PICHIA PASTORIS

Yeast is another traditional, powerful tool for expressing recombinant proteins and has been used successfully to express a multitude of proteins. Yeast has many of the advantageous features of E. coli such as a short doubling time and a readily manipulated genome, but also has the additional benefits of a eukaryote that includes improved folding and most posttranslational modifications. The first yeast routinely used for recombinant protein expression was Saccharomyces cerevisiae. However, in the last 15 years, P. pastoris has become the yeast of choice because it typically permits higher levels of recombinant protein expression than does S. cerevisiae. P. pastoris is a methyltropic yeast, and can use methanol as its only carbon source. The growth of P. pastoris in methanol-containing medium results in the dramatic transcriptional induction of the genes for alcohol oxidase (AOX) and dihydroxyacetone synthase. After induction, these proteins comprise up to 30% of the P. pastoris biomass. Investigators have exploited this methanol-dependent gene induction by incorporating the strong, yet tightly regulated, promoter of the alcohol oxidase I (AOX1) gene into the majority of vectors for expressing recombinant proteins. The P. pastoris expression vectors integrate in the genome whereas by contrast, S. cerevisiae vectors use the more unstable method of replicating episomally. The length of time to assess recombinant gene expression with the P. pastoris method is approximately 3–4 weeks which includes the transformation of yeast, screening the transformants for integration, and an expression timecourse. An appealing feature of P. pastoris is the extremely high cell densities achievable under appropriate culture conditions. Using inexpensive medium, the P. pastoris culture can reach 120 g/l of dry cell weight density. An important caveat is that the induction medium requires a low percentage of methanol. In large-scale cultures, the amount of methanol becomes a fire hazard requiring a new level of safety conditions.

P. pastoris has been used to obtain both intracellular and secreted recombinant proteins. Like other eukaryotes, it efficiently generates disulfide bonds and has successfully been used to express proteins containing many disulfide bonds. To facilitate secretion, the recombinant protein must be engineered to carry a signal sequence. The most commonly used signal sequence is the pre-pro sequence from S. cerevisiae a-mating factor. Because P. pastoris secretes few endogenous proteins, purification of the recombinant protein from the medium is a relatively simple task. If proteolysis of the recombinant protein is a concern, expression can be completed using the pep4 protease-deficient strain of P. pastoris selecting an Expression System 135. This strain has reduced vacuole peptidase A activity which is responsible for activation of carboxypeptidase Y and protease B1.

Yeast has the posttranslational capacity to add glycans at both specific asparagine residues (N-linked) and serine/threonine residues (O-linked). These glycan structures are substantially different from the modifications added by insect and mammalian cells. In P. pastoris the N-linked glycan is a high mannose type and usually contains 8–17 mannoses, which is quite different from S. cerevisiae structures that consist of approximately 50–150 mannose residues. Similar to insect and mammalian cells, the consensus sequence for N-linked glycans in yeast is Asn-Xaa-Ser/Thr. Two groups have completed extensive engineering to create P. pastoris strains that produce complex N-linked glycan structures comparable to those produced by mammalian cells. Wever, only the strains developed by Roland Contreras' group are available to investigators and must be licensed through Research Corporation

Technologies. The O-linked structures in P. pastoris have not been studied comprehensively but are known to be formed by the addition of one to four mannose residues to serines/threonines. Several reports have indicated that expression of certain proteins in P. pastoris resulted in the addition O-linked glycans not observed when the protein was expressed endogenously in mammalian cells.

