UNIVERSIDADE NOVA DE LISBOA

FACULDADE DE CIÊNCIAS E TECNOLOGIA

DEPARTAMENTO DE QUÍMICA

Ana Margarida Gonçalves Carvalho Dias

Development of an Affinity Pair "Tag-Receptor" for Recombinant Protein Expression and Purification

Dissertação apresentada para a obtenção do Grau de Mestre em Biotecnologia, pela Universidade Nova de Lisboa, Faculdade de Ciências e Tecnologia

Orientadoras:

Prof.ª Doutora Ana Cecília Roque
Doutora Ana Sofia Coroadinha

LISBOA

2010

Nº DE ARQUIVO

UNIVERSIDADE NOVA DE LISBOA

FACULDADE DE CIÊNCIAS E TECNOLOGIA

DEPARTAMENTO DE QUÍMICA

Ana Margarida Gonçalves Carvalho Dias

Development of an Affinity Pair "Tag-Receptor" for Recombinant Protein Expression and Purification

Orientadoras:

Prof.ª Doutora Ana Cecília Roque
Doutora Ana Sofia Coroadinha

LISBOA

2010

ACKNOWLEDGMENTS:

In this year, I have been developing my master thesis, I would like to thank to my coordinator Cecília Roque for the guidance and encouragement. Also, thanks for opening my scientific horizons, since I had contact with different scientific areas and techniques, and to challenge myself to take further steps in my academic formation.

I would like to acknowledge my colleagues in the Biomolecular Engineering Laboratory (FCT-UNL) for the good work environment, helpful discussion as well as the enthusiasm of making science. Particularly to Ana Pina, who helped me in the development of the thesis, by spending part of her time teaching the methods and responding to my doubts and for the helpful discussion of the results; but also to make the work environment more friendly and fun, even in the more stressful moments.

During this year I had the possibility to expand my knowledge in mammalian cells culture and in molecular biology methods, for that was important the work that I developed in Cell Line Development and Molecular Biotechnology Laboratory (ITQB-UNL) coordinated by Ana Coroadinha, who I would like to thank for the disponibility to teach the methods and the discussion of the results. To the other members of this group, specially my master colleagues, that made me feel welcome and were helpful in everything I needed, I would like to express my gratefulness.

Finally, I would specially acknowledge to my parents that have been my backup throughout my years of study. They have been my support and inspired me to develop my academic knowledge and skills but also help me grow as a person. They demonstrate that hard work and perseverance compensate.

To all, who in different situations, ways and moments taught and impelled me to achieve my goals, my sincere thanks!

ABSTRACT:

The main objective of this work was the development of an affinity pair for the purification of recombinant proteins.

In this work, ligands based on the Ugi Reaction and the 1,3,5-Triazine scaffold were synthesised in solid-phase and screened for binding to an affinity tag, an hexapeptide constituted by asparagine aminoacid (N). The ligands were tested against pure solutions of the hexapeptide and Green Fluorescence Protein (GFP), used as a model protein. The ligands that had the highest affinity for the hexapeptide and lowest affinity for the protein were selected for further studies with cellular extracts. The cellular extracts were produced in HEK 293T cells transfected with two designed vectors: one containing the GFP tagged with the affinity tag, and the other containing GFP without tag. The efficient expression of a recombinant GFP fused with the designed affinity tag was demonstrated. The cellular extracts were then loaded onto chromatographic columns containing the lead ligands immobilised onto agarose, and the amount of total protein and GFP bound and eluted noted. The results demonstrated that the Ugi ligands were less selective than the Triazine ligands for the hexapeptide. The triazine ligand 7,4 has been considered as the most selective for the designed affinity tag.

In addition, preselected lead ligands for another hexapeptide (RW) of interest were studied. Mammalian cells HEK 293T were transfected with a vector expressing for GFP tagged with this peptide. The ligands immobilized onto agarose were loaded with cellular extracts, being noted that the lead A6C3 showed a high selectivity for the tag tested.

RESUMO:

Este trabalho teve como objectivo o desenvolvimento de um par de afinidade para purificação de proteínas recombinantes.

Neste trabalho foram sintetizadas duas bibliotecas de ligandos em fase sólida, baseadas na reacção de Ugi e na estrutura da 1,3,5-Triazina para formar um par com o hexapéptido constituído por asparaginas (N). A afinidade dos ligandos for testada usando uma solução

pura de hexapéptido e a "Green Fluorescence Protein" (GFP), proteína modelo. Os ligandos que tinham maior afinidade para o hexapéptido foram seleccionados para estudos com extractos celulares. Os extractos celulares foram produzidos em células HEK 293T, células animais que foram transfectadas com dois vectores: um a GFP recombinante com uma cauda de afinidade e outro com GFP. Foi demonstrado que é possível produzir GFP recombinante com cauda de afinidade funcional.

Os resultados demonstraram que os ligandos do Ugi são menos selectivos do que os da Triazina para o péptido. O ligando 7,4 da biblioteca da triazina demonstrou mais afinidade para o hexapéptido.

Foram também estudados ligandos de Ugi pré-seleccionados para outro hexapéptido (RW). Estes foram testados com extractos celulares de HEK 293T transfectadas com um vector com GFP recombinante com uma cauda de afinidade com este péptido. Os resultados permitiram concluir que o ligando A6C3 é muito selectivo para este péptido.

CONTENTS:

INDEX OF FIGURES

INDEX OF TABLES

INDEX OF ABREVIATIONS

a.a. - aminoacids
Amp - Ampicillin
Arg - Arginines
Asn – Asparagines (N)
Asp – Aspartic Acid
Ax, By – Ligand based on the Ugi Multicomponent Reaction (A – Amine, B – Carboxylic acid)
(X, Y) – Amine X (R1), Amine Y (R2) – Triazine scaffold based ligands
BSA – Bovine Serum Albumine
cDNA – circular DNA, a vector
$[DNA]_f$ –Final Concentration of DNA
Eq - equivalents
E. coli – Escherichia coli
FITC – Fluorescein Isothiocyanate
GFP – Green Fluorescent Protein
Hexapeptide NN – hexapeptide composed by asparagines (N)
Hexapeptide RW – hexapeptide composed by 3 tryptophan (W) and 3 arginines (R)
HEK293T cells – Human Embryonic Kidney cells
His-Tag – Histidine Affinity Tag
IMAC – Immobilized Metal Ion Affinity Chromatography
IgG – Immunoglobuline G

LMW Marker – Low Molecular Weight SDS Calibration Kit
mAb – Monoclonal Antibody
NMR – Nuclear Magnetic Resonance
PBS – Phosphate Buffered Solution
PEI - Polyethylenimine
PSA - Ammonium Persulphate
pRWTagMGFP – Vector that expresses fused GFP with a hexapeptide RW Affinity Tag
p6NTagMGFP – Vector that expresses fused GFP with a hexapeptide NN Affinity Tag
pMDISGFP – Vector that expresses GFP without Affinity Tag
TAE – Tris-Acetate (EDTA) buffer
TEMED - N,N,N,N-Tetramethylethylenediamine
TLC – Thin Layer Chromatography
V_{plate} – Petri Plate Volume

CHAPTER 1 - INTRODUCTION

1.1 Production of Recombinant Proteins

In recent years important therapeutics as proteins drugs, monoclonal antibodies (mAbs), enzymes, interferons or vaccines, have been developed. The medical areas in expansion regarding the search for new biotherapeutics are the cancer, followed by the infections diseases (PhaRM 2008).

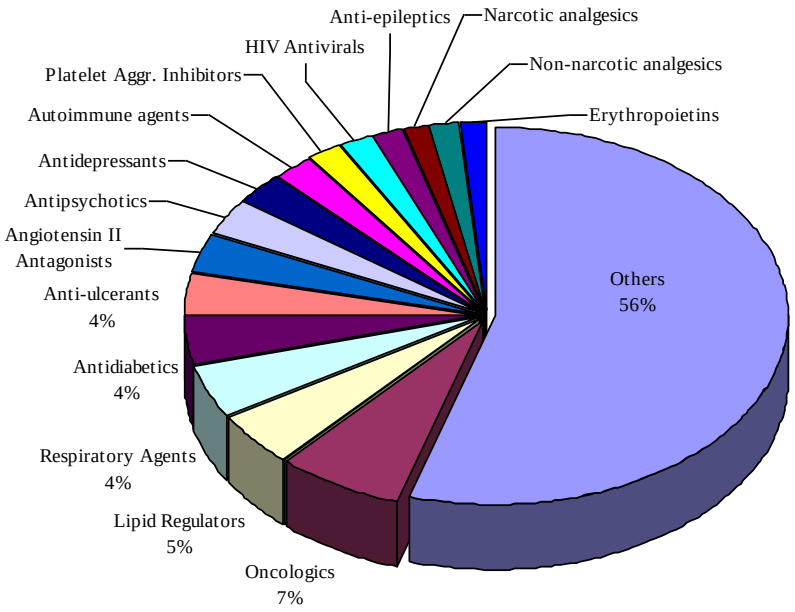


Figure 1.1- Top 15 Medicine Sales 2009, adapted from report of (Midas 2009)

In 2009 the market of pharmaceutics was estimated in 75202 US$. The oncologic therapeutics alone accounted for 52372 US$ (7% of the market). The second most rentable medicines were the lipid regulators, they were evaluated in 35281US$ (5% of the market). The infectious drug related diseases represented 13758 US$ (2% of the market) (Figure 1.1) (Midas 2009).

If in the past, some protein drugs (e.g. insulin) were produced from scarce animal sources, the technology of recombinant DNA allowed for the production of larger amounts of proteins in order to deal with the market demand. The industry started using recombinant proteins that were expressed in bacteria, mainly in *E. coli* (Graumann and Premstaller 2006), or even more recently mammalian cells – Chinese hamster ovary (CHO) cells, mouse myeloma cells or human embryo kidney cells (HEK-293) (Gupta and Lee 2007; Seth et al. 2007). The main interests in bacterial expression systems are the high growth level, easiness of manipulation,

wide implementation of expression vectors and experimental methods, as well as high productivity rates (Graumann and Premstaller 2006). However, these organisms do not make complex post-translational modifications in proteins, as glycosylation, phosphorylation, misfolding, oxidation and deamination of aminoacids, and proteolysis (Jenkins et al. 2008). These modifications are fundamental for secretion, stability and efficiency of the therapeutic drug, in order to avoid immune system responses.

Mammalian cells cultures are able to produce more complex biotherapeutics as proteins from eukaryotic sources that need glycosylation (Sevastsyanovich et al. 2009). Despite of these advantages, these hosts present lower growth rates, require specific culture medium and laboratorial technological requisites. Furthermore, the time to develop a cell line producing a therapeutic product at an acceptable yield is longer than in bacteria (Seth et al. 2007). These factors have increased the production costs but with the demand of biotherapeutics for chronic diseases, which require high doses, the production level in mammalian has progressively improved and nowadays it can reach up to 5g/l (Shukla and Thömmes 2010). In consequence, the pressure has passed from the production methods to the downstream processes (purification and polishing).

1.2 Purification of Recombinant Proteins

The downstream processing is responsible for 50% to 80% of the total cost of production for a recombinant protein, particularly when the purity required for the final product is very high. To introduce a product in the market is necessary to insure its quality, safety and efficacy, which is related with product purity (Clonis 2006).

In downstream processes, chromatography methods, e.g. affinity, hydrophobic interactions and ion-exchange, are the most commonly used in bioseparation. Affinity chromatography is one of the best methods to obtain high quality biotherapeutics, because it allows the direct application of the cellular extracts and also is a technique that is already implemented in the industry (Clonis 2006). However, the high cost of resins, the space needed for the equipment, and the cleaning and sanitization of these supports fostered the search for new separation processes (Low et al. 2007). Non-chromatographic methods, as affinity precipitation, aqueous

two-phase systems (Roque et al. 2007), crystallization and magnetic separations (Low et al. 2007) are becoming more popular.

1.3 Affinity Ligands for Bioseparation Processes

To improve the efficacy, affinity and selectivity of affinity-based purification methods, new resins have been developed. There have been progresses in the design of affinity ligands to increase the purification of different biomolecules. These ligands can be classified as biological or synthetic. The former take advantage of the natural affinity, like antibody-antigen, e.g. Protein A and IgG, these ligands can be classified as biospecific (Roque et al. 2007). As consequence of their biological origin the costs of production and purification are high. Also ligand leaching can occur causing contamination of the end product and the sanitization of the resins is limited, due to the possibility of ligands degradation. To overcome these disadvantages, these ligands were resynthesized to maintain their good features and improve their strength against degradation, and can be described as pseudospecific (Roque et al. 2007). More recently fully synthetic ligands have been developed, nevertheless possess similar features to the biological ligands, e.g. *de novo* designed, but also ligands that are bioengineered, e.g. peptides or small protein domains; they can be named biomimetic ligands (Roque et al. 2007).

Biomimetic ligands are usually synthesized based on the knowledge of the structure of the target protein and some important interactions that can occur in nature with other proteins or natural ligands. By using the knowledge about the structure of the target protein it is possible to construct a ligand, study the possible binding interactions between them (*in silico*), and then construct a library of ligands that have interesting features and can have different affinity and selectivity to the target protein (Clonis 2006).

This technique uses Molecular Modelling softwares, to understand the interactions between ligands and the molecule of interest (protein, antibody or peptide) (Forster 2002). Nowadays, there are structure databases with information about known protein structures, e.g. Protein Data Bank that are essential for these studies (Smith and Sternberg 2002).

Molecular modelling studies can give information about the sites of ligand-protein binding and the geometry of the complex, but also can calculate the free energy involved in the binding process (Forster 2002). The basic principle of these studies is that the lowest the free

energy of the conformation between the ligand and the protein, the more favourable is the interaction (Thomsen and Christensen 2006).

In silico studies can therefore direct the search for ligands with the desired properties as chemical structure, functional groups, hydrophobicity or possibility to form hydrogen bonds, polarity, and charge. Other advantage of this rational design is the lowering of the cost and time consuming in experimental resources, during synthesis and screening (Davies et al. 2006). Even though *in silico* results are important, they are only a first step in the process of new ligands development, since the results must be demonstrated by experimental data (Figure1.2).

The next step, involves the synthesis of libraries of ligands. The libraries are normally synthesized in a solid-support - the resin, that should be chemical stable, have good mechanical strength and low non-specific adsorption, should maximize the surface area for the target product and, at the same time, have low surface for contaminants (Tozzi et al. 2003). The materials normally used are: natural polymers (e.g. agarose, cellulose or cross-linked dextran), synthetic polymers (e.g. polyacrylamide), or inorganic materials (e.g. silica) (Tozzi et al. 2003). Each support has different functional groups, therefore to choose the support is necessary to analyse each functional groups are interesting to start the activation and functionalization of the support.

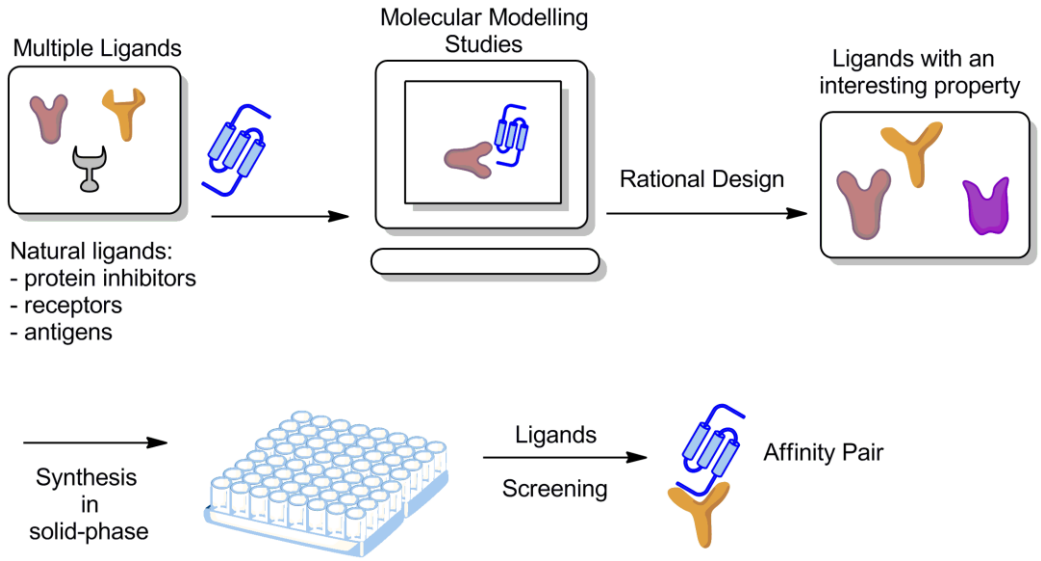


Figure 1.2 – Strategy to develop an affinity ligand using rational design. First step is the discovery of a property of interest and develop a ligand using molecular modelling studies; synthesis in solid-phase of ligands with the property of interest; finally, screening of the ligands with the molecule of interest, e.g. protein; development of an affinity pair (ligand – molecule).

The functionalization of the support can be based in combinatorial chemistry. Combinatorial chemistry is one method for the synthesis of libraries of compounds, by using reactions involving multi-component condensations, as one-pot reactions. In these reactions the core structure is kept constant in all ligands although the lateral groups can be changed which gives variability to the library of ligands, even though the characteristic of interest can be maintained (Armstrong et al. 1996; Dömling 2006).

An example of one-pot reactions is the Ugi-four component reaction, which combines an aldehyde, amine, isocyanide and a carboxylic acid compound (Ugi 2001). By changing one of the compounds or the four compounds it is possible to generate libraries with different number of ligands. This reaction has been used for the synthesis of different ligands, as a cysteine protease inhibitor (Nakamura et al. 2000); serine protease thrombin inhibitor (Illgen et al. 2000); drug discovery (Akritopoulou-Zanze 2008); or affinity ligands for separation of immunoglobulins (Haigh et al. 2009).

Another type of quite popular for combinatorial chemistry are the 1,3,5-Triazine based ligands. These ligands mimic the binding of natural anionic heterocyclic substrates like nucleic acid, nucleotides, coenzymes or vitamins in consequence they have high affinity for proteins (Lowe 2001). The libraries are simple to construct and take advantage of the temperature to sequentially substitute the chloride atoms by different nucleophiles (Lowe et al. 2001). These groups can be synthesized in supports with amino or hydroxyl groups (Palanisamy et al. 2000). In addition, the nitrogen atoms that are electron-rich can form hydrogen bonds with aminoacids in the binding site of the target protein (Labrou 2003). Some triazine based affinity ligands have been developed for purification of recombinant insulin precursor MI3 purification (Sproule et al. 2000); antibodies (Roque et al. 2005); anti-human immunodeficiency virus 2F5 (mAb 2F5) (Platis et al. 2006); human tissue plasminogen activator (Wu and Yu 2007); or human anti-HIV mAb 4E10 (Platis et al. 2009).

Once generated the libraries of ligands need to be tested in order to select potential lead ligands. Several methods have been used in High-Throughput Screening (HTS) and can be automatic or partially automatic. The screening methodology can be designed considering different properties of the target molecule and ligand as fluorescence of the complex, colorimetric assays to quantify proteins (Bradford Method, Lowry Method or BCA assay) or

mass spectrometry. The fluorescent methods explore intrinsic properties of some proteins, that possess aromatic aminoacids residues e.g. tryptophan, tyrosine, phenylalanine. The Green Fluorescent Protein (GFP) from *Aequorea victoria* is a particular case because the fluorescence is due to a chromophore (Thor et al. 2005). The GFP is used as a protein model in several studies including: genetics (e.g. demonstrate the expression of a new vector or gene expression marker (Inouye and Tsuji 1994;Zhuang et al. 2008)); affinity chromatography (by fusion with tags to confirm affinity and selectivity of ligands in solid support for the tag), or in cell biology, to understand cellular mechanisms (Chung et al. 2009). There are small molecules (e.g. fluorescein isothiocyanate (FITC)) that can be conjugated with a target protein and then used on the screening of the affinity ligands, which as proved to be an inexpensive technique (Roque et al. 2004). The methods that use fluorescence are normally more sensitive, less expensive and less time consuming than the colorimetric assays, as they do not requires reagents.

1.4 Affinity Tags in Protein Purification

As a result of the difficulty of developing affinity ligands for each specific protein, other purification strategies have been found. Small peptides can be fused with the target protein and after the expression purified in a chromatographic column containing an affinity ligand specific for the small peptide (Waugh 2005). An example is the use of the Histidine-tag, a small peptide (2-10 Histidines) that interacts with adsorbents (Ni(II)-nitrilotriacetic acid resin) also known as a Immobilized Metal Ion Affinity Chromatography (IMAC) resin. The IMAC method uses the affinity of certain aminoacid side chains to chelated metal ions. The imidazol group plays the important role in the binding interaction to the metal ions in the surface of the resin (Mondal and Gupta 2006). This technique have a low cost, can resist to multiple regeneration processes and exhibits high binding capacity (Waugh 2005), although this strategy is less useful when the proteins have metal ions (Arnau et al. 2006).

Other tags exploit different types of affinity interactions (Hedhammar et al. 2005) (Table 1.1). The FLAG-tag, a hydrophilic octapeptide, is isolated by an antibody coupled to the chromatographic resin (Arnau et al. 2006); the natural affinity complex of streptavidin-biotin is explored for development of a Streptavidin binding peptide tag that is purified using a streptavidin column, this tag is eluted using biotin.

Table 1.1 – Affinity tags commonly used for purification of recombinant proteins.

Affinity Tag	Comments	References
His-Tag	5 -15 Histidines uses IMAC resin	(Derewenda 2004)
FLAG tag	Antibody based purification	(Einhauer and Jungbauer 2001)
Streptag II	Streptavidin column	(Cass et al. 2005)
c-myc	Antibody based purification	(Terpe 2003)
T7-tag	Antibody based purification	(Chatterjee and Esposito 2005)
Arg	5 to15 Arginines uses a cation-exchange resin	(Terpe 2003)
Asp	5-15 Asparagines uses a anion-exchange resin	(Hedhammar et al. 2005)
S-tag	Based on the affinity of S-Protein (RnaseA)	(Terpe 2003)

The cost and the scalability of the chromatographic supports are crucial for the selection of the affinity tag pair, since natural products, (e.g. antibodies) are more costly and less stable than synthetic ones. In addition, some tags can improve the production of the protein when placed at the N-terminal position or even enhance the protein solubility. Nevertheless, it is important to analyse the effect of the tag in the structure of the protein of interest (Waugh 2005).

The tags derived from peptides offer multiple advantages because the methods of synthesis have improved and the variability is ensured by combinatorial methods in solid phase or through phage display methods, allow the production of tags with different chemical and physical properties or functional groups (Tozzi et al. 2003), that can be studied for development of new ligands for application in affinity purification.

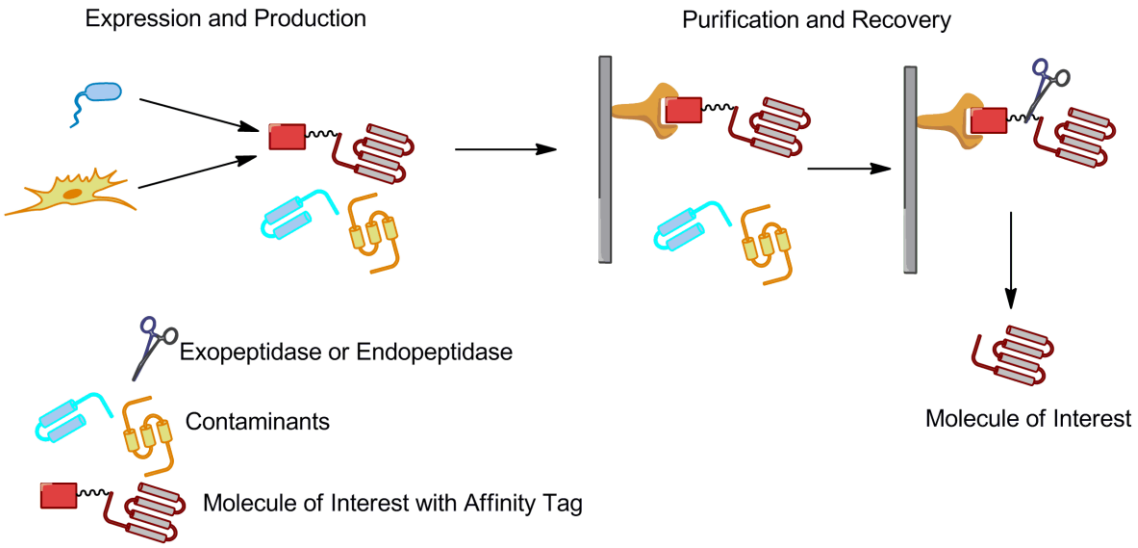


Figure 1.3 – Expression of fused protein in different hosts, purification and recovery of the fused protein using an Affinity Tag with a sequence for recognition of an enzyme. This sequence is used for remove the affinity tag from the protein.

However, after purification and elution of the fusion protein, it is often necessary to remove the tag from the protein (Figure 1.3).

This can be performed by chemical or enzymatic methods. By chemical methods, harsh conditions are used which can denature the proteins (Hedhammar et al. 2005). As a consequence, enzymatic methods are preferentially used. For this purpose, a sequence for enzyme recognition is added to the tag. This sequence has to be considered in the design of the vector for expression of the fusion protein (Arnau et al. 2006). There are several endoproteases (e.g. TEVprotease, Enterokinase, 3CProtease) that are used for this purpose nevertheless the specificity of the cleavage is a problem, as enzymatic cleavage can result in protein degradation and low yields. Exopeptidases are another option (e.g. Aminopeptidase M, Carboxypeptidase A and B), although there are concerns about contamination of therapeutic protein samples (Arnau et al. 2006). To select the appropriate enzyme it is necessary to evaluate the specificity of cleavage but also the costs and the conditions to maintain the functionality of the target protein (Kenig et al. 2006) This subject is reviewed in detail by several authors (Hedhammar et al. 2005; Waugh 2005; Arnau et al. 2006; Kenig et al. 2006; Mondal and Gupta 2006).

