The discovery of innovative therapeutic approaches:

Under the street light is not necessarily the right place to search.

Dr. François Iris, Founder & C.S.O. BIO-MODELING SYSTEMS

BIT's 10th Annual Congress
International Drug Discovery Science and Technology 2012
November 8-10, 2012, Nanjing, China.

Contact: www.bmsystems.net email: manuel.gea@bmsystems.net
The discovery of therapeutics.

It is first and foremost a matter of integrating huge masses of information.

But there are first THREE problems & ONE paradox to resolve.

**Problem 1. The nature of information**

All that goes by the name « *Information* » is not necessarily *Useful* and/or *Utilisable*!

In the bio-medical realm, the information available is *ALWAYS*:

- **incomplete**, to an unknown extent;
- **biased**, to an unknown extent; and
- **erroneous**, to an unknown extent.
This takes particular significance in drugs development

- **Sequence analysis**
- **Annotations**
- **Comparative genomics**
- **Pharmacogenomics**
- **Functional genomics**
- **Expression analysis**
- **Proteomics**
- **Structural genomics**
- **Pathway discovery**
- **Expression (arrays)**
- **Proteomics**
- **Pathway discovery**
- **HTS**
- **Combi. chemistry**
- **Expression (arrays)**
- **Proteomics**
- **Structural genomics**
- **Animal models**
- **In silico simulations**
- **Expression (arrays)**
- **Proteomics**
- **PK/PD**
- **Expression (arrays)**
- **Pharmacogenomics**
- **PK/PD**

A process generating a flood of heterogeneous information

© Bio-Modeling Systems  www.bmsystems.net
Problem 2. Therapeutic success.

The success of a therapeutic approach largely arises from the coherent manipulation of a physiological system as a whole and not from that of a target in a molecular context.

Therefore, any given medical problem must be approached from a systems standpoint.
Problem 3. The data analysis & integration tools we must utilise in systems-level approaches.
These approaches result in a highly misleading vision of protein interactions & networks.

How can a single hub protein bind so many different partners?

The problem is largely non-existent and resides in the construction and the representation of protein interaction networks within data-bases.

Proteins derived from a single gene, even if different, are clustered in maps into a single node.

This leads to the impression that a single protein binds to a very large number of partners.

In reality, it does not.

Protein networks reflect confusions involving combinations of functional plasticity addressing a same protein together with distinct physiological roles of different proteins encoded by one gene.
And the more complex the organism being analysed, the worse it gets.

Protein-Stress Association Network

- 3145 proteins linked to 32 stresses by 10777 relations
- On average
  - each protein associated with 3.4 stresses
  - each stress associated with 337 proteins
- Filtering associations based on three scoring metrics IP, M and N
- Which metric and cut-off are most suited for filtering noise?
This significantly affects the **Classical analytical Process**

What constitutes a **GOOD** filter?

What constitutes a **GOOD** indexation strategy?

**ALL** information entered into the DB is **ALWAYS biased, incomplet, erroneous**, etc...

The challenge is clearly not a question of technologies only!
The paradox that must be resolved.

If you dream to create the first operational bird model...

Be sure to use the appropriate modeling concepts & tools. If not...

... a “basic” living Complex system that not only flies...

...you get a Complicated “Cartesian” system. It does fly, but...

How to escape the paradox whereby we have NO OTHER CHOICE but utilise tools & approaches that CANNOT enable us to reach our goal?

By changing our intellectual approach!

© Bio-Modeling Systems www.bmsystems.net
Another way of thinking:
The CADI™ Integration & Modelling principles.

Data of Interest

- Injection
- Indexing
- Indexing & visualisation
- Identified components
- Nucleic acids, Proteins, etc.
- Séquences 1, 2, 3 ... n

SPAD
OMIM
MEDLine

Filters

Indexed DB

Data acquisition

Identified components

The DB contents are analytically IRRELEVANT

Negative Selection of hypotheses

This constitutes a heuristic approach!

© BMSystems Confidential www.bmsystems.net
The differences between « heuristic » and « mathematical » approaches.

