Protein Arrays

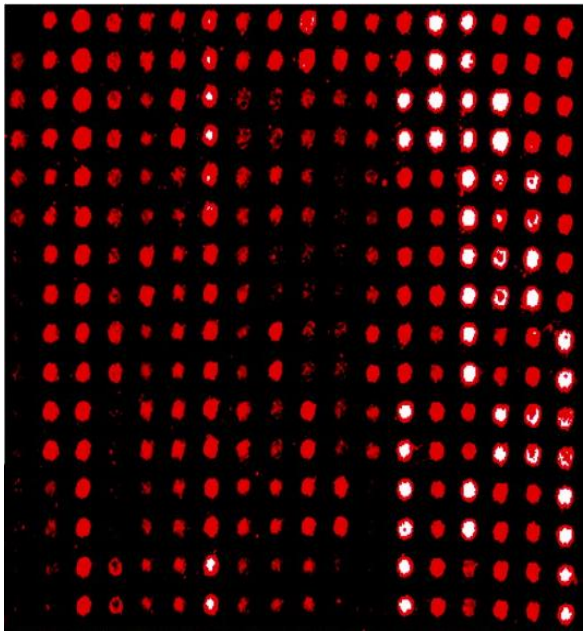
Dr. S. Sadeghi

Methods for Protein Analysis

- Gel electrophoresis, northern/western blot (fluorescence/radioactive label)
- X-ray crystallography
- 2D - mass spectrometry
- Protein microarrays
- SELDI MS protein chips

Protein Microarray

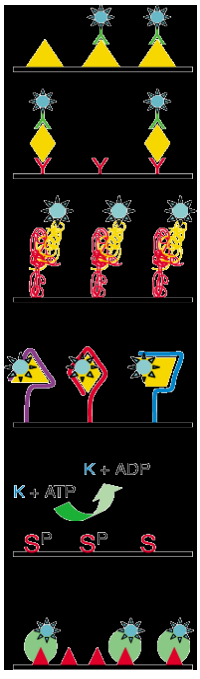
1. High throughput analysis of hundreds of thousands of proteins.
2. Proteins are immobilized on glass or other surface chip.
3. Various probes (protein, lipids, DNA, peptides, etc) are used.



Protein microarrays

- Different capture molecules must be used to study different interactions
- Examples
  - Antibodies (or antigens) for detection
  - Proteins for protein-protein interaction
  - Enzyme-substrate for biochemical function

REF: Benfey & Protopapas, 2005



Antigen–antibody

Protein–protein

Aptamers

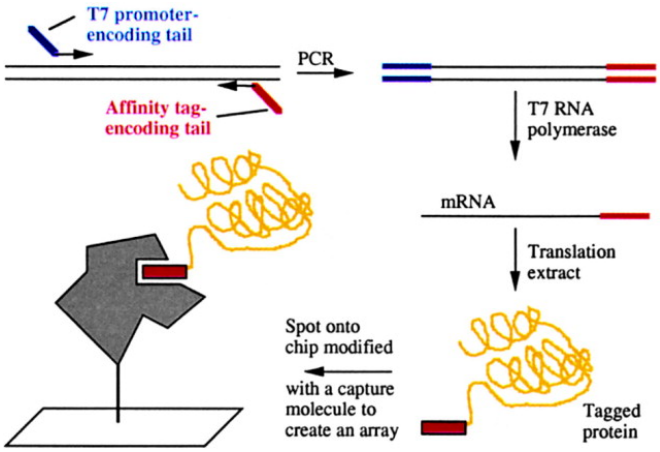
Enzyme–substrate

Receptor–ligand

Aptamers are oligonucleotide or peptide molecules that bind to a specific target molecule.

Protein function arrays

- PCR: amplify the gene of interest from cDNA using primers that would allow a T7 promoter to be incorporated 5′ to the gene and an affinity tag at the 3′ end.
- In vitro transcription/translation - desired protein containing the affinity tag.
- protein immobilized by spotting onto a glass slide impregnated with a molecule able to bind the affinity tag tightly.
- All other proteins in the translation mixture would be washed away.
- The figure depicts a small epitope (red box) that is recognized by a monoclonal antibody.



http://dx.doi.org/10.1016/S1074-5521(00)90067-X

Protein-protein interactions

Example of a model protein function microarray and its use in probing protein–protein interactions. Protein G, p50 and the FKBP-rapamycin binding (FRB) domain of FRAP were spotted onto a chemically derivatised glass slide, four spots each. The slide was then probed with various fluorescently labeled proteins.

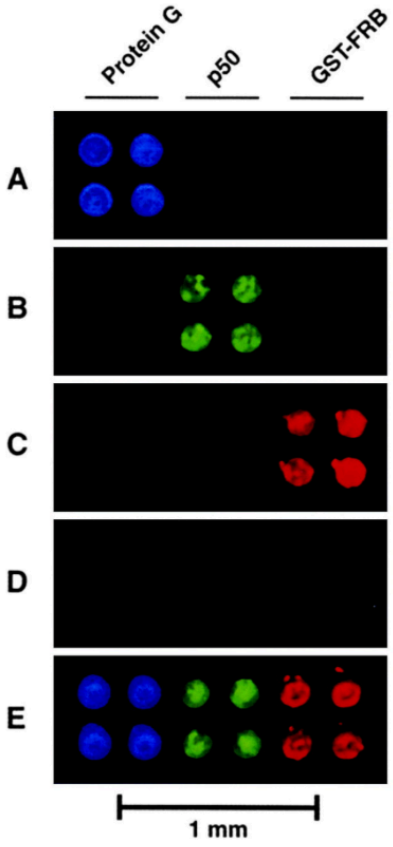
(A) Slide probed with BODIPY-FL-IgG, binds protein G.

(B) Slide probed with Cy3-labeled IkBa, binds to p50.

(C) Slide probed with Cy5-labeled FKBP12, known to bind FRB in the presence of rapamycin (included in the buffer).

(D) Same as (C), except rapamycin was not present.

(E) Slide probed with all 3 labeled proteins in the presence of rapamycin.



Activity-based protein profiling (ABPP)

a chemical strategy that utilizes active site directed covalent probes to profile the functional state of enzymes in complex proteomes.

Detection strategies for activity-based proteomics

Examples

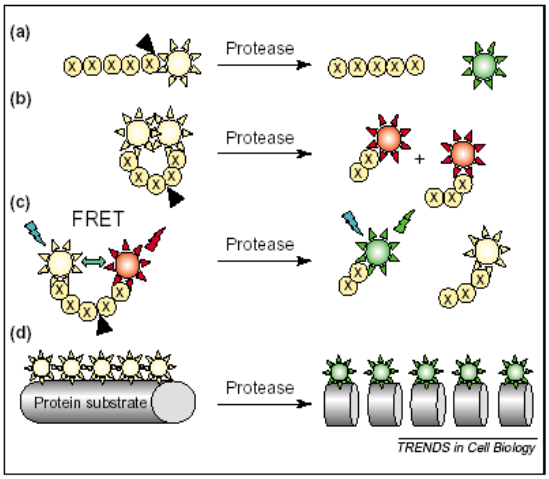
→1. Small-molecule substrate reporters of enzymatic activity

→ 2. Protein-based reporters of enzymatic activity

→ 3. Activity-based probe

1. Small-molecule substrate reporters of enzymatic activity

These reagents carry fluorescent groups, and thus energy emission upon their enzymatic conversion to product can be monitored over time



REF: TRENDS in Cell Biology, Vol.14 No.1 January 2004

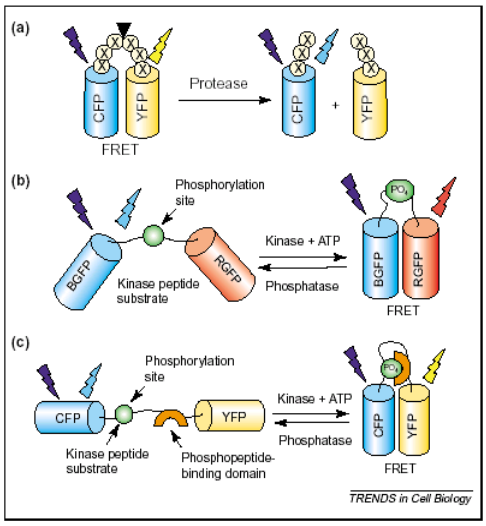
Disadvantages:

