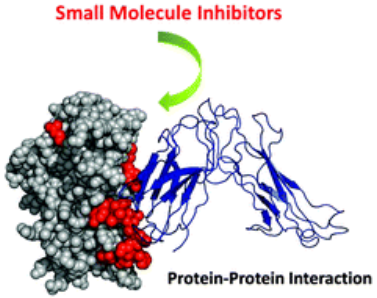
Calculating Label Distances Within Biomolecules Using Anomalous Small Angle X-ray Scattering

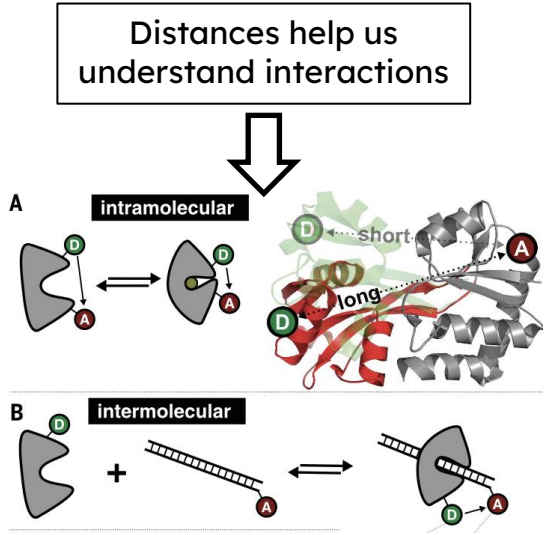
Student: Benson Chan
Mentors: Joseph Curtis, Susan Krueger
School: Thomas Sprigg Wootton High School

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

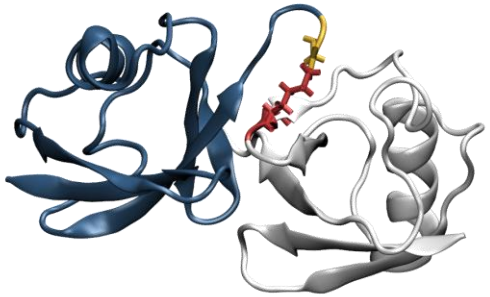
Distances in Biomolecules

- Interaction between proteins is important
- Learning how proteins interact helps us learn their function and can lead to medicines for diseases



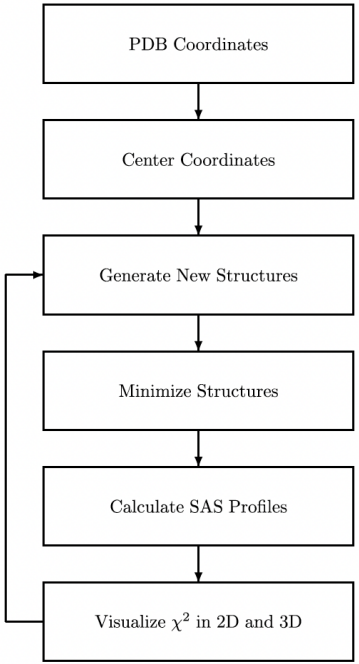


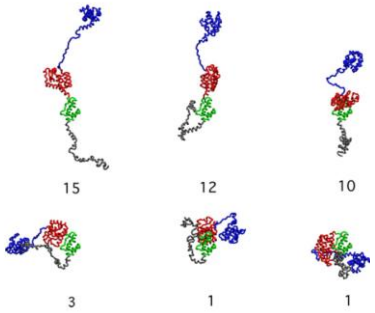
- Proteins and nucleic acids are dynamic
- Many possible structures, an ensemble
- Distances are dynamic as well

