Evotec Registration System
Catherine Reisser

Evotec

- A drug research and development company with an IND engine that can deliver innovation to its partners
- Focused on neuroscience and related areas for our proprietary programmes
- A strong drug discovery platform that drives
  - A range of research collaborations in various therapeutic areas
  - Our proprietary discovery programmes within neuroscience
- Our core business is collaborative research and licensing

IND = Investigational New Drug

Evotec worldwide

Aim of the Evotec Registration System

- Track all the compounds we have ever made for an Internal Project or for an External Customer
- Track all the samples sent off site, to a Customer or to Hamburg for Screening
- Enforce Exclusivity for each compound
- Each piece of the data entered should be extractable to provide the data to each client in the format they need

Why did we choose to write a Registration System in-house?

- Flexibility
- Within the last 12 months
  - Evotec has sold its scale-up & pilot-plant facilities to Aptuit
  - Evotec has formed a joined venture with RSIL in India
  - Evotec has put in an offer to buy Renovis
- What has changed since we have started using our current Registration tool?
  - Less need for Large Combinatorial Library Enumeration
  - More focussed Small Libraries & more medicinal Chemistry projects
  - Security, e.g. making sure our own Internal Research Division cannot access external customers data

How are we different from a Pharmaceutical Company?

- Exclusivity – A compound or a small series of compounds is granted an Exclusivity which varies from Project to Project
  - A customer owns the rights for that compound for a limited period or forever.
  - Evotec cannot make the same compound for either any internal drug discovery program or another customer during that period
- We do not only store our data but some of our customer's data
  - Library Identifier / Compound Identifier / Sample Identifier

Why did we choose the ChemAxon tools?

- Marvin and JChem have been successfully incorporated into the following tools
  - ELN — Electronic Lab Notebook
  - E-Req — Chemical Requisition System
  - ESD / SSD — Evotec Supplier Database / Screening Database
  - EVOseek™ — Evotec Structure Activity Database
  - ESMA — EVOseek™ Searching and Modelling Application
  - Intranet Chemistry Forums
- Our license agreement will allow us to give access to all our scientists across all sites
- Java is our preferred development platform
- Single vendor toolkit

Our Development Approach

- Modular Approach
  - Requirements Meetings with Users
  - Functional Design Document where all functionality and screens are described
  - Database design
  - Technical Design Document & Build Plan
  - Testing

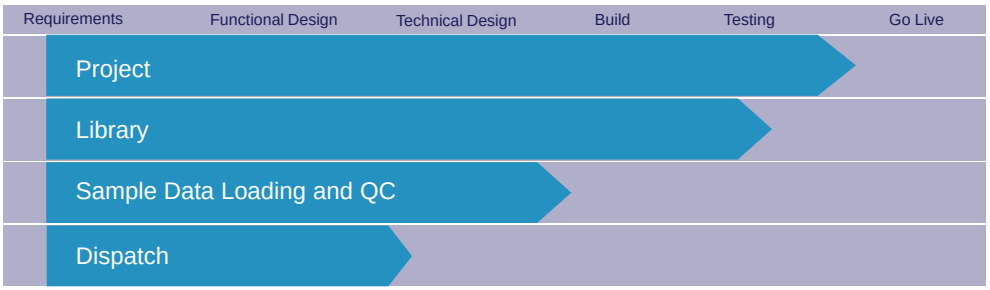
Registration Application Modules

- Project Module
  - All our data is organised around Contract / Project and Users and Library Templates
  - A Library Template is a sub-project around a biological target or a common sub-structure
- Library Module – Compound pre-registration
  - Compounds can be added one by one using Marvin Sketch
  - Compounds can be imported from an SD File
  - Compounds can be the product of a Combinatorial Library Enumeration