Baculovirus/Insect Cells

Baculovirus-mediated expression in insect cells offers another useful tool for generating recombinant proteins. Baculovirus is a lytic, large (130 kb), double-stranded DNA virus, and the Autographacalifornica virus is the most commonly used baculovirus isolate for recombinant expression. Baculovirus is routinely amplified in insect cell lines derived from the fall armyworm Spodopterafrugiperda (Sf 9, Sf 21), and recombinant protein expression is completed either in the aforementioned lines or in a line derived from the cabbage looperTrichoplusiani (High-Five). Originally, creating recombinant baculoviruses involved cotransfecting the gene of interest flanked by baculovirus sequence with baculovirus DNA into insect cells, and screening for rare homologous recombination events. Recombinants were identified by screening plaques with a modified morphology, and often additional rounds of plaque screening were required to ensure that the recombinant viral preparation was not contaminated with wild-type virus. This lengthy and laborious process for generating recombinant viruses has been largely replaced by using site-specific transposition (Bac-to-Bac or BaculoDirect, Invitrogen) or an improved homologous recombination method with an engineered 136 William H. Brondykbaculovirus containing a lethal mutation in orf1629 (flashBAC from Oxford Expression Technologies or BacMagic from EMD-Novagen). Both of these approaches overcome the requirement to isolate plaques because the efficiency of recombination is 100%. Following one or two rounds of amplifying the recombinant baculovirus, the investigator can quantify the baculovirus concentration stock either by the plaque assay or by using the newer, more rapid real-time PCR or antibody-based assays. The improvements in creating and quantifying recombinant baculoviruses have dramatically reduced the time for evaluating baculovirus expression to approximately 3 weeks, including a time-course study for optimizing expression.

The most common promoters used with baculovirus expression are the polH and p10 promoters, both of which induce a high level of expression in the very late phase of the baculovirus infection. During this phase, cells undergo cell death with the concomitant release of proteases, which can result in degradation of the expressed recombinant protein. To reduce proteolysis of the recombinant protein, promoters active in earlier phases of the lytic cycle such as the basic promoter have been used. Alternatively proteolytic activity can be minimized by using constructs deleted in the chiA and v-cath genes, which encode chitinase and a cathepsin protease, respectively. Baculovirus-mediated expression is routinely used to generate both cytoplasmic and secreted recombinant proteins. Efficient secretion generally requires the presence of a signal peptide. Both insect and mammalian signal sequences can promote entry into the insect cell secretory pathway. Insect cells were originally grown in serum-containing medium which complicated purification of the secreted proteins. Recent advances in media development permit the replacement of serum with protein hydrolysates derived from either animal tissues or plants, thereby greatly simplifying protein purification. However, the high cost of this specialized media can limit its

use for large-scale bioproduction. Insect cells efficiently generate disulfide bonds in recombinant proteins. They also produce the majority of the posttranslational modifications found in mammalian cells. However, the N-linked glycan structure formed in most insect cells is the predominantly fucosylatedpaucimannose structures (Man3GlcNAc2-N-Asn). This finding has prompted the recent generation of insect cell lines that produce glycoproteins with the complex N-linked glycans normally found in mammalian cells. A transgenic Sf-9 insect line expressing several glycosyltransferases is commercially available (Mimic cell line, Invitrogen) and produces N-linked glycanscontaining a biantennary, sialylated structure. There are only a few reports describing the O-linked glycans structures generated by insect cells.

MAMMALIAN CELLS

Mammalian expression methods have conventionally been considered to be the least efficient vehicle for expressing recombinant proteins. However, recent advances have significantly improved the expression levels from mammalian cell lines. For example, stably transfected Chinese hamster ovary (CHO) cells have been reported to express recombinant antibodies up to a level of a few grams per liter. While many cell lines and expression strategies have been tested, this chapter will focus on transient transfection in human embryonic kidney (HEK293) cells and stable transfection with CHO cells.

The HEK293 cell line was derived from human embryonic kidney cells transformed with adenovirus. HEK293 cells can be transiently transfected with a high efficiency (>80%) using certain cationic lipids, calcium phosphate, or polyethyleneimine as transfection reagents. For large-scale transient transfections (>100 ml), calcium phosphate or polyethyleneimine reagents are more cost-effective options when compared to cationic lipids. Transient transfections have been performed at even the bioreactor level but for most laboratories this scale is technically challenging. The transient transfection method is relatively easy, and the evaluation for a given recombinant protein can be made in less than 2 weeks.

