Data Deposition, Metadata, ISO Standards, Calibration Standards, Publication Standards for Small-Angle Scattering

– a personal perspective on some ongoing developments

Andrew Allen
Materials Measurement Science Division
Material Measurement Laboratory,
NIST, Gaithersburg, MD, USA
andrew.allen@nist.gov

National Institute of Standards and Technology
U.S. Department of Commerce

United Nations Educational, Scientific and Cultural Organization

2014 international year of crystallography

International Union of Crystallography

IUCr Commission on Small-Angle Scattering 2014-2017

Continuing Members and Members from 2014:

Jill Trewhella (Australia, 2005, Chair)	jill.trewhella@sydney.edu.au
David Babonneau (France, 2011)	david.babonneau@univ-poitiers.fr
Pete Jemian (USA, 2008)	jemian@anl.gov
U-Ser Jeng (Taiwan, 2011)	usjeng@nsrrc.org.tw
Ritva Serimaa (Finland, 2008)	ritva.serimaa@helsinki.fi
Vladimir Volkov (Russia, 2008)	vvo@ns.crys.ras.ru
Toshiji Kanaya (Japan)	kanaya@scl.kyoto-u.ac.jp
Duncan McGillivray (New Zealand)	d.mcgillivray@auckland.ac.nz

Consultants:

Andrew Allen (USA, Immediate Past Chair)	andrew.allen@nist.gov
Elliot Gilbert (Australia)	epg@ansto.gov.au
Naota Yagi (Japan, Retiring Member)	yagi@spring8.or.jp
Yoshiyuki Amemiya (Japan)	amemiya@k.u-tokyo.ac.jp
Gernot Kostorz (Switzerland)	gk-iucr@ethz.ch
Jan Skov Pedersen (Denmark)	jsp@chem.au.dk
Dmitri Svergun (Germany)	dmitri.svergun@embl-hamburg.de
Iris Torriani (Brazil)	torriani@ifi.unicamp.br

All canSAS efforts of interest to, and encouraged by, SAS Commission!

International Year of Crystallography (IYCr 2014):

(1) IYCr presence at international symposia, and a record number of SAS microsymposia at 2014 IUCr Congress in Montreal.

(2) Chapter on "Disordered and Heterogeneous Materials" for new International Crystallography Table H – *out November 2015.*

(3) Guidelines issued in open letter to major research journal editors regarding publication standards for Biological SAS papers.

(4) SAS Commission web-page revised and linked with "SAS Portal".

(5) Online SAS2012 Special Issue published February 2014 in the *Journal of Applied Crystallography*. Similar plan for SAS 2015.

(6) NIST SAXS intensity calibration reference material in 2015.

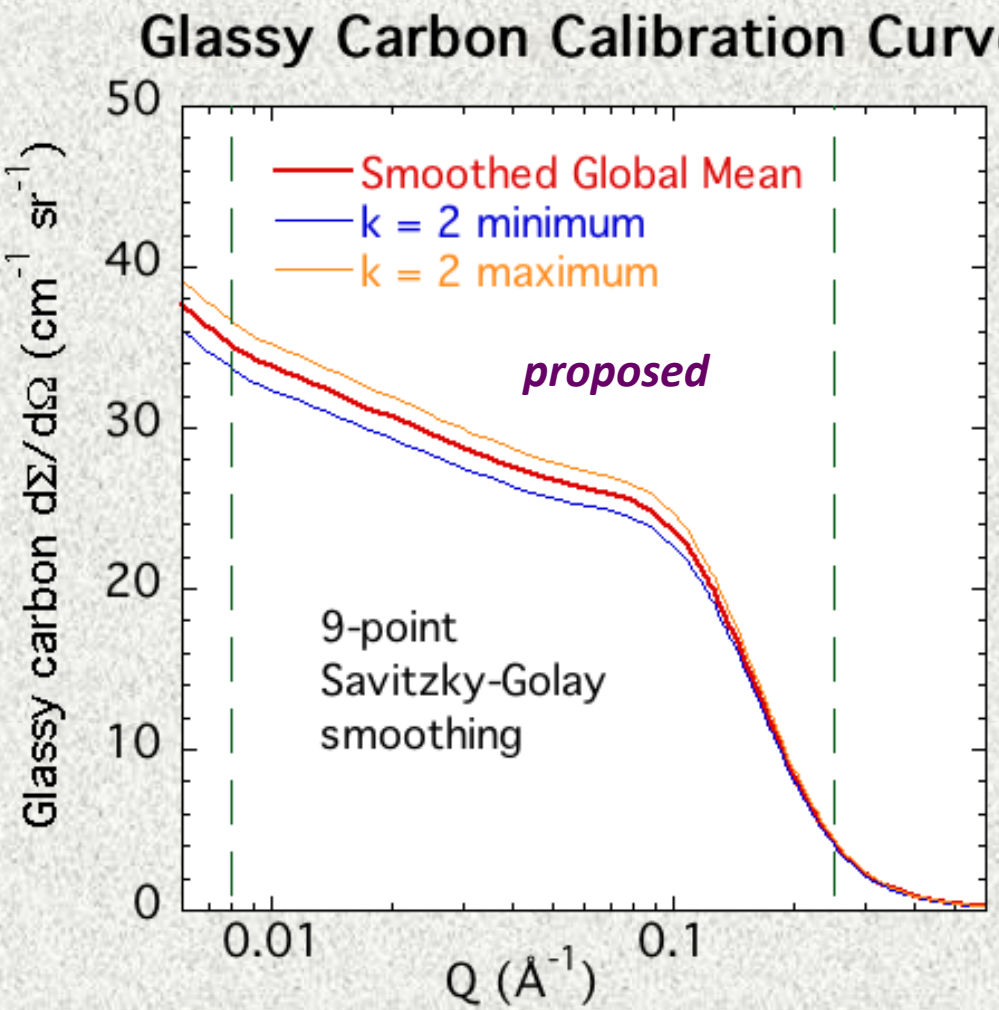
(7) ISO standard on SAXS-based particle size in 2015.