→1.The majority of basic fluorogenic probes cannot be directly applied to complex cellular environments

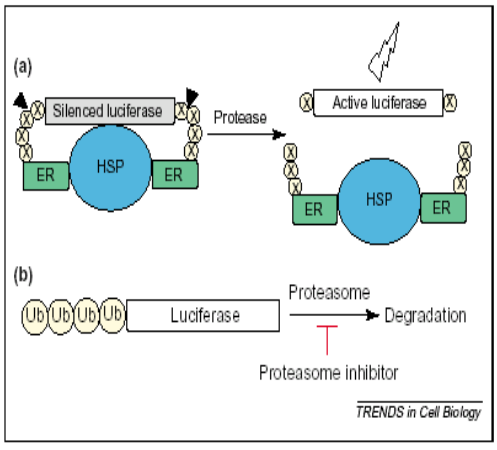
→2. Challenging to generate probes that are specific for an individual enzyme **(a peptide has the potential to function as a substrate for more than one class of proteolytic enzymes**)

2. Protein-based reporters of enzymatic activity

Fluorescent reporters



Bioluminescent reporters



TRENDS in Cell Biology, Vol.14 No.1 January 2004

Disadvantage:

They all suffer from the selectivity of probes for a specific enzyme target

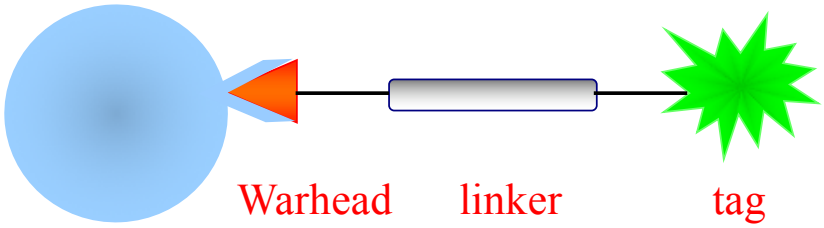
3. Activity-Based probe (ABP)

The activity-based probes (ABPs): they generally contain **three** main functional groups:

→1. The chemical reactive group or warhead (covalently modifies an active-site residue of the enzyme of interest)

→2. A linker region, which can be specific for different enzymes

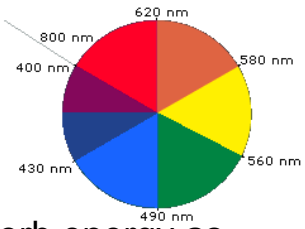
→3. A tag, which is used to visualize the modified enzyme



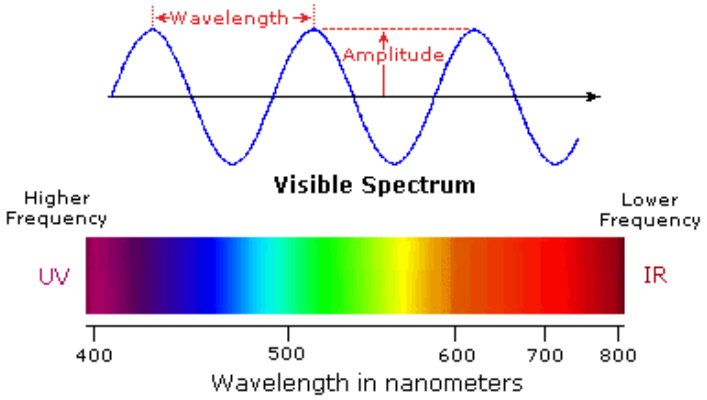
Tags used

- Antibodies
- Fluorescent tags
  - Markers that emit light at specific wavelengths where enhancement in optical signal indicates a binding reaction
    - GFP
- Luminescence tags
  - Light emission from a living organism
    - Luciferase

Absorption



- Basic quantum mechanics requires that molecules absorb energy as quanta (photons) based upon a criteria specific for each molecular structure
- Absorption of a photon raises the molecule from ground state to an excited state
- The structure of the molecule dictates the likelihood of absorption of energy to raise the energy state to an excited one



- **Violet:** 400 - 420 nm
- **Indigo:** 420 - 440 nm
- **Blue:** 440 - 490 nm
- **Green:** 490 - 570 nm
- **Yellow:** 570 - 585 nm
- **Orange:** 585 - 620 nm
- **Red:** 620 - 780 nm

Parameters

- Extinction Coefficient
  - ε refers to a single wavelength (usually the absorption maximum)
- Quantum Yield
  - Is a measure of efficiency of a reaction
  - A quantum yield of 1 indicates that 1 event occurs for every photon absorbed
- Quantum Yield

$$Q = \frac{\text{photons emitted}}{\text{photons absorbed}}$$

Fluorescence Lifetime (τ)

is the time delay between the absorbance and the emission

- Photon emission, as an electron returns from an excited state to ground state

Properties of fluorescent molecules

- Large extinction coefficient at the region of excitation
- High quantum yield
- Optimal excitation wavelength
- Photostability
- Excited-state lifetime
- Minimal perturbation by probe

FRET

Fluorescence resonance energy transfer (FRET) is a **distance-dependent** interaction between the electronic excited states of two dye molecules.

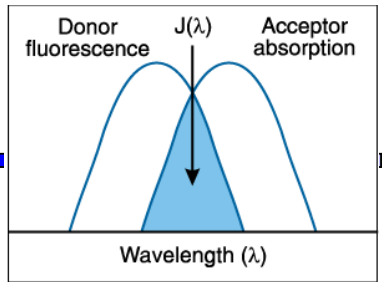
Excitation is transferred from a donor molecule to an acceptor molecule *without emission of a photon*.

FRET is also called Föster RET, as he was the first person to publish the theory behind this method (1946)

FRET is dependent on the inverse 6th power of the intermolecular separation, making it useful over distances comparable with the dimensions of biological macromolecules.

FRET is an important technique for investigating a variety of biological phenomena that produce changes in molecular proximity.

FRET

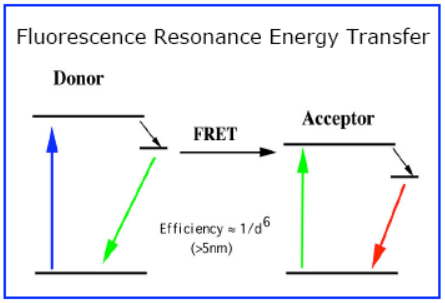


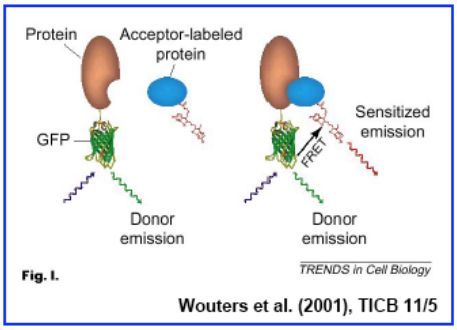
Primary Conditions for FRET

Donor and acceptor molecules must be in close proximity (typically 10–100 Å).

The absorption spectrum of the acceptor must overlap fluorescence emission spectrum of the donor (see figure).

Donor and acceptor transition dipole orientations must be approximately parallel.





