Title	A zinc-binding site by negative selection induces metallodrug susceptibility in an essential chaperonin
Author(s)	Cun, S; Sun, H
Citation	Proceedings Of The National Academy Of Sciences Of The United States Of America, 2010, v. 107 n. 11, p. 4943-4948
Issued Date	2010
URL	http://hdl.handle.net/10722/168445
Rights	Creative Commons: Attribution 3.0 Hong Kong License

A zinc-binding site by negative selection induces metallodrug susceptibility in an essential chaperonin

Shujian Cun and Hongzhe Sun[1]

Department of Chemistry and Open Laboratory of Chemical Biology, University of Hong Kong, Pokfulam Road, Hong Kong, China

Edited* by Harry B. Gray, California Institute of Technology, Pasadena, CA, and approved January 27, 2010 (received for review December 4, 2009)

GroES is an indispensable chaperonin virtually found throughout all life forms. Consequently, mutations of this protein must be critically scrutinized by natural selection. Nevertheless, the homolog from a potentially virulent gastric pathogen, *Helicobacter pylori*, strikingly features a histidine/cysteine-rich C terminus that shares no significant homology with other family members. Additionally, three more (H45, C51, and C53) are uniquely present in its apical domain. The statistical analyses show that these residues may have originated from negative selection, presumably driven by either dependent or independent amino acid mutations. In the absence of the C-terminal metal-binding domain, the mutant protein still exhibits a substantial capacity for zinc binding in vivo. The biochemical properties of site-directed mutants indicate that H45, C51, and C53 make up an oxidation-sensitive zinc-binding site that may donate the bound metal to a zinc acceptor. Of interest, bismuth antiulcer drugs strongly bind at this site (K_d of approximately 7×10^{-26} M), replacing the bound zinc and consequently inducing the disruption of the quaternary structure. Because biological features by negative selection are usually inert to change during evolution, this study sheds light on a promising field whereby medicines can be designed or improved to specifically target the residues that uniquely evolved in pathogenic proteins so as to retard the emergence of drug resistance.

bioinorganic chemistry | bismuth susceptibility | GroES chaperonin | *Helicobacter pylori* | natural selection

New mutations of a protein are distinguishably chosen by natural selection. Positive selection amplifies a mutant that benefits the majority of the population. Therefore, a positively selected mutation presents a significantly higher occurrence frequency than others. In contrast, when a mutation results in unfitted physicochemical modification to the original, it will eventually impair protein functions or even become deleterious to the host cell. Thus, negative selection operates the damage control to shrink the mutant population until its extinction. Nevertheless, the mutation can be retained in a few individuals that do not lethally suffer, or even benefit from the exceptional properties evoked by the mutation in certain circumstances. Consequently, the proportion of this type of mutation is remarkably low within the populations. Unlike the directional selections from above, neutral selection, also known as the random genetic drift (1), has arguably been regarded as the major form of natural selection in general. Neutral selection usually works on slightly mutated proteins that exhibit few phenotypic characters and thus introduce neither advantage nor disadvantage. As a result, neutral selection contributes a plentiful source to amino acid polymorphism at a specific position, which therefore may represent selective pressure there. All three selective modes are coexisting during the evolution, and any one or none of them could be the decisive force in the evolution of a protein, unless explicitly given a historical period and a geographic environment. Relevant topics have been comprehensively reviewed recently (2).

Point mutations at the protein level can be considered as the minimal unit of phenotypes that are directly exposed to natural selection. Two distinct types of intrinsic force may appreciably affect the output: the genetic code per se and the adjacent amino acids, referred to as an independent substitution and dependent coevolution, respectively. If a point mutation solely relies on its codon, then the variance at this site is independent from other linked residues. On the other hand, because most residues are accompanied by two others at both amino and carboxyl ends, this possibly results in the dependence on the adjoining neighbors. It has been demonstrated that certain amino acids are particularly favored or disfavored in the formation of secondary structures (3). Formation of specific secondary structures requires cooperativity of a series of adjacent residues, and this is the prerequisite of functional tertiary and quaternary structures in most cases. Hence, the neighbors can either exempt or restrict the available options of amino acid substitutions. Selective pressure and mode may vary remarkably when mutations happen to different sites on different proteins. Up to this point, sophisticated models are continually being established to analyze site-directed mutations, but it remains difficult to ascertain which type of mechanisms actually worked on amino acid polymorphism of a given site on a given protein. Hence, the conclusion should be more reliable if the contrastive assumptions of independent substitution and dependent coevolution are both examined rather than either one of them being arbitrarily accepted during the analysis on amino acid substitution.

GroEL/S chaperone system is highly conserved both structurally and functionally in evolution, and has been comprehensively studied among all chaperonins to date (4). Natural selection on GroEL/S must be very cautious so that the system adopts as few ad hoc changes as possible, because functional development of chaperonins needs to compromise on specialization and generalization during the evolution (5). The heptameric GroES serves as a lid to facilitate the formation of a hydrophobic cavity in the tetradecameric GroEL, where nonnative proteins are encapsulated and refolded rapidly (6). The necessity of GroES chaperonin implicates that dysfunction of the protein may be potentially lethal for its host organism.

Being a pathogen connected to peptic ulcers, gastritis, duodenitis, and gastric cancers (7), *Helicobacter pylori* features an unusual version of GroES (i.e., *Hp*GroES), also referred to as the heat-shock protein A (HspA), that owns a unique C terminus rich in histidines and cysteines. The protein has been reported to interact with several metal ions, including ubiquitously existing Zn^{2+} (8) and rarely present Bi^{3+} (9). Because *H. pylori* infection triggers the immune response to rapidly release abundant zinc as well as oxidative radicals (10), the zinc-binding abilities may let the chaperonin contribute to the bacterial pathogenesis in oxidative situations, particularly considering both intra- and extracellular locations of *Hp*GroES (11). Conversely, bismuth plays no role in all known organisms. With an exceptionally strong affinity

Author contributions: S.C. and H.S. designed research; S.C. performed research; S.C. and H.S. analyzed data; and S.C. and H.S. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

[1]To whom correspondence should be addressed. E-mail: hsun@hku.hk.

