Recombinant Expression of Functional Mouse AhR LBD in *E coli*

By

Wail A. M. Elhawari, BSc, MSc

Thesis submitted to the University of Nottingham for the Degree of Master of Philosophy (M.Phil)

March 2011

Abstract

2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) is a toxic halogenated aromatic hydrocarbon, which is a potent toxin to different species such as fish, birds, and mammalians. TCDD exposure causes induction of cytochrome P4501A1 (CYP1A1), and that is controlled through the Aryl hydrocarbon receptor (AhR). AhR is a ligand-dependent cytosolic protein. It is a protein containing basic helix-loop-helix and Per-ARNT-Sim (PAS) domains. It is present in the cytosol as a complex containing two molecules of heat shock protein (hsp90) and AhR interacting protein (AIP).

When TCDD binds to AhR, the ligand–AHR complex translocates to the nucleus and dissociates from chaperone proteins and binds to another protein called ARNT (AhR nuclear translocator). The AhR-ARNT heterodimer then activates transcription. The three-dimensional structure of AhR is unknown, as is its interaction with ligands.

The aim of this work is trying to produce a high level expression and purification of an AhR-GFP (AhR–green florescent protein) fusion protein functional and sufficient for fluorescence analytical techniques to study the ligand binding and the change in AhR conformation.

Escherichia coli strain BL21 was used to express the recombinant protein since it is easy, cheap and yields a high level of protein. However, the attempts were not successful to express GST-EGFP-AhR (GGA) and GST-EGFP-AhR-EGFP (GGAG) recombinants in a functional conformation. By lowering the cultivation temperature, the proteins could be expressed and fold correctly in E. coli. However, E. coli does not contain the chaperon proteins essential for ligand binding. Therefore, the bacterially expressed protein was refolded in human reticulocyte lysate.

This study established that GFP AhR (LBD) recombinant protein can be obtained in the ligand binding conformation through the expression of the protein in E coli followed by refolding in reticulocyte lysate.

Acknowledgement

I would like to thank my supervisors Prof. John Armour and Prof Paul O'Shea for their constant help and support. I also would like to show my gratitude to Dr David Bell my former supervisor for showing me the way at the beginning of my journey.

I would like to thank Mr Declan Brady who was always there to give help and advice in the laboratory, as I would like to thank my wife Abeer for supporting me all the way in the office as well as at home.

I also want to show my appreciation to all my friends and family in the UK as well as back home in Libya.

Dedication

To my Mom and Dad, to my daughter Alaa and sons Monder and Mohammed

To my lovely wife Abeer

This thesis is dedicated to the martyrs for the cause of Libya’s freedom

Declaration

I declare;

- This work was accomplished within the period of my M.Phil studies at the University of Nottingham.
- This work is my unique effort unless otherwise stated (information from other resources has been fully recognized).
- No part of this work has been submitted for evaluation leading to a degree.

Wail Elhawari

March 2011

Contents

List of table

List of Figurs

Abbreviations

AhR	Aryl Hydrocarbon Receptor
AIP	AhR interaction protein
APS	Ammonium Persulphate
ARNT	Aryl Hydrocarbon Receptor Nuclear Translocator
bHLH	Basic helix-loop-helix
BSA	Bovine Serum Albumin
CYP	Cytochrome P450
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
DRE	Dioxin Response Elements
EDTA	Ethylene diamine tetra acetic acid
EGFP	Enhanced green fluorescent protein
GST	Glutathione S-transferase
HAH	Halogenated aromatic hydrocarbons
His Tag	Histidine Tag
Hsp90	Heat shock protein 90
IPTG	Isopropyl-?-D-thiogalactopyranoside
Kb	Kilobase
kDa	kilo Dalton
LB	Luria-Bertani broth
LBD	Ligand binding domain
OD	Optic Density
P450	Cytochrome P450
PAGE	Polyacrylamide Gel Electrophoresis
PAH	Polycyclic aromatic hydrocarbon
P23	An Hsp90-associated protein
PCBs	Polychlorinated biophenyls

RNA	Ribonucleic acid
SDS	Sodium Dodecyl Sulphate
TCAOB	3,3',4,4'-tetrachloroazoxybenzene
TCDD	2,3,7,8-tetrachlorodibenzo-p-dioxin
TEMED	N,N,N',N'-tetramethylethylenediamine
Tris	Tris(hydroxymethyl)methylamine
UHP	Ultra High Purity
XAP2	The hepatitis virus X protein-associated protein 2
XRE	Xenobiotic Response Elements

Introduction

Humans and animals are daily exposed to a large number of chemical compounds and environmental pollutants present in the environment, such as drugs, agrochemicals and other environmental contaminants. All of these chemicals are detoxified and eventually should be excreted from the body. Many of these chemicals require modification to be in a soluble form to enable excretion. The Cytochrome P450 family are the major enzymes involved in these metabolic mechanisms (Lash, 1994)

Aromatic hydrocarbons are environmental contaminants, which cause toxic effects or diseases in both humans and other animals. There are many examples of the effects of these hydrocarbon compounds in humans and animals such as immunosuppression, cleft palate, chloracne, cancer and heart disease (Birnbaum, 1994).

Many of these chemicals have a similar toxic response, and act by a common mechanism through cytochrome P450, such as CYP1A1 to enhance detoxification which is controlled by a cytosolic protein receptor called Aryl hydrocarbon receptor (AhR) (Knutson and Poland, 1982; Whitlock, et al. 1996; Hahn, 1998).

1 Xenobiotic metabolism

A series of metabolic reactions modifies the chemical structure of xenobiotics, to enable excretion of these compounds. These can be divided into three phases. Phase I reactions are catalysed by enzymes such as cytochrome P450. Phase II reactions are

catalysed by transferase enzymes such as glutathione S-transferases. Finally, in phase III, the conjugated xenobiotics resulting from Phase I and Phase II need to be further processed, before being recognised and excreted from cells, for example by transforming glutathione conjugates to acetylcysteine.

1.1.2 Cytochromes P450

Cytochrome P450s form a group of enzymes involved in the metabolism of many clinical and toxic chemicals. They are also involved in many metabolic reactions of various endogenous products such as steroids, and fat-soluble vitamins (Keeney and Waterman, 1993; Higashi et al.,1991; White et al., 1994).

This group of enzymes is responsible for catalysing a number of detoxification reactions called Phase I xenobiotic reaction. These include epoxidation, N-dealkylation, O-dealkylation, S-oxidation and hydroxylation.

The main goal of the detoxification process is to convert the inactive compound to a form to be readily excreted out of the body. However, many of these reactions can convert inactive molecules to highly polar forms which can be carcinogenic. One of the P450 members is CYP1A1, also known as AHH (aryl hydrocarbon hydroxylase); this enzyme is involved in the metabolic activation of aromatic hydrocarbons. For example, benzo[a]pyrene is a hydrocarbon compound with a link to causing lung cancer (Denissenko et al 1996). In the metabolic reaction of benzo[a]pyrene, the benzo[a]pyrene oxidation is catalysed by CYP1A1 to form BaP-7,8-epoxide; hydrolysis of BaP-7,8-epoxide by epoxide hydrolase (EH) forms BaP-7,8-dihydrodiol which is finally oxidized by CYP1A1 to form BaP-7,8-dihydrodiol-9,10-epoxide, which is the ultimate carcinogen and can covalently bind to DNA (Figure 1.1) (Beresford, 1993). The expression of the CYP1A1 is regulated by the aryl hydrocarbon receptor (Poland and Knutson, 1982).

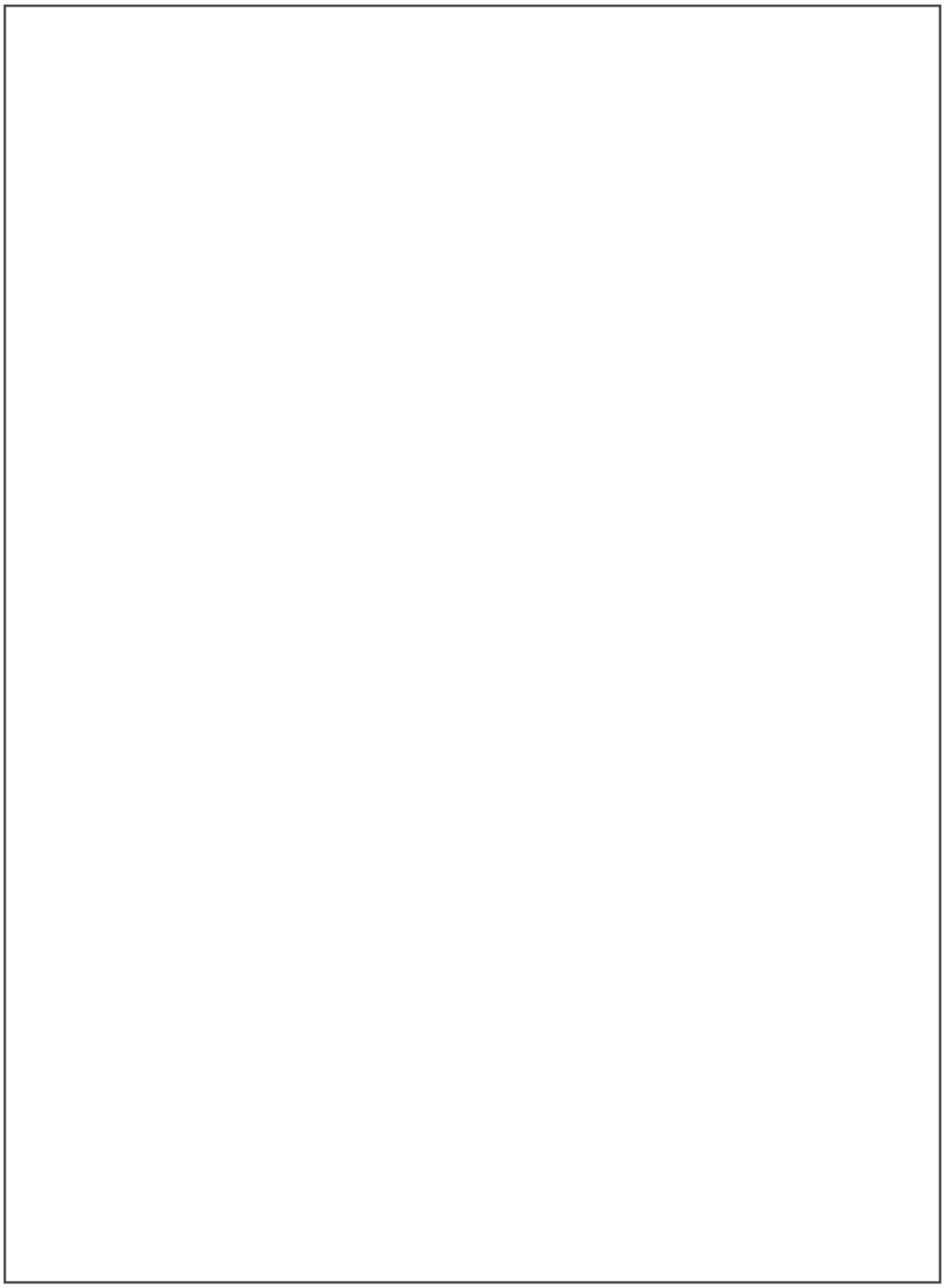


Figure 2.1 Formation of an Ultimate Carcinogen from benzo[a]pyrene

1.2 Aryl hydrocarbon Receptor (AhR)

AhR is a ligand-dependent cytosolic transcription factor that mediates the induction of cytochrome P450 family members. AhR is classified as a member of the basic helix-loop-helix/Per-ARNT-Sim (bHLH/PAS) family (Fukunaga et al., 1995; Gu et al., 2000; Hankinson, 1995). It is present in the cytosol as a complex and interacts with co-chaperone proteins, including AhR interacting protein (AIP) (Gu et al., 2000), which facilitate signalling by the AhR and play an important role in regulating nuclear translocation (Ma and Whitlock, 1997; Whitlock et al., 1996). Two molecules of heat shock protein (hsp90) and the co-chaperone p23 were found to be associated with the non activated form of the Aryl hydrocarbon receptor (Kazlauskas et al., 1999; Ma and Whitlock, 1997).

1.2.1 Identification of the Ah receptor

Several studies defined the role of AhR in the toxicity and the mechanism of action of HAHs (Halogenated Aromatic Hydrocarbons) and PAH (Polycyclic Aromatic Hydrocarbons) including the toxicity of TCDD and related HAHs (Safe 1990; Poland and Knutson 1982).

The responses to dioxin, related halogenated hydrocarbons compounds and polycyclic aromatic hydrocarbons are not the result of a covalent chemical reaction of the toxin itself, and the Ah receptor mediates the toxic effects of TCDD (Poland and Knutson, 1982). The AhR has been reported to exist in several human and animal tissues including placenta, liver, lung, thymus, and fetal kidney. [^{3}H] TCDD ligand-binding assay experiments confirmed the existence of a cytosolic protein that binds specifically to radio-labelled TCDD and mediates the induction of CYP1A1 (Allen et al., 1975; Poland et al., 1986). In 1996 Gonzalez and Fernandez-Salguero reported that AhR-Deficient Mice are resistant to TCDD toxicity. These studies provided the evidence of the existence of the AhR (Fernandez-Salguero et al., 1996).

1.2.2 The AhR locus

There is genetic variation in the inducibility of CYP1A1 the response to 3-methylcholanthrene (MC) varies significantly among different mouse strains, such that in some strains there is no induction of CYP1A1 (nonresponsive mice) (Poland and Knutson 1982). These genetic differences in the inducibility of CYP1A1 are controlled by the Ah locus and two alleles were defined, Ah^b in responsive strains and Ah^d in nonresponsive strains (Nebert and Jensen, 1979). TCDD and HAH related compounds were capable of inducing CYP1A1 in both responsive and nonresponsive mice. It was suggested that a mutation in nonresponsive mice results in a defect in the structural gene that encodes for a receptor for TCDD (Poland and Glover, 1979).

The AhR locus encodes the structural gene for the AhR, and the human AhR gene has been localized to human chromosome 7 (Bertrand et al., 1995). In mouse, a region homologous to the human AhR gene has been localized on chromosome 12 (Poland et al., 1987).

1.2.3 AhR ligands

AhR is also known as the dioxin receptor because of the ability to activate the AhR in presence of very small amounts of dioxin. However, there are many chemicals which are considered as AhR ligands. AhR ligands could be exogenous or could be endogenous.

1.2.3.1 Exogenous ligands

These compounds are produced through incomplete combustion of various carbon sources such as diesel exhaust, cigarette smoke, and fried foods (Schmidt et al., 1996; Sorensen et al., 1996). These compounds can be divided into halogenated aromatic hydrocarbons (HAHs) and polyaromatic hydrocarbons (PAHs).

1.2.3.1.1 Halogenated Aromatic Hydrocarbons (HAHs)

HAH compounds are extremely dangerous and can have a significant impact on health because of their ubiquitous distribution, fat solubility, and resistance to degradation. HAHs have a relatively high binding affinity for the AhR (in the pM to nM range). However, the majority of these chemicals are relatively weaker inducers compared with TCDD. These include the polychlorinated dibenzo-p-dioxins (PCDDs), dibenzofurans and biphenyls and related chemicals (Safe, 1990). There are 75 possible congeners of chlorinated dioxin and their properties depending on the number and position of chlorine atoms (Poland and Knutson, 1982). TCDD is one of the important dioxin congeners which is extremely toxic to different species such as fish, birds, and mammals.

1.2.3.1.2 Polycyclic Aromatic hydrocarbons (PAHs)

These chemicals are known to be less potent than HAHs and have lower binding affinity

to the AhR compared to the HAHs. PAHs are among the most widespread organic pollutants, formed by incomplete combustion of carbon-containing fuels such as wood, coal, diesel, fat, tobacco, or incense. These include benzo(a)pyrene, 3-methylcholanthrene (3MC) and aromatic amines and other related chemicals (Poland et al., 1976).

1.2.3.2 Endogenous ligands

Activation of the AhR signalling pathway was observed in the absence of exogenous ligands in several studies (Chang and Puga 1998). Using AhR-null mice, the role of the AhR in normal development was investigated. Various physiological changes and developmental abnormalities were detected in these mice, for example early hepatovascular defects lead to a number of physiological and morphological changes in the adult mouse liver (Schmidt and Bradfield 1996).

A number of natural compounds have been considered as AhR ligands including derivatives of tryptophan such as indigo and indirubin (Adachi et al., 2001), tetrapyrroles such as bilirubin (Sinal and Bend, 1997), the arachidonic acid metabolites lipoxin A4 and prostaglandin G (Seidel et al., 2001), and a modified low-density lipoprotein (McMillan and Bradfield, 2007). The role of these chemicals remains unclear but these chemicals are present in many species and tissues.

1.2.4 Toxicity of AhR ligands

Polychlorinated dioxins are normally stored in fatty tissues and it has been estimated that the half-life of the polychlorinated dioxins in a human adult's body is approximately 7 years (Geyer et al. 2002). Their toxicity depends on a number of factors, such as dose, age, sex, animal strain and species (Poland and Knutson 1982). Toxicity of TCDD differs among species. For example, the acute oral LD_{50} dose for the guinea pig is 1 ?g/kg whereas for hamsters it is 5000 ?g/kg. In humans, a high- level exposure to TCDD leads to chloracne which is the common effect in human. This effect was also observed in some laboratory animals such as rhesus monkeys and rabbits after TCDD exposure. However, it was not reported in other laboratory animals such as rats, hamsters and mice (Poland and Knutson 1982).

It has been confirmed that TCDD is carcinogenic in rats and it is a very potent tumour promoter. Pitot and his colleagues in 1980 studied the possible effect of TCDD on rats as a hepatocarcinogenic promoter; the animals were exposed firstly to a single dose of 10 mg/kg diethylnitrosamine followed by doses of TCDD between (0.14 to 1.4 µg/kg) each 14 days for 7 months. Nearly 70% of the treated rats were noticed to have hepatocellular carcinomas. However, TCDD is not considered as genotoxic since it does not bind covalently to protein or nucleic acid (Greenlee and Poland 1979).

