



Transcribing Genes the Hard Way: *In Vitro* Reconstitution of Nanoarchaeal RNA Polymerase Reveals Unusual Active Site Properties

Sven Nottebaum^{1,2} and Robert O. J. Weizierl^{1*}

¹Department of Life Sciences, Imperial College London, London, United Kingdom, ²Orthomol Pharmazeutische Vertriebs GmbH, Langenfeld, Germany

OPEN ACCESS

Edited by:

Francisco Navarro,
University of Jaén, Spain

Reviewed by:

Thomas J. Santangelo,
Colorado State University,
United States

Zachary Burton,
Michigan State University,
United States

*Correspondence:

Robert O. J. Weizierl
r.weizierl@imperial.ac.uk

Specialty section:

This article was submitted to
Protein and RNA Networks,
a section of the journal
Frontiers in Molecular Biosciences

Received: 18 February 2021

Accepted: 26 April 2021

Published: 11 May 2021

Citation:

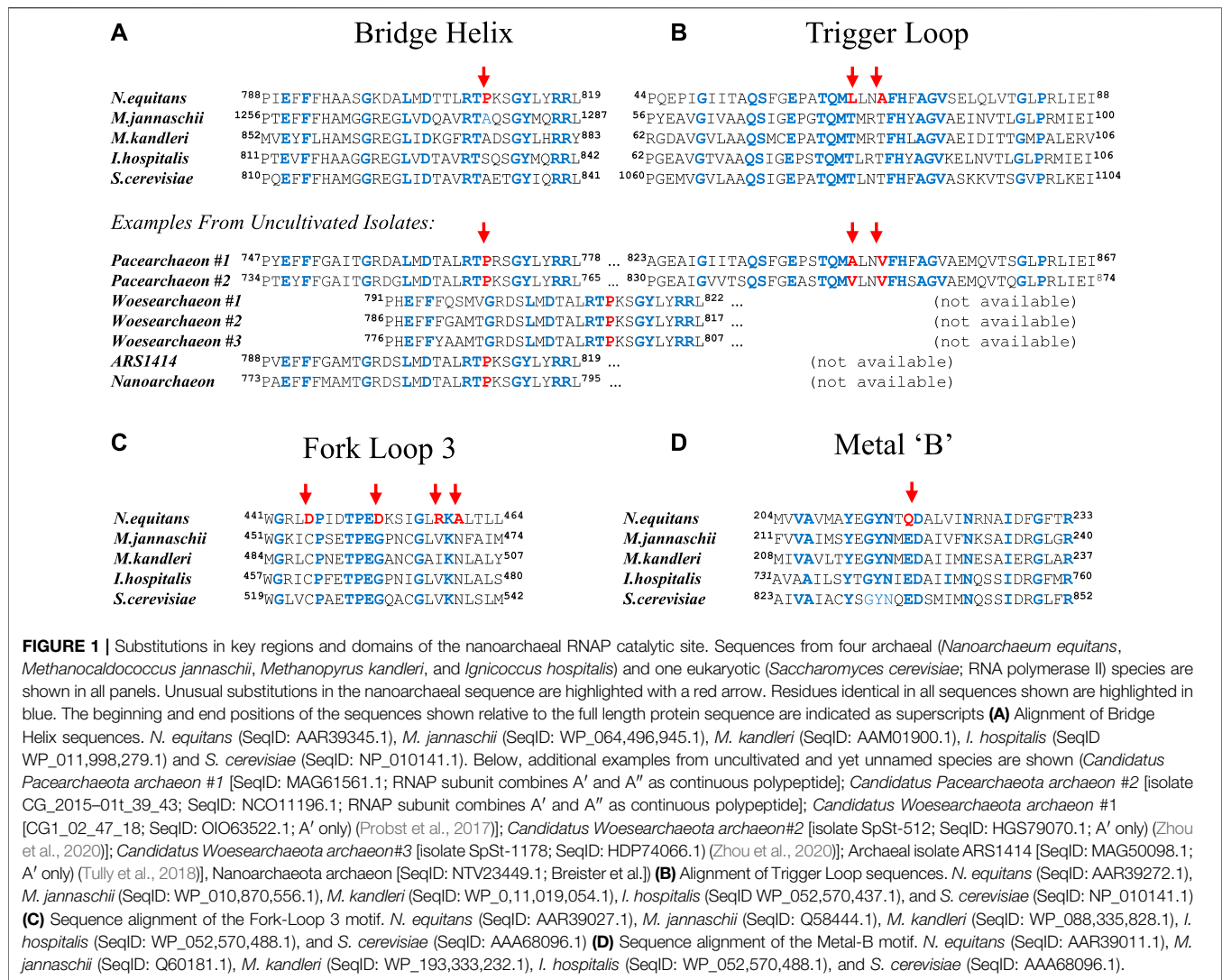
Nottebaum S and Weizierl ROJ
(2021) Transcribing Genes the Hard
Way: *In Vitro* Reconstitution of
Nanoarchaeal RNA Polymerase
Reveals Unusual Active
Site Properties.
Front. Mol. Biosci. 8:669314.
doi: 10.3389/fmolb.2021.669314

Nanoarchaea represent a highly diverged archaeal phylum that displays many unusual biological features. The *Nanoarchaeum equitans* genome encodes a complete set of RNA polymerase (RNAP) subunits and basal factors. Several of the standard motifs in the active center contain radical substitutions that are normally expected to render the polymerase catalytically inactive. Here we show that, despite these unusual features, a RNAP reconstituted from recombinant *Nanoarchaeum* subunits is transcriptionally active. Using a sparse-matrix high-throughput screening method we identified an atypical stringent requirement for fluoride ions to maximize its activity under *in vitro* transcription conditions.

Keywords: archaea, nanoarchaea, RNA polymerase, catalytic center, active site, high-throughput assay, sparse matrix sampling, fluoride

INTRODUCTION

The basal transcriptional machineries of Archaea are intriguingly similar to the core components of the eukaryotic RNA polymerase II (RNAPII) transcriptional machinery (Cramer et al., 2001). This close similarity to eukaryotic systems, combined with the greater experimental accessibility, has established archaeal systems as key model systems for in-depth structure/function analyses of the transcriptional machinery (Werner and Weizierl 2002; Ouhammouch et al., 2004; Werner et al., 2006; Naji et al., 2007; Hirata et al., 2008; Tan et al., 2008; Thomm et al., 2009; Weizierl 2013; Fouqueau et al., 2018; Blombach et al., 2019; Wenck and Santangelo, 2020). Apart from serving as model systems for eukaryotic systems, archaea also include numerous examples of extremophiles that do not fit the general pattern (Adam et al., 2017). Such species often provide unusual examples of molecular organization that have the capacity of enlarging our understanding of fundamental molecular mechanisms by illustrating the degree of flexibility that is possible, or by providing examples for achieving the same goal in a variety of alternative ways (Coker, 2019). Some of the best-known examples include the adaption of enzymes to operate in high-salt environments (halophiles), over a wide range of temperatures (psychrophile, mesophiles, thermophiles, hyperthermophiles), or at low or high pH (acidophiles and alkaliphiles, respectively). Another interesting class of archaea are the evolutionary “outliers”, such as *Methanopyrus kandleri*, *Cenarchaeum symbiosum*, and *Nanoarchaeum equitans*. The phylogenetic classification of these species is contentious, and their protein sequences frequently contain unique and unusual substitutions that are not shared by other archaea. Such unorthodox features raise many, yet unanswered, questions regarding the evolutionary origin of such species (deep-branching evolutionary ancestry or recent degeneracy?) and often

