PROTEIN CHIPS

FROM GENOMICS TO PROTEOMICS

PROTEOME is the group of protein encoded by the genome.

The study of proteome is known as **PROTEOMICS**

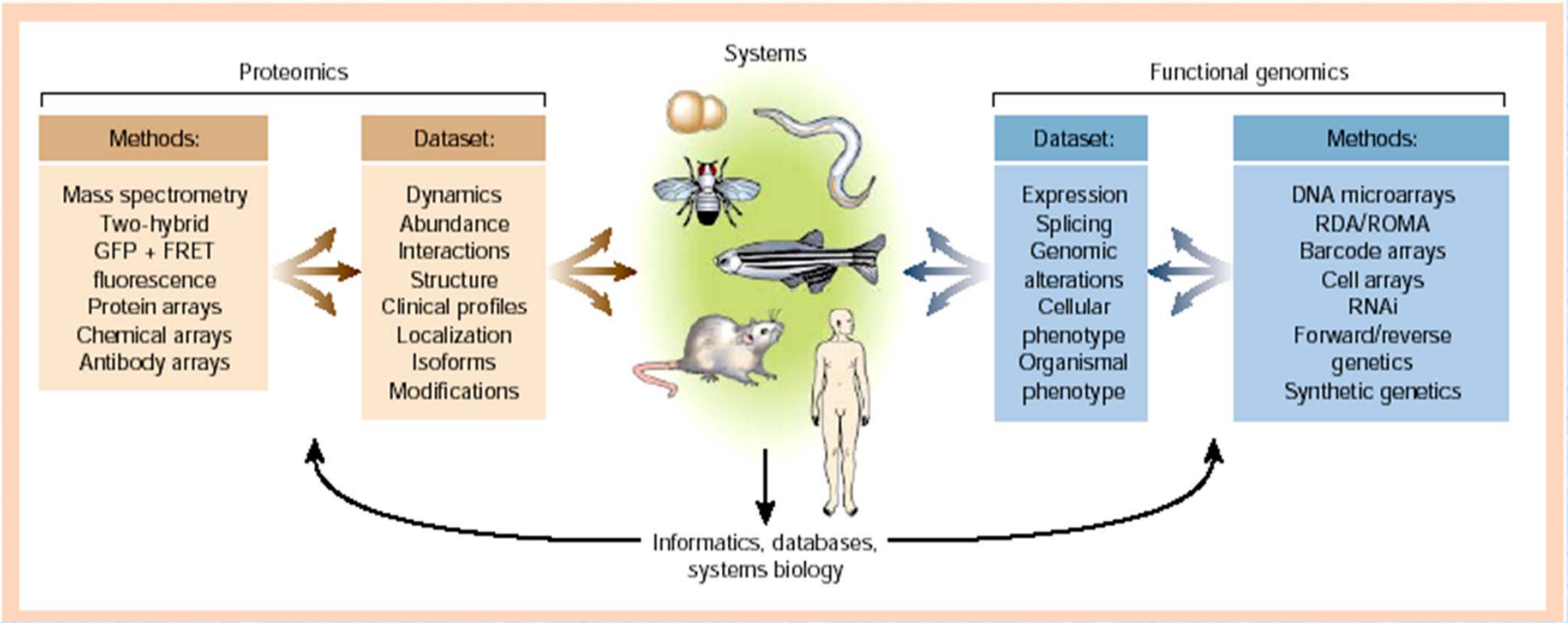
- Proteins from any cell
- Isoforms
- Modified proteins
- Interactions among proteins
- Structural description of proteins
- Structural description of protein complexes
- Moreover: any post-genome thing

High-throughput Biochemistry

Proteomics (nowadays): Protein biochemistry on a great scale, high precision and performance.

