Supporting information

Hyperstability and substrate promiscuity in laboratory resurrections of Precambrian β-lactamases

Valeria A. Risso,*,† Jose A. Gavira, ‡ Diego F. Mejia-Carmona, † Eric A. Gaucher, § and Jose M. Sanchez-Ruiz*,†

†Facultad de Ciencias, Departamento de Química Física, Universidad de Granada, 18071-Granada, Spain

‡Laboratorio de Estudios Cristalográficos, Instituto Andaluz de Ciencias de la Tierra (Consejo Superior de Investigaciones Científicas—Universidad de Granada), Avenida de las Palmeras 4, 18100 Armilla, Granada, Spain.

§Georgia Institute of Technology. School of Biology, School of Chemistry, and Parker H. Petit Institute for Bioengineering and Biosciences, 30332 Atlanta, Georgia, USA.

MATERIALS AND METHODS

Phylogenetic analysis and ancestral sequence reconstruction

Chromosomal sequences of extant β-lactamases (Tables S1 and S2) were retrieved from the **October**-2011 release of the complete genomes database available at NCBI (www.ncbi.nlm.nih.gov/genome) and, therefore, plasmidic sequences were not included. To further avoid potential complications during the reconstruction process that may arise from recent evolution during the antibiotic era, we determined that none of the sequences came from clinical isolates. Sequences were aligned using MUSCLE (Figure S1). The phylogeny of the Class A β-lactamases was constructed by using the Bayesian method[1] as implemented in the program MrBayes 3.1.2. This analysis used two independent Markov-chain Monte Carlo runs, with four chains each, and was performed for 4,000,000 generations to ensure convergence. Trees were sampled every 1000 generations for each run resulting in a total of 8000 trees, from which a consensus tree was obtained with high probabilities at the nodes (Figure S2 and S3).

Horizontal gene transfer (HGT) is known to be responsible for the spread of antiobiotic resistance determinants and the rate at which resistance genes are mobilized from chromosomes to plasmids has been reported to have increased during the last 60 years[2]. Remarkably, the consensus tree derived from the set of chromosomal sequences (Figures S2 and S3) does not reveal large HGT distortions. First of all, the Bayesian analysis places most sequences of extant lactamases in their "correct" phylogenetic groups (the groups to which the corresponding species belong). Secondly, the overall topology of the consensus tree is reasonably close, not only to the lactamase tree published by Hall and Barlow (ref. 3a in the main text), but also with the Eubacteria tree reported in the Time Tree of Life (ref. 12 in the main text). To make this latter point clear, we include in figure S4 a schematic representation of the Eubacteria tree which can be compared with our consensus lactamase tree of Figures S2 and S3. The only clear evidence of HGT appears to be the presence of a few sequences from Gram negative bacteria that cluster with Terrabacteria group, but this result is in agreement with a previous report (ref. 3a in the main text). There is also some difference in the attachment of the CFB group, although this is mainly related to the fact that we have followed Hall and Barlow (ref. 3a in the main text) and used the CFB group as an outgroup.

The above differences notwithstanding, the general agreement between the topology of the inferred phylogenetic with an accepted phylogeny of the organisms allows us to target well-defined Precambrian nodes: specifically those corresponding to the last common ancestor of enterobacteria (ENCA), the last common ancestor of gamma-proteobacteria (GPBCA), the last common ancestor of various Gram negative bacteria (GNCA) and the last common ancestor of various Gram positive and Gram negative bacteria (PNCA). These nodes are defined in terms of the Eubacteria tree taken from the Time Tree of Life (see Figure S4) and, therefore, can be dated using the divergence times provided in the Time Tree of Life (ref. 6 in the main text), but they can also be easily identified in our inferred consensus tree for lactamases (Figures S2 and S3), opening up the possibility ot targeting the sequence reconstruction exercise to these Precambrian nodes. It is also important that the Bayesian,

Markov-chain Monte Carlo analysis assigns high posterior probability values to the targeted nodes (PP≥0.9, see Figure S2). This result is relevant because it has been shown that sequence reconstructions at nodes with high posterior probability are robust to uncertainties about the underlying tree[3]. Finally, it must be noted that the reported experimental stability determinations with the resurrected proteins (Figure 2) are fully consistent with the claim that our sequence reconstruction process targets Precambrian nodes. Thus, we find a rate of increase in denaturation temperature of about 10 degrees per billion years, which agrees with the the rates found in laboratory resurrection of Precambrian elongation factors (ref. 5d in the main text) and thiroedoxins (ref. 5e in the main text), and also with the rate of change of the temperature of the oceans in the Precambrian as inferred from the isotopic composition of rocks (ref. 5d in the main text).

Ancestral sequences were reconstructed using PAML[4] version 4.4e with the Wag substitution model and incorporating the gamma distribution for variable replacement rates across sites. β-lactamases are rather homogeneous in length and only a few sequence insertions were detected and, furthermore each one of these insertions occurred in a very small number of sequences (2-3 sequences and 1 insertion of 1 amino acid apperring in 7 sequences). These insertions were deleted because they are highly unlikely to have appeared in the common ancestors and, in any case, they are poorly determined in the reconstruction process. Removal of those insertions in the extant sequences would lead to lenghts spanning the range 262-267 amino acids for the mature proteins, with most of the sequences actually being 262 amino acids in length. Therefore, the reconstructred sequences were reduced to a common length of 262 amino acid residues. This is very unlikely to have any signficant effect because the few external residues deleted occur at the highly flexible C-terminal end, which is removed from the active-site (in any case, those positions are again poorly determined in the recontruction process). Profiles of probability for the most probabilistic residue versus position are given in Figure S2 for the sequences reconstruted for the ENCA, GPBCA, GNCA and PNCA nodes (Figure S5). These most probabilistic sequences at each node were selected for laboratory resurrection. The reconstructed sequences have substantial amino acid differences with extant β-lactamases and also to a consensus sequence computed from extant ß-lactamases (Table S3). Nevertheless, it is plausible that consensus mutations could make a significant contribution to the enhancement found, but this may only mean that consensus is an approximate/heuristic way to do ancestral sequence reconstruction, as some authors have suggested[5]. It must be noted, nevertheless, that consensus sequence changes have been explored with the TEM-1 lactamase (ref. 12 in the main text) and the Tm enhancements reported are much smaller that those found here.

To test the phenotypic robustness of the laboratory resurrection, alternative protein sequences were also obtained (and subsequently resurrected in the laboratory) by sampling from Monte Carlo runs based on the posterior probability distribution at the GNCA node (Figure S6, S7, S8, S9 and S10, and Table S4).

Protein Expression and Purification

Genes encoding the ancestral β-lactamase enzymes were synthesized and codon-optimized for expression in *E. coli* cells. The genes were cloned into a pET24 vector (Novagen) and transformed in *E. coli* BL21 (DE3) cells (Stratagene, CA). The β-lactamases were purified by osmotic shock followed by gel-filtration chromatography on a Superdex-75 column in 25 mM Hepes, pH 7.0, 150 mM NaCl. Protein purity was verified by SDS-PAGE. In all cases, stock protein solutions were prepared by exhaustive dialysis against 25 mM Hepes buffer, pH 7.0.

Enzyme kinetics

K_m, k_{cat} and k_{cat}/K_m values for the lactamase-catalyzed hydrolysis of the antibiotics (Figure S10) benzylpenicillin (BZ), cefotaxime (CTX) and ceftazidime (CAZ) were determined[6] at 25°C in 50 mM sodium phosphate buffer pH 7.0 (Table S4 an S5). Initial rates were calculated from the change in UV absorbance that ensued from the consumption of no more than 10% of the substrate. k_{cat} and K_m values were calculated by fitting the Michaelis–Menten equation to the profiles of velocity versus substrate concentration (Figures S12 and S13). In a few instances in which K_m was very large, only k_{cat}/K_m values were determined from the linear profiles of velocity versus substrate concentration. In all cases, experiments were done in triplicate.

Antibiotic susceptibility

E. coli DH5-α cells transformed with a pBR322 plasmid encoding TEM-1 or a ancestral β-lactamases were used for antibiotic resistance tests. Minimum inhibitory concentrations (MICs) were determined by broth micro-dilutions[7]. The antibiotics tested (Figure S11) were: cefotaxime (CTX), ceftazidime (CAZ), ceftriaxone (CRO) and ampicillin (AMP) (Sigma-Aldrich Inc., St. Louis, MO, EUA). Plates were incubated at 37°C for 18-20 hours and examined for cell growth. The MICs of each transformant, as well as that of the control strain *E. coli* ATCC 25922, were assayed in triplicate with serially diluted (2-fold) antibiotic stocks. In all cases, identical values were obtained and, therefore, the variability between different experiments is less than twofold (Table S6).

DSC experiments

Thermal stabilities of ancestral and modern β-lactamase enzymes (Figures 2, S10, S14 and S15) were measured with a VP-Capillary DSC [Microcal (GE Healthcare) Northhampon, USA] as previously described[8-10]. Protein solutions for calorimetric experiments were exhaustively dialyzed against 25 mM Hepes buffer, pH 7.0 and the buffer from the last dialysis step was used as reference in the calorimetric experiments. Prior to scans with protein solutions, several buffer-buffer baselines were obtained to ensure proper equilibration of the calorimeter. For each protein, experiments at different scan rates (60, 90, 120, 150 and 200 K/h) and at different protein concentrations (within the 0.1-1.5 mg/mL range) were typically performed. Concentrations were determined spectrophotometrically at 280 nm using extinction coefficients and molecular weights calculated from the sequences. No significant concentration effects on thermal denaturation were detected. Small effects of scanning rate were found (Figures 2 and S15) and described in terms of the change in denaturation temperature associated to a ten-fold change in scanning rate (kinetic distortion ΔT_m in inset Figure 2A). These scanning rate effects were

certainly much smaller that the differences in T_m between the extant and resurrected Precambrian β-lactamases.

DSC transitions were calorimetrically irreversible in all cases (i.e., no transitions were detected in re-heating runs performed after stopping the first run at temperature immediately above the calorimetric transtion). This indicattes that irreversible alterations of protein states (likely unfolded or partially unfolded states) take place upon thermal denaturation (ref. 8 in the main text) and may potentially give rise to kinetic distortions in the DSC thermograms, as we have extensively discussed in the literature. However, such distortions are unlikely to be substantial in the case of the GNCA and PNCA lactamases because the scanning rate effect on denaturation temperature (a metric of kinetic distortions) is rather small in these cases (see Inset in Figure 2A and S15). Therefore, analysis of DSC transitions for the GNCA and PNCA lactamases in terms of equilibrium thermodynamics models may be permissible and Figure S15 shows the fit of an equilibrium pseudo-two-state model to these transitions. The model has the general mathematical form of a two-state equilibrium model, but allows for a van`t Hoff enthalpy value (the enthalpy change that determines the temperature dependence of the unfolding equilibrium constant) that may differ from the true unfolding enthalpy (the so-called calorimetric enthalpy). As has been extensively discussed in the field[11], the comparison between the two enthalpy values provides a stringent test of the two-state character of the denaturation process (provided that the low-temperature and high-temperature baselines are well defined in the experimental thermograms, as is the case with the calorimetric profiles reported here). Fits of the pseudo-two-state model to the transitions for the GNCA and PNCA lactamases were excellent and yielded van't Hoff enhtalpy values close to the corresponding calorimetric enthalpy values (Figure S15), strongly supporting the two-state character of the thermal unfolding in these cases. Actually, excellent fits can be obtained using a true two-state model in which only one unfolding enthalpy value (determining both the are under the transition and the temperature-dependence of the equilibrium constant) is used. This latter type of fit is that given in Figure 2B of the main text.

It is to be noted that the analyses described above (using equilibrium thermodynamics models) cannnot be applied to the DSC profiles for GPCA, ENCA and TEM-1 lactamases because, in these cases, the scanning-rate effects are larger (see inset in Figure 2A of the main text and Figure S15) indicating substantial kinetic distortions. In these cases, therefore, the potential deviations from two-state behavior cannot be ascertained from the DSC transitions in a straight forward manner, although it is possible that a more elaborate analysis (along the general lines provided in Sanchez-Ruiz[12]) could lead to a reliable description of the thermal denaturation process (work in progress).

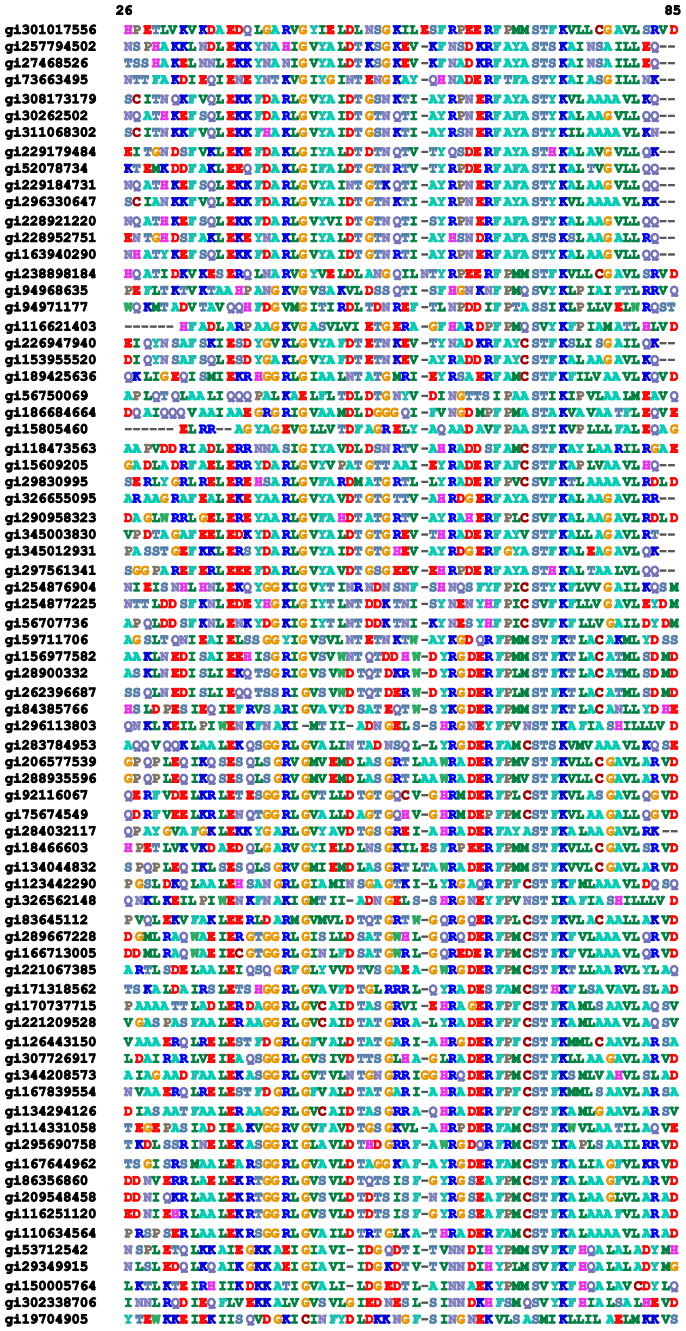
Protein crystallization and structure determination

Crystals of the ENCA and GNCA β-lactamase where grown using the counter-diffusion technique[13]. Good-quality diffracting crystals were obtained at protein concentrations of 30 mg/mL in 25 mM Hepes pH 7.0. Resurrected β-lactamases were crystallized under the following conditions: ENCA: 2.0 M litium sulphate, 0.1M sodium Hepes (pH 7.50) and GNCA: 5.0 M sodium formate, 0.1 M sodium acetate (pH 4.0). The crystal structures

(Figures 1 and S16) were determined by the molecular-replacement method, using as search model the GES-1 β-lactamase structure from *K. Pneumoniae* (PDB.ID 2QPN) solved at 1.1 Å. X-ray crystal structure coordinates have been deposited in the PDB with the following accession codes: ENCA (PDB.ID 3zdj) and GNCA (PDB.ID 4b88). Data collection and refinement statistics are detailed in Table S7.

Although it is often assumed that high stability implies high conformational rigidity, we have not found any clear correlation between B factors and stability. However, although the variation of B-factors along the polypeptide chain can be interpreted in dynamical terms[14], proteins exhibit individual atomic anisotropic motion and collective, large-scale motion over a range of time scales. Dynamics and heterogeneity remains in the crystalline form[15]. Crystal contacts[16] data resolution limit[15], radiation damage, etc., can definitively affect the values determined for the B-factors making difficult to correlate B-factor and stability. We also note that the relation between conformational rigidity and stability, although plausible, does not necessarily hold in all cases, as Rainer Jaenicke discussed several years ago[17] .

