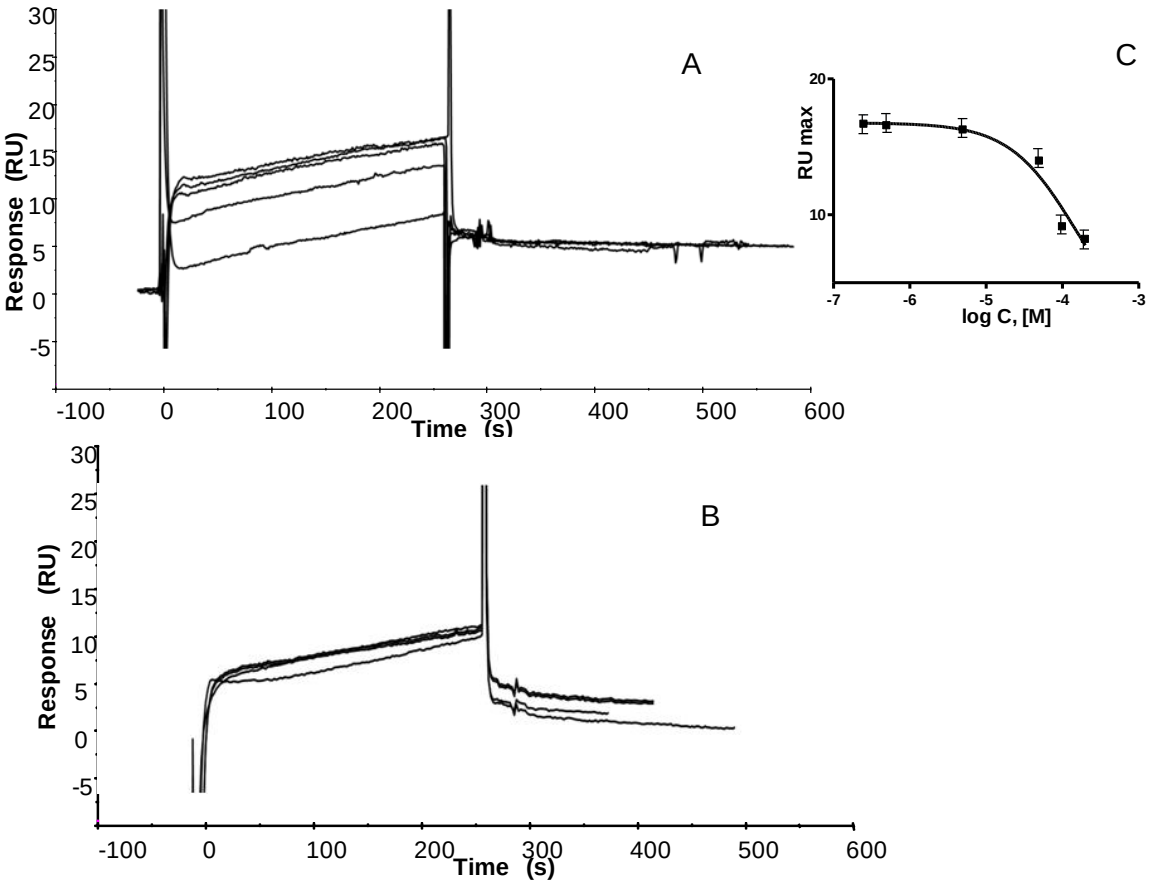


***Fig* 3.6: Competitive SPR experiments of lead peptides:** Overlay of sensorgrams relative to the competitive binding of (A) inhibitor peptide, EGGYPHY (B) negative control to the PED-PEA15/D4α complex. Peptides equivalents were between 0 and 200 respect to D4α concentration of 1.0 μM. Experiments were carried out in duplicates, at a 25° C, at a constant flow rate of 20 μL/min using HBS-0.1 mM TCEP, as running buffer (90 μL injected for each experiment).

Further, we investigated the molecular basis of the inhibition mechanism, carrying out direct binding experiments on both PED/PEA15 and D4α proteins. As shown in *Fig.* 3.7 A, inhibitor peptide did not bind to immobilized PED/PEA15, while it clearly recognized the immobilized D4α, in a dose-response way (concentrations range 200-2000 μM, *Fig.* 3.7 B).

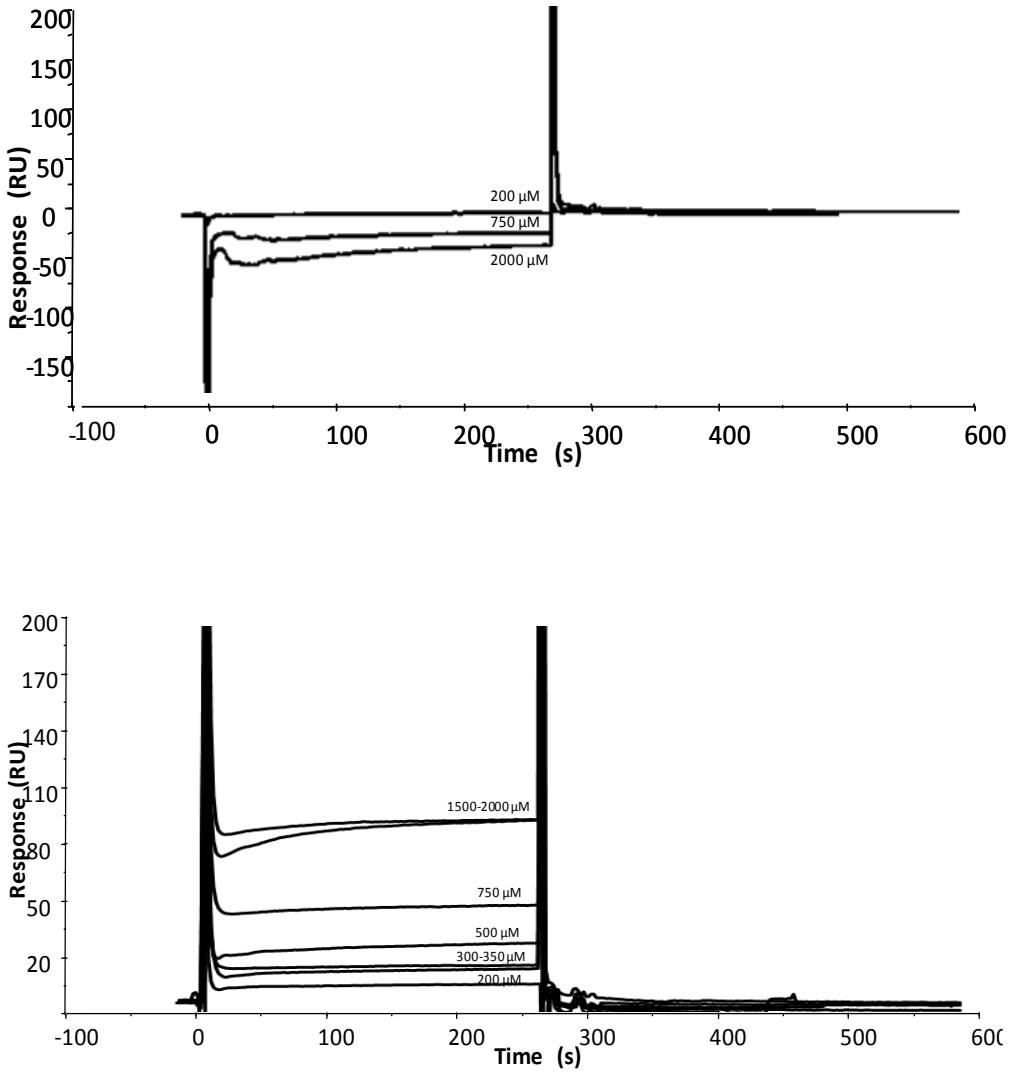


***Fig*. 3.7: Direct SPR experiments of lead peptides:** Overlay of sensorgrams relative to the direct binding of inhibitor peptide to immobilized (**A**) PED-PEA15 and (**B**) D4α-MBP. Concentrations were between 200 and 2000 µM. Experiments were carried out in duplicate, at a 25° C, at a constant flow rate of 20 µL/min using HBS-0.1 mM TCEP as running buffer (90 µL injected for each experiment).

4. DISCUSSION

Type 2 diabetes is a chronic disease characterized by high blood glucose levels generally associated to insulin resistance and is the most common form of diabetes (41).

Disease onset and progression is caused by many concurrent alterations, among which the protein complex PED/PEA15-PLD1 in individuals expressing high levels of PED, is recognized as a major cause of insulin-resistance (23,11), rendering the disruption of this complex a good target for therapeutic intervention.

Small molecules, which are cheaply produced and have better pharmacological profiles compared to biotherapeutics, are ideal candidates as drugs, but, when the target is a protein-protein interaction, their use is a great challenge (42,43), due to the characteristics of protein interfaces, particularly the lack of grooves and pockets in which they could accommodate. An easier practicable way to inhibit protein-protein interactions is represented by the use of peptides and peptidomimetics that, being more flexible, can adopt multiple conformations and are able to adapt to and to mime protein interfaces.

Several studies reported the identification by random screening peptide agonists/antagonists that bind to proteins (44,45).

Our simplified approach, based on the use of short peptides selected from random libraries and constructed with small sets of amino acids, seeks to interpose between peptide and small-molecule synthetic approaches, since it preserves the chemical robustness and diversity of synthetic peptide-based screening and, contemporaneously, aims to find out short active peptides that could work as starting template for the design of new compounds with improved affinity and stability.

Here we present peptides with inhibitory activity that can be rapidly and easily identified by this approach, especially in the PS-LTS format. Indeed, it leads, after one single screening round, to the unambiguous assignment of residues on each randomized position. The obtained results are further validated by the use of two distinct peptide libraries, differing only for the charge at their N-terminus, and by the convergence of the screenings on very similar sequences.

Our study has led to the identification of several bioactive peptides having on each of the four explored positions, residues with very similar chemical-physical features.

As shown, a H-bonding residue on position P4 (Ser, Tyr), a residue with a sterically hindered side chain on P5 (Pro) and aromatic residues on positions P6 (His, Phe) and P7 (Tyr) are likely required for activity.

It should be underscored that sequences selected from the Lys-tagged libraries are not identical to those identified from the Glu-tagged one, suggesting the hypothesis that the activity of these molecules could be partly influenced by unspecific contributions from the positive or negative charges at their N-terminus. This is as also demonstrated by the fact that addition of functional tags on the N-terminal end, leads to peptides inactivation (see below). The observation that replacement of Tyr with Trp does not alter in either way the activity of the peptides, suggests that the recognition with the target protein is likely modulated by both aromatic and H-bond donating interactions; this latter presumably prevails at the C-terminus where replacement of Tyr with Gln does not substantially impair the activity.

Since the inhibitor peptide binds to D4α, as assessed by direct SPR analysis, we looked for sequence similarity between the sequence EGGYPHY and the primary sequence of PED/PEA15. However, no similarity exists between the peptide sequences and PED/PEA15 regions involved in binding to PLD1, that are PED[1-24] (33) and PED[53-112] (46), nor with the whole sequence of the protein. Instead, comparing them with that of D4α, we surprisingly found that some peptide stretches on the protein, such as SYPF (spanning 714-717) or, more strikingly, SRHY (corresponding to segment 772-775, whereby only an arginine in place of a proline is present), are pretty identical to the identified peptides. This observation suggests that the peptides we have selected, could mimic a region of D4 involved in the dimerization of PLD1 (28,47) and could also likely explain the capacity of the peptides to block the interaction between the two proteins, indeed the segment SRHY is located within the region of D4α mainly involved in the binding with PED-PEA15, that spans from 762 to 801 residues (35).

Dose-response competitive ELISAs allowed to evaluate IC_{50} values of selected peptides in the micromolar range. Any attempt to evaluate the binding with the target protein by tagging the peptides on the N-terminus (with biotin or fluorescent tags) were unsuccessful since we experienced a total loss of activity (data not shown).

SPR competitive experiments carried out with the peptides completely confirmed the capacities of selected peptides to block complex formation and indeed the

micromolar IC_{50} values were consistent with those measured by ELISA. Direct binding assays to the immobilized proteins showed that the peptide recognized only D4α, although with a K_D in a millimolar range. This discrepancy could be explained by a different conformational state of D4α between the PED/PEA15-bound form and the free one, presumably the former has a conformation more favorable for peptide binding. Several structural studies are in progress and preliminary data suggest that D4α and, in particular, the region 762-801 undergoes major conformational changes upon PED/PEA15 recognition.

In conclusion, our results indicate that PED/PEA15-PLD1 protein complex appears to be susceptible to peptides with relatively high aromaticity and H-donor capacity, which were readily isolated from our simplified libraries. These small sequences can be used as promising scaffolds into higher-affinity compounds both for in vivo applications and SAR studies that could aid the design of more efficient inhibitors through medicinal-chemistry approaches, preparing small focused libraries of single compounds where proline on position 5 is maintained, while residues on position 4, 6 and 7 are substituted with H-bond donor/acceptor and aromatic non-natural amino acids.

REFERENCES

1. Malecki M.T. and Klupa T. (2005). Type 2 diabetes mellitus: from genes to disease. Pharmacological Reports; 57:20-32
2. Gillespie K.M. (2006). Type 1 diabetes: pathogenesis and prevention. CMAY; 175(2):165-70
3. Del Guerra S., Lupi R., et al (2005). Functional and molecular defects of pancreatic islets in human type 2 diabetes. Diabetes; 54:727-35
4. Pessin JE, Saltiel AR. Signaling pathways in insulin action: molecular targets of insulin resistance. J Clin Inv 2000; 106 (2):165-69.
5. Chang L, Chiang SH, Saltiel AR. Insulin signalling and the regulation of glucose transport. Mol Med 2004; 10(7-12):65-71.
6. Patti M.E. and Kahn C.R. (1998). The insulin receptor-a critical link in glucose homeostasis and insulin action. J. Basic Clin. Physiol. Pharmacol.; 9: 89–109
7. Withers DJ, White M, Perspective. The insulin signaling system — a common link in the pathogenesis of type 2 diabetes. Endocrinology 2000; 141: 1917–1921.
8. Holten MK, Zacho M, Gaster M et al. Strength Training Increases Insulin Mediated Glucose Uptake, GLUT4 Content and Insulin Signaling in Skeletal Muscle in Patients With Type 2 Diabetes. Diabetes 2004; 53: 204–305.
9. Shepherd PR, Kahn BB. Glucose transporter and insulin action- implications for insulin resistance and diabetes mellitus. NEJM 1999; 341: 248–257.
10. Prokopenko I, McCarthy MI, Lindgren CM. Type 2 diabetes: new genes, new understanding. Trends Genet. 2008 Dec;24(12):613-21. Epub 2008 Oct 25.
11. Condorelli G, Vigliotta G, Iavarone C, Caruso M, Tocchetti CG, Andreozzi F, Cafieri A, Tecce MF, Formisano P, Beguinot L, Beguinot F. PED/PEA-15 gene controls glucose transport and is overexpressed in type 2 diabetes mellitus. EMBO J 17: 3858–3866, 1998.
12. Valentino R, Lupoli GA, Raciti GA, Oriente F, Farinaro E, Della Valle E, Salomone M, Riccardi G, Vaccaro O, Donnarumma G, Sesti G, Hribal ML, Cardellini M, Miele C, Formisano P, Beguinot F. The PEA15 gene is overexpressed and related to insulin resistance in healthy first-degree relatives of patients with type 2 diabetes. Diabetologia 49: 3058–3066, 2006.
13. Zhang Y, Redina O, Altshuller YM, Yamazaki M, Ramos J, Chneiweiss H, Kanaho Y, Frohman MA. Regulation of expression of phospholipase D1 and D2 by PEA-15, a novel protein that interacts with them. J Biol Chem 275: 35224–35232, 2000.
14. Condorelli G, Vigliotta G, Trencia A, Maitan MA, Caruso M, Miele C, Oriente F, Santopietro S, Formisano P, Beguinot F. Protein kinase C (PKC)-alpha activation inhibits PKC-zeta and mediates the action of PED/PEA-15 on glucose transport in the L6 skeletal muscle cells. Diabetes 50: 1244–1252, 2001.
15. Vigliotta G, Miele C, Santopietro S, Portella G, Perfetti A, Maitan MA, Cassese A, Oriente F, Trencia A, Fiory F, Romano C, Tiveron C, Tatangelo L, Troncone G, Formisano P, Beguinot F. Overexpression of the ped/pea-15 gene causes diabetes by impairing glucose-stimulated insulin secretion in addition to insulin action. Mol Cell Biol 24: 5005–5015, 2004.

16. Estelles A, Yokoyama M, Nothias F, Vincent JD, Glowinski J, Vernier P, Chneiweiss H. The major astrocytic phosphoprotein PEA-15 is encoded by two mRNAs conserved on their full length in mouse and human. J Biol Chem 1996; 2 71(25):14800-06
17. Formstecher E, Ramos JW, Fauquet M, Calderwood DA, Hsieh JC, Canton B, Nguyen XT, Barnier JV, Camonis J, Ginsberg MH, Chneiweiss H. PEA-15 mediates cytoplasmic sequestration of ERK MAP kinase. Dev Cell 2001; (2):239-50.
18. Kubes M, Cordier J, Glowinski J, Girault JA, Chneiweiss H. Endothelin induces a calcium-dependent phosphorylation of PEA-15 in intact astrocytes: identification of Ser104 and Ser116 phosphorylated, respectively, by protein kinase C and calcium/calmodulin kinase II in vitro. J Neurochem 1998; 71(3):1307-14.
19. Trencia A, Perfetti A, Cassese A, Vigliotta G, Miele C, Oriente F, Santopietro S, Giacco F, Condorelli G, Formisano P, Beguinot F. Protein kinase B/Akt binds and phosphorylates PED/PEA-15, stabilizing its antiapoptotic action. Mol Cell Biol 2003; 23(13):4511-21.
20. Hill JM, Vaidyanathan H, Ramos JW, Ginsberg MH, Werner MH. Recognition of ERK MAP kinase by PEA-15 reveals a common docking site within the death domain and death effector domain. EMBO J 2002; 21(23):6494-504.
21. Eberstadt M, Huang B, Chen Z, Meadows RP, Ng SC, Zheng L, Lenardo MJ, Fesik SW. NMR structure and mutagenesis of the FADD (Mort1) death-effector domain. Nature. 1998 Apr 30;392(6679):941-5.
22. Yang JK, Wang L, Zheng L, Wan F, Ahmed M, Lenardo MJ, Wu H. Crystal structure of MC159 reveals molecular mechanism of DISC assembly and FLIP inhibition. Mol Cell. 2005 Dec 22;20(6):939-49.
23. Francesca Fiory, Pietro Formisano, Giuseppe Perruolo, and Francesco Beguinot Frontiers: PED/PEA-15, a multifunctional protein controlling cell survival and glucose metabolism AJP-Endocrinol Metab 2009 vol 297
24. Jenkins GM, Frohman MA, Phospholipase D: a lipid centric review. Cell Mol Life Sci 2005; 62: 2305-2316.
25. Liscovitch M, Czarny M, Ciucci G, Tang X. Phospholipase D: molecular and cell biology of a novel family. Biochem J 2000; 345: 401-415.
26. Exton JH. Phospholipase D-structure, regulation and function. Rev Physiol Biochem Pharmacol 2002; 144:1-94.
27. Sung TC, Roper RL, Zhang Y, Rudge SA, Temel R, Hammond SM, Morris AJ, Moss B, Engebrecht JA, Frohman MA. Mutagenesis of phospholipase D defines a superfamily including a trans-Golgi viral protein required for poxvirus pathogenicity. EMBO J 1997; 16(15):4519-30.
28. Sung TC, Altshuller YM, Morris AJ, Frohman M.A, Molecular analysis of mammalian Phospholipase D2. J Biol Chem 1999; 274(1):494-502.
29. Xie Z, Ho WT, Exton JH. Conserved amino acids at the C-terminus of rat phospholipase D1 are essential for enzymatic activity. Eur J Biochem. 2000; 267(24):7138-46.
30. Sarri E, Pardo r, Fensome-Green A, Cockcroft S. Endogenous phospholipase D2 localizes to the plasma membrane of RBL-2H3 mast cells and can be distinguished

from ADP ribosylation factor-stimulated phospholipase D1 activity by its specific sensitivity to oleic acid. Biochem J 2003; 369:319-29.
31. Emoto M, Klarlund JK, Waters SB, Hu V, Buxton JM, Chawla A, Czech MP. Arole for phospholipase D in GLUT4 glucose transporter translocation. J Biol Chem 2000; 275(10):7144-51.
32. Zhang Y, Redina O, Altshuller YM, Yamazaki M, Ramos J, Chneiweiss H, Kanaho Y, Frohman MA. Regulation of expression of phospholipase D1 and 50 D2 by PEA-15, a novel protein that interacts with them. J Biol Chem 2000; 275(45):35224-32.
33. Viparelli, F., Cassese, A., Doti, N., Paturzo, F., Marasco, D., Dathan, N.A., Monti, S.M., Basile, G., Ungaro, P., Sabatella, M., Miele, C., Teperino, R., Consiglio, E., Pedone, C., Beguinot, F., Formisano, P., and Ruvo, M. (2008) Targeting of PED/PEA-15 molecular interaction with phospholipase D1 enhances insulin sensitivity in skeletal muscle cells. J Biol Chem 283: 21769-21778.
34. Viparelli F, Doti N, Sandomenico A, Marasco D, Dathan NA, Miele C, Beguinot F, Monti SM, Ruvo M. Expression and purification of the D4 region of PLD1 and characterization of its interaction with PED-PEA15. Protein Expr Purif 59: 302–308, 2008.
35. Doti, N., Cassese, A., Marasco, D., Paturzo, F., Sabatella, M., Viparelli, F., Dathan, N., Monti, S.M., Miele, C., Formisano, P., Beguinot, F., Ruvo, M. (2010) Residues 762-801 of PLD1 mediate the interaction with PED/PEA15, Molecular Biosyst, 6:2039-48.
36. Fields, G.B., and Noble, R.L. (1990). Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids. Int J Pept Protein Res 35: 161-214.
37. Marasco, D., Perretta, G., Sabatella, M., and Ruvo, M. (2008) Past and future perspectives of synthetic peptide libraries. Curr Protein Pept Sci 9 : 447-467.
38. Geysen, H.M., Rodda, S.J., and Mason, T.J. (1986) A priori delineation of a peptide which mimics a discontinuous antigenic determinant. Mol Immunol 23: 709-715.
39. Ruvo, M., Scardino, P., Cassani, G. Fassina, G. (1994) Facile Manual Synthesis of Peptide Libraries. Peptide and Protein Letters, 1-187-192.
40. Johnsson, B., Lofas, S., and Lindquist, G. (1991) Immobilization of proteins to a carboxymethyldextran-modified gold surface for biospecific interaction analysis in surface plasmon resonance sensors. Anal Biochem 198: 268-277.
41. Flier, J.S. (1983) Insulin receptors and insulin resistance Annu Rev Med 34:145-60.
42. Cheng, A.C., Coleman, R.G., Smyth, K.T., Cao, Q., Soulard, P., Caffrey, D.R., Salzberg, A.C., and Huang, E.S. (2007) Structure-based maximal affinity model predicts small-molecule druggability. Nat Biotechnol 25: 71-75.
43. Arkin, M.R., Randal, M., DeLano, W.L., Hyde, J., Luong, T.N., Oslob, J.D., Raphael, D.R., Taylor, L., Wang, J., McDowell, R.S., Wells, J.A., and Braisted, A.C. (2003) Binding of small molecules to an adaptive protein-protein interface. Proc Natl Acad Sci U S A 100: 1603-1608
44. Sidhu, S.S., Fairbrother, W.J., and Deshayes, K. (2003) Exploring protein-protein interactions with phage display. Chembiochem 4: 14-25.
45. Pillutla, R.C., Hsiao, K.C., Beasley, J.R., Brandt, J., Ostergaard, S., Hansen, P.H., Spetzler, J.C., Danielsen, G.M., Andersen, A.S., Brissette, R.E., Lennick, M.,

Fletcher, P.W., Blume, A.J., Schaffer, L., and Goldstein, N.I. (2002) Peptides identify the critical hotspots involved in the biological activation of the insulin receptor. J Biol Chem 277: 22590-22594.

46. Besterman, J.M., Duronio, V., Cuatrecasas, P. (1986) Rapid formation of diacylglycerol from phosphatidylcholine: a pathway for generation of a second messenger. Proc Natl Acad Sci U S A 83:6785-9.
47. Xie, Z., Ho, W.T., Exton, J. (1998) Association of the N- and C-terminal domains of Phospholipase D is required for catalytic activity. J. Biol. Chem. 273: 34679–34682

CHAPTER III

1. INTRODUCTION

1.1. Psoriasis

Psoriasis is an inflammatory genetically disease of the skin, that is accompanied by an alteration of the proliferation and differentiation of keratinocytes that are the major cellular components of the epidermis. This disease shows several clinical forms, the most common of which is characterized by evidence of erythematous lesions, isolated or confluent.

They are selectively affected limb extensor face, especially the elbows and knees, the sacrococcygeal region and scalp. Psoriasis has an incidence of 2% in the population and 5-7% of all dermatoses. It affects both sexes with equal frequency and can occur in childhood or later until the ninth decade. It occurs most frequently between 5 and 40 years, with the highest peak between 10 and 30 years. In psoriatic lesions are demonstrable in several biochemical, histopathological and immunological alterations but none of them is so specific that it is considered a marker of psoriasis.

1.1.1. The causes of Psoriasis

Although the primary cause of psoriasis is still controversial whether, this disease resulted from an intrinsic abnormality of the epidermal keratinocyte and/or altered regulation of the immune system. Numerous studies have shown unequivocally that T cells play a crucial role in the turnover and psoriatic keratinocyte growth (1). Studies in immunocompromised mice (SCID) (in which psoriatic skin asymptomatic and healthy grafts were transplanted) showed that intrinsic alterations of the keratinocyte are necessary to induce psoriasis (1). Indeed only the psoriatic skin is able to evolve pathologically, once transplanted immune-competent cells of the patient in the dermis of the animals. In addition psoriasis stems from defects in endogenous keratinocyte has been recently obtained through the use of genetically engineered animals (2-4). In particular, transgenic mice that express specifically high levels in the epidermis of the protein STAT3 (Signal Transducer and Activator of Transcription 3) (2) or lacking IKK2 inhibitor (Inhibitor of IK kinase 2) (3) develop skin lesions that resemble to those of psoriatic patients. Similarly, the elimination of

the protein in keratinocytes AP-1 (JunB/activator protein 1) (4) induces a phenotype in the skin of mice with histological features of psoriasis, including epidermal hyperplasia and a marked presence of an abundant inflammatory infiltrate dermis. In addition the psoriatic keratinocyte shows defects in the control of the expression of pro-inflammatory factors displaying altered proliferation and differentiation ability. For example, patients with psoriatic keratinocytes produce abnormal amounts of certain chemotactic substances, known to be involved in the recruitment of inflammatory cells, including IL-8, MCP-1 and IP-10 (1). The system of production of cytokines, proteins essential for intercellular communication, is severely defective in psoriatic keratinocytes, thus profound alterations in receptor-interleukin 1 (IL-1) and expression of cytokines IL-15, IL-18 and VEGF (Vascular Endothelial Growth Factor), were detected in the epidermis of psoriatic skin (1).

The improper answer of psoriatic keratinocytes against IFN (interferon)-γ (5) is one of their intrinsic defects. IFN-γ is one of the most powerful inflammatory stimuli, abundantly present in psoriatic skin. The intradermal injection of IFN-γ in psoriatic asymptomatic skin determines the proliferation of keratinocytes and the development of psoriatic lesions (5). Therefore the increased and inappropriate expression of inflammatory molecules (chemotactic substances, cytokines) as well as the alteration of proliferative processes in the epidermis of psoriatic patients depends heavily on the local production of IFN-γ (5).

1.2. Cytokines and IFN-γ

Cytokines is a highly heterogeneous group of small and secret messenger proteins, involved in intercellular communication in multicellular organisms and some of them remain cell bound. Two crucial features characterize the cytokine network: their pleiotropy and redundancy. Pleiotropy indicates that one cytokine can cause a broad range of responses depending on the cell type or the differentiation stage; thus cytokines can orchestrate a coordinated response of different cellular processes. Redundancy implies that different cytokines can exert similar biological activities. In this way, important cellular mechanisms are preserved by a back-up mechanism, in which one cytokine can compensate for the loss of another. These compensatory mechanisms can be explained in part by the common use of certain receptor chains and signaling molecules by the different cytokines.

Originally, cytokines were classified on the basis of their biological responses, but, given the apparent functional pleiotropy and low sequence homology, current cytokine classification relies on structural similarities. According to the Cytokine Web (http://cmbi.bjmu.edu.cn/cmbidata/cgf/CGF_Database/cytweb/index.html) different fold families can be distinguished (Table 1.1).

α-helical structure	**long chain 4-α-helix bundle superfamily** interleukin-6 (IL-6), growth hormone (GH), Leptin, erythropoietin (Epo), prolactin (PRL), granulocyte-colony stimulating factor (G-CSF), myelomonocytic growth factor, leukemia inhibitory factor (LIF), oncostatin M (OSM), ciliary neurotrophic factor (CNTF), cholinergic differentiation factor (CDF) **short chain 4-α-helix bundle superfamily** IL-2, IL-3, IL-4, IL-5, granulocyte macrophage colony-stimulating factor (GM-CSF), macrophage-colony stimulating factor (M-CSF) **dimeric 4-α-helix bundle superfamily** interferon γ (IFNγ), interferon β (IFNβ), IL-10
β-sheet structure	**B-Trefoil** IL1-α, IL1-β, fibroblast growth factor (FGF) **β-sandwich** tumour necrosis factor α (TNFα), tumour necrosis factor β (TNFβ) **β-EGF-like** transforming growth factor α (TGF-α) **β-Cystine knot dimerization domains** gonadotropin, nerve growth factor (NGF), platelet-derived growth factor (PDGF), transforming growth factor-β2 (TGF-β2)
α/β-structure	IL-8, platelet factor 4 (PF-4), macrophage inflammatory protein 1 alpha (MIP-1α), MIP-1β, melanoma growth stimulating activity (MGSA)

Table 1.1: Classification of the cytokines based on similarity in protein folding

IFN-γ is a key cytokine that regulates a variety of cellular activities, including antiviral immunity, apoptosis, and cell cycle progression. In the skin, it plays a

pivotal role in the development of inflammatory and immune responses. Interferons allow communication between cells to trigger the protective defenses of the immune system that eradicate pathogens or tumors and are named after their ability to "interfere" with viral replication within host cells. Furthermore they activate immune cells, such as natural killer cells and macrophages, increase recognition of infection or tumor cells by up-regulating antigen presentation to T lymphocytes; and the ability of uninfected host cells to resist new infection by virus. Certain host symptoms, such as aching muscles and fever, are related to the production of IFNs during infection. About ten distinct IFNs have been identified in mammals; seven of these have been described for humans. They are typically divided among three IFN classes: Type I IFN, Type II IFN, and Type III IFN. IFNs belonging to all IFN classes are very important for fighting viral infections.

Based on the type of receptor through which they signal, human interferons have been classified into three major types.

- Interferon type I: All type I IFNs bind to a specific cell surface receptor complex known as the IFN-α receptor (IFNAR) that consists of IFNAR1 and IFNAR2 chains.(6) The type I interferons present in humans are IFN-α, IFN-β and IFN-ω.(7)
- Interferon type II: Binds to IFNGR that consists of IFNGR1 and IFNGR2 chains. In humans this is IFN-γ.
- Interferon type III: Signal through a receptor complex consisting of IL10R2 (also called CRF2-4) and IFNLR1 (also called CRF2-12). Acceptance of this classification is less universal than that of type I and type II, and unlike the other two, it is not currently included in Medical Subject Headings.(8)

Keratinocytes, which constitutively bear the IFN- γ receptor complex, are a primary target of IFN-γ (9). During immune-mediated skin diseases, after exposure to IFN-γ, keratinocytes become a source of a plethora of inflammatory mediators involved in the initiation and amplification of pathogenetic processes (10). In particular, IFN-γ induces the expression of numerous chemokines, including CCL2 and CXCL10, which drive immigration of T cells, monocytes, and dendritic cells into inflamed skin. IFN- γ-treated keratinocytes are also a source of CXCL8, a pleiotropic

chemokine showing a strong chemoattractant activity on neutrophils as well as tissue-protective and prosurvival functions (11).

1.3. The JAK-STAT pathway

IFN-γ intracellular signaling activates a number of molecular cascades initiated by Janus-activated kinases 1 (Jak1) and Jak2 phosphorylation and culminates in the activation of transcription factors, mainly Signal transducer and Activator of Transcription (STAT) 1 and interferon regulatory factor-1, which can induce the IFN-stimulated gene (ISG) expression (12).

The cell membrane receptor for IFN-γ is composed of two subunits, IFN-γR1 and IFN-γR2. Upon binding to IFN-γ, the IFN- γ-Receptor complex induce clustering and/or reorganization of the receptor chains in such a way that the associated Janus tyrosine kinases (JAKs) will be brought in close proximity, allowing them to activate each other by cross-phosphorylation. JAK enzymes phosphorylate one another and then subsequently phosphorylate the IFN-γ receptor, which results in the formation of a docking site for the latent cytoplasmic transcription factor named STAT-1, a member of the STAT (signal transducer and activator of transcription) protein family (13). Upon phosphorylation, STAT-1 homodimerizes, translocates to the nucleus, and regulates gene transcription by binding to IFN-γ-activated sequences (GAS) in the IFN-γ-inducible genes. Homodimerization of STAT-1 is mediated by the binding of the phosphorylated tyrosine 701 of one STAT-1 monomer the Src homology 2 domain of another. However, maximal transcriptional activity by active STAT-1 homodimers also requires STAT-1 phosphorylation at serine 727 (14). It has been found that STAT-1 phosphorylation plays a critical role in IFN-mediated innate immunity to microbial infection. STAT-1 signaling can also be negatively regulated by the protein inhibitor of activated STAT-1 (PIAS1) and suppressor of cytokine signaling (SOCS) (15). More interestingly, IFN-γ can also regulate expression of its inducible genes in a STAT-1-independent manner, suggesting that multiple signaling pathways in parallel play important roles in the biological response to IFN-γ (*Fig*.1.1).

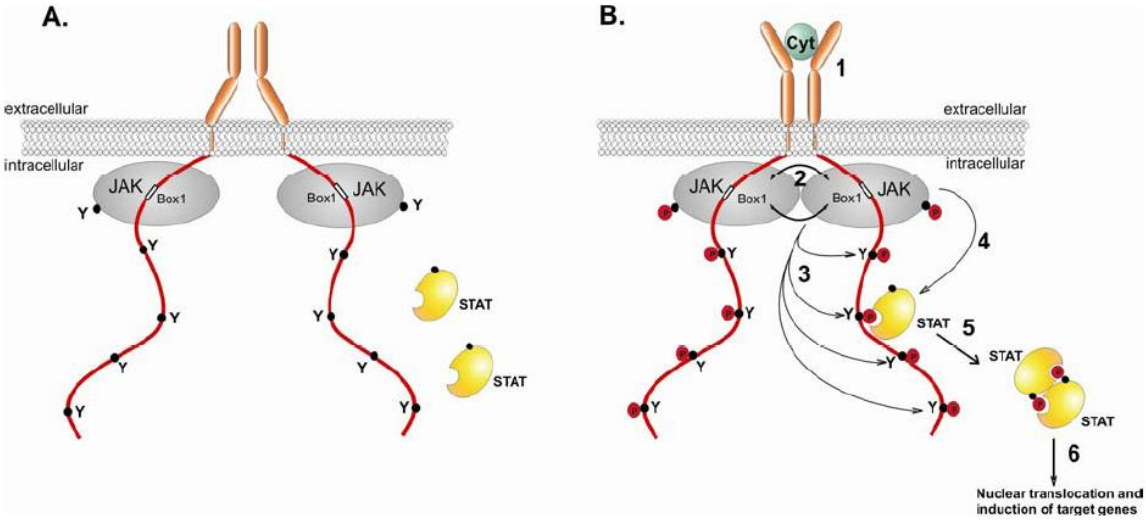


***Fig.* 1.1: Schematic representation of the JAK/STAT pathway**

1.3.1. The JAK protein family

This kinase family is named after Janus, the Roman god of gates and doors, beginnings and endings. Therefore Janus is represented with a double-faced head, each looking in an opposite direction. This is reminiscent of the distinctive feature of JAK structure: their kinase and pseudokinase domain. In mammals, the JAK family comprises four members: JAK1-3 and Tyk2. JAK1, JAK2 and Tyk2 are expressed ubiquitously, while the expression of JAK3 is restricted to cells of haematopoietic origin (16). Targeted disruption of the JAK2 gene results in embryonic lethality due to failure of erythropoiesis. Cells from these mice showed that this kinase is essential for different cytokine responses including IL-3, IL-5, GM-CSF and IFN-γ (17). JAKs are large proteins of approximately one thousand amino acids, the comparison of JAK sequences reveals seven regions of high similarity, called JAK homology (JH) domains 1 to 7 (*Fig.* 1.2). The C-terminal JH1 and JH2 domains encode respectively a kinase and pseudokinase domain. Although this latter domain contains structural features of a tyrosine kinase, it is devoid of any catalytic activity. Reports suggest that this domain modulates the catalytic activity of the kinase domain (18). The JH1 kinase domain behaves like a classical tyrosine kinase: it contains tyrosine residues that become phosphorylated upon activation, thereby inducing conformational changes that allow binding of the substrates in the catalytic site of the JAK (19).

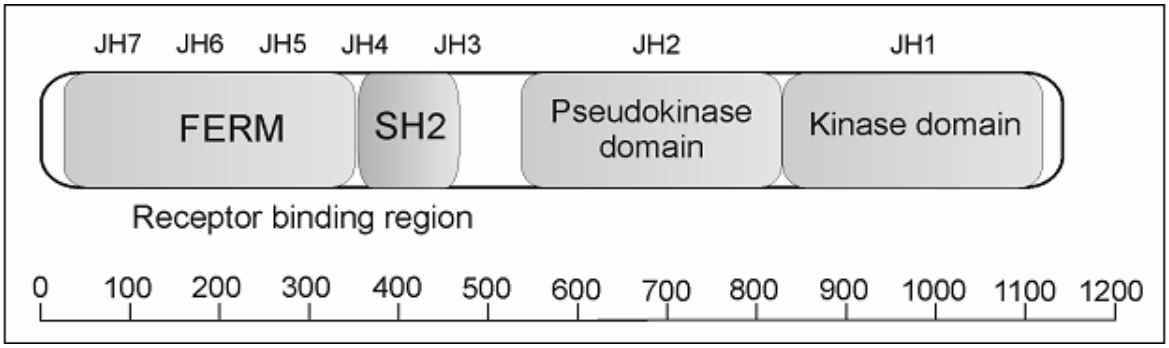


***Fig*. 1.2: Schematic representation of the JAK structure**: JAKs share seven regions of high similarity (JH1-7). N-terminal domains mediate receptor association; JH1 and JH2 encode respectively a kinase and pseudokinase domain. The ruler underneath indicates the number of AA.

The N-terminal part of the JAKs, containing JH3-7, is involved in association with the receptor. Specifically, JAKs associate with the proline-rich, membrane-proximal box1 and box2 domains of class I cytokine receptors. The JH3-JH4 regions form a structural domain resembling a SH2 domain, but appear not to be implicated in phosphotyrosine-dependent interactions, as SH2 domains typically do. However, this region is structurally important for receptor association and receptor surface expression (20). Finally, the structure of the N-terminal (JH5-JH7) domains resembles that of Four-point-one, Ezrin, Radixin and Moesin (FERM) domains, which are known to mediate protein-protein interactions with for example phosphatidylinositolbisphosphate (PIP2) and inositoltrisphosphate (IP3) (21). This region mediates receptor binding and is involved in the maintenance of catalytical activity (22).

1.3.2. The STAT protein family

The family of mammalian STAT proteins consists of seven members: STAT1, STAT2, STAT3, STAT4, STAT5a, STAT5b and STAT6. They function as transcription factors which reside predominantly in the cytoplasm in unstimulated cells, most probably as preformed dimers (23). STAT1, STAT3 and STAT4 can form both homo- and heterodimers while for STAT5 and STAT6 only homodimers have been observed. STAT2 is only functional when complexed with STAT1 or

STAT4. Receptor binding and subsequent phosphorylation on tyrosine and serine residues will activate STATs followed by their nuclear translocation. In the nucleus, they initiate transcription by associating with specific response elements in the promoter of target genes. The mechanisms underlying the transport of the STATs between the cytoplasm and the nucleus are only partially understood. The predominantly cytosolic localisation for inactive STATs has been shown to reflect a steady-state, where continuous basal nuclear import is balanced by continuous basal nuclear export. After activation, the balance is shifted toward nuclear accumulation and during signal decay toward nuclear export. Translocation of STAT dimers to the nucleus is mediated by specific nuclear localisation signals (NLS) and involves an active nuclear import mechanism depending on Ran and importin-α5 (24). A nuclear export signal (NES) drives the nuclear export of STATs mediated by Ran and exportin 1. STAT dephosphorylation and dissociation from the DNA unmasks the NES sequence and activates the export mechanism (25).

STATs are about 800 amino acid long and therein six structurally and functionally conserved domains have been identified (*Fig.* 1.3). The N-terminus contains a dimer-dimer interaction domain that apparently allows STAT dimers to form tetramers that can cooperatively associate with multiple tandem STAT response elements in the promoter (26). It was also reported that this domain is involved in nuclear translocation and STAT deactivation (27). The N-terminal region is linked to the DNA binding domain by a coiled-coil domain which consists of four helices. This domain can interact with other transcription factors and is also implicated in receptor binding, tyrosine phosphorylation and nuclear export (28).

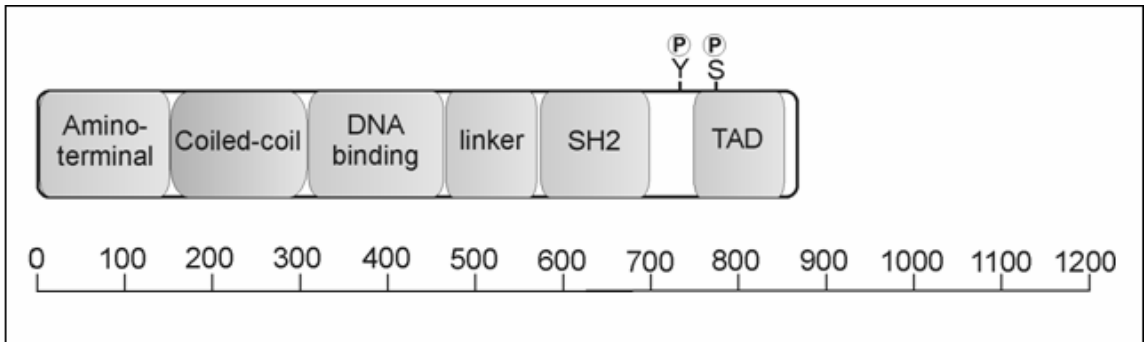


***Fig.* 1.3: Schematic representation of the STAT structure:** P indicates phosphorylated residues. The ruler underneath indicates the number of AA.

The centrally located DNA binding domain harbours a typical β-barrel with an Ig fold, a structure found in various transcription factors, like NF-κB and p53. Obviously, it is implicated in DNA association but there are only little direct interaction sites (29).

The SH2 domain is crucial for docking of the STAT to the phosphorylated tyrosine motifs in the receptor or in JAK kinases and it also mediates dimerisation of the STATs. The SH2 domain consists of a β-sheet flanked by two α-helices, which form a pocket structure. A conserved arginine residing in this pocket is essential for the interaction with phosphotyrosine residues (29). C-terminal to the SH2 domain, STATs have a conserved tyrosine residue which becomes phosphorylated by the JAK kinases upon receptor binding. Dimerization of the activated STATs is based on the reciprocal interaction of the SH2 domain with this

phosphorylated tyrosine. The more variable C-terminus of STATs encodes a transcription activation domain (TAD). This divergence provides an opportunity to associate with distinct transcriptional regulators. This domain can be serine phosphorylated, which enhances the transcription of some genes (30).

STAT1 deficient mice confirmed the pivotal role that STAT1 plays in the biological response to both type I and type II IFNs. These mice are highly susceptible to bacterial and viral infections (31).

Indeed naturally occurring mutations in human STAT1 exhibit increased susceptibility to viral and bacterial infections (32). STAT1 seems to be also implicated in non immune-responses, like chondrocyte proliferation (33) and IFNγ-mediated growth inhibition, and was revealed by the enhanced tumor susceptibility of STAT1-/- mice (34).

1.4. SOCS family

SOCSs are a family of intracellular proteins that play a major role in the regulation of cytokine responses, it counts eight members: CIS and SOCS1 through SOCS7 (35). SOCS proteins are often expressed at low or undetectable levels in resting cells. They are rapidly up-regulated in response to a broad spectrum of cytokines and, in turn, control the duration and intensity of cytokine responses by blocking various aspects of the signalling pathways. SOCS proteins function in a typical

negative feedback loop, since they can down-modulate the signalling pathway that stimulates their production. The pattern of SOCS expression by a particular cytokine tends to vary according to the cell type or tissue. The induced SOCS proteins can attenuate responses of various cytokines and may be involved in inhibitory cross-talk between different cytokine systems, providing a mechanism by which concurrent signaling processes can modulate each other. Although SOCS proteins are induced by a range of cytokines and show a high structural and functional overlap, large evidence indicates the potential of specific SOCS proteins to modulate cytokine signalling with exquisite control.

STAT proteins are the major regulators of SOCS gene expression. STAT binding sequences were identified in the SOCS promoter region and electrophoretic mobility shift assays confirmed STAT association to these motifs. In some cases, STAT induced SOCS expression is indirect. STAT1 is for example indirectly implicated in IFNγ mediated upregulation of SOCS1 since it drives expression of the interferon regulatory factor-1 (IRF-1) transcription factor which is responsible for SOCS1 induction (36). The general structure of SOCS proteins is evolutionary well preserved and common features include an N-terminal pre-SH2 domain, a central SH2 domain and a C-terminal SOCS box (*Fig.* 1.4). The N-terminal part varies in length and sequence similarity while the SH2 domain and SOCS box are most conserved.

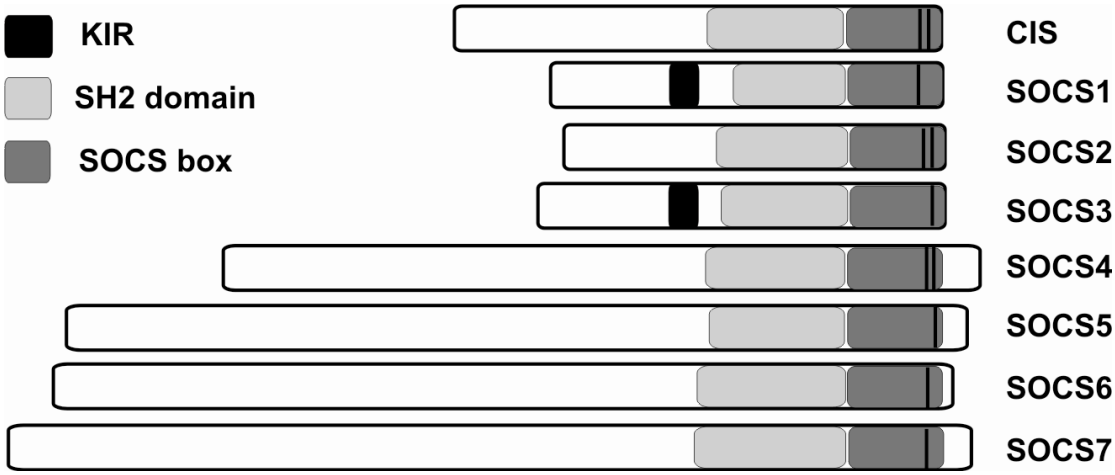


***Fig.* 1.4: Domain structure of the SOCS protein family:** The major structural characteristic of the SOCS family is the presence of two domains with relatively well conserved AA sequence: an SH2 domain in the middle portion and a SOCS box at the C-terminus. Only SOCS1 and -3 possess a KIR immediately upstream of the SH2 domain. Conserved tyrosines in the SOCS box are represented by a black line.

1.4.1. The N-terminal domain: JAK kinase inhibition

The N-terminus varies greatly in length among the SOCS members and contains no homology to known structural domains. Nevertheless, an extended SH2 subdomain (ESS) helix was identified as a conserved structural element in CIS and SOCS1-3 and appeared to be critical for high affinity SH2 substrate interactions (37). Both SOCS1 and 3 can be distinguished from the other SOCS proteins by an additional small kinase inhibitory region (KIR) located in the N-terminal portion and involved in inhibition of the JAK kinases. This region was found functionally interchangeable between the two SOCS suggesting a common inhibitory mechanism (38). The KIR region displays some sequence similarity with the activation loop of JAK2, suggesting that it acts as a pseudosubstrate by mimicking the activation loop that regulates access to the catalytic groove (37). Based on a structural model of SOCS1 in complex with JAK2 it was proposed that the KIR region suppresses JAK activity by obstructing the access of both ATP and substrate to their respective binding sites (39) (*Fig.* 1.5). Although the SH2 domain of SOCS3 does not have a high affinity for JAK kinases, the KIR domain of SOCS3 shows a stronger potential for both binding and inhibition of JAKs than that of SOCS1 (40). The wide range of actions exhibited by SOCS1 and SOCS3 is most probably due to their additional ability to inhibit the catalytic activity of JAK kinases.