Objetives: Full description of celular function

Platforms for proteomics and functional genomics



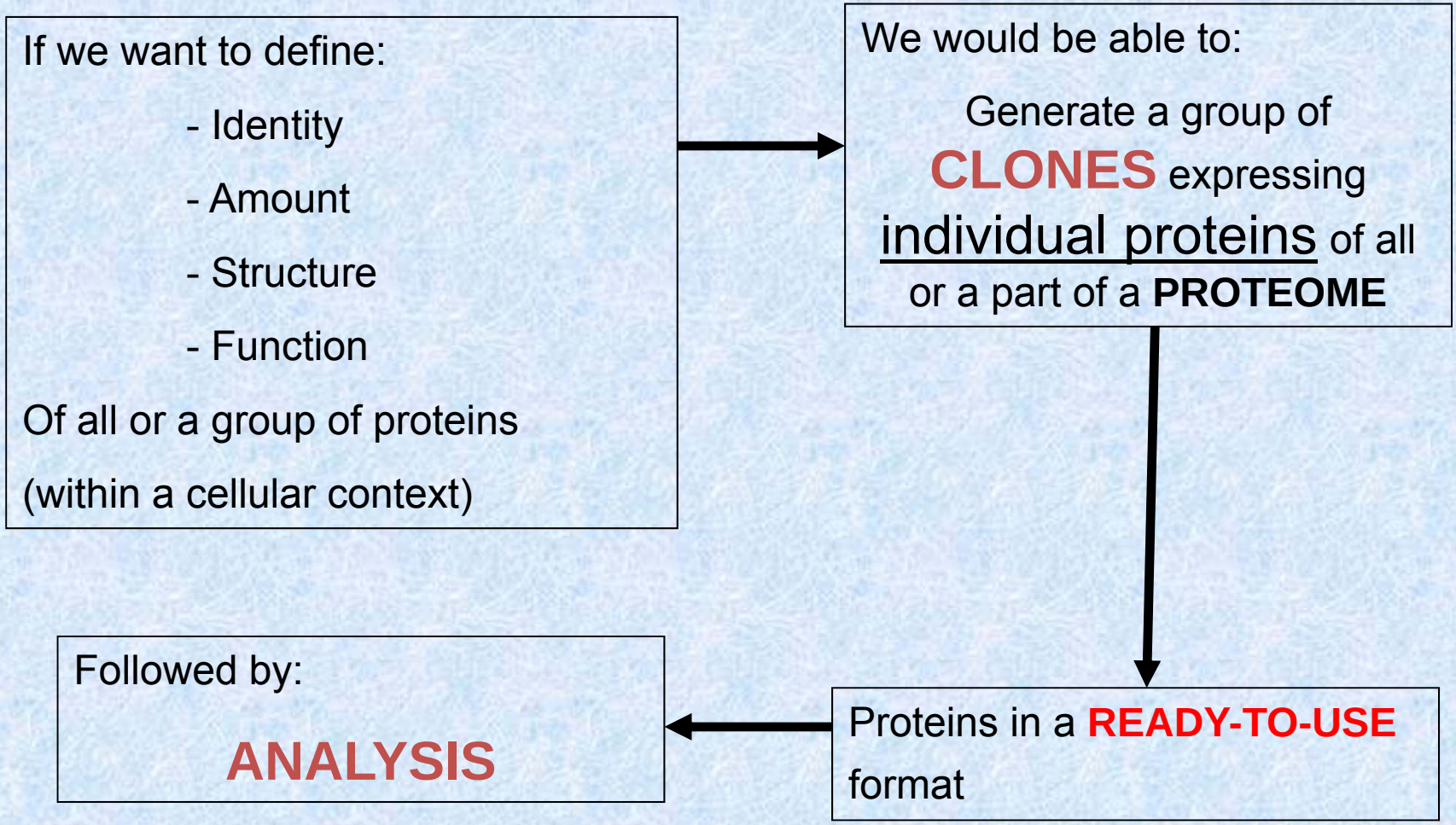
Proteomics scalability

Genomics: High scalability → PCR, automatic sequencers, etc.

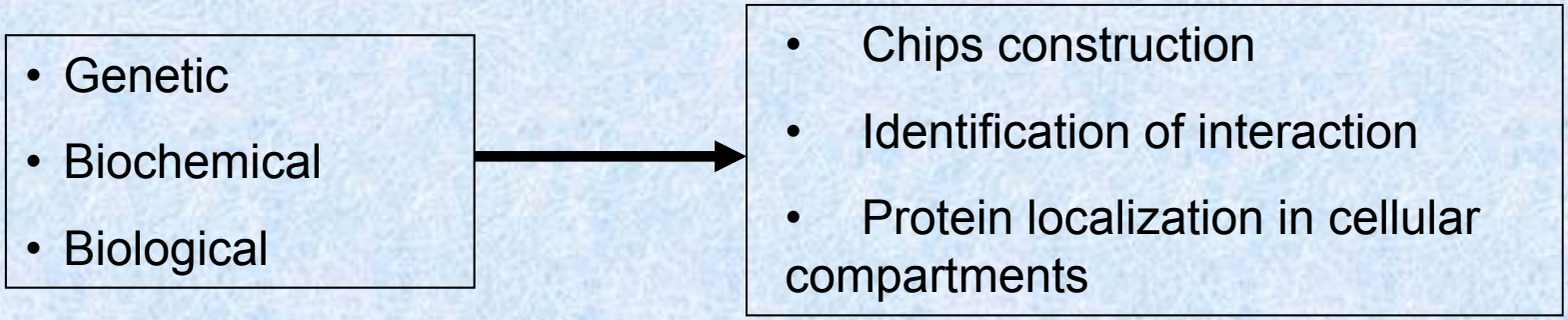
Proteomics: LOW scalability. Reasons:

- Variable and limited material
- Sample degradation
- Very high dynamic range abundance (10^6)
- Frequent post-translational modifications
- Temporal and on-going specificity
- Pathological disruptions
- Pharmacological disruptions

Protein analysis at proteomic scale



NECESSARY TECHNIQUES



1. **Expression and purification of proteins**
2. **Activity tests at proteomic scale**

a) Biochemical genomic approximation

b) Microarrays
- Analytic
- Functional

1. Expression and protein purification

ORF Localization

(open reading frame)

→

Problems:

- Difficult localization (sometimes)
- Splicing signals
- Polyadenylation

ORF Clonation

→

Problems:

- It is a simplification
- A mRNA is selected and isoforms are discarded
- Post-translational modifications (phosphorilation, glycosylation, methylation, acetylation) should not be considered
- Inadequate for membrane proteins

ORF Clonation

Requires:

- PCR primers
- Vector insertion

→

HT synthesis methods:

