Computational Chemistry looking forward: How can we as a community improve our field?

Catherine E. Peishoff
D3R workshop
March 10-11, 2016

First a quick intro...

- PhD, Purdue University, 1984, Jorgensen group
- Practicing computational chemist at Lederle Labs (now Pfizer, 1985-1987) and SmithKline Beckman (now GSK, 1987-present)
- Current role: Vice President, Chemical Sciences, GSK R&D
  - Computational chemistry and Cheminformatics
  - Structural biology and Biophysics
  - Biological Mass Spectrometry
  - Analytical Chemistry
  - Medicinal Chemistry

- Some fun along the way...
  - Associate Editor, *Journal of Medicinal Chemistry*
  - Scientific Advisory Board, Keystone Symposia
  - External Advisory Board, NIH Molecular Libraries Screening Centers Network
  - Protein Structure Initiative reviews
  - NIH grant CD4/gp120 (with Hendrickson lab)
  - NY Times Technology section (February , 2013) *'And Now, From IBM, Chef Watson'*

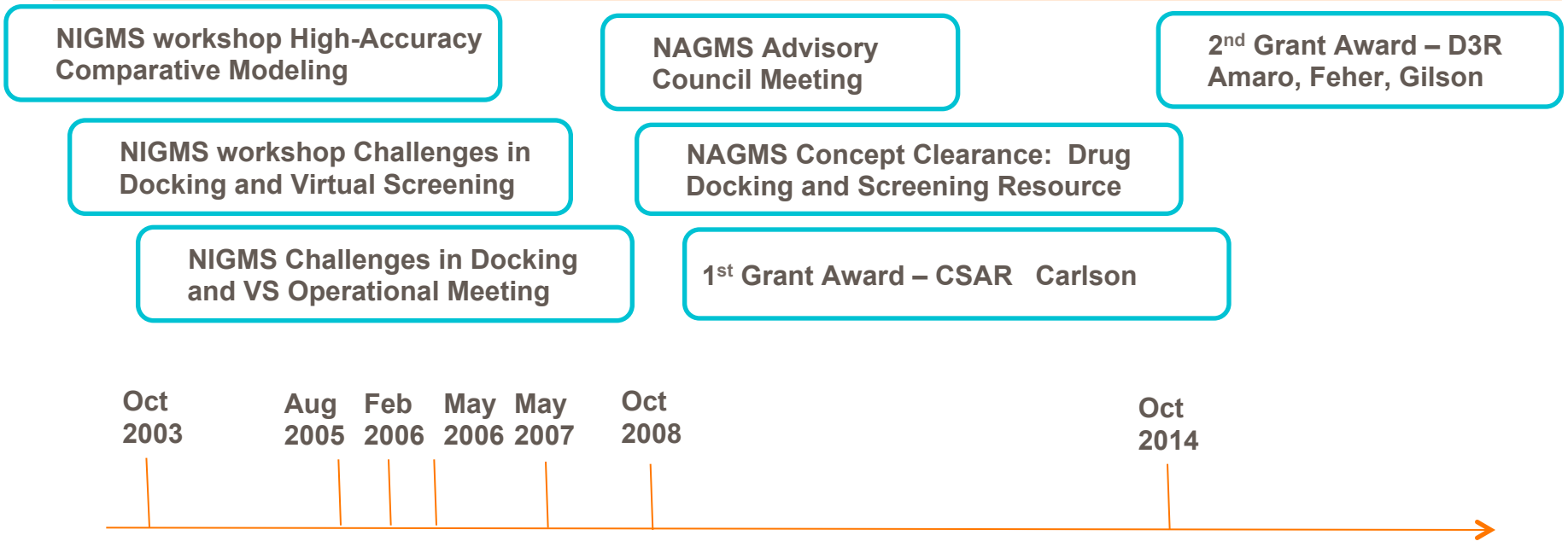
First a quick intro...

- PhD, Purdue University, 1984, Jorgensen group
- Practicing computational chemist at Lederle Labs (now Pfizer, 1985-1987) and SmithKline Beckman (now GSK, 1987-present)
- Current role: Vice President, Chemical Sciences, GSK R&D
  - Computational chemistry and Cheminformatics
  - Structural biology and Biophysics
  - Biological Mass Spectrometry
  - Analytical Chemistry
  - Medicinal Chemistry

RETIRED 12/2015

- Some fun along the way...
  - Associate Editor, *Journal of Medicinal Chemistry*
  - Scientific Advisory Board, Keystone Symposia
  - External Advisory Board, NIH Molecular Libraries Screening Centers Network
  - Protein Structure Initiative reviews
  - NIH grant CD4/gp120 (with Hendrickson lab)
  - NY Times Technology section (February , 2013) *'And Now, From IBM, Chef Watson'*

A timeline...