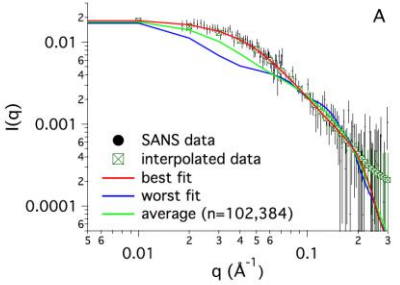


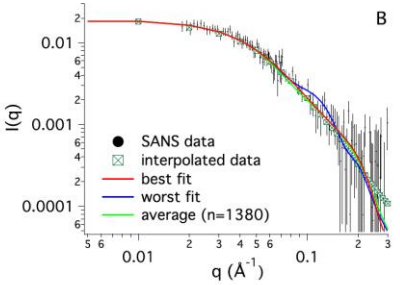
SASSIE / SASSIE-web

- Provides simulation and modeling tools
- Program process:

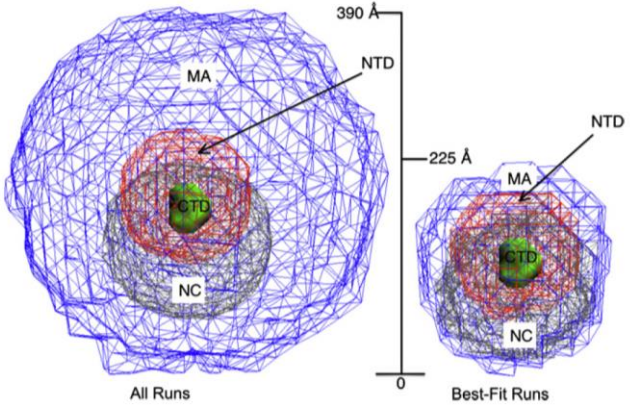








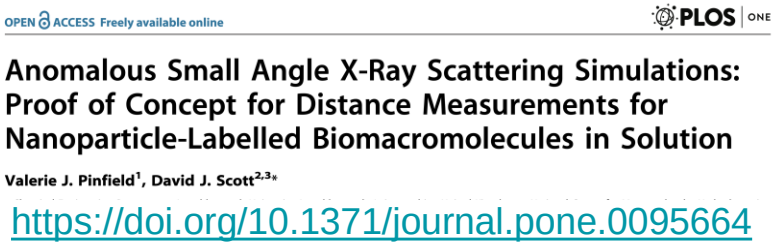
Example of SASSIE calculation process: HIV-1 Gag Protein



Curtis, J. E., Raghunandan, S., Nanda, H., & Krueger, S. (2012). SASSIE: A program to study intrinsically disordered biological molecules and macromolecular ensembles using experimental scattering restraints. *Computer Physics Communications*, *183*(2), 382–389.

Use of Anomalous Small Angle X-ray Scattering (ASAXS)

- Used in materials science for determining distances
- Uncommon for biological systems
  - Limited by sample preparation labeling methods
  - Lack of adequate tools to generate and evaluate ensembles and their theoretical ASAXS data



OPEN ACCESS Freely available online PLOS ONE

Anomalous Small Angle X-Ray Scattering Simulations: Proof of Concept for Distance Measurements for Nanoparticle-Labelled Biomacromolecules in Solution

Valerie J. Pinfield[1], David J. Scott[2,3]*

https://doi.org/10.1371/journal.pone.0095664

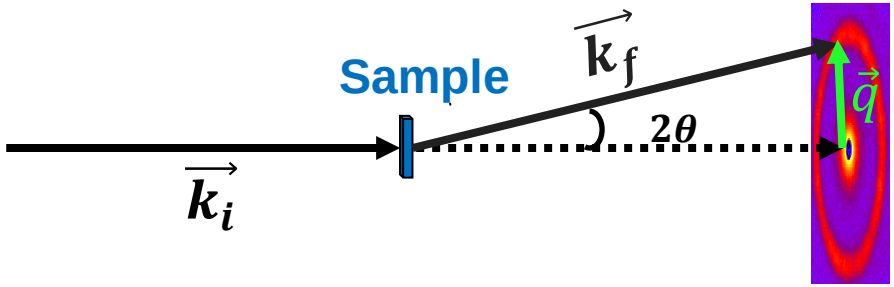
Program to calculate the intramolecular distance of a single labeled biomolecule by simulating ASAXS

But, biological systems are dynamic and often flexible.

