doi:10.1016/j.jmb.2008.04.024 J. Mol. Biol. (2008) 379, 1029–1044

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Intense Neutral Drifts Yield Robust and Evolvable Consensus Proteins

Shimon Bershtein, Korina Goldin and Dan S. Tawfik*

Department of Biological Chemistry, Weizmann Institute of Science, Rehovot, 76100, Israel

Received 23 February 2008; received in revised form 8 April 2008; accepted 9 April 2008
Available online 16 April 2008

What changes occur when a natural protein that had been under low mutation rates is subjected to a neutral drift at high mutational loads, thus generating genetically diverse (polymorphic) gene ensembles that all maintain the protein's original function and structure? To address this question we subjected large populations of TEM-1 β-lactamase to a prolonged neutral drift, applying high mutation rates and purifying selection to maintain TEM-1's existing penicillinase activity. Purging of deleterious mutations and enrichment of beneficial ones maintained the sequence of these ensembles closer to TEM-1's family consensus and inferred ancestor. In particular, back-to-consensus/ancestor mutations that increase TEM-1's kinetic and thermodynamic stability were enriched. These acted as global suppressors and enabled the tolerance of a broad range of deleterious mutations, thus further increasing the genetic diversity of the drifting populations. The probability of a new function emerging (cefotaxime degradation) was also substantially increased in these ensembles owing to the presence of many gene variants carrying the global suppressors. Our findings indicate the unique features of large, polymorphic neutral ensembles generated under high mutational loads and prompt the speculation that the progenitors of today's proteins may have evolved under high mutational loads. The results also suggest that predictable back-to-consensus/ancestor changes can be used in the laboratory to generate highly diverse and evolvable gene libraries.

Keywords: protein evolution; neutral drift; mutational robustness; consensus/ancestor sequence; protein evolvability

Edited by J. Karn

Introduction

The focus of molecular and experimental evolution studies has been primarily adaptive evolution, namely, the acquisition of new protein functions. However, most of the genetic diversity on this planet is the outcome of neutral drifts, i.e., gradual sequence changes that occur under selection to maintain the existing function and structure. However, the majority of mutations undermine protein stability and folding,[1–3] and a smaller fraction of active-site mutations directly diminishes its function.[4,5] Purifying (or negative) selection constantly purges such mutations from neutrally drifting populations.[6] Living organisms are generally maintained at a mutation-selection equilibrium whereby the rate of mutations and the selection pressure dictate a certain level of genetic diversity (polymorphism) and fitness.[7] The mutation rates that underline the contemporary organisms and their proteins are generally very low (in the order of 10^{-9} substitutions per base per generation).[7] But these rates need not be so slow in all organisms, or throughout evolution. RNA viruses, where mutation rates can be as high as 10^{-4}, are a notable exception. The result is highly polymorphic populations where evolution acts not only on single, wild-type genes, but also on whole ensembles, networks, or quasi-species of sequences that stem from a given "wild type".[8–12]

What would be the effects of increasing the mutation rate on a natural protein that has been constantly

*Corresponding author. E-mail address: tawfik@weizmann.ac.il.

Abbreviations used: ORF, open reading frame; GdmCl, guanidinium chloride; dPTP, 6(2-deoxy-beta-D-ribofuranosyl)-3,4-dihydro-8H-pyrimido[4,5-c][1,2]-oxazin-7-one triphosphate; 8-oxo-dGTP, 8-oxo-2′-deoxyguanosine triphosphate.

evolving under low mutational input? What would be the features of the ensemble of highly polymorphic sequences that stem from such a rapidly drifting protein?

To address these questions we applied laboratory evolution methodologies and performed an accelerated neutral drift of a model protein, TEM-1 β-lactamase, at a rate that is over a millionfold higher than observed in the bacteria in which this enzyme is found in nature. This rate allows a considerable sequence drift to be achieved within the timescale of a laboratory experiment. Several other protein drifts have been recently performed in the laboratory by subjecting different enzymes to iterative rounds of mutation and purifying selection that removes deleterious mutations.[13–16] The outcome of these accelerated drifts was highly polymorphic populations, namely, large ensembles of genes that are all derived from the same parental wild-type gene each containing a different set of mutations, while maintaining the original structure and function. The results indicate large changes in latent, promiscuous functions that occur under selection to maintain the protein's original activity.[13,15] Indications that highly polymorphic protein populations evolve towards higher thermodynamic stability and mutational robustness were also reported.[16]

Previously, we reported 10 rounds of mutational drift of TEM-1 and the changes in fitness (expression and activity levels) that followed the accumulation of mutations.[14] Here we report the outcome of 18 rounds of drift, with a focus on the sequence changes that occurred under the intense mutational rates applied, i.e., a sequence "signature" that may characterize proteins that are subjected to an accelerated neutral drift. We also sought to examine how traits such as the ability to tolerate highly deleterious mutations change under an intense mutational drift. Because mutations that alter protein function are particularly destabilizing,[4,5] the ability to compensate for a wider range and a higher number of mutations is also expected to increase the rate of adaptation (evolvability).[5,11,17] Our data indicate that the response to higher mutation rates involves a known and predictable type of sequence changes that bring the sequence of the drifting variants closer to their family consensus and inferred ancestor (*back-to-consensus/ancestor mutations*). These mutations act as global suppressors that mitigate the destabilizing effects of many different mutations, including mutations that endow the drifting protein a new function.

Results

The experimental system

TEM-1 β-lactamase, an enzyme that efficiently degrades penicillins such as ampicillin, served as our model protein. As described,[14] TEM-1's gene was cloned into a pUC19 plasmid under its endogenous promoter. We developed a highly reproducible *in vitro* random mutagenesis protocol and calibrated it to an average mutational rate of two mutations per gene per round of mutagenesis. Recloning after each round of mutagenesis confined the mutational drift to TEM-1's open reading frame (ORF). Throughout this experiment, we maintained three populations of drifting genes. Each population, or plasmid library, was separately mutated, ligated into an empty vector, transformed into *Escherichia coli* host cells, and then underwent purifying selection by plating the transformed bacteria on agar plates containing an appropriate concentration of ampicillin: 250 μg/ml of ampicillin comprised the "high" selection pressure (Lib250), since this concentration approaches the maximum at which wild-type TEM-1 confers 100% survival to freshly transformed *E. coli* cells. The concentration of the "low" selection pressure was 20-fold lower and above the lowest effective concentration of ampicillin (12.5 μg/ml; Lib12.5). No ampicillin was applied for the Lib0. After growth on selection plates, plasmid DNA was extracted from the surviving *E. coli* colonies and the TEM-1 gene was subjected to the next round of mutagenesis.[14]

Altogether, we now describe 18 successive rounds of mutagenesis and purifying selection performed under the previously described conditions.[14] We ensured the critical absence of bottlenecks throughout: The loss of diversity observed was $\leq 50\%$ per round, and $\geq 10^6$ independent clones per library were maintained throughout. The large, polymorphic drifting populations thus met the conditions that were predicted to induce mutational robustness.[11,16,18,19]

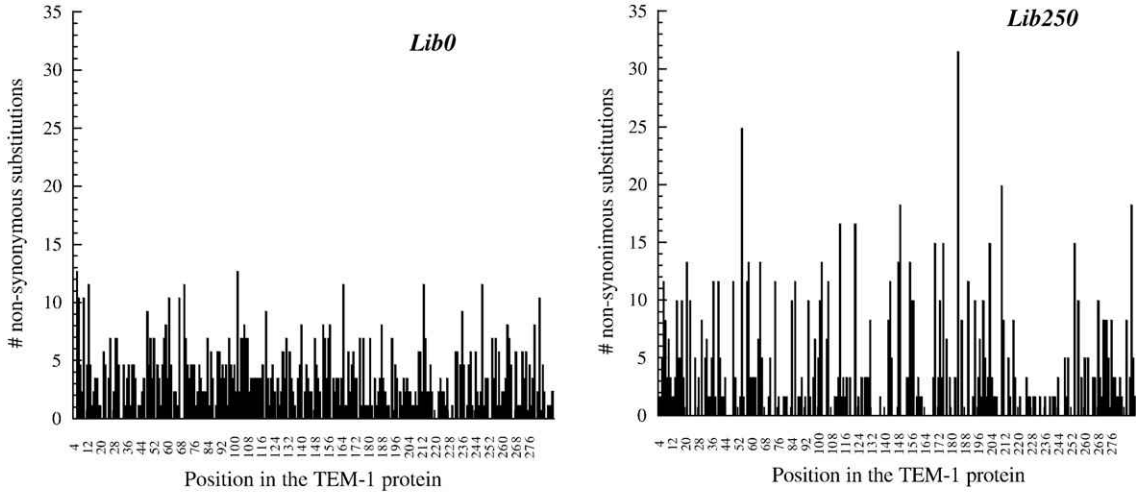
Sequence changes that follow the neutral drift

Variants from each of the three libraries (Lib0, Lib12.5, and Lib250) were randomly picked and sequenced after the 5th, 10th, 15th, and 18th round of mutagenesis and selection. In total, >1200 mutations per library were registered and analyzed. Obviously, the number of mutations per gene increased as the drift progressed. However, since our analysis regarded the frequency of mutations per individual position (see below), we aimed at a similar number of analyzed mutations, rather than a similar number of genes (e.g., 46 clones were sequenced from the 5th round of Lib250, as opposed to 14 clones from the 18th round). As expected, purifying selection caused a reduction in the fraction of non-synonymous mutations—from 71% in the unselected library (Lib0) to about 50% in the libraries under selection (Lib12.5 and Lib250) (Table 1). Statistical analysis performed on the distributions of the non-synonymous mutations also showed a significant difference between the selected and unselected libraries. In particular, the number of sites in which no amino acid substitutions could be detected (15.7% in Lib0) increased following selection to 32.4% in Lib12.5 and 37.6% in Lib250 (Fig. 1). These phenomena are indicative of negative, purifying selection acting on the neutrally drifting TEM-1 populations.

