BIOGRAPHICAL SKETCH

Provide the following information for the Senior/key personnel and other significant contributors.

Follow this format for each person. DO NOT EXCEED FIVE PAGES.

NAME: Craig, Paul

eRA COMMONS USER NAME (agency login): PCRAIG

POSITION TITLE: Professor

EDUCATION/TRAINING

INSTITUTION AND LOCATION	DEGREE (if applicable)	Completion Date MM/YYYY	FIELD OF STUDY
Oral Roberts University, Tulsa, OK	BS	05/1979	Chemistry
The University of Michigan, Ann Arbor, MI	PHD	05/1985	Biological Chemistry
Henry Ford Hosptial, Detroit, MI	Postdoctoral Fellow	08/1988	Physical biochemistry of blood clotting

A. Personal Statement

My formal training is in biochemistry and I have skills in communicating with computer scientists. While I am not a programmer, I consider myself a computational biochemist and that is my research focus - developing computational tools to explore protein structure and function, then implementing them for researchers ranging from students to professionals. A summer research project at Brookhaven National Labs led to collaboration with Herbert Bernstein, which has resulted in the development of three plugins for PyMOL: EZ-Viz [Grell, 2006], ConScript [Mottarella, 2010], and ProMOL [Hanson, 2014]. This has continued for 10 years, most recently in our 2015 publication, “Annotation of Proteins of Unknown Function: Initial Enzyme Results“ [McKay T, 2015]. Herbert is a mathematician and computer scientist with extensive experience in structural biology and biological databases. I combine a biological and functional perspective with many years of experience in the use and development of software. Our complementary skill sets have led to a fruitful collaboration that has been highly beneficial to students on both campuses. On a personal level, I find great fulfillment in having published six manuscripts that included 25 undergraduate student authors, as well as five manuscripts with four different pre-tenure faculty members.

1. Grell L*, Parkin C*, Slatest L, Craig PA. EZ-Viz, a tool for simplifying molecular viewing in PyMOL. Biochem Mol Biol Educ. 2006 Nov;34(6):402-7. PMID: 21638731.
2. Mottarella SE*, Rosa M*, Bangura A*, Bernstein HJ, Craig PA. Conscript: RasMol to PyMOL script converter. Biochem Mol Biol Educ. 2010 Nov;38(6):419-22. PMID: 21567873; PMCID: PMC3134254.
3. Hanson B*, Westin C*, Rosa M*, Grier A, Osipovitch M, MacDonald ML*, Dodge G*, Boli PM*, Corwin CW*, Kessler H*, McKay T*, Bernstein HJ, Craig PA. Estimation of protein function using template-based alignment of enzyme active sites. BMC Bioinformatics. 2014 Mar 27;15:87. PMID: 24669788; PMCID: PMC4229977.
4. McKay T*, Hart K*, Horn A*, Kessler H*, Dodge G*, Bardhi K*, Bardhi K*, Mills JL, Bernstein HJ, Craig PA. Annotation of proteins of unknown function: initial enzyme results. J Struct Funct Genomics. 2015 Mar;16(1):43-54. PMID: 25630330; PMCID: PMC4332402.

* indicates an undergraduate student and ‡ indicates a pre-tenure faculty member

B. Positions and Honors

Positions and Employment

1988 - 1993 Analytical Biochemist, BioQuant, Ann Arbor, MI
1993 - 1999 Assistant Professor, Rochester Institute of Technology, Dept. of Chemistry, Rochester, NY
1999 - 2003 Associate Professor, Rochester Institute of Technology, Dept. of Chemistry, Rochester, NY
2001 - 2002 Visiting Scholar, UCSD, San Diego Supercomputing Center, San Diego, CA
2003 - Professor, Rochester Institute of Technology (RIT), Dept. of Chemistry, Rochester, NY
2011 - 2012 Associate Head, RIT, Dept. of Chemistry, Rochester, NY

2012 - Head, RIT, School of Chemistry & Materials Science, Rochester, NY

Other Experience and Professional Memberships

1976 - Member, American Chemical Society
1993 - Member, American Society for Biochemistry & Molecular Biology
2007 - Member, American Crystallographic Association

