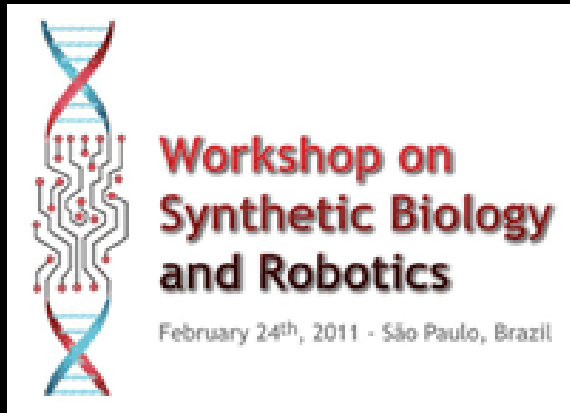
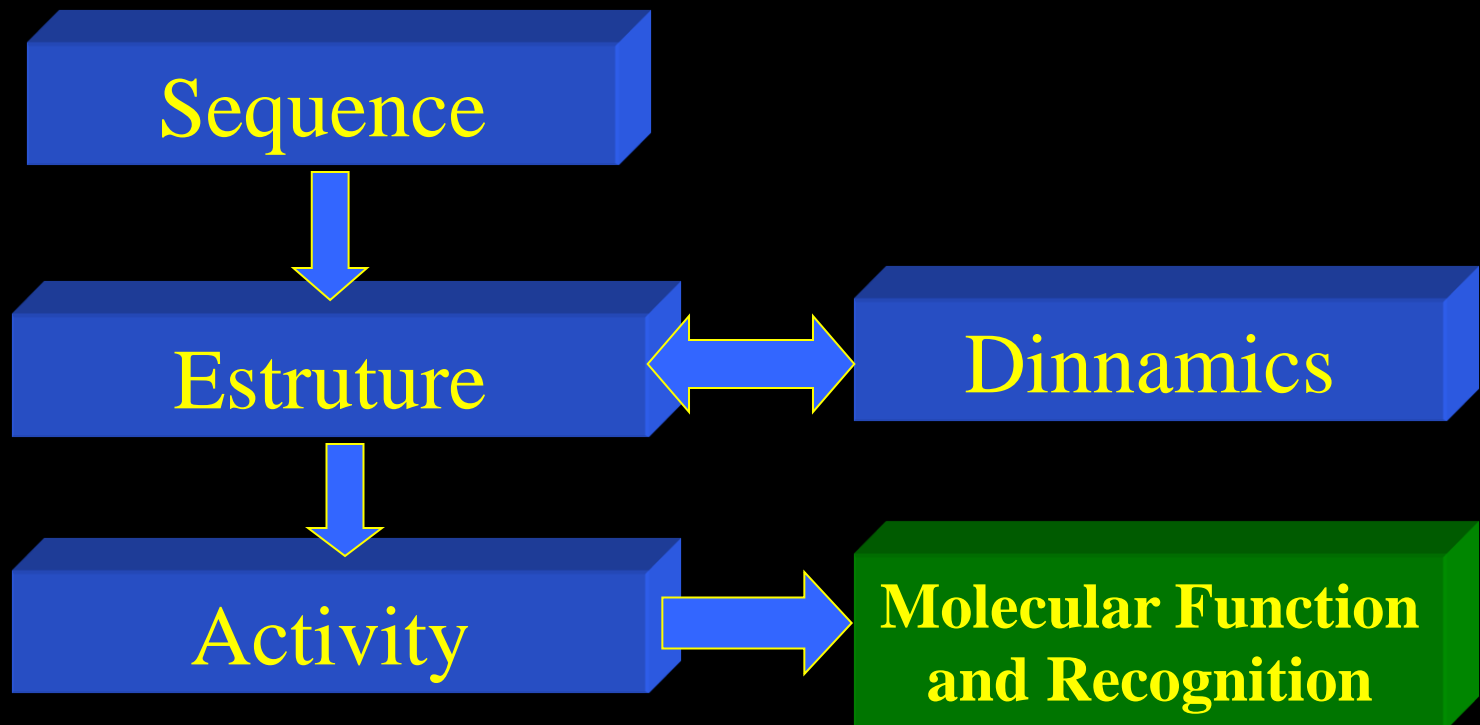


# Automation in protein crystallography: advances and new frontiers.



Otavio H. Thiemann  
February 24, 2011

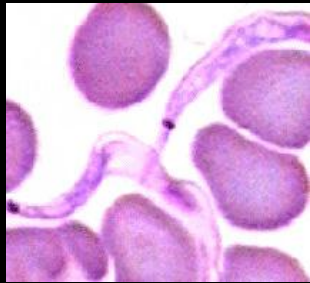
# Structural Biology in a Nutschell



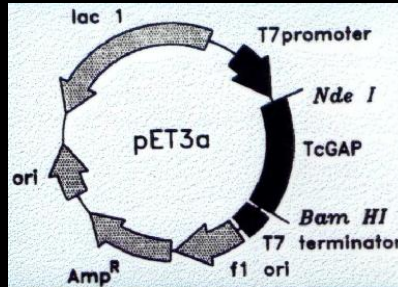
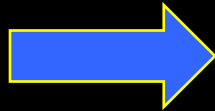
# Steps in protein crystallization and analysis

- Protein Production
- Crystallization
- X-Ray data collection
- Diffraction data Refinement

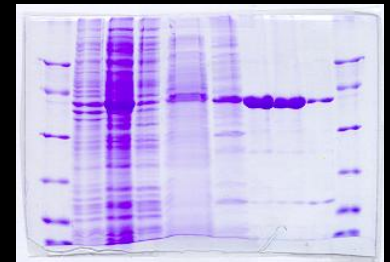
# Estrategy



**cloning**



**expression**

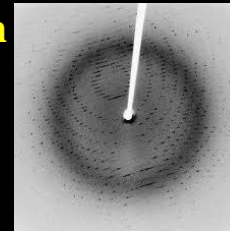


**purification**

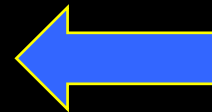
**crystallization**



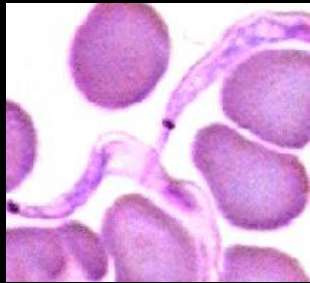
**Resolution**



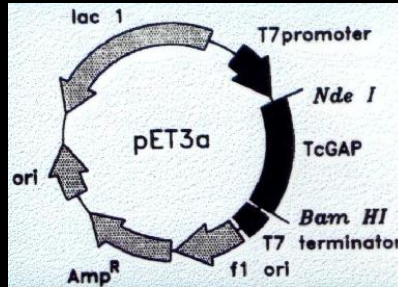
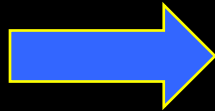
**Diffraction**



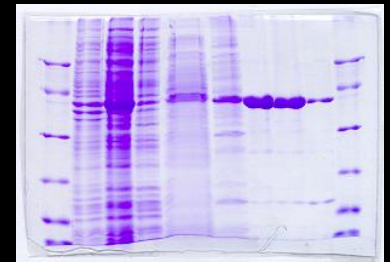
# Estrategy



cloning



expression

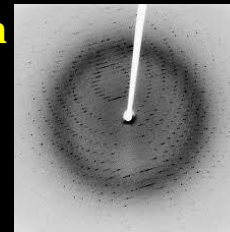
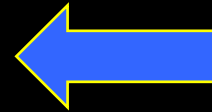


purification

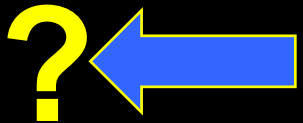
crystallization



Diffraction



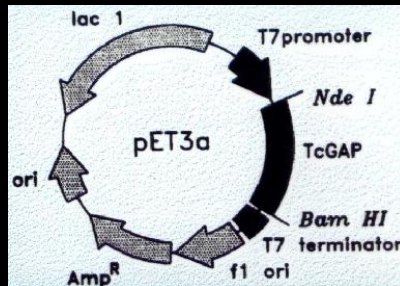
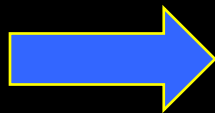
Resolution



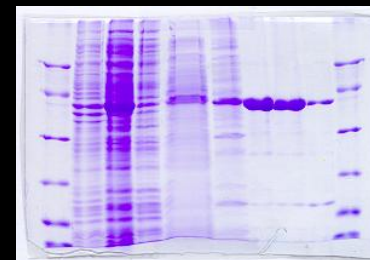
# Estrategy



cloning



expression

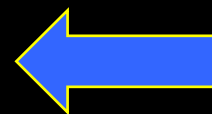


purification

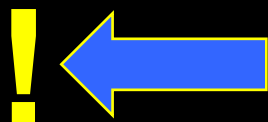
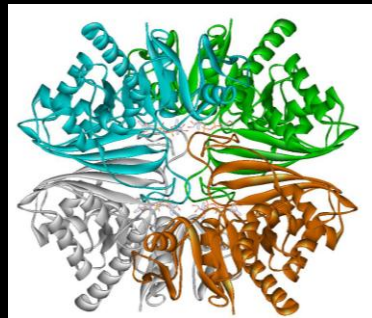
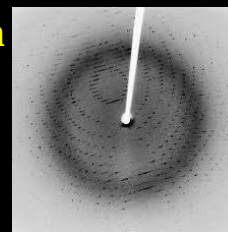
crystallization



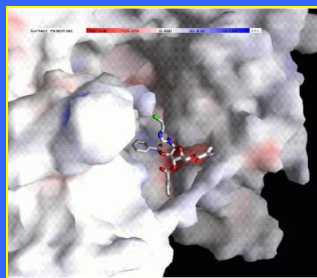
Diffraction



Resolution



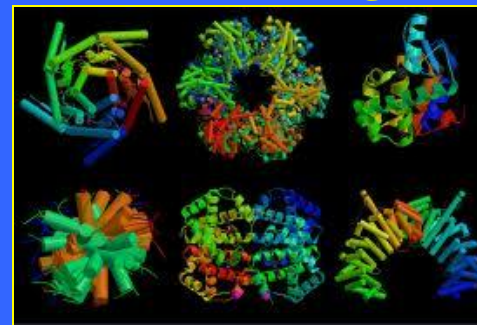
Rational Drug Design



Enzyme Catalysis



Protein Folding



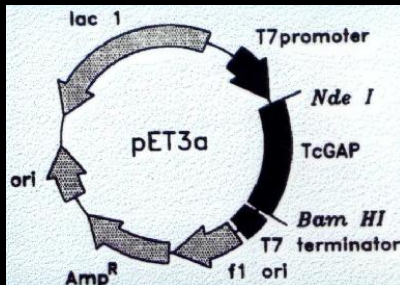
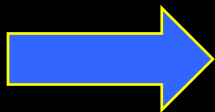
and more...