Table 1. Sequence analysis of clones randomly picked from the drifting libraries

	Lib0					Lib12.5					Lib250				
	Round					Round					Round				
	5	10	15	18	Total	5	10	15	18	Total	5	10	15	18	Total
Sequenced clones (no.)	17	14	12	11	54	38	32	20	14	104	46	35	22	14	117
Total no. mutations	200	288	336	400	1224	244	376	352	300	1272	257	373	337	294	1261
Average no. mutations per gene	12	21	28	36		6	12	18	21		6	11	15	21	
No. non-synonymous mutations					866 (71%)					641 (50%)					654 (52%)

For a more detailed analysis, we compared the frequency of the non-synonymous substitutions in the selected and unselected libraries per position, using a chi-square test ($P<0.05$) (Supplementary Table 1). To further ensure that the results are of statistical significance, we analyzed only sites in Lib12.5 and Lib250 where the mutation frequency was above 12.7 substitutions per site per 10^3 non-synonymous mutations (12.7 is the highest number of substitutions per *any* site of Lib0; Fig. 1). The analyzed TEM-1 positions were divided into two groups. The first group included positions in which a significant decrease in the number of non-synonymous substitutions in Lib12.5 and Lib250 was observed compared to Lib0, and probably reflects the outcome of negative, purifying selection (Table 2 and Fig. 2a). The second group comprises positions in which non-synonymous mutations were enriched, i.e., their frequency in Lib12.5 and Lib250 is much higher than the frequency in the same position in Lib0 and, therefore, might reflect positive, adaptive selection (Table 3 and Fig. 2b), particularly because under the populations sizes applied here, fixation of random mutations by pure chance is not expected.[21]



	Lib0	*Lib12.5*	*Lib250*
% of positions in which no non-synonymous substitutions have occurred	15.7	32.4	37.6
maximum number of substitutions detected per site (per 1,000 non-synonymous mutations)	12.7	20.3	32.1

	P-value (KS-test)	**D-value (KS-test)**	**P-value (Wilcoxon rank-sum test)**
Lib0-Lib12.5	$1.4 \cdot 10^{-6}$	0.19	0.14
Lib0-Lib250	$4 \cdot 10^{-9}$	0.25	$8.9 \cdot 10^{-5}$
Lib12.5-Lib250	$5 \cdot 10^{-7}$	0.23	0.02

Fig. 1. Distribution of the non-synonymous mutations identified in the unselected (Lib0) and selected (Lib250; "high" level of the purifying selection) libraries. Amino acid substitutions that have occurred in every position of the protein in Lib0 and Lib250 throughout the 5th, 10th, 15th, and 18th round of mutagenesis and selection were calculated and expressed as the number of non-synonymous substitutions per site per 1000 mutations. The statistical significance of the distribution was determined using Kolmogorov–Smirnov test and Wilcoxon rank sum tests. Both tests determine whether two data sets differ significantly despite their distributions being non-parametric, and non-normal. *D* value defines the maximum difference between cumulative distribution in Kolmogorov–Smirnov test.

Table 2. Purged mutations

Position	Wild-type TEM-1 sequence	Mutations found in Lib0	Consensus/ Ancestor	*P* value Lib12.5	*P* value Lib250
27	P	L,S (6)	P/I	0.03	0.03
46	Y	H,D,C (8)	Y/V	0.02	0.01
50	D	G,E,N (6)	D/D		0.03
57	L	F,A,R,P (7)	L/L	0.02	
66	F	L,S,V,C (9)	F/F		0.01
68	M	T,L,V (10)	M/M		0.02
102	L	T,S,V (6)	L/L		0.03
105	Y	W,C,H (6)	Y/Y		0.03
106	S	P,L (6)	S/S		0.03
138	L	F,P,R,V (7)	L/L	0.02	0.02
152	L	S,V,T (6)	L/L	0.03	
163	D	G,A,E,N (10)	D/D	0.01	0.01
179	D	G,V,N (6)	D/D	0.03	0.03
186	M	T,V,L (7)	M/M		0.02
211	M	V,T,R,L (10)	M/M	0.02	0.01
263	I	T,S,L (7)	I/I		0.02
264	Y	N,H,C,S (6)	Y/Y	0.03	0.03

Listed are all positions in which the mutation frequency was significantly lower (chi-square P value<0.05) in the selected libraries (Lib12.5 and Lib250) *versus* the unselected library (Lib0). The numbers in parenthesis represent the actual number of mutations identified.

Drifting TEM-1 is maintained closer to its family consensus

Mutations in highly conserved residues usually cause large decreases in stability and activity.[22] Conversely, reverting protein residues that deviate from the consensus amino acid can increase stability.[23–28] Mutating residues to match the inferred ancestral sequence (which is very close, but not necessarily identical to the consensus) has also been shown to increase stability.[29–31] Indeed, almost all the purged mutations fall in the away-from-consensus category, whilst half of the enriched substitutions are back-to-consensus/ancestor (Fig. 2c and Tables 2 and 3). For example, positions 152 and 153 are highly conserved in the β-lactamase family. In TEM-1, position 152 is occupied by a conserved Leu, while position 153 is exceptionally a His (Arg in all TEM-1's closest homologues; Fig. 2c). Under selection, mutations at positions 152 to Ser, Trp, and Val that accumulated in Lib0 were purged, whereas a mutation that reverted position 152 to consensus (His152Arg) was enriched.

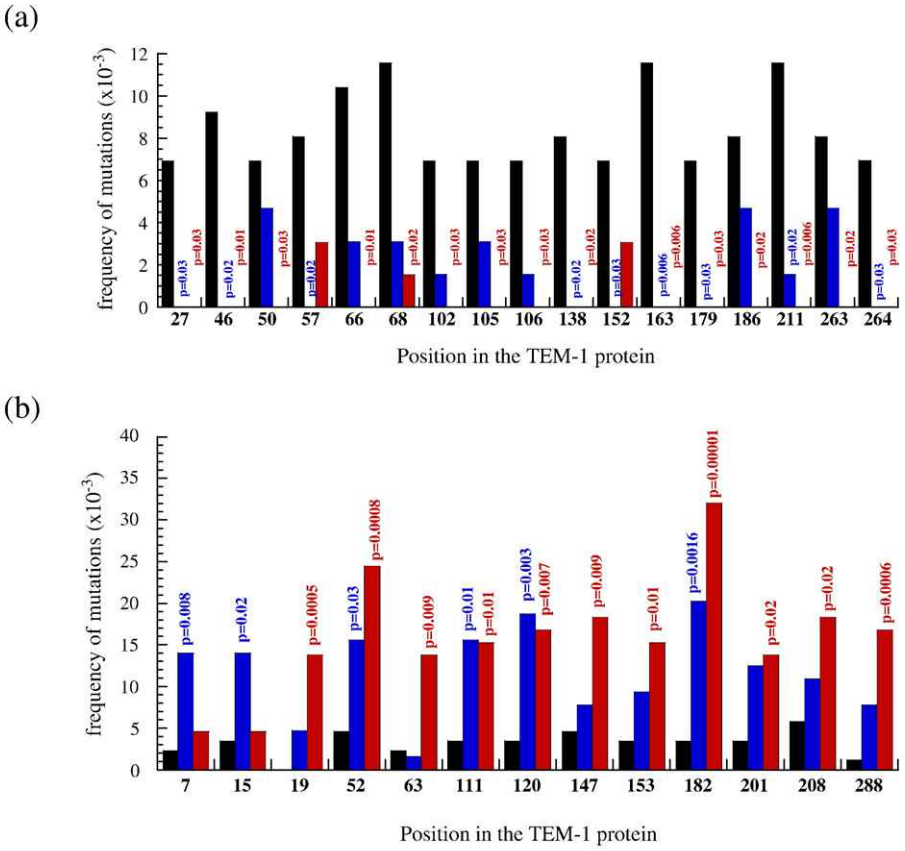
The enriched mutations increase TEM-1's stability

The mutations under positive selection could be related to adaptation of TEM-1 to the particulars of our selection system by providing, e.g., improved levels of expression in *E. coli*, or higher catalytic efficiency. Alternatively, the enriched mutations might be compensating for the loss of TEM-1's stability due to the high mutation rates. The role of the enriched mutations in increasing TEM-1's stability is suggested by their identification as back-to-consensus/ancestor mutations of the type that has been shown to stabilize numerous proteins,[24,26,27,29–32] and by the previously identified role of Met182Thr in increasing the thermodynamic stability of TEM-1 by 2.7 kcal/mol and compensating for a wide range of destabilizing mutations.[5,17,33] The Met182Thr mutation, along with Arg120Gly, Glu147Gly, and His153Arg (also enriched in our experiment), was also observed when a destabilized form of TEM-1 was evolved to regain stability.[34]

The overall impact of the enriched and purged mutations on protein stability was also confirmed by the FoldX analysis of the stability changes of the mutations. The FoldX empirical computation has been experimentally validated, and, although the values of individual mutations can vary, overall trends are accurately predicted.[2,35] FoldX indicated that destabilizing mutations frequently observed in the unselected library Lib0 were purged by the purifying selection. Highly destabilizing mutations ($\Delta\Delta G$ >4 kcal/mol) were almost completely purged, and the average $\Delta\Delta G$ value declined from 1.3 kcal/mol in Lib0 to 0.74 kcal/mol for Lib12.5 and 0.6 kcal/mol for Lib250 (Fig. 3). The predicted $\Delta\Delta G$ values of the identified purged and enriched mutations also revealed pronounced shifts in stability (an average of 2 kcal/mol for the purged mutations and ~0 kcal/mol for the enriched mutations).

To study the precise nature of the enriched back-to-consensus mutations, we generated five variants corresponding to the most common enriched mutations: Arg120Gly, Glu147Gly, His153Arg, Met182Thr, and Leu201Pro (Table 3). An additional variant composed the most frequently observed combination of Arg120Gly *plus* Met182Thr.