1.5 Aims and Objectives

This thesis is part of a larger research project aiming at the development of affinity pairs for the purification of recombinant proteins. Specifically, this work focused on the development of two combinatorial libraries of affinity ligands, one based on the Ugi reaction and other in the 1,3,5-Triazine scaffold, for binding to a tag composed by a hexapeptide of Asparagines. These libraries were screened against the hexapeptide and GFP, used as a model protein. The most promising ligands, presenting highest affinity for the hexapeptide and lowest affinity for the GFP, were screened against cellular extracts, produced in HEK 293T cells transiently transfected with a vector for expression of a GFP fusion protein with an affinity tag.

CHAPTER 2 – MATERIALS AND METHODS

2.1 Materials

Chemical Compounds

The reagents utilized were of the highest grade available. 1-Amino-2-Propanol (11,024-8), 4-Aminobenzamide (98%, 284576), 5-Aminoisophthalic acid (94%, 18679), 9-Anthracenecarboxylic acid (99%, A8,940-5), Cyanuric Chloride (99%, C95501), 1,4-Diaminobutane (99%, D13208), Epichlorohydrin (99%), Glutaric Acid (99%, G3407), 3-(4-Hydroxyphenyl)propionic acid (98%; H5, 240-6), 3-Indoleacetic acid (98%, I375-0), Isopropyl Isocyanide (97%), Phenylacetic acid (P16621), Phenethylamine (12,894-5), 1-Pyrenemethylamine hydrochloride (95%, 401633), Sodium Periodate ($NaIO_4$) 99,8% A.C.S. Reagent, Succinamic acid (97%, 134376), Tyramine (T9,034-4) were acquired from Aldrich (Sintra, Portugal). N_α,N_α-Bis(carboxymethyl)-L-lysine hydrate (I4580), L-Glutamine (49419) acquired from Fluka (Sintra, Portugal). ß- Arabic Acid (98%, A3006), N-acetyl-D-phenylalanine (A3876), Deuterium Water, Glutamic Acid (G1763), 3-methyl-1-Butanamine (98%), Propionaldehyde (P6889), Sodium Hydroxide (NaOH) 97% beads, Tryptamine were purchased from Sigma (Sintra, Portugal). Ethanol Absolute PA 99,5% 121086 from Panreac (Barcelona, Spain); N,N-Dimethylformamide (DMF) 99,5%, Hydrochloric acid 1mol/l, Methanol 99%, 2-Propanol 99.5% were from Roth (Navarra, Spain). The Sodium Chloride (NaCl) PA ACS ISO 99.5%, 131659.1211; Sodium di-Hydrogen Phosphate 1-hydrate, 131965; di-Sodium Hydrogen Phosphate 2-Hydrate 122507.1211; Tris (Hidroxymetil) Aminomethane 131940 99.8% were purchase from Panreac (Barcelona, Spain). Glycine 99.7% was purchased from Merck (Lisbon, Portugal).

For mammalian cells transfection was used Polyethylenimine, Linear, 25000Mw from Polysciences (Pennsylvania, USA) (Stock solution of 1mg/ml).

The protein quantification assay used was BCA kit Bicinchoninic Acid Solution from Sigma.

For the nucleic acid electrophoresis was used Agarose, SeaKem® LE Agarose, was purchase from Lonza (Basel, Switzerland). The loading buffer, BlueJuice™ Gel was purchase from Invitrogen (Barcelona, Spain). To stain the agarose gel was used Gel Red - GelRed™ Nucleic Acid Gel Stain, 10000X in water from Biotium, Inc. (California, USA)

For the SDS-PAGE the reagents used were: Acrylamide-Bisacryalamide was purchase from Biorad (California,USA). Ammonium Persulphate (PSA) and N,N,N,N-Tetramethylethylenediamine (TEMED) were from NZYTech (Lisbon, Portugal).

CHAPTER 2 – MATERIALS AND METHODS

Chromatographic and Screening Material

Cross-linked agarose, (Sepharose ^{TM} CL-6B 17-0160-01), was purchased from GE Healthcare (Carnaxide, Portugal). The Captiva 96 well 20µm polypropylene filtration plates (0.8x0.6cm), columns Empty bond elut reservoir, Reservoir-3ML Capacity and Bond elut clean polyethylene frits, Frits -3ML, 3/8'', 20µm were purchased from Varian (Porto, Portugal). The samples were analysed in different microplates: for UV measurements UV Star Plate 96 well Flat Bottom Half Area Greiner Bio-one (Frickenhausen, Germany) were used; for colorimetric assays the Microtest plate 96 well Flat bottom Transparent Polystyrene, Sarstedt (Nümbrecht, Germany); in the fluorescence studies, BRAND Plates – Imunogrades, Brand Tech Scientific (Carnaxide, Portugal) were used.

Biochemical Reagents and Cells

The hexapeptide NNNNNN 99.46% (w/w) was purchased from GeneCust (Rudelang, Luxembourg) and the Green Fluorescent Protein, rTurboGFP from Evrogen (Moscow, Russia).

The Strains of Bacteria used were Chemically Competent *E. coli* GT115 cells purchased from InvivoGen (Toulouse, France) and Library Efficiency® DH5α™ Competent Cells purchased from Invitrogen (Barcelona, Spain).

The Mammalian Cells used were Human Embryonic Kidney 293T cells (HEK 293T cells).

Table 2.1 – Vectors used during the work.

Vectors	Description	Source
pRWTagMGFP	GFP tagged with hexapeptide RW expression under control of CMV promotor.	GenArt (Regensburg, Germany)
p6NTag	It has the Affinity Tag with hexapeptide NN	GenArt (Regensburg, Germany)
p6NTagMGFP	Express GFP tagged with hexapeptide NN under control of CMV promotor.	Constructed during the thesis work
pMDISGFP	Express GFP under control of CMV promotor	----------

For Bacterial Growth, after transformation, the SOC Medium was used to grow bacteria; Amp (Ampicillin) Agar Solid Medium - Fast-Media® Amp Agar was used for clone picking and TB Amp Liquid Medium - Fast-Media® TB Amp that was used for plasmid production all were purchased from InvivoGen (Toulouse, France).

CHAPTER 2 – MATERIALS AND METHODS

For restriction of the DNA plasmids the following enzymes were used: EcoRI, HindIII, KpnI and NheI with appropriated buffer 1 and 2, was also used BSA. For the vector ligation were used: 10x Antarctic Phosphatase Buffer and Antarctic Phosphatase Buffer (5Units); T4DNA Ligase Buffer and T4DNA Ligase (4U/ µl) all were purchased from New England Biolabs (Massachusettes, USA).
For DNA purification the Maxi Prep Kit from Roche (Mannheim, Germany); Gel Band Purification Kit was purchase from GE Heathcare (Germany); Mini Prep Kit was purchased from Qiagen (Hilden, Germany) were used.
In the nucleic acid electrophoresis the DNA Ladders: DNA Ladder 1kb G5711 and PCR Marker G3161 were purchased from Promega (Wisconsin, USA).

For Mammalian Cells Growth and maintenance was used a Dulbecco's modified Eagles's medium (DMEM) – +4,5g/l Glucose; +L-Glutamine; + Pyruvate Gibco®; supplemented with 10% (v/v) of Fetal Bovine Serum (FBS) Gibco®; Dulbesco's Phosphate-Buffered Saline 1x (DPBS) 1X - (-)$CaCl_2$; (-) $MgCl_2$ and Trypsin, 0.05% 1X that were purchase from Invitrogen (Barcelona, Spain).

The protein marker used in the SDS-PAGE was Low Molecular Weight SDS Kit, NZYTech (Lisbon, Portugal).

Equipment

A Rotisserie oven – Big Shot III Hybridization Oven Model 230402, Boekel Scientific (Pennsylvania, USA) was utilized for the synthesis of the combinatorial libraries. Laboratory Medical Centrifuge LMC-3000, BioSan (Riga, Latvia). The Microplate Reader – Tecan Infinite F200 from Tecan (Männedorf, Switherzland) was used for all spectrophotometric and spectrofluorometric assays. To determine the melting point was used Reichert Thermovar Thermometer (30ºC-230ºC), Austria.

Software

Molecular Modelling and Docking studies were performed with Molegro Virtual Docker software from Molegro ApS – Bioinformatics Solutions (Aarhus, Denmark), Marvin Sketch 4.1.6. Pre1 from ChemAxon, PyMol DeLano Scientific LLC (California, USA) and ChemBioDraw Ultra 12.0 CambridgeSoft, UK .

2.2 Methods

2.2.1 Synthesis of a Ligand Library based on the Ugi Reaction

2.2.1.1 Epoxy Activation of Sepharose CL-6B

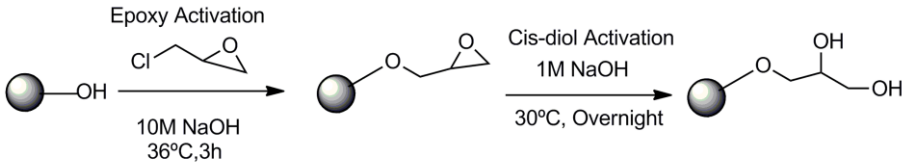


Figure 2.1 - Reaction of Activation of the Sepharose CL-6B with epoxy group and preparation for activation with an aldehyde group.

Initially the required amount of Sepharose CL-6B was washed with distilled water (10 x volume of the resin's weight). The sample was filtered under gravity without applying vacuum. After this, 7,2ml of NaOH 10 M (with a proportion of ~ 40 ml per 1000 gram of gel) was added and incubated at 30 ºC during 30 min with orbital agitation. In order to introduce the epoxy group in the gel, 12,96 ml of epichlorohydrin (in concordance with the optimize proportion 72 ml/ 1 Kg of gel) was also added to the gel, and then incubated during 3 h at ~36ºC with constant stirring (Figure 2.1). After this period, the mixture was washed thoroughly with distilled water, being the first litre of solution taken into the chlorinated solvents container. The epoxy group content was determined as described in 2.2.12.1.

2.2.1.2 Cis-diol Activation of the Epoxy-activated Support

A volume of 160 ml of 1 M NaOH (1 ml/g gel) was added to 160 g of epoxy-activated gel (17 μmol of epoxy/g of gel). and then incubated overnight at 30ºC (Figure 2.1). After this, the gel was washed as described in 2.2.1.1.

2.2.1.3 Aldehyde Functionalization of the Activated Support

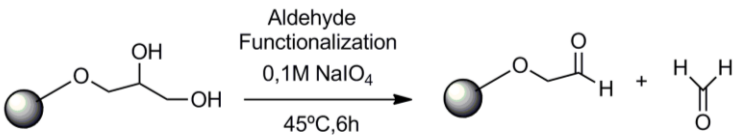


Figure 2.2 - Reaction of the functionalization of the support with aldehyde group.

The aldehyde functionalization was performed by adding 160 ml of 0.1 M $NaIO_4$ (with the proportion of 1 ml of solvent / 1 g of gel) and then incubated during 6h at 45ºC with constant stirring (Figure 2.2). After functionalization was necessary to proceed to the washing as described in 2.2.1.1. The presence of aldehydes was detected using the Tollens assay (2.2.12.2).

2.2.1.4 Solid-phase Synthesis of Ugi Library

Subsequent to the activation of the support with the Aldehyde group, one of the compounds of the Ugi reaction was possible to add the other three compounds to the reaction Figure 2.3.

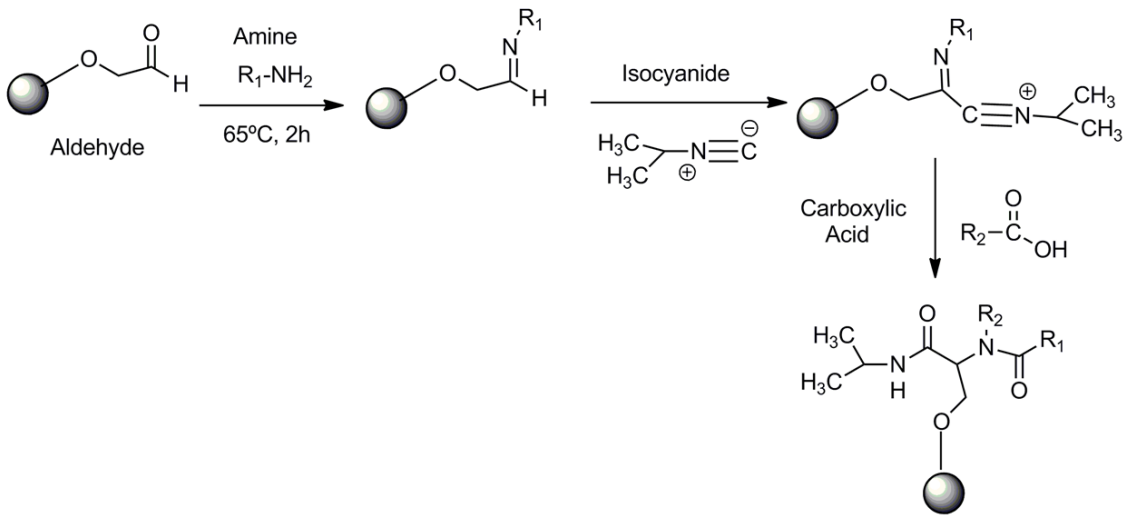
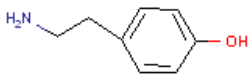
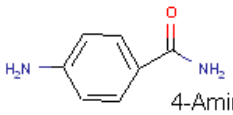
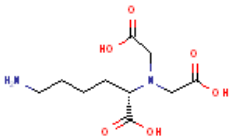
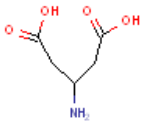
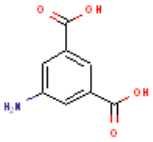
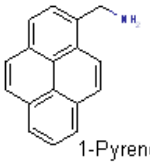
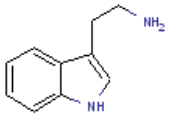


Figure 2.3 - Ugi Reaction, the functionalized support with aldehyde is the base of the reaction; the second compound is an amine that reacts for two hours to let the formation of the imine group. After that the isocyanide and the carboxylic acid is added. Reaction continues for 48h.

The aldehyde functionalized gel was washed with several concentrations of methanol, from 20% (v/v) to 100 % (v/v) methanol, at 20% increments. This procedure is made to improve the dissolution of apolar groups in support that is composed of ~ 90% of water. Once the resin was washed, 35 ml of 100% methanol (v/v) was added to the resin. The slurry was divided in 0.4ml (corresponding to 0,250mg of resin) per each well of the reaction block Captiva 96-well (were used only 64 wells). The flexible end cap of the reaction block was removed in order to drain the solvent and to settle the resin. After this, the end cap was again placed in the bottom and it was possible to start adding the Ugi compounds, (5 molar excess of epoxy groups) dissolved in methanol for a total volume of 250 μl/per well.

A)

Amine Compounds for Ugi Reaction Library			
N°	Compound Name and Structure	N°	Compound Name and Structure
A1	 Tyramine	A5	 4-Aminobenzamide
A2	 Nα,Nα-Bis (carboxymethyl-L-Lysine Hydrate)	A6	 1,4-Diaminobutane
A3	 ß– Glutamic acid	A7	 5 – Aminoisophthalic acid
A4	 1-Pyrenemethylamine hydrochloride	A8	 Tryptamine

B)

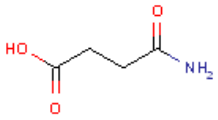
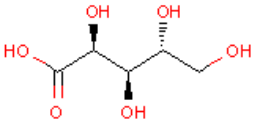
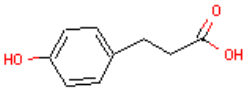
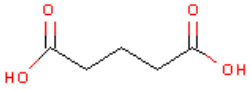
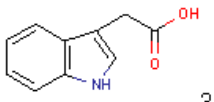
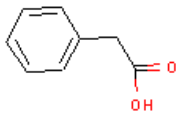
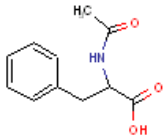
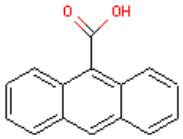
Carboxylic Acids Compounds for Ugi Reaction Library			
N°	Compound Name and Structure	N°	Compound Name and Structure
C1	 Succinamic acid	C5	 Arabic acid
C2	 3-(4-Hydroxyphenyl) propionic acid	C6	 Glutaric Acid
C3	 3-Indoleacetic acid	C7	 Phenylacetic acid
C4	 N-acetyl-D-phenylalanine	C8	 9-Anthracenecarboxylic acid

Figure 2.4 - Chemical Structures of the A) Amines and B) Carboxylic Acids for Ugi Reaction Library.

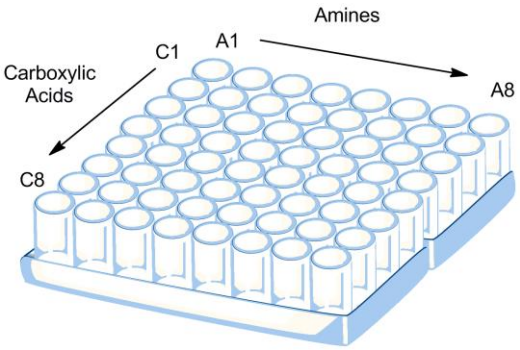


Figure 2.5 - Ugi Combinatorial 64 Ligands Library. Strategy to add Amines and Carboxylic acids to construct the library.

The first compound added was the amine, being the addition done in each column of the reaction block from A1 (first column) until A8 (last column) (Figure2.5). Amines A2, A3 and A7 possess both amine and carboxylic acid functionalities (Figure 2.4), and for this, it was necessary to block the carboxylic acid with NaOH (3 eq. the amount of the carboxylic acid (A2) and 2 eq. for A3 and A7, considering the three and two COOH groups of the molecule, respectively). The aldehyde and amine compound were incubated for 2h at 60ºC, in order to guarantee the formation of the imine compound.

The following compound added was the isopropyl isocyanide, being this addition performed in all wells (16µl per well in an excess of 8 eq.).

The last compound, carboxylic acid compound, was added with the same volume 250 µl / well). The addition of this compound was performed in this way due to the fact that the carboxylic acid compound was kept constant in each row (C1 in first row and C7 in last row) of the reactional block (Figure 2.5).

All these quantities were determined considering the molecular weight of each compound and the extent of epoxy activation of the agarose that as determined in 2.2.12.1.

After the addition of all components, where only two compounds involved were varied, the upper cap was fixed in the top of the reaction block. Then, the reaction block was incubated at 60ºC for 48h in a rotisserie oven to ensure a good mixing of the compounds.

At the end of this reaction, the end cap was carefully removed to drain each well and then a washing procedure was performed to guarantee that all unreacted reagents were removed. Each well was washed with 1 ml per each step as follows: (1) 100 % (v/v) methanol (2) 50% DMF : 50 % methanol (v/v) (3) 50% (v/v) DMF (4) water (5) 0.1 M HCl (6) water (7) 0.2 M

NaOH in 50 % (v/v) isopropanol (8) 2x water and (9) 20% (v/v) ethanol. Then the washed beads were stored with 20% (v/v) ethanol at 4ºC.

In parallel, a positive control ligand was also synthesized, where the amine compound was the 1-Pyrene methylamine hydrochloride (pyrene). From the same batch of agarose functionalized with the aldehyde as described in 2.2.1.4, 1g of the functionalized agarose was taken for a reactional flask and then 1 ml of methanol was added to perform the synthesis of the ligand. The four components of the ugi reaction were added (5 molar excess of epoxide groups) dissolved in 1 ml of methanol. In first, the amine was added to the slurry, followed by the addition of the isopropyl isocyanide and in the last a random carboxylic acid, 4 hydroxybenzoic acid, was added.

To analyse the fluorescence on the beads, 500mg of blank agarose and 500mg of gel functionalized with pyrene were washed several times with deionised water, and then 1ml of deionised water was added to the resin. Samples with 200µl of Agarose solution and Agarose functionalized with pyrene were taken and the fluorescence intensity on the agarose beads was measured using the microplate reader fluorescence filter (excitation 485(20)nm and emission: 535(25)nm).

2.2.2 Synthesis of Ligands Library based on the Triazine scaffold

2.2.2.1 Amination of the epoxy-activated support

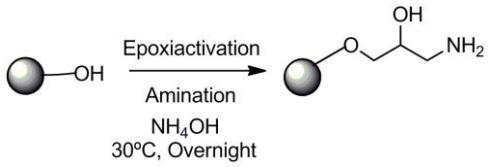


Figure 2.6 – Activation of the Sepharose CL-6B with epoxy groups and amine groups.

The epoxy-activated agarose beads (as in 2.2.1.1 and in 2.2.12.1; 29 µmol of epoxy/g of gel) were aminated by using the optimal conditions of 1.5ml of Ammonia hydroxide solution (NH_4OH) per gram of gel. The slurry was incubated overnight at 30ºC wit orbital agitation (Figure 2.6). After this, the aminated agarose beads were thoroughly washed (about 10x gel volume) with distilled water and the amine groups content was determined (see 2.2.12.3).

2.2.2.2 Triazine Library Synthesis

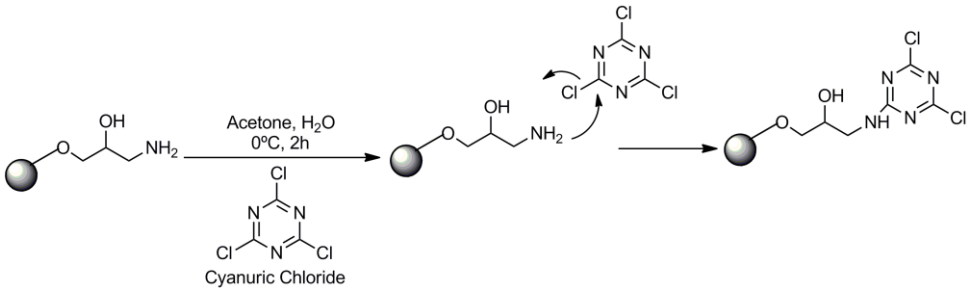


Figure 2.7 – Activation of the 1,3,5-Triazine based ligands, by addition of the cyanuric chloride.

To the aminated gel (25g, 20 µmol of Amines/g of gel), 25ml of 50% (v/v) Acetone/Water mixture were added and then aminated at 0ºC in an ice bath shaker. After this, cyanuric chloride (5 molar excess to the aminated groups) was dissolved in 3.96ml of Acetone (8.6ml Acetone/g of Cyanuric Chloride) and added to the agarose slurry with constant shaking at 0ºC and with maintenance of neutral pH. The reaction occurred at 0ºC for 2h (Figure 2.7). The gel was washed with a mixture of water:acetone (v/v)with the following ratios: 1:1,1:3,0:1 and then 1:1, 3:1 and 1:0 until reach 100% water. The slurry was divided in 0.4ml portions (corresponding to 250mg of resin) and dispersed in the individual wells of the reaction block Captiva 96-well (utilized 64 wells).

Amine Compounds for Triazine Library

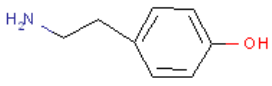
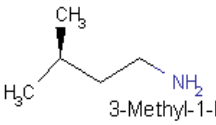
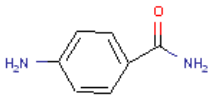
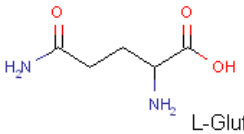
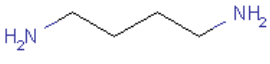
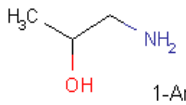
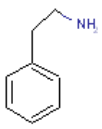
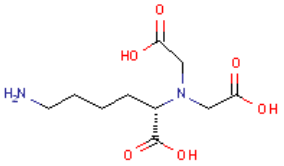
N°	Compound Name and Structure	N°	Compound Name and Structure
1	 Tyramine	5	 3-Methyl-1-butanamine
2	 4-Aminobenzamide	6	 L-Glutamine
3	 1,4-Diaminobutane	7	 1-Amino-2-Propanol
4	 Phenethylamine	8	 Nα,Nα-Bis (carboxymethyl-L-Lysine Hydrate)

Figure 2.8 - Chemical structure of the amines used in the synthesis of the Triazine library.

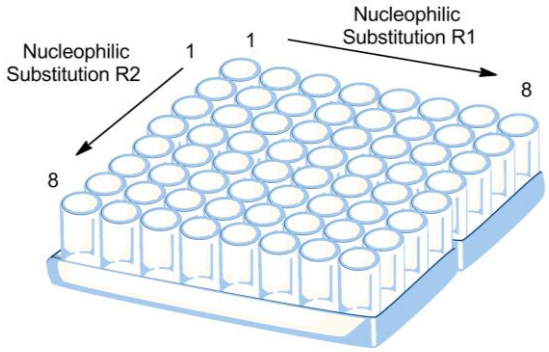


Figure 2.9 - Triazine Combinatorial 64 Ligands Library. Strategy to add amines in the Nucleophilic Substitution R1 and Nucleophilic Substitution R2.

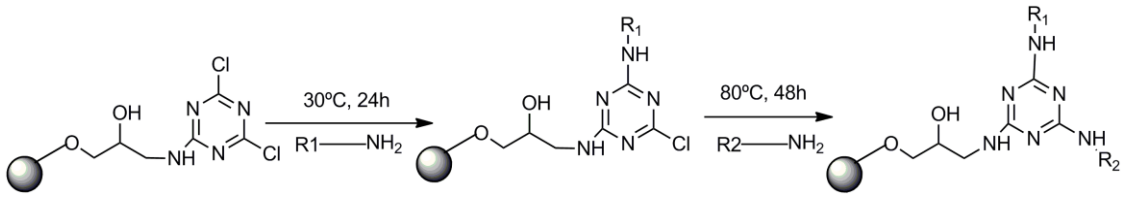


Figure 2.10 – Nucleophilic Substitutions of the positions 1 and 3 of the Triazine ring.

For the Nucleophilic Substitution R1

All the amines were used with a 2 molar excess relative to the amount of amine-groups and dissolved in 50%DMF:50%Water(v/v) in a total volume to give 250µl/per well. Amines 6 and 8 possess both amines and carboxylic acid functionalities (Figure 2.8), and for this, it was necessary to block the carboxylic acid wit NaOH (2molar excess the quantity of the carboxylic acid, considering the number of the COOH groups of the molecule). Each amine was added only in the columns of the reaction block, starting with the addition of the A1 amine in the first column until the A8 in the last column (Figure 2.9). Then the reactional block was incubated at 30ºC for 24h with orbital agitation (Figure 2.10). After this, several washes in each well were performed, starting with 1ml of 50%DMF:50%Water(v/v) and then four times with 1ml of distilled water in each well.

