New Web-based Approaches to Communicating and Teaching in Chemistry & Crystallography

Simon Coles
EPSRC National Crystallography Service
School of Chemistry
University of Southampton
“The internet wasn't created for mockery! It was created so scientists from different universities could share datasets....”


Social Networks for Chemists

Google generation: new behaviour and approach
• Video + Paper = Pubcast
New Approaches to ‘Sharing Experiments’

- Specialised domain-oriented innovations
Formation of Open Communities

- New approaches surfacing and growing FAST
Open Notebook Science

• Immediate sharing of experimental information & data
New Information Exchange Environments

• Immersive alternative to conventional browsing & interaction
The primary cause is the current data publication process, which is tied to journal articles and peer review.
The Solution

Intellect & Interpretation (Journal article, report, etc)

Underlying data (Institutional data repository)
The eCrystals Data Repository

- Quick & simple to deposit
- Software tools
- Laboratory archive
- Community involvement
- ‘Embargo’ facility
- Structured foundations
- Discoverable & harvestable

Available Files

<table>
<thead>
<tr>
<th>Final Result</th>
<th>Available Files</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005sg0007.cif</td>
<td>11k</td>
</tr>
<tr>
<td>2005sg0007.cml</td>
<td>4k</td>
</tr>
<tr>
<td>2005sg0007.bat</td>
<td>120k</td>
</tr>
</tbody>
</table>

Validation

- 2005sg0007_checksum.txt: 9k

Retrieval

- 2005sg0007.res: 9k
- 2005sg0007_vh.txt: 39k

Solution

- 2005sg0007.psd: 6k
- 2005sg0007_ubt.txt: 44k

Processing

- 2005sg0007_hkl: 532k
- 2005sg0007_hms.txt: 114k
- 2005sg0007_nh3.jpg: 95k
- 2005sg0007_nh3.jpg: 97k
- 2005sg0007_nh3.png: 79k

Data Collection

- 2005sg0007_crystal.png: 17k

Other Files

- 2005sg0007_dec.png: 188k
- 2005sg0007_mcm.png: 1k
- 2005sg0007_dif.png: 4k
- 2005sg0007_msd.png: 2k
- 2005sg0007_psd.png: 1k
- 2005sg0007_plipso.png: 27k

http://ecrystals.chem.soton.ac.uk
A Thorough Approach to Dissemination

- Using simple Dublin Core protocol (OAI-PMH)
  - Crystal structure
  - Title (Systematic IUPAC Name)
  - Authors
  - Affiliation
  - Creation Date
- Additional chemical information through Qualified Dublin Core
  - Empirical formula
  - International Chemical Identifier (InChI)
  - Compound Class & Keywords
- Specifies which ‘datasets’ are present in an entry

- Application Profile http://www.ukoln.ac.uk/projects/ebank-uk/schemas/
- DOI links http://dx.doi.org/10.1594/ecrystals.chem.soton.ac.uk/145
- Rights & Citation http://ecrystals.chem.soton.ac.uk/rights.html
Poster (Thurs/Fri): P24.02.04 (C627)
Simple Integration into Teaching

- Component of MChem course...using REAL data

Devise → Search → Discover → Analyse → Manipulate
***Open Science Experiment***

http://ecrystals.chem.soton.ac.uk/544/
http://ecrystals.chem.soton.ac.uk/543/
http://ecrystals.chem.soton.ac.uk/547/

Susanne Huth Talk (Sat): MS.98.3 (C163)
Welcome to the e-malaria Project!

What's the problem?

Malaria kills over a million people a year, and it's spreading. The world needs new drugs to combat the disease. As resistance to existing drugs grows and global warming increases the range of the mosquitoes that need it are increasing.

So what has this got to do with you?

Well, you need to design the drugs! This project asks you to design a small polypeptide, which we run a series of calculations on to determine its possible use as an anti-malarial drug. At this stage your four schools are signed up as pilot centres for the project. That means you are one of only about 100 people taking part in the project. You have a huge chance to change the fate of the project. This project has the potential to get promising compounds to the Phase 1/2a trials and to be a starting point for drug development.