The catalytic parameters of the mutated variants exhibited no considerable or consistent improvement in response to the back-to-consensus mutations (Table 4). This result is in agreement with the levels of ampicillin resistance and the periplasmic expression levels, being identical, within experimental error, to wild-type TEM-1 (data not shown). By contrast, the biophysical measurements indicated substantial and consistent changes in stability. The latter could be divided into changes in thermodynamic *versus* kinetic stability.[36,37] In cases where the native state is a global energy minimum that is reached regardless of the starting configuration, increasing the energy gap between the unfolded and native states (ΔG_{U-N}) increases the protein's overall stability. However, because TEM-1's folding pathway is populated by multiple intermediate states that can misfold, different final states might be reached depending on the starting configuration.[38,39] Mutations may affect the kinetics by which certain intermediates form or disappear and thereby affect protein stability. Upon thermal unfolding, all the tested variants appeared to be more stable with apparent transition midpoint temperatures (T_m^{app}) higher than that of wild-type TEM-1 (from 1.4 °C higher for the Leu201Pro mutant up to 7.3 °C for the Arg120Gly-Met182Thr double



(c)

Position		Purged mutations																	Enriched mutations												
		27	46	50	57	66	68	102	105	106	138	152	163	179	186	211	263	264	7	15	19	52	63	111	120	147	153	182	201	208	288
TEM-1 (*Klebsiella pneumoniae*)	100%	P	Y	D	L	F	M	L	Y	S	L	L	D	D	M	M	I	Y	H	F	F	N	E	K	R	E	H	M	L	I	K
Most frequent mutation		S	H	G	P	L	V	W	C	P	V	S	A	G	T	V	S	H	R	C	L	D	A	E	G	G	R	T	P	L	R
Pla-6 (*Raoultella planticola*)	72%	Q	Y	D	L	F	M	L	Y	S	L	L	D	D	M	M	I	Y	Q	L	L	A	D	K	G	G	R	T	S	Q	R
ORN-5 (*Raoultella ornithinolytica*)	71%	E	Y	D	L	F	M	L	Y	S	L	L	D	D	M	M	I	Y	Q	L	L	A	G	K	G	G	R	T	P	Q	R
AAB (*Klebsiella pneumoniae*)	68%	P	M	D	L	F	L	L	Y	S	L	L	D	D	M	M	I	Y	Y	L	L	A	S	K	G	G	R	T	A	L	E
SHV-1 (*Klebsiella pneumoniae*)	65%	P	M	D	L	F	M	L	Y	S	L	L	D	D	M	M	I	Y	Y	L	L	A	D	K	G	G	R	T	A	L	E
Len-1 (*Klebsiella pneumoniae*)	64%	P	I	D	L	F	M	L	Y	S	L	L	D	D	M	M	I	Y	Y	L	L	A	D	K	G	G	R	T	A	L	E
Ohio-1 (*Enterobacter cloacae*)	61%	L	M	D	L	F	M	L	Y	S	L	L	D	D	M	M	I	Y	Y	L	A	A	D	K	G	G	R	T	A	L	E
CKO-1 (*Citrobacter koseri*)	48%		L	D	T	F	L	L	Y	S	A	L	D	D	I	M	I	Y	K		L	A	E	K	A	A	R	T	P	M	S
CARB-5 (*Acinetobacter calcoaceticus*)	47%	D		D	W	F	L	L	Y	S	L	L	D	D	M	L	L	Y	S	F	A	E	N	K	A	A	R	T	I	E	K
HERA-2 (*Escherichia hermannii*)	45%	P	V	N	V	F	L	L	Y	A	L	M	D	D	M	M	I	F	L	L		Q	D	N	R	A	R	S	P	A	A
AST-1 (*Nocardia asteroides*)	45%		L	D	V	F	F	L	Y	S	L	I	D	D	L	L	V	L				G	D	T	R	G	R	T	A	K	D
AmpC (*Vibrio fischeri*)	44%	G	F	N	W	F	M	I	W	S	L	L	D	D	M	M	I	Y	L	N	G	A	D	P	K	S	R	T	P	K	N
blaF (*Mycobacterium fortuitum*)	43%	A	V	N	I	F	L	L	N	S	L	A	D	D	L	M	T	R	L		P	Q	D	A	A	A	R	T	P	E	P
ABPS-1 (*Burkholderia pseudomallei*)	42%	V	V	D	I	F	M	L	Y	S	I	A	D	D	M	M	V	Y			A	A	D	Q	A	A	R	T	A	N	R
L2 (*Stenotrophomonas maltophilia*)	42%	I	I	D	I	F	M	L	H	A	L	L	D	D	M	L	A	Y	F	S	V	A	D	R	R	A	R	T	P	A	R
PSE-4 (*Pseudomonas aeruginosa*)	42%	Q	V	D	W	F	M	L	Y	S	L	L	D	D	I	M	I	Y	L			Q	N	K	D	G	R	T	E	E	D
NB64 (*Xanthomonas axonopodis*)	41%	D	V	D	L	F	M	M	H	A	F	L	D	D	M	M	T	Y	A	A	A	A	H	R	G	G	R	T	P	T	K
JN56		I	V	D	L	F	M	I	Y	S	L	L	D	D	M	M	I	Y	L	I	A	E	D	K	A	A	R	T	P	T	E

Fig. 2. Mutations purged and enriched during the neutral drift. Shown are the frequencies of non-synonymous mutations that were significantly (chi-square P value <0.05) purged (a) or enriched (b), in Lib0 (black bars), Lib12.5 (blue bars), and Lib250 (red bars). (c) Away-from- and back-to-consensus mutations. Sequences of class A β-lactamases (numbers to the left represent percent of amino acid identity to TEM-1) were collected from a phylogenetic analysis[20] and aligned by TCoffee. The consensus residues are marked in red. Also aligned is the inferred ancestor of the subgroup of class A β-lactamases to which TEM-1 belongs (JN56), and inferred ancestors that reside deeper in the β-lactamase phylogeny can be found in Supplementary Fig. 1. Shown are representative positions of the alignment in which mutations were purged or enriched. The most frequent mutation reflects the most frequently occurring mutations in Lib0 (for the purged mutations), or in Lib250 (for the enriched mutations). Information regarding the pattern and the frequencies of all purged and enriched positions is provided in Tables 2 and 3.

mutant; Fig. 4a, Table 4, and Supplementary Fig. 2). However, due to its kinetically controlled folding, TEM-1's thermal denaturation is not completely reversible (at least not under the conditions applied here). Thus, increases in apparent T_m do not necessarily reflect increases in thermodynamic stability *per se*, although the tested mutants also showed a substantial increase in the fraction of refolded protein (≥2.5-fold; Table 4 and Supplementary Fig. 2).

The effects of the enriched back-to-consensus mutations on thermodynamic *versus* kinetic stability were also seen in the chemical denaturation experiments in which one of the intermediate states along TEM-1's folding pathway can be detected.[34,40] The

Table 3. Enriched mutations

Position	Lib12.5		Lib250		Consensus/ Ancestor
	Mutation, %	*P* value	Mutation, %	*P* value	
7	H→R, 56 (5) H→Y, 22 (2) H→P, 11 (1) H→Q, 11 (1)	0.008			Q,Y/L
15	F→C, 56 (5) F→L, 33 (3) F→V, 11 (1)	0.02			L/L
19			F→C, 44 (4) F→L, 44 (4) F→V, 12 (1)	0.0005	L/A
52	N→D, 57 (6) N→S, 29 (3) N→T, 14 (1)	0.03	N→S, 38 (6) N→D, 31 (5) N→H, 25 (4) N→T, 6 (1)	0.0008	A/E
63			E→A, 67 (6) E→G, 11 (1) E→K, 11 (1) E→D, 11 (1)	0.009	D/D
111	K→E, 56 (6) K→R, 22 (2) K→Q, 22 (2)	0.01	K→R, 70 (7) K→Q, 20 (2) K→T, 10 (1)	0.01	K,Q,R/K
120	R→G, 44 (5) R→S, 44 (5) R→E, 12 (2)	0.003	**R→G, 91 (10)** R→S, 9 (1)	0.007	G/A
147			**E→G, 75 (9)** E→A, 17 (2) E→K, 8 (1)	0.009	G/A
153			**H→R, 100 (10)**	0.01	R/R
182	**M→T, 90 (12)** M→V, 10 (1)	0.0016	M→T, 62 (13) M→V, 19 (4) M→R, 9 (2) M→K, 5 (1) M→L, 5 (1)	0	T/T
201			**L→P, 78 (7)** L→Q, 11 (1) L→R, 11 (1)	0.02	A,P/P
208			I→L, 33 (4) I→V, 33 (4) I→M, 25 (3) I→T, 9 (1)	0.02	L,M,Q/T
288			K→R, 45 (5) K→E, 27 (3) K→T, 18 (2) K→Q, 10 (1)	0.0006	E,R/E

Listed are all the positions in which the mutation frequency in the selected libraries (Lib12.5 and Lib250) was ≥12.7 per 1000 mutations, and was also significantly higher (chi-square *P* value <0.05) than the frequency at the same position in Lib0. Enriched positions in which the same amino acid substitution occurred in ≥75% of the cases are in bold.
The numbers in parenthesis represent the actual number of mutations identified.

fluorescence curves were fitted to a three-state model (N↔H↔U), and the midpoints for the intermediate native transition (Cm_{H-N}) and the unfolded intermediate transition (Cm_{U-H}) were derived as described[40] (Fig. 4b, Table 4, and Supplementary Fig. 3). The Met182Thr mutation increases TEM-1's thermostability by 2.7 kcal/mol[5] and, as expected, exhibits a higher Cm_{H-N} value, whereas the midpoint of the unfolded intermediate transition (Cm_{U-H}) remains unchanged (Fig. 4b, Table 4). In contrast, the influence of the Arg120Gly substitution on the intermediate state was prominent (an increased

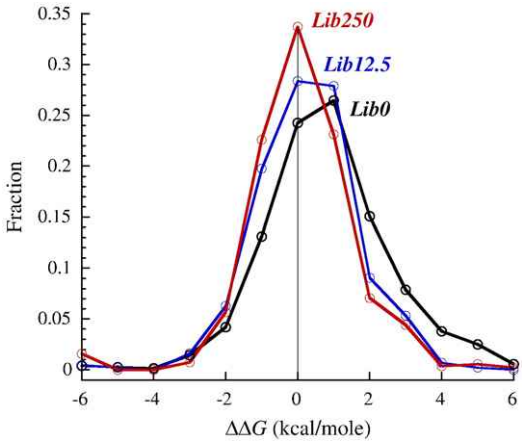


Fig. 3. Distribution of stability effect of non-synonymous mutations accumulated in Lib0, Lib12.5, and Lib250 throughout 18 rounds of mutagenesis and selection. The stability changes induced by these mutations (predicted $\Delta\Delta G$ values, in kcal/mol) were individually computed with FoldX, a structure-based algorithm that has been shown to predict $\Delta\Delta G$ values with considerable fidelity ($R = 0.83$ for a data set of 1030 mutations in 27 proteins).[2,35] The predicted $\Delta\Delta G$ values were arranged in 1-kcal/mol bins. Plotted are the distributions of all non-synonymous mutations for the unselected library (Lib0; black), and the two selected libraries: Lib12.5, "low" selection stringency (in blue), and Lib250, "high" selection stringency (in red).

Cm_{U-H} value), while the Cm_{H-N} value remained unchanged. It appears therefore that the Met182Thr mutation primarily influences TEM-1's thermodynamic stability, whereas Arg120Gly affects the intermediate state and thus increases TEM-1's kinetic stability. The other mutations seem to affect both features to various extents. Indeed, the behavior of the double mutant Arg120Gly-Met182Thr indicated that the kinetic and thermodynamic effects can be additive, as this mutant shows an increase in both Cm_{H-N} (as in Met182Thr) and Cm_{U-H} (as in Arg120Gly) (Table 4). Alternatively, the other mutations may affect additional intermediate states that remain undetected in the chemical denaturation experiments.