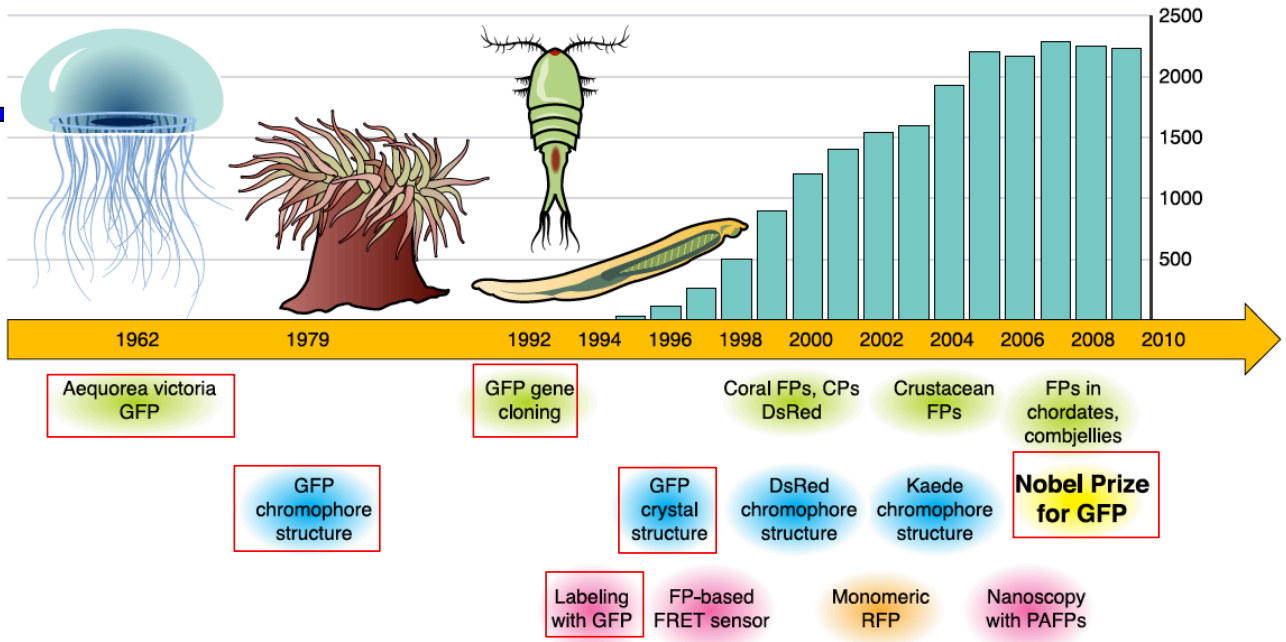
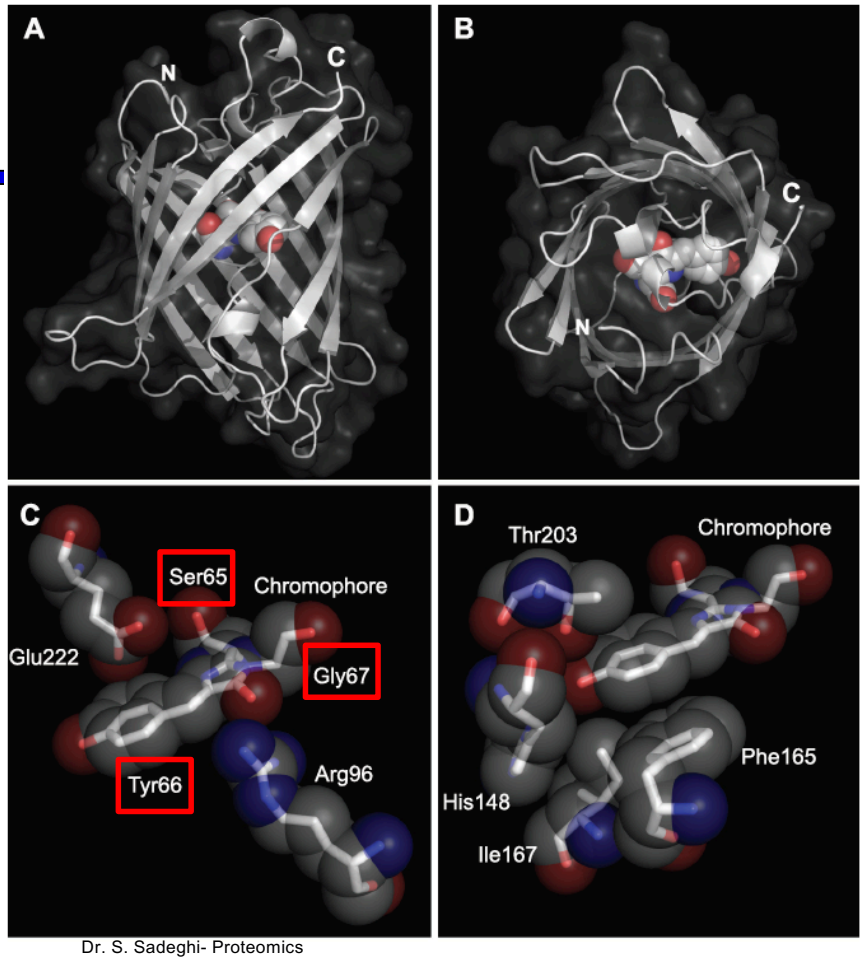


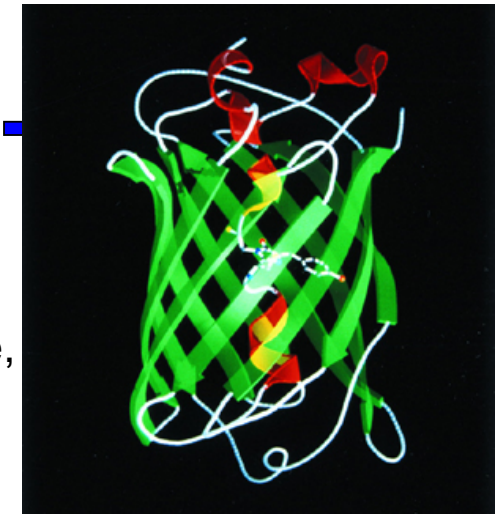
FIG. 1. Green color below the text highlights basic studies of natural diversity of green fluorescent protein (GFP)-like proteins; blue, structural insights; orange, development of novel fluorescent protein (FP) variants; magenta, appearance of FP-based technologies.
Columns above the timeline show the number of scientific articles per corresponding year that can be found searching PubMed with the term green fluorescent protein

Ref: Chudakov et al. 2010. Physiol Rev 90: 1103–1163

Structure of GFP.
A and B: overall structure of GFP -barrel with semitransparent surface is shown from the side (A) and from the top (B). Chromophore is shown in a spacefill representation. C and D: GFP chromophore and selected nearby residues in sticks and semitransparent spacefill representation. Carbon =gray, nitrogen= blue, and oxygen =red.

Ref: Chudakov et al. 2010. Physiol Rev 90: 1103–1163





- the tertiary structure is a barrel made from 11 beta sheets, capped with helices.
- At the centre of this lies the chromophore, a short chain of altered aminoacids (Ser65, Tyr66, Gly67) responsible for the light emission.
- The barrel structure keeps the chromophore away from solvents, making GFP capable of fluorescing under almost any conditions, being able to fluoresce nearly to the point at which the protein is denatured by heat and pH

Green Fluorescent Protein

GFP from the chemiluminescent jellyfish Aequorea victoria

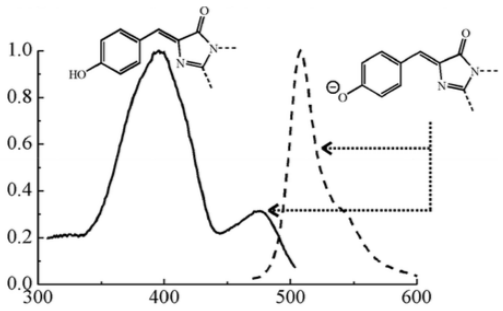
238 amino acids protein.