This article contains supporting information online at www.pnas.org/cgi/content/full/0913970107/DCSupplemental.

BIOCHEMISTRY

to the thiolate sulfur of selective proteins, the metal can be empowered to block cysteine-related functions once the residue is accessible. Due to this biochemical character, bismuth drugs have shown antimicrobial and antiviral effects in a number of cases (12, 13).

In this report, we propose three metal-binding residues of *Hp*GroES (i.e., H45, C51, and C53) as the result of negative selection among the GroES family, and we experimentally demonstrate their importance in Zn^{2+} binding and Bi^{3+} susceptibility by comparative study. The potential application correlated with our findings is also discussed.

Results

Negative Selection of H45, C51, and C53 Assuming Either Independent or Dependent Substitution. Based on the sequence alignment of GroES homologs selected from a wide range of model organisms (Fig. S1), the position-specific polymorphic proportions are estimated with *Hp*GroES for reference. To evaluate the statistical frequency of amino acid substitution for H45, C51, and C53 of *Hp*GroES, a total of 474 sequences of GroES proteins was used to build up aligned blocks, from which a GroES-specific scoring matrix (namely, GSM474) was constructed (Fig. S2). Each score individually indicates the observed frequency for a certain substituent either higher (as positive), equal (as zero), or lower (as negative) than the expected by random. With respect to the identified substitution modes (Fig. 1*A*), the polymorphic proportions were multiplied by the corresponding substitution scores to give an overall evaluation for H45, C51, and C53 as −5.879, −4.059, and −1.354, respectively. According to the mathematical essence of GSM474, the probability of the concurrent appearance of H45, C51, and C53 on the given positions is calculated from $10^{[(-5.879)+(-4.059)+(-1.354)]/10}$, i.e., 13.5-fold less likely than random. Consequently, the occurrence of H45, C51, and C53 is inferred to be negatively selected among widespread species, if presumed to have originated independently from other residues.

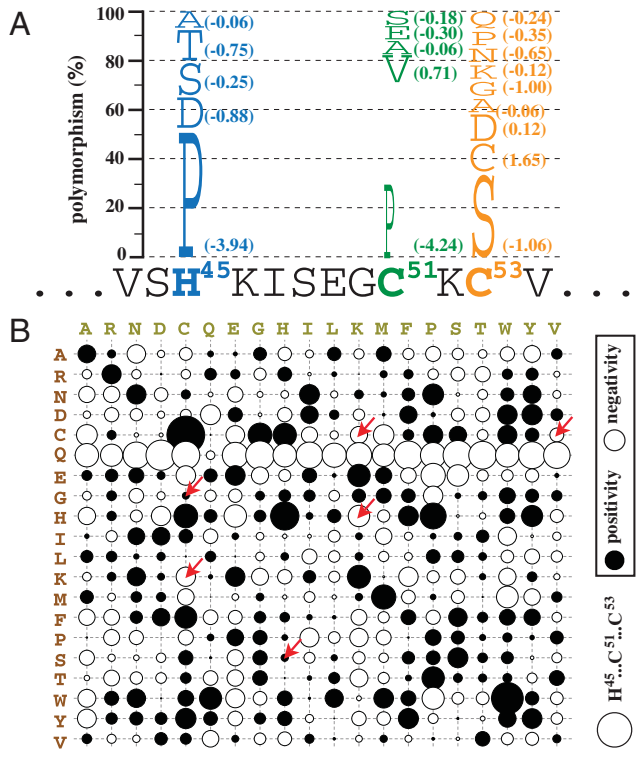

Fig. 1. (*A*) Position-specific polymorphic proportions related to the three metal-binding residues. A fraction of the GroES from *H. pylori* (*Hp*GroES) is shown as the reference of residue position. The substitution modes are identified from the alignment of multiple GroES homologs, and the corresponding proportions of each substitute are represented by the height of the letters. The numbers in brackets are the substitution scores for relevant substitution modes, individually given by GSM474 score times polymorphic proportion. (*B*) Data plot of neighbor preference based on the UniProtKB database. Filled circles are used to represent positive values and open circles indicate negative values. The circular area equals the absolute value of corresponding data, and the thresholds are given by the statistical analysis. The six amino acid pairs discussed in this work are marked by arrows, of which the total score is shown below the thresholds. The *y* axis symbols are appointed as the leader residue of an amino acid pair, and the *x* axis as the follower.

On the contrastive assumption that amino acid substitutions may have been codominated by adjacent residues, the neighbor preference can thus be applied to examine the popularity of a certain residue in between its amino- and carboxyl- neighbors. The preferential data were generated using 392,750 protein sequences (Table S1). The dataset was found to fit a transformed normal distribution (Fig. S3). Using 95% as the confidence interval (CI), the critical values for positivity and negativity were then identified as 0.1680 (the upper 95% CI) and −0.1969 (the lower 95% CI), respectively. Each preferential datum of an amino acid pair "*i-j*" indicates that amino acid *j* follows amino acid *i* favorably if the *i-j* value is greater than 0.1680, or disfavorably if lower than −0.1969. The constructed data table was then employed to evaluate the neighbor preference of the six amino acid doublets relevant to the three metal-binding residues, i.e., S-H, H-K, G-C, C-K, K-C, and C-V. In accordance with the neighbor preference, the preferential data are illustrated in Fig. 1*B*. Altogether, the sum of the six scores (−0.5136) is significantly lower than the critical value of negativity (−0.1969), indicating a fairly lower frequency of the observed concurrence than that predicted by random. This implicates that the neighbors of the three metal-binding residues show an overall disfavor toward their appearance on the observed positions, suggesting a negative selection in the context of dependent coevolution.