Accounting ensembles and direct comparisons to experiments are needed.

Our goal: extend these ideas and use the simulation and experimental comparison tools of SASSIE-web toward Ensemble-ASAXS

Small Angle X-ray Scattering (SAXS)



Momentum transfer ($Å^{-1}$)

$$|\vec{q}| = q = \frac{4\pi \sin(\theta)}{\lambda}$$

X-ray wavelength (Å)

What is measured?

Scattering intensity: $$I(q) = \sum_{i=1}^{N} \sum_{j=1}^{N} f_i f_j \frac{\sin(q r_{ij})}{q r_{ij}}$$ **Debye Equation**

Number of atoms in the scattering particle

Atom-dependent scattering lengths

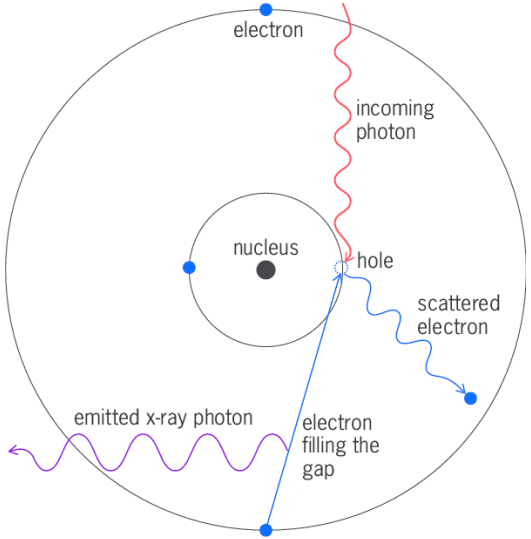
$f \sim Z$

Distance between atoms i and j

Anomalous Small Angle X-ray Scattering (ASAXS)

- Fluorescence X-ray photon emitted at very specific energy (absorption edge).

Relatively high-Z atom, e.g. Au

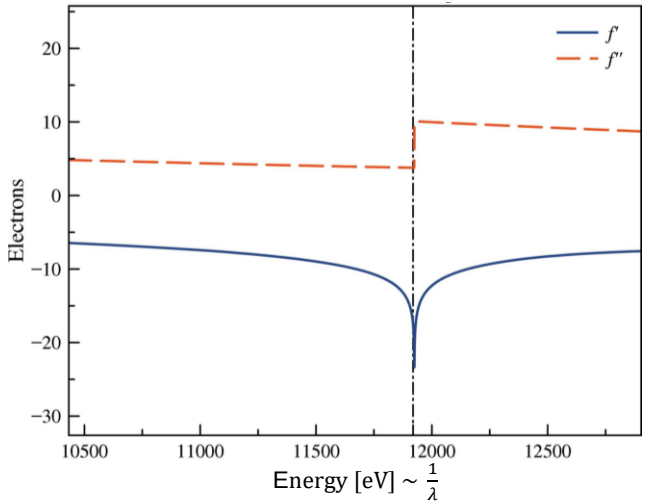


Mclaughlin, Brendan. (2013). McGraw Hill Science Yearbook 2013.

- Scattering length now has new energy-dependent terms near the absorption edge.

$$f = f_0 + f'(E) + if''(E)$$

Au absorption edge: 11919 eV



Gruzinov *et al.* (2021). *J. Synchrotron Rad.* **28**, 812-823.

