# From sequence to function: Bridging single-molecule kinetics and molecular diversity

**A. N. Kapanidis[1]*, L. Muras[2], K. Sreenivasa[3], J. Hazra[1], J. van Noort[4], C. Joo[3], S. Deindl[2]***

[1]Department of Physics, Kavli Institute for Nanoscience Discovery, University of Oxford, Oxford, UK.
[2]Department of Cell and Molecular Biology, Science for Life Laboratory, Uppsala University, 75105 Uppsala, Sweden.
[3]Department of BioNanoScience, Kavli Institute of Nanoscience, Delft University of Technology, 2629 HZ Delft, Netherlands.
[4]Biological and Soft Matter Physics, Huygens-Kamerlingh Onnes Laboratory, Leiden University, Niels Bohrweg 2, 2333 CA Leiden, Netherlands.

*Corresponding authors. Email: kapanidis@physics.ox.ac.uk (A.K.) and sebastian.deindl@icm.uu.se (S.D.)

**Abstract:**

**Biological function is fundamentally determined by nucleic acid and protein sequence. Beyond encoding genetic information, nucleic acids also display complex physicochemical parameters that shape structure, dynamics, and interactions. Understanding how sequence variation sculpts the energetic landscapes underlying these properties requires methods that capture both molecular diversity and dynamic behavior. Single-molecule techniques are ideally suited to this task, but conventional formats remain time- and cost-intensive. Recent breakthroughs have enabled highly multiplexed approaches for observing molecular dynamics across millions of individual molecules representing thousands of sequences or barcoded entities. Though still in development, these methods have begun to bridge sequence, structure, dynamics, and function at scale - opening new opportunities in drug discovery, molecular diagnostics, and functional genomics.**

**Teaser:**

A Review of emerging strategies for multiplexed single-molecule analysis of biomolecular structure, function, and dynamics.

**Main text:**

## Sequence-dependent structure and function in nucleic acids

Biological function is fundamentally determined by the sequence of nucleic acids and proteins (Fig. 1). This principle is especially evident in the role of DNA and RNA in encoding genetic information, which in turn dictates the sequences and structural properties of proteins (Fig. 1A). Beyond serving as carriers of genetic instructions, nucleic acids also exhibit complex physicochemical properties that critically influence their structure, dynamics, and interactions (Fig. 1B-D). Differences in the chemical properties of the DNA bases enable specific interactions with proteins, such as transcription factors, through contacts in both the major and minor grooves of the DNA (Fig. 1D).

Within double-stranded DNA, variations in hydrogen bonding between base pairs and differences in base-stacking interactions between the aromatic rings of adjacent bases make the mechanical properties of DNA sequence-dependent (Fig. 1B,C). Such variations can have important biological consequences, facilitating melting bubbles that enable transcription and recombination, introducing bends and kinks, and giving rise to alternative structures such as G-quadruplexes. Furthermore, local variations in the mechanical properties of DNA suggest the existence of a mechanical code for positioning nucleosomes (*1*, *2*), adding another layer of transcription regulation. RNAs, being single-stranded, allow for an even wider spectrum of sequence-dependent structures and dynamic conformational exchanges. As a result, DNA and RNA store and transmit information beyond the explicit genetic code. Achieving a comprehensive understanding of how sequence variations shape the energetic landscapes (Fig. 1E) that govern structural dynamics and molecular interactions - and ultimately modulate biological function - often requires exploration of a broad sequence space.

Recent ensemble-based methods for high-throughput *in vitro* analysis have substantially advanced mechanistic molecular biology (*2–19*). Many of these approaches rely on DNA sequencing to distinguish between different library members. In these cases, DNA sequences can serve as direct readouts or indirectly report on RNA or protein sequences via transcription, translation, or DNA barcoding. For example, massively parallel reporter assays link DNA barcodes to regulatory elements to quantify their effects on gene expression (*20*, *21*). Deep mutational scanning (DMS) enables systematic mapping of the functional consequences of mutations across a protein or RNA sequence (*22*, *23*). In another approach, DNA-barcoded protein-binding assays can reveal sequence determinants of transcription factor binding with high resolution (*4*, *24*). More broadly, DNA barcoding extends these strategies to encompass diverse types of chemical variation (*6*, *7*). While these methods offer significant advantages, their capacity to resolve dynamic processes is inherently constrained by ensemble averaging.

**Single-molecule measurements and the power of multiplexing**

Single-molecule techniques provide a powerful means to study these dynamic processes, as they circumvent ensemble averaging, which can obscure transient intermediates and alternative kinetic pathways (*25–43*). However, single-molecule methods are often limited in their ability to observe multiple distinct molecular species, typically requiring separate labor-intensive experiments for each species, one at a time. This constraint often restricts single-molecule investigations to a small subset of possible samples, making it impractical to study large libraries.

Recent studies have demonstrated the simultaneous analysis of mixtures containing two (*44*) or nine (*45*) distinct DNA sequences using fluorophore-labeled detection oligonucleotides. More recently, Squires et al. demonstrated an elegant solution-phase alternative in which nanostructured FRET barcodes, held in an anti-Brownian electrokinetic trap, discriminate on the order of two-dozen DNA species (*46*). While the barcodes markedly expand the distinguishable palette, the approach does not provide a direct readout of each molecule's sequence and is currently limited to tens of variants. Collectively, these studies established a proof-of-concept for correlating the functional characteristics of an individual DNA molecule - referred to as its 'single-molecule phenotype' - with its corresponding DNA sequence. This phenotype may encompass structural or dynamic features of molecular interactions, such as the kinetics of a biochemical reaction at the single-molecule level, the equilibrium dynamics of binding and unbinding events, or conformational changes in a protein-DNA complex. For surface-immobilized libraries, the workflow involves immobilization of a library of distinct

DNA sequences on a solid support, a single-molecule fluorescence measurement that evaluates a single-molecule phenotype of each DNA molecule, and a step where the library complexity is interrogated via sequential transient DNA hybridization. In contrast, the solution-phase barcode strategy decouples the phenotype measurement from any surface attachment, relying instead on photophysical tag combinations for molecular identity. Both routes, however, still face scalability challenges when applied to truly large libraries.
In another study, single-molecule fluorescence imaging was combined with *in situ* single-molecule DNA sequencing to provide an endpoint measurement for identifying histone modifications and genomic positions of individual nucleosomes (*47*, *48*). However, the protocol used did not include capturing single-molecule dynamics, limiting its application to mapping populations. These pioneering efforts relied on custom methodologies, and their progression is hampered by the absence of commercial platforms for *in situ* single-molecule sequencing.

**Linking single-molecule kinetics to nucleic acid sequence**
The development of DNA sequencing approaches has often been closely intertwined with advancements in single-molecule methods, with each technology driving the other's growth and impact. For example, the pursuit of single-molecule DNA sequencing catalyzed early advances in single-molecule fluorescence spectroscopy in solution. Pioneering efforts in R. Keller's lab at Los Alamos involved attempts to sequence DNA by digesting immobilized, fluorescently labeled DNA fragments using an exonuclease, then detecting the cleaved fluorescent nucleotides as they passed through a laser-interrogated detection zone (*49*). Similarly, the original aim of the company Solexa in the 1990s was to achieve fluorescence-based single-molecule DNA sequencing. Although these initial efforts were hampered by sensitivity challenges, they ultimately led to the development of clonal-cluster DNA next-generation sequencing (NGS), later commercialized by Illumina.

Other efforts, motivated by the push for rapid and affordable genome sequencing during the Human Genome Project, were driven by single-molecule biophysics groups and led to several single-molecule technologies that made third-generation DNA sequencing (based on single-molecule detection) a reality. These efforts began with seminal proof-of-concept studies (Fig. 2) demonstrating the feasibility of reading sequence information from single DNA molecules using either fluorescence-based or force-based approaches (*50–52*) (Fig. 2A), followed by the first demonstration of single-molecule DNA sequencing (*53*) (Fig. 2B). This work culminated in technologies, such as PacBio sequencing, which relies on the real-time incorporation of labeled nucleotides in zero-mode waveguides (*54*) (Fig. 2C), and nanopore sequencing, which reads DNA sequence information by detecting electrical signal changes as individual bases of nucleic acid strands pass through a protein nanopore (*55*) (Fig. 2D).

The interplay between single-molecule methods and DNA sequencing, however, extends beyond decoding the one-dimensional sequence information for nucleic acids. From their inception, single-molecule experiments have explored the mechanical properties of biological heteropolymers - particularly DNA and RNA - and how these properties influence function and biological roles. This interest stemmed both from the rich information content of nucleic acids and from practical considerations, such as the ease of preparation, labeling, modification, and surface immobilization for single-molecule studies. From the earliest single-molecule fluorescence resonance energy transfer (FRET) measurements (*56*, *57*), researchers explored sequence variation and modifications - initially to validate methodologies, then to test hypotheses, and ultimately to uncover unexpected behaviors using powerful single-molecule techniques. Sequence dependence played a central role in many studies on ribozymes and RNA

folding kinetics (*57*, *58*), DNA hybridization (*59*), and protein-nucleic acid interactions, such as exonucleases and nucleic acid polymerases.

A key example involves the bacterial RNA polymerase (RNAP), which catalyzes DNA-templated RNA synthesis. Early single-molecule optical trapping experiments measured the force and velocity of RNAP during transcription on DNA and revealed that the RNAP motion on the DNA during transcription elongation (and the associated transcription kinetics) was influenced by DNA sequence (*60*). Subsequent studies demonstrated frequent polymerase pausing, independent of RNAP backtracking (*61*). Early single-molecule work by Greenleaf and Block (*52*) began to establish a link between RNAP pausing and DNA sequence (Fig. 2A), anticipating later genome-wide approaches. However, the sequence dependence of this ubiquitous pausing remained unclear until single-molecule measurements were combined with NGS (*62*, *63*), leading to the identification of a consensus DNA sequence associated with frequent elongation pauses and to the mechanistic dissection of these events. The power of single-molecule approaches in probing sequence-dependent functional effects was further exemplified by recent high-resolution nanopore tweezers studies, which track RNAP motion along the DNA template with unprecedented temporal and spatial resolution - resolving translocation events smaller than a single base-pair step (<3.4 Å). These studies revealed a complex kinetic landscape in which RNAP transitions between multiple transcriptional states, with these transitions being highly DNA-sequence dependent and modulated by interactions between DNA and structural elements of RNAP (*64*) (Fig. 2D). Nanopore tweezers strategies have also been applied to helicases, providing detailed mechanistic insights(*65–67*), and helicases of this type also serve as processive motors in nanopore sequencing platforms to regulate translocation through the pore (*68*). Extending beyond RNAP and helicases, similar advances have been achieved for other linear molecular machines that operate on nucleic acids, such as DNA polymerase (*69*) and the ribosome (*70*). However, the requirement for linear motion along a template constrains the scope of these technologies, necessitating the development of additional, more general methods to study the broader landscape of protein-DNA interactions and biomolecular interactions in general.