SAS Derived Envelopes for Biomolecules in PDB

- **World-wide Protein Data Bank (wwPDB) Small-Angle Scattering Task Force (SAStf) –**
  (http://www.wwpdb.org/workshop/sasTaskForce.html).
- **Draft publication guidelines: structural modeling of SAS data from biomolecules in solution –**
  (http://journals.iucr.org/d/issues/2012/02/00/me0456/index.html).
  Also see: *Acta Cryst. D* (Jacques *et al*., 2012) and:
  http://journals.iucr.org/d/issues/2012/06/00/be5200/index.html.
- **Letter sent on behalf of CSAS to editors of relevant journals indicating publication guidelines –**
  as set out in: *Structure* (Trewhella *et al*., 2013, and at:
  http://www.cell.com/structure/abstract/S0969-2126%2813%2900150-0.

SAS derived envelopes allowed in PDB if requirements met.

NIST Standard Reference Material
SRM 3600 – *Draft*



SAXS intensity calibration standard:

- Primary calibration using APS USAXS – *with Jan Ilavsky*;
- Based on glassy carbon;
- Certified for SAXS in *Q*-range from 0.008 $Å^{-1}$ to 0.25 $Å^{-1}$;
- Validated with SANS at NCNR;
- Calibration curve will be provided with (k = 2) 95 % confidence uncertainty bands;
- Primarily for industry.

International Organization for Standardization (ISO) Best Practice Standard for SAXS

ISO: Technical Committee TC 24:
Particle characterization including sieving:
Subcommittee SC 4: *Particle characterization*

ISO/FDIS 17867:
Particle size analysis — Small-angle X-ray scattering

– initiated by Alan Rawle (Malvern Instr.), who taught himself SAS in process!

– describes best practice for SAXS measurements and data analysis for SAXS determination of nanoparticle size.

Developed over several years and released in 2015, this is a basic best practice standard for SAXS-based determination of particle size, and is the first ISO standard for small-angle scattering. It is primarily aimed at industry.

Editorial board of *Journal of Applied Crystallography*

Data deposition and validation issues are beginning to cross our desks!
e.g., IUCr Diffraction Data Deposition Working Group (DDDWG)

Some Data Issues in Our Future

- U.S. Government funding agencies (NSF, DOE, etc.) are requiring data management plans as condition for funded research.
- Data must be safeguarded & may ultimately be made available.
- This is becoming an international trend. ***What does it mean?***

- ❑ Raw data (instrument-dependent) are not very useful without a lot of explanation of how data were obtained.
- ❑ Reduced data (instrument-independent) are potentially much more useful, together with modeled / fitted data ***– but guidance needed!***
- ❑ Different researchers may ultimately get results and publish using previous researchers' data. Credit to be divided between those obtaining data, and those developing new results from data.
- ❑ ***Especially for a field like SAS, metadata will be critical!***
  – IUCr Executive asking Commissions for metadata requirements.

