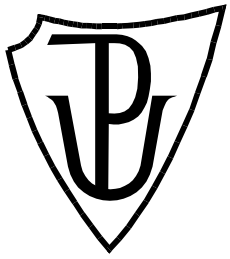
PALACKY UNIVERSITY OLOMOUC

Faculty of Natural Sciences

Department of Biochemistry



Protein modifications

MASTER THESIS

Author:	**Bc. Lukáš Najdekr**
Study program:	B1406 Biochemistry
Study branch:	Biochemistry
Study mode:	Full-time
Advisor:	**prof. Mgr. Marek Šebela, Dr.**
Co-adviser:	Eva Estebanez-Perpina, Ph.D.
Submitted:	29.4.2011

I hereby declare that this thesis has been written solely by myself and that the sources used in this thesis are cited and included in the references part.

In Olomouc .....

Lukáš Najdekr

Acknowledgement

I would like say thank you to my advisor Marek Šebela for his willingness, guidence, valuable advices and patience, to Jan Frömmel for his help with enzyme kinetics and production of AMADH2 and to all department staff for their help and goodwill. I would like say also thank you to all laboratory staff at IBUB especially to Eva Estébanez-Perpiñá, Laia Rodríguez-Carbó and Victor Buzón for their willingness, advice and all their help during my internship. And special thanks to Radka Končitíková for her great coffee-breaks during my stay at department of biochemistry.

Bibliografická identifikace:

Jméno a příjmení autora:	Lukáš Najdekr
Název práce:	Proteinové modifikace
Typ práce:	Diplomová
Pracoviště:	Katedra Biochemie; IBUB
Vedoucí práce:	prof. Mgr. Marek Šebela, Dr.
Rok obhajoby práce:	2011

Abstrakt: Proteinové modifikace jsou velmi důležitou součástí jak v přírodě tak i v laboratoři. Spousta moderních vědeckých metod je založena na proteinových modifikacích. Biotinylace je jednou z nejběžnějších modifikací používaných v afinitní chromatografii pro purifikaci a imobilizaci proteinů. V této práci byla biotinylována aminoaldehyd dehydrogenása z hrachu setého (*Pisum Sativum*) pomocí reakčního činidla Sulfo-NHS-LC-LC-Biotin. S pomocí MALDI-TOF-MS byla úspěšně určena poloha pěti modifikaci (LC-LC-biotinu) v proteinové struktuře AMADH2. Biotinylace měla značný dopad na enzymatickou activitu AMADH2, která byla zredukována na 16% activity nemodifikované AMADH2.

Druhá část této práce byla věnována studiu modifikací využívaných při purifikaci proteinů. Androgenní receptory, které patří mezi skupinu nukleárních receptorů, byli cílovou skupinou pro studium těchto modifikací. Konkrétně bylo pracováno s koregulátorem androgeních receptorů 70 (ARA70) a byla provedena příprava ARA70 1-170 pro měření retgenovou krystalografii. Pro tvorbu tohoto proteinu bylo využito tří rozdílných expresních

vektorů – pDEST15, pDEST17 a pDEST-HisMBP. Pro separaci ARA70 ze směsi proteinů bylo využito His-Trap kolon a konečné purifikace byly prováďeny s pomocí iontově výměnné chromatografie. Nakonec protein ARA70 1-170 byl úspěšně vyprodukován a purifikován jako vzorek pro rentgenovou krystalografii.

Klíčová slova: Aminoaldehyd dehydrogenasa 2; MALDI; biotinylace; androgenní receptory, ARA70

Počet stran: 107

Počet příloh: 0

Jazyk: Anglický

Bibliographical identification:

Author´s first name and surname:	Lukáš Najdekr
Title:	Protein modifications
Type of thesis:	Master
Department:	Department of Biochemistry, IBUB
Supervisor:	prof. Mgr. Marek Šebela, Dr.
The year of presentation:	2011

Abstract:

Protein modifications are very important in the nature and in the laboratory. Many modern scientific methods are based on artificial protein modifications.

Biotinylation is one of the most common modification used in affinity chromatography for purification and immobilization. In this thesis, aminoaldehyde dehydrogenase 2 from *Pisum Sativum* was biotinylated by using the Sulfo-NHS-LC-LC-Biotin reagent. Using MALDI-TOF-MS, localization of five LC-LC-biotins was successfully determined in the protein structure of AMADH2. Biotinylation had significant impact on AMADH2 activity which was reduced to 16% of activity of non-modified AMADH2.

In the second part of this thesis, modifications necessary for protein purification were studied. In detail, androgen receptor-associated co-regulator 70 (ARA70) and its preparation for X-ray crystallography was studied. Three different expression vectors were used for its

production – pDEST15, pDEST17 and pDEST-HisMBP. For separation, His-Trap column was used and for the final purification ion exchange chromatography was performed. Finally, part of the ARA70 1-170 was purified and prepared for further analysis and crystallization conditions screening.

Keywords: Aminoaldehyde dehydrogenase 2; MALDI; biotinylation; androgen receptors; ARA70

Number of pages: 107

Number of appendices: 0

Language: English

Table of content

Aims of work:

- Elaboration of theoretical research
- Preparation of the modified aminoaldehyde dehydrogenase 2
- Immobilization of the modified enzyme
- Expression of the ARA70 1-170
- Purification of the ARA70 1-170

Theoretical part

1. Protein modification

1.1 Some naturally occurring modifications

If scientists study cellular structure or biological processes, they almost always encounter proteins. Most of the dry cell's mass is created by proteins. The specificity of proteins depends on their binding abilities, because many functions of proteins are given by specific interactions with ligands (word is derived from Latin *ligare,* meaning "to bind"), their three dimensional structure and protein-protein interactions (Alberts et al, 2008).

In nature there are many very important modifications, which are made after translation – post-translation modifications. Post-translation modification affect many different aspects of protein "life cycle" (stability, intracellular localization, their folding, activity and interactions with other biomolecules). Essentially, every modification gives the cell information about the state of proteins inside. For example, phosphorylation is something like a „switch on" press button of enzymes. Poly-ubiquitination on the other hand gives to the cell a signal, that the protein is ready for degradation (Alberts et al, 2008).

Glycosylations form one of the biggest groups of protein modifications. Carbohydrates themselves form a separate group of biopolymers including small oligosaccharides as well as large macromolecules like cellulose and glycogen. Glycosylation influences some important aspects of proteins, like a cellular localization, turnover and protein quality control (Parodi, 1999). The complexity of glycosylation is quite huge, it requires a lot of enzymes (glycosyltransferases, glycosidases) which allow to generate a specific glycan pattern in relation to actual state of the cell (Marino et al, 2010). Glycosylation takes place in endoplasmatic reticulum (ER). N-linked oligosaccharide structures are the most common in glycopeptides. The rest of the glycans are linked to the protein through a hydroxyl group of serine, threonine or hydroxylysine side chains. These O-linked oligosaccharides are formed in the Golgi apparatus (Alberts et al, 2008).

Protein modification can also influence protein folding and conformational changes. For example, during phosphorylation, phosphate brings a negative charge to

the protein chain, which can attract positively charged groups of amino acid side chains and thus can reveal some other interacting place on the surface of the protein. On the other hand, phosphate can be part of binding or interacting side of protein and may mediate an interaction. Through this effect, many of protein complexes and other biochemical pathways are regulated or co-regulated by phosphorylation (Alberts et al, 2008).

Acetylation represents another very important protein modification in eukaryotes. Some arguments point to the fact that acetylation is at the same level of importance as phosphorylation (Kouzarides, 2000). There are two kinds of protein acetylation: acetylation of the alpha-amino group of the N-terminal amino acid of proteins and acetylation of the epsilon-amino group of an internal lysine. The main biological reason for acetylation is still not clear. It was reported that two different versions of the same protein, one with acetylation and second without it, differ in protein stability, activity and biological half life. An ability, which is often discussed, is a protection against proteolytic degradation (Soppa et al, 2010).

Nevertheless, protein modifications in the living cell are really complex and difficult and it is important to study them for better understanding of the cell mechanisms.

1.2 Artificial protein modifications

When scientists started to work with proteins, they began to develop methods for a separation and visualization of proteins of their interest. Since in the cell, there is an enormous amount of ballast proteins, it was quite difficult to separate only one protein, especially if those were proteins with a low concentration. Using protein modification to introduce different tags to the protein of interest makes the job much easier. However, not all of the tags were universal for every experiment and that is why a research of new suitable tags began. Nowadays one can go by two different ways: first, may be incorporated tags into DNA sequence and the protein is expressed with the desired modifications (HisTag, GST) (Waugh, 2005; Bashiri et al, 2007), or you can modify an already existing protein can be modified using of specially designed reagents (NHS-Biotin, etc.) (Elia, 2008). It is also possible to incorporate radioactive tags into proteins

by letting cells grow on a medium a with radioactive isotope (Werner and Mohammad, 1966).

If specific tags are required for protein purification, will mostly encounter tags based on affinity interaction (chromatography). Represents one of the most common protein affinity tags (His_6) using a poly-histidine tail line. Together with immobilized metal affinity chromatography (IMAC) gives a powerful tool for protein purification (Nallamsetty and Waugh, 2007). Another quite common tag is glutathione S-transferase (GST). Here fusion proteins are purified by glutathione affinity chromatography (Scheich et al, 2003).

A modification can be used to enhance specific property of a protein. For X-ray crystallography or nuclear magnetic resonance (NMR) a large amount of soluble protein is needed and sometimes the protein is not as soluble according to wishes. By coupling it with some other more soluble proteins like maltose binding protein (MBP) (Nallamsetty and Waugh, 2007) or NusA (Davis et al, 1999), you can achieve your goal.

ICAT (Isotope coded affinity tags) is a good example of the second way of protein modification . This method uses very specific reagents, which are designed to bring a maximum benefit from one modification (Fig. 1) (Turecek, 2002).

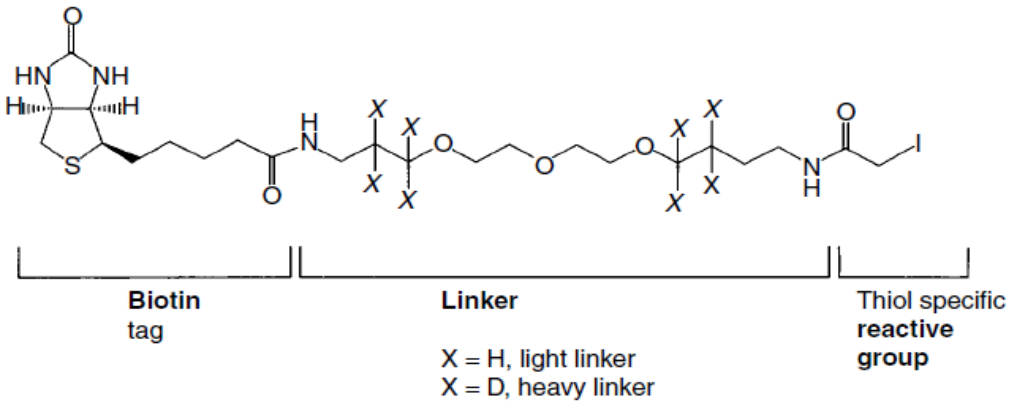


Figure 1: Example of an ICAT reagent, combining biotin tag for purification and light and heavy linker for quantification. (Turecek, 2002).

Biotinylation may be considered both a natural an artificial modification in the same time. It is normally appears in living organisms (Dupuy D'ngeac et al, 2006; Kothapalli et al, 2005) and it is also used for purification, isolation (Desarnaud et al,

1992) and immobilization of proteins (Elia, 2008).

Fluorescent tags represent another possible type of protein modification. The most frequently used ones are coupled or fusion proteins with some other protein or peptide containing fluorochrome. One of the most frequent tags is the green fluorescent protein the (GFP). GFP is a protein which exhibits green fluorescence when exposed to the blue light. It was first isolated from a jellyfish *Aequorea victoria*, and nowadays thousands of scientific articles can be found based on using this protein or its homologs providing different colours for visualizing proteins (Fig.2) (Shimomura et al, 1962; Tsien, 1998).

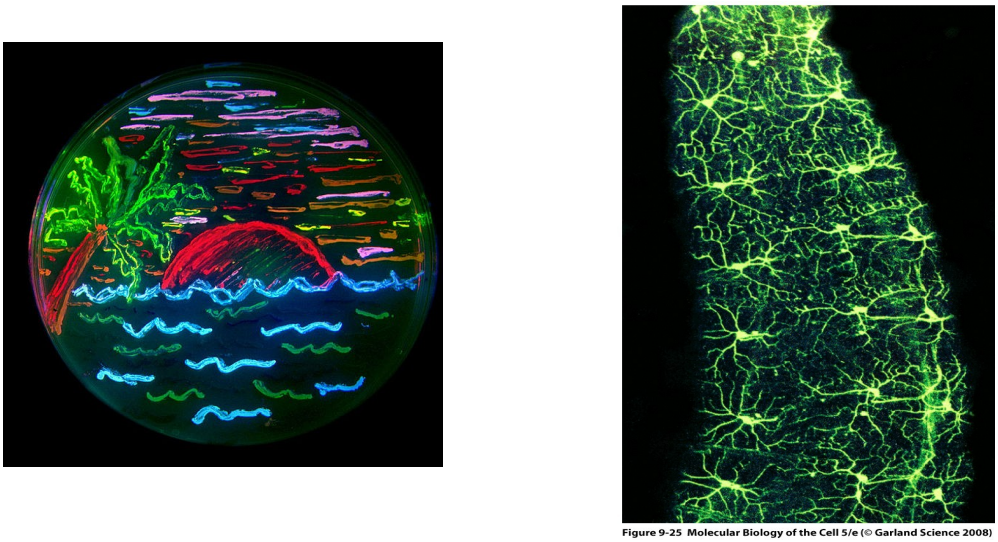


Figure 2: Picture on the left - Picture of a beach using eight different GFP modifications (http://en.wikipedia.org 20.3.2011). Picture on the right – neurons of a fruit fly embryo coupled with GFP (Alberts et al, 2008).

Some analytical techniques are based on fluorescent modifications, especially those for studying protein-protein interaction or visualizing protein location in the cell. For visualization of these modifications a fluorescent microscope, which is able to illumine a sample with a single wavelength to excite florescence is needed. An example of such analytical technique, Fluorescence resonance energy transfer (FRET) can be considered. FRET is used for studying protein-protein interactions. Using confocal microscope and GFP-fused proteins we can study protein kinetics directly in the cell.

This method is called fluorescence recovery after photo-bleaching (FRAP) (Alberts et al, 2008).

Some modifications are smaller and they take place on the functional groups of side chains of amino acid in a protein. Alkylation may serve as an example of such modifications. Iodoacetamide is used as an alkylation reagent, which reacts with thiol groups and prevents cysteins to create disulfide bonds (Anson, 1940).

Protein modification is quite an active research field and any interesting discovery at in this field may be fundamental for new experimental methods or better understanding of protein structure, interactions and function. There is also a huge potential in optimizing older methods and making scientific research "easier".

2 Biotinylation:

2.1 Biotin:

Biotin is a hydrophobic molecule of a molecular weight of 244,31 Da. It acts as a vitamin and co-enzyme essential to all life forms. It is composed of a tetrahydroimidizalo ring fused with a tetrahydrothiophene ring. Valeric acid is attached to one of the carbons at the tetrahydrothiophene ring (Fig. 3). The empiric formula was determined by Du Vigneaud et al. (1941) and its structure by the same group in 1942 (du Vigneaud et al, 1942; Melville et al, 1942).

It is synthesized by plants, most of bacteria, some fungi and it plays an important role in metabolic pathways in many more complex organisms. In the metabolism, biotin can react only when it is covalently bound to protein. Biotin binds itself to enzymes that transfer carbon dioxide from bicarbonate to organic acids to form cellular metabolites. This group of enzymes has a key role in gluconeogenesis, lipogenesis, amino acid metabolism and energy transduction (Chapman-Smith et al., 1999a). In fatty acid synthesis acetyl-CoA carboxylase α (ACC)α is an important enzyme, which catalyses the binding of bicarbonate to acetyl-CoA and thus generates malonyl-CoA, which is involved in further steps of fatty acid synthesis. Acetyl-CoA carboxylase β (ACC)β in organisms serves also as a reservoir for biotin. Propionyl-CoA carboxylase (PCC), 3-methylcrotonyl-CoA carboxylase (MCC) and pyruvate carboxylase (PC) are other

enzymes where biotin is used as a co-enzyme, all these enzymes are located in the mitochondrial matrix (Zempleni et al, 2008).

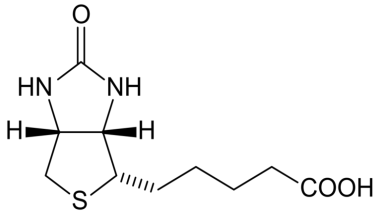


Figure 3: Biotin (vitamin B7, a.k.a. vitamin H).

Another function of biotin in living organisms involves in the modifications of histones. DNA chromatin is comprised mainly of DNA and DNA-binding proteins – histones and non-histone proteins. There are five types of histones: H1, H2A, H2B, H3, H4. All histones are rich in the amino acids lysine (K) and arginine (R), which are the foremost targets for biotinylation. Those amino acids also mediate the positive charge of histones at physiological conditions (Hassan et al, 2008). Histones contain a flexible tail, where many of the modifications take place. Specifically, these modifications take place at ε-amino groups, guanidino groups, carboxyl groups and hydroxyl groups. This tail is modified by covalent acetylation (Lee et al, 1993), methylation, phosphorylation, ubiquitination, poly(ADP-ribosylation) (Boulikas et al, 1990) and biotinylation (Kothapalli et al, 2005) and those modifications have influence on interactions between negatively charged DNA strand and positively charged groups of the amino acids in the histone tail. Depending on the way, how the histone is modified, for example trimethylation of K4 on histone H3 is associated with transcriptional activation of surrounding DNA, while dimethylation of K9 is an epigenetic mechanism for the transcription silencing (Hassan et al, 2008), biotin is bound to the ε-amino group through actions of holocarboxylase synthetase (HCS) in two step ATP-dependent reaction (Fig. 4). Recycling of biotin is provided by the enzyme biotinidase (BTD). This enzyme cleaves biotin from biotinyl-lysine or small peptides containing biotin (Wolf et al, 1983).

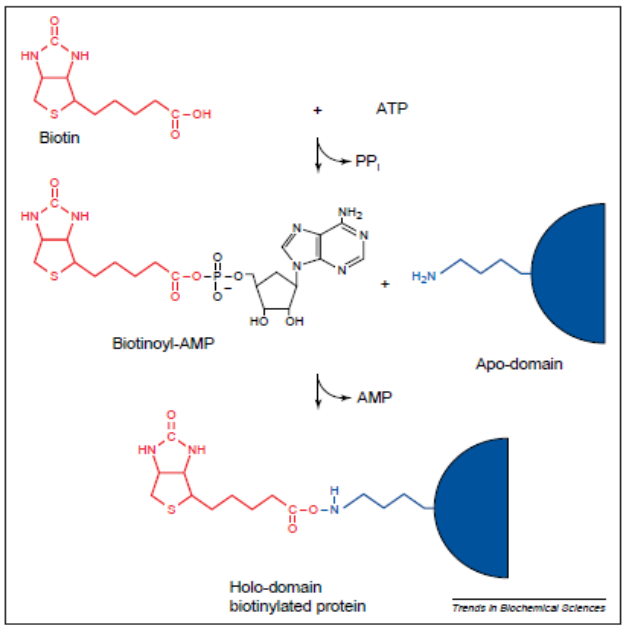


Figure 4: Holocarboxylase synthetase (HCS; a.k.a biotin protein ligase (BPL)) reaction.

2.2 Analytical use of biotin

Next advantage of biotin, for which it applicable in many analytical procedures, is its very-high-specific interaction with protein avidin and streptavidin. Thanks to that, we can use many technologies based on this interaction for detection, purification and immobilization of proteins in biochemistry, immunology, cell biology and biotechnology (Watanabe et al, 2007). Biotin is quite a small molecule, so there is no problem with its binding to any of the four equal places at avidin or streptavidin tetramer. This bond is about four orders of magnitude stronger ($K_d = 1{,}3 \times 10^{-15}$ M at pH 5.0) than a typical antigen-antibody interaction (Birkert et al, 2000). A plates or test tubes with avidin or streptavidin immobilized on the surface (Fig. 5) may be used in many cases. There is growing pressure on biochemical sensors these days. The scientists want cheaper sensors and faster methods for analysis, so now one can buy and use many types of biosensors and set things as desired. For measuring, the following methods can be used: surface plasmon resonance spectroscopy (SPR), quartz microbalance (QMB), grating coupler (GCP), resonant mirror (RM), or reflectometric interference spectroscopy (RifS). All these methods require specific gold- or glass-type surface of

the biosensors (Brecht et al, 1995). Usually there is no problem with distinguishing specific and non-specific interactions on the biosensor surface. However, if label-free samples are used, a problem may occur. Surfaces therefore must be modified with chained thiols, dextrans, ethylene glycols, and with other polymers and co-polymers (Birkert et al, 2000).

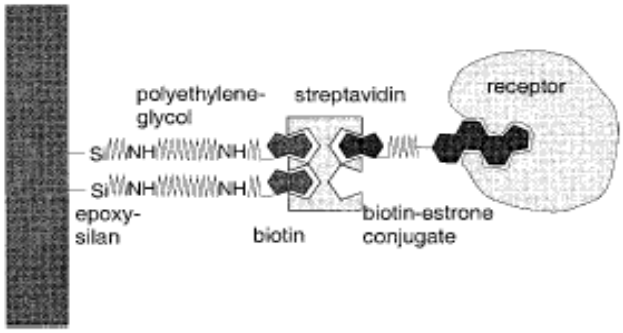


Figure 5: Example of immobilized streptavidin biosensor with bound analyte (Birkert et al, 2000)

Biotin – streptavidin interaction is examined on nanoparticles too. The scientists also consider the use of nanoparticles as a drug-delivery method. In one study they report that they have prepared small gelatin nano particles with avidin surface as a potential delivery method for lung cancer treatment and for accumulation of cis-platine in cancerous lungs (Tseng et al, 2007; Tseng et al, 2009).

Other studies describe the preparation of ZnO nanoparticles with a biotin-covered surface (SelegArd et al, 2010). Biotin can also be also used as a powerful tool for detection in a new field of studying 3-D structures of proteins and protein-protein interactions. When using modified cross-linkers of different lengths, it is easier to separate cross-linked proteins from proteins without cross-linker which simplify reading of mass spectra (Kang et al, 2009). Probes based on biotin can be used for the determination of virus glycoproteins in capsid structure. Maybe this technique will be efficient for the identification of amines which are involved in virus – host interactions or to localize conformational changes playing role in infection process (Sharp et al, 2006).

2.3 Biotinylation and compounds used for biotinylation

We can prepare biotinylated proteins by two different ways: first, is a solution by mixing a solution of a biotinyl reagent (in different solvents) and a solution of protein of our concern. Secondly, is its immobilized form e.g. protein exposed on the cell surface or the cells lining endothelia in vivo and ex vivo (Elia, 2008).

The biotinylation is mostly performed in solution. For doing that we can buy or obtain many different reagents from various suppliers. It depends on the desired conditions, but all reagents contain the bi-cyclic ring of biotin on one end and the other end they contain the pentanoic acid side chain, to which a linear spacer may (or may not) be attached through formation of an amide bond (Fig. 6).

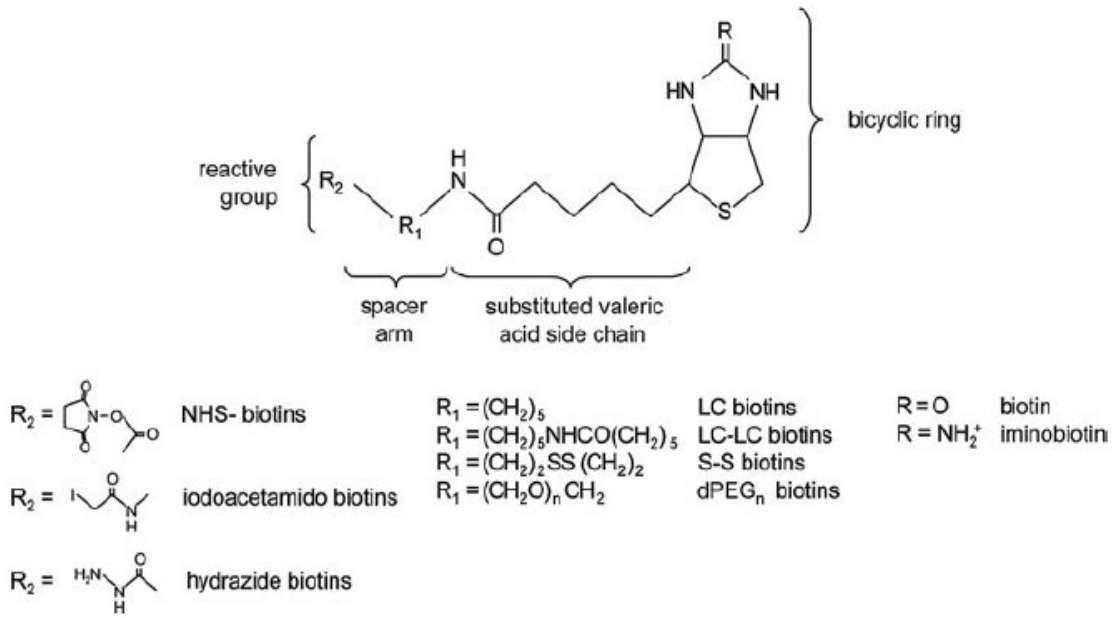


Figure 6: Some reagents for protein biotinylation (Elia, 2010)

Nowadays, many protocols describing protein labelling by biotin can be found on the internet.

Biotin primarily reacts with free ε-amino groups of lysine (K) and arginine (R) residues. Those amino-acids are normally abundant in most animal proteins and in many cases they occupy accessible places. Lysine is mostly not involved in protein function or in some protein-protein interactions, so the modification at this place does

not produce detrimental effects on the biological activity or other interactions (Elia, 2010).

N-hydroxy-succinimide esters (NHS-esters) of biotin are frequently used for biotinylation. For better solubility in water sulfo-NHS-esters are available, which exclude using dimethyl sulfoxide (DMSO) or dimethylformamide (DMF) as solvents. Special spacer arms were added to decrease steric hindrance, so now one can buy e.g. commercial NHS-LC-Biotin (spacer arm 22,4 Å) and NHS-LC-LC-Biotin (spacer arm 30,5 Å) (Fig. 7) (Diamandis et al, 1991).

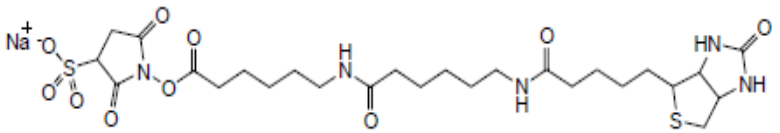


Figure 7: The structure of Sulfo-NHS-LC-LC-Biotin (spacer arm 30,5 Å) (*Pierce*)

The nature of the LC arm is quite hydrophobic so after cleavage of Sulfo-NHS moiety it starts to seek hydrophobic parts of the protein and it is lost for an effective interaction with avidin. This problem can be solved by using a spacer arm based on polyethylene glycol. The reagent can be purchased in many variants with the modification of highly folded regions of proteins (http://quantabiodesign.com). There is an advantage of using cleavable spacers. Those spacers are designed to be broken by reducing agent like 2-mercaptoethanol or dithiothreitol. They are available in both basic versions with an alkyl spacer or PEG spacer. (e.g. Sulfo-NHS-SS-Biotin; Sulfo-NHS-SS-dPEG4-Biotin). There are also other modifications which can be found, such as photo-cleavable linker or linker with a chromophoric tag (Quanta BioDesign).

Recommended pH for biotinylation is between 7 and 9, but the optimal value is arround pH = 7,5 (product protocols www.piercenet.com (25.1.2011); (Volkmann et al, 2009; Elia, 2010). Fig. 8 shows a scheme of the reaction.

For making experiments with biotinylated proteins under in vivo conditions we need to stabilize them. In plasma, the biotin-amide bond is quickly cleaved by

hydrolytic enzymes. To prevent this, a carboxylate group can be introduced at the alpha position to biotin-amide bond (Bogusiewicz et al, 2005).

Biotinylation can be also used for labelling other reactive groups in a protein. Via adding a reactive group to the reagent we can target other exposed residues of amino acid in protein. For example using biotin-LC-hydrazide or biotin hydrazide we can label carbohydrates, glycoproteins or some other glycoconjugates. At first, sample reacts with $NaIO_4$ to oxidize vicinal hydroxyl groups of sugars to aldehyde groups. The oxidized sample then react with hydrazide-containing biotin labelling reagent (Diamandis et al, 1991).

