Saritha Marella[1]

Antidiabetic plant proteins/peptides as complementary and alternative medicine – analytical perspectives

[1] Nanotechnology Laboratory, Institute of Frontier Technology, Regional Agricultural Research Station, Acharya N.G. Ranga Agricultural University, Tirupati 517502, Andhra Pradesh, India, E-mail: sarimarella@gmail.com

Abstract:
Many of the plant proteins are used as medicinal agents in the treatment of various diseases/disorders as they are produced by using molecular tools of biotechnology. Each protein is unique in its amino acid composition, sequence, subunit structures, size, shape, net charge, *iso electric* point, solubility, heat stability and hydrophobicity known to play a major role in the isolation and characterization procedures. The study of the protein of interest out of a large number is not possible unless it is obtained in its highly purified and intact form. Extraction, purification and characterization of proteins for different sample types are useful in determining structural, functional and other biological information in the field of pharmacy. Hence, the present review focuses on the sources, isolation, purification and characterization of natural proteins which are proven to be antidiabetic so as to commercialize these drugs (neutraceuticals) to compete with insulin, an ultimate in the treatment of diabetes mellitus.

Keywords: characterization, diabetes mellitus, neutraceuticals, proteins, purification

DOI: 10.1515/revac-2017-0025

Received: November 18, 2017; **Accepted:** August 4, 2018

Introduction

The second most prevailing metabolic disease/disorder globally is diabetes mellitus. The high blood glucose level is a consequence of either non-production of insulin by pancreas or non-perception of cells to insulin that is available. Herbal products are seeking vast support from the public due to less adverse effects compared to synthetic drugs. Herbal proteins/peptides were being reported recently to have antihyperglycemic activity and pharmacists are in trials of drug designing which can compete with insulin. Many of these act as therapeutic agents for the treatment of diabetes on and some of the proteins/peptides of herbal origin which are proven to have antidiabetic activity are listed in Table 1.

Table 1: Antidiabetic peptides and proteins isolated from medicinal plants.

Hypoglycemic agent	Source	Plant part	Properties
1. Polypeptide	*Panax ginseng* Meyer (Araliaceae)	Root	Different hypoglycemic and insulin-like principles have been isolated from the roots including various glycans, designated as panaxans A to P and ginsenoside Rb1. Some compounds are reported with antilipolytic activity: adenosine, a carboxylic acid, a peptide with a molecular weight of 1400 that can inhibit catecholamine-induced lipolysis in rat epididymal fat pads (Ng & Yeung, 1985)
2. Polypeptide	*Cystoseira barbata* P Cystoseiroceae (Rhodophyta)	Leaf	Two polypeptide fractions were isolated from C. *barbata* by gel filtration. One of these fractions had lipolytic effect, the other hypoglycemic effects [Gueven et al, 1974]

Saritha Marella is the corresponding author.

3. Polypeptide p-insulin	*Momordica charantia*\L. (Cucurbitaceae)	Fruit, Seeds	A polypeptide p-insulin was isolated from the fruit, seeds and tissue of *Momordica charantia* (bitter gourd). Amino acid analysis indicated 166 residues composed of 17 amino acids with a molecular weight of 11,000. In gerbils, langurs and humans with diabetes, p-insulin exhibited hypoglycemic activity after subcutaneous administration. In juvenile diabetic patients, the peak effect was observed within 4–8 h as compared with 2 h for bovine insulin [Khanna P et al, 1981]
4. Proteins	*Acacia melanoxylon* Wild (Leguminosae) *Bauhinia retusa* L. (Leguminosae)	Seeds	The effect of protein, isolated from seeds *A. milanoxylon* and *B. retusa*, was studied in normoglycemic and hyperglycemic rats. This principle lowered the fasting plasma glucose level in normoglycemic and hyperglycemic rats [Singh and Chandra, 1977]
5. Protein	*Momordica cymbalaria*\Hook f. (Fenzl) or M. tuberosa (Roxb)	Fruits, Seeds	The protein named Mcy isolated from the fruits and seeds of *M. cymbalaria* lowered the fasting plasma glucose levels in streptozotocin-induced diabetic rats [Rajasekhar et al, 2010]. It also exhibited good antihyperlipidemic action [Marella et al, 2015] and inhibitory actions on aldose reductase and α-glucosidase activities [Marella et al, 2016]
6. Protein	*Coix lacryma-jobi* L. var. *ma-yuen* Stapf.	Grains	Coixin, an alcohol-soluble protein belonging to the prolamine group, is the major protein in Adlay grains. It was proven to have antidiabetic and antihyperlipidemic activities [Watanabe et al, 2012]
7. Protein	*Glycine max*	Grains	Soy protein isolated from soya beans exhibited antihyperglycemic and antioxidant activities [Mendes et al, 2014]
8. Protein	*Triticum aestivum*	Bran	Gluten, a constituent of wheat bran proven to reduce the risk of type 2 diabetes [Hadjivassilious et al, 2003]
9. Protein	*Oryza sativa*	Rice bran	Consumption of rice bran protein as a functional food improved insulin resistance [Boonloh et al, 2015]

Herbal proteins known for antihyperglycemic effect.

World Health Organization also recommended the identification, characterization and development of effective therapeutic agents from folklore medicine for treatment of chronic diseases in remote areas to compensate the modern drugs. Distinction in abundance of proteins and heterogeneity in terms of physicochemical properties are the major limitations of proteome analysis.

Salting out, isoionic precipitation, etc. are some of the isolation methods, but purification of proteins is quite challenging, and, therefore, several approaches such as sodium dodecyl sulfate/polyacrylamide gel electrophoresis (SDS-PAGE), native gel electrophoresis and chromatography techniques such as gel filtration and affinity chromatography are available. Characterization of proteins can be performed by mass spectrometry [matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF)]/liquid chromatography-mass spectrometry. The amino acid sequence of a protein can be detected by using tandem mass spectrometry, Edman degradation, etc.

Heterogeneity in abundance and physicochemical properties of proteins and peptides hurdle the proteome analysis and also have a profound effect on the final outcome of protein/peptide separation. Proteomics field represents the structural and functional characterization of a specific protein as well (Martínez-Maqueda et al. 2013).

Protein extraction and identification, purification and characterization are the major issues in proteomics (Carpentier et al. 2008). Two-dimensional gel electrophoresis (2DGE), mass spectrometry and bioinformatics are the key components of current proteomics technology. 2DGE is the only method currently available which is capable of simultaneously separating thousands of proteins.

Methods of protein extraction/isolation

Difficulty in plant proteome analysis lies in a relatively low protein concentration in plant tissues compared to bacteria and animals. Besides this, protein analysis is interfered by the polysaccharides, lipids, proteolytic and oxidative enzymes and polyphenols present in cell wall and vacuole. Various extraction methods for protein have been developed, including the use of acetone, trichloroacetic acid (TCA) or phenol extraction with ammonium acetate precipitation (Ashoub et al. 2011 and Boschetti et al. 2009). High-abundant proteins in plants like ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCo) or cruciferins also make the low-abundant proteins difficult to analyze (Boschetti et al. 2009). Therefore, various fractionation techniques are necessary to enrich these proteins prior to analysis (Kim et al. 2001). Separation and isolation of proteins from a complex biological sample is a very difficult task.

Extraction buffers

For all protein extraction methods, frozen tissues were ground to fine powder in liquid nitrogen using a pre-cooled mortar and pestle. Protein fraction of interest can be obtained through different technologies imposing on cell disruption, solubilization/precipitation and enrichment systems. A lysis buffer is used that allows efficient extraction of protein without protein degradation and interference with proteins' biological activity. The nature and pH of the buffer system depend on the stability of the protein of interest with respect to pH and also the purification procedure. Table 2 represents the buffers and their pH range (normally used at concentrations of 20–50 mM).

Table 2: Buffers commonly used in protein extraction.

Buffer	pH range
Citric acid-NaOH	2.2–6.5
Sodium citrate-citric acid	3.0–6.2
Sodium acetate-acetic acid	3.6–5.6
Cacodylic acid sodium salt-HCl	5.0–7.4
MES-NaOH	5.6–6.8
Sodium dihydrogen phosphate-disodium hydrogen phosphate	5.8–8.0
Imidazole-HCl	6.2–7.8
MOPS-KOH	6.6–7.8
Triethanolamine hydrochloride – NaOH	6.8–8.8
Tris-HCl	7.0–9.0
HEPES-NaOH	7.2–8.2
Tricine-NaOH	7.6–8.6
Sodium tetraborate-boric acid	7.6–9.2
Bicine-NaOH	7.7–8.9
Glycine-NaOH	8.6–10.6

Specific pH range of various buffers for fruitful protein extraction.

In general, cell disruption leads to the release of proteolytic enzymes which could lower the overall yield, and hence a cocktail of protease inhibitors (Table 3) is to be added to control this undesirable proteolysis.

Table 3: Protease inhibitors commonly used in protein extraction processes.

Protease inhibitor	Type of protease inhibited
Aprotinin	Serine proteases
Benzamidine	Serine proteases
Ethylene glycol-bis(β-aminoethyl ether)-N,N,N′,N′-tetraacetic acid (EDTA)	Metallo proteases
Leupeptin	Cysteine and serine proteases
Phenyl methyl sulfonyl fluoride	Serine proteases
Pepstatin A	Aspartic proteases

Compounds inhibiting specific proteases that interfere in protein extraction.

Since many of these compounds are not very stable in aqueous solutions, it is important that they are added to the lysis buffer from a stock solution in an organic solvent (methanol, ethanol, isopropanol or dimethyl sulfoxide) immediately before use.