Some Proposed Metadata Needs for SAS

– A.J. Allen with Fan Zhang and others, 2015

Common contents:

- Descriptive title of experiment
- Names of experimenters
- Instrument name and location
- Type of source:
  X-ray tube, rotating anode or synchrotron; neutron reactor or pulsed source
- X-ray energy/wavelength or neutron wavelength
- Estimated wavelength spread
- Types of attenuator and amount of attenuation for transmissions etc.
- Sample identifier, nature of sample, preparation, history, etc.
- Sample thickness
- Sample chamber and sampe environment (temperature, pressure, pH, etc.)
- Measurement duration and time stamp (start, end, etc.)
- File location
- Data reduction package
- Data analysis / modeling package

Proposed Metadata Needs for SAS (continued)

Pinhole-collimated SAXS and SANS:

- Horizontal & vertical beam defining slit apertures (or diameter)
- Horizontal & vertical guard slit apertures (or diameter)
- Source-to-sample distance collimation details
- Type, number, thicknesses of X-ray or neutron windows in beam path
- Quality of vacuum in the beam path; amount of ambient beam path
- Type of 2D detector (& relevant set-up parameters, readout mode, etc.)
- Detector horizontal & vertical pixel sizes and binning
- Sample-to- detector distance & collimation details
- Detector tilt angle & offset from incident beam path
- Horizontal & vertical beam center coordinates (in pixels)
- Polarization correction factor
- *Q* calibration & intensity calibration standard methods & mode
- Incident & transmitted beam intensities, sample transmission
- SAS measurement / exposure / image acquisition times
- Location of flat field / detector sensitivity 2D correction file
- Location of dark field / electronic background 2D correction file
- Location of 2D detector mask file applied

Proposed Metadata Needs for SAS (continued)

Additional for Bonse-Hart USAXS and USANS:

- 1D or 2D collimation (for USAXS) & pre-monochromator (for USANS)
- Incident beam collimating monochromator crystal type & reflection
- Incident beam number of collimating monochromator reflections
- Incident beam transverse collimating crystal, reflection, number (if any)
- Scattered beam transvers analyzer crystal, reflection, number (if any)
- Scattered beam analyzer crystal type & reflection
- Scattered beam number of analyzer reflections
- Darwin curve FWHM width in $Å^{-1}$ or nm^{-1} (Q units) & Q resolution
- Primary beam intensity at $Q = 0$ without / with sample; sample transmission
- Pre-sample ion chamber / fission monitor details & readout
- Main USAXS or USANS detector type; readout mode, gains, backgrounds
- Main detector active area & dimensions
- Sample-to-analyzer & sample-to-detector distances; slit-length in Q (1D)
- Angular offset of scan at $Q = 0$ (especially for USANS)
- Start (negative) and end positions of scan in Q
- Number of data points, Q-steps & dwell times, or fly-scan parameters
- Other coordinated stage motions (1D or 2D USAXS)

Other Metadata Considerations

At least two other SAS configurations have additional detailed metadata requirements:

(1) SANS at pulsed neutron sources – *clearly metadata need to detail neutron time-binning and all associated effects; must be provided in a way that is transparent for deposited data subsequently used by other researchers.*

(2) GI-SAXS and NS-SANS – *details of grazing angle incident geometry & collimation, critical angle information, slit dimensions, sample length along beam, substrate details, etc., need to be included in the metadata.*

(3) Also: magnetic SANS, SE-SANS, anomalous SAXS, etc.

It is critical that all deposited data have corresponding, comprehensive metadata tightly attached:

(1) Good metadata will reduce incidence of data fraud.
(2) Good metadata should ensure deposited data are not abused (accidentally or purposefully) in subsequent use by others.
(3) Good metadata will lead naturally to improved publication standards.

Concluding Comments and Suggestion

(1) In a future world where SAS data supporting publications are deposited and available, metadata, best-practice standards, common data formats, standard reference materials, and publication standards, will become *highly-interrelated* issues.

(2) Good dedicated SAS researchers probably do not need the infrastructure emerging to support these issues, but will use it!

(3) Others who seek to employ SAS methods as a tool *will* need the infrastructure; and the whole SAS field / profession can be put on *significantly* firmer ground if things are done well.

(4) While all of the above issues need work in parallel, metadata issues are probably the most urgent for impending data deposition.

Should *canSAS* and other activities be configured into task forces charged with working / reporting at the triennial SAS Conferences?