CHO cells are commonly used for mammalian expression when large quantities of recombinant protein are needed. For example, most therapeutic antibodies currently on the market are manufactured using this method. The standard method for stable CHO expression involves transfecting dihydrofolatereductase (DHFR)-deficient CHO cells with a DHFR selection cassette along with an expression cassette containing the gene of interest. Dihydrofolatereductase converts dihydrofolate into tetrahydrofolate which is required for the de novo synthesis of purines, certain amino acids, and thymidylic acid. Methotrexate, which binds and inhibits DHFR, is used as a selection agent and only those cells that have integrated the DHFR selection cassette will survive. Sequentially increasing the concentration of methotrexate will result in amplification of the DHFR gene along with the linked gene of interest. Following at least one round of selection with the drug methotrexate, the stably transfected pools are subcloned using limiting dilution cloning into multiwell plates. Typically only a small percentage of the screened subclones will be expressing the recombinant gene at a high level since in the majority of the clones, the expression cassette has integrated into the heterochromatin region which is transcriptionally inactive. Unfortunately, the entire selection and screening process takes at least 2–3 months, making this the major drawback of the CHO 138 William H. Brondyk method. However, recent high-throughput methods based on flow cytometry or automation have increased the ease in rapidly screening and selecting high expressing

clones. Another development has been to use specific cis-acting DNA elements flanking the recombinant gene cassette that confer active transcription to integration sites. Unfortunately, the majority of these DNA elements is owned by companies and must be licensed for use in the laboratory, and, even with the aforementioned advances in CHO expression, the timelines for generating a high expressing CHO clone have not changed considerably.

Mammalian expression systems are used primarily to generate secreted rather than intracellular recombinant proteins. Serum-free media have been developed for both the CHO and HEK293 cell lines, which simplifies the purification of secreted recombinant proteins. However, the cost of the media is quite high, making large-scale bioproduction rather costly. Mammalian cells contain the most superior folding and disulfide bond formation when compared to other expression hosts. The N-linked and O-linked glycan structures formed by mammalian cells are extremely varied and are not only dependent on the protein but also on the mammalian cell type used as the expression host. Furthermore, the cell culture conditions such as nutrient content, pH, temperature, oxygen levels and ammonia concentration can significantly affect the glycosylation profile. N-linked glycosylation can result in oligomannose, hybrid, and complex structures, and the structures all contain the Man3GlcNac2 core. The oligomannoseglycans can have two to six additional mannoses and the mannoses can be phosphorylated or sulfated. The most common complex structures have two to four Gal b1,4-GlcNac2 attached to the mannoses which result in bi-, tri-, and tetra-antennary branches. The branches can terminate with sialic acid, and fucose can also be attached to the structures. Hybrid structures contain features of both the oligomannose and complex structures. O-Glycosylation structures can be classified into eight types based on their core structures: O-GalNAc-type glycosylation, O-GlcNAc-type glycosylation, O-fucosylation, O-mannosylation, O-glucosylation, phosphoglycosylation, O-glycosaminoglycan-type glycosylation, and collagen-type glycosylation.

Proteins structure: Crystallography

X-ray crystallography is essentially a form of very high resolution microscopy. It enables us to visualize protein structures at the atomic level and enhances our understanding of protein function. Specifically we can study how proteins interact with other molecules, how they undergo conformational changes, and how they perform catalysis in the case of enzymes. Armed with this information we can design novel drugs that target a particular protein, or rationally engineer an enzyme for a specific industrial process.

In all forms of microscopy, the amount of detail or the resolution is limited by the wavelength of the electro-magnetic radiation used. With light microscopy, where the shortest wavelength is about 300 nm, one can see individual cells and sub-cellular organelles. With electron microscopy, where the wavelength may be below 10 nm, one can see detailed cellular architecture and the shapes of large protein molecules. In order to see proteins in atomic detail, we need to work with electro-magnetic radiation with a wavelength of around 0.1 nm or 1 Å, in other words we need to use X-rays.