Determination of Zn^{2+} Binding in Vivo. Apart from the histidine/cysteine-rich C terminus, three more metal-binding residues (as H45, C51, and C53) are located in the apical domain according to the structural superposition (Fig. S4). Furthermore, the three metal-binding residues are completely conserved among *Hp*GroES proteins from various *H. pylori* strains (8). Considering that the C-*X*-C motif has often been observed in zinc-binding proteins (12), the metal may bind at this site in a cellular environment as well. To examine the possibility, various wild types and mutants of GroES were expressed and isolated from *Escherichia coli* cell extracts via a denaturing cycle by heat treatment (Fig. S5), and small molecules were removed from the solutions after a desalting step. The metal content of the prepared GroES proteins was then determined by inductively coupled plasma mass spectrometry (ICP-MS). The ratio of Zn^{2+} versus *Hp*GroESΔMBD (where MBD signifies metal-binding domain) was 0.71 ± 0.18 per monomer. The binding ratios of *H45A*, *C51A*, and *C53A* decreased to 0.3–0.5 Zn^{2+} per monomer. In contrast, the double-cysteine mutant (*C51A*/*C53A*) and *E. coli* GroES (*Ec*GroES) gave only 0.12 ± 0.05 and 0.04 ± 0.02 Zn^{2+} per monomer, respectively, indicative of being incapable of zinc binding (Fig. 2). The binding of other transition metals, including manganese, iron, cobalt, nickel, and copper, was also investigated for comparison, and none was detected under the identical conditions.

Oxidative Sensitivity of Zn^{2+} Transfer. To investigate the properties of the Zn^{2+}-protein interactions, 4-(2-pyridylazo)resorcinol reagent (PAR) was used as a colorimetric indicator in the biochemical assays (Fig. S6). To mimic an oxidative stress, Zn^{2+}-supplemented GroES proteins were treated with hydrogen peroxide (H_2O_2) in a concentration gradient and then desalted into a buffer free of metals and oxidative radicals. The bound Zn^{2+} of *Hp*GroESΔMBD was proportionally released from

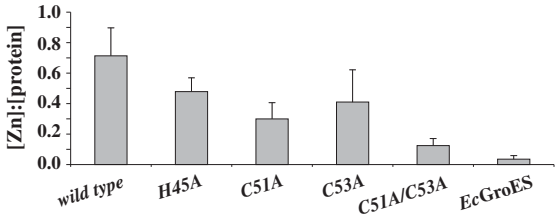

Fig. 2. Determination of in vivo zinc binding for GroES variants. The wild-type and point mutants of *Hp*GroESΔMBD as well as the wild-type *Ec*GroES were isolated from *E. coli* cells grown in media consisting of different nutrients, and then desalted into zinc-free buffer containing 50 mM Tris-HCl (pH 7.5) and 100 mM NaCl. The protein concentration was estimated by Bradford assay, and the zinc content was quantitated by ICP-MS with reference to the standard solution of zinc. Each column represents the mean ± the standard deviation in triplicate.

the mutant, whereas *Ec*GroES-MBD lost nearly none with the same $[H_2O_2]$:[protein] ratios (Fig. 3*A*). To monitor the potential transfer of Zn^{2+} between proteins, alkaline phosphatase (AP) from *E. coli*, a common model in zinc transfer (14), was chosen as the zinc acceptor, of which the roles of Zn^{2+} in the catalytic sites have been well demonstrated (15). The hydrolysis of *p*-nitrophenyl phosphate (*p*-NPP) is catalyzed by holo-AP, and the produced *p*-nitrophenolate leads to an absorption around 400 nm, which can therefore be used to monitor the enzymatic activity, i.e., the extent of metal transfer from GroES proteins to the apoenzyme. The AP activity was represented by the half-life ($t_{1/2}$) for hydrolysis of *p*-NPP. As shown in Fig. 3*B*, the hydrolysis of half of the substrate needed 2 min for the holo-AP, but more than 1 h for the apoform. The Zn^{2+}-bound *Hp*GroESΔMBD restored the activity of the enzyme with a half-life of 8 min. In contrast, the *Hp*GroESΔMBD samples upon addition of 0.5 or 2 molar equivalents of H_2O_2 showed distinguishably impaired capacity of reactivating the apo-AP, with $t_{1/2}$ of 26 min or over 1 h, possibly due to the partial loss of zinc-binding capacity by the H_2O_2-induced metal release. In comparison, the *Ec*GroES-MBD samples with or without H_2O_2 treatment exhibited similar restoration of AP activity ($t_{1/2}$ = 13 and 9 min, respectively), despite being slightly slower than that by *Hp*GroESΔMBD. The observation possibly implies a higher Zn^{2+} affinity of the MBD, as described below, which raises up the thermodynamic threshold of metal delivery from the chaperonin to the apo-AP and thus may delay zinc reconstitution of the enzyme. As expected, the reactivation by the mutant *C51A/C53A* as well as wild-type *Ec*GroES, both of which failed to bind zinc in vivo (Fig. 2), resembled the activity of the apoenzyme; non-zinc-binding proteins cannot reconstitute the enzyme. Clearly, the observed recovery of AP catalytic activity was due to the zinc transfer from the chaperonin to the enzyme, depending on oxidative atmosphere for the C-*X*-C motif but not for the MBD.

Different Roles of H45, C51, and C53 in Metal Binding in Vitro. The wild-type *Hp*GroES was found to bind two Zn^{2+} per monomer with a dissociation constant (K_d) of $1.23 \pm 0.51 \times 10^{-10}$ M by PAR assay (Table 1). To examine the role of each residue individually and meanwhile minimize the interference from the C-terminal MBD, site-directed alanine scanning has been applied to *Hp*GroESΔMBD, an *Hp*GroES mutant with the MBD deleted. The mutants with single-histidine, single-cysteine, or double-cysteine substituted were thus produced, respectively, referred to as *H45A*, *C51A*, *C53A*, or *C51A/C53A* in this report. The competitive assays with PAR showed that zinc coordinated moderately ($K_d = 2.00 \times 10^{-9}$ M) presumably by all three residues in a 1:1 molar ratio. However, when the histidine was replaced by an alanine, Zn^{2+}-binding affinity was augmented over 10-fold ($K_d = 1.85 \times 10^{-10}$ M) with the same binding ratio. With respect to the results from *C51A*, *C53A*, and *C51A/C53A*, both cysteine residues should be necessary in Zn^{2+} binding, because Zn^{2+} binding was undetectable for the mutants even in excess molar equivalents.