Figure S1 Multiple alignment of the amino acid sequence from the 75 class A β-lactamases used in our ancestral sequence reconstruction. Residue numbering for enzymes follows Ambler et al[18]. Only the mature protein sequences are shown here. Analyses were actually carried out with the preprotein sequences which included the ~20 amino acids long signal peptide. However, signal peptides have been omitted here because they show little sequence similarity and their alignment includes large numbers of gaps. Insertions that occur in very small numbers of sequences have also been removed (see "Materials and Methods" for details).



```
              86                                                            145
gi301017556   AGQ EQL GRR IHY SQN DLV EYS PVT EKH LTD GMT VRE LCS AAI TMS DNT AAN LLL TTI GGP
gi257794502   IPY NKL NKK IHI NKD DIV AYS PIL EKY VGK DIT LKE LIE ASM TYS DNT ANN KII KEI GGI
gi27468526    APY NKL NKK VHI NKD DIV AYS PIL EKY VGK DIT LKE LIE ASM KYS DNT ANN KII NEI GGI
gi73663495    VAP SEL NKK VEI NES EIV ANS PVT EQY IGK TMS LKA LIK ASM LQS DNT ANN KIM QEL GGV
gi308173179   KPI EKL NDV IRY TKE DLV TYS PIT EKH LDT GMS LKE ISE AAI RYS DNT AGN ILL QQL GGP
gi30262502    NST KKL DEV ITY TKE DLV DYS PVT EKH VDT GMT LGE IAE AAV RYS DNT AGN ILF HKI GGP
gi311068302   NSI EKL NEV IHY SKD DLV TYS PIT EKH LDT GMS LKE ISE AAI RYS DNT AGN ILL QQL GGP
gi229179484   KSI EDL NQR VLY TRE DLV NYN PIT EKY VDS GMT LKE LAD ASL RYS DNT AQN LIL KQL GGP
gi52078734    KSI EDL NQR ITY TRD DLV NYN PIT EKH VDT GMT LKE LAD ASL RYS DNT AQN LIL KQI GGP
gi229184731   NST MKL DEV ITY TKE DLV DYS PVT EKH VDT GMT LGE IAE AAV RYS DNT AGN ILF NKI GGP
gi296330647   NPI EKL NEV MHY SKD DLV TYS PIT EKH VDT GMS LKE ISE AAI RYS DNT AGN ILL QQL GGP
gi228921220   NSI DSL NEV ITY TKE DLV DYS PVT EKH VDT GMK LGE IAE AAV RSS DNT AGN ILF NKI GGP
gi228952751   NSI EAL DEG ITY TRE DLS NYN PIT KKH VYT GMT LKE LAD ASV RYS DST AHN LIL KKL GGP
gi163940290   NSI DKL NEV ITY TKD DLV EYS PIT EKH VDT GMK LGE IAE AAV RSS DNT AGN ILF NKI GGP
gi238898184   AGL EQL DRR IKY RQS DLV KYS PMT KQH LAD GMT VAE LCS AAI TVS DNT AAN LLL STI GGP
gi94968635    EGK IKL DQV VKV DED DLA ANS PLR DKH PRG EFT IED LLR RAI VDS DGS ASD VLL KQI GGT
gi94971177    GTG QKL TDL YTF RKE DVV PDS AIM ENL TPS KVT NRD LAG MVV AVS DNS ATN VLI DRV -GF
gi116621403   TGK LKL DQP VPI AKA ELP GHS PIR DQH PEA TLS LRE VIR FAV AES DGT ASD VLL RLS GGP
gi226947940   YSS DQL KQV IKY SPK DVL SYA PVT KNH VDK GMT IEE LCD AAV RFS DNT AAN LLI NLI GGP
gi153955520   DSL EQL KQL VKY KKE DVL SYA PIA KDN VDK GMT IEE ICS AAI RFS DNT AAN LLL NHI GGP
gi189425636   EKK ISL DRG IPY TTA DLL DWA PIT KKH VSG AMT VER LCA AAI EYS DNT AAN LLL DLL GGP
gi56750069    QGK VRL DEQ LTM TQD LVA SAG AMQ YQP VGS KFT VLE TAI EMI RIS DNT ATN LLI ARL GGI
gi186684664   KGN FRL DQQ FPM MVR -VA ETG RVA EAP LRT VMT AQS LME LMI TRS HNE ATD GLI HAV GGY
gi15805460    RGD LDL IER VTL RAE DRV PGA GVH ELG LGL ALT WQD VLT LMI VVS DNT ATN LLI ERL -GQ
gi118473563   RGE LSL DDR VFV DPA ALL SNS PIT ETH AGG EMT LAE LCQ AAL QRS DNA AAN LLL KQI GGP
gi15609205    NPL THL DKL ITY TSD DIR SIS PVA QQH VQT GMT IGQ LCD AAI RYS DGT AAN LLL ADL GGT
gi29830995    HDG EFL ARR IRY TEQ YVS GYG PVT GRA ENA GMT VAE LCA AAV SDS DNA AGN LLL REL GGP
gi326655095   YGL GGL ERV VTY RRE DLV DHS PVT EKH VAT GMS LGA LCD AAV RFS DNT AGN LLF DAV GGP
gi290958323   RDG AYL GRR IFY TKE YAA GYS PST DEN VAA GMT VGE LCA ATV SHS DNG AGN LLL REL GGP
gi345003830   YTV GGT DRL ITY SGD DVV PDS PVA EKH VDT GMT LDA LCD AAV RYS DNT AAN LLL GQL GGP
gi345012931   VKL GGL DRV IKY SED DLI DNS PVT EKH VDT GMS LGA LCD AAV RYS DNT AAN LLF DEL GGP
gi297561341   NTP EEM EEV VTF TEE DLV DYS PVT EQH VDT GMT LME VAD AAV RYS DNT AAN LLF EEL GGP
gi254876904   TDN NLL NKE VKI SAN QII GYS PVT KKH INQ TMT VSE LSK AAI Q-S DNT ATN LLI EKL GGL
gi254877225   QHK GFL DKK ILI TQD DIL GYA PVT GKN IGK TLT ISQ LNY AAI -LS DNP AAN ILV REI GGL
gi56707736    HNQ GFL DKK IPI NQD DIL GYA PIT AKN VGK TLT ISQ LNY AAI -LS DSP ASN ILV REL GGL
gi59711706    VGG IDK YKT TSI TKE QIM PWS PVT EPL IGN MIT TQK ACE ATM LMS DNT AAN IVL NEI GGP
gi156977582   NEK LDK NAT AKV EER NMV VWS PVM DRM AGQ TTR IEH ACE ATM LMS DNT AAN IVL RSI GGP
gi28900332    SGK LNK NAT AKI DER NIV VWS PVM DKL AGQ STR IEH ACE AAM LMS DNT AAN LVL NEI GGP
gi262396687   SGK LSK NAT AKI DER SIV VWS PVM DKL AGQ STR VEH ACE AAM LMS DNT AAN LVL NEI GGP
gi84385766    NEG LYL ESK VGV KSD ELI AWS PMT KAF VGK EIS LTG ACA ATM EMS DNT AAN IVL AGI DGP
gi296113803   KEK LDL NEK III KES DLI EYS PVC KKY FDK PIS ISE LCE ATI TLS DNG SAN ILL DKI GGL
gi283784953   TRP AIL QQK IPI KKA DLT NWN PVT EKY VGK EMT LAE LSA ATL QYS DNT AMN KLI AHL GGP
gi206577539   AGL EQL DRR IHY RQQ DLV DYS PVS EKH LVD GMT IGE LCA AAI TLS DNS AGN LLL ATV GGP
gi288935596   AGL EQL DRR IHY RQQ DLV DYS PVS EKH LVD GMT IGE LCA AAI TLS DNS AGN LLL ATV GGP
gi92116067    AGK ESL ARR IYF NEA DLV THS PET HKR VGI GIT VAE LCK AAI TLS DNT AGN LLL ASI GGP
gi75674549    AGK EDL ARR IYF NEA DLV AHS PET RKH VGT GMT VAE LCK AAI TLS DNT AGN LLL ASI GGP
gi284032117   YTL QGM GKV VKY SES DLL ANS PIT EKH VAT GMT LRE LCD AAV RYS DNT ASN LLF DAL GGP
gi18466603    AGQ EQL GRR IHY SQN DLV EYS PVT EKH LTD GMT VRE LCS AAI TMS DNT AAN LLL TTI GGP
gi134044832   AGD EQL ERK IHY RQQ DLV DYS PVS EKH LAD GMT VGE LCA AAI TMS DNS AAN LLL ATV GGP
gi123442290   SQP NLL NKH INY HES DLL SYA PIT RKN LAC GMT VSE LCA ATI QYS DNT AAN LLI KEL GGL
gi326562148   KEK LDL NEK III KES DLI EYS PVC KKY FDK PIS ISE LCE ATI TLS DNG SAN ILL DKI GGL
gi83645112    AGE ENL ERR VPI HAR DIV TYS PVT KDR IGE GMT LRE LCQ AAM TKS DNT AAN LVL KAV GGP
gi289667228   QGK LML AQR VKI RAS DML EYA PVT ERH IGG SLS VGE LCR ATM IYS DNP AAN LLF PLV GDP
gi166713005   QGK LTL TQR VKI RAS DML EHA PVT ERH VGG SLS VGE LCR ATM IYS DNP AAN VLF PLV GDP
gi221067385   RGE IRL WRK LYY SPA EVV AWS PIS EKR AGG GMT VQE LCE AMV VVS DNT AAN VML EAS GGP
gi171318562   QGK LDL DKP VKY GKT DIL DYA PVA KKH VDG FMT VDE LCA AAT VWS DNT AAN LLI ELL GGP
gi170737715   ERP GLL QQR VTY TKA DLV NYS PVS EKH VGS GMT VAA LCE AAI QYS DNS AAN LLM KLI GGP
gi221209528   AHP GLL QQR VTY GRS DLV NYS PVT ERH VDT GMT VAE LCA ATI QYS DNT AAN ELM KRI GGP
gi126443150   GEP ALL QRR IAY AKG DLI RYS PIT EQH VGA GMS VAE LCA ATL QYS DNT AAN LLI ALL GGP
gi307726917   RGE ENL QRR IVY SQS ELV PNS PHT RER TGA GMS IAE LCK AAI TLS DNT AAN LLL QSF GGP
gi344208573   AGR VSL DTR VPI AEK DLL SYA PVA RRH VGK DLT VRD LCR GTL TTS DNT AAN LLL DVV GGP
gi167839554   DDA ALL QRR IPY AKR DLV RYS PIT ERH VGA GMT VAE LCA ATL QYS DNT AAN LLI DLL GGP
gi134294126   EHP GLL QQR VTY AQA DLV NSS PVS GRH VDT GMT VAQ LCE AAI QYS DNT AAN LAM KLI GGP
gi114331058   QNL LSL SSH VLY GQA DLL AYA PVT SMH AKG FMT VSE LTR AAI VVS DNT AAN LLL SKL GGP
gi295690758   QGR ERL DRR VTF GPE VLM DNS PIV EKH VQD GMT IGQ LCE ATI TLS DNA AAN LLF EAL GGP
gi167644962   QGK ERL DRR ITY GKG VLL AHS PAT EKH VGD GMT VGE LCE ATI TLS DNA AAN LLL RTF GGP
gi86356860    KGE ENL ERR VTY GRD KLV DHS PLS EKH AGD GMT IAE LCE AAV TVS DNT AGN LLL ESF GGP
gi209548458   KGE ESL DRR VTY GKD KLV DYS PLS EKH ADD GMT ITE LCE AAV TVS DNT AGN LLL ESF GGP
gi116251120   KGE ENL DRR VSY GKD KLV DYS PIS EKH AGD GMT LAE LCE AAV TVS DNT AGN LLL ESF GGP
gi110634564   AGK EDL NRR ISY GRE KLI FWS PVT KNH ADP GMT LAE LCE AAV TMS DNT AGN LLL ESL GGP
gi53712542    HQK QPL ETR LLI KKS DLD TYS PLR ETY PQI EMS IAD LLK YTL QQS DNN ACD ILF NYQ GGP
gi29349915    KQQ QSL NFE LTI KKE DLD TYS PLR DSF PQF NID IAD LLN YTL QQS DNN ACD ILF QYQ GGV
gi150005764   KRH IPL STS LYL DKK YFD TYS PLR DKY PQL ELP ISE LLV YTM QSS DNV ACD ILF DYI GGV
gi302338706   TGN LSL EDS IML KES DLN TWS PLR ETY LAK EIT LAE LIN YTV ALS DNN GSD ILL RLI GGT
gi19704905    ENK FSL SDT ITI TNF MKT GGD GVL KEL NAH HFT LKE LAT LMI IVS DNQ ATN ILI DFL -GM
```

```
             146                                                              205
gi301017556  KE LTA FLH NMG DHV TRL DSW EPE LNE AIP NDE RDTT MP AAM ATT LRK LLT GELL TL ASR Q
gi257794502  KK VKQ RLK ELG DKV TNP VRY EIE LNYY SP KSKK DTST P AAF GKT LNK LIA NGK LSK ENKK
gi27468526   KK IKK RLKK LG DKV TNP VRY EIE LNYY SP KSKK DTST P AAF GKT LNK LIA NGK LSKK NK N
gi73663495   NG LKH ELV QLG DDV SEP QRL EPE LNY FDP NSK ADTTT P RAAA QT LNS ILTS NE MNE SNL S
gi308173179  KG FEK SLK QIG DHV TKA DRF ETD LNS AIP GDI RDTST A KAL ATD LKA FTL GNT LTT DKR T
gi30262502   KG YEK ALR QMG DRV TMS DRF ETE LNE AIP GDI RDTST A KAI ATN LKA FTA GNA LPN HKR N
gi311068302  NG FEK SLK QIG DHV TKS NRF ETD LNS AIP GGI RDTST A KAL ATD LKA FTL DNT LTT DKR T
gi229179484  SE FKK SLR EVG DTV TNP ERF EPE LNE VPP GDV RDTST P KAL ATS LQA FTL GDVL ST EKR D
gi52078734   ES LKK ELR KIG DEV TNP ERF EPE LNE VNP GET QDTST A RAL ATS LQA FAL EDK LPS EKR E
gi229184731  KG YEK ALR QMG DRV TMS DRF ETE LNE AIP GDI RDTST A KAI ATN LKA FTA GNA LPN HKR N
gi296330647  KG FEK SLK QIG DHV TKA NRF ETD LNS AIP GDI RDTST A KAL ASD LKA FTV DNT LTT DQR T
gi228921220  KG YEK ALR HMG DRI TMS DRF ETE LNE AIP GDI RDTST A KAI ATN LKA FTV GNA LPA EKR K
gi228952751  SA FEK ILR ETG DTV TKS ERF EPE LNE VHP GET RDTST P EAI AKT LQS FTL GSA LPT EKR E
gi163940290  KG YEK ALR QMG DRV TIA DRF EPE LNE ATP GDI RDTST A KAI ATN LKA FTV GNA LPA DKR K
gi238898184  KG LTS FLQ KTG DQV TRL DRW ETE LNQ ALP GDE RDTTT P KIM AQT LHN LLT GKIL SV KSR Q
gi94968635   RE VRR YLK SQK IKG IRV QHT EEQ LIE NTH LQY SDS AKP DAMV DLL ER LQQ GKLL DT ENT A
gi94971177   DN VNK MLD AQG LHK TRL RRH MMD LAAA KA GNE -NV STP HEM TQLL EA VYR GKIL NP QMS E
gi116621403  SA VTA YLR GLG ING VIV ATS ELE MSR GPMV QY RNW ATP DSMV DLL RA FHA GRG LSP AAR T
gi226947940  NG FKS ALN QLG DTV TEP ARI EPE LNV ATP GDN RDTST P RQL SID LKE YTT GNIL SDD KKK
gi153955520  KG FKS ALN QLG DSV TQP VHI EPE LNE GIP GDI GDTST P RQL ATD LQA YTT GNIL TED KKK
gi189425636  KA VTA YAR SLG DSV TRL DRN EPS LND NLP NDD RDTTS P SSMV ET MNK VLS GHA LSP SSR K
gi56750069   EA VNQ RFAA WG LQTT RI RNLL PD L------ EG TNSTT P KEMV ELL AK VEN GSLL DP RSR D
gi186684664  EN VNR WLT RNG ITG QRL DHT MAT LVR DDG RID RTSST P RAMI ALL AA IDR GGVL SP ESR A
gi15805460   DH FNL WLT ARG LNST RLI GP LER QNE AQRR GE RNR TTA RDQ VALLL S LLR GDG LTP QMQ H
gi118473563  AE ITA FAR SIG DQR TRL DRW ETE LNS AVP GDP RDTST P AAL AGG FRA VLT GDVL APP QR Q
gi15609205   AA FTG YLR SLG DTV SRL DAEE PE LNR DPP GDE RDTTT P HAI ALVL QQ LVL GNA LPP DKR A
gi29830995   TA ITG FCR SIG DGR TRL DRW EPD LNS AEP WRD TDTTS P RAI GKT YAR LIL GDVL PP QDR T
gi326655095  RK LQA VLA GLG DEV TRMV RR ETE LNE WTP GAT RDTST P RAL AED LRA FVL GDA LRG PER A
gi290958323  TA ITR FCR SLG DRA TRL DRW EPE LNT AEP WRI TDTTT P RAI GGT YAR LLL GRA LPD ADR E
gi345003830  EA LNA LLK RLG DDV TRM ERL EPD LSR WDP DST RDTST P RAF AEN LRA FVL GDVL GD AER A
gi345012931  KG LDA TLE KMG DDVI QM ERR EPE LSR WVP GST RDTST P RAM AKD LRA FVL GDVL GD GER A
gi297561341  DA FEDE LE AIG DDV TDP ERI ETE LND WTP GDTS DTST P RAM AGN LRE FTL GDA LPED RR E
gi254876904  NN LNN FIL SLH DHA TKV AGL EPV ANQ VSL TTN QNK TTP KIM AKD INK LAF SNIL DKK HR L
gi254877225  EK LNK LVA KLG DKD TII KNDE PK INH TKP DSN INK TTP KAI TQD IYN LAF GNIL DKK HR E
gi56707736   QN LNK FIKK LG DND TII TADE PE INY TQP HSN INK TTP KAI TKD IYK LAF GNIL DKK HK D
gi59711706   SS LTQ FLR AMG DTK TRL DRI EPE LNE SKN GDK RDTTT P VAM SQT LNA LLF GNT LNS QDE Q
gi156977582  RG VTT FLR SVG DKA TRL DRF EPR LNE ANP GDK RDTTT P NAMV NT LH TLL EGD ALS YES RI
gi28900332   KA VTQ FLR SIG DKA TRL DRL EPR LNE AKP GDK RD TTT PNA MVN TLH TLM EDN ALS YES RT
gi262396687  KA VTM FLR SIG DKA TRL DRL EP RLN EAT PGD NRD TTT PNA MVN TLR TLI EGE TLS YES RV
gi84385766   NG LTE FLR SI GDR ETR LDH IEP DLN HAR PGDE RD TTT PNT MVK TLN ELV YGN VLS EES KA
gi296113803  TA FN QFL KEI GAD MV- LANN EP LLN RSH YGE TSD TAK PIP YTK SLK ALIV GN ILS NQS KE
gi283784953  GK VT AFA RST GDTT FR LDR QEP ELN TAI PGDE RD TSS PRA MAD SLR NLT LGD ALA DTQ RA
gi206577539  AG LT AFL RQI GDN VTR LDR WET ALN EAL PGD ARD TTT PAS MAA TLR KLL TAQ HLS PRS QQ
gi288935596  AG LT AFL RQI GDN VTR LDR WET ALN EAL PGD ARD TTT PAS MAA TLR KLL TAQ HLS ARS QQ
gi92116067   QG LT AFV RKL GDD VTR LDR IET ELN EAA PGD PRD TTT PNA MAS DLR ALA VGD VLS AKS RA
gi75674549   QG LT AFA RRL GDE VTR LDR IET ELN ESV PGD PRD TTT PSA MAA NLR ALA LGD VLS VKS RT