PAHs were shown to be AhR ligands and exposure to PAH may contribute to induction of mouse mammary carcinoma through inducing cytochrome p450 family members which oxidize PAH to mutagenic and carcinogenic intermediates (Trombino et al. 2000).

Creation of AhR deficient mice confirmed that toxicity of TCDD depends on AhR. These mice were resistant and had no response to 10 times the dose of TCDD which was highly toxic in normal mice. However, these mice exhibited a number of abnormal phenotypes including reduced liver weight and portal fibrosis associated with accelerated rates of apoptosis (Schmidt et al. 1996).

1.2.5 Structure of the AhR

AhR is classified as a member of the basic helix-loop-helix/Per-ARNT-Sim (bHLH/PAS) family. AhR contains several domains essential for AhR function and classification. The bHLH region located in the N terminus of the protein is a common region in a number of bHLH super family members. bHLH is a region involved in AhR-ARNT complex dimerization, binding to heat shock protein and DNA binding (Gu et al., 2000).

bHLH region has Nuclear Localisation Sequence (NLS) and Nuclear Export Sequence (NES) domains located in this region (Berg and Pongratz 2001; Ikuta al., 1998). This domain is also necessary for heterodimerization to ARNT and binding of the AhR-ARNT complex to the dioxin response elements (DRE) and activation of target genes (Landers and Bunce, 1991; Fukunaga et al., 1995).

The PAS regions contain two PAS domain (Figure 1.2); the PAS-A and PAS-B domains are known to enhance interactions with other PAS domain containing proteins, such as interaction between AhR and ARNT. The PAS-B domain has an important function, namely to bind to ligand (Denison et al. 2002).

A Q-rich domain is located in the C-terminal region of the AhR protein and is responsible for interactions with the general transcriptional activation machinery (Fukunaga et al., 1995). AhR proteins have been characterized in several mammals and there are considerable variations in the structure among different species, and among strains within species (Poland et al., 1976).

Figure 1.2 cartoon representation of the domain structure of the AhR

1.2.5.1 Ligand binding domain (LBD)

The function of the Ligand Binding Domain region is critical for binding both the ligand and hsp90. It is believed that LBD region is responsible for the correct folding of the AhR and the maintenance of the three dimensional structure in a high ligand affinity conformation.

In a study to identify the functional domains of the AhR , deletion of the PAS-A region (amino acids 121-182) was shown to reduce ligand binding by a third, whereas deletion of the PAS-B region (amino acids 259-374) entirely eliminated binding to the ligand, the same response as complete deletion of PAS region. Accordingly, the LBD region was localised between amino acids 230 and 397 which includes PAS-B region (Fukunaga et al., 1995). The LBD has similar affinity for TCDD binding as the full-length AhR (Coumailleau et al., 1995). Therefore, LBD could be a useful to use LBD as a tool in studying the function of the AhR.

1.2.5.2 AhR Signaling Pathway

Ligand-free AhR is present in the cytoplasm as an inactive protein complex containing the AhR (figure 1.3), two molecules of Hsp90, and a single molecule of AhR interacting protein (AIP) (Gu et al. 2000). Following the ligand binding AhR conformational changes take place leading to unveiling of the nuclear localization sequence, and then the ligand–AhR complex translocates to the nucleus (Pollenz et al. 1994). Once the ligand–AhR complex enters the nucleus, the AhR dissociates from the complex and dimerises with ARNT (Dzeletovic et al. 1997).

This dimerisation alters the AhR into a high affinity DNA binding structure and enables AhR to bind DNA Xenobiotic Response Elements XRE (Dioxin Response Elements,DRE) of target promoters. (Dzeletovic et al. 1997). Binding of the Ligand-AhR-ARNT complex to this region enhances the transcription of several genes, including CYP1A1, known as the AhR gene battery (Durrin and Whitlock 1989).

However, the AhR signaling pathway is still not fully understood because of the lack of information about the molecular mechanism of how the AhR signalling pathway is activated and structural information about AhR LBD and the interaction between AhR and chaperone proteins such as AIP and Hsp90.

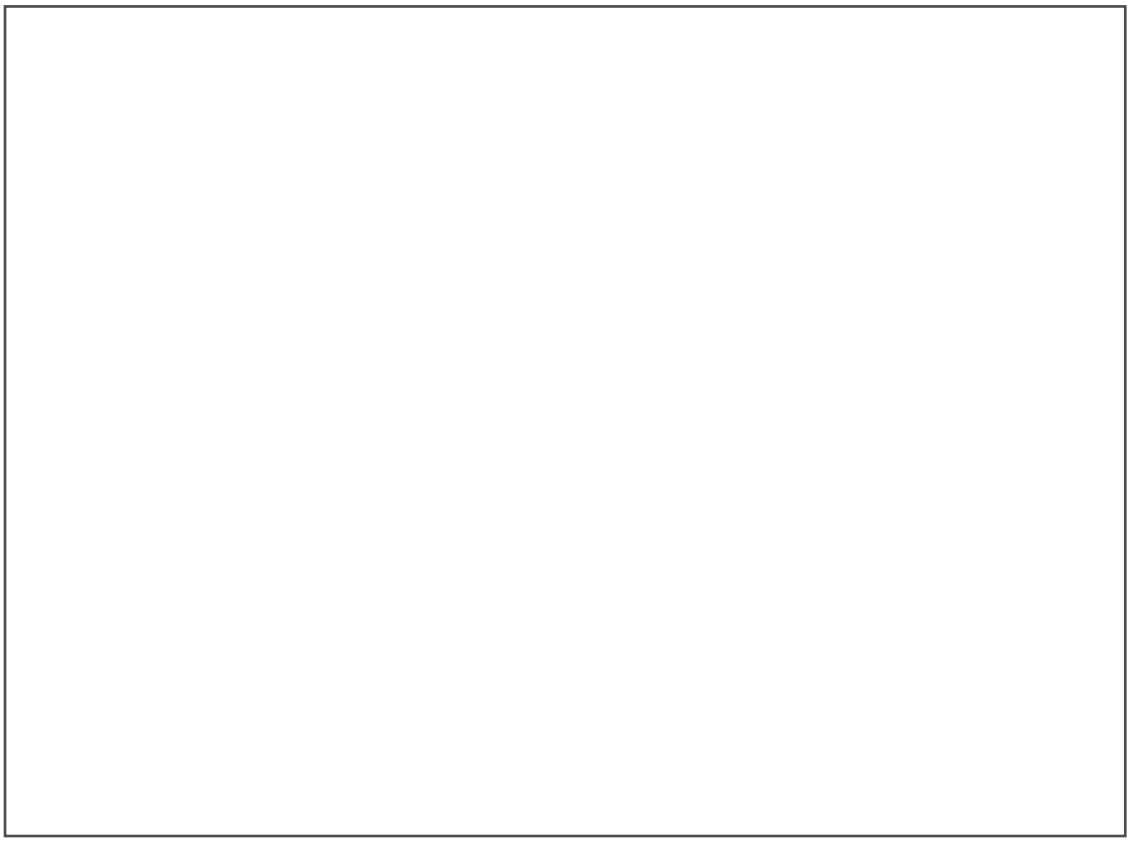


Figure 1.3 Cartoon explaining the AhR Signaling Pathway (Gu, Hogenesch et al. 2000). (A) Unbound AhR protein complex containing two molecule of Hsp90, and a single molecule of AIP. (B) Ligand binding to the AhR complex. (C) AhR dissociates from the complex and dimerises with ARNT (D) AhR/ARNT complex binds to DNA in DRE region

1.3 Chaperones of Non-ligand bound AhR

The Cytosolic AhR interacts with a number of chaperone proteins, including AIP, which facilitates signalling by the AhR and plays an important role in regulating nuclear translocation (Ma and Whitlock 1997) and two molecules of Hsp90 and the co-chaperone p23. Hsp90 interacts with p23 forming a dimer, this dimer has several crucial roles, including protection of the AhR from proteolysis, conservation of the AhR pocket conformation to enable ligand binding and to prevent untimely ARNT binding (Kazlauskas et al., 1999).

1.3.1 Heat shock protein 90 (Hsp90)

Hsp90 is a chaperone protein abundant in the cytoplasm, and represents about 1-2% of the total expressed proteins in the cell. Hsp90 was found in eukaryotic cells and it is necessary for cell viability (Pearl and Prodromou, 2001; Richter and Buchner, 2001). Hsp90 acts as chaperone protein and is essential for protein folding and function (Pratt et al., 2004).

1.3.1.1 The role of hsp90 in AhR signalling

It has been established that Hsp90 is a critical protein for the AhR ligand signaling pathway, and is consistent with the normal function of the Hsp90 as a chaperone protein (Pongratz et al., 1994). Separation of the Hsp90 from the AhR leads to incapability of the AhR to bind the ligand; when Hsp90 was reduced by nearly 95%, AhR failed to response to the ligand (Carver and Bradfield, 1997; Whitelaw et al., 1995; Perdew, 1988).

There are two regions in the AhR where Hsp90 interacts with AhR; these are the LBD region (the region between amino acids 230 to 421) and the bHLH region (Whitelaw et al., 1994; Antonsson et al., 1995).

Hsp90 has a number of important functions in term of AhR signalling and binding to the ligand. The first one is to maintain the AhR in a high-affinity conformation state and properly fold the receptor in a conformation suitable for ligand binding (Swanson et al., 1993; Pollenz et al., 1994). Exposure of the Hsp90 to a high temperature and high salt dissociated Hsp90 from the AhR and abolished binding of ligand (Grenert et al., 1997).

The Second function of Hsp90 is to prevent the AhR from binding to ARNT (McGuire et al., 1994). The other function for the Hsp90 is to unmask the nuclear localisation signal (NLS) which exists in the N-terminus of the AhR targeting the AhR to the nucleus. This step is important to translocate the AhR to the nucleus (Ikuta et al., 1998).

1.3.2 AhR interacting Protein

AIP is also called X-Associated Protein 2 (XAP2) (Kuzhandaivelu et al., 1996) or AhR associated protein 9 (ARA9) (Carver and Bradfield, 1997). This protein belongs to the Immunophilin family, which has peptidylprolyl isomerase (PPIase) activity.

In the N-terminal region of this protein, there is a peptidylprolyl isomerase domain and in the C-terminal region, there are three tetratricopeptide (TPR) repeat motifs; these are 34-amino acid residue repeats which provide the ?-helical structure for this unit and form the protein/ protein interaction domain (Pratt et al., 2004; Pratt et al., 2004) .

The C-terminal region is necessary for the interaction with both the AhR and other chaperone proteins (Bell and Poland 2000). A study performed by Bell and Poland in 2000, mutating the conserved basic residue in TPR of AhR interacting protein, proposed that this region was essential for binding both AhR and Hsp90 (Bell and Poland, 2000). The LBD region in the AhR is essential for binding to AIP. In the Hsp90, the C terminus region is essential to bind AIP (Bell and Poland, 2000).

1.3.2.1 The role of AIP in AhR signalling

AIP is important in enhancing the AhR transcriptional activity through increasing the stability of the AhR and protecting the unbound AhR from degradation (LaPres et al., 2000), (Kazlauskas et al., 2000), (Ma and Baldwin, 2000), (Lees and Whitelaw, 2002). In addition, AIP localizes AhR in the cytoplasm. Recent studies showed that over expression of AIP redistributes the AhR to the cytoplasm (LaPres et al., 2000; Petrulis et al., 2000; Petrulis and Perdew, 2002). AIP is thought to connect the unbound AhR to the actin filaments to stabilize the AhR in the cytoplasm (Petrulis et al., 2003). Moreover; AIP reduces the binding of the nuclear localisation signal (NLS) to Importin and prevents AhR from entering the nucleus (Kazlauskas et al., 2001; Petrulis and Perdew, 2002).

1.3.3 P23 or Hsp90-associated protein

The p23 is a ubiquitous, highly conserved protein, which interacts with Hsp90 (Pratt and Toft, 1997), and is also called Hsp90 binding protein (Johnson et al., 2000). The p23 binds to the N-terminal region of Hsp90 to stabilize the interaction between the Hsp90 and the receptor which enhances the frequency of receptor in the high-affinity conformation (Pratt and Toft, 1997; Cox and Miller, 2002).

p23 is necessary for dimerizing the two units of Hsp90 and formation of the cytosolic complex; also it has an important role in dimerizing AhR to ARNT. p23 binds to the ATP binding region in Hsp90 and acts as substrate releasing factor (Kazlauskas et al., 2001); releasing Hsp90 form the AhR depends on the Hsp90 ATPase activity which is stimulated by p23 (Kazlauskas et al., 1999; Young and Hartl, 2000). Also, it interacts with other proteins to maintain them in a folding-form stage (Freeman et al., 1996;

Freeman and Morimoto, 1996). However, a recent study has demonstrated that p23 is not required *in vivo* as co-chaperone protein in AhR binding or targeting (Murray et al., 2009).

1.4 AhR activation and protein folding

Non-functional proteins result from incorrect folding (misfolding) (Shortle, 1997). AhR responses to different ligands depend on the primary structure of the AhR. This fact was demonstrated by site-directed mutations within the LBD of mouse AhR in specific residues; one mutation changed the response to antagonist (Henry and Gasiewicz, 2008). A study conducted by Backlund and Ingelman-Sundberg (2004) on mutated residues in the LBD of Gal4-AhR, revealed that high- and low-affinity ligands interact with different residues in the AhR ligand-binding pocket. There are specific structural and chemical requirements for ligand binding. A previous study has attempted to investigate the structure of the AhR; (Pandini et al., 2007) developed a model of the mouse AhR PAS B domain by comparative modelling techniques.

The PAS domain structures of the functionally related hypoxia-inducible factor 2R (HIF-2R) and AhR nuclear translocator (ARNT) proteins, which show the highest degree of sequence similarity with AhR, were chosen as a model to develop a two-template model of the PAS B domain. Moreover, these models were tested by using site-directed mutagenesis of amino acids in key positions within these models then a model was proposed and it suggesting some key features required for ligand binding (Pandini et al., 2007).

A similar study was conducted, based on the similarity of PAS domain structure of hypoxia-inducible factor and AhR LBD, to generate a homology model of the mouse AhR-LBD based on its similarity (Bisson et al., 2009). However, these attempts were based on virtual ligand screening and did not demonstrate the real structure and the exact features of the LBD AhR binding cavity. Also, these studies proposed that AhR LBD receptor has the highest affinity binding conformation, and it may be that the LBD AhR has lower affinity than the ideal proposed.

The structure of AhR has not been determined. No X-ray or NMR- study has demonstrated the structure of bound or unbound AhR. The difficulty is the expression of the functional AhR in sufficient amounts. Trials to express AhR in bacteria were carried out since bacterial systems are cheaper, on the other hand, AhR can be expressed with ligand-binding functionality in reticulocyte lysates, but the amount of AhR produced was very minimal - approximately 20 ?g of purified protein (Burbach et al., 1992).

1.5 AhR-GFP (AhR –green florescent protein) fusion protein

This work followed the Shaikh-Omar (2006) study, in order to obtain a high yield expression of functional AhR in bacteria. In that study three plasmids were constructed based on pET41b. The first one (GG) contained enhanced green fluorescent protein (EGFP) tagged to glutathione S-transferase (GST), the second (GGA) contained GST-tagged AhR LBD with EGFP and the third construct (GGAG) contained GST-tagged mouse AhR LBD with two EGFP (Figure 1.4). Shaikh-Omar (2006) attempted to express GGA and GGAG recombinants in BL21 D3 at 37^{o}C and his efforts to express and purify

these recombinants did not produce any success in obtaining functional protein; SDS PAGE showed no band in the soluble fraction, possibly because there was no soluble protein, or a very low concentration of the expressed protein. However, expression and the purification of GG recombinant were successful.

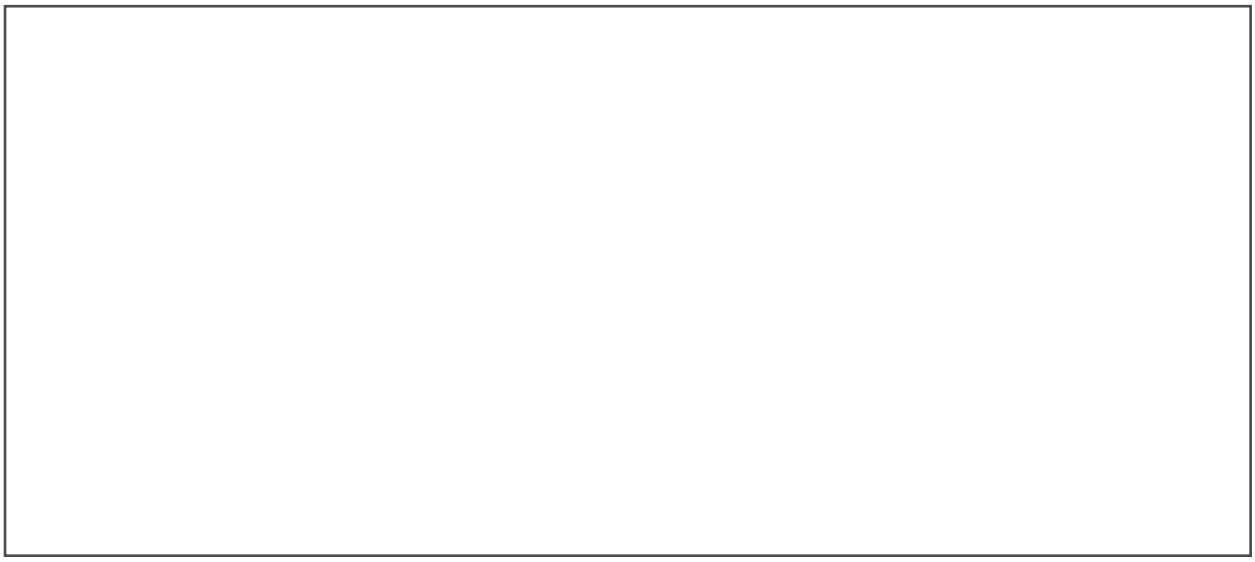


Figure 1.4 Cartoon of GG,GGA and GGAG protein constructs . These constructs contain Enhanced green fluorecent protein (EGFP) tagged to glutathione S-transferase (GST) and AhR ligand binding Domain (AhR-LBD)

1.6 The aim of this work

The aim of this work is to investigate the structure of AhR by examining an AhR-GFP (AhR–green fluorescent protein) fusion protein which has enhanced solubility compared with other bacterially expressed AhR constructs.

