Archives of Biochemistry and Biophysics 606 (2016) 64–72

Contents lists available at ScienceDirect

Archives of Biochemistry and Biophysics

journal homepage: www.elsevier.com/locate/yabbi

Putative binding mode of *Escherichia coli* exopolyphosphatase and polyphosphates based on a hybrid *in silico*/biochemical approach

Cristhian Boetsch [a], Daniel R. Aguayo-Villegas [b], Fernando D. Gonzalez-Nilo [b], Á. Teresita Lisa [a], Paola R. Beassoni [a, *]

[a] Departamento de Biología Molecular, Facultad de Ciencias Exactas Físico-Químicas y Naturales, Universidad Nacional de Río Cuarto, Río Cuarto, Argentina
[b] Centro de Bioinformática y Biología Integrativa, Universidad Andres Bello, Santiago, Chile

ARTICLE INFO

Article history:
Received 10 June 2016
Received in revised form
5 July 2016
Accepted 8 July 2016
Available online 14 July 2016

Keywords:
Exopolyphosphatase
Molecular dynamics
Processivity
Binding
Polyphosphate

ABSTRACT

The exopolyphosphatase of *Escherichia coli* processively and completely hydrolyses long polyphosphate chains to ortho-phosphate. Genetic surveys, based on the analysis of single *ppx*$^-$ or *ppk*$^-$ mutants and on the double mutant, demonstrate a relationship between these genes and the survival capacity. The exopolyphosphatase belongs to the ASKHA protein superfamily, hence, its active site is well known; however, the knowledge of the way in which this enzyme binds polyP remains incomplete. Here we present different computational approaches, site-direct mutagenesis and kinetic data to understand the relationship between structure and function of exopolyphosphatase. We propose H^{378} as a fundamental gatekeeper for the recognition of long chain polyphosphate.

1. Introduction

Inorganic polyphosphate (polyP) is a linear anionic polymer of phosphate molecules which was found in all living organisms and may form aggregates. The phosphate molecules within polyP are held together by high-energy phosphoanhydride bonds. The length of this polymer may vary between ten to hundreds of units, depending on the organism and its physiological stage.

The metabolism of polyP in prokaryotes is mainly governed by two enzymes: i. polyphosphate kinase (PPK) which synthesizes polyP from ATP; and ii. exopolyphosphatase (Ppx) which releases inorganic phosphate (P_i) from polyP. In *Escherichia coli*, Ppx (from now on *ec*Ppx) hydrolyses terminal residues of long chain polyP into P_i in a highly processive manner [1–3]. *ec*Ppx activity is Mg^{2+} dependent and it is activated by certain monovalent cations, being K^+ the best positive effector at an optimal concentration of 175 mM [1,4].

polyP is an energy and phosphate reservoir, playing an important role in bacterial survival and stress responses [5,6]. It is also involved in the virulence of several pathogenic microorganisms such as *Helicobacter pylori*, *Pseudomonas aeruginosa*, *Escherichia coli* and *Vibrio cholerae* [1,5,7–9].

In *Mycobacterium tuberculosis* [10], *Bacillus cereus* [11] and *Neisseria meningitidis* [12], the deletion of the *ppx* gene affects pathogenesis while in *Corynebacterium glutamicum* produces deficiencies to achieve growth in P_i poor media [13]. Recently, our laboratory research group has demonstrated that deletion of *ppx* in *P. aeruginosa* affects motility (both swarming and swimming) and quorum sensing [7].

Ppx structure was solved by two independent groups [14,15]. Ppx is comprised by two monomers, with four domains –I to IV- each, positioned in a head to tail orientation. The first two domains contribute to the formation of the N-terminal subunit and resemble a characteristic fold of the ASKHA superfamily. Domains III and IV comprise the C-terminal subunit [14,15]. Both subunits are linked together by a protease-sensitive loop [1]. The active site (AS) is located in the N-terminal subunit and it is formed by residues E^{121}, D^{17}, S^{20}, N^{21}, S^{22}, D^{143}, G^{146}, G^{147}, S^{148} and E^{150} [14,15]. The C-

Abbreviations: AS, active site; *aa*Ppx, *A. aeolicus* exopolyphosphatase; ADT, AutoDockTools; ATP, adenosine tri-phosphate; CERELA, Centro de Referencia para Lactobacillis; *ec*Ppx, *E. coli* exopolyphosphatase; MD, molecular dynamics; *pa*Ppx, *P. aeruginosa* exopolyphosphatase; P_i, inorganic phosphate; PMF, potential of mean force; polyP, Inorganic polyphosphate; PP_n, polyphosphate of *n* residues; PPK, polyphosphate kinase; Ppx, polyphosphate phosphatase exopolyphosphatase; SMD, steered molecular dynamics.

* Corresponding author.
E-mail address: pbeassoni@exa.unrc.edu.ar (P.R. Beassoni).

http://dx.doi.org/10.1016/j.abb.2016.07.005
0003-9861/

terminal subunit was proposed to be directly involved in long chain polyP binding, possessing an affinity 100 times higher than the N-terminal one [1]. Furthermore, from biochemical experiments it was observed that *ec*Ppx is highly processive but along the enzymatic reaction releases polyP of ~40 residues long [3]. Two other interaction points for polyP chain were proposed to be present in the protein in addition to the AS under non-processivity conditions (2–10 times of maximum [K^+]) since polyP of ~14 and ~50 residues were detected [1].

To date, there is no available data on the structure of the Ppx-polyP complex; however, the number of sulphate ions bound to Ppx in one of the reported structures is greater than any other structure deposited in the Protein Data Bank [14]. Thus, the authors proposed a model in which the phosphate chain follows a path marked by the sulphate ions bound to the protein. In this way, polyP would not only bind to the AS (or site i), but it would also bind to a second moiety (R^{138} and N^{208}; site ii) and traverse an aqueduct formed by II, III and IV domains (N^{119} E^{163} R^{165} T^{200} E^{371} N^{377} H^{378} S^{379} D^{393} N^{447} and R^{449}) to then reach a channel occupied by a sulphate triad (R^{166} K^{197} H^{378} and R^{413}; site iii) and finally polyP would cross over to another sulphate tetrad (S^{316} H^{356} Q^{358} R^{364} W^{365} ^{428}KKK^{430}; site iv).

On the other hand, Rangarajan et al. [15] emphasized the binding model in two aspects. Electrostatic potential calculations suggest the formation of an electropositive channel where polyP would bind. Indeed, this channel is coincident with the region occupied by the sulphate triad in the crystal structure mentioned above. These authors also compare domains I and II with the Ppx/GppA structure of *Aquifex aeolicus* (*aa*Ppx). They proposed that *ec*Ppx would have an open state as *aa*Ppx, and polyP would bind first to the electropositive channel to reach then directly the AS.

The literature review and comparison between the structures described by those authors show that it is unclear the mode in which polyP binds to Ppx. Recently, Wei et al. [16] have proposed that positively charged semi-tunnels in polyP-related proteins are of significance for polyP binding, based on their electrostatic properties.

As polyP plays a remarkable role in pathogenesis, survival and stress tolerance, we believe that the study of this system with a theoretical approach through molecular simulations and complemented with experiments is crucial to understand the action mode of this enzyme, which could be a target for therapeutic action against this and other bacteria.

2. Materials and methods

2.1. Electrostatic calculation

Side chains pKa were calculated using the propKa method [17] at pH 8.0 (optimal activity conditions). Partial charges and vdW radius were assigned using PDB2PQR [18,19] according to CHARMM parameters. The electrostatic potential was calculated using the APBS method [20] on a 160 × 160 × 190 Å box size with 0.75 Å of resolution centered at the protein center of mass. The calculated electrostatic potential was mapped into the protein surface calculated with the MSMS method in VMD.

2.2. Molecular docking assays

All docking assays and grid calculations were performed using Autodock4 and Autogrid4, respectively [21]. AutoDockTools 1.5.6 (ADT) platform was used for system preparation, partial charges assignment and hydrogen atoms adding as described elsewhere [21] using protonation states as calculated with propKa. A polyP molecule composed by five phosphate group was sketched with Molefacture from VMD and prepared with the Ligand utility of ADT, where partial atomic charges were added using the Gasteiger-Marsili method [22]. The phosphodiester backbone was used as rotational axis. The conformational search was performed using the Lamarckian genetic algorithm [23] as conformational-search engine with all parameters as default. The protein surface was sampled using 165 overlapping energy scoring grids of 15 × 15 × 15 $Å^3$ with a resolution of 0.625 Å. The minimum binding energies around 5 Å of the protein were mapped into Ppx surface calculated using MSMS form VMD platform [24]. For molecular dynamics simulations (MD), 200 runs were performed with a maximum number of energy eval of 1.10^7 in a 22.5 × 22.5 × 15 $Å^3$ grid box with a resolution of 0.325 Å which was centered into the electropositive channel (the rest of parameters were set as default).

2.3. Evolutionary trace

*ec*Ppx sequence was used as query in pHMMer [25] search on Uniprot rp55 databank. Retrieved sequences were considered orthologous if they retrieved the original *ec*Ppx sequence as top hit when reciprocal pHMMer query into *E. coli* predicted proteome was performed. Sequences which did not cover more than 77% of the query or E-values greater than 10^{-14} were discarded. Multiple sequence alignment was carried out using MAFFT method and BLOSUM62 as substitution matrix. The maximum likelihood phylogenetic tree was constructed using the PROML method from the PHYLIP platform and the JTT substitution model [26]. The ConSurf server (http://consurf.tau.ac.il/) [27] was used to estimate the evolutionary conservation of amino acid positions in the *ec*Ppx molecule, based on the phylogenetic relationships between homologous sequences. The conservation scores are divided into a discrete scale of nine grades for visualization, from the most variable positions (grade 1), through intermediately conserved positions (grade 5), to the most conserved ones (grade 9).

2.4. Molecular dynamics simulations

All systems were assembled using VMD [24] and MD were carried out with NAMD software [28] using the CHARMM27 force field [29]. Briefly, the protein alone or complexed with the docked ligand was hydrated with a TIP3P cubic water box of 27 nm^3. K^+ and Cl^- counterions were added to a final concentration of 175 mM in order to neutralize charges and to represent the estimated optimal ion concentration for *in vitro* assays of Ppx activity [1]. The systems were submitted to an energy minimization of 10,000 steps using a conjugate gradient algorithm followed by a 100 ps heating to 310 K through Langevin dynamics in a NVT ensemble. For equilibration, 5 ns of simulation at 1 atm were done in an isobaric-isothermal ensemble using the LangevinPiston method. The backbone was restrained with a harmonic potential of 10, 5, 2, and 1 kcal/mol $Å^2$ by 100 ps per stage to finally set it at 0.5 kcal/mol $Å^2$. For steered molecular dynamics simulation (SMD) all the atoms restraints around 15 Å of the aqueduct were disabled. MD simulations were done with a timestep of 1 fs with a 12 Å spherical cut-off for nonbonding interactions, a switching function from 10 Å for the van der Waals term, and force-shifted electrostatics. For long-range interactions, the particle-mesh Ewald method was used with a grid spacing of 1 Å or less. Periodic boundary conditions were employed for all simulations.

2.5. Free energy calculation

To locate polyP inside the Ppx cavity, a polyP of 6 residues (PP_6) was docked inside the channel and it was submitted to 5 ns molecular dynamics simulation. The resulted PP_6 was then further

extended to obtain polyP chains 20 residues (PP_{20}) to represent a short to medium range of polyP lengths. While PP_6 allows the interaction between all phosphates with protein residues within the channel, PP_{20} is long enough to consider the case when phosphate residues also interact with Ppx residues near the entry channel. For non-equilibrium SMD, polyP was pulled with a constant velocity of 10 Å/ns, from the first phosphorus atom when using PP_6 and from each of the phosphorus atoms when using PP_{20} to retain its structure through the pulling. As polyP pathway remains unknown the aqueduct principal axis was oriented into the Cartesian z-axis and used as reaction coordinate of pulling the polyP through the Ppx aqueduct. Eight to ten SMD replica were then used to calculate the potential of mean force using the Jarzinsky equality as described elsewhere [30,31].

2.6. Bacterial strains and growth conditions

E. coli strains were grown in LB medium at 37˚C. For recombinant *E. coli*, LB medium was supplemented with ampicillin 150 mg L^{-1}. *E. coli* XL10-Gold strain (Stratagene) was used for plasmid maintenance, *E. coli* BL21 Codon Plus strain (Stratagene) was used for protein expression.

2.7. DNA manipulation

Genomic and plasmidic extractions were both performed using commercial kits (Promega or Qiagen, resp.). Briefly, primer 1 and 2, from Table 1, were used to amplify *ecppx* with *E. coli* BL21 genomic DNA. Restriction enzymes *Nde*I and *Bam*HI were used in plasmid pET15b and PCR products, according to the manufacturer's instructions (Promega). DNA fragments were separated by agarose gel electrophoresis and purified from the gels employing a Zymoclean Gel DNA Recovery kit (Zymo Research). To obtain pET15-b:*ecppx* plasmid construct, fragments were ligated with T4 ligase and incorporated to *E. coli* XL10 by transformation. Point mutations were performed using the QuickChange Lightning Site-Directed Mutagenesis kit (Agilent Technologies) with the plasmid pET15:*ecppx* as template as described in Ref. [32]. The mutagenic primers used are listed in Table 1. To confirm the successful introduction of the gene or mutation, plasmids were sequenced using the services of CERELA (Tucumán, Argentina).

2.8. Protein purification and enzyme activity

Wild-type and H378A mutant variant of *ec*Ppx were expressed in *E. coli* BL21 CodonPlus (Stratagene) as a N-terminal fusion to a $MGSS(H)_6SSGLVPRGSH$ tag, and purified by affinity chromatography with a Ni^{2+} resin as described in Ref. [32]. The imidazole needed in the purification step was removed by dialysis in a Tris-HCl 10 mM buffer pH 8.0, NaCl 100 mM and EDTA 100 mM to avoid protein precipitation. After washing the imidazole, EDTA was reduced to 50 mM. Protein concentration was spectrophotometrically determined at λ_{280}, using the correspondent theoretical molar extinction coefficient calculated with the ProtParam tool [20] for physicochemical parameter prediction, which is available at the Expasyserver (http://www.expasy.com/).

All the kinetic determinations were performed under initial velocity conditions; typically 1–2 ng of enzyme was used in each assay. EDTA present in the enzyme stock was negligible compared to $MgCl_2$ according to Maxchelator calculations (http://maxchelator.stanford.edu/) [33], due to dilution effect. The standard assay to measure polyP activity was performed in 10 mM Tris-HCl buffer, pH 8.0 at 37˚C, with 1 mM Mg^{2+} and 160 mM KCl, at varying substrate concentration (Sodium Phosphate Glass, Sigma, with 25 -PP_{25}-, 65 -PP_{65}- or PPK synthesized polyP with 700 -PP_{700}- average number of residues). After polyP hydrolysis, the released phosphate was detected by the malachite green method at 630 nm as described in Ref. [34]. Data were fitted in Michaelis-Menten equation using Solver of LibreOffice Calc software [32].

Table 1
Cloning and mutagenic primers used in this work.

Name	Sequence
Up-NdeI-ecPpx[a]	5′- AACCTGAATAACCATATGCCA -3′
Dw-ecPpx-BamHI[a]	5′- CGGGATCCGAAAGTGCCTGAATAAT -3′
H378A[b]	5′- CGGGTTGAATATCAACGCCAGCGGTTTGCATCGCC -3′

[a] Primers for cloning *ppx* gen, the enzyme restriction site added is underlined.
[b] A pair of complementary primers were used to create each mutant, the sequence of the sense strand is shown, changes in the sequences are shown in underlined.

3. Results

We hypothesized that the terminal phosphate of the polyP chain would first bind to the electropositive channel. From here, the polyP chain would go across the aqueduct to finally reach the AS. To evaluate this proposal, we performed *in silico* assays; including docking, electrostatic potential calculations, molecular dynamics and evolutionary trace analysis, and also we complemented the study with *in vitro* assays such as site-directed mutagenesis and kinetic assays.

3.1. Favorable binding sites

As long chain polyP is highly negatively charged, its interaction with *ec*Ppx could be mainly driven by electrostatic energies. Electrostatic calculation and several docking assays were performed to find possible regions within *ec*Ppx with high affinity for polyP. In these assays, polyPs were represented as a short chain of 5 residues long (PP_5).

It was consistent the best conformations docked with the presence of positive electrostatic potential. The frontal surface of the protein had a clear predominance of electropositive potential (Fig. 2A) and the minimum binding energies were also obtained in that surface (Fig. 1A). This was favored by interaction with R^{166}, K^{197}, H^{382}, G^{380} and K^{414} (Fig. S1A). Other favorable region was formed by residues K^{353}, K^{428}, K^{429}, K^{430} and Q^{431} in the joint of domains I and IV (Fig. S1B). On the contrary, the binding energies of the back side of the protein surface were unfavorable to bind polyP (Fig. 1B), consistent with a predominance of electronegative potential (Fig. 2B). Indeed, *ec*Ppx was characterized by a clear division between electropositive (frontal) and electronegative (back) potential.

As expected, the conformations obtained within the AS had the most favorable binding energies in the N-terminal subunit (Fig. 1C). Remarkably, strong interactions were observed with residues S^{22}, H^{24}, K^{42}, Q^{43}, R^{44}, R^{75}, Q^{77} and R^{298} from domain I, which may form a electropositive semi-tunnel; and K^{221} and R^{267} from domain II in the area adjacent to the AS; resulting in very favorable binding energies (Fig. S1C). When comparing Figs. 1C and 2C it could be observed that the binding energies and the electrostatic potential are consistent.

3.2. Implication of H^{378} protonation state

To understand the hypothesis of Alvarado et al. [14], we paid attention to the amino acids forming the aqueduct of each monomer. When both aqueducts were separately analyzed, an interesting change in the protonation state of H^{378} (pKa 7.92 and 3.55)

was noted. This may probably be caused by the distance between H^{378} and the acidic residue E^{163} (2.9–7.5 Å). Furthermore, this difference was present in both crystal structures (monomer A and B or C and D in PDB:2FLO [15]; and A and B in PDB:1U6Z [14]). Fig. 2D, E, illustrate the effect of H^{378} protonation state on the electrostatic potential. When the residue was protonated, the potential varied gradually from positive to negative (Fig. 2E); on the contrary, when H^{378} was deprotonated, the variation in the electrostatic potential on each side of the aqueduct was abrupt (Fig. 2D).

3.3. Residues conservation

The characteristics and conserved residues of the ASKHA superfamily are well described and are present in the *ec*Ppx AS. On the other hand, as we pointed out, residues beyond the AS, including those involved in the interaction with polyP, have not yet been characterized. In order to predict the conserved residues along the surface of *ec*Ppx enzyme, we performed an evolutionary trace analysis.

Most of the residues belonging to the electropositive channel are well conserved, with values of 7 or 8 (Fig. 3A), i.e. R^{165}, K^{197}, H^{382}, R^{413} and N^{447}, although several differences were observed along the phylogenetic tree (Fig. S2). As it happens with the Ppx of *P. aeruginosa* (*pa*Ppx), residues of the channel from the N-terminal subunit are substituted by non-basic residues or even hydrophobic ones, i. e. R^{165} by leucine (Fig. S2A). In this sense, the total amount of basic residues may vary between different Ppx's. However, most of the arginines or lysines conserved belong to domain III (R^{410}, R^{412} and K^{413}, and in minor degree the R^{166}).

H^{370} and E^{371} from domain III are highly conserved, forming part of the same loop (H^{370} to G^{380}), and both belonging to the aqueduct. H^{378} is well conserved; although it varies among the proteins of the phylogenetic tree (Fig. S2B), which is divided into two distinct groups according to the residue present in that position: having histidine or having large hydrophobic or aromatic residues. On the other hand, the back view of *ec*Ppx, outside the aqueduct (mainly the H^{327}) is well conserved, and also the interface between domains I and II (Fig. 3B). Nevertheless, R^{288} and the region around it are highly variable.

Remarkably, the residues S^{316}, K^{353}, H^{356}, Q^{358}, R^{364}, W^{365}, the ^{428}KKK^{430} triplet and Q^{431} (site iv), which are interacting with the last tetrad of sulphate ions in the structure PDB:1U6Z [14], are highly variable and within them, only W^{365} is conserved.

The Fig. 3C) shows a conserved region of the N-terminal where AS is highly conserved. In addition, adjacent to AS, there are R^{29}, Q^{36}, R^{40} and R^{75} residues markedly conserved.

3.4. Effect of the polyP length on the traversal through the aqueduct

Several SMD simulations were performed to evaluate the

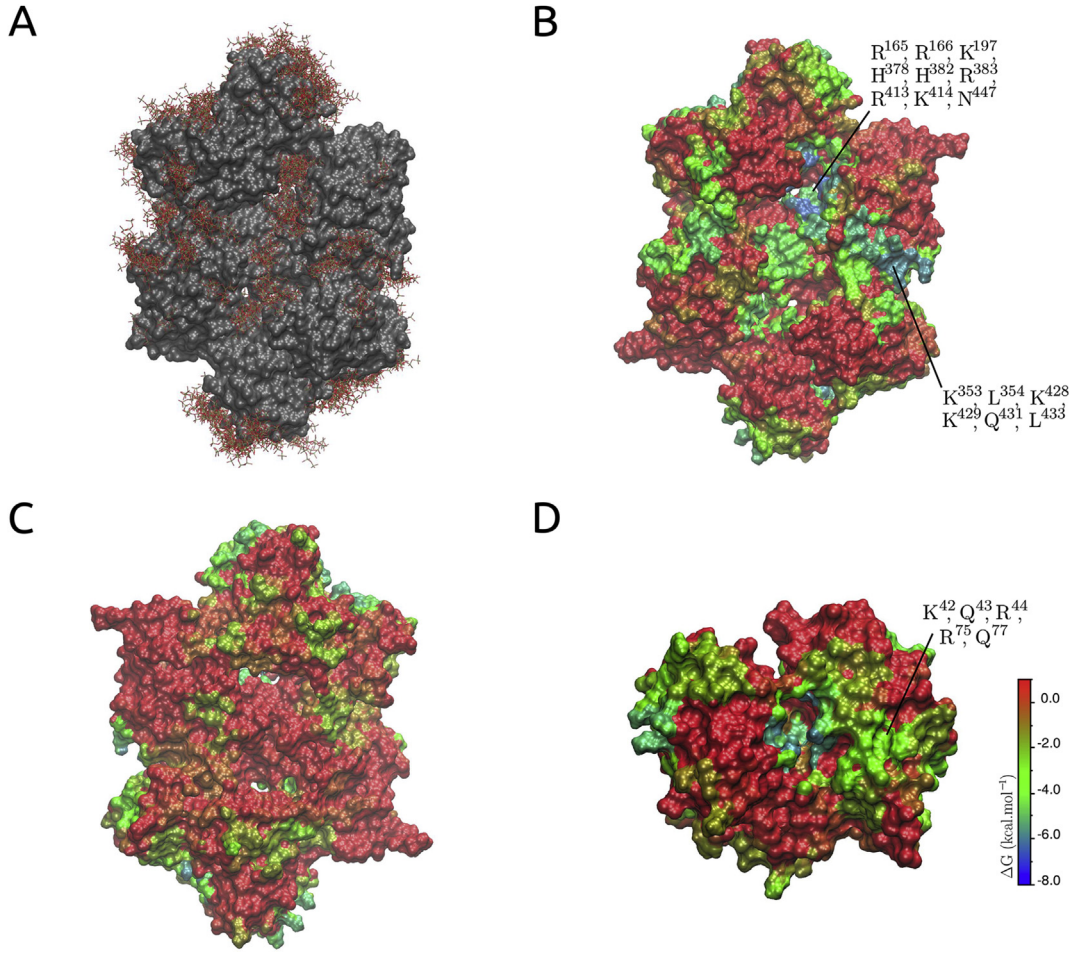

Fig. 1. Docking simulations assayed all over the protein surface. (A) conformations with $\Delta G \leq 0$. SSA of *ec*Ppx colored by minimal energy docked ligand: (B), (C) and (D) are front, back and AS views, respectively. Scale from −8 (blue) to 1 kcal/mol (red). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

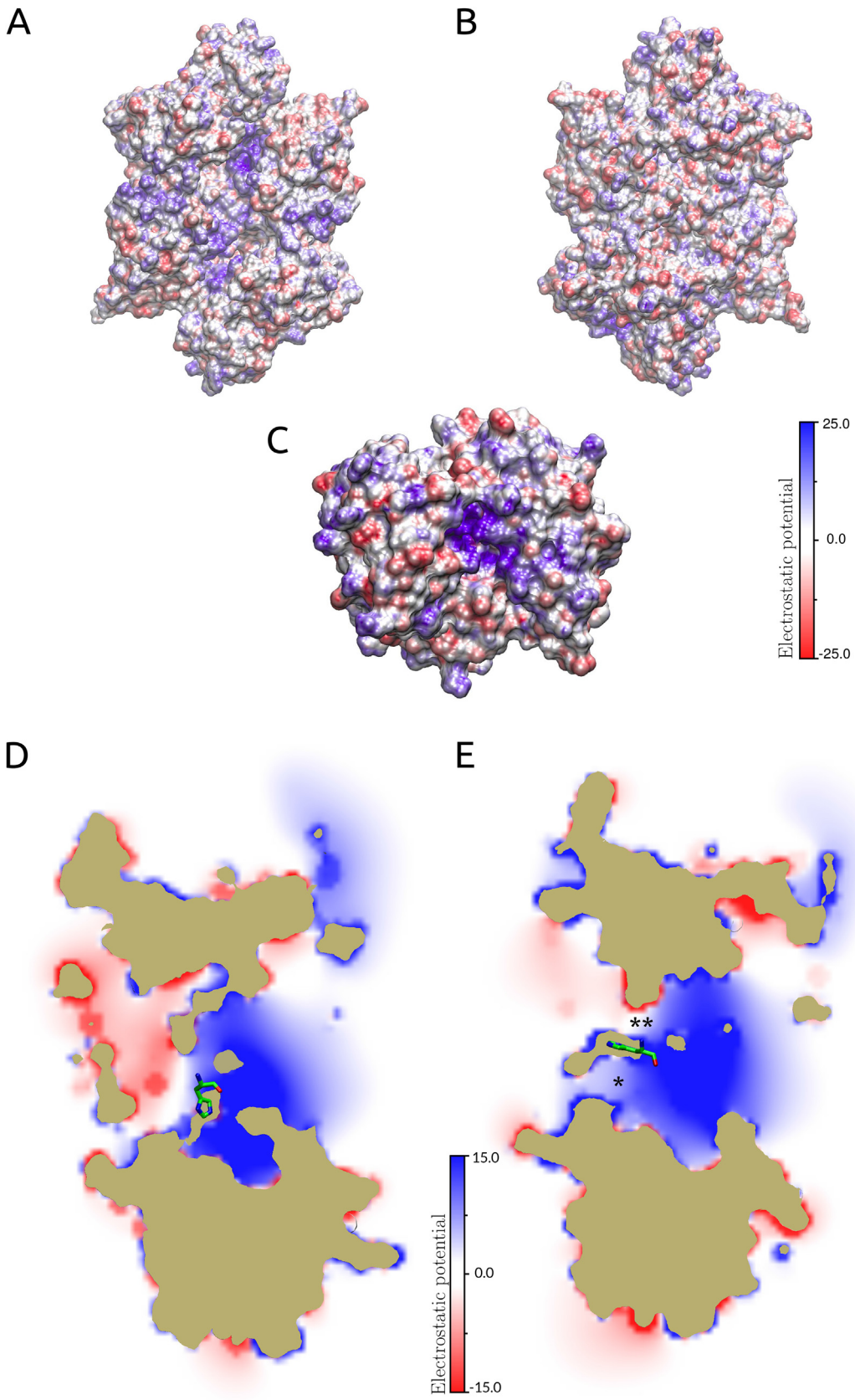

Fig. 2. SSA of *ec*Ppx colored by electrostatic potential: (A), (B) and (C) are front, back and AS views, respectively; (D) and (E) transversal slice from the aqueducts with H^{378} deprotonated and protonated, respectively. Asterisks in (E) indicates 2 pores separated by H^{378} and E^{163} salt bridge.

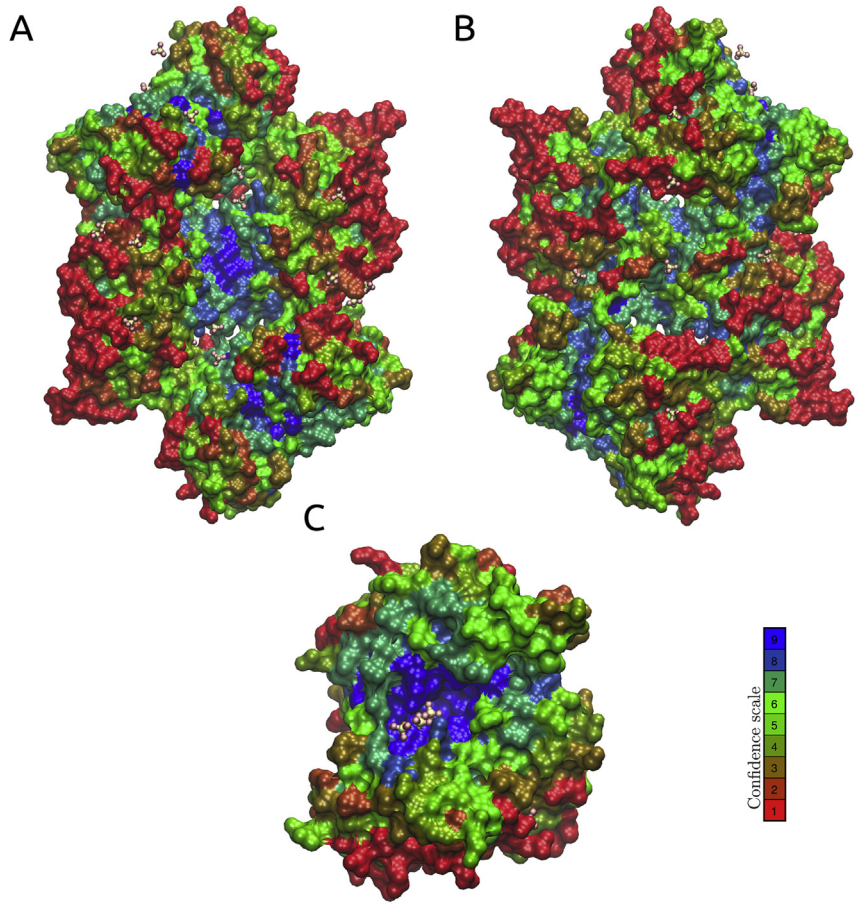

Fig. 3. SSA of *ec*Ppx colored by evolutionary conservation rate (from 1, for the less conserved, to 9, for the most conserved ones). (A), (B) and (C) are front, back and AS views, respectively.

potential of mean force (PMF) for passing through the aqueduct. After minimization and equilibration of the Ppx-polyP system obtained by docking assays, 8 to 10 runs of SMD simulation were performed, pulling out polyP on the phosphorus atom of the first residue from the channel towards the aqueduct and opposite to the binding direction. A ΔG ~100 kcal mol^{-1} was observed when a short chain of polyP (PP_6) was pulled toward and across the aqueduct. This barrier was reduced once PP_6 got close to H^{378} and sharply decreased when the first polyP residue passed this residue position. Two energy wells were observed when polyP interacted with basic residues. The first energy well corresponded to the loss of contact between PP_6 and R^{383} and to the interaction with T^{200}. The second corresponded to the unbinding of polyP and T^{200}, which was followed by the loss of interaction with any other residues on its path through the aqueduct. The energy required to bind PP_6 to the channel (ΔG_b) (from 35 to 5 Å) was significantly lower than the one necessary for passing through the aqueduct (ΔG_a) (Fig. 4A).

To understand the implication of the polyP length, SMD simulations were performed to describe the PMF of a PP_{20} mol, pulling all phosphorus atoms, instead of the phosphorus of the first residue as used for PP_6, in order to avoid polyP denaturation. Interestingly, the terminal phosphate was not the first residue in reaching the H^{378} location, remaining attached to the R^{413} at the first stages of the simulations. As shown in the Fig. 4B), there was an energy barrier of ~250 Kcal.mol^{-1} for PP_{20} along the pathway to the other side of the aqueduct, which became favorable once PP_{20} reached H^{378} and passed over it, maintaining a negative ΔG until the other extreme of the polyP left the channel.

3.5. H^{378} works as a gatekeeper

After cloning and purification of the WT *ec*Ppx and of the variant (H378A), enzyme activity was assayed with polyP of different chain length: PP_{25}, PP_{65} and PP_{700} which represents the physiological substrate.

The kinetic parameters showed that *ec*Ppx had preference for larger substrates since the higher the number of polyP residues, the smaller the K_m despite that the turnover number remained almost constant (Table 2). In consequence, the catalitic efficiency of Ppx was much higher with the physiological substrate.

The substitution of H^{378} by an alanine residue resulted in an active enzyme, but far less efficient than the WT variant and also unable to distinguish the length of polyP (Table 2). Even more, we were unable to measure K_m for PP_{700} since no saturation was achieved.

4. Discussion

Studies on the binding mechanism of polyP to Ppx are difficult to perform, as the enzyme has at least 2 binding sites in addition to the AS; only the later has been well studied, structurally and functionally. Since no structure could be obtained in the presence of substrate, in this work several *in silico* and biochemical studies were performed to clarify how long chain polyP binds to Ppx.