Usually, the lysis buffer consists of a dialyzable mild detergent at a low concentration, specific buffer and salt. Some of the detergents used in proteomics are mentioned in Table 4.

Table 4: Ionic and nonionic detergents used in protein extraction and purification.

Sodium dodecyl sulfate	0.1–0.5% denatures proteins (used for SDS-PAGE)
Triton™ X-100	0.1% (nonionic detergent for membrane solubilization)
NP-40	0.05–2% (nonionic detergent for membrane solubilization)
Dodecyl β D-maltoside	1%
Octyl β D-glucoside	1–1.5%

Detergents used in protein extraction processes at specific concentrations.

Tris-acetone extraction

The tissue is ground using Tris-HCl buffer (extraction buffer) – pH 8.0 containing sucrose, ascorbic acid, cysteine-HCl and polyvinyl pyrrolidone and the cell debris is removed by centrifugation at 20,800 *g* at 4°C for 15 min. The centrifugation is repeated twice at 20,800 *g* at 4°C for 60 min to collect the protein-rich supernatant. Proteins were precipitated by adding two volumes of ice-cold acetone and stored overnight at –20°C and pelleted by centrifuging samples at 20,800 *g* at 4°C for 10 min (Pavokovic, Kriznik & Krsnik-Rasol, 2012).

Proteins were also extracted using Tris buffer consisting of 50 mM Tris pH 8.0, 1% Triton X-100 and 0.1–1 mM PMSF. The tissue is ground in liquid N_2 and mixed with extraction buffer. It is left for 30 min at 4°C and then centrifuged at 25,000 *g* for 20 min. The supernatant is collected, the pellets were re-extracted and the supernatants were pooled. Proteins in the supernatant are then precipitated with 5 volumes (V/V) of 100% acetone and incubated at –20°C for 2 h. After centrifugation, the pellet is washed twice with 80% acetone and then resuspended in appropriate buffer (Marella et al. 2016).

The pellet is washed twice with acetone and subjected to centrifugation. Acetone is discarded, and pellets are dried under air at room temperature.

TCA/acetone extraction

The tissue is homogenized in ice-cold extraction buffer consisting of Tris-HCl (pH 8.8), SDS, glycerol and 2-mercaptoethanol, and the cell debris was removed by centrifuging at 500 *g* at 4°C for 15 min. The proteins in supernatant are precipitated by adding four volumes of ice-cold acetone containing trichloroacetic acid (10%), and 2-mercaptoethanol (0.07%) and the samples are stored at –20°C for at least 1 h. Pellets obtained by centrifugation at 15,000 *g* at 4°C for 45 min are washed three times with an ice-cold solution of 2-mercaptoethanol (0.07%) in water/acetone (20%) and repeated washes by centrifuging at 15,000 *g* for 15 min. Supernatants are discarded, and pellets are dried in air at room temperature (Caruso et al. 2009).

Phenol/methanol extraction

The homogenized tissue is resuspended in 3 ml of extraction buffer containing 500 mmol Tris (pH 8.0), 50 mmol ethylene diamine tetra acetic acid, 700 mmol sucrose, 100 mmol KCl and 2 mercaptoethanol (0.07%). Samples are vortexed and incubated by shaking for 10 min on ice. Equal volume of Tris-buffered phenol (pH 8.0) is added to the solutions, and then the solutions are incubated on a shaker for 10 min at room temperature. Cell debris is eliminated by centrifuging at 5500 *g* at 4°C for 10 min and the upper phenolic phase is re-extracted with 3 ml of extraction buffer. The samples are incubated on a shaker for 3 min, vortexed and centrifuged at 3200 *g* at 4°C for 20 min. The phenol phase is mixed with four volumes of precipitation buffer, consisting of 100 mmol ammonium acetate in ice-cold methanol, mixed by inversion, and the samples are incubated overnight at –20°C. Proteins are pelleted by centrifugation at 3200 *g* at 4°C for 15 min and washed three times with ice-cold precipitation buffer and finally with ice-cold acetone. After each washing step, the sample is centrifuged for 5 min at 5500 *g* at 4°C and pellets are dried in air at room temperature (Faurobert, Pelpoir & Chaïb, 2007).

Methods of quantification of proteins

The yield of proteins can be quantified by different methods at each step in proteomic analysis and also downstream processing. Table 5 depicts different quantification methods each one having its own specifications (Deutscher 1990).

Table 5: Protein quantification methods and its detectable concentration.

Method	Detectable range
Bradford method	1 μg/ml to 1.5 mg/ml
Lowry method	1 μg/ml to 1.5 mg/ml
UV spectral method	50 ng/ml to 2 mg/ml
BCA method	0.5 μg/ml to 1.5 mg/ml

Quantifying methods of protein.

Bradford method

This colorimetric protein assay is based on an absorbance shift of the dye Coomassie Brilliant Blue G-250. Under acidic conditions, the red form of the dye is converted into its bluer form, binding to the protein being assayed. The dye forms a strong, non-covalent complex with the protein's carboxyl group by Vander Waals force and amino group through electrostatic interactions (Bradford 1976). This Bradford protein assay is less susceptible to interference by various chemicals that may be present in protein samples and so more sensitive.

Lowry method

It is based on the reaction of Cu^{+}, produced by the oxidation of peptide bonds, with Folin-Ciocalteu reagent (a mixture of phosphotungstic acid and phosphomolybdic acid in the Folin-Ciocalteu reaction). The mechanism involves the reduction of the Folin-Ciocalteu reagent and oxidation of aromatic residues (mainly tryptophan, tyrosine), and the concentration of the reduced Folin reagent is measured by absorbance at 750 nm. As a result, the total concentration of protein in the sample can be deduced from the concentration of tryptophan and tyrosine residues that reduce the Folin-Ciocalteu reagent (Lowry et al. 1951).

UV spectral method

A simple spectrometer can quantify the amount of protein in a solution and the absorption of radiation in the near-UV by proteins depends on the Tyr (tyrosine) and Trp (tryptophan) content and to a very small extent on the amount of Phe (phenylalanine and disulfide bonds). Therefore, the A280 varies greatly between different proteins.

Bicinchoninic acid method

Proteins reduce alkaline Cu (II) to Cu (I) in a concentration-dependent manner. Bicinchoninic acid is a highly specific chromogenic reagent for Cu (I) forming a complex with an absorbance maximum at 562 nm. Because of this property, the resultant absorbance at 562 nm is directly proportional to the protein concentration. Bovine serum albumin is used as a protein standard.

Methods of separation and purification of proteins

Precipitation of proteins

Ammonium sulfate $(NH_4)_2SO_4$ precipitation

Several methods are available for precipitation of proteins utilizing changes in pH and temperature, or the addition of salts and organic solvents. During precipitation techniques, ammonium sulfate is added to the solution and protein samples are gathered and concentrated as they precipitate. This method is often used for bulk protein extraction. $(NH_4)_2SO_4$ is the most commonly used precipitant for salting out of proteins. Specific amounts of $(NH_4)_2SO_4$ should be added gradually in small amounts to the protein solution to achieve required saturation (20%, 40%, 60%, etc.); the concentration relevant to the percentage of saturation is represented in Table 6.

Table 6: Quantities of ammonium sulfate required to reach given degrees of saturation at 20°C.

	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100
Starting percent saturation amount of ammonium sulfate to add (g) per liter of solution at 20°C																	
0	113	144	176	208	242	277	314	351	390	430	472	516	561	608	657	708	761
5	85	115	146	179	212	246	282	319	358	397	439	481	526	572	621	671	723
10	57	86	117	149	182	216	251	287	325	364	405	447	491	537	584	634	635
15	28	58	88	119	151	185	219	255	293	331	371	413	456	501	548	596	647
20	0	29	59	39	121	154	133	223	260	293	337	373	421	465	511	559	609
25	0	29	60	91	123	157	191	228	265	304	344	386	429	475	522	571	
30	0	30	61	92	125	160	195	232	270	309	351	393	438	485	533		
35	0	30	62	94	128	163	199	236	275	316	358	402	447	495			
40	0	31	63	96	130	166	202	241	281	322	365	410	457				
45	0	31	64	98	132	169	206	245	286	329	373	419					
50	0	32	65	99	135	172	210	250	292	335	381						
55	0	33	66	101	138	175	215	256	298	343							
60	0	33	67	103	140	179	219	261	305								
65	0	34	69	105	143	183	224	267									
70	0	34	70	107	146	186	228										
75	0	35	72	110	149	190											
80	0	36	73	112	152												
85	0	37	75	114													
90	0	37	76														
95	0	38															

Ammonium sulfate saturation for protein precipitation.

The precipitated protein is removed by centrifugation at 12,000 *g* for 10 min and resuspended in 5–10 ml of suitable buffer. The solution is dialyzed in cold room against several changes of the same buffer over a 16-h period to remove residual $(NH_4)_2SO_4$, centrifuged at 12,000 *g* for 10 min to remove insoluble particulate matter and the supernatant is tested for the presence of the target protein (Nehete et al. 2013).

Values calculated according to protein purification (Scopes 1993).

Dialysis

Protein purification should be preceded by the elimination of small molecular weight substances such as reducing agents, nonreacted cross-linking or labeling reagents or preservatives that might interfere with the subsequent steps besides exchange of the protein sample into a different buffer system for downstream application. Dialysis is one method for accomplishing both contaminant removal and buffer exchange for macromolecular samples such as proteins. This is a separation technique that facilitates the removal of small, unwanted compounds from macromolecules in solution by selective and passive diffusion through a semipermeable membrane where the sample and buffer solution (called the dialysate, usually 200–500 times the volume of the sample) are placed on opposite sides.