Binding of Bi^{3+} to *Hp*GroES variants was also determined by equilibrium dialysis together with ICP-MS. The collected data were fitted to a one-site binding (hyperbola) model. Despite an extremely high affinity, Bi^{3+} demonstrated similar binding patterns to Zn^{2+}. All three residues were involved in Bi^{3+} coordination, as the mutation of H45A seemingly increased the metal binding ($K_d = 7.67 \times 10^{-27}$ M) by approximately 10-fold higher than to the wild type ($K_d = 7.3 \times 10^{-26}$ M), leaving the binding ratio unchanged. The precipitation of *C51A* and *C53A* in the presence of bismuth hindered our effort to quantitatively measure the physicochemical diversity of the two cysteine residues in Bi^{3+} coordination.

Distinguishable Zn^{2+}/Bi^{3+}-Binding Affinities Between the MBD and the Apical Domain. To emulate metal-binding characteristics that solely rely on the MBD, *E. coli* GroES, which shows no specific affinity to transition metals with respect to current knowledge, was fused with the histidine/cysteine-rich MBD at the C terminus, and Zn^{2+}/Bi^{3+}-binding properties of the mutant (i.e., *Ec*GroES-MBD) were examined similarly. The MBD showed distinct

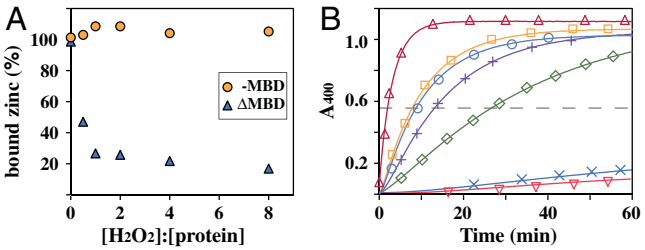

Fig. 3. Zinc transfer conducted by the two domains of the chaperonin, with or without oxidative sensitivity. (*A*) The bound Zn^{2+} of the H-X_5-C-*X*-C motif and the MBD was separately determined with increasing molar ratios of H_2O_2 versus the protein. (*B*) The hydrolysis of *p*-NPP by active AP was monitored by the changes of the absorption at 400 nm (A_{400}) with or without the presence of several GroES variants, which were incubated with zinc and H_2O_2 (or not) before desalting. The examined samples are holo-AP (up-pointing triangle), apo-AP with *Hp*GroESΔMBD (square), apo-AP with 2 molar equivalents H_2O_2-treated *Ec*GroES-MBD (circle), apo-AP with *Ec*GroES-MBD (plus), apo-AP with 0.5 molar equivalents H_2O_2-treated *Hp*GroESΔMBD (diamond), apo-AP with *C51A/C53A* (cross), and apo-AP (down-pointing triangle). The half of maximal A_{400} is shown by a dashed line. All tests were performed in 50 mM Tris-HCl (pH 7.5) and 100 mM NaCl.

Table 1. Maximal binding ratios (B_{max}) and dissociation constants (K_d) of the studied proteins

(A) Zn^{2+}-binding parameters

Protein	B_{max}/monomer	K_d, M
*Hp*GroES	2.0 ± 0.3	$1.23 \pm 0.51 \times 10^{-10}$
*Hp*GroESΔMBD	1.0 ± 0.1	$2.00 \pm 0.93 \times 10^{-9}$
*Hp*GroESΔMBD(H45A)	0.9 ± 0.1	$1.85 \pm 0.38 \times 10^{-10}$
*Ec*GroES-MBD	1.3 ± 0.2	$1.08 \pm 0.51 \times 10^{-10}$

(B) Bi^{3+}-binding parameters

Protein	B_{max}/monomer	K_d, M
*Hp*GroES *	2.0 ± 0.1	$5.90 \pm 0.03 \times 10^{-25}$
*Hp*GroESΔMBD	1.0 ± 0.1	$7.30 \pm 1.48 \times 10^{-26}$
*Hp*GroESΔMBD(H45A)	1.0 ± 0.1	$7.67 \pm 2.27 \times 10^{-27}$
*Ec*GroES-MBD	1.0 ± 0.2	$3.18 \pm 1.00 \times 10^{-24}$

Zn^{2+}-binding affinities were determined by competitive assay with PAR reagent, and Bi^{3+} binding is investigated by ICP-MS using BiNTA. The data were shown as the median ± the error. All samples were determined in 50 mM Tris-HCl (pH 7.5) and 100 mM NaCl, at ambient temperature.

*Data derived from our published work (9).

BIOCHEMISTRY

preference to the two metals. As shown in Table 1, the zinc coordination by C51 and C53 can be moderately relaxed by H45, as the mutation of H45A increased the binding affinity toward the metal around 10-fold, whose biological significance remains unclear at present. And *Ec*GroES-MBD exhibited a 10-fold higher affinity (K_d of 1.08×10^{-10} M) to Zn^{2+} than that of *Hp*GroESΔMBD. Correlated with the discrete oxidative sensitivities, there seemingly exist two metal-binding sites in the apical domain and at the C terminus, separately. For the MBD, such a domain bound Bi^{3+} approximately 40-fold more weakly than *Hp*GroESΔMBD (3.18×10^{-24} M versus 7.30×10^{-26} M), in spite of extremely strong affinities in both domains. Reasonably, Bi^{3+} binding should inevitably prevent a reversible binding of Zn^{2+} under cellular conditions, and most probably disrupt any potential zinc transfer that is conducted by either one or both.