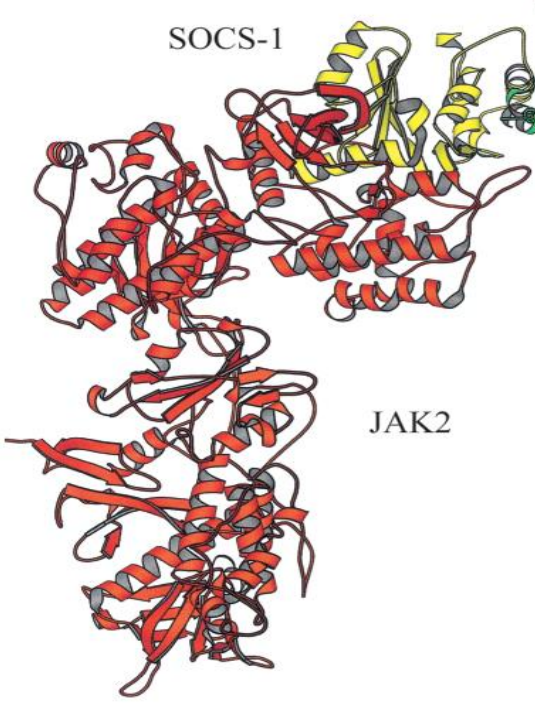


***Fig.* 1.5: Structural model of the complex between JAK2 (red) and SOCS1 (yellow).**

1.4.2. The SH2 domain: substrate specificity and competition for receptor motifs

The central SH2 domain determines the target of the SOCS protein. It allows interaction with phosphorylated tyrosine residues of other proteins, like receptors, JAKs or adaptors. In this way, SOCS proteins can exert their inhibitory effects by competing with other signalling molecules including STATs for phosphorylated tyrosine motifs in the receptor complex.

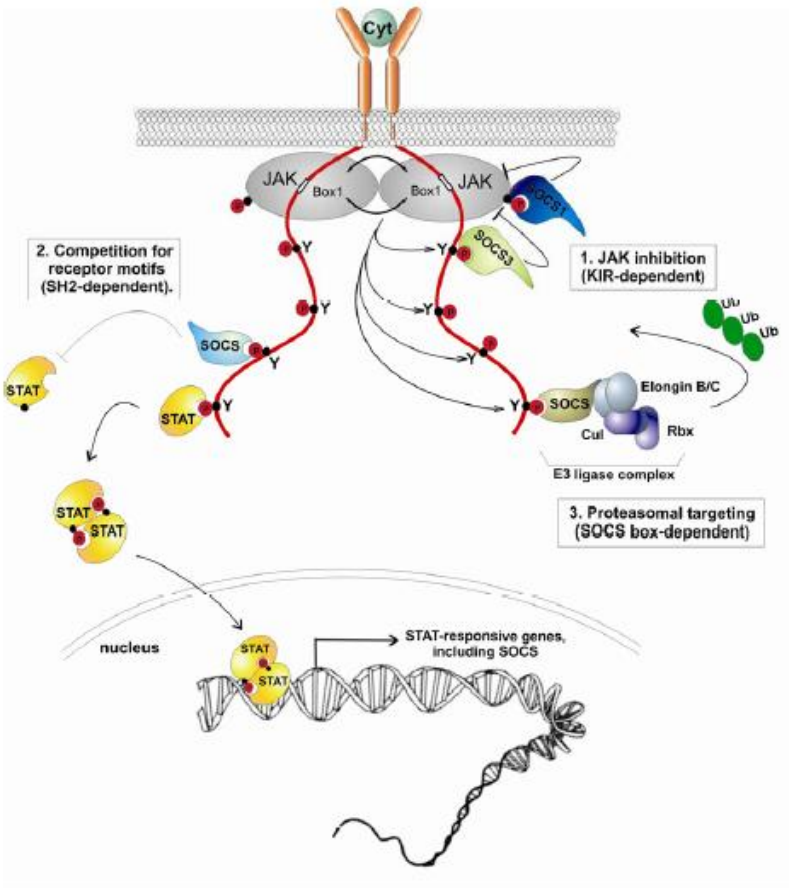


***Fig*. 1.6: Schematic representation of molecular mechanisms of SOCS actions**

SOCS1 directly interacts with all JAK members, thereby inhibiting their catalytic activity (38). It targets the phosphotyrosine at position Y1007, which lies within the activation loop of JAK2 (39,37). Activation of the kinase is dependent on phosphorylation of this particular tyrosine. The SOCS1 SH2 domain is sufficient for JAK2 interaction. However, as above mentioned, an additional region of

approximately 30 residues immediately N-terminal to the SH2 domain (the ESS) and the KIR are additionally required for high affinity binding and inhibition of JAK2 activity (39,37). Remarkably, SOCS1 has also been shown to bind directly to the type I and type II IFN receptors, which might provide a very efficient inhibitory effect of SOCS1 on IFN signaling (41) (*Fig*. 1.6).

1.4.3. The SOCS box: proteasomal targeting and protein stability

The C-terminal SOCS box is conserved throughout the SOCS family, thus suggesting an important role for this region in the function and regulation of SOCS proteins. Different SOCS box containing protein families have been identified: clues to understand the function of the SOCS box came from the initial finding that the Von Hippel-Lindau (VHL) tumor suppressor protein interacts with its Elongin B/C complex via the SOCS box, linking VHL to the ubiquitin/proteasome degradation system (42). The Elongin B/C-Cullin-SOCS box (ECS) complex was characterized as a family of E3 ubiquitin ligases in which the SOCS box protein acts as the substrate recognition unit of the complex (43). The E3 ubiquitin ligase participates in the polyubiquitin tagging of associated proteins and is responsible for substrate specificity. In this way, SOCS box proteins may suppress signalling by linking associated signalling components for degradation (44). Consistent with this idea, proteasomal inhibitors block the inhibitory functions of SOCS and induce sustained JAK-STAT signalling (45). The functional importance of the SOCS box was confirmed by the generation of transgenic mice expressing a C-terminal truncated variant of SOCS1 or SOCS3. These mice exhibit impaired regulation of respectively INFγ and G-CSF signalling, suggesting that the SOCS box is essential for complete SOCS activity (46).

1.5. SOCS proteins in disease: Inflammatory diseases

SOCS proteins are crucially implicated in the regulation of JAK/STAT signalling in inflammation. Accordingly, alterations in SOCS protein levels, and more specific SOCS1, has been associated with the pathogenesis of various inflammatory diseases including psoriasis. SOCS1 expression is induced by a wide variety of inflammatory and anti-inflammatory cytokines, including IFNγ, IL-3, IL-6 and IL-10 and will function to counteract STAT1 activation associated with inflammatory disorders. Several data show that the over-expression of SOCS1 in keratinocyte

clones, abrogates the IFN-γ-induced expression of many pro-inflammatory genes such as ICAM-1 (Inter-Cellular Adhesion Molecule-1) and the MHC (Major Histocompatibility Complex) class II molecules, and the release of related chemokines, by inhibiting JAK2 and STAT1 protein activities. Furthermore, the protection against SOCS1-mediated cytokine-induced cytotoxicity is correlated with a significant inhibition of the transcription factor interferon regulatory factor-1 (IRF-1) (47,48). The inactivation of kinase activity of JAK2 by SOCS1 impedes the IFN-γ-induced tyrosine phosphorylation of IFN-γ receptor (IFN-γ R)α subunit and consequently STAT1 activation (49,50) (*Fig.* 1.7). Hence the lack of STAT1 function in keratinocytes over-expressing SOCS1 impedes the expression of pro-inflammatory proteins such as ICAM-1 and MHC class II membrane molecules, as well as CXCL10, CXCL9, and CCL2 in response to IFN-γ, blocking the inflammatory process (50).

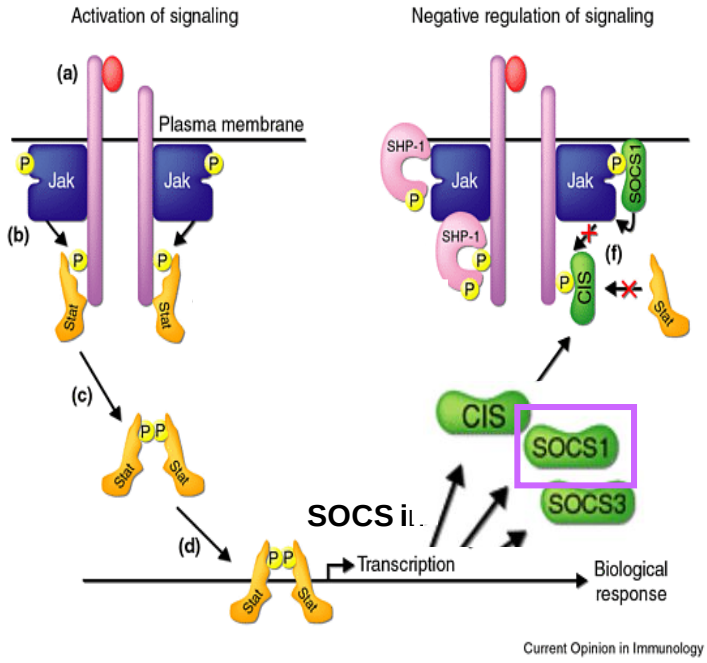


Fig*. 1.7: Cytokine signal transduction.** (**a**) Cytokines induce the dimerization of their cognate receptors, which leads to juxtaposition and activation of Janus kinases (JAKS). (**b**) JAKs phosphorylate the cytoplasmic domains of receptors, creating docking sites for cytoplasmic proteins such as signal transducers and activators of transcription (STATs). (c*)** Following phosphorylation by JAKs, STATs form dimers and (**d**) migrate to the nucleus to activate target genes, including those that encode the suppressors of cytokine signaling (SOCS). Once produced, members of the SOCS protein family (including SOCS-1, SOCS-3 and CIS) act back on the JAK–STAT pathway to temper signal transduction.

1.6. Therapeutic applications of SOCS

Therapeutic application of SOCS proteins are valuable for the treatment of diseases in which overshoot cytokine signalling is involved, such as those related with inflammation or cancer. One approach would be the overexpression of SOCS molecule, indeed it has been demonstrated that the adenovirus-mediated administration of SOCS3 prevents the development of RA in experimental mouse models (51) and fusion of SOCS proteins with membrane-permeable peptides might be an alternative way to introduce them into cells. In this context, intracellular delivery of a cell-penetrating form of SOCS3 was proven to be effective for treatment of various types of inflammation and septic shock (52). A second approach is related to the identification of small-molecule mimetics of SOCS proteins. The tyrosine kinase inhibitor peptide (Tkip) is an example of a mimetic of SOCS1 and effectively inhibits JAK2-mediated phosphorylation of STATs. This peptide prevented the development of experimental allergic encephalomyelitis (EAE) in mice (53) and blocked the proliferation of prostate cancer cell lines (54). A third approach might be blocking SOCS degradation in vivo. Alternatively, the structural insights in SOCS interactions may provide a basis for the design of SOCS inhibitors.

1.7. Aim of the work

Johnson and coworkers investigated the anti-inflammatory effects of SOCS1 employing peptides mimicking KIR domain: they did not inhibit JAK2 autophosphorylation but they were able to block STAT1 activation (similar to the pattern or profile of SOCS-1 inhibition of phosphorylation) in vitro and in vivo inhibiting IFN-γ activation of Raw 264.7 murine macrophages and Ag-specific splenocyte proliferation (55). Thus, they have shown in that studies (55) that the KIR region of SOCS-1 can directly bind to the autophosphorylation site of JAK2.

The purpose of the project was the development of new peptide sequences mimicking KIR capable of to mime SOCS-1 function by inhibiting the JAK2/STAT1 cascade induced by IFN-γ in human keratinocytes through binding and inhibition of JAK2 kinase activity. In order to select new peptides with improved binding properties for the target protein, a SARs studies of new peptides were carried out by using multiple approaches to chemical synthesis and biochemical assays.

These findings with SOCS-1 mimetics and antagonists have implications for novel therapeutic approaches to mimicking SOCS-1 for treatment of inflammatory diseases and for suppressing SOCS-1 to enhance the immune response against cancer and infectious diseases.

2. EXPERIMENTAL SECTION

2.1. Materials

N-α-protected Fmoc-amino acid derivatives and coupling reagents for peptide synthesis were from Inbios (Pozzuoli, Napoli). DIEA, Rink amide MBHA resin, tri-isopropyl-silane (TIS) and reagents for ELISA assays were from Sigma-Aldrich (Milano, Italy). HPLC-grade solvents and trifluoroacetic acid (TFA) were from LabScan (Stillorgan, Dublin, Ireland) and were used as received unless otherwise stated. Other reagents and chemical suppliers are indicated in the section of methods. Solid phase peptide syntheses was performed on a fully automated multichannel peptide synthesizer Syro I (Multisynthech, Germany). Preparative RP-HPLC were carried out on a Shimadzu LC-8A, equipped with a SPD-M10 AV detector and with a Phenomenex C18 Jupiter column (50x22 mm ID; 10 μm). LC-MS analyses were carried out on an LCQ DECA XP Ion Trap mass spectrometer equipped with an OPTON ESI source, operating at 4.2 kV needle voltage and 320 °C and with a complete Surveyor HPLC system, comprising a MS pump, an autosampler and a photo diode array (PDA). Narrow bore 50x2 mm C18 BioBasic LC-MS columns were used for analyses. All ELISA assays to screen the synthetic peptides, were carried out using a fully automated system (Hamilton Robotics, Milano, Italy) comprising a liquid handler, an automatic arm, a washer and an automated Synergy 4 multi-wavelength reader (BIOTEK Instruments, Inc. Highland Park, VT, USA).

2.2. Peptide Synthesis

Peptide libraries and single sequences were prepared by the solid phase method on a 50 μmol scale following the Fmoc strategy and using standard Fmoc-derivatized amino acids RINK AMIDE resin (substitution 0.5 mmol/g) was used as solid support. Activation of amino acids was achieved using HBTU/HOBt/DIEA (1:1:2), whereas Fmoc deprotection was carried out using a 40% (v/v) piperidine solution in DMF. All couplings were performed for 15 minutes and deprotections for 10 minutes. Randomized positions were coupled using equimolar mixtures of 12 building blocks employing a 100-fold excess of each amino acid. The focused simplified peptide library was prepared in the positional scanning format following

pre-MIX methodologies (47). Peptides were removed from the resin by treatment with a TFA:TIS:H_2O (90:5:5, v/v/v) mixture, then they were precipitated in cold diethylether and lyophilized. Single peptides were purified by preparative RP-HPLC, and their identity was evaluated by LC-MS analysis. Peptide products were stored at -20 °C until use.

The MW distributions analysis of single sub-libraries was performed with LC-ESI mass spectrometry (averaging the mass spectra over the retention time window where peptides were eluted) and appeared consistent with the theoretical distribution of MWs calculated for the sequences present within single mixture and the homogeneous composition of each mixture was further assessed by amino acidic analysis. To perform cellular assays CPPs (Cell penetrating Peptides) conjugated peptides were synthesized and purified. Conjugated peptides were also fluoresceinated at their N-termini with Fluoresceine-βAla.

2.3. ELISA Binding and competition experiments

To perform ELISA assays, 96-well plates (Nunc, Milan, Italy) were coated with single Fmoc-peptides and Fmoc-peptide mixtures at fixed concentrations (0.0100 mM) in carbonate buffer 100 mM pH 9 for 16 h at 4°C. Some wells were filled with buffer alone and were used as blank. After coating, the plates were washed three times with PBS-T buffer (PBS containing 0.04% Tween-20) and blocked with 300 µL of PBS containing 1% BSA, for 2 h at 37°C. After washing, solutions at increasing concentrations of biotinylated Jak2 peptides between 0.010 and 0.600 mM in carbonate buffer 100 mM were added to coated wells and incubated 1 h at 37°C. After washing again with PBS-T, 100 µL of SA-HRP conjugated (Santa Cruz Biotechnology, California, USA) diluted 1:10000 with PBS was added. After 1 h incubation at 37°C, the plates were washed again with PBS-T and 100 µL of chromogenic substrate, o-phenylendiamine 0.4 mg/mL in 50 mM sodium-phosphate-citrate buffer pH 5.0 containing 0.4 mg/mL H2O2, were added and the reaction was stopped by adding 50 µL of 2.5 M H2SO4 in each well. The absorbance was measured at 490 nm. Data points were in triplicates or duplicates and reported experiments are from three independent assays. Competitive ELISA experiments were carried out coating with Fmoc-KIR (0.010 mM) and pre-incubating biotinylated JAK2 peptides at 0.100 mM with PS-peptides in dose-response experiments in concentrations range $1*10^{-5}$ ÷ 0.100 mM. Data were

reported subtracting signals deriving by non-coated wells. Data were averaged by three independent experiments. Experimental data were fitted with GraphPad Prism, vers. 4.00, GraphPad Software (San Diego, California).

2.4. Stability Assay

Tat-PS-5 and Tat-KIR were dissolved in water at a 10 mg/mL concentration. The samples were then serially diluted in PBS, pH 7.3, containing 10% fetal calf serum (FCS) to obtain 100 µg/mL solutions and incubated at 37 °C for 20 h. 10 µL aliquots (1.0 µg of total peptide) were removed every two hours. Samples were immediately centrifuged for 5 min at 14,000 rpm and analyzed by reverse phase-HPLC after discarding the pellet. The experiment was carried out in duplicate and the data were reported as residual peptide in solution (in percentages) versus time.

2.5. Peptide processing in cellular extracts

Keratinocyte protein extracts for MS analysis were prepared using keratinocyte cultures pre-treated with PS-4, PS-5 and PS-KIR peptides and, then, treated with IFN-*γ* for 3h. Cells were lysed using a buffer containing 20 mM Tris-HCl, 150 mM NaCl, 1% Triton, and, then, analyzed by RP-HPLC-MS using a C18 column and mass spectrometer in SIM (Single Ion Monitoring)-mode.

2.6. Keratinocyte cultures and treatments with peptides

Normal human keratinocytes were obtained from skin biopsies of healthy volunteers (n = 3), as previously reported [28]. Briefly, cells were cultured in the serum-free medium, keratinocyte growth medium (KGM; Clonetics, Walkersville, MD, USA), for at least 3–5 days (at 60 – 80% confluence). Stimulation with 200 U/ml rh IFN-*γ* (R&D Systems, Minneapolis, MN, USA) was performed in keratinocyte basal medium (KBM; Clonetics). Cultures were treated with 10 µM of PS-4, PS-5 and KIR for two hours before adding IFN-γ, and then processed for the different experiments.

2.7. Fluorescence microscopy and FACS analysis

Keratinocyte cultures treated for 24 h with peptides tagged with TAT or pVec sequences and conjugated with the fluorocrome FITC were analyzed by

fluorescence microscopy, prior fixation with 3% paraformaldehyde/PBS. Coverslips were mounted and slides analyzed using a Zeiss LSM 510 meta-confocal microscope (Zeiss, Oberkochen, Germany). Otherwise, peptide-treated keratinocytes were analyzed by flow cytometry with a FACScan equipped with Cell Quest software (Becton Dickinson, Mountain View, CA).

2.8. Western blotting analysis

Subconfluent keratinocyte cultures were pre-treated for 2 h with 10 mM PS-4, PS-5 and KIR, then, stimulated for 15 min or 3 h with rh IFN-γ. Cells were lysed in a buffer containing 20 mM Tris-HCl, 150 mM NaCl, 1% Triton. Cell lysates were centrifuged, and supernatants directly evaluated for total protein amount with the BioRad Protein Assay (BioRad, Hercules, CA). Equal quantity of total proteins were separated by 10% SDS-gel electrophoresis and transferred onto a nitrocellulose filter (Hybond-ECL, GE Healthcare). Protein detection was performed using the following antibodies: anti-phosphotyrosine (PY) STAT1 (Tyr701) (Santa Cruz Biotechnology, Santa Cruz, CA), anti-STAT1 (Santa Cruz Biotechnology), anti–IRF-1 (M-20; Santa Cruz Biotechnology), and anti-β-actin (C-11; Santa Cruz Biotechnology), and a chemiluminescence detection system (GE Healthcare, Amersham, UK). Relative intensity of signals was quantified using a GS-710 densitometer (Bio-Rad).

2.9. Circular Dichroism (CD) Spectroscopy

CD spectra were recorded on a Jasco J-810 spectropolarimeter (JASCO Corp, Milan, Italy). CD spectra were registered at 25 °C in the far UV region from 190 to 260 nm. Each spectrum was obtained averaging three scans, subtracting contributions from corresponding blanks and converting the signal to mean residue ellipticity in units of deg*cm^{2}*dmol^{-1}* res^{-1}. Other experimental settings were: 20 nm/min scan speed, 2.0 nm band width, 0.2 nm resolution, 50 mdeg sensitivity, and 4 sec response. The concentration of peptides was kept at 10×10^{-5} M and a 0.1 cm path-length quartz cuvette was used. Spectra were acquired in 10 mM phosphate buffer at pH 7.0.

3. RESULTS

3.1. Identification of crucial residues of KIR involved in JAK2 catalytic site recognition

The ability of KIR (SOCS1 52-67) domain to bind to the JAK2 peptide (JAK2 1001-1013) reproducing the JH1 domain both in the phosphorylated (Tyr1007) pJAK2, and in the non-phosphorylated form, JAK2 (Table 3.1) was already reported in previous studies (55).

Peptide	Sequence
JAK2	^{1001}LPQDKEYYKVKEP^{1013}
pJAK2	^{1001}LPQDKE**Y**YKVKEP^{1013}
MuIFN-γ	^{95}AKFEVNNPQVQR^{106}
KIR	^{52}DTHFRTFRSHSDYRRI^{67}
New KIR	^{52}DTHFRTFRSH^{61}
PS-5	DTC(Acm)RQTFRSH
PS-11	DTFPQTFRSH
PS-12	DTFFQTFRSH

Table 3.1: Peptide sequences utilized in this study. Bold Y indicates a phosphorylated tyrosine

In order to identify new sequences able to bind to JAK2, here we report the affinity of the binding employing an easy ELISA procedure that involves the immobilization of the synthetic peptides on the plate by means of the Fmoc group on their N-termini (56). Dose-dependent experiments confirmed that KIR has the same affinity against both JAK2 peptides (54), providing K_D values of 76 and 81 μM for JAK2 and pJAK2, respectively (Table 3.2 and *Fig*. 3.1).

	KIR	New KIR	PS-5		PS-12	
	K_D (μM)	K_D (μM)	K_D (μM)	IC50 (μM)	K_D (μM)	IC50 (μM)
JAK2	76 ± 3	27 ± 6	5.0 ± 1.0	3.3 ± 1.0	3.6 ± 1.0	2.5 ± 1.2
pJAK2	81 ± 8	16 ± 4	0.8 ± 0.1	2.4± 1.7	4.5± 0.4	2.8± 1.0

Table 3.2: K_D and IC_{50} values of KIR and PS peptides in the complex with JAK2 catalytic region (in phosphorylated and non-phosphorylated form).

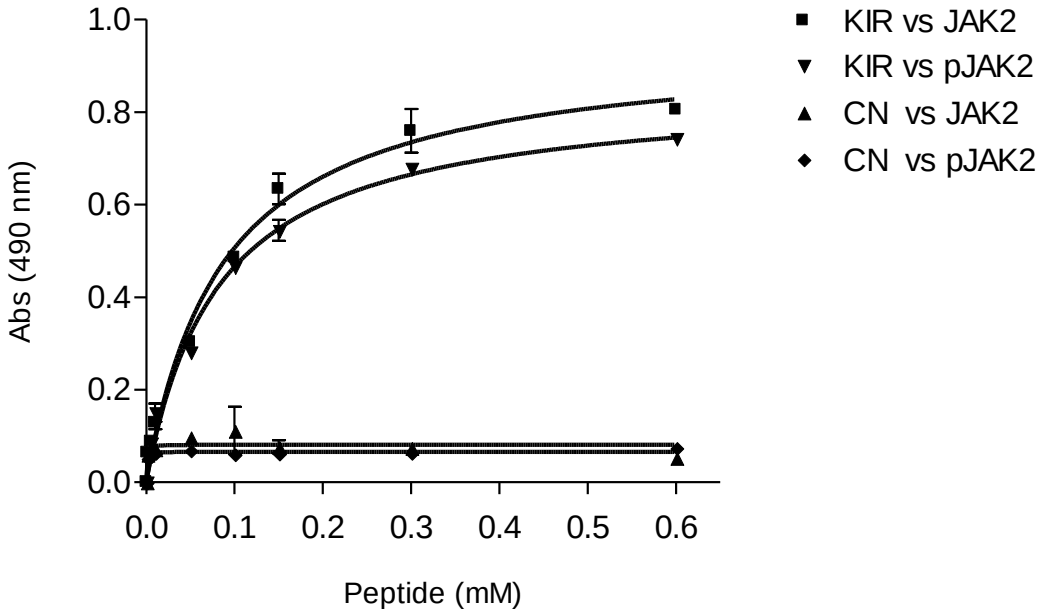


***Fig* 3.1: Direct ELISA assay**: Biotinylated JAK2 and pJAK2 at the concentrations range of 0.01-0.6 mM were added in triplicate to a 96-well plate coated with either KIR, negative control peptide (CN) (MuIFNγ (95–106)) at the concentration of 0.0100 mM, or buffer alone. Data have been reported subtracting signals deriving by no-coated wells.

In order to investigate crucial residues of KIR involved in JAK2 recognition, we synthesized and analyzed Ala-scan peptides of this region, in which the wild-type residues were systematically changed to alanine. Their binding capacities to both JAK2 peptides were analyzed through ELISAs, carried out at two different concentrations of coated Fmoc-peptides as reported in *Fig*. 3.2. The analysis of

these experiments revealed that peptides in which residues in positions 52 and 53 (D and T) and in 57-61 (TFRSH) were changed to alanine had lower affinity, while changes in positions 54-56 (HFR) and 62-67 (SDYRRI) did not affect the binding. These data suggested that the ability of KIR to recognize JAK2 catalytic site is confined within the region 52-61. To proof this hypothesis we synthesized and analyzed the binding properties of this shorter region of KIR, named new KIR (Table 3.1), and its direct ELISAs vs both JAK2 peptides, in comparison with KIR, are reported in *Fig*. 3.3. As KIR, also new KIR showed dose-dependent binding and data fitting provided values of K_D similar for JAK2 and pJAK2, 27 μM and 16 μM, respectively (Table 3.2).

Thus the deletion of six residues in the C-terminal region of KIR led to a shorter sequence able to bind to JAK2 catalytic site with greater affinity.

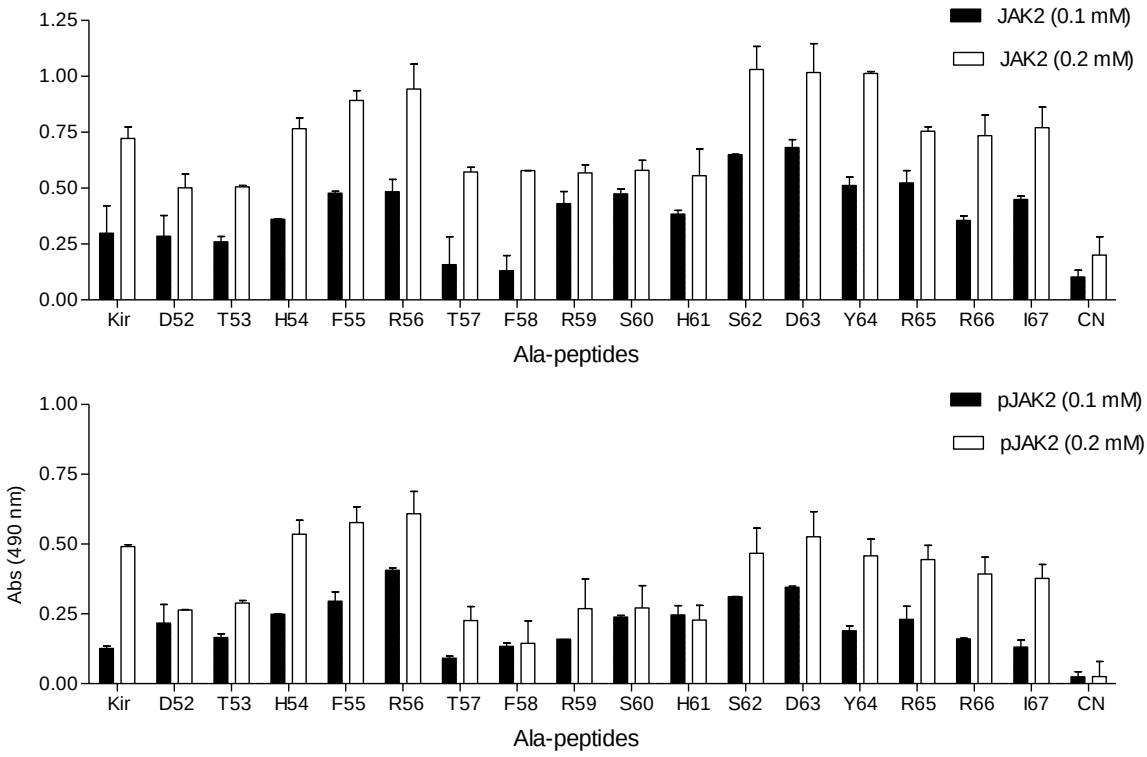


***Fig* 3.2: Direct ELISA assay:** Biotinylated JAK2 (**upper)** and pJAK2 (**lower**), at two different concentrations 0.100 – 0.200 mM were added in triplicate, to a 96-well plate coated with either KIR, Ala-scan peptides, CN (negative control: MuIFNγ) at the concentration of 0.010 mM, or buffer alone.

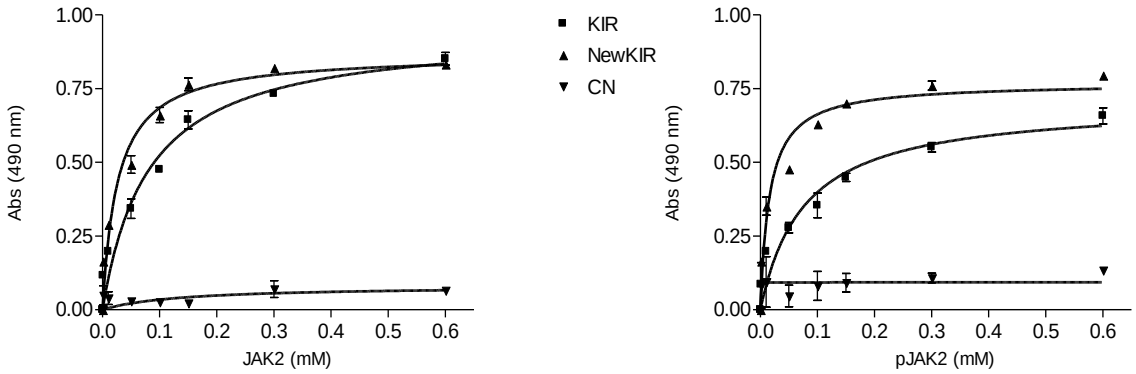


***Fig* 3.3: Direct ELISA assay:** Biotinylated JAK2 (**left**) and pJAK2 (**right**) catalytic site in a concentration range of 0.010 – 0.600 mM were added in triplicate to a 96-well plate coated with either KIR, new KIR, CN peptides at the concentration of 0.010 mM, or buffer alone.

3.2. Generation of new KIR-mimetics through a focused simplified peptide library

In order to identify new peptides able to mime KIR activity, we screened a focused simplified peptide library (56-57). This library was designed on the basis of Ala-scanning results, randomizing residues in positions 54-56, that appeared not directly responsible of the interaction with JAK2 catalytic region.

Our approach employs a minimum number (12 instead of 20) of non redundant amino acids as building blocks (Table 3.3).

	Selected amino acid	**Side chain**
1	Arg	basic
2	Asp	acidic
3	Gln	amide
4	Leu	aliphatic
5	Ala	hydrophobic
6	Pro	hydrophobic
7	Met	hydrophobic
8	Ser	polar
9	Phe	hydrophobic, aromatic
10	Cys(Acm)	aliphatic
11	Tyr	hydrophobic, polar, aromatic
12	His	hydrophobic, polar

Table 3.3: List of building blocks used to assemble the peptide library, Cys was chosen as acetamidomethyl derivative to avoid polymerization.

It is simplified in terms of a reduced number of screened compounds still capable to cover a wide chemical diversity, in a positional scanning format (PS-SSPL: Positional Scanning-Simplified Synthetic Peptide Library) (58).

The diversity of the library consisted of 123=1728 different compounds and the PS format implied the synthesis of 12x3=36 sub-libraries, each composed of 144 diverse sequences. Each sub-library was characterized by LC-MS and amino acid analysis (data not shown). Their experimental compositions were in agreement with theoretical ones. Peptide sub-libraries were tested to direct binding ELISAs, at one fixed concentration of coated mixtures (*Fig.* 3.4).

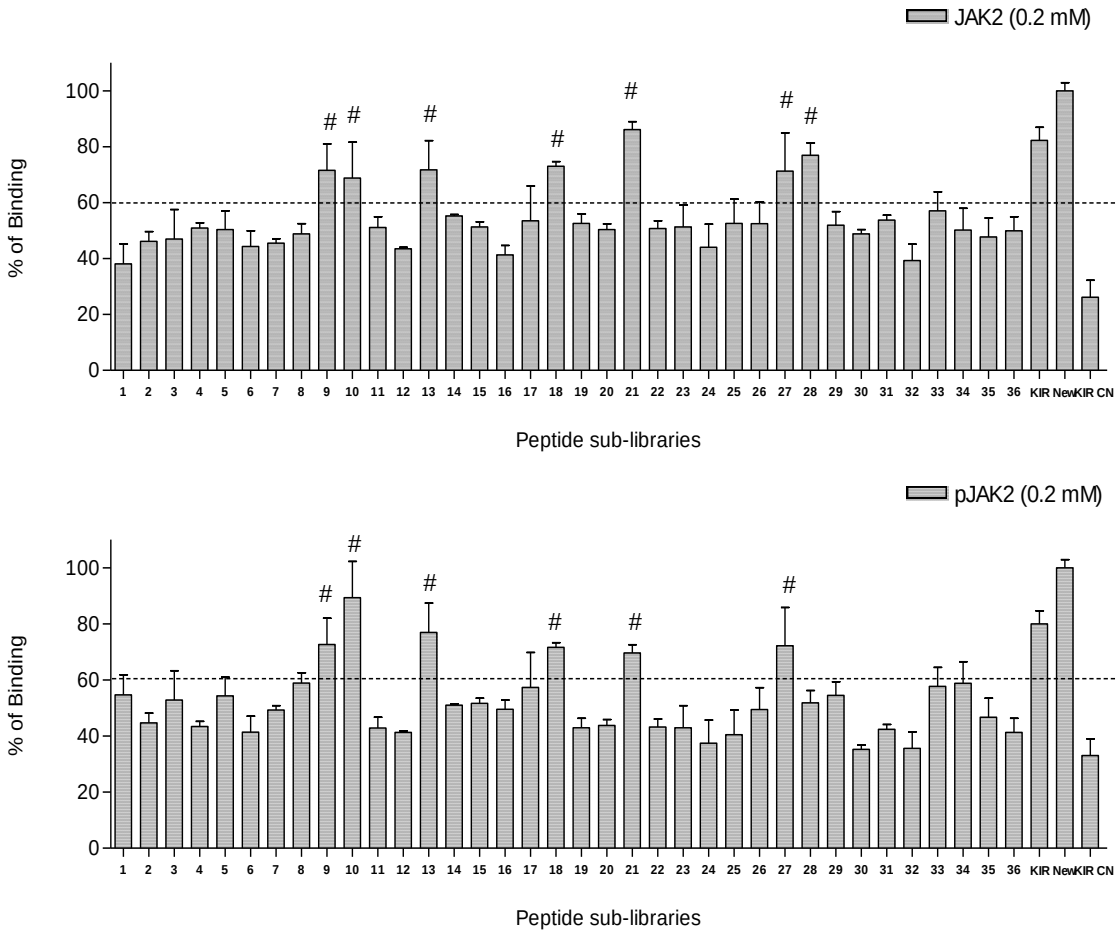


***Fig* 3.4: Direct ELISA assay:** Biotinylated JAK2 (**upper**) and pJAK2 (**lower**) at the concentration of 0.200 mM were added in duplicate to a 96-well plate coated with either KIR, new KIR, CN and PS-SSPL peptide sublibraries at the concentration of 0.010 mM, or buffer alone. The numeration from 1 to 12 indicates peptide mixtures having known residue at position 55, from 13 to 24 indicates known residue at position 56 and from 25 to 36 indicates known residue at position 57, (the number of each building block is reported in Table 3.3).

Results were expressed as percentage of binding considering new KIR as 100%. (0.01 mM of each Fmoc sub-library that is $6,9*10^{-5}$ mM of each single sequence).

By data analysis, none of the screened sub-libraries showed an improved affinity toward JAK2 peptides respect to new KIR and KIR and this can be ascribed to a concentration effect. Indeed each sequence in the mixtures is 144-fold lesser concentrated respect to wild-type KIR single peptides. Nevertheless, from one screening cycle, we have selected several amino acids for each position, combining experiments involving JAK2 and pJAK2: Phe and Cys(Acm) on position 54 Arg, Phe and Pro on P55 and Gln and Leu on P56 were selected.

The combination of identified amino acids in each position led to the design of 12 single peptides whose sequences are reported in Table 3.4.

Name	Sequence
New KIR	DTHFRTFRSH
PS-1	DTC(Acm)PLTFRSH
PS-2	DTC(Acm)RLTFRSH
PS-3	DTC(Acm)FLTFRSH
PS-4	DTC(Acm)PQTFRSH
PS-5	DTC(Acm)RQTFRSH
PS-6	DTC(Acm)FQTFRSH
PS-7	DTFPLTFRSH
PS-8	DTFRLTFRSH
PS-9	DTFFLTFRSH
PS-10	DTFRQTFRSH
PS-11	DTFPQTFRSH
PS-12	DTFFQTFRSH

Table 3.4: Composition of peptide sequences deriving from the screening of focused PS-SSPL.

They were synthesized and purified by RP-HPLC and identified by LC-MS (data not shown). Their binding capacities to JAK2 peptides were analyzed through direct ELISA at three fixed concentrations of coated peptides as reported in *Fig*. 3.5. This experiment allowed to select sequences with an enhanced affinity respect to new KIR toward JAK2 peptides: PS-11 and PS-12 were selected from the experiment involving JAK2, while PS-5 from that with pJAK2; from both assays PS-4 was chosen as control sequence for further analyses.

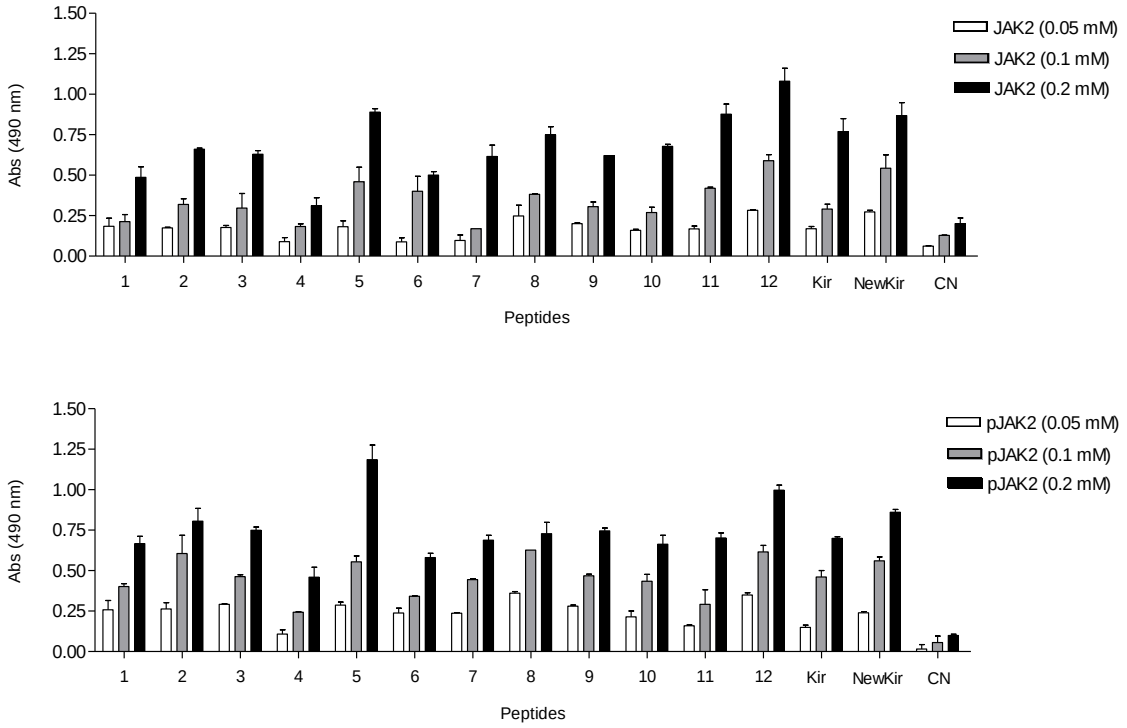


***Fig* 3.5: Direct ELISA assay:** Biotinylated JAK2 (**upper**) and pJAK2 (**lower**) were added in duplicate at three concentrations (0.050, 0.100, 0.200 mM) to a 96-well plate coated with either KIR, new KIR, CN and single PS-peptides at the concentration of 0.010 mM or buffer alone.

3.3. Affinity of new KIR mimetic peptides toward JAK2 catalytic site

Single peptides selected from PS-SSPL screening were analyzed in dose-response direct binding ELISA and related experiments are reported in *Fig*. 3.6 A.

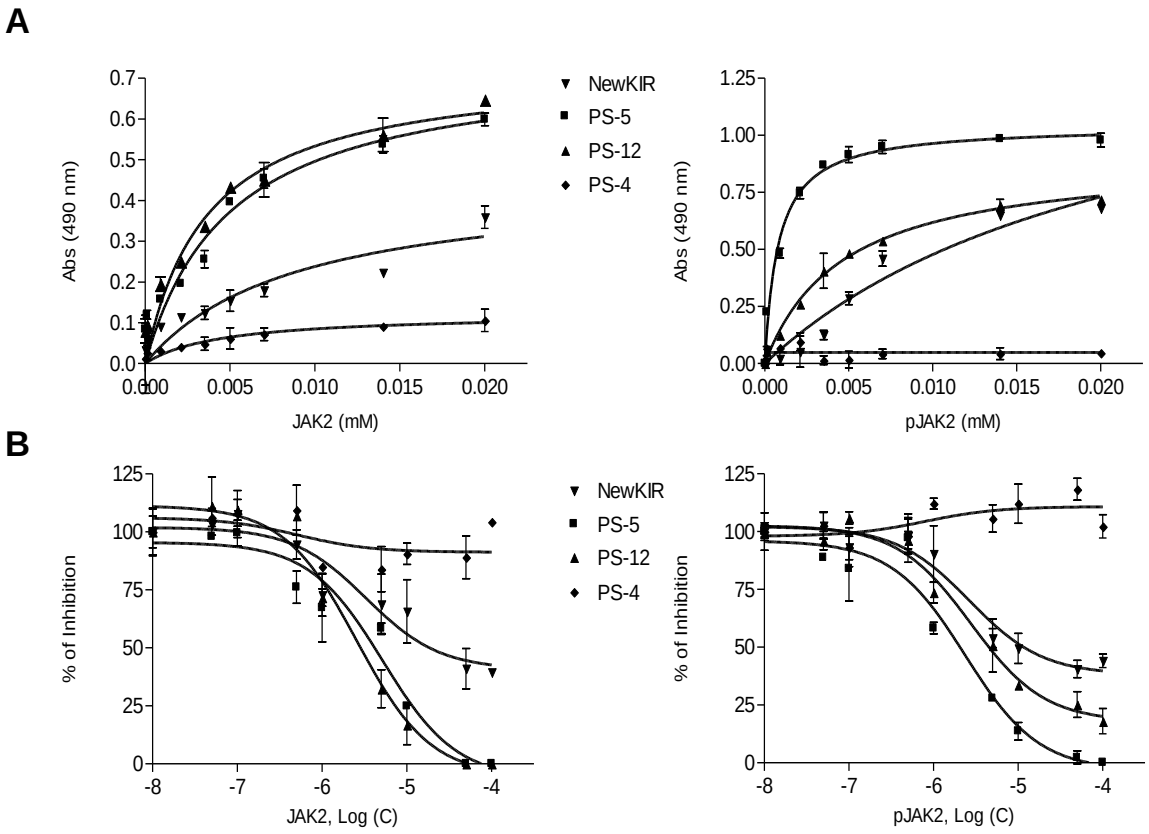


***Fig.* 3.6: Direct ELISA assay:** (**A**) Biotinylated JAK2 (**left**) and pJAK2 (**right**) in a concentrations range of 0.001 – 0.020 mM were added, in duplicate, to a 96-well plate coated with single peptides deriving from PS-screening, PS-5, -12, -4, at the concentration of 0.003 mM, or buffer alone. (**B**) Biotinylated JAK2 (**left**) and pJAK2 (**right**) at pre-saturation concentration (0.100 mM) were incubated with PS-5, -12, -4 in a concentrations range of 1*10-5 – 0.100 mM, and were added in duplicate to a 96-well plate coated with Fmoc-KIR at a concentration of 0.01 mM, or buffer alone.

Data fitting provided K_D values in the high nanomolar range, as reported in Table 3.2. PS-5 and PS-12 presented K_D values 15-fold lower respect to wild type KIR, and noteworthy, PS-5 showed a 5-fold higher affinity toward pJAK2 compared to JAK2. PS-11 sequence, that bears a Proline residue at position 54, showed lesser affinity and seems to be unable to bind pJAK2 in the explored concentration range (data not shown). In order to further validate our findings, we inverted ELISA

experiments, immobilizing Fmoc-JAK2 peptides and evaluating the binding with biotinylated PS-sequences: dose-dependent experiments provided similar values for K_Ds (data not shown). Also in this experiment PS-11 was not able to bind to coated JAK2 peptides and was neglected in further studies.

With the aim to assess binding specificity and to determine whether PS-peptides and natural sequences bind to the same site of JAK2 peptides, competition experiments were carried out employing PS-5 and PS-12 as competitors of the KIR-JAK2 peptide complexes. For this purpose, Fmoc-KIR was immobilized on the ELISA plate, then JAK2 peptides at a pre-saturation concentration of 0.100 mM were pre-incubated with new KIR, PS-5, -12 and -4 at different molar ratios. As shown in *Fig.* 3.6 B, both PS-5 and -12 peptides inhibited the binding of JAK2 peptides to KIR in a dose-dependent fashion and data fitting provided IC_{50} values in the low micromolar range which are consistent with K_D values estimated by direct binding (Table 3.2); new KIR peptide resulted less effective in competing with KIR.

	Sequence
PS-5	DTC(Acm)RQTFRSH
PS13	DTC(Acm)**K**QTFRSH
PS14	DTC(Acm)R**N**TFRSH
PS15	DTC(Acm)**KN**TFRSH
PS12	DTFFQTFRSH
PS16	DT**W**FQTFRSH
PS17	DT**WW**QTFRSH
PS18	DTF**W**QTFRSH
PS19	DT**W**F**N**TFRSH
PS20	DT**WWN**TFRSH
PS21	DTF**WN**TFRSH
PS22	DTFF**N**TFRSH

Table 3.4: Composition of new sequences analogs of lead peptides (underlined) deriving from the screening of PS-SSPL, in bold "new" residues are reported

In order to follow through the simplified approach, we designed 10 new peptides on the basis of first generation lead compounds: PS-5 and PS-12 (Table 3.4).