Dialysis works by diffusion, a process that results from the thermal, random movement of molecules in solution and leads to the net movement from areas of higher to lower concentration (until equilibrium is reached). Changing the dialysate buffer removes the small molecules that are no longer in the sample and allows more contaminants to diffuse into the dialysate. Using a high buffer-to-sample volume ratio helps to maintain the concentration gradient. The number of dialysate buffer changes and the dialysis time also affect the outcome achieved in dialysis.

Centrifugation

The use of centrifugation is one of the simplest methods used for isolation and enrichment/fractionation of proteins. It forms the first step to separate different cell substructures like mitochondria, membrane or nucleus from our proteins of interest. This process involves multiple centrifugation steps where the cellular homogenate is separated into different layers based on the molecular weight, size and shape of each component. Centrifugation is also commonly used to fractionate a protein mixture into different fractions that is based on the coefficient of sedimentation of the proteins. This coefficient is usually expressed in Svedberg units (S), and the smaller the S value, the slower a molecule moves in a centrifugal field. Separation will depend on the mass, the shape and the protein density (Sharma et al. 2010).

Chromatography techniques

Chromatography is an important biophysical technique that enables the separation, identification and purification of the components of a mixture for qualitative and quantitative analysis. It separates a mixture of substances into their components on the basis of their molecular structure and molecular composition involving a stationary phase (a solid or a liquid supported on a solid) and a mobile phase (a liquid or a gas). The mobile phase flows through the stationary phase and carries the components of the mixture with it.

Sample components that display stronger interactions with the stationary phase will move more slowly through the column than components with weaker interactions which causes the separation of various components. Proteins can be purified based on characteristics such as size and shape, total charge, hydrophobic groups present on the surface and binding capacity with the stationary phase (Table 7.). Chromatography methods exploit the physical properties of either the target protein or the other elements in the solution. The dialyzed protein fraction was loaded on to the column using a specific matrix and gone for a chromatographic procedure relevant to the protein to be separated.

Table 7: Chromatographic separation techniques based on various characteristic features.

Specific characteristic	Method of separation
Molecular weight	Gel filtration chromatography
Electric charge	Ion exchange chromatography
Hydrophobic domains	Hydrophobic interaction chromatography
Bio-specific affinity	Affinity chromatography

Chromatography techniques for separation of speific proteins of interest.

Gel filtration chromatography

Gel filtration chromatography separates proteins, peptides and oligo nucleotides on the basis of size. Molecules move through a bed of porous beads, diffusing into the beads to greater or lesser degrees. Smaller molecules diffuse further into the pores of the beads and therefore move through the bed more slowly, while larger molecules enter less or not at all and thus move through the bed more quickly. Both molecular weight and three-dimensional shape contribute to the degree of retention. One to two milliliters of precipitated protein fraction is loaded on the column. The matrices generally used are Sephadex, Sepharose, Sephacryl, Superdex, Ultragel, etc.

Ion-exchange chromatography

This chromatography separates ions and polar molecules based on their affinity to the ion exchanger (cationic/anionic). The water-soluble and charged molecules such as proteins, amino acids and peptides bind to moieties which are oppositely charged by forming ionic bonds to the insoluble stationary phase (Mohammad 2012). The equilibrated stationary phase consists of an ionizable functional group where the targeted molecules of a mixture to be separated and quantified can bind while passing through the column – a cationic stationary phase is used to separate anions and an anionic stationary phase is used to separate cations.

The dialyzed fraction is applied to the column packed with an anion-exchanger, DEAE-cellulose or DEAE-Sepharose, previously equilibrated against the specific buffer. The column is connected to a Pharmacia P-1 pump and a Frac-100 fraction collector and is washed with ~60–100 ml of buffer to remove unbound proteins. The protein fraction bound to the matrix (including the target protein) is eluted with 150 ml of a linear 0–1.75 M NaCl or KCl gradient, prepared in the same buffer. Fractions rich in target protein are pooled and subjected to further studies.

Hydrophobic interaction chromatography

It is widely used for the separation and purification of proteins, in their native state, as well as for isolating protein complexes and studying protein folding and unfolding. This is initially termed as "salting out." It utilizes the reversible interaction of protein and the hydrophobic ligand for the separation of protein mixtures. Proteins separate out based on the increasing order of hydrophobicity. The protein sample is applied to a column containing a hydrophobic matrix, such as phenyl- or octyl-Sepharose, pre-equilibrated with 20 mM Tris-HCl (pH 7.5). The salt in the buffer reduces the salvation of the sample solutes and as salvation decreases, the hydrophobic regions that become exposed are adsorbed by the media. The more hydrophobic the molecule, the less salt is needed to promote binding. Usually, a decreasing salt gradient is used to elute samples from the column in order of increasing hydrophobicity. Finally, the protein sample is eluted with the original buffer-desired protein fractions combined and concentrated using Centricon filter concentrators (10 or 30 kDa cutoff).

Affinity chromatography

It is a type of chromatographic laboratory technique used for purifying biological molecules within a mixture by exploiting molecular properties. Biological macromolecules such as enzymes and other proteins interact with other molecules with high specificity through several different types of bonds and interaction. Such interactions include hydrogen bonding, ionic interaction, disulfide bridges, hydrophobic interaction and more. The high selectivity of affinity chromatography is caused by allowing the desired molecule to interact with the stationary phase and be bound within the column in order to be separated from the undesired material which will not interact and elute first (Ninfa, Ballou & Marilee, 2009). The molecules no longer needed are first washed away with a buffer, while the desired proteins are let go in the presence of the eluting solvent (of higher salt concentration). Different affinity matrices are used for separation of different proteins. The column is equilibrated with the appropriate buffer and the protein sample is loaded on to the column.

The column is packed with Cibacron Blue-agarose 3GA resin (for NAD (P)-binding proteins) or ADP-, ATP-agarose or Con A-Sepharose (containing immobilized adenosine nucleotides or Concanavalin A) can be employed for ATP/ADP-binding proteins or glycoproteins, respectively. Proteins bound to the matrix are eluted with ~80 ml of a linear 0–1.75 M NaCl gradient, fractions of the protein of interest are pooled and centrifuged to pellet insoluble proteins.

High-performance liquid chromatography (HPLC)

It is used to separate, identify and quantify each component in a mixture. The principle involves pumping a pressurized liquid solvent containing the sample mixture through a column filled with a solid adsorbent material where each component in the sample interacts slightly differently with the adsorbent material, causing different flow rates for the different components and leading to separation of the components as they flow out the column. The adsorbent in the column is typically a granular material made of solid particles (e.g., silica, polymers, etc.), 2–50 μm in size and the pressurized liquid is typically a mixture of solvents (e.g., water, acetonitrile and/or methanol), referred to as a "mobile phase." Common mobile phases used include any miscible combination of water with various organic solvents (the most common are acetonitrile and methanol), acids (such as formic, phosphoric or trifluoroacetic acid) or salts to assist in the separation of the sample components. The eluting strength of the mobile phase is reflected by analyte retention times with high eluting strength producing fast elution (short retention times). A typical gradient profile in reversed phase chromatography might start at 5% acetonitrile (in water or aqueous buffer) and progress linearly to 95% acetonitrile over 5–25 min.

There are four types of HPLC depending on the stationary phase used:

- Normal phase HPLC: this method separates analytes on the basis of polarity. NP-HPLC uses the polar stationary phase and non-polar mobile phase. Therefore, the stationary phase is usually silica and typical mo-

bile phases are hexane, methylene chloride, chloroform, diethyl ether and mixtures of these. Polar samples are thus retained on the polar surface of the column packing longer than less polar materials.

- Reverse phase HPLC: the stationary phase is non-polar (hydrophobic) in nature, while the mobile phase is a polar liquid, such as mixtures of water and methanol or acetonitrile. It works on the principle of hydrophobic interactions; hence the more non-polar the material, the longer it will be retained.
- Size-exclusion HPLC: the column is filled with material having precisely controlled pore sizes, and the particles are separated according to their molecular size. Larger molecules are rapidly washed through the column; smaller molecules penetrate inside the porous of the packing particles and elute later.
- Ion-exchange HPLC**:** the stationary phase has an ionically charged surface of opposite charge to the sample ions. This technique is used almost exclusively with ionic or ionizable samples. The stronger the charge on the sample, the stronger it will be attracted to the ionic surface and, thus, the longer it will take to elute.

Polyacrylamide gel electrophoresis

Electrophoresis separates mixtures of proteins based on charge, charge/mass ratio, size or shape. This technique is mainly used as an analytical and preparative tool, especially one-dimensional separation, often employed as a pre-fractionating technique (Guttman, Varoglu & Khandurina, 2004; Jorgenson & Evans, 2004). The state of purity of the sample is judged by PAGE; the number of stained polypeptide bands will decrease with progressive removal of contaminating proteins. Ultimately, a single, stained band should result, indicating a nearly homogeneous sample.

The following are different ways of separating proteins of distinct characteristics.

SDS-PAGE

SDS-PAGE (a function of the length of a polypeptide chain and its charge) describes a collection of related techniques to separate protein according to their electrophoretic mobility in the denatured (unfolded state).

In most proteins, the binding of SDS to the polypeptide chain imparts an even distribution of charge per unit mass, thereby resulting in a fractionation by approximate size during electrophoresis. SDS, a strong detergent, when heated with protein mixture at 100°C makes proteins become rod-like structures (Figure 1) possessing a uniform charge density, that is the same net negative charge per unit length. The electrophoretic mobilities of these proteins will be a linear function of the logarithms of their molecular weights (Laemmli 1970). Percentage of gel (acrylamide) to be prepared depends on the molecular weight of the protein to be resolved and is as shown in Table 8.

Figure 1: Protein denaturation on the addition of SDS.