Registration Application Modules

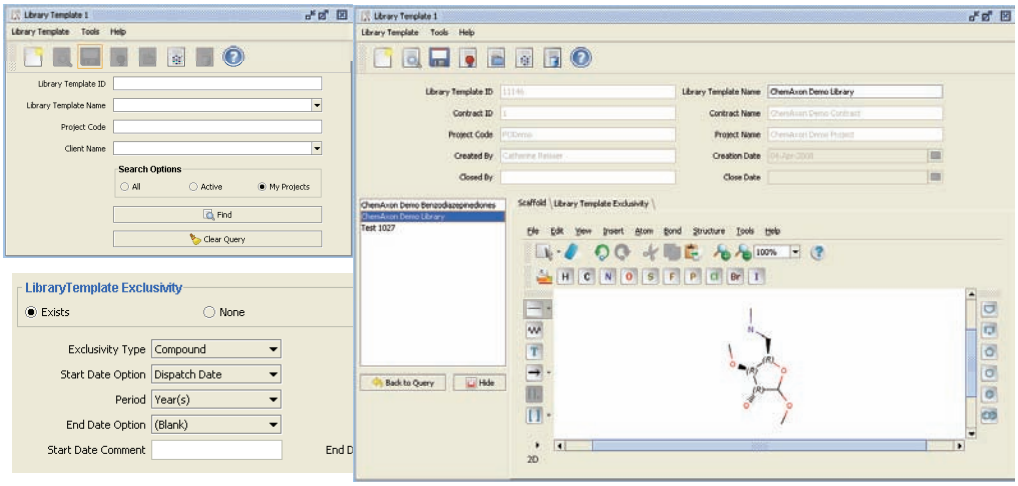
- Sample Data Loading & QC Module
  - Import Tool - Loading all sample data
  - Searching data by most fields including structure and Displaying the results
- Dispatch Module
  - Overlap Tool – Checking prior to Registration that these compounds do not belong to someone else
  - Registration – Assign a unique Evotec ID to each new molecule
  - Export Tool – Export all data in a flexible format (e.g. rename fields, format date and numeric fields, structure)

Current Project Status



- Started in January 2007
- Expected to go live by November 2008

Project Module

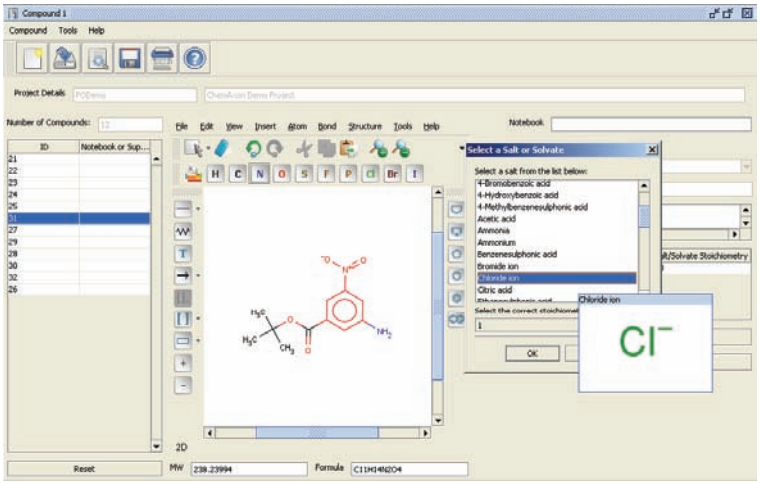


What ChemAxon Tools have we used here?

- Marvin Sketch
- Molecule API to check that a valid molecule has been entered

FUTURE PLANS – Search across all Scaffolds

Compound Window



What ChemAxon Tools have we used here?

