Biochemical Society Transactions (2021) **49** 2711–2726
https://doi.org/10.1042/BST20210674

PORTLAND PRESS

Review Article

Using cryo-EM to uncover mechanisms of bacterial transcriptional regulation

David M. Wood[1], **Renwick C.J. Dobson**[1,2] and **Christopher R. Horne**[3,4]

[1]Biomolecular Interaction Centre and School of Biological Sciences, University of Canterbury, Christchurch, New Zealand; [2]Bio21 Molecular Science and Biotechnology Institute, Department of Biochemistry and Pharmacology, University of Melbourne, Parkville, VIC, Australia; [3]Walter and Eliza Hall Institute of Medical Research, 1G Royal Parade, Parkville, VIC 3052, Australia; [4]Department of Medical Biology, University of Melbourne, Parkville, VIC 3052, Australia

Correspondence: Christopher R. Horne (horne.c@wehi.edu.au)

Transcription is the principal control point for bacterial gene expression, and it enables a global cellular response to an intracellular or environmental trigger. Transcriptional regulation is orchestrated by transcription factors, which activate or repress transcription of target genes by modulating the activity of RNA polymerase. Dissecting the nature and precise choreography of these interactions is essential for developing a molecular understanding of transcriptional regulation. While the contribution of X-ray crystallography has been invaluable, the 'resolution revolution' of cryo-electron microscopy has transformed our structural investigations, enabling large, dynamic and often transient transcription complexes to be resolved that in many cases had resisted crystallisation. In this review, we highlight the impact cryo-electron microscopy has had in gaining a deeper understanding of transcriptional regulation in bacteria. We also provide readers working within the field with an overview of the recent innovations available for cryo-electron microscopy sample preparation and image reconstruction of transcription complexes.

Introduction

The expression of a gene to assemble a protein (the Central Dogma) is a fundamental process of life. Transcription is the first step in gene expression, which is coordinated by a complex of proteins that cooperate at the promoter region to transcribe the DNA gene sequence into mRNA. The lead actor of transcription, RNA Polymerase (RNAP), comprises a core subunit of proteins labelled $\alpha_2\beta\beta'\omega$ (Figure 1A) [1]. Upon interaction with a sigma (σ)-factor, RNAP forms the active holoenzyme [2,3]. Sigma factors are large, multi-domain proteins that bind various sites across the core RNAP and promoter DNA [4], and are responsible for guiding RNAP to the transcription start sites by locating the −35 element (consensus sequence TTGACA) and the −10 element (consensus sequence TATAAT) within the promoter [5] (Figure 1A). These two DNA elements constitute major features of a bacterial promoter and serve as notable controllers of transcriptional activity. Bacterial promoters are also decorated with other regulatory components, such as −10 extension (EXT) [6], discriminator (DISC) [7,8], and the upstream element (UP element) [9]. Once RNAP holoenzyme binds DNA, a series of conformational changes serve to manipulate the DNA and unwind a 13 base pair region (promoter melting) to initiate transcription.

In bacteria, the regulation of gene expression occurs primarily at transcription initiation, allowing bacteria to maintain homeostasis and adapt to changing environmental conditions, such as nutrient availability, by changing which genes are expressed, and which are not [excellent reviews on gene regulation focused on initiation can be found here [10,11]]. Transcriptional regulation in bacteria is predominantly modulated by transcription factors, an important class of *trans*-acting factor, that bind DNA and function as either activators or repressors of gene expression (Figure 1B). Whereas

Received: 1 October 2021
Revised: 10 November 2021
Accepted: 15 November 2021

Version of Record published:
2 December 2021

PORTLAND PRESS

Biochemical Society Transactions (2021) **49** 2711–2726
https://doi.org/10.1042/BST20210674

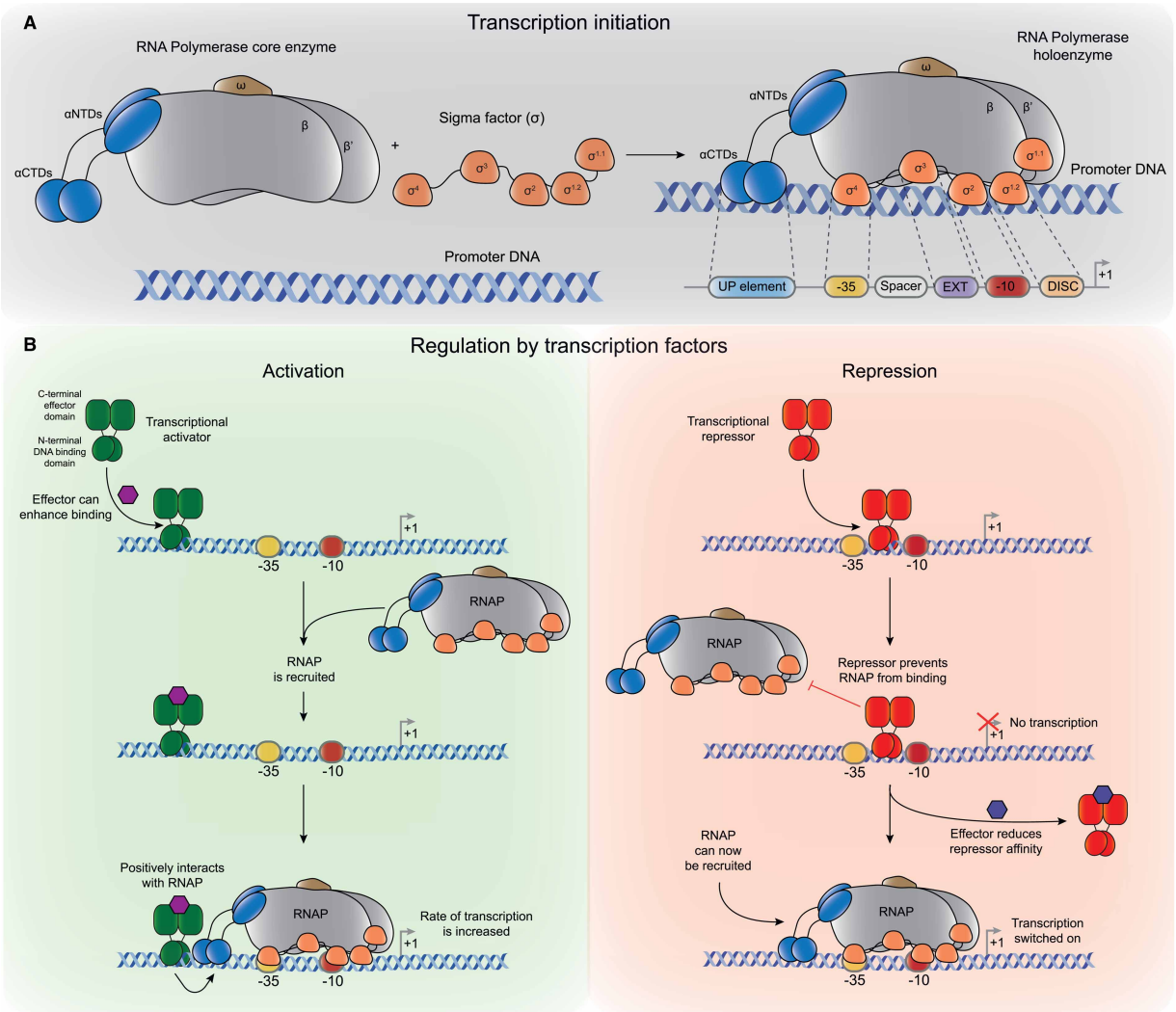

Figure 1. Schematic model of bacterial transcription initiation and regulation by transcription factors.

(**A**) In bacteria, transcription is initiated by the formation of the RNA polymerase (RNAP) holoenzyme, which comprises the RNAP core and a sigma (σ) factor (e.g. σ^{70}). Upon interaction with promoter DNA, the RNAP holoenzyme forms a closed transcription complex. The subunits of the RNAP core, σ factor and sequence-specific interactions with the promoter DNA are illustrated. RNAP core: α N-terminal/C-terminal domains (αNTD/CTD) — blue; β and β′ — grey; ω — black. σ factor — orange, linker regions (black lines). Promoter DNA (grey): UP element — cyan; –35 element — yellow; spacer — white; extended –10 (EXT) — violet; –10 element — dark red; discriminator (DISC) — peach; transcription start site (+1, arrow). Sequence-specific interactions are highlighted with dash lines. (**B**) Transcription factors, known as activators or repressors can activate or repress transcription initiation, respectively. The general mechanism of each is shown here. Transcriptional activators (left panel, green) bind to a site upstream of the promoter (–35 and –10 element) and transcription start site (+1, arrow) where they can recruit the RNAP holoenzyme by interacting with the C-terminal of the α-subunit (αCTD) of RNAP. This process can be enhanced by small molecules or effectors (purple hexagon) to increase the rate of transcription. In contrast, transcriptional repressors (right panel, red) bind to a site that overlaps the core –35 and –10 elements of the promoter to directly block the binding of RNAP to the promoter, switching gene transcription off (shown by red cross). In the presence of an effector (blue hexagon) the DNA affinity is reduced, and the repressor dissociates from the promoter. This allows RNAP to be recruited, switching gene transcription on.

transcriptional activators generally bind upstream of the RNAP-binding site to co-opt RNAP and enhance activity, transcriptional repressors bind to the operator region of target genes to directly obstruct the binding and activity of the RNAP [10,12].

Transcription factors largely share a common domain architecture, comprising an N-terminal DNA-binding domain and a C-terminal effector-binding domain, typically connected by a flexible linker (Figure 1B). The

DNA-binding domain recognises a specific DNA sequence and most often contains the highly conserved, helix-turn-helix motif, while the effector-binding domain functions as a signal sensor [13]. In general, the effector is a small molecule or pathway metabolite that allosterically binds the protein to trigger a conformational change that alters the affinity of the DNA-binding domain to the target sequence [12,13]. This collective function enables transcription factors to act as molecular switches, enabling bacteria to rapidly respond to sudden environmental challenges [14].

In addition to transcription factors that directly modulate RNAP activity, σ-factors also serve as an important class of *trans*-acting factor. By acting in complex with RNAP, σ-factors facilitate broader changes, such as the expression of genes required for bacterial cell viability [10,12] (reviewed here [10,15,16]). Overall, RNAP serves as the core element of gene regulation, combining information from an assortment of sensory systems to appropriately modulate gene expression — the net outcome is to determine which genes are transcribed, and to what extent, under any specific growth condition [17].