**Multiplexed single-molecule measurements**

In addressing this need, two distinct strategies for highly multiplexed single-molecule measurements (see Box 1) have recently emerged: the closely related MUSCLE (multiplexed single-molecule characterization at the library scale) (*71*) and SPARXS (single-molecule parallel analysis for rapid exploration of sequence space) (*72*) methods on the one hand, and SPIN-Seq (single-molecule phenotyping and *in-situ* sequencing) (*73*) on the other (Fig. 3).

*Integrating single-molecule fluorescence with Illumina sequencing*

The MUSCLE and SPARXS methods integrate single-molecule fluorescence microscopy with Illumina next-generation sequencing (*74*) to monitor complex dynamic behaviors at the single-molecule level while subsequently determining their sequence (*71*, *72*). The workflow begins with the design of a molecular library that includes sequencing adapters for immobilization on the surface of an Illumina MiSeq flow cell, ensuring compatibility with both single-molecule imaging and sequencing (Fig. 3A). The MiSeq flow cells are first examined by single-molecule fluorescence imaging of surface-bound constructs using objective-type total internal reflection fluorescence (TIRF) microscopy. With highly automated TIRF microscopy, a large number of single molecules can be tracked over time - typically on the millisecond to minute timescale - revealing dynamic processes and structural changes.

Following this, Illumina's bridge amplification process generates clusters of identical copies from single template molecules, which are then subjected to sequencing by synthesis. This sequencing method produces a FASTQ file, a widely used format for storing sequence data alongside quality scores. In the case of Illumina sequencing, the FASTQ file also records the position of each sequencing tile and cluster positions within the tile. In the final stage, each molecule's sequence is coupled to its corresponding fluorescence by linking the sequencing position to the spatial coordinates of the single-molecule time traces, yielding a dataset containing the sequence of hundreds of thousands to millions of single-molecule fluorescence time trajectories. These data can then be analyzed using various computational and visualization approaches to elucidate sequence-dependent molecular dynamics in libraries containing up to tens of thousands of members. To facilitate adoption of MUSCLE/SPARXS, standardize experimental procedures, and enhance reproducibility across laboratories, we have made the accompanying software and test datasets freely available and have published detailed protocol papers (*75*, *76*).

The most direct application of this approach is to investigate how DNA sequence influences conformational equilibria and dynamic behavior of structured DNA elements. For example, SPARXS has enabled the systematic analysis of the sequence dependence of the Holliday junction core across the complete sequence space (*72*) (Fig. 3B). While the influence of core and flanking sequences on junction dynamics has been recognized for decades (*77*, *78*), a quantitative high-resolution experimental study had not been made possible until recently. By applying SPARXS, the authors examined all 4096 possible core sequences of the junction, generating a rich dataset that allowed them to construct a thermodynamic model capturing the underlying kinetics with high accuracy. This analysis revealed that both the core and penultimate bases contribute significantly to the dynamic behavior of the junction - a concept that had been qualitatively proposed, but not quantitatively validated, until now. Furthermore, the model enabled direct estimation of dinucleotide stacking energies, which showed strong agreement with values obtained from bulk assays. As another example involving only DNA, MUSCLE has been used to systematically characterize DNA hairpin structures using libraries of up to 4096 distinct sequences, examining their properties at several different temperatures (*71*).

To illustrate how this approach can be extended to protein-nucleic acid interactions, MUSCLE has also been applied to investigate the unwinding-rewinding dynamics of target DNA induced by Cas9 (*71*), a programmable nuclease widely used in research, biotechnology, and gene editing (*79*, *80*) (Fig. 3C). These analyses revealed that the rate of DNA rewinding is the primary determinant of the equilibrium between unwound and rewound states. Target sequences that remained longer in the unwound conformation were more efficiently cleaved, consistent with a key role for the unwinding-rewinding equilibrium in Cas9 activity. Furthermore, by exploring a broad sequence space, we identified several target sequences with unexpected behaviors that substantially deviated from the general trend. This highlights a major advantage of comprehensive, high-throughput analyses: their independence from *a priori* assumptions, which enhances their capacity to reveal unanticipated phenomena.

*SPIN-Seq: Single-molecule base calling and reaction monitoring*

Large-scale single-molecule profiling of DNA libraries can, however, also be achieved without cluster formation and subsequent sequencing by synthesis. In a novel sequencing-by-hybridization method, called Gap-Seq, DNA sequences featuring short single-stranded gaps can be identified from transient binding of short fluorescent DNA oligos by automated single-molecule fluorescence microscopy (*44*). Building upon this, SPIN-Seq provides an additional

approach for single-molecule multiplexing (*73*). Demonstrating the ability to work with libraries, Hazra et al. related the single-molecule phenotype to its corresponding DNA sequence by performing both the functional and the sequencing assay on the same solid support and instrument (Fig 3D). The protocol starts with a DNA library featuring permutations of the sequence segment under investigation that are immobilized randomly on a glass surface. The DNA library is then interrogated for interactions with a protein of interest (or other biomolecules) using single-molecule fluorescence imaging, enabling the observation of thousands of individual molecules. These interactions also include real-time, non-equilibrium reactions driven by large multi-subunit proteins. Finally, SPIN-Seq is used to determine the sequence of each DNA molecule and link it to the corresponding single-molecule phenotype. The sequencing assay currently achieves high accuracy for sequences up to five nucleotides in length, thereby covering up to 1,024 distinct DNA sequences. This approach enabled detailed investigation of RNA polymerase kinetics during initial transcription (Fig. 3D), providing substantial insight into the mechanisms underlying key transitions in this complex reaction pathway.

**Expanding the impact of multiplexed single-molecule studies**

The ability to probe large sequence libraries with single-molecule techniques marks a major advance in molecular biology, enabling the connection of sequence identity with structural information and functional kinetics across millions of molecules in parallel. To fully realize its potential, future efforts should focus on further increasing throughput, refining approaches to study non-equilibrium enzymatic processes, and broadening applications beyond nucleic acids (Fig. 4).

*Overcoming current limitations in throughput*

One of the main challenges–despite highly automated implementations–remains throughput, particularly when compared to NSG platforms used in MUSCLE/SPARXS. The Illumina MiSeq, for example, can yield several times $10^7$ reads. However, when combined with sequential single-molecule fluorescence imaging, the yield reduces to only ~$10^6$ of linked molecules, due to incomplete cluster formation, bleaching events, overlapping positions, and stringent filtering of sequencing reads and time traces. Alternatively, SPIN-Seq has demonstrated high accuracy for sequences up to five nucleotides in length, thereby covering up to 1,024 distinct DNA sequences. However, as sequence length increases, the number of possible sequences grows exponentially. For example, a 10-nucleotide sequence space encompasses over 1,048,576 unique sequences. This rapid expansion makes complete sequence sampling beyond 10 nucleotides impractically time-consuming, even with increased throughput.

A key strategy for increasing the efficiency of data collection is to expand the field of view (FOV) (Fig. 4A). Modern sCMOS cameras offer higher resolution than the EMCCD cameras used in early multiplexed single-molecule implementations. Utilizing the full FOV could enhance acquisition efficiency by up to an order of magnitude. Additionally, for slower kinetic processes occurring on timescales of several seconds or more, imaging efficiency can be further increased by cycling through multiple FOVs sequentially at each time point. Acquisition efficiency can also be enhanced by methods such as DyeCycling (*81*) and REFRESH (*82*), which sustain donor fluorescence in single-molecule FRET by continuously exchanging photobleached dyes during TIRF excitation, thereby enabling longer observation windows for kinetic analysis. Advances in hardware and software automation, alongside optimized imaging protocols to maximize the fraction of molecules that can be successfully linked, will be critical

in improving the accessibility, efficiency and throughput of multiplexed single-molecule studies.

It is important to note, however, that the strength of multiplexed single-molecule approaches lies not in replacing high-throughput sequencing but in complementing it by providing kinetic and mechanistic insights that bulk sequencing cannot capture. As such, single-molecule sequence profiling may be best suited for targeted studies, focusing on biologically relevant sequences rather than random libraries. A genome-referenced approach, as demonstrated in loop-seq (*2*) provides a promising strategy for prioritizing relevant sequences. By anchoring analysis to a well-defined reference framework, such as the yeast genome or specific regulatory elements, kinetic parameters from tens of thousands of sequences involved in key processes such as transcription, translation, and chromatin dynamics can be efficiently extracted.

When it is impractical to exhaustively sample the vast sequence space, machine learning (ML) offers a powerful tool for extracting meaningful insights from incomplete datasets. ML can identify patterns, classify molecular behaviors, and infer missing kinetic parameters from sparsely sampled data.
Nonetheless, broad sequence scans remain essential, as they often uncover rare behaviors that ML models trained on limited data may miss. Recent work shows that large language models (LLMs) trained on biological sequences and kinetic measurements can capture rich sequence-function relationships (*83*), suggesting that the high-throughput datasets produced by MUSCLE, SPARXS, and SPIN-Seq could serve as particularly powerful training material for such models. Conversely, single-molecule sequence profiling can serve as an experimental platform to validate ML-predicted interaction kinetics and binding specificities. The iterative feedback between data-driven prediction, continuous experimental testing, and refined modeling will ultimately enhance the interpretability and utility of these technologies.

*Advancing methodology: non-equilibrium studies*
Many fundamental biological processes - such as transcription, translation, and ligand binding - occur under non-equilibrium conditions and involve transient molecular states that are difficult to capture using ensemble methods. A key challenge is the study of nucleoside triphosphate (NTP)-hydrolysing enzymes, such as helicases and polymerases, which drive irreversible modifications of biological macromolecules. In single-molecule experiments, these non-equilibrium processes are typically initiated by perfusing the flow cell with enzymes and/or NTP cofactors. Only one field of view (FOV) can be imaged because, after data acquisition, the reaction has already occurred across the entire surface of the flow cell. Consequently, only a small fraction of the substrates can be observed. To overcome this issue, photo-uncaging strategies such as LAGOON (*84*) provide spatiotemporal control over enzymatic reactions, enabling precise synchronization between reaction initiation and imaging. Integrating LAGOON or other light-based reaction control strategies (*85–87*) will therefore enable multiplexed single-molecule observations of complex non-equilibrium dynamics.