Competition Between Bi^{3+} and Zn^{2+} Binding. To investigate whether Bi^{3+} directly binds at the same site as Zn^{2+}, a competition experiment between the two metals was carried out by monitoring the UV absorption at 360 nm. *Hp*GroESΔMBD was preequilibrated with 2 molar equivalents of Bi^{3+} [as $Bi(NO_3)_3$]. As the result, the absorption remained almost unchanged ($A_{360} \sim 0.455$) even upon addition of up to 10-fold of Zn^{2+} (as $ZnSO_4$) into the protein (Fig. 4*A*). In contrast, for the *Hp*GroESΔMBD solution that was presaturated with Zn^{2+}, a new absorption at 360 nm appeared, and the absorbance increased up to 0.337 upon addition of increasing amounts of Bi^{3+}. Similar features were observed for *Ec*GroES-MBD under identical conditions (Fig. 4*B*). The absorption of 360 nm is characteristic for the formation of Bi-S bond, indicative of the coordination of Bi^{3+} to cysteines (16).

Different Effects of Zn^{2+}/Bi^{3+} Binding on the Quaternary Structure. The oligomeric structures are essential for GroEL/S-assisted protein folding (17), and therefore it leaves no doubt that any dramatic change of the conserved structures can be devastating once unpredictable consequences happen either instantly or latently. To study the possibility of metal-induced structural changes, oligomeric states of *Hp*GroESΔMBD variants with or without metals were revealed by gel filtration chromatography. The protein elution was monitored by the absorption at 215 nm (A_{215}) in this work rather than the more usual A_{280} to improve the sensitivity of the protein samples abundant in arginine, cysteine, histidine, and phenylalanine but scarce in tryptophan and tyrosine (18). As shown in Fig. 5*A*, all apo- and zinc-supplemented proteins were eluted at around 12.8 mL in majority, corresponding to a molecular mass of 70 kDa, in agreement with a heptamer (70.7 kDa for *Hp*GroESΔMBD). Nevertheless, the *Hp*GroESΔMBD variants exhibited various oligomeric behaviors when loaded with 1 molar equivalent of Bi [as $Bi(NO_3)_3$], excluding *C51A/C53A* that retained the elution volume independently

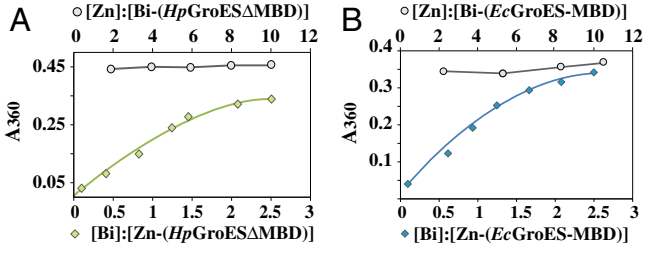

Fig. 4. Competition between Zn^{2+} and Bi^{3+} to the chaperone domain of *Hp*GroES (represented by *Hp*GroESAΔMBD) (*A*) and the histidine/cysteine-rich MBD (represented by *Ec*GroES-MBD) (*B*). The apoproteins were presaturated by $ZnSO_4$ or $Bi(NO_3)_3$ and then titrated by $Bi(NO_3)_3$ or $ZnSO_4$, accordingly. The absorption at 360 nm (A_{360}) was monitored for the formation of Bi-S bond. The experiments were performed at ambient temperature and in a buffer containing 50 mM Tris-HCl (pH 7.5) and 100 mM NaCl.

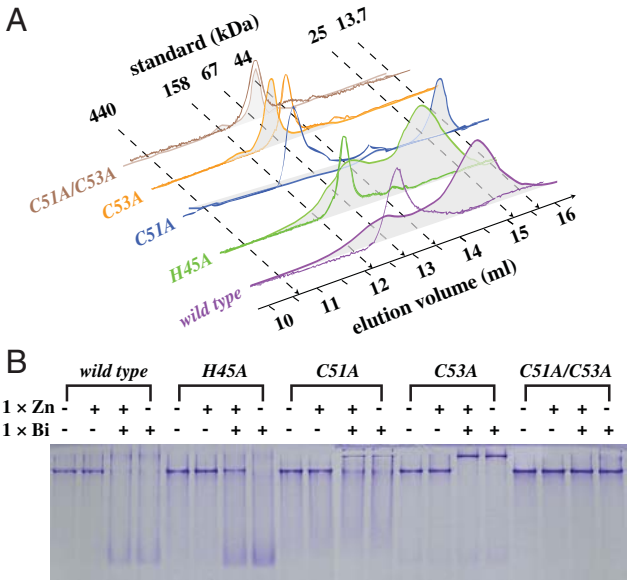

Fig. 5. Oligomeric states of apo- and metal-bound *Hp*GroESΔMBD variants. (*A*) One molar equivalent of $ZnSO_4$ (solid line) or $Bi(NO_3)_3$ (shaded curve) was loaded into proteins, and the samples were run on a Superose 12 column. The molecular masses were estimated by standard globular proteins, and A_{215} was monitored to indicate the protein elution. (*B*) Apo-, Zn-, and Bi-bound proteins were visualized by Coomassie Blue staining in 15% polyacrylamide gel under nondenaturing conditions. Both were carried out at room temperature.

from the presence of bismuth. The elution profiles of bismuth-loaded *Hp*GroESΔMBD and *H45A* proteins differed obviously from the apo- (or zinc-loaded) form, predominantly around 15 mL, corresponding to a molecular mass of 25 kDa. The elution of bismuth-loaded *C51A* mutant was delayed until 15.6 mL, implicating a molecular mass around 12 kDa, whereas the single mutant *C53A* demonstrated a slightly enlarged molecular mass than the apoform, deduced from the 13.5-mL elution volume. To further investigate the possible structural changes of the proteins, *Hp*GroESΔMBD variants in 50 ~ 100 μM were prepared in apo- or 1 molar equivalent of metal-supplemented forms [as $ZnSO_4$, $ZnSO_4$ in the presence of $Bi(NO_3)_3$, and $Bi(NO_3)_3$], and subjected to polyacrylamide gel electrophoresis under nondenaturing conditions. As exhibited in Fig. 5*B*, the molecular masses of Zn^{2+}-loaded samples resembled their apoforms. In contrast, the Bi^{3+}-bound proteins migrated from a larger size to a distinctly smaller one for the wild-type, *H45A*, and *C51A*, or enlarged the molecular mass (for *C53A*). As expected, both cysteine-mutated *C51A/C53A* showed no apparent size migration regardless of whether Zn^{2+} was supplemented. The overall results are clearly in agreement with those from gel filtration chromatography. Moreover, it is also noticeable that aged *Hp*GroES proteins may demonstrate a complicated equilibrium of oligomeric states in vitro under unestablished conditions, presumably due to oxidation of the cysteine residues and/or transient folding states (19, 20). We found that the appropriate supplement of tris(2-carboxyethyl)phosphine and transition metals (e.g., Zn^{2+}) greatly facilitated restoration and stabilization of the heptameric structure.

