CHAPTER THREE

Quantifying Nucleic Acid Ensembles with X-ray Scattering Interferometry

Xuesong Shi*, Steve Bonilla†, Daniel Herschlag*,[1], Pehr Harbury*,[1]
*Department of Biochemistry, Stanford University, Stanford, California, USA
†Department of Chemical Engineering, Stanford University, Stanford, California, USA
[1]Corresponding authors: e-mail address: herschla@stanford.edu; harbury@stanford.edu

Contents

Abstract

The conformational ensemble of a macromolecule is the complete description of the macromolecule's solution structures and can reveal important aspects of macromolecular folding, recognition, and function. However, most experimental approaches determine an average or predominant structure, or follow transitions between states that each can only be described by an average structure. Ensembles have been extremely difficult to experimentally characterize. We present the unique advantages and capabilities of a new biophysical technique, X-ray scattering interferometry (XSI), for probing and quantifying structural ensembles. XSI measures the interference of scattered waves from two heavy metal probes attached site specifically to a macromolecule. A Fourier transform of the interference pattern gives the fractional abundance of different probe separations directly representing the multiple conformation states populated by the macromolecule. These probe–probe distance distributions can then be used to define the structural ensemble of the macromolecule. XSI provides accurate, calibrated distance in a model-independent fashion with angstrom scale sensitivity in distances. XSI data can be compared in a straightforward manner to atomic coordinates

Methods in Enzymology, Volume 558
ISSN 0076-6879
http://dx.doi.org/10.1016/bs.mie.2015.02.001

determined experimentally or predicted by molecular dynamics simulations. We describe the conceptual framework for XSI and provide a detailed protocol for carrying out an XSI experiment.

1. ENSEMBLES AND ENERGY SURFACES

The first high-resolution DNA and protein structures reported in the 1950s suggested that biological macromolecules fold into static, well-defined conformations. Fifty years later, it is clear that a continuum of order exists, ranging from rigid structures to intrinsically disordered ones. Moreover, even the most rigid macromolecules populate a dynamic ensemble of conformational states. Measuring and quantitatively modeling these ensembles remains an open challenge in biophysics.

Why care about structural ensembles? At the most basic level, they are what macromolecules look like at atomic resolution. This information is then essential for developing an accurate mental picture of macromolecules, and an intuition for how they behave. Indeed, most structured nucleic acids and proteins must adopt a multiplicity of conformations to carry out their biological functions. The ribosome is a prime example (Fischer, Konevega, Wintermeyer, Rodnina, & Stark, 2010; Frank & Agrawal, 2000). Beyond that, ensembles play a central role in biological interactions, particularly in phenomena such as allostery and conformational capture. They are linked to macromolecular folding, embodied in concepts such as folding funnels and folding intermediates (Fenwick, Esteban-Martin, & Salvatella, 2011; Salmon, Yang, & Al-Hashimi, 2014). Ensembles determine the mechanical properties of macromolecules, such as their elastic response to stretching, bending, and twisting (Mathew-Fenn, Das, & Harbury, 2008; Olson, Gorin, Lu, Hock, & Zhurkin, 1998; Shi, Herschlag, & Harbury, 2013). In this capacity, understanding ensembles will also be important in the area of engineered nanostructures (Feldkamp & Niemeyer, 2006). Perhaps most importantly, experimental characterization of ensembles will lead to conceptual advances in the physical potentials that we use for molecular modeling. Computed conformational ensembles are the simplest and most meaty predictions of molecular dynamics simulations, and experimental measurements of ensemble structure are the corresponding ground truth.

Nucleic acid conformational ensembles are particularly rich and interesting. In structured RNA, the stable secondary structures (helices) are

connected through a variety of types of linkers. Each linker type imposes a different tertiary structure, or ensemble, and the resulting ensembles can be diffused, with highly populated conformations differing dramatically from one another. In addition to their intrinsic structural heterogeneity, cellular metabolism imposes diverse conformations of nucleic acids, such as D-loops, mismatched bulges, kinks, hairpins, and tight spirals (Bacolla & Wells, 2004; Palecek, 1991). The energy landscape over this complex conformational space impacts basic biology. Sequence preferences for deformed DNA states, for example, influence the patterning of nucleosomes on the genome and the binding specificity of transcription factors (Kaplan et al., 2009; Rohs et al., 2010). Despite the importance of nucleic acid ensembles, our understanding of them is quite limited.

2. THE EXPERIMENTAL CHALLENGE

Ensembles are extremely hard to study experimentally. In part, this is due to technical challenges. First, the direct study of ensembles requires a sample in aqueous solution. This is incompatible with high-resolution, solid-state approaches such as X-ray crystallography. Second, most biophysical tools provide ensemble-averaged measurements of an observable. Thus, the measurement does not provide direct information about the individual substructures that make up the ensemble. Third, for many types of measurements, extracting structural information from the raw data is indirect and model dependent. This leads to uncertainty in the conclusions, and means that many different models of the ensemble are compatible with the data. Finally, it is sometimes difficult to compare a measured observable to results from a molecular simulation, because computing the predicted value for the observable is impractical or impossible. Despite these challenges, a variety of physical techniques have been applied to the ensemble problem. The different techniques yield complementary types of information. Here, we highlight two particular approaches that have been productive in recent years: nuclear magnetic resonance (NMR) and molecular rulers. We then present methodological details for a recently developed approach.

Several different NMR experiments are useful for studying ensembles. First, NMR relaxation measurements provide incisive information about the existence, extent, and timescale of motions within an ensemble (Palmer, 2004). Specifically, in simple two-state systems, relaxation dispersion measurements can detect a well-defined alternate conformation that is

infrequently populated (down to the 0.5% level) (Dethoff, Petzold, Chugh, Casiano-Negroni, & Al-Hashimi, 2012; Mittermaier & Kay, 2006). In the general multistate case, however, relaxation experiments are blind to the specific subconformations that make up the ensemble. A second type of NMR experiment measures residual dipolar couplings (RDCs) between nuclear spins (Lange et al., 2008; Zhang, Stelzer, Fisher, & Al-Hashimi, 2007). RDCs provide ensemble-averaged values for the angles between common magnetic field and interspin vectors. RDCs also provide order parameters for these polar angles. Thus, it is possible to infer averaged angular relationships between structural elements of a macromolecule and the extent to which the angles vary within an ensemble. Finally, NMR is used to measure the rate of hydrogen-deuterium exchange of labile protons in a macromolecule (Woodward, Simon, & Tuchsen, 1982). These experiments can detect high-energy states, from which exchange occurs, as well as the protons that become solvent accessible in those states. Collectively, the great strengths of the NMR approaches are their dense information content, the ability to determine timescales for motions, and the ability to determine calibrated structural parameters (i.e., angles) in a model-free way. A limitation is that the observables are often averaged over heterogeneous conformations.

Molecular rulers represent a different attack on the ensemble problem. They are used to determine the distance between probes that are attached to a macromolecule. The use of probes provides high signal-to-noise for the experimental measurements, but it reduces the information density of the data relative to NMR approaches. Intramolecular distances are complementary to the orientation angles derived from RDCs. For macromolecules with a single dominant conformation, the distance data alone can specify a unique spatial arrangement of the probes relative to each other (allowing for global rotation, translation, or reflection).