*Broadening applications*
While multiplexed single-molecule studies have primarily been applied to DNA libraries and can be readily adapted to RNA, its methodology can also be extended to protein libraries, thereby broadening its impact in molecular biology (Fig. 4B). A key application is the study of protein-protein and peptide-protein interactions, which play a crucial role in critical processes such as signal transduction and immune recognition. By incorporating cDNA display or ribosome display for peptide and protein libraries (*88*, *89*), MUSCLE/SPARXS or SPIN-Seq could enable high-throughput kinetic analyses of millions of individual peptides and proteins

in parallel. Recent advances have paved the way for such approaches, including the profiling of DNA-barcoded protein interactions at the single-molecule level (*90*), real-time kinetic screening of thousands of protein interactions using multiplexed biosensors (*91*), and the integration of ribosome display with deep sequencing for rapid antibody discovery (*92*).

Multiplexed single-molecule profiling also holds promise for investigating protein-ligand interactions, which play central roles in receptor signaling and drug binding (Fig. 4C). A notable example is the use of nucleic acid aptamers - short, structured DNA or RNA molecules selected for high-affinity binding to specific protein targets - which can be integrated into libraries for parallel interaction profiling. Traditional HTS methods rely primarily on affinity-based measurements, which fail to capture critical kinetic parameters, such as association and dissociation rates, which are key determinants of drug efficacy and specificity. Multiplexed single-molecule studies offer an alternative, enabling real-time kinetic screening of ligand-target interactions across large target and chemical libraries. For example, while DNA-encoded chemical libraries (DELs) (*93–95*) contain millions to billions of small molecules, their integration with single-molecule kinetic measurements remains an unexplored frontier. Bridging this gap could allow multiplexed single-molecule approaches to identify compounds with optimal kinetic profiles for therapeutic development.

While the MUSCLE/SPARXS and SPIN-Seq approaches thus far have been utilized primarily for smFRET experiments, they are also compatible with other single-molecule fluorescence techniques, including colocalization analysis and protein-induced fluorescence enhancement (PIFE) (*96*). Furthermore, single-molecule localization microscopy methods can be integrated, such as PALM, STORM, and DNA-PAINT (*97–99*), enabling the exploration of how sequence variations influence nano- and microscale structural organization. Many biological processes involve structural transitions that span hundreds of nanometers to micrometers. Investigating such large-scale dynamics often requires techniques capable of probing molecular mechanics over extended distances while maintaining single-molecule resolution. Integrating force-spectroscopy methods, particularly magnetic tweezers (*100–103*), into the MUSCLE/SPARXS and SPIN-Seq approaches would enable high-throughput studies of these long-range biomolecular interactions (Fig. 4D). Additionally, large rigid DNA scaffolds constructed via DNA origami (*104–107*) can be readily combined with oligo-nucleotide libraries and integrated, enabling the study of large-scale dynamics beyond the spatial range accessible by smFRET. Building on the Origami-rotor-based imaging and tracking (ORBIT) approach (*108*), DNA origami can also be used within multiplexed single-molecule studies to access a qualitatively different mode of motion - DNA rotation - that is not readily detectable by smFRET. However, overly large structures may interfere with cluster formation, and experimental validation of these approaches may reveal additional practical limitations.

While the MUSCLE/SPARXS and SPIN-Seq approaches enable the direct linkage of complex molecular dynamics to sequence identity, they are inherently *in vitro*-based and may not fully capture the dynamic behaviors occurring within native cellular environments. One strategy to bridge this gap involves the use of cellular extracts applied to immobilized nucleic acid libraries, thereby recreating more physiologically relevant conditions. Additionally, the emergence of *in situ* genotyping and sequencing techniques - such as rolling circle amplification (RCA), barcoding strategies, and related methods (*109–112*) - has opened new avenues for coupling genotypic information with multiplexed phenotypic readouts. Proof-of-concept studies for optical pooled screening in both bacterial and mammalian systems have already demonstrated the feasibility of combining live-cell phenotyping with *in situ* genotyping or barcoded readouts (*113–116*).

Naturally, the *in vivo* environment presents major experimental challenges, including limited spatiotemporal resolution, lower signal-to-noise ratios, and constraints on the labeling and detection of molecular species (*117–121*). Substantial technological advances will be required to bring *in vivo* single-molecule approaches closer to their *in vitro* counterparts in terms of data quality. However, such developments may one day provide similarly granular mechanistic insights within native cellular contexts (Fig. 4E).

**A powerful reunion of sequencing and single-molecule biophysics**

Taken together, these developments underscore the ongoing convergence of two foundational technologies in modern biology: sequencing and single-molecule biophysics. By fusing the throughput and scalability of advanced sequencing technologies with the mechanistic and dynamic insight of single-molecule imaging, we are now poised to interrogate complex biological phenomena with previously unattainable depth and precision. Whether on a flow cell or within a living cell, the ability to link molecular identity with dynamic behavior - across millions of distinct molecules - ushers in a new era of discovery. This represents the latest chapter in a recurring narrative - where sequencing and single-molecule biophysics repeatedly intersect to advance our understanding of life's molecular underpinnings.

**Boxes:**

**Box 1. Defining "Multiplex" in Single-Molecule Methods**

The term multiplex in the context of single-molecule techniques can refer to distinct experimental strategies. One common usage simply denotes the parallel measurement of many individual molecules of the same type. These approaches enhance throughput and statistical resolution and are broadly applied in high-throughput single-molecule studies (*122–124*).

In this review, we adopt a more specific definition: multiplex single-molecule methods that resolve and monitor multiple distinct molecular species within the same experiment. Here, multiplexing refers to molecular identity - such as differences in sequence - being discerned at the single-molecule level. These methods enable direct, comparative analyses of heterogeneous populations under identical conditions. Notably, commercial single-molecule sequencing platforms (e.g., PacBio and nanopore) already operate in a massively parallel fashion with respect to sequence readout. However, they have generally not been used to multiplex dynamic single-molecule phenotypes in the sense defined here - our focus is on methods that explicitly couple sequence identity to kinetic or conformational readouts across many distinct molecules in the same experiment.

**Figures:**

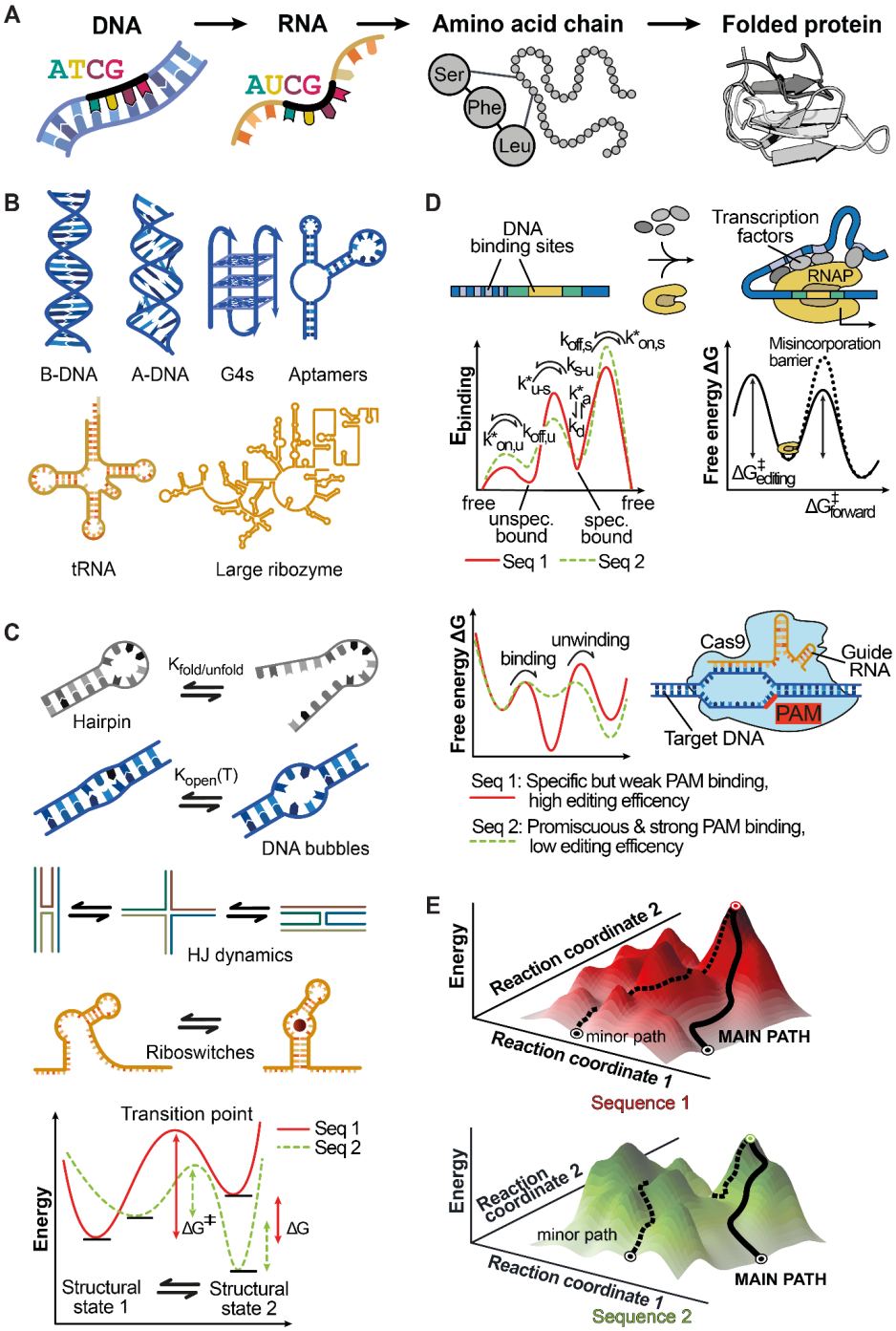


**Fig. 1. Sequence variation modulates structure, dynamics, and interactions of biomolecules across multiple levels.** (**A**) Schematic overview of the hierarchical information flow from DNA to folded protein according to the central dogma of molecular biology, illustrating how DNA sequence not only encodes amino acid identity but also shapes structural and dynamic properties at each stage of molecular processing.
(**B**) The sequences of DNA and RNA encode their characteristic secondary and tertiary structures. Shown are canonical duplex forms such as B-DNA and A-DNA, as well as noncanonical motifs including G-quadruplexes and structured aptamers. Common RNA secondary structures include stem-loop hairpins, while fully folded tRNAs and ribozymes exhibit complex tertiary architectures shaped by sequence context and base identity. tRNA structure adapted from ToGo Picture Gallery (DOI:10.7875/togopic.2024.06), CC BY 4.0.
(**C**) Sequence variation shapes energetic landscapes and transition kinetics of folding and unfolding pathways. Shown are examples of force-induced hairpin unfolding, DNA bubble formation, and conformational exchange within structured RNA or DNA segments, such as Holliday junctions (HJ). (**D**) DNA sequence controls interactions with proteins, including the sequence-dependent recognition by transcription factors and RNA polymerase (top) as well as the sequence-dependent Cas9 binding and target DNA unwinding (bottom). Energy landscapes illustrate sequence-specific kinetic barriers and intermediates during enzyme function. *In vivo*

transcription factor binding energy landscape adapted from (*125*). Indicated are energy differences between bound and free states (including transition barriers) and the kinetic rates of target-site search by facilitated diffusion. Genome editing energy landscape adapted from (*126*) (**E**) In complex multi-step biological processes, different sequences give rise to distinct energy topographies. Sequence variation reshapes both the geometry of the energy surface and the accessibility of kinetic pathways, favoring distinct major and minor pathways depending on barrier heights, cooperative transitions, and environmental constraints.

