Recording and Analyzing Nucleic Acid Distance Distributions with X-Ray Scattering Interferometry (XSI)

Thomas Zettl,[1,2] Rhiju Das,[2,3] Pehr A. B. Harbury,[2,6] Daniel Herschlag,[2,4,6] Jan Lipfert,[1,6] Rebecca S. Mathew,[5] and Xuesong Shi[2]

[1]Department of Physics, Nanosystems Initiative Munich, and Center for Nanoscience, LMU Munich, Munich, Germany
[2]Department of Biochemistry, Stanford University, Stanford, California
[3]Department of Physics, Stanford University, Stanford, California
[4]Department of Chemical Engineering, Stanford University, Stanford, California
[5]Department of Cell Biology, Harvard Medical School, Harvard University, Boston, Massachusetts
[6]Corresponding authors: *harbury@stanford.edu*, *herschla@stanford.edu*, *Jan.Lipfert@lmu.de*

Most structural techniques provide averaged information or information about a single predominant conformational state. However, biological macromolecules typically function through series of conformations. Therefore, a complete understanding of macromolecular structures requires knowledge of the ensembles that represent probabilities on a conformational free energy landscape. Here we describe an emerging approach, X-ray scattering interferometry (XSI), a method that provides instantaneous distance distributions for molecules in solution. XSI uses gold nanocrystal labels site-specifically attached to a macromolecule and measures the scattering interference from pairs of heavy metal labels. The recorded signal can directly be transformed into a distance distribution between the two probes. We describe the underlying concepts, present a detailed protocol for preparing samples and recording XSI data, and provide a custom-written graphical user interface to facilitate XSI data analysis.

Keywords: ensemble determination • energy landscape • gold nanocrystals • molecular ruler • structure determination • small angle X-ray scattering • X-ray scattering interferometry

How to cite this article:
Zettl, T., Das, R., Harbury, P. A. B., Herschlag, D., Lipfert, J., Mathew, R. S., & Shi, X. (2018). Recording and analyzing nucleic acid distance distributions with X-ray scattering interferometry (XSI). *Current Protocols in Nucleic Acid Chemistry*, e54. doi: 10.1002/cpnc.54

INTRODUCTION

Richard Feynman famously said "everything that is living can be understood in terms of the jiggling and wiggling of atoms" (Feynman, Leighton, & Sands, 1977, p. 3–6). Biological macromolecules such as unfolded or partially folded RNAs or intrinsically disordered proteins, are especially dynamic, given the noncovalent forces that hold them together, their aqueous surroundings, and physiological temperature that provides thermal energy. Moreover, significant conformational changes of molecules can be triggered by external stimuli and are typically integrally involved in the functions of biomolecules

Current Protocols in Nucleic Acid Chemistry e54
Published in Wiley Online Library (wileyonlinelibrary.com).
doi: 10.1002/cpnc.54

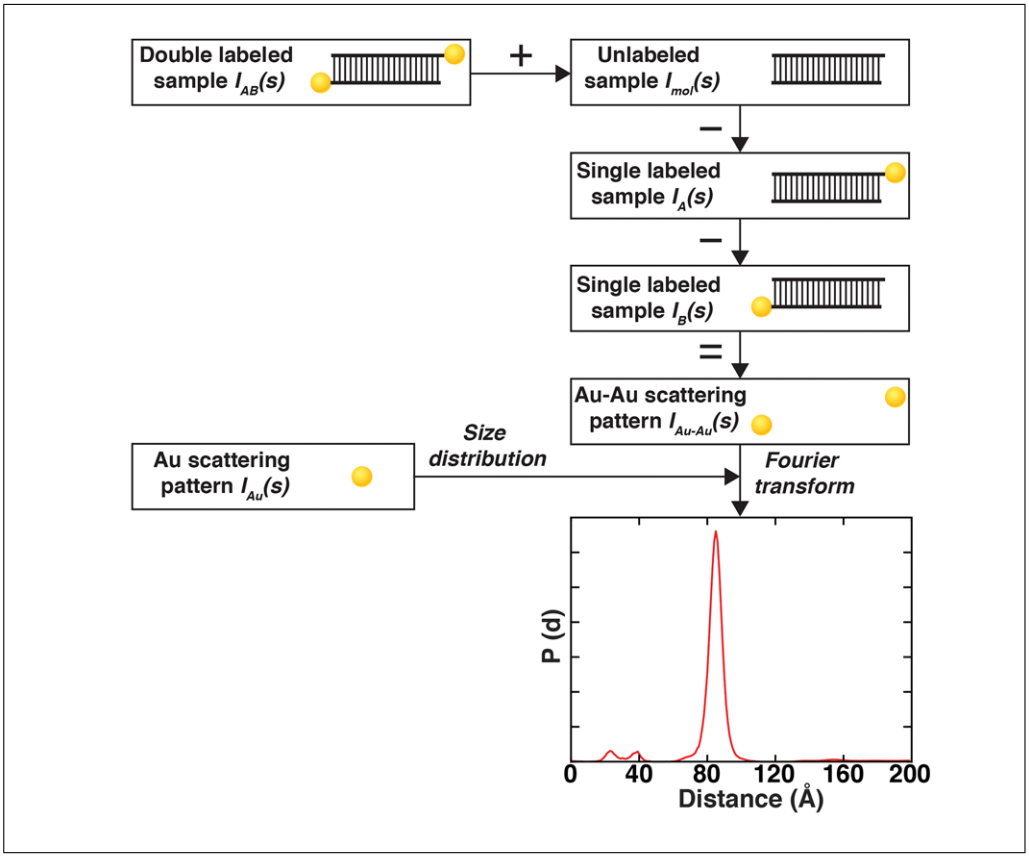

Figure 1 Schematic of AuSAXS workflow to determine gold label-gold label distance distributions. Scattering intensity equation for a single-labeled molecule. The scattering signal can be decomposed into a sum of the individual scattering contributions: Double-labeled sample, the macromolecule only, two gold-macromolecule cross-terms, and the interference from the gold labels. Schematic of the workflow to determine the Au-Au distance distribution. The SAXS profiles of the shown samples are used to extract the gold-gold interference scattering profile. The interference pattern is Fourier transformed into a distance distribution using basic profiles generated with the size distribution of Au nanocrystals. SAXS, small-angle X-ray scattering.

(Fischer, Konevega, Wintermeyer, Rodnina, & Stark, 2010; Frauenfelder, 1989; Müller et al., 2016). Thus, conformational changes play an important role in understanding the basic mechanics and are key in reconstructing causes from the molecular level to macromolecular systems.

This article describes an emerging molecular ruler, termed X-ray scattering interferometry (XSI) (Mathew-Fenn, Das, & Harbury, 2008a; Mathew-Fenn, Das, Silverman, Walker, & Harbury, 2008b; Shi, Beauchamp, Harbury, & Herschlag, 2014; Shi, Bonilla, Herschlag, & Harbury, 2015; Shi, Herschlag, & Harbury, 2013; Shi, Huang, Lilley, Harbury, & Herschlag, 2016; Shi, Walker, Harbury, & Herschlag, 2017), which can be used to generate whole distance distributions at Ångström resolution. XSI measures the interference of scattered X-rays between two specifically attached gold nanocrystals (Fig. 1). The strengths of XSI are that it provides: (1) distance information in solution; (2) the distance information that is unperturbed by temporal averaging because scattering is fast relative to atomic motions; (3) the direct mathematical relationship between scattering and distance, formally related by Fourier transformation, allows XSI data to be unambiguously converted into a calibrated distance distribution; and (4) while sample preparation is time consuming, it is straightforward, highly reliable, and highly reproducible.

Over the past decade, XSI has been successfully applied to nucleic acids and nucleic acid/protein complexes (Hura et al., 2013; Mathew-Fenn et al., 2008a; Mathew-Fenn

et al., 2008b; Shi et al., 2014; Shi et al., 2015; Shi et al., 2013; Shi et al., 2016; Shi et al., 2017). Labeling strategies employing gold nanocrystals of various sizes, diverse macromolecules, and variable attachment positions have been reported. So far, demonstrated labeling strategies include: (1) end-labeled DNA nanocrystal conjugates (Ackerson, Sykes, & Kornberg, 2005; Alivisatos et al., 1996; Mathew-Fenn et al., 2008a; Mathew-Fenn et al., 2008b; Shi et al., 2014; Shi et al., 2015; Shi et al., 2013), (2) end-labeled RNA molecules (Shi et al., 2015; Shi et al., 2016; Shi et al., 2017), (3) gold labels positioned at defined internal sites of DNA or RNA helices (Mathew-Fenn et al., 2008a; Shi et al., 2014; Shi et al., 2015; Shi et al., 2013; Shi et al., 2016; Shi et al., 2017), and (4) protocols to form single-labeled protein constructs (Aubin-Tam & Hamad-Schifferli, 2005; Aubin-Tam, Hwang, & Hamad-Schifferli, 2009; Azubel & Kornberg, 2016).

This article covers the design and preparation of end-labeled nucleic acid samples. Below we present the procedures required to prepare samples, to acquire XSI data, and to generate ensemble distance distributions. Basic Protocol 1 describes sample preparation for end-labeled nucleic acid gold conjugates and includes a protocol for gold nanocrystal synthesis. Basic Protocol 2 describes the acquisition of a full data set at a synchrotron radiation facility for XSI analysis. Basic Protocol 3 describes the data analysis of XSI data and the use of a custom-written graphical user interface (GUI) in MATLAB. The detailed protocols and the user interface presented in this chapter will enable scientists interested in molecular distance measurements to perform and analyze XSI measurements easily.

SAMPLE PREPARATION

BASIC PROTOCOL 1

In this article, we focus on the preparation of end-labeled nucleic acid samples for XSI. An additional protocol on labeling proteins will be forthcoming.

