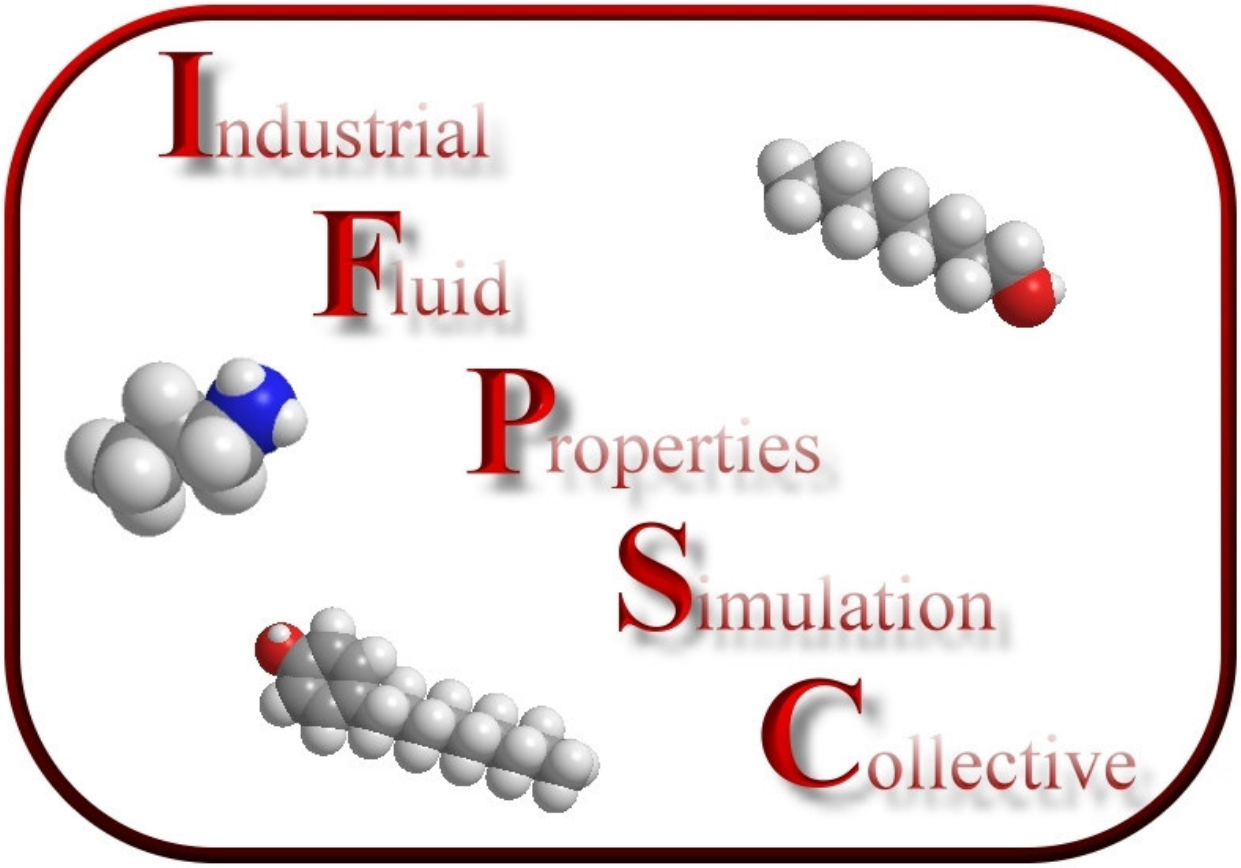

Please provide feedback online at:

http://fluidproperties.org
or
http://ifpsc.org

Industrial Fluid Properties Simulation Collective (IFPSC) Vision and Strategic Plan

VISION STATEMENT

A robust, accurate, and easy-to-use set of modeling tools will be widely available for the prediction of physical properties of fluids and obtaining insight into the connections between molecular structure and properties. As a part of this tool set, molecular simulation will become a breakthrough technology that is widely accepted in the chemical industry and applied in conjunction with other predictive methods to meet the industry's evolving fluid property data needs. Through an international collaboration (IFPSC) between industry, academia, and national labs (coordinated by the National Institute of Standards and Technology), we will develop Standard Reference Simulations, validation of methods, quantification of uncertainty, force field and simulation databases, communication standards between computer programs, and recommendations regarding the use of other predictive methods, thus enabling users to select the appropriate tools to achieve results with requisite accuracy and insight.

THE TEAM/STEERING COMMITTEE

Thomas Allison (NIST), John Brennan (Army Research Lab), Fiona Case (Case Scientific), Anne Chaka (NIST), Kerwin Dobbs (DuPont), Daniel Friend (NIST), Dave Frurip (Dow), Peter Gordon (ExxonMobil), Russ Johnson (NIST), Jonathan Moore (Dow), Ray Mountain (NIST), Jim Olson (Dow), Rick Ross (3M), Martin Schiller (DuPont), Vince Shen (NIST)

SUMMARY OF DELIVERABLES

The following deliverables are proposed:

- A series of Standard Benchmark Reference Simulation examples with model protocols to illustrate techniques for both expert developers and novice users to test and develop their codes.
- A series of test problems for developers to gauge transferability and how well force fields capture essential chemistry and physics to predict properties with the degree of accuracy required by industry using generic established simulation methods.
- An online repository for simulation codes and simulation-related subroutines.
- An online force field repository whereby forcefield developers submit their force fields to the repository in a flexible, extensible, generic format; forcefields in the repository are then available and can be retrieved in a file format appropriate for several of the most popular simulation codes.
- An online database of simulation and related experimental results.
- A periodic challenge (contest) to test and stimulate development of methods and force fields.
- A flow chart/web-based tool for recommended methods (simulation and non-simulation) based on property, material and conditions.
- A set of standard criteria for journal articles that report simulation results and methods to enable archiving, quantification of uncertainty, reproducibility, and maximize scientific leverage.

TABLE OF CONTENTS

BACKGROUND AND PROBLEM STATEMENT

The members of the organizing committee and many others [1,2,3] view molecular simulation as a very promising research tool for increasing research productivity and providing knowledge to guide strategic decision making in an industrial setting. By bridging the gap between experiment and theory, molecular simulation provides an unambiguous means of testing theoretical assumptions and leads to a better understanding of microscopic structure and transport mechanisms. The reward for incorporating molecular simulation in the industrial research process is a substantial improvement in the acquisition of quantifiable, accurate, and pertinent technical information that will allow the creation and design of new products to meet specific marketplace demands. In addition, molecular simulations challenge physical assumptions and familiar ways of thinking about chemical processes. This challenging environment stimulates the flow of ideas between experimentalists and modelers, leading to the development of new insights and new conceptual models. However, the potential benefits of molecular simulation have not been fully realized in the chemical industry despite several decades of development within the scientific community.

---

[1] "Vison 2020. Chemical Industry of the Future. Technology Roadmap for Computational Chemistry." (1999). http://www.chemicalvision2020.org/pdfs/compchem.pdf

[2] P. R. Westmoreland, P. A. Kollman, A. M. Chaka, P. T. Cummings, K. Morokuma, M. Neurock, E. B. Stechel and P. Vashishta. Applying Molecular and Materials Modeling: An International Comparative Study (Kluwer Academic, New York, 2002). http://www.wtec.org/loyola/molmodel/mm_final.pdf

[3] "Computational Science: Ensuring America's Competitiveness." President's Information Technology Advisory Committee (2005). http://www.nitrd.gov/pitac/reports/20050609_computational/computational.pdf

Many circumstances contribute to this situation. Molecular simulation is still very demanding in terms of the required expertise and computational resources. Uncertainty exists about the level of accuracy that can be routinely expected from molecular simulation for the range of properties and chemistries of importance to industry. Most current industrial problems of interest require simultaneous consideration of multiple components, multiple phases, multiple relevant length and time scales, and/or multiple properties. Though numerous force fields are available in the literature, an industrial scientist will most likely be unable to find all the necessary force field parameters for the system of interest, even if the requirement that parameters provide sufficient quantitative accuracy is removed for the problem at hand. Force field development is recognized as being important among modelers, but convincing funding agencies to support such a development effort is a daunting task. Even though a multitude of potentially useful molecular simulation technologies exist and are being developed within the scientific community, the timely transfer of these technologies to industry is sorely lacking. "New" methods are incorporated in commercial software but usually many years (>5) after they were first introduced.

In recent years, the increasing interests and opportunities in the design of materials at the molecular (nano) level are resulting in an urgent need for improving the capabilities of molecular simulation methods. Unfortunately, the simulation technologies are currently scattered across a number of different fields, the simulation codes themselves lack a high degree of interoperability, and the software development is not properly coordinated with practical needs in mind. In the context of our attempts to address these issues, we are focusing on fluid properties because they are sensitive gauges of molecular simulation accuracy.

VISION ELEMENTS

SCIENCE

Force Fields

Description

Our vision is that force fields will be readily available in a standardized format, well-studied and characterized, and well-understood for property or performance prediction.

Team members

Fiona Case, Anne Chaka*, Peter Gordon, Jonathan Moore

3-year objectives

Year 1

1. Educate ourselves regarding other previous and on-going efforts that are similar to ours.
2. Identify and obtain buy-in to strategic plan by key stakeholders.

3. Develop a test set of properties and molecules/systems to bracket range of nonbonded interactions, i.e. increasingly polar, polarizable, hydrogen bonding donors and acceptors, ionic systems, increasing electrolyte concentration, pH, etc. and increasingly complex combinations of functional groups within the molecule.
4. Choose a test molecule on which to run the entire range of methods - from the quantum to classical to continuum to empirical - to serve as a benchmark and reference and to serve as a model to design an interface framework and proof-of-concept demonstration.
5. Delineate all of the characteristics required to specify a force field (e.g., functional forms, combining rules, treatment of electrostatics, etc.). Identify/define a standard reference format for force fields to encourage formation of a publicly-accessible force field repository.
6. Complete a high-level design of interface to database and [reference] simulations, and delineate the type of tools needed to enable users and developers to test and assess reliability and transferability of force fields and interpret results.
7. Offer a periodic challenge in some form to test methods and stimulate development of new force fields.

Year 2

1. Create a database of force fields, populating for each one the characteristics identified in Year 1.
2. Identify and publicize industrially-relevant classes of chemistries requiring improved descriptions and different formalisms and potential functions and make these assessments available via the web. Develop a sicklist and archive of known problems.
3. Develop a series of test problems for developers to gauge transferability and how well force fields capture essential chemistry and physics to predict properties with the degree of accuracy required by industry using generic established simulation methods.

Year 3

1. Develop an online forcefield repository/tool whereby forcefield developers submit their force fields to the repository in a flexible, extensible, generic format; forcefields in the repository are then available and can be retrieved in a file format appropriate for several of the most popular simulation codes. Include a mechanism for user comments or ratings, as well as an indication of how many times each force field has been cited in the literature.

Simulation Methods

Description

Our vision is that simulation codes will be modular and compatible in open source or

Standard Reference Simulation (SRS) Framework,* efficient and robust, and easy to use. Furthermore, tools will be available for complete workflow from problem set up, through convergence, to analysis for a broad range of problems. In addition, simulation and relevant experimental results will be stored in a shared database accessible to program developers and researchers to be used for benchmarking, developing code, and learning both modeling protocols and additional applications (*SRS can include commercial and proprietary code linked by standard input/output formats.)

Team members

Tom Allison, Anne Chaka, Kerwin Dobbs, Peter Gordon, Jonathan Moore, Ray Mountain, Rick Ross*, Vincent Shen

3-year objectives

Year 1

1. Educate ourselves regarding other previous and on-going efforts that are similar to ours.
2. Identify and obtain buy-in to strategic plan by key stakeholders.
3. Identify commercially- and publicly-available resources; provide links and test cases.
4. Identify properties and classes of molecules and systems that are routine.
5. Identify gaps in workflow tools and publicize.
6. Identify standards for input and output, and identify which software modules are currently available to enable interoperability between codes/steps.

Year 2

1. Establish the necessary characteristics of a graphical user interface (GUI) for setting up, launching, and monitoring a simulation as well as for the analysis of the end results.
2. Identify a problem-oriented simulation language (GUI or text line editing) for tying simulation tasks together to solve a problem.
3. Determine how to establish error bars for calculations.
4. Delineate all steps and criteria required to predict properties to a specific level of accuracy within established error bars for a basic set of simulation tasks which are readily amenable to code modularization.
5. Develop a series of Standard Benchmark Reference Simulation examples with model protocols to illustrate techniques for both expert developers and novice users to test and develop their codes. In so doing, enable the accurate comparison of the results from different codes on a systematic basis via a well-defined protocol. Include a set of coordinate files for a variety of specific systems along with a complete listing of the numerical values of each contribution to the potential energy for a given force field (non-bonded, angles, bonds, torsions, electrostatics) for a variety of systems spanning very simple (Lennard-Jones) to

more complex (proteins) for use in validating methods/codes for the calculation of potential energy/forces. Include the consideration of quantum-chemical-based methods including criteria to establish when they have been improved to the extent necessary to sufficiently reproduce non-bonded interactions for fluid simulations.

Year 3

1. Establish a repository of short, explanatory articles about methods and algorithms. Each article should focus on a particular algorithm, contain a "pseudo-code" section which describes its steps in plain terms, and highlight the key papers from the literature which provide further information.
2. Establish a repository for simulation codes and simulation-related subroutines (analysis routines, property calculation routines, etc); Establish curatorship protocols for accepting and storing routines; Educate simulation users on the benefits of sharing codes and subroutines; Encourage simulation users to adopt standards for facilitating straightforward integration.
3. Establish a database of simulation and related experimental results. Delineate and develop standards for storing data; Evaluate and recommend use of centralized or distributed databases or a combination of both; Establish curatorship protocols for accepting and integrating data; Educate simulation users on the benefits of sharing data and encourage them to adopt standards for facilitating automated data capture and integration.
4. Develop a primer on writing good molecular simulation routines, a tutorial including guides regarding topics such as the best way to parameterize molecular variables, subdivide tasks, speed performance, and enhance portability from one problem to another.
5. Offer a periodic challenge to test methods and stimulate development of new methods.

Benchmark and Non-simulation Predictive Methods

Description

Our vision is that prediction methods will be an important part of the industrial toolbox for property prediction and will be used in conjunction with Simulation methods to solve important industrial problems. Innovative predictive methods must be encouraged and established methods improved. Simulation methods will be routinely and effectively used for validation of these predictive methods.

Definition of non-simulation methods: any method which does not neatly fit into the definition of molecular simulation (MS): *Molecular simulation refers to any method that generates a trajectory through phase space for the system of interest by executing a series of deterministic and/or stochastic steps that obey the thermodynamic constraints.* (Ilja Siepmann, 2006)

Team members
Tom Allison, Fiona Case, Dan Friend, Dave Frurip, Jim Olson, Martin Schiller*.

3-year objectives

Year 1

1. Obtain buy-in to strategic plan by key stakeholders.
2. Identify publicly available resources (commercial and academic); provide links and test cases.
3. Explore the possibility of developing a flow chart/web tool for best methods based on property, material and conditions.
4. Develop a process to provide benchmark data for periodic Challenges.
5. Offer a periodic challenge in some form to test methods and stimulate development of new methods.

Year 2

1. Develop a flow chart/web tool for best methods based on property, material and conditions.

Literature

Description
Our vision is that journal publications will report simulation results and methods according to established criteria to enable archiving, reproducibility, and maximize scientific leverage. These criteria will be established by a partnership of journal editors, developers, and industrial experts.

Team members
Fiona Case, Dan Friend, Jonathan Moore, Ray Mountain*

3-year objectives

Year 1

1. Identify key journals.
2. Obtain buy-in to strategic plan by key journal editors.
3. Develop and vet criteria for presenting results and methods with users and developers of simulations and associated methods.
4. Develop an agreement with journals to publicize validation criteria and to implement a policy for authors to adhere to these criteria prior to acceptance and publication of relevant articles in these journals. *J. Chem. Phys.* **108**, 6109 (1998) contains examples of validation criteria.

Year 2

- Develop a repository/database of refereed simulation results
  - Define more carefully what types of simulation results to be incorporated and criteria to judge validity of results
  - http://www.biosimgrid.org/ is an example of a simulation database
  - Provide framework for researchers to submit results (check with Tom Allison who is doing this for kinetics)

Years 3

- Monitor acceptance of criteria

STAKEHOLDERS

Industrial Parties

Description:
Our vision is that industry stakeholders will be well organized with international representation and recognized as a force for change. They will effectively partner with all stakeholders to ensure the success (commercial if appropriate) of effective computational tools. They will offer a periodic challenge in some form to test methods and stimulate development of new methods. Industry will communicate needs, priorities, sample problems, and prediction quality criteria and offers incentives to drive academic development.

Developers

Description:
Our vision is that developers – academic, government, and commercial – will be "sensitive" to industrial needs and routinely seek direct input about them. They will embrace the concept of the Industrial Fluid Properties Simulation Challenge or other comparable methods to compare results for different methods and force fields. Their developments - whether methods, force fields, or algorithms – will be made available and structured in such a way that their validation and practical use are facilitated.

Users

Description:
Our vision is that users of molecular simulation, whether expert or non-expert, will have sufficient information and tools readily available to enable the user to select and apply appropriate simulation methods for a given problem to achieve results with requisite accuracy on an acceptable time scale.

NIST

Description:
Our vision is that NIST will be the center/focus of the activities described in this document, will lead this effort with a program that is well funded and well staffed, and will coordinate appropriate databases of methods, force fields, and critical and objective reviews. NIST will delineate the essential physics that models need to capture for classes of chemistry and properties and develop sets of test problems (including obtaining experimental data and performing benchmark calculations). NIST will broadly communicate industrial needs and priorities. It will archive results and evaluations to capture what is learned and to leverage scientific impact. It will coordinate activities with other national labs and technical societies such as the American Chemical Society, American Institute of Chemical Engineers, and American Physical Society.

Federal Funding Agencies

Description:
Our vision is that federal funding agencies will be "sensitive" to industrial needs and support the development of methods, force fields, and infrastructure in alignment with these needs.

Other Organizations (DIPPR®, DECHEMA, AIChE, ACS, etc.)

Description:
Our vision is that all categories of industrial, academic, and governmental stakeholders will work effectively together to systematically drive improvement of the science of molecular simulation. Organizations such as DIPPR(R), DECHEMA, etc. will be well-informed about the activities described here and will be invited to actively participate.