In light microscopy, the subject is irradiated with light and causes the incident radiation to be diffracted in all directions. The diffracted beams are then collected, focused and magnified by the lenses in

the microscope to give an enlarged image of the object. The situation with electron microscopy is similar only in this case the diffracted beams are focused using magnets. Unfortunately it is not possible to physically focus an X-ray diffraction pattern, so it has to be done mathematically and this is where the computers come in. The diffraction pattern is recorded using some sort of detector which used to be X-ray sensitive film, but nowadays is usually an image plate or a charge-coupled device (CCD).

The diffraction from a single molecule would be too weak to be measurable. So we use an ordered three-dimensional array of molecules, in other words a crystal, to magnify the signal. Even a small protein crystal might contain a billion molecules. If the internal order of the crystal is poor, then the X-rays will not be diffracted to high angles or high resolution and the data will not yield a detailed structure. If the crystal is well ordered, then diffraction will be measurable at high angles or high resolution and a detailed structure should result. The X-rays are diffracted by the electrons in the structure and consequently the result of an X-ray experiment is a 3-dimensional map showing the distribution of electrons in the structure.

A crystal behaves like a three-dimensional diffraction grating, which gives rise to both constructive and destructive interference effects in the diffraction pattern, such that it appears on the detector as a series of discrete spots which are known as reflections. Each reflection contains information on all atoms in the structure and conversely each atom contributes to the intensity of each reflection. As with all forms of electro-magnetic radiation, X-rays have wave properties, in other words they have both an amplitude and a phase. In order to recombine a diffraction pattern, both of these parameters are required for each reflection. Unfortunately, only the amplitudes can be recorded experimentally all phase information is lost. This is known as "the phase problem". When crystallographers say they have solved a structure, it means that they have solved "the phase problem". In other words they have obtained phase information sufficient to enable an interpretable electron density map to be calculated.

Crystal structure determination

Firstly we need to obtain a pure sample of our target protein. We can do this by either isolating it from its source, or by cloning its gene into a high expression system. The sample then needs be assessed for suitability according to the following criteria:

1. Is it pure and homogeneous? we can test this by various electrophoretic methods and mass spectrometry .
2. Is the protein soluble and folded? if protein estimations suggest that a lot of protein is being lost, then it may be due to precipitation. The degree of ordered secondary structure can be tested with circular dichroism if this is very low then the protein may be misfolded. This may occur if the protein is being produced faster than it can fold and may result in the formation of insoluble inclusion bodies. Attenuating the induction can alleviate this problem e.g. using a lower temperature.
3. Is the sample monodisperse? in other words is the sample free from aggregation? This can be monitored using a dynamic light scattering (DLS) device.

4. Is the protein still active? check with activity assays
5. Is the sample stable? Occasionally good protein crystals will form overnight at room temperature, but usually it may take several days to one or two weeks before suitable crystals can grow. Therefore, ideally the sample needs to remain stable over that period

If the sample fails one or more of the above criteria, it may be worthwhile returning to the expression and purification protocols and trying something different, such as the addition of ligands known to interact with the protein, or adding extra purification steps. In extreme cases it may be worthwhile switching to a different expression system altogether or working with a mutated or truncated construct. It may be possible to refold protein successfully using chaotropic reagents such as urea. Aggregated or polydisperse samples may be made monodisperse by simply changing pH or adding some salt. However, without DLS, this is very difficult to assess.

Crystallization

Before beginning trials the sample needs to be concentrated and transferred to dilute buffer containing little or no salt if the protein is happy under these conditions. This can easily be achieved using centrifugal concentrators. In order to screen a reasonable number of conditions we need at least 200 m l of protein at 10 mg/ml. If this is not the case then you may need to scale up the expression and purification to make it so.

If a similar protein has already been crystallized then it is definitely worth trying the conditions used to grow crystals of this protein. In any case if you have enough material one would normally subject it to one or more sparse matrix screens. To date the total number of different conditions in our repertoire of screens comes to about 400.