That the outcome of TEM-1's folding largely depends on the starting configuration was also demonstrated by refolding the protein from different initial states induced by pre-incubation in different guanidinium chloride (GdmCl) concentrations.[38] Wild-type TEM-1 refolds rather well from the unfolded state (>70% refolding after pre-incubation in ≥3 M GdmCl), but tends to misfold to a larger degree when refolded from an intermediate state (40% refolding from 1 M GdmCl) (Fig. 4c and Supplementary Fig. 4). The refolding of all tested mutants was less sensitive to the starting configuration, with His153Arg being the most kinetically stable single mutant as appears from this experiment. Notably, the double mutant Arg120Gly-Met182Thr exhibited ≥70% refolding regardless of the starting configuration.

The enriched mutations compensate for many destabilizing mutations

The stabilizing nature of the enriched mutations was also confirmed by their ability to compensate for highly deleterious mutations incorporated into TEM-1. As the first test, we chose Leu76Asn that results in dramatically reduced levels of ampicillin resistance (<50 μg/ml *versus* ~3000 μg/ml for wild type) that are largely alleviated by Met182Thr.[33] To test the compensatory capability of the intensely drifting TEM-1 libraries, we incorporated the Leu76Asn mutation on the various libraries at the 2nd to 18th rounds and selected the mutated libraries at a range of ampicillin concentrations (500–2500 μg/ml) that are well above the concentration at which the Leu76Asn mutant survives. Starting with the 5th round, a subpopulation of ~1% of the variants within the TEM-1 population

Table 4. Catalytic and biophysical parameters of the TEM-1 variants

	Catalytic parameters			Thermal denaturing		Chemical denaturing	
Variant	k_{cat} (s^{-1})	K_M ($\times 10^{-6}$ M)	k_{cat}/K_M ($\times 10^{6}$ M^{-1} s^{-1})	T_m^{app} (°C)	Reversibility of unfolding (%)[a]	Cm_{H-N}[b] (M)	Cm_{U-H}[b] (M)
wt TEM-1	199±2[c]	45.5±1.4[c]	4.4	54.2±0.03[d]	23	1.3	2.9
R120G	226±3	30.7±1	7.4	56±0.02	41	1.3	3.3
E147G	248±0.3	50.8±1.2	4.9	56.8±0.02	31	1.4	2.9
H153R	222±2.1	38.3±1.3	5.8	57.5±0.03	35	1.3	3
M182T	169±5	35.8±0.9	4.7	59.2±0.03	50	1.7	2.9
L201P	214±3.2	38.7±1.8	5.5	55.6±0.03	51	1.2	3
R120G+ M182T	149±1.5	28.5±0.9	5.2	61.5±0.03	58	1.8	n.d.[e]

[a] Reversibility of unfolding was tested by estimating the percentage of refolding obtained by cooling the purified variants after graduate heating to 80 °C. Since each variant might be populated by different intermediate states that behave differently (in terms of misfolding and aggregation) throughout graduate heating, the comparison of the reversibility of unfolding might be biased by the gradual heating that preceded the cooling.
[b] The transition midpoints values were derived by fitting the obtained data to a three-state model (see Materials and Methods).
[c] Catalytic parameters were obtained by measuring the rate of hydrolysis of nitrocefin substrate (see Materials and Methods). The standard deviations result from two independent measurements.
[d] The standard deviation refers to the deviation from the two-state model for a single data set (see Materials and Methods).
[e] Cm_{U-I} is >5 indicates that the intermediate could not be detected in the range of GdmCl concentration tested.

drifting under the high selection stringency (Lib250) possessed compensatory mutations capable of buffering Leu76Asn to different extents (Supplementary Fig. 6). Interestingly, the fraction of compensated clones in the selected libraries, and Lib250 in particular, remained largely unchanged throughout the drift, although the overall "fitness" (activity and expression levels) of the drifting population declined (Supplementary Fig. 5). (The fitness of the libraries subjected to purifying selection remained unchanged at concentrations under the applied selection thresholds. At concentrations exceeding the selection thresholds, constant decreases in fitness were observed. For instance, whereas wild-type TEM-1 withstands >2500 μg/ml ampicillin, only ~40% of Lib250 variants remained viable at ≥2500 μg/ml ampicillin after 18 rounds of mutagenesis and selection.) Several other highly deleterious mutations chosen from a pool of random mutations were tested for compensation (Ile246Asp, Arg222Cys, Leu199Pro, Gln205Pro, Asn136Ser).[14] These mutations occurred in highly conserved residues and resulted in much reduced levels of ampicillin resistance (100–250 μg/ml ampicillin *versus* ~3000 μg/ml for wild type). They were individually introduced onto the drifting TEM-1 libraries after the 2nd and/or the 5th rounds, and variants capable of escaping their deleterious effects were isolated. Similarly to Leu76Asn, all five deleterious mutations tested were compensated, primarily in the neutrally drifting libraries Lib12.5 and Lib250 (Supplementary Fig. 6).

The compensated genes carried the very same consensus/ancestor mutations that were enriched in the drifting libraries, the most frequently occurring mutations being Met182Thr, Arg201Gly, and Leu201Pro (91%, 54%, and 20% of the survived variants, respectively; Fig. 5). The compensatory mutations appeared either on their own or often in combinations that exhibited additive effects (Table 5). For example, the highly destabilizing Leu76Asn mutation is only partially compensated for by Met182Thr (Table 5, clone 1), but in combination with Arg120Gly, its effect is almost completely alleviated (Table 5, clone 2). Additive effects were also exhibited by Glu147Gly and Leu201Pro that, in combination with Met182Thr, were capable of restoring wild-type-like resistance (e.g., clones 6, 10, and 11).

In summary, data described in this and the previous sections support the hypothesis that the intense drift enriched mutations that endow an excess of thermodynamic and kinetic stability and compensatory power towards deleterious mutations. The excess of stability and tolerance to mutations relative to wild-type TEM-1 is observed in a small, yet consistent, subpopulation of the drifting variants and is mediated primarily by the stabilizing consensus/ancestor mutations that act as global suppressors towards a wide range of highly destabilizing, deleterious mutations.

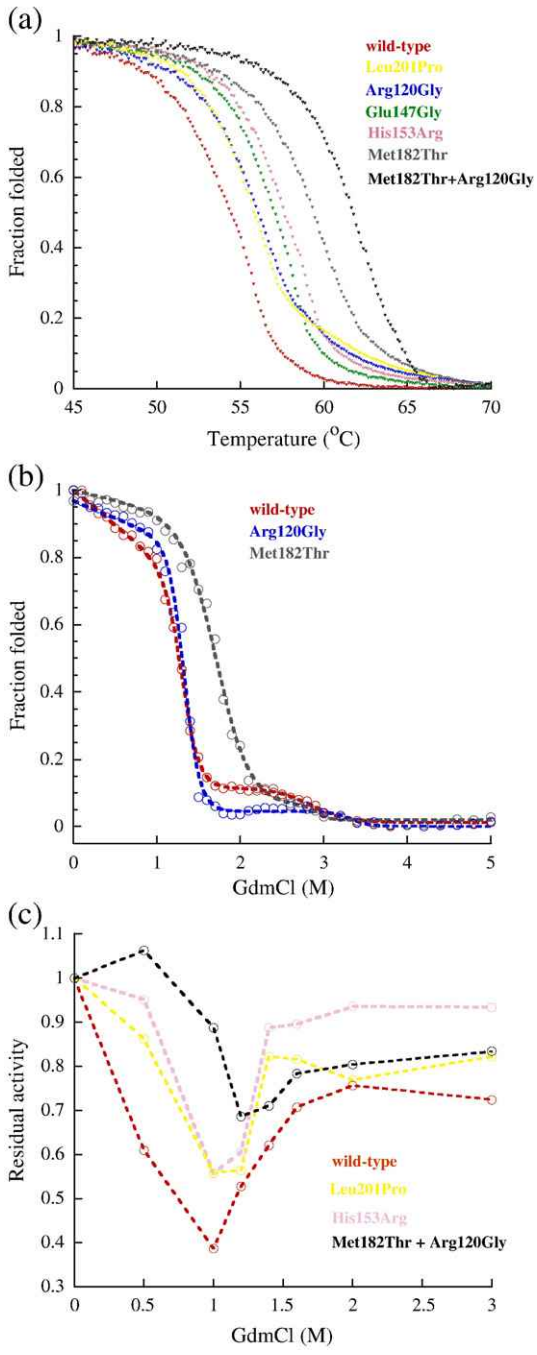


Fig. 4. Stability effects of the enriched mutations. (a) Engineered TEM-1 variants, each carrying a different enriched mutation, or a combination of R120G and M182T, were subjected to thermal denaturation at 25–80 °C at a rate of 1.5 °C/min, while changes in tryptophan fluorescence were monitored. (b) The TEM-1 variants were subjected to chemical denaturation at a range of GdmCl concentrations (0–5 M) while changes in tryptophan fluorescence were monitored. The obtained data were fitted to a three-state model (see Materials and Methods). (c) Refolding from partially denaturated states. The TEM-1 variants were incubated for half an hour at 37 °C in a range of GdmCl concentrations (0–3 M) and subsequently refolded by diluting to near-zero denaturant concentration. The fraction of properly refolded protein was determined by measuring enzymatic activity using nitrocefin substrate at 100 μM. Presented are average values of the two to three independently repeated experiments. Errors were in a range of 4–6%.

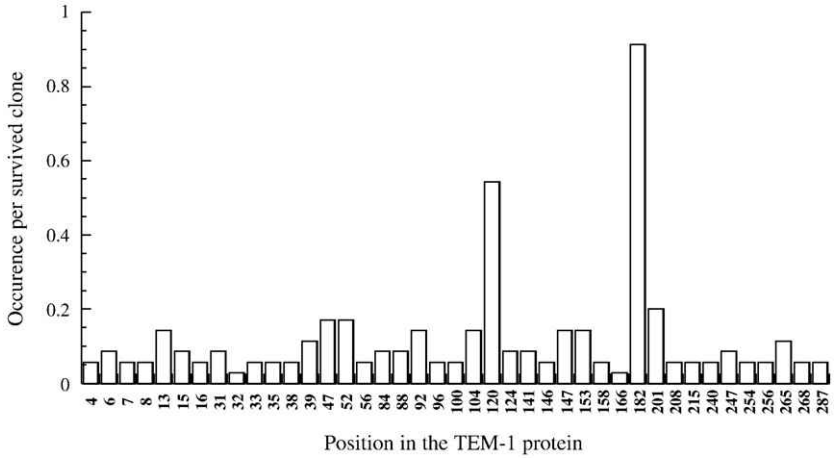


Fig. 5. The most frequent mutations found in Lib250 variants that compensated for the severely destabilizing Leu76Asn mutation. The Leu76Asn mutation was introduced into the Lib250 after the various rounds of the neutral drift (2nd, 5th, 10th, and 18th). Variants capable of compensating for it were isolated by plating the mutated libraries at a range of ampicillin concentrations (500–2500 μg/ml) that were well above the concentration at which wild-type Leu76Asn-TEM-1 survives (50 μg/ml). Shown is the frequency of occurrence per survived clone for all mutations that appeared at least twice in 35 randomly picked surviving clones.

