LECTURE-10

Sample preparation for proteomics applications:

Serum and bacterial proteome

TRANSCRIPT

Welcome to the proteomics course. In today's lecture we will talk about sample preparation for various proteomic applications. I will give you some specific examples for serum and bacterial proteome analysis. In the previous lectures we were discussing about the workflow, how to make a good protein sample preparation. I described cell disruption or lysis, protection by applying different protease inhibitors during the proteolysis step, how to fractionate the samples and reduce the complexity of the proteome, how to extract and solubilise the proteins. So in the previous lecture we tried to give different types of commonly available methods for each of the steps. And as I recommended none of the methods you can directly follow in your own protocol. You have to think about your biological question which you want to ask. Depending on that you can apply combination of these methods.

So let's look back at the last lecture. In the previous lecture, I described how to disrupt the cell using different methods, how to protect the proteins from proteolysis using various protease inhibitors, sample pre-fractionation methods were discussed and then I described protein extraction and solubilisation.

In today's lecture I will talk about few precipitation methods which are commonly used for protein sample preparation. Then we will talk about the removal of interfering substances. These small substances interfere during the proteomic applications whether these are salts, nucleic acids or other types of contaminants often they make the whole experiment obsolete if you do not remove them in the beginning. So I will talk to you about some of the common interfering compounds and how one can try to get rid of them. Then we will move on to specific examples, sample preparation for serum proteome analysis and bacterial proteome analysis. So then you will realize that we

need to integrate lot of concepts which we have talked about in the previous lecture and will discuss in today's lecture. Then we need to apply that for proteome analysis of serum and bacteria. We will continue of specific examples in the following lecture and we will talk about the protein quantification.

So now, let's talk about the precipitation procedures. There are different types of precipitation methods available. I will go one by one and in the end I will give you some recommendation about which ones can be used more commonly.

So let's talk about ammonium sulfate precipitation which is one of the most previously used methods for the classical way of performing the experiments. Although it's usage is not so common when you are preparing samples for the proteomic applications but this one still remains as a good choice. Due to the high salt concentration, in ammonium sulfate precipitation, the proteins lose water in the hydration shell, they aggregate and precipitate out of the solution. If you add >50% concentration of ammonium sulfate and upto its full saturation the protein precipitation occurs which can be recovered by centrifugation.

Now let's talk about acetone precipitation. In this method many organic-solvent soluble contaminants such as detergents, lipids they are left in solution. So it is very effective. If add an excess of at least 3 or 4 volume of ice cold acetone in your extract and incubate it at -20°C for 1-2 hour to allow for protein precipitation. By performing this step the proteins can be pelleted down during the centrifugation step and then subsequently you can remove acetone and dry it out. So acetone precipitation can be performed very easily and is very effective.

Now let's talk about TCA or trichloroacetic acid precipitation. TCA is one of the very effective proteins precipitant. One can add 10-20% TCA to the sample and allow to precipitate on ice for 30-60 min. Protein pellet should be washed by adding acetone or other organic solvents such as ethanol. This method is very effective for sample recovery. Almost 99-100% sample recovery can be expected in this method.

Now since both TCA and acetone alone are very effective, people have tried to combine both the methods i.e., addition of both TCA and acetone. So this combination has demonstrated that it can precipitate the proteins more efficiently, which could not be achieved either by using TCA alone or Acetone alone. So a recommended concentration one can try is to lyse the sample in 10% TCA in acetone containing 15-20 mM DTT. Allow the protein sample to precipitate for 1-2 hours at -20°C and then the sample can be pelleted by centrifugation. The pellet can be washed with acetone alone. Try performing the whole step in cold condition so that you can avoid the protein degradation by proteolysis. Even acetone with 20 mM DTT can be used for an effective washing of the pellet. So this is a very easy method to precipitate out the protein. First add 10% TCA with acetone and after centrifugation wash the pellet to remove the TCA which could be present there and then wash 3-4 times with acetone containing DTT. After that you need to dry out your pellet so that any residual acetone is dried out from pellet.