For the Nucleophilic Substitution R2

All the amines were used with a 5molar excess relative to the amount of amine-groups and dissolved in 50%DMF:50%Water(v/v) in a total volume to give 250µl/per well. Amines 6 and 8 possess both amines and carboxylic acid functionalities, and for this, it was necessary to

block the carboxylic acid wit NaOH (5molar excess the quantity of the carboxylic acid, considering the number of the COOH groups of the molecule). Each amine was added only in the columns of the reaction block, starting with the addition of the A1 amine in the first row until the A8 in the last row. Then the reactional block was incubated at 80ºC for 48h with orbital agitation. In the end of the R2 activation, several washes were performed in each well, starting with 2X 1ml of 50%DMF:50%Water (v/v); 2x Water; 0.1M HCl; Water; 0.2M NaOH in 50%(v/v) Isopropanol; 2x Water and 20% Ethanol. Then the block was stored with 20%Ethanol at 4ºC.

In parallel, a negative and a positive control were synthesized. For the synthesis of the controls 1g of cyanuric chloride-activated gel was used. The negative control or 0/0 ligand was synthesized by adding 1ml of ammonium hydroxide and 1ml of water for both R1 and R2 substitution. The positive control was synthesized by adding 1-pyrenemethylamine hydrochloride as amine compound for both R1 and R2 substitutions. Then, fluorescence studies were performed with control ligands as was previously described 2.2.1.4.

2.2.3 Ligands Screening with Hexapeptide NN and GFP

The prepared combinatorial libraries of ligands were screened to determine the affinity of each ligand to the hexapeptide NN and the protein GFP.

2.2.3.1 Screening with Hexapeptide NN

The supports of the Ugi ligands and the Triazine based ligands were washed with 2x1ml distilled water, followed by 3x1ml regeneration buffer (0.1M NaOH in 30%(v/v) isopropanol) alternated with 3x1ml distilled water. Subsequently, 10x1ml of equilibration buffer (PBS buffer: 10mM sodium phosphate, 150mM NaCl, pH 7.4) were added in order to equilibrate the columns. The columns were washed with equilibration buffer until the absorbance of the samples reached $A_{280nm} \leq 0.005$ (approximately 10x1ml).

The peptide was dissolved in equilibration buffer to yield a 1mg/ml solution. The peptide solution (250µl) was loaded in each well of the 96-well filtration plate. Then the blocks were incubated at 25ºC for 20min with agitation.

The first fraction of unload peptide was collected in a 96-deep well collection plate, the block with the collector was centrifuged at 10000rpm for 20sec in the Centrifuge LMC-3000 BioSan.

After this the wells were washed with equilibration buffer for 8x250 µl. Then was made the same procedure described for the first fraction.

The bound peptide was eluted by adding 5x250 µl of elution buffer (0.1M Glycine-HCl, pH3.5). 50 µl of Tris-HCl pH9 were immediately added to each fraction. After the elution step the columns were washed with regeneration buffer (5x1ml), followed by distilled water (5x1ml) and stored at 4ºC in 20% (v/v) ethanol. The factions loaded, washed and eluted were analyzed by the BCA assay (2.2.12.4) using a calibration curve Y=0.3671X+0.0593, R^2=0.99.

2.2.3.2 Screening with Green Fluorescent Protein

The blocks were regenerated in between the screening of the hexapeptide and the GFP by washing the wells as in (2.2.3.1).

For the screening of GFP a solution of $1.75x10^{-2}$mg/ml in equilibration buffer (250µl) was loaded in each well of the 96-well filtration plate. Then the blocks were incubated at 25ºC for 20min with agitation.

The washing and the elution procedures were repeated as described for the hexapeptide (2.2.3.1). From each sample were analysed (150µl) by fluorescence in wavelengths of 485nm (excitation) and 535nm (emission) in the microplate reader to determined the concentration of GFP in the samples (Y=2E6*X-1200.7, R^2=0.99) using a calibration curve of GFP between 2.5E-3 to 1.75E-2mg/ml.

2.2.4 Scale-up Synthesis and Screening of the Lead Ligands

After screening of the combinatorial libraries (Ugi and Triazine libraries) with the hexapeptide NN and GFP, some ligands were selected for further studies.

From previous work with another hexapeptide RW some ligands from the Ugi library were selected for studies with cellular extracts.

The possible lead ligands for binding to hexapeptide NN and RW were synthesized in larger scale (5g). The ligands synthesized are summarized in Table 2.2.

Table 2.2 – Selected lead ligands for the hexapeptide NN and RW.

Combinatorial Library	Ligands	Target
Ugi	A3C1, A6C1, A5C2, A1C5, A6C5, A7C6, A3C7, A6C7	Hexapeptide NN
Triazine	6,3; 8,3; 7,4; 8,6	Hexapeptide NN
Ugi	A4C3, A6C3, A7C3, A4C4, A4C5, A4C6, A4C7, A7C8	Hexapeptide RW

The resin for the Ugi ligands synthesis was activated using the procedure as described before (see 2.2.1.1. to 2.2.1.4) The Ugi ligands for hexapeptide NN had 41µmol epoxy groups/g gel and the ligands hexapeptide RW a 20µmol epoxy groups/g gel. Each compound in the Ugi Reaction was added in the appropriated proportion to the molar ratio of epoxy groups in the resin. In a 5ml of Methanol each compound was added to the reaction flask. In the end of the reaction the slurry was washed with a volume of 5ml as described before (2.2.1.4). After that were collected in to a flask and stored at 4ºC and 20% Ethanol.

The Triazine ligands of hexapeptide NN were synthesized with the appropriate molar excess, knowing that the amination activation was 17µmol amine groups/g support. The components of the reaction were dissolved in 5ml of 50%DMF: 50%Water. When the reaction ended the slurry was washed with a volume of 5ml as in 2.2.2.2. After that, the agarose functionalized was collected into a flask and stored at 4ºC and 20% Ethanol.

2.2.4.1 Scale-up Screening of the Lead Ligands with the Hexapeptide NN

The lead ligands for hexapeptide NN were screened for binding to the peptide. 500mg of agarose (~750µl of slurry) modified with each ligand was packed in a column. The supports were regenerated as described for the 96 wells block (2.2.3.1).

After equilibration of the adsorbent with equilibration buffer 500µl of a hexapeptide NN solution of 1mg/ml was added to each column. The columns were incubated at 25ºC for 20min, after each washes and the elution steps were performed by adding 500µl of the required solution (see 2.2.3.1).

From each step was taken a sample of 25µl to analyze using the BCA assay (2.2.12.1), with a calibration curve of $Y=0.373X+0.079$, $R^2=0.99$.

2.2.5 Molecular Modelling Studies - Possible Lead Ligands *in silico*

The lead ligands for the hexapeptideNN (A3C1, A6C1, A5C2, A1C5, A6C5, A7C6, A3C7 and A6C7) were analyzed *in silico* for interactions with: the Affinity Tag composed by hexapeptideNN – Three Prolines - Sequence for Recognition of Enterokinase; and, the subunit A of GFP (PDB structure of Wild Type GFP of *Aequorea Victoria* – 1W7S).

The construction of the ligands and the Affinity Tag was performed in the program *Marvin Sketch*. The structures were edited in 3D clean and saved with a *.pdb extension. The GFP structure was downloaded from PDB Protein Data Bank site with *.pdb extension. To download only the structure of the subunit A it was necessary to enter in the CATH site and choose the subunit of interest.

The molecules saved as *pdb files were imported for the docking software (Molegro Virtual Docker) work space and prepared for the docking (File>*Import Molecules*>*Preparation (Always)*). For the protein it was first necessary to apply the *Detect Cavities* option. Then was used the option, *Docking Wizard*, to start the process of docking between the ligands and affinity tag or the subunit A, where was selected *Score* (*Moldock Score* and the *Radius* of interaction was optimized) and then selected the option *Algorithm (Moldock Optimizer; Number of Runs: 10; Population number: 200);* the results were recorded in a specific folder. In the end of the running in the option *Settings* (select the Energies: *Interaction and HBond*>*Recalculate the Energies*); then were *Select the Poses and Energies*>*OK*; and in the *Pose Organizer* selected the *Rerank Score* with the lowest energy, since this value gives idea of the energy envolved in the affinity between the elements of the complex.

In the results it was possible to evaluate the affinity of each ligand and the protein (Affinity tag or Subunit A GFP) based on the free energy and the interactions (Hydrogen Bond and Hydrophobic interactions) that are formed between the molecules.

2.2.6 Solution Phase Synthesis and Characterization of Ugi Ligand, A6C5

In a flask was added 79 µl of Aldehyde (Propionaldehyde), 2.5ml of methanol and 111µl of Amine (1,4-Diaminobutane), and left incubating for 2h at 60ºC, in order to guarantee the formation of the imine compound. After that, 100µl of Isocyanide (Isopropyl Isocyanide) and the solution of Carboxylic acid (219mg of Arabic Acid in 3.5ml of methanol and 17µl of NaOH) were added. The compounds were let to react for more 48h at 60ºC.

The white solid formed was washed with small volumes of methanol and filtered and left in the desiccator for further drying.

After the synthesis was necessary to characterize the crystalline solid. To this purpose was made the TLC technique and determined the melting point of the solid (Solomons and Fryhle 2004). In addition was analysed a sample in 1H (10mg) and 13C NMR (30mg) this sample was dissolved in 0.5ml of Deuterium Water (D_2O).

2.2.7 Ligands Nomenclature

The Ugi ligands in this work are represented by AxCy, A-Amines and C-Carboxylic Acid, following the nomenclature normally use in this tip of works (Haigh, Hussain et al. 2009). In the case of the Triazine ligands are represented by numbers that correspond to the amines envolved in the Nucleophilic Substitution: Amine x (R1), Amine y (R2).

2.2.8 Expression of GFP fusion proteins in Mammalian Cells

2.2.8.1 Amplification of the vectors pRWTagMGFP and p6NTag

2.2.8.1.1 E.coli GT115 Competent cells Transformation

The protocol was followed as described by the supplier (InvivoGen). After incubation in the plates with solid medium Agar Amp, the six colonies were picked using a pipette tip that was used to inoculate 3ml of TB Amp medium and were cultivated overnight at 37ºC with agitation at 180rpm. In the next day were determined the optic density of the cultures clones at 600nm (OD_{600nm}), the clone selected to inoculate the culture in 200ml of TB Amp, had an OD_{600nm} higher than 5.

The 200ml TB Amp medium was inoculated with 200µl of the selected clone bacteria suspension. This culture was incubated overnight at 37ºC with agitation at 180rpm. In the next day, the OD_{600nm} was determined and normally was between 2 and 6. From this culture was took a sample of 8,5ml was added 1,5ml of sterile Glycerol to make a thaw solution that was divided in cryovials and stored at -85ºC.

Remain suspension of the start culture was used for purification of the plasmid DNA from *E.coli* GT115.

2.2.8.1.2 Isolation and Purification of cDNA from E.coli

To isolate the plasmid DNA was used the Maxi Prep Kit (Roche) and the procedure for High copy Number Plasmid was followed as described by the supplier. The plasmid DNA was resuspended in 300µl of TE buffer pre-warmed at 37ºC.

2.2.8.1.3 Spectrophotometric Quantification of cDNA

The proper dilutions of the plasmid DNA sample were made, to a total volume of 50µl, to determine the absorbance at diferents wavelengths: 260nm, 230nm, 280nm and 320nm using a Spectrophotometer. The concentration of the plasmid DNA was determined and also it purity, by the quotients of absorbance: 260nm/230nm and 260nm/280nm. The purified plasmid DNA sample was stored at -20ºC.

2.2.8.1.4 Enzymatic Modifications of cDNA

The volumes of enzyme, BSA, plasmid, Pure Water and appropriated Buffer were calculated to digest 1µg of plasmid.

To digest pRWTagMGFP the following restriction enzymes were used: EcoRI and HindIII. The enzymes were used in Buffer 2, where they have the highest enzyme activity.

Three different experiments were made: EcoRI; HindIII and EcoRI + HindIII. In the experiment with only one enzyme was used 1µl plasmid DNA, 2µl Buffer 2, 1µl Enzyme and the volume was complete to 20 µl with sterile water; in the experiments of two enzymes were added 1µl of each enzyme and the volume was completed with sterile water.

The enzymatic reaction occurred at 37ºC for 90min with agitation.

2.2.8.1.5 Electrophoresis

The restriction samples were analysed in agarose gel to confirm the presence of the expected fragments.

The electrophoresis gel was made by mixing agarose with the appropriated volume of TAE 1x buffer to make the agarose total percentage of 1 – 2%. The electrophoresis was let to run for 90min at 90V.

To reveal the gel it was incubated in the dark for 45min with Gel Red 1x (diluted in TAE buffer) and observed in a UV transilluminator.

2.2.8.2 Construction of vector p6NTagMGFP

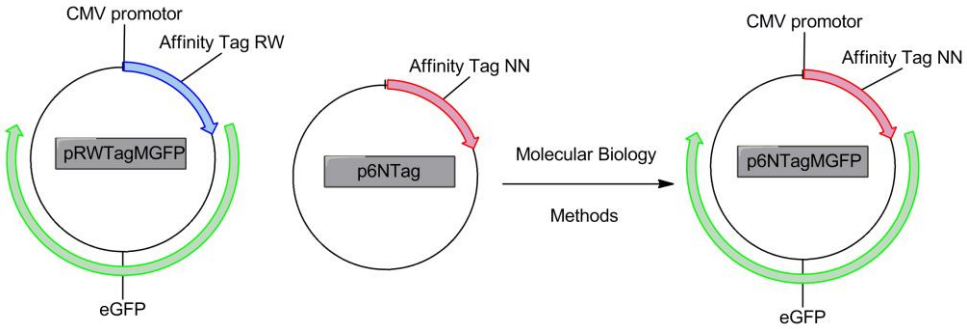


Figure 2.11 - Construction of the vector p6NTAgMGFP, that expresses in mammalian cells the fused protein GFP tagged with hexapeptide NN. To construct this vector were used the pRWTagMGFP and the p6NTag vectors.

2.2.8.2.1 Extraction and Purification of DNA from Agarose Gels

After the digestion of the plasmids p6NTag and pRWTagMGFP with NheI and KpnI an electrophoresis gel was made. To this gel two DNA markers were added: PCR Marker and DNA Ladder 1kb. In the gel, one well had a small sample of the restriction reaction (2 µl) for analysis purpose and the other one had the all sample of the restriction reaction (16 µl) to recover. The Electrophoresis was made in the same conditions as described before (2.2.8.1.5).

In the end the gel was cut in four fragments: one for pRWTagMGFP (with two markers and a sample with the restriction reaction with NheI and KpnI), other with all sample of pRWTagMGFP; one for p6NTag (with two markers and a sample with the restriction reaction with NheI and KpnI) and other with all sample of p6NTag. These two fragments with restriction sample and markers were incubated in the dark for 30min with Gel Red.

These two fragments were compared with the other fragments that were not revealed with Gel Red. By comparing the two gel fragments for the same plasmid digestion reaction (small sample and all sample) was possible to select and cut the bands in agarose gel that had the fragments of interest. In the plasmids the fragments of interest were: for pRWTagMGFP 4126bp and for p6NTag 78bp.

Following to the recovery of the bands from the agarose gel was necessary to purify the fragments of DNA from the gel. For that purpose was necessary to use a mini prep, Gel Band Purification Kit. The elution of the DNA was made by addition to the purification column of 60µl of sterile water.

The concentration and purity of the fragments was determined in 50µl of sample (2.2.8.1.3).

2.2.8.2.2 Dephosphorylation of a Vector

The all sample of purified DNA vector was dephosphorylated by adding 5µl of 10x Antarctic Phosphatase Buffer and 1µl of Antarctic Phosphatase Buffer (5Units). After mixing it was incubated for 60min at 37ºC.
Following the enzyme was heat inactivated at 65ºC for 10min. The sample was put at RT and subsequently was made the ligation of the vector with the insert.

2.2.8.2.3 DNA Ligation

To the dephosphorilation mixture was added 5,9µl of T4DNA Ligase Buffer, 2 µl of T4DNA Ligase (4U/ µl) and 1 µl of DNA insert. The reaction was incubated overnight at RT.
In the next day the enzyme was heat inactivated at 65ºC for 20min.

2.2.8.2.4 Amplification of the p6NTagMGFP

In this case, the product of the ligation was used for transformation of Library Efficiency DH5α^{TM} competent cells the volume that was added was 5µl (10ng) to 100µl of cells. The transformation protocol was followed as described by the supplier (Invitrogen).
Subsequently to the transformation the mixture was spread (100 µl) in Agar Amp medium plates and incubated overnight at 37ºC.
In the next day there were some colonies, they were picked with a pippet tip that was used for inoculation of culture tubes with 3ml of TB Amp medium that were incubated (2.2.8.1.1).

2.2.8.2.5 Purification in small scale of the plasmid DNA and confirmation of the correct ligation

The protocol of the supplier for MiniPrep kit (Qiagen) was followed. In the end was determined the concentration and purity of the plasmid (2.2.8.1.3).
After this procedure the plasmid DNA was fragmented with the appropriated restriction enzymes (KpnI and NheI). The restriction reaction occurred in the same conditions as 90min at 37ºC with agitation (2.2.8.1.4).
The result samples were added to a gel of agarose 2% with appropriated DNA markers (2.2.8.1.5).
Due to the confirmation of the plasmid p6NTagGFP the amplification of the plasmid continued (2.2.8.1.1), as well as the purification and quantification of the DNA (2.2.8.1.2 and 2.2.8.1.3).

2.2.8.2.6 Enzymatic Modifications of plasmid DNA

The same procedure was used for this plasmid and even the same restriction enzymes (KpnI and NheI), although to confirm the presence of only one insert in the plasmid was necessary to cut it with Hind III, that has a restriction site in the middle of the insert.
The restriction reaction occurred as (2.2.8.1.4). The samples of the restriction reaction were added to an agarose gel and the electrophoresis occurred as described before (2.2.8.1.5)
From the start culture was took a sample of 8.5ml to make a stock of cells with the p6NTagMGFP (2.2.8.1.1).

2.2.8.3 Amplification of vector pMDISGFP

To start the amplification of this plasmid was used a stock sample stored at -85ºC. The solution was partially thawed and a sample was spread in a plate of Agar Amp medium that was incubated overnight at 37ºC.
Than was followed the same procedure that was described for amplification (2.2.8.1.1) and purification (2.2.8.1.2) of the pRWTagMGFP and p6NTag.

2.2.9 Transfection of Mammalian cells with pRWTagMGFP, p6NTagMGFP and pMDISGFP

The mammalian cells were subcultured regularly and grown in an adherent monolayer in T-flasks at 37ºC in a humidified 5% CO_2 atmosphere. The cells were grown in a mixture of DMEM medium supplemented with 10% of FBS. To subculture the cells (80-90% confluence) Trypsin was used, to detach the cells from the support, and passed the appropriated volume to a new T-flask. When was necessary to determine the cell concentration and viability a hemocytometer was used to count cells stained with trypan blue dye exclusion.
The day before transfection of the HEK 293T cells, a petri plate (58.1cm^2) was inoculated with the $5x10^4$ cells/cm^2 (~60% confluence) and was incubated with DMEM at 37ºC and 5%CO_2.
At the day of the transfection, the medium was changed 2h before the transfection and the transfection mixture was prepared. The required quantity of DNA ($[DNA]_f$=0,0167µg/µl) was added to the required volume of DMEM without FBS (V_{plate}=10ml). The required amount of PEI (3 PEI: 1 DNA) was added to the mixture drop-wise and mix in the vortex. The mixture

was incubated at RT for 10min to let the formation of DNA:PEI complex. The mixture was added dropwise along the plate to distribute the complex by the culture. The plate was incubated at 37ºC and 5% CO_2 for 48h. The plate was observed at 24h to determine the confluence of the culture and efficiency of the transfection.

After 48h of transfection the plates were observed and the cells were collected. To collect the cells from the petri plates a cellscraper was used. First the medium was removed and to each plate was added 500µl of PBS and the cells were scraped from the dish. The suspension was collected to a flask and the cells were resuspended in the total volume of 3ml PBS. These samples were stored at -20ºC.

2.2.10 Extraction and Quantification of Proteins from Cellular Extracts

To extract the proteins from the cells a freeze-thaw method was used. The cellular extracts were frozen in Liquid Nitrogen and than thawn in a water bath at 37ºC. Three cycles of freeze and thaw were made.

The samples were divided in 1ml alliquots. These samples were centrifuged at 12851g, 4ºC for 15min. The supernatant was collected and stored at -20ºC.

2.2.10.1 Quantification of Proteins in the Cellular Extracts

The total protein content of the samples obtained in 2.2.10 using the BCA Kit, was determined by using BSA as a standard protein (calibration curve 0.2 to 1mg/ml ($Y=0.914X+0.106$, $R^2=0.99$)). Samples were diluted 1:10 and 1:100 in a total volume of 25µl. The assay was prepared as described (2.2.12.4).

In the samples containing GFP, the concentration of this protein was determined based on the fluorescence intrinsic property. The samples were diluted in 1:10 and 1:100, calibration curve was the same used as previously described (2.2.3.2)

2.2.11 Screening of Lead Ligands with Cellular Extracts

The scale-up ligands were screened with three different samples of cellular extracts: GFP tagged with an Affinity Tag (hexapeptide NN or hexapeptide RW); Positive Control (GFP expressed in mammalian cells, but without Affinity Tag); and, Negative Control (mammalian cells cellular extracts without GFP).

For each ligand three columns were packed with 500mg of support (~750µl of slurry).
Subsequently was initiated the regeneration process of agarose for each column, using the procedure described for the 96 wells block (2.2.3.1).
The solutions of GFP tagged with Affinity Tag (hexapeptide RW or hexapeptide NN) and Positive Control were prepared for having a concentration of GFP 1.75E-2mg/ml. The solutions of Negative Control were prepared to have a similar quantity of Total Protein as the GFP tagged solutions. The negative control for hexapeptide RW was prepared with a concentration of 1.48mg/ml for the hexapeptide NN was prepared a solution with 0.435mg/ml.
The solutions were added (500µl) to the columns. The columns were incubated at 25ºC and for 20min, followed the washes and elution procedure as described before (2.2.3.1).
From each step was taken a sample of 25µl to determine, using the BCA assay, the total protein using BSA protein as reference (2.2.12.4). Also was took a sample of 100µl to analyze the fluorescence of GFP in the samples that had this protein (2.2.1.3.2).

2.2.11.1 SDS-PAGE of the Purified Samples

The samples that had higher total protein concentration (first wash with PBS) were analysed in SDS-PAGE. The gel was polymerized for a 12.5% of Acryalmide. The samples from hexapeptide RW ligands were more concentrated, and from each sample 8µg of Total Protein (15µl) were added per well. The samples of hexapeptide NN ligands were more diluted and 20µl of each sample (without dilution) was added per well. Before loading the samples in the gel, they were mixed with loading buffer (dye, SDS and ß-Mercaptoethanol) and were denatured at 100ºC for 2 min.
The samples were run in the gel for 90min at 150V and 25mA. For detection of the protein bands, the gel was incubated with Coomassie Blue solution for 30min with agitation and discoloured overnight (Roskams and Rodgers 2002).

2.2.12 Analytical Assays

2.2.12.1 Determination of the Epoxy content

The extension of epoxy activation was determined by adding 3 ml of 1.3 M sodium thiosulfate into 1 g of epoxy activated gel and then incubated during 20 min at room temperature. This mixture was neutralized by 0.1 M hydrochloride acid. The volume used to reach pH 7 corresponds to the number of moles of OH released (10 μmoles per each 100 μl added), which equals to μmole epoxy-groups/g gel.

2.2.12.2 Tollens Test for the Qualitative Analysis of Aldehydes

To 1g of agarose gel from 2.2.1.4. was added 1 ml of freshly prepared Tollens reagent. The presence of a silver mirror or a black precipitate revealed the presence of aldehydes.
The positive control was 1 ml of glutaraldehyde and the negative control was 1 g of Sepharose CL-6B unmodified. The Tollens reagent was added to the controls as described. The Tollens reagent was freshly prepared by adding into a test tube, (previously cleaned with 3 M sodium hydroxide) 2 ml of 0. 2 M silver nitrate and a drop of 3 M sodium hydroxide. After this, it was added drop by drop with a constant shaking a 2.8 % ammonia solution until almost all precipitate of silver oxides has dissolved. In order to remove all of the obtained precipitate, a 8.8% ammonia solution was added dropwise.

2.2.12.3 Determination of the Amine content by the Kaiser test

The extension of the amination was determined by the Kaiser test. This test is a colorimetric assay to determine free amine groups. It is based on the reaction of ninhydrin with primary amines, which gives a characteristic dark blue colour. To an aminated agarose sample (100mg of gel to a total volume of 1ml) were added 50μL of each of the following reagents: 80% crystalline phenol in ethanol (w/v), 2% 0.001M aqueous solution of potassium cyanide in pyridine (v/v) and 5% ninhydrin in ethanol (w/v). The samples were then placed in a water-bath at 100°C for 5 min. The calibration curve was represented by standard solutions of glycine (0-5 μmol/mL) and the absorbance measurements of the samples (diluted 1:18) were performed at 560 nm ($Y=4.108X-0.098$, $R^2=0.99$).

2.2.12.4 BCA Assay for the Quantification of Peptides and Proteins

The BCA assay is a sensitive colorimetric assay widely employed on the quantification of total peptide/protein. The principle of the Bicinchoninic acid relies on the formation of a Cu^{2+}-protein complex under alkaline conditions, followed by reduction of the Cu^{2+} to Cu^{1+}. The amount of Cu^{2+} reduced is proportional to the amount of peptide/protein present in solution. Then, two molecules of bicinchonic acid chelate with each Cu^{1+} ion, forming a purple coloured product that strongly absorbs light at any wavelength between 550-570nm with minimal (less than 10%) loss of signal. The rate of BCA colour formation is dependent of the incubation temperature and the relative amounts of reactive amino acids contained in the proteins. The BCA reacts with protein residues like cysteine, tyrosine and tryptophan. At higher temperatures (37ºC to 60ºC), peptide bonds assist in the formation of the reaction product.