The aim of this work is to

- To express AhR LBD in *E. coli* (arctic express), which has chaperone proteins that could help AhR folding.
- To examine the solubility and functionality of the LBD-AhR constructs
- In the case of non-functional protein, to refold the soluble AhR in reticulocyte lysate
- To purify the fusion proteins and with any associated chaperone protein using GST affinity tags.

Materials and methods

2.1 Material

2.1.1 Chemical

All reagents and chemicals were obtained at highest standard commercially available. Bovine serum albumin (BSA), dithiothreitol (DTT), ethidium bromide, sodium dodecyl sulphate (SDS), and TEMED (tetramethylethylenediamine) were obtained from Sigma Chemical Co. Glacial acetic acid, methanol, NaCl, and NaOH, were bought from Fisher Scientific UK Limited. Glycerol was from Courtin & Warner. bromophenol blue, and ethanol were obtained from BDH. Agarose was from Boehringer Mannheim.

Brilliant blue G was from Aldrich. ammonium persulphate, coomassie brilliant blue R-250 and hydrochloric acid were obtained from ICN Flow. 30% acrylamide/bisacrylamide was from Severn Biotech Ltd. 1Kb plus DNA ladder was from Gibco BRL. 6H protein marker was from Sigma.

2.1.2 Buffers

MN stock buffer (pH 7.5 at 4ºC) consists of: 25 mM MOPS and 0.02% sodium azide.

MEN stock buffer (pH 7.5 at 4ºC) consists of: MN buffer with 1 mM EDTA.

MDENG stock buffer (pH 7.5 at 4ºC) consists of: MEN buffer with 10 % (w/v) glycerol,1 mM dithiothreitol (DTT), (DTT is freshly supplemented to the buffer before the protein preparation) and 20 mM EDTA.

1X TBS buffer consists of 20 mM Tris-HCl (pH 7.5), 500 mM NaCl (pH8.0).

1X PBS buffer consists of 137 mM NaCl, 2.7 mM KCl, 100 mM Na_2HPO_4, and 2 mM KH_2PO_4 (pH7.5).

2.1.3 Media

Lysogeny broth (LB) 1 liter of LB contain10 g tryptone, 5 g yeast extract and 10 g NaCl.

LB plates LB broth + 15 g/litre.

SOC medium (Super Optimal Broth) one litre of SOC contains Bacto-tryptone 20 g, Bacto-yeast extract 5 g NaCl 0.5 g,1M KCl 2.5 ml and ddH2O 1000 ml.

2.2 Methods

2.2.1.1 Miniprep of Plasmid DNA

Plasmids were transformed into JM109 as described in section 2.2.2 below. A single colony was picked up into 25 ml LB broth containing appropriate antibiotic and incubated overnight at 37^{o}C shaking at 240 rpm. QIA prep Spin Miniprep kit (Qiagen) was used to purify the plasmids. Following the manufacturer's instructions, 1 ml was taken from the overnight incubated culture of JM109 cells containing appropriate selective antibiotics. The cells were then transferred to an Eppendorf tube and centrifuged.

The pellet was resuspended with 250 ?l P1 and then 250?l of P2 buffer was added and the tube was inverted gently 4-6 times. Immediately, 350?l of N3 buffer was added to the mixture and inverted gently 4-6 times. The mixture was then spun at 14,000 g for 10 minutes. The supernatant was transferred to the QIA prep spin column and centrifuged at 14,000 g for 1 minute and the flow-through was discarded.

Using 750?l of PE Buffer, the column was washed and centrifuged for 1 minute. After the flow-through was discarded the QIA column was spun additionally for 1 minute to remove any remaining washing buffer (PE). The DNA was eluted by adding 50?l of sterile UHP water to the centre of the column.

2.2.1.2 Restriction digestion of the Plasmids

The plasmids were digested using *Eco*R? restriction digestion enzyme with 1 µl 10x Buffer 4, 6.5 µl H_2O, 2 µl DNA, 0.5 µl enzyme and 0.1mg/ml BSA according to the manual. The mixture was incubated at 37 ^{o}C in a water bath for one hour.

2.2.1.3 Agarose Gel Electrophoresis

50 ml of 1% agarose gel was prepared by adding 0.5 g agarose powder to 50 ml of TBE buffer, and then the mixture was microwaved at 50% power for 1-5 minutes. The solution was transferred into a water bath at 60 ^{o}C then loaded into the gel cassette and left for 20 minutes at room temperature to settle. The plasmid samples were prepared by adding 1 volume of 5x DNA loading buffer per 5 volumes of plasmid samples. The sample and DNA Marker (1Kb plus) were loaded on to the gel and run for 30 min at 100 V. The gel was then stained with ethidiim bromide stain and visualized under UV.

2.2.2 Transformation

2.2.2.1 Transformation of calcium chloride competent cells

A Single colony of the bacteria was picked up into 5 ml LB broth containing appropriate antibiotic and incubated overnight at 37 ^{o}C shaking at 240 rpm. Next day the culture was moved to 50 ml LB broth for 3 to 4 hours with shaking at 240 rpm. The culture was cooled on ice for 15 minutes then spun for 10 min at 4000g at 4 ^{o}C. The pellet was re-suspended in 50ml 0.1 M $CaCl_2$. The cell centrifugation was repeated four times. After the final centrifugation it was resuspended in 2 ml 0.1 M $CaCl_2$.

A 1 µl sample of each construct was added to 100 µl of the competent cells and incubated on ice for 30 min followed by 90 seconds at 42 ^{o}C and again on ice for two minutes. A 1 ml aliquot of SOC medium was added to the cells and incubated at 37 ^{o}C for 45 minutes. 200 µl of the cells were spread on LB plate containing appropriate antibiotic and incubated at 37 ^{o}C overnight.

2.2.2.2 Transformation of electro-competent cells

A single colony of the bacteria was picked up into 5 ml LB broth containing appropriate antibiotic and incubated overnight at 37 ^{o}C shaking at 240 rpm. The next day the culture was diluted to 50 ml LB broth and left for 3 to 4 hours with shaking at 240 rpm. The culture was cooled on ice for 15 min and then centrifuged for 10 minutes at 4000g at 4 ?C. The pellet was re-suspended in 250 ml ice cold ultra highly purified water three times.

After final centrifugation the pellet was re-suspended in twice its volume of ice cold UHP-water. A 1 µl sample of each construct was added to 70 µl of the competent cells and placed in a 1 mm gap electroporation cuvette and electroporated at a voltage of 1.5 kV.

The cells were then incubated in SOC medium at 37 ?C for 45 minutes. A 200 µl aliquot of the cells was spread on an LB plate containing appropriate antibiotic and incubated at 37 ?C overnight.

2.2.3 1 Protein Expression

A single colony of the bacteria was picked up into 25 ml LB broth containing appropriate antibiotic and incubated overnight at 37?C shaking at 240 rpm. Next day the culture was moved to 200 ml LB broth for 4 to 5 hours with shaking at 240 rpm at 30 ?C. The protein induction was applied using 1mM of IPTG. The cells were then incubated at 12 ?C overnight.

The cells were spun at 10,000 g for 10 minutes re-suspended in Tris-NaCl buffer and

they were then subjected to lysis by lysozyme, followed by freezing at -80 ^{o}C and thaw at room temperature (three times) and then sonication for 30 second at maximum power repeated 6 times.

The proteins were then centrifuged at 10,000g and the supernatant was taken and centrifuged at 200,000g for 30 minutes.

2.2.4 Bradford protein assay

The Bradford assay is a colorimetric assay for measuring total protein concentration. The principle of this assay is measuring the absorbency at 595 nm for the protein when it binds to the Coomassie brilliant blue G-250.

100 mg Coomassie brilliant blue G-250 was dissolved in 50 ml 95% ethanol, 100 ml 85% (w/v) phosphoric acid was added then diluted to 1 litre. Ranges of 5 to 100 micrograms of BSA protein were prepared in 20 µl volume and 1 ml dye reagent was added to each one and incubated for 5 minutes, and the absorbance was measured at 595 nm. A standard curve of absorbance data was plotted for BSA range of protein concentrations and used to determine the amounts of unknown protein from the curve.

2.2.4 SDS Sample preparation

A BIO-RAD Mini-PROTEAN II Cell electrophoresis kit for home made gels was used to perform this experiment. The resolving and stacking gel solutions were prepared previously and stored at 4 °C until required; all the components were added with the exception of TEMED and 10% APS (to start the polymerisation), which were added immediately prior to pouring into the gel plate (see table 1).

Components (ml)	Resolving (10 ml)			Stacking 10 ml
	8%	10%	12 %	5%
H2O	4.6	4.0	3.3	6.8
30% Acrylamide	2.7	3.2	4.0	1.7
1.5 M Tris pH 8.8	2.5	2.5	2.5	-
10% SDS	0.1	0.1	0.1	0.1
10 % APS	0.1	0.1	0.1	0.1
1.0 M Tris, pH 6.8	-	-	-	1.25
TEMED	0.006	0.004	0.004	0.01

Table 1 SDS sample preameration.

5x SDS Running Buffer: (1 L) Tris 15 g, Glycine 72 g and SDS 5 g Coomassie Blue Stain: 10% (v/v) acetic acid, 0.006% (w/v) Coomassie Blue dye and 90% ddH_2O. Destaining Solution: 10% (v/v) acetic acid, 30% (v/v) methanol and 65% ddH_2O.

2.2.4.1 Preparing an SDS gel

To prepare SDS gels, two glass plates were placed into the rack and 5 ml of the gel was loaded between the two plates. 100 µl of 0.1 % SDS was poured to cover the surface of the running gel, and the gel was left for 20 minutes for polymerisation at room temperature. The overlaying SDS solution was removed and 2 ml of stacking gel was loaded. Rapidly, the comb was placed into the space between the two glass plates.

5 X SDS sample loading buffer (40 ml) was prepared by adding ddH_2O (16 ml) into 0.5M Tris-HCl, pH 6.8 (5 ml), 50% Glycerol (8 ml), 0.02% bromophenol blue and 10% SDS (8 ml) with Dithiothreitol (DTT) 77 mg/ml added immediately before use. 5 µl 5 X of SDS loading buffer was added to 15 µl of each sample and heated at 95?C for 10 minutes, in order to denature the proteins. The BioRad kit was used for the electrophoresis, and the gel cassette was immersed in the tank and covered with 1 X running buffer.

The samples were loaded into the gel and electrophoresis was applied with voltage of 80 volts for 2.5 hours. The gel was stained with Coomassie Brilliant Blue R250 and de stained with 10% acetic acid 30% methanol destain and dried at 80 ^{o}C in vacuum dryer for 2 hours.

2.2.5 Western blotting

Western Blotting Detection Kit (Amersham Life Science) was used for immunodetection. Following SDS-PAGE separation, the protein was blotted to the nitrocellulose membrane. The gel was placed onto 2 pieces of 3MM Whatman paper, and the nitrocellulose membrane was placed on top of the gel and covered with 2 pieces of 3MM Whatman paper. The cassette was covered with 1X transfer buffer [25 mM Tris?HCl, 192 mM glycine, 20% methanol (v/v) and 0.1% SDS] and run in a BioRad tank for 1 hour at 90 V.

The nitrocellulose membrane was then blocked with 100 ml of 1×TBS [20 mM Tris-HCl (pH 7.5), 500 mM NaCl] containing 5% BSA for 1 hour blocking solution overnight. The nitrocellulose membrane was incubated with either anti- his HRP tag or anti-GST HRP for 45 minutes with dilution of 1:10,000 in TTBS (1XTBS+0.1% Tween 20, v/v) then, the nitrocellulose was washed with TTBS for 5 minutes 5 times. ECL was added to the nitro cellulose membrane in the gel doc machine to detect the chemoluminescence reflecting the bands of the target protein.

2.2.6 Determination of EGFP fluorescence

50 ?l of bacterially expressed protein was transferred to a crystal cuvette and diluted with 450 ?l of water. Then fluorescence emission (without polarisation) was measured to determine the optimal wavelengths. EGFP has minimum emission intensity between 488 and 509 nm. Fluorescence was detected by measuring the intensity of the emission for each expressed protein.

2.3 Examination of the recombinant protein activity

3.3.1 [^{3}H] TCDD ligand binding assay

The principle of this assay is to determine and quantify the functionality of the AhR LBD; all the steps of this assay were performed at 4°C.

2.3.1.1 Preparation of the mouse liver cytosol

Mouse liver cytosol was used as positive control for the AhR ligand binding assay. The mouse was killed by cervical dislocation and the liver was taken out and placed into 20 ml of ice-cold buffer (150 mM KCl pH 7.4) to remove the blood from the liver sample. The liver was cut into small pieces using scissors and placed in 3-fold (w/v) of lysis buffer (MENG buffer (25 mM MOPS, pH 7.5, 1 mM EDTA, 0.02% sodium azide, and 10% glycerol) containing 10 mM sodium molybdate, 2 mM DTT and 1mM PMSF). The liver pieces were then homogenized and the homogenate was centrifuged at 10,000 g for 15 minutes and the supernatant was collected .The supernatant was further centrifuged for 1 hour in the ultracentrifuge and the supernatant was collected. The protein concentration was measured using the Bradford assay. The liver cytosol was stored at –80 ^{o}C until use.

2.3.1.2 Preparation of the bacterially expressed protein

The induced cells were spun and 1 g of the cell pellet was resuspended with 10 ml of ice-cold lysis buffer (MENG buffer containing 2 mM DTT, 100 ?l/ml protease inhibitor cocktail, 1 mM PMSF, and 10 mM sodium molybdate). The cells were lysed using lysozyme followed by freezing at -80 ^{o}C and thawing at 4 °C three times followed by sonication for 30 minutes at maximum power repeated 6 times. The lysate was then centrifuged at 10,000g and supernatant was taken and further centrifuged at 200,000g for 30 minutes in the ultracentrifuge. The supernatant was collected the protein concentration was determined and the sample was stored at -80 ^{o}C.

2.3.1.3 Preparation of receptor dilution

The bacterially expressed proteins were diluted into a range of concentrations. Each protein was diluted into 1.0 to 0.1 mg /ml using ice-cold lysis buffer (MENG buffer containing 2 mM DTT, 100 ?l/ml protease inhibitor cocktail, 1 mM PMSF, and 10 mM sodium molybdate). A 4 mg/ml solution of BSA was added to each sample to reduce nonspecific binding of [^{3}H]TCDD to the AhR receptor.

2.3.1.4 Preparation of the radioactive ligand

The original stock of radioactive ligand, [^{3}H]TCDD has initial specific activity 34.7 Ci/mmol, and a concentration of 0.929 mCi/ml. An aliquot of this stock was diluted to 535 nM with p-dioxane and was used as a master stock stored at -20°C. Further dilution to a final concentration 200 nM was freshly prepared prior to each experiment. For the master stock with p-dioxane, TCAOB was used as non radioactive ligand. 3 mM TCAOB in p-Dioxane was prepared as a master stock and stored at room temperature. Prior to each experiment further dilution of this master stock was made to 40 ?M in p-dioxane.

2.3.1.5 Setting up the binding reaction

Triplicate samples of 200 ?l aliquots of each receptor preparation were transferred into an Eppendorf tube. 1 ?l [^{3}H]TCDD was added to each sample. To test nonspecific binding, the same samples were prepared at the same time with 200 nM of the non radioactive ligand .The samples were thoroughly mixed and incubated for 16 hours at 4°C.

The samples were then treated with dextran-coated charcoal to remove the loosely bound and unbound radioactive ligand; each sample was treated with 30?l of dextran-coated charcoal (33 mg charcoal/ml in MDENG buffer) to each sample to a concentration of 0.2 mg charcoal/mg protein. The sample was mixed and incubated for 10 minutes on ice. The samples then were spun at 14,000g to remove the charcoal/dextran for 10 minutes at 4 °C. 150 ?l of supernatant was transferred from each sample into a scintillation vial containing 5 ml scintillation fluid. The sample was mixed vigorously with scintillation fluid before counting.

The radioactivity of ligand [^{3}H]TCDD was detected by liquid scintillation counting using a Packard Tri-carb Model 2100TR Liquid Scintillation Analyser. The total bound was described as the radioactivity of ligand that bound to the receptor in solution after charcoal treatment. Non-specific binding was described as the bound radioactive ligand in the presence of a 200-fold excess of non-radioactive competitor. The specific binding was described as the difference between the total and non-specific binding. The measurement of bound ligand (d.p.m.) was converted to molar concentration as follows (Hulme and Birdsall 1993):

RL* = B/(V ? SA ? 2220) nM

Where B is the radioligand bound (d.p.m.) corrected for counter background, V is the volume of radioligand assayed (ml), SA is the specific activity of the radioligand.

2.4 Refolding the bacterially expressed GGA and GGAG in human reticulocyte lysate.

The recombinant proteins were prepared as described in section 2.2.3.1. 5 ?l aliquots of each protein sample were incubated with 25 ?l of human reticulocyte lysate at room temperature for one hour, and then the sample was diluted using MENG buffer containing 2 mM DTT, 100 ?l/ml protease inhibitor cocktail, 1mM PMSF, and 10 mM sodium molybdate buffer to 200 ?l. [^{3}H] TCDD specific binding assay was preformed to test the functionality each protein as described in section 2.3.1.5.