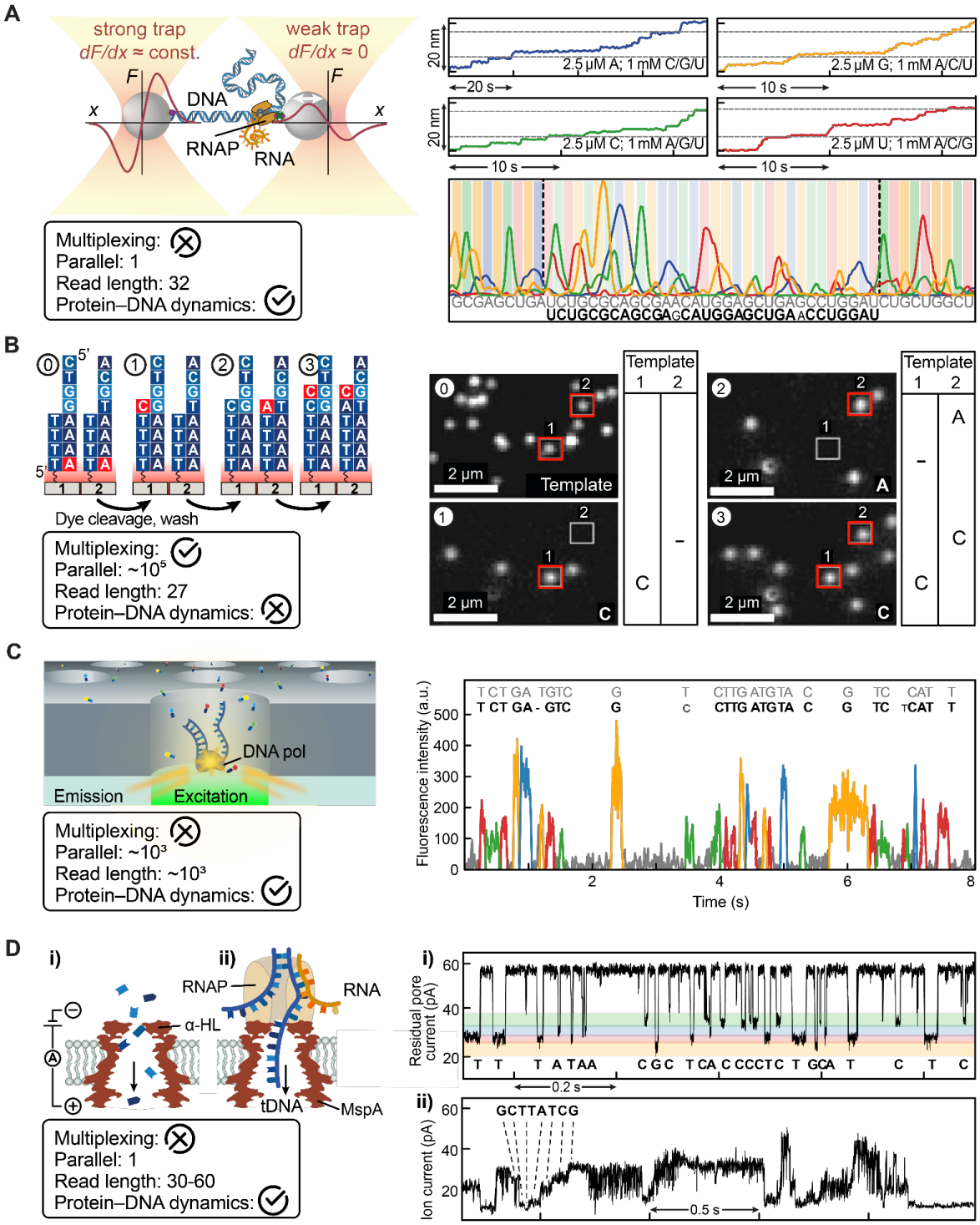


**Fig. 2. Early single-molecule approaches linking DNA sequence to molecular function.** (**A**) Motion-based DNA sequencing using optical tweezers. Greenleaf and Block (*52*, *127*) tracked single RNA polymerase (RNAP)

molecules transcribing in dual optical traps. Limiting the concentration of one type of nucleotide at a time caused site-specific pauses, producing distinct patterns that reveal sequence. Aligned transcription traces under different conditions generate position histograms, from which base identity is inferred. Adapted from (*52*, *127*). (**B**) Early sequencing-by-synthesis with asynchronous base incorporation. Harris et al. (*53*) hybridized poly(dA)-tailed templates to surface-tethered poly(dT) primers. Each cycle added a single fluorescently labeled nucleotide, followed by polymerase extension, imaging, dye cleavage, and washing. Fluorescence at fixed positions across cycles allowed direct single-molecule reads, laying groundwork for today's multiplexed platforms. Adapted from (*53*). (**C**) Real-time sequencing in zero-mode waveguides (ZMWs). Eid et al. (*54*) used immobilized DNA polymerase and fluorescently labeled dNTPs confined in ZMWs. Fluorophores on terminal phosphates allowed real-time detection without inhibiting enzymatic activity. Early traces show color-coded incorporation events across a ~28-base read, capturing nucleotide identity and polymerase kinetics. Adapted from (*54*). Commercial implementations monitor tens of thousands of ZMWs in parallel, but a single ZMW is depicted for clarity. (**D**) Nanopore-based methods linking sequence with dynamics. (i) Clarke et al. (*55*) demonstrated sequencing through a modified α-hemolysin pore with an adapter, detecting individual nucleotides via ionic current signatures - an early foundation for modern nanopore sequencing. Adapted from (*55*). Modern nanopore sequencers use large arrays of pores and molecular barcoding to achieve massive parallelization; here, a single pore is shown to illustrate the underlying principle. (ii) Nova et al. (*64*) combined nanopore sensing with force application. Using engineered templates and transcription complexes, RNAP was held at the pore rim while voltage-induced translocation revealed sub-nucleotide steps and sequence-dependent pausing, capturing both dynamic and mechanical features of transcription. Reproduced from (*64*).

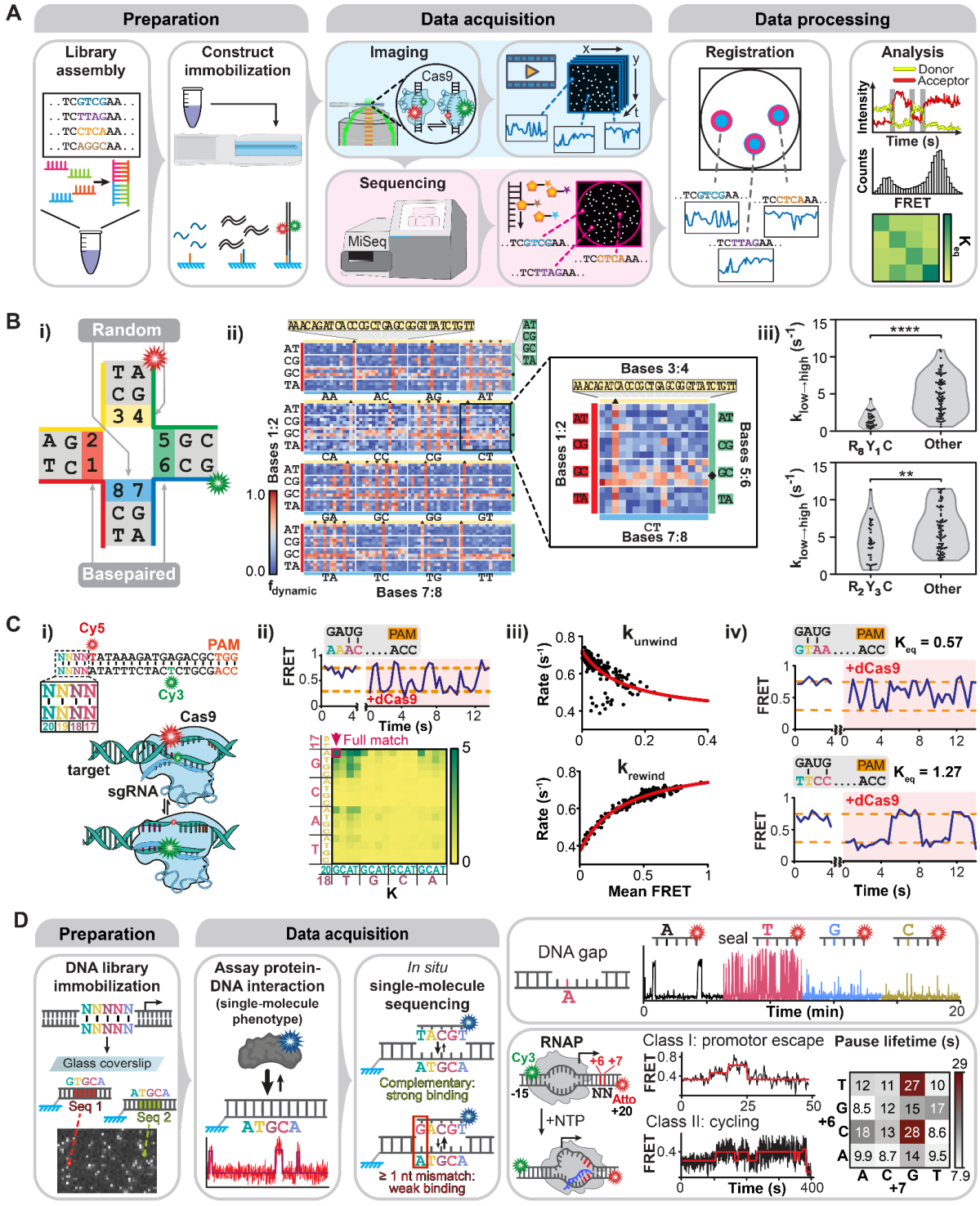


**Fig. 3. Quantitative insights from multiplexed single-molecule platforms reveal sequence-dependent variation in kinetic and thermodynamic behavior.** (**A**) Overview of a MUSCLE/SPARXS experiment (example: Cas9 DNA unwinding). A DNA sequence library is synthesized and immobilized on a flow cell. Single-molecule fluorescence imaging yields time traces from individual molecules. The same flow cell undergoes next-generation sequencing to identify the sequence and spatial position of each DNA cluster. Fluorescence and sequencing data are aligned to extract kinetic or equilibrium parameters - such as FRET states or transitions - per sequence. Adapted from the authors' previous work (*9*, *71*, *76*).