gi284032117  KG LD AVL AEL GDN VTQ MEQ REL ELS VWD PSN PRD TST PRA FAR NLR TFVL GD ALR APE RA
gi18466603   K ELT AFL HNM GDH VTR LDR WEP ELN EAI PNDE RD TTM PAA MATT LR KLL TGE LLT LAS RQ
gi134044832  A GLT AFL RQI GDN VTR LDR WET ELN EAL PGD ARD TTT PAS MAA TLR KLL TSQ RLS ARS QR
gi123442290  AA VN QFA RSI GDQ MFR LDR WEP DLN TAL PND PRD TTT PAA MAA SMN KLVL GD ALR PAQ RS
gi326562148  T AFN QFL KEI GAD MV- LANN EP LLN RSH YGE TSD TAK PIP YTK SLK ALIV GN ILS NQS KE
gi83645112   P RLT EYL RGL GDE TTR LDR WET ELN EAT PGD PRD TTS PAA MAST LQ TLF LNS PLS RAS RQ
gi289667228  A GVT AFL RSI GDA QTC SDR HEP DMN RFAA GD PRD TTT PAA MAN TLR TLLL GD ALR PAS RK
gi166713005  A GLT GFL RGI GDA QTR SDR YEP EMN RFAA GD PRD TTT PAA MAA TLR TLLL GD ALQ PSS RT
gi221067385  AA LT QWL REM GDG TTR LDR NEP SLN TAL PGDE RD TTT PRA MVQ SLQ KLLL GD VLE GYA RA
gi171318562  A GWT RFA RSI GDN ASR LDR TEP ALN SAI PGD PRD TTT PDA MLA NLQQ TLL GT VLS DSS RE
gi170737715  S AVT AYA RSI GDD AFR LDR WET ELN TAL PGD PRD TTT PAA MAA SLR VLT LGD ALP AAQ RA
gi221209528  AA VT AYA RSI GDD TFR LDR WET ELN TAL PGD LRD TTT PAA MAA SLR VLVL GD ALPP AQ RA
gi126443150  Q AVT AYA RSI GDA TFR LDRR EP ELN TAL PGDE RD TTT PAA MAA SVH RLLV GD ALG AAQ RA
gi307726917  AA LT AFA RSL GDG MTR LDR NEP TLN EAA PGD PRD TTT PNA MLG NLR ALVM GE QLSSTS RV
gi344208573  S ALT AFL RGQ GDS VTR NDR NEP DVN LFA KGD PRD TTS PAA MATS MA RFA VGN GLQ PAS RQ
gi167839554  Q AVT AYA RSI GDA TFR LDRR ET ELN TAI PGDE RD TTT PAA MAA SVRR LL AGD ALG AAQ RA
gi134294126  AA VT AYA RSI GDD TFR LDR WET ALN TAL PGD LRD TTT PAA MAA SIR VLT LGD ALP AAQ RT
gi114331058  S VVT EFV RSC GDA VTR LDR NEP TLN DND AGD LRD TTS PRA MAT LMG NVL CGN RLSSTS RD
gi295690758  AA LT RFL RAI GDQ TTR SDR LEP ELN RGPP DD PRD TTT PAA MLA TWK TLLL DD ALS PAS RQ
gi167644962  P ALT AFL RGL GDK VTR TDG YEP GIN VVE PGE VHD TTS PNA MLA TLR TLT LGS ALA PAS RQ
gi86356860   A GLT DWL RSI GDG TTR LDR TEP TLN EGR KDD PRD TTS PDA MLD TLG NLT LGS VLT EASS D
gi209548458  A GLT DWL RSI GDG TTR LDR TEP KLN EGR KDD PRD TTT PDA MLD TLG NLT LGS VLS EASS D
gi116251120  A GLT DWL RSI GDG TTR LDR TEP TLN EGR KGD PRD TTT PDA MLD TLG NLT LGS VLT EAS CD
gi110634564  K GLT AFF RSL GDE VSR LDR WEP ELN DVPP GD LRD TTT PEA MLG SLR SVVL GD ALS PNS RQ
gi53712542   D AVN KYL HSL GIR ECA VIH TEN DMH KNL EFC YQN WTT PLAAA K LLE IFR NEN LFD KEY KN
gi29349915   D TVN QYI HSL GVT DCA IVC TEN DMH QDE SLC YQN WTT PLAAA R LLE IFR KEA LFP QEY KD
gi150005764  N VVDE YI HSL GIN DVS ITA TEDE MH QGM DDC YKN WTT PME AAN LLE LFM TQD FMR NEY TD
gi302338706  E KVN EYI HGL GII DVQ ITT DEE NMH KES DLQ FEN WTT PEAA TE LLT LFY DGK VLS KQST D
gi19704905   E NIN FLG KEL GLR ETF LERR MM D-A EAS KNG YDN YTS ADD ISLLL K LIY QEK LIN KETS Q
```

```
                206                                                          267
gi301017556     QLIDWMEADKVAGPLLRSALPAGWFIADKSGAGKRGSRGIIAALGPDGKPSRIVVIYTTG
gi257794502     FLLDLMLNNKSGDTLIKDGVSKDCKVADKSGQAIYASRNDVAFVYPGQSEPIVLVIFTNK
gi27468526      FLLDLMLNNKNGDTLIKDGVPKDYKVADKSGQAIYASRNDVAFIYPNQSEPIILVIFTNK
gi73663495      LLKQTMIENKTGDTLIKAGMPNSYTVGDKSGQALYATRNDLAFIYPGQDKPIILAIYSKQ
gi308173179     ILTDWMRGNATGDELIRAGAPAGWEVGDKSGAGSYGTRNDIAIVWPPDRAPIVLAILTKR
gi30262502      ILTKWMKGNATGDKLIRAGVPTNWVVADKSGAGSYGTRNDIAIVWPPNRAPIIIAILSSK
gi311068302     ILTDWMRGNATGDELIRAGAPIGWEVGDKSGAGSYGTRNDIAIVWPPNRAPIVLAILSNR
gi229179484     LLIDWMKRNTTGDNLIRAGVPGGWEVADKTGSGSYGTRNDIAIIWPPNKKPIVLAILSNH
gi52078734      LLIDWMKRNTTGDALIRAGVPEGWEVADKTGAGSYGTRNDIAIIWPPKGDPVVLAVLSSR
gi229184731     ILTEWMKGNATGDKLIRAGVPTDWIVADKSGAGSYGTRNDIAIVWPPNRAPIIIAILSSK
gi296330647     ILTDWMRGNATGDELIRAGAPTGWEVGDKSGAGSYGTRNDIAIVWPPNRAPIVLAILSKR
gi228921220     ILTEWMKGNATGDKLIRAGVPTDWVVGDKSGAGSYGTRNDIAIVWPPNRAPIIIAILSSK
gi228952751     LLVDWMKRNTTGDKLIRAGVPKGWEVADKTGAGSYGTRNDIAIIWPPNKKPIVLAILSNH
gi163940290     VLTEWMKGNATGDKLIRAGVPTDWIVGDKSGAGSYGTRNDIAIVWPPNRAPIIIAILSSK
gi238898184     QLQTWMVMNKVGGPLLRSVLPANWFIADKTGAGNHGSRGIIAALGADGKPSRIVVIYLTE
gi94968635      RLVGWMKETDTGKDRIRAKLPSG-TVADKTGSSGKGLTNDVALVTLPNGNHFAIAIFVTD
gi94971177      DF--WNLMCTKKESPLPRLLPEDLRIANKPGELD-GVRTDSGLVFLKNR-PYVISVMTTY
gi116621403     LLEDFMAQSTPGPKRLKGLLPPGTRVAHKTGTSGTLTRNDVGIVTLPNGRHLAIAVFVSD
gi226947940     ILINWMSGNATGDKLIRAGAPKDWMVSDKSGTGSYGTRNDIAIVIPPNKKPIFIAILSSK
gi153955520     ILIDWMAGNTTGNTLIRAGAPKSWIVADKSGTGPYGRRNDIAIVMPPNKKPIIIAILSTH
gi189425636     ILENYLLANTTGANRLRAGIPLDWRVGDKTGTGSNGAVNDIAVMWPPDRSPIFVAVYYSG
gi56750069      RFFAVLRTTET-NTLLSRGLGEEATIAHKTGDIA-MVVGDVGLIDMPSGKRYIAAMLVRR
gi186684664     VLLDTMTRTSTGKTRIRAGLPVGTLVAHKTGTLS-GVTDDVGIIRLPDGRHLAVAFFVTG
gi15805460      LALDILSRQHLRD-LIGRGVPAGYRVASKSGELR-GVHHDVGLLWAP--RPLVVALLSQG
gi118473563     LLDEWMRANETSS--LRAGLPDGWTSADKTGSGDYGSTNDVGIAYGPQGQRILLMVRTRG
gi15609205      LLTDWMARNTTGAKRIRAGFPADWKVIDKTGTGDYGRANDIAVVWSPTGVPYVVAVMSDR
gi29830995      RLTNWLIANTTNTARFRAGLPADWILADKTGGGNYGVANDVGVVWPPNRSPIVMTVLTTK
gi326655095     RLTQWLTTNTTGGELIRAGVPKGWTVGDKTGAGSYGTRNDIAVVWPPDAAPLVLAVLSNR
gi290958323     LLTGWLIANSTNVERFRAGLPAGWTLADKTGGGSYGVANDVGVVWPPGRPPLVLSVLSTK
gi345003830     RLTHWLRTNTTGDEVIRAGVPKGWTVGDKTGTGIYGVRNDIAVLWPPDRAPVVVAVMTSR
gi345012931     QLTKWLRTNTTGDALIRAGVPKDWQVGDKTGSGSYGGRNDIAVVWPPDSAPIVVAIMSNR
gi297561341     VLNDLLVNNTTGDALIRAGVPEDWRVGDKTGAGSHGTRNDIAVIWPPEGDPIVLAVMSSR
gi254876904     LFKKWLLENDSGNNRIAAEIPDEWEIGNKTGTCAYGSTNDVAIIWPDDDKAIIMAIFYTQ
gi254877225     IFIGYLQKNNTGANRIAYSMPKNWTIGDKTGTCGYAATNDVAIIWPQNKPPFALSILYTN
gi56707736      IFIKYLQDNNTGANRIAFSMPKDWIIGDKTGTCGYAATNDVAIIWPKNQQPIALGILYTN
gi59711706      TLKSWMMNNKVSDPLLRSILPNDWSIADRSGAGGFGSRGITAAIWNDKHQPIIISIYLTQ
gi156977582     QLKIWMQDNKVSDSLMRSVLPKGWSIADRSGAGGFGSRGITAMIWKENHKPIYISIYITE
gi28900332      QLKIWMQDNKVSDSLMRSVLPKGWSIADRSGAGNYGSRGISAMIWKDNYKPVYISIYVTD
gi262396687     QLKIWMQDNKVSDSLMRSVLPTGWSIADRSGAGGHGSRGINAIIWKENHRPVYISIYVTE
gi84385766      QLKTWMMDNKVSDGMIRSILPDGWNIADRSGAGAYGSRAITAIVWSETRAPLIISISLTE
gi296113803     QLITWLINDKVADNLLRKYLPKNWRIGDKTGTGS-ESKNIIAVIWNENNKPYFISLFHDG
gi283784953     QLVAWLKGNTTGGQSIRAGVPAHWVVGDKTGAGEYGTTNDIAIIWPENRPPLILVTYFTQ
gi206577539     QLLQWMVDDRVAGPLIRAVLPAGWFIADKTGAGERGARGIVALLGPDGKPERIVVIYLRD
gi288935596     QLLQWMVDDRVAGPLIRAVLPPGWFIADKTGAGERGARGIVALLGPDGKPERIVVLYLRD
gi92116067      QLVGWLAANTTGGKRLRAGLPAGWRVGDKTGTGERGTANDVAVIWPPDRAPFIITAYLTG
gi75674549      QLVGWLVANTTGDKRLCAGLPAGWRAGDKTGTGERGTANDVAVIWPPDRAPFVVTAYLTG
gi284032117     QLTTWLRTNLTGAELIRAGMPKGWVVGDKTGTAAYGGRNDIAVVWPPGRAPIVLAIFSNR
gi18466603      QLIDWMEADKVAGPLLRSALPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTG
gi134044832     QLLQWMVDDRVAGPLIRSVLPAGWFIADKTGAGERGARGIVALLGPNNKAERIVVIYLRD
gi123442290     QLAAWLKGNTTGDATIRAGAPTDWIVGDKTGSGDYGTTNDIAVLWPTKGAPIVLVVYFTQ
gi326562148     QLITWLINDKVADNLLRKYLPKNWRIGDKTGTGS-ESKNIIAVIWNENNKPYFISLFHDG
gi83645112      QLTAWFLANETGDAKLRAGLPQDWRIGDRTGGGAYGSTNDVAVIWPPGREPLIVTVYITE
gi289667228     QLTDWMINNRTGDDCLRAGLTRDWKIGDKTGSNGIDTRNDIAILWPKGRAPLLLTAYLNG
gi166713005     QLTDWMIDNRTGDDCLRAGLTREWKIGDKTGSNGTDTRNDIAILWPKGRAPLLLTTYLNG
gi221067385     LLQQWLADSRTGDKRVRSGMPGDWKVGGKTGSGERGTACDTLIVWPAQSAPLLVTAYLTG
gi171318562     KLQNWMLSGTITDTLLRAGVPLTWRVGDKSGAGRNGTRNDIGIIFPPNASPILAAIYYTE
gi170737715     QLVAWLRGNKVGDKRMRAGVPAGWVVGDKTGTGDYGTTNDAGVIWPTSRAPIVLAVYYTQ
gi221209528     QLIEWLRGNKVGDKRIRAGVPTGWRVGDKTGTGDYGTTNDAGVLWPPSRAPIVLAVYYTQ
gi126443150     QLNAWMLGNKTGDARIRAGVPADWRVADKTGTGDYGTANDIGVAYPPNRAPIVFIIYTTM
gi307726917     QLLAWLAANETGGARIRARLPKDWGVGDKTGTGGHGTANDVAILWPPGRGPILVAVYLTG
gi344208573     QFADWLIDNQTGDACLRAGLGKRWRVGDKTGSNGEDTRNDIAVLWPAGGAAWVVTAYLQG
gi167839554     QLNAWMLGNKTGDARLRAGVPAGWRVADKTGTGDYGTGNDIGVAYPPDRAPIVFVVYTTM
gi134294126     QLVAWMRGNKVGDKRIRAGVPAGSQVADKTGTGDYGTTNDAGVVWLPSRAPIVLTVYYTQ
gi114331058     LLLQWLRDCETGYDRLRAKLPKDWLVGDKTGTGEHGAVNDVAIAIPPGRSPILIAAFLSE
gi295690758     QFRDWLIANKTGDRRLRAGLPPGWRVGDKTGNNGRDITNDVAIAWPPGRGPILIAAFHDR
gi167644962     RLVGWLVANQTGDKRLRAGLLAGWTVGDKTGSWGDGTTNDIAVIWPTGRPPLLITAYLAD
gi86356860      RLIAWLVASTTGKERLRAGLPQDWKVGDKTGTGQNGSLGDIAVIWPPDRGPIVAAVYIAE
gi209548458     RLIAWLVASTTGGERLRAGLPAEWKVGDKTGTGPNGSLGDIAVIWPPNRGPIVAAVYIAE
gi116251120     RLIAWLVASTTGKERLRAGLPADWKVGDKTGTGPNGSLGDIAVIWPPDRGPIVAAVYISE
gi110634564     RLAGWLIGCKTGDKRLRAGFPAGWTVGDKTGTGS-TVASDIAVAWARSGATLVVSAYLEA
gi53712542      FIYQTMVECQTGQGRLIAPLLKKVTMGHKTGTGDRGQQNDIGFILLPDGHVYSIAVFVKD
gi29349915      FIYQTMTECQTGQDRLVAPLLKEVTIGHKTGTGDRGQQNDIGFVLLPDGHAYSIAVFVKD
gi150005764     FLKHIMIECGTGKDRLPASLPSEVKIGHKTGTSDNGEYNDIGFVILPDGSRYVVAVFVKD
gi302338706     FLWKVMAGTETGGGRIKGKLPSGTLVAHKTGTSAPGVTNDIGIVTLPNNNHFVMSVFVAN
gi19704905      LMLDILLRQQQGE-RLQRYLPPDIKIAHKCGDLD-NLENDGGIIWLGNKI-YILVVLTSG
```

```
                  268                 290
gi301017556       SQATMDERNRQIAEIGASLIKHW----
gi257794502       DNKSDKPNDKLISETAKSVMKEF----
gi27468526        DNKSDKPNDKLISETAKNVINKF----
gi73663495        DQKDAKPNDKVISDSAREVIKYLK---
gi308173179       FTKDAEYDNALIAEAAKVALDDLK---
gi30262502        DEKGATYDNQLIAEAAEVIVNAFR---
gi311068302       FTKDANYDNALIAEAAKVALNDLK---
gi229179484       DKEDAKYDDKLIAEATKIVLDALKVTN
gi52078734        DKKDAKYDDKLIAEATKVVVKALNME S
gi229184731       DEKEATYDNQLIAEAAEVIVNAFR---
gi296330647       FTKDANYDNALIAEAAKVVLNDLK---
gi228921220       DEKEATYDNQLIAEATKVIVKALR---
gi228952751       DKKDAEYDDKLIADTTKVVLDTLKVTN
gi163940290       DEKEAAYDNQLIAEATEVIVKALK---
gi238898184       TQATMDERNKQIAEIGASLIKNWYS--
gi94968635        SKADTKTRNLVIANIAHEVWSCWAEE T
gi94971177        AA-NEREASAAIAQISLIAYNYFHAVS
gi116621403       STASLDLRKEAIAKIARAAWDAFQ---
gi226947940       NDKDAKYDDKTISEASKIVFDYFINTR
gi153955520       DTKEAKYDDKLIAKASKIIFDSFTTTE
gi189425636       SPSPNSDREAVLAEVGKLIAIEFNKRS
gi56750069        PDNDLKA-NELIRQVSQRTYQFLNSLP
gi186684664       PE-GHAAHAALIARMTRVIYNGYSTPV
gi15805460        GEDREHPGNRDVTRLSAALWPLLAELG
gi118473563       DDPNADGFRPLIGELTALVLPELGVH-
gi15609205        AGGDAEFREALLAEAATCVAGVLA---
gi29830995        HAAAGPSDSPLVARTAELLAAELT---
gi326655095       TTADADHDNTLIAKAASAAVTALTR--
gi290958323       YDPAGPTDNALVARAAGLVAGALTA--
gi345003830       DREDDTHDNRLLAEAASVVTGALS---
gi345012931       GDKDATSDDKLIADAAHVVADTLS---
gi297561341       EAEDAEYDDALIARAAEIVAKELG---
gi254876904       SQKDAKPNSKIIREVTKILLDKLEQQN
gi254877225       PSDKAPSNEKIIQQASKLVSESIAKKD
gi156707736       PNDNAPSNEKIIQQAAKLIANDLTNTY
gi59711706        TKLDMAERNQVIVEVGNAIFKEFKIN-
gi156977582       TDLSLQARDQIIAQVSQLILDEYNTI-
gi28900332        TDLSLQARDQLIAQISQLILEHYKES-
gi262396687       TELSLQARDQLIAQISQLILQKYKDN-
gi84385766        TELTIQERDTVINEIGQLIFDAYRVQ-
gi296113803       KSLFKNQKDEIMAQIGKEIYPFL----
gi283784953       RQQDAKNRKDVLAAAAKIVTEGR----
gi206577539       TPASMAERNQHIAGIGAALIEHWQR--
gi288935596       TPASMAERNQHIAGIGAALIEHWQR--
gi92116067        ATVSADRQNAVMAAVGRAVATALA---
gi75674549        ATVPADRQNAVMAAIGRAVTTVLA---
gi284032117       PAAEAEHDNRLLAEAARVVAAAYAR--
gi18466603        SQATMDERNRQIAEIGASLIKHW----
gi134044832       TPASMAERNQQIAGIGAALIEHWQR--
gi123442290       REKDAKPRRDVLASATQIILSQIS---
gi326562148       KSLFKNQKDEIMAQIGKEIYPFL----
gi83645112        TEAAFDDRNKAIADVGRAVGAAFQP--
gi289667228       AKVEDAARDAALKAVAVAVRDSIAA--
gi166713005       AKVDDAARDAALKAVAVAVRESIAGCV
gi221067385       SPLDGAGREAVLARAGEAIKRWYYAI-
gi171318562       STGSLASRKEVIAGVGGIIAKTFGA--
gi170737715       TRADARAKDDVIASVARIVAQTFG---
gi221209528       ARADAKAKDDVIAAATRIAIATLG---
gi126443150       RNPNAQARDDVIASATRIAARAFA---
gi307726917       TAGTAARCNAAIADVGALVVESV----
gi344208573       ASVDEDQRGAVLARVGALADAMIG---
gi167839554       RSQNAPVRDDVIAAAARIAARAFA---
gi134294126       ARADAKAKENVIAAAARIAVATFA---
gi114331058       GRADRPALVAAHASISHLVAHHFA---
gi295690758       GG-DEAARNAVHANVARAVVAAGFGA-
gi167644962       SRLDSDGRDAVIAEVARLVVAAGYAT-
gi186356860       ATASVKDLNSIFAEVGKMITEMV----
gi209548458       ATAPVKELNPVFAEVGRMIMEMV----
gi116251120       ATVPAKELNPVFAEVGRMIVEMV----
gi110634564       PQLPAPKRDGIFREIGEIASEMADLN-
gi53712542        SEADNREN SKIIAEISRIVYEYVTQQI
gi29349915        SEKNNQEN SKIIADISRIVYEYVTBQ-
gi150005764       SKENMETNAKIISDISAAVYRYAGNR-
gi302338706       SKESYETNEKIIADITKLVWNYFTHKA
gi19704905        MS-NLQCRQT-IGKISKFVYDKMEES L
```

