UNIVERSITA' DEGLI STUDI DI PARMA

Dottorato di ricerca in Scienze del Farmaco, delle Biomolecole e dei Prodotti per la Salute

Ciclo XXIX

Chemical modification and characterization of proteins and peptides for pharmaceutical and biomedical applications

Coordinatore:

Chiar.mo Prof. Marco Mor

Tutor:

Chiar.mo Prof. Stefano Bettati

Chiar.mo Prof. Luca Ronda

Dottorando: Riccardo Piano

Contents

Chapter 1

Introduction

Based on the building block of modern molecular biology, proteins are the products of the genetic information preserved in the DNA. They are involved in a plethora of vital biological functions, i.e. from the catalysis of biochemical reactions to the formation of receptors, carriers, and membrane channels [1]. 25,000–40,000 different genes are currently estimated in the human genome and the number of functionally distinct proteins is much higher [2,3]. Proteins can be the subject of focused studies aimed at the understanding of disease mechanisms and/or at their exploitation for therapeutic purposes.

To date, more than 130 different proteins or peptides are approved for clinical use by the US Food and Drug Administration (FDA), and many more are in development [4]. They are mostly produced by recombinant DNA technology and purified from a wide range of organisms, whose choice is mainly based on post-translational modification required for the protein activity and on the production costs [4]. The best example of trends in the production and use of protein therapeutics is provided by the history of insulin in the treatment of diabetes mellitus type I (DM‑I) and type II (DM-II). In 1922 insulin was first purified from bovine and porcine pancreas with high protein purification costs and immunological side effects; in 1982 the US FDA approved recombinant insulin, for the therapy of DM-I, as the first commercially available recombinant protein therapeutic [5,6]. Thanks to the recombinant DNA technology, proteins can be more efficiently produced and modified to improve their function or to provide a novel activity.

The currently approved protein therapeutics can be classified based on their pharmacological action (from Group I to Group IV) [4].

Protein therapeutics classified in Group I are enzymes and regulatory proteins whose function can be related to the replacement of a protein that is deficient or abnormal and

to the enhancement of an existing pathway; the administration of proteins can provide also novel functions when the body is lacking a particular activity. Proteins such as immunoadhesins and monoclonal antibodies, belonging to Group II, exert their targeting activity by binding specifically to molecules or organisms and blocking their function directly or through the activation of signaling pathways. Group III and Group IV proteins are respectively protein vaccines, that provide protection against infectious agents, and protein diagnostics that are critical to medical decision before therapy [4].

Challenges in protein therapeutics

Protein therapeutics dramatically increases in number and frequency of use over the years and their pharmacokinetic evaluation plays an important role in their development. In fact, potential protein therapy failure is often related to a number of challenges that are faced in the development and use of these compounds. Protein solubility, stability, route of administration and distribution are factors that can hinder the successful application of a protein therapy [1,7]. Proteins may undergo different degradation pathways during their lifecycle including expression, purification, storage and administration. One of the most common instability challenges is aggregation, often due to the loss of native, three-dimensional structure resulting in the formation of high molecular weight (HMW) species and possibly sub-visible aggregates. Moreover, poor solubility can lead to protein precipitation, formation of visible particles and surfaces absorption resulting in loss of a soluble, therapeutically active, protein [8]. Proteins can be prone to chemical degradation reactions such as deamidation (asparagine, less often glutamine), isomerization (aspartic acid), oxidation (methionine, tryptophan, tyrosine), etc. The peptide bonds can be hydrolyzed and different protein chemical degradation pathways include loss of sugar moieties for glycosylated proteins or loss of disulfide bonds, resulting in the formation of compounds that are not biologically active [8]. Most of the therapeutic proteins are not administered via the oral route due to gastrointestinal enzymatic degradation. The intravenous and the subcutaneous are the two most frequently used routes of administration; however, for high molecular weight proteins,

even the extravascular route of administration may not be possible [1]. Following subcutaneous administration, a protein drug may enter into systemic circulation via the blood capillaries or lymphatic system. For a protein administered through the subcutaneous route, its molecular weight may play an important role in the absorption process [9]. For high molecular weight proteins the lymphatic system becomes the predominant pathway for absorption resulting in the enhancement of the time to reach maximum concentration compared to a protein drug which is absorbed via the blood capillaries [4,1]. Another important challenge is that the body may react with an immune response against the therapeutic protein. The administration of proteins obtained from non-human sources can result in an immunogenic response and the advent of human recombinant proteins does not completely overcome this problem [10].

The formation of antibodies in response to exogenous protein administration may have an impact on the pharmacokinetics of drugs, above all on protein biological activity, safety and clearance [1]. Clearance of protein drugs represents, by itself, a significant limitation for their administration. Several pathways are involved in the removal of therapeutic proteins from circulation or interstitial fluid such as degradation by proteolysis, Fcγ receptor-mediated clearance, target-mediated clearance, nonspecific endocytosis, formation of immune-complexes (ICs) followed by complement- or Fc receptor-mediated clearance mechanisms and direct glomerular crossing in the kidney without previous modifications [11]. Renal filtration of endogenous compounds is a complex process that involves passage in the Bowman capsule, tubular re-absorption and secretion, in which solutes are exchanged owing to their size and physicochemical properties. The first process is a filtration step within the glomerule where unbound plasma components with a molecular size less than 50–60 kDa are filtered into the primitive urine. This cut-off is a strong determinant in the clearance of biotherapeutics because small proteins such as interferons or growth factors are quickly filtered and have half-lives of only a few hours, while large proteins such as antibodies have an elimination rate 10- to 50-fold lower. A short half-life in blood, common for proteins having a molecular weight below the kidney excretion threshold, poses a significant limitation in

their use because they have to be administered at high doses to reach the desired effect, simultaneously exacerbating side effects. On the other hand, some of the mechanisms of protein elimination, degradation, and post-translational modifications affect also administered therapeutic proteins thus influencing the response/dose ratio and the efficacy [12,13]. Researchers have several tools for the development of better therapeutic proteins, such as genetic engineering, advanced protein formulation and delivery. The development of protein drug delivery systems (DDSs) has been and will be essential for the full exploitation of protein-based drugs. For example, genetic engineering is strictly at the base of some delivery approaches like the expression of fusion or hyper-glycosylated proteins. In addition other DDSs have been applied to proteins such as incorporation into micro/nanoparticles [14,15], liposomes [16] or hydrogels [17], and chemical modification with small molecules (e.g. acetylation [18]) or polymers[19]. In this context, the covalent attachment of the polymer polyethylene glycol (PEG) to bioactive substances has become the leading approach for overcoming most of the aforementioned limits. Since the first PEGylated product was approved by the Food and Drug Administration in 1990, PEGylation has been widely used as a post-production modification methodology for improving biomedical efficacy and physicochemical properties of therapeutic proteins [13,20].

PEGylation of therapeutic proteins

Protein conjugation with polyethylene glycol (PEG) is a largely employed technology that improves the pharmacokinetics behavior of protein therapeutics. In the 1970s the first attempts of protein PEGylation were carried out by Abuchowski et al., and in 1990 the first PEGylated biopharmaceutical appeared on the market: a PEGylated form of adenosine deaminase, Adagen® (Enzon Pharmaceuticals, USA), for the treatment of severe combined immunodeficiency disease (SCID). Since then, about a dozen of bioconjugate therapeutics currently approved for human administration consist of PEGylated proteins [20,21].

PEG has been approved by the US FDA as 'generally recognized as safe'. It is synthetized by ring opening polymerization of monomeric ethylene oxide. Water or ethanol can be used to initiate the polymerization reaction, leading respectively to bifunctional polymer chains or monofunctional PEG. The last one, mono- methoxylated PEG (mPEG), is often preferred for protein conjugation either in its linear or branched form, and can be further functionalized through its free hydroxyl end group. There are various chemical groups on the amino acid side chains that could in principle be exploited for the reaction with the PEG active moiety, such as -NH2, -NH-, COOH, -OH, -SH groups as well as disulfide (-S-S-) bonds. As lysine is one of the most abundant amino acids in protein sequences, many PEGs were designed for reacting with amino groups. Nowadays several other routes are available on the market [20].

PEG is a rather unique flexible polymer, soluble in both organic and aqueous solvents. It is able to coordinate about 6-7 water molecules per monomer unit and for this reason its hydrodynamic volume is higher than that of a globular protein of the same molecular weight [22]. Interestingly, the PEG/protein hydrodynamic volume ratio increases in correlation with the polymer molecular weight (e.g. 10 and 40 kDa PEGs have hydrodynamic volumes corresponding to that of proteins of 65.4 and 670.7 kDa, respectively) [23].

Conjugation of PEG to proteins results in a new macromolecule with significantly changed physicochemical features. PEGylation, by increasing the molecular weight of a molecule, can impart several significant pharmacological advantages over the unmodified form, such as:

- Improved drug solubility
- Reduced dosage frequency (with potentially reduced toxicity)
- Extended circulating lifetime
- Increased drug stability
- Enhanced protection from proteolytic degradation
- Reduced immunogenic response

PEGylation is typically an additional step, implemented at the end of an already existing process for the production of a given protein. Therefore, the additional manufacturing costs introduced by PEGylation, caused by supplementary reactions and purification steps, should be minimized. The reactor for PEGylation in the pharmaceutical industry is typically a batch reactor. Small scale batch reactors are likely to be more economical, and easy to develop, control and maintain, without requiring too many auxiliary equipment [24,20].

PEGylation strategies overview

Looking back into the history of PEGylation, it can been seen that, until recently, a "random" conjugation approach has been carried out targeting the amino groups on the protein, most frequently the ε-amino groups on the side chains of lysine residues [20].

In fact, lysines, widely expressed on protein surfaces, are polar amino acid residues prone to chemical reactions with properly derivatized PEG reagents such as PEG-chlorotriazine [25], succinimidyl succinate (SS-PEG) [26] and succinimidyl carbonate PEG reagents (SC-PEGs) [27]. Although other chemistries have also been tried in the past, most of the random PEGylation reagents possess the activated carbonyl group in the form of *N*-hydroxy-succinimide esters that form stable protein-PEG conjugates via amide linkages. However, due to reactions with different nucleophilic groups on the protein, even mono-PEGylation leads to positional isomers that can differ substantially in their biological and biomedical properties. A clear consequence of such approach is the resulting complex mixture of conjugates, with different degrees of PEGylation due also to the abundance of lysine residues and to the reactions with other protein nucleophiles (i.e. *N*-terminal amino groups, the imidazolylnitrogen of histidine residues and even the side chains of serine, threonine, tyrosine and cysteine residues). The first PEGylated pharmaceuticals, Adagen® (pegademase) and Oncaspar® (PEGylated asparaginase, pegaspargase), were actually complex mixtures of various PEGylated species [28,29]. Also subsequently approved drugs, such as PegIntron® and Pegasys®, are produced by random PEGylation. Both the latter biopharmaceutical drugs are mixtures of mono-PEGylated positional isomers:

PegIntron® contains linear 12 kDa PEG chains bound to different sites of IFN-α2b and Pegasys® includes branched 40 kDa PEG chains bound mainly to four Lys residues of IFN-α2a [24].

PEGylation technology, aimed at the development of better defined products, shifted from random to site-specific PEGylation reactions. Two milestones of this kind of approach are N-terminal and cysteine-specific PEG conjugation.

For N-terminal PEGylation, PEG aldehydes were developed that under acidic conditions (pH 4-5) are more selective for N-terminal α-amine; a reducing agent was used (i.e. sodium cyanoborohydride) [30] to obtain a stable secondary amine from a Schiff base intermediate [20]. An example is Neulasta®, which is an N-terminally mono-PEGylated G-CSF bearing a 20 kDa PEG. The improved pharmacokinetic behavior enables administration only once per chemotherapy cycle, to be compared to the first generation molecule, Neupogen®, which was administered daily [31].

Cysteines, usually forming a disulfide bridge, when unpaired on protein surface or genetically introduced exhibit good selectivity. A variety of thiol-specific reagents are available, such as maleimide, pyridyl disulfide, vinyl sulfone, thiol reagents. Genetically introduced cysteines represent an opportunity to direct the PEG moiety to a pre-defined site in the protein. A very promising group of proteins for cysteine-specific PEGylation are Fab' fragments. PEGylation appears to be an ideal method to reduce their antigenicity and prolong their *in vivo* circulation times [32].

To overcome free unpaired cysteine scarcity in the amino acid sequence, strategies based on disulfide bridge opening have been adopted. TheraPEG™ is a PEGylation technology from PolyTherics that exploits special PEG monosulfone reagents. The site-specific bis-alkylation of both sulfur atoms in the natural disulfide bond that is sufficiently exposed on the protein surface results into the insertion of the PEG linker into the disulfide bond and formation of a three-carbon PEGylated bridge [33].

Taking into account that histidine affinity tags are among the most frequently used tools for easy and rapid purification of recombinant proteins, a strategy targeting histidine tag for PEG attachment has been developed by PolyTherics [34].

Another alternative strategy, aimed at the development of homogeneous products, is PEGylation targeting non-natural amino acids, requiring the genetic manipulation of the protein and of the host organism to allow incorporation of non-natural amino acids (the so-called Amber technology), which can be specifically conjugated with appropriate PEG reagents [35].

Finally, instead of traditional conjugation reactions through chemical procedures, enzymes can also be employed to achieve specific PEGylation. For instance, transglutaminase has been used to catalyze the addition of PEG amine to the acyl group of glutamine residues. The reaction offers a high degree of specificity since only those glutamine residues that are encompassed in a flexible or unfolded region are modified [20].

Peptides therapeutics and diagnostics

More than 7000 naturally occurring peptides have been recognized for having a wide variety of roles in human physiology [36,37]. In general, peptides are efficacious signaling molecules that bind to specific cell surface receptors or ions channels, with high selectivity, triggering intracellular effects. Peptides are in the sweet spot between small molecules and biopharmaceuticals in terms of production complexity and costs.

For these reasons peptides represent an excellent starting point for the design of novel therapeutics, not only for their selectivity and efficacy but also for their safety and tolerability profile.

At present, there are more than 60 US FDA-approved peptide medicines on the market and therapeutic peptide research is currently growing. In fact approximately 140 peptide drugs are in clinical trials and more than 500 therapeutic peptides are in preclinical development [38]. The main disease areas currently driving the therapeutic use of peptide drugs are metabolic diseases and oncology [39].

Naturally occurring peptides are often not directly suitable for use as therapeutics agents because they have intrinsic weakness, including poor chemical and physical stability, and a short circulating plasma half-life. Consequently traditional rational design of peptides

has focused on techniques to mitigate these weaknesses and/or to enhance the peptides affinity for their biological target. Rational design can start with a known crystal structure of the peptide giving the secondary and tertiary structure. Then, based on the input from various analyses, such as alanine substitutions (Ala-scan), and small focused libraries, the structure-activity relation (SAR) is built in sequential steps that lead to the identification of essential amino acids and also sites for possible substitution [40,41].

At present small-medium size peptides can be chemically synthesized using the combination of solid phase peptide synthesis with microwave irradiation (MW-SPPS) with higher yield and purity with respect to biotechnological techniques.

Thanks to their specific mode of action, peptides are often excellent biomarkers and therefore can also be used in diagnostic assays [42]. For example peptide microarrays have become increasingly accessible in recent years; beyond their initial utility in substrate profiling to sense protein activity, the peptides on the arrays can act as molecule ligands in detection systems. A collection of peptides displayed on a solid surface (chip) can be directly incubated with a variety of biological samples.

In the development of a detection system, peptides can be used to recognize a given target in a biological sample. Moreover, peptides can be synthetically produced and modified by conjugation with an extrinsic chromophoric probe, giving an easily detectable signal and acquiring a vital role in the development of biosensors.

References

[1] I. Mahmood and M. D. Green, “Pharmacokinetic and Pharmacodynamic Considerations in the Development of Therapeutic Proteins,” *Clin. Pharmacokinet.*, vol. 44, no. 4, pp. 331–347, 2005.

[2] E. Pennisi, “Gene Counters Struggle to Get the Right Answer,” *Science*, vol. 301, no. 5636, pp. 1040-1041, Aug. 2003.

[3] J. C. Venter et al., “The Sequence of the Human Genome,” *Science*, vol. 291, no. 5507, pp. 1304–1351, Feb. 2001.

[4] B. Leader, Q. J. Baca, and D. E. Golan, “Protein therapeutics: a summary and pharmacological classification,” *Nat. Rev. Drug Discov.*, vol. 7, no. 1, pp. 21–39, Jan. 2008.

[5] F. G. Banting, C. H. Best, J. B. Collip, W. R. Campbell, and A. A. Fletcher, “Pancreatic Extracts in the Treatment of Diabetes Mellitus.,” *Can. Med. Assoc. J.*, vol. 12, no. 3, pp. 141–6, Mar. 1922.

[6] D. V Goeddel *et al.*, “Expression in Escherichia coli of chemically synthesized genes for human insulin.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 76, no. 1, pp. 106–10, Jan. 1979.

[7] S. D. Putney and P. A. Burke, “Improving protein therapeutics with sustained-release formulations,” *Nat. Biotechnol.*, vol. 16, no. 2, pp. 153–157, Feb. 1998.

[8] P. Arora, R. Gandhi, and T. Kelly, “Overview of Drug Product Development for Protein Therapeutics: Present Strategies and Future Perspectives,” *American Pharmaceutical Review*, vol. 14, no. 7, pp. 1-5, Dec. 2011.

[9] S. Toon, “The relevance of pharmacokinetics in the development of biotechnology products.,” *Eur. J. Drug Metab. Pharmacokinet.*, vol. 21, no. 2, pp. 93–103.

[10] S. E. Fineberg, J. A. Galloway, N. S. Fineberg, M. J. Rathbun, and S. Hufferd, “Immunogenicity of recombinant DNA human insulin,” *Diabetologia*, vol. 25, no. 6, pp. 465–469, Dec. 1983.

[11] Y. Vugmeyster, X. Xu, F.-P. Theil, L. A. Khawli, and M. W. Leach, “Pharmacokinetics

and toxicology of therapeutic proteins: Advances and challenges.," *World J. Biol. Chem.*, vol. 3, no. 4, pp. 73–92, Apr. 2012.

[12] K. Petrak and P. Goddard, "Transport of macromolecules across the capillary walls," *Adv. Drug Deliv. Rev.*, vol. 3, no. 2, pp. 191–214, 1989.

[13] G. Pasut and F. M. Veronese, "State of the art in PEGylation: The great versatility achieved after forty years of research," *J. Control. Release*, vol. 161, no. 2, pp. 461–472, Jul. 2012.

[14] C. Pinto Reis, R. J. Neufeld, A. J. Ribeiro, and F. Veiga, "Nanoencapsulation II. Biomedical applications and current status of peptide and protein nanoparticulate delivery systems," *Nanomedicine Nanotechnology, Biol. Med.*, vol. 2, no. 2, pp. 53–65, Jun. 2006.

[15] A. Almeida and E. Souto, "Solid lipid nanoparticles as a drug delivery system for peptides and proteins" *Adv. Drug Deliv. Rev.*, vol. 59, no. 6, pp. 478–490, Jul. 2007.

[16] S. Martins, B. Sarmento, D. C. Ferreira, and E. B. Souto, "Lipid-based colloidal carriers for peptide and protein delivery--liposomes versus lipid nanoparticles.," *Int. J. Nanomedicine*, vol. 2, no. 4, pp. 595–607, 2007.

[17] J. K. Tessmar and A. M. Göpferich, "Matrices and scaffolds for protein delivery in tissue engineering," *Adv. Drug Deliv. Rev.*, vol. 59, no. 4, pp. 274–291, 2007.

[18] M. C. Jones and M. Patel, "Insulin detemir: A long-acting insulin product," *Am. J. Heal. Pharm.*, vol. 63, no. 24, pp. 2466–2472, Dec. 2006.

[19] F. M. Veronese and G. Pasut, "PEGylation: Posttranslational bioengineering of protein biotherapeutics," *Drug Discov. Today Technol.*, vol. 5, no. 2, pp. e57–e64, 2008.

[20] D. Pfister and M. Morbidelli, "Process for protein PEGylation," *J. Control. Release*, vol. 180, pp. 134–149, 2014.

[21] S. Kontos and J. A. Hubbell, "Drug development: longer-lived proteins," *Chem. Soc. Rev.*, vol. 41, no. 7, p. 2686, 2012.

[22] J. M. Harris, "Introduction to Biotechnical and Biomedical Applications of Poly(Ethylene Glycol)," in *Poly(Ethylene Glycol) Chemistry*, Boston, MA: Springer

US, 1992, pp. 1–14.

[23] C. J. Fee and J. M. Van Alstine, “Prediction of the Viscosity Radius and the Size Exclusion Chromatography Behavior of PEGylated Proteins,” *Bioconjug. Chem.*, vol. 15, no. 6, pp. 1304–1313, Nov. 2004.

[24] S. JevsÌŒevar, M. Kunstelj, and V. G. Porekar, “PEGylation of therapeutic proteins,” *Biotechnol. J.*, vol. 5, no. 1, pp. 113–128, Jan. 2010.

[25] A. Abuchowski, T. van Es, N. C. Palczuk, and F. F. Davis, “Alteration of immunological properties of bovine serum albumin by covalent attachment of polyethylene glycol.,” *J. Biol. Chem.*, vol. 252, no. 11, pp. 3578–81, Jun. 1977.

[26] A. Abuchowski *et al.*, “Cancer therapy with chemically modified enzymes. I. Antitumor properties of polyethylene glycol-asparaginase conjugates.,” *Cancer Biochem. Biophys.*, vol. 7, no. 2, pp. 175–86, Jun. 1984.

[27] S. Zalipsky, R. Seltzer, and S. Menon-Rudolph, “Evaluation of a new reagent for covalent attachment of polyethylene glycol to proteins.,” *Biotechnol. Appl. Biochem.*, vol. 15, no. 1, pp. 100–14, Feb. 1992.

[28] C. Booth and H. B. Gaspar, “Pegademase bovine (PEG-ADA) for the treatment of infants and children with severe combined immunodeficiency (SCID).,” *Biologics*, vol. 3, pp. 349–58, 2009.

[29] C. H. Fu and K. M. Sakamoto, “PEG-asparaginase,” *Expert Opin. Pharmacother.*, vol. 8, no. 12, pp. 1977–1984, Aug. 2007.

[30] O. Kinstler, G. Molineux, M. Treuheit, D. Ladd, and C. Gegg, “Mono-N-terminal poly(ethylene glycol)–protein conjugates,” *Adv. Drug Deliv. Rev.*, vol. 54, no. 4, pp. 477–485, 2002.

[31] G. Molineux, “The design and development of pegfilgrastim (PEG-rmetHuG-CSF, Neulasta).,” *Curr. Pharm. Des.*, vol. 10, no. 11, pp. 1235–44, 2004.

[32] D. P. Humphreys *et al.*, “Alternative antibody Fab’ fragment PEGylation strategies: combination of strong reducing agents, disruption of the interchain disulphide bond and disulphide engineering,” *Protein Eng. Des. Sel.*, vol. 20, no. 5, pp. 227–234, Apr. 2007.

[33] S. Balan *et al.*, “Site-Specific PEGylation of Protein Disulfide Bonds Using a Three-Carbon Bridge,” *Bioconjug. Chem.*, vol. 18, no. 1, pp. 61–76, Jan. 2007.

[34] Y. Cong *et al.*, “Site-Specific PEGylation at Histidine Tags,” *Bioconjug. Chem.*, vol. 23, no. 2, pp. 248–263, Feb. 2012.

[35] A. Deiters, T. A. Cropp, D. Summerer, M. Mukherji, and P. G. Schultz, “Site-specific PEGylation of proteins containing unnatural amino acids,” *Bioorg. Med. Chem. Lett.*, vol. 14, no. 23, pp. 5743-5745, Dec. 2004.

[36] A. Padhi, M. Sengupta, S. Sengupta, K. H. Roehm, and A. Sonawane, “Antimicrobial peptides and proteins in mycobacterial therapy: Current status and future prospects,” *Tuberculosis*, vol. 94, no. 4, pp. 363–373, Jul. 2014.

[37] S. D. Robinson, H. Safavi-Hemami, L. D. McIntosh, A. W. Purcell, R. S. Norton, and A. T. Papenfuss, “Diversity of Conotoxin Gene Superfamilies in the Venomous Snail, Conus victoriae,” *PLoS One*, vol. 9, no. 2, p. e87648, Feb. 2014.

[38] A. A. Kaspar and J. M. Reichert, “Future directions for peptide therapeutics development,” *Drug Discov. Today*, vol. 18, no. 17, pp. 807–817, 2013.

[39] K. Fosgerau and T. Hoffmann, “Peptide therapeutics: current status and future directions,” *Drug Discov. Today*, vol. 20, no. 1, pp. 122–128, 2015.

[40] M. C. Manning, D. K. Chou, B. M. Murphy, R. W. Payne, and D. S. Katayama, “Stability of Protein Pharmaceuticals: An Update,” *Pharm. Res.*, vol. 27, no. 4, pp. 544–575, Apr. 2010.

[41] I. W. Hamley, “Peptide Fibrillization,” *Angew. Chemie Int. Ed.*, vol. 46, no. 43, pp. 8128–8147, Nov. 2007.

[42] I. Schulte, H. Tammen, H. Selle, and P. Schulz-Knappe, “Peptides in body fluids and tissues as markers of disease.,” *Expert Rev. Mol. Diagn.*, vol. 5, no. 2, pp. 145–57, Mar. 2005.

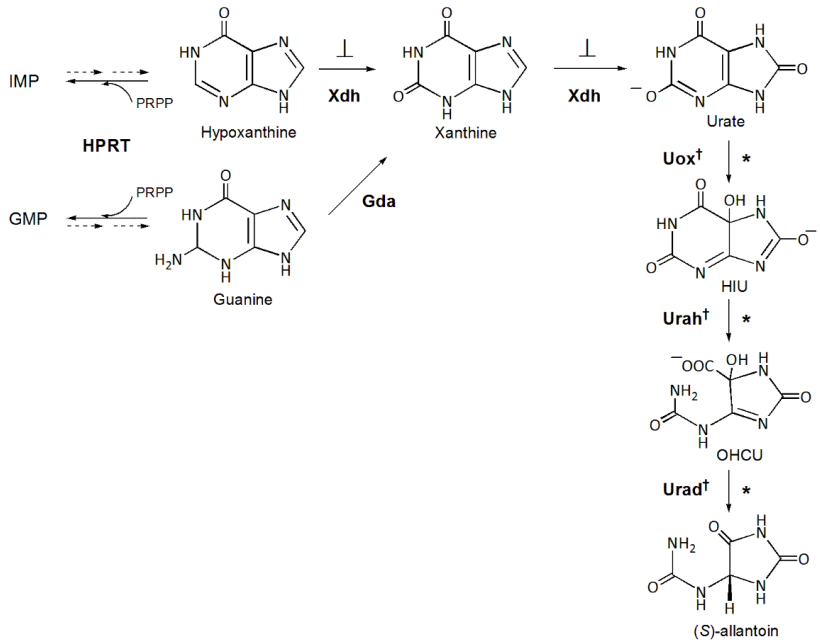
Chapter 2

Rebuilding the urate degradation pathway in humans: a triad of PEGylated enzymes for uricolysis

Introduction

Hypoxanthine-Guanine phosphoribosyltransferase deficiency

Purine nucleotides can be synthesized in two distinct pathways. First, purines are biologically synthesized de novo in particular as ribotides (bases attached to ribose 5-phosphate), starting from simple materials such as amino acids and bicarbonate. A key regulatory step is the production of 5-phospho-α-D-ribosyl 1-pyrophosphate (PRPP) by ribose phosphate pyrophosphokinase. Alternatively, purine bases, released by the hydrolytic degradation of nucleic acids and nucleotides or from diet, can be salvaged and recycled. The so called purine salvage pathways are known for the energy that they save and the remarkable effects of their attenuation and/or absence [1]. Hypoxanthine-guanine phosphoribosyltransferase (HPRT) is a key enzyme involved in these recycle pathways, catalyzing the conversion of hypoxantine to inosine monophosphate (IMP) and guanine to guanosine monophosphate (GMP) via PRPP-dependent transfer reactions (Scheme 1).