- Automatic synthesis
- Reasonable cost
- Speed and precision

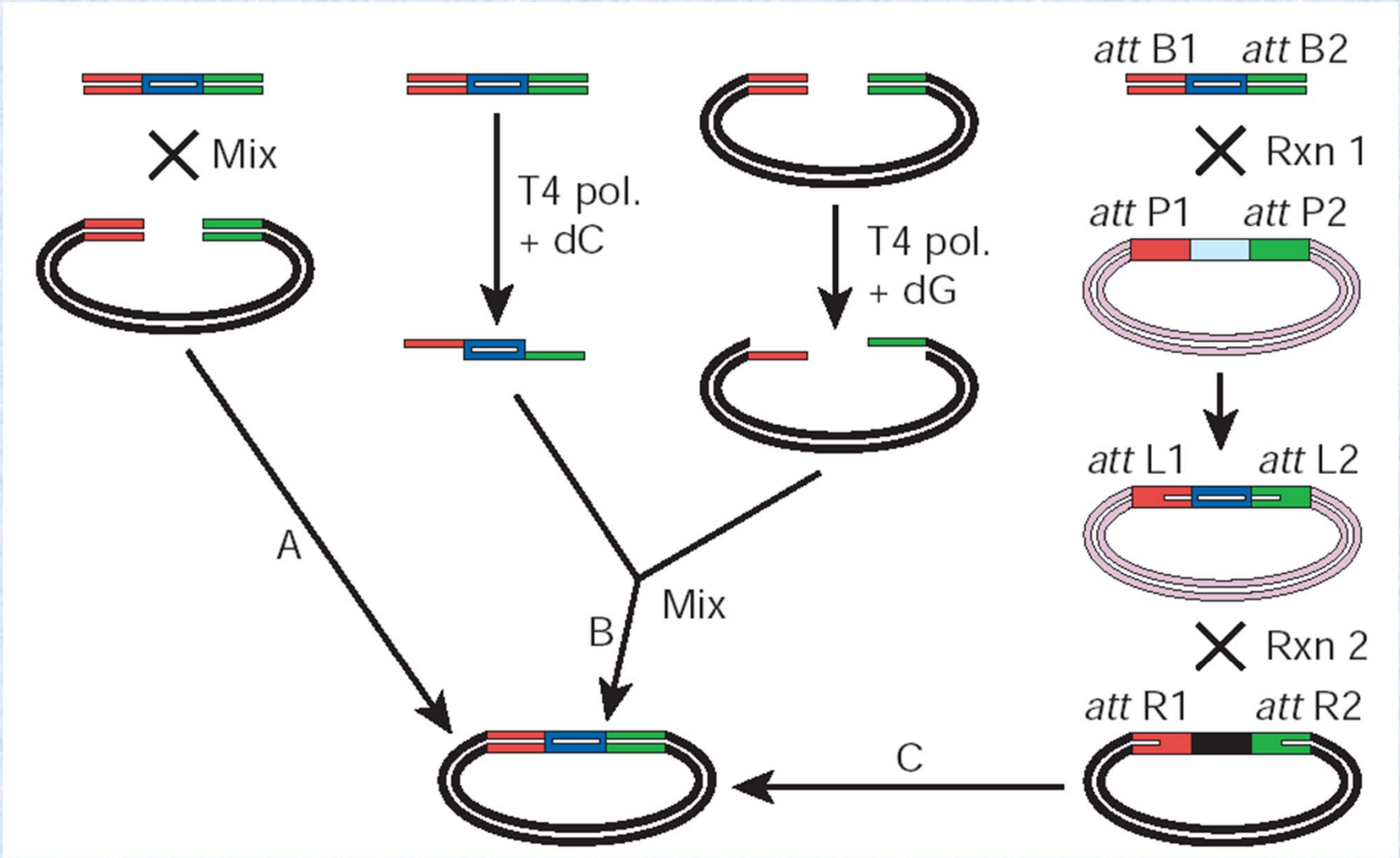
Vector insertion

A) Recombination mediated by gap repairing

B) Ligation independent clonage (T4 polimerase + dCTP or dGTP)

C) Invitrogen Gateway System (from phage λ integration/split)

Vector insertion



Expression

Homologue systems → The best (post-translational modifications, interactions with protein pairs, etc)

Heterologue systems → Necessary for a great amount of organisms. One alternative is insect cells (similar modifications to mammals)

Marker incorporation

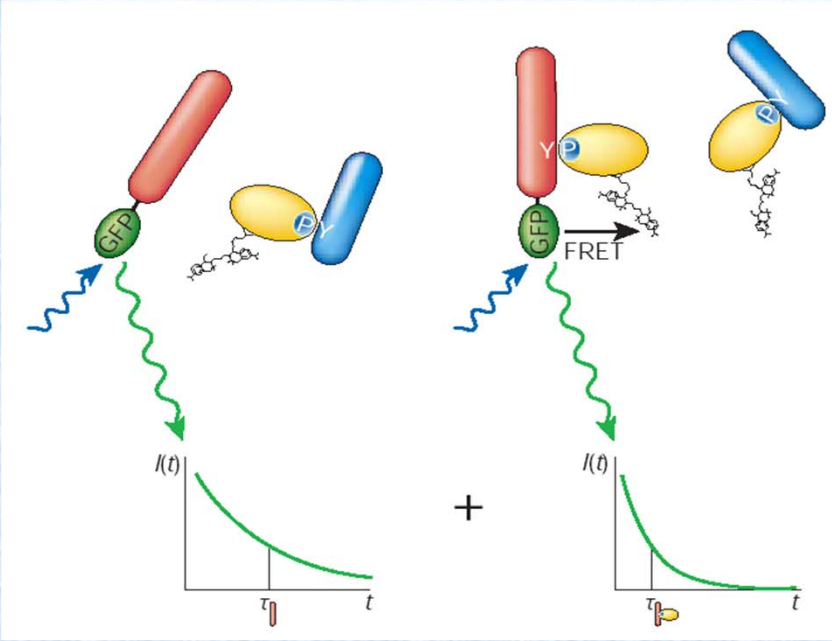
Peptide or protein fusion to selected proteins