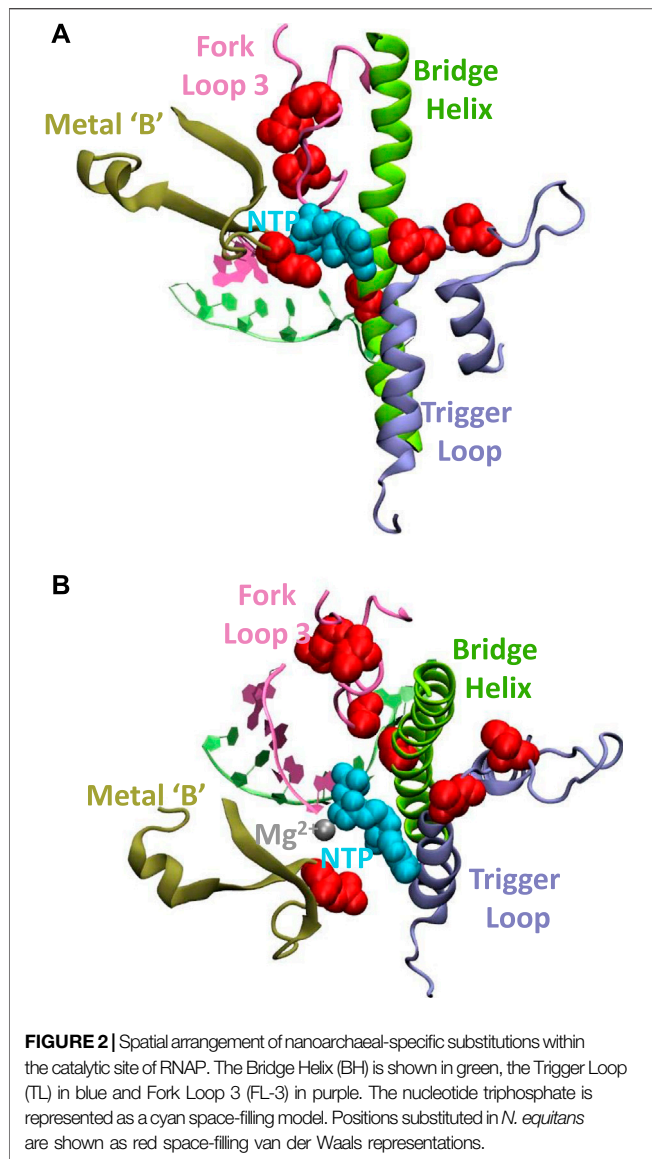


challenge fundamental concepts of apparently well-understood enzymatic pathways and mechanisms (Randau et al., 2005; Randau et al., 2008).

Here we focus on the molecular organization and properties of the RNAP from the hyperthermophile *Nanoarchaeum equitans* (from here on abbreviated as *N. equitans*, or “n” as a prefix). *N. equitans* is a highly unusual archaeon because of its diminutive size (0.35–0.5 μm cell diameter), drastically reduced genome and parasitic lifestyle (Huber et al., 2002; Waters et al., 2003; Forterre et al., 2009; Rawle et al., 2017). The genome lacks most genes that are required to produce cellular precursors, such as amino acids, nucleotides, cofactors, and lipids. These are most likely imported directly from the host cell, the crenarchaeote *Ignicoccus hospitalis*. Depending on the criteria chosen, *N. equitans* has been plausibly classified as a new and early diverging archaeal phylum (the “Nanoarchaeota” (Huber et al., 2003)), a sister branch of the Crenarchaea (Ciccarelli et al., 2006), or as a fast-evolving Euryarchaeon (Brochier et al., 2005). Recent studies have demonstrated that Nanoarchaea are widespread and occur in a variety of locations, including mesophilic and halophilic

environments (Hohn et al., 2002; McCliment et al., 2006; Probst et al., 2017; Tully et al., 2018; Zhou et al., 2020).

The *N. equitans* genome encodes a full complement of all RNA polymerase (RNAP) subunits and basal factors TBP, TFB, TFE, and TF-S (Huber et al., 2002; Waters et al., 2003). Considering the minimal size of the genome, the presence of a set of genes encoding a complete basal transcriptional machinery strongly suggests that *N. equitans* is fully capable of transcribing its own genome. We observe, however, a distinct set of substitutions in several key positions of the *ne*RNAP catalytic center that are of a unique and radical nature and raise the question whether such an enzyme could display a substantial amount of catalytic activity. The Bridge Helix (BH), Trigger Loop (TL), Fork Loop-3 (FL-3), as well as the “Metal B” binding domain (Me-B; responsible for positioning one of the two catalytically active Mg²⁺ ions) display substitutions in positions that are typically absolutely or highly conserved in all other archaeal and eukaryotic RNAPs (Figure 1). Some of these substitutions (such as the presence of a proline in the Bridge Helix (BH); Figure 1A) are predicted to have highly disruptive, non-local effects by destabilizing the *a*-helical integrity



of such a key element in a particularly critical position (Tan et al., 2008; Weinzierl, 2010b; Weinzierl, 2010a; Weinzierl, 2011). Although proline substitutions in particular places of BH results in a substantial increase the specific activity of the structurally closely related euryarchaeal RNAP from *Methanocaldococcus jannaschii* (*mj*RNAP) (Tan et al., 2008; Weinzierl, 2010b), a proline located in the position characteristic for *neA'* causes a substantial drop in activity in *mj*RNAP (Tan et al., 2008). Several other unusual substitutions in other key elements of the catalytic site (Cramer et al., 2001) are evident, including the Trigger Loop (TL; **Figure 1B**), Fork Loop-3 (FL3; **Figure 1C**) and the Metal-B motif required to coordinate the Mg^{2+} ions facilitating the various types of catalytic chemistries (Sosunov et al., 2003); Me-B; **Figure 1D**). All these nanoarchaeal substitutions are spatially in close vicinity within the catalytic site of RNAP (**Figure 2**). Based on our current understanding of the structural basis of the nucleotide addition cycle, such substitutions

would be predicted to have a substantially deleterious effect on the catalytic function of the *ne*RNAP active site. In comparison, the RNAP of the archaeon *I. hospitalis* - the host to *N. equitans* - does not encode any of these unusual substitutions found in the *ne*RNAP (**Figure 1**), thus essentially ruling out that the substitutions are required to survive in a particular environment.

Recent large-scale sequencing efforts have demonstrated that similar unusual substitution patterns can be found in hundreds of sequence samples derived from fresh- and marine water sources (Probst et al., 2017; Tully et al., 2018; Zhou et al., 2020). Several data base entries, labeled as yet unnamed representatives of Woeearchaea or Pacearchaea, show the same types of substitutions as originally found in *N. equitans* (**Figure 1A**). Although the Bridge Helix and Trigger Loop are usually encoded by separate subunits of archaeal RNAPs (*A'* and *A''*, respectively) - and can therefore usually not be allocated to the same species in environmental sequencing samples - there are two pacearchaeal sequences where *A'* and *A''* appear to be fused into a single subunit (**Figure 1A**; directly comparable to the eukaryotic large RNAP subunits). We can therefore see from these examples, that - like in *N. equitans* - the specific substitution pattern in both Bridge Helix and Trigger Loop are encoded within an RNAP subunit from the same species. This suggests that, although for a long time considered unusual, *N. equitans* is actually a fairly typical representative of a larger group of archaea (including Pacearchaea, Woeearchaea etc.) that display comparable, but structurally diverged RNAP active site architectures.