The enriched mutations increase the adaptive potential

Mutations endowing new enzymatic functions are generally destabilizing.[4,5] The compensatory power endowed by the consensus mutations is therefore expected to have an effect on TEM-1's rate of adaptation, or evolvability.[41] Our test for the rate of adaptation made use of cefotaxime—a "third-generation" cephalosporin antibiotic—with which wild-type TEM-1 possesses only negligible activity.[5] However, since the introduction of this new antibiotic, extended-spectrum TEM-1 variants have evolved in the clinic. These carry two primary mutations, Gly238Ser and Glu104Lys, that increase TEM-1's catalytic efficiency towards cephalosporins by ~800-

Table 5. Representative variants with compensated L76N

Clone	Identified mutations	Library	IC (μg/ml)[a]
	Wild-type TEM-1		3000
	L76N		<50
1	*L76N*, **M182T**	Lib0 (round 2)	1000
2	*L76N*, **R120G, M182T**	Lib0 (round 2)	2500
3	I13V, *L76N*, **R120G**, **M182T**	Lib250 (round 2)	3000
4	*L76N*, E104G, I247V, T265M, **M182T**	Lib250 (round 5)	2000
5	F15L, F19L, L21P, V31A, D50E, N52S, *L76N*, **M182T**	Lib250 (round 10)	2000
6	S4N, I13T, *L76N*, S124R, **E147G**, **M182T, L201P**	Lib250 (round 10)	3000
7	I56V, P62S, *L76N*, I84V, H96Y, **R120G**, I173L, **M182T**, I287V	Lib250 (round 10)	3000
8	A36T, I47V, K55E, *L76N*, I84T, Q99R, N100G, **R120G**, S124G, **E147G**,N175D, **M182T**, **L201P**, S243G, T271A,	Lib250 (round 18)	2500
9	F60W, *L76N*, T114A, **R120G**, H158Y, **M182T**, K215R, I247V	Lib250 (round 18)	2500
10	K32R, D35A, Q39R, *L76N*, I84V, **R120G**, N154S, **M182T**, V184M, **L201P**, D254G, Q278R, S285T	Lib250 (round 18)	3000
11	H7Y, F16C, N52D, *L76N*, H96Y, **R120G, H153R, M182T**, D209G	Lib250 (round 18)	3000

Identified global suppressors are marked in bold.
[a] Inhibitory concentration (IC) of ampicillin.

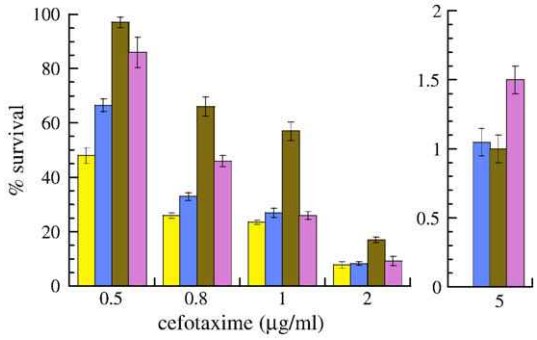


Fig. 6. The frequency of TEM-1 clones conferring new function (cefotaxime resistance at different levels) in various randomly mutated libraries. Libraries were generated by introducing random mutations at a rate of approximately two mutations per gene into (i) wild-type TEM-1 (yellow); (ii) TEM-1 carrying the Met182Thr enriched mutation (brown); (iii) a mixture of four different TEM-1 genes, each carrying a different enriched mutation (Arg120Gly, Glu147Gly, His153Arg, or Leu201Pro) (blue); (iv) Lib250 after the 10th round of mutagenesis and selection (magenta). About 10^6 transformants from each gene library were plated on agar containing 0.3 μg/ml cefotaxime (wild-type TEM-1 only survives ≤0.02 μg/ml) and 100 surviving colonies were randomly picked. The resistance profile (IC values) of these colonies was evaluated by measuring their survival rates on a range of cefotaxime concentrations (0.5–5 μg/ml; no clones from the wild-type library survived 5 μg/ml cefotaxime). Error bars represent standard deviation from two independent experiments.

fold. However, the newly evolved cephalosporinase activity considerably compromises TEM-1's thermodynamic stability, and most cephalosporin-resistant clinical isolates therefore carry the compensatory Met182Thr mutation.[5]

The frequency of cefotaxime-resistant variants was compared in libraries derived from (i) wild-type TEM-1, (ii) the engineered TEM-1 variants carrying individual consensus mutations described earlier, and (iii) Lib250 after the 10th round of mutagenesis and selection. The *in vitro* mutagenesis protocol used throughout the drift was applied to incorporate an average of approximately two mutations per gene, and the resulting libraries were selected on 0.3 μg/ml cefotaxime (wild-type TEM-1 only survives ≤0.02 μg/ml cefotaxime). A hundred surviving colonies were randomly picked from each library, and the resistance profile of the picked colonies was evaluated by measuring their survival rate on a range of cefotaxime concentrations (Fig. 6). The levels of cefotaxime resistance were substantially higher in the library derived from Met182Thr-TEM-1 (≤2.5-fold) and also (although to a lesser extent) in a library derived from combining genes each carrying one of the remaining four global suppressors (Arg120Gly, Glu147Gly, His153Arg, and Leu201-Pro). Notably, at the highest cefotaxime levels (≥5 μg/ml), libraries stemming from wild-type TEM-1 gave no resistant variants, whereas the other libraries yielded a small but reproducible number of survivors. Further, when mutated to the same degree, the drifting genes taken from the 10th round of Lib250 showed higher levels of adaptation relative to that of wild-type TEM-1.

Sequencing indicated that, as previously seen,[5] resistance towards high cefotaxime concentrations (≥2 μg/ml) required a combination of the two most common cefotaxime conferring mutations (typically Gly238Ser, Glu104Lys) and at least one global suppressor. The latter was Met182Thr in genes derived from the Met182Thr-TEM-1, and Arg120Gly or His153Arg in libraries derived from the combination of other back-to-consensus/ancestor mutations. Cefotaxime-resistant variants isolated from Lib250 were found to carry at least one (and often more than one) of the above-mentioned global suppressors.

Discussion

Data collected through 18 rounds of mutagenesis and selection indicated that the sequence of the drifting TEM-1 genes is maintained close to the family consensus and inferred ancestor (the sequence of the latter being close, but not identical, to the consensus). This trend is seen in the purging of mutations in highly conserved residues (selection against deviation from the consensus) and in the enriched mutations that largely occurred in positions where wild-type TEM-1 deviates from its consensus/ancestor (selection for back-to-consensus/ancestor mutations) (Fig. 2). Maintaining the sequence of the drifting TEM-1 close to the consensus had a profound impact on stability as indicated by the FoldX predictions—the stability effects of the mutations are strongly shifted towards neutrality (Fig. 3).

As expected, the traits of negative, purifying selection are clearer than those of positive, adaptive selection. Indeed, purging is a process that eliminates all variants that carry deleterious mutations, unless they are compensated for as described above. Under a neutral drift, and provided that large populations are maintained throughout, no fixation of mutations is expected to occur. Indeed, we have observed no fixation (the same mutation appearing in all variants), although many mutations were significantly enriched—i.e., their frequency in the neutrally drifting populations (i.e., the populations drifting under ampicillin selection) is much higher than expected. The lack of fixation of one particular mutation may actually indicate that the most pronounced selection force was negative, or purifying, selection that removed deleterious mutations—namely, the drift was primarily neutral. Any positive selection occurring in the drifting libraries is relatively weak. In addition, positive selection, e.g., a selection for a new enzymatic function, typically leads to the fixation of one, or few, specific mutations. In this case, many different mutations at a variety of protein positions seem to be capable of increasing stability. Diverse solutions are also available for certain individual positions. For instance, position 288 is occupied by Lys in TEM-1, and was enriched in four different substitutions (Arg, Glu, Thr, and Gln) all of which correspond to "allowed" residues at this position (Fig. 2c and Table 3). Enrichment of more than one mutation in the same position was also observed in highly conserved positions, e.g., Met182Val and Met182Arg, and these additional mutations, although not as stabilizing as the Met182Thr back-to-consensus mutation, are likely to have a compensatory effect. In addition, certain back-to-consensus substitutions such as Asn52Ala (Fig. 2c) cannot be reached by a single nucleotide exchange. However, enrichment towards other amino acids (primarily Asp52 and Ser52; Table 3) brought the sequence within a single nucleotide exchange from Ala (consensus) and Glu (ancestor). The Asn-to-Asp mutation also brings this position closer to the ancestor in its physico-chemical character.

The sequence changes described here also provide new insights regarding the neutral network that surrounds, or includes, the sequence of wild-type TEM-1. This known sequence is clearly not singular, but comprises only one edge within a whole network of sequences that are connected by single-amino-acid exchanges and encode proteins with same structure and function.[10,42,43] By virtue of keeping very large populations ($\geq 10^5$ TEM-1 polymorphs or variants per round) throughout the neutral drift, our experiment represents a significant sampling of this network. One clear indication is that the neutral network can be smoothly traversed by exchanges away from, and back into, residues that represent the family consensus and/or inferred ancestor. The latter compensates for the former, and

in this way a whole network, or cloud, of neutral variants is formed that deviates from the consensus/ancestor to a given, and largely constant, extent. This manner of expansion can be easily reproduced by computationally generated neutral networks with various models.[10,43] Thus, the data generated by this experiment comprise an opportunity to further explore structure–function relationships within neutral networks by computational means.

As anticipated by their nature, all the enriched back-to-consensus/ancestor mutations we analyzed appear to increase TEM-1's stability, as is the case with back-to-consensus/ancestor mutations in many other proteins.[23,25,27,28,44,45] What was less expected are the different and complex ways by which these mutations affect TEM-1's stability. Met182Thr (as previously reported)[5] appears to be the only mutation that affects TEM-1's thermodynamic stability as such, namely, by increasing the free-energy difference between the unfolded and native states. Other mutations, most notably Arg120Gly and His153Arg, seem to affect various partially folded intermediates that are prolific in TEM-1,[38,39] thereby affecting the kinetic stability of TEM-1. Back-to-consensus mutations have been previously associated with increased kinetic stability,[22] although in this case, it appears that both kinetic and thermodynamic stabilities are affected, with the major compensator (Met182Thr) belonging to the thermodynamic category. Interestingly, the kinetic and thermodynamic effects can be additive, as the biophysical analysis of the Met182Thr-Arg120Gly double mutant demonstrates (Fig. 4 and Table 4). Indeed, the combination of thermodynamic and kinetic mutations evidently possesses substantial stabilizing and compensatory powers (Table 5).