The most broadly used molecular ruler is based on fluorescence resonance energy transfer (FRET) (Jares-Erijman & Jovin, 2003). Energy transfer results from dipolar coupling between two fluorophores that are attached site specifically to a macromolecule. The efficiency of energy transfer depends strongly on the distance between the fluorophores. However, FRET data only provide an ordinal estimate of intramolecular distances. This is because energy transfer depends on fluorophore orientation and dynamics in addition to distance, and because the mapping of the FRET signal-to-distance is highly nonlinear. The FRET ruler possesses two powerful attributes. First, it can be used in complex biological environments such

as whole cells. This distinguishes it from most high-resolution structural techniques. Second, it can be applied to single molecules. Different values of the FRET observable indicate the existence of different conformations of the single molecule. The approach is capable of detecting long-lived conformations that are infrequently populated. Rapidly interconverting conformations within an ensemble, however, are generally invisible to single-molecule FRET techniques.

Below, we present methods for a new class of molecular ruler based on X-ray scattering interferometry (XSI) (Mathew-Fenn, Das, Silverman, Walker, & Harbury, 2008; Shi, Beauchamp, Harbury, & Herschlag, 2014). XSI measures the interference of scattered waves from atoms in a macromolecule, which is the same physical principle that underlies X-ray crystallography. In XSI, the scattering is dominated by two heavy metal probes attached site specifically to the macromolecule. Probe scattering interference causes the intensity of the interference pattern to oscillate sinusoidally with increasing scattering angle. The frequency of these oscillations is proportional to the distance between the probes. XSI data include signals from multiple different interprobe distances, because the macromolecule populates multiple conformations with different probe separations. A Fourier transform of the interference pattern directly gives the fractional abundance of each one. These data are referred to as a distance distribution. Because scattering occurs faster than any atomic motion, the distance distribution is an unaveraged snapshot of the intramolecular probe separations that coexist in solution. Distance distributions between many different pairs of probe attachment sites can help to define the ensemble of shapes that the macromolecule adopts.

The strength of XSI is that it provides accurate, calibrated distances in a model-independent fashion. The data resolve distinct substates of an ensemble, even when the substates are rapidly interconverting. The technique is sensitive to distance differences on the angstrom scale. Calibrated distance is a simple physical observable that is directly related to three-dimensional structure. Thus, XSI data can be compared straightforwardly to atomic coordinates that have been determined experimentally or have been predicted by molecular dynamics simulations. XSI does not provide timescale information, it cannot be applied at the single-molecule level, and it is not compatible with complex environments. Its strengths and limitations are complementary to those of the FRET ruler.

Below, we describe the conceptual framework for XSI, followed by a detailed protocol for running an experiment.

3. CONCEPTUAL FRAMEWORK FOR XSI

X-ray interferometry, short for XSI, is a variant of solution X-ray scattering (SAXS). In a conventional SAXS experiment, the summed scattering inference from all atoms in a macromolecule is measured. Standard SAXS can be used to determine the overall size and rough shape of a macromolecule but provides no site-specific distance information. Point-to-point distance measurements can be obtained if a macromolecule is labeled at two locations (denoted A and B) with strongly scattering spherical probes (Fig. 1). The scattering intensity profile of the double-labeled macromolecule, $I_{AB}(S)$, includes the probe-to-probe scattering interference pattern, $I_{\Delta}(S)$. Here, I denotes the intensity of X-ray scattering, and S is the momentum transfer, which is proportional to the scattering angle. $I_{\Delta}(S)$ is a weighted sum of interference patterns corresponding to different discrete interprobe distances D:

$$I_{\Delta}(S) = \sum_{D=1-1000\,\text{Å}} P(D) * f^2{}_{\text{probe}}(S) * \frac{\sin(2\pi SD)}{2\pi SD}.$$

The distance distribution function, $P(D)$, is the fractional abundance of probe pairs separated by a center-to-center distance D. $P(D)$ can be determined by factoring $I_{\Delta}(S)$ into a sum of cardinal sine functions. The amplitude of the function with frequency D is the fractional abundance of the interprobe distance D. The $f(S)$ term is the scattering form factor of the spherical probe, which is determined experimentally; this term describes

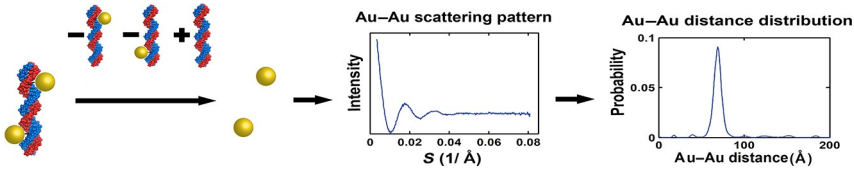


Figure 1 Obtaining a probe-probe distance distribution from X-ray scattering interferometry. (Left) A DNA duplex is labeled with gold nanocrystal probes. Subtracting the scattering profiles of the two singly labeled helices from the profiles of the doubly labeled and unlabeled helices gives the pattern of scattering interference between the two isolated gold probes (middle graph). This interference pattern is Fourier transformed into the probability distribution for the center-to-center distance between the two probes (right graph). The data shown are for two gold probes separated by 15 base steps within a 26 base-pair duplex. *Figure 1 is modified from figure 1 of Shi et al. (2013).*

the scattering properties of the probe due to its size and shape. Given an appropriate choice of distance bin resolution, there is a unique solution for $P(D)$. Reference Mathew-Fenn, Das, Silverman, et al. (2008) presents a more detailed discussion.

One complication is that the measured scattering intensity profile, $I_{AB}(S)$, has components other than the interprobe scattering interference pattern, specifically the intramacromolecule and probe-macromolecule interference patterns. But these additional components can be independently determined from the scattering profiles of the isolated macromolecule, $I_{macro}(S)$, and of macromolecules that are singly labeled at site A and site B, $I_A(S)$ and $I_B(S)$, and the extra components can then be subtracted off:

$$I_\Delta(S) = I_{AB}(S) - c_A I_A(S) - c_B I_B(S) + c_{macro} I_{macro}(S).$$

Due to uncertainty in sample concentration, the scaling factors c_A, c_B, and c_{macro}, can deviate from unity and need to be determined by an optimization procedure (Mathew-Fenn, Das, Silverman, et al., 2008). The optimal coefficients are chosen so that the oscillations in an S-weighted $I_\Delta(S)$ profile sum to zero and so that unphysical negative components of $P(D)$ are minimized.

The information content of XSI distance distributions depends on several experimental factors. The first factor is the scattering power of the probes relative to the macromolecule. This ratio determines the highest value of S at which $I_\Delta(S)$ can be measured over noise, and sets the distance resolution of the corresponding distance distribution. A second factor is how spherical and uniformly sized the probes are. These properties influence the extent to which macromolecular conformation can be inferred from probe center-to-center distances. A third factor is how rigidly the probes are affixed to the macromolecule. Compact and highly localized probes will provide the finest level of structural detail. In pioneering XSI work, heavy metal probes of just 1–4 atoms were used (Miakelye, Doniach, & Hodgson, 1983; Vainshtein et al., 1980). Alternatively, deuterated ribosomal subunits were used as neutron-scattering probes (Capel, Kjeldgaard, Engelman, & Moore, 1988). The marginal scattering power of these probes and the large size of the deuterated subunits limited the resolution.