Honors

1979 National Merit Scholar, National Merit Scholar Corporation
1986 Fellow, American Heart Association

C. Contribution to Science

1. *Software Development and Implementation in Computational Biochemistry.* Since 1993, I have worked with students and faculty members on projects at the interface between life science and computer science programs. In 2004, our group at RIT began studying 3D modeling of macromolecules with PyMOL. Undergraduates from RIT then wrote the EZ-Viz plugin for PyMOL [Grell, 2006] to overcome the steep learning curve for PyMOL. This led to a collaboration with Herbert Bernstein that has resulted in the creation of a script convertor that allows PyMOL to accept Jmol/RasMOL scripts in the PyMOL command line [Mottarella, 2010]. Subsequently, our students developed ProMOL to help people propose protein function based on comparison to a series of active site templates [Hanson, 2014] and have used ProMOL to propose functions for a number of proteins [McKay, 2015].

    a. Grell L*, Parkin C*, Slatest L, Craig PA. EZ-Viz, a tool for simplifying molecular viewing in PyMOL. Biochem Mol Biol Educ. 2006 Nov;34(6):402-7. PMID: 21638731.
    b. Mottarella SE*, Rosa M*, Bangura A*, Bernstein HJ, Craig PA. Conscript: RasMol to PyMOL script converter. Biochem Mol Biol Educ. 2010 Nov;38(6):419-22. PMID: 21567873; PMCID: PMC3134254.
    c. Hanson B*, Westin C*, Rosa M*, Grier A, Osipovitch M, MacDonald ML*, Dodge G*, Boli PM*, Corwin CW*, Kessler H*, McKay T*, Bernstein HJ, Craig PA. Estimation of protein function using template-based alignment of enzyme active sites. BMC Bioinformatics. 2014 Mar 27;15:87. PMID: 24669788; PMCID: PMC4229977.
    d. McKay T*, Hart K*, Horn A*, Kessler H*, Dodge G*, Bardhi K*, Bardhi K*, Mills JL, Bernstein HJ, Craig PA. Annotation of proteins of unknown function: initial enzyme results. J Struct Funct Genomics. 2015 Mar;16(1):43-54. PMID: 25630330; PMCID: PMC4332402.

2. *Biochemistry Education*. Early in my RIT career, I was tasked with creating an updated biochemistry lab. I received NSF support for this effort, primarily because of mentoring by Chris Rollman, who offered significant guidance as I prepared the proposal. This led to my first publication at RIT [Craig, 1999]. Since that time, I have been heavily engaged in the use [Craig, 2013] and development of molecular visualization software [Grell, 2006]. Among my most memorable experiences was my 2001-2002 sabbatical with Philip Bourne, who was an associate director of the Protein Data Bank. Since then I have given invited talks at ACS and ASBMB conferences and recently led a workshop on the assessment of student learning with molecular visualization at a 2013 ASBMB conference, “Transforming Undergraduate Education in Molecular Life Sciences”. I am grateful to Jenny Loertscher for the invitation and her very helpful guidance on running a dynamic, interactive workshop. One of my proudest moments occurred at the 2015 ASBMB meeting when one of my undergraduates gave an invited talk [Hart, 2015].

    a. Craig PA. A Project Oriented Biochemistry Laboratory Course. Journal of Chemical Education. 1999 August 01; 76(8):1130.
    b. Grell L*, Parkin C*, Slatest L, Craig PA. EZ-Viz, a tool for simplifying molecular viewing in PyMOL. Biochem Mol Biol Educ. 2006 Nov;34(6):402-7. PMID: 21638731.
    c. Craig PA, Michel LV‡, Bateman RC. A survey of educational uses of molecular visualization freeware. Biochem Mol Biol Educ. 2013 May-Jun;41(3):193-205. PMID: 23649886; PMCID: PMC4098825.
    d. Hart K*, McKay T*, Tedla-Boyd W*, Mills JL, Bernstein HJ, Craig PA. Protein Function Prediction Using ProMOL and PyMOL. FASEB Journal. ASBMB; 2015 March 29; Boston, MA, USA. Rockville Pike, MD.