Table 8: Percentage of acrylamide used for specific molecular weight of a protein.

% gel	Molecular weight range
7	50–500 kDa
10	20–300 kDa
12	10–200 kDa
15	3–100 kDa

Resolution of specific protein with acrylamide concentration.

Proteins can then be detected by a variety of means, but the most commonly used stains are silver and Coomassie Brilliant blue staining. In the former case, a silver colloid is applied to the gel. The silver binds to cysteine groups within the protein. The silver is darkened by exposure to ultraviolet light. The amount of silver can be related to the darkness and, therefore, the amount of protein at a given location on the gel. This measurement can only give approximate amounts, but is adequate for most purposes. Silver staining is 100× more sensitive than Coomassie Brilliant Blue with a 40-fold range of linearity (Switzer, Merril & Shifrin, 1979).

Native gel electrophoresis

Native gels, also known as non-denaturing gels, analyze proteins that are still in their folded state. Thus, the electrophoretic mobility depends not only on the charge-to-mass ratio, but also on the physical shape and size of the protein. Native PAGE also has the potential for separating proteins of identical molecular weight which cannot be resolved with SDS-PAGE. In addition, proteins on native PAGE usually retain their activity. This allows enzymes to be detected by sensitive and specific activity stains and delicate proteins to be resolved and recovered in a biologically active form.

The equation governing protein mobility in native gels is as follows:

$$\log R_f = \log(Y_o) - K_R T$$

where

- R_f = relative mobility, normalized to the dye front or some other standard.
- Y_o = relative mobility of the protein in the absence of any sieving matrix.
- K_R = "retardation coefficient," the extent to which the gel matrix affects mobility.
- T = % monomer of the gel matrix.

Two-dimensional gel electrophoresis (2DGE)

Mixtures of proteins are separated by two properties in two dimensions on 2D gels. In electrophoresis in the first dimension, molecules are separated linearly according to their *iso electric* point. In the second dimension, the molecules are then separated at 90° from the first electrophoretogram according to molecular mass. Since it is unlikely that two molecules will be similar in two distinct properties, molecules are more effectively separated in 2D electrophoresis than in 1D electrophoresis.

Iso electric focusing

Separation of the proteins by the *iso electric* point is called *iso electric* focusing (IEF). Thereby, a gradient of pH is applied to a gel and an electric potential is applied across the gel, making one end more positive than the other. At all pH values other than their *iso electric* point, proteins will be charged. The proteins applied in the first dimension will move along the gel and will accumulate at their *iso electric* point, that is, the point at which the overall charge on the protein is 0 (a neutral charge).

SDS-PAGE

After completion of the first dimension, the complexes are destroyed by applying the denaturing SDS-PAGE in the second dimension, where the proteins of the complexes are separated by their mass. Since the SDS molecules are negatively charged, the result of this is that all of the proteins will have approximately the same mass-to-charge ratio as each other. In addition, proteins will not migrate when they have no charge (a result of the IEF step); therefore, the coating of the protein in SDS (negatively charged) allows migration of the proteins in the second dimension. The proteins will be attracted to the more positive side of the gel (because SDS is negatively charged) proportionally to their mass-to-charge ratio.

Methods of characterization

The progress of purification can be monitored by evaluation of biological activity or an assay procedure based on enzymatic activity specific to the protein of interest. The progressive increase in specific activity can be measured by the spectrophotometric or colorimetric method which could be an excellent indicator of the efficacy of the purification step. Alternately, immunochemical procedures like western blotting or enzyme-linked immunosorbent assay (ELISA) can be carried out for proteins lacking a measurable biological activity provided suitable antibodies are available. *In vitro* and *in vivo* experiments could also be performed to check the biological activity during sequential steps of purification. X-ray crystallography and NMR spectroscopy were used to determine the 3D structure of proteins, but nowadays, through bioinformatics, structure and modeling of protein can be predicted by computer programs by using the chemical properties of amino acids and structural properties of known proteins.

UV-visible absorption spectrum

UV-visible spectroscopy (radiations with wavelengths between 10 and 1000 nm) offers information about the transition of the most external electrons of the atoms. Since atoms or molecules absorb UV-visible radiation at a different wavelength, spectroscopy/spectrometry is often used in physical and analytical chemistry for the identification of substances through the spectrum emitted from or absorbed by them. This technique is also used to assess the concentration or amount of a given species using the Lambert-Beer law.

$$\log T = \log I / \log I_0 = l\alpha$$

where T is the transmissivity of light, l is the path length and α is the absorption coefficient of the substance.

The wavelength of maximum absorption and the intensity of absorption are determined by the molecular structure and the effect on the absorption spectrum of a compound when diluted in a solvent but vary depending on the chemical structures involved. UV-visible spectroscopy can be used to determine many physicochemical characteristics of compounds and thus can provide information to identify a particular compound. The absorbance spectra of proteins result largely from the presence of aromatic amino acids tryptophan, tyrosine and phenylalanine. A protein at room temperature has a specific tertiary structure or conformation that in turn creates a specific electronic environment for the aromatic amino acids, and so as the protein is heated, it will, at a certain temperature, unfold or melt and lose its structure. In this process, the electronic environment of the aromatic amino acids changes, which in turn results in spectral changes or shifts (Levine and Federici 1982).

The chemical composition (number as well as type of amino acids) of the protein will affect the absorption that causes variation. Protein absorption at 280 nm is dependent on the amount of tyrosine and especially tryptophan: the aromatic ring of phenylalanine absorbs well at 260 nm, but not 280 nm. So proteins of similar molecular weight can have quite different absorbances, since they can have completely different tryptophan and tyrosine content. UV absorbance of aromatic side chains is also affected by protein structure, temperature, pH, ionic strength or the presence of detergents that can affect the ability of aromatic residues to absorb light at 280 nm, and change the value of the protein's extinction coefficient.

The A260/A280 ratio can be used as a guide to the purity of the sample (GE Healthcare Life Sciences-Spectrophotometry handbook) (GE Health care, Chicago, IL, USA).

Circular dichroism

3D configuration constitutes the biomolecular structure of the protein which aids in the identification of the protein's interactions and function. Circular dichroism (CD) spectroscopy is the most widespread technique used for estimating the secondary structures of proteins and polypeptides in solution (Greenfield 1996). This technique can be used to distinguish between unordered (random coil) and ordered (α-helix or β-sheet) structures (Johnson 1999, 1988). CD detects wavelength-dependent differences in the absorption of right and left circularly polarized light by optically active molecules such as peptides and proteins. The CD spectrum of unordered peptides is usually characterized by a single band below 200 nm, whereas α-helical structures usually present two negative bands at 208 and 222 nm along with one positive band at 192 nm; β-sheet structures typically show a negative band at 217 nm and a positive band at 195 nm. Proteins possess a number of chromophores which can give rise to CD signals. The secondary structural features like α-helix and β-sheet analyzed from the CD spectrum (Kelly, Jess & Price, 2005) correspond to peptide bond absorption in the far-UV region

(240–180 nm) (Figure 2). The tertiary structure of the protein can be obtained in the near-UV region (320–260 nm) reflecting the proximities of aromatic amino acid side chains.

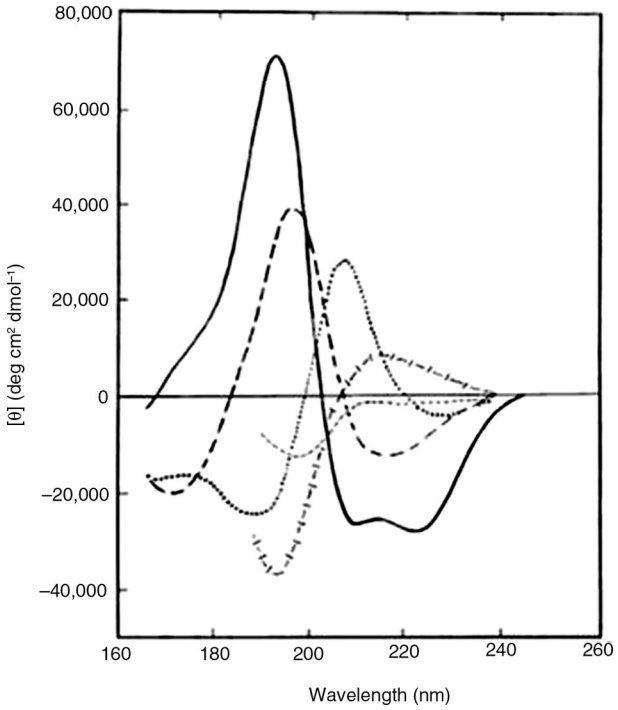

Figure 2: CD spectrum with various curves representing specific protein folding.
Far UV CD spectra associated with various types of secondary structure.
Note: Solid line: α-helix; long dashed line: anti-parallel β-sheet; dotted line: type I β-turn; cross dashed line: poly II helix; short dashed line: irregular structure.

CD spectroscopy has the advantage that it can be applied to molecules in solution and is remarkably sensitive to the secondary structure of biological polymers. CD signals of other non-protein chromophores such as flavin and heme moieties depend on the precise environment of the chromophore concerned and also their interactions. CD is an extremely useful technique for assessing the structural integrity of membrane proteins during extraction and characterization procedures.