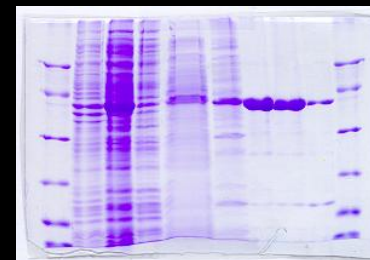
# Estrategy



cloning



expression

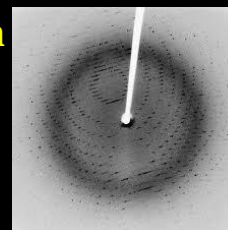
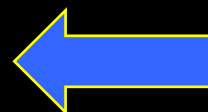


purification

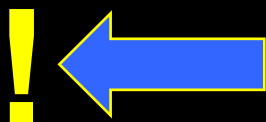
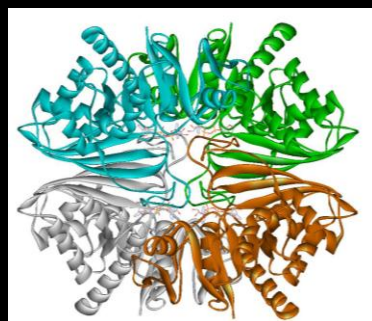
crystallization



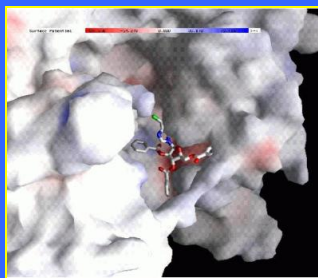
Diffraction



Resolution



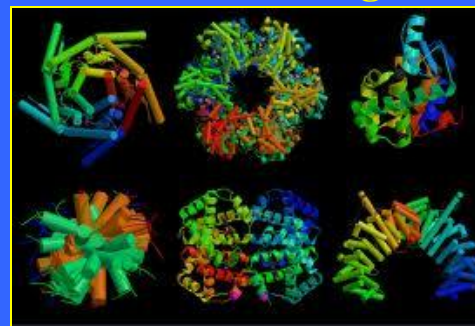
Rational Drug Design



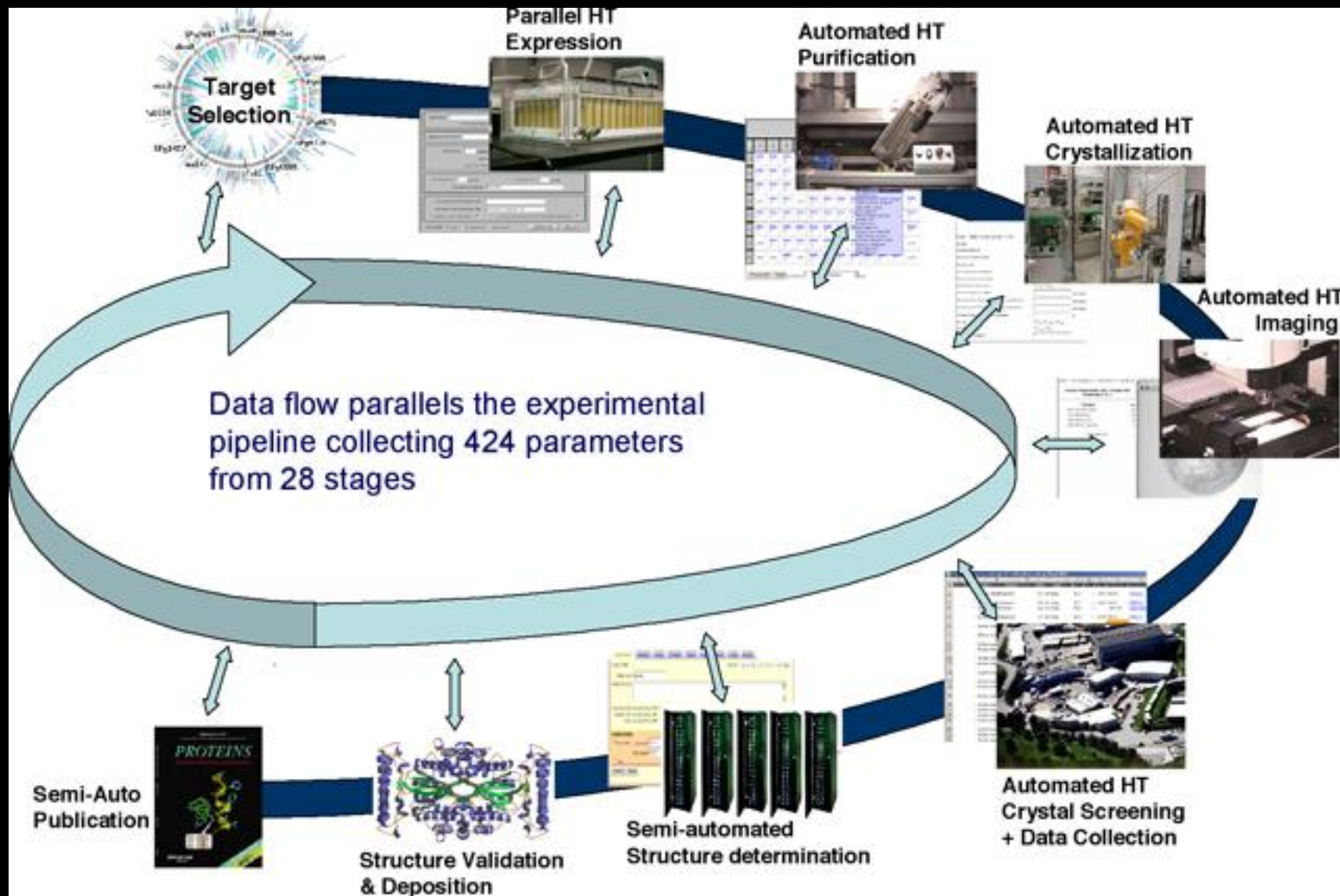
Enzyme Catalysis



Protein Folding

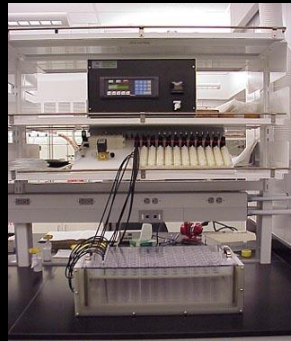


and more...





# Cloning



# Fermentation

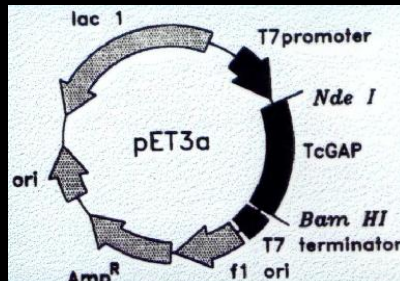
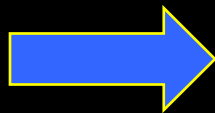


# Purification

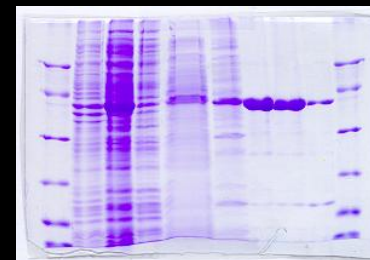
# Estrategy



cloning



expression

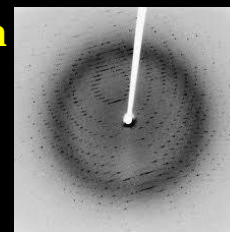
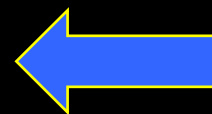


purification

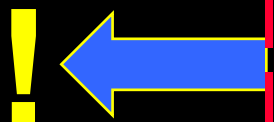
crystallization



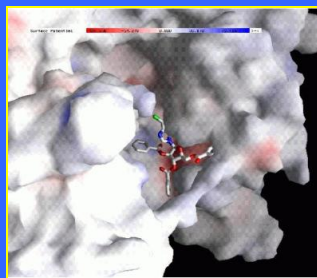
Diffraction



Resolution



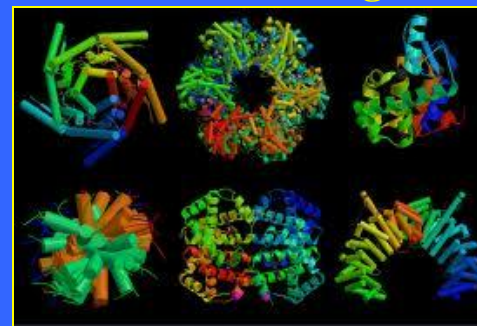
Rational Drug Design



Enzyme Catalysis

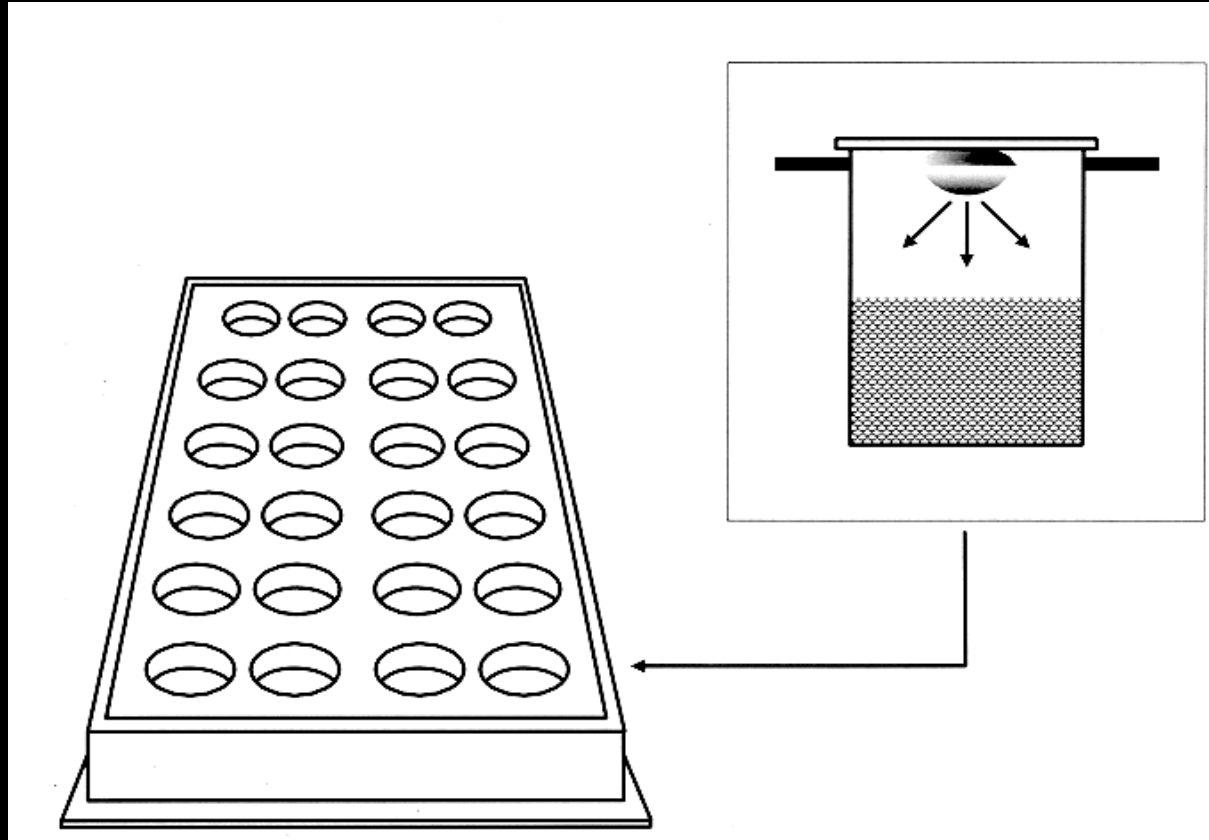


Protein Folding



and more...