Excitation maxima at 395 and 470 nm

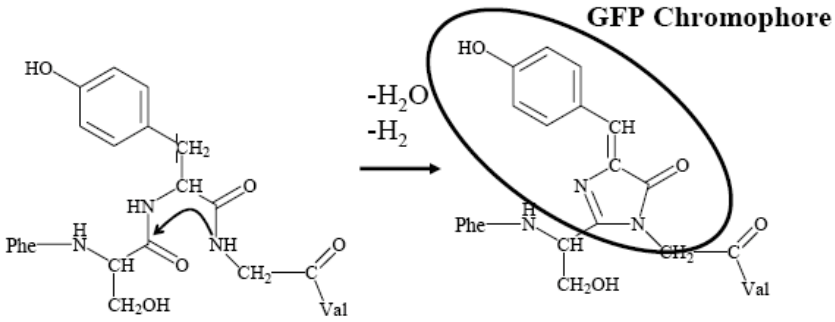
(quantum efficiency is 0.8)

Peak emission at 509 nm

The intrinsic fluorescence of the protein is due to a unique covalently attached chromophore, which is formed translationally within the protein upon cyclisation and oxidation of residues 65-67, Ser-Tyr-Gly.



Fluorescence excitation (full-line curve) and emission (dashed curve) spectra of native GFP from *Aequorea victoria* (Tsien et al., 1998).



The Nobel Prize in Chemistry 2008

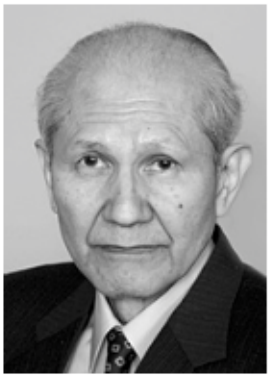


Photo: U. Montan

Osamu Shimomura

Prize share: 1/3

Photo: U. Montan

Martin Chalfie

Prize share: 1/3

Photo: U. Montan

Roger Y. Tsien

Prize share: 1/3

The Nobel Prize in Chemistry 2008 was awarded jointly to Osamu Shimomura, Martin Chalfie and Roger Y. Tsien *"for the discovery and development of the green fluorescent protein, GFP".*

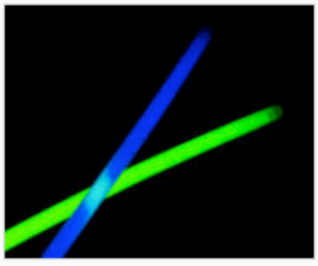
The cover of *Science* for February 11, 1994 showing GFP in *C. elegans* neurons (Chalfie et al.).

Luminescence

Chemiluminescence

- Emission of light (luminescence) without emission of heat as a result of a chemical reaction.
- The decay of the excited state to a lower energy level is responsible for the emission of light.
- In theory one photon of light should be given off for each molecule of reactant.
- One of the oldest reactions
  - White phosphorus oxidising in moist air, producing a green glow.

A chemoluminescent reaction carried out in an erlenmeyer flask producing a large amount of light.



Bioluminescence

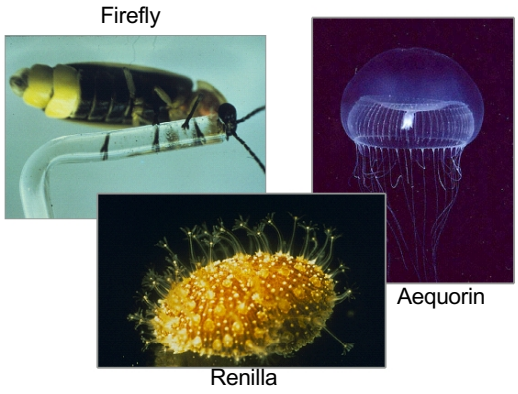
- Bioluminescence is a natural phenomenon in which organisms produce and emit light via an enzyme-mediated biochemical process.
- This visible light is “cold”- meaning that, unlike other sources, light generated by bioluminescence does not give off much energy in the form of heat.
- The performance and capabilities of bioluminescence can be extended by engineering the luciferases and luciferins

Luciferase enzyme

- Luciferase is the general name applied to the class of enzymes that are responsible for converting chemical energy into light energy.
- The substrate molecule upon which luciferase acts is called a luciferin.
- Widely distributed in nature
  - Diverse set of organisms:
    - Bacteria, fungi, fish, insects, shrimp etc
- There are 5 basic luciferase systems
  - Two are studied in detail
    - Beetle (firefly) Gene= *luc* yellow light 560nm
    - Bacteria Gene=*luxA/B* blue/green light 490nm

Bioluminescence used in biological assays

- Firefly luciferase (beetle luciferases)
  - Reporter gene
    - most common application
  - Many new applications emerging
- Renilla luciferase
  - Reporter gene - mostly in dual repo assays
- Bacterial luciferase
  - Reporter gene - mostly in bacterial hosts
- Aequorin
  - Intracellular calcium



General reaction

Luciferin + Luciferase + ATP + Mg^{2+}

↓

Luciferyladenylate-luciferase + Pyrophosphate

Luciferyladenylate-luciferase + O_2

↓

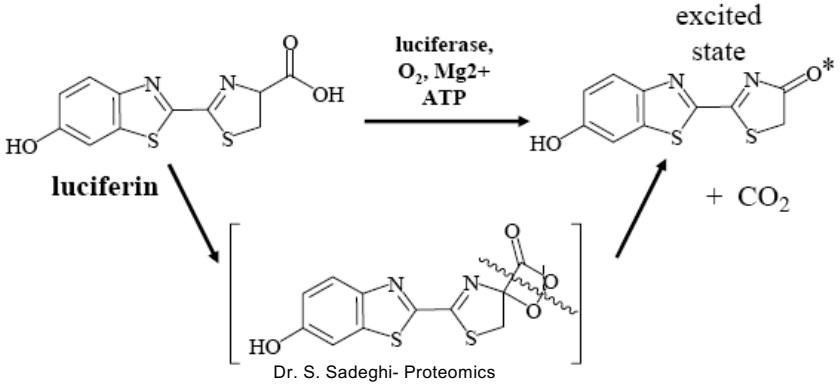
Luciferase + Oxyluciferin + AMP + Light + CO_2

Bioluminescence

- firefly luciferase:
- is a monomeric 61 kDa enzyme
- Found in the light-emitting organ known as lantern within the abdomen.
- Firefly emits flashes of light to attract its mate.
- Luciferase has many characteristics that make it ideal for a reporter:
  - Its activity is not dependent on any post-translational modification, making it immediately available for quantitation.
    - GPF requires minutes to form chromophore
  - The luminescent is very bright, having the highest quantum efficiency for any bioluminescence reaction
    Quantum yield=0.9

Firefly luciferase

- Firefly luciferase catalyses a two-step oxidation of luciferin, which yields light at 560 nm.
- The first step involves the activation of the protein by ATP to produce a reactive mixed anhydride intermediate.
- In the second step, the active intermediate reacts with oxygen to create a transient dioxetane, which quickly breaks down to the oxidised product oxyluciferin and carbon dioxide along with a burst of yellow light.