High **AFINITY** and **SELECTIVITY**

ACTIVITY preservation

- Glutation-S-transferase / glutation agarose
- Hisx6 marker / Ni agarose
- Peptide to calmoduline linking / calmoduline agarose
- Protein A / γ-inmunoglobuline
- Protease sites (trombine, TEV, etc), in addition to the aforementioned
- Maltose binding protein (MBP)
- Epitopes: hemaglutinine, Myc and FLAG
- GFP derivates. FRET Detection

FRET GFP detection



PROTEIN FUSION: GENERAL PROBLEMS

- Each fusion protein requires a different scheme and protocol
- There are proteins which cannot be purified
- There are proteins which become inactive after fusion

2. Protein activity at proteomic scale

a) Biochemical genomic approximation
b) Microarrays (Chips)

a) Biochemical genomic approximation

Use of biochemical analysis in parallel to groups of proteins coming from a proteome. The objective is **identifying** a protein as responsible of a **function.**

GST-ORF Clonation → Expression and purification in groups of 96→

Biochemical assay→ Positive groups→ Subgroups →

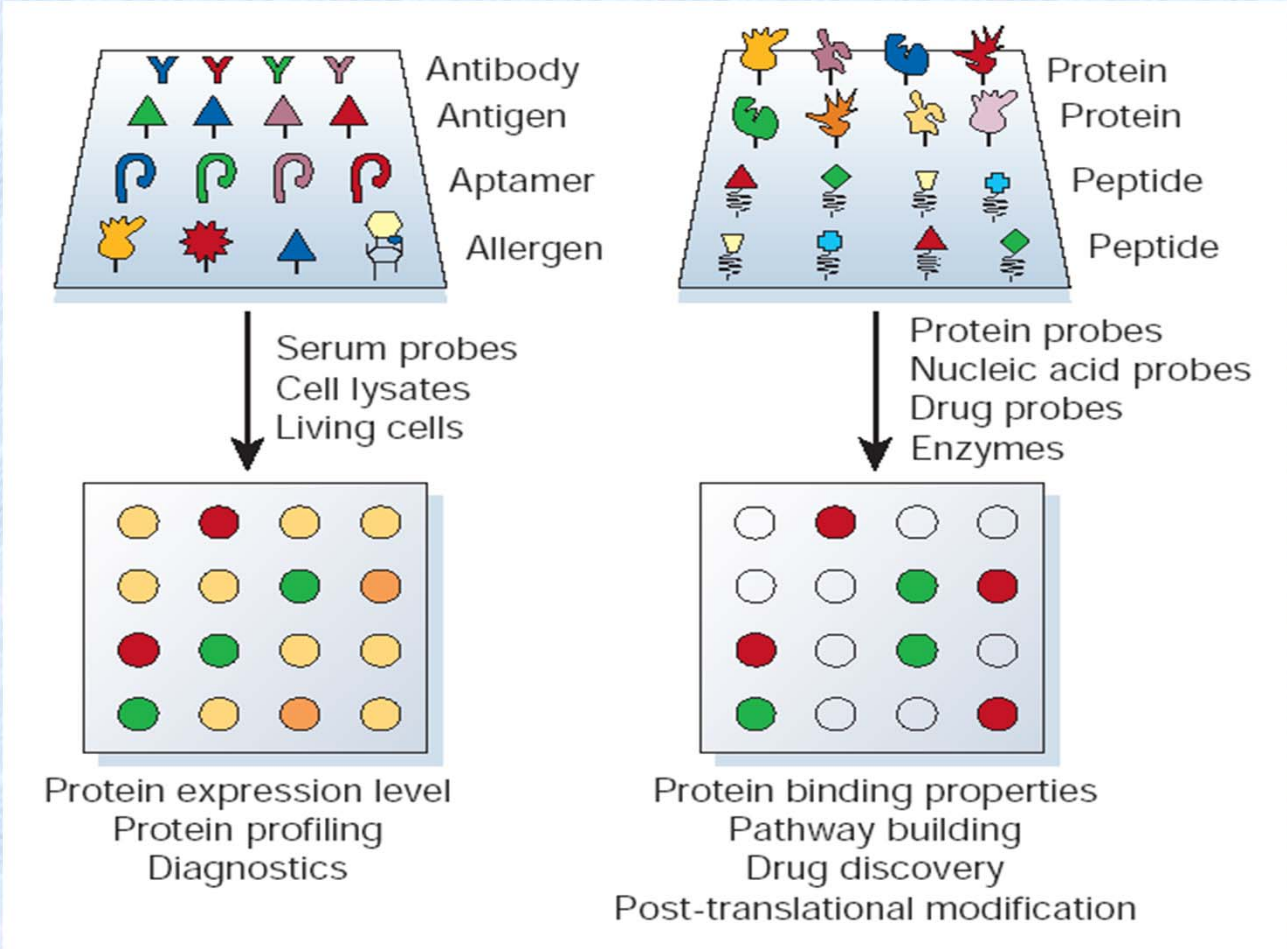
Subgroup assay→ Positive (deconvolution)

ADVANTAGES:
- Speed in function assignation to ORFs.
- Valid for any type of activity assay
- High sensibility
- Allows detection of protein complexes

b) Microarrays

Objective: High performance analysis of genomic function, at a high scale, quick and cheap.