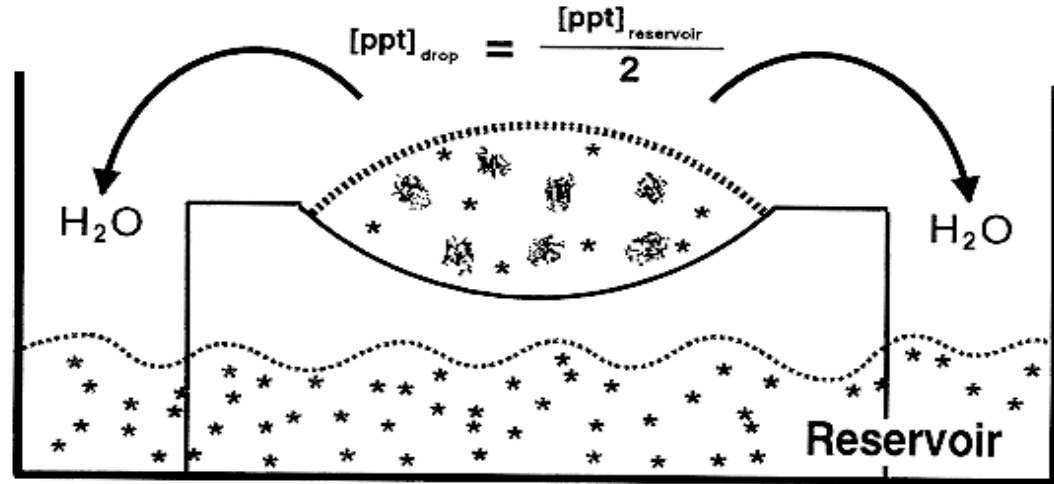
# Protein Crystallization: manual



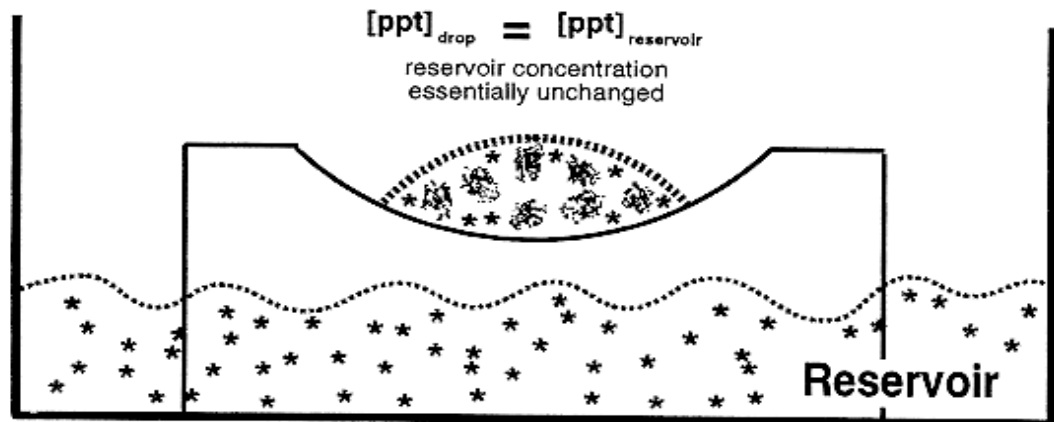
# Crystallization by Vapour Diffusion

Macromolecule

\* Precipitant Ions  
or Molecules (ppt)



(A) Equilibration proceeds through vapor phase



(B) Drop volume decreases, increasing concentration  
of both precipitant and protein

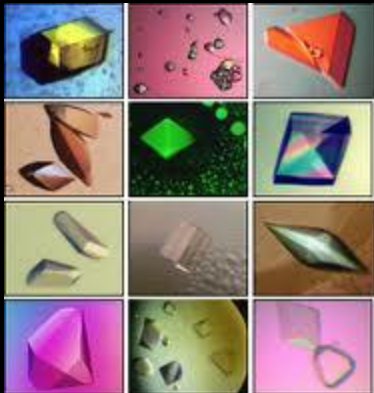


# Protein Crystallization: manual



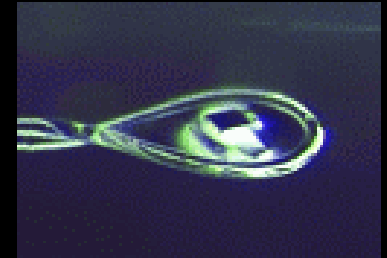
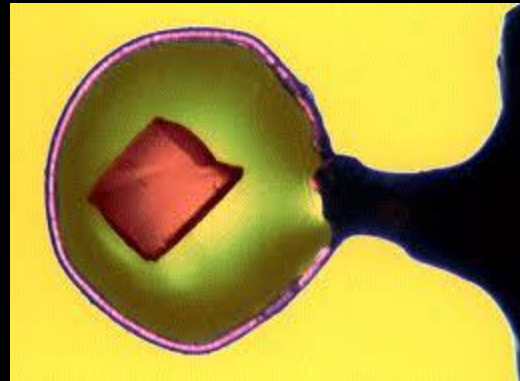
Typical experiment:

- 2 to 3  $\mu\text{L}$  protein
- $>6$  mg/mL
- 2 to 3  $\mu\text{L}$  crystallization solution
- 300 to 500  $\mu\text{L}$  crystallization solution in well
- Initial screening:
  - 300 – 500 conditions
  - i.e. 3.6 to 6 g of pure protein !!!!!

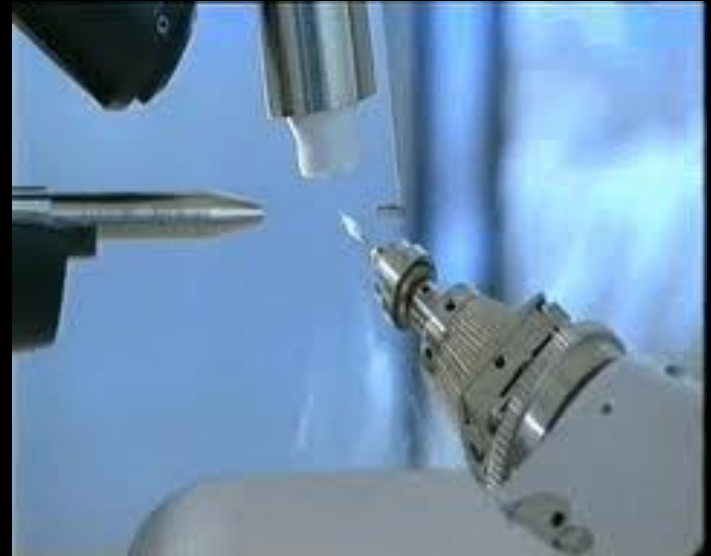
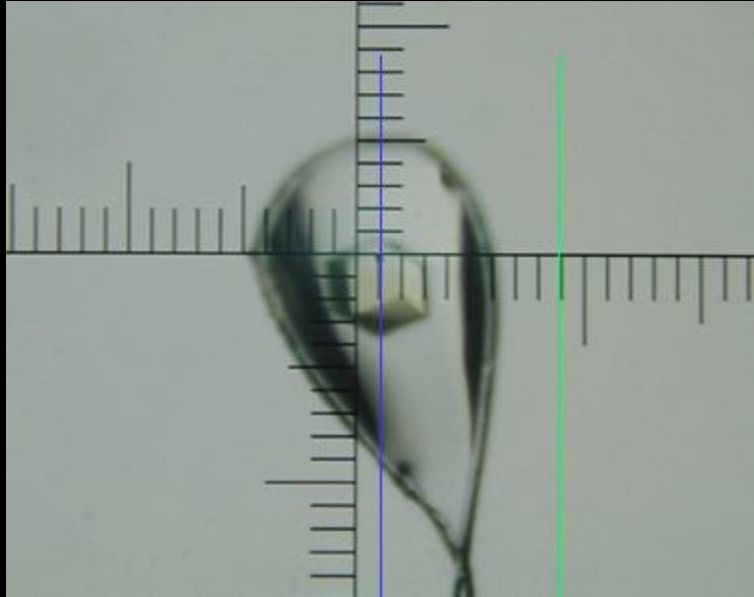




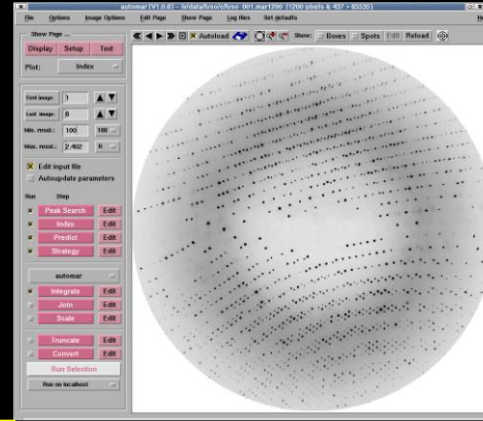
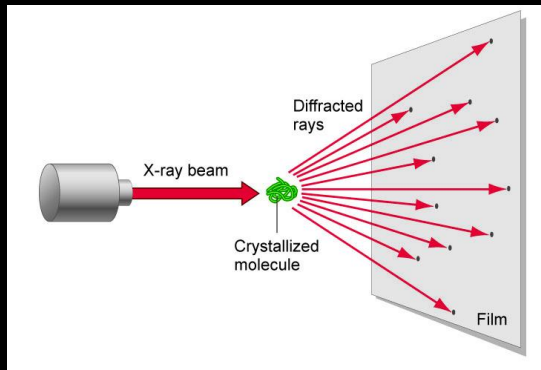
# Protein Crystallization: mounting