Heuristics:
A problems solving approach evaluating each step in a process, searching for satisfactory solutions rather than for optimal solutions, using all available qualitative information instead of quantitative information.

Thus,

**Heuristic modelling** starts from accumulated information to produce a model capable of describing the mechanisms that generated the observed outcome / data and predict their modifications associated with a different outcome;

*It plays the role of an architect*

While

**mathematical (Bayesian) modelling** starts from quantitative data to produce models capable of reiterating this data and predict the outcome of a different experimental paradigm.

*It plays the role of an engineer*

and

*Far from being incompatible, these two approaches can be complementary.*
Bayesian and Heuristic approaches can be complementary, provided they are harnessed in the proper order.

Bayesian approaches are of limited usefulness when applied to poorly defined multicellular physiological systems because they cannot efficiently reveal & define the functional states within such a system (cross-talks alterations, etc...).

But heuristic approaches are very efficient at doing precisely this.

Heuristic models are of limited usefulness when addressing the dynamics of defined complex physiological pathways structures and cross-talks because they are not open to mathematical manipulations.

But Bayesian models are very efficient at doing precisely this.

To efficiently address the translation of systems biology to clinical & medical interventions (dominated by patient's data heterogeneity and largely unstructured documents), ways to achieve synergy between Heuristic and Bayesian approaches can be effectively designed.
The heuristic analytical process must follow a « relativistic » approach.

Within this framework, **Non-linearity, Irrelevance, Wear, Relative weights & Contexts** are key concepts.
Why? Because of a very simple set of rules which imposes « relativistic » approaches.

**Events** tell **contexts** how to **evolve**

**Contexts** tell **components** how to **behave**

**Components** tell **events** how to **arise**

Analyses in terms of biological components and functions are now **IRRELEVANT**.

**EVENT-DRIVEN** (relativistic) analytical approaches become necessary.

This, in turn, imposes analytical procedures based upon the **negative selection** of working hypotheses.
Why “negative selection” of working hypotheses?

“While it is not always possible to demonstrate that a statement is true, it is always possible to demonstrate it to be false” Karl Popper, 1963.

Mathematical approaches are based on “positive selection”: it is assumed that every dataset/statement is actually valid.

Yet, “an estimated 85% of current published research findings are false or exaggerated”


Positive selection becomes a killer!

How to identify what is NOT false and/or exaggerated?

By doing every thing possible to destroy working hypotheses!

Only hypotheses that resist destruction are worth retaining.

Hence, what is demonstrated “False” can now be used to discover what could be “True”.
The *CADI*™ Integration workflow

What is **FALSE** allows to discover what could be **TRUE**!

- Data acquisition
  - Experimental verification
- Biological Validation
- Hypothetic Physiological Mechanism
- Biological Modeling
- Components Interactions maps
- Destructive hypotheses testing
- Production of working hypotheses
- Database Searching
- Identified biological events
- Literature
- Nucleic acids, Proteins, Etc...

© BMSystems  Confidential    www.bmsystems.net
MCF-7/MCF-7 ras Differentially Expressed Genes

TCP 1 θ
Agrin / Follistatin
PTK7
ARP-1
Rab 6 homologue
Cathepsin D
HnRNP-M
Hps 1
Thymosin b4
ARP 3
Cytokeratin 18 and 19
Podocalxin
Ubiquitin-Thiolesterase
AP-2 α
Rab 6 homologue
HnRNP-E1
HsN3
HSP 70
L-Plastin
Ubiquitin
PP4 (regulatory subunit)
Protein 14.3.3. ε
DKK-1
SAP49
Reticulocalbin
TFF-1
TACE
AP-2 α
ARP 3

Data acquisition
Gene identification
Data verification
Physiological mechanism hypothesis
Biological validation
Interaction maps
Database searching
Scientific literature
Hypothesis testing

HSP 27
SSRP1
HSP 27
67LR
HDAC-3
FKBP-52
Il d-1
TGF β-masking protein

Calpain (small subunit)
Dipeptidyl peptidase III
Annexin VI
Olfactomedin-like
LAR