The acquisition of back-to-consensus/ancestor also had a marked effect on TEM-1's evolvability. Function-altering mutations can be severely destabilizing[5] and are, on average, more destabilizing than mutations that characterize neutral drifts.[4] They may not be tolerated in the absence of a suitable compensatory mutation. Under the intense drift applied here, adaptation towards high levels of cefotaxime proceeded *via* the simultaneous acquisition of as many as three mutations, primarily because of the compensatory power of the global mutations. Thus, consensus/ancestor sequences can diverge "downhill"[46] while exploiting their excess of thermostability and compensatory power to rapidly evolve and adopt new functions. In contrast, ordinary proteins possess only marginal stability and little compensatory power and thus diverge "uphill" while awaiting the acquisition of compensatory mutations after the fixation of function-altering mutations.[5]

It is worth noting that our data do not indicate that the acquisition of the stabilizing global suppressors led to a global increase in the stability and robustness to mutations of the drifting populations. Rather, the appearance of these global suppressors seems to have enabled a higher number of mutations to accumulate (genes carrying a global suppressor, and Met182Thr, in particular, carried on average a higher number of mutations; data not shown) that, in turn, resulted in a parallel reduction in stability. Overall, as suggested by the FoldX predictions, the drifting genes are on average less stable, and consequently more sensitive to mutations, than wild-type TEM-1 (average $\Delta\Delta G$ values are 0.74 kcal/mol per mutation in Lib12.5, and 0.6 kcal/mol for Lib250). Nonetheless, because the drifting populations exhibit a very wide distribution with regard to the stability and number and type of mutations, there exists a significant subpopulation of variants that are more stable and more tolerant to mutations than wild-type TEM-1 (as indicated, e.g., by the compensation experiments). Similarly, a subpopulation of variants that carry a combination of a global suppressor and a new function mutation is responsible for the higher evolvability of the drifting populations relative to wild-type TEM-1. Thus, as demonstrated for RNA viruses,[9] higher robustness and evolvability are properties of the highly polymorphic populations as such, and not of all, or even most, of the individual genes that comprise them.

An interesting, although speculative, implication of our results relates to the fact that by default, the consensus sequence is highly related to the inferred, or predicted, ancestral sequence. As can be seen in TEM-1's alignment, 6 of the 13 enriched mutations are found in inferred ancestors for the subgroup of class A β-lactamases to which TEM-1 and its closest homologues belong (JN56; Fig. 2), and/or in ancestors that reside deeper in the β-lactamase phylogeny (Supplementary Fig. 1). To what degree the long gone real ancestor resembles the inferred ancestor, and hence the consensus, is obviously unknown. Nonetheless, inferred, or resurrected, ancestors have been shown to possess high thermostability, and this property has been associated with primordial high-temperature environments.[29,30] However, the stability of reconstructed ancestors might be overestimated[47] also because of the inherent convergence of inferred ancestors and family consensus sequences and the higher stability of the latter. On the other hand, the unique stability of the ancestor sequence relative to the consensus has also been demonstrated.[31] Putting aside the caveats regarding the validity of ancestral predictions for a moment, one might speculate that this excess of stability can be associated not only with thermophilic environments, but also with intense mutational loads. This because robustness, or neutrality, implies a flat position in genotype–phenotype space and a minimized effect of mutations, and is thus, in a way, counterproductive to adaptation.[11,48] For example, excessive stability can become inhibitory to adaptation by limiting structural plasticity.[49,50] It therefore follows that a large excess of stability is advantageous only, or primarily, under high mutational loads where the excess of stability can be exchanged with an excess of mutational tolerance. The "hyperthermophilic universal ancestor" hypothesis[30,51] might therefore be complemented by a "hypermutating universal ancestor" hypothesis. In the course of evolution, the gradual decrease in the planet's temperature[29] may

have also allowed themophilic ancestors to diverge downhill and rapidly accommodate new functions while losing their excess of thermostability.

Finally, our results demonstrate new ways of performing protein evolution *in vitro*. Neutral drifts provide gene libraries in which all variants are active and stable. As indicated by the compensation experiment (Table 5) and the FoldX analysis (Fig. 3), a substantial subfraction of neutral variants are even more stable than wild type. Neutral drifts have also been shown to foster the potential for new functions in the form of latent adaptations towards a range of promiscuous activities and thus allow the isolation of mutants exhibiting new enzymatic activities from very small libraries.[13] As seen in Fig. 6, the incorporation of global suppressors onto wild-type TEM-1 has greatly increased the potential for a new function emerging, in particular, at the high cefotaxime levels. Thus, under the high mutation rates that are typically used in laboratory evolution, genes carrying back-to-consensus/ancestor stabilizing mutations of the types described here (i.e., affecting both kinetic and thermodynamic stability) are better starting points for library making. Spiking back-to-consensus/ancestor mutations into gene libraries (using, e.g., synthetic oligo shuffling protocols[52,53]) can also boost the directed evolution of enzymes.

Materials and Methods

Library construction

The TEM-1 β-lactamase gene under the control of its endogenous promoter (originating in Tn3) was recloned into a modified pUC19 plasmid (NEB) containing the pMB1 origin of replication from pBR322, as previously described.[14] Briefly, pUC19 was modified by introducing new restriction sites (NcoI and NotI) flanking the start and stop codons of TEM-1's ORF, and by incorporating a chloramphenicol resistance gene (cam^r) at the XbaI site. Random mutagenesis was performed by error-prone PCR using the wobble base analogues 6(2-deoxy-beta-D-ribofuranosyl)-3,4-dihydro-8H-pyrimido[4,5-c][1,2]-oxazin-7-one triphosphate (dPTP) and 8-oxo-2'-deoxyguanosine triphosphate (8-oxo-dGTP).[54] The protocol was optimized to achieve a representative fraction of transversions (30%). The overall number of mutations incorporated was controlled by separating the introduction of transversions (by 8-oxo-dGTP) and transitions (by dPTP) and by calibrating the level of mutagenesis *via* the number of PCR cycles and base analogue concentration. The first 18 PCR cycles were performed with BioTaq polymerase (Bioline) in 50-μl reactions using primers flanking TEM-1's ORF, 25 μM 8-oxo-dGTP (TriLink), and 10 ng of plasmid template. The resulting PCR product was treated with DpnI (NEB, 1 h, 37 °C) to destroy the template plasmid, purified, and then used as a template (7 ng) for another 17 PCR cycles using the same primers and 4 μM dPTP (TriLink). The second PCR product was used as a template (7 ng) for 17 further PCR cycles, this time performed with nested primers and BioTaq polymerase. The product of this final PCR was restricted with NcoI and NotI (NEB, 3 h, 37 °C) and ligated into the modified pUC19 described above. This approach created with high reproducibility a library carrying an average of two mutations per gene and a typical pattern of nucleotide changes.[14]

Purifying selection of the libraries (neutral drift)

The pUC19 ligation products were electroporated into *E. coli* cells (E-cloni, Lucigen) and grown overnight at 37 °C in 500 ml LB medium containing 34 μg/ml chloramphenicol. DNA extracted from these liquid cultures (Maxi, Qiagen) was retransformed into XL-1 Blue cells (Stratagene) and plated on agar plates containing 34 μg/ml chloramphenicol and either 12.5 μg/ml (Lib12.5) or 250 μg/ml ampicillin (Lib250). After 7 h of incubation at 37 °C, surviving colonies were collected from the plates and their plasmid DNA was extracted. To ensure a diversity of $>10^6$ individual transformants, a sample from each transformation was serially diluted and plated onto chloramphenicol plates, and the number of colonies was counted.

Phylogenetic analysis

Tree construction

Sequence alignment was performed on the MAAFT server† using the L-INS-i strategy, the BLOSUM 62 scoring matrix and gap opening penalty 1.53, and extracted in Fasta format. The Fasta alignment was then transformed into Phylip format using the WWW READSEQ Sequence Conversion Server‡. The tree was constructed by PhyML (PhyML Online Server§ using the JTT substitution model, performing 100 bootstrap calculations, underestimation of the proportion of invariable sites and the Gamma distribution parameter, using the topology optimization and branch lengths and rate parameter optimization options. The PhyML trees were inspected using FigTree (freeware from the Evolutionary Biology Group, Oxford University‖.

Ancestral sequence calculation

Calculation of the ancestral sequences was performed by the FastML Software¶ using the JTT model, under branch length optimization, including gamma distribution as well as alpha optimization. (Nodes used to infer the ancestral sequences are designated with arrows).

The alignment

The ancestral sequences were aligned to the beta-lactamase sequences using the TCoffee program.[a]

Protein expression and purification

TEM-1 mutants were generated by site-directed mutagenesis, as described,[55] and expressed using OmpA-TEM-1 fusion within pET-24 plasmid (NdeI, NotI) carrying

† http://align.bmr.kyushu-u.ac.jp/mafft/online/server
‡ http://www-bimas.cit.nih.gov/molbio/readseq/
§ http://atgc.lirmm.fr/phyml
‖ http://evolve.zoo.ox.ac.uk/software.html?id=figtree
¶ Server version, http://fastml.bioinfo.tau.ac.il
[a] http://igs-server.cnrs-mrs.fr/Tcoffee/tcoffee_cgi/index.cgi

kanamycin resistance (a generous gift of Prof. J. Pelletier).[56] The OmpA leading sequence fused to a TEM-1 (instead of its endogenous signal peptide) is responsible for the excretion of the mature TEM-1 protein into the extracellular medium. Constructs were transformed into BL21 (DE3) cells and grown in the LB media at 37 °C in the presence of kanamacin (25 μg/ml) up to OD_{600} of 0.8. The cells were then washed with M9 minimal media (to eliminate the contaminants present in the rich media), resuspended into M9 minimal media in the presence of kanamycin (25 μg/ml) and glucose (4 g/L), induced with 0.4 mg/ml IPTG, and grown for 18 h at 25 °C. The supernatant was collected and saturated to 45% with ammonium sulfate. The soluble fraction was collected and concentrated by precipitating with 85% ammonium sulfate. After extensive dialysis of the resolubilized pellet against 25 mM Tris–HCl (pH 8.4), samples were applied to an anion-exchange column (HiTrap Q FF, GE Healthcare) pre-equilibrated with 25 mM Tris–HCl (pH 8.4), washed, and the TEM-1 mutants eluted with a 0–600 mM linear gradient of NaCl. The eluted proteins were dialyzed against 25 mM Tris–HCl (pH 8.4), and finally purified (to >95%) by gel filtration (HiLoad Superdex 75 column, Amersham Biosciences) in the presence of 160 mM NaCl. The specific activity of the TEM-1 proteins (hydrolysis of the nitrocefin substrate) was measured throughout each purification step.