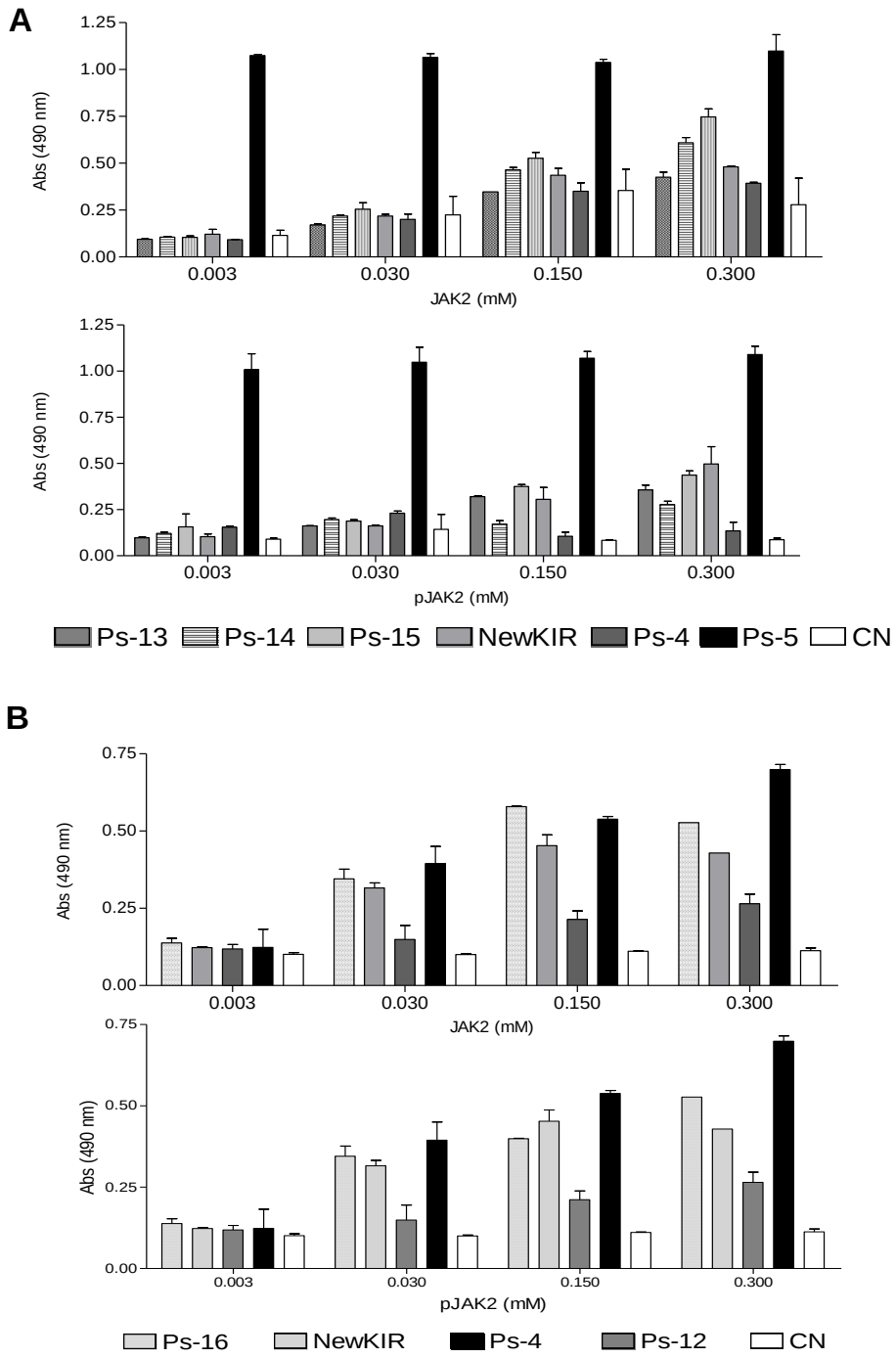


***Fig.* 3.7: Direct ELISA assay:** Biotinylated JAK2 (**upper**) and pJAK2 (**lower**) at the concentrations of 0.003, 0.030, 0.150 and 0.300 mM, were added in triplicate to a 96-well plate coated with new PS sequences designed on the basis of (**A**) PS-5 and (**B**) PS-12 (to simplify only PS-16 signals are reported, those of PS- 17-22 are almost the same), or buffer alone. PS-4 and CN were used as negative control.

They were built substituting Arg with Lys, Gln with Asn and Phe with Trp. The analysis of binding data showed that new sequences had less affinities toward JAK2 peptides compared with PS-5 and PS-12 (*Fig.* 3.7 A-B), even if PS- 16-22, at 0.150 and 0.300 mM, provided signals close to those of PS-12. From these experiments and those reported in *Fig.* 3.6 A, PS-12 seemed to be less effective and specific to bind to JAK2 catalytic site, respect to PS-5.

3.4. Binding Properties of PS sequences toward full-length JAK2

To validate the specificity of the recognition mechanism shown by PS mimetics of KIR, we carried out direct binding ELISAs of PS-5 in comparison with KIR, employing full-length JAK2 protein. Dose-response experiments are reported in *Fig.* 3.8; data fitting resulted in a K_D value of 0.43 ± 0.09 μM, similar to that showed toward JAK2 peptides, while KIR gave no measurable responses.

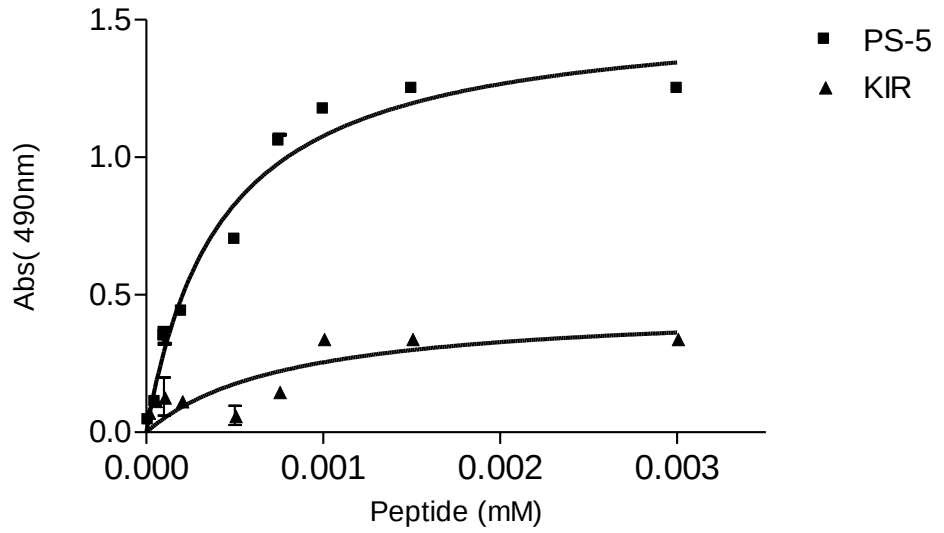


***Fig.* 3.8: Direct ELISA assay:** Biotinylated KIR and PS-5 in a concentration range of 0.0100 – 300 μM were added in triplicate to a 96-well plate coated with JAK2 protein at the concentration 0.1 μM, or buffer alone.

3.5. Structural characterization of new KIR mimetic peptides by Circular Dichorism

To evaluate the impact of intrinsic structural preferences of KIR mimetics on their ability to bind to JAK2 peptides, KIR, new KIR, PS-4 and PS-5 sequences were characterized by far-UV circular dichroism spectroscopy (CD).

In a predicted model of SOCS1, KIR-ESS domain contained two α- helices (R57–R69 and L74–A77) linked by a small coil segment (39).

The CD spectrum of KIR presents a minimum at ~ 200 nm and a shoulder at ~ 220 nm (*Fig.* 3.9); these features suggest, as expected, mixed conformational states, in which an unfolded state coexists with a helical conformation. The propensity of this domain to adopt helical conformation was also confirmed by TFE titration (*Fig.* 3.10 A).

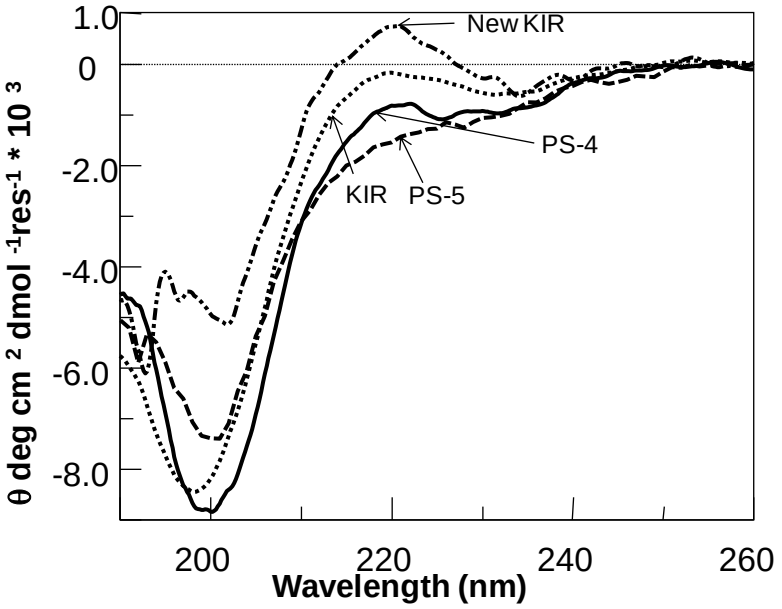


***Fig.* 3.9**: **CD spectra of the PS peptides in phosphate buffer 10 mM, PS-5, PS-4, KIR new KIR, as reported in figure legend.**

The truncated sequence of new KIR showed a similar CD profile also upon the addition of TFE (*Fig*. 3.10 B). Punctual mutations in PS-peptides did not dramatically change the spectrum (*Fig.* 3.9); and furthermore while PS-5 showed a clear propensity to adopt helical conformations upon TFE addition(*Fig.* 3.10 C), the negative control sequence PS-4 does not (*Fig.* 3.10 D).

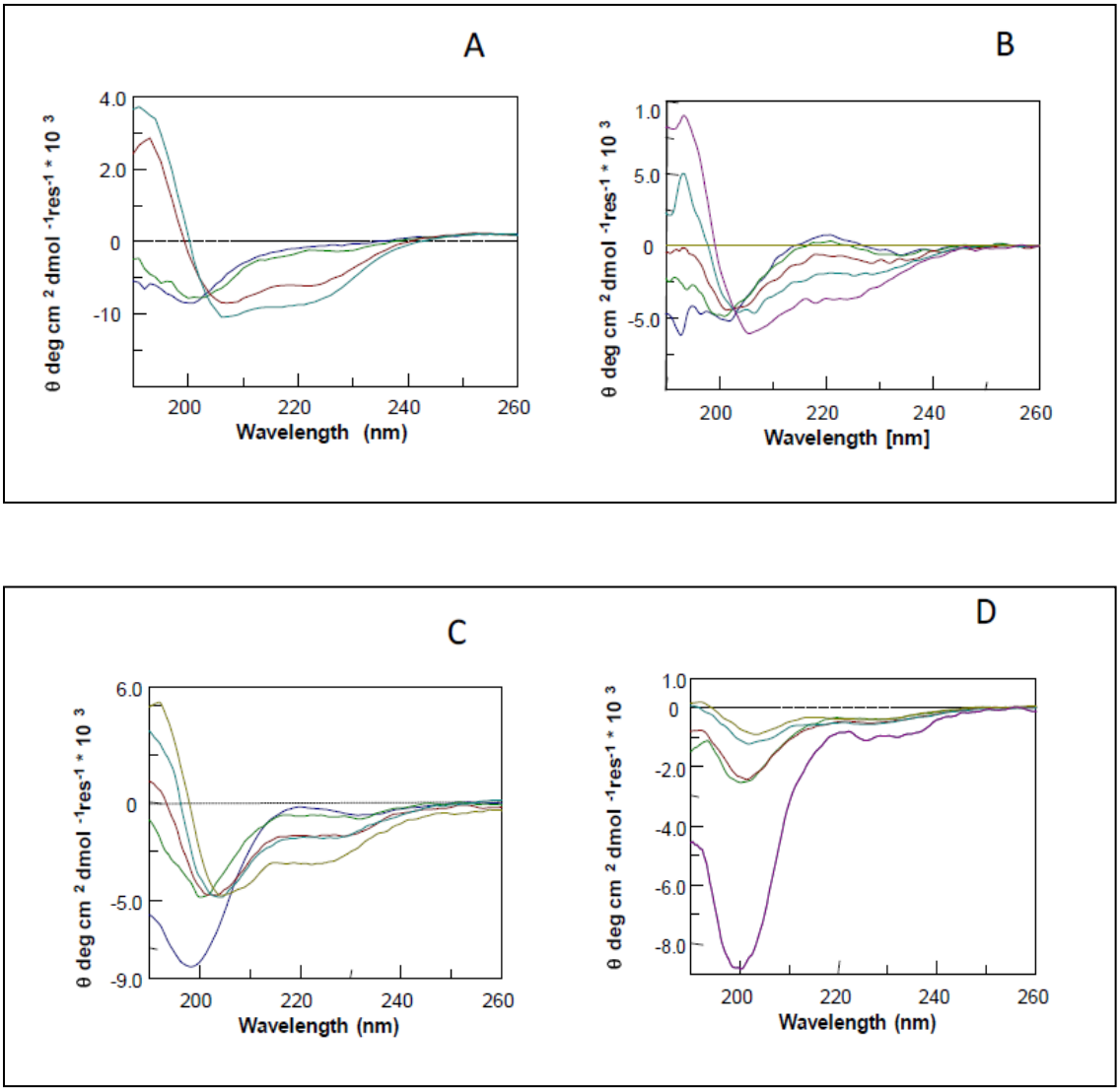


***Fig.* 3.10: Overlay of CD spectra of the peptides in phosphate buffer 10 mM with increasing amounts of TFE, from 0 to 70 % (v/v): (A) KIR, (B) NewKIR, (C) PS-5, (D) PS-4.**

3.6. Peptide uptake and stabilities studies in primary cultures of human keratinocytes

To confirm the efficacy of the selected new peptides to bind JAK2 and to be valuable mimetics of SOCS1, we planned cellular experiments using PS peptides conjugated to two different cell penetrating peptides (CPP): the fragment 48-60 of the HIV Tat protein, and pVec, derived from the murine vascular endothelium cadherin (58). They were positioned at the N-terminus and C-terminus, respectively, of PS-peptides in order to speculate on a potential directional preference in the recognition mechanism between KIR mimetics and JAK2 peptides.

Firstly, we assessed whether the Tat and pVec conjugated peptides were still capable to bind to JAK2 peptides, testing them in ELISA dose-response experiments, in a concentration range of 0.0500-0.100 mM for biotinylated JAK2 peptides on plate coated with conjugated peptides PS-5, -12, -4 at the concentration of 5.0 µM. CPP-conjugated peptides provided K_D values similar to those of non conjugated peptides, as reported in *Fig*. 3.11 and table 3.5.

	TatPS-5	**TatPS-12**	**pVecPS-5**	**pVecPS-12**
	KD (µM)	**KD** (µM)	**KD** (µM)	**KD** (µM)
JAK2	3.0± 0.7	8±0.9	4.1± 0.6	4.7± 0.3
pJAK2	1.8± 0.1	6.5± 0.2	6.0± 0.3	5.9± 0.4

Table 3.5: KD values of KIR mimetic peptides conjugated whit CPPs in the complex with JAK2 and pJAK2.

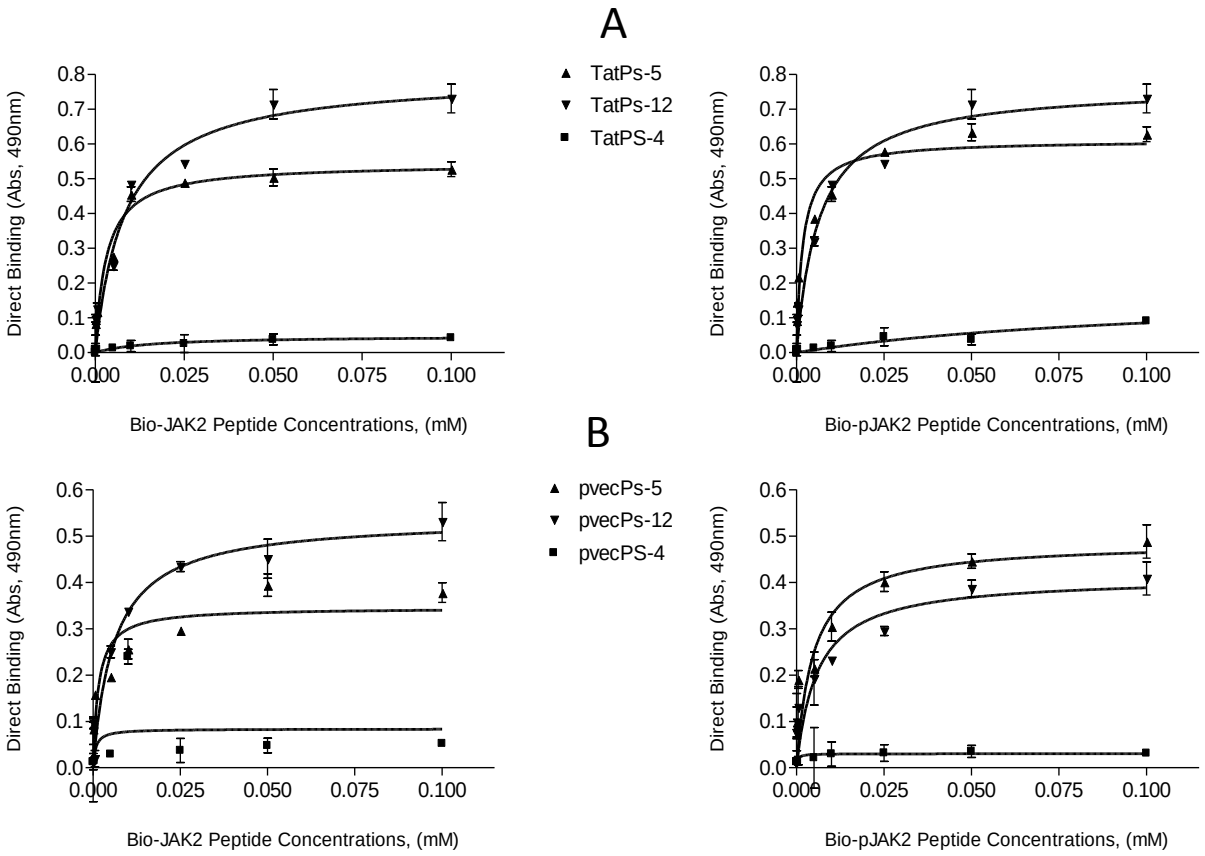


***Fig.* 3.11: Direct ELISA assay:** Biotinylated JAK2 (**right**) and pJAK2 (**left**) at the concentrations range of 0.005-0.1 mM were added in triplicate to 96-well plates coated with conjugated PS- sequences: (**A**) Tat fused PS-5, -12, -4 and (**B**) pVec fused to PS-5, -12, -4, at the concentration of 0.005 mM, or buffer alone. Data have been reported subtracting signals deriving by no-coated wells.

These data suggested that the interaction between PS and JAK2 peptides was not affected by the presence of CPPs and that any directionality in recognition mechanism was not detected.

Subsequently, we evaluated the entry of PS peptides in cultured keratinocytes using Tat and pVEC-conjugated PS peptides. Cells were treated for 6 hours at room temperature with different concentrations of FITC-conjugated peptides, PS-5, PS-12 and PS-4, then the cellular localization of peptides was analyzed by fluorescence microscopy and the uptake of conjugated peptides was quantified by FACS analysis (*Fig.* 3.12 A-B). By this experiment we established that N-term-Tat-conjugated peptides penetrated cells efficiently, with a percentage of penetration of 95%. The

penetration ability of pVec-conjugated peptides was less efficient (data not shown) and they were not further analyzed.

Peptide treatment of cells over time was in agreement with proteases stability assays carried out on PS-5; indeed in parallel we verified the stability of Tat- PS-5 in serum in comparison with Tat-KIR.

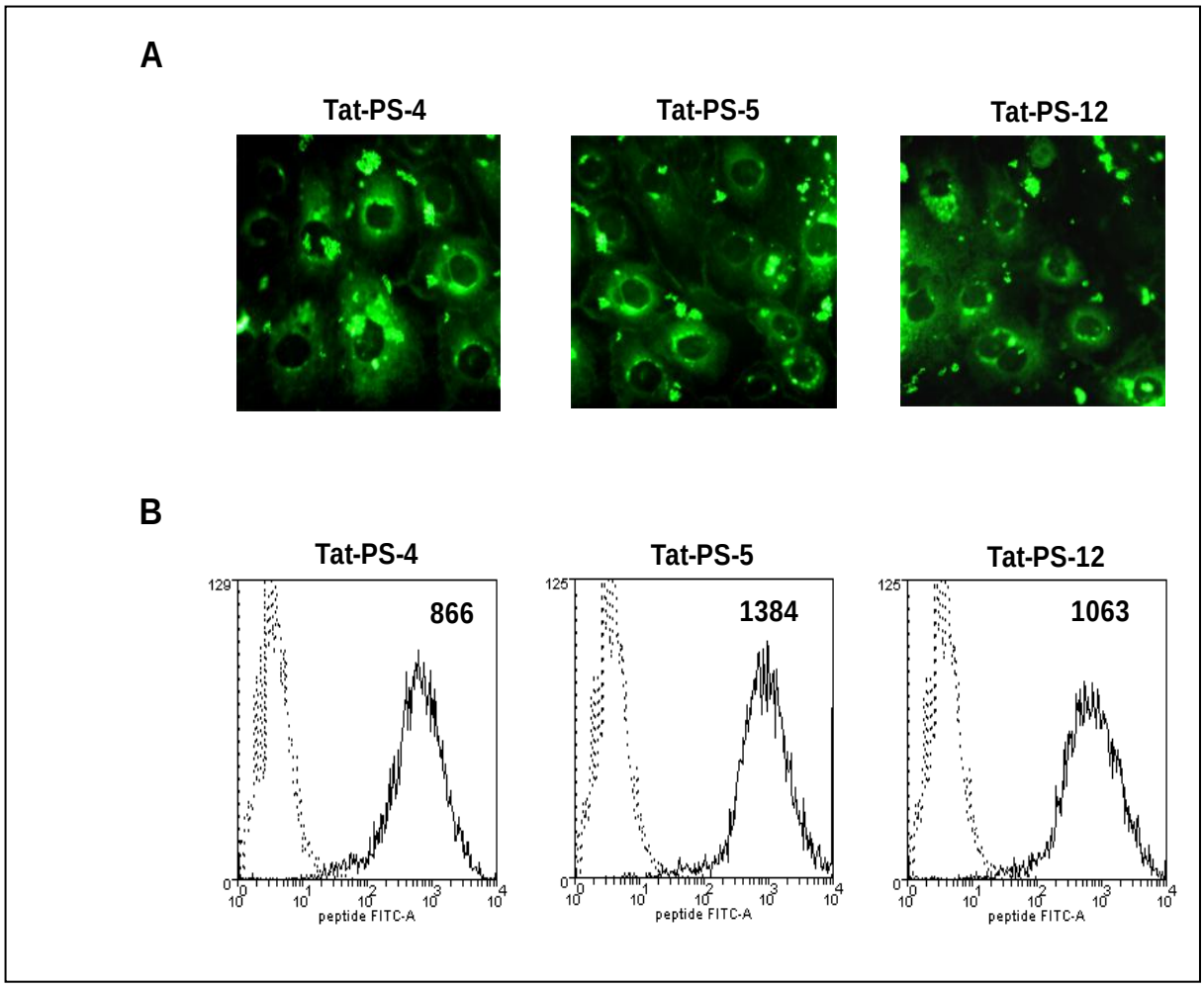


***Fig.* 3.12: TAT-conjugated PS-4, PS-5 and KIR efficiently enter primary cultures of human keratinocytes.** (**A**), Keratinocyte cultures treated for 24 h with PS-4, PS-5 and PS-12 peptides at 0.001 mM final concentration, all conjugated with TAT sequences and tagged to N-terminus with the fluorocrome FITC, were analyzed by fluorescence microscopy. Pictures of a representative experiment are shown at a 20X magnitude. (**B**), Keratinocytes treated with Tat-tagged- or not-tagged-PS-4, PS-5 and KIR were analyzed by flow cytometry with a FACScan.

As shown in *Fig.* 3.13 A-B, after 6h treatment Tat-PS-5 has a residual concentration greater than 75% of the initial one, while Tat-KIR has been degraded by 75% respect to initial concentration. The higher level of stability shown by Tat-PS-5 can be due to the presence of non-natural side-chain (Cys-(Acm)). Peptide stability was also checked in cells; extracts prepared from keratinocyte cultures incubated with peptides were analyzed by LC-MS, and unmodified peptides were detected by HPLC and mass spectrometry (data not shown).

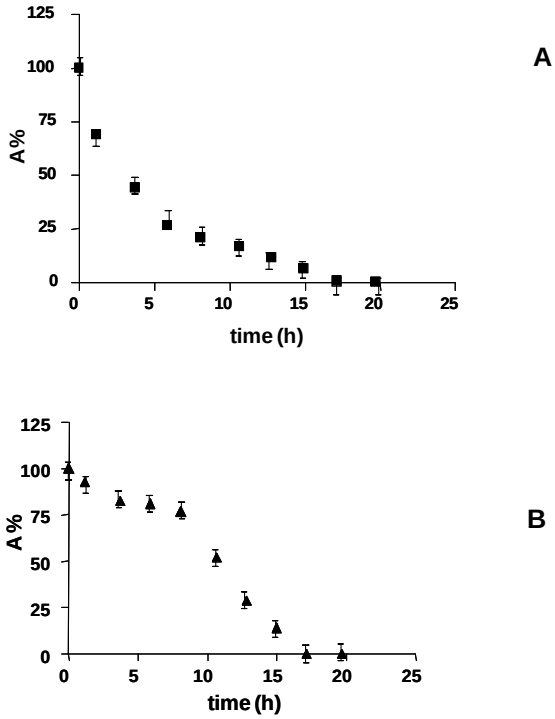


***Fig.* 3.13: Stability Assay**: The resistance to enzyme degradation of (**A**) Tat-KIR and (**B**) Tat-PS-5 was assessed by incubation in 10% fetal calf serum for 20h. At time 0, and after two hours, three aliquots were removed, centrifuged to remove proteins, and analyzed by HPLC. Residual peptide quantity, expressed as the percentage of the initial amount versus time, was plotted.

3.7. Evaluation of effects of KIR mimetic peptides on STAT1 signalling

To evaluate the effect of PS peptides on the JAK2-STAT1 signalling, we compared the ability of PS-5, KIR and PS-4 to regulate the activity of endogenous JAK2 in keratinocytes, in terms of inhibition of STAT1 activation in response to IFN-γ treatment. We found that treatments with PS-5, and less efficiently with PS-KIR, inhibited the STAT1 phosphorylation on Tyr 701 residue in keratinocyte cultures

stimulated with rh IFN-γ (*Fig.* 3.14), as compared to treatment with the negative control PS-4. The percentage of STAT1 phoshorylation inhibition by PS-5 and KIR was about 80 and 65, respectively, as assessed by densitometric analyses of Western blotting experiments (*Fig.* 3.14). In contrast, PS-4 peptide did not show significant effects on IFN-γ-induced activation of STAT1 (*Fig.* 3.14 A). Surprisingly, in spite of biochemical data previously shown, PS-12 appeared less able respect to PS-5 to block the STAT1 activation and was not further investigated. Furthermore, as a possible consequence of STAT1 inhibition by PS-5 and KIR, we observed a strong inhibition of the IFN-γ-induced accumulation in keratinocytes of IRF-1, that is a transcription factor whose expression is strictly dependent on STAT1 activity (*Fig.* 3.14 B).

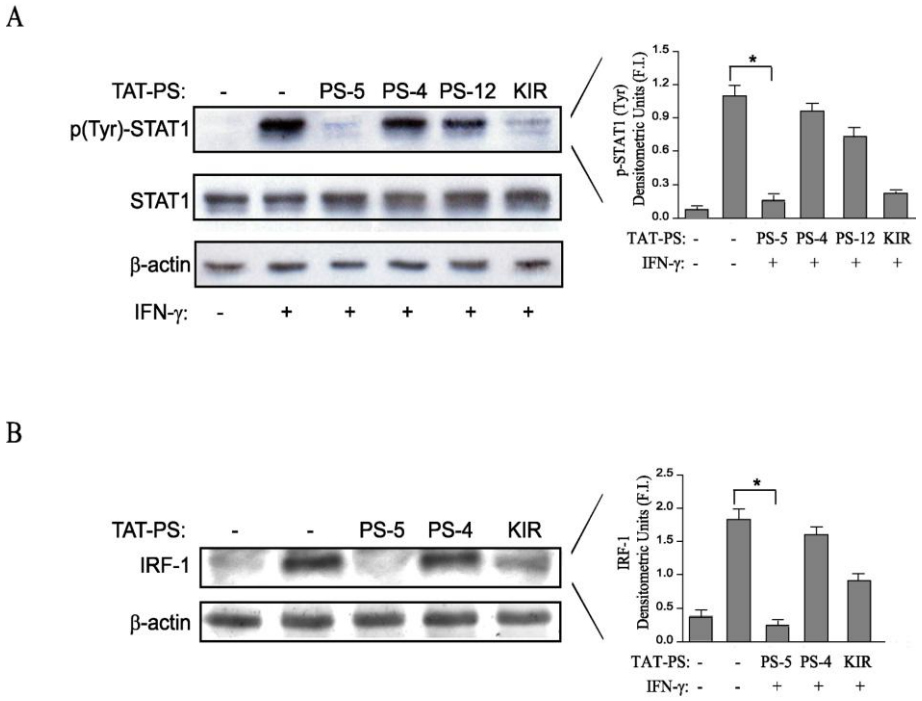


***Fig.* 3.14: Densitometric analyses of Western blotting experiments:** (**A**) Subconfluent keratinocyte cultures were pre-treated for 2 h with 10 μM of PS-4, PS-5, PS-12, KIR peptides, and, then, stimulated for 15 min with IFN-γ. Cells were lysed and the whole-cell extracts were resolved by 10% SDS-gel electrophoresis, transferred to nitrocellulose membranes, and examined using specific Abs to phosphorylated STAT1 (p-Tyr 701). The membrane was stripped and reprobed with Abs specific to unphosphorylated STAT1. (**B**) Similarly, keratinocyte cultures were pre-treated with 10 μM of PS-4, PS-5, KIR peptides, stimulated for 3 hours with IFN-γ, and finally, lysated. IRF-1 expression was detected by Western blotting analysis by using anti-IRF-1 antibody. Equal protein loading was assessed by probing with anti-β-actin. Data are representative of at least three separate experiments. Relative intensity of signals was quantified using a densitometer and reported as densitometric units. In all panels, *p< 0.001.

4. DISCUSSION

IFN-γ is a key cytokine that regulates many cellular functions: antiviral immunity, apoptosis and cell cycle progression. In keratinocytes it is involved in inflammatory and immune responses (9). SOCS proteins play a critical role in modulating the activities of inflammatory cytokines that use tyrosine kinases in their signalling pathways (50). In particular SOCS1 and SOCS-3 regulate JAK2 enzymatic activity by at least two mechanisms: one involves binding to the activation loop of JAK2 and the other facilitates proteasomal degradation (59). Given the primary role of SOCS1 to negatively regulate signalling pathways mediated by cytokine binding, the development of molecules able to mimic SOCSs functions represents a challenging target in the field of drug discovery, both in allergy and other immune-based diseases (60).

Johnson and coworkers demonstrated that SOCS1-like effects can be obtained both employing the KIR domain (encompassing residues 52-67), and by a rationally designed mimetic peptide named TKip that inhibits JAK2 function by binding to its activation loop (54,55).

In this framework, here we present new mimetic peptides of KIR identified through the screening of a focused peptide library (56,57). After the evaluation of the affinity between KIR and JAK2 peptides we carried out an Ala-scanning investigation of KIR in order to evaluate residues crucial for JAK2 recognition. Noteworthy our peptide-based approach was supported by previous mutation studies (37) carried out on the full length and on deletion mutants of SOCS1 (60).

Our results demonstrate that the shorter fragment new KIR, binds to JAK2 peptides more tightly: the deletion of the last six residues of KIR (^{62}SDYRRI^{67}) improves the affinity to JAK2 peptides, suggesting that they are mainly involved in stabilizing the α-helical secondary structure of this domain (39), as also suggested by CD studies and that, once taken out from protein architecture, they negatively interfere with JAK2 recognition. On the basis of the Ala-scan results, we have designed a focused peptide library in which residues directly involved in binding JAK2 remained unchanged, while positions 54-56 were randomized following a simplified approach. From the screening of PS-SSPL we identified new ligands with increased affinity for JAK2 catalytic site with K_D values in the high nanomolar range. It can be observed that these peptides had a residue with a sterically hindered side-chain on position 54 (Cys(Acm) in PS-5 or Phe for PS-11-12) (His in

the w-t sequence); in position 55 natural Phe was conserved in all PS sequences except for PS-5 where there is Arg, in position 56, a Gln residue resulted for all PS binding peptides, confirming the need of a polar side chain in this position (Arg in w-t sequence). In order to speculate on potential conformational differences among PS and natural sequence: we carried out CD studies on these peptides, that confirmed a strong tendency of KIR protein region to adopt helical conformation, which is only slightly affected by C-term truncation and by mutations on positions 54-56 and suggested a role of secondary structure of the SOCS1 domain in JAK2 recognition mechanism. Competitive experiments confirmed PS sequences recognition specificity and that they bind to the same site of JAK2 peptides. Among PS peptides, PS-5 resulted the most interesting and effective and it was employed for preliminary cellular experiments. Indeed it showed the lowest value of dissociation constants and a remarkable difference between the binding to JAK2 and pJAK2: it has a stronger affinity (5-fold greater) for the phosphorylated 1007Y at the JAK2 site. This evidence suggest that wild-type peptides and PS-5 recognize the same site on JAK2 but not in the same way: presumably the presence a positive charge (R55) and a H-bond-prone side chain (Q56) establish more intense electrostatic or polar interactions with the negative phosphate moiety on Y^{1007}. Furthermore PS-5 contains a non-natural residue (Cys(Acm)) that renders this sequence more stable to protease degradation and the most suitable to be tested in experiments involving keratinocytes cells.

In these experiments, PS-5 caused a strong decrease of the phosphorylation levels of STAT1 induced by IFN-γ, and in parallel, of the STAT1-dependent gene IRF-1. On these data, it can be assumed that PS-5 peptide can effectively inhibit pro-inflammatory genes induced by IFN-γ, and, ultimately suppress the IFN-γ-dependent immune responses in keratinocytes, similarly to the findings of Johnson and co-workers that demonstrated that KIR peptide negatively controls type 1 T cell-induced inflammatory diseases in mice, including experimental allergic encephalomyelitis, and, therefore, has a strong therapeutic effectiveness (53).

In this study we demonstrated that PS-5 binds to JAK2 more efficiently than KIR and in keratinocyte reverts the IFN-γ-induced activation of STAT1 and downstream effects. Future studies using the SOCS1 mimetic peptide PS-5 in experimental models of skin inflammation should better define the therapeutic efficacy of this new peptide.

REFERENCES

1. Pastore S., Mascia F. and Girolomoni G. (2006). The contribution of keratinocytes to the pathogenesis of atopic dermatitis. *Eur J Dermatol.;* 16(2):125-31. Review.
2. Owaki T., Asakawa M., Morishima N., Mizoguchi I., Fukai F., Takeda K., Mizuguchi J. and Yoshimoto T. (2007). STAT3 is indispensable to IL-27-mediated cell proliferation but not to IL-27-induced Th1 differentiation and suppression of proinflammatory cytokine production. *J. Immunol.*;1;180(5):2903-1
3. Pasparakis M., Courtois G., Hafner M., Schmidt-Supprian M., Nenci A., Toksoy A., Krampert M., Goebeler M., Gillitzer R., Israel A., Krieg T., Rajewsky K and Haase I. (2002). TNF-mediated inflammatory skin disease in mice with epidermis-specific deletion of IKK2. *Nature;* 20;417(6891):861-6
4. Zenz R., Eferl R., Kenner L., Florin L., Hummerich L., Mehic D., Scheuch H., Angel P., Tschachler E. and Wagner E.F. (2005). Psoriasis-like skin disease and arthritis caused by inducible epidermal deletion of Jun proteins. *Nature;* 15;437(7057):369-75
5. Arakura F., Hida S., Ichikawa E., Yajima C., Nakajima S., Saida T. and Taki S. (2007). Genetic control directed toward spontaneous IFN-alpha/IFN-beta responses and downstream IFN-gamma expression influences the pathogenesis of a murine psoriasis-like skin disease. *J. Immunol*.; 1;179(5):3249-57
6. de Weerd NA, Samarajiwa SA, Hertzog PJ (2007). "Type I interferon receptors: biochemistry and biological functions". J Biol Chem 282 (28): 20053–20057.
7. Liu YJ (2005). "IPC: professional type 1 interferon-producing cells and plasmacytoid dendritic cell precursors". Annu Rev Immunol 23: 275–306.
8. Vilcek, Novel interferons. Nature Immunology, 2003, Volume 4, pages 8-9.
9. Bos, J. D., de Rie, M. A., Teunissen, M. B., and Piskin, G. (2006) Psoriasis: dysregulation of innate immunity. Br. J. Dermatol. 152, 1098–1107.
10. Albanesi, C., Scarponi, C., Giustizieri, M. L., and Girolomoni, G. (2005) Keratinocytes in inflammatory skin diseases. Curr. Drug Targets Inflamm. Allergy 4, 329–334
11. Zhang, X., Chen, X., Song, H., Chen, H. Z., and Rovin, B. H. (2005) Activation of the Nrf2/antioxidant response pathway increases IL-8 expression. Eur. J. Immunol. 35, 3258–3267
12. Platanias, L. C. (2005) Mechanisms of type-I- and type-IIinterferon-mediated signalling. Nat. Rev. Immunol. 5, 375–385.
13. O'Sullivan L.A., Liongue C., Lewis R.S., Stephenson S.E., Ward A.C. (2007). Cytokine receptor signalling through the Jak-Stat-Socs pathway in disease. Mol. Immunol.; 44(10):2497-506
14. Hebenstreit D, Horejs-Hoeck J and Duschl A (2005). "JAK/STAT-dependent gene regulation by cytokines". Drug News Perspect 18 (4): 243–249.
15. Regulation of cytokine signaling pathways by PIAS proteins". Cell Research 16 (2): 196–202.
16. Leonard, W. J., and O'Shea, J. J. (1998). Jaks and STATs: biological implications. Annu Rev Immunol 16, 293-322.

17. Neubauer, H., Cumano, A., Muller, M., Wu, H., Huffstadt, U., and Pfeffer, K. (1998). Jak2 deficiency defines an essential developmental checkpoint in definitive hematopoiesis. Cell 93, 397-409.
18. Saharinen, P., Takaluoma, K., and Silvennoinen, O. (2000). Regulation of the Jak2 tyrosine kinase by its pseudokinase domain. Mol Cell Biol 20, 3387-3395.
19. Rane, S. G., and Reddy, E. P. (2000). Janus kinases: components of multiple signaling pathways. Oncogene 19, 5662-5679.
20. Radtke, S., Haan, S., Jorissen, A., Hermanns, H. M., Diefenbach, S., Smyczek, T., Schmitz-Vandeleur, H., Heinrich, P. C., Behrmann, I., and Haan, C. (2005). The Jak1 SH2 domain does not fulfill a classical SH2 function in Jak/STAT signaling but plays a structural role for receptor interaction and up-regulation of receptor surface expression. J Biol Chem 280, 25760-25768.
21. Girault, J. A., Labesse, G., Mornon, J. P., and Callebaut, I. (1998). Janus kinases and focal adhesion kinases play in the 4.1 band: a superfamily of band 4.1 domains important for cell structure and signal transduction. Mol Med 4, 751-769.
22. Hilkens, C. M., Is'harc, H., Lillemeier, B. F., Strobl, B., Bates, P. A., Behrmann, I., and Kerr, I. M. (2001). A region encompassing the FERM domain of Jak1 is necessary for binding to the cytokine receptor gp130. FEBS Lett 505, 87-91.
23. Mertens, C., Zhong, M., Krishnaraj, R., Zou, W., Chen, X., and Darnell, J. E., Jr. (2006). Dephosphorylation of phosphotyrosine on STAT1 dimers requires extensive spatial reorientation of the monomers facilitated by the N-terminal domain. Genes Dev 20, 3372-3381.
24. Fagerlund, R., Melen, K., Kinnunen, L., and Julkunen, I. (2002). Arginine/lysine-rich nuclear localization signals mediate interactions between dimeric STATs and importin alpha 5. J Biol Chem 277, 30072-30078.
25. Bhattacharya, S., and Schindler, C. (2003). Regulation of Stat3 nuclear export. J Clin Invest 111, 553-559.
26. John, S., Vinkemeier, U., Soldaini, E., Darnell, J. E., Jr., and Leonard, W. J. (1999). The significance of tetramerization in promoter recruitment by Stat5. Mol Cell Biol 19, 1910-1918.
27. Strehlow, I., and Schindler, C. (1998). Amino-terminal signal transducer and activator of transcription (STAT) domains regulate nuclear translocation and STAT deactivation. J Biol Chem 273, 28049-28056.
28. Begitt, A., Meyer, T., van Rossum, M., and Vinkemeier, U. (2000). Nucleocytoplasmic translocation of Stat1 is regulated by a leucine-rich export signal in the coiled-coil domain. Proc Natl Acad Sci U S A 97, 10418-10423.
29. Chen, X., Vinkemeier, U., Zhao, Y., Jeruzalmi, D., Darnell, J. E., Jr., and Kuriyan, J. (1998). Crystal structure of a tyrosine phosphorylated STAT-1 dimer bound to DNA. Cell 93, 827-839.
30. Decker, T., and Kovarik, P. (2000). Serine phosphorylation of STATs. Oncogene 19, 2628-2637.
31. Durbin, J. E., Hackenmiller, R., Simon, M. C., and Levy, D. E. (1996). Targeted disruption of the mouse Stat1 gene results in compromised innate immunity to viral disease. Cell 84, 443-450.

32. Chapgier, A., Boisson-Dupuis, S., Jouanguy, E., Vogt, G., Feinberg, J., Prochnicka-Chalufour, A., Casrouge, A., Yang, K., Soudais, C., Fieschi, C., et al. (2006). Novel STAT1 alleles in otherwise healthy patients with mycobacterial disease. PLoS Genet 2, e131.
33. Sahni, M., Ambrosetti, D. C., Mansukhani, A., Gertner, R., Levy, D., and Basilico, C. (1999). FGF signaling inhibits chondrocyte proliferation and regulates bone development through the STAT-1 pathway. Genes Dev 13, 1361-1366.
34. Shankaran, V., Ikeda, H., Bruce, A. T., White, J. M., Swanson, P. E., Old, L. J., and Schreiber, R. D. (2001). IFNgamma and lymphocytes prevent primary tumour development and shape tumour immunogenicity. Nature 410, 1107-1111.
35. Hilton, D. J., Richardson, R. T., Alexander, W. S., Viney, E. M., Willson, T. A., Sprigg, N. S., Starr, R., Nicholson, S. E., Metcalf, D., and Nicola, N. A. (1998). Twenty proteins containing a C-terminal SOCS box form five structural classes. Proc Natl Acad Sci U S A 95, 114-119.
36. Saito, H., Morita, Y., Fujimoto, M., Narazaki, M., Naka, T., and Kishimoto, T. (2000). IFN regulatory factor-1-mediated transcriptional activation of mouse STAT-induced STAT inhibitor-1 gene promoter by IFN-gamma. J Immunol 164, 5833-5843.
37. Yasukawa, H., Misawa, H., Sakamoto, H., Masuhara, M., Sasaki, A., Wakioka, T., Ohtsuka, S., Imaizumi, T., Matsuda, T., Ihle, J. N., and Yoshimura, A. (1999). The JAK-binding protein JAB inhibits Janus tyrosine kinase activity through binding in the activation loop. Embo J 18, 1309-1320.
38. Nicholson, S. E., Willson, T. A., Farley, A., Starr, R., Zhang, J. G., Baca, M., Alexander, W. S., Metcalf, D., Hilton, D. J., and Nicola, N. A. (1999). Mutational analyses of the SOCS proteins suggest a dual domain requirement but distinct mechanisms for inhibition of LIF and IL-6 signal transduction. Embo J 18, 375-385.
39. Giordanetto, F., and Kroemer, R. T. (2003). A three-dimensional model of Suppressor Of Cytokine Signalling 1 (SOCS-1). Protein Eng 16, 115-124.
40. Sasaki, A., Yasukawa, H., Suzuki, A., Kamizono, S., Syoda, T., Kinjyo, I., Sasaki, M., Johnston, J. A., and Yoshimura, A. (1999). Cytokine-inducible SH2 protein-3 (CIS3/SOCS3) inhibits Janus tyrosine kinase by binding through the N-terminal kinase inhibitory region as well as SH2 domain. Genes Cells 4, 339-351.
41. Qing, Y., Costa-Pereira, A. P., Watling, D., and Stark, G. R. (2005). Role of tyrosine 441 of interferongamma receptor subunit 1 in SOCS-1-mediated attenuation of STAT1 activation. J Biol Chem 280, 1849-1853.
42. Kibel, A., Iliopoulos, O., DeCaprio, J. A., and Kaelin, W. G., Jr. (1995). Binding of the von Hippel-Lindau tumor suppressor protein to Elongin B and C. Science 269, 1444-1446.
43. Kile, B. T., Schulman, B. A., Alexander, W. S., Nicola, N. A., Martin, H. M., and Hilton, D. J. (2002). The SOCS box: a tale of destruction and degradation. Trends Biochem Sci 27, 235-241.
44. Kamura, T., Sato, S., Haque, D., Liu, L., Kaelin, W. G., Jr., Conaway, R. C., and Conaway, J. W. (1998). The Elongin BC complex interacts with the conserved SOCS-box motif present in members of the SOCS, ras, WD-40 repeat, and ankyrin repeat families. Genes Dev 12, 3872-3881.

45. Callus, B. A., and Mathey-Prevot, B. (1998). Interleukin-3-induced activation of the JAK/STAT pathway is prolonged by proteasome inhibitors. Blood 91, 3182-3192.
46. Boyle, K., Egan, P., Rakar, S., Willson, T. A., Wicks, I. P., Metcalf, D., Hilton, D. J., Nicola, N. A., Alexander, W. S., Roberts, A. W., and Robb, L. (2007). The SOCS box of suppressor of cytokine signaling-3 contributes to the control of G-CSF responsiveness in vivo. Blood.
47. Noon-Song, E. N.; Ahmed, C. M.; Dabelic, R.; Canton, J.; Johnson, H. M. (2011) Controlling nuclear JAKs and STATs for specific gene activation by IFNgamma. Biochem Biophys Res Commun. 410, 648
48. Solomon, M.; Flodstrom-Tullberg, M. and Sarvetnick, N. (2011) Beta-cell specific expression of suppressor of cytokine signaling-1 (SOCS-1) delays islet allograft rejection by down-regulating Interferon Regulatory Factor-1 (IRF-1) signaling. Transpl. Immunol. 24, 181-188
49. Madonna, S.; Scarponi, C.; Sestito, R.; Pallotta, S.; Cavani, A. and Albanesi, C.. (2011) The IFN-gamma-dependent suppressor of cytokine signaling 1 promoter activity is positively regulated by IFN regulatory factor-1 and Sp1 but repressed by growth factor independence-1b and Kruppel-like factor-4, and it is dysregulated in psoriatic keratinocytes. J. Immunol. 185, 2467-2481.
50. Madonna, S.; Scarponi, C.; De Pita, O. and Albanesi, C. (2008) Suppressor of cytokine signaling 1 inhibits IFN-gamma inflammatory signaling in human keratinocytes by sustaining ERK1/2 activation. FASEB J. 22, 3287-3297
51. Shouda, T., Yoshida, T., Hanada, T., Wakioka, T., Oishi, M., Miyoshi, K., Komiya, S., Kosai, K., Hanakawa, Y., Hashimoto, K., et al. (2001). Induction of the cytokine signal regulator SOCS3/CIS3 as a therapeutic strategy for treating inflammatory arthritis. J Clin Invest 108, 1781-1788.
52. Jo, D., Liu, D., Yao, S., Collins, R. D., and Hawiger, J. (2005). Intracellular protein therapy with SOCS3 inhibits inflammation and apoptosis. Nat Med 11, 892-898.
53. Mujtaba, M. G., Flowers, L. O., Patel, C. B., Patel, R. A., Haider, M. I., and Johnson, H. M. (2005). Treatment of mice with the suppressor of cytokine signaling-1 mimetic peptide, tyrosine kinase inhibitor peptide, prevents development of the acute form of experimental allergic encephalomyelitis and induces stable remission in the chronic relapsing/remitting form. J Immunol 175, 5077-5086.
54. Flowers, L. O., Subramaniam, P. S., and Johnson, H. M. (2005). A SOCS-1 peptide mimetic inhibits both constitutive and IL-6 induced activation of STAT3 in prostate cancer cells. Oncogene 24, 2114-2120.
55. Waiboci, L. W.; Ahmed, C. M.; Mujtaba, M. G.; Flowers, L. O.; Martin, J. P. Haider, M. I. Johnson, H. M.. (2007) Both the suppressor of cytokine signaling 1 (SOCS-1) kinase inhibitory region and SOCS-1 mimetic bind to JAK2 autophosphorylation site: implications for the development of a SOCS-1 antagonist. J. Immunol. 178, 5058-5068.
56. Marasco, D.; Perretta, G.; Sabatella, M. and Ruvo, M. (2008) Past and future perspectives of synthetic peptide libraries. Curr. Protein Pept. Sci. 9, 447-467.

57. Scognamiglio, P. L.; Doti, N.; Grieco, P.; Pedone, C.; Ruvo, M. and Marasco, D. (2011) Discovery of Small Peptide Antagonists of PED/PEA15-D4alpha Interaction from Simplified Combinatorial Libraries. Chem. Biol. Drug Des. 77, 319-327.
58. Elmquist, A.; Lindgren, M.; Bartfai, T. and Langel, U. VE-cadherin-derived cell-penetrating peptide, pVEC, with carrier functions. Exp Cell Res 2001, 269, 237-244.
59. Dalpke, A.; Heeg, K.; Bartz, H. and Baetz, A. (2008) Regulation of innate immunity by suppressor of cytokine signaling (SOCS) proteins. Immunobiology 213, 225-235.
60. Kubo, M.; Hanada, T. and Yoshimura, A. (2003) Suppressors of cytokine signaling and immunity. Nat. Immunol. 4, 1169-1176.