The XSI approach detailed below utilizes spherical gold nanocrystals composed of 68 atoms (Fig. 2; Mathew-Fenn, Das, & Harbury, 2008; Shi et al., 2013). These nanocrystals scatter as strongly as a 20 kDa nucleic acid. Larger ellipsoidal nanocrystals composed of ~10,000 gold atoms have been utilized in other work (Hura et al., 2013; Mastroianni, Sivak, Geissler, &

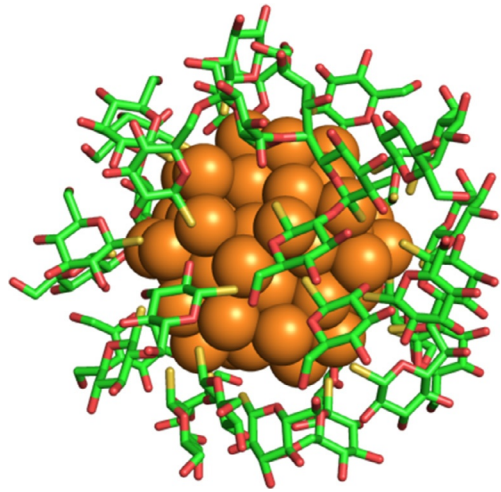


Figure 2 An illustration of the Au nanocrystal (spheres) with its thioglucose protection shell. The nanocrystal coordinates are based on a substructure of the nanocrystal reported in Jadzinsky, Calero, Ackerson, Bushnell, and Kornberg (2007) and the experimental analysis of Dass (2009). *Figure 2 is reproduced from figure 2C of Shi et al. (2013).*

Alivisatos, 2009). The variable size and shape of the ellipsoids reduced the resolution of the XSI distance measurements. Nevertheless, they scatter 15,000-fold more strongly than the 68-atom nanocrystals and are thus compatible with larger macromolecular complexes and enable measurements at sample concentrations down to 100 n*M* (Hura et al., 2013). At the opposite extreme, it should be possible to implement XSI with very small clusters. Gold atoms rigidly embedded in the side chain of an amino acid, for example, could provide high-resolution information about protein structure and fluctuations. This approach would require accurate matching of the solvent electron density to the protein electron density to minimize scattering by the protein.

Over the last few years, XSI has been fruitfully applied. Initially, distance distribution measurements on short 3′-end-labeled DNA helices revealed a previously unknown cooperative stretching motion (Mathew-Fenn, Das, & Harbury, 2008). A subsequent study analyzed long DNA fragments approaching one persistence length (Mastroianni et al., 2009). XSI was then used to measure the twisting and bending elasticity of DNA helices at short length scales (Shi et al., 2013). More recent work probed the solution conformational ensemble of a DNA bulge (Fig. 3; Shi et al., 2014). This model helix–junction–helix motif adopts heterogeneous structures in solution.

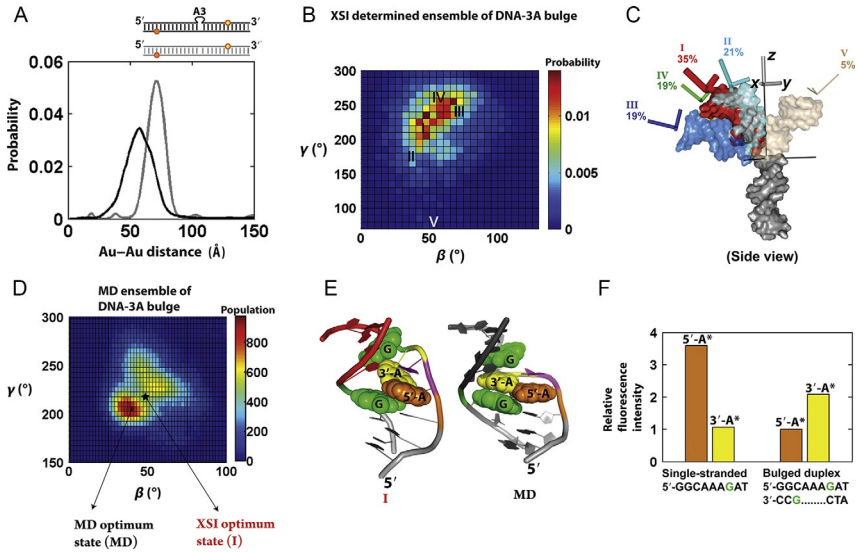


Figure 3 An ensemble model for a three adenosine DNA (DNA-3A) bulge based on XSI data. (A) Comparison of the Au–Au center-to-center distance distributions for a 26-base pair DNA helix (gray) and for DNA-3A, a 26-base pair DNA helix with a three adenosine bulge (black). (B) A heat map showing the conformational ensemble of the DNA-3A bulge determined by XSI. β and γ represent, respectively, the bending angle and the bending direction of the top helix relative to the bottom helix (see panel C). The model is based on six XSI distance distributions with different labeling positions for the Au nanocrystal probes. The distribution in panel A is one of the six data sets. (C) Five representative conformations of the DNA-3A ensemble (I–V) are shown as three-dimensional models. (D) A heat map showing the conformational ensemble of the DNA-3A bulge generated by an molecular dynamics (MD) ensemble simulation. The MD ensemble differs significantly from the XSI ensemble. For example, the most highly populated conformer in the XSI ensemble (state I, black star) is sparsely populated in the MD ensemble. (E) Atomic-level models of the most populated conformer from the XSI ensemble (state I) and from the MD. The XSI model predicts that the 5′ adenosine of the bulge is stacked on a flanking guanine, whereas the MD model predicts that the 3′ adenosine of the bulge is stacked on the flanking guanine. (F) Test of atomic-level models. The 5′ and 3′ adenosines of the bulge were replaced with fluorescent 2-aminopurine analogs. Stacking of 2-aminopurine on a guanine base quenches its fluorescence. Thus, the XSI model predicts that the 5′ 2-aminopurine fluorophore (A*) should be quenched when the bulge is formed from single-stranded oligonucleotide precursors, whereas the MD model predicts that the 3′ 2-aminopurine fluorophore (A*) should be quenched. Experimentally, quenching is observed at the 5′ position, consistent with the XSI model. *Panels (A–F) are reproduced from figures 1D, S6A, 4C, S12, and 5 of Shi et al. (2014).*

The DNA bulge experiments established the utility of XSI for studying complex ensemble systems. XSI has also been successfully applied to RNA and RNA-protein complexes (Shi et al., manuscript in preparation). Finally, large structural intermediates in bacterial DNA mismatch repair have been characterized by XSI (Hura et al., 2013).

