Helicase-like functions in phosphate loop containing beta-alpha polypeptides

Pratik Vyas[a], Olena Trofimyuk[a], Liam M. Longo[b,c], Fanindra Kumar Deshmukh[a], Michal Sharon[a], and Dan S. Tawfik[a,1]

[a]Department of Biomolecular Sciences, Weizmann Institute of Science, 7610001 Rehovot, Israel; [b]Tokyo Institute of Technology, Earth-Life Science Institute, Tokyo 152-8550, Japan; and [c]Blue Marble Space Institute of Science, Seattle, WA 98104

Edited by James M. Berger, Johns Hopkins Medical Institute, Baltimore, MD, and approved February 27, 2021 (received for review July 31, 2020)

The P-loop Walker A motif underlies hundreds of essential enzyme families that bind nucleotide triphosphates (NTPs) and mediate phosphoryl transfer (P-loop NTPases), including the earliest DNA/RNA helicases, translocases, and recombinases. What were the primordial precursors of these enzymes? Could these large and complex proteins emerge from simple polypeptides? Previously, we showed that P-loops embedded in simple βα repeat proteins bind NTPs but also, unexpectedly so, ssDNA and RNA. Here, we extend beyond the purely biophysical function of ligand binding to demonstrate rudimentary helicase-like activities. We further constructed simple 40-residue polypeptides comprising just one β-(P-loop)-α element. Despite their simplicity, these P-loop prototypes confer functions such as strand separation and exchange. Foremost, these polypeptides unwind dsDNA, and upon addition of NTPs, or inorganic polyphosphates, release the bound ssDNA strands to allow reformation of dsDNA. Binding kinetics and low-resolution structural analyses indicate that activity is mediated by oligomeric forms spanning from dimers to high-order assemblies. The latter are reminiscent of extant P-loop recombinases such as RecA. Overall, these P-loop prototypes compose a plausible description of the sequence, structure, and function of the earliest P-loop NTPases. They also indicate that multifunctionality and dynamic assembly were key in endowing short polypeptides with elaborate, evolutionarily relevant functions.

protein evolution | P-loop | Walker A | multifunctionality | polyphosphate

Protein machines, such as ATP (adenosine triphosphate) synthetases, helicases, or RecA recombinases, perform remarkable actions and are immensely complex (1, 2). Among other factors, complexity is manifested in the concerted action of multiple functional elements, the productive orchestration of which demands large, well-defined structures. This complexity is, however, at odds with the assumption that proteins emerged by duplication and fusion of relatively short and simple polypeptides (3, 4). Albeit, these polypeptides predated the last universal common ancestor (LUCA) and as such cannot be reconstructed by conventional phylogenetic methods nonetheless; despite nearly 4 billion years of evolution and extensive diversification of sequence, structure, and function, these ancestral polypeptides are traceable by virtue of being the most conserved and functionally essential motifs of modern-day proteins (5). Perhaps the most widely spread motif is the Walker-A P-loop (6). Defined as GxxxxGK(T/S) or GxxGxGK, this motif underlies the most abundant and diverse protein class: the P-loop NTPases (6–9). Structurally, the P-loop NTPase domain comprises a tandem repeat of at least five β-(loop)-α elements arranged in a αβα three-layer sandwich architecture (7, 8, 10). The loops connecting the C termini of the β strands to the *N* termini of the following α helices compose the active site, while short loops link the β-(loop)-α elements to one another (here, unless stated otherwise, loops refer to the former). The Walker A P-loop is the key element mediating NTP binding and catalysis, and it uniformly resides within the first β-(loop)-α element, henceforth designated as β1-(P-loop)-α1.

P-loop NTPases are among the most ancient enzyme families—if not the most ancient enzyme family itself—and the first P-loop NTPase likely emerged at the transition from the RNA world to the primordial RNA-protein world (5, 11–13). Binding of phospho-ligands, foremost ATP and GTP (guanosine triphosphate), is the founding function of P-loop NTPases (14). In agreement with early emergence, the key to phospho-ligand binding is not only the P-loop, but also a "crown" of hydrogen bonds realized by the end of the Walker A motif (GK(T/S)) that comprises the first turn of α1 (r 14, 15). It has been accordingly hypothesized that polypeptides composing the P-loop and its flanking secondary structural elements, namely, a β-(P-loop)-α segment, were the seed from which modern P-loop NTPases emerged (5, 16–18). However, that relatively short polypeptides can confer an evolutionarily meaningful function is far from obvious. Polypeptides generally lack structural volume and complexity, and cannot, in principle, align multiple functional elements as do intact protein domains (19). Hence, identifying functional polypeptides, and especially polypeptides that are assigned as the seeding elements of the earliest enzyme families, is of paramount importance to our understanding of protein evolution.

Previously, we showed that an ancestral β1-(P-loop)-α1 motif, inferred via phylogenetic analysis of all known P-loop NTPase families, could be grafted onto a simple repeat scaffold composed of four consecutive βα elements (16). This grafting resulted in simple proteins that contained two P-loops. These bound via the

Significance

It is widely assumed that today's large and complex proteins emerged from much shorter and simpler polypeptides. Yet the nature of these early precursors remains enigmatic. We describe polypeptides that contain one of the earliest protein motifs, a phosphate-binding loop, or P-loop, embedded in a single beta-alpha element. These simple P-loop polypeptides show intriguing characteristics of a primordial world composed of nucleic acids and peptides. They are "generalists" capable of binding different phospho-ligands, including inorganic polyphosphates and single-stranded DNA. Nonetheless, in promoting double-stranded DNA unwinding and strand-exchange, they resemble modern P-loop helicases and recombinases. Our study describes a missing link in the evolution of complex proteins—prototypes that tangibly relate to contemporary P-loop enzymes in sequence, structure, and function.

Author contributions: P.V. and D.S.T. designed research; P.V., O.T., and F.K.D. performed research; P.V., L.M.L., and M.S. contributed new reagents/analytic tools; P.V., L.M.L., F.K.D., M.S., and D.S.T. analyzed data; and P.V. and D.S.T. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: dan.tawfik@weizmann.ac.il.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2016131118/-/DCSupplemental.

Published April 12, 2021.

P-loop motif ATP and GTP, but rather unexpectedly, also RNA and single-stranded DNA (ssDNA). It is therefore likely that the primordial P-loop was a multifunctional phospho-ligand binder that could function in the absence of other auxiliary residues (16). In the contemporary P-loop NTPases, ssDNA binding is mostly conferred by specialized domains (20–23) that are distinct from the ATP binding P-loop (9). And yet, in support of the P-loop's multifunctionality, we searched for and revealed examples of ssDNA binding mediated by the Walker A P-loop (detailed in *Discussion*). Further, many of the P-loop NTPases families that date back to LUCA are involved in RNA/DNA remodeling, including helicases, RecA like recombinases, and translocases (24). ATP synthetase, for example, is thought to have emerged from such nucleic acid–remodeling P-loop NTPases (25).

That ssDNA binding can be mediated by our P-loop prototypes, and the possible origins of P-loop NTPases as nucleic acid remodelers, guided us to extend our previous study (16) beyond the realm of ligand binding per se. We now i) examine the potential of P-loop prototypes to exert biologically and evolutionarily relevant helicase-like functions, i.e., to mediate unwinding of double-strand DNA (dsDNA) and strand exchanges, and ii) identify the minimal structural context in which the P-loop can confer these relevant functions.

Results

P-Loop Prototypes Mediate Strand Separation. dsDNA undergoes spontaneous local fluctuations referred to as "DNA breathing," (26, 27) thus giving access to ssDNA binding proteins (28). For instance, T4 helicases bind to and stabilize these transient open regions to unwind dsDNA (27)—a mode of unwinding also referred to as "passive unwinding" (29). To test the potential of P-loop prototypes to facilitate such passive unwinding, we used a standard, fluorescence-based, molecular beacon helicase assay (30) (Fig. 1*A*). This assay uses a dsDNA segment composed of a "beacon oligo" with a fluorophore at the 5′ end and a quencher at the 3′ end (the beacon sense strand) hybridized to an "unlabeled" complementary antisense strand. Upon addition of a protein that promotes strand separation, owing to its complementary ends, the beacon strand collapses into an intramolecular hairpin (stem-loop structure). This brings together the fluorophore and the quencher, thus quenching the beacon's fluorescent signal. As shown later, this stem-loop product is likely stabilized by the fluorophore-quencher interaction (31), thus providing a simple way of monitoring DNA unwinding (30).

For the initial round of experiments, the originally described 110 amino acid proteins that present two P-loops [D-PLoop (16)], hereafter referred to as the "intact prototype," was used. This intact prototype bound ssDNA preferably over dsDNA (16). As shown later, the P-loop prototypes also exhibit preference for TC-rich ssDNA while GA-rich sequences are barely bound. Accordingly, the dsDNA beacon comprised a TC-rich unlabeled antisense strand to promote binding and strand separation, while the beacon sense strand was GA-rich to avoid interference with the stem-loop formation and nonspecific quenching due to protein binding (oligonucleotide sequences are provided in *SI Appendix*, Table S1). When added to the beacon dsDNA, the intact prototype induced strand separation as indicated by the decrease in fluorescence (Fig. 1*B*). The P-loop prototype did not induce quenching of the fluorophore itself as indicated by a negligible change in fluorescence intensity with digested beacon dsDNA (Fig. 1*B*; nonetheless, this nonspecific quenching was subtracted to give the quenching traces reported below; see *Materials and Methods*). Strand separation, as reported by this beacon dsDNA, was applied for testing various fragments of the intact P-loop prototype as described in the next section. Subsequently, using the most active fragment, the kinetics and mechanism of strand separation were further investigated as described below (see The N-*αβα* P-loop Prototype Shows Avid and Cooperative Strand Separation).

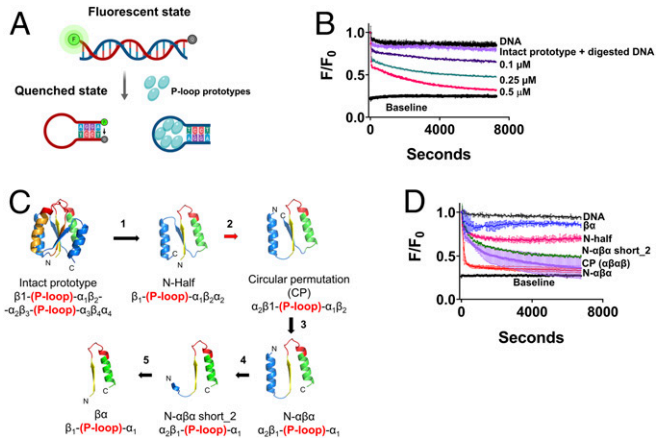

Fig. 1. The molecular beacon assay reports the strand separation of P-loop prototypes. (*A*) Simplified schematic of the strand-separation molecular beacon assay (30). In the initial dsDNA state, the quencher of the beacon strand is held apart from the fluorophore resulting in energy transfer (high fluorescence). The preferential binding of P-loop prototypes to ssDNA induces strand separation, allowing the beacon strand to assume a hairpin state wherein the fluorophore is quenched. The oligos used in this assay and the assays described in the subsequent figures are listed in *SI Appendix*, Table S1. (*B*) A representative strand separation experiment using the intact 110 residue P-loop prototype (D-P-loop; ref. 16). Strand separation is reported by the change in FRET (Förster Resonance Energy Transfer) signal (i.e., fluorescence quenching in our experimental setup) upon addition of the P-loop prototype at increasing protein concentrations. All assays were performed with 5 nM beacon dsDNA, in 50 mM Tris (pH 8), at 24 °C. Shown are normalized F/F_0 values, whereby the initial fluorescence of beacon dsDNA prior to protein addition takes the value of 1. Baseline represents the signal of the fully quenched hairpin beacon. Digested DNA represents the signal upon addition of 0.5 μM P-loop prototype to the beacon dsDNA pretreated with Benzonase nuclease. Traces were fitted to a biphasic exponential decay model (*SI Appendix*) and the apparent rate constants are given in *SI Appendix*, Table S4. (*C*) New P-loop prototypes were constructed by systematic truncation and circular permutation of the intact prototype. The numbered arrows indicate sequential steps in the engineering of new constructs as follows: 1) Truncation of the intact 110 residue prototype into half (16) (*N*-half indicates N-terminal half). 2) Circular permutation (red arrow) of *N*-half to a construct with 'αβαβ' architecture. 3) Truncation of the C-terminal β-strand to give *N*-αβα. 4 and 5) Incremental truncations of N-terminal helix of *N*-αβα down to a βα fragment. The structural models indicate the ancestral P-loop element in yellow (β1), red (the Walker A P-loop) and green (α1), while the remaining parts are in blue. (*D*) Strand separation by truncated P-loop prototypes, at 1 μM protein concentration, and under the stringent condition (with 100 mM NaCl; other assay conditions as in *A*). The lines represent the average from two to six independent experiments, and the error bars represent the SD values.

Structural Minimization of the P-Loop Prototype. The intact P-loop prototypes were based on an "ideal fold" (32) that reproduces the three-layered αβα sandwich architecture of P-loop NTPases (7, 8, 10, 24). They are composed of four βα elements with the ancestral β-(P-loop)-α element replacing the first and third elements (16). We asked whether the intact prototype could be truncated while retaining its function. In other words, what is the minimal "stand-alone" P-loop fragment that possesses phospho-ligand binding and strand-separation activity?

The repeat nature of our prototypes allowed multiple options for single P-loop constructs with varied structural topologies. In the first step, following our original report, the first half of the intact P-loop prototype (*N*-half) was tested (16). Other "half" constructs comprising two strands and two helices were also tested

(Fig. 1*C* and *SI Appendix*, Fig. S1; DNA and amino acids sequences of all constructs are provided in *SI Appendix*, Tables S2 and S3). Secondly, we used circular permutation to generate half fragments that possessed an αβαβ architecture (Fig. 1*C*). All of these constructs, and especially the circular permutants, showed high expression and purification yields (via a His-tag and Ni-NTA affinity chromatography; *SI Appendix*, Fig. S2*A*). Nonetheless, to maintain solubility at high protein concentrations, an osmolyte such as L-arginine was needed. We then truncated the circularly permuted "half" constructs further to obtain the minimal constructs described below (Fig. 1*C* and *SI Appendix*, Fig. S1).

All "half" constructs, and most of their truncated versions, mediated strand separation to some degree (*SI Appendix*, Fig. S2*B*). However, nucleic acid binding is known to be highly sensitive to ionic strength, and indeed the activity of most constructs was diminished in the presence of 100 mM NaCl (Fig. 1*D* and *SI Appendix*, Fig. S2*C*). We therefore set our assay conditions at 100 mM NaCl as a benchmark for robust activity. Intriguingly, under this stringent binding condition, a truncated variant dubbed "*N*-αβα," comprising the ancestral β-(P-loop)-α element with one additional helix that precedes it, showed the most efficient and rapid strand separation (Fig. 1*D*, red trace).

The Minimal P-Loop Fragment. Can the *N*-αβα construct be further shortened? Given its polar nature, the primary role of N-terminal helix might be to enhance protein solubility (Fig. 1*C*, steps 4 through 6, *SI Appendix*, Table S3). Indeed, partial truncations of this helix showed very similar activity to the intact *N*-αβα (*N*-αβα short_2, Fig. 1*D*, green trace), while its complete truncation resulted in loss of the strand-separation activity (βα, Fig. 1*D*, blue trace). Nonetheless, this shortest construct, which in effect comprises a β-(P-loop)-α fragment, tended to copurify with nucleic acids, suggesting that it retained some binding capability. The functional prototype *N*-αβα short_2 contains only a short N-terminal segment (KRRGV) linked to the β-(P-loop)-α fragment by a short linker (GSG; *SI Appendix*, Table S3). It therefore appears that given a charged solubility tag, a single β-(P-loop)-α fragment can confer strand-separation activity. Also, worth noting is that *N*-αβα possesses low sequence complexity—it is made of only 12 amino acid types. Ten of these are considered abiotic amino acids, plus two cationic amino acids, Lys and Arg (excluding the 6xHistag for purification, and a tryptophan introduced for determining concentration by absorbance at 280 nm; *SI Appendix*, Table S3).

Overall, we found that the core β-(P-loop)-α motif can be placed in a variety of structural contexts, with varying strand topologies, and yet retain or even show improved biochemical function (Fig. 1*D* and *SI Appendix*, Figs. S1 and S2). Given its minimal size and sequence complexity, its high expression, purity, and foremost its high strand-separation activity, the *N*-αβα construct was employed as the representative P-loop prototype for the studies described below.

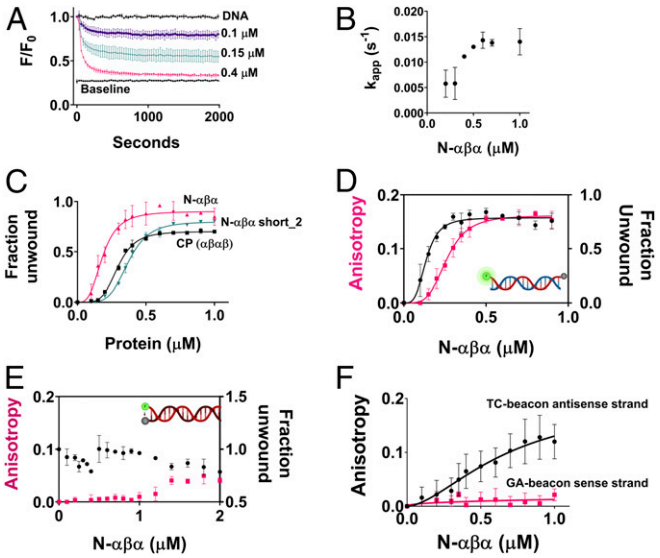

Fig. 2. Binding affinities and kinetics of strand separation by P-loop prototypes. (*A*) A representative strand-separation experiment with varying concentrations of *N*-αβα (the beacon dsDNA is described in Fig. 1*A*, and assay conditions are as in Fig. 1*D*). The values shown are the average from two to six independent experiments, and the error bars represent the SD values (the number of experiments in the subsequent panels is denoted as *n*). (*B*) Traces shown in *A* were fitted to a one phase exponential decay model (*SI Appendix*, Eq. 1), and the apparent rate constants were plotted against protein concentration (*SI Appendix*, Table S5). (*C*) Binding isotherms of P-loop prototypes (for their topology, see Fig. 1*C*). End point F/F_0 values (2 h), from strand separation experiments (e.g., *A*) were normalized (fluorescence of the starting beacon dsDNA equals 0, and of the fully quenched ssDNA hairpin equals 1) to derive the relative fraction of unwound dsDNA and were then plotted versus protein concentration (n = 2 to 6; error bars represent SD values). (*D*) Simultaneous monitoring of the changes in fluorescence intensity and in anisotropy of the beacon dsDNA. Fluorescence polarization and quenching were monitored for 2 h after incubating the beacon dsDNA with varying concentrations of *N*-αβα. Changes in anisotropy (pink trace) and in the fraction unwound (black trace; as in *C*) were measured as described in *Materials and Methods* (n = 4 to 8; error bars represent SD values). Shown here are 2 h end point values plotted against *N*-αβα concentration. (*E*) As above but with a dsDNA construct having the fluorophore and quencher on the opposite strands (here, strand separation should result in an increase in fluorescence rather than decrease as in *D* and in all other beacon assays; n = 4 to 8; error bars represent SD values). (*F*) Binding to the individual strands of the dsDNA beacon as monitored by fluorescence anisotropy assay. Fluorescence polarization was monitored for 2 h after incubation of the GA-beacon sense strand and the TC-beacon antisense strand, with varying concentrations of *N*-αβα Changes in anisotropy were measured as described in *Materials and Methods* (n = 4 to 8; error bars represent SD values). Shown here are 2 h end point values for GA-beacon sense strand (pink trace) and the TC-beacon antisense strand (black trace) plotted against *N*-αβα concentration.

The *N*-αβα P-Loop Prototype Shows Avid and Cooperative Strand Separation. Strand separation was dependent on the concentration of P-loop prototypes, in terms of both the rate and amplitude of fluorescence decay. At higher concentrations, unwinding was faster and complete (i.e., reached baseline), whereas at lower concentrations unwinding was slower and partial (Fig. 2*A* and *SI Appendix*, Tables S4 and S5). Isotherms obtained by plotting the end-point fluorescence-decay values versus protein concentration indicated that strand separation is highly cooperative (sigmoidal curves and high Hill's coefficient) and occurs with apparent K_D values in the sub μM range (Fig. 2*C* and *SI Appendix*, Table S6). The unwinding kinetics are complex since the dependence of the apparent rate constants on protein concentration was not linear (Fig. 2*B*). Thus, strand separation appears to be a multiple step process. Specifically, the rate determining step changes nonlinearly with protein concentration and at >0.5 μM becomes essentially concentration independent (Fig. 2*B*). As shown later, both of these phenomena (cooperative isotherms and complex kinetics) are in agreement with changes in quaternary structure upon ligand binding. Mutations in the key P-loop residues of *N*-αβα reduced both the rate and the amplitude of fluorescence decay and accordingly showed loss of binding to ssDNA by ELISA (enzyme-linked immunosorbent assay) (*SI Appendix*, Fig. S3).

Unwinding of the beacon dsDNA is likely driven by the binding preference of the P-loop prototypes for ssDNA over dsDNA. This preference was originally observed with the intact prototype using an ELISA-like assay (with immobilized DNA and detection of binding with antibodies to the 6xHis tag) (16). This result was confirmed here, also with the *N*-αβα fragment and the same

oligonucleotides used in the beacon assay (*SI Appendix*, Fig. S4). However, the ELISA exhibited high background. Thus, to more accurately measure binding to ssDNA versus dsDNA, and to gain further insight into the mechanism of strand separation, we employed an assay, following a previously described setup (33), to simultaneously measure strand separation by fluorescence quenching (as in the beacon assay, Fig. 1) and binding to the fluorescently labeled DNA strand(s) by fluorescence anisotropy.

We first tested the beacon dsDNA. Upon addition of the *N*-αβα prototype, a concentration-dependent change in anisotropy was observed that was concomitant with the quenching signal (Fig. 2*F*). The change in anisotropy occurs in parallel with quenching suggesting that strand separation and binding occur in a concerted manner and that the anisotropy signal reports binding to ssDNA or, in other words, the dsDNA bound states do not accumulate. Indeed, under the conditions applied here, the P-loop prototypes do not exhibit dsDNA binding, as confirmed with a dsDNA construct in which the quencher resides on the complementary strand. This construct should show an increase in fluorescence upon strand separation (as opposed to the decrease observed with the beacon dsDNA). However, no such increase was observed upon prototype addition (Fig. 2*E*), likely due to the to the fluorophore-quencher interaction inducing higher duplex stability (31) (also indicated by strand exchange assays described below). Accordingly, this construct also showed a negligible change in anisotropy (Fig. 2*E*).

The second principle governing dsDNA unwinding is the selectivity in binding the TC-rich antisense strand relative to the GA-rich sense strand of the beacon dsDNA. This selectivity was observed by ELISA (*SI Appendix*, Fig. S4) as well as by anisotropy (Fig. 2*F* and *SI Appendix*, Table S7).

Finally, *N*-αβα does not quench the fluorescently labeled dsDNA in the absence of the quencher-containing complementary strand (*SI Appendix*, Fig. S5). This rules out the posibility that fluorescence quenching is driven by spurious interactions of the P-loop prototype with the fluorophore.

The P-Loop Prototypes Facilitate Strand Exchange. Next, we assessed the potential of *N*-αβα to mediate strand-exchange, namely, to accelerate the rate of exchange between a duplex bound ssDNA and an identical competing strand added in excess. We started with the dsDNA construct, which, due to higher stability, did not show strand separation (Figs. 2*E* and 3*A*). This construct also showed no exchange when mixed with 100-fold excess of the unlabeled complementary strand, not even after 24 h incubation (Fig. 3*B*, pink bar). However, addition of *N*-αβα led to an increase of fluorescence, indicating the displacement of the quencher strand by the unlabeled one, in a concentration and time dependent manner (6 to 24 h; Fig. 3*B*).

That the fluorophore-quencher interaction stabilizes the duplex DNA and hinders strand separation and exchange was indicated in the swapped setup showing rapid exchange, namely, displacement of an unlabeled strand by a quencher-containing strand (Fig. 3*C*). In contrast to the setup shown in Fig. 3*A*, *N*-αβα mediated strand exchange at the minute time scale, with only 10-fold excess of the complementary strand (Fig. 3*D* and *SI Appendix*, Table S8). Acceleration of strand exchange by *N*-αβα was also also observed in the same setup yet with "linear" ssDNA oligos that do not form a hairpin (Fig. 3 *E* and *F*). Here, too, no siginficant exchange occurred within the time scale of the experiment in the absence of *N*-αβα, while addition of ≤2.5 μM *N*-αβα induced complete exchange within minutes (Fig. 3 *D* and *F*). Note that the concentration of *N*-αβα remains in large excess (≥0.5 μM) over the excess of competing ssDNA (50 nM). Overall, as expected, the energy gap between the start and end states dictates whether binding by the P-loop prototypes can shift the equilibrium between dsDNA and ssDNA and also how fast the shift is. However, the *N*-αβα prototype also accelerated strand exchange in a more demanding setup (Fig. 3*B*).

Strand exchange mediated by P-loop prototypes was further validated by EMSA (electrophoretic mobility shift assay)

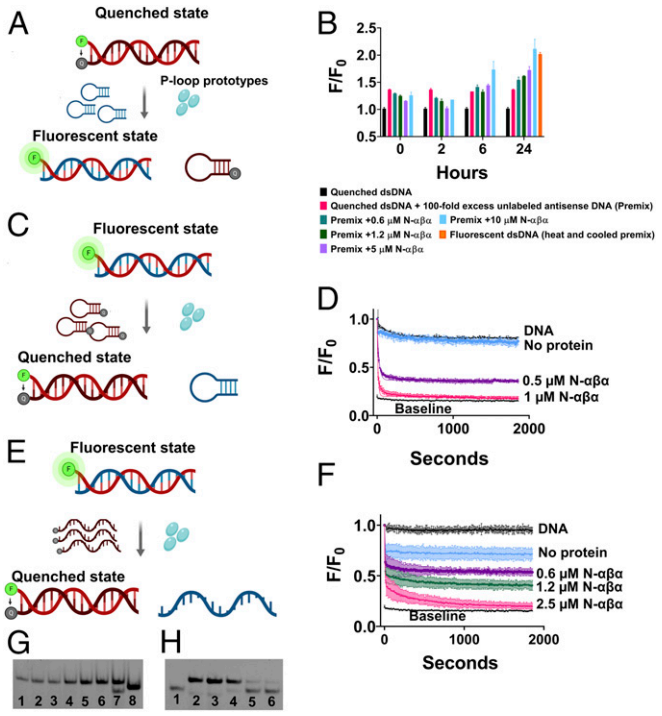

Fig. 3. The *N*-αβα P-loop prototype mediates strand exchanges. (*A*) A schematic of the strand exchange reaction between a quenched dsDNA and a nonlabeled complementary strand (hairpin-forming). (*B*) The exchange reaction was monitored by increase in fluorescence upon addition of *N*-αβα to a premix of quenched dsDNA and 100-fold excess of unlabeled complementary antisense strand. Fluorescence was normalized (F/F_0) to quenched dsDNA (equals 1) and plotted versus time ($n = 2$ to 4; error bars represent SD values). (*C*) A schematic of the strand exchange reaction between a fluorescent dsDNA and a quencher-containing complementary strand (hairpin-forming). (*D*) Fluorescence quenching (F/F_0) was monitored upon addition of *N*-αβα to a premix of fluorescently labeled dsDNA and 10-fold excess of a quencher-containing competing strand (hairpin-forming) ($n = 2$ to 4; error bars represent SD values). Data were fit to a one-phase exponential decay (*SI Appendix*, Table S8). (*E*) A schematic of the strand exchange reaction between a fluorescent dsDNA and a quencher-containing complementary strand (linear). (*F*) Fluorescence quenching (F/F_0) was monitored upon addition of *N*-αβα to a premix of fluorescently labeled dsDNA and 10-fold excess of a linear quencher-containing competing strand ($n = 2$; error bars represent SD values). Data were fit to a two-phase exponential decay. (*G*) Native EMSA gel of the strand-exchange reaction shown in *B*. The reactions were allowed to reach completion (24 h) and DNA products were resolved on native polyacrylamide TBE gels (see *Materials and Methods*). The lanes are as follows: 1) Quenched dsDNA (FAM-GA sense strand plus BHQ-1 antisense strand); 2) DNA Premix (quenched dsDNA with 100-fold excess of unlabeled antisense strand); 3 through 6) DNA Premix with *N*-αβα (0.6, 1.2, 5, and 10 μM, respectively); 7) Fluorescent beacon dsDNA; 8) Fluorescent ssDNA (FAM-GA sense strand) (*SI Appendix*, Table S1). Short dsDNA duplexes comprising of A/T rich regions are prone to improper annealing and/or formation of internal hairpins due to self-complementary regions. This is likely the reason for the visible lower band in lane 7. (*H*) Native EMSA gel of a strand-exchange reaction as shown in *E* but with both strands fluorescently labeled. Strand-exchange reactions were carried out as described above, allowed to reach steady state (14 h), and analyzed on a native polyacrylamide TBE gel. The lanes are as follows: 1) Fluorescent ssDNA (FAM-TC-linear antisense strand); 2) Fluorescent dsDNA (FAM-GA-linear sense strand plus FAM-TC-linear antisense strand); 3) DNA premix (fluorescent dsDNA with 10-fold excess of BHQ-1-linear antisense strand); 4) DNA Premix with 0.6 μM *N*-αβα and 5) with 1.25 μM *N*-αβα; 6) Quenched dsDNA (FAM-GA-linear sense strand plus BHQ-1-linear antisense strand) with FAM-TC-linear antisense strand (5 nM; 1:1 ratio).

experiments. Once the strand-exchange reactions described above reached steady state, the bound proteins were removed by addition of hexametaphosphate, and the DNA products were resolved on native polyacrylamide gels (Fig. 3 *G* and *H*). Fig. 3*G* shows the transition from the initial quenched dsDNA to a fluorescent dsDNA, in a protein concentration dependent manner, indicating DNA unwinding and exchange between the quencher-labeled strand and an unlabeled complementary strand. However, this assay does not track down the change in state of the quencher-containing duplex DNA to ssDNA state. To this end, we used an alternate experimental setup that follows Fig. 3*E*, except that both strands of the duplex DNA were fluorescently labeled. Strand-exchange reactions with a complementary quencher-carrying strand were performed as described above, and the DNA products were resolved on a native EMSA gel. The gel analysis (Fig. 3*H*) indicated the transition of fluorescent dsDNA to a quenched dsDNA state (due to exchange of the bottom fluorescent strand with the quencher-containing strand) and a concomitant accumulation of the displaced flurescent strand (as indicated by the lower band in lanes 4 and 5). Overall, the observed changes in intensity of dsDNA bands (Fig. 3 *G* and *H*) and the parallel accumulation of ssDNA products (Fig. 3*H*) occur only in the *N*-αβα-containing reactions, consistent with the fluorescent assays, thus corroborating the helicase-like activity of P-loop prototypes.

ssDNA Release upon NTP and Polyphosphate Binding. Strand-separation as shown so far is, in essence, a shift in equilibrium in favor of P-loop prototype–bound ssDNA. However, helicases are enzymes that turnover; namely, they also release the bound DNA, typically upon ATP hydrolysis. The P-loop prototypes bind ssDNA and ATP via the same P-loop motif as suggested by mutations in the P-loop residues that abrogated binding of both (16). Based on this observation, we asked if ATP addition could displace the ssDNA, thus also allowing the DNA to relax to its initial dsDNA state (Fig. 4*A*). To this end, the *N*-αβα prototype was mixed with the beacon dsDNA and strand separation was allowed to reach equilibrium. Upon subsequent addition of ATP or GTP, fluorescence reverted to its initial state indicating the release of the bound prototype and reversion to the starting dsDNA state (Fig. 4*B*). Release by ATP occurred with an apparent K_D value in the range expected for a small ligand (K_D^{App} = 2.8 mM), and GTP was a more potent inducer of reversion with twofold tighter K_D^{App} (1.4 mM; Fig. 4*C*). It seems, however, that release of the ssDNA is primarily conferred by the triphosphate group of these NTPs and not the nucleoside moiety, as indicated by even tighter binding of triphosphate (K_D^{App} = 790 μM). Given that inorganic phosphoanhydrides were proposed to have preceded NTPs as life's energy coin, we tested polyphosphate—Kornberg's energy fossil (34)—as well as hexametaphosphate. *N*-αβα showed a strong preference for both polyphosphates (K_D^{App} = 13.7 μM; calculated for an average of 18 phosphates per molecule and 5.6 μM for hexametaphosphate; Fig. 4*C*). Addition of phosphate at molar concentrations that are >100-fold higher than hexametaphosphate did not release the bound protein (5.6 mM phosphate barely had any effect on releasing the bound ssDNA, whereas 5.6 μM hexametaphosphate, equivalent to 33.6 μM phosphate, induced complete release; Fig. 4*C*). Thus, the regain of fluorescence seems to be driven by specific ligand binding rather than by nonspecific effects such as changes in ionic strength. The efficient release of prototype-bound ssDNA by hexametaphosphate was further corroborated by ELISA experiments showing that prior incubation of *N*-αβα with hexametaphosphate inhibited binding to ssDNA (Fig. 4*D*).

Quaternary Structural Plasticity. The intact P-loop prototypes tend to form dimers (16). We assumed that their fragments are more likely to do so and possibly form even higher-order oligomers.

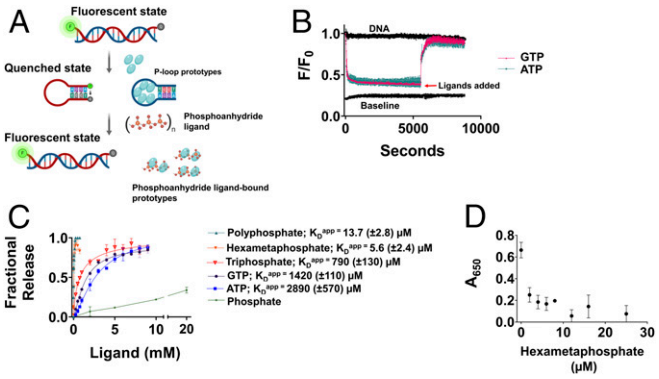

Fig. 4. Release of bound ssDNA by phosphoranhydride ligands. (*A*) A schematic description of the induction of strand separation by a P-loop prototype (first step) followed by the displacement of the bound ssDNA by various phosphoanhydrides (P_n) and relaxation to the initial dsDNA state (second step). (*B*) Representative strand separation experiments with 0.4 μM *N*-αβα added at time 0. Subsequent addition of ATP (turquoise) and GTP (pink, both at 9 mM concentration) after 90 min, leads to the release of the bound ssDNA and its relaxation to the initial dsDNA state. (*C*) The apparent affinities (K_D^{App}) of binding of various phosphoanhydride ligands to P-loop prototype *N*-αβα. The fraction of released ssDNA upon addition of phospholigands was calculated by normalizing the F/F_0 values from the plot in *B*. Complete release (= 1) corresponds to the initial unbound dsDNA state and no release (= 0) corresponds to the steady-state value of F/F_0 prior to ligand addition. The apparent binding affinities (K_D^{app}) were calculated by *SI Appendix*, Eq. **3**. Vertical error bars, and values in parenthesis, represent SD from two independent experiments. (*D*) Inhibition of ssDNA binding. Preincubation of 0.125 μM *N*-αβα with increasing concentrations of hexametaphosphate shows abrogation of binding to 24 base biotinylated ssDNA (*SI Appendix*, Table S1) in an ELISA-format as detected by anti-Histag antibodies (n = 2 to 4; error bars represent SD values).

Despite extensive attempts, we could not obtain crystals of *N*-αβα or of any other functional construct described here. However, we performed chemical crosslinking experiments that revealed that *N*-αβα forms higher-order oligomers, spanning from dimers to hexamers (*SI Appendix*, Fig. S2*D*). Native mass spectrometry (MS) revealed that *N*-αβα self-assembles to form an ensemble of even higher-order forms, with the predominant and unambiguously assigned assemblies being 10-mer and 30-mer (Fig. 5*A*). Furthermore, MS/MS experiments confirmed the identity of these species, as indicated, for example, by isolating the 30+ charge state of the 30-mer species and inducing its dissociation into a highly charged monomer and a stripped 29-mer complex (Fig. 5*B*). Dynamic light scattering (DLS) measurements also indicated that *N*-αβα exists in higher oligomeric form. The addition of inorganic polyphosphates to *N*-αβα results in an increase in the intensity of scattered light which translates to an increase in the average particle diameter (*SI Appendix*, Fig. S6 and Table S9). In agreement with its lower apparent binding affinity (Fig. 4*C*), GTP induced a smaller effect, and the shift demanded higher concentrations. However, the diameter reported by DLS is only a proxy of the actual size, and the precise nature of the changes in the oligomeric assembly of *N*-αβα in response to ligand binding needs to be further investigated.

Discussion

Previous studies have shown that NTP binding and modest catalysis can arise in polypeptides that comprise fragments of modern NTPases (16, 35) or of other ancient domains (36). This study extends our previous report (16) to demonstrate that polypeptides comprising the ancestral β-(P-loop)-α motif, with some minimal sequence additions to facilitate solubility, confer helicase-like functions. Our findings have several implications regarding the primordial P-loop NTPases, as discussed in the sections below.

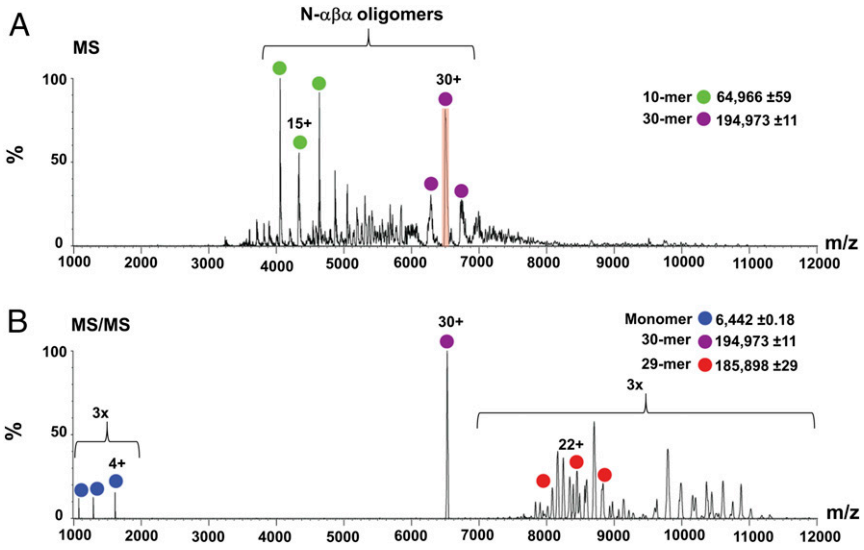

Fig. 5. Quaternary structural characterization of *N*-αβα. (*A*) Native MS analysis of *N*-αβα prototype. Under nondissociative conditions, charge series corresponding to 10 and 30 *N*-αβα oligomers were unambiguously assigned in the MS spectrum (green and purple dots, respectively). The 30+ charge state that was selected for tandem MS analysis is shadowed in red. (*B*) Tandem MS of the *N*-αβα 30-mer oligomer releases a highly charged monomer (blue; at the lower *m/z* range) and a stripped 29-mer oligomer (red), confirming the oligomer stoichiometry. The spectra are magnified threefold above 7,000 and below 2,000 *m/z*.

From a Generalist Phospho-Ligand Binder to Specialized Enzymes. In contemporary helicases, translocases and RecA proteins, the P-loop mediates NTP binding and hydrolysis, while DNA binding is mediated by other surface loops or even by a separate domain (20–23). In contrast, in our prototypes, the P-loop binds NTPs and inorganic polyphosphates, as well as nucleic acids. This multifunctionality is the key to the prototype's helicase-like action. Specifically, the strand-exchange assays demonstrate the ability of the P-loop prototypes to accelerate the exchange between a free and a duplex-bound DNA strand (Fig. 3). However, in the unwinding experiments (Fig. 2), binding of the P-loop prototypes merely shifts the equilibrium toward ssDNA. Nonetheless, the prototypes' action can be reversed by addition of phospho-ligands such as ATP, and foremost by inorganic polyphosphates. These binding-release cycles are not enzymatic turnovers, yet may be a first step toward a bona fide helicase.

This multifunctionality is surprising, as although nucleic acids have phosphate groups, they fundamentally differ from NTPs. Can vestiges of a generalist P-loop, and specifically of ssRNA/DNA binding, be found in extant P-loop NTPases? We searched the PDB for domains that belong to the P-loop NTPase lineage and have ssRNA/DNA bound in proximity to the P-loop (see *Materials and Methods*). This search identified at least two extant P-loop NTPase families: XPD (xeroderma pigmentosum group D) helicase (Fig. 6*A*) and polynucleotide kinase (Fig. 6*B*), where the P-loop interacts with ssDNA, as it does in our P-loop prototypes. In the case of XPD helicase, the canonical Walker A motif diverged and phosphate binding is mediated by a short motif, SGR, at the tip of α1. Remarkably, in case of polynucleotide kinase (Fig. 6*B*), a canonical P-loop binds GTP/GDP (guanosine diphosphate) as well as ssDNA (oligonucleotide) by a network of interactions facilitated by the P-loop residues. This, to our knowledge, is the only instance of an extant multifunctional P-loop that not only catalyzes phosphoryl transfer but also binds ssDNA.

At present, in the absence of structures at atomic resolution, the precise mode of binding by our P-loop prototypes remains unknown. It seems, however, that the Walker A motif is not critical per se. Specifically, we observed that the first Gly of the Walker A motif of our P-loop prototypes plays no role and that the last three residues (GK(S/T)) seem most critical; however, even at these three positions, mutations reduce but not completely abolish binding (see ref. 16 and *SI Appendix*, Fig. S3). Indeed, the most rudimentary forms of phosphate binding make use of backbone amides, of glycine as well as other residues, and

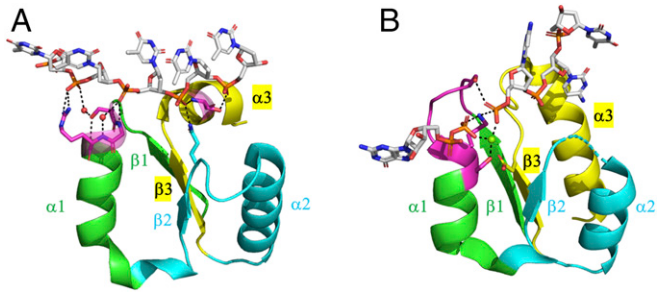

Fig. 6. Instances of P-loops binding to ssDNA in extant P-loop NTPases. (*A*) The Helicase_C_2 domain of XPD helicases (F-groups, ECOD: 2004.1.1.106, Pfam: PF13307). Shown here is a fragment taken from a representative structure (ATP dependent DinG helicase; ECOD domain ID: e6fwrA1, residues 448 to 703; the fragment shown, residues 535 to 599, spans from β1 to α3). The strand topology of this domain follows the simplest P-loop NTPase topology (2-3-1-4-5). Its P-loop resides at the tip of α1, in the very same location as the Walker A motif, but its sequence is noncanonical (SGR, in magenta). Shown are direct as well as water-mediated interactions between the residues of the P-loop and phosphate groups of the ssDNA oligonucleotide (waters shown as red spheres). A glutamine at the tip of β2 (in cyan) and a serine at the N terminus of α3 (in magenta) provide additional anchoring points for the ssDNA. (*B*) Bacterial polynucleotide kinase (F-groups, ECOD: 2004.1.1.32; Pfam: PF13671; strand topology: 2–3-1-4-5). Shown here is a fragment of the P-loop NTPase domain from a representative structure (*Clostridium thermocellum* polynucleotide kinase; ECOD domain ID: e4mdeA1, residues 1 through 170; the shown fragment, residues 8 through 37 and 62 through 98, spans from β1 to α3, with residues 38 through 61 truncated for clarity). The canonical Walker A motif resides between β1 and α1 (in magenta, GSSGSGKS) and binds GTP. This representative structure shows the product complex of GDP•Mg^{+2} and the phosphorylated ssDNA. The lysine and serine residues of the P-loop motif at the tip of α1 provide bridging interactions, coordinated by an Mg^{2+} ion (green sphere), between the β phosphate of the bound GDP and the phosphate group at the 5′-OH of the ssDNA.

most critically, of bidentate interactions involving both backbone and side-chain H-bonds, foremost by Ser/Thr (14). These elements are seen in XPD helicase (Fig. 6*A*) and are likely to be the key to a generalist P-loop.

That the ancient P-loop was a "generalist" that mediates multiple functions is in line with the hypothesis that the earliest enzymes were multifunctional (37). A plausible explanation, in accord with Dayhoff's hypothesis (3, 4), is that such multifunctional primordial P-loop fragments were, at a later stage, duplicated and fused. This resulted in a tandem repeat of β-(P-loop)-α elements, where the P-loops compose the active site. Further divergence allowed one P-loop to retain NTP binding and develop the ability to catalyze phosphoryl transfer, while the other P-loops were repurposed to bind nucleic acids. This scenario is illustrated by XPD helicases wherein a P-loop ATPase domain, the DEAD domain (ECOD domain ID: e6fwrA2), is fused to the ssDNA binding helicase C_2 domain.

A Minimal Structural Context for P-Loop Function. Bioinformatics analyses have pointed out the β-(P-loop)-α−β fragment as the recurrent structural/functional unit in dinucleotide binding proteins (5, 38). While we found that this fragment mediates strand separation, it was aggregation prone, likely due to the highly hydrophobic β1 strand (*SI Appendix*, Table S3), and active only under nonstringent conditions. In contrast, the α−β-(P-loop)-α fragment was soluble and functional in the presence of salt. Having a helix preceding the β-(P-loop)-α fragment seems to present an advantage, primarily by enhancing solubility of this exceedingly hydrophobic βα fragment. We also note that although in most P-loop NTPase domains β1, from which the P-loop extends, comprises the N terminus, an alpha helix that precedes β1 is also observed in some P-loop families (e.g., in the F-type ATP synthases [F-groups, ECOD: 2004.1.1.53; Pfam: PF00006], Transcription termination factor Rho [F-groups, ECOD: 2004.1.1.53; Pfam: PF00006], and RecA [F-groups, ECOD: 2004.1.1.237; Pfam: PF00154]). Moreover, the helicase-like function is retained despite truncation of most of this N-terminal helix, and even the β-(P-loop)-α fragment that could not be isolated was copurified with bound nucleic acids. These results, along with the observation that ancestral β-(P-loop)-α motif can be embedded in various structural contexts while retaining function, suggest that β-(P-loop)-α is likely the minimal stand-alone seed from which P-loop NTPases could have emerged.

Dynamic Self-Assembly Is Key to Function. We hypothesized that self-assembly is a critical bridging step that enables short polypeptides to be functional on their own prior to their duplication and fusion—a step that likely required advanced genetic and protein translation machineries (4). Shorter polypeptides are unlikely to confer biochemical functions such as small ligands binding (39). However, self-assembly, even in rudimentary forms such as amyloid fibers (40, 41) or coacervates (42), can provide the operative volume and network of interactions necessary for biochemical function (41, 43). Although further studies are required to elucidate the mechanistic and structural basis of these assemblies, the propensity of the P-loop prototypes to self-assemble is evident (Fig. 5). Foremost, oligomerization confers avidity (i.e., binding via multiple P-loops per functional unit). Otherwise, binding at μM affinity by a solvent exposed loop would have been impossible. Changes between oligomeric states are also indicated by the cooperativity of the strand separation reactions, as manifested in high Hill coefficients and complex kinetics (Fig. 2). Similar DNA-induced cooperative functions are documented for SSB proteins (44), helicases (45, 46), and RecA proteins (47–50), and the latter two are also P-loop NTPases. Further, RecA has been shown to also possess strand-separation activity on short dsDNAs (51), and its recombinase function involves changes in oligomeric states, from monomers-dimers up to long filaments (1, 2).

The Potential Role of P-Loop Prototypes in the Pre-LUCA World. Whether proteins (metabolism) or nucleic acids (information) came first has been extensively debated. However, these scenarios are not necessarily mutually exclusive—both elements had to coexist in some rudimentary form well before the LUCA (52). It is widely accepted that P-loop NTPases emerged early, possibly within an RNA-protein world (5, 11–13, 17, 53, 54). Accordingly, RNA/DNA remodelers such as SFII (superfamily II) helicases (54) and RecA constitute a major class of P-loop NTPases ascribed to the LUCA (7, 8). P-loop polypeptides may have therefore marked the early stages of cooperation between proteins and nucleic acids [other examples include polypeptides that interacted with ribosomal RNA (55–57)]. Specifically, in the absence of protein-based polymerases, nucleic acids must have replicated either nonenzymatically (58, 59) and/or in a self-catalytic manner (60). However, in such scenarios, a fundamental impediment for multiple rounds of replication is the stability of duplex RNA (61) [or DNA for that matter (62)]. Heat can separate the strands, yet reannealing would be faster than any abiotic replication (63). A plausible solution for unwinding RNA duplex products is short "invader" oligonucleotides (64); however, how strand displacement would initiate remains unresolved (64). Thus, for multiround replications, certainly of templates that are dozens of nucleotides long, an unwinding polypeptide is likely necessary. The duplex unwinding and strand displacement activities of our P-loop prototypes provide a plausible solution to these challenges.

Our results suggest that the strand separation and exchange mediated by the P-loop prototypes is a thermodynamically driven process. The feasibility and rate of this process are dictated by the energetic gap between the initial and final DNA states. It is therefore likely that prior to the emergence of enzymatic function, primordial helicase-like polypeptides were fundamentaly ssDNA/RNA binding proteins that "passively" unwound DNA duplexes by relying on transient exposure of ssDNA due to thermal fluctuations and/or DNA fraying (26, 27). Considering that the prototypes' affinity to the GA-rich strand of the beacon is low, it is likely that strand separation is primarily driven by binding to the TC-rich strand (Fig. 2*F*). In fact, this is a commonly reported mode of unwinding for hexameric replicative helicases where the helicase-ring encircles one strand while the other strand is excluded (65). The ability to bind ssDNA likely undelies the strand-exchange process as well (Fig. 3), thus resembling the mode of action of extant RecA proteins (66). Thus, although further studies will be required to decipher the precise mechanism of action of these P-loop prototypes, it appears that these rudimentary P-loop helicases hold tangible links to their modern descendants.

Inorganic Polyphosphates as Energy Fossils. Our P-loop prototypes bind phosphate moieties with no apparent interactions with the nucleoside (ATP, GTP, and triphosphate bind similarly; Fig. 4). Foremost, long chain polyphosphate binds with 1,000-fold higher affinity than ATP. Indeed, inorganic phosphoanhydrides were proposed to have been the abiotic energy precursors of NTPs (67). Of particular interest is hexametaphosphate—a cyclic ring of six phosphates that also binds with μM affinity (Fig. 4). In the primitive biotic world, polyphosphates could have served not only as an energy source but also as a scaffold that facilitated the assembly and orientation of phospholipids, nucleic acids, and proteins (68). Further, trimetaphosphate has been extensively explored as an abiotic condensation reagent that promotes synthesis of peptides and nucleic acids (69). Trimetaphosphate binds weakly to our P-loop prototypes, but it coexists in equilibrium with hexametaphosphate (69) that binds at μM affinity. Thus, the mode of

action of these P-loop prototypes is tantalizingly tailored to the requirements of a primordial world.

To conclude, Charles Darwin's immortal statement echoes: "from so simple a beginning endless forms most beautiful and most wonderful have been, and are being, evolved." The P-loop prototypes described here provide a glimpse of how these simple beginnings might have looked—short and simple polypeptides that are nonetheless linked to their modern descents in sequence, structure, and function.

Materials and Methods

P-loop prototypes were cloned, expressed, and purified as described (16) with some modifications (*SI Appendix*, Supplementary text and *Materials and Methods*). For strand-separation assays, fluorescence quenching was measured upon titrating the beacon dsDNA with varying concentrations (0.1 to 1 μM) of P-loop prototypes in 50 mM Tris (pH 8), or in 50 mM Tris with 100 mM NaCl (pH 8). Reactions were monitored for 2 h using the Infinite M Plex microplate reader (TECAN), at 24 °C, with an excitation/emission wavelength of 495/540 nm. The fluorescence decay values were normalized to the initial fluorescence (before protein addition; F_0) to derive F/F_0 values that were fit to standard one- or two-phase exponential decay equations (*SI Appendix*, Eqs. **1** and **2**, Supplementary text and *Materials and Methods*) to derive the apparent rate constants (k_{app}). The apparent binding affinities (K_D^{app}) and Hill's coefficient (h) (for strand separation and release by phospho-ligands) were determined by plotting the normalized end point F/F_0 values (denoted as "Fraction Unwound") against protein concentration and by fitting the data to a cooperative binding equation (*SI Appendix*, Eq. **3**, Supplementary text and *Materials and Methods*).

Strand-exchange assays were performed by titrating DNA premixes (dsDNA with excess of complementary single strand) (*SI Appendix*, Table S1) with varying concentrations of the *N*-αβα prototype (0.5 to 10 μM). The reactions were carried out in the stringent binding conditions (50 mM Tris with 100 mM NaCl [pH 8]) and monitored using the Infinite M Plex microplate reader for 2 h as above. The normalized end point F/F_0 values were fit to standard one or two exponential decay equation (*SI Appendix*, Eqs. **1** and **2**).

For fluorescence anisotropy measurements, various ssDNA and dsDNA constructs (*SI Appendix*, Table S1) were titrated with *N*-αβα prototype (0.1 to 2 μM) in buffer containing 50 mM Tris with 100 mM NaCl (pH 8) at 24 °C. Fluorescence polarization was monitored for 2 h using the Cytation 5 multimode reader (BioTek) with the green filter set (485/20 nm for excitation, and 528/20 nM for emission) and a top optic probe with a 510 nm dichroic mirror. The steady-state anisotropy values were derived from the parallel and perpendicular emission intensities using the standard anisotropy equation (*SI Appendix*, Eq. **4**, Supplementary text and *Materials and Methods*), and binding affinities (K_D) and Hill's coefficient (h) were derived by fitting to a cooperative binding equation (*SI Appendix*, Eq. **5**, Supplementary text and *Materials and Methods*). For anisotropy-quenching dual assays, the change in anisotropy and fraction unwound were measured as described above and plotted against protein concentration. All data fitting was carried out using GraphPad Prism (8.3.0) software.

For native MS measurements, the *N*-αβα polypeptide was dialyzed against 150 mM ammonium acetate (pH 7.5) and diluted to a final concertation of 10 μM. Nanoflow electrospray ionization MS and tandem MS experiments were conducted under nondenaturing conditions on a Q Exactive UHMR Hybrid Quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific).

Identification of ssDNA binding by extant P-loop NTPases was carried out using a short script implemented in Python to identify structures, from the Protein Data Bank, in which a P-loop domain and ssDNA come within 3.5 Å of each other (*SI Appendix*, Supplementary text and *Materials and Methods*).

Data Availability. All study data are included in the article and/or *SI Appendix*.

ACKNOWLEDGMENTS. This research was funded by a Minerva Foundation grant for scientific cooperation between Germany and Israel. We are grateful to Prof. Ita Gruic-Sovulj for her insightful suggestion to simultaneously monitor quenching and anisotropy. We thank Dr. Guy Shmul for help with DLS analysis and for providing size standards and Dr. Gili Ben-Nissan for help with native MS experiments. Schematics for strand separation, strand exchange, and release assays were created with http://BioRender.com.

1. J. C. Bell, S. C. Kowalczykowski, A. Rec, RecA: Regulation and mechanism of a molecular search engine. *Trends Biochem. Sci.* **41**, 491–507 (2016).
2. F. Bleichert, M. R. Botchan, J. M. Berger, Mechanisms for initiating cellular DNA replication. *Science* **355**, eaah6317 (2017).
3. R. V. Eck, M. O. Dayhoff, Evolution of the structure of ferredoxin based on living relics of primitive amino acid sequences. *Science* **152**, 363–366 (1966).
4. M. L. R. Romero, A. Rabin, D. S. Tawfik, Functional proteins from short peptides: Dayhoff s hypothesis turns 50. *Angew. Chem. Int. Ed. Engl.* **1980**, 15966–15971 (2016).
5. V. Alva, J. Söding, A. N. Lupas, A vocabulary of ancient peptides at the origin of folded proteins. *Elife* 1–19 (2015).
6. J. E. Walker, M. Saraste, M. J. Runswick, N. J. Gay, Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* **1**, 945–951 (1982).
7. D. D. Leipe, Y. I. Wolf, E. V. Koonin, L. Aravind, Classification and evolution of P-loop GTPases and related ATPases. *J. Mol. Biol.* **317**, 41–72 (2002).
8. D. D. Leipe, E. V. Koonin, L. Aravind, Evolution and classification of P-loop kinases and related proteins. *J. Mol. Biol.* **333**, 781–815 (2003).
9. M. Saraste, P. R. Sibbald, A. Wittinghofer, The P-loop–A common motif in ATP- and GTP-binding proteins. *Trends Biochem. Sci.* **15**, 430–434 (1990).
10. T. Ogura, A. J. Wilkinson, AAA+ superfamily ATPases: Common structure–diverse function. *Genes Cells* **6**, 575–597 (2001).
11. E. V. Koonin, Y. I. Wolf, L. Aravind, Protein fold recognition using sequence profiles and its application in structural genomics. *Adv. Protein Chem.* **54**, 245–275 (2000).
12. J. Söding, A. N. Lupas, More than the sum of their parts: On the evolution of proteins from peptides. *BioEssays* **25**, 837–846 (2003).
13. B. G. Ma *et al.*, Characters of very ancient proteins. *Biochem. Biophys. Res. Commun.* **366**, 607–611 (2008).
14. L. M. Longo, D. Petrović, S. C. L. Kamerlin, D. S. Tawfik, Short and simple sequences favored the emergence of N-helix phospho-ligand binding sites in the first enzymes. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 5310–5318 (2020).
15. J. D. Watson, E. J. Milner-White, A novel main-chain anion-binding site in proteins: The nest. A particular combination of phi,psi values in successive residues gives rise to anion-binding sites that occur commonly and are found often at functionally important regions. *J. Mol. Biol.* **315**, 171–182 (2002).
16. M. L. Romero Romero *et al.*, Simple yet functional phosphate-loop proteins. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E11943–E11950 (2018).
17. I. N. Berezovsky, Towards descriptor of elementary functions for protein design. *Curr. Opin. Struct. Biol.* **58**, 159–165 (2019).
18. P. Laurino *et al.*, An ancient fingerprint indicates the common ancestry of Rossmann-fold enzymes utilizing different ribose-based cofactors. *PLoS Biol.* **14**, e1002396 (2016).
19. B. W. Matthews, C. S. Craik, H. Neurath, Can small cyclic peptides have the activity and specificity of proteolytic enzymes? *Proc. Natl. Acad. Sci. U.S.A.* **91**, 4103–4105 (1994).
20. G. M. Weinstock, K. McEntee, I. R. Lehman, ATP-dependent renaturation of DNA catalyzed by the recA protein of Escherichia coli. *Proc. Natl. Acad. Sci. U.S.A.* **76**, 126–130 (1979).
21. J. P. Menetski, D. G. Bear, S. C. Kowalczykowski, Stable DNA heteroduplex formation catalyzed by the Escherichia coli RecA protein in the absence of ATP hydrolysis. *Proc. Natl. Acad. Sci. U.S.A.* **87**, 21–25 (1990).
22. A. E. Gorbalenya, E. V. Koonin, Helicases: Amino acid sequence comparisons and structure-function relationships. *Curr. Opin. Struct. Biol.* **3**, 419–429 (1993).
23. M. Prentiss, C. Prévost, C. Danilowicz, Structure/function relationships in RecA protein-mediated homology recognition and strand exchange. *Crit. Rev. Biochem. Mol. Biol.* **50**, 453–476 (2015).
24. L. M. Iyer, D. D. Leipe, E. V. Koonin, L. Aravind, Evolutionary history and higher order classification of AAA+ ATPases. *J. Struct. Biol.* **146**, 11–31 (2004).
25. A. Y. Mulkidjanian, K. S. Makarova, M. Y. Galperin, E. V. Koonin, Inventing the dynamo machine: The evolution of the F-type and V-type ATPases. *Nat. Rev. Microbiol.* **5**, 892–899 (2007).
26. P. H. von Hippel, N. P. Johnson, A. H. Marcus, Fifty years of DNA "breathing": Reflections on old and new approaches. *Biopolymers* **99**, 923–954 (2013).
27. C. Phelps, W. Lee, D. Jose, P. H. von Hippel, A. H. Marcus, Single-molecule FRET and linear dichroism studies of DNA breathing and helicase binding at replication fork junctions. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 17320–17325 (2013).
28. J. Fei, T. Ha, Watching DNA breath one molecule at a time. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 17173–17174 (2013).
29. T. Lionnet, M. M. Spiering, S. J. Benkovic, D. Bensimon, V. Croquette, Real-time observation of bacteriophage T4 gp41 helicase reveals an unwinding mechanism. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 19790–19795 (2007).
30. C. A. Belon, D. N. Frick, Monitoring helicase activity with molecular beacons. *Biotechniques* **45**, 433–440, 442 (2008).
31. Z. A. Zimmers, N. M. Adams, W. E. Gabella, F. R. Haselton, Fluorophore-quencher interactions effect on hybridization characteristics of complementary oligonucleotides. *Anal. Methods* **11**, 2862–2867 (2019).
32. N. Koga *et al.*, Principles for designing ideal protein structures. *Nature* **491**, 222–227 (2012).
33. D. A. Hiller *et al.*, Simultaneous DNA binding and bending by EcoRV endonuclease observed by real-time fluorescence. *Biochemistry* **42**, 14375–14385 (2003).
34. A. Kornberg, Inorganic polyphosphate: Toward making a forgotten polymer unforgettable. *J. Bacteriol.* **177**, 491–496 (1995).
35. W. J. Chuang, C. Abeygunawardana, A. G. Gittis, P. L. Pedersen, A. S. Mildvan, Solution structure and function in trifluoroethanol of PP-50, an ATP-binding peptide from F1ATPase. *Arch. Biochem. Biophys.* **319**, 110–122 (1995).