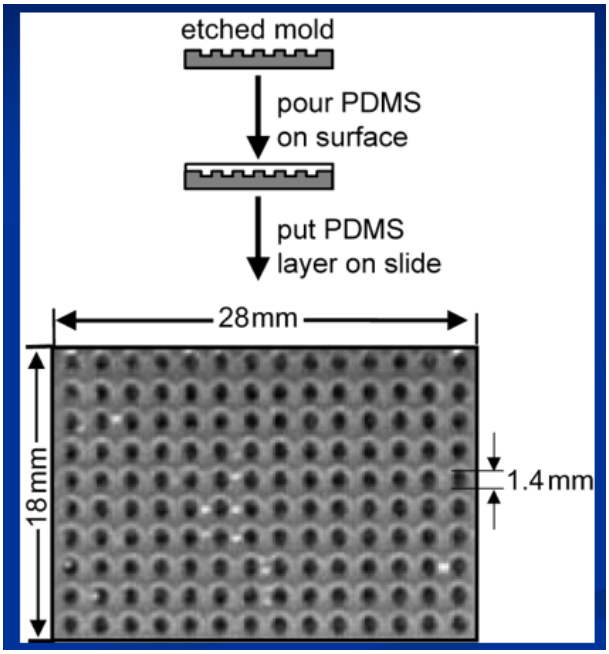


Protein Array

1. Surface fabrication
2. Protein immobilisation
3. Detection

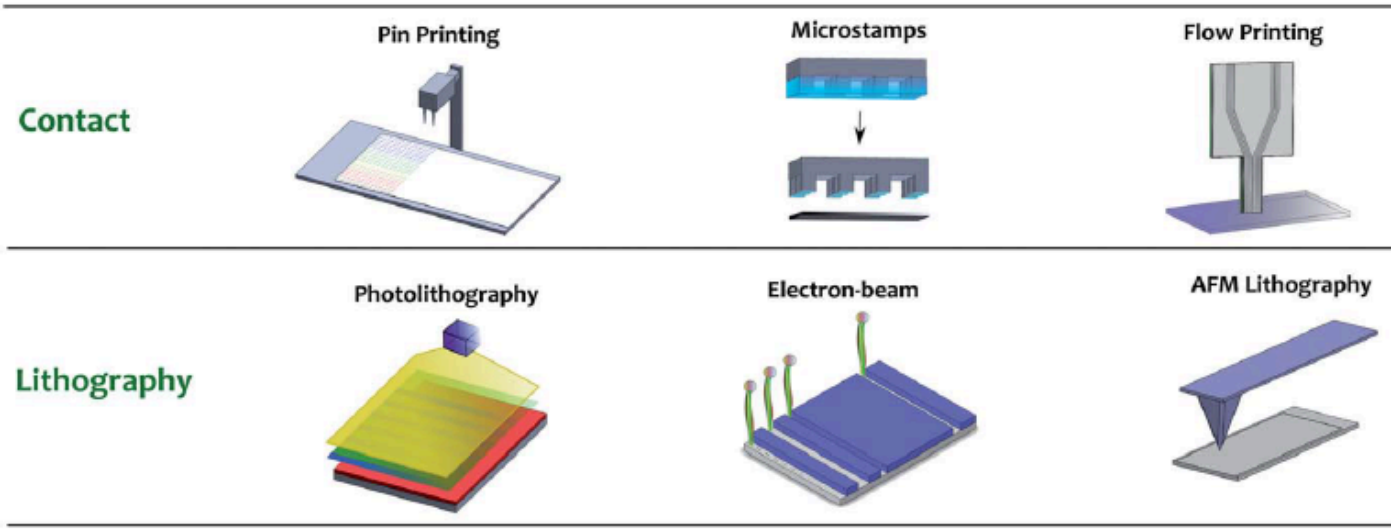
1. Protein Array Fabrication

- Protein substrates
  - Commonly used physical supports include glass slides, silicon, microwells, nitrocellulose or PVDF membranes,
- Magnetic microbeads.
- Proteins deposited on chip surface by robots



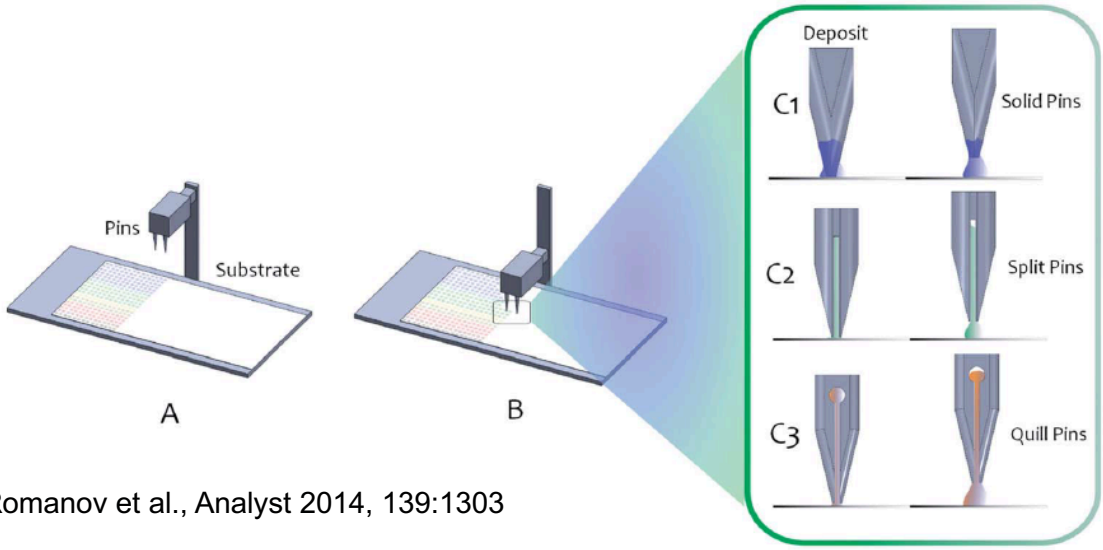
Benfey & Protopapas, 2005

Protein Array Fabrication



Ref: Romanov et al., Analyst 2014, 139:1303

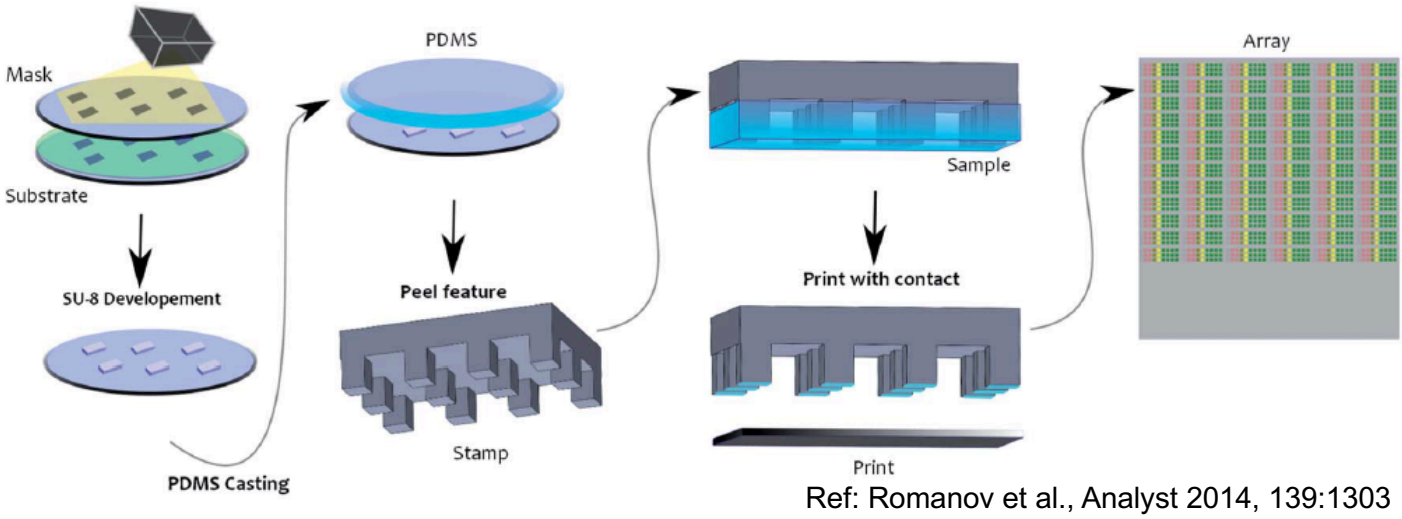
Pin Printing



Ref: Romanov et al., Analyst 2014, 139:1303

(A) A robotic print head with multiple printing pins is loaded with print solutions from a source plate and then contacts the substrate surface to deposit protein solution in (B). Various types of pins: C1 is a solid pin. C2 is a slotted pin. C3 is a quill pin, distinguished from the split pin by the inclusion of a reservoir.