The future directions and unsolved problems for XSI are numerous and include developing in-depth knowledge of the ensembles of RNA motifs and structured RNAs and RNA-protein complexes across a variety of solution conditions, all problems inaccessible by traditional approaches. Principal among unexplored areas is the application of XSI to proteins, which awaits an advance in probe attachment strategies. Another outstanding problem is to create large, spherical, and monodisperse nanoparticle probes. If they existed, these probes could be used to accurately measure structural rearrangements in biological supercomplexes, such as the ribosome. The scattering power of spherical XSI probes grows as the sixth power of their diameter. Thus, a modest increase in probe size can allow probe scattering to overwhelm the scattering of large biological assemblies. A third possible direction is single-molecule XSI using a free electron laser. This could provide correlated distance measurements between multiple sites on a macromolecule. Finally, new motifs for rigidly attaching scattering probes to nucleic acid and protein secondary structures could simplify the setup and interpretation of XSI experiments.

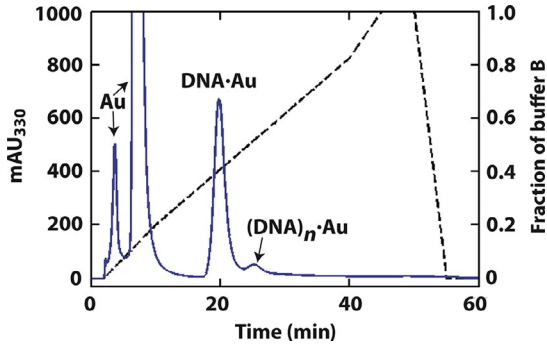
4. A GENERAL PROTOCOL FOR XSI MEASUREMENT OF NUCLEIC ACID CONSTRUCTS

This protocol is aimed to provide a detailed guide for carrying out an XSI experiment. The protocol is divided into several sections that roughly correspond to the chronological order of a typical project.

4.1 Getting Started: The Choice of Nucleic Acid System and Construct Design

The current protocol uses thioglucose passivated gold nanocrystals with a diameter of 1.2 nm (Fig. 2). These probes are adequate for nucleic acids up to 40 base pairs in size. We are not aware of larger spherical nanocrystals that are monodisperse and suitable for XSI. Making such probes would be a significant technical advance.

Note 1: *Proteins scatter X-rays more weakly than nucleic acids, and can be contrast matched with buffers containing sucrose or heavy salts* (Stuhrmann & Miller,



Scheme 1 Two sets of hypothetical 15-mer DNA constructs. The top quartet is labeled with gold nanocrystal probes at its ends, and the bottom quartet is labeled with gold nanocrystals at internal thymidine residues. The quartets are assembled from four single-stranded oligonucleotides. Within each quartet, "W" denotes the Watson strand and "C" denotes the Crick strand. The superscripts "3′-OH" and "3′-Au" indicate that the 3′-terminus of the respective oligonucleotide is a free hydroxyl group or is coupled to a gold nanocrystal. The superscripts "T-H" and "T-Au" indicate that the 5-methyl group of the internal thymine base is coupled to a proton or is coupled to a gold nanocrystal.

1978). *In protein–nucleic acid complexes, protein size is less limiting than nucleic acid size on a per kilo Dalton basis.*

Gold nanocrystals can be attached at the 3′- or 5′-end of a nucleic acid strand (end labeled) or to a modified base within the strand (internally labeled). To avoid undesirable fraying of terminal base pairs, it is advisable to put a G–C cap at the ends of helices. For internally labeled constructs, we recommend that labeling sites be placed at least three base pairs inward from the helix ends.

In the following protocol, we describe two hypothetical 15-mer DNA constructs, one end labeled, and the other internally labeled (Scheme 1).

4.2 Stage 1: Synthesis and Purification of Au Nanocrystals

Steps 1.1 and 1.2 prepare FPLC columns.

Two FPLC columns are used for the purification of gold nanocrystals, a Sephadex G15 column (GE Healthcare) for rapid desalting and a Superdex 30 column (GE Healthcare) for size exclusion purification.

1.1. The G15 column (50 mL in a 26/10 column housing) is first washed with two column volumes of water, followed by four column volumes of cleaning solution (20 m*M* Tris–HCl, pH 8.0, and 100 m*M* DTT) at 2 mL/min. The column is then equilibrated with four column volumes of water.

1.2. The Superdex 30 column (120 mL in a 16/60 column housing) is first washed with two column volumes of water at 0.75 mL/min, followed by 4–5 column volumes of cleaning solution and another 4–5 column volumes of water to remove the residual DTT. The column is then equilibrated with two column volumes of 150 m*M* ammonium acetate, pH 5.6, the running buffer for nanocrystal purification.

Note: *The DTT in the cleaning solution removes gold residue left behind by previous samples. The residue causes a brown discoloration of the white chromatography resin. The column should turn white again after it is treated with the cleaning solution.*

Steps 1.3–1.9 synthesize Au nanocrystals.

1.3. Prepare a 250-mL round-bottom flask, washed, dried, and cooled. Add a stir bar and cap the flask with a vented septum. Mount the flask above a magnetic stir plate inside a dark fume hood.

1.4. Prepare a 72 mL solution of 5:1 by volume methanol–acetic acid.

1.5. Weigh 0.544 g of hydrogen tetrachloroaurate (III) hydrate (Strem Chemicals 79-0500). The gold salt is light sensitive. Immediately transfer the powder to the round-bottom flask. Add 36 mL of the 5:1 methanol–acetic acid solution to the flask. The stirred solution should be clear with a bright orange color.

Note: *The hydrogen tetrachloroaurate (III) hydrate bottle should be prewarmed to room temperature on a bench before it is opened. This minimizes the condensation of water inside the bottle. The bottle should be flushed with nitrogen or argon after use.*

1.6. Dissolve 1 g of 1-thio-β-D-glucose (Sigma T6375 or Santa Cruz Biotech sc-216128) in the remaining 36 mL of the 5:1 methanol–acetic acid solution using a vortexer. Add the thioglucose solution to the stirred reaction flask. The mixture in the round-bottom flask should become cloudy.

1.7. Add 0.9 g of sodium borohydride to 20 mL of water in a 50 mL Falcon tube. Dissolve the solid by inverting the tube two to three times, or by vortexing briefly. *How you perform the next step is critical to the quality of the nanocrystal synthesis*! Add the sodium borohydride solution to the reaction flask in a dropwise manner over 3–4 min. Use a 1-mL pipettor.

Note: *To maintain a more constant addition speed, a glass addition funnel with a teflon stopcock can be used instead of the pipettor. The funnel needs to be uncapped during the addition and the teflon stopcock needs to preadjusted to ensure a suitable dropwise draining rate.*

1.8. Allow the reaction to stir at ambient temperature for another 30 min.

1.9. Reduce the volume of the crude reaction product to about 12 mL using a rotary evaporator.

Steps 1.10–1.15 purify the Au nanocrystals.

1.10. Keep the crude Au nanocrystal preparation on ice or in a cold room. Filter the nanocrystal solution with a 0.22-μm filter to remove particulates. The filtered Au nanocrystals solution is then ready for FPLC purification.

1.11. Desalt the crude preparation on a G15 FPLC column (see step 1.1). Load a maximum of 6 mL of the Au nanocrystals solution, and then pump water over the column at 2 mL/min. The nanocrystal peak can be detected by absorption at any wavelength between 260 and 360 nm. Au nanocrystals of 1.2–1.4 nm in diameter elute in the 9–12 min window. Collect the main nanocrystal peak, and avoid the salt front by monitoring conductivity. Repeat step 1.11 until the entire volume of the crude material is desalted.

