Small-angle scattering of x-rays (or neutrons) tells us about the size and shape of macromolecules.

$Q = \frac{4\pi \sin \theta}{\lambda}$

Sample - randomly oriented particles

Scattering particle

$P(r)$: probable distribution of inter-atomic distances ($R_g$, $M$, $D_{max}$)

Shape restoration, Rigid body modeling

3Cp RNA complex, Claridge et al. (2009) J. Struct. Biol. 166, 251-262
SAS data represent a time and ensemble average of randomly oriented structures; the rotational-averaging of 3D structures yields a 1D profile.

The SAS experiment is conceptually simple, but practically demanding; both instrumentally and with respect to samples.

SAS data primarily tell us about shape; as shapes become more complex, different shapes can yield the same scattering profile.

Nonetheless, certain parameters, can be determined both accurately and precisely ($R_g$, Molecular mass, distance distributions over a wide range 5 – 1000’s Å) and 3D models developed or tested against SAS data can advance our understanding of bio-molecular structure/function relationships.
For an ensemble of identical, randomly oriented particles:

\[ I(q) = N(\Delta\rho V)^2 P(q) S(q) \]

\(N\) = particles/unit volume
\(V\) = particle volume (\(V^2\) dependence = highly sensitive to large particle contaminants, incl. aggregation)

\(\Delta\rho = \langle \rho(r) - \rho_s \rangle\) is the average contrast, the scattering density difference between the scattering particle and solvent

\(P(q)\) = form factor \(\Rightarrow\) intra-particle distances

\(S(q)\) = structure factor \(\Rightarrow\) inter-particle distances
Inter-particle distance correlations between molecules:

\[ D = \frac{2\pi}{q} \]

\[ S(q) \]

\[ P(q) \]

\[ I(q) \]

..... yields a non-unity \( S(q) \) term that is concentration dependent and impacts the lowest angle data.
Determining the size of your scattering particle; a critical check

- Place data on an absolute scale (water scattering) and use:

\[ M = I(0) \frac{N_A}{C \Delta \rho_M^2} \quad \text{where} \quad \Delta \rho_M = \Delta \rho \nu \]
\[ \nu = \text{partial specific volume} \]


- Use a known mono-disperse protein scatterer (such as lysozyme) and:

\[ M_{\text{unknown}} = \frac{I(0)_{\text{unknown}}}{I(0)_{\text{lysozyme}}} \frac{C_{\text{lysozyme}}}{C_{\text{unknown}}} M_{\text{lysozyme}} \]

Krigbaum and Kugler (1970) *Biochemistry* 9, 1216

- Use the scattering invariant

\[ Q = \int_0^\infty I(q)q^2 dq \quad \text{and} \quad V = \frac{2\pi^2}{Q} \]


AutoPorod – in ATSAS suite of programs more generally
Good quality, reliable scattering data

- Reliable scattering data are those you can demonstrate are from the particle you are interested and have been demonstrated to be free from instrumental and sample state biasing effects.

- Once obtained they provide:
  - long range distance constraints that complement NMR distance and orientational constraints and can aid in refinement of NMR structures
  - opportunities for constrained rigid body modeling of large multi-domain or multi-subunit structures/assemblies
  - the ability to characterize structures with inherent flexibility
Improving the accuracy of solution structural models

<table>
<thead>
<tr>
<th>Backbone rmsd to crystal structure, Å</th>
<th>-SAXS</th>
<th>+SAXS</th>
</tr>
</thead>
<tbody>
<tr>
<td>N-terminal domain (6-85)</td>
<td>0.63±0.05</td>
<td>0.56±0.05</td>
</tr>
<tr>
<td>C-terminal domain (94-175)</td>
<td>1.09±0.09</td>
<td>0.90±0.04</td>
</tr>
<tr>
<td>Both domains (6-85,94-175)</td>
<td>1.96±0.07</td>
<td>1.31±0.04</td>
</tr>
</tbody>
</table>