- Application Client
  - Marvin Sketch for users who can enter new molecules
  - Marvin View for users who have read-only access to the data
  - Molecule API to check that a valid molecule has been entered
  - MolImporter / MolExporter / MolConverter when adding compounds via an SD File
- Server
  - Standardizer
  - jc_exact to remove salts and solvates (a dictionary is used and can be extended)
  - jc_compare to search compounds within a Project

Standardization

- 2 level approach
  - Compound Standardization (a compound can contain more than one fragment)
    - Removing charges to apply the Evotec Compound Business Rules
    - Removing salts and solvates
    - Standardizing the structures
  - Molecule Standardization (a molecule can only contain one fragment) – Compound Exclusivity is applied here
    - Remove stereochemistry
    - Separate fragments

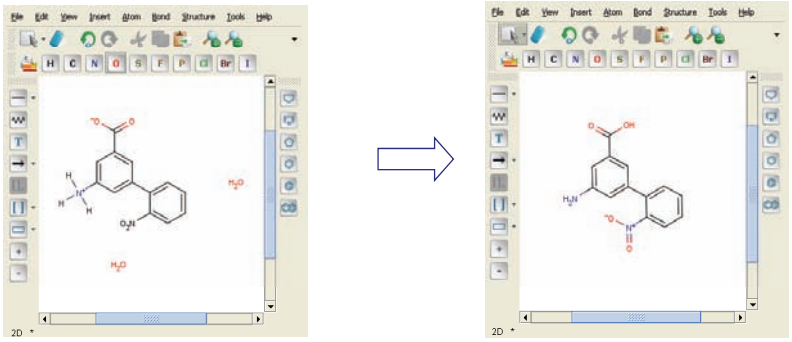
Compound Standardization - Removing Charges

```
-- Attempt nitrogen first.
vl_TransformSMILES := JCF_React('[N;+1:1]([H])>>[N:1]', vl_ReturnSMILES,
      'sep=.. method:a..outFormat:cxsmiles');

-- Then attempt oxygen.
vl_TransformSMILES := JCF_React('[O&-1:1][*&+0:2]>>[O:1][*&+0:2]',
      vl_ReturnSMILES, 'sep=.. method:a..outFormat:cxsmiles');

-- Then attempt Sulfur.
vl_TransformSMILES := JCF_React('[S;-1:1][*&+0:2]>>[S:1][*&+0:2]',
      vl_ReturnSMILES, 'sep=.. method:a..outFormat:cxsmiles');
```


Compound Standardization – Standardizing

```
vl_Reactions   :=  '[O:1][N:2]=O>>[O:1-][N+:2]=O..' ||  '[O:1]=[N:2]=O>>[O-:1][N+:2]=O..'
 ||  '[C][N:1](=C)=[O:2]>>[C][N+:1]([O-:2])=C..' ||  '[C][N:1]([O:2])=C>>[C][N+:1]([O-:2])=C..'
 ||  '[C]:[n:1]([O:2])c>>[C][n+:1]([O-:2])c..' ||  '[c][n:1](=[O:2])c>>[c][n+:1]([O-:2])c..'
 ||  '[c][n:1]([O:2])c>>[c][n+:1]([O-:2])c..' ||  'C[N+:1]([O:2])=C>>C[N+:1]([O-:2])=C..'
 ||  '[C][N:1](C)(C)[O:2]>>[C][N+:1]([C])([C])[O-:2]..' ||  '[C][N:1]([C])([C])=[O:2]>>[C][N+:1]([C])([C])[O-
      :2]..'
 ||  '[O:1]=[N:2]#[C:3]>>[O-:1][N+:2]#[C:3]..'   ||  '[C:1][N:2][N+:3]#[N:4]>>[C:1][N:2]=[N+:3]=[N-:4]..'
 ||  '[O-:1][S+:2][O-:3]>>[O:1]=[S:2]=[O:3]..' ||  '[O-:1][S:2][O-:3]>>[O:1]=[S:2]=[O:3]..'
 ||  '[O:1][S:2][O:3]>>[O:1]=[S:2]=[O:3]..' ||  '[O-:1]-[S+:2]>>[O:1]=[S:2]..'
 ||  '[#6][N:1]=[N+:2]=[N:3]>>[#6][N:1]=[N+:2]=[N-:3]..' ||  '[#6]N[N+]#N>>[#6]N=[N+]=[#7-]..'
 ||  '[N;H1,H2:1][C:2]=[C:3]>>[C:1][C:2]=[N:3]..' ||  '[O;H1:1][C:2]=[C:3]>>[C:1][C:2]=[O:3]';

      vl_StandardSMILES := jcf_Standardize(pi_SMILES, 'sep=!
      config:dearomatize..dehydrogenize..expand..tautomerize..' || vl_Reactions ||
      '..aromatize!cleaningTemplate:select "null" from dual!outFormat:cxsmiles');
```

