Molecular Biology II: Protein function

Dr Nick Plant

Centre for Toxicology, Faculty of Health and Medical Sciences, University of Surrey, Guildford,

Surrey, GU2 7XH, UK

Email: N.Plant@Surrey.ac.uk

Nick Plant is a Reader in Molecular Toxicology within the Centre for Toxicology at the University of

Surrey, UK. His research interest include the use of systems modelling to understand the role of

nuclear receptors in co-ordinating body responses to therapeutic drugs

Words (Total): 2528

Words (Abstract): 173

Figures: 2 (300 word equivalent)

Tables: 0

Keywords

Co-localisation; GFP; FRET; RNAi; transgenic animals; ES cells

Conflicts of Interest

None declared

Abstract

DNA may be viewed as the blueprint for the body, with mRNA/protein the component parts made from this blueprint. However, it is only when proteins are allowed to interact with each other and their surroundings that true biological complexity is achieved. Thus, while it is informative to study transcriptional control and mRNA transcript levels, it is equally important to assess the impact of the encoded proteins on the total cellular environment. For example, expression of a ligand-activated receptor such as the epidermal growth factor receptor is of no biological consequence if no EGF is present in the system. It is thus important to be able to study protein interactions and modification, enabling a comprehensive understanding of the biological mechanisms that underlie any particular phenotype. This article will outline the basic technologies to both visualise protein localisation and interaction between co-localised proteins. In addition, the manipulation of protein levels, both in vitro and in vivo, will be described, as this provides an important tool for the further examination of protein functionality within biological systems.

Protein Quantification

The initial step in any investigation of protein expression is to be able to accurately determine the level of expression of a target protein. This is particularly important as mRNA levels do not always correlate with protein levels and hence measurement at the mRNA level alone may be misleading (1). The simplest way of measuring protein expression is through an immunoassay (2). Assay plates can be purchased that have antibodies specific to your target protein incorporated into each well. Initially, the wells are 'blocked' with an albumin mixture (usually powdered milk) to prevent non-specific binding to the antibody in the next stage. Following blocking, protein samples are added to the wells, incubated to allow interaction between antibody and protein, and then several washes are undertaken to remove non-specifically bound protein. Finally, the antibody is detected, usually through the use of colourimetric or fluorometric substrates to enzymes linked a second, detection antibody.

An alternative to this procedure is the Western blot (3): In this technique samples are separated according to size by electrophoresis through a polyacrylamide gel (PAGE), with smaller fragments moving faster and hence further in a given time. As proteins can adopt three dimensional structures, which could impact on their migration rate through the gel, it is usual to add the surfactant sodium dodecyl sulphate (SDS) to the gel and hence the procedure is often referred to as SDS-PAGE. Once the samples have been separated by size it is necessary to identify the protein of interest from within the multitude of proteins separated on the gel. To achieve this, proteins are first transferred from the polyacrylamide gel to a more robust support, usually nitrocellulose or PVDF membrane. Next, the membrane is treated with blocking agent, antibody, washed and specific binding detected in a similar manner to that seen with immunoassays. This procedure will result in a band corresponding to your target protein, with the intensity of the band being correlated to the amount of protein present.

For both immunoassay and Western blotting, it is important to ensure even loading of samples into the wells, for if this is not achieved then differential band strengths may be observed as an artefact. Western blots are 'stripped' using the surfactants SDS and tween-20 and then reprobed with an antibody for a general protein such as actin; this can then be used to normalise numbers for each lane and ensure a robust quantification (4).

An advantage of Western blotting over immunoassays is the ability to include a recombinantly-expressed version of the target protein in one lane of the gel, thus providing proof that the detected band is specific for the target proteins; this is important given the questionable specificity of some antibodies. It should also be remembered that this disadvantage of reduced specificity must be weighed against the increased biological relevance of the answer. This is particularly true when you consider that phospho-specific antibodies are now available for many proteins; as phosphorylation is a central post-translational mechanism for switching proteins on and off this means that it is possible to not only measure the level of a protein but also distinguish between active and inactive forms.