Briefly, thioglucose-protected gold nanocrystals are synthesized using the method of Schaaff and coworkers (Schaaff, Knight, Shafigullin, Borkman, & Whetten, 1998). DNA or RNA oligonucleotides are ordered with a C3-thiol modification, from a commercially available source, for end labeling or a C2 dT amino modification for internal labeling. High-performance liquid chromatography (HPLC), performed either in-house or by the oligonucleotide vendor, is used to purify the oligonucleotides. In the case of the internal C2 dT amino modification, the amino group is converted to a thiol group using the commercially available N-succinimidyl 3-(2-pyridyldithio)propionate (SPDP) cross-linker. For a detailed protocol of internal label attachment see Shi et al., 2015. The gold nanocrystals couple to thiol groups, forming stable conjugates. A second HPLC purification step is used to purify 1:1 nanocrystal-nucleic-acid conjugates, eliminating nanocrystals coupled to multiple oligonucleotides and excess gold particles. Finally, modified and unmodified single-stranded molecules are mixed in various combinations to form a sample quartet, which consists of one unmodified construct, two complementary single-labeled molecules with a single gold nanocrystal attached to one of the two labeling sites respectively, and one double-labeled construct. After HPLC purification and desalting, these duplexed constructs can be stored at –20°C for several months. A full set of samples is required for the data analysis to work as explained in detail in Basic Protocol 3.

The protocol below describes: (1) the synthesis of monodisperse, thiol-passivated gold nanocrystals with 0.7 nm radius; (2) the preparation of end-labeled gold-oligonucleotide conjugates; and (3) the preparation of a sample quartet that is ready for XSI data acquisition. The concentrations cited below are based on ordering a 200 nmol scale nucleic acid synthesis, and can be adjusted for alternate quantities of starting material.

CAUTION: Some of the chemicals and reagents used are flammable. Refer to material safety data sheets prior to use. Conduct reactions in a well-ventilated fume hood and use standard laboratory protective equipment.

NOTE: Use ultrapure water in all solutions and protocol steps.

Materials

Fast protein liquid chromatography (FPLC) cleaning solution (see recipe)
Size exclusion running buffer (see recipe)
Isopropanol
Methanol
Acetic acid
Hydrogen tetrachloroaurate(III) hydrate (Sigma Aldrich, cat. no. 50790)
1-Thio-β-D-glucose (Sigma Aldrich, cat. no. T6375)
Sodium borohydride
3′-Thiol modified oligonucleotides [Integrated DNA Technologies (IDT) or Stanford Protein and Nucleic Acid (PAN) Facility]
Oligonucleotides
2 M ammonium acetate, pH 5.6
Low salt borate buffer (see recipe)
High salt borate buffer (see recipe)
Ethanol
Magnesium chloride ($MgCl_2$)
1 M Tris·Cl, pH 9.0
Low salt ammonium acetate buffer (see recipe)
High salt ammonium acetate buffer (see recipe)

G25 column, 26/10 housing (Sigma Aldrich, cat. no. GE17-5087)
FPLC system (dual wavelength detector recommended)
Superdex 30 column, 16/600 housing (Sigma Aldrich, cat. no. GE28-9893-31)
250-mL round-bottom flask
Optional: Addition funnel with metering valve (e.g., Chemglass Life Sciences, cat. no. CG-1714)
Magnetic stirrer/hotplate
Vortexer
0.22-μm syringe filter units
3 kDa and 10 kDa Amicon spin filtration units (Sigma Aldrich)
Dionex DNAPac Pa200 column, 9/250 housing (Thermo Fisher Scientific, cat. no. 063421)
HPLC system (dual wavelength detector recommended)
Rotary evaporator equipped with water bath, dry ice condenser, and connected to an oil pump
–20° or –80°C freezer
Nanodrop ND-1000 spectrophotometer or other UV spectrophotometer

Preparation of FPLC columns (start 2 days in advance)

1. Using a flow rate of 2 mL/min, wash G25 column with two column volumes water, followed by four column volumes FPLC cleaning solution. Equilibrate column with four column volumes water.

2. Using a flow rate of 0.75 mL/min, wash Superdex 30 column with two column volumes water, followed by five column volumes FPLC cleaning solution. Remove FPLC cleaning solution with five volumes water. Finally, equilibrate Superdex 30 with two column volumes size exclusion running buffer.

The gold nanocrystals cause the column resin to turn brown. This coloration is reversed by the DTT in the cleaning solution. If the resin does not revert to off-white after five column volumes, apply additional cleaning solution before equilibrating the column. Start the cleaning and equilibration of both columns at least 1 day before the synthesis.

Synthesis of gold nanocrystals (4 hr)

3. Rinse a 250-mL round-bottom flask with isopropanol, dry in a heated oven, and cool (to room temperature). Add stir bar, cap flask with Parafilm, and mount flask above a magnetic stir plate.
4. Prepare 72 mL 5:1 (v/v) methanol/acetic acid (60 mL methanol, 12 mL acetic acid).
5. Weigh 0.544 g hydrogen tetrachloroaurate(III) hydrate and immediately transfer to round-bottom flask. Add 36 mL 5:1 methanol/acetic acid solution to the flask; color should be a clear, bright orange.
6. Dissolve 1 g 1-thio-β-D-glucose in 36 mL 5:1 methanol/acetic acid mixture and vortex solution until the powder is fully dissolved.
7. Add dissolved 1-thio-β-D-glucose to the 250-mL round-bottom flask; solution should turn cloudy. Stir mixture 20 min at room temperature.
8. Weigh 0.9 g sodium borohydride and dissolve in 20 mL water (H_2O) by vortexing.
9. Carefully add sodium borohydride solution dropwise to reaction flask over 12 to 15 min. Use a 1-mL pipet for this step or alternatively an addition funnel with a metering valve. Set the stirring rate to allow rapid intermixing of the droplets.

   It is critical to keep the addition rate and droplet volume constant to ensure monodisperse and high-quality gold nanocrystals.

10. Stir mixture 30 min at room temperature.
11. Concentrate resulting solution (~92 mL) to 12 to 20 mL using a rotary evaporator. Maintain water bath at room temperature.

Purification of gold nanocrystals (10 hr)

12. Filter crude nanocrystal solution using a 0.22-μm filter unit and store solution on ice.
13. Desalt filtered nanocrystals with the prepared G25 column. Split total volume into two to three runs. Load 1 aliquot solution (at most 5 mL with a standard sample loop) while applying a flow rate of 2 mL/min using water as the running buffer. Set FPLC system up to monitor 260 or 280 nm to detect Au nanocrystals.

    Particles should elute between 8 and 14 min (see Fig. 2). Also monitor conductance of the solution to detect the salt peak, which elutes after the nanocrystals.

    Repeat desalting procedure for additional aliquots.

    The desalted Au nanocrystals are stable and can be stored at 4°C up to a few days or at –20°C for months.

14. Concentrate desalted Au nanocrystal solution to 10 mL with centrifugal filter units (3 kDa cutoff, 15 mL, 3000 × *g*, 40 min) at 4°C.

    The flow through should be clear.

15. Purify a monodisperse population of Au nanocrystals on Superdex 30 size exclusion column (see Fig. 3). Set up FPLC system to monitor 260 or 280 nm to detect the nanocrystals; up to 5 mL of sample can be injected per run. After loading

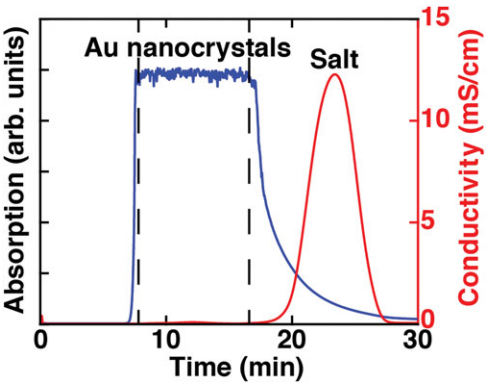

Figure 2 A sample FPLC chromatogram of Au nanocrystal desalting. The absorption of the Au nanocrystals (blue) is monitored at 260 nm and separated from the salt front (red) observed by a peak in conductivity. The sample eluting between the two dashed lines was used. FPLC, fast protein liquid chromatography.

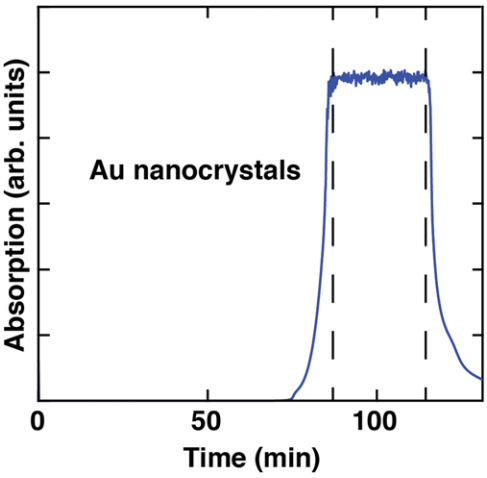

Figure 3 A sample FPLC chromatogram of Au nanocrystal size exclusion. The absorption of the Au nanocrystals (blue) is monitored at 260 nm. The sample between the two dashed lines can be used for a highly uniform particle size distribution. FPLC, fast protein liquid chromatography.

an aliquot of filtered, desalted, and concentrated nanocrystal solution, apply size exclusion running buffer at 0.75 mL/min for at least 210 min. Collect only the center of largest Au nanocrystal elution peak (see Fig. 3), and discard lower and upper shoulders. Immediately desalt solution using H_2O and centrifugal filter units (3 kDa cutoff, 15 mL) at 3000 × *g*, 4°C. Repeat centrifugal desalting three times and pool concentrated particles after the final run.

16. Determine final gold nanocrystal concentration by measuring UV absorption; extinction coefficient is 0.076 μm/cm at 360 nm. Store solution at –20°C.

Typically one synthesis yields 3 to 10 μmol of purified gold nanocrystals.

Preparation of gold nanocrystal-nucleic acid conjugates

17. Prepare oligonucleotides (~8 hr for four oligonucleotides): Purchase 3′-thiol modified oligonucleotides (C3-S-S) and unmodified oligonucleotides with the same sequences.

It is critical to use the short three-carbon linker for end-labeled samples to ensure high-quality results for the measurements of the distance distributions.

Design construct such that the terminating bases are GC base pairs to minimize end fraying.

18. Purify ordered oligonucleotides using the Dionex DNAPac 200 and anion exchange HPLC. Inject up to 100 nmol oligonucleotide onto the column and apply a flow rate of 3 mL/min.

