UNIVERSITY OF Southampton
School of Chemistry

The data explosion

…and the need to manage diverse data sources in scientific research

Simon Coles (s.j.coles@soton.ac.uk)

Director, UK National Crystallography Service

Why manage?

UNIVERSITY OF Southampton
School of Chemistry

- Volume
  - Day to day coping
  - How to "publish"??
- Scientific Responsibility
  - Accurate recording of the whole experiment
  - Archival & curation
  - Enabling future science
- Diversity
  - Heterogeneous data
  - Disciplinary differences
  - Institutional boundaries
- Accountability
  - Auditing and charging
  - Funder mandates

What can structures be used for?

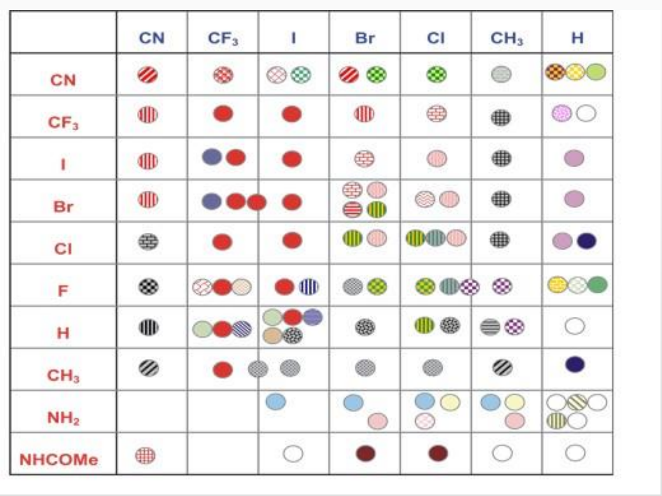
University of Southampton
School of Chemistry

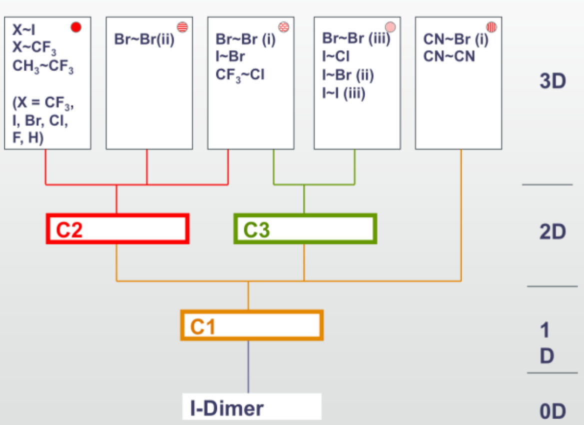
- For example…
  - Materials science – link to properties, starting point for calculations
  - Insight into / control of crystallisation & crystallisability
  - Starting point for computational chemistry
- Data management is key, because its all about context and linking…
- Good crystallographic practice requires information about the experiment not covered wholly by CIF
- Not possible to predict how a crystal structure might be used in the future – measure everything you possibly can!

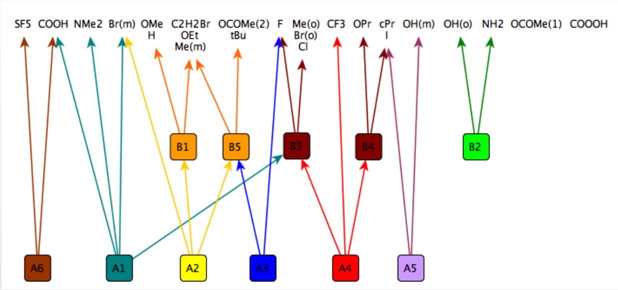
Increasing complexity

UNIVERSITY OF Southampton
School of Chemistry

- Even for pure crystallographic studies volumes of data are getting bigger
- ‘Informatics’ approaches being adopted in order to develop fundamental understanding, rules etc
- Must be able to reliably evaluate quality (provenance) to incorporate results of others

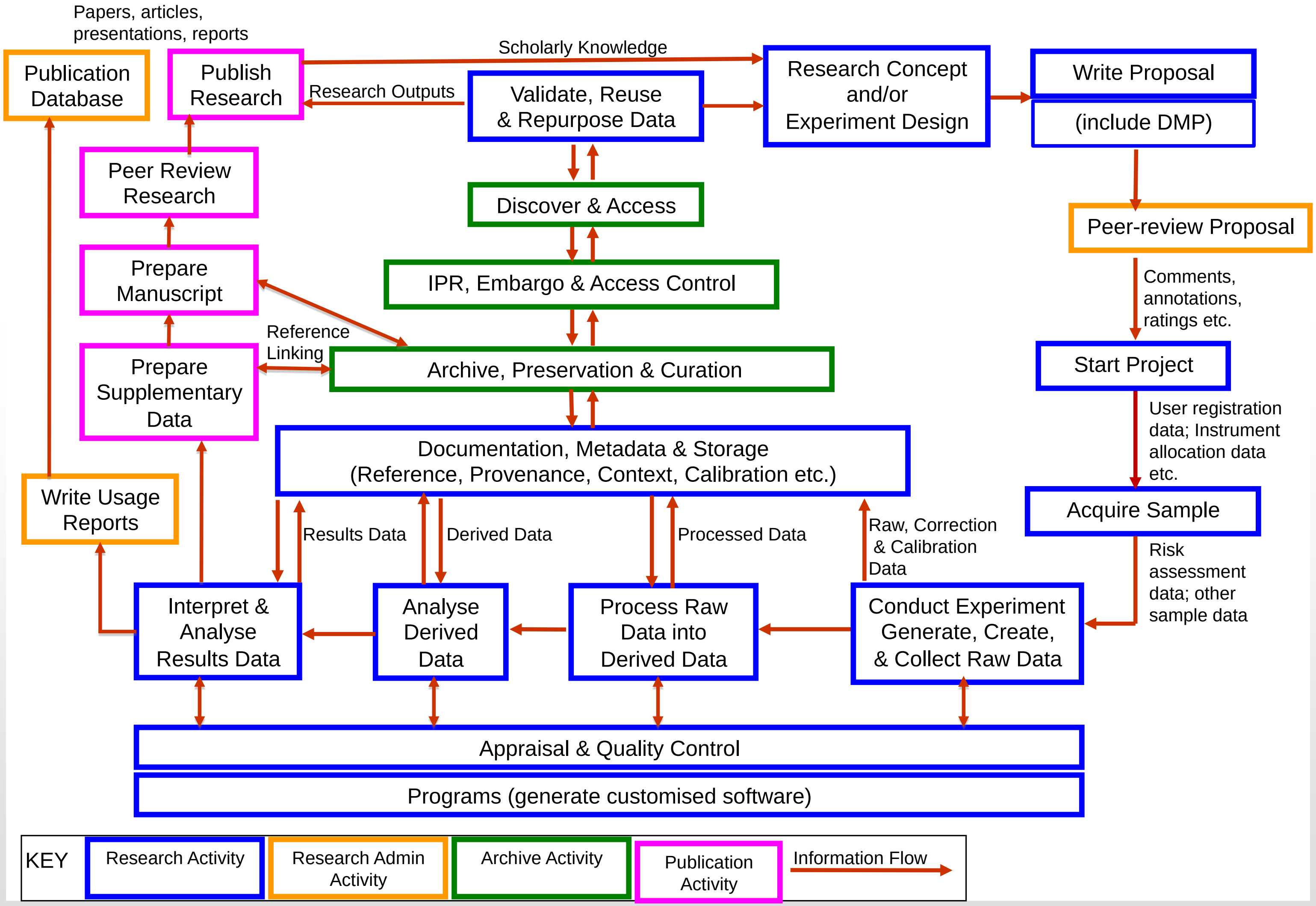






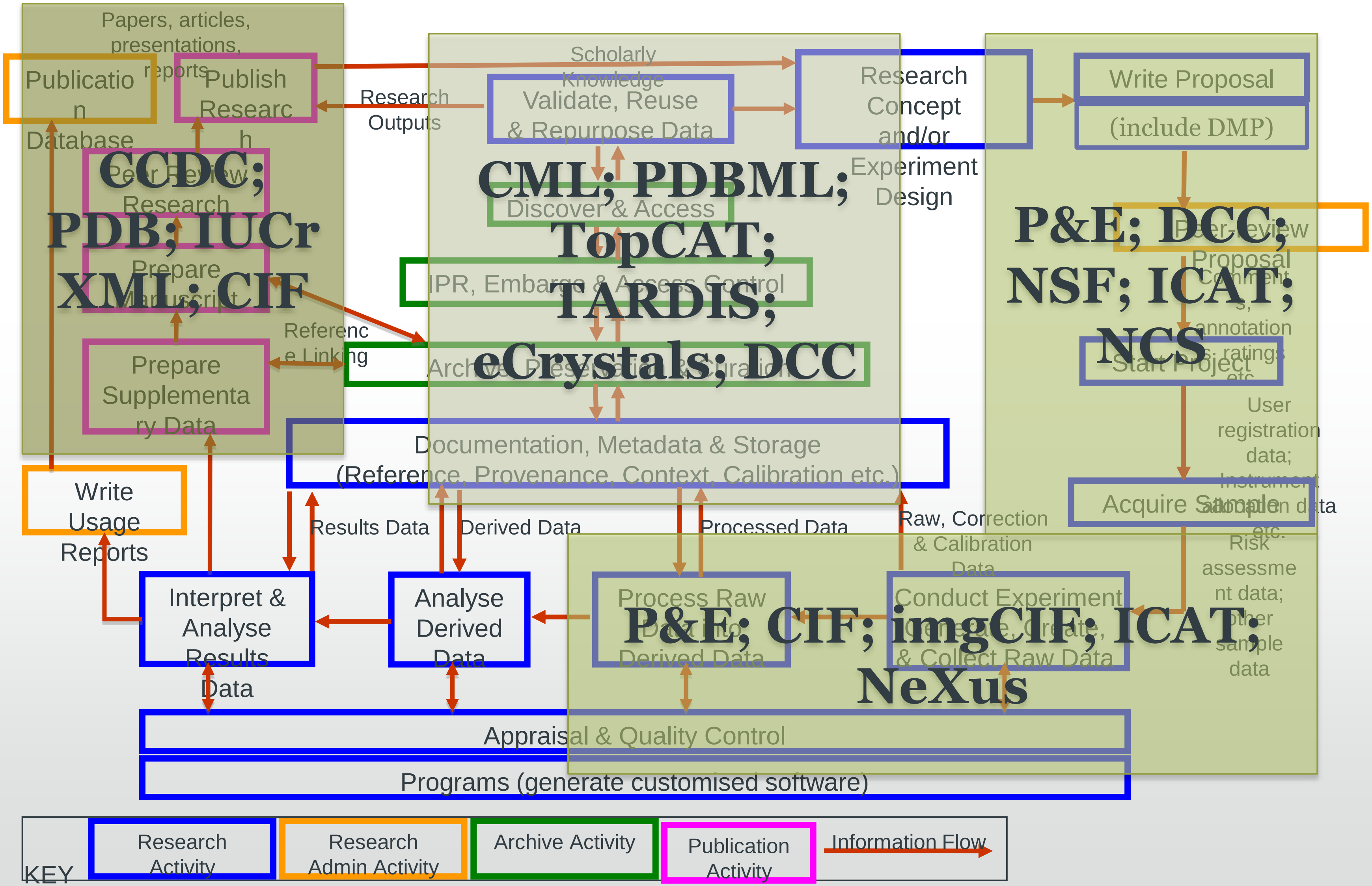
The whole research process

UNIVERSITY OF Southampton
School of Chemistry



How many standards?!

UNIVERSITY OF Southampton
School of Chemistry



Data management (eco)systems

University of Southampton
School of Chemistry

• Publication Repository	[Results Data]
• LIMS	[User & Sample Data]
• ELN	[Supporting Experiment Data]
• Datastores	[Underlying Diffraction Data]
• Desktop/Laptop	[Everything!]
• iPad/iPhone	[Anything!]

- How to integrate these??

UNIVERSITY OF Southampton
School of Chemistry

Structured Data.

i.e. using LIMS, Datastores & Repositories

LIMS: Lifecycle Management

UNIVERSITY OF Southampton
School of Chemistry

ncs UK National Crystallography Service

Current User Log out
Service Account (info@ncs.ac.uk)

Application for an Allocation

Period: 1st May 2013 - 31st October 2013

Personal Details

Name: Dr Mike Coogan
Email Address: m.coogan@lancaster.ac.uk
Department School of Chemistry
Institution: Lancaster University
Address: B6 Faraday Building
Lancaster University
Lancaster
LA1 4YB
United Kingdom

Funding

Funding Source: Not Funded

Local Facilities

Local Facilities: No

Reason(s) why additional facilities are required:

No local facilities, this is a new department just starting up and while we intend to buy a diffractometer, this will not be until next year at the earliest.