Fluorescent spectrum

Proteins and peptides, when excited with UV light, are intrinsically fluorescent due to the presence of aromatic amino acids besides many enzymatic cofactors, such as flavin mononucleotide, flavin adenine dinucleotide, nicotinamide adenine dinuceotide and porphyrins. These moieties have a common trait in that they all contain aromatic ring structures that absorb UV light for excitation. The inherent fluorescence of proteins is primarily contributed by the three amino acid residues like tryptophan, tyrosine and phenylalanine. These residues have distinct absorption and emission wavelengths and differ in the quantum yields. Tryptophan is much more fluorescent than either tyrosine or phenylalanine but its fluorescent properties are solvent dependent. Signals in the region from 250 to 270 nm are attributable to phenylalanine residues, from 270 to 290 nm to tyrosine and those from 280 to 300 nm to tryptophan (Table 9). Disulfide bonds give rise to broad weak signals throughout the near-UV spectrum.

Table 9: Absorbance and fluorescent properties of aromatic amino acids.

Amino acid	Absorbance		Fluorescence	
	λ_{max} (nm)	ε_{max} (M^{-1} cm^{-1})	λ_{max} (nm)	φ_F [b]
Tryptophan	280	5600	355	0.13
Tyrosine	275	1400	304	0.14
Phenylalanine	258	200	282	0.02

Specific absorption and fluorescent wavelengths of aromatic aminoacids.

As the polarity of the solvent decreases, the spectrum shifts to shorter wavelengths and increases in intensity and this phenomenon has been utilized to study protein denaturation (Lakowicz 1999). Tyrosine can be excited at wavelengths similar to that of tryptophan, but emits at a distinctly different wavelength. Tyrosine provides significant signal, as it is often present in large numbers in many proteins but is less compared to tryptophan. Tyrosine fluorescence has been observed to be quenched by the presence of nearby tryptophan moieties via resonance energy transfer, as well as by ionization of its aromatic hydroxyl group. Phenylalanine is very weakly fluorescent and can only be observed in the absence of both tryptophan and tyrosine. However, the fluorescence spectrum of a protein containing the three amino acids usually resembles that of tryptophan due to its greater absorptivity, higher quantum yield and resonance energy transfer.

Tryptophan, the most significant fluorescence emitter, will have an emission peak at lower wavelengths if it is buried within the hydrophobic inner regions of a protein (Lakowicz 1999). Tyrosine moieties will often transfer their energy to adjacent tryptophan amino acids, while ionized tyrosine also demonstrates wavelengths similar to tryptophan, suggesting that for many proteins a good starting point for excitation and emission wavelengths are those for tryptophan.

Gene expression studies

Updated technology using more sophisticated, fast and robust tools needs to be used at the best in order to cope up with the increasing incidence of diabetes. DNA microarray is one such technique where we can study the expression of a number of genes in a single reaction and have quick and efficient analysis (Bosch et al. 2000). This technology has empowered the scientific community to explore the genetic causes of anamalies in humans (Kane et al. 2000). It is a complementary hybridization between nucleic acids where pieces of known DNA are anchored to a solid support and targets are fluorescently labeled, free-floating amplified RNA or complementary DNA from specific samples (Cheung et al. 1999). The key regulatory genes that are mainly involved in diabetes are phosphoenol pyruvate carboxykinase (PEPCK), glucokinase, sterol regulatory element binding protein-1c (SREBP) and glucose transporter protein (GLUT-2).

PEPCK is the rate-limiting enzyme in the gluconeogenic pathway by converting oxaloacetate into phosphoenolpyruvate and its expression is controlled by the rate of transcription of its gene (Sasaki et al. 1984). Dysregulation of PEPCK is a contributing factor to the pathogenesis of both type 1 and type 2 diabetes mellitus (DeFronzo, Bonadonna & Ferrannini, 1992; Sun et al., 2002).

Hepatic glucokinase is a crucial enzyme in maintaining glucose homeostasis and its activity in liver is low during fasting and in diabetes mellitus. The activity increases after carbohydrate ingestion or in diabetic animals after treating with insulin secretagogue agents (in T_2DM only) (Sharma, Majeshwar & Weinhouse, 1963). Once glucose flux is increased, it can simultaneously promote glycolysis and lipogenesis, and decrease gluconeogenesis by inhibiting PEPCK. GK (hexokinase IV) is the major glucose-phosphorylating enzyme in the liver and its abundance is regulated by insulin and glucagon, and post-translationally by the GK regulatory protein (Iynedjian 1993).

GLUT2 is membrane protein, transports glucose across the hepatic plasma membrane bidirectionally. GLUT2 is mainly expressed in the liver, β-cells of the pancreas and the baso lateral membrane of kidney proximal tubules and plays an important role in glucose homeostasis in living organisms. The molecular mechanism involved in the regulation of GLUT2 expression was postulated by Im et al. (2005). According to their findings, GLUT2 expression is dependent on SREBP-1c levels, higher levels of SREBP-1c triggers the upregulation of GLUT2 and vice versa.

SREBP-1c belongs to a family of transcription factors involved in cholesterol and fatty acid metabolism. Three isoforms of SREBPs are known to date: SREBP-1a, SREBP-1c and SREBP-2. SREBP-2 plays a crucial role in regulating cholesterol synthesis, whereas SREBP-1c controls gene expression of lipogenic enzymes (Horton et al. 1998 and Shimano 2001). Insulin is known to be a trigger for enhancing the SREBP-1c transcription levels in livers of streptozotocin-induced diabetic rats (Shimomura et al. 2000). Yamamoto et al. (2004) postulated insulin-independent induction of SREBP-1c expression in the re-fed state in the livers of STZ-induced diabetic mice.

Immunological characterization

Antibodies to particular proteins or to their modified forms have been used in biochemistry and cell biology studies and have been playing a pivotal role in protein identification as they are among the most common tools in molecular biology techniques. Modified proteins can be studied by developing an antibody specific to

that modification which can be used to determine the set of proteins that have undergone the modification of interest.

Western blotting

The Western blot can be used for detection and quantification of individual proteins, where in an initial step a complex protein mixture is separated using SDS-PAGE and then the protein of interest is identified using an antibody. The procedure can be broken down into a series of steps:

- Size separation of the proteins in the mixture by PAGE.
- Transfer of the separated proteins to a membrane while retaining their relative position.
- Detecting the protein under investigation by its specific reaction with an antibody and determination of its size relative to standard proteins of known size.

Usually, for protein blotting, either nitrocellulose or polyvinylidene fluoride membranes are used, but nitrocellulose is more compatible with most protein stains that enable visual inspection of transfer efficiency. While proteins are transferred to a membrane, blocking reagents like non-fat milk, bovine serum albumin, etc., are used to reduce antibodies non-specific binding on the membrane.

Transfer techniques

There are different methods used to transfer proteins electrophoretically to a membrane (Kurien and Scofield 2003).

Wet electro blotting (tank transfer)

The gel is first equilibrated with the transfer buffer and then placed in the "transfer sandwich" (filter paper-gel membrane-filter paper), cushioned by pads and pressed together by a support grid. The gel sandwich is placed vertically in a tank between stainless/platinum wire electrodes and filled with transfer buffer.

Semidry electro blotting (semidry transfer)

The transfer sandwich is placed horizontally between two plate electrodes in a semidry transfer apparatus. The gel is pre-equilibrated with transfer buffer so that maximum current passes through the gel instead of around the gel.

Dry electro blotting (dry transfer)

A unique gel matrix (transfer stack) that incorporates buffer is used instead of buffer tanks or soaked filter papers. The high ionic density in the gel matrix enables rapid. Here, the transfer time is reduced compared to conventional transfer techniques.

Vacuum blotting (vacuum capillary blotting)

This method uses a slab gel dryer system or other suitable gel drying equipment to draw polypeptides from gel to membrane. Gels dry out after some time under vacuum requiring plenty of reserve buffers.

Gel staining is used to monitor whether proteins have separated properly on the gel and membrane staining is used to monitor the efficiency of the protein transfer to the membrane from the gel. Membrane staining is important to ensure that the results seen are not unduly influenced by the failure of proteins in a particular area of the gel from transferring.

Detection

A variety of detection methods are available (chemiluminescence, fluorescence, colorimetric) and the choice of technique should match the experimental requirement.

Protein detection

Membrane-bound proteins are generally detected using secondary antibodies that are labeled with radioisotopes or colloidal gold or conjugated to fluorescent molecules (fluorophore) or an enzyme such as alkaline phosphatase (AP) or horseradish peroxidise (HRP).

Colorimetric detection

Enzymes such as AP and HRP convert several substrates to a colored precipitate when accumulated on the blot, develop a colored signal and are readily visible. The enzymatic reaction can be monitored and stopped when the desired signal on background is reached and so easier to use than film-based detection methods which require trial and error in determining appropriate exposure times and use of costly materials such as X-ray film and darkroom chemicals.

Fluorescence detection

In this, a primary or secondary antibody labeled with a fluorophore is used during immune detection. A light source excites the fluorophore and the emitted fluorescent signal is captured by a camera to produce the final image.

Immunodetection with enhanced chemiluminescence

This technique is based on the detection of light emission from a secondary antibody linked to horseradish peroxidase which catalyzes the oxidation of luminal and the light emission is detected on X-ray film. Although this system is more expensive and less convenient than the colorimetric alkaline phosphatase detection system, the sensitivity can be incredible.

Bioluminescence detection involves incubation of the membrane (with bound antigen-antibody-enzyme complex) in a bioluminogenic substrate and simultaneous measurement of emitted light. The substrate here is a luciferin-based derivative where light detection is performed using a photon counting camera, and the blotted proteins are visualized as bright spots.

The γ-emitting radioisotope ^{125}I is used to label lysines in immunoglobulins for radiometric antigen detection, and the radiolabeled blots are detected using X-ray film (autoradiography).

Immunogold labeling

This method utilizes gold-labeled secondary antibodies for antigen detection. As this method has relatively low sensitivity and the signal is not permanent, silver enhancement methods similar to that of colloidal gold total protein stains are developed.