Over the past few decades, structural biology has been instrumental for defining these regulatory mechanisms and the biological function of RNAP during transcription [18]. Yet, despite a focused effort to delineate these processes, principally through X-ray crystallography, technical limitations have restricted our molecular understanding of transcriptional regulation. Notably, these include the intrinsic flexibility, conformational/compositional heterogeneity, and transient nature of the transcription complexes [19,20]. Crystal growth is usually easiest when the macromolecules are stably folded, can be concentrated, are homogenous (i.e. purified away from all other contaminants and in a stable oligomeric state), and have limited flexibility. Fuelled by technological breakthroughs in data collection and imaging processing (reviewed in [21]), cryo-electron microscopy (cryo-EM) now offers a powerful alternative to overcome these challenges. This capacity has granted researchers the power to 'see' transient or heterogeneous complexes that are unamenable to crystallisation and determine their structure at or near atomic-level resolution.

Herein, we review the contribution of cryo-EM to further our understanding of transcriptional regulation in bacteria, with a focus on studies that have provided key mechanistic insights into transcription initiation. We also highlight state-of-the-art sample preparation and 3D reconstruction strategies for structure determination with a particular focus on 'tricks' for protein–nucleic acid complexes.

Recent cryo-EM structures advance our understanding of bacterial transcriptional regulation

Visualising transcription complexes at the atomic level is essential for unravelling their mechanism of function. Over the past few years, cryo-EM has been indispensable for resolving large and heterogeneous complexes, where previous crystallographic studies have come up short, providing over 65 structures to date (summarised in Table 1). Here, we highlight pivotal transcription complexes active during transcription initiation, which contain transcriptional activators and repressors that until the advent of cryo-EM weren't fully understood.

Transcriptional activators

Activators serve to increase transcription by binding at, or upstream, of a promoter region, where they can positively interact with and recruit RNAP to initiate transcription of target genes (Figure 1B, left panel). This process can be achieved by the activator distorting promoter DNA to facilitate RNAP binding, or by directly tethering RNAP to the promoter region. To illustrate how each regulatory mechanism enhances RNAP binding, we outline two recent cryo-EM structures, respectively below.

The cryo-EM structure of MerR family regulator EcmrR provides our first example of DNA distortion (Figure 2A). Promoters that bind this family contain an additional 2 to 3 base pairs (or non-canonical space) between the −35 and −10 elements, which prevents optimal promoter recognition by RNAP and transcription initiation [48,49]. In contrast, a canonical promoter contains a 17 base pair spacer region between the −35 and −10 elements. MerR regulators bind and twist the non-optimal DNA spacer, such that the DNA promoter elements are readily recognisable by the RNAP holoenzyme (Figure 2A, inset). While previous crystal structures of MerR regulators in the absence of RNAP reported this DNA distortion [50,51], multiple cryo-EM structures of EcmrR in complex with RNAP provided a deeper molecular understanding of this promoter remodelling [22]. Similarly, cryo-EM was also used to dissect the DNA distortion mechanism from another MerR family regulator, BmrR [23].

Biochemical Society Transactions (2021) **49** 2711–2726
https://doi.org/10.1042/BST20210674

Table 1 Summary of bacterial transcription complex structures resolved by cryo-EM Part 1 of 2

Transcription complex	Family	Organism	EMDB ID	PDB ID	Reference
EcmrR-RP^2_o	MerR	*Escherichia coli*	EMD-22234	6XL5	[22]
EcmrR-RP^2_o (EcmrR-spacer DNA complex)			EMD-22235	6XL6	
EcmrR-RP^3_{int} with 3 nt RNA transcript			EMD-22236	6XL9	
EcmrR-RP^3_{int} with 3 nt RNA transcript (EcmrR-spacer DNA complex)			EMD-22237	6XLA	
EcmrR-RP^3_{int} with 4 nt RNA transcript			EMD-22245	6XLJ	
EcmrR-RP^3_{int} with 4 nt RNA transcript (EcmrR-spacer DNA complex)			EMD-22246	6XLK	
EcmrR-RP^2_o (clearer σ^{70} density)			EMD-23291	-	
BmrR-RNA polymerase complex	MerR	*Bacillus subtilis*	EMD-30390	7CKQ	[23]
CueR-RNA polymerase complex (without RNA transcript)	MerR	*Escherichia coli*	EMD-22184	6XH7 6XH8	[24]
CueR-RNA polymerase complex (with RNA transcript)			EMD-22185		
CueR-RNA polymerase complex (clearer σ^{70} density)			EMD-22289		
CueR- RNA polymerase complex	MerR	*Escherichia coli*	EMD-30268	6LDI	[25]
CueR- RNA polymerase complex (with fully duplex DNA)			EMD-0874	7C17	
NanR-$dimer_1$/DNA complex	GntR	*Escherichia coli*	EMD-21652	6WFQ	[26]
NanR-$dimer_3$/DNA complex			EMD-21661	6WG7	
BusR-$tetramer_1$/pAB DNA complex	GntR	*Streptococcus agalactiae*	EMD-13119	7OZ3	[27]
BusR-$tetramer_1$/pAB_1 DNA complex			EMD-12051	7B5Y	
TraR-Eσ^{70} (state I)	LuxR	*Escherichia coli*	EMD-0348	6N57	[28]
TraR-Eσ^{70} (state II)			EMD-0349	6N58	
TraR-Eσ^{70} (state III)			EMD-20231	N/A	
MmfR-$dimer_2$/DNA complex	TetR	*Streptomyces coelicolor*	EMD-20781	N/A	[29]
Class-II CAP-TAC^1 without RNA transcript (state I)	CAP	*Escherichia coli*	EMD-20287	6PB5	[30]
Class-II CAP-TAC^1 without RNA transcript (state II)			EMD-20288	6PB6	
Class-II CAP-TAC^1 with RNA transcript (state II)			EMD-20286	6PB4	
Class-I CAP-TAC^1	CAP	*Escherichia coli*	EMD-7059	6B6F	[31]
Class-I CAP-TAC^1 (focused map on αCTD-CAP region)			EMD-7060		
Crl-Eσ^S-RNA polymerase complex	Crl	*Escherichia coli*	EMD-200090	6OMF	[32]
Spx-RNA polymerase complex	Spx	*Bacillus subtilis*	EMD-31485	7F75	[33]
WhiB7-RP^2_o	WhiB	*Mycobacterium tuberculosis*	EMD-22886	7KIF	[34]
WhiB7-RP^4_c			EMD-22887	7KIM	
Rgg2-short hydrophobic peptide complex	Rgg	*Streptococcus thermophilus*	EMD-22341	7JI0	[35]
GreB-RNA polymerase elongation complex (pre-RNA cleavage)	Gre	*Escherichia coli*	EMD-4892	6RIN	[36]
GreB-RNA polymerase elongation complex (post-RNA cleavage)			EMD-4885	6RI7	
GreB-RNA polymerase reactivated complex (before RNA extension)			EMD-4882	6RH3	
CarD-RP^2_o	CarD	*Mycobacterium tuberculosis*	EMD-9037	6EDT	[37]
CarD-RNA polymerase intermediate (with 8-nt RNA transcript)			EMD-9039	6EE8	

Continued

Table 1 Summary of bacterial transcription complex structures resolved by cryo-EM Part 2 of 2

Transcription complex	Family	Organism	EMDB ID	PDB ID	Reference
CarD-RP_o^2 (with corallopyronin A)			EMD-9041	6EEC	
CarD-RNA polymerase holoenzyme (with corallopyronin A)			EMD-9047	6M7J	
CarD-RP_o^2 (with Sorangicin A)	CarD	*Mycobacterium tuberculosis*	EMD-21407	6VVY	[38]
CarD-$^{S456L}RP_o^2$ (with Sorangicin A)			EMD-21408	6VW0	
CarD-RP_o^2 (with Sorangicin A) (with 8-nt RNA transcript)			EMD-21406	6VVX	
CarD-$^{S456L}RP_o^2$ (with Sorangicin A) (with 8-nt RNA transcript)			EMD-21409	6VVZ	
SspA-σ^{70}-RP_o^2	GST	*Escherichia coli*	EMD-30307	7C97	[39]
DksA-RP_o^2 (State I) with guanosine tetraphosphate (ppGpp)	DksA	*Escherichia coli*	EMD-21881	7KHI	[40]
DksA-RP_o^2 (State II) with guanosine tetraphosphate (ppGpp)			EMD-21883	7KHE	
NusG-*ops*EC	NusG	*Escherichia coli*	EMD-7351	6C6U	[41]
RfaH-NusG-N-Term-*ops*EC	RfaH		EMD-7350	6C6T	
RfaH-full-length-*ops*EC	RfaH		EMD-7349	6C6S	
RNAP-HelD	HelD	*Bacillus subtillus*	EMD-21921	6WVK	[42]
Msm HelD–RNAP complex State I	HelD	*Mycobacterium smegmatis*	EMD-10996	6YXU	[43]
Msm HelD–RNAP complex State II			EMD-11004	6YYS	
Msm HelD–RNAP complex State III			EMD- 11026	6Z11	
Spt4/5-RNAP complex (with antibodies)	Spt4/5	*Pyrococcus furiosus*	EMD- 1840	N/A	[44]
Mfd-dependent transcription termination complex	MfD	*Thermus thermophilus*	EMD- 30117	6M6A	[45]
Mfd-dependent transcription termination complex with ATPγS	MfD		EMD- 30118	6M6B	
Mfd-bound RNA polymerase elongation complex — L1 state (with ATP)	MfD	*Escherichia coli*	EMD-21996	6X26	[46]
Mfd-bound RNA polymerase elongation complex — L2 state (with ADP)			EMD-22006	6X2F	
Mfd-bound RNA polymerase elongation complex — I state			EMD-22012	6X2N	
Mfd-bound RNA polymerase elongation complex — II state			EMD-22039	6X43	
Mfd-bound RNA polymerase elongation complex — III state			EMD-22043	6X4W	
Mfd-bound RNA polymerase elongation complex — IV state			EMD-22044	6X4Y	
Mfd-bound RNA polymerase elongation complex — V state			EMD-22045	6X50	
σ^{70}-RP_o^2	σ-factor	*Klebsiella pneumoniae*	EMD-0001	6GH5	[47]
σ^{70}-RNA polymerase (intermediate partially loaded) complex			EMD-0002	6GH6	
σ^{70}-RNA polymerase (initially transcribing) complex			EMD-4397	6GFW	

[1]CAP-TAC, cAMP receptor protein-dependent transcription activation complex.
[2]RP_o, RNA polymerase-promoter open complex.
[3]RP_{int}, RNA polymerase-promoter initial transcribing complex.
[3]RP_c, RNA polymerase-promoter closed complex.
N/A, Not available.