Our docking and electrostatic potential calculations pointed the AS (interface between domains I and II) and the electropositive channel (site iii) as the best binding sites. This result supports the

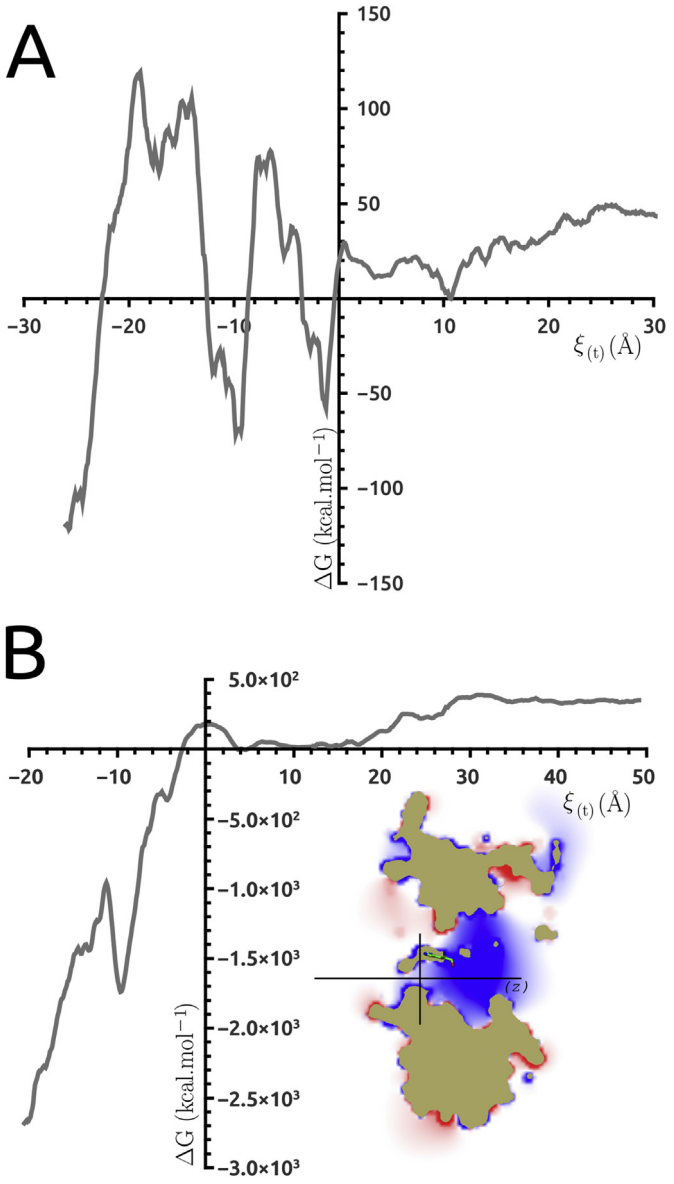

Fig. 4. PMF for a PP_6 (A) and PP_{20} (B) traversing the aqueduct. ΔG was calculated from 10 SMD simulations and averaged with cumulant equation [30,31].

suggestions made by Alvarado et al. [14] and Rangarajan et al. [15].

In a previous research, our laboratory working group has demonstrated differences between *ec*Ppx and *pa*Ppx, indicating that the later requires half of the KCl concentration to reach exopolyphosphatase maximal activity. One of the significant differences between both proteins is the electrostatic potential of the channel region. While in *ec*Ppx the channel is highly positive, by contribution of residues from domains II and III; residues corresponding to *pa*Ppx from domain III present positively charged residues but the residues from domain II at this position are essentially neutral (L^{171}, Q^{172} and E^{203} for *pa*Ppx instead of R^{165}, R^{166} and K^{197} for *ec*Ppx) [32]. This residue conservation pattern suggests that these differences are evolutionarily guided (Fig. S2), and may be related to enzyme regulation mediated by $[K^+]$.

On the other hand, four sulphates sequentially bounded to Ppx at site iv were observed into the SO_4/Ppx complex (PDB:1U6Z), therefore it was proposed as a binding site [14]. Docking and electrostatic assays may confirm such possibility; nevertheless, the

Table 2
Enzymatic parameters.

WT	PP25	PP65	PP700
K_m (nM)	260 ± 90	90 ± 10	3.5 ± 0.3
k_{cat} (s^{-1})	34 ± 2	35 ± 6	36 ± 6
Cat. eff. (s^{-1}. μM^{-1})	131	389	10,256
H378A			
K_m (μM)	3.0 ± 0.1	6.6 ± 0.9	ND
k_{cat} (s^{-1})	85.2 ± 0.7	38 ± 9	ND
Cat. eff. (s^{-1}. μM^{-1})	28.4	5.76	ND

evolutionary analysis done herein demonstrated a very low conservation of the site, suggesting that it is highly unlikely that polyP binds that patch.

4.1. Aqueduct and channel

Since polyP is an anionic polymer, its interaction with Ppx should be dominated by electrostatic interactions within the channels and protein surface. Moreover, Alvarado et al. [14] proposed that polyP should go across an aqueduct formed between the two Ppx monomers and due to the complexity of the binding complex, its interaction should occur at specific locations of the protein. To test this hypothesis, the physicochemical properties of Ppx were characterized by means of the surface electrostatic potential and polyP docking assays. Docking and electrostatic potential assays allowed the identification of 4 regions (2 for each monomer) on the Ppx surface located at the AS cleft and in the channel formed between domains II and III (Figs. 1 and 2). The highly electropositive characteristic together with the lowest favorable polyP binding energies, suggest that those sites are primary binding sites.

These results do not fit with the hypothesis of Alvarado et al. [14]. Considering that polyP should first bind to the electropositive channel and then pass through the aqueduct to finally reach the AS, and based on the results shown above, a negatively charged particle as polyP would have to move from an electropositive binding site, to a negative and unfavorable region. This would discard the polyP binding path proposed as a possible pathway of polyP through Ppx. However, a well conserved region was observed at the exit of the aqueduct (Fig. 3) and at the dimer and domain I and II interfaces. Obtained results encouraged us to continue studying Alvarado et al. [14] model and aqueduct function.

Transversal cuts of the aqueduct show different electrostatic potential depending on the H^{378} position and therefore on its protonation states. As shown in Fig. 2D, when histidine is unprotonated, the electrostatic potential changes from highly positive and favorable for polyP binding to a negative one and unfavorable for binding. However, the only change in the protonation state of histidine modifies the global potential and therefore the potential on both sides of the aqueduct varies gradually (Fig. 2E). Thus, H^{378} together with E^{163} are in the narrower region of the aqueduct forming a bottleneck and within the limit between electropositive/negative potential, which makes H^{378} a good candidate as a gatekeeper, changing its protonation state to bind or release polyP. When analyzing domain III separately, a semi-tunnel is found, which ranges from the electropositive channel (R^{440}) to the output of the aqueduct (H^{378}), (data not shown), suggesting the importance of the site and its affinity on polyP.

4.2. Only large polyP can transverse the aqueduct

According to suggestions performed in Alvarado et al. [14], based on site directed mutagenesis assays of Ppx (E371A), polyP

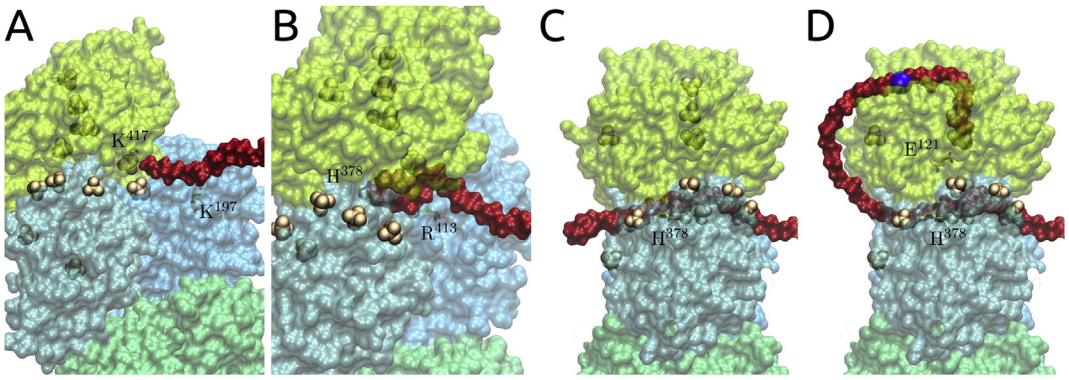

Fig. 5. Proposed binding mode for polyP. The scheme shows polyP (red) in the proximity of the channel (A); it follows the loop (B); then it traverses the aqueduct (C) and finally, it reaches the AS (D) (residue 14 of polyP counting from the AS is highlighted as a blue sphere. Sulphate ions are shown as light ocher spheres. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

would cross the aqueduct. Nevertheless, docking assays and electrostatic potential calculations presented herein dismiss that possibility. Several SMD simulations, in which polyP was pulled through the aqueduct, were performed in order to understand the function of the aqueduct and the residues in the surroundings on the polyP binding mechanism. To set the initial polyP coordinates inside Ppx a short polyP, PP_6, was docked inside the channel vestibule. As polyp could present a different angle-strain due to its highly flexible conformation, several reaction coordinates through the channel were probed from this position. Firstly, a PP_6 was pulled using as reaction coordinate the axis formed between the aquaduct and the adjacent pore located at the other side of the H^{378}, near the III-III domains interface (marked by two asterisks in Fig. 2E), however it traversed the aqueduct in the same way as when it is not forced (marked by an asterisk in Fig. 2E). Furthermore, when polyP is located near to the N-terminal subunit, an electronegative potential zone probably acts as a repulsive barrier and the ligand rapidly reaches the electropositive potential zone. On the contrary and in agreement with the positively-charged semi-tunnel proposed by Wei et al. [16], the lower energy conformations were found into the channel, where PP_6 interacts with R^{165}, R^{166}, K^{197}, H^{378}, H^{382}, R^{383}, K^{414} and K^{488} residues position in conjunction with the marked electropositive potential inside the channel suggest this zone as a possible polyP pathway to the aqueduct.

The PMF over the channel axis shows an unfavorable energy barrier which would impede the pass of polyP through the aqueduct, probably driven by a transversion of the electrostatic potential promoted by H^{378}. The energetic barrier sensed by polyP traversing the aqueduct should depend on its unbounded-phosphates within the protein surface; furthermore, the driven energy to move forward could be provided by the residues that immediately after bind the electropositive channel. To test this hypothesis, the PMF of pulling polyP through the aqueduct was calculated for a polyP of 20 residues long. The length of the chosen polyP ensures that whereas the first residues are bound to the Ppx surface, the remaining phosphates are still unbound, while PP_{20} advance through the aqueduct. When pulled from channel to solvent in opposite direction, a small energetic barrier of ΔG_{aq} ~250 kcal mol^{-1}, compared with the reverse pulling (ΔG_b ~ 350 kcal mol^{-1}), was noted. After the first residue of PP_{20} crossed the aqueduct narrowest region (E^{163}, H^{378}, N^{167}) and reached residue T^{200}, the landscape energy decreased to negative values. This differs from PP_6 where abrupt changes in the PMF were observed once the ligand traverse the aqueduct probably due to the absence of more phosphate residues binding to the channel.

One interesting fact was the way in which polyP approximated to H^{378}, not directly with the terminal phosphate, but forming a loop. This observation is in agreement with the structure obtained by Ref. [14] (PDB:1U6Z), as could be seen by the superposition of sulphate ions with the polyP residues linked by a loop (Fig. S3), which could fix the 6th to 10th residues and induce the first 5 residues to reach the H^{378} and then through the aqueduct, a critical number of residues being necessary to ensure that they modify the electrostatic potential of the environment, which could be a mechanism to induce the polyP to reach the histidine residue and then traverse the aqueduct.

When E^{371} was mutated by alanine, a 37% decrease on the activity was obtained [14]. However, such decrease does not necessarily explain an interruption in the processivity. Moreover, a DALI search [35] of structural similarity of Ppx domain III showed well conserved E^{371} and R^{449} residues forming a salt bridge in Ppx (data not shown). This arrangement would be functional to the maintenance of the integrity of the domain, rather than its function related to the traversing of polyP through the aqueduct.

Considering H^{378} as a gatekeeper between the N-terminal AS and the C-terminal long chain polyP binding site, we studied the enzymatic activity of Ppx against the substrates PP_{25}, PP_{65} and PP_{700}. In the same way as it was described by Beassoni et al. [32] for *pa*Ppx, *ec*Ppx preferred larger substrates, similarly to its homolog in *P. aeruginosa*. This is due to the fact that short polyP only binds to the AS for being hydrolyzed, while longer polyP could bind both sites at the same time, the electropositive channel to reach then the AS. In Beassoni et al. [32] it was also described a variant which contained only the N-terminal of *pa*Ppx and was unable to distinguish the length of the substrate based on the K_m parameter. In this sense, Bolesch and Keasling [1] described that mutant lacking the C-terminal released polyP of 14 residues long. Herein, we showed that the mutation of the H^{378} led to the interruption in the long chain polyP binding, which produced the loss of processivity with apparently no modification of the stronger binding site.

Taking into account results of our present study, we propose a polyP binding mode summarized in the following steps:

1. polyP binds to the electropositive channel, interacting with residues R^{166}, R^{410}, R^{413}, K^{414} and K^{417} (Fig. 5A).
2. While polyP binds to the electropositive channel and reaches the aqueduct, the terminal phosphate remains interacting with R^{413}, forming a loop (Fig. 5B).
3. The terminal phosphate enters into the aqueduct and interacts with H^{378} to then traverse it.

4. As polyP advances through the aqueduct, the next polyP residue binds the channel, making the advance energetically possible. In this step, the presence of K^+ is crucial.
5. Once the first residues leave the aqueduct, the negative surface on the back of the protein drives polyP away from it (Fig. 5C).
6. polyP continues to reach the surroundings of the AS (residues R^{42}, K^{44} and R^{75}), to then binds properly for catalysis at the AS (Fig. 5D).
7. polyP is hydrolyzed until the other end of polyP reaches the channel and is released from Ppx. The remaining PP_{40-50} are hydrolized by the N-terminal subunit (without C-terminal binding).

The points of agreement between this proposed mechanism and the ones of Alvarado et al. [14] and Rangarajan et al. [15] is that polyP binds to the electropositive channel and to the AS, and when compared with to Alvarado et al. [14], the commonality is the fact that polyP traverses the aqueduct to reach the AS. However, the low conservation of site iv and the way in which polyP binds to the channel suggest that this site is not implicated in polyP interaction. In the same way, site ii, in the electronegative face of Ppx, may be not involved in the binding of polyP, or at least, strong interactions may not occur, since the overall electronegative potential may promotes the advancement of polyP to reach the surroundings of AS. At this point, the relevance of the residues H^{24}, K^{42}, R^{44} and R^{75} near the AS (Fig. S1C) is clear, since they interact near residue 14 of polyP and could explain the appearance of fragments of 14 residues when only N-terminal subunit was used Bolesch and Keasling [1].

Finally, we do not emphasize on the AS, since its sequence and structure are very conserved and extensive to prokaryotic and eukaryotic organisms. However, the overall sequence is unique of Bacteria. The knowledge of the structure/function of the aqueduct and more precisely of H^{378} may be very useful for new drug discovery, focused not so much in the interruption of polyP catalysis, but in the diminution of polyP metabolism.

Author contributions

Cristhian Boetsch designed part of the research and carried out the *in silico* and biochemical studies and also wrote part of the manuscript. F. Danilo Gonzalez-Nilo designed part of the research and discussed the results. Daniel R. Aguayo and Á. Teresita Lisa discussed and analyzed part of the data. Paola R. Beassoni designed the research, analyzed and discussed the data and wrote the manuscript.

Funding

This work was supported by grants from SECYT-UNRC, CONICET (Argentina) and FONDECYT 11130576 (Chile).

Acknowledgments

Paola R. Beassoni and Angela T. Lisa are Career Members of the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). Cristhian Boetsch would like to acknowledge fellowship support from CONICET.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.abb.2016.07.005.

References

[1] D.G. Bolesch, J.D. Keasling, J. Biol. Chem. 275 (2000) 33814–33819, http://dx.doi.org/10.1074/jbc.M002039200.
[2] M. Akiyama, E. Crooke, A. Kornberg, J. Biol. Chem. 268 (1993) 633–639.
[3] J.D. Keasling, L. Bertsch, A. Kornberg, PNAS 90 (1993) 7029–7033.
[4] D.G. Bolesch, J. Keasling, Biochem. Biophys. Res. Commun. 274 (2000) 236–241.
[5] N.N. Rao, M.R. Gómez-García, A. Kornberg, Annu. Rev. Biochem. 78 (2009) 605–647.
[6] E. Crooke, M. Akiyama, N.N. Rao, A. Kornberg, J. Biol. Chem. 269 (1994) 6290–6295.
[7] L.A. Gallarato, D.G. Sánchez, L. Olvera, E.D. Primo, M.N. Garrido, P.R. Beassoni, E. Morett, A.T. Lisa, Microbiology 160 (2014) 406–417.
[8] N. Ogawa, C. Tzeng, C.D. Fraley, A. Kornberg, J. Bacteriol. 182 (2000) 6687–6693.
[9] M.H. Rashid, A. Kornberg, PNAS 97 (2000) 4885–4890.
[10] S.M. Thayil, N. Morrison, N. Schechter, H. Rubin, P.C. Karakousis, PLoS One 6 (2011) e2876.
[11] X. Shi, N.N. Rao, A. Kornberg, PNAS 101 (2010) 17061–17065.
[12] Q. Zhang, Y. Li, C.M. Tang, J. Biol. Chem. 285 (2010) 34259–34268.
[13] S.N. Lindner, S. Knebel, H. Wesseling, S.M. Schoberth, V.F. Wendisch, Appl. Environ. Microbiol. 75 (2009) 3161–3170.
[14] J. Alvarado, A. Ghosh, T. Janovitz, A. Jauregui, M.S. Hasson, D.A. Sanders, Structure 14 (2006) 1263–1272.
[15] S. Rangarajan, G. Nadeau, Y. Li, J. Wagner, M.N. Hung, J.D. Schrag, M. Cygler, A. Matte, J. Mol. Biol. 359 (2006) 1249–1260.
[16] Z.Z. Wei, G. Vatcher, A.H.Y. Tin, J.L. Teng, J. Wang, Q.H. Cui, J.G. Chen, A.C.H. Yu, PloS One 10 (2015) e0123713.
[17] H. Li, A.D. Robertson, J.H. Jensen, Proteins 61 (2005) 704–721.
[18] T.J. Dolinsky, P. Czodrowski, H. Li, J.E. Nielsen, J.H. Jensen, G. Klebe, N.A. Baker, Nucleic Acids Res. 35 (2007) 522–525.
[19] T.J. Dolinsky, J.E. Nielsen, J.A. McCammon, N.A. Baker, Nucleic Acids Res. 32 (2004) W665–W667.
[20] N.A. Baker, D. Sept, S. Joseph, M.J. Holst, J.A. McCammon, PNAS 98 (2001) 10037–10041.
[21] G.M. Morris, R. Huey, A.J. Olson, Curr. Protoc. Bioinforma. (2008) 8–14.
[22] J. Gasteiger, M. Marsili, Tetrahedron 36 (1980) 3219–3228.
[23] G.M. Morris, D.S. Goodsell, R.S. Halliday, R. Huey, W.E. Hart, R.K. Belew, A.J. Olson, J. Comput. Chem. 19 (1998) 1639–1662.
[24] W. Humphrey, A. Dalke, K. Schulten, J. Molec. Graph. 14 (1996) 33–38.
[25] R.D. Finn, J. Clements, S.R. Eddy, Nucleic Acids Res. (2011) gkr367, http://dx.doi.org/10.1093/nar/gkr367.
[26] D.T. Jones, W.R. Taylor, J.M. Thornton, Comput. Appl. Biosci. 8 (1992) 275–282.
[27] F. Glaser, T. Pupko, I. Paz, R.E. Bell, D. Bechor-Shental, E. Martz, N. Ben-Tal, Bioinf. Appl. Note 19 (2003) 163–164.
[28] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, L. Skeel, K. Schulten, J. Comput. Chem. 26 (2005) 1781–1802.
[29] A.D. MacKerell, D. Bashford, M. Bellott, R.L. Dunbrack, J.D. Evanseck, M.J. Field, S. Fisher, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, F. Kuczera, F.T.K. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, M. Karplus, J. Phys. Chem. B 102 (1998) 3586–3616.
[30] G.E. Crooks, J. Stat. Phys. 90 (1998) 1481–1487.
[31] S. Park, F. Khalili-Araghi, E. Tajkhorshid, K. Schulten, J. Chem. Phys. 119 (2003) 3559–3566.
[32] P.R. Beassoni, L.A. Gallarato, C. Boetsch, M.N. Garrido, A.T. Lisa, Enzyme Res. 2015 (2015), http://dx.doi.org/10.1155/2015/404607.
[33] C. Patton, Pacific Grove, 1994.
[34] A.A. Baykov, O.A. Evtushenko, S.M. Avaeva, Anal. Biochem. 171 (1988) 266–270, http://dx.doi.org/10.1016/0003-2697(88)90484-8.
[35] L. Holm, C. Sander, Trends Biochem. Sci. 20 (1995) 478–480.