Standardization Example



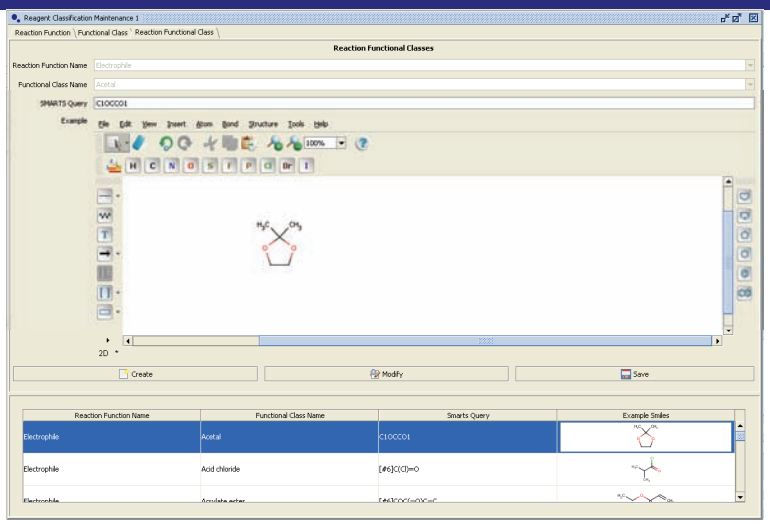
- Charges removed
- Nitro Expanded and Converted
- Water Molecules no longer depicted

Combinatorial Library Enumeration

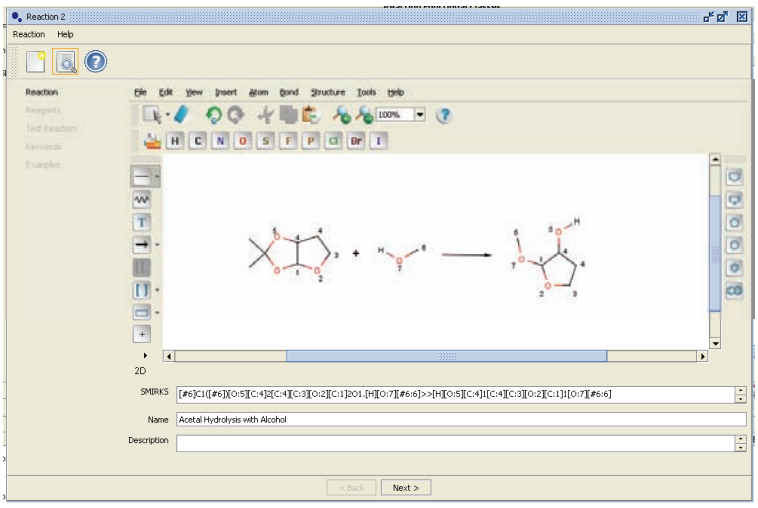
- Reaction based enumeration
- Small reaction database maintained by a team of Chemical Data Scientists

- Each reaction is described by
  - Keyword(s)
  - Example(s)
- Each reagent type is described by
  - Category
  - Functional Class
  - Smarts Query
  - Example Structure