Last Allocation

	FSA Full Structure	DSO Data Collection Only
Your last allocation was:	5	5
Your usage over the last period was:	0	2

Could you indicate the percentage of these:

Have not/cannot be worked up
Have been fully refined
Have been written up into a report or thesis
Have been published (please give full details of paper(s) in the next section)

Request for this Allocation	5	5

Outline of Scientific Program

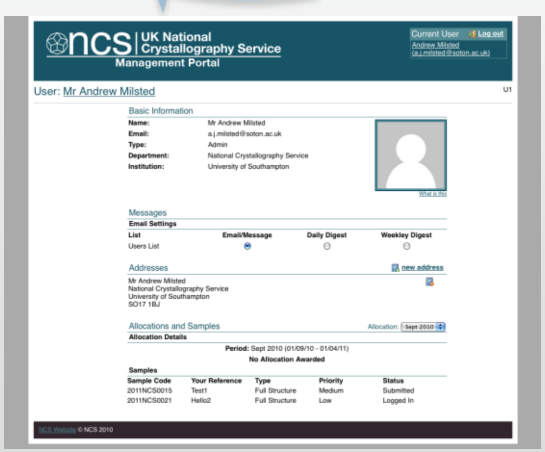
We have developed a range of complexes based around luminescent transition metal fragments which are useful cell imaging applications. Some of these are macromolecular structures which can act as hosts for smaller molecules or ions and the luminescence is in many cases modulated upon encapsulation. As well as expanding the range of metals which we can use in cell imaging to the early TMs (Zr, Hf) as well as the more traditional late TMs we are looking at new host macromolecules: hetero-cycle appended calixarenes which form interesting hydrogen bonded structures as well as complexes with a range of transition metals; also complexes based around expanded structures based on polypyridyls in dimeric, trimeric and larger assemblies. All these complexes show interesting photophysical properties, e.g. acting as sensors, acting as imaging agents. The balance between crystalinity and solubility / lipophilicity (essential for crossing cell membranes) is difficult to build in or predict (especially in the macromolecules) with such macromolecules and in many cases it is difficult to obtain high quality single crystals of these species, and thus as well as access for data collections only, we have requested full structure solution for difficult cases where the expertise of the service will be essential.

Publications

Additional Information

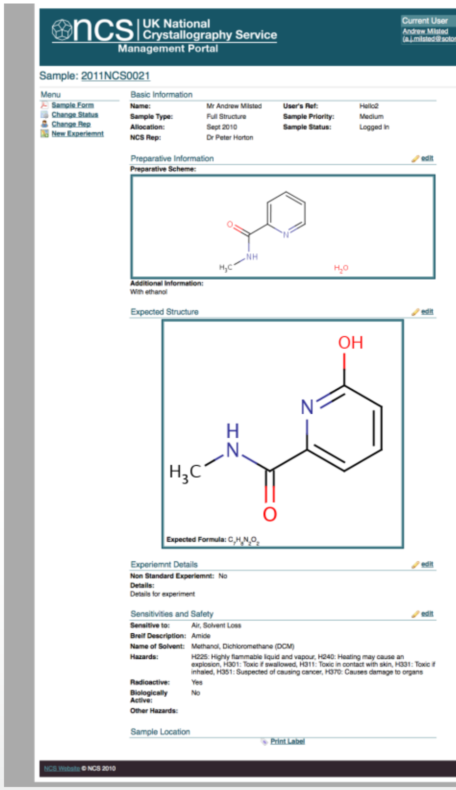
Please accept my apologies for the lack of samples outputs for the last period- this is a result of the development of the new chemistry department at Lancaster having been more complex that originally anticipated, so no lab-work has been possible until this week as major refurbishments were undertaken to bring the facilities to standard.

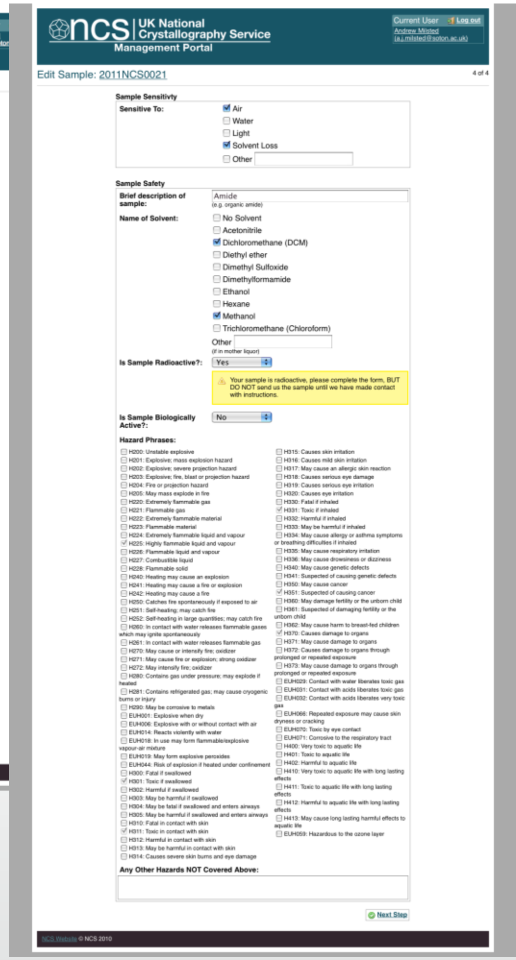
Attached Files

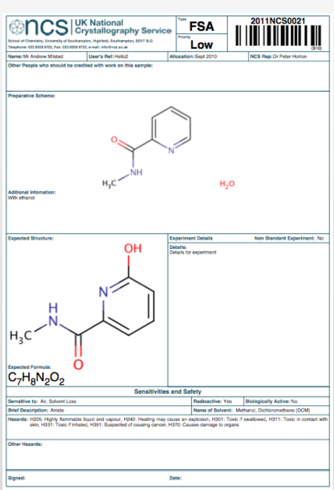


LIMS: Lifecycle Management

UNIVERSITY OF Southampton
School of Chemistry

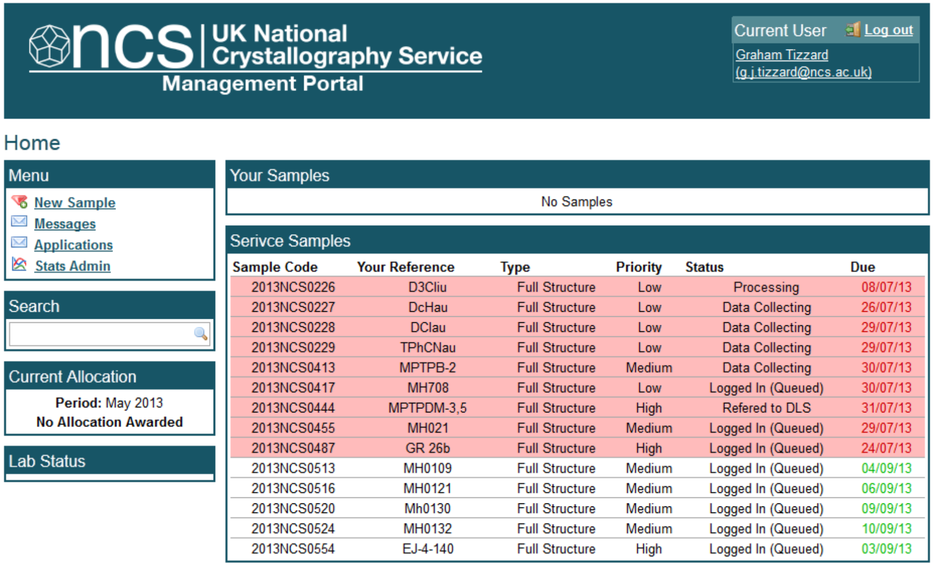






LIMS: Lifecycle Management

UNIVERSITY OF Southampton
School of Chemistry



ncs UK National Crystallography Service
Management Portal

Current User Log out
Graham Tizzard
(g.j.tizzard@ncs.ac.uk)

Home

Menu
- New Sample
- Messages
- Applications
- Stats Admin

Search

Current Allocation
Period: May 2013
No Allocation Awarded

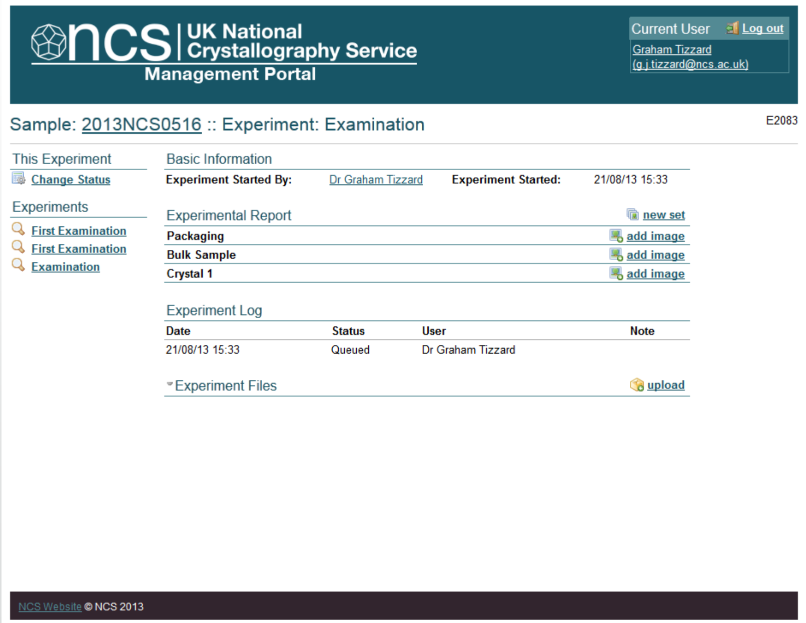
Lab Status

Your Samples

No Samples

Serivce Samples

Sample Code	Your Reference	Type	Priority	Status	Due
2013NCS0226	D3Cliu	Full Structure	Low	Processing	08/07/13
2013NCS0227	DcHau	Full Structure	Low	Data Collecting	26/07/13
2013NCS0228	DClau	Full Structure	Low	Data Collecting	29/07/13
2013NCS0229	TPhCNau	Full Structure	Low	Data Collecting	29/07/13
2013NCS0413	MPTPB-2	Full Structure	Medium	Data Collecting	30/07/13
2013NCS0417	MH708	Full Structure	Low	Logged In (Queued)	30/07/13
2013NCS0444	MPTPDM-3,5	Full Structure	High	Refered to DLS	31/07/13
2013NCS0455	MH021	Full Structure	Medium	Logged In (Queued)	29/07/13
2013NCS0487	GR 26b	Full Structure	High	Logged In (Queued)	24/07/13
2013NCS0513	MH0109	Full Structure	Medium	Logged In (Queued)	04/09/13
2013NCS0516	MH0121	Full Structure	Medium	Logged In (Queued)	06/09/13
2013NCS0520	Mh0130	Full Structure	Medium	Logged In (Queued)	09/09/13
2013NCS0524	MH0132	Full Structure	Medium	Logged In (Queued)	10/09/13
2013NCS0554	EJ-4-140	Full Structure	High	Logged In (Queued)	03/09/13



ncs UK National Crystallography Service
Management Portal

Current User Log out
Graham Tizzard
(g.j.tizzard@ncs.ac.uk)

Sample: 2013NCS0516 :: Experiment: Examination

E2083

This Experiment
- Change Status

Experiments
- First Examination
- First Examination
- Examination

Basic Information

Experiment Started By: Dr Graham Tizzard **Experiment Started:** 21/08/13 15:33

Experimental Report — new set

Packaging	add image
Bulk Sample	add image
Crystal 1	add image

Experiment Log

Date	Status	User	Note
21/08/13 15:33	Queued	Dr Graham Tizzard	

Experiment Files — upload

NCS Website © NCS 2013

LIMS: Lifecycle Management

UNIVERSITY OF Southampton
School of Chemistry

UK National Crystallography Service
Management Portal

Current User Log out
Graham Tizzard
(g.j.tizzard@ncs.ac.uk)

Sample: 2012NCS0714

S1338

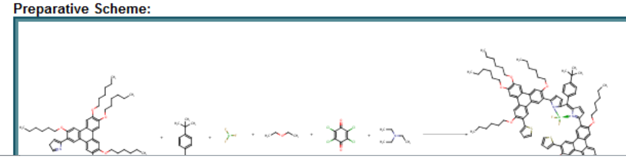
Menu
- Sample Form
- Change Status
- Change Rep
- New Sample (User)

Basic Information edit

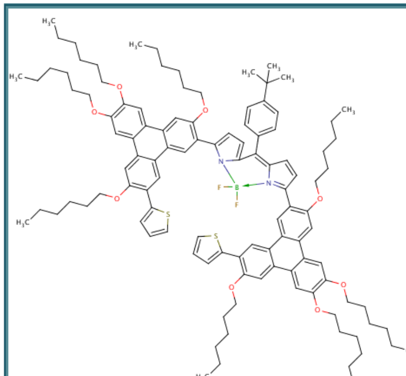
Name:	Prof Andy Cammidge	**User's Ref:**	HG815 BODIPY
Sample Type:	Full Structure	**Sample Priority:**	Reserved
Allocation:	May 2012	**Sample Status:**	Completed
NCS Rep:	Dr Graham Tizzard	**Next Step:**	First Examination

Preparative Information edit

Preparative Scheme:



Expected Structure edit



Expected Formula: $C_{111}H_{139}BF_2N_2O_8S_2$

Experiment Details edit

Non Standard Experiemnt: No
Details:

Sensitivities and Safety edit

Sensitive to:

Breif Description:	Triphenylene-BODIPY Hybrid
Name of Solvent:	Methanol, THF
Hazards:	H225: Highly flammable liquid and vapour, H301: Toxic if swallowed, H311: Toxic in contact with skin, H331: Toxic if inhaled, H370: Causes damage to organs
Radioactive:	No
Biologically Active:	No
Other Hazards:	

Sample History

Date	Status	User	Note
24/08/12 10:18	Draft	Miss Qiying Huang	
24/08/12 10:18	Submitted	Miss Qiying Huang	
24/08/12 10:18	Logged In (Queued)	Miss Qiying Huang	
05/12/12 10:35	Refered to DLS	Dr Graham Tizzard	v low res data only collected at Soton, try DLS
06/06/13 15:51	Completed	Dr Graham Tizzard	

Sample Stats

Stat	Days	Quota
Logged in - 1st Exam	59	over
Exam - Complete	130	over

Sample Location

Print Label

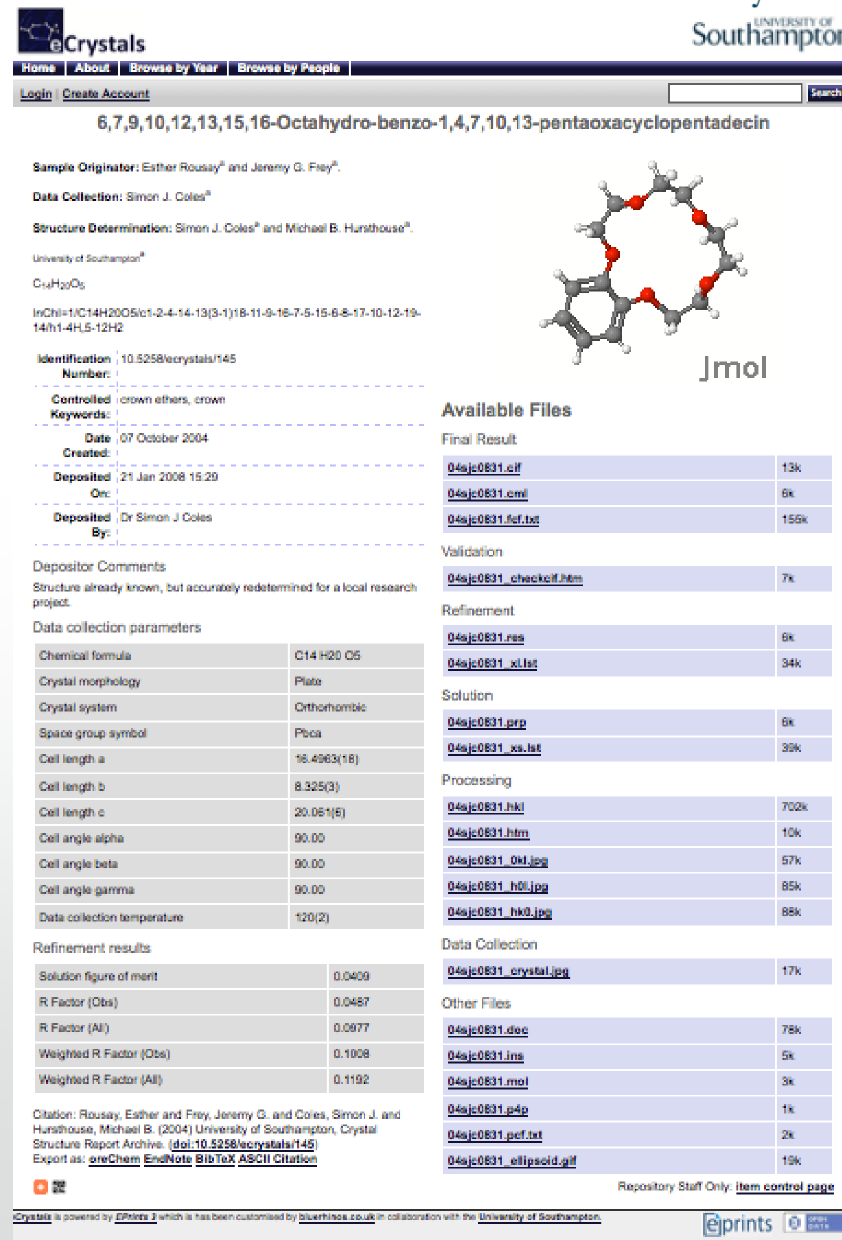
Experiments add

First Examination	Queued Since: 24/08/12 10:18
Examination	Queued Since: 21/08/13 16:07

NCS Website © NCS 2013

LIMS: Lifecycle Management

UNIVERSITY OF Southampton
School of Chemistry



eCrystals
Southampton

Home | About | Browse by Year | Browse by People

Login | Create Account

Search

6,7,9,10,12,13,15,16-Octahydro-benzo-1,4,7,10,13-pentaoxacyclopentadecin

Sample Originator: Esther Rousay[a] and Jeremy G. Frey[a].

Data Collection: Simon J. Coles[a]

Structure Determination: Simon J. Coles[a] and Michael B. Hursthouse[a].

University of Southampton[a]

$C_{14}H_{20}O_5$

InChI=1/C14H20O5/c1-2-4-14-13(3-1)18-11-9-16-7-5-15-6-8-17-10-12-19-14/h1-4H,5-12H2

Identification Number:	10.5258/ecrystals/145
Controlled Keywords:	crown ethers, crown
Date Created:	07 October 2004
Deposited On:	21 Jan 2008 15:29
Deposited By:	Dr Simon J Coles

Depositor Comments

Structure already known, but accurately redetermined for a local research project.

Data collection parameters

Chemical formula	C14 H20 O5
Crystal morphology	Plate
Crystal system	Orthorhombic
Space group symbol	Pbca
Cell length a	16.4963(18)
Cell length b	8.325(3)
Cell length c	20.061(6)
Cell angle alpha	90.00
Cell angle beta	90.00
Cell angle gamma	90.00
Data collection temperature	120(2)

Refinement results

Solution figure of merit	0.0409
R Factor (Obs)	0.0487
R Factor (All)	0.0977
Weighted R Factor (Obs)	0.1008
Weighted R Factor (All)	0.1192

Citation: Rousay, Esther and Frey, Jeremy G. and Coles, Simon J. and Hursthouse, Michael B. (2004) University of Southampton, Crystal Structure Report Archive. (doi:10.5258/ecrystals/145)
Export as: oreChem EndNote BibTeX ASCII Citation

Jmol

Available Files

Final Result

04sjc0831.cif	13k
04sjc0831.cml	6k
04sjc0831.fcf.txt	155k

Validation

04sjc0831_checkcif.htm	7k

Refinement

04sjc0831.res	6k
04sjc0831_xl.lst	34k

Solution

04sjc0831.prp	6k
04sjc0831_xs.lst	39k

Processing

04sjc0831.hkl	702k
04sjc0831.htm	10k
04sjc0831_0kl.jpg	57k
04sjc0831_h0l.jpg	85k
04sjc0831_hk0.jpg	88k

Data Collection

04sjc0831_crystal.jpg	17k

Other Files

04sjc0831.doc	78k
04sjc0831.ins	5k
04sjc0831.mol	3k
04sjc0831.p4p	1k
04sjc0831.pcf.txt	2k
04sjc0831_ellipsoid.gif	19k

Repository Staff Only: item control page

eCrystals is powered by EPrints 3 which is has been customised by bluerhinos.co.uk in collaboration with the University of Southampton.

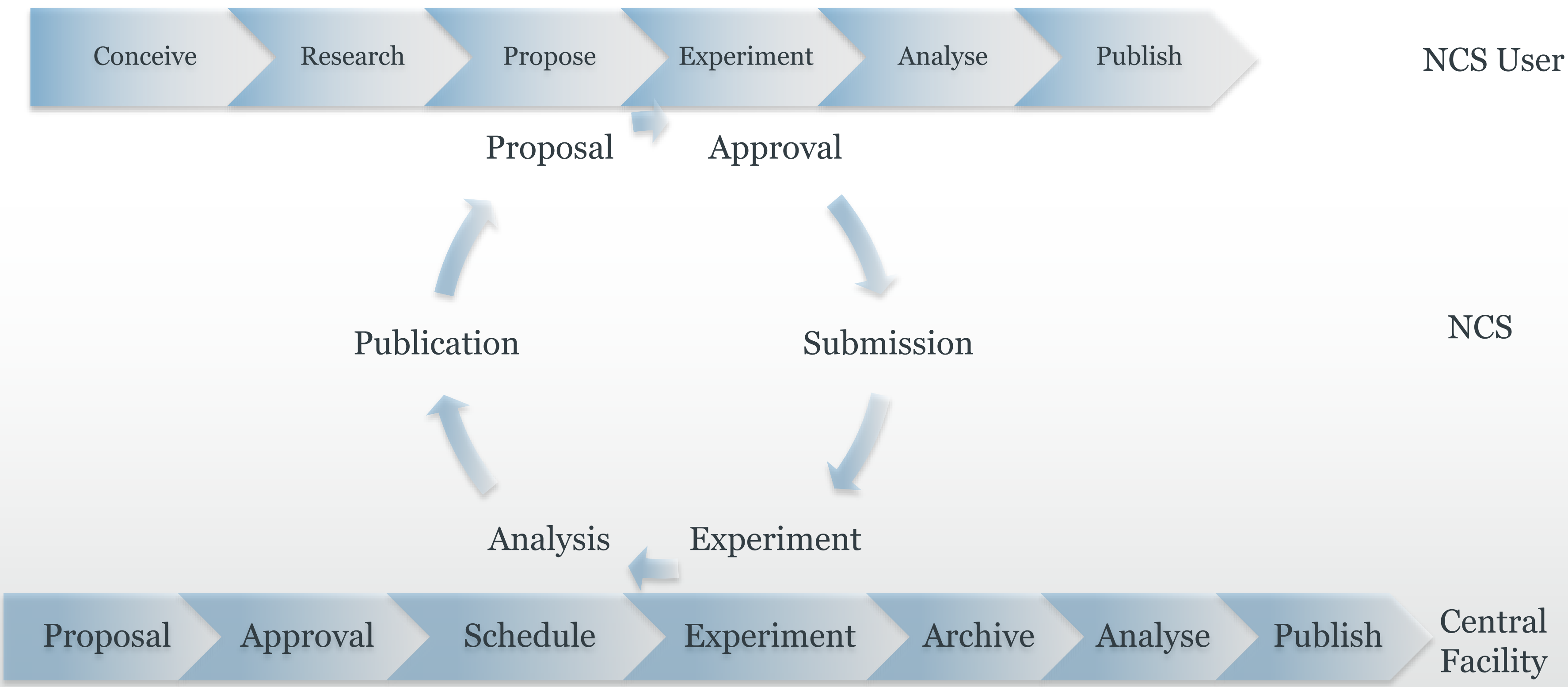
eprints

http://ecrystals.chem.soton.ac.uk

Management Across Boundaries

UNIVERSITY OF Southampton
School of Chemistry

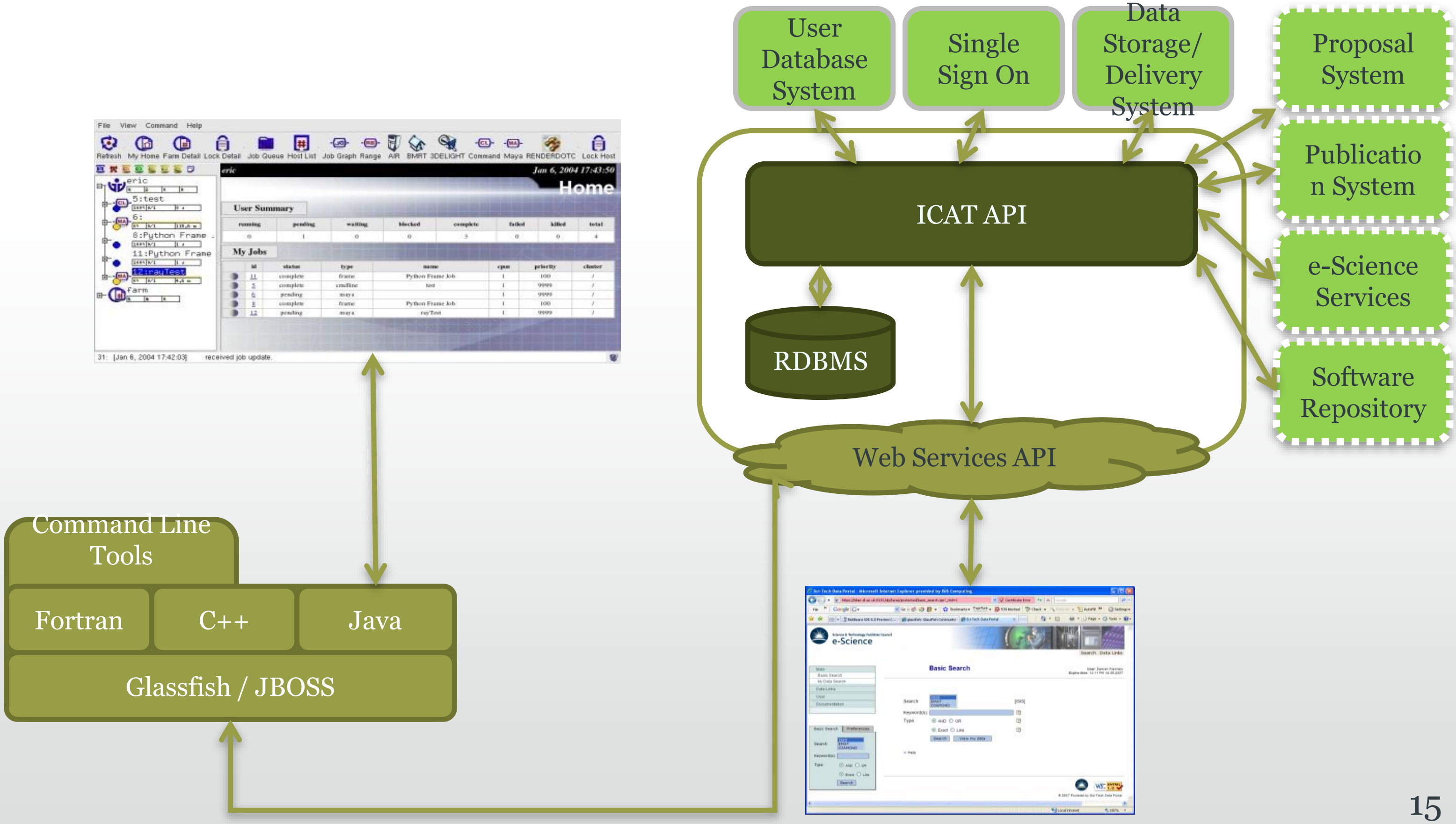
- Management across facilities (ICAT Information Model)



