Open Archive TOULOUSE Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

The InBioSynSolv Project

Virtual laboratory for synthons and biosolvent design based on reverse molecular engineering

Vincent.Gerbaud@ensiacet.fr

Financing: French ANR program CP2D 2009

Computer and product modeling leader

InBioSynSolv

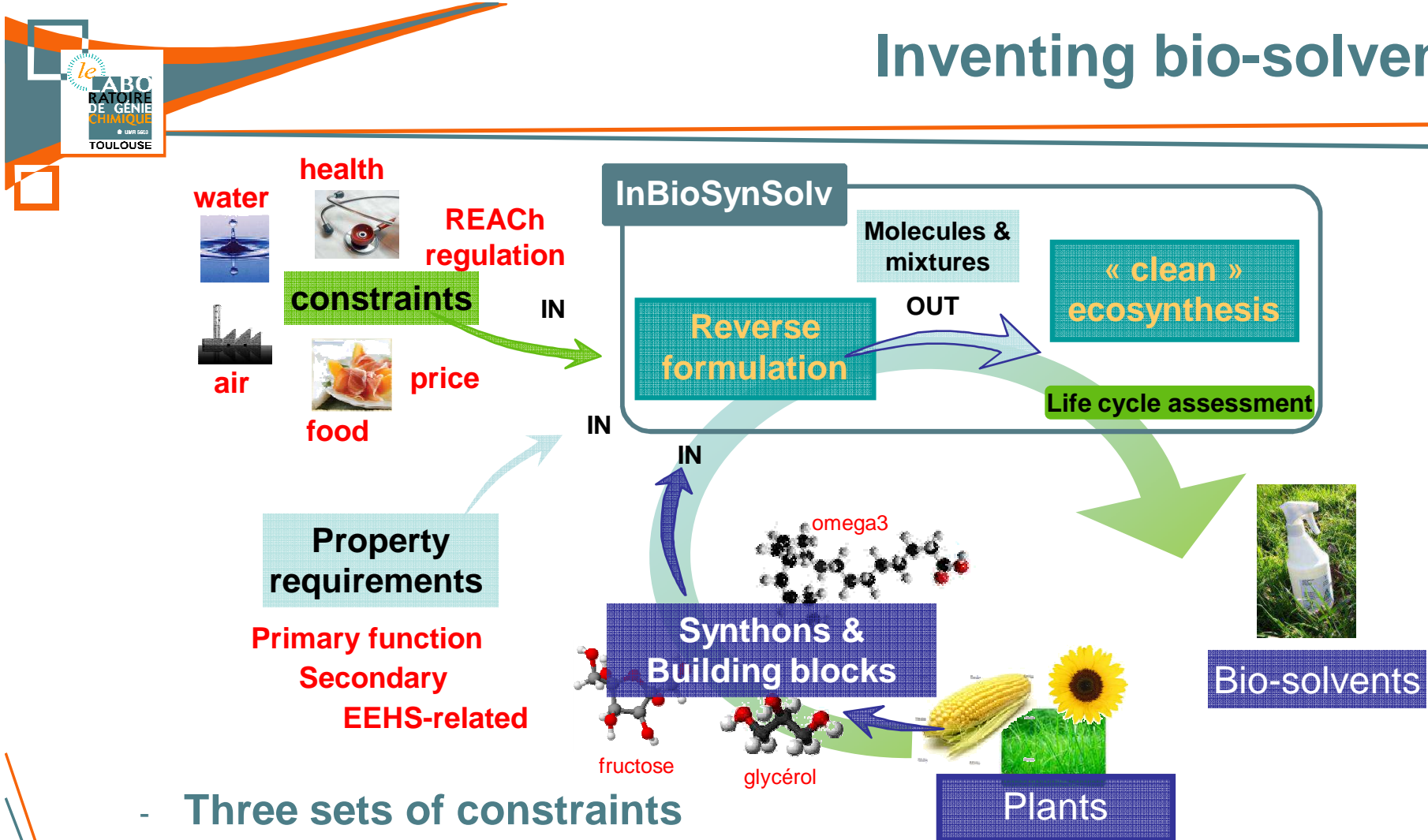**Solvent chemistry and formulation**

Solvents and applications

Agro-sourced solvents

- *Project accredited by the french pôles de compétivitité*
  - AXELERA, MAUD
- *4 years*
  - Start: 01/10/2009 end: 30/09/2013
- *Budget: 1.2 M€ financed. total 2.9 M€*