Experimental Strategy Using ASAXS

1. Incorporate high-Z atoms (Au) into biological molecule
2. Perform SAXS measurements at several energies near the Au absorption edge
3. Separate $I(q)$ into resonant (R) and non-resonant (NR) terms
4. Solve for $I_R(q)$, $I_{NR}(q)$ and $I_{R,NR}(q)$

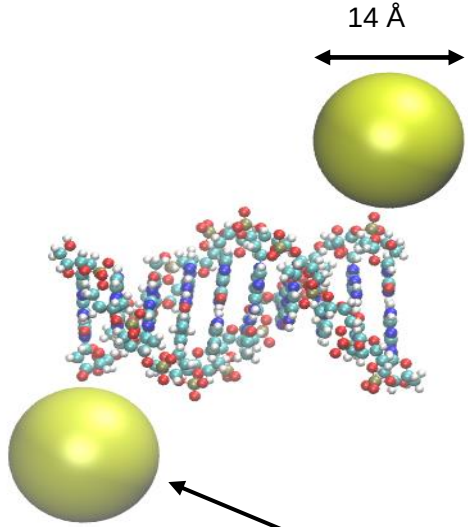
$$f(E) = f_0 + f'(E) + if''(E)$$

Inherent scattering factor

Absorption edge dependent

$$I(q) \sim f_R^2 I_R(q) + f_{NR}^2 I_{NR}(q) + 2 f_R f_{NR} I_{R,NR}(q)$$

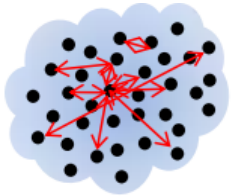
Au



Probability Distance Distribution Function

- Related to $I(q)$ by a Fourier Transform:

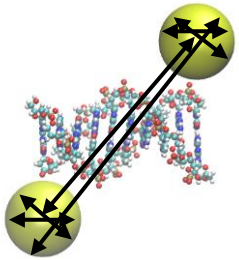
$$I(q) \sim \sum_{r=0}^{D_{max}} P(r) \frac{\sin(qr)}{qr}$$

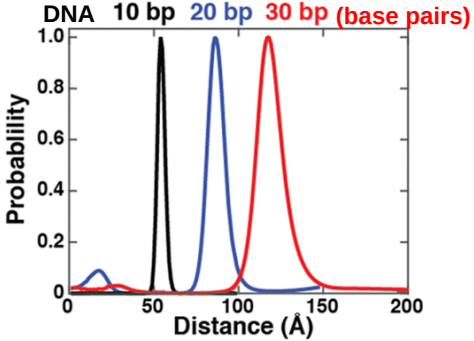


D_{max}: maximum distance in the molecule

$$P_R(r) \sim \sum_q I_R(q)\, q\, sin(qr)$$

R: resonant term (Au)





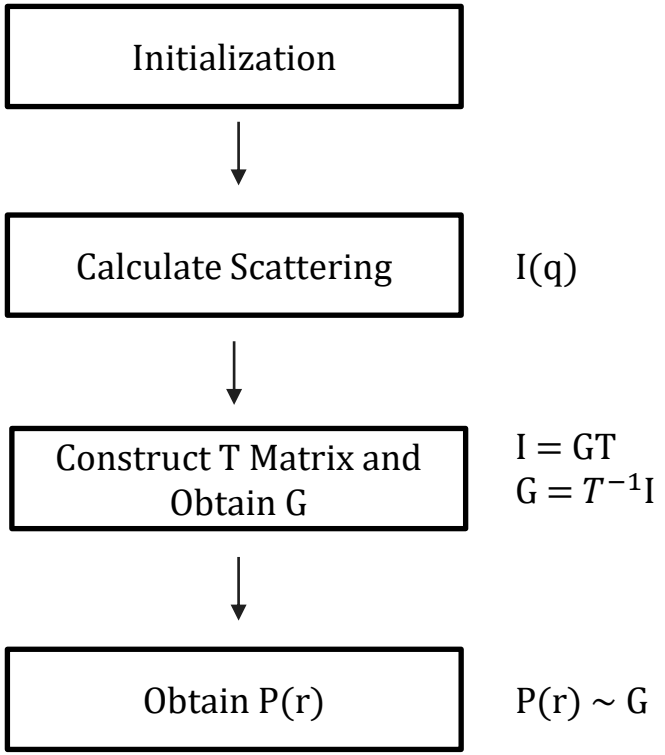
Zettl et al. (2016). *Nano Letters* **16**, 5353-5357.

