Proteomics and Rational Drug Design

Dr Gilda Tachedjian
Molecular Interactions Group
Dec 2006
Topics to be Covered in this Presentation

- What is Proteomics?
- Tools used to study the Proteome.
- Application in medical diagnosis
- Application in rationale drug design.
- Application in vaccines
What is Proteomics?

- **Proteome:**
  The complete set of proteins in existence in an organism throughout its life cycle,
  OR on a smaller scale the entirety of proteins found in a particular cell type under a particular type of stimulation.

- **Proteomics: Study of the Proteome**
  The effort to establish the identities, quantities, structure, and biochemical and cellular functions of complete complements of proteins in an organism, organ or organelles and how these properties vary in space, time and/or physiological state.
Traditional

GENE → mRNA → PROTEIN

20,000 to 25,000 genes in human genome

Contemporary

GENOME → TRANSCRIPTOME → PROTEOME

1,000,000 proteins in human proteome
Information Obtained from Genomic & Proteomic Analysis

- **Genome**
  - Sequencing
  - Gene identification
  - Gene structure
  - SNP (single nucleotide polymorphism)
  - Predict protein structure

- **Transcriptome**
  - DNA Microarrays
  - Gene expression patterns in normal and diseased tissue

- **Proteome**
  - Mass spectrometry
  - Protein arrays/chips
  - Protein identification
  - Protein quantitation
  - Post-translational modification
  - Protein-protein interaction
Need for Proteomics

- Functions of proteins depends on the structure and interaction, which cannot be predicted accurately from sequence information.
- Abundance of a given RNA transcript may not reflect the abundance of the corresponding protein.
- Protein diversity is generated post-transcriptionally.
- Protein activity often depends on post-translational modifications, which are not predictable from the level of corresponding mRNA.
- The function of a protein often depends on its localization.
- Proteins are the most therapeutically relevant molecules.
A single genome can hypothetically give rise to an infinite number of qualitatively and quantitatively different proteomes.

The advantage of the proteome analysis rests with the ability to monitor:

- effect of stage of the cell cycle;
- effect of growth
- effect of nutrient conditions
- effect of temperature & stress responses
- pathological changes...etc
Significance of Proteomics

- Proteins are the commercial endpoints of most research in the Biological/Biomedical Sciences.
- Proteins carry out nearly all controlled biological functions.
- Protein-protein interactions control most cellular processes.
- Most diseases are treated at the protein level.
- Proteins are extremely valuable products for the pharmaceutical, food, environmental and related biotechnological industries.
What does Proteomics study?

Protein and subcellular localization

Protein function and interactions

- Protein-ligand interactions
- Protein complexes (machines)
- Protein families (activity or structural)
- Protein X
- Protein Y
- Ions/small molecule
- DNA
- Transcription factors
- Nuclear pore complexes
- Ribosomes
- Proteases
- Secreted
- Transmembrane
- Membrane associated
- Post-translationally modified proteins
- Lipids
- Phospho-arginine (P Arg)
- Hydroxylation (Hx)

Burnet Institute
Working towards a healthy world
Key Technologies in Proteomics
Protein Separation

Reproducible separation & characterisation of large numbers of proteins from complex biological mixtures.

**Electrophoretic techniques**
- 1D Gel Electrophoresis (separation on size)
- 2D Gel Electrophoresis (separation on charge and size)

**Chromatographic techniques**
- Liquid Chromatography (size separation)
Protein Separation: 2D SDS-PAGE

1st Dimension
Isoelectric Focusing (IEF)

Equilibration
Reduction/Alkylation

Protein separated based on electric charge in a pH gradient - protein stops migrating in gel when reaches isoelectric point (i.e. not charged)
Cut Out Protein from 2D Gel

- Picker head
- Gel tray
- Camera and light assembly
- Rinse station
- Racks
Mass Spectrometry Analysis of Tryptic Peptide Fragments

Electric field accelerates ion - lighter ions (if have same charge) reach detector first
Peptide Mapping and Mass Fingerprint by MALDI-TOF

- The most commonly used technique (MALDI = matrix assisted laser desorption ionization; TOF = time of flight)

- Proteins are identified by matching a list of experimental peptide masses with the calculated list of all peptides masses of each entry in a proteomics databases.