Microstamping

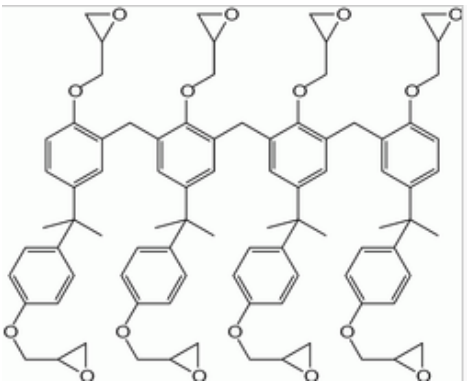


Ref: Romanov et al., Analyst 2014, 139:1303

After the creation of a mask and subsequent exposure onto the substrate, features are developed with the aid of SU-8.
Conformal sealing with PDMS creates chambers in the final product where SU-8 created raised indentations. Once the stamp is created, sample may be transferred onto its surface for subsequent printing

SU-8 photoresist

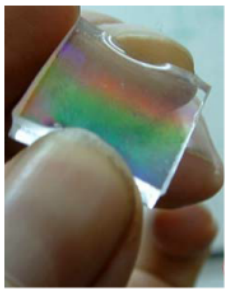
- SU-8 is a commonly used epoxy-based negative photoresist. Negative refers to a photoresist whereby the parts exposed to UV become cross-linked, while the remainder of the film remains soluble and can be washed away during development.
- SU-8 derives its name from the presence of 8 epoxy groups. It is these epoxies that cross-link to give the final structure.



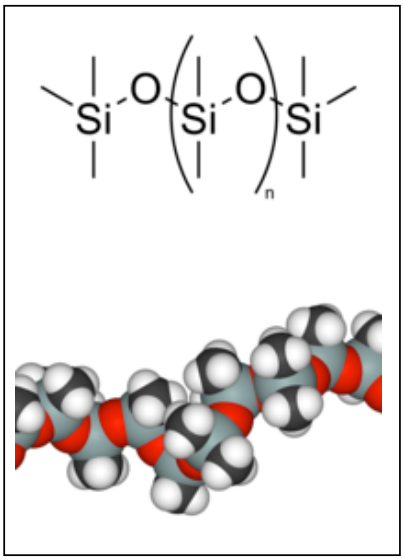
Micro stamp

- Hundreds of spots can be printed in parallel, leading to high-throughput microarray fabrication.
- The process is:
  - inexpensive + simple;
    - Sample is adsorbed on the patterned surface of stamp
    - Sample is transferred to substrate by physical contact
  - For good contact, stamps are made of *elastomeric* material
    - These compounds conform to surface roughness under an applied load

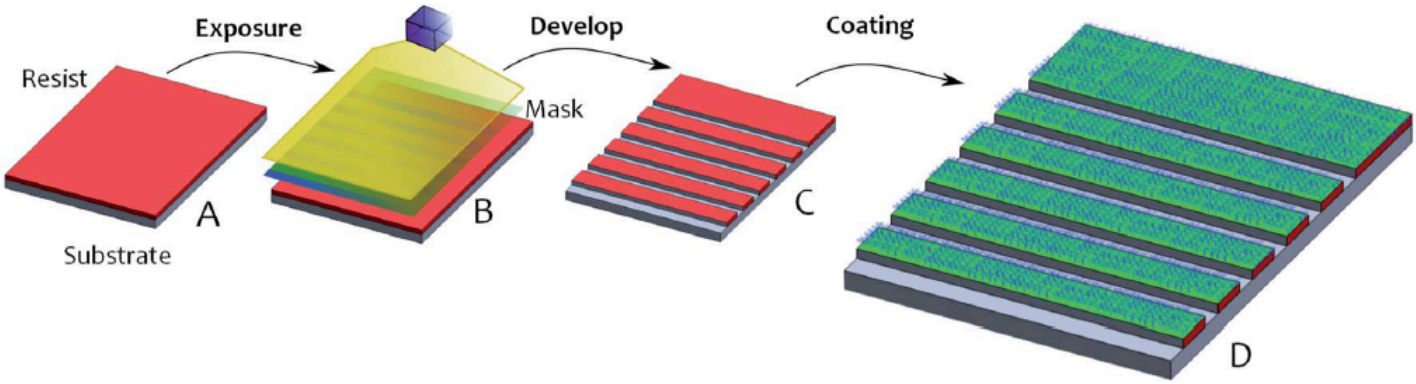
PDMS



- **Polydimethylsiloxane** (PDMS) is the most widely used silicon-based organic polymer.
- Its applications range from contact lenses and medical devices to elastomers, lubricating oils and heat resistant tiles.
- PDMS is optically clear, inert, non-toxic and non-flammable.
- PDMS has a flexible polymer backbone due to its siloxane linkages.
- Such flexible chains become loosely entangled when MW is high, which results in PDMS having an unusually high level of viscoelasticity.



Photolithography

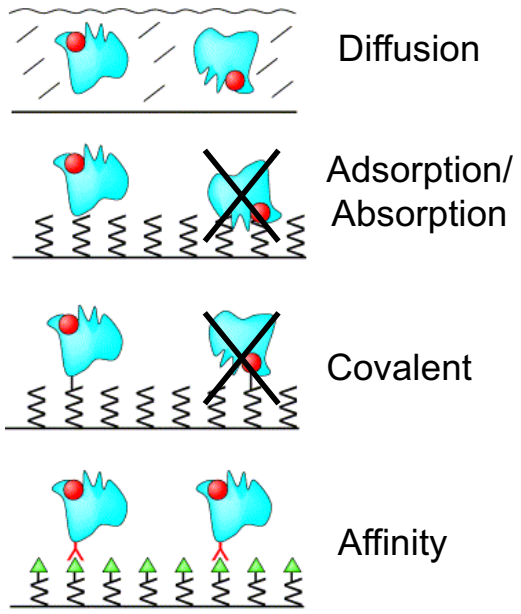


Ref: Romanov et al., Analyst 2014, 139:1303

Schematic of Photolithography.
(A) Undeveloped photoresist (red).
(B) Photoresist is exposed to light (yellow) through the photomask.
(C) Development removes the exposed, softened photoresist and a nanopatterned photoresist is generated.
(D) Microarray is generated by attachment of proteins to patterned photoresist.

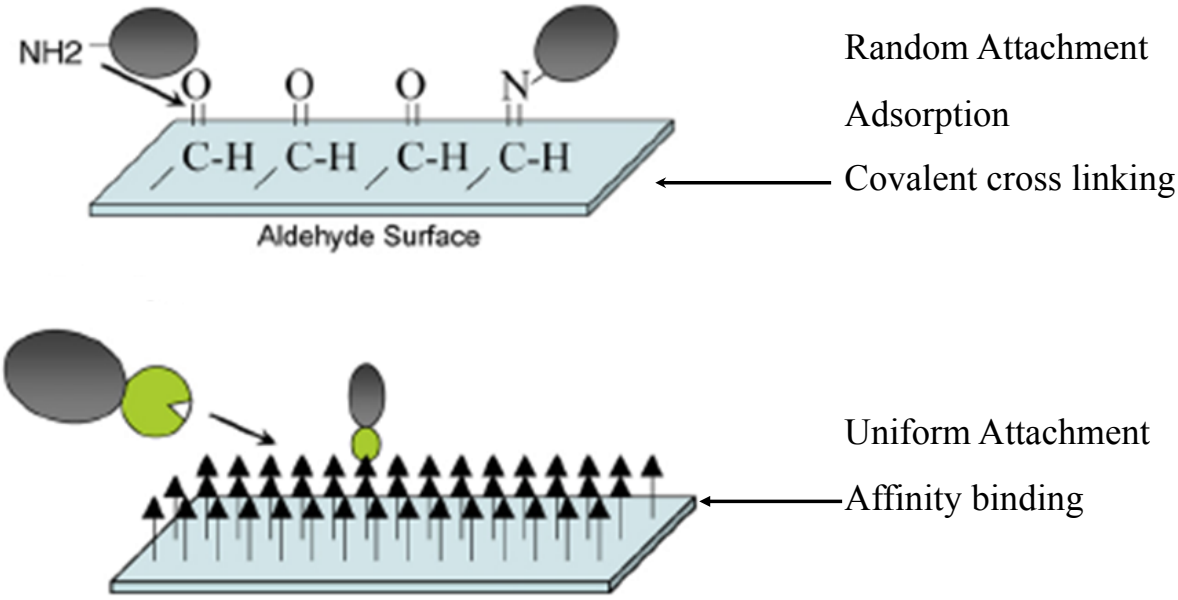
2. Protein Attachment / Immobilisation

- Diffusion
  - Protein suspended in random orientation, but presumably active
- Adsorption/Absorption
  - Some proteins inactive
- Covalent attachment
  - Some proteins inactive
- Affinity
  - Orientation of protein precisely controlled