Inventing bio-solvents

water
health
REACh regulation
constraints
air
price
food

InBioSynSolv

Molecules & mixtures

« clean » ecosynthesis

Reverse formulation

IN

OUT

IN

IN

Life cycle assessment

Property requirements

Primary function
Secondary
EEHS-related

omega3

Synthons & Building blocks

fructose

glycérol

Plants

Bio-solvents

- **Three sets of constraints**
  - **EEHS & REACh regulation**
    - *finding or subtituting greener and safer molecules*
  - **Products**
    - *often single purpose but must comply with many properties*
  - **« green » is marketable**
    - *Renewable resources*

InBioSynSolv Goals & Challenges

- Use **reverse engineering of molecules (CAMD Computer Aided Molecular Design)**

- **Develop a software tool** to perform CAMD

- **Synthesise and formulate** the most promising molecules

InBioSynSolv Goals & Challenges

- Use **reverse engineering of molecules (CAMD Computer Aided Molecular Design)**
  - **Challenges**
    - *Avoid combinatorial explosion while searching molecules from building blocks*
    - *Use predictive property models (and accurate)*

- **Develop a software tool** to perform CAMD
  - **Challenges**
    - *Easy-to-use*
      - **Identify end-users, either experts or basic**
    - *Flexible for any future change*
      - **In regulations, in resources, in synthesis pathways in modelling**

- **Synthesise and formulate** the most promising molecules
  - **Challenges**
    - *Find molecules that can be synthesized at minimum cost*
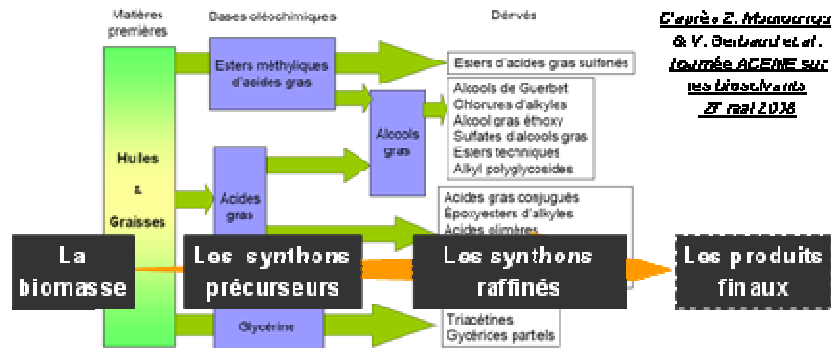   - *Satisfy green chemistry principles for molecule synthesis*

- Use of synthons rather than raw biomass
 - **Synthons are more reliable**

- Price
  - **The formulated solvent should be affordable**
    - *Start from cheap synthons*
    - *Keep only manageable and cheap synthesis path ways*