36. Y. Pham *et al*., Tryptophanyl-tRNA synthetase urzyme: A model to recapitulate molecular evolution and investigate intramolecular complementation. *J. Biol. Chem.* **285**, 38590–38601 (2010).
37. A. R. Jensen, Enzyme recruitment in evolution of new function. *Annu. Rev. Microbiol.* **30**, 409–425 (1976).
38. R. K. Wierenga, M. C. H. De Maeyer, W. G. J. Hoi, Interaction of pyrophosphate moieties with α-helixes in dinucleotide binding proteins. *Biochemistry* **24**, 1346–1357 (1985).
39. D. R. Corey, M. A. Phillips, Cyclic peptides as proteases: A reevaluation. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 4106–4109 (1994).
40. J. Greenwald, R. Riek, On the possible amyloid origin of protein folds. *J. Mol. Biol.* **421**, 417–426 (2012).
41. C. M. Rufo *et al*., Short peptides self-assemble to produce catalytic amyloids. *Nat. Chem.* **6**, 303–309 (2014).
42. L. M. Longo *et al*., Primordial emergence of a nucleic acid-binding protein via phase separation and statistical ornithine-to-arginine conversion. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 15731–15739 (2020).
43. O. V. Makhlynets, P. M. Gosavi, I. V. Korendovych, Short self-assembling peptides are able to bind to copper and activate oxygen. *Angew. Chem. Int. Ed. Engl.* **55**, 9017–9020 (2016).
44. K. Dubiel *et al*., Structural mechanisms of cooperative DNA binding by bacterial single-stranded DNA-binding proteins. *J. Mol. Biol.* **431**, 178–195 (2019).
45. N.-R. Lee *et al*., Cooperative translocation enhances the unwinding of duplex DNA by SARS coronavirus helicase nsP13. *Nucleic Acids Res.* **38**, 7626–7636 (2010).
46. A. K. Byrd, K. D. Raney, Increasing the length of the single-stranded overhang enhances unwinding of duplex DNA by bacteriophage T4 Dda helicase. *Biochemistry* **44**, 12990–12997 (2005).
47. I. R. Lehman, M. M. Cox, recA protein-promoted DNA strand exchange. Stable complexes of recA protein and single-stranded DNA formed in the presence of ATP and single-stranded DNA binding protein. *J. Biol. Chem.* **257**, 8523–8532 (1982).
48. M. M. Cox, Motoring along with the bacterial RecA protein. *Nat. Rev. Mol. Cell Biol.* **8**, 127–138 (2007).
49. S. C. Kowalczykowski, Structural biology: Snapshots of DNA repair. *Nature* **453**, 463–466 (2008).
50. K. Ragunathan, C. Joo, T. Ha, Real-time observation of strand exchange reaction with high spatiotemporal resolution. *Structure* **19**, 1064–1073 (2011).
51. M. Bianchi, B. Riboli, G. Magni, E. coli recA protein possesses a strand separating activity on short duplex DNAs. *EMBO J.* **4**, 3025–3030 (1985).
52. M. Preiner *et al*., The future of origin of life research: Bridging decades-old divisions. *Life (Basel)* **10**, E20 (2020).
53. A. Goncearenco, I. N. Berezovsky, Protein function from its emergence to diversity in contemporary proteins. *Phys. Biol.* **12**, 045002 (2015).
54. V. Anantharaman, E. V. Koonin, L. Aravind, Comparative genomics and evolution of proteins involved in RNA metabolism. *Nucleic Acids Res.* **30**, 1427–1464 (2002).
55. A. N. Lupas, V. Alva, Ribosomal proteins as documents of the transition from unstructured (poly)peptides to folded proteins. *J. Struct. Biol.* **198**, 74–81 (2017).
56. K. A. Lanier, P. Roy, D. M. Schneider, L. D. Williams, Ancestral interactions of ribosomal RNA and ribosomal proteins. *Biophys. J.* **113**, 268–276 (2017).
57. M. Frenkel-Pinter *et al*., Mutually stabilizing interactions between proto-peptides and RNA. *Nat. Commun.* **11**, 3137 (2020).
58. C. Deck, M. Jauker, C. Richert, Efficient enzyme-free copying of all four nucleobases templated by immobilized RNA. *Nat. Chem.* **3**, 603–608 (2011).
59. K. Adamala, J. W. Szostak, Nonenzymatic template-directed RNA synthesis inside model protocells. *Science* **342**, 1098–1100 (2013).
60. D. P. Bartel, J. A. Doudna, N. Usman, J. W. Szostak, Template-directed primer extension catalyzed by the Tetrahymena ribozyme. *Mol. Cell. Biol.* **11**, 3390–3394 (1991).
61. P. D. Ross, J. M. Sturtevant, The kinetics of double helix formation from polyriboadenylic acid and polyribouridylic acid. *Proc. Natl. Acad. Sci. U.S.A.* **46**, 1360–1365 (1960).
62. J. S. Teichert, F. M. Kruse, O. Trapp, Direct prebiotic pathway to DNA nucleosides. *Angew. Chem. Int. Ed. Engl.* **58**, 9944–9947 (2019).
63. J. W. Szostak, The eightfold path to non-enzymatic RNA replication. *J. Syst. Chem.* **3**, 2 (2012).
64. L. Zhou *et al*., Non-enzymatic primer extension with strand displacement. *Elife* **8**, e51888 (2019).
65. Z. Yuan *et al*., DNA unwinding mechanism of a eukaryotic replicative CMG helicase. *Nat. Commun.* **11**, 688 (2020).
66. T. van der Heijden *et al*., Homologous recombination in real time: DNA strand exchange by RecA. *Mol. Cell* **30**, 530–538 (2008).
67. A. Kornberg, N. N. Rao, D. Ault-riché, Inorganic polyphosphate: A molecule of many functions. *Annu. Rev. Biochem.* **68**, 89–125 (1999).
68. M. R. W. Brown, A. Kornberg, Inorganic polyphosphate in the origin and survival of species. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 16085–16087 (2004).
69. G. Baccolini, The possible role of cyclic pentacoordinate phosphorus intermediates in the origin and evolution of life. Are phosphoric anhydride and trimetaphosphates prebiotic reagents? *Phosphorus Sulfur Silicon Relat. Elem.* **190**, 2173–2186 (2015).