3. Results

3.1 Optimization of cell lysis

Cell lysis is an essential step for successful purification of any expressed protein. Many different methods are commonly used to release recombinant protein from bacterial cells such as sonication, enzymatic lysis using lysozyme, and freeze–thaw. Sonication is a wildely used technique for cell lysis but, one of the disadvantages of sonication is overheating of the sample, resulting in protein denaturation, and for that reason it is important to place the sample in ice while sonicating. Freeze–thaw lysis, by freezing at −20^{o}C or −80 °C then thawing at room temperature, is a gentle method, but can often result in incomplete lysis (Peti and Page 2007). It was decided to use a combination of different methods; lysozyme, freeze-thaw and sonication for 30 seconds. Therefore, four different lysis methods were evaluated for their ability to release the proteins from the cells; the first method (A) was using lysozyme (B) was using lysozyme and freeze and thaw three times, the second method (C) was using lysozyme and then sonication, and the last method (D) was using lysozyme and freeze then thaw three times following by sonication. Bradford Assay was performed as described in the Material and Methods to determine the amount of protein obtained from each method.

When comparing the efficiency of the three methods, method (C and D) were clearly more efficient than (A) or (B) and this allowed us to obtain the highest protein yields from the cells. Generally, the condition in which lysozyme is combined with sonication, gives higher amount of proteins extracted from the cells compared to the freeze thaw cycles, when applied independently or in combination with lysozyme.

The result showed that using lysozyme followed by freeze then thaw three times or/ and followed by sonication gives the highest yield of protein. Figure 3.1 shows that using lysozyme and then sonication gives 1.87mg/ml while using lysozyme and freeze then thaw three times followed by sonication gives 1.96mg/ml.

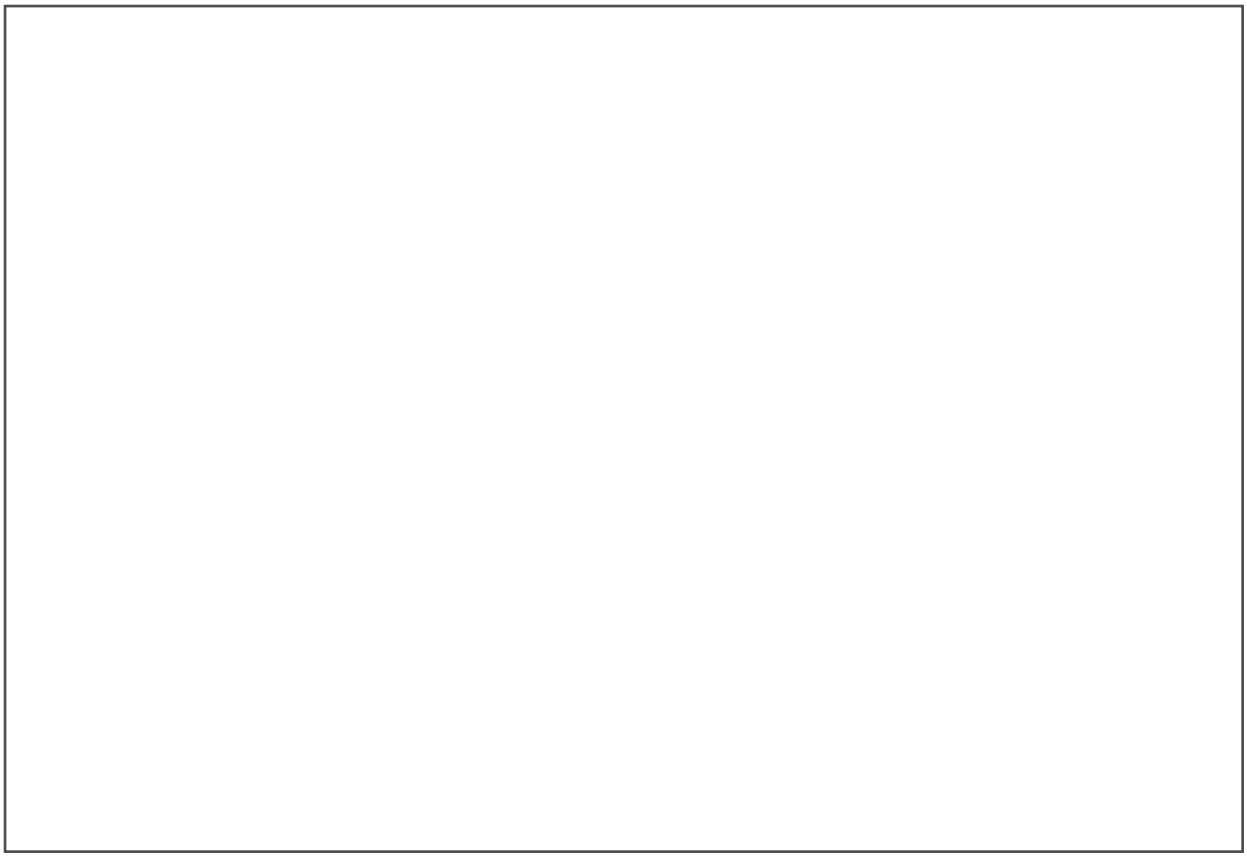


Figure 3.1 Protein concentration result from different methods of cell lysis: 5ml of overnight bacterial culture was diluted in 50 ml LB broth and then incubated for three hours. The cells were centrifuged and resupended with NaCl-Tris buffer. Different lysing methods were used to lyse the cells. The protein concentration was determined using Bradford assay as described in Material and Methods. (A) lysozyme (B) lysozyme and freeze and thaw three times (C) lysozyme and then sonication, and the last method (D) lysozyme and freeze then thaw three times following by sonication.

3.2 Expression of GG, GGA and GGAG in *E. coli* at 15?C

The attempts to express GGA and GGAG recombinants in BL21 D3 at 37^{o}C and purify these recombinants did not produce any success in obtaining functional protein. It was decided to examine if the expression of GG, GGA and GGAG in *E. coli* Arctic Express at 15^{o}C can yield high levels of these proteins. Therefore the three constructs were transformed into *E. coli* Arctic Express.

The three constructs GG, GGA and GGAG were transformed into *E. coli* Arctic Express cells. 5 ml of overnight bacterial culture were diluted with 50 ml LB broth and incubated for 3 hours, and then the cells were induced with 1 mM IPTG and incubated overnight at 15 ^{o}C with shaking (see Materials and Methods).

The cells were centrifuged and resuspended in 10 ml of Tris-NaCl lysing buffer and lysed by incubating the cells with lysozyme for 30 minutes on ice, followed by freeze and thaw three times, followed by sonication. The total protein concentrations of the samples were determined using Bradford protein assay. The total protein samples were analyzed using SDS PAGE.

Figure 3.2 shows three bands in the induced samples compared with un-induced samples and the control: the band in the GG induced sample is around 65 kDa and the induced band in the GGA sample is about 80 kDa. In the induced GGAG the band is around 109 kDa . These are the expected sizes for the three proteins.

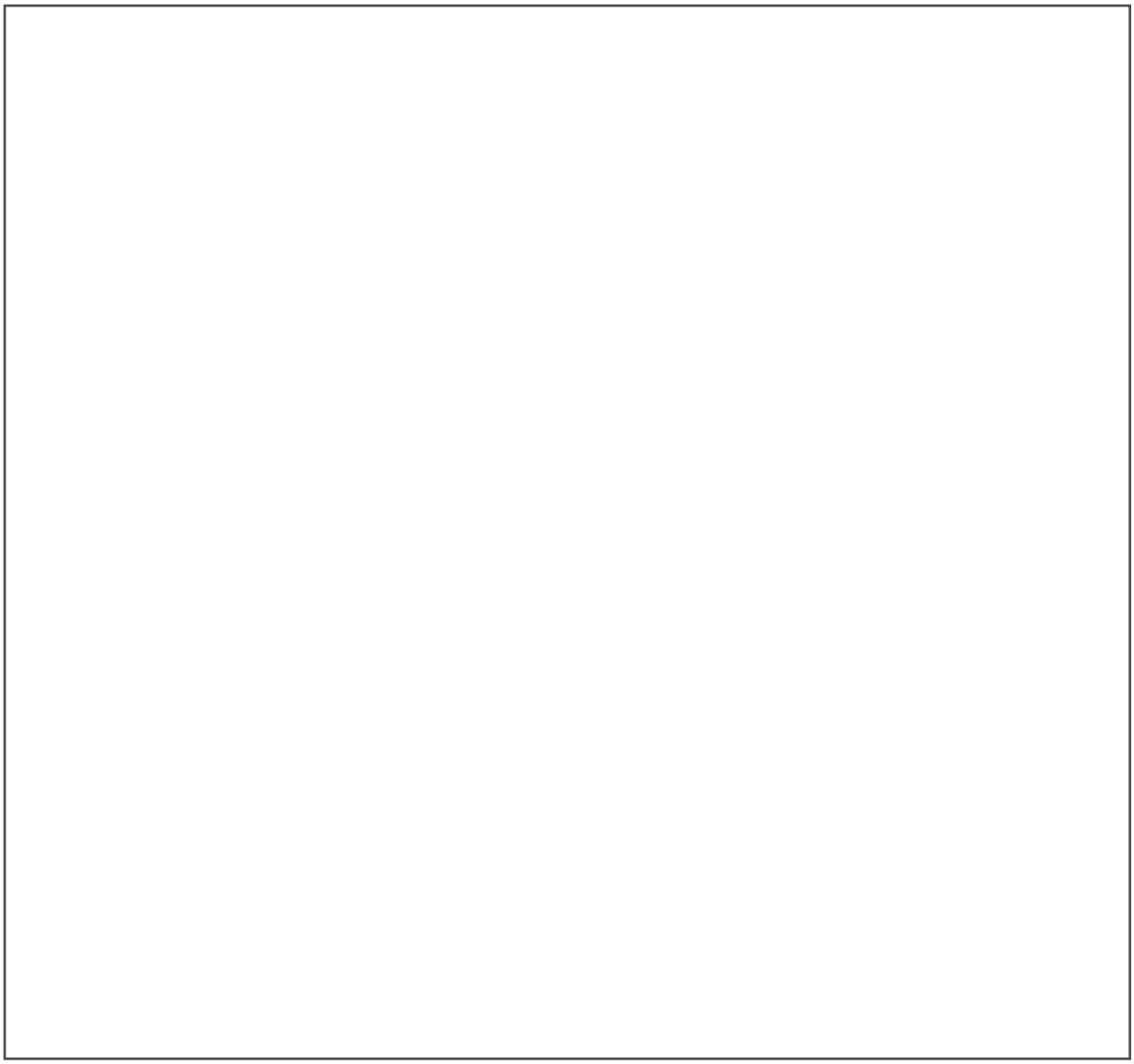


Figure 3.2 Expression of GG, GGA and GGAG in arctic express cells. The total protein was loaded onto SDS PAGE: In first lane 6H molecular weight marker ;(29, 45, 66, 97, 116 and 205 kDa), lane 2 contained untransformed *E coli* arctic express cells, I= induced, U = uninduced.

3.3 Solubility of AhR fusion protein in bacteria

In a previous study performed in the same lab, using the same constructs in order to obtain soluble proteins, the constructs were expressed in *E coli* BL21 at 37^{o}C and SDS PAGE showed no band in the soluble fraction, to examine the solubility of AhR that was expressed in Arctic Express at 15 ^{o}C, the bacterial cells were induced, lysed and the homogenate centrifuged at 10000 g for 10 minutes followed by further centrifugation at 200000g for 30 minutes, The supernatant was taken and loaded onto an SDS PAGE gel (Figure 3.3).

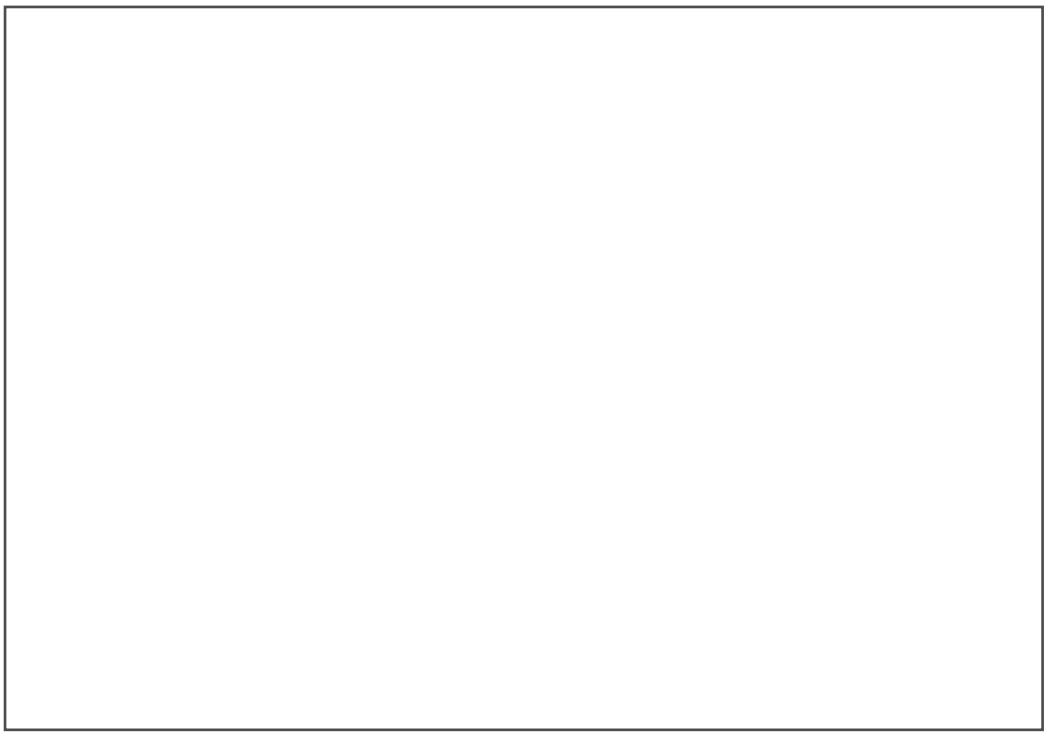


Figure 3.3 Expression of Induced GG, GGA and GGAG in Arctic express at 15°C. Total protein and supernatant were loaded onto SDS PAGE : lane 6H molecular weight marker showing the position of 29, 45, 66, 97, 116 and 205 kD lane 2 contained *E coli* arctic express cells, lane 3 contained induced GG total protein , lane 4 contained induced GGA total protein, lane 5 contained induced GGAG total protein, lane 6 contained induced GG supernatant, lane 7 contained induced GGA supernatant, lane 8 contained induced GGAG supernatant.

From Figure 3.3 there is a band in lane 8 with induced GGAG in the supernatant compared with total protein samples and the control at about 109 kDa, this is the

expected size for the GGAG which is suggesting the solubility of GGAG protein. However, the band for GGA in the soluble fraction is not clear, and the proteins might be present in the soluble fractions at low concentration.

3.4 Expression of GG, GGA and GGAG in *E. coli* Arctic express at 37?C

To examine if the expression of GG, GGA and GGAG in 37?C could yield a high level of soluble AhR, the cells were induced by 1mM of IPTG, incubated for 24 hours at 37^{o} C in the shaker and the samples were analyzed by SDS PAGE. Expression of GG, GGA and GGAG at 37?C showed three expected bands in the total proteins as well as the pellet (after 10000g centrifugation) at sizes 66, 84 and 109 kDa. Figure 3.4.

Figure 3.4 Expression of Induced GG, GGA and GGAG in E. coli Arctic Express at 37°C. Total protein, pellet and supernatant were loaded into SDS PAGE: lane1 6H molecular weight marker showing the position of 29, 45, 66, 97, 116 and 205 kDa, lane 2 contained *E coli* Arctic Express cells, lane 3 contained induced GG total protein, lane 4 contained induced GGAG total protein, lane 5 contained induced GGA total protein, lane 6 contained induced GG supernatant, lane 7 contained induced GGA supernatant, lane 8 contained induced GGAG supernatant, lane 9 contained induced GG pellet, lane 10 contained induced GGA pellet, lane 11 contained induced GGAG pellet

Figure 3.4 shows three bands in the induced samples in the total protein and the pellet compared with the control: the first band in the GG induced sample is around 60 kDalton, the second induced band in the GGA sample is about 84 kDa and the induced GGAG band about 109 kDa. These are the expected sizes for the three proteins. The supernatant samples did not show any solubility of these proteins presumably because 37 ^{o}C expression can decrease the cell's ability to fold the protein properly.

3.5 Expression of GG, GGA and GGAG in *E. coli* BL21 D3 at 37?C

To examine if the expression of GG, GGA and GGAG in a different strain of *E coli* could obtain soluble AhR compared to Arctic Express, the three constructs were transformed into BL21 D3 cells. The cells were induced by 1 mM IPTG and incubated for 24 hours at 37^{o}C in the shaker (see Materials and Methods) and the samples were analyzed by SDS PAGE.

The cells were spun at 10000 g for 10 minutes, re-suspended in Tris-NaCl buffer, subjected to lysis by lysozyme and freezing at -80^{o}C and thaw one time, then sonication for 30 seconds at maximum power. The samples were then centrifuged at 10000g, the supernatant was taken and centrifuged at 200.000g for 30 minutes and all the samples were loaded into SDS PAGE.

Expression of GG, GGA and GGAG at 37?C showed three expected bands only in the total protein are 66, 84 and 109 kDa (Figure 3.5).The supernatant samples did not show any solubility of these proteins.

Figure 3.5 Expression of Induced GG, GGA and GGAG in BL21 D3 at 37ºC. Total protein and supernatant were loaded into SDS PAGE : lane 6H molecular weight marker showing the position of 29, 45, 66, 97, 116 and 205 kDa, lane 2 contained E coli BL21 cells, lane 3 contained induced GG total protein, lane 4 contained induced GGA total protein, lane 5 contained induced GGAG total protein, lane 6 contained induced control BL21 supernatant, lane7 GG supernatant, lane 8 contained induced GGA supernatant, lane 9 contained induced GGAG supernatant.

3.6.1 Purification of GG, GGA, and GGAG

The GST SpinTrap Purification kit was used to purify the proteins from the soluble fraction. Based on the affinity for GST tag were used to purify these proteins. 5 ml of each overnight Arctic express culture was diluted to 200 ml and incubated for 4 hours at 30^{o}C.