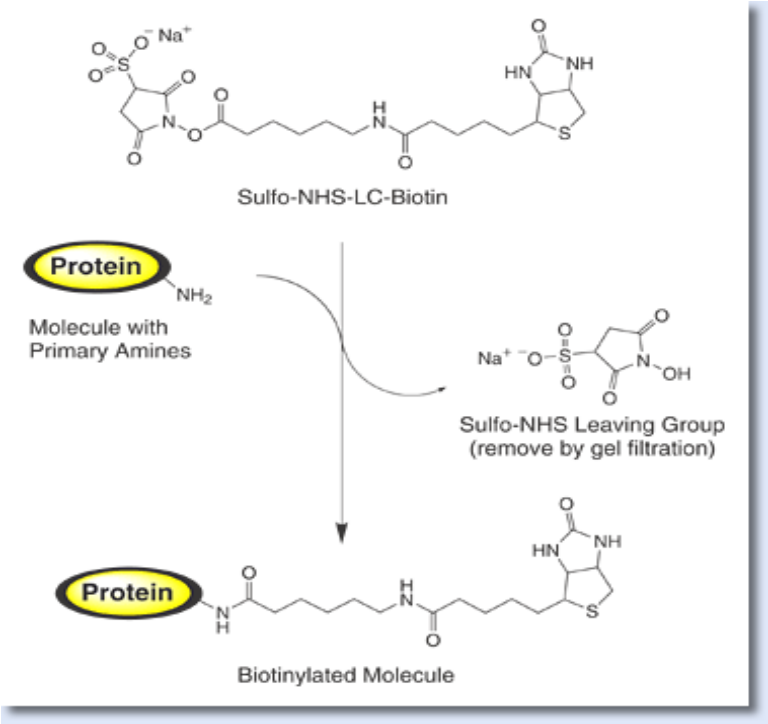


Figure 8: Reaction of Sulfo-NHS-LC-Biotin with protein primary amines. (www.piercenet.com 25.1.2011)

Another usual reactive group for protein derivatization is free sulfhydryl group on cystein residues. For such a derivatization, we can use three different reagent concepts. First, maleimide is very useful as it is super-reactive with free sulfhydryls at pH 7. The reaction depends highly on pH, because at higher pH values maleimide reacts

with preferentially primary amines and undergoes degradation. The second concept involves iodoacetate or iodoacetamide. This compounds react with sulfhydryls to produce a thioester group at pH 7,5-8,5. By lowering pH down to 6,9-7,0, we can observe a reaction with imidazole groups, but the reaction time has to be increased up to a week. The last reagent uses a pyridithiol group, which provides a mixed disulfide bond. Normally this reaction runs at neutral pH, but we can use wide-pH-range buffers (Elia, 2008).

For cell surface biotinylation serine residues of acyl carrier protein (ACP) can be used. ACP is bound to the protein of interest and then is biotinylated. Biotin is bound to co-enzyme-A (CoA) to form CoA-Biotin. Through phosphopantetheine transferase (PPTase), the 4'-phosphopantethein in the biotinylated CoA, is transferred to serine residue at acyl carrier protein (George et al, 2004). Nowadays almost every membrane protein or floating receptor can be labeled and this brings big benefits to studies of the membrane proteins.

2.4 Avidin and streptavidin

2.4.1 Avidin

Avidin is a glycoprotein that can be found mainly in egg white. It is a tetrameric protein with a molecular weight of about 68 kDa, its primary structure contains 128 amino acids and each subunit contains one biotin binding site (Fig. 9) (DeLange and Huang, 1971).

Each monomeric part of the avidin tetramer is based of a β-barrel constructed of eight anti-parallel β-strands. Most differences in their tertiary structure are thanks to six extended loops, which are connected together by β-strands together (Livnah et al, 1993). Two amino acid residues play the key-role in biotin binding to avidin (namely: Trp-70; Trp-110) (Gitlin et al, 1988). Lately Livnah et al. published how biotin is bound to avidin. They found that biotin replaces a part of solvent molecule and it is locked in by one of the exposed loops (residues 36 – 44; this loop connects β3 to β4 strands). Biotin is thus fixed inside of the β-barrel and than is stabilized by hydrophobic pocket, which is composed of two tryptophans (Trp-70; Trp-97) and two phenylalanines (Phe-

79; Phe-72) as depicted in Fig. 10. Trp-110 stabilizes an interaction on other monomeric subunit (Livnah et al, 1993). Gitlin et al also, according to previous studies, found that a modification of Tyr-33 by specific tyrosine reagent (*p*-nitrobenzensulphonyl fluoride (Nbs-F)) decreases the effect of binding biotin to avidin, which suggests that Tyr-33 is also involved in the binding pocket (Gitlin et al, 1990).

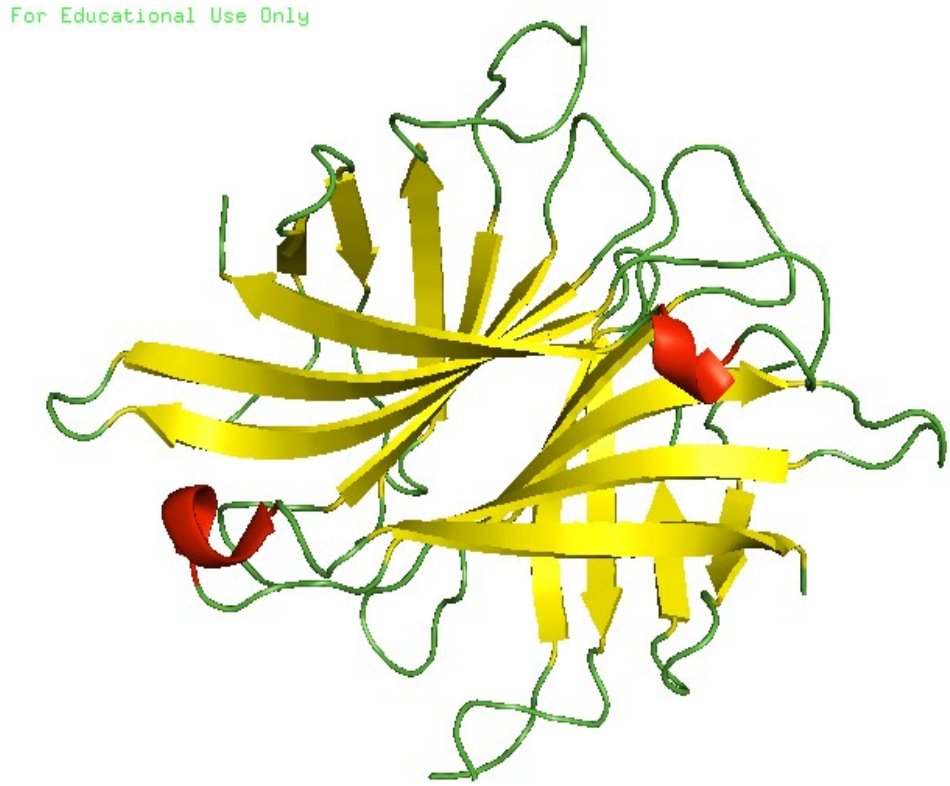


Figure 9: The crystal structure of avidin (PDB file 3FDC; Barker et al, 2009)

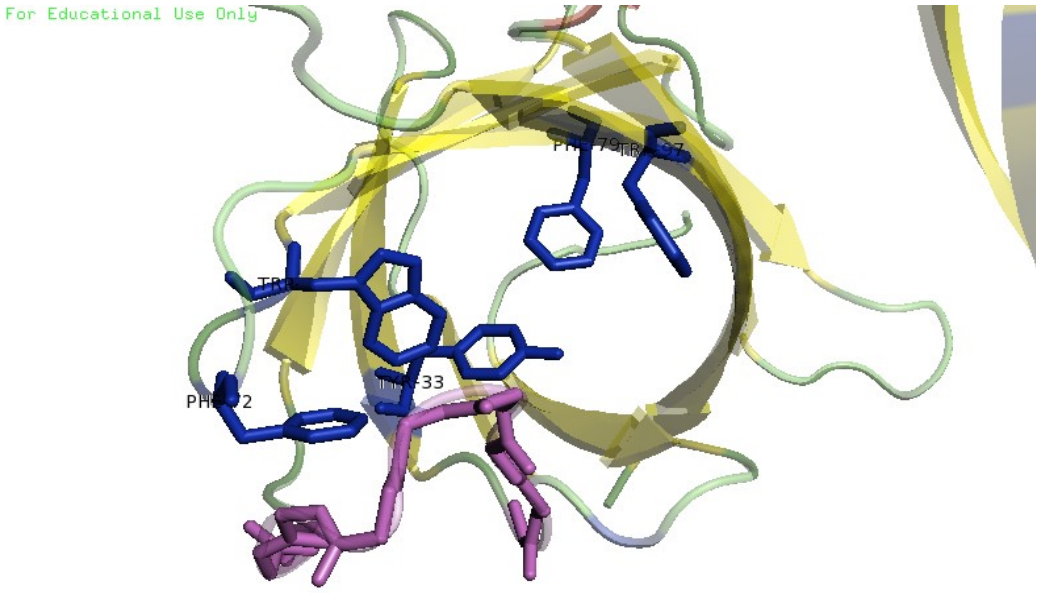


Figure 10: The binding pocket of avidin for biotin consists of a loop (residues 36-44; marked *purple*) and residues which stabilize biotin – Tyr-33; Trp-70; Phe-72; Phe-79 and Trp-97; marked in *blue*) (PDB file 2AVI; Livnah et al, 1993)

After the mutation of Trp-110 to Lys it was found that Trp-110 residue is important in avidin oligomerization, more 1 → 2 interaction (interaction between monomer 1 and monomer 2) (Laitinen et al, 1999). In another study they tried to find some answers to the question of how can ligand bound to the monomeric protein help in assembling of the multi-protein complex. Interactions between monomeric avidin subunits 1 → 3 are based on hydrophobic interactions of Val-115; Met-96 and Ile-117. Interaction 1 → 4 appears through Asn-69 on one monomer and Asn-54 on the other one, among which there is water molecule. After mutations V115A, M96A, I117A, it was found, that avidin is not able to form tetramer. But after adding biotin to the solution, avidin started to assemble its tetrameric normal form. This was a proof that the ligand can influence the quaternary structure of proteins. In this case it was probably caused by an interaction of biotin with Trp-110 from adjacent monomer. This points out importance of 1 → 2 interaction (Laitinen et al, 2001).

Avidin is glycosylated on Asn-17 and the carbohydrate moiety on each subunit is

composed of four or five residues of mannose and three residues of N-acetyl-glucosamine (DeLange, 1970). But this composition of sacharides changes a lot, as has been demonstrated by Bruch and White (Bruch and white, 1982). The carbohydrate moieties do not influence biotin binding to avidin (Hillar et al, 1987).

2.4.2 Streptavidin

Streptavidin is a bacterial analogue of avidin. Its molecular weight is 52,8 kDa. It was isolated from the bacterium *Streptomyces avidinii*. The sequences of avidin and streptavidin show 30% identity and 41% similarity (64% (74%) within homologous segments and 7% (17%) outside of these segments (Livnah et al, 1993). Streptavidin forms tetramers like avidin as can be seen in Fig. 11 and contains one binding site for biotin on each monomeric subunit.

The tertiary structure of streptavidin and avidin are almost identical, both are based on eight β-strands which are folded to an anti-parallel β-barrel. The binding place for biotin is also based on the same principal. Both proteins differ only in a few residues and their exact position in streptavidin is function of Phe-79 substitution by Trp-92 and in avidin it is moreover Phe-72, which does not have equivalent in streptavidin (Livnah et al, 1993).

There are more avidin-like proteins known. In *Bradyrhizobium japonicum*, a nitrogen fixing bacterium which is symbiotic on the soybean, this analogue is named bradividin. There is no significant similarity in the amino acid sequence, but for the tight binding of biotin is also mediated by Trp-110, which interacts with biotin on the other subunit (Helppolainen et al, 2008). Frog avidin, called xenavidin, was discovered in *Xenopus tropicalis,* its tertiary structure and binding pocket are also very similar to those of avidin and streptavidin (Maatta et al, 2009).

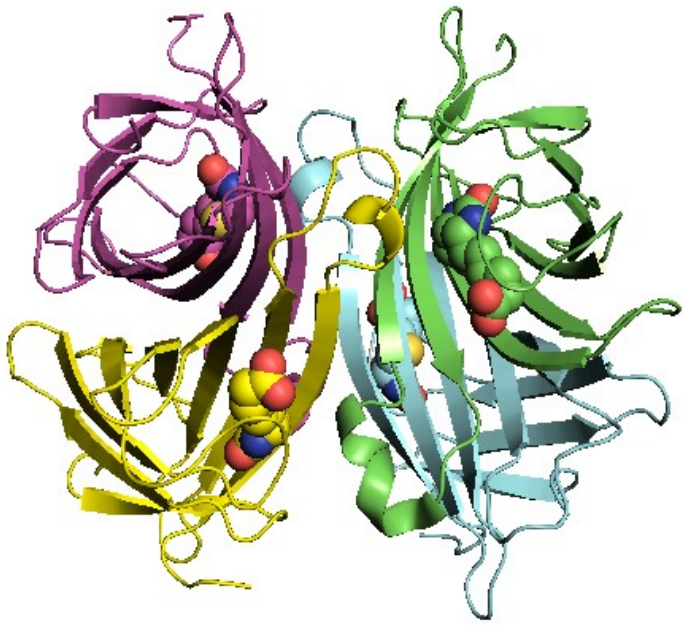


Figure 11: The tetrameric structure of avidin binding biotin (red). (PDB structure 3MG5)

3 Androgen receptors:

3.1 Physiological function:

Androgen receptors form a part of the nuclear receptor super-family belonging to group of the most important transcriptional regulators in animals (metazoans). Nuclear receptors are involved in different functions such as reproduction, development and metabolism (Laudet and Gronemeyer, 2002). All receptors are activated through the small signal molecules which differ greatly in chemical structure and function. Among the signal molecules, there are steroid hormones, thyroid hormones, retinoids and vitamin D (Fig 12.).

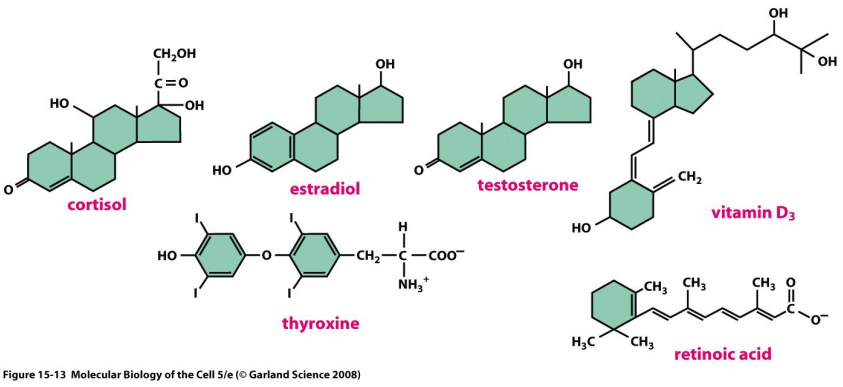


Figure 12: Selected signal molecules (Alberts et al, 2008**)**

Nuclear receptors act as a ligand-depended transcriptional factor and through that they provide a link between signal molecules and transcriptional response in the cells and organisms. There are many nuclear receptors with the function of which we don't know. Those receptors are called orphan receptors and finding of their ligands is a very active research field nowadays. Since signal molecules are quite small and easy to modify, scientists also want to discover their analogues as potential drugs for curing diseases (Heemers and Tindall, 2007).

Among the enormous variety of proteins, there are only five steroid receptors: glucocorticoid and mineralocorticoid receptors, estrogen (α; β^2), progesterone and androgen. Androgen receptor is a water soluble protein, which is a common attribute to steroid receptors and its function is regulated by androgens.

Human androgen receptors are mostly expressed in skeletal muscles, central nervous system (CNS), liver and prostate, there is the highest observed level ever. High levels were also found in adrenal gland and epididymis. The most important activators are testosterone and 5α-dihydrotestosteron (5α-DHT, DHT) (Gao et al, 2005). Physiological function of androgen receptors is to control male sexual differential *in utero* and secondary sexual signs in puberty. In adults, they are responsible for functioning libido, bone mineral density, muscle mass and strength, erythropoisis and spermato-genesis (Heemers and Tindall, 2007).

3.2 Structure of androgen receptor:

Androgen receptor contains 919 amino acids and its molecular mass is approximately 99 kDa. Some structure domains are common to all nuclear and androgen receptors. Most conserved is the DNA-binding domain (DBD) which includes zinc-finger motives for binding to major grooves at DNA. The second part is ligand-binding domain (LBD), which is very well conserved through nuclear super-family with only few differences. LBD is located at the C-terminal end of the androgen receptor (Duff et al, 2006). The biggest part of the receptor is N-terminal domain (NTD). This domain is more divergent than the other ones and it contains a ligand-independent transcription activation function 1 (AF1). Second ligand-independent transcription activation function (AF2) is located at LDB. LBD in nuclear receptor consists of twelve α-helices (Haelens et al, 2007). The last and also the smallest part of androgen receptor is a hinge region (LBDh), located at androgen receptor residues 628-669 aa. Hinge region is highly flexible and it is between LBD and DBD (Fig 13.) (Deeb et al, 2008).

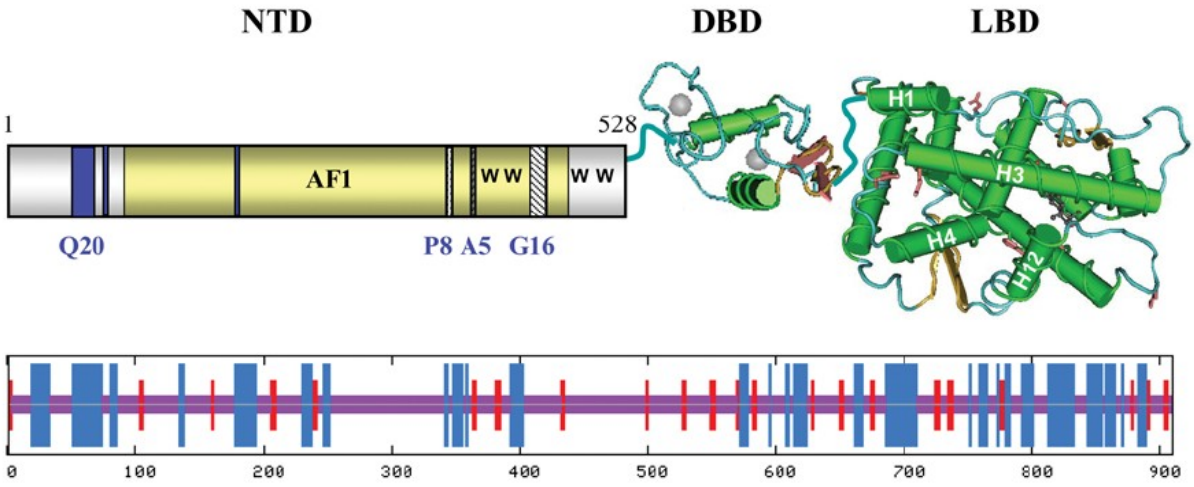


Figure 13: The first panel shows the domain organization of androgen receptor. The second panel shows prediction of androgen receptor structure (blue bars = alfa-helix; red bars = beta-strands) (Duff et al, 2006).

3.3 Principles of androgen receptor function:

Across all amino-acid sequence there are several small structural motifs, which are involved in protein-protein and protein-ligand interactions. Without presence of ligand the androgen receptor is located in the cytoplasm, complexed with heat-shock proteins and immunophilin chaperones – Hsp90 Hsp70, Hsp56 and p23 (Fang et al, 1996; Heemers and Tindall, 2007). After ligand binding, androgen receptors undergo dimerization and then the homodimer is translocated to the nucleus (Fig 14.). Nuclear localization sequence (NLS) is recognized by α-importin which serves as an adaptor for β-importin. This complex is moved to the nucleus through nuclear-pore complexes (NPCs). In nucleus, androgen receptor is released thanks to the Ras family GTPase Ran (Cutress et al, 2008). It was reported that some mutations in highly conserved NLS region are associated with androgen insensitivity syndrome or prostate cancer (Tilley et al, 1996). In nucleus, androgen receptor binds to the DNA through several mechanisms. The DNA-binding specificity is mainly due to P-box at zinc finger of DBD. Residues in this box recognize sequence of 5'-TGTTCT-3'. Another interaction is provided by androgen-response elements (AREs) which are located in the promoter and enhancer regions of target genes. To this complex can bind parts of transcription pre-initiation complex or some other elements and which may enhance an interaction (Haelens et al, 2003).

3.3.1 Intra-molecular interactions

Intra-molecular interactions are also very important for a correct functioning of androgen receptor. Transactivation function 1 (a.k.a. AF1), located at NTD of androgen receptor, is inhibited by interaction with LBD. When those two domains are separated, androgen receptor remains fully active. Two motifs on androgen receptor NTD – 23-FQNLF-27 and 433-WHTLF-473 – are underlying interaction with AR LBD and those short sequences are also important for transcription of some genes (Schaufele et al, 2005; He et al, 2002). AF-1 consists of two transactivation units (TAUs), TAU 1 and TAU 5 (Jenster et al, 1995).

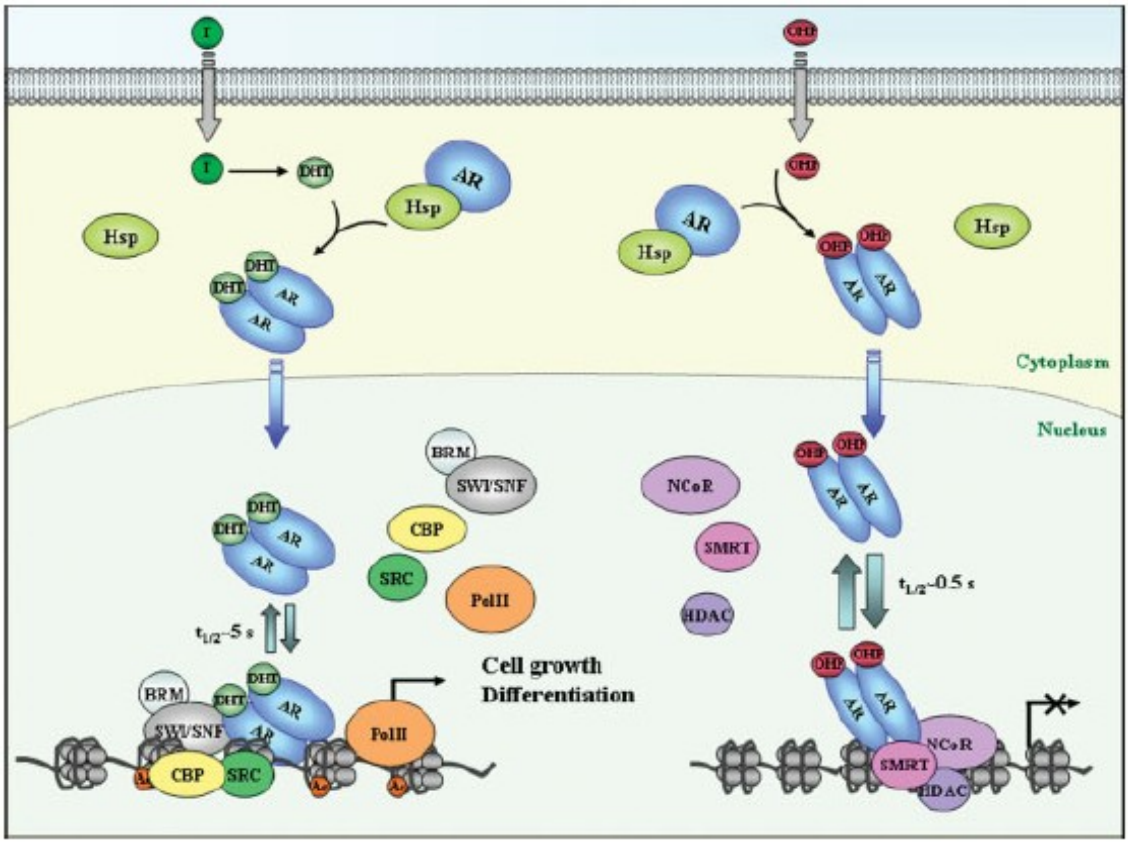


Figure 14: Hit-and-run model for AR action. The left panel of the picture depicts an agonist of the receptor, in this case 5α-dihydrotestosteron (DHT). Upon entering of DHT to the cell, androgen receptor undergoes dimerization and enters the nucleus. Than the dimer interacts with other proteins and DNA, which is required for transcription. The right panel on the picture shows the inhibition of this pathway by binding of antagonist hydroxyflutamide (OHT). Binding of OHT also triggers dimerization, but after entering the nucleus, different proteins are involved in the reaction and transcription is consequently, jam up (Kaarbø et al, 2007).

The pivotal role in NTD-LBD interaction is played by the residues 3-36 in the NTD (AR_{3-36}) including the above-mentioned 23-FQNLF-27 motif. More details were related by pull-down assays, which showed that segment AR_{3-13} is not essential but it modulates NTD-LBD interaction. Conversely, the segment AR_{16-36} proved to be critical for this interaction. This can be caused by the fact that those residues can be fold in the

long amphipathic helix (Fig. 15.) and it also contains two important hydrophobic motifs – 23-FXXLF-27 and 30-VXXVI-34 (He et al, 2000; Steketee et al, 2002). On LBD, there are essential residues K720 and E897, but further experiments showed that K720 can be replaced with another amino acid without significant effects (Berrevoets et al, 1998).

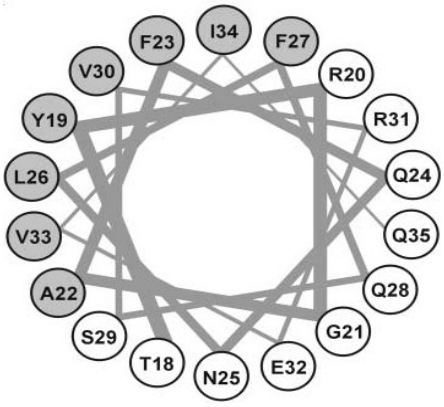


Figure 15: Interacting amphipathic helix AR_{18-35}, grey circles show hydrophobic amino acids (Steketee et al, 2002).

The polyglutamine part of NTD is also involved in a proper functioning of androgen receptor. Its regular length is between 8 to 31 repeats, on average the value is twenty. When the chain is extended up to 40 and more amino acids, it starts to cause problems. It may result in spinal and bulbar muscular atrophy (a.k.a. Kennedy disease), an X-linked pathology characterized by neurological features and late onset symptoms of androgen insensitivity syndrome (Lumbroso et al, 1997; la Spada et al, 1991). Beyond, shortening of the polyglutamine chain can cause predisposition to prostatic neoplasia (Giovannucci et al, 1997). Another polymorphic repeat is a polyglycine motif, containing 10-27 glycine residues in normal population (Buchanan et al, 2004). Activation function 2 (AF2) (Fig. 16) represents a small hydrophobic groove on LBD which is formed by a ligand binding and shunt and stabilization of helix-12. This small pocket mediates major protein-protein interaction within the nuclear receptor super family. Nevertheless, androgen receptors are little bit different. While nuclear receptors strongly bind motifs LXXLL from other receptor or co-activators, in androgen receptor

motifs FXXLF are bound more tightly (this sequence corresponds with the motif 23-FQNLF-27, which can be found on AR NTD, Fig. 17) (He et al, 2000). A precise mechanism of interaction and competition between NTD and LBD has not been elucidated. Eva Estebanez-Perpina et al. surprisingly found a new binding site for small molecules in LBD. They tested new surface-interacting compounds for blocking the active site of AF2 and using X-ray crystallography they found that some compounds were bound to the new site BF3. BF3 is made by a junction of helix 1, loops of helices 3-5 and helix 9 (Estebanez-Perpina et al, 2007). Such results indicate that in this field still there is a lot of questions that remain without answer.