Regardless of the specificity of the antibody used, one important limitation of Western blotting is that it can only really be seen as semi-quantitative, looking for the relative expression of a target in different samples, for example (5). For true quantification of protein species it is necessary to turn to methods that combine chromatography with mass spectrometry of NMR, which are both more sensitive/specific and able to provide absolute expression levels for a protein (6). This enhanced robustness comes at a price, however; MS-based quantification technologies are both more technologically challenging and more expensive (7). It is thus important to decide what level of information is required. If we wish to determine if a particular protein is over-expressed in a tumour, then a semi-quantitative relative measure will suffice, but for complex pharmacokinetic/dynamic simulations of drug fate in the body, absolute protein levels are preferable (6).

Examining Protein:Protein Interactions

It is often necessary for two, or more, proteins to interact before they can exert their biological effect. Such effects include the interaction of dimerisation partners prior to DNA binding (8), the interaction of proteins in signalling cascades (9), or interaction of proteins to permit sub-cellular localisation (10, 11). It is thus necessary to examine the interaction of proteins to fully understand their biological functioning, and a large number of techniques exist for this purpose. Herein, we will consider some of the more commonly used approaches, detailing their pros and cons.

Perhaps the simplest technology is co-localisation, which works on the premise of guilt-by-association. In this approach, the sub-cellular localisation of two proteins is determined, either using an antibody-based technique similar to Western blotting but undertaken on whole cells (immunocytochemistry). Alternatively, molecular cloning can be used to create recombinant proteins fused to fluorophores, meaning these proteins will emit light of a characteristic wavelength within cells, allowing their localisation. If the signal for two proteins occurs within the same region of the cell, it can be assumed that they are in close proximity and may well be interacting (3, 12).

In a similar vein, it is possible to look at movement of proteins through cell. Altered sub-cellular localisation is often an important part of the cells response to chemical challenge, allowing new protein:protein and protein:DNA interactions to form. Through the use of flurogenic fusion proteins it is possible to watch the movement of proteins within a cell in real time. For example, figure 2 shows the movement of a glucocorticoid receptor-GFP fusion protein from the cytoplasm of the cell to the nucleus following addition of its ligand (11). This movement is the first step in the activation of gene transcription, resulting in the cellular response to heightened glucocorticoid levels.

However, this is not proof positive and further steps are required to prove interaction. Whereas co-localisation experiments show that proteins are within micrometers (10^{-6}m), fluorescence resonance enegy transfer (FRET) determines co-localisation within angstroms (10^{-10}m). At such

distances it is highly likely that the two proteins are interacting. FRET uses a similar approach to co-localisation in that both proteins are tagged with a fluorophore, but in this instance the fluorophores are chosen such that the emission spectrum of one fluorophore (A) overlaps the excitation spectrum of the other (B). If the proteins localise separately within the cell then excitation of fluorophore A will result in the characteristic light emission of A, and the same for fluorophore B. However, if the proteins are within 100Å, which almost certainly means that they are interacting, then excitation of flurophore A will result in emission of light characteristic of fluorophore B as the light energy is transferred (3).

Finally, interaction can be positively proven through co-precipitation assays: This assay is very similar to the chromatin immunoprecipitation (ChIP) assay used to identify protein:DNA interactions, and described in the preceding article. However, in this case the precipitation of one protein with a specific antibody, is used to co-precipitate any interacting proteins as well (3). For such co-precipitation to occur an interaction between the two proteins must exist.

Altering protein expression levels

The examination of protein levels by immunoassay, Western blotting or LC-MS, coupled with interaction assays in the cellular context (co-localisation, FRET and co-immunoprecipitation) lend great support to understanding the roles of proteins within the cell. However, a useful technology to further explore the functionality of proteins is the ability to artificially alter their expression levels. Over-expressing or knock-down of protein expression can help determine which processes in a cell a particular protein is involved in. To achieve this manipulation different, but related, technologies are used depending on whether you wish to alter protein expression in cell lines (in vitro) or in whole animals (in vivo).