Original Code Process

"are there vocaloids in other languages" yes. english, chinese, spanish, korean "does vocaloid have an anime" the vocaloid china project had a promotional anime and miku has cameo'd in a few animes "what is project diva" rhythm game developed by sega "is miku still a vocaloid" crypton is now developing their own singing synthesizer called piapro studio and all future miku banks will be on that engine instead from now on, though her vocaloid voicebanks still exist "i heard vocaloid is dying" yamaha has taken v4 off the market and v5 is very expensive and it is also now absurdly expensive for any company to make a voicebank so we haven't gotten a new vocaloid in like 2 years "who is kasane teto" an utau "what is utau" a freeware voice synthesizer program "is teto a vocaloid" no "what is voiceroid" basically TTS and some vocaloid characters are there "are rin and len siblings" if you want them to be "what is project sekai" mobile gacha game developed by sega "who is the vocaloid in porter robinson's songs" avanna "who is the vocaloid used in bee in puppycat" oliver "are gumi and miku friends" if you want them to be "who is miku's boyfriend" no one "how old is gumi" she doesnt have a canon age "how old is IA" she doesn't have a canon age "is IA still a vocaloid" she's been moved to CeVio now mostly "what is CeVio" another singing synthesizer "how is that pronounced" cheh-vee-oh "really?" yes "are there any black vocaloids" leon, lola, cyber songman, amy, chris "is gumi's last name megpoid" no, megpoid is the name of the software. her name is gumi "who is the first chinese vocaloid" luo tianyi "which is the first name in 'luo tianyi'" tianyi is her first name "can i just call her luo because it's shorter" it would be weird if you did "what is shiteiyanyo" a funny miku derivative "are neru and haku vocaloids" no "what happened to SeeU's voice provider" she went to jail "why does vocaloid sound like auto-tune" because it is a synthesizer "was miku the first vocaloid" no the first vocaloids were Leon and Lola. miku is a vocaloid2 "was miku the first vocaloid2" no that was sweet ann "was miku the first japanese vocaloid" no that was meiko "why was miku the first commercial success" because she was cute "is kaito miku's boyfriend" no "why is miku associated with a leek" probably originated from the leek spin meme "why is rin associated with oranges" her bow look like orange leaves "why is len associated with bananas" his hair looks like bananas "how many spanish vocaloids are there" 3 "which kind of spanish" spain spanish "how many male vocaloids are there" 23, not counting vy2 "wait kanshi yonezu is hachi?" yes "is vocaloid expensive" vocaloid 5 is $225 "how much do voicebanks usually cost" about 90 USD on average "does vocaloid tune itself" not usually no "is miku a hologram" screen projection technology is used for miku's live concert performances. miku herself is still just a software though "what's with that guy that legally married miku" idk "is miku a vtuber" no (not yet at least) "what is mikumikudance" freeware 3d animation program "how many languages does miku know" miku has voicebanks in japanese, english, and chinese "what is mikuexpo" world concert tour events organized by crypton future media "who provided the voice for miku" saki fujita "who manages miku" crypton future media "does miku manage all the vocaloids" no vocaloids are managed by their own companies "what does hatsune miku mean" first sound of th

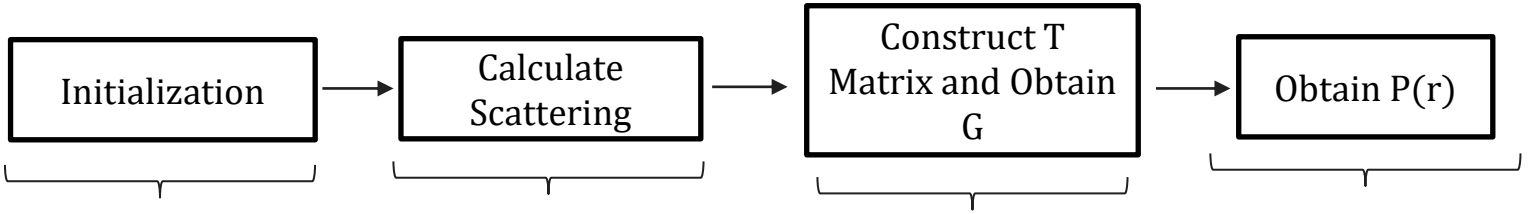
Original Script

*Not actually, though



The new code will extend the process to handle ensembles and experimental data.