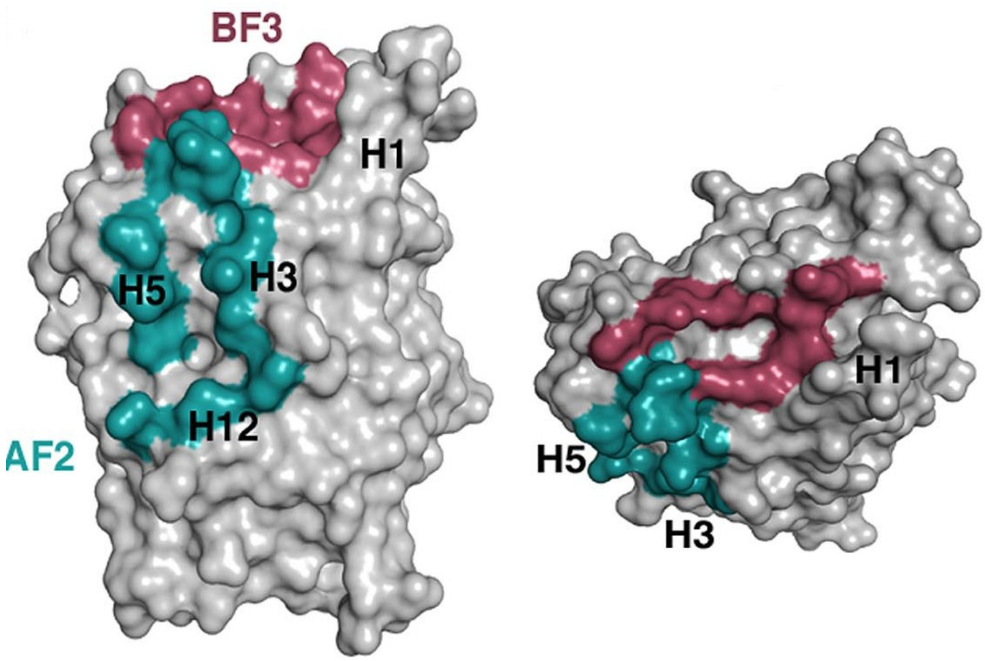


Figure 16: Binding pocket of AR AF2 and BF3. It is formed by specific residues of the helices 3, 5 and 12. (Estebanez-Perpina et al, 2007)

The significant importance of AF2 for binding FXXLF motifs is prove by the fact that in nature there are some mutations in this motif causing resistance to androgens. Unfortunately such a resistance has no effect on high affinity binding of androgens resulting in androgen insensitivity syndrome (AIS). AIS leads to various stages of incomplete masculinization of external genitalia, which depend on the extent of mutation (Langley et al, 1998). Crystal structures show that the AF2 binding site can

be used also for LXXLL motifs located at the steroid receptor co-activator (SRC)/p160. Whereas AF2 prefers FXXLF motifs, LXXLF binding motifs are competitively inhibited by FXXLF (He et al, 2004a). If protein p160 is in abundance in the cell, it can highly increase the transcriptional activity of androgen receptor, pointing that malfunction of androgen receptor is not necessarily caused by mutations in the receptor amino acid sequence.

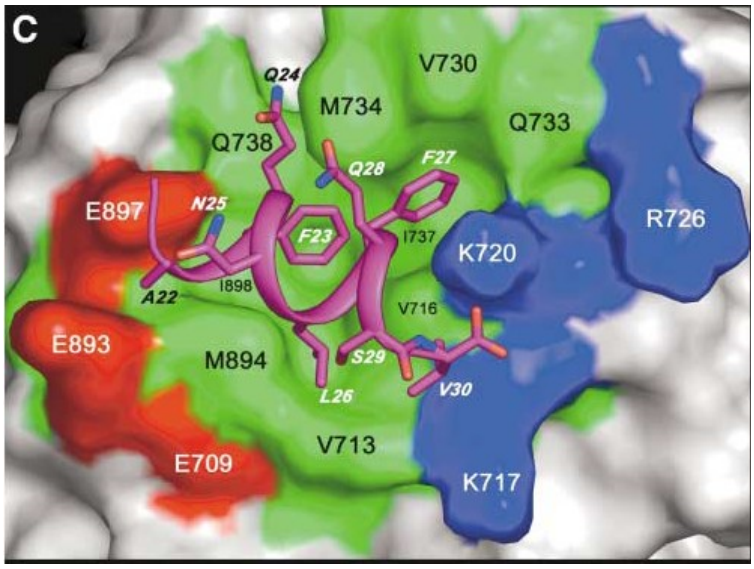


Figure 17: Surface of AF2 with AR_{20-30} (contains 23-FONLF-27) motif bound. Blue colour represents positively charged amino acid residues and red colour negatively charged ones (He et al, 2004b).

More unique attribute is an interaction with two high affinity biologically active hormones – testosterone and 5α-dihydrotestosteron (DHT). It is hard to say which one is more active, because of the similar instability of androgen receptor with the bound testosterone to androgen receptor without ligand. Nevertheless, a ten-times higher concentration of testosterone is required compared with DHT for the same activation (Grino et al, 1990). Testosterone dissociates from androgen receptor three-times faster than DHT. This leads to the weaker binding of FXXFL motifs, which explains less activity of testosterone (Askew et al, 2007).

3.3.2 Ligand-binding pocket in LBD

Ligand-binding pocket consists of many amino acid residues. Still, it is not clear how many of them are really involved in the binding of androgens and total stabilization of LBD. Jasuja et al discovered that AR-LBD passes through an intermediate state where it is partly unfolded, suggesting that this is probably important for loading co-activators and/or co-represors. Their data also reveal that binding of a ligand e.g. DHT stabilizes some other residues, which are not directly involved in binding of DHT, but they participate in forming co-activators binding grooves (Jasuja et al, 2009). LBD is mainly composed of hydrophobic amino acids, but the ligand binding pocket contains also a number of hydrophilic amino acids through which can establish hydrogen bonds with the steroid base of androgen moiety and hold it in its place. The binding site consists of residues belonging to helices H3, H4, H5 and H11 and a β-strand, which is located between helices H5 and H6. An answer to the question "how can androgen receptor bind structurally different ligands?" may reside in mobility of side chains of hydrophilic amino acids (Pereira de Jesus et al, 2006). The most important residues providing androgen stability in the binding pocket are Arg752 and Gln711 (Fig.18). These two residues are in a direct contact with the cycle A of steroid moiety and they are able to create interaction with oxygen atom. The importance of Arg752 is supported by the fact that in the case of mutation (e.g. Arg752Gln) androgen receptor is affected activity. Males with this mutation are affected by genetic disorder – androgen insensitivity syndrome (AIS) (Sakai et al, 2000; Pereira de Jesus et al, 2006).

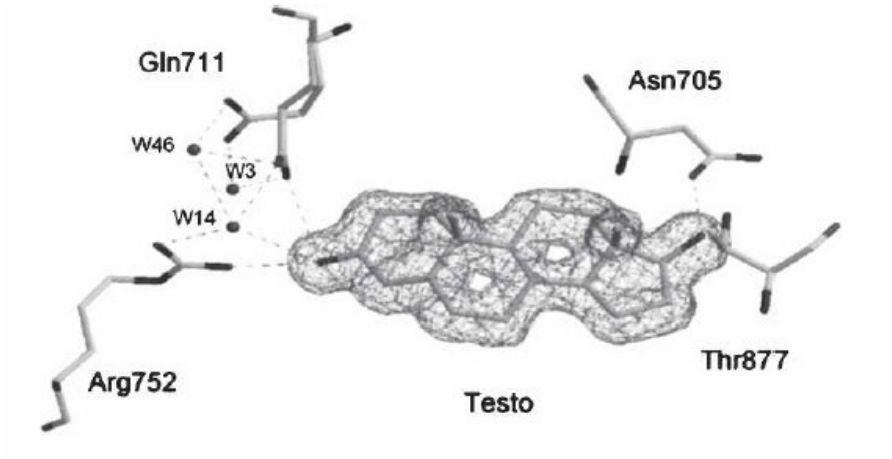


Figure 18: Electron density map of the ligand-binding pocket of LBD and residues of interest with testosterone (Pereira de Jesus et al, 2006).

Other studies suggested that also Met745 is involved in interactions with the ring A. An interaction with the ring C of steroid moiety is provided by main chain of Leu704 and side chain Asn705. The contact with C18 of DHT and the Oγ1 of Thr877 is unique for wild-type. Asn705 creates a hydrogen bond with hydroxyl group on the ring D in the case of DHT. The last interactions are provided by water molecules inside the binding pocket (Sack et al, 2001; He et al, 2006).

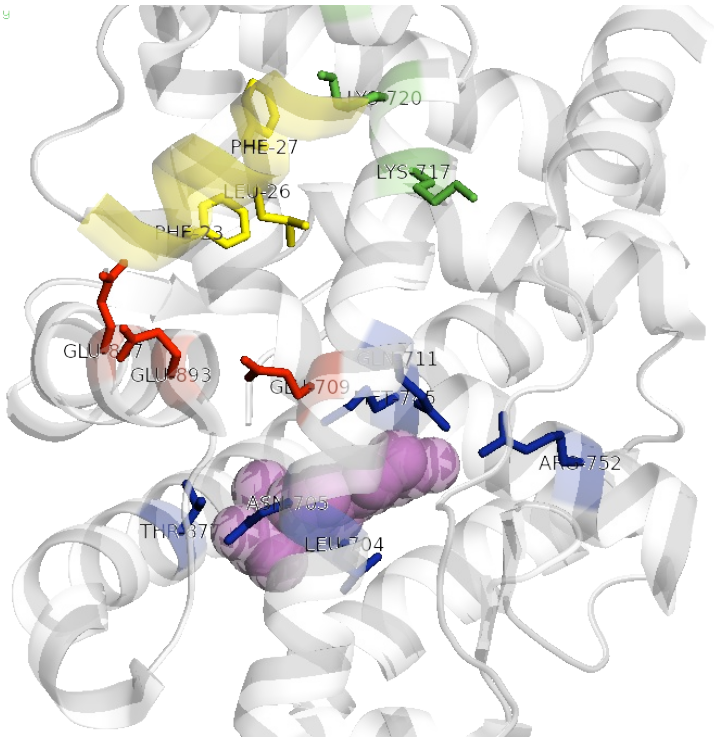


Figure 19: Survey of important residues in LBD involved in interactions with ligands. *Blue* – residues at ligand binding pocket – Leu704, Asn705, Gln711, Met745, Arg752, Thr847. *Red* – negatively charged residues in AF2 – Glu709, Glu893, Glu897. *Green* – positively charged residues in AF2 – Lys717, Lys720, Arg726. *Yellow* – residues in the interacting FXXLF motif. (PDB file 1XOW – He et al, 2004b).

3.4 Co-activators and co-represors of androgen receptors – basic overview:

Among many proteins interacting with androgen receptor, there are also transcriptional factors that are necessary for loading of RNA polymerase II and initiation of transcription. These proteins form one large group.

The next group of interacting proteins are androgen receptor co-regulators. Nowadays we know around two hundred different co-regulators. The first discovered activator (1995) was SRC-1 (steroid receptor co-activator 1) (Oñate et al, 1995). Details can be found in online databases – the AR gene mutation database (http://androgendb.mcgill.ca/) and the Nuclear receptor Signalling Atlas (NURSA)

database (http://www.nursa.org/). One of the subgroups of co-regulators is represented by several proteins involved in chromatin remodelling complex. As an example, androgen receptor interacting protein (ARIP) 4 in an interaction with androgen receptor zinc-finger region (Domanskyi et al, 2006). BAF57 (BRG1-associated factor 57) is another component of chromatin remodelling complex. It binds directly to the androgen receptor upon ligand binding. Function of BAF57 is associated with SWI/ANF ATPase activity and cooperation with other co-activators (Link et al, 2005). Basically, we can say that the large majority of this group is represented by activators and they are somehow associated with the transcriptional machinery.

Co-activators, another group of interacting proteins, have histone acetylase activity (HAT) and histone deacetylase activity (HDAC). Tat interactive protein 60 kDa (Tip60) may serve as an example. It interacts with androgen receptor ligand binding domain and enhances transcription by histone acetylation, it acetylates androgen receptor as well (Brady et al, 1999). On the other hand, some HDACs of classes I and II act as repressors of androgen receptor transactivation. The recruitment of HDACs to the androgen receptor transcriptional complex is usually vicarious. It is mediated by association with multi-subunit co-repressor complexes, such as NCoR and silencing mediator of retinoid and thyroid receptors (SMRT) (Wang et al, 2005).

A few co-regulators are involved in the ubiquitination pathway and they posses E3 ligase activity (E6-AP, Mdm2, PIRH2, SNURF/RNF4, Chip). Proof of the ligand-dependent recruitment of promoter region of androgen receptor genes is E3 ligase E6-associated protein (E6-AP). E6-AP interacts with ligand binding domain of androgen receptor after binding of ligand and through that increases transactivation function of androgen receptor (Khan et al, 2006). Some proteins that act in splicing interact with AF-1 – PSF (polypyrimidine tract-binding protein-associated splicing factor), PSP1 (paraspeckle protein 1) and PSP2. Component of U1snRNP (U1 spliceosomal RNA) prespliceasome directly interacts with androgen receptor NTD (Ishitani et al, 2003).

Methyltransferases and demethylases are also associated with androgen receptor transcriptional functions. Co-activator-associated arginine methyltransferase 1 (CARM-1), protein arginine methyltransferase (PRMT-5), and histone methyltransferase acting as a H3R17 all interact with steroid receptor co-activators (SRC). Due to the fact that they act like this, they were marked as second co-activators of androgen receptor

(Heemers and Tindall, 2007). The small ubiquitin-related modifier SUMO is also involved in enhancing of the transcriptional activity of androgen receptor. Three different SUMO enzymes influence the function of the androgen receptor. SUMO 1 decreases transcriptional activity, while SUMO 2 and SUMO 3 enhance the activity (Poukka et al, 2000). Few members of SUMO-specific proteases also influence the control of androgen receptors in almost the same way like SUMO. SUMO1/sentrin-specific protease 1 (SENP1) has also a positive effect on the activity, whereas SENP2 and SEPN3 show just a modest effect on androgen receptor (Cheng et al, 2004). Among proteins that control the cell cycle, there are some, which act as androgen receptor co-regulators. One of them is proapoptotic caspase-8, which represses androgen receptor-dependent gene expression. Caspase-8 directly interacts with FXXLF and WXXLF motifs in NTD (Qi et al, 2007).

Another group of interacting partners are kinases and phosphatases. Male germ cell-associated kinase (MAK) may serve as an example. MAK directly interacts with androgen receptor and it enhance transactivation potential in an androgen- and kinase-dependent manner. AR-interacting nuclear protein kinase (ANPK) interacts with DBD-hinge region (Ma et al, 2006; Moilanen et al, 1998).

Androgen receptors are of course influenced by other nuclear receptor activators and repressors. Amidst activators, there are Asc-1 (activating signal cointegrator-1) (Lee et al, 2002), Asc-2 (Goo et al, 2004), components of the Trap/Mediator complex (Wang et al, 2002), CoCoA (coiled-coiled co-activator) (Yang et al, 2006), NRIP (nuclear receptor co-regulatory protein) (Tsai et al, 2005), PNRC (proline-rich nuclear receptor co-regulatory protein) (Zhou et al, 2000), TIF-1α (Teyssier et al, 2006), MRF1 (modulator recognition factor 1) (Georgescu et al, 2005), PDIP1 (PPARγ-DBD-interacting protein 1) (Tomaru et al, 2006), Zac1 (Huang et al, 2000), GT198 (Ko et al, 2002) and ARA70 (Yeh et al, 1996). Co-repressors which are common to nuclear receptors and androgen receptors include: Alien (Moehren et al, 2007), AES (amino-terminal enhancer of split) (Yu et al, 2001), components of SMRT and NCoR repressor complexes (Zhu et al, 2006), RIP140 (receptor interacting protein 140 kDa) (Carascossa et al, 2006), PATZ (POZ-AT hook-zinc finger protein) (Pero et al, 2002) and TGIF (5'TG3' interacting factor) (Sharma et al, 2001).

Many more transcriptional factors play a role of androgen receptor co-regulators.

Table 1 shows a summary of such compounds (Table 1). Some of those factors bind directly to special motifs of DNA, which are scattered between binding sites of androgen receptor and AREs. Together they co-regulate transcription of androgen receptor genes (Heemers and Tindall, 2007).

Transcriptional factor	A/R	Direct/indirect
AML3/CBFα1	A	Direct
AP-1	R	Direct-NTD, LBD
ATF2	R	Direct-DBD
Brn-1	A/R	Direct
c-Jun	A/R	Direct-DBDh, DBD-LBD
C-rel	R	Direct
c/EBPα	R	Direct-NTD, DBDh
Dax1	R	Direct-LBD
EGR1	A	Direct-NTD
ERα	R	Direct-NTD
FKHR	R	Direct-NTD, LBD
Foxa1	A	Direct-DBDh
Foxa2	A/R	Direct-DBD
FoxH1	R	Direct-NTD-DBD
GATA-2	A	ND
GATA-3	A	ND
GR	R	Direct-DBD
HoxB13	R	Direct
NF1	A	ND
Oct-1	A/R	Direct-DBD
Oct-2	ND	Direct-DBD
Pod-1	R	Direct-DBDh
p53	R	Indirect
PDEF	A	Direct-DBD
RelA	R	Direct-NTD-DBD
RXR	A/R	Direect-LBD
Sox9	A/R	Direct-DBD
Sp1	A	Direct
SRY	R	Direct-DBD
SHP	R	Direct-NTD, LBD
SF1	A	Direct-DBD
TR2	R	Direct
TR4	R	Direct-NTD, DBD, LBD
USF2	A	Direct

Table 1: A summary of transcription factors involved in the activity of androgen receptor. A/R = activator (A), repressor (R) or both. Direct/indirect = transcriptional

factor interacts directly or indirectly with androgen receptor (ND = not determined); DBDh = DBD-hinge region (table based on Heemers and Tindall, 2007).

Finally, also interactions with chaperones and co-chaperones are very important. They upkeep androgen receptors stable in their in-active form in the cytoplasm. Thanks to this fact, they down-regulate transcription of androgen receptor genes and they are an important part of the life cycle of the androgen receptor (Prescott and Coetzee, 2006). Proteins binding to the LBD maintain androgen receptor in a high ligand affinity state. At the beginning, the activation of proteins as Hsp40 (Ydj1), Hsp70 (Hsc70), Hip, Hps90, Hop and p23 and Hsp40 is necessary for ligand binding (Fliss et al, 1999).

3.5 Androgen receptor pathology:

3.5.1 Prostate cancer

The most frequently diagnosed non-skin cancer in men is prostate cancer. It is also the third leading cause in deaths of men. First demonstrations on the importance of androgens in the formation and progress of prostate cancer appeared about seventy years ago (Huggins and Hodges, 1941). Prostate cancer belongs to the group of hormonally regulated malignancies, in this case androgen receptors play a major role. Most troubling part of the progression of prostate cancer is a state when tumor passes from the androgen-dependent state to androgen-independent (AI), which now resists to any effective treatment. In the late stages rise hormone-refractory (HR) tumors and androgen receptors are activated by very low levels of androgens (those levels correspond to the castration levels) (Kung et al, 2009). Latest observations summarize over ninety missense mutations from clinical specimens (Fig. 20) (Gottlieb et al, 2004). Nowadays, standard treatments are based on blocking androgen receptor activity by two ways: reducing of expression of androgen receptors and its inactivation by androgens antagonist. However, this therapy usually fails, indicating higher levels of prostate-specific antigen (PSA) in tumors switching to an alternative mechanism of activation (Steinkamp et al , 2009). In tumors we can observe also over-expression of TIF2 and SRC1 that enhances androgen receptor activity at physiological levels of androgens.

One of the most common missense mutation is T877A identified in the LNCaP cell line. This mutation allows androgen receptor to be activated by other steroid hormones (e.g. progesterone, estrogen, adrenal androgens) (Han et al, 2005). Another examples of pathological mutations are A229T and E231G in short motif of ARNSM. This motif is unique to androgen receptor NTD. Both mutations increase the ligand-independent basal activity, and E231G also enhances response to ARA160 and ARA70 co-activators (Han et al, 2001).

3.5.2 Androgen insensitivity syndrome

There are two versions of androgen insensitivity syndrome (AIS), depending on their complexity. Both complete AIS (CAIS) and partial (PAIS) leads to the organ insensitivity to androgens. Adult patients suffering from CAIS are tall women with female distribution of adipose tissue, small or no sexual hair, female breasts and external genital development. At an early age, patients can be present with testes descend into the inguinal canals or labia. During adolescence, people affected by this disease, develop like women, but without menses and sexual hair, in contrast with normal breast development (Wisniewski et al, 2009). Those patients are treated as women, although they are genetically males with testes.

Androgen receptors contain two polymorphic repeats - polyglutamine and polyglycine motifs with different length. It was discovered that a shortening of polyglycine repeat and elongation of polyglutamine repeat together with A645D mutations within the hinge region leads to AIS. Nevertheless, A645D appeared in many tested subjects, from normal and healthy people to cases with CAIS (Werner et al, 2006).

3.5.3 Kennedy disease

Kennedy disease, also known as spinal bulbar muscular atrophy (SBMA) is another disease caused by malfunctioning of androgen receptor. It is a progressive neuromuscular disease. The mechanism of action is still not clear, but it seems that elongation of poly glutamine motif in androgen receptor is toxic for motoneurons

resulting in muscular atrophy (Mo et al, 2010). This toxicity is also increased by activation of unfolded protein response and transcriptional dysregulation (Yu et al, 2006). Patients suffering from symptoms have facial and perioral contraction fasciculation and dysarthria (motor speech disorder resulting from neurological injury), dysphagia (medical term for the symptom of difficulty in swallowing, related to dysphasia) and dysphonia (disorders of the voice, sound organs). They also have full-fledged sings of partial androgen insensitivity syndrome (Lund et al, 2001).

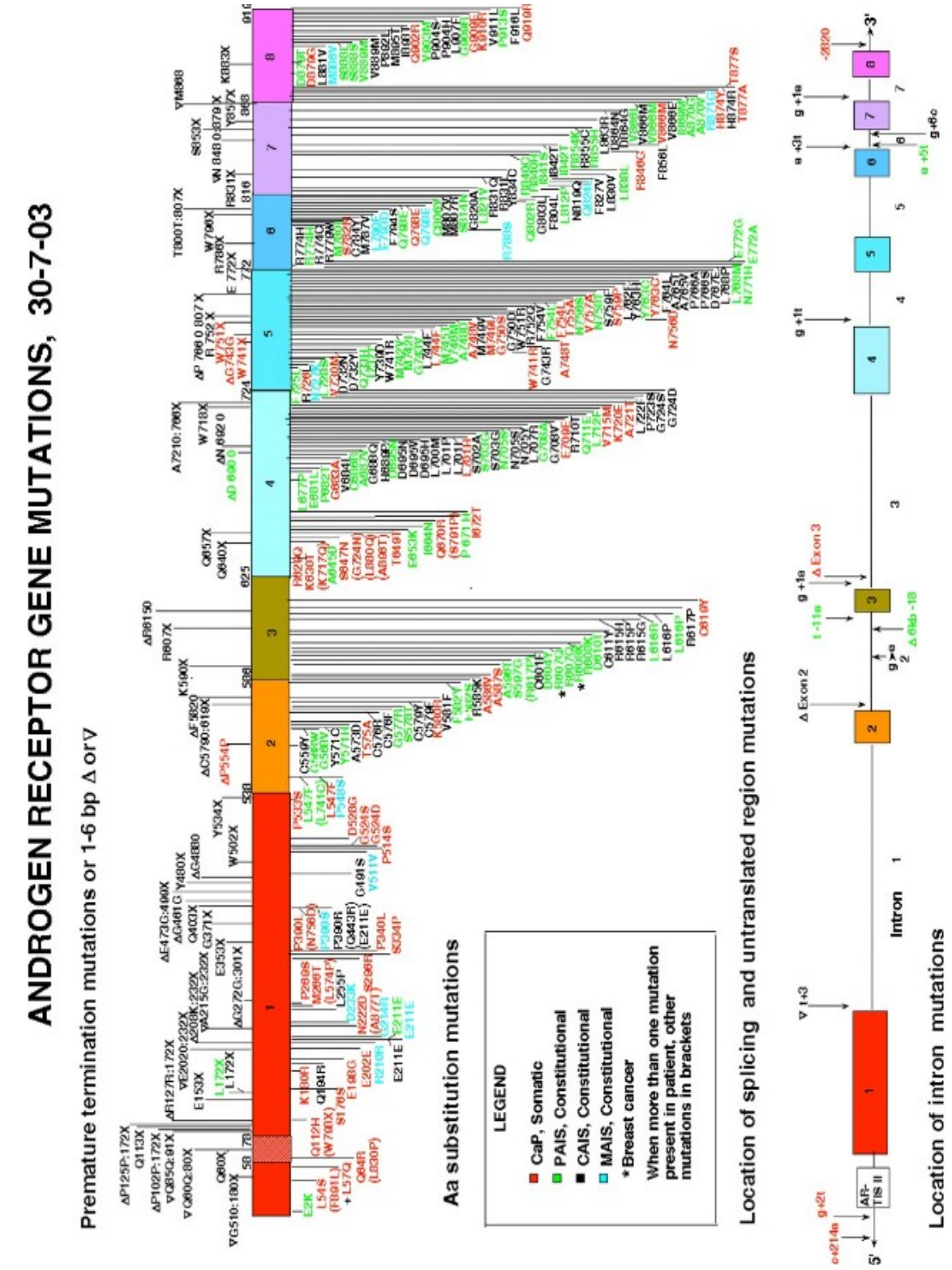


Figure 20: Summary of androgen receptor gene mutations based on (Gottlieb et al, 2004).

3.6 ARA 70

Androgen receptor-associated co-regulator 70 (ARA70) is one of the first identified co-regulators of androgen receptors. It is a protein which consists of 614 amino acid residues (70 kDa). ARA70 is expressed in a lot of tissues including prostate, testes, thymus and adrenal gland (Yeh et al, 1996). Several previous studies have shown that ARA70 plays an important role in androgen ablation and it can also function as an agonist of anti-androgens. It has been shown that a mutation P308S change the role of ARA70 to inhibitor of androgen receptor activity without any change in androgen receptor structure. This ability of ARA70 can be used in a new research related to prostate cancer treatment (Rahman et al, 2003). ARA70 interacts with androgen receptor via FXXLF motifs. The presence of LXXLL motif supports the fact that ARA70 interacts also with other nuclear receptors. Androgen receptor protein is more stable after binding of ARA70, suggesting that this is the way how ARA70 enhances the activity of androgen receptor (Hu et al, 2004). ARA70 occurs in the following isoforms: ARA70α (full length of 70 kDa) and an internally spliced 35-kDa ARA70β. ARA70β was reported as an oncogene in prostate cancer cell growth (Peng et al, 2008). Ligr et al reports that the full length of ARA70α inhibits growth when over-expressed in prostate cancer cell line LNCaP. This effect was observed by increased activity of apoptotic pathways. An increase in apoptotic gene Bax and cleaved caspase 3 and decrease in antiapoptotic gene Bcl-XL is androgen and androgen receptor dependent. A mutation T877A reduces interactions between ARA70α and androgen receptor (Ligr et al, 2010). There is still an unsolved question, why in prostate cancer cells the level of ARA70α expression is reduced and it is very important to solve this problem for the future treatment of prostate cancer.

4 Aminoaldehyde dehydrogenase 2 from *Pisum Sativum*

Aminoaldehyde dehydrogenase 2 (AMADH2) (EC 1.2.1.19) is an enzyme involved in polyamines catabolism. It belongs to the group of oxidoreductases subgroup of enzymes that act on aldehydes (EC 1.2.) and subclass of enzymes with NAD^+ and $NADP^+$ acting as an acceptors (EC 1.2.1.). The enzyme has been known under different names, but the officially accepted name is aminobutaraldehyde dehydrogenase, but there are also other alternatives 4-aminobutanal dehydrogenase (ABALDH) or 1-pyrroline dehydrogenase. The plant enzyme also oxidises 4-guanidinobutanal as a substrate and thus it has been classified under the name gamma-guanidinobutaraldehyde dehydrogenase (EC 1.2.1.54) (http://expasy.org/enzyme/1.2.1.19 23.2.2011). It has been demonstrated that peroxisomal targeting signal protoplastic localization of aminoaldehyde dehydrogenase in *Pisum sativum* is in the intracellular spaces of cambium, pericycle and endodermal cells. The highest activity was observed in lateral meristem (Sebela et al, 2001). However, the intracellular localization of AMADH is still unknown.

4.1 Substrate specificity

This enzyme has a broad substrate specificity in some organisms, but in general we can say that it prefers aliphatic chain aldehydes up to seven carbon atoms. In plants, AMADH is involved in polyamines catabolism and this pathway includes also copper amine oxidases (CAO, EC 1.4.3.6.) and FAD-containing polyamine oxidases (PAO, EC 1.5.3.11) (Bouchereau et al, 1999). CAOs catalyse the oxidation of diamine and polyamine substrates (e.g. putrescine, spermidine) at their primary amino group resulting in the formation of the corresponding aminoaldehydes, ammonia and hydrogen peroxide. The oxidation of spermine and spermidine by PAO produces 4-(3-aminopropylamino)butaraldehyde and 4-aminobutaraldehyde (ABAL), respectively, 1,3-diaminopropane and hydrogen peroxide (Sebela et al, 2000). 1,3-Diaminopropane can undergo transformation to 3-aminopropionaldehyde (APAL) (Awal and Hirasawa, 1995). APAL can be also trimethylated and transformed into the osmoprotectant β-alanine betaine. The oxidation of ABAL leads to 4-aminobutyric acid (GABA), which is

involved in numerous physiological processes (Petrivalsky et al, 2007; Roberts, 2007). The best substrates for AMADH are APAL, ABAL and 4-guanidinobutalaldehyde (GBAL) (i.e. C3 and C4 aminoaldehydes). The basic reaction, which is catalysed by AMADH, requires substrate (e.g. ABAL), NAD^+ and water. The product compounds are, for ABAL, 4-aminobutanoate and NADH (Sebela et al, 2000).