CHAPTER IV

1.INTRODUCTION

1.1. Ischemia

Ischemia is a restriction in blood supply, generally due to factors in the blood vessels, with resultant damage or dysfunction of tissue. Rather than hypoxia (a more general term denoting a shortage of oxygen, usually a result of lack of oxygen in the air being breathed), ischemia is an absolute or relative shortage of the blood supply to an organ, i.e. a shortage of oxygen, glucose and other blood-borne fuels. A relative shortage means the mismatch of blood supply (oxygen/fuel delivery) and blood request for adequate metabolism of tissue. Ischemia results in tissue damage because of a lack of oxygen and nutrients (1).

In very aerobic tissues such as heart and brain, at body temperature necrosis due to ischemia usually takes about 3–4 minutes before becoming irreversible. Ischemia is a feature of heart diseases, transient ischemic attacks, cerebrovascular accidents, ruptured sensitive to inadequate blood supply. Ischemia in brain tissue, for example due to stroke or head injury, causes a process called the ischemic cascade to be unleashed, in which proteolytic enzymes, reactive oxygen species, and other harmful chemicals damage and may ultimately kill brain tissue.

Restoration of blood flow after a period of ischemia can actually be more damaging than the ischemia. Reintroduction of oxygen causes a greater production of damaging free radicals as well as allowing, via removal of the extracellular acidotic conditions, influx of calcium and thus calcium overloading.

1.2.Neuronal Cell Death

Cell death involving a cell-autonomous active contribution of catabolic enzymes, apoptosis, plays a prominent role in the evolution of hypoxic-ischemic (HI) injury in the neonatal brain, and apoptosis is at least as important for the loss of neurons as unregulated cell death, necrosis (2,3). Caspases are a class of specific cysteine proteases that mediate apoptotic death in a variety of cellular systems. Caspases are activated after HI, in particular in the immature brain (3). The molecular pathway leading to caspase activation is initiated by mitochondrial outer membrane permeabilization (MOMP), presumably as a result of the permeability transition (4) and/or a process regulated by proteins from the Bcl-2 family (5). MOMP results in the release of cytochrome *c* from the mitochondrial intermembrane space to the

cytosol, where cytochrome *c* can interact with Apaf-1, triggering the formation of the apoptosome, the caspase-9 and caspase-3 activation complex. Prevention of caspase activation or inhibition of caspases affords neuroprotective effects in the immature brain (6). Recent data demonstrate that when caspase activation is inhibited at or downstream of the apoptosome, neurons can undergo a delayed, caspase-independent death (7). One of the key components of the caspase independent cell death pathway is Apoptosis-Inducing Factor (AIF) (8–10) (*Fig*. 1.1).

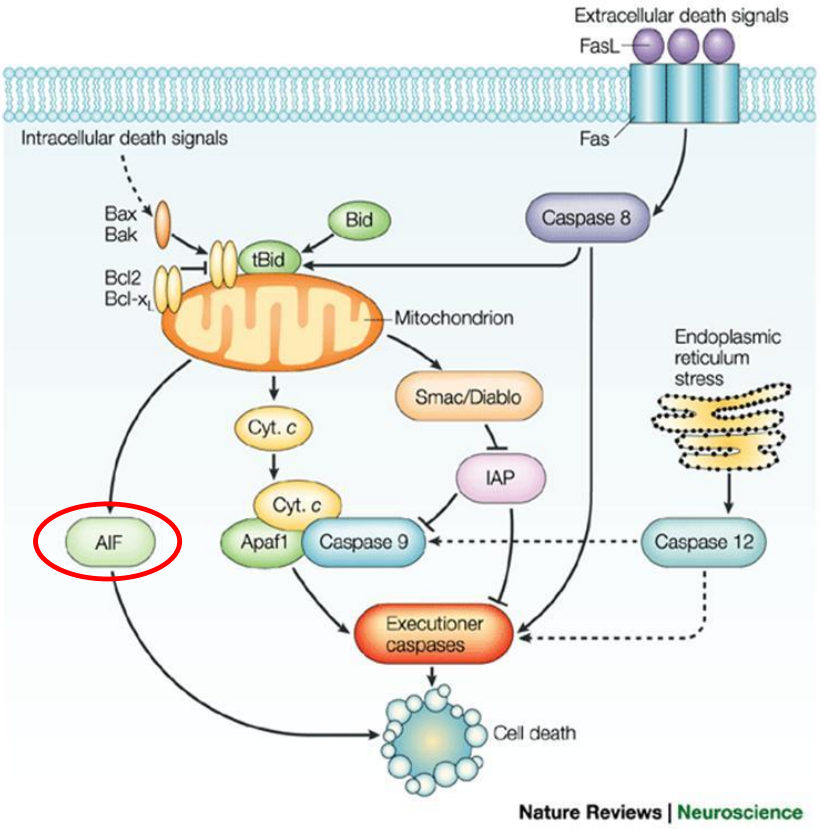


***Fig*. 1.1: Schematic representation of apoptotic pathways mediated by mitochondrial proteins.**

1.3. Apoptosis-inducing factor (AIF) protein: structure and function

Mature apoptosis-inducing factor (AIF) is a 57 kDa flavoprotein that is confined to mitochondria. The AIF protein includes three domains: an amino-terminal

mitochondrial localization sequence (MLS) of 110 amino acids; a spacer sequence of 27 amino acids; and a carboxiterminal 485 amino acids domain with strong homology to oxidoreductases from other species. The overall structure of AIF displays a glutathione reductase (GR)-like fold and includes one FAD molecule per monomer. Similarly to the enzymes of the GR family, AIF is composed of three domains: an FAD-binding domain (residues 122−262 and 400−477), an NADH-binding domain (263−399) and a C-terminal domain (478−610), which in GR constitutes most of its dimer interface. Both the FAD- and NADH-binding domains display the classical Rossmann fold, whereas the C-terminal domain is composed of five antiparallel β-strands (residues 477−579) followed by two α-helices (residues 580−610) (11) (*Fig.* 1.2).

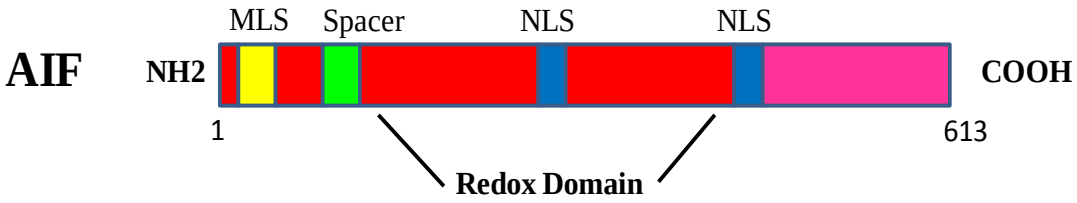


***Fig.* 1.2: Primary structure of apoptosis-inducing factor (AIF).** AIF has an oxidoreductase domain, mitochondrial localization sequence (MLS) and nuclear localization sequence (NLS).

AIF has a dual function in the control of stress-induced cell death (12). This protein is a phylogenetically old mitochondrial NADH oxidase whose local redox function is essential for optimal oxidative phosphorylation and for an efficient anti-oxidant defense. The absence of AIF can cause degeneration of cerebellar granule neurons and retinal ganglion cells, as has been shown for harlequin (Hq) mice (13). Nonetheless, AIF has also been implicated in neuronal apoptosis as a cell death effector. In several models of apoptosis, AIF translocates to the nucleus, where it induces chromatin condensation and large-scale DNA degradation by an unknown molecular mechanism. The nuclear translocation of AIF can be inhibited by blocking upstream signals of apoptosis (such as poly(ADPribose) polymerase-1) (14), inhibition of non-caspase cysteine proteases that are involved in the detachment of AIF from the inner mitochondrial membrane (15), suppression of MOMP (16) or

overexpression of heat-shock protein 70 (Hsp70), which can intercept AIF in the cytosol (17).

This lethal action seems to be independent from its oxidoreductase activity (18). Thus, as seen in cytochrome *c*, AIF may behave as a bifunctional protein with dissociable apoptogenic and redox properties. However, the nuclear events caused by AIF apparently depend on the apoptosis inducer and the cell type, and are reversible at low apoptotic insult (19).

1.4. *AIF-Mediated Cell Death*

Upon apoptosis induction, AIF translocates from mitochondria to the cytosol and to the nucleus. During early nuclear condensation, AIF is found to be associated with chromatin (20). AIF protein binds to DNA and causes purified nuclei to undergo chromatin condensation, DNA degradation to ~50 kbp fragments and DNA loss (8). Although the mitochondrial release of AIF is controlled by caspases in some models (21), it is caspase-independent in several important paradigms of apoptosis induction. Thus, DNA damage can induce the caspase-independent AIF release either through p53 (22), or via PARP-mediated NAD depletion (23). AIF is also released in several in vivo models of anoikis (24), HIV-1 infection (25), and ischemia/reperfusion (2), even in conditions in which caspases are suppressed. Additional caspase-independent enzymes that have been suggested to participate in apoptotic chromatinolysis include endonuclease G (26), DNAse I (27), and cyclophilins (28). Apparently, the apoptogenic activity of AIF does not depend on its NADH oxidase activity. Chemical inactivation of the flavine adenine nucleotide (FAD) moiety required for AIF redox activitydoes not block its apotogenic function in cell free systems (7), and mutations that destroy the FAD binding site do not affect the apoptogenic function of AIF in transfection assays. Rather, mutation of amino-acid residues required for DNA binding (but not for redox activity) abrogate the apoptogenic potential of AIF in the cell-free system as well as in intact cells (20). In C. elegans, AIF reportedly cooperates with endonuclease G, a DNAse that also undergoes a mitochondrio-nuclear translocation process (26). However, no such evidence has been obtained in the mammalian system, where endonuclease G appears to cooperate with DNAse I (27).To explore the mechanism through which AIF induces DNA degradation, Cande´ et al (29) reported that AIF can cooperate

with cyclophilin A (CypA) to induce chromatinolysis both in vitro and in vivo. In that studies they show that, in the nucleus, AIF undergoes electrostatic interactions with DNA and tethers CypA to chromatin. Formation of such a trimolecular complex (CypA/AIF/DNA) is possible because the DNA- and CypA-binding domains of AIF are distinct. The AIF/CypA complex then causes DNA degradation.

1.5. Cyclophilin A

Cyclophilins constitute a family of phylogenetically conserved proteins found in prokaryotes as well as in humans. Cyclophilins have peptidyl-prolyl isomerase activity in vitro (30), indicating that they influence the conformation of proteins in cells, but their functions are largely unknown (31). The founding member of the cyclophilin family (15 members in humans), cyclophilin A (CypA), is an abundant, ubiquitously expressed protein originally discovered as an intracellular ligand of the immunosuppressive drug cyclosporin A (CsA) (32). CypA $^{-/-}$ embryonic stem cells grow normally and differentiate into hematopoietic precursor cells in vitro, indicating that CypA is not essential for mammalian cell viability (33). Immunosuppression is mediated by the CypA – CsA complex that inhibits calcineurin, a phosphatase crucial for T cell activation (34). CypA expression is particularly high in the brain, in neurons, where it is expressed mainly in the cytoplasm but also in the nucleus (35). Pioneering work by Montague et al. revealed that CypA possesses a latent, apoptosis related DNase activity (28). CypA has been demonstrated to participate in excitotoxin-induced apoptosis (36) and to interact with a limited number of proteins, including dynein (37), the antioxidant protein Aop1 (38), the nuclear pore protein Nup358 (39), and apoptosis-inducing factor (AIF) (29).

1.6. AIF/Cyp A interaction

Several studies have shown that CypA interacts with AIF by mass spectroscopy, coimmunoprecipitation, pull-down assays, and molecular modeling. During the initial, caspase-independent stage of chromatin condensation that accompanies apoptosis, AIF and CypA were found to co-immunolocalize in the nucleus. Recombinant AIF and CypA proteins synergized in vitro in the degradation of

plasmid DNA, as well as in the capacity to induce DNA loss in purified nuclei. The apoptogenic cooperation between AIF and CypA did not rely on the CypA peptidyl-prolyl cis–trans isomerase activity. In Cyp-expressing cells, AIF overexpression augmented apoptotic chromatinolysis. The AIF-dependent large-scale DNA fragmentation was less pronounced in CypA knockout cells as compared to controls. AIF mutants lacking the CypA-binding domain were inefficient apoptosis sensitizers in transfection experiments. Moreover, AIF failed to sensitize CypA knockout cells to apoptosis induction, and this defect in the AIF response was reversed by reintroduction of the CypA gene into CypA-deficient cells. In summary, AIF and CypA collaborate in chromatinolysis.

In particular, deletion of amino acids (aa) 367–466, 269–442, or 263–399 abolished the interaction between AIF and CypA. Thus, the CypA interaction domain is located within (or close to) a stretch of ~30 amino acids (367–399) of AIF, which is different from its HSP70 interacting domain (aa 150–228) (17) and from its DNA-binding sites (aa 255/265 and aa 510/518) (20). Accordingly, a molar excess of Hsp70 or DNA failed to block the AIF–CypA interaction. Addition of cyclosporin A, which inhibits the peptidyl-prolyl cis–trans isomerase activityof CypA had no effect on the AIF/CypA interaction, which thus occurs independently of its chaperone function.

The notion that AIF interacts with CypA through a novel interaction domain was confirmed by molecular modeling (29). An optimal model of docking between CypA and AIF indicates that CypA and AIF establish extensive molecular contacts. AIF contributes to the interface with one of its a-helices (P345–R358) and several β-strands (V361–N366, T328–F334, R387–A397) and turns (E359–G360, A367–S371, K382–G386) that form together a β-sheet bulge. CypA contributes to the interface with one of its two a-helices (M136–G146), the following turn and β-strand (S147–G150), as well as part of its central β-barrel (R55–Q63, H92–A101, Q111–T119). (*Fig*. 1.3). The predicted interface is stabilized by several molecular interactions.

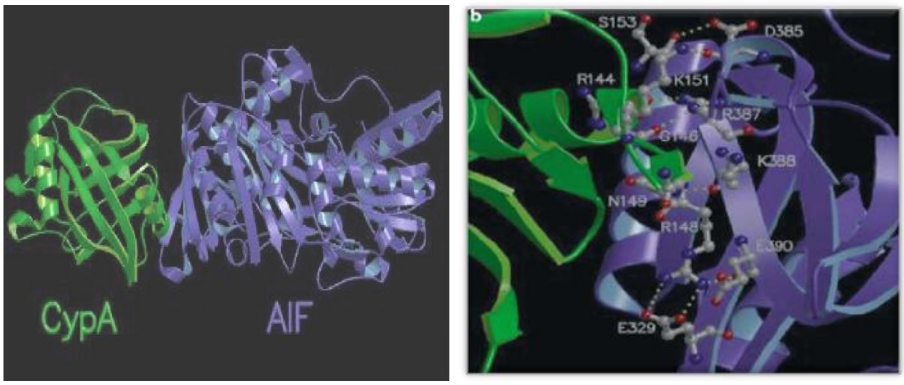


***Fig.* 1.3: Molecular model of the complex between AIF and CypA.** (**left**) Global view of the interaction between AIF and CypA, obtained by homology modeling on the known crystal structures and energy minimization calculations. (**right**) Close-up of the interface of interaction between AIF (blue) and CypA (green).

1.7. Aim of the study

Many data indicate that the lethal translocation of AIF to the nucleus requires interaction with CypA, suggesting a model in which two proteins that normally reside in separate cytoplasmic compartments acquire novel properties when moving together to the nucleus. In spite of the considerable amount of works performed on AIF, the mechanisms underlying the nuclear action of AIF are, however, largely unknown. The purpose of this study is the elucidation of the mechanism underlying the neuronal cell death mediated by AIF and CypA proteins by the selective disruption of the CypA-AIF complex with peptides mimicking the protein surfaces involved in mutual binding. Importantly, the selective interference in the formation of this target complex should be an ideal strategy to suppress the pro-apoptotic action of AIF without interfering with its implication in bioenergetic and redox metabolism and in the immunosuppressive features of CypA.

Starting from structural data, peptides that mimic the mutual binding region of AIF and CypA have been designed and synthesized. Peptides synthesized have been tested by using cell-free and cell based assays.

In particular for the screening of peptide direct binding and competition experiments have been developed by Surface Plasmon Resonance (SPR) techniques. For cell based assays we used a model of glutamate toxicity in HT-22 cells. In these immortalized hippocampal neurons, glutamate induces glutathione depletion and increased formation of reactive oxygen species (ROS) (40). Glutamate toxicity resulted in mitochondrial fragmentation and release of AIF from mitochondria to

nuclei where it induced chromatinolisys. This model is very useful to investigate the effects of the peptide identified on AIF-mediated neuronal cell death.

2. EXPERIMENTAL SECTION

2.1. Materials

N-α-protected Fmoc-amino acid derivatives and coupling reagents for peptide synthesis were from Inbios (Pozzuoli, Napoli). DIEA, Rink amide MBHA resin, tri-isopropyl-silane (TIS) and reagents for ELISA assays were from Sigma-Aldrich (Milano, Italy). HPLC-grade solvents and trifluoroacetic acid (TFA) were from LabScan (Stillorgan, Dublin, Ireland) and were used as received unless otherwise stated. Other reagents and chemical suppliers are indicated in the section of methods. Solid phase peptide syntheses was performed on a fully automated multichannel peptide synthesizer Syro I (Multisynthech, Germany). Preparative RP-HPLC were carried out on a Shimadzu LC-8A, equipped with a SPD-M10 AV detector and with a Phenomenex C18 Jupiter column (50x22 mm ID; 10 μm). LC-MS analyses were carried out on an LCQ DECA XP Ion Trap mass spectrometer equipped with an OPTON ESI source, operating at 4.2 kV needle voltage and 320 °C and with a complete Surveyor HPLC system, comprising a MS pump, an autosampler and a photo diode array (PDA). Narrow bore 50x2 mm C18 BioBasic LC-MS columns were used for analyses. All ELISA assays to screen the synthetic peptides, were carried out using a fully automated system (Hamilton Robotics, Milano, Italy) comprising a liquid handler, an automatic arm, a washer and an automated Synergy 4 multi-wavelength reader (BIOTEK Instruments, Inc. Highland Park, VT, USA).

2.2. AIF expression and purification

The expression of recombinant truncated AIF (lacking the amino acids 1-121, tAIF) fused to a His tag was optimised in the BL21(DE3) bacterial strain and finally induced in the presence of 0.1 mM IPTG for 3 h at 37 °C. 200 mL pellets were re-suspended in 10 mL cold lysis buffer (50 mM Tris, 150 mM NaCl, Triton 0.25%, 1 mM TCEP, pH 8) supplemented with a protease inhibitor mixture (1 mM PMSF, 1.0 μg mL^{-1} aprotinin, 1.0 μg mL^{-1} leupeptin, 1.0 μg mL^{-1} pepstatin, 1.0 mg mL^{-1} lysozime) and incubated at room temperature for 30 min. Cells were disrupted by sonication on ice with 20 s on/off cycles for a total sonication time of 12 min. After centrifugation at 15000 rpm for 20 min at 4°C, supernatants were loaded onto a 1

mL Ni-NTA-resin equilibrated with 50 mM Tris, 250 mM NaCl, 1 mM TCEP, 0.1 mM PMSF, pH 8.0 (buffer A) and washed extensively. The protein was then eluted with 200 mM imidazole dissolved in the same buffer. Then the product was further purified by size exclusion chromatography (gel filtration s200). Protein elution was monitored by 10% SDS-PAGE electrophoresis (*Fig.* 2.1).

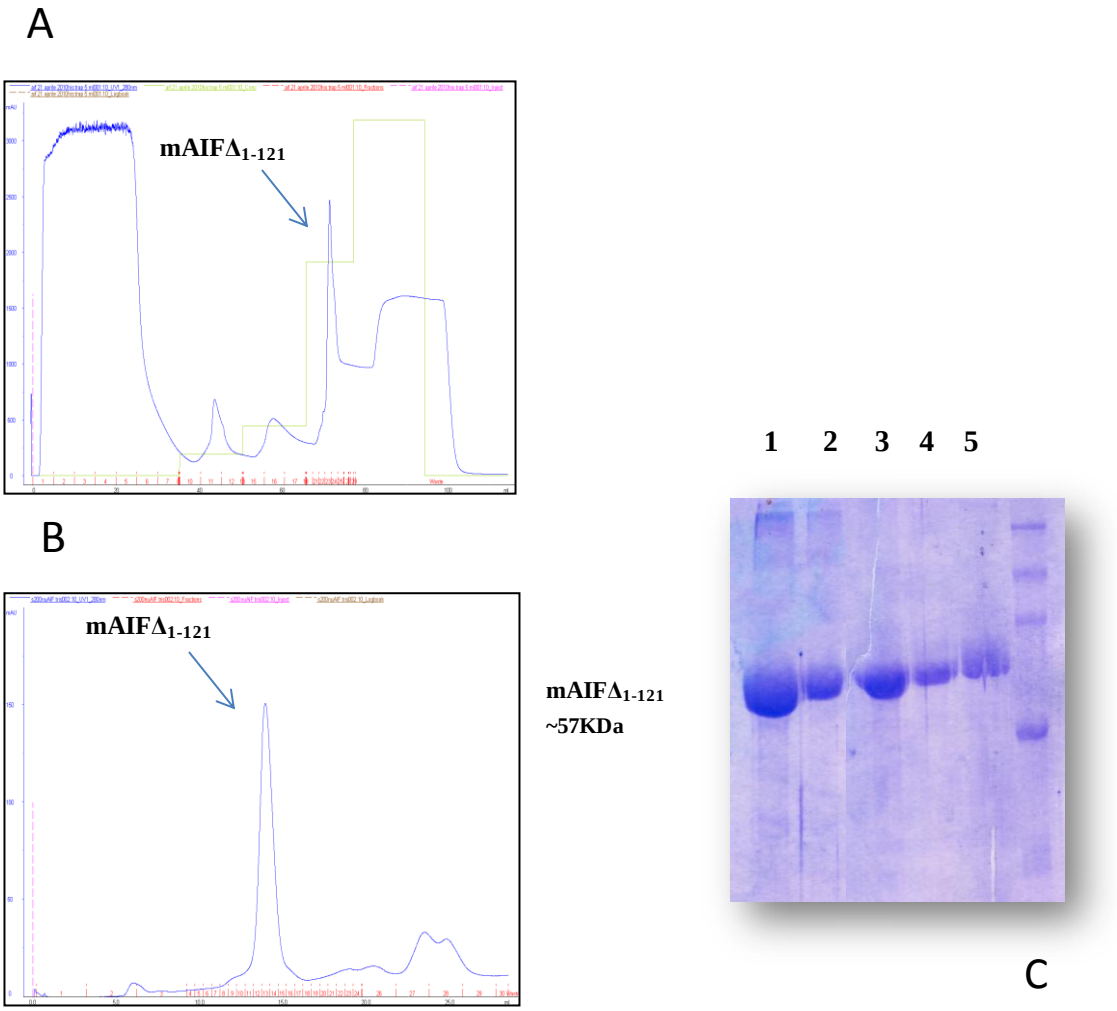


***Fig.* 2.1: Purification of tAIF.** (**A**) First step of purification of tAIF by affinity chromatography on Ni-NTA resin. (**B**) The material corresponding to peak showed in figure 1A was further purified by Size Exclusion Chromatography and analyzed on SDS–PAGE gel (**C**).

2.3. Peptide Synthesis

Peptides were prepared by the solid phase method on a 50 µmol scale following the Fmoc strategy and using standard Fmoc-derivatized amino acids RINK AMIDE

resin (substitution 0.5 mmol/g) was used as solid support. Activation of amino acids was achieved using HBTU/HOBt/DIEA (1:1:2), whereas Fmoc deprotection was carried out using a 40% (v/v) piperidine solution in DMF. All couplings were performed for 15 minutes and deprotections for 10 minutes. Peptides were removed from the resin by treatment with a TFA:TIS:H_2O (90:5:5, v/v/v) mixture, then they were precipitated in cold diethylether and lyophilized. Single peptides were purified by preparative RP-HPLC, and their identity was evaluated by LC-MS analysis. Peptide products were stored at -20 °C until use. The MW was determined by LC-ESI mass spectrometry. To quantify the percentuage of transfection in neuronal cells, peptides were also fluoresceinated at their N-termini with Fluoresceine-βAla.

2.4. SPR analysis

The BIAcore 3000 SPR system for Real time binding assay and related reagents were from GE Healthcare (Milano, Italy).

Human CypA protein (purchased from SIGMA C-3805) was immobilized at a concentration of 50 µg/mL in 10 mM acetate buffer pH 4.0 (flow rate 5 µL/min, time injection 7 min) on a CM5 Biacore sensor chip, using EDC/NHS chemistry following the manufacturer's instructions (41). Residual reactive groups were deactivated by treatment with 1 M ethanolamine hydrochloride, pH 8.5. Reference channel was prepared by activating with EDC/NHS and deactivating with ethanolamine. Binding assays were carried out at 20 µL/min, with 4,5 min contact-time; tAIF protein and peptides were diluted in the running-buffer, HBS (10 mM Hepes, 150 mM NaCl, 3 mM EDTA, pH 7.4) 0.1 mM TCEP. Analyte injections of 90 µL were performed at the indicated concentrations. BIAevaluation analysis package (version 4.1, GE Healthcare, Milano, Italy) implemented by instrument software was used to subtract the signal of the reference channel and to evaluate kinetic and thermodynamic parameters of complexes. Competitive SPR experiments with AIF peptides were carried out pre-incubating each peptides with tAIF at several concentrations.

2.5. Circular Dichroism (CD) Spectroscopy

CD spectra were recorded on a Jasco J-810 spectropolarimeter (JASCO Corp, Milan, Italy). CD spectra were registered at 25 °C in the far UV region from 190 to

260 nm. Each spectrum was obtained averaging three scans, subtracting contributions from corresponding blanks and converting the signal to mean residue ellipticity in units of deg*cm^{2}*dmol^{-1}* res^{-1}. Other experimental settings were: 20 nm/min scan speed, 2.0 nm band width, 0.2 nm resolution, 50 mdeg sensitivity, and 4 sec response. The concentration of peptides was kept at 10×10^{-5} M and a 0.1 cm path-length quartz cuvette was used. Spectra were acquired in 10 mM phosphate buffer at pH 7.0.

2.6. HT-22 cells

HT-22 hippocampal-derived cells were cultured in Dulbecco's modified Eagle containing-calcium medium (DMEM, Invitrogen, Karlsruhe, Germany) or in DMEM without calcium. Both culture media were supplemented with 10% fetal calf serum, 100U/mL penicillin, 100µg/mL streptomycin and 2mM glutamine.

2.7. Transfection

Cellular loading of peptides was performed with cationic lipid mixture Pro-Ject™ Protein Transfection Reagent kit according to the manufacturer's instructions (Pierce Prod# 89850). Briefly, HT-22 cells were transfected with thre different concentrations of peptides (10, 25 and 50 µM) in 24 well plates (60000 cells/well) for 24 h. Afterwards, cells were seeded in 96 well plates and after another 24 h HT-22 cells were treated with glutamate (2 mM). The quantification of the transfection was carried out by Leica microscope The evaluation of the impact of these peptides was carried out by MTT assays and by xCELLigence technique.

2.8. Cell viability assessment

Quantification of cell viability was performed by 3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) reduction assay at 0.5 mg/mL for 1 h. The reaction was terminated by removing the media and freezing the plate at -80° C for at least 1 h. DMSO solvent was added to each well for 1h under shaking conditions at 37° C. The absorbance of each well was determined with an automated FLUOstar Optima reader (BMG Labtechnologies GmbH, Inc., Offenburg, Germany) at 570 nm with a reference filter at 630 nm. In addition, real time detection of cell death was performed by measurements of cellular impedance by the xCELLigence

system (Roche, Penzberg, Germany).

2.9. xCELLigence impedance-based system

HT-22 cells were seeded at a density of 8.000-10.000 cells/well in 96-well E-plate (Roche Diagnostics GmbH). The impedance, depicted as cell index (CI) was used to monitor the real time kinetics of cellular growth and alteration of cell morphology. Fourthy-eight hours after transfection, the cells were treated with glutamate 2 mM.

2.10. AnnexineV-PI assay

For detection of apoptotic cells, HT-22 were loaded with 2µM AnnexineV-FITC-PI (Invitrogen). Flow cytometry was performed using a FACScan (BD Bioscience, Heidelberg, Germany). 488nm UV line argon laser was used for excitation and BODIPY emission was recorded on channels FL1 at 530nm (green) and FL2 at 585nm (red). Data were collected from at least 10,000 cells (n=3).

2.11.Statistical analysis

All data are given as means ± S.D. For statistical comparisons between two groups, Student's *t*-test was used; multiple comparisons were performed by ANOVA followed by Scheffé's *post hoc* test. Calculations were made with the Winstat standard statistical software package (Robert Fitch Software, Bad Krozingen, Germany). A statistically significant difference was assumed at $*p < 0.05$, $**p < 0.01$, $***p < 0{,}001$.

3.RESULTS

3.1. SPR experiment: analysis of AIF protein binding to CypA

Since this interaction has not been assessed employing any experiment in vitro, we analyzed it by a technique that employs unlabelled bio-molecules. For this purpose, Recombinant hCypA protein was immobilized on the chip surfaces, achieving, under the reported conditions (see Experimental Section for details), 800 RU immobilization levels. Direct binding between hCypA and recombinant truncated AIF(Δ1-121), the nuclear apoptogenic form of AIF, were performed by injecting AIF(Δ1-121) (tAIF) solutions at increasing concentrations from 10 to 110 nM . In particular, concentrations of tAIF were injected over the CypA-derivatized CM5 sensor chip in 50 mM sodium phosphate, 150 mM NaCl, 1 mM TCEP, pH 7.4, buffer at 25 °C. This signal was reproducibly proportional to the concentration of tAIF that was injected. A specific SPR resonance signal was detected in the flow-cell, indicating that tAIF was bound to CypA and the concentration-dependence of the steady-state SPR signals showed a hCypA/tAIF dissociation constant (KD) of 3±3 10^{-8} M (*Fig.* 3.1). These results show that the hCypA/tAIF interaction is direct and does not require additional proteins.

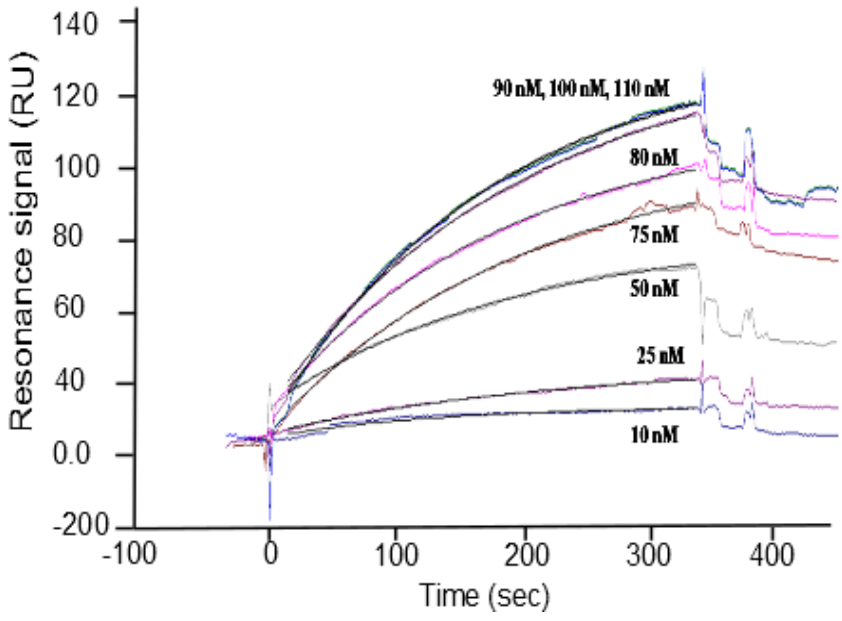


***Fig.* 3.1: Evaluation of interaction between tAIF and CypA.** SPR analysis of the interaction between tAIF and CypA. Results were reported as (RU/RU0)*100.

3.2. Design and characterization of AIF peptides

The notion that AIF interacts with CypA through a novel interaction domain was confirmed by mass spectroscopy, coimmunoprecipitation, pull-down assays, and molecular modeling (29). These previous results showed that, in CypA-expressing cells, AIF overexpression augmented apoptotic chromatinolysis. The AIF-dependent large-scale DNA fragmentation was less pronounced in CypA knockout cells as compared to controls. AIF mutants lacking the CypA-binding domain were inefficient apoptosis sensitizers in transfection experiments. Moreover, AIF failed to sensitize CypA knockout cells to apoptosis induction, and this defect in the AIF response was reversed by reintroduction of the CypA gene into CypA-deficient cells (29).

An optimal model of docking between CypA and AIF indicates that CypA and AIF establish few molecular contacts. These data showed that the formation of CypA/AIF complex may involve the AIF 370-394 fragment and one of the helices (aa 343-360). On the base of these data two potential blocking peptides AIF(370-394), AIF(343-360) and a set of a shorter peptides AIF(370-376), AIF (376-383) and AIF(383-394) were designed, synthesized and characterized by Circular dicroisme.

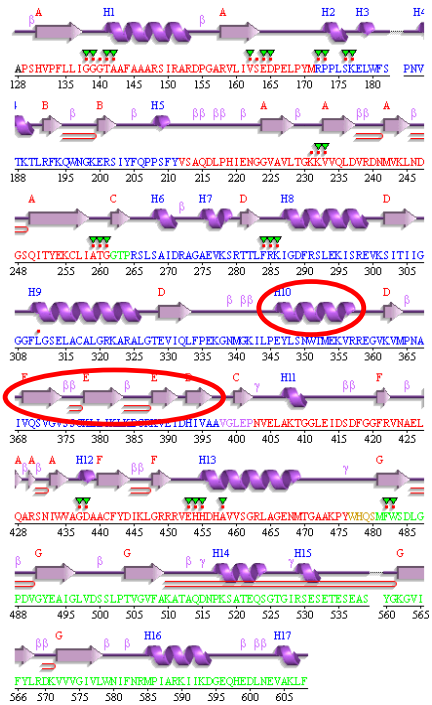


***Fig*. 3.2: Schematic representation of secondary structural regions of AIF,** where the potential regions involved in the interaction are circled in red.

Description	Sequence	Structural propensity
α-helix - AIF 343-360	ILPEYLSNWTMEKVRREG	α-helix
β-sheet-turns AIF 370-394	QSVGVSSGKLLIKLKDGRKVETDHI	α-helix
β-strand AIF 370-376	QSVGVSS	none
β-strand AIF 376-383	SGKLLIKL	none
β-strand AIF 383-394	KDGRKVETDHI	none

Table 3.1: Peptide sequences utilized in this study and their structure secondary propensity deriving from CD studies

All peptides were chemically synthesized in good yields by SPPS, using Fmoc methodologies as N-terminal acetylated and C-terminal amidated derivatives and purified by RP-HPLC; their identity and purity were assessed by LC-MS (data not shown). To evaluate the impact of intrinsic structural preferences of these fragments on their ability to binding, all peptides were characterized by far-UV circular dichroism spectroscopy (CD). As expected, the peptides corresponding to the shorter regions (AIF(370-376), AIF(376-383) and AIF(383-394)) were unfolded, although these peptides contain fragments that are structured in the complex. (*Fig.* 3.3). Similarly, the CD spectra of AIF(343-360) present a structural propensity to assume an α-helix secondary structure, in according with the secondary structure present in the native protein. The CD experiment carried out on AIF(370-394) peptide shows a propensity to assume a β-sheet structure.

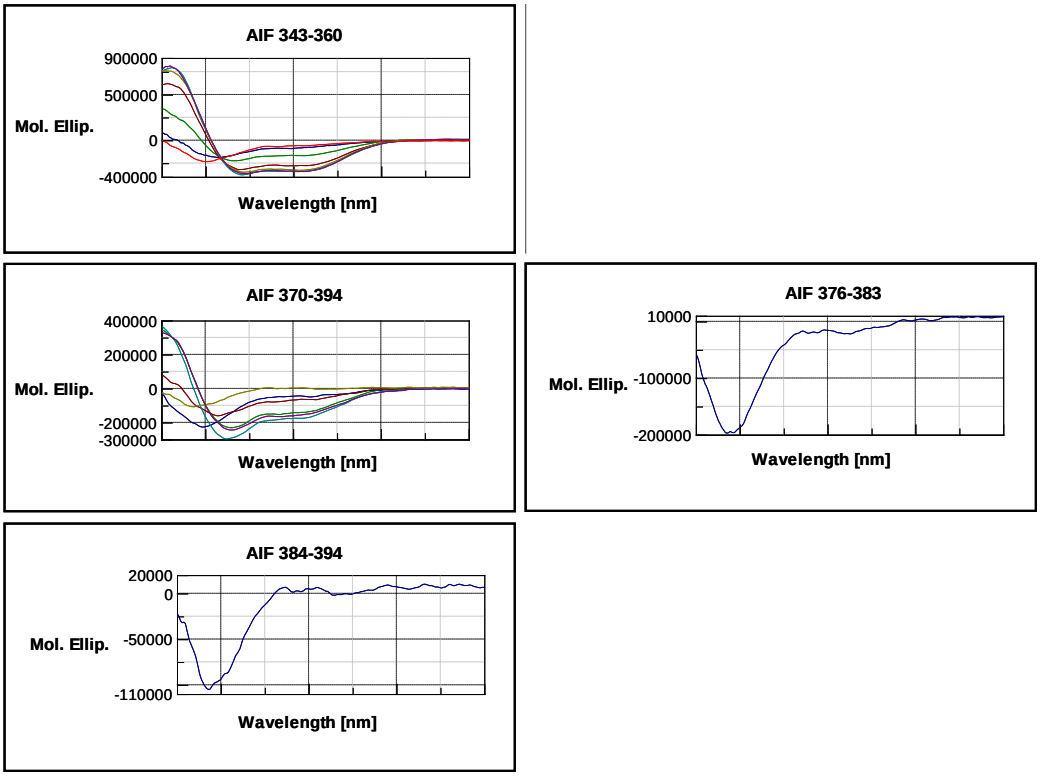


***Fig*. 3.3: Overlay of CD spectra of the AIF peptides in phosphate buffer 10 mM with increasing amounts of TFE, from 0 to 70 % (v/v).**

3.3. Analysis of AIF peptides binding to CypA

In order to identify the AIF amino acidic region involved in the interaction with CypA, we set up SPR-based binding assays in which the peptide singly were injected on the recombinant CypA-captured chip at different concentrations. As shown in *Fig*. 3.1 B, the peptide mimicking the region 370-394 is able to bind CypA in a dose-dependent manner showing a K_D value in micromolar range (70 ± 0,5* 10^{-6}M). Instead the peptide AIF(343-360) and the shorter peptides bind CypA with less affinity showing a K_D value in millimolar range. In particular the K_D value for the peptide AIF(343-360) is of 3.5 ± 0,7* 10^{-3} M, for AIF(376-383) is of 3.8 ± 0,3* 10^{-3} M and for AIF(383-394) is of 3.2 ± 0,3* 10^{-3} M (*Fig.* 3.4, B, C and D, respectively), while the peptide AIF(370-376) is not able to bind CypA (data not shown). All data suggest that the amino acidic region of AIF (spanning residues 370-394) is much involved in the interaction with CypA.

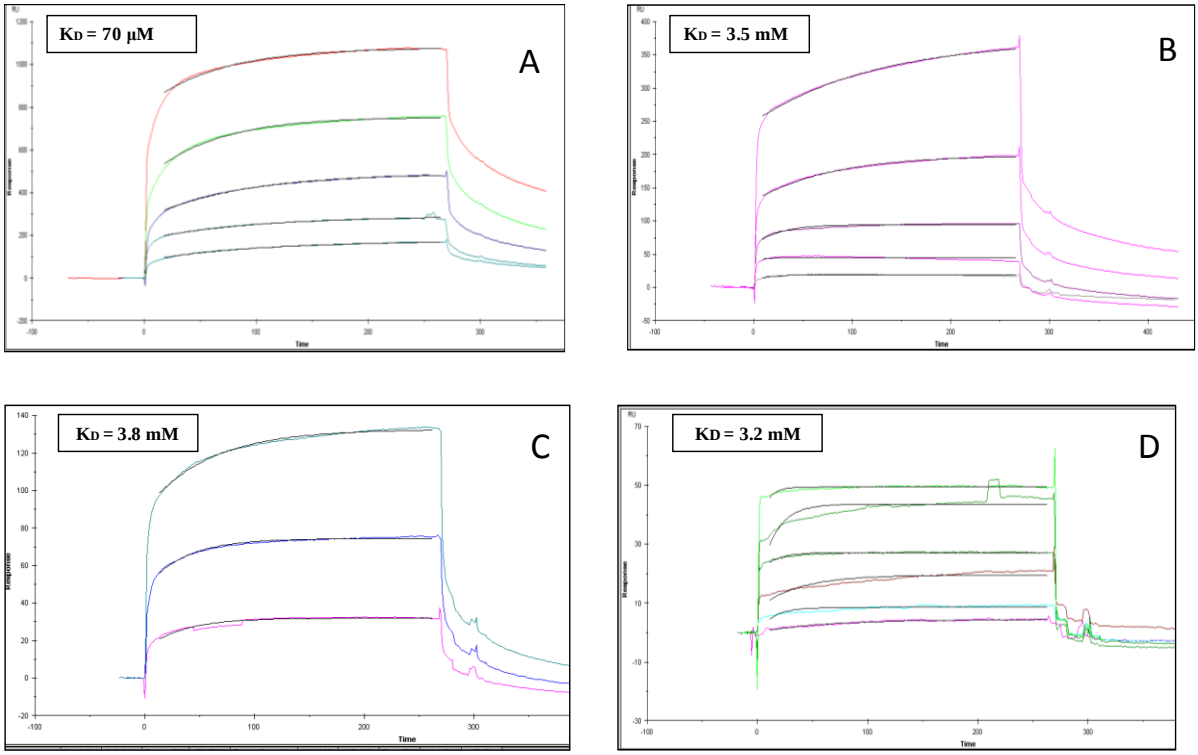


***Fig*. 3.4: SPR-binding experiments of the interaction between CypA and AIF peptides. A, AIF(370-394), B, AIF(343-360); C, AIF(376-383); D, AIF(383-394)).** Increasing concentrations of AIF peptides (1 ÷ 100 µM) were injected over the CypA-derivatized CM5 sensor chip in the same conditions previously reported.

In the following we tested if these regions are able to block the interaction between AIF and CypA setting up SPR-based completion assays. In each experiments, before being injected into the CypA-captured chip, tAIF, at the concentration of 75 nM, was mixed with peptides at different concentrations (in the range 1÷ 100 µM) (*Fig.* 3.5), including the peptides AIF(370-394) and AIF (343-360).

In agreement to what was observed on SPR-binding experimets, the tAIF/CypA binding signal was reduced only when AIF was pre-incubated with the peptide AIF(370-394). All together these data indicated that AIF specifically associates with CypA with the region 370-394.

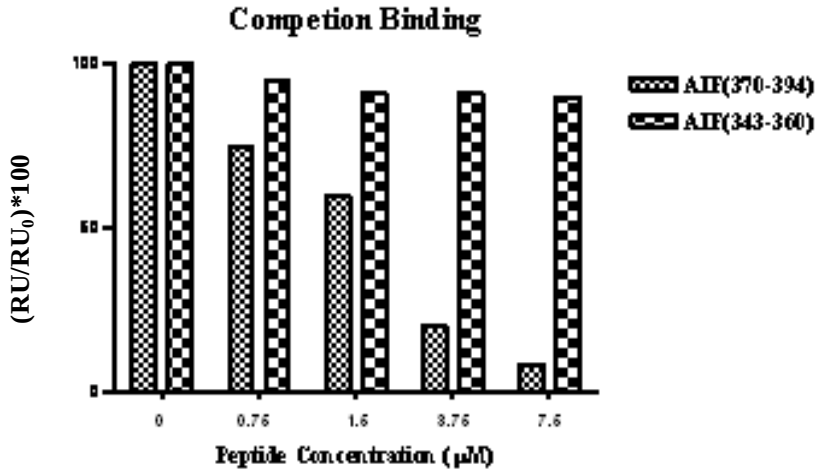


***Fig.* 3.5: AIF(370-394) peptide is able to block tAIF/CypA complex formation.** SPR-based competition experiments using AIF peptides (AIF(370-394) and AIF(343-360)). Results were reported as (RU/RU0)*100, where RU is the maximum response (RUmax) for a given competitor and RU0 is the RUmax without competitors.

3.4. *The peptide AIF(370-394) shows protection in in vitro models of AIF-mediated cell death*

To evaluate the effect of peptide selected in vitro models of AIF-mediated cell death, HT-22 cells were transfected with varying doses of peptides AIF(370-394) and AIF(343-360) (10 µM, 25 µM, 50 µM) and an unrelated peptide at the highest concentration tested (50 µM), used as negative control (40). Cell viability after glutamate treatment (2 mM for 12 h) was assessed by MTT assays. As depicted in *Fig.* 3.6 (upper panel) the peptide AIF(370-394) showed a significant dose dependent protective effect compared the control treated with glutamate. While the peptide AIF(343-360) and the negative control did not work.

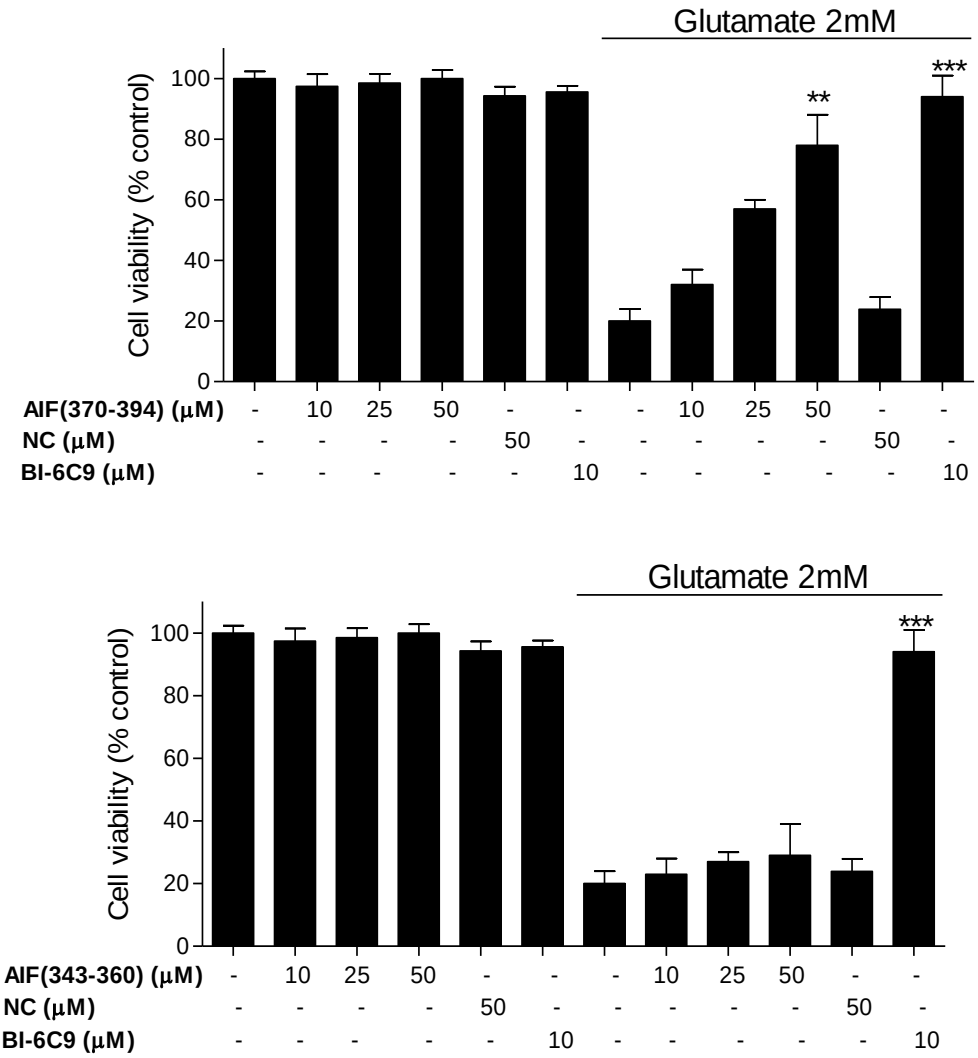


***Fig.* 3.6: The peptide AIF(370-394) shows protection in *in vitro* models of AIF-mediated cell death.** HT-22 cells were transfected with varying doses of peptides AIF(370-394) (upper panel) and AIF(343-360) (lower panel) (10 μM, 25 μM, 50 μM) and an unrelated peptide (as negative control) at the highest concentration tested (50 μM), used as negative control. The Bid-inhibitor BI-6c9 (10 μM) was used as positive control. Cell viability after glutamate treatment (2 mM for 12 h) was assessed by MTT assays.

Moreover, we also evaluated the protective effect of the peptide AIF(370-394) after glutamate treatment by AnnexinV-FITC/PI double staining. HT-22 cells transfected with AIF(370-394) peptide and unrelated peptide were untreated (control) or treated with glutamate 2 mM for 12 h. Then labeled with AnnexinV-FITC and PI, and

analyzed by flow cytometry. In *Fig.* 3.7 the results were reported as percentages of apoptotic cells (double positive).