The goal of this study was to determine whether the RNAP encoded by the *N. equitans* genome was 1) enzymatically active and 2) to what extent the substitution pattern resulted in altered catalytic properties.

MATERIALS AND METHODS

Identification of *ne*RNAP Subunits and Basal Transcription Factors

*ne*RNAP subunit-encoding open reading frames were identified using existing data base annotations and tblastn searches of the *Nanoarchaeum equitans* genome sequence (SeqID: AE017199.1; see **Supplementary Table S1** for more details).

Markov chain Monte Carlo Simulations

Markov Chain Monte Carlo MCMC simulations were carried out as described previously (Sullivan and Weinzierl, 2020). Briefly, the simulations employed the PROFASI forcefield in the PHAISTOS package (Boomsma et al., 2013). Due to the origin of the proteins from hyperthermophilic organisms, the simulation temperature was set to 355 K (81.85°C). The resulting trajectory data (based on 50,000 calculated states per simulation) was analyzed for secondary structure elements using cptraj (Roe and Cheatham, 2013) and processed/visualized with custom scripts on *Python* Jupyter notebooks.

In Vitro Reconstitution *ne*RNAP

The protein-coding portions of RNAP subunits *A'*, *A''*, *B'*, *B''*, *D*, *H*, *L*, *N* and *p* were PCR amplified from purified *N. equitans*

genomic DNA (a gift from Prof. M. Thomm, University of Regensburg) as full-length, non-tagged sequences and cloned as NdeI-BamHI (*neA''*, *neB'*, *neB''*, *neF*, *neK*), or NdeI-EcoRI (*neA'*, *neD*, *neE*, *neH*, *neL*, *neN*, *neP*) fragments into the bacterial expression vector pET21a. Recombinant proteins were expressed with IPTG-induction in *E. coli* BL21-DE3 Rosetta 2 (Merck) under standard conditions (Werner and Weinzierl, 2002). Subunits *neA'*, *neA''*, *neB'* and *neB''* were purified as insoluble inclusion bodies. Briefly, bacterial cells expressing these recombinant subunits were resuspended in T/G₀ (25 mM Tris-base, 200 mM glycine, 10 mM magnesium acetate, 100 μM zinc acetate, 14 mM β-mercaptoethanol and 10% glycerol at pH7.5) and sonicated. The inclusion bodies were washed extensively with 1 x deoxycholate buffer (1 mg/ml deoxycholate, 15 mM β-mercaptoethanol) and water/15 mM β-mercaptoethanol before solubilizing them in T/G₀ in the presence of saturating urea or 6 M guanidine-hydrochloride. Subunits *neD*, *neL*, *neH*, *neN* and *neP* were expressed similarly as soluble recombinant proteins. Bacterial cells expressing these recombinant subunits were resuspended in P300 Buffer (300 mM potassium acetate, 20 mM Tris-acetate pH 7.9, 10 mM magnesium acetate, 100 μM zinc acetate, 14 mM β-mercaptoethanol and 10% glycerol) and sonicated. The supernatant containing the solubilized proteins were heat-inactivated at 70°C for 10 min to precipitate the bacterial proteins present in the extract (the hyperthermophilic *ne* subunits remain completely soluble during this treatment).

The urea-solubilized inclusion bodies, or the soluble subunits, were passed over ~5 ml SP- or Q-Sepharose (Fast flow, Amersham) in chromatography columns. Proteins were eluted in a salt gradient from T/G₀ to T/G1000 using a DuoFlow BioLogic FPLC system (BioRad). The purified subunits were assembled by mixing them in the presence of 8 M urea in a dialysis cell (Slide-A-Lyzer 3500MCOW frames [Pierce], or 96-well microdialyser (SpectraPor) on a Theonix robotic platform (Aviso) for high-throughput assembly (Nottebaum et al., 2008; Weinzierl, 2013)), followed by lowering the urea concentration by gradual dilution in the dialysis buffer (Werner and Weinzierl, 2002; Naji et al., 2007; Nottebaum et al., 2008; Weinzierl, 2013). Equimolar amounts of the large subunits were mixed with small subunits, which were in at least four-fold excess to the large subunits, under denaturing conditions. The highest yield of enzymatically active *ne*RNAP (due to increased folding efficiency) was obtained in the presence of 500 mM salt (either sodium chloride, potassium- or sodium acetate) in the refolding buffer (Supplementary Figure S1). The assembly of large complexes was monitored by analytical size exclusion chromatography on Superose-6 and Superose-12 columns (Amersham) at a flow-rate of 0.5 ml/min. When required, soluble protein complexes were concentrated further using centrifugal YM-50 Centricon (Millipore) units according to the manufacturers instructions.

In Vitro Transcription Assay

Refolded RNAPs were assayed for transcriptional activity by measuring incorporation of *a*-³²P-UTP into RNA. Refolded RNAP was added to 1 x transcription buffer (1 x TB) containing 500 μM ATP, CTP, GTP, 1 μM UTP, 27 nM *a*-³²P-UTP (6000 Ci/mmol, Amersham), 1.5 μg nuclease-