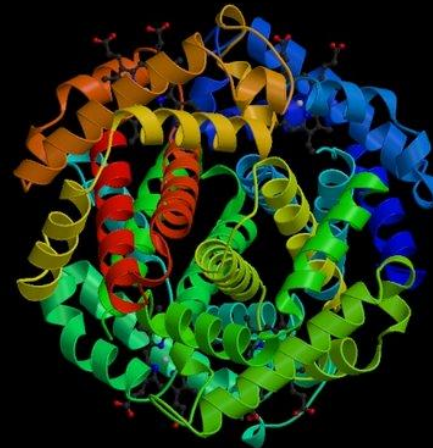
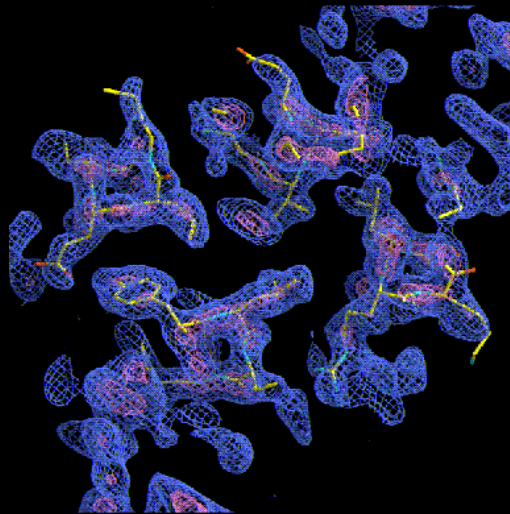
# Protein Crystallization: aligning



# Protein Crystallization: Diffraction data collection



**Refinement/interpretation**



# Protein Crystallization: Potential for automation

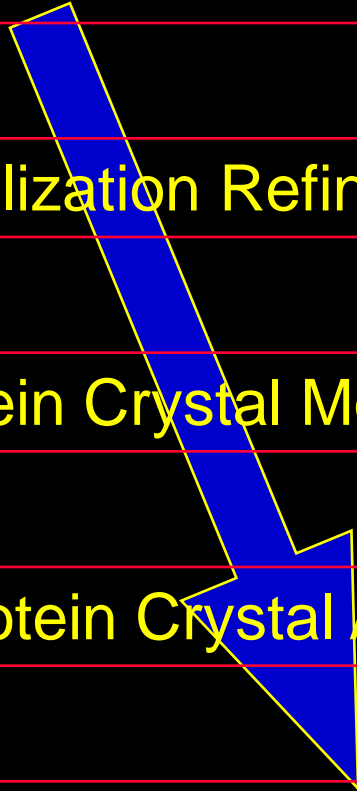
Crystallization Screening

Crystallization Refinement

Protein Crystal Mounting

Protein Crystal Alignment

Data Collection and Refinement



# Protein Crystallization: Potential for automation

Crystallization Screening

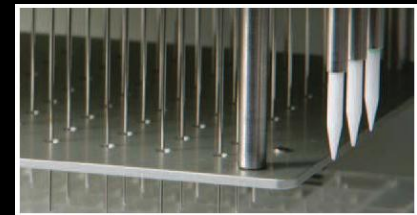
Crystallization Refinement

Protein Crystal Mounting

Protein Crystal Alignment

Data Collection and Refinement

Research lab. scale



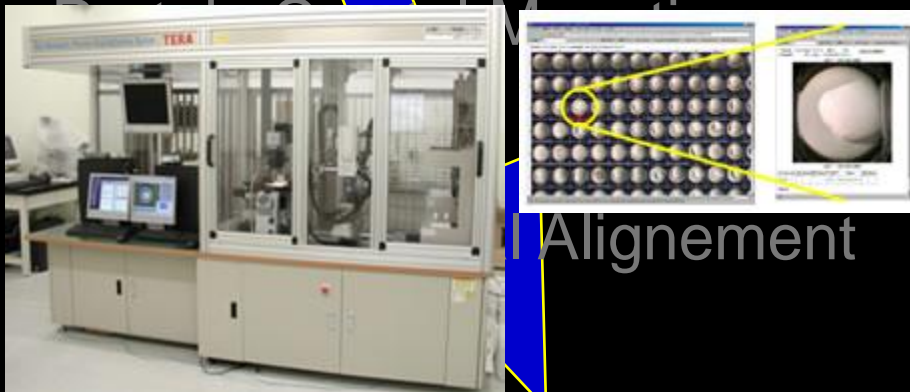


# Protein Crystallization: Potential for automation

Crystallization Screening

Crystallization Refinement

Large Scale (Structural Genomics)



Alignment



Data Collection and Refinement

# Protein Crystallization: Potential for automation

Crystallization Screening

Crystallization Refinement

**Protein Crystal Mounting**

Protein Crystal Alignment

Data Collection and Refinement

Large Scale (Structural Genomics)



# Protein Crystallization: Potential for automation

Crystallization Screening

Crystallization Refinement

Protein Crystal Mounting

**Protein Crystal Alignment**

**Data Collection and Refinement**

Large Scale (Structural Genomics)



**The long-term goal of the NYSGRC is to determine the 10,000 plus three-dimensional protein structures envisaged by the National Institute of General Medical Sciences Protein Structure Initiative.**

**The specific aims are:**

- **Identify target protein sequences for structural genomics**
- **Develop high-throughput E.coli expression of soluble target proteins**
- **Develop high-throughput production of target proteins**
- **Develop high-throughput biophysical characterization of target proteins**
- **Develop high-throughput crystallization of target proteins**
- **Develop efficient experimental strategies for MAD crystallography**
- **Develop high-throughput synchrotron data collection with target protein crystals**
- **Develop and use an Internet-based computational pipeline for protein crystallography**
- **Develop high-throughput molecular replacement tools for crystallography**
- **Develop high-throughput comparative modeling for structural genomics**
- **Develop efficient annotation and dissemination of protein structures and models**
- **Develop an Internet-based "web-book" for use by the NYSGRC**

# Structural Genomes

## **Center for Eukaryotic Structural Genomics**

Develop the methodologies and technology necessary for high-throughput, genome scale, eukaryotic protein production, characterization and structure determination, focusing on proteins from Arabidopsis.

## **Joint Center for Structural Genomics**

Develop technology and structural studies of human and *C. elegans* proteins involved in signal transduction.

## **Midwest Center for Structural Genomics**

Develop robotic technology and synchrotron-based X-ray crystallography methods; structural studies of archaea, bacteria and eukarya proteins.

## **Northeast Structural Genomics Consortium**

Develop technology; analyze complementarity of NMR and crystallographic methods; and structural studies of roundworm, fly and human proteins.

## **Southeast Collaboratory for Structural Genomics**

Develop automated NMR and crystallographic structure determination; structural studies of *Pyrococcus furiosus*, *C. elegans* and human proteins.

## **Structural Genomics of Pathogenic Protozoa**

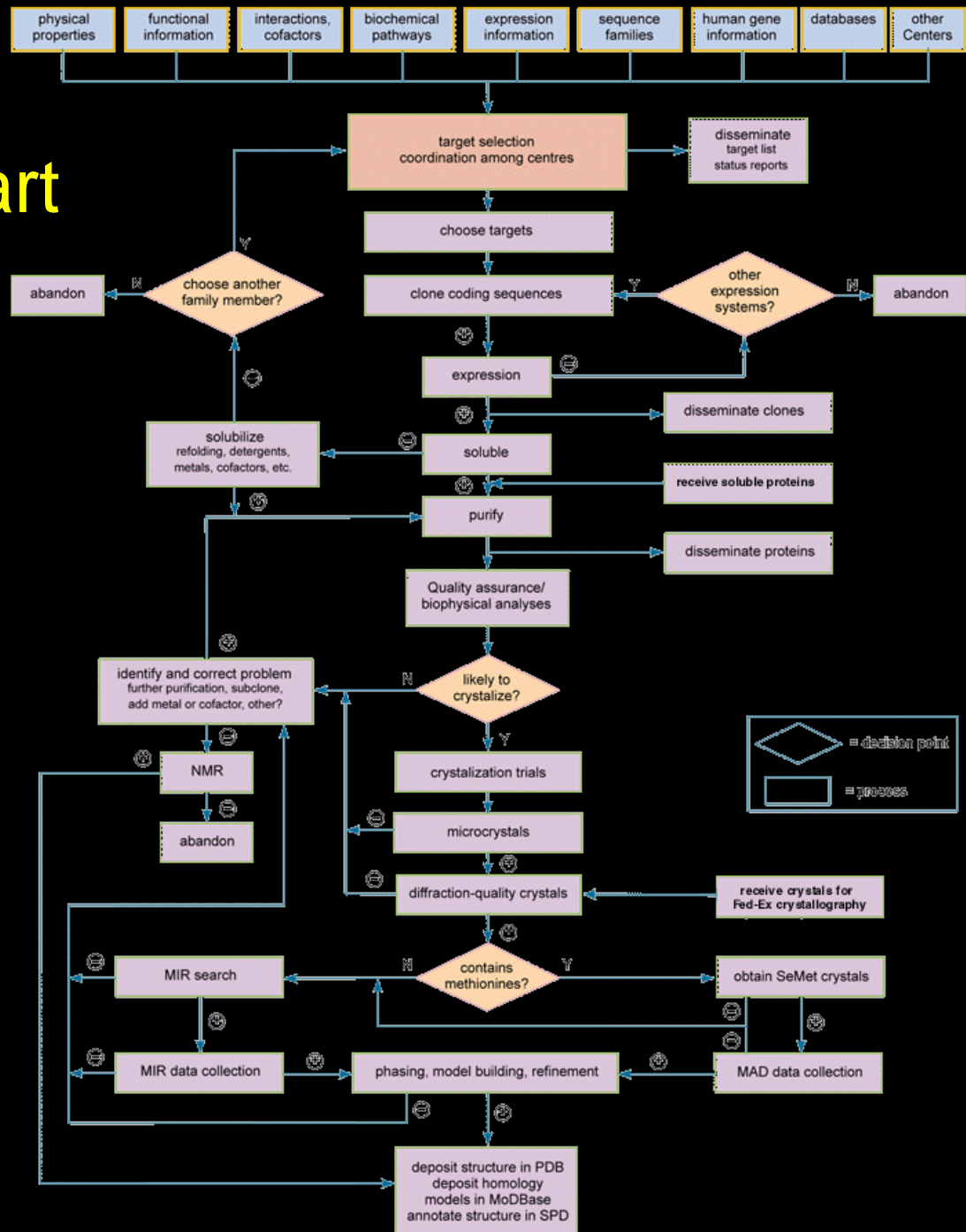
Develop technologies and structure determination of proteins from major global pathogenic protozoa, *Leishmania major*, *Trypanosoma brucei*, *Trypanosoma cruzi* and *Plasmodium falciparum*.

## **TB Structural Genomics Consortium**

Develop technology and structural study of *Mycobacterium tuberculosis* proteins.