PORTLAND PRESS

Biochemical Society Transactions (2021) **49** 2711–2726
https://doi.org/10.1042/BST20210674

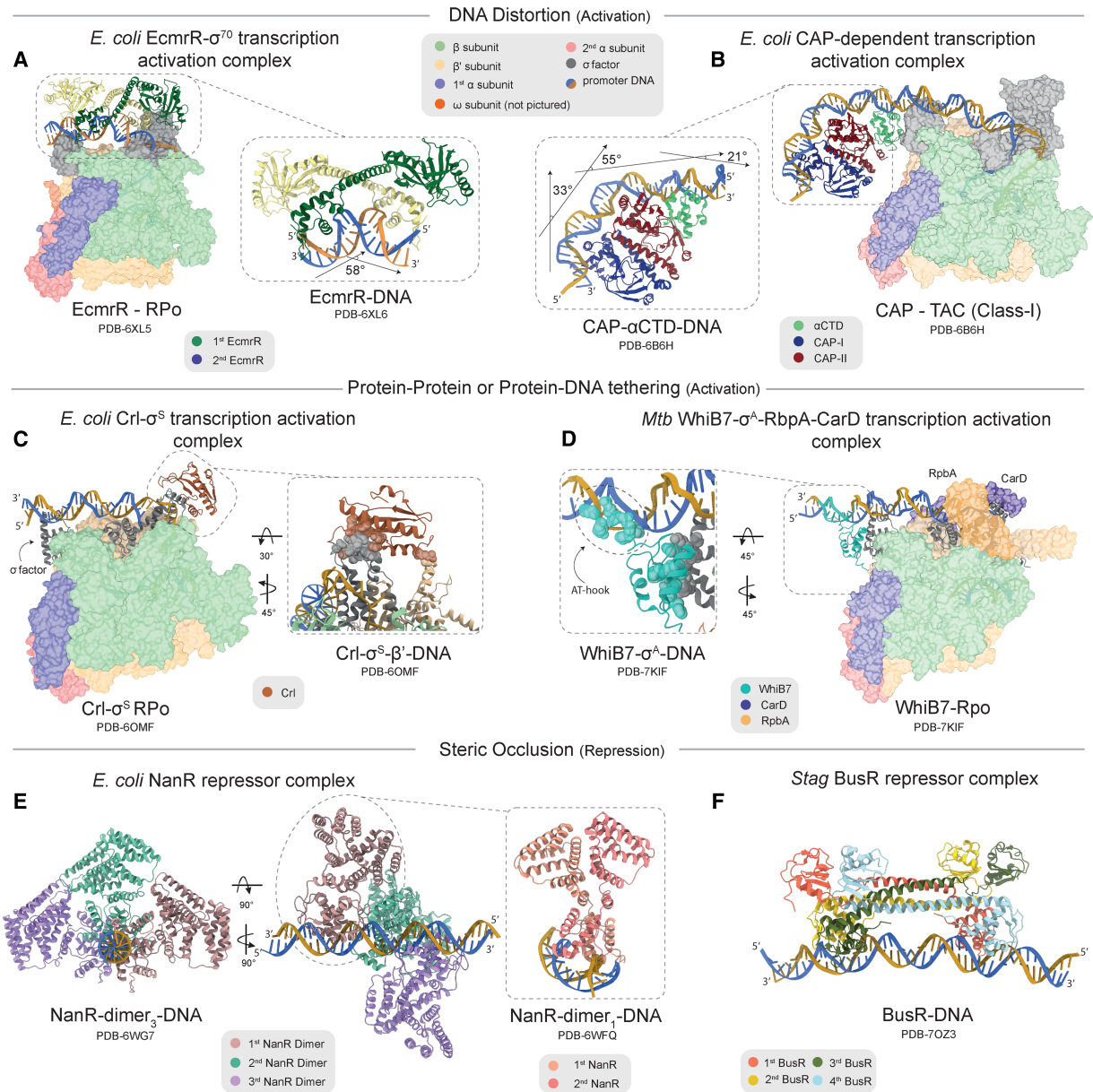

Figure 2. Recent bacterial transcription complexes solved by cryo-EM. Part 1 of 2

Examples (**A**,**B**) of transcriptional activator complexes that distort DNA. (**A**) Dimeric EcmrR in complex with promoter DNA (cartoon inset, PDB-6XL6), remodels (58° kink) the promoter DNA to create the optimal promoter architecture for *E. coli* holoenzyme [σ^{70} (grey surface) RNAP ($\alpha_2\beta\beta'\omega$ surface colour shown in shaded box)] to form the EcmrR-RPo (PDB-6XL5) [22]. (**B**) Overview of the *E. coli* class-1 CAP-TAC (PDB-6B6H). The cyclic adenosine 3′,5′-monophosphate receptor (CAP) protein dimer (dark blue and dark red, cartoon inset), binds its cognate DNA and αCTD of RNAP (green) to introduce three DNA kinks (33°, 55°, 21°). This results in a full 92° turn, optimally orienting the promoter DNA for σ^{70}-RNAP to bind [31]. Examples of transcription factors that stabilise aspects of the transcription complex are shown in **C** and **D**. (**C**) Crl (dark orange cartoon) binds residues of the β′clamp (beige cartoon) on *E. coli* RNAP and alternative σ factor σ^S (grey cartoon) through a distinct interface (shown by spheres) in the left inset. This tethering action creates a Crl-σ^S-RNAP complex that binds alternate promoter DNA to form the Crl-σ^S-RPo complex (PDB-6OMF) shown to the right of the inset [32]. (**D**) WhiB7 (cyan, cartoon) is a transcriptional activator in *Mycobacterium tuberculosis* (*Mtb*) that binds an AT-rich 'hook' sequence of DNA (shown by arrow) and σ^A (cartoon inset). By binding the active *Mtb* RNAP holoenzyme (surfaces coloured as *E. coli* RNAP), it creates the WhiB7-RPo (PDB-7KIF) [34]. Examples of steric occlusion in transcriptional repressors are shown in **E** and **F**. (**E**) Three *E. coli* NanR dimers binds three GGTATA repeats to form a NanR-dimer$_3$/DNA complex. Their close proximity allows intramolecular protein–protein interactions to stabilise the multimeric assembly (PDB-6WG7). The 70.5 kDa cryo-EM structure of dimeric NanR in complex with

Figure 2. Recent bacterial transcription complexes solved by cryo-EM. Part 2 of 2

cognate DNA (PBD-6WFQ) [26]. (**F**) *Streptococcus agalactiae* (*Stag*) BusR binds palindromic promoter DNA as a tetramer to repress transcription (PDB-7OZ3) [27]. The 5′ and 3′ DNA strands have been annotated throughout.

Our second example of DNA distortion is illustrated by the role of the global transcription factor, cyclic AMP (cAMP) receptor protein (CAP) to promote transcription. Two major classes of CAP exist, each of which can activate and initiate transcription by bending DNA to optimise RNAP recruitment [52]. These classes are differentiated by their promoter site and interaction mode with RNAP; class-I CAPs bind a −61 site and interact predominantly with αCTD subunit of RNAP (Figure 1A), while class-II CAPs bind at a −41 site and interact with multiple RNAP subunits. As the CAP–RNAP interactions are small, and the full CAP–RNAP–DNA complex is dynamic, any CAP-induced conformational changes in the presence of RNAP are difficult to capture by the freeze-frame feature of crystallography [52,53]. Recently, an intact transcription activation complex, containing a class-I CAP along with RNAP was resolved by cryo-EM (Figure 2B). This structure revealed extensive remodelling of the promoter DNA (∼90° kink) induced by CAP-binding (Figure 2B, inset) to wrap upstream DNA and co-opt RNAP via αCTD binding [31]. An analogous study reported the cryo-EM structure of the class-II CAP–RNAP complex [30]. Taken together, these structures shed light into how class-I and -II CAP activation complexes assemble to activate transcription.

Alternatively, an unconventional mode of activation involving protein tethering can stabilise and activate RNAP via a DNA-independent or -dependent process. DNA-independent modes involve activators that stabilise σ-factor and RNAP association solely through protein–protein interactions. This unusual mode of activation was first hypothesised using crystallography, but many interactions were absent due to crystal packing [54,55]. However, numerous cryo-EM structures of transcriptional activators, Crl [32], RbpA/CarD [37,38], Spx [33] and SspA [39], can now detail the precise interactions and conformational changes between the σ-factor that facilitate the formation of the RNAP holoenzyme. To highlight this protein–protein tethering mode, we present the structure of Crl bound to σ^S and a small domain of the β′ subunit to stabilise the holoenzyme (Figure 2C). Conversely, structural insight into DNA-dependent tethering was revealed through cryo-EM structures of WhiB7, which play a role in antibiotic resistance in mycobacteria. In addition to forming protein–protein contacts with the σ-factor, WhiB7 was observed to interact with promoter DNA via an AT-hook motif [34] (Figure 2D). This was unexpected as AT-hooks are rare in bacteria, yet common in eukaryotes [56]. Thus, these structures expand our understanding of how WhiB7 serves to regulate antibiotic resistance in mycobacteria, but also unearths a novel mode of transcriptional regulation in bacteria.

Transcriptional repressors

Repressors function to sterically occlude RNAP binding to DNA by occupying a site that overlaps the -35 and -10 promoter elements to prevent σ-factor recognition, switching gene transcription off [11] (Figure 1B, right panel). We illustrate this mode of regulation below.

Recently, it was shown that *Escherichia coli* NanR, which regulates bacterial sialic acid metabolism [57–60], cooperatively binds a three-repeat sequence overlapping the −10 element [26]. Through cryo-EM, three NanR dimers were observed to assemble in close proximity across the promoter, where intramolecular protein–protein interactions stabilise the repressor complex (Figure 2E). This multimeric assembly is unique among reported GntR-type regulators [26]. The lower-order NanR-$dimer_1$/DNA complex (70.5 kDa) was also resolved within the study at near atomic resolution, demonstrating the power of cryo-EM in this so-called 'resolution revolution' (Figure 2E, inset). Similarly, cooperative binding was also observed for the TetR-type regulator, MmfR, where two dimers bound DNA at an obtuse angle of 140° in the cryo-EM structure [29].

BusR is a transcriptional repressor that binds the c-di-AMP molecule; a vital molecule in normal cellular growth conditions and a target for antibiotic development [61]. A recent cryo-EM structure of BusR revealed how it binds bipartite DNA motifs (Figure 2F) as a tetramer and proposes a new regulator family, as the protein architecture is unlike any other transcriptional regulator described [27].