Grishaev et al. JACS 127, 16621 2005
Comparison of structures for 82 kDa Malate Synthase G from NMR-only data and joint fit of SAXS-NMR data

- NMR/SAXS refinement improves backbone rmsd values with respect to the crystal structure from 4.5 to 3.3 Å, largely due to more accurate translational positioning of domains.
- The mid-Q scattering range had most influence.

\[ \chi = 3.05 \]
\[ \chi = 1.01 \]
\[ \chi = 0.97 \]

Grishaev et al. J. Biomol. NMR 376, 95, 2008
Neutron contrast variation by hydrogen (\( ^1H \))/deuterium (\( ^2H \)) exchange adds a powerful dimension to scattering data from bio-molecular complexes.
Histidine kinase-antikinase, KinA$_2$-2D$_{Sda}$ example

The sensor histidine kinase KinA - response regulator spo0A in *Bacillus subtilis*

- Failure to initiate DNA replication
  - DNA damage
- Change in N₂ source
- Environmental signal
- Sda
- KipA
- KipI
- KinA
  - Spo0F
  - Spo0B
  - Spo0A
  - Sporulation
Our molecular actors

**KinA**
*Based on H853 Thermotoga maritima*

to sensor domains

**Sda**

**KipI**
*Pyrococcus horikoshi*

**Pro**

**CA**

**His**

**DHp**

**Trp**
HK853 based KinA model predicts the KinA X-ray scattering data. Sda is a dimer in solution. KinA$_2$ contracts upon binding 2 Sda molecules.
Sda is a trimer in solution

\[ \chi^2 = 0.85 \]

KipI dimerizes via its N-terminal domains and 2 KipI molecules bind KinA$_2$

KipI$_2$ \( R_g = 31.3 \, \text{Å}, \quad d_{max} = 100 \, \text{Å} \)

KinA$_2$ \( R_g = 29.6 \, \text{Å}, \quad d_{max} = 80 \, \text{Å} \)

KinA$_2$-2KipI \( R_g = 33.4 \, \text{Å}, \quad d_{max} = 100 \, \text{Å} \)
MONSA: 3D shape restoration for $\text{KinA}_2:2^{\text{DSda}}$
Component analysis

\[ I(Q) = \Delta \rho_1^2 I_1(Q) + \Delta \rho_2^2 I_2(Q) + \Delta \rho_1 \Delta \rho_2 I_{12}(Q) \]
Rigid-body refinement

KinA$_2$-2Sda components

KinA$_2$-2KipI

Pull down assays and Trp fluorescence show mutation of Pro$^{410}$ abolishes KipI binding to KinA but Sda can still bind.

Trp fluorescence confirms that the C-domain of KipI interacts with KinA
KipI-C domain has a cyclophilin-like structure

Hydrophobic groove

3Å crystal structure KipI-C domain

Overlay with cyclophilin B
Aromatic side chain density in the hydrophobic groove

Jacques, Langely, Jeffries et al, in review *J. Mol. Biol.* 2008
The KinA helix containing Pro$^{410}$ sits in the KipI-C domain hydrophobic groove
A possible role for *cis-trans* isomerization of Pro^{410} in tightening the helical bundle to transmit the KipI signal to the catalytic domains?

Or is the KipI cyclophilin-like domain simply a proline binder?
Sda and KipI bind at the base of the KinA dimerization phosphotransfer (DHp) domain. Sda binding does not appear to provide for a steric mechanism of inhibition. KipI interacts with that region of the DHp domain that includes the conserved Pro sidechain. Sda and KipI induce the same contraction of KinA upon binding (4 Å in *Rg*, 15 Å in *Dmax*).

DHp helical bundle is a critical conduit for signaling.
..and the relationship between the Kip I inhibitor and its regulatory binding partner Kip A

The N-terminal regulatory domains of the cardiac myosin binding protein C (cMyBP-C) influence motility.