Methodology: individual purified proteins are deposited in a surface, generally a glass sheet to analyze its activity



Comparison with DNA chips

DNA chips:

- No direct information of genic function
- Easy to make

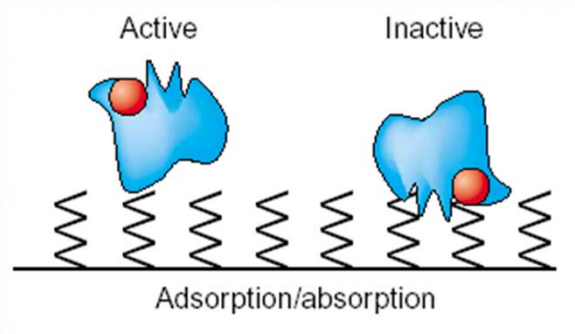
Protein chips:

- Direct information of function
- Complex to make, as protein function is dependent of state, like post-translational modifications, interaction with other proteins, subcellular localization, reversible covalent modifications, etc.

Chips manufacturing

Must retain protein function
Must be compatible with manufacturing technologies
Must be in a wet environment

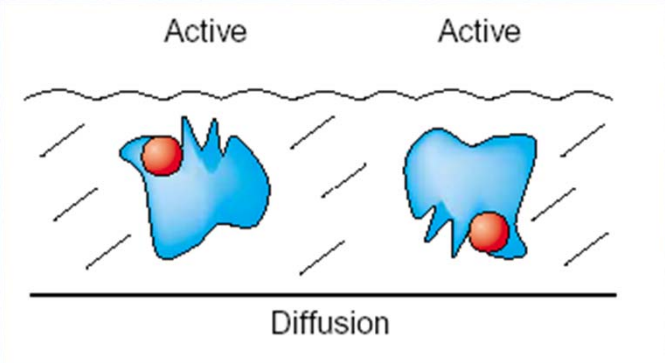
Soft substrates



Polystyrene, nitrocellulose or PVDF surfaces are used.
Do not allow high density packing
Application points disseminate the surface
Very low signal-noise
Random orientation

3D surface structures

- A thin polyacrylamide or agarose layer is prepared on glass
- Proteins are deposited by means of photolitography
- They are immobilised by cross-linking



Advantages

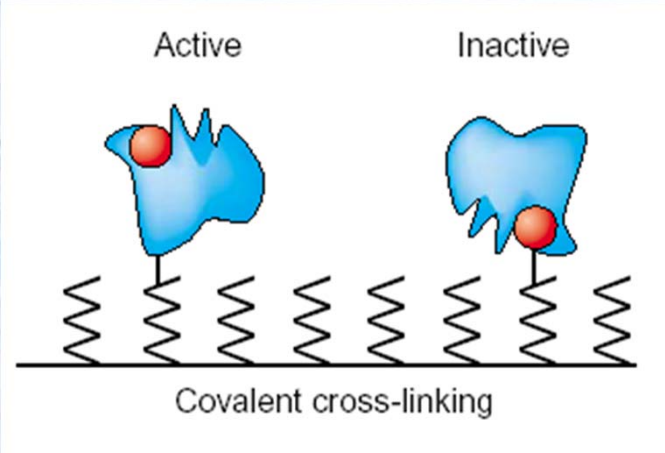
High immobilization capacity
Water environment
Active proteins

Disadvantages

Difficult buffer change
Difficult molecule recovery
Random orientation

Nanowells

Glass covered by PDMS
Nanowells are made, in which proteins are deposited
Proteins are immobilized by cross-linking



Advantages

- High immobilization capacity
- Evaporation reduction
- Cross contamination minimized
- Background noise minimized
-They are open, so they could be treated sequentially with buffers, washes, etc.
- Easy sample recovery