4.2 Structure of aminoaldehyde dehydrogenase 2

AMADH2 is a protein of 503 amino acids giving a molecular mass value of 54.6 kDa. In nature, it probably occurs as a dimer (Fig. 21) with an approximate molecular mass of 117 kDa (according to gel filtration results) (Tylichova et al, 2010) There are two isoforms. According to Sebela et al the pI values were calculated 5.44 for the main isoform and 5.39 for the other one (Sebela et al, 2001). Both isoforms (AMADH1 and AMADH2) show 80% sequence identity (Brauner et al, 2003).

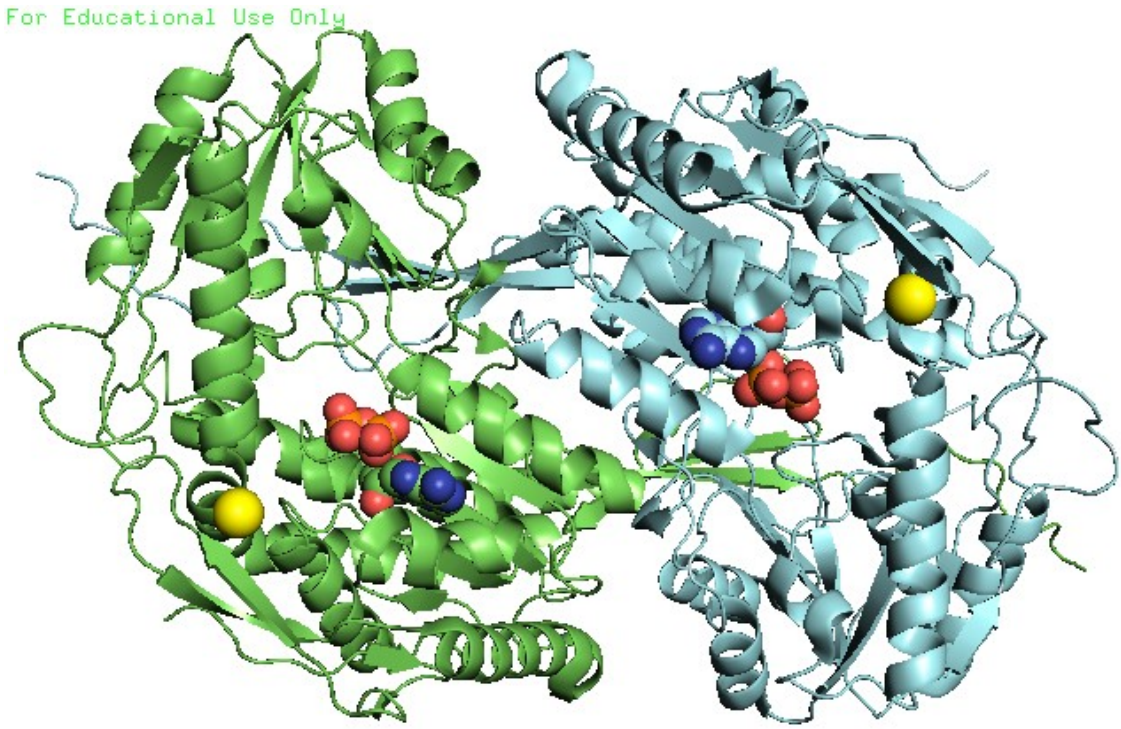


Figure 21: The crystal structure of pea aminoaldehyde dehydrogenase isoforms with NAD^+ and Na^+ (yellow spheres) – PDB 3IWJ (Tylichova et al, 2010).

The secondary structure consists mainly of alpha helices (20 for one chain, 40% of the structure) and from 20% of beta sheets (24 for one chain) (Fig. 22).

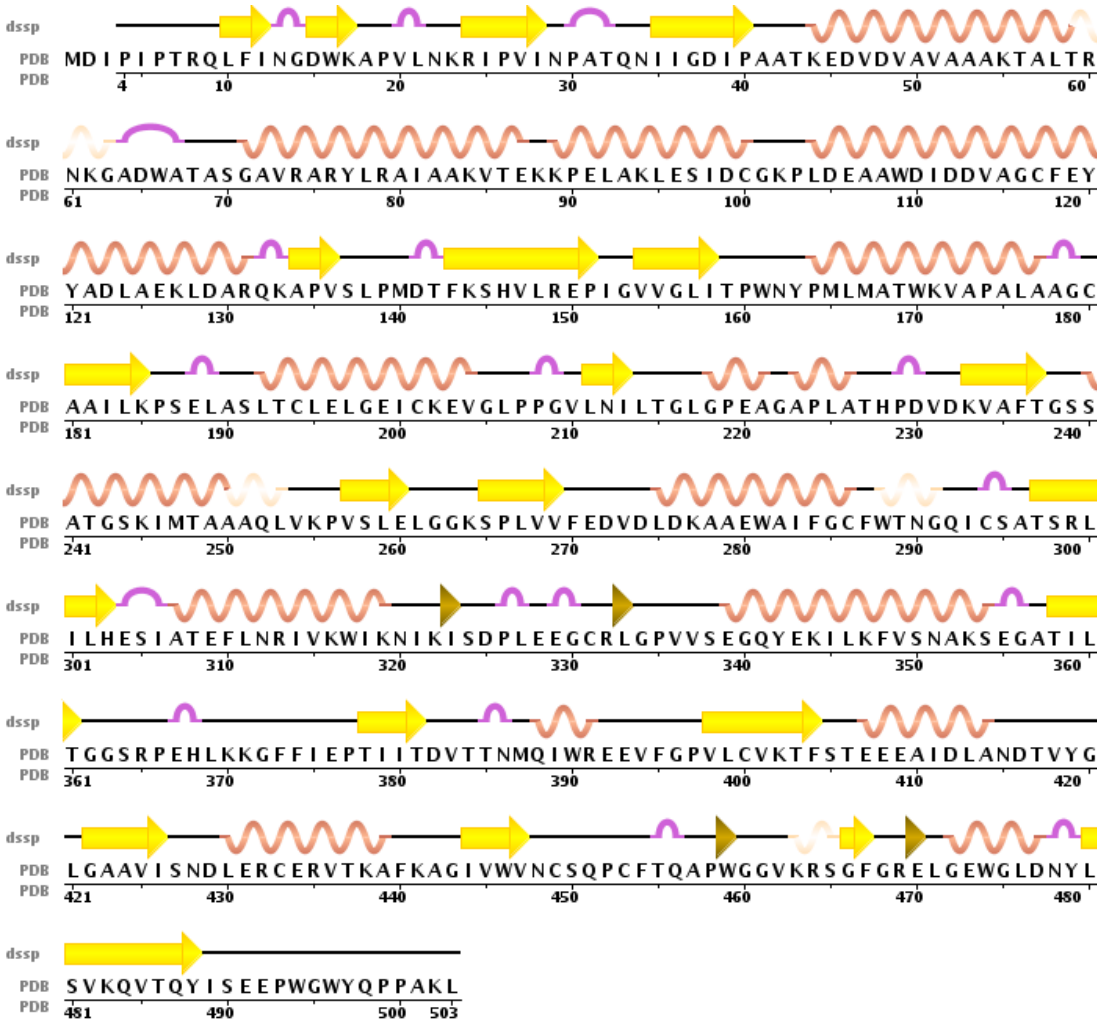


Figure 22: A detail of the secondary structure of AMADH2. (*brown swung dash* = α-helices; *yellow arrows* = β-strands; *purple loops* = hydrogen bonded turn) (server using accession 3IWJ – 24.3.2011).

AMADH2 contains three binding sites – a cation binding site, an NAD^+ binding site and a substrate binding site. The cation binding site is an intra-subunit cavity and it is highly conserved across aldehyde dehydrogenase structures (González-Segura et al, 2009). In AMADH2 cation binding cavity is formed by equivalent carbonyl groups of Ile28, Asp99 and Leu189 and the side chain of Asp99 (Fig. 23). Unfortunately, the role

of the monovalent cation in aldehyde dehydrogenases is not clear, it may stabilize loop from the NAD^+ binding site (Tylichova et al, 2010).

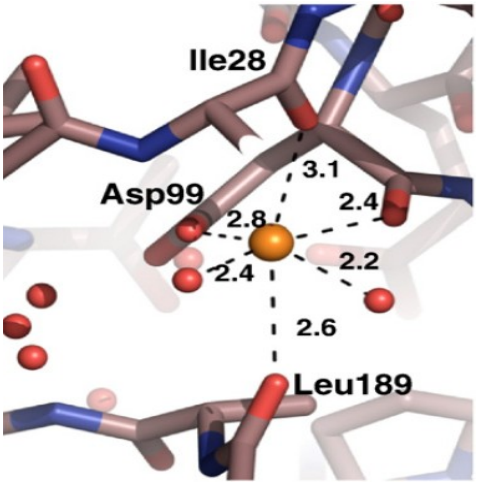


Figure 23: Binding site of Na^+ in AMADH2 (Tylichova et al, 2010).

The co-enzyme binding site for NAD^+ is placed in the hydrophobic pocket created by α-helices D and E. The ribose moiety of NAD^+ is stabilized by hydrogen bonds between oxygens O2B, O3B and the side chain of Glu188 (O2B) and the side chain oxygen of Thr159 (O3B). O2B and O3B simultaneously interact with the side chain of Lys185. Alpha-phosphate oxygens (O1A,O2A) interact with OG1 of Thr242. O2A also interact with oxygen atom of Ser239. The whole binding site is depicted in Fig. 24. (Tylichova, et al, 2010).

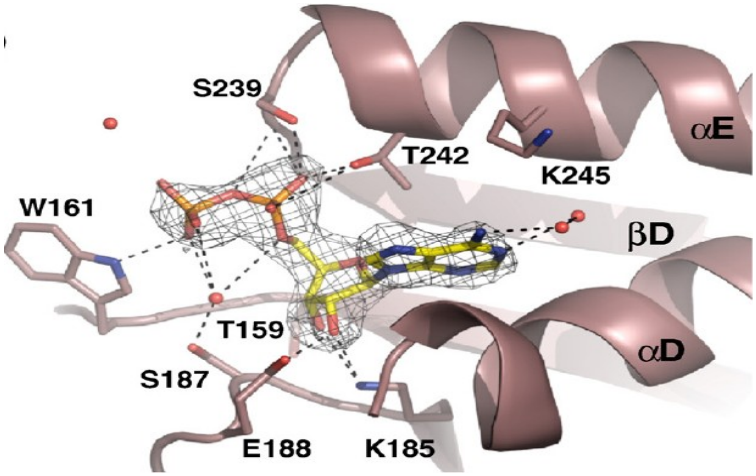


Figure 24: Binding site of NAD^+ (Tylichova et al, 2010).

The third binding site is represented by the substrate channel. Since there is only one published structure of AMADH2 and in this publication, they were unsuccessful to obtain crystal structure with a substrate, we have information only with glycerol molecule. Nevertheless, Tylichova and her team were able to describe parameters of the substrate channel. It is 14 Å long and 5-8 Å wide. The entrance is created by Trp109, Asp113, Pro452, Cys453 and its internal space is formed by side chains of Asp110, Asn162, Tyr163, Leu166, Met167, Trp170, Phe284, Trp288, Ile293, Cys294, Ser295, Gln451 and Trp459 (Fig. 25) (Tylichova et al, 2010).

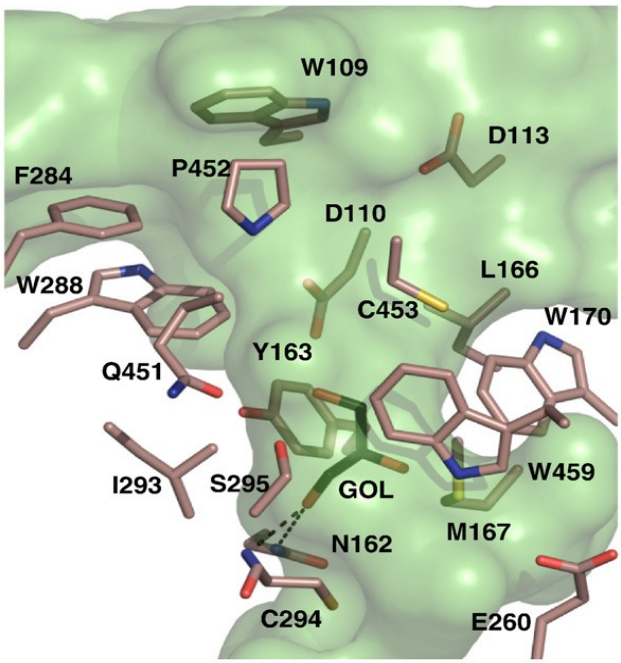


Figure 25: Substrate binding site of AMADH2 with co-crystallized glycerol (GOL) (Tylichova et al, 2010).

The active site consists of the catalytic Cys294, which lies amidst NAD^+ binding site and substrate channel and of Ans162 and Glu260 located at the bottom of the substrate channel (Tylichova et al, 2010).

Experimental Part:

5 Material and Methods

5.1 Chemicals and equipment

Chemicals:

- α-4-cyano 4-hydroxycinamic acid *Fluka (St. Louis, USA)*
- 4-aminobutyraldehyde-diethylacetal *Sigma (St. Louis, USA*)
- Acetic acid *Sigma (St. Louis, USA)*
- Acetonitrile *Sigma (St. Louis, USA)*
- Acrylamide *Sigma (St. Louis, USA)*
- Ammonium bicarbonate *Sigma (St. Louis, USA)*
- Ampicilin *Sigma (St. Louis, USA)*
- 3-aminopropionalaldehyde diethylacethal *Sigma (St. Louis, USA)*
- Ammonium persulfate *Fluka* (St. Louis, USA)
- Avidin-sepharose beads *Sigma (St. Louis, USA)*
- N,N'-methylene-bis-acrylamide *Sigma (St. Louis, USA)*
- Chloramphenicol *Sigma (St. Louis, USA)*
- Chymotrypsin *Sigma (St. Louis, USA)*
- Coomassie Brilliant blue *Sigma (St. Louis, USA)*
- Dithiothreitol *Sigma (St. Louis, USA)*
- Gateway® Cloning system *Invitrogen (California, USA)*
- Hydrochloric acid *Sigma (St. Louis, USA)*
- Imidazole *Sigma (St. Louis, USA)*
- Iodoacetamide *Sigma (St. Louis, USA)*
- IPTG *Sigma (St. Louis, USA)*
- Methanol *Sigma (St. Louis, USA)*
- NAD^+ *Sigma (St. Louis, USA)*
- Natrium azide *Sigma (St. Louis, USA)*
- NHS-LC-LC-Biotin *Pierce (Ili, USA)*

- pDEST-HisMBP *addGene (New England Biolabs, USA)*
- pDEST15 *Invitrogen (Ca, USA)*
- pDEST17 *Invitrogen (Ca, USA)*
- *E. Coli - Rosetta* cells
- Sample - Aminoaldehyde dehydrogenase 2 from *Pisum Sativum* (prepared by Mgr. Jan Frömmel)
- Sample ARA70 1-170 aa (made on my own)
- Sinapinic acid *Sigma (St. Louis, USA)*
- Sodium dodecyl sulphate *Fluka (St. Louis, USA)*
- N,N,N',N'-tetra-methylethylenediamine (TEMED) *Bio-Rad (USA)*
- TEV Protease *Invitrogen (Ca, USA)*
- Trifluor-acetic acid *Sigma (St. Louis, USA)*
- Trypsin *Sigma (St. Louis, USA)*

Equipment:

- Mass spectrometer Microflex MALDI-TOF LRF20, *Bruker Daltonics (Bremen,D)*
- MALDI target AnchorChip 600/96, *Bruker Daltonics (Bremen, D)*
- Pipettes 5000, 1000, 200, 100, 20, 10 and 2.5 µL, *Eppendorf (Hamburg, D)*
- Laboratory plastics (tips and tubes), *Eppendorf (Hamburg, D)*
- Vortex Stuart, *Bibby Sterilin (Stone, UK)*
- Centrifuges: miniSpin, *Eppendorf (Hamburg, SRN);* Spectrafuge mini, *Labnet (Korea)*; other centrifuges available in Parc Científic de Barcelona *Beckman (Ca, USA)*
- HisTrap™ HP Columns *GE Healthcare*
- Q Sepharose™ *GE Healthcare*
- Electrophoresis system Bio-Rad *(USA*)
- Peristaltic pump
- Spectrophotometer Beckman DU 7500 *Beckman*

5.2 Buffers and media

Growing media:

Luria-Bertani (LB) medium: 1 % (w/v) peptone, 0.5 % (w/v) yeast extract, 1 % (w/v) NaCl, pH 8.5

Buffer for biotinylation:

PBS buffer pH 7,4: 3.2 mM Na_2HPO_4, 0.5 mM KH_2PO_4, 1.3 mM KCl, 135 mM NaCl

Buffers and gels for SDS-PAGE electrophoresis with AMADH2:

Sample buffer – Laemmli buffer: 0.5 M Tris, 30 % (v/v) glycerol, 10 % SDS (w/v), 0.6 DTT, 0.012 M Bromophenol blue

Running buffer: 0.025 M Tris, 0.192 M glycine, pH 8.3

Acrylamide – bis-acrylamide: 30 % (w/v) acrylamide, 0,8 % (w/v) bis-acrylamide

4X Buffer for running gel: 1.5 M Tris/HCl, pH 8.8

4X Buffer for stacking gel: 0.5 M Tris/HCl, pH 6.8

SDS – stock solution: 10 % (w/v) SDS

Butanol for overlay of running gel: 50 % (v/v) butanol,

Ammonium persulphate $(NH_4)_2S_2O_8$ (APS): 10 % (w/v) ammonium persulfate composition

Running gel: 4 mL acrylamide – bis-acrylamide, 2.5 mL running gel buffer, 3.2 mL H_2O, 0.1 mL SDS – stock solution, 0.01 mL TEMED, 0.05 mL APS.

Stacking gel: 0.65 mL acrylamide – bis-acrylamide, 1.25 mL stacking gel buffer, 2.95 mL H_2O, 0.1 mL SDS – stock solution, 0.005 mL TEMED, 0.050 mL APS

Buffers and gels for SDS-PAGE electrophoresis with ARA70:

Sample buffer – Laemmli buffer: 0.5 M Tris, 30 % (v/v) glycerol, 10 % SDS (w/v), 0.6 DTT, 0.012 M Bromophenol blue

Running buffer: 0.025 M Tris, 0.192 M glycine, pH 8.3

Acrylamide – bis-acrylamide: 30 % (w/v) acrylamide, 0.8 % (w/v) bis-acrylamide

Buffer for running gel: 3.0 M Tris/HCl, pH 8.8

Buffer for stacking gel: 0.5 M Tris/HCl, pH 6.8

SDS – stock solution: 10 % (w/v) SDS

Butanol for overlay of running gel: 50 % (v/v) butanol,

Ammonium persulphate $(NH_4)_2S_2O_8$ (APS): 10 % (w/v) ammonium persulfate composition

Running gel: 6 mL acrylamide – bis-acrylamide, 3,75 mL running gel buffer, pH 8.8, 5.25 mL H_2O, 0.1 mL SDS – stock solution, 0.01 mL TEMED, 0.05 mL APS

Stacking gel: 0,65 mL acrylamide – bis-acrylamide, 1.25 mL stacking gel buffer, 3,05 mL H_2O, 0.1 mL SDS – stock solution, 0.005 mL TEMED, 0.050 mL APS

Buffers for Coomassie Brilliant Blue R-250 staining:

Coomassie Brilliant Blue R-250 solution: 0.1 % (w/v) Coomassie Brilliant Blue R-250, 15 % (v/v) acetic acid, 45 % (v/v) methanol

Destaining buffer: 40 % (v/v) methanol, 10 % (v/v) acetic acid

Buffer for measuring of enzymatic activity:

Activity assay buffer: 150 mM Tris-HCl, pH 9.0

Buffers for ARA70 1-170 purification:

Re-suspension buffer: 50 mM Tris-HCl, 150 mM NaCl, 25 mM imidazole pH 8.2

Dialysis buffer: 50 mM Tris-HCl, 150 mM NaCl pH 8.2

His-Trap buffers:

Washing buffer A: 50 mM Tris-HCl, 1 M NaCl, 25 mM imidazole pH 8.2

Washing buffer B: 50 mM Tris-HCl, 150 mM NaCl, 40 mM imidazole pH 8.2

Elution buffer: 50 mM Tris-HCl, 150 mM NaCl, 250 mM imidazole pH 8.2

Ion exchange buffers:

Buffer A: 50 mM Tris-HCl, 150 mM NaCl, pH 8.2

Buffer B: 50 mM Tris-HCl, 500 mM NaCl, pH 8.2

Buffer C: 50 mM Tris-HCl, 1 M NaCl, pH 8.2

MALDI in-gel digestion buffers & solutions:

Washing buffer: 0.1 M ammonium bicarbonate

Reducing solution: 0.1 M ammonium bicarbonate, 10 mM dithiothreitol

Alkylation solution: 0.1 M ammonium bicarbonate, 55 mM iodoacetamide

Digestion Buffer A: 25 mM ammonium bicarbonate, 5mM $CaCl_2$, 4 μM trypsin

Digestion Buffer B: 25 mM ammonium bicarbonate, 4 μM chymotrypsin

Digestion Buffer C: 25 mM ammonium bicarbonate

5.3 Methods

5.3.1 Biotinylation of AMADH2

AMADH sample concentration of 1,54 mg/mL was measured by the BCA method. According to this information, biotinylation was made following the protocol by the manufacture of *EZ-Link® Sulfo-NHS-LC-LC-Biotin (Pierce, USA)*. Based on an calculation 3 ml of protein solution were mixed with 84 μL of the biotin reagent and left it for 2 hours on ice.

5.3.2 SDS-PAGE electrophoresis

SDS-PAGE was used for a partial purification and separation of biotinylated samples from the reagent mixture and for visualization of ARA70 results. Usually 12% gels were used. First, running gel was polymerized in the gel caster, the polymerizing

solution was covered by 50% butanol (v/v). Afterwards, stacking gel was cast and with a comb placed to create wells. Samples (5 µL) were mixed with Laemmli buffer in ratio 1:1, boiled for ten minutes and loaded to the wells. Finally, the apparatus was connected to a power supply to run gels first 10 minutes at 110 V and than at 140 V. Afterwards, gels were stained and destained according to standard protocol with Coomassie Brilliant Blue staining.

5.3.3 MALDI sample preparation - In-gel digestion

Destained gels were rinsed with water. Bands of interest containing AMADH2 were excised from the gel, cut into small cubes and put into 0,5 mL Eppendorf tubes. First the gel particles were washed with water. Then 100 µL of NH_4HCO_3 and 100 µL of acetonitrile were added and the mixture was incubated it at laboratory temperature until gel particles became destained (30 – 60 minutes). Afterwards, liquid was aspirated and 100 µL of acetonitrile was added to shrink the pieces. After approximately 10 minutes acetonitrile was removed and the gel pieces were incubated with 10 mM dithiothreitol/0.1 M NH_4HCO_3 for 30 minutes at 56°C to reduce protein. Reduced samples were dried and shrunk again by acetonitrile. A acetonitrile was replaced with 55 mM iodoacetamide/0.1 M NH_4HCO_3 and incubated in the dark at laboratory temperature for 20 minutes. Later on, gel particles were washed with 100 µL of 0.1 M NH_4HCO_3 (15 minutes) and shrunk again with 100 µL of acetonitrile. Finally, acetonitrile was replaced by the digestion buffer containing trypsin or chymotrypsin and the samples were incubated for 1 hour on ice. In the last step, digestion buffer was removed and replaced by 20 µL of 25 mM of NH_4HCO_3. The digestion proceed at 37°C overnight.

5.3.4 Measuring at the mass spectrometer MALDI Microflex LRF 20

A digest aliquot or protein sample (0,8 µL) was pipetted on a MALDI target AnchorChip™ 600/96, mixed with 0,8 µL of matrix (CHCA or sinapic acid) and dried

by spontaneous evaporation at laboratory temperature.

Mass spectra were gained by measuring in linear positive mode at a Microflex MALDI-TOF-MS LRF20 apparatus. Spectra were accumulated from 800 – 1500 laser shots at pulse frequency of 10 Hz and recorded in the range of m/z 500 – 6000 for peptides and m/z 4000 – 250 000 for protein mass. Mass spectrometer was calibrated by external calibration by peptide standard (peptide measuring) or BSA (intact mass measuring). Data were processed by original software flexAnalysis v 3.0 and mMass v 3.0.0 and 3.10.0. (Strohalm et al, 2008; Strohalm et al, 2010).

5.3.5 Measurement of enzyme activity

Enzyme activity was measured to evaluate the impact of biotinylation to the enzyme function. The measurement were performed with two substrates: 3-aminopropionaldehyde (APAL) and 4-aminobutarladehyde (ABAL). Stock solutions of the substrates were made by mixing 0.2 M HCl with a precise volume of ABAL or APAL to create final concentration of 100 mM. Afterwards, the substrate solution was boiled at 100°C for 10 minutes. The reaction mixture (2 mL) was composed of 1.55 mL of 150 mM Tris-HCl buffer pH 9; 50 μL of 20 mM NAD^+; 370 μL of ddH_2O and 10 μL of enzyme AMADH2. All solutions were always freshly prepared. Substrate solution (20 μL) was added to the mixture just before measurement.

Each measurement was performed on a Beckman DU 7500 spectrophotometer at the wavelength of 340 nm wing 15^{-s} intervals in a total time of 20 minutes. The cuvette was thermostated at 30°C. Kinetic parameters were determined by linearisation of initial speed dependencies on substrate concentration according to Lineweaver and Burk and according to Eadie and Scatchard. From such results, arithmetic averages were calculated which were considered output values.

5.3.6 Immobilization of AMADH2 on avidin-agarose beads

Avidin-agarose beads were purchased from *Sigma* and according to previous protocols (Schriemer and Ling, 1996; Schmitt et al, 1990; Swack et al, 1978) several

different conditions were used for immobilization. For all conditions, we mixed 100 µL of agarose beads and 500 µL of AMADH2 together. Different incubation times and temperatures were used (more in Results and discussion section).

5.3.7 Transfection of the Rosetta cells (ARA70)

A vector pDEST-HisMBP containing sequence of ARA70 1-170 was used for transfection of the *Rosetta* cell line. DNA sample (2 µL) was mixed together with 50 µL of cells and incubated on ice for 30 minutes. Afterwards, a heat shock was performed at 42°C for 30 seconds. For regeneration of the cells, 1 mL of LB media was added to the solution and incubated at 37°C for 1 hour. Then aliquot of the suspension (100 µL) was placed on the LB medium plates with chloramphenicol and ampicilin. Finally, the plates were left for incubation at 37°C overnight.

5.3.8 Preparation of the cell culture in a big volume (ARA70)

For purification of ARA70 1-170 a large amount of cells was needed to extract the protein of interest. On the first day, a colony was picked up from the plate and put into 5 mL of LB media containing chloramphenicol and ampicilin. This culture was left to grow at 37°C overnight in an incubator. Next day, 3 mL of already grown culture were put into 50 mL of fresh LB media (also containing chloramphenicol and ampicilin) and let grow in the incubator as above. The final step was an inoculation of 2 L of LB media (with chloramphenicol and ampicilin). The culture was grown to an OD_{600} = 0,6. Afterwards, IPTG was added to final concentration of 1 mM and the culture was incubated at 30°C for 6 hours with shaking. After induction, the cultures were spin down and re-suspended in 80 mL of the re-suspension buffer.

5.3.9 Purification of ARA70 1-170

First, a cell disruption was performed by adding lysozyme, protease inhibitors cocktail and the culture was repeadly frozen and thawn to achieve a high yield of free

proteins. Before loading sample to a His-Trap column, RNAse and DNAse were added to prevent clogging of the column. It is also highly recommended to use a 50 μM filter because of the same reason.

To equilibrate the column (HisTrap HP Column *GE Healthcare*) several column volumes of the re-suspension buffer were pass. After equilibration filtered protein sample was left to pass through the column. For washing steps, first washing buffer A (again several column volumes) was used followed by buffer B. To be sure that protein binding to the affinity column is effective it is recommended to check the flow-through fraction for the presence of any protein by a protein detection kit. The washing buffer B contains higher concentration of imidazole to prevent non-specific protein interactions on the column. When the flow-through fraction is free of proteins it is possible to start with elution of the bound protein. The majority of the bond protein should appear in the first fraction (in this case the volumje was 1 mL, but it depends on the volume of the column), but it is better to take more consecutive fractions to be sure that all desired protein is collected (in this case it was 10 fractions, 1 mL each).

An approach in case of purification after TEV cleavage is a little bit different. Your protein of interest is in the flow-through. Only a tag-part of the protein remains in the column (in our case MBP-His).

5.3.10 TEV cleavage and dialysis of ARA70 1-170

After His-trap column chromatography first two fractions were pooled and mixed with TEV-protease. The mixture was poured into a dialysis bag and let to dialyze against 1.5 L of 50 mM Tris-HCl, 150 mM NaCl buffer at 4°C for 24 hours.