We normally use these tissue culture trays to set up crystallizations with up to 24 different conditions per tray. The method used is hanging drop vapourdiffusion it has the advantage of being the least expensive on protein. The set up is as follows:

The well is prepared first and usually contains 1ml of a buffered precipitant solution such as polyethylene glycol or ammonium sulfate or even a mixture of PEG and salt. Sometimes additives are also included such as detergents or metal ions which may enhance the crystallization. Then 1 m l of the concentrated protein sample is pipetted onto a siliconized coverslip, followed by 1 m l of the well solution. The coverslip is then inverted over the well and sealed using a bead of vacuum grease. This is then left undisturbed for at least 24 hours to equilibrate. At the start of the experiment, the precipitant concentration in the drop is half that of the well. Equilibration then takes place via the vapour phase. Given the relatively large volume of the well, its concentration effectively remains the same. The drop however loses water vapour to the well until the precipitant concentration equals that of the well. Hopefully, if the conditions have been favourable, at some point during this process the protein has become supersaturated and been driven out of solution in the form of crystals. All too often however these trials result in precipitate or the formation

of salt crystals, or nothing happens at all and the drops remain clear. I would estimate that the success rate at this stage is less than 0.1%.

If no promising leads are found then there are several possible courses of action. We can add various things to the sample which may affect crystallization. We can work at a different temperature, temperaturecan have a profound affect on protein solubility. Temperatures of 4° C and 18° C are typically used. If we have already been round this cycle more than once, it may be time to go back to the purification and expression and try something different, such as working with a fragment of our target protein.

If however we are lucky enough to get one or more "hits" in the screens, then we do follow-up experiments which will be variations on a theme where the theme is the successful set of conditions. Essentially we need to refine all variables and possibly introduce some new ones in order to achieve our goal, which is large, single crystals (see below). Things to try at this stage include varying the concentrations of all components in the crystallization, slight pH changes, using additives, switching to similar buffers or precipitants, or even using different crystallization methods (e.g. dialysis). Occasionally good crystals will form overnight, but more typically they will take from several days to several weeks to grow.

X-raydiffraction

As mentioned above, X-rays are electromagnetic waves of the same nature as visible light or radio waves, the only difference being the very short wavelength of around 1 Å (Ångström, which is 10-10 meter). For comparison, the wavelength of visible light is approximately between 400 and 700 nm (one nm is 10 Å). X-rays may be generated using various laboratory sources or at synchrotrons, where very high intensity and highly focused X-rays can be generated. To obtain X-ray data from a crystal, it needs to be placed in a monochromatic (single wavelength) X-ray beam. Subsequently, it is repeatedly exposed to the X-ray beam, while changing its orientation (usually rotating). Each exposure provides an image, similar to that shown above. Each spot on the image is a diffracted X-ray beam, which emerged from the crystal and was registered by the X-ray detector. Thousands of diffraction spots need to be collected to solve a protein structure. Depending on the type of the crystal (cell dimensions and symmetry), different strategies for data collection are followed and a different amount of data is collected. Usually the crystal is rotated in the X-ray beam one degree a time, and exposed to X-rays for a short period (seconds to minutes, depending on the intensity of the X-ray source). The intensities of these spots are subsequently used to calculate the electron density of the molecules within the crystal. The electron density, in turn, will tell us where the atoms are located, information which can be used to build a model of the molecule or molecules in the crystal.

Using crystallographic terminology, this process is called **X-ray data collection**. When the X-rays

hit the crystal, a phenomenon called **X-ray diffraction** takes place. Diffraction is a common physical phenomenon and occurs when a wave (of any nature) encounters an obstacle, which can be any material object. This results in bending of the wave around that object, also called scattering of waves. Another way for diffraction to occur is when a wave encounters a small opening, a small hole or a slit. This causes spreading of the wave in all directions. In practice, in both cases, the obstacle and the hole/slit start to act as a new wave source, sending around waves with slightly different direction of propagation, as compared to the original wave. The "new" scattered waves interact with each other, resulting in another physical phenomena called interference, which translated to normal language simply means addition of waves.