High Ca²⁺

Low Ca²⁺

Controls
+cMyBP-C reg domains

movies courtesy of Samantha Harris, UC Davis
cMyBP-C in Muscle Contraction

- cMyBP-C plays **structural** and **regulatory** roles in striated muscle sarcomeres. However, the specific details of how it interacts with actin and myosin are unclear.
cMyBP C: a modular protein

- C0C2 N-terminal “regulatory” domains
- C0C2 C-terminal myosin binding domains

- Ig (blue) and (red) Fn modules

- 42% of clinical cases of familial hypertrophic myopathies are attributable to cMyBP-C dysfunction
SAXS data + crystal and NMR structures of individual modules show the N-terminal domains of mouse cMyBP-C form an extended structure with a defined disposition of the modules.

Mixing mono-disperse solutions of cMyBP-C with actin results in a dramatic increase in scattering signal due to the formation of a large, rod-shaped assembly.
Neutron contrast variation on actin thin-filaments with deuterated C0C2 show they bind actin and stabilize filaments.
SANS data show regulatory cMyBP-C domains (mouse) stabilise F-actin and provide a structural hypothesis for the observed Ca$^{2+}$-signal buffering effect.

Whitten et al. (2008) *PNAS* 105, 18360
SAXS data show significant species differences

Correlation between % Pro/Ala composition in the C0-C1 linker and heart rate from different organisms (Shaffer and Harris (2009) J. Muscle Res. Cell Motil. 30:303-306.)
SAXS data cannot define relative positions of human C0 and C1. NMR relaxation data show human PA$_L$ is flexible.
2D reconstruction of human C0C1-actin assembly from neutron contrast series consistent with C0 binding with a flexible and extended P/A_L

Lu et al., J. Mol. Biol. 413, 908-913, 2011
Human EM and mouse SANS comparison

Orlova, Galkin, Jeffries, Egelman and Trewhella (2011) *JMB* 412, 379-386
NMR data identify residues involved in (human) C0-actin interaction
Actin Binding Hot-spots

Lu, Kwan, Trewhella, Jeffries (2011) JMB, 413, 908-913
Shared Actin and Myosin Binding Sites

C0

SSKVK

Myosin RLC Binding

C1

Myosin ΔS2 Binding

Lu, Kwan, Trewhella, Jeffries (2011) JMB, 413, 908-913
By combining EM, Crystallography, SANS, SAXS and NMR, we show that:

- Human C0C1 interacts with actin specifically and promotes formation of regular assemblies of F-actin decorated by C0C1.
- Human C0 and C1 interact with myosin and actin using a common set of binding determinants.
- NMR and SAXS data indicate that P/A linker is flexible and can facilitate N-terminal domains spanning the interfilament distances.
- The switching could be regulated by phosphorylation of the motif.
The Motif of Human cMyBP-C is Required for its Ca$^{2+}$-Dependent Interaction with Calmodulin
Ca\textsuperscript{2+}-CaM addition to C1C2 blue-shifts the Trp emission peak from $\lambda_{\text{max}}$ ~345 to ~336 nm; and the intensity decreases.

Ca\textsuperscript{2+}-CaM addition to \textsuperscript{15}N-C1C2 results in significant intensity changes for ~66% of the amide resonance peaks mapped to the structured region of the motif (W322 in the motif disappears on the first addition).
Selecting likely CaM binding region

<table>
<thead>
<tr>
<th></th>
<th>Human 261</th>
<th>Mouse 259</th>
<th>Human 311</th>
<th>Mouse 307</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pept</td>
<td>GTGDLDDLSSAFRRTSRAGGGRRISDSDHTGILDFSSLKRRDSFRTPRD</td>
<td>S..........A......A.T..............</td>
<td>SKLEAPAEEDVWEILRQAPPSEYERIAFQYGVTDLRGMLKRLKGMRR</td>
<td>S..................H.............KQ</td>
</tr>
<tr>
<td></td>
<td>310</td>
<td>306</td>
<td>358</td>
<td>353</td>
</tr>
</tbody>
</table>