REF: Benfey & Protopapas, 2005

Random and Affinity Attachment



REF: Zhu, H; Snyder, M: *Current Opinion in Chemical Biology*, **2003**, *7*, 55-63

3. Detection

- The preferred method of detection is fluorescence detection.
  - is compatible with standard microarray scanners,
  - the spots on the resulting image can be quantified by commonly used microarray quantification software packages.
  - some minor alterations to the analysis software may be needed.
- Other common detection methods include
  - colorimetric techniques
  - chemiluminescent
  - Surface Plasmon Resonance

Examples of a protein biochip

Bunger et al., BMC Cancer 2012, 12:393

CEA- a colon cancer biomarker

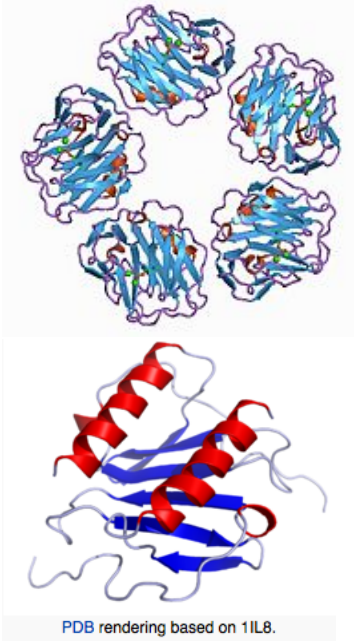
- More than 1.2 million new cases of colorectal cancer are reported each year worldwide.
- Despite actual screening programs, about 50% of the patients are diagnosed at advanced tumor stages presenting poor prognosis.
- Innovative screening tools could aid the detection at early stages and allow curative treatment interventions.
- A nine target multiplex serum protein biochip was generated and evaluated using validation-set of 317 highly standardized, liquid nitrogen preserved serum samples comprising controls, adenomas (benign tumor), and colon cancers.

Ref: Bunger et al., BMC Cancer 2012, 12:393

- The aim of this study was to apply biochip array technology to colon cancer screening.
- first study to report the development of a serum biochip array for the simultaneous assessment of nine serum biomarkers for clinical application to colon cancer screening in a large serum sample cohort.
- A biochip array was designed and developed for the multiplex determination of 9 serum markers allowing for low inter-analysis variability, decreased workload and faster processing time as well as lower costs due to high-throughput automation.
- The performance of the biochip arrays for colon cancer screening was evaluated in a validation set consisting of 317 highly standardized, liquid nitrogen preserved serum samples.

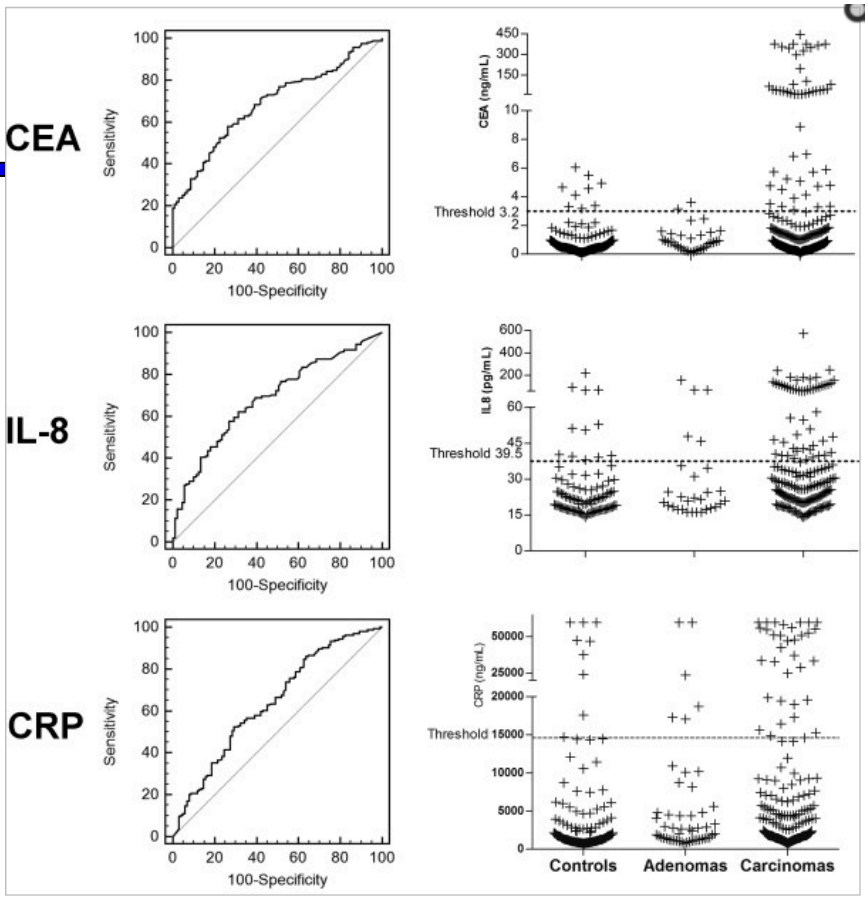
3 of the main biomarkers

- Carcinoembryonic antigen (CEA) describes a set of highly related glycoproteins involved in cell adhesion.
  - CEA is normally produced in gastrointestinal tissue during fetal development, but the production stops before birth. Therefore, CEA is usually present only at very low levels in the blood of healthy adults.
  - serum levels are raised in some types of cancer, which means that it can be used as a tumor marker in clinical tests. CEA critical to the metastatic dissemination of colon carcinoma cells.
- C-reactive protein (CRP) is an annular (ring-shaped), pentameric protein found in blood plasma, whose levels rise in response to inflammation.
- Interleukin 8 (IL-8) is a chemokine produced by macrophages
  - implied to have a role in colorectal cancer by acting as an autocrine growth factor for colon carcinoma cell lines or the promotion of division and possible migration
  - Potent promoter of angiogenesis



PDB rendering based on 1IL8.

Biochip revealed combinations of CEA + IL-8 and CEA + CRP to show the best screening performance for colon cancer with 47% sensitivity, for early carcinomas with 33% sensitivity, and adenomas with 18% sensitivity at an overall specificity of 86%.



Dr. S. Sadeghi- Proteomics

"Global Analysis of Protein Activities Using Proteome Chips"

Yale University

Ref: Science, 2001, Vol 293, 2101-2105

"Global Analysis of Protein Activities Using Proteome Chips"

Snyder Lab, Yale University, *2101-2105, Vol 293, Science, 2001*

Objectives

1. Construct yeast proteome chip containing 80% of yeast proteins in high throughput manner.
2. Study protein interactions at cell level using the proteome chip.

Protein Immobilization on Surface

1. Cloning of 5800 ORFs.
2. Production of fusion proteins (GST- HisX6).
3. Printing on glass chip.
4. Verification by anti-GST.

"Global Analysis of Protein Activities Using Proteome Chips"

Snyder Lab, Yale University, *2101-2105, Vol 293, Science, 2001*

Protein-Lipid Interactions

1. Phospholipids-Biotin.
2. About 150 proteins interacted with phospholipid probes.
3. Several of them were unknown, and some related to glucose metabolism.