Let's talk about one of the less commonly used methods, precipitation with ammonium acetate in methanol. This is commonly used for plant samples which are rich in polyphenol and other interfering substances. By using this precipitation method proteins are extracted in phenol and subsequently precipitated by adding 0.1 M acetate in methanol. Pellet can be washed with acetone. As I mentioned, it is less commonly used method and for specific application people try different types of precipitation and washing methods.

Let's talk about the removal of interfering substances. As I mentioned the interfering substances are very detrimental for any proteomic application. Whether you want to perform 2-DE or whether you want to go directly for LC-MS based applications or you want to do surface Plasmon resonance label-free based proteomics or you want to apply on protein microarrays. In all of these methods the interfering substances can be very detrimental. So let's talk about different types of interfering substances and how we can get rid of them. It is probable not possible to completely remove these interfering

substances but even the partial removal of those interfering substances will ensure the analysis of your further experiment.

So we need to remove the contaminants and these contaminants include salts, small ionic components, polysaccharides, nucleic acids and lipids. So if your aim is to perform 2-DE experiment, please ensure that you have removed salt very efficiently. Otherwise it is going to interfere in the isoelectric focusing step. There are types of contaminants which may also affect the quality of your proteomic experiment such as polysaccharides, lipids and nucleic acids. These components can form complexes along with the proteins by electrostatic interactions. And when you are separating the proteins using gels they can form clog on the gel. So in the gel-based proteomic methods these types of components or artifacts are going to affect the quality of your experiment very much.

Let's first talk about salts and buffers. During your entire processing you use lot of buffers and residual buffers are always there which could affect the sample preparation. Salt is present in the sample depending on the nature of the sample itself. For example, if you are talking about biological fluids such as urine, plasma or serum, these samples are already very rich in the salt content. Similarly there different types of plant cells which are quiet rich in the salt contents. So if you want to remove this salt you have to follow different types of salt removal methods. These methods dialysis, spin dialysis, gel filtration, precipitation and resolubilisation. Dialysis is one of the most commonly used methods where samples can be added to the dialysis membrane along with salt or other interfering components and in water or different buffer condition so that the salt can be eliminated out slowly. Only problem here is your sample volume can be very dilute and it can become very much. So if your application requires concentrated solutions and a small quantity then this may not be the popular choice for doing the proteomics experiment. There are other methods based on precipitation of proteins with dyes. Those are also commonly used depending upon your proteomic application.

Then there are nucleic acid contaminations. If nucleic acids are present in the protein extract they can increase the sample viscosity and later on if you are separating the proteins on 2-DE gels it is going to show background smear or different type of streaking. The high molecular weight nucleic acids such as DNA or RNA can clog the gel pores which will be used for gel-based proteomics applications. These nucleic acids can bind to the proteins through electrostatic interactions and it will interfere in the isoelectric focusing step and it may ultimately result in to severe streaking. The nucleic acid can also form complexes with the carrier ampholytes which are added during the isoelectric focusing step.

How to remove the nucleic acid contamination? To remove the nucleic acid contamination, your sample should be treated with protease-free DNase or RNase mixtures. And you can accomplish this by adding one-tenths of sample volume of a solution containing 1 mg/mL DNase, 0.25 mg/mL RNase and 50 mM $MgCl_2$. Please perform these steps in cold condition. Try to keep this reaction mixture on ice so that you are effectively performing the nucleic acid contamination removal.

So now let's talk about polysaccharide contamination. Similar to the nucleic acids, polysaccharides may also cause problems. However, the severity will be less as compared to the nucleic acid contamination. There are different types of uncharged polysaccharides such as starch and glycogen. These are very large molecules and they clog the pores of polyacrylamide matrices, similar to nucleic acids.

How to remove the polysaccharides contamination? During the precipitation step itself lot of polysaccharides get removed. TCA, ammonium sulfate or phenol/ammonium acetate precipitation are efficient ways of removing the polysaccharide contamination.

Lipids are very important and they are used to probe various biological questions. But if your aim is to study the protein you would like to get rid of lipids or any other nucleic acids or any other inferring components. Because you want to analyse proteins only. Since we are talking about sample preparation of the proteins for proteomic applications you would like to get rid of lipids. In membranous material lipids bind to specific proteins

such as lipid-carriers and lead to artefactual heterogeneity. If very low amount of protein is present in your protein sample the presence of detergents in solubisation solution should disaggregate the lipids, delipidate and solubilise the proteins.