Disadvantages

Specialized, sophisticated and expensive equipment required
Random orientation

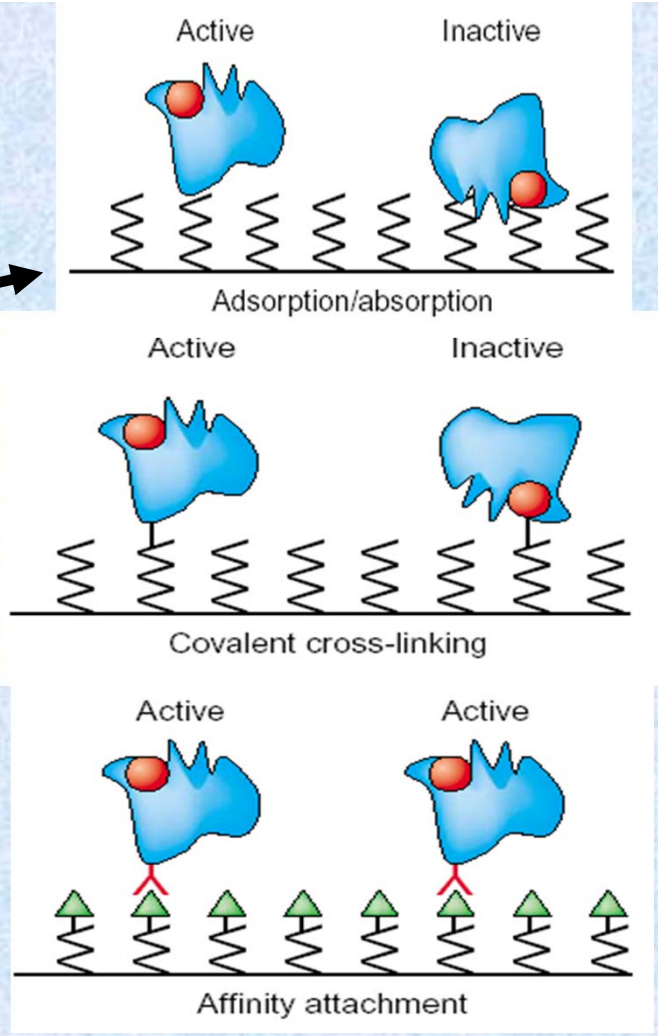
Glass

Direct protein linking in the glass surface or after chemical modification:

- Poly-L-Lysine (adsorption)
- Aldehyde activation (covalent)
- Epoxy activation (covalent)
- Gold cover (covalent)

Or

- Biotine (afinity)
- Hisx6 (afinity)



Advantages

-Covalent bond decreases probability of alteration of native conformation

- Affinity: correct orientation

Disadvantages

–NH2 groups in side chains can modify activity.

Affinity: All proteins should be biotinilated or Hisx6 tagged

Chip types

	Surface	Advantages	Disadvantages
2D soft substrates	Polystyrene Nitrocellulose membranes PVDF (Polyvinylidene fluoride) (Adsortion and absortion)	It does not require protein modifying. High capability of protein retention	Non-specific linking to protein. Random orientation. Backgroupd noise. Low density chips
3D surface structures	Polyacrylamide and agarose (difusion)	It does not require protein modifying. High capability of protein retention	Not commercialised
Nanowells	PDMS covered glass (polydimethylsilosane) (cross-linking)	Strong link and high packing density	Protein random orientation
Glass	Only glass poly-L-Lys cover aldehide activation Epoxy activation Gold cover Afinity union: Avidin or resina Ni2+ resin (adsortion, absortion, cross-linking or link afinity)	Idem as 2D Idem as 2D High density chips High density and resolution SPR and MS coupling High density, specific, strong linking	Idem as 2D Idem as 2D Random orientation Random orientation Random orientation Proteins should be biotinilated or Hisx6 tagged

Protein deposit systems

Low density: DOT-BLOT systems of 96 wells

High density: Robots able to deposit >30000 proteins

- Direct contact with the surface
- Ink-jet technology. Surface not touched
- Electrospray. Decreases spots from 150 to 30 μm.

Protein detection systems

Fluorescence (preferred because)
- Simple
- Safe
- Very sensitive
- Can have very high resolution

Fluorescence detection:
- 1 step: fluorescent molecule
- 2 steps: using an activity marker (f.e. biotine + Fluorescent streptavidin)
- RCA (rolling circle amplification)

ELISA
Radioisotopic marking

Protein-protein, DNA or drug interaction study

Protein detection systems (2)

SELDI-MS: Mass spectroscopy with laser desorption ionization increased on surface

- No marking needed: direct detection

- Used in chips covered with metallic layers, where laser vaporises proteins and MS reveals the identity of the protein

AFM: Atomic Force Microscopy

- Takes advantage of topological changes on the surface to identify captured proteins

SPR: Surface Plasmon Resonance

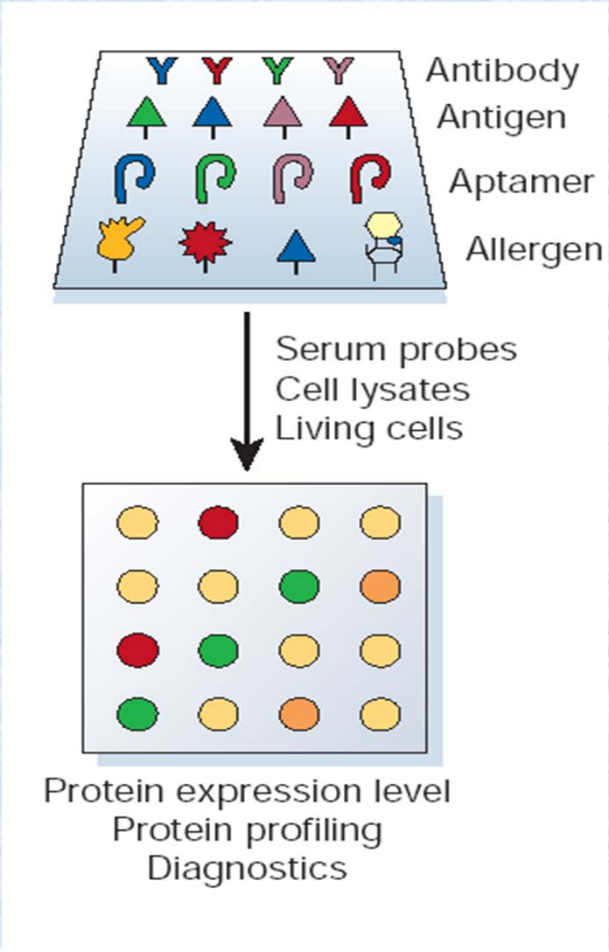
- Allows detection in real time

- Monito5rs interaction kinetics (i.e.) antigen-antibody

- Wide range of molecular weights, affinity and linking speed

Analytic chips

Different ligand types are bound, as antibodies, antigens, DNA or RNA aptamers, carbohydrates and small molecules



Protein samples from two biological states are marked separately with red and green fluorescent agents, mixed and incubated in the chip. Spots indicate overexpression

Find out protein expression levels.