The cells were induced by adding 1 mM IPTG and incubated overnight at 15^{o} C. The cultures were centrifuged at 10000g to sediment the cells and re-suspended in 10 ml of ice-cold 1× PBS, and then the cells were lysed using lysosyme, freeze and thaw, and sonication.

The homogenate was spun to remove insoluble material. 600 ?l of the supernatant was transferred to Glutathione Sepharose 4B MicroSpin columns and spun for 1 minute at 735 × g. The columns were washed with 600 ?l PBS and spun. This step was repeated three times. 100–200 ?l of glutathione elution buffer was added to each column, and incubated at room temperature for 5–10 minutes.

Each eluate was collected by centrifugation for one minute at 735 × g. The proteins were analysed using SDS PAGE.

From figure 3.6.1 it can be observed that the three proteins were successfully purified from the soluble fraction, with the expected sizes for the GG, GGA, and GGAG purified proteins compared with unpurified proteins.

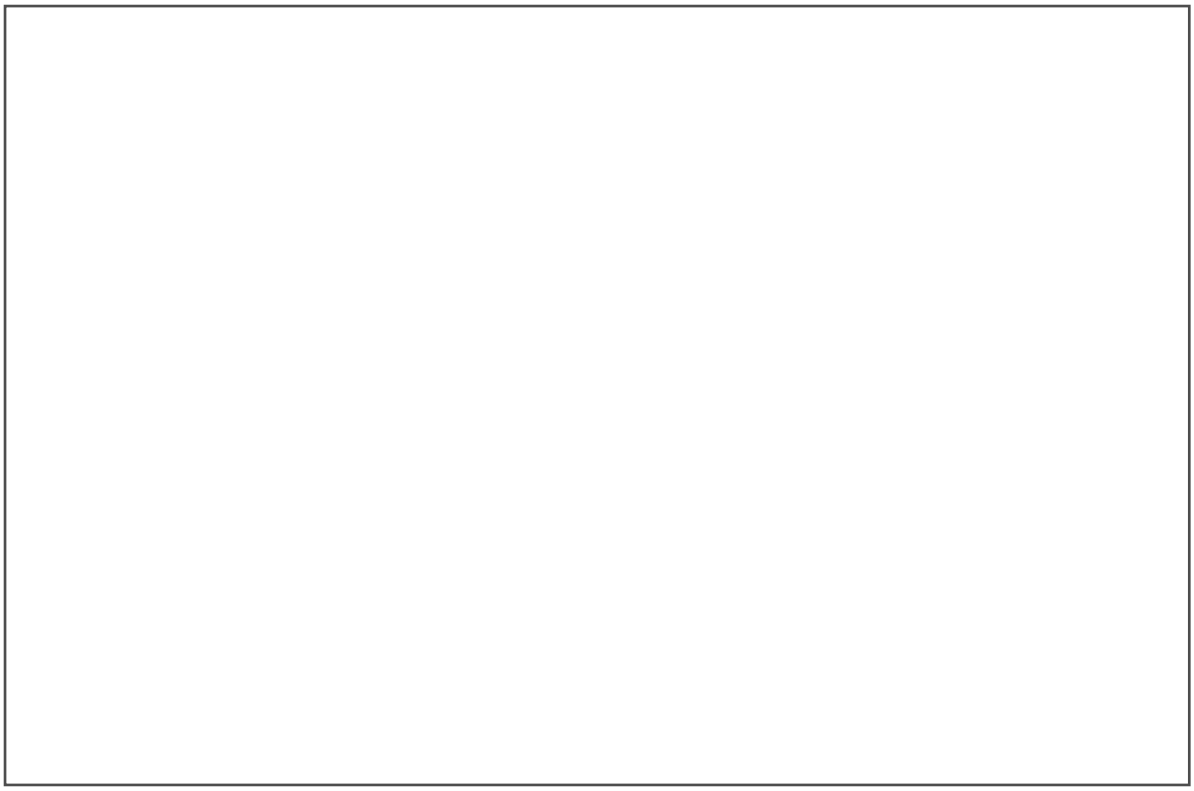


Figure 3.6.1 Purification of the three constructs using GST Tag: lane 1 6H molecular weight marker showing the position of 29, 45, 66, 97, 116 and 205 kDa, lane 2 contained GG purified protein, lane 3 contained GGA purified protein, lane 4 lane 3 contained GGAG, lane 5 contained induced GG unpurified protein, lane 6 contained GGAG unpurified protein, and lane 7 contained GGA unpurified protein.

3.6.2 Detection of the protein fluorescence

The fluorescence intensities of the purified expressed proteins were measured to confirm the presence of EGFP. The emissions of the protein samples were scanned between 500 to 600 nm with excitation at 450 nm. EGFP-AhR constructs showed peaks at 510 nm unlike the controls which had a peak around 540.

50 µl of the purified protein was diluted to 500 µl with water and the fluorescence intensity was scanned. Arctic express bacterial lysate was used as a negative control.

The expressed proteins had a fluorescence emission with peak around 510 nm (Figure 3.6.2).

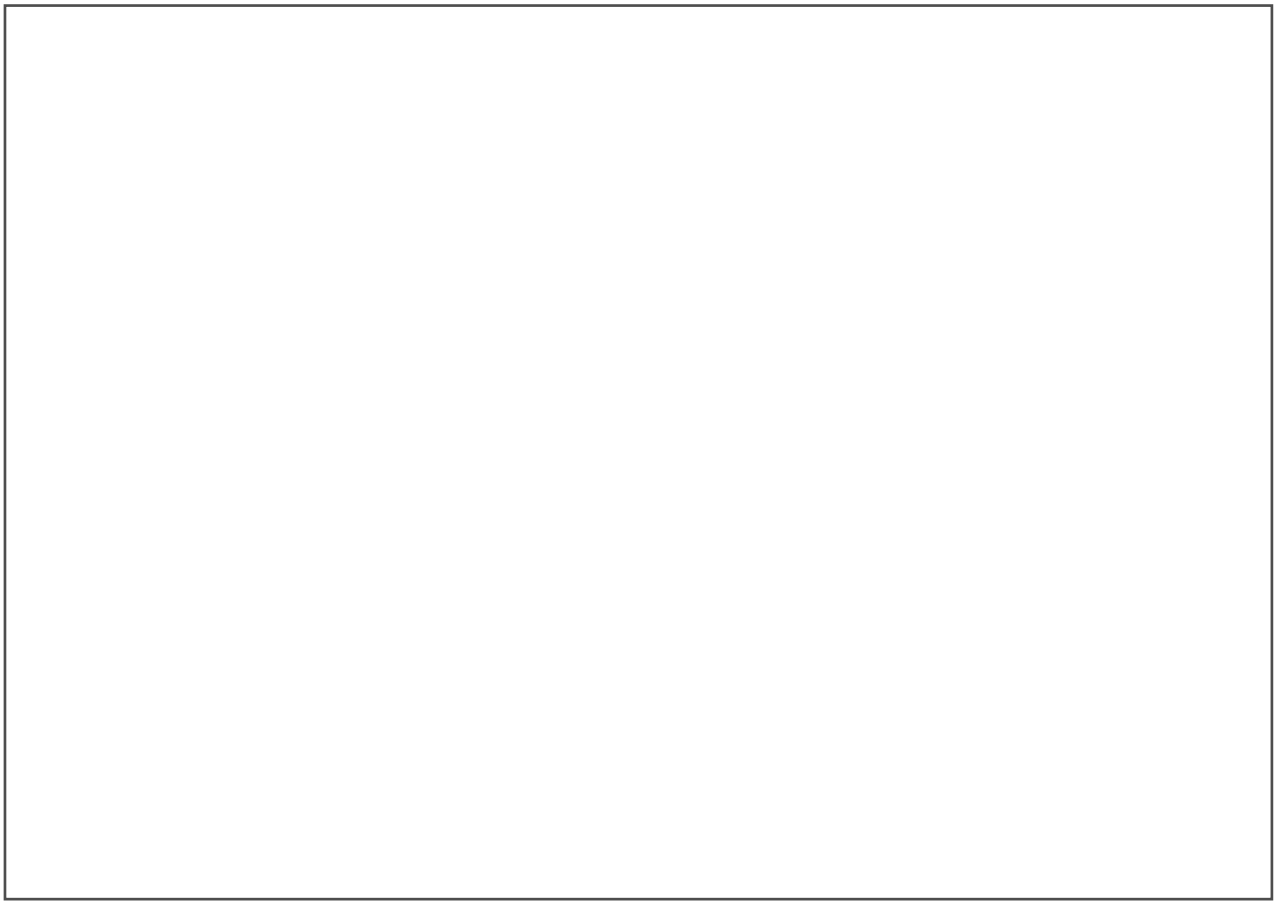


Figure 3.6.2 Fluorescence spectra for the purified proteins. 50 µl of the purified protein was diluted to 500 µl and the fluorescence intensity was scanned between 500 to 600 nm and excitation at 450nm. Arctic express lysate was used as negative control

3.6.3 Mass spectrometry analysis for GG, GGA, and GGAG

GG, GGA, and GGAG proteins were analyzed by mass spectrometry in order to confirm the identity of these expressed proteins. The purified proteins were loaded onto 10% SDS-PAGE.

The expected bands of the proteins on the gel were cut. The samples were subjected to tryptic digestion and the sizes of tryptic fragments were identified by the molecular mass of the ion within 0.1Da mass unit of the predicted size. Table 3.2 shows a significant matching of the tryptic fragment with the three proteins ($p<0.05$).

In GG protein analysis, there were four specific fragments matched with the GFP and three fragments matched with the GST, which confirms the band was GG protein. In the second protein which was expected to be GGA, there were two fragments matched to GST, two fragments matched to the LBD and one fragment matched to the GFP which also confirmed the existence of GGA protein. In GGAG proteins there were two fragments matched to the LBD, one matched to the GST and another matched to GFP, which is indicating that this band is GGAG protein.

Peptide	Mr(expt)	Mr(calc)	Residues	Protein matched
GG				
K.SAMPEGYVQER.T	1265.5682	1265.5710	397-410	GFP
K.FSVSGEGEGDATYGK.L	1502.6522	1502.6525	338-356	GFP
R.TQISSSSFEFCSR.R + Carbamidomethyl (C)	1534.6736	1534.6722	550-567	GFP
K.LTQSMAIIR.Y + Oxidation (M)	1047.5718	1047.5746	63-74	GST
.MSPILGYWK.I	1093.5608	1093.5630	1-9	GST
R.AEISMLEGAVLDIR.Y	1515.8004	1515.7966	88- 102	GST
GGA				
R.IEAIPQIDK.Y	1025.5694	1025.5757	181-192	GST
K.LTQSMAIIR.Y + Oxidation (M)	1047.5788	1047.5746	63-74	GST
K.TGESGMTVFR.L	1083.5014	1083.5019	640-651	AhR
K.TGESGMTVFR.L + Oxidation (M)	1099.4928	1099.4968	640-651	AhR
K.SAMPEGYVQER.T	1265.5682	1265.5710	397-410	GFP
GGAG				
K.SAMPEGYVQER.T	1265.5682	1265.5710	397-410	GFP
K.LTQSMAIIR.Y + Oxidation (M)	1047.5788	1047.5746	63-74	GST
K.TGESGMTVFR.L	1083.5014	1083.5019	640-651	AhR
K.TGESGMTVFR.L + Oxidation (M)	1099.4928	1099.4968	640-651	AhR

Table 3.2 summary of TOF-MS. Tryptic analysis of the three expressed AhR constructs (GG, GGA and GGAG) .Sample were tryptic digested and analysed by TOF-MS. The size of Tryptic fragments were predicted depending on the molecular mass of the ion within 0.1Da mass. The table shows the specific matching of the peptides to sequence within the proteins. The hits were matched using Mascot Search http://www.matrixscience.com p<0.05

3.6.4 Conformation of the protein specificity by western blot

The proteins were separated on 10% SDS-PAGE as described in the Material and Methods section and the gel was transferred onto nitrocellulose membrane. The membrane was blocked with the blocking solution, the membrane was then washed three times in TTBS, and then incubated in His-tag antibody solution in TTBS for 1 hour then washed three times. The blot was detected using the Western Blotting Detection kit.

In figure 3.6.3 three bands of the expected size are observed for the three proteins which have His-tag compared to the negative control.

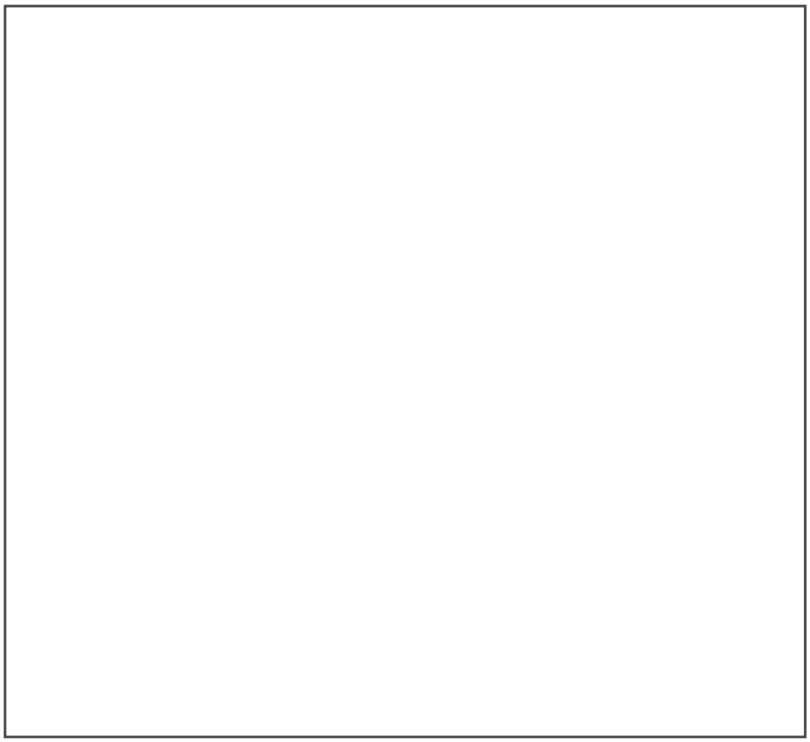


Figure 3.6.3 Detection of Recombinant AhR GG, GGA and GGAG by Western Blotting. Induced Arctic express cells, GG, GGA and GGAG proteins were run on 10% SDS-PAGE then blotted to nitrocellulose membrane. Lane 1 (control) contained induced cells. Lane 2 GG protein, Lane 3 GGA and Lane 4 GGAG.

3.7 The Ligand Binding Assay

This experiment aimed to examine the functionality of the AhR LBD constructs, and to quantitate the ligand binding activity of these constructs using an *in vitro* radioligand binding assay.

The main idea of this binding assay is to perform the experiment in the presence of a constant concentration of ligand and receptor. The unbound and loosely bound ligand should be removed and at equilibrium the binding activity is calculated. Therefore it is important to optimise experimental conditions.

3.7.1 Optimization Ligand Binding Assay

Mouse liver cytosol was used as a source of AhR receptor. Triplicate samples of 200?l of the mouse cytosol protein were used in this assay. 1 ?l aliquot of 1 nM [^{3}H]TCDD was added to each sample.

To test non-specific binding, a 1 ?l aliquot of [^{3}H]TCDD and 200-fold excess of TCAOB competitor were added to each sample; the samples were mixed and then incubated for 16 hours at 4 °C. A 30 ?l aliquot of each sample was transferred to a scintillation vial containing 5ml of scintillation fluid and the samples were analysed by liquid scintillation counting.

The remaining sample in each tube was treated with dextran-coated charcoal to remove unbound and loosely bound ligands. 30 ?l of dextran-coated charcoal (33 mg charcoal/ml in MDENG buffer) was added to each sample; the sample was mixed, and incubated on ice for 10 minutes followed by centrifugation for 10 minutes at 14,000g at 4 °C. A 150 ?l aliquot of supernatant was transferred to a scintillation vial containing 5 ml scintillation fluid. The sample was mixed vigorously and analysed by liquid scintillation counting.

3.7.1.1 The effect of protein concentration

To determine the amount of protein receptor binding to TCDD, and to decide how much protein concentration could be used in the binding assay, different concentrations of mouse liver cytosol (5, 7, and 10 mg/ml) were incubated at 4 °C with 1 nM [^{3}H]TCDD in the presence or absence of a 200-fold excess of competitor TCAOB.

Samples were incubated for 16 hours, followed by charcoal-dextran treatment to remove unbound ligand. A 150 ?l aliquot of each sample was analysed by scintillation counting. Total, nonspecific and specific binding were determined.

Figure 3.7.1 shows that there was a specific binding of [^{3}H]TCDD to mouse liver cytosol, at different protein concentrations ranging from 0.5 to 7 mg/ml. The determined binding capacities of mouse cytocol AhR using different protein concentration were 45.2, 31.5, 22.4 and 16.7 fmol/mg for 0.5, 2, 5, and 7 mg protein concentrations. Figure 3.7.1 concludes using lower protein concentrations appears to give a higher binding capacity.