Scheme 1 Urate biosynthesis and degradation. *Enzymes names are in bold, and abbreviated as follows: HPRT, hypoxanthine-guanine phosphoribosyltransferase; Xdh, xanthine oxidase/dehydrogenase; Gda, guanine deaminase; Uox, urate oxidase; Urah, 5-hydroxyisourate (HIU) hydrolase; Urad, 2-oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline (OHCU) dehydrogenase. Genes that are lost in hominoids are denoted by a dagger symbol. Reactions inhibited by allopurinol and febuxostat are indicated by inhibition symbols. The proposed therapeutic approach is based on the facilitation of purine degradation through the administration of uricolytic enzymes (asterisks), and is anticipated to normalize the level of upstream metabolites.*

The lack or deficiency of HPRT catalytic activity leads to the augmented degradation of free purine bases, hypoxantine and guanine, that cannot be recycled, and consequently to the overproduction of the end product of purine degradation in humans, uric acid [2]. This slightly soluble metabolite accumulates in plasma (hyperuricaemia) and urine (hyperuricosuria) and precipitates in joints and kidney causing juvenile gout and renal failure. Among disorders of purine metabolism, the Lesch–Nyhan disease (LND) is also characterized by neurological manifestations. The LND is a rare X-linked disease characterized by severe deficiency of HPRT enzyme caused by mutations of the human gene HPRT1 diminishing or abrogating the catalytic activity of the encoded enzyme. LND

patients are affected by motor disabilities such as severe action dystonia, involuntary movement, by cognitive impairments such as moderate mental retardation and by unusual self-injurious behavior as a result of obsessive-compulsive behavior [3]. While overproduction of uric acid is present in all HPRT-deficient patients, the neurological manifestations of the disease show a continuum spectrum of gravity depending on the degree of the enzymatic deficiency. Different classifications of HPRT-deficiency exist according to the severity of the neurological symptoms. At the higher end of the spectrum is the classic LND with the full clinical phenotype. At the lower end of the spectrum is HPRT-gout with no apparent neurological deficits. Inbetween are Lesch-Nyhan variants (LNV) with different degrees of abnormalities [4].

Traditional uricostatic therapeutic approach

The most common current therapeutic approaches to uric acid-related symptoms of HPRT deficiency rely on uricostatic drugs (i.e. allopurinol, febuxostat). These drugs exploit their action blocking the purine degradation and reducing serum urate and urine uric acid levels. Allopurinol (Figure 1), an analogue of hypoxanthine, acts as a competitive inhibitor of xanthine oxidase, the enzyme responsible of the hypoxanthine conversion in xanthine and finally in uric acid (Scheme 1).

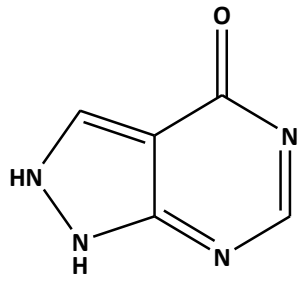


Figure 1 *Allopurinol chemical structure*

Control of uric acid overproduction with allopurinol reduces the risk for nephrolithiasis, urate nephropathy, gouty arthritis, and tophi, but without effects on behavioral and

neurologic symptoms [5]. The uricostatic approach, at the same time, causes the accumulation of urate precursors xanthine and hypoxanthine resulting in alteration of their physiological concentrations. The increased concentration of hypoxanthine and xanthine represents the most evident biochemical dysfunction in the central nervous system. A toxic effect of hypoxanthine has been implicated in the severe neurological manifestation of LND probably through inhibition of the adenosine transport [6]; an increase in hypoxanthine and xanthine excretion rates of about 5-10 fold is observed in allopurinol-treated patients [5,6]. So, paradoxically, a therapeutic treatment contrasting uric acid production through a block of the degradation pathway is expected to aggravate rather than normalize the purine imbalance caused by HPRT deficiency.

Uric acid degradation pathway

At variance with humans (and hominoids in general), most mammals are able to produce enzymes catalyzing the conversion of uric acid in its much more soluble metabolite (S)-allantoin, easily excreted with urine. For this reason they are not prone to gout or urate kidney stones formation. It has long been thought that urate oxidase (Uox) is the unique enzyme involved in this conversion [7,8]. Nevertheless bioinformatic and biochemical studies have revealed a multi-enzymatic pathway (Scheme 1): Uox is responsible for the oxidation of uric acid to (S)-5-hydroxyisourate (HIU); HIU is then converted into (S)-2-oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline (OHCU) by HIU hydrolase (Urah), and the last step gives (S)-allantoin from OHCU decarboxylation thanks to OHCU decarboxylase (Urad) catalysis [9].

The metabolite HIU, *in vitro*, is able to spontaneously degrade to racemic allantoin on a time scale of several hours, while the reaction catalyzed by the enzymatic triad directly leads to (S)-allantoin production in seconds, confirming the presence of the enzymatic pathway in cells [10]. Starting from these knowledge an uricolytic treatment has the possibility to reduce uric acid levels and also to normalize the concentrations of upstream metabolites by increasing the metabolic flux towards degradation. The first enzyme of the pathway, Uox, has been approved for human use in US and Europe in two different

formulations: Rasburicase® (Sanofi-Aventis, for prevention and treatment of tumor lysis syndrome) [11] which is not in an appropriate form for a prolonged use (as in the case of LND) and Pegloticase® (Savient Pharmaceuticals, for treatment-refractory, chronic gout), with augmented half-life [12]. However, the treatment with the full enzymatic pathway would be preferable because bioinformatic data suggest that the two enzymes following Uox in the degradation pathway have a detoxification role [9]; it is also known that in the animal model, the presence of urate oxidase in the absence of the downstream enzyme (Urah) has a potent oncogenic effect [13].

Research purpose

The current therapeutic approach to uric acid overproduction related to HPRT-deficiency is characterized by the inhibition of xanthine oxidase using uricostatic drugs (i.e. allopurinol). For those cases irresponsive or intolerant to uricostatic therapies [14], Uox can be administered. When a chronic treatment is required (as in the case of LND), the therapy with only Uox shows important side effects. It is known that in animal models, the product of Uox catalysis, HIU, can have toxic and oncogenic effects [13]. This knowledge together with bioinformatic data suggests that the two enzymes following Uox in the degradation pathway, Urah and Urad, have a detoxification role [9].

The enzymatic triad (Uox, Urah and Urad), present in other mammals but not in humans, catalyzes the conversion of uric acid to its more soluble metabolite (S)-allantoin. The treatment with the full enzymatic set would be preferable because it controls the upstream oxypurines concentration by facilitating the metabolic flux towards degradation. For this reason the uricolytic treatment could be effective on behavioral and neurologic symptoms in HPRT-deficient patients. Anyway the proposed enzymatic treatment would also represent an alternative treatment for those cases irresponsive or intolerant to uricostatic drugs without purine imbalance-derived side effects.

This research line has the purpose to contribute to the development of a more sound treatment for HPRT-deficiency and LND, facilitating rather than blocking purine degradation. Our strategy is to obtain Uox, Urah and Urad in recombinant form, and to increase their half-life and decrease their immunogenicity while preserving enzymatic activity and stability, through polyethylene glycol (PEG) modification of enzymes surface.

Materials and Methods

Enzymes PEGylation

All PEG preparations were of the best available purity and polydispersivity. For protein modification PEG functionalized with maleimido-group (MAL-PEG) of three different molecular masses (5, 20 and 40 kDa) was used. PEG with nominal molecular weight of 5 kDa was purchased from NOF Corporation (Figure 2).

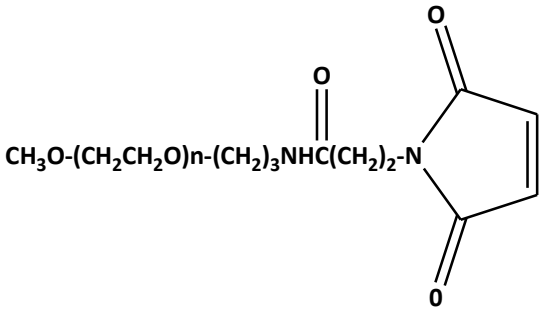


Figure 2 *MAL-PEG (Sunbright, NOF) chemical structure*

20 kDa and 40 kDa MAL-PEG were obtained from Iris Biotech GMBH (Figure 3).

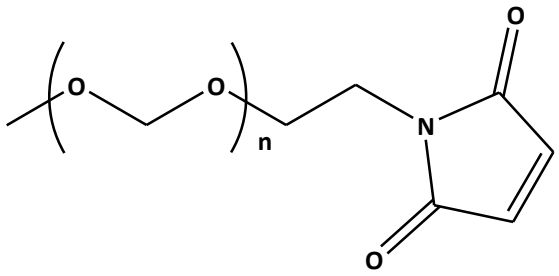


Figure 3 *MAL-PEG(Iris Biotech) chemical structure*

Different conjugation schemes were followed by varying PEG molecular weight, the stoichiometry of the reactants, the molar ratio between protein and PEG and the time of reaction.

DrUox and DrUrah PEGylation

Uox and Urah enzymes from Zebrafish (*Danio rerio*) were expressed in *E. coli* and purified at the Biochemistry and Molecular Biology laboratory at the University of Parma. DrUox was purified by affinity chromatography onto an agarose resin functionalized with xanthine, an analog of the enzyme substrate that cannot be oxidized by Uox. The protein was then eluted by adding an uric acid solution. DrUrah was purified by subsequent cationic and anionic exchange chromatography at different pH values. The two enzymes solutions were finally diafiltered and concentrated before storage.

For DrUox and DrUrah the same conjugation scheme was followed. The reactions were carried out in vials adopting a direct PEG attachment based on the number of free cysteines on protein surface.

2 mg/ml enzyme stock solution was reacted with fresh reagents solutions in 100 mM potassium phosphate, 150 mM NaCl, pH 7.4 at 20 °C. As a control, a sample not treated with PEG (added with the same volume of buffer) was used for each enzyme. The conjugation was stopped and monitored over the time by adding a molar excess (6-fold with respect to the protein concentration) of free cysteine. The PEGylated samples were then dialyzed at 4 °C in 100 mM potassium phosphate, 150 mM NaCl, pH 7.4 in order to eliminate unreacted MAL-PEG and other reagents. Dialysis was carried out using Millipore Amicon Ultra ultrafiltration tubes supplied by Merck Millipore; a membrane with a cut-off of 10 kDa was used for the unmodified protein, 30 KDa for the 5 and 20 kDa PEGylated proteins and a cut-off of 100 kDa for the 40 KDa PEGylated enzymes.

DrUrad PEGylation

Urad enzyme from Zebrafish was expressed in *E. coli* and purified at the Biochemistry and Molecular Biology laboratory at the University of Parma. DrUrad was purified by affinity chromatography on cobalt resin (Talon, Clontech), thanks to its N-terminal poly-histidine tag. The protein was eluted using a buffer containing imidazole. The enzyme solutions were finally diafiltered and concentrated before storage. In this case, for protein modification, a two-step procedure was followed. The first step involved the modification

of the amino groups of the enzyme by reaction with 2-imino thiolane (IMT) (Figure 4), adding a spacer arm ending with a thiol group which is able to react with the maleimido ring of MAL-PEG (second step) [15].

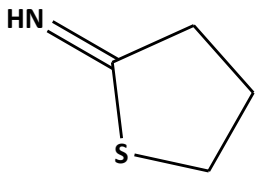


Figure 4 *2-iminothiolane (IMT) chemical structure*

The reactions were quenched by the addition of an excess of lysine (10-fold with respect to the IMT concentration) and cysteine (6-fold with respect to the PEG concentration). Finally the PEGylated samples were dialyzed at 4 °C in 100 mM potassium phosphate, 150 mM NaCl, pH 7.4 in order to eliminate unreacted PEG and IMT using Micro Float-A-Lyzer® micro dialysis devices (cut-off membrane 20 kDa, 50 kDa and 100 kDa) purchased from Spectrum Labs.

Lyophilization

The PEGylated enzymes and the control samples (unmodified proteins) were split in small aliquots and added with 4% mannitol and 1% sucrose (w/v, final concentration) as lyoprotectant agents [16]. All the aliquots were then flash-frozen in liquid nitrogen and then lyophilized O/N using a Modulyo® Freeze Dryer apparatus (Edwards). Lyophilized samples were then stored at -20 °C.

SDS PAGE of PEGylated enzymes

In order to evaluate the degree of protein derivatization, denaturing polyacrylamide gel electrophoresis was used. Protein samples were added with acetone (1 protein volume : 5 acetone volume) and left overnight at - 20 °C. After centrifugation (20 minutes at 13000 rpm) the pellet was dried using a SpeedVac concentrator (Christ RVC 2-18) and

resuspended in sample buffer. For each sample 5 to 10 µg were usually loaded on gels. Staining was carried out with Biosafe® Coomassie (destaining with water) for the detection of protein bands. The SDS PAGE gels were scanned using a ChemiDoc® imager (BIO-RAD) and bands intensity was evaluated by densitometric analysis through Image Lab software. In addition, the subunits to which PEG molecules are attached were revealed using a PEG specific dye (barium iodide) solution. 50 mL of 5% $BaCl_2$ solution and 2.5 ml of 0.05 M iodine solution were added to the previously Coomassie-stained gel for 5-10 minutes; after a quick rinse with deionised water the gel was electronically scanned.

Activity assays

The activity assays for the three enzymes were performed at 37 °C in 100 mM potassium phosphate, pH 7.4. The enzymatic activity of DrUox was analyzed by directly monitoring at 292 nm the uric acid consumption in the presence of 100 µM urate and an excess of HIU hydrolase. The secondary enzyme is required in order to eliminate HIU, one of the Uox reaction products, that absorbs in the same UV region of urate and affects the uricase activity with a product inhibition kinetics.

The substrates of Urah and Urad, HIU and OHCU, respectively, are not commercially available because of their instability. It was therefore necessary to produce them during the assay, by adding the enzymes responsible of their synthesis. The activity of DrUrah and DrUrad was monitored by measuring absorbance at 312 nm and 257 nm, respectively, in the presence of an excess of DrUox (for DrUrah), and DrUrah (for DrUrad) and 100 µM urate.

Circular dichroism spectroscopy

Circular dichroism spectra were carried out with a Jasco J715 spectropolarimeter equipped with a Peltier thermostatic cell set at 20 °C. Spectra were collected using a 1 mm quartz cuvette, in 20 mM potassium phosphate, pH 7.4.

Results and discussion

The purpose of this project was to obtain all the three enzymes involved in the uricolytic degradation pathway of purine catabolism (DrUox, DrUrah and DrUrad), in a form suitable for therapy after recombinant expression and covalent conjugation with PEG. In view of a future scaling-up of the PEGylation procedure for testing in animal models and perspective therapy, a common PEGylation chemistry for DrUox, DrUrad and DrUrah was adopted. PEG derivatized with maleimido-group (MAL-PEG) was used for the reaction with the thiol functional groups exposed at the protein surface. All the three sequences of recombinant enzymes used in this study are from zebrafish. This source has been selected because of the high solubility and stability of the isolated proteins and of the availability of three-dimensional structures [17,18]. For Uox the three-dimensional structure (PDB ID: 5LL1) has been recently solved. Anyway structure knowledge was helpful for the rational design of the chemical modification protocol.

DrUox, DrUrah and DrUrad sequences contain 5, 2 and 5 reduced cysteine residues per monomer, respectively (20, 8 and 10 cysteines per biological unit). Visual inspection of the DrUox structure indicates that under oxidizing conditions 4 Cys side chains are engaged in disulfide bridges and 16 are available for chemical conjugation.

In the table below (Table 1) the molecular weights, the quaternary structure and the content in free cysteine residues of each enzyme are summarized.

Table 1 *DrUox, DrUrah and DrUrad molecular weights, quaternary structures and number of free cysteine residues per biological unit.*

	Molecular Weight (kDa)	**Quaternary structure**	**Free Cysteines**
DrUox	136748	tetramer	16
DrUrah	52992	tetramer	8
DrUrad	44000	dimer	10

DrUox PEGylation

DrUox is a homotetrameric enzyme with 16 available cysteines for chemical conjugation. As a starting point, the reaction conditions were optimized investigating the time-evolution of the DrUox conjugation with 5 kDa MAL-PEG at two different PEG/free cysteines molar ratios (1.25 and 7.5). The reaction was carried out at 20 °C and monitored over the time by adding a molar excess of cysteine respectively after 15, 30 and 45 minutes from PEG addition. In Figure 5 the SDS-PAGE and its densitometric analysis are reported, indicating that the degree of conjugation strongly depends on PEG molar excess, as expected; in fact the highest PEG molar excess (7.5) already leads to significant protein derivatization after 15 minutes.

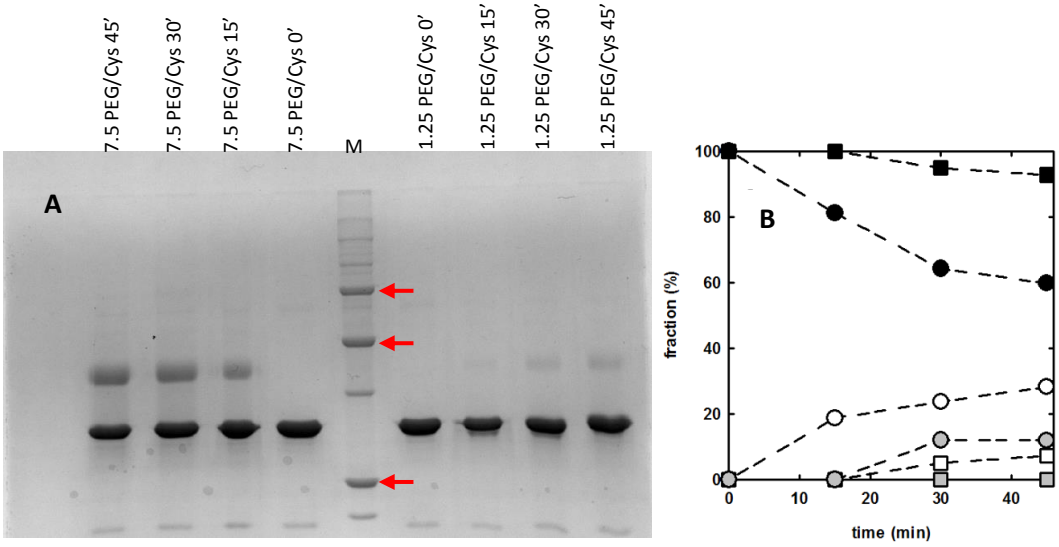


Figure 5 *Panel A: SDS-PAGE of DrUox PEGylation as a function of time at two different MAL-PEG/free cysteine ratios. Unstained Precision Plus Protein(R) standards (Biorad) were used as MW markers (M) (25, 50 and 75 kDa bands are highlighted in the gel). Panel B: PEGylation kinetics of DrUox at different PEG/cysteine molar ratios. The conjugation reaction was carried out at 20 °C, at a 5 kDa MAL-PEG/reactive cysteines molar ratio of 7.5 (circles) and 1.25 (squares). Grey symbols refer to the fraction of di-PEGylated monomers, open symbols refer to the fraction of mono-PEGylated monomers, and closed symbols to the fraction of unmodified monomers. Dashed lines through data points are just drawn for eye guidance.*

Considering the modest progression of the conjugation reaction at times longer than 30 minutes a 7.5 PEG/cysteine molar ratio and a 30 minutes reaction time were chosen in order to assure both a sufficient degree of PEGylation and to limit the incubation time for concerns relative to protein stability. Next, the effect of PEG chain length on the degree of covalent modification and the retention of enzymatic activity were investigated. To this end, DrUox was conjugated with MAL-PEG of three different molecular weights, 5, 20 and 40 kDa, using the same protocol. Figure 6 reports SDS-PAGE of DrUox reacted with the three different PEGs, in comparison with the unmodified protein.

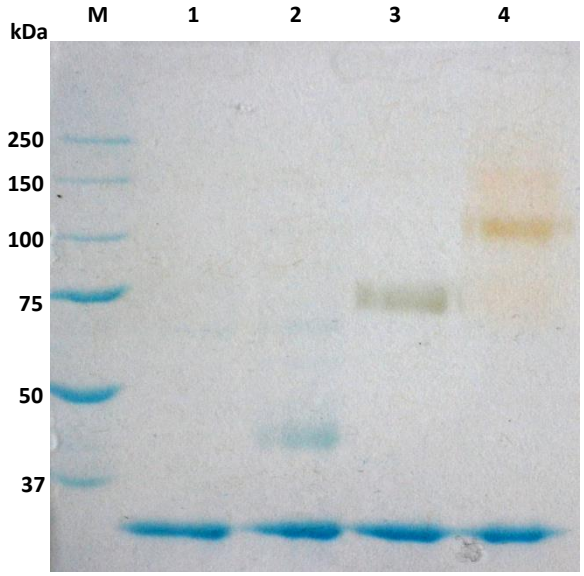


Figure 6 *PEGylation of DrUox. SDS-PAGE gel of unmodified and PEGylated DrUox. Lane M, MW standards; Lane 1, unmodified DrUox; Lane 2, DrUox + 5 kDa PEG; Lane 3, DrUox + 20 KDa PEG; Lane 4, DrUox + 40 kDa PEG. The gel underwent a double staining procedure, first with Coomassie blue (blue bands), and then with the PEG-specific dye barium iodide (yellow-brown bands).*

Yellow-brown bands in lanes 2, 3 and 4 are due to gel staining with a dye specific for PEG (barium iodide). Barium iodide staining was carried out after Coomassie blue staining to confirm the attribution of slowly migrating bands to PEGylated forms of DrUox monomers. Starting from the densitometric analysis of protein bands to evaluate the degree of PEGylation of single monomers, an estimate of the degree of tetramers

PEGylation and fraction of un-modified tetramers was obtained following the procedure previously reported for PEGylated hemoglobin [19]. DrUox PEGylation with MAL-PEG of different molecular weight resulted in similar PEGylation yields (Table 2).

Table 2 *Distribution of un-reacted and PEGylated DrUox monomers and tetramers, and enzymatic activity, after reaction with 5, 20 or 40 kDa MAL-PEG*

	% non-PEGylated monomers	% mono-PEGylated monomers	% di-PEGylated monomers	Average no. of PEG bound (per tetramer)	% Unmodified tetramers	% Enzymatic activity [a]
No PEG	100	0	0	0	100	87.7±19.2[a]
5 kDa PEG	60	28	12	2.1	13	87.9±3.5
20 kDa PEG	46	54	0	2.2	5	124.1±10.7
40 kDa PEG	71	21	8	1.5	25	141.6±2.0

[a] The activity is expressed as percentage with respect to freshly prepared DrUox (mean ± SE).

The enzymatic activity of all samples was checked right after the conjugation reaction. The results, shown in Table 2, indicate a good retention of the enzymatic activity after PEGylation, very close to the native enzyme for the 20 kDa or 40 kDa PEGylated protein. The enzymatic activity of the un-conjugated protein was measured on samples representing the control that underwent the same procedure of PEGylated ones, without the addition of MAL-PEG in the reaction mixture. The value significantly lower than 100 % is consistent with the observation that recombinant DrUox is marginally stable at room temperature.

DrUrah PEGylation

DrUrah is a tetrameric protein possessing 2 cysteines per monomer available for PEG attachment.

For DrUrah derivatization, we first tried to define the optimal PEGylation conditions measuring as a function of time the degree of PEGylation upon reaction with 5 kDa MAL-PEG, at a 1.25 or 7.5 PEG/free cysteines molar ratio, as seen for DrUox. In Figure 7 the SDS-PAGE and its densitometric analysis are reported.

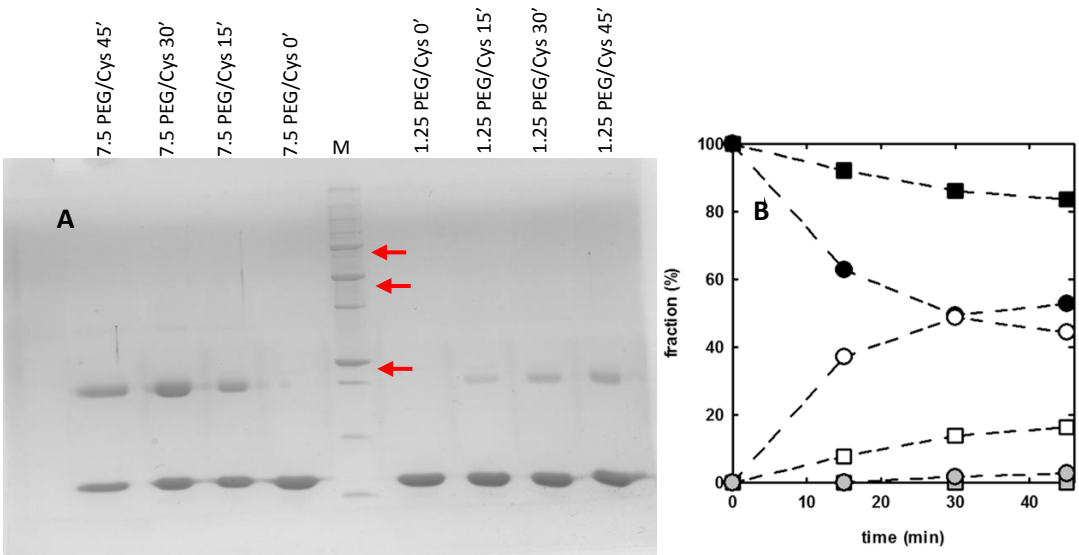


Figure 7 *Panel A: SDS PAGE of DrUrah PEGylation as a function of time at two different MAL-PEG/free cysteine ratios. Unstained Precision Plus Protein® standards (Biorad) were used as MW markers (25, 50 and 75 kDa bands are highlighted in the gel). Panel B: PEGylation kinetics of DrUrah at different PEG/cysteine molar ratios. The conjugation reaction was carried out at 20 °C, at a 5 kDa MAL-PEG/reactive cysteines molar ratio of 7.5 (circles) and 1.25 (squares). Grey symbols refer to the fraction of di-PEGylated monomers, open symbols refer to the fraction of mono-PEGylated monomers, and closed symbols to the fraction of unmodified monomers. Dashed lines through data points are just drawn for eye guidance.*

As already observed with DrUox, also for DrUrad a PEG/free cysteines molar ratio of 7.5 resulted in a good degree of derivatization, and the reaction apparently reached

completion within 30 minutes. Hence, in all following experiments DrUrah was PEGylated at 20 °C at a PEG/cysteine ratio of 7.5. The reaction with 5, 20 or 40 kDa MAL-PEG always yielded a good degree of derivatization, showing the presence of mono- and di-PEGylated hydrolase monomers as shown by the SDS-PAGE (Figure 7B). Densitometric analysis of the gel and probabilistic calculations allowed to determine the average number of PEG chains conjugated to DrUrah tetramers, as well as the probability to have unmodified tetramers, an undesired occurrence since the latter would be rapidly filtered through the kidneys (Table 3).