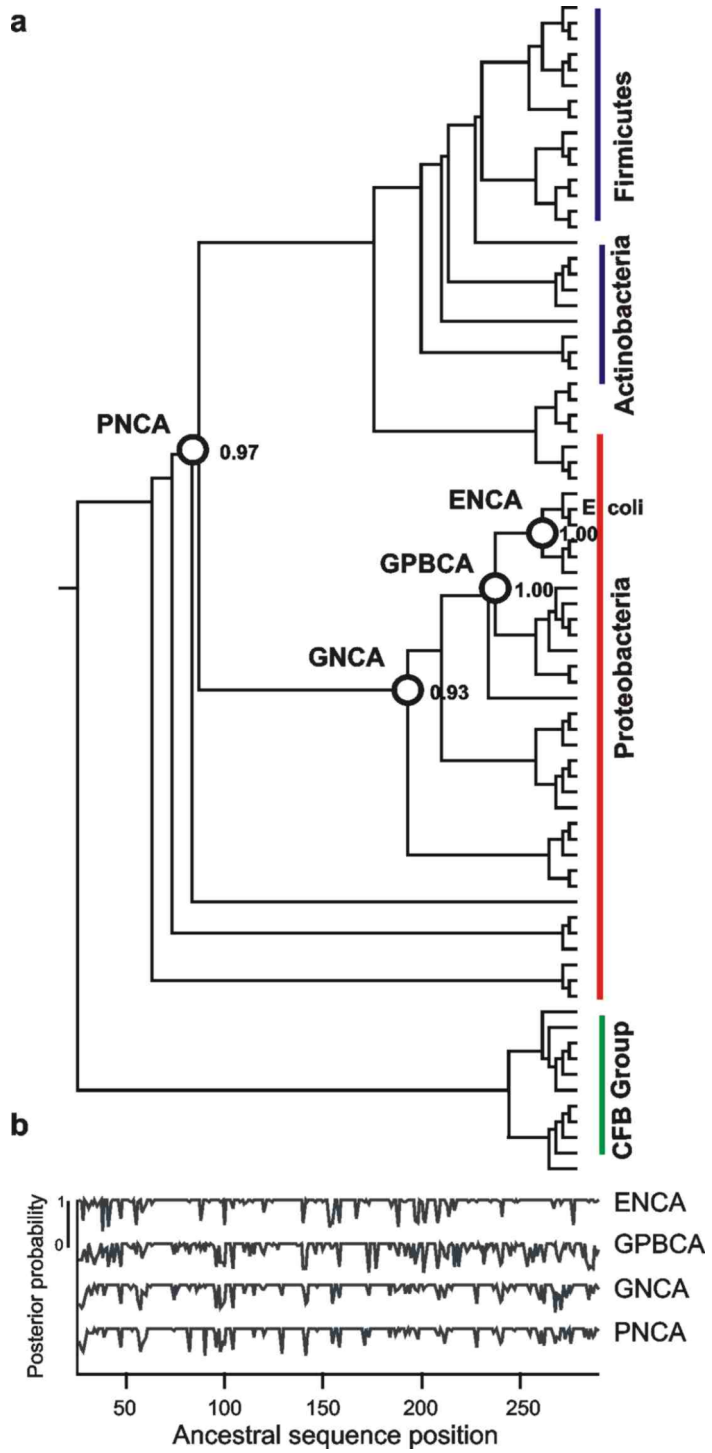


Figure S2 Phylogenetic analysis of β-lactamases and ancestral sequence reconstruction. (a) Cladogram of the tree used as a basis for the laboratory resurrection of Precambrian β-lactamases. The internal nodes targeted are labeled and their posterior probabilities shown. **(b)** Posterior probability profiles of the inferred amino acids across the 262 positions for the targeted internal nodes. The amino acid incorporated at each position of the ancient protein is that with the highest posterior probability.

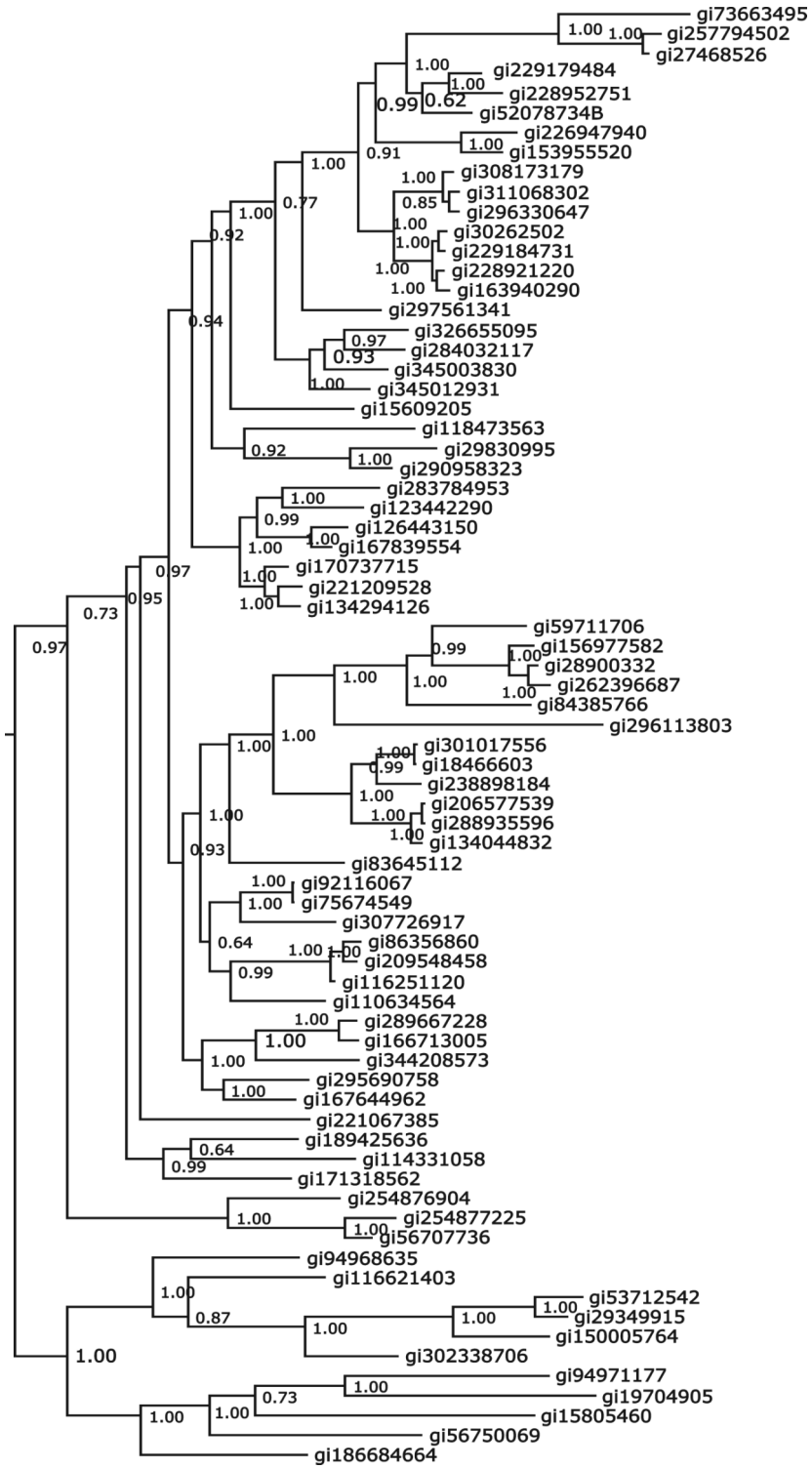


Figure S3 Cladogram of the tree used as a basis for the reconstruction of the sequences of Precambrian β-lactamases. This tree is identical with that in Figure S2, except that the posterior probabilities indicating nodal support are shown for all nodes and the GI numbers for the extant sequences are included (see Table S1). Branch lengths are poportional to distances as measured in amino acid substitutions.

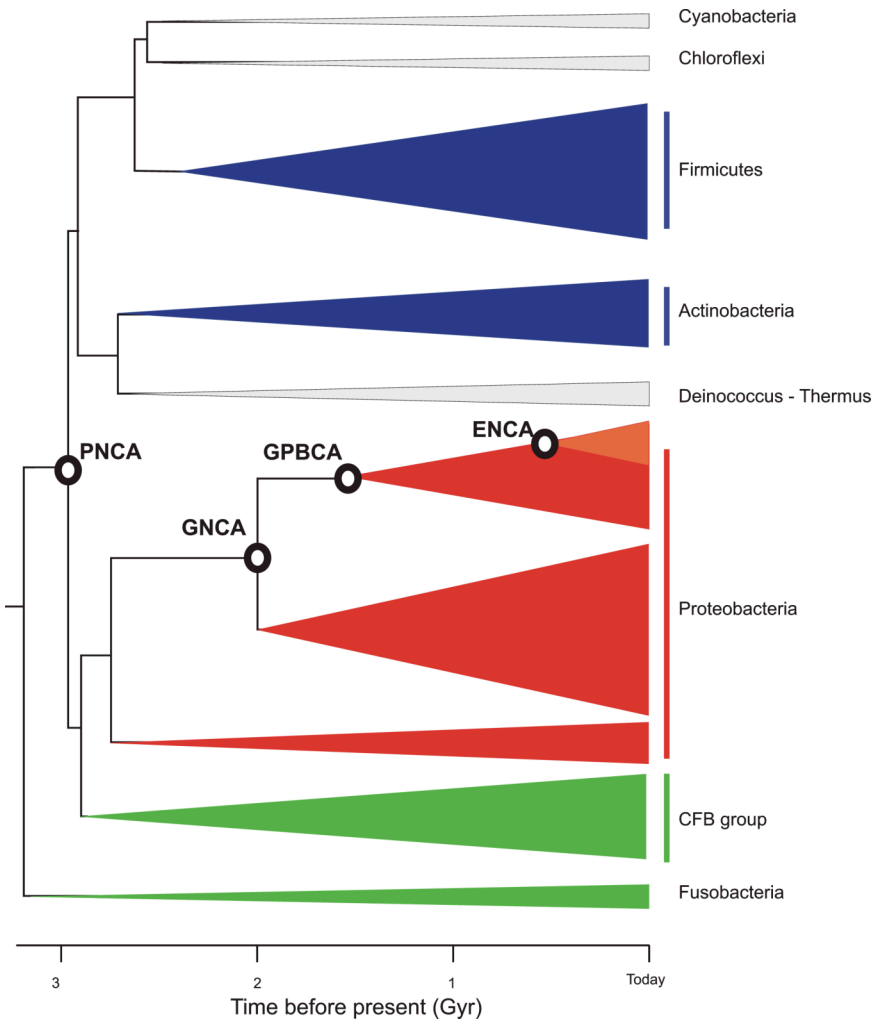


Figure S4 Schematic representation of the tree for Eubacteria given in the Time Tree of Life. The Precambrian nodes targeted are labeled and geological time scale has been added.

```
TEM-1      A GQE QLG RRI HYS QND LVE YSP VTE KHL TDG MTV REL CSAA IT MSD NTAA NLLL TT IGG P
ENCA       A GLE QLD RRI HYR QQD LVE YSP VTE KHL ADG MTV AEL CAAA IT MSD NTAA NLLL A TIGG P
GPBCA      A GKE NLD RRI SYR KED LVD YSP VTE KHV GDG MTV AEL CEAA IT MSD NTAA NLLL E AIGG P
GNCA       Q GKE RLD RRI TYG KED LVD YSP VTE KHV GDG MTV AEL CEAA IT LSD NTAA NLLL E ALGG P
PNCA       Q GKE RLD RRI TYS KAD LVD YSP VTE KHV GGG MTV AEL CEAA IT YSD NTAA NLLL E ALGG P
           * *.*.*** * : ***:********: .**** *** **** ********** ::***

TEM-1      K ELT AFL HNM GDH VTR LDR WEP ELN EAI PNDE RD TTM PVA MATT LR KLL TGE LLT LAS RQ
ENCA       A GLT AFL RNI GDN VTR LDR WET ELN EAL PGDE RD TTT PAA MAA TLR KLL TGE ILS AAS RQ
GPBCA      A GLT AFL RSI GDE VTR LDR WEP ELN EAA PGD PRD TTT PAA MAA TLR TLLL GD VLS PAS RQ
GNCA       AA LT AFL RSI GDE VTR LDR WEP ELN EAA PGD PRD TTT PAA MAA TLR TLLL GD ALS PAS RQ
PNCA       AA LT AFL RSI GDE VTR LDR WEP ELN EAL PGD PRD TTT PRA MAA TLR KLLL G DAL SPA SRE
            *****:.:**.********.***** *.* **** * ***:***.** *: *: ***:

TEM-1      Q LID WME ADK VAG PLL RSA LPA GWF IAD KSG AGE RGS RG IIAA LG PDG KPS RIVVI Y TTG
ENCA       Q LIT WMV ADK VAG PLL RSVL PA GWF IAD KTG AG ERG SRG IIA VLG PDG KPS RIVVI Y LTE
GPBCA      Q LVA WLV ANK TGD KRL RAG LP ADW RIG DKT GAGG YG STN DIA VIW PPG RAP IIV TVY LTE
GNCA       Q LVD WLV ANK TG DKR LRA GLP ADW RVG DKT GTGG HG TTN DIA VIW PPG RAP IVV TVY LTE
PNCA       Q LVD WLL AN KTG DKR LRA GLP ADW RVG DKT GTGG HG TTN DIA VIW PPN RAP IVV AVYY TE
          **: *: *:*...  **: ***.* :.**:*:*  *: . **.: * .:.. :*.:* *

TEM-1      S QAT MDE RN RQI AEI GAS LIK HW
ENCA       T QAS MDE RN QQI AEI GAA LIE HW
GPBCA      T QAS MDE RN KVI AEV GRA VVE AF
GNCA       S QVD AD ARD AVI AEV GRLVV E AF
PNCA       S QAD AE ARD AVI AEV GRLV AE AF
          :*.  : *:  ***:*  :     : :
```

Figure S5 Sequence of TEM-1 β-lactamase and reconstructed sequences for the targeted Precambrian nodes (see Figures S2 and S4). The reconstructed sequences shown are the most probabilistic from each node (i.e., the sequences with the most probable amino acid at each position).

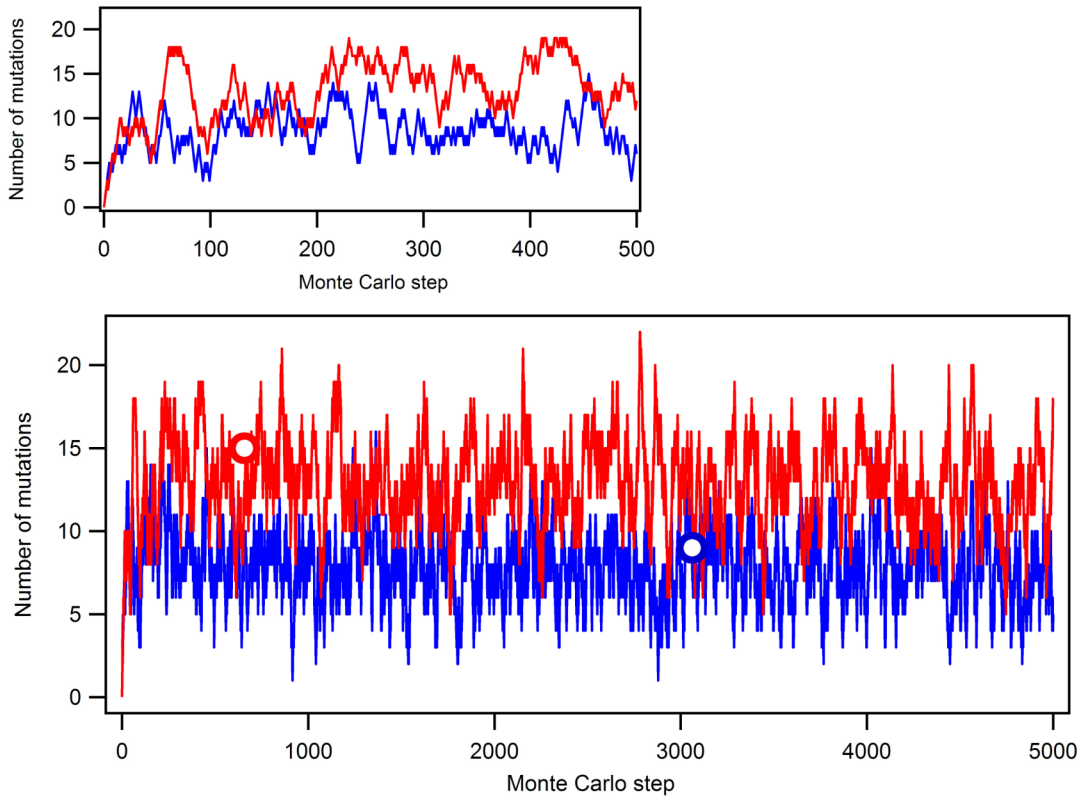


Figure S6 Sequence Monte Carlo sampling based on the posterior amino acid probability distribution for the GNCA node. At each step, a position and amino acid are randomly selected and a replacement to a particular amino acid is accepted or rejected according to its posterior probability at the selected position. This procedure may occasionally introduce very-low probability mutations which are likely to be disruptive (and which would be rejected during the course of natural evolution). To prevent this from taking place, a probability threshold is imposed in the Monte Carlo runs, in such a way that mutations with probabilities below the threshold are always rejected. Thresholds were actually defined in terms of the ratio between the probability and the highest probability at the position. The Monte Carlo runs shown were obtained with ratios of 0.5 (blue) and 0.2 (red) and were used to obtain alternative sequences at the GNCA node in order to test phenotypic robustness (see Table S4 and Figure S9). These sequences were derived from the steps labeled with circles and are shown in Figure S7.

```
GNCA-0     AAQLSEQLAELEKRSGGRLGVAVLDTATGRRIAYRGDERFPMCSTFKALLAAAVLARVDQ
GNCA-1     ASQLSEQLAELEKRSGGRLGVAVLDTATGRRIAYRGDERFPMCSTFKALLAAAVLARVDQ
GNCA-2     AAALSEQLAELEKRSGGRLGVAVLDTATGRRFGYRGDERFPMCSTFKALLAAAVLARVDQ
           *: ****************************:.****************************

GNCA-0     GKERLDRRITYGKEDLVDYSPVTEKHVGDGMTVAELCEAAITLSDNTAANLLLEALGGPA
GNCA-1     GKERLDRRITYSKADLVDYSPVTEKHVGDGMTVAELCEAAITMSDNTAANLLLESLGGPA
GNCA-2     GKENLDRRITYGKEDLVDYSPVTEKHVGDGMTVAELCEAAITYSDNTAANLLLEALGGPA
           ***.*******.* **************************** ***********:*****

GNCA-0     ALTAFLRSIGDEVTRLDRWEPELNEAAPGDPRDTTTPAAMAATLRTLLLGDALSPASRQQ
GNCA-1     ALTAFLRSIGDEVTRLDRWEPELNEAAPGDPRDTTTPAAMAATLRTLLLGDALSPASRQQ
GNCA-2     ALTAFLRSIGDNVTRLDRWEPELNTAAPGDPRDTTTPAAMAATLRTLLLGDVLSPASRQQ
           ***********:************ **************************.********

GNCA-0     LVDWLVANKTGDKRLRAGLPADWRVGDKTGTGGHGTTNDIAVIWPPGRAPIVVTVYLTES
GNCA-1     LVDWLIANKTGDKRLRAGLPADWRVGDKTGTGGHGTTNDIAVIWPPGRAPIVVTVYLTEA
GNCA-2     LVDWLIANKTGDKRLRAGLPADWRVGDKTGTGEHGTTNDIAVIWPPNRAPILVAVYLTES
           *****:*************************** *************.****:*:*****:

GNCA-0     QVDADARDAVIAEVGRLVVEAF
GNCA-1     QADADARDAVIAEVARLVVEAF
GNCA-2     QVDADARDAVIAEVARLVVAAF
           *.************.**** **
```

Figure S7 Reconstructed sequences of β-lactamases corresponding to the GNCA node. The most probabilistic sequence is shown (GNCA-0), together with the sequences of two alternative reconstructions (GNCA-1 and GNCA-2) derived from a Monte Carlo sampling based on the posterior amino acid distribution at the node (see Figure S6). The number of amino acid differences between GNCA-1 and GNCA-2 sequences compared with the GNCA-0 are 9 and 15 sites, respectively. Experimental measurements on the proteins corresponding to these three sequences were used to test the phenotypic robustness of the laboratory resurrection process (see Table S4 and Figure S9).