Briefly, we would be remiss not to mention how cryo-EM has had a marked impact on functionally understanding the transcription factor TraR, which functions both as an activator and repressor. Using cryo-EM, Chen et al. [28] resolved a series of structures that, alongside *in vitro* experiments, helped elucidate the transition of active RNAP from RPc (closed complex) to RPo (open complex) in the presence of TraR. The

Biochemical Society Transactions (2021) **49** 2711–2726
https://doi.org/10.1042/BST20210674

deconvolution of heterogeneous intermediate conformations allowed researchers to propose a mechanism, and use biochemical techniques to validate it [62]. The understanding of this conformational landscape by cryo-EM is a pivotal discovery within the field, as it has structurally illuminated the complicated and multifaceted mechanism of transcription initiation.

Contemporary cryo-EM strategies for resolving protein–DNA complexes

Having highlighted how cryo-EM has transformed our understanding of transcriptional regulation in bacteria, we now outline contemporary cryo-EM strategies for determining the structure of these dynamic macromolecular assemblies, particularly protein–DNA complexes. This overview will include the recent and ongoing developments in sample preparation and structure determination (summarised in Figure 3). Further discussion on data acquisition, image processing, and refinement are beyond the scope of this review (but are reviewed in [63,64]).

Sample preparation

Sample preparation for cryo-EM is paramount. In an ideal case, upon vitrification, the sample would be free of contamination, randomly oriented and evenly distributed in a monolayer of thin ice [74,75]. In reality, two main elements must be optimised to achieve this outcome — sample preparation and grid preparation. Sample preparation involves the purification of the protein or macromolecular complex in an intact and stable manner [74,76]. Given this process isolates the sample from its cellular environment, the buffering conditions must be optimised (e.g. salt, pH) to emulate native conditions. This can be achieved systematically and in high throughput using thermal stability assays [77], such as ProteoPlex [78], which is based on differential scanning

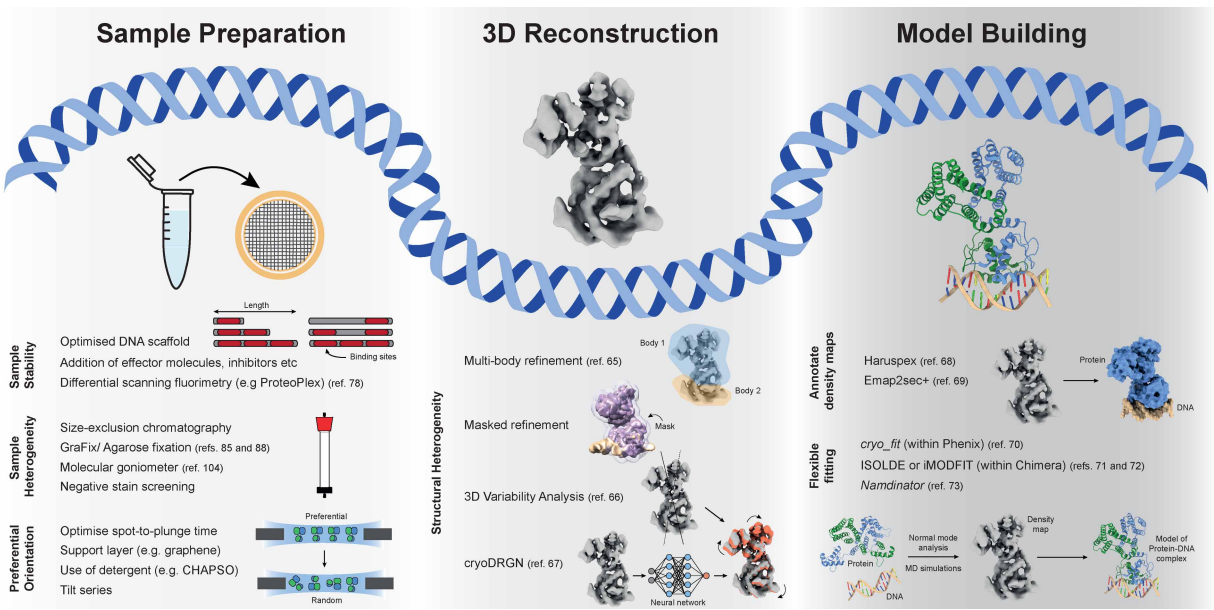

Figure 3. Contemporary cryo-EM strategies for solving protein–DNA complexes.

Recent innovations in sample preparation (left panel), 3D reconstruction (middle panel) and model building (right panel) are illustrated. During sample preparation (left panel) there are three facets that must be optimised: sample stability (e.g. optimising the length and number of binding sites within the DNA scaffold); sample homogeneity (using size exclusion chromatography); and preferential orientation (optimising spot-to-plunge time). To combat structural heterogeneity during 3D reconstruction (middle panel) and explore dynamics, multi-body refinement [65], masked refinement, 3D variability analysis [66] or cryoDRGN, using a neural network [67] can be employed. Density map is shown in grey, with hypothetical flexibility shown in orange. To annotate density maps during model building (right panel), Haruspex [68] or Emap2sec+ [69] can be utilised. As transcription complexes are dynamic assemblies, flexible fitting tools (*cryo_fit* [70], ISOLDE [71], iMODFIT [72] or *Namdinator* [73]) are required to fit PDB components (e.g. protein and DNA) into a cryo-EM density map (shown in grey) to generate a model (shown on the bottom right of panel). Cryo-EM data, maps and models are from ref. 63 (EMD-21652, PDB-6WFQ).

fluorimetry. Common additives such as glycerol, which mimic a crowded environment for stability is largely avoided within the cryo-EM community, because it significantly decreases contrast during data collection [76,79,80]. However, recent evidence suggests ≤20% glycerol can improve the stability of large complexes that are prone to disassembly, without compromising data quality and therefore should not be fully discounted as an additive in cryo-EM [81]. Enhanced stability can also be afforded by the addition of small effector molecules (particularly for transcriptional activators), co-factors or inhibitors that may stabilise or biochemically arrest the macromolecule in a unique functional state [22,82].

Knowledge of the specific DNA sequence that your protein–DNA system binds is critical. This includes the kinetics of your system (K_D, k_{on}, k_{off}) and stoichiometry of the protein–DNA interaction, which can vary by the length or number of binding sites within the DNA scaffold and by protein concentration. Thus, these elements must be carefully evaluated both biochemically and biophysically to inform the optimal DNA scaffold and concentrations you should use for grid preparation.

To perform their function, protein–DNA complexes undergo dynamic conformational rearrangements across one or more subunits [18,83]. While this conformational heterogeneity can typically be tackled during image processing through independent 3D classification and masked refinement (discussed below), sources of compositional heterogeneity must be mitigated. These sources include variations in the stoichiometry of the interaction partners, partially assembled complexes or the presence of assembly intermediates [75,76,80]. At the biochemical level, these phenomena can be addressed through sample preparation procedures, mostly commonly via size exclusion chromatography to remove aggregates or unbound components and isolate a target complex [22,23,26,29]. If complexes are inherently fragile, researchers have successfully employed a fractionation technique, named GraFix, to prepare homogeneous cryo-EM samples [39,84–86]. Here, using centrifugation, a density gradient (e.g. glycerol) is combined with weak chemical fixation (e.g. glutaraldehyde), which leads to the formation of monodisperse and chemically stabilised complexes [85,87]. The use of a weak fixation reagent is advantageous as it largely favours the formation of intramolecular crosslinks, which can prevent complex dissociation. To avoid reduced contrast, the density solution is removed by buffer exchange (Zeba) spin columns [85]. If the sample is scarce or unstable following removal of the density solution, agarose fixation offers an alternative strategy [88].

Following grid preparation [reviewed in [74,89]], which is largely dependent on user expertise and experience, sample stability, heterogeneity and particle distribution can be assessed by iterative negative stain experiments or under cryogenic temperatures [74,75,79]. When vitrified, protein–DNA complexes often preferentially adhere to the air–water interface or grid support [90]. This preferred particle orientation can lead to under sampling of some structural features, sample denaturation and anisotropic resolution in the density map [91,92]. Experimentally, this can be addressed by reducing the time interval between sample application to the grid and vitrification (spot-to-plunge time) [90,93,94], the use of support layer (e.g. graphene) to sequester the complex from the air–water interface [92,95,96] or an affinity support, such as streptavidin to immobilise biotinylated molecules [97], and by data collection at a fixed tilt-angle [98]. When acquiring data at a single tilt, gold foil grids can be used to minimise beam-induced movement during imaging [99,100]. Notably, bacterial RNAP transcription complexes suffer from severe preferential orientation [101], however, this is routinely combated by the addition of the zwitterionic detergent CHAPSO during sample preparation [22,23,39,40]. Other detergents, such as β-octyl glucoside have also been used to promote random particle distribution and discourage complex dissociation [27,102,103]. Recently, molecular goniometers have been constructed using DNA origami, which enable the DNA-binding protein to bind and be precisely oriented via a sequence-specific DNA stage [104]. As proof-of-concept, this nanoscale technology was utilised to resolve the 82 kDa DNA-binding protein, BurrH [104]. Moving forward, this concept can be adaptable to other small (<100 kDa) or asymmetric protein–DNA complexes.

Structure determination

3D reconstruction and model building represent the final hurdle towards determining a structure within the cryo-EM workflow. As previously discussed, protein–DNA complexes are driven by functionally relevant conformational and compositional changes as part of their dynamic modes of action [18,83]. While this intrinsic feature poses challenges for structure determination by cryo-EM, recent innovations make it possible to study these dynamic assemblies and gain unique insights into their molecular mechanism. In this section, we highlight these key innovations, which include algorithms to visualise molecular motions and identify interacting components, along with tools for flexible fitting in cryo-EM maps by molecular dynamics.