SWAP 2
ORS-1
p80 coilin
PRP 16
HSP 47

RCC-1

© BMSystems Confidential www.bmsystems.net
Information Retrieval & Extraction

Set of genes

TCP 1 θ
HPS 1
hnRNPM-1
Thymosin β4
51-PK(s)
ARP 3

HSP 70
hnRNP-E1
PTK7
ARP-1

Calpain (small subunit)
Dipeptidyl peptidase III
Cathepsin D
Ubiquitin-Thiolesterase

Podocalixin
Ubiquitin
RMB-6
PP4 (regulatory subunit)

Protein 14.3.3. ε
Bax-α
DAGK

MA-3
Nh3RF

ATPase

p38 δ kinase
FKBP-52

STE5

MDC15
CD 24

dead Box

X caveolin

Ese-1a


© BMSystems
Confidential
www.bmsystems.net
Prions are thought to consist of infectious proteins that cause transmissible spongiform encephalopathies. According to overwhelming evidence, the pathogenic prion protein PrPSc converts its host encoded isoform PrPC into insoluble aggregates of PrPSc, concomitant with pathological modifications (for review, see refs. 1-3). Although the physiological role of PrPC is poorly understood, studies with PrP knockout mice demonstrated that PrPC is required for the development of prion diseases. Using the yeast two-hybrid technology in Saccharomyces cerevisiae, we identified the 37-kDa laminin receptor precursor (LRP) as interacting with the cellular prion protein PrPC. Mapping analysis of the LRP-PrP interaction site in S. cerevisiae revealed that PrP and laminin share the same binding domain (amino acids 161 to 180) on LRP. The LRP-PrP interaction was confirmed in vivo in insect (Sf9) and mammalian cells (COS-7). The LRP level was increased in scrapie-infected murine N2a cells and in brain and spleen of scrapie-infected mice. In contrast, the LRP concentration was not significantly altered in these organs from mice infected with the bovine spongiform encephalopathic agent (BSE), which have a lower PrPSc accumulation. LRP levels, however, were dramatically increased in brain and pancreas, slightly increased in the spleen and not altered in the liver of crapie-infected hamsters. …
Prions are thought to consist of infectious proteins that cause transmissible spongiform encephalopathies. According to overwhelming evidence, the pathogenic prion protein PrPSc converts its host encoded isoform PrPC into insoluble aggregates of PrPSc concomitant with pathological modifications (for review, see refs. 1-3). Although the physiological role of PrPC is poorly understood, studies with PrP knockout mice demonstrated that PrPC is required for the development of prion diseases. Using the yeast two-hybrid technology in Saccharomyces cerevisiae, we identified the 37-kDa laminin receptor precursor (LRP) as interacting with the cellular prion protein PrPC. Mapping analysis of the LRP PrP interaction site in S. cerevisiae revealed that PrP and laminin share the same binding domain (amino acids 161 to 180) on LRP. The LRP-PrP interaction was confirmed in vivo in insect (Sf9) and mammalian cells (COS-7). The LRP level was increased in scrapie-infected murine N2a cells and in brain and spleen of scrapie-infected mice. In contrast, the LRP concentration was not significantly altered in these organs from mice infected with the bovine spongiform encephalopathic agent (BSE), which have a lower PrPSc accumulation. LRP levels, however, were dramatically increased in brain and pancreas, slightly increased in the spleen and not altered in the liver of scrapie-infected hamsters. …
Generating hypotheses & systematic destructive testing.

Index visualisation & manipulation

RNA Splicing & expression

Apoptosis & cell survival

© BMSystems  Confidential  www.bmsystems.net
Interaction maps describe the pathways that have become functional and those that have become forbidden in response to local conditions imposed by the activation of defined biological mechanisms.

Specific biological events do not occur because they are fated to. They occur because other events could not!
The CADI™ Integration & Modelling Process.

This iterative process does three things:
- It largely resolves the coherence issues attached to the classical approach;
- It reveals hitherto unknown mechanisms/processes, and
- It allows the translation of systems biology to clinical & medical interventions.
As a result,

Searching under the street lamp because that is where there is light may be much less helpful than expected.