Expression profiles

Clinical diagnosis

Analytic chips: antibodies

- It is the most used
- Antibodies are deposited at high density in the chip
- A cell extract, serum or living cells are deposited on the chip
- The antigen is linked
- Detection with tagged antigen (extract), or with a second antibody

Disadvantages:

- Having agents recognizing the right protein: antibodies
- Non-specific and expensive polyclonal antibodies
- Very expensive monoclonal antibodies, difficult to produce and time-consuming

Alternative to production of antibodies or similar

- Ab presentation in phages
- Ribosome presentation
- SELEX (Systematic Evolution of Ligands by Exponential Enrichment)
-mRNA presentation
- Affibody display

Construction of a **wide collection** of regions with potential **binding activity**

Analytic chips: allergens or antigens

- Opposite to the previous one
- Antigens are deposited at high density in the chip
- Antibodies are detected (usually serum) on the chip
- Antibody detection

Example 1: Hiller et al.
94 purified alergens linked to glass
Reactivity profiles allergic patient IgE
Small amount of serum
Allergy test as an alternative to skin test

Example 2: Robinson et al.
>300 autoantigens from 8 autoimmune diseases on glass
Small amounts of serum

Analytic chips: peptides

Objective: Detection of epitopes in proteins which define their basic activity.

Example 1: Houseman et al
- Immobilization in gold-covered glass of >3000 peptides from 9-mer amino acids
- Localization of c-Src tyrosine kinase substrates
- SPR, fluorescence and phosphorus image detection

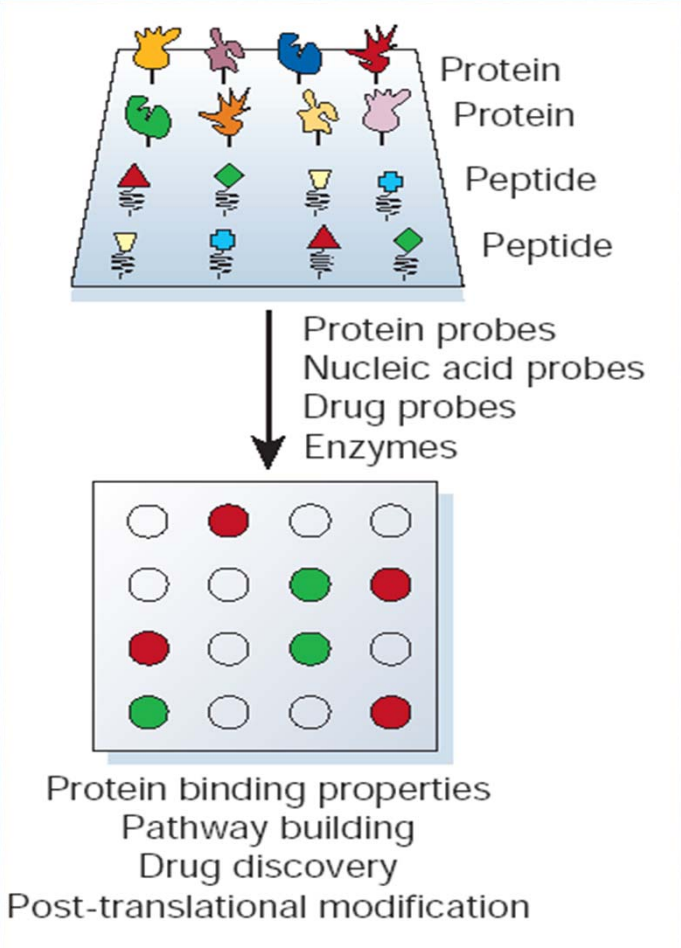
Advantages:

- Shorter peptides
- Peptides more stable than proteins
- Allow high density chip manufacturing
- Peptide synthesis *in situ* using photolitography or light-directed synthesis (cost-saving and few materials are needed)

Analytic chips: carbohydrates

Functional chips

Proteins and peptides coming from an organism's proteome are purified or sinthetised individually using high performance methods or high throughput (HT)



Protein samples from two biological states are marked separately with red and green fluorescent agents, mixed and incubated in the chip. Spots indicate overexpression

Allows to analyse:

- Protein activity
- Ligand-binding properties
- Post-translational modifications
- Pharmacological target identification
- Construction of biological networks