Table 3 *Distribution of un-reacted and PEGylated DrUrah monomers and tetramers, and enzymatic activity, after reaction with 5, 20 or 40 kDa MAL-PEG*

	% non-PEGylated monomers	% mono-PEGylated monomers	% di-PEGylated monomers	Average no. of PEG bound (per tetramer)	% Unmodified tetramers	% Enzymatic activity [a]
No PEG	100	0	0	0	100	85.6±28.5
5 kDa PEG	26	71	3	3.1	1	55.3±6.6
20 kDa PEG	26	61	13	3.5	1	183.3±43.9
40 kDa PEG	42	15	43	4.0	3	166.5±33.3

[a] The activity is expressed as percentage with respect to freshly prepared DrUrah (mean ± SE).

Compared to DrUox, the average number of conjugated PEG molecules per tetramer resulted to be higher, and the fraction of non-conjugated tetramers appeared to be almost negligible. Enzymatic assays showed that PEGylation, regardless of the molecular weight of the conjugated PEG, did not significantly impair enzymatic activity. Increased activity of PEGylated enzymes, though apparently counterintuitive, could be rationalized

taking into account quaternary stabilization of oligomeric proteins [20], selective stabilization of active conformations and, possibly, perturbations of the active site environment causing favorable ligand partitioning and binding.

DrUrad PEGylation

In view of a common PEGylation chemistry, DrUrad was firstly modified adopting the same conjugation protocol used for the two upstream enzymes. PEGylation with 5, 20 and 40 kDa MAL-PEG (7.5 PEG/free cysteines molar ratio, 30 min reaction time) occurred with good yields (Figure 8), but PEGylated DrUrad was almost totally inactive in all three cases.

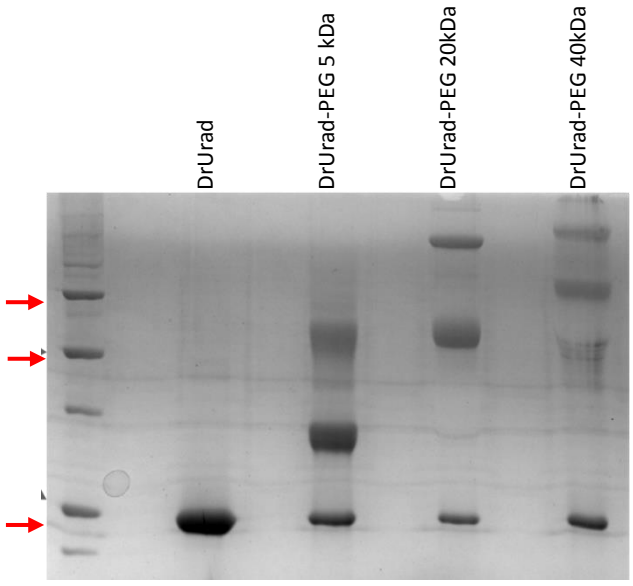


Figure 8 *SDS-PAGE of DrUrad before and after reaction with 5, 20 and 40 kDaMAL-PEG. The reaction was carried out at a PEG/cysteine ratio of 7.5, at 20 °C for 30 minutes. Unstained Precision Plus Protein(R) standards (Biorad) were used as MW markers (25, 50 and 75 kDa bands are highlighted in the gel).*

The analysis of DrUrad structure [17] indicates that the enzyme has five cysteine residues; interestingly, two are close to the dimer interface, and two, Cys-25 and Cys-66, are very close to the active site, not directly involved in catalysis (Figure 9).

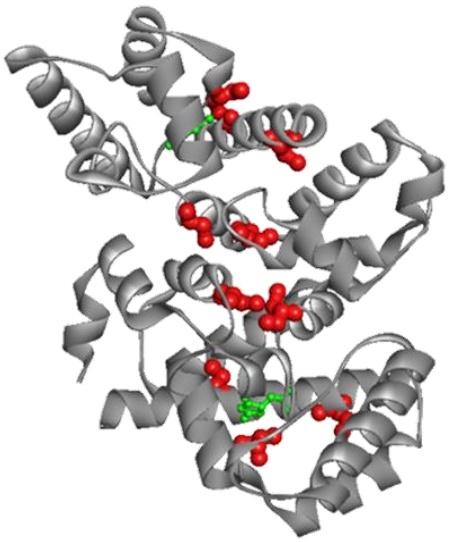


Figure 9 *3D structure of dimeric DrUrad (PDB ID: 2O73). Cysteine residues (red) and the competitive inhibitor (S)-allantoin (green) bound to the active site are shown, respectively, in sphere and ball and stick mode.*

Therefore, it is possible that PEG derivatization of cysteine residues could interfere with DrUrad enzymatic activity either by destabilizing the dimer interface, or by perturbing the accessibility of the substrate to the active site. To better investigate the situation, the PEGylation reaction was carried out in the presence of allantoin, the final product of DrUrad catalysis; allantoin binds to the catalytic site as an inhibitor allowing the protection of Cys-25 and Cys-66 from PEG conjugation. This test resulted in a completely inactive PEGylated enzyme after removal of the inhibitor. We further investigated the role of Cys-25 and Cys-66 by exploiting their proximity, likely to favor the formation of a disulphide bridge under non-reducing conditions, as an alternative way to protect them from PEGylation. Incubation of DrUrad at 20 °C for 5 hours, corresponding to the preparation time of the PEGylated enzyme (reaction and dialysis), without the presence of a reducing agent results in an enzymatic activity that is almost completed abolished (Figure 10). Such loss of specific activity, that is not observed after incubation in ice for an equivalent time, is almost completely reverted after the addition of the reducing agents glutathione or tris(2-carboxyethyl)phosphine (TCEP) (Figure 10). This reversible

inactivation of DrUrad has to be attributed to oxidation of the cysteines close to the active site, and not those participating to the dimer interface, as demonstrated by inspection of the crystallographic structure, that does not show any inter-subunit disulphide bridge [17].

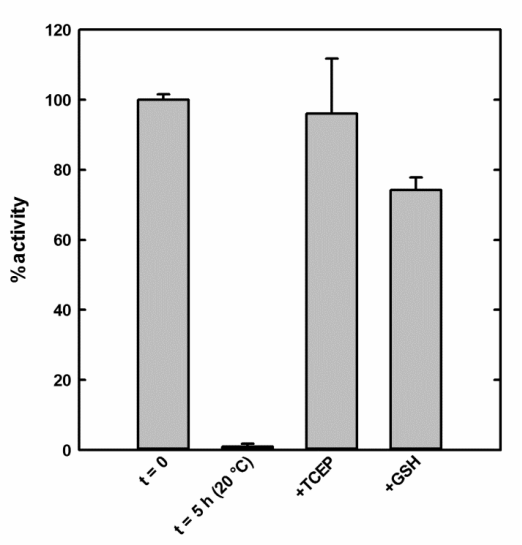


Figure 10 *Enzymatic activity of DrUrad under oxidizing and reducing conditions. From left to right, bars refer to freshly prepared DrUrad (chosen as a reference), the enzyme spontaneously oxidized after an incubation of 5 hours at 20 °C, and the recovery of activity after treatment of the inactivated (oxidized) enzyme with 1 mM tris(2-carboxyethyl)phosphine (TCEP) or 1mM glutathione (GSH).*

In view of a future administration as a protein therapeutic the presence in the plasma of sufficient reducing agents such as glutathione should assure that the protein will be in the fully active form. However, PEGylation trials carried out on inactive (oxidized) DrUrad yielded an inactive enzyme, even after the addition of reducing agents. The fact that the reduction of oxidized cysteines is not sufficient for the recovery of the enzymatic activity indicates that chemical conjugation of residues involved in inter-subunit contacts is an important determinant of function loss upon PEGylation, likely because of dimer destabilization.

Overall, our results led to change PEGylation strategy. In order to use the same PEG derivative (MAL-PEG) changing the target aminoacid a two-step procedure was chosen. The first step involved the modification of the amino groups of the enzyme by reaction with 2-imino thiolane (IMT), adding a spacer arm ending with a thiol group which is able to react with the maleimido ring of MAL-PEG (second step). This technique has already been used in PEGylation of human hemoglobin to increase the number of reaction sites [21,22] while avoiding the derivatization of cysteines with crucial role in protein activity [15]. To further reduce the possible reaction between cysteines and PEG, DrUrad was incubated O/N at 15 °C to increase the thiol groups oxidation before PEGylation. Different reaction conditions were investigated, performing the reactions at 10, 5 or 2.5 MAL-PEG/DrUrad molar ratio (Figure 11), in the presence of a 20x or 40x IMT molar excess with respect to protein dimer concentration.

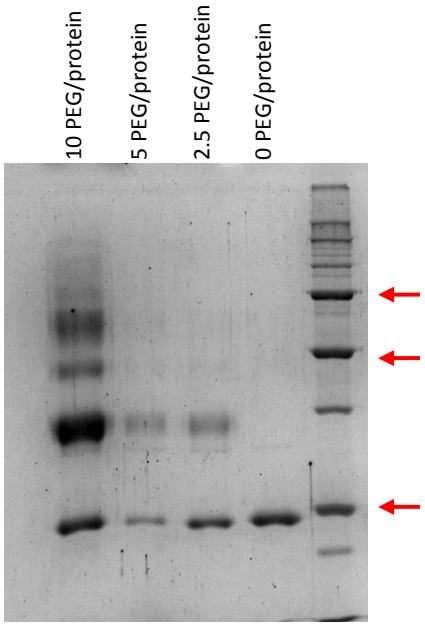


Figure 11 *SDS-PAGE of DrUrad PEGylation after reaction with iminothiolane. PEGylation reactions were carried out at an IMT/protein ratio of 20, at different MAL-PEG/protein ratios (2.5, 5 and 10). Unstained Precision Plus Protein(R) standards (Biorad) were used as MW markers (25, 50 and 75 kDa bands are highlighted in the gel).*

The degree of PEGylation was evaluated by densitometric analysis of SDS-PAGE bands. While the molar excess of IMT had a scarce effect on the PEGylation yield (data not shown), since IMT is highly reactive and a 10x molar excess is enough to assure that most reactive amino groups are functionalized, significant differences were observed between a MAL-PEG/DrUrad ratio of 2.5 or 5, and much less between 5 and 10 (Figure 12A).

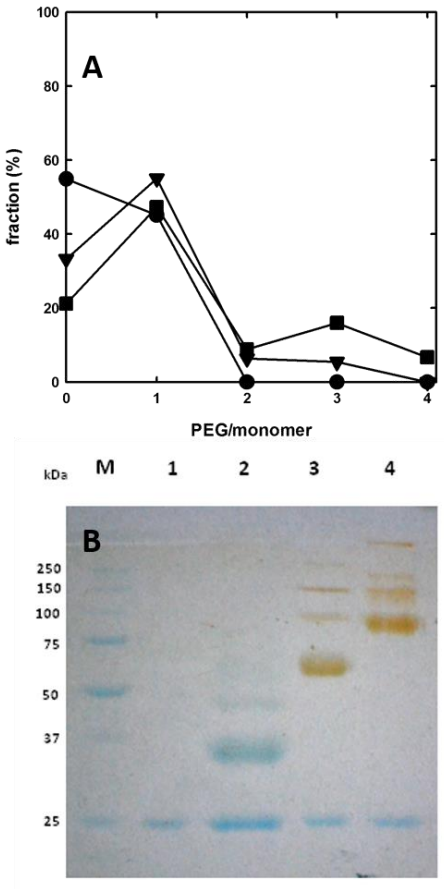


Figure 12 *PEGylation of DrUrad. Panel A: Effect of MAL-PEG/DrUrad ratio on the degree of PEGylation. Circles, triangles and squares refer to a MAL-PEG/DrUrad ratio of 2.5, 5 and 10, respectively. Panel B: SDS-PAGE gel of unmodified and PEGylated DrUrad. Lane M, MW standards; Lane 1, unmodified DrUrad; Lane 2, DrUrad + 5 kDa PEG; Lane 3, DrUrad + 20 kDa PEG; Lane 4, DrUrad + 40 kDa PEG. The gel underwent a double staining procedure, first with Coomassie blue (blue bands), and then with the PEG-specific dye barium iodide (yellow-brown bands).*

Based on this result, and keeping in mind that an excessive PEGylation rate might involve cysteine side chains and prove detrimental for the maintenance of enzyme activity, DrUrad was reacted with MAL-PEG of 5, 20 and 40 kDa molecular weight, keeping constant the 20x molar excess of IMT with respect to protein concentration and a MAL-PEG/DrUrad molar ratio of 5. In all cases, the conjugation reaction resulted in a distribution mainly of unmodified and monoPEGylated monomers, with smaller fractions of di-, tri- and tetraPEGylated monomers (Figure 12B). As previously reported for DrUox and DrUrah, densitometric analysis of the SDS-PAGE bands and combinatorial calculations yielded the average number of PEG chains conjugated to DrUrad monomers and dimers, and the probability to have unmodified dimers (Table 4).

Table 4 *Distribution of un-reacted and PEGylated DrUrad monomers and dimers, and enzymatic activity, after reaction with 5, 20 or 40 kDa MAL-PEG*

	% non PEGylated monomers	% mono-PEGylated monomers	% di-PEGylated monomers	% tri-PEGylated monomers	% tetra-PEGylated monomers	Average no. of PEG bound (per dimer)	% Unmodified dimers	% Enzymatic activity [a]
No PEG	100	0	0	0	0	0	100	24.8±1.6
5 kDa PEG	29	42	10	19	0	2.4	17	39.2±2.4
20 kDa PEG	27	52	9	8	4	2.3	5	17.7±3.1
40 kDa PEG	48	35	9	4	4	1.7	25	23.1±1.7

[a] The activity is expressed as percentage with respect to freshly prepared DrUrad (mean ± SE).

In all cases, and at variance with respect to DrUox and DrUrah, the activity of the PEGylated enzyme was significantly lower than the freshly prepared protein, though comparable to that of samples that underwent the same treatment, with the exclusion of MAL-PEG from the reaction solution (24.8 ± 1.6%). It has to be noted that data reported in Table 4 were originated by experiments carried out in the presence of reducing agents. The only partial retention of enzymatic activity of the unPEGylated protein does not contradict the full functional recovery induced by TCEP, because the activities reported in Table 4 are from samples that underwent O/N dialysis, a procedure that is necessary to eliminate excess reagents, but is badly tolerated by DrUrad. On the other hand, the fact that the activity of PEGylated DrUrad is close to that of the unmodified enzyme indicates that PEG conjugation by extension arm chemistry does not cause any further perturbation of enzyme functionality. In our experimental setting, conjugation with 20 kDa PEG appears to be the most promising approach for future development of a therapeutic form of DrUrad, in spite of a slightly lower residual enzyme activity, since it appears to yield a lower fraction of unmodified dimers (Table 4).

Lyophilization of DrUox, DrUrah and DrUrad: effect of PEG on enzyme activity and stability.

Lyophilization represents the elite strategy for storage of protein therapeutics allowing to produce a stable formulation, in form of a powder to be resuspended at the moment of use, with long shelf life. Therefore, for each enzyme, both unmodified and conjugated with 5, 20 or 40 kDa PEG, the preservation of the enzymatic activity upon lyophilization and resuspension was evaluated and monitored over the time. Each sample was separately flash-frozen and lyophilized in the presence of 4% w/v mannitol and 1% w/v sucrose as lyoprotectants [16]. Once lyophilized, samples were stored at -20 °C. After resuspension in water for parenteral uses, the enzymatic activity of PEGylated DrUox, DrUrah and DrUrad was monitored up to several days and compared with that of the unmodified protein. Resuspended enzymes samples were stored at 4 °C.

In the case of DrUox, PEGylation allows to retain a higher degree of enzymatic activity, about 60-80 % that of a freshly prepared, unmodified sample (Figure 13A). PEG molecular weight does not affect the enzyme activity that is quantitatively maintained throughout a storage time after resuspension of 8 days at 4 °C.

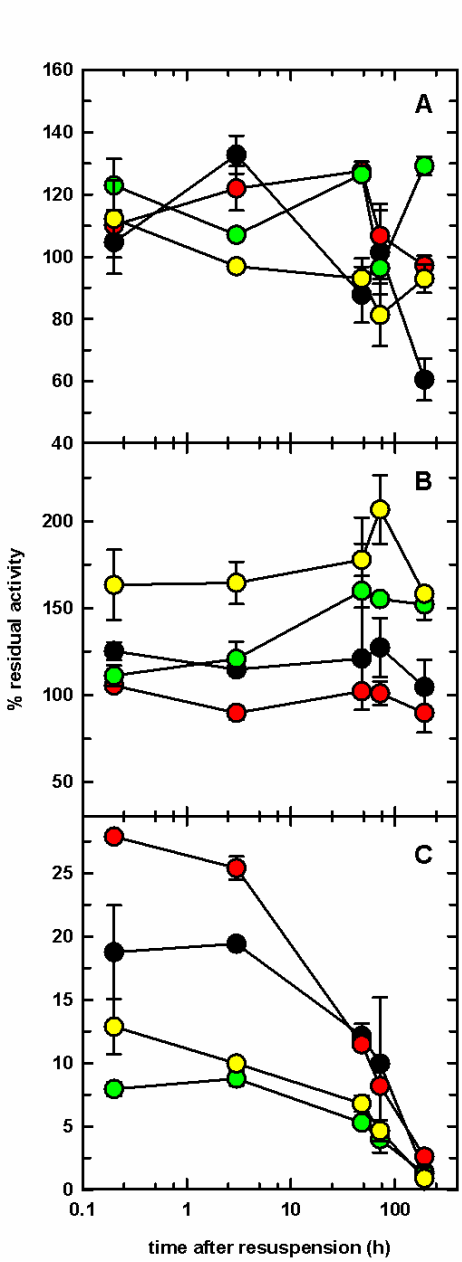


Figure 13 *Activity of PEGylated enzymes after lyophilization and resuspension. Panel A, DrUox; Panel B, DrUrah; Panel C, DrUrad. Different colors correspond to the different molecular weight PEG conjugated to the enzymes: black, no PEG (control); red, 5 kDa PEG; green, 20 kDa PEG; yellow, 40 kDa PEG*

In the case of DrUrah, all PEGylated forms exhibited native-like enzymatic activity that was fully maintained for more than a week (Figure 13B), with no significant differences related to PEG chain length.

DrUrad confirmed to be the more labile component of the enzymatic triad. Not only the initial enzymatic activity after lyophilization and resuspension is significantly lower than that observed for DrUox and DrUrah, but it gradually decays to marginal values within 8 days storage at 4 °C (Figure 13C). PEG does not appear to confer increased stability upon long term storage.

Anyway we can say that lyophilization and resuspension of PEGylated proteins after storage at -20 °C do not significantly affect enzymatic activity. Even in the case of the less stable DrUrad, the activity just after PEGylation (last column in Table 3) is comparable to that initially retrieved after resuspension of lyophilized samples (Figure 13C); despite the apparent lability of DrUrad catalytic activity, the lyophylized enzyme retains a measurable activity for at least a few days upon resuspension, and upon appropriate dosage can effectively complete the set of PEGylated proteins required for full enzymatic conversion of urate to soluble (S)-allantoin, in a formulation suitable for prolonged storage and intravenous administration.

Catalytic competence of a combination of PEGylated DrUox, DrUrah and DrUrad

The final goal of this work was to provide evidence that the three PEGylated enzymes DrUox, DrUrah and DrUrad conserve catalytic activity when co-present in the same reaction solution, allowing the full conversion of urate to (S)-allantoin with no accumulation of reaction intermediates. Since the results exposed in the previous sections indicate that in our experimental conditions the 20 kDa PEG is always a valuable solution (granting good retention of enzyme activity and low amounts of unreacted proteins), for symmetry reason a 20 KDa-PEGylated DrUox-DrUrah-DrUrad combination was chosen for the following tests.

Circular dichroism spectroscopy is able to discriminate between the products of this enzymatic pathway: HIU, OHCU and enantiomeric (S)-allantoin [8]. Therefore, a solution

containing PEGylated DrUox, DrUrah and DrUrad (0.26 μM, 0.02 μM, and 0.12 μM, respectively) was prepared and circular dichroism spectra after addition of a 200 μM urate concentration (a concentration of the same order as the physiological range [23]) were recorded as a function of time. It is clearly visible, over the time, the formation of a species with a positive peak at 208-210 nm and a less intense, broader negative band at around 240-250 nm (Figure 14).

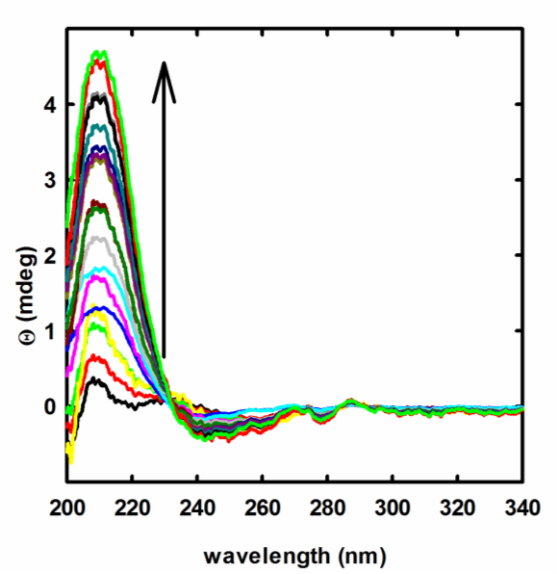


Figure 14 *Time-evolution of far-UV circular dichroism spectra of an urate solution in the presence of DrUox, DrUrah and DrUrad. The solution contained 200 μM urate and 0.26 μM DrUox, 0.02 μM DrUrah and 0.12 μM DrUrad, all conjugated with 20 kDa PEG, in 20 mM potassium phosphate, pH 7.4. The time sequence of the spectra, collected every 5 minutes, is indicated by the arrow.*

These spectral features correspond to the S enantiomer of allantoin [8] demonstrating that using the three enzymes urate is rapidly converted to the soluble metabolite (S)-allantoin, without the accumulation of oxidized purines.

Conclusions

The current therapeutic approaches to uric acid overproduction related to HPRT deficiency are characterized by the inhibition of xanthine oxidase using uricostatic drugs (i.e. allopurinol) or by the uricolytic treatment through the administration of PEGylated urate oxidase, favoring the conversion of the insoluble urate to HIU, that slowly converts to OHCU and racemic allantoin. Both therapies show side effects: in the first case an accumulation of the upstream metabolites occur and in the second one an increased concentration of HIU, known to be toxic [13].

Based on this knowledge the current therapeutic standard could be greatly improved if Urah and Urad, the two enzymes that complete the purine degradation pathway, could be administered in combination with Uox in a biocompatible and bioavailable form.

Our results show that all the three enzymes that constitute the urate degradation pathway present in most animals and lost in humans, can be obtained in recombinant form and maintain an acceptable degree of activity after PEGylation, necessary to reduce immunogenicity and to increase circulation lifetimes of protein therapeutics, and lyophilization, granting long term storage of the formulation. A combination of the three enzymes, in appropriate ratios, demonstrated to effectively convert uric acid to (S)-allantoin in vitro, with no accumulation of intermediate metabolites. These encouraging results pose the basis for future tests in a mouse model of disease.

References

[1] J. M. Berg, J. L. Tymoczko, and L. Stryer, “Purine Bases Can Be Synthesized de Novo or Recycled by Salvage Pathways,” *Biochemestry. 5th edition,* Section 25.2, 2002.

[2] W. Y. Fujimoto and J. E. Seegmiller, “Hypoxanthine-guanine phosphoribosyltransferase deficiency: activity in normal, mutant, and heterozygote-cultured human skin fibroblasts.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 65, no. 3, pp. 577–84, Mar. 1970.

[3] R. A. Gibbs, P. N. Nguyen, L. J. McBride, S. M. Koepf, and C. T. Caskey, “Identification of mutations leading to the Lesch-Nyhan syndrome by automated direct DNA sequencing of in vitro amplified cDNA.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 86, no. 6, pp. 1919–23, Mar. 1989.

[4] H. A. Jinnah *et al.*, “Attenuated variants of Lesch-Nyhan disease.,” *Brain*, vol. 133, no. Pt 3, pp. 671–89, Mar. 2010.

[5] W. L. Nyhan, J. P. O’Neill, H. A. Jinnah, and J. C. Harris, "*Lesch-Nyhan Syndrome*." 1993.

[6] R. J. Torres, C. Prior, and J. G. Puig, “Efficacy and safety of allopurinol in patients with hypoxanthine-guanine phosphoribosyltransferase deficiency.,” *Metabolism.*, vol. 56, no. 9, pp. 1179–86, Sep. 2007.

[7] Kalju Kahn, and Peter Serfozo, and P. A. Tipton, “Identification of the True Product of the Urate Oxidase Reaction,” J. Am. Chem. Soc., vol. 119, no. 23, pp 5435-5442, Jun. 1997.

[8] S. Pipolo *et al.*, “Absolute stereochemistry and preferred conformations of urate degradation intermediates from computed and experimental circular dichroism spectra,” *Org. Biomol. Chem.*, vol. 9, no. 14, pp. 5149–5155, 2011.

[9] I. Ramazzina, C. Folli, A. Secchi, R. Berni, and R. Percudani, “Completing the uric acid degradation pathway through phylogenetic comparison of whole genomes.,” *Nat. Chem. Biol.*, vol. 2, no. 3, pp. 144–8, Mar. 2006.

[10] K. Kahn and P. A. Tipton, “Spectroscopic characterization of intermediates in the

urate oxidase reaction.," *Biochemistry*, vol. 37, no. 33, pp. 11651–9, Aug. 1998.

[11] D. K. L. Cheuk, A. K. S. Chiang, G. C. F. Chan, and S. Y. Ha, "Urate oxidase for the prevention and treatment of tumour lysis syndrome in children with cancer.," *Cochrane database Syst. Rev.*, no. 8, p. CD006945, 2014.

[12] M. K. Reinders and T. L. T. A. Jansen, "New advances in the treatment of gout: review of pegloticase.," *Ther. Clin. Risk Manag.*, vol. 6, pp. 543–50, 2010.

[13] W. S. Stevenson *et al.*, "Deficiency of 5-hydroxyisourate hydrolase causes hepatomegaly and hepatocellular carcinoma in mice.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 107, no. 38, pp. 16625–30, Sep. 2010.

[14] E. B. Weiss, P. Forman, and I. M. Rosenthal, "Allopurinol-induced arteritis in partial HGPRTase deficiency. Atypical seizure manifestation.," *Arch. Intern. Med.*, vol. 138, no. 11, pp. 1743–4, Nov. 1978.

[15] I. Portöro *et al.*, "Towards a novel haemoglobin-based oxygen carrier: Euro-PEG-Hb, physico-chemical properties, vasoactivity and renal filtration," *Biochim. Biophys. Acta - Proteins Proteomics*, vol. 1784, no. 10, pp. 1402–1409, Oct. 2008.

[16] R. E. Johnson, C. F. Kirchhoff, and H. T. Gaud, "Mannitol-sucrose mixtures--versatile formulations for protein lyophilization.," *J. Pharm. Sci.*, vol. 91, no. 4, pp. 914–922, 2002.

[17] L. Cendron, R. Berni, C. Folli, I. Ramazzina, R. Percudani, and G. Zanotti, "The structure of 2-oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline decarboxylase provides insights into the mechanism of uric acid degradation.," *J. Biol. Chem.*, vol. 282, no. 25, pp. 18182–9, Jun. 2007.

[18] G. Zanotti, L. Cendron, I. Ramazzina, C. Folli, R. Percudani, and R. Berni, "Structure of zebra fish HIUase: insights into evolution of an enzyme to a hormone transporter.," *J. Mol. Biol.*, vol. 363, no. 1, pp. 1–9, Oct. 2006.

[19] D. Caccia *et al.*, "PEGylation Promotes Hemoglobin Tetramer Dissociation," *Bioconjug. Chem.*, vol. 20, no. 7, pp. 1356–1366, Jul. 2009.

[20] H. Tian, Y. Guo, X. Gao, and W. Yao, "PEGylation enhancement of pH stability of uricase via inhibitive tetramer dissociation.," *J. Pharm. Pharmacol.*, vol. 65, no. 1,

pp. 53–63, Jan. 2013.