But if your samples are very much rich in the lipid contamination, few samples which are intrinsically rich in the lipids such as plant seeds or algae. So you need to treat your samples using chemical delipidization prior to the sample solubilisation. This process of delipidation can be achieved by extraction with organic solvents containing chlorinated solvents or ethanol or acetone alone. But this step becomes very crucial if you are analysing the proteome of those biological samples which are very rich in lipid components.

Now let's talk about ionic detergents. Sodium dodecyl sulfate (SDS) is one of the ionic detergents which form very strong complexes with the proteins. We will talk about SDS and how it can be used for gel electrophoresis such as SDS-PAGE. But in this context when you are talking about protein preparation SDS is one of the very efficient compounds. However, if your aim is to perform protein separation by IEF or other gel-based methods it is going to create problems. Because SDS will bind to the proteins and will result in the negative charged complex, which will not focus unless the SDS is removed from the protein sample mixture. SDS solubilised sample can be diluted by using high concentration of non-ionic or zwitterionic detergents such as CHAPS and triton X-100 and we have talked about different types of detergents in the last lectures. So you can try different type of non-ionic or zwitterionic detergents. This step will ensure that the final SDS concentration is less than 0.25%. Otherwise your isoelectric focusing will be hampered by excess of SDS molecules.

So we have talked about different types of interfering components so far but this is not end of the list. There are many other interfering compounds present and depending upon your unique biological sample, you may be encountered with more and more interfering compounds and you have to come up with new creative ways to remove those interfering compounds so that your proteomic study can be performed with very

high quality. There are few interfering compounds which are present in the plant extracts such as lignins, polyphenols, tannins, alkaloids and pigments. I will talk about some of these in more detail when we talk in the next class about a case study where we will discuss about performing plant proteome analysis.

So far we have covered different strategies of preparing very good samples for proteomic applications. Now let's try to apply this knowledge for various specific applications. I will show these applications with serum proteome analysis and bacterial proteome analysis.

So we will take examples from human serum, bacteria and plant leaf. This gives you the diversity of biological samples, different sets of challenges being imposed by each of these unique sample types and during the process we will see how we can integrate the knowledge which we have acquired during this lecture and the previous lecture of different type of sample preparation strategies.

So now let's move on to the first application; the serum or plasma proteome analysis.

So first of all you may ask why to analyse serum or plasma for any proteomic applications. As you know that the blood proteome is one of the most complicated components of the human proteome. The liquid portion of the blood is referred to as plasma and the removal of fibrinogen and the other clotting factors from the plasma results in serum.

So human serum or plasma proteins mostly originate from a variety of tissue and blood cells as a result of secretion or leakage from the neighboring tissue or the blood cells. The rapid alteration in the expression pattern of various serum proteins due to response of disease condition or an external stimulus. It is true reflection of physiological conditions occurring in an individual. So to get the feel about all the physiological changes happening in a patient due to a disease people analyze serum or plasma proteome very often. Because blood removal for various types of tests are being performed in clinics. So blood is very easily accessible sample and performing serum or

plasma proteome becomes very easy as compared to dissecting out tissue for further analysis.

Although sample removal is easy but the serum or plasma proteome analysis is not so easy. There are major challenges in serum or plasma proteome analysis.

Let's talk about some of these challenges.

Dynamic range of the proteins concentration: In serum there is a large diversity of proteins, which provides a very dynamic environment of almost 10^{10} magnitude. So the concentration of serum proteins ranges more than 10 orders of magnitude. If you want to obtain the full spectrum of serum or plasma by applying any of the conventional proteomic techniques, it is very challenging because the typical dynamic range for any of these techniques on any these platforms will be much smaller ranging from 10^2 to 10^4. So how to capture all the dynamic events, which are happening in the serum, if the dynamic events in the serum are very large and your detection techniques are not able to capture the whole dynamic range. So to avoid these issues people try to remove some of these abundant proteins from very complex serum proteome so that overall dynamic range can be reduced and minimized.