Note: *As the crude Au nanocrystals are relatively unstable at high salt, high pH, and high temperature, try to minimize unnecessary downtime in steps 1.10 and 1.11. Before loading onto the G15 column, check the solution for visible precipitates and refilter if necessary.*

Note: *Once the Au nanocrystal solution is desalted, it is safe to store it at 4 °C for days or at −20 °C for months.*

1.12. Reduce the volume of the desalted Au nanocrystal solution to about 12 mL by either centrifugal filtration (3K Amicon) or rotary evaporation.

1.13. Inject up to 6 mL of the concentrated Au nanocrystal solution onto the Superdex 30 column. Pump 150 m*M* ammonium acetate, pH 5.6, over the column at 0.75 mL/min for 205 min. Collect only the center of the largest UV-absorbing peak and avoid the shoulder region to ensure nanocrystal size homogeneity. Repeat step 1.13 until the entire volume of the desalted Au nanocrystal solution is purified. Immediately after each iteration of step 1.13, repeat step 1.12 to reduce the solution volume.

1.14. Immediately repeat step 1.11 to desalt the purified Au nanocrystals.

1.15. The concentration of the final purified and desalted Au nanocrystals is determined by UV absorption using an extinction coefficient of 0.076 μ*M*/cm at 360 nm.

Note: *The size distribution of the Au nanocrystals can be determined from their X-ray scattering profile* (*see step 4.1 below*). *The purity and homogeneity of the Au nanocrystals can also be checked by 15% denaturing PAGE.*

4.3 Stage 2: Generating Au Nanocrystal Labeled Constructs

We describe the procedures below using an end-labeled and an internally labeled duplex (Scheme 1) as examples. Although the examples are in DNA, the same procedure also applies to RNA. Suggested minor modifications to the procedure for RNA are noted at steps 2.9 and 2.17.

Steps 2.1–2.3 synthesize and purify single-stranded oligonucleotides.

2.1. Purchase amine-modified oligonucleotides at the 200-nmol scale from a commercial vendor. Larger quantities may be needed for repeated measurements. Alternatively, practitioners can synthesize these oligonucleotides using a DNA/RNA synthesizer (e.g., ABI 394). As illustrated in Scheme 1, oligonucleotides $W^{3'\text{-OH}}$, $C^{3'\text{-OH}}$, $W^{T\text{-H}}$, and $C^{T\text{-H}}$ are unmodified. Oligonucleotides $W^{3'\text{-Au}}$ and $C^{3'\text{-Au}}$ incorporate a 3′-thiol modification (Glen Research C3-S-S, 20-2933). Oligonucleotides $W^{T\text{-Au}}$ and $C^{T\text{-Au}}$ incorporate an amine-modified T base (Glen Research amino-modifier C2 dT, 10-1037) at positions intended for gold labeling.

Note: *The 3′-thiol modification (C3-S-S) contains a short three-carbon linker between the terminal phosphate and the sulfur atom. We are not aware of a comparably short 5′-thiol modification. Longer six-carbon thiol modifications are commercially available for both the 3′- and 5′-ends of oligonucleotides.*

Note: *For the 3′-thiol modification (C3-S-S), leave the disulfide bond intact.*

Note: *It is advised to order oligonucleotides as "DMT-on." These oligonucleotides retain their 5′-terminal dimethoxytrityl protecting group and can be cleaned up over a reverse phase cartridge (e.g., Glen Research Poly-Pak cartridges).*

2.2. Purify the single-stranded oligonucleotides by anion exchange HPLC using a Dionex DNAPac semipreparative column (9×250 mm) at 3 mL/min. Oligonucleotides are eluted with a salt gradient formed from buffer A (10 m*M* NaCl, 20 m*M* Na-borate, pH 7.8) and buffer B (1500 m*M* NaCl, 20 m*M* Na-borate, pH 7.8).

Note: *Borate buffer is used instead of an amine-containing buffer, which would interfere with the reaction in step 2.6 below.*

Note: *Before each HPLC run, particulates should be removed. The samples can be spun in a microfuge for 4 min at 4 °C, followed by transfer of the cleared solution to a new tube. Alternatively, the samples can be passed through a 0.22 μm filter.*

2.3. Desalt and concentrate the HPLC-purified oligonucleotides by centrifugal filtration (3K Amicon) or ethanol precipitation. Check the purity by PAGE, analytical HPLC, or capillary electrophoresis. Typical yields from the HPLC purification are 30–60%, depending on the purity of the synthesized oligonucleotides.

Note: *One typically gets ~120 nmol of crude oligonucleotide from a 200 nmol scale synthesis* (*step 2.1*). *Following step 2.3*, *~60 nmol of purified single-stranded oligonucleotide is recovered*.

Steps 2.4–2.11 introduce a thiol moiety into oligonucleotides with internal amine modifications.

Note: *The internally amino modified* $W^{T\text{-}Au}$ and $C^{T\text{-}Au}$ *must be reacted with SPDP* (*N-succinimidyl 3-*[*2-pyridyldithio*]*-propionate*) *to introduce a disulfide group. This is described in steps 2.4–2.8. The oligonucleotides* $W^{3'\text{-}Au}$ and $C^{3'\text{-}Au}$ *arrive from the vendor with 3′ disulfide modifications*, *so they do not require any additional chemical steps*.

2.4. Reduce the volume of $W^{T\text{-}Au}$ and $C^{T\text{-}Au}$ (~60 nmol) to less than 160 μL. Add water to bring the total volume to 160 μL. Then, add 20 μL of 1 *M* borate buffer (pH 7.8).

2.5. Make an SPDP (*N*-succinimidyl 3-[2-pyridyldithio]-propionate) solution by dissolving SPDP in DMSO (1 mg SPDP per 10 μL of DMSO). A total of 40 μL of SPDP solution is required for each oligonucleotide (a total of 80 μL of SPDP solution for the $W^{T\text{-}Au}$ and $C^{T\text{-}Au}$ pair).

2.6. Preheat the oligonucleotide solutions to 37 °C. Add 20 μL of SPDP solution to each tube. Mix by pipetting up and down, and then incubate for 30 min at 37 °C.

2.7. After the 30 min, add another 20 μL of SPDP solution and incubate for another 30 min at 37 °C.

2.8. Add 2 μL of 2 *M* $MgCl_2$ and 1 mL of cold ethanol. After gentle mixing, incubate the reaction tubes in dry ice for 40 min. Spin the tubes for 30 min at maximum speed using a microfuge at 4 °C, and remove the supernatant. Wash by adding 1 mL of cold ethanol. Spin for another 15 min at 4 °C, and then remove the supernatant.

Note: $W^{T\text{-}Au}$ and $C^{T\text{-}Au}$ *are now internally labeled with a disulfide group. The typical yield for steps 2.4–2.7 is close to 100%*.