What's it in for you?

Well, we've provided some materials for you connected with the science involved in the project, which should help with revision. There are also opportunities for visits to Southampton and in school talks from project members. But the best thing by far is you could discover a brand new drug. The project has the potential to get promising compounds to the Phase 1/2a trials and to be a starting point for drug development.

Bonding

Dipoles are two equal and opposite partial charges on a molecule. They arise from polar bonds. An opposite charge attracts the dipole and repels the same charge, keeping the molecule apart. A hydrogen bond is a weak dipole-dipole interaction. These are only a few of the possible interactions between a drug molecule and its target. Click here to see a list of others.

The Drug Design Process

Drug design is a long complicated process. It takes approximately 15 years to get a drug from concept to market, and it costs a staggering £500 million pounds. From somewhere in the region of 10,000 candidates only one drug will make it to the shops. So what happens that costs so much, and takes so long?

Step 1

The first thing is to determine and improve on existing environmental conditions for which the treatments are being designed and fix treatment, h

Step 2

The next question is how does a drug get into the body? How does this drug get to its target? The answer is usually that the drug is absorbed into the bloodstream and is transported to the target organ. These are only a few of the possible interactions between a drug molecule and its target. Click here to see a list of others.

Step 3

Next you need to conduct experiments to determine the effectiveness of the drug. This is often done by testing the drug on animal models. If the drug is effective, it will be tested on humans. These are only a few of the possible interactions between a drug molecule and its target. Click here to see a list of others.

Hydrophilicity Test

Check your understanding of hydrophilicity below.

Task: You must sort Polar and Non-Polar.
**Open Science Experiment**
General Chemistry Issues: Data Generation

Characterisation

Synthesis
Shortfalls in Data Management

“Data from experiments conducted as recently as six months ago might be suddenly deemed important, but those researchers may never find those numbers – or if they did might not know what those numbers meant”

“Lost in some research assistant’s computer, the data are often irretrievable or an undecipherable string of digits”

“To vet experiments, correct errors, or find new breakthroughs, scientists desperately need better ways to store and retrieve research data”

“Data from Big Science is … easier to handle, understand and archive. Small Science is horribly heterogeneous and far more vast. In time Small Science will generate 2-3 times more data than Big Science.”

Spectroscopic analysis is often performed to ensure a reaction is proceeding according to plan – as a result <5% are published (via a process with heavy information loss).
A General Chemistry Laboratory Repository

Create new compound (parent record)

Add new experiment type

Add metadata and upload data files
***Open Science Experiment***

http://r4l.eprints.org
A New Kind of Electronic Lab Notebook

Welcome

moreTea is an electronic experiment planning, recording and reviewing tool. It enables you to create experiments, helps you to carry them out, and then lets you view the results and add to your notes.

1. Sign Up and Sign In

Before you can get started using moreTea, you must first create a user profile using the form on the left. You will be sent an e-mail asking you to activate your account. You only have to do this once. If you have already created a profile then just sign in instead.

2. Create and Plan

Once signed in, you can start creating experiments and adding materials and steps to your experiment.

3. Record and Observe

When you've finished planning, it's time to go into the lab and do it for real. The moreTea Lab Tool connects to the server to let you access your experiments in the lab. You can use the Lab Tool to record the amount of each material that you use and make observations by drawing or writing notes.

4. Review and Write-up

Once the experiment is complete you can return to the experiment browser and view the results online. You can also add to the observation notes you made in-lab to expand them as part of your write-up.
A New Kind of Electronic Lab Notebook

Procedure

1. The 1,8-naphthyridine 2,7-disulphonic dibromide was added and the flask was cooled to a magnetic stirrer.
2. Chloroform (250 mL) was added dropwise via a needle and syringe.
3. Following the addition of chloroform, the mixture was added dropwise.
4. The mixture was left to stir overnight and was monitored by TLC.
5. Once the TLC showed the completion of the reaction, it was passed into 35 mL of water.
6. The organic layer was then separated and washed with further water (3 x 50 mL).
7. The organic layer was then separated and washed with 50 mL of ethyl acetate.
8. The solid product was then recrystallized from a DMF-ethanol mixture.
9. The product was characterized.