Workflow



Original code

- Validate independent code w/o GUI output to original code output

Independent code section

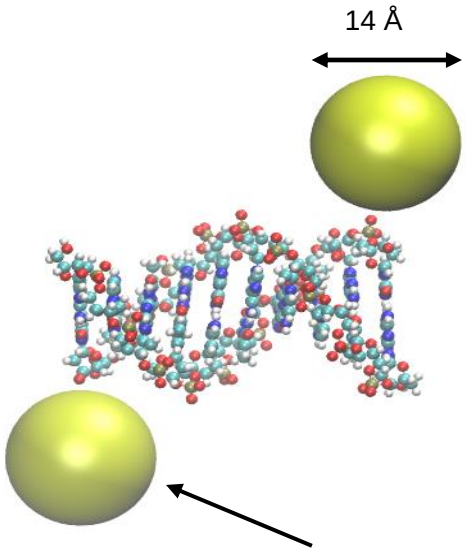
- Validate prototype script output to independent code output

Prototype script

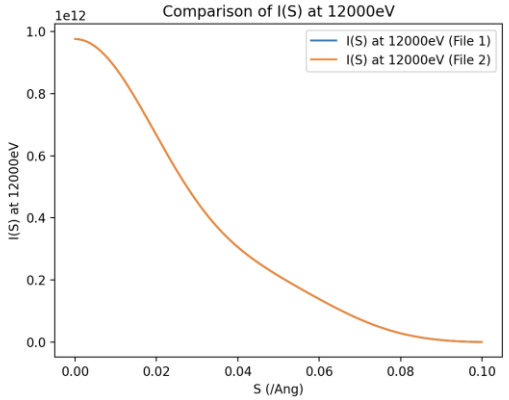
- Implement prototype script to Sassie-web