# Typical flowchart



Selection of potential crystallographic and NMR targets



Complete cDNA sequencing



Cloning



Protein over-expression



Protein purification



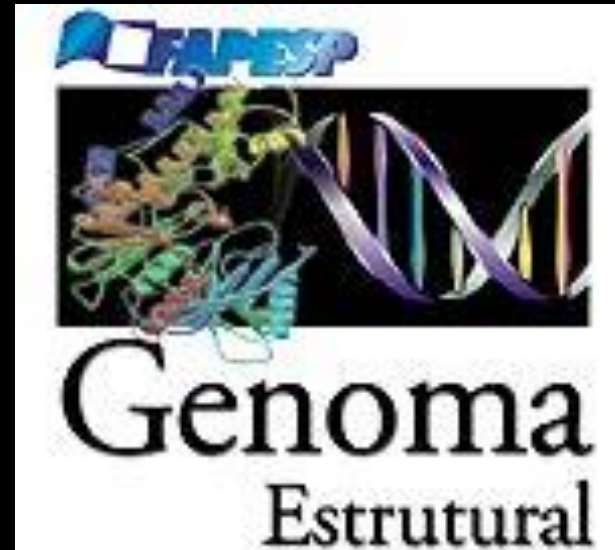
NMR data collection or protein crystallization

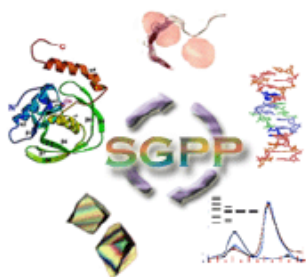


X-ray data collection



Structure determination.





## STRUCTURAL GENOMICS OF PATHOGENIC PROTOZOA

# SGPP

[Click to see status of SGPP Targets](#)

[MSGPP Targets](#)

[CHTSB Targets](#)

[Project Description](#)

[Flow Diagrams](#)

[Contact Information](#)

[Consortium Members](#)

[Diseases Under Investigation](#)

\* *Chagas' Disease*

[*American*

*Trypanosomiasis*]

\* *African Sleeping Sickness*

[*African Trypanosomiasis*]

\* *Leishmaniasis*

\* *Malaria*

[Target Organisms](#)

\* *Trypanosoma cruzi*

\* *Trypanosoma brucei*

\* *Leishmania spp.*

\* *Plasmodium falciparum*

\* (*Plasmodium vivax*)

[Send Suggestions for  
Protein Targets](#)

[Genome Status](#)

[SGPP Progress](#)

[3-D Structures](#)

[Structures w/Ligands](#)

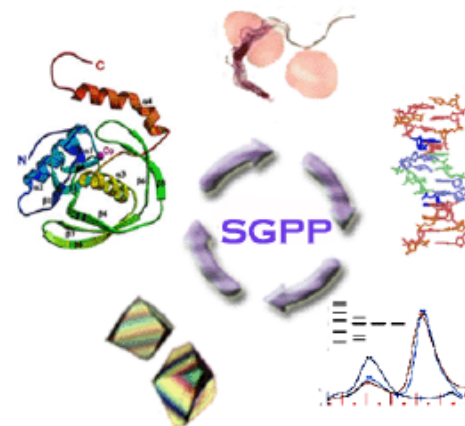
[Papers by SGPP](#)

[Related Links/Resources](#)

[Employment Opportunities](#)

[SGPP News Articles](#)

[Feedback on SGPP web site](#)



**Structural Genomics Method Development**

**Discovery of New Folds**

**Medicinal Drug Design**

Supported by:

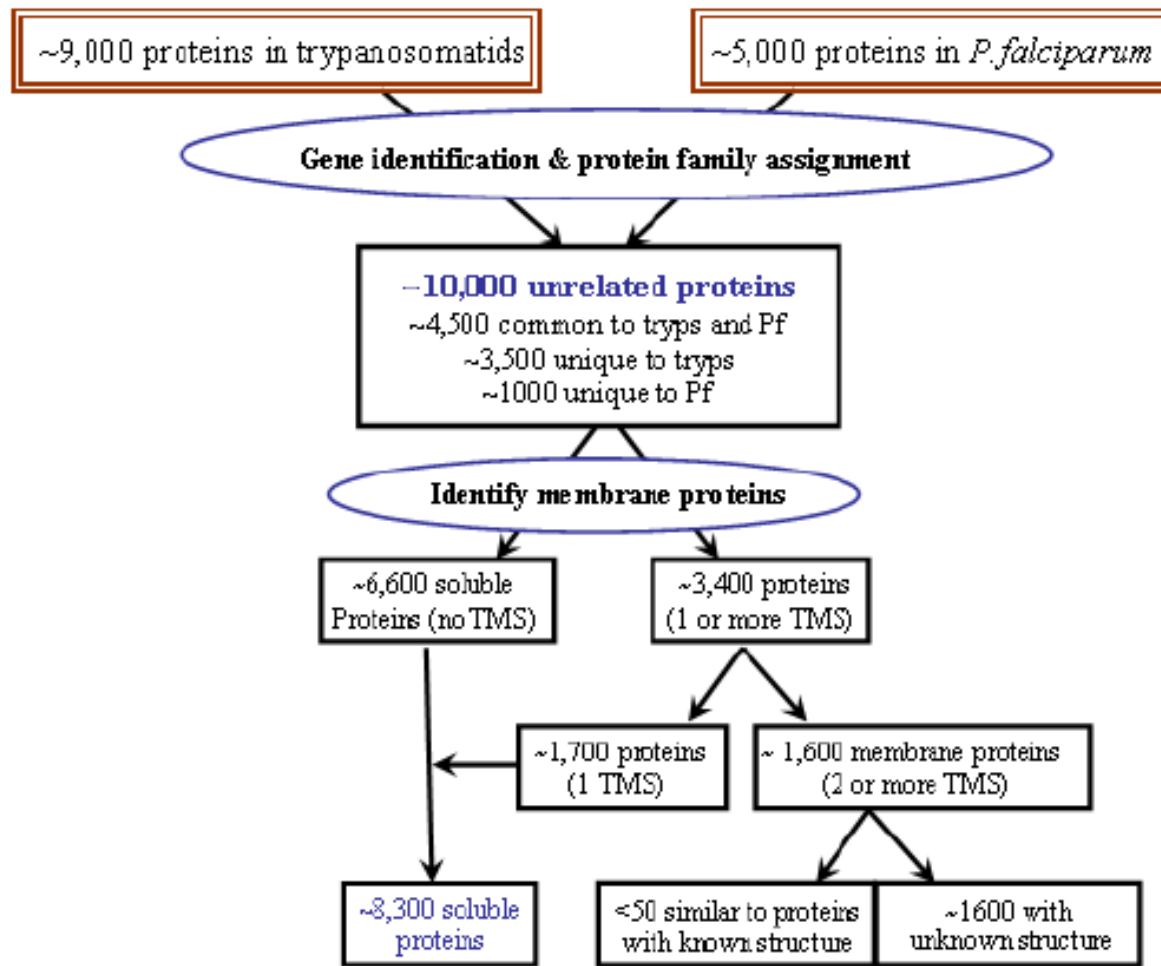


This website is no longer being actively updated. For further progress please see  
SGPP's successors:

**MSGPP**  
**Medical Structural Genomics of**  
**Pathogenic Protozoa**

**CHTSB**  
**Center for High-Throughput**  
**Structural Biology**

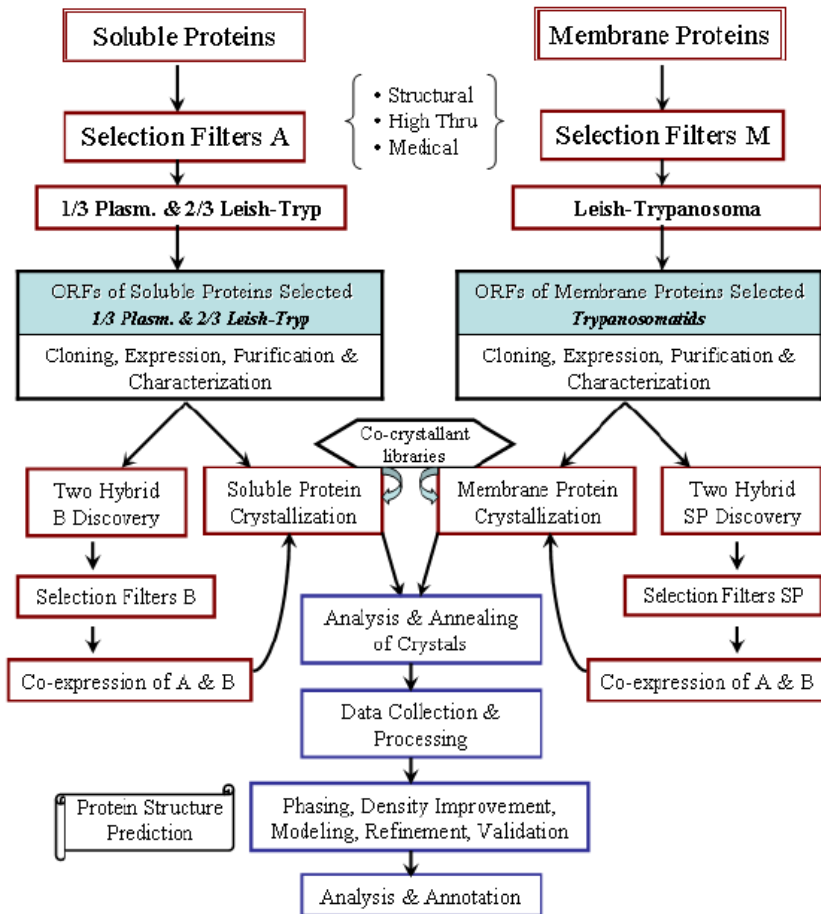
# Structural Genomics of Pathogenic Protozoa Target Selection Scheme



# STRUCTURAL GENOMICS OF PATHOGENIC PROTOZOA

## FLOW DIAGRAMS

### Structural Genomics of Pathogenic Protozoa - Flow Diagram



*B's are soluble protein partners of selected soluble A's;  
SP's are soluble partners of membrane proteins M*



# SGPP Target Status - Page 1 of 1391

<<first <previous <all> next> last>>

Last updated  
May 22, 2007

Search in  \* for

\***title**: SGPP ID, source, original source ID, updated ID if known, protein name, function, PDB ID, tags;

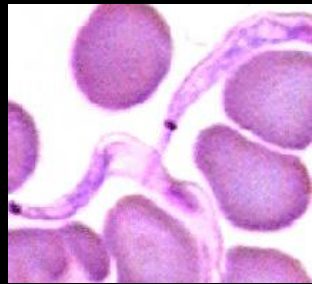
**aa\_seq, nt\_seq**: amino acid or nucleotide sequence - exact match only

Targets are sorted by progress. See [all targets](#), [15 targets per page](#), [nucleotide sequences](#), [amino acid sequences](#) or [XML](#).