[21] D. Li, B. N. Manjula, and A. S. Acharya, “Extension arm facilitated PEGylation of hemoglobin: correlation of the properties with the extent of PEGylation.,” *Protein J.*, vol. 25, no. 4, pp. 263–74, Jun. 2006.

[22] K. D. Vandegriff and R. M. Winslow, “Hemospan: design principles for a new class of oxygen therapeutic.,” *Artif. Organs*, vol. 33, no. 2, pp. 133–8, Feb. 2009.

[23] G. Desideri *et al.*, “Is it time to revise the normal range of serum uric acid levels?,” *Eur. Rev. Med. Pharmacol. Sci.*, vol. 18, no. 9, pp. 1295–306, 2014.

Chapter 3

PEGylation of engineered human hemoglobin A towards the development of a hemoglobin-based blood substitute

Introduction

Blood transfusion represents a key life-saving therapy in many chronic or acute diseases, during or upon surgical interventions and in many traumatic events. The fast recovery of oxygen delivery to organs makes blood transfusion the elite treatment to restore circulation volume, thus maintaining a physiological blood pressure and an efficient blood flux. While the blood pressure control can be fulfilled to some extent by using plasma expanders, up to now no approved alternatives to red blood cells exists for restoring oxygen delivery capacity [1].

There is a worldwide increased demand for blood transfusions. In Western countries, this trend is due to the aging of the population, whereas in Third World countries this is due to the increase of infectious and parasitic diseases. Moreover, blood supply in Western countries is not increasing, and in Third World countries is not sufficient due to the spreading of infections, such as HIV [2].

Thanks to the voluntary action of blood donors, transfusion represents an effective and safe treatment when accurate procedures are strictly followed [3], but this oxygen-based therapy is exposed to a series of limitations. For example, blood transfusion usually requires a careful cross-matching between donor and acceptor; blood units are not always promptly available in emergency events such as accidents and battle fields. In order to avoid potential health risks, expensive analyses are required to detect infective agents in blood [1]. Furthermore, in stored blood units the intracellular levels of effectors

such as 2-3 bisphosphoglycerate (2,3-BPG), potassium ions, lactate, together with the intracellular pH of red cells, change as a function of time affecting the efficacy of hemoglobin as oxygen carrier [4]. There are examples of clinical needs unmet by blood transfusions such as the supply of oxygen to ischemic tissues where red blood cells cannot penetrate, and the treatment of patients that are immunoreactive to all blood types or that refuse blood transfusions for religious concerns.

Based on this knowledge, the development of an "artificial blood" or a "blood substitute" represents one of the main challenges of biotechnology and modern medicine. Attempts to develop substances that could replace blood span more than 7 decades, and exploited several strategies many of which are based on cell-free hemoglobin-based oxygen carriers (HBOCs) [5].

Hemoglobin (Hb), contained in red blood cells, efficiently delivers oxygen from lungs to tissues thanks to its four heme groups incorporated in a tetrameric protein structure. Oxygen binds to Hb cooperatively, with an affinity increased as a function of saturation and modulated by specific allosteric effectors [6]. However, Hb, when free in the plasma, causes hypertension and renal failure. Inside the erythrocytes Hb is almost entirely in a tetrameric form, while free in solution at low μM concentration dissociates in dimers, that are small enough to be filtered by glomerulus causing nephrotoxicity [1]. The heme iron of the Hb dimer is more prone to oxidation than in the tetramer and the formation of ferric ions triggers a cascade of reactions that generate reactive oxygen and nitrogen species involved in the oxidative damage [7] that cannot be counterbalanced by the plasma antioxidant activity [8]. Hb also binds NO, either in its deoxy form through the reaction with the heme group, leading to the formation of nitrosyl-hemoglobin, or when oxygenated forming cysteine β93 S-nitroso oxyhemoglobin or methemoglobin [9,10]. Cell-free Hb is much more effective in scavenging NO with respect to Hb inside red blood cells because the interactions are not limited by NO diffusion through the erythrocyte membrane [11]. Furthermore, Hb both in dimeric and tetrameric forms extravasates and scavenges NO directly in its site of action [12], causing hypertension for the loss of NO regulatory activity on the vascular tone.

Taking into account the wide knowledge about the oxygen transport physiology and Hb structure and function, different strategies aimed at the development of Hb derivatives have been carried out via either genetic or chemical modifications of Hb.

Recombinant technology allowed to modulate functional properties through the production of engineered Hb molecules by amino acid substitutions, for example to prevent NO scavenging [13]. Another genetic approach is based on mutations that generate a Hb with a single cysteine on the surface, able to autopolymerize, as in the case of hemoglobin Polytaur [14,15,16].

Different strategies involved also chemical modifications of Hb using reagents, such as glutaraldehyde, that cross-link tetramers intra- or inter-molecularly, generating stable tetramers or polymers [17,18] (as in the case of Hemopure and Polyheme) or biocompatible organic compounds that decorate Hb surface such as polyethylene glycol (PEG) [2,19,20] (Hemospan and Euro-PEG-Hb). Structural and functional investigations on PEGylated Hb highlighted two major shortcomings: heterogeneity in the degree of PEGylation and enhanced tetramer dissociation upon PEG conjugation that may cause vasoconstriction due to enhanced HBOC extravasation [19]. Basically, HBOCs show adverse effects that limit their applicability; a meta-analysis of the clinical studies carried out with HBOCs indicated that their use is associated with a significantly increased risk of myocardial infarction and death [21]. The HBOCs are designed to be an alternative to blood transfusion mimicking just one component of the whole blood, hemoglobin; for this reason interferences with other functions are observed. Despite the failure of many HBOCs, based on the knowledge accumulated over the years a combination of different approaches (*i.e* hemoglobin genetic and chemical manipulation) could be useful to overcome HBOCs shortcomings or to better understand the toxicity mechanisms.

Research purpose

Over the years, several attempts have been made to produce an alternative to blood transfusion. Different strategies have been adopted and after almost fifty years, a safe product has not been developed yet. Starting from the knowledge that molecular homogeneity of a HBOC is a critical issue to evaluate the relationship among its chemical structure, biological activity and adverse effects, the genetic manipulation of human hemoglobin has been combined with surface modification using poly-ethylene glycol to create a PEGylated Hb more homogeneous than the previously available products.

Hb mutants, expressed in recombinant form in *E.coli*, have been engineered by replacing the cysteine β93 residue with an alanine residue, obtaining a sample named A11 Hb with no exposed cysteines residues; A11 was further engineered by replacing with cysteine a surface alanine residue on alfa or beta chains, obtaining the A12 and A13 Hb samples, respectively. In the present work, the specific sites of the mutations are not mentioned because the Hb mutants are under evaluation at the patent office. These genetically added cysteines will be used for the direct attachment of PEG molecules functionalized with a maleimido group aiming to obtain homogeneous products.

The Hb mutants will be then characterized in terms of oxygen affinity and cooperativity before and after PEGylation.

Materials and Methods

Protein concentration measurement

Absorption spectra were collected using a Cary 4000 UV-Vis spectrophotometer. A buffer containing 100 mM HEPES, 100 mM NaCl, 1.2 mM Na phosphate, 1 mM ETDA, pH 7 was used for PEGylation and p50 measurements. The buffer was bubbled with carbon monoxide before adding protein samples. For each sample, an absorption spectrum was collected in the range 450-700 nm to check concentration and oxidation state of the proteins: firstly, each sample was diluted with CO-equilibrated buffer (5 µl of stock protein samples + 115 µl of CO-equilibrated buffer) to check the fraction of oxidized Hb. Then, sodium dithionite (powder) was added to the sample to remove oxidized Hb, thus obtaining a pure carbon monoxide Hb (HbCO). Another spectrum was then acquired to determine protein concentration by using the molar extinction coefficient of HbCO (ε= 13.4 mM^{-1} cm^{-1}). All the samples were brought to 1 mM concentration (on monomer basis) with the CO-equilibrated PEGylation buffer.

Protein PEGylation under CO-saturated and anaerobic conditions

The HbA mutants A11, A12, A13, and the recombinant wt Hb A1 were directly conjugated with 20 kDa maleimido-functionalized PEG (MAL-PEG) (MeO-PEG-mal, Iris Biotech) in CO-saturated and anaerobic conditions. MAL-PEG stock solution (10 mM) was freshly prepared from powder just before the reactions. The final volume of the samples to be PEGylated was 200 µl.

In the first case, the Hb samples (0.25 mM on tetramer basis), previously equilibrated with CO gas, were reacted with 20 kDa MAL-PEG in 12:1 PEG:tetramer molar ratio, at 25 °C for 60 min. The reactions were quenched by adding 2 µl of a 0.9 M cysteine solution to each sample. The PEGylation under anaerobic conditions was carried out in vials fluxed with inert gas, at 25 °C, by adding deoxygenated solutions of reagents and using the same protein/reagents ratio used in the previous protocol. Unreacted reagents and unmodified Hb were removed by dialysis (100 kDa MWCO centrifuge filters). A colored flow through

was observed for all the samples due to unmodified and/or monoPEGylated Hb. After dialysis, absorption spectra were collected to check concentration and oxidation state. The samples were then freeze-thawed and stored at -80 °C.

SDS-PAGE of PEGylated samples

Denaturating protein electrophoresis (SDS-PAGE, 12% acrilamide/bisacrilamide) was carried out loading 10 μg of unmodified and PEGylated samples. Protein samples were added with sample buffer and brought to 100 °C for 10 minutes. Protein staining was carried out with Biosafe® Coomassie. The SDS-PAGE gels were scanned using a ChemiDoc® imager (Biorad) and bands intensity was evaluated by densitometric analysis.

Mass spectrometry

Spectra were acquired using a 48000 Plus MALDI TOF/TOF™ AbSciex mass spectrometer. The samples were prepared by dried droplet method using a 10 mg/Ml α-cyano-4-hydroxycinnamic acid (CHCA) solution prepared in acetonitrile 75 % - TFA 0.05 %.

Oxygen binding measurement

For each sample (equilibrated in air), the absorption spectrum was collected immediately after thawing. Ascorbate and catalase were added to the solution before titrations to reduce metHb and to limit its formation during titration. The samples were deoxygenated using a helium flow and then equilibrated with oxygen/helium gas mixtures at different oxygen partial pressures. Each titration required about 5 hours.

p50 measurements were carried out diluting the stock Hb solutions in 100 mM HEPES, 100 mM sodium chloride, 1.2 mM sodium phosphate, 1 mM EDTA, pH 7.0. The final protein concentration was 100 μM (on heme basis).

The determination of oxygen binding curves of Hb samples was carried out at 25 °C using a tonometer; the sample chamber consists of a 2 mm cuvette fused to a 2.5 ml open-top threaded reservoir, where a screw cap with fittings for inlet and outlet gas lines was

mounted. In this way, the sample chamber is continuously fluxed by the gas mixture, maintaining a defined oxygen partial pressure, with no pressure changes within the chamber [2]. Gas mixtures were prepared by using an Environics 200 gas mixture generator with a gas flow of 50 ml/min. During sample equilibration (30-40 minutes), the tonometer is thermostatted at 25 °C in a shaking bath. The fractional oxygen saturations and the fraction of oxidized hemes are determined by fitting the observed spectra in the 470-670 nm range to a linear combination of reference spectra [22,23], i.e. the spectra of pure deoxygenated, oxygenated and oxidized Hb.

Results and discussion

Human Adult hemoglobin (HbA) was expressed in recombinant form in *E. coli* obtaining A1 Hb sample; A1 sequence was engineered by replacing the cysteine β93 residue with an alanine residue, in order to delete the reactive cysteine constitutively present in human hemoglobin, obtaining A11. A11 was further engineered by replacing an alanine residue on the surface of either α or β chain with a cysteine, obtaining A12 and A13, respectively. Summarizing, four recombinant Hb samples have been obtained and produced (by collaborators at the University of Essex, UK):

- A1, the recombinant form of HbA
- A11, with no reactive cysteine residues
- A12, with two exposed cysteine residues per tetramer (on α subunits)
- A13, with two exposed cysteine residues per tetramer (on β subunits).

The number of reactive cysteine residues were evaluated for the mutants A11, A12 and A13 by reaction with PMB. As expected, it was found that A12 and A13 contain two reactive cysteines per tetramer whereas A11 contains no reactive cysteines.

Engineered HbA characterization

The purity of recombinant HbAs, as assessed by SDS-PAGE, is reported in Figure 1. The visual inspection and the densitometric analysis of gels indicate that predominantly A12, but also A11 and A13 preparations contain traces of a band at around 26-30 kDa. Band excision, followed by digestion and peptide mass fingerprint allowed the identification of the band as being constituted by beta subunits. This contaminant has already been observed during recombinant expression of heme proteins due to metal-catalyzed oxidative deamination and identified as beta dimers [24].

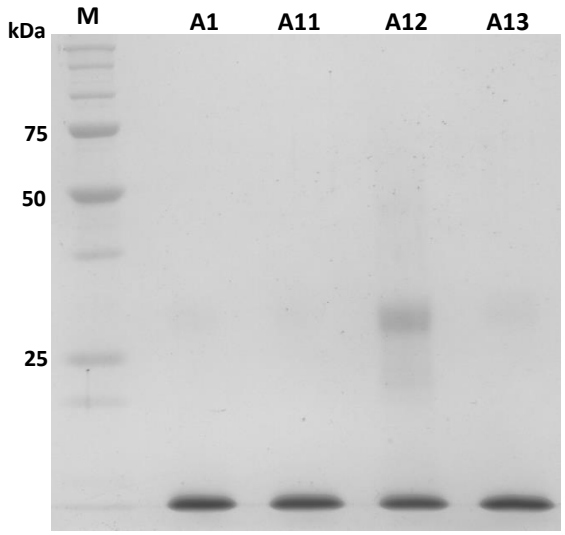


Figure 1 *SDS PAGE of recombinant HbAs: A1, A11, A12, A13. Unstained Precision Plus Protein(R) standards (Biorad) were used as MW markers (25, 50 and 75 kDa bands are highlighted in the gel).*

The oxygen binding curves of A1, A11, A12, and A13 were determined (Figure 2). A comparison of the oxygen binding curve of human purified Hb (A0) with recombinant Hbs, clearly showed that recombinant Hbs exhibit a higher affinity and a moderately reduced cooperativity.

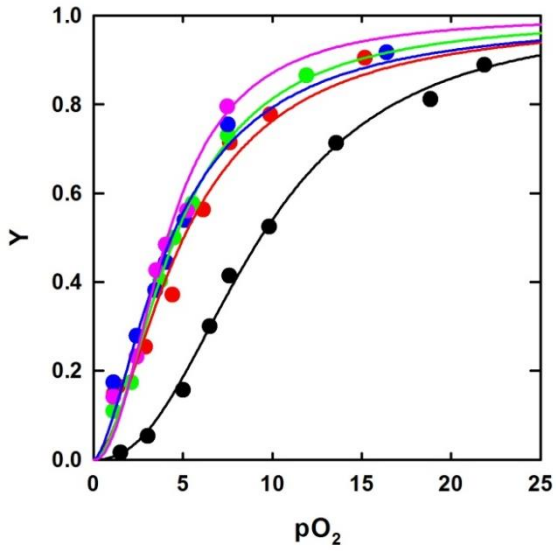


Figure 2 *Oxygen binding curves in 100 mM HEPES, 100 mM sodium chloride, 1.2 mM sodium phosphate, 1 mM EDTA, ascorbate and catalase as a reducing system, pH 7 at 25 °C. Curves refer to Hb A0 (black), A1 (red), A11 (green), A12 (blue), A13 (pink).*

The p50 and Hill coefficients (n) are reported in Table 1. The comparison of oxygen binding properties of the different mutants indicates that these mutations do not significantly affect Hb oxygen affinity and cooperativity.

Table 1 *Oxygen binding parameters for non-PEGylated recombinant HbAs*

	A0	**A1**	**A11**	**A12**	**A13**
p50 (torr)	9.3 ± 0.2	5.0 ± 0.3	4.6 ± 0.1	4.4 ± 0.2	4.2 ± 0.2
n	2.4 ± 0.1	1.7 ± 0.2	1.9 ± 0.1	1.6 ± 0.2	1.9 ± 0.3

PEGylation of recombinant HbA forms

Recombinant HbAs were conjugated with maleimido-functionalized PEG (MAL-PEG) (20 kDa MW) in either CO-saturated or anaerobic conditions. A direct conjugation approach was chosen, using as reactive protein groups the genetically cysteines introduced on A12 and A13 mutants, aimed at the development of homogeneously PEGylated products. The

recombinant Hb A1 and A11 mutant, with no exposed cysteines, were also PEGylated as control. The degree of PEGylation was assessed by 1D-PAGE (12%) (Figure 3).

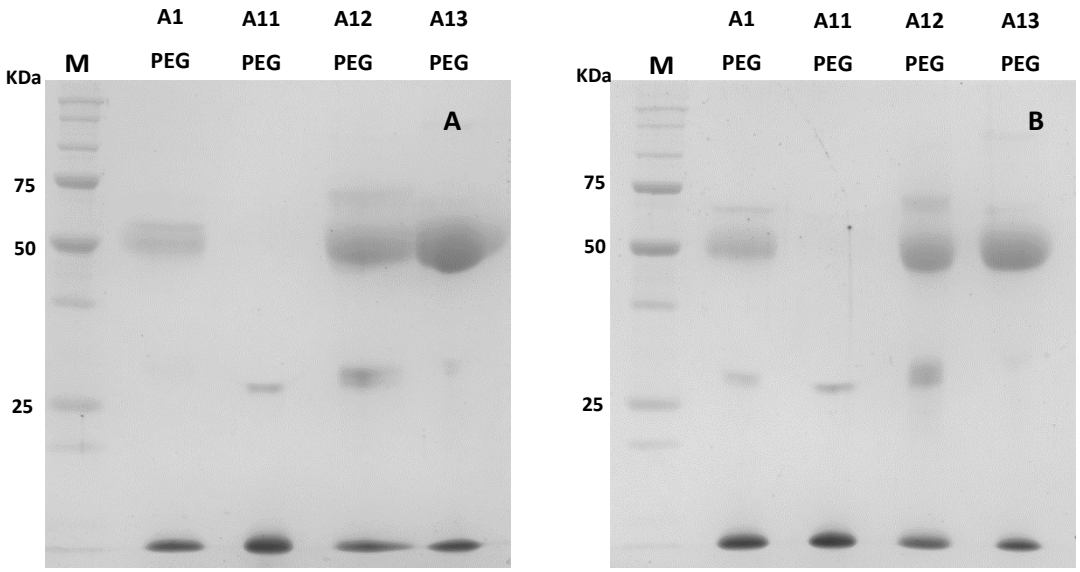


Figure 3 *SDS PAGE of PEGylated Hb A1, A11, A12 and A13 in CO-saturated (Panel A) and anaerobic (Panel B) conditions. Unstained Precision Plus Protein(R) standards (Biorad) were used as MW markers (25, 50 and 75 kDa bands are highlighted in the gel).*

The comparison between the SDS-PAGE of samples PEGylated in CO-saturated and anaerobic conditions clearly shows that PEGylation occurs under both conditions with similar yields. Hb A1 PEGylation under anaerobic conditions should not occur because it is known that cysteine β93 is only exposed and reactive in the R state whereas it is buried in the T state. However, the relatively high amount of metHb, that is known to be in the R state, might be responsible for the observed PEGylation. Furthermore, A11 does not react under either anaerobic or CO-saturated conditions, as expected, because no exposed cysteines are present.

PEGylation in the R state occurred for A1, A12 and A13 with good yield, as estimated from the densitometric analysis, which detected a 50% of unmodified and 50% of mono-PEGylated globin bands for PEGylated A12 and A13. For Hb A1 a 30% of mono-PEGylated

globins and a 70% of unmodified globin bands was observed, demonstrating that the cysteine residues introduced by site-specific mutagenesis are more exposed and hence more reactive than betaCys93.

In order to confirm PEG conjugation under CO-saturated and anaerobic conditions for A12 and A13 mutants, a mass spectrometry analysis was carried out using MALDI TOF/TOF instrumentation (Figure 4).

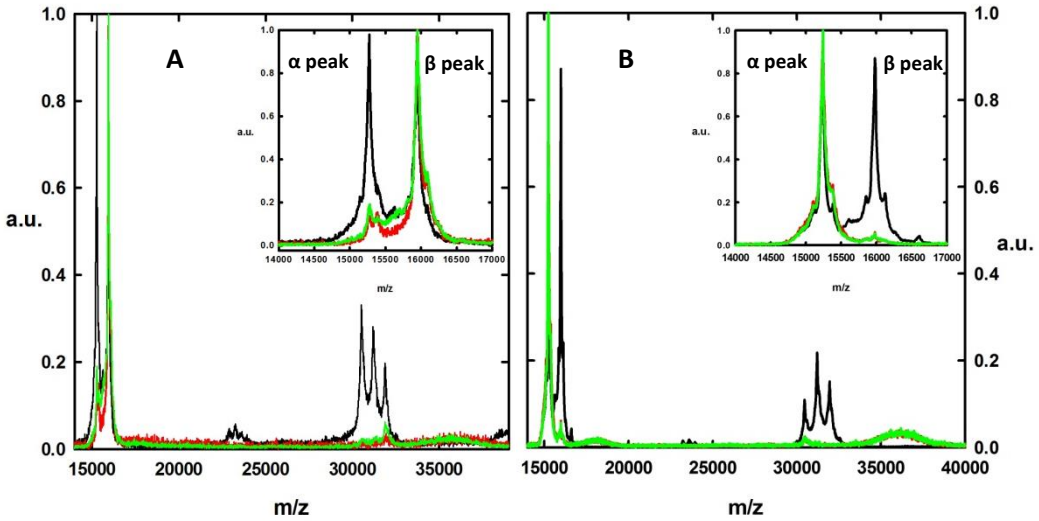


Figure 4 *Comparison of the mass spectra of A12 (Panel A) and A13 (Panel B) either unmodified (black line), or PEGylated under CO-saturated conditions (red line) or under anaerobic conditions (green line). Inset: α-peak is associated to the Hb α-monomer and β-peak to the Hb β-monomer.*

The mass spectra of unmodified or PEGylated A12 (Figure 4A) and A13 (Figure 4B), either obtained under CO saturation or in anaerobic conditions, were compared. In the region 14000-17000 m/z (Figure 4, insets), A12 and A13 showed both α and β peaks, while in the region 30000-34000 three peaks appear, corresponding to α-α, α-β and β-β complexes, possibly due to artifacts generated by mass spec experimental conditions, as also observed for human HbA. Upon PEGylation in conditions favoring the R or the T state, the A12 sample (Figure 4A) showed an almost complete disappearance of the peak

corresponding to the α chain. This is due to the derivatization of the cysteine, genetically introduced in the alpha subunit, with 20 KDa MAL-PEG that exhibited a broad peak around 36000 m/z. It is known that PEG-conjugated proteins exhibit a low propensity to fly and, consequently, to give signal in mass spectrometers. The same behavior was observed for A13 (Figure 4B) where the peak corresponding to the beta subunit disappears and a peak at 36000 m/z appears. This means that, independently of the T or R quaternary state, cysteines react giving similar PEGylation yields.

In order to investigate the oxygen binding profile of A12 and A13 mutants after PEGylation in both CO-saturated (R-state) and anaerobic (T-state) conditions, the oxygen binding curves were determined in the same experimental conditions used for the characterization of the unmodified Hb samples. In table 2 the p50 and Hill coefficients (n) values are reported.

Table 2 *Oxygen binding parameters for recombinant A12 and A13 Hb PEGylated in both CO-saturated (R state) and anaerobic (T state) conditions*

	A12-PEG R state	**A12-PEG T state**	**A13-PEG R state**	**A13-PEG T state**
p50 (torr)	4.1 ± 0.2	3.0 ± 0.2	4.6 ± 0.2	4.0 ± 0.2
n	1.6 ± 0.1	1.3 ± 0.2	2.0 ± 0.2	1.6 ± 0.1

Comparing the oxygen binding parameters of recombinant Hbs before (Table 1) and after PEGylation (Table 2), it is interesting to notice that the conjugation with PEG did not cause a significant change in p50 and cooperativity for both A12 and A13, independently of the conjugation conditions.

Looking back into the literature, the retention of oxygen binding properties after PEG attachment represents a novelty in the development of PEGylated HBOCs. It is largely appreciated that PEGylation affects Hb in different ways: perturbation of the T and R state [25], perturbation of the quaternary transition [26], and influence on the tetramer-dimer equilibrium [27]. PEGylated HBOCs usually change their oxygen binding profile after

conjugation, as in the case of MP4 [20] and Euro-PEG-Hb [2]. MP4 is obtained via the conjugation of human Hb with an average of 6-7 5 kDa PEG chains under oxygenated conditions in a two step procedure characterized by thiolation and PEG-maleimido derivative conjugation [20,28]; Euro-PEG-Hb is a 7-PEG molecules derivative Hb obtained using the same approach but in deoxygenated conditions [2]. The altered oxygen affinity and cooperativity, observed in both cases, might not be due merely to lysines modifications. It is known that interface mutants or chemically modified Hbs characterized by an increased oxygen affinity and reduced cooperativity exhibit a perturbed tetramer-dimer equilibrium, favoring the latter species [29]. Furthermore the same behavior is observed for non-human Hb species, such as for the tetra-PEGylated canine Hb [30]. The proposed recombinant Hb mutants, A12 and A13, may be homogenously modified without affecting oxygen binding profile. Given the complexity of the protocol for PEGylation under anaerobic conditions, their preparation can be more easily carried out under aerobic conditions. Moreover, the presence of CO will prevent oxidation that might occur during processing. Finally their further genetic manipulation, for example to reduce or avoid autoxidation or NO scavenging, should lead to mutants in which PEGylation will not affect the oxygen binding properties.

Conclusions

In this work the genetic manipulation of recombinant human hemoglobin has been combined with protein surface chemical modification, for the development of homogeneous hemoglobin-based blood substitutes.

Starting from the production of human recombinant Hb, expressed in *E. coli*, Hb mutants have been engineered by replacing the reactive cysteine β93 with an alanine residue, obtaining the A11 Hb sample. A11 was further engineered by replacing a surface alanine residue on alfa or beta chains, obtaining A12 and A13 Hb samples, respectively. The genetically introduced cysteines were used for the attachment of PEG aiming to develop PEGylated compounds that are homogeneous in terms of derivatization. These compounds were PEGylated under CO-saturated and anaerobic conditions and characterized in terms of oxygen affinity and cooperativity.

The cysteine residues introduced in A12 and A13 mutants allowed us to have a better control of PEG conjugation, obtaining homogeneous products as characterized through 1D-PAGE and mass spectrometry. Interestingly, the oxygen binding profile remains almost the same after PEGylation in both conditions. Thus, given the higher complexity of the protocol for PEGylation under anaerobic conditions, the preparation of these mutants can be more easily carried out under aerobic conditions.

The production of these PEGylated Hb mutants will be carried out in large scale for studies in animal models. Based on our preliminary data, these studies might demonstrate that PEGylated Hb, once proper mutations will be introduced, may allow to overcome the main shortcomings that have impeded so far the development of safe and effective HBCOCs, such as autooxidation, vasoactivity, NO scavenging.

References

[1] A. Mozzarelli, L. Ronda, S. Faggiano, S. Bettati, and S. Bruno, “Haemoglobin-based oxygen carriers: research and reality towards an alternative to blood transfusions.,” *Blood Transfus.*, vol. 8 Suppl 3, no. Suppl 3, pp. s59-68, Jun. 2010.