So the second point.

High abundance proteins: There are different high abundant proteins which are present in serum and plasma which makes it's analysis very complicated. There are almost 22 highly abundant proteins present in serum, which represent about 99% of the total protein mass in plasma or serum. These high abundance proteins prevent the detection of very low abundant proteins and often these low abundant proteins are target which you are looking for as a part of biomarker discovery. So how to get rid of high abundant proteins? I will describe some of the strategies, which can be used to remove high abundant proteins in the next couple of slides. But let's first talk about what are the other challenging factors for the serum or plasma analysis.

Third point.

The presence of high salt and other interfering compounds:

As we have talked about different types of interfering compounds, salt is one of the components which could interfere a lot during your proteomic applications. Now salts are present in the bloods which are required for various functions such as maintenance of osmotic balance, acid-base balance etc. Few salts such as sodium chloride or potassium chloride are also added when you are processing the serum sample. During the whole sample manipulation, due to the intrinsic salt present in the blood as well as extrinsic salt added during the sample processing, the overall salt component becomes very high and that creates problems for various types of proteomic applications. So high salt and interfering components should be removed.

Now presence of excessive salts, detergents and other contaminants can tremendously influence the electrophoretic separation of proteins. If your target technology is 2-DE or other gel-based method you have to ensure that salt is very low in the serum or plasma. It also affects the direct determination of the proteins and peptides by mass spectrometry based techniques. So regardless of whether you use gel-based or gel-free techniques you have to ensure that the overall salt component is removed efficiently from the serum or plasma.

Now let's talk about the fourth point.

Variations among individuals and lack of reproducibility:

These are some of the major issues in clinical studies, where you have inter- and intra-individual variations. Intra-individual variations are obviously more expected but even with in one person due to the diet, medication the serum or plasma proteins can be changed. So how to avoid these inter- and intra-individual variations and if you refer to the previous lecture when we talked about how to minimize these effects of these factors which are going to ensure the success of clinical studies probably you will be able to keep an eye on different type of extrinsic and intrinsic influences which one should try to remove as much as one can during the clinical proteomic analysis. So the

drastic heterogeneity or the large biological variations such as gender, age, genetic factors, dietary considerations, environmental factors and drug treatments are going to affect the reproducibility of your experiments. So you are careful in designing the experiments one can at least minimize these variations and increase the reproducibility of the proteome analysis.

So now I am giving you an overview here for serum sample preparation. There are various steps shown in the images such as withdrawal of intravenous blood, blood collection in the tube, centrifugation step of whole blood, how to remove the serum from the whole blood, serum can be transferred in to fresh tubes. For the serum proteome analysis different types of modifications can be performed such as sonication. Sonication can be used for the disruption of high abundant proteins and it also helps in better resolution. Then depletion strategies to remove the abundant proteins, precipitation of these proteins by adding acetone and then drying out the pellet and reconstituting in suitable buffer for the proteomic applications.

So let me show you the details of serum protein sample preparation in this animation.

Serum proteome analysis:
In this interactive animation I will show you step wise procedure of collection of serum and the procedure of proteome sample preparation.

1. Blood collection

The blood proteome is one of the most complex components of the human proteome. It fluctuates depending on the physiological and pathological conditions of the patient. Blood samples (5.0 mL) collect from the antecubital vein of healthy or diseased participants using serum separation tubes.

2. Immediately after blood collection the tubes should be kept in ice for 30 minutes for clotting. After clotting, the samples are centrifuged at 1200 rpm at 20^0C for 10 minutes.

3. Serum is separated and stored in multiple aliquots at -80^{0}C.

4. Crude serum dilute five times with phosphate buffer (pH 7.4) and perform uniform mixing by vortexing the tubes for 30 sec.

5. Mild sonication is performed using a sonicator for 6 cycles of 5 sec pulse with 30 sec gap in between; at 20% amplitude. Mild sonication of serum sample is found to be effective in improving the gel quality and resolution.

6. There are commercially available depletion columns. Affinity binders (ligands) are immobilized to a solid support (i.e., chromatographic medium) and used to specifically bind abundant proteins from a complex protein solution. The depletion strategy effectively enhances the resolution and provides maximum number of spots, and thereby better coverage of the whole serum proteome. I'll describe this step in detail in lecture.