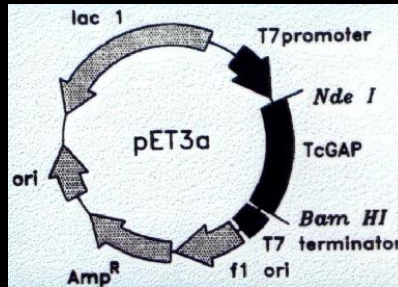
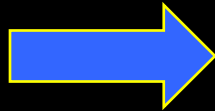
Target			Status									
Organism & Target ID	Src	Database Identifier	Cloned	Expressed	Soluble	Purified	Crystals		Diff. Qual. Xtals	Diffraction Data, Å	Xtal Structure	In PDB
							Screening	Growth Lab				
Lmaj004091AAA	Gd	LmjF30.0810	+	+	+	+	+	+	+	1.9	+	<a href="#">1xrp</a>
Lmaj004144AAA	Gd	LmjF30.1890	+	+	+	+	+	+	+	1.7	+	<a href="#">1y63</a>
Lmaj005461AAB	Gd	LmjF32.0700	+	+	+	+	+	+	+	2.0	+	<a href="#">1yf9</a>
Lmaj007771BAB	Gd	LmjF23.0050	+	+	+	+	+	+	+	2.0	+	<a href="#">2hq9</a>
Lmaj008024AAA	Gd	LmjF01.0480	+	+	+	+	+	+	+	2.1	+	<a href="#">1svv</a>
Lmaj01134AAC	Gd	LmjF13.1460	+	+	+	+	+	+	+	1.95	+	<a href="#">1y1x</a>
Pfal000304AAA	P	PF10_0225	+	+	+	+	+	+	+	2.1	+	<a href="#">2f84</a>
Pfal004331AAA	P	MAL13P1.257	+	+	+	+	+	+	+	2.2	+	<a href="#">1zso</a>
Pfal004546AAA	P	MAL6P1.148	+	+	+	+	+	+	+	2.2	+	<a href="#">1y13</a>
Pfal005984AAA	P	PF11_0208	+	+	+	+	+	+	+	2.6	+	<a href="#">1xq9</a>
Pfal006645AAA	P	PF13_0349	+	+	+	+	+	+	+	3.05	+	<a href="#">1xiq</a>
Pfal007201AAA	P	PF14_0545	+	+	+	+	+	+	+	2.9	+	<a href="#">1syf</a>
Pfal007254AAA	P	PF14_0598	+	+	+	+	+	+	+	2.5	+	<a href="#">2b4r</a>
Pfal008421AAA	P	PFE0660c	+	+	+	+	+	+	+	1.8	+	<a href="#">1sq6</a>
Pfal008434AAA	P	PFE0730c	+	+	+	+	+	+	+	2.9	+	<a href="#">2f8m</a>

Target			Status									
Organism & Target ID	Src	Database Identifier	Cloned	Expressed	Soluble	Purified	Crystals		Diff. Qual. Xtals	Diffraction Data, Å	Xtal Structure	In PDB
							Screening	Growth Lab				
Tbru019101AAA	Gd	Tb09.211.3420	+	+	+	+	+	+	+	2.5	+	<a href="#">3buw</a>
Tcru019078AAA	Gd	Tc00.1047053510339.50	+	+	+	+	+	+	+	2.8	+	<a href="#">3buw</a>
Pfal000066AAA	P	PF10_0022	+	+	+	+	+	+	+	2.8		
Pfal006821AAA	P	PF14_0164	+	+	+	+	+	+	+	3.5		
Lmaj010130AAA	Gd	LmjF10.0560	+	+	+	+	+	+	+	2.0		
Lbra003107AAA		LbrM25.2080	+	+	+	+	+	+	+			
Lmaj002537AAA	Gd	LmjF16.0230	+	+	+	+	+	+	+	2.8		
Lmaj004542AAA	Gd	LmjF31.0560	+	+	+	+	+	+	+	2.1		
Lmaj004655AAA	Gd	LmjF31.2410	+	+	+	+	+	+	+			
Lmaj00521AAA	Gd	LmjF32.2980	+	+	+	+	+	+	+	3.8		
Lmaj00592AAA	Gd	LmjF23.0360	+	+	+	+	+	+	+	3.0		
Lmaj00847AAA	Gd	LmjF14.1370	+	+	+	+	+	+	+			
Lmaj00884AAA	Gd	LmjF14.0240	+	+	+	+	+	+	+	1.9	+	<a href="#">3ksv</a>
Lmaj01188AAA	Gd	LmjF04.0070	+	+	+	+	+	+	+	3.1		
Pviv005676AAA	P	Pv111245	+	+	+	+	+	+	+	3.3	+	<a href="#">2pgf</a>

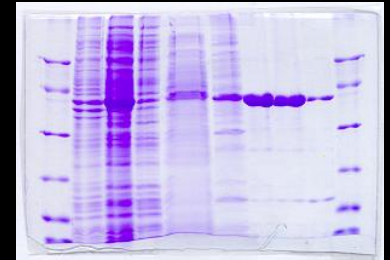
# Estrategy



cloning



expression

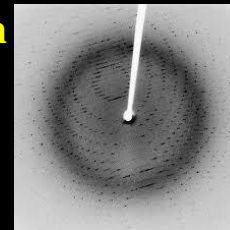
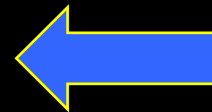


purification

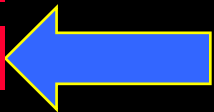
crystallization



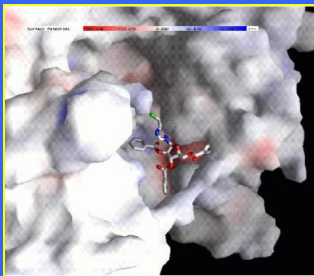
Diffraction



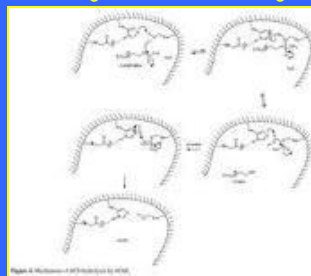
Resolution



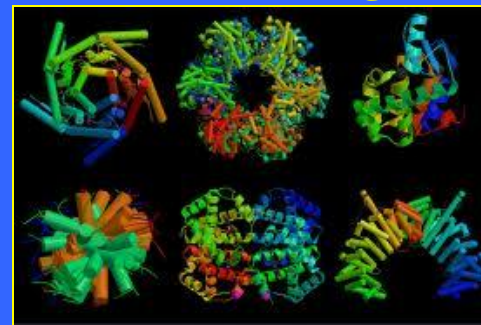
Rational Drug Design



Enzyme Catalysis

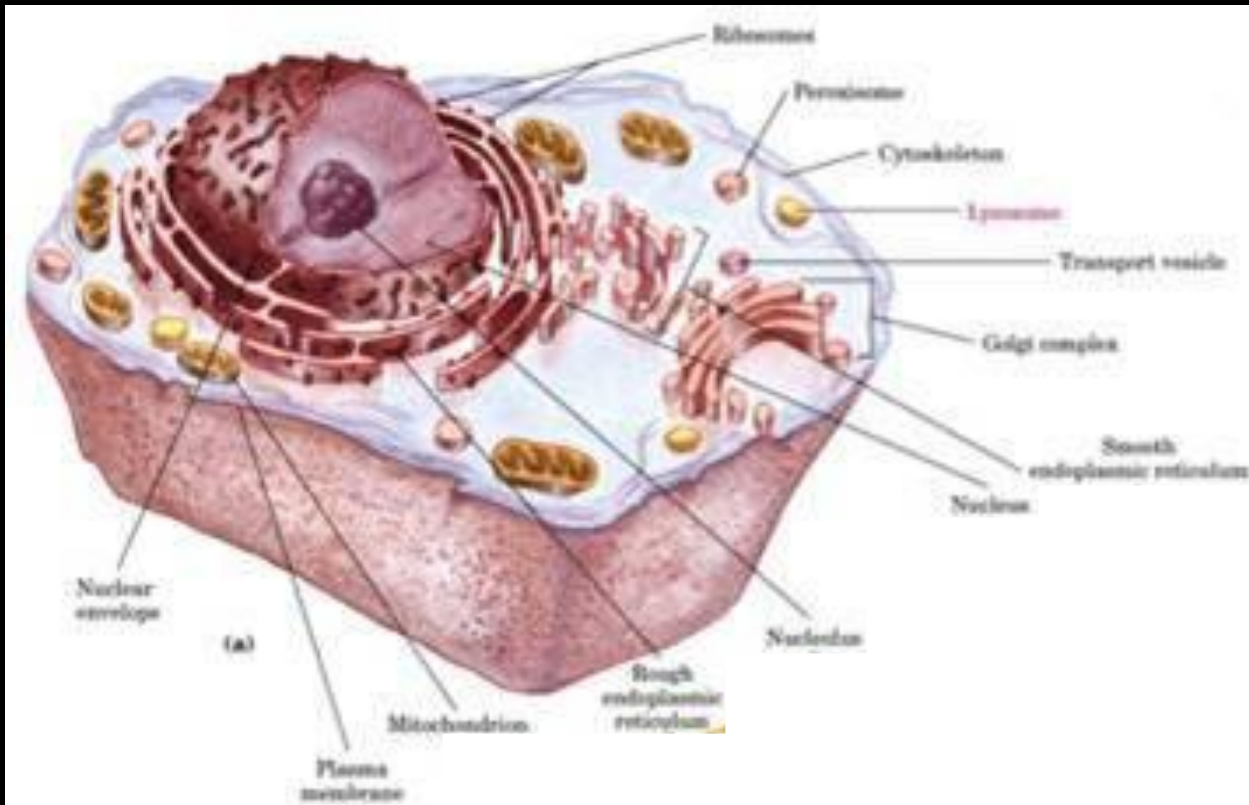


Protein Folding



and more...