[2] I. Portörő *et al.*, “Towards a novel haemoglobin-based oxygen carrier: Euro-PEG-Hb, physico-chemical properties, vasoactivity and renal filtration,” *Biochim. Biophys. Acta - Proteins Proteomics*, vol. 1784, no. 10, pp. 1402–1409, Oct. 2008.

[3] L. M. Napolitano *et al.*, “Clinical practice guideline: Red blood cell transfusion in adult trauma and critical care,” *Crit. Care Med.*, vol. 37, no. 12, pp. 3124–3157, Dec. 2009.

[4] E. Bennett-Guerrero *et al.*, “Evolution of adverse changes in stored RBCs,” *Proc. Natl. Acad. Sci.*, vol. 104, no. 43, pp. 17063–17068, Oct. 2007.

[5] J.-Y. Chen, M. Scerbo, and G. Kramer, “A review of blood substitutes: examining the history, clinical trial results, and ethics of hemoglobin-based oxygen carriers.,” *Clinics (Sao Paulo).*, vol. 64, no. 8, pp. 803–13, 2009.

[6] W. A. Eaton, E. R. Henry, J. Hofrichter, and A. Mozzarelli, “Is cooperative oxygen binding by hemoglobin really understood?,” *Nat. Struct. Biol.*, vol. 6, no. 4, pp. 351–358, Apr. 1999.

[7] P. W. Buehler and A. I. Alayash, “Toxicities of hemoglobin solutions: in search of in-vitro and in-vivo model systems,” *Transfusion*, vol. 44, no. 10, pp. 1516–1530, Oct. 2004.

[8] H. Sies, “Total antioxidant capacity: appraisal of a concept.,” *J. Nutr.*, vol. 137, no. 6, pp. 1493–5, Jun. 2007.

[9] R. P. Patel, “Biochemical aspects of the reaction of hemoglobin and NO: implications for Hb-based blood substitutes.,” *Free Radic. Biol. Med.*, vol. 28, no. 10, pp. 1518–25, May 2000.

[10] C. E. Cooper, “Nitric oxide and iron proteins,” *Biochim. Biophys. Acta - Bioenerg.*, vol. 1411, no. 2, pp. 290–309, 1999.

[11] M. W. Vaughn, K. T. Huang, L. Kuo, and J. C. Liao, “Erythrocytes possess an intrinsic barrier to nitric oxide consumption.,” *J. Biol. Chem.*, vol. 275, no. 4, pp. 2342–8, Jan. 2000.

[12] J. S. Olson, E. W. Foley, C. Rogge, A.-L. Tsai, M. P. Doyle, and D. D. Lemon, “No scavenging and the hypertensive effect of hemoglobin-based blood substitutes,” *Free Radic. Biol. Med.*, vol. 36, no. 6, pp. 685–697, Mar. 2004.

[13] D. H. Doherty *et al.*, “Rate of reaction with nitric oxide determines the hypertensive effect of cell-free hemoglobin,” *Nat. Biotechnol.*, vol. 16, no. 7, pp. 672–676, Jul. 1998.

[14] K. M. Bobofchak *et al.*, “A recombinant polymeric hemoglobin with conformational, functional, and physiological characteristics of an in vivo O_2 transporter,” *Am. J. Physiol. - Hear. Circ. Physiol.*, vol. 285, no. 2, pp. H549–H561, Aug. 2003.

[15] C. Fronticelli, D. Arosio, K. M. Bobofchak, and G. B. Vasquez, “Molecular engineering of a polymer of tetrameric hemoglobins.,” *Proteins*, vol. 44, no. 3, pp. 212–22, Aug. 2001.

[16] C. Fronticelli and R. C. Koehler, “Design of Recombinant Hemoglobins for Use in Transfusion Fluids,” *Crit. Care Clin.*, vol. 25, no. 2, pp. 357–371, Apr. 2009.

[17] T. M. Chang, “Modified hemoglobin-based blood substitutes: crosslinked, recombinant and encapsulated hemoglobin.,” *Vox Sang.*, vol. 74 Suppl 2, pp. 233–41, 1998.

[18] P. Keipert, J. Minkowitz, and T. M. Chang, “Cross-linked stroma-free polyhemoglobin as a potential blood substitute.,” *Int. J. Artif. Organs*, vol. 5, no. 6, pp. 383–5, Nov. 1982.

[19] D. Caccia *et al.*, “PEGylation Promotes Hemoglobin Tetramer Dissociation,” *Bioconjug. Chem.*, vol. 20, no. 7, pp. 1356–1366, Jul. 2009.

[20] K. D. Vandegriff, A. Malavalli, J. Wooldridge, J. Lohman, and R. M. Winslow, “MP4, a new nonvasoactive PEG-Hb conjugate.,” *Transfusion*, vol. 43, no. 4, pp. 509–16, Apr. 2003.

[21] C. Natanson, S. J. Kern, P. Lurie, S. M. Banks, and S. M. Wolfe, “Cell-Free Hemoglobin-Based Blood Substitutes and Risk of Myocardial Infarction and Death,” *JAMA*, vol. 299, no. 19, p. 2304, May 2008.

[22] S. Bettati and A. Mozzarelli, “T state hemoglobin binds oxygen noncooperatively with allosteric effects of protons, inositol hexaphosphate, and chloride.,” *J. Biol. Chem.*, vol. 272, no. 51, pp. 32050–5, Dec. 1997.

[23] C. Rivetti, A. Mozzarelli, G. L. Rossi, E. R. Henry, and W. A. Eaton, “Oxygen binding by single crystals of hemoglobin,” *Biochemistry*, vol. 32, no. 11, pp. 2888–2906, Mar. 1993.

[24] J. Levine *et al.*, “Identification of a Nickel(II) Binding Site on Hemoglobin Which Confers Susceptibility to Oxidative Deamination and Intramolecular Cross-linking,” *J. Biol. Chem.*, vol. 273, no. 21, pp. 13037–13046, May 1998.

[25] F. E. Lui, P. Dong, and R. Kluger, “Polyethylene glycol conjugation enhances the nitrite reductase activity of native and cross-linked hemoglobin.,” *Biochemistry*, vol. 47, no. 40, pp. 10773–80, Oct. 2008.

[26] D. I. Svergun *et al.*, “Solution structure of poly(ethylene) glycol-conjugated hemoglobin revealed by small-angle X-ray scattering: implications for a new oxygen therapeutic.,” *Biophys. J.*, vol. 94, no. 1, pp. 173–81, Jan. 2008.

[27] T. Hu, D. Li, B. N. Manjula, M. Brenowitz, M. Prabhakaran, and S. A. Acharya, “PEGylation of Val-1(alpha) destabilizes the tetrameric structure of hemoglobin.,” *Biochemistry*, vol. 48, no. 3, pp. 608–16, Jan. 2009.

[28] K. D. Vandegriff, A. Malavalli, G. M. Mkrtchyan, S. N. Spann, D. A. Baker, and R. M. Winslow, “Sites of Modification of Hemospan, a Poly(ethylene glycol)-Modified Human Hemoglobin for Use As an Oxygen Therapeutic,” *Bioconjug. Chem.*, vol. 19, no. 11, pp. 2163–2170, Nov. 2008.

[29] V. J. LiCata, P. C. Speros, E. Rovida, and G. K. Ackers, “Direct and indirect pathways of functional coupling in human hemoglobin are revealed by quantitative low-temperature isoelectric focusing of mutant hybrids.,” *Biochemistry*, vol. 29, no. 42, pp. 9771–83, Oct. 1990.

[30] S. A. Acharya, V. N. Acharya, N. D. Kanika, A. G. Tsai, M. Intaglietta, and B. N. Manjula, “Non-hypertensive tetraPEGylated canine haemoglobin: correlation between PEGylation, O2 affinity and tissue oxygenation,” *Biochem. J.*, vol. 405, no. 3, 2007.

Chapter 4

Development of an innovative user-friendly device for the detection of *Staphylococcus aureus* from biological sources

Introduction

Staphylococcus aureus (*S. aureus*), a common commensal of skin and nares, is one of the leading causes of health care-associated infections and a cause of mortality and morbility [1,2]. It is a Gram-positive pathogen, which affects humans and animals, causing different kinds of diseases, ranging from minor skin to life threatening infections such as pneumonia, osteomyelitis, acute endocarditis, meningitis and abscesses. The prognosis of these infections is dramatically worsened when the etiological agent is resistant to methicillin [3]. For these reasons, there is currently a huge demand for rapid methods to detect *S. aureus* in biological samples. Conventional methods to detect *S. aureus*, such as cell cultures, require more than one day and cannot distinguish between strains belonging to other species such as *S. intermedius*, *S. caprae*, *S. simulans* and *S. capitis* [4]. Rapid *S. aureus* agglutination tests have been developed as an alternative for routine diagnosis, but their accuracy, specificity and predictive capacity have been questioned [5,6,7]. The advent of real-time PCR led to the development of rapid methods (few hours), not requiring the bacteria isolation [8,9,10], but this technique requires expensive equipment and trained personnel, hence is not suitable for bedside point-of-care use. MALDI-TOF MS-based identification of bacteria is a fast and accurate technology [11,12,13,14] and could be an alternative to molecular tests if the test accuracy is proven [15], but requires an expensive instrument as well. The future of bacterial detection seems to reside in the development of biosensors, which are devices where a biological

component, such as proteins or oligonucleotides, is coupled with a transducer for obtaining a readable signal. For this application, the research is focused on designing novel and user-friendly devices, allowing the rapid detection of bacteria without the need of expensive and fixed equipments. The most common recognition tools in the development of detection systems are: antibodies [16], bacteriophages [17], phage display peptides and phage's receptor binding proteins and nucleic acids. The last case is mainly represented by a novel and highly performing approach for diagnostic applications, i.e. aptamers [18]; these are short single-stranded DNA or RNA sequences that fold into well-defined three-dimensional structures forming pockets and clefts for the tight and specific binding of any given molecular target. Therefore, aptamers represent tailored nucleic acids segments acting as artificial recognition elements that are able to recognize conserved epitopes on the surface of a bacterium. Aptamers libraries are continuously growing thanks to the recent progress in the aptamer selection procedure. Moreover, thanks to the relatively easy prediction of their secondary structure, the possibility to modulate affinity towards a target by a rational modification of the sequence has been recently demonstrated [19]. Aptamers interacting with *S. aureus* are already known [20,21,22,23] and hence a set of the molecular recognition probes are available, but the development of a proper transducer for a biosensor is still calling for an intensive research in the field. One of the technological challenges to the use of aptamers-based biosensors in the clinical practice is the lack of materials allowing their high-density immobilization while maintaining their binding properties. Nanostructured surfaces are very promising for this kind of applications [24,25,26,27] because they can favor the high density immobilization of biorecognition elements while retaining their structure and functionality. In particular, it has been demonstrated that the surface nanoscale morphology promotes protein adsorption [23,28]. Thus, the capability of producing nanostructured substrates having well controlled and scalable roughness represents a unique asset for the implementation of aptamers-based devices suitable for point-of-care diagnostics.

Taking advantage of this knowledge, a new device was proposed based on the conjugation of three building blocks: aptamers, nanostructured surfaces and peptides labeled with new solvatochromic probes (Figure 1).

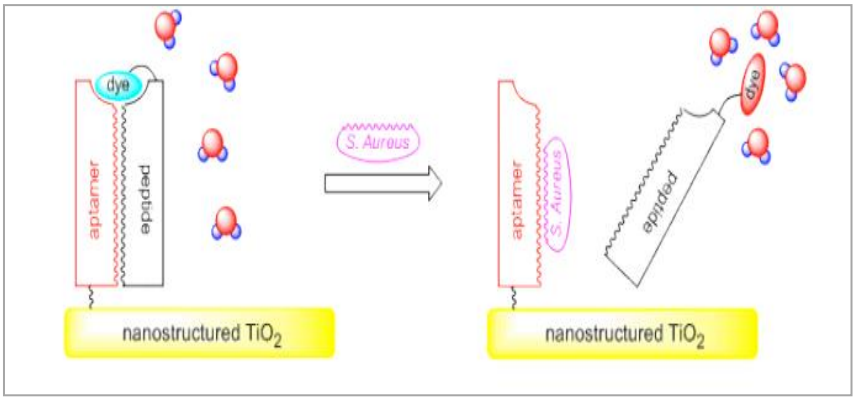


Figure 1 *Schematic representation of the proposed user-friendly device for the detection of S. aureus in biological samples.*

The concept of the device is to immobilize, on a nanostructured support, aptamers hybridized with labeled peptides that could be displaced in the presence of *S. aureus*, inducing a colorimetric switch of the probe in the visible region. As a surface for aptamers immobilization a nanostructured support was chosen for a high density derivatization. The rationale for the selection of peptides able to form specific complexes with aptamers, is based on the concept of 'energetic amino acid-base recognition code' [29]. The code was achieved by determining, in a dataset consisting of 100 functionally heterogeneous high resolution protein-DNA structures, the frequency and the energy of interaction between each amino acid and base. Based on the selected aptamer(s), proteins able to interact with nucleic acids will be screened and mutated accordingly to the recognition code.

Research purpose

Staphylococcus aureus (*S. aureus*) is a Gram-positive pathogen, which affects humans and animals, causing different kind of diseases, ranging from minor skin to life threatening infections. The prognosis of these infections is dramatically worsened when the etiological agent is resistant to methicillin [3]. Currently, the most important detection methods are classical cultures and biochemical tests. However, they are time-consuming and not highly sensitive. Therefore, there is currently a huge demand for rapid methods to detect *S. aureus* in biological samples.

This project is aimed at the development of an innovative user-friendly device for the detection of *S.aureus* that does not require any expensive detection system. The project aims at the development of a device consisting of three elements: an aptamer sequence, a nanostructured support, and a peptide labeled with a solvatochromic dye. The idea is to immobilize, on the nanostructure support, aptamers hybridized with labeled peptides, that can be displaced in the presence of *S. aureus* inducing a colorimetric switch in the visible region.

Within this multicenter project, our role is i) to characterize the spectroscopic properties of the synthesized polar-sensitive dyes, choosing the best candidate to be used in the colorimetric detection system and, more generally, for biological applications; ii) to characterize, using circular dichroism and fluorescence spectroscopy, the interaction between the selected aptamer(s) and the peptides that are designed and synthesized. Finally, the resulting construct will be tested for its capability to give a naked-eyes detectable signal as a consequence of the interaction with the aptamer.

Materials and methods

1. Spectroscopic characterization of the MediaChrom polar-sensitive dyes

The small library of six polarity-sensitive fluorescent dyes, nicknamed MediaChrom, was designed and synthesized by the DISFARM-Section of General and Organic Chemistry "A. Marchesini", University of Milan [30]. This class of dyes is characterized by a pyrimidoindolone core fitted out with a conjugated push-pull system and a carboxy linker for a conceivable coupling with biomolecules (Figure 2).

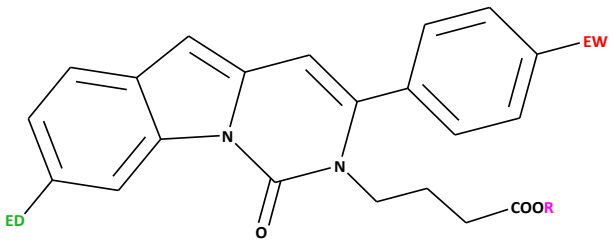


Figure 2 *General formula of MediaChrom dyes (ED= electron-donating group, EW= electron-withdrawing group, R= Bn, H).*

The electron-donating (ED) and the electron-withdrawing (EW) groups were selected based on literature data. The most used ED groups in push–pull solvatochromic dyes are secondary amines and in this study diethylamino and morpholine groups were used. Different EW groups are present in most effective polarity sensitive dyes, ranging from carbonyl to sulfonic, nitro, and cyanogroups. The effect of the latter three groups was investigated for their strong mesomeric EW effect, and the trifluoromethyl group was also studied, being a strong inductive EW group.

In table 1 the combination of ED and EW groups for each MediaChrom dye was reported.

Table 1 *ED and EW substituent combination for MediaChrom dyes.*

MediaChrom	ED substituent	EW substituent
15 a	Et_2N-	-SO_2-Me
15 b	Et_2N-	-NO_2
15 c	Et_2N-	-CF_3
15 d	Et_2N-	-CN
15 e	morpholine	-SO_2-Me
15 f	morpholine	-CF_3

1.2 Absorption spectroscopy

Absorption spectra were collected using a Cary 4000 UV-Vis spectrophotometer. MediaChrom stock solutions (5 mg/mL) were prepared by dissolving lyophilized powders in DMSO, while Prodan stock solution was prepared in ethanol. To collect absorption spectra, 2 µL of stock solutions were layered in a vial, and DMSO or ethanol was evaporated by a vacuum concentrator. 500 mL of different solvents were then added and transferred to the cuvette to obtain a final dye concentration of 0.02 mg/mL. The cuvette was kept protected from light. Extinction coefficients were determined by acquiring absorption spectra at four different concentrations(MediaChrom 15c from 139 to 13.9 µM, Prodan from 100 to 10µM) in solvents with different polarity (hexane, n-octanol, ethanol, DMF).

1.3 Fluorescence spectroscopy

Fluorescence spectra were collected using a FluoroMax-3 spectrofluorimeter. Dye stock solution was diluted in ethanol to obtain a 0.2 mg/mL solution that was further diluted in ethanol at different concentrations (from 2.5 to 5 µM). The fluorescence intensity values were recorded by exciting the samples at a 380 nm wavelength. Fluorescence quantum yields were determined by taking Prodan in ethanol (QY = 71%) as a reference, using an excitation wavelength of 380 nm. The quantum yield values were corrected for the solvent refractive index.

1.4 Fluorophore characterization

Onsager cavity radii were calculated for MediaChrom dyes and Prodan by applying the following equation [31][30]:

$$a = \sqrt[3]{\left(\frac{3M}{4\pi\delta N_A}\right)} \quad (1)$$

where M is the molecular weight of the fluorophore, NA is Avogadro's number, and δ is the compound density. Dipole moment changes upon excitation were assessed through the general solvent effects described by the Lippert-Mataga equation (equation 2) [32]:

$$\overline{\upsilon_a} - \overline{\upsilon_f} = \frac{2}{hc}\left(\frac{\varepsilon - 1}{2\varepsilon + 1} - \frac{n^2 - 1}{2n^2 + 1}\right)\frac{(\mu^* - \mu)^2}{a^3} + const \quad (2)$$

where υ_a and υ_f are the wavenumbers in cm^{-1} of absorption and emission peaks, h is Planck's constant, c is the speed of light, a is the Onsager cavity radius, and μ* and μ are the dipole moment of the molecule in the excited and ground state, respectively.

1.5 Photodegradation test

In the photodegradation tests, a 5 μM solution in ethanol of a given dye in a quartz cuvette (path length 0.5 cm) was illuminated at 380 nm by the light of a xenon lamp of a spectrofluorometer (excitation slits = 8 nm, emission slits = 1 nm). During the time of illumination (100 min), the fluorescence signal was recorded as a function of time. A THORLABS power meter (PM100USB) was used to measure the energy on the sample, corresponding to 1.01 mW on an area of 15 mm^2 for this experimental setup.

1.6 Dipalmitoylphosphatidylcholine (DPPC) vesicle preparation

DPPC vesicles were prepared by sonication for 5 min of a 1 mg/mL DPPC solution in double distilled water.

2. Peptides-aptamers interactions studies

The aptamers oligonucleotide sequences were purchased in lyophilized form from Sigma Aldrich and stock solutions were prepared using sterile water. The interaction studies were carried out both on SA23 aptamer and on the shorter versions, SA23 short 1 and SA23 short 2, in which the double strand region used as a scaffold for peptides design is present (Figure 3, Results and discussion); as a control, a mismatching sequence of SA23 short 2, named SFAV-IIA, was specifically designed to have no interaction with the synthetic peptides.

The sequences of the ʎ-Cro peptide mutants designed and synthesized to interact with the aptamers are reported in Table 2.

Table 2 *ʎ-Cro peptide mutants sequences*

Peptides	Sequences
IA	Ac-GQTKTAKDLGVYKSAIEEAIHAG
IAser	Ac-GQTKTAKDLGVYKDAIEEAIHAG
IIb2	Ac-GQTKTAKDLGVYDSAIEEAIHAG
IIA	Ac-GQTKTAKDLGVYESAIEEAIHAG
IIA2	Ac-GQTKTAKDLGVYEDAIEKAIHAG
IIA3	Ac-GQTKTAKDLGVYEDAIEFAIHAG

The peptides were designed by the molecular modeling laboratory, Food Science Dept., University of Parma and synthesized by the DISFARM, University of Milan.

2.1 Circular dichroism studies

Circular dichroism experiments were carried out on a Jasco J-715 spectropolarimeter equipped with a Peltier thermostatic unit set at 20°C. Secondary structures analysis were performed by using Dichroweb server (CDSSTR algorithm-reference set 7).

2.2 Fluorescence analysis

Fluorescence spectra of each peptide-aptamer complex were recorded on a Jasco FP-777 spectrofluorimeter. The SA23 aptamer tagged with carboxyfluorescein was titrated with increasing amounts of peptide. The peptide-aptamer mixture was incubated at room temperature for 15 min to allow the formation of the complexes before the registration of each fluorescence spectrum. An excitation wavelength of 492 nm was used.

The fluorescence emission spectrum of peptide IAser labeled with MediaChrom 15c in the presence and absence of SA23 was collected at a concentration of 50 µM. An excitation wavelength of 380 nm was used.

Results and discussion

As previously mentioned, the proposed device is characterized by four building blocks: aptamers, peptides, new synthesis solvathocromic dyes and a nanostructured support.

Aptamers specific for *S. aureus* have been selected from literature [21]; the five aptamers presented by Cao and coworkers [21] were taken under consideration for selecting the best candidate for *in silico* design of the peptide sequence able to interact. SA23 was chosen among these aptamers mainly because of a structural peculiarity: based on the predicted secondary structure, it presents a relatively long trait in double strand configuration (the hairpin made by nucleotides 16-44), with high GC content in its inner portion (Figure 3A).

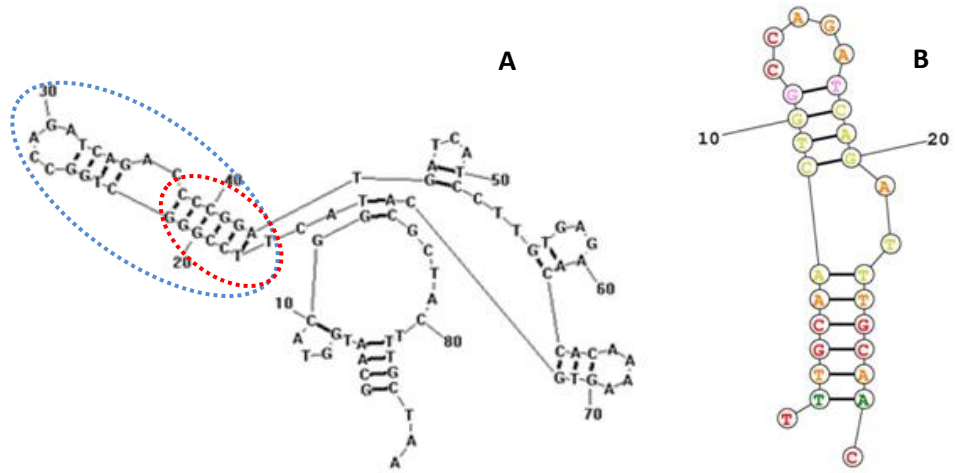


Figure 3 *Panel A: Representation of SA23 aptamer folded into its classical three-dimensional structure and the shorter version of SA23 are highlighted in light blue (SA23 short 2) and in red (SA23 short1); Panel B: mismatching sequence of SA23 short 2 (SFAV-IIA) designed to have no interaction with the synthetic peptides.*

This double strand DNA region of the aptamer was then selected as the binding region of the peptides. Moreover, the SA23 aptamer results as the one with the higher affinity within this group [21]. The rational design for the selection of peptides able to form

specific complexes with aptamers, is based on the concept of 'energetic amino acid-base recognition code' [29]. The code was achieved by determining, in a dataset consisting of 100 functionally heterogeneous high resolution protein-DNA structures, the frequency and the energy of interaction between each amino acid and base. Based on the selected aptamer(s), proteins able to interact with nucleic acids will be screened and mutated accordingly to the recognition code. Cro was selected as a scaffold for the design of the aptamer-interacting peptides. Cro is a 66 aa protein which plays a pivotal role in the switch from lysogenic to lytic phase in the growth cycle of phage λ and represents a suitable scaffold for several reasons: i) the availability of three-dimensional structures both in the presence and absence of its cognate DNA [33] allows detailed structural evaluations; ii) the helix-turn-helix motif of the DNA binding domain is relatively small and all the interactions between the nucleobases of the consensus sequence and the peptidic backbone are well characterized; iii) the minimum functional portion of the consensus sequence is quite short and made by contiguous nucleobases along the two strands of the tract of the DNA target; iv) previous works reported successful examples of Cro reprogramming in order to bind consensus sequences that differ from the wild type [34]. Moreover, the helix-turn-helix peptide domain of the Cro protein showed to well mimic full-length protein for binding [35]. Based on these observations, the sequence of the helix-turn-helix motif of Cro was mutated, accordingly to the previously reported recognition code, in order to bind the double strand region of SA23, and a library of six peptides were synthesized using stepwise solid phase synthesis (SPPS) (see Materials and Methods). For the colorimetric detection, a new class of six polarity-sensitive fluorescent dyes, nicknamed MediaChrom, was designed and synthesized by the DISFARM-Section of General and Organic Chemistry "A. Marchesini", University of Milan, for a conceivable coupling with the aforementioned peptides. Finally as a surface for aptamers immobilization a nanostructured zirconium dioxide support was chosen for a high density derivatization and to test bacteria-aptamer interactions using microarray techniques.

1. Photophysical evaluation of MediaChrom dyes

The spectroscopic properties of MediaChrom dyes were firstly characterized by collecting absorption spectra (Table 3). For comparison, the commercially available polarity-sensitive dye Prodan was chosen [36].

Table 3 *Onsanger radius and absorption properties of MediaChrom dyes 15a-f compared to Prodan (20 µg/mL, ethanol)*

dye	Osanger radius (Å)	$\lambda_{absorption}$[a](nm)	ε (mM^{-1} cm^{-1})
15a	5.7162	401	8.5
15b[b]	5.5919	456	7.6
15c	5.6943	393	13.8
15d	5.6234	409	15.4
15e[c]	5.6323	369	17.1
15f	5.6100	377	17.5
Prodan	4.3628[d]	360[d]	18.4[d]

[a]As the wavelength of the most red-shifted peak. [b]C = 0.5 mg/mL. [c] Poor solubility. [d]From ref. [37]

All MediaChrom dyes showed intense absorption peaks, spanning from 369 to 456 nm. MediaChrom 15b, the dye with the most red-shifted absorption band, exhibits lower absorption intensity. In order to obtain acceptable quality spectra, a 25-fold more concentrated solution was prepared. The presence of the stronger EW group (nitro group) is responsible of the marked red shift, typically observed when the conjugation of the chromophoric system is increased. The comparison of the spectra of MediaChrom 15a and 15c, characterized by the presence of the same ED group (−NEt_2) but different EW groups (−SO2Me vs −CF3), indicates that the presence of a mesyl resulted in a 8 nm red-shifted absorption peak, with respect to the trifluoromethyl group. A similar effect was observed by comparison of the absorption peaks of MediaChrom 15e and 15f, where the same ED group (morpholine) was introduced. In these dyes, the presence of the morpholine substituent causes a blue shift of the absorption peaks. MediaChrom 15d seemed to be the most promising fluorophore for imaging applications since an

interesting red-shifted absorption peak is accompanied by good absorption intensity (Table 3). In fact, the availability of a dye that can be excited in the visible range with many cheap and accessible light excitation sources represents a remarkable benefit. Next, all dyes were characterized for their fluorescence solvatochromic properties by collecting fluorescence spectra at20 °C in solvents at different polarities, such as hexane, n-octanol, ethanol, and DMF (Table 4).