7. Depleted serum samples are mixed with ice-cold acetone containing 10% TCA and vortexed for 15 sec for uniform mixing. The mixture is incubated at -20^{0}C for 2 hrs for protein precipitation.

8. After incubation, tubes were centrifuged at 1000g for 15 min at 4^{0}C. Supernatant was separated and pellets were dissolved in lysis buffer (8 M urea, 2 M thiourea, 4% (w/v) CHAPS, 2% (v/v) IPG buffer, 40mM DTT and traces of bromophenol blue).

Store this extract in -80C until further use.

So I hope in the animation you were able to understand how to go step by step to perform the serum proteome analysis. Now as I promised previously I will talk to you about how to remove high abundant proteins, which are present in the serum.

So as we talked there are more than 20 abundant proteins, which are present in the serum and albumin alone constitutes about 50% of these abundant proteins, IgG makes upto 15-25% of the abundant proteins. Then IgA, haptoglobin, transferrin and anti-

trypsin are also present as a part of high abundant proteins in the serum. So in the diagram I have shown you the 6 high abundance serum proteins.

People have tried different ways to remove these high abundant proteins such as molecular weight cut-offs and different types of chromatography methods have been used. The affinity chromatography methods are highly efficient for the specific removal of these abundant proteins, which are present in the serum. Because there is always a fear with the molecular weight cut-offs that along with albumin and all other proteins you also might get rid off various high molecular weight proteins, which are non-abundant. So the affinity based methods ensure the specific targeting of albumin, IgG and other abundant proteins. The antibody affinity ligands are used for albumin, IgG and other abundant proteins and it results a very specific depletion. These resins can selectively bind to these proteins and the unbound proteins can be eluted in suitable buffers.

By applying the affinity resin based fractionation method and different types of depletion strategies, one can get rid of high abundant proteins as shown in this image of SDS-PAGE gel. The left lane is loaded with untreated serum and the right lane is loaded with treated serum. As you can see some of the high abundant proteins such as albumin, IgG are efficiently removed which allowed some of the low abundant proteins to appear on the gel.

Now I am showing you image of 2-DE gel, which we will discuss in the next module of 2-DE. But in these gels from left and right, you can see that the left panel is shown for crude serum and the right for depleted serum. In the crude serum albumin and other abundant proteins have resulted into very much masking of other proteins and the isoelectric focusing and the gel quality is also not so good. So streaking and different types of artifacts can be seen on this gel. Whereas, on the right and depleted serum 2-DE gel shows better pattern.

Salt removal is one of the very crucial steps. I hope by now you are convinced that it is important to remove all the salts present in your serum samples. The salts can be removed by various types of dialysis, gel filtration and precipitation. You need to ensure

that the overall salt level is less than 10 mM if your aim is to perform isoelectric focusing. There are various commercial clean-up kits available and since they avoid the dilution of samples, at the end you can obtain high concentrated samples. You can elute your sample in very less volume. These commercial methods are also very efficient. Here I have shown untreated and desalted sample gels and as you can see that there is lot of streaking in the left image where there is high amount of salt present. But the right image from the desalted sample no streaking pattern can be seen and different proteins appear on the same region of the gel.

Now let me show the laboratory demonstration of serum proteome sample preparation. By watching this video you can learn minute details on performing the serum sample preparation and one can use that for further serum proteome analysis.

Laboratory demonstration

Serum protein extraction
Serum processing:

Processing of serum involves blood and separation of serum, sonication of serum, depletion of high abundant proteins and precipitation of proteins.

Blood collection

1. Carefully withdraw around 4 mL of intravenous blood into a vacutainer tube.
2. Place the tube on ice for an hour immediately after collection to allow the blood to coagulate. Centrifuge the tube for 10 min at 2500 rpm to separate the coagulated blood. The blood cell along with the clotting factors form the pellet, while the serum containing proteins of interest forms a clear supernatant.
3. Transfer the supernatant containing serum into fresh tubes. These can be stored at -80°C until required for further use.
4. Remove the serum samples from the freezer at least 15 min before performing the experiment.
5. Place it on ice and allow it to thaw before use.