For the assay was used, 25µl of sample (fractions collected in the loading, washes and elution obtained in the microscale affinity assay and calibration curve solutions) was added in each well of a 96-well microplate. Then the BCA working reagent was prepared by mixing 50 parts of reagent A and 1 part of reagent B. After, 200µl of BCA working reagent was added in all the wells. The microplates were covered by aluminum foil and incubated at 37ºC during 30min and then the absorbance at 560nm was measured in the microplate reader. A calibration curve was prepared for the hexapeptide NN using concentrations between 0.2 to 1mg/ml of a total peptide solution. To determination of the Total Protein in the cellular extracts was used a calibration curve of 0.2 to1mg/ml with BSA protein.

CHAPTER 3 – SYNTHESIS AND SCREENING OF AFFINITY LIGANDS LIBRARIES

The combinatorial process is the best method to produce a large library of ligands, by changing two or more components in the reaction. Prior to the synthesis of combinatorial libraries of ligands it is necessary to select the properties (e.g. chemical structure, functional groups, hydrophobicity, polarity or charge) that the ligands should possess. This is accomplished by a rational design approach, where the possible interactions between the target and the designed ligands are investigated *in silico*. Despite the *in silico* results, it is necessary to validate the affinity of the ligands to the target biomolecule by experimental data in order to select the putative lead ligands for further studies and optimisation.

This chapter discusses the results of the synthesis of two libraries (64 ligands each) based on Ugi Multicomponent Reaction and in the Triazine scaffold Reaction. These libraries were screened for binding to the Hexapeptide NN and the model protein, GFP.

3.1 Solid-phase Synthesis of Combinatorial Libraries of Affinity Ligands

In the synthesis of the combinatorial libraries two different ligands were developed: one based on Ugi Reaction and one based in 1,3,5-Triazine scaffold.

The hexapeptide that is being studied is composed by asparagines (Asn, N). This aminoacid is classified as hydrophilic and uncharged. Since the hexapeptide is only composed by this aminoacid, the groups in the side chains are amide groups. This hexapeptide can make intermolecular hydrogen bonds and in high concentration aqueous solutions it can make gels (Issaq et al. 2009). The expected ligands to interact strongly with this hexapeptide are the ligands that can form hydrogen bonds, the ones that have hydroxyl groups or nitrogen atoms, that are electron-rich and can make hydrogen bonds with the aminoacid side chain (Labrou 2003).

Using this information several ligands were studied using rational design (Pina 2008) and some were selected to make the library.

For the synthesis of ligands based on the Ugi Reaction, eight different Amines and eight different Carboxylic Acids were the basis for the variability of ligands in the library. These compounds were added to an agarose resin activated with an aldehyde group, another component of the Ugi Reaction. From the combination of the amines and carboxylic acids compounds a library of 64 different ligands was generated.

CHAPTER 3 – SYNTHESIS AND SCREENING OF AFFINITY LIGANDS LIBRARIES

The combinatorial library of 1,3,5-Triazine based scaffold was constructed by exploring the affinity between the hexapeptide, and compounds with groups that have the possibility to form hydrogen bonds. In this library eight different amines were combined between them and the library was composed by 64 different ligands.

In order to assess the success of the solid-phase synthesis of the Ugi and Triazine ligands, a positive control was generated. The pyrene molecule was used as the amine compound for both syntheses. The fluorescence intensity in the resin was then quantified (Figures 3.1) and the samples were illuminated with UV light transilluminator (290-365nm) (Figure 3.2).

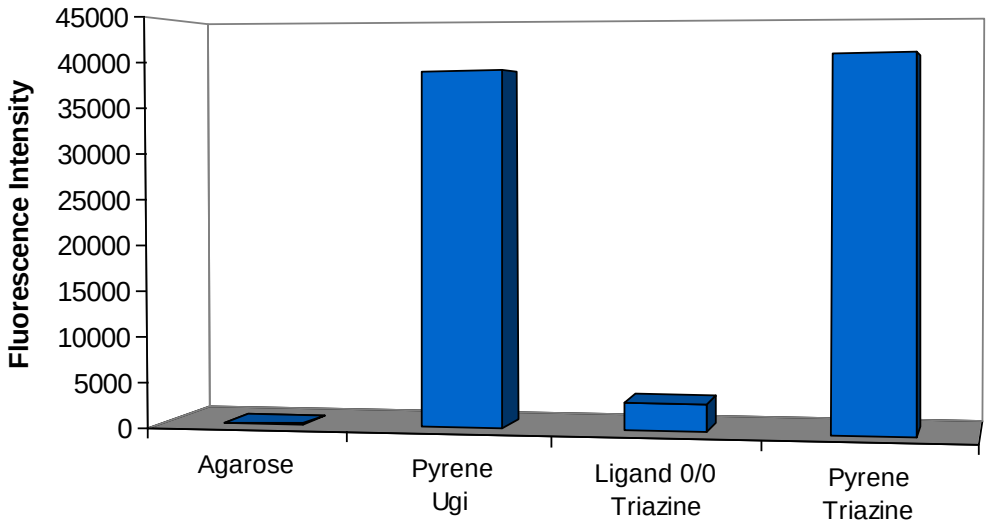


Figure 3.1 - Fluorescence Intensity of Pyrene Control Ligand Synthesized by Ugi and Triazine Reaction.

Figure 3.2 - Samples of Agarose (Left) and Pyrene Control Ligand (Right) Synthesized by Ugi Reaction illuminated with a UV transilluminator (290-365nm).

CHAPTER 3 – SYNTHESIS AND SCREENING OF AFFINITY LIGANDS LIBRARIES

3.1.1 Screening of Ligands with Hexapeptide NN and GFP

The two solid-phase combinatorial libraries were screened for binding to the Hexapeptide NN and GFP.

The detection of peptides can be made by UV/Vis, fluorescence, or mass spectrometry. Despite the low selectivity peptides can be detected by UV absorption at 185-220nm (range of peptide bond) or at 280nm when containing aromatic groups. Another method that can be explored for these aminoacids that have aromatic rings is the intrinsic fluorescence, being the tryptophan the most selective. Other methods can be applied for detection of peptides without intrinsic fluorescence, like labelling peptides with fluorescent substances or radio labelling (Issaq et al. 2009).

In this work the detection method was the BCA assay, which is a colorimetric assay, that can detect proteins or peptides in a range of concentrations of 0.2 to 1mg/ml.

The concentration of the peptide in the loading sample, washes and elution were determined by the BCA method, being possible to calculate the mass of peptide bound to the columns.

With the information form the Figure 3.3 it was possible to determine the ligands possessing higher binding percentage for the hexapeptide, and that corresponds to the ligands that have the highest affinity for the hexapeptide.

In the Ugi library there are some ligands have higher binding percentage (30-45%) of binding and that stand out between the 64 ligands (Figure 3.3 – A). For the Triazine library all the ligands have similar affinities for the hexapeptide (20%) (Figure 3.3 – B).

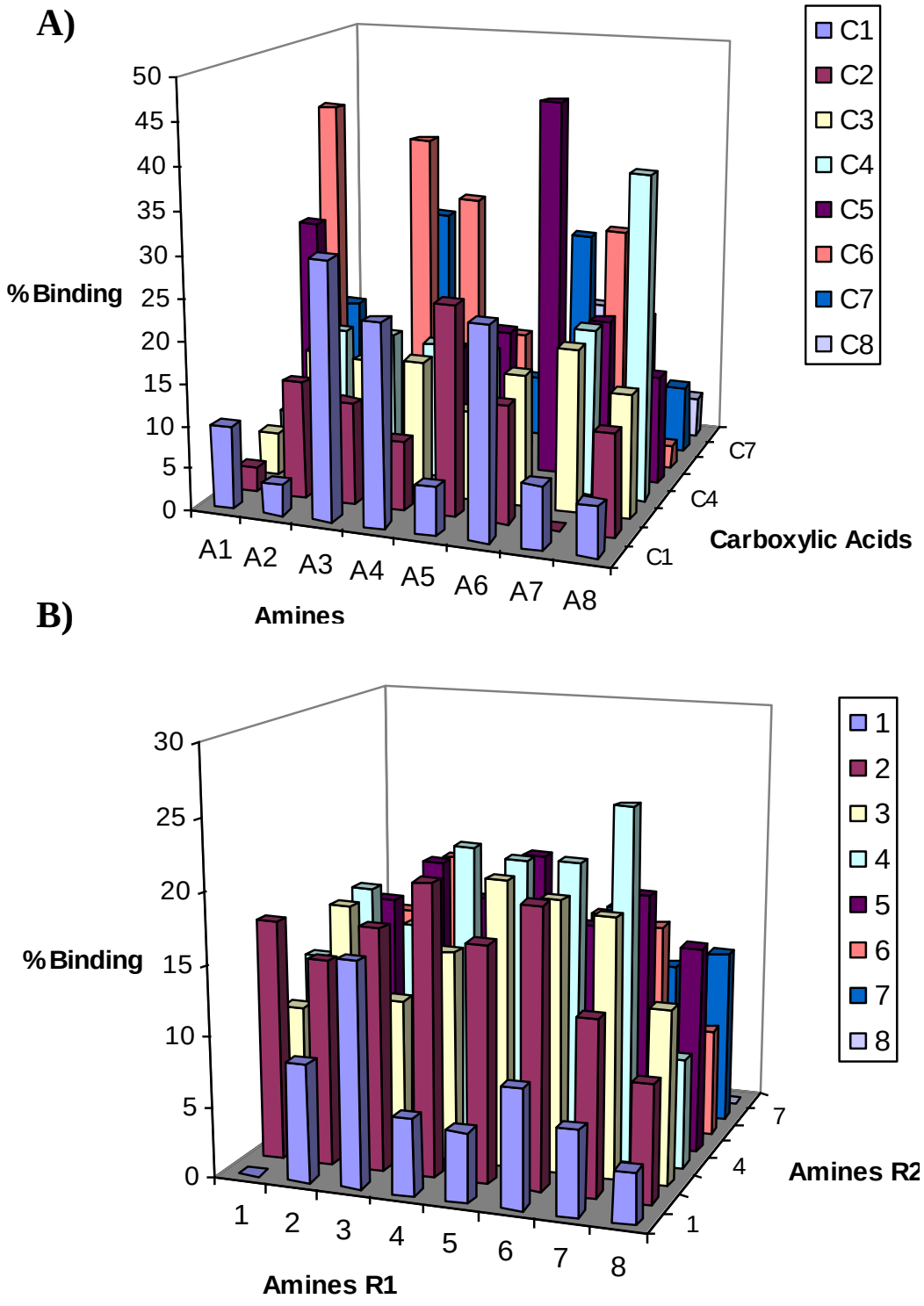


Figure 3.3 – Results of the screening of the combinatorial libraries with the Hexapeptide NN: A) Ugi ligands and B) Triazine based Ligands.

CHAPTER 3 – SYNTHESIS AND SCREENING OF AFFINITY LIGANDS LIBRARIES

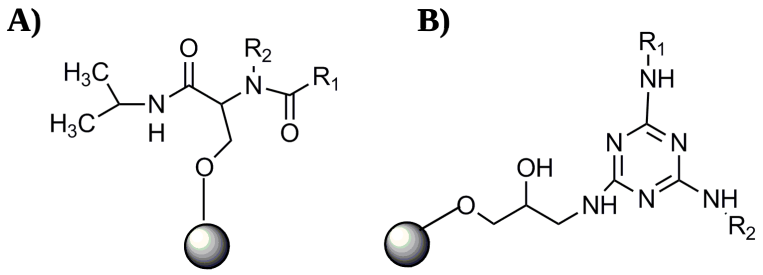


Figure 3.4 - Structures of the ligands: A) Ugi ligands and B) Triazine based ligands.

These distinct binding results between the Ugi and the Triazine ligands can be due to the structural difference between them. The Ugi ligands have a single carbon bond structure that allow the free rotation of the atoms or the side chain groups, in contrast with the Triazine Ligands, which are based in the 1,3,5-Triazine ring and are more rigid (Figure 3.4 - B). The flexibility can be useful when the ligands interact with the hexapeptide, that would allow different forms of interactions in the case of the Ugi ligands and less interactions for the Triazine ligands. On the other hand, an increase in the flexibility can generate a decrease in the specificity.

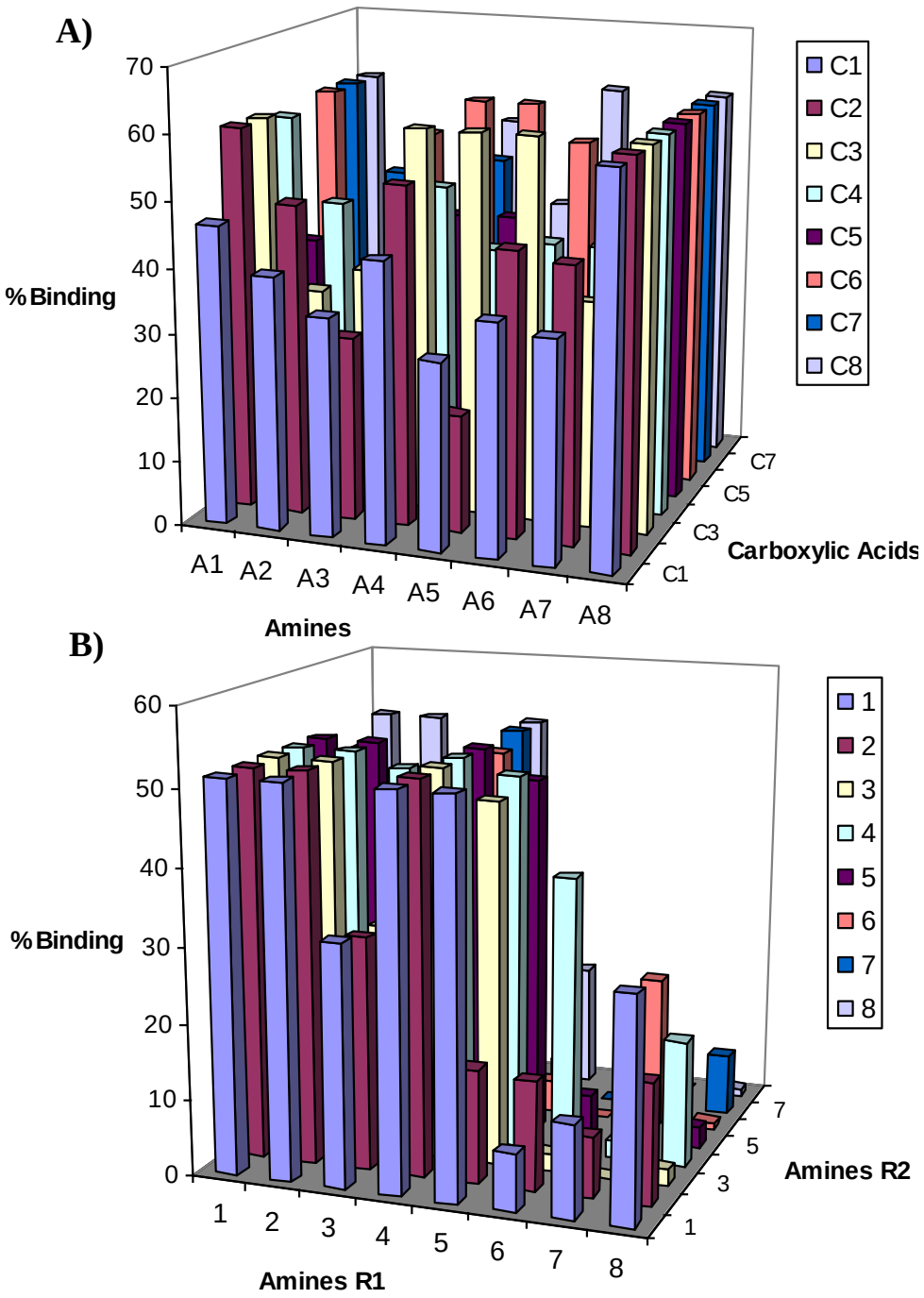
The samples from elution were analysed and was concluded that no peptide was being eluted from the columns.

The screening with GFP allowed understands which ligands are selective for the hexapeptide. The intrinsic fluorescent of GFP has been explored for this purpose.

This method is very sensitive and more selective than any colorimetric assay because the colorimetric assays normally explore the presence of some characteristic of proteins, as in the BCA assay.

The fluorescence intensity of the samples from loading, washes and elution were analysed and the percentage of bond GFP to each ligand was determined.

The results demonstrated that the Ugi ligands have much more affinity than the Triazine ligands for GFP (Figure 3.5).



Figures 3.5 - Results of the screening of the combinatorial libraries with the GFP: A) Ugi ligands and B) Triazine based Ligands.

The ratio between the mols number of Hexapeptide and mols number of GFP that were bonded in each resin, (Figure 3.6). In this figure is compared the results of the affinity between each ligand and the hexapeptide NN or the GFP. The ligands that have the highest ratio have consequently the highest selectivity for the hexapeptide NN.

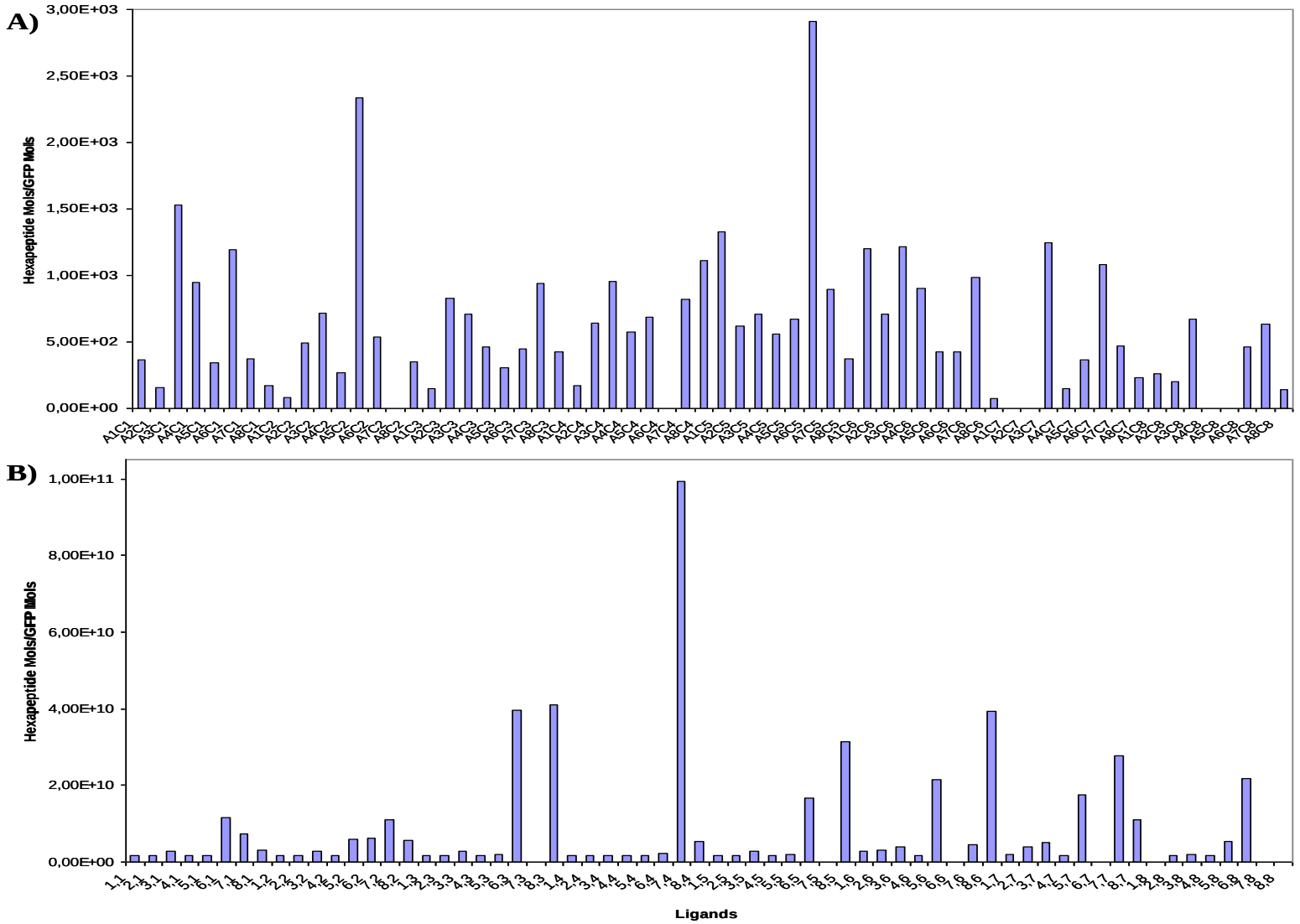


Figure 3.6 - Ratio between the Number of Mols Hexapeptide:GFP bonded to the solid-phase ligands A) Ugi ligands and B) Triazine based Ligands.

For the Ugi library, (Figure 3.6 –A), the results demonstrated that despite the binding affinity of the ligands for the hexapeptide these ligands are not very selective for it. In contrast, in the Triazine library, although the library ligands appeared to had the same affinity for the peptide, there some ligands, that can be very selective for it, (Figure 3.6 –B).

From these (Figure 3.6) were selected the ligands in both libraries that had the highest selectivity for the hexapeptide.

Table 3.1 – Summary of the possible lead ligand for binding to Hexapeptide NN selected.

Combinatorial Library	Ligands
Ugi	A3C1, A6C1, A5C2, A1C5, A6C5, A7C6, A3C7, A6C7
Triazine	6,3; 8,3; 7,4; 8,6

3.1.2 Scale-up Synthesis and Screening of Possible Lead Ligands

The lead ligands (Table 3.1) were synthesized at a larger scale and 0.5mg of this resin was packed in chromatographic columns. The ligands were screened for binding to the hexapeptide and the samples were analysed by the BCA assay (Figure 3.7).

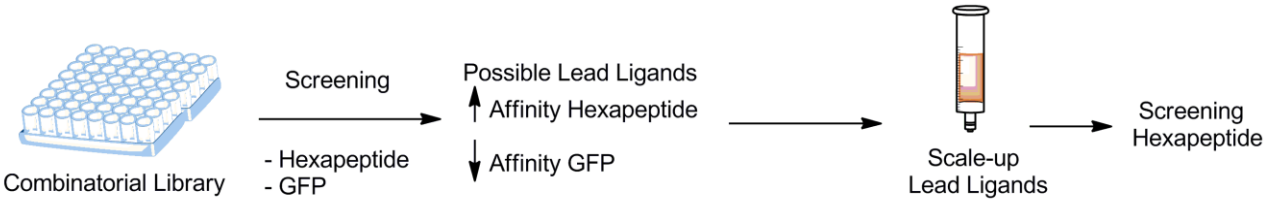


Figure 3.7 – Strategy for the detection of lead ligands.

The binding affinity was determined as explained previously (3.1.1) in the graphs are compared the ratio Hexapeptide mass: Support mass, this gives the idea of the quantity of hexapeptide bonded in the support, in the 96 wells blocks and in the chromatographic columns (scale-up samples).

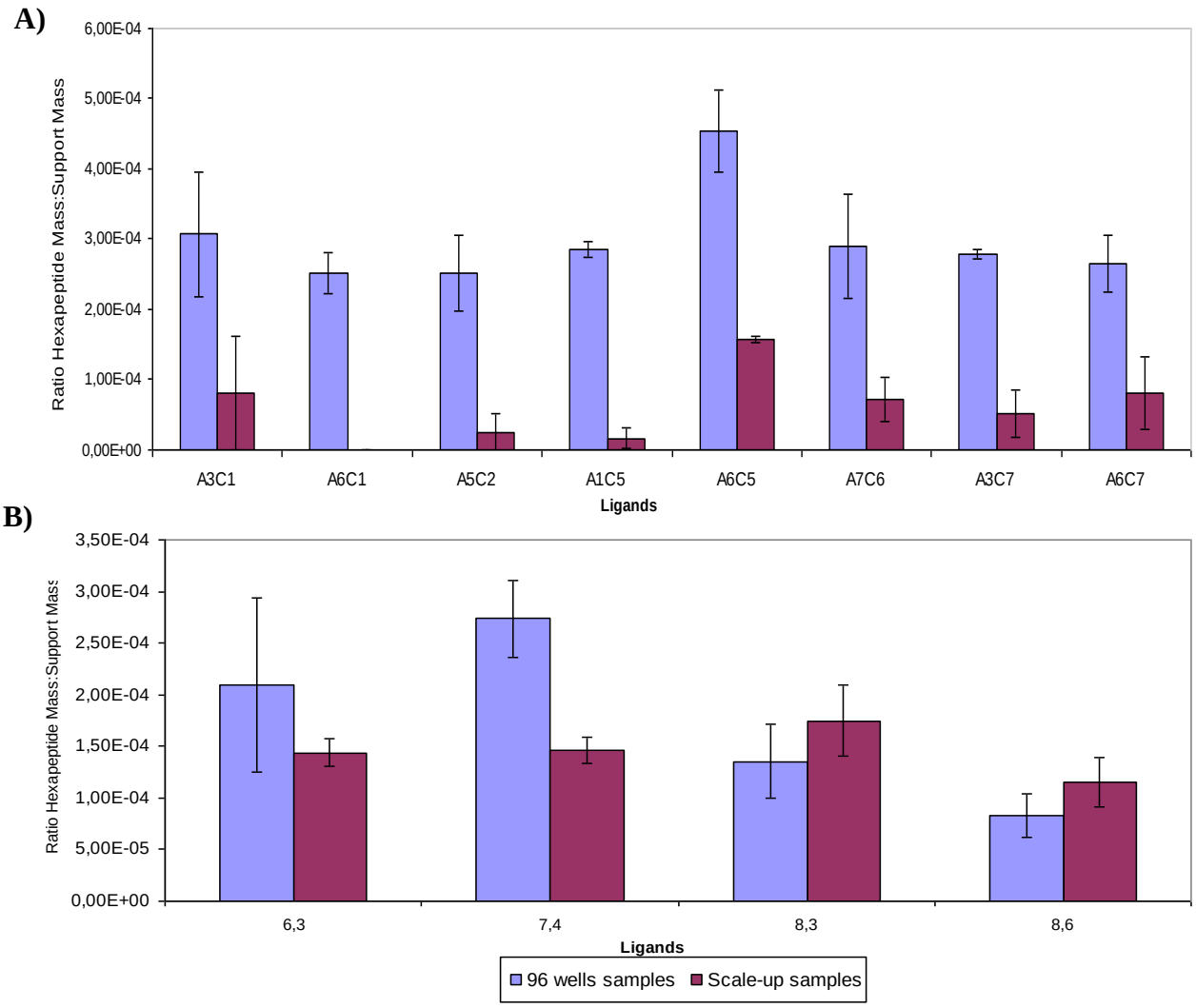


Figure 3.8 - Comparison between Ratio of Hexapeptide Mass and Support Mass results of 96 wells blocks and Chromatographic columns: A) Ugi Ligands and B) Triazine based Ligands.