    The salt gradient is formed from low salt borate buffer and high salt borate buffer.

    Tune the period of the gradient according to the length of your oligonucleotide. Perform an analytical run before the preparative runs to determine elution time of the product. For analytical runs, adjust injected sample according to the instrument detection sensitivity.

19. Desalt purified oligonucleotides by buffer exchange into water using centrifugal filter units and H_2O (3 kDa cutoff, 4 mL, 4,000 × *g* in a swinging basket centrifuge or 7500 × *g* in a fixed angle rotor, 30 min, 4°C). Repeat this step three times. Reduce volume to ~40 μL using centrifugal filter units (3 kDa cutoff, 0.5 mL, 14,000 × *g* using a benchtop centrifuge, 30 min, 4°C).

    Typically, ~60 nmol of oligonucleotide remain after purification of a 200 nmol scale synthesis.

    To protect against loss of oligonucleotide from a broken filter unit, keep flow through and check its absorbance at 260 nm; the oligonucleotides can be stored in a freezer at –20°C.

Thiol-modified oligonucleotides (~3 hr)

20. Immediately before coupling oligonucleotides to gold nanocrystals, ensure pendant thiols are fully reduced by adding 150 μL 200 mM DTT and 50 mM Tris·Cl, pH 9.0. Incubate 30 min, 60°C for DNA or 50°C for RNA.

    The DTT can be replaced by immobilized TCEP Reducing Gel (Thermo Fisher Scientific, cat. no. 77712). Follow the manual provided by the vendor to reduce and extract the oligonucleotides and proceed with step 23.

21. Purify oligonucleotide by ethanol precipitation. Add 2 μL 2 M $MgCl_2$ and 1 mL cold ethanol and mix solution. Incubate mixture on dry ice 40 min. Spin mixture (15,000 × *g*, 30 min, 4°C). Remove supernatant and wash with 1 mL ethanol. Spin mixture again (15,000 × *g*, 15 min, 4°C) and remove supernatant. Be careful not to disturb the precipitated pellet on the bottom.

22. Dissolve precipitated pellet in 500 μL H_2O and add it to a filter unit (3 kDa, 0.5 mL). Spin solution (14,000 × *g*, 30 min, 4°C) to remove residual DTT and determine final concentration of oligonucleotide by UV absorption at 260 nm. Use extinction coefficient provided by the manufacturer or calculate it using the nucleic acid sequence (e.g., OligoAnalyzer 3.1 *http://www.idtdna.com/calc/analyzer*). If a strong absorbance at 230 nm and a shoulder peak above 300 nm is observed, repeat this step to remove excess DTT. Keep flow-through to test for broken filter units. Immediately proceed to step 23.

Conjugate oligonucleotides and nanocrystals (~8 hr)

23. Add a 6-fold molar excess purified and desalted gold nanocrystals to the reduced oligonucleotide and vortex mixture (i.e., add 300 nmol Au particles to 50 nmol oligonucleotides). Add 20 μL 1 M Tris·Cl, pH 9.0 and vortex again. Incubate solution 2 hr at room temperature.

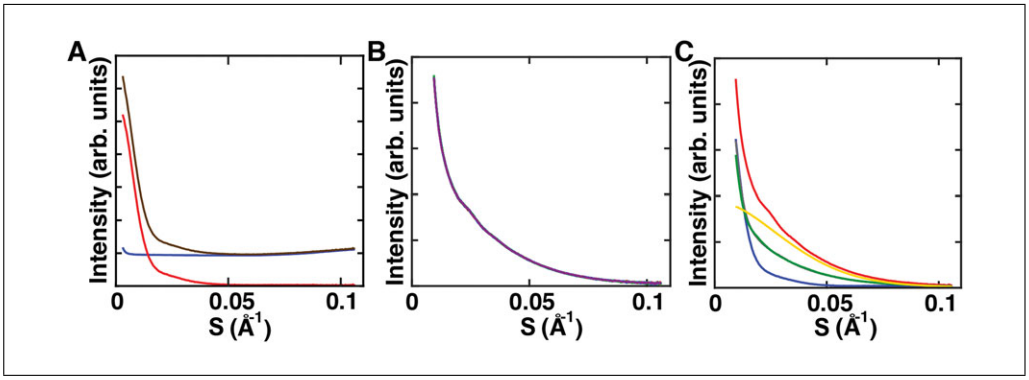

Figure 4 Example scattering profiles of XSI samples. (**A**) Buffer-only scattering profile used for buffer subtraction (blue), scattering profile of an unlabeled sample without (brown) and with buffer subtraction (red). (**B**) Ten individual exposures for a double-labeled sample recorded in one run. All profiles match and thus radiation damage can be excluded. These scattering profiles are buffer subtracted. (**C**) One full set of samples including one double-labeled sample (red), two orthogonal single-labeled samples (green, gray), bare gold nanocrystals (yellow), and unlabeled sample (blue). All scattering profiles are buffer subtracted. XSI, X-ray scattering interferometry.

24. Add 15 μL 2 M ammonium acetate, pH 5.6 to stop reaction and store mixture on ice.

25. Purify solution using the Dionex DNAPac 200 and anion exchange HPLC. Elute conjugates with a salt gradient using a low salt acetate buffer and a high salt acetate buffer. Tune period of gradient according to the length of your oligonucleotide.

    Typically, 15-mer DNA-gold nanocrystal conjugates elute at around 40% high salt buffer.

    Set flow to 3 mL/min and monitor absorbance at 260 and 360 nm.

    The oligonucleotides only absorb at 260 nm whereas the gold nanocrystals absorb at both wavelengths. Typically, oligonucleotides with a single gold nanocrystal elute earlier than unlabeled oligonucleotides of the same length. Free gold nanocrystals elute very early and gold nanocrystals with multiple oligonucleotides elute later than 1:1 conjugates. For a detailed chromatogram see Shi et al., 2015.

    Use a small amount of sample to perform an initial analytical run, so that you can to make adjustments to the salt gradient if required.

26. Desalt gold-oligonucleotide conjugates using centrifugal filter units (3 kDa, 4000 × *g*, 35 to 40 min, 4°C) and H_2O. Repeat this step three times. Determine concentration of purified conjugate by measuring absorption at 360 nm (0.076 μM/cm).

    The desalted conjugates are stable and can be stored at –20°C for months. Typically, 12 nmol of sample can be recovered.

Preparation of final duplex conjugates for a sample quartet (~5 hr)

27. Mix pairs of complementary single-stranded oligonucleotides in equimolar ratio and incubate DNA at room temperature or RNA at 40°C, 30 min.

    A samples quartet consists of a double-labeled sample with two modified strands, two single-labeled samples with a single modified strand, and an unlabeled sample (see Fig. 1 and Fig. 4). Use unmodified oligonucleotides for unlabeled sample or as the complementary strand for single-labeled samples.

    Use desalted oligonucleotides from step 26. If your structure does not form in H_2O and room temperature only, add the required buffer and salt to the solution and perform thermal annealing.

28. Purify annealed constructs by anion exchange HPLC using the same approach as in step 25. Use a small amount of sample to perform an initial analytical run to allow adjustments to salt gradient if required.

    The duplex constructs typically elute later then single-stranded conjugates.

29. Collect desired HPLC fractions and immediately desalt them using centrifugal filter units (10 kDa, 4000 × *g*, 15 min, 4°C) and H_2O. Repeat this step three times. Determine concentration of purified conjugate by measuring the absorption at 360 nm (0.076 μM/cm for single-labeled samples and 0.152 μM/cm for double-labeled samples).

    Typically, 2 to 3 nmol of each final duplexed construct should be obtained. The desalted conjugates are stable and can be stored at –20°C for months.

COLLECTING X-RAY SCATTERING INTERFEROMETRY DATA

BASIC PROTOCOL 2

To date, XSI data has been successfully recorded at beamline 4-2 of the Stanford Synchrotron Radiation Lightsource (SSRL; Mathew-Fenn et al., 2008a; Mathew-Fenn et al., 2008b; Shi et al., 2014; Shi et al., 2013; Shi et al., 2016; Shi et al., 2017), beamline 12-ID of the Advanced Photon Source (APS; Mathew-Fenn et al., 2008a; Mathew-Fenn et al., 2008b), the SIBLYS beamline of the Advanced Light Source (ALS; Hura et al., 2013), and the BM29 beamline of the European Synchrotron Radiation Facility (ESRF; see Fig. 5). In general, measurements can be carried out at any synchrotron with beam lines set up for small-angle X-ray scattering (SAXS) measurements that meet the following requirements:

(1) *S*-range: 0.0015 to 0.11 $Å^{-1}$ (optimal, for details see below) corresponding to a *q*-range 0.01 to 0.7 $Å^{-1}$
(2) X-ray energy: 9 to 15 keV (9 to 11 keV is optimal). This is the tested energy range used in experiments to date, for details see below.
(3) Sample volumes: 16 to 40 μL is the typical amount used at state-of-the-art synchrotrons (Lipfert, Millett, Seifert, & Doniach, 2006). This amount allows for ten independent exposures without requiring large quantities of sample.