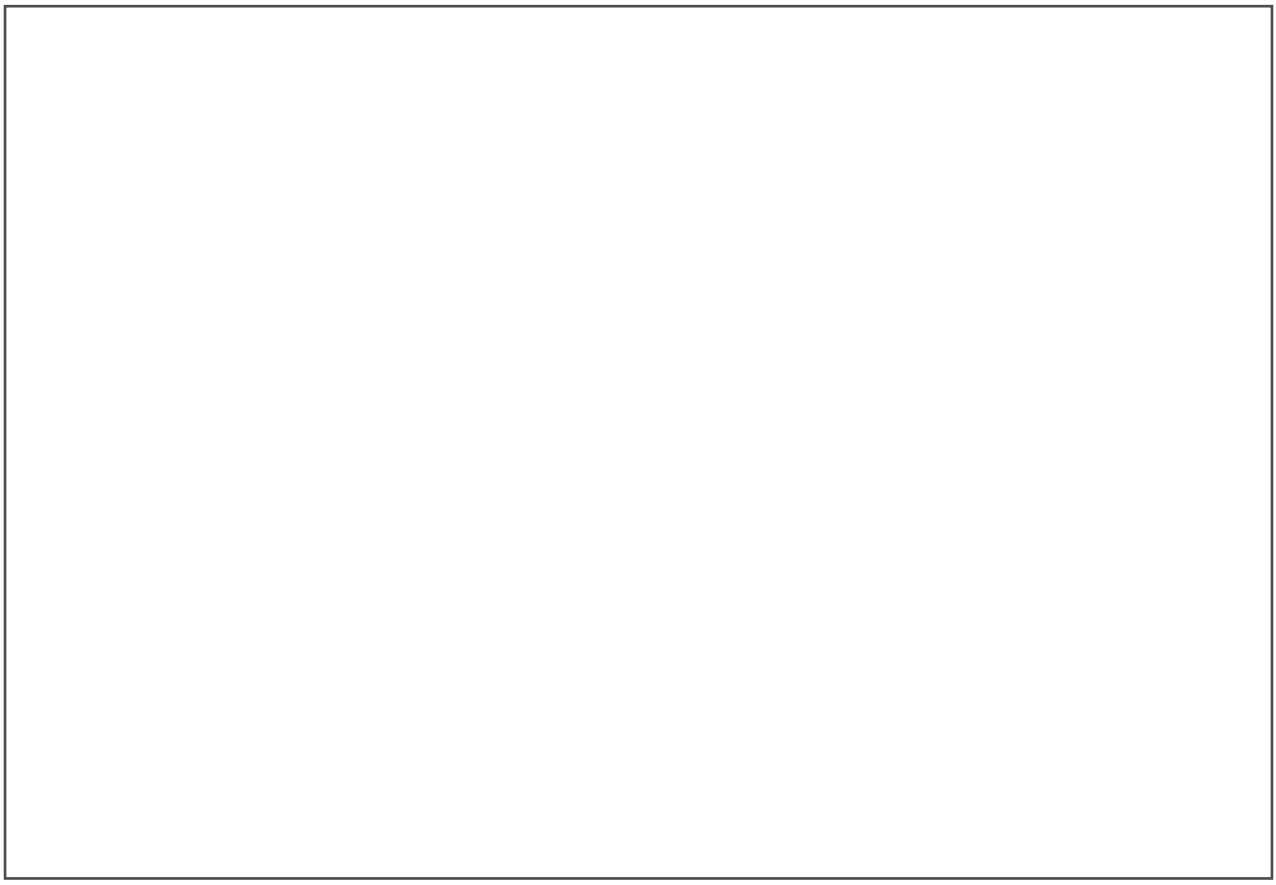


Figure 3.7.1 the effect of protein concentration. Different concentration of mouse liver cytosol (0.5, 2, 5 and 7 mg/ml) were incubated at 4°C with 1nM [3H] TCDD in the presence or absence of a 200-fold excess of competitor TCAOB as described in Material and Method. Samples were incubated for 16h, followed by charcoal-dextran treatment to remove unbound ligand. 150?l Aliquot of each sample was analysed by scintillation counting. Total, nonspecific and specific binding were determined.

3.7.2 Binding assay standard Curve

A [^{3}H]TCDD standard binding assay was performed to determine the apparent affinity constant (Kd) and the maximum concentration of specific binding sites, Bmax. 5 mg/ml of mouse liver cytosol aliquot was incubated with several concentrations of [^{3}H]TCDD (0-2.0 nM) in the presence and absence of 200-fold molar excess TCAOB and incubated for 16 h at 4 ºC. 30 ?l of dextran-coated charcoal (33 mg charcoal/ml in MDENG buffer) was added after the 16-hours incubation and radioactivity measured.

The analysis of saturation binding curve was plotted using nonlinear regression for single site binding hyperbola analysis in GraphPad Prism version 5.

Figure 3.7.2 shows that [^{3}H]-TCDD specific binding to mouse cytosol started to be stable

at [^{3}H]-TCDD concentration ? 1 nM. The apparent Kd for specific binding of [^{3}H]TCDD was 0.50 ± 0.07 nM (103 fmol/mg) and the concentration of binding sites, Bmax was 0.23 ± 0.01 nM (46 fmol/mg).

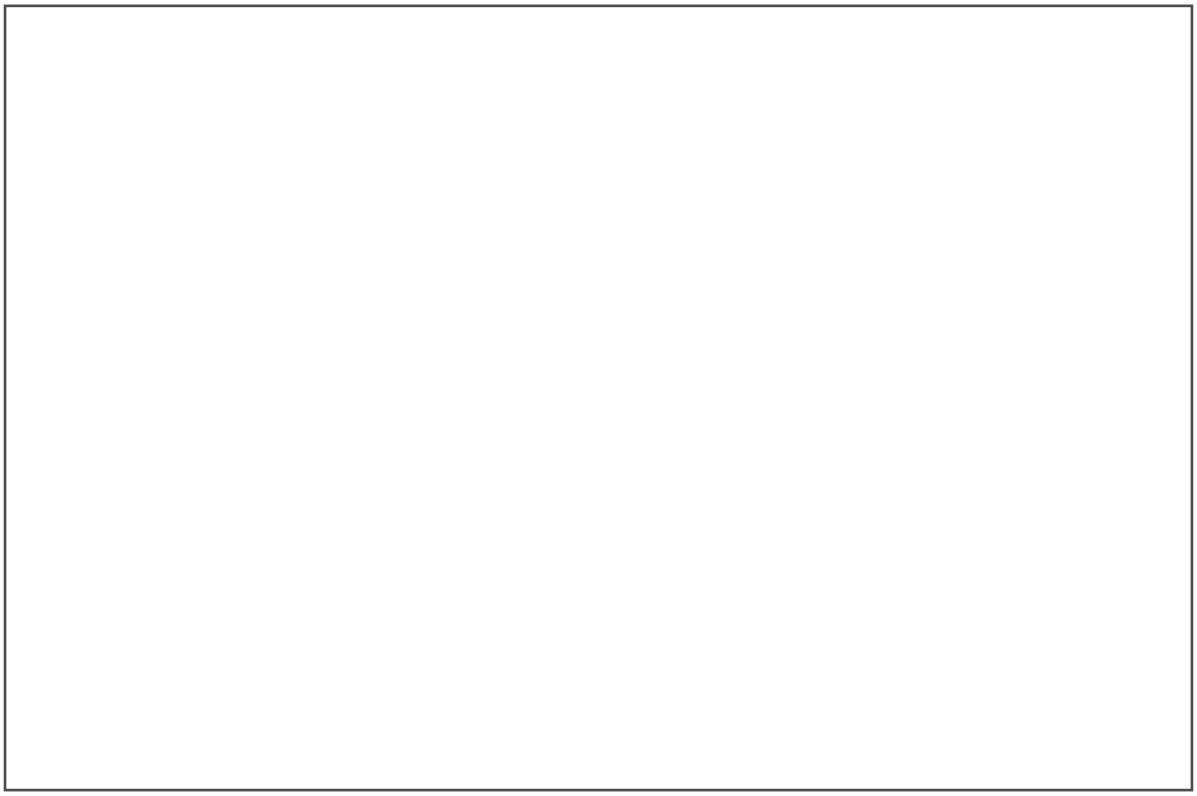


Figure 3.7.2 Specific Binding of [^{3}H]TCDD to mouse liver cytosol: Range of 0 to 2.0 nM [3H]TCDD were incubated with 5mg/ml of mouse liver cytosol in the presence and absence of a 200-fold excess of competitor TCAOB. After 16h incubation at 4C, unbound and loosely bound radioligand was removed by charcoal-dextran treatment, specific binding data were plotted to show specific binding capacities (fmol/mg cytosol protein). Each point represents the average of triplicate samples. Data were analysed using one site binding / GraphPad Prism version 5.

3.7.3 The specific Binding of [^{3}H]TCDD to GG, GGA and GGAG

The aim of this experiment is to quantify the functionality of the bacterially expressed proteins.

200 ?l aliquots of each protein sample were incubated with 1 ?l [^{3}H]TCDD in triplicate. To test non-specific binding, the same samples were prepared at the same time with 200 nM of the non radioactive ligand and incubated for 16 hours at 4 °C.

The samples were then treated with Dextran-coated charcoal to remove the loosely bound and unbound radioactive ligand; each sample was treated with 30 ?l of dextran-coated charcoal

In Figure 3.7.3, it can be seen there is no specific binding within the three protein samples compared to the controls. The binding capacity for RCL (Rat liver cells) was 45.2 fmol/mg whereas in the GGA the binding capacity was very low (5 fmol/mg) and no binding capacity was detected in GGAG.

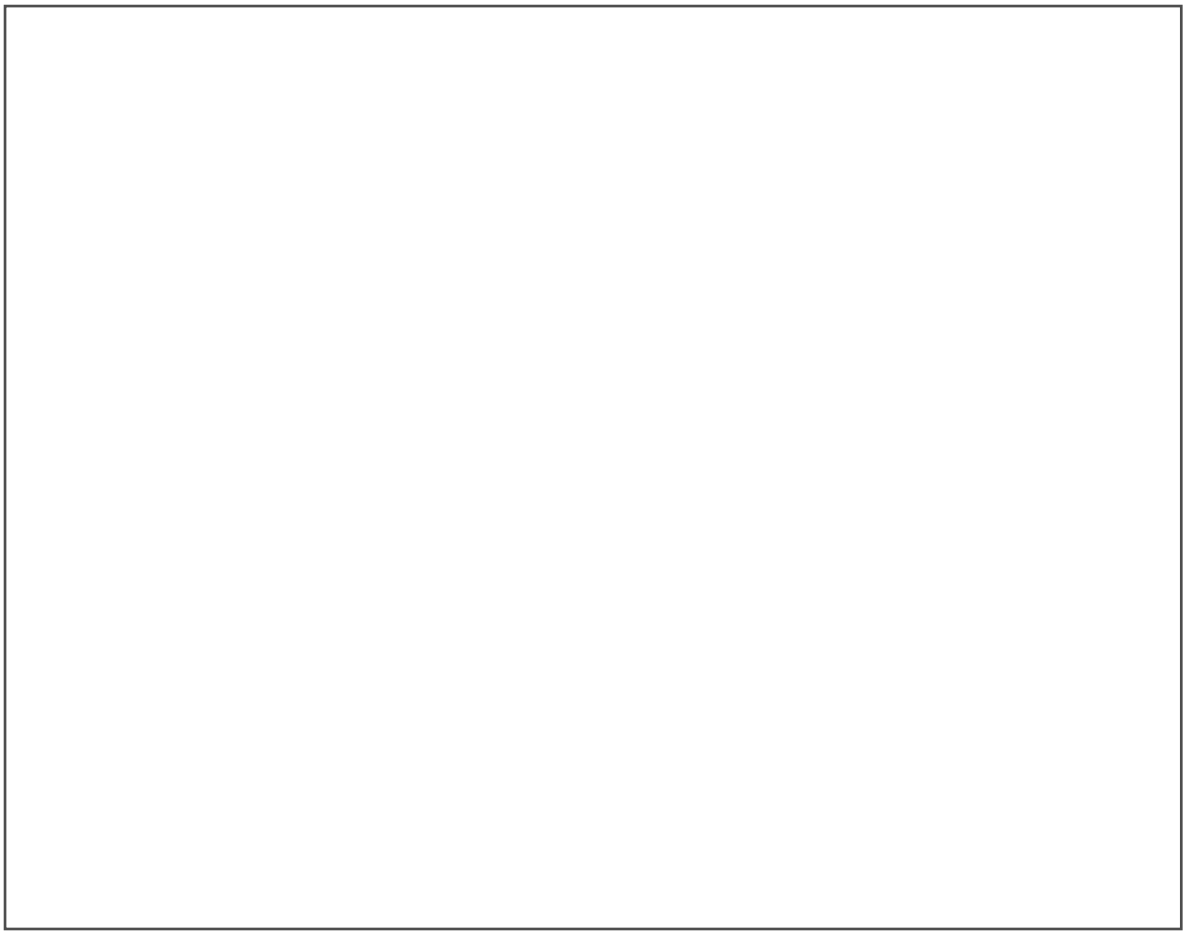


Figure 3.7.3 Specific binding of [^{3}H]TCDD to GG, GGA, and GGAG: liver cytosol was used as positive control, BSA and GG as negative control. The proteins were incubated with [3H]TCDD in presence and absence of a 200-fold excess of TCAOB as a competitor for 16h incubation at 4 ^{o}C, free radioligand was removed by charcoal-dextran treatment and the specific binding was detected

3.8 Refolding the bacterially expressed GGA and GGAG in Human reticulocyte lysate.

The recombinant proteins were prepared as described in section 2.2.3.1. 5 ?l aliquots of each protein sample were incubated with 25 ?l of human reticulocyte lysate at room temperature for one hour, and then the sample was diluted using MENG buffer to 200 ?l. [^{3}H]TCDD specific binding assay was preformed to test the functionality each protein.
Figure 3.8 shows the radioligand binding activity in each protein; there was a binding activity around 0.022 nmol in both GGA and GGAG sample compared with control samples

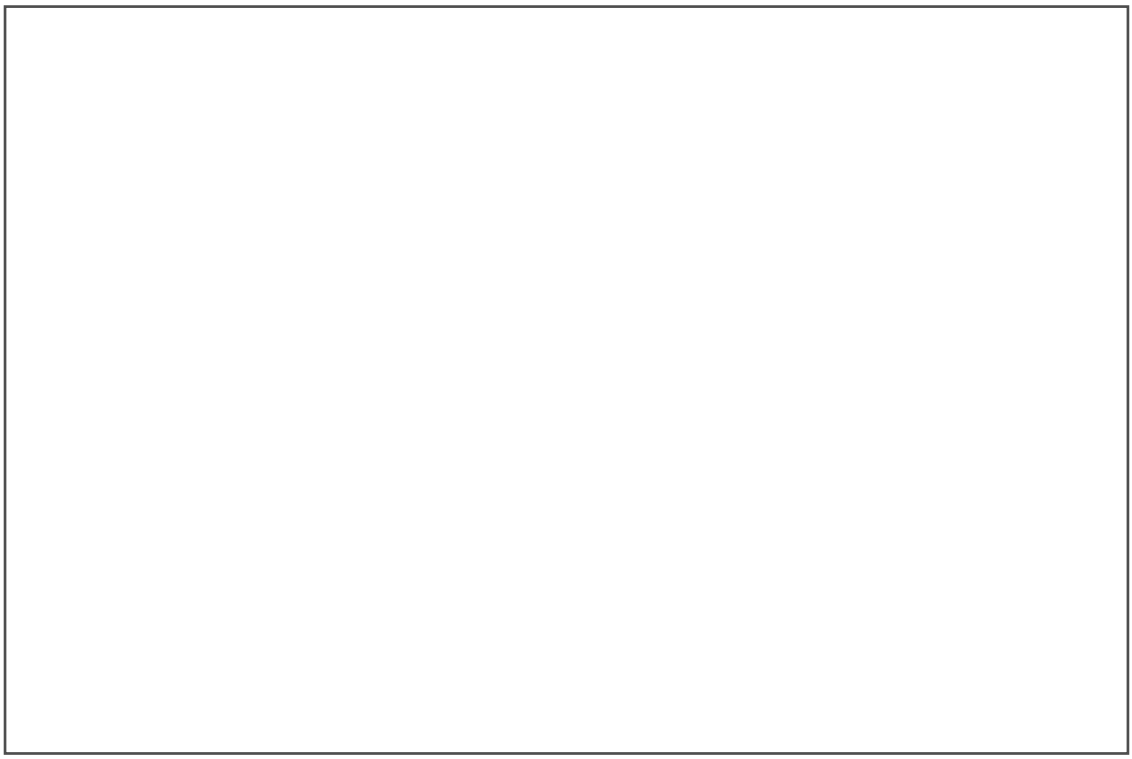


Figure 3.8 Refolding the bacterially expressed GGA and GGAG in Human reticulocyte lysate. 5 ?l aliquots of each protein sample were incubated with 25 ?l of human reticulocyte lysate at room temperature for one hour, and then the sample was diluted to 200 ?l. [^{3}H]TCDD specific binding assay was preformed to test the functionality each protein

4- Discussion

4.1 GGA and GGAG expression

Fusion proteins are used to facilitate the identification and the purification of proteins. Examples include fusing to a GST tag or His tag (Uhlen et al., 1983), to improve the yield of a recombinant protein (Butt et al., 1989) and to improve the solubility of a recombinant protein by fusing to a highly soluble partner that would help the proteins fold independently and functionally (Schein and Noteborn, 1988). GGA and GGAG recombinant proteins were originally made by Shaikh-Omer (2006) by fusing one or two EGFP proteins to mouse AhR LBD and tagged with GST protein. EGFP was used to enhance the solubility of the LBD and GST-EGFP was used as a control.

E. coli strain BL21 D3 is widely used to express recombinant proteins since it is easy, cheap and yields a high level of protein. However, the attempts to express GGA and GGAG recombinants in BL21 D3 at 37 ^{o}C did not produce any success in obtaining functional protein (Shaikh-Omer 2006).

The difficulty when trying to express LBD recombinant proteins in BL21 D3, results in incorrectly folded proteins and aggregations called inclusion bodies. Normal cultivation temperature for the *E. coli* could decrease the cell's ability to fold the protein properly (Miyake et al., 2007). Therefore, by lowering the cultivation temperature, the proteins could be expressed and fold correctly in *Escherichia coli* (Annamalai et al., 2009).

It has been demonstrated that expression at a low temperature condition, increases the stability of the fusion protein. In addition, it also limits the toxic effect of T7 driven expression (Mujacic et al., 1999). Arctic express strain (Stratagene) was chosen for this purpose because of tt's ability to grow at lower temperatures. Arctic express strain has been genetically engineered in order to introduce two proteins known as cold-adapted Chaperonins. This strain specifically produced Chaperonin Cpn60 and co-chaperonin Cpn10 from *Oleispira antarctica*. It was demonstrated that co-expression of the Chaperonins facilitates grown at lower temperatures (4-15 ^{o}C) and helps the recombinant proteins to fold correctly; this was proven to be 180 fold more than that expressed in an ordinary *E. coli* strain (Ferrer et al., 2004).