Altering Protein levels in vitro

To increase the level of a protein in a cell line is a relatively simple task. First, the DNA encoding the protein of interest is cloned into an expression plasmid, a circular DNA that can be transfected into the cell line of choice for expression. This plasmid will then produce the protein encoded by the DNA, either constitutively or following the addition of stimulating chemical. It should be noted that, in general, the size of mammalian genes is such that the DNA cloned cannot be the genomic sequence (i.e. including exons and introns), but is generally the complimentary DNA (cDNA) produced from the mRNA (i.e. containing only the protein coding exons).

Following the transfection of this expression plasmid, Western blotting is commonly used to demonstrate that the desired protein has indeed been overexpressed. Following this important control, experiments can be undertaken to examine the effect of this over-expression on any particular biological process, showing how the target protein impacts upon it (13). Knock down of protein expression is achieved through a similar methodology, except that instead of transfecting an expression plasmid into the cell line we transfect small interfering RNAs (siRNA): siRNAi mimics naturally occurring micro RNAs, which cause mRNA degradation and are used by the body to regulate mRNA transcript pool size (14). Thus, transfection of a specifi siRNA against our target will lead to the degradation of the mRNA transcript pool for the target protein and ultimately a reduction in protein levels. However, it is important to note that as protein levels will persist for some time after the mRNA transcript pool has been depleted, and that depletion is seldom 100%, it is important that Western blotting is used to show that suppression of protein levels has truly been achieved (4).

Altering Protein levels in vivo

Alteration of protein levels in vivo is considerably more complex than in vitro. Although it is possible to use over-expression plasmids and RNAi as described previously these approaches are technically demanding, and seldom 100% successful. By far and away the more common approach is the production of transgenic animals, predominantly mice. The reason for mice being the animal of

choice for mammalian transgenics reflects their ease of husbandry, relatively fast generation time and, most importantly, ability to harvest embryonic stem (ES) cells. ES cells are derived from the inner cell mass of a blastocyst and are the totipotent cells that go on to form the foetus. These cells can be easily isolated from mice and if grown on a differentiation inhibiting medium can be cultured almost indefinitely. These cells are then manipulated to contain the desired genetic material (described in the next paragraph), including a gene encoding antibiotic resistance allowing selection of those ES cells where integration has been successful. Following selection of successful integrants in vitro, these cells can be injected back into the core of a second blastocyst, which in turn can be implanted into a pseudopregnant female mouse. The resultant pups will be chimeras, derived from both the inserted, modified, ES cells and those ES cells that were already present in the blastocyst. Further breeding of the chimaeric transgenic pups with wild type mice will result in fully transgenic mice, where every cell contains the transgene; if this is present on only one sister chromosome the mice are heterozygous, but if both sister chromosomes contain the transgene then the animal is homozygous (3).

The process for inserting a transgene into genomic DNA is relatively simple, and can be carried out in a non-targeted or targeted fashion. For non-targeted integration the exact location that the transgene inserts into the genomic DNA is not controlled. This has the advantage that no prior knowledge is required about the integration site and is common for the integration of plasmids to over-express a protein (knock-in). However, the lack of control over the integration site can lead to problems if, for example, the transgene integrates into a region of compressed chromatin (heterochromatin), leading to reduced expression of the transgenic protein. Alternatively, the transgene could integrate in to another gene, disrupting its expression in the process, and thus producing a phenotype that is difficult to interpret (3).

By comparison, targeted integration provides total control over the integration site, and has the added advantage that this now allows disruption of protein expression from genes already present

in the genome (knock-out). For targeted integration the targeting plasmid is designed such that it includes two regions of DNA that exactly match the genomic DNA; these then act as the points for homologous recombination and integration to occur (Figure 2). To create a knock-out it is simply a matter of designing the homologous regions so that they interrupt the genomic copy of the target gene (Figure 2A)(15). Alternatively, the sequence for a gene you wish to over-express can be included between these homolgous regions, producing a knock-in (Figure 2B). In addition to an antibiotic resistance gene for integration, the targeting plasmid usually includes a second marker, thymidine kinase (TK). As the TK gene is outside of the region subject to homologous recombination then if targeting is successful it will not be included in to the genomic DNA; if however, the targeting plasmid integrates randomly then the TK will also integrate. A simple TK assay can hence be used to determine the correct targeting and integration of the construct into the genomic DNA (Figure 2B).