Indeed, since it is the presence/absence of specifiable events which govern phenotypic characteristics, it follows that

The mechanisms that characterise a given phenotype (be it pathological or not) within a given biological system can be utilised as analytical tools to unravel those associated with any other phenotypic transition affecting the same biological system.

Here, symptomatology ceases to function as defining criteria to become contextual and relative functional end-products.

Therefore, the solution to a given biological problem can be efficiently obtained through entirely indirect investigations.

This is particularly true for problems resistant to direct approaches because characterised by

- High symptomatologic heterogeneity, and/or
- High functional/phenotypic uncertainty.
Example 1:
Understanding the *in vivo* mechanisms of Creutzfeldt-Jakob Disease pathogenesis & progression leads to the discovery of improved treatment for non-degenerative psychiatric disorders.

CJD: A few facts.

- Progressive neurodegenerative disease, invariably fatal.
- Long asymptomatic incubation phase (30 years + in man)
- Short clinical phase (death within 6 months - 2 years)
- Pathological agent: abnormally folded form of the PrP protein
- Pathological mechanisms: Unknown
- Clinical progression mechanisms: Unknown
- Mode of propagation within the CNS: Unknown
- Therapeutic or preventive means: Unknown
- Pathogenesis Biomarkers: Clinical symptoms + CNS spongiosis (post-mortem).

© Bio-Modeling Systems  www.bmsystems.net
The pathogenesis & clinical progression model

**Early Phases**

- Prolonged opening of Calcium channels
- Activation of PLA2-PLC pathways
- Secretion-accumulation of Pro-inflammatory molecules
- IL-1β, IL-6, TNFα
- NFκB, Jnk pathways activated

**Microglial cells**

**Proliferation**

**Late Phases**

- Prolonged opening of Calcium channels
- Traf6-PI3K-PDK1-PKC pathways activated
- RhoA-ROCK-FA pathways inhibited
- Cytoskeleton modifications
- Moesin+Ezrin
- GFAP, Vim, Traf6
- PI3K-PDK1-PKC pathways activated
- RhoA-ROCK-FA pathways inhibited

**Astrocytes**

**Proliferation**

**PrPres or derived peptides**

- Accumulation of PrPres
- Loss of PrPc function (Antioxidative protection)
- NMDA Activation
- CYP + Cox1/2 pathways activated
- Excitotoxicity
- Free radicals
- Lipids peroxydation
- Cytoskeleton alterations
- De-afferentation
- Neuronal death

**Neurons**

- Glutamate
- Prolonged activation AMPAR GLUR1
- Integrons pathways activated
- Partial loss of inter-astrocyte & Astrocyte-synapses contacts

**NMDAR**

- Iris F (2012); Pharmacopsych. 45 (Suppl. 1): S12-21.

© Bio-Modeling Systems www.bmsystems.net
The main driving mechanism.

II-1β & TGFβ-mediated signalling in hippocampus astrocytes

The pathways through which chronic neuronal stress signalling and concurrent glial pro-inflammatory responses lead to reactive astrocyte activation (GFAP + Vim) associated with cytoskeleton reorganisation (ezrin). This leads to a major switch in Cx targeting & distribution, resulting in the formation of a syncytium with massively altered diffusive properties and neurotrophic functions.
Effects of the modifications of glial connexions upon intercellular diffusion processes.

Healthy hamster

Stereotaxic Injection of 500nl fluorescent tracer < 1000d

280µm

Scrapies hamster

Stereotaxic Injection of 500nl fluorescent tracer < 1000d

1080µm

The effects are much more wide-spread than anticipated!
Functional alterations linked with modifications of glial connections.

Healthy mice

Scrapies mice (asymptomatic phase).

© Bio-Modeling Systems  www.bmsystems.net
Practical consequences.

One of the roles of connexins is to dampen neuronal synchronisation.

In healthy animals, pharmacological blockade of Cx activity results in quantitative EEG patterns closely resembling an epileptic crisis (frequency range-specific hyper-synchronisations).