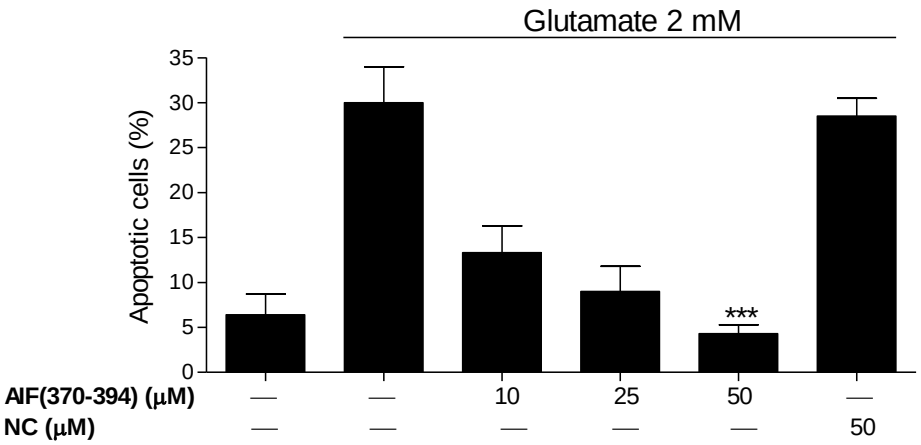


***Fig.* 3.7: The peptide AIF(370-394) blocks the apoptotic process in HT-22 after glutamate treatment.** HT-22 cells were transfected with varying doses of peptides AIF(370-394) (10 µM, 25 µM, 50 µM) and an unrelated peptide (NC) at the highest concentration tested (50 µM), used as negative control. Data are shown as percentage of apoptotic cells (double positive). Experiments were independently repeated three times with n=3 per condition.

Finally, we tested the effect of AIF(370-394) peptide during oxidative stress in HT-22 by monitoring the cellular impedance in real time by the Xcelligence system. HT-22 cells after transfection with AIF(370-394) and an unrelated peptide used as negative control (NC) were treated with glutamate 2 mM.

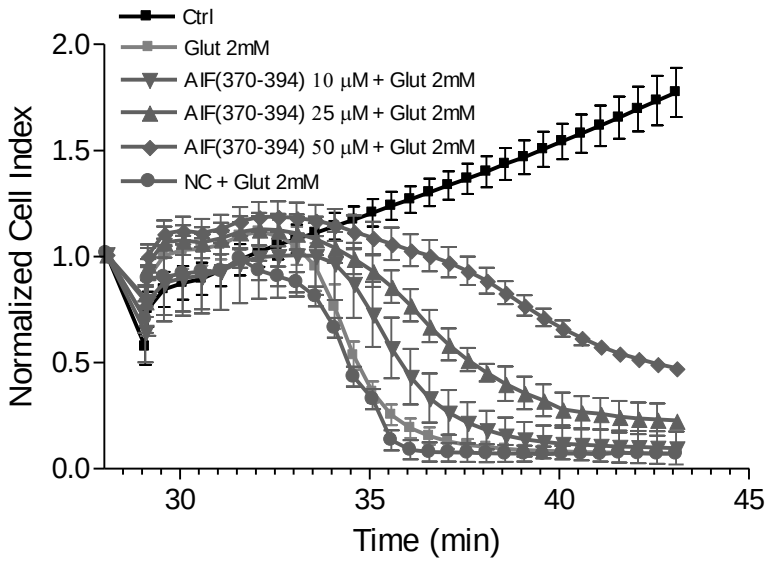


***Fig.* 3.8: AIF(370-394) peptide-treatement of HT-22 cells attenuates glutamate-induced oxidative stress.** HT-22 cells after transfection with AIF(370-394) and an unrelated peptide used as negative control (NC) were treated with glutamate 2 mM. The cellular vitality was continuously monitored for 42h by the XCELLingence system. Experiments were independently repeated three times with n=8 per condition.

So in conclusion we demonstrated that the administration of the peptide mimicking the aminoacidic regions of AIF involved in the interaction with CypA (AIF(370-394)) prevents the glutamate-mediated apoptotic process in HT-22 cells.

4. DISCUSSION

Several prominent neurological disorders manifest symptoms that result from degeneration and death of neurons. The genetic and environmental factors that initiate neurodegeneration differ among diseases; a shared biochemical cascade of events appears to carry out the cell death process. Recently, the importance of cytochrome *c* release and caspase activation in neuronal apoptosis has been supplemented by novel insights into a causal role of caspase-independent death signaling, such as the mitochondrial release of apoptosis inducing factor (AIF), in delayed neuronal death after cerebral ischemia or brain injury (9). Several studies in vitro and in vivo convincingly demonstrated that the pro-apoptotic action of AIF in neurological disorders occurs following its release through mitochondrial membrane and its translocation to the nucleus where induces chromatinolysis (14). Beside, recent data show that translocation of AIF to the nucleus after cerebral hypoxia-ischemia, requires interaction with cyclophilin A (CypA), suggesting a model in which two proteins, that normally reside in separate cytoplasmic compartments, acquire novel properties when moving together to the nucleus (29).

In spite of the considerable amount of work performed on AIF, the mechanisms underlying the nuclear action of AIF are, however, largely unknown. The aim of this project is the further elucidation of the mechanism underlying the neuronal cell death mediated by AIF and CypA proteins by the selective disruption of this complex with peptides mimicking the protein surfaces involved in mutual binding. Using Surface Plasmon resonance (SPR) approach we have demonstrated that Δ1-121AIF (tAIF) and CypA interact directly and with high affinity (K_D=30nM). Then, starting from the molecular model of AIF-CypA complex a set of potential blocking peptides AIF(370-394), AIF(343-360) and the shorter variants AIF(370-376), AIF (376-383) and AIF(383-394) have been designed and synthesized. The precise domain of tAIF involved in the interaction with CypA was mapped by competition and direct binding SPR experiments. We have found that the aminoacidic region 370-394 of mouse AIF specifically associates to CypA (showing IC_{50} and a K_D values in the micromolar range). Data have shown also that the AIF aminoacidic region 343-360 is less involved in the recognition of CypA exhibiting a K_D value in the millimolar range. The peptide is also not able to block the interaction between AIF and CypA. We have developed cell-based assays to evaluate the effect of the peptide AIF(370-396) in neuronal cell lines. The peptide AIF(343-360) and another unrelated peptide have been used as negative controls. In

particular, these peptides have been tested in the glutamate-mediated cell death model system using neuronal HT-22 cell lines (40).

More experiments have been also done to evaluate the protective effects of peptides in HT-22 cells by MTT assays. Results suggest that the peptide AIF(370-394) has a significant dose dependent protective effect while peptides AIF(343-360) and the negative controls (Neg Ctrl) do not work, in good agreement with biochemical observations. Cell death has been evaluated also by FACS analysis using Annexin V-FITC and PI double staining. Data obtained confirm that the peptide AIF(370-394) has a significant dose-dependent protective effect after injury, while the peptide AIF(343-360) and the unrelated peptides don't work, in good agreement with the previous results. The same result has been obtained monitoring the kinetics of cellular apoptosis in neuronal cell lines HT-22 by the XCELLingence system.

The selective disruption of the CypA-AIF complex with this peptide mimicking the protein surfaces involved in mutual binding, should be an ideal strategy to suppress the pro-apoptotic action of AIF without interfering with its implication in bioenergetic and redox metabolism and in the immunosuppressive features of CypA.

REFERENCES

1. World Health Organization Department of Health Statistics and Informatics in the Information, Evidence and Research Cluster (2004). The global burden of disease 2004 update. Geneva: WHO.
2. Blomgren K, Zhu C, Wang X, Karlsson JO, Leverin AL, Bahr BA et al. (2001) Synergistic activation of caspase-3 by m-calpain after neonatal hypoxia-ischemia: a mechanism of pathological apoptosis. J Biol Chem; 276: 10191–10198.
3. Hu BR, Liu CL, Ouyang Y, Blomgren K, Siesjo BK. (2000). Involvement of caspase-3 in cell death after hypoxia-ischemia declines during brain maturation. J Cereb Blood Flow Metab; 20: 1294–1300.
4. Green DR, Kroemer G. The pathophysiology of mitochondrial cell death. (2004) Science; 305: 626–629.
5. Kluck RM, Esposti MD, Perkins G, Renken C, Kuwana T, Bossy-Wetzel E et al. (1999). The proapoptotic proteins, Bid and Bax, cause a limited permeabilization of the mitochondrial outer membrane that is enhanced by cytosol. J Cell Biol; 147: 809–822.
6. Cheng Y, Deshmukh M, D'Costa A, Demaro JA, Gidday JM, Shah A et al. (1998). Caspase inhibitor affords neuroprotection with delayed administration in a rat model of neonatal hypoxic-ischemic brain injury. J Clin Invest; 101: 1992–1999.
7. Susin SA, Daugas E, Ravagnan L, Samejima K, Zamzami N, Loeffler M et al. (2000). Two distinct pathways leading to nuclear apoptosis. J Exp Med; 192: 571–580.
8. Susin SA, Lorenzo HK, Zamzami N, Marzo I, Snow BE, Brothers GM et al. (1999). Molecular characterization of mitochondrial apoptosis-inducing factor. Nature; 397: 441–446.
9. Cheung EC, Melanson-Drapeau L, Cregan SP, Vanderluit JL, Ferguson KL, McIntosh WC et al. (2005). Apoptosis-inducing factor is a key factor in neuronal cell death propagated by BAXdependent and BAX-independent mechanisms. J Neurosci; 25: 1324–1334.
10. Kroemer G, Martin SJ. Caspase-independent cell death. Nat Med 2005; 11: 725–730.
11. Maté MJ., Ortiz-Lombardía M, Boitel B, Haouz A, Tello D, Susin SA., Penninger J, Kroemer G, Alzari PM. (2002). The crystal structure of the mouse apoptosis-inducing factor AIF. Nature Structural Biology 9, 442 – 446.
12. Modjtahedi N, Giordanetto F, Madeo F, Kroemer G. (2006). Apoptosis-inducing factor: vital and lethal. Trends Cell Biol; 16: 264–272.
13. Klein JA, Longo-Guess CM, Rossmann MP, Seburn KL, Hurd RE, Frankel WN et al. (2002). The harlequin mouse mutation downregulates apoptosis-inducing factor. Nature; 419: 367–374.
14. Culmsee C, Zhu C, Landshamer S, Becattini B, Wagner E, Pellechia M et al. (2005) Apoptosisinducing factor triggered by poly(ADP-Ribose) polymerase and bid mediates neuronal cell death after oxygen-glucose deprivation and focal cerebral ischemia. J Neurosci; 25:10262–10272.
15. Polster BM, Basanez G, Etxebarria A, Hardwick JM, Nicholls DG. (2005). Calpain I induces cleavage and release of apoptosis-inducing factor from isolated mitochondria. J Biol Chem; 280: 6447–6454.

16. Ferrand-Drake M, Zhu C, Gido G, Hansen AJ, Karlsson JO, Bahr BA et al. (2003). Cyclosporin A prevents calpain activation despite increased intracellular calcium concentrations, as well as translocation of apoptosis-inducing factor, cytochrome c and caspase-3 activation in neurons exposed to transient hypoglycemia. J Neurochem; 85: 1431–1442
17. Ravagnan L, Gurbuxani S, Susin SA, Maisse C, Daugas E, Zamzami N et al. (2001). Heat-shock protein 70 antagonizes apoptosis-inducing factor. Nat Cell Biol; 3: 839–843.
18. Miramar MD, Costantini P, Ravagnan L, Saraiva LM, Haouzi D, Brothers G, Penninger JM, Peleato ML, Kroemer G, Susin SA. (2001). NADH oxidase activity of mitochondrial apoptosis-inducing factor. J Biol Chem. 11;276(19):16391-8.
19. Dumont C, Dürrbach A, Bidère N, Rouleau M, Kroemer G, Bernard G, Hirsch F, Charpentier B, Susin SA, Senik A. (2000). Caspase-independent commitment phase to apoptosis in activated blood T lymphocytes: reversibility at low apoptotic insult. Blood. 1;96(3):1030-8.
20. Ye H, Cande C, Stephanou NC, Jiang S, Gurbuxani S, Larochette N, Daugas E, Garrido C, Kroemer G and Wu H. (2002). DNA binding is required for the apoptogenic action of apoptosis inducing factor. Nat. Struct. Biol., 9, 680–684.
21. Arnoult D, Parone P, Martinou J-C, Antonsson B, Estaquier J and Ameisen JC. (2002). Mitochondrial release of apoptosis-inducing factor occurs downstream of cytochrome c release in response to several proapoptotic stimuli. J. Cell Biol., 59, 923–929.
22. Cregan SP, Fortin A, MacLaurin JG, Callaghan SM, Cecconi F, Park DS, Dawson TM, Kroemer G and Slack RS. Apoptosis-inducing factor is involved in the regulation of caspase-independent neuronal cell death. (2002). J. Cell Biol., 158, 507–517.
23. Yu SW, Wang H, Poitras MF, Coombs C, Bowers WJ, Federoff HJ, Poirier GG, Dawson TM and Dawson VL. (2002). Mediation of poly(ADP-ribose) polymerase-1-dependent cell death by apoptosis-inducing factor. Science, 297, 259–263.
24. Hisatomi T, Sakamoto T, Murata T, Yamanaka I, Oshima Y, Hata Y, Ishibashi T, Inomata J, Susin SA and Kroemer G. (2001). Relocalization of apoptosis-inducing factor in photoreceptor apoptosis induced by retinal detachment in vivo. Am. J. Pathol., 158, 1271–1278.
25. Ferri KF, Jacotot E, Blanco J, Este´ JA, Zamzami A, Susin SA, Brothers G, Reed JC, Penninger JM and Kroemer G. (2000). Apoptosis control in syncytia induced by the HIV type 1-envelope glycoprotein complex: role of mitochondria and caspases. J. Exp. Med., 192, 1081–1092.
26. Li LY, Luo X and Wang X. (2001). Endonuclease G is an apoptotic DNase when released from mitochondria. Nature, 412, 95–99.
27. Widlak P, Li LY, Wang X and Garrard WT. (2001). Action of recombinant human apoptotic endonuclease G on naked DNA and chromatin substrates: cooperation with exonuclease and DNase I. J. Biol. Chem., 276, 48404–48409.
28. Montague JW, Hughes FM and Cidlowski JA. (1997). Native recombinant cyclophilins A, B, and C degrade DNA independently of peptidylprolyl cis-trans-isomerase activity. Potential roles of cyclophilins in apoptosis. J. Biol. Chem., 272, 6677–6684.

29. Candé C, Vahsen N, Kouranti I, Schmitt E, Daugas E, Spahr C, Luban J, Kroemer RT, Giordanetto F, Garrido C, Penninger JM, Kroemer G. (2004). AIF and cyclophilin A cooperate in apoptosis-associated chromatinolysis. Oncogene. 2004 Feb 26;23(8):1514-21.
30. Fischer , G. , B. Wittmann-Liebold , K. Lang , T. Kiefhaber , and F.X. Schmid . 1989 . Cyclophilin and peptidyl-prolyl cis-trans isomerase are probably identical proteins. Nature . 337 : 476 – 478.
31. Schiene , C. , and G. Fischer . 2000 . Enzymes that catalyse the restructuring of proteins. Curr. Opin. Struct. Biol. 10 : 40 – 45 .
32. Handschumacher , R.E. , M.W. Harding , J. Rice , R.J. Drugge , and D.W. Speicher . 1984 . Cyclophilin: a specifi c cytosolic binding protein for cyclosporin A. Science . 226 : 544 – 547 .
33. Colgan , J. , M. Asmal , and J. Luban . 2000 . Isolation, characterization and targeted disruption of mouse ppia: cyclophilin A is not essential for mammalian cell viability. Genomics . 68 : 167 – 178 .
34. Liu , J. , J.D. Farmer Jr ., W.S. Lane , J. Friedman , I. Weissman , and S.L. Schreiber . 1991 . Calcineurin is a common target of cyclophilincyclosporin A and FKBP-FK506 complexes. Cell . 66 : 807 – 815 .
35. Goldner , F.M. , and J.W. Patrick . 1996 . Neuronal localization of the cyclophilin A protein in the adult rat brain. J. Comp. Neurol. 372 :283 – 293 .
36. Capano , M. , S. Virji , and M. Crompton . 2002 . Cyclophilin-A is involved in excitotoxin-induced caspase activation in rat neuronal B50 cells. Biochem. J. 363 : 29 – 36 .
37. Galigniana , M.D. , Y. Morishima , P.A. Gallay , and W.B. Pratt . 2004 . Cyclophilin-A is bound through its peptidylprolyl isomerase domain to the cytoplasmic dynein motor protein complex. J. Biol. Chem. 279 : 55754 – 55759 .
38. Jaschke , A. , H. Mi , and M. Tropschug . 1998 . Human T cell cyclophilin18 binds to thiol-specifi c antioxidant protein Aop1 and stimulates its activity. J. Mol. Biol. 277 : 763 –769 .
39. Wu , J. , M.J. Matunis , D. Kraemer , G. Blobel , and E. Coutavas . 1995 . Nup358, a cytoplasmically exposed nucleoporin with peptide repeats, Ran-GTP binding sites, zinc fi ngers, a cyclophilin A homologous domain, and a leucine-rich region. J. Biol. Chem. 270 : 14209 – 14213.
40. Grohm J, Plesnila N, Culmsee C. (2010). Bid mediates fission, membrane permeabilization and peri-nuclear accumulation of mitochondria as a prerequisite for oxidative neuronal cell death. Brain Behav Immun.;24(5):831-8.
41. Johnsson, B., Lofas, S., and Lindquist, G. (1991). Immobilization of proteins to a carboxymethyldextran-modified gold surface for biospecific interaction analysis in surface plasmon resonance sensors. Anal Biochem 198, 268-277.

CHAPTER V

1. INTRODUCTION

1.1. HIV/AIDS epidemy

Human immunodeficiency virus (HIV) is a *lentivirus* (a member of the retrovirus family) that causes acquired immunodeficiency syndrome (AIDS) (1,2) which in humans causes a progressive failure of the immune system and allows life-threatening opportunistic infections and cancers to thrive. Infection with HIV occurs by the transfer of blood, semen, vaginal fluid, pre-ejaculate, or breast milk. Within these fluids HIV is present as both free particles virus and within infected immune cells. The four major routes of transmission are unsafe sex, contaminated needles, breast milk, and transmission from an infected mother to her baby at birth (perinatal transmission). Screening of blood products for HIV has largely eliminated transmission through blood transfusions or infected blood products in the developed world.

HIV infection (*Fig.* 1.1) in humans is considered pandemic by the World Health Organization (WHO), complacency may play a key role in HIV risk, (3,4) from its discovery in 1981 to 2006, AIDS killed more than 25 million people. HIV infects about 0.6% of the world's population (5). In 2009, AIDS claimed an estimated 1.8 million lives, down from a global peak of 2.1 million in 2004. Approximately 260,000 children died of AIDS in 2009. A disproportionate number of AIDS deaths occur in Sub-Saharan Africa, retarding economic growth and exacerbating the burden of poverty. An estimated 22.5 million people (68% of the global total) live with HIV in sub-Saharan Africa, which is also home to 90% of the world's 16.6 million children orphaned by HIV (6). Treatment with antiretroviral drugs reduces both the mortality and the morbidity of HIV infection. Although antiretroviral medication is still not universally available, expansion of antiretroviral therapy programs since 2004 has helped to turn the tide of AIDS deaths and new infections in many parts of the world. Intensified awareness and preventive measures, as well as the natural course of the epidemic, have also played a role. Nevertheless, an estimated 2.6 million people were newly infected in 2009 (6).

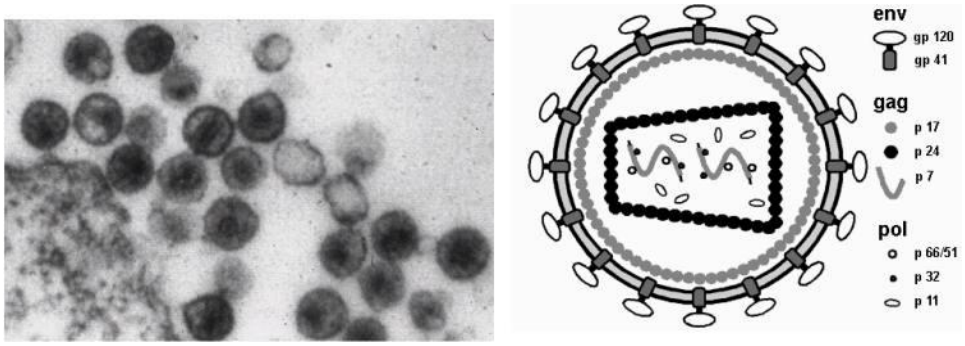


***Fig.* 1.1: Human immunodeficiency virus (HIV)**

1.2. HIV's structure and genome

HIV is different in structure from other retroviruses. It is roughly spherical (7) with a diameter of about 120 nm, around 60 times smaller than a red blood cell It is composed of two copies of positive single-stranded RNA that codes for the virus's nine genes enclosed by a conical capsid composed of 2,000 copies of the viral protein p24 (8) The single-stranded RNA is tightly bound to nucleocapsid proteins, p7, and enzymes needed for the development of the virion such as reverse transcriptase, proteases, ribonuclease and integrase. A matrix composed of the viral protein p17 surrounds the capsid ensuring the integrity of the virion particle. This in turn, is surrounded by the viral envelope that is composed of two phospholipidic layers taken from the membrane of a human cell when a newly formed virus particle buds from the cell. Embedded in the viral envelope are proteins from the host cell and about 70 copies of a complex HIV protein that protrudes through the surface of the virus particle (8) This protein, known as Env, consists of a cap made of three molecules called glycoprotein (gp) 120, and a stem consisting of three gp41 molecules that anchor the structure into the viral envelope. This glycoprotein complex enables the virus to attach to and fuse with target cells to initiate the infectious cycle (9). Both these surface proteins, especially gp120, have been considered as targets of future treatments or vaccines against HIV (*Fig.*1.2).

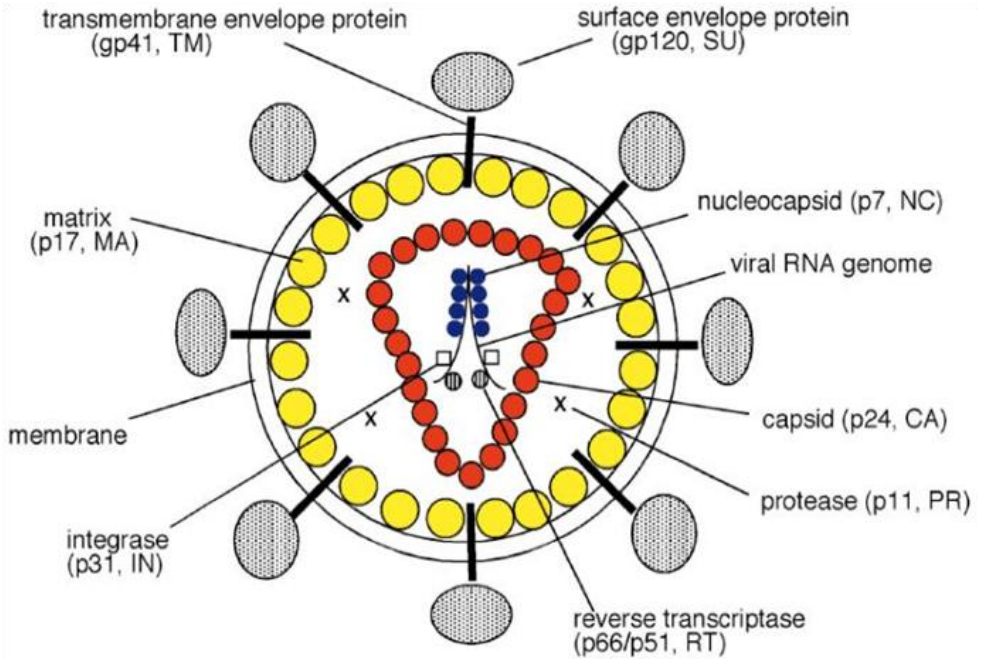


***Fig.* 1.2: Schematic representation of a mature HIV-1 particle**. Positions of the major viral proteins, the lipid bilayer, and the genomic RNA are indicated

The RNA genome consists of at least seven structural landmarks (LTR, TAR, RRE, PE, SLIP, CRS, and INS), and nine genes (gag, pol, and env, tat, rev, nef, vif, vpr, vpu, and sometimes a tenth tev, which is a fusion of tat env and rev), encoding 19 proteins (*Fig.*1.3). Three of these genes, gag, pol, and env, contain information needed to make the structural proteins for new virus particles (8). For example, env codes for a protein called gp160 that is broken down by a viral enzyme to form gp120 and gp41. The six remaining genes, tat, rev, nef, vif, vpr, and vpu (or vpx in the case of HIV-2), are regulatory genes for proteins that control the ability of HIV to infect cells, produce new copies of virus (replicate), or cause disease (8).

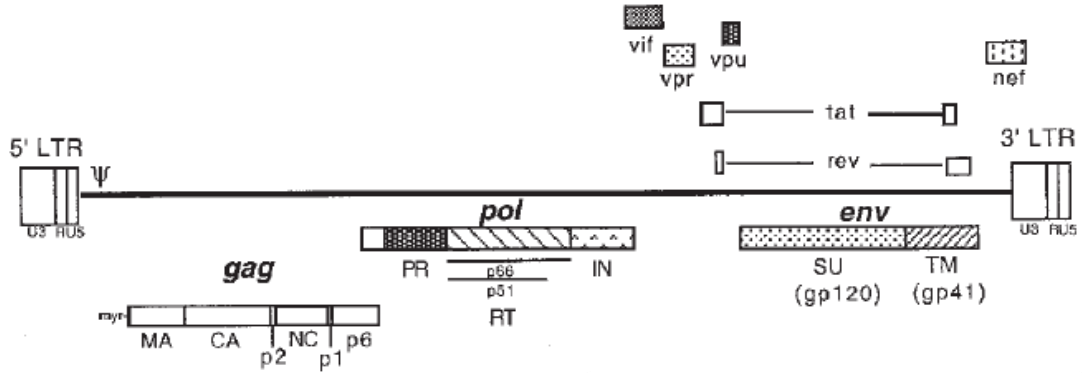


***Fig.* 1.3: Organization of the HIV-1 genome:** The relative locations of the HIV-1 open reading frames *gag*, *pol*, *env*, *vif*, *vpr*, *vpu*, *nef*, *tat*, and *rev* are indicated. The 5' and 3' LTRs are shown, with U3, R, and U5 regions noted. The ψ indicates the position of the RNA packaging signal. The major Gag domains (MA, CA, NC, p6) and the Gag spacer peptides (p2 and p1) are shown under the *gag* gene. The site of Gag N-terminal myristylation is denoted as "myr". Under the *pol* gene are indicated the PR, RT (p66 and p51 subdomains), and IN coding regions. The SU and TM Env glycoproteins (gp120 and gp41, respectively) are enlarged to show the position in gp120 of the major conserved (C1-C5) and variable (V1-V5) regions and in gp41 the location of the fusion peptide, the N- and C-helices, membrane-spanning domain, and the cytoplasmic tail.

1.3. The virus life cycle

The HIV-1 replication cycle (10) may be divided into 15 discrete steps, as depicted in *Fig.* 1.4. Starting from the integration of the viral genome into a host chromosome the order of events leads to expression of the viral gene products, production of virus particles, infection of a new cell, and reintegration of the viral genome. In step 1, viral transcripts are expressed from the promoter located in the 50 long terminal repeat (LTR), with Tat greatly enhancing the rate of transcription. In step 2, a set of spliced and genomic-length RNAs are transported from the nucleus to the cytoplasm, where they can be translated or packaged in this Rev has a regulatory action. In step 3, viral mRNAs are translated in the cytoplasm, and the Gag and Gag-Pol polyproteins become localized to the cell membrane. The Env mRNA is translated at the endoplasmic reticulum (ER). In step 4, the core particle is assembled from the Gag and Gag-Pol polyproteins (later processed to MA, CA, NC, p6, PR, RT, and IN), Vif, Vpr, Nef, and the genomic RNA, and an immature virion begins to bud from the cell surface. To provide SU and TM proteins for the outer

membrane coat during budding, the Env polyprotein must first be released from complexes with CD4 (the cell surface HIV-1 receptor), which is coexpressed with Env in the ER. Vpu assists this process by promoting CD4 degradation, as shown in step 5. Env is then transported to the cell surface (step 6), where again it must be prevented from binding CD4. Nef promotes endocytosis and degradation of surface CD4 (step 7). As the particle buds and is released from the cell surface coated with SU and TM (step 8), the virion undergoes a morphologic change known as maturation (step 9). This step involves proteolytic processing of the Gag and Gag-Pol polyproteins by PR and a less well defined function of Vif. The mature virion is then ready to infect the next cell, which is targeted by interactions between SU and CD4 and CC or CXC chemokine coreceptors (step 10). Following binding, TM undergoes a conformational change that promotes virus-cell membrane fusion, thereby allowing entry of the core into the cell (step 11). The virion core is then uncoated to expose a viral nucleoprotein complex, which contains MA, RT, IN, Vpr, and RNA (step 12). This complex is transported to the nucleus (step 13), where the genomic RNA is reverse transcribed by RT into a partially duplex linear DNA (step 14). Then IN catalyzes integration of the viral DNA into a host chromosome and the DNA is repaired (step 15), thereby completing the viral replication cycle.

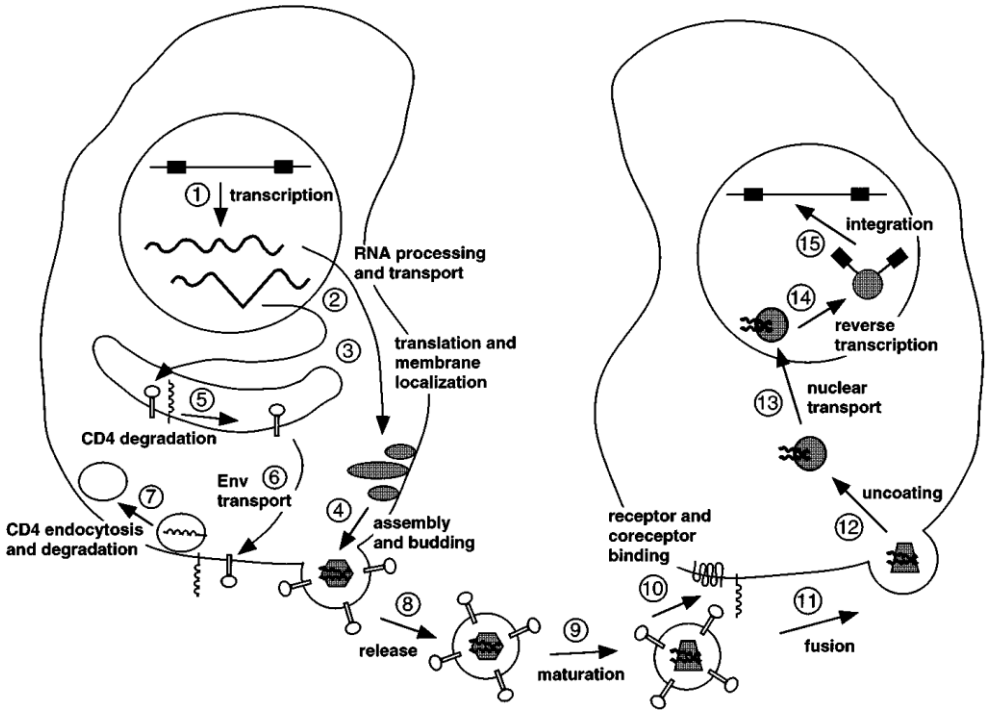


Fig. 1.4: HIV-1 replication cycle. Steps 1–15 are described in the text.

1.3.1. Replication cycle

HIV enters macrophages and CD4+ T cells by the adsorption of glycoproteins on its surface to receptors on the target cell followed by fusion of the viral envelope with the cell membrane and the release of the HIV capsid into the cell (11,12).

Entry into the cell begins through the interaction of the trimeric envelope complex (gp160 spike) and both CD4 and a chemokine receptor (generally either CCR5 or CXCR4) on the cell surface (11,12). gp120 binds to integrin α4β7 activating LFA-1 the central integrin involved in the establishment of virological synapses, which facilitate efficient cell-to-cell spreading of HIV-1 (13). The gp160 spike contains binding domains for both CD4 and chemokine receptors (11,12).

The first step in fusion involves the high-affinity attachment of the CD4 binding domains of gp120 to CD4. Once gp120 is bound with the CD4 protein, the envelope complex undergoes a structural change, exposing the chemokine binding domains of gp120 and allowing them to interact with the target chemokine receptor (11,12).This allows for a more stable two-pronged attachment, which allows the N-terminal fusion peptide gp41 to penetrate the cell membrane (11,12). Repeat sequences in gp41, HR1, and HR2 then interact, causing the collapse of the extracellular portion of gp41 into a hairpin. This loop structure brings the virus and cell membranes close together, allowing fusion of the membranes and subsequent entry of the viral capsid (11,12).

After HIV has bound to the target cell, the HIV RNA and various enzymes are injected into the cell (11). During the microtubule-based transport to the nucleus, the viral single-strand RNA genome is transcribed into double-strand DNA, which is then integrated into a host chromosome (*Fig.* 1.5).

Upon the viral capsid entry into the cell, the enzyme reverse transcriptase liberates the single-stranded (+) RNA genome from the attached viral proteins and copies it into a complementary DNA (cDNA) molecule (14). The process of reverse transcription is extremely error-prone, and the resulting mutations may cause drug resistance or allow the virus to evade the body's immune system. The reverse transcriptase also has ribonuclease activity that degrades the viral RNA during the synthesis of cDNA, as well as DNA-dependent DNA polymerase activity that creates a sense DNA from the antisense cDNA. The cDNA along with its complement form a double-stranded viral DNA that is transported into the cell

nucleus. The integration of the viral DNA into the host cell's genome is carried out by the viral integrase enzyme (14).

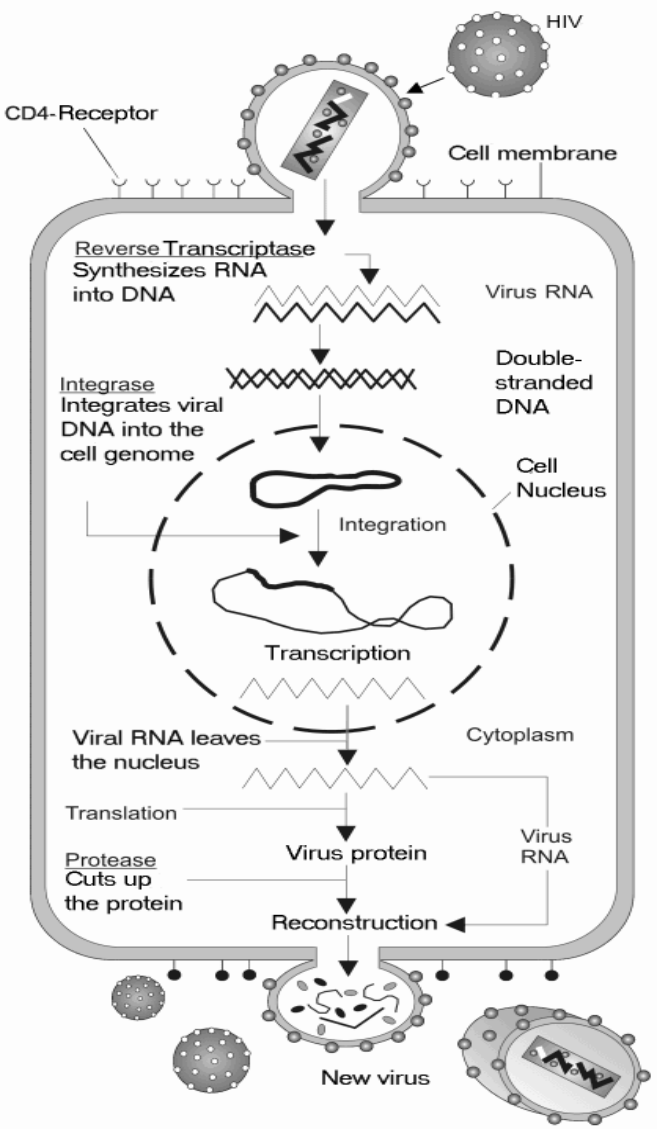


***Fig.* 1.5: The HIV replication cycle**

This integrated viral DNA may then lie dormant, in the latent stage of HIV infection (14) but to actively produce the virus, certain cellular transcription factors need to be present, the most important of which is NF-κB, which is upregulated when T-cells become activated (15). This means that those cells most likely to be killed by HIV are those currently fighting infection.

During viral replication, the integrated DNA provirus is transcribed into mRNA, which is then spliced into smaller pieces. These small pieces are exported from the nucleus into the cytoplasm, where they are translated into the regulatory proteins Tat and Rev. As the newly produced Rev protein accumulates in the nucleus, it binds to viral mRNAs and allows unspliced RNAs to leave the nucleus, where they are otherwise retained until spliced (16). At this stage, the structural proteins Gag and Env are produced from the full-length mRNA. The full-length RNA is actually the virus genome; it binds to the Gag protein and is packaged into new virus particles.

The final step of the viral cycle that is the assembly of new HIV-1 virons, begins at the plasma membrane of the host cell. The Env polyprotein (gp160) goes through the endoplasmic reticulum and is transported to the Golgi complex where it is cleaved by protease and processed into the two HIV envelope glycoproteins gp41 and gp120. These are transported to the plasma membrane of the host cell where gp41 anchors the gp120 to the membrane of the infected cell. The Gag (p55) and Gag-Pol (p160) polyproteins also associate with the inner surface of the plasma membrane along with the HIV genomic RNA as the forming virion begins to bud from the host cell. Maturation occurs either in the forming bud or in the immature virion after it buds from the host cell. During maturation, HIV proteases cleave the polyproteins into individual functional HIV proteins and enzymes. The various structural components then assemble to produce a mature HIV virion that is then able to infect another cell.

1.4. Trans-Activator of Transcription(TAT) protein

The HIV-1 promoter is located in the 5’ LTR and contains a number of regulatory elements important for RNA polymerase II transcription. The LTR is composed of three regions: U3 (for unique, 3' end), R (for repeated) and U5 (for unique, 5' end). Transcription initiates at the U3/R junction. U3 contains a variety of elements that direct the binding of RNA polymerase II (pol II) to the DNA template. A TATA element, to which transcription factor IID (TFIID) binds, is located approximately 25 nucleotides upstream of the transcription start site. Located 5' of the TATA box are three Sp1 and two NF-ϰB binding sites. Although mutational analyses of the HIV-1 LTR reveal that the Sp1 and NF-ϰB sites are variably important, depending upon

the cell type, removal of all Sp1 and NF-ϰB sites abolishes virus replication (17). Upstream of the NF-ϰB sites is a domain, sometimes referred to as the “modulatory region”, which contains binding sites for several additional transcriptional factors, including LEF, Ets, and USF. Despite the importance of these factors, transcription complexes initiated at the HIV-1 promoter are rather inefficient at elongation; RNA synthesis is greatly increased (by more than two logs) when the transcriptional transactivator protein Tat is present (18,19) (*Fig*. 1.6).

In the absence of Tat, polymerases generally do not transcribe beyond a few hundred nucleotides, though they do not appear to terminate at specific sites. It is still unclear how Tat causes transcribing polymerases to become sufficiently processive to completely transcribe the ~9-kb viral genome, but recent experiments suggest that Tat may assemble into transcription complexes and recruit or activate factors that phosphorylate the RNA polymerase II C-terminal domain (CTD), including the general transcription factor TFIIH and other novel kinases (20).

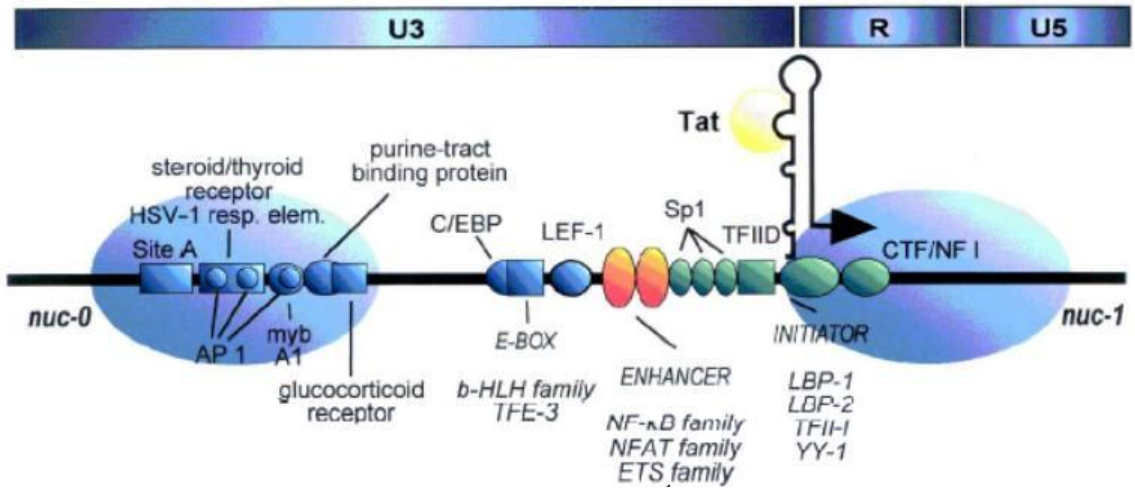


***Fig.* 1.6: Schematic representation of the HIV-1 LTR**: The position of binding sites for host factors (LBP-1, NF-κB, LEF, Ets, USF-1, and NFAT-1) are shown at and 5' of the transcription start site. The TAR stem/loop structure, with bulge, is represented at the 5' end of a nascent mRNA.

Tat consists of between 86 and 101 amino acids depending on the subtype encoded by a two-exon RNA; Tat contains several distinct functional domains: an

activation domain, which lies within the N-terminal 48 residues of the protein and which itself is comprised of an acidic domain (2-11 residues), a Cys-rich region (22-37 residues), and a hydrophobic core element (38-48 residues); a highly basic RNA binding domain (Arginine-rich domain); and an overlapping nuclear localization signal (*Fig.* 1.7).

***Fig*. 1.7: HIV Tat protein**: in red is reported Cys-rich domain and in blue Arg-rich domain

Unlike typical transcriptional activators, Tat binds not to a DNA site but rather to an RNA hairpin known as TAR (trans-activating response element), located at the 50 end of the nascent viral transcripts. TAR, which is present at the 5' end of all viral RNAs, consists of a base-paired stem, a small (trinucleotide) non-base-paired bulge, and a 6-nucleotide G-rich loop. The arginine-rich domain of Tat helps mediate binding to a three-nucleotide bulge region of TAR, with one arginine residue being primarily responsible for recognition. NMR studies of TAR complexed to arginine (21) show a base-specific contact between the arginine side chain and a guanine in the RNA major groove. The complex is stabilized by additional contacts to the phosphate backbone and by formation of a U-A:U base triple between a bulge nucleotide and a base pair above the bulge. NMR studies of the full-length 86–amino acid Tat protein have suggested that a hydrophobic core region of about 10 amino acids adopts a defined structure but that the rest of the molecule, including the arginine-rich RNA-binding domain, is relatively disordered (22). It seems likely that Tat requires interactions with cellular proteins in addition to TAR to adopt a stable structure.

This prediction was borne out recently with the identification of a cellular protein that interacts, via its activation domain, with Tat (23). The protein in question is cyclin T1 (cycT1), which forms a heterodimer with a member of the cyclin-dependent kinase family (CDK9). The cycT1/CDK9 heterodimer is itself part of a large protein complex

related to the Drosophila positive-transcriptional elongation factor b (P-TEF-b). Tat recruits the human P-TEFb complex to TAR, resulting in the phosphorylation of the C-terminal domain (CTD) of RNA Pol II and a dramatic stimulation of transcriptional processivity (*Fig.* 1.8).

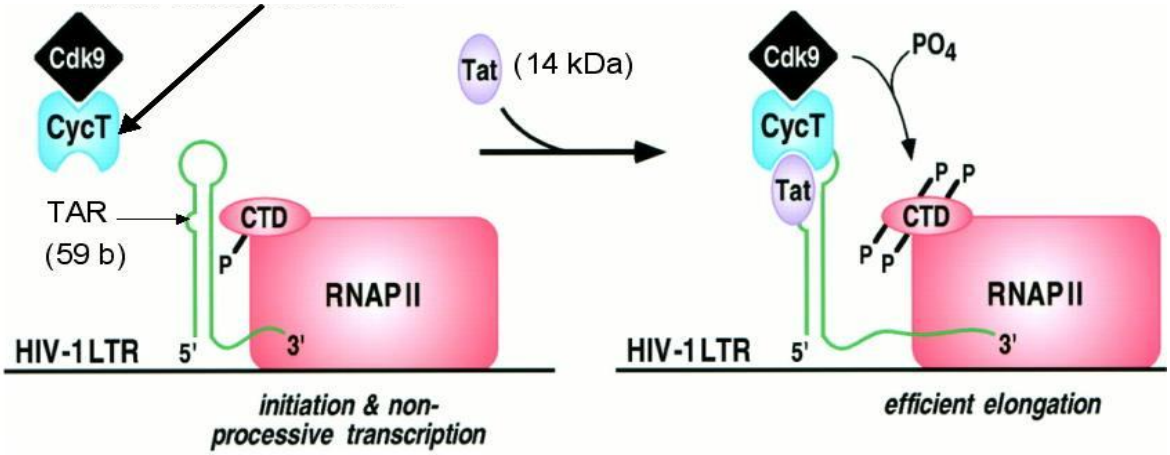


***Fig.* 1.8: Mechanisms of HIV-1 Tat function.** The structure of nascent HIV-1 RNA containing the 5′ TAR element, with bound Tat and cellular proteins is shown.

1.5. IκB-α/ NF-κB complex

The NF-κB transacting factors enhance the HIV-1 transcription by binding to the κB sites of the HIV-1 LTR (24). The NF-κB proteins, namely p105/p50, p100/52, p65/RelA, c-Rel, and RelB, share a Rel homology domain that is required for subunit dimerization, nuclear localization, and DNA binding (25). The NF-κB transcriptional activity is negatively regulated through the association with the IκB proteins (26). IκB-α (nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha), is the best characterized and ubiquitous member of the IκB family. In unstimulated cells, IκB-α associates with the p50/p65 NF-κB complex, preventing its nuclear accumulation, the association with DNA and transcriptional activity of NF-κB (25,26). The IκB-α/NF-κB complex shuttles in and out of the nucleus and is prevalently retained in the cytoplasm (27). Upon NF-κB activating stimuli, IκB-α is phosphorylated at serines 32 and 36 by the κB kinase complex, ubiquitinated at lysines 21 and 22, and proteolyzed by the 26 S proteasome (25, 26). Following the proteolysis of IκB-α, the free NF-κB complex translocates to the

nucleus, where it binds to the ϰB sites and activates the transcription of the NF-ϰB dependent genes, including the IϰB-α gene (25, 26). The newly synthesized IϰB-α migrates to the nucleus, where it displaces the NF-ϰB complex from DNA and promotes its nuclear export, thus terminating the NF-ϰB transcriptional activity (28) (*Fig*. 1.9).

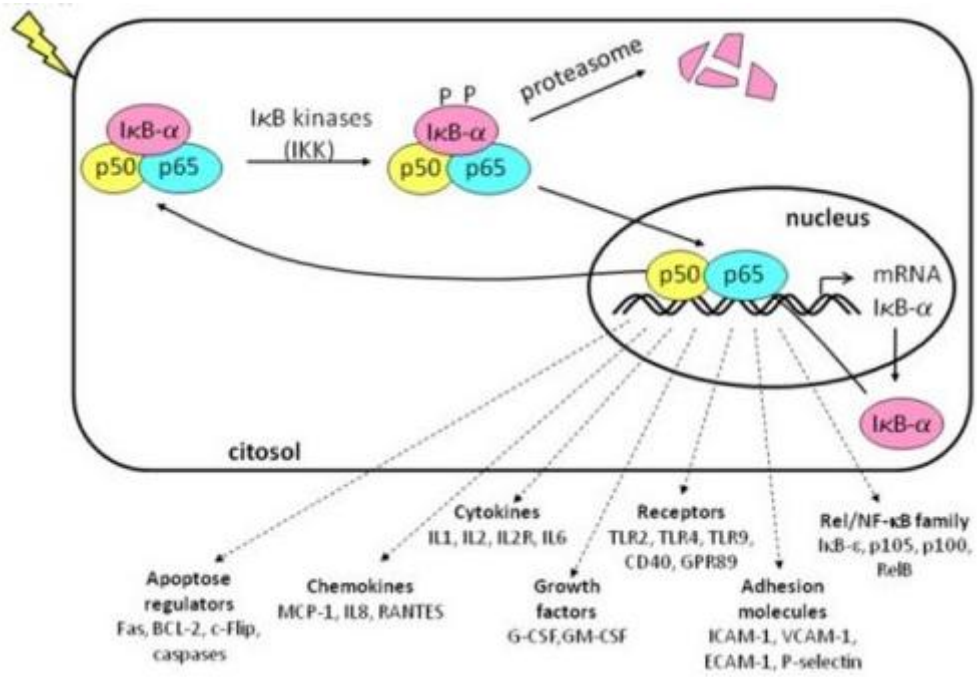


***Fig.* 1.9: Classical pathway of NF-ϰB activation**

From a molecular point of view, IκB-α is a 36 kDa protein composed of: a) a N-terminal signal response region, where phosphorylation and ubiquitination occur in response to NF-κB activation signals; b) an ankyrin (AR)-rich region made of six ankyrin repeats; c) a C-terminal PEST sequence rich in proline, glutamic acid, serine, and threonine that is involved in basal degradation of free IκB-α (29). Each AR repeat of IκB-α has a length of approximately 35 amino acids, and is composed of a β-hairpin followed by two anti-parallel α-helices and a variable loop. AR repeats represent a protein-protein interaction domain that is shared by different proteins with highly varied functions (30,31). In fact, IκB-α binds to the NF-κB dimer via its AR domain (spanning residues 67-287) interfering with the NF-κB interaction with DNA. The binding affinity of IκB-α to p50/p65 hetero- and p65/p65 homo-dimers is

in the picomolar range, and recent studies revealed that binding interface has two hot spots one at either end of the interface (32,33). Notably, in contrast to its remarkable stability in the NF-κB-bound state, free IκB-α is intrinsically unstable and is rapidly degraded through a process that does not require phosphorylation and ubiquitination (34). In free state, ARs 1–4 of IκB-α, are compactly folded while ARs 5–6 are presumably weakly folded and highly flexible; however, ARs 5-6 adopt a fully folded conformation when IκB-α binds to NF-κB (35,36,37). The coupled folding and binding of ARs 5–6 in IκB-α has been proposed to modulate the binding affinity of IκB-α to NF-κB (38) as well as the switch between the basal and stimulated degradation mechanisms (36). Furthermore, ARs 5-6 appear to be involved in the mechanism by which IκB-α increases the dissociation rate of NF-κB from the DNA (39).

