Evidence for the Existence of Elaborate Enzyme Complexes in the Paleoarchean Era

Bernd Reisinger,[†] Josef Sperl,[†] Alexandra Holinski,[†] Veronika Schmid,[†] Chitra Rajendran,[†] Linn Carstensen,[†] Sandra Schlee,[†] Samuel Blanquart,[‡] Rainer Merkl,*[,†] and Reinhard Sterner*[,†]

[†]Institute of Biophysics and Physical Biochemistry, University of Regensburg, Universitätsstraße 31, D-93053 Regensburg, Germany
[‡]Equipe Bonsai, Institut National de Recherche en Informatique et en Automatique, INRIA Lille Nord Europe, 40 avenue Halley, 59650 Villeneuve d'Ascq, France

 S *Supporting Information*

ABSTRACT: Due to the lack of macromolecular fossils, the enzymatic repertoire of extinct species has remained largely unknown to date. In an attempt to solve this problem, we have characterized a cyclase subunit (HisF) of the imidazole glycerol phosphate synthase (ImGP-S), which was reconstructed from the era of the last universal common ancestor of cellular organisms (LUCA). As observed for contemporary HisF proteins, the crystal structure of LUCA-HisF adopts the $(\beta\alpha)_8$-barrel architecture, one of the most ancient folds. Moreover, LUCA-HisF (i) resembles extant HisF proteins with regard to internal 2-fold symmetry, active site residues, and a stabilizing salt bridge cluster, (ii) is thermostable and shows a folding mechanism similar to that of contemporary $(\beta\alpha)_8$-barrel enzymes, (iii) displays high catalytic activity, and (iv) forms a stable and functional complex with the glutaminase subunit (HisH) of an extant ImGP-S. Furthermore, we show that LUCA-HisF binds to a reconstructed LUCA-HisH protein with high affinity. Our findings suggest that the evolution of highly efficient enzymes and enzyme complexes has already been completed in the LUCA era, which means that sophisticated catalytic concepts such as substrate channeling and allosteric communication existed already 3.5 billion years ago.
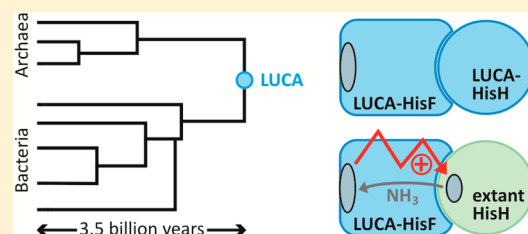
INTRODUCTION

Modern enzyme complexes are elaborate molecular machineries that have been optimized in the course of evolution for the efficient and specific processing of their substrates. One prominent example is the imidazole glycerol phosphate synthase (ImGP-S), a bienzyme complex which belongs to the family of glutamine amidotransferases[1] and constitutes a branch point connecting amino acid and nucleotide biosynthesis. ImGP-S consists of the cyclase subunit HisF and the glutaminase subunit HisH. HisF binds the substrate N'-[(5'phosphoribulosyl)formimino]-5-aminoimidazole-4-carboxamide-ribonucleotide (PRFAR) and performs a cycloligase/lyase reaction that generates imidazole glycerol phosphate (ImGP) and 5-aminoimidazole-4-carboxamide ribotide (AICAR), which are further used in histidine and de novo purine biosynthesis, respectively[2] (Figure 1). The ammonia molecule required for this transformation is produced by the glutaminase subunit HisH and transported to the active site of HisF through an extended molecular channel. This channeling hampers diffusion of ammonia into bulk solvent and thus presumably prevents its protonation to the nonproductive ammonium ion. Another specific feature of the HisF/HisH complex is the tight coordination of the two enzymatic activities: Binding of PRFAR (or its analogue N'-[(5'-phosphoribosyl)formimino]-5-aminoimidazole-4-carboxamide-ribonucleotide, ProFAR) to HisF results in an allosteric signal that leads to a several-hundred-fold stimulation of the

glutaminase activity of HisH.[3−5] This property precludes the hydrolysis of glutamine by HisH in the absence of an acceptor substrate at the active site of HisF.[6]

We were interested to find out whether the characteristics of modern HisF enzymes were already present in those species that colonized Earth in a very early phase of biological evolution. A straightforward answer to this question is difficult due to the lack of macromolecular fossils. However, computational techniques of amino acid sequence reconstruction[7,8] make it possible to travel back in time and to study extinct proteins.[9−16] In extreme cases, these algorithms enable us to study enzymes from the last universal common ancestor of cellular organisms (LUCA), which preceded the diversification of life and existed in the Paleoarchean era, i.e., at least 3.5 billion years ago.[17]

Along these lines, we have previously computationally reconstructed the amino acid sequence of HisF from the LUCA era (LUCA-HisF).[18] To this end, a set of 87 extant HisF and HisH proteins from the seven phylogenetic clades Crenarchaeota, Actinobacteria, Chlorobi, Cyanobacteria, Firmicutes, Proteobacteria, and Thermotogae has been used to determine a phylogenetic tree t_{HisF_HisH} based on the CAT model[19] (Supporting Information, Figure S1). After having rooted this tree between the superkingdoms Archaea and
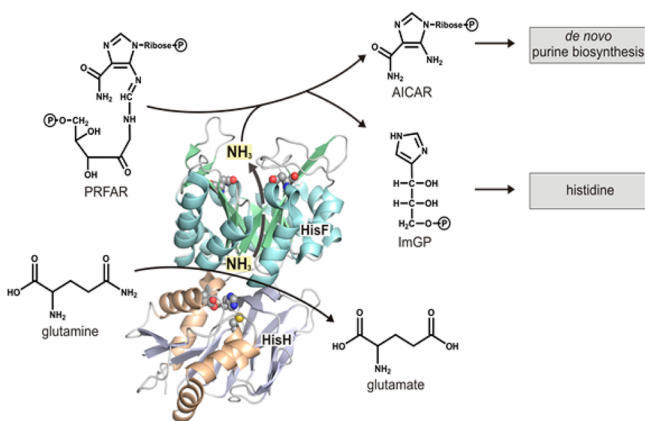
Figure 1. Reaction catalyzed by the heterodimeric ImGP synthase complex. The synthase subunit HisF catalyzes the reaction of N'-[(5′phosphoribulosyl)formimino]-5-aminoimidazole-4-carboxamide-ribonucleotide (PRFAR) with ammonia (NH_3) to imidazole glycerol phosphate (ImGP) and 5-aminoimidazole-4-carboxamide ribotide (AICAR). ImGP is further utilized in the synthesis of histidine, whereas AICAR is an intermediate in de novo purine biosynthesis, rendering HisF a branch-point enzyme of amino acid and nucleotide biosynthesis. The ammonia molecule required for the HisF reaction is produced by the glutaminase subunit HisH (catalytic triad residues are depicted as spheres) and subsequently channeled to the active site of HisF (catalytic aspartate residues are depicted as spheres). In the absence of HisH, HisF can also use external ammonia that is added as ammonia salt. HisF adopts the $(\beta\alpha)_8$-barrel fold, an ubiquitous and catalytically versatile protein architecture,[21] which is considered one of the three most ancient protein folds.[22] HisH adopts the α/β hydrolase fold.[23]

Bacteria, we reconstructed a predecessor of HisF from Bacteria and Crenarchaeota as described.[19,20] Thus, although the precise lineage of the three superkingdoms is still under debate,[24] LUCA-HisF is among the oldest so far reconstructed proteins, if not the oldest hitherto calculated predecessor.[16,25,26] Among the 87 descendants of LUCA-HisF used for reconstruction, 78 of the 250 residues are less than 50% conserved, whereas 49 residues are strongly conserved. Accordingly, LUCA-HisF differs in 55 amino acids (22%) from the closest BLAST match,[27] which is HisF from *Thermovibrio ammonificans*. The nucleotide and amino acid sequences of LUCA-HisF are given in the Supporting Information.

We have now produced LUCA-HisF in *Escherichia coli*, and analyzed its crystal structure, conformational stability, folding mechanism, and catalytic activity. The observed molecular characteristics of LUCA-HisF turned out to be similar to contemporary HisF proteins. Moreover, LUCA-HisF activates an extant HisH protein and thus comprises all elements required for allosteric interaction. Finally, we have also reconstructed and produced a LUCA-HisH protein and could show that it binds to LUCA-HisF with high affinity. Taken together, our results suggest that the protein inventory of the LUCA already contained elaborate enzyme complexes.

■ METHODS

Cloning, Expression, and Purification of LUCA-HisF. The gene coding for LUCA-HisF was optimized for its expression in *E. coli*, synthesized (GeneArt), and cloned into the vector pET24a(+) (Stratagene) using the terminal restriction sites for *Nde*I and *Xho*I. Since the addition of a C-terminal hexahistidine tag to LUCA-HisF might influence its interaction with HisH proteins, a stop codon was integrated at the end of the gene. The gene was expressed in *E. coli* T7-

Express cells (New England Biolabs) transformed with pET24a(+)-LUCA-*his*F. To this end, 4 L of Luria broth (LB) medium supplemented with 75 μg/mL kanamycin were inoculated with a preculture and incubated at 37 °C. After an OD_{600} of 0.6 was reached, the temperature was lowered to 30 °C. Expression was induced by adding 0.5 mM IPTG, and growth was continued overnight. Cells were harvested by centrifugation (Sorvall/RC5B, GS3, 15 min, 4000 rpm, 4 °C), washed with 50 mM potassium phosphate, pH 7.5, and centrifuged again. The cells were suspended in the same buffer, lysed by sonication (Branson Sonifier W-250D, 2 × 2 min in 15 s intervals, 45% pulse, 0 °C), and centrifuged again (Sorvall/RC5B, SS34, 30 min, 13.000 rpm, 4 °C) to separate the soluble from the insoluble fraction of the cell extract. In a first step, the soluble supernatant was subjected to ion exchange chromatography using a MonoQ column (HR 16/10, 20 mL, Pharmacia), which had been equilibrated with 50 mM potassium phosphate, pH 7.5. The column was washed with equilibration buffer, and bound LUCA-HisF was eluted by applying a linear gradient of 0−1.5 M NaCl. Protein-containing fractions were pooled, dialyzed against 50 mM potassium phosphate, pH 7.5, and subjected to ammonia sulfate precipitation. After 80% saturation with ammonia sulfate, precipitated protein was centrifuged (Sorvall/RC5B, SS34, 30 min, 13.000 rpm, 4 °C), dissolved in 50 mM potassium phosphate, pH 7.5, and 300 mM potassium chloride, and finally purified via size exclusion chromatography. For this purpose a Superdex200 column (HiLoad 26/60, 320 mL, GE Healthcare) was operated with 50 mM potassium phosphate, pH 7.5, and 300 mM potassium chloride at 4 °C. Fractions with pure protein were pooled and dialyzed against 50 mM Tris·HCl, pH 7.5. According to SDS-PAGE (12.5% acrylamide), LUCA-HisF was more than 95% pure. About 30 mg of protein was obtained per liter of culture.

In order to determine the binding properties of LUCA-HisF to HisH proteins via fluorescence titration, all tryptophan residues of LUCA-HisF were replaced by tyrosines. Hence, LUCA_*his*F_W138Y+W156Y was generated via overlap extension polymerase chain reaction (PCR)[28] using pET24a(+)-LUCA-*his*F as a template (see the Supporting Information for oligonucleotide sequences), and subsequently cloned into pET24a(+) via the terminal restriction sites for *Nde*I and *Xho*I. Expression and purification were performed as described for LUCA-HisF, yielding a comparable amount and purity of LUCA-HisF_W138Y+W156Y.

Cloning, Expression, and Purification of zmHisH. Genomic DNA of *Zymomonas mobilis* (DSM424) was ordered from the Leibniz Institute DSMZ. In order to remove the internal restriction site for *Nde*I, the zm*his*H gene was amplified by overlap extension PCR[28] (see the Supporting Information for oligonucleotide sequences) and cloned into pET24a(+) using the terminal restriction sites for *Nde*I and *Xho*I. After transformation of *E. coli* strain BL21(DE3) (Stratagene), expression was carried out at 30 °C overnight in 4 L of LB medium, supplemented with 75 μg/mL kanamycin. Protein purification was performed as described for LUCA-HisF including ion exchange chromatography using 50 mM Tris·HCl, pH 9, as buffer, ammonia sulfate precipitation, size exclusion chromatography, and final dialysis against 50 mM Tris·HCl, pH 7.5. According to SDS-PAGE (12.5% acrylamide), zmHisH was more than 95% pure. About 8 mg of protein was obtained per liter of culture.

Sequence Reconstruction, Cloning, Expression and Purification of LUCA-HisH. As for LUCA-HisF, the reconstruction of LUCA-HisH was based on the tree t_{HisF_HisH} (Supporting Information, Figure S1), which is close to an accepted organism phylogeny. In comparison to the multiple sequence alignment (MSA) of extant HisF sequences, the 87 extant HisH sequences exhibit a significantly higher variability. In fact, 140 of 226 residues are less than 50% conserved. Furthermore, the MSA ($HisH_{ext}$) contains several gaps. Recently, it has been shown that a novel algorithm for the phylogeny-aware gap placement named PRANK[29] improves MSA quality. This is why we used PRANK with the option −showanc to deduce LUCA-HisH from the MSA $HisH_{ext}$ under the control of t_{HisF_HisH} (Supporting Information, Figure S2). The nucleotide and amino acid sequences of LUCA-HisH are given in the Supporting Information. The protein

shares 123 of 226 residues (54%) with the closest BLAST[27] match, which is HisH from *Syntrophothermus lipocalidus*.

The gene coding for LUCA-HisH was optimized for its expression in *E. coli*, synthesized (GeneArt), and cloned into the vector pET24a(+) (Stratagene) using the terminal restriction sites for *Nde*I and *Xho*I. (The gene encodes a C-terminal hexahistidine tag; see the Supporting Information.) Subsequently, pET24a(+)-LUCA-*his*H was used to transform *E. coli* strain BL21-Gold (DE3) (Stratagene). Protein expression, harvesting of cells, and cell lysis were performed as described for LUCA-HisF. As LUCA-HisH showed a high thermal stability, most of the host proteins could be removed by heat denaturation (70 °C, 15 min) followed by centrifugation (Sorvall/RC5B, SS34, 30 min, 13.000 rpm, 4 °C). For further purification, the supernatant of the heat step was loaded onto a HisTrapFF crude column (5 mL; GE Healthcare), which had been equilibrated with 50 mM potassium phosphate, pH 7.5, 300 mM potassium chloride, and 10 mM imidazole. After washing with equilibration buffer, the bound protein was eluted by applying a linear gradient of 10−375 mM imidazole. Fractions with pure protein were pooled, and LUCA-HisH was dialyzed against 10 mM potassium phosphate, pH 7.5. As judged by SDS-PAGE, the protein was more than 95% pure. About 26 mg of LUCA-HisH was obtained per liter of culture.

Crystallization, Data Collection, and Refinement of LUCA-HisF. Crystallization trials were carried out using the PEG/Ion screen (Hampton Research). The hanging drop vapor diffusion method was performed in 96-well plates (Greiner) at 291 K. Drops contained 300 nL of the respective reservoir buffer mixed with 300 nL of LUCA-HisF (13.9 mg/mL) in 10 mM potassium phosphate, pH 7.5. In each well equilibration was performed against 100 μL of reservoir buffer. Crystals were obtained with 0.2 M sodium phosphate monobasic monohydrate, pH 4.7, and 20 wt %/vol PEG 3350. After flash freezing in liquid nitrogen, data of single crystals were collected at the synchrotron beamline PX2 (SLS) at 100 K. Data were processed using XDS,[30] and the data quality assessment was done using phenix.xtriage.[31] Molecular replacement was performed with MOLREP within the CCP4i suite.[32] A homology model of LUCA-HisF with HisF from *Thermotoga maritima* (*tm*HisF) (PDB ID 1THF) was built with MODELLER[33] and served as a search model. Initial refinement was performed using REFMAC.[34] The model was further improved in several refinement rounds using automated restrained refinement with the program PHENIX[31] and interactive modeling with Coot.[35] The data collection and refinement statistics are summarized in Table S1 in the Supporting Information. The final model was analyzed using the program MolProbity.[36]

Analysis of the Thermal Stability of LUCA-HisF and LUCA-HisH. Differential scanning calorimetry (DSC) was performed with LUCA-HisF in 50 mM potassium phosphate, pH 7.5, by heating the sample in a CSC 5100 Nano differential scanning calorimeter with a scan rate of 1 °C min^{-1}. The DSC data were analyzed with the program CpCalc (version 2.1; Calorimetry Sciences Corp., 1995) to determine apparent melting temperatures (T_M^{app} DSC). Thermal denaturation traces of LUCA-HisF and LUCA-HisH in 50 mM potassium phosphate, pH 7.5, were monitored with a JASCO J-815 circular dichroism (CD) spectrometer in a 0.1 cm cuvette by following the loss of ellipticity at 220 nm. Unfolding was induced by raising the temperature in 1 °C increments at a ramp rate of 1 °C min^{-1} with a Peltier-effect temperature controller. The midpoint temperatures of the unfolding transitions (T_M^{app} CD) were determined. Data are shown in Figure S3 in the Supporting Information. The irreversibility of the denaturation traces precluded the thermodynamic analyses of the DSC and CD unfolding measurements.

Equilibrium Unfolding/Refolding Transitions and Formation of a Burst-Phase Intermediate by LUCA-HisF. The thermodynamic stability of LUCA-HisF was determined by guanidinium chloride (GdmCl) induced equilibrium unfolding transitions. The loss of tertiary structure was probed by protein fluorescence; the loss of secondary structure was probed by far-UV CD. Samples with 2 μM protein were prepared in 50 mM Tris·HCl buffer (pH 7.5) containing different concentrations of GdmCl. GdmCl (ultrapure) was purchased from MP Biomedicals (Illkirch, France), and its concentration was

determined by the refractive index of the solution.[37] To reach equilibrium, LUCA-HisF was preincubated at the indicated concentration of GdmCl for 24 h at 25 °C.

The fluorescence emission signal at 320 nm (bandwidth 5 nm) after excitation at 280 nm (bandwidth 3 nm) was monitored with a JASCO Model FP-6500 spectrofluorimeter. The equilibrium unfolding transition of LUCA-HisF obtained by monitoring fluorescence is shown in Figure S4A in the Supporting Information in comparison to *tm*HisF and its artificially designed precursors Sym1 and Sym2, which were constructed by duplication and fusion of the C-terminal half-barrel HisF-C followed by the optimization of the initial construct.[38−40] The transitions were analyzed according to the two-state equilibrium model, assuming a linear dependency of the free-energy of unfolding on the GdmCl concentration.[41] The obtained values for ΔG_D, m, and $[D]_{1/2}$ are listed in Table S2 in the Supporting Information.

The far-UV circular dichroism (CD) signal at 225 nm was monitored using a JASCO Model J815 CD spectrophotometer (path length 5 mm; bandwidth 1 nm). The equilibrium unfolding/refolding transitions of LUCA-HisF obtained by monitoring the far-UV CD signal are shown in Figure S4B in the Supporting Information.

Kinetics of refolding of LUCA-HisF in Figure S4B in the Supporting Information were obtained by following the far-UV CD signal for 200 s in manual mixing experiments at various concentrations of GdmCl and extrapolating the exponential curve to zero time. The observed amplitude was plotted as a function of the GdmCl concentration and is shown in Figure S4B in the Supporting Information. It was significantly lower than the amplitude observed in the refolding equilibrium transitions, indicating that the major part of the CD change occurred within the dead time of the manual mixing experiment. This is interpreted with the formation of a compact burst-phase refolding intermediate with a high content of secondary structure.

Fluorescence Titration of *zm*HisH and LUCA-HisH with LUCA-HisF. Fluorescence titration was used to determine the binding stoichiometry and affinity of the LUCA-HisF/*zm*HisH and LUCA-HisF/LUCA-HisH complexes, as in the course of complex formation a tryptophan residue lying at the HisH interface is shielded from the solvent.[5] Hence, when titrating either 7 μM *zm*HisH or 5 μM LUCA-HisH in 50 mM potassium phosphate, pH 7.5, with LUCA-HisF_W138Y+W156Y, the emission maxima shifted from 345 to 325 nm and from 345 to 329 nm, respectively (excitation at 295 nm). The decreases in fluorescence emission at 318 nm were plotted against the added amounts of LUCA-HisF_W138Y+W156Y, and the resulting curves were analyzed with a quadratic fit. Both titrations were performed in triplicate.

Analysis of Enzymatic Activity in Vitro and in Vivo. In vitro enzymatic activities were determined by steady-state kinetics. The ammonia- and glutamine-dependent conversions of PRFAR into ImGP and AICAR (HisF reaction) were measured spectrophotometrically at 300 nm as previously described.[5] At 25 °C, entire progress curves at four different PRFAR concentrations were recorded either in 50 mM Tris-acetate, pH 8.5, in the presence of 100 mM ammonium acetate (*ammonia-dependent cyclase reaction*) or in 50 mM Tris-acetate, pH 8.0, in the presence of 15 mM glutamine and 2 μM *zm*HisH (*glutamine-dependent cyclase reaction*). In both cases an excess of HisA from *T. maritima* was added in order to synthesize PRFAR in situ from ProFAR,[44] and 0.5 μM LUCA-HisF was used to initiate the measurements. Data were analyzed with the integrated form of the Michaelis−Menten equation using the program COSY[45] to obtain k_{cat} and K_M^{PRFAR}. In case of the LUCA-HisF/LUCA-HisH complex (10 μM), no *glutamine-dependent cyclase activity* could be determined in the presence of 10 mM glutamine and 100 μM ProFAR. The glutaminase activity of *zm*HisH (1 μM) in complex with liganded LUCA-HisF (2 μM; 40 μM ProFAR) was measured in a coupled enzymatic assay as previously described.[5] At 25 °C, produced glutamate was oxidized by a molar excess of glutamate dehydrogenase (Roche) in 50 mM Tricine hydroxide, pH 8.0. Thus, the reduction of the coenzyme NAD$^+$ to NADH could be monitored spectrophotometrically at 340 nm. Three glutamine saturation curves were recorded and fitted with the
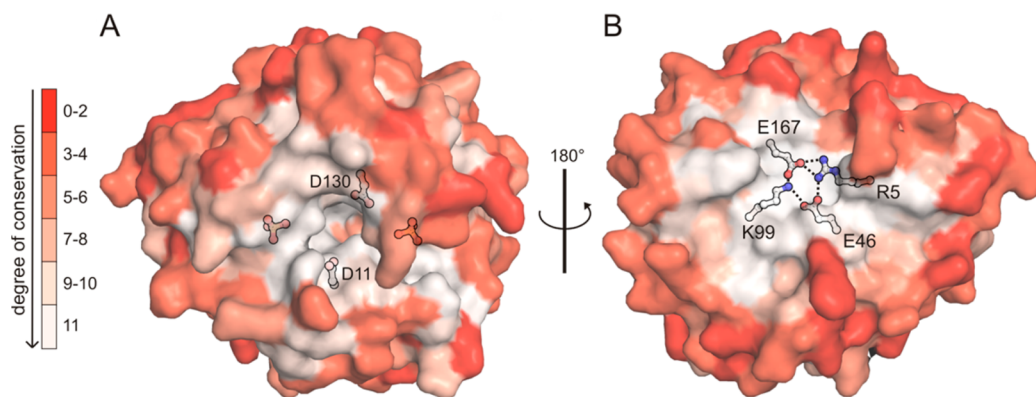
Figure 2. Crystal structure of LUCA-HisF. The surface is color coded according to residue conservation deduced from the MSA used for reconstruction. Conservation values [0−11] were determined by means of Jalview;[42] strictly conserved residues are white. (A) Catalytic face of HisF and view along the ammonia channel. The catalytically important aspartate residues D11 and D130 as well as two bound phosphate ions, which mimic the phosphate moieties of the substrate PRFAR, are shown as sticks. (B) Stability face and ammonia tunnel gate at the bottom of the β-barrel. The salt bridge cluster between the residues R5, E46, K99, and E167 (depicted as sticks; electrostatic interactions indicated by dashed lines) defines the entrance to the ammonia channel.[43]

Michaelis−Menten equation to obtain k_{cat} and K_M^{Gln}. In an identical setup, no glutaminase activity (12 mM glutamine) could be detected for LUCA-HisH (20 μM) in complex with ligand-bound LUCA-HisF (20 μM; 200 μM ProFAR). The extent to which LUCA-HisF liganded with ProFAR activates zmHisH had to be determined in a discontinuous assay, since NAD⁺ also exhibits a stimulating effect on glutaminase activity.[6] To this end, 10 mM glutamine was incubated at 25 °C with 0.5 μM zmHisH and 5 μM LUCA-HisF either in the absence of ProFAR or in the presence of 40 μM ProFAR. Aliquots of 150 μL of the reaction mixture were collected after 15, 30, 45, and 60 (only in the absence of ProFAR) min and spun through a 10 kDa filter (Roth) to remove the enzymes. The V_{max} value was calculated from the linear increase of glutamate production with time, which was determined with the help of 1 mg/mL glutamate dehydrogenase and 0.7 mM APAD⁺ (Sigma) (the reaction mixture was diluted 1:7.5, and absorption was measured at 363 nm). All measurements were performed in triplicates.

ProFAR to PRFAR isomerization activity (HisA reaction) was measured with the enzymatic assay described for the ammonia-dependent HisF reaction, however, in the presence of an excess of HisF.[46] PRA to CdRP isomerization activity (TrpF reaction) was followed at 25 °C by a fluorimetric assay (excitation at 350 nm, emission at 400 nm).[47,48] The substrate PRA was generated in situ by 1 μM yeast anthranilate phosphoribosyl transferase from anthranilate and PRPP, which was provided in a 30-fold molar excess. Moreover, 2.5 μM indole-3-glycerol phosphate synthase from *T. maritima* was added to prevent product inhibition.

To test for enzymatic activity in vivo, the gene coding for LUCA-HisF was subcloned into the pTNA vector, which allows for constitutive expression in *E. coli*.[49] The resulting pTNA-LUCA-*hisF* plasmid was used to transform cells of auxotrophic Δ*hisF*, Δ*hisA*, or Δ*trpF E. coli* strains.[50,51] These strains lack the *hisF*, *hisA*, or *trpF* gene on their chromosome and are, therefore, unable to grow on medium without histidine or tryptophan, respectively. Growth experiments and controls were performed as described.[50]

RESULTS AND DISCUSSION

Structure Determination of LUCA-HisF. The gene coding for LUCA-HisF was synthesized, cloned into a plasmid, and expressed in *E. coli*. The LUCA-HisF protein was predominantly found in the soluble fraction of the host cell extract, and purified in a three-step process using ion exchange chromatography, ammonia sulfate precipitation, and size exclusion chromatography. Purified LUCA-HisF was crystallized and its three-dimensional structure was determined at 1.48

Å resolution by molecular replacement based on the structure of tmHisF[52] (Figure 2, Table S1 in the Supporting Information). LUCA-HisF (PDB ID 4EVZ) adopts the conserved $(\beta\alpha)_8$-barrel structure observed in the three extant HisF proteins from *Pyrobaculum aerophilum* (PDB ID 1H5Y), *Thermus thermophilus* (PDB ID 1KA9), and *T. maritima* (PDB ID 1THF), for which crystal structures have been previously determined. The superposition of LUCA-HisF with each of these structures by means of STAMP[53] resulted in an overall root-mean-square deviation (rmsd) ranging from 1.14 to 1.43 Å. In agreement with the postulated evolution of the $(\beta\alpha)_8$-barrel fold from a $(\beta\alpha)_4$-half-barrel,[54,55] LUCA-HisF displays a clear 2-fold symmetry: The superposition of its N-terminal $[(\beta\alpha)_{1−4}]$ and C-terminal $[(\beta\alpha)_{5−8}]$ halves yielded an rmsd of 1.68 Å, which is similar to the corresponding values for the three extant HisF proteins (1.27 Å for 1H5Y, 1.52 Å for 1KA9, and 1.69 Å for 1THF). Consistent with the internal symmetry, the two catalytically important aspartate residues[5] are found on opposite sides of the active site at the C-terminal ends of β-strand 1 and β-strand 5. Likewise, the two cocrystallized phosphate groups, which represent the two phosphate groups of the substrate PRFAR (Figure 1), are anchored by the C-terminal ends of β-strands 3 and 4, and β-strands 7 and 8, respectively (Figure 2A). Moreover, a stabilizing salt-bridge cluster at the N-terminal end of the β-barrel, which contains four charged and invariant residues that form the gate to the cyclase ammonia channel,[6,38,43] is also present in LUCA-HisF (Figure 2B).

Stability and Folding Mechanism of LUCA-HisF. The thermal stability of LUCA-HisF was determined by differential scanning calorimetry (DSC), which monitors overall unfolding, and the heat-induced decrease of the far-UV CD signal, which indicates the loss of secondary structure. The combination of both methods showed that thermal unfolding of LUCA-HisF is a two-step process with apparent transition midpoints of about 70 and 100 °C (Supporting Information, Figure S3A,B). These results characterize LUCA-HisF as an enzyme with a high resistance to heat. Interestingly, even higher denaturation temperatures were previously observed for enzymes from the common ancestors of Bacteria, Archaea, and Archaea/Eukaryota.[16] These findings and our results are interesting in the light of rRNA and protein sequence analyses which have

provided independent support for the increase of thermotolerance from the LUCA to the ancestors of Bacteria and Archaea/Eukaryota.[56] Furthermore, the conformational stability of LUCA-HisF was analyzed by GdmCl-induced equilibrium unfolding and refolding transitions. The loss or gain of tertiary structure was probed by protein (Tyr/Trp) fluorescence. The equilibrium unfolding and refolding curves superpose well, which proves the reversibility of unfolding (Figure 3).
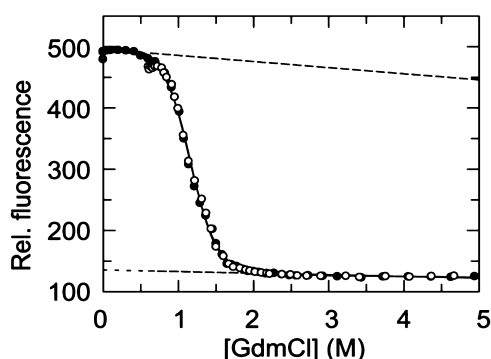

Figure 3. GdmCl-induced equilibrium unfolding/refolding transitions of LUCA-HisF. The transitions were followed by Trp/Tyr fluorescence (excitation at 280 nm; emission at 320 nm) in 50 mM Tris·HCl buffer, pH 7.5. Closed symbols represent the unfolding experiment, started with folded protein, and open symbols represent the refolding experiment, started with protein that was previously unfolded in 6.0 M GdmCl. The continuous line represents a fit to the unfolding transition on the basis of the two-state model. The dashed lines indicate the baselines for the pure N and U states. The thermodynamic parameters deduced from the analysis are given in the text and listed in Table S2 in the Supporting Information.

Moreover, the transitions are adequately described by the two-state model,[41] indicating that no significant amounts of stable equilibrium intermediates are populated. The analysis yielded an m-value of ~15 kJ mol^{-1} M^{-1}, a transition midpoint ($[D]_{1/2}$) at 1.2 M GdmCl, and a free energy of unfolding in the absence of denaturant (ΔG_D) of 18 kJ mol^{-1} (Supporting Information, Table S2). LUCA-HisF has a lower ΔG_D but a comparably high m-value as tmHisF and its artificially designed precursors Sym1 and Sym2,[39,40] indicating that it is comparably compact as these proteins but less stable (Supporting Information, Figure S4A, Table S2). Folding and unfolding

kinetics followed by Tyr/Trp fluorescence showed that the reduced stability of LUCA-HisF is due to strongly increased *un*folding rates of LUCA-HisF in comparison to *T. maritima* HisF (tmHisF) (Supporting Information, Figure S5A). The comparison of the *re*folding kinetics of LUCA-HisF, tmHisF, Sym1, and Sym2 followed by fluorescence and far-UV CD (Supporting Information) showed that all four proteins share a common sequential folding mechanism including a non-productive burst-phase intermediate (Supporting Information, Figure S4B) and two productive intermediates (Supporting Information, Figure S5). The rate-limiting step that synchronizes folding is conserved (Supporting Information, Figure S6).

Catalytic Activity of LUCA-HisF. The enzymatic activity of LUCA-HisF was measured in vitro using steady-state kinetics. The analysis of PRFAR to ImGP/AICAR progress curves obtained in the presence of saturating concentrations of externally added ammonia (*ammonia-dependent cyclase activity*) yielded a catalytic efficiency (k_{cat}/K_M^{PRFAR}) of 2.8 × 10^5 M^{-1} s^{-1}, which is similar to the catalytic efficiency of 3.3 × 10^5 M^{-1} s^{-1} that was obtained for tmHisF (Table 1). As ancient enzymes have been proposed to be less specific (more promiscuous) than their modern descendants,[57] we tested LUCA-HisF for its ability to catalyze related metabolic reactions. The homologous enzyme HisA, which precedes HisF in the histidine biosynthesis pathway, catalyzes the Amadori rearrangement of ProFAR to PRFAR. HisA shares with HisF the overall $(\beta\alpha)_8$-barrel fold as well as the location of the two symmetry-related catalytic aspartate residues and phosphate binding sites.[52] Phosphoribosyl anthranilate (PRA) isomerase (TrpF) catalyzes an Amadori rearrangement in tryptophan biosynthesis analogous to HisA in histidine biosynthesis.[49] Remarkably, a single amino acid exchange in the HisA and HisF proteins from *T. maritima* leads to TrpF activity, suggesting that these three phosphate-binding $(\beta\alpha)_8$-barrel proteins have evolved from a common precursor.[48,58] We examined LUCA-HisF for the isomerization activity toward ProFAR and PRA. However, no substrate turnover could be detected, even in the presence of 50 μM protein. These findings were complemented by assessing catalytic activity in vivo using metabolic selection. For this purpose, a plasmid harboring the LUCA-HisF gene was used to transform auxotrophic *E. coli* strains lacking either the intrinsic *his*F, *his*A, or *trp*F gene. When plated on minimal medium without histidine or tryptophan, the ΔhisF cells formed visible colonies within 24 h, whereas the ΔhisA and ΔtrpF cells did not

Table 1. Steady-State Kinetic Constants of the ImGP Synthase Pairs LUCA-HisF/zmHisH and tmHisF/tmHisH

ammonia-dependent cyclase activity[a]	k_{cat}, s^{-1}	K_M^{PRFAR}, μM	k_{cat}/K_M^{PRFAR}, M^{-1} s^{-1}
LUCA-HisF	0.078 (±0.003)	0.29 (±0.04)	2.8 (±0.3) × 10^5
tmHisF[b]	1.2	3.6	3.3 × 10^5
glutamine-dependent cyclase activity[c]	k_{cat}, s^{-1}	K_M^{PRFAR}, μM	k_{cat}/K_M^{PRFAR}, M^{-1} s^{-1}
LUCA-HisF/zmHisH	0.058 (±0.006)	0.36 (±0.07)	1.6 (±0.3) × 10^5
tmHisF/tmHisH[b]	1.1	2.0	5.5 × 10^5
glutaminase activity[d]	k_{cat}, s^{-1}	K_M^{Gln}, mM	k_{cat}/K_M^{Gln}, M^{-1} s^{-1}
LUCA-HisF/zmHisH	0.21 (±0.03)	1.9 (±0.9)	1.2 (±0.3) × 10^2
tmHisF/tmHisH[b]	0.1	0.8	1.3 × 10^2
stimulation of glutaminase activity[d]	k_{cat}, s^{-1} (without ProFAR)	k_{cat}, s^{-1} (ProFAR satd)	stimulation factor[e]
LUCA-HisF/zmHisH	3.85 (±0.04) × 10^{-2}	0.483 (±0.006)	13
tmHisF/tmHisH[b]	3.3 × 10^{-4}	0.1	303

[a]Reaction conditions: 50 mM Tris-acetate buffer, pH 8.5, at 25 °C. [b]Data taken from ref 6. [c]Reaction conditions: 50 mM Tris-acetate buffer, pH 8.0, at 25 °C. [d]Reaction conditions: 50 mM Tricine hydroxide, pH 8.0, at 25 °C. [e]The stimulation factor is the quotient k_{cat}(ProFAR saturated)/ k_{cat}(without ProFAR).

grow within 1 week. Taken together, these results suggest that LUCA-HisF is a monofunctional enzyme.

Formation of LUCA-HisF/HisH Complexes. In order to test whether LUCA-HisF contains all structural elements required for complex formation, substrate channeling, and allosteric communication, we assayed its functional interaction with the extant *zm*HisH enzyme from *Zymomonas mobilis*. For this purpose, *zm*HisH was produced in *E. coli* and purified. The binding of *zm*HisH to LUCA-HisF was analyzed via fluorescence titration,[5] which showed that the two proteins form a stoichiometric complex with a thermodynamic dissociation constant (K_D) of 113 nM (Figure 4A).

The steady-state kinetic constants k_{cat} and K_M^{PRFAR} of LUCA-HisF in the presence of *zm*HisH and saturating concentrations of glutamine (*glutamine-dependent cyclase activity*) compare well with the above-reported *ammonia-dependent cyclase activity* (Table 1). This outcome confirms the functionality of the LUCA-HisF/*zm*HisH complex, as ammonia produced at the active site of HisH by means of glutamine hydrolysis is used as efficiently by LUCA-HisF as externally added ammonia. Moreover, this finding suggests that ammonia is transported from HisH to the active site of the synthase through a molecular channel formed by the central β-barrel of LUCA-HisF, as observed for extant HisF enzymes.[43,59] Furthermore, glutamine hydrolysis by *zm*HisH in the presence of LUCA-HisF and saturating concentrations of the substrate analogue ProFAR (*glutaminase activity*) is as efficient as glutaminase activity of HisH from *T. maritima* (*tm*HisH) in complex with ProFAR-liganded *tm*HisF[4] (Table 1). The comparison of the *zm*HisH activity in the presence and absence of ProFAR indicates a 13-fold stimulation by the HisF-ligand in this non-native complex (Figure 4B), which is 23-fold lower than the stimulating effect of ProFAR in the native *tm*HisF/*tm*HisH complex (Table 1).

Following the characterization of LUCA-HisF, we also reconstructed the amino acid sequence of the corresponding glutaminase LUCA-HisH. Again, we used the tree t_{HisF_HisH} but opted for the phylogeny-aware gap placement of PRANK[29] to deduce LUCA-HisH from the MSA $HisH_{ext}$, which contains several insertions and deletions. The gene coding for LUCA-HisH was synthesized, cloned into a plasmid, and expressed in *E. coli*. The produced protein was soluble and could be purified by a combination of heat denaturation and Ni^{2+} affinity chromatography. As observed for LUCA-HisF, LUCA-HisH exhibits a high thermotolerance. Unfolding followed by CD resulted in a single transition with a midpoint of about 79 °C (Supporting Information, Figure S3C). Complex formation between LUCA-HisH and LUCA-HisF was probed by fluorescence titration.[5] Both proteins interacted stoichiometrically with very high affinity as demonstrated by a K_D value of 4 nM (Figure 4C). However, when testing the LUCA-HisF/LUCA-HisH complex for *glutamine-dependent cyclase activity* or LUCA-HisH for the hydrolysis of glutamine in the presence of LUCA-HisF and saturating concentrations of ProFAR, no enzymatic turnover could be determined. Thus, unlike LUCA-HisF, LUCA-HisH is catalytically inactive. As outlined in the following, uncertainties in the reconstruction process are probably responsible for this finding.

The evolutionary models underlying reconstruction consider each residue position independently of all other positions. Thus, the reliability of a given reconstruction is limited not by sequence length, but by the composition of the MSA and the topology of the deduced phylogenetic tree. In the case of
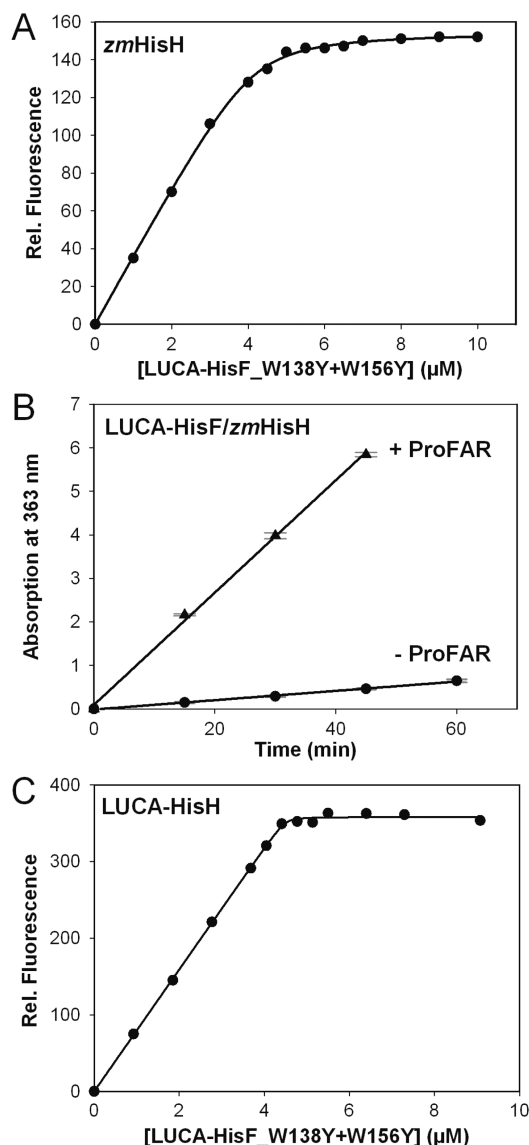

Figure 4. Fluorescence titration curve of *zm*HisH and LUCA-HisH with LUCA-HisF and activation of *zm*HisH by LUCA-HisF. (A) LUCA-HisF_W138Y+W156Y was added to 7 μM *zm*HisH in 50 mM potassium phosphate, pH 7.5, and 25 °C. Fluorescence emission at 318 nm was determined following excitation at 295 nm. Quadratic fits of the obtained data points resulted in a K_D value of 113 (±1) nM. The stoichiometry is slightly deviating from a 1:1 complex, indicating that a small fraction of *zm*HisH is not active. (B) Glutaminase activity of the LUCA-HisF/*zm*HisH complex in the absence (circles) and presence (triangles) of ProFAR was tested in a discontinuous assay (see the Supporting Information for detailed information). Mean values and standard deviations of triplicate measurements are shown. Glutamine (10 mM) was incubated with 0.5 μM *zm*HisH and 5 μM LUCA-HisF at 25 °C in both cases. Glutaminase activity is enhanced 13-fold in the presence of ProFAR (see Table 1). (C) Titration of 5 μM LUCA-HisH with LUCA-HisF_W138Y+W156Y was performed and analyzed analogous to (A), yielding a stoichiometric complex with a K_D value of 4 (±2) nM.

LUCA-HisF, 49 out of 250 residues are strictly conserved; among them are the two active site aspartate residues[5] and amino acids contributing to the central ammonia channel[38,43] (Figure 2). Furthermore, the four central nodes of t_{HisF_HisH} (Supporting Information, Figure S1) possess posterior probabilities of ≥0.88. Taken together, these features suggest

that tree topology and choice of the most likely residues for the corresponding predecessors and LUCA-HisF is largely unambiguous. In contrast, in the case of LUCA-HisH, only 21 out of 226 residues are strictly conserved, which makes the reconstruction much more prone to uncertainties.

Even more than a certain conservation of amino acid sequence composition, conservation of sequence length is an important prerequisite for a valid reconstruction. Along these lines, the sequence lengths of extant and reconstructed thioredoxins, which is the only other example for a fully functional enzyme from the LUCA era,[16] are very similar. In contrast, MSA $HisH_{ext}$ contains several gaps. Nonetheless, the phylogeny-aware gap placement by means of PRANK did allow us to reconstruct a stable LUCA-HisH protein with a fully functional protein−protein interface, albeit lacking enzyme activity. Obviously, t_{HisF_HisH} was sufficiently informative to reconstruct ancestral residues at positions whose role did not change during evolution such as the catalytic triad[3] and residues involved in binding of the substrate glutamine, as deduced from the structure of tmHisH.[6] In contrast, reconstruction seems to have failed at residue positions that underwent frequent changes during evolution due to insertions and deletions. It has to be shown that highly articulated phylogenetic trees will enable us to reconstruct the correct series of indels and to further improve reconstruction for such difficult cases.

CONCLUSIONS

Taken together, LUCA-HisF, which presumably existed about 3.5 billion years ago, is similar to extant HisF proteins with respect to structure, stability, folding, and activity. Since similar results were obtained for predecessors of thioredoxin,[16] experimental evidence accumulates for the existence of highly effective enzymes in the LUCA era. In addition, LUCA-HisF forms a stable complex with LUCA-HisH and a functional enzyme complex with the extant glutaminase zmHisH. It is therefore plausible to assume that the evolution of the ImGP-S complex, including ammonia channeling and allosteric communication, had been completed in the LUCA era. Thus, our experimental findings are in line with the hypothesis that the LUCA had already a rather diverse metabolism, which was as sophisticated as the metabolisms of its archaeal and bacterial successors are.[60]

ASSOCIATED CONTENT

ⓢ Supporting Information

Nucleotide and amino acid sequences of LUCA-HisF and LUCA-HisH. Oligonucleotides used for the construction of LUCA_hisF_W138Y+W156Y and for the genomic amplification of zmHisH. Comparison of the folding mechanism of LUCA-HisF, $T.$ $maritima$ HisF, and its artificial precursors Sym1 and Sym2. Crystal structure determination of LUCA-HisF (PDB ID 4EVZ): data collection and refinement statistics (Table S1). Thermodynamic unfolding parameters of LUCA-HisF, $T.$ $maritima$ HisF, Sym1, and Sym2 (Table S2). Phylogenetic tree t_{HisF_HisH} used for the reconstruction of LUCA-HisF and LUCA-HisH (Figure S1). PRANK output for the reconstruction of LUCA-HisH (Figure S2). Thermal denaturation of LUCA-HisF and LUCA-HisH (Figure S3). Equilibrium unfolding transitions of LUCA-HisF, $T.$ $maritima$ HisF, Sym1, and Sym2, and formation of a burst-phase intermediate by LUCA-HisF (Figure S4). Apparent rate constants (λ) and amplitudes of refolding and unfolding kinetics of LUCA-HisF (Figure S5). Unifying folding

mechanism for LUCA-HisF, $T.$ $maritima$ HisF, and Sym1 and Sym2 (Figure S6). References for Supporting Information. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
Rainer.Merkl@ur.de; Reinhard.Sterner@ur.de
Author Contributions
B.R. and J.S. contributed equally. All authors have given approval to the final version of the manuscript.
Notes
The authors declare no competing financial interest.

ACKNOWLEDGMENTS

REFERENCES

(1) Zalkin, H. $Adv.$ $Enzymol.$ $Relat.$ $Areas$ $Mol.$ $Biol.$ **1993**, 66, 203.
(2) Chittur, S. V.; Chen, Y.; Davisson, V. J. $Protein$ $Expression$ $Purif.$ **2000**, 18, 366.
(3) Myers, R. S.; Jensen, J. R.; Deras, I. L.; Smith, J. L.; Davisson, V. J. $Biochemistry$ **2003**, 42, 7013.
(4) Klem, T. J.; Davisson, V. J. $Biochemistry$ **1993**, 32, 5177.
(5) Beismann-Driemeyer, S.; Sterner, R. $J.$ $Biol.$ $Chem.$ **2001**, 276, 20387.
(6) List, F.; Vega, M. C.; Razeto, A.; Haeger, M. C.; Sterner, R.; Wilmanns, M. $Chem.$ $Biol.$ **2012**, 19, 1589.
(7) Hanson-Smith, V.; Kolaczkowski, B.; Thornton, J. W. $Mol.$ $Biol.$ $Evol.$ **2010**, 27, 1988.
(8) Harms, M. J.; Thornton, J. W. $Curr.$ $Opin.$ $Struct.$ $Biol.$ **2010**, 20, 360.
(9) Benner, S. A.; Sassi, S. O.; Gaucher, E. A. $Adv.$ $Enzymol.$ $Relat.$ $Areas$ $Mol.$ $Biol.$ **2007**, 75, 1.
(10) Thomson, J. M.; Gaucher, E. A.; Burgan, M. F.; De Kee, D. W.; Li, T.; Aris, J. P.; Benner, S. A. $Nat.$ $Genet.$ **2005**, 37, 630.
(11) Chang, B. S.; Jonsson, K.; Kazmi, M. A.; Donoghue, M. J.; Sakmar, T. P. $Mol.$ $Biol.$ $Evol.$ **2002**, 19, 1483.
(12) Malcolm, B. A.; Wilson, K. P.; Matthews, B. W.; Kirsch, J. F.; Wilson, A. C. $Nature$ **1990**, 345, 86.
(13) Stackhouse, J.; Presnell, S. R.; McGeehan, G. M.; Nambiar, K. P.; Benner, S. A. $FEBS$ $Lett.$ **1990**, 262, 104.
(14) Jermann, T. M.; Opitz, J. G.; Stackhouse, J.; Benner, S. A. $Nature$ **1995**, 374, 57.
(15) Finnigan, G. C.; Hanson-Smith, V.; Stevens, T. H.; Thornton, J. W. $Nature$ **2012**, 481, 360.
(16) Perez-Jimenez, R.; Inglés-Prieto, A.; Zhao, Z. M.; Sanchez-Romero, I.; Alegre-Cebollada, J.; Kosuri, P.; Garcia-Manyes, S.; Kappock, T. J.; Tanokura, M.; Holmgren, A.; Sanchez-Ruiz, J. M.; Gaucher, E. A.; Fernandez, J. M. $Nat.$ $Struct.$ $Mol.$ $Biol.$ **2011**, 18, 592.
(17) Nisbet, E. G.; Sleep, N. H. $Nature$ **2001**, 409, 1083.
(18) Richter, M.; Bosnali, M.; Carstensen, L.; Seitz, T.; Durchschlag, H.; Blanquart, S.; Merkl, R.; Sterner, R. $J.$ $Mol.$ $Biol.$ **2010**, 398, 763.
(19) Lartillot, N.; Philippe, H. $Mol.$ $Biol.$ $Evol.$ **2004**, 21, 1095.
(20) Blanquart, S.; Lartillot, N. $Mol.$ $Biol.$ $Evol.$ **2008**, 25, 842.
(21) Wierenga, R. K. $FEBS$ $Lett.$ **2001**, 492, 193.
(22) Caetano-Anollés, G.; Kim, H. S.; Mittenthal, J. E. $Proc.$ $Natl.$ $Acad.$ $Sci.$ $U.S.A.$ **2007**, 104, 9358.
(23) Ollis, D. L.; Cheah, E.; Cygler, M.; Dijkstra, B.; Frolow, F.; Franken, S. M.; Harel, M.; Remington, S. J.; Silman, I.; Schrag, J.; Sussman, J. L.; Verschueren, K. H. G.; Goldman, A. $Protein$ $Eng.$ **1992**, 5, 197.
(24) Gribaldo, S.; Poole, A. M.; Daubin, V.; Forterre, P.; Brochier-Armanet, C. $Nat.$ $Rev.$ $Microbiol.$ **2010**, 8, 743.