Conclusions

The ability to study protein function within a living cell, both in vitro and in vivo, is central to our ability to understand how the body functions. Such studies will allow us to understand the biological interactions that occur within cells and how these alter under different stimuli. Finally, the ability to alter expression levels of proteins, either increasing or decreasing them, adds the tantalizing prospect of being able to recreate disease conditions, thus providing further insights into the pathogenesis of disease and novel potential therapies.

References

1.   Greenbaum D, Colangelo C, Williams K, Gerstein M. Comparing protein abundance and mRNA expression levels on a genomic scale. Genome Biol. 2003;4(9):117. .

2.   Mattiasson B, Teeparuksapun K, Hedstrom M. Immunochemical binding assays for detection and quantification of trace impurities in biotechnological production. Trends in Biotechnology. [Review]. 2010 Jan;28(1):20-7.

3.   Dale JW, von Schantz M, Plant N. From Genes to Genomes: Concepts and Applications of DNA Technology. 3 ed. Oxford: Wiley-Blackwell; 2011.

4.   Plant KE, Anderson E, Simecek N, Brown R, Forster S, Spinks J, et al. The neuroprotective action of the mood stabilizing drugs lithium chloride and sodium valproate is mediated through the up-regulation of the homeodomain protein Six1. Toxicol Appl Pharm. 2009;235:124-34.

5.   Howe K, Gibson GG, Coleman T, Plant N. In silico and in vitro modelling of hepatocyte drug transport processes: Importance of ABCC2 expression levels in the disposition of carboxydichlorofluroscein. Drug Met Disp. 2009;37(2):391-9.

6.   Li N, Singh P, Mandrell KM, Lai Y. Improved Extrapolation of Hepatobiliary Clearance from in Vitro Sandwich Cultured Rat Hepatocytes through Absolute Quantification of Hepatobiliary Transporters. Molecular Pharmaceutics. 2010 May-Jun;7(3):630-41.

7.   Ewles M, Goodwin L. Bioanalytical approaches to analyzing peptides and proteins by LC-MS/MS. Bioanalysis. [Review]. 2011 Jun;3(12):1379-97.

8.   Plant N, Aouabdi S. Nuclear receptors: the controlling force in drug metabolism of the liver? Xenobiotica. 2009;39(8):597-605.

9.   Budihardjo I, Oliver H, Lutter M, Luo X, Wang XD. Biochemical pathways of caspase activation during apoptosis. Annual Review of Cell and Developmental Biology. 1999;15:269-90.

10.   Plant KE, Everett DM, Gibson GG, Lyon J, Plant NJ. Transcriptomic and phylogenetic analysis of Kpna genes: a family of nuclear import factors modulated in xenobiotic-mediated liver growth. Pharmacogenet Genomics. 2006 Sep;16(9):647-58.

11.   Kolodkin AN, Bruggeman FJ, Plant N, Mone MJ, Bakker BM, Campbell MJ, et al. Design principles of nuclear receptor signaling: how complex networking improves signal transduction. Mol Sys Biol. 2010;6:446.

12.   Huang H, Starodub O, McIntosh A, Atshaves BP, Woldegiorgis G, Kier AB, et al. Liver fatty acid-binding protein colocalizes with peroxisome proliferator activated receptor alpha and enhances ligand distribution to nuclei of living cells. Biochemistry. 2004 Mar 9;43(9):2484-500.

13.   El-Sankary W, Plant N, Gibson G. Regulation of the CYP3A4 gene by hydrocortisone and xenobiotics: role of the glucocorticoid and pregnane X receptors. Drug Met Disp. 2000;28(5):493-6.

14.   Pasquinelli AE. MicroRNAs: deviants no longer. Trends Genet. 2002 Apr;18(4):171-3.

15.   Xie W, Barwick J, Downes M, Blumberg B, Simon C, Nelson M, et al. Humanized xenobiotic in mice expressing nuclear receptor SXR. Nature. 2000;406:435-8.