It is important to pay attention to the definition of the magnitude of the momentum transfer vector *S* as two different conventions are used. In this protocol *S* is defined as $S = 2sin(\theta/\lambda)$, *q* is defined as $4 \cdot \pi \cdot sin(\theta/\lambda)$, where λ is the X-ray wavelength and θ is half the total scattering angle. *S* is alternatively reported in units of $Å^{-1}$ and nm^{-1}. We report *S* in $Å^{-1}$ in this protocol. Set up a sample-to-detector distance that covers an *S*-range

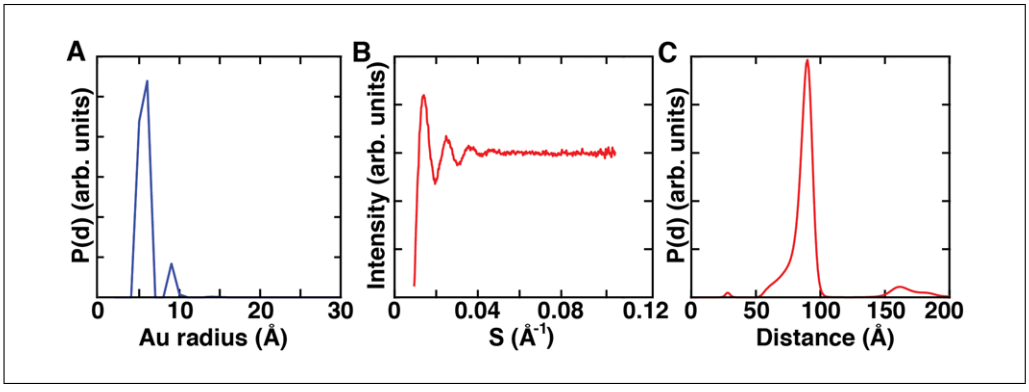

Figure 5 Maximum entropy analysis of XSI data. (**A**) Radius distribution of gold nanocrystals used to generate basis profiles *I(S,D)*. (**B**) Calculated interference pattern for the example data set, and (**C**) final distance distribution averaged over 10 maximum entropy fitting runs. XSI, X-ray scattering interferometry.

from 0.0015 to 0.11 $Å^{-1}$ (for example, this corresponds to a sample-to-detector distance of 1.1 meter for the Pilatus 300K detector at 11 keV on Stanford beamline 4-2). Typically, the sample and detector configuration must be arranged with beamline scientists well in advance of data collection, since it requires hardware alignment and calibration. If the beamline cannot cover the full "optimal" *S*-range (0.0015 to 0.11 $Å^{-1}$), S_{max} should not be <0.095 $Å^{-1}$ as the labeled samples and the bare gold particles contribute scattering intensities up to 0.085 $Å^{-1}$. This is crucial for obtaining a valid interference profile by the analysis procedure described below. If the chosen beamline cannot reach this S_{max} for lower X-ray energies, one solution is to extend the *S*-range by selecting higher energies (for example, using 15 keV instead of 11 keV). However, X-ray energies close to gold absorption edges (L-III at 11.92 keV, L-II at 13.73 keV, and L-I at 14.35 keV) should be avoided and energies below L-III are preferable to minimize X-ray fluorescence from these edges. Be aware that important details of the scattering profile can be lost in the low *S*-range for X-ray energies chosen too low or depleted by a low signal-to-noise ratio at X-ray energies set too high. (For general protocols on SAXS sample preparation, data collection, and data analysis see Burke & Butcher, 2012; Doniach & Lipfert, 2012; Dyer et al., 2014; Grishaev, 2012; Hura et al., 2013; Jeffries et al., 2016; Lipfert & Doniach, 2007; Rozners, 2010; Lipfert & Doniach, 2007; Tuukkanen, Spilotros, & Svergun, 2017.)

NOTE: Use ultrapure water in all solutions and protocol steps.

Materials

Scattering standard sample (e.g., cytochrome *c*)
Tris·Cl buffer, pH 7.4
Sodium ascorbate
Sodium chloride (NaCl)
Purified gold nanocrystal sample for titration series (10 nmol; see Basic Protocol 1)
Purified double-labeled sample (1 nmol per buffer condition; see Basic Protocol 1)
Purified 2× orthogonal single-labeled samples (1 nmol per buffer condition each; see Basic Protocol 1)
Purified unlabeled sample (1 nmol per buffer condition; see Basic Protocol 1)
Sample buffer (e.g., 10× buffer: 700 mM Tris·Cl, pH 7.4, 100 mM sodium ascorbate, 1.5 M NaCl, and 10 mM $MgCl_2$)
MilliQ water

UV/vis spectrometer
Vortex mixer
−20 or −80°C freezer
Microcentrifuge
0.5 mL centrifugal filters (Amicon Ultra)

NOTE: The protocol below is a suggested workflow for data acquisition at a synchrotron facility. It may vary based on the instrument set up at the facility.

Sample preparation and set up initialization

1. If you are not familiar with the beamline and settings, record data for a standard, e.g., cytochrome *c* (Lipfert et al., 2006) and compare it to benchmark profiles (see Small Angle Scattering Biological Data Bank, *https://www.sasbdb.org/*).

2. Prepare 10× buffer mixture containing Tris·Cl, pH 7.4, and sodium ascorbate together with desired amount of additional salt and other components (e.g., ligands) required for the experiment.

   An example 10× buffer solution for near-physiological conditions is 700 mM Tris·Cl, pH 7.4, 100 mM sodium ascorbate, 1.5 M NaCl, and 10 mM $MgCl_2$.

Use Vortex mixer to ensure proper mixing of components.

It is important to use sodium ascorbate and Tris as radical scavengers in the buffer solution to capture free radicals and thus to reduce radiation damage to your sample during X-ray exposure; this allows longer total exposure times and therefore a better signal-to-noise ratio.

Replace sodium ascorbate stock solution every 3 hr to ensure good scavenger capability. Cover sodium ascorbate with aluminum foil or store in a dark place.

3. Total exposure time depends on the flux at the synchrotron beamline used. In a typical scheme used at beamline 4-2, set total exposure time to 30 sec as a series of ten independent repeats of 3 sec each for data collection. Screen each trace for radiation damage, which can be detected by a gradual change in scattering intensity especially in the low *S*-range region in subsequent X-ray exposures. Exclude scattering profiles with oxidative damage determined from subsequent analysis; do not reuse samples that have been exposed to the X-ray beam. If photon flux is much less than 10^{12} photons/sec, extend total exposure time.

4. Determine concentration of your sample using a UV/vis spectrometer (see Basic Protocol 1, step 29).

   The extinction coefficient of the gold nanocrystals is 0.076 μM/cm at 360 nm.

5. Record a titration series of gold nanocrystal for every beam time as a scattering standard and to obtain nanocrystal size distribution required for further analysis.

   A typical concentration series is 200, 100, 50, and 25 μM gold particles (include higher concentrations if they are necessary for your experiment). The shape of the scattering profile should not change with concentration and the scattering profiles should be superimposable after normalization. Interparticle scattering should be avoided. It can be detected by a concentration-dependent change in the scattering profile at low S.

6. Store gold-labeled samples on dry ice or in a freezer at −20°C until the beamline is set up for experiments. During data acquisition, store sample stock solutions on ice or in fridge.

7. Thaw required amount of sample to room temperature and vortex before measurements.

8. Combine 10× buffer, water, and concentrated sample to achieve a 30 μM final sample concentration (e.g., 0.9 nmol sample in 30 μL). If it is not possible to prepare buffer as a 10× stock solution (e.g., due to solubility limitations) or if concentration of charge of the macromolecule is comparable to the concentration of counterions in solution at very low ionic strength, prepare sample by buffer exchange using centrifugal filter units with a suitable molecular weight cutoff (e.g., 10 kDa Amicon, three repeats, 35 min each).

9. Spin final sample mixture 2 min at 10,000 × *g* at 4°C to sediment out any large contaminant particles.

   Large particles strongly perturb the scattering signal, as the forward scattering intensity of an object grows quadratically with its molecular mass.

Data recording

10. Prepare five samples for data collection.

    The full set of samples consists of an unlabeled molecule sample, two complementary single-labeled samples, one double-labeled macromolecule sample, and a buffer-only sample (Fig. 4). Again, the concentration of the macromolecule should be at least 30 μM in 1× buffer to provide a good signal-to-noise ratio.

Measure scattering profiles of the five samples on the same set up, in direct succession, to keep conditions as similar as possible. Always record buffer-only scattering profile (at least) twice, once at the beginning of the acquisition sequence and once at the end (e.g., buffer, unlabeled molecule sample, two complementary single-labeled samples, double-labeled macromolecule sample, and buffer again).

An automated sample changer installed at the beamline can aid data collection for such a series.

11. Repeat five-sample data acquisition sequence with each macromolecule construct and/or condition in your experiment (e.g., at varying salt concentrations, with and without ligand binding partners).

BASIC PROTOCOL 3

ANALYZING X-RAY SCATTERING INTERFEROMETRY DATA

The data obtained at the beamline can be processed either by applying individual scripts step by step or using the custom written graphical user interface (GUI) in Matlab (au_saxs_gui.m, see Materials). The underlying principles are described in detail by Mathew-Fenn et al. (2008b) and a brief summary follows.

After standard SAXS data processing, as outlined below, the radius distribution of the spherical gold nanocrystals is determined first, from scattering data of the unconjugated gold labels. To accomplish this, the recorded scattering profile of the gold nanocrystals is decomposed into a volume-weighted sum of scattering profiles of spheres with radii ranging from 1 to 100 Å. Using the nanocrystal synthesis protocol described above, (Mathew-Fenn et al., 2008a; Mathew-Fenn et al., 2008b; Shi et al., 2015) the nanocrystal size distribution should have a radius centered at 6 to 7 Å. The obtained radius distribution is then used to calculate the precise basis scattering functions $I(S,D)$, which are the scattering interference patterns for two nanocrystals separated by a fixed center-to-center distance D, where D is varied from 1 to 200 Å. These basis functions will be used to decompose $I_{Au\text{-}Au}(S)$, which is the experimentally determined scattering interference pattern for the two gold nanocrystals attached to the macromolecule. Importantly, the measured scattering profile from the double-labeled macromolecule includes contributions from the macromolecule itself and from the cross-scattering terms between the gold labels and the macromolecule, in addition to $I_{Au\text{-}Au}(S)$.