Thermal denaturation

Enzymes (5 μM, 200 mM potassium phosphate buffer, pH 7) were denatured by raising the temperature from 25 to 80 °C and cooling back to 25 °C at 1.5 °C/min ramping rates, using a Cary Eclipse fluorescence spectrophotometer. Changes in tryptophan fluorescence (280 nm excitation, 340 nm emission) were followed. All melts were apparent two-state but not fully reversible (23–58%). The collected spectroscopic data were normalized to eliminate temperature-dependent changes in tryptophan fluorescence of the native and unfolded states. The apparent midpoint temperature of melting (T_m^{app}) was derived by fitting the normalized data to Eq. (1)[57]:

$$F_{obs} = ((\alpha_N + p^*T) + (\alpha_U - q^*T)^*\exp((m_{N-U}^*T - T_m^{app})/RT))/(1 + \exp((m_{N-U}^*T - T_m^{app})/RT)) \quad (1)$$

where α_N is a fluorescence at $T = 25$ °C, p is a slope of fluorescence change of the native state, α_U is a fluorescence at $T = 80$ °C, q is a slope of fluorescence change of the unfolded state, and m_{N-U} is the slope of transition.

Chemical denaturation

Chemical denaturation (5 μM enzyme in 200 mM potassium phosphate buffer, pH 7, 25 °C) was induced by GdmCl (0 to 5 M range in 0.1 M intervals), and changes in tryptophan fluorescence (280 nm excitation, 340 nm emission) were followed in a Cary Eclipse spectrophotometer. The obtained transition curves were fitted using Eq. (2), assuming a three-state model (N↔H↔U), where N is the native state, H is the intermediate, and U is the fully unfolded state[40]:

$$F_{obs} = ((F_N + p^*[\mathrm{GdmCl}]) + (F_H + F_U^*\exp((-\Delta G_{U-H}^0 + m_{U-H}^*[\mathrm{GdmCl}]/RT))^*\exp((-\Delta G_{H-N}^0 + m_{H-N}^*[\mathrm{GdmCl}]/RT))/(1 + \exp((-\Delta G_{H-N}^0 + m2^*[\mathrm{GdmCl}]/582)^*(1 + \exp((-\Delta G_{U-H}^0 + m_{U-H}^*[\mathrm{GdmCl}]/RT) \quad (2)$$

where F_N is a fluorescence at 0 M GdmCl, p is a slope of fluorescence change of the folded state, F_U is a fluorescence at 5 M GdmCl, q is a slope of fluorescence change of the unfolded state, F_H is a fluorescence at the intermediate stage, ΔG^0_{H-N} and ΔG^0_{U-H} are the differences in free energy between H and N and between U and H states, respectively, in the absence of denaturant, and m_{H-N} and m_{U-H} are the slopes of the respective transitions.

Enzyme kinetics

The kinetic parameters (1 nM enzyme in 100 mM potassium phosphate buffer, pH 7, 25 °C) were assayed spectrophotometrically at 486 nM using the chromogenic substrate nitrocefin ($\Delta\varepsilon_{486} = 20{,}000\ M^{-1}\ cm^{-1}$). The initial rates were obtained with substrate concentrations ranging from 20 to 300 μM. The kinetics parameters K_M and k_{cat} were determined by initial velocity non-linear regression analysis [Eq. (3)] using KaleidaGraph (Synergy Software):

$$V_0 = k_{cat} \cdot [S]/(K_M + [S]) \quad (3)$$

where V_0 is initial velocity and [S] is substrate concentration.

Recovery of enzymatic activity after partial chemical denaturation

TEM-1 variants (36 μM, 100 mM potassium phosphate buffer, pH 7) were incubated for 30 min at 37 °C in the presence of varying GdmCl concentrations (0, 0.5, 1, 1.2, 1.4, 1.6, 2, and 3 M) and then diluted (with 100 mM potassium phosphate buffer, pH 7, 37 °C) to a final denaturant concentration of 0.02 M. Recovery of the enzymatic activity

Table 6. The deleterious mutations and the primers used to incorporate them

Deleterious mutation		Primer
L76N	Forward	GATATCGTAACA**GGTCTC**[a]GAAC[b]TGTGGCGCGGTATTATCCCG
	Reverse	TAACTGACATCA**GGTCTC**CAGTTCAGAACTTTAAAAGTGCTCATCATTG
N136S	Forward	GATATCGTAACA**GGTCTC**CGTCTACTTCTGACAACGATCGGAG
	Reverse	TAACTGACATCA**GGTCTC**CAGACTGGCCGCAGTGTTATCACTC
I246D	Forward	GATATCGTAACA**GGTCTC**TGACATTGCAGCACTGGGGCCA
	Reverse	TAACTGACATCA**GGTCTC**ATGTCACCGCGACTCCCACGCTC
Q206P	Forward	GATATCGTAACA**GGTCTC**CCGCTAATAGACTGGATGGAGGCGG
	Reverse	TAACTGACATCA**GGTCTC**CAGCGGTTGCCGGGAAGCTAGAGTAAG
L199P	Forward	GATATCGTAACA**GGTCTC**CCGACTCTAGCTTCCCGGCAAC
	Reverse	TAACTGACATCA**GGTCTC**AGTCGGTAGTTCGCCAGTTAATAG
R222C	Forward	GATATCGTAACA**GGTCTC**CCTCTGCTCGGCCCTTCCGGCTGG
	Reverse	TAACTGACATCA**GGTCTC**CAGAGAAGTGGTCCTGCAACTTTATC

[a] BsaI restriction site.
[b] Mutated nucleotides.

(1 nM refolded enzyme, 100 mM potassium phosphate buffer, pH 7, 25 °C) was spectroscopically monitored by nitrocefin hydrolysis (100 μM) at 486 nm. The experiment was performed two to three times, and the average values and the standard deviations were calculated.

Inhibitory concentration determination

To determine the ampicillin inhibitory concentrations (IC) of various TEM-1 clones, overnight colonies (grown at 37 °C on agar plates bearing chloramphenicol only) were picked and individually grown in 96-well plates (200 μl LB medium with 34 μg/ml chloramphenicol, 3 h, 37 °C with shaking). The 96-well plates were then replica plated on a set of chloramphenicol agar plates with varying concentrations of ampicillin (0, 12.5, 25, 50, 100, 250, 500, 1000, 2000, and 2500 μg/ml). A similar protocol was applied for cefotaxime with the following differences. Various TEM-1 libraries were transformed into XL1-Blue cells and plated (>100,000 transformants per plate) on agar plates bearing 34 μg/ml chloramphenicol and 0.3 μg/ml cefotaxime. A hundred surviving colonies were collected and individually grown overnight in 96-well plates. The overnight culture was then diluted 1:10 and grown for several hours to obtain OD_{600} of 0.5. The 96-well plates were then replica plated on agar plates containing chloramphenicol and a range of cefotaxime concentrations (0.5–15 μg/ml). The IC values correspond to the highest concentration of antibiotic on which a given colony could grow.

Introducing deleterious mutations

Deleterious mutations (Table 6) were individually introduced to a given TEM-1 library by the following procedure. All-around PCR (25 cycles) was performed in 50 μl with Expand Long polymerase (Roche) using ~5 ng of the plasmid library and a pair of mutagenesis primers. These primers were designed to carry the designated deleterious mutation and a tail recognized by a type IIs restriction enzyme (BsaI) (Table 6). DpnI treatment (NEB, 1 h, 37 °C) digested the template, and the resulting product was restricted with BsaI (Fermentas, 3 h, 37 °C), ligated, and transformed into E-cloni. The transformants ($>10^6$) were grown overnight at 37 °C in 500 ml LB medium containing 34 μg/ml chloramphenicol, and the plasmid DNA was extracted from the liquid culture (Maxi, Qiagen).

Computation of stability effects (ΔΔ*G*) by FoldX

We applied the published four-step procedure[35]: (i) the 3-D structure of TEM-1 (Protein Data Bank code 1BTL) was optimized; (ii) mutant 3-D structures were created by introducing one mutation at a time and locally optimizing the structure; (iii) the energy functions of the mutants were computed and compared with the wild type's function; (iv) the actual free-energy values were calculated based on FoldX's correlation function between experimental and calculated data [$\Delta\Delta G^{\text{Experimental}} = (\Delta\Delta G^{\text{FoldX}} + 0.078)/1.14$] (Ref. 2).

Acknowledgements

Financial support by the Minerva Foundation, the EU MiFEM network, and the Estate of Fannie Sherr are gratefully acknowledged. We are very grateful to Chaparro-Riggers and Andreas Bommarius for their insightful consensus analysis of TEM-1, and to Roy Bekerman, Michal Segal, Hannes Hchtenhagen, Shalev Itzkovitz, Dana Reichmann, and Adrian Hugenmatter for their valuable assistance. D.S.T. and S.B. designed the research, analyzed the data, and wrote the article. S.B. and K.G. performed the experiments.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2008.04.024

References

1. Camps, M., Herman, A., Loh, E. & Loeb, L. A. (2007). Genetic constraints on protein evolution. *Crit. Rev. Biochem. Mol. Biol.* **42**, 313–326.
2. Tokuriki, N., Stricher, F., Schymkowitz, J., Serrano, L. & Tawfik, D. S. (2007). The stability effects of protein mutations appear to be universally distributed. *J. Mol. Biol.* **369**, 1318–1332.
3. Yue, P., Li, Z. & Moult, J. (2005). Loss of protein structure stability as a major causative factor in monogenic disease. *J. Mol. Biol.* **353**, 459–473.
4. Tokuriki, N., Stritcher, F., Serrano, L. & Tawfik, D. S. How protein stability and new functions tradeoff. *PLoS Comput. Biol.* In press. doi:10.1371/journal.pcbi.1000002.
5. Wang, X., Minasov, G. & Shoichet, B. K. (2002). Evolution of an antibiotic resistance enzyme constrained by stability and activity trade-offs. *J. Mol. Biol.* **320**, 85–95.
6. DePristo, M. A., Weinreich, D. M. & Hartl, D. L. (2005). Missense meanderings in sequence space: a biophysical view of protein evolution. *Nat. Rev. Genet.* **6**, 678–687.
7. Drake, J. W., Charlesworth, B., Charlesworth, D. & Crow, J. F. (1998). Rates of spontaneous mutation. *Genetics*, **148**, 1667–1686.
8. van Nimwegen, E. (2006). Epidemiology. Influenza escapes immunity along neutral networks. *Science*, **314**, 1884–1886.
9. Vignuzzi, M., Stone, J. K., Arnold, J. J., Cameron, C. E. & Andino, R. (2006). Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature*, **439**, 344–348.
10. Bornberg-Bauer, E. & Chan, H. S. (1999). Modeling evolutionary landscapes: mutational stability, topology, and superfunnels in sequence space. *Proc. Natl. Acad. Sci. USA*, **96**, 10689–10694.
11. Wagner, A. (2005). *Robustness and Evolvability in Living Systems.* Princeton University Press, Princeton, NJ.
12. Wilke, C. O., Wang, J. L., Ofria, C., Lenski, R. E. & Adami, C. (2001). Evolution of digital organisms at high mutation rates leads to survival of the flattest. *Nature*, **412**, 331–333.
13. Amitai, G., Devi-Gupta, R. & Tawfik, D. S. (2007). Latent evolutionary potentials under the neutral mutational drift of an enzyme. *HFSP J.* **1**, 79–87.
14. Bershtein, S., Segal, M., Bekerman, R., Tokuriki, N. & Tawfik, D. S. (2006). Robustness–epistasis link shapes