2.9. Reduce the disulfide bonds in $W^{3'\text{-}Au}$, $C^{3'\text{-}Au}$, $W^{T\text{-}Au}$, and $C^{T\text{-}Au}$ by dissolving them in 150 μL of 200 m*M* DTT, 50 m*M* Tris–HCl, pH 9.0. Incubate the solutions at 50–70 °C for 30 min.

Note: *50 °C for RNA*.

2.10. Purify the oligonucleotides by ethanol precipitation as in step 2.8.

2.11. Dissolve the pellets in 500 μL of water, and remove residual DTT by centrifugal filtration at 4 °C (3K Amicon, 14,000 × *g*, 30 min). Determine the concentration of each oligonucleotide by UV absorption (using a Nanodrop for example).

Note: *Excess DTT destabilizes nanocrystal-DNA conjugates. If a significant amount of residual DTT is observed, indicated by a strong absorbance at 230 nm and a shoulder peak at ≥300 nm, repeat step 2.11.*

Steps 2.12–2.15 label the single-stranded oligonucleotides with Au nanocrystals.

Perform step 2.12 immediately after completing step 2.11.

2.12. This step uses a 6-to-1 molar ratio of purified and desalted Au nanocrystals to oligonucleotide. For example, use 300 nmol of Au nanocrystals for 50 nmol of oligonucleotide from step 2.11. Add the oligonucleotide solution to the Au nanocrystal solution and gently vortex. Add 20 μL of 1 *M* Tris–HCl, pH 9.0, and gently vortex again. Incubate for 2 h at room temperature.

2.13. Stop the reaction with 15 μL of 2 *M* ammonium acetate, pH 5.6, and store the samples on ice.

2.14. Purify the Au-labeled oligonucleotides by anion exchange HPLC using a Dionex DNAPac column. The conjugates are eluted with a salt gradient formed from buffer A (10 m*M* NaCl, 20 m*M* ammonium acetate, pH 5.6) and buffer B (1500 m*M* NaCl, 20 m*M* ammonium acetate, pH 5.6). Monitor the absorbance at 260 nm, where nanocrystals and nucleic acids both absorb. Also monitor absorbance at 330 nm, where only nanocrystals absorb. A sample chromatogram is provided in Fig. 4.

Note: *The 1:1 Au-oligonucleotide conjugates elute from the DNAPac column earlier than unlabeled oligonucleotides of the same length. Depending on oligonucleotide length, the conjugates come off of the column in the 25-50%B range. The conjugates absorb at 330 nm whereas the free oligonucleotides do not. There should be almost no free oligo remaining after step 2.12. Unmodified gold nanocrystals do not bind strongly to the column.*

Note: *There will be by-products consisting of gold nanocrystals coupled to multiple oligonucleotides. These species elute later than the 1:1 nanocrystal-oligonucleotide conjugate.*

2.15. Buffer exchange the nanocrystal-oligonucleotide fractions into water by centrifugal filtration at 4 °C (3K Amicon, 3500 × *g* with a swinging bucket rotor for 30–40 min; repeat three times).

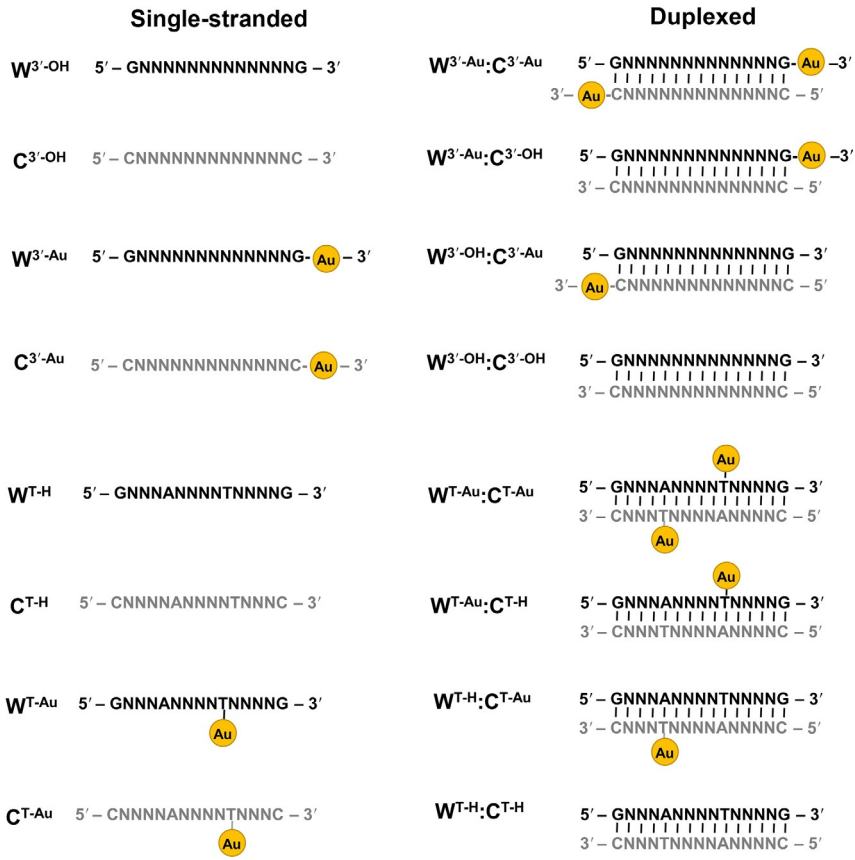


Figure 4 A sample HPLC chromatogram of Au-labeled oligonucleotides. An ion-exchange HPLC chromatogram of a crude coupling reaction between gold nanocrystals and thiol-modified DNA. The different products and biproducts include uncoupled gold nanocrystals (Au), a 1:1 complex of a gold nanocrystal and a 21-mer single-stranded DNA (DNA·Au), and 1:multiple complexes of gold nanocrystals and 21-mer single-stranded DNAs (DNA_n·Au).

2.16. Determine the concentration of the purified conjugates by absorption at 260 nm. Calculate the extinction coefficient of the conjugate as the sum of the extinction coefficient for the oligonucleotide component and an extinction coefficient of 0.215 μM/cm for the gold nanocrystal component.

Note: *After desalting, the samples can be stored at −20 °C for extended periods of time before step 2.17.*

Note: *A typical yield for steps 2.12–2.16 is 20%. Starting with 50 nmol at step 2.12, roughly 10 nmol are recovered after step 2.16.*

Steps 2.17–2.22 prepare the final duplexed constructs for XSI measurements.

2.17. Mix pairs of complementary single-stranded oligonucleotides in a 1:1 molar ratio. Incubate the solutions at room temperature for 30 min.

Note: *For RNA, incubate, the solutions at 40 °C for 30 min.*

2.18. Purify the annealed samples by anion exchange HPLC as in step 2.14. The duplex DNA elutes roughly 5% Buffer B (See Step 2.2 above for the definition of Buffer B) later than the single-stranded starting materials.

2.19. Buffer exchange the duplex DNA into water by centrifugal filtration at 4 °C (10K Amicon, $3500 \times g$ with a swinging bucket rotor for 15 min, repeat three times).

2.20. Reduce the volume of the desalted samples to about 30–40 μL by centrifugal filtration at 4 °C (10K Amicon, $10{,}000 \times g$ with a 0.5 mL filter unit for 20 min).