3. *Classical Enzymology*. I received my Ph.D. in Biological Chemistry with Eugene E. Dekker (who earned his Ph.D. with William C.M. Rose), where we focused on enzymology and protein chemistry of amino acid metabolism [Craig, 1986]. I was also fascinated by the role of metal ions in protein structure and function [Craig, 1990]. My post-doctoral training was in the physical biochemistry of blood clotting with Joseph D. Shore and Steven T. Olson. I studied enzyme kinetics with factor X_a (a serine protease) in its interactions with antithrombin III and heparin using UV-Visible spectroscopy and stopped-flow fluorescence [Craig, 1989; Olson, 1992].

    a. Craig PA, Dekker EE. L-threonine dehydrogenase from Escherichia coli K-12: thiol-dependent activation by Mn^{2+}. Biochemistry. 1986 Apr 22;25(8):1870-6. PMID: 3518793.
    b. Craig PA, Olson ST, Shore JD. Transient kinetics of heparin-catalyzed protease inactivation by antithrombin III. Characterization of assembly, product formation, and heparin dissociation steps in the factor Xa reaction. J Biol Chem. 1989 Apr 5;264(10):5452-61. PMID: 2925612.
    c. Craig PA, Dekker EE. The sulfhydryl content of L-threonine dehydrogenase from Escherichia coli K-12: relation to catalytic activity and Mn^{2+} activation. Biochim Biophys Acta. 1990 Jan 19;1037(1):30-8. PMID: 2104757.
    d. Olson ST, Björk I, Sheffer R, Craig PA, Shore JD, Choay J. Role of the antithrombin-binding pentasaccharide in heparin acceleration of antithrombin-proteinase reactions. Resolution of the antithrombin conformational change contribution to heparin rate enhancement. J Biol Chem. 1992 Jun 25;267(18):12528-38. PMID: 1618758.

4. *Analytical Biochemistry and Proteomics*. I spent five year at a startup company in Ann Arbor, MI, called BioQuant, where we studied noninvasive detection of small molecules in alternative body fluids, including the analysis of cocaine metabolites in blood, urine and saliva [Schramm, 1993]. I also gained experience with synthesis of labeled bioconjugates for use in immunoassays [Schramm, 1990]. When I moved to academia, I became interested in proteomics, which led to an article about proteomics for undergraduates [Kim, 2010] and a simulation of 2D electrophoresis and Tandem MS, which provides undergraduates with training with technology that their institutes may not be able to afford [Fisher, 2012].

    a. Schramm W, Smith RH, Jackson TM, Craig PA, Grates HE, Minton LL. Rapid solid-phase immunoassay for 6-keto prostaglandin F1 alpha on microplates. Clin Chem. 1990 Mar;36(3):509-14. PMID: 2311222.
    b. Schramm W, Craig PA, Smith RH, Berger GE. Cocaine and benzoylecgonine in saliva, serum, and urine. Clin Chem. 1993 Mar;39(3):481-7. PMID: 8448861.
    c. Kim TD‡, Craig PA. Introducing proteomics in the undergraduate curriculum: A simple 2D gel electrophoresis exercise with serum proteins. Biochem Mol Biol Educ. 2010 Jan;38(1):29-34. PMID: 21567787.
    d. Fisher A*, Sekera E*, Payne J, Craig P. Simulation of two dimensional electrophoresis and tandem mass spectrometry for teaching proteomics. Biochem Mol Biol Educ. 2012 Nov-Dec;40(6):393-9. PMID: 23166029.

Complete List of Published Work in MyBibliography:
http://www.ncbi.nlm.nih.gov/sites/myncbi/paul.craig.2/bibliography/42699640/public/?sort=date&direction=descending

D. Research Support

Ongoing Research Support

2015/06/01-2017/05/31
1503811, NSF DUE-IUSE-Engaged Student Learning: Exploration
Craig, Paul (PI)
Collaborative Research: Using protein function prediction to promote hypothesis-driven thinking in undergraduate biochemistry education

Students on six campuses (California Polytechnic San Luis Obispo, Hope College, Oral Roberts University, Rochester Institute of Technology, St. Mary's University, and Ursinus College) will participate in authentic research experiences in their undergraduate biochemistry lab courses. They will integrate computational (*in silico*) and wet lab (*in vitro*) techniques as they characterize proteins whose three dimensional structures are known but to which functions have not been previously ascribed. Their learning as students and their growth as scientists will be assessed in terms of research methods, visualization, biological context, and mechanism of protein function. The modules that are developed will be disseminated to the scientific community via a web site (promol.org). Over the course of the project, we will also be following changes in faculty and teaching assistant competence in two areas: effective teaching with structural biology tools and the development of skills in the area of measuring learning gains by students.
Role: PI

2011/09/01-2015/08/31
2R15GM078077-02, NIH
Craig, Paul (PI)
Algorithmic assignment of probable function to proteins of previously unknown function
Objectives and Specific Aims: The goal of this project is to extend and apply algorithms that show promise in assigning a probable function for PDB entries of currently unknown function. This should contribute to deriving benefit from the Protein Structure Initiative by "help[ing] researchers illuminate structure-function relationships and thus formulate better hypotheses and design better experiments." *Research Design and Methods*: New protein structures are being determined at a rate faster than their biological function can be assigned. There are currently 2939 entries in the Protein Data Bank with the classification "Unknown Function". We have developed a plugin for the PyMOL molecular graphics environment called ProMOL that relies on the geometric relationships conserved in enzyme catalytic sites. Motifs in ProMOL were created from the active site specifications found in the Catalytic Site Atlas (CSA) (http://www.ebi.ac.uk/thornton-srv/databases/CSA/). Our approach explicitly searches for CSA- defined catalytic site residues according to specific atomic geometry, similar in concept to the CSA JESS templates. This dispenses with the need to filter out confounding elements such as conserved folding domains or ligand binding regions. Extensive testing of structural files from the serine protease and peroxidase families confirmed that the geometric relationships of catalytic residues alone are effective and sufficient for function prediction in protein structures. In addition to extensive characterization of serine proteases and peroxidases, we also performed a preliminary study of 39 PDB entries classified as "Structural Genomics, Unknown Function" using the Motif Finder in ProMOL, which contains 22 "native" ProMOL motifs, along with the corresponding CSA JESS C1C2 motifs and CSA Functional Atom motifs. Of the 39 entries studied, 26 (67%) yielded prediction values of 1 (exact match to an existing template). An active site lacking one residue or containing an extra (outlier) residue was identified for 36 (92%) of the structures. No match was reported in only three of the test cases. We will extend the number of motifs in ProMOL's Motif Finder, using newly created ProMOL motifs and existing JESS motifs to include representatives from the most prominent protein families, increase automation of the process and then evaluate all PDB entries described as having "unknown function". Entries that show positive correlation will then be further explored using sequence and structure alignment tools. Both software and results will be openly released to the community.
Role: PI

2014/08/08-2015/08/31
3R15GM078077-02S3, NIH
Craig, Paul (PI)
Algorithmic assignment of probable function to proteins of previously unknown function
Supplement to NIH Award 2R15GM078077-02
Role: PI

Completed Research Support

2011/11/14-2014/08/31
3R15GM078077-02S1, NIH
Craig, Paul (PI)
Algorithmic assignment of probable function to proteins of previously unknown function
Supplement to award 2R15GM078077-02
Role: PI

2013/12/19-2014/08/31
3R15GM078077-02S2, NIH
Craig, Paul (PI)
Algorithmic assignment of probable function to proteins of previously unknown function
Supplement to NIH Award 2R15GM078077-02
Role: PI

2006/07/01-2010/06/30
1R15GM078077-01, NIH
Craig, Paul (PI)
AREA: Structural Biology Extensible Visualization Scripting Language
The goal of the SBEVSL project is to create a new extensible scripting language for molecular graphics, as used in structural biology, by combining the intuitive expressive power of the scripting language created by Roger Sayle for RasMol with the general object-oriented extensibility of the Python scripting of PyMOL. Major existing open source molecular graphics programs, including RasMol, Jmol and PyMol will be adapted to accept scripts written in the new scripting language. Rather than impose a language on any program, the SBEVSL project will extract all the concepts used in the command languages of major molecular graphics programs and gather them in one master ontology. Relevance: This unification of an essential component of the infrastructure used in understanding and communicating the structure and function of biologically significant molecules will increase the efficiency of many activities in structural biology, such as drug design. Time now being lost in the struggle to move descriptions of molecular renderings among such programs as RasMol, Jmol, PyMOL, Molscript and Raster3D will be available for more productive activities.
Role: PI

2009/04/27-2010/06/30
3R15GM078077-01S1, NIH
Paul Craig (PI)
SBEVSL - Structural Biology Extensible Visualization Scripting Language
Supplement to NIH Award 1R15GM078077-01
Role: PI

2004/02/01-2007/01/31
DUE-0402408, NSF
Craig, Paul (PI)
Building a Cross-Institutional Collaboratory for 3D Visualization in Technical Education and Training
The NSF ATE program supported a joint program between Brookhaven National Lab and several community colleges and PUIs. The focus was on using 3D visualization for training the workforce. Projects included architecture, biochemistry, and engineering.
Role: PI