# Targets for Drug Action



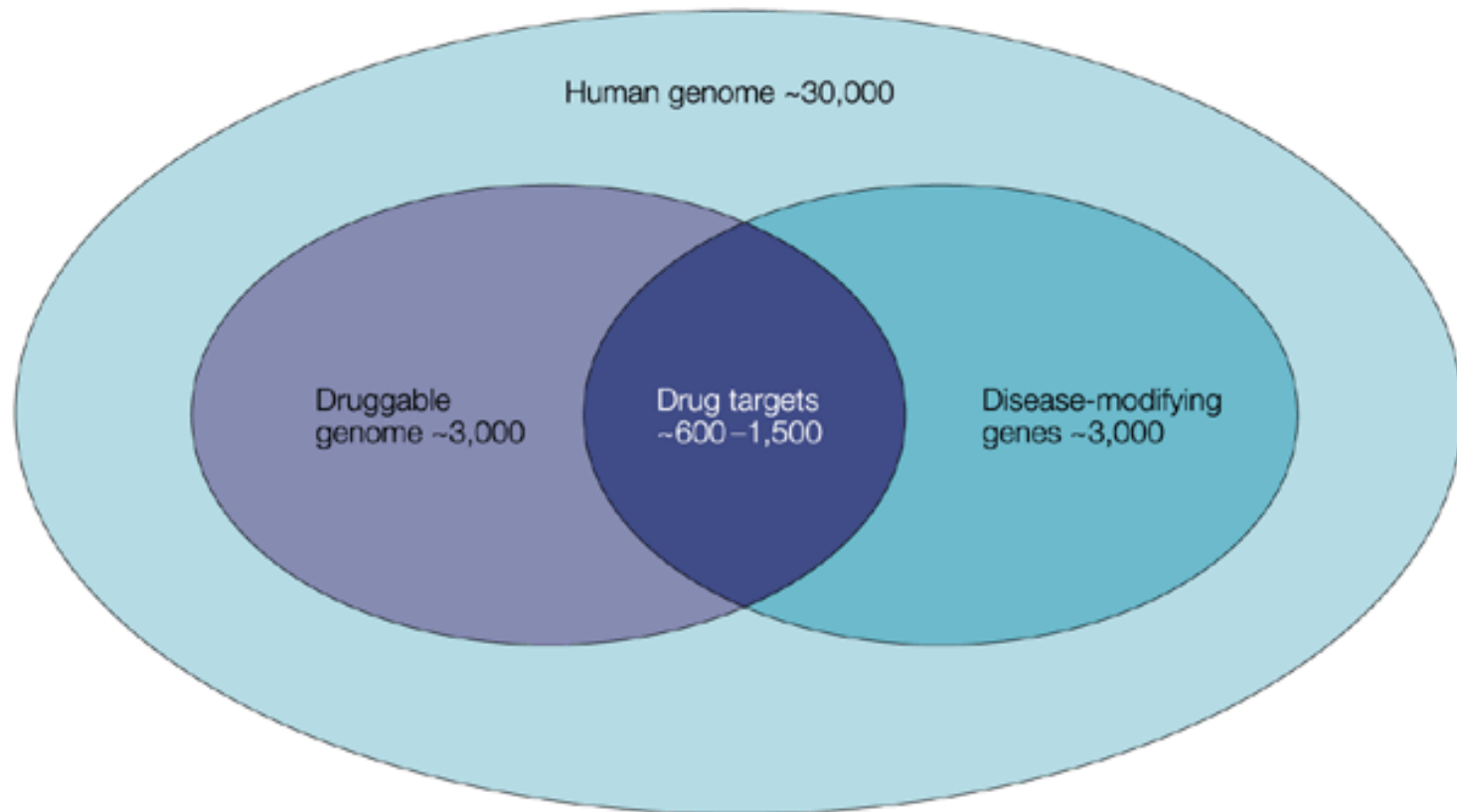
- Proteins > 98% pharmaceuticals
  - Nucleic Acids
  - Lipids (membranes)
  - sugars
- } < 2% pharmaceuticals

# The “druggable” genome

revisão por A.L.Hopkins & C.R.Groom,  
Nature Reviews Drug Discovery, 2002, vol.1,#9, pp 727-730

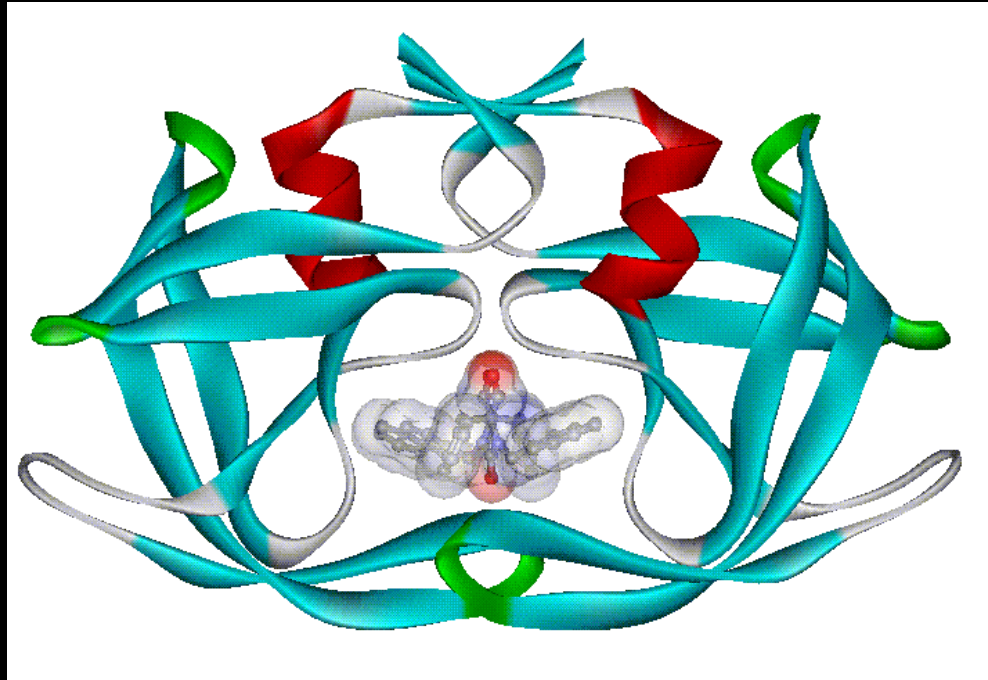
- = of the ~30000 genes in the human genome capable of expressing a proteins and binding to a small molecule (Drug).

# Estimated number of drug targets



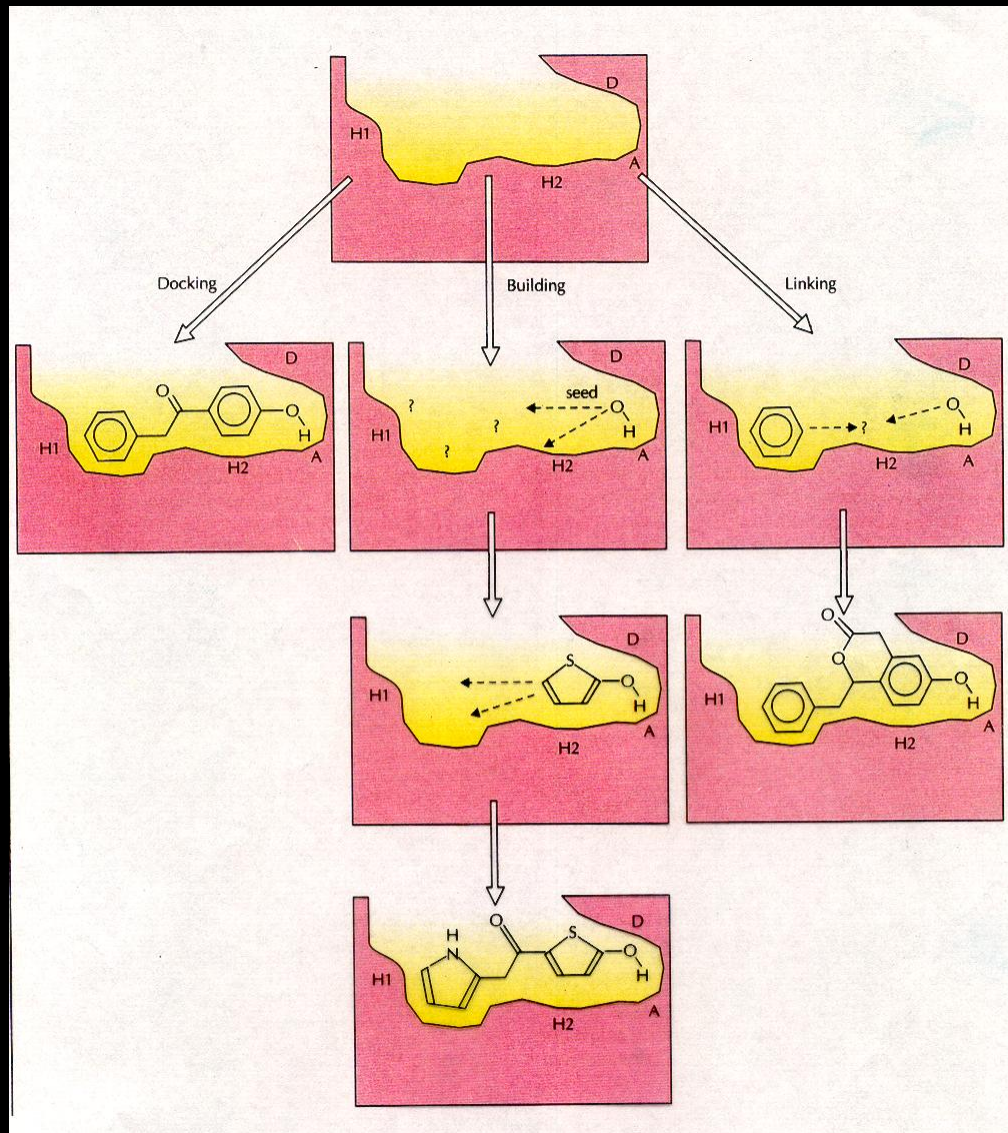


# Paradime in strcuture based drug design

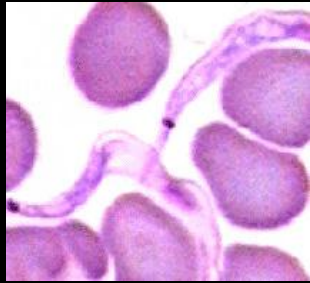


A good inhibitor should have a significant structural and chemical comlemetarity to the target receptor.

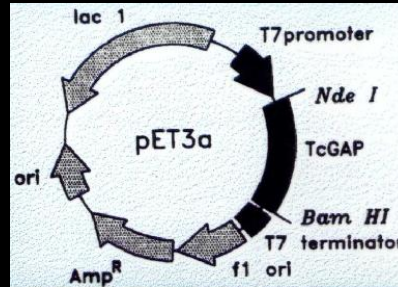
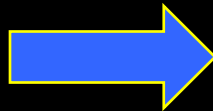
# Structure based drug target design



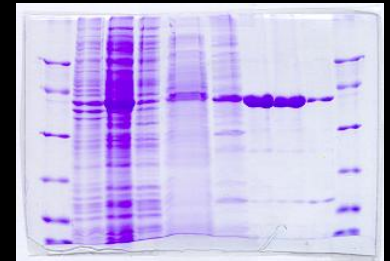
# Estratégia (1)