SELDI MS-based ProteinChip

- Utilizes *Surface Enhanced Laser Desorption/Ionization Mass Spectrometry* (1993)
- MALDI MS combined with chromatography (Bioaffinity): surface-MALDI

SELDI

- SELDI-TOF-MS is a variation of matrix-assisted laser desorption/ionization (MALDI) that uses a target modified to achieve biochemical affinity with the analyte compound.
- In MALDI, a protein or peptide sample is mixed with the matrix molecule in solution and small amounts of the mixture are deposited on a surface and allowed to dry. The sample and matrix co-crystallize as the solvent evaporates.
- In SELDI, the protein mixture is spotted on a surface modified with a chemical functionality. Some proteins in the sample bind to the surface, while the others are removed by washing.
- After washing the spotted sample, the matrix is applied to the surface and allowed to crystallize with the sample peptides.
- Binding to the SELDI surface acts as a separation step and the subset of proteins that bind to the surface are easier to analyze.

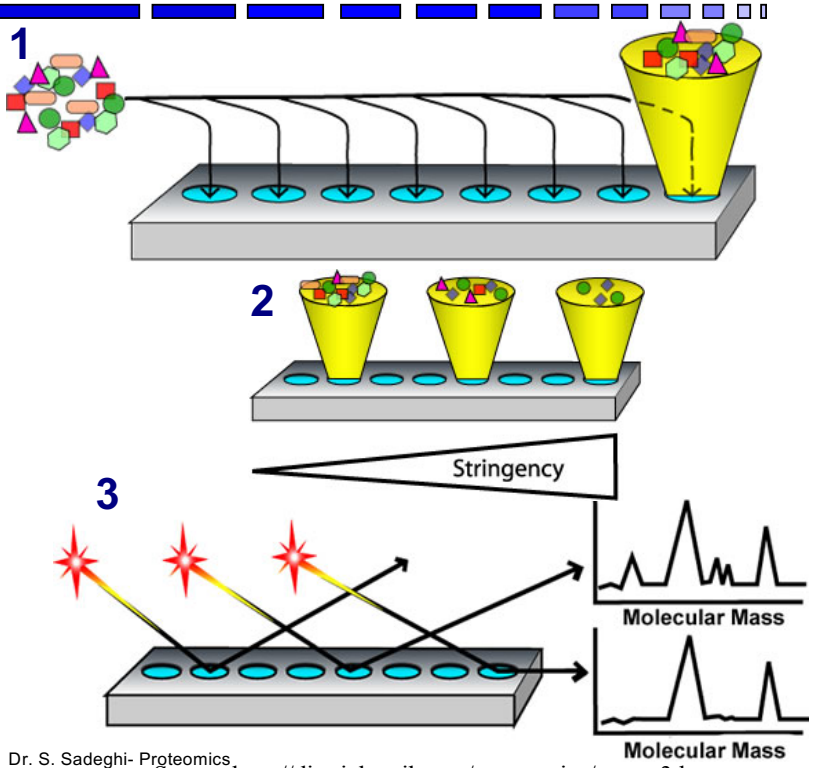
- Samples spotted on a SELDI surface are typically analyzed using time-of-flight mass spectrometry TOF.
- A laser ionizes peptides from crystals of the sample/matrix mixture.
- The ions are accelerated through an electric potential and down a flight tube.
- A detector measures ions as they reach the end of the tube.
- The mass-to-charge ratio of each ion can be determined from the length of the tube, the kinetic energy given to ions by the electric field, and the time taken to travel the length of the tube.

Protein Analysis by SELDI-MS

1) Apply sample (serum, tissue extract, etc.) to ProteinChip® array.

2) Wash sample with increasing stringency to remove non-specific proteins.

3) Energy absorbing molecules are added to retained proteins. Following laser desorption and ionization of proteins, Time-of Flight (TOF) mass spectrometry accurately determines their masses



Source:http://dir.niehs.nih.gov/proteomics/emerg3.htm

Advantages & Applications of SELDI MS

- Extraction, fractionation, clean-up and amplification of samples on surface
- High throughput, high level multiplexing
- Large scale/ Low sample volume
- High sensitivity
- Various molecules on surface to capture probes
- Discovery of protein biomarkers
- Purification of target proteins

"SELDIProteinChip Array Technology: Protein-Based Predictive Medicine and Drug Discovery Applications"

Ref: Ciphergen Biosystems Inc. J. Biomed. & Biotechnol., 2003, Vol 4, 237-241

Application 1:
Identification of HIV Replication Inhibitor

Some individuals infected with HIV-1 remain clinically stable for years:
CD8+ of these individuals secrete a soluble factor= CAF
Supresses HIV-1 replication

1. CAF (CD8+ antiviral factor) thought to be related to AIDS development
2. Determined the identity of CAF with SELDI techniques: alpha-defensin -1, -2 and -3
3. Demonstrated de novo discovery of biomarker and multimarker patterns, identification of drug candidates and determination of protein functions

Application 2:
Multimarker Clinical Assays for Cancer

1. Early detection of cancer – critical in effective cancer treatment
2. Cancer biomarker – massive protein expression profiling
3. High throughput assay for multimarker provided by SELDI array and multivariate software algorithms produced high sensitivity and specificity.

Data from three different research groups

- Combining 9 different proteins generates an assay with better sensitivity (83%) and specificity (97%) for diagnosis of prostate cancer than the prostate specific antigen (PSA) test.

- Three newly discovered biomarkers for breast cancer used in combination, the SELDI assay has a significantly higher sensitivity (93%) and specificity (91%) relative to CA15.3, the best available protein marker.

- For ovarian cancer, a SELDI multimarker profile has sensitivity (100%) and specificity (95%), compared to the poor performance of the CA125 test.

Application 3:
Biomarker and Drug Discovery
Applications in Neurological Disorders

1. SELDIProteinChip for Alzheimer's Disease
2. Wide range of samples
   Small sample amount
3. SELDI using antibody protein array : Ab against N-terminal sequence of target peptides (beta-amyloid)
4. Discovered candidate biomarkers, related inhibitors and their functions and peptide expression levels

- As the most prevalent form of neurodegenerative disorders, AD affects about 4 million people in the USA, generally people over the age of 65.
- Clinical features of AD include beta-amyloid deposits in brain, neuritic plaques, and degeneration of synapses
- All of the beta-amyloid peptides share a common N-terminal sequence, so an antibody that is specifically raised against this N-terminal sequence can be immobilised on the ProteinChip Array and used to capture beta-amyloid peptides from complex samples.
- The SELDITOF-MS analysis of such a capture can be used to monitor the relative amounts of peptides of various lengths, including many from beta-amyloid 1–15 to 1–42, some of which correlate more strongly with the development of AD than others.

Finally, Protein chips

1. **Diagnostics:** detection of antigens and antibodies in blood samples; profiling of sera to discover new disease markers;

2. **Proteomics:** protein expression profiling; organ and disease specific arrays.

3. **Protein functional analysis:**
protein-protein interactions;
ligand-binding properties of receptors; enzyme activities;
antibody cross reactivity and specificity, epitope mapping.

Learning outcome

- From this lecture you should understand:
  - Protein chip:
    - Fabrication
    - Type of tags used
      - GFP
      - Luciferase
  - Examples of protein chips:
    - Cancer research
    - Protein activity
  - Protein analysis by SELDI-MS
    - Examples of its applications.

References

- Benfey & Protopapas, 2005
- Bunger et al., BMC Cancer 2012, 12:393
- TRENDS in Cell Biology, Vol.14 No.1, 2004
- Chalfie et al. Science 1994
- Ciphergen Biosystems Inc. J. Biomed. & Biotechnol., 2003, Vol 4, 237-241
- Chudakov et al. 2010. Physiol Rev 90: 1103–1163
- Tsien et al., 1998
- Romanov et al., Analyst 2014, 139:1303
- Zhu, H; Snyder, M: Current Opinion in Chemical Biology, **2003**, *7*, 55-63
- Science, 2001, Vol 293, 2101-2105