The Figure 3.8 A and B show a decrease in the affinity of the ligands to the hexapeptide. These results can be related with two factors, one is the quantity of ligands in the resin and the other the screening method used.

To functionalize the resin with a ligand in the first step is necessary to activate by adding a epoxy-group, in the case of the Ugi ligands, the resin of the 96-well plate had a 17µmol epoxy groups/g support, although the activation of the epoxy group in the resin of the chromatographic columns had a concentration of 40 µmol epoxy groups/g support. The quantity of epoxy groups have been discussed by different authors, since the concentration of the ligands in the surface of the resin can improve the affinity of the resin or an have an inhibitory effect. For Ugi ligands for purification of IgG the optimum concentration of epoxy groups was determined in 20 µmol/g (Haigh et al. 2009). For Triazine ligands, the decrease in

the affinity is not so accentuated as in the Ugi ligands probably because in this library the concentration of ligands in the surface of the resin was similar in the resin of the 96 wells and the chromatographic columns, 20 and 17 µmol epoxy groups/g respectively, however some authors demonstrated that the best concentration for purification of glycoproteins was approximately 24 µmol/g (Palanisamy et al. 2000). The low concentration of ligands, as explained by Palanisamy and co-workers inhibits the formation of a complex between the ligands and the protein, since the co-operativity between ligands may be an important factor in the ligands binding to the glycoproteins. For IgG was observed that the interactions in the resin occurred among the protein and more than one ligand (Haigh et al. 2009). Despite the difference in size between a Hexapeptide and an IgG or a glycoprotein the affinity of the ligands to the Hexapeptide can be lower, when the concentration is higher, because these ligands can make interactions between them and be less available to interact with the Hexapeptide.

Other question that has to be addressed in this discussion is the screening process. In the 96 wells blocks the resin is not packed, so after adding the hexapeptide solution the block is incubated, during the time of incubation the interactions between the ligands and the peptide can be formulated and there is an equilibrium that is established. Subsequent to the incubation the solution is drained, using a centrifugation step, and collected from the well and the washes start. In the columns the resin is packed in between two frits, when the peptide solution is added the time to establish the equilibrium and formulate the interactions between the hexapeptide and the ligands is dependent on the gravity and the viscosity of the solution; despite that the period of incubation is respected and only after that the washes start. From these results is possible to conclude that the flow rate of the samples in the column is limiting for the interaction between the ligands and the hexapeptide. This step must be improved to promote these interactions, probably the best method is to add a solution to an approximately rate of 0,5ml/min (Platis et al. 2006) or low 0,4ml/min (Palanisamy et al. 2000) as referred in the literature as the optimum rate for these chromatographic columns.

Despite the apparent affinity decrease between the ligands and the hexapeptide the affinity proportion between the ligands is the same. In the Ugi library the ligand with the highest affinity is the A6C5 and in the Triazine library is the 7,4.

3.1.3 Automated Docking Studies

Molecular modelling studies were performed to understand the interaction between the Ugi lead ligands and the Affinity Tag, but also with the ligands and GFP (model protein).

These docking studies, allow the analysis *in silico* of the interactions that occur between a ligand and a protein when they are in contact. Subjacent to these studies is the basic thermodynamic idea that every system in nature tend to minimize the state of energy. For example in the protein structure prediction methods, the native state of the protein is the state of lowest free energy (Lazaridis and Karplus 2000). In the molecular modelling softwares, the most favourable pose, complex hypothetical conformation generated by a molecular model, will be the one with the lowest free energy (Mobley and Dill 2009).

Before using the molecules was necessary to construct them. The structures of the Ugi ligands that were used are present above:

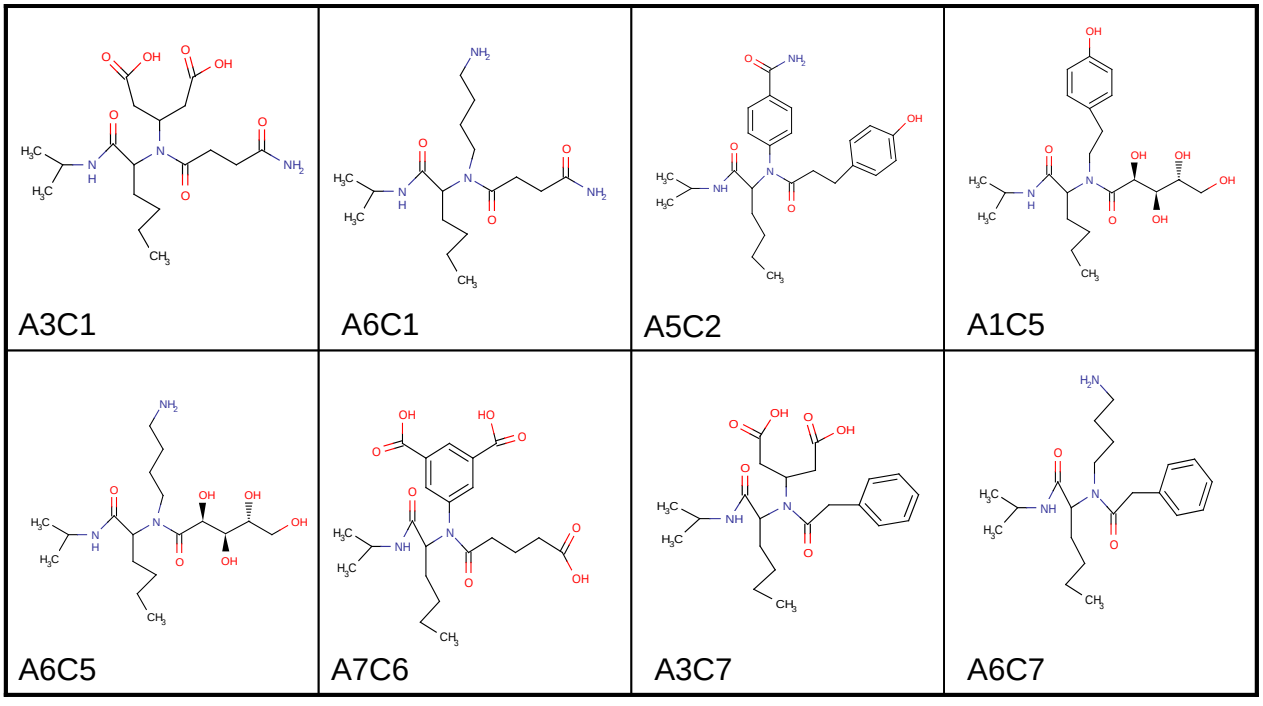


Figure 3.9 - Structures of the Ugi ligands used in the *in silico* Studies.

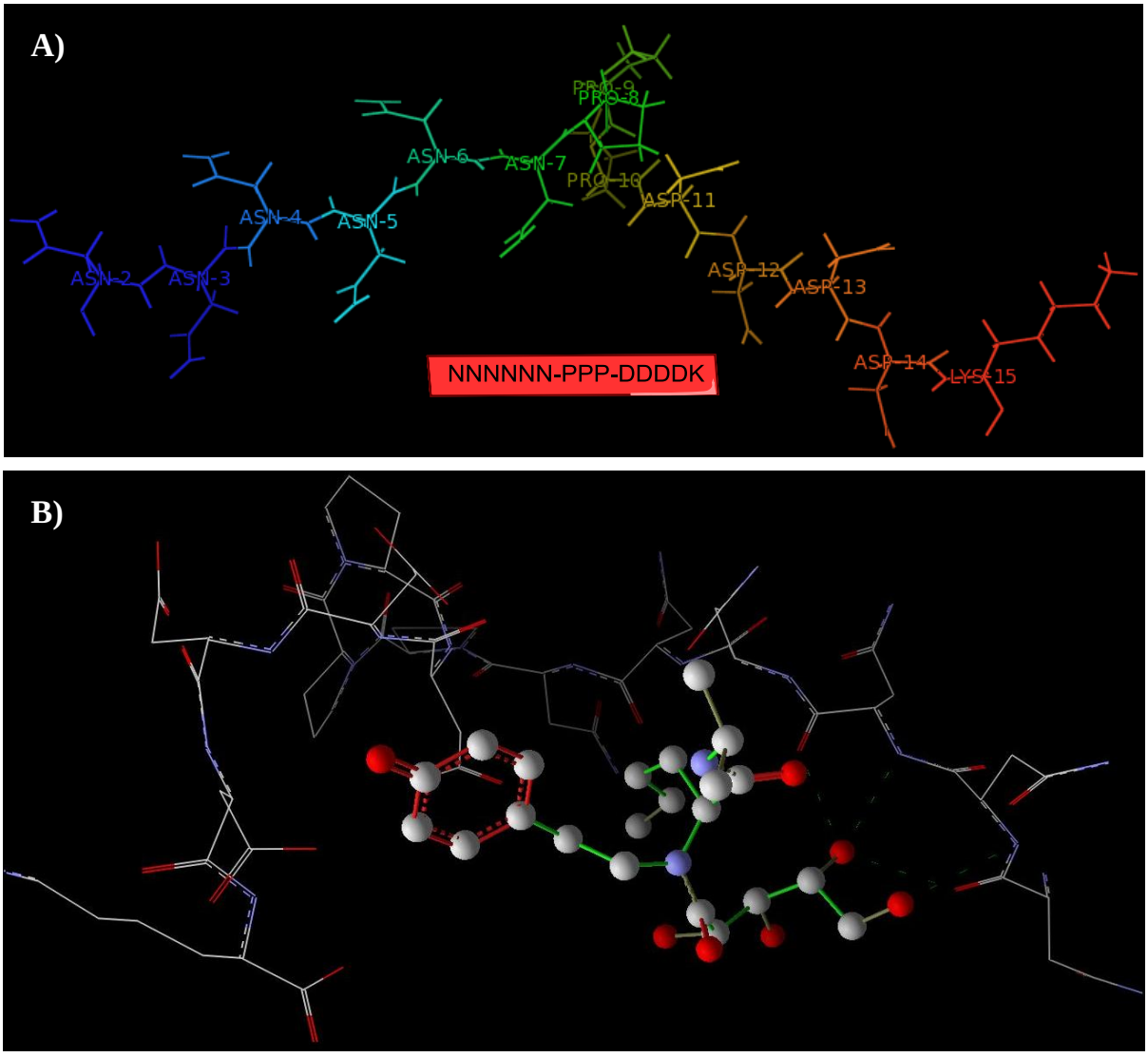


Figure 3.10 - Structure of the Affinity Tag used *in silico* studies A) (Hexapeptide+3xProlines+Enterokinase Recognition Sequence); B) Example of docking result between a ligand and the affinity tag.

The main interactions between the ligands and the affinity tag are made with Asn-5, Asn-6 and Asn-7 (Figure 3.10). Some ligands make also interactions with the sequence of the Enterokinase. One example is the A6C5 it has the highest affinity for the hexapeptide in the experimental results, this results *in silico* can indicate that this ligand is not highly selective for the hexapeptide NN when in the protein are present sequences as the Enterokinase (Table 3.3).

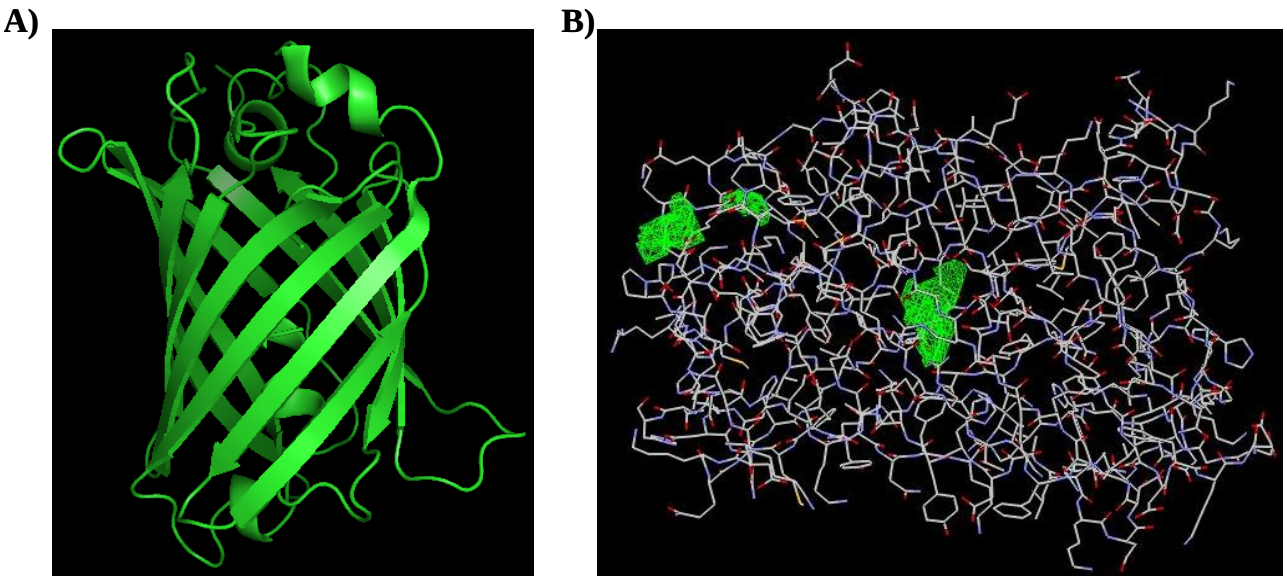


Figure 3.11 – Subunit A of GFP (1W7S, PDB code): A) – Structure and B) Structure with the cavities the cavities determined.

The structure of the subunit A of GFP, (Figure 3.11 - A), was analysed in the docking software for the presence of cavities (Figure 3.11 - B). However, the subunit has three cavities (two in the surface and one in the core of the structure (Table 3.2) and the ligands will be attached to the support, only the interactions in the cavity with the large area in the surface of the subunit were analysed (Figure 3.13).

Table 3.2 - Areas of the Cavities in the Subunit A

Cavities	Volume and Position
1	41,472 A^3; Intern
2	23,04 A^3; Superficial
3	12,8 A^3; Superficial

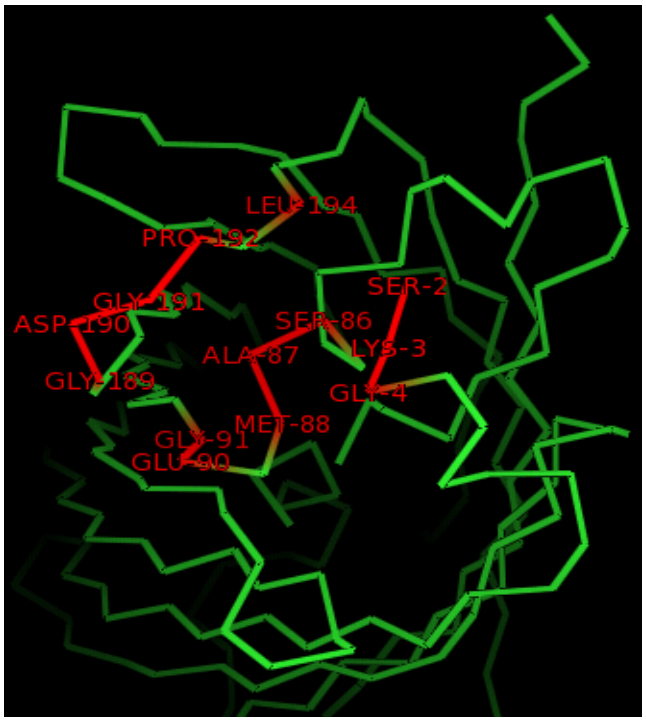


Figure 3.12 - The larger cavity of GFP Subunit A with the aminoacids identified.

The larger cavity is constituted mainly by apolar and hydrophobic aminoacids. The results demonstrated that the ligands mainly interact with the SER-2 (polar and uncharged), Lys-3 (positively charged) and Gly-4 (apolar) (Figure 3.12).

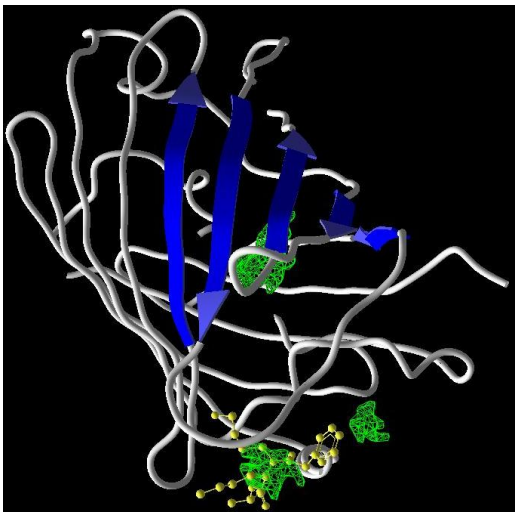


Figure 3.13 – Typical docking between the subunit A and a ligand.

By the docking result was possible to determine the best position for the ligands and the GFP but also the free energy involved in that process (Figure 3.13).

By analysing the docking and the experimental results (Table 3.3), a correlation between them was observed. The ligands that form more hydrogen bond have strongest interactions with the hexapeptide and present a higher percentage of hexapeptide binding.
Furthermore, with these studies was possible to confirm the affinity of each ligand to the GFP, the free energy involved in the conformation between the GFP and the ligands is lower, consequently more favourable, than the energy of the complex formed by the ligands and the Affinity Tag. However, the ligands A6C5 and A7C6 have the lowest Rerank scores with GFP which means that they have the highest affinity for the GFP, what was not observed in the experimental results. The ligand that has the highest score (-91.2), more positive, for the GFP is the A6C7, but this result is similar to the score of the affinity tag which means that it has almost the same affinity for the GFP than the Affinity Tag.

Table 3.3– Results of the docking studies for the Ugi Ligands between the GFP or the Affinity Tag.

Ligands	GFP			Affinity Tag (Hexapeptide NN-3Prolines-Enterokinases Recognition Sequence)		
	Rerank Score*	HBond	% Binding	Rerank Score*	HBond	% Binding
A3C1	-94.18	-5.4	34.04	-88.36	-5	§
A6C1	-96.03	-7.41	35.84	-66.14	-6.09	25.00
A5C2	-99.04	-0.42	18.27	-82.94	-8.80	24.99
A1C5	-98.83	-4.35	36.31	-92.44	-7.65	28.31
A6C5	-102.74	-12.92	26.44	-88.75	-20.13	§
A7C6	-100.93	-0.72	49.90	-87.94	0	§
A3C7	-82.92	-5.38	37.99	-75.64	-2.5	§
A6C7	-91.17	-3.83	41.37	-92.76	-6.73	26.32

Notes:

*) The Rerank Score is a specific scoring function of Molegro Docking Software, to denote the best docking solution from the docking solutions and does not have any specific units (Thomsen and Christensen 2006).

§) Ligands that do not interact with the hexapeptide NN. They interact with the Enterokinase Sequence.

To understand the correlation between the affinity percentages determined in the experimental assays and the automated docking results for GFP the results were plot (Figure 3.14).

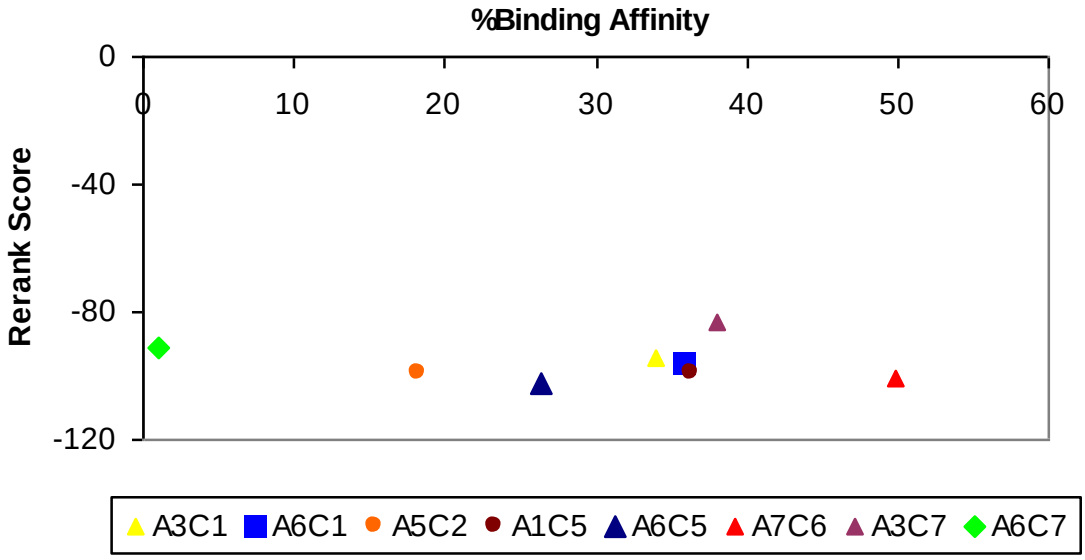


Figure 3.14 – Correlation between the experimental assays and the automated scores.

From the data is possible to determine that the relation between the experimental assays and the automated results is the same for all the ligands (Figure 3.14). The study of interactions *in silico* presents several limitations, because neither the support (e.g. agarose resin) or the solvatation of protein-ligand are considered (Platis et al. 2006). These points are important, since when the ligand is attached to a support the conformation and the mobility of the ligand are limited, besides it can only interact with residues that are exposed in the protein surface and not in the core, although in these study was considered the spacer-arm - epoxide group activated in the resin, like suggested by other works (Haigh et al. 2009). Furthermore, the consideration of the presence of a solvent is important for the mobility of the structure and it is able of mediate some interactions, e.g. hydrogen bonds in the complex, because the angles of the complex are more flexible (Haigh et al. 2009). However, when the flexibility of the ligands is taken into account that would increase the number of possible interactions, to analyse all that binding possibilities, that would required even with computers more time and improved softwares (Thomsen and Christensen 2006). Furthermore, in this particular study was not considered the secondary structure of the Affinity Tag that can interfere in the interaction between this molecule and the ligands, since the more or less rigidity of the molecule can restrict or improve the interactions.

3.1.4 Conclusions

Two solid-phase combinatorial libraries of affinity ligands have been synthesized one based on the Ugi Reaction and other based in the Triazine scaffold. The two libraries demonstrated different affinity for the hexapeptide NN. In the Ugi library a higher percentage of binding to the hexapeptide (30-45% compared with 25% for Triazine based library) was observed. However, the Triazine based ligands demonstrated that they are more selective for the hexapeptide NN than the GFP (model protein).

In addition, even though the promising results of the 96 wells platforms, when the Ugi ligands were tested using more quantity of resin and packed in the chromatographic columns, the affinity of the ligands decreased. This decrease in the affinity of the triazine based ligands was not so distinct. These results can be related with the process of incubation of the peptide solution in the columns, that must be improved to allow the formation of interactions between the ligands and the hexapeptide. In addition was not observed elution of the hexapeptide or the GFP in the conditions that are currently used.

Some ligands presenting highest affinity for the hexapeptide and less for the GFP, had been selected to carry out more studies: 8 Ugi ligands and 4 triazine based ligands.

Finally, the Ugi ligands selected were analysed by automated docking studies. The docking results with the ligands and the Affinity Tag or the GFP, demonstrated that the interactions that are involved in binding of hexapeptide and the ligands are mainly hydrogen bonds, and the ligands that have the best scores are the ones that can form these bonds with the hexapeptide. The affinity of the ligands to GFP was confirmed in these studies and in general the docking scores for the GFP are higher, the complex formed ligand-protein is more stable, than for the Affinity Tag.

CHAPTER 4 – EXPRESSION IN MAMMALIAN CELLS

Nowadays the production of certain biopharmaceutical is made in mammalian cells, e.g.monoclonal antibodies (mAb) (Shukla et al. 2007), since these cells make posttranslational modifications that are essential for the correct functionality of proteins used in human health. Despite the costs of production and the low yields, in recent years have been developed cells lines with higher productions rates. As a consequence was necessary to improve the downstream methods to recover the biomolecules of interest. For purification is normally used a Affinity Chromatography, that can be based in natural affinities e.g. Protein A chromatography for purification of antibodies (Shukla et al. 2007), or in specific interactions between a ligand for a protein or a fused protein e.g. His-Tag, using IMAC. These fused proteins are expressed with a tag, a sequence of aminoacids that has affinity for a specific ligand. After purification is possible to remove the tag and recover the protein of interest.

In this work were designed vectors to express in mammalian cells, HEK 293T cells, the production of a fused recombinant protein. The model protein used was GFP and the affinity tag codified a sequence of aminoacids for a hexapeptide, three Prolines and an Enterokinase Recognition Sequence.

The first vector had the hexapeptide RW, and using this as base was constructed a second vector with the hexapeptide NN in the affinity tag. Was also expressed a vector with GFP without any affinity tag.

4.1 Amplification, Construction and Purification of Vectors

The vectors pRWTagMGFP and p6NTag were design in the laboratory and the sequence ordered and purchased, in consequence they were ready to use for bacteria transformation. Subsequent to amplification the vectors these were purified with a concentration of 1.661µg/µl and 2.212 µg/µl and the purity was observed by the ratio A_{260}/A_{280} of 1.81 and 1.80 respectively, that demonstrate the good purification results. The restriction reaction electrophoresis gel allowed confirming vector correct sequence.