Biochemical Society Transactions (2021) **49** 2711–2726
https://doi.org/10.1042/BST20210674

During data acquisition, millions of snapshots across a conformational landscape are captured for the molecule of interest. This structural heterogeneity has typically been approached using various '3D classification' or 'heterogeneous refinement' tools, implemented in cryo-EM software packages such as RELION [105], cryoSPARC [106] or cisTEM [107]. These tools effectively divide the data into a small number of independent and discrete states, each of which are assumed to be structurally homogeneous. However, in scenarios where macromolecular complexes exhibit continuous conformational transitions of single domains or motions across multiple domains, these discrete classification algorithms are ineffective for heterogeneous reconstruction as they often omit functionally relevant or transient states [83]. As a result, more focused approaches have evolved to deal with continuous flexibility in cryo-EM data. These include, multi-body refinement, which uses a discrete number of independently moving, rigid bodies to model the dynamics within a protein complex and improve density maps of flexible regions [65,108]. Implemented in RELION, this multi-body approach has been utilised to characterise the TraR-induced structural changes in *E. coli* RNAP to regulate transcription. An analogous strategy, masked 3D refinement, can also employed to combat structural heterogeneity by applying a mask that excludes contents outside a region of interest, local resolution can be improved [26]. To avoid introducing artefacts or overfitting, generated mask are often low-pass filtered with soft edges [83]. 3D Variability Analysis (3DVA), available in cryoSPARC, is an algorithm that fits a linear subspace model to visualise molecular motions of macromolecules at high resolution [66]. 3DVA has since been utilised to resolve the dynamic interaction between RNAP and the transcription factor, HelD from *Bacillus subtilis* [42]. A similar tool, cryoDRGN (http://cryodrgn.csail.mit.edu), utilises deep neural networks to reconstruct density maps that model both discrete compositional heterogeneity and continuous conformational changes [67]. While these methods primarily combat motions across multiple domains, they can provide a potential trajectory to additionally refine single domain motions.

Following 3D reconstruction, atomic models provide the basis to structurally and functionally interpret the cryo-EM density maps. Most commonly, this process involves the input of existing X-ray or NMR structures from the PDB [22,26–29], however in their absence, the neural network-based, structure prediction programs, AlphaFold [109] or RoseTTA [110] now offer an alternative. To guide this initial structure into the target density map, various flexible fitting tools, such as *cryo_fit* within phenix [70], ISOLDE [71] or iMODFIT [72] in Chimera, *Namdinator* [73] among others [29,111] can be used to accommodate conformational heterogeneity by utilising molecular dynamics simulations or normal mode analysis [112]. In scenarios where it is difficult to annotate protein and DNA in the density maps, the recent tools Haruspex [68] and Emap2sec+ [69] can be employed to detect these structures in high-resolution maps (>4 Å) and lower resolution maps 5–10 Å, respectively using neural networks. Aside from flexible fitting, atomic models can also be constructed *de novo* when the resolution is better than 3.5 Å [113]. Although the Rosetta refinement strategy can be implemented as a *de novo* model-building approach for cryo-EM maps at 3–5 Å [114]. For tools that permit further refinement and validation of these models, we direct readers to a more detailed review [115].

Conclusion

The advent of cryo-EM, driven by advances in hardware and data processing, has revolutionised our understanding of transcription regulation in bacteria. This is evidenced by the rapidly expanding structural repertoire of bacterial transcription complexes over the past two years, with >65 resolved by cryo-EM to date. As illustrated within this review, these structures have shed light on unique conformational changes not seen in previous crystallographic studies, which include: promoter remodelling to stabilise intermediate complexes of transcription initiation [22,23,37]; insights into the conformational plasticity during the transition from transcription initiation to elongation [22,28]; the cooperative assembly of the transcriptional repressor, NanR [26]; the effector-induced reconfiguration of BusR to bind a bipartite DNA motif [27]; and the interaction of Crl and WhiB7 with RNAP to tether the σ-factors that they regulate through protein–protein or protein–DNA interactions, respectively [32,34]. However, despite these advancements reviewed here, gaps in our knowledge remain.

To yield a comprehensive model of transcription regulation for a given bacterial system, researchers must employ an integrative structural, biophysical and biochemical approach [116]. While cryo-EM is powerful at resolving large, conformationally dynamic assemblies, which are difficult to be captured by crystallography due to crystal packing, there are still limitations in size and resolution. In contrast, crystallography is better suited to yield atomic coordinates of macromolecules under 100–200 kDa and can attain higher resolutions (<2 Å) as molecules are constrained to a crystal lattice [117]. It is now routine to dock these smaller, more ordered

structures into cryo-EM maps using flexible fitting to accommodate minor conformational differences and allow a more in-depth interpretation of the model [112]. Hence, crystallography remains a useful tool in structural biology to complement cryo-EM studies.

In keeping with the theme of an integrative approach, small angle X-ray scattering (SAXS) can 'observe' the conformational landscape of transcription complexes in solution and compare this with the cryo-EM model to evaluate biological relevance [118]. Likewise, the stoichiometry and molecular mass of protein–DNA complexes can be determined in solution, using an emerging, label-free analytical ultracentrifugation method that features multi-wavelength detection to deconvolute the spectral signals of protein and DNA based on their unique optical properties [26,119,120]. Single molecule mass photometry can also be employed as a tool to determine molecular mass of complexes in solution [121]. Other tools for characterising protein–DNA interactions (reviewed in [122]), include cryo-electron tomography, NMR, mass spectrometry/cross linking, Förster resonance energy transfer (FRET) and hydrogen–deuterium exchange.

Moving forward, future structural endeavours using cryo-EM will no doubt build on the contributions highlighted here to deconvolute the complicated and often multifaceted molecular mechanisms of transcriptional regulation in bacteria. Alongside regular hardware and software improvements, we anticipate machine learning methods, such as cryoDRGN [67], along with the prospect of time-resolved cryo-EM [123,124] will enable researchers to explore more transient intermediate complexes and thus gain a deeper understanding of the molecular choreography that drives these regulatory mechanisms.

Perspectives

- Transcription complexes are dynamic assembles whose function is often intertwined with their many structural configurations. The precise choreography and nature of these motions remains incompletely understood. This knowledge is essential to understand the molecular mechanisms of transcription regulation in bacteria.
- Fuelled by the 'resolution revolution', cryo-EM has emerged to provide researchers a means of probing these larger and structurally heterogeneous macromolecules, which are sensitive to crystallisation. To date, these studies have contributed >65 complex structures and provided unprecedented insights into bacterial transcription regulation.
- The prospect of temporally linking the dynamic nature of transcription complexes remains of immense interest. Excitingly, the evolution of machine learning and time-resolved cryo-EM applications represent a future avenue to explore these transient intermediates.

Competing Interests
The authors declare that there are no competing interests associated with the manuscript.

Funding
R.C.J.D. acknowledges the following for funding support, in part: 1) the New Zealand Royal Society Marsden Fund (contract UOC1506); 2) a Ministry of Business, Innovation and Employment Smart Ideas grant (contract UOCX1706); and 3) the Biomolecular Interactions Centre (University of Canterbury).

Open Access
Open access for this article was enabled by the participation of University of Melbourne in an all-inclusive *Read & Publish* pilot with Portland Press and the Biochemical Society under a transformative agreement with CAUL.

Author Contributions
D.M.W., R.C.J.D., and C.R.H. reviewed the literature, wrote the manuscript and created the figures.

Biochemical Society Transactions (2021) **49** 2711–2726
https://doi.org/10.1042/BST20210674

Acknowledgements

We thank the editors from the Biochemical Society for the invitation to contribute.

Abbreviations

3DVA, 3D Variability Analysis; CAP, cyclic AMP (cAMP) receptor protein; CryoDRGN, Cryo-Deep Reconstructing Generative Networks; Cryo-EM, Cryo-electron microscopy; CTD/NTD, C-terminal/ N-terminal domain; DNA, Deoxyribonucleic acid; EMDB, Electron Microscopy Data Bank; kDa, Kilodalton; PDB, Protein Data Bank; RNAP, RNA polymerase; RPc, RNA polymerase closed complex; RPo, RNA polymerase open complex; σ-factor, Sigma factor.