Instructions

- The organic layer was noted and was washed with further water (3 x 50 mL).
- The solid product was noted and then characterized.
Analysis & Discussion: Blogging Experiments

A repository can…

- Allow one to *put*, *store* and *get*
- Provide *search* and *browse* functionality
- **NOT** provide the *presentation* and *discussion* functions essential to working up a scientific study

- Enables ‘geographically distributed collaborative research’
- Can be open or private
- A useful approach for sharing ‘failed’ experiments?

http://chemtools.chem.soton.ac.uk/projects/blog/
Automatic Blogging by Machines
Automatic Blogging by Sensors

- Continuous log of ‘environmental’ conditions in a laboratory
- Instant detection of erroneous events
- Correlate with inconsistencies in datasets
Comments and Collaborative Tools

- Annotation tools allow comments and foster collaboration and/or communication
- Need for more advanced Blog tools/technology around data
Acetamide project eMalaria docking results

Creator: Simon Coles
Experiment Type: eMalaria docking_job
Total energy = 37.59, 37.01, 41.74 kJ mol⁻¹ for Me, OMe, OEt respectively

Broken down to:
Me = 5.53, 26.87, -4.89
OMe = 5.41, 27.57, -6.31
OEt = 5.76, 31.44, -7.26

for f-bond, Van Der Waals and Strain contributions respectively.

This reflects the greater degree of freedom in the OEt chain, whereas OMe and Me are relatively similar.

Acetylation of amines - general synthetic route
N-(4-methoxyphenyl)acetamide crystal structure

N-(4-methoxyphenyl)acetamide

[Additional content not legible due to image quality]
Open Science Experiment

Experiment Type: Ordered Comparison

1. Acetanilide comparison

- N-acetyl-4-methylphenylacetamide
- N-acetyl-4-methylphenylacetamide (HCl salt)

There are few differences in the IR's of N-acetyl-4-methylphenylacetamide and N-acetyl-4-methylphenylacetamide (HCl salt). In the crystal structures, both contain some water, however this is not present in the N-acetyl-4-methylphenylacetamide (HCl salt) for Me. Ome & OE.

Crystal structure evidence suggests that the N-acetyl-4-methylphenylacetamide (HCl salt) and N-acetyl-4-methylphenylacetamide have similar packing motifs, however the molecule structure of the N-acetyl-4-methylphenylacetamide (HCl salt) shows a higher degree of hydrogen bonding and spatial separation compared to the N-acetyl-4-methylphenylacetamide.

- Summary:
  - Melting points agree with high stage behaviour in any of the N-acetyl-4-methylphenylacetamide (HCl salt) for Me. Ome & OE.
  - X-ray analysis is required for spatial separation.
Packaging, Describing and Sharing

- New moves in Digital Libraries community to enable distributed repositories to exchange content
- OAI-ORE (Open Archives Initiative – Object Reuse and Exchange)
  - http://www.openarchives.org/ore/
- Describes an aggregation of objects in an exchangeable format
- Microsoft funded eChemistry testbed project
Towards a New Model for Chemical Information Exchange
‘Sharing Experiments’

http://www.myexperiment.org/
A solid foundation for Open/Self-Publishing and Communication of Chemistry Data???

Thanks to:

• Mike Hursthouse, Jeremy Frey, Cameron Neylon, Andrew Milsted, Richard Stephenson, Jamie Robinson, Steven Wilson, Andrew Bailey, Mark Borkum

• Dave DeRoure, Les Carr, monica schraefel, Chris Gutteridge, Tim Myles-Board, Arouna Woukei, Dave Tarrant, Stuart Middleton

• Liz Lyon, Rachel Heery, Monica Duke, Michael Day, Traugott Koch, Manjula Patel, Pete Cliff