In this study it has been confirmed that the expression of GGA and GGAG in *E. coli* Arctic Express at 15 ^{o}C yields soluble proteins. Figure 3.3 shows three bands in the induced samples compared with un-induced samples and the control. There can be seen clearly in the soluble fraction a band at 112 kD, consistent with GGAG's size, and this suggests that band is GGAG.

To guarantee that the protein in the soluble fraction was not a contamination from the pellet, the soluble fraction was further centrifuged at 240,000 x g for 30 minutes. This result suggests the solubility of this protein. However, the GGA band in the soluble fraction was not clear; this could be explained by the possibility that protein present in the soluble fractions was at a low concentration, in agreement with Alshikh- Omer (2006) who also suggested that the GGA was not very

soluble.

The soluble protein co-expressed with the chaperonin related to the lower temperature expression not the chaperonin-mediated the folding (Ferrer et al., 2004). Expression of GGA and GGAG in *E. coli* Arctic Express at 37 ^{o}C did not show any protein solubility compared with the total protein in the pellet, which suggests that the solubility of the protein was not because of the chaperonin but because of the low temperature (figure 3.4).

4.2 Optimization of cell lysis

Three different lysis methods were evaluated in order to release the highest amount of the recombinant expressed protein. These methods were enzymatic lysis by lysozyme treatment, sonication, and freeze–thaw.

Freeze–thaw is an effective method, done by re-suspending the cells in lysis buffer and then freezing at −80 ^{o}C followed by thawing on ice. This was repeated three times. This method is gentle and it is recommended for *E. coli* expression. However, this method does not guarantee a complete lysis of the cells (Peti and Page, 2007). On the other hand, sonication is a powerful method widely used for cell lysis. Although the maximum yield of protein can be produced by this technique, the disadvantage of this procedure is that the sample could overheat if the sonication was applied for a long time, resulting in protein denaturation. To overcome this problem, it was necessary to reduce the sonication time and to place the sample on ice while sonicating (Joersbo and Brunstedt, 1990).

Therefore, it was decided to use a combination of different methods; lysozyme, freeze-thaw and sonication for 30 seconds. When comparing the efficiency of the three methods, using lysozyme followed by freeze then thaw repeated three times or/ and followed by sonication, gave the highest yield of protein 1.96 mg/ml.

4.3 Purification of GG, GGA, and GGAG

GST tag facilitates a high-quality purified protein from cell extracts. Glutathione S-transferase fusion protein can be purified easily from the bacterial lysates without denaturing the protein by immobilising the GST in affinity column (Johnson et al., 1988).

The GST Spin Trap Purification kit is a GST pre packed column (from GE Healthcare). This system was used to purify the soluble proteins from the bacterial crude extraction. In Figure 3.6.1 it can be observed from the SDS PAGE analysis, that the three proteins were successfully purified from the soluble fraction, with the expected sizes for the GG, GGA, and GGAG purified proteins compared with un-purified proteins. Protein eluted was quantified by Bradford assay and nearly half of the soluble protein was recovered in both GGA and GGAG. This proves that the protocol used was successful and both proteins could be found to be soluble GGA and GGAG. Comparing this result with Skaikh-Omer (2006), he did not purify any bacterially expressed proteins, probably because of the very low level of the

soluble protein. The purification of GGA and the GGAG was confirmed by western blot, fluorescent detection and Mass spectrometry analysis

In figure 3.6.3 three bands of the expected size are observed for the three proteins which have His-tag compared to the negative control. The fluorescence intensities of the purified expressed proteins were also measured to confirm the presence of EGFP. The emissions of the protein samples were scanned between 500 to 600 nm with excitation at 450 nm. EGFP- AhR constructs showed peaks at 510 nm unlike the controls which had a peak around 540.

Mass spectrometry analysis confirms the identity of GG, GGA, and GGAG proteins. There was a significant matching of the tryptic fragments with the three proteins ($p<0.05$). In GG protein analysis, there were four specific fragments matched with the GFP and three fragments matched with GST, which is consistent with the expected structure of the GG protein, the second band which was expected to be GGA, there were two fragments matched to GST, two fragments matched to the LBD and a fragment matched to the GFP, consistent with GGA protein. In GGAG bands there were two fragments matched to the LBD, one matched to the GST and another matched to GFP indicating that this band is GGAG protein.

4.4 The Ligand Binding Assay

The *in vitro* AhR binding has been reported in different papers as an approach to investigate the functionality of the AhR. A number of methods have been described to characterize the *in vitro* AhR binding, such as using [^{3}H]TCDD (Poland et al 1976; Bradfield and Poland 1988) or using [^{125}I] 2-iodo-2,3-dibromodibenzo-p-dioxin (Bradfield et al., 1988). In this study, AhR was treated with [^{3}H]TCDD followed by charcoal adsorption and then centrifugation to remove the free radioligand and the non specifically bound radioligand, and determining the binding kinetics of AhR binding. Triplicate samples of Mouse liver cytosol were used as a source of AhR receptor and protein concentration was determined to relate protein concentration to the binding assay.

Protein concentration is an important parameter affecting the result of the binding assay. The specific binding of [^{3}H]TCDD to mouse liver cytosol was increased by lowering the concentration of protein. The binding capacities of mouse AhR using different protein concentration were measured as 45.2, 31.5, 22.4 and 16.7 fmol/mg for 0.5, 2, 5, and 7 mg/ml protein concentrations (finger 3.7.1). This result was consistent with Bradfield and colleagues‘(1988) work, which found that lowering the protein concentrations improves the specificity of AhR binding the using the radioligand [^{3}H]TCDD . (Bradfield et al, 1988)

The apparent Kd for specific binding of [^{3}H]-TCDD was 0.52±0.06 nmol (103 fmol/mg) and the concentration of binding sites, Bmax, was 0.23± 0.01 nmol (46 fmol/mg).This assay is robust and it‘s results are reliable with other research that used mouse AhR giving Kd of 0.37 nmol af Bmax 40 fmol/mg (Bazzi 2008) and a Kd 0.27 nmol Bmax of 84 fmol/mg (Poland et al,1976).

4.5 Refolding the bacterially expressed GGA, GGAG in Human reticulocyte lysate

The bacterially expressed AhR LBD was not functional because of the lack of the chaperones; on the other hand expression of the AhR LBD in a eukaryotic system such as reticulocyte lysate was functional (Shaikh-Omar, 2006). However, expression of AhR LBD in mammalian cells yields a very low amount of funcional AhR protein (Shaikh-Omar, 2006).

It has been established that chaperone proteins are necessary for AhR protein to behave functionally, by preventing the aggregation of the AhR protein. Hsp90 is one of the most important proteins it is required for proper folding AhR and regulation of the AhR function (Pongratz et al. 1992; Carver et al. 1994; Whitelaw et al. 1995).

Hsp90 shows similar regulatory activities to the AhR and glucocorticoid receptors, and both of them depend on Hsp90 to fold in the correct conformation (Pratt, 1997). In an experiment to refold the glucocorticoid receptor in reticulocyte lysate, incubating it in reticulocyte lysate refolds the receptor in a binding conformation that is needed to bind

the hormone (Dittmar et al., 1997).

In this study, bacterially expressed AhR LBD was refolded in Human reticulocyte lysate; the GGA and GGAG refolded proteins were tested using a ligand binding assay. Both GGA and GGAG were functional and bound to the radioactive TCDD.

Finally, this study established the expression of GFP AhR (LBD) recombinant protein to enable study the interaction of AhR with associated chaperone protein by pulldown experiments to separate all the proteins associated with the AhR LBD. The associated proteins can be separated on SDS PAGE and then sequence them using MS-MS sequencing.

Also enable to investigate the change in AhR conformation after the ligand binding of the AhR and using fluorescence analytical techniques such as Fluorescence Correlation Spectroscopy (FCS) and Fluorescence Resonance Energy Transfer (FRET).

References

Adachi, J., Mori, Y., Matsui, S., Takigami, H., Fujino, J., Kitagawa, H., Miller, C.A., 3rd, Kato, T., Saeki, K., and Matsuda, T. (2001). Indirubin and indigo are potent aryl hydrocarbon receptor ligands present in human urine. J Biol Chem *276*, 31475-31478.

Allen, G.S., Glover, A.B., McCulloch, M.W., Rand, M.J., and Story, D.F. (1975). Modulation by acetylcholine of adrenergic transmission in the rabbit ear artery. Br J Pharmacol *54*, 49-53.

Annamalai, T., Dani, N., Cheng, B., and Tse-Dinh, Y.C. (2009). Analysis of DNA relaxation and cleavage activities of recombinant Mycobacterium tuberculosis DNA topoisomerase I from a new expression and purification protocol. BMC Biochem *10*, 18.

Antonsson, C., Whitelaw, M.L., McGuire, J., Gustafsson, J.A., and Poellinger, L. (1995). Distinct roles of the molecular chaperone hsp90 in modulating dioxin receptor function via the basic helix-loop-helix and PAS domains. Mol Cell Biol *15*, 756-765.

Backlund, M., and Ingelman-Sundberg, M. (2004). Different structural requirements of the ligand binding domain of the aryl hydrocarbon receptor for high- and low-affinity ligand binding and receptor activation. Mol Pharmacol *65*, 416-425.

Bell, D.R., and Poland, A. (2000). Binding of aryl hydrocarbon receptor (AhR) to AhR-interacting protein. The role of hsp90. J Biol Chem *275*, 36407-36414.

Beresford, A.P. (1993). CYP1A1: friend or foe? Drug Metab Rev *25*, 503-517.

Bertrand, R., Beauchemin, M., Dayan, A., Ouimet, M., and Jolivet, J. (1995). Identification and characterization of human mitochondrial methenyltetrahydrofolate synthetase activity. Biochim Biophys Acta *1266*, 245-249.

Birnbaum, L.S. (1994). Evidence for the role of the Ah receptor in response to dioxin. Prog Clin Biol Res *387*, 139-154.

Bisson, W.H., Koch, D.C., O'Donnell, E.F., Khalil, S.M., Kerkvliet, N.I., Tanguay, R.L., Abagyan, R., and Kolluri, S.K. (2009). Modeling of the aryl hydrocarbon receptor (AhR) ligand binding domain and its utility in virtual ligand screening to predict new AhR ligands. J Med Chem *52*, 5635-5641.

Bradfield, C.A., Kende, A.S., and Poland, A. (1988). Kinetic and equilibrium studies of Ah receptor-ligand binding: use of [125I]2-iodo-7,8-dibromodibenzo-p-dioxin. Mol Pharmacol *34*, 229-237.

Burbach, K.M., Poland, A., and Bradfield, C.A. (1992). Cloning of the Ah-receptor cDNA reveals a distinctive ligand-activated transcription factor. Proc Natl Acad Sci U S A *89*, 8185-8189.

Butt, T.R., Jonnalagadda, S., Monia, B.P., Sternberg, E.J., Marsh, J.A., Stadel, J.M., Ecker, D.J., and Crooke, S.T. (1989). Ubiquitin fusion augments the yield of cloned gene

products in Escherichia coli. Proc Natl Acad Sci U S A *86*, 2540-2544.

Carver, L.A., and Bradfield, C.A. (1997). Ligand-dependent interaction of the aryl hydrocarbon receptor with a novel immunophilin homolog in vivo. J Biol Chem *272*, 11452-11456.

Cox, M.B., and Miller, C.A., 3rd (2002). The p23 co-chaperone facilitates dioxin receptor signaling in a yeast model system. Toxicol Lett *129*, 13-21.

Dittmar, K.D., Demady, D.R., Stancato, L.F., Krishna, P., and Pratt, W.B. (1997). Folding of the glucocorticoid receptor by the heat shock protein (hsp) 90-based chaperone machinery. The role of p23 is to stabilize receptor.hsp90 heterocomplexes formed by hsp90.p60.hsp70. J Biol Chem *272*, 21213-21220.

Fernandez-Salguero, P.M., Hilbert, D.M., Rudikoff, S., Ward, J.M., and Gonzalez, F.J. (1996). Aryl-hydrocarbon receptor-deficient mice are resistant to 2,3,7,8-tetrachlorodibenzo-p-dioxin-induced toxicity. Toxicol Appl Pharmacol *140*, 173-179.

Ferrer, M., Chernikova, T.N., Timmis, K.N., and Golyshin, P.N. (2004). Expression of a temperature-sensitive esterase in a novel chaperone-based Escherichia coli strain. Applied and Environmental Microbiology *70*, 4499-4504.

Freeman, B.C., and Morimoto, R.I. (1996). The human cytosolic molecular chaperones hsp90, hsp70 (hsc70) and hdj-1 have distinct roles in recognition of a non-native protein and protein refolding. EMBO J *15*, 2969-2979.

Freeman, B.C., Toft, D.O., and Morimoto, R.I. (1996). Molecular chaperone machines: chaperone activities of the cyclophilin Cyp-40 and the steroid aporeceptor-associated protein p23. Science *274*, 1718-1720.

Fukunaga, B.N., Probst, M.R., Reisz-Porszasz, S., and Hankinson, O. (1995). Identification of functional domains of the aryl hydrocarbon receptor. J Biol Chem *270*, 29270-29278.

Grenert, J.P., Sullivan, W.P., Fadden, P., Haystead, T.A., Clark, J., Mimnaugh, E., Krutzsch, H., Ochel, H.J., Schulte, T.W., Sausville, E.*, et al.* (1997). The amino-terminal domain of heat shock protein 90 (hsp90) that binds geldanamycin is an ATP/ADP switch domain that regulates hsp90 conformation. J Biol Chem *272*, 23843-23850.

Gu, Y.Z., Hogenesch, J.B., and Bradfield, C.A. (2000). The PAS superfamily: sensors of environmental and developmental signals. Annu Rev Pharmacol Toxicol *40*, 519-561.

Hahn, M.E. (1998). The aryl hydrocarbon receptor: a comparative perspective. Comp Biochem Physiol C Pharmacol Toxicol Endocrinol *121*, 23-53.

Hankinson, O. (1995). The aryl hydrocarbon receptor complex. Annu Rev Pharmacol Toxicol *35*, 307-340.

Henry, E.C., and Gasiewicz, T.A. (2008). Molecular determinants of species-specific agonist and antagonist activity of a substituted flavone towards the aryl hydrocarbon receptor. Archives Of Biochemistry And Biophysics *472*, 77-88.

Higashi, Y., Hiromasa, T., Tanae, A., Miki, T., Nakura, J., Kondo, T., Ohura, T., Ogawa, E., Nakayama, K., and Fujii-Kuriyama, Y. (1991). Effects of individual mutations in the P-450(C21) pseudogene on the P-450(C21) activity and their distribution in the patient genomes of congenital steroid 21-hydroxylase deficiency. J Biochem *109*, 638-644.

Ikuta, T., Eguchi, H., Tachibana, T., Yoneda, Y., and Kawajiri, K. (1998). Nuclear localization and export signals of the human aryl hydrocarbon receptor. J Biol Chem *273*, 2895-2904.

Joersbo, M., and Brunstedt, J. (1990). Protein synthesis stimulated in sonicated sugar beet cells and protoplasts. Ultrasound Med Biol *16*, 719-724.

Johnson, B.D., Chadli, A., Felts, S.J., Bouhouche, I., Catelli, M.G., and Toft, D.O. (2000). Hsp90 chaperone activity requires the full-length protein and interaction among its multiple domains. J Biol Chem *275*, 32499-32507.

Johnson, E.W., Blalock, J.E., and Smith, E.M. (1988). ACTH receptor-mediated induction of leukocyte cyclic AMP. Biochem Biophys Res Commun *157*, 1205-1211.

Kazlauskas, A., Poellinger, L., and Pongratz, I. (1999). Evidence that the co-chaperone p23 regulates ligand responsiveness of the dioxin (Aryl hydrocarbon) receptor. J Biol Chem *274*, 13519-13524.

Kazlauskas, A., Poellinger, L., and Pongratz, I. (2000). The immunophilin-like protein XAP2 regulates ubiquitination and subcellular localization of the dioxin receptor. J Biol Chem *275*, 41317-41324.

Kazlauskas, A., Sundstrom, S., Poellinger, L., and Pongratz, I. (2001). The hsp90 chaperone complex regulates intracellular localization of the dioxin receptor. Mol Cell Biol *21*, 2594-2607.

Keeney, D.S., and Waterman, M.R. (1993). Regulation of steroid hydroxylase gene expression: importance to physiology and disease. Pharmacol Ther *58*, 301-317.

Knutson, J.C., and Poland, A. (1982). Response of murine epidermis to 2,3,7,8-tetrachlorodibenzo-p-dioxin: interaction of the ah and hr loci. Cell *30*, 225-234.

Kuzhandaivelu, N., Cong, Y.S., Inouye, C., Yang, W.M., and Seto, E. (1996). XAP2, a novel hepatitis B virus X-associated protein that inhibits X transactivation. Nucleic Acids Res *24*, 4741-4750.

LaPres, J.J., Glover, E., Dunham, E.E., Bunger, M.K., and Bradfield, C.A. (2000). ARA9

modifies agonist signaling through an increase in cytosolic aryl hydrocarbon receptor. Journal of Biological Chemistry *275*, 6153-6159.

Lash, L.H. (1994). Role of renal metabolism in risk to toxic chemicals. Environ Health Perspect *102 Suppl 11*, 75-79.

Lees, M.J., and Whitelaw, M.L. (2002). Effect of ARA9 on dioxin receptor mediated transcription. Toxicology *181-182*, 143-146.

Ma, Q., and Baldwin, K.T. (2000). 2,3,7,8-tetrachlorodibenzo-p-dioxin-induced degradation of aryl hydrocarbon receptor (AhR) by the ubiquitin-proteasome pathway. Role of the transcription activaton and DNA binding of AhR. J Biol Chem *275*, 8432-8438.