5.3.11 Ion exchange chromatography (ARA70)

First the column (Sepharose Q *GE* Healthcare) was washed with ddH20 and then with a high-salt concentration buffer. After these preparation steps protein sample was loaded on the column and than eluted with salt gradient buffers. Ten mililitres of each elution buffer (150 mM, 500 mM and 1 M salt concetration) was taken by 1 mL

aliquotes from the column. From each, the respective fraction was collected for further analysis.

5.3.12 SDS-PAGE sample preparation (ARA70)

For SDS-PAGE analysis several different approaches were involved for sample preparation. For all samples prepared from cell cultures the same approach was used. 1 mL of the cell culture was spin down and re-suspended in 50 µL of 50 mM Tris-HCl, 150 mM buffer. Then an aliquot of 15 µL of the cell culture was taken out, mixed it with 3 µL of Laemmli buffer and boiled for 5 minutes. Samples with soluble proteins were obtained by a similar approach, but with one additional step. After re-suspending in the above buffer cells were disrupted in the sonicator and then spun down again. Only the supernatant was taken out, containing soluble proteins. Other samples were prepared by mixing a sample solution and Laemmli buffer in a ratio of 1:5 and boiling the mixture for 5 minutes.

Pictures edited by using PDB structures were made in PyMol software (Schrodinger, LLC, 2010).

6 Results

6.1 Results of AMADH2 biotinylation

AMADH2 biotinylation was performed for the purpose of enzyme immobilization on chips through the interaction between biotin and avidin. Aminolaldehyde dehydrogenase 2 contains 35 lysine residues in its amino acid sequence (Table 2), most of them are localized on the surface (Fig. 26). All these surface lysines represent possible targets for binding of LC-LC-biotin.

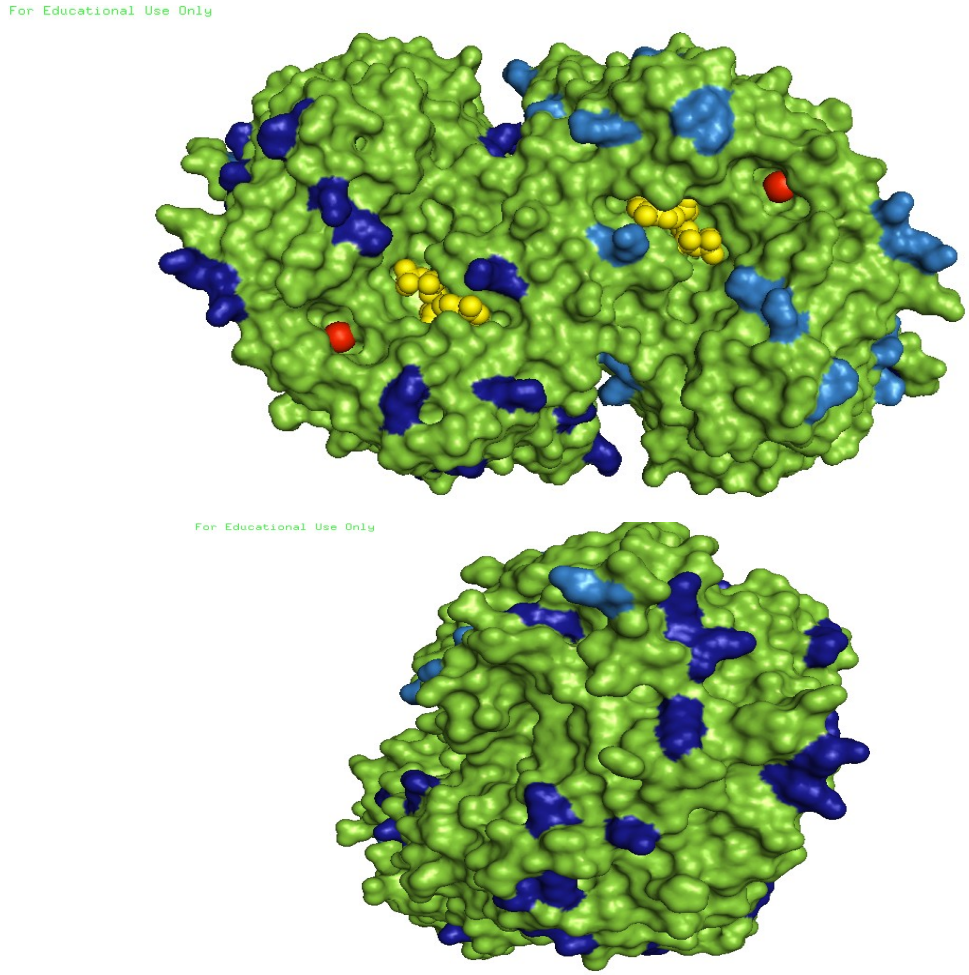


Figure 26: Surface of AMADH2, Blue and light blue – surface lysines from chain A and B, respectively; yellow spheres– NAD^+; red spheres– Na^+ cation (PDB

structure 3IWJ).

```
1       MDIPIPTRQL FINGDWKAPV LNKRIPVINP ATQNIIGDIP AATKEDVDVA
51      VAAAKTALTR NKGADWATAS GAVRARYLRA IAAKVTEKKP ELAKLESIDC
101     GKPLDEAAWD IDDVAGCFEY YADLAEKLDA RQKAPVSLPM DTFKSHVLRE
151     PIGVVGLITP WNYPMLMATW KVAPALAAGC AAILKPSELA SLTCLELGEI
201     CKEVGLPPGV LNILTGLGPE AGAPLATHPD VDKVAFTGSS ATGSKIMTAA
251     AQLVKPVSLE LGGKSPLVVF EDVDLDKAAE WAIFGCFWTN GQICSATSRL
301     ILHESIATEF LNRIVKWIKN IKISDPLEEG CRLGPVVSEG QYEKILKFVS
351     NAKSEGATIL TGGSRPEHLK KGFFIEPTII TDVTTNMQIW REEVFGPVLC
401     VKTFSTEEEA IDLANDTVYG LGAAVISNDL ERCERVTKAF KAGIVWVNCS
451     QPCFTQAPWG GVKRSGFGRE LGEWGLDNYL SVKQVTQYIS EEPWGWYQPP
501     AKL
```

Table 2: AMADH2 amino acid sequence, lysines are marked in red.

6.1.1 Mass spectra

With the help of mass spectrometry it was possible to identify 5 LC-LC-biotin modifications in AMADH2 (Fig. 27 – 31) in the positions K133, K316, K438, K502 and K370 or K371. The last mentioned place of modification is not clear, because tryptic digest peptide 354 - 371 (SEGATILTGGSRPEHLKK) contains 2 lysine residues and without using tandem mass spectrometry it is impossible to distinguish which lysine is modified. All registered modifications are summarized in Figure 33. Each modification increases a non-modified peptide mass by 452.6 Da.

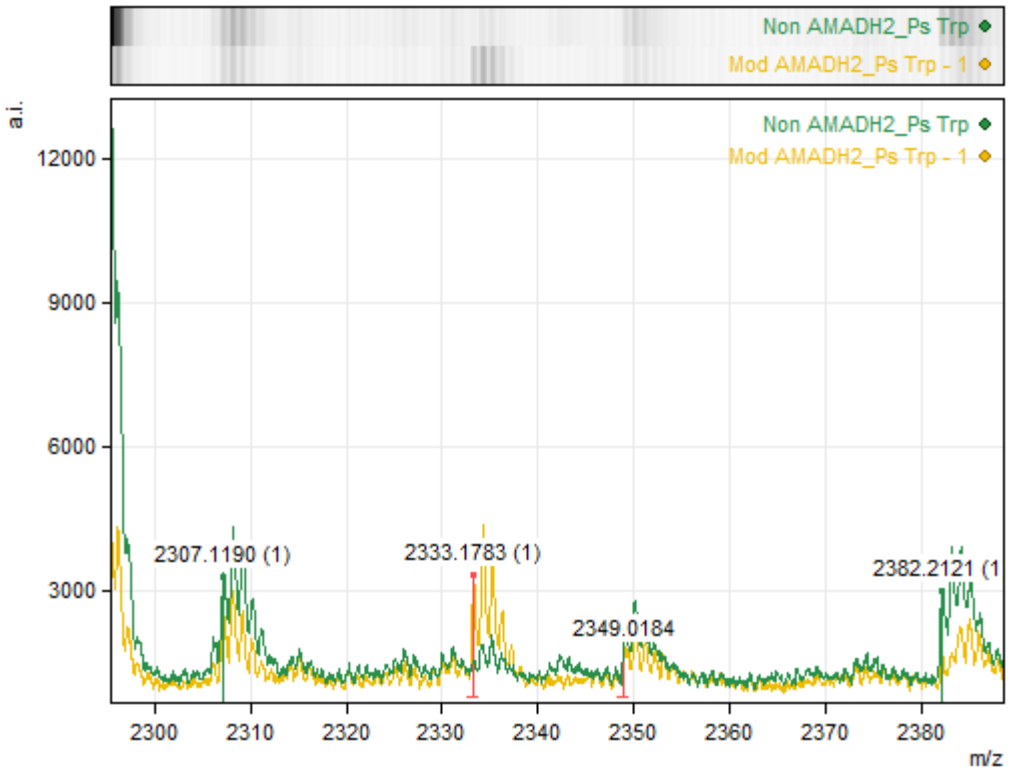


Figure 27: MALDI-TOF mass spectrum of AMADH2, tryptic digest. A modified peptide with m/z 2333.1783 is shown (teoretical m/z 2333.6189) which corresponds to a predicted tryptic peptide 354 - 371 with sequence SEGATILTGGSRPEHLKK and teoretical non-modified m/z 1881.0189. Yellow spectrum – modified AMADH2; green spectrum – non-modified AMADH2.

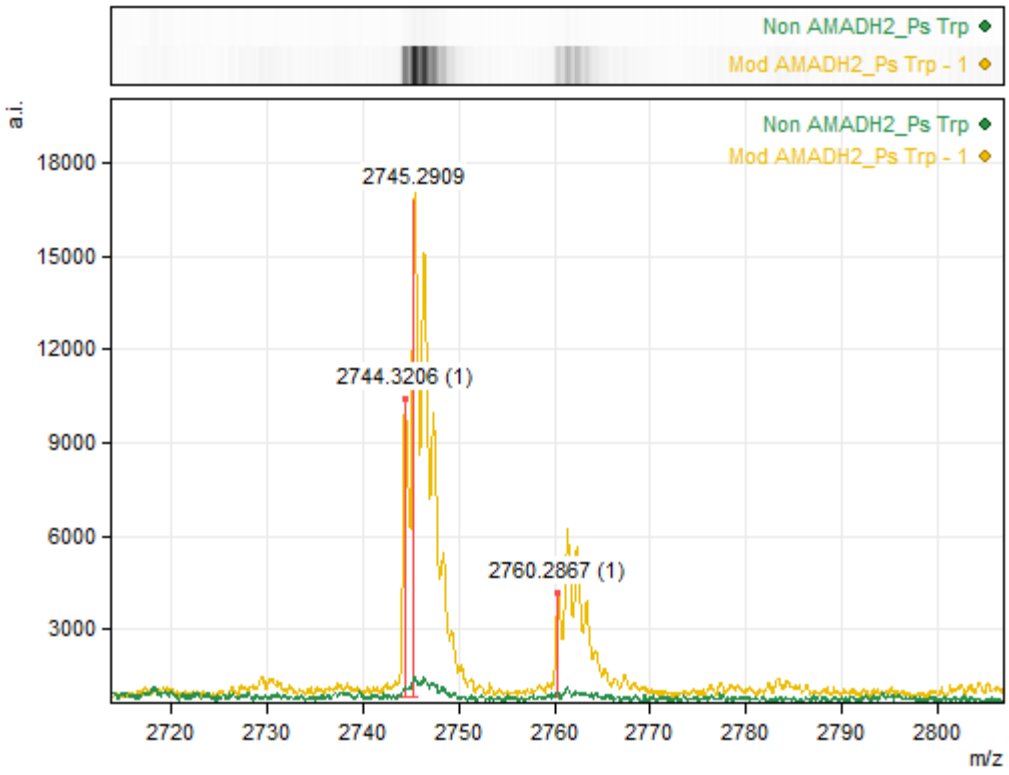


Figure 28: MALDI-TOF mass spectrum of AMADH2, tryptic digest. A modified peptide with m/z 2760.2867 is shown (teoretical m/z 2759.7080) which corresponds to a predicted tryptic peptide 484 - 502 with sequence QVTQYISEEPWGWYQPPAK and teoretical non-modified m/z 2307.1081. Yellow spectrum – modified AMADH2; green spectrum – non-modified AMADH2. The peak with m/z 2744.3206 may reflects the same peptide without methionine oxidation (m/z 2760.2867), but this peptide has no Met in its sequence. This means that peak with m/z 2744.3206 is possibly a fragmentation product.

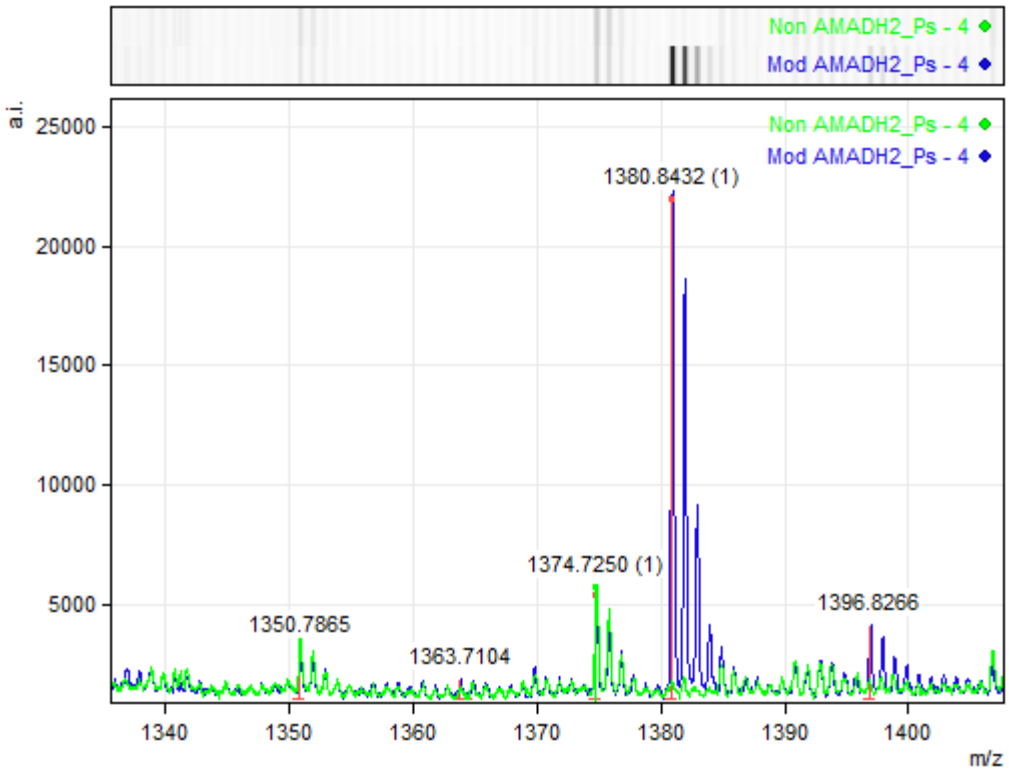


Figure 29: MALDI-TOF mass spectrum of AMADH2, chymotryptic digest. A modified peptide with m/z 1380.8432 is shown (teoretical m/z 1381.1727) which corresponds to a predicted chymotryptic peptide 311 - 317 with sequence LNRIVKW and teoretical non-modified m/z 928.5727. Blue spectrum – modified AMADH2; green spectrum – non-modified AMADH2.

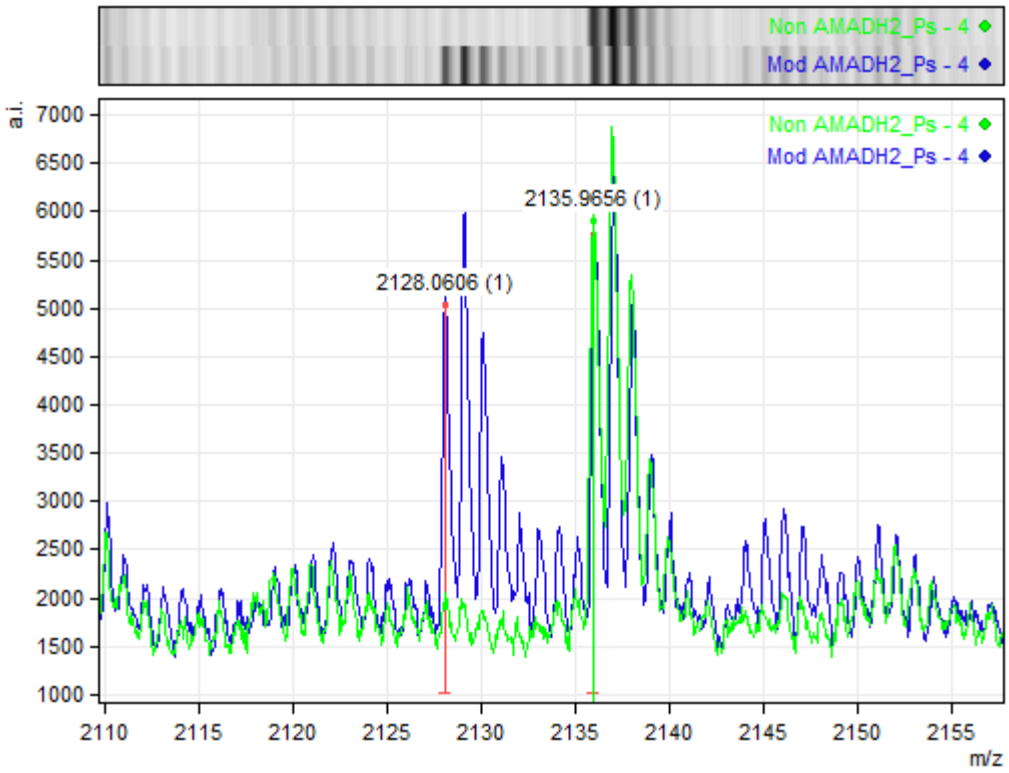


Figure 30: MALDI-TOF mass spectrum of AMADH2, chymotryptic digest. A modified peptide with m/z 2128.0606 is shown (teoretical m/z 2128.4472) which corresponds to a predicted chymotryptic peptide 129 - 143 with sequence DARQKAPVSLPMDTF and teoretical non-modified m/z 1675.8472. Blue spectrum – modified AMADH2; green spectrum – non-modified AMADH2.

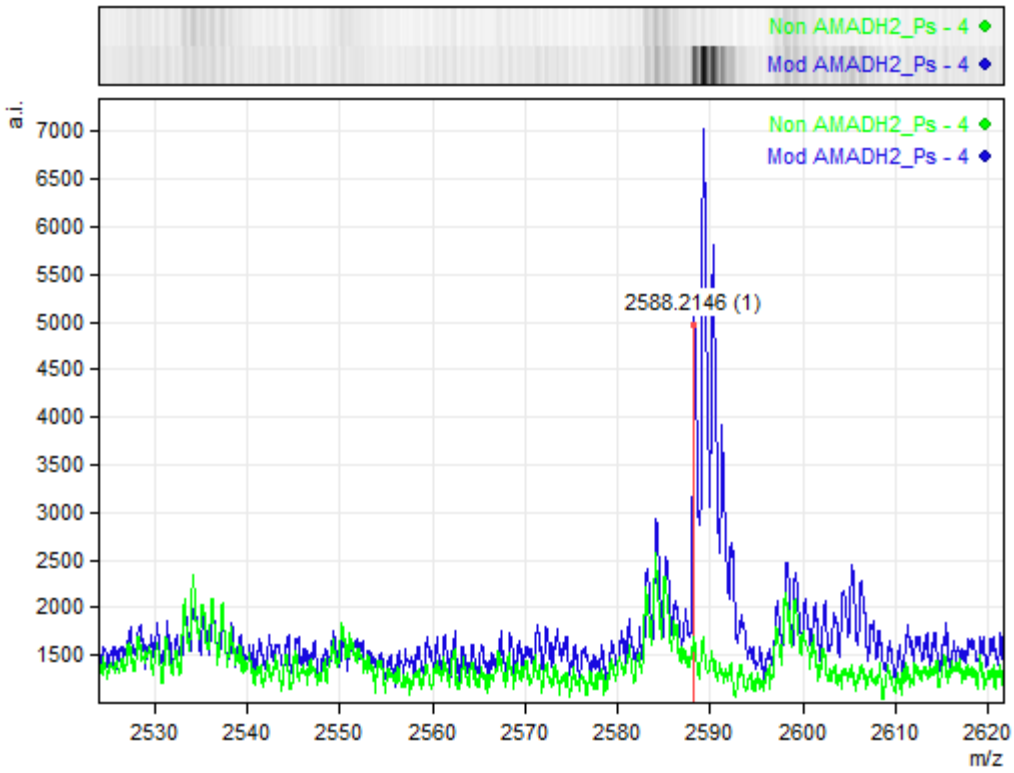


Figure 31: MALDI-TOF mass spectrum of AMADH2, chymotryptic digest. A modified peptide with m/z 2588.2146 is shown (teoretical m/z 2588.6866) which corresponds to a predicted chymotryptic peptide 422 - 440 with sequence GAAVISNDLERCERVTKAF and teoretical non-modified m/z 2136.0866. Blue spectrum – modified AMADH2; green spectrum – non-modified AMADH2.

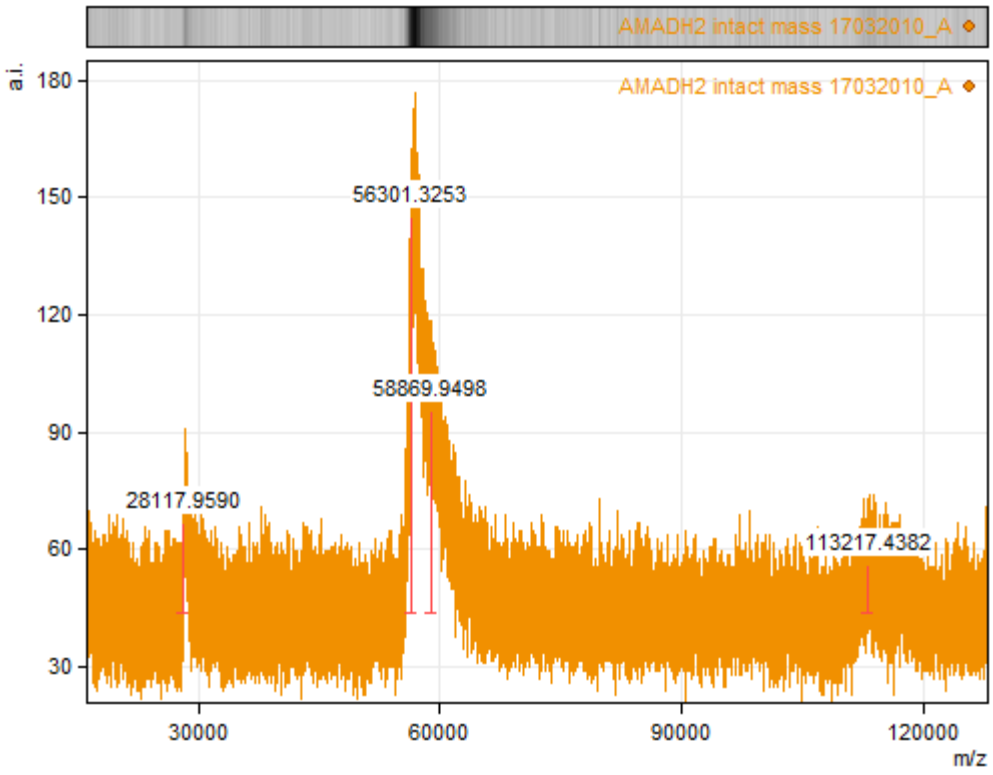


Figure 32: MALDI-TOF mass spectrum of intact AMADH2. The peak m/z 56301 corresponds with predicted non-modified AMADH2 mass; the shoulder at m/z 58869 corresponds to an intact mass of AMADH2 carrying five LC-LC-biotins, the peak at m/z 28117 corresponds to a doubly charged non-modified AMADH2, peak at m/z 113217 corresponds to a dimer of AMADH2. Low intensity may be caused by low concentration of AMADH2.

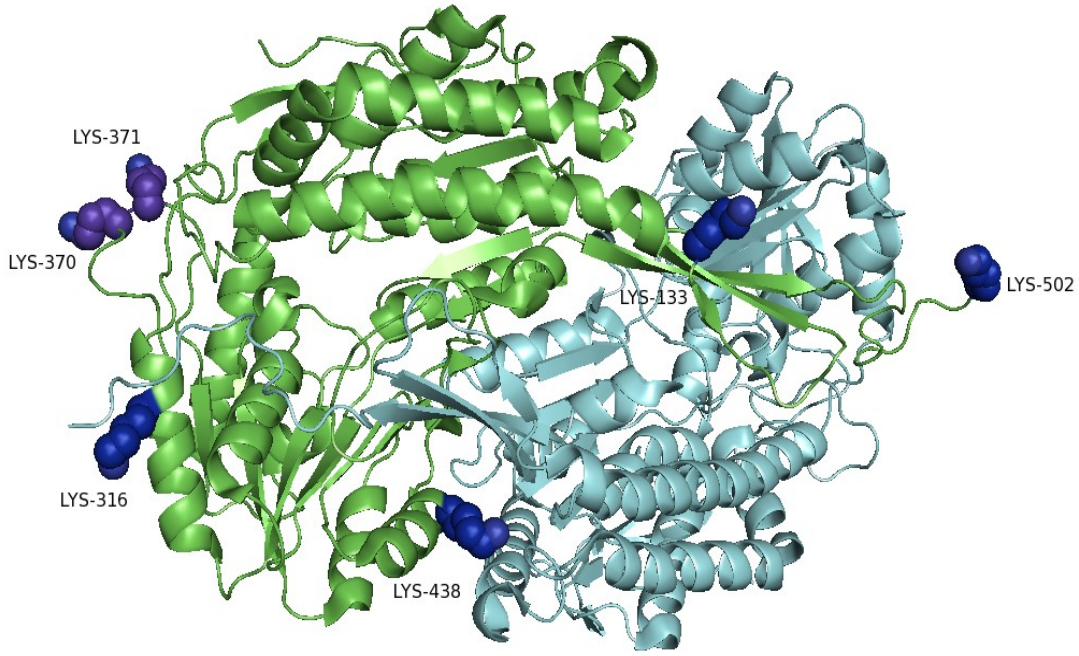


Figure 33: Modified lysine residues in AMADH2 structure. Blue spheres– Lys133, Lys316, Lys438, Lys502. Purple spheres – two options of the fifth modification: Lys370 and Lys371 (PDB structure 3IWJ).

6.1.2 Enzyme activity

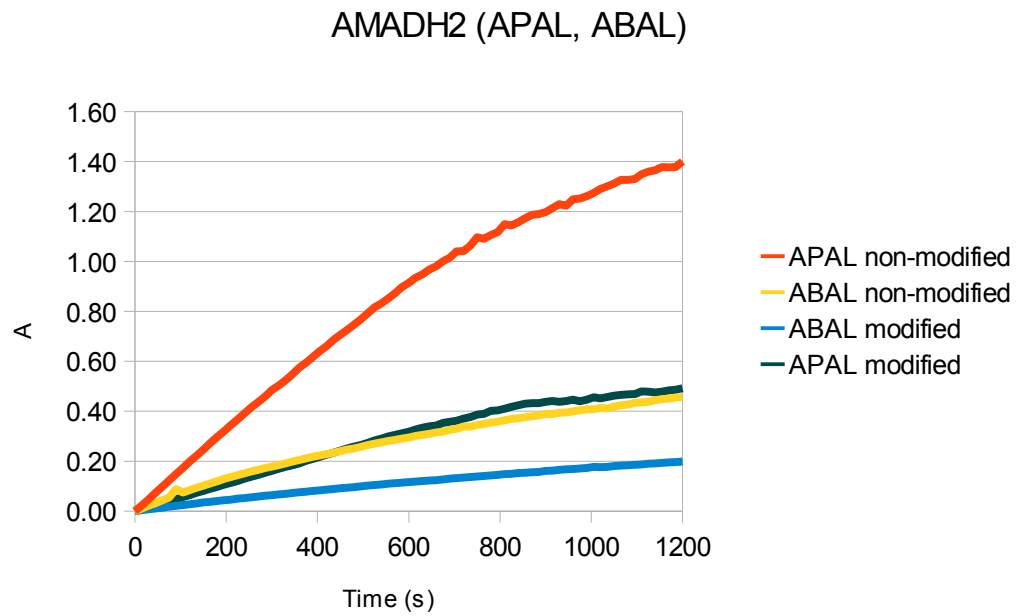


Figure 34: Activity of AMADH2 with two different substrates – APAL and ABAL. Orange and yellow lines represent the reaction of non-modified AMADH2 with APAL (APAL non-modified) and ABAL (ABAL non-modified), respectively. Blue and green lines are modified AMADH2 with APAL (APAL modified) and ABAL (ABAL modified), respectively. Concentration of the substrates was 300 μM.