References

1 Burgess, R.R. (1969) Separation and characterization of the subunits of ribonucleic acid polymerase. *J. Biol. Chem.* **244**, 6168–6176 https://doi.org/10.1016/S0021-9258(18)63521-5
2 Vassylyev, D.G., Sekine, S.I., Laptenko, O., Lee, J., Vassylyeva, M.N., Borukhov, S. et al. (2002) Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature* **417**, 712–719 https://doi.org/10.1038/nature752
3 Murakami, K.S. (2002) Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science* **296**, 1285–1290 https://doi.org/10.1126/science.1069595
4 Feklístov, A., Sharon, B.D., Darst, S.A. and Gross, C.A. (2014) Bacterial sigma factors: a historical, structural, and genomic perspective. *Annu. Rev. Microbiol.* **68**, 357–376 https://doi.org/10.1146/annurev-micro-092412-155737
5 Campbell, E.A., Muzzin, O., Chlenov, M., Sun, J.L., Olson, C.A., Weinman, O. et al. (2002) Structure of the bacterial RNA polymerase promoter specificity σ subunit. *Mol. Cell* **9**, 527–539 https://doi.org/10.1016/S1097-2765(02)00470-7
6 Barne, K.A., Bown, J.A., Busby, S.J.W. and Minchin, S.D. (1997) Region 2.5 of the *Escherichia coli* RNA polymerase σ70 subunit is responsible for the recognition of the 'extended –10' motif at promoters. *EMBO J.* **16**, 4034–4040 https://doi.org/10.1093/emboj/16.13.4034
7 Feklistov, A., Barinova, N., Sevostyanova, A., Heyduk, E., Bass, I., Vvedenskaya, I. et al. (2006) A basal promoter element recognized by free RNA polymerase σ subunit determines promoter recognition by RNA polymerase holoenzyme. *Mol. Cell* **23**, 97–107 https://doi.org/10.1016/j.molcel.2006.06.010
8 Haugen, S.P., Berkmen, M.B., Ross, W., Gaal, T., Ward, C. and Gourse, R.L. (2006) rRNA promoter regulation by nonoptimal binding of σ region 1.2: an additional recognition element for RNA polymerase. *Cell* **125**, 1069–1082 https://doi.org/10.1016/j.cell.2006.04.034
9 Estrem, S.T., Ross, W., Gaal, T., Chen, Z.W., Niu, W., Ebright, R.H. et al. (1999) Bacterial promoter architecture: subsite structure of UP elements and interactions with the carboxy-terminal domain of the RNA polymerase alpha subunit. *Genes Dev.* **13**, 2134–2147 https://doi.org/10.1101/gad.13.16.2134
10 Browning, D.F. and Busby, S.J.W. (2016) Local and global regulation of transcription initiation in bacteria. *Nat. Rev. Microbiol.* **14**, 638–650 https://doi.org/10.1038/nrmicro.2016.103
11 Browning, D.F. and Busby, S.J. (2004) The regulation of bacterial transcription initiation. *Nat. Rev. Microbiol.* **2**, 57–65 https://doi.org/10.1038/nrmicro787
12 Seshasayee, A.S.N., Sivaraman, K. and Luscombe, N.M. (2011) An Overview of Prokaryotic Transcription Factors. In *A Handbook of Transcription Factors* (Hughes, T.R., ed.), pp. 7–23, Springer Netherlands, Dordrecht
13 Perez-Rueda, E., Hernandez-Guerrero, R., Martinez-Nuñez, M.A., Armenta-Medina, D., Sanchez, I. and Ibarra, J.A. (2018) Abundance, diversity and domain architecture variability in prokaryotic DNA-binding transcription factors. *PLoS ONE* **13**, e0195332 https://doi.org/10.1371/journal.pone.0195332
14 Weisberg, R.A. (2004) A genetic switch: phage lambda revisited. Third Edition. By Mark Ptashne. *Q. Rev. Biol.* **79**, 427–428 https://doi.org/10.1086/428187
15 Ruff, E., Record, M. and Artsimovitch, I. (2015) Initial events in bacterial transcription initiation. *Biomolecules* **5**, 1035–1062 https://doi.org/10.3390/biom5021035
16 Lee, D.J., Minchin, S.D. and Busby, S.J. (2012) Activating transcription in bacteria. *Annu. Rev. Microbiol.* **66**, 125–152 https://doi.org/10.1146/annurev-micro-092611-150012
17 Helmann, J.D. (2009) RNA polymerase: a nexus of gene regulation. *Methods (San Diego, Calif)* **47**, 1–5 https://doi.org/10.1016/j.ymeth.2008.12.001
18 Hanske, J., Sadian, Y. and Müller, C.W. (2018) The cryo-EM resolution revolution and transcription complexes. *Curr. Opin. Struct. Biol.* **52**, 8–15 https://doi.org/10.1016/j.sbi.2018.07.002
19 Nogales, E. and Scheres, S.H. (2015) Cryo-EM: a unique tool for the visualization of macromolecular complexity. *Mol. Cell* **58**, 677–689 https://doi.org/10.1016/j.molcel.2015.02.019
20 Nogales, E., Louder, R.K. and He, Y. (2016) Cryo-EM in the study of challenging systems: the human transcription pre-initiation complex. *Curr. Opin. Struct. Biol.* **40**, 120–127 https://doi.org/10.1016/j.sbi.2016.09.009
21 Benjin, X. and Ling, L. (2020) Developments, applications, and prospects of cryo-electron microscopy. *Protein Sci.* **29**, 872–882 https://doi.org/10.1002/pro.3805
22 Yang, Y., Liu, C., Zhou, W., Shi, W., Chen, M., Zhang, B. et al. (2021) Structural visualization of transcription activated by a multidrug-sensing MerR family regulator. *Nat. Commun.* **12**, 2702 https://doi.org/10.1038/s41467-021-22990-8
23 Fang, C., Li, L., Zhao, Y., Wu, X., Philips, S.J., You, L. et al. (2020) The bacterial multidrug resistance regulator BmrR distorts promoter DNA to activate transcription. *Nat. Commun.* **11**, 6284 https://doi.org/10.1038/s41467-020-20134-y
24 Shi, W., Zhang, B., Jiang, Y., Liu, C., Zhou, W., Chen, M. et al. (2021) Structural basis of copper-efflux-regulator-dependent transcription activation. *iScience* **24**, 102449 https://doi.org/10.1016/j.isci.2021.102449
25 Fang, C., Philips, S.J., Wu, X., Chen, K., Shi, J., Shen, L. et al. (2021) Cuer activates transcription through a DNA distortion mechanism. *Nat. Chem. Biol.* **17**, 57–64 https://doi.org/10.1038/s41589-020-00653-x

26 Horne, C.R., Venugopal, H., Panjikar, S., Wood, D.M., Henrickson, A., Brookes, E. et al. (2021) Mechanism of NanR gene repression and allosteric induction of bacterial sialic acid metabolism. *Nat. Commun.* **12**, 1988 https://doi.org/10.1038/s41467-021-22253-6
27 Bandera, A.M., Bartho, J., Lammens, K., Drexler, D.J., Kleinschwärzer, J., Hopfner, K.P. et al. (2021) Busr senses bipartite DNA binding motifs by a unique molecular ruler architecture. *Nucleic Acids Res.* **49**, 10166–10177 https://doi.org/10.1093/nar/gkab736
28 Chen, J., Gopalkrishnan, S., Chiu, C., Chen, A.Y., Campbell, E.A., Gourse, R.L. et al. (2019) *E. coli* TraR allosterically regulates transcription initiation by altering RNA polymerase conformation. *eLife* **8**, e49375 https://doi.org/10.7554/eLife.49375
29 Zhou, S., Bhukya, H., Malet, N., Harrison, P.J., Rea, D., Belousoff, M.J. et al. (2021) Molecular basis for control of antibiotic production by a bacterial hormone. *Nature* **590**, 463–467 https://doi.org/10.1038/s41586-021-03195-x
30 Shi, W., Jiang, Y., Deng, Y., Dong, Z. and Liu, B. (2020) Visualization of two architectures in class-II CAP-dependent transcription activation. *PLoS Biol.* **18**, e3000706 https://doi.org/10.1371/journal.pbio.3000706
31 Liu, B., Hong, C., Huang, R.K., Yu, Z. and Steitz, T.A. (2017) Structural basis of bacterial transcription activation. *Science* **358**, 947–951 https://doi.org/10.1126/science.aao1923
32 Cartagena, A.J., Banta, A.B., Sathyan, N., Ross, W., Gourse, R.L., Campbell, E.A. et al. (2019) Structural basis for transcription activation by Crl through tethering of σ(S) and RNA polymerase. *Proc. Natl Acad. Sci. U.S.A.* **116**, 18923–7 https://doi.org/10.1073/pnas.1910827116
33 Shi, J., Li, F., Wen, A., Yu, L., Wang, L., Wang, F. et al. (2021) Structural basis of transcription activation by the global regulator Spx. *Nucleic Acids Res.* **49**, 10756–10769 https://doi.org/10.1093/nar/gkab790
34 Lilic, M., Darst, S.A. and Campbell, E.A. (2021) Structural basis of transcriptional activation by the *Mycobacterium tuberculosis* intrinsic antibiotic-resistance transcription factor WhiB7. *Mol. Cell* **81**, 2875–2886.e5 https://doi.org/10.1016/j.molcel.2021.05.017
35 Capodagli, G.C., Tylor, K.M., Kaelber, J.T., Petrou, V.I., Federle, M.J. and Neiditch, M.B. (2020) Structure-function studies of Rgg binding to pheromones and target promoters reveal a model of transcription factor interplay. *Proc. Natl Acad. Sci. U.S.A.* **117**, 24494–24502 https://doi.org/10.1073/pnas.2008427117
36 Abdelkareem, M., Saint-André, C., Takacs, M., Papai, G., Crucifix, C., Guo, X. et al. (2019) Structural basis of transcription: RNA polymerase backtracking and Its reactivation. *Mol. Cell* **75**, 298–309.e4 https://doi.org/10.1016/j.molcel.2019.04.029
37 Boyaci, H., Chen, J., Jansen, R., Darst, S.A. and Campbell, E.A. (2019) Structures of an RNA polymerase promoter melting intermediate elucidate DNA unwinding. *Nature* **565**, 382–385 https://doi.org/10.1038/s41586-018-0840-5
38 Lilic, M., Chen, J., Boyaci, H., Braffman, N., Hubin, E.A., Herrmann, J. et al. (2020) The antibiotic sorangicin A inhibits promoter DNA unwinding in a *Mycobacterium tuberculosis* rifampicin-resistant RNA polymerase. *Proc. Natl Acad. Sci. U.S.A.* **117**, 30423–30432 https://doi.org/10.1073/pnas.2013706117
39 Wang, F., Shi, J., He, D., Tong, B., Zhang, C., Wen, A. et al. (2020) Structural basis for transcription inhibition by *E. coli* SspA. *Nucleic Acids Res.* **48**, 9931–9942 https://doi.org/10.1093/nar/gkaa672
40 Shin, Y., Qayyum, M.Z., Pupov, D., Esyunina, D., Kulbachinskiy, A. and Murakami, K.S. (2021) Structural basis of ribosomal RNA transcription regulation. *Nat. Commun.* **12**, 528 https://doi.org/10.1038/s41467-020-20776-y
41 Kang, J.Y., Mooney, R.A., Nedialkov, Y., Saba, J., Mishanina, T.V., Artsimovitch, I. et al. (2018) Structural basis for transcript elongation control by nusG family universal regulators. *Cell* **173**, 1650–1662.e14 https://doi.org/10.1016/j.cell.2018.05.017
42 Newing, T.P., Oakley, A.J., Miller, M., Dawson, C.J., Brown, S.H.J., Bouwer, J.C. et al. (2020) Molecular basis for RNA polymerase-dependent transcription complex recycling by the helicase-like motor protein HelD. *Nat. Commun.* **11**, 6420 https://doi.org/10.1038/s41467-020-20157-5
43 Kouba, T., Koval', T., Sudzinová, P., Pospíšil, J., Brezovská, B., Hnilicová, J. et al. (2020) Mycobacterial HelD is a nucleic acids-clearing factor for RNA polymerase. *Nat. Commun.* **11**, 6419 https://doi.org/10.1038/s41467-020-20158-4
44 Klein, B.J., Bose, D., Baker, K.J., Yusoff, Z.M., Zhang, X. and Murakami, K.S. (2010) RNA polymerase and transcription elongation factor Spt4/5 complex structure. *Proc. Natl Acad. Sci. U.S.A.* **108**, 546–550 https://doi.org/10.1073/pnas.1013828108
45 Shi, J., Wen, A., Zhao, M., Jin, S., You, L., Shi, Y. et al. (2020) Structural basis of Mfd-dependent transcription termination. *Nucleic Acids Res.* **48**, 11762–11772 https://doi.org/10.1093/nar/gkaa904
46 Kang, J.Y., Llewellyn, E., Chen, J., Olinares, P.D.B., Brewer, J., Chait, B.T. et al. (2021) Structural basis for transcription complex disruption by the Mfd translocase. *eLife* **10**, e62117 https://doi.org/10.7554/eLife.62117
47 Glyde, R., Ye, F., Jovanovic, M., Kotta-Loizou, I., Buck, M. and Zhang, X. (2018) Structures of bacterial RNA polymerase complexes reveal the mechanism of DNA loading and transcription initiation. *Mol. Cell* **70**, 1111–1120.e3 https://doi.org/10.1016/j.molcel.2018.05.021
48 Hobman, J.L. (2007) Merr family transcription activators: similar designs, different specificities. *Mol. Microbiol.* **63**, 1275–1278 https://doi.org/10.1111/j.1365-2958.2007.05608.x
49 Julian, D.J., Kershaw, C.J., Brown, N.L. and Hobman, J.L. (2009) Transcriptional activation of MerR family promoters in Cupriavidus metallidurans CH34. *Antonie Van Leeuwenhoek* **96**, 149–159 https://doi.org/10.1007/s10482-008-9293-4
50 Heldwein, E.E. and Brennan, R.G. (2001) Crystal structure of the transcription activator BmrR bound to DNA and a drug. *Nature* **409**, 378–382 https://doi.org/10.1038/35053138
51 Philips, S.J., Canalizo-Hernandez, M., Yildirim, I., Schatz, G.C., Mondragón, A. and O'Halloran, T.V. (2015) TRANSCRIPTION. Allosteric transcriptional regulation via changes in the overall topology of the core promoter. *Science* **349**, 877–881 https://doi.org/10.1126/science.aaa9809
52 Schultz, S.C., Shields, G.C. and Steitz, T.A. (1990) Crystallization of *Escherichia coli* catabolite gene activator protein with its DNA binding site. *J. Mol. Biol.* **213**, 159–166 https://doi.org/10.1016/S0022-2836(05)80128-7
53 Benoff, B., Yang, H., Lawson, C.L., Parkinson, G., Liu, J., Blatter, E. et al. (2002) Structural basis of transcription activation: the CAP-alpha CTD-DNA complex. *Science* **297**, 1562–1566 https://doi.org/10.1126/science.1076376
54 Hubin, E.A., Fay, A., Xu, C., Bean, J.M., Saecker, R.M., Glickman, M.S. et al. (2017) Structure and function of the mycobacterial transcription initiation complex with the essential regulator rbpA. *eLife* **6**, e22520 https://doi.org/10.7554/eLife.22520
55 Liu, B., Zuo, Y. and Steitz, T.A. (2016) Structures of E. coli σS-transcription initiation complexes provide new insights into polymerase mechanism. *Proc. Natl Acad. Sci. U.S.A.* **113**, 4051–4056 https://doi.org/10.1073/pnas.1520555113
56 Morris, R.P., Nguyen, L., Gatfield, J., Visconti, K., Nguyen, K., Schnappinger, D. et al. (2005) Ancestral antibiotic resistance in *Mycobacterium tuberculosis*. *Proc. Natl Acad. Sci. U.S.A.* **102**, 12200–12205 https://doi.org/10.1073/pnas.0505446102