To extract $I_{Au\text{-}Au}(S)$, the scattering profiles for the quartet of samples must be summed and the summation requires that the profile intensities are accurately scaled relative to each other. The most difficult part of the data processing is finding the correct scaling factors, denoted c_U, c_{A+B}, and c_{Buf} (see Eqn. 2 and 3). To do this, the measured scattering profiles of the double-labeled, single-labeled, and unlabeled constructs and of the buffer ($I_{AB}(S)$, $I_A(S)$, $I_B(S)$, $I_U(S)$, and $I_{Buf}(S)$) are first transformed into interatomic Patterson distributions $P_{AB}(D)$, $P_A(D)$, $P_B(D)$, $P_U(D)$, and $P_{Buf}(D)$, using point-scatterer basis functions (see Eqn. 1). Using the measured scattering profiles and the Patterson distributions, the scaling factors are optimized to satisfy two constraints: (1) that the integral of the sinusoidal function $S \bullet I_{Au\text{-}Au}(S)$ sums to zero (Eqn. 4), and (2) that none of the computed interatomic distances between gold labels are negative, which should not be possible (Eqn. 5). Deviations from these constraints are summed together in a target function T (see Critical Parameters and Troubleshooting), and the scaling factors that minimize T are determined. Using the optimized scaling factors, the measured profiles are summed to obtain $I_{Au\text{-}Au}(S)$.

$$I(S) = \sum_{D_{min}}^{D_{max}} P(D) \cdot \frac{\sin(2\pi DS)}{(2\pi DS)}$$

Equation 1

$$I_{Au-Au}(S) = I_{AB}(S) + c_U \cdot I_U(S) - c_{A+B} \cdot (I_A(S) + I_B(S)) + c_{Buf} \cdot I_{Buf}(S)$$

Equation 2

$$P_{Au-Au}(D) = P_{AB}(D) + c_U \cdot P_U(D) - c_{A+B} \cdot (P_A(D) + P_B(D)) + c_{Buf} \cdot P_{Buf}(D)$$

Equation 3

$$\sum_{S_{\min}}^{S_{\max}} I_{Au-Au}(S) \cdot S \approx 0$$

Equation 4

$$P_{Au-Au}(D) > 0 \; for \; D \in [D_{\min};\; D_{\max}]$$

Equation 5

Finally, $I_{Au\text{-}Au}(S)$ is decomposed into a sum of the $I(S,D)$ basis functions using a maximum entropy algorithm, resulting in a center-to-center distance distribution between the two gold nanocrystals of the double-labeled sample (Eqn. 1). Alternatively, the $I_{Au\text{-}Au}(S)$ decomposition can be performed using non-negative least squares algorithms that are available in most scientific programming languages.

Some beamlines, such as the beamline 4-2 at SSRL and beamline BM29 at ESRF, provide beamline software packages that perform radial averaging and buffer subtraction of scattering profiles, which allows for immediate detection of radiation damage or other technical problems.

Materials

Data set including scattering profiles from:
- Gold nanocrystal sample (e.g., see Basic Protocol 2)
- Double-labeled sample (e.g., see Basic Protocol 2)
- Orthogonal single-labeled samples (two individual scattering profiles; e.g., see Basic Protocol 2)
- Unlabeled sample (e.g., see Basic Protocol 2)
- Sample buffer (e.g., see Basic Protocol 2)
- Computer (minimum requirements: Any Intel or AMD x86-64 processor, 2.5 GB Disk space, 2 GB RAM)
- MATLAB license, graphical user interface (GUI) support from version 2015b guaranteed
- Au-SAXS graphical user interface (GUI) (*https://gitlab.physik.uni-muenchen.de/Jan.Lipfert/AuSAXSGUI.git*)
- Example files (*https://gitlab.physik.uni-muenchen.de/Jan.Lipfert/AuSAXSGUI.git*)

A step-by-step guide on how to obtain the scattering interference pattern $I_{Au\text{-}Au}(S)$, including an example set of data (see exemplary files), is given below.

Data preparation

1. If it has not already been done automatically by the beamline software, reduce 2D scattering matrix into a one-dimensional scattering profile by radial averaging: The output should be a matrix with three columns for scattering momentum transfer vector S, corresponding scattering intensity, and variance/standard deviation in scattering intensity at different radial positions.

   If the incident X-ray beam is polarized, simple radial averaging cannot be performed. See Pauw, 2014 and Svergun & Koch, 2003 for further instruction on how to process the raw data into one-dimensional scattering profiles.

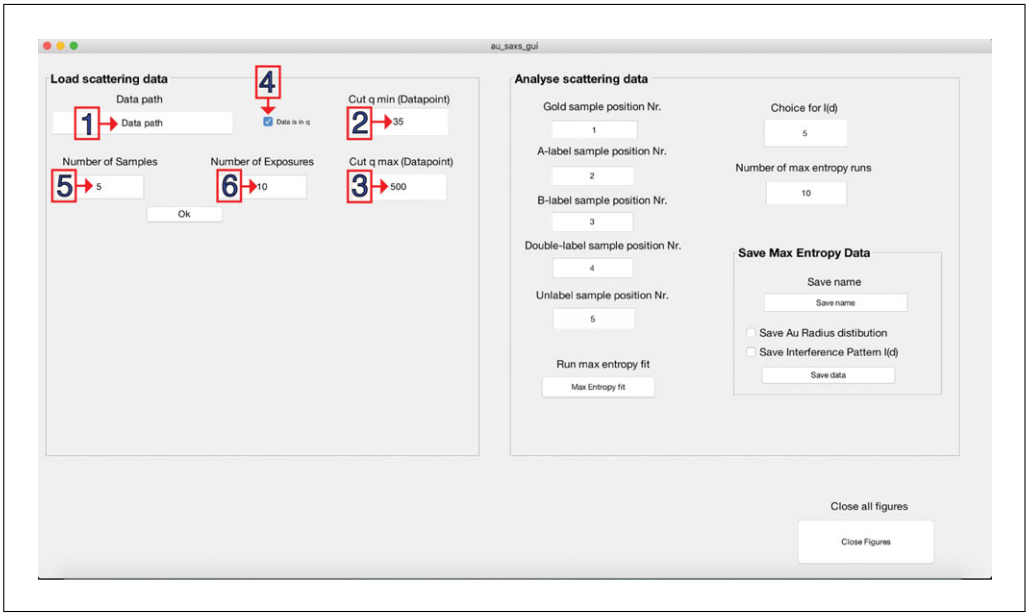

Figure 6 GUI to analyze XSI data. Default values can be modified to adjust the data analysis. The GUI contains panels to specify the following: (1) storage path of the data; (2) adjust q_{min}; (3) adjust q_{max}; (4) specify the momentum transfer convention according to the scattering data; (5) specify the number of samples; and (6) specify the number of exposures per sample. GUI, graphical user interface; XSI, X-ray scattering interferometry.

2. Before starting the custom written au_saxs_gui.m GUI and loading data, rename data files '*_i.dat' where * is any name for the sample and i is the i^{th} exposure, i.e., ranging from 01 to 10 for 10 exposures per molecule ('AB_01.dat', 'AB_02.dat', 'AB_03.dat', . . . , for double-labeled samples). Structure data files so that the scattering momentum vector occupies the first column, the corresponding recorded scattering intensity occupies the second column, and the variance/standard deviation occupies the third column. Separate column entries by a single blank space (see example files for comparison).

3. Initialize GUI by executing 'au_saxs_gui.m' script.

   For proper execution, files 'au_saxs_gui.m', 'au_saxs_gui.fig', and folder 'subroutines' have to be stored in the same directory to allow the main script to find the required subroutines.

Data initialization

4. Enter the full path into the field 'Data path' (Fig. 6, panel 1).

5. Optional: Manipulate scattering momentum by setting the lower (q_{min}; Fig. 6, panel 2) and the upper (q_{max}; Fig. 6, panel 3) limit for the scattering angle.

   The default input is the 35^{th} data point of the initial data up to data point 500; however, this strongly depends on the settings of the beamline and type of sample.

6. Optional: Untick 'Data is in q' box (Fig. 6, panel 4) to switch scattering momentum vector to S [$2sin(\theta/\lambda)$ in $Å^{-1}$].

   The scattering momentum vector is set to q per default [$4\pi\ sin(\theta/\lambda)$, in $Å^{-1}$] since q is the common output format at synchrotrons.

7. Optional: Modify number of samples (Fig. 6, panel 5), if required.

   Default is 5 for one full set of samples.

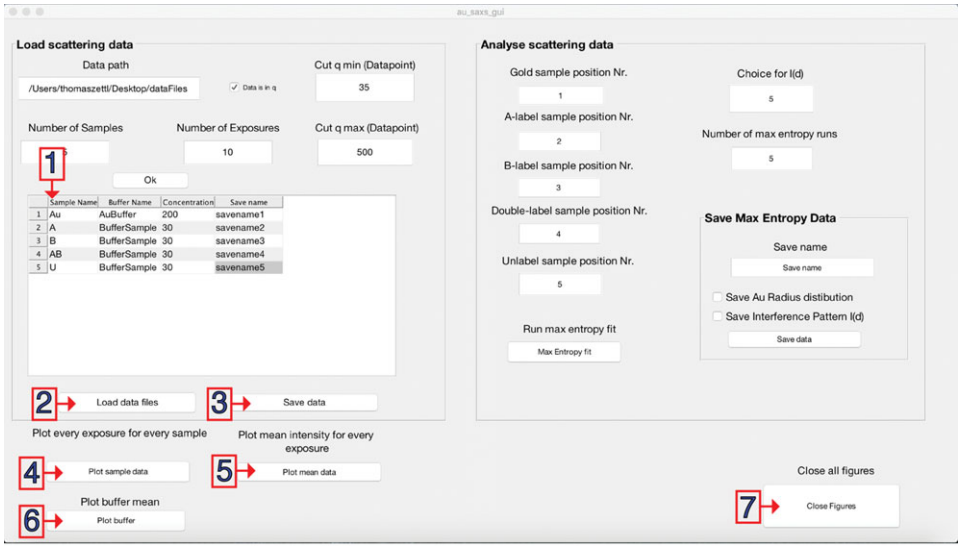

Figure 7 GUI to analyze XSI data. Sample and corresponding buffer filenames, sample concentration, and save as filename (optional) are entered for a full set of data (1). Initial files are loaded (2) and can be plotted (4–6; optional). Unified and truncated scattering profiles can be saved (3). All figures except for the main GUI can be closed (7). GUI, graphical user interface; XSI, X-ray scattering interferometry.

8. Optional: Change number of exposures (Fig. 6, panel 6) to number of files read per sample.

   Default is 10 as described in the protocol.

9. Press 'Ok' to initialize script.

   New buttons and a table will appear on the left side of the window (see Fig. 7).