Npep (18aa) RDSKLEAPAEEDVWEILR
Cpep (23aa) EDVWEILRQAPPSEYERIAFQYG
Ca$^{2+}$-CaM addition to motif alone blue-shifts the Trp emission peak, and the intensity decrease as for C1C2. The Cpep peptide from the motif sequence also changes the Trp emission spectrum, but not Npep.
Trp fluorescence and NMR yield similar $K_d$ values for Ca$^{2+}$-CaM-C1C2/C1C2EEE, Ca$^{2+}$-CaM-motif and Ca$^{2+}$-Cpep

<table>
<thead>
<tr>
<th></th>
<th>C1C2</th>
<th>C1C2 (+100 mM NaCl)</th>
<th>C1C2 (+EGTA)</th>
<th>C1C2 EEE</th>
<th>Motif</th>
<th>Motif (+50 mM NaCl)</th>
<th>Cpep</th>
<th>Npep</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K_d$ (µM)</td>
<td>13.5</td>
<td>8.7</td>
<td>&gt; 120</td>
<td>6.6</td>
<td>4.3</td>
<td>2.4</td>
<td>4.3</td>
<td>&gt; 400</td>
</tr>
<tr>
<td>±1SD</td>
<td>2.5</td>
<td>3.0</td>
<td>N/A</td>
<td>1.7</td>
<td>1.2</td>
<td>0.3</td>
<td>1.3</td>
<td>N/A</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.998</td>
<td>0.996</td>
<td>0.994</td>
<td>0.994</td>
<td>0.999</td>
<td>0.999</td>
<td>0.998</td>
<td>0.999</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Residues</th>
<th>F16</th>
<th>V55</th>
<th>A57</th>
<th>T70</th>
<th>R74</th>
<th>K115</th>
<th>L116</th>
<th>V121</th>
<th>I130</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K_d$ (µM)</td>
<td>47.3</td>
<td>59.2</td>
<td>4.9</td>
<td>28.4</td>
<td>24.4</td>
<td>9.3</td>
<td>5.3</td>
<td>9.4</td>
<td>4.6</td>
</tr>
<tr>
<td>±1SD</td>
<td>8.1</td>
<td>16.2</td>
<td>2.3</td>
<td>9.6</td>
<td>5.6</td>
<td>5.2</td>
<td>2.3</td>
<td>4.4</td>
<td>2.2</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.997</td>
<td>0.994</td>
<td>0.99</td>
<td>0.99</td>
<td>0.995</td>
<td>0.983</td>
<td>0.990</td>
<td>0.986</td>
<td>0.990</td>
</tr>
</tbody>
</table>
Small-angle X-ray Scattering indicates CaM is in an semi-extended conformation when bound to Cpep
Identification of CaM residues affected by motif binding (NMR chemical shifts).
An ordered hierarchy of phosphorylation exists among the cMyBP-C phosphorylation sites. S282 (mouse) is phosphorylated by CaMKII which induces further phosphorylation at S273 and S302 by PKA.

CaMKII-regulated phosphorylation is strictly Ca\(^{2+}\)-CaM-dependent and can be inhibited by the Ca\(^{2+}\) chelator EGTA or MLCK-I.

CaMKII inhibition reduces both cMyBP-C and TnI phosphorylation and decreases maximum force through a cross-bridge feedback mechanism.

When directly isolated from muscle tissue, cMyBP-C is purified with endogenous CaMKII activity.

The Ca\(^{2+}\)-CaM-dependent MLCK phosphorylates the myosin regulatory light chain (RLC).

The CaM-dependent phosphorylation of cMyBP-C and RLC both contribute to the contraction/relaxation cycle by modifying the local concentration of cross-bridges at the interface with actin.

The C0 domain of cMyBP-C can directly interact with the RLC, while the cMyBP-C motif, in an unphosphorylated state, can interact with myosin ΔS2.

The Ca\(^{2+}\)-CaM-cMyBP-C interaction is independent of cMyBP-C phosphorylation.
CaM may act as a structural conduit that links cMyBP-C with Ca\(^{2+}\) signaling pathways to help coordinate phosphorylation events and synchronise the multiple interactions between cMyBP-C, myosin and actin during the heart muscle contraction.