(**B**) Sequence-dependent Holliday junction (HJ) dynamics in a 4096-member SPARXS library. (i) Bases 3, 4, 7, and 8 are randomized. (ii) Heatmap of the dynamic molecule fraction across variants, with clustering seen at bases 7:8. Stars: fully matched junctions. (iii) Transition rates are reduced in sequences with the RYC motif. Adapted from (*72*).

(**C**) Cas9 unwinding and cleavage dynamics *via* MUSCLE. (i) smFRET assay with a DNA library varying four PAM-distal bases. (ii) FRET traces for a mismatched sequence (top) and equilibrium constants for all sequences (bottom), organized by identity at positions 17–20. (iii) Unwinding/rewinding rates plotted *versus* mean FRET, with model fits. (iv) FRET traces and cleavage for selected mismatched sequences show impaired cleavage correlating with FRET, linking kinetics to outcomes. Adapted from (*71*).
(**D**) SPIN-Seq. Left: Immobilized DNA libraries are imaged for biomolecular interactions (e.g., protein binding kinetics), denatured, and rehybridized. Sequence is read by hybridizing fluorescent "seals" to gapped DNA; seal binding distinguishes complementary from mismatched sequences. Top right: sequential interrogation of a DNA gap by four seals, with T-seal as the match. Bottom right: smFRET monitoring of transcription initiation reveals distinct promoter DNA pathways and measures pause lifetimes across a 16-sequence library (+6/+7 degenerate bases). Adapted from the authors' work (*44*, *73*).

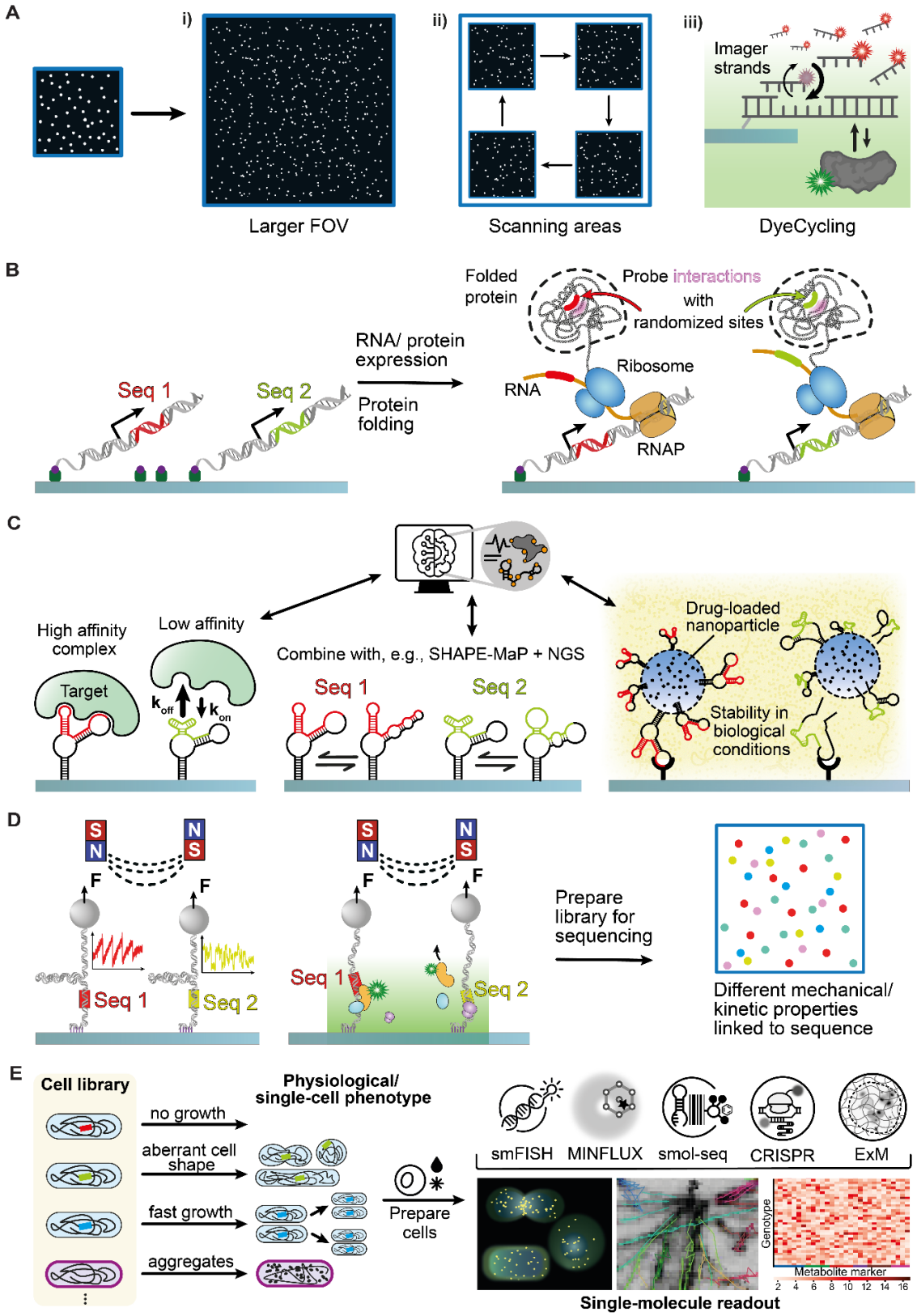


**Fig. 4. Expanding multiplexed single-molecule analysis across molecular scales and applications.** (**A**) Increasing data collection efficiency. Throughput can be maximized by using full sCMOS camera fields of view (i) or cycling through multiple FOVs sequentially for slower kinetic processes (ii). DyeCycling (*81*) sustains donor

fluorescence in single-molecule FRET by continuously exchanging photobleached dyes during TIRF excitation, enabling longer observation windows for kinetic analysis and increasing the statistical power of equilibrium binding measurements from fewer molecules. (iii). (**B**) Peptide/protein libraries. DNA-barcoded peptide and protein libraries allow single-molecule kinetic screening of binding, folding, or enzymatic activity. Techniques such as ribosome and cDNA display link behavior to sequence, enabling functional profiling across millions of variants. Applications span signaling, immunity, and protein engineering. (**C**) Aptamer characterization. High-throughput screening of nucleic acid aptamers at the single-molecule level can uncover folding pathways, binding kinetics, and structural dynamics tied to sequence. This enables identification of aptamers with high specificity, adaptability, and functional stability for therapeutic, diagnostic, and analytical uses. (**D**) Sequence multiplexing in other single-molecule methods, such as magnetic tweezers. Incorporating sequence-encoded tethers enables parallel force spectroscopy on multiple DNA or RNA constructs, facilitating comparison of mechanical responses or enzymatic processing. These platforms can also include colocalization-based fluorescence readouts for monitoring protein dynamics or substrate conformational changes alongside mechanical data. (**E**) Expansion to cell libraries. Linking single-molecule readouts to cell-based assays involves using genetically defined cell libraries combined with phenotypic sorting (e.g., growth, stress). Cells can be analyzed for molecular structure, localization, and dynamics using FRET, MINFLUX (*128*), Expansion Microscopy (*129*), and smFISH (*130*, *131*), connecting genotype, phenotype, and mechanism in a unified experimental pipeline. Visual examples are stylized illustrations inspired by reported single-molecule observations in cells.

## References:


1. E. Oberbeckmann, N. Krietenstein, V. Niebauer, Y. Wang, K. Schall, M. Moldt, T. Straub, R. Rohs, K.-P. Hopfner, P. Korber, S. Eustermann, Genome information processing by the INO80 chromatin remodeler positions nucleosomes. *Nat Commun* **12**, 3231 (2021).
2. A. Basu, D. G. Bobrovnikov, Z. Qureshi, T. Kayikcioglu, T. T. M. Ngo, A. Ranjan, S. Eustermann, B. Cieza, M. T. Morgan, M. Hejna, H. T. Rube, K.-P. Hopfner, C. Wolberger, J. S. Song, T. Ha, Measuring DNA mechanics on the genome scale. *Nature* **589**, 462–467 (2021).
3. J. M. Tome, A. Ozer, J. M. Pagano, D. Gheba, G. P. Schroth, J. T. Lis, Comprehensive analysis of RNA-protein interactions by high-throughput sequencing-RNA affinity profiling. *Nat. Methods* **11**, 683–688 (2014).
4. R. Nutiu, R. C. Friedman, S. Luo, I. Khrebtukova, D. Silva, R. Li, L. Zhang, G. P. Schroth, C. B. Burge, Direct measurement of DNA affinity landscapes on a high-throughput sequencing instrument. *Nat. Biotechnol.* **29**, 659–664 (2011).
5. C. Jung, J. A. Hawkins, S. K. Jones Jr, Y. Xiao, J. R. Rybarski, K. E. Dillard, J. Hussmann, F. A. Saifuddin, C. A. Savran, A. D. Ellington, A. Ke, W. H. Press, I. J. Finkelstein, Massively Parallel Biophysical Analysis of CRISPR-Cas Complexes on Next Generation Sequencing Chips. *Cell* **170**, 35–47.e13 (2017).
6. R. E. Kleiner, C. E. Dumelin, G. C. Tiu, K. Sakurai, D. R. Liu, In vitro selection of a DNA-templated small-molecule library reveals a class of macrocyclic kinase inhibitors. *J. Am. Chem. Soc.* **132**, 11779–11791 (2010).
7. U. T. T. Nguyen, L. Bittova, M. M. Müller, B. Fierz, Y. David, B. Houck-Loomis, V. Feng, G. P. Dann, T. W. Muir, Accelerated chromatin biochemistry using DNA-barcoded nucleosome libraries. *Nat. Methods* **11**, 834–840 (2014).
8. G. P. Dann, G. P. Liszczak, J. D. Bagert, M. M. Müller, U. T. T. Nguyen, F. Wojcik, Z. Z. Brown, J. Bos, T. Panchenko, R. Pihl, S. B. Pollock, K. L. Diehl, C. D. Allis, T. W. Muir, ISWI chromatin remodellers sense nucleosome modifications to determine substrate preference. *Nature* **548**, 607–611 (2017).
9. I. Severins, C. Joo, J. van Noort, Exploring molecular biology in sequence space: The road to next-generation single-molecule biophysics. *Mol. Cell* **82**, 1788–1805 (2022).
10. J. B. Kinney, D. M. McCandlish, Massively Parallel Assays and Quantitative Sequence-Function Relationships. *Annu. Rev. Genomics Hum. Genet.* **20**, 99–127 (2019).