This CEA/BMSystems collaborative research in CNS (psychiatric and neurological disorders) led to the co-owned patent WO201029131 (Use of anti-connexin agents for modulating the therapeutic effect of psychotropic drugs). Sept, 2008.
The net results.

CJD is not a neurological disease *stricto sensus*. It is a disease that primarily affects astrocytes structures and functions which, over time, lethally affects healthy glial & neuronal cells through « bystander effects », leading to widespread CNS disorganisation (spongiosis) and functional failure.

But this model also provides an understanding of key mechanisms associated with psychiatric & neurology disorders.

An entirely new approach for their effective treatment was designed, tested *in vivo* and validated.

Patent covering novel therapeutics for psychiatry & neurodegenerative disorder (CEA/BMSystems).

This CEA/BMSystems collaborative research in CNS (psychiatric and neurological disorders) led to the co-owned patent WO201029131 with a worldwide exclusive license to Theranexus CEA’s spin-off currently in Phase II.

Neither of which have much to do with CJD *per se*...

This work received a Bio-IT World « Best Practices » award from the Cambridge HealthTech Institute (USA).

AND

Was selected as 1 of the 3 pan-European « state of the art examples of systems biology approaches of benefit to medicine » by the European Commission’s DG Research, Directorate of Health (June 2010).
Example 2:

Understanding the co-evolutionary interplays between bacteria and bacteriophages leads to the discovery of the means whereby _undefined_ multi-resistant bacterial pathogens can be efficiently controlled.

The questions (French Defence)

- How to _rapidly_ (less than 30 min) and _efficiently_ _detect_ the presence of any given _LIVE_ bacterial pathogen?

- How to _rapidly_ and _efficiently_ _destroy_ any _unknown_ bacterial pathogen or emerging strain _without using_

  A) _Antibiotics_: too many resistant strains, and very rapid resistance acquisition.

  B) _Vaccines_: much too slow to act, and small strain variations often lead to inefficacy.

In other words, what is required is a “detector-killer”.

© Bio-Modeling Systems  www.bmsystems.net
Bacteriophages, the natural predators of bacteria, could present the best potential to act as detectors-killers.

- Many are very host-specific,
- They only replicate in LIVE bacterial cells,
- Many kill the cells in which they replicate,
- As the phage progeny population increases that of the target diminishes (in a « closed » environment, few targets, if any, should escape), and
- They are extremely numerous and varied (they probably represent the most numerous « life forms » on the planet).

BUT the matter is NOT as simple as it first appears!
Co-evolution versus unidirectional predatory pressure.

Natural Hosts pool → Natural Phage pool → Target Hosts Pop. → Isolated Phage X

Continuous co-evolution → Ever increasing evolutionary divergence → Ever decreasing chances to find a new efficient lytic phage.

Bacteria have existed for nearly **4 BILLION Years**. They have so far **resisted to EVERYTHING**. And it is certainly **NOT** for lack of phages!
The model-derived solution.

• **What, in essence, is the problem?**

The bacterial targets will try anything to escape predation and we have no idea what will be the successful strategy. Furthermore, this strategy is likely to vary between locations (populations) for a same target.

• **What do we need to achieve?**

We must be capable of always preceding the targets escape strategies, no matter what they could be.

• **The best-fit solution** (model-derived):

  we MUST adopt a stochastic approach.

  It becomes necessary to
  • abandon all idea of « natural phage pools » and,
  • stochastically engineer phage banks in order to produce particles capable of targeting anything and everything while maintaining their capacity to replicate in the face of targets evasion attempts.
The problems:

phage T4

How to modify any of these proteins in \( N \) different regions, at \( X \) different sites, in \( Z \) different manners, all this simultaneously and then recombine the multitude of variants generated in a population of obligate lytic phages?
The technological answers.

Three proprietary technologies (invented at BMSystems) allowing the production of stochastically engineered phage banks.

**TAPE** (WO 2008/093009):
A technology allowing to rapidly & simultaneously introduce defined densities of random mutations in any number of selected regions within a gene while conserving intact any number of defined coding domains in this same gene.