In the construction of the vector p6NTagMGFP was used the pRWTagMGFP was used as backbone vector, and the local of the affinity tag, with the hexapeptide RW was removed and the sequence for the affinity tag with the hexapeptide NN was introduced.

Despite the low levels of purification for the two fragments of interest, from the agarose gel, they were sufficient for the reaction to occur.

The constructed vector, p6NTagMGFP, was use for bacteria transformation and the presence of colonies demonstrated the ligation of the plasmid, despite these results was necessary to confirm the correct insertion of the sequence of the affinity tag fragment. The vector was purified in small scale, the restriction reaction results showed the fragments of interest (78bp for the insert and 4126bp for the vector). The production was completed and the vector purified, the final product had a concentration of 7.77 µg/µl and A_{260}/A_{280} of 1.96. To confirm the presence of only one insert in the vector was used a restriction enzyme HindIII, the results of the restriction reaction demonstrated that there was only one insert.
The vector pMDISGFP, that contained the GFP without affinity tag, was amplified and purified with a concentration of 1.72 µg/µl and the purity was determined by the ratio A_{260}/A_{280} of 1.88, that represents a good result. The analysis of the restriction reaction by the electrophoresis gel allowed confirming the expected configuration of the vectors.

Subsequent to the production and purification of the vectors they were used for transfection of HEK 293T cells.

4.2 Transfection of HEK 293T cells

To insert the vectors in mammalian cells different methods have been used, however the Transient Gene Expression (or transient transfection) is the most frequently used method for protein expression for research purposes. For protein production in large scale is possible to develop a cell line, but as was referred before this strategy is time consuming (Geisse and Fux 2009).
Different methods can be used to transfer the DNA to the cells, like: lipofection reagents, with the highest percentage of transduced cells – 90-100%; Calcium-phosphate or Polyethylenimine (PEI), these two in general with lower percentage of transduced cells (80%) (Geisse and Fux 2009). The PEI reagent is a cationic polymer that neutralizes the negative charge of DNA, and allows the entrance of the nucleic acid in cell membrane by endocytosis, and protects the nucleic acid until it enters in the nucleus (Bertschinger et al. 2006). This was the reagent used for transfection in this work (Figure 4.1).

CHAPTER 4 – EXPRESSION IN MAMMALIAN CELLS

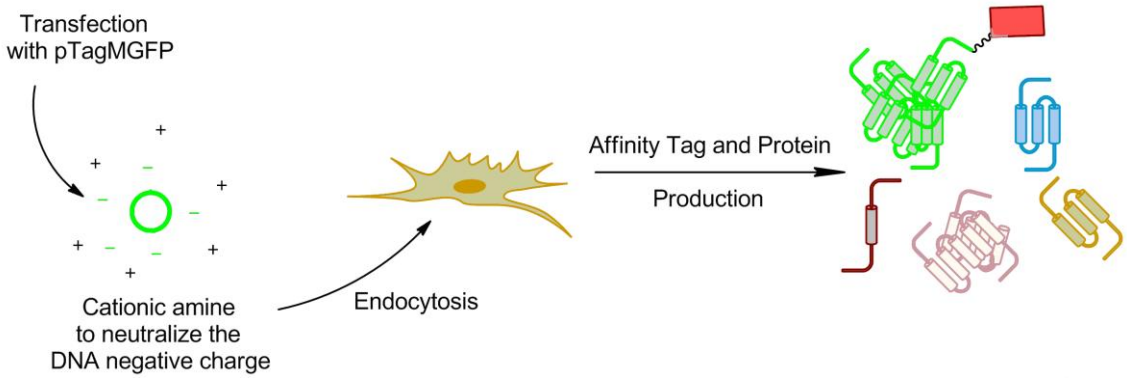


Figure 4.1 – Strategy for transfection of HEK 293T cells for production of a fused recombinant protein.

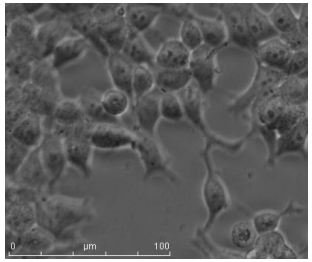
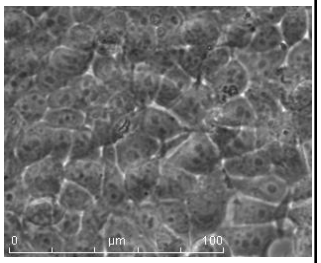
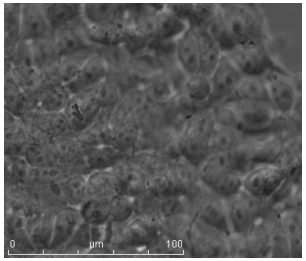
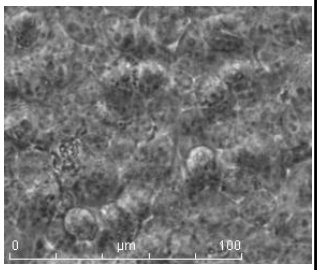
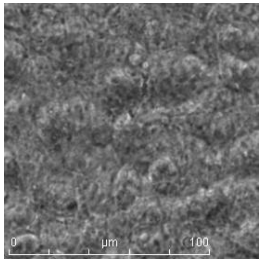
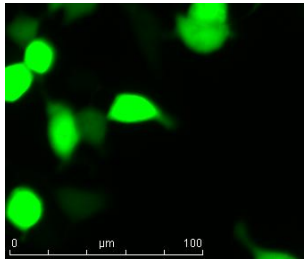
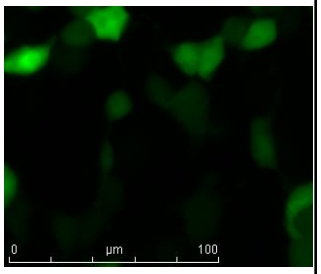
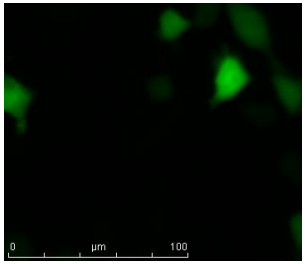
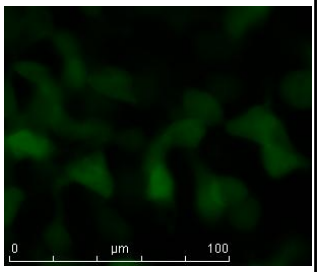
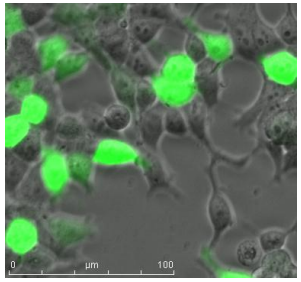
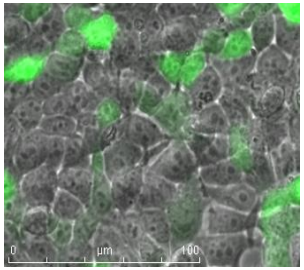
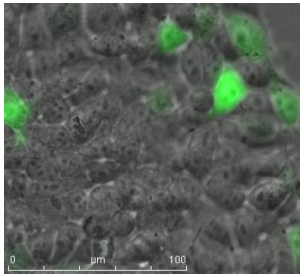
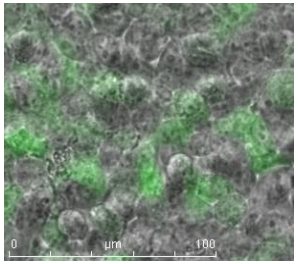
Table 4.1 - The transfection results and the confluence of the culture.

Vector	24h After	48h After
pRWTagMGFP	60% Confluence and 50% Transfection	70% Confluence
p6NTagMGFP	80% Confluence and 60% Transfection	90% Confluence
pMDISGFP	80% Confluence and 60% Transfection	90% Confluence

The transfection percentage (Table 4.1) was determined based on the observation of the cells with fluorescence that indicates the production of GFP, and the ones without fluorescence.
The transfection was more efficient for the last two vectors this can be related to the formation of the complex between the PEI and the DNA. During these transfection procedures were seeded cells in the same conditions but were not transfected – these samples were labelled "untransfected cells", was observed that in these plates the growing rate was higher, having at 48h a confluence of 100%. These results give the idea that the transfection process is always a stressful procedure for the cells that have to over express a recombinant protein, as consequence the energy supply is used for production of the protein and not for growing. The production of the GFP can be toxic for the cells and slows down the cellular growth. By observing fluorescence in the samples is possible to say that the conformation that allows the formation of the GFP chromophore was not altered, consequently the introduction of the affinity tag in the N-terminal of the protein apparently did not affect the structure of the protein (Table 4.2).

CHAPTER 4 – EXPRESSION IN MAMMALIAN CELLS

Table 4.2 - Transfection Results (24h and 48h post transfection) for cells transfected with p6NTagMGFP and pMDISGFP. Compared with "untransfected cells" The Magnification of the images is 20x.

Filter	p6NTagMGFP		pMDISGFP		Untransfected Cells
	24h	48h	24h	48h	48h
Without Fluorescence					
Fluorescence Filter					(No fluorescence Signal)
Contrast					----------

Subsequent to the transfection the cells were scrapped from the plates, this method was used since it does not require the use of Trypsin, that would increase the complexity of the mixture, and the cellular extracts were collected in the less volume possible to concentrate the proteins and stored at -20ºC. The freeze-thaw method was used to open the membranes of the cells and collected the intracellular proteins. The samples were centrifuged, the supernatant was collected and analysed to determine the concentration of total protein and GFP.

Table 4.3 – Results of the extraction of proteins from HEK 293T cells.

Samples	[Protein Total] mg/ml	[GFP] mg/ml	% GFP
pRWTagMGFP	3.010	3.55E-2	1.18
p6NTagMGFP	7.008	28.2E-2	4.03
pMDISGFP	7.926	9.38E-2	1.18
Untransfected Cells	5.775	----------	----

From the results (Table 4.3) is possible to observe differences in the production of the cells that is related with the: efficiency of the transfection procedure, confluence that was obtained in each experience and number of plates that were used.

The cellular extracts from pRWTagMGFP (GFP tagged with Hexapeptide RW) were used to test the affinity of the RW lead ligands previously selected; the p6NTagMGFP (GFP tagged with Hexapeptide NN) were used to test the NN ligands (Ugi and Triazine ligands); the pMDISGFP (GFP without tag) were used to test the affinity and selectivity of the ligands previously selected for the hexapeptide RW and NN, it was used as a control positive; and, finally, the untransfected cells extracts were used as another control (negative control), since these extracts do not have any GFP.

4.3 Conclusions

The amplification and purification of the plasmids in bacteria was efficient. The construction of a new plasmid p6NTagGFP based on the pRWTagMGFP and p6NTag was successful with a high concentration and purification.

CHAPTER 4 – EXPRESSION IN MAMMALIAN CELLS

Subsequent to purification of the plasmids was necessary to transfect the mammalian cells, the transfection percentage with the different plasmid was between 50-60%, that is considered in the normal interval for 293T cells transfected using PEI.

Since was observed fluorescence in the fused GFP that indicates that the affinity tag doesn't affect the conformation of the protein and does not destroy the chromophore of the protein.

In this work was demonstrated the efficient expression of a fused GFP with two different affinity tags, one with mix mode character hexapeptide RW and other more hydrophilic the hexapeptide NN.

To collect the cells from the culture plates was used the cell scraper and to extract the intracellular proteins the Freeze-Thaw method was used.

It was determined the concentration of the total protein in the cellular extracts and the concentration of GFP, de results demonstrated that the GFP present in the samples were 1% or 4% of the proteins in the cellular extracts.

Finally is possible to conclude that the vectors can be improved to increase the productivity of fused GFP recombinant protein.

CHAPTER 5 – CHARACTERIZATION OF POSSIBLE LEAD LIGANDS

Following the selection of lead ligands from two combinatorial libraries of affinity ligands for binding to the hexapeptide, it was necessary to test these ligands in the real conditions in order to assess the selectivity of the ligands for the target molecules. For that purpose cellular extracts from HEK 293T cells that contained the recombinant fused GFP were used. Despite the high content in GFP tagged with a hexapeptide, the cellular extracts obtained also contain other proteins that can work as competitors for the ligand.

5.1 Screening of Ligands with Cellular Extracts

The ligands tested were the ones that had affinity for the i) hexapeptide NN and ii) hexapeptide RW. Each ligand was tested against the cellular extracts: that had the GFP tagged with Affinity Tag (Hexapeptide-3xProlines-Sequence for Enterokinase); a Positive Control (PC), that had the GFP without tag; and, a Negative Control (NC), without GFP (Figure 5.1).

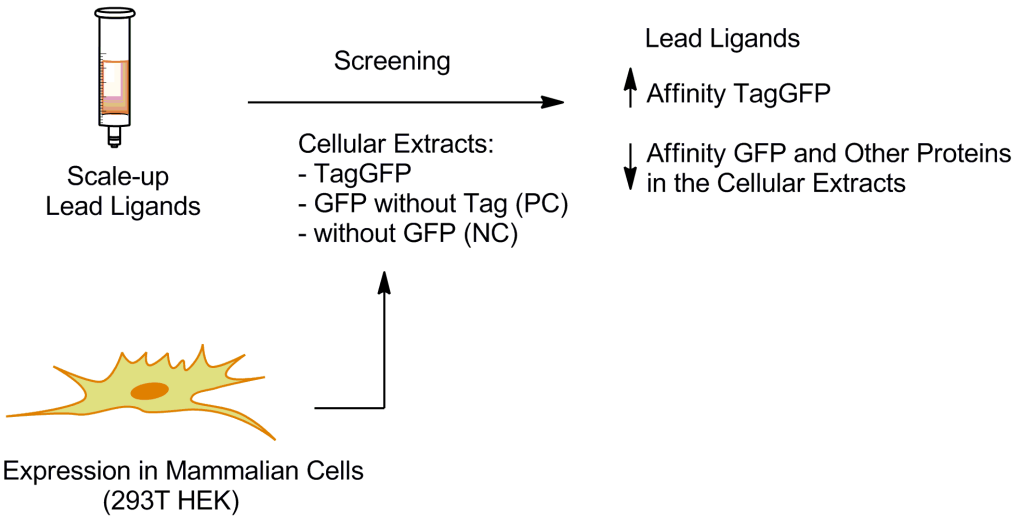


Figure 5.1 – Strategy to screen and select the lead ligands with cellular extracts.

The samples from loading, washes and elution steps were analysed to determine the concentration of Total Protein and GFP, (only in the samples of the GFP tagged Affinity Tag and PC).

Table 5.1 - Concentration of the loading samples for each cellular extracts, in Total Protein and GFP.

Cellular Extracts	Protein Total (mg/ml)	GFP (mg/ml)	%GFP
GFP tagged RW	1.337	2.49E-2	1.86
GFP tagged NN	0.257	1.59E-2	6.19
PC (GFP without tag)	1.209	1.99E-2	1.65
NC (without GFP) (RW/NN)	(1.240/0.566)	----------	----------

The loading samples were made to have a GFP concentration of 1.75E-2mg/ml, the results demonstrated similar values for PC and GFP tagged with NN in GFP concentration. The GFP tagged with RW had a higher concentration than the others loading samples. The percentage of GFP in the solutions is similar for the PC and GFP tagged RW (2% of GFP), but the GFP tagged NN is 6% (Table 5.1). The GFP tagged compared with the other proteins should be high as possible to improve the detection of the affinity ligands of the affinity tag, maximizing the selectivity of the ligands. The concentration of the total proteins in the NC was different for the RW and NN ligands due to the differences in the content of total protein in the GFP tagged RW and GFP tagged NN cellular extracts (Figure 5.2).

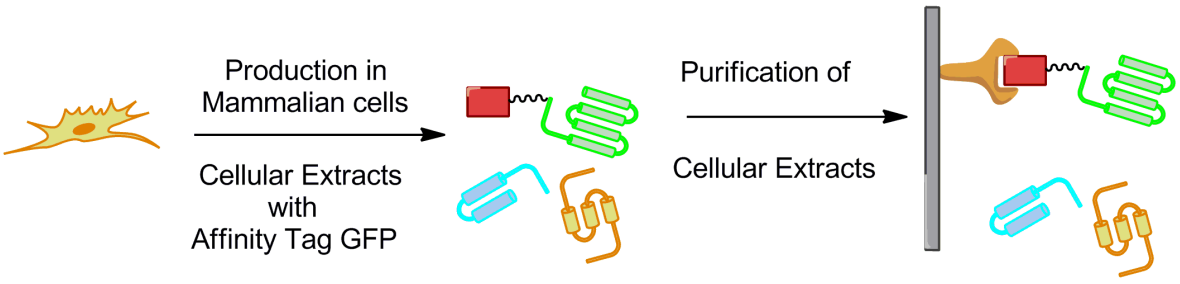


Figure 5.2 – Production and purification of a fused recombinant protein. The affinity and selectivity of the ligands in the purification step is important for the final yield of the fused protein.

The total protein and the GFP was determined for the washes samples and was analysed the binding affinity using the ratio of Total Protein Mass:Support Mass and the ratio GFP Mass: Support Mass. In one hand, the ratio of Total Protein Mass:Support Mass give the idea of the quantity of the proteins that stay in the columns, the ligands that have the lowest quantity in this ratio are the ones that can be more interesting, in principle they will adsorb less contaminants, and adsorb the protein of interest; also, if the ratio total protein is low in the

samples with GFP tagged and higher for the PC and NC that means that the ligands are more specific for the GFP tagged. In another hand, the ratio GFP Mass:Support Mass, gives the idea of the ligands that are specific for the affinity tag, in consequence high ratio GFP in support for the GFP tagged and low ratio of GFP in the support for the PC is what should be observed for a ligand that is specific.

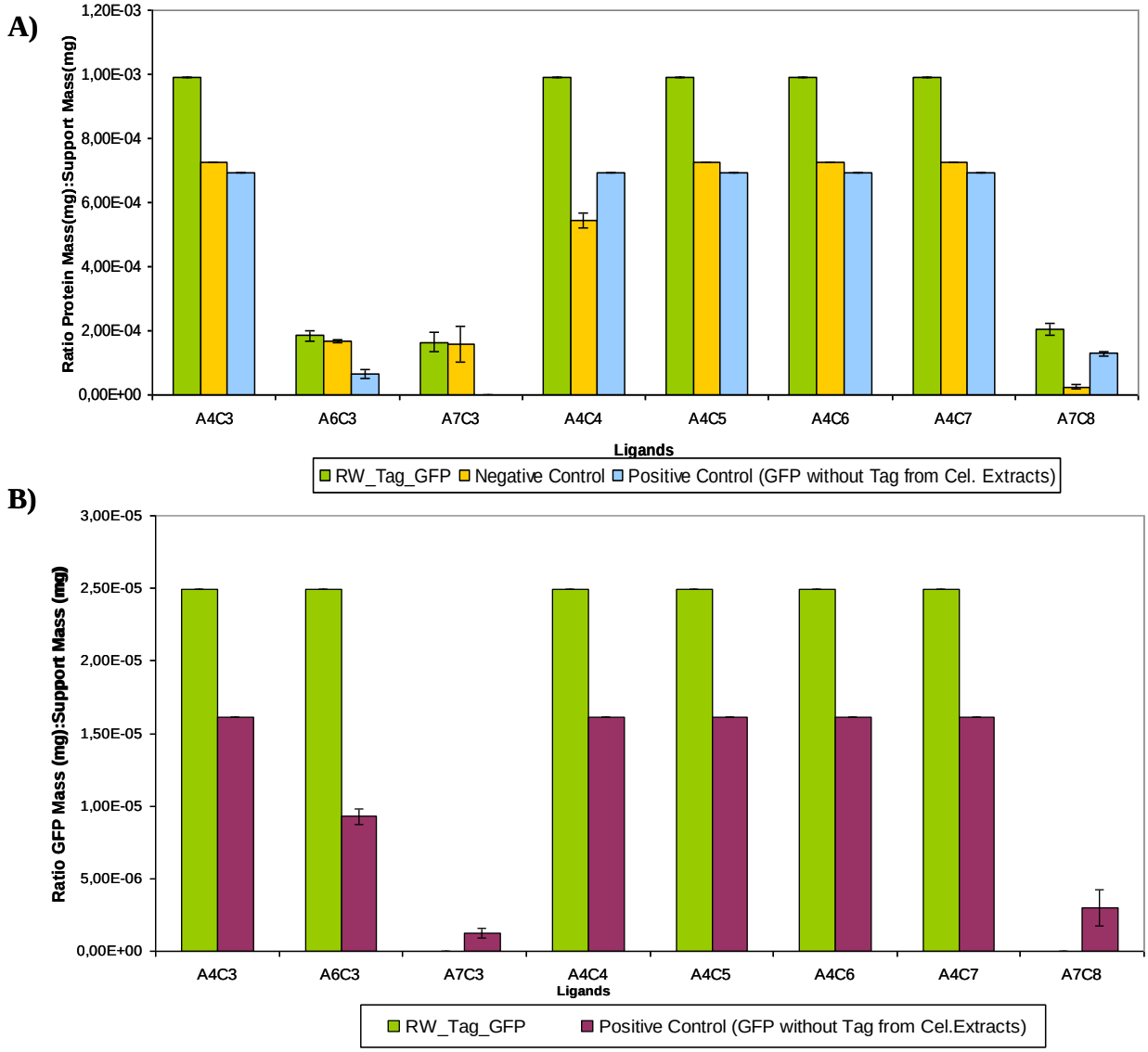


Figure 5.3 - Comparison between the cellular extracts of mammalian cells screened with hexapeptide RW ligands: A) Results of Ratio of Total Protein Mass: Support Mass and B) Results of Ratio GFP Mass: Support Mass.

The results for the RW ligands (Figure 5.3) demonstrate that these ligands retain the main quantity of protein when in the medium exists GFP tagged with hexapeptide RW. The quantity of Total Proteins retained with these extracts is higher than with NC extracts or the

PC. This could indicate high affinity for the hexapeptide RW. Despite these results the ligands A6C3, A7C3 and A7C8, have lowest binding results for all cellular extracts.

When in the Figure 5.3 the graph A is compared with B is possible to observe that for the majority of the ligands, the affinity for the GFP tagged with RW is higher than GFP without tag. Interestingly the ligand A6C3 is the ligand where the difference between what stays in the column for the GFP tagged RW and GFP is significantly different; then, is possible to say that for this ligand what is retained in the column is mostly the GFP tagged RW and the rest of the proteins are washed from the column. The ligands A7C3 and A7C8 do not have much affinity for the proteins in the mixture even when in the extract is present the hexapeptide RW.

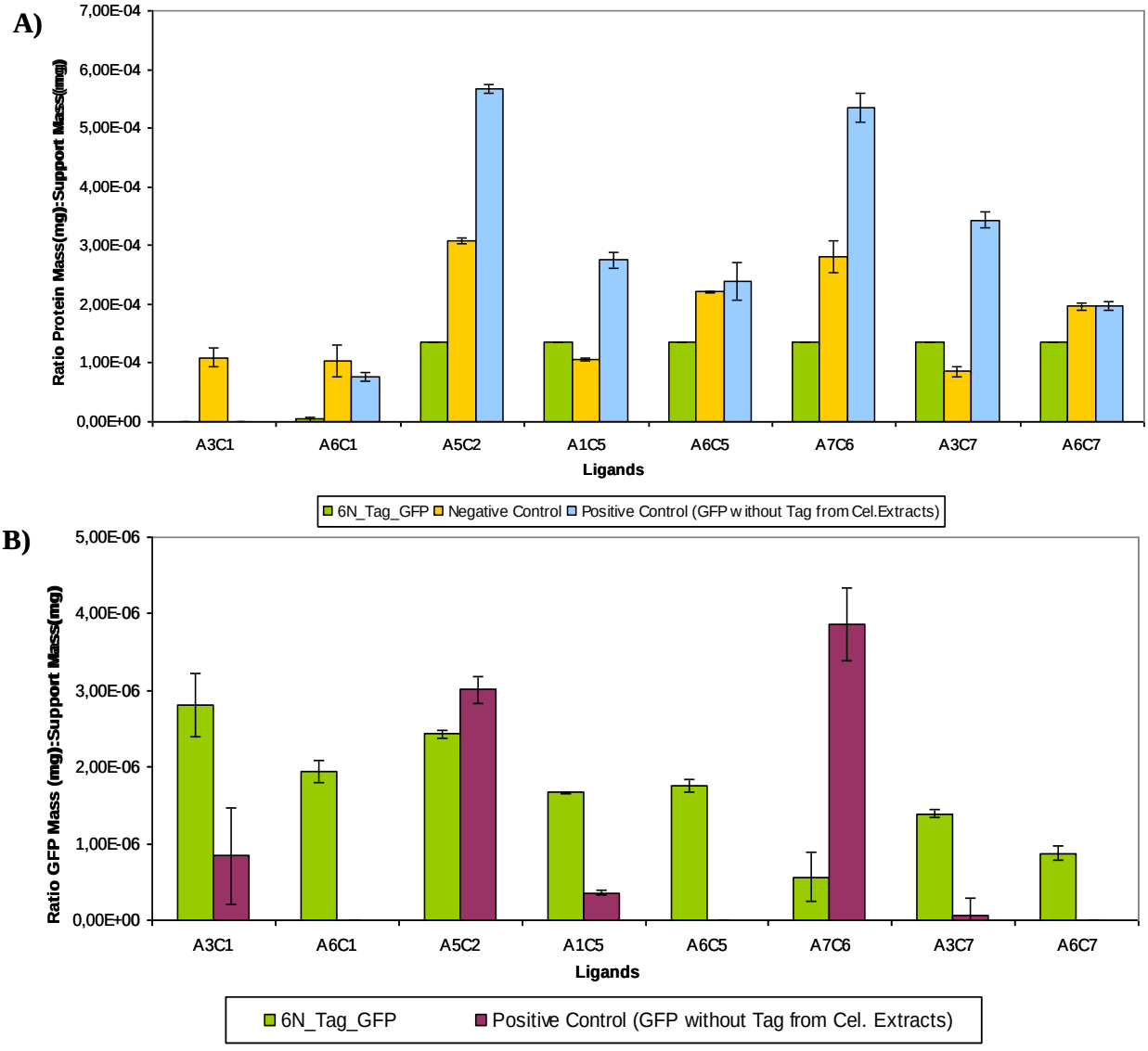


Figure 5.4 - Comparison between the cellular extracts of mammalian cells screened with Ugi Ligands for the hexapeptide NN: A) Results of Ratio of Total Protein Mass: Support Mass and B) Results of Ratio GFP Mass: Support Mass.