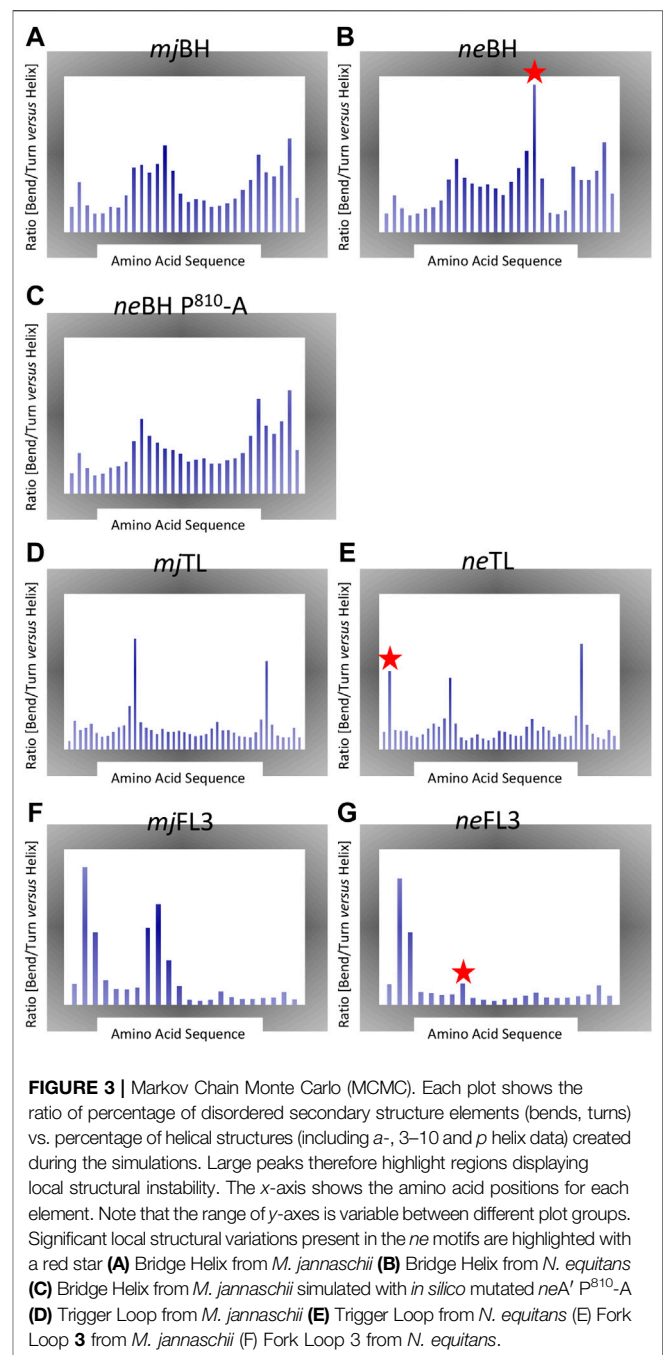


FIGURE 3 | Markov Chain Monte Carlo (MCMC). Each plot shows the ratio of percentage of disordered secondary structure elements (bends, turns) vs. percentage of helical structures (including *a*-, 3-10 and *p* helix data) created during the simulations. Large peaks therefore highlight regions displaying local structural instability. The x-axis shows the amino acid positions for each element. Note that the range of y-axis is variable between different plot groups. Significant local structural variations present in the *ne* motifs are highlighted with a red star (A) Bridge Helix from *M. jannaschii* (B) Bridge Helix from *N. equitans* (C) Bridge Helix from *M. jannaschii* simulated with *in silico* mutated *neA'* P⁸¹⁰-A (D) Trigger Loop from *M. jannaschii* (E) Trigger Loop from *N. equitans* (F) Fork Loop 3 from *M. jannaschii* (G) Fork Loop 3 from *N. equitans*.

activated calf thymus DNA (Fluka), 120 mM potassium acetate, 10 mM magnesium acetate, 10 mM Tris-acetate pH 7.5 and 10 mM DTT, which was incubated at 37–65°C for 45 min. The final reaction volume was 50 μL. The reactions were stopped by addition of 15% (w/v) trichloroacetic acid followed by 30 min incubation on ice. The precipitate was collected on 96-well GF/F glass fiber filter plates (Whatman), washed twice with excess 10% TCA, once with 95% ethanol, and quantitated in a scintillation counter in presence of scintillant fluid (Opti-fluor, Packard Bioscience). These steps were fully automated on a Theonix

liquid handling platform (Aviso) (Nottebaum et al., 2008; Weinzierl, 2013). Independent repetitions (“biological replicates”) of the same transcription reaction are reproducible within a 12% error margin. Transcripts originating solely from abortive initiation are not precipitated using this method. Therefore, only transcripts from elongation-competent RNAPs (longer than ~20 nucleotides) give rise to a signal in this assay.

High-Throughput “Sparse-Matrix” Sampling

Crystallization buffer sets ICL-1, -3, -4, and -5 (Hampton Research, Aliso Viejo, United States; **Supplementary Figures S2A–F**) were used as 10 x concentrates for high-throughput transcription assays based on nicked DNA templates as described previously (Werner and Weinzierl, 2002; Nottebaum et al., 2008; Tan et al., 2008; Boomsma et al., 2013). Briefly, the assay measures the incorporation of α - 32 P-UTP into acid-insoluble RNA by liquid scintillation counting in a robotically implemented workflow. Similar to the strategy used when employing such buffer sets for crystallization screens, the initial screen was only carried out with one assay per buffer set. Buffer sets that gave high levels of activity were subsequently tested in triplicate to confirm the result.

RESULTS

Computational Simulations

The structural consequences of several of the substitutions were determined by comparing secondary structure propensities of sequences from *N. equitans* to equivalent domains from *M. jannaschii* (*mj*). Both species are hyperthermophiles thus containing similar sequence-encoded features that stabilize their protein structures at elevated temperatures. Markov Chain Monte Carlo (MCMC) simulations is the method of choice for a systematic and comprehensive exploration of conformational space (Boomsma et al., 2013) (**Figure 3**). As expected, the presence of a proline in the *ne*BH causes a substantial disruption of a region that displays high α -helical propensity in the *mj*BH (Tan et al., 2008). On a structural level, the presence of proline in *neA'* in position 810 (*neA'* P⁸¹⁰) is predicted to cause a substantial destabilization of the α -helical conformation of the Bridge Helix in a slightly more N-terminal location (mostly affecting *neA'* R⁸⁰⁸; **Figure 3B**). Simulating a *ne* Bridge Helix with a “corrected” *in silico* point mutation (*neA'* P⁸¹⁰-A) restores the predicted conformational population to one that is very close to the *mj* Bridge Helix (**Figure 3C**). This proves that the unusual conformational properties are predominantly due to *neA'* P⁸¹⁰ position, rather than any of the other differences in the primary amino acid sequence. The region in the *ne* Bridge Helix most distorted corresponds to the orthologous region in *M. jannaschii* (*mjA'*-R⁸²⁰) which is a structure with one of the highest α -helical propensities of the entire domain (**Figure 3A**). High-throughput mutagenesis studies of *mjA'*-R⁸²⁰ in *mj*RNAP have shown it to be highly sensitive to point mutations, with only phenylalanine and tryptophane substitutions not resulting in substantial loss of catalytic activity (Tan et al., 2008; Weinzierl, 2013).