1.6. IκB-α /TAT complex

Recent studies found that IκB-α represses the Tat-mediated transactivation of the HIV-1 LTR upon cell transfection. IκB-α physically interacts with Tat and determines its nuclear export to the cytoplasm, which results in the inhibition of Tat-dependent HIV-1 expression (40). Consistently, recombinant HIV-1 and SIVmac239 retroviruses that expressed the IκB-αS32/36A mutant, which is resistant to the signaling-induced proteolysis were highly attenuated in cell culture (41). They have shown that IκB-α represses the Tat activity independently of the NF-κB inhibitory activity by physical association and displacement of Tat from the nucleus to the cytoplasm. The association of IκB-α with the arginine rich domain of Tat requires both the NLS (amino acids 110–120) and the C-NES (amino acids 265–277) together with the binding site for Tat (amino acids 263–269) in order to interfere with the nuclear distribution and the transcriptional activity of Tat. So IκB-α binds to Tat in the nucleus and exports the viral transactivator to the cytoplasm, where the complex IκB-α/Tat is mostly retained. IκB-α is subjected to persistent proteolysis in the course of HIV-1 infection (42), in fact the HIV-1 entry through the gp120 envelope protein binding to CD4 receptor activates the IκB kinase complex, which promotes the proteolysis of IκB-α (43). This event leads to the transcriptional activation of NF-κB-dependent genes, including the HIV-1 genome and pro-

inflammatory genes, which in turn sustains the proteolysis of IκB-α and the activation of NF-κB (44). For this reason, the endogenous IκB-α does not counteract the viral expression in HIV-1-infected cells. The physical and functional interaction of IκB-α with Tat discloses a novel mechanism of HIV-1 transcriptional regulation. Indeed the ratio between IκB-α and Tat could determine the level of expression of the target genes, including HIV-1 (*Fig.* 1.10).

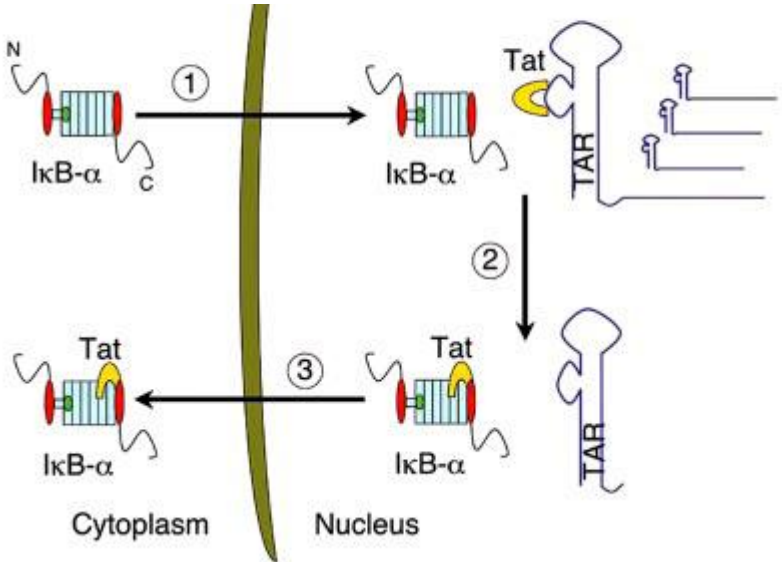


***Fig.* 1.10: Schematic representation of the mechanism of Tat inhibition by IκB-α**. The IκB-α repressor enters in the nucleus (step 1), where it associates to Tat (step 2) and exports the viral transactivator to the cytoplasm (step 3). The nuclear localization signal, the carboxyl-terminal nuclear export signal, and the Tat-binding site of IκB-α are required for the nuclear export of Tat.

In this regard, the novel mechanism of Tat inhibition by IκB-α described here could assist in the development of a novel class of HIV-1 inhibitors and thus compounds that increase the stability of endogenous IκB-α, such as inhibitors of IκB kinase and proteasomes, could be additional tools of conventional anti-retroviral therapies by inhibiting the NF-κB-Tat-dependent HIV-1 transcription.

1.7. Aim of the work

The purpose of the project was the quantification of the binding between IκB-α and Tat proteins through *in vitro* Surface Plasmon Resonance (SPR) experiments. Moreover, the shorter regions of the two proteins directly involved in the IκB-α–Tat complex formation should be investigated through the development of new IκB-α-based peptides. The characterization of IκB-α peptides will provide novel insights into the intrinsic properties of IκB-α fragments and structural determinants of IκB-α/Tat recognition. These studies were carried out by using multiple approaches to chemical synthesis, biochemical assays and mutagenesis studies.

2. EXPERIMENTAL SECTION

2.1. Materials

Reagents for peptide synthesis (Fmoc-protected amino acids and resins, activation and deprotection reagents) were from Novabiochem (Laufelfingen, Switzerland) and InBios (Napoli, Italy). Solvents for peptide synthesis and HPLC analyses were from Romil (Dublin, Ireland); reversed phase columns for peptide analysis and the LC-MS system were from ThermoFisher (Milan, Italy). Solid phase peptide syntheses were performed on a fully automated multichannel peptide synthesizer Syro I (Multisynthech, Germany). Preparative RP-HPLC were carried out on a Shimadzu LC-8A, equipped with a SPD-M10 AV detector and with a Phenomenex C18 Jupiter column (50x22 mm ID; 10 µm). LC-MS analyses were carried out on a LCQ DECA XP Ion Trap mass spectrometer equipped with a OPTON ESI source, operating at 4.2 kV needle voltage and 320 °C with a complete Surveyor HPLC system, comprised of MS pump, an autosampler and a photo diode array (PDA). Narrow bore 50x2 mm C18 BioBasic LC-MS columns were used for these analyses.

2.2. Protein and peptide synthesis

The entire HIV-1 Tat protein [1-86] was prepared as already reported [40]. Peptides were synthesized employing the solid phase method on a 50 µmol scale following standard Fmoc strategies (45). RINK AMIDE resin (substitution 0.5 mmol/g) was used as solid support. Activation of amino acids was achieved using HBTU/HOBt/DIEA (1:1:2), whereas Fmoc deprotection was carried out using a 20% (v/v) piperidine solution in DMF. All couplings were performed for 15 min and deprotections for 10 min and at the end of peptide chain assembly peptides were acetylated at their N-termini. Peptides were removed from the resin by treatment with a TFA:TIS:H2O (90:5:5, v/v/v) mixture for 90 min at room temperature; then crude peptides were precipitated in cold ether, dissolved in a water/acetonitrile (1:1, v/v) mixture and lyophilized.

Products were purified by RP-HPLC applying a linear gradient of 0.1% TFA CH3CN in 0.1% TFA water from 5% to 65% over 12 min using a semi-preparative 2.2x5 cm C18 column at a flow rate of 20 mL/min. Peptides and purity and identity were

confirmed by LC-MS. Purified peptides were lyophilized and stored at -20 °C until use.

2.3. Plasmids

The nucleotide sequence of IκB-α was amplified by PCR from pcDNA-3HA- IκB-α (40) with the forward primer GGCGCTAGCATGTTCCAGGCGGCCG and reverse primer CGGGAATTCTCATAACGTCAGACGCTG. The PCR product was digested with NheI/EcoRI and ligated to pPet28c (Novagen, Darmstadt, Germany) to generate pPet28c-IκB-α. The eukaryotic expression plasmids for IκB-α RPSTRIQQQL/A, IκB-α TRIQQQL/A, IκB-α QQQ/A and R264/S were generated by site-directed mutagenesis of pcDNA-3HA-IkB-α (40). pRc/CMV-p65, p3xFLAG-CMV-Tat and pNF-κB-deleted-LTR-Luc were previously described (46); pκBluc and pSV-β-Gal were purchased from Promega (Madison, WI, USA).

2.4. Production and purification of recombinant IκB–α

Escherichia coli BL21 (DE3) culture was transformed with pPet28c-IκB-α by calcium chloride and transformants were grown in LB medium (1 L) containing kanamycin (250 μg/L) up to OD_{600nm} = 0.5. Then, the bacterial culture was induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) for 6 h. Bacteria were harvested by centrifugation (5000 x *g)* for 15 min at 4°C. The pellet was resuspended into 10 mL of Native Binding Buffer (50mM NaH_2PO_4, 500mM NaCl, 1mM DTT and 5 mM Imidazole, pH 8.0), and sonicated 10 times for 10 s by using Bandelin Sonoplus GM70 (Bandelin Electronic, Berlin, Germany). The lysate was clarified by centrifugation (15000 x *g*) for 20 min and incubated with 5 mL of ProBond Nickel chelating Resin (Invitrogen, Carlsbad, CA, USA) at 4°C overnight. The resin was extensively washed with Wash Buffer (50mM NaH_2PO_4, 500mM NaCl, 20 mM Imidazole, pH 8.0). and, the recombinant protein 6xHis-IκB-α was eluted in Elution Buffer (50mM NaH_2PO_4, 500mM NaCl, 1mM DTT, 500 mM Imidazole, pH 8.0). 6xHis-IkB-a recombinant protein was desalted with a Sephadex G-25 column (GE Healthcare, UK) in a buffer containing 1x PBS, 1mM DTT and 250 mM NaCl. The protein was concentrated at 0.2 mg/mL (5.36 μM) by using Centricon columns YM-10 (Millipore, Bedford, MA, USA).

2.5. SPR analysis

The BIAcore 3000 SPR system for Real time binding assay and related reagents were from GE Healthcare (Milano, Italy).

Tat protein was immobilized at a concentration of 50 µg/mL in 10 mM acetate buffer pH 4.0 (flow rate 5 µL/min, time injection 7 min) on a CM5 Biacore sensor chip, using EDC/NHS chemistry following the manufacturer's instructions (47). Residual reactive groups were deactivated by treatment with 1 M ethanolamine hydrochloride, pH 8.5. Reference channel was prepared by activating with EDC/NHS and deactivating with ethanolamine. Binding assays were carried out at 20 µL/min, with 4,5 min contact-time, IκB-α protein and peptides were diluted in the running-buffer, HBS (10 mM Hepes, 150 mM NaCl, 3 mM EDTA, pH 7.4) 0.1 mM TCEP. Analyte injections of 90 µL were performed at the indicated concentrations. BIAevaluation analysis package (version 4.1, GE Healthcare, Milano, Italy) implemented by instrument software was used to subtract the signal of the reference channel and to evaluate kinetic and thermodynamic parameters of complexes.

Competitive SPR experiments with Tat peptides were carried out pre-incubating Tat peptides with IκB-α full-length in a 2:1 ratio (Tat peptides: IκB-α 160: 80 nM).

Pep $_{260\text{-}272}$ and Pep $_{262\text{-}280}$ were analyzed in the concentration ranges 1-10 µM and 1-20 µM, respectively, while Pep $_{262\text{-}287}$ in 1-40 µM and Pep $_{243\text{-}263}$ in 10-700 µM, respectively.

2.6. Circular Dichroism (CD) Spectroscopy

CD spectra were recorded on a Jasco J-810 spectropolarimeter (JASCO Corp, Milan, Italy). CD spectra were registered at 25 °C in the far UV region from 190 to 260 nm. Each spectrum was obtained averaging three scans, subtracting contributions from corresponding blanks and converting the signal to mean residue ellipticity in units of deg * cm2 * dmol -1* res^{-1}. Other experimental settings were: 20 nm/min scan speed, 2.0 nm band width, 0.2 nm resolution, 50 mdeg sensitivity, and 4 sec response. The concentration of peptides was kept at 10×10^{-5} M and a 0.1 cm path-length quartz cuvette was used. Spectra were acquired in 10 mM phosphate buffer at pH 7.0.

Thermal denaturation profiles were obtained by measuring the temperature dependence of the ellipticity at 220 nm in the 20-90 °C range with a resolution of 0.5 °C and 1.0 nm bandwidth. The heating rate was 1°C/min and the response at 16 s with a Peltier temperature controller.

2.7. Cell culture and transfection

HeLa cells were grown in DMEM supplemented with 10% fetal calf serum (FCS) and 100 µg/mL penicillin/streptomycin, and maintained at 37°C with 5% CO_2. For plasmid transfection, cells were transfected by FUGENE HD reagent Roche Diagnostic GmbH, Mannheim, Germany) according to the manufacturer's instructions. Total DNA amounts were equalized by transfection of pRc/CMV empty vector (Invitrogen, Carlsbad, CA, USA).

2.8. *In vitro* translation

Wild type and mutant 3HA-IκB-α proteins were *in vitro* translated by using as templates pCDNA plasmids expressing the IκB-α sequence under T7 promoter in the TnT Quick Coupled Transcription/Translation Systems (Promega, Madison, WI, USA) according to the manufacturer's protocol. For *in vitro* translation the FLAG-Tat, the FLAG Tat sequence was PCR-amplified from p3xFLAG-CMV-Tat using the 5'T7-FLAG forward primer CGCCGGTAATACGACTCACTATAGGGACGCCA CCATGGACTACAAAGACCATGAC and the 3'TAT reverse primer GCTCTAGACTATTCCTTCGGGCCTGTCG; the PCR product was *in vitro* translated under the T7 promoter by using the TnT Quick Coupled Transcription/Translation Systems (Promega).

2.9. Immunoprecipitation and Western blotting

Wild type and mutant HA-IκB-α (20 µL) were incubated with or without FLAG-Tat (20 µL) in 500 µL of modified RIPA buffer [20 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% NP40, 1 mM PMSF, 1x Complete Protease Inhibitor (Roche Diagnostic GmbH)] in presence of protein G-Sepharose-coupled anti-FLAG antibody overnight at 4°C. Immuno-complexes were recovered by centrifugation at 200x *g* for 10 min at 4° C, extensively washed in modified RIPA buffer, resuspended

in SDS gel loading buffer, separated by 12% SDS-PAGE, and analyzed by Western blotting with anti-HA and anti-FLAG antibodies.

2.10. Luciferase assay

For testing Tat transcriptional activity, HeLa cells ($5x10^5$) were transfected with pNF-κB-deleted-LTR-Luc (0.5 μg), pSV-β-Gal (0.2 μg) with or without p3xFLAG-Tat (0.5 μg), pcDNA-3HA-IκB-α wild type (2 μg) or indicated mutants (2 μg).

For testing p65 transcriptional activity, HeLa cells ($5x10^5$) were transfected with pκB-Luc (0.5 μg), pSV-β-Gal (0.2 μg) with or without pRc/CMVp65 (0.5 μg), pcDNA-3HA-IκB-α wild type or mutants (2 μg). Forty eight h post-transfection, cells were lysed in lysis buffer (Dual Light Luciferase System, Tropix, Bedford, MA, USA), and the luciferase and β-galactosidase activities were evaluated by using Dual Light Luciferase System (Tropix) in a bioluminometer (Turner Biosystem, Sunnyvale, CA, USA). The ratio of firefly luciferase activity to β-galactosidase activity was expressed as relative light units.

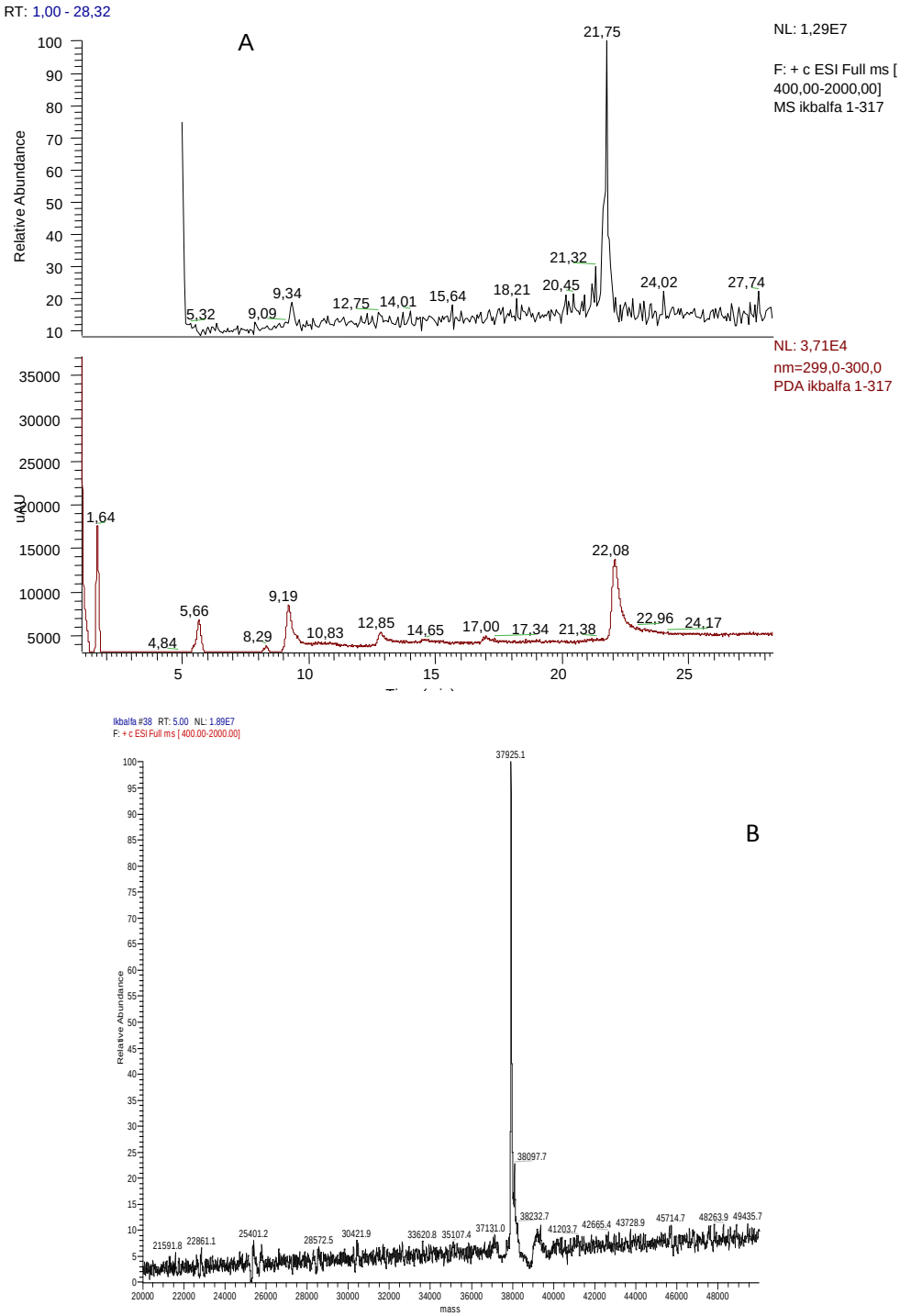
2.11. Statistical analysis

The data were reported as the mean ± S.E. using the two-tail unpaired Student's t test to assess the statistical significance of differences between means. Differences between the means were accepted as statistically significant at the 99% level (p <0.01).

3. RESULTS

3.1. SPR analysis of the Tat- IκBα protein complex

In order to quantify the entity of Tat-IκB-α binding, SPR experiments were carried out. Both proteins were preliminary obtained with a high degree of purity as suggested by LC-MS analysis (*Fig.* 3.1).



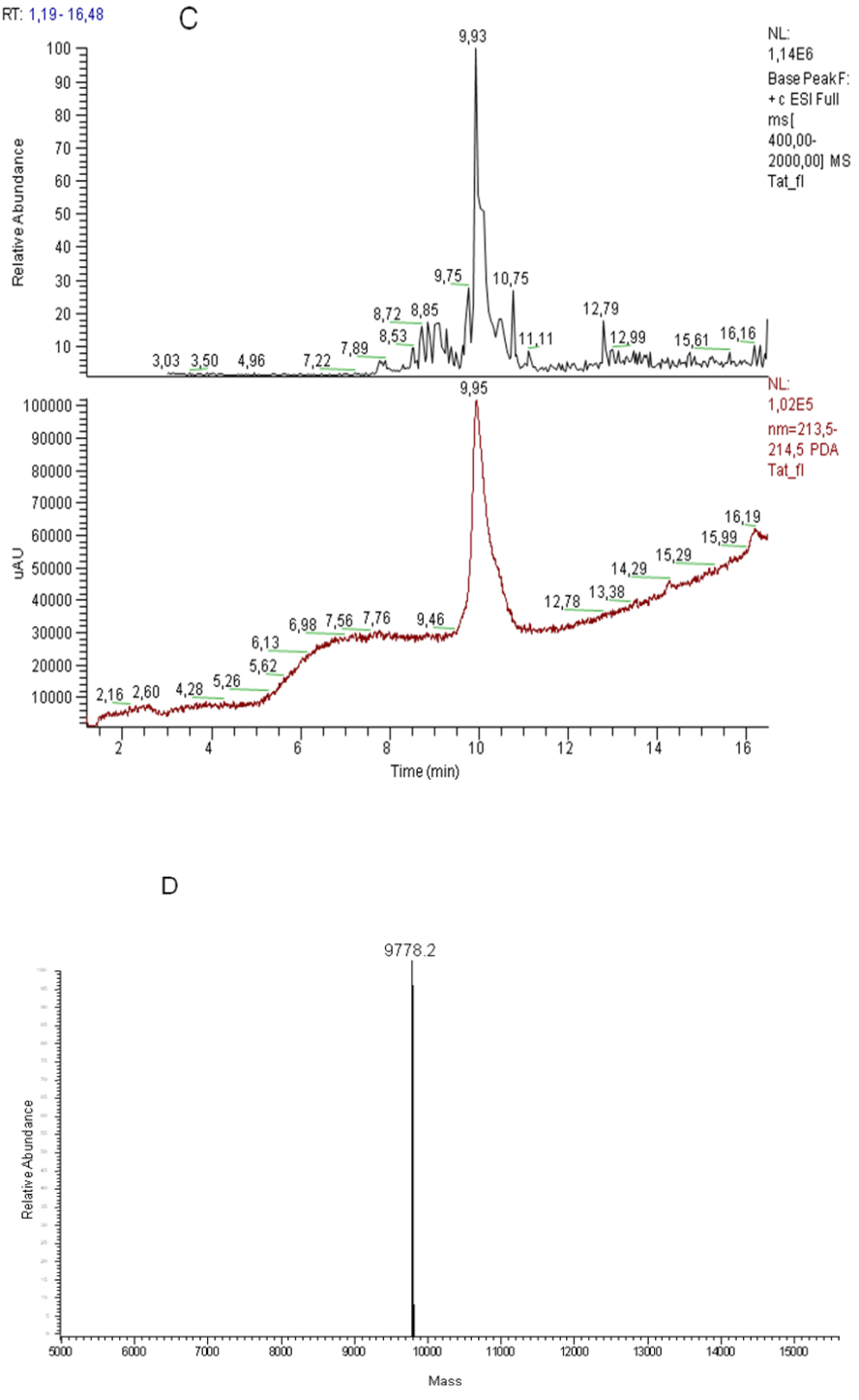


***Fig.* 3.1**: **MS spectra of IκB-α and HIV-Tat protein**. LC-MS analysis of purified (**A**) 6xHis-tagged IκB-α 1-317 and (**C**) HIV-Tat 1-86. Upper panels are UV chromatograms lower panels are ESI-mass profiles. Deconvoluted mass spectra of purified (**B**) IκB-α 1-317 (MWth/exp 37917/37920 Da) and (**D**) HIV-Tat 1-86 (MWth/exp 9777/9778 Da) proteins.

To this end, the full length His-IκB-α (1-317) was expressed in the BL21(DE3) bacterial strain and purified by affinity chromatography using nickel-agarose columns. HIV-1 Tat 1-86 was chemically synthesized following standard Fmoc protocols (48) and purified by RP-HPLC.

Tat protein was efficiently immobilized on CM5 (averaged immobilization level 4000 RU) and IκB-α was employed as analyte (concentration range 0-1000 nM). Dose-response sensorgrams revealed that IκB-α (1-317) avidly bound Tat; the association rate appeared to be rapid (ka= $7.39*10^4$ M^{-1} s^{-1}) whereas the dissociation was slow (kd=7.86 10^{-4} s^{-1}) thus providing a K_D value of 44.1±1.2 nM (*Fig.* 3.2 A).

The analysis of Tat sequence unveils the presence of two characteristic domains: the cysteine-rich motif (residues 20-39) comprising seven Cys residues present in the protein and a basic region (residues 48-60) containing several arginine and lysine residues (*Fig.* 3.3) (49)**.**

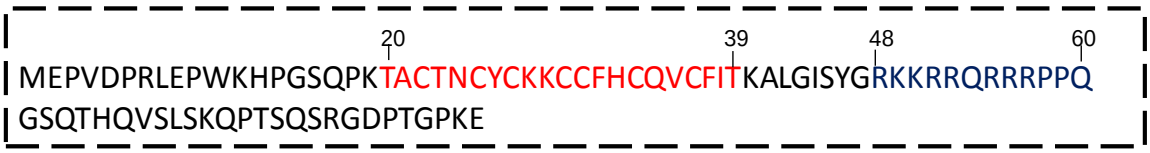


***Fig.* 3.3: HIV Tat BH10 protein,** in red is reported Cys-rich domain and in blue Arg-rich domain

In order to identify Tat regions involved in IκB-α recognition, we carried out competition experiments by using the peptides Tat_{20-39} and Tat_{48-60} corresponding to the cysteine-rich and arginine-rich region of Tat, respectively. The presence of pre-incubated Tat Cys domain with IκB-α (2:1 molar ratio, 160: 80 nM) did not affect the entity of IκB-α/Tat binding (*Fig.* 3.2 B). On the other hand, the presence of the peptide corresponding to the basic region (used in the same 2:1 ratio) significantly lowered the entity of IκBα binding to Tat (*Fig.* 3.2 C).

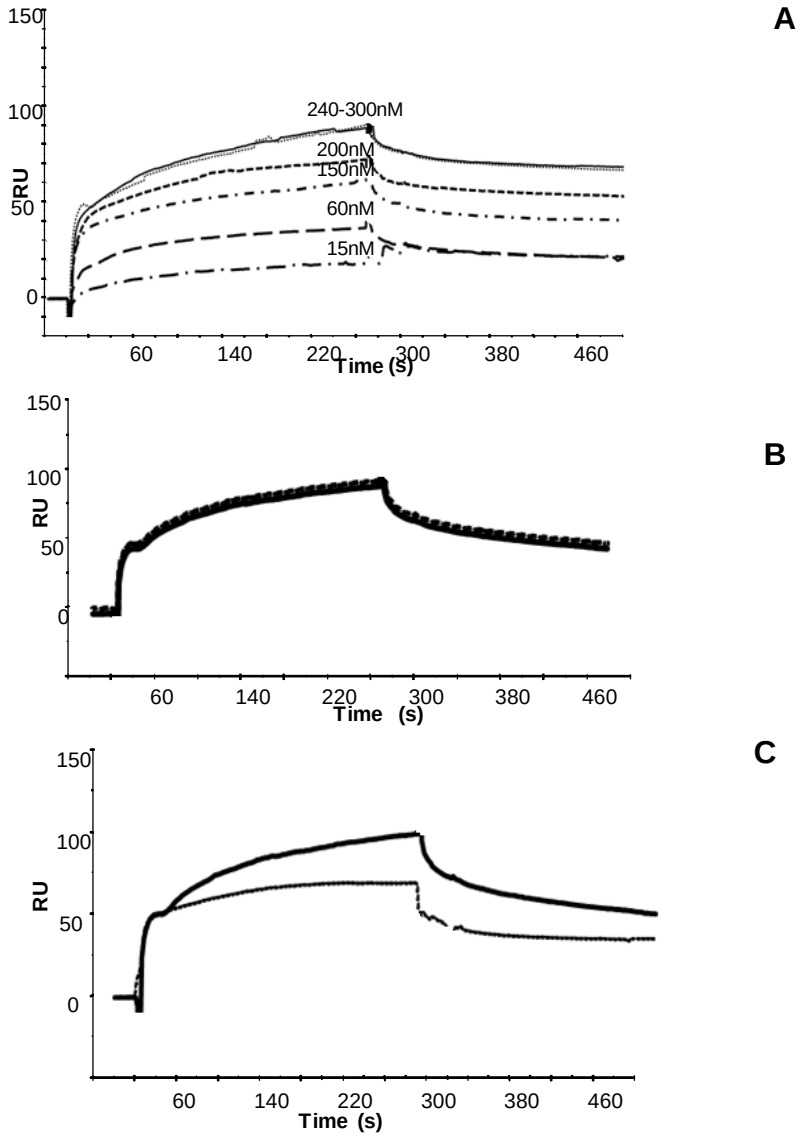


***Fig.* 3.2**: **SPR experiments of Tat-IκB-α.** Overlay of sensorgrams relative to the binding to immobilized Tat of (**A**) IκB-α protein, (concentration range 5-1000 nM); (**B**) IκB-α (black line) at 80 μM and a mix of IκB-α at 80 μM pre-incubated with Tat 20-39 at 160 μM (red line) (**C**) IκB-α (plain line) at 0.80 μM and a mix of IκB-α at 0.80 μM pre-incubated with Tat 40-68 at 1.60 μM (dotted line). Experiments were carried out in duplicate at a 25° C, at a constant flow rate of 20 μL/min using HBS 0.1mM TCEP, as running buffer.

3.2. Identification of IκB-α regions involved in Tat recognition through the characterization of IκB-α -based peptides

The identification of the IκB-α regions involved in Tat binding was achieved through the characterization of *ad-hoc* designed IκB-α peptides. Previous biochemical data suggested that a region close to the IκB-α C-terminus was involved in the binding to

Tat (40). To further characterize IκB-α fragments involved in the IκB-α /Tat complex, the C-terminal region of IκB-α (residues 243-317) was dissected in smaller fragments to unveil the role of specific determinants in the recognition process such as local structure and sequence and electrostatic charge. Although structural information on the free IκB-α protein is missing, clues on the conformation adopted by specific regions of this protein may be inferred from the structures of its complexes with NF-κB (29). In the structure of the IκB-α/p50/p65 complex, residues 243-280 of IκB-α constitute the sixth ankyrin repeat of the protein. The C-terminus of the AR6 region is followed by a flexible negatively charged motif denoted as PEST (residues 281-287). The very C-terminal end of the protein is composed by an additional negatively charged fragment (residues 288-302), and by the region 303-317. We designed peptides by dissecting and combining fragments of the AR6, the PEST region, and the C-terminus of the protein (*Fig.* 3.4 A-B).

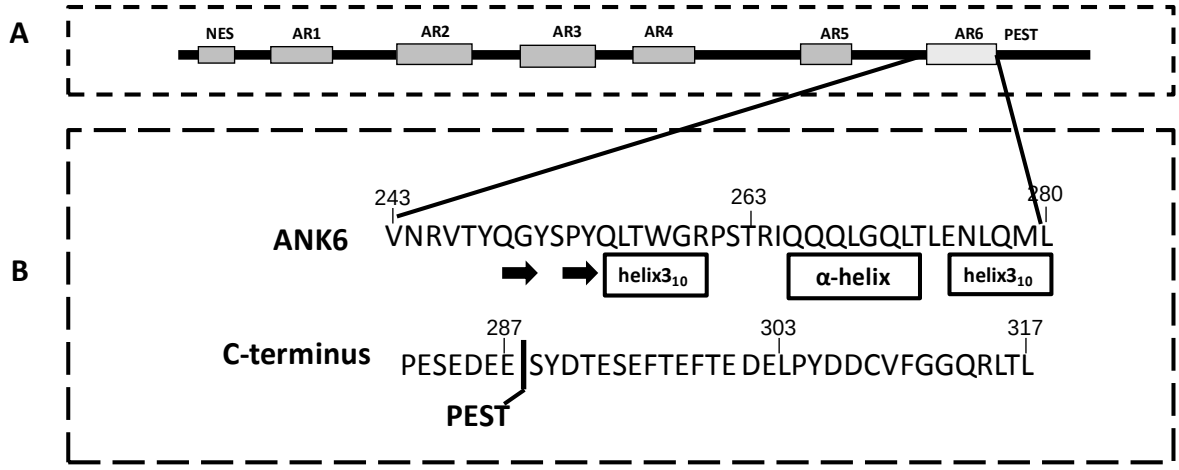


***Fig.* 3.4**. **HIV Tat and IκB-α proteins. (A)** Schematic representation of modular structure of IκB-α protein; Primary structures of (**B**) Ankyrin 6 and C-terminus of IκB-α protein (herein some C- and N-termini of IκB-α-based designed peptides are evidenced.

Indeed, AR6 was divided in two large peptides corresponding to the regions 243-263 ($Pep_{243-263}$) and 262-280 ($Pep_{262-280}$). In the structure of the IκB-α/NF-κB complex $Pep_{243-263}$ comprises a β-hairpin motif and a 3_{10} helix, whereas $Pep_{262-280}$ contains an α- and a 3_{10}- helix (*Fig.* 3.4 B and Table 3.1).

IkB-a peptide	Peptide sequence	SPR Binding to Tat, K_D
243-263 harpin helix 3_{10}	VNRVTYQGYSPYQLTWGRPST	+ 400 µM
262-287 alpha, 3_{10} helices and PEST domain	STRIQQQLGQLTLENLQMLPESEDEE	+ 200 nM
262-280 alpha, 3_{10} helices	STRIQQQLGQLTLENLQML	-
260-272alpha helix	RPSTRIQQQLGQL	-
269-287 3_{10} helix and PEST domain	LGQLTLENLQMLPESEDEE	-
281-302 negatively charged domain	PESEDEESYDTESEFTEFTEDE	-
303-317	LPYDDCVFGGQRLTL	-

Table 3.1: IκB-α-based peptides analyzed in this study

This latter region was further reduced by considering the peptide corresponding only to the α-helix of AR6 ($Pep_{260-272}$). The 262-280 region was also combined to the PEST motif in the peptide $Pep_{262-287}$. Moreover, the PEST region was linked to the 3_{10} helix located at the C-terminus of AR6 ($Pep_{269-287}$). Finally, the highly negatively charged peptide $Pep_{281-302}$ (total charge -12) and the C-terminal fragment $Pep_{303-317}$ were also considered.

All peptides were chemically synthesized in good yields by SPPS, using Fmoc methodologies as N-terminal acetylated and C-terminal amidated derivatives and purified by RP-HPLC; their identity and purity were assessed by LC-MS (data not shown). To evaluate the impact of intrinsic structural preferences of these fragments on their ability to bind Tat, all peptides were characterized by far-UV circular dichroism spectroscopy (CD). As expected, the peptides corresponding to the C-terminal end of the protein ($Pep_{281-302}$ and $Pep_{303-317}$) were unfolded (data not shown). Similarly, the CD spectra of $Pep_{243-263}$ and $Pep_{269-287}$ present a minimum at a wavelength lower than 200 nm (*Fig.* 3.5 A-B).

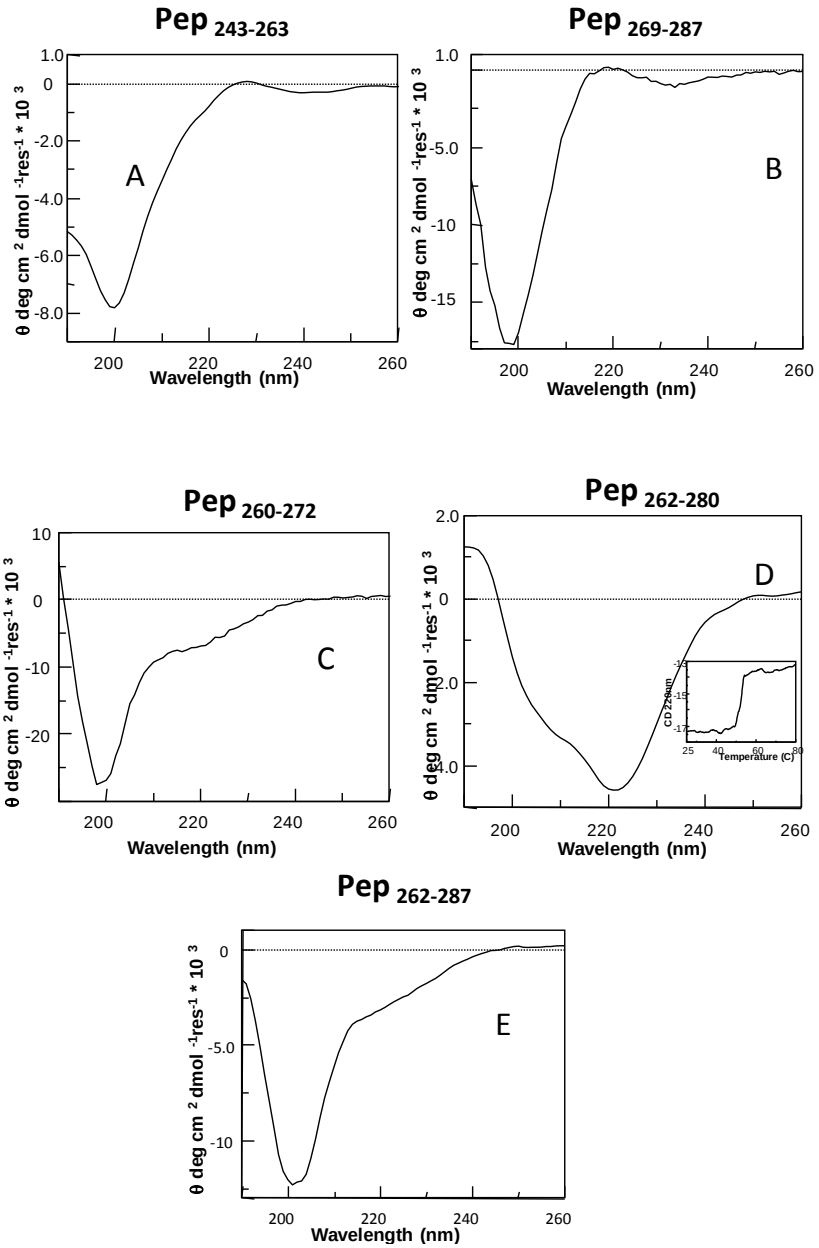


***Fig.* 3.5: CD analysis of IκB-α peptides.** CD spectra of the IκB-α synthetic peptides in phosphate buffer 10 mM, (**A**) $Pep_{243-263}$, (**B**) $Pep_{269-287}$ (**C**) $Pep_{260-272}$, (**D**) $Pep_{262-280}$ Inset:Thermal denaturation at 220 nm (**E**) $Pep_{262-287}$.

Although these peptides contain fragments that are structured in the IκB-α/NFκB complex (29), these findings indicate that both $Pep_{269-287}$ and $Pep_{243-260}$ are intrinsically unfolded. The spectrum of $Pep_{260-272}$, which is helical in the complex, presents a minimum at ~ 200 nm and a shoulder at ~ 220 nm (*Fig.* 3.5 C); these features are suggestive of mixture of conformational states, in which a predominant unfolded state coexists with a helical conformation. The addition of eight residues at the C-terminus of $Pep_{260-272}$ to generate the peptide $Pep_{262-280}$ produces a strong

negative CD signal at 220 nm (*Fig.* 3.5 D). The analysis of the thermal dependence of the CD signal at 220 nm of the peptide $Pep_{262-280}$ shows a clear cooperative transition at 52° C (inset of *Fig.* 3.5 D). Although the $Pep_{262-280}$ spectrum does not match canonical spectra corresponding to individual secondary structure elements, these findings indicate that $Pep_{262-280}$ is more structured than $Pep_{260-272}$. Interestingly, the addition of the PEST domain to $Pep_{262-280}$ to generate the peptide $Pep_{262-287}$ drastically affects the CD spectrum (*Fig.* 3.5 E). Indeed, the spectrum of $Pep_{262-287}$ becomes similar to that displayed by $Pep_{260-272}$. Altogether these findings provide clear evidence of the mutual structural influence of the different fragments of this IκB-α region.

3.3. Analysis of Tat binding to IκB-α-based peptides

The ability of IκB-α-based peptides to bind Tat was checked by using SPR direct binding assays. These experiments were carried out in the same way of those involving the full-length IκB-α protein by using Tat as ligand and the IκB-α-based peptides as analytes. Among the IκB-α-based peptides, only $Pep_{262-287}$ and $Pep_{243-263}$ showed detectable binding to Tat (Table 3.1). The overlay of the sensorgrams corresponding to the peptide $Pep_{262-287}$ (*Fig.* 3.6), showed a clear a dose-response binding in the concentration range 1-40 μM and the estimation of value K_D is 207 ±5 nM on the basis of the related kinetic data.

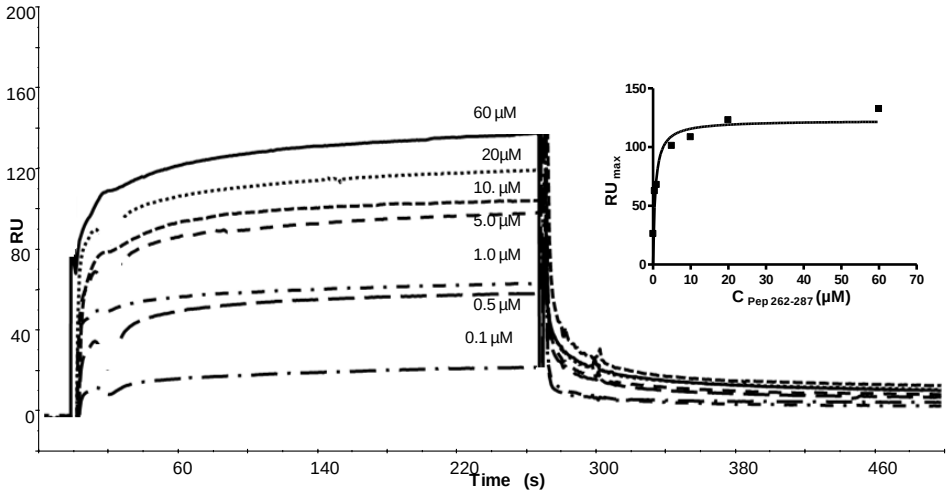


***Fig*. 3.6**: **SPR experiments of IκB-α peptides**. Overlay of sensorgrams relative to the direct binding of IκB-$α_{262-287}$ to immobilized Tat. Concentrations were between 1.0 and 40 μM. Experiments were carried out in duplicate, at a 25° C, at a constant flow rate of 20 μL/min using HBS-0.1 mM TCEP as running buffer.

The affinity of $Pep_{243-263}$ for Tat was much lower as dose-response experiments performed in a larger concentration range (10-700 μM) provided K_D value of 409 ±7 μM (*Fig.* 3.7).

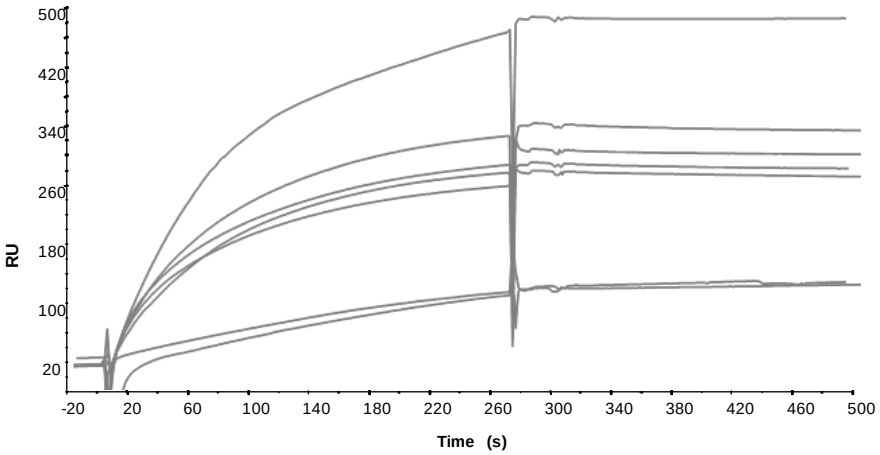


***Fig.* 3.7: SPR experiments of IκB-$α_{243-263}$ peptide**. Overlay of sensorgrams relative to the direct binding of IκB-$α_{243-263}$ to immobilized Tat. Concentrations were between 10 and 700 μM. Experiments were carried out in duplicate, at a 25° C, at a constant flow rate of 20 μL/min using HBS-0.1 mM TCEP as running buffer. Estimated K_D value = 409 ±7 μM.

The remarkable difference of magnitude of the K_D values displayed by the two peptides indicated that a region embedded in the fragment 262-287 was likely responsible for the binding of IκB-α to Tat. Notably, the reduction of the length of this region by cutting either its N-terminal or C-terminal ends abolished the binding to Tat, as shown by the inability of $Pep_{262-280}$, $Pep_{260-272}$, and $Pep_{269-287}$ to bind to protein (Table 3.1).

3.4. Characterization of IκB-α mutants in the region 260-269

We previously suggested that the IκB-α region encompassing the amino acids 262-269 could be important for the binding of this protein to Tat (40).

This observation has been partially corroborated by the SPR analysis on the IκB-α -based peptides. Although the isolated peptide $Pep_{260-272}$ is unable to bind Tat, the removal of the region 260-269 from the peptide $Pep_{262-287}$ abolishes the binding to

Tat as shown by the data obtained for $Pep_{269-287}$. In this framework, the role of specific fragments within the region 260-269 has been evaluated by using an Ala scanning approach. To further address the role of distinct amino acids within the region 260-269 in the binding to Tat, we generated the full-length mutants IκB-α $_{260}$(RPSTRIQQQL)$_{269}$/A, IκB-α $_{263}$(TRIQQQL)$_{269}$/A, IκB-α $_{266}$(QQQ)$_{268}$/A, where the indicated amino acids were substituted with alanines. These mutants were tested for their ability to bind to Tat by immunoprecipitation of *in vitro* translated proteins. Consistently with previous observation (40), the substitution of amino acids 260-269 or amino acids 263-269 abolished the binding of IκB-α to Tat (*Fig.* 3.8, lanes 4, 6). A significant reduction of the binding was observed for IκB-α $_{266}$(QQQ)$_{268}$/A, (*Fig.* 3.8, lane 8). These results indicated that the minimal sequence of IκB-α essentially required for Tat binding mapped within amino acids 263-269 with a partial contribution of the glutamine triplet at 266-268.

As further characterization, wild type and IκB-α mutants were analysed for their inhibition of Tat transcriptional activity. To this end, Jurkat cells were transfected with the expression vectors of the IκB-α genes together with Tat and the NF-κB-deleted Long Terminal Repeats (LTR) of HIV-1 fused to the *luciferase* gene. The use of the NF-κB-deleted LTR allowed determining the effect of IκB-α on the Tat trans-activating ability independently from its inhibition of NF-κB activity. As compared to wild type, IκB-α $_{260}$(RPSTRIQQQL)$_{269}$/A and IκB-α $_{263}$(TRIQQQL)$_{269}$/A completely loss the ability to repress the Tat trans-activation of the LTR, which corresponded to their inability to associate to Tat; a sligth but significant inhibition was exhibited by IκB-α $_{266}$(QQQ)$_{268}$/A, which retained a partial ability to bind Tat (*Fig.* 3.9 A).