10. Enter all sample names in the first column of table (see Fig. 7).

11. Enter corresponding buffer in the second column.

12. Enter determined sample concentration (in μM) in the third column.

13. Optional: Enter save as filenames for all samples in the fourth column (see Fig. 7, panel 1).

14. Load data files by pressing 'Load data files' button (Fig. 7, panel 2).

    If successful, a window indicating 'All data was successfully imported!' will be displayed. In case a file could not be found or a wrong sample name was entered a warning will be generated noting the incorrect position. If the concentration was not entered properly, a similar error message will be produced.

Data testing

15. Optional: Inspect scattering profiles for all samples either by plotting all exposures per sample as individual traces into one figure per sample (Fig. 7, panel 4) or by plotting the averaged profile over all exposures (Fig. 7, panel 5).

16. Optional: Overlay these plots by pressing 'Plot sample data' first and 'Plot mean data' second.

17. Optional: Plot averaged buffer for each sample using 'Plot buffer' button (Fig. 7, panel 6).

Figure 8 GUI to analyze XSI data. Panels to set the sample type (1) according to the position in the sample loading table (2). The choice of minimization function (3) and the number of individual maximum entropy fitting runs (4) can be specified prior to starting the routine (5). Final distance distributions [Au radius distribution and interference pattern *I(S)*, optional] can be saved (6). GUI, graphical user interface; XSI, X-ray scattering interferometry.

18. Optional: In case too many MATLAB figures are open, close all except the main GUI by pressing 'Close Figures' button (Fig. 7, panel 7).

19. Optional: Save scattering data for all loaded samples as 'YOURFILENAME-HERE_scattering_data.mat' files using 'Save data' (Fig. 7, panel 3).

Maximum entropy fitting

20. Set options by specifying sample positions according to row number in the table (see Fig. 8, panel 1), i.e., Gold sample position = 1, A-label sample position = 2, B-label sample position = 3, Double-label sample position = 4, and Unlabeled sample position = 5.

21. Optional: Change number of runs for the maximum entropy fit (default is 10, Fig. 8, panel 3); the output is one high-resolution distance distribution per run. Average distributions over all runs to obtain the final distribution. Lower the number to shorten the required computational time for test purposes.

22. Optional: Change minimization function option to extract gold-gold interference pattern $I_{Au\text{-}Au}(S)$ (default is 5 ranging from 1 to 7, see Critical Parameters and Troubleshooting for more details, Fig. 8, panel 2).

23. Start maximum entropy fit by pressing 'Max Entropy fit' (Fig. 8, panel 4).

    A progress bar in the lower left corner of the main GUI will display progress and vanish as soon as the calculations are finished. Three new figures showing the gold nanocrystal radius distribution, gold-gold scattering interference signal $I_{Au\text{-}Au}(S)$, and final distance distribution determined via maximum entropy fit will appear (Fig. 5).

24. Enter a save as filename (e.g., 'YOURSAMPLENAME') for distance distribution (Save Max Entropy Data, right side, Fig. 8, panel 6).

    This file has one column per maximum entropy run (10 as default) and a 1 Å spaced distance distribution ranging from 1 to 200 Å.

25. After setting the name, press 'Save data' (Save Max Entropy Data, right side); a file in the format of 'YOURSAMPLENAME_Distance_Distribution.mat' will be saved in your current folder.

Optional: Save gold nanocrystal radius distribution and/or the Au-Au interference pattern$I_{Au\text{-}Au}(S)$ by checking individual boxes.

REAGENTS AND SOLUTIONS

Ammonium acetate stock buffer

2 M ammonium acetate (Sigma Aldrich, cat. no. A1542), pH 5.6, in ultrapure water
Store up to 1 month at 4°C

Fast protein liquid chromatography (FPLC) cleaning solution

100 mM dithiothreitol (DTT; Thermo Fisher Scientific, cat. no. R0861), 20 mM Tris·Cl, pH 8.0
Prepare fresh

High salt ammonium acetate buffer

1.5 M NaCl, 20 mM ammonium acetate, pH 5.6, diluted from ammonium acetate stock buffer (see recipe)
Store up to 1 month at room temperature

High salt borate buffer

1.5 M NaCl, 20 mM sodium borate, pH 7.8
Store up to 1 month at room temperature

Low salt ammonium acetate buffer

10 mM NaCl, 20 mM ammonium acetate, pH 5.6, diluted from ammonium acetate stock buffer (see recipe)
Store up to 1 month at room temperature

Low salt borate buffer

10 mM NaCl, 20 mM sodium borate, pH 7.8
Store up to 1 month at room temperature

Size exclusion running buffer

150 mM ammonium acetate, pH 5.6, diluted from ammonium acetate stock buffer (see recipe)
Prepare fresh

COMMENTARY

Background Information

Basic principle

X-ray scattering interferometry (XSI) measures the interference between X-rays scattered by two gold nanocrystals attached to a macromolecule or macromolecular complex. In addition to the desired nanocrystal-nanocrystal interference pattern, the recorded scattering profile includes signals from scattering interference between pairs of atoms in the macromolecule, between pairs of atoms in a single gold nanocrystal, and between pairs of atoms in the macromolecule and the nanocrystal. It is critical to isolate the nanocrystal-nanocrystal interference pattern from the other signals, since it alone contains direct information about the distribution of center-to-center distances between the nanocrystal probes. This distance distribution is obtained by Fourier transformation of the nanocrystal-nanocrystal interference pattern. A limitation of XSI is the time-consuming sample preparation, which only allows low throughput. The strength of XSI is that it provides accurate ensemble distance distributions even for dynamic and rapidly interconverting macromolecules, with precise external distance calibration

from the wavelength of the incident X-ray radiation. These absolute distances can be directly compared to high-resolution atomic models from NMR spectroscopy, X-ray crystallography, electron microscopy, or molecular simulations. XSI is very complementary to spectroscopic rulers such as fluorescence resonance energy transfer (FRET), and combined application of the two types of ruler presents interesting possibilities.

Application of XSI to nucleic acids and nucleic acid-protein complexes

Over the last decade, XSI has been successfully applied to dsDNA (Mathew-Fenn et al., 2008a; Mathew-Fenn et al., 2008b; Shi et al., 2013), nucleic acid two-way junctions (Shi et al., 2014; Shi et al., 2017), and nucleic acid-protein complexes (Hura et al., 2013; Shi et al., 2016). It has been used to characterize the stretching (Mathew-Fenn et al., 2008a), twisting, and bending elasticity (Shi et al., 2013) of short DNA helices. Recently, DNAs and RNAs containing bulge sequences have been studied, revealing a complex multistate behavior that responds to the solution conditions (Shi et al., 2014; Shi et al., 2017). In addition, an RNA-protein complex has been investigated (Shi et al., 2016). The use of metal clusters as structural probes makes XSI scalable. Clusters of a single or few metal atoms are suitable for smaller systems (Miake-Lye, Doniach, & Hodgson, 1983; Vainshtein et al., 1980), while clusters containing thousands of atoms can be applied to large complexes (Hura et al., 2013). While small labels are desirable for accurate position determination and to minimize the perturbation of the molecular structures of interest, for large macromolecules noise from the macromolecule signal can prevent accurate Au-Au distance determination and larger gold nanocrystals are needed.

A variant of the XSI interference technique has recently been demonstrated, which uses anomalous SAXS to directly extract the $I_{Au\text{-}Au}(S)$ interference pattern from double-labeled samples without the need to record any single- or unlabeled samples (Zettl et al., 2016). This approach will be in particular relevant for samples where it is difficult to prepare matching single-labeled constructs.

Future directions

The future challenges and possible applications of XSI are numerous. Using the labeling and measuring protocols described here, the conformational ensembles of diverse nucleic-acid motifs and nucleic acid-protein complexes can be determined, including their dependence on environmental conditions. The application of XSI to nanocrystal-labeled proteins will provide a new window into protein conformational ensembles, especially for folding intermediates and intrinsically disordered proteins (IDPs). Finally, the application of next-generation free electron lasers with very high brilliance could allow measurement of correlated and time resolved distances between multiple sites in a macromolecule, with a time resolution of tens of femtoseconds (Arnlund et al., 2014; Bada, Walther, Arcangioli, Doniach, & Delarue, 2000; Ball, 2017; Doniach, 2000; Mendez et al., 2014; Mendez et al., 2016; Pile, 2010; Schenk et al., 2016; Smolsky et al., 2007).

Critical Parameters and Troubleshooting

A key aspect for successful determination of intramolecular distance distributions using XSI is sample preparation and sample quality. The samples have to be highly homogeneous (e.g., no free gold nanocrystals in solution) and free of degradation. Furthermore, buffer matching should be as precise as possible. Given that XSI requires very pure samples, it is crucial to monitor and adjust the sample quality in advance. To assess sample quality, non-denaturing polyacrylamide gel electrophoresis (PAGE; Andrus & Kuimelis, 2001) or ion exchange chromatography (Sinha & Jung, 2015) can be used. In addition, we advise the user to perform additional experiments (e.g., UV melting curve analysis and circular dichroism spectroscopy (see Mathew-Fenn et al., 2008b) to test whether the attached gold nanocrystals perturb the macromolecular structure and ideally to establish the absence of such perturbations. In addition, G/C base pairs at the blunt ends of nucleic acids are recommended to avoid fraying of nucleic acid helices.

Another requirement of the XSI technique is high signal-to-noise SAXS measurements of the set of double-labeled, single-labeled, and unlabeled samples. We recommend testing the beamline set up with readily available and characterized samples, such as cytochrome *c*, BSA, or unlabeled double-stranded DNA (see *https://www.sasbdb.org/*), prior to an XSI experiment. The measurements of the scattering standard samples should give the expected scattering profiles and radii of gyration. Such

control measurements should be repeated every time the SAXS set up is reconfigured and can help to detect misalignment of the X-ray beam, problems with parasitic scattering, and issues in detector calibration and conversion of the detector image to the 1D scattering profile.

Sometimes, it is not possible to obtain a high-quality distance distribution even when the sample preparation and recording are performed correctly. Potential issues can be the scattering intensity ratio between the molecule and the gold nanocrystals resulting in poor signal-to-noise ratio for larger nucleic acid structures or highly flexible molecules. Additionally, larger metal clusters should be considered in cases of insufficient signal-to-noise.