Functional chips: examples

Ejemplo 1: Zhu et al

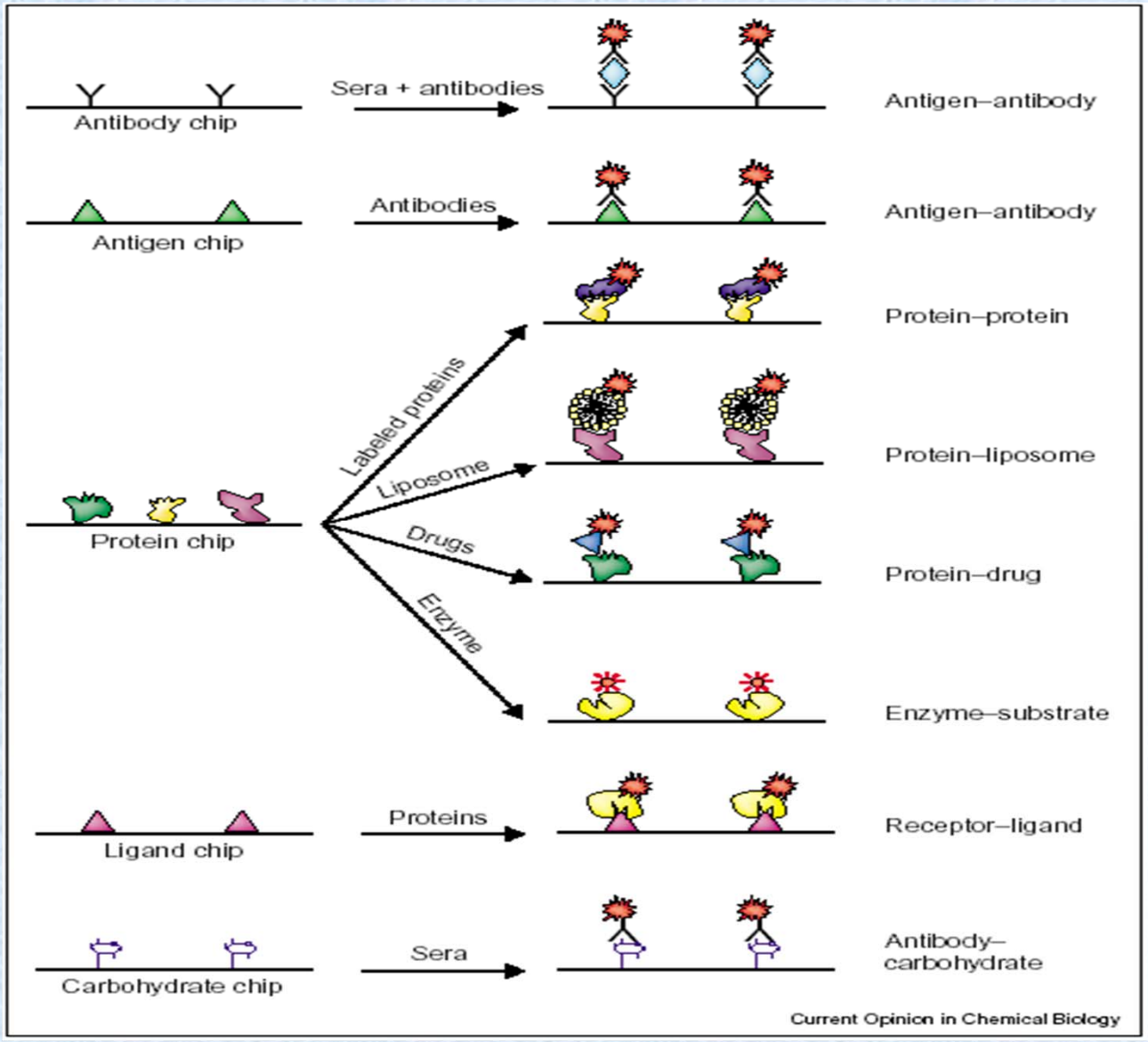
Interaction analysis of protein families
-119 kinase proteins and 17 substrates
-Substrates are immobilized in nanowells
-Kinases are incubated with substrates in the presence of ATP*
-Marked substrates are detected with phosphorus image

Ejemplo 2: Zhu et al

Analysis of a complete proteome

- 5800 from 6200 yeast ORFs
- Proteins are marked with GST and Hisx6 (both). A performance of 80% functional proteins was obtained
- They are linked to the Ni-NTA-covered chip with His x6
- Tests of calmoduline and phosphoinositides (PIs)
- 6 known interactions were obtained together with 33 new from calmoduline and more than 150 proteins interacting with PIs

Chip general applications



Biochemical Genomics vs Chips

Biochemical Genomics:

- It requires 64 essays to meet yeast genome (groups of 96 proteins)
- Very flexible for most types of biochemical assays
- Particularly useful for enzymatic activity assays

Disadvantages:

The use of protein pools does not guarantee protein quality
The other 95 proteins can interfere with measurement
No liable measurement of fluorescence linking
Cannot manage multiple positive at the same time

Protein chips:

Allow checking individual quality of proteins
Immediate ORF identification responsible for specific activities
Identification of multiple positive in one round
High performance analysis from protein activities
Automatation of construction, assay and results reading

Disadvantages:

6000 strains (yeast) should be cultivated, and 6000 protein purifications
Substrates should be immobilized for fluorescence-linking assays

Conclusions

- Chips can be a powerful tool in high scale biology
- Manufacturing on glass methodology is validated for protein analysis from a complete proteome
- Manufacturing methodology is robotized, as well as assay's and result reading's
- Antibody-improved production as reactive will influence chip improvement as tools for proteome analysis