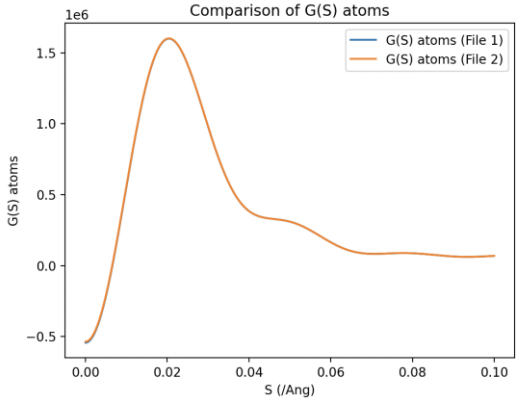
Ensemble-ASAXS

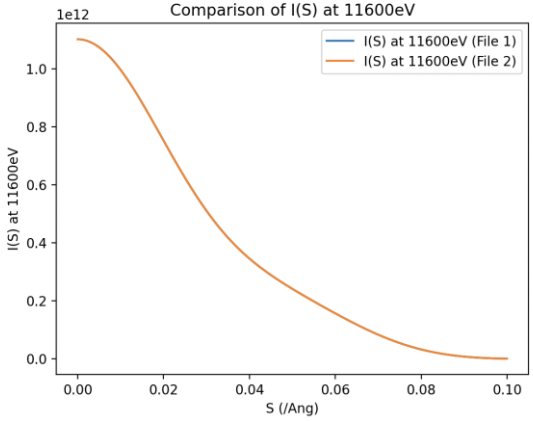
Validating Visualization

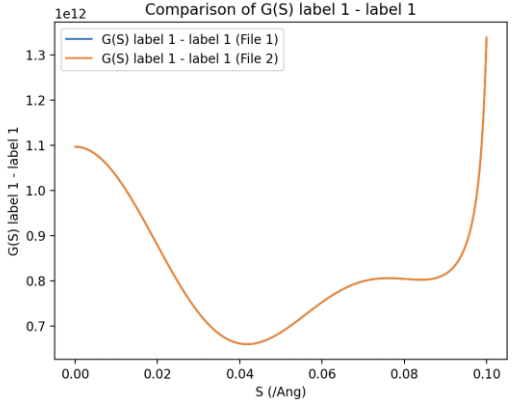


Au nanoparticle (78 Au atoms)









Progress

	Original Code	Independent Script	Prototype Code	Ensemble-ASAXS
Initialization	✓	✓	✓	✓
Calculate Scattering	✓	✓	✓	✓
Construct T and Obtain G	✓	✓	✓	
Obtain P(r)	✓	✓		

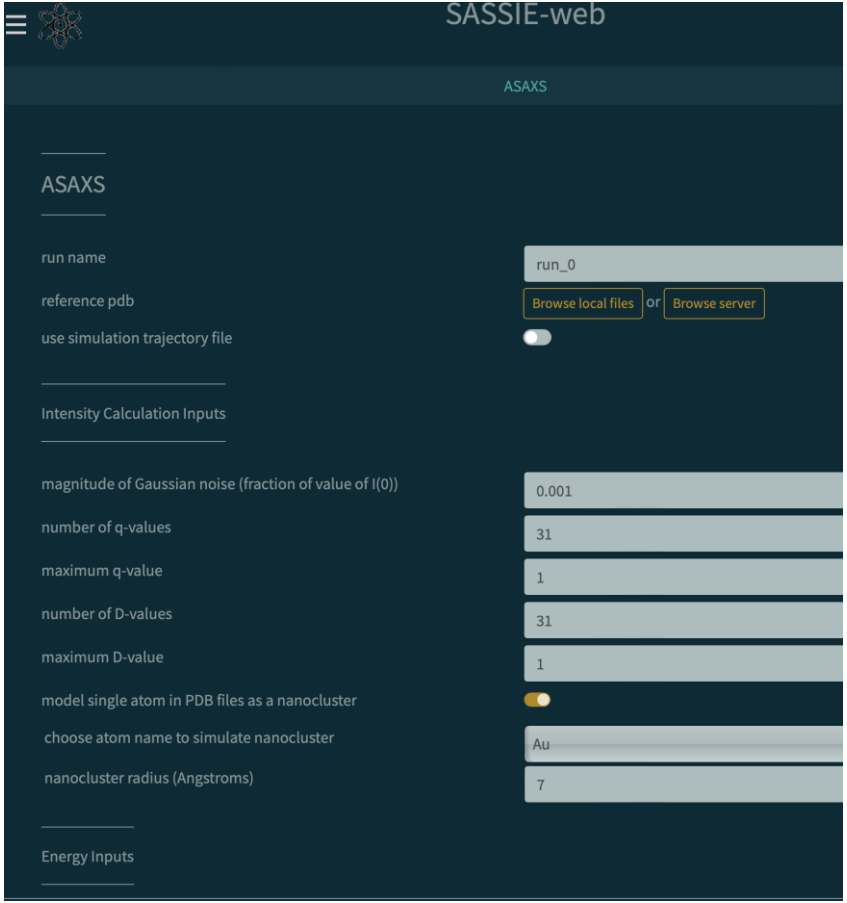
Future Work

1. Complete the remaining bits for the calculation process steps & enable CUDA/GPU.

2. Successfully use Ensemble-ASAXS for DNA-protein experimental data

3. Release Ensemble-ASAXS to the community in the Beta stage

4. Manuscript!



Thank you for listening!

Special thanks to Joseph Curtis and Susan Krueger!