Some additional factors that can affect the data analysis and the final distance distribution and how to address them are listed below:

(1) Parasitic X-ray scattering in the low *S*-range as well as high noise levels in the high *S*-range can occur and influence the normalization and calculation of the interference pattern.

Solution: We recommend measuring scattering standards to test and optimize the X-ray set up before running XSI samples. In the post-processing of recorded data sets, one should vary S_{min} and/or S_{max} multiple times and test the impact on the resulting distance distribution.

(2) The scattering profile can be distorted by various effects such as radiation damage or bubbles passing through the X-ray beam.

Solution: Carefully compare scattering profiles from subsequent exposures of the same sample and exclude all traces differing from the majority.

(3) Choice of minimization function *T* to extract to gold-gold scattering interference Patterson $I_{Au\text{-}Au}(S)$ where *S* is the scattering momentum vector, $I_U(S)$ the scattering intensity vector of the unlabeled sample, and $P_{Au\text{-}Au}(D)$ the label-label interatomic radial Patterson function. The gold-gold scattering interference Patterson $I_{Au\text{-}Au}(S)$ looks like a sinc (i.e., a decaying oscillation) function (see Fig. 5B). If that is not the case, redo the analysis with a different choice for *T*. For guidance and comparison, a multiple interference pattern of static and dynamic systems can be found in the published literature (Mathew-Fenn et al., 2008a; Mathew-Fenn et al., 2008b; Shi et al., 2014; Shi et al., 2015; Shi et al., 2013; Shi et al., 2016; Shi et al., 2017) for comparison.

Choice 1: Original function described by Mathew-Fenn, Das, and coworkers (Equation 6; Mathew-Fenn et al., 2008b)

$$T = \frac{\sum_S \left[I_{Au-Au}(S)\cdot S^2\right]^2}{\sum_S \left[I_U(S)\cdot S^2\right]^2} + \frac{\sum_{P_{Au-Au}(D)<0} P^2_{Au-Au}(D)}{\sum_D P^2_U(D)}$$

Equation 6

The minimization function is similar to Kratky analysis, dividing out the decay of the scattering intensity by weighting the interference pattern $I_{Au\text{-}Au}(S)$ by S^2.

Choice 2 (Equation 7):

$$T = \frac{\sum_S \left[I_{Au-Au}(S)\cdot S\right]^2}{\sum_S \left[I_U(S)\cdot S^2\right]^2} + \frac{\sum_{P_{Au-Au}(D)<0} P^2_{Au-Au}(D)}{\sum_D P^2_U(D)}$$

Equation 7

Au-Au interference pattern is only weighted by *S*, thus $I_{Au\text{-}Au}(S)$ minimization is dominated by values in the low *S*-region.

Choice 3 (Equation 8):

$$T = \frac{\sum_S \left[I_{Au-Au}(S)\cdot S\right]^2\cdot S}{\sum_S \left[I_U(S)\cdot S^2\right]^2} + \frac{\sum_{P_{Au-Au}(D)<0} P^2_{Au-Au}(D)}{\sum_D P^2_U(D)}$$

Equation 8

The low *S*-region of $I_{Au\text{-}Au}(S)$ is weighted more compared to choice 1, but less than choice 2 and the high *S*-region is weighted less compared to choice 1 and more than choice 2.

Choice 4 (Equation 9):

$$S_{min} = Min(S)\ for\ S_i \in S \geq 0.06\,\text{Å}^{-1}$$

$$S_2 = \frac{S_i}{S_{\min}}\ for\ S_i \in S < 0.06\,\text{Å}^{-1}$$

$$S_2 = 1\ for\ S_i \in S \geq 0.06^{-1}$$

$$T = \frac{\sum_{S_2}\left[\sum_S \left[I_{Au-Au}(S)\cdot S\right]^2\cdot S_2\right]}{\sum_S \left[I_U(S)\cdot S^2\right]^2} + \frac{\sum_{P_{Au-Au}(D)<0} P^2_{Au-Au}(D)}{\sum_D P^2_U(D)}$$

Equation 9

Very similar to choice 3 in the high *S*-range, however the high *S*-range is weighted slightly less.

Choice 5 (Equation 10):

$$S_{\min,1} = Min(S_i)\, for\, S_i \in S < 0.06 \mathring{A}^{-1}$$

$$S_{\min,2} = Min(S_i)\, for\, S_i \in S \geq 0.06 \mathring{A}^{-1}$$

$$S_2 = \frac{S_i}{S_{\min,2}}\, for\, S_i \in S < 0.06 \mathring{A}^{-1}$$

$$S_2 = \frac{S_{\min,1}}{S_{\min,2}}\, for\, S_i \in S \geq 0.06 \mathring{A}^{-1}$$

$$T = \frac{\sum_{S_2}[\sum_S [I_{Au-Au}(S) \cdot S]^2 \cdot S_2]}{\sum_S [I_U(S) \cdot S^2]^2} + \frac{\sum_{P_{Au-Au}(D)<0} P^2_{Au-Au}(D)}{\sum_D P^2_U(D)}$$

Equation 10

Very similar to choice 3 in the high *S*-range, however the high *S*-range contributes very little.

Choice 6 (Equation 11):

$$T = \frac{\sum_{P_{Au-Au}(D)<0} P^2_{Au-Au}(D)}{\sum_D P^2_U(D)}$$

Equation 11

Only $P_{Au\text{-}Au}(D)$ is minimized such that it does not include negative values, however the oscillations in $I_{Au\text{-}Au}(S)$ are not minimized.

Choice 7 (Equation 12):

$$T = \frac{\sum_{S>0.06\ \mathring{A}^{-1}} [I_{Au-Au}(S) \cdot S]^2}{\sum_S [I_U(S) \cdot S^2]^2}$$

Equation 12

Only the high *S*-region of $I_{Au\text{-}Au}(S)$ is minimized similar to a base line correction. $P_{Au\text{-}Au}(D)$ is allowed to include negative values.

Anticipated Results

The protocols presented here guide experimenters to synthesize gold nanocrystals, prepare and purify gold-labeled nucleic acid constructs, measure a complete set of small-angle X-ray scattering (SAXS) profiles for the labeled samples, and extract intra-molecular distance distributions from the data.

Following the steps presented in Basic Protocol 1, the following yields are typically achieved: The gold nanocrystal synthesis yields 3 to 10 µmol after final purification and desalting; the gold nanocrystal attachment protocol results in a yield of 2 to 3 nmol of each labeled and purified duplexed construct for XSI measurements.

Basic Protocols 2 and 3 describe a reliable procedure to obtain SAXS scattering profiles and a step-by-step guide on how to use the provided GUI to extract high-resolution distance distributions on an absolute scale (Fig. 5C). To study more complex macromolecules that can undergo conformational changes or to disentangle complex geometries and motions, multiple label pairs attached to different positions can be prepared and analyzed to increase the information content.

Time Considerations

The sample preparation should be performed about 1 month in advance. The final purification (steps 28 and 29 in Basic Protocol 1) should be carried out shortly before data collection to ensure the highest sample quality. The overall data acquisition time varies depending on the synchrotron facility. Typically, data collection for several samples can take up to 1 day. The run time of the custom-written Matlab routine to compute distance distributions that is provided with this protocol is on the order of minutes on a standard personal computer. The total time required for data analysis, validation, and interpretation is variable and can take several days.

Acknowledgements

We thank T. Matsui and T. Weiss at beamline 4-2 of the Stanford Synchrotron Radiation Lab (SSRL) and Gabriele Giachin at beamline BM29 of the European Synchrotron Radiation Facility (ESRF) for technical support in synchrotron small-angle X-ray scattering experiments, and Steffen Sedlak for helpful discussions and comments on the manuscript. This work was supported by Sonderforschungsbereich (SFB) 1032 and the National Institutes of Health (NIH) P01 GM066275 to D.H. and DP-DO000429 to P.A.B.H.

Literature Cited

Ackerson, C. J., Sykes, M. T., & Kornberg, R. D. (2005). Defined DNA nanoparticle conjugates. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 13383–13385. doi: 10.1073/pnas.0506290102.

Alivisatos, A. P., Johnsson, K. P., Peng, X., Wilson, T. E., Loweth, C. J., Bruchez, M. P., & Schultz, P. G. (1996). Organization of 'nanocrystal molecules' using DNA. *Nature*, *382*, 609–611. doi: 10.1038/382609a0.

Andrus, A. & Kuimelis, R. G. (2001). Polyacrylamide gel electrophoresis (PAGE) of synthetic nucleic acids. *Current Protocols in*

Nucleic Acid Chemistry, *1*, 10.4.1–10.4.10. doi: https://doi.org/10.1002/0471142700.nc1004s01.

Arnlund, D., Johansson, L. C., Wickstrand, C., Barty, A., Williams, G. J., Malmerberg, E., ... Neutze, R. (2014). Visualizing a protein quake with time-resolved X-ray scattering at a free-electron laser. *Nature Methods*, *11*(9), 923–926. doi: 10.1038/nmeth.3067.

Aubin-Tam, M. E., & Hamad-Schifferli, K. (2005). Gold nanoparticle-cytochrome c complexes: The effect of nanoparticle ligand charge on protein structure. *Langmuir*, *21*, 12080–12084. doi: 10.1021/la052102e.

Aubin-Tam, M. E., Hwang, W., & Hamad-Schifferli, K. (2009). Site-directed nanoparticle labeling of cytochrome c. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 4095–4100. doi: 10.1073/pnas.0807299106.

Azubel, M., & Kornberg, R. D. (2016). Synthesis of water-soluble, thiolate-protected gold nanoparticles uniform in size. *Nano Letters*, *16*(5), 3348–3351. doi: 10.1021/acs.nanolett.6b00981.

Bada, M., Walther, D., Arcangioli, B., Doniach, S., & Delarue, M. (2000). Solution structural studies and low-resolution model of the *Schizosaccharomyces pombe* sap1 protein. *Journal of Molecular Biology*, *300*(3), 563–574. doi: 10.1006/jmbi.2000.3854 S0022-2836(00)93854-3 [pii]

Ball, P. (2017). Europe's X-ray laser fires up. *Nature*, *548*, 507–508. doi: 10.1038/548507a.