*Applicable to any known coding sequence.*

**RipH** (WO 2009/090081):
A technology allowing to reversibly inactivate the genome of an obligate lytic phage within its host and carry out high efficiency homologous recombinations targeting multiple genes simultaneously without adversely affecting the host bacteria and the replicative capacity of the phage.

**Ab-ACCUS** (WO 2008/093010):
A recombination technology allowing the rapid & efficient production of lytic phage banks in which every individual differs from all others for any number of selected genes or other sequences.

*Applicable to any phage and to any known sequence.*
The results.

While T4 is specific to a narrow range of E. coli strains, An engineered T4 bank contains variants capable of detecting and killing gram\(^-\) bacteria far removed from E. coli.

**Yersinia sp.**

- **A** (LPs)
- **B** (No lysis)

<table>
<thead>
<tr>
<th>Dilution</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
</tbody>
</table>

**Pseudomonas sp.**

- **D** (No lysis)
- **E** (LPs)

- **F** (Mutation levels: 0.6%, 1%, 2%)

Practical consequences
(besides high-profile publications)

• Three technological patents with broad applications,

• Creation and financing of a Bio-Pharma company (Pherecydes Pharma) specialised in biodefense & biosecurity (40% owned by BM-Systems),

• Research program with French civil and defence health-services,

• Discussions with US food industry firms,

• Contract with anti-bodies producer & discussions with enzymes producers.

All this in less than 3 years.

From our point of view, under the street light is definitely not the right place to search!
Nevertheless, it MUST be remembered that Models are AIDS to thought NOT a replacement for it.
Our collaborative R&D programs & their outputs
This list excludes our contractual research programs with our clients


Max Planck Institute (Munich): Project “Chronic Anxiety”.

Successfully completed; 3 publications & a Reference Book “Biomarkers for Psychiatric disorders” chapter 19.

INSERM: 3 Projects “Tumoral Progression”; “Therapeutic Resistance”; “RGD 15 & Metastasis”. All 3 successfully completed, 3 publications.

CNRS: Project “Müllerian Regression” Tissue differentiation Successfully completed, 1 publication.

Foundation FondaMental: Project “Bipolar Disorders & Schizophrenia”. Immuno-inflammatory hypothesis. On going, 1 publication pending.

L’OREAL Arkema, Rhodia/Solvay ARD: “Synthons” Government funded feasibility Program at IAR cluster Industrial Biotech Feasibility study Completed 16 molecules evaluated, 2 strains built, 1 program with 1 patent (industrial partner only) - Skin Homeostasis: Reference book “Computational Biophysics of Skin” chapter 15 with Dr. Querleux (L’Oréal)

Centre of excellence in Epigenetics IISER Pune India: Project “Etiology & Epigenetic for metabolic disorders” Etiology & Epigenetic for metabolic disorders, on going 1 publication pending.
Useful Downloads

For more information about information quality & reliability

- A new evidence published in Sciences confirms the poor reproducibility (less than 1/3) of studies published in peer-reviewed.
- An estimated 85% of current published research findings are false or exaggerated: How to Make More Published Research True. Published in PLOS Medicine by John P. A. Ioannidis Meta-Research Innovation Center at Stanford (METRICS), Stanford University.
- Diagnosing the decline in pharmaceutical R&D efficiency. Published Nature Review Drug Discovery. The diagnostic is clear for our industry.
- Believe it or not: how much can we rely on published data on potential drugs targets? Their title is crystal clear. Published Nature Review Drug Discovery

Heuristic modeling principles and case studies

- The discovery of Innovative Therapeutic Approaches: Under the street light is not the right place to search BIT’s 10th Annual Congress International Drug Discovery Science and Technology 2012 November 8-10, 2012, Nanjing, China
- The Differences & Complementarities Between «Heuristic» and «Mathematical» approaches. The scientific presentation given by Dr. François IRIS (CSO BMSystems) during the EPA (European Psychiatric Association) conference in 2011.

For more information Email: manuel.gea@bmsystems.net
Questions

Contact:  www.bmsystems.net
Email: manuel.gea@bmsystems.net