Bi^{3+} Susceptibility in Vivo. To explore potential cytotoxicity of protein instability caused by Bi^{3+} binding that was determined in vitro, *E. coli* BL21 cells harboring expression vectors of individual *Hp*GroESΔMBD variants were induced by isopropyl β-D-1-thiogalactopyranoside (IPTG) in LB medium to express the mutants in a similar level to each other (ranging from 28 to 38% of total proteins; see Fig. S7) and then grown in M9 minimal medium in the presence of 50 μM bismuth(III) citrate. As

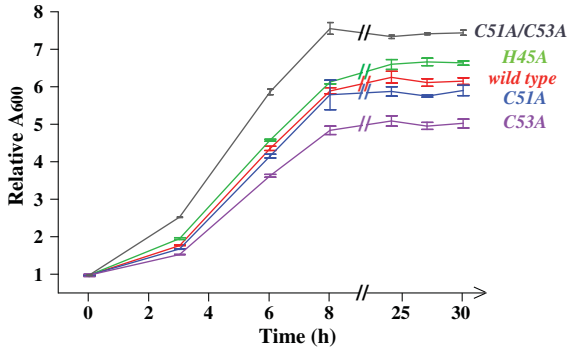

Fig. 6. Growth rates of *E. coli* BL21 cells hosting *Hp*GroESΔMBD variants in M9 minimal medium supplemented with 50 μM bismuth(III) citrate. The bacteria were induced by IPTG in nutrient-rich medium in advance, and appropriately diluted into M9 medium (with $A_{600} = 0.1–0.2$) as the starter culture. At each selected time point, A_{600} of an aliquot of bacterial culture was determined to indicate the cell number, and the A_{600} values were normalized to that at the beginning. The data were collected in triplicate and represented as the mean ± the standard deviation.

shown in Fig. 6, in the first 3 h after aerobic incubation at 37 °C, *E. coli* expressing the mutant *C51A/C53A* grew markedly faster, reaching 2.6-fold higher cell density than the starter culture, whereas others only proliferated to 1.7-fold. Keeping the fastest growth among all samples, *E. coli* expressing *C51A/C53A* finally amplified the cell density over 7-fold higher than that at the beginning. In contrast, the growth rate of the cells having *C53A* was most retarded, having merely 64% cell density of *C51A/C53A*. *E. coli* cells hosting the wild-type, *H45A*, or *C51A* proliferated almost indistinguishably from each other, approaching final cell densities approximately 70% as that of *E. coli* hosting *C51A/C53A*. Subsequently, the profiles of the bacterial growths remained stable even after 30-h overtime incubation.

Discussion

As the position-specific conservation of certain residues usually hints at the functional importance of a protein, a thorough knowledge of a conserved chaperone system may enlighten a broad range of understanding about protein evolution and structure-function relationship. It has long been suggested that random mutations may accumulate in bacterial genome if under low pressure of natural selection (21), which implies a possible correlation between selective force and mutational rate for a certain residue of a protein. Nevertheless, other genetic procedures can cause amino acid polymorphism as well, such as insertion, deletion, or recombination. These mutagenic factors are bypassed in the present work, because it could be technically problematic to model such processes (22), even if it is empirically claimed that some instances of amino acid polymorphism could have originated by these means (23). In general, directional selection, neutral drift and mutational bias have been proposed to be the driving forces for the evolution at a molecular level (2), leaving aside the controversy about which one plays the major role. Because enormous supportive instances could be provided by each side, it seems likely that everyone might be the decisive power, depending on what constraint is specified. Therefore, it should result in more comprehensive knowledge by comparing the analytic outputs derived from contrastive assumptions to find out if a common agreement can be reached. Hence, as demonstrated in the present work (Fig. 1), a GroES-specific scoring matrix named GSM474 (Fig. S2) and the scoring table of neighbor preference (Table S1) were constructed to investigate the substitutions of the three metal-binding amino acids of *H. pylori* GroES.

Our combined data show that H45, C51, and C53 form a zinc-binding site (Fig. 2), and such a binding is oxidation-sensitive. Importantly, we demonstrated Zn^{2+} trafficking between *Hp*GroES and zinc enzymes using AP as an example (Fig. 3). On the contrary, the Zn^{2+} capacity of the MBD is seemingly unaffected under the oxidative stress, perhaps due to the richness in metal-binding residues nearby (e.g., histidines) that are insensitive toward oxidation. This is partially supported by the enhanced reconstitution of apo-AP by the H_2O_2-treated *Ec*GroES-MBD when compared to the unoxidized sample, which can probably be attributed to the weakened affinity of Zn^{2+}, and consequently the metal transfer became easier. Besides the oxidation-sensitive metal binding, the C-*X*-C motif may also function as a mimic of protein disulfide isomerase (24). In agreement with the observed oxidation-sensitive AP reactivation, oxidative functions of the investigated motif may work separately or in parallel to the identified Zn^{2+} binding under physiological conditions. Unsurprisingly, even more functions can be related to the H-X_5-C-*X*-C motif as the research on this field goes further. On the other hand, from a practical point of view, further understanding of parasitic origin and consequence of natural selection can greatly strengthen our ability to cope with intractable diseases more strategically in future.