Burke, J. E. & Butcher, S. E. (2012). Nucleic acid structure characterization by small angle x-ray scattering (SAXS). *Current Protocols in Nucleic Acid Chemistry*, *51*, 7.18.1–7.18.18. doi: https://doi.org/10.1002/0471142700.nc0718s51.

Doniach, S. (2000). Fourth-generation X-ray sources: Some possible applications to biology. *Journal of Synchrotron Radiation*, *7*(Pt 3), 116–120. doi: S0909049500004143 [pii] 10.1107/S0909049500004143.

Doniach, S., & Lipfert, J. (2012). Small and wide angle X-ray scattering from biological macromolecules and their complexes in solution. *Comprehensive Biophysics*, *1*, 376–397. doi: 10.1016/B978-0-12-374920-8.00122-3.

Dyer, K. N., Hammel, M., Rambo, R. P., Tsutakawa, S. E., Rodic, I., Classen, S., ... Hura, G. L. (2014). High-throughput SAXS for the characterization of biomolecules in solution: A practical approach. *Methods in Molecular Biology*, *1091*, 245–258. doi: 10.1007/978-1-62703-691-7_18.

Feynman, R. P., Leighton, R. B., & Sands, M. (1977). *The Feynman lectures on physics: Mainly mechanics, radiation, and heat. Sixth edition*. New York: Pearson.

Fischer, N., Konevega, A. L., Wintermeyer, W., Rodnina, M. V., & Stark, H. (2010). Ribosome dynamics and tRNA movement by time-resolved electron cryomicroscopy. *Nature*, *466*(7304), 329–333. doi: 10.1038/nature09206.

Frauenfelder, H. (1989). New looks at protein motions. *Nature*, *338*, 623-624. doi: 10.1038/338623a0.

Grishaev, A. (2012). Sample preparation, data collection, and preliminary data analysis in biomolecular solution X-ray scattering. *Current Protocols in Protein Science*, *70*, 17.14.1–17.14.18. doi: https://doi.org/10.1002/0471140864.ps1714s70.

Hura, G. L., Budworth, H., Dyer, K. N., Rambo, R. P., Hammel, M., McMurray, C. T., & Tainer, J. A. (2013). Comprehensive macromolecular conformations mapped by quantitative SAXS analyses. *Nature Methods*, *10*, 453. doi: 10.1038/nmeth.2453.

Hura, G. L., Tsai, C. L., Claridge, S. A., Mendillo, M. L., Smith, J. M., Williams, G. J., ... Tainer, J. A. (2013). DNA conformations in mismatch repair probed in solution by X-ray scattering from gold nanocrystals. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(43), 17308–17313. doi: 10.1073/pnas.1308595110.

Jeffries, C. M., Graewert, M. A., Blanchet, C. E., Langley, D. B., Whitten, A. E., & Svergun, D. I. (2016). Preparing monodisperse macromolecular samples for successful biological small-angle X-ray and neutron scattering experiments. *Nature Protocols*, *11*(11), 2122–2153. doi: 10.1038/nprot.2016.113.

Lipfert, J., & Doniach, S. (2007). Small-angle X-ray scattering from RNA, proteins, and protein complexes. *Annual Review of Biophysics and Biomolecular Structure*, *36*, 307–327. doi: 10.1146/annurev.biophys.36.040306.132655.

Lipfert, J., Millett, I. S., Seifert, S., & Doniach, S. (2006). Sample holder for small-angle X-ray scattering static and flow cell measurements. *Review of Scientific Instruments*, *77*(4), 26–28. doi: 10.1063/1.2194484.

Mathew-Fenn, R. S., Das, R., & Harbury, P. A. (2008a). Remeasuring the double helix. *Science*, *322*, 446–449. doi: 10.1126/science.1158881.

Mathew-Fenn, R. S., Das, R., Silverman, J. A., Walker, P. A., & Harbury, P. A. (2008b). A molecular ruler for measuring quantitative distance distributions. *PLoS One*, *3*(10), e3229. doi: 10.1371/journal.pone.0003229.

Mendez, D., Lane, T. J., Sung, J., Sellberg, J., Levard, C., Watkins, H., ... Doniach, S. (2014). Observation of correlated X-ray scattering at atomic resolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1647), 20130315. doi: 10.1098/rstb.2013.0315.

Mendez, D., Watkins, H., Qiao, S., Raines, K. S., Lane, T. J., Schenk, G., ... Doniach, S. (2016). Angular correlations of photons from solution diffraction at a free-electron laser encode molecular structure. *IUCrJ*, *3*(Pt 6), 420–429. doi: 10.1107/S2052252516013956.

Miake-Lye, R. C., Doniach, S., & Hodgson, K. O. (1983). Anomalous X-ray scattering from terbium-labeled parvalbumin in solution. *Biophysical Journal*, *41*(3), 287–292.

doi: S0006-3495(83)84440-3 [pii] 10.1016/S0006-3495(83)84440-3.

Müller, J. P., Löf, A., Mielke, S., Obser, T., Brützel, L. K., Vanderlinden, W., ... Benoit, M. (2016). pH-dependent interactions in dimers govern the mechanics and structure of von Willebrand factor. *Biophysical Journal*, *111*(2), 312–322. doi: 10.1016/j.bpj.2016.06.022.

Pauw, B. R. (2014). Everything SAXS: Small-angle scattering pattern collection and correction. *Journal of Physics: Condensed Matter*, *26*, 239501. doi: 10.1088/0953-8984/26/23/239501.

Pile, D. F. P. (Editor) (2010). Building X-ray lasers. Interview with Paul Emma. *Nature Photonics*, *4*, 802–803. doi: 10.1038/nphoton.2010.276.

Rozners, E. (2010). Determination of nucleic acid hydration using osmotic stress. *Current Protocols in Nucleic Acid Chemistry*, *43*, 7.14.1–7.14.13. doi: 10.1002/0471142700.nc0714s43.

Schaaff, T. G., Knight, G., Shafigullin, M. N., Borkman, R. F., & Whetten, R. L. (1998). Isolation and selected properties of a 10.4 kDa gold:glutathione cluster compound. *Journal of Physical Chemistry B*, *102*(52), 10643–10646. doi: 10.1021/jp9830528.

Schenk, G., Krajina, B., Spakowitz, A., & Doniach, S. (2016). Potential for measurement of the distribution of DNA folds in complex environments using correlated X-ray scattering. *Modern Physics Letters B* , *30*(8), 1650117. doi: 10.1142/S0217984916501177.

Shi, X., Beauchamp, K. A., Harbury, P. B., & Herschlag, D. (2014). From a structural average to the conformational ensemble of a DNA bulge. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(15), E1473–E1480. doi: 10.1073/pnas.1317032111.

Shi, X., Bonilla, S., Herschlag, D., & Harbury, P. (2015). Quantifying nucleic acid ensembles with X-ray scattering interferometry. *Methods in Enzymology*, *558*, 75–97. doi: 10.1016/bs.mie.2015.02.001.

Shi, X., Herschlag, D., & Harbury, P. A. (2013). Structural ensemble and microscopic elasticity of freely diffusing DNA by direct measurement of fluctuations. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(16), E1444–E1451. doi: 10.1073/pnas.1218830110.

Shi, X., Huang, L., Lilley, D. M., Harbury, P. B., & Herschlag, D. (2016). The solution structural ensembles of RNA kink-turn motifs and their protein complexes. *Nature Chemical Biology*, *12*, 146–152. doi: 10.1038/nchembio.1997.

Shi, X., Walker, P., Harbury, P. A., & Herschlag, D. (2017). Determination of the conformational ensemble of the TAR RNA by X-ray scattering interferometry. *Nucleic Acids Research*, *45*(8), e64. doi: 10.1093/nar/gkw1352.

Sinha, N. D. & Jung, K. E. (2015). Analysis and purification of synthetic nucleic acids using HPLC. *Current Protocols in Nucleic Acid Chemistry*, *61*, 10.5.1–10.5.39. doi: 10.1002/0471142700.nc1005s61.

Smolsky, I. L., Liu, P., Niebuhr, M., Ito, K., Weiss, T. M., & Tsuruta, H. (2007). Biological small-angle X-ray scattering facility at the Stanford synchrotron radiation laboratory. *Journal of Applied Crystallography*, *40*, S453–S458. doi: 10.1107/s0021889807009624.

Svergun, D., & Koch, M. (2003). Small-angle scattering studies of biological macromolecules in solution. *Reports on Progress in Physics*, *66*, 1735–1782. doi: 10.1088/0034-4885/66/10/R05.

Tuukkanen, A. T., Spilotros, A., & Svergun, D. I. (2017). Progress in small-angle scattering from biological solutions at high-brilliance synchrotrons. *IUCrJ*, *4*(Pt 5), 518–528. doi: 10.1107/S2052252517008740.

Vainshtein, B. K., Feigin, L. A., Lvov, Y. M., Gvozdev, R. I., Marakushev, S. A., & Likhtenshtein, G. I. (1980). Determination of the distance between heavy-atom markers in haemoglobin and histidine decarboxylase in solution by small-angle X-ray scattering. *Febs Letters*, *116*(1), 107–110. doi: 10.1016/0014-5793(80)80539-4.

Zettl, T., Mathew, R. S., Seifert, S., Doniach, S., Harbury, P. A., & Lipfert, J. (2016). Absolute intra-molecular distance measurements with Ångström-resolution using anomalous small-angle X-ray scattering. *Nano Letters*, *16*(9), 5353–5357. doi: 10.1021/acs.nanolett.6b01160.

Zettl, T., Mathew, R. S., Shi, X., Doniach, S., Herschlag, D., Harbury, P. A. B., & Lipfert, J., (2018). Gold nanocrystal labels provide a sequence -to -3D structure map in SAXS reconstructions. *Science Advances*, 4, eaar4418. In press.

Internet Resources

https://eu.idtdna.com/calc/analyzer

Online tool to calculate melting temperature, extinction coefficient for a given DNA sequence.

https://www.sasbdb.org/

Small Angle Scattering Biological Data Bank providing experimental, measured SAXS profiles.

https://gitlab.physik.uni-muenchen.de/Jan.Lipfert/AuSAXSGUI.git

Git repository providing the Matlab AuSAXS GUI and example data files.