57 Coombes, D., Davies, J.S., Newton-Vesty, M.C., Horne, C.R., Setty, T.G., Subramanian, R. et al. (2020) The basis for non-canonical ROK family function in the N-acetylmannosamine kinase from the pathogen *Staphylococcus aureus*. *J. Biol. Chem.* **295**, 3301–3315 https://doi.org/10.1074/jbc.RA119.010526

58 North, R.A., Horne, C.R., Davies, J.S., Remus, D.M., Muscroft-Taylor, A.C., Goyal, P. et al. (2018) "Just a spoonful of sugar…": import of sialic acid across bacterial cell membranes. *Biophys. Rev.* **10**, 219–227 https://doi.org/10.1007/s12551-017-0343-x

59 Currie, M.J., Manjunath, L., Horne, C.R., Rendle, P.M., Subramanian, R., Friemann, R. et al. (2021) N-Acetylmannosamine-6-phosphate 2-epimerase uses a novel substrate-assisted mechanism to catalyze amino sugar epimerization. *J. Biol. Chem.* **297**, 101113 https://doi.org/10.1016/j.jbc.2021.101113

60 Horne, C.R., Kind, L., Davies, J.S. and Dobson, R.C.J. (2020) On the structure and function of *Escherichia coli* YjhC: an oxidoreductase involved in bacterial sialic acid metabolism. *Proteins* **88**, 654–668 https://doi.org/10.1002/prot.25846

61 de Lencastre, H., Dengler, V., McCallum, N., Kiefer, P., Christen, P., Patrignani, A. et al. (2013) Mutation in the C-Di-AMP cyclase dacA affects fitness and resistance of methicillin resistant *Staphylococcus aureus*. *PLoS ONE.* **8**, e73512 https://doi.org/10.1371/journal.pone.0073512

62 Chen, J., Chiu, C., Gopalkrishnan, S., Chen, A.Y., Olinares, P.D.B., Saecker, R.M. et al. (2020) Stepwise promoter melting by bacterial RNA polymerase. *Mol. Cell* **78**, 275–288.e6 https://doi.org/10.1016/j.molcel.2020.02.017

63 Lyumkis, D. (2019) Challenges and opportunities in cryo-EM single-particle analysis. *J. Biol. Chem.* **294**, 5181–5197 https://doi.org/10.1074/jbc.REV118.005602

64 Costa, T.R.D., Ignatiou, A. and Orlova, E.V. (2017) Structural analysis of protein complexes by cryo electron microscopy. *Methods Mol. Biol.* **1615**, 377–413 https://doi.org/10.1007/978-1-4939-7033-9_28

65 Nakane, T., Kimanius, D., Lindahl, E. and Scheres, S.H. (2018) Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION. *eLife* **7**, e36861 https://doi.org/10.7554/eLife.36861

66 Punjani, A. and Fleet, D.J. (2021) 3D variability analysis: resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM. *J. Struct. Biol.* **213**, 107702 https://doi.org/10.1016/j.jsb.2021.107702

67 Zhong, E.D., Bepler, T., Berger, B. and Davis, J.H. (2021) CryoDRGN: reconstruction of heterogeneous cryo-EM structures using neural networks. *Nat. Methods* **18**, 176–185 https://doi.org/10.1038/s41592-020-01049-4

68 Mostosi, P., Schindelin, H., Kollmannsberger, P. and Thorn, A. (2020) Haruspex: a neural network for the automatic identification of oligonucleotides and protein secondary structure in cryo-electron microscopy maps. *Angew. Chem. Int. Ed. Engl.* **59**, 14788–14795 https://doi.org/10.1002/anie.202000421

69 Wang, X., Alnabati, E., Aderinwale, T.W., Maddhuri Venkata Subramaniya, S.R., Terashi, G. and Kihara, D. (2021) Detecting protein and DNA/RNA structures in cryo-EM maps of intermediate resolution using deep learning. *Nat. Commun.* **12**, 2302 https://doi.org/10.1038/s41467-021-22577-3

70 Kim, D.N., Moriarty, N.W., Kirmizialtin, S., Afonine, P.V., Poon, B., Sobolev, O.V. et al. (2019) Cryo_fit: democratization of flexible fitting for cryo-EM. *J. Struct. Biol.* **208**, 1–6 https://doi.org/10.1016/j.jsb.2019.05.012

71 Croll, T.I. (2018) ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr. D Struct. Biol.* **74**, 519–530 https://doi.org/10.1107/S2059798318002425

72 Lopéz-Blanco, J.R. and Chacón, P. (2013) iMODFIT: efficient and robust flexible fitting based on vibrational analysis in internal coordinates. *J. Struct. Biol.* **184**, 261–270 https://doi.org/10.1016/j.jsb.2013.08.010

73 Kidmose, R.T., Juhl, J., Nissen, P., Boesen, T., Karlsen, J.L. and Pedersen, B.P. (2019) Namdinator - automatic molecular dynamics flexible fitting of structural models into cryo-EM and crystallography experimental maps. *IUCrJ.* **6**, 526–531 https://doi.org/10.1107/S2052252519007619

74 Weissenberger, G., Henderikx, R.J.M. and Peters, P.J. (2021) Understanding the invisible hands of sample preparation for cryo-EM. *Nat. Methods* **18**, 463–471 https://doi.org/10.1038/s41592-021-01130-6

75 Cheng, Y., Grigorieff, N., Penczek, P.A. and Walz, T. (2015) A primer to single-particle cryo-electron microscopy. *Cell* **161**, 438–449 https://doi.org/10.1016/j.cell.2015.03.050

76 Stark, H. and Chari, A. (2016) Sample preparation of biological macromolecular assemblies for the determination of high-resolution structures by cryo-electron microscopy. *Microscopy (Oxf)* **65**, 23–34 https://doi.org/10.1093/jmicro/dfv367

77 Seabrook, S.A. and Newman, J. (2013) High-throughput thermal scanning for protein stability: making a good technique more robust. *ACS Comb. Sci.* **15**, 387–392 https://doi.org/10.1021/co400013v

78 Chari, A., Haselbach, D., Kirves, J.M., Ohmer, J., Paknia, E., Fischer, N. et al. (2015) Proteoplex: stability optimization of macromolecular complexes by sparse-matrix screening of chemical space. *Nat. Methods* **12**, 859–865 https://doi.org/10.1038/nmeth.3493

79 Drulyte, I., Johnson, R.M., Hesketh, E.L., Hurdiss, D.L., Scarff, C.A., Porav, S.A. et al. (2018) Approaches to altering particle distributions in cryo-electron microscopy sample preparation. *Acta Crystallogr. D Struct. Biol.* **74**, 560–571 https://doi.org/10.1107/S2059798318006496

80 Thompson, R.F., Walker, M., Siebert, C.A., Muench, S.P. and Ranson, N.A. (2016) An introduction to sample preparation and imaging by cryo-electron microscopy for structural biology. *Methods* **100**, 3–15 https://doi.org/10.1016/j.ymeth.2016.02.017

81 Basanta, B., Hirschi, M.M., Grotjahn, D.A. and Lander, G.C. (2021) A case for glycerol as an acceptable additive for single particle cryoEM samples. *bioRxiv* 1–10 https://doi.org/10.1101/2021.09.10.459874

82 Nwanochie, E. and Uversky, V.N. (2019) Structure determination by single-particle cryo-electron microscopy: only the Sky (and intrinsic disorder) is the limit. *Int. J. Mol. Sci.* **20**, 4186 https://doi.org/10.3390/ijms20174186

83 Serna, M. (2019) Hands on methods for high resolution cryo-Electron microscopy structures of heterogeneous macromolecular complexes. *Front. Mol. Biosci.* **6**, 33 https://doi.org/10.3389/fmolb.2019.00033

84 De Wijngaert, B., Sultana, S., Singh, A., Dharia, C., Vanbuel, H., Shen, J. et al. (2021) Cryo-EM structures reveal transcription initiation steps by yeast mitochondrial RNA polymerase. *Mol. Cell* **81**, 268–280.e5 https://doi.org/10.1016/j.molcel.2020.11.016