Ma, Q., and Whitlock, J.P., Jr. (1997). A novel cytoplasmic protein that interacts with the Ah receptor, contains tetratricopeptide repeat motifs, and augments the transcriptional response to 2,3,7,8-tetrachlorodibenzo-p-dioxin. J Biol Chem *272*, 8878-8884.

McGuire, J., Whitelaw, M.L., Pongratz, I., Gustafsson, J.A., and Poellinger, L. (1994). A cellular factor stimulates ligand-dependent release of hsp90 from the basic helix-loop-helix dioxin receptor. Mol Cell Biol *14*, 2438-2446.

McMillan, B.J., and Bradfield, C.A. (2007). The aryl hydrocarbon receptor is activated by modified low-density lipoprotein. Proc Natl Acad Sci U S A *104*, 1412-1417.

Miyake, R., Kawamoto, J., Wei, Y.L., Kitagawa, M., Kato, I., Kurihara, T., and Esaki, N. (2007). Construction of a low-temperature protein expression system using a cold-adapted bacterium, Shewanella sp strain Ac10, as the host. Applied and Environmental Microbiology *73*, 4849-4856.

Mujacic, M., Cooper, K.W., and Baneyx, F. (1999). Cold-inducible cloning vectors for low-temperature protein expression in Escherichia coli: application to the production of a toxic and proteolytically sensitive fusion protein. Gene *238*, 325-332.

Murray, I.A., Morales, J.L., Flaveny, C.A., Dinatale, B.C., Chiaro, C.R., Gowdahalli, K., Amin, S.G., and Perdew, G.H. (2009). Evidence for ligand-mediated selective modulation of aryl hydrocarbon receptor activity. Mol Pharmacol.

Nebert, D.W., and Jensen, N.M. (1979). The Ah locus: genetic regulation of the metabolism of carcinogens, drugs, and other environmental chemicals by cytochrome P-450-mediated monooxygenases. CRC Crit Rev Biochem *6*, 401-437.

Nebert, D.W., and Jensen, N.M. (1979). Benzo[a]pyrene-initiated leukemia in mice. Association with allelic differences at the Ah locus. Biochem Pharmacol *28*, 149-151.

Pandini, A., Denison, M.S., Song, Y., Soshilov, A.A., and Bonati, L. (2007). Structural and functional characterization of the aryl hydrocarbon receptor ligand binding domain by homology modeling and mutational analysis. Biochemistry *46*, 696-708.

Perdew, G.H. (1988). Association of the Ah receptor with the 90-kDa heat shock protein. J Biol Chem *263*, 13802-13805.

Peti, W., and Page, R. (2007). Strategies to maximize heterologous protein expression in Escherichia coli with minimal cost. Protein Expr Purif *51*, 1-10.

Petrulis, J.R., Hord, N.G., and Perdew, G.H. (2000). Subcellular localization of the aryl hydrocarbon receptor is modulated by the immunophilin homolog hepatitis B virus X-associated protein 2. J Biol Chem *275*, 37448-37453.

Petrulis, J.R., Kusnadi, A., Ramadoss, P., Hollingshead, B., and Perdew, G.H. (2003). The hsp90 Co-chaperone XAP2 alters importin beta recognition of the bipartite nuclear localization signal of the Ah receptor and represses transcriptional activity. J Biol Chem *278*, 2677-2685.

Petrulis, J.R., and Perdew, G.H. (2002). The role of chaperone proteins in the aryl hydrocarbon receptor core complex. Chem Biol Interact *141*, 25-40.

Poland, A., and Glover, E. (1979). An estimate of the maximum in vivo covalent binding of 2,3,7,8-tetrachlorodibenzo-p-dioxin to rat liver protein, ribosomal RNA, and DNA. Cancer Res *39*, 3341-3344.

Poland, A., Glover, E., Ebetino, F.H., and Kende, A.S. (1986). Photoaffinity labeling of the Ah receptor. J Biol Chem *261*, 6352-6365.

Poland, A., Glover, E., and Kende, A.S. (1976). Stereospecific, high affinity binding of 2,3,7,8-tetrachlorodibenzo-p-dioxin by hepatic cytosol. Evidence that the binding species is receptor for induction of aryl hydrocarbon hydroxylase. J Biol Chem *251*, 4936-4946.

Poland, A., Glover, E., and Taylor, B.A. (1987). The murine Ah locus: a new allele and mapping to chromosome 12. Mol Pharmacol *32*, 471-478.

Poland, A., and Knutson, J.C. (1982). 2,3,7,8-tetrachlorodibenzo-p-dioxin and related halogenated aromatic hydrocarbons: examination of the mechanism of toxicity. Annu Rev Pharmacol Toxicol *22*, 517-554.

Pollenz, R.S., Sattler, C.A., and Poland, A. (1994). The aryl hydrocarbon receptor and aryl hydrocarbon receptor nuclear translocator protein show distinct subcellular localizations in Hepa 1c1c7 cells by immunofluorescence microscopy. Mol Pharmacol *45*, 428-438.

Pongratz, A., Schwarzkopf, A., Hahn, H., Heesemann, J., Karch, H., and Doll, W. (1994). [The effect of the pipe material of the drinking water system on the frequency of Legionella in a hospital]. Zentralbl Hyg Umweltmed *195*, 483-488.

Pratt, W.B. (1997). The role of the hsp90-based chaperone system in signal transduction by nuclear receptors and receptors signaling via MAP kinase. Annu Rev Pharmacol Toxicol *37*, 297-326.

Pratt, W.B., Galigniana, M.D., Harrell, J.M., and DeFranco, D.B. (2004). Role of hsp90 and the hsp90-binding immunophilins in signalling protein movement. Cell Signal *16*, 857-872.

Pratt, W.B., Galigniana, M.D., Morishima, Y., and Murphy, P.J. (2004). Role of molecular chaperones in steroid receptor action. Essays Biochem *40*, 41-58.

Pratt, W.B., and Toft, D.O. (1997). Steroid receptor interactions with heat shock protein and immunophilin chaperones. Endocr Rev *18*, 306-360.

Safe, S. (1990). Polychlorinated biphenyls (PCBs), dibenzo-p-dioxins (PCDDs), dibenzofurans (PCDFs), and related compounds: environmental and mechanistic considerations which support the development of toxic equivalency factors (TEFs). Crit Rev Toxicol *21*, 51-88.

Schein, C.H., and Noteborn, M.H.M. (1988). Formation of Soluble Recombinant Proteins in Escherichia-Coli Is Favored by Lower Growth Temperature. Bio-Technology *6*, 291-294.

Schmidt, J.V., Su, G.H.T., Reddy, J.K., Simon, M.C., and Bradfield, C.A. (1996). Characterization of a murine Ahr null allele: Involvement of the Ah receptor in hepatic growth and development. Proceedings Of The National Academy Of Sciences Of The United States Of America *93*, 6731-6736.

Seidel, S.D., Winters, G.M., Rogers, W.J., Ziccardi, M.H., Li, V., Keser, B., and Denison, M.S. (2001). Activation of the Ah receptor signaling pathway by prostaglandins. J Biochem Mol Toxicol *15*, 187-196.

Shaikh-Omar, O.A. (2006). RECOMBINANT EXPRESSION OF THE ARYL HYDROCARBON RECEPTOR In School of Biology (Nottingham, University of NottinghamSchool of Biology).

Shortle, D. (1997). Structure prediction: folding proteins by pattern recognition. Curr Biol *7*, R151-154.

Sinal, C.J., and Bend, J.R. (1997). Aryl hydrocarbon receptor-dependent induction of cyp1a1 by bilirubin in mouse hepatoma hepa 1c1c7 cells. Mol Pharmacol *52*, 590-599.

Sorensen, A.G., Buonanno, F.S., Gonzalez, R.G., Schwamm, L.H., Lev, M.H., Huang-Hellinger, F.R., Reese, T.G., Weisskoff, R.M., Davis, T.L., Suwanwela, N.*, et al.* (1996). Hyperacute stroke: evaluation with combined multisection diffusion-weighted and hemodynamically weighted echo-planar MR imaging. Radiology *199*, 391-401.

Swanson, H.I., Tullis, K., and Denison, M.S. (1993). Binding of transformed Ah receptor complex to a dioxin responsive transcriptional enhancer: evidence for two distinct heteromeric DNA-binding forms. Biochemistry *32*, 12841-12849.

Uhlen, M., Nilsson, B., Guss, B., Lindberg, M., Gatenbeck, S., and Philipson, L. (1983). Gene fusion vectors based on the gene for staphylococcal protein A. Gene *23*, 369-378.

White, R.D., Hahn, M.E., Lockhart, W.L., and Stegeman, J.J. (1994). Catalytic and immunochemical characterization of hepatic microsomal cytochromes P450 in beluga whale (Delphinapterus leucas). Toxicol Appl Pharmacol *126*, 45-57.

Whitelaw, M.L., Gustafsson, J.A., and Poellinger, L. (1994). Identification of transactivation and repression functions of the dioxin receptor and its basic helix-loop-helix/PAS partner factor Arnt: inducible versus constitutive modes of regulation. Mol Cell Biol *14*, 8343-8355.

Whitelaw, M.L., McGuire, J., Picard, D., Gustafsson, J.A., and Poellinger, L. (1995). Heat shock protein hsp90 regulates dioxin receptor function in vivo. Proc Natl Acad Sci U S A *92*, 4437-4441.

Whitlock, J.P., Okino, S.T., Dong, L.Q., Ko, H.S.P., ClarkeKatzenberg, R., Qiang, M., and Li, H. (1996). Cytochromes P450 .5. Induction of cytochrome P4501A1: A model for analyzing mammalian gene transcription. Faseb Journal *10*, 809-818.

Young, J.C., and Hartl, F.U. (2000). Polypeptide release by Hsp90 involves ATP hydrolysis and is enhanced by the co-chaperone p23. EMBO J *19*, 5930-5940.

C

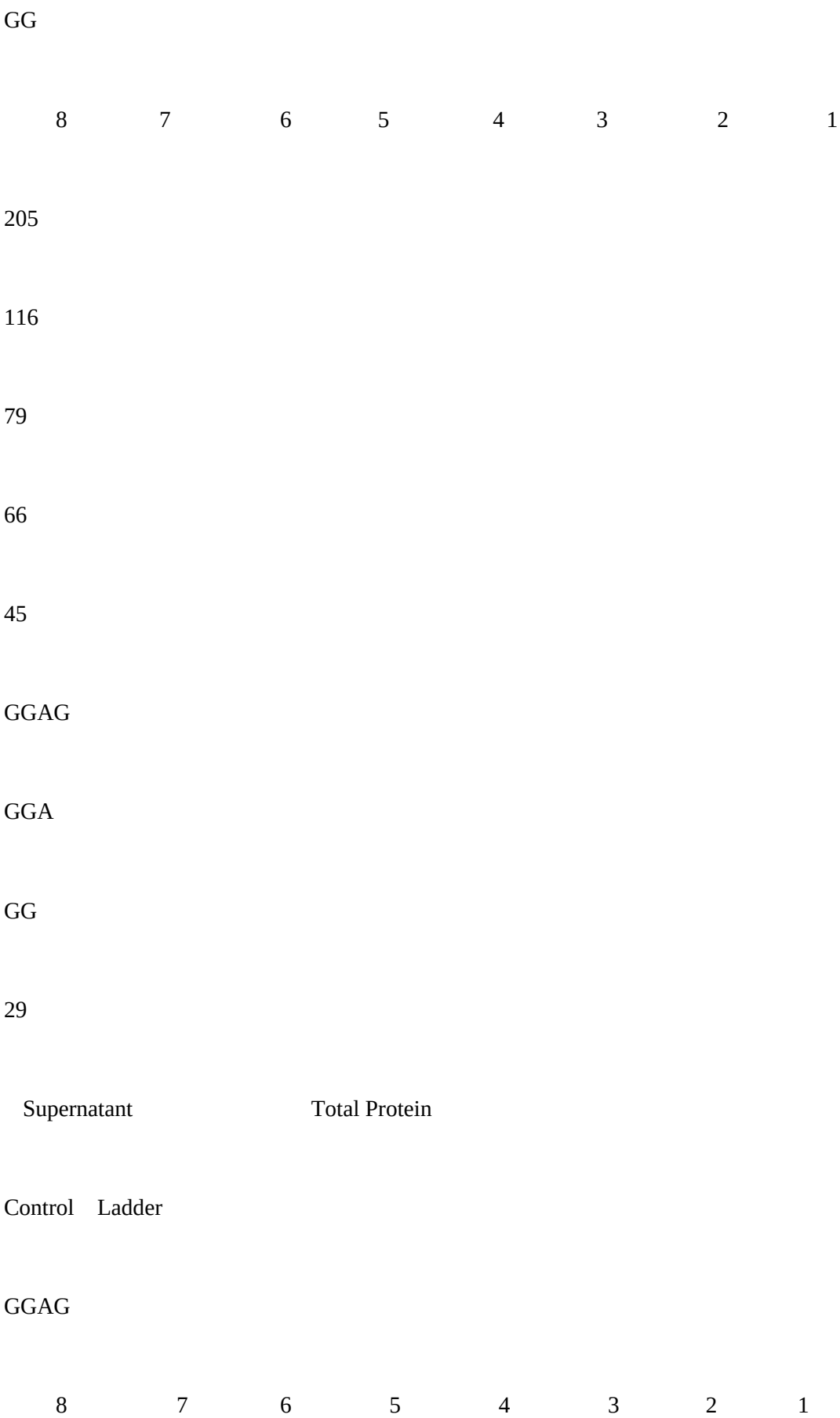
B

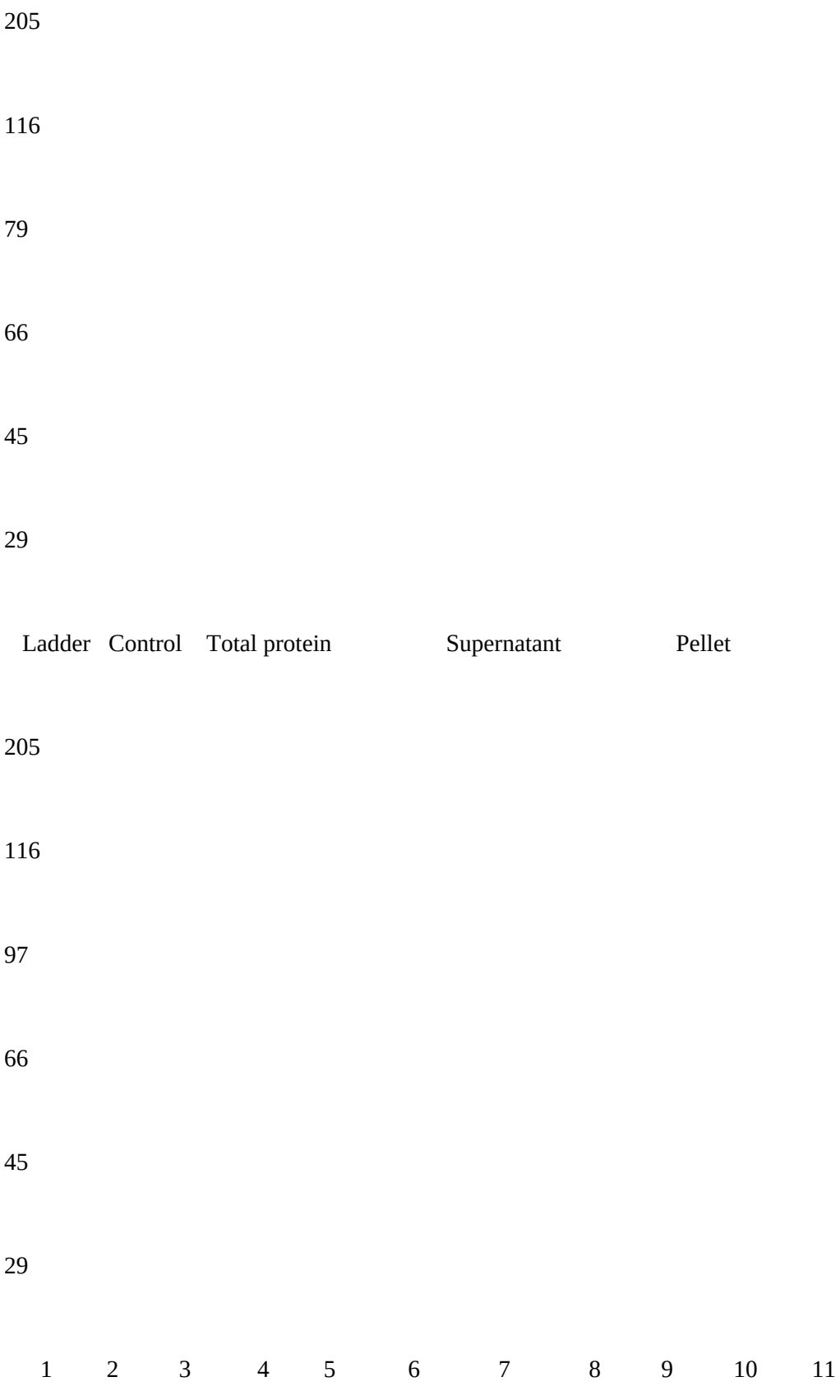
A

D

GGAG

GGA


GG
8 7 6 5 4 3 2 1
205
116
79
66
45
GGAG
GGA
GG
29
Supernatant
Total Protein
Control Ladder
GGAG
8 7 6 5 4 3 2 1


205
116
79
66
45
29
Ladder
Control
Total protein
Supernatant
Pellet
205
116
97
66
45
29
1
2
3
4
5
6
7
8
9
10
11
Control

Total protein

Supernatant

Ladder Control

205

166

97

66

45

29

1 2 3 4 5 6 7 8 9

Ladder

purified protein

unpurified protein

7 6 5 4 3 2 1

GG

GGA

205

166

97

66

45

29

GGAGG

116

97

66

1 2 3 4