- Solution
  - **10 main pools of renewable resources have been identified from the literature**
    - *eg glycerol (C3 source)*
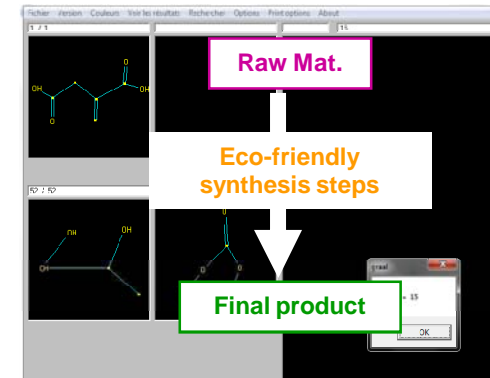
(Angew. chem. Int. Ed. 2007)
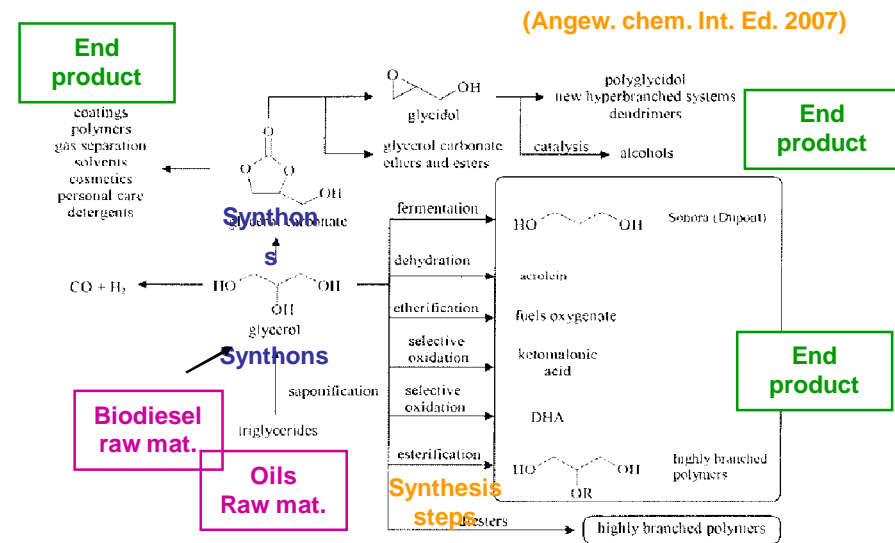
- **Bottom up (the common way)**

  - **Identification of renewable resources pools**
    - *glycerol, itaconic acid, …*
  - **Definition of 30-50 affordable and eco-friendly synthesis pathways**
  - **Automatic generation of possible molecules**
  - *A posteriori* **evaluation of the molecules performance**
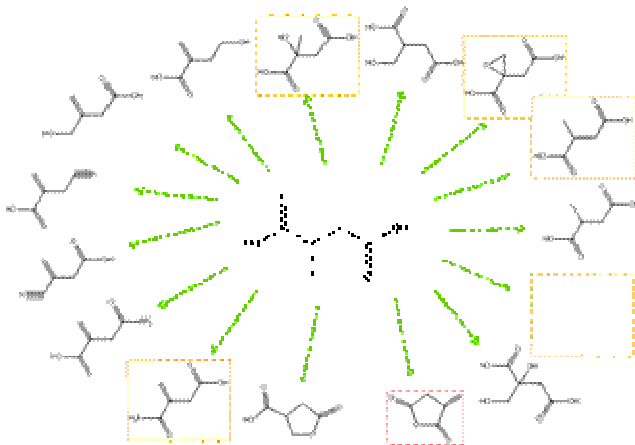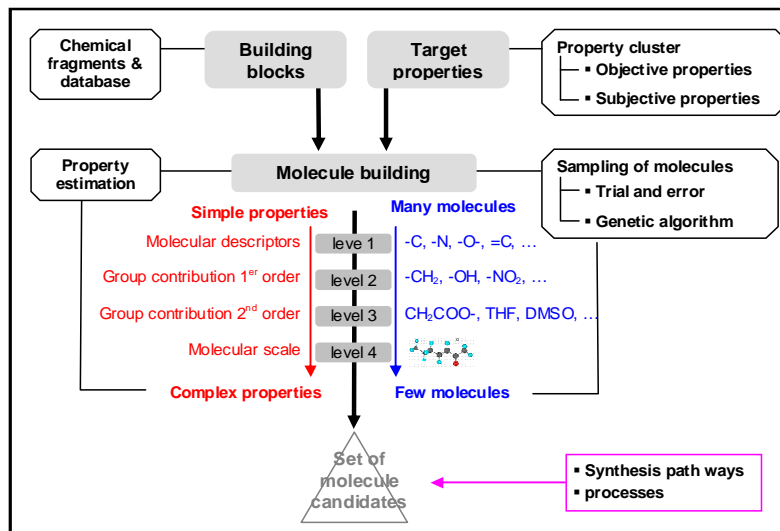