- Mass mapping require a purified protein the technique is generally used in conjunction with prior protein fractionation
What can Proteomics be used for:

Research: Studies of basic cell function and molecular organisation

The discovery of novel drug targets i.e. Tamiflu and Relenza for Influenza Virus

The discovery of molecular markers (biomarkers) for diagnosis and monitoring disease i.e. prostate specific antigen in blood as marker for prostate cancer

The discovery of antigens expressed in cancer cells that can be targeted for vaccines development
Biomarker Discovery Process

Discovery ➔ Characterization ➔ Validation

- Rapid Screening
- Protein Diff. Display
- Biological Fluids
- Cell / Bact. Lysates
- Whole Cells (LCM)
- Digest Map Profiling
- Partial Sequence Det.
- Post-Translational Mods.
- Data Base Mining
- I.D. of Known Protein
- I.D. of Novel Protein
- Throughput
- Intensive
- Antibody Valid.
- Phage Display

LCM: Laser Capture Microdissection
Cancer Development and Progression

- Food
- Environmental factors
- Chemicals
- Mutations
  - Sporadic
  - Inherited

Clonal expansion
Pre-malignant lesion
Early cancer
Primary tumour
Advanced metastatic cancer

Intervention efficiency
High
Low

Early detection
Biomarkers
Improved survival
Likelihood of progression
Low
High
Role of serum markers in cancer

I. Early diagnosis & differential diagnosis of space-occupying lesions.

II. Follow-up & therapeutic outcome & early detection of relapse.

III. Assessment of biological malignancy and prognosis.

IV. Selection of therapeutic interventions.

V. Monitoring the patients who are at risk of developing cancer.
Proteomic Analysis Identifies Breast Cancer Biomarkers in Nipple Aspirate Fluid

NAF from Normal Breast

NAF from Cancer Breast

Identity of spots determined by mass spectrometry
Spot 1 = prolactin-induced protein
Spot 2 = apolipoprotein D
Spot 3 = α1-acid glycoprotein

Alexander et al 2004 Clinical Cancer Research 10:7500
Proteomic Analysis Identifies Breast Cancer Biomarkers in Nipple Aspirate Fluid

Significance of presence of proteins validated by testing for their presence in the NAF from 53 benign breasts and 52 from breasts with cancer using ELISA (enzyme linked immunosorbant assay) - plate coated with antibodies to proteins- binding detected by a color reaction.

$\alpha_1$-acid glycoprotein levels were higher in women with breast cancer ($P=0.002$)
Path to Drug Discovery

Target structure-function studies (X ray, NMR, computer)

Cell-virus molecular biology research

Medical need identified → Relevant mechanism

“Hits” → Lead compound → Drug candidate → Clinical testing

Mechanisms-based screens

High-throughput screens

Libraries: natural products, compound collections, combinatorial chemistry

Medicinal chemistry, analog synthesis, combinatorial chemistry

Preclinical development

Animal pharmacology, toxicology, metabolism, pharmacokinetics

Compounds synthesized or purified

Antiviral effect in cells

Not acceptable in cells

Representative numbers of compounds

100,000s

Rejected

Antiviral effect in animals

Not acceptable in animals

Antiviral effect in humans

Not acceptable in humans

Compound approved for general use

Time required

0

5-10 years

Rejected

Rejected

Rejected

Rejected

Rejected

Rejected

Principles of Virology, Flint et al
Screening for Antiviral Compounds

Cell based screen - Not Rational Drug Design

- can assay cytotoxicity and antiviral activity simultaneously
- virus replication in cell culture
- multiple targets in single screen
- suitable for high-throughput
- follow up studies of active compound more complicated
Screening for Antiviral Compounds

Cell based screen

Selectivity index (SI) = \frac{50\% \text{ cytotoxic conc. (CC}\_\text{50})}{50\% \text{ inhibitory conc. (IC}\_\text{50})}