11. E. Marklund, Y. Ke, W. J. Greenleaf, High-throughput biochemistry in RNA sequence space: predicting structure and function. *Nat. Rev. Genet.*, doi: 10.1038/s41576-022-00567-5 (2023).
12. C. J. Layton, P. L. McMahon, W. J. Greenleaf, Large-Scale, Quantitative Protein Assays on a High-Throughput DNA Sequencing Chip. *Mol. Cell* **73**, 1075–1082.e4 (2019).
13. E. A. Boyle, J. O. L. Andreasson, L. M. Chircus, S. H. Sternberg, M. J. Wu, C. K. Guegler, J. A. Doudna, W. J. Greenleaf, High-throughput biochemical profiling reveals sequence determinants of dCas9 off-target binding and unbinding. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 5461–5466 (2017).
14. N. Svensen, O. B. Peersen, S. R. Jaffrey, Peptide Synthesis on a Next-Generation DNA Sequencing Platform. *Chembiochem* **17**, 1628–1635 (2016).
15. L. C. Szymczak, H.-Y. Kuo, M. Mrksich, Peptide Arrays: Development and Application. *Anal. Chem.* **90**, 266–282 (2018).
16. D. Wu, T. Feagin, P. Mage, A. Rangel, L. Wan, D. Kong, A. Li, J. Coller, M. Eisenstein, H. Soh, Flow-Cell-Based Technology for Massively Parallel Characterization of Base-Modified DNA Aptamers. *Anal. Chem.* **95**, 2645–2652 (2023).
17. Z. Li, X. Wang, D. Xu, D. Zhang, D. Wang, X. Dai, Q. Wang, Z. Li, Y. Gu, W. Ouyang, S. Zhao, B. Huang, J. Gong, J. Zhao, A. Chen, Y. Shen, Y. Dong, W. Zhang, X. Xu, C. Xu, Y. Jiang, DNB-based on-chip motif finding: A high-throughput method to profile different types of protein-DNA interactions. *Sci Adv* **6**, eabb3350 (2020).
18. N. Mamet, I. Rusinek, G. Harari, Z. Shapira, Y. Amir, E. Lavi, A. Zamir, N. Borovsky, N. Joseph, M. Motin, D. Saban, I. Bachelet, Ab-initio discovery of tumoricidal oligonucleotides in a DNA sequencing machine, *bioRxiv* (2019)p. 630830.
19. J. D. Buenrostro, C. L. Araya, L. M. Chircus, C. J. Layton, H. Y. Chang, M. P. Snyder, W. J. Greenleaf, Quantitative analysis of RNA-protein interactions on a massively parallel array reveals biophysical and evolutionary landscapes. *Nat. Biotechnol.* **32**, 562–568 (2014).
20. A. Melnikov, A. Murugan, X. Zhang, T. Tesileanu, L. Wang, P. Rogov, S. Feizi, A. Gnirke, C. G. Callan Jr, J. B. Kinney, M. Kellis, E. S. Lander, T. S. Mikkelsen, Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay. *Nat Biotechnol* **30**, 271–277 (2012).
21. R. Tewhey, D. Kotliar, D. S. Park, B. Liu, S. Winnicki, S. K. Reilly, K. G. Andersen, T. S. Mikkelsen, E. S. Lander, S. F. Schaffner, P. C. Sabeti, Direct Identification of Hundreds of Expression-Modulating Variants using a Multiplexed Reporter Assay. *Cell* **165**, 1519–1529 (2016).
22. D. M. Fowler, C. L. Araya, S. J. Fleishman, E. H. Kellogg, J. J. Stephany, D. Baker, S. Fields, High-resolution mapping of protein sequence-function relationships. *Nat Methods* **7**, 741–746 (2010).
23. C. L. Araya, D. M. Fowler, Deep mutational scanning: assessing protein function on a massive scale. *Trends Biotechnol* **29**, 435–442 (2011).
24. A. Jolma, J. Yan, T. Whitington, J. Toivonen, K. R. Nitta, P. Rastas, E. Morgunova, M. Enge, M. Taipale, G. Wei, K. Palin, J. M. Vaquerizas, R. Vincentelli, N. M. Luscombe, T. R. Hughes, P. Lemaire, E. Ukkonen, T. Kivioja, J. Taipale, DNA-binding specificities of human transcription factors. *Cell* **152**, 327–339 (2013).
25. S. Mohapatra, C.-T. Lin, X. A. Feng, A. Basu, T. Ha, Single-Molecule Analysis and Engineering of DNA Motors. *Chem. Rev.* **120**, 36–78 (2020).
26. L. Bacic, A. Sabantsev, S. Deindl, Recent advances in single-molecule fluorescence microscopy render structural biology dynamic. *Curr. Opin. Struct. Biol.* **65**, 61–68 (2020).
27. F. R. Hill, E. Monachino, A. M. van Oijen, The more the merrier: high-throughput

single-molecule techniques. *Biochem. Soc. Trans.* **45**, 759–769 (2017).
28. J. Choi, R. Grosely, E. V. Puglisi, J. D. Puglisi, Expanding single-molecule fluorescence spectroscopy to capture complexity in biology. *Curr. Opin. Struct. Biol.* **58**, 233–240 (2019).
29. M. Orrit, T. Ha, V. Sandoghdar, Single-molecule optical spectroscopy. *Chem. Soc. Rev.* **43**, 973–976 (2014).
30. F. Kulzer, M. Orrit, Single-molecule optics. *Annu. Rev. Phys. Chem.* **55**, 585–611 (2004).
31. M. Dangkulwanich, T. Ishibashi, L. Bintu, C. Bustamante, Molecular mechanisms of transcription through single-molecule experiments. *Chem. Rev.* **114**, 3203–3223 (2014).
32. J. Zhou, V. Schweikhard, S. M. Block, Single-molecule studies of RNAPII elongation. *Biochim. Biophys. Acta* **1829**, 29–38 (2013).
33. J. Michaelis, B. Treutlein, Single-molecule studies of RNA polymerases. *Chem. Rev.* **113**, 8377–8399 (2013).
34. L. Bai, T. J. Santangelo, M. D. Wang, Single-molecule analysis of RNA polymerase transcription. *Annu. Rev. Biophys. Biomol. Struct.* **35**, 343–360 (2006).
35. M. Stracy, A. N. Kapanidis, Single-molecule and super-resolution imaging of transcription in living bacteria. *Methods* **120**, 103–114 (2017).
36. E. Lerner, T. Cordes, A. Ingargiola, Y. Alhadid, S. Chung, X. Michalet, S. Weiss, Toward dynamic structural biology: Two decades of single-molecule Förster resonance energy transfer. *Science* **359** (2018).
37. D. Dulin, B. A. Berghuis, M. Depken, N. H. Dekker, Untangling reaction pathways through modern approaches to high-throughput single-molecule force-spectroscopy experiments. *Curr. Opin. Struct. Biol.* **34**, 116–122 (2015).
38. M. F. Juette, D. S. Terry, M. R. Wasserman, Z. Zhou, R. B. Altman, Q. Zheng, S. C. Blanchard, The bright future of single-molecule fluorescence imaging. *Curr. Opin. Chem. Biol.* **20**, 103–111 (2014).
39. J. H. Felce, S. J. Davis, D. Klenerman, Single-Molecule Analysis of G Protein-Coupled Receptor Stoichiometry: Approaches and Limitations. *Trends Pharmacol. Sci.* **39**, 96–108 (2018).
40. J. Camunas-Soler, M. Ribezzi-Crivellari, F. Ritort, Elastic Properties of Nucleic Acids by Single-Molecule Force Spectroscopy. *Annu. Rev. Biophys.* **45**, 65–84 (2016).
41. S. Park, T. Ha, G. D. Bowman, Nucleosome sliding by the Chd1 chromatin remodeler relies on the integrity of the DNA duplex, *bioRxiv* (2022)p. 2022.04.19.488687.
42. A. A. Hoskins, J. Gelles, M. J. Moore, New insights into the spliceosome by single molecule fluorescence microscopy. *Curr. Opin. Chem. Biol.* **15**, 864–870 (2011).
43. S. Deindl, X. Zhuang, Monitoring conformational dynamics with single-molecule fluorescence energy transfer: applications in nucleosome remodeling. *Methods Enzymol.* **513**, 59–86 (2012).
44. R. Andrews, H. Steuer, A. El-Sagheer, A. Mazumder, H. el Sayyed, A. Shivalingam, T. Brown, A. N. Kapanidis, Transient DNA binding to gapped DNA substrates links DNA sequence to the single-molecule kinetics of protein-DNA interactions, *bioRxiv* (2022)p. 2022.02.27.482175.
45. K. Makasheva, L. C. Bryan, C. Anders, S. Panikulam, M. Jinek, B. Fierz, Multiplexed Single-Molecule Experiments Reveal Nucleosome Invasion Dynamics of the Cas9 Genome Editor. *J. Am. Chem. Soc.* **143**, 16313–16319 (2021).
46. J. Chu, A. Ejaz, K. M. Lin, M. R. Joseph, A. E. Coraor, D. A. Drummond, A. H. Squires, Single-molecule fluorescence multiplexing by multi-parameter spectroscopic detection of nanostructured FRET labels. *Nat Nanotechnol* **19**, 1150–1157 (2024).
47. E. Shema, D. Jones, N. Shoresh, L. Donohue, O. Ram, B. E. Bernstein, Single-molecule