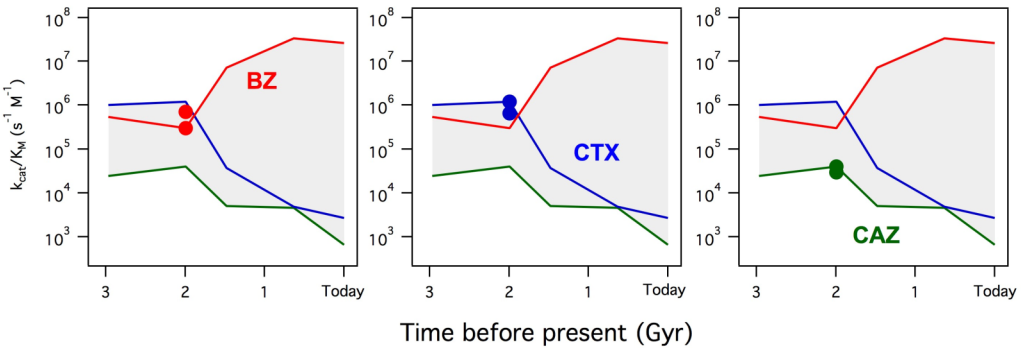


Figure S8 . Second-order rate constant for the in vitro hydrolysis of antibiotics catalyzed by three alternative reconstructions (GNCA-0, GNCA-1 and GNCA-2: see Figure S7) at the GNCA node. The data are plotted in a schematic representation of Figure 3A in the main text to emphasize that the uncertainty associated to the reconstruction does not signficantly affect the overall phenotypic picture.

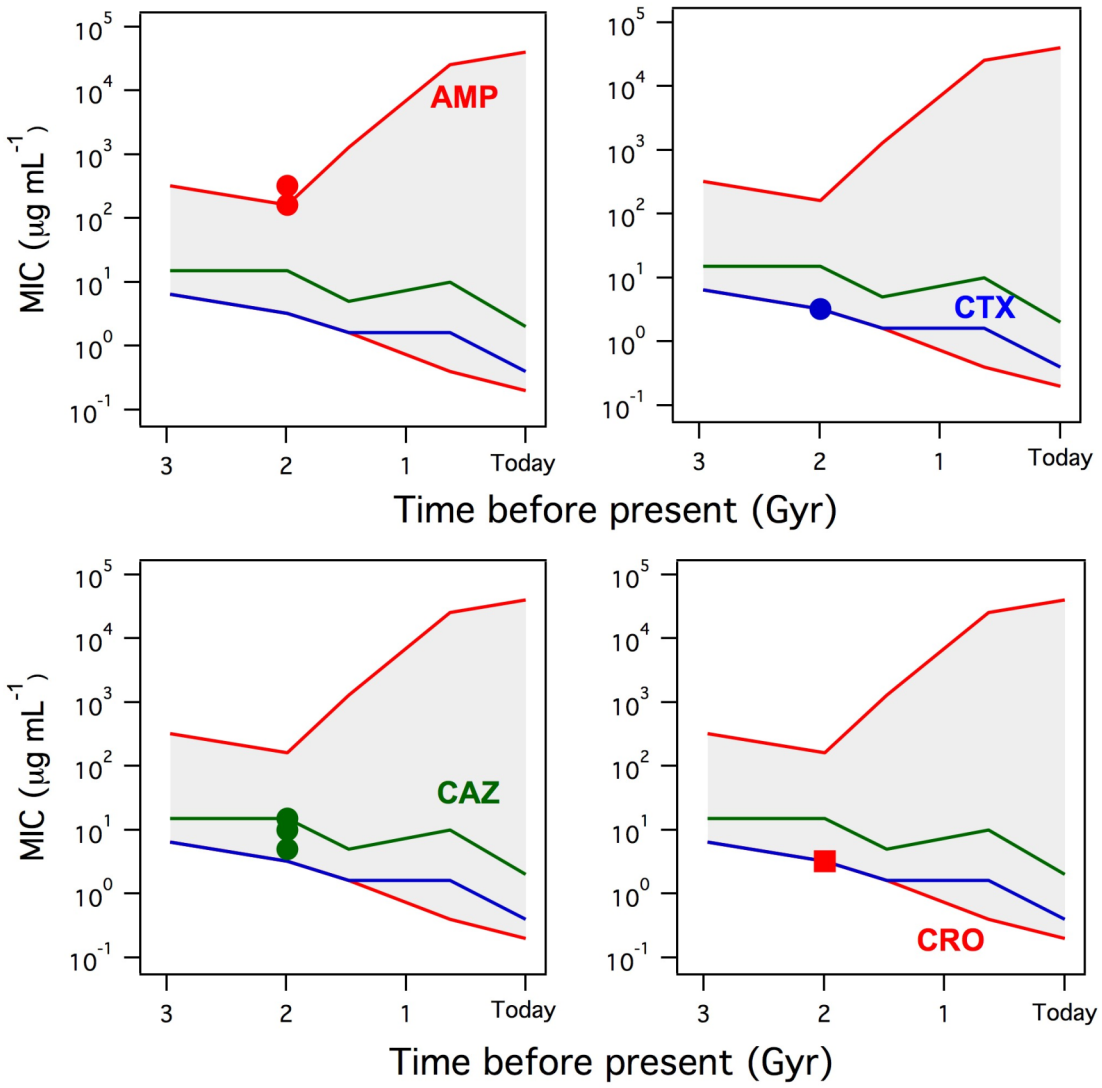


Figure S9 MIC values determined using an *E. coli* strain tranformed containing three alternative reconstructions (GNCA-0, GNCA-1 and GNCA-2: see Figure S7) at the GNCA node. The data are plotted in a schematic representation of Figure 3D in the main text to emphasize that the uncertainty associated to the reconstruction does not signficantly affect the overall phenotypic picture.

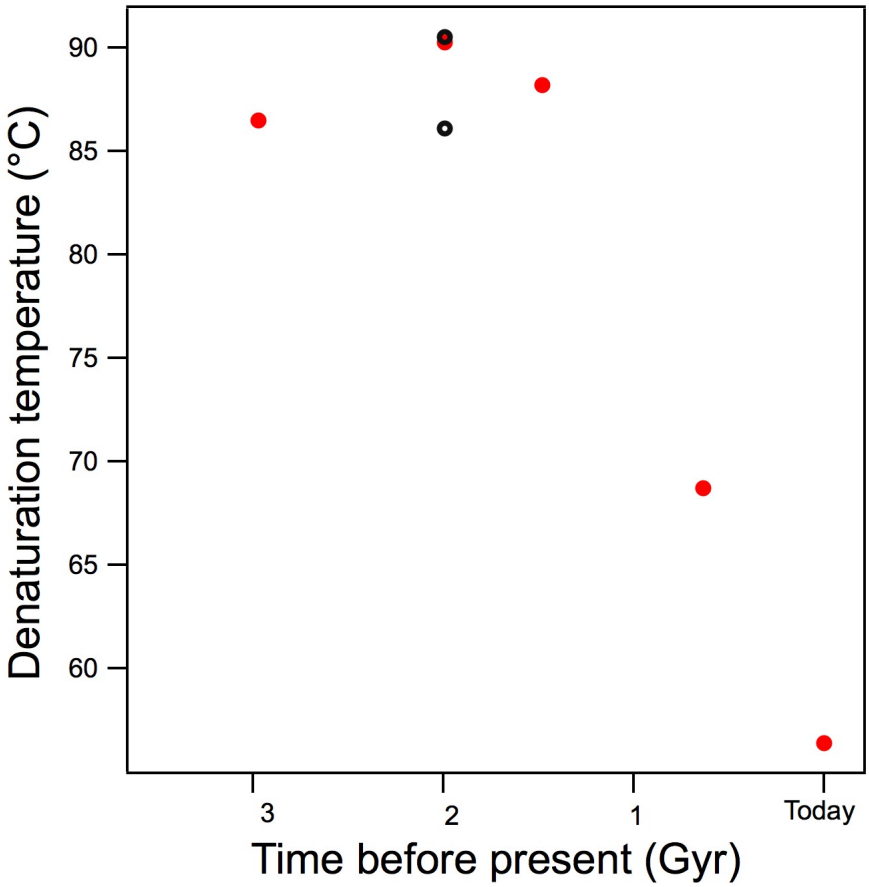


Figure S10 Denaturation temperature values for three alternative reconstructions (GNCA-0, GNCA-1 and GNCA-2: see Figure S7) at the GNCA node. The data are plotted in a schematic representation of Figure 2A in the main text to emphasize that the uncertainty associated to the reconstruction does not signficantly affect the overall phenotypic picture.

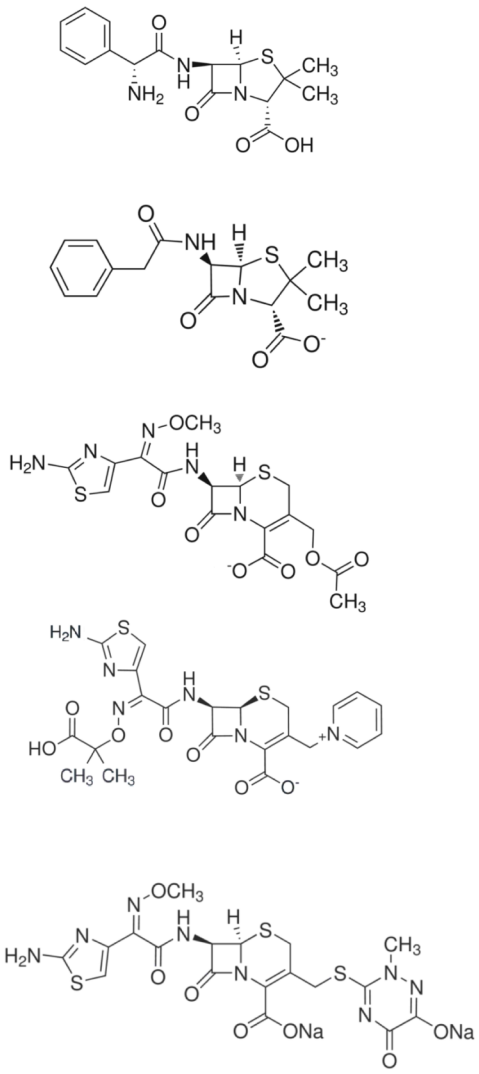


Figure S11 Chemical structures of the β-lactam substrates used in this work.

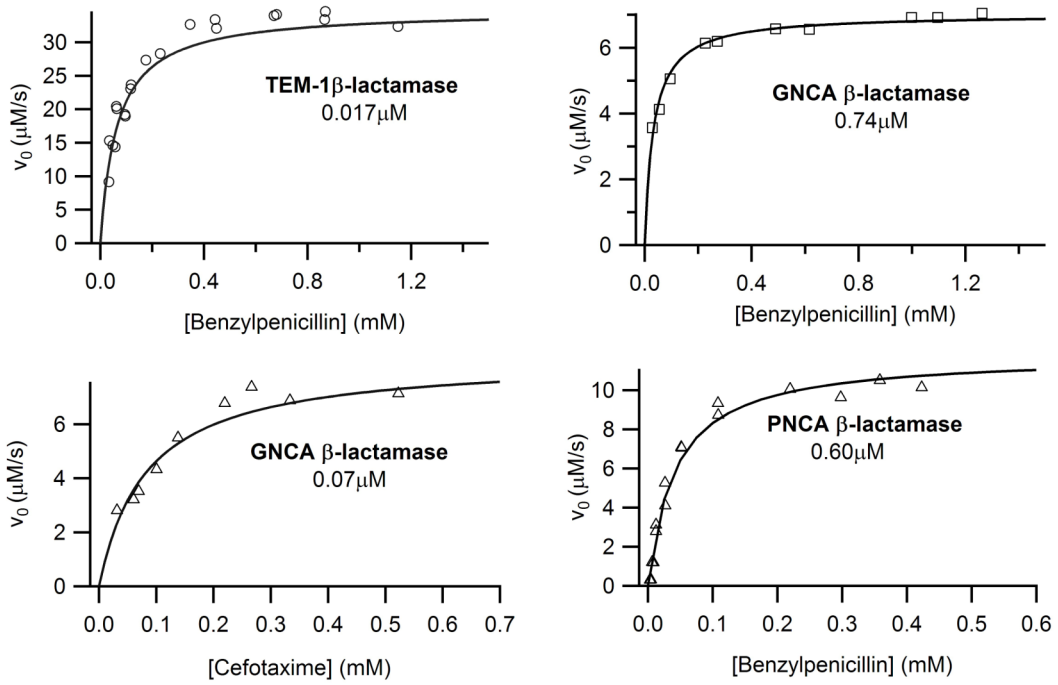


Figure S12 Catalysis of *in vitro* hydrolysis of antibiotics by extant and laboratory resurrections of Precambrian β-lactamases. Representative plots of initial rate versus substrate concentration. The rate values were derived from the time dependence of the UV absorbance associated with antibiotic hydrolysis. For experiments at cefotaxime concentrations > 250 μM, 0.2 cm pathlength cells were used to avoid absorbance values that were too high. The continuous lines are the best fits of the Michaelis-Menten equation. The values obtained for the Michaelis-Menten parameters are collected in Tables S4 and S5.

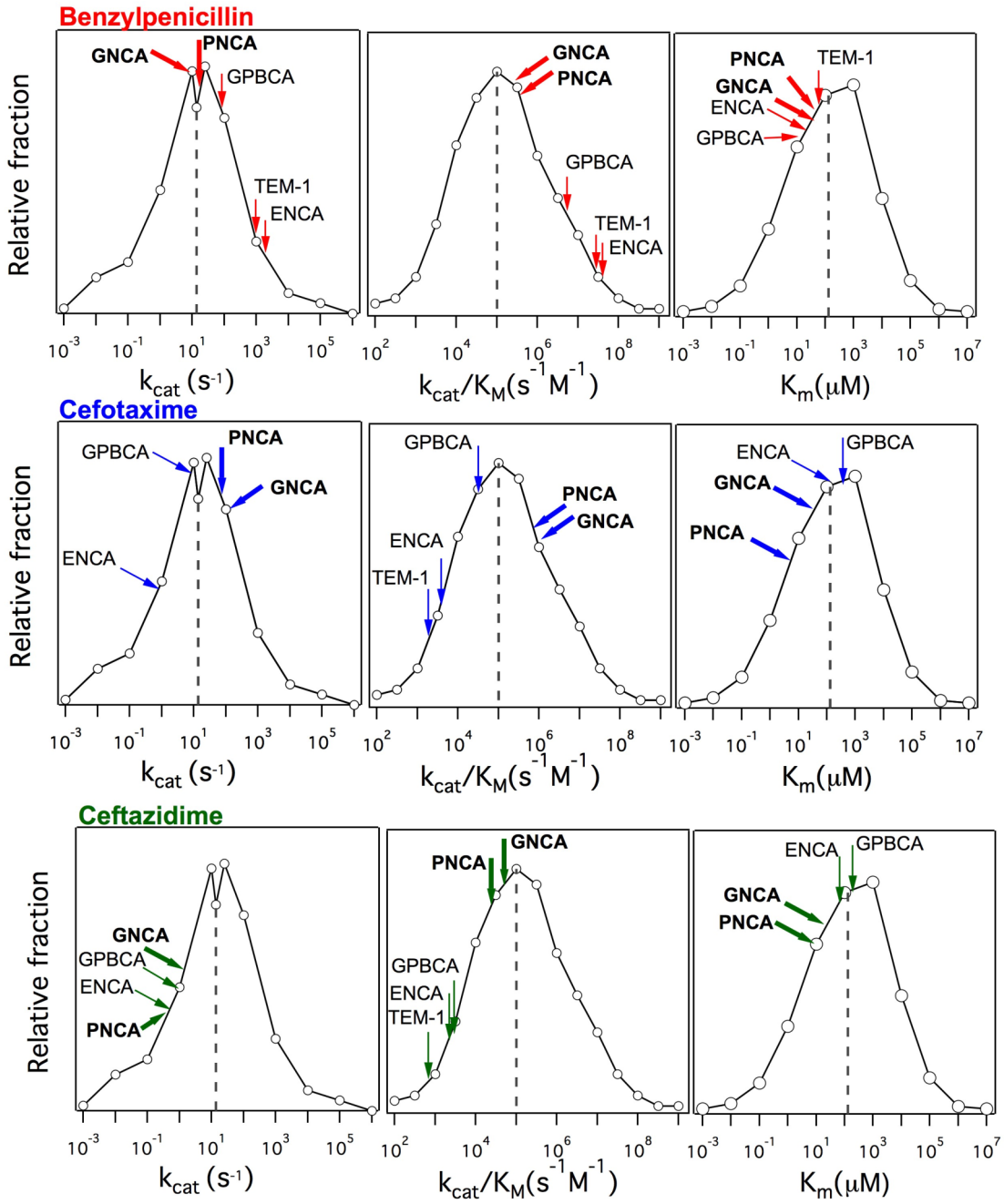


Figure S13 Distribution of kinetics parameters for moderm enzymes with added arrows that locate the values for the TEM-1 and the resurrected β-lactamases. Open circles represent values taken from Bar-Even et al. (ref. 15 in the main text). The antibiotics studied are: benzylpenicillin, and the third-generation cefotaxime and ceftazidime (Figure S11). Thicker arrows are used for the laboratory resurrections of the 2-3 billion-years-old promiscuos lactamases.