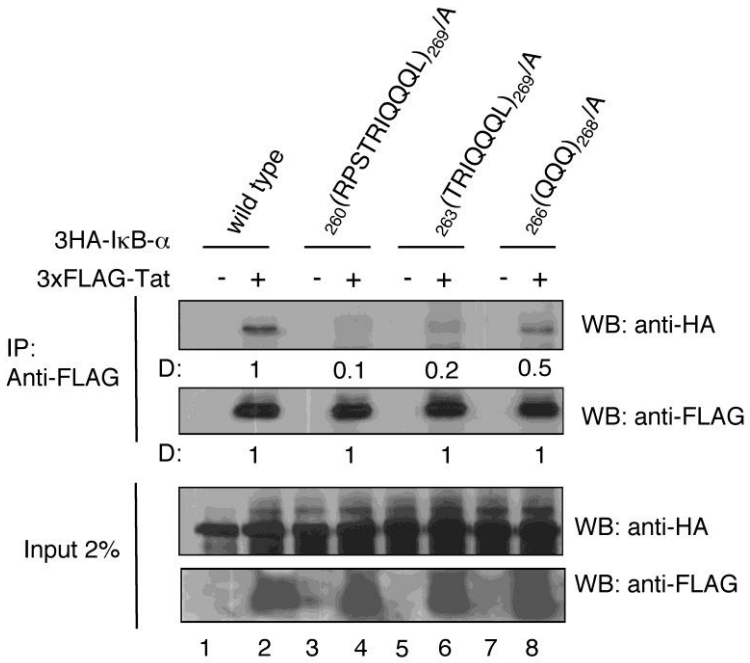


***Fig.* 3.8: Characterization of amino acid residues of IkB-a involved in the *in vivo* binding to Tat.** Wild type and mutant HA-IκB-α proteins were *in vitro* translated and equivalent aliquots were incubated with or without FLAG-Tat; protein complexes were immunoprecipitated with protein G-Sepharose-coupled anti-FLAG antibody. The immuno-complexes were separated on 12% SDS-PAGE and analyzed by Western blotting with anti-HA and anti-FLAG antibodies.

It is worthwhile to note that the base-pair mutations of IκB-α did not affect the inhibition of the p65/NF-κB transcriptional activity. In fact, all IκB-α mutants inhibited the p65-mediated transactivation of the NF-κB-dependent *luciferase* expression at the same extent as the wild type protein, suggesting that these mutations did not affect the global folding of IκB-α (*Fig.* 3.9 B). These results also demonstrate that the IκB-α region 263-269, which is required for the binding and inhibition of Tat, is not involved in the inhibition of p65 transcriptional activity.

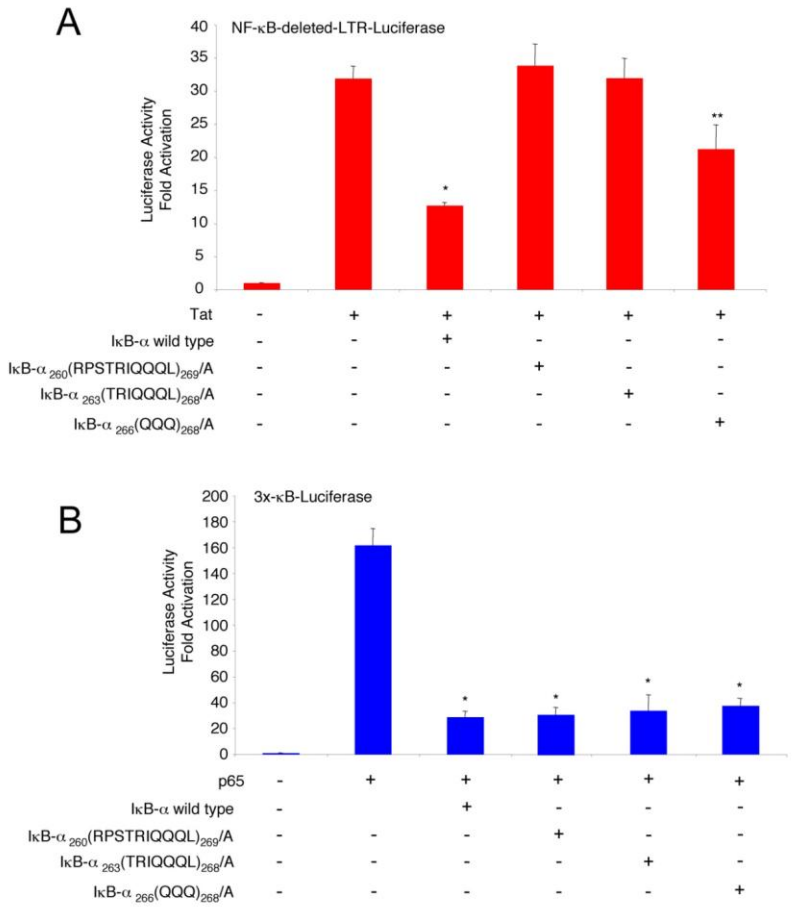


***Fig*. 3.9: Characterization of the amino acid residues of IκB-α involved in the inhibition of Tat or NF-κB transcriptional activities.** HeLa cells were transfected with (**A**) the reporter plasmid pNF-κB-deleted LTR-Luc with or without p3xFLAG-Tat, pcDNA-3HA-IκB-α wild type or indicated mutants; (**B**) the reporter plasmid pκB-Luc with or without pRc/CMV-p65, pcDNA-3HA-IκB-α wild type or indicated mutants. Forty-eight h post-transfection, the luciferase activity was measured in cell extracts. The asterisks indicate a statistically significant inhibition according to Student's t test (*, $p < 0.01$).

4. DISCUSSION

Protein-protein interactions play a fundamental role in the vast majority of physiological or pathological processes. An emerging paradigm of recent investigations is the involvement of the same protein in multiple and diversified biological partnerships. The identification and characterization of alternative partnerships has enormous implications for understanding the molecular mechanisms underlying pathological events, and for the development of new molecular entities able to regulate these processes (50).

In this framework, we have recently discovered that IκB-α, in addition to the known partners including p50/p65 NF-kB complex (29), p53 (51) and HDAC (52), also binds the HIV-1 Tat transcriptional activator (49). This interaction has a remarkable impact on the mechanism of HIV infection as shown by the high attenuation of viral growth upon expression of proteolysis-resistant IκB-α mutant (41,46). As this finding may be exploited for the design of new strategies for combating the widespread HIV-1 infection, here we have characterized in detail the Tat/IκB-α interaction. Binding experiments carried out by using SPR techniques indicates that the Tat and IκB-α tightly bind each other with a K_D constant in the nanomolar range. The entity of the binding affinity between the two proteins corroborates the proposal that the IκB-α–Tat binding plays a significant role in HIV-1 transcriptional regulation. Moreover, we characterized the regions of the two proteins directly involved in the IκB-α–Tat complex formation. To this aim two alternative approaches were adopted: (a) the dissection of the two proteins into smaller fragments, and (b) the design and the characterization of IκB-α mutants. The fragmentation of Tat clearly demonstrated that the arginine-rich region of the protein (residues 48 – 60) was directly involved in the protein-protein recognition; consistently with our previous observations (40). Although this result could suggest that electrostatic interactions could play a fundamental role in the binding process, the identification of the “hot spot” regions of IκB-α proved to be quite intricate.

Indeed, the peptide $Pep_{281-302}$ designed by considering essentially the negatively charged region 281-302 of IκB-α was unable to bind to Tat. This suggests that mere electrostatic interactions cannot account for the tight association of the two proteins. Along this line, no binding was also observed for the peptide $Pep_{269-287}$, which also contains the negatively charged PEST region. However, a remarkable binding was observed when the N-terminus of the peptide $Pep_{269-287}$ was extended up to residue

262 ($Pep_{262-287}$). This indicates that residues of the fragment 262-268, which assumes a helical conformation in the complex between IκB-α -NFκB (29), are important for the binding.

Noticeably, the context in which this fragment is embedded plays a role in Tat recognition. Indeed, the peptide $Pep_{260-272}$, which contains the minimal extension of this fragment, is unable to bind Tat. Collectively these data suggest that the PEST domain, although not directly involved in the binding, plays a role in selecting the conformation of the region 262-268 that is prone to bind Tat. This is not surprising since the CD characterization of the residual structure of isolated IκB-α fragments here reported indicates a mutual influence of the different regions.

In order to validate, these findings *in vivo* and to gain further insights into IκB-α-Tat recognition domain, three IκB-α mutants were designed and characterized. The attention was focused on the region 260-269, with a particular emphasis to the three consecutive Gln residues of the region 266-268, that could bind the positively charged arginine-rich Tat domain. The trans-activation data confirmed a critical role of this region in Tat recognition, which involves the three glutamine and few nearby residues. Interestingly, the three IκB-α mutants displayed an unaltered ability to inhibit p65. This is in line with the crystallographic data available for the IκB-α–p50/p65 complex and with the biochemical characterization of IκB-α–p50/p65 complex that identified two distinct “hot spot” regions of IκB-α for p65 binding helix 4 and residues 276-287 (29). Therefore, distinct IκB-α regions might mediate the interaction with the p50/p65 dimer or Tat.

In conclusion, the findings here reported provide a detailed characterization of IκB-α-Tat recognition. The identification of a peptide that present a good affinity for Tat represents a promising starting point for the development of molecules able to sequestrate Tat to inhibit its transcriptional activity, which could be tested as potential therapeutics.

REFERENCES

1. Weiss RA (May 1993). "How does HIV cause AIDS?". Science 260 (5112): 1273–9.
2. Douek DC, Roederer M, Koup RA (2009). "Emerging concepts in the immunopathogenesis of AIDS". Annu. Rev. Med. 60: 471–84.
3. "CDC – HIV/AIDS – Resources – HIV Prevention in the United States at a Critical Crossroads". Cdc.gov. Retrieved 2010-07-28.
4. "HIV and AIDS among Gay and Bisexual Men" (PDF). Retrieved 2010-07-28.
5. Joint United Nations Programme on HIV/AIDS (2006). "Overview of the global AIDS epidemic" (PDF). 2006 Report on the global AIDS epidemic. ISBN 9291734799. Retrieved 2006-06-08.
6. Joint United Nations Programme on HIV/AIDS (2010). "Overview of the global AIDS epidemic". UN report on the global AIDS epidemic 2010.
7. McGovern SL, Caselli E, Grigorieff N, Shoichet BK (2002). "A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening". J Med Chem 45 (8): 1712–22.
8. HIV Sequence Compendium 2008 Introduction. Retrieved 2009-03-31.
9. Chan, DC., Fass, D., Berger, JM., Kim, PS. (1997). "Core Structure of gp41 from the HIV Envelope Glycoprotein" (PDF). Cell 89 (2): 263–73.
10. Frankel AD and Young JAT. (1998). HIV-1: Fifteen Proteins and an RNA. Annu. Rev. Biochem. 67:1–25.
11. Chan D, Kim P (1998). "HIV entry and its inhibition". Cell 93 (5): 681–4.
12. Wyatt R, Sodroski J (1998). "The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens". Science 280 (5371): 1884–8.
13. Arthos J, et al. (2008) HIV-1 envelope protein binds to and signals through integrin alpha(4)beta(7), the gut mucosal homing receptor for peripheral T cells. Nat Immunol 9:301–309.
14. Zheng, Y. H., Lovsin, N. and Peterlin, B. M. (2005). "Newly identified host factors modulate HIV replication". Immunol. Lett. 97 (2): 225–34.
15. Hiscott J, Kwon H, Genin P. (2001). "Hostile takeovers: viral appropriation of the NF-kappaB pathway". J Clin Invest. 107 (2): 143–151.
16. Pollard, V. W. and Malim, M. H. (1998). "The HIV-1 Rev protein". Annu. Rev. Microbiol. 52: 491–532.
17. Ross EK, Buckler-White AJ, Rabson AB et al. Contribution of NF-kB and Sp1 binding motifs to the replicative capacity of human immunodeficiency virus type 1: Distinct patterns of viral growth are determined by T-cell types. J Virol 1991; 65:4350-4358.
18. Dayton AI, Sodroski JG, Rosen CA et al. (1986). The trans-activator gene of the human T cell lymphotropic virus type III is required for replication. Cell; 44:941-947.
19. Fisher AG, Feinberg MB, Josephs SF et al. (1986). The trans-activator gene of HTLV-III is essential for virus replication. Nature; 320:367-371.
20. Jeang K-T, Xiao H, and Rich EA. Multifaceted Activities of the HIV-1 Transactivator of Transcription, Tat. (1999) Journal of Biological Chemistry. Vol. 274, No. 41; 28837–28840.

21. Puglisi JD, Tan RY, Calnan BJ, Frankel AD, Williamson JR. (1992). Role of RNA structure in arginine recognition of TAR RNA. Science 257: 76–80.
22. Bayer P, Kraft M, Ejchart A, Westendorp M, Frank R, Rosch P. (1995). Structural studies of HIV-1 Tat protein. J. Mol. Biol. 247:529–35.
23. Wei P, Garber ME, Fang SM et al. (1998) A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA. Cell; 92:451-462.
24. Meyer BE, Malim MH. (1994). The HIV-1 Rev trans-activator shuttles between the nucleus and the cytoplasm. Genes Dev. 8:1538–47.
25. Bogerd HP, Fridell RA, Madore S, Cullen BR. (1995). Identification of a novel cellular cofactor for the Rev/Rex class of retroviral regulatory proteins. Cell 82:485–94.
26. Kjems J, Sharp PA. (1993). The basic domain of Rev from human immunodeficiency virus type 1 specifically blocks the entry of U4/U6.U5 small nuclear ribonucleoprotein in spliceosome assembly. J. Virol. 67: 4769–76.
27. Lee PP, Linial ML. (1994). Efficient particle formation can occur if the matrix domain of human immunodeficiency virus type 1 Gag is substituted by a myristylation signal. J. Virol. 68:6644–54
28. Gallay P, Swingler S, Song JP, Bushman F, Trono D. (1995). HIV nuclear import is governed by the phosphotyrosine-mediated binding of matrix to the core domain of integrase. Cell 83:569–76.
29. Huxford, T., Huang, D.B., Malek, S., and Ghosh, G. (1998). The crystal structure of the IkappaBalpha/NF-kappaB complex reveals mechanisms of NF-kappaB inactivation. Cell 95, 759-770.
30. Mosavi, L.K., Cammett, T.J., Desrosiers, D.C., and Peng, Z.Y. (2004). The ankyrin repeat as molecular architecture for protein recognition. Protein Sci 13, 1435-1448.
31. Li, J., Mahajan, A., and Tsai, M.D. (2006). Ankyrin repeat: a unique motif mediating protein-protein interactions. Biochemistry 45, 15168-15178.
32. Bergqvist, S., Ghosh, G., and Komives, E.A. (2008). The IkappaBalpha/NF-kappaB complex has two hot spots, one at either end of the interface. Protein Sci 17, 2051-2058.
33. Sue, S.C., Cervantes, C., Komives, E.A., and Dyson, H.J. (2008). Transfer of flexibility between ankyrin repeats in IkappaB* upon formation of the NF-kappaB complex. J Mol Biol 380, 917-931.
34. Mathes, E., O'Dea, E.L., Hoffmann, A., and Ghosh, G. (2008). NF-kappaB dictates the degradation pathway of IkappaBalpha. Embo J 27, 1357-1367.
35. Ferreiro, D.U., Cervantes, C.F., Truhlar, S.M., Cho, S.S., Wolynes, P.G., and Komives, E.A. (2007). Stabilizing IkappaBalpha by "consensus" design. J Mol Biol 365, 1201-1216.
36. Truhlar, S.M., Torpey, J.W., and Komives, E.A. (2006). Regions of IkappaBalpha that are critical for its inhibition of NF-kappaB.DNA interaction fold upon binding to NF-kappaB. Proc Natl Acad Sci U S A 103, 18951-18956.

37. Croy, C.H., Bergqvist, S., Huxford, T., Ghosh, G., and Komives, E.A. (2004). Biophysical characterization of the free IkappaBalpha ankyrin repeat domain in solution. Protein Sci 13, 1767-1777.
38. Bergqvist, S., Croy, C.H., Kjaergaard, M., Huxford, T., Ghosh, G., and Komives, E.A. (2006). Thermodynamics reveal that helix four in the NLS of NF-kappaB p65 anchors IkappaBalpha, forming a very stable complex. J Mol Biol 360, 421-434.
39. Bergqvist, S., Alverdi, V., Mengel, B., Hoffmann, A., Ghosh, G., and Komives, E.A. (2009). Kinetic enhancement of NF-kappaBxDNA dissociation by IkappaBalpha. Proc Natl Acad Sci U S A 106, 19328-19333.
40. Puca, A., Fiume, G., Palmieri, C., Trimboli, F., Olimpico, F., Scala, G., and Quinto, I. (2007). IkappaB-alpha represses the transcriptional activity of the HIV-1 Tat transactivator by promoting its nuclear export. J Biol Chem 282, 37146-37157.
41. Quinto, I., Mallardo, M., Baldassarre, F., Scala, G., Englund, G., and Jeang, K.T. (1999). Potent and stable attenuation of live-HIV-1 by gain of a proteolysis-resistant inhibitor of NF-kappaB (IkappaB-alphaS32/36A) and the implications for vaccine development. J Biol Chem 274, 17567-17572.
42. McElhinny, J. A., MacMorran, W. S., Bren, G. D., Ten, R. M., Israel, A., and Paya, C. V. (1995) Regulation of I kappa B alpha and p105 in monocytes and macrophages persistently infected with human immunodeficiency virus. J. Virol. 69, 1500–1509
43. Bossis, G., Salinas, S., Cartier, C., Devaux, C., and Briant, L. (2002) NF-kappaB activation upon interaction of HIV-1 envelope glycoproteins with cell surface CD4 involves IkappaB kinases. FEBS Lett. 516, 257–264
44. Hiscott, J., Kwon, H., and Genin, P. (2001) Hostile takeovers: viral appropriation of the NF-kappaB pathway. J. Clin. Investig. 107, 143–151.
45. Fields, G.B., and Noble, R.L. (1990). Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids. Int J Pept Protein Res 35, 161-214.
46. Quinto, I., Puca, A., Greenhouse, J., Silvera, P., Yalley-Ogunro, J., Lewis, M.G., Palmieri, C., Trimboli, F., Byrum, R., Adelsberger, J., Venzon, D., Chen, X., and Scala, G. (2004). High attenuation and immunogenicity of a simian immunodeficiency virus expressing a proteolysis-resistant inhibitor of NF-kappaB. J Biol Chem 279, 1720-1728.
47. Johnsson, B., Lofas, S., and Lindquist, G. (1991). Immobilization of proteins to a carboxymethyldextran-modified gold surface for biospecific interaction analysis in surface plasmon resonance sensors. Anal Biochem 198, 268-277.
48. Ruvo, M., Scarallo, A., Vecchio, G., Palombo, G., Fassina, G (1996). Synthesis and biological characterization of HIV-TAT [1-86] BH10 isola. Peptides, 771-772.
49. Brady, J., and Kashanchi, F. (2005). Tat gets the "green" light on transcription initiation. Retrovirology 2, 69.
50. Betzi, S., Restouin, A., Opi, S., Arold, S.T., Parrot, I., Guerlesquin, F., Morelli, X., and Collette, Y. (2007). Protein protein interaction inhibition (2P2I) combining high throughput and virtual screening: Application to the HIV-1 Nef protein. Proc Natl Acad Sci U S A 104, 19256-19261.

51. Chang, N.S. (2002). The non-ankyrin C terminus of Ikappa Balpha physically interacts with p53 in vivo and dissociates in response to apoptotic stress, hypoxia, DNA damage, and transforming growth factor-beta 1-mediated growth suppression. J Biol Chem 277, 10323-10331.
52. Aguilera, C., Hoya-Arias, R., Haegeman, G., Espinosa, L., and Bigas, A. (2004). Recruitment of IkappaBalpha to the hes1 promoter is associated with transcriptional repression. Proc Natl Acad Sci U S A 101, 16537-16542.

CHAPTER VI

1. INTRODUCTION

1.1. Protein splitting-Protein reassembly

Reassembly of proteins also known as ‘‘Protein fragment complementation’’ represent a structure based dissection of a protein through the expression or chemical synthesis of two or more fragments and the their re-assembling by means of non-covalent interactions. This approach have had several analytical and biochemical applications including study of protein–protein interactions, antibody screening, immunoassays and high-throughput screening.

After protein dissection, protein regions can reassemble spontaneously and/or assisted by the interaction with a host partner. Several proteins including ribonuclease, chymotrypsin inhibitor-2, t-RNA synthase and inteins undergo spontaneous reassembly from their peptide fragments to yield active protein (1–4).

However, formation of unordered aggregates of peptides is a major drawback of spontaneous reassembly of proteins (*Fig*. 1.1).

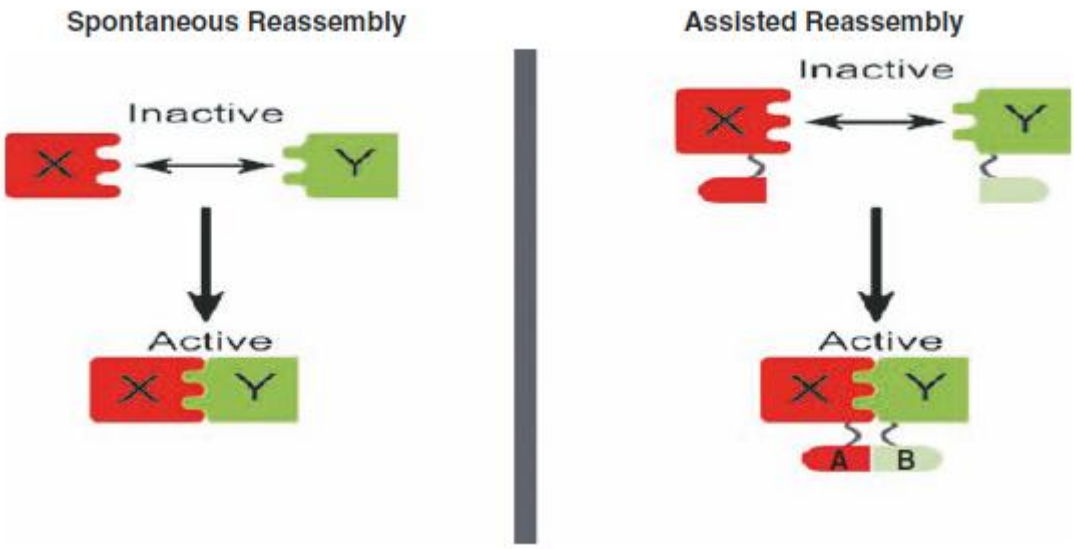


***Fig.* 1.1: Schematic representation of spontaneous and assisted reassembly where X and Y are peptide fragments of the reporter protein and A and B are the interacting domains**

To overcome this drawback, soluble oligomerization domains are attached to the peptide fragments, aiding the reassembly of proteins. This strategy has been successfully used in the reconstitution of several reporter proteins such as, dihydrofolate reductase, β-lactamase, aminoglycoside kinase, and green fluorescent protein (GFP) (5–10).

In this strategy, an enzyme or fluorescent protein (the reporter) is rationally dissected into two fragments and the fragments fused to two proteins that are thought to bind to each other. Folding of the reporter enzyme from its fragments is catalyzed by the binding of the test proteins to each other, and is detected as reconstitution of enzyme activity. The strategies described above refocus chemical genetics and therapeutic discovery from protein targeting to 'network-based targeting'. There are some considerable advantages of network targeting that are summarized in this points:

- Optimal target(s) for a network of interest can be identified in a simple cell-based assay.
- The development of this approach is simple and necessitate neither the production of recombinant proteins nor protein purification, as is required for the development of large-scale in vitro based assays of drug targets; often an intractable problem.
- This method enables the analysis of combinations of drugs that may impact multiple targets simultaneously. The 'drug cocktail' procedure is becoming particularly relevant in the oncology arena, where recent data demonstrated highly synergistic effects of kinase inhibitors (e.g. in combination with traditional cytotoxic chemotherapeutic agents) (11).

The most critical issue in the success of the reassembly is the site of the cleavage in the protein and the availability of the 3-D structure and circular permutation studies can aid in the design of the protein fragments. In planning successful protein reassembly strategies it is important to focus on small size proteins with monomeric structure and addressable in a biochemical assay. For overcoming a part of these problems, *Host-guest chemistry* could represent a valid strategy for the reassembly of two protein fragments.

1.2. Host-Guest Chemistry

Supramolecular chemistry approaches have generated a vast variety of multicomponent architectures by virtue of noncovalent interactions. Host-guest chemistry encompasses the idea of molecular recognition and interactions through non covalent bonding that is critical in maintaining the three-dimensional structure of large molecules such as proteins and is involved in many biological processes in which large molecules bind specifically but transiently to one another.

Insights gained from studies of noncovalent interactions have been of practical value in a wide range of applications including drug delivery, self-assembly and molecular recognition, chromatographic stationary phases, nanotechnology, sequestration of contaminants from solution, and the development of catalysts and chemical sensors. These applications require the availability of low-molecular-weight receptors (12,13) natural or non-natural oligomers and polymers,(14) or solid-state materials (15) that interact with their analytes in high affinity, highly selective binding processes.

Commons host molecules are macrocycles, as they provide cavities that can surround guest molecules and may be chemically modified to fine-tune their properties. Usually used macrocycles are (*Fig*.1.2):

- Cyclodextrins, calixarenes, cucurbiturils and crown ethers, that are readily synthesized in large quantities.
- Complex cyclophanes, and cryptands that can be synthesized to provide more tailored recognition properties.
- Metallacycles, that are macrocyclic aggregates with metal ions in the ring, often formed from angular and linear modules. Common metallacycle shapes as triangles, squares, and pentagons, as ligand field and bearing functional groups as coordinating moieties that "self-assembly around metal ion.
- Metallacrowns are metallamacrocycles, that are generated via a similar self-assembly approach from fused chelate-rings.

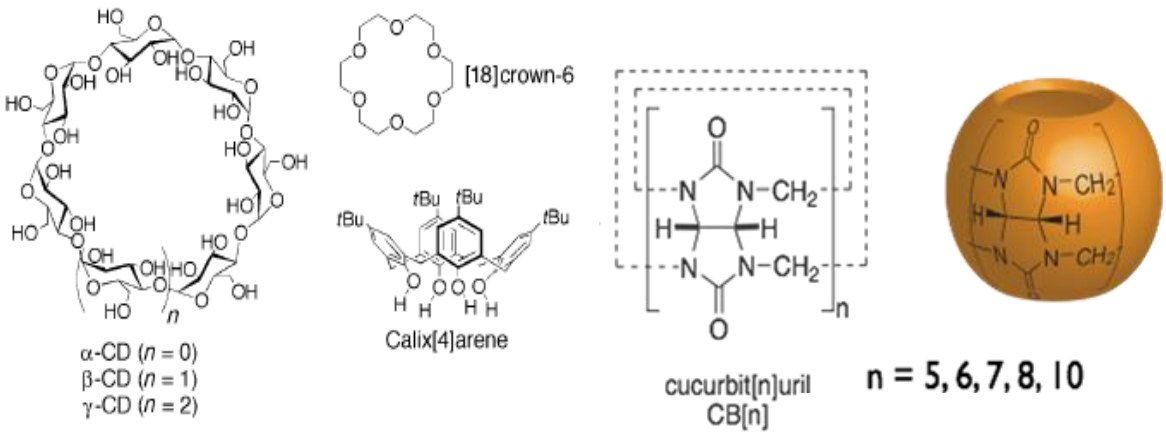


***Fig*. 1.2: Structural formulas of α-, β-, and γ-cyclodextrin, [18]crown-6, a calix[4]arene and cucurbit[n]uril**

1.2.1. The Cucurbit[n]uril Family

Although cyclodextrins (CDs) remain the macrocyle of choice for industrial applications on account of their commercial availability and low cost, several limitations arising from their low affinity and low selectivity exists. To this end, a new class of supramolecular host molecules, cucurbit[n]uril (CB[n]), has emerged as a versatile platform for fundamental and applied molecular recognition and self-assembly studies. All the members of the CB[n] family possesses a confluence of properties that suggest their high potential in nanotechnology as components of molecular machines.

These properties now include:

- commercial availability in four different sizes
- binding interactions of high affinity
- high selectivity of binding
- synthetic control over size, shape, and functional-group placement
- high structural integrity
- solubility in both organic and aqueous solution
- association and dissociation with controlled kinetics
- control of the molecular recognition processes by suitable electrochemical, photochemical, and chemical stimuli.

In 1981, during the reinvestigation of the condensation reaction of glycoluril and formaldehyde in concentrated HCl first reported by Behrend in 1905 (16), *(Fig*. 1.3) Mock et al. confirmed the remarkable macrocyclic structure consisting of six glycoluril units and twelve methylene bridges (CB[6]) and named it “cucurbituril” owing to its resemblance of a pumpkin, a prominent member of the cucurbitaceae family (17).

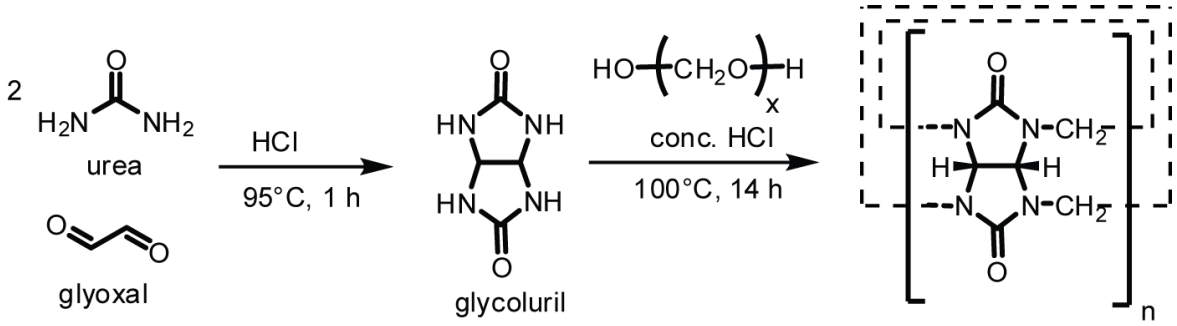
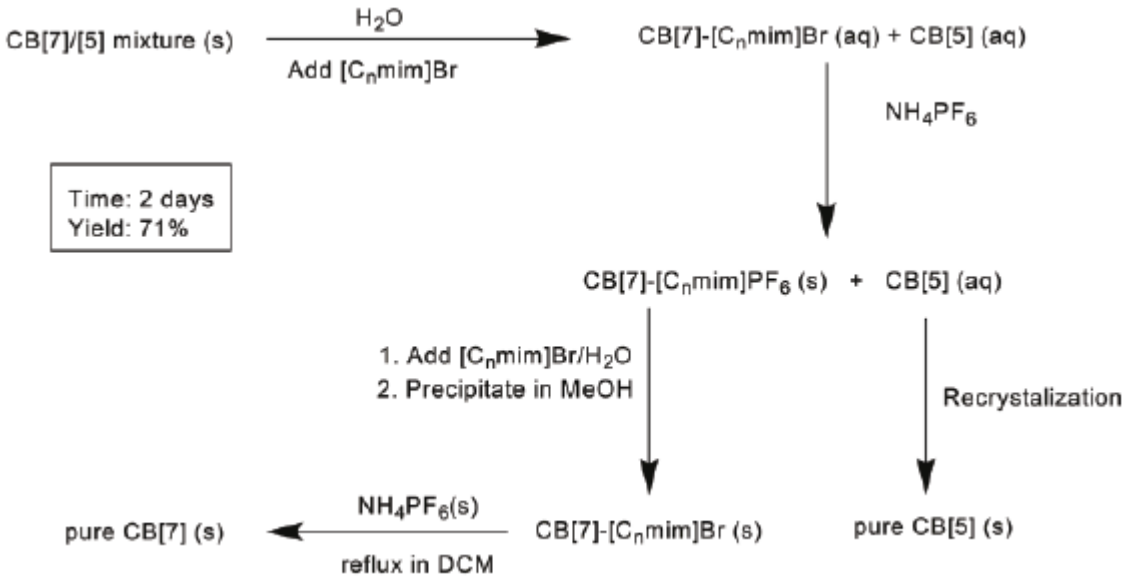


Fig. **1.3: Synthesis of cucurbit[n]uril**: from glycoluril under reported conditions: CH_2O, HCl, 100°C, 18 h.

Recently Scherman’s group reported a new “greener” method for the isolation of CB[5] and CB[7] based on ionic liquid binding and a solid state metathesis reaction (18). In this new strategy (Scheme 1.1), 1-ethyl-3-methylimidazolium bromide [C2mim]Br (Ka > 105 M^{-1} for CB[7]·[C2mim]Br,; it does not bind to CB[5] at all) (19) was added to an aqueous mixture of CB[5] and CB[7] followed by addition of NH_4PF_6 leading to the precipitation of the CB[7]·[C2mim]PF_6 complex. The CB[7]·[C2mim]PF_6 complex was isolated leaving an aqueous solution of CB[5] which was readily recrystalized to give pure CB[5]. The solid CB[7]·[C2mim]PF_6 complex was converted back to its bromide form and a solid state metathesis reaction using solid CB[7]·[C2mim]Br complex and solid NH_4PF_6 in dichloromethane yielded CB[7] in its pure form. The new technique not only provided a greener method for CB isolation, and high yield (71%), but also reduced the separation time from several weeks to a few days. A similar approach has also been applied to separate CB[6] and CB[8] (20).



Scheme 1.1: Isolation of CB[5] and CB[7] using ionic liquid and a solid state metathesis reaction

CB[n] are cyclic methylene-bridged glycoluril oligomers whose shape resembles a pumpkin. *Fig.* 1.4 shows the X-ray crystal structures for CB[5]–CB[8] and CB[5]-CB[10]. The cavity of CB[6] in the solid state contains three H-bonded H_2O molecules which can be released upon guest binding. The distinctive features of CB[5]–CB[10] are their two portals lined by ureido carbonyl groups that provide entry to their hydrophobic cavity (21). Similar to the cyclodextrins, the various CB[n] have a common depth (9.1 Å), but their equatorial widths, annular widths a, and volumes vary systematically with ring size (Table 1.1).

The portals guarding the entry to CB[n] are approximately 2 Å narrower than the cavity itself which results in constrictive binding that produces significant steric barriers to guest association and dissociation (22).

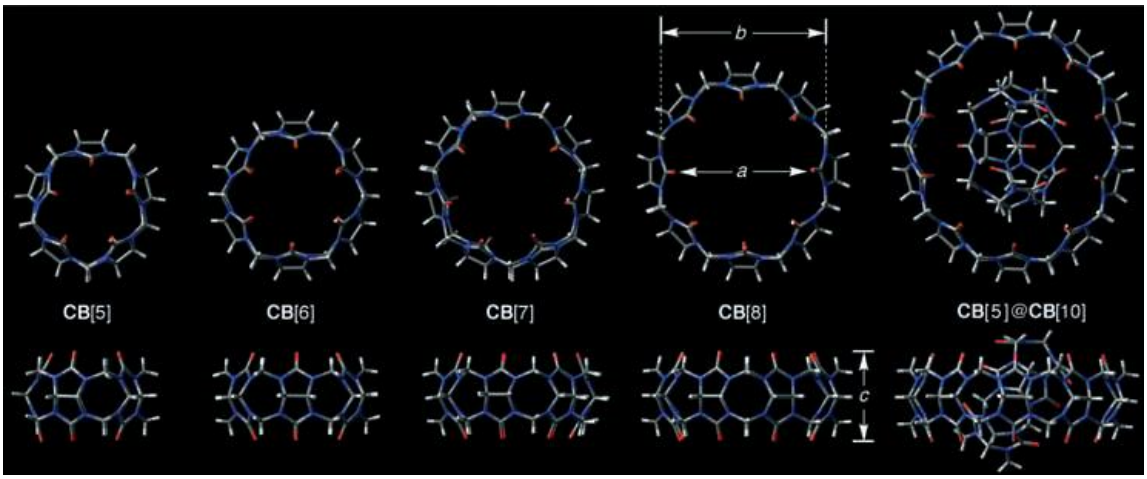


***Fig*. 1.4: Top and side views of the X-raycry stal structures of CB[5], CB[6], CB[7], CB[8], and CB[5]/CB[10].** The various compounds are drawn to scale.

One of the potential limitations of the CB[n] family is their relatively poor solubility in water: CB[6] and CB[8] are essentially insoluble, whereas CB[5] and CB[7] possess modest solubility in water (Table 1.1). The solubility of the CB[n] family is generally lower than the cyclodextrins. Like urea itself, however, the carbonyl groups lining the portals of CB[n] are weak bases: the pKa value of the conjugate acid of CB[6] is 3.02. Although the pKa values for CB[5], CB[7], and CB[8] have not been measured, they are likely to be similar to that of CB[6]. Accordingly, the solubility of CB[5]–CB[8] increase dramatically in concentrated aqueous acid.

Host	Portal Diameter (Å)	Interior Cavity Diameter (Å)	Height (Å)	Cavity Volume (Å^3)	Solubility in Water (mM)
CB[5]	2.4	4.4	9.1	82	20-30
CB[6]	3.9	5.8	9.1	164	0.018
CB[7]	5.4	7.3	9.1	279	20-30
CB[8]	6.9	8.8	9.1	479	<0.01
CB[10]	9.5-10.6	11.3-12.4	9.1	870	Not reported

Table 1.1: Dimensions and physical properties of CB[n]

Electrostatic effects play a crucial role in molecular recognition events in both aqueous and organic solution (23). *Fig*. 1.5 shows the electrostatic potentials of β-CD and CB[7]. Clearly, the electrostatic potential at the portals and within the cavity

of CB[7] is significantly more negative than for β-CD. This difference in electrostatic potential has significant consequences for their recognition behavior: CB[n] exhibit a pronounced preference to interact with cationic guests whereas β-CD prefers to bind to neutral or anionic guests.

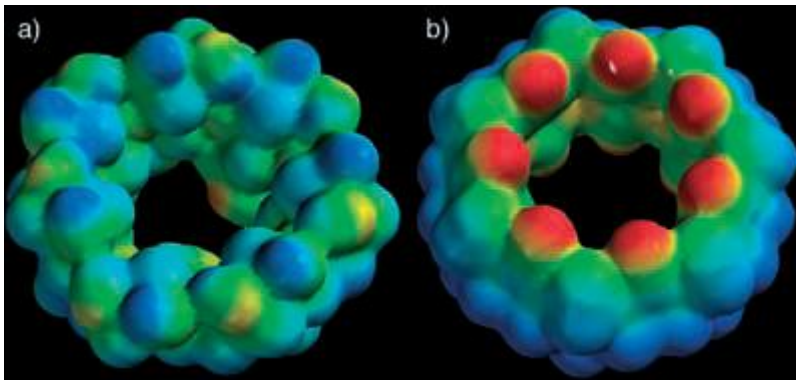


***Fig*. 1.5: Electrostatic potential maps for a) β-CD and b) CB[7].** The red to blue color range spans -80 to 40 kcal mol^{-1}

Owing to their rigid hydrophobic cavity and two hydrophilic portals, the CB[n] family emerged as an interesting and promising supramolecular host (24). A plethora of molecules have been studied as potential guests for these supramolecular hosts (25). The supramolecular chemistry of CB[5] is controlled by the narrow portals lined with carbonyl groups which provide entry to a cavity of low volume. Consequently, much of the supramolecular chemistry of CB[5] has been limited to the binding of protons as well as metal and ammonium ions at their portals. The pioneering work of Buschmann et al. (26) has established equilibrium constants and in many cases the enthalpic (ΔH) and entropic (ΔS) contributions to ΔG for the binding of CB[6] to ω-amino acids and ω-amino alcohols, aliphatic alcohols, acids, and nitriles, bipyridine derivatives, aromatic compounds, non-ionic surfactants and poly(ethylene glycols), cyclodextrins, diamides, and a-amino acids and dipeptides (26).

CB[7] is slightly more voluminous than β-CD and thus can bind a wider range of guests than CB[6] or CB[5]. CB[7] binds a variety of positively charged aromatic compounds. CB[7] also binds the metal complexes which suggests the use of CB[7] to reduce toxicity in cancer treatment. The cavity of CB[8] is similar in terms of volume to γ-CD, but is less conformationally flexible. CB[8] behaves like a larger

version of CB[5]–CB[7] in many ways, but also exhibits more complex recognition behavior. Just like CB[5]–CB[7], CB[8] prefers to bind to positively charged guests by ion–dipole interactions (27). CB[8] readily binds single guest molecules that partially or completely fill its cavity. Although the smaller homologues of the CB family can bind only one guest truly inside the cavity, CB[8] is unique in its ability to bind two guests as has been shown by Kim and others (28) CB[8] can serve to stabilize a charge transfer (CT) complex in its cavity. This result demonstrates there is cooperativity between the binding of the first and second aromatic rings. This recognition motif has been used to control intramolecular folding processes and the formation of vesicles (29). Kim and coworkers have also reported that CB[8]·Methylviologen (MV) can bind to a variety of other aromatic guests such as tryptophan, tyrosine, and dopamine (28). Urbach et al. demonstrated a systematic study on the thermodynamic parameters of amino acid binding to CB[8] (30). (*Fig.* 1.6)

guest	K_a $(M^{-1})^a$
Trp (1)	$4.3\ (\pm 0.3) \times 10^4$
Phe (2)	$5.3\ (\pm 0.7) \times 10^3$
Tyr (3)	$2.2\ (\pm 0.1) \times 10^3$
His (4)	no binding observed

***Fig.* 1.6: Equilibrium association constants for CB[8]-MV with aromatic amino acids**

1.3. Aim of the work

The goal of this project is a proof of concept study of protein reassembly upon designed dissection through non covalent interaction in CB[8] cavity.

In this direction the attention was focused toward Ubiquitin. It is a small regulatory protein that has been found in almost all tissues (ubiquitously) of eukaryotic organisms, it functions attaching to proteins and labelling them for their degradation. The ubiquitin-tag brings proteins to the proteasome, that is a large protein complex in the cell that has the prominent role of destroy and recycle unneeded proteins.

Ubiquitin has a small size with a globular three-dimensional structure in which secondary structures are well-localized in several regions. It was divided in two fragments, N-terminal and C-terminal region that were suitably labeled for their incapsulation in CB[8] cavity.

2. EXPERIMENTAL SECTION

2.1. Materials

O-(benzotriazol-1-yl)-N,N,N,N-tetramethyluronium-hexafluoro-phosphate (HBTU) was purchased from CEM. Amino acids, DMF and NMP were purchased from AGTC Bioproducts. NovaPEG Rink Amide LL resin was purchased from NovaBiochem. All other materials were purchased from Alfa Aesar and Sigma Aldrich and used as received unless stated otherwise.

2.2. Instrumentation

^{1}H NMR (400 MHz) spectra were recorded on a Bruker Avance QNP 400. HPLC was performed on a Varian 940LC. ESI-MS was performed on a Fischer Thermo Scientific LTQ Velos Ion Trap Mass Spectrometer with ESI source.

2.3. Synthesis of Fmoc-Asp(Boc-Nmec-Hmb)Gly-OH

2.3.1. Synthesis of Hmb-Gly-OBzl:

H-Gly-OBzl, free base, 3,1 gram (18,8 mmol) was dissolved in 15 ml ethanol and a solution of 2,57 gram (17 mmol) 2-hydroxy-4-methoxybenzaldehyde dissolved in 40 ml ethanol was added portion-wise. A strong yellow color was developed and after 5 minutes a thick precipitation was formed that no longer could be stirred with a magnetic stirring bar and the solution was therefore stirred manually for 25 minutes with a teflon rod. The solution was put on ice and 0,64 gram (17 mmol) solid $NaBH_4$ was added portion wise during 10 minutes and precipitated material started to disappear and a turbid white solution was formed. After 30 minutes the reaction was terminated by addition of 200 ml water and the pH adjusted to a stable pH of 3,0 with 1 M HCl solution. After acidification all solid material dissolved and a clear solution was obtained. The bulk of the ethanol was removed by reduced pressure, the solution diluted by 100 ml water and extracted once with diethyleter. The pH of the water phase was raised by addition of sodium hydroxide solution and gradually a precipitation was formed which dissolved upon addition of 100 ml DCM. The pH of the aqueous layer was adjusted to pH 8, the DCM phase isolated and the water phase extracted with 3x5 ml DCM. The pooled DCM phases were washed with 150 ml water and the water phase back-washed

with 3x5 ml DCM. The pooled DCM phases were washed with 2x with saturated sodium chloride solution, dried with Na_2SO_4 and evaporated to an oil that was that was dissolved in 75 ml ethylacetate and filtered through 2 cm packed uncoated TLC silica to remove unreacted H-Gly-OBzl and colored material. The solvent was evaporated and resulting oil was crystallized from EtOAc/petroleum ether. Yield 3,22 gram (62,9%).

2.3.2. Synthesis of Fmoc-Asp-HmbGly-OBzl:

1,87 gram (5,5 mmol) Fmoc-Asp-OH was added to 10 ml DCM and cooled to 0°C on ice and 0,694 gram. (5,5 mmol) diisopropylcarbodimide in 5 ml DCM was added portion-wise during 5 minutes. This solution was added to an ice cold solution of 0,755 gram (free base, 5 mmol) H-Hmb-Gly-OBzl in 5 ml DCM and the solution was stirred for 30 minutes on ice followed by 12 hours at room temperature. To the reaction mixture was added 0,2 gram ethylendiamine and the solution stirred for 5 minutes followed by addition of a mixture of 50ml DCM and 100 ml water containing 1% HCl. The solution was shaken and filtered through Celite to remove the remove precipitated diisopropylurea. The DCM phase was washed with 2x200 ml 0,1M HCl, and once with a saturated NaCl solution, dried with Na_2SO_4 and the solvent removed by reduced pressure. The resulting oil was dissolved in 20 ml EtOAc and petroleum ether was added until the solution became turbid (≈22ml). The solution was purified added to a dry flash chromatography column on to with a step-wise gradient of ethyl acetate in petroleum ether (20-40%). Rf (TLC; EtOAc/n-hexane 35/65, v/v) 0,34. Yield 2,95 gram (solid foam).

2.3.3. Synthesis of Fmoc-Asp-Hmb(BocNmec)Gly-OBzl:

2,85 gram Fmoc-Asp-HmbGly-OBzl (2) was added to 20 ml dry DCM. The solution was cooled to 0°C and 1,14 gram p-nitrophenylchloroformate was added. Diisopropylethylamine 0,65 gram in 5 ml DCM was added drop wise during 5 minutes. The solution was stirred for an additional 25 minutes on ice followed by addition of 1.72 gram mono-Boc-N,N'-dimetylethylendiamine and 1,18 gram diisopropylethylamine in 5 ml DCM. The solution was stirred for 1 hour on ice, transferred to a separation funnel and washed with 2x200 ml 0,1 % aqueous HCl, 2x200 ml 50 mM Na_2CO_3 and once with 200 ml water. All aqueous solution were back-washed three times with 5 ml DCM. The solution was dried with anhydrous Na_2SO_4 and the solvent removed by reduced pressure. The resulting oil was dissolved 20 ml EtOAc and 25 ml petroleum ether, applied to a dry flash

chromatography and the product eluted with a step-wise gradient of ethyl acetate in petroleum ether (30-50%). Rf (TLC; EtOAc/n-hexane 60/40, v/v), 0,37 and 0,48. No attempts were made to resolve these products with additional chromatography as the desired product was obtained with high purity after the next step. Yield 3,2 gram.

2.3.4. Synthesis of Fmoc-Asp-(Boc-Nmec-Hmb)Gly-OH:

3 grams Fmoc-Asp-Hmb(BocNmec)Gly-OBzl was dissolved in 50 ml EtOH, the solution flushed with nitrogen and 450 mg 10 % Pd(C) was added. The solution was bubbled with a gentle stream of hydrogen and the reaction monitored by TLC by withdrawing samples at regular intervals. After 40 minutes almost all starting material had reacted and fast moving spot that probably represented 9-methyl fluorene started to appear. The reaction mixture was flushed with nitrogen, filtered through Celite and the solvents removed by reduced pressure. The resulting solid was dissolved ethylacetate-peter 1:1 and purified with dry flash chromatography. The column was initially eluted with EtOAc:petroleumether; 1:1 to remove traces of starting material followed EtOAc-petroleum ether containing 2% acetic acid. The fractions containing the product were pooled and washed eight times with water and the pooled water solutions extracted with EtOAc and washed eight times with water. The EtOAc solution was dried and evaporated to an oil that still contained traces of acetic acid. The product was dissolved in diethyleter and washed six times with water and this procedure was repeated for the pooled water solutions. The diethyleter phases were pooled, dried with Na_2SO_4 and evaporated to a solid foam that was further dried at high vacuum overnight. Rf (TLC; EtOAc/n-hexane/HOAc 78/20/2, v/v/v) Yield 2,43 gram (3,25 mmol).

2.3.5. Synthesis of mono-Boc-N,N'-dimethylethylendiamine:

90 mmol (7,93 gram) N,N'-dimetylethylendiamine was dissolved in 300 ml ethylacetate and a solution of 30 mmol (6,5 gram) di-tert-butyl dicarbonate dissolved in 50 ml ethylacetate was added portion-wise. When about half of the solution had been added the solution became turbid. The solution was stirred for 1 hour and transferred to a separation funnel and shaken with 2x20 ml water and 1x50 saturated sodium chloride solution. The solution was dried with anhydrous Na_2SO_4 and the solvent removed by reduced pressure. The resulting oil was dissolved 50 ml chloroform, evaporated, redissolved in 50 ml

dietylether and evaporated to remove residual traces of N,N'-dimetylethylendiamine (bp 119°C). Residual solvents were removed under high vacuum for 1 hour. Yield 4, 9 gram (oil)

2.4. Synthesis of C-terminal UBIQUITIN (residues 47-76)

Solid phase peptide synthesis was performed on a Liberty Microwave Peptide Synthesizer (CEM Corporation) using Fmoc-strategy. The Fmoc group was removed in each step with 20 % (v/v) piperidine in DMF with 0.1M HOBt for 3 min using a microwave power of 45 W. The maximum temperature was set to 75°C. The coupling step was performed with 5 equiv. of Fmoc protected amino acid in DMF (0.2 M), 4.5 equiv. of DIC in DMF (0.45 M) and 10 equiv. of DIPEA in NMP (2 M). All couplings were performed for 10 min at 25 W at a maximum temperature of 75° C. Following completion of the sequence the peptide was released from the resin with concomitant removal of side chain protecting groups by treatment with TFA//H_2O/TIS (95/2.5/2.5, v/v/v) at room temperature for 2 h. The crude peptide was precipitated in cold Et_2O. The crude product was purified on a Varian 940-LC using a preparative Varian Polaris C8 column applying a linear gradient of 10 % to 60 % B in 30 min. The mobile phases were 0.1 % TFA in H_2O (eluent A) and 0.1 % TFA in Acetonitrile (eluent B).