Datastores, ICAT & CSMD

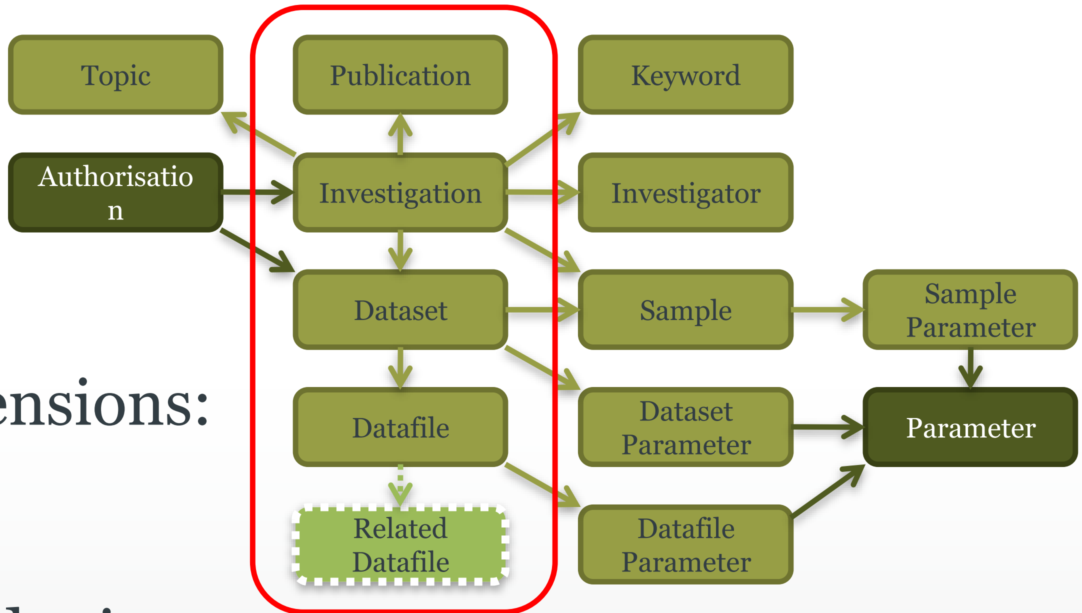
UNIVERSITY OF Southampton
School of Chemistry

The Core Scientific Metadata model forms the information model for ICAT & is designed to describe facilities-based experiments



CSDM as a Starting Point

UNIVERSITY OF Southampton
School of Chemistry

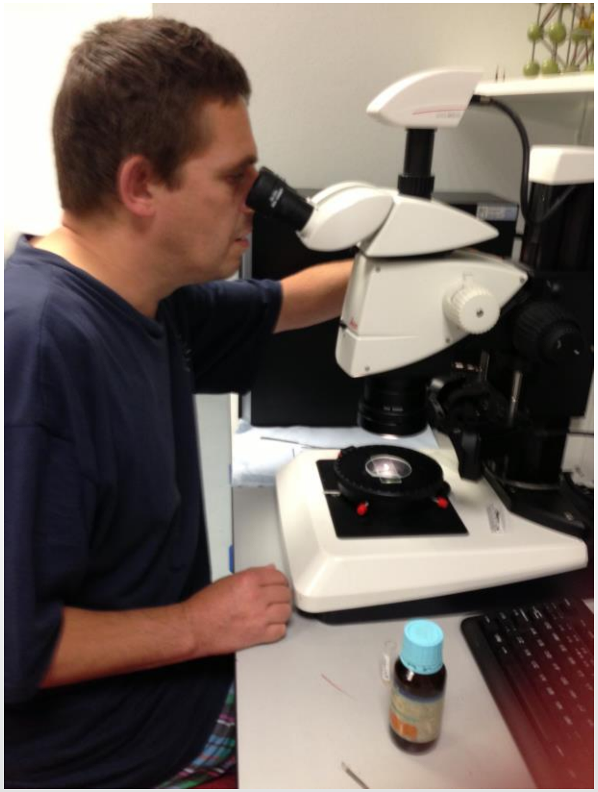


- Forms the basis for extensions:
  - To derived data
  - To laboratory based science
  - To secondary analysis data
  - To preservation information
  - To publication data

Getting Mobile in the Lab

University of Southampton
School of Chemistry

- But every crystallographic experiment has 'unstructured' data



Getting Mobile in the Lab

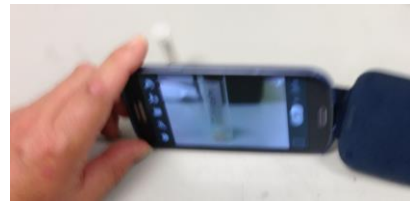
UNIVERSITY OF Southampton
School of Chemistry

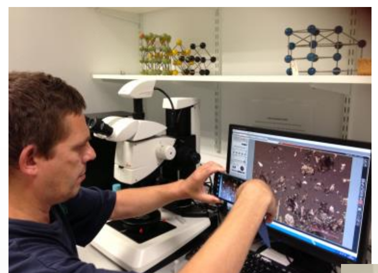
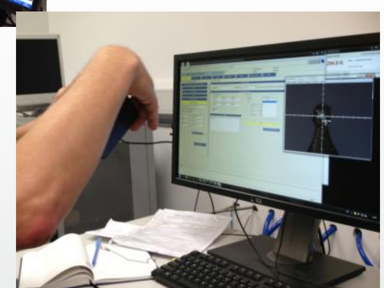
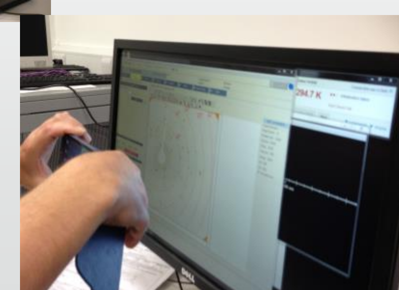
Package & Sample Tube

Bulk Sample & Manipulation

Mounted Sample

Trial Diffraction Pattern

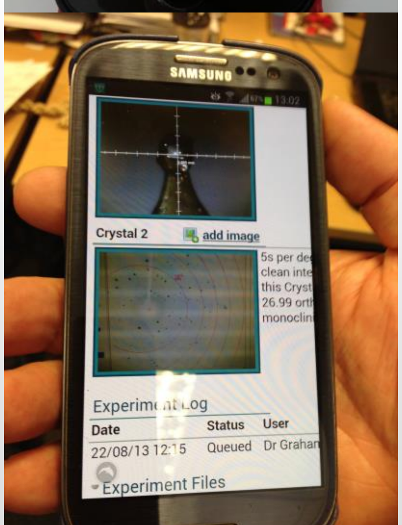



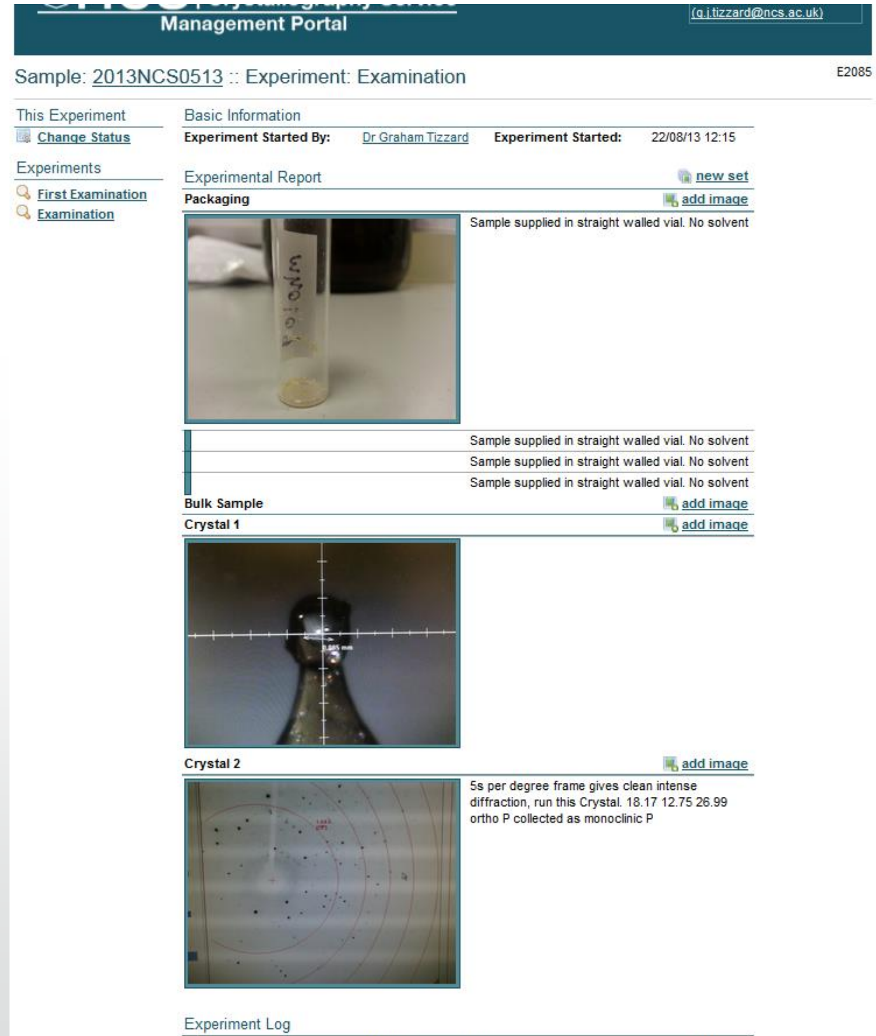





Getting Mobile in the Lab

UNIVERSITY OF Southampton
School of Chemistry

The completed record…





UNIVERSITY OF Southampton
School of Chemistry

Unstructured data.

Invariably the context for a crystal structure

Or

The context a crystal structure has to fit into!

Laboratory Notebooks

UNIVERSITY OF Southampton
School of Chemistry

- All this is not a new problem…

Phys. Educ. **26** (1991). Printed in the UK

Faraday's notebooks: the active organization of creative science

Ryan D Tweney

Faraday's notebooks constitute one of the largest and most revealing archives left to us by a major scientist. These records reveal a good deal of systematic invention and exploration of recording techniques by Faraday, work that reveals much about his thinking about science, as well as of the role of memory in general in scientific thinking.

Scientists are students—students of nature, to be sure, but, like all students, dependent for their success on the taking of notes. In even the most routine of scientific research, scientists must preserve external records of their work. Most externalize far more than just data, making records of their hypotheses, readings of the literature, wild speculations and the like. Thus, scientific diaries, laboratory notebooks, indeed the entire range of recording techniques, constitute an important topic for a full understanding of just what scientists do (Holmes 1987).

Michael Faraday has left us a richer documentary legacy of thought than exists for perhaps any other scientific figure in history. Faraday's daily laboratory notebooks, diaries and commonplace books, almost all of which were carefully bound by Faraday himself, and almost all of which are held in the Archives of the Royal Institution of Great Britain and the Institution of Electrical Engineers are a rich source for the historian. As a rough guess, he left us records of about 30 000 experiments, both successful and unsuccessful, as well as a large number of speculative idea books, bibliographies, indexes, scrap-books, etc etc. What is known as 'the' Diary has even been published (Martin 1932–6). Though only a part of the archival holdings, this work does cover most of his famous discoveries (and lots of lesser ones as well). But it is very much a laboratory diary, and thus gives a somewhat misleading picture of the whole. This paper will present a broader perspective, one that sheds light on the entire range of Faraday's records. In this way, we will be able to gain some insights into the uniquely creative mind of a genius.

Why are Faraday's records so extensive? In part, it is because Faraday was mistrustful of his own memory (see Williams 1965, pp 473, 491–501; Hare 1974). Faraday more than once repeated an experiment that he had earlier completed and apparently forgotten about, and his use of elaborate memory-retrieval devices (see below) makes a similar point.

Memory weaknesses aside, however, Faraday was very much part of a cultural tradition of ideas deriving from John Locke (1632–1704) that placed central importance on memory as an essential cognitive process in the acquisition of knowledge. In the Lockean theory of ideas, knowledge is built upon *stored* sensations and ideas: imagination, reason, and such things as Descartes' 'clear and distinct ideas' are built on the foundation of memory. Locke himself saw that this account had strong implications for the importance of memory aids. If memory was to be the foundation of knowledge, then the *weaknesses* of human memory—forgetting, distortion, the vagaries of the retrieval process—constituted serious problems. True knowledge needed accurate memories, and this required the use of accurate records.

Faraday derived his Lockean view of memory from Isaac Watts (1674–1748) whose *Improvement of the Mind* (1809/1741) was read by Faraday in 1809 and credited by him with 'having taught me to think'. It inspired Faraday's first surviving memory aid, the *Philosophical Miscellany* (Faraday 1809–10). His use of memory aids evolved subsequently during the course of his career, culminating, after 1831, in the mature recording and retrieval system that is described in the next section. It is worth noting that one important respect in which Faraday was *not* Lockean was in his reliance on the power of experiment. Neither

Ryan D Tweney is Professor of Psychology at Bowling Green State University, Ohio, USA. His research in cognitive psychology has focused on the understanding of scientific thinking, with a special interest in the working methods of Faraday.

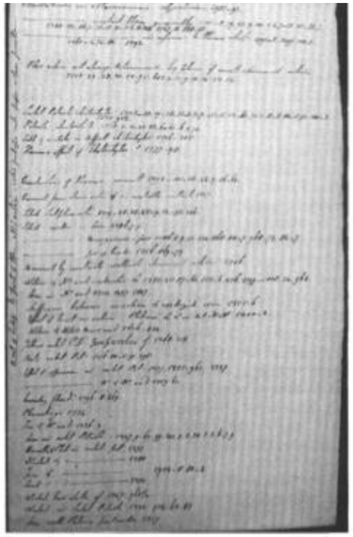


Figure 2. An example of a retrieval sheet (IEE Archives, Misc. Mss. SC2).

own Diary entries or on his speculative thinking. His use of the retrieval aids was a dynamic one.

Looking at the retrieval sheets once suggested a kind of experiment to me. I took one of the sheets and photocopied the relevant sections referred to in the Diary, pasted *these* up on long sheets, and tried to read the result, to see if I could get insights into what Faraday was after in making the retrieval sheet. Unfortunately, the result made little sense, even when it dealt with a part of the Diary that I was fairly familiar with. The reason is clear: Faraday didn't just need references to the particular facts recorded in the Diary; he needed *cues* to the entire context of his memories about the incidents in question. In fact, when I abandoned the long paste-ups that I had made and simply read the Diary in the order in which references were made, the whole made much more sense. My eye could pick up closely associated references in adjacent parts of the Diary, and these were helpful in reconstructing the context. How much richer this would have been for Faraday, for whom these adjacent 'reminders' would have been full of the concrete memories, images, emotions and so on of the original incident.

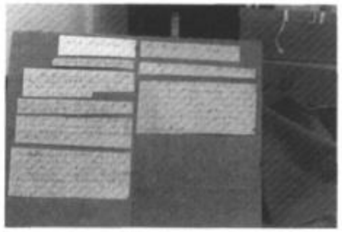


Figure 3. A 'paste-up' (IEE Archives, Misc. Mss. SC2).

Among the many types of retrieval sheets that survive, some stand out because they don't refer directly to the Diary itself. Instead, they seem to be indexes of indexes, or 'menus' to use a modern computerese term. These are striking because they suggest that Faraday used so many retrieval devices that he needed to organize these as well. Anyone who has lost a file on a hard disk will appreciate the need!

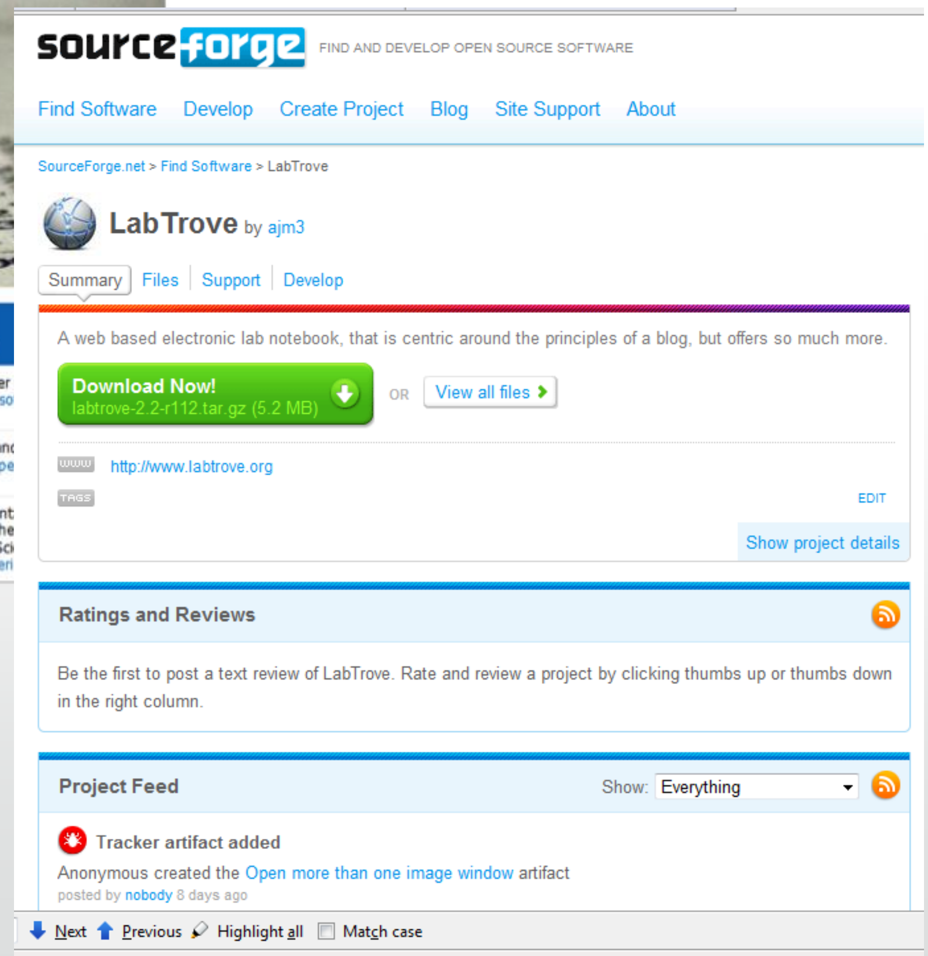
Constructing retrieval aids of the sort described above can be seen as analogous to an encoding process in which a retrieval cue (an index tag, say) is encoded in a physical form with an address to the full diary entry. Once the tag exists, it can be used in one of two ways. First, obviously, it can serve as a finding aid for the retrieval of specific diary entries. Second, and perhaps even more importantly, it can serve as a mnemonic cue useful in the structuring of diary-based knowledge. By sorting slips one can impose one or another organization on diary entries, and vary that organization in the service of other goals.

Development of the memory aids

So far, the discussion has focused on Faraday's use of retrieval aids in their mature form, i.e. after 1832. By looking at earlier records, however, it is possible to trace the evolution of Faraday's use of external memory aids. Faraday's earliest surviving notebook is the *Philosophical Miscellany* kept in 1809–10, which was used mainly to copy out interesting things Faraday had read. It contains an ordinary alphabetical index evidently prepared after the book was completed; while such an index is useful, note that it can only be prepared *after* all

Electronic Lab Notebooks

UNIVERSITY OF Southampton
School of Chemistry



http://www.labtrove.org/

LabTrove

UNIVERSITY OF Southampton
School of Chemistry

Login | All Blogs | Help | Support | About

ourexperiment alpha

Synthesis of 4-substituted methylidene oxindoles

Project E-Lab Notebook for the synthesis of five 4-substituted methylidene oxindole from oxindole and their corresponding aromatic aldehydes.

Older Posts >>

Single Crystal X-ray crystallography

27th June 2012 @ 23:09

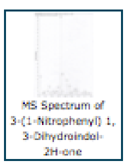
Single-crystal X-ray diffraction analyses were performed using a Bruker APEXII CCD diffractometer mounted at the window of a Bruker FR591 rotating anode (MoKα = 0.71073 Å) and equipped with an Oxford Cryosystems cryostream device. Data were processed using the Collect package and unit cell parameters were refined against all data. An empirical absorption correction was carried out using SADABS . The structures were solved by direct methods using SHELXS-97 and refined on Fo^2 by full-matrix least-squares refinements using SHELXL-97. All non-hydrogen atoms were refined with anisotropic displacement parameters. All hydrogen atoms were added at calculated positions and refined using a riding model with isotropic displacement parameters based on the equivalent isotropic displacement parameter (Ueq) of the parent atom. Figures were produced using OLEX2

Graham Tizzard | View Source | Analytical Procedures | Comments (0)

MS Spectrum of (3E)-3-(4-Nitrobenzylidene)-1,3-dihydro-2H-indol-2-one

6th May 2012 @ 18:16

Spectroscopic Method: MS-ESI
Substituent: Nitro
MS Spectrum of (3E)-3-(4-Nitrobenzylidene)-1,3-dihydro-2H-indol-2-one:



The mass spectrum of (3E)-3-(4-Nitrobenzylidene)-1,3-dihydro-2H-indol-2-one has been obtained by negative electrospray ionization (ESI). The peak at m/e = 265.2 confirms the molecular mass of this compound as the molecular ion gains a proton.

Interpretation of MS Spectrum of (3E)-3-(4-Nitrobenzylidene)-1,3-dihydro-2H-indol-2-one:

Peak Position	Diff. between molecular mass and peak	Suspected molecules or ions	Inference
265.2	265 – 264 = 1	(M+H)+	Parent compound plus a proton

Nicola Knight | View Source | Spectroscopic Data | Comments (0)

Search

Archives

Sections

Substituent

Spectroscopic Method

Tools

Show/Hide Keys

Login | All Blogs | Help | Support | About

ourexperiment alpha

Pictet-Spengler route to Praziquantel

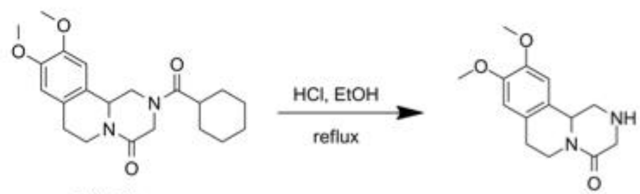
Synthesis of intermediates and derivatives of PZQ

Older Posts >>

Hydrolysis of MNR11-18 to give MNR26-7

23rd April 2013 @ 01:09

Mnr: 21-30
As for
Hydrolysis of MNR11-16 to MNR26-4
Hydrolysis of MNR11-17 to MNR26-5
Starting material from
Synthesis of MNR11-18



MNR11 MNR26

Compound	FW (g/mol)	Density (g/mL)	mmol	Mass (g)	Volume (mL)	eq	conc. (mmol/mL)
MNR11-18	372.46		10.63	3.96		1.00	
1M HCl		0.97	106.32		110	10.00	
EtOH					21		0.50
MNR26-7	262.30						

Hazard and Risk Assessment:

Procedure:
MNR11-8 (3.96 g, 10.63 mmol) was dissolved in EtOH (21 mL) and HCl (1M) (110 mL) and heated to reflux for 2.5 hours. The solution was allowed to cool to room temperature then cooled in an ice bath, basified with NaOH pellets (approx 8 g, pellets used to minimise volume of aqueous material) to pH 12-13 and extracted with DCM (4 x 100 mL). The organic fractions were combined, dried over magnesium sulphate, filtered and concentrated under reduced pressure to give an orange crispy solid (1.008 g).

TLC
Reaction mixture after 2.5 hours ran in 100% EtOAc

Search

Archives

Sections

Mnr

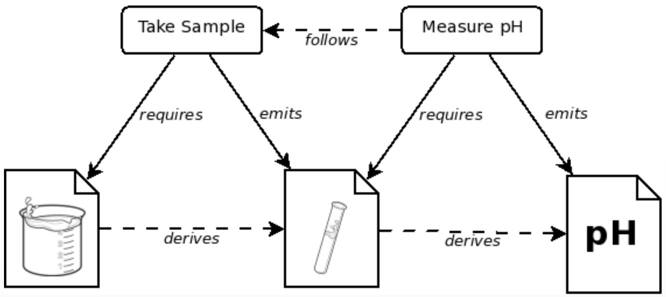
Sc

Tools

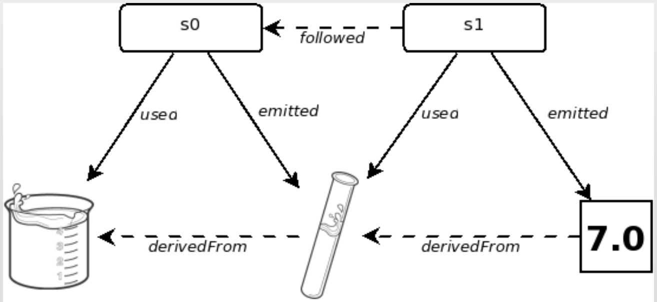
Show/Hide Keys

Imposing structure – Planning & Enactment Ontology

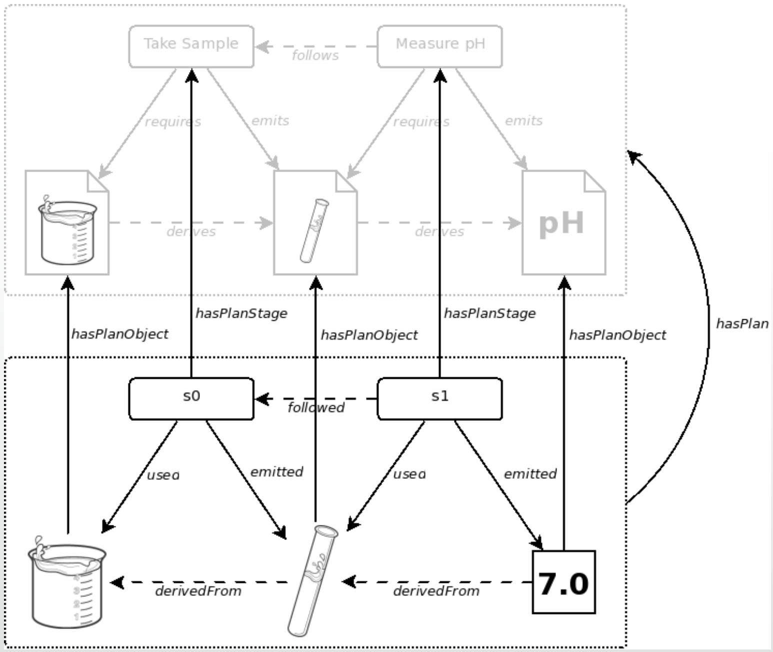
- Plan (Prospective provenance)



- Enactment (Retrospective provenance)



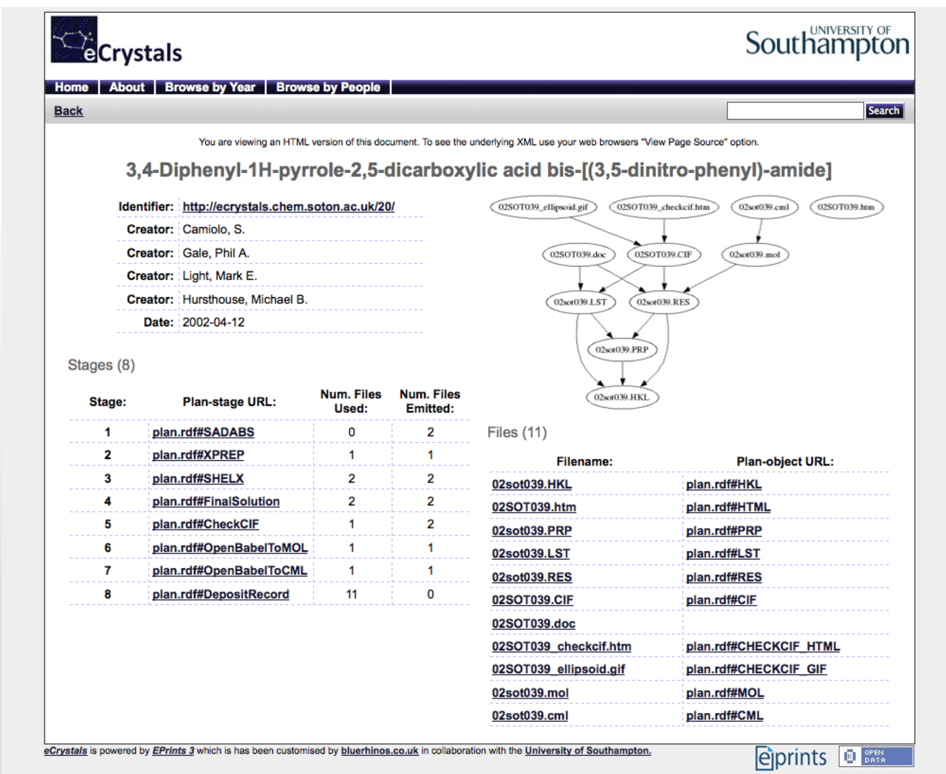
- Realisation



oreChem Plan for eCrystals

UNIVERSITY OF Southampton
School of Chemistry

- Machine-readable representation of methodology
- Describes requirements for software and data products


eCrystals

UNIVERSITY OF Southampton

Home | About | Browse by Year | Browse by People

Back

Search

You are viewing an HTML version of this document. To see the underlying XML use your web browsers "View Page Source" option.

3,4-Diphenyl-1H-pyrrole-2,5-dicarboxylic acid bis-[(3,5-dinitro-phenyl)-amide]

Identifier:	http://ecrystals.chem.soton.ac.uk/20/
Creator:	Camiolo, S.
Creator:	Gale, Phil A.
Creator:	Light, Mark E.
Creator:	Hursthouse, Michael B.
Date:	2002-04-12

Stages (8)

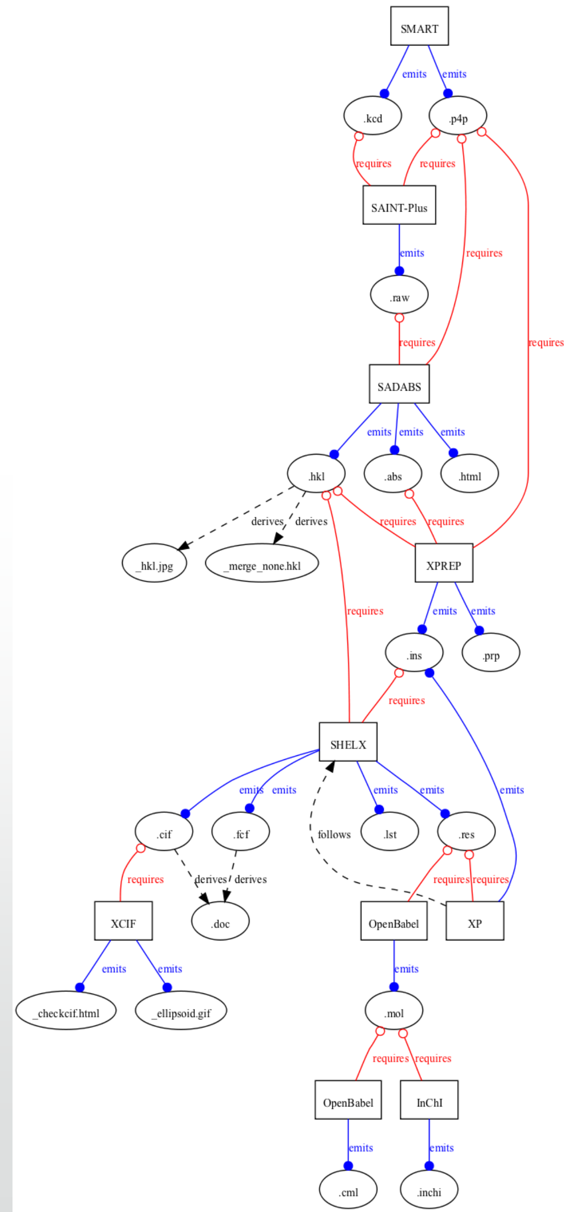
Stage:	Plan-stage URL:	Num. Files Used:	Num. Files Emitted:
1	plan.rdf#SADABS	0	2
2	plan.rdf#XPREP	1	1
3	plan.rdf#SHELX	2	2
4	plan.rdf#FinalSolution	2	2
5	plan.rdf#CheckCIF	1	2
6	plan.rdf#OpenBabelToMOL	1	1
7	plan.rdf#OpenBabelToCML	1	1
8	plan.rdf#DepositRecord	11	0

Files (11)

Filename:	Plan-object URL:
02sot039.HKL	plan.rdf#HKL
02SOT039.htm	plan.rdf#HTML
02sot039.PRP	plan.rdf#PRP
02sot039.LST	plan.rdf#LST
02sot039.RES	plan.rdf#RES
02SOT039.CIF	plan.rdf#CIF
02SOT039.doc	
02SOT039_checkcif.htm	plan.rdf#CHECKCIF_HTML
02SOT039_ellipsoid.gif	plan.rdf#CHECKCIF_GIF
02sot039.mol	plan.rdf#MOL
02sot039.cml	plan.rdf#CML

eCrystals is powered by EPrints 3 which is has been customised by bluerhinos.co.uk in collaboration with the University of Southampton.

eprints OPEN DATA

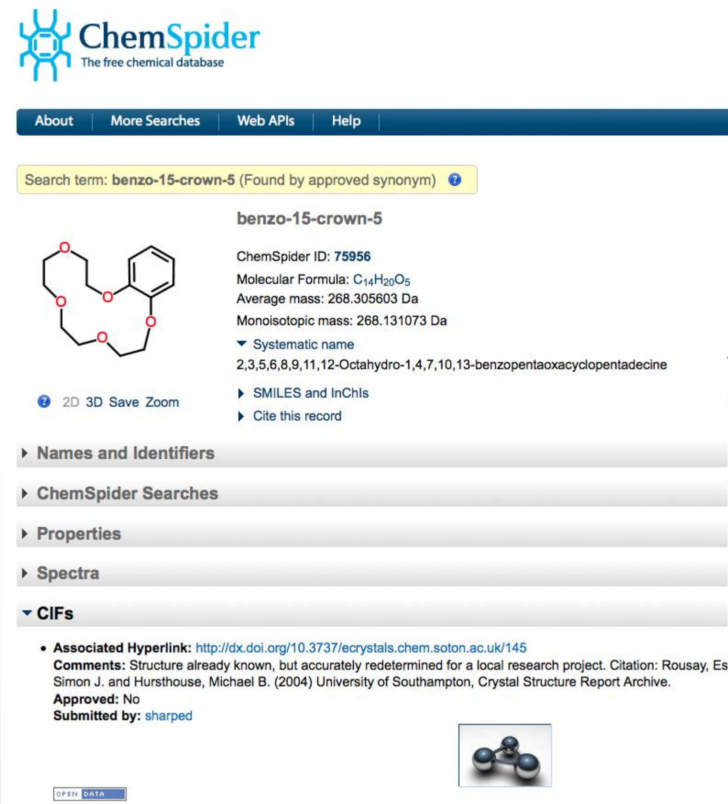


UNIVERSITY OF Southampton
School of Chemistry

Bringing it all together.

Structures in context of compound, property, etc data

University of Southampton
School of Chemistry



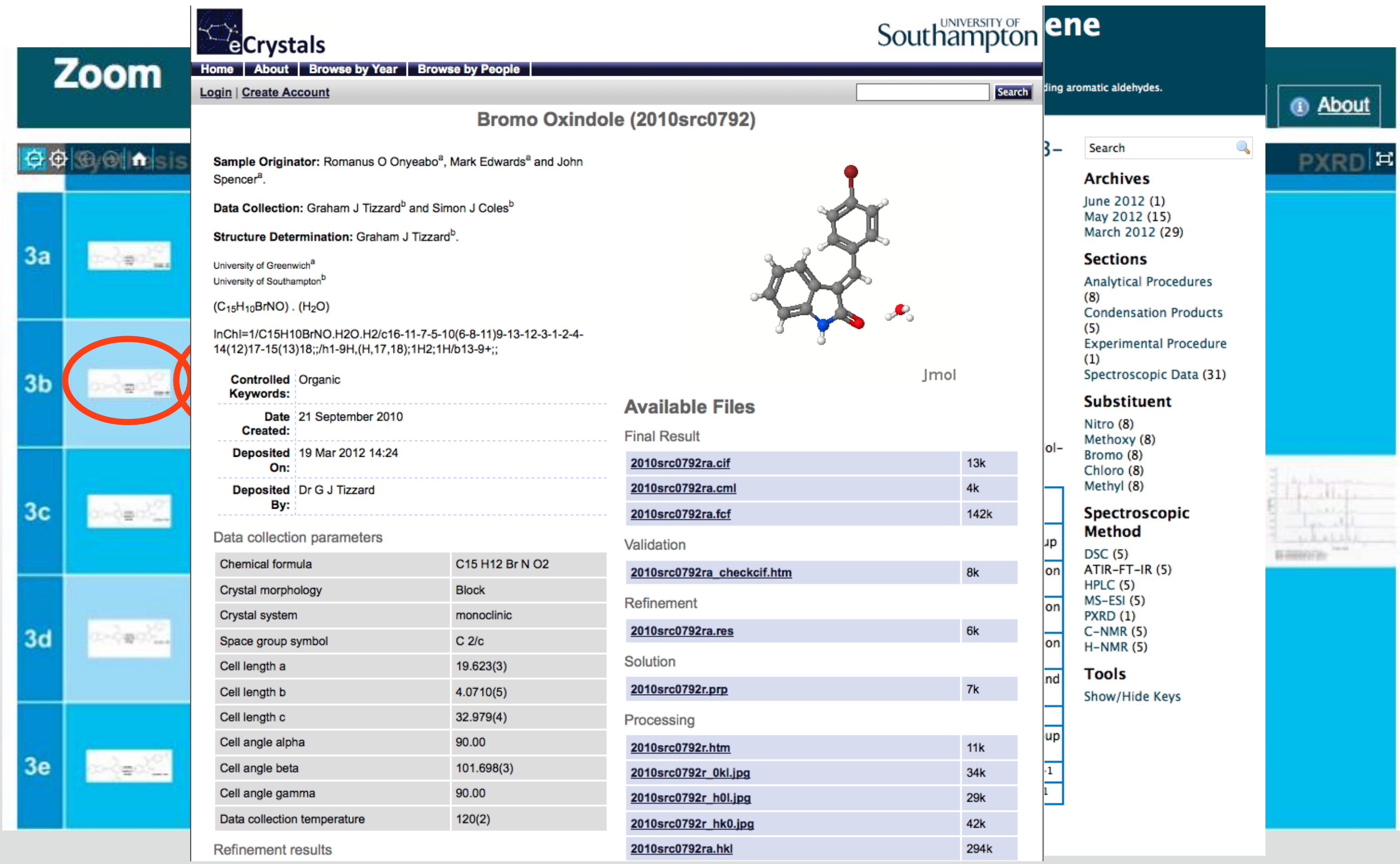
```
- <rdf:RDF>
 - <rdf:Description rdf:about="http://www.chemspider.com/Chemical-Structure.75956.rdf">
    <dc:format>application/rdf+xml</dc:format>
    <dc:rights>http://www.chemspider.com/Disclaimer.aspx</dc:rights>
    <dcterms:creator rdf:parseType="Literal">ChemSpider</dcterms:creator>
    <dcterms:modified rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2011-04-20</dcterms:modified>
  - <foaf:primaryTopic>
   - <ore:ResourceMap rdf:about="http://www.chemspider.com/Chemical-Structure.75956.rdf#ResourceMap">
    - <ore:describes>
     - <ore:Aggregation rdf:about="http://www.chemspider.com/Chemical-Structure.75956.rdf#Aggregation">
      - <ore:aggregates>
       - <rdf:Description rdf:about="http://www.chemspider.com/Chemical-Structure.75956.html">
          <dc:format>text/html</dc:format>
         </rdf:Description>
        </ore:aggregates>
      - <ore:aggregates>
       - <rdf:Description rdf:about="http://www.chemspider.com/ImagesHandler.ashx?id=75956">
          <foaf:depicts rdf:resource="http://www.chemspider.com/Chemical-Structure.75956.rdf#Compound" />
          <dc:format>image/png</dc:format>
         </rdf:Description>
        </ore:aggregates>
      - <ore:aggregates>
       - <rdf:Description rdf:about="http://www.chemspider.com/Chemical-Structure.75956.rdf">
        - <foaf:primaryTopic>
         - <chemdomain:NamedChemicalSpecies rdf:about="http://www.chemspider.com/Chemical-Structure.75956.rdf#Compound">
          - <chemdomain:hasPart>
```


RDF (ChemAxiom)

Structures in the context of 'traditional' publishing

UNIVERSITY OF Southampton
School of Chemistry

- ELN providing Electronic Supplementary Information for conventional publication (Chemistry Central: accepted)



DOI's for Data

UNIVERSITY OF Southampton
School of Chemistry

- DataCite Institutional registration of DOI's
- Promotes rigorous 'self publishing' of data
- Soton exemplars
- eCrystals (Repository/structured data

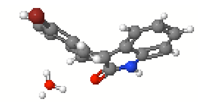
Citation: Onyeabo, Romanus O and Edwards, Mark and Spencer, John and Tizzard, Graham J and Coles, Simon J (2010) University of Southampton, Crystal Structure Report Archive. (**doi:10.5258/ecrystals /1505**)
Export as: **oreChem** **EndNote** **BibTeX** **ASCII Citation**

Structure Determination: Graham J Tizzard[b].

University of Greenwich[a]
University of Southampton[b]

$(C_{15}H_{10}BrNO) . (H_2O)$

InChI=1/C15H10BrNO.H2O.H2/c16-11-7-5-10(6-8-11)9-13-12-3-1-2-4-14(12)17-15(13)18;;/h1-9H,(H,17,18);1H2;1H/b13-9+;;



Identification Number:	10.5258/ecrystals/1505
Controlled Keywords:	Organic
Date Created:	21 September 2010
Deposited On:	28 Mar 2012 15:40
Deposited By:	Dr G J Tizzard

Data collection parameters

Chemical formula	C15 H12 Br N O2
Crystal morphology	Block
Crystal system	monoclinic
Space group symbol	C 2/c
Cell length a	19.623(3)
Cell length b	4.0710(5)
Cell length c	32.979(4)
Cell angle alpha	90.00
Cell angle beta	101.698(3)
Cell angle gamma	90.00
Data collection temperature	120(2)

Refinement results

Solution figure of merit	
R Factor (Obs)	0.0792
R Factor (All)	0.1204
Weighted R Factor (Obs)	0.1462
Weighted R Factor (All)	0.1688

Citation: Onyeabo, Romanus O and Edwards, Mark and Spencer, John and Tizzard, Graham J and Coles, Simon J (2010) University of Southampton, Crystal Structure Report Archive. (**doi:10.5258/ecrystals /1505**)
Export as: **oreChem** **EndNote** **BibTeX** **ASCII Citation**

Available Files

Final Result

2010src0792ra.cif	13k
2010src0792ra.cml	4k
2010src0792ra.fcf	142k

Validation

2010src0792ra_checkcif.htm	8k

Refinement

2010src0792ra.res	6k

Solution

2010src0792r.prp	7k

Processing

2010src0792r.htm	11k
2010src0792r_0kl.jpg	34k
2010src0792r_h0l.jpg	29k
2010src0792r_hk0.jpg	42k
2010src0792ra.hkl	294k

Other Files

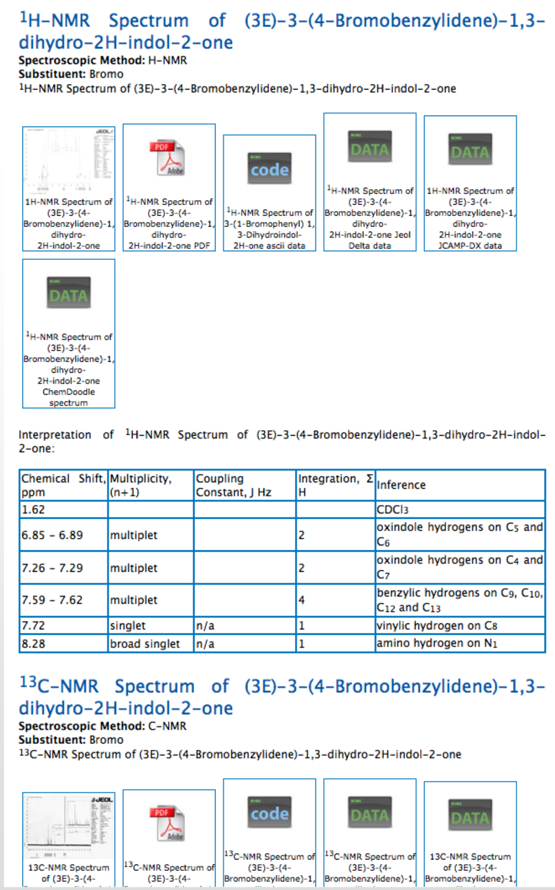
2010src0792r.doc	611k
2010src0792r_1kl.jpg	33k
2010src0792r_h1l.jpg	28k
2010src0792r_hk1.jpg	42k
2010src0792ra.inchi	2k
2010src0792ra.ins	5k
2010src0792ra.mol	3k
shelxl.lst	52k

This is the latest version of this item.

Repository Staff Only: **item control page**

DOI's for Data

- LabTrove (unstructured data)
- Locked down HTML export of selected records



^{1}H-NMR Spectrum of (3E)-3-(4-Bromobenzylidene)-1,3-dihydro-2H-indol-2-one

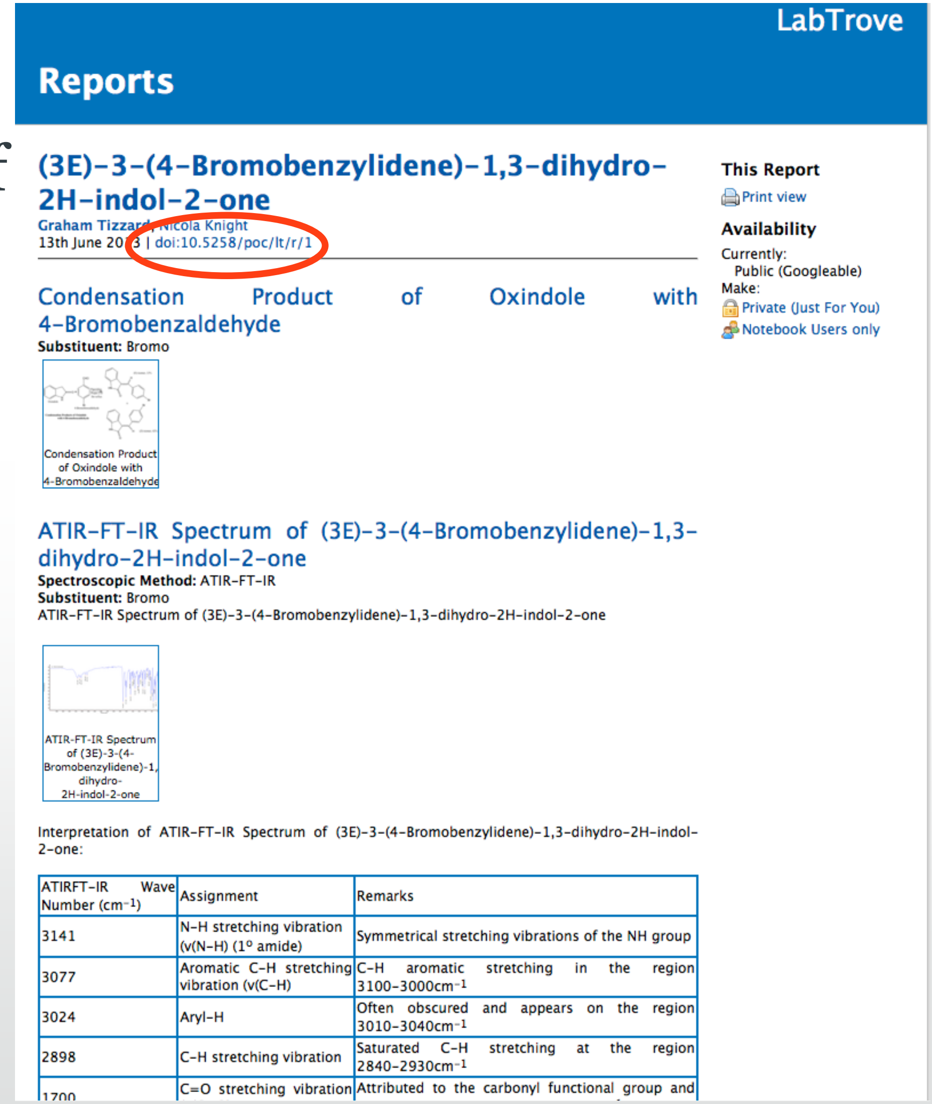
Spectroscopic Method: H-NMR
Substituent: Bromo
^{1}H-NMR Spectrum of (3E)-3-(4-Bromobenzylidene)-1,3-dihydro-2H-indol-2-one

Interpretation of ^{1}H-NMR Spectrum of (3E)-3-(4-Bromobenzylidene)-1,3-dihydro-2H-indol-2-one:

Chemical Shift, ppm	Multiplicity, (n+1)	Coupling Constant, J Hz	Integration, Σ H	Inference
1.62				$CDCl_3$
6.85 - 6.89	multiplet		2	oxindole hydrogens on C_5 and C_6
7.26 - 7.29	multiplet		2	oxindole hydrogens on C_4 and C_7
7.59 - 7.62	multiplet		4	benzylic hydrogens on C_9, C_{10}, C_{12} and C_{13}
7.72	singlet	n/a	1	vinylic hydrogen on C_8
8.28	broad singlet	n/a	1	amino hydrogen on N_1

^{13}C-NMR Spectrum of (3E)-3-(4-Bromobenzylidene)-1,3-dihydro-2H-indol-2-one

Spectroscopic Method: C-NMR
Substituent: Bromo
^{13}C-NMR Spectrum of (3E)-3-(4-Bromobenzylidene)-1,3-dihydro-2H-indol-2-one



(3E)-3-(4-Bromobenzylidene)-1,3-dihydro-2H-indol-2-one

Graham Tizzard, Nicola Knight
13th June 20[illegible]3 | doi:10.5258/poc/lt/r/1

This Report
Print view

Availability
Currently:
Public (Googleable)
Make:
Private (Just For You)
Notebook Users only

Condensation Product of Oxindole with 4-Bromobenzaldehyde

Substituent: Bromo

ATIR-FT-IR Spectrum of (3E)-3-(4-Bromobenzylidene)-1,3-dihydro-2H-indol-2-one

Spectroscopic Method: ATIR-FT-IR
Substituent: Bromo
ATIR-FT-IR Spectrum of (3E)-3-(4-Bromobenzylidene)-1,3-dihydro-2H-indol-2-one

Interpretation of ATIR-FT-IR Spectrum of (3E)-3-(4-Bromobenzylidene)-1,3-dihydro-2H-indol-2-one:

ATIRFT-IR Wave Number (cm^{-1})	Assignment	Remarks
3141	N-H stretching vibration (ν(N-H) (1° amide)	Symmetrical stretching vibrations of the NH group
3077	Aromatic C-H stretching vibration (ν(C-H)	C-H aromatic stretching in the region 3100-3000cm^{-1}
3024	Aryl-H	Often obscured and appears on the region 3010-3040cm^{-1}
2898	C-H stretching vibration	Saturated C-H stretching at the region 2840-2930cm^{-1}
1700	C=O stretching vibration	Attributed to the carbonyl functional group and

SIMS:
Sample Information Management System

- A standard/format for crystallographic sample and experiment data management and archival

- Supported by CrystalClear and NCS Portal, providing interaction between facility, instruments and CIF, ImgCIF etc


```xml
<?xml version="1.0"?>
<FileType Type = "Sample">
  <User Name = "Russ.Athay"/>
  <ProjectName Name = "P1" />
  <SampleName Name = "EPSIMS11" />
  <ImageDirectory Directory = "C:\Images\RussCyt"/>
  <ImageTemplate Template = "EPSIMS11"/>
  <ProjectInfo>
    <DatabaseID ID = "AX3T45"/>
    <UserID ID = "AX2456"/>
    <FacilityID ID = "CAMSM33"/>
    <Priority Priority = "Medium"/>
    <SubmissionForm Form = "Standard OnLine"/>
    <BulkProperties Properties = "Multiple Crystals in
    <BulkMorphology Morphology = "Mostly rods"/>
    <SolutionGoal Goal = "Publishable Structure"/>
    <UserFormula Formula = "C9O5N3H13"/>
    <LabelingScheme Scheme = "Standard"/>
    <Sensitivity Sensitivity = "Light Sensitive"/>
    <Safety Safety = "Caustic medium"/>
    <DateProjectSubmission Date = "2012.01.04"/>
    <DateProjectReport Date = "2012.01.06"/>
  </ProjectInfo>
  <SampleInfo>
    <CrystalNumber Number = "4"/>
    <Crystallographer Name = "Graham"/>
    <PreparationScheme Scheme = "Chip of large crystal
    <FragmentCut Cut = "Yes"/>
    <CrystalColor Color = "Colorless"/>
    <CrystalSize X = "0.20" Y = "0.20" Z = "0.20"/>
    <CrystalMorphology Morphology = "Prism"/>
    <Solvent Solvent = "PEG200"/>
    <ManipulationMedium Medium = "Stuff"/>
    <MountType Type = "Fiber"/>
    <SampleMountID ID = "NCS2765"/>
    <VisualQuality Quality = "Unknown"/>
    <OpticalProperties Properties = "Transparent"/>
    <VisuallyTwinned Twinned = "No"/>
    <DateDatabaseExtraction Date = "2012.01.06"/>
    <DateMounted Date = "2012.01.07"/>
    <DateScreened Date = "2012.01.07"/>
    <DateCollected Date = "2012.01.08"/>
    <DateProcessed Date = "2012.01.09"/>
    <DateSolved Date = "2012.01.09"/>
    <DateDatabaseDeposit Date = "2012.01.10"/>
  </SampleInfo>
  <Screen>
    <ScreenProgram Program = "CrystalClear3.1a4"/>

  <Screen>
    <ScreenProgram Program = "CrystalClear3.1a4"/>
    <Environment TempC = "128.00" Pressure = "29.89" Humidity = "62.30"/>
    <ColdStreamTemp TempC = "-148.40"/>
    <Instrument Instrument = "AFC12 Saturn 724+"/>
    <Wavelength Wavelength = "0.7070000"/>
    <Element Element = "Moly"/>
    <ScreenScans>
      <Scan ScanNumber = "1" Axis = "Omega">
        <ScanRange Start = "-60.0" End = "-59.0" Width = "1.0" Step = "1.0"
        <GonioAxes ChiOrKappa = "54.0" PhiOrOmega = "0.0"/>
        <DetectorSettings Distance = "48.8" Exposure = "10.0" TwoTheta = "2
        <DetectorMode Binning = "2" Dezinger = "Yes"/>
      </Scan>
    </ScreenScans>
    <Rank Rank = "135"/>
    <IndexProgram Program = "CrystalClear3.1a4"/>
    <IndexAlgorithm Algorithm = "d*TREK"/>
    <CSDMatches>
      <CSDMatch Match = "Someone in Cambridge did this years ago"/>
      <CSDMatch Match = "St Andrews did this in 1994"/>
    </CSDMatches>
    <State Type = "Screen">
      <Cell A = "5.123" B = "13.987" C = "14.764" Alpha = "90.000" Beta = "
      <CellESD A = "0.000000" B = "0.000000" C = "0.000000" Alpha = "0.0000
      <CrystalSystem System = "Orthorhombic"/>
      <SpaceGroupName Name = "P 2 2 2"/>
      <LaueClass Class = "m m m"/>
      <Lattice Lattice = "P"/>
      <Centricity Centricity = "acentric"/>
      <SpaceGroupNumber Number = "16"/>
      <Rmerge Rmerge = "4.30"/>
      <Completeness Completeness = "98.70"/>
      <Redundancy Redundancy = "4.30"/>
      <TotalReflections Reflections = "34987"/>
      <UniqueReflections Reflections = "27654"/>
      <RejectedReflections Reflections = "2456"/>
      <MaxResolution Resolution = "0.83"/>
    </State>
  </Screen>
  <Collect>
    <StrategyProgram Program = "CrystalClear 3.1a4"/>
    <CollectSchedule Schedule = "Default_Collect"/>
    <PredictedCompleteness Completeness = "99.80"/>
    <PredictedRedundancy Redundancy = "4.40"/>
    <CollectProgram Program = "CrystalClear 3.1a4"/>
    <Environment TempC = "128.00" Pressure = "29.89" Humidity = "62.30"/>

    <EmpiricalTmin Tmin = "0.04"/>
    <EmpiricalTmax Tmax = "0.66"/>
    <Twins>
      <Twin Number = "1" TwinLaw = "Slide 1/2"/>
      <Twin Number = "2" TwinLaw = "Twist 1/3"/>
    </Twins>
  </Process>
  <Solve>
    <SolutionProgram Program = "Olex2/Shelx"/>
    <SolutionMethod Method = "Direct"/>
    <Restraints>
      <Restraint Parameter = "Param1" Value = "0.0000" Tolerance = "0.1000"/>
      <Restraint Parameter = "Param2" Value = "1.0000" Tolerance = "0.2000"/>
    </Restraints>
    <Constraints>
      <Constraint Parameter = "Param3" Value = "0.0000" Tolerance = "0.1000"/>
      <Constraint Parameter = "Param4" Value = "-1.0000" Tolerance = "0.0010"/>
    </Constraints>
    <Structure Type = "Solve">
      <R1 Percentage = "4.30"/>
      <Atoms>
        <Atom>
          <AtomCoord Label="O1" Symbol="O" x="0.000000" y="0.100000" z="0.200000" Occupancy="1.000000"/>
          <AtomAniso U11="0.000000" U22="0.010000" U33="0.020000" U23="0.000000" U13="0.003000" U12="0.006000"/>
        </Atom>
        <Atom>
          <AtomCoord Label="N1" Symbol="N" x="0.000000" y="0.200000" z="0.400000" Occupancy="1.000000"/>
          <AtomAniso U11="0.000000" U22="0.020000" U33="0.040000" U23="0.000000" U13="0.006000" U12="0.012000"/>
        </Atom>
      </Atoms>
    </Structure>
  </Solve>
  <History>
    <Step Date = "12.01.11" Type = "Screen"></Step>
    <Step Date = "12.01.11" Type = "Note">The sample was abandoned here becaue the reflections were very weak and
it could not be indexed</Step>
  </History>
  <Results>
    <ValidationProgram Program = "CheckCif"/>
    <ValidationErrors>
      <ValidationError Error = "Rmerge too high"/>
      <ValidationError Error = "Resolution too low"/>
    </ValidationErrors>
  </Results>
  <EmailAddress Email = "some_user@localhost"/>
  <Status Status = "Ready"/>
  <LastError Error = "None"/>
  <Rank Rank = "135"/>
```

elnItemManifest

UNIVERSITY OF Southampton
School of Chemistry

- 3 layer metadata model for description, export & packaging
- This is the first (information) layer – leads into knowledge
- Published through Dial-a-Molecule athttp://wp.me/p2JoQ6-xF & submitted to J. ChemInf

```xml
<?xml version="1.0"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">

<xs:annotation>
  <xs:documentation>Change history</xs:documentation>
  <xs:documentation>==============</xs:documentation>
  <xs:documentation>15 June 2012 - [CLB] Created from enDataDescription prototype</xs:documentation>
</xs:annotation>

<xs:annotation>
  <xs:documentation>Definitions of data types used in the manifest of the ELN item</xs:documentation>
  <xs:documentation>~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~</xs:documentation>
</xs:annotation>

<xs:complexType name="keywordSet">
  <xs:annotation>
     <xs:documentation>A list of terms</xs:documentation>
  </xs:annotation>
  <xs:sequence>
    <xs:element name="keyword" type="xs:string" minOccurs="0" maxOccurs="unbounded"/>
  </xs:sequence>
</xs:complexType>

<xs:complexType name="identifierSet">
  <xs:sequence>
    <xs:element name="primaryLocalIdentifier" type="xs:string">
      <xs:annotation>
        <xs:documentation>Primary string, URI, or item in any other format that enables</xs:documentation>
        <xs:documentation>this record to be located uniquely in the originating system</xs:documentation>
      </xs:annotation>
    </xs:element>
    <xs:element name="otherLocalIdentifier" type="xs:string" minOccurs="0" maxOccurs="unbounded">
      <xs:annotation>
        <xs:documentation>[Optional] Alternative means of locating record in the originating system</xs:documentation>
      </xs:annotation>
    </xs:element>
    <xs:element name="accessIdentifier" type="xs:anyURI" minOccurs="0">
      <xs:annotation>
        <xs:documentation>[Optional] URI that provides a direct link to the content.</xs:documentation>
        <xs:documentation>If included, must be a 'linked data' URI giving open access</xs:documentation>
      </xs:annotation>
    </xs:element>
  </xs:sequence>
</xs:complexType>
```

```xml
<xs:complexType name="contentInformation">
  <xs:annotation>
     <xs:documentation>Text that describes what the item is, with a descriptor of the</xs:documentation>
     <xs:documentation>digital type, or "undefined" if no corresponding MIME format</xs:documentation>
  </xs:annotation>
  <xs:sequence>
    <xs:element name="description" type="xs:string"/>
    <xs:element name="mimeType" type="xs:string" default="undefined"/>
  </xs:sequence>
</xs:complexType>

<xs:complexType name="relatedID">
  <xs:annotation>
     <xs:documentation>Nature of the related information, for example, publication or related work</xs:documer
     <xs:documentation>Id can be any string, but DOI preferred if the related item is a publication</xs:docume
  </xs:annotation>
  <xs:sequence>
    <xs:element name="relationship" type="xs:string"/>
    <xs:element name="id" type="xs:string"/>
  </xs:sequence>
</xs:complexType>

<xs:complexType name="relatedItemSet">
  <xs:annotation>
     <xs:documentation>Zero or more item(s) of related information</xs:documentation>
  </xs:annotation>
  <xs:sequence>
    <xs:element name="item" type="relatedID" minOccurs="0" maxOccurs="unbounded"/>
  </xs:sequence>
</xs:complexType>

<xs:complexType name="contributorInformation">
  <xs:annotation>
     <xs:documentation>For example, Author, Funding Body, PI, institution, ...</xs:documentation>
     <xs:documentation>Plain text, but name ideally complemented by unique identifiers</xs:documentation>
  </xs:annotation>
  <xs:sequence>
    <xs:element name="role" type="xs:string"/>
    <xs:element name="name" type="xs:string"/>
  </xs:sequence>
</xs:complexType>

<xs:complexType name="contributorSet">
  <xs:annotation>
```

Thanks

DCC
RSC | Advancing the Chemical Sciences
UK National Crystallography Service
Pioneering research and skills
EPSRC
eCrystals
Rigaku
Unilever Cambridge Centre For Molecular Science Informatics
UKOLN
LabTrove
hefce
I_2S_2 Infrastructure for Integration in Structural Sciences
Chemical Database Service
R4L Repository for the Laboratory
Science & Technology Facilities Council
ChemistryCentral
DataCite Helping you to find, access, and reuse data
eprints
JISC
ncs
UNIVERSITY OF Southampton
ROYAL SOCIETY OF CHEMISTRY
Jisc
eBank UK
diamond
ICAT
IUCr
Microsoft
e-Science
ChemSpider The free chemical database