2.21. Determine the concentration of the purified samples by absorption at 260 nm and 360 nm. Calculate the extinction coefficient of the double-stranded conjugate as the sum of the extinction coefficients for the DNA component and the gold nanocrystal component.

2.22. Store the samples in a –20 or a -80 °C freezer. Desalted samples free of DNase and RNase contamination are stable for at least months.

Note: *The typical yield for steps 2.17–2.22 is about 50%. Starting with 10 nmol of each single-stranded oligonucleotide at step 2.17, 5 nmol of duplexed product is obtained. Each XSI measurement requires 0.9 nmol of sample.*

4.4 Stage 3: Collecting the X-ray Scattering Data

The XSI measurements are carried out at a synchrotron beamline configured for small-angle X-ray scattering. An example is beamline 4-2 of the Stanford Synchrotron Radiation Lightsource (SSRL).

3.1. Use a sample-to-detector distance that covers a q range from 0.01 to 1 Å^{-1}. This corresponds to 1.1 m at 11 keV on SSRL beamline 4-2.

3.2. Make a 10 × buffer solution containing Tris–HCl, sodium ascorbate, the desired concentration of additional salt, and any other components required for the experiment. An example of a 1 × buffer is 70 m*M* Tris–HCl, pH 7.4, 10 m*M* sodium ascorbate, 150 m*M* NaCl, and 1 m*M* $MgCl_2$.

Note: *Ascorbate functions as radical scavenger, as does Tris to a lesser extent. A radical scavenger is essential; it protects the DNA from oxidative damage.*

3.3. Before measurement, the samples are thawed and vortexed. Prepare a 200 μM solution of Au nanocrystals in 1 × buffer. Spin the solution in a microfuge for 2 min at 10,000 × *g* and 4 °C. This removes potential large particle contaminants that can dominate the scattering.

3.4. Set up the data collection as 10 repeats of a 3 s exposure. In general, it is advisable not to reuse the sample.

Note: *Oxidative damage to the sample can be detected by a gradual change in scattering intensity over the ten repeats. Alternatively, samples can be recovered after exposure to X-rays and analyzed by HPLC. Do not use data collected under conditions that cause oxidative damage.*

3.5. Measure the scattering of the pure Au nanocrystal solution. Repeat step 3.5 for a solution of the buffer alone.

Note: *The scattering profiles of Au nanocrystals under different salt conditions are nearly identical.*

Note: *It is advisable to measure scattering for a dilution series of the Au nanocrystals, for example, at 200, 100, 50, and 25 μM. The shape of the scattering profiles should be independent of concentration. A concentration dependence of the profile at low q can result from interparticle scattering. Work in a concentration range where interparticle scattering is negligible.*

3.6. Prepare a quartet of samples for data collection. The quartet consists of an unlabeled duplex, two singly labeled duplexes, and a doubly labeled duplex (e.g. see Fig. 1 and Scheme 1). The nanocrystal–oligonucleotide conjugates should be at 30 μM concentration in 1 × buffer. Spin the samples as in step 3.3. Measure the scattering profiles of the quartet in a direct sequence, so that the data collection conditions are as closely matched as possible.

Note: *Measurements typically require 30 μL of solution (0.9 nmol of sample). In cases, where the concentration of phosphate in the nucleic acid backbone is comparable to the concentration of a counterion in the buffer, for example, at very low concentrations of* $MgCl_2$, *the samples should be prepared by exchange into 1× buffer using a centrifugal filter (10K Amicon).*

3.7. Repeat step 3.6 with additional quartets, with alternate buffers, and with additional experimental components such as binding partners.

4.5 Stage 4: Analyzing the Data

Data processing is carried out using beamline software and custom MATLAB scripts. The conceptual basis for the processing algorithms is described in detail elsewhere (Mathew-Fenn, Das, Silverman, et al., 2008).

A brief overview of the analysis pipeline is summarized here. Following regular SAXS data processing (step 4.1), the radius distribution of the Au nanocrystals is first determined. This is done by decomposing the measured gold auto-scattering signal into a sum of scattering profiles for solid spheres of varying radius (step 4.2). The radius distribution is then used to calculate IqD, which are basis functions for scattering interference between two nanocrystals separated by varying center-to-center distances (step 4.3). Independently, the scattering profiles of the double-, single-, and non-labeled constructs are each decomposed into a sum of simple cardinal sine functions (these are denoted $P(D)$). These crude transforms are used to optimize scaling coefficients so that the sum $P_{\Delta}(D) = P_{\mathrm{AB}}(D) - c_{\mathrm{A}}P_{\mathrm{A}}(D) - c_{\mathrm{B}}P_{\mathrm{B}}(D) + c_{\mathrm{macro}}P_{\mathrm{macro}}(D)$ does not include negative values. Following optimization of the scaling coefficients, a final high-resolution probability distribution for nanocrystal-nanocrystal center-to-center distances is obtained by decomposing the $I_{\Delta}(S)$ interference profile into a sum of the IqD basis functions using a maximum entropy procedure (step 4.5).

Below, we provide the step-by-step guide for obtaining the nanocrystal-nanocrystal scattering interference profile, and the center-to-center distance distribution between a pair of nanocrystals attached to a nucleic acid construct. The custom MATLAB scripts can be downloaded from the open-access website for (Mathew-Fenn, Das, Silverman, et al., 2008). The procedure to compute an ensemble from distance distribution data is described elsewhere (Shi et al., 2014, 2013).

4.1. Reduce the raw data to a one-dimensional X-ray scattering profile. The raw scattering data are a two-dimensional matrix of photon intensities as a function of position on the X-ray detector. Using beamline-specific software, the matrix is integrated into a one-dimensional scattering profile. The data reduction program at SSRL beamline 4-2 is named sastool. The processed data consist of ordered pairs of scattering intensity and magnitude of the scattering vector. There are two conventions for expressing the magnitude of the scattering vector, represented by the symbols q and S. Specifically, q is calculated as $4\pi\sin(\theta/\lambda)$, where θ is half of the scattering angle and λ is the wavelength of the X-ray radiation. S is calculated as $q/(2\pi)$. q and S have units of $\mathrm{\AA}^{-1}$. Both are linearly proportional to scattering angle for small angles. Ten separate scattering profiles are measured for each sample (step 3.4). The variance in the measured scattering intensities for each value of q (or S) is included in the data set. The data should be

truncated at very low values of q to remove beamline artifacts. This is generally done by visual inspection.

4.2. Obtain a volume-weighted Au nanocrystal radius distribution, denoted PAU6. First, subtract the buffer scattering profile from the pure Au nanocrystal scattering profile (these data are obtained in step 3.5). The buffer-corrected profile is denoted I_Au. Then, obtain PAU6 with the MATLAB function call: PAU6 = lsqnonneg(gen_IqR(*S*, 1:100), I_Au). *S* is a vector containing the magnitudes of the scattering vectors for each intensity value in I_Au.

4.3. Obtain basis functions for the scattering interference profile of a pair of Au nanocrystals at different center-to-center distances. These are stored in a matrix denoted IqD. Obtain IqD with the MATLAB function call: IqD = gen_IqD(gen_Ic2c0(PAU6, 1:100, *S*), *S*, 1:200). The resulting IqD matrix includes interference basis functions for inter-nanocrystal distances of 1-200 Å in 1 Å increments.