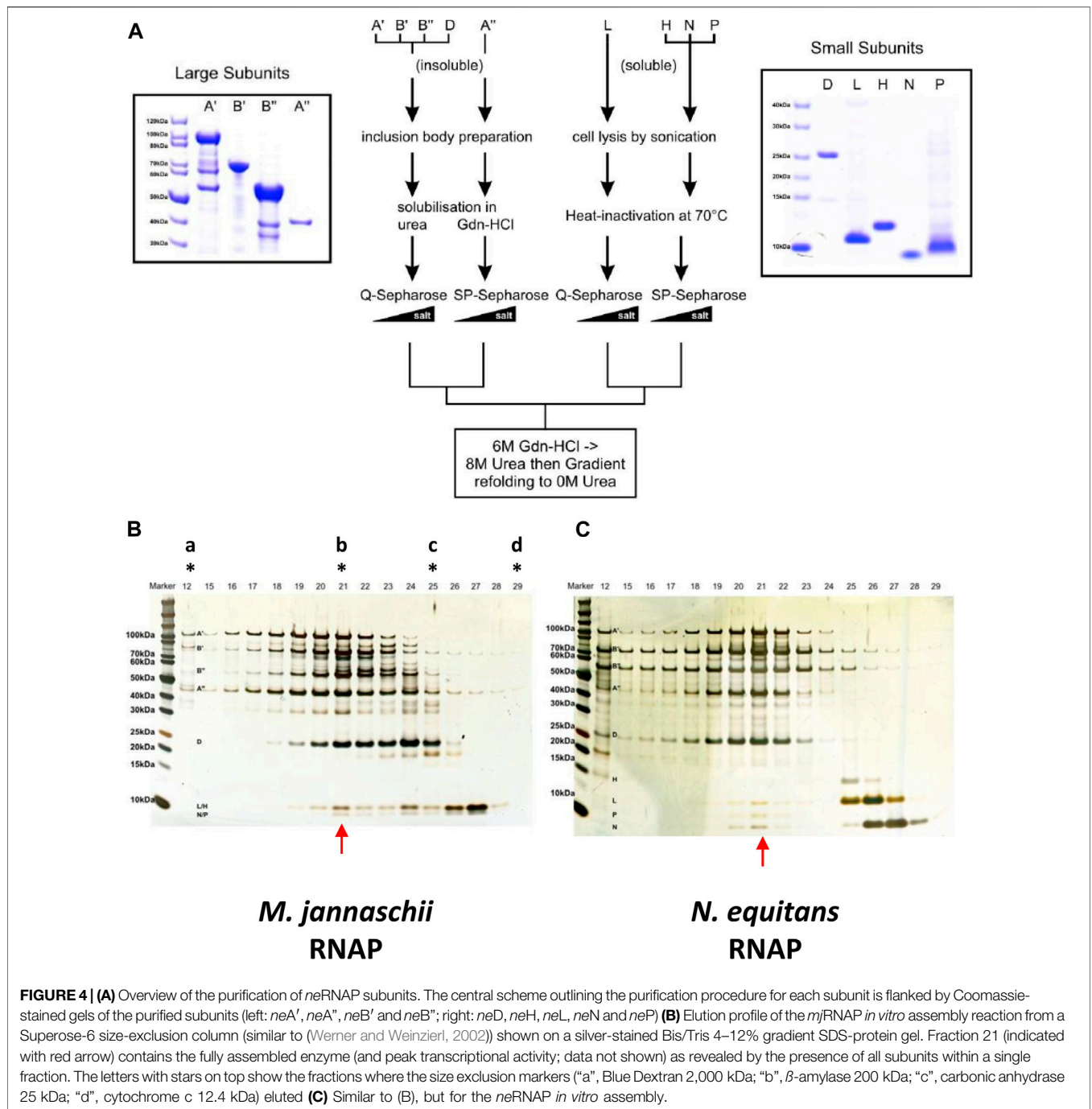
Similar comparisons of the *ne* and *mj* Trigger Loop conformations yield a less clear-cut result (**Figure 3D,E**), although the unusual position of a proline in the *ne* motif near the edges of the domain (*neA'*-P⁴⁷) again is likely to contribute a destabilizing influence (**Figure 3E**).

For *ne*FL3, the structural consequences of replacing highly conserved residues that are virtually invariant in other archaeal and eukaryotic polymerases in a non-conservative manner (for example, in FL3: C-D⁴⁴⁵, G-D⁴⁵², V/I-R⁴⁵⁸, and N-A⁴⁶⁰; **Figure 1C**) suggest that this would cause distinct changes in the functional contributions of these residues to catalysis. Especially one of these substitutions *neB'* G⁴⁵²-D is predicted to reduce the flexibility of the central region of *ne*FL3 considerably in comparison to the orthologous sequence of *mj*FL3 (**Figures 3F,G**). Similarly, Metal B contains two highly conserved acidic residues that coordinate of binding of the Mg²⁺ ion brought along by the incoming NTP, but in *N. equitans* one of them is converted to glutamine (*neA'* Q217) and thus is predicted bind the metal less strongly (**Figure 1D**).

Overall, based on previous insights from a range of structural and functional studies from archaeal and eukaryotic RNAPs representative of the majority of such organisms, a picture of a structurally diverged catalytic site in nanoarchaeal RNAPs emerges that suggests that the catalytic site may be more flexible in some areas (prolines in the *ne*Bridge Helix and *ne*Trigger Loop domains, stiffer in the diverged *ne*FL-3 domain and potentially compromised electrostatically by a diminished *ne*Metal-B motif).

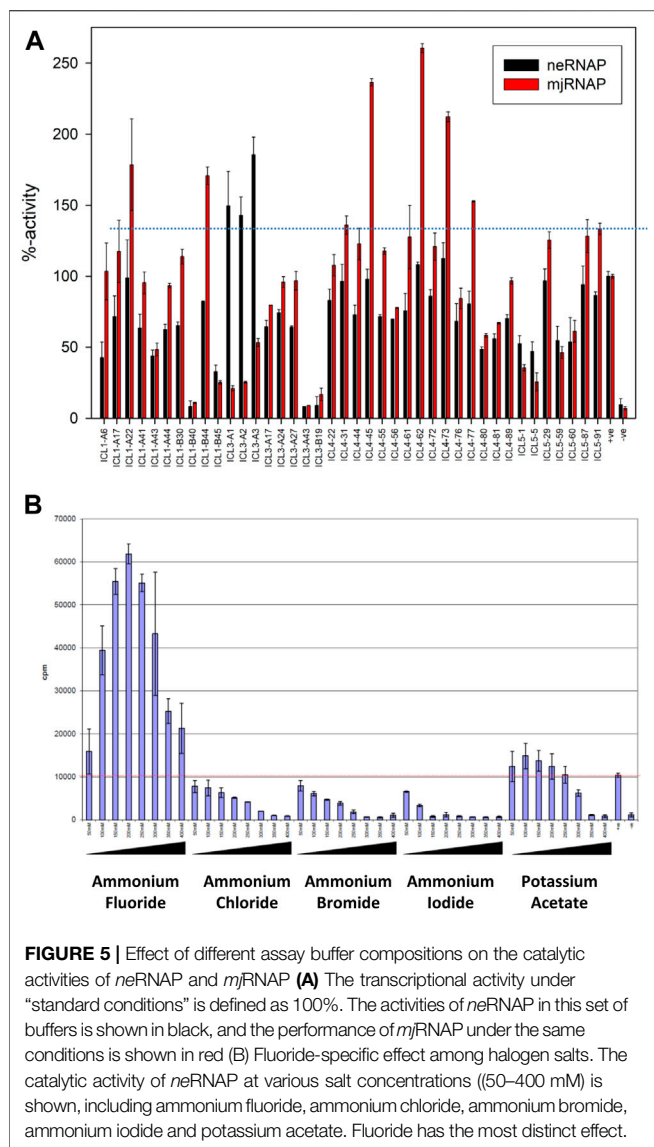
In Vitro Assembly of *ne*RNAP and High-Throughput “Sparse-Matrix” Sampling of *ne*RNAP Assay Conditions

The conformational distortions caused by potentially disruptive radical substitution suggest that the *Nanoarchaeum* RNAP may display only very low - or even no - catalytic activity. On the other hand, the presence of all known RNAP subunits in an otherwise minimal genome implies selective pressure responsible for maintaining an active transcriptional machinery. Technical problems with obtaining *N. equitans* in quantities sufficient for biochemical analysis preclude a direct purification of native enzymes from cells. We therefore decided to investigate this question by adopting the *in vitro* assembly approach that has been applied successfully for the assembly of RNAPs from other hyperthermophilic archaea (Werner and Weinzierl, 2002; Naji et al., 2007). The *in vitro* assembly of *ne*RNAP followed essentially the same procedure that we employed successfully in the past for *mj*RNAP (Werner and Weinzierl, 2002). Each of the subunits essential for catalytic activity was expressed as a recombinant protein in *E. coli*, followed by chromatographic purification and *in vitro* assembly by controlled dialysis from denaturing conditions (**Figure 4A**). Under these conditions, a portion of the *ne*RNAP subunits assembled into a complex that - comparable to *mj*RNAP (**Figure 4B**) - eluted as a distinct peak of activity during size exclusion chromatography (**Figure 4C**). As expected from its hyperthermophilic origin, the temperature optimum for catalytic activity was around



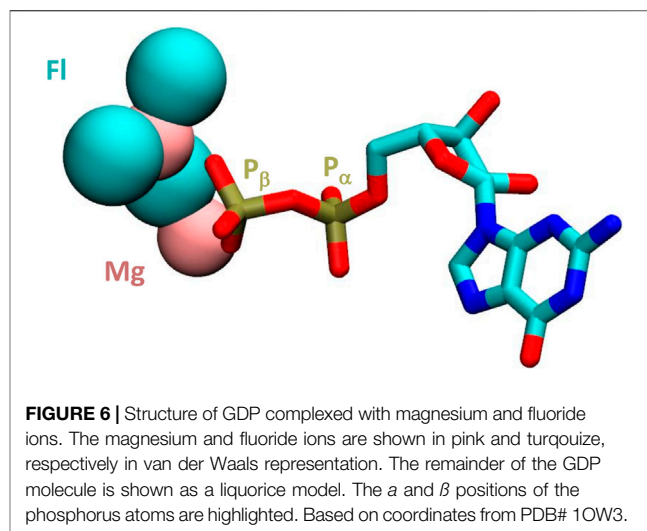
76°C (Supplementary Figures S3). Initial transcription experiments with *ne*RNAP suggested that the standard buffer conditions (120 mM potassium acetate, 10 mM magnesium acetate, 20 mM Tris-acetate, pH 8.6) were probably suboptimal because we observed a ~7-fold lower specific activity for *ne*RNAP as compared to *mj*RNAP when assembled in parallel. We therefore attempted to optimize the assay conditions over a wider range of pH values, salt concentrations and in the presence of various additives. The concept of “sparse-matrix”

sampling is well established in the macromolecular crystallization community where the method is used to identify the optimal (yet initially unknown) conditions to obtain macromolecular crystals for structural studies (Jancarik and Kim, 1991). Such approaches have also been employed usefully to identify optimal renaturation conditions (Hofmann et al., 1995), or for stabilizing macromolecular complexes (Chari et al., 2015). Here, we employed such a strategy to identify the best assay conditions for *ne*RNAP that included a wide range of different concentrations of



various cations and anions, buffers at different pHs, and the presence of a variety of detergents and stabilizing reagents. A series of buffer sets (ICL-1, ICL-3, ICL-4, and ICL-5; Hampton Research), comprised of 386 different cocktails (see **Supplementary Figures S2A–F** for composition), were used as 10 x stock solutions after supplementing them with Mg^{2+} and Zn^{2+} in automated high-throughput transcription assays. Most of the mixtures include an inorganic or organic salt, a buffering compound (with pH ranges from 4.5 to 9.5) and a “precipitant”, such as polyethylene glycol. In our assays, the precipitant may display stabilizing effects on protein structure - especially quaternary structures - under hyperthermophilic assay conditions.

A summary of the results (see **Supplementary Figure S4** for the complete data set) shows that *neRNAP* had a clear preference for a group of three buffers (ICL-3 #A1, A2, A3) that contained 20 mM sodium fluoride, potassium fluoride and ammonium fluoride, respectively (**Figure 5A**). This apparent preference for fluoride



is unique to *neRNAP* because *mjRNAP* only performed moderately (30–50% in comparison to standard conditions) in these buffers (**Figure 5A**). To test this potential requirement for fluoride further, *neRNAP* activity was assayed in transcription buffers containing varying amounts of fluoride salts. Optimal *neRNAP* stimulation was achieved with 200–300 mM potassium fluoride or ammonium fluoride (**Figure 5B** and **Supplementary Figure S5**). The stimulation of *neRNAP* activity by fluoride ions raised the question of whether other halogen ions (chloride, bromide, or iodide) would have a similar effect on *neRNAP*. This, however, was not the case, suggesting that the stimulating effect on the catalytic activity of *neRNAP* is indeed highly specific for fluoride.

DISCUSSION

Nanoarchaea are, in many ways, puzzling organisms. Their unique parasitic lifestyle has substantial effects on their cell- and genome size, which are both greatly minimized (Huber et al., 2002; Huber et al., 2003; Waters et al., 2003). Therefore, the cells depend on their host, *I. hospitalis* for many metabolites and precursors (Rawle et al., 2017). Analysis of the *N. equitans* genome has, however, revealed the presence of orthologs of all RNAP subunits and other components of the basal transcriptional machinery (TBP, TFB, and TFS; (Huber et al., 2002; Waters et al., 2003). It therefore looks as if *N. equitans* is capable of transcribing its own genome without help from its host cell in terms of imported basal transcription factors. Nevertheless, a number of key domains and motifs that constitute the active site of RNAP contains a distinct set of highly unusual and radical substitutions that appear to be deleterious to its catalytic activity.

Here we show, by *in vitro* assembly of nanoarchaeal RNAP from recombinant subunits expressed in and purified individually from *E. coli*, that the resulting enzyme displays catalytic activity. The temperature, pH optimum and specific activity are within the expected range of a hyperthermophilic organism and comparable to a similar enzyme assembled from *M. jannaschii* (Werner and

Weinzierl 2002). We therefore conclude that the changes in sequence, unusual as they may be, do not preclude catalytic activity. In a search for optimal assay conditions involving a sparse matrix approach, we discovered, however, an unexpected property: *ne*RNAP responded favourably to the presence of a high concentration of fluoride ions in the reaction buffer (optimal fluoride concentration for *ne*RNAP ~200–300 mM). Reports in the research literature from the 1970s describe a similar stimulatory effect of fluoride on adenylate cyclase (Drummond et al., 1971; Stalmans and Hers, 1975). These biochemical analyses showed that the reaction velocity (V_{\max}) of adenylate cyclase increased in the presence of fluoride but had no effect on the affinity (K_m) for substrate molecules. It later became apparent that it was a regulatory subunit that was the target of the fluoride stimulation, and not adenylate cyclase itself (Hebdon et al., 1978; Sahyoun et al., 1981). The identity of the regulatory protein turned out to be a subunit of a membrane bound, heterotrimeric G-protein complex. This G-protein is a gtpase and upon binding of GTP activates adenylate cyclase activity. The stimulatory effect of fluoride is believed to be the result of the ability of fluoride to form multi-fluorinated complexes with metal ions, such as Mg^{2+} (Antonny et al., 1993). Such “MgFx” complexes are capable of mimicking the γ -phosphate of a GTP molecule (Higashijima et al., 1991) and are thus able to occupy the phosphate binding pocket of the nucleotide-binding site of the G-protein. Several other G-protein dependent regulatory enzymes (such as Erk, Rho, Ras) have been shown to respond to fluoride in such a way (Bogatcheva et al., 2006). Fluoride has also been shown to bind to pyrophosphate (Baykov et al., 2000). We therefore hypothesize that the stimulation of transcription by high levels of fluoride ions may have a comparable cause in nanoarchaeal RNAP. The presence of mono- or multi-fluorinated NTP complexes (see **Figure 6** for a GDP-based example) may assist with binding of NTPs to a structurally more flexible active site in *ne*RNAP and/or help to stabilize some transition complexes in the nucleotide addition cycle. It is possible that especially the binding of Mg^{2+} ions to the divergent Metal B motif could be influenced in such a manner. According to such a model, the observed lack of effect of fluoride on the catalytic activity *mj*RNAP would reflect the fact that “conventional” RNAPs do not require this kind of assistance for their catalytic sites to operate.