X-ray diffraction is caused by the interaction of electromagnetic waves with the matter inside the crystals, and particularly with the electrons. These waves get scattered by the electrons, or each electron becomes a small X-ray source of its own. Scattered waves from all the electrons within each atom are added to each other, giving diffracted waves from each atom, etc. When the scattered waves are added, they may either get stronger or cancel each other. Those which get stronger are registered by the X-ray detector, as in the figure above. Interestingly, we do not necessarily need X-rays to observe interference, we can, for example go to a lake nearby, through two stones into the water and then observe how the waves from the two stones either reinforce eachotherorbecomeweaker.

Nuclear Magnetic Resonance Spectroscopy

NMR spectroscopy allows structure determination in solution under conditions that approximate the physiological environment of a protein. It is based on the observation of physical phenomena exhibited when nuclei absorb energy from a radio frequency source at certain characteristic frequencies in the presence of strong external magnetic fields. The position of the nuclei in the molecule effects the electronic environment of the nucleus and thus affects the absorption frequency. The frequency differences observed in the resultant spectrum can be used to determine the molecular structure of the sample. NMR has low sensitivity and the data obtained is noisy. It is used for smaller proteins.

NMR spectroscopy is one of only two techniques that can provide detailed structural information about macromolecules at atomic resolution. This detailed view of molecular structure results from a laborious examination of a number of conformationally sensitive parameters and the application of distance geometry programs to provide high-resolution structures of peptides and proteins to about 30,000 MW. However, unlike the organic chemist, who has long characterized small molecules by applying empirical “rules” associating the chemical shift with structure and conformation, the use of chemical shift as a tool to understand biological conformation has not been widely employed. The major difficulty has been a poor understanding of the link between chemical shifts and structural parameters. While the theoretical difficulties remain largely unsolved, there now exists a large body of NMR chemicalshift data for peptides and small proteins that can be used to develop empirical relationships. In an early statistical study, Szilagyi and Jardetzky identified a significant correlation between aH chemical shifts and helical and b-sheet structures. In the absence of other effects, helical conformations produce up-field shifts while b-structures

shift the a proton downfield. A smoothed plot of 1 Ha chemical shifts as a function of sequence can be used to readily identify.

Regions of secondary structure. A closely related method assigns a chemical shift index (CSI) to each residue in a protein, by comparison with a table of chemical shifts corresponding to random structure. Regions where the CSI is clustered with negative values are assigned as a-helical; those with positive values are assigned to b-structure. In contrast to optical methods for determination of protein structure, NMR provides information on the location of secondary structural elements within the protein sequence. Oldfield has even suggested that suf- ficiently accurate data may be used to predict the three-dimensional structure of the protein from chemical-shift data alone. In order to apply chemicalshift information to predictions of secondary structure, the chemical shifts must be assigned to particular residues in the protein. This is a tedious task that requires the measurement and analysis of 2-D and often 3-D spectra. In addition, the limit of about 30K for a protein that produces sufficiently narrow lines seriously hampers the general application of the method. However, the chemical-shift index serves as a useful check on further model refinement in high-resolution NMR studies of small proteins.

Ultraviolet-Visible (UV-Vis) absorption spectroscopy, Circular Dichroism (CD), Fourier Transform Infrared (FTIR) and fluorescence spectroscopy. These techniques can be utilised to study the structure of proteins as structural changes can have a major impact on their activity, stability and toxicity, and consequently can compromise the efficacy and shelf life of products.

UV-Vis can be used as a sensitive measure of subtle changes in protein structure and also to determine the protein concentration and purity of a protein solution.

CD in the far UV region (180–260nm) provides information regarding different forms of regular secondary structure found in proteins whereas the near UV region (240-360nm) can provide a detailed fingerprint of the tertiary structure. It can provide information about interaction between ligands or cofactors for e.g. DNA-protein interaction. CD can also be very useful in the comparison of batches of pharmaceuticals and we can provide some additional help with the analysis using objective pattern recognition techniques. We also provide advice and consultancy on obtaining a good quality CD spectral measurement

.