decoding of combinatorially modified nucleosomes. *Science* **352**, 717–721 (2016).
48. V. Fedyuk, N. Erez, N. Furth, O. Beresh, E. Andreishcheva, A. Shinde, D. Jones, B. B. Zakai, Y. Mavor, T. Peretz, A. Hubert, J. E. Cohen, A. Salah, M. Temper, A. Grinshpun, M. Maoz, A. Zick, G. Ron, E. Shema, Multiplexed, single-molecule, epigenetic analysis of plasma-isolated nucleosomes for cancer diagnostics. *Nat. Biotechnol.* **41**, 212–221 (2023).
49. Rapid DNA sequencing based upon single molecule detection. *Genetic Analysis: Biomolecular Engineering* **8**, 1–7 (1991).
50. I. Braslavsky, B. Hebert, E. Kartalov, S. R. Quake, Sequence information can be obtained from single DNA molecules. *Proc Natl Acad Sci U S A* **100**, 3960–3964 (2003).
51. M. J. Levene, J. Korlach, S. W. Turner, M. Foquet, H. G. Craighead, W. W. Webb, Zero-mode waveguides for single-molecule analysis at high concentrations. *Science* **299**, 682–686 (2003).
52. W. J. Greenleaf, S. M. Block, Single-molecule, motion-based DNA sequencing using RNA polymerase. *Science* **313**, 801 (2006).
53. T. D. Harris, P. R. Buzby, H. Babcock, E. Beer, J. Bowers, I. Braslavsky, M. Causey, J. Colonell, J. Dimeo, J. W. Efcavitch, E. Giladi, J. Gill, J. Healy, M. Jarosz, D. Lapen, K. Moulton, S. R. Quake, K. Steinmann, E. Thayer, A. Tyurina, R. Ward, H. Weiss, Z. Xie, Single-molecule DNA sequencing of a viral genome. *Science* **320**, 106–109 (2008).
54. J. Eid, A. Fehr, J. Gray, K. Luong, J. Lyle, G. Otto, P. Peluso, D. Rank, P. Baybayan, B. Bettman, A. Bibillo, K. Bjornson, B. Chaudhuri, F. Christians, R. Cicero, S. Clark, R. Dalal, A. Dewinter, J. Dixon, M. Foquet, A. Gaertner, P. Hardenbol, C. Heiner, K. Hester, D. Holden, G. Kearns, X. Kong, R. Kuse, Y. Lacroix, S. Lin, P. Lundquist, C. Ma, P. Marks, M. Maxham, D. Murphy, I. Park, T. Pham, M. Phillips, J. Roy, R. Sebra, G. Shen, J. Sorenson, A. Tomaney, K. Travers, M. Trulson, J. Vieceli, J. Wegener, D. Wu, A. Yang, D. Zaccarin, P. Zhao, F. Zhong, J. Korlach, S. Turner, Real-time DNA sequencing from single polymerase molecules. *Science* **323**, 133–138 (2009).
55. J. Clarke, H.-C. Wu, L. Jayasinghe, A. Patel, S. Reid, H. Bayley, Continuous base identification for single-molecule nanopore DNA sequencing. *Nat Nanotechnol* **4**, 265–270 (2009).
56. T. Ha, T. Enderle, D. F. Ogletree, D. S. Chemla, P. R. Selvin, S. Weiss, Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 6264–6268 (1996).
57. X. Zhuang, L. E. Bartley, H. P. Babcock, R. Russell, T. Ha, D. Herschlag, S. Chu, A single-molecule study of RNA catalysis and folding. *Science* **288**, 2048–2051 (2000).
58. J. Liphardt, B. Onoa, S. B. Smith, I. Tinoco Jr, C. Bustamante, Reversible unfolding of single RNA molecules by mechanical force. *Science* **292**, 733–737 (2001).
59. G. Bonnet, O. Krichevsky, A. Libchaber, Kinetics of conformational fluctuations in DNA hairpin-loops. *Proc Natl Acad Sci U S A* **95**, 8602–8606 (1998).
60. M. D. Wang, M. J. Schnitzer, H. Yin, R. Landick, J. Gelles, S. M. Block, Force and velocity measured for single molecules of RNA polymerase. *Science* **282**, 902–907 (1998).
61. K. C. Neuman, E. A. Abbondanzieri, R. Landick, J. Gelles, S. M. Block, Ubiquitous transcriptional pausing is independent of RNA polymerase backtracking. *Cell* **115**, 437–447 (2003).
62. M. H. Larson, R. A. Mooney, J. M. Peters, T. Windgassen, D. Nayak, C. A. Gross, S. M. Block, W. J. Greenleaf, R. Landick, J. S. Weissman, A pause sequence enriched at translation start sites drives transcription dynamics in vivo. *Science* **344**, 1042–1047

(2014).
63. I. O. Vvedenskaya, H. Vahedian-Movahed, J. G. Bird, J. G. Knoblauch, S. R. Goldman, Y. Zhang, R. H. Ebright, B. E. Nickels, Interactions between RNA polymerase and the "core recognition element" counteract pausing. *Science* **344**, 1285–1289 (2014).
64. I. C. Nova, J. M. Craig, A. Mazumder, A. H. Laszlo, I. M. Derrington, M. T. Noakes, H. Brinkerhoff, S. Yang, H. Vahedian-Movahed, L. Li, Y. Zhang, J. L. Bowman, J. R. Huang, J. W. Mount, R. H. Ebright, J. H. Gundlach, Nanopore tweezers show fractional-nucleotide translocation in sequence-dependent pausing by RNA polymerase. *Proc Natl Acad Sci U S A* **121**, e2321017121 (2024).
65. J. M. Craig, A. H. Laszlo, H. Brinkerhoff, I. M. Derrington, M. T. Noakes, I. C. Nova, B. I. Tickman, K. Doering, N. F. de Leeuw, J. H. Gundlach, Revealing dynamics of helicase translocation on single-stranded DNA using high-resolution nanopore tweezers. *Proc Natl Acad Sci U S A* **114**, 11932–11937 (2017).
66. J. M. Craig, M. Mills, H. C. Kim, J. R. Huang, S. J. Abell, J. W. Mount, J. H. Gundlach, K. C. Neuman, A. H. Laszlo, Nanopore tweezers measurements of RecQ conformational changes reveal the energy landscape of helicase motion. *Nucleic Acids Res* **50**, 10601–10613 (2022).
67. A. H. Laszlo, J. M. Craig, M. Gavrilov, R. Tippana, I. C. Nova, J. R. Huang, H. C. Kim, S. J. Abell, M. deCampos-Stairiker, J. W. Mount, J. L. Bowman, K. S. Baker, H. Higinbotham, D. Bobrovnikov, T. Ha, J. H. Gundlach, Sequence-dependent mechanochemical coupling of helicase translocation and unwinding at single-nucleotide resolution. *Proc. Natl. Acad. Sci. U. S. A.* **119**, e2202489119 (2022).
68. M. MacKenzie, C. Argyropoulos, An introduction to nanopore sequencing: Past, present, and future considerations. *Micromachines (Basel)* **14**, 459 (2023).
69. M. Palla, S. Punthambaker, B. Stranges, F. Vigneault, J. Nivala, D. Wiegand, A. Ayer, T. Craig, D. Gremyachinskiy, H. Franklin, S. Sun, J. Pollard, A. Trans, C. Arnold, C. Schwab, C. Mcgaw, P. Sarvabhowman, D. Dalal, E. Thai, E. Amato, I. Lederman, M. Taing, S. Kelley, A. Qwan, C. W. Fuller, S. Roever, G. M. Church, Multiplex Single-Molecule Kinetics of Nanopore-Coupled Polymerases. *ACS Nano* **15**, 489–502 (2021).
70. S. Uemura, C. E. Aitken, J. Korlach, B. A. Flusberg, S. W. Turner, J. D. Puglisi, Real-time tRNA transit on single translating ribosomes at codon resolution. *Nature* **464**, 1012–1017 (2010).
71. J. Aguirre Rivera, G. Mao, A. Sabantsev, M. Panfilov, Q. Hou, M. Lindell, C. Chanez, F. Ritort, M. Jinek, S. Deindl, Massively parallel analysis of single-molecule dynamics on next-generation sequencing chips. *Science* **385**, 892–898 (2024).
72. I. Severins, C. Bastiaanssen, S. H. Kim, R. B. Simons, J. van Noort, C. Joo, Single-molecule structural and kinetic studies across sequence space. *Science* (2024).
73. Hazra J,...,Kapanidis A, Unraveling single-molecule reactions via multiplexed *in-situ* DNA sequencing.
74. Bentley, D. R., Balasubramanian, S., Swerdlow, H. P., Smith, G. P., Milton, J., et al., Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **456**, 53–59 (2008).
75. C. Bastiaanssen, I. Severins, J. van Noort, C. Joo, Single-molecule parallel analysis for rapid exploration of sequence space. *Nature Protocols*, 1–29 (2025).
76. M. Panfilov, G. Mao, J. Guo, J. Aguirre Rivera, A. Sabantsev, S. Deindl, Multiplexed single-molecule characterization at the library scale. *Nature Protocols*, 1–26 (2025).
77. S. M. Miick, R. S. Fee, D. P. Millar, W. J. Chazin, Crossover isomer bias is the primary sequence-dependent property of immobilized Holliday junctions. *Proc Natl Acad Sci U S A* **94**, 9080–9084 (1997).
78. D. M. Lilley, R. M. Clegg, The structure of branched DNA species. *Q Rev Biophys* **26**,

131–175 (1993).
79. M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, E. Charpentier, A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
80. A. V. Anzalone, L. W. Koblan, D. R. Liu, Genome editing with CRISPR-Cas nucleases, base editors, transposases and prime editors. *Nat Biotechnol* **38**, 824–844 (2020).
81. B. Vermeer, S. Schmid, Can DyeCycling break the photobleaching limit in single-molecule FRET? *Nano Res* **15**, 9818–9830 (2022).
82. M. Kümmerlin, A. Mazumder, A. N. Kapanidis, Bleaching-resistant, Near-continuous Single-molecule Fluorescence and FRET Based on Fluorogenic and Transient DNA Binding. *Chemphyschem* **24**, e202300175 (2023).
83. G. Wei, X. Ran, R. Ai-Abssi, Z. Yang, Finding the dark matter: Large language model-based enzyme kinetic data extractor and its validation. *Protein Sci* **34**, e70251 (2025).
84. A. Sabantsev, G. Mao, J. Aguirre Rivera, M. Panfilov, A. Arseniev, O. Ho, M. Khodorkovskiy, S. Deindl, Spatiotemporally controlled generation of NTPs for single-molecule studies. *Nat. Chem. Biol.* **18**, 1144–1151 (2022).
85. A. Korman, H. Sun, B. Hua, H. Yang, J. N. Capilato, R. Paul, S. Panja, T. Ha, M. M. Greenberg, S. A. Woodson, Light-controlled twister ribozyme with single-molecule detection resolves RNA function in time and space. *Proc Natl Acad Sci U S A* **117**, 12080–12086 (2020).
86. C. Brieke, F. Rohrbach, A. Gottschalk, G. Mayer, A. Heckel, Light-controlled tools. *Angew Chem Int Ed Engl* **51**, 8446–8476 (2012).
87. Y. Liu, R. S. Zou, S. He, Y. Nihongaki, X. Li, S. Razavi, B. Wu, T. Ha, Very fast CRISPR on demand. *Science* **368**, 1265–1269 (2020).
88. J. Yamaguchi, M. Naimuddin, M. Biyani, T. Sasaki, M. Machida, T. Kubo, T. Funatsu, Y. Husimi, N. Nemoto, cDNA display: a novel screening method for functional disulfide-rich peptides by solid-phase synthesis and stabilization of mRNA-protein fusions. *Nucleic Acids Res* **37**, e108 (2009).
89. J. Hanes, A. Plückthun, In vitro selection and evolution of functional proteins by using ribosome display. *Proc Natl Acad Sci U S A* **94**, 4937–4942 (1997).
90. L. Gu, C. Li, J. Aach, D. E. Hill, M. Vidal, G. M. Church, Multiplex single-molecule interaction profiling of DNA-barcoded proteins. *Nature* **515**, 554–557 (2014).
91. C. V. Agu, R. L. Cook, W. Martelly, L. R. Gushgari, M. Mohan, B. Takulapalli, Multiplexed proteomic biosensor platform for label-free real-time simultaneous kinetic screening of thousands of protein interactions. *Commun Biol* **8**, 468 (2025).
92. B. T. Porebski, M. Balmforth, G. Browne, A. Riley, K. Jamali, M. J. L. J. Fürst, M. Velic, A. Buchanan, R. Minter, T. Vaughan, P. Holliger, Rapid discovery of high-affinity antibodies via massively parallel sequencing, ribosome display and affinity screening. *Nat Biomed Eng* **8**, 214–232 (2024).
93. M. C. Needels, D. G. Jones, E. H. Tate, G. L. Heinkel, L. M. Kochersperger, W. J. Dower, R. W. Barrett, M. A. Gallop, Generation and screening of an oligonucleotide-encoded synthetic peptide library. *Proc Natl Acad Sci U S A* **90**, 10700–10704 (1993).
94. S. Melkko, J. Scheuermann, C. E. Dumelin, D. Neri, Encoded self-assembling chemical libraries. *Nat Biotechnol* **22**, 568–574 (2004).
95. S. Brenner, R. A. Lerner, Encoded combinatorial chemistry. *Proc Natl Acad Sci U S A* **89**, 5381–5383 (1992).
96. H. Hwang, H. Kim, S. Myong, Protein induced fluorescence enhancement as a single molecule assay with short distance sensitivity. *Proceedings of the National Academy of Sciences* **108**, 7414–7418 (2011).
97. E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino,