Following graphs (Fig. 35 – 37) show the determination of kinetic parameters of modified AMADH2 measured with increasing concentration of APAL as a substrate (8 – 400 μM).

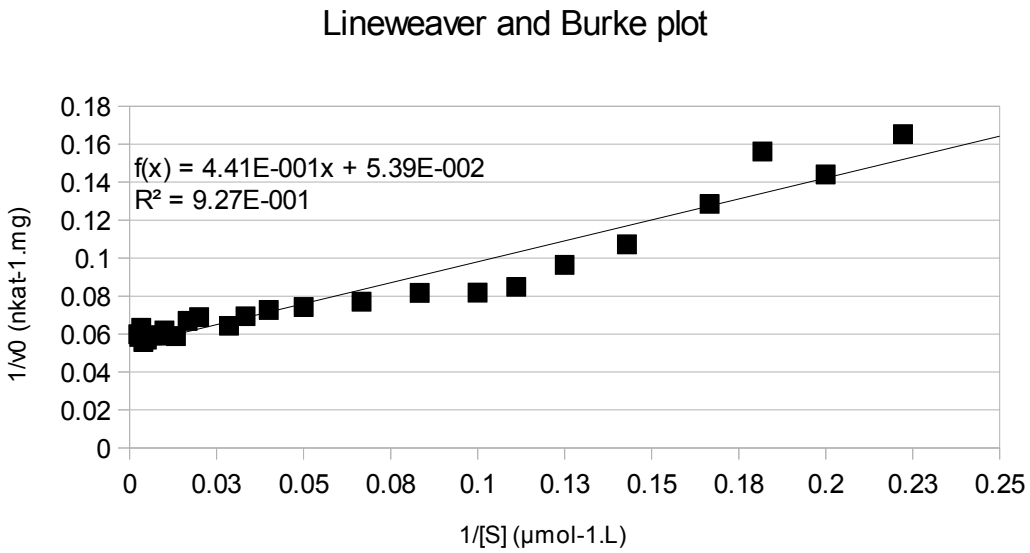


Figure 35: A Lineweaver and Burke plot of the reaction of modified AMADH2 with the substrate APAL.

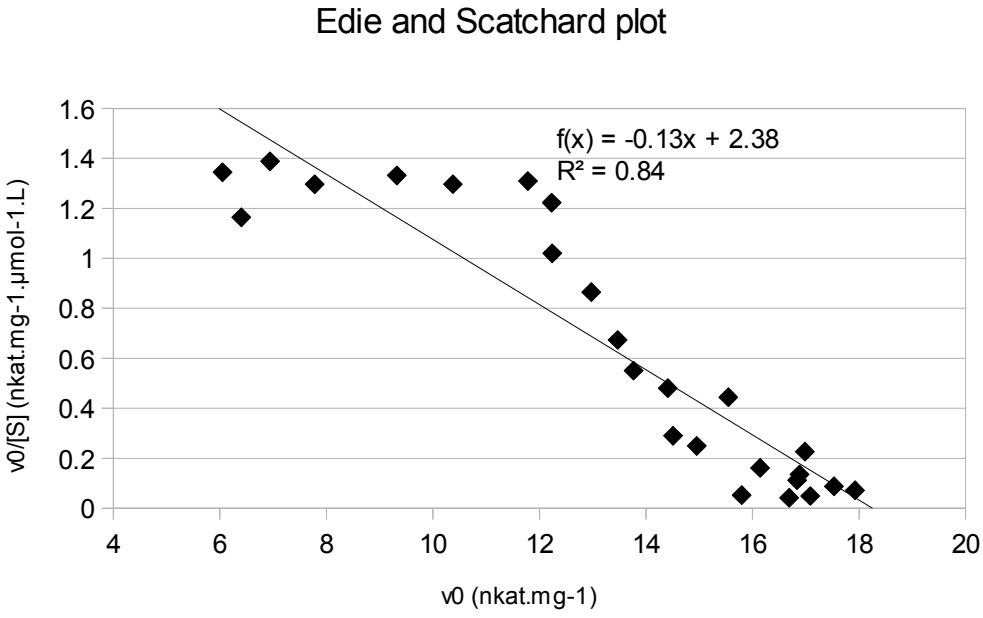


Figure 36: A Edie and Scatchard plot of the reaction of modified AMADH2 with the substrate APAL.

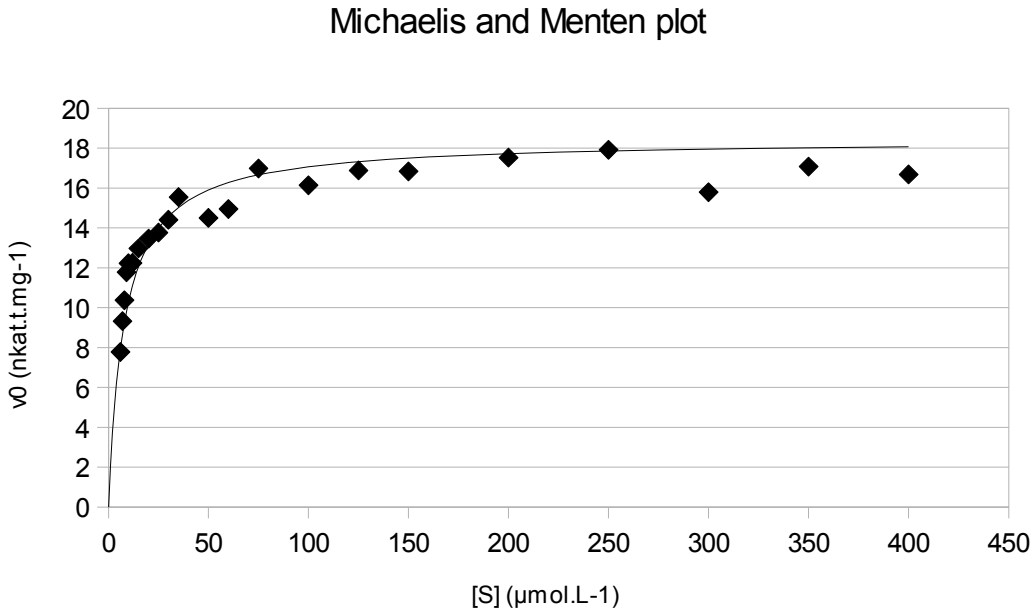


Figure 37: Michaelis and Menten plot of the reaction of modified AMADH2 with the substrate APAL. The plot shows dependence of the initial rate on substrate concentration.

According to the experiments, the following kinetic parameters were determined: K_m = 7.94 µmol.L^{-1} and V = 18.4 nmol.s^{-1}.mg^{-1} (APAL as a substrate). For non-modified AMADH2, K_m (12 µmol.L^{-1}) and V (175 nmol.s^{-1}.mg^{-1}) were previously measured (Frömmel, 2010). The results are summarized in Tab. 3.

	Non-modified AMADH2	**modified AMADH2**	V/K_m
K_m(µmol.L^{-1})	12	7.94	100.00%
V (nmol.s^{-1}.mg^{-1})	175	18.4	15.89%

Table 3: The determined values of K_m, V and relative values V/K_m for APAL oxiditaion with recombinant pea AMADH2 and its biotinylated derivate.

6.2 ARA70 1-170 results

The following pictures show results from production and purification of the first part of ARA70. This part consists of the first 170 amino acids from the N-terminus of ARA70 and its molecular mass is 20 kDa. ARA70 1-170 with a GST tag (Fig. 38) is equivalent to molecular mass of 47 kDa and ARA70 1-170 with a His-tag (Fig. 39) is equivalent to molecular mass of 22 kDa. Protein expressed from the vector pDEST-HisMBP-ARA70 1-170 (Fig. 40) corresponds a molecular mass of 66 kDa. All samples were obtained according to the protocols in chapter 5.3 and purification was performed as shown in Figure 41.

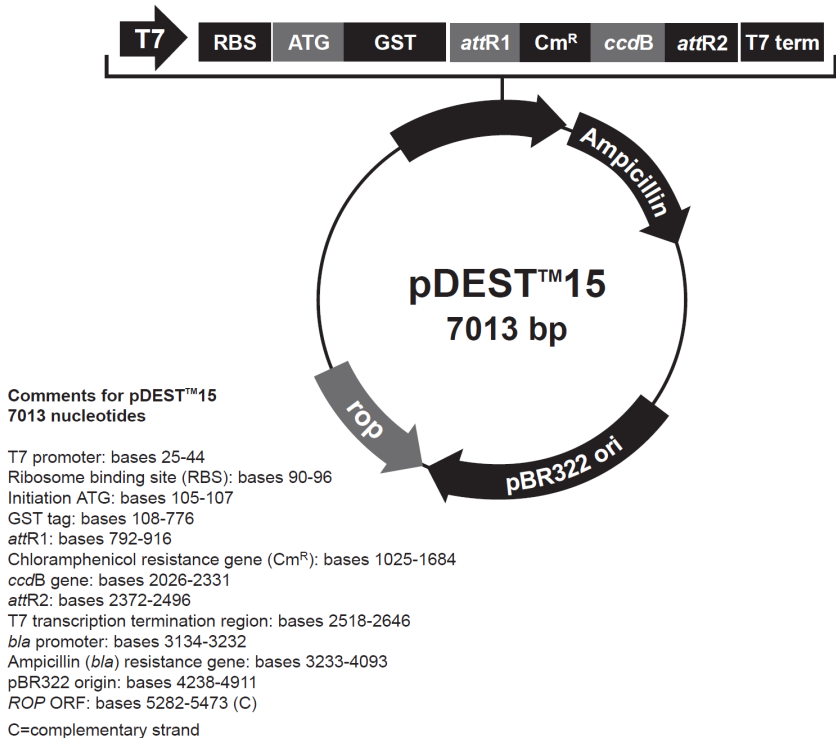


Figure 38: Map of the vector *pDEST 15*. ARA70 sequence was inserted between the segments attR1 and attR2 under control of T7 promotor.

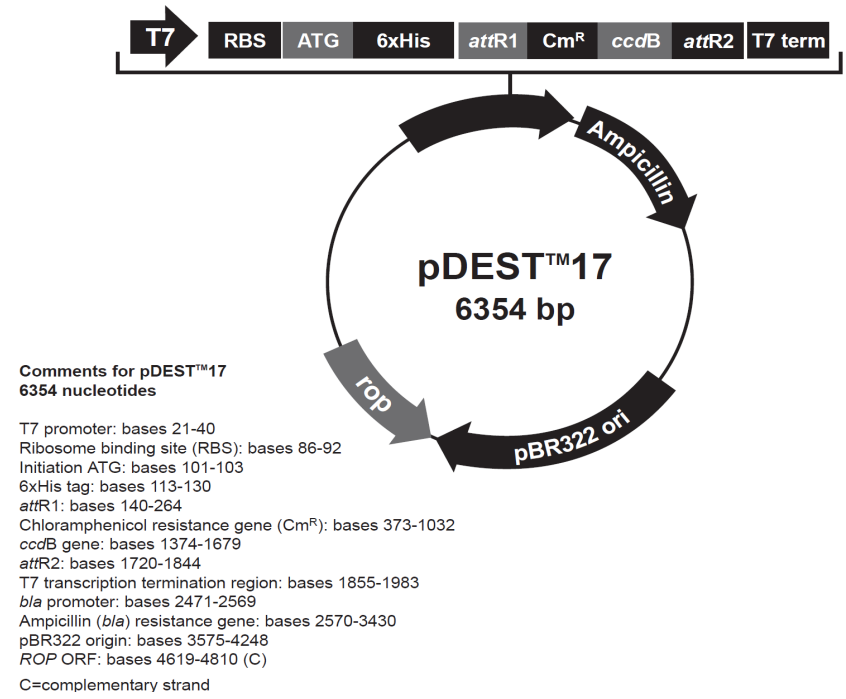


Figure 39: Map of the vector *pDEST 17*. ARA70 sequence was inserted between the segments attR1 and attR2 under control of T7 promotor.

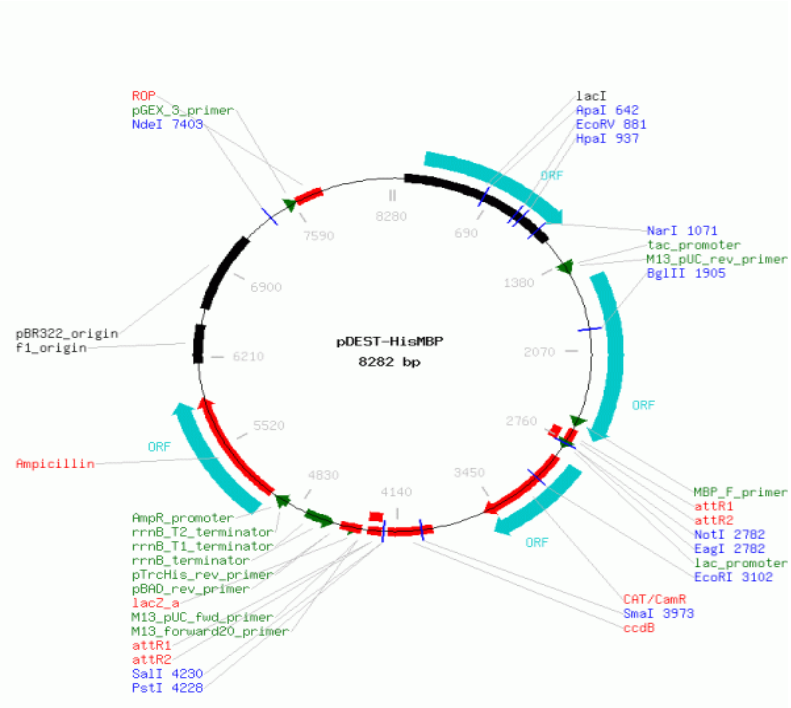


Figure 40: Map of the vector *pDEST-HisMBP*. ARA70 sequence was inserted

between the segments attR1 and attR2 under control of T7 promotor.

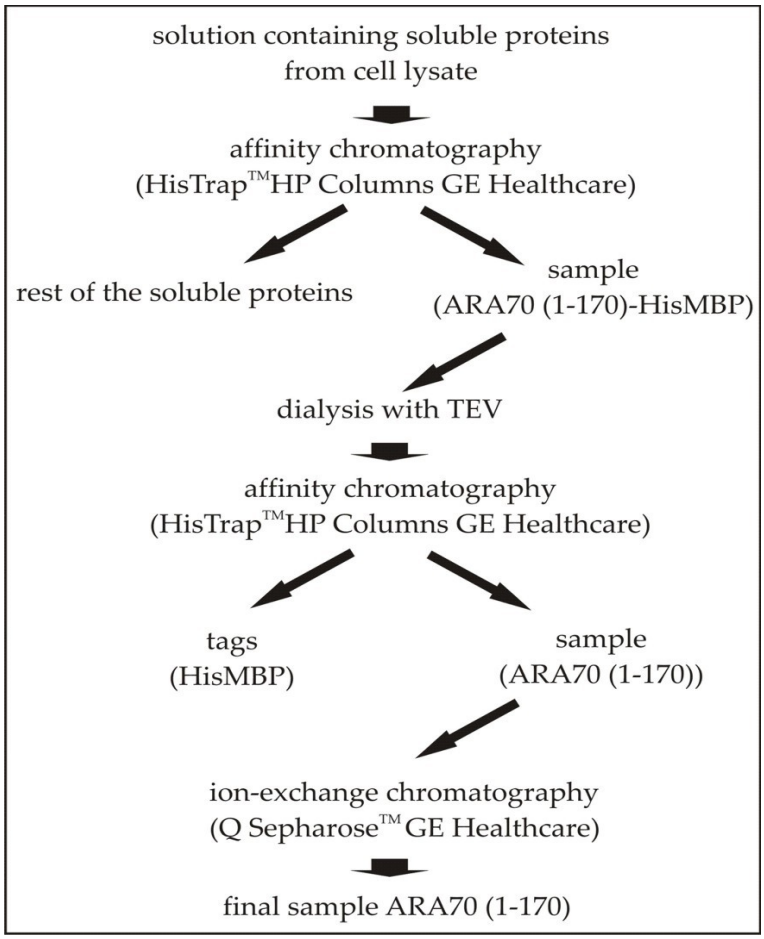


Figure 41: Scheme of purification procedure of ARA70 (1-170).

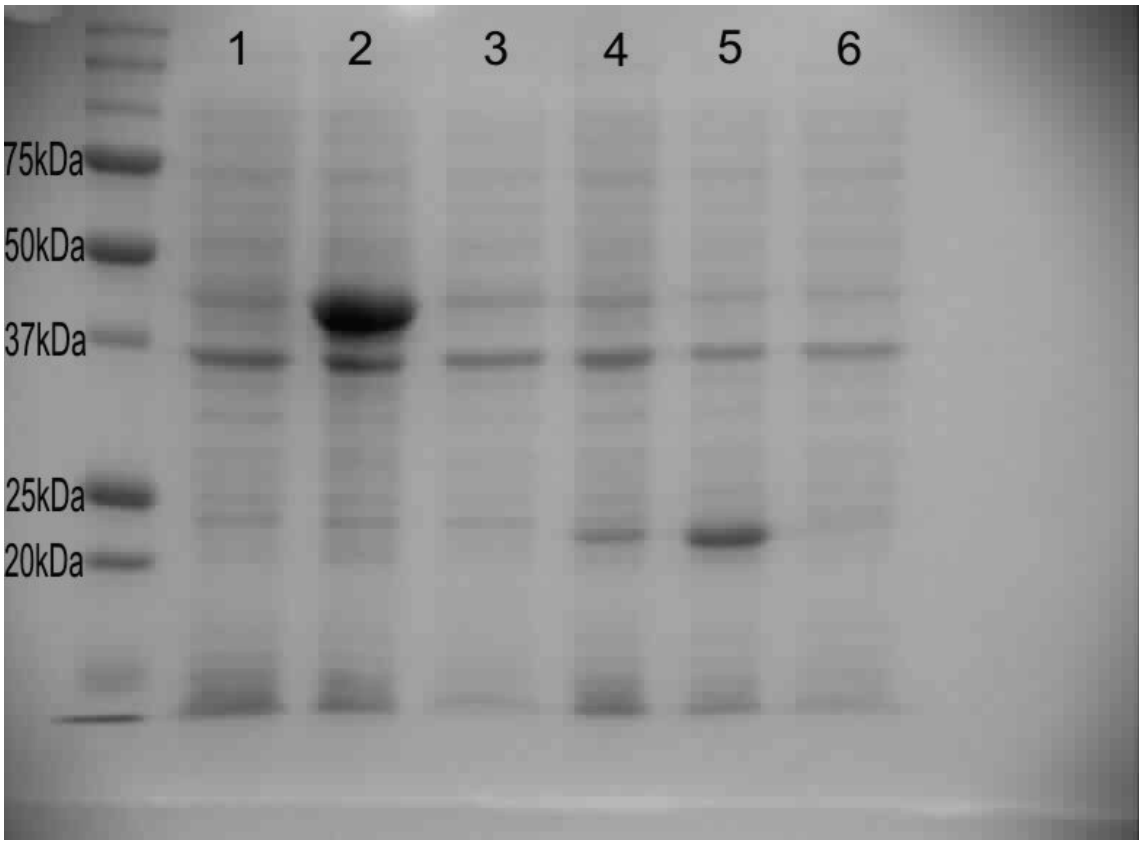


Figure 42: SDS-PAGE gel, showing protein separation; **line 1** – cell lysate from *Rosetta* cell line, not induced, pDEST15-ARA70 1-170; **line 2** – cell lysate from *Rosetta* cell line, induced at 37°C for 3 hours, 1 mM IPTG, pDEST15-ARA70 1-170 (47 kDa); **line 3** - cell lysate from *Rosetta* cell line, induced at 37°C for 3 hours, 1 mM IPTG, pDEST15-ARA70 1-170, soluble proteins; **line 4** - cell lysate from *Rosetta* cell line, not induced, pDEST17-ARA70 1-170; **line 5** - cell lysate from *Rosetta* cell line, induced at 37°C for 3 hours, 1 mM IPTG, pDEST17-ARA70 1-170 (22 kDa); **line 6** - cell lysate from *Rosetta* cell line, induced at 37°C for 3 hours, 1 mM IPTG, pDEST15-ARA70 1-170, soluble proteins.

In initial experiments ARA70 1-170 was cloned into two vectors with different tags – pDEST15 (Fig. 38) containing a GST tag and pDEST17 (Fig. 39) with a His-Tag. Unfortunately, the desired protein was expressed as insoluble molecule in these vectors (Fig. 42 - lines 3,6). There were different induction times as well as different concentration of IPTG tested to prevent from the formation of the inclusion bodies, but

without any effect (data not shown). An explanation could reside in the backbone of the chosen vector.

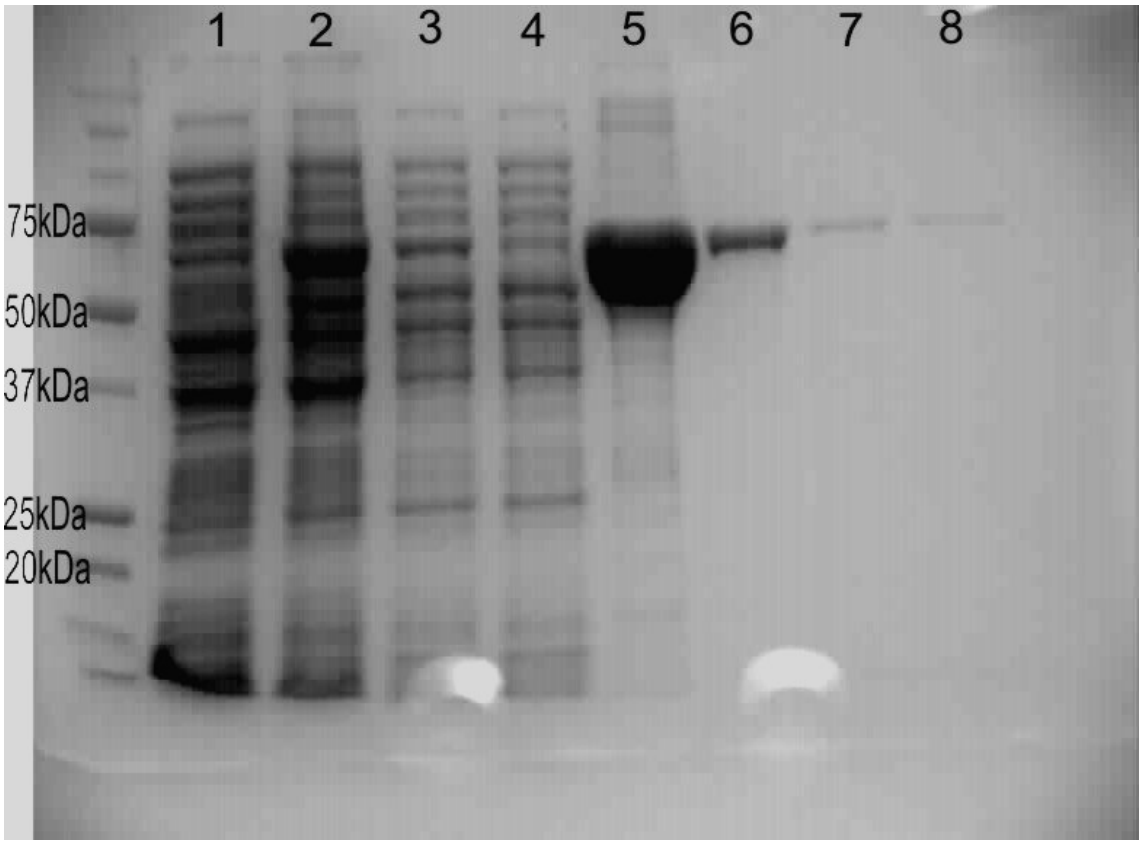


Figure 43: SDS-PAGE gel, showing protein separation; **line 1** - cell lysate from *Rosetta* cell line, not induced, pDEST-HisMBP-ARA70 1-170; **line 2** - cell lysate from *Rosetta* cell line, induced at 30°C for 6 hours, 1 mM IPTG, pDEST-HisMBP-ARA70 1-170 (66 kDa); **line 3** – solution loaded to the His-Trap column, cell lysate containing pDEST-HisMBP-ARA70 1-170; **line 4** – flow-through from the His-Trap column; **line 5** – elution from His-Trap 1st fraction containing pDEST-HisMBP-ARA70 1-170 (66 kDa); **line 6** – elution from His-Trap 2nd fraction containing pDEST-HisMBP-ARA70 1-170 (66 kDa); **line 7** – elution from His-Trap 3rd fraction containing pDEST-HisMBP-ARA70 1-170 (66 kDa); **line 8** – elution from His-Trap 4th fraction containing pDEST-HisMBP-ARA70 1-170 (66 kDa).

Vector pDEST-HisMBP was quite similar to pDEST17, it was compatible with

Gateway cloning system and it contained His-Tag. Additionaly, the construct contained maltose binding protein (MBP), which enhanced solubility. With this vector good results were achieved upon induction (Fig. 43). Figure 43 shows the results from the first step of purification of ARA70 (Fig. 43 – lines 3-8). In the line 5, one can see a quite pure and already concentrated modified protein.

In the next step, ARA70 1-170 was separated from the backbone construct with MBP and His-tag. During the dialysis, according to the protocol in chapter 5.3.10, TEV-protease was used to separate affinity tags. TEV-protease (Tobacco Etch Virus protease; 27 kDa) is a very specific protease which distinguishes 7 amino acid long sequence (ENLYFQG) for cleavage. Dialysis was important to clear solution from imidazole and to prepare it in this way for the chromatography on the second His-Trap column (Fig. 44).

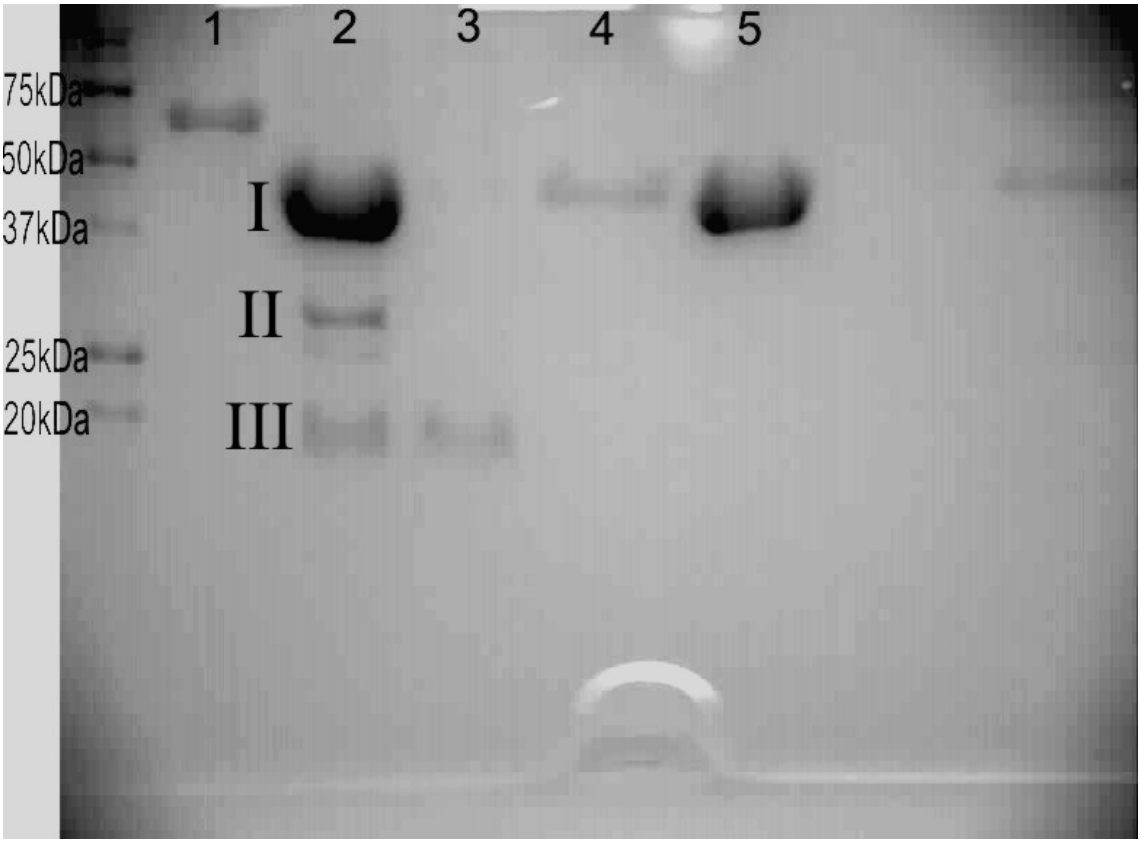


Figure 44: SDS-PAGE gel, showing protein separation; **line 1** – MBP-His-ARA70 1-170 (66 kDa) before dialysis with TEV-protease; **line 2** – solution after

dialysis, I - MBP-His (46 kDa); II - TEV-protease (27 kDa); III - ARA70 1-170 (20 kDa) (respectively); **line 3** – flow-through after 2nd His-Trap column containing ARA70 1-170 (20 kDa); **line 4** – solution after washing of the 2nd His-Trap column containing some unbound MBP-His (46 kDa); **line 5** – elution from 2nd His-Trap column containing MBP-His (46 kDa).