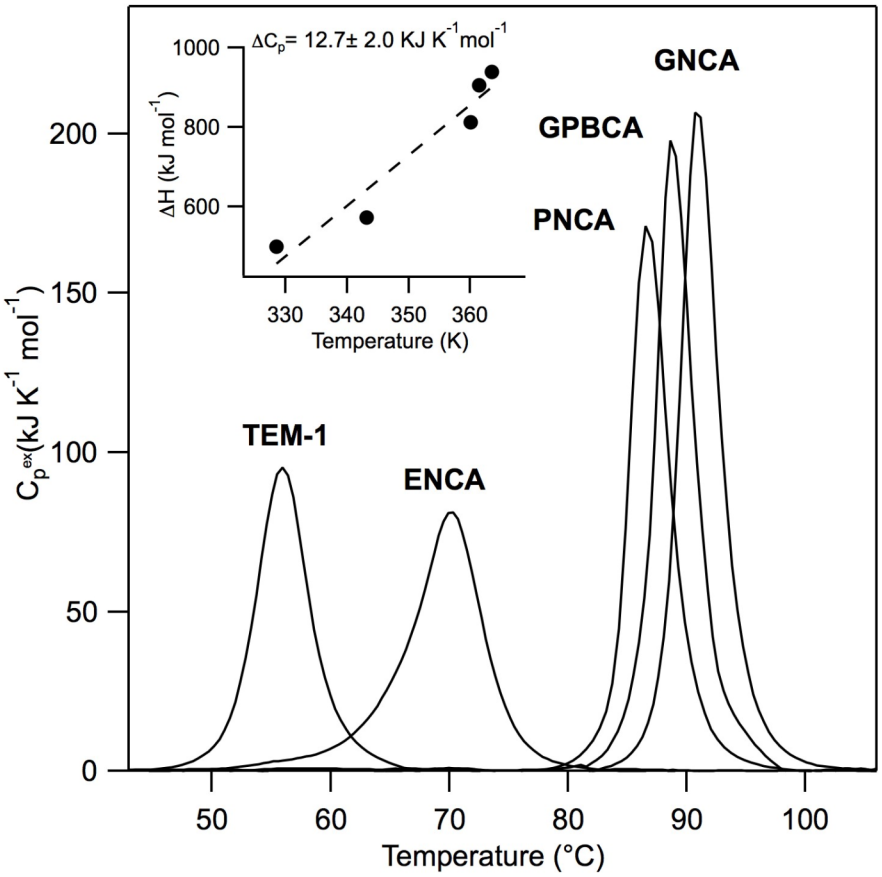


Figure S14 Thermal denaturation of extant and laboratory resurrected Precambrian β-lactamases as followed by differential scanning calorimetry (DSC). The DSC profiles shown were obtained at a scan-rate of 200 K/hour and a protein concentration of 1 mg/mL. Additional experiments within the range 0.1-1.5 mg/mL indicated the absence of a significant concentration dependence of the DSC profiles. On the other hand, experiments at different scanning rates within the range 60-200 K/hour did reveal changes in denaturation temperature (T_m). These scan-rate effects on T_m are significant, in particular for the more recent β-lactamases (see Figure S15), but they are still much smaller that the differences in T_m between the extant and laboratory resurrections of Precambrian β-lactamases. The difference in transition enthalpy (area under the heat capacity profile) visually apparent in the plots shown here is due to the increase in denaturation enthalpy with temperature associated with a positive value of the unfolding heat capacity change. Actually, a plot of denaturation enthalpy versus denaturation temperature (inset) gives a slope consistent with the unfolding heat capacity change expected for a protein of the size of β-lactamase.

Denaturation temperature values given in Fig 2A for the TEM-1 and Precambrian lactamases are actually derived from the transitions shown here. T_m values for the extant β-lactamases from *B. licheniformis* (unpublished work from our lab), *E. cloacae*[19] and *B. cereus*[20] are also included in Fig 2A. *B. licheniformis* is a thermophilic organims, which accounts for its somewhat enhanced Tm value (see Fig. 2A) which is, nevertheless, about

25 degrees below the Tm values for the GNCA β-lactamase. Published kinetic parameters [21] support that the β-lactamase from *B. licheniformis* is a pencillin specialist.

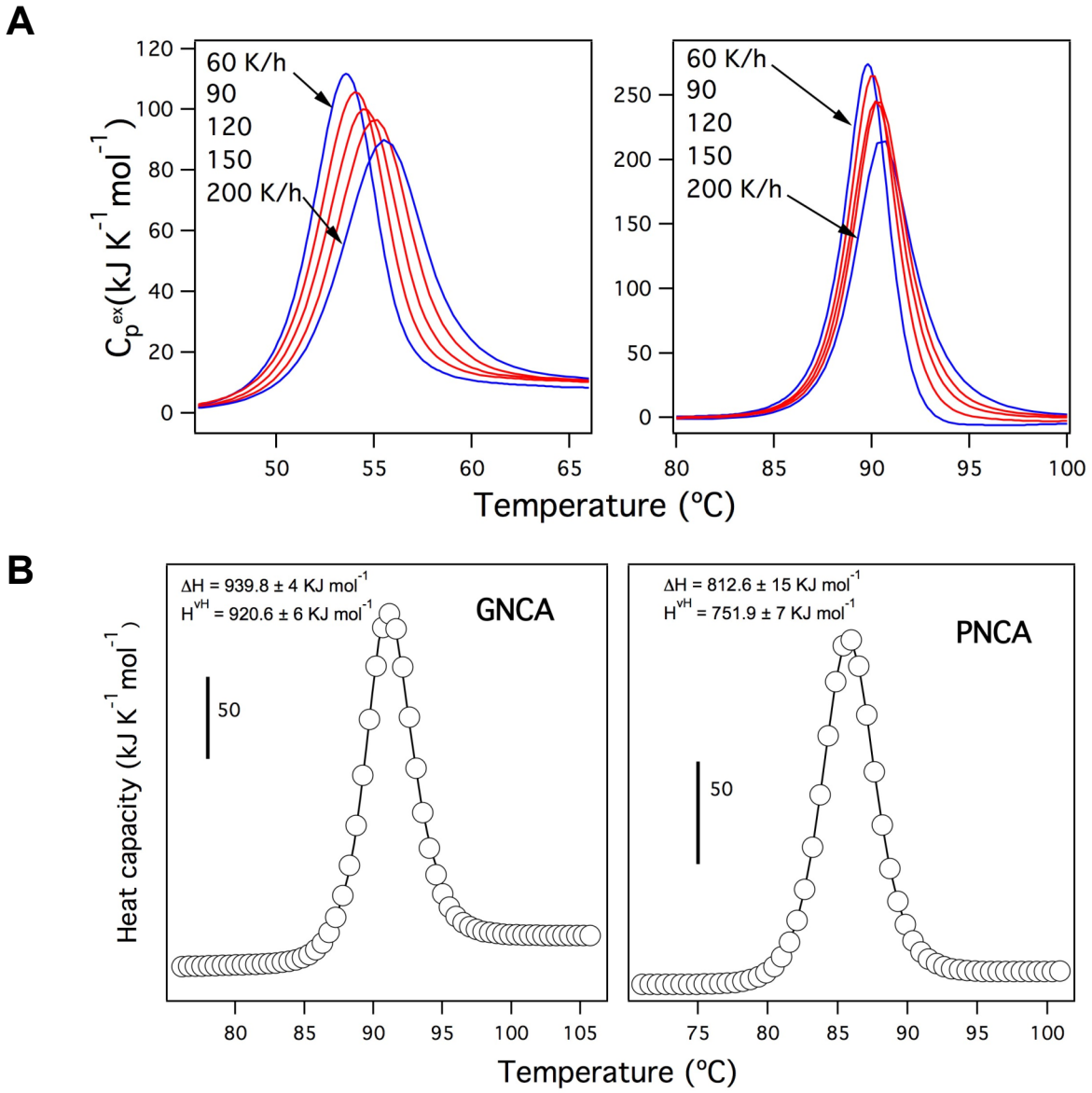


Figure S15 Thermal denaturation of extant and laboratory resurrected Precambrian β-lactamases as followed by differential scanning calorimetry (DSC). (A) Representative DSC profiles illustrating the effect of scanning rate (left panel: TEM-1, right panel: GNCA). (B) Fits of a pseudo two-state model to the DSC profiles corresponding to the GNCA and PNCA β-lactamases. Circles are the experimental data and the continuous line represents the best fits. The values of the calorimetric and van't Hoff enthalpy values derived from the fits are shown. Note that the agreement between the two enthalpy values supports the adequacy of the two-state model. The DSC profiles analyzed are those obtained at a scan-rate of 200 K/hour, since a high scan-rate minimizes kinetic distortions[12]

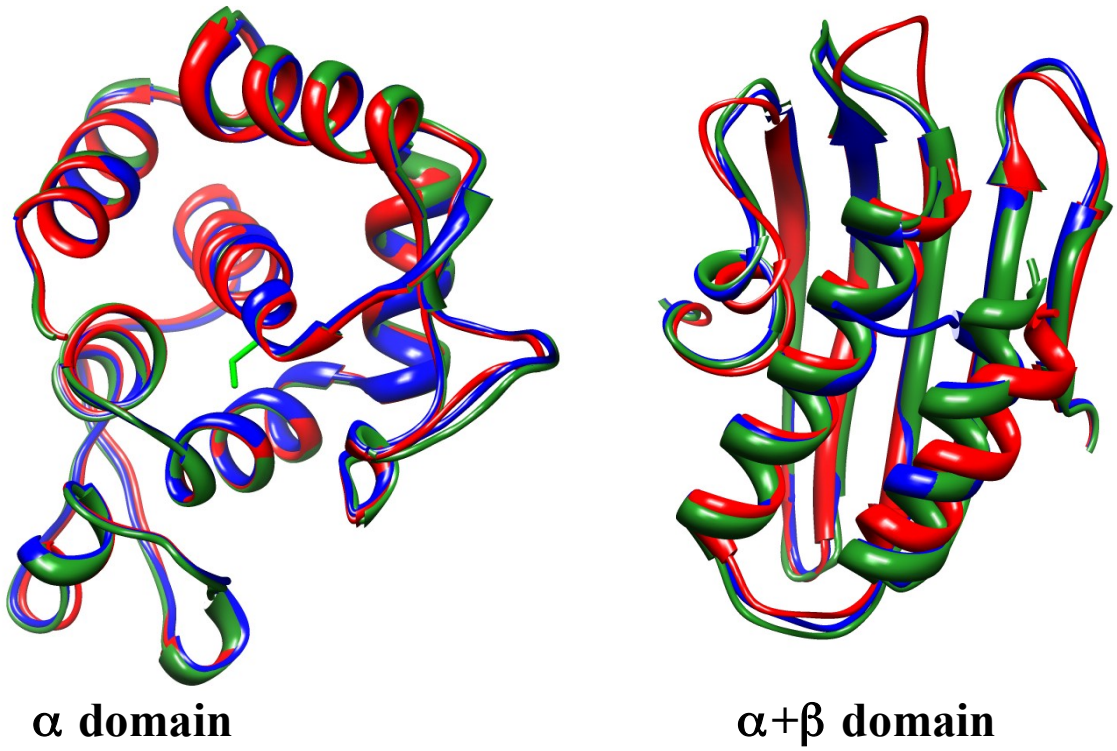


Figure S16 Structural comparison of the α and α+β domains of TEM-1 β-lactamase from *E. coli* (PDB.ID 1btl; green) and the laboratory resurrected β-lactamase corresponding to the last common ancestor of Gram-negative bacteria (GNCA) (PDB.ID 4b88; red) and the last cammon ancestor of enterobacteria (ENCA) (PDB.ID 3zdj; blue). Note that the differences are mostly confined to the α+β domain, while little disparity is observed in the active-site α domain.

Table S1 Sequences of extant β-lactamases class A used for ancestral sequence reconstruction. GI numbers were accessed from the GenBank. The names of the hosting organisms are also provided. Numbers are the amino acid length of the preprotein and the mature protein sequence in the alignment of Figure S1. A pubmed search (lactamase class A "organism name") yields specificity information for only 4 of the extant enzymes: GI301017556 (Frère JM. Mol Microbiol. (1995);(3):385-95), GI52078734 (Brown NG, Pennington JM, Huang W, Ayvaz T, Palzkill T. J Mol Biol. (2010);404(5):832-46), GI257794502 (Gebhard LG, Risso VA, Santos J, Ferreyra RG, Noguera ME, Ermácora MR. J Mol Biol. (2006) ,358(1):280-8) and GI56707736 (Antunes NT, Frase H, Toth M, Vakulenko SB. Antimicrob Agents Chemother. (2012) ;56(2):666-71). These 4 enzymes are reported to have high activity for penicillins and low activity for other antibiotics.

GI 110634564 Chelativorans (290/264)
GI 114331058 Nitrosomonas (319/263)
GI 116251120 Rhizobium (304/262)
GI 116621403 Solibacter (281/257)
GI 118473563 Mycobacterium (293/263)
GI 123442290 Yersinia (294/263)
GI 126443150 Burkholderia (314/263)
GI 134044832 Yersinia (286/265)
GI 134294126 Burkholderia (297/263)
GI 150005764 Bacteroides (319/264)
GI 153955520 Clostridium (310/264)
GI 15609205 Mycobacterium (303/261)
GI 156977582 Vibrio (283/265)
GI 15805460 Deinococcus 1 (265/253)
GI 163940290 Bacillus (309/261)
GI 166713005 Xanthomonas (302/266)
GI 167644962 Caulobacter (300/265)
GI 167839554 Burkholderia (294/263)
GI 170737715 Burkholderia (312/263)
GI 171318562 Burkholderia (305/264)
GI 18466603 Salmonella (286/263)
GI 186684664 Nostoc (339/263)
GI 189425636 Geobacter (300/266)
GI 19704905 Fusobacterium (264/259)
GI 206577539 Klebsiella 342 (286/265)
GI 94971177 Koribacter (293/259)
GI 209548458 Rhizobium (296/262)
GI 221067385 Comamonas (301/265)
GI 221209528 Burkholderia (302/263)
GI 226947940 Clostridium (308/264)
GI 228921220 Bacillus (312/261)
GI 228952751 Bacillus (320/264)
GI 229179484 Bacillus (301/264)
GI 229184731 Bacillus (312/261)
GI 238898184 Hamiltonella (286/265)

GI 254876904 Francisella (291/265)
GI 254877225 Francisella (287/265)
GI 257794502 Staphylococcus (281/260)
GI 262396687 Vibrio (283/265)
GI 27468526 Staphylococcus (284/260)
GI 283784953 Citrobacter (295/262)
GI 284032117 Kribbella (268/262)
GI 288935596 Klebsiella (286/265)
GI 28900332 Vibrio (283/265)
GI 289667228 Xanthomonas (297/264)
GI 290958323 Streptomyces (296/264)
GI 29349915 Bacteroides (293/264)
GI 295690758 Caulobacter (296/264)
GI 296113803 Moraxella (309/259)
GI 296330647 Bacillus (304/261)
GI 297561341 Nocardiopsis (305/261)
GI 29830995 Streptomyces (321/263)
GI 301017556 Escherichia coli (286/263)
GI 302338706 Spirochaeta (305/266)
GI 30262502 Bacillus (309/261)
GI 307726917 Burkholderia (315/262)
GI 308173179 Bacillus (304/261)
GI 311068302 Bacillus (304/261)
GI 326562148 Moraxella (307/257)
GI 326655095 Streptomyces (312/262)
GI 344208573 Burkholderia (304/264)
GI 345003830 Streptomyces (311/261)
GI 345012931 Streptomyces (319/261)
GI 52078734 Bacillus (307/264)
GI 53712542 Bacteroides (300/ 265)
GI 56707736 Francisella (287/265)
GI 56750069 Synechococcus (310/257)
GI 59711706 Vibrio (283/266)
GI 73663495 Staphylococcus (290/260)
GI 75674549 Nitrobacter (296/263)
GI 83645112 Hahella (299/264)
GI 84385766 Vibrio (319/265)
GI 86356860 Rhizobium (296/262)
GI 92116067 Nitrobacter (297/263)
GI 94968635 Koribacter (291/265)

Table S2 Sequence differences between the 75 class A β-lactamases used in sequence reconstruction. For each lactamase, the differences with all the extant lactamases included in our analyses are shown in the order specified below. Sequence differences are given as number of aminoacids (total length is 262 amino acids).