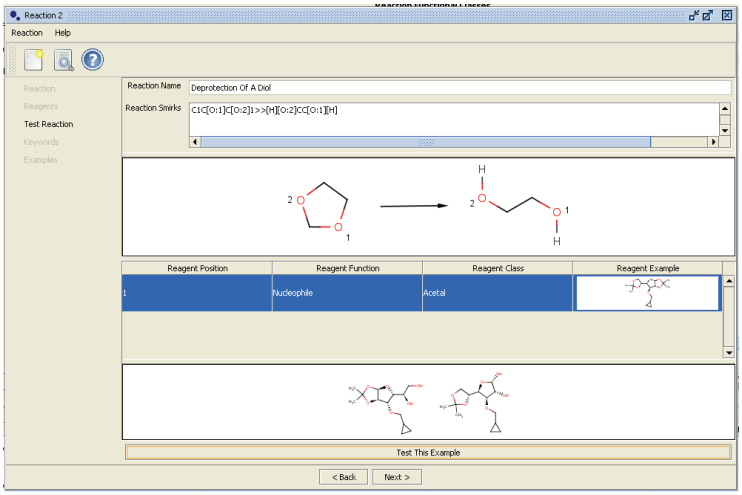
Reagent Type Maintenance Screen



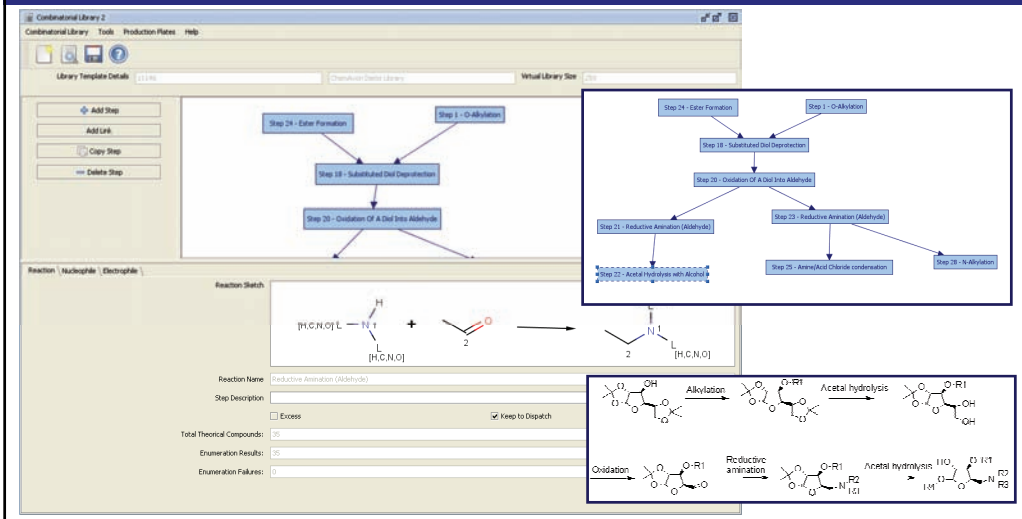
Reaction Maintenance Screen



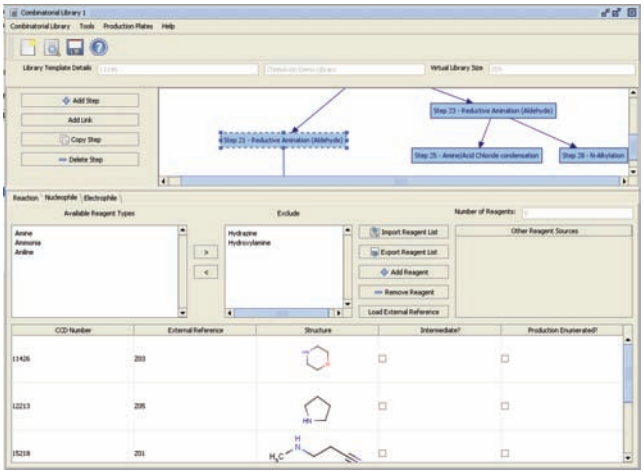
Reaction Maintenance Screen

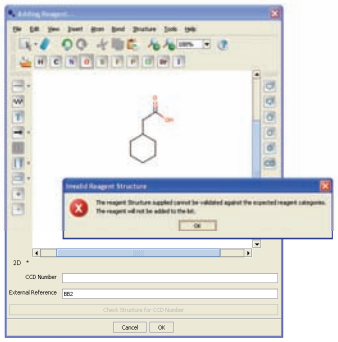


Combinatorial Library Screens



Combinatorial Library Screens

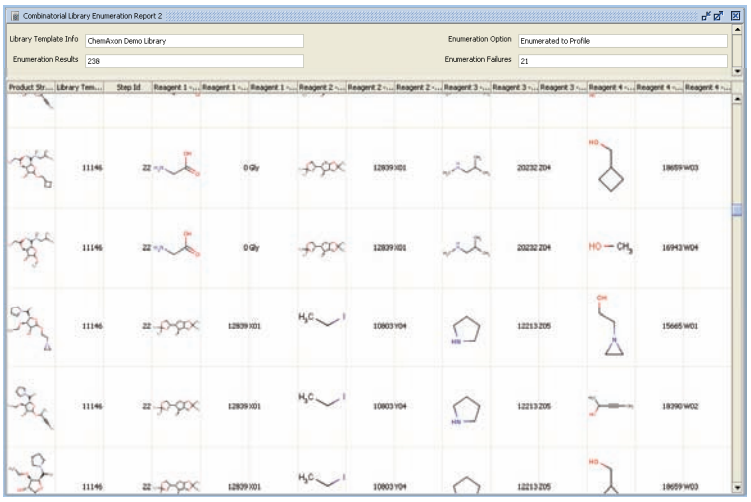




Enumeration Options

- Cross-Section Enumeration
  - Out of 10*10*10 = 1000, we enumerate 1*1*10 + 1*10*1 + 10*1*1 = 30
  - This is used to check that all reagents will react and produce the correct products
- Enumeration for Profiling
  - The full Virtual Library is enumerated here, but the compounds are not available for Exclusivity check
  - This is to prevent some projects from booking a large area of virtual chemical space
- Enumeration for Production
  - The Full Library is enumerated here, but the reagents have been optimised to produce a smaller set of compounds

Enumeration Results



What ChemAxon Tools have we used here?