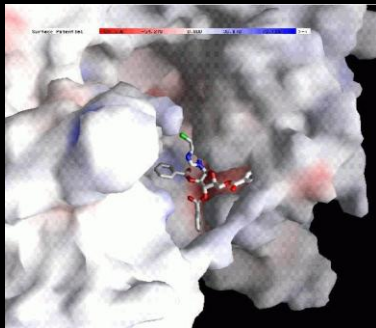
clonagem



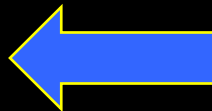
expressão



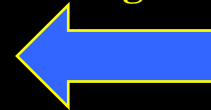
cristalização



Planejamento Racional

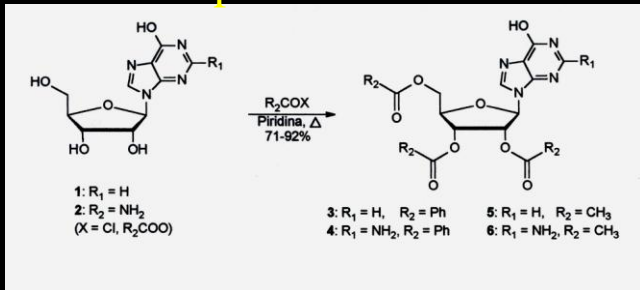


Estrutura Cristalográfica

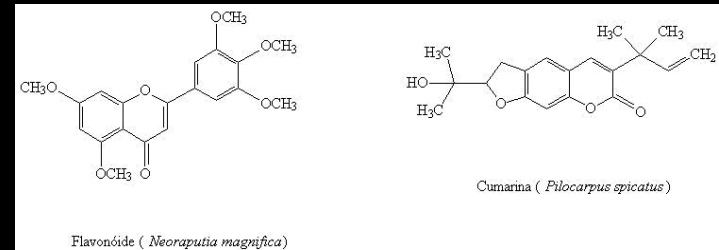


# Estratégia (2)

## Compostos sintéticos



## Produtos naturais

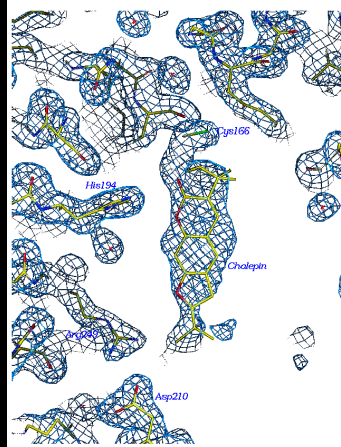


## Ensaio bioquímicos e biológicos

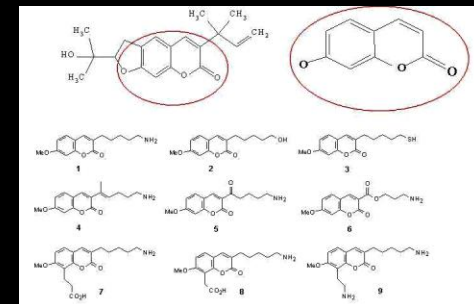
## Co-cristalização



## *T. cruzi* GAPDH-Chaperin complex

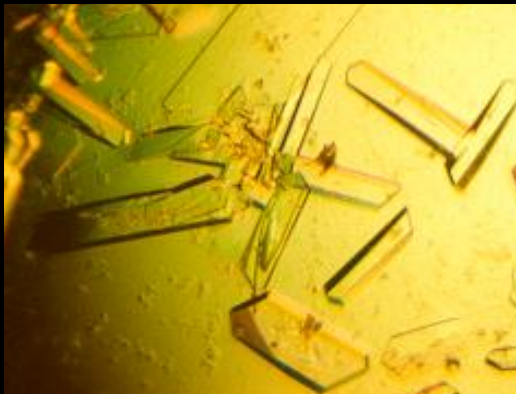
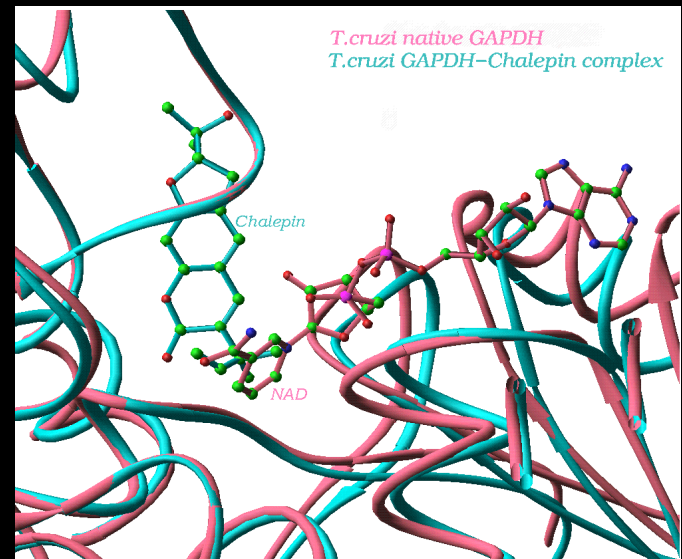
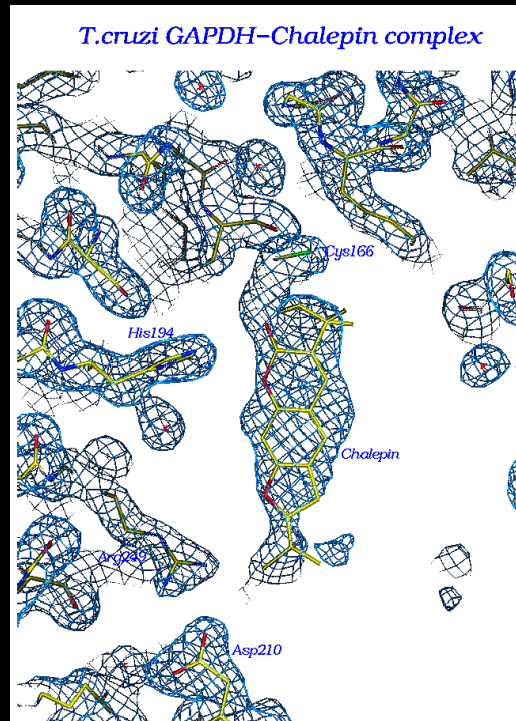
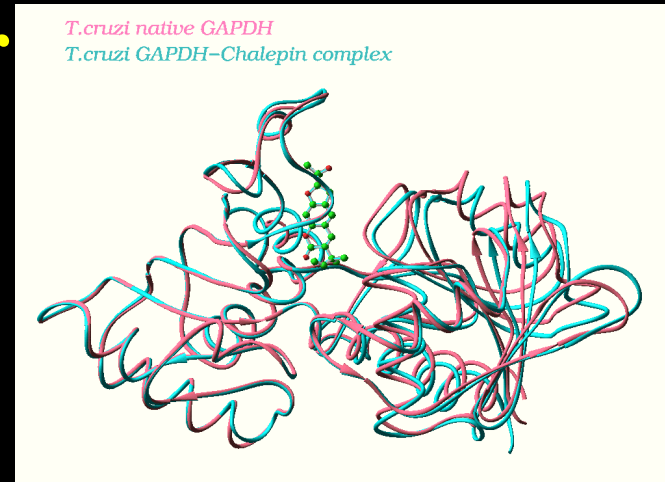
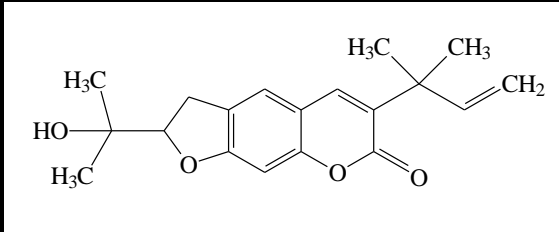


## Otimização Molecular





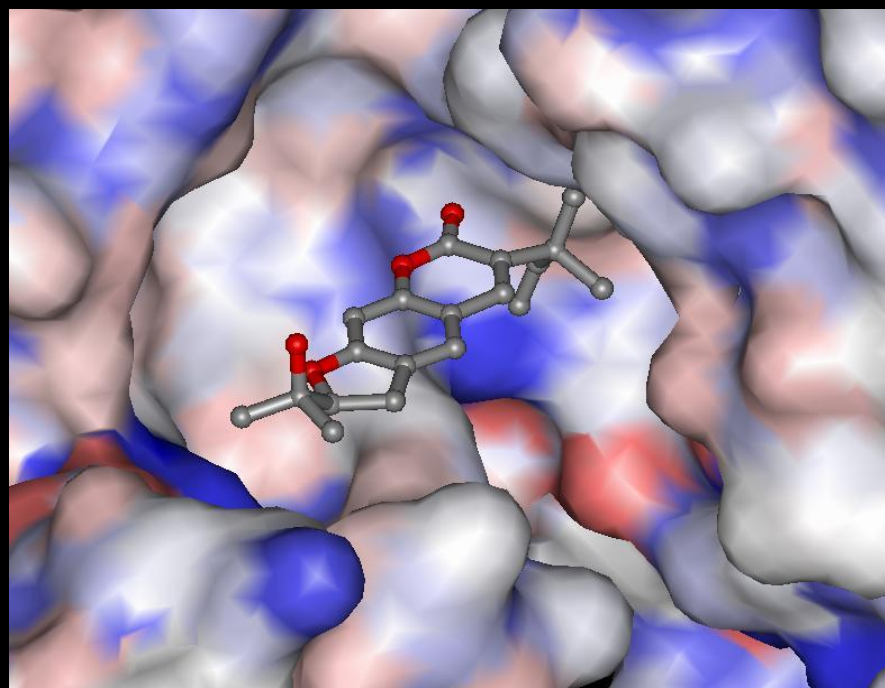
# Crystal structure of *T.cruzi* GAPDH in complex with natural product competitive inhibitor



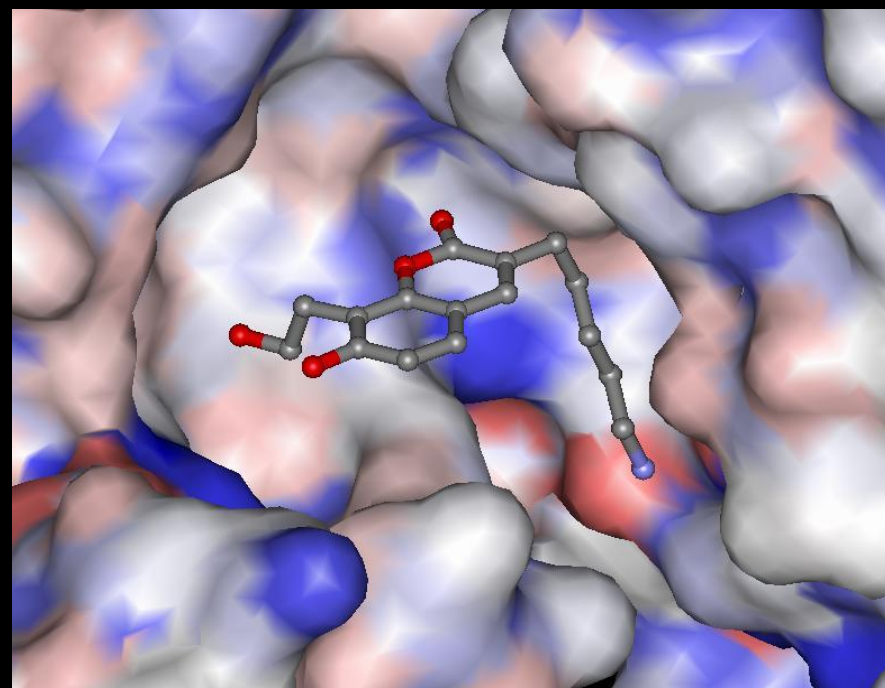
Space grown  
crystals



# Structure-based design of modified natural products derivatives



Chalepin at *T. cruzi*  
GAPDH active site



Compound SCG001

**Thank you**