Seq->	gi301017556	gi257794502	gi27468526	gi73663495	gi308173179
	gi30262502	gi311068302	gi229179484	gi52078734	gi229184731
	gi296330647	gi228921220	gi228952751	gi163940290	gi238898184
	gi94968635	gi94971177	gi116621403	gi226947940	gi153955520
	gi189425636	gi56750069	gi186684664	gi15805460	gi118473563
	gi15609205	gi29830995	gi326655095	gi290958323	gi345003830
	gi345012931	gi297561341	gi254876904	gi254877225	gi56707736
	gi59711706	gi156977582	gi28900332	gi262396687	gi84385766
	gi296113803	gi283784953	gi206577539	gi288935596	gi92116067
	gi75674549	gi284032117	gi18466603	gi134044832	gi123442290
	gi326562148	gi83645112	gi289667228	gi166713005	gi221067385
	gi171318562	gi170737715	gi221209528	gi126443150	gi307726917
	gi344208573	gi167839554	gi134294126	gi114331058	gi295690758
	gi167644962	gi86356860	gi209548458	gi116251120	gi110634564
	gi53712542	gi29349915	gi150005764	gi302338706	gi19704905

gi301017556	ID	180	182	192	157	164	164	174	171	165	164
165	179	165	69	192	212	198	177	168	163	213	204
203	166	174	169	166	167	162	167	177	188	197	202
155	164	156	155	163	180	164	89	91	140	140	165
2	87	168	179	142	161	165	164	161	157	156	156
140	159	152	161	162	162	148	151	149	151	153	212
217	214	209	217								
gi257794502	180	ID	25	132	156	157	155	150	154	157	155
160	155	157	189	203	214	206	160	161	188	217	207
214	182	177	190	172	188	164	163	162	187	189	181
183	191	181	186	187	190	174	193	194	184	186	162
179	191	179	189	177	189	194	189	189	174	181	175
191	197	172	181	191	186	185	189	187	185	192	211
210	212	200	213								
gi27468526	182	25	ID	128	158	154	157	151	155	153	155
161	158	157	188	205	214	206	160	158	192	217	207
210	182	178	191	172	188	164	162	160	182	185	175
181	192	181	184	186	192	174	190	191	179	184	162
181	187	182	191	174	190	193	190	185	174	179	177
191	197	174	182	190	184	185	188	187	184	191	211
214	214	202	211								
gi73663495	192	132	128	ID	167	168	164	157	162	167	163
167	163	164	190	209	228	210	162	168	195	221	216
214	191	177	200	168	191	171	160	162	175	185	188
175	189	186	191	187	198	177	193	194	179	181	166
191	192	178	197	181	187	186	190	186	174	174	181
182	192	182	180	189	188	187	195	192	191	187	204
207	205	208	218								
gi308173179	157	156	158	167	ID	76	23	104	112	78	22
78	109	74	163	197	206	197	132	118	163	217	196
204	162	138	157	115	155	125	118	114	179	188	185
178	180	174	175	176	186	145	165	167	150	150	123
156	167	152	185	163	171	174	172	159	146	143	149
157	174	146	148	163	168	161	157	155	149	157	206
208	207	203	215								

gi30262502	164	157	154	168	76	ID	78	111	107	8	76
23	102	31	161	195	203	194	125	119	164	217	195
205	169	139	163	115	150	129	115	113	189	186	186
171	174	172	172	168	187	148	165	166	150	151	125
163	167	160	186	159	175	176	173	156	144	141	148
162	178	144	147	166	167	160	160	153	153	160	207
210	211	204	214								
gi311068302	164	155	157	164	23	78	ID	105	113	80	19
81	106	77	171	194	211	197	129	120	170	216	203
206	167	145	163	117	160	128	120	120	179	186	183
184	183	176	179	178	184	152	169	170	157	157	124
163	170	156	183	167	174	176	175	162	153	148	155
160	178	154	153	169	170	165	162	158	155	164	208
212	210	205	216								
gi229179484	174	150	151	157	104	111	105	ID	61	113	106
110	62	109	172	194	208	198	118	115	167	208	197
203	153	145	169	126	160	133	125	118	177	178	179
179	177	177	178	178	188	149	179	180	154	155	135
173	180	161	187	159	172	173	170	162	154	145	152
168	174	154	157	172	169	161	160	156	155	157	204
206	211	200	209								
gi52078734	171	154	155	162	112	107	113	61	ID	109	114
105	81	101	171	199	203	202	134	131	169	204	198
204	158	140	168	124	158	124	124	110	184	182	187
183	184	182	182	186	189	153	177	177	156	153	140
170	179	158	188	158	176	175	175	167	153	141	155
168	174	154	154	174	170	160	163	162	159	161	204
208	214	204	204								
gi229184731	165	157	153	167	78	8	80	113	109	ID	78
22	104	28	163	195	204	194	125	118	162	218	198
206	170	138	162	116	151	130	116	113	190	185	183
169	173	172	172	169	188	149	166	167	151	152	126
164	168	157	187	159	174	177	173	157	146	141	148
162	177	145	148	166	168	161	160	154	153	161	206
209	211	206	213								
gi296330647	164	155	155	163	22	76	19	106	114	78	ID
78	111	72	169	195	210	198	129	120	167	216	202
207	165	141	163	116	158	126	116	116	181	188	183
182	180	172	174	175	185	148	169	170	150	153	122
163	170	155	184	160	173	174	175	159	149	144	149
159	176	149	152	167	168	164	160	156	153	163	206
210	209	206	216								
gi228921220	165	160	161	167	78	23	81	110	105	22	78
ID	104	21	165	199	207	199	125	122	165	219	200
206	170	141	166	117	154	133	119	116	185	186	187
171	172	169	168	166	183	149	171	171	147	151	128
164	171	158	182	158	173	177	172	156	147	140	149
160	177	151	152	168	169	163	157	152	152	161	204
205	210	204	211								
gi228952751	179	155	158	163	109	102	106	62	81	104	111
104	ID	104	172	192	212	194	129	126	163	207	199
205	163	135	163	125	154	137	123	121	178	177	178
183	183	180	183	182	188	152	180	181	162	160	133
178	181	158	187	163	176	177	177	164	154	149	158
171	178	161	161	173	170	157	167	165	163	162	209
211	214	201	209								

gi163940290	165	157	157	164	74	31	77	109	101	28	72
21	104	ID	167	198	206	197	122	118	165	217	199
208	168	139	161	117	151	130	117	114	187	187	184
170	174	170	169	169	184	153	169	169	149	152	124
164	169	155	183	159	172	176	169	157	148	140	145
158	177	150	152	166	168	161	158	153	151	160	206
209	210	208	216								
gi238898184	69	189	188	190	163	161	171	172	171	163	169
165	172	167	ID	196	205	201	171	167	163	209	204
207	166	167	167	163	165	170	164	175	189	190	199
157	159	155	154	154	181	165	89	91	137	137	169
68	91	166	180	139	160	163	164	159	158	154	151
140	163	151	159	161	165	151	153	149	151	149	211
216	212	206	220								
gi94968635	192	203	205	209	197	195	194	194	199	195	195
199	192	198	196	ID	210	157	201	200	198	207	181
215	191	193	202	194	199	200	194	205	205	199	206
206	199	196	197	202	203	197	192	193	189	189	194
192	194	193	202	198	201	201	200	199	187	194	199
193	205	200	187	186	190	185	187	185	188	205	184
192	184	171	214								
gi94971177	212	214	214	228	206	203	211	208	203	204	210
207	212	206	205	210	ID	208	210	209	215	194	199
191	208	212	207	210	211	203	215	201	217	218	217
203	212	210	212	216	213	212	210	211	211	211	214
211	210	212	212	210	211	213	208	208	205	202	200
208	211	198	205	209	208	218	213	209	209	214	207
206	217	208	186								
gi116621403	198	206	206	210	197	194	197	198	202	194	198
199	194	197	201	157	208	ID	202	196	188	211	190
204	194	187	201	187	195	195	197	198	208	210	205
210	205	202	199	209	205	200	197	197	182	181	194
198	197	208	204	186	183	185	192	184	187	184	187
186	186	182	181	191	186	185	189	187	186	191	190
197	191	160	216								
gi226947940	177	160	160	162	132	125	129	118	134	125	129
125	129	122	171	201	210	202	ID	71	161	214	207
211	174	151	176	143	170	144	131	141	174	175	175
178	179	176	178	178	189	163	178	179	164	169	141
176	178	171	188	171	175	176	180	155	163	158	161
165	182	163	163	165	180	169	175	172	172	165	209
213	210	203	215								
gi153955520	168	161	158	168	118	119	120	115	131	118	120
122	126	118	167	200	209	196	71	ID	164	213	201
208	170	144	171	136	162	146	137	142	178	176	175
178	178	174	177	175	190	161	173	174	160	167	138
168	175	168	189	170	179	181	181	153	165	158	163
164	181	163	163	168	180	169	164	160	161	170	209
209	213	203	224								
gi189425636	163	188	192	195	163	164	170	167	169	162	167
165	163	165	163	198	215	188	161	164	ID	205	194
213	167	143	154	162	159	155	158	161	181	178	177
178	172	173	173	176	188	149	162	163	140	139	159
162	163	137	187	153	152	155	154	134	141	144	138
143	145	135	150	134	148	139	144	144	141	141	206
213	206	204	212								

gi56750069	213	217	217	221	217	217	216	208	204	218	216
219	207	217	209	207	194	211	214	213	205	ID	184
195	205	209	208	208	206	203	212	208	204	207	210
216	213	206	206	219	214	213	211	212	206	206	216
213	211	207	214	208	212	211	209	210	209	208	211
210	211	214	213	205	199	205	206	204	204	209	214
214	214	217	195								
gi186684664	204	207	207	216	196	195	203	197	198	198	202
200	199	199	204	181	199	190	207	201	194	184	ID
189	194	194	201	195	200	191	198	199	209	215	216
209	204	204	201	206	212	187	196	197	193	188	198
204	199	195	211	196	198	196	188	195	192	187	195
188	201	194	191	187	186	191	193	189	188	195	199
203	205	191	206								
gi15805460	203	214	210	214	204	205	206	203	204	206	207
206	205	208	207	215	191	204	211	208	213	195	189
ID	206	201	206	200	201	203	211	197	213	218	219
211	214	214	212	216	218	211	195	196	200	197	209
203	196	207	217	208	205	204	194	199	202	197	202
203	206	199	204	204	197	209	205	205	203	206	224
225	224	217	200								
gi118473563	166	182	182	191	162	169	167	153	158	170	165
170	163	168	166	191	208	194	174	170	167	205	194
206	ID	162	156	159	153	163	166	165	194	199	195
180	179	179	178	175	200	162	164	165	156	154	163
165	167	158	199	158	165	168	162	161	144	144	151
152	169	147	149	153	154	161	167	164	158	157	204
208	204	200	215								
gi15609205	174	177	178	177	138	139	145	145	140	138	141
141	135	139	167	193	212	187	151	144	143	209	194
201	162	ID	149	138	141	120	127	130	189	178	177
184	183	179	182	177	193	151	171	173	147	149	136
174	175	142	192	155	161	162	155	155	134	131	136
150	158	139	133	159	150	152	153	152	150	145	204
204	205	205	221								
gi29830995	169	190	191	200	157	163	163	169	168	162	163
166	163	161	167	202	207	201	176	171	154	208	201
206	156	149	ID	156	77	149	153	154	192	186	183
184	177	173	172	178	191	160	164	166	142	144	161
168	166	150	190	148	163	169	161	158	147	141	142
144	168	139	146	160	163	154	148	148	147	151	210
213	211	203	218								
gi326655095	166	172	172	168	115	115	117	126	124	116	116
117	125	117	163	194	210	187	143	136	162	208	195
200	159	138	156	ID	135	91	87	107	191	186	189
184	185	181	183	180	195	150	156	156	132	135	84
165	156	159	194	145	163	161	160	163	138	133	148
147	162	142	142	167	157	145	145	147	146	141	205
205	207	203	224								
gi290958323	167	188	188	191	155	150	160	160	158	151	158
154	154	151	165	199	211	195	170	162	159	206	200
201	153	141	77	135	ID	142	148	152	193	178	182
181	176	174	175	176	193	157	165	166	140	145	146
166	168	154	192	147	162	167	160	162	141	141	136
142	174	133	149	164	152	138	152	150	152	146	209
206	205	202	218								

gi345003830	162	164	164	171	125	129	128	133	124	130	126
133	137	130	170	200	203	195	144	146	155	203	191
203	163	120	149	91	142	ID	83	113	191	182	186
192	192	185	185	186	189	155	169	169	142	139	91
161	171	155	188	157	166	169	161	161	138	132	147
148	159	142	142	162	155	149	148	148	149	147	207
210	212	203	222								
gi345012931	167	163	162	160	118	115	120	125	124	116	116
119	123	117	164	194	215	197	131	137	158	212	198
211	166	127	153	87	148	83	ID	104	183	182	179
185	183	181	182	180	183	152	169	170	140	141	89
166	170	147	182	156	160	163	162	156	137	133	140
150	166	143	145	164	162	156	158	157	154	149	208
210	206	208	225								
gi297561341	177	162	160	162	114	113	120	118	110	113	116
116	121	114	175	205	201	198	141	142	161	208	199
197	165	130	154	107	152	113	104	ID	194	195	197
185	186	184	180	189	184	150	176	176	146	145	116
176	174	153	183	158	166	169	166	156	133	131	146
155	164	144	143	157	160	157	156	155	152	152	207
211	212	207	220								
gi254876904	188	187	182	175	179	189	179	177	184	190	181
185	178	187	189	205	217	208	174	178	181	204	209
213	194	189	192	191	193	191	183	194	ID	147	142
193	201	196	198	198	194	173	199	200	180	180	186
188	200	169	193	185	196	199	197	184	177	178	187
182	194	188	185	186	185	186	185	184	184	185	198
200	203	205	207								
gi254877225	197	189	185	185	188	186	186	178	182	185	188
186	177	187	190	199	218	210	175	176	178	207	215
218	199	178	186	186	178	182	182	195	147	ID	61
196	196	194	197	204	188	183	199	200	174	181	186
197	201	176	187	190	186	186	195	190	186	185	192
181	188	192	188	187	190	179	193	190	193	193	200
202	207	191	205								
gi56707736	202	181	175	188	185	186	183	179	187	183	183
187	178	184	199	206	217	205	175	175	177	210	216
219	195	177	183	189	182	186	179	197	142	61	ID
197	199	200	201	203	189	185	199	200	179	183	182
202	200	173	188	192	195	193	195	191	190	189	187
186	191	189	195	188	187	187	194	193	192	201	201
201	205	195	210								
gi59711706	155	183	181	175	178	171	184	179	183	169	182
171	183	170	157	206	203	210	178	178	178	216	209
211	180	184	184	184	181	192	185	185	193	196	197
ID	115	118	121	127	186	172	164	165	167	171	197
154	162	176	185	155	170	173	160	169	175	176	175
167	176	178	181	181	167	167	170	167	162	168	218
220	218	212	220								
gi156977582	164	191	192	189	180	174	183	177	184	173	180
172	183	174	159	199	212	205	179	178	172	213	204
214	179	183	177	185	176	192	183	186	201	196	199
115	ID	47	54	119	175	174	162	163	165	168	191
163	159	177	174	155	162	167	164	160	168	169	166
161	176	172	172	176	172	167	164	164	164	165	216
217	216	210	223								

gi28900332	156	181	181	186	174	172	176	177	182	172	172
169	180	170	155	196	210	202	176	174	173	206	204
214	179	179	173	181	174	185	181	184	196	194	200
118	47	ID	31	125	175	175	153	154	164	166	187
155	150	175	174	150	167	168	165	159	166	170	167
160	177	172	169	174	169	166	163	164	163	165	213
213	215	213	218								
gi262396687	155	186	184	191	175	172	179	178	182	172	174
168	183	169	154	197	212	199	178	177	173	206	201
212	178	182	172	183	175	185	182	180	198	197	201
121	54	31	ID	122	177	176	157	158	161	165	189
154	150	175	176	148	164	167	165	157	165	169	169
158	178	172	170	173	172	166	166	165	165	166	217
217	219	209	218								
gi84385766	163	187	186	187	176	168	178	178	186	169	175
166	182	169	154	202	216	209	178	175	176	219	206
216	175	177	178	180	176	186	180	189	198	204	203
127	119	125	122	ID	177	169	163	164	165	167	185
163	158	180	176	154	173	173	165	172	178	169	165
167	185	174	178	176	180	168	173	170	168	168	206
213	216	203	217								
gi296113803	180	190	192	198	186	187	184	188	189	188	185
183	188	184	181	203	213	205	189	190	188	214	212
218	200	193	191	195	193	189	183	184	194	188	189
186	175	175	177	177	ID	188	187	186	173	171	195
180	186	181	1	188	187	191	193	188	185	182	187
179	181	191	194	187	183	184	182	180	183	177	203
207	208	198	210								
gi283784953	164	174	174	177	145	148	152	149	153	149	148
149	152	153	165	197	212	200	163	161	149	213	187
211	162	151	160	150	157	155	152	150	173	183	185
172	174	175	176	169	188	ID	166	167	137	139	156
162	165	117	187	155	153	152	149	149	114	113	119
145	156	117	125	156	160	141	150	149	146	145	206
211	210	199	217								
gi206577539	89	193	190	193	165	165	169	179	177	166	169
171	180	169	89	192	210	197	178	173	162	211	196
195	164	171	164	156	165	169	169	176	199	199	199
164	162	153	157	163	187	166	ID	3	137	135	169
87	23	167	186	141	155	155	159	154	152	151	153
135	165	151	153	158	150	149	151	146	150	151	210
216	212	209	220								
gi288935596	91	194	191	194	167	166	170	180	177	167	170
171	181	169	91	193	211	197	179	174	163	212	197
196	165	173	166	156	166	169	170	176	200	200	200
165	163	154	158	164	186	167	3	ID	137	136	169
89	24	168	185	141	156	156	159	155	152	152	154
135	166	151	153	158	150	151	151	147	151	153	210
216	212	209	219								
gi92116067	140	184	179	179	150	150	157	154	156	151	150
147	162	149	137	189	211	182	164	160	140	206	193
200	156	147	142	132	140	142	140	146	180	174	179
167	165	164	161	165	173	137	137	137	ID	39	143
138	138	152	172	120	140	142	138	147	123	128	130
105	134	124	133	145	137	121	126	126	123	126	205
205	209	196	219								

gi75674549	140	186	184	181	150	151	157	155	153	152	153
151	160	152	137	189	211	181	169	167	139	206	188
197	154	149	144	135	145	139	141	145	180	181	183
171	168	166	165	167	171	139	135	136	39	ID	144
138	137	148	170	125	138	138	141	146	128	127	135
108	139	129	132	148	139	121	122	123	118	118	201
207	205	200	219								
gi284032117	165	162	162	166	123	125	124	135	140	126	122
128	133	124	169	194	214	194	141	138	159	216	198
209	163	136	161	84	146	91	89	116	186	186	182
197	191	187	189	185	195	156	169	169	143	144	ID
165	169	156	194	155	161	164	166	163	140	139	149
149	164	145	146	157	154	152	159	157	156	152	207
212	206	203	227								
gi18466603	2	179	181	191	156	163	163	173	170	164	163
164	178	164	68	192	211	198	176	168	162	213	204
203	165	174	168	165	166	161	166	176	188	197	202
154	163	155	154	163	180	162	87	89	138	138	165
ID	85	167	179	141	160	164	162	160	156	155	155
139	158	151	160	160	161	148	150	148	150	152	212
217	214	209	216								
gi134044832	87	191	187	192	167	167	170	180	179	168	170
171	181	169	91	194	210	197	178	175	163	211	199
196	167	175	166	156	168	171	170	174	200	201	200
162	159	150	150	158	186	165	23	24	138	137	169
85	ID	165	185	143	155	154	155	155	149	150	150
140	168	149	152	157	156	156	155	151	155	154	211
218	212	208	220								
gi123442290	168	179	182	178	152	160	156	161	158	157	155
158	158	155	166	193	212	208	171	168	137	207	195
207	158	142	150	159	154	155	147	153	169	176	173
176	177	175	175	180	181	117	167	168	152	148	156
167	165	ID	180	153	143	151	148	150	110	99	110
146	143	115	119	159	154	153	157	153	150	148	200
202	208	210	221								
gi326562148	179	189	191	197	185	186	183	187	188	187	184
182	187	183	180	202	212	204	188	189	187	214	211
217	199	192	190	194	192	188	182	183	193	187	188
185	174	174	176	176	1	187	186	185	172	170	194
179	185	180	ID	187	186	190	192	187	184	181	186
178	180	190	193	186	182	183	181	179	182	176	202
206	207	197	210								
gi83645112	142	177	174	181	163	159	167	159	158	159	160
158	163	159	139	198	210	186	171	170	153	208	196
208	158	155	148	145	147	157	156	158	185	190	192
155	155	150	148	154	188	155	141	141	120	125	155
141	143	153	187	ID	143	147	152	156	137	138	142
128	144	137	144	157	140	136	125	132	127	146	203
207	209	203	211								
gi289667228	161	189	190	187	171	175	174	172	176	174	173
173	176	172	160	201	211	183	175	179	152	212	198
205	165	161	163	163	162	166	160	166	196	186	195
170	162	167	164	173	187	153	155	156	140	138	161
160	155	143	186	143	ID	30	154	151	146	138	143
137	124	142	147	146	135	135	153	156	151	155	205
210	202	208	220								