the fitness landscape of a randomly drifting protein. *Nature*, **444**, 929–932.
15. Bloom, J. D., Romero, P. A., Lu, Z. & Arnold, F. H. (2007). Neutral genetic drift can alter promiscuous protein functions, potentially aiding functional evolution. *Biol. Direct*, **2**, 17.
16. Bloom, J. D., Lu, Z., Chen, D., Raval, A., Venturelli, O. S. & Arnold, F. H. (2007). Evolution favors protein mutational robustness in sufficiently large populations. *BMC Biol.* **5**, 29.
17. Bloom, J. D., Silberg, J. J., Wilke, C. O., Drummond, D. A., Adami, C. & Arnold, F. H. (2005). Thermodynamic prediction of protein neutrality. *Proc. Natl. Acad. Sci. USA*, **102**, 606–611.
18. Forster, R., Adami, C. & Wilke, C. O. (2006). Selection for mutational robustness in finite populations. *J. Theor. Biol.* **243**, 181–190.
19. van Nimwegen, E., Crutchfield, J. P. & Huynen, M. (1999). Neutral evolution of mutational robustness. *Proc. Natl. Acad. Sci. USA*, **96**, 9716–9720.
20. Hall, B. G. & Barlow, M. (2004). Evolution of the serine beta-lactamases: past, present and future. *Drug Resist. Update*, **7**, 111–123.
21. Kimura, M. (1983). *The Neutral Theory of Molecular Evolution.* Cambridge University Press, Cambridge.
22. Godoy-Ruiz, R., Ariza, F., Rodriguez-Larrea, D., Perez-Jimenez, R., Ibarra-Molero, B. & Sanchez-Ruiz, J. M. (2006). Natural selection for kinetic stability is a likely origin of correlations between mutational effects on protein energetics and frequencies of amino acid occurrences in sequence alignments. *J. Mol. Biol.* **362**, 966–978.
23. Steipe, B., Schiller, B., Pluckthun, A. & Steinbacher, S. (1994). Sequence statistics reliably predict stabilizing mutations in a protein domain. *J. Mol. Biol.* **240**, 188–192.
24. Lehmann, M. & Wyss, M. (2001). Engineering proteins for thermostability: the use of sequence alignments *versus* rational design and directed evolution. *Curr. Opin. Biotechnol.* **12**, 371–375.
25. Lehmann, M., Kostrewa, D., Wyss, M., Brugger, R., D'Arcy, A., Pasamontes, L. & van Loon, A. P. (2000). From DNA sequence to improved functionality: using protein sequence comparisons to rapidly design a thermostable consensus phytase. *Protein Eng.* **13**, 49–57.
26. Forrer, P., Binz, H. K., Stumpp, M. T. & Pluckthun, A. (2004). Consensus design of repeat proteins. *ChemBioChem*, **5**, 183–189.
27. Nikolova, P. V., Henckel, J., Lane, D. P. & Fersht, A. R. (1998). Semirational design of active tumor suppressor p53 DNA binding domain with enhanced stability. *Proc. Natl. Acad. Sci. USA*, **95**, 14675–14680.
28. Wang, Q., Buckle, A. M., Foster, N. W., Johnson, C. M. & Fersht, A. R. (1999). Design of highly stable functional GroEL minichaperones. *Protein Sci.* **8**, 2186–2193.
29. Gaucher, E. A., Govindarajan, S. & Ganesh, O. K. (2008). Palaeotemperature trend for Precambrian life inferred from resurrected proteins. *Nature*, **451**, 704–707.
30. Miyazaki, J., Nakaya, S., Suzuki, T., Tamakoshi, M., Oshima, T. & Yamagishi, A. (2001). Ancestral residues stabilizing 3-isopropylmalate dehydrogenase of an extreme thermophile: experimental evidence supporting the thermophilic common ancestor hypothesis. *J. Biochem. (Tokyo)*, **129**, 777–782.
31. Watanabe, K., Ohkuri, T., Yokobori, S. & Yamagishi, A. (2006). Designing thermostable proteins: ancestral mutants of 3-isopropylmalate dehydrogenase designed by using a phylogenetic tree. *J. Mol. Biol.* **355**, 664–674.
32. Polizzi, K. M., Chaparro-Riggers, J. F., Vazquez-Figueroa, E. & Bommarius, A. S. (2006). Structure-guided consensus approach to create a more thermostable penicillin G acylase. *Biotechnol. J.* **1**, 531–536.
33. Sideraki, V., Huang, W., Palzkill, T. & Gilbert, H. F. (2001). A secondary drug resistance mutation of TEM-1 beta-lactamase that suppresses misfolding and aggregation. *Proc. Natl. Acad. Sci. USA*, **98**, 283–288.
34. Hecky, J. & Muller, K. M. (2005). Structural perturbation and compensation by directed evolution at physiological temperature leads to thermostabilization of beta-lactamase. *Biochemistry*, **44**, 12640–12654.
35. Guerois, R., Nielsen, J. E. & Serrano, L. (2002). Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. *J. Mol. Biol.* **320**, 369–387.
36. Baker, D. & Agard, D. A. (1994). Kinetics *versus* thermodynamics in protein folding. *Biochemistry*, **33**, 7505–7509.
37. Watters, A. L., Deka, P., Corrent, C., Callender, D., Varani, G., Sosnick, T. & Baker, D. (2007). The highly cooperative folding of small naturally occurring proteins is likely the result of natural selection. *Cell*, **128**, 613–624.
38. Georgiou, G., Valax, P., Ostermeier, M. & Horowitz, P. M. (1994). Folding and aggregation of TEM beta-lactamase: analogies with the formation of inclusion bodies in *Escherichia coli*. *Protein Sci.* **3**, 1953–1960.
39. Vanhove, M., Raquet, X. & Frere, J. M. (1995). Investigation of the folding pathway of the TEM-1 beta-lactamase. *Proteins*, **22**, 110–118.
40. Vanhove, M., Guillaume, G., Ledent, P., Richards, J. H., Pain, R. H. & Frere, J. M. (1997). Kinetic and thermodynamic consequences of the removal of the Cys-77-Cys-123 disulphide bond for the folding of TEM-1 beta-lactamase. *Biochem. J.* **321**(Pt. 2), 413–417.
41. Bloom, J. D., Labthavikul, S. T., Otey, C. R. & Arnold, F. H. (2006). Protein stability promotes evolvability. *Proc. Natl. Acad. Sci. USA*, **103**, 5869–5874.
42. Schuster, P., Fontana, W., Stadler, P. F. & Hofacker, I. L. (1994). From sequences to shapes and back: a case study in RNA secondary structures. *Proc. Biol. Sci.* **255**, 279–284.
43. Govindarajan, S. & Goldstein, R. A. (1997). Evolution of model proteins on a foldability landscape. *Proteins*, **29**, 461–466.
44. Maxwell, K. L. & Davidson, A. R. (1998). Mutagenesis of a buried polar interaction in an SH3 domain: sequence conservation provides the best prediction of stability effects. *Biochemistry*, **37**, 16172–16182.
45. Ohage, E. C., Graml, W., Walter, M. M., Steinbacher, S. & Steipe, B. (1997). Beta-turn propensities as paradigms for the analysis of structural motifs to engineer protein stability. *Protein Sci.* **6**, 233–241.
46. Giver, L., Gershenson, A., Freskgard, P. O. & Arnold, F. H. (1998). Directed evolution of a thermostable esterase. *Proc. Natl. Acad. Sci. USA*, **95**, 12809–12813.
47. Williams, P. D., Pollock, D. D., Blackburne, B. P. & Goldstein, R. A. (2006). Assessing the accuracy of ancestral protein reconstruction methods. *PLoS Comput. Biol.* **2**, e69.
48. Kirschner, M. & Gerhart, J. (1998). Evolvability. *Proc. Natl. Acad. Sci. USA*, **95**, 8420–8427.
49. Ancel, L. W. & Fontana, W. (2000). Plasticity, evolvability, and modularity in RNA. *J. Exp. Zool.* **288**, 242–283.

50. James, L. C. & Tawfik, D. S. (2003). Conformational diversity and protein evolution-a 60-year-old hypothesis revisited. *Trends Biochem. Sci.* **28**, 361–368.
51. Di Giulio, M. (2001). The universal ancestor was a thermophile or a hyperthermophile. *Gene*, **281**, 11–17.
52. Fox, R. J., Davis, S. C., Mundorff, E. C., Newman, L. M., Gavrilovic, V., Ma, S. K. *et al.* (2007). Improving catalytic function by ProSAR-driven enzyme evolution. *Nat. Biotechnol.* **25**, 338–344.
53. Herman, A. & Tawfik, D. S. (2007). Incorporating synthetic oligonucleotides *via* gene reassembly (ISOR): a versatile tool for generating targeted libraries. *Protein Eng. Des. Sel.* **20**, 219–226.
54. Zaccolo, M., Williams, D. M., Brown, D. M. & Gherardi, E. (1996). An approach to random mutagenesis of DNA using mixtures of triphosphate derivatives of nucleoside analogues. *J. Mol. Biol.* **255**, 589–603.
55. Khersonsky, O. & Tawfik, D. S. (2006). The histidine 115–histidine 134 dyad mediates the lactonase activity of mammalian serum paraoxonases. *J. Biol. Chem.* **281**, 7649–7656.
56. Sosa-Peinado, A., Mustafi, D. & Makinen, M. W. (2000). Overexpression and biosynthetic deuterium enrichment of TEM-1 beta-lactamase for structural characterization by magnetic resonance methods. *Protein Expression Purif.* **19**, 235–245.
57. Fersht, A. (1998). Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding. In (Freeman, W. H., ed.), Macmillan, New York.