The results (Figure 5.4 - A) for NN ligands demonstrated that A3C1, A6C1 are the ligands that have the lowest ratio of total protein for the GFP tagged NN, that could indicate that the ligand is not selective. However, because in the quantification of GFP (Figure 5.4 - B) these ligands have the highest values for GFP tagged NN, which correspond to a high ratio GFP Mass:Support Mass, and represent that the ligands retain the GFP tagged NN preferentially. As a consequence we can say that these ligands are more selective for the GFP tagged than the GFP, especially the ligand A6C1.

For the A5C2 and A7C6 (Figure 5.4 - A) is possible to observe high values in the total protein for PC samples. Although the A5C2 the results for GFP quantification (Figure 5.4 – B) demonstrated that it is not selective when in the presence of GFP, because in the samples with GFP (GFP tagged and PC samples) it has approximately the same pattern. In contrast the A7C6 when in the loading sample is present GFP tagged NN and GFP, the last protein stays in the column, so the ligand has more affinity for GFP than the tag with the hexapeptide NN (Figure 5.4 – B).

The other ligands have lower values of ratio Total Protein: Support Mass, mainly for the samples that have GFP. Furthermore, when the GFP is quantified (Figure 5.4 – B), the quantity of GFP tagged NN that stays in the columns is higher than the GFP. With these results is possible to conclude that the ligands A1C5, A6C5, A3C7 and A6C7 are more selective for the tag with hexapeptide NN.

Overall the ligand A6C1 is the one that binds more to the GFP tagged, as a result is the ligand with the highest specificity for GFP tagged with the hexapeptide NN.

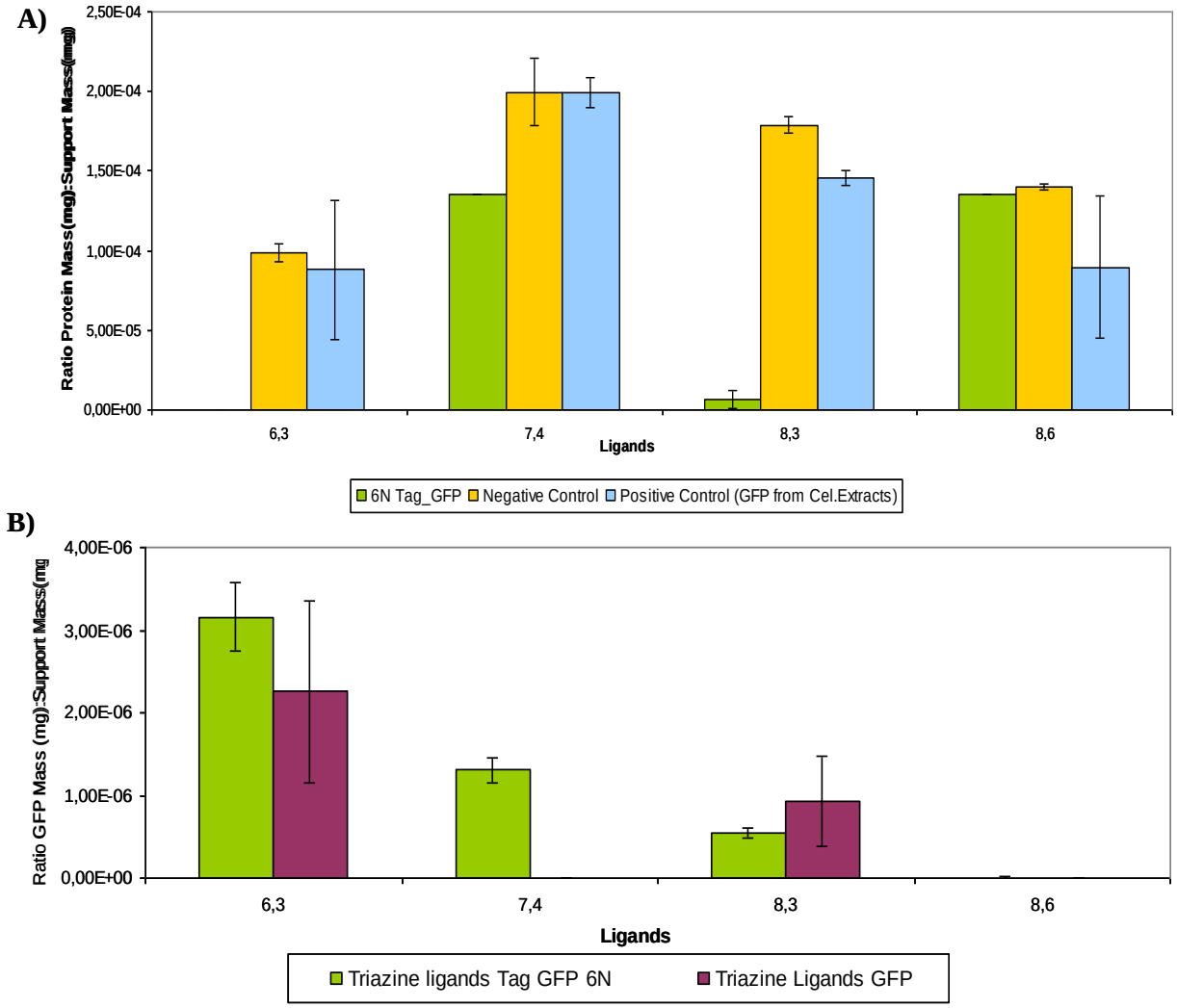


Figure 5.5 - Comparison between the cellular extracts of mammalian cells screened with Triazine Ligands for the hexapeptide NN: A) Results of Ratio of Total Protein Mass: Support Mass and B) Results of Ratio GFP Mass: Support Mass.

The results for Triazine Ligands show that the 6,3 is the ligand with lowest ratio Total Protein Mass:Support Mass (Figure 5.5 - A) for the samples of GFP tagged, although when the results of the fluorescence (Figure 5.5 -B) are observed is possible to conclude that GFP tagged was retained in the column. However it has also affinity for the GFP without tag.

Other ligand that have low values of ratio for total protein is the 8,3 (Figure 5.5 – A), although when is determined the mass of GFP that is retained in the column, the value for GFP without tag is higher than GFP tagged (Figure 5.5 – B).

The 7,4 ligand binds to proteins but not specifically to the GFP (Figure 5.5 – A), the PC and NC have a similar value, however when in the loaded sample there is GFP tagged it has a lower value.

Finally, the 8,6 ligand binds to the mixture of proteins from the loaded sample of GFP tagged, but has no specificity for the GFP tagged, since these protein is washed from the column (Figure 5.5 – B).
Despite these results, the Triazine ligands bind less total protein than the Ugi ligands and the ligand 7,4 is the ligand that binds more to the GFP tagged NN and noting to the GFP.
By the analysis of the elution results was not recovered any protein and the conditions of this procedure must be optimize.

5.2 SDS-PAGE Analysis

The samples of the most concentrated wash (first wash with PBS buffer) were analysed by 1D SDS-PAGE to understand the behaviour observed in the graphs of binding affinity.
This method allows the separation of proteins in each sample, based in the weight of the proteins, because this method occurs in denaturant conditions - the samples suffer a heat treatment in presence of ß-mercaptoethanol and SDS, that give to the proteins the same charge and shape eliminating the secondary and tertiary structure.
The GFP band will appear at 27kDa for the pure protein and the expected band for the GFP tagged with the hexapeptides is 29kDa.

5.2.1 Ligands for Hexapeptide RW

By comparing the results of 1D SDS-PAGE for the hexapeptide RW ligands of the different samples of cellular extracts, GFP tagged with hexapeptide RW, PC and NC is possible to observe that (Figure 5.6):
- For ligand A4C3 the proteins stay in the column so this ligand is not selective for any protein;
- Ligands A7C3 and A7C8 are letting the loaded solutions to pass freely through the resin, they are not retaining any protein;
- The ligands A4C4, A4C5, A4C6, A4C7 have a similar behaviour. For the loaded samples of PC and NC they bind to all the proteins. In the presence of the GFP tagged with the hexapeptide RW the ligands have more affinity and the GFP band does not appear in the gel and some proteins are washed from the column.

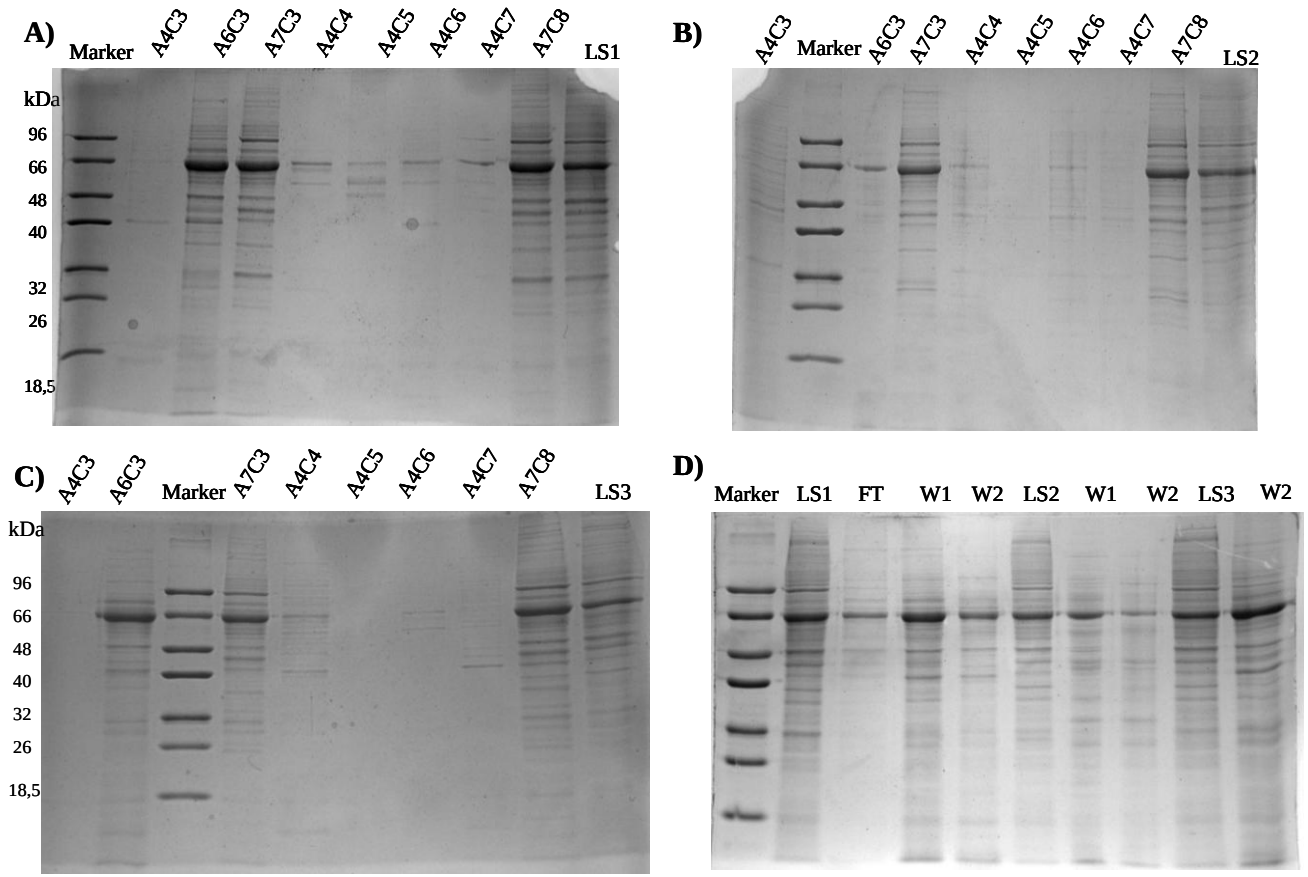


Figure 5.6 – 1D SDS-PAGE analysis of samples from the first wash with PBS, hexapeptide RW ligands. A) Columns Screened with GFP tagged RW cellular extracts, LS1 – Loaded Sample; B) Columns Screened with PC cellular extracts, LS2 – Loaded Sample; C) Columns screened with NC cellular extracts, LS3 – Loaded Sample. D) Samples from chromatographic columns with resin activated with ligand A6C3 and screened with different cellular extracts. Column screened with GFP tagged RW: LS1 – Loaded Sample with GFP tagged RW; FT- Flow through sample; W1 – First wash with PBS; W2 – Second Wash with PBS. Column screened with PC cellular extracts: LS2 – Loaded Sample PC cellular extracts; W1 – First wash with PBS; W2 – Second Wash with PBS. Column screened with NC cellular extracts: LS3 – Loaded Sample NC cellular extracts; W2 – Second Wash with PBS. Marker – LMW marker.

- Finally, ligand A6C3 has the most interesting behaviour because in the NC gel all the proteins are washed away; in the PC gel only a small band at 66 kDa is observed it has affinity for the most of the proteins and GFP; however, when the sample have GFP tagged with the hexapeptide it is very selective and all the proteins are washed off the column but the band of the GFP tagged is almost invisible, therefore is possible to conclude that this is the most specific ligand for the hexapeptide RW.

The results of the SDS-PAGE are in agreement with the Fluorescence results and the analysis of the SDS-PAGE only makes more clear the information provided by the fluorescence.

To better understand the selectivity of ligand A6C3 for GFP tagged RW other washes were analysed (Figure 5.6 - D). The results in this gel confirm the observation that this ligand is very selective comparing with the other ligands for the hexapeptide RW.

The second wash with PBS was analysed for the ligands (A4C4, A4C5, A4C6 and A4C7), although the proteins are not washed of the columns, then the ligand interact with all kind of proteins and are not selective for the hexapeptide RW.

5.2.2 Ligands for Hexapeptide NN

The results of the 1D SDS-PAGE for the Hexapeptide NN ligands based on the Ugi reaction (Figure 5.7) demonstrated that all the ligands do not bind proteins in the NC (Figure 5.7 - C). For the PC extracts (Figure 5.7 - B) the ligands bind to all the material loaded. For the ligands A3C1, A6C1 and A7C6 some proteins are washed out of the column specially one with a molecular weight of 66kDa (molecular weight of Albumin from bovine serum, used as a marker).

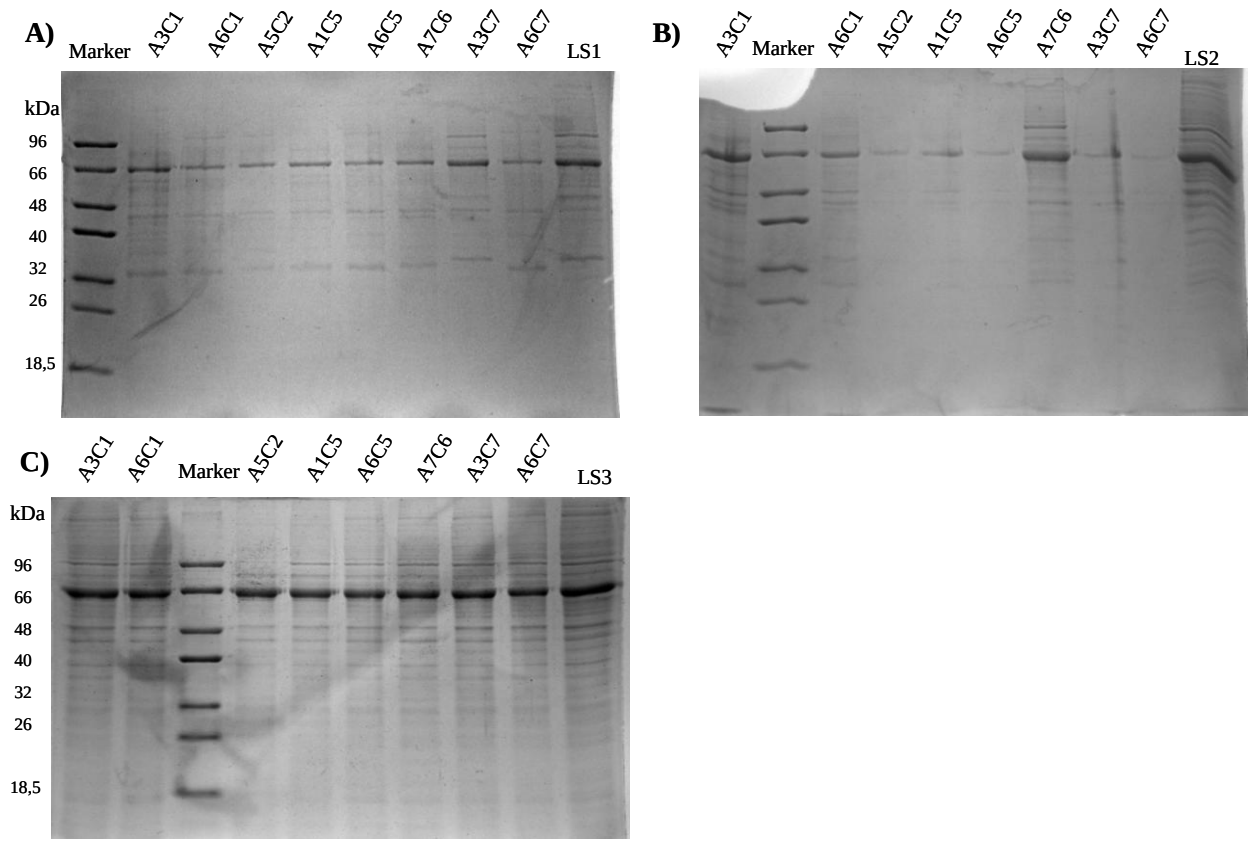


Figure 5.7 – 1D SDS-PAGE analysis of hexapeptide NN ligands (Ugi Reaction) samples. The samples loaded in the gel were represented the first wash with PBS after the screening of the ligands with different cellular extracts: A) Ligands columns were screened with GFP tagged NN cellular extracts, LS1 – Loaded Sample; B) Ligands columns were screened with PC cellular extracts, LS2 – Loaded Sample; C) Ligands columns were screened with NC cellular extracts, LS3 – Loaded Sample. Marker – LMW marker.

CHAPTER 5 – CHARACTERIZATION OF POSSIBLE LEAD LIGANDS

In the results for the GFP tagged with hexapeptide NN, all ligands retain the proteins. From these results is difficult to identify the band of GFP, probably because it is in a low concentration and the resolution of the gel does not allow observing that band. These results bring concerns about the specificity of the ligands for the hexapeptide NN.

The results for the hexapeptide NN ligands based on the Triazine scaffold (Figure 5.8) show for NC samples that the ligands do not bind to the proteins. For the PC samples the ligands have more affinity for the protein when there is GFP, although 6,3 ligand has the lowest affinity for the proteins in this mixture and the GFP is also washed from the column. In the samples with GFP tagged NN there is more protein washed from the columns although the apparent band of GFP tagged NN (32kDa) appears in all the ligands (Figure 5.8 – A). The ligand that has the lowest affinity for the other proteins and the highest affinity for the GFP tagged NN is the 6,3 ligand.

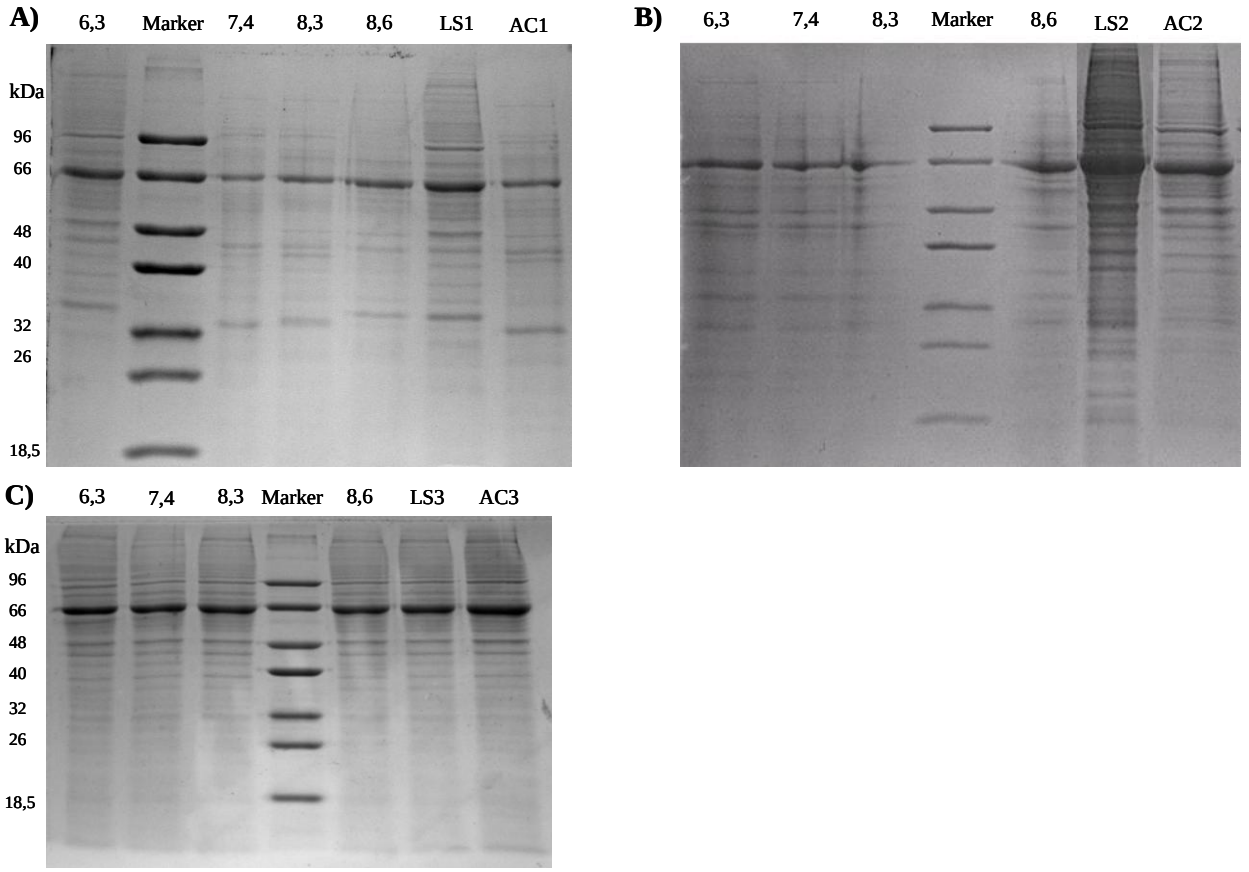


Figure 5.8 – 1D SDS-PAGE analysis of hexapeptide NN ligands (Triazine based Reaction) samples. The samples loaded in the gel were represented the first wash with PBS after the screening of the ligands with different cellular extracts: A) Ligands columns were screened with GFP tagged NN cellular extracts, LS1 – Loaded Sample; B) Ligands columns were screened with PC cellular extracts, LS2 – Loaded Sample; C) Ligands columns were screened with NC cellular extracts, LS3 – Loaded Sample. Marker – LMW marker.

In these SDS-PAGE gels were added samples from the first wash with PBS of a control (AC1, AC2 ad AC3, Figure 5.8). This control is blank agarose that was packed in a chromatographic column in the same conditions as the ligands and was screened with the cellular extracts. This control helps to understand the contribution of the agarose without functionalization in the process of purification. In the ideal situation the support does not interact with the products of the cellular extracts. From these results is possible to infer that the agarose has some affinity for the proteins in the cellular extracts, since part of the proteins are retained in the column.

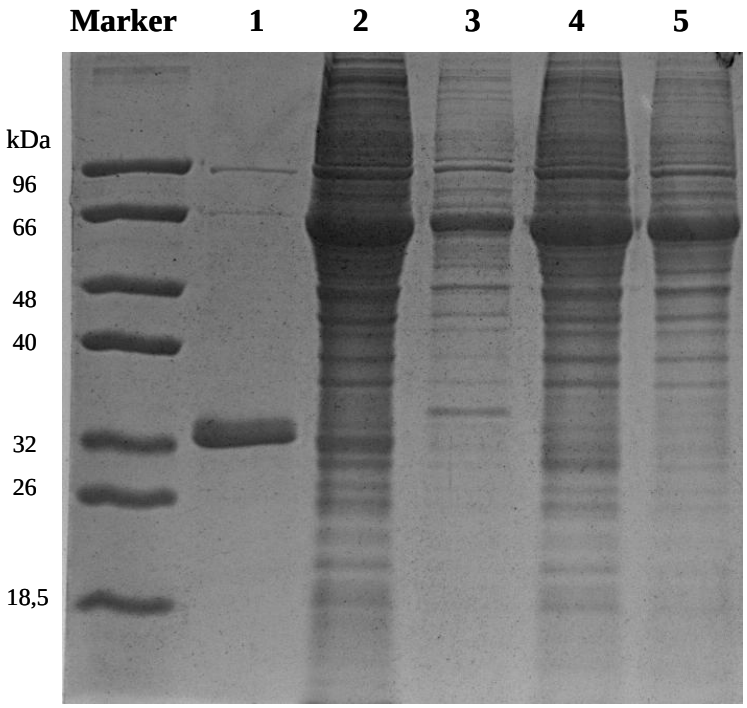


Figure 5.9 – 1D SDS-PAGE Results for the following loaded samples: 1) TurboGFP (1µg/µl) (Evrogen), 2) Loaded Sample Cellular Extracts GFP tagged RW; 3) Loaded Sample cellular extracts GFP tagged NN; 4) Loaded Sample cellular extracts PC (GFP without tag); 5) Loaded Sample cellular extracts NC (cellular extract without GFP).

By comparing the results (Figure 5.9) is possible to observe that the concentration of GFP in the samples is lower than pure GFP (TurboGFP), and the GFP tagged NN has the lowest concentration. To correctly identify the band of GFP in the cellular extracts it could be use different approaches: concentrate the samples before applying in the gel, using other method of staining or make a Western-Blot procedure and incubate with an antibody specific for GFP.