4.4. Obtain the nanocrystal-nanocrystal scattering interference profile, denoted I_delta. I_delta is computed by subtracting the two singly labeled scattering profiles from the doubly labeled scattering profile. The profiles must be properly scaled relative to each other. Obtain I_delta with the MATLAB function call: I_delta = gen_Idelta(*S*, I_U, I_A, I_B, I_AB, I_Au, I_Buf, 1). I_U, I_A, I_B, I_AB, I_Au, and I_Buf are, respectively, the scattering profiles of the unlabeled duplex, the duplexes that are singly labeled at sites A and B, the doubly labeled duplex, the Au nanocrystals and the buffer. The gen_Idelta function also returns the optimal coefficients for scaling the scattering profiles. These are denoted c_U, c_A, c_B, c_AB, c_Au, and c_Buf. The variance in I_delta for each value of *S*, denoted V_delta, is returned by the script.

Note: *The coefficient* c_Au *was originally included to correct for free Au nanocrystals in the experimental samples. With the improved sample-preparation protocols described above, free Au nanocrystals are negligible. Therefore,* c_Au *can be set to 0.*

4.5. Obtain the center-to-center distance distribution between the pair of nanocrystals, denoted $P(D)$. The script is based on a maximum entropy optimization. Obtain $P(D)$ with the MATLAB function call: [~,P_D] = jacknife(*S*, I_delta, V_delta, IqD, ICF).

Note: *A canonical ICF* (*intrinsic correlation function matrix*) *is included with the MATLAB script download. If desired, alternative ICF matrices can be generated by modifying the script gen_ICF.*

ACKNOWLEDGMENTS

We thank H. Tsuruta (deceased), T. Matsui, and T. Weiss at beamline 4-2 of the Stanford Synchrotron Radiation Lightsource (SSRL) for their technical support on synchrotron experiments and during the development of the X-ray scattering interferometry technique reported here. This work was supported by NIH grants PO1 GM066275 (D.H.) and DP-OD000429-01 (P.B.H.).

REFERENCES

Bacolla, A., & Wells, R. D. (2004). Non-B DNA conformations, genomic rearrangements, and human disease. *Journal of Biological Chemistry*, *279*(46), 47411–47414.

Capel, M. S., Kjeldgaard, M., Engelman, D. M., & Moore, P. B. (1988). Positions of S2, S13, S16, S17, S19 and S21 in the 30-S-ribosomal subunit of escherichia-coli. *Journal of Molecular Biology*, *200*(1), 65–87.

Dass, A. (2009). Mass spectrometric identification of Au-68(SR)(34) molecular gold nanoclusters with 34-electron shell closing. *Journal of the American Chemical Society*, *131*(33), 11666–11667.

Dethoff, E. A., Petzold, K., Chugh, J., Casiano-Negroni, A., & Al-Hashimi, H. M. (2012). Visualizing transient low-populated structures of RNA. *Nature*, *491*(7426), 724–728.

Feldkamp, U., & Niemeyer, C. M. (2006). Rational design of DNA nanoarchitectures. *Angewandte Chemie-International Edition*, *45*(12), 1856–1876.

Fenwick, R. B., Esteban-Martin, S., & Salvatella, X. (2011). Understanding biomolecular motion, recognition, and allostery by use of conformational ensembles. *European Biophysics Journal*, *40*(12), 1339–1355.

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.

Frank, J., & Agrawal, R. K. (2000). A ratchet-like inter-subunit reorganization of the ribosome during translocation. *Nature*, *406*(6793), 318–322.

Hura, G. L., Tsai, C. L., Claridge, S. A., Mendillo, M. L., Smith, J. M., Williams, G. J., et al. (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.

Jadzinsky, P. D., Calero, G., Ackerson, C. J., Bushnell, D. A., & Kornberg, R. D. (2007). Structure of a thiol monolayer-protected gold nanoparticle at 1.1 A resolution. *Science*, *318*(5849), 430–433.

Jares-Erijman, E. A., & Jovin, T. M. (2003). FRET imaging. *Nature Biotechnology*, *21*(11), 1387–1395.

Kaplan, N., Moore, I. K., Fondufe-Mittendorf, Y., Gossett, A. J., Tillo, D., Field, Y., et al. (2009). The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature*, *458*(7236), 362–366.

Lange, O. F., Lakomek, N.-A., Fares, C., Schroeder, G. F., Walter, K. F. A., Becker, S., et al. (2008). Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution. *Science*, *320*(5882), 1471–1475.

Mastroianni, A. J., Sivak, D. A., Geissler, P. L., & Alivisatos, A. P. (2009). Probing the conformational distributions of subpersistence length DNA. *Biophysical Journal*, *97*(5), 1408–1417.

Mathew-Fenn, R. S., Das, R., & Harbury, P. A. (2008). Remeasuring the double helix. *Science*, *322*(5900), 446–449.

Mathew-Fenn, R. S., Das, R., Silverman, J. A., Walker, P. A., & Harbury, P. A. B. (2008). A molecular ruler for measuring quantitative distance distributions. *PLos One*, *3*(10), e3229.

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

Mittermaier, A., & Kay, L. E. (2006). Review—New tools provide new insights in NMR studies of protein dynamics. *Science*, *312*(5771), 224–228.

Olson, W. K., Gorin, A. A., Lu, X. J., Hock, L. M., & Zhurkin, V. B. (1998). DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(19), 11163–11168.

Palecek, E. (1991). Local supercoil-stabilized DNA structures. *Critical Reviews in Biochemistry and Molecular Biology*, *26*(2), 151–226.

Palmer, A. G. (2004). NMR characterization of the dynamics of biomacromolecules. *Chemical Reviews*, *104*(8), 3623–3640.

Rohs, R., Jin, X., West, S. M., Joshi, R., Honig, B., & Mann, R. S. (2010). Origins of specificity in protein-DNA recognition. *Annual Review of Biochemistry*, *79*, 233–269.

Salmon, L., Yang, S., & Al-Hashimi, H. M. (2014). Advances in the determination of nucleic acid conformational ensembles. *Annual Review of Physical Chemistry*, *65*(65), 293–316.

Shi, X. S., 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.

Shi, X. S., Herschlag, D., & Harbury, P. A. B. (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.

Stuhrmann, H. B., & Miller, A. (1978). Small-angle scattering of biological structures. *Journal of Applied Crystallography*, *11*(Oct), 325–345.

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 hemoglobin and histidine-decarboxylase in solution by small-angle X-Ray-scattering. *FEBS Letters*, *116*(1), 107–110.

Woodward, C., Simon, I., & Tuchsen, E. (1982). Hydrogen-exchange and the dynamic structure of proteins. *Molecular and Cellular Biochemistry*, *48*(3), 135–160.

Zhang, Q., Stelzer, A. C., Fisher, C. K., & Al-Hashimi, H. M. (2007). Visualizing spatially correlated dynamics that directs RNA conformational transitions. *Nature*, *450*(7173), 1263–1267.