NIGMS workshop High-Accuracy Comparative Modeling
NAGMS Advisory Council Meeting
2nd Grant Award – D3R Amaro, Feher, Gilson
NIGMS workshop Challenges in Docking and Virtual Screening
NAGMS Concept Clearance: Drug Docking and Screening Resource
NIGMS Challenges in Docking and VS Operational Meeting
1st Grant Award – CSAR Carlson
Oct 2003
Aug 2005
Feb 2006
May 2006
May 2007
Oct 2008
Oct 2014

Another timeline...

NIGMS workshop High-Accuracy Comparative Modeling

NAGMS Advisory Council Meeting

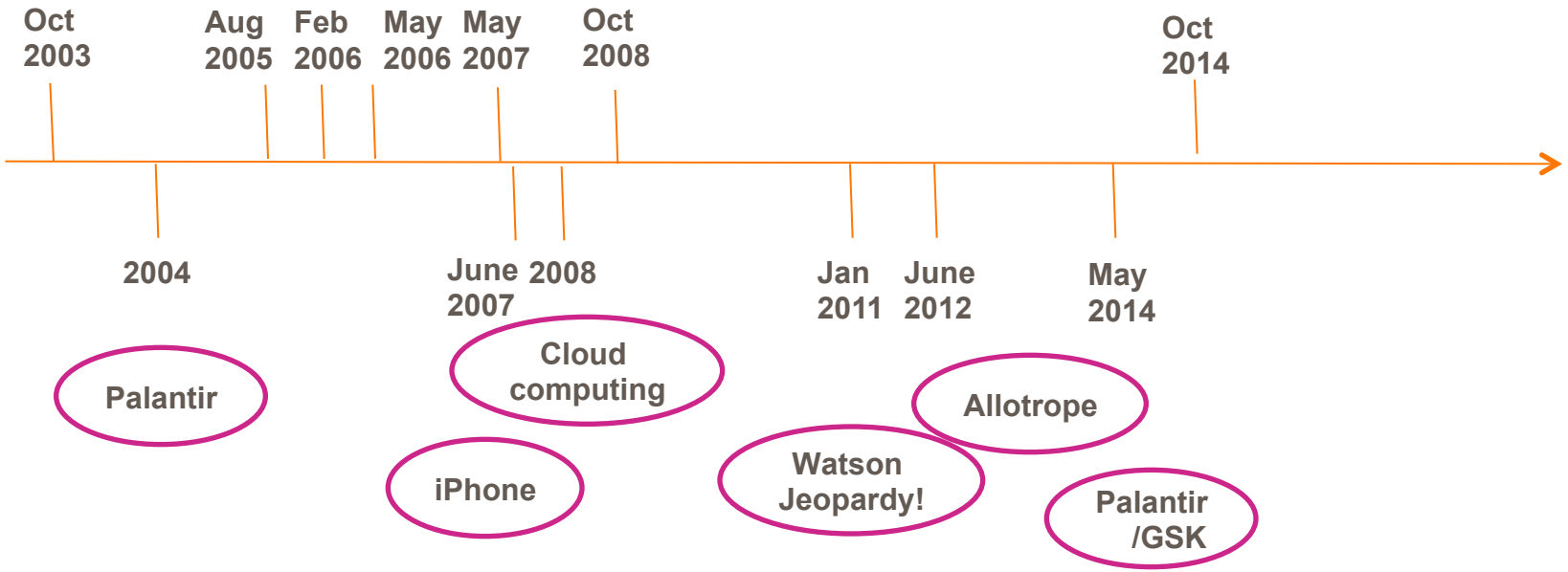
2nd Grant Award – D3R Amaro, Feher, Gilson

NIGMS workshop Challenges in Docking and Virtual Screening

NAGMS Concept Clearance: Drug Docking and Screening Resource

NIGMS Challenges in Docking and VS Operational Meeting

1st Grant Award – CSAR Carlson



A little help from my friends...

Thinking about both the science and culture of computational chemistry today, develop a Strengths/Weaknesses/ Opportunities/Threats analysis

>100 participants

Pharma companies – individual and group discussions

Academics – personal reflections

D3R workshop exercise

Computational Chemistry Strengths:

- Many examples of documented impact on biomedical problems, an accepted field
- Increasingly clever applications, e.g. molecular interaction predictions applied to protein maturation for vaccines, biopharms, and enzymes for manufacturing
- Metrics that influence other fields, e.g. molecular similarity, ligand efficiency
- Directly benefit from other fields, e.g. informatics, computer science in general
- Skills are problem agnostic…just need a few atoms and some data
- Ability to work on many different problems leads to cross-fertilization
- People in this field love what they're doing...creative, interdisciplinary
- Field encourages and accommodates a diversity of backgrounds, there are few dogmas that are inviolable and the field is not inward-looking
- A relatively young field (40ish) with a high degree of personal connection
- Tasks such as docking, conformational analysis, homology model building, QM calculations have become faster and easier to run
- Tools are now in the hands of experimentalists

- We understand the limits of our tools, have had enough usage to know strengths and weaknesses

Computational Chemistry Weaknesses:

- A relatively young field (40ish) with a broad range of applications
- Lack of rigorous statistics, experimental design, head to head evaluations, data standards and governance, good data sets, open source software, general sharing
- Key partners, e.g. medicinal chemistry still rely a lot on art thus not the best problems or data for developing quantitative analysis
- Reliance on experimentalists coupled to lack of rigor increases the weight of opinion and need for salesmanship
- Proliferation of tools creating a generation of super-users, not computational chemists
- The perception that comp chem is “easy” when it’s the software tools that are easy
- Number of projects that can be positively impacted by qualitative analysis/intuition distracts from the core science and keeps us overly reliant on the individual
- Our data and knowledge is fragmented, no educational standards, documentation, or systematic way of addressing
- Publication bias (little sharing of failures) and lack of high quality publications

- Many methods appear to have reached a plateau, e.g. docking, protein folding, homology model building, DFT, QSAR,etc…variations more than advances

Computational Chemistry Threats:

- The dynamics of the talent pool are starting to shift and will shift significantly
  - Projected retirements
  - Fewer basic research groups graduating students
  - Career opportunities increasing at all levels
- Loss of institutional knowledge could be significant over the next decade
- Contraction of pharma yields fewer jobs (albeit not as hard hit as e.g. medicinal chemistry) leading to a negative impression of the industry with students, knock on effect to software
- Projects are becoming more complex (less amenable to xray) and therapy modalities are changing (cell and gene based) leading to a perceived need for the same or fewer computational chemists
- Students have little to no engagement with real time pharma problems and are much more exposed to wildly growing fields such as data analysis and to entrepreneurial training
- There is increased computational work in other industries, e.g. energy, materials science

- Fewer centers of technology excellence in academia and academic groups are increasingly emphasizing application over method development…is this because we don't know what to do?
- Loss of the excitement for methods development and the comradery/competition it generated

Computational Chemistry Environment for Opportunity:

- Data is significantly easier to come by…automation for synthesis and assay, CROs
- Experimental techniques are becoming more sensitive and opening a window to understanding functional dynamics reinvigorating the need for e.g. simulation studies
- Data aggregation/integration/visualization tools coupled to deployment of our most solid methods will continue to drive usage of computational chemistry
- New methods in data processing, informatics, and machine learning coupled to Moore's law will drive the need for strong computational scientists in pharma
- Public datasources have emerged, e.g. PDB, ChEMBL, BindingDB, PDBbind
- Large gap in graduate student training...can make knowledge visible/available
- Pre-competitive working is becoming more the norm in the pharma industry, e.g.
  - Transcelerate for clinical trial design; Allotrope for analytical data
  - Structural Genomics Consortium for protein crystal structures and tool molecules
  - The Center for Therapeutic Target Validation and Altius for target/disease linkage
  - SGC – University of North Carolina Center for Kinase Research
- Partners exist to support/accelerate discussion, e.g. ACS, NIH
- Examples from which to learn, e.g. Center for Selective C-H Functionalization (NSF funded, 23 research groups across 15 institutions)

Themes emerge

- Opportunity abounds – many biomedical problems can benefit from current computational chemistry methods

- Return on Investment (ROI) demands better, cheaper, and/or faster – we're doing well on cheaper and faster, but only incrementally changing better

- We (or at least some of us) understand the strengths and weaknesses of our methods very well, but struggle to come together to solve the really hard problems

- Demand for skills associated with comp chem is likely to increase – people who can research, develop hypotheses, and code have many, many options

- The people who were in the trenches 30 years ago developing the methods we use today are highly likely to have left the field within 10 years – facing a significant loss of institutional knowledge

- We are not providing foundational education/training to our students (comp and synthetic chemists)

The time is now

- There are evolving community efforts to create infrastructure on which to improve our science and culture
  - D3R (Drug, Design, Data Resource) www.drugdesigndata.org
    - SAR and crystallography data, blind prediction challenges, computational chemistry education
    - Contacts: Mike Gilson, Rommie Amaro, Vicki Feher UCSD
  - ACS Pharma Leaders Meeting and subgroups (C&E News, Vol 94/4 p 36)
    - Computational/chemistry and structural biology; physical compound sets
    - Contacts: Cathy Peishoff (peishoffc@gmail.com) or Brad Sherborne (brad.sherborne@merck.com)
    - sessions planned at the Fall ACS meeting in Philadelphia
- We need to partner and use our strengths to accelerate the pace of change
- Please engage
- Please be generous
- With thanks to GSK, Merck, Pfizer, BMS, Vertex, Roche, UCSD, Kate/Chuck/Frank/Ajay/Bill