This last method is frequently used when is necessary to identify the protein of interest in a gel with a mixture of unknown molecular weight proteins (Walker 2005).
From these results of SDS-PAGE for Ugi and Triazine ligands of the hexapeptide NN the concentration of the samples is too low, is difficult to identify the bands of the proteins. To improve the results would be useful to use more concentrated samples.

These results from the cellular extracts were unexpected; since some ligands that had highest affinities for the hexapeptide, namely A6C5, A3C1 or A7C6, had in these assays low affinity for the GFP tagged NN or had more affinity for the GFP. This is the realistic test for the selection of the lead ligands for each hexapeptide since this will be the normal conditions that they will face, if any of these ligands became a separation pair for the hexapeptide. Furthermore, the purification of cellular extracts is important to understand the behaviour of these ligands in situations where there are molecules that compete with protein labelled with the affinity tag (Bradner et al. 2006). One of the conclusions that come out with these results is the necessity to increase the quantity of protein fused with the hexapeptide compared to the overall proteins since that would improve the sensivity and affinity of the ligand to the protein of interest. In consequence more work should be done on the improvement of fused protein production.

5.3 Characterization of the Lead Ligands

5.3.1 Hexapeptide RW Lead Ligand

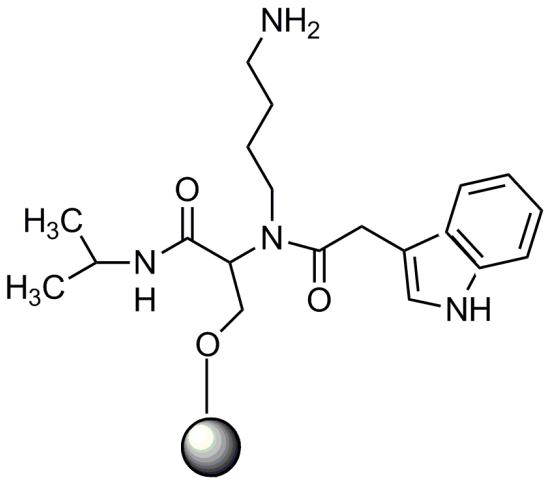


Figure 5.10 - Structure of ligand A6C3.

The lead ligand for the hexapeptide RW is the A6C3 that is composed by amine 1,4-Diaminbutane and the carboxylic acid 3-Indoleacetic acid.
The hexapeptide RW has tryptophan (W) and arginine (R), the first is hydrophobic and the second one has an hydrophilic character, this gives to the peptide a mix character.
The lead ligand has the same mixture so it can interact with the hydrophilic and hydrophobic aminoacids of the hexapeptide (Figure 5.10).

5.3.2 Hexapeptide NN Lead Ligands

For the hexapeptide NN, eight ligands were selected from the affinity screening of 96 wells plate that had the highest affinity for the hexapeptide NN and lowest affinity for the GFP, from these ligands one had the best results the A6C5 (Figure 5.11). This ligand was synthesized in solution phase (yield of 18%) and was characterized using the TLC, and 1H and 13C NMR as well as melting point analysis.

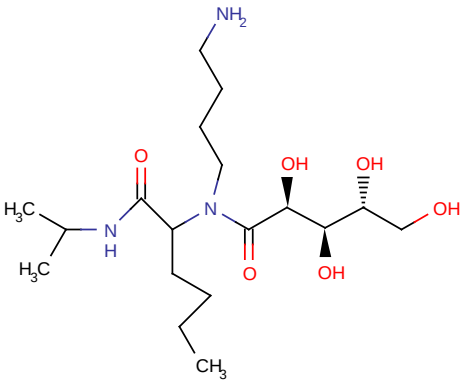


Figure 5.11 - Structure of Ligand A6C5.

To determine the melting point a sample of the compound was heated, with controlled temperature, until 200ºC. At this temperature the degradation of the product was observed, since the aspect of the compound was the same, white pounder, and there were no alterations during the heating process is possible to say that the compound was relatively pure. The TLC did not show any results, because this ligand does not have aromatic ring nor double carbon bonds.

Finally, the product was analysed by NMR 1H at 65ºC, as the structure of the compound is too mobile and an increase in the temperature increases the rotational energy and have less variability in position of the hydrogen groups. The results demonstrated that some characteristic groups are present in the compound like three CH_3 groups (1,5ppm), the H protons near NH groups (2.1ppm), the proton of the CH_2–CH_2 (4ppm) and the protons near the OH groups (5 – 5.5 ppm) (Appendix 1 and figure 5.8). The NMR 13C demonstrated that the characteristic groups of this compound are present in the sample, like C-$(CH_3)_2$ (30ppm), the C-OH (70ppm) and HN-C=O (175ppm) (Appendix 2). The results of the 13C NMR demonstrated that the sample was not completely pure since is possible to observe a noise in the spectrum.

The results for the Triazine based ligands demonstrated that the 6,3 or the 7,4 ligands are the more selective for the hexapeptide NN (Figure 5.12).

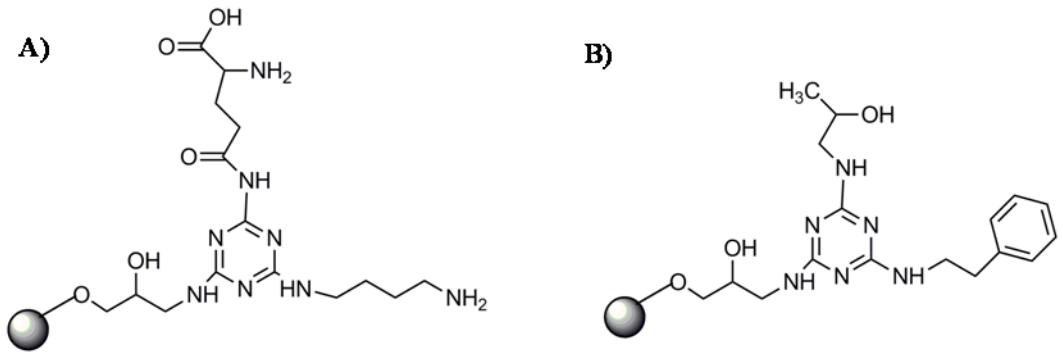


Figure 5.12 – Structures of the Triazine based ligands: A) 6,3 ligand and B) 7,4 ligand.

The 6,3 ligand is composed by two amines, L-Glutamine and 1,4-Diaminobutane, these two amines have hydroxyl groups and nitrogen that can make hydrogen bonds with the hexapeptide NN. The 7,4 ligand is composed by amines, 1-Amino-2-Propanol and Phenethylamine that has the possibility to make hydrogen bonds with the hexapeptide.

5.4 Conclusions

A few lead ligands were tested to determine the selectivity for the hexapeptide with cellular extracts.

The results demonstrated that the ligands that had more affinity for the hexapeptide RW in a pure solution were less selective when it is present in a mixture of proteins. Although was identified one ligand A6C3 that is selective for the hexapeptide RW, since when the affinity tag is present all the proteins are washed from the column and not the GFP tagged RW.

For the hexapeptide NN were selected 8 Ugi ligands and 4 Triazine based ligands from the studies with pure solutions of hexapeptide NN and GFP. From the results with the cellular extracts is possible to conclude that all the Ugi ligands appear to have more affinity for the GFP, because the proteins are washed from the columns when it is present. What was already observed when as used the pure GFP. When the GFP tagged NN is present the majority of proteins are washed but also the GFP tagged, then the ligands are not being selective for the

tag. The Triazine ligands are more selective for the GFP tagged NN specially the 6,3 and 7,4 ligands. Despite these results is not possible to say that the hexapeptide NN has an optimum ligand. The concentration of the GFP in the cellular extracts can be increased to make easier the visualization in the SDS-PAGE method. To overcome this situation more concentrated solutions should be use, or other staining methods that are more selective for lower concentration of proteins. Furthermore to understand the interactions between these ligands and the hexapeptide automated docking can be made.

Once more the elution of the proteins from the columns was not observer, then this method needs to be optimized.

CHAPTER 6 - CONCLUSIONS

In this work the affinity between synthetic ligands and an hexapeptide constituted by asparagines (N) for application in the affinity purification of recombinant fusion proteins has been studied.

Two combinatorial libraries with 64 affinity ligands, one based in the Ugi Multicomponent Reaction and other in the 1,3,5-Triazine scaffold, were synthesized in solid-phase on a 96 well plate format.

The libraries were screened with pure solutions of hexapeptide NN and GFP (model protein). The hexapeptide concentration was determined by the BCA assay and the GFP by its intrinsic fluorescence. From these results, 8 Ugi ligands and 4 Triazine based ligands presented preferential high binding to the hexapeptide and low for the GFP, and were selected for further studies. The Ugi ligands bound approximately to 40% of the loaded hexapeptide, whereas the triazine ligands bound only 20% of the loaded peptide. However, the selectivity of the triazine ligands seemed higher in comparison to the Ugi ligands. The lead ligands were produced in more quantity and screened with a pure solution of hexapeptide NN, in these results was observed a decrease in the affinity, that can be related with the screening mode, the quantity of ligands in the surface of the resin and the geometry of the chromatographic columns. The Ugi ligands were synthesized with higher concentration in the surface of the resin (40µmol/g support) that can have an inhibitory effect in the ligands affinity. The Triazine based ligands that had approximately the same concentration of ligands in the resin (20 µmol/g support) in every study slightly decrease the affinity for the hexapeptide.

To study the affinity of the ligands in the real conditions, recombinant fusion proteins (GFP tagged with the hexapeptide) were expressed in mammalian HEK 293T cells.

The mammalian cells were transfected with vectors by the PEI method. Three vectors were designed and used which codified for: hexapeptide NN fused with GFP; hexapeptide RW fused with GFP (used for screening of RW lead ligands - another affinity tag under study); and, a GFP without tag. The transfection was successful with a 50-60% rate of transfection, and was observed fluorescence of GFP which indicates that the affinity tag does not interfere with the conformation of the protein. To collect the cells from the culture plates cell scrape and to extract the intracellular proteins the Freeze-Thaw method was used. The concentration

of the total protein in the cellular extracts and the concentration of GFP was determined. The results demonstrated that the GFP present in the samples was 1- 4% of the total proteins in the samples. Despite these results the design of the expression vector can be improved to increase the efficiency of the production of the fused GFP.

Crude extracts from the mammalian cell growth were then loaded onto small chromatographic columns containing agarose modified with the lead ligands. The results demonstrated that the ligands that had more affinity for the hexapeptide RW in a pure solution were less selective when it was present in a mixture of proteins. However, ligand A6C3 has been identified as the most selective for this hexapeptide. More studies must be conducted in the future in order to confirm these data and determine affinity constants and optimize the ratio of ligands in the surface of the resin, to improve this affinity pair (ligand A6C3 and hexapeptide RW). For the hexapeptide NN, the Ugi ligands presented high affinity for the GFP, as was already observed in the screenings with pure GFP and confirmed by the automated docking results. When the GFP tagged NN is present the other proteins are washed out of the column, but this protein was not completely retained in the column. However, the Triazine lead ligands showed to be more selective for the GFP tagged NN, specially the ligands 6,3 and 7,4. Despite these results was not possible to determined one ligand with high specificity for the hexapeptide NN. The concentration of the GFP tagged in the samples must be higher than 1,75E-2mg/ml, in order to facilitate the analysis of the SDS-PAGE gels. Also to improve the detection of the protein bands it could be used other method of staining or the samples can be concentrated. In general, the elution procedures also need to be studied in detail in the future.

CHAPTER 7 - REFERENCES

1. Akritopoulou-Zanze, I. (2008). "Isocyanide-based multicomponent reactions in drug discovery." Current Opinion in Chemical Biology **12**: 324-331.
2. Armstrong, R. W., A. P. Combs, et al. (1996). "Multiple Component Condensation Strategies for Combinatorial Library Synthesis." Acc. Chem. Res. **29**: 123-131.
3. Arnau, J., C. Lauritzen, et al. (2006). "Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins." Protein Expression and Purification **48**: 1-13.
4. Bradner, J. E., O. M. McPherson, et al. (2006). "A Robust Small-Molecule Microarray Platform for Screening Cell Lysates." Chemistry and Biology **13**: 493-504.
5. Bertschinger, M., G. Backliwal, et al. (2006). "Disassembly of polyethylenimine-DNA particles in vitro: Implications for polyethylenimine-mediated DNA delivery." Journal of Controlled Release **116**(1): 96-104.
6. Cass, B., P. L. Pham, et al. (2005). "Purification of recombinant proteins from mammalian cell culture using a generic double-affinity chromatography." Protein Expression and Purification **40**: 77-85.
7. Chatterjee, D. K. and D. Esposito (2005). "Enhanced soluble protein expression using two new fusion tags." Protein Expression and Purification.
8. Chung, H. J., A. Steplewski, et al. (2009). "Fluorescent protein markers to tag collagenous proteins: The paradigm of procollagen VII." Biochemical and Biophysical Research Communications **390**: 662-666.
9. Clonis, Y. D. (2006). "Affinity chromatography matures as bioinformatic and combinatorial tools develop." Journal of Chromatography A **1101**: 1-24.
10. Davies, J. W., M. Glick, et al. (2006). "Streamlining lead discovery by aligning in silico and high-throughput screening." Current Opinion in Chemical Biology **10**: 343-351.
11. Derewenda, Z. S. (2004). "The use of recombinant methods and molecular engineering inprotein crystallization." Methods **34**: 354-363.
12. Dömling, A. (2006). "Recent developments in Isocyanide Based Multicomponent Reactions in Applied Chemistry " Chem. Rev. **106**: 17-89.
13. Einhauer, A. and A. Jungbauer (2001). "The FLAG peptide, a versatile fusion tag for the purification of recombinant proteins." Journal of Biochemical and Biophysical Methods **49**: 455-465.
14. Forster, M. J. (2002). "Molecular modelling in structural biology." Micron **33**: 365-384.
15. Geisse, S. and C. Fux (2009). Chapter 15 Recombinant Protein Production by Transient Gene Transfer into Mammalian Cells. Methods in Enzymology. R. B. Richard and P. D. Murray, Academic Press. **Volume 463:** 223-238.
16. Graumann, K. and A. Premstaller (2006). "Manufacturing of recombinant therapeutic proteins in microbial systems." Biotechnology Journal **1**: 164-186.
17. Gupta, P. and K. H. Lee (2007). "Genomics and proteomics in process development: opportunities and challenges." Trends in Biotechnology **25**(7): 324-330.
18. Haigh, J. M., A. Hussain, et al. (2009). "Affinity ligands for immunoglobulins based on the multicomponent Ugi reaction." Journal of Chromatography B **877**: 1440-1452.
19. Hedhammar, M., T. Gräslund, et al. (2005). "Protein Engineering Strategies for Selective Protein Purification." Chem. Eng. Technol. **28**(11).
20. Illgen, K., T. Enderle, et al. (2000). "Simulated molecular evolution in a full combinatorial library." Chemistry & Biology **7**: 433-441.

CHAPTER 7 - REFERENCES

21. Inouye, S. and F. I. Tsuji (1994). "Aequorea green fluorescent protein expression of the green fluorescence characteristics of the recombinant protein." FEBS Letters **341**: 277-280.
22. Issaq, H. J., K. C. Chan, et al. (2009). "Separation, detection and quantification of peptides by liquid chromatography and capillary electrochromatography." Journal of Chromatography A **1216**: 1825-1837.
23. Jenkins, N., L. Murphy, et al. (2008). "Post-translational Modifications of Recombinant Proteins: Significance for Biopharmaceuticals." Molecular Biotechnology **39**: 113-118.
24. Kenig, M., S. Peternel, et al. (2006). "Influence of the protein oligomericity on final yield after affinity tag removal in purification of recombinant proteins." Journal of Chromatography A **1101**: 293-306.
25. Labrou, N. E. (2003). "Design and selection of ligands for affinity chromatography." Journal of Chromatography B **790**: 67-78.
26. Lazaridis, T. and M. Karplus (2000). "Effective energy function for protein structure prediction." Current Opinion in Structural Biology **10**: 139-145.
27. Li, Y., X. Zhang, et al. (2004). "Synthesis of cluster mannosides via a Ugi four-component reaction and their inhibition against the binding of yeast mannan to concanavalin A." Carbohydrates Research **339**: 873-879.
28. Low, D., R. O'Leary, et al. (2007). "Future of antibody purification." Journal of Chromatography B **848**(1): 48-63.
29. Lowe, C. R. (2001). "Combinatorial approaches to affinity chromatography." Current Opinion in Chemical Biology **5**: 248-256.
30. Lowe, C. R., A. R. Lowe, et al. (2001). "New developments in affinity chromatography with potential application in the production of biopharmaceuticals." Journal of Biochemical and Biophysical Methods **49**(1-3): 561-574.
31. Midas, I. H. (2009). Top 15 Global Therapeutic Classes, 2009, Total Audited Markets 2009 Global Prescription Sales Information IMS Health Incorporated**:** 1.
32. Mondal, K. and M. N. Gupta (2006). "The affinity concept in bioseparation: Evolving paradigms and expanding range of applications." Biomolecular Engineering **23**: 59-76.
33. Nakamura, M., J. Inoue, et al. (2000). "A Two-Step, One-Pot Synthesis of Diverse N-Pyruvoyl Amino Acid Derivates Using the Ugi Reaction " Bioorganic & Medical Chemistry Letters **10**: 2807-2810.
34. Palanisamy, U. D., D. J. Winzor, et al. (2000). "Synthesis and evaluation of affinity adsorbents for glycoproteins: an artificial lectin." Journal of Chromatography B: Biomedical Sciences and Applications **746**(2): 265-281.
35. PhaRM (2008). "Biotecnology Research Continuos to Bolster Arsenal Against Disease with 633 medicines in development." Report - Medicines in Development Biotecnology 1-60.
36. Pina, A. S. (2008). "Synthesis and Screening of Rationally Designed Combinatorial Libraries for Proteomics Applications". Chemistry Department. Lisbon, New University of Lisbon, Faculty of Science and Technology. Master.
37. Platis, D., A. Maltezos, et al. (2009). "Combinatorial *de novo* design and application of a biomimetic affinity ligand for the purification of human anti-HIV mAb 4E10 form trangenic tobacco." Journal of Molecular Recognition **22**: 415-424.
38. Platis, D., C. A. Sotriffer, et al. (2006). "Lock-and-key motif as a concept for designing affinity adsorbents for protein purification " Journal of Chromatography A **1128**: 138-151.

CHAPTER 7 - REFERENCES

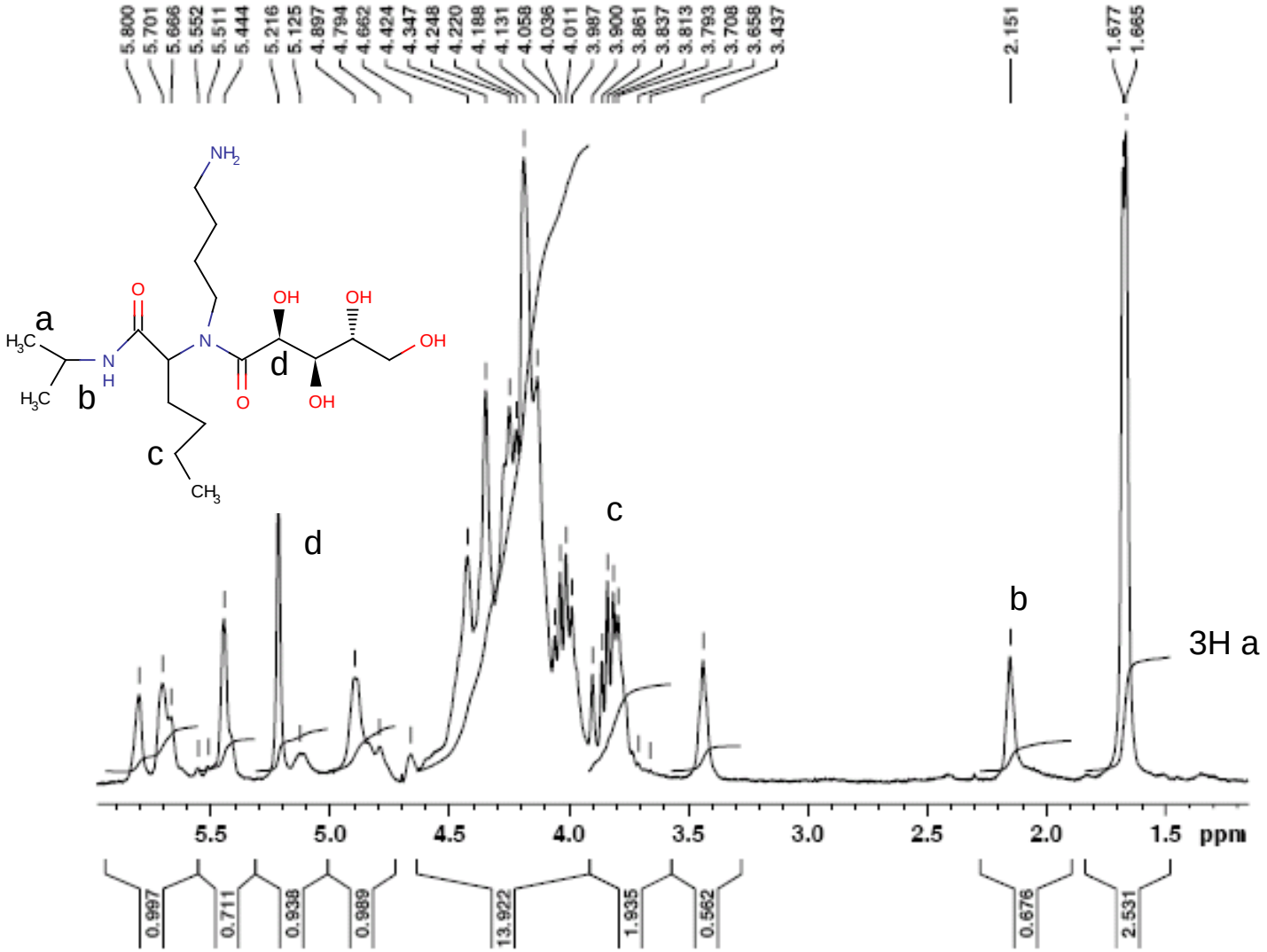
39. Roque, A. C. A., S. O. C. Silva, et al. (2007). "Affinity based methodologies and ligands for antibody purification: Advances and perspectives." Journal of Chromatography A **1160**: 44-55.
40. Roque, A. C. A., M. Â. Taipa, et al. (2004). "A new method for screening of solid-phase combinatorial libraries for affinity chromatography." Journal of Molecular Recognition **17**: 262-267.
41. Roque, A. C. A., M. Â. Taipa, et al. (2005). "Synthesis and screening of a rationally designed combinatorial library of affinity ligands mimicking protein L from *Peptostreptococcus magnus*." Journal of Molecular Recognition **18**: 213-224.
42. Roskams, J. and L. Rodgers (2002). Electrophoresis of DNA, RNA, and Proteins. Lab Ref - A Handbook of Recipes, Reagents, and Other Reference Tools for Use at the Bench. J. Roskams and L. Rodgers. Cold Spring Harbor Cold Spring Harbor Laboratory Press
43. Seth, G., S. Charaniya, et al. (2007). "In pursuit of a super producer - alternative paths to high production recombinant mammalian cells." Current Opinion in Biotechnology **18**: 557-564.
44. Sevastsyanovich, Y., S. Alfasi, et al. (2009). "Recombinant protein production: a comparative view on host physiology." New Biotechnology **25**(4): 175-180.
45. Solomons, T. W. G. and C. B. Fryhle (2004). Organic Chemistry, John Wiley & Sons, Inc.
46. Shukla, A. A. and J. Thömmes (2010). "Recent advances in large-scale production of monoclonal antibodies and related proteins." Trends in Biotechnology **28**(5): 253-261.
47. Shukla, A. A., B. Hubbard, et al. (2007). "Downstream processing of monoclonal antibodies-Application of platform approaches." Journal of Chromatography B **848**(1): 28-39.
48. Smith, G. R. and M. J. E. Sternberg (2002). "Prediction of protein-protein interactions by docking methods." current Opinion in Structural Biology **12**: 28-35.
49. Sproule, K., P. Morrill, et al. (2000). "New strategy for the design of ligands for the purification of pharmaceutical proteins by affinity chromatography." Journal of Chromatography B **740**: 17-33.
50. Terpe, K. (2003). "Overview of tag protein fusions: from molecular and biochemical fudamentals to commercial systems." App. Microbiol. Biotechnol. **60**: 523-533.
51. Thomsen, R. and M. H. Christensen (2006). "MolDock: A New Technique for High-Accuracy Molecular Docking." J. Med. Chem. **49**: 3315-3321.
52. Thor, J. J. v., G. Y. Georgiev, et al. (2005). "Ultrafast and Low barrier Motions in the Photoreactions of the Green Fluorescent Protein." The Journal of Biological Chemistry **280**(39): 33652-33659.
53. Tozzi, C., L. Anfossi, et al. (2003). "Affinity Chromatography techniques based on the immobilization of peptides exhibiting specific binding activity." Journal of Chromatography B **797**: 289-304.
54. Ugi, I. (2001). "Recent progress in chemistry of multicomponent reactions." Pure Appl. Chem. **73**(1): 187-191.
55. Walker, J. M. (2005). Electrophoretic techniques. Principles and Techniques of Biochemistry and Molecular Biology K. Wilson and J. Walker, Cambridge University Press.
56. Waugh, D. S. (2005). "Making the most of affinity tags." TRENDS in Biotecnology **23**(6).

CHAPTER 7 - REFERENCES

57. Wu, F. and J. Yu (2007). "Novel biomimetic affinity ligands for human tissue plasminogen activator." Biochemical and Biophysical Research Communication **355**: 673-678.
58. Zhuang, R., Y. Zhang, et al. (2008). "Purification of GFP fusion proteins with high purity and yield by monoclonal antibody-coupled affinity column chromathography." Protein Expression and Purification **59**: 138-143.

APPENDIX

Appendix 1 – Spectrum 13C NMR – Affinity Ligand A6C5 from Ugi Combinatorial Library:



APPENDIX

Appendix 2 – Spectrum 13C NMR - Affinity Ligand A6C5 from Ugi Combinatorial Library: