CCP-SAS Project

A Collaborative Computational Project for Small Angle Scattering

Grant No.
SI2/CHE-1265821

An SI² cyberinfrastructure project addressing Grand Challenges in the Chemical Sciences

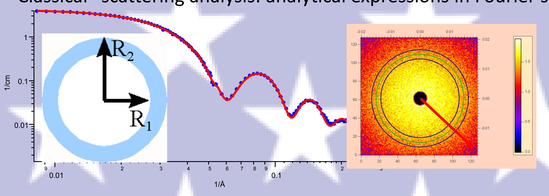
- **Paul Butler (PI)**
  University of Tennessee Knoxville
  NIST Center for Neutron Research
- Emre Brookes
  University of Texas Health Science Center San Antonio
- Jianhan Chen
  Kansas State University
- Joseph Curtis
  NIST Center for Neutron Research
- Tom Irving
  Advanced Photon Source

+ other collaborators

KICKOFF WORKSHOP: Feb 7-9, 2014

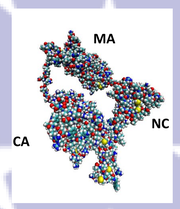
BACKGROUND

"Classical" scattering analysis: analytical expressions in Fourier space



Can get quite computer intensive with 2D oriented scattering including advanced sampling of error surface in parameter space etc.

Functionality provided by sasview.org
- Currently maintained by 5 facilities
- Code camp held 31 Mar-6 Apr 2014



BUT …. what about the increasingly complex systems that have little symmetry and where the possible variations could not be captured by a single analytical model with parameters?

➔ Real space generation of model candidates and FFT into scattering space, compare and iterate.

√ Pioneered by Dmitri Svergun distributed through ATSAS suite

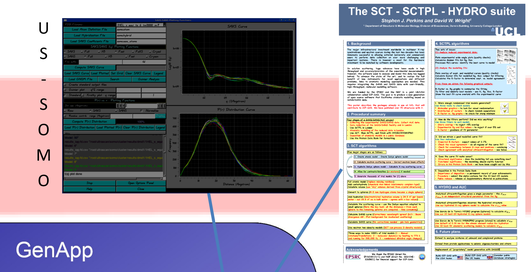
CCP-SAS in a NUTSHELL

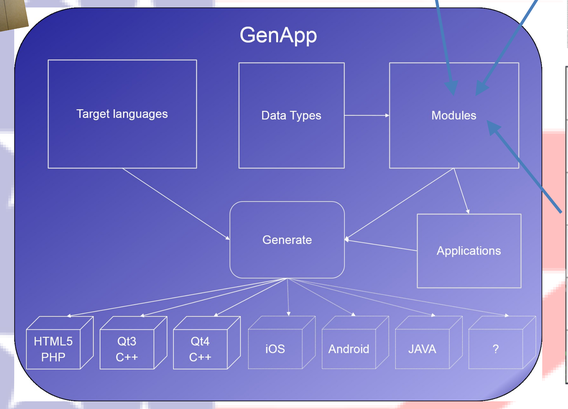
Create new and enable existing computational tools to model scattering data in real space to dramatically improve accessibility by non-experts.

Our approach: Generate ensembles of possible structures using as much a priori information as possible with high throughput computing methods to screen for reasonable structures that match the data.

A priori information = constraints on system.
- Connectivity imposes constraints
- MD/MC use the physical chemistry/chemical physics of the system in order to provide a representative sampling of phase space
- Data from other experimental methods such as AUC, NMR, etc. also provide constraints

- Include infrastructure to
  - Provide as transparent an access to "HPC" resources as possible
  - Provide as transparent an access to advanced techniques and algorithms as possible (building experience into software)
  - Allow "simple" plugging in of new modules that add new tools
- Continue to improve and extend modelling tools
  - Extend existing modeling of protein solution scattering to larger classes of problems
  - Adding new MD and MC sampling techniques
- Fully open source which encourages community contribution
- Long term support and maintenance





SASSIE

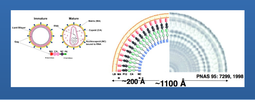
YEAR 1 GOALS

1. Core Software:
   - Deploy web prototype & begin alpha testing w/ grant members
   - Preliminary design HPC (core & gateway) [access & usage]
   - Publish APIs for web framework and SASMOL [grow developer community]
2. Chemical Physics
   - Implementation of an interface to CHARMM and a torsional angle molecular dynamics (TAMD) module
   - Test of various atomistic implicit solvent models and simulation protocols for proteins using model systems
   - Initiate testing of sampling protocols and force field options for nucleic acids and glycosylated proteins.
   - Identify best non protein target problems to address
3. Testing:
   - Identify candidate test projects appropriate for the state of the software and kickoff first testing project
   - Identify technologies/build infrastructure to: track project queue and status, simplify feedback loop and documentation effort
4. Dissemination
   - Begin engagement activities
   - Identify technologies and build infrastructures (capturing user feedback, providing video tutorials, user mailing lists, FAQS, etc.)

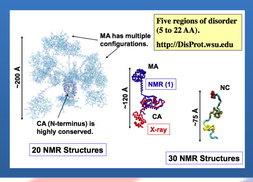
APPLICATIONS: EXAMPLES FROM SASSIE

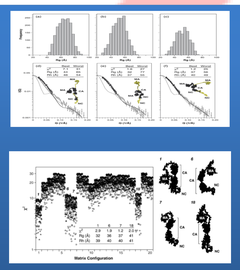
HIV-1 Gag

Gag is the main protein component in HIV-1. Once assembled, it is known to form a regular structure. The protein is composed of several domains separated by disordered linkers.

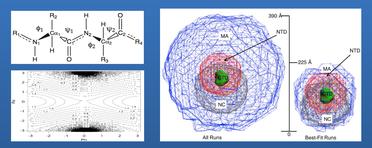


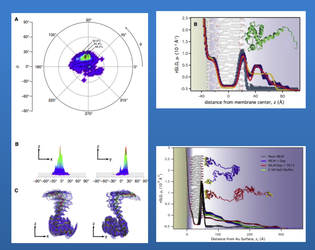
Complete models can be created using existing NMR and X-ray coordinates and scattering profiles compared to data.





More thorough sampling of inter-domain configurations can be obtained by generating structures using atomistic models with flexibility dictated by CHARMM force-field parameters. HIV-1 Gag is compact in solution.





Further studies defined the orientation of the Gag matrix domain on membrane surfaces and that nucleic acid causes compact Gag to extend on the membrane surface. Thus providing step-wise insight into the assembly of HIV-1.

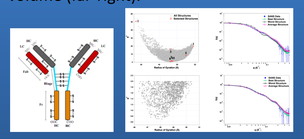
w/ H. Nanda et al.

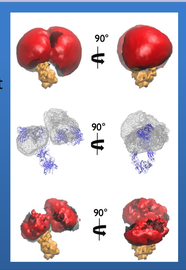
J. Mol. Biol. **365**, 812 (2007), Biophys. J. **99** (2010), J. Mol. Biol. **406** (2011), Comp. Phys. Comm. **183** , 382 (2012).

Monoclonal Antibodies

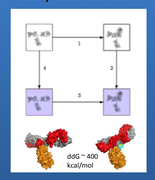
Manufacturing of therapeutic monoclonal antibodies (mAb) supports a > $40B/yr global market. mAb proteins are flexible and thus do not adopt a single structure in solution. mAb products often have low specific activity thus requiring formulation at high concentration. Computational methods to analyze scattering data of mAbs is of increasing utility.

Using the SASSIE Monte Carlo module ensembles of mAb can be created and theoretical profiles can be compared to experimental data (below) with the best fitting ensemble covering a reduced volume (far-right).

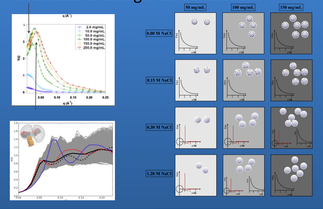




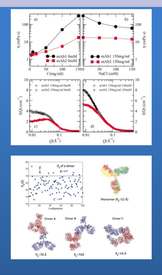
Solvation free-energy analysis indicates most-likely structures



Ensemble modeling allows for precise colloidal representation of mAb interactions at high concentration.



At high concentration mAbs can have wide range of intrinsic viscosities. SASSIE was used to determine energetically viable dimers to understand bulk properties.



Additional studies to explore mAb isoforms, excipient effects, liquid-liquid phase separation, and solid phase behavior are ongoing. Further enhancements to SASSIE to model scattering data are under development.

AMGEN Genentech *A Member of the Roche Group*

w/ N. Clark et al, and Y. Liu et al.

J.Phys. Chem B **117**, 14029 (2013) and Biophys. J. (accepted).

EPSRC

Collaborative Computational Project for advanced analyses of structural data in chemical biology and soft condensed matter

Grant No.
EP/K039121/1

- **Stephen Perkins (PI)**
  University College London
- David Barlow
  Kings College London
- Karen Edler
  University of Bath
- Richard Heenan & Steve King
  ISIS Pulsed Neutron & Muon Source
- David Scott
  Nottingham University
- Nick Terrill
  Diamond Light Source

+ other collaborators

STRUCTURE

- The Executive team is responsible for co-ordinating activities, tracking overall progress, organizing meetings and generally trying to make sure nothing falls through the cracks.
- The core software team team is responsible for developing and deploying the core plugin framework, web interface, and for porting existing code bases into the framework. Chair Joseph Curtis.
- The Chemical physics team is responsible for developing new algorithms and new plugins. Chair: Jianhan Chen
- Testing team is responsible for testing the software with real world applications, feeding back new ideas, usability and bug reports in a continuous cycle. Experience is also used to commence documentation Chair: Steve Perkins
- Dissemination is responsible for promotion of the project and its activities , education of the community, and engagement with other projects and other facilities. Chair: Steve King.

STATUS

Organizational
- Web presence established with ccpsas domain registered
- Joined CCP steering panel
- Several papers published or in press more to come.
- Talks: NIBB, ACA(2), ACNS, XSEDE
- Started engaging other facilities
- Working groups created team documents describing scope, vision, and milestone for each area.
- Fortnightly Executive team GoTo Meetings
- First full project GoTo Webinar meeting in December (plan monthly)

Software development
- GENAPP framework prototype created and tested with SASSIE modules
- SCT-SCTPL scattering curve simulation software made open source
- Web prototype implemented: roll out March 2014 to begin alpha testing w/ grant members
- Preliminary design HPC (core & gateway) [access & usage]
- Publish APIs for web framework and SASMOL [grow developer community]
- Glycoprotein Builder Prototype
- CHARMM interface implemented
- New modules:
  - Contrast Calculator released and published
  - SLD-MOL released and submitted[reflectivity of ensembles on/in surfaces]
  - SASCALC prototype
  - TAMD prototype

http://www.ccpsas.org