REFERENCES

- Adam, P. S., Borrel, G., Brochier-Armanet, C., and Gribaldo, S. (2017). The Growing Tree of Archaea: New Perspectives on Their Diversity, Evolution and Ecology. *Isme J.* 11, 2407–2425. doi:10.1038/ismej.2017.122
- Antonny, B., Sukumar, M., Bigay, J., Chabre, M., and Higashijima, T. (1993). The Mechanism of Aluminum-independent G-Protein Activation by Fluoride and Magnesium. 31P NMR Spectroscopy and Fluorescence Kinetic Studies. *J. Biol. Chem.* 268, 2393–2402. doi:10.1016/s0021-9258(18)53789-3
- Baykov, A. A., Fabrichniy, I. P., Pohjanjoki, P., Zyryanov, A. B., and Lahti, R. (2000). Fluoride Effects along the Reaction Pathway of Pyrophosphatase: Evidence for a Second Enzyme-Pyrophosphate Intermediate†. *Biochemistry* 39, 11939–11947. doi:10.1021/bi000627u
- Blombach, F., Matelska, D., Fouqueau, T., Cackett, G., and Werner, F. (2019). Key Concepts and Challenges in Archaeal Transcription. *J. Mol. Biol.* 431, 4184–4201. doi:10.1016/j.jmb.2019.06.020

Future studies will focus on the potential interplay between fluoride, Mg^{2+} and NTPs, as well as defining in more detail which of the diverged motifs is most susceptible to this effect. By replacing some of the substitutions - either individually or in groups - with residues that are normally found in their position in other RNAPs, we will be able to study which of them are most likely to be responsible for this unusual behavior of *ne*RNAP in presence of fluoride.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

RW designed the project. SN carried out the experimental work and analyzed the data. The manuscript is a joint effort of SN and RW.

FUNDING

The project was funded by the Wellcome Trust (Grant 078043/Z/05/Z).

ACKNOWLEDGMENTS

We would like to thank Prof. Michael Thomm for providing us a sample of purified *N. equitans* genomic DNA.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2021.669314/full#supplementary-material>

- Bogatcheva, N. V., Wang, P., Birukova, A. A., Verin, A. D., and Garcia, J. G. N. (2006). Mechanism of Fluoride-Induced MAP Kinase Activation in Pulmonary Artery Endothelial Cells. *Am. J. Physiol.-Lung Cell Mol. Physiol.* 290, L1139–L1145. doi:10.1152/ajplung.00161.2005
- Boomsma, W., Frellsen, J., Harder, T., Bottaro, S., Johansson, K. E., Tian, P., et al. (2013). PHAISTOS: A Framework for Markov Chain Monte Carlo Simulation and Inference of Protein Structure. *J. Comput. Chem.* 34, 1697–1705. doi:10.1002/jcc.23292
- Brochier, C., Gribaldo, S., Zivanovic, Y., Confalonieri, F., and Forterre, P. (2005). ‘Nanoarchaea: Representatives of a Novel Archaeal Phylum or a Fast-Evolving Euryarchaeal Lineage Related to Thermococcales? *Genome Biol.* 6. doi:10.1186/gb-2005-6-5-r42
- 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. doi:10.1038/nmeth.3493
- Ciccarelli, F. D., Doerks, T., von Mering, C., Creevey, C. J., Snel, B., and Bork, P. (2006). Toward Automatic Reconstruction of a Highly Resolved Tree of Life. *Science* 311, 1283–1287. doi:10.1126/science.1123061