IC\_\text{50} = 2.8
CC\_\text{50} = 295
SI = 105
Screening for Antiviral Compounds

Mechanism based screen (Rational)

• Well defined assay for single target
• Assays for viral proteases, polymerases, helicases
• Need purified protein and specific substrate
• Will not determine cytotoxicity
• Amenable to high-throughput
  - 1536 well trays with µl reactions
Sources of Compounds Used for Screening

- Libraries of chemical compounds (500,000 cmpd)
- Combinatorial libraries: all possible combinations of a basic set of molecular components - tagged microbeads or chemical supports so active compounds in mixtures can be traced and identified easily
- Natural products - diverse mixtures of unknown compounds from plants and marine animals

Principles of Virology, Flint et al
Zanamivir (Relenza) bound to Influenza virus neuraminidase (NA)

• Zanamivir -Australian designed drug to treat influenza based on 3D structure of viral NA complexed with substrate.
  • Peter Colman, Mark von Itzstein, Graeme Laver (VCP/CSIRO/ANU/Biota Holdings/GlaxoSmithKline)

• Inhibitor designed to fit in highly conserved deep cavity in NA
• Active against influenza A and B
• Zanamivir (Relenza) - nasal spray for therapeutic use and prevention.
• Oseltamivir (Tamiflu) - oral formulation (Gilead/Roche)
  Stockpiled by Australia - against H5N1 Avian influenza
Chemical Structures of Neuraminidase Inhibitors

- Neuraminic acid derivative mimics the geometry of the transition state during the enzymatic reaction.
- To increase interaction between the substrate and enzyme a guanidinyl group was substituted for a hydroxyl group - zanamivir.

Neuraminidase Inhibitors Block Influenza Release From Cells

Drugs targeting HIV protease

- **Saquinavir**
- **Indinavir**
- **Ritonavir**
- **Nelfinavir**
- **Amprenavir**
- **Lopinavir**
- **Atazanavir**
- **Tipranavir**
- **Darunavir**

**Protease Inhibitors**

• First inhibitors were peptidomimetics (mechanism based)

• Based on 7 amino acid substrate with noncleavable bond

Principles of Virology, Flint et al
“Virtual” or “in Silico” Screening

• Structure of target must be available

• Identify pockets at functionally important sites (i.e. catalytic site in enzyme or at protein:protein interface)

• Dock molecules into pocket - libraries available to public

• Order compounds, test in vitro for activity

• Repeat docking
**In silico (virtual) screening - an overview**

1. **Protein Database**
2. Visualise X-ray crystal structure of protein
3. Define binding site (pocket), setting up a grid box.
4. Fit a compound into the pocket (compound is flexible - protein is rigid)
5. Score binding energy of each compound
6. Thousands of compounds from databases
7. Results sorted by Energy (G-score)
Application in Therapeutic Cancer Vaccines

• A cell membrane glycoprotein Mucin 1 is highly expressed in breast cancer cells

• Mucin 1 shown to induce immune responses in healthy subjects and cancer patients - however responses are variable and weak.

• Modified immunogenic epitopes of Muc 1 being examined for their capacity to elicit an immune response for use as cancer vaccines (i.e. oxidized mannan-Muc 1)
Acknowledgments

Dr John Lee and A/Prof Mibel Aguilar, Department of Biochemistry and Molecular Biology, Monash University for making available their lecture material for use in this presentation
Australian Society for Medical Research (ASMR) www.asmr.org.au

- Peak Professional Society Representing Health and Medical Research
  - Political, scientific and public advocacy

- Public Relations
  - ASMR Medical Research Week: June 4th-8th 2007
    - Online school quiz for high school students - 20 multiple choice questions designed to find those students with a good knowledge in medical science - with prizes for winners (gildat@burnet.edu.au) (years 7-9 and years 10-12)

  - High school visits by Medical Researchers
    - Visits to High School Science classes and Career nights
    - Regional "Caravan of Science" Tours