ELISA

The ELISA has been used for decades to detect and quantitatively measure proteins in samples. ELISA is a test that uses antibodies and color change to identify a substance. Antigens from the sample are attached to a surface. Then, a further specific antibody is applied over the surface so it can bind to the antigen. This antibody is linked to an enzyme, and, in the final step, a substance containing the enzyme's substrate is added. The subsequent reaction produces a detectable signal, most commonly a color change in the substrate.

ELISA also known as the enzyme immunoassay is a plate-based assay technique designed for detecting and quantifying substances such as peptides, proteins, antibodies and hormones. In an ELISA, an antigen must be immobilized to a solid surface and then complexed with an antibody that is linked to an enzyme. Detection is accomplished by assessing the conjugated enzyme activity via incubation with a substrate to produce a product that can be measured. The most crucial element of the detection strategy is a highly specific antibody-antigen interaction.

In "Direct ELISA," only a labeled primary antibody is used, and in "Indirect ELISA," the antigen is bound by the primary antibody which is then detected by a labeled secondary antibody (Schmidt et al. 2012). In "Sandwich ELISA," the capture antibody is first added followed by the detecting antibody which binds to an antigen. Later, the enzyme-linked secondary antibody is added which binds to the detecting antibody succeeded by a substrate to be converted by the enzyme to a detectable form.

Mass spectroscopy

MALDI-TOF mass spectrometry has become an important tool of choice for large molecular analyses, especially for proteins. First, the samples are mixed with an organic compound that acts as a matrix to facilitate the desorption and ionization of compounds in the sample. The analyte molecules are distributed throughout the matrix so that they are completely isolated from each other. Some of the energy incident on the sample plate is absorbed by the matrix, causing rapid vibrational excitation. The analyte molecules can become ionized by simple protonation by the photo-excited matrix, leading to the formation of the typically singly charged ions. Some multiply charged ions are also formed, but this is rarer. The analyte ions are then accelerated by an electrostatic field to common kinetic energy. If all the ions have the same kinetic energy, the ions with low mass to charge ratio (m/z) travel faster than those with higher m/z values; therefore, they are separated in the flight tube and the number of ions reaching the detector at the end of the flight tube is recorded as the intensity of the ions. For MALDI, normally the charge is equal to 1 or 2.

MALDI-TOF mass spectrometry is the technique of choice for peptide mass fingerprinting. The peptide mass fingerprint of a protein is believed to be fastest and cheapest and specific enough to identify the protein solely by the comparison of the peptide mass values measured with those calculated by applying the corresponding enzyme cleavage rules, using an appropriate scoring algorithm. The protein of interest is first cleaved into smaller peptides by appropriate protease, whose absolute masses can be accurately measured with a mass spectrometer such as MALDI-TOF or electrospray ionization (ESI)-TOF (Clauser, Baker & Burlingame, 1999). The peptide masses are compared to either a database containing known protein sequences or even the genome. This is achieved by using computer programs that translate the known genome of the organism into proteins, then theoretically cut the proteins into peptides, and calculate the absolute masses of the peptides from each protein. They then compare the masses of the peptides of the unknown protein to the theoretical peptide masses of each protein encoded in the genome. The results are statistically analyzed to find the best match.

Structural information can be generated using certain types of mass spectrometers, usually those with multiple analyzers which are known as tandem mass spectrometers. This is achieved by fragmenting the sample inside the instrument and analyzing the products generated. The sample has to be introduced into the ionization source of the instrument. Once inside the ionization source, the sample molecules are ionized, because ions are easier to manipulate than neutral molecules. These ions are extracted into the analyzer region of the mass spectrometer where they are separated according to their m/z ratios. The ionization methods used for the majority of biochemical analyses are ESI and MALDI. There are a number of mass analyzers currently available, the better known of which include quadrupoles, TOF analyzers, magnetic sectors and both Fourier transform and quadrupole ion traps.

Tandem (MS-MS) mass spectrometers are instruments that have more than one analyzer and so can be used for structural and sequencing studies. The detector monitors the ion current, amplifies it and the signal is then transmitted to the data system where it is recorded in the form of mass spectra. The m/z values of the ions are plotted against their intensities to show the number of components in the sample, the molecular mass of each component and the relative abundance of the various components in the sample.

ESI is one of the atmospheric pressure ionization techniques and is well suited to the analysis of polar molecules ranging from less than 100 Da to more than 1,000,000 Da in molecular mass (Yamashita and Fenn 1984). In this, the sample is dissolved in a polar, volatile solvent and pumped through a narrow, stainless steel capillary (75–150 μm i.d.) at a flow rate of between 1 and 1 ml/min. A high voltage of 3 or 4 kV is applied to the tip of the capillary, which is situated within the ionization source of the mass spectrometer, and as a consequence of this strong electric field, the sample emerging from the tip is dispersed into an aerosol of highly charged droplets, a process that is aided by a co-axially introduced nebulizing gas flowing around the outside of the capillary. This gas, usually nitrogen, helps to direct the spray emerging from the capillary tip toward the mass spectrometer. The charged droplets diminish in size by solvent evaporation, assisted by a warm flow of nitrogen known as the drying gas which passes across the front of the ionization source which is seen in Figure 3. Eventually charged sample ions, free from solvent, are released from the droplets, some of which pass through a sampling cone or orifice into an intermediate vacuum region, and from there through a small aperture into the analyzer of the mass spectrometer, which is held under high vacuum. The lens voltages are optimized individually for each sample.

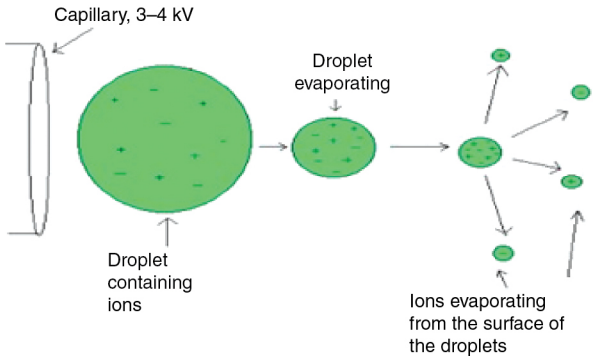

Figure 3: Principle of ESI.
Ion emission (positive and negative) from sample droplet.

Nanospray ionization (Wilm and Mann 1996) is a low flow rate version of ESI. A small volume (1–4 µl) of the sample dissolved in a suitable volatile solvent, at a concentration of ca. 1–10 pmol/µl, is transferred into a miniature sample vial. A reasonably high voltage (ca. 700–2000 V) is applied to the specially manufactured gold-plated vial resulting in sample ionization and spraying. This technique is used to analyze a protein digest mixture generating a list of molecular masses for the components present, and then analyzed further by tandem mass spectrometric (MS-MS) amino acid sequencing techniques.

In positive ionization mode, a trace of formic acid is often added to aid protonation of the sample molecules; in negative ionization mode a trace of ammonia solution or a volatile amine is added to aid deprotonation of the sample molecules. Proteins and peptides are usually analyzed under positive ionization conditions where the m/z scale must be calibrated by analyzing a standard sample of a similar type to the sample being analyzed (e.g., a protein calibrant for a protein sample), and then applying a mass correction.

Databases available for Id of MS spectra:

- SWISS-PROT-nr database of annotated protein sequences. It contains additional information on protein function, protein domains, known post-translational modifications, etc. (http://us.expasy.org/sprot).
- TrEMBL-computer-annotated supplement of Swiss-Prot that contains all the translations of EMBL nucleotide sequence entries not yet integrated in Swiss-Prot.
- PIR-International-nr annotated database of protein sequences (http://www-nbrf.georgetown.edu/).
- NCBInr-translated GenBank DNA sequences, Swiss-Prot, PIR.
- ESTdb-expressed sequence tag database (NIH/NSF).
- UniProt-proposed new database. It will joint Swiss Prot, TrEMBL, PIR (http://pir.georgetown.edu/uniprot).

Discussion

The peptide hormone, insulin first discovered in 1921 (Banting and Best 1922), is still the ultimate treatment of diabetes mellitus. Many secondary metabolites have been implicated in the anti-diabetic properties of plants (Grover, Yadav & Vats, 2002). Although the presence of insulin-like substances/proteins was reported in plant materials like onions, lettuce, green bean leaves, etc., way back in 1923, not much attention was paid to these results (Collip 1923). Collip isolated insulin from plant extracts and named it as "Glucokinin" which showed a significant decrease in blood glucose levels when tested on experimental rabbits. However, there was a renewed interest in plant-derived insulin when the presence of insulin was reported from the fruits of *Momordica charantia* (bitter gourd) (Khanna et al. 1981, 1976).

Recently a protein showing distinct functional and structural homology to insulin was characterized from embryo axes of maize. This 5.7 kDa protein, with a well-defined a-helix structure, induced selective synthesis of DNA as well as ribosomal proteins, just like mammalian insulin (Rodriguez-Lopez et al. 2011). Insulin-like protein was also isolated from the leaves (chloroplast) of *Bauhinia variegate* (Azevedo et al. 2006). Evidence based on immunological cross-reactivity and anti-diabetic properties has suggested the presence of insulin-like proteins in plants. However, isolated insulin-like protein is found to have a synergistic effect in reducing the blood

glucose level in experimental animals and most of the insulin-like proteins of different plant species in Table 10 show peptide sequence homology with insulin (Khursheed, Anwer & Fatma, 2012).

Table 10: Similarities of plant insulin sequence with bovine insulin.