- **Results**

  - **GRASS: software for oriented synthesis**

  - **Pro:**
    - *Affordable molecules can be produced*

  - **Against:**
    - *Exponential number as reactions steps increase*
      - **>4000 molecules after 3 steps!**
    - *Post-evaluation = high rejection and pb for many properties*

- Top Down (CAMD)

- • **Target a set of many properties and find (build) candidate molecules and mixtures that satisfy them**
- • **Choice of building blocks from renewable resources**
- • **Specification of MANY target properties and values**
  - • *(solvent, VOC, FP, toxicity,)*

- Results

- • **Design of a flexible CAMD software**
- • **Solutions are chemically functionnalized synthons from renewable resources**

perf. 0.9985   perf. 0.9981   perf. 0.9973

- • **Pro:**
  - • *Candidates molecule satisfy the properties*
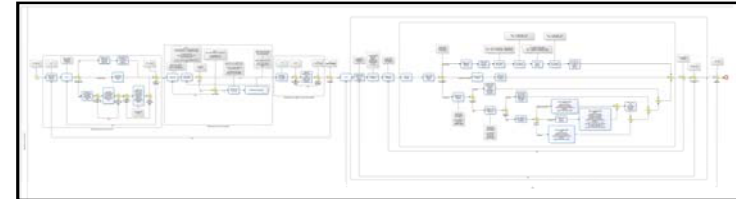- • **Against:**
  - • *No clue about synthetizability of molecules*

- **More on the Top Down approach (CAMD)**

  - **Design of a component-based software from UML**
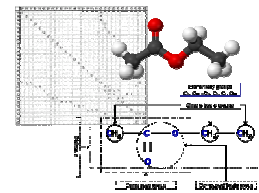
    - *Search and evaluation component*
    - *Molecule building component*
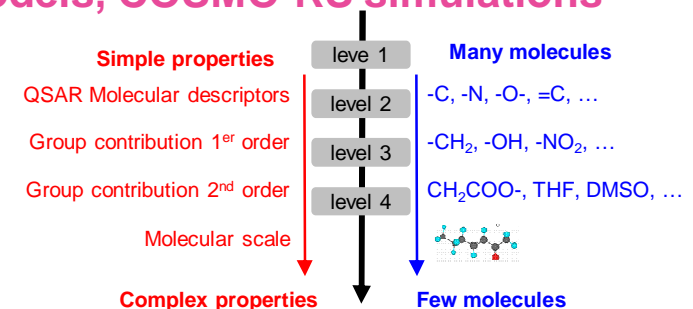   - *Property calculation component*

      – property component is also used for fast checking of the thousands of candidates of the Bottom Up approach
      – Easy addition of new property models

  - **Multiscale modelling of molecules for many property**

    - *Depict molecules in a format suitable for any property model*

      – We use QSAR, knowledge based models, COSMO-RS simulations

      Ethyl acetate

      | Simple properties | leve 1 | Many molecules |
      | QSAR Molecular descriptors | level 2 | -C, -N, -O-, =C, … |
      | Group contribution 1er order | level 3 | -CH$_2$, -OH, -NO$_2$, … |
      | Group contribution 2nd order | level 4 | CH$_2$COO-, THF, DMSO, … |
      | Molecular scale | | |
      | Complex properties | | Few molecules |

  - **Flexible software tool**

    - *Candidates can be mixtures*
    - *Easy management of building blocks list*

      – *The C3 source, the C5 source, …*

InBioSynSolv conclusions

- ## Main result:

  - **Combine bottom-up and top-down (CAMD) approaches in *in silico* tool to suggest property-matching solvent candidates and evaluates their synthesis easiness.**

- ## The cooperative effect

  - **No single partner can provide the solution alone**

    - ### *Chemist expert*
      - Solvent formulation
      - Definition of building blocks
      - Synthesis pathways and Lab synthesis

    - ### *Industrial expert*
      - Application and property values
      - Industrial synthesis pathways

    - ### *Modelling expert*
      - Property models choice and evaluation
      - Software engineering

Thank you for your attention

Vincent.Gerbaud@ensiacet.fr