FTIR facilitates the structural analysis of proteins in different chemical environments, which makes it a valuable tool for the biotechnology and pharmaceutical industry. This technique can be utilised to analyse the structure of protein therapeutics at higher concentrations, than CD. The facilities available in house - ATR accessories specially designed for protein solutions and powders - enable the analysis of proteins in formulation buffer as well as in powder form.

Fluorescence spectroscopy can also provide tertiary structural information. Changes in the local environment of tryptophan residues can be followed by changes in the emission spectra.

Proteins are highly diversified class of biomolecules Differences in their chemical properties, such as charge shape, size and solubility, enable them to perform many biological functions. These functions include – enzyme catalysts, metabolic regulation, binding and transport of small molecules, gene regulation, immunological defense and cell structure.

The cellular activities and functions involve one or more proteins. Their central place in the cell is reflected in the fact that genetic information is ultimately expressed as proteins,

The basic building blocks of proteins are amino acids. There are about 20 amino acids found in proteins, all of which share certain structural features. These features are:

Carboxyl (acid) (-COOH) group

An amino (basic) (-NH) group

They differ from each other with respect to their side chains. Amino acids of proteins are linked together by peptide bonds between their carboxyl and –amino group to form linear polymers. Proteins have 3 or 4 levels of structural organization. and complexity. The primary structure of a protein is the sequence of amino acids in its polypeptide chain or chains. Secondary structure is formed and stabilized by the interaction of amino acids that are fairly close to one another on the polypeptide chain. The polypeptide with its primary and secondary structure can be coiled or organized along three axes to form a more complex, three dimensional shape. Thus, level of organization is the tertiary structure

A number of colorimetric and photometric methods are used for the determination of proteins. Photocolorimetric methods are based on the so called “colour” reactions for functional group of protein molecules. Among these are reactions for peptide groups and folin’s test for amino acid aromatic radicals (tyrosin and trypthophan). The biuret test is more specific since peptide bond occurs only in proteins and peptides. It is widely used in clinico-biochemical examination. The Lowry’s method, based on folin’s reaction is highly sensitive but of low specificity, since free aromatic amino acids and numerous materials containing a phenolic group produce a similar colouration. Photonephelometric methods for protein 2 concentration determination are based on the estimation of the degree of turbidity (or clouding) of a protein suspension in solution. These methods have not gained wide acceptance in practice.

Spectrophotometric methods are sub-divided into direct and indirect methods. The latter method represents a sensitive and accurate variant of the photocolorimetrically techniques. After the induction of the colour reaction of a protein, the coloured solution is measured spectrophotometrically and the protein. Concentration is estimated by the percentage of monochromatic light energy absorbed by the colour solution.

The direct method is based on the measure of light absorption by protein solution in the ultra violet spectra region at 200-220nm (characteristic absorption due to aromatic amino acid radicals, chiefly tryphotophan and tyrosine). These methods are easy to handle and require no preliminary colouration of the solution to be induced by a chromogenic agent. The 200-220nm spectrophotomery is more specific than that at 230nm. Since in the latter case, the additional absorption due to various low molecular aromatic compound, which are found in biological materials that interferes with the measurement accuracy.

The local dye “*Uri isi*” which was purchased at Nsukka market is used locally for dying grey hair. It is believed to undergo some reactions with certain chemical components of the hair in the presence of hydrogen peroxide. When applied on the hair in the presence of hydrogen peroxide, the grey colour of the hair is changed to dark colour. Preliminary screening showed that the dye reacts with proteins to produce a

change in colour. The present study attempts to design a new colorimetric method for estimation of proteins based on the colour reaction between local dye "*uriisi*" and proteins.

References

1. Guide to Protein Purification, 2nd Edition Edited by Richard R. Burgess, Murray P. Deutscher Volume 463,

2. Production of recombinant proteins in fermenter cultures of the yeast Pichia pastoris - Geoff P.Lin Cereghino, Joan Lin Cereghino, Christine Ilgen, James MCregg.

3. Chapter 13 - Expression in the Yeast Pichia pastoris - James M.Cregg, Ilya Tolstorukov, Anasua Kusari, Jay Sunga, Knut Madden, Thomas Chappell