Insulin	Sequence
Bovine insulin (α-chain)	1GIVEQCCASVCSLYQLENYCN 21
Bovine insulin (β-chain)	1FVNQHLCGSHLVEALYLVCGERGFFYTPKA 30
C. ensiformis I-SC	1 GIVEQCCASVCSLYQLENYCN 21
C. ensiformis I-LC	1 FVNQHLCGSHLVEALYLVCGERGFFYTPKA 30
V. unguiculata I-SC	1 GIVEQXXASVXSLYQLENYXN 21
V. unguiculata I-LC	2 FVNQHLXGSHLVEALYLVXGERGFFYTPKA 30
B. variegata a	1 GIVEQ 5
B. variegata b	1 FVNQH 5
Spirulina platnsis S5	44 GER 46

Sequence similarity of plant and animal insulin.

Venancio et al. (2003) identified that developing fruits of cow pea plant contained proteins of similar mass and amino acid sequence as bovine insulin and also exhibits metabolic functions as those of animal insulin by promoting several metabolic activities through glucose transportation into the cell and by phosphorylating proteins regulating carbohydrate metabolism. The secondary structure of glycoproteins (alpha amylase inhibitor from macrotyloma uniflorum and alpha amylase inhibitor from vigna unguiculata) isolated and purified from seeds of *Macrotyloma uniflorum* and *Vigna unguiculata* with molecular weights of 26.91 and 21.89 kD respectively were determined using CD studies (Gupta et al. 2013). The secondary structure of an antidiabetic protein isolated from Spirulina was also predicted from the CD spectrum (Anwer et al. 2012). The peptide sequence of glycoprotein from *M. uniflorum* was identified using MALDI-TOF/TOF analysis (Gupta et al. 2013). MS/MS has been effectively used for amino acid sequencing and investigating the structural information of bioactive peptides derived from different proteins, such as soy protein (Kodera and Nio 2006), casein (Hernández-Ledesma et al. 2004), collagen (Li et al. 2007) and ovotransferin (Majumder and Wu 2010).

Conclusion and future prospects

As of now, recombinant insulin from various animals including humans has been in use. When the insulin is directly available through plant products, the majority of which are neutraceuticals, there can be a hope of minimizing the use of synthetic drugs in the treatment of diabetes mellitus. However, a lot of research is needed to commercialize and formulate drugs from those sources but definitely will be a boon to diabetics in the near future.

References

Anwer, R.; Alam, A.; Khursheed, S.; Kashif, S. M.; Kabir, H.; Fatma, T. Spirulina: possible pharmacological evaluation for insulin-like protein. *J Appl Phycol.* **2012**, *25*, 883–889.

Ashoub, A.; Berberich, T.; Beckhaus, T.; Bruggemann. W. A competent extraction method of plant proteins for 2-D gel electrophoresis. *Electrophoresis* **2011**, *32*, 2975–2978.

Azevedo, C. R.; Maciel, F. M.; Silva, L. B.; Ferreira, A. T.; da Cunha, M.; Machado, O. L.; Fernandes, K. V.; Oliveira, A. E.; Xavier-Filho, J.; Isolation and intracellular localization of insulin-like proteins from leaves of Bauhinia variegate. *Braz. J. Med. Biol. Res* **2006**, *39*, 1435–1444.

Banting, F.G.; Best, C. H. The internal secretion of the pancreas. *J. Lab. Clin. Med.* **1922**, *7*, 465–480.

Boonloh, K., Kukongviriyapan, V., Kongyingyoes, B., Kukongviriyapan, U., and Pannangpetch, S. T. P,. Rice bran protein hydrolysates improve insulin resistance and decrease pro-inflammatory cytokine gene expression in rats fed a high carbohydrate-high fat diet. *Nutrients* 2015, *7*, 6313–6329. .

Bosch, J. T.; Seidel, C.; Batra, S.; Lam, H.; Tuason, N.; Saljoughi, S.; Saul, R. *Validation of sequence-optimized 70 base oligonucleotides for use on DNA microarrays.* TIGR Genome Sequencing and Analysis Conference, 12-15 September, Miami Beach, FI, 2000.

Boschetti, E.; Bindschedler, L. V.; Tang, C.; Fasoli, E.; Righetti, P. G. Combinatorial peptide ligand libraries and plant proteomics: a winning strategy at a price. *J. Chromatogr. A* **2009**, *1216*, 1215–1222.

Bradford, M. M. Rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **1976**, *72*, 248–254.

Carpentier, S. C.; Panis, B.; Vertommen, A.; Swennen, R.; Sergeant, K.; Renaut, J.; Laukens, K.; Witters, E.; Samyn, B.; Devreese, B. Proteome analysis of non-model plants: a challenging but powerful approach. *Mass Spectrom Rev*. **2008**, *27*, 354–377.
Caruso, G.; Cavaliere, C.; Foglia, P.; Gubbiotti, R.; Guarino, C.; Samperi, R.; Laganà, A. Analysis of drought responsive proteins in wheat (*Triticum aestivum* durum) by 2D-PAGE and MALDITOF mass spectrometry. *Plant Sci*. **2009**, *177*, 570–576.
Cheung, V. G.; Morley, M.; Aguilar, F.; Massimi, A.; Kucherlapati, R.; Childs, G. Making and reading microarrays. *Nat. Genet*. **1999**, *21*, 15–19.
Clauser, K. R.; Baker, P.; Burlingame, A. L. Role of accurate mass measurement (±10 ppm) in protein identification strategies employing MS or MS/MS and database searching. *Anal Chem*. **1999**, *71*, 2871–2882.
Collip, J. B. Glukokinin: a new hormone present in plant tissue. *J. Biol. Chem*. **1923**, *56*, 513–543.
DeFronzo, R. A.; Bonadonna, R. C.; Ferrannini, E. Pathogenesis of NIDDM. A balanced overview. *Diabetes Care* **1992**, *15*, 318–368.
Deutscher, M. P. *A Guide to protein purification: methods in enzymology*, Vol. 182. Academic Press: San Diego, CA, USA, 1990.
Faurobert, M.; Pelpoir, E.; Chaïb, J. Phenol extraction of proteins for proteomic studies of recalcitrant plant tissues. *Methods Mol. Biol*. **2007**, *355*, 9–14.
Greenfield, N. J. Methods to estimate the conformation of proteins and polypeptides from circular dichroism data. *Anal. Biochem*. **1996**, *235*, 1–10.
Grover, J. K.; Yadav, S.; Vats, V. Medicinal plants of India with anti-diabetic potential. *J. Ethnopharmacol*. **2002**, *1*, 81–100.
Gueven, K. C.; Aktin, E.; Koyuncuoglu, H.; Bergisadi, N. Lipolytic and hypoglycemic compounds obtained from *Cystoseria barbata*. *Arzneim-Forsch*. **1974**, *24*,144–147.
Gupta, L. H.; Chougale, A.; Kulkarni, M.; Sabharwal, S. G. Characterization of the α-amylase inhibitor from the seeds of *Macrotyloma uniflorum* and *Vigna unguiculata*. *Int. J. Pharm. Bio. Sci*. **2013**, *4*, 127–137.
Guttman, A.; Varoglu, M.; Khandurina, J. Multidimensional separations in the pharmaceutical arena. *Drug Discov. Today* **2004**, *9*, 136–144.
Hadjivassiliou, M.; Grunewald, R.; Sharrack, B.; Sanders, D.; Lobo, A.; Williamson, C.; Woodroofe, N.; Wood, N.; Davies-Jones, A. Gluten ataxia in perspective: epidemiology, genetic susceptibility and clinical characteristics. *Brain* **2003**, *126*, 685–691.
Hernández-Ledesma, B.; Amigo, L.; Ramos, M.; Recio, I. Angiotensin converting enzyme inhibitory activity in commercial fermented products. Formation of peptides under simulated gastrointestinal digestion. *J. Agric. Food Chem*. **2004**, *52*, 1504–1510.
Horton, J. D.; Bashmakov, Y.; Shimomura, I.; Shimano, H. Regulation of sterol regulatory element binding proteins in livers of fasted and re fed mice. *Proc. Natl. Acad. Sci USA*. **1998**, *95*, 5987–5992.
Im, S. S.; Kang, S. Y.; Kim, S. Y.; Kim, H. I.; Kim, J. W.; Kim, K. S.; Ahn, Y. H. Glucose-stimulated up regulation of GLUT2 gene is mediated by sterol response element-binding protein-1c in the hepatocytes. *Diabetes* **2005**, *54*, 1684–1691.
Iynedjian, P. B. Mammalian glucokinase and its gene. *Biochem. J*. **1993**, *293*, 1–13.
Johnson, W. C. Analyzing protein circular dichroism spectra for accurate secondary structures. *Proteins* **1999**, *35*, 307–312.
Johnson, W. C. Secondary structure of proteins through circular dichroism spectroscopy. *Annu. Rev. Biophys. Biophys. Chem*. **1988**, *17*, 145–166.
Jorgenson, J. W.; Evans, C. R. Multidimensional LC-LC and LC-CE for high-resolution separations of biological molecules. *Anal. Bioanal. Chem*. **2004**, *378*, 1952–1961.
Kane, M. D.; Jatkoe, T. A.; Stumpf, C. R.; Lu, J.; Thomas, J. D.; Madore, S. J. Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays. Department of Molecular Biology and Genomics and Department of Infectious Diseases, Pfizer, Global Research and Development, Ann Arbor, MI. 2000.
Kelly, S. M.; Jess, T. J.; Price, N. C. How to study proteins by circular dichroism. *BBA Proteins Proteom*. **2005**, *1751*, 119–139.
Khanna, P.; Nag, T. N.; Chandrajain, S.; Mohan, S. Process for isolation of insulin from plant source. Granted US Patent No. 3945988, 1976.
Khanna, P.; Jain, S. C.; Panagariya, A.; Dixit, V. P. Hypoglycemic activity of polypeptide-P from a plant source. *J. Nat. Prod*. **1981**, *44*, 648–655.
Khursheed, S.; Anwer, R.; Fatma, T. Insulin like antigen: sources other than pancreas. *Int. J. Curr. Pharm. Res*. **2012**, *4*, 24–28.
Kim, S. T.; Cho, K. S.; Jang, Y. S.; Kang, K. Y. Two-dimensional electrophoretic analysis of rice proteins by polyethylene glycol fractionation for protein arrays. *Electrophoresis* **2001**, *22*, 2103–2109.
Kodera, T.; Nio, N. Identification of an angiotensin I-converting enzyme inhibitory peptides from protein hydrolysates by a soybean protease and the antihypertensive effects of hydrolysates in 4 spontaneously hypertensive model rats. *J. Food Sci*. **2006**, *71*, C164–C173.
Kurien, B. T.; Scofield, R. H. Protein blotting: a review. *J. Immunol. Methods* **2003**, *274*, 1–15.
Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **1970**, *227*, 680–685.
Lakowicz, J. R. Ed. Principles of Fluorescence Spectroscopy. 2nd edn. Kluwer Academic/Plenum Publishers: New York, 1999.
Levine, R. L.; Federici, M. M. Quantification of aromatic residues in proteins; model compounds for second derivative spectroscopy. *Biochemistry* **1982**, *21*, 2600–2606.
Li, B.; Chen, F.; Wang, X.; Ji, B.; Wu, Y. Isolation and identification of antioxidative peptides from porcine collagen hydrolysate by consecutive chromatography and electrospray ionization–mass spectrometry. *Food Chem*. **2007**, *102*, 1135–1143.
Lowry, O. H.; Rosebrough, N. J.; Farr, A. L.; Randall, R. J. Protein measurement with the Folin phenol reagent. *J. Biol. Chem*. **1951**, *193*, 265–275.
Majumder, K.; Wu, J. A new approach for identification of novel antihypertensive peptides from egg proteins by QSAR and bioinformatics. *Food. Res. Int*. **2010**, *43*, 1371–1378.
Marella, S.; Maddirela, D. R.; Badri, K. R.; Jyothi Kumar, M. V.; Chippada A. Antihyperlipidemic and biochemical activities of Mcy protein in streptozotocin induced diabetic rats. *Cell Physiol. Biochem*. **2015**, *35*, 1326–1334.
Marella, S.; Maddirela, D. R.; Kumar, E. G.; Tilak, T. K.; Badri, K. R.; Chippada, A. Mcy protein, a potential antidiabetic agent: evaluation of carbohydrate metabolic enzymes and antioxidant status. *Int. J. Biol. Macromolec*. **2016**, *86*, 481–488.
Martínez-Maqueda, D.; Hernández-Ledesma, B.; Amigo, L.; Miralles, B.; Gómez-Ruiz, J. A. Chapter 2: Proteomics in Foods, Principles and Applications. In *Extraction/Fractionation Techniques for Proteins and Peptides and Protein Digestion*; Toldra, F., and Nollet , L. M. L. Eds.; Springer Science+Business Media: New York, 2013.
Mendes, R. H.; Hagen, M.K.; Barp, J.; de Jong, E. V.; Moreira, J. D.; Reischak-Oliveira, Á.; Irigoyen, M. C.; Belló-Klein, A. Isolated soy protein-based diet ameliorates glycemia and antioxidants enzyme activities in streptozotocin-induced diabetes. *Food Nutr. Sci*. **2014**, *5*, 2089–2096.
Mohammad, L.; Inamuddin. Ion exchange technology II. Springer: The Netherlands, 2012, p. 1.