85 Kastner, B., Fischer, N., Golas, M.M., Sander, B., Dube, P., Boehringer, D. et al. (2008) Grafix: sample preparation for single-particle electron cryomicroscopy. *Nat. Methods* **5**, 53–55 https://doi.org/10.1038/nmeth1139

86 Narayanan, A., Vago, F.S., Li, K., Qayyum, M.Z., Yernool, D., Jiang, W. et al. (2018) Cryo-EM structure of *Escherichia coli* σ(70) RNA polymerase and promoter DNA complex revealed a role of σ non-conserved region during the open complex formation. *J. Biol. Chem.* **293**, 7367–7375 https://doi.org/10.1074/jbc.RA118.002161

87 Stark, H. (2010) Grafix: stabilization of fragile macromolecular complexes for single particle cryo-EM. *Methods Enzymol.* **481**, 109–126 https://doi.org/10.1016/S0076-6879(10)81005-5
88 Adamus, K., Le, S.N., Elmlund, H., Boudes, M. and Elmlund, D. (2019) Agarfix: simple and accessible stabilization of challenging single-particle cryo-EM specimens through crosslinking in a matrix of agar. *J. Struct. Biol.* **207**, 327–331 https://doi.org/10.1016/j.jsb.2019.07.004
89 Goswami, P., Locke, J. and Costa, A. (2018) Preparing frozen-hydrated protein-Nucleic acid assemblies for high-resolution cryo-EM imaging. *Methods Mol. Biol.* **1814**, 287–296 https://doi.org/10.1007/978-1-4939-8591-3_17
90 Noble, A.J., Wei, H., Dandey, V.P., Zhang, Z., Tan, Y.Z., Potter, C.S. et al. (2018) Reducing effects of particle adsorption to the air-water interface in cryo-EM. *Nat. Methods* **15**, 793–795 https://doi.org/10.1038/s41592-018-0139-3
91 Glaeser, R.M. and Han, B.G. (2017) Opinion: hazards faced by macromolecules when confined to thin aqueous films. *Biophys. Rep.* **3**, 1–7 https://doi.org/10.1007/s41048-016-0026-3
92 D'Imprima, E., Floris, D., Joppe, M., Sánchez, R., Grininger, M. and Kühlbrandt, W. (2019) Protein denaturation at the air-water interface and how to prevent it. *eLife* **8**, e42747 https://doi.org/10.7554/eLife.42747
93 Ashtiani, D., Venugopal, H., Belousoff, M., Spicer, B., Mak, J., Neild, A. et al. (2018) Delivery of femtolitre droplets using surface acoustic wave based atomisation for cryo-EM grid preparation. *J. Struct. Biol.* **203**, 94–101 https://doi.org/10.1016/j.jsb.2018.03.012
94 Dandey, V.P., Wei, H., Zhang, Z., Tan, Y.Z., Acharya, P., Eng, E.T. et al. (2018) Spotiton: new features and applications. *J. Struct. Biol.* **202**, 161–169 https://doi.org/10.1016/j.jsb.2018.01.002
95 Liu, N., Zhang, J., Chen, Y., Liu, C., Zhang, X., Xu, K. et al. (2019) Bioactive functionalized monolayer graphene for high-resolution cryo-Electron microscopy. *J. Am. Chem. Soc.* **141**, 4016–4025 https://doi.org/10.1021/jacs.8b13038
96 Wang, F., Liu, Y., Yu, Z., Li, S., Feng, S., Cheng, Y. et al. (2020) General and robust covalently linked graphene oxide affinity grids for high-resolution cryo-EM. *Proc. Natl Acad. Sci. U.S.A.* **117**, 24269–24273 https://doi.org/10.1073/pnas.2009707117
97 Glaeser, R.M., Han, B.-G., Watson, Z., Ward, F. and Cate, J.H.D. (2018) Streptavidin affinity grids for cryo-EM. *Biophys. J.* **114**, 163a https://doi.org/10.1016/j.bpj.2017.11.911
98 Tan, Y.Z., Baldwin, P.R., Davis, J.H., Williamson, J.R., Potter, C.S., Carragher, B. et al. (2017) Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nat. Methods* **14**, 793–796 https://doi.org/10.1038/nmeth.4347
99 Russo, C.J. and Passmore, L.A. (2014) Electron microscopy: ultrastable gold substrates for electron cryomicroscopy. *Science* **346**, 1377–1380 https://doi.org/10.1126/science.1259530
100 Naydenova, K., Peet, M.J. and Russo, C.J. (2019) Multifunctional graphene supports for electron cryomicroscopy. *Proc. Natl Acad. Sci. U.S.A.* **116**, 11718–11724 https://doi.org/10.1073/pnas.1904766116
101 Chen, J., Noble, A.J., Kang, J.Y. and Darst, S.A. (2019) Eliminating effects of particle adsorption to the air/water interface in single-particle cryo-electron microscopy: bacterial RNA polymerase and CHAPSO. *J. Struct. Biol. X.* **1**, 100005 https://doi.org/10.1016/j.yjsbx.2019.100005
102 Ha, B., Larsen, K.P., Zhang, J., Fu, Z., Montabana, E., Jackson, L.N. et al. (2021) High-resolution view of HIV-1 reverse transcriptase initiation complexes and inhibition by NNRTI drugs. *Nat. Commun.* **12**, 2500 https://doi.org/10.1038/s41467-021-22628-9
103 Larsen, K.P., Mathiharan, Y.K., Kappel, K., Coey, A.T., Chen, D.H., Barrero, D. et al. (2018) Architecture of an HIV-1 reverse transcriptase initiation complex. *Nature* **557**, 118–122 https://doi.org/10.1038/s41586-018-0055-9
104 Aksel, T., Yu, Z., Cheng, Y. and Douglas, S.M. (2021) Molecular goniometers for single-particle cryo-electron microscopy of DNA-binding proteins. *Nat. Biotechnol.* **39**, 378–386 https://doi.org/10.1038/s41587-020-0716-8
105 Scheres, S.H. (2012) RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 https://doi.org/10.1016/j.jsb.2012.09.006
106 Punjani, A., Rubinstein, J.L., Fleet, D.J. and Brubaker, M.A. (2017) cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 https://doi.org/10.1038/nmeth.4169
107 Grant, T., Rohou, A. and Grigorieff, N. (2018) cisTEM, user-friendly software for single-particle image processing. *eLife* **7**, e35383 https://doi.org/10.7554/eLife.35383
108 Nakane, T. and Scheres, S.H.W. (2021) Multi-body refinement of cryo-EM images in RELION. *Methods Mol. Biol.* **2215**, 145–160 https://doi.org/10.1007/978-1-0716-0966-8_7
109 Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O. et al. (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 https://doi.org/10.1038/s41586-021-03819-2
110 Bradley, P., Misura, K.M. and Baker, D. (2005) Toward high-resolution de novo structure prediction for small proteins. *Science* **309**, 1868–1871 https://doi.org/10.1126/science.1113801
111 Kulik, M., Mori, T. and Sugita, Y. (2021) Multi-scale flexible fitting of proteins to cryo-EM density maps at medium resolution. *Front. Mol. Biosci.* **8**, 631854 https://doi.org/10.3389/fmolb.2021.631854
112 Trabuco, L.G., Villa, E., Mitra, K., Frank, J. and Schulten, K. (2008) Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics. *Structure* **16**, 673–683 https://doi.org/10.1016/j.str.2008.03.005
113 Alnabati, E. and Kihara, D. (2020) Advances in structure modeling methods for cryo-electron microscopy maps. *Molecules* **25**, 82 https://doi.org/10.3390/molecules25010082
114 Wang, R.Y.-R., Kudryashev, M., Li, X., Egelman, E.H., Basler, M., Cheng, Y. et al. (2015) De novo protein structure determination from near-atomic-resolution cryo-EM maps. *Nat. Methods* **12**, 335–338 https://doi.org/10.1038/nmeth.3287
115 Beckers, M., Mann, D. and Sachse, C. (2021) Structural interpretation of cryo-EM image reconstructions. *Prog. Biophys. Mol. Biol.* **160**, 26–36 https://doi.org/10.1016/j.pbiomolbio.2020.07.004
116 Rout, M.P. and Sali, A. (2019) Principles for integrative structural biology studies. *Cell* **177**, 1384–1403 https://doi.org/10.1016/j.cell.2019.05.016
117 Shoemaker, S.C. and Ando, N. (2018) X-rays in the cryo-electron microscopy Era: structural biology's dynamic future. *Biochemistry* **57**, 277–285 https://doi.org/10.1021/acs.biochem.7b01031
118 Mertens, H.D. and Svergun, D.I. (2010) Structural characterization of proteins and complexes using small-angle X-ray solution scattering. *J. Struct. Biol.* **172**, 128–141 https://doi.org/10.1016/j.jsb.2010.06.012

Biochemical Society Transactions (2021) **49** 2711–2726
https://doi.org/10.1042/BST20210674

119 Gorbet, G.E., Pearson, J.Z., Demeler, A.K., Cölfen, H. and Demeler, B. (2015) Next-Generation AUC: analysis of multiwavelength analytical ultracentrifugation data. *Methods Enzymol.* **562**, 27–47 https://doi.org/10.1016/bs.mie.2015.04.013
120 Horne, C.R., Henrickson, A., Demeler, B. and Dobson, R.C.J. (2020) Multi-wavelength analytical ultracentrifugation as a tool to characterise protein-DNA interactions in solution. *Eur. Biophys. J.* **49**, 819–827 https://doi.org/10.1007/s00249-020-01481-6
121 Li, Y., Struwe, W.B. and Kukura, P. (2020) Single molecule mass photometry of nucleic acids. *Nucleic Acids Res.* **48**, e97 https://doi.org/10.1093/nar/gkaa632
122 Ziegler, S.J., Mallinson, S.J.B., St John, P.C. and Bomble, Y.J. (2021) Advances in integrative structural biology: towards understanding protein complexes in their cellular context. *Comput. Struct. Biotechnol. J.* **19**, 214–225 https://doi.org/10.1016/j.csbj.2020.11.052
123 Dandey, V.P., Budell, W.C., Wei, H., Bobe, D., Maruthi, K., Kopylov, M. et al. (2020) Time-resolved cryo-EM using spotiton. *Nat. Methods* **17**, 897–900 https://doi.org/10.1038/s41592-020-0925-6
124 Frank, J. (2017) Time-resolved cryo-electron microscopy: recent progress. *J. Struct. Biol.* **200**, 303–306 https://doi.org/10.1016/j.jsb.2017.06.005