Figures
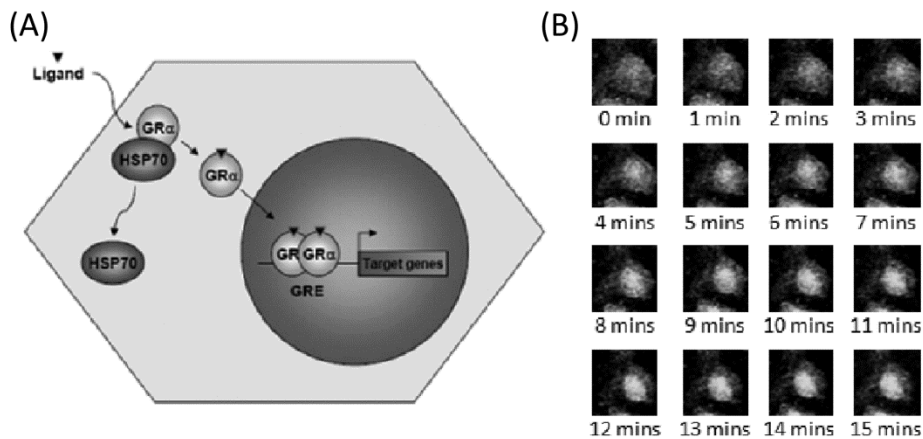
Figure 1: Visualising the translocation of the Glucocorticoid Receptor. (A) Glucocorticoids (ligand) enter cells and activate the glucocorticoids receptor (GR) in the cytoplasm. This causes disassociation from HSP70 and translocation to the nucleus where the activated GR can bind to DNA and initiate transcription of target genes. (B) A fusion protein of GR and green fluorescent protein was expressed in a human liver cell line, demonstrating even distribution of GR within the cytoplasm. Following exposure of the cells to 1μM dexamethasone (an artificial glucocorticoid), time-lapse photography demonstrates the translocation of the GR into the nucleus within 15 minutes.
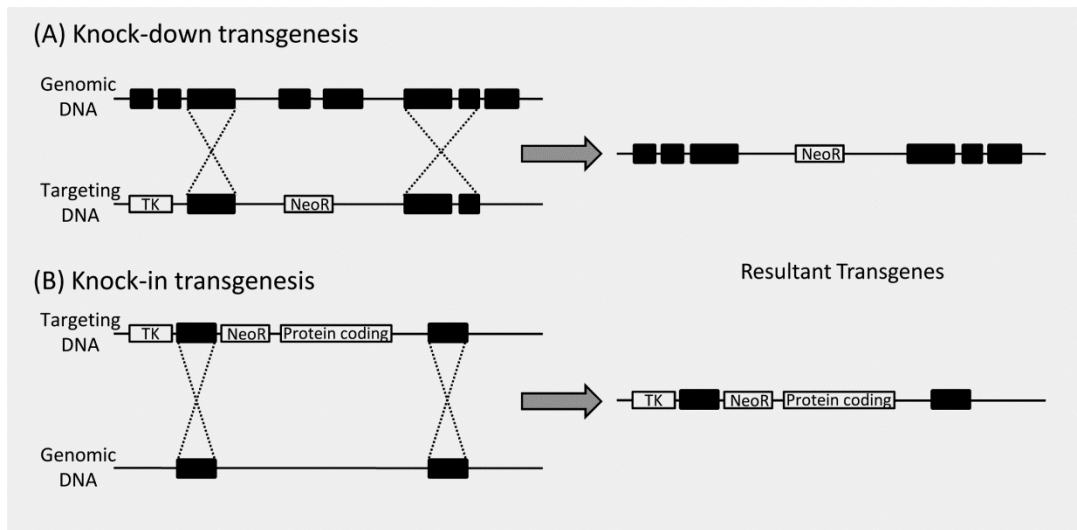
Figure 2: Construction of targeted transgenes. (A) To create a transgenic knock-out the targeting constructs are designed to match the exons within the target gene. These then undergo homologous recombination and integrate the selection marker, antibiotic resistance, into the genomic DNA. (B) To create a knock-in the targeting constructs are designed to integrate both the antibiotic resistance and coding region for the protein to be expressed into a region of DNA that will allow expression, but not disrupt any other genes. In both cases inclusion of antibiotic resistance (NeoR) and a selectable marker (TK) are used to ensure specific integration of the targeting construct.