Nehete, J. Y.; Bhambar, R. S.; Narkhede, M. R.; Gawali, S. R. Natural proteins: sources, isolation, characterization and applications. *Pharmacogn. Rev*. **2013**, *7*, 107–116.
Ng, T. B.; Yeung, H. W. Hypoglycemic constituents of *Panax ginseng*. *Gen. Pharmacol*. **1985**, *16*, 549–552.
Ninfa, A. J.; Ballou, D.P.; Marilee, B. *Fundamental Laboratory Approaches for Biochemistry and Biotechnology*, 2nd ed. Wiley, 2009, p. 133.
Pavokovic, D.; Kriznik, B.; Krsnik-Rasol, M. Evaluation of protein extraction methods for proteomic analysis of non-model recalcitrant plant tissues, Croat. *Chem Acta*. **2012**, *85*, 177–183.
Rajasekhar, M. D.; Badri, K. R.; Vinay Kumar, K.; Babu, K. R.; Fatima, S. S.; Sampath Kumar, M. T.; Appa Rao, C. Isolation and characterization of a novel antihyperglycemic protein from the fruits of *Momordica cymbalaria*. *J. Ethnopharmacol*. **2010**, *128*, 58–62.
Rodriguez-Lopez, C. D.; Rodriguez-Romero, A.; Aguilar, R.; de Jimenez, E. S. Biochemical characterization of a new maize (*Zea mays L.*) peptide growth factor. *Protein Pept. Lett*. **2011**, *18*, 84–91.
Sasaki, K.; Cripe, T. P.; Koch, S. R.; Andreone, T. L.; Petersen, D. D.; Beale, E. G.; Granner, D. K. Multihormonal regulation of phosphoenolpyruvate carboxykinase gene transcription: the dominant role of insulin. *J. Biol. Chem*. **1984**, *259*, 15242–15251.
Schmidt, S. D.; Mazzella, M. J.; Nixon, R. A.; Mathews, P. M. Aβ measurement by enzyme-linked immunosorbent assay. *Methods Mol. Biol*. **2012**, *849*, 507–527.
Scopes, R. K. *Protein Purification, Principles and Practice*, 3rd ed. Springer-Verlag: New York, 1993, p. 346.
Sharma, C.; Majeshwar, R.; Weinhouse, S. Effects of diet and insulin on glucose-adenosine triphosphate phosphotransferases of rat liver. *J. Biol. Chem*. **1963**, *238*, 3840–3845.
Sharma, G. M.; Mundoma, C.; Seavy, M., Roux; K. H.; Sathe, S. K. Purification and biochemical characterization of Brazil nut (*Bertholletia excelsa* L.) seed storage proteins. *J. Agric. Food Chem*. **2010**, *58*, 5714–5723.
Shimano, H. Sterol regulatory element-binding proteins (SREBPs): transcriptional regulators of lipid synthetic genes. *Prog. Lipid Res*. **2001**, *40*, 439–452.
Shimomura, I.; Matsuda, M.; Hammer, R. E.; Bashmakov, Y.; Brown, M. S.; Goldstein, J. L. Decreased IRS-2 and increased SREBP-1c lead to mixed insulin resistance and sensitivity in livers of lipodystrophic and ob/ob mice. *Mol. Cell*. **2000**, *6*, 77–86.
Singh, K. N.; Chandra, Y. Hypoglycemic and hypocholesterolemic effects of proteins of *Acacia milanoxylon* and *Bauhinia retusa* wild leguminous seeds in young albino rats. *J. Indian Med. Assoc*. **1977**, *68*, 201–203.
Handbooks from GE ealthcare lifesciences. n.d. Nucleic Acid Sample Preparation for Downstream Analyses Principles and Methods 28-9624-00. In *Spectrophotometry Handbook*. GE Healthcare Lifesciences: Chicago, IL, USA.
Sun, Y.; Liu, S.; Ferguson, S.; Wang, L.; Klepck, P.; Yun, J. S.; Friedman, J. E. Phosphoenolpyruvate carboxykinase over expression selectively attenuates insulin signalling and hepatic insulin sensitivity in transgenic mice. *J. Biol. Chem*. **2002**, *277*, 23301–23307.
Switzer, R. C. III; Merril, C. R.; Shifrin, S. A highly sensitive silver stain for detecting proteins and peptides in polyacrylamide gels. *Anal. Biochem*. **1979**, *98*, 231–237.
Venancio, T. M.; Oliveira, A. E.; Silva, L. B.; Machado, O. L.; Fernandes, K. V.; Xavier Filho, J. A protein with amino acid sequence homology to bovine insulin is present in the legume *Vigna unguiculata* (Cow pea). *Braz. J. Med. Biol. Res*. **2003**, *36*, 1167–1173.
Watanabe, M.; Kato, M.; Ayugase, J. Anti-diabetic effects of adlay protein in type 2 diabetic *db/db* mice. *Food Sci. Technol. Res*. **2012**, *18*, 383–390.
Wilm, M.; Mann, M. Analytical properties of the nanoelectrospray ion source. *Anal. Chem*. **1996**, *68*, 1–8.
World Health Organization. WHO Meeting on the Promotion and Development of Traditional Medicine (1977: Geneva) & World Health Organization. The promotion and development of traditional medicine : report of a WHO meeting, 28 November to 2 December, 1978. http://www.who.int/iris/handle/10665/40995.
Yamamoto, T.; Shimano, H.; Nakagawa, Y.; Ide, T.; Yahagi, N.; Matsuzaka, T.; Nakakuki, M.; Takahashi, A.; Suzuki, H.; Sone, H.; Toyoshima, H.; Sato, R.; Yamada, N. Author information. SREBP-1 interacts with hepatocyte nuclear factor-4α and interferes with PGC-1 recruitment to suppress hepatic gluconeogenic genes. *J. Biol. Chem*. **2004**, *279*, 12027–12035.
Yamashita, M.; Fenn, J. B. Negative ion production with the electrospray ion source. *J. Phys. Chem*. **1984**, *88*, 4671–4675.