6. Then transfer the required amount of serum into fresh clean tube.

Sonication

1. Dilute the serum sample 5 times with phosphate buffer, pH 7.4.
2. Vortex the sample to allow complete and uniform mixing.
3. Place the tube on ice to provide sonication of 6 cycles of 5 sec pulses at 20% amplitude with 59 sec gap in between each cycle.

Depletion of high abundant proteins

Serum contains several proteins in a wide range of concentrations. Of these albumin and IgG are found to be most abundant. These proteins may interfere with the gel pattern in the experimental analysis while trying to detect other proteins at lower concentrations. It is, therefore, preferred to remove these high abundance proteins before electrophoresis.

1. Depletion columns for the removal of high abundant proteins are now available. Open the cap of the column and place it a collection tube and centrifuge it after the addition of phosphate buffer.
2. Discard the flow-through in the collection tube.
3. Then add the binding buffer provided in the kit to the column and centrifuge the tube.
4. Add the serum sample to the column and incubate on ice for 5 min. This allows the high abundant proteins to bind to the affinity matrix of the column while the rest of the proteins remain unbound.
5. Centrifuge the column and collect the serum at the bottom of the collection tube for further processing.

Proteins precipitation

1. Transfer the serum depleted of high abundant proteins into a fresh tube.
2. Add a mixture of trichloro-acetic acid and acetone to the depleted serum and mix the contents well. The solution gradually becomes turbid due to precipitation of proteins.

3. Place the tube at -20°C for at least 4 hours before centrifugation to enable complete precipitation of proteins.
4. Then remove the tube and centrifuge the contents to allow the proteins to settle down as a pellet.
5. Discard the supernatant and dry the pellet at room temperature.
6. Reconstitute this dried pellet with rehydration buffer.
7. Vortex the tube for a uniform mixing.
8. Store the reconstituted protein mix at -80°C until performing the next step of protein quantification.

So I hope your video demo was effective for you to demonstrate the serum proteome analysis. Now let's move on to the second example; how to perform bacterial proteome analysis.

Let's talk about the sample preparation for bacterial proteome analysis. Similar to other samples the aim is to solubilize all the proteins because if you are aiming for a proteome level analysis you would like to obtain all the proteins so that you can represent all the possible proteins on the gel map or by using a LC-MS approach. Bacterial lysis is often very challenging for different bacteria. But people have tried different types of lysis methods such as constituents of lysis buffer can be used for lysis, sonication is very efficient, enzymatic digestion methods. Sometimes a combination of these methods has been applied for bacterial lysis.

Let's talk about sonication, which is very commonly used if you are lysing the bacterial cells. Sonication is a physical method, which breaks a very complex inter- and intra-protein interactions. The ultrasonic waves generated by a sonicator lyse the cells through shear forces. One has to take care while performing this step and should be performed in cold conditions so that the heating and the foaming effects can be minimized.

So now I am giving you a pictorial overview of bacterial sample preparation. First you need to inoculate the culture and after obtaining the right colonies you need to grow the

culture to obtain the bacterial pellets. Then you need to wash the bacterial pellet and sonicate the bacterial pellets in the preferred lysis buffer and then add trizol. Then perform centrifugation and then you can see different layers of DNA, RNA and proteins. If the correct layer isn't removed properly then you may have contaminations such as phenol or interfering substances like DNA or RNA. We will talk about these in detail in the next lecture.

So in summary today we have discussed about the sample preparation workflow, which was continued from the last lecture. We talked about different types of precipitation methods, which can be used for different types of applications. Then we talked about removal of interfering components and also discussed the pros and cons of those components. Then we moved on to specific examples. I gave demonstration to perform human proteome sample preparation and analysis. Then we moved onto bacterial proteome sample preparation. We will discuss in the next lecture about the plant proteome analysis followed by quantification and once you have obtained quality and quantity of the proteins then you are ready to perform the proteomic experiment by using different types of techniques. Then we will start discussing about 2-DE, other gel-based proteomic technologies and in the subsequent lectures we will talk about mass spectrometry based methods.