gi166713005	165	194	193	186	174	176	176	173	175	177	174
177	177	176	163	201	213	185	176	181	155	211	196
204	168	162	169	161	167	169	163	169	199	186	193
173	167	168	167	173	191	152	155	156	142	138	164
164	154	151	190	147	30	ID	153	151	150	141	148
141	126	146	149	151	132	134	156	159	156	157	207
213	206	210	222								
gi221067385	164	189	190	190	172	173	175	170	175	173	175
172	177	169	164	200	208	192	180	181	154	209	188
194	162	155	161	160	160	161	162	166	197	195	195
160	164	165	165	165	193	149	159	159	138	141	166
162	155	148	192	152	154	153	ID	152	149	147	147
135	153	150	149	155	147	140	144	143	137	143	216
220	214	213	222								
gi171318562	161	189	185	186	159	156	162	162	167	157	159
156	164	157	159	199	208	184	155	153	134	210	195
199	161	155	158	163	162	161	156	156	184	190	191
169	160	159	157	172	188	149	154	155	147	146	163
160	155	150	187	156	151	151	152	ID	140	140	142
149	157	141	146	147	149	148	143	135	137	149	210
215	208	205	222								
gi170737715	157	174	174	174	146	144	153	154	153	146	149
147	154	148	158	187	205	187	163	165	141	209	192
202	144	134	147	138	141	138	137	133	177	186	190
175	168	166	165	178	185	114	152	152	123	128	140
156	149	110	184	137	146	150	149	140	ID	52	90
134	148	86	49	144	142	134	141	137	134	138	197
206	204	210	223								
gi221209528	156	181	179	174	143	141	148	145	141	141	144
140	149	140	154	194	202	184	158	158	144	208	187
197	144	131	141	133	141	132	133	131	178	185	189
176	169	170	169	169	182	113	151	152	128	127	139
155	150	99	181	138	138	141	147	140	52	ID	89
131	140	86	43	145	139	136	140	136	135	129	199
205	206	205	226								
gi126443150	156	175	177	181	149	148	155	152	155	148	149
149	158	145	151	199	200	187	161	163	138	211	195
202	151	136	142	148	136	147	140	146	187	192	187
175	166	167	169	165	187	119	153	154	130	135	149
155	150	110	186	142	143	148	147	142	90	89	ID
135	148	29	94	152	144	144	149	148	141	140	198
203	207	205	220								
gi307726917	140	191	191	182	157	162	160	168	168	162	159
160	171	158	140	193	208	186	165	164	143	210	188
203	152	150	144	147	142	148	150	155	182	181	186
167	161	160	158	167	179	145	135	135	105	108	149
139	140	146	178	128	137	141	135	149	134	131	135
ID	145	136	137	132	136	128	115	119	114	129	203
208	209	197	217								
gi344208573	159	197	197	192	174	178	178	174	174	177	176
177	178	177	163	205	211	186	182	181	145	211	201
206	169	158	168	162	174	159	166	164	194	188	191
176	176	177	178	185	181	156	165	166	134	139	164
158	168	143	180	144	124	126	153	157	148	140	148
145	ID	144	146	149	140	134	156	153	149	146	210
215	217	214	226								

gi167839554	152	172	174	182	146	144	154	154	154	145	149
151	161	150	151	200	198	182	163	163	135	214	194
199	147	139	139	142	133	142	143	144	188	192	189
178	172	172	172	174	191	117	151	151	124	129	145
151	149	115	190	137	142	146	150	141	86	86	29
136	144	ID	88	153	142	142	147	147	139	137	196
200	203	205	218								
gi134294126	161	181	182	180	148	147	153	157	154	148	152
152	161	152	159	187	205	181	163	163	150	213	191
204	149	133	146	142	149	142	145	143	185	188	195
181	172	169	170	178	194	125	153	153	133	132	146
160	152	119	193	144	147	149	149	146	49	43	94
137	146	88	ID	151	145	143	147	142	141	139	206
209	213	206	227								
gi114331058	162	191	190	189	163	166	169	172	174	166	167
168	173	166	161	186	209	191	165	168	134	205	187
204	153	159	160	167	164	162	164	157	186	187	188
181	176	174	173	176	187	156	158	158	145	148	157
160	157	159	186	157	146	151	155	147	144	145	152
132	149	153	151	ID	149	146	150	151	149	151	204
208	203	199	213								
gi295690758	162	186	184	188	168	167	170	169	170	168	168
169	170	168	165	190	208	186	180	180	148	199	186
197	154	150	163	157	152	155	162	160	185	190	187
167	172	169	172	180	183	160	150	150	137	139	154
161	156	154	182	140	135	132	147	149	142	139	144
136	140	142	145	149	ID	109	135	132	133	142	207
211	209	209	219								
gi167644962	148	185	185	187	161	160	165	161	160	161	164
163	157	161	151	185	218	185	169	169	139	205	191
209	161	152	154	145	138	149	156	157	186	179	187
167	167	166	166	168	184	141	149	151	121	121	152
148	156	153	183	136	135	134	140	148	134	136	144
128	134	142	143	146	109	ID	124	123	122	123	199
209	214	200	218								
gi86356860	151	189	188	195	157	160	162	160	163	160	160
157	167	158	153	187	213	189	175	164	144	206	193
205	167	153	148	145	152	148	158	156	185	193	194
170	164	163	166	173	182	150	151	151	126	122	159
150	155	157	181	125	153	156	144	143	141	140	149
115	156	147	147	150	135	124	ID	28	27	126	200
206	204	198	218								
gi209548458	149	187	187	192	155	153	158	156	162	154	156
152	165	153	149	185	209	187	172	160	144	204	189
205	164	152	148	147	150	148	157	155	184	190	193
167	164	164	165	170	180	149	146	147	126	123	157
148	151	153	179	132	156	159	143	135	137	136	148
119	153	147	142	151	132	123	28	ID	23	122	199
202	202	199	218								
gi116251120	151	185	184	191	149	153	155	155	159	153	153
152	163	151	151	188	209	186	172	161	141	204	188
203	158	150	147	146	152	149	154	152	184	193	192
162	164	163	165	168	183	146	150	151	123	118	156
150	155	150	182	127	151	156	137	137	134	135	141
114	149	139	141	149	133	122	27	23	ID	118	199
204	204	203	219								

gi110634564	153	192	191	187	157	160	164	157	161	161	163
161	162	160	149	205	214	191	165	170	141	209	195
206	157	145	151	141	146	147	149	152	185	193	201
168	165	165	166	168	177	145	151	153	126	118	152
152	154	148	176	146	155	157	143	149	138	129	140
129	146	137	139	151	142	123	126	122	118	ID	208
209	210	203	223								
gi53712542	212	211	211	204	206	207	208	204	204	206	206
204	209	206	211	184	207	190	209	209	206	214	199
224	204	204	210	205	209	207	208	207	198	200	201
218	216	213	217	206	203	206	210	210	205	201	207
212	211	200	202	203	205	207	216	210	197	199	198
203	210	196	206	204	207	199	200	199	199	208	ID
61	130	166	210								
gi29349915	217	210	214	207	208	210	212	206	208	209	210
205	211	209	216	192	206	197	213	209	213	214	203
225	208	204	213	205	206	210	210	211	200	202	201
220	217	213	217	213	207	211	216	216	205	207	212
217	218	202	206	207	210	213	220	215	206	205	203
208	215	200	209	208	211	209	206	202	204	209	61
ID	119	163	210								
gi150005764	214	212	214	205	207	211	210	211	214	211	209
210	214	210	212	184	217	191	210	213	206	214	205
224	204	205	211	207	205	212	206	212	203	207	205
218	216	215	219	216	208	210	212	212	209	205	206
214	212	208	207	209	202	206	214	208	204	206	207
209	217	203	213	203	209	214	204	202	204	210	130
119	ID	155	207								
gi302338706	209	200	202	208	203	204	205	200	204	206	206
204	201	208	206	171	208	160	203	203	204	217	191
217	200	205	203	203	202	203	208	207	205	191	195
212	210	213	209	203	198	199	209	209	196	200	203
209	208	210	197	203	208	210	213	205	210	205	205
197	214	205	206	199	209	200	198	199	203	203	166
163	155	ID	210								
gi19704905	217	213	211	218	215	214	216	209	204	213	216
211	209	216	220	214	186	216	215	224	212	195	206
200	215	221	218	224	218	222	225	220	207	205	210
220	223	218	218	217	210	217	220	219	219	219	227
216	220	221	210	211	220	222	222	222	223	226	220
217	226	218	227	213	219	218	218	218	219	223	210
210	207	210	ID								

Table S3 Sequence differences between the β-lactamases studied in this work. This table includes the differences among extant TEM-1 β-lactamase, the laboratory resurrections of Precambrian β-lactamases, and also, for each lactamase, the difference with the best hit sequence (sequences with highest similarity with the query) obtained from BLAST searches in the non-redundant protein database. Sequence differences are given as number of aminoacids (total length is 262 amino acids). Diffrences of the sequences of the β-lactamases studied with the consensus sequence from the 75 class A β-lactamases used in our phylogenetic analysis are also provided; these differences with the consensus, however, are given as the percentage of ancestral vs. TEM-1 differences that differ from the consensus prediction.

Sequence Difference	**TEM-1**	**ENCA**	**GPBCA**	**GNCA**	**PNCA**
TEM-1	-	55	111	122	124
ENCA	55	-	79	98	104
GPBCA	111	79	-	30	43
GNCA	122	98	30	-	17
PNCA	124	104	43	17	-
Best hit	-	42	79	79	84
Consensus sequence	-	64%	41%	42%	39%

Table S4 Testing the phenotypic robustness of the laboratory resurrection of Precambrian β-lactamases. Denaturation temperatures (from DSC experiments at 200 K/h), MIC values and Michaelis-Menten parameters for the GNCA β-lactamases corresponding the most probabilistic sequence and two alternative Monte Carlo sampling reconstructions (Figures S6-S10). For comparison, data for the extant TEM-1 lactamase are also included.

	TEM-1	GNCA-0	GNCA-1	GNCA-2
T_m (°C)	55.37	90.39	90.51	86.10
Log_{10}(MIC μg ml^{-1})				
AMP	3.602	1.20	1.50	1.20
CTX	-1.39	-0.49	-0.49	-0.49
CAZ	-0.69	0.17	0	-0.30
CRO	-1.69	-0.49	-0.49	-0.49
Kinetic Parameters				
BZ				
k_{cat} (s^{-1})	1760±100	9.5±1.1	12.5±2	36±2
K_m (μM)	66±8	32±12	17±3	158±56
k_{cat}/ K_M (s^{-1} μM^{-1})	26±4.7	0.3±0.1	0.7±0.2	0.2±0.1
CTX				
k_{cat} (s^{-1})	ND	116±7	140±25	126±35
K_M (μM)	ND	97±20	114±15	197±26
k_{cat}/ K_m (s^{-1} μM^{-1})	$2.6 \times 10^{-3} \pm 1 \times 10^{-3}$	1.2±0.3	1.2±0.4	$6.4 \times 10^{-1} \pm 0.3$
CAZ				
k_{cat} (s^{-1})	ND	1.5±0.1	1.7±0.2	0.9±0.2
K_m (μM)	ND	37±8	56±7	75±13
k_{cat}/ K_M (s^{-1} μM^{-1})	$6.5 \times 10^{-4} \pm 3 \times 10^{-4}$	$40 \times 10^{-3} \pm 11 \times 10^{-3}$	$29 \times 10^{-3} \pm 7 \times 10^{-3}$	$11 \times 10^{-3} \pm 5 \times 10^{-3}$

Table S5 Michaelis-Menten parameters for the hydrolysis of several antibiotics catalyzed by the β-lactamases studied in this work. In all cases, three independent determinations were performed and the values given are the corresponding averages and standard deviations.

Substrates	Variants	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/ K_m (s^{-1} μM^{-1})
BZ	TEM-1	1760±100	66±8	26±4.7
	ENCA	878±38	26±4	34±6.6
	GPBCA	85±20	12±2	7±2
	GNCA	9.5±1.1	32±12	0.3±0.1
	PNCA	18.9±0.5	35±1	0.5±0.03
CTX	TEM-1	ND	ND	$2.6 \times 10^{-3} \pm 1 \times 10^{-3}$
	ENCA	0.77 ±0.1	160±39	$4.8 \times 10^{-3} \pm 1 \times 10^{-3}$
	GPBCA	8.9±1	240±34	$37 \times 10^{-3} \pm 9 \times 10^{-3}$
	GNCA	116±7	97±20	1.2±0.3
	PNCA	73±16	72±15	1.0±0.4
CAZ	TEM-1	ND	ND	$6.5 \times 10^{-4} \pm 3 \times 10^{-4}$
	ENCA	0.4±0.02	80±20	$4.6 \times 10^{-3} \pm 1 \times 10^{-3}$
	GPBCA	0.9±0.03	180±50	$5 \times 10^{-3} \pm 1 \times 10^{-3}$
	GNCA	1.5±0.1	37±8	$40 \times 10^{-3} \pm 11 \times 10^{-3}$
	PNCA	0.3±0.05	10±1	$30 \times 10^{-3} \pm 8 \times 10^{-3}$

Table S6 Minimum inhibitory concentration (MIC) values determined using an E. coli strain tranformed with plasmid containing the ancestral β-lactamases. The MICs of each transformant, as well as that of the control strain *E. coli* ATCC 25922, were assayed in triplicate with serially diluted (2-fold) antibiotic stocks. In all cases, identical values were obtained and, therefore, the variability between different experiments is less than twofold

	TEM-1	ENCA	GPBCA	GNCA	PNCA
$Log_{10}(MIC/\mu g\ ml^{-1})$					
AMP	3.60	3.40	2.11	1.20	1.51
CTX	-1.39	-0.79	-0.79	-0.49	-0.19
CAZ	-0.69	0	-0.30	0.17	0.17
CRO	-1.69	-1.40	-0.80	-0.49	-0.19

Table S7 X-ray data collection and refinement statistics.

	GNCA	ENCA
Wavelength (Å)	1.54	1.54
Resolution range (Å)	18.66 - 2.05 (2.12 - 2.05)	48.13 - 2.4 (2.49 - 2.4)
Space group	I 41	P 64
Unit cell	94.833 94.833 95.34 90 90 90	111.144 111.144 93.722 90 90 120
Total reflections	196012 (27973)	546677 (2678)
Unique reflections	26446 (2628)	25827 (1246)
Multiplicity	7.4 (7.2)	21.17 (20.93)
Completeness (%)	100.00 (100.00)	99.68 (100.00)
Mean I/sigma(I)	13.8 (3.5)	15.38 (1.73)
Wilson B-factor	19.73	34.2
R-factor	0.1266 (0.1699)	0.1599 (0.2649)
R-free	0.1656 (0.2066)	0.1786 (0.2799)
Number of atoms	4336	8644
macromolecules	2054	4185
ligands	9	4
water	164	171
Protein residues	260	526
RMS(bonds)	0.010	0.013
RMS(angles)	1.36	1.30
Ramachandran favored (%)	98	98
Ramachandran outliers (%)	0	0
Clashscore	5.57	7.43
Average B-factor	24.80	29.80
macromolecules	23.60	29.30
solvent	39.90	36.90

Statistics for the highest-resolution shell are shown in parentheses.

References

1. Huelsenbeck, J.P. & Ronquist, F. Mr. Bayes: Bayesian inference of phylogeny. *Bioinformatics* **17,** 754-755 (2001).
2. Barlow, M., What antimicrobial resistance has taught us about horizontal gene transfer, *Methods Mol Biol.* **532**, 397-411 (2009).
3. Hanson-Smith V, Kolaczkowski B, Thornton JW. Robustness of ancestral sequence reconstruction to phylogenetic uncertainty. *Mol Biol Evol.* **27**,1988-99 (2010).
4. Yang, Z. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.* **13,** 555-556 (1997).
5. Bershtein S, Goldin K, Tawfik DS. Intense neutral drifts yield robust and evolvable consensus proteins. *J Mol Biol.* **379,**1029-44 (2008).
6. Risso, V.A., Primo, M.E. & Ermácora, M.R. Re-engineering a beta-lactamase using prototype peptides from a library of local structural motifs. *Protein Sci.* **2,** 440-449 (2009).
7. Wiegand, I., Hilpert, K. & Hanconck, R.E. Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. *Nat. Protoc.* **3,** 163-175 (2008).
8. Guzman-Casado, M., Parody-Morreale, A., Robic, S., Marqusee, S., Sanchez-Ruiz, J.M. Energetic evidence for formation of a pH-dependent hydrophobic cluster in the denatured state of Thermus thermophilus ribonuclease H. *J Mol Biol.* **329**,731-43 (2003).
9. Ibarra-Molero, B. and Sanchez-Ruiz, J.M. , *Emerging Techniques in Biophysics*, ed. J. L. Arrondo and A. Alonso, Elsevier, Amstadam, 27-48 (2006).

10. Ibarra-Molero, B. and Sanchez-Ruiz, J.M., *Statistical Differential Scanning Calorimetry: Probing Protein Folding-Folding Ensambles*,ed. V. Muñoz, RSC Publishing, Cambridge, 85-105 (2008).

11. Naganathan AN, Perez-Jimenez R, Sanchez-Ruiz JM, Muñoz V. Robustness of downhill folding: guidelines for the analysis of equilibrium folding experiments on small proteins. *Biochemistry.* **44**,7435-49 (2005).

12. Sanchez-Ruiz JM. Theoretical analysis of Lumry-Eyring models in differential scanning calorimetry. *Biophys J.* **61**, 921-35 (1992).

13. Otálora, F. Gavira, J.A., Ng, J.D. & Garcia-Ruiz, J.M. Counterdiffusion methods applied to protein crystallization. *Progr. Biophys. Mol. Biol.* **101,** 26-37 (2009).

14. Halle B. Flexibility and packing in proteins. *Proc Natl Acad Sci U S A*. **99**, 1274-9 (2002) .

15. DePristo MA, de Bakker PI, Blundell TL. Heterogeneity and inaccuracy in protein structures solved by X-ray crystallography. *Structure*. **12**,831-8 (2004)

16. Eyal E, Gerzon S, Potapov V, Edelman M, Sobolev V. The limit of accuracy of protein modeling: influence of crystal packing on protein structure. *J. Mol. Biol.* **351**, 431-42 (2005).

17. Jaenicke, R. Do ultrastable proteins from hyperthermophiles have high or low conformational rigidity?, *Proc Natl Acad Sci U S A.* **97**, 2962-4 (2000) .

18. Ambler RP. The structure of beta-lactamases. *Philos Trans R Soc Lond B Biol Sci.* **289**, 321-31 (1980).

19. Main, N.; Liu, A. D.; Ramer, S.; Aehle, W.; Meijer, D.; Metin, M.; Wong, S.; Gualfetti, P.; Main, N.; Liu, A. D.; Ramer, S.; Aehle, W.; Meijer, D.; Metin, M.; Wong, S.; Gualfetti, P.; Schellenberger, V. *Protein Eng.* 17, 787-93 (2004).

20. Arriaga, P. Menéndez, M., Villacorta, J.M. & Laynez, J. Differential scanning calorimetric study of the thermal unfolding of beta-lactamase I from *Bacillus cereus. Biochemistry* **31,** 6603-6607 (1992)

21. Matagne A, Misselyn-Bauduin AM, Joris B, Erpicum T, Granier B, Frère JM. The diversity of the catalytic properties of class A beta-lactamases. *Biochem J.* **265**,131-46 (1990).