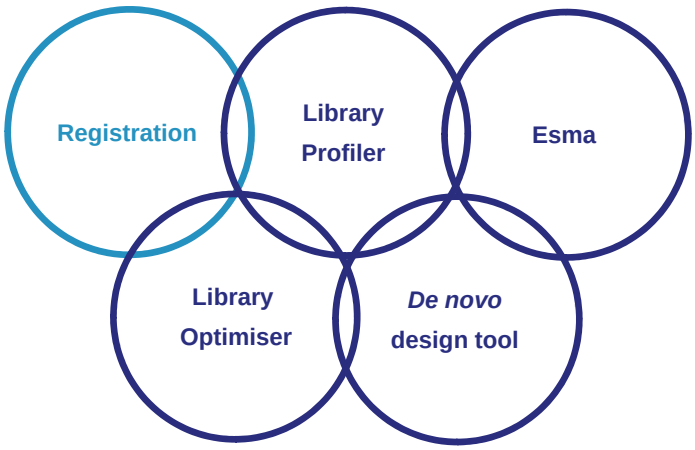
- Client
  - Marvin View / Marvin Sketch
  - Searching to validate the Reagent added against the expected reagent type
  - MolImporter / MolExporter / MolConverted to import and export reagent SD Files
- Server
  - jc_react
  - jc_compare to validate the reaction intermediates against the expected reagent types
  - we have written our own ‘enumeration’ PL/SQL procedure in order to track the reagents used to make each product

Reaction Maintenance Screen

FUTURE PLANS

- Allow the scientists to submit new reactions using the reaction addition wizard
- Allow Reaction and Reaction Example Sub-Structure Searching
- Display the reaction on each Step Label

Future – Integrated Drug Design Application



Thanks

- Alexander Roberts
- Rajeev Halder
- Phil Cochrane
- Alistair Sedwell
- Paul Barrett

- Chemistry User Group – Geoffroy Maillol, Bob Marmon, Mike Potterton, Jane Couzens, Anne Pintat

Evotec Registration System
Catherine Reisser