2.5. Synthesis of N-terminal UBIQUITIN (residues 1-46)

Solid phase peptide synthesis was performed on a Liberty Microwave Peptide Synthesizer (CEM Corporation) using Fmoc-strategy. The Fmoc group was removed in each step with 5 % (v/v) piperazine in DMF with 0.1M HOBt for 3 min using a microwave power of 45 W. The maximum temperature was set to 75°C. The coupling step was performed with 5 equiv. of Fmoc protected amino acid in DMF (0.2 M), 4.5 equiv. of DIC in DMF (0.45 M) and 10 equiv. of DIPEA in NMP (2 M). All couplings were performed for 10 min at 25 W at a maximum temperature of 75° C. Following completion of the sequence the peptide was released from the resin with concomitant removal of side chain protecting groups by treatment with TFA//H_2O/TIS (95/2.5/2.5, v/v/v) at room temperature for 2 h. The crude peptide was precipitated in cold Et_2O.

2.6. Synthesis of MV-C2-Benz-Mal

A solution of p-Maleimidobenzoyl azide (0.10 g, 0.41 mmol) in 10 mL anhydrous DMF was heated to 145°C for 2 h under nitrogen atmosphere in a 50 mL two-neck RBF, equipped with a reflux condenser. After cooling to room temperature MV-C_2H_4OH $2PF_6$ (0.16 g, 0.32 mmol) was added and the solution was stirred over night. The solution was concentrated to a few millilitres in vacuo and diethylether was added. The precipitate was collected by suction filtration, redissolved in 10 mL of acetone and filtered. A solution of tetrabutylammonium chloride (0.33 g, 1.2 mmol) in 5 mL of acetone was added and the solution was allowed to stand overnight in the fridge. The resulting precipitate was collected by suction filtration, washed with cold acetone, to yield the title compound (0.10 g, 60%) as a yellowish solid.

2.7. Synthesis of MV- C2-Benz-Mal-N UBQ

To a suspension of MV-C2-Benz-Mal (10 mg, 0.012 mmol) in 100 mL of 100 mM phosphate buffer (pH = 7) was added N-terminal fragment Ubiquitin (30 mg, 0.0026 mmol) and the mixture was stirred at room temperature for 24 hours. $Th_{e\ s}olutio_n$ was concentrated by lyophilisation and ready for purification.

3. RESULTS

3.1. Design of Dissection of UBIQUITIN.

Ubiquitin, a small protein (8.5 kDa) present in all eukaryotic cells. It contains a highly conserved sequence of 76 amino acids that is identical in a wide variety of sources including humans, fish, and insects. It participates in diverse cellular functions, such as protein degradation, chromatin structure, and heat shock, by conjugation to other proteins through the carboxy terminus. This protein is highly conserved in evolution: yeast and human ubiquitin differ at only 3 of 76 residues. The carboxyl-terminal glycine of ubiquitin becomes covalently attached to the amino group of lysine residues of proteins destined to be degraded. The human ubiquitin primary sequence is as follows:

^{1}MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTL SDYNIQKESTLHLVLRLRGG^{76}

Its crystal structure has been refined at 1.8 Å resolution (31) and the overall structure of ubiquitin is extremely compact and tightly hydrogen-bonded; approximately 87% of the polypeptide chain is involved in hydrogen-bonded secondary structure. Prominent secondary structural features include three and one-half turns of α-helix, a short piece of 3_{10}-helix, a mixed β-sheet that contains five strands, and seven reverse turns. There is a marked hydrophobic core formed between the β-sheet and α-helix. The protein features a number of unusual secondary structural features, including a parallel Gl β-bulge, two reverse Asx turns and a symmetrical hydrogen-bonding region that involves the two helices and two of the reverse turns (*Fig.* 3.1).

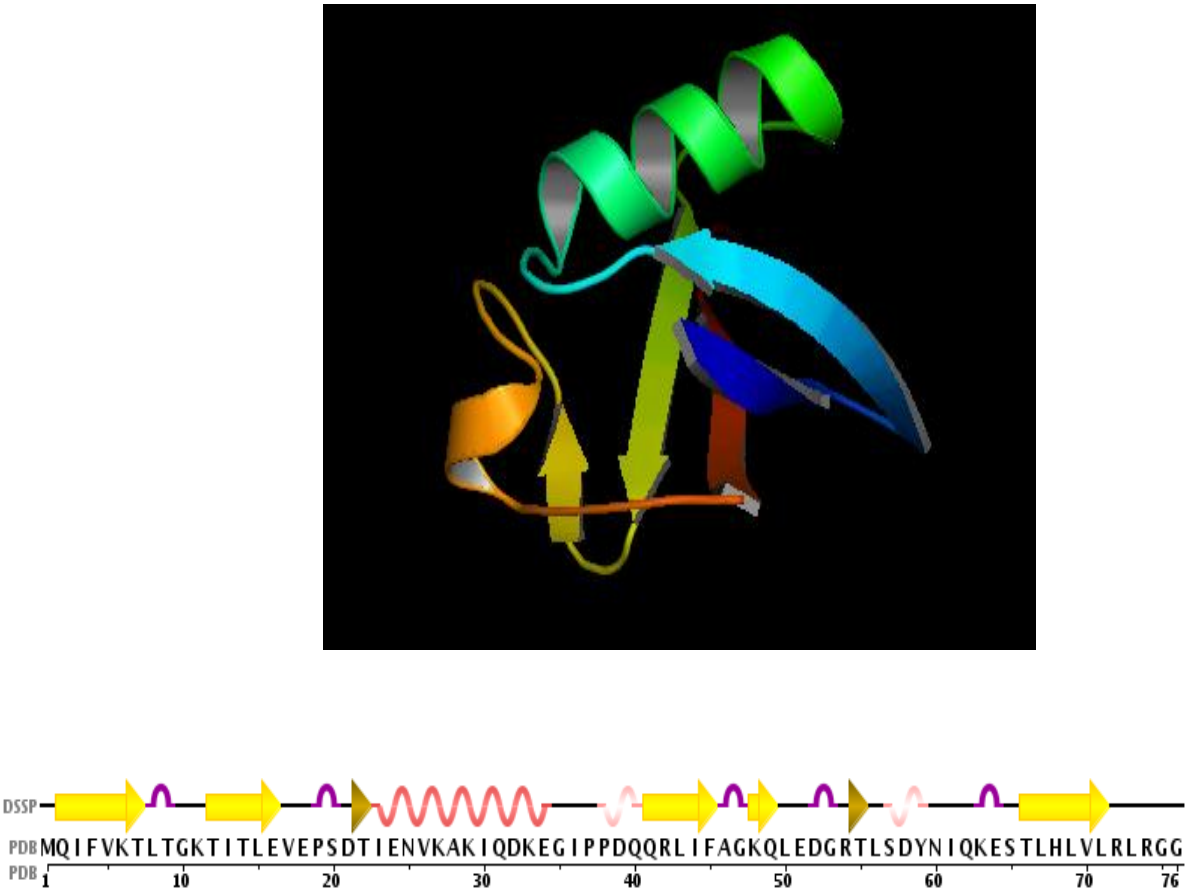


***Fig.* 3.1: Structural representation of Ubiquitin protein (upper) and schematic representation of secondary structural regions (lower)**

In protein dissection strategy the most critical issue is represented by the choice of the site of the cleavage in the protein, and by the visual inspection of the protein structure, a region of surface loop between residues 46 (alanine) and 47 (glycine) was chosen as splitting site (*Fig*. 3.2).

[1]MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTL SDYNIQKESTLHLVLRLRGG[76]

***Fig*. 3.2: Primary sequence of Ubiquitin.** The site of the cleavage is represented in red.

The dissection resulted in N- and C-terminal fragments, designated N-UBQ (spanning residues 1-46) and C-UBQ (spanning residues 47-76). At the C-terminal of the N-UBQ fragment a residue of cysteine was inserted to allow the coupling of methylviologen (*Fig*. 3.3).

N-UBQ : ^{1}**MQIFVKTLTGKTITLEVEPSDTIENVKAIQDKEGIPPDQQRLIFA**^{46}**-Cys**

C-UBQ : **Trp-Gly-**^{47}**GKQLEDGRTLSDYNIQKESTLHLVLRLRGG**^{76}

***Fig*. 3.3: Sequence of splitted Ubiquitin fragments**

On the other hand, in the C-UBQ fragment a dipeptide was inserted at the N-terminal: WG, tryptophan residue was introduced to allow the formation of aromatic interaction inside the cavity of CB[8], while glycine was inserted as spacer (*Fig.* 3.3).

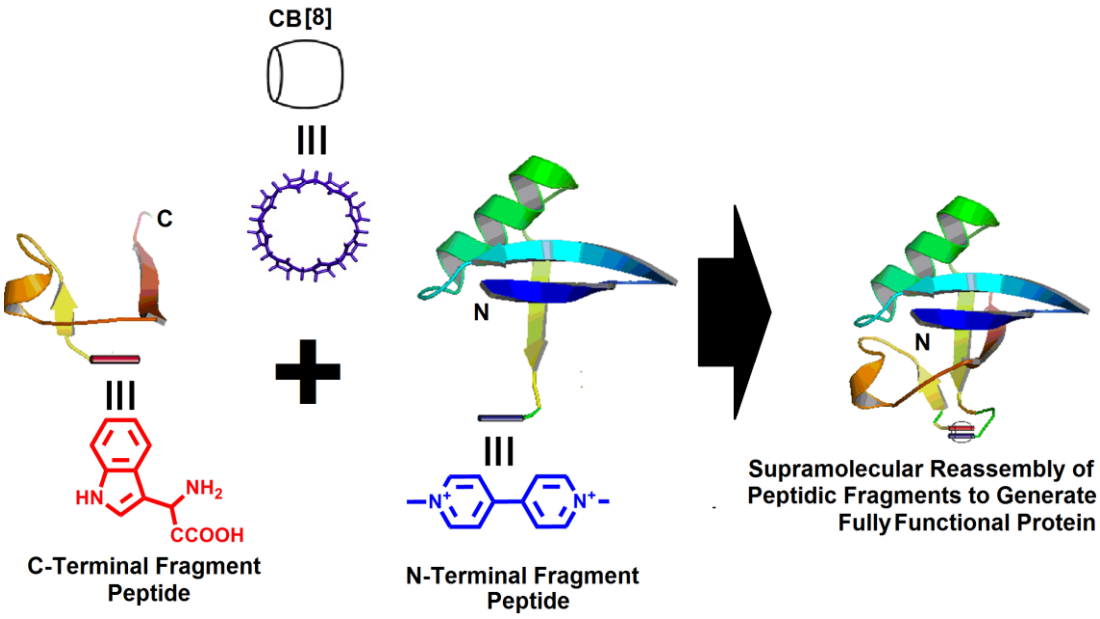


***Fig*. 3.4: Schematic representation of reassembly of fragments with CB[8]**

Designed fragments should self-assemble in the presence of CB[8] to refold native ubiquitin. Indeed CB[8] is able to bind two aromatic guests in this case, methylviologen and tryptophan, and to stabilize the complex in its cavity through electronic charge transfer (CT) mechanism, as described in *Fig.* 3.4.

3.2. Synthesis of C-UBQ fragment

Aggregation of the peptide chain during peptide synthesis is a major problem in both Boc and Fmoc chemistries. It results in incomplete acylation of the growing peptide chain and is the major source of deletion peptides. Incomplete acylation is usually caused by β-sheet formation, which has been studied extensively. By exploring the secondary structure of the C-UBQ fragment, it is evident that two consistent β-strands are present in this region. Furthermore the presences of many hydrophobic residues make the peptide chain very aggregating in the assembly phase and presumably insoluble after the cleavage. HPLC purification of poorly soluble peptides is a major practical problem resulting in low loading of the peptide onto the column, poor resolution and low yields. In addition, hydrophobic and aggregated peptides are often hard to characterize by mass spectrometry because of poor ionization ability.

Moreover, exploring the primary sequence of the C-UBQ fragment, the Asp-Gly sequence is present: this is particularly prone to “Aspartimide formation”, that is one of the most frequently encountered side reactions affecting Fmoc SPPS, which results from a ring-closure between the β-carboxy side-chain of aspartic acid and the nitrogen of the α-carboxamide.

Such aspartimides are very susceptible to base-catalyzed epimerization (32) and readily undergo ring-opening reactions giving mixtures of often inseparable a-, b-aspartyl peptides and α- and β-piperidides (*Fig.* 3.5) (33).

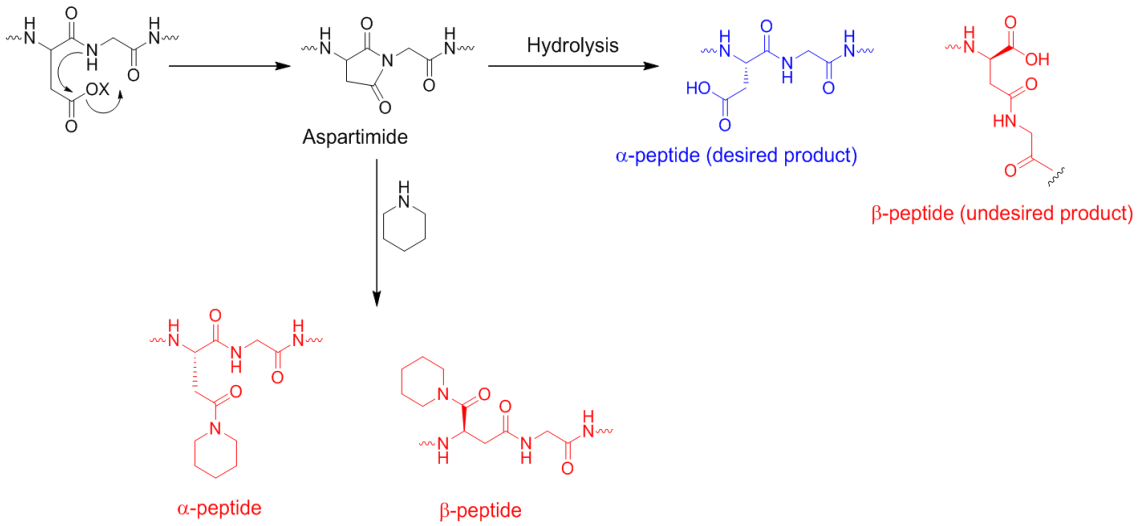


***Fig*. 3.5: Mechanism of aspartimide formation**

An efficient general strategy to counteract β-sheet formation and aspartimide formation during the synthesis of a peptide is reversible modification of selected peptide bonds. Reversible protection of the peptide backbone using modified N-(2-hydroxy-4-methoxybenzyl) (Hmb) groups was introduced by Wahlstrom et al. (34). In particular the phenolic oxygen in Hmb is protected with a Boc-N-methyl-N-[2-(methylamino)ethyl] carbamoyl group (Boc-Nmec). Thus it is possible to introduce modified Hmb groups into a peptide chain as dipeptides. (*Fig.* 3.6)

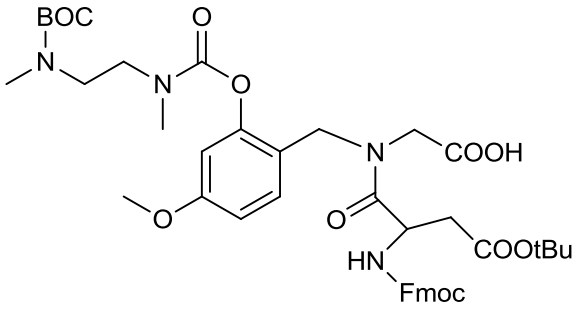


***Fig*. 3.6: Structure of the dipeptide Fmoc-Asp-(Boc-NmecHmb)Gly-OH**

During synthesis of the peptide, Hmb-protected peptides attached to a solid phase are less prone to β-sheet and aspartimide formation, and moreover coupling of the remaining residues usually proceeds with significantly enhanced efficiency.

During cleavage of the peptide from the resin with TFA, the Boc group was removed but the protonated methyl-(2-methylamino-ethyl) carbamoyl group remained attached to the Hmb group **3**. In this form, peptide have a greatly increased solubility in aqueous solutions as a result of both the β-sheet-disrupting effect of the benzylated peptide bond, and the solubilizing effect of the cationic charge. After purification, the peptide was exposed to slightly alkaline reaction conditions during which an intramolecular cyclization reaction produced the Hmb-protected peptide **4** and *N,N'*-dimethylimidazoldione (**5**). The Hmb group can be removed with TFA. In the fully deprotected form, the peptide **6** could be isolated either by an additional chromatographic step or simply by precipitation (*Fig*. 3.7).

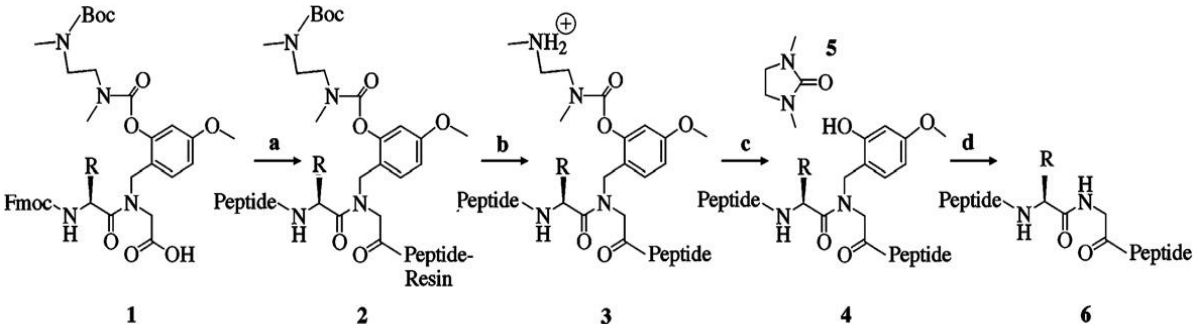


***Fig*. 3.7: General scheme for the use of Fmoc dipeptides with peptide bonds protected with modified 2-hydroxy-4-methoxybenzyl groups, R=CH(CH_3)$_2$.**
(a) Peptide synthesis; (b) TFA, RP-HPLC purification; (c) DMF/water, N-methylmorpholine; (d) TFA, diisopropylsilane

We therefore synthesized Fmoc-Asp-(Boc-NmecHmb)Gly-OH as outlined in Scheme 3.1.

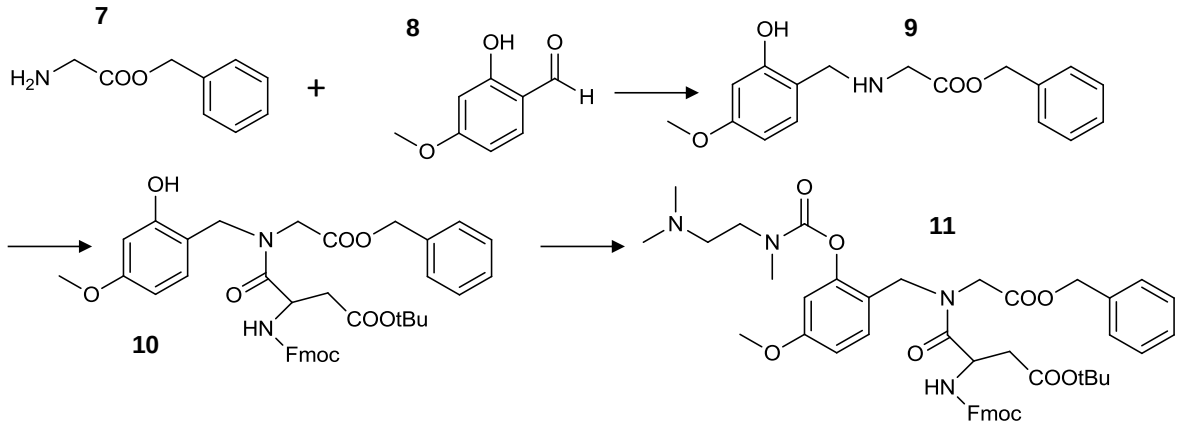
3.2.1. Synthesis of Fmoc-Asp-(Boc-Nmec-Hmb)Gly-OH

The Hmb derivative of the benzyl ester of glycine **9** was obtained by the reductive alkylation of H-Gly-OBzl (**7**) with hydroxymethoxybenzyl aldehyde (**8**) in ethanol, and the free base was isolated as a crystalline compound. Acylation of **9** with Fmoc-Asp-OH/*N,N'*-diisopropylcarbodiimide gave the dipeptide Fmoc-Asp-(Hmb)Gly-OBzl (**10**) as a solid foam after purification by dry flash chromatography.

The phenolic hydroxyl group was activated with p-nitrophenylchloroformate followed by treatment with mono-Boc-*N,N'*-dimethylethylenamine to give carbamate **11**.

Introduction of the Boc-Nmec group resulted in a large increase in retention time of product **11** in TLC systems, and facilitated purification by dry flash chromatography.

The benzyl ester was removed by catalytic hydrogenation, and the extent of the reaction was followed by TLC. The reaction was terminated when traces of 9-methylfluorene were detected by TLC.



Scheme 3.1: **Synthesis of Fmoc-Asp-(Boc-Nmec-Hmb)Gly-OH**

After purification by dry flash chromatography, the protected dipeptide was obtained as a solid foam in an overall yield of 45%. The ^{1}H NMR and ^{13}C NMR spectra were in accordance with the expected structure (data not shown). Moreover, the compound was characterized with MALDI-TOF (MW_{th}=816.3) (*Fig*. 3.8).

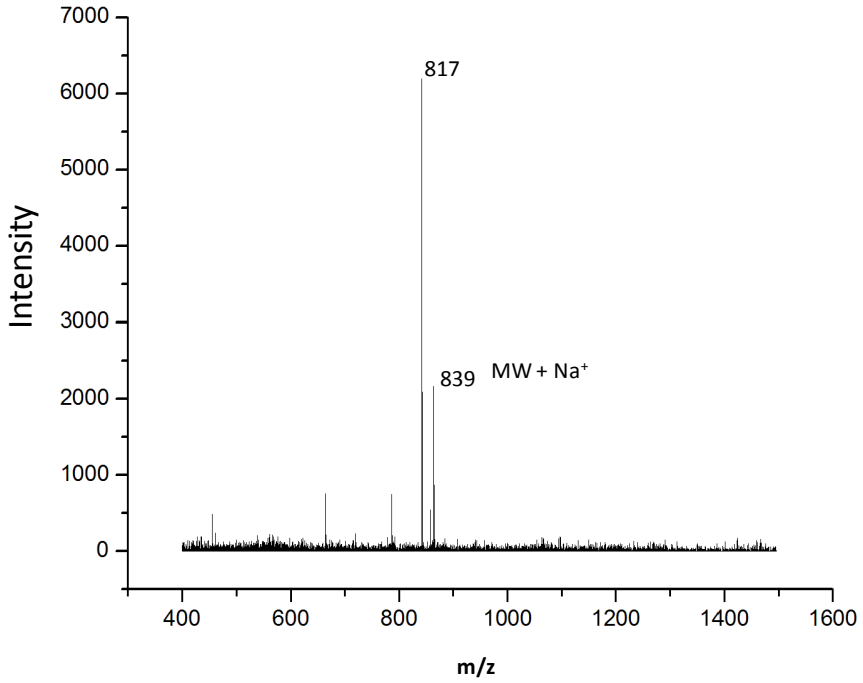


***Fig*. 3.8: MALDI spectrum of Fmoc-Asp(Boc-Nmec-Hmb)Gly**
(MWth=816.5/MWexp=817.0)

3.2.2. Peptide Synthesis of C-UBQ fragment and purification

Initially the C.terminal fragment 53-76 (23 amino acids) was synthesized, before coupling of the dipeptide of Fmoc-Asp(Boc-Nmec-Hmb)Gly-OH. The Fmoc-amino acids were coupled using PyBOP/DIEA in DMF, and a solution of 20% piperidine/0.1M HOBt was used for the deprotection of Fmoc group.

The dipeptide Fmoc-Asp-(Boc-Nmec-Hmb)Gly-OH was coupled to fragment C-UBQ 53-76 using DIC a stronger activator in DMF. The reaction was carried out for 30 min under microwave followed by 60 min at room temperature. For the couplings of the following seven 7 residues DIC was used.

The product C-UBQ was cleaved with TFA and precipitated five times with cold diethyl ether. The resulting Hmb-protected peptide was re-dissolved in acetonitrile/water, lyophilized, and analyzed by RP-HPLC and MALDI experiment (MW_{th}=3905.6) (*Fig*. 3.9).

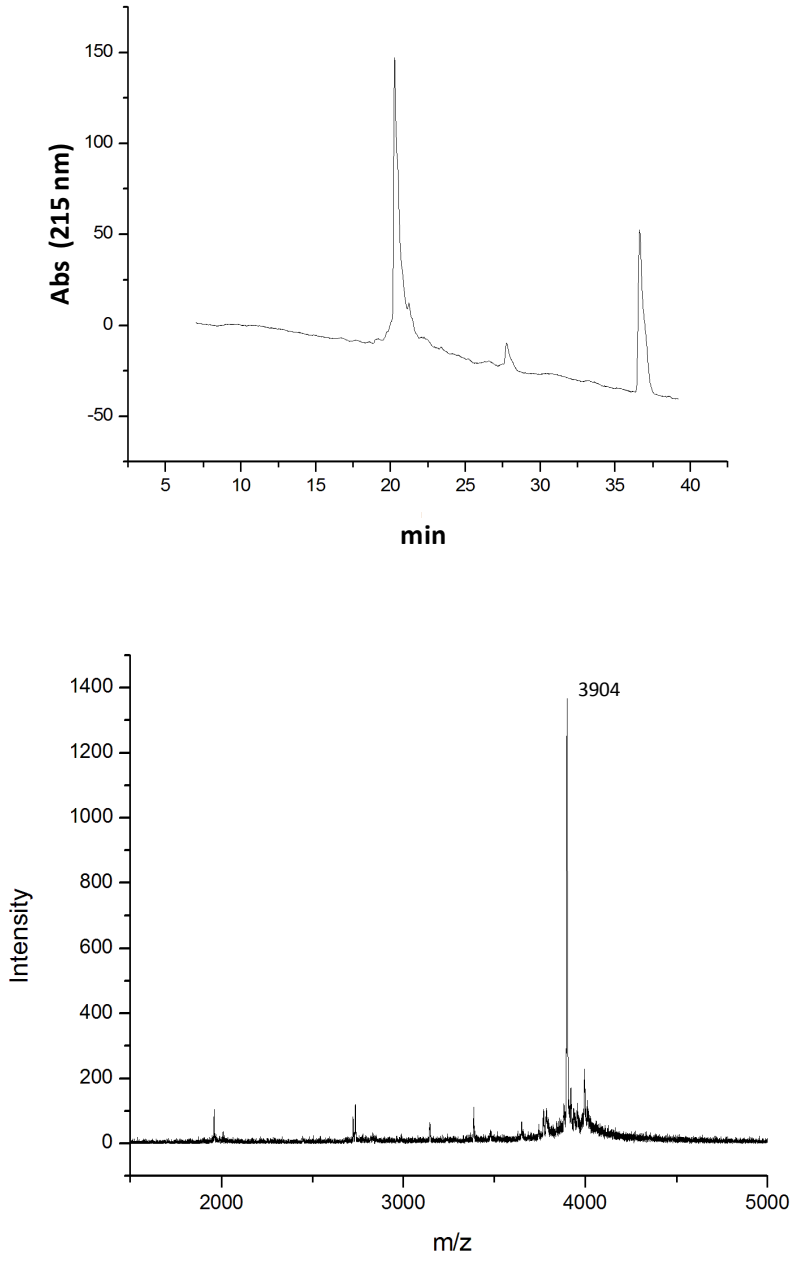


***Fig.* 3.9: Analytical chromatogram at 210 nm of crude C-UBQ (upper) and MALDI spectrum of crude C-UBQ (lower)** (MWth=3905.5/MWexp=3904.0)

Purification was carried out using a preparative Varian Polaris C8 column applying a gradient of acetonitrile (10 % to 60 % B in 30 min). The pure product (C-UBQ) was analyzed at MALDI-TOF (*Fig.* 3.10).

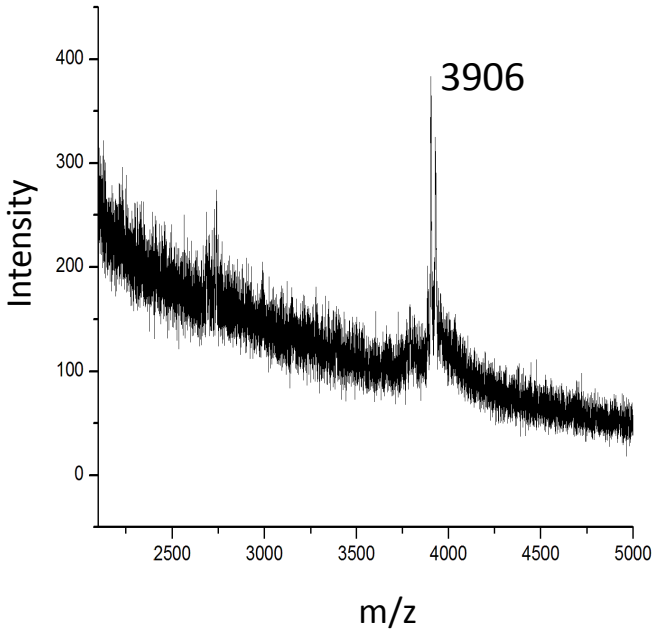


***Fig.* 3.10: MALDI spectrum of the pure C-UBQ** (MWth=3905.5/MWexp=3906.0)

3.3. Peptide Synthesis of N-UBQ fragment

By visual inspection of primary sequence also this fragment shows possible problems associated with aggregation and aspertimmide formation.

But alternative ways were done for the synthesis of this fragment. It is well-known that the incorporation of HOBt (0.1 M) in the piperidine solution allows a reduced level of aspartimide formation. In addition, the use of piperazine in place of piperidine has demonstrated significantly lower levels of aspartimide formation (35). Piperazine is less odorous and toxic than piperidine, but with a pKa of 9.8 compared to 11.1 for piperidine, piperazine is a slower deprotection reagent. For conventional synthesis of hydrophobic sequences, the use of piperazine may not be desirable. But according to Stacey et al. (35), microwave energy is able to accelerate Fmoc deprotection substantially with piperazine and that complete Fmoc removal can be accomplished with piperazine in 3 min without the presence of deletion products.

On these basis the use of piperazine coupled with microwave synthesis was employed to reduce aspartimide formation. Furthermore deprotection was performed in two stages using a fresh solution of 5 % (v/v) piperazine in DMF-0.1M

HOBt. An initial deprotection of 30 s at 50 W was followed by a 3-min deprotection at 50 W, with a maximum temperature of 80 °C. The coupling step was performed with 5 equiv. of Fmoc protected amino acid in DMF (0.2 M), 4.5 equiv. of DIC in DMF (0.45 M) and 10 equiv. of DIPEA in NMP (2 M). All couplings were performed for 10 min at 25 W at a maximum temperature of 75° C. Following completion of the sequence the peptide was released from the resin with concomitant removal of side chain protecting groups by treatment with TFA//H_2O/TIS (95/2.5/2.5, v/v/v) at room temperature for 2 h. The crude peptide was precipitated in cold Et_2O.

3.3.1. Synthesis of MV-(N-UBQ) fragment

To a suspension of MV-C2-Benz-Mal (10 mg, 0.012 mmol) in 100 mL of 100 mM phosphate buffer (pH = 7) was added the crude N-terminal fragment (30 mg, 0.0026 mmol) and the mixture was stirred at room temperature for 24 hours. The solution was concentrated by lyophilisation and analyzed with ESI experiment (MW_{th}=5724). (*Fig.* 3.11).

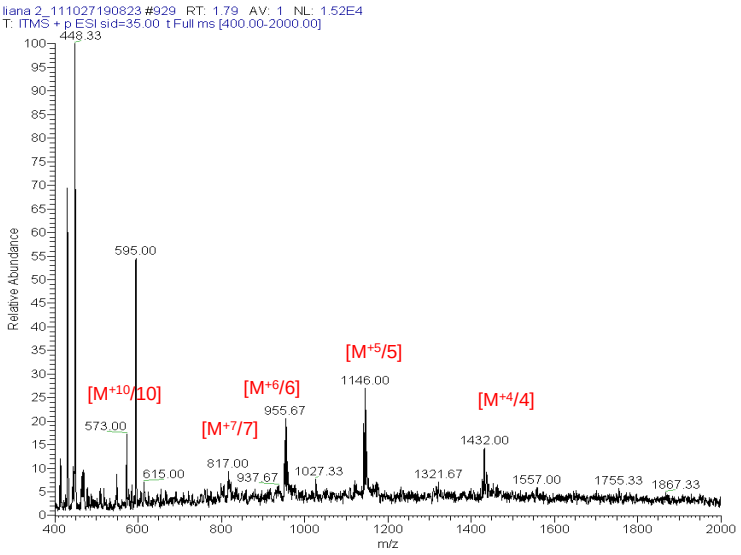


***Fig.* 3.11: ESI spectrum of pure MV-(N-UBQ) fragment**

4. DISCUSSION

Methods developed based on protein reassembly have several advantages such as rapidity, homogeneous format, and ability to perform in vivo screening (11). In addition when the reassembly is assisted by interacting pairs of proteins used such as leucine zippers (36) it reduces the possibility of false positives compared to the strategy where self-complementation is used for the reassembly of proteins as in the case of beta-galactosidase (10). Limitations of this method include, high background signal, possible interference in reassembly by the matrix components.

The biggest challenge of splitting strategy is designing fragments of reporter with affecting the 3-D structure of the protein drastically. For overcoming the problems linked to the reassembly assisted by the interaction with a host partner, in this study the host-guest chemistry for protein reassembly has been exploited , through non covalent interaction in CB[8] cavity of splitted fragments. CB [8] shows greater affinity to positively charged guest and its cavity is able to bind simultaneously two aromatic rings due to the formation of a cooperative bond between rings which results in greater interaction in the charge transfer between them.

Here we focuesd on ubiqutin both for its essential regulatory biological function and for its small size. The separated protein regions were obtained by chemical synthesis: at the C-terminal fragment (residues 47-76) was inserted a glycine as a spacer and then a tryptophan indole group, for the binding to the CB[8]. This fragment was synthesized using a microwave procedure, but the synthesis was rather complex due to the formation of aspartimmide. In order to avoid this side reaction, the dipeptide Fmoc-Asp-Gly-OH was synthesized, in it the amide bond is further protected with a derivative of the Hmb (2-hydroxy-4-methoxybenzyl) and employed for the development of the synthesis. At the C-terminus of N-terminal fragment (residues 1-46), a cysteine was inserted to allow the connection with "methylviologen", which has an aromatic ring strongly depleted of electrons for the formation of the ternary complex with CB[8]. The raw product was conjugated to methylviologen via preferential binding of the thiolate group with the maleimide function and was confirmed by ESI-MS. These preliminary synthetic results are very promising for the future development of this project that will be structural studies of individual fragments by solution state spectroscopy (NMR, CD, Fluorescence) and the structural and functional characterization of the ternary complex through the same techniques and with biochemical assays such as SPR, ITC and ELISA.

The protein reassembly through host-chemistry strategy can have a wide variety of applications: primarily in the study of protein folding determinants and then in the design of specific modulators of protein folding and biological function.

REFERENCES

1. Richards FM, Vithayathil PJ (1959) The preparation of subtilisn-modified ribonuclease and the separation of the peptide and protein components. J Biol Chem 234:1459–1465.
2. Gay GP, Fersht AR (1994) Generation of a family of protein fragments for structure-folding studies. 1. Folding complementation of two fragments of chymotrypsin inhibitor-2 formed by cleavage at its unique methionine residue. Biochemistry 33:7957–7963.
3. Shiba K, Schimmel P (1992) Functional assembly of a randomly cleaved protein.Proc Natl Acad Sci U S A 89:1880–1884.
4. Southworth MW, Adam A, Panne D, Byer R, Kautz R, Perler FB (1998) Control of protein splicing by intein fragment reassembly. EMBO J 17:918–926.
5. Pelletier JN, Campbell-Valois FX, Michnick SW (1998) Oligomerization domain-directed reassembly of active dihydrofolate reductase from rationally designed fragments. Proc Natl Acad Sci USA 95:12141–12146.
6. Remy I, Michnik SW (1999) Clonal selection and in vivo quantitation of protein interactions with protein-fragment complementation assays.Proc Natl Acad Sci U S A 96:5394–5399.
7. Pelletier JN, Arndt KM, Pluckthun A, Michnick SW (1999) An in vivo library-versus-library selection of optimized protein-protein interactions. Nat Biotechnol 17:683–690.
8. Karimova G, Pidoux J, Ullman A, Ladant D (1998) A bacterial two-hybrid system based on a reconstituted signal transduction pathway.Proc Natl Acad Sci U S A 95:5752–5756.
9. Johnsson N, Varshavsky A (1994) Split ubiquitin as a sensor of protein interactions in vivo. Proc Natl Acad Sci U S A 91:10340–10344.
10. Rossi F, Charlton CA, Blau HM (1997) Monitoring protein-protein interactions in intact eukaryotic cells by beta-galactosidase complementation. 94:8405–8410.
11. Sapna K. Deo. Exploring bioanalytical applications of assisted protein reassembly. (2004) Anal Bioanal Chem. 379: 383–390.
12. Comprehensive Supramolecular Chemistry, Vol. 1 Molecular Recognition: Receptors for Cationic Guests (Ed.: G. W. Gokel), Pergamon, Oxford, 1996.
13. Comprehensive Supramolecular Chemistry, Vol. 2 Molecular Recognition: Receptors for Molecular Guests (Ed.: F. VWgtle), Pergamon, Oxford, 1996.
14. D. J. Hill, M. J. Mio, R. B. Prince, T. S. Hughes, J. S. Moore, (2001). A field guide to foldamers. Chem. Rev. 101, 3893 – 4011.
15. G. Wulff, (2002), Enzyme-like catalysis by molecularly imprinted polymers. Chem. Rev. 102, 1 – 27.
16. R. Behrend, E. Meyer, F. Rusche, Justus Liebigs (1905) Ann. Chem., 339, 1 – 37.
17. W. A. Freeman, W. L. Mock, N.-Y. Shih, J. Am. Chem. Soc. 1981, 103, 7367 – 7368.
18. D. Jiao, N. Zhao, O. A. Scherman, (2010), A "green" method for isolation of cucurbit[7]uril via a solid state metathesis reaction. Chem. Commun. 46, 2007 – 2009.

19. Z. Miskolczy, L. Biczok, M. Megyesi, I. Jablonkai, (2009). Inclusion complex formation of ionic liquids and other cationic organic compounds with cucurbit[7]uril studied by 4',6-diamidino-2-phenylindole fluorescent probe. J. Phys. Chem. B, , 113, 1645 – 1651.
20. D. Jiao, O. A. Scherman, PCT/GB2010/002330.
21. Thiocarbonyl analogues of cucurbiturils have been discussed in the patent literature. For DFT calculations, see: F. Pichierri, Chem. Phys. Lett. (2004), 390, 214 – 219.
22. C. Marquez, R. R. Hudgins, W. M. Nau, (2004) Mechanism of host-guest complexation by cucurbituril. J. Am. Chem. Soc., 126, 5806 – 5816.
23. B. Honig, A. Nicholls, Science (1995), Classical electrostatics in biology and chemistry. 268, 1144 – 1149.
24. J. Lagona, P. Mukhopadhyay, S. Chakrabarti, L. Isaacs, (2005). The cucurbit[n]uril family. Angew. Chem., 117, 4922 – 4949; Angew. Chem. Int. Ed. 44, 4844 – 4870.
25. J. W. Lee, S. Samal, N. Selvapalam, H.-J. Kim, K. Kim, (2003) Cucurbituril homologues and derivatives: new opportunities in supramolecular chemistry. Acc. Chem. Res, 36, 621 – 630.
26. H. Buschmann, E. Cleve, E. Schollmeyer, Inorg. Chim. Acta,(1992), 193, 93.
27. J. Kim, I.-S. Jung, S.-Y. Kim, E. Lee, J.-K. Kang, S. Sakamoto, K. Yamaguchi, K. Kim, (2000) J. Am. Chem. Soc., 122, 540 – 541.
28. U.S. Patent: K. Kim, J. Kim, I.-S. Jung, S.-Y. Kim, E. Lee, J.- K. Kang, Pohang University of Science and Technology Foundation, 2000.
29. J. W. Lee, K. Kim, S. Choi, Y. H. Ko, S. Sakamoto, K. Yamaguchi, K. Kim, (2002). Unprecedented host-induced intramolecular charge-transfer complex formation. Chem. Commun., 2692 – 2693.
30. M. E. Bush, N. D. Bouley, A. R. Urbach, (2005). Charge-mediated recognition of N-terminal tryptophan in aqueous solution by a synthetic host.J. Am. Chem. Soc., 127, 14511 – 14517.
31. Vijay-Kumar S, Bugg CE, Cook WJ. (1987). Structure of ubiquitin refined at 1.8 A resolution. J Mol Biol.;194(3):531-44.
32. I. Schön, et al. (1991) J. Chem. Soc., Chem. Commun., 3213.
33. E. Nicolás, et al. (1989) Tetrahedron Lett., 30, 497.
34. Wahlstrom K, Planstedt O, Unde´n A. (2008). Synthesis and purification of aggregation-prone hydrophobic peptides by the incorporation of an Fmoc dipeptide with the peptide bond protected with a modified 2-hydroxy-4-methoxybenzyl (Hmb) group. Tetrahedron Letters. 49 3921–3924.
35. Stacey A. Palasek, Zachary j. Cox and Jonathan M. Collins. (2007) Limiting racemization and aspartimide formation in microwave-enhanced Fmoc solid phase peptide synthesis. J. Pept. Sci.; 13: 143–148.
36. Ghosh I, Hamilton AD., and Regan L. (2000). Antiparallel Leucine Zipper-Directed Protein Reassembly: Application to the Green Fluorescent Protein. J. Am. Chem. Soc. 122, 5658-5659

CONCLUSIONS

The innovative concept that proteins exert their function by establishing intricate contact networks rather than acting as independent entities is of outstanding importance both in the comprehension of molecular mechanisms underlying biological processes and in modern drug discovery. Protein-protein interactions are the ensemble of fine tuned recognition events that take place at protein surfaces. These contacts frequently involve large protein surface areas, which comprise several contact sites but, alternatively, they can be mediated by “hot spots” represented by few crucial amino acids. Many studies have highlighted the role of peptide molecules as powerful tools for the characterization and the regulation of these interactions. Peptides are particularly suitable as models for proteins that are not fully folded in their isolated states but that achieve conformational stability upon the formation of complexes with other partners. The characterization of the structural determinants of protein-protein recognition represents an important step for the development of molecular entities able to modulate these interactions.
The identification of compounds able to modulate protein-protein interactions represents a major goal for modern drug discovery. Both structure-based design and screening of combinatorial collections have been successfully applied, in this instance often small peptides and peptidomimetics have been selected as inhibitors of protein complexes. In many cases, the lack of sufficient structural information on protein interacting regions renders the combinatorial approach the most suitable method, but the use of completely random peptide libraries often produces redundant chemical species that complicate the deconvolution phases of screening and puzzles the way to the identification of active compounds. Thus, in order to avoid the synthesis of too large combinatorial libraries in two case-studies during this doctoral thesis a simplified approach was followed in it libraries of short peptides, made of a small subset of amino acids were prepared and screened. This approach has been applied to the identification of inhibitors of the protein complex between PED/PEA15 and PLD1 that is involved in the molecular mechanisms of insulin resistance occurring in type 2 diabetes. This study led to the identification of several short peptide inhibitors rich in aromatic and H-donor donor features through the employment of ELISA and SPR binding assay. These short sequences will be converted in small organic molecules following medicinal chemistry rules and

actually can be considered as interesting starting scaffold for the design of new therapeutic agents in type 2 diabetes.

On the other hand, the screening of small but focused libraries is increasingly used to target specific systems. This method has been applied to the identification of mimicking peptides of KIR region of SOCS1 that is involved in JAK2 recognition. This interaction plays a key role in molecular mechanisms of psoriasis. After the identification of a shorter SOCS1-KIR domain, through an Ala-scanning investigation, focused peptide library in which non essential residues were randomized in a simplified approach were built and assayed. This screening led to the selection of new peptides more effective than natural sequences; indeed one decapeptide resulted binds to JAK2 with a dissociation constant in the nanomolar range and was able to block STAT1 phosphorylation. Actually cellular and *in vivo* experiments are ongoing in order to investigate its therapeutic applicability.

Rational approaches for the identification of protein interacting regions provide deeper insights into the recognition molecular mechanism This rational approach has been applied to other two projects in this thesis. The complex between proteins CypA and AIF is involved in apoptotic cell death, but at structural level only a docked theoretical model is available. In this model few molecular contacts between proteins are detectable, thus a set of sequences were designed in order to map crucial residues of AIF domains involved in the interaction with CypA. The selective disruption of the CypA-AIF complex with these peptides mimicking the protein surfaces was analysed by competition and direct binding SPR experiments. These experiments allowed to identify an active sequence that, once properly modified, could be as a potential suppress the pro-apoptotic action of AIF.

This minimalistic approach was successfully applied to another protein complex IκB-α/Tat, two proteins involved in the replication of HIV-1. Here IκB-α-based peptides were designed and tested in SPR assays. Also in this project the interacting region of IκB-α was restricted to a shorter region (26 residues) representing a promising starting point for the development of molecules able to block HIV-1 infection. The characterization of IκB-α peptides has provided novel insights into the intrinsic properties of IκB-α fragments and structural determinants of IκB-α/Tat recognition.

In the investigation of protein folding determinants and then in the design of specific modulators of proteic folding and biological function, the protein reassembly through host-chemistry strategy could represent a valid approach. During a period of six months spent at Cambridge University, a protein splitting approach was applied to the small a regulator protein Ubiqutin. Two dissected protein regions were obtained by chemical synthesis: at the C-terminal fragment (residues 47-76) was inserted a glycine as a spacer and then a tryptophan indole group, for the binding to the macrocycle CB[8]; instead at the C-terminus of N-terminal fragment (residues 1-46), a cysteine was inserted to allow the connection with "methylviologen", which has an aromatic ring strongly depleted of electrons for the formation of the ternary complex with CB[8]. The preliminary synthetic results obtained resulted promising for the future development of this project that will be structural studies of individual fragments and the structural and functional characterization of the ternary complex.

ABBREVIATIONS:

TIS:	Triisopropylsilane
TFA:	Trifluoroacetic acid
Fmoc:	Fluorenylmethoxycarbonyl
HPLC:	High Performance Liquid Chromatography
LC-MS:	Liquid Chromatography Mass Spectrometry
SPR:	Surface Plasmon Resonance
Acm:	acetamidomethyl
TCEP:	Tris(2-Carboxyethyl) phosphine
IC_{50}:	Concentration giving 50% of Inhibition
PED/PEA15:	Phosphoprotein Enriched in Diabetes/Phosphoprotein Enriched in Astrocytes15
PLD:	Phospholipase D
ELISA:	Enzyme-Linked ImmunoSorbent Assay
EDC:	1-Ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride
NHS:	N-hydroxy-succinimide
DMF:	Dimethylformamide
DCM:	Dichloromethane

PUBLICATIONS

1. Scognamiglio PL, Doti N, Grieco P, Pedone C, Ruvo M, Marasco D. Discovery of small peptide antagonists of PED/PEA15-D4α interaction from simplified combinatorial library. Chemical Biology & Drug Design. 2011 May;77(5):319-27. Epub 2011 Mar 1.

2. Vitagliano L, Fiume G, Scognamiglio PL, Doti N, Cannavò R, Puca A, Pedone C, Scala G, Quinto I, Marasco D. Structural and functional insights into IκB-α/HIV-1 Tat interaction. Biochimie. 2011 Sep;93(9):1592-600. Epub 2011 May 31.

3. Doti N, Scognamiglio PL, Madonna S , Scarponi C, Ruvo M, Albanesi C, Marasco D. New mimetic peptides of Kinase Inhibitory Region (KIR) of SOCS1 through focused peptide libraries. Biochemical Journal. 2011. (under revision).

4. Scognamiglio PL, Perretta G, Marasco D. Automation for the identification of bioactive compounds. Automation, book by In-tech. 2011. (submitted).

5. Scognamiglio PL, Perretta G, Marasco D. From peptides to small molecules: a natural but puzzled way to the development of new drugs. 2011. (review in preparation).