(25) Gaucher, E. A.; Govindarajan, S.; Ganesh, O. K. *Nature* **2008**, *451*, 704.

(26) Risso, V. A.; Gavira, J. A.; Mejia-Carmona, D. F.; Gaucher, E. A.; Sanchez-Ruiz, J. M. *J. Am. Chem. Soc.* **2013**, *135*, 2899.

(27) Boratyn, G. M.; Camacho, C.; Cooper, P. S.; Coulouris, G.; Fong, A.; Ma, N.; Madden, T. L.; Matten, W. T.; McGinnis, S. D.; Merezhuk, Y.; Raytselis, Y.; Sayers, E. W.; Tao, T.; Ye, J.; Zaretskaya, I. *Nucleic Acids Res.* **2013**, *41*, W29.

(28) Ho, S. N.; Hunt, H. D.; Horton, R. M.; Pullen, J. K.; Pease, L. R. *Gene* **1989**, *77*, 51.

(29) Löytynoja, A.; Goldman, N. *Science* **2008**, *320*, 1632.

(30) Kabsch, W. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 125.

(31) Adams, P. D.; Grosse-Kunstleve, R. W.; Hung, L. W.; Ioerger, T. R.; McCoy, A. J.; Moriarty, N. W.; Read, R. J.; Sacchettini, J. C.; Sauter, N. K.; Terwilliger, T. C. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2002**, *58*, 1948.

(32) Potterton, L.; McNicholas, S.; Krissinel, E.; Gruber, J.; Cowtan, K.; Emsley, P.; Murshudov, G. N.; Cohen, S.; Perrakis, A.; Noble, M. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 2288.

(33) Sali, A.; Blundell, T. L. *J. Mol. Biol.* **1993**, *234*, 779.

(34) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1997**, *53*, 240.

(35) Emsley, P.; Cowtan, K. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 2126.

(36) Davis, I. W.; Leaver-Fay, A.; Chen, V. B.; Block, J. N.; Kapral, G. J.; Wang, X.; Murray, L. W.; Arendall, W. B., 3rd; Snoeyink, J.; Richardson, J. S.; Richardson, D. C. *Nucleic Acids Res.* **2007**, *35*, W375.

(37) Pace, C. N. *Methods Enzymol.* **1986**, *131*, 266.

(38) Höcker, B.; Lochner, A.; Seitz, T.; Claren, J.; Sterner, R. *Biochemistry* **2009**, *48*, 1145.

(39) Carstensen, L.; Sperl, J. M.; Bocola, M.; List, F.; Schmid, F. X.; Sterner, R. *J. Am. Chem. Soc.* **2012**, *134*, 12786.

(40) Carstensen, L.; Zoldak, G.; Schmid, F. X.; Sterner, R. *Biochemistry* **2012**, *51*, 3420.

(41) Santoro, M. M.; Bolen, D. W. *Biochemistry* **1988**, *27*, 8063.

(42) Waterhouse, A. M.; Procter, J. B.; Martin, D. M.; Clamp, M.; Barton, G. J. *Bioinformatics* **2009**, *25*, 1189.

(43) Douangamath, A.; Walker, M.; Beismann-Driemeyer, S.; Vega-Fernandez, M. C.; Sterner, R.; Wilmanns, M. *Structure* **2002**, *10*, 185.

(44) Thoma, R.; Obmolova, G.; Lang, D. A.; Schwander, M.; Jeno, P.; Sterner, R.; Wilmanns, M. *FEBS Lett.* **1999**, *454*, 1.

(45) Eberhard, M. *Comput. Appl. Biosci.* **1990**, *6*, 213.

(46) Jürgens, C.; Strom, A.; Wegener, D.; Hettwer, S.; Wilmanns, M.; Sterner, R. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 9925.

(47) Hommel, U.; Eberhard, M.; Kirschner, K. *Biochemistry* **1995**, *34*, 5429.

(48) Leopoldseder, S.; Claren, J.; Jürgens, C.; Sterner, R. *J. Mol. Biol.* **2004**, *337*, 871.

(49) Henn-Sax, M.; Thoma, R.; Schmidt, S.; Hennig, M.; Kirschner, K.; Sterner, R. *Biochemistry* **2002**, *41*, 12032.

(50) Claren, J.; Malisi, C.; Höcker, B.; Sterner, R. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 3704.

(51) Sterner, R.; Dahm, A.; Darimont, B.; Ivens, A.; Liebl, W.; Kirschner, K. *EMBO J.* **1995**, *14*, 4395.

(52) Lang, D.; Thoma, R.; Henn-Sax, M.; Sterner, R.; Wilmanns, M. *Science* **2000**, *289*, 1546.

(53) Russell, R. B.; Barton, G. J. *Proteins* **1992**, *14*, 309.

(54) Höcker, B.; Beismann-Driemeyer, S.; Hettwer, S.; Lustig, A.; Sterner, R. *Nat. Struct. Biol.* **2001**, *8*, 32.

(55) Sterner, R.; Höcker, B. *Chem. Rev.* **2005**, *105*, 4038.

(56) Boussau, B.; Blanquart, S.; Necsulea, A.; Lartillot, N.; Gouy, M. *Nature* **2008**, *456*, 942.

(57) Jensen, R. A. *Annu. Rev. Microbiol.* **1976**, *30*, 409.

(58) List, F.; Sterner, R.; Wilmanns, M. *ChemBioChem* **2011**, *12*, 1487.

(59) Chaudhuri, B. N.; Lange, S. C.; Myers, R. S.; Chittur, S. V.; Davisson, V. J.; Smith, J. L. *Structure* **2001**, *9*, 987.

(60) Glansdorff, N.; Xu, Y.; Labedan, B. *Biol. Direct* **2008**, *3*, 29.