After the second His-Trap column, almost pure protein ARA70 1-170 was obtained without any tags or other protein modifications. Nevertheless, in solution, there were still remains of non-cleaved MBP-His-ARA70 1-170, MBP-His and TEV-protease (Fig 45.). Thus a decision was made to perform ion-exchange chromatography (IEC).

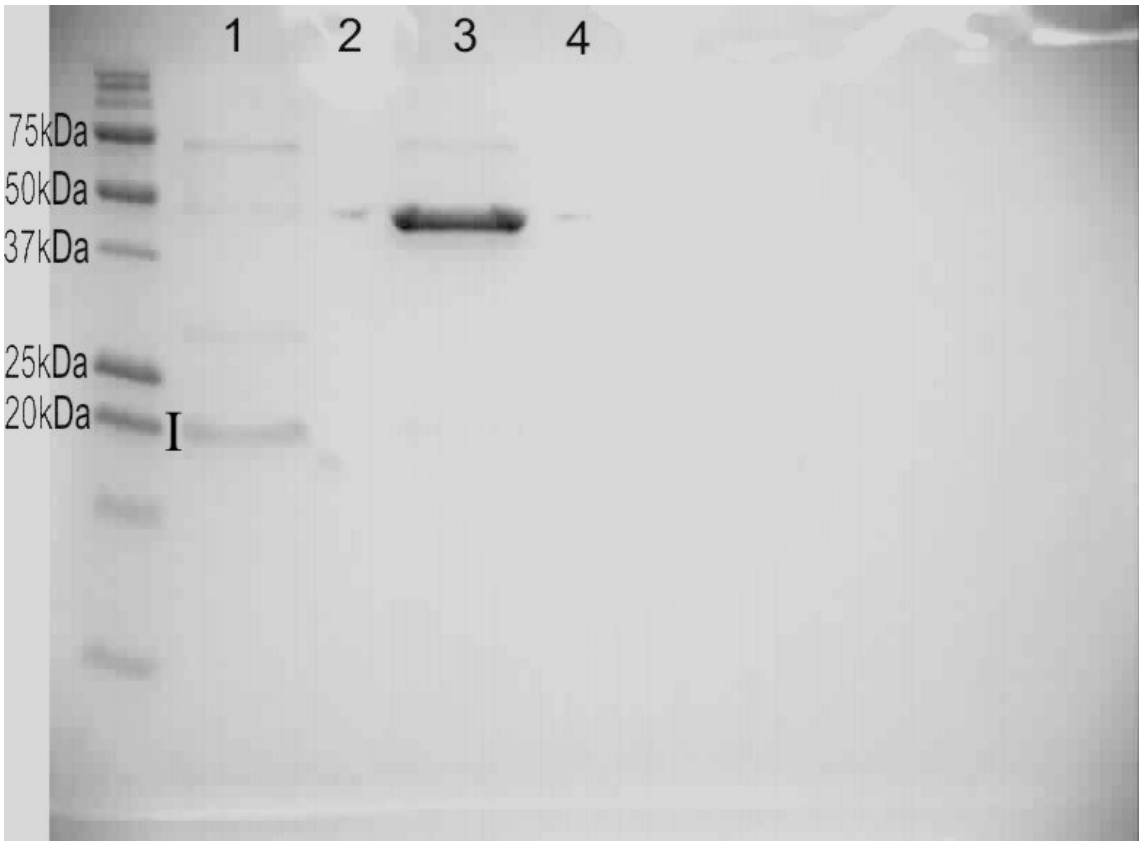


Figure 45: SDS-PAGE gel, showing protein separation; **line 1** – flow-through from 2nd His-Trap column containing ARA70 1-170 (20 kDa) (I) and impurities. **Line 3** – elution from 2nd His-Trap column containing MBP-His (46 kDa). **Line 2 and 4** –

MBP-His contamination from overloading of well 3.

To purify ARA70 1-170 protein several IEC columns were tested until satisfying result was achieved (Fig. 46). Also a gel filtration on ÄKTA apparatus (*GE Healthcare*) was tried (data not shown), but all protein was eluted in the bed volume, the reason of which is uknown, probably because of forming inclusion bodies. As a decision was made to avoid urea, the procedure was finally skipped.

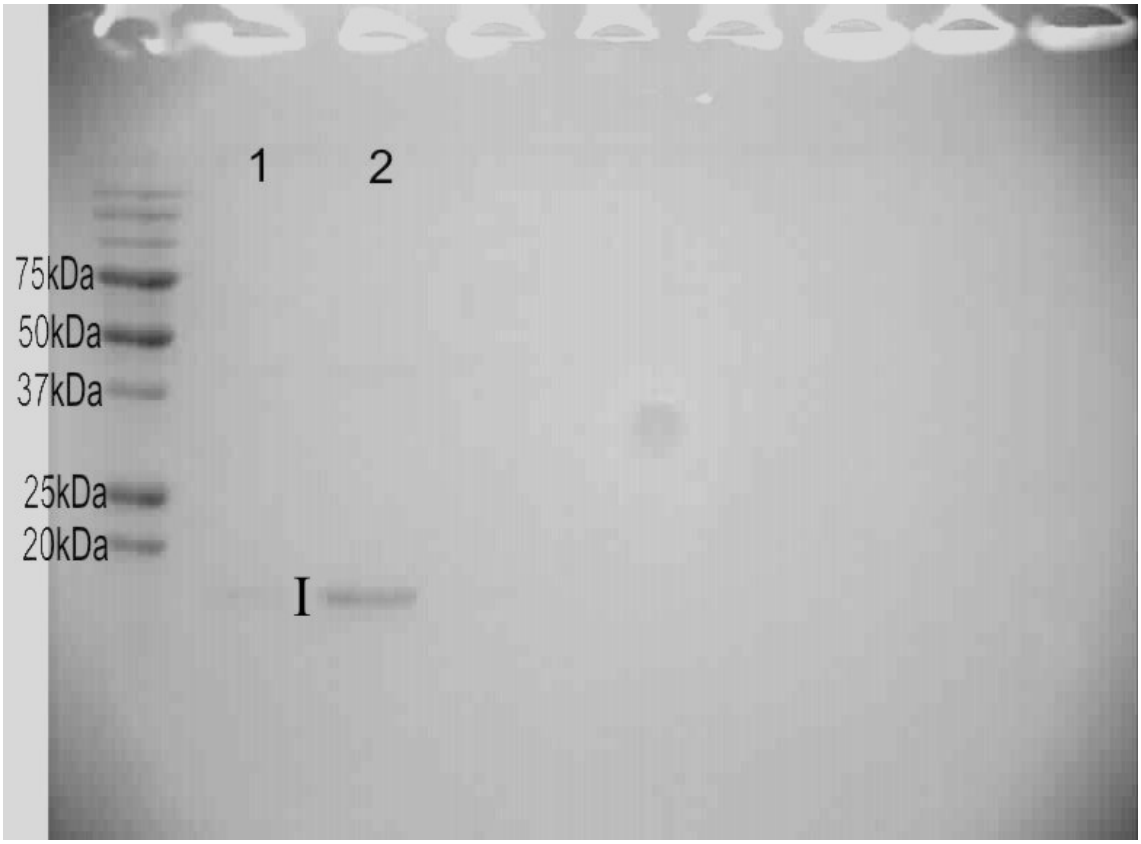


Figure 46: SDS-PAGE gel, showing protein separation; **line 1** – load to the IEC column; **line 2** – fraction eluted with 200 mM NaCl containing ARA70 1-170 (20 kDa) (I); **the other lanes** – fractions eluted with salt gradient, but without any protein.

7 Discussion

7.1 AMADH2 biotinylation

As a result of modification experiments with AMADH2 several biotin groups were introduced into the enzyme. All five modifications were localized at surface lysine residues, suggesting they are probably easy to modify. Basically, we can say that Lys133, Lys316, Lys370, Lys371, Lys438 and Lys502 are easiest accessible places for modification. Probably, by increasing volume of the biotinylation reagent added to the protein solution, we can introduce more biotin groups into the structure. In this work, Sulfo-NHS-LC-LC-biotin reagent was used, which should minimize steric hindering between reagent and protein surface thanks to the long aliphatic chain between NHS group and biotin. From the MALDI mass spectrum of intact AMADH2 (Fig. 32) it can be deduced that the reaction mixture contained both unmodified and modified protein.

Another question was how the introduced LC-LC-biotin groups influence AMADH2 activity. According to the data from kinetic measurements it is clear that the enzyme was not inactivated, but its activity was significantly reduced down to 16 % of activity of non-modified AMADH2 (Tab. 3.) which is more than six times lower (Frömmel, 2010). This reduction may by caused by biotinylation of Lys370 or Lys371. These two residues are close to the Na^+ cation binding side and thanks to the long linker, they might also interfere with NAD^+ binding, which is in a close proximity (Fig. 47). Other modifications are mostly on the other site of the enzyme molecule, but still far enough not to interfere with substrate binding channel (Fig. 48).

Another possible explanation of the observed decrease in activity of the modified enzyme resides in partial inactivation of a major portion of enzymatically active molecules during the modification process. This could occur due to a high susceptibility of the catalytic cystein residue to air oxidation (Sebela et al, 2000). As the reaction mixture was incubated for a relatively long time, the cysteine was exposed to an unfavorable enviroment.

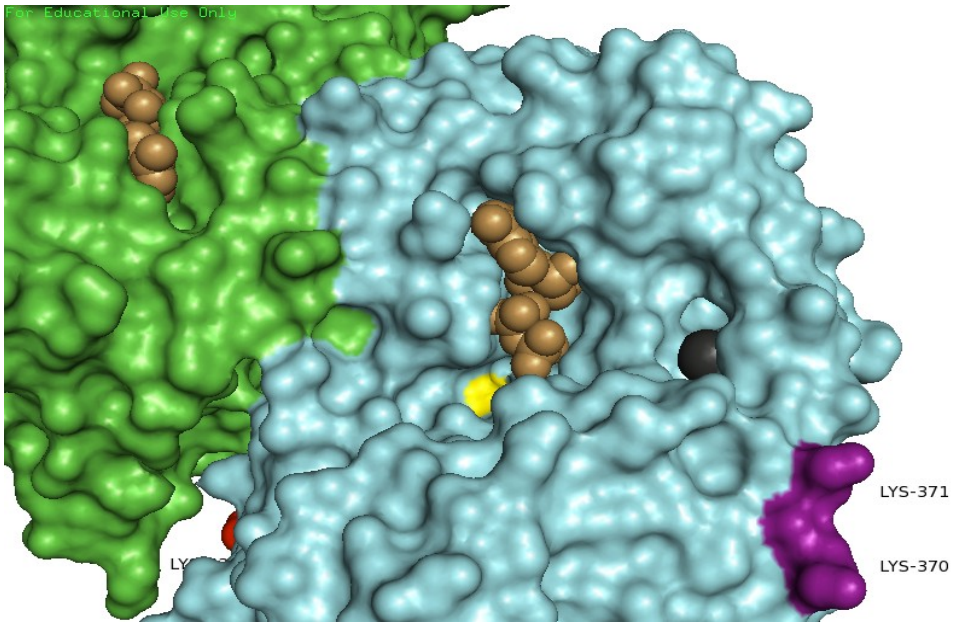


Figure 47: Closer view on surface of NAD^+ and Na^+ binding site of AMADH2. *Brown* - NAD^+ in its binding place; *black sphere* represent Na^+ cation; *purple* – modified lysine residues 371 and 370; *yellow* – surface of the catalytic Cys294 (PDB structure 3IWJ)

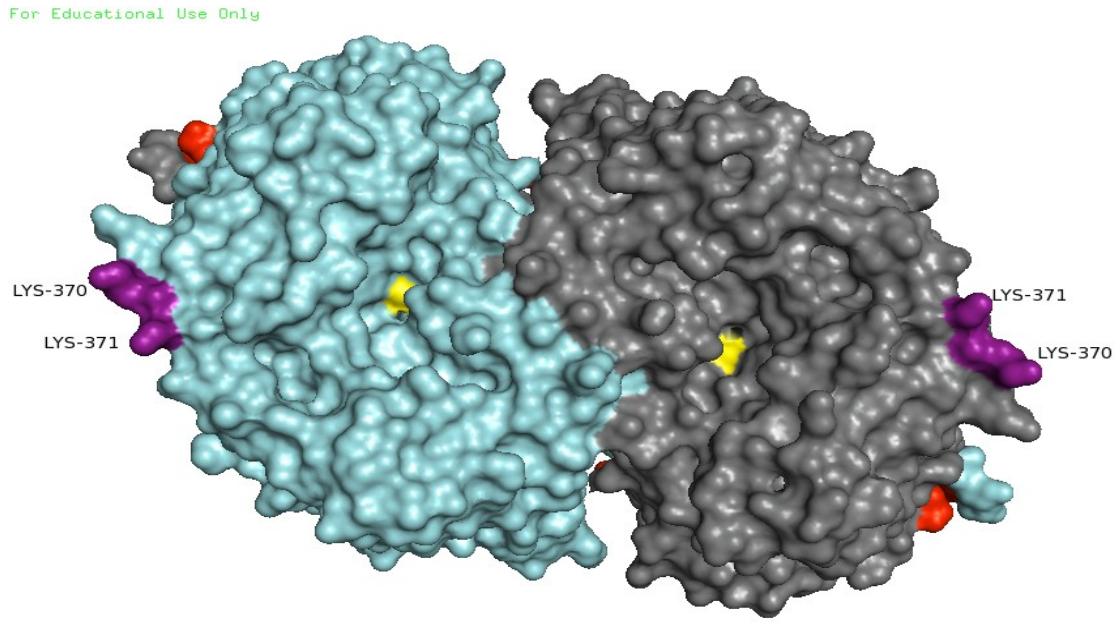


Figure 48: AMADH2 dimer with highlighted modifications. *Yellow* – surface of

the Cys294 residue in the bottom of the substrate binding channel; *red* and *purple* – the most closely modified lysine residues.

This observation also suggests that further biotinylation may completely abolish enzymatic activity of AMADH2.

Attempts were made in order to immobilize the biotinylated enzyme on avidin-sepharose beads (*Sigma, St. Louis, USA*). First trial was according to the protocol of Schriemer and Li, 1996 with some modifications. Such protocol modification were included to avoid using vortex. Instead, a prolonged incubation in the shaker for 20 minutes was performed. Afterwards enzymatic activity of the beads supernatant was measured which should be lower than control, to reflect immobilization of AMADH2 on the beads. Unfortunately, no changes were found (data not shown). After the first unsuccessful attempt, a thermo-mixer was replaced by mixer with top-to-bottom rotor which should enhance access of the enzyme to the beads. Reaction conditions were also changed. Overnight incubations at room-temperature and 4°C were tested, but did not provide positive results at all. Despite established protocols from the literature (Schriemer and Li, 1996; Schmitt et al, 1990; Swack et al, 1978) and experience of other colleagues, no positive results were achieved. To refute or confirm, whether biotinylated AMADH2 is bound to the beads, we performed MALDI-MS of the trypsin digest directly from the beads, but spectra were completely free of meaningful signals (data not shown). In my opinion, failure of the immobilization may be caused by low concentration of modified AMADH2 and poor accessibility of the biotin label to the beads surface. Since we purchased beads from Sigma company, it is unlikely that they would not work.

7.2 Production and purification of ARA70 1-170

As regards to the second part of this thesis it has been completed at IBUB (Institut de Biomedicina de la Universitat de Barcelona) in Barcelona, Spain. I was working in a team preparing protein samples for X-ray crystallography. Specifically, in this laboratory, they work with androgen receptors and androgen receptors related

proteins.

At the beginning, some issues were solved with respect to induction and solubility of the expressed protein. Solubility was in this case the most important, because crystallization is performed in water-based (polar-based) solutions. There was an assumption that the higher content of the hydrophilic residues in ARA70 1-170 protein would be enough for a good solubility preventing from the creation of the inclusion bodies. However, it was not. Most of the protein produced in pDEST15 and pDEST17 was in the inclusion bodies did not appear solution.

To solve this problem, ARA70 1-170 was cloned into a different expression vector (pDEST-HisMBP) containing a coding sequence for maltose binding protein (MBP). MBP is more than two times bigger than ARA70 1-170 protein and it is highly soluble. Using this vector the desired protein was expressed without any difficulties. In looking for the best conditions for production of ARA70 1-170, were tried many different conditions (concentration of IPTG, induction time, temperature), but eventually the optimalization ended up with conditions previously described by Nallamsetty and Waugh, 2007.

The final task was to purify the protein to remove impurities. First attempts were to purify the protein by gel filtration, but all protein was eluted already in the bed volume, the reason of which is unknown. The reason might reside in protein aggregation. Finally, a decision was made to perform ion exchange chromatography and after several cycles, a pure protein sample was obtained, as you can see on Fig. 46. Unfortunately, the sample when separeted by SDS-PAGE, provided a molecular mass, which did not correspond to a calculated prediction. Because sequence of the vector with insert was verified, there was no significant doubts about correctness or substitution.

Finally a pure sample of ARA70 1-170 was obtained for crystallization conditions screening and further analysis.

8 Conclusion

The theoretical part of this thesis was focused on basic protein modifications and the process of protein biotinylation. Also the biotin-avidin interaction is reported in detail. In the second part of theory, the focused was targeted on the issue of androgen receptors and its significance in the cell, mainly on the androgen receptor-associated co-regulator 70 (ARA70).

In the experimental part, a biotinylated derivative of AMADH2 was successfully prepared and its kinetic parameters were determined. Unfortunately, attempts to immobilize this modified enzyme to an avidin-sepharose resin failed.

Finally, a segment of ARA70 comprising the amino acid 1-170 was expressed and purified for further analysis and crystallization for X-ray crystallography.

9 List of abbreviations

5α-DHT	5α-dihydrotestosteron
ABAL	4-aminobutaraldehyde
ABALDH	4-aminobutanal dehydrogenase
ACC	acetyl-CoA carboxylase
ACP	acyl carrier protein
AES	amino-terminal enhancer of split
AF1	activation function 1 (related to the androgen receptors)
AF2	activation function 2 (related to the androgen receptors)
AI	androgen-independent
AIS	androgen insensitivity syndrome
AMADH2	aminoaldehyde dehydrogenase 2
ANPK	AR-interacting nuclear protein kinase
APAL	3-aminopropionaldehyde
AR	androgen receptor
ARA70	androgen receptor-associated co-regulator 70
AREs	androgen-response elements
ARIP	androgen receptor interacting protein
Asc-1	activating signal cointegrator-1
BAF57	BRG1-associated factor 57
BF3	binding function 3 (related to the androgen receptors)
BPL	biotin protein ligase
BTD	biotinidase
CAIS	complete androgen insensitivity syndrome
CARM-1	co-activator-associated arginine methyltransferase 1
CHCA	4-cyano hydroxycinamic acid
CNS	central nervous system
DBD	DNA-binding domain
DHT	5α-dihydrotestosteron

DMF	dimethylformamide
DMSO	dimethyl sulfoxide
E6-AP	E6 associated protein
FRAP	fluorescence recovery after photo-bleaching
FRET	fluorescence resonance energy transfer
GABA	4-aminobutyric acid
GBAL	4-guanidinobutalaldehyde
GCP	grating coupler
GFP	green fluorescent protein
GST	gluthathio-S-transferase
HAT	histone acetylase
HCS	holocarboxylase synthetase
HDAC	histone deacetylase
HR	hormone-refractory
ICAT	isotope coded affinity tag
IMAC	immobilized metal affinity chromatography
IPTG	isopropyl β-D-1-thiogalactopyranoside
LB	Luria-Bertani media
LBD	ligand-binding domain
LBDh	hinge region
MAK	male germ cell-associated kinase
MALDI-TOF-MS	matrix assisted laser desorption-ionization time of flight mass spectrometry
MBP	maltose binding protein
MCC	methylcrotonyl-CoA carboxylase
MDM2	murine double minute oncogene
MRF1	modular recognition factor 1
Nbs-F	p-nitrobenzensulphonyl fluoride
NCor	nuclear receptor corepressor
NHS	*N*-Hydroxy-succinimide
NLS	nuclear localization sequence

NMR	nuclear magnetic resonance
NPCs	nuclear-pore complexes
NRIP	nuclear receptor co-regulatory protein
NTD	N-terminal domain
NURSA	nuclear receptor Signaling Atlas database
OHT	hydroxyflutamide
PAIS	partial androgen insensitivity syndrome
PATZ	POZ-AT hook-zinc finger protein
PC	pyruvate carboxylase
PCC	propionyl-CoA carboxylase
PDIP1	PPARγ-DBD-interacting protein 1
PEG	polyethylene glycol
PMMA	poly(methyl methacrylate)
PNRC	proline-rich nuclear receptor co-regulatory protein
PPTase	phosphopantetheine transferase
PRMT-5	protein arginine methyltransferase
PSA	prostate-specific antigen
PSF	polypyrimidine tract-binding protein-assciated splicing factor
PSP1	paraspeckle protein 1
PSP2	paraspeckle protein 2
QMB	quartz microbalance
RifS	reflectometric interference spectroscopy
RIP140	receptor interacting protein 140 kDa
RM	resonant mirror
SBMA	spinal bulbar muscular atrophy
SENP1	SUMO1/sentrin-specific protease 1
SMRT	silencing mediator of retinoid and thyroid receptors
SPR	plasmon surface resonance
SRC	steroid receptor co-activator
SUMO	small ubiquitin-related modifier

TAUs	transactivation units
TEV	tobacco etch virus

10 References

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P., *Molecular biology of the cell*. 5th edn, (Garland Science, 2008).

Anson, M. L. The Reactions of Iodine and Iodoacetamide with Native Egg Albumin. *J Gen Physiol* **23**, 321-331 (1940).

Askew, E. B., Gampe, R. T., Jr., Stanley, T. B., Faggart, J. L. & Wilson, E. M. Modulation of androgen receptor activation function 2 by testosterone and dihydrotestosterone. *J Biol Chem* **282**, 25801-25816, doi:M703268200 [pii] 10.1074/jbc.M703268200 (2007).

Awal, H. M. A., Hirasawa, E., 1,3-Diaminopropane is a suicide substrate for pea diamine oxidase. *Phyt* **39,** 489-490, doi:10.1016/0031-9422(95)00003-P, (1995).

Barker, K. D. *et al.* Protein binding and the electronic properties of iron(II) complexes: an electrochemical and optical investigation of outer sphere effects. *Bioconjug Chem* **20**, 1930-1939, doi:10.1021/bc900270a (2009).

Bashiri, G., Squire, C. J., Baker, E. N. & Moreland, N. J. Expression, purification and crystallization of native and selenomethionine labeled Mycobacterium tuberculosis FGD1 (Rv0407) using a Mycobacterium smegmatis expression system. *Protein Expr Purif* **54**, 38-44, doi:S1046-5928(07)00025-3 [pii]10.1016/j.pep.2007.01.014 (2007).

Berrevoets, C. A., Doesburg, P., Steketee, K., Trapman, J. & Brinkmann, A. O. Functional interactions of the AF-2 activation domain core region of the human androgen receptor with the amino-terminal domain and with the transcriptional

coactivator TIF2 (transcriptional intermediary factor2). *Mol Endocrinol* **12**, 1172-1183 (1998).

Birkert, O. *et al.* A streptavidin surface on planar glass substrates for the detection of biomolecular interaction. *Anal Biochem* **282**, 200-208, doi:10.1006/abio.2000.4615 S0003-2697(00)94615-3 [pii] (2000).

Bogusiewicz, A., Mock, N. I. & Mock, D. M. A biotin-protein bond with stability in plasma. *Anal Biochem* **337**, 98-102, doi:S0003-2697(04)00841-3 [pii] 10.1016/j.ab.2004.10.023 (2005).

Bouchereau A., Aziz A., Larher F., Martin-Tanguy J., Polyamines and environmental challenges: recent development. *Plant Sci*. **140,** 103–125 (1999).

Boulikas, T., Bastin, B., Boulikas, P. & Dupuis, G. Increase in histone poly (ADP-ribosylation) in mitogen-activated lymphoid cells. *Exp Cell Res* **187**, 77-84, doi:0014-4827(90)90119-U [pii] (1990).

Brady, M. E. *et al.* Tip60 is a nuclear hormone receptor coactivator. *J Biol Chem* **274**, 17599-17604 (1999).

Brauner, F. *et al.* Pea seedling aminoaldehyde dehydrogenase: primary structure and active site residues. *Plant Phsiol. Biochem.***41**, 1-10, (2003).

Brecht, A. & Gauglitz, G. Optical probes and transducers. *Biosens Bioelectron* **10**, 923-936 (1995).

Bruch, R. C. & White, H. B., 3rd. Compositional and structural heterogeneity of avidin glycopeptides. *Biochemistry* **21**, 5334-5341 (1982).

Buchanan, G. *et al.* Structural and functional consequences of glutamine tract variation

in the androgen receptor. *Hum Mol Genet* **13**, 1677-1692, doi:10.1093/hmg/ddh181ddh181 [pii] (2004).

Carascossa, S. *et al.* Receptor-interacting protein 140 is a repressor of the androgen receptor activity. *Mol Endocrinol* **20**, 1506-1518, doi:me.2005-0286 [pii] 10.1210/me.2005-0286 (2006).

Chapman-Smith, A. & Cronan, J. E., Jr. The enzymatic biotinylation of proteins: a post-translational modification of exceptional specificity. *Trends Biochem Sci* **24**, 359-363, doi:S0968-0004(99)01438-3 [pii] (1999a).

Cheng, J., Wang, D., Wang, Z. & Yeh, E. T. SENP1 enhances androgen receptor-dependent transcription through desumoylation of histone deacetylase 1. *Mol Cell Biol* **24**, 6021-6028, doi:10.1128/MCB.24.13.6021-6028.2004 24/13/6021 [pii] (2004).

Cutress, M. L., Whitaker, H. C., Mills, I. G., Stewart, M. & Neal, D. E. Structural basis for the nuclear import of the human androgen receptor. *J Cell Sci* **121**, 957-968, doi:jcs.022103 [pii] 10.1242/jcs.022103 (2008).

Davis, G. D., Elisee, C., Newham, D. M. & Harrison, R. G. New fusion protein systems designed to give soluble expression in Escherichia coli. *Biotechnol Bioeng* **65**, 382-388, doi:10.1002/(SICI)1097-0290(19991120)65:4<382::AID-BIT2>3.0.CO;2-I [pii] (1999).

Deeb, A. *et al.* A novel mutation in the human androgen receptor suggests a regulatory role for the hinge region in amino-terminal and carboxy-terminal interactions. *J Clin Endocrinol Metab* **93**, 3691-3696, doi:jc.2008-0737 [pii]10.1210/jc.2008-0737 (2008).

DeLange, R. J. & Huang, T. S. Egg white avidin. 3. Sequence of the 78-residue middle cyanogen bromide peptide. Complete amino acid sequence of the protein subunit. *J Biol Chem* **246**, 698-709 (1971).

DeLange, R. J. Egg white avidin. I. Amino acid composition; sequence of the amino- and carboxyl-terminal cyanogen bromide peptides. *J Biol Chem* **245**, 907-916 (1970).

Desarnaud, F. *et al.* Protein purification using combined streptavidin (or avidin)-Sepharose and thiopropyl-Sepharose affinity chromatography. *J Chromatogr* **603**, 95-104 (1992).

Diamandis, E. P. & Christopoulos, T. K. The biotin-(strept)avidin system: principles and applications in biotechnology. *Clin Chem* **37**, 625-636 (1991).

Domanskyi, A., Virtanen, K. T., Palvimo, J. J. & Janne, O. A. Biochemical characterization of androgen receptor-interacting protein 4. *Biochem J* **393**, 789-795, doi:BJ20050823 [pii] 10.1042/BJ20050823 (2006).

Duff, J., Davies, P., Watt, K. & McEwan, I. J. Structural dynamics of the human androgen receptor: implications for prostate cancer and neurodegenerative disease. *Biochem Soc Trans* **34**, 1098-1102, doi:BST0341098 [pii]10.1042/BST0341098 (2006).

Dupuy D'Angeac, A. *et al.* Biotinylation of glycan chains in beta2 glycoprotein I induces dimerization of the molecule and its detection by the human autoimmune anti-cardiolipin antibody EY2C9. *Biochem J* **393**, 117-127, doi:BJ20050932 [pii]10.1042/BJ20050932 (2006).

Elia, G. Biotinylation reagents for the study of cell surface proteins. *Proteomics* **8**, 4012-4024, doi:10.1002/pmic.200800097 (2008).