- Coker, J. A. (2019). Recent Advances in Understanding Extremophiles. *FI000Res* 8, F1000 Faculty Rev-1917. doi:10.12688/f1000research.20765.1
- Cramer, P., Bushnell, D. A., and Kornberg, R. D. (2001). Structural Basis of Transcription: RNA Polymerase II at 2.8 Angstrom Resolution. *Science* 292, 1863–1876. doi:10.1126/science.1059493
- Drummond, G. I., Severson, D. L., and Duncan, L. (1971). Adenyl Cyclase. Kinetic Properties and Nature of Fluoride and Hormone Stimulation. *J. Biol. Chem.* 246, 4166–4173. doi:10.1016/s0021-9258(18)62068-x
- Forster, P., Gribaldo, S., and Brochier-Armanet, C. (2009). Happy Together: Genomic Insights into the Unique Nanoarchaeum/Ignicoccus Association. *J. Biol.* 8, 7. doi:10.1186/jbiol110
- Fouqueau, T., Blombach, F., Cackett, G., Carty, A. E., Matelska, D. M., Ofer, S., et al. (2018). The Cutting Edge of Archaeal Transcription. *Emerg. Top. Life Sci.* 2, 517–533. doi:10.1042/etls20180014
- Hebdon, M., Le Vine, H., 3rd, Sahyoun, N., Schmitges, C. J., and Cuatrecasas, P. (1978). Properties of the Interaction of Fluoride- and Guanylyl-5'-Imidodiphosphate-Regulatory Proteins with Adenylate Cyclase. *Proc. Natl. Acad. Sci.* 75, 3693–3697. doi:10.1073/pnas.75.8.3693
- Higashijima, T., Graziano, M. P., Suga, H., Kainosho, M., and Gilman, A. G. (1991). 19F and 31P NMR Spectroscopy of G Protein Alpha Subunits. Mechanism of Activation by Al3+ and F-. *J. Biol. Chem.* 266, 3396–3401. doi:10.1016/s0021-9258(19)67806-3
- Hirata, A., Klein, B. J., and Murakami, K. S. (2008). Erratum: The X-Ray Crystal Structure of RNA Polymerase from Archaea. *Nature* 452, 248. doi:10.1038/nature06844
- Hofmann, A., Tai, M., Wong, W., and Glabe, C. G. (1995). A Sparse Matrix Screen to Establish Initial Conditions for Protein Renaturation. *Anal. Biochem.* 230, 8–15. doi:10.1006/abio.1995.1429
- Hohn, M. J., Hedlund, B. P., and Huber, H. (2002). Detection of 16S rDNA Sequences Representing the Novel Phylum "Nanoarchaeota": Indication for a Wide Distribution in High Temperature Biotopes. *Syst. Appl. Microbiol.* 25, 551–554. doi:10.1078/07232020260517698
- Huber, H., Hohn, M. J., Rachel, R., Fuchs, T., Wimmer, V. C., and Stetter, K. O. (2002). A New Phylum of Archaea Represented by a Nanosized Hyperthermophilic Symbiont. *Nature* 417, 63–67. doi:10.1038/417063a
- Huber, H., Hohn, M. J., Stetter, K. O., and Rachel, R. (2003). The Phylum Nanoarchaeota: Present Knowledge and Future Perspectives of a Unique Form of Life. *Res. Microbiol.* 154, 165–171. doi:10.1016/s0923-2508(03)00035-4
- Jancarik, J., and Kim, S. H. (1991). Sparse Matrix Sampling: a Screening Method for Crystallization of Proteins. *J. Appl. Cryst.* 24, 409–411. doi:10.1107/s0021889891004430
- McCliment, E. A., Voglesonger, K. M., O'Day, P. A., Dunn, E. E., Holloway, J. R., and Cary, S. C. (2006). Colonization of Nascent, Deep-Sea Hydrothermal Vents by a Novel Archaeal and Nanoarchaeal Assemblage. *Environ. Microbiol.* 8, 114–125. doi:10.1111/j.1462-2920.2005.00874.x
- Naji, S., Grünberg, S., and Thomm, M. (2007). The RPB7 Orthologue E' Is Required for Transcriptional Activity of a Reconstituted Archaeal Core Enzyme at Low Temperatures and Stimulates Open Complex Formation. *J. Biol. Chem.* 282, 11047–11057. doi:10.1074/jbc.m611674200
- Nottebaum, S., Tan, L., Trzaska, D., Carney, H. C., and Weinzierl, R. O. J. (2008). The RNA Polymerase Factory: a Robotic In Vitro Assembly Platform for High-Throughput Production of Recombinant Protein Complexes. *Nucleic Acids Res.* 36, 245–252. doi:10.1093/nar/gkm1044
- Ouhammouch, M., Werner, F., Weinzierl, R. O. J., and Geiduschek, E. P. (2004). A Fully Recombinant System for Activator-dependent Archaeal Transcription. *J. Biol. Chem.* 279, 51719–51721. doi:10.1074/jbc.c400446200
- Probst, A. J., Castelle, C. J., Singh, A., Brown, C. T., Anantharaman, K., Sharon, I., et al. (2017). Genomic Resolution of a Cold Subsurface Aquifer Community Provides Metabolic Insights for Novel Microbes Adapted to High CO2 concentrations. *Environ. Microbiol.* 19, 459–474. doi:10.1111/1462-2920.13362
- Randau, L., Münch, R., Hohn, M. J., Jahn, D., and Söll, D. (2005). Nanoarchaeum Equitans Creates Functional tRNAs from Separate Genes for Their 5'- and 3'-halves. *Nature* 433, 537–541. doi:10.1038/nature03233
- Randau, L., Schröder, I., and Söll, D. (2008). Life without RNase P. *Nature* 453, 120–123. doi:10.1038/nature06833
- Rawle, R. A., Hamerly, T., Triplet, B. P., Giannone, R. J., Wurch, L., Hettich, R. L., et al. (2017). Multi-omics Analysis Provides Insight to the Ignicoccus Hospitalis-Nanoarchaeum Equitans Association. *Biochim. Biophys. Acta (Bba) - Gen. Subjects* 1861, 2218–2227. doi:10.1016/j.bbagen.2017.06.001
- Roe, D. R., and Cheatham, T. E. (2013). PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theor. Comput.* 9, 3084–3095. doi:10.1021/ct400341p
- Sahyoun, N. E., LeVine, H., 3rd, Davis, J., Hebdon, G. M., and Cuatrecasas, P. (1981). Molecular Complexes Involved in the Regulation of Adenylate Cyclase. *Proc. Natl. Acad. Sci.* 78, 6158–6162. doi:10.1073/pnas.78.10.6158
- Sosunov, V., Sosunova, E., Mustaev, A., Bass, I., Nikiforov, V., and Goldfarb, A. (2003). Unified Two-Metal Mechanism of RNA Synthesis and Degradation by RNA Polymerase. *EMBO J.* 22, 2234–2244. doi:10.1093/emboj/cdg193
- Stalmans, W., and Hers, H.-G. (1975). The Stimulation of Liver Phosphorylase B by AMP, Fluoride and Sulfate. A Technical Note on the Specific Determination of the a and B Forms of Liver Glycogen Phosphorylase. *Eur. J. Biochem.* 54, 341–350. doi:10.1111/j.1432-1033.1975.tb04144.x
- Sullivan, S. S., and Weinzierl, R. O. J. (2020). Optimization of Molecular Dynamics Simulations of C-Myc(1-88)-An Intrinsically Disordered System. *Life (Basel)* 10 (7), 109. doi:10.3390/life10070109
- Tan, L., Wiesler, S., Trzaska, D., Carney, H. C., and Weinzierl, R. O. (2008). Bridge Helix and Trigger Loop Perturbations Generate Superactive RNA Polymerases. *J. Biol.* 7, 40. doi:10.1186/jbiol98
- Thomm, M., Reich, C., Grünberg, S., and Naji, S. (2009). Mutational Studies of Archaeal RNA Polymerase and Analysis of Hybrid RNA Polymerases. *Biochem. Soc. Trans.* 37, 18–22. doi:10.1042/bst0370018
- Tully, B. J., Graham, E. D., and Heidelberg, J. F. (2018). The Reconstruction of 2,631 Draft Metagenome-Assembled Genomes from the Global Oceans. *Sci. Data* 5, 170203. doi:10.1038/sdata.2017.203
- Waters, E., Hohn, M. J., Ahel, I., Graham, D. E., Adams, M. D., Barnstead, M., et al. (2003). The Genome of Nanoarchaeum Equitans: Insights into Early Archaeal Evolution and Derived Parasitism. *Proc. Natl. Acad. Sci.* 100, 12984–12988. doi:10.1073/pnas.1735403100
- Weinzierl, R. O. J. (2010a). Nanomechanical Constraints Acting on the Catalytic Site of Cellular RNA Polymerases. *Biochem. Soc. Trans.* 38, 428–432. doi:10.1042/bst0380428
- Weinzierl, R. O. J. (2010b). 'The Nucleotide Addition Cycle of RNA Polymerase Is Controlled by Two Molecular Hinges in the Bridge Helix Domain'. *BMC Biol.* 8, 134. doi:10.1186/1741-7007-8-134
- Weinzierl, R. O. J. (2011). The Bridge Helix of RNA Polymerase Acts as a Central Nanomechanical Switchboard for Coordinating Catalysis and Substrate Movement. *Archaea* 2011, 608385. doi:10.1155/2011/608385
- Weinzierl, R. O. J. (2013). The RNA Polymerase Factory and Archaeal Transcription. *Chem. Rev.* 113, 8350–8376. doi:10.1021/cr400148k
- Wenck, B. R., and Santangelo, T. J. (2020). Archaeal Transcription. *Transcription* 11, 199–210. doi:10.1080/21541264.2020.1838865
- Werner, F., and Weinzierl, R. O. J. (2002). A Recombinant RNA Polymerase II-like Enzyme Capable of Promoter-specific Transcription. *Mol. Cell* 10, 635–646. doi:10.1016/s1097-2765(02)00629-9
- Werner, F., Wiesler, S., Nottebaum, S., and Weinzierl, R. O. J. (2006). Modulation of RNA Polymerase Core Functions by Basal Transcription Factor TFB/TFIIB. *Biochem. Soc. Symp.* 73, 49–58. doi:10.1042/bss0730049
- Zhou, Z., Liu, Y., Xu, W., Pan, J., Luo, Z. H., and Li, M. (2020). Genome- and Community-Level Interaction Insights into Carbon Utilization and Element Cycling Functions of Hydrothermarchaeota in Hydrothermal Sediment. *mSystems* 5, e00795–19. doi:10.1128/mSystems.00795-19

Conflict of Interest: SN is currently employed by Orthomol Pharmazeutische Vertriebs GmbH.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Nottebaum and Weinzierl. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.