Table 4 *Solvatchromic properties of MediaChrom dyes 15a-f*

			λ emission	**(nm)[a]**		
solvent	**15a**	**15b**	**15c**	**15d**	**15e**	**15f**
hexane	525	594	490	525	512	482
n-octanol	585	-	525	585	568	514
ethanol	595	-	540	595	582	528
DMF	605	-	565	615	597	553

[a] Excitation wavelength.

All MediaChrom dyes exhibited fluorescence emissions in all solvents with a solvatochromic shift spanning over 90 nm from hexane to DMF. The only exception is 15b, which showed significant fluorescence only in hexane (Table 4). This high sensitivity to solvent polarity is comparable to that observed for many commercially available solvatochromic probes (e.g. Prodan, Laurdan, Dansyl). For each MediaChrom dye, the integration of the emission peak in different solvents was calculated and normalized to the peak with the maximum intensity to evaluate the relative intensity of fluorescence emission in solvents with different polarity (Figure 4).

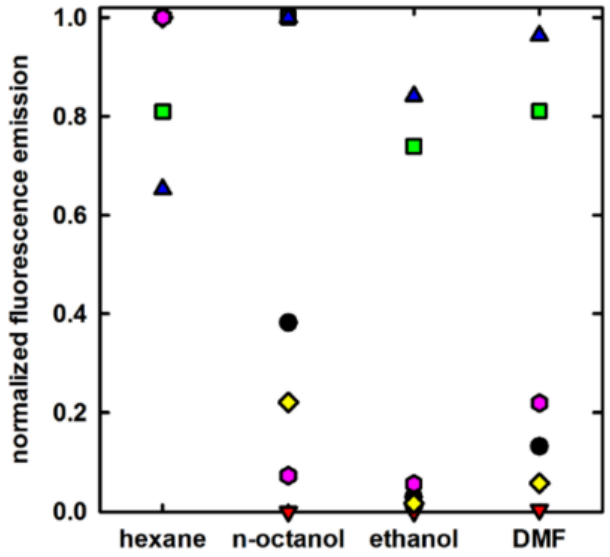


Figure 4 *Normalized fluorescence emission (integrated peak area) of MediaChrom dyes at 1.6 µg/mL concentration in different solvents at 20 °C: 15a (black circles), 15b (red triangles), 15c (green squares), 15d (yellow diamonds), 15e (magenta hexagon), and 15f (blue triangles).*

Figure 4 shows that only MediaChrom 15c and 15f preserve an almost constant fluorescence emission in all tested solvents over a wide range of polarity, together with a simultaneous large change in the Stokes shift. These properties make both of these molecules good candidates for several purposes, ranging from structural and dynamic studies on biological macromolecules to cell imaging applications. When absorption and emission properties are compared, 15c is considered the best candidate because it requires a red-shifted excitation with respect to 15f (Table 3). In fact, in applications such as fluorescent labeling for user-friendly detection kits or dyes for cell imaging, the possibility to work at longer wavelength guarantees a reduced biological tissue autoflorescence and scattering and/or decreased cell damage. For these reasons, the following analysis was focused on MediaChrom 15c, and its properties were compared to the properties of the commercial solvatochromic probe Prodan. The polarity-sensitive properties of 15c were investigated by collecting absorbance and fluorescence spectra in solvents at different polarities. Fluorescence quantum yields were determined by taking

Prodan in ethanol (quantum yield, QY = 71%) as a reference [37]. For spectroscopic measurements, solutions of the dyes at different concentrations (from 2.5 to 5 µM) were both excited at 380 nm, an average excitation wavelength between the absorbance λ_{max} of Prodan and MediaChrom 15c.

In Table 5 absorbance and fluorescence parameters are reported for both MediaChrom 15 c and Prodan.

Table 5 *Comparison of the absorbance and fluorescence properties of 15c and Prodan in four solvents with different polarities*

	λ_{max}(abs), nm		ε ($mMcm^{-1}$)[a]		λ_{max}(fluor), nm[b]		QY (%)[c]	
solvent	**Prodan**	**15c**	**Prodan**	**15c**	**Prodan**	**15c**	**Prodan**	**15c**
hexane	346	396	36.8	14.6	412	480	60.4	51.5
n-octanol	362	397	33.6	14.0	473	522	59.9	52.4
ethanol	368	393	19.9	13.8	487	540	71.0	39.9
DMF	355	396	39.3	13.7	455	562	64.2	36.9

[a]Absorption spectra acquired at four different concentrations (15c from 139 to 13.9 µM, Prodan from 100 to 10 µM). [b]Excitation wavelength: 380 nm, [dye] = from 2.5 to 5 µM. [c]Quantum yield values were corrected for the solvent refractive index.

The absorption spectra of 15c in different solvents did not show significant differences (Figure 5A). In all investigated solvents, the most red-shifted absorption peak lies around 395 ± 2 nm. This property guarantees that 15c can be excited at the same excitation wavelength (for example, with a common 405 nm diode laser), irrespective of the polarity of the solvent.

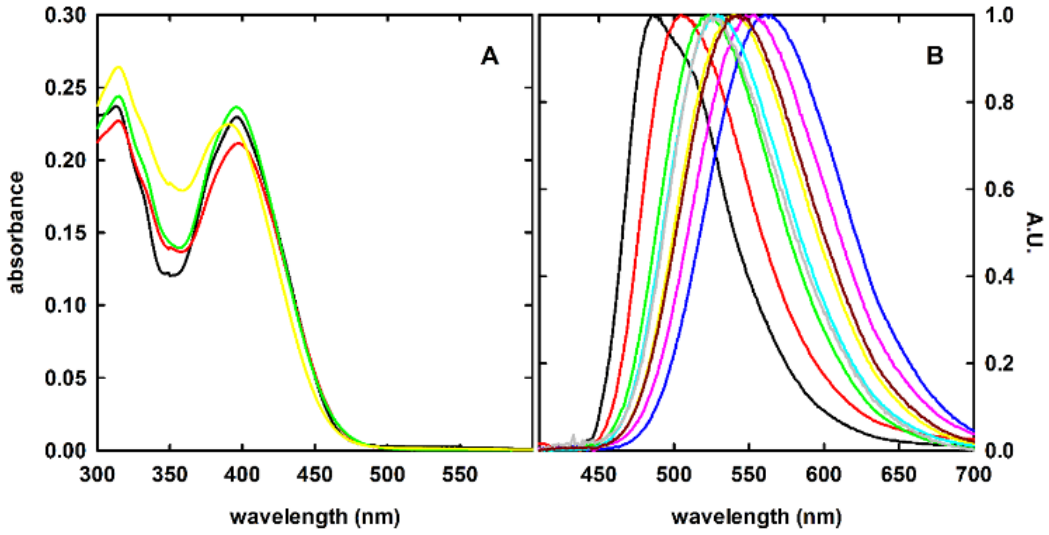


Figure 5 *Absorption (A) and normalized fluorescence (B) spectra of 15c in different solvents at 20 °C. (A) 15c absorbance spectra in hexane (black line), n-octanol (red line), DMF (green line), and methanol (yellow line). (B) 15c normalized fluorescence emission spectra (excitation wavelength = 380 nm) in hexane (black line), triethylamine (red line), n-octanol (green line), chloroform (gray line), n-butanol (light blue line), ethanol (yellow line), methanol (brown line), acetone (magenta line), and DMF (blue line).*

Conversely, Prodan displays a pronounced difference in the maximum of absorbance depending on the nature of the solvent (from 346 to 368 nm), together with an unfavorable blue-shifted absorption maximum (ranging from −25 nm in ethanol to −50nm in hexane, compared to 15c). The absorption coefficients (ε) of 15c in the four solvents tested are from 13.7 to 14.6 $mM^{-1}cm^{-1}$. In all solvents (except for ethanol), they are lower than those observed for Prodan. The fluorescence emission spectra of MediaChrom 15c in solvents with different polarities show that the dye exhibits a Stokes shift from 84 nm in hexane to 166 nm in DMF (Table 5 and Figure 5B). This large solvatochromic shift (82 nm) suggests that the chromophore is able to detect even small polarity changes. Moreover, the emission maxima are also strongly red-shifted with respect of those of Prodan. Despite being lower than those of Prodan, the quantum yields of MediaChrom 15c in all examined solvents are good.

To describe quantitatively the effects of the physical properties of the solvent on the fluorescent emission spectra of the dye, the Lippert-Mataga equation was used (equation 2, Materials and methods) [32].This correlation is based on the assumption that the solvent is a continuum in which the fluorophore is contained, and solvent-specific interactions are not considered. It can be approximated that the energy difference between the ground and the excited states is a property of the refractive index (n) and the dielectric constant (ε) of the solvents. In Figure 6, the orientational polarizability (Δf) is plotted against the Stokes shift (in cm^{-1}) for MediaChrom 15c and compared to Prodan under the same experimental conditions.

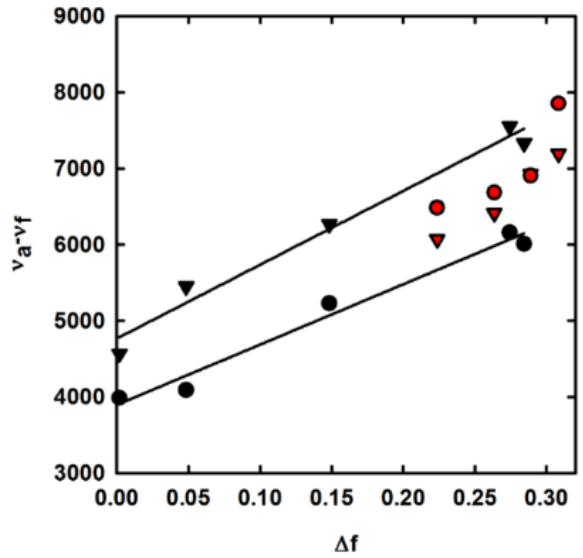


Figure 6 *Lippert plot for MediaChrom 15c (triangles) and Prodan (circles) in aprotic (black) and protic (red) solvents.*

For both fluorophores, a clear dependence of the Stokes shift on orientational polarizability of the solvent was observed. It is worth noting that the Stokes shifts for MediaChrom 15c in protic solvents are weaker than those for aprotic solvents of similar polarity (Figure 6, red triangles vs black triangles), while Prodan displays exactly the opposite trend (Figure 6, red circles vs black circles). These phenomena are usually

retraced to specific interactions of the fluorophore with the solvent. For example, in Prodan, this phenomenon was related to H-bonding of the protic solvents with the carbonyl group of the fluorophore [38]. The reverse effect observed in MediaChrom 15c could be related to a particular interaction with the solvent, probably due to the presence of the uncommon trifluoromethyl group in the D-π−A system. Since the Lippert-Mataga equation usually well describes fluorophore behavior in aprotic solvents, only data for both dyes in aprotic solvents were used in the fitting to determine the change in dipole moment upon excitation (μ*−μ). MediaChrom 15c and Prodan Stokes shifts showed a well defined dependence to orientational polarizability, considering all solvents. In particular, data obtained in aprotic solvents showed a good correlation, with a μ*−μ of 13.3 ± 0.6 and 8.1 ±0.4 D, respectively.

To estimate the photostability of MediaChrom 15c, a photodegradation test was performed (Figure 7).

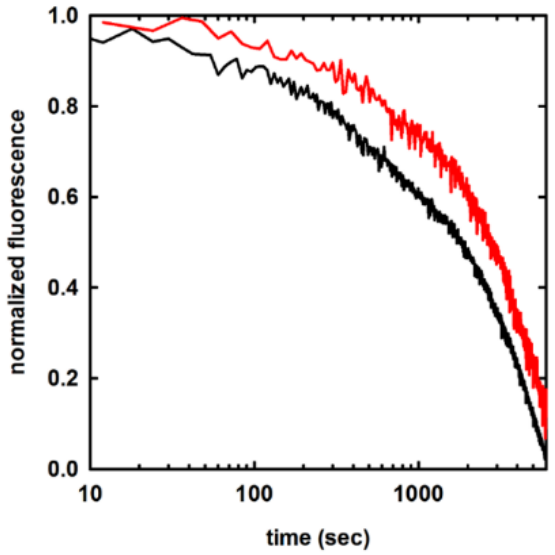


Figure 7 *Photodegradation test for MediaChrom 15c (red line) and Prodan (black line) in ethanol (c = 5 μM, excitation wavelength = 380 nm).*

Ethanol solutions of 15c and Prodan were prepared in two different quartz cuvettes and illuminated at 380 nm with a xenon lamp for 100 min, while the fluorescence signal was recorded as a function of time. The power density applied to the sample was 6.7 mW/cm^2. As shown in Figure 7, MediaChrom 15c decays more slowly than Prodan, showing a halved emission after 2800 seconds (1750 seconds for Prodan). Based on these results, the comparison between MediaChrom 15c and Prodan suggests that, in spite of exhibiting a slightly reduced quantum yield, the former possesses photochemical properties comparable to those of the commercial dye, with the advantage of a strong, red-shifted absorption and fluorescence emission and an increased photostability.

1.2 Behavior of MediaChrom 15c in Dipalmitoylphosphatidylcholine (DPPC) vesicles

Prodan dye is widely used for probing the membrane state, including transition phases and oxidation processes of phospholipid systems, due to its capability of penetrating in the lipophilic layer and sensing polarity changes [38,39]. In order to evaluate the potential of MediaChrom 15c as a membrane probe, we collected fluorescence spectra of both Prodan and MediaChrom 15c after their addition to a DPPC vesicles suspension at different temperatures. It has already been observed that an increase in temperature causes a change in the fluorescence emission spectrum of Prodan. This is due to the transition of the phospholipid system from the gel to the liquid phase that allows more water molecules to enter into the lipid phase [40]. The fluorescence spectra of MediaChrom 15c embedded in DPPC vesicles showed an emission peak centered at 528 nm (Figure 8), very similar to that observed when it is dissolved in n-octanol (525 nm) at the same temperature (20 °C).

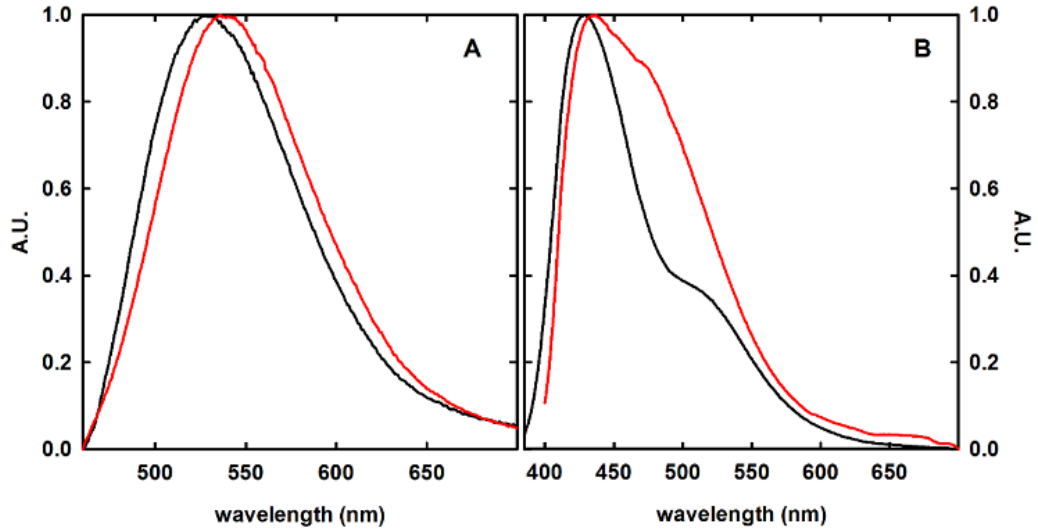


Figure 8 *Fluorescence spectra of 1.6 µg/mL MediaChrom 15c (A) and Prodan (B) in DPPC vesicles at different temperatures: 20 °C (black lines) and 37 °C (red lines). Excitation wavelength = 380 nm.*

This result suggests that the fluorophore lies in the lipid phase of DPPC vesicles, probing this region. The comparison of the behavior of MediaChrom 15c and Prodan clearly demonstrates that the former is able to sense the transition phase of phospholipid systems similarly to Prodan, with the advantage of an emission wavelength red-shifted by 100 nm, significantly reducing scattering effects that are considerably high in micellar suspensions as well as when cellular membranes are present.

1.3 MediaChrom 15' as a peptide label

MediaChrom dyes 15 were designed with a benzyl ester to be easily deprotected, transforming them into carboxy-free MediaChrom labels 15'. This transformation makes these dyes suitable for an easy conjugation with proteins, peptides, or an amino-modified oligonucleotide. These labeled bioconjugates could thus be used for the detection of interactions involving a change of polarity of the system, such as protein/DNA binding. As a proof of concept, we selected the protein Cro, whose interaction with its consensus DNA sequence (OR3) is known and widely investigated [33,41]. In particular, Cro is a 66

amino acids dimeric protein that plays a key role in the switch from the lysogenic to the lytic cycle in bacteriophage λ. Its interaction with DNA is mainly restricted to a small helix-turn-helix motif spanning from residue 15 to 38, with a K_d of around 25 µM [34][35]. The wild-type DNA binding sequence (GQTKTAKDLGVYQSAINKAIHAG) was prepared by microwave-assisted solid-phase synthesis [35] and labeled at the N-terminus with the MediaChrom 15'c using standard protocols [42]. Furthermore the peptide Cro has been selected as model for the design of the aptamer-interacting peptides for the development of the proposed device. The fluorescence spectra of 15'c-labeled Cro:1 in the presence and in the absence of OR3 consensus DNA sequence were recorded (Figure 9).

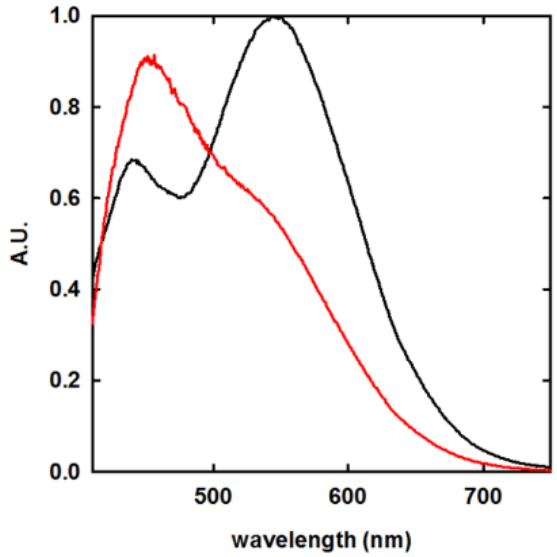


Figure 9 *Normalized fluorescence spectra of wt DNA-binding peptide of Cro in the presence (red line) and in the absence (black line) of OR3 consensus sequence (50 µM 15'c-Cro:1, 50 µM OR3, PBS buffer, pH 7.4, 20 °C). Excitation wavelength = 380 nm*.

In the absence of a cognate DNA sequence, the labeled peptide 15'c-Cro:1 displayed a maximum of fluorescence at 545 nm (Figure 9, black line). Upon addition of the OR3 consensus sequence, a strong hypsochromic shift (93 nm) was observed (Figure 9, red line), thus indicating that the MediaChrom 15'c sensitively detects the change in the

polarity of the environment when the peptide interacts with the DNA sequence. The minor peak present in the 15'c-Cro:1 spectrum and the shoulder in the peak of the 15'c-Cro:1-OR3 complex suggest that there is an equilibrium of two discrete peptide conformations, clearly signaled by the dye. The binding of OR3 consensus sequence shifts the peptide conformational equilibrium from one state to the other.

2. Peptides-aptamers interaction studies

Circular dichroism (CD) spectroscopy has been carried out on λ-Cro peptide mutants optimized to interact with the double strand portion of the SA23 aptamer in 20 mM phosphate buffer, pH 7.4. Circular dichroism has been used to characterize the secondary structure of synthesized peptides, and hence the effect of mutation on the stability of the structure, by collecting far-UV CD spectra (Figure 10).

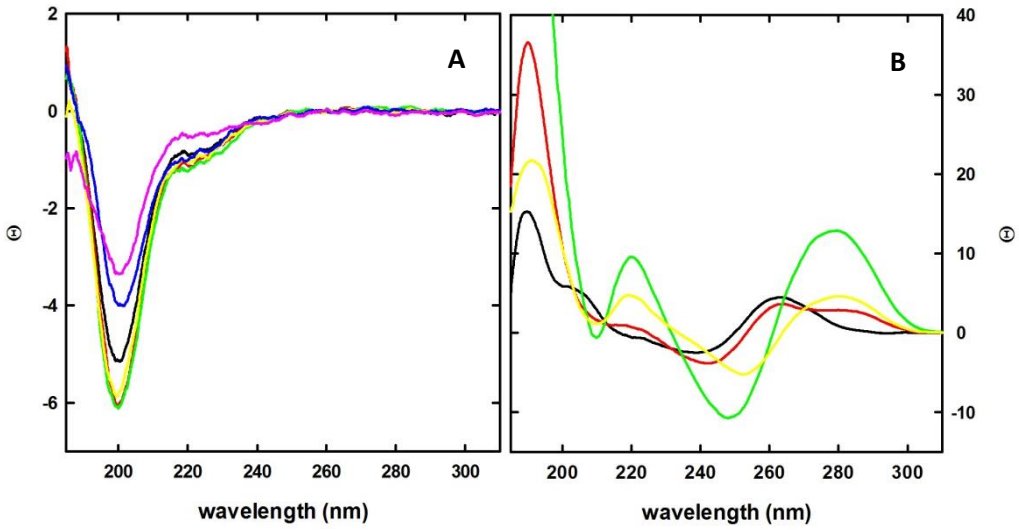


Figure 10 *Panel A: CD spectra of 20 μM peptides solutions in 20mM phosphate buffer, pH 7.4: peptide IA (black line), peptide IAser (red line), peptide IB2 (green line), peptide IIA (yellow line), peptide IIA2 (blue line), peptide IIA3 (purple line). Panel B: CD spectra of 20 μM aptamers in 20 mM phosphate buffer, pH 7.4: SA23 short 1 (black line), SA23 short2 (red line), SA23 (green line), SFAV-IIA (yellow line).*

It has already been demonstrated that it is possible to mimic the DNA binding behavior of λ-Cro by shorter, chemically synthesized peptides representing the helix-turn-helix region of the protein [35]. These peptides, also in dimeric form, showed significant helical content only when α-amino isobutyric acid is introduced in the sequence [35]. To quantify the λ-Cro peptide mutants helical content, the spectra of the peptides have been analyzed using Dichroweb server (CDSSTR algorithm-reference set 7).

Table 6 *Dichroweb analysis of peptides CD spectra (CDSSTR algorithm-reference set 7)*

Peptide	Helices (%)	Strands (%)	Turns (%)	Unordered (%)
IA	17%	28%	18%	36%
IAser	21%	24%	20%	35%
IB2	21%	26%	21%	32%
IIA	19%	24%	20%	37%
IIA2	7%	32%	23%	37%
IIA3	5%	36%	17%	40%

The spectra showed a negative band centered at 200 nm, a feature typical of random coil folding (Figure 10A), as confirmed by spectra deconvolution (Table 6) yielding a low percentage of helical content. This suggests that the isolated peptides in solution are unstructured in the absence of the target nucleotides sequence. To verify that, even at higher concentration, the peptides do not tend to form aggregates or oligomeric species with different structural features, CD spectra have been collected as a function of peptide concentration (10-100 μM). The CD spectra in this range of concentration appeared identical. Peptides solutions demonstrated to be stable even after one or two freeze-thawing cycles. CD spectra of the SA23 aptamer and its shorter versions containing the double strand interacting region, showed well-defined bands, demonstrating that they have a significant content of ordered secondary structure in solution (Figure 10B).

2.1 Circular dichroism studies on peptides-aptamers interactions

The peptide-aptamer interaction has been firstly studied by near-UV CD spectroscopy. In this region, only aptamers (DNA) contribute to the spectra. The analysis of this spectral region allowed to observe the effect of the peptides on aptamers structure without any interferences from the conformational changes of the peptides.

Solutions containing aptamers and peptides at the same concentration (20 µM) were incubated before recording the spectra. A comparison between the spectra of the aptamer-peptide mixture and the arithmetic sum of the two separated components would highlight potential interactions. The collected data showed that the aptamer SA23 interacts with four peptides, promoting a destructuring process associated with a decrease of CD bands intensity (Figure 11A).

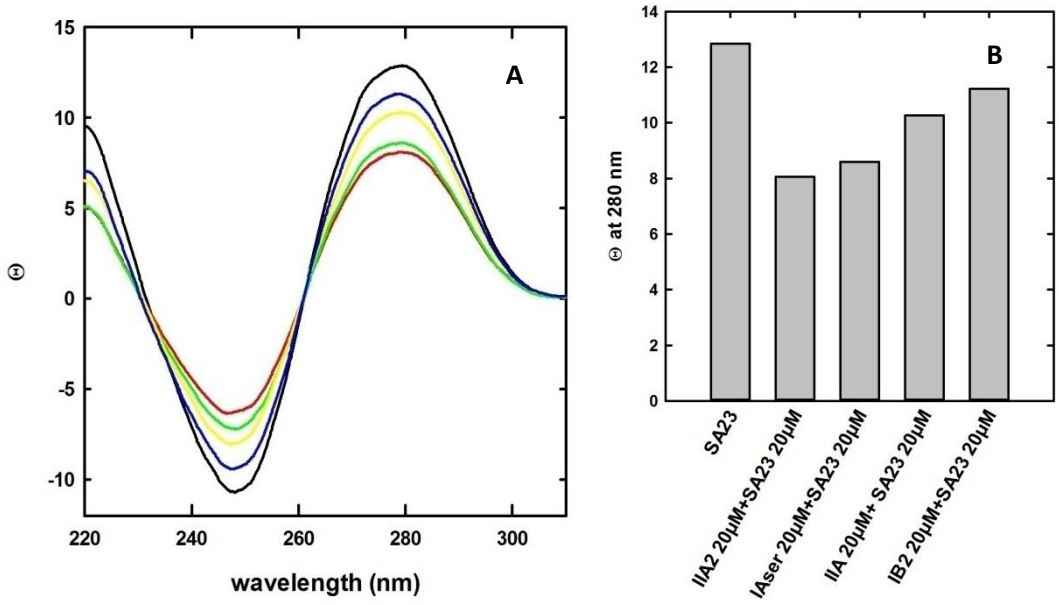


Figure 11 *Panel A: Near UV CD spectra of SA23 in the absence and presence of peptides. Spectra were recorded for solutions containing 20 mM phosphate buffer, pH 7.4 at a 20 µM concentration of aptamer and peptides-aptamer mixtures: SA23 (black line), SA23-peptide IIA2 mixture (red line), SA23-peptide IAser mixture (green line), SA23-peptide IIA mixture (yellow line), SA23-peptide IB2 mixture (blue line). Panel B: Ellipticity at 280 nm of SA23 in the absence and presence of interacting peptides.*

Comparing the ellipticity at 280 nm of SA23 alone and in mixture with the interacting peptides, the most evident effect is observed for peptides IAser and IIA2 (Figure 11B).

The CD spectra have been also collected at lower concentration (1 μM peptides and aptamer); IAser is the only peptide that still showed an effect, suggesting that it should be the one with the higher affinity.

The same studies have been performed for the shorter versions of SA23 aptamer, SA23 short 1 and SA23 short 2. The absence of differences between CD spectra of the aptamer-peptide mixtures and the arithmetic sum of the component spectra may derive from the fact that these sequences, shorter than SA23, upon peptide binding do not undergo conformational changes, and then the circular dichroism signal does not significantly change. For this reason, the absence of differences observed for the aptamer SFAV-IIA does not guarantee the absence of interaction.