Elia, G. Protein biotinylation. *Curr Protoc Protein Sci* **Chapter 3**, Unit 3 6, doi:10.1002/0471140864.ps0306s60 (2010).

Estebanez-Perpina, E. *et al.* Structural insight into the mode of action of a direct inhibitor of coregulator binding to the thyroid hormone receptor. *Mol Endocrinol* **21**,

2919-2928, doi:me.2007-0174 [pii]10.1210/me.2007-0174 (2007).

Fang, Y., Fliss, A. E., Robins, D. M. & Caplan, A. J. Hsp90 regulates androgen receptor hormone binding affinity in vivo. *J Biol Chem* **271**, 28697-28702 (1996).

Fliss, A. E., Rao, J., Melville, M. W., Cheetham, M. E. & Caplan, A. J. Domain requirements of DnaJ-like (Hsp40) molecular chaperones in the activation of a steroid hormone receptor. *J Biol Chem* **274**, 34045-34052 (1999).

Frömmel, J., *Příprava nových syntetických substrátů aminoaldehyddehydrogenasy (Master thesis)*, (Palacky university Olomouc, Faculty of Natural Sciences, Olomouc, CZ, 2010).

Gao, W., Bohl, C. E. & Dalton, J. T. Chemistry and structural biology of androgen receptor. *Chem Rev* **105**, 3352-3370, doi:10.1021/cr020456u (2005).

George, N., Pick, H., Vogel, H., Johnsson, N. & Johnsson, K. Specific labeling of cell surface proteins with chemically diverse compounds. *J Am Chem Soc* **126**, 8896-8897, doi:10.1021/ja048396s (2004).

Georgescu, S. P. *et al.* Modulator recognition factor 1, an AT-rich interaction domain family member, is a novel corepressor for estrogen receptor alpha. *Mol Endocrinol* **19**, 2491-2501, doi:me.2004-0311 [pii] 10.1210/me.2004-0311 (2005).

Giovannucci, E. *et al.* The CAG repeat within the androgen receptor gene and its relationship to prostate cancer. *Proc Natl Acad Sci U S A* **94**, 3320-3323 (1997).

Gitlin, G., Bayer, E. A. & Wilchek, M. Studies on the biotin-binding site of avidin. Tryptophan residues involved in the active site. *Biochem J* **250**, 291-294 (1988).

Gitlin, G., Bayer, E. A. & Wilchek, M. Studies on the biotin-binding sites of avidin and

streptavidin. Tyrosine residues are involved in the binding site. *Biochem J* **269**, 527-530 (1990).

Gonzalez-Segura, L., Rudino-Pinera, E., Munoz-Clares, R. A. & Horjales, E. The crystal structure of a ternary complex of betaine aldehyde dehydrogenase from Pseudomonas aeruginosa Provides new insight into the reaction mechanism and shows a novel binding mode of the 2'-phosphate of NADP+ and a novel cation binding site. *J Mol Biol* **385**, 542-557, doi:S0022-2836(08)01364-8 [pii] 10.1016/j.jmb.2008.10.082 (2009).

Goo, Y. H. *et al.* Interactions between activating signal cointegrator-2 and the tumor suppressor retinoblastoma in androgen receptor transactivation. *J Biol Chem* **279**, 7131-7135, doi:10.1074/jbc.M312563200 M312563200 [pii] (2004).

Gottlieb, B., Beitel, L. K., Wu, J. H. & Trifiro, M. The androgen receptor gene mutations database (ARDB): 2004 update. *Hum Mutat* **23**, 527-533, doi:10.1002/humu.20044 (2004).

Grino, P. B., Griffin, J. E. & Wilson, J. D. Testosterone at high concentrations interacts with the human androgen receptor similarly to dihydrotestosterone. *Endocrinology* **126**, 1165-1172 (1990).

Haelens, A. *et al.* DNA recognition by the androgen receptor: evidence for an alternative DNA-dependent dimerization, and an active role of sequences flanking the response element on transactivation. *Biochem J* **369**, 141-151, doi:10.1042/BJ20020912 BJ20020912 [pii] (2003).

Haelens, A., Tanner, T., Denayer, S., Callewaert, L. & Claessens, F. The hinge region regulates DNA binding, nuclear translocation, and transactivation of the androgen receptor. *Cancer Res* **67**, 4514-4523, doi:67/9/4514 [pii]10.1158/0008-5472.CAN-06-1701 (2007).

Han, G. *et al.* Hormone status selects for spontaneous somatic androgen receptor variants that demonstrate specific ligand and cofactor dependent activities in autochthonous prostate cancer. *J Biol Chem* **276**, 11204-11213, doi:10.1074/jbc.M008207200 M008207200 [pii] (2001).

Han, G. *et al.* Mutation of the androgen receptor causes oncogenic transformation of the prostate. *Proc Natl Acad Sci U S A* **102**, 1151-1156, doi:0408925102 [pii] 10.1073/pnas.0408925102 (2005).

Hassan, Y. I. & Zempleni, J. A novel, enigmatic histone modification: biotinylation of histones by holocarboxylase synthetase. *Nutr Rev* **66**, 721-725, doi:NURE127 [pii] 10.1111/j.1753-4887.2008.00127.x (2008).

He, B. *et al.* An androgen receptor NH2-terminal conserved motif interacts with the COOH terminus of the Hsp70-interacting protein (CHIP). *J Biol Chem* **279**, 30643-30653, doi:10.1074/jbc.M403117200 M403117200 [pii] (2004a).

He, B. *et al.* Probing the functional link between androgen receptor coactivator and ligand-binding sites in prostate cancer and androgen insensitivity. *J Biol Chem* **281**, 6648-6663, doi:M511738200 [pii] 10.1074/jbc.M511738200 (2006)

He, B. *et al.* Structural basis for androgen receptor interdomain and coactivator interactions suggests a transition in nuclear receptor activation function dominance. *Mol Cell* **16**, 425-438, doi:S109727650400588X [pii]10.1016/j.molcel.2004.09.036 (2004b).

He, B., Kemppainen, J. A. & Wilson, E. M. FXXLF and WXXLF sequences mediate the NH2-terminal interaction with the ligand binding domain of the androgen receptor. *J Biol Chem* **275**, 22986-22994, doi:10.1074/jbc.M002807200 M002807200 [pii] (2000).

He, B., Minges, J. T., Lee, L. W. & Wilson, E. M. The FXXLF motif mediates androgen receptor-specific interactions with coregulators. *J Biol Chem* **277**, 10226-10235, doi:10.1074/jbc.M111975200 M111975200 [pii] (2002a).

Heemers, H. V. & Tindall, D. J. Androgen receptor (AR) coregulators: a diversity of functions converging on and regulating the AR transcriptional complex. *Endocr Rev* **28**, 778-808, doi:er.2007-0019 [pii]10.1210/er.2007-0019 (2007).

Helppolainen, S. H. *et al.* Bradavidin II from Bradyrhizobium japonicum: a new avidin-like biotin-binding protein. *Biochim Biophys Acta* **1784**, 1002-1010, doi:S1570-9639(08)00128-3 [pii] 10.1016/j.bbapap.2008.04.010 (2008).

Hiller, Y., Gershoni, J. M., Bayer, E. A. & Wilchek, M. Biotin binding to avidin. Oligosaccharide side chain not required for ligand association. *Biochem J* **248**, 167-171 (1987).

Hu, Y. C. *et al.* Functional domain and motif analyses of androgen receptor coregulator ARA70 and its differential expression in prostate cancer. *J Biol Chem* **279**, 33438-33446, doi:10.1074/jbc.M401781200 M401781200 [pii] (2004).

Huang, S. M. & Stallcup, M. R. Mouse Zac1, a transcriptional coactivator and repressor for nuclear receptors. *Mol Cell Biol* **20**, 1855-1867 (2000).

Huggins, C., Stephens, R. C. & Hodges, C. V. (1941) *Arch. Surg.* **43,** 209.

Ishitani, K. *et al.* p54nrb acts as a transcriptional coactivator for activation function 1 of the human androgen receptor. *Biochem Biophys Res Commun* **306**, 660-665, doi:S0006291X03010210 [pii] (2003).

Jasuja, R. *et al.* Kinetic and thermodynamic characterization of dihydrotestosterone-induced conformational perturbations in androgen receptor ligand-binding domain. *Mol*

Endocrinol **23**, 1231-1241, doi:me.2008-0304 [pii] 10.1210/me.2008-0304 (2009).

Jenster, G., van der Korput, H. A., Trapman, J. & Brinkmann, A. O. Identification of two transcription activation units in the N-terminal domain of the human androgen receptor. *J Biol Chem* **270**, 7341-7346 (1995).

Kaarbo, M., Klokk, T. I. & Saatcioglu, F. Androgen signaling and its interactions with other signaling pathways in prostate cancer. *Bioessays* **29**, 1227-1238, doi:10.1002/bies.20676 (2007).

Kang, S. *et al.* Synthesis of biotin-tagged chemical cross-linkers and their applications for mass spectrometry. *Rapid Commun Mass Spectrom* **23**, 1719-1726, doi:10.1002/rcm.4066 (2009).

Khan, O. Y. *et al.* Multifunction steroid receptor coactivator, E6-associated protein, is involved in development of the prostate gland. *Mol Endocrinol* **20**, 544-559, doi:me.2005-0110 [pii] 10.1210/me.2005-0110 (2006).

Ko, L., Cardona, G. R., Henrion-Caude, A. & Chin, W. W. Identification and characterization of a tissue-specific coactivator, GT198, that interacts with the DNA-binding domains of nuclear receptors. *Mol Cell Biol* **22**, 357-369 (2002).

Kothapalli, N. *et al.* Biological functions of biotinylated histones. *J Nutr Biochem* **16**, 446-448, doi:S0955-2863(05)00098-7 [pii] 10.1016/j.jnutbio.2005.03.025 (2005).

Kouzarides, T. Acetylation: a regulatory modification to rival phosphorylation? *EMBO J* **19**, 1176 1179, doi:10.1093/emboj/19.6.1176 (2000).

Kung, H. J. & Evans, C. P. Oncogenic activation of androgen receptor. *Urol Oncol* **27**, 48-52, doi:S1078-1439(08)00156-7 [pii] 10.1016/j.urolonc.2008.06.002 (2009).

La Spada, A. R., Wilson, E. M., Lubahn, D. B., Harding, A. E. & Fischbeck, K. H. Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature* **352**, 77-79, doi:10.1038/352077a0 (1991).

Lagunas, A., Comelles, J., Martinez, E. & Samitier, J. Universal chemical gradient platforms using poly(methyl methacrylate) based on the biotin-streptavidin interaction for biological applications. *Langmuir* **26**, 14154-14161, doi:10.1021/la102640w (2010).

Laitinen, O. H. *et al.* Biotin induces tetramerization of a recombinant monomeric avidin. A model for protein-protein interactions. *J Biol Chem* **276**, 8219-8224, doi:10.1074/jbc.M007930200 M007930200 [pii] (2001).

Laitinen, O. H. *et al.* Mutation of a critical tryptophan to lysine in avidin or streptavidin may explain why sea urchin fibropellin adopts an avidin-like domain. *FEBS Lett* **461**, 52-58, doi:S0014 5793(99)01423-4 [pii] (1999).

Langley, E., Kemppainen, J. A. & Wilson, E. M. Intermolecular NH2-/carboxyl-terminal interactions in androgen receptor dimerization revealed by mutations that cause androgen insensitivity. *J Biol Chem* **273**, 92-101 (1998).

Lee, D. Y., Hayes, J. J., Pruss, D. & Wolffe, A. P. A positive role for histone acetylation in transcription factor access to nucleosomal DNA. *Cell* **72**, 73-84, doi:0092-8674(93)90051-Q [pii] (1993).

Lee, Y. S. *et al.* Activating signal cointegrator 1 is highly expressed in murine testicular Leydig cells and enhances the ligand-dependent transactivation of androgen receptor. *Biol Reprod* **67**, 1580-1587 (2002).

Ligr, M. *et al.* Tumor suppressor function of androgen receptor coactivator ARA70alpha in prostate cancer. *Am J Pathol* **176**, 1891-1900, doi:S0002-9440(10)60501-9 [pii] 10.2353/ajpath.2010.090293 (2010).

Link, K. A. *et al.* BAF57 governs androgen receptor action and androgen-dependent proliferation through SWI/SNF. *Mol Cell Biol* **25**, 2200-2215, doi:25/6/2200 [pii] 10.1128/MCB.25.6.2200-2215.2005 (2005).

Livnah, O., Bayer, E. A., Wilchek, M. & Sussman, J. L. Three-dimensional structures of avidin and the avidin-biotin complex. *Proc Natl Acad Sci U S A* **90**, 5076-5080 (1993).

Lumbroso, S. *et al.* Molecular analysis of the androgen receptor gene in Kennedy's disease. Report of two families and review of the literature. *Horm Res* **47**, 23-29 (1997).

Lund, A. *et al.* Multiple founder effects in spinal and bulbar muscular atrophy (SBMA, Kennedy disease) around the world. *Eur J Hum Genet* **9**, 431-436, doi:10.1038/sj.ejhg.5200656 (2001).

Ma, A. H. *et al.* Male germ cell-associated kinase, a male-specific kinase regulated by androgen, is a coactivator of androgen receptor in prostate cancer cells. *Cancer Res* **66**, 8439-8447, doi:66/17/8439 [pii] 10.1158/0008-5472.CAN-06-1636 (2006).

Maatta, J. A. *et al.* Structural and functional characteristics of xenavidin, the first frog avidin from Xenopus tropicalis. *BMC Struct Biol* **9**, 63, doi:1472-6807-9-63 [pii] 10.1186/1472-6807-9-63 (2009).

Marino, K., Bones, J., Kattla, J. J. & Rudd, P. M. A systematic approach to protein glycosylation analysis: a path through the maze. *Nat Chem Biol* **6**, 713-723, doi:nchembio.437 [pii] 10.1038/nchembio.437 (2010).

Mo, K. *et al.* Microarray analysis of gene expression by skeletal muscle of three mouse models of Kennedy disease/spinal bulbar muscular atrophy. *PLoS One* **5**, e12922, doi:e12922 [pii] 10.1371/journal.pone.0012922 (2010).

Moehren, U., Papaioannou, M., Reeb, C. A., Hong, W. & Baniahmad, A. Alien interacts with the human androgen receptor and inhibits prostate cancer cell growth. *Mol Endocrinol* **21**, 1039-1048, doi:me.2006-0468 [pii] 10.1210/me.2006-0468 (2007).

Moilanen, A. M., Karvonen, U., Poukka, H., Janne, O. A. & Palvimo, J. J. Activation of androgen receptor function by a novel nuclear protein kinase. *Mol Biol Cell* **9**, 2527-2543 (1998).

Nallamsetty, S. & Waugh, D. S. A generic protocol for the expression and purification of recombinant proteins in Escherichia coli using a combinatorial His6-maltose binding protein fusion tag. *Nat Protoc* **2**, 383-391, doi:nprot.2007.50 [pii]10.1038/nprot.2007.50 (2007).

Onate, S. A., Tsai, S. Y., Tsai, M. J. & O'Malley, B. W. Sequence and characterization of a coactivator for the steroid hormone receptor superfamily. *Science* **270**, 1354-1357 (1995).

Parodi, A. J. Reglucosylation of glycoproteins and quality control of glycoprotein folding in the endoplasmic reticulum of yeast cells. *Biochim Biophys Acta* **1426**, 287-295, doi:S0304-4165(98)00130-5 [pii] (1999).

Pereira de Jesus-Tran, K. *et al.* Comparison of crystal structures of human androgen receptor ligand-binding domain complexed with various agonists reveals molecular determinants responsible for binding affinity. *Protein Sci* **15**, 987-999, doi:15/5/987 [pii] 10.1110/ps.051905906 (2006).

Pero, R. *et al.* PATZ attenuates the RNF4-mediated enhancement of androgen receptor-dependent transcription. *J Biol Chem* **277**, 3280-3285, doi:10.1074/jbc.M109491200 M109491200 [pii] (2002).

Petrivalsky, M., Brauner, F., Luhova, L., Gagneul, D. & Sebela, M. Aminoaldehyde

dehydrogenase activity during wound healing of mechanically injured pea seedlings. *J Plant Physiol* **164**, 1410-1418, doi:S0176-1617(07)00158-7 [pii] 10.1016/j.jplph.2007.01.018 (2007).

Poukka, H., Karvonen, U., Janne, O. A. & Palvimo, J. J. Covalent modification of the androgen receptor by small ubiquitin-like modifier 1 (SUMO-1). *Proc Natl Acad Sci U S A* **97**, 14145-14150, doi:10.1073/pnas.97.26.14145 97/26/14145 [pii] (2000).

Prescott, J. & Coetzee, G. A. Molecular chaperones throughout the life cycle of the androgen receptor. *Cancer Lett* **231**, 12-19, doi:S0304-3835(05)00006-6 [pii] 10.1016/j.canlet.2004.12.037 (2006).

Qi, W., Wu, H., Yang, L., Boyd, D. D. & Wang, Z. A novel function of caspase-8 in the regulation of androgen-receptor-driven gene expression. *EMBO J* **26**, 65-75, doi:7601483 [pii] 10.1038/sj.emboj.7601483 (2007).

Rahman, M. M. *et al.* Reducing the agonist activity of antiandrogens by a dominant-negative androgen receptor coregulator ARA70 in prostate cancer cells. *J Biol Chem* **278**, 19619-19626, doi:10.1074/jbc.M210941200 M210941200 [pii] (2003).

Roberts, M. R. Does GABA Act as a Signal in Plants?: Hints from Molecular Studies. *Plant Signal Behav* **2**, 408-409 (2007).

Robinson-Rechavi, M., Escriva Garcia, H. & Laudet, V. The nuclear receptor superfamily. *J Cell Sci* **116**, 585-586 (2003).

Sack, J. S. *et al.* Crystallographic structures of the ligand-binding domains of the androgen receptor and its T877A mutant complexed with the natural agonist dihydrotestosterone. *Proc Natl Acad Sci U S A* **98**, 4904-4909, doi:10.1073/pnas.081565498 98/9/4904 [pii] (2001).

Sakai, N. *et al.* Bilateral testicular tumors in androgen insensitivity syndrome. *Int J Urol* **7**, 390-392 (2000).

Schaufele, F. *et al.* The structural basis of androgen receptor activation: intramolecular and intermolecular amino-carboxy interactions. *Proc Natl Acad Sci U S A* **102**, 9802-9807, doi:0408819102 [pii]10.1073/pnas.0408819102 (2005).

Scheich, C., Sievert, V. & Bussow, K. An automated method for high-throughput protein purification applied to a comparison of His-tag and GST-tag affinity chromatography. *BMC Biotechnol* **3**, 12, doi:10.1186/1472-6750-3-12 (2003).

Schmitt, S., Goetschel, J. D., Rommelt, E. M. & Adler, C. One-step purification of trypsin and alpha-chymotrypsin by affinity chromatography on Eupergit-aprotinin, a novel carrier for purification of serine proteases. *J Chromatogr* **510**, 239-242 (1990).

Schriemer, D. C. & Li, L. Combining avidin-biotin chemistry with matrix-assisted laser desorption/ionization mass spectrometry. *Anal Chem* **68**, 3382-3387 (1996).

Schrodinger, LLC. *The PyMOL Molecular Graphics System, Version 1.3r1* (2010).

Sebela, M. *et al.* Characterisation of a homogeneous plant aminoaldehyde dehydrogenase. *Biochim Biophys Acta* **1480**, 329-341, doi:S0167-4838(00)00086-8 [pii] (2000).

Sebela, M. *et al*. Light microscopic localisation of aminoaldehyde dehydrogenase activity in plant tissues using nitroblue tetrazolium-based staining method. *Plant Phsiol. Biochem.* **39**, 831-839, (2001)

SelegArd, L. *et al.* Biotinylation of ZnO nanoparticles and thin films: a two-step surface functionalization study. *ACS Appl Mater Interfaces* **2**, 2128-2135, doi:10.1021/am100374z (2010).

Sharma, M. & Sun, Z. 5'TG3' interacting factor interacts with Sin3A and represses AR-mediated transcription. *Mol Endocrinol* **15**, 1918-1928 (2001).

Sharp, J. S., Nelson, S., Brown, D. & Tomer, K. B. Structural characterization of the E2 glycoprotein from Sindbis by lysine biotinylation and LC-MS/MS. *Virology* **348**, 216-223, doi:S0042-6822(05)00823-8 [pii]10.1016/j.virol.2005.12.020 (2006).

Shimomura, O., Johnson, F. H. & Saiga, Y. Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea. *J Cell Comp Physiol* **59**, 223-239 (1962).

Soppa, J. Protein acetylation in archaea, bacteria, and eukaryotes. *Archaea* **2010**, doi:820681 [pii] 10.1155/2010/820681 (2010).

Steinkamp, M. P. *et al.* Treatment-dependent androgen receptor mutations in prostate cancer exploit multiple mechanisms to evade therapy. *Cancer Res* **69**, 4434-4442, doi:0008-5472.CAN-08-3605 [pii] 10.1158/0008-5472.CAN-08-3605 (2009).

Steketee, K. *et al.* Amino acids 3-13 and amino acids in and flanking the 23FxxLF27 motif modulate the interaction between the N-terminal and ligand-binding domain of the androgen receptor. *Eur J Biochem* **269**, 5780-5791, doi:3276 [pii] (2002).

Strohalm, M., Hassman, M., Kosata, B. & Kodicek, M. mMass data miner: an open source alternative for mass spectrometric data analysis. *Rapid Commun Mass Spectrom* **22**, 905-908, doi:10.1002/rcm.3444 (2008).

Strohalm, M., Kavan, D., Novak, P., Volny, M. & Havlicek, V. mMass 3: a cross-platform software environment for precise analysis of mass spectrometric data. *Anal Chem* **82**, 4648-4651, doi:10.1021/ac100818g (2010).

Swack, J. A., Zander, G. L. & Utter, M. F. Use of avidin-sepharose to isolate and idenify biotin polypeptides from crude extracts. *Anal Biochem* **87**, 114-126 (1978).

Teyssier, C., Ou, C. Y., Khetchoumian, K., Losson, R. & Stallcup, M. R. Transcriptional intermediary factor 1alpha mediates physical interaction and functional synergy between the coactivator-associated arginine methyltransferase 1 and glucocorticoid receptor-interacting protein 1 nuclear receptor coactivators. *Mol Endocrinol* **20**, 1276-1286, doi:me.2005-0393 [pii] 10.1210/me.2005-0393 (2006).

Tilley, W. D., Buchanan, G., Hickey, T. E. & Bentel, J. M. Mutations in the androgen receptor gene are associated with progression of human prostate cancer to androgen independence. *Clin Cancer Res* **2**, 277-285 (1996).

Tomaru, T. *et al.* Isolation and characterization of a transcriptional cofactor and its novel isoform that bind the deoxyribonucleic acid-binding domain of peroxisome proliferator-activated receptor-gamma. *Endocrinology* **147**, 377-388, doi:en.2005-0450 [pii] 10.1210/en.2005-0450 (2006).

Tsai, T. C., Lee, Y. L., Hsiao, W. C., Tsao, Y. P. & Chen, S. L. NRIP, a novel nuclear receptor interaction protein, enhances the transcriptional activity of nuclear receptors. *J Biol Chem* **280**, 20000-20009, doi:M412169200 [pii] 10.1074/jbc.M412169200 (2005).

Tseng, C. L. *et al.* Development of gelatin nanoparticles with biotinylated EGF conjugation for lung cancer targeting. *Biomaterials* **28**, 3996-4005, doi:S0142-9612(07)00388-2 [pii] 10.1016/j.biomaterials.2007.05.006 (2007).

Tseng, C. L., Su, W. Y., Yen, K. C., Yang, K. C. & Lin, F. H. The use of biotinylated-EGF-modified gelatin nanoparticle carrier to enhance cisplatin accumulation in cancerous lungs via inhalation. *Biomaterials* **30**, 3476-3485, doi:S0142-9612(09)00275-0 [pii]10.1016/j.biomaterials.2009.03.010 (2009).

Tsien, R. Y. The green fluorescent protein. *Annu Rev Biochem* **67**, 509-544, doi:10.1146/annurev.biochem.67.1.509 (1998).

Turecek, F. Mass spectrometry in coupling with affinity capture-release and isotope-coded affinity tags for quantitative protein analysis. *J Mass Spectrom* **37**, 1-14, doi:10.1002/jms.275 [pii] 10.1002/jms.275 (2002).

Tylichova, M. *et al.* Structural and functional characterization of plant aminoaldehyde dehydrogenase from Pisum sativum with a broad specificity for natural and synthetic aminoaldehydes. *J Mol Biol* **396**, 870-882, doi:S0022-2836(09)01520-4 [pii] 10.1016/j.jmb.2009.12.015 (2010).

Volkmann, G. & Liu, X. Q. Protein C-terminal labeling and biotinylation using synthetic peptide and split-intein. *PLoS One* **4**, e8381, doi:10.1371/journal.pone.0008381 (2009).

Wang, L., Hsu, C. L. & Chang, C. Androgen receptor corepressors: an overview. *Prostate* **63**, 117-130, doi:10.1002/pros.20170 (2005).

Wang, Q., Sharma, D., Ren, Y. & Fondell, J. D. A coregulatory role for the TRAP-mediator complex in androgen receptor-mediated gene expression. *J Biol Chem* **277**, 42852-42858, doi:10.1074/jbc.M206061200 M206061200 [pii] (2002).

Watanabe, T., Muranaka, N., Iijima, I. & Hohsaka, T. Position-specific incorporation of biotinylated non-natural amino acids into a protein in a cell-free translation system. *Biochem Biophys Res Commun* **361**, 794-799, doi:S0006-291X(07)01593-8 [pii]10.1016/j.bbrc.2007.07.099 (2007).

Waugh, D. S. Making the most of affinity tags. *Trends Biotechnol* **23**, 316-320, doi:S0167-7799(05)00084-3 [pii]10.1016/j.tibtech.2005.03.012 (2005).

Werner, G. & Mohammad, N. [Radioactive tagging of tropan-alkaloids. V. Synthetic incorporation of tritium in (-)-cocaine, (+)-pseudococaine and (-)-scopolamine]. *Justus Liebigs Ann Chem* **694**, 157-161 (1966).

Werner, R. *et al.* The A645D mutation in the hinge region of the human androgen receptor (AR) gene modulates AR activity, depending on the context of the polymorphic glutamine and glycine repeats. *J Clin Endocrinol Metab* **91**, 3515-3520, doi:jc.2006-0372 [pii] 10.1210/jc.2006-0372 (2006).

Wisniewski, A. B. & Mazur, T. 46,XY DSD with Female or Ambiguous External Genitalia at Birth due to Androgen Insensitivity Syndrome, 5alpha-Reductase-2 Deficiency, or 17beta-Hydroxysteroid Dehydrogenase Deficiency: A Review of Quality of Life Outcomes. *Int J Pediatr Endocrinol* **2009**, 567430, doi:10.1155/2009/567430 (2009).

Wolf, B., Grier, R. E., Allen, R. J., Goodman, S. I. & Kien, C. L. Biotinidase deficiency: the enzymatic defect in late-onset multiple carboxylase deficiency. *Clin Chim Acta* **131**, 273-281 (1983).

Yang, C. K., Kim, J. H., Li, H. & Stallcup, M. R. Differential use of functional domains by coiled-coil coactivator in its synergistic coactivator function with beta-catenin or GRIP1. *J Biol Chem* **281**, 3389-3397, doi:M510403200 [pii] 10.1074/jbc.M510403200 (2006).

Yeh, S. & Chang, C. Cloning and characterization of a specific coactivator, ARA70, for the androgen receptor in human prostate cells. *Proc Natl Acad Sci U S A* **93**, 5517-5521 (1996).

Yu, X., Li, P., Roeder, R. G. & Wang, Z. Inhibition of androgen receptor-mediated transcription by amino-terminal enhancer of split. *Mol Cell Biol* **21**, 4614-4625, doi:10.1128/MCB.21.14.4614-4625.2001 (2001).

Yu, Z. *et al.* Androgen-dependent pathology demonstrates myopathic contribution to the Kennedy disease phenotype in a mouse knock-in model. *J Clin Invest* **116**, 2663-2672, doi:10.1172/JCI28773 (2006).

Zempleni, J., Hassan, Y. I. & Wijeratne, S. S. Biotin and biotinidase deficiency. *Expert Rev Endocrinol Metab* **3**, 715-724, doi:10.1586/17446651.3.6.715 (2008).

Zhou, D. *et al.* PNRC: a proline-rich nuclear receptor coregulatory protein that modulates transcriptional activation of multiple nuclear receptors including orphan receptors SF1 (steroidogenic factor 1) and ERRalpha1 (estrogen related receptor alpha-1). *Mol Endocrinol* **14**, 986-998 (2000).

Zhu, P. *et al.* Macrophage/cancer cell interactions mediate hormone resistance by a nuclear receptor derepression pathway. *Cell* **124**, 615-629, doi:S0092-8674(06)00006-7 [pii] 10.1016/j.cell.2005.12.032 (2006).