M. W. Davidson, J. Lippincott-Schwartz, H. F. Hess, Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **313**, 1642–1645 (2006).
98. M. J. Rust, M. Bates, X. Zhuang, Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat Methods* **3**, 793–795 (2006).
99. R. Jungmann, M. S. Avendaño, J. B. Woehrstein, M. Dai, W. M. Shih, P. Yin, Multiplexed 3D cellular super-resolution imaging with DNA-PAINT and Exchange-PAINT. *Nat Methods* **11**, 313–318 (2014).
100. S. B. Smith, L. Finzi, C. Bustamante, Direct mechanical measurements of the elasticity of single DNA molecules by using magnetic beads. *Science* **258**, 1122–1126 (1992).
101. T. R. Strick, J. F. Allemand, D. Bensimon, A. Bensimon, V. Croquette, The elasticity of a single supercoiled DNA molecule. *Science* **271**, 1835–1837 (1996).
102. J. P. Cnossen, D. Dulin, N. H. Dekker, An optimized software framework for real-time, high-throughput tracking of spherical beads. *Rev Sci Instrum* **85**, 103712 (2014).
103. A. Huhle, D. Klaue, H. Brutzer, P. Daldrop, S. Joo, O. Otto, U. F. Keyser, R. Seidel, Camera-based three-dimensional real-time particle tracking at kHz rates and Ångström accuracy. *Nature Communications* **6**, 1–8 (2015).
104. P. W. K. Rothemund, Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297–302 (2006).
105. S. M. Douglas, H. Dietz, T. Liedl, B. Högberg, F. Graf, W. M. Shih, Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature* **459**, 414–418 (2009).
106. D. Han, S. Pal, J. Nangreave, Z. Deng, Y. Liu, H. Yan, DNA origami with complex curvatures in three-dimensional space. *Science* **332**, 342–346 (2011).
107. N. C. Seeman, Nucleic acid junctions and lattices. *J Theor Biol* **99**, 237–247 (1982).
108. P. Kosuri, B. D. Altheimer, M. Dai, P. Yin, X. Zhuang, Rotation tracking of genome-processing enzymes using DNA origami rotors. *Nature* **572**, 136–140 (2019).
109. C. Larsson, I. Grundberg, O. Söderberg, M. Nilsson, In situ detection and genotyping of individual mRNA molecules. *Nat Methods* **7**, 395–397 (2010).
110. R. Ke, M. Mignardi, A. Pacureanu, J. Svedlund, J. Botling, C. Wählby, M. Nilsson, In situ sequencing for RNA analysis in preserved tissue and cells. *Nat Methods* **10**, 857–860 (2013).
111. J. H. Lee, E. R. Daugharthy, J. Scheiman, R. Kalhor, J. L. Yang, T. C. Ferrante, R. Terry, S. S. F. Jeanty, C. Li, R. Amamoto, D. T. Peters, B. M. Turczyk, A. H. Marblestone, S. A. Inverso, A. Bernard, P. Mali, X. Rios, J. Aach, G. M. Church, Highly multiplexed subcellular RNA sequencing in situ. *Science* **343**, 1360–1363 (2014).
112. C. Xia, J. Fan, G. Emanuel, J. Hao, X. Zhuang, Spatial transcriptome profiling by MERFISH reveals subcellular RNA compartmentalization and cell cycle-dependent gene expression. *Proc Natl Acad Sci U S A* **116**, 19490–19499 (2019).
113. M. J. Lawson, D. Camsund, J. Larsson, Ö. Baltekin, D. Fange, J. Elf, genotyping of a pooled strain library after characterizing complex phenotypes. *Mol Syst Biol* **13**, 947 (2017).
114. G. Emanuel, J. R. Moffitt, X. Zhuang, High-throughput, image-based screening of pooled genetic-variant libraries. *Nat Methods* **14**, 1159–1162 (2017).
115. A. Farrar, P. Turner, H. El Sayyed, C. Feehily, S. Chatzimichail, S. Ta, D. Crook, M. Andersson, S. Oakley, L. Barrett, C. Nellåker, N. Stoesser, A. Kapanidis, Ribosome phenotypes for rapid classification of antibiotic-susceptible and resistant strains of Escherichia coli. *Commun Biol* **8**, 319 (2025).
116. D. Feldman, A. Singh, J. L. Schmid-Burgk, R. J. Carlson, A. Mezger, A. J. Garrity, F. Zhang, P. C. Blainey, Optical Pooled Screens in Human Cells. *Cell* **179**, 787–799.e17 (2019).

117. A. Gahlmann, W. E. Moerner, Exploring bacterial cell biology with single-molecule tracking and super-resolution imaging. *Nat Rev Microbiol* **12**, 9–22 (2014).
118. J. Elf, I. Barkefors, Single-Molecule Kinetics in Living Cells. *Annu Rev Biochem* **88**, 635–659 (2019).
119. M. Lelek, M. T. Gyparaki, G. Beliu, F. Schueder, J. Griffié, S. Manley, R. Jungmann, M. Sauer, M. Lakadamyali, C. Zimmer, Single-molecule localization microscopy. *Nat Rev Methods Primers* **1** (2021).
120. N. K. Podh, S. Paliwal, P. Dey, A. Das, S. Morjaria, G. Mehta, In-vivo Single-Molecule Imaging in Yeast: Applications and Challenges. *J Mol Biol* **433**, 167250 (2021).
121. T. Ha, C. Kaiser, S. Myong, B. Wu, J. Xiao, Next generation single-molecule techniques: Imaging, labeling, and manipulation in vitro and in cellulo. *Mol Cell* **82**, 304–314 (2022).
122. N. Ribeck, O. A. Saleh, Multiplexed single-molecule measurements with magnetic tweezers. *Rev Sci Instrum* **79**, 094301 (2008).
123. D. Yang, A. Ward, K. Halvorsen, W. P. Wong, Multiplexed single-molecule force spectroscopy using a centrifuge. *Nat Commun* **7**, 11026 (2016).
124. R. Agarwal, K. E. Duderstadt, Multiplex flow magnetic tweezers reveal rare enzymatic events with single molecule precision. *Nat Commun* **11**, 4714 (2020).
125. D. Huynh, P. Hoffmeister, T. Friedrich, K. Zhang, M. Bartkuhn, F. Ferrante, B. D. Giaimo, R. A. Kovall, T. Borggrefe, F. Oswald, J. C. M. Gebhardt, Effective in vivo binding energy landscape illustrates kinetic stability of RBPJ-DNA binding. *Nat Commun* **16**, 1259 (2025).
126. H. Shi, N. Al-Sayyad, K. M. Wasko, M. I. Trinidad, E. E. Doherty, K. Vohra, R. S. Boger, D. Colognori, J. C. Cofsky, P. Skopintsev, Z. Bryant, J. A. Doudna, Rapid two-step target capture ensures efficient CRISPR-Cas9-guided genome editing. *Mol. Cell* **85**, 1730–1742.e9 (2025).
127. E. A. Abbondanzieri, W. J. Greenleaf, J. W. Shaevitz, R. Landick, S. M. Block, Direct observation of base-pair stepping by RNA polymerase. *Nature* **438**, 460–465 (2005).
128. F. Balzarotti, Y. Eilers, K. C. Gwosch, A. H. Gynnå, V. Westphal, F. D. Stefani, J. Elf, S. W. Hell, Nanometer resolution imaging and tracking of fluorescent molecules with minimal photon fluxes. *Science* **355**, 606–612 (2017).
129. F. Chen, P. W. Tillberg, E. S. Boyden, Optical imaging. Expansion microscopy. *Science* **347**, 543–548 (2015).
130. A. M. Femino, F. S. Fay, K. Fogarty, R. H. Singer, Visualization of single RNA transcripts in situ. *Science* **280**, 585–590 (1998).
131. A. Raj, P. van den Bogaard, S. A. Rifkin, A. van Oudenaarden, S. Tyagi, Imaging individual mRNA molecules using multiple singly labeled probes. *Nat Methods* **5**, 877–879 (2008).

**Acknowledgments:**

We thank Mikhail Panfilov and Anton Sabantsev for critical reading of the manuscript.

**Funding:** European Research Council (ERC) Advanced Grant ERC-ADG-101092623 (S.D.); Knut and Alice Wallenberg Foundation grants KAW/WAF 2019.0306 and KAW 2024.0012 (S.D.); Cancerfonden grant 25 4453 Pj (S.D.); Swedish Research Council project grant VR 03255 (S.D.); ERC Consolidator Grant 819299 (C.J.); Bio&Medical Technology Development Program of the National Research Foundation RS-2025-02217909 (C.J.) and the Basic Research Laboratory Program NRF-2023R1A2C2004745 (C.J.) by the Ministry of Science and ICT, Korea; Wellcome Trust grants 110164/Z/15/Z and 226662/Z/22/Z (A.N.K.); Leverhulme Trust award grant RPG-2024-037 (A.N.K., J.P.H.).

**Competing interests:** The work discusses single-molecule fluorescence measurements, which can be performed using commercial microscopes from Oxford Nanoimaging, a company in which A.N.K. is a co-founder and shareholder. J.P.H. and A.N.K. have financial interests in patent applications related to the SPIN-Seq methodology discussed in this manuscript.