In addition to structural disruption that causes cytotoxicity, Bi^{3+} binding may functionally interfere with *Hp*GroES by other means. As reported, the cysteine residues at MBD have been found to coordinate Ni^{2+} (9). The appearance of UV-visible absorption at 360 nm of *Ec*GroES-MBD indicates that the domain binds Bi^{3+} via cysteine residues (Fig. S8). Hence, the in vivo Ni^{2+} interaction may be interrupted in the presence of Bi^{3+}, and this may disable functions mediated by the domain. Similarly, the functional interruption might occur in the H-X_5-C-*X*-C motif in the apical domain as well. It would be of general interest for future work to investigate whether Bi^{3+}-induced structural disruption represents a universal character of the GroES fold family, a superfamily sharing GroES-like modules (25). For instance, other than the GroES chaperonin from *H. pylori*, another zinc-binding protein, alcohol dehydrogenase (ADH) from *Saccharomyces cerevisiae*, was also found to shift from its native tetramer to a dimer upon bismuth binding (26), and it is noticeable that ADH can be a potential target against pathogens (27). Nevertheless, we cannot rule out any variable reaction of *Hp*GroES in response to the presence of bismuth in the cytoplasmic environment. Also of note, although a typical Zn^{2+}-binding motif (i.e., C-*X*-C) together with a histidine was identified, no attempt has been made yet to reveal the fourth donor, which completes the coordination sphere. For example, it might be glutamate, aspartate, or a water molecule, depending on the surroundings.

It holds obvious appeal to engineer simple-to-produce as well as difficult-to-resist drugs for virulent infections. Inspired by the model-based drug development (28), we are keen to incorporate an evolution-driving module as a supplement to the current models, which may rapidly locate unique and important residues from numerous parasitic proteins as potential drug targets. Although the molecular mechanism of bismuth inhibition on microbial pathogens is not clear yet, certain proteins and enzymes have been believed to be potential targets of the metallodrug (12, 13). Unlike H45, C51, and C53, which are completely conserved within *H. pylori* strains, the C terminus of *Hp*GroES diverged so frequently that they can be used as a signature to characterize geographic isolates of *H. pylori* (29). Such a contrast strongly suggests an uneven selective pressure on the two domains, and thus disproves the MBD as an appropriate target for drugs. Besides *Hp*GroES discussed in this work, other *H. pylori* proteins may also be selectively targeted by bismuth drugs. For instance, another histidine/cysteine-rich protein, Hpn, also interacts with several metals including Zn^{2+} and Bi^{3+} (30), even though the physiological significance of such interactions have not yet clarified. Moreover, a few urease homologs have been demonstrated to be inhibited by bismuth binding to a conserved cysteine (31). Encouragingly, such a property

seems broadly practicable in addition to *H. pylori* treatment. For example, the arterivirus replicases possess 13 conserved histidine and cysteine residues to form a putative MBD that distinguishes the viral replicases from human homologs (32). This protein was later demonstrated to be an effective drug target to inhibit the severe acute respiratory syndrome (SARS) coronavirus, as it was shown that Bi^{3+} binds tightly to the MBD of the protein (33).

The stability of a protein also exhibits an important impact on the "speed" of molecular evolution (34). As far as structural and functional modification is concerned, either inclusion or exclusion of a metal-binding site is a serious choice for a protein, particularly for a versatile chaperonin. Having the second lowest frequency in proteins, cysteine residues were negatively selected in most cases, perhaps due to their aggressive reactivity that may interfere with the metabolism of essential metals if randomly introduced. But, noticeably, the residue plays irreplaceable roles in protein functionality, usually in forms of C-X-C and C-X_2-C motifs (12), and thus is often inert to change in evolution. To current knowledge, resistance against bismuth does not occur yet in *H. pylori* (35), although the metal has been used for centuries in clinical practice (12, 13). More excitingly, for those patients who failed to eradicate the bacterium after treatment with antibiotics, bismuth compounds plus proton-pump inhibitors exhibit satisfactory results with low toxicity to the human body (36). For this reason, the application of bismuth in medicinal practice shall be further explored. Inspired by the correlation of these evolutionary and physicochemical attributes, we are trying to find out whether the vulnerability to bismuth is common among pathogenic microorganisms that rely on cysteine-associated functions, and then figure out an effective way to harness these resourceful bugs. As the present study instances the "imperfection" in microbial evolution, it manifests that, when evolved to utilize transition metals by natural selection, some bugs may "accidentally" sensitize their proteomes to a few heavy metals rarely present in nature, as the price of choice.

Materials and Methods

GroES records were retrieved from the UniProtKB database using a characteristic signature of the protein family, and the sequences were aligned to build up the scoring matrix. The neighbor preference was calculated from total protein sequences in the database. Alanine scanning on the metal-binding residues was done by site-directed mutagenesis of the wild-type gene. The mutant proteins were overexpressed, purified, and restored to apoforms prior to metal-binding determination. Bismuth susceptibility was tested under minimal-nutrient conditions, using A_{600} of bacterial cultures as the indicator of cell density. Additional details are in *SI Text*.

ACKNOWLEDGMENTS. This work was supported by Research Grants Council of Hong Kong (HKU1/07C, HKU2/06C, HKU7042/07P, HKU7038/08P, and HKU7049/09P) and the University Grants Committee of Hong Kong Special Administration Region of China under the scheme of the Area of Excellence. We thank C. N. Tsang for constructing the *Ec*GroES-MBD vector, X. Kong for his helpful comments on statistical analysis, and University of Hong Kong for their support.