In Figure 12, the comparison between the arithmetic sum of component spectra and the observed spectra of experimental mixtures is shown for the IAser peptide.

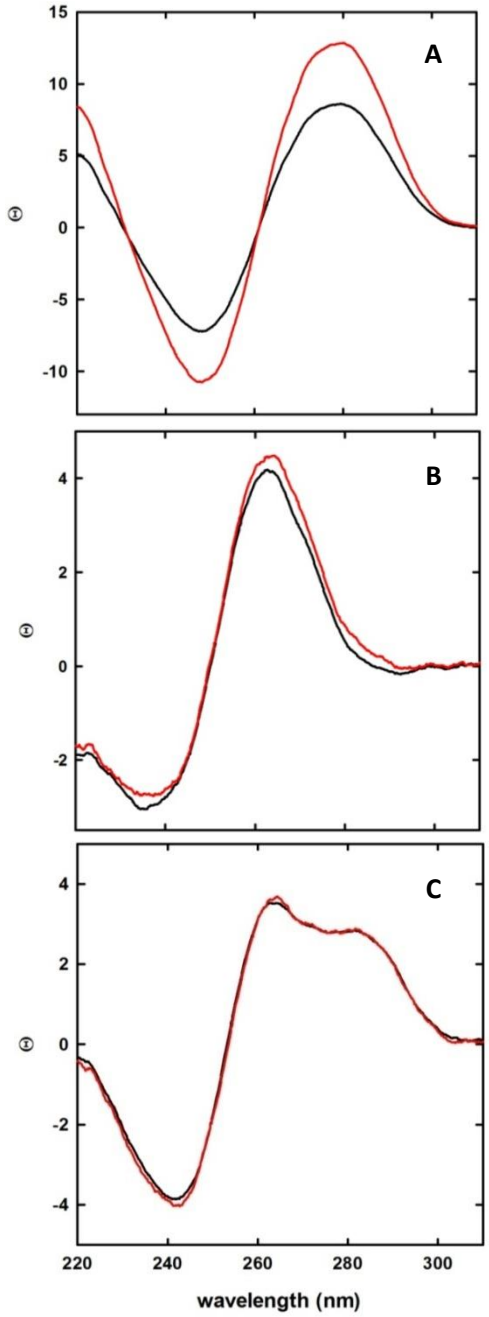


Figure 12 *Comparison between the arithmetic sum of component spectra (red line) and the spectrum of the experimental mixture (black line) for peptide IAser and SA23 (A), SA23 short 1 (B) and SA23 short 2 (C). CD spectra were recorded at a 20 μM concentration of each component.*

To demonstrate that peptide binding is determined by the specific mutation of λ-Cro peptide mutants, we synthesized and checked sequence-scrambled peptides, and a peptide where mutation causing unfavorable interaction were inserted. The comparison of the CD spectrum of the mixture of these peptides and SA23 with the arithmetic sum of the components spectra gave overlapping results, demonstrating the absence of interaction.

2.3 Fluorescence studies

Following the results of circular dichroism characterization, further analyses were carried out by fluorescence spectroscopy using the IAser peptide in order to calculate the dissociation constants (Kd) of the peptide-aptamer complex, which determination is prevented in circular dichroism measurements by the experimental limits of acquiring CD spectra at low concentrations.

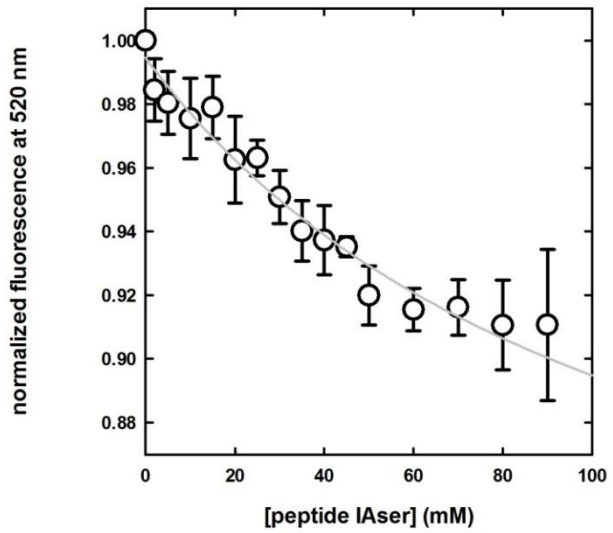


Figure 13 *Fluorescence emission at 520 nm upon excitation at 492 nm of SA23 aptamer as a function of IAser concentration; a K_d of 83 ± 30 µM is measured*

The SA23 aptamer tagged with the commercially available 3' carboxyfluorescein was titrated with increasing amounts of peptide using an excitation wavelength of 492 nm. The analysis of triplicate data allowed to obtain a K_d of 83 ± 30 µM.

The peptide IAser was also labeled with the most promising chromophore of the MediaChrom series, 15c', and tested for its interaction with the aptamer SA23. The fluorescence spectra of 15'c-labeled IAser in the presence and in the absence of SA23 are reported in Figure 14.

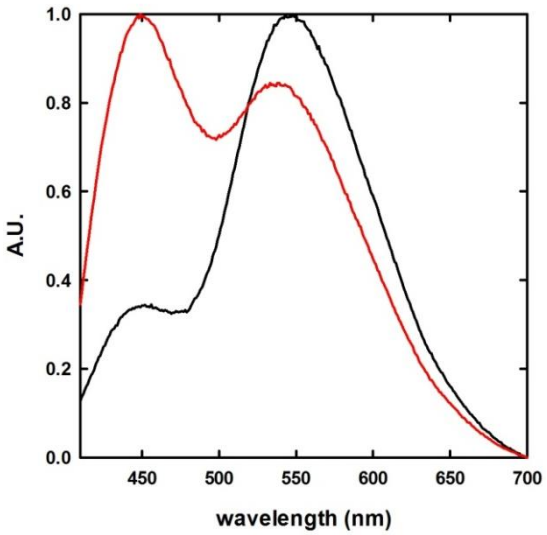


Figure 14 *Normalized fluorescence emission spectra of MediaChrom 15'c-labeled IAser peptide in the presence (red line) and in the absence (black line) of SA23 (50 µM MediaChrom 15'c-IAser, 50 µMSA23, 20 mM phosphate buffer, pH 7.4).*

In the absence of the cognate aptamer, the labeled 15'c-IAser peptide displayed a maximum of fluorescence emission at 548 nm (Figure 14, black line). When in mixture with SA23, a strong hypsochromic shift was observed (Figure 14, red line), indicating that the MediaChrom 15'c dye sensitively detects the change in the polarity of the environment when the peptide interacts with the aptamer.

As shown in the following picture (Figure 15), the illumination with a commercial ultraviolet LED (emitting at 400 nm) of solutions containing the IAser peptide labeled with MediaChrom 15'c in the absence and presence of the SA23 aptamer highlighted a different color of the emitted light detectable by naked eyes.

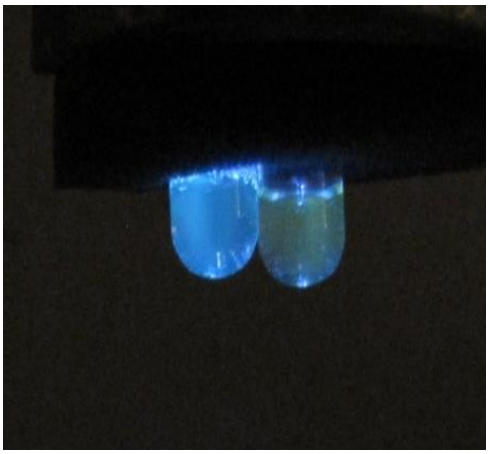


Figure 15 *IAser peptide labeled with MediaChrom 15'c in the absence (on the right) and presence (on the left) of SA23 aptamer under ultraviolet LED illumination*

Conclusions

The idea of the proposed device for the dection of *S. aureus* is to use aptamers immobilized on a nanostructured support and hybridized with labeled-peptides, that could be displaced by *S. aureus* producing an easily detectable signal in the visible region.
Within this project, we firstly characterized a small library of new polar sensitive dyes, nicknamed MediaChrom. MediaChrom 15c shows the best features for biological applications: a high absorption coefficient in the visible range that is almost constant in solvents with different polarities, a wide solvatochromic effect combined with an almost constant fluorescence emission, a good QY, and a noteworthy photostability if compared to the commercial solvatochromic fluorophore Prodan [30]. MediaChrom 15c was demonstrated to be a versatile solvatochromic dye that can be used as a membrane/lipophilic probe, and its parent carboxy-free MediaChrom 15'c was demonstrated to be a highly sensitive label for peptide tagging, useful for studying the interaction between peptides (or proteins) and other target biomolecules, such as DNA.
Circular dichroism and fluorescence studies have been carried out to test the interaction between SA23 aptamer and the computationally designed interacting peptides. The peptide IAser was the one that proved to be effective at lower concentration and was choosen as the best candidate for a biosensing device. Finally, once the most promising peptide and solvatochromic dye have been chosen, the resulting construct Mediachrom 15c'-IAser has been tested, with positive results, for its capability to give a naked-eyes detectable signal as a consequence of the interaction with the aptamer.

References

[1] J. S. Ferreira, W. L. R. Costa, E. S. Cerqueira, J. S. Carvalho, L. C. Oliveira, and R. C. C. Almeida, "Food handler-associated methicillin-resistant Staphylococcus aureus in public hospitals in Salvador, Brazil," *Food Control*, vol. 37, pp. 395–400, 2014.

[2] E. Török *et al.*, "Staphylococcal and streptococcal infections," *Medicine (Baltimore).*, vol. 33, no. 5, pp. 97–100, May 2005.

[3] B. N. Green, C. D. Johnson, J. T. Egan, M. Rosenthal, E. A. Griffith, and M. W. Evans, "Methicillin-resistant Staphylococcus aureus: an overview for manual therapists()," *J. Chiropr. Med.*, vol. 11, no. 1, pp. 64–76, Mar. 2012.

[4] K. Becker and C. von Eiff, "Staphylococcus, Micrococcus, and Other Catalase-Positive Cocci," *Manual of Clinical Microbiology, 10th Edition*, American Society of Microbiology, pp. 308–330, 2011.

[5] A. L. Truant and A. T. Evangelista, "Rapid Systems and Instruments for Antimicrobial Susceptibility Testing of Bacteria," in *Manual of Commercial Methods in Clinical Microbiology*, American Society of Microbiology, 2002, pp. 413–428.

[6] P. P. Bourbeau, A. L. Truant, and A. T. Evangelista, "Rapid Systems and Instruments for the Identification of Bacteria," in *Manual of Commercial Methods in Clinical Microbiology*, American Society of Microbiology, 2002, pp. 22–49.

[7] K. Weist, "Evaluation of six agglutination tests for Staphylococcus aureus identification depending upon local prevalence of meticillin-resistant S. aureus (MRSA)," *J. Med. Microbiol.*, vol. 55, no. 3, pp. 283–290, Mar. 2006.

[8] A. M. Frye *et al.*, "Clinical Impact of a Real-Time PCR Assay for Rapid Identification of Staphylococcal Bacteremia," *J. Clin. Microbiol.*, vol. 50, no. 1, pp. 127–133, Jan. 2012.

[9] L. R. Peterson, "Molecular Laboratory Tests for the Diagnosis of Respiratory Tract Infection Due to Staphylococcus aureus," *Clin. Infect. Dis.*, vol. 52, no. Supplement 4, pp. S361–S366, May 2011.

[10] H. -y. Wang, S. Kim, J. Kim, S.-D. Park, Y. Uh, and H. Lee, "Multiplex Real-Time PCR Assay for Rapid Detection of Methicillin-Resistant Staphylococci Directly from Positive Blood Cultures," *J. Clin. Microbiol.*, vol. 52, no. 6, pp. 1911–1920, Jun. 2014.

[11] F. Fleurbaaij, H. C. van Leeuwen, O. I. Klychnikov, E. J. Kuijper, and P. J. Hensbergen, "Mass Spectrometry in Clinical Microbiology and Infectious Diseases," *Chromatographia*, vol. 78, no. 5–6, pp. 379–389, Mar. 2015.

[12] S. Sauer and M. Kliem, "Mass spectrometry tools for the classification and identification of bacteria," *Nat. Rev. Microbiol.*, vol. 8, no. 1, pp. 74–82, Jan. 2010.

[13] F. Szabados, J. Woloszyn, C. Richter, M. Kaase, and S. Gatermann, "Identification of molecularly defined Staphylococcus aureus strains using matrix-assisted laser desorption/ionization time of flight mass spectrometry and the Biotyper 2.0 database," *J. Med. Microbiol.*, vol. 59, no. 7, pp. 787–790, Jul. 2010.

[14] O. Ueda, S. Tanaka, Z. Nagasawa, H. Hanaki, T. Shobuike, and H. Miyamoto, "Development of a novel matrix-assisted laser desorption/ionization time-of-flight mass spectrum (MALDI-TOF-MS)-based typing method to identify meticillin-resistant Staphylococcus aureus clones," *J. Hosp. Infect.*, vol. 90, no. 2, pp. 147–155, Jun. 2015.

[15] K. Bernardo *et al.*, "Identification and discrimination of *Staphylococcus aureus* strains using matrix-assisted laser desorption/ionization-time of flight mass spectrometry," *Proteomics*, vol. 2, no. 6, pp. 747–753, Jun. 2002.

[16] T. Majumdar, R. Chakraborty, and U. Raychaudhuri, "Development of PEI-GA modified antibody based sensor for the detection of S. aureus in food samples," *Food Biosci.*, vol. 4, no. 4, pp. 38–45, Dec. 2013.

[17] S. Balasubramanian, I. B. Sorokulova, V. J. Vodyanoy, and A. L. Simonian, "Lytic phage as a specific and selective probe for detection of Staphylococcus aureus—A surface plasmon resonance spectroscopic study," *Biosens. Bioelectron.*, vol. 22, no. 6, pp. 948–955, Jan. 2007.

[18] S. Tombelli, M. Minunni, and M. Mascini, "Analytical applications of aptamers,"

Biosens. Bioelectron., vol. 20, no. 12, pp. 2424–2434, Jun. 2005.

[19] L. R. Schoukroun-Barnes and R. J. White, “Rationally designing aptamer sequences with reduced affinity for controlled sensor performance.,” *Sensors (Basel).*, vol. 15, no. 4, pp. 7754–67, Mar. 2015.

[20] A. Abbaspour, F. Norouz-Sarvestani, A. Noori, and N. Soltani, “Aptamer-conjugated silver nanoparticles for electrochemical dual-aptamer-based sandwich detection of staphylococcus aureus,” *Biosens. Bioelectron.*, vol. 68, pp. 149–155, Jun. 2015.

[21] X. Cao *et al.*, “Combining use of a panel of ssDNA aptamers in the detection of Staphylococcus aureus.,” *Nucleic Acids Res.*, vol. 37, no. 14, pp. 4621–8, Aug. 2009.

[22] Y. Huang *et al.*, “Selection and characterization of DNA aptamers against Staphylococcus aureus enterotoxin C1,” *Food Chem.*, vol. 166, pp. 623–629, Jan. 2015.

[23] Y. Lian, F. He, H. Wang, and F. Tong, “A new aptamer/graphene interdigitated gold electrode piezoelectric sensor for rapid and specific detection of Staphylococcus aureus,” *Biosens. Bioelectron.*, vol. 65, pp. 314–319, Mar. 2015.

[24] J. Kohn, “New approaches to biomaterials design,” *Nat. Mater.*, vol. 3, no. 11, pp. 745–747, Nov. 2004.

[25] R. Langer and D. A. Tirrell, “Designing materials for biology and medicine,” *Nature*, vol. 428, no. 6982, pp. 487–492, Apr. 2004.

[26] Y. Lu and J. Liu, “Functional DNA nanotechnology: emerging applications of DNAzymes and aptamers,” *Curr. Opin. Biotechnol.*, vol. 17, no. 6, pp. 580–588, Dec. 2006.

[27] S. Mitragotri and J. Lahann, “Physical approaches to biomaterial design,” *Nat. Mater.*, vol. 8, no. 1, pp. 15–23, Jan. 2009.

[28] P. E. Scopelliti *et al.*, “The effect of surface nanometre-scale morphology on protein adsorption.,” *PLoS One*, vol. 5, no. 7, p. e11862, Jul. 2010.

[29] A. Marabotti *et al.*, “Energy-Based Prediction of Amino Acid-Nucleotide Base Recognition,” 2008.

[30] M. Dell’Acqua *et al.*, “MediaChrom: Discovering a Class of Pyrimidoindolone-Based

Polarity-Sensitive Dyes," *J. Org. Chem.*, vol. 80, no. 21, 2015.

[31] P. Suppan, "Excited-state dipole moments from absorption/fluorescence solvatochromic ratios," *Chem. Phys. Lett.*, vol. 94, no. 3, pp. 272–275, 1983.

[32] N. Mataga, Y. Kaifu, and M. Koizumi, "Solvent Effects upon Fluorescence Spectra and the Dipolemoments of Excited Molecules," *Bull. Chem. Soc. Jpn.*, vol. 29, no. 4, pp. 465–470, Apr. 1956.

[33] S. L. Svenningsen, N. Costantino, D. L. Court, and S. Adhya, "On the role of Cro in lambda prophage induction.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 102, no. 12, pp. 4465–9, Mar. 2005.

[34] D. H. Ohlendorf, D. E. Tronrud, and B. W. Matthews, "Refined structure of Cro repressor protein from bacteriophage lambda suggests both flexibility and plasticity.," *J. Mol. Biol.*, vol. 280, no. 1, pp. 129–36, Jul. 1998.

[35] S. Pellegrino et al., "Expedient chemical synthesis of 75mer DNA binding domain of MafA: an insight on its binding to insulin enhancer.," *Amino Acids.*, vol. 43, no. 5, pp. 1995–2003, Nov. 2012.

[36] G. Weber and F. J. Farris, "Synthesis and spectral properties of a hydrophobic fluorescent probe: 6-propionyl-2-(dimethylamino)naphthalene.," *Biochemistry*, vol. 18, no. 14, pp. 3075–8, Jul. 1979.

[37] Z. Lu, S. J. Lord, H. Wang, W. E. Moerner, and R. J. Twieg, "Long-wavelength analogue of PRODAN: synthesis and properties of Anthradan, a fluorophore with a 2,6-donor-acceptor anthracene structure.," *J. Org. Chem.*, vol. 71, no. 26, pp. 9651–7, Dec. 2006.

[38] L. Bagatolli, E. Gratton, T. K. Khan, and P. L. Chong, "Two-photon fluorescence microscopy studies of bipolar tetraether giant liposomes from thermoacidophilic archaebacteria Sulfolobus acidocaldarius.," *Biophys. J.*, vol. 79, no. 1, pp. 416–25, Jul. 2000.

[39] L. A. Bagatolli and E. Gratton, "Two photon fluorescence microscopy of coexisting lipid domains in giant unilamellar vesicles of binary phospholipid mixtures.," *Biophys. J.*, vol. 78, no. 1, pp. 290–305, Jan. 2000.

[40] T. Parasassi, E. K. Krasnowska, L. Bagatolli, and E. Gratton, “Laurdan and Prodan as Polarity-Sensitive Fluorescent Membrane Probes naphthalene derivatives for the study of solvent dipolar relaxation,” *J. Fluoresc.*, vol. 8, no. 4, 1998.

[41] M. Ptashne *et al.*, “How the lambda repressor and cro work.,” *Cell*, vol. 19, no. 1, pp. 1–11, Jan. 1980.

[42] S. Pellegrino *et al.*, “Synthetic peptides containing a conserved sequence motif of the Id protein family modulate vascular smooth muscle cell phenotype.,” *Bioorg. Med. Chem. Lett.*, vol. 19, no. 22, pp. 6298–302, Nov. 2009.

Chapter 5

Molecular insights into dimerization inhibition of c-Maf transcription factor

Introduction

Basic region-leucine zipper (bZIP) proteins constitute one of the largest families of transcription factors in eukaryotes. The dimeric bZIPs represent the simplest DNA binding motif consisting of i) a region rich in basic residues binding, through hydrogen bonds and hydrophobic interactions, the major groove of DNA (DNA binding domain), followed by ii) heptad repeats of hydrophobic residues forming a coiled coil structure, consisting of two parallel α-helices (leucine zipper - LZ) involved in the formation of homo- or hetero-dimers [1]. These two long α-helices grip DNA like a pair of chopsticks [2] (Figure 1).

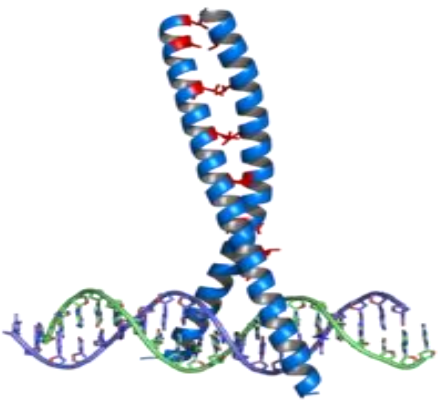


Figure 1 *Basic Leucine Zipper Domain in complex with consensus DNA sequence.*

The Maf protein family belongs to a subgroup of bZIP proteins, the activator protein 1 (AP-1) superfamily. The basic domain of the Maf family proteins has significant homology (up to 40%) to other bZIP AP-1 members. However, Maf proteins show a higher degree of similarity when compared with each other, both within and outside of the basic and zipper domains. Thus, from a structural point of view, the Maf factors form a distinct bZIP subfamily [1]. Mafs are characterized by a unique DNA-binding domain containing a highly conserved extended homology region (EHR), immediately preceding the basic region, that allows to recognize longer DNA sequences than other AP-1 family members, termed Maf recognition element (MARE). MARE is a 13–14 bp palindromic consensus sequence (TGCTGAC(G)TCAGCA) containing the canonical 7–8 bp core motif similar to TRE or CRE recognized by the other subgroups of bZIP proteins such as AP1 and CREB/ATF families, respectively, flanked by extended GC sequences [3]. The Maf family members can be subdivided into two classes based on their molecular weight, structure and function: Small Mafs (17-18 kDa) and Large Mafs (26-39 kDa), structurally differing for an N-terminal transactivation domain preserved in Large Mafs and not in Small Mafs [1] (Figure 2).

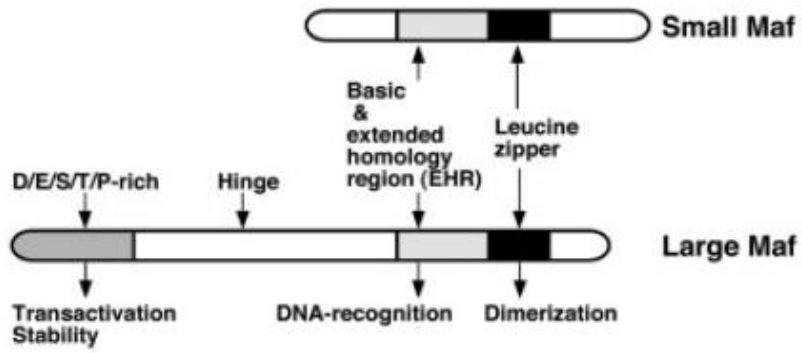


Figure 2 *Schematic representation of Large and Small Mafs domains and functions*

Maf family transcription factors (Maf TFs) regulate tissue-specific gene expression and cell-differentiation in a wide variety of tissues and are also involved in human diseases and oncogenic transformation [4]. Large Maf genes are *bona fide* oncogenes as their products can transform primary cells and their oncogenic activity has been validated *in vivo* both in birds and mammals [5]. c-Maf belongs to the large Maf subgroup and is a major regulator of gene expression and cell differentiation during development [4]. Recurrent translocations of the genes for the Maf family transcription factors, especially for c-Maf, are involved in the genesis of tumors, as in the case of multiple myeloma (MM) [6]. MM is a malignant disorder characterized by the clonal proliferation of plasma cells in the bone marrow, recently proposed as a model system for the comprehension of tumor biology [7]. Molecular approaches, e.g. genomics, transcriptomics and molecular cytogenetics, have allowed improving classification of MM to obtain molecular portraits [8] leading to the identification of novel potential biological targets. Over-expression of Maf TFs is observed in about 50% of MM cases and is associated to poor prognosis, high proliferation index and low median survival [9]. Despite the proven critical role of Mafs, recently recognized as innovative biological targets for MM therapy [10], no synthetic molecules aiming to directly inhibit their transcriptional activity have been designed and tested so far.

One of the most recurrent themes in transcriptional events is the interplay between regulation of function and oligomeric assembly. The basic region-leucine zipper constitutes the largest super family of TFs, exclusively operating in eukaryotes. Several types of TFs bind DNA as homo- or hetero-dimers, thus dimerization is a crucial event for exploiting their function. The DNA binding process involves homo- or hetero-dimerization via LZ α-helices, with dimerization constants in the micromolar range [11]. In these TFs, chain flexibility is a fundamental feature for exerting physiological function [11]. Eukaryotes TFs are reported to possess extended disordered regions accounting for at least 82% of the sequence [12], in contrast to their prokaryotic counterparts, which rarely show unstructured regions [13]. Intrinsically disordered proteins (IDPs) are involved in regulating enzymes, or providing sites for post-translational modification or for signaling

interactions with nucleic acids. Many are reported to become (more) structured upon binding to their biological targets, thus sustaining coupling of folding and binding events as a mechanism for molecular recognition and fine regulation [14]. A number of IDPs are known to be over-expressed in major diseases (cancer, diabetes, cardiovascular diseases), which makes them good prospective drug targets [14]. A relevant feature of IDPs as therapeutic targets is that they can interact with a variety of molecules, some of them showing high specificity. In fact, the interactions of IDPs are characterized by a high specificity with a low affinity, an effective property for reversible binding in multiple interaction networks.

Research purpose

The physiological role of Maf TFs as regulators of tissue-specific gene expression and cell-differentiation has been validated in a wide variety of experimental models. Maf TFs activate expression of specific genes whose products are involved in vital cellular functions. The deregulation of their activity is involved in the development of several human diseases and their oncogenic activity has been validated *in vivo* both in birds and mammals [5]. c-Maf belongs to the large Maf subgroup. When over-expressed is involved in the genesis of about 50% of Multiple Mieloma cases characterized by poor prognosis, and low median survival [9]. Despite Mafs have been recently recognized as innovative biological targets for MM therapy [10], there is a lack of synthetic molecules aiming to directly inhibit their transcriptional activity. From a structural point of view, Mafs DNA binding process involves homo- or hetero-dimerization via LZ α-helices. Therefore, dimerization represents a crucial event for exerting their function [11].

This research line aims at the development of a simple experimental model to study potential dimerization inhibitors as a way for interfering with c-Maf function. The c-Maf LZ motif, obtained by solid phase peptide synthesis, has been characterized in terms of structural features and dimerization properties. This synthetic dimerization domain is used to test potential synthetic peptidic inhibitors of c-Maf dimerization, rationally designed starting from a homology model of the whole DNA binding domain.

Materials and methods

Peptides and protein domain

All protein domains and inhibitors peptides were synthesized using Microwave-assisted Solid Phase Peptide Synthesis (MW-SPPS) and purified using RP-HPLC [15] by the DISFARM-Section of General and Organic Chemistry "A. Marchesini", University of Milan. After lyophilization, protein domains and peptides stock solutions (100 µM) were prepared using 20 mM phosphate buffer, pH 7.4 and stored at -20 °C. In table 1 leucine zipper sequences of c-Maf and MafA, and the sequences of potential c-Maf LZ dimerization inhibitors are reported.

Table 1 *Sequences of c-Maf and MafA LZ domain and potential c-Maf dimerization inhibitors (mutations in red)*

c-Maf LZ	QRHVLESEKNQLLQQVDHLKQEISRLVRERDAYKEKYEKLV
MafA LZ	QRHILESEKCQLQSQVEQLKLEVGRLAKERDLYKEKYEKLA
N-terminus wild type sequence wt1	RFKRVQQRHVLESEKNQLLQQV
1	EFKEVQERHVRLSEKNQLIQQV
2	EFKDVQERHVKLSEKNQLIQQV
3	EFKEVQERHVRLSEKNQLI
4	EFKDVQERHVRLSEKNQLLQQI
5	EFKDVQERHVLESEKNQLL
C-terminus wild type sequence wt2	ISRLVRERDAYKEKYEKL
6	LSRLLRERDAYKEKYEKI
7 (negative control from wt1)	RFKRVQQRHVAESEANQAL

Circular dichroism

Circular dichroism (CD) spectra were collected using a Jasco J-715 spectropolarimeter equipped with a Peltier thermostatic unit set at 20 °C. The spectral scans were collected

between 280 and 190 nm, 0.1 nm data pitch, at 50 nm/min scanning speed. Each spectrum is the result of 3 averaged accumulations.
Depending on the peptides concentration, cuvettes with 0.1, 0.5 and 1 cm path length were used. Secondary structure analysis was carried out using Dichroweb server (CONTIN algorithm, reference set 7).

Sample preparation for CD experiments

Leucine zipper and dimerization inhibitor peptides stored in lyophilized form were solubilized in 20 mM phosphate buffer, pH 7.4. 100 µM stock solutions were prepared for single experiments. CD spectra on LZ-inhibitor mixtures were collected after overnight incubation.

Mass spectrometry

Cross-linked samples were dialyzed against deionized water using diafiltration devices (Amicon Ultra—0.5mL, cut-off 3000 Da) to remove glutaraldehyde and sodium borohydride. Cross-linked and unmodified samples were prepared by dried droplet method using a 10 mg/ml α-cyano-4-hydroxycinnamic acid (CHCA) solution prepared in acetonitrile 50% – TFA 0.05%. The samples were mixed in equal volume with the matrix and then 1 µl of the mixture was spotted onto MALDI plate. Spectra were acquired using a 4800 Plus MALDI TOF/TOF™ AbSciex mass spectrometer. The instrument was calibrated with cytochrome c and apomyoglobin (Sigma-Aldrich®).

Cross-linking

Solutions containing 50 µM LZ alone or a mixture of 50 µM LZ and 50 µM dimerization inhibitors were reacted with glutaraldehyde (the latter, after overnight incubation). Glutaraldehyde bivalently reacts with amino groups such as N-termini and lysine or arginine side chains giving covalent cross-linking.

Different protocols for glutaraldehyde reaction were tested, varying glutaraldehyde concentration and incubation time, as listed below:

- Glutaraldehyde 5% for 15'
- Glutaraldehyde 2% for 15'
- Glutaraldehyde 0.5% for 60'
- Glutaraldehyde 0.05% for 30'
- Glutaraldehyde 0.005% overnight

The reaction was stopped by adding a 150 mg/ml sodium borohydride solution in 0.1 M sodium hydroxyde to reach a 4:1 excess molar ratio with respect to glutaraldehyde.

Results and discussions

Inhibition of LZ dimerization was previously investigated for other classes of transcription factors [16]. Compound library screenings were performed searching for dimerization inhibitors, and several peptidomimetic molecules were identified as in the case of c-Myc LZ inhibitors [16]. Moreover, NMR data indicated that the binding region of bZIP domain, when in complex with these molecules, is more structured relative to the uncomplexed form, although conserving a substantial flexibility.

As a starting point, a homology model of the c-Maf homodimer, both free and bound to the consensus DNA sequence, was created using the crystal structure of the homodimeric MafA (4EOT) as the template. A rational design of a small library of peptides as putative inhibitors of c-Maf dimerization was attempted using molecular dynamic (MD) simulations followed by Molecular Mechanics-Generalized Born Surface Area (MM-GBSA) analysis [15]. The c-Maf and MafA LZ fragments and putative inhibitors were chemically synthesized using MW-SPPS (Table 1 - Materials and Methods).

Circular dichroism spectroscopy on isolated leucine zipper domains

The secondary structure of c-Maf and MafA LZ was characterized by far-UV CD spectroscopy in order to evaluate if the isolated fragment, obtained by solid phase peptide synthesis, was able to adopt in solution a helical structure, similar to the expected secondary arrangement of the LZ dimerization domain of other members of the large Maf family [3]. The LZ of MafA shows a completely unordered structure (Figure 3).

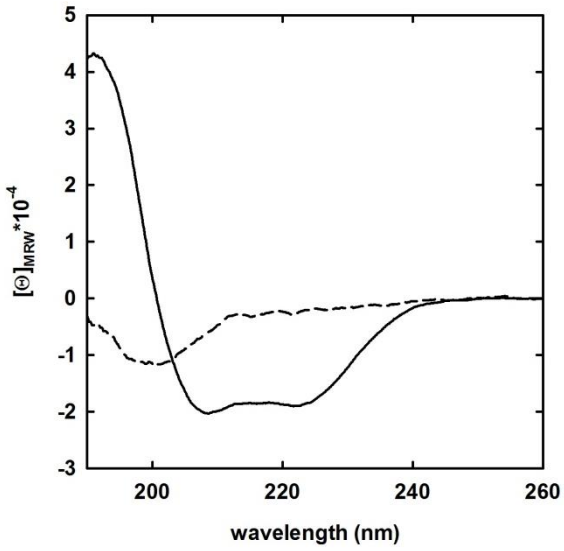


Figure 3 *Circular dichroism spectra of 20 μM c-Maf (solid line) and MafA (dashed line) isolated LZ fragments in 20 mM phosphate buffer, pH 7.4, 20 °C.*

This result is in agreement with data reported in the literature indicating that, in the absence of the consensus DNA sequence, both the LZ and the EHR of MafA are crucial for the complete helical folding of the DNA binding domain. Moreover, post-translational modifications are crucial for the correct folding of MafA [3]. Opposite to MafA, the c-Maf LZ motif shows a CD spectrum with two negative bands centered at 208 and 222 nm, typical of α-helices (Figure 3). To quantify the helical content, the spectrum was analyzed using the CONTIN algorithm, as previously done for the bZIP domain of other Mafs [17]. The analysis yields a 43% of helical structure, indicating that the chemically synthesized, isolated LZ fragment of c-Maf achieves a significant degree of secondary structure in solution, in the absence of both the remaining sections of DNA binding domain in large Mafs [3], and of the partner DNA, which is required for most of Maf members folding. This not obvious result indicates that solid phase peptide synthesis of the isolated c-Maf LZ region is a valuable strategy to set up a simplified model and experimental platform for testing putative LZ dimerization inhibitors. Obviously, the capability of the LZ domain to form dimers in solution is a necessary condition for its use as a model system to study

dimerization inhibitors. Therefore, based on the expected dependence of secondary structure content on the degree of dimerization, we further exploited CD spectroscopy of c-Maf LZ at different concentrations to evaluate its oligomerization state in solution (Figure 4).

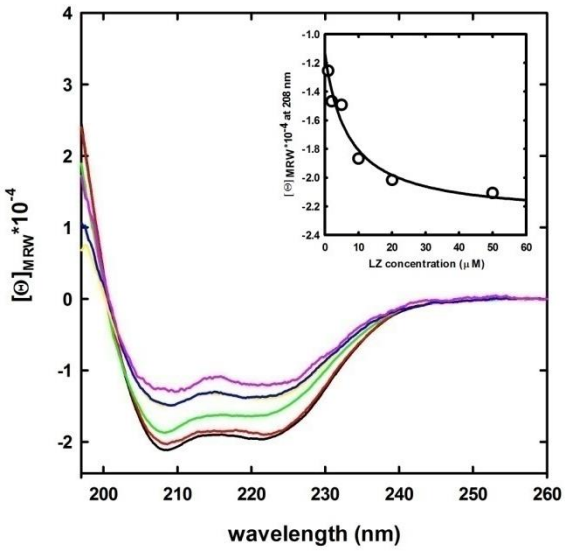


Figure 4 *Concentration dependence of c-Maf LZ CD spectra. Spectra were recorded for solutions containing 20 mM phosphate buffer, pH 7.4, 20 °C, and different concentrations of c-Maf LZ: 50 μM (black line), 20 μM (red line), 10 μM (green line), 5 μM (yellow line), 2 μM (blue line) and 1 μM (magenta line). Inset: signal change monitored at 208 nm. Data were fitted to a simple binding isotherm, with a Kd of 7.0 ± 4.3 μM (solid line).*

The normalized CD spectra show that the intensity of the two bands at 208 and 222 nm increases as a function of c-Maf LZ concentration spanning from 1 to 50 μM (Figure 4), indicating that protein–protein interactions promote an increase of ordered helical structure. In principle, we cannot exclude that under our experimental conditions the isolated LZ region undergoes oligomerization processes other than a simple monomer–dimer equilibrium. However, the dependence on LZ concentration of the ellipticity at 208 nm is well described by a simple binding isotherm with a dimerization constant Kd of 7.0 ± 4.3 μM (Figure 4, inset).

Mass spectrometry on c-Maf leucine zipper domain

To better characterize the oligomerization state of c-Maf LZ, and exclude that higher order oligomers, other than homodimers, are present at equilibrium, a mass spectrometry analysis was carried out at a LZ concentration of 50 μM. Under these conditions, the concentration dependence of helical content, as evaluated from CD spectra (Figure 4), is close to plateau. As reported in Figure 5, panel A, the mass spectrum shows a major peak at 5,076 m/z (peak 1), corresponding to the monomeric form of c-Maf LZ, and only a small peak at 10,131 m/z (peak 2), attributable to a dimeric form, both in the absence and presence of dimerization inhibitors.

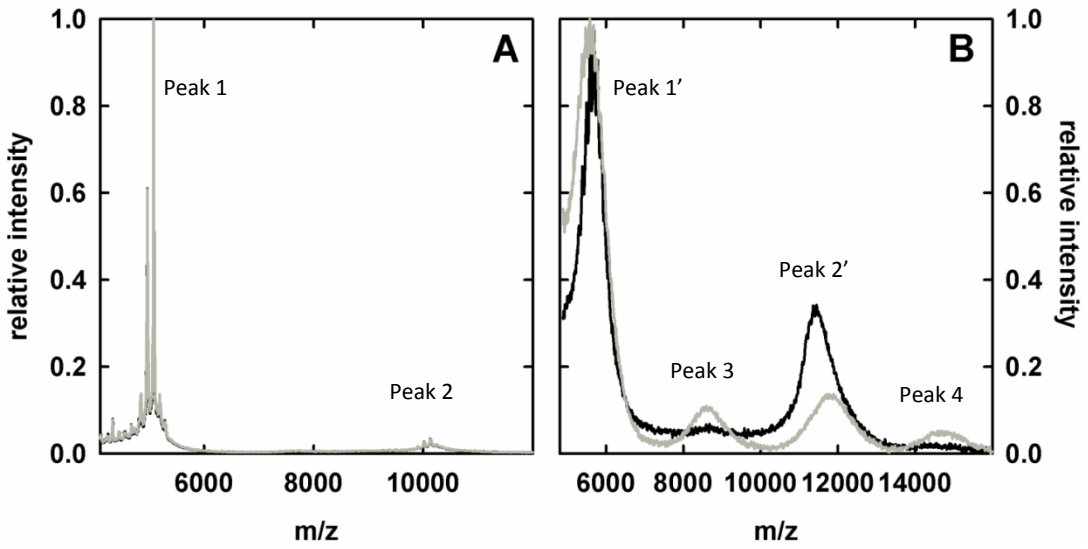


Figure 5 Panel A: mass spectrum of 50 μM c-Maf LZ in the absence (black line) and presence (gray line) of peptide 5. Panel B: mass spectrum of c-Maf LZ in the absence (black line) and presence (gray line) of peptide 5, upon treatment with 0.5 % glutaraldehyde for 15 min. Peaks attribution is as follows (see main text): peak 1, c-Maf LZ monomer; peak 1', c-Maf LZ monomer reacted with glutaraldehyde; peak 2, c-Maf LZ dimer; peak 2', c-Maf LZ dimer reacted with glutaraldehyde; peak 3, complex between c-Maf LZ monomer and peptide 5 (reacted with glutaraldehyde); peak 4, complex between c-Maf LZ dimer and peptide 5 (reacted with glutaraldehyde).

This result depends on the sample preparation protocol for mass spectra analysis, that encompasses denaturing conditions that are expected to cause dimers disassembly and LZ unfolding. To stabilize the equilibrium distribution of quaternary complexes, LZ solutions were reacted with glutaraldehyde, a cross-linker covalently binding the amino groups of Lys and Arg side chains as well as protein N-terminus. The reaction was stopped using the reducing agent sodium borohydride, which also stabilizes the crosslink. Different protocols, at varying glutaraldehyde concentrations and incubation times, were tested (described in Materials and Methods section). The example spectrum reported in Figure 5, panel B, was obtained reacting 50 μM LZ in the presence of 0.5% glutaraldehyde, with a 15 min incubation time, the condition showing the best signal intensity. Indeed, the reaction of c-Maf LZ with glutaraldehyde produced a clear change in the mass spectrum. As compared to the unreacted solution (Figure 5, panel A), the monomer peak (peak 1') is right shifted at about 5,700 m/z and is more heterogeneous, due to LZ reaction with a variable number of glutaraldehyde molecules which cause an increase of the molecular weight. The dimer peak at about 11,400 m/z (peak 2') is definitely more intense than in the absence of glutaraldehyde crosslinking (peak 2). Peaks at higher m/z, attributable to trimers or large oligomers, are not present. These evidences highlight that the LZ domain is present in solution as a concentration-dependent equilibrium of monomers and dimers, and dimerization favors the structuring of the domain. The absence of higher order oligomers confirms the possibility to use the LZ domain of c-Maf as a model system to test dimerization inhibitors [15].

Circular dichroism spectroscopy and mass spectrometry in the presence of dimerization inhibitors

The interaction between the designed peptidic dimerization inhibitors and c-Maf LZ was studied by far-UV CD spectroscopy, as already done for other Mafs [17]. All synthetic inhibitor peptides show CD spectra indicative of a completely unstructured fold (e.g. peptide 5, Figure 6A, green line), thus suggesting that these peptides are not prone to form structured dimers or oligomers upon self-association.

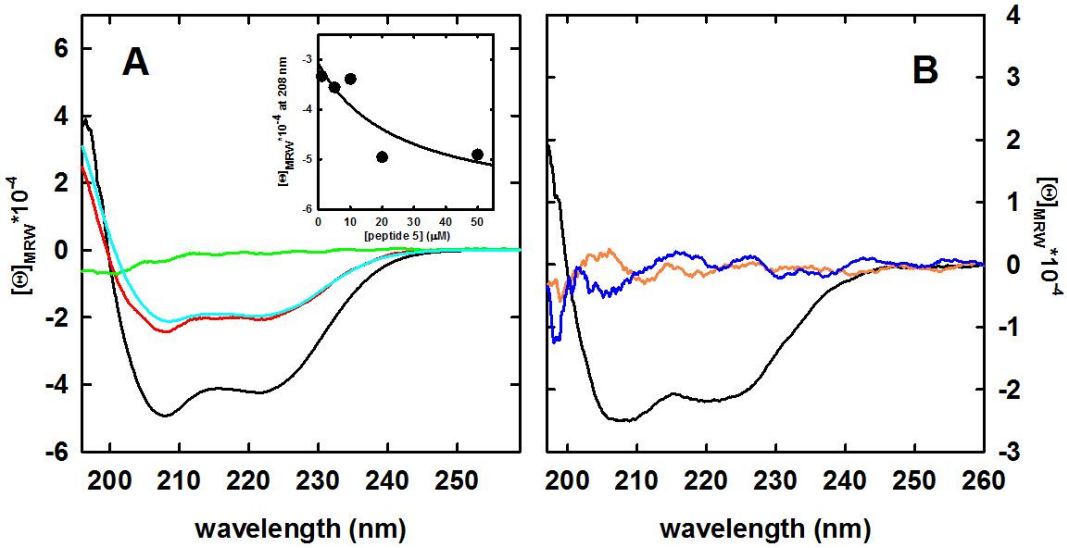


Figure 6 *Panel A: binding of peptide 5 to c-Maf LZ. CD spectra were collected on solutions containing 20 mM phosphate buffer, pH 7.4, 20 °C, and either 50 μM peptide 5 (green line), 50 μM c-Maf LZ (light blue line) or a mixture of 50 μM c-Maf LZ and 50 μM peptide 5 (black line). c-Maf LZ concentrations are always expressed on a monomer basis. The red line is the arithmetic sum of the CD spectra of peptide 5 alone and c-Maf LZ alone. Inset: signal change at 208 nm at increasing concentration of peptide 5 in the presence of 50 μM c-Maf LZ. Panel B: CD difference spectra between the observed spectrum of reacted mixture and the arithmetic sum of the spectra of isolated components for c-Maf LZ and peptide 5 (black line), MafA LZ and peptide 5 (orange line) and c-Maf LZ and peptide 7 (blue line).*

Solutions containing LZ domain and peptide inhibitors at the same concentration were incubated overnight before recording spectra. A comparison between the spectra of the reacted LZ-inhibitor peptide mixture and the arithmetic sum of the two separated components (at the same concentrations) highlights that the mixtures of LZ zipper domain and peptide inhibitor have a higher helical content, as shown in Figure 6A for the complex with peptide 5, demonstrating an interaction which promotes a structuring process. The CD spectra of mixtures of 20 μM LZ domain and 20 μM peptidic inhibitors were analyzed using the CONTIN algorithm and the results of the analysis of triplicate experiments are shown in Figure 7.

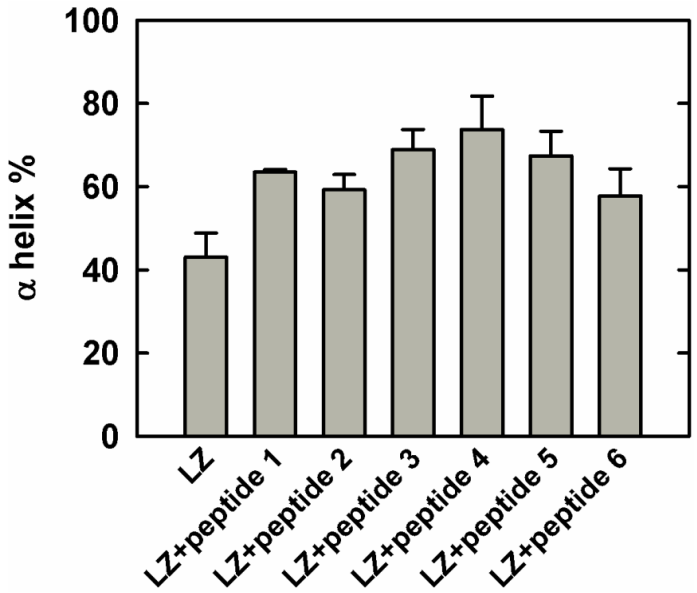


Figure 7 *Dimerization inhibitors effect on c-Maf LZ secondary structure. The α-helical content of LZ in the absence and in the presence of 20 μM dimerization inhibitors was estimated using the CONTIN algorithm. All experiments were carried out in triplicate.*

In comparison with LZ domain alone, which exhibits an estimated helical fraction of 43%, the presence of all tested peptide inhibitors promotes an increase of helical structure, spanning from 58 ± 7% (peptide 6) to 74 ± 8% (peptide 4). A similar increase of the helical content was observed at a 2 μM LZ concentration, the lower value yielding reliable CD spectra with our experimental apparatus. The inset in Figure 6A shows the dependence on inhibitor concentration (peptide 5) of the signal at 208 nm, where α-helices make a dominant contribution to molar ellipticity, of a c-Maf LZ-inhibitor mixture.

We are aware that changes in CD signal arise from the overlapping contribution of i) loss of helical content due to LZ dimer displacement and ii) gain of secondary structure content of LZ and, possibly, peptidic inhibitor upon LZ-inhibitor complex formation. Therefore, from CD experiments it is not possible to retrieve reliable thermodynamic parameters for binding of inhibitors to LZ. However, the observed concentration dependence shows that a maximal effect is almost reached at an inhibitor concentration

of 50 μM (Figure 6A, inset, peptide 5). From this dependence, and considering the LZ dimerization constant of 7.0 μM, a semi-quantitative conclusion can be drawn that the binding constant resides in the low micromolar range, which encourages further development of more and more effective ligands. All the peptides, that have been specifically designed to interact with c-Maf LZ, were also tested for their capability to bind to MafA LZ domain, but no one proved to be able to promote helical secondary structure formation. With this respect, Figure 6B shows that the spectrum of a MafA LZ-peptide 5 mixture overlaps with the sum of the spectra of the single isolated components. It should be noted that the results of the latter control experiments do not prove specificity for c-Maf of our inhibitors, since while the isolated c-Maf LZ is substantially structured in solution and competent for concentration-dependent dimerization, MafA LZ is not (Figure 3), and there are several claims in the literature that dimerization of MafA i) is dependent on post-translational modifications on the N-terminal domain and ii) occurs after a coupled DNA binding/disorder-to-order transition. This hampers the possibility to carry out a proper positive control by designing peptidic inhibitors targeting MafA LZ sequence and testing them on the isolated MafA LZ fragment [15].

The interaction between c-Maf LZ domain and peptidic inhibitors has been also investigated using mass spectrometry. Complexes between c-Maf LZ and peptidic inhibitors were covalently stabilized by glutaraldehyde cross-linking and analyzed with mass spectrometry to characterize the oligomerization state, as described above. The resulting spectra showed a different m/z pattern with respect to the mass spectrum of glutaraldehyde-treated LZ in the absence of dimerization inhibitors (Figure 5, panel B), with a marked decrease of the peak corresponding to the LZ dimer (peak 2’), and the appearance of broad bands at m/z values of about 8,600 (peak 3) and 14,700 m/z (peak 4), compatible with masses of LZ monomer-inhibitor and LZ dimer-inhibitor complexes. Obviously, the latter has not to be intended as a preexisting ternary complex stabilized by glutaraldehyde, but rather as the product of an aspecific crosslinking between a LZ dimer and a free inhibitor peptide molecule. Although mass spectra analysis cannot be considered a quantitative technique and the glutaraldehyde reaction is not able to

stabilize all the complexes present in solution, these data demonstrate that the peptidic inhibitors that we have designed and tested are able to selectively interact with c-Maf LZ sequence. Such interaction results in the destabilization of dimers, the functional form of Maf TFs. The increase in the content of ordered structure elicited by binding of dimerization inhibitors to monomeric LZ domains appears to be a key feature in the mechanism of dimerization inhibition.

Conclusions

In this work, the LZ motif of c-Maf transcription factor was obtained by solid phase peptide synthesis and characterized in terms of structural features and dimerization properties. The results highlight the tight coupling between dimerization and folding, adding new insights in the poorly characterized Maf proteins family. Potential peptidic inhibitors of c-Maf dimerization were tested *in vitro* for their effect on dimerization of the isolated LZ fragment. Some of these compounds proved to bind c-Maf LZ, promoting a disordered-to-structured transition through an increase in helical content and inhibiting c-Maf LZ dimerization [15]. Our results showed that solid phase peptide synthesis of protein domains and analysis of secondary and quaternary structure by CD and MALDI TOF techniques allow to build up a simple experimental platform to test LZ dimerization inhibitors. Rationally designed peptide inhibitors provide a promising lead for further development of more selective, stable and bioavailable peptidic or peptidomimetic drugs targeting MM and other diseases whose molecular basis relies on protein–protein recognition.

References

[1] V. Blank and N. C. Andrews, “The Maf transcription factors: regulators of differentiation.,” *Trends Biochem. Sci.*, vol. 22, no. 11, pp. 437–41, Nov. 1997.

[2] T. E. Ellenberger, C. J. Brandl, K. Struhl, and S. C. Harrison, “The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted alpha helices: crystal structure of the protein-DNA complex.,” *Cell*, vol. 71, no. 7, pp. 1223–37, Dec. 1992.

[3] X. Lu, G. P. Guanga, C. Wan, and R. B. Rose, “A novel DNA binding mechanism for maf basic region-leucine zipper factors inferred from a MafA-DNA complex structure and binding specificities.,” *Biochemistry*, vol. 51, no. 48, pp. 9706–17, Dec. 2012.

[4] K. Kataoka, “Multiple mechanisms and functions of maf transcription factors in the regulation of tissue-specific genes.,” *J. Biochem.*, vol. 141, no. 6, pp. 775–81, Jun. 2007.

[5] A. Eychène, N. Rocques, and C. Pouponnot, “A new MAFia in cancer.,” *Nat. Rev. Cancer*, vol. 8, no. 9, pp. 683–93, Sep. 2008.

[6] C. S. Mitsiades, N. Mitsiades, N. C. Munshi, and K. C. Anderson, “Focus on multiple myeloma.,” *Cancer Cell*, vol. 6, no. 5, pp. 439–44, Nov. 2004.

[7] W. Matsui, “Perspective: A model disease,” *Nature*, vol. 480, no. 7377, pp. S58–S58, Dec. 2011.

[8] H. A. F. Stessman *et al.*, “Profiling bortezomib resistance identifies secondary therapies in a mouse myeloma model.,” *Mol. Cancer Ther.*, vol. 12, no. 6, pp. 1140–50, Jun. 2013.

[9] R. Fonseca *et al.*, “International Myeloma Working Group molecular classification of multiple myeloma: spotlight review.,” *Leukemia*, vol. 23, no. 12, pp. 2210–21, Dec. 2009.

[10] Y.-T. Tai *et al.*, “CS1 promotes multiple myeloma cell adhesion, clonogenic growth, and tumorigenicity via c-maf-mediated interactions with bone marrow stromal

cells.," *Blood*, vol. 113, no. 18, pp. 4309–18, Apr. 2009.

[11] M. Miller, "The importance of being flexible: the case of basic region leucine zipper transcriptional regulators.," *Curr. Protein Pept. Sci.*, vol. 10, no. 3, pp. 244–69, Jun. 2009.

[12] J. Liu, N. B. Perumal, C. J. Oldfield, E. W. Su, V. N. Uversky, and A. K. Dunker, "Intrinsic disorder in transcription factors.," *Biochemistry*, vol. 45, no. 22, pp. 6873–88, Jun. 2006.

[13] Y. Minezaki, K. Homma, A. R. Kinjo, and K. Nishikawa, "Human transcription factors contain a high fraction of intrinsically disordered regions essential for transcriptional regulation.," *J. Mol. Biol.*, vol. 359, no. 4, pp. 1137–49, Jun. 2006.

[14] H. J. Dyson and P. E. Wright, "Coupling of folding and binding for unstructured proteins.," *Curr. Opin. Struct. Biol.*, vol. 12, no. 1, pp. 54–60, Feb. 2002.

[15] S. Pellegrino *et al.*, "Molecular insights into dimerization inhibition of c-Maf transcription factor.," *Biochim. Biophys. Acta*, vol. 1844, no. 12, pp. 2108–15, Dec. 2014.

[16] T. Berg *et al.*, "Small-molecule antagonists of Myc/Max dimerization inhibit Myc-induced transformation of chicken embryo fibroblasts.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 99, no. 6, pp. 3830–5, Mar. 2002.

[17] S. Pellegrino, C. Annoni, A. Contini, F. Clerici, and M. L. Gelmi, "Expedient chemical synthesis of 75mer DNA binding domain of MafA: an insight on its binding to insulin enhancer.," *Amino Acids*, vol. 43, no. 5, pp. 1995–2003, Nov. 2012.