1. Kimura M (1968) Evolutionary rate at the molecular level. *Nature* 217:624–626.
2. Nei M (2005) Selectionism and neutralism in molecular evolution. *Mol Biol Evol* 22:2318–2342.
3. Minor DL, Kim PS (1996) Context-dependent secondary structure formation of a designed protein sequence. *Nature* 380:730–734.
4. Horwich AL, Farr GW, Fenton WA (2006) GroEL-GroES-mediated protein folding. *Chem Rev* 106:1917–1930.
5. Wang JD, Herman C, Tipton KA, Gross CA, Weissman JS (2002) Directed evolution of substrate-optimized GroEL/S chaperonins. *Cell* 111:1027–1039.
6. Tang Y-C, et al. (2006) Structural features of the GroEL-GroES nano-cage required for rapid folding of encapsulated protein. *Cell* 125:903–914.
7. Covacci A, Telford JL, Del Giudice G, Parsonnet J, Rappuoli R (1999) *Helicobacter pylorivirulence* and genetic geography. *Science* 284:1328–1333.
8. Kansau I, Guillain F, Thiberge JM, Labigne A (1996) Nickel binding and immunological properties of the C-terminal domain of the*Helicobacter pylori* GroES homologue (HspA). *Mol Microbiol* 22:1013–1023.
9. Cun S, Li H, Ge R, Lin MC, Sun H (2008) A histidine-rich and cysteine-rich metal-binding domain at the C terminus of heat shock protein A from *Helicobacter pylori*: Implication for nickel homeostasis and bismuth susceptibility. *J Biol Chem* 283:15142–15151.
10. Dovhanj J, et al. (2009) *Helicobacter pylori*, zinc and iron in oxidative stress-induced injury of gastric mucosa. *Mini-Rev Med Chem* 9:26–30.
11. Vanet A, Labigne A (1998) Evidence for specific secretion rather than autolysis in the release of some *Helicobacter pylori* proteins. *Infect Immun* 66:1023–1027.
12. Bertini I, Gray H, Stiefel E, Valentine J (2006) *Biological Inorganic Chemistry: Structure and Reactivity* (University Science Books, Mill Valley, CA).
13. Ge R, Sun H (2007) Bioinorganic chemistry of bismuth and antimony: Target sites of metallodrugs. *Acc Chem Res* 40:267–274.
14. Jacob C, Maret W, Vallee B (1998) Control of zinc transfer between thionein, metallothionein, and zinc proteins. *Proc Natl Acad Sci USA* 95:3489–3494.
15. Coleman JE (1992) Structure and mechanism of alkaline phosphatase. *Annu Rev Biophys Biomol Struct* 21:441–483.
16. Sun H, Li H, Harvey I, Sadler PJ (1999) Interactions of bismuth complexes with metallothionein(II). *J Biol Chem* 274:29094–29101.
17. Weber F, Keppel F, Georgopoulos C, Hayer-Hartl M, Hartl F (1998) The oligomeric structure of GroEL/GroES is required for biologically significant chaperonin function in protein folding. *Nat Struct Biol* 5:977–985.
18. Goldfarb AR, Saidel LJ, Mosovich E (1951) The ultraviolet absorption spectra of proteins. *J Biol Chem* 193:397–404.
19. Perham M, Wittung-Stafshede P (2007) Folding and assembly of co-chaperonin heptamer probed by forster resonance energy transfer. *Arch Biochem Biophys* 464:306–313.
20. Perham M, Chen M, Ma J, Wittung-Stafshede P (2005) Unfolding of heptameric co-chaperonin protein follows "fly casting" mechanism: Observation of transient nonnative heptamer. *J Am Chem Soc* 127:16402–16403.
21. Ochman H, Davalos LM (2006) The nature and dynamics of bacterial genomes. *Science* 311:1730–1733.
22. Thorne JL (2000) Models of protein sequence evolution and their applications. *Curr Opin Genet Dev* 10:602–605.
23. Podlaha O, Zhang J (2003) Positive selection on protein-length in the evolution of a primate sperm ion channel. *Proc Natl Acad Sci USA* 100:12241–12246.
24. Woycechowsky KJ, Raines RT (2003) The CXC motif: A functional mimic of protein disulfide isomerase. *Biochemistry* 42:5387–5394.
25. Murzin AG (1996) Structural classification of proteins: New superfamilies. *Curr Opin Struct Biol* 6:386–394.
26. Jin L, Szeto KY, Zhang L, Du W, Sun H (2004) Inhibition of alcohol dehydrogenase by bismuth. *J Inorg Biochem* 98:1331–1337.
27. Mukherjee PK, et al. (2006) Alcohol dehydrogenase restricts the ability of the pathogen *Candida albicans* to form a biofilm on catheter surfaces through an ethanol-based mechanism. *Infect Immun* 74:3804–3816.
28. Lalonde RL, et al. (2007) Model-based drug development. *Clin Pharmacol Ther* 82:21–32.
29. Raymond J, et al. (2004) Genetic and transmission analysis of *Helicobacter pylori* strains within a family. *Emerg Infect Dis* 10:1816–1821.
30. Ge R, et al. (2006) Thermodynamic and kinetic aspects of metal binding to the histidine-rich protein, Hpn. *J Am Chem Soc* 128:11330–11331.
31. Zhang L, et al. (2006) Inhibition of urease by bismuth(III): Implications for the mechanism of action of bismuth drugs. *BioMetals* 19:503–511.
32. van Dinten L, van Tol H, Gorbalenya A, Snijder E (2000) The predicted metal-binding region of the arterivirus helicase protein is involved in subgenomic mRNA synthesis, genome replication, and virion biogenesis. *J Virol* 74:5213–5223.
33. Yang N, et al. (2007) Bismuth complexes inhibit the SARS coronavirus. *Angew Chem Int Ed Engl* 46:6464–6468.
34. Zeldovich KB, Chen P, Shakhnovich EI (2007) Protein stability imposes limits on organism complexity and speed of molecular evolution. *Proc Natl Acad Sci USA* 104:16152–16157.
35. Graham DY, Shiotani A (2008) New concepts of resistance in the treatment of *Helicobacter pylori* infections. *Nat Clin Pract Gastr* 5:321–331.
36. Axon A (1999) Treatment of *Helicobacter pylori*: Where are we now? What are the key questions?. *Eur